The Russian State Scientific Center for Robotics and Technical Cybernetics

Proceedings
of the International Scientific and Technological Conference

EXTREME ROBOTICS

June 13-15, 2019, Saint-Petersburg, Russia

Saint-Petersburg, 2019

er.rtc.ru

УДК 004.896:007.5

Э41

Proceedings of the International Scientific and Technological Conference EXTREME ROBOTICS. – SaintPetersburg: OOO "Izdatel'sko-poligraficheskii kompleks "Gangut" Publ., 2019. – 584 p.

Proceeding's materials highlight an array of challenging issues and tasks in the sphere of robotic systems and safety facilities discussed at the 30th International Scientific and Technological Conference «EXTREME ROBOTICS».

DOI: 10.31776/ConfER.30.2019

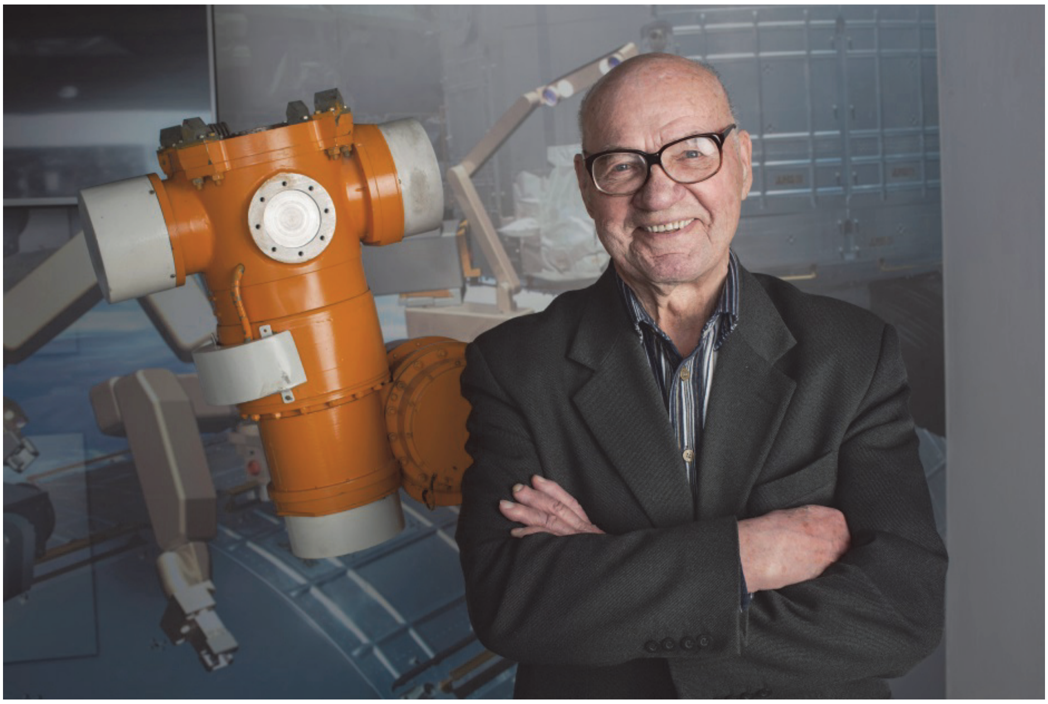
Papers are published with author's edition.

Труды международной научно-технической конференции ЭКСТРЕМАЛЬНАЯ РОБОТОТЕХНИКА (Proceedings of the International scientific and technological conference EXTREME ROBOTICS). – Санкт-Петербург: ООО "Издательско-полиграфический комплекс "Гангут", 2019. – 584 с.

Материалы сборника отражают круг актуальных проблем и задач в сфере робототехнических систем и средств безопасности, представленных на 30-й Международной научно-технической конференции «ЭКСТРЕМАЛЬНАЯ РОБОТОТЕХНИКА».

Доклады опубликованы в авторской редакции.

ISSN 2658-7645



Founder and First Director of the Russian Scientific Center for Robotics and Technical Cybernetics (RTC), Doctor of Technical Sciences, Professor, Honored Science and Technology Worker of the Russian Federation, Honorary Chief Designer of RTC, Honorary Co-Chairman of the Program Committee of EXTREME ROBOTICS Conference

Evgeny Yurevich

Dear Colleagues!

I am glad to welcome participants of the 30th International scientific and technical conference «Extreme Robotics», organized by the Russian State Scientific Center for Robotics and Technical Cybernetics. In the last decades, achievements in the field of robotics largely determine the successes in the development of outer space, the depths of the World Ocean, in the implementation of advanced medical technologies. This year the scope of our conference is devoted to practical questions of application of robotic means of new generation for solving problems of space exploration, depths of the World Ocean, medicine, nuclear power, production. Further development of the above mentioned areas is directly related to the development and implementation of new technologies, the expansion of mutually beneficial cooperation between domestic and foreign research centers and companies.

I am sure that the level of the forthcoming conference will allow you to get acquainted with the latest achievements in the field of extreme robotics, to cover current issues in research and development and to identify the main market trends in robot industry.

I wish all participants new discoveries, solutions, expansion of the circle of friends and partners!

Alexander Lopota

ORGANIZER

- *The Russian State Scientific Center for Robotics and Technical Cybernetics (RTC), Saint-Petersburg, Russia*

WITH SUPPORT OF

- *Ministry of Science and Higher Education of the Russian Federation*
- *Military-Industrial Commission of the Russian Federation*
- *EMERCOM of Russia*
- *Ministry of Health of the Russian Federation*
- *Russian Academy of Sciences*
- *State Scientific Centers of the Russian Federation Association*
- *All-Russian branch association of employers «Russian Engineering Union»*
- *Government of St. Petersburg*
- *Peter the Great St. Petersburg Polytechnic University*
- *Almazov National Medical Research Centre of the Ministry of Health of the Russian Federation*

INFORMATION SUPPORT

- *Journal «Mechatronics, Automation, Control», Moscow, Russia*
- *Journal «Proceedings SPIIRAS», Saint-Petersburg, Russia*
- *Journal «Robotics and Technical Cybernetics», Saint-Petersburg, Russia*
- *Journal «New Defensive Order. Strategy», Saint-Petersburg, Russia*

CONFERENCE PROGRAM COMMITTEE

Chairman:

- ***Gryaznov Nikolay**, Candidate of Physical and Mathematical Sciences, Deputy Director for Science of RTC, Saint-Petersburg*

Honorary Co-Chairman:

- ***Yurevich Evgeny**, Doctor of Technical Sciences, Professor, Honorary Chief Designer of RTC, Saint-Petersburg*

Academic Secretary:

- ***Spassky Boris**, Candidate in Technical Sciences, Head of Department, RTC, Saint-Petersburg*

Members of Program Committee:

- ***Albu-Schäffer Alin Olimpiu**, Doctor, Professor, Director of DLR Institute of Robotics and Mechatronics, Germany*
- ***Antsev Georgy,** General Director - General Designer, Concern Morinformsystem-Agat JSC*
- ***Bagnenko Sergey,** Academician of the RAS, Rector of Pavlov First Saint Petersburg State Medical University*
- ***Borovkov Alexey,** Vice-rector for perspective projects, Peter the Great Saint-Petersburg Polytechnic University*
- ***Ji Sup Yoon,** Doctor, Research Advisor, KAERI, Korea*
- ***Kalyaev Igor,** Academician of the RAS, Chairman of the Council on the priority of scientific and technological development of the Russian Federation*
- ***Kaprin Andrey,** Academician of the RAS, Director General of National Medical Research Radiological Center of the Ministry of Health of the Russian Federation*
- ***Kononov Alexey,** Head of Priority Technological Direction for Robotics Technologies of the National Center for the Development of Technologies and Basic Elements for Robotics of the Advanced Research Foundation*
- ***Mikrin Eugeny,** Chief Designer - First Deputy General Director, RSC Energia*
- ***Peschehonov Vladimir,** Academician of the RAS, General Director of Concern CSRI Elektropribor, JSC*
- ***Ploeger Paul**, Doctor, Professor Bonn-Rhein-Sieg University of Applied Science, Germany*
- ***Saeed Sarkar,** Doctor, Professor, Director of Research Center for Science & Technology in Medicine, Iran*
- ***Schneider Frank Eugen**, Doctor, Head of department, Fraunhofer Institute for Communication, Information Processing and Ergonomics (FKIE), Germany*
- ***Shlyakhto Evgeny**, Academician of the RAS, Director General of the Almazov National Medical Research Centre*
- ***Tsarichenko Sergey**, Doctor of Technical Sciences, Head of the proof ground of Research Institute «Geodeziya»*
- ***Uemura Kensuke,** Doctor, CEO, ShinMaywa Industries, Japan*
- ***Verba Vladimir,** Corresponding member of the RAS, Radio Engineering Corporation «VEGA», JSC*
- ***Vizil'ter Yury,** Professor of the RAS, Head of Division, FGUP GosNIIAS*
- ***Yushchenko Arkady,** Doctor of Technical Sciences, Professor, Baumann State Technical University*
- ***Zheltov Sergey,** Academician of the RAS, FGUP GosNIIAS*

CONFERENCE ORGANIZING COMMITTEE

Chairman:

- ***Lopota Alexander**, Doctor of Technical Sciences, Director and Chief Designer of RTC, Saint-Petersburg*

Deputy Chairman:

- ***Korenko Natalia**, Head of Center for Information and Analysis, RTC, Saint-Petersburg*

Secretary:

- ***Nikolaev Alexey**, Head of Department, RTC, Saint-Petersburg*

Members of Organizing Committee:

- ***Antsev Ivan,** Executive director, «NPP» Radar MMS, JSC*
- ***Chemezov Sergey,** CEO, Rostec Corporation*
- ***Emelyanov Sergey,** Doctor of Technical Sciences, Professor, Rector of Southwest State University*
- ***Kalabin Yuriy,** Chairman of the Committee on Industrial Policy and Innovations of St. Petersburg*
- ***Katalinic Branko,** Doctor, President of the DAAAM International Association, Austria*
- ***Katenev Vladimir,** Chairman of the Board St. Petersburg Commerce and Industry Chambers*
- ***Konyukhovskaya Alisa,** CEO, Russian Association of Robotics*
- ***Kudzh Stanislav,** Doctor of Technical Sciences, Rector of Moscow Technological University*
- ***Lobin Mikhail,** General director of executive directorate, first vice-president of Union of Industrialists and Entrepreneurs of Saint Petersburg*
- ***Maksimov Andrey,** Chairman of Committee for Science and Higher School of Saint-Petersburg*
- ***Martyanov Oleg,** Head of the National Center for the Development of Technologies and Basic Elements for Robotics of the Advanced Research Foundation*
- ***Medvedev Vadim,** Head of Department, Ministry of Science and Higher Education of the Russian Federation*
- ***Pshikhopov Vyacheslav,** Director, Research Institute of Robotics and Control Processes, SFU*
- ***Rudskoy Andrey,** Academician of the Russian Academy of Science, Rector of Peter the Great St. Petersburg Polytechnic University*
- ***Tsyganov Dmitry,** Deputy head of Department, Ministry of Science and Higher Education of the Russian Federation*
- ***Turichin Gleb,** Rector, SPbGMTU*
- ***Uiba Vladimir,** Head of the Federal Medical and Biological Agency of the Russion Federation*
- ***Vorob'eva Zhanna,** Chairman of the Committee on Education of St. Petersburg*

CONTENT

H. Senov, Yu. Bolgov

DEVELOPMENT OF A DIGITAL MAP OF MUDFLOW AND AVALANCHE AREAS OF KBR

Kabardino-Balkarian State University, Nalchik, Russia
XMSenov@mail.ru, yuriy6601@mail.ru

Abstract

The results of creating a digital model of the relief of the territory of the Kabardino-Balkaria Republic (KBR) according to radar interferometric survey are presented. The format of the initial data and the algorithm of the relief model construction are described. The algorithm provides rapid construction of a digital model of different parts of the earth's surface at different scales and levels of detail and can be used to build three-dimensional thematic maps for various purposes, such as mudflow and avalanche areas, and in the creation of navigation support systems of mobile robotics, including automatic drones.

Keywords: digital elevation model, radar interferometric survey.

Modern geographic information systems (GIS) based on three-dimensional models have recently become widespread. In addition to latitude and longitude, altitude data are the main coordinates in these GIS models. At the same time, such systems provide visibility of information and operate with tens and hundreds of thousands of elevations, not with units and tens, which was possible using the methods of "paper" cartography. Due to the availability of fast computer processing of large amounts of high-altitude data, the task of creating a realistic digital elevation model (DEM) becomes feasible. On the basis of DEM, in turn, information systems for various purposes are created, for example, thematic maps - steepness and exposure of slopes, erosion hazard, geochemical migration of elements, landscape stability, etc. Such data are very important in the development and design of monitoring systems for dangerous and especially dangerous natural phenomena [1, 2], for example, the construction of digital maps of mudflow or avalanche areas.

Of particular interest is the use of DEM in the tasks of navigation support of mobile robotics systems, including automatic drones. At the same time, if ready-made software packages [3] are used for the simple construction of the DEM, for more specialized tasks (including navigation in mobile robotics), the development of specialized software is necessary.

The aim of this work was to develop an algorithm and software for the construction of a digital map of the territory, allowing the data of radar interferometric measurements to build geographical maps of different areas and at different scales.

Initial data for construction of DEM

The initial data for the construction of the relief was the information of radar interferometric survey obtained from the Board of a reusable spacecraft (Shuttle) and called in the technical literature SRTM (Shuttle radar topographic mission) [4, 5]. Measurements were made from 11 to 22 February 2000 as part of the SRTM project with the objective of obtaining a digital model of the planet's surface with a resolution of about 30 m (1").

The selected flight parameters (orbit altitude of 233 km, inclination of 57°, orbital period of 89.2 min) provided the radar survey of 85% of the earth's surface, enclosed between 60° North latitude and 54° South latitude. [4]. For shooting we used two installed on-Board radar sensor SIR-C and X-SAR, performing the location of the planet's surface at C-band (wavelength $\lambda = 3{,}75 \div 7{,}5$ cm) and X ($\lambda = 2{,}5 \div 3{,}75$ cm), respectively, with a resolution time of 1 second. In just 11 days and 5.5 hours, the Shuttle made 182 turns, from which about 12 terabytes of radar data were obtained [5].

There is a preliminary version of the SRTM data (2003) and the final version (February 2005), and the 2005 data have undergone additional processing, consisting in the allocation of coastlines, water bodies, filtering of erroneous values, etc. They are provided to users free of charge, in one of the following options [4]:

– option 1 (for the United States) in the form of a regular grid with a resolution of one angular second (about 30 m) and a size of 1° 1° (3601 3601 elements);

– option 2 (for the territories of other countries, including Russia) in the form of a regular grid with a resolution of three angular seconds (about 90 m) and a size of 1° 1° (1201 1201 elements);

– option 3 in the form of ARC GRID files, as well as ARC ASCII and GeoTIFF format, containing a regular grid with a discreteness of three angular seconds (~90 m) and a size of 5°×5° (6001×6001 elements).

Option 3 data were obtained by processing the original elevation data to provide smooth topographical surfaces and interpolated values for areas where the original data were not available.

In all variants, the lower additional rows (1201, 3601, 6001) and the right additional columns (1201, 3601, 6001) are repeated on the adjacent matrix [4, 5]. The heights of the nodes of the regular grid are measured relative to the surface of the geoid EGM-96, rounded to one meter and represented by integers two bytes long (16 bits).

The data is divided into separate files, the names of which contain information about the coordinates of the represented area of the earth's surface. For example, for the second data presentation option, each file covers a 1-degree latitude and 1-degree longitude block of the earth's surface. The characters in the file name show the southwest corner of the block, the letters N, S, E, and W, denote North, South, East, and West. Thus, "N34W119.hgt" file covers from 34 to 35 degrees North latitude and from 118 to 119 degrees West longitude. File extension ".hgt" means the word "height". These files contain 16-bit integers, the height is measured in meters above sea level, in the "geographic" (latitude and longitude) projection, the lack of measurement data is indicated by the number -32768. International 3-second files have 1201 column and 1201 row data, the total file size is 2884802 bytes (1201×1201×2).

The declared measurement error values are 20 meters in plan and 16 meters in height [6]. The actual accuracy turned out to be slightly higher than the calculated accuracy (especially for the range X with a shorter wavelength), and the numerical error values for different regions of the planet differ by two or more times [6]. However, studies have shown that in some cases (in particular - in flat areas) the data have higher accuracy, and in mountainous areas - lower accuracy, and contain systematic errors caused by averaging the heights in the radar spots, moreover, the heights of the peaks are always underestimated, and the bottom of narrow gorges - overestimated [7].

The algorithm for constructing the DEM

To create a digital elevation model, software was developed in C# using DirectX. DirectX includes several components that underlie the programming of modern computer graphics:

– Direct3D - a low-level graphics API (software interface for applications) that allows you to display three-dimensional objects using hardware accelerators of three-dimensional graphics. Direct3D can be represented as an intermediary between an application and a graphics device (graphics card hardware).

– DirectDraw - allows you to manage computer hardware, providing direct and quick access to video memory.

The basis for building a three-dimensional image is a scene - a set of objects or models. The object is represented by a grid with triangular cells. Individual grid triangles are the elements by which objects are modeled. The camera and the light source of the objects are added to the scene.

The camera determines which part of the object the viewer can see and, therefore, for which part it is necessary to create an image. The camera is positioned and oriented in space and determines the visible area, which allows you to consider both objects as a whole and its individual fragments. The area of visible space is a truncated pyramid and is determined by the angles of the field of view, the front and rear planes. Objects that are not within the specified space are invisible and are excluded from further processing.

To build a terrain model, SRTM data is written to a vertex buffer, a special memory area that contains the coordinates of the grid vertices, surface normals, and texture coordinates. Vertex buffers are used to store data from the corresponding arrays because they can be stored in the graphics card's memory. Visualizing data stored in the graphics card memory is much faster than visualizing data stored in the computer's system memory. Fig. 1 shows the sequence of formation of three-dimensional digital elevation models of the earth's surface – a representation of a surface mesh with triangular cells, the calculation of the normals to the surfaces of the triangular cells and the cast of the normals to the vertices of the triangles, overlaid on a texture surface (in the form of satellite imagery or topographic maps of the area of the given surface). Fig. 2 shows, for example, the distribution of snow depth against a three-dimensional surface topography.

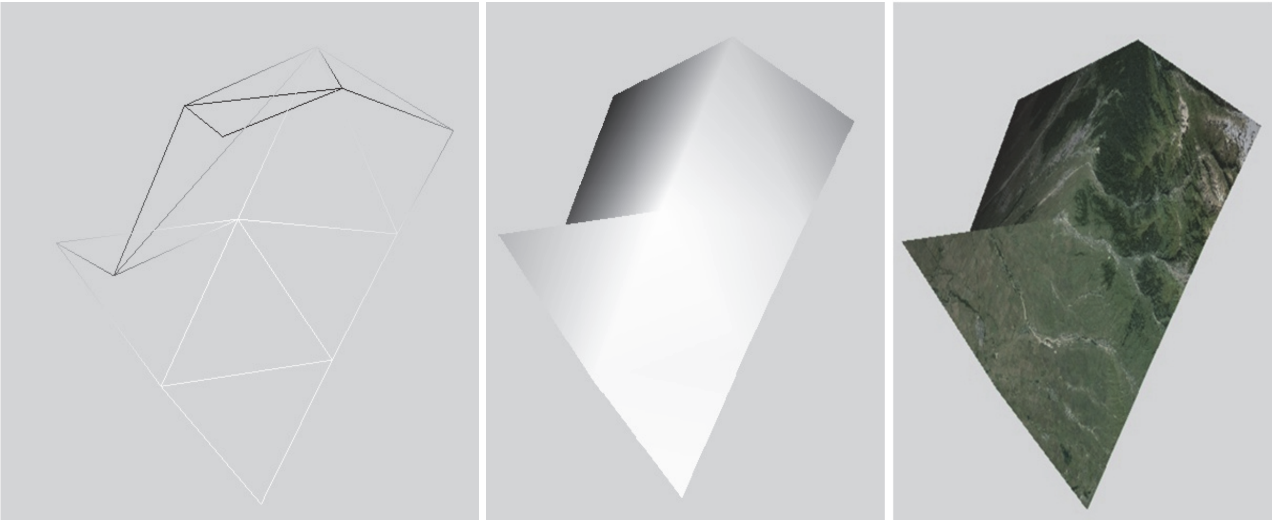


Figure 1 – Sequence of formation of a three-dimensional digital model of the earth's surface relief

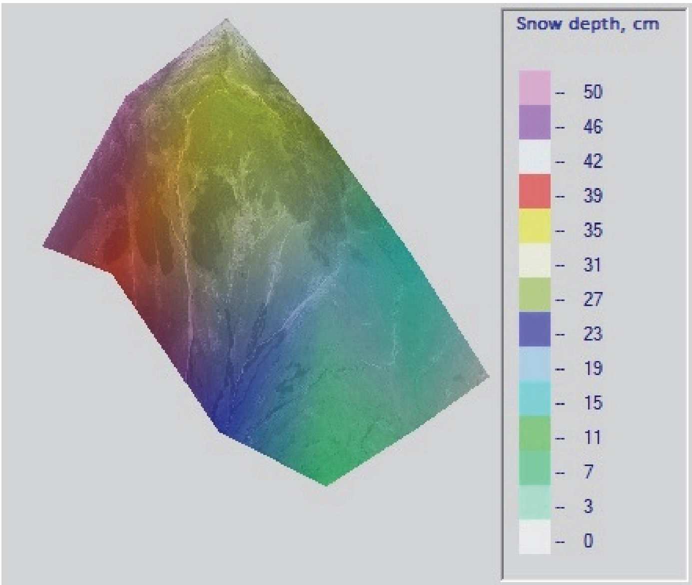


Figure 2 – Representation of snow depth distribution against a three-dimensional surface topography

Fig. 3 shows the mesh models of real terrain – views of the Baksan gorge (the most mudflow dangerous area of the KBR), built according to the SRTM.

Fig. 4 shows the elevation model of the same site, but based on the normals. Normal to the surface is used to determine the direction of reflected light, so the brightness of the triangles forming the surface depends on its position relative to the light source and the camera.

For full realism of the representation, the resulting relief model is superimposed texture. Textures, as a rule, are stored in graphic files of jpg, png and other formats. In this case, space images of the specified areas of the territory were used. To overlay textures to the mesh triangles it is necessary that their vertices contain texture coordinates. Texture coordinates are a pair of numbers, typically varying between 0 and 1, that uniquely point to a texture element called a Texel. Texture coordinates are defined for each vertex of the triangle, thus binding to some triangular area on the texture. Fig. 5 shows a terrain model with a superimposed texture (satellite image). In this case, the figure shows the area of the KBR (view of the Baksan gorge).

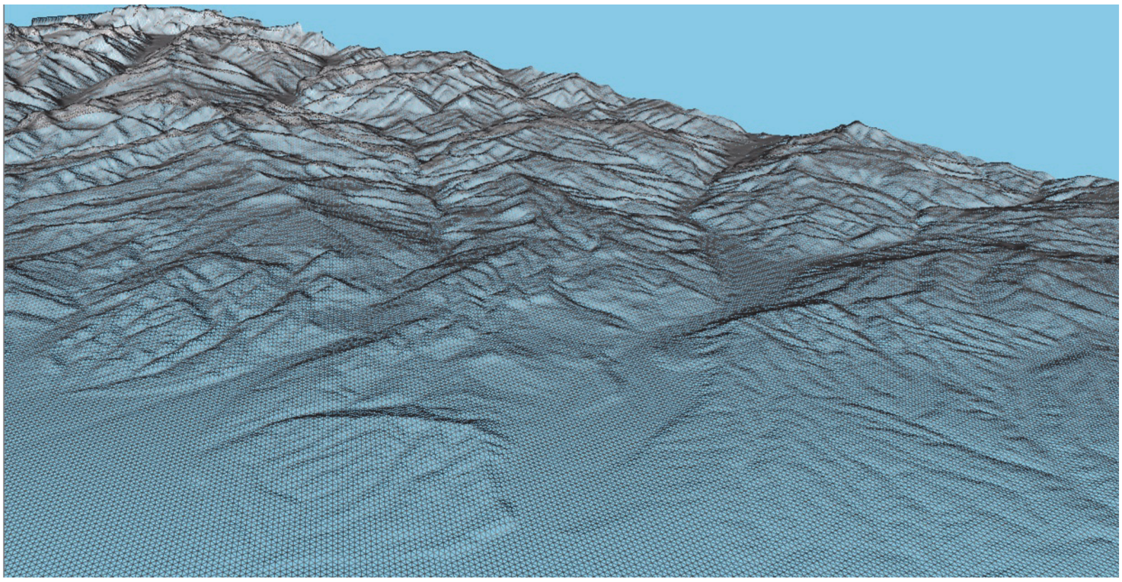


Figure 3 – Terrain model Grid

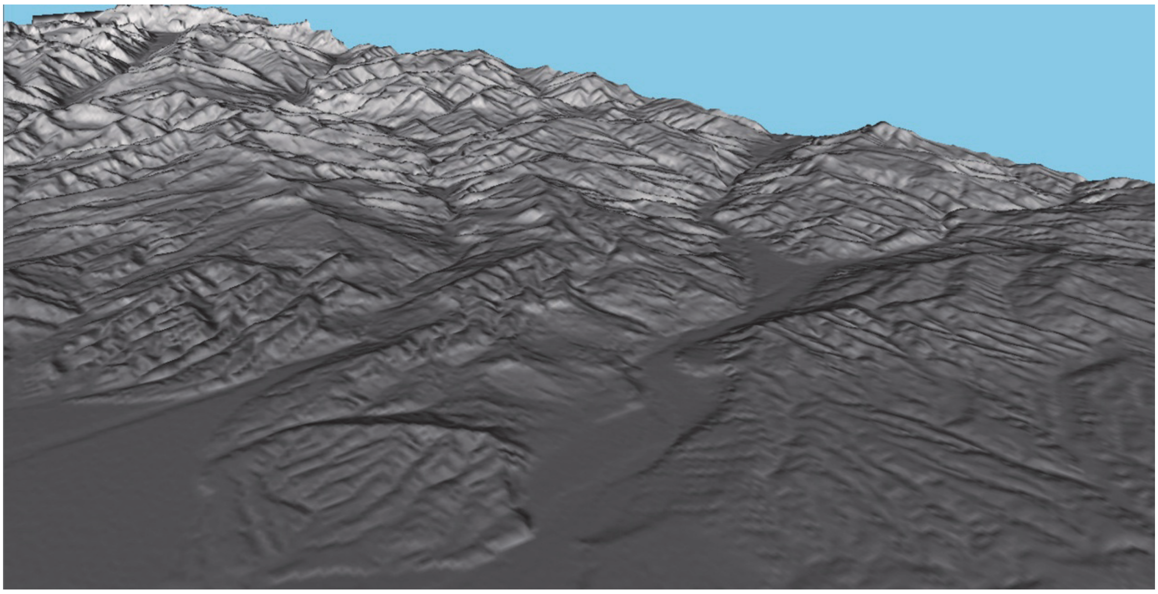


Figure 4 – Relief Model based on the calculated normals

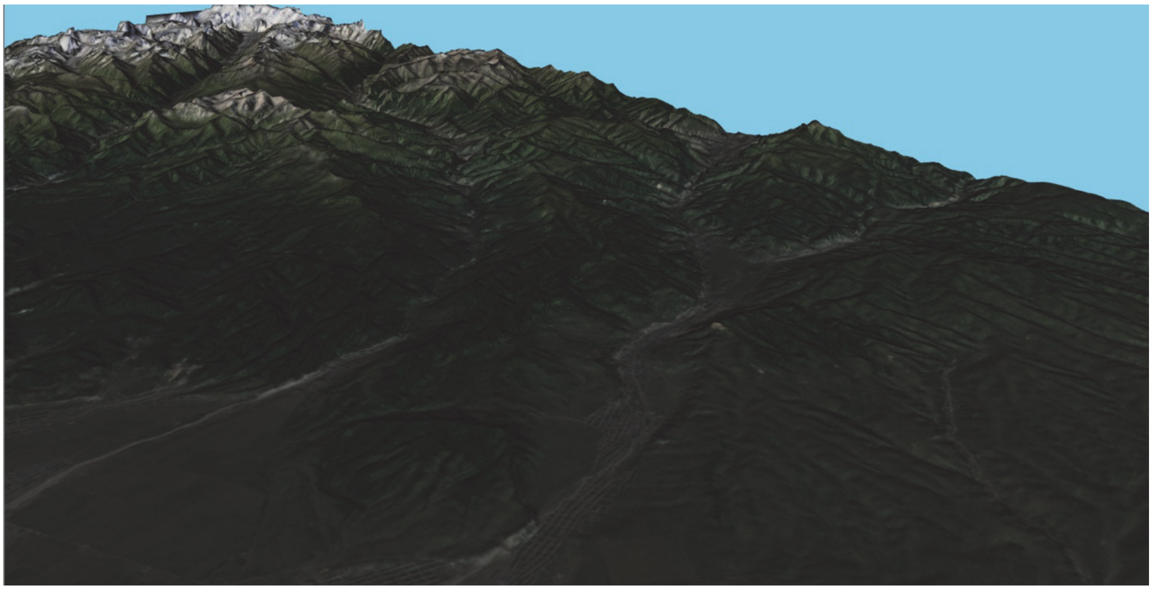


Figure 5 – Terrain Model with superimposed texture

Fig. 6 and 7 show, for example, a 1-degree latitude and 1-degree longitude elevation model for the KBR area. The upper (southwestern) corner of the rectangular region has coordinates of 43 degrees North latitude and 43 degrees East longitude. Data sources for building a terrain is depicted in figures 4 and 5 contains the file - N43E043.hgt. As a texture in fig. 7, a satellite image of this area of the earth's surface was used.

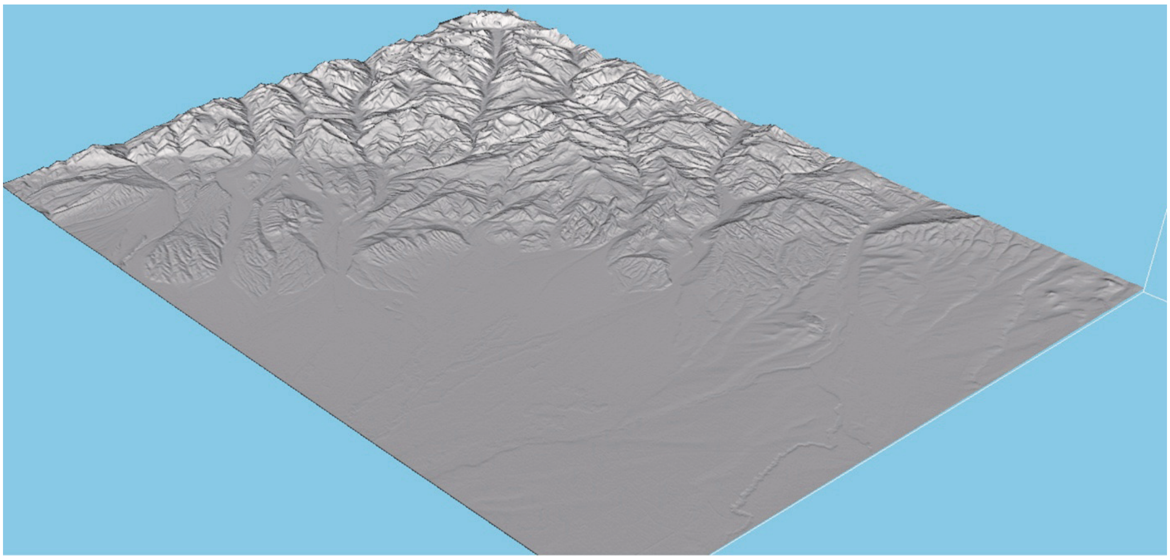


Figure 6 – Terrain Model of the land area of the KBR is built
with consideration of the calculated normal

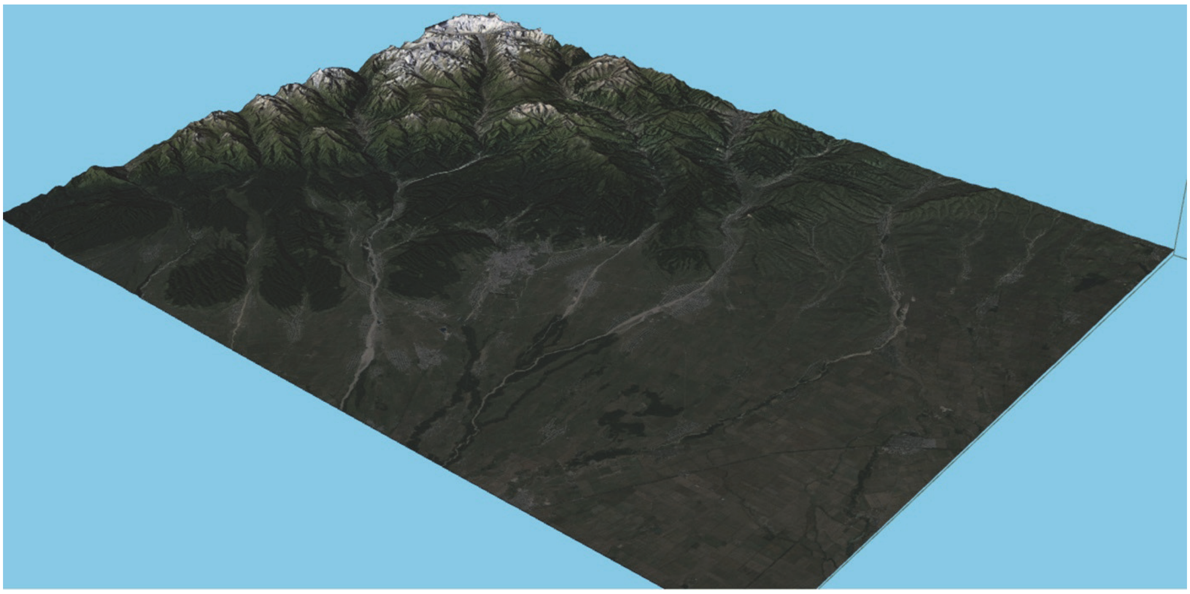


Figure 7 – Relief Model of the KBR area with
superimposed texture (satellite images)

Fig. 8 shows a three-dimensional digital terrain model of the area of the glacier (glacier Cosidon) at various angles. The North-South direction is marked on the relief.

In conclusion, fig. 9 shows an example of building a digital model of the relief of a large area of the earth's surface – part of the territory of the Caucasus range (in the center is visible mount Elbrus).

The presented algorithm provides for the rapid construction of Dems of different parts of the earth's surface at different scales and levels of detail and can be used for the construction of three-dimensional thematic maps for various purposes, and in the creation of navigation support systems of mobile robotics, including automatic drones.

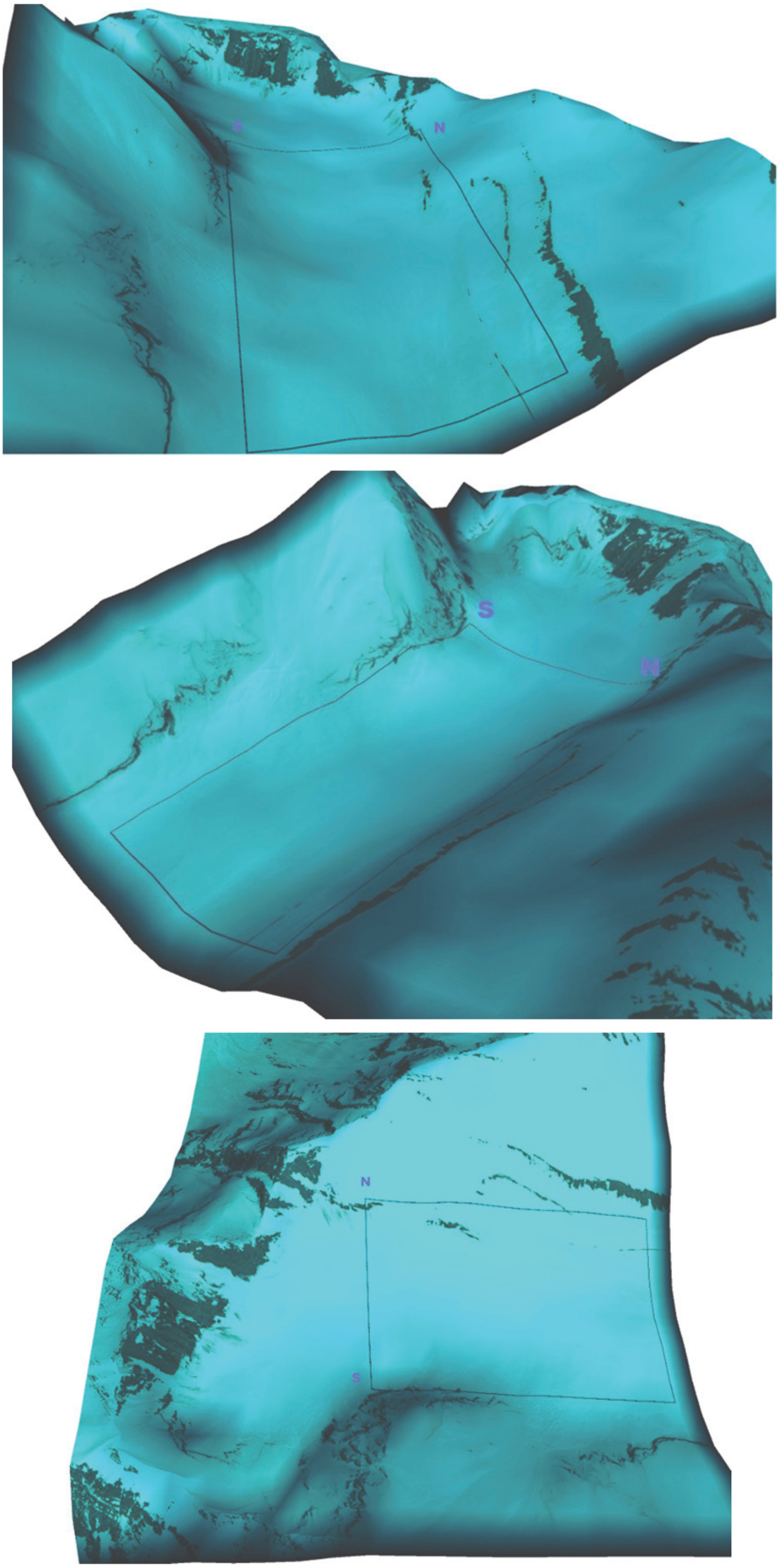


Figure 8 – Three-dimensional Digital terrain model of the area of the glacier (glacier Cosidon) at various angles

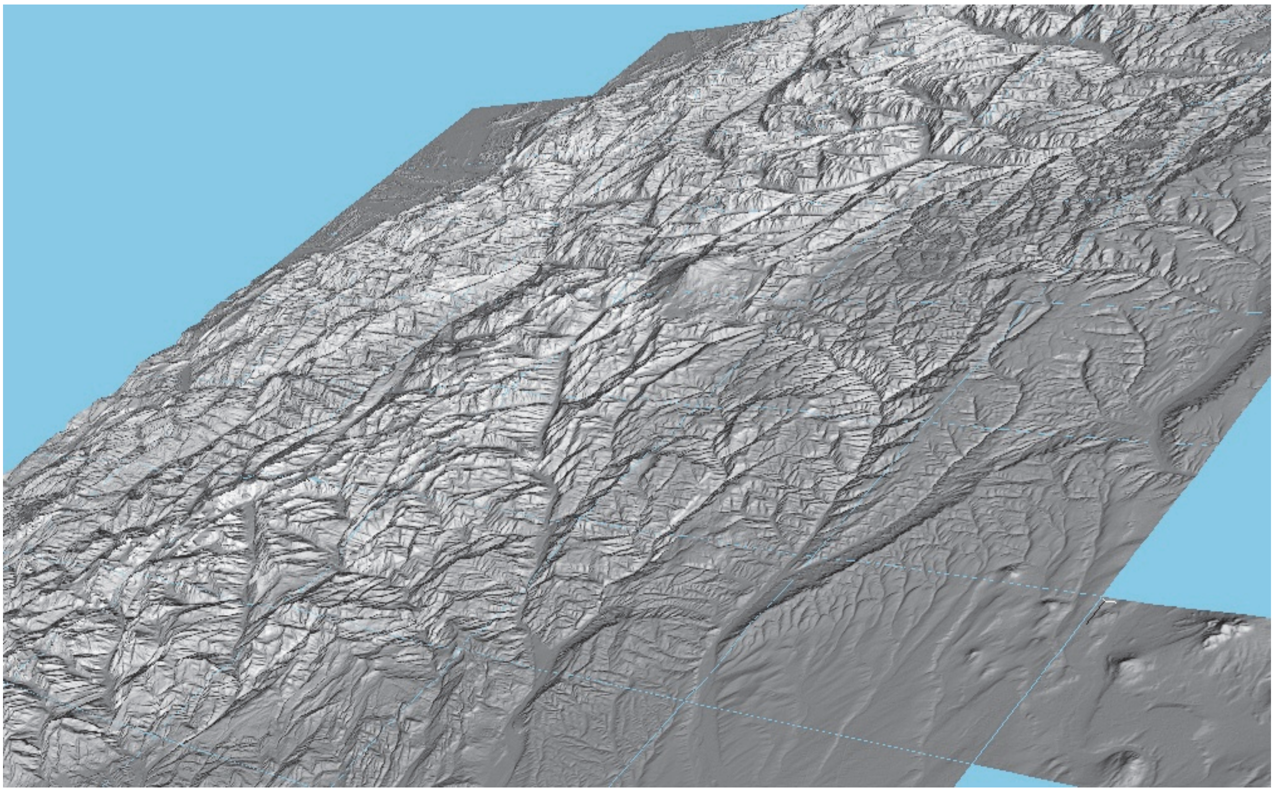


Figure 9 – An Example of building a digital model of
the relief of a large area of the earth's surface

References

1. Adzhiev A.H., Bolgov Ju.V., Senov H.M., Kondrat'eva N.V. Avtomatizirovannaja sistema distancionnogo monitoringa seleopasnyh ochagov. Kachestvo. Innovacii. Obrazovanie. Vserossijskaja konferencija «Informacionnye tehnologii, menedzhment kachestva, informacionnaja bezopasnost». 2015. № 5. t. II. s. 205 – 209. (in Russian).
2. Adzhiev A.H., Bolgov Yu.V., Kondratyeva N.V., Senov H.M. A Hardware–Software Complex for Remote Monitoring of Debris Flows //Pribory i Tekhnika Eksperimenta, 2016, No. 5, pp. 138–146. DOI: 10.1134/S002044121604014X.
3. About SRTM data and importing it with Arcinfo Workstation. http://gis-lab.info/qa/srtm.html
4. The shuttle radar topography mission. / Farr Tom G., Hensley Scott, Rodriguez Ernesto, Martin Jan, Kobrick Mike. // CEOS SAR Workshop. Toulouse 26-29 Oct. 1999. Noordwijk. 2000, c. 361-363.
5. Description and retrieval of SRTM data. http://gis-lab.info/qa/strm.html.
6. Karionov Ju.I. Ocenka tochnosti matricy SRTM. http://www.racurs.ru. (in Russian).
7. Muravev L. Vysotnye dannye SRTM protiv topograficheskoj semki.. http://web.ru/db/msg.html?mid=1177761. (in Russian).

H. Senov, Yu. Bolgov

THE DEVELOPMENT OF ROBOTIC SYSTEM FOR REMOTE MONITORING OF AVALANCHE AND MUDFLOW HOTBEDS MOUNTAIN AREAS OF KBR

Kabardino-Balkarian State University, Nalchik, Russia
XMSenov@mail.ru, yuriy6601@mail.ru

Abstract

The development of mountain areas is complicated by a number of natural hazards, which primarily include mudflows and avalanches. The article provides an overview of the existing monitoring systems of mudflow and avalanche hotbeds and presents a description of the project of a robotic complex for remote monitoring of avalanche and mudflow hotbeds in the mountainous territory of KBR. The complex includes a network of stationary monitoring devices, a system for receiving and processing information and a mobile component of the system (includes a robotic platform of high permeability and a quadrocopter with a video recording system).

Keywords: robotic complex, remote monitoring, mudflow, snow avalanche, mountain areas.

The development of mountain areas is complicated by a number of natural hazards, which primarily include mudflows and avalanches. Mudflow processes have a significant negative impact on the objects of economic activity in mountainous areas. The average annual damage caused by villages amounts to tens of millions of dollars and is often associated with loss of life. Due to ill-considered economic activities, failure to take timely protective and preventive measures against mountain erosion, the scale of mudflow danger is increasing. Therefore, the problem of network monitoring is relevant and economically feasible for a number of regions of Russia (Krasnodar and Stavropol territories, the republics of the North Caucasus, etc.). For accurate prediction of mudflow hazard it is necessary to observe both the development of exogenous processes in mudflows, and the accompanying meteorological factors and the regime of water and mudflow. The territory of mountainous areas of KBR is one of the most dangerous in Russia. Currently, monitoring and comprehensive studies of mudflow basins are carried out irregularly and only in local areas. Observations of mudflow basins from stationary points have no effect due to the lack of stable funding, only occasional route surveys of basins are carried out with the development of forecasts of possible scenarios of mudflows.

Human activities in cold mountain regions are always at risk from avalanches. The expansion of mountain development in Russia and, as a consequence, the increased risk of snow avalanches leads to an increase in the requirements for conceptual tools for the proper planning of infrastructure and design of safe structures. Construction organizations working on issues related to avalanche danger, in need, in addition to the qualitative assessments of hydrological and geomorphological conditions of the plots in quantitative methods for the prediction of avalanches and impact assessments. Such forecasts, as a rule, are based on mathematical models of snow mass movement, but due to the complexity of the phenomenon, there is still no unambiguous solution to the problem within the framework of the use of the mathematical apparatus of classical mechanics, in this regard, in practice, a variety of approaches are used, including systems for monitoring avalanche areas.

This paper presents a description of the project complex robotic remote monitoring of avalanche and mudflow hotbeds mountain areas of the KBR.

Mudflow processes

Villages are one of the most dangerous natural phenomena in mountainous regions and are characteristic of different climatic conditions. Their destructive capacity is due to various factors: transportation and removal of a huge amount of solid mass (up to 106 m3), fragments of which can reach large sizes (up to 10 m), high speed of movement (up to 10-15 m/c) [1]. Fig. 1 shows a picture of a mudflow dam destroyed in 2000 in the vicinity of the city Tyrnyauz. Advance warning of mudflow hazard is usually based on empirical relationships reflecting the correlation between precipitation and mudflow activity. Such systems use information from meteorological radars and ground-based rainwater networks [2]. Empirical relationships may include different parameters – rainfall intensity, duration of precipitation, amount of precipitation and its accumulation over a certain period of time. On the basis of observations of mudflow hotbeds entered the critical threshold values of meteorological parameters, under exceeding of which a warning message is issued about the mudflow hazard. During the operation of such systems, there is a high probability of false alarms – in fact, not every precipitation in excess of critical parameters causes mudflow [3]. Warning systems are placed directly in the areas of mudflow foci and allow you to control various parameters of the mudflow (including video image of mudflow in real time) both in the stage of its development and in the process of descent. The warning systems use ultrasonic sonars, ground vibration sensors, video cameras, photocells, etc. Such systems provide prompt warning of mudflow danger, reduce the probability of false alarms [1].

Figure 1 – Protective dam destroyed in 2000 in the vicinity of Tyrnyauz

Monitoring systems typically include a number of sensors and devices.

Ultrasonic sonars are used to monitor the flow level [4]. Sonars are used to record the hydrograph of the mudflow and allow you to control the erosion of the channel or alluvium at the installation site. Ultrasonic sensors are located above the flow channel and measure the distance from its surface to the sensor. Figure 2 shows, for example, the hydrograph recorded with the help of an ultrasonic sonar, which came down in Italy on July 8, 1996 (Moscardo Torrent) [5]. Radar and laser sensors are also used to measure the height of the flow [6].

Seismic activity sensors. The passage of the mudflow causes vibration of the soil (signals of seismic activity) and can be recorded by accelerometers, velocimeters, geophones [1]. Seismic activity sensors can be installed at a safe distance from the channel and thus are less prone to damage when the mudflow. The output signal of seismic sensors is proportional to the rate of displacement of their fixing. They are usually rigidly fixed on the rock.

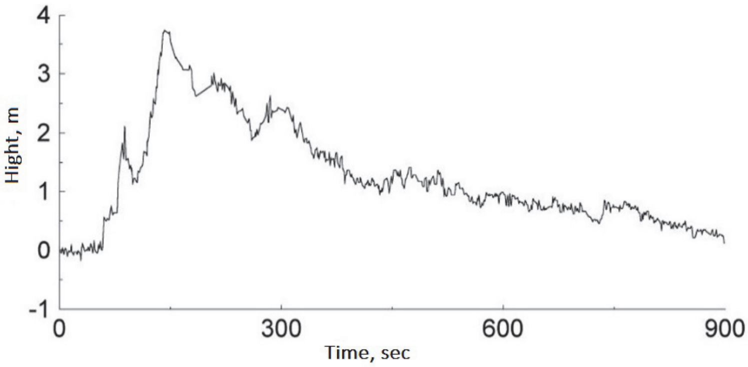


Figure 2 – Hydrograph (Moscardo Torrent, Italy, 8.07.1996) [5]

When processing signals from sensors of the seismic activity occurs is the problem of selection of a signal caused by the landslide, from the signals subject to other sources of fluctuations in the passage of trains or trucks, landslides rocks, etc. In [7], the analysis of the experimental data (recordings of readings of seismic activity in the landslide) and the possibility of filtering the signal caused by the landslide.

Wired sensors and photo sensors. A number of wires stretched at different levels above the channel are used to measure the maximum height of mudflows [1]. The height of the wave is fixed by the level of the highest wire, which was cut off by the flow, and the time of occurrence of the flow in the place of installation of sensors is recorded. After the wires break, these devices cannot give information about the height of the waves that follow the main mass of the flow. Photovoltaic cells and infrared sensors are used as a tool to record the mudflow. As in the case of wired sensors, several solar cells are installed at different heights in the cross section of the channel. The mudflow mass interrupts the beam emitted by the sensor, which allows to detect the passage of the mudflow and fix the height of the flow wave. When using sensors of this type, there is a risk of false positives from random factors (passage of animals, falling trees, etc.).

Camera-recorders. Video surveillance devices are a mandatory element of automated monitoring systems of mudflow foci [1]. The image of the observation area (usually obtained in real time) is used for visual assessment of the state of the mudflow hazard (Fig. 3) and interpretation of information received from other sensors of the system. In this case, the video data can be used for quantitative analysis of the speed, maximum flow rate and the total volume of material transported by the mudflow – is calculated based on the previously removed geometry of the cross section of the channel (according to topographic survey), the height of the flow and the speed of its movement.

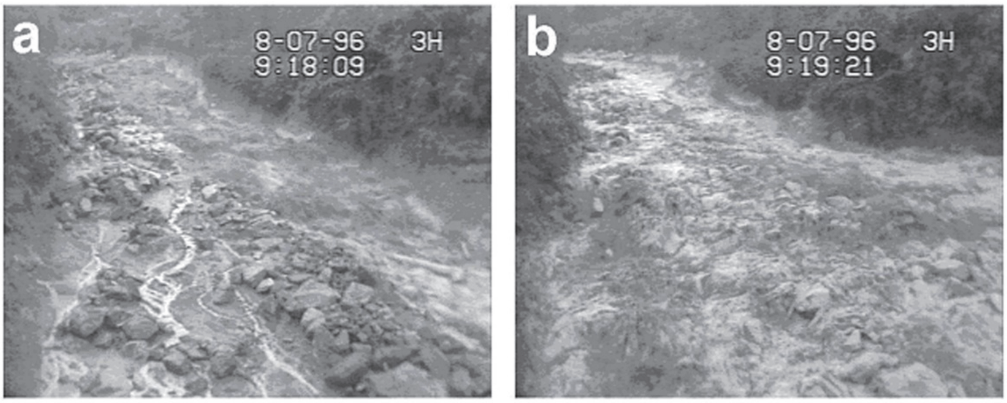


Figure 3 – Video footage of mudflow waves [5]:
a – initial stage, b – peak of mudflow

For example, in Fig. 3 shows some footage of the village, descended in the Italian Alps July 8, 1996 [5]. At the initial stage, the appearance of the mudflow is accompanied by a surge in the turbulence of the flow, mainly consisting of turbid water (Fig. 3a). After about one minute, the flow reaches its peak, with a height of about 4 m and a high content of transported material (Fig. 3b). Fog or low light at night can make it difficult to use a video surveillance system.

Pluviographs. Since a significant part of the mudflows has a storm genesis, the monitoring systems include automatic pluviographs, which can simultaneously be an element of the ground rainwater network. Data from pluviographs can also be used by advance warning systems for mudflow hazards as a supplement to radar data of the rains.

Pressure sensor. Mudflows have a devastating impact on all obstacles in their path, such as piles of bridges, dams, etc. Assessment of the impact of mudflows is important in the development of protective structures. The pressure generated by the flow during the passage of the mud is measured by sensors installed directly in the channel. Sensors are used to measure normal and shear loads generated by the mudflow, flow overpressure sensors, which are installed in the lower part of the channel. A number of devices for measuring the impact force of mudflows the principle of action, which is based on the measurement of the degree of deformation of structural elements under the influence of mudflow. Also used more complex tensometric sensors to register the change of the shock load of the mudflow at the time [7].

Sensors involved in the flow. They are strong sealed containers that are captured by the mudflow and transported by the stream. Inside the container, sensors of various types are placed to measure acceleration, impact force from collision with solid elements of mud mass and fluid pressure inside the flow [7]. Such sensors are used for scientific purposes to study the internal dynamics of mudflows.

Example of implementation of the system of monitoring and warning of mudflow hazard

For example, in Fig. 4 the layout of elements of the system of warning of mudflow danger (Eastern Italy, Alps) is presented [4].

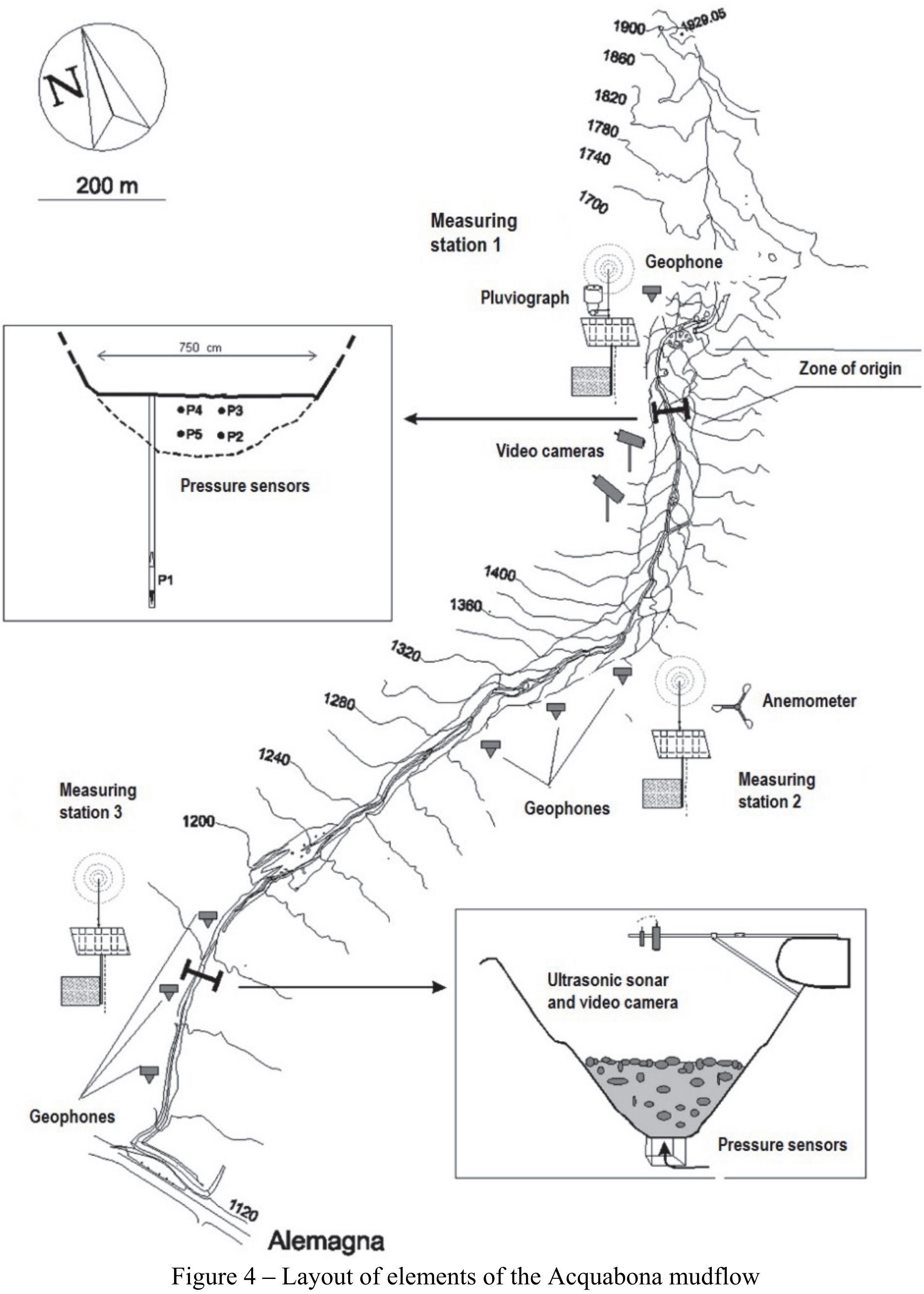


Figure 4 – Layout of elements of the Acquabona mudflow hazard warning system (Eastern Italy, Alps) [4]

The system includes three measuring stations located along the channel. The first measuring station is installed in the zone of origin of the mudflow and contains an automatic pluviograph, geophone, pressure sensors installed in the lower part of the channel, and video cameras. The second station is installed in the middle of the mudflow basin and includes a system of geophones placed along the channel. The third station is installed in the lower part of the mudflow basin and contains an ultrasonic sonar for recording the flow hydrograph, a geophone system, a video camera and a pressure sensor located in the lower part of the channel. Information from the measuring stations is transmitted by radio channel to the processing center for analysis and issuing a warning signal about the mudflow. A similar warning system is used in Taiwan [8]. Such systems contain measuring stations and sensors, the number of which is determined by the specific feature of debris flow area.

Snow avalanche

The main dynamic characteristics of an avalanche are considered to be: the speed of the leading edge; the range of the ejection; the force of the avalanche impact on the obstacle; the volume. All these characteristics vary depending on physical and geographical conditions [9]. In fact, the movement of the avalanche is mainly due to gravity, not wind as in the formation of snow deposits, and is characterized by a high speed of movement, measured in m/s, in contrast to the movement of sliding and sliding, which are measured by speeds from mm/day to cm/day.

The site of the slope and the valley where the snow avalanche originates, moves and stops is called an avalanche. In avalanches can be divided into three zones – the origin, transit and deposition of avalanches. The boundaries between these zones are fuzzy and are conditional [10]:

The zone of avalanches origin (avalanche hearth) is located in the upper part of the avalanche and is a section of the mountain slope, where there is an accumulation of snow mass, which can lose stability and form an avalanche. Most often it is a bowl-shaped depression in the apical part of the mountains of the Kara type or an extended part of the erosion incision. However, avalanches can also form on flat slopes. Areas of avalanches are usually confined to the smoothed slopes without forest steepness of more than 30°. Frequent cases, when avalanches are born on more gentle slopes. Avalanches can also occur on wooded slopes. The process of avalanche formation in addition to the steepness and nature of the slope surface is influenced by the amount and condition of snow.

Disturbance of stability and avalanche formation is observed on the slopes of the steepness from 15° to 60°. On the steeper slopes of the snow is poorly maintained, most snowflakes during snowfall rolls down and large masses of snow are deposited relatively rarely. Therefore, the rocks bordering the upper part of the zone of avalanches and steep slopes can serve as a source of additional snow accumulation. The most avalanche are considered steep slopes from 25 to 50°. Depending on the features of this avalanche zone, avalanches of loose snow may be classified (Fig. 5a) or an avalanche in the form of a slab (Fig. 5.b).

[a] [b]

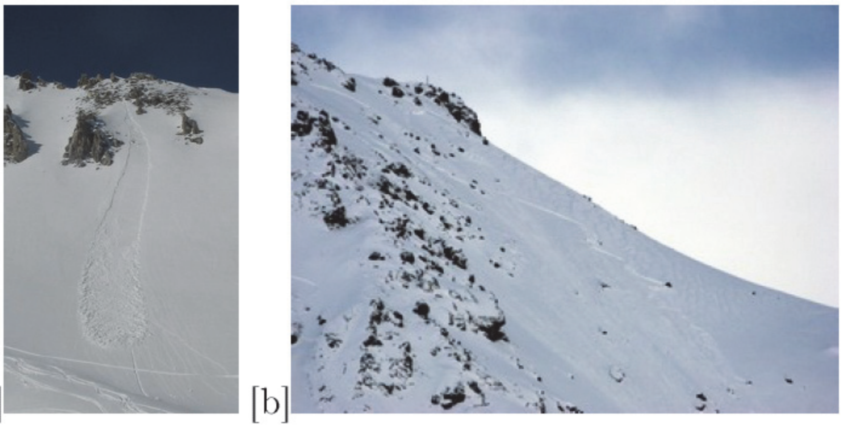


Figure 5 – a) Avalanche of loose snow, b) avalanche in the form of a plate

Avalanche formation begins at a point and gains mass, developing into a fan-like shape, usually of small size. This type of avalanche is formed in the presence of wet or dry unbound snow. In the second case, there is a well-defined fault line, which limits the release zone. If its top (upper bound) and its flanks (side bounds) are clearly visible, then the lower bound is not always recognizable. A slab is formed when a large layer of bound snow is superimposed on a weak layer, or when there is a weak bond between layers. Typically, avalanches formed in the form of plates lead to more catastrophic consequences than avalanches of unrelated snow, as they involve large snow masses.

Transit zone - the area between the starting zone and the sediment zone. The zone can be of channel type (channel avalanche) or in the form of an open slope (avalanche of an open slope). In this area, the avalanche is fully formed and reaches its maximum speed. Avalanche can involve a significant amount of snow mass, while its deposits are insignificant and increase when crossing ravines and behind rocks. The path of the avalanche can be channeled in a clearly defined bed or bed (avalanche tray), but can also be located on a relatively flat slope between the zones of origin and deposition of the avalanche. The length of the zones has transit channeled avalanches more than not channeled. An avalanche path can have several branches, when lateral channels flow into the main channel, each of which begins in a separate zone of avalanches. There may be cases when several avalanche routes are fed from one avalanche hearth.

Sediment zone - an area where an avalanche slows down quickly and deposits occur. For sewer avalanches, this zone often has a cone-shaped appearance (avalanche removal cone). At the place where the avalanche flow goes out on a gentle slope, the speed of the flow due to the increase in resistance forces slows down, the flow expands and there is a partial deposition of snow. Due to these deposits (especially long-term), involved in avalanches in the movement of rock particles, formed mineral cone removal with its characteristic shape — convex, expanding in terms of the shaft with decreasing steepness and convex in the transverse profiles. In the narrow valleys of the avalanche from a slope can pass in its bottom and climb the opposite slope, where it will be located area of the deposits. In non-channeled avalanches, the sediment zone is located at the foot or on a gentle slope at the bottom of the avalanche.

Three types of avalanches on the condition of snow are considered:

1) From dry snow – dust avalanche (Fig. 6). When the snow layer moves, its debris can collapse and form a dust cloud.

Figure 6 – Dust avalanche

2) From dry snow - snow plate (Fig. 7). A layer of fine-grained snow or snow Blizzard density of 250-600 kg/m3, lying on the surface of less dense snow. Under the snow Board often there are voids, which leads to its subsidence and destruction. The collapse of the snow layer takes place on a large area. Line break of snow avalanches constitutes the arch of the stage, perpendicular to the surface of the slope.

Figure 7 – Snow plate

3) From wet and wet snow - avalanche from "point" (Fig. 8). It has a teardrop-shaped avalanche beginning from the end of a rocky ledge on a snowy slope. The rocks, heated, fuel moisture grip snow from the rocky base, and it offers an avalanche. It is typical for the spring period.

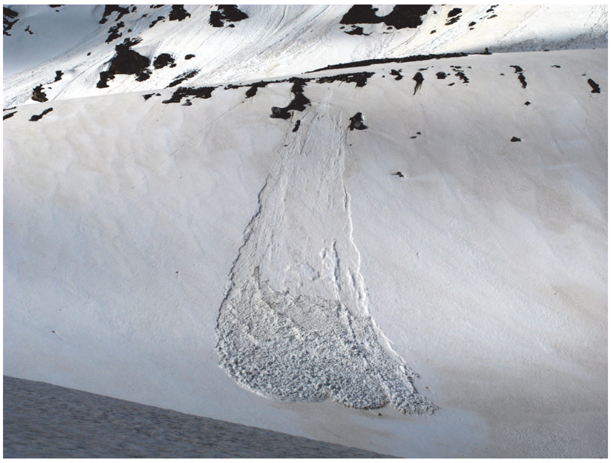


Figure 8 – Avalanche of wet snow

Depending on the state and properties of avalanche snow, six forms of avalanche movement are distinguished [10]:

1. The motion of a cloud of dusty (powdery) snow that may accompany or overtake a denser core.
2. The turbulent movement of low-cohesive dry snow is similar to the movement of bulk materials. During the movement of the friction between the particles decreases and there is a kind of "liquefaction" of the flow.
3. Movement of fragments of snow plates. The effect of air lubrication (air cushion) may occur between the bottom surface of the plate debris and the sliding surface.
4. The movement of the lumps of snow — sliding, rolling, stirring.
5. The movement of a consolidated mass of wet or wet snow is similar to the flow of a plastic substance or viscous liquid.
6. Turbulent movement of snow-water mixture, sometimes with admixture of soil particles and stones captured by the flow, like the movement of mudflows.

There may be intermediate cases and the transition of one form of movement to another during the avalanche path.

Robotic complex for remote monitoring of avalanche and mudflow dangerous mountain areas

To solve the problems of remote monitoring for the purpose of forecasting and early warning of the danger of avalanches and mudflows, it is proposed to create a robotic complex. The block diagram of the complex is shown in Fig. 9.

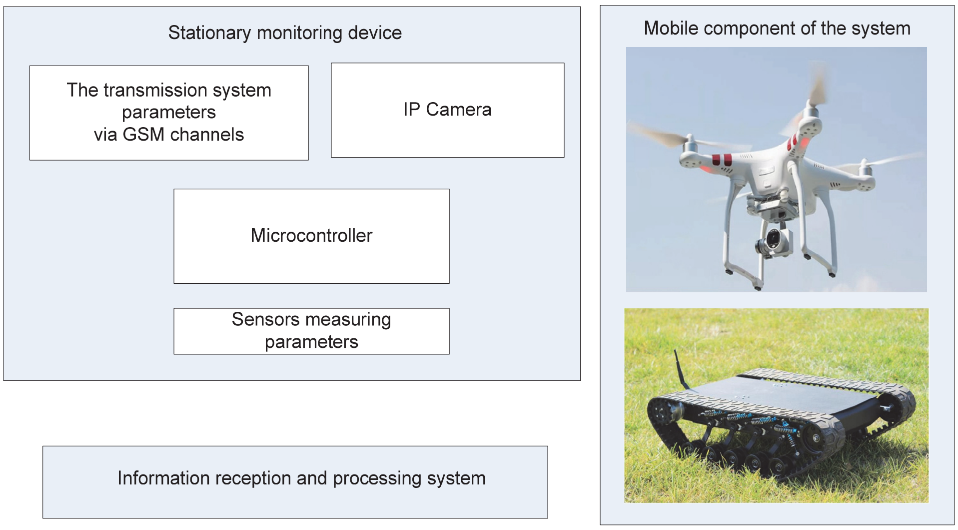


Figure 9 – Block diagram of a robotic remote monitoring complex

The complex includes a network of stationary monitoring devices; a system for receiving and processing information; a mobile component of the system (includes a high-pass robotic platform and a quadrocopter with a video camera).

The system of reception and processing of information consists of a remote computer connected to the Internet and a GSM modem. Video image from video cameras is received via the Internet, data from sensors is received via GSM modem, and remote control of the devices operation mode is performed.

The stationary monitoring device consists of an IP video camera, a control microcontroller, a GSM modem and a sensor system. The tasks of the device include receiving and transmitting video (in real time) and telemetry about the meteorological and hydrological situation in the monitoring areas. To transmit video and information, the device uses the channels of mobile operators. This type of communication imposes a restriction on the use of the device (placement is possible only in the coverage area of mobile operators, usually it is developed mountainous areas), but it gives a number of advantages: small size of equipment and low power consumption, low cost of the transceiver system, no restrictions on the range of information transmission. The block diagram of the device is shown in Fig. 10.

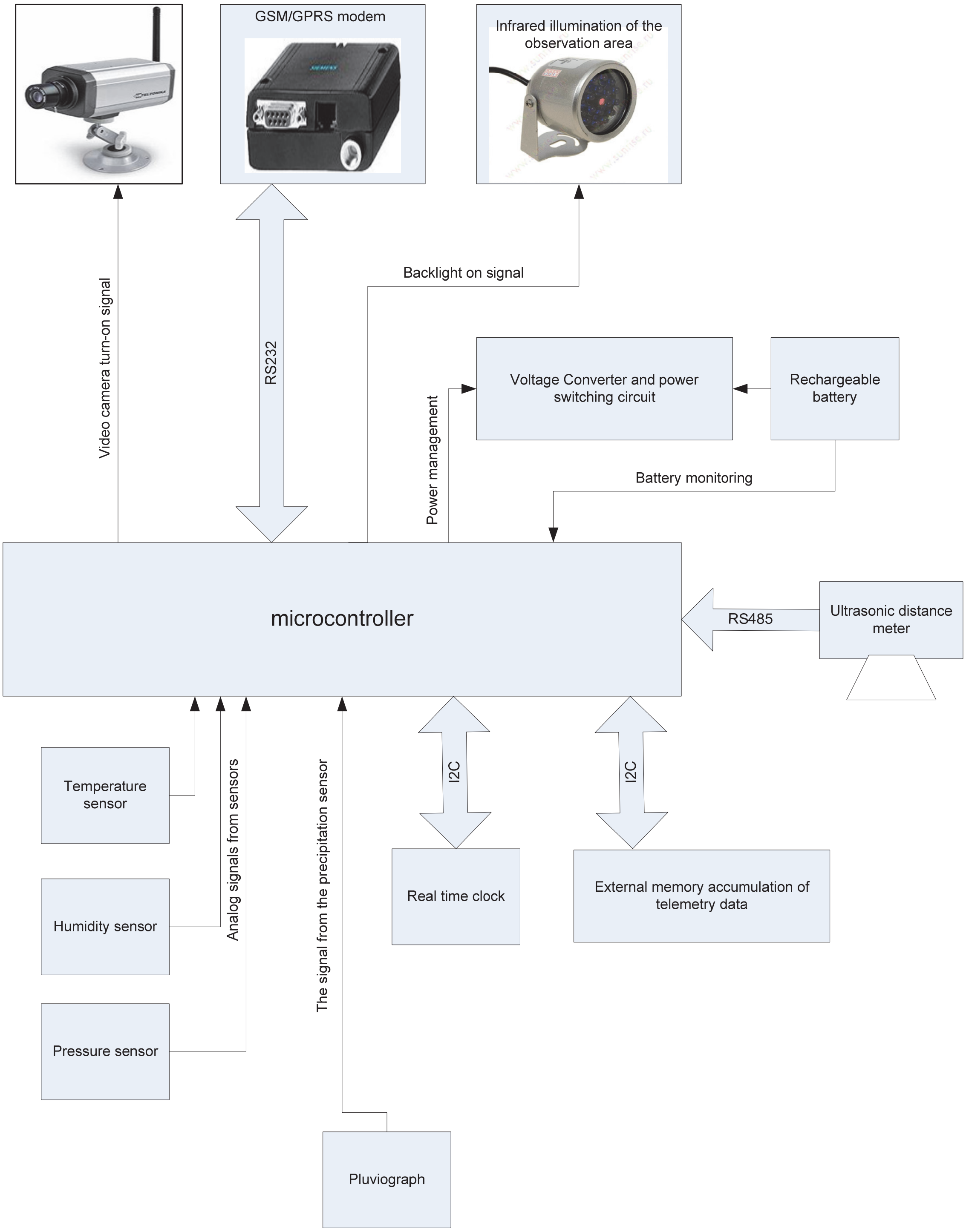


Figure 10 – Block diagram of a stationary monitoring device

The device is controlled by a microcontroller (MC). The objectives of the MC included:

– Inclusion of video cameras to transmit video monitoring station (at night time, provided by the inclusion of infrared light surveillance zone).

– Obtaining information from sensors of air temperature, atmospheric pressure, air humidity, measurement of the amount and intensity of precipitation.

– Obtain information about the water level in the river, or the height of the snow cover. To measure the level of the device uses an ultrasonic distance meter, the information from which the RS485 interface is supplied to the MC. Using the RS485 interface allows you to install the sensor in any convenient location - the maximum distance of 1000 ÷ 1200 m.

– All information received by the device is recorded on the external memory chip (connected to the interface I2C MC), for accumulation and subsequent transfer to the system of reception and processing of information.

– Communication with modem via RS 232 interface. The MC transmits telemetry data and receives control commands via the modem from the information reception and processing system.

– Monitoring the status of the battery, and control the supply of power to peripheral devices (video camera, infrared illumination, ultrasonic range meter).

A real-time clock module is installed in the device to record the time of measurements. The device is fully autonomous and is designed to operate without recharging batteries for 1-3 months (depending on the intensity of the system). Setting the mode of operation of the device is carried out remotely through the system of reception and processing of information.

The mobile component of the system includes a high-pass robotic platform and a quadrocopter with a video camera. The platform provides delivery of the aircraft to the specified zone. The use of a quadrocopter as part of the complex provides operational monitoring of long stretches and hard-to-reach areas.

References

1. Adzhiev A. H., Bolgov Y. V., Kondratyeva N. V., Senov H. M. A Hardware–Software Complexfor Remote Monitoring of Debris Flows. Pribory i Tekhnika Eksperimenta, 2016, No. 5, pp. 754–761. doi: 10.1134/S002044121604014X.
2. Wilson R.C., Mark R.K., Barbato G. Operation of a Real–time warning system for debris flows in the S. Francisco Bay Area, California. Proceedings of the ASCE 1993 National Conference on Hydraulic Engineering and International Symposium on Engineering Hydrology, San Francisco, 25–30 July 1993; Shen H. W.; Su S.T.; Wen F., Eds.; ASCE: New York, 1993; pp. 1592–1597.
3. Sejnova I.B., Tatjan L.V. Kriticheskie znachenija meteorologicheskih parametrov seleopasnyh situacij vysokogornogo rajona Centralnogo Kavkaza //Meteorologija i gidrologija. 1977. №12. s. 74–81. (in Russian).
4. Arattano M., Deganutti A.M., Marchi L. Debris Flow Monitoring Activities in an Instrumented Watershed of the Italian Alps. Proceedings of the First International Conference on Debris-flow Hazards Mitigation: Mechanics, Prediction, and Assessment, San Francisco, August 7–9, 1997; Chen C.; Ed.; Water Resources Engineering Division – ASCE: New York, 1997; pp. 506–515.
5. Arattano M., Marchi L. Systems and Sensors for Debris-flow Monitoring and Warning //MDPI. (www.mdpi.org/sensors) Sensors 2008, 8, p. 2436–2452. doi:10.3390/s8042436.
6. Hürlimann, M., Rickenmann D., Graf C. Field and monitoring data of debris-flow events in the Swiss Alps, Canadian Geotechnical Journal 2003, 40, 161–175. doi:10.1139/T02-087.
7. Abancó C., Hürlimann M., Moya J. Analysis of the ground vibration produced by debris flows and other torrential processes at the Rebaixader monitoring site (Central Pyrenees, Spain). Natural Hazards and Earth Systems Sciences 2013, 1, 4389–4423. doi:10.5194/nhess-14-929-2014.
8. Yin H.Y. Debris-flow monitoring and warning in Taiwan. International Workshop Monitoring Bedload and debris flows in Mountain Basins, Bozen-Bolzano (Italy) 10–12 October 2012.
9. Bozhunskij A.N., Losev K.S. Osnovy lavinovedenija. L.: GIMIZ, 1987. 280 s. (in Russian).
10. Vojtkovskij K.F. Lavinovedenie. –M.: Iz-vo MGU, 1989. -158 s. (in Russian).

M.V. Miroshkina, E.C. Briskin, V.A. Serov, S.A. Ustinov

ABOUT ENERGETICALLY EFFICIENT MODES OF THE WALKING ROBOTS MOVEMENT AT ITS DISPLACEMENT ALONG A SURFACE WITH OBSTACLES

Volgograd State Technical University, Volgograd, Russia
mariatiminen@yandex.ru

Abstract

There is the solution of one of the problems arising in the design of mobile walking robots: the problem of reducing energy costs on displacement of a walking robot, increasing the efficiency of a multi-legged mechanism and a reduction of the developed power due to a change in the mode of motion of the multi-legged walking robot.

Keywords: multi-legged walking robot, the gait of a multi-legged walking robot, minimum energy consumption.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (project No. 19-48-340018).

Mobile robots with walking movers have the ability to contactlessly overcome obstacles, which is one of the characteristic features of such machines [1]. The view of the trajectory of the moving walking propulsion, shown in fig. 1, is explained by the need to organize the interaction of the stop movers with the supporting surface without horizontal slipping along it in the absence of information about its profile. But with such information about the support surface, you can change the law of transfer of the foot. The combinations of these indicators can determine the boundaries of the Pareto-optimal traffic regimes for quality indicators.

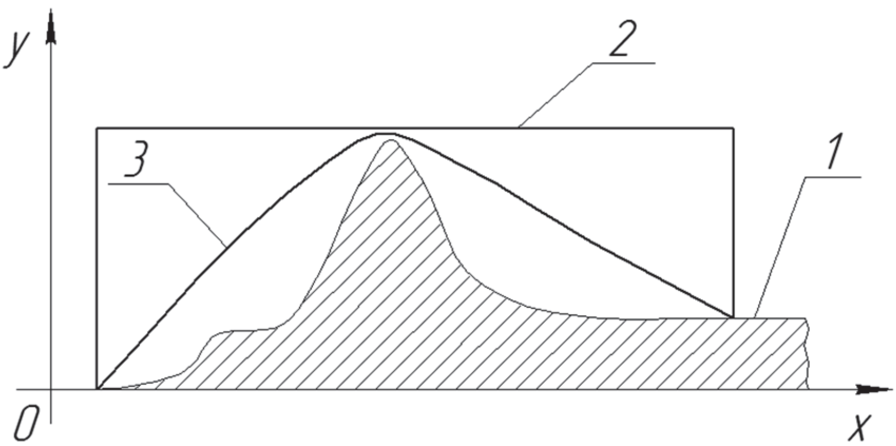


Figure 1 – Trajectory of transfer of robot foot
1 – profile of the support surface; 2 – absolute trajectory according to N.V. Umnov; 3 – possible trajectory

Formulation of the problem

We consider translational, with constant speed, movement of the robot body with orthogonal walking movers, such as Ortonog. The profile of the support surface, characterized by the distance S from the foot entering the transfer phase to the obstacle in the form of a projection H height, the distance L to the position of its projection at height h from the initial level (fig. 2) is considered known according to the information-measuring system of the robot. The task is to determine the laws of the transfer of the foot from the initial position to the final position, ensuring the minimum of the complex optimality criterion I, which is formed from the linear additive convolution of dimensionless partial indicators I_j.

Differential equations of motion of the foot in the horizontal and vertical direction and the equation describing the uniform translational motion of the robot body to determine the private dimensionless indicators I_j are

$$\begin{cases} m\ddot{x} = P \\ m\ddot{y} = T - mg \\ 0 = Q - F - P \end{cases} \tag{1}$$

where g is the acceleration of free fall, Q is the force of resistance to the movement of the robot; F, P are the forces developed by the directional motion drives, respectively supporting the movement of the body and ensuring the horizontal transfer of the foot to a new position, T is the force developed by the drive of the vertical movement of the foot.

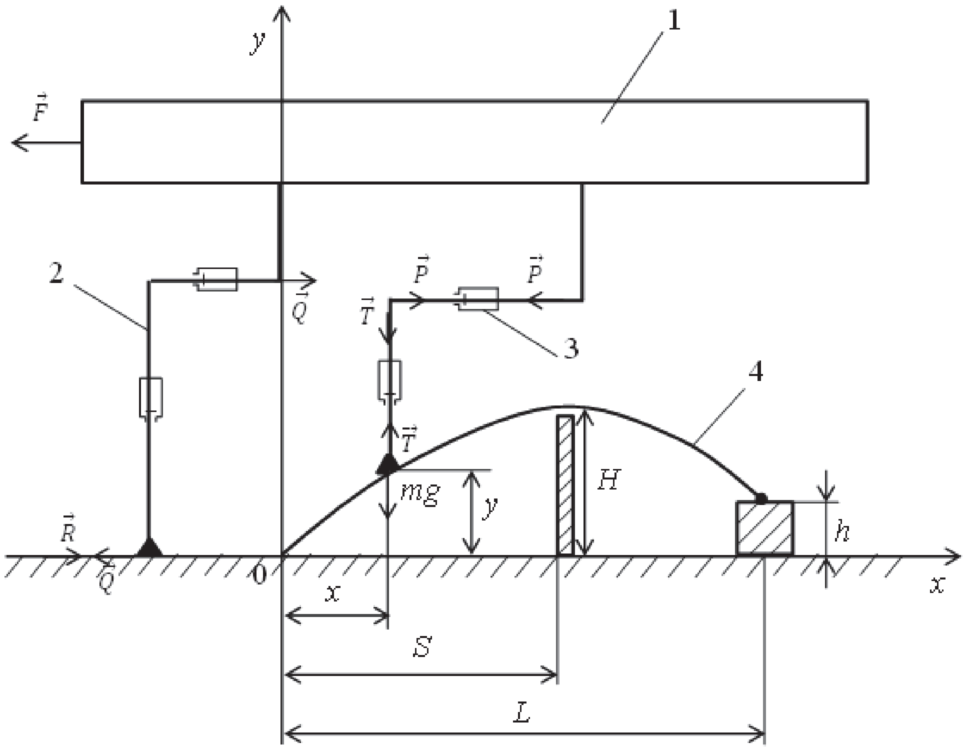


Figure 2 – Scheme for settlement
1 – the body of the robot; 2 – the mover, which is interacting with the support surface; 3 – the transferable walking mover; 4 – the trajectory of the movement the foot of the walking mover

Among the private indicators, the dimensionless level of heat losses in engines is considered

$$I_1 = \frac{2}{FL}\left(\alpha \int_0^{\tau} Q^2 dt + \alpha \int_0^{\tau} P^2 dt + \beta \int_0^{\tau} T^2 dt\right), \tag{2}$$

dimensionless root-mean-square accelerations of the foot during the transfer of I_2, I_3 are also considered

$$I_2 = \frac{1}{\tau}\int_0^{\tau} \ddot{x}^2 dt = \frac{1}{g}\sqrt{\frac{1}{\tau}\int_0^{\tau} \ddot{x}^2 dt},$$
$$I_3 = \frac{1}{\tau}\int_0^{\tau} \ddot{y}^2 dt = \frac{1}{g}\sqrt{\frac{1}{\tau}\int_0^{\tau} \ddot{y}^2 dt}, \tag{3}$$

where τ is the time of transfer of the foot, sec; L - step length, m; α, β - known engine characteristics.

The task is to determine the laws of motion x (t) and y (t), which will ensure both the overcoming of obstacles and the unstressed interaction of the portable foot with the supporting surface, and the minimum of the complex quality criterion

$$I = k_1 I_1 + k_2 I_2 + k_3 I_3 \tag{4}$$

where k_1, k_2, k_3 – subjectively introduced weight coefficients.

Mathematical model

The mathematical model is based on splitting the movement into two stages:

− the first stage corresponds to raising the foot on the section to the obstacle $0 < t < \tau_1$:

$$x(0) = 0, \quad \dot{x}(0) = 0, \quad y(0) = 0, \quad \dot{y}(0) = 0,$$
$$x(\tau_1) = S, \quad \dot{x}(\tau_1) = \dot{S}, \quad y(\tau_1) = H, \quad \dot{y}(\tau_1) = 0; \tag{5}$$

– the second stage corresponds to lowering the foot in the area after the obstacle $0 < t < \tau - \tau_1$:

$$x(0) = 0, \quad \dot{x}(0) = \dot{S}, \quad y(0) = H, \quad \dot{y}(0) = 0,$$
$$x(\tau - \tau_1) = L - S, \quad \dot{x}(\tau - \tau_1) = 0, \quad y(\tau - \tau_1) = h, \quad \dot{y}(\tau - \tau_1) = 0. \tag{6}$$

At each stage, the Euler-Poisson equations [5] are compiled for the variational problem

$$\delta I = \delta \int_0^{\tau} \Phi \, dt + \delta \int_0^{\tau - \tau_1} \Phi \, dt \tag{7}$$

with integrand function

$$\Phi = \frac{2k_1}{FL}\{\alpha[(m\ddot{x} + F)^2 + (m\ddot{x})^2] + \beta(m\ddot{y} + mg)^2\} + \frac{k_2}{g^2 r}\ddot{x}^2 + \frac{k_3}{g^2 r}\ddot{y}^2. \tag{8}$$

Then the Euler-Lagrange equations (9) are reduced to the form (10)

$$\frac{d^2}{dt^2}\left(\frac{\partial \Phi}{\partial \ddot{x}}\right) - \frac{d}{dt}\left(\frac{\partial \Phi}{\partial \dot{x}}\right) + \frac{\partial \Phi}{\partial x} = 0$$
$$\frac{d^2}{dt^2}\left(\frac{\partial \Phi}{\partial \ddot{y}}\right) - \frac{d}{dt}\left(\frac{\partial \Phi}{\partial \dot{y}}\right) + \frac{\partial \Phi}{\partial y} = 0 \tag{9}$$

$$\ddddot{x} = 0; \quad \ddddot{y} = 0 \; . \tag{10}$$

The solution at each of the j stages (j = 1, 2) has the form

$$x_j = x_{j0} + \dot{x}_{j0} t + C_j t^2 + D_j t^3$$
$$y_j = y_{j0} + \dot{y}_{j0} t + E_j t^2 + G_j t^3 \tag{11}$$

where the constants $x_{j0}, \dot{x}_{j0}, C_j, D_j, y_{j0}, \dot{y}_{j0}, E_j, G_j$ are determined from the initial conditions (5, 6)

$$x_{10} = 0; \quad \dot{x}_{10} = 0; \quad C_1 = \frac{3S}{\tau_1^2} - \frac{\dot{S}}{\tau_1}; \quad D_1 = \frac{\dot{S}}{\tau_1^2} - \frac{2S}{\tau_1^3};$$
$$y_{10} = 0; \quad \dot{y}_{10} = 0; \quad E_1 = \frac{3H}{\tau_1^2}; \quad G_1 = -\frac{2H}{\tau_1^3};$$
$$x_{20} = 0; \quad \dot{x}_{20} = \dot{S}; \quad C_2 = \frac{3(L-S)}{(\tau - \tau_1)^2} - \frac{2\dot{S}}{\tau - \tau_1}; \quad D_2 = \frac{\dot{S}}{(\tau - \tau_1)^2} - \frac{2(L-S)}{(\tau - \tau_1)^3};$$
$$y_{20} = H; \quad \dot{y}_{20} = 0; \quad E_2 = \frac{3(h-H)}{(\tau - \tau_1)^2}; \quad G_2 = -\frac{2(h-H)}{(\tau - \tau_1)^3} \tag{12}$$

A characteristic feature of the obtained equations (10) and their solutions is independence from the subjectively introduced weight coefficients k_j.

When substituting (12) in (11), particular indices are determined

$$I_1 = \frac{2}{FL}\int_0^{\tau_1} \alpha[(m\ddot{x}_1 + F)^2 + (m\ddot{x}_1)^2] + \beta(m\ddot{y}_1 + mg)^2 dt +$$
$$+\frac{2}{FL}\int_0^{\tau-\tau_1} \alpha[(m\ddot{x}_2 + F)^2 + (m\ddot{x}_2)^2] + \beta(m\ddot{y}_2 + mg)^2\, dt;$$
$$I_2 = \frac{1}{g^2\tau}\int_0^{\tau_1} \ddot{x}_1{}^2 dt + \frac{1}{g^2\tau}\int_0^{\tau-\tau_1} \ddot{x}_2{}^2 dt\,;$$
$$I_3 = \frac{1}{g^2\tau}\int_0^{\tau_1} \ddot{y}_1{}^2 dt + \frac{1}{g^2\tau}\int_0^{\tau-\tau_1} \ddot{y}_2{}^2 dt \quad (13)$$

and complex criterion I at the subjective assignment of weights coefficients k_1, k_2, k_3.

Mathematical Model Analysis

The developed mathematical model allows to conduct a fairly large range of research, to determine the influence of the external geometric characteristics of the support surface and the parameters characterizing the movement of the robot on the quality indicators. Indeed, setting the information and measuring system with the height of the obstacle H to be overcome and the lowering level h of the carried foot on the step length L characterizes the profile of the supporting surface, which affects on this quality indicators (13). Setting the total step time τ together with its length L determines the speed of the robot V, and also affects on indicators

$$V = \frac{L}{2\tau} \quad (14)$$

The physicomechanical properties of the support surface in the model are taken into account by the resistance force F. As is known [6], for walking machines, the average resistance force is determined by the expression

$$F = \sum_{j=1}^{N} \frac{P_{j\max}^2}{2cL} \quad (15)$$

where $P_{j\ \max}$ is the maximum normal reaction under the foot, c is the normal stiffness of the "foot-bearing surface" system, N is the number of propulsive devices.

The nature of the movement of the portable foot also affects the quality indicators, which is estimated by the time τ_1 and the horizontal speed of the foot at the moment it is above the obstacle. The parameters of the drive motors α, β, affecting the level of heat loss are also taken into account. These parameters are proportional to the active resistance of the windings for DC motors. In addition to the objective parameters, the indicators of motion quality of the robot are also affected by subjectively introduced weights coefficients k_1, k_2, k_3 or the intrinsic properties of the robot's control system embedded in its software, which forms together with the information-measuring system the type and nature of movement [6].

The results of mathematical modeling

Influence of the translational movement speed V of the walking robot "Ortonog" on the movement quality indicators

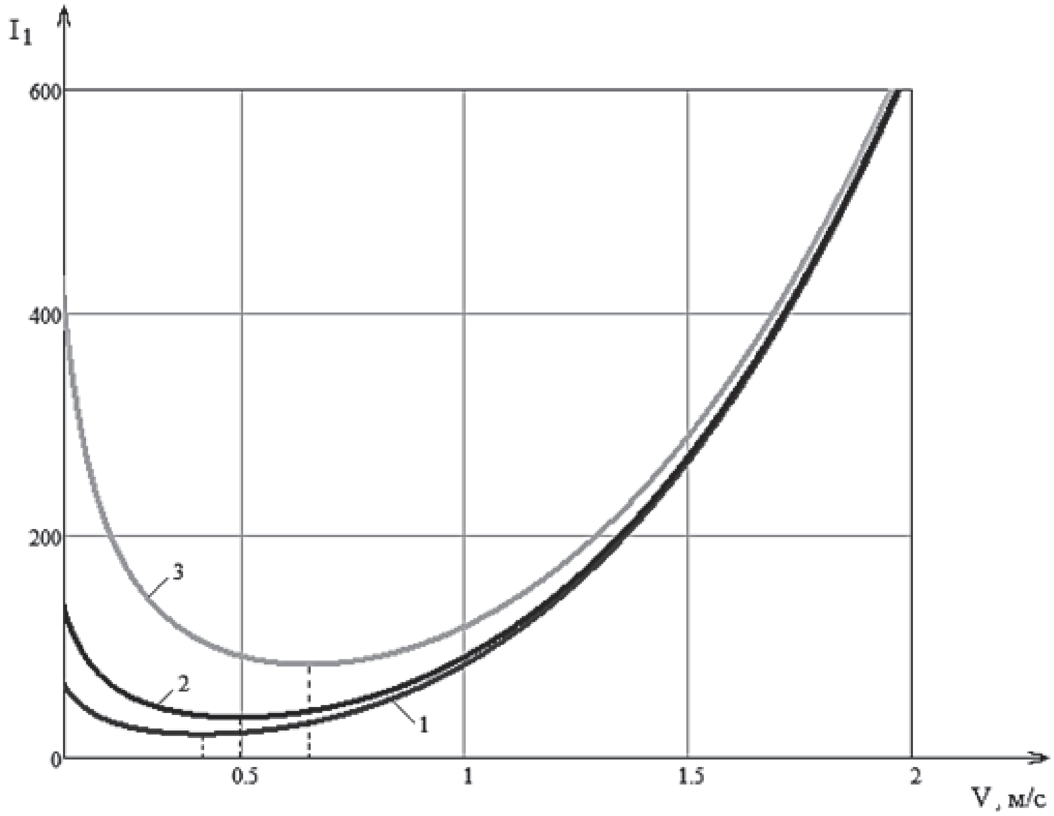


Figure 3 – Graphs of the dependence of heat loss (indicator I_1)
on the speed V at various forces of resistance F
1 – at F = 500 N; 2 – at F = 1000 N; 3 – at F = 2000 N

The translational motion of the walking robot "Ortonog" with parameters L = 0.91 m, m = 70 kg is considered. The robot moves along various horizontal deformable surfaces (h = 0) with different resistance forces F. At the same time, it overcomes an obstacle of height H located at a distance of $S = 0.3L$ from the foot of the propulsor entering the transfer phase. The speed of the foot at the time ($\tau_1 = \beta\tau$) of the passage of the obstacle $\dot{S} = \gamma V$. The total foot transfer time τ is determined in accordance with (14).

The graphs (fig. 3) show the dependences of heat losses (I_1) on the speed V for various resistance forces F.

Analysis of the graphs (fig. 3) shows that the level of heat loss depends on the physicomechanical properties of the supporting surface. The lower the soil stiffness, the lower the rate of heat loss is achieved at a lower speed. Therefore, when driving on different soils, there are optimal speeds that correspond to this minimum level.

Based on the above, it can be concluded that integral indicators have been formulated that characterize the quality of the laws of foot transfer in terms of the minimum of unproductive thermal losses in drive engines.

A mathematical model has been developed that allows to conduct a fairly large set of research, to determine the effect of the physical and mechanical properties of the soil, the external geometric characteristics of the reference surface and the parameters characterizing the movement of the robot on the quality indicators.

A method for determining the law of transfer of the foot of an orthogonal propulsion, which takes into account both the geometric and physical-mechanical properties of the environment, is also proposed.

The influence of the speed of movement of the robot on the quality indicators has been defined, that allows for each bearing surface, characterized by its physicomechanical properties and geometrical parameters, to determine the optimal speed by a complex optimality criterion.

The movement parameters of the walking robot "Ortonog" are determining at overcoming an obstacle to ensure the optimal value of the complex criterion.

The dependences of the level of heat losses on the velocity of the center of mass of the body of the robot with different resistance forces has been received, a characteristic feature of which is the presence of modes that provide a minimum of thermal loss.

A task was set and an algorithm was proposed for selecting Pareto-optimal software modes for moving the propulsion foot from the current position to the new robot by control system during motion self-planning.

References

1. Okhotsimskiy DE Mekhanika i upravlenie dvizheniem avtomaticheskogo shagayushchego apparata [Mechanics and Control of Automatic Walking Machine Movement]. Moscow, Nauka Publ., 1984. 312 p.
2. Ob upravlenii pokhodkoy shagayushchey mashiny «Vos'minog» [On the Control of the Walk of Step Machine "Eight-Feet"] 2008 E S Briskin. Mekhanika. Avtomatizatsiya. Upravleniye [Mechanics. Automation. Control] no 5 pp 6-10.
3. Ob energeticheski effektivnykh algoritmakh dvizheniya shagayushchikh mashin s tsiklovymi dvizhitelyami [On energy efficient motion algorithms for walking machines with cyclic thrusters] 2011 E S Briskin, Ya V Kalinin Izv. RAN. Teoriya i sistemy upravleniya [Journal of Computer and Systems Sciences International] no 2 pp 170-176.
4. Ob upravlenii adaptatsiyey ortogonal'nykh shagayushchikh dvizhiteley k opornoy poverkhnosti [On controlling the adaptation of orthogonal walking movers to the supporting surface] 2017 E S Briskin, Ya V Kalinin et al Izvestiya Rossiyskoy akademii nauk. Teoriya i sistemy upravleniya [Journal of Computer and Systems Sciences International] no 3 pp 184-190.
5. Elsgolts L E [Differential equations and calculus of variations]. Moscow, Nauka, 1969. 423 p.
6. Ob otsenke effektivnosti shagayushchikh robotov na osnove mnogokriterial'noy optimizatsii ikh parametrov i algoritmov dvizheniya [Assessment of the performance of walking robots by multicriteria optimization of their parameters and algorithms of motion] 2017 E S Briskin, Ya V Kalinin, et al Izvestiya RAN. Teoriya i sistemy upravleniya [Journal of Computer and Systems Sciences International] no 2 pp 168-176.

V.E. Pavlovsky[1, 2], M.V. Andreeva[2], E.Yu. Kolisnechenko[1],
I.A. Orlov[1, 2], A.P. Aliseychik[1], A.V. Podoprosvetov[1]

THE LOGISTICS SYSTEM CONSTRUCTED BY GROUP OF TRIANGULAR ROBOTS WITH OMNI-WHEELS

[1]M.V. Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences (KIAM), Moscow, Russia
[2]Moscow Automobile and Road Construction State Technical University (MADI), Moscow, Russia
vlpavl@mail.ru

Abstract

In article the control of the robot with three omni-wheels for creation of a specialized logistics system is studied. The feature of the elaborated robot is that he presents himself the platform in the form of a rectangular triangle. In work function of control is investigated and obvious formulas of the torques of forces, which need to be put to wheels for the movement along the given trajectory are given. Two special cases of the movement are considered: forward and the movement on a tangent to a trajectory in relation to the offered problem of logistics.

Keywords: omni-wheels, triangular platform, omnidirectional movement, control of the omni-robot.

Acknowledgments

Work is performed with assistance of the Russian Federal Basic Research Fund, projects 16-08-00880-a, 16-01-00131-a, 15-07-07483-a, 16-29-04412-ofi-m, and with assistance of the Program of RAS 1.31, the section "Current Problems of Robotics".

1. Introduction

Many works are devoted to robots on three omni-wheels [1–5]. Platforms in the form of an equilateral triangle are generally studied. This work considers a rectangular triangle. In this case, the equations have more irregular shape. In work, the mathematical model of such robot is studied. In particular, the torques which need to be put to wheels for the movement along any curve are investigated. At the expense of omni-wheels along a trajectory, it is possible to move in the different ways. Here two cases will be considered. The first is a progress when the robot does not turn in the course of movement. The second is the movement on a tangent to the chosen curve when the robot turns according to curvature of a trajectory. Both cases realize logistic function based on the considered robots**.**

2. Kinematics of the triangular omni-robot

Let us analyze kinematics of the robot on three omni-wheels. Let the case be a rectangular triangle with the parties of *a; b; c* (for definiteness of $c^2 = a^2 + b^2$) and wheels are located in its tops. Axes of wheels are directed along triangle medians (see an example to Fig. 1).

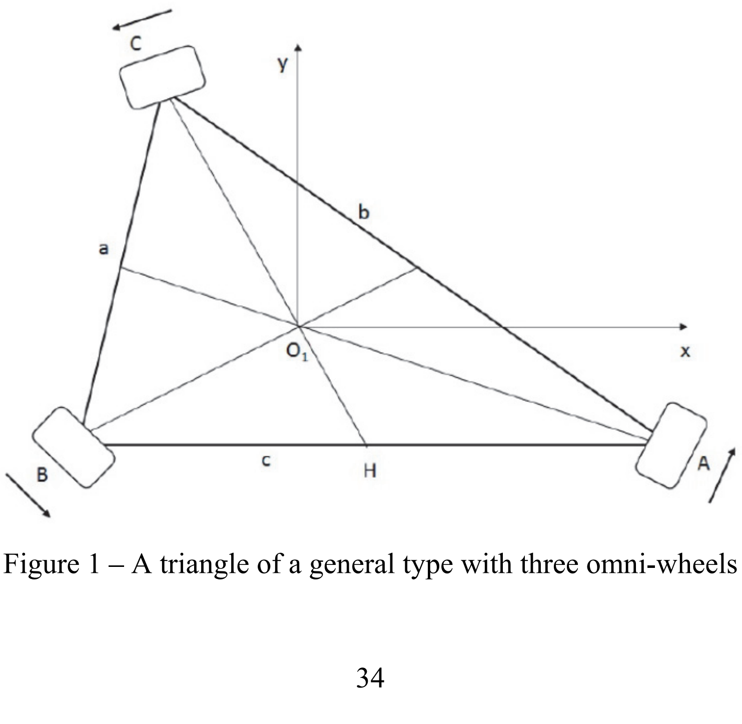


Figure 1 – A triangle of a general type with three omni-wheels

The geometrical center of a triangle (i.e. a point of intersection of medians) we will designate O_1, and the centers of wheels — A; B; C. Let us connect with the robot the system of coordinates O_1xy so that the axis of O_1x was parallel to the party of AB, as shown in Fig. 1. Thus, coordinates of the centers of wheels will be:

$$A=\left(\frac{a^2+2b^2}{3c},-\frac{a\sqrt{c^2-a^2}}{3c}\right),\ B=\left(-\frac{2a^2+b^2}{3c},-\frac{a\sqrt{c^2-a^2}}{3c}\right),$$

$$\left(C=-\frac{b^2-a^2}{3c},\frac{2a\sqrt{c^2-a^2}}{3c}\right).$$

Axes of wheels with an axis of O_1x make angles α_1; α_2; α_3. It is supposed that positive angular velocities rotate wheels counterclockwise:

$$\alpha_1=\pi-\arccos\frac{2b^2+a^2}{c\sqrt{4b^2+a^2}},\ \alpha_2=\arccos\frac{b^2+2a^2}{c\sqrt{4b^2+a^2}},\ \alpha_3=-\arccos\frac{b^2-a^2}{c\sqrt{a^2+b^2}}.$$

According to [6, 7] constraints are imposed on the device.

$$\begin{pmatrix}\dot q_1\\ \dot q_2\\ \dot q_3\end{pmatrix}=-\frac{1}{r}M\begin{pmatrix}v_x\\ v_y\\ \omega\end{pmatrix}$$

For the robot considered in this work

$$M=\begin{pmatrix}-\frac{ab}{c\sqrt{a^2+4b^2}} & -\frac{a^2+2b^2}{c\sqrt{a^2+4b^2}} & -\frac{\sqrt{a^2+4b^2}}{3}\\ -\frac{ab}{c\sqrt{b^2+4a^2}} & \frac{b^2+2a^2}{c\sqrt{a^2+4b^2}} & -\frac{\sqrt{b^2+4a^2}}{3}\\ \frac{2ab}{c^2} & \frac{b^2-a^2}{c^2} & -\frac{c}{3}\end{pmatrix}$$

M^{-1} matrix is necessary for the solution of a direct problem of kinematics (i.e. calculations of velocities of the movement of the robot and of a trajectory on angular speeds of rotation of wheels). It in this task this matrix is as follows.

$$M^{-1}=\begin{pmatrix}-\frac{a^3\sqrt{a^2+4b^2}}{3bc^3} & -\frac{b^3\sqrt{b^2+4a^2}}{3ac^3} & \frac{a^4+3a^2b^2+b^4}{3abc^2}\\ -\frac{(b^2+3a^2)\sqrt{a^2+4b^2}}{6c^3} & \frac{(a^2+3b^2)\sqrt{a^2+4b^2}}{6c^3} & \frac{b^2-a^2}{6c^2}\\ -\frac{\sqrt{a^2+4b^2}}{2c^2} & -\frac{\sqrt{a^2+4b^2}}{2c^2} & -\frac{1}{2c}\end{pmatrix}$$

3. Dynamics of the triangular omni-robot

Let us designate $\vec q=(q_1,q_2,q_3)^T$, $\vec z=(v_x,v_y,\omega)^T$.

Kinetic energy (it is Lagrangian in this task) has the following appearance:

$$T=\frac{1}{2}mv^2+\frac{1}{2}I\varpi^2+\frac{1}{2}\sum_{i=1}^{3}I_i\dot q_i^2$$

where m – is the full mass of the device, v2 = vx2 +vy2, I – the full moment of inertia concerning a point of O1, Ii, i = 1; 2; 3 – moment of inertia of i-th wheel of rather corresponding axis of a wheel.

We assume that wheels are identical each other therefore I1 = I2 = I3.

Using Lagrange's equations with uncertain multipliers (see [8]); we will receive dynamics equations for this device:

$$\left(A+\frac{I_1}{r^2}M^TM\right)\dot{\vec{z}}+\Gamma\omega\vec{z}=-\frac{1}{r}M^T\vec{\mu},$$

where – $\vec{\mu}=(\mu_1,\mu_2,\nu_3)^{\mathrm{T}}$ is a vector of the moments attached to wheels $A=diag\left(m,m,I\right)$,

$$\Gamma=\begin{pmatrix}0 & -m & 0\\ m & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Last equation

$$I+\frac{I_1\left(2a^2+2b^2\right)}{3r^2}\dot{\omega}=\frac{1}{3r}\left(\sqrt{4b^2+a^2}\mu_1+\sqrt{4a^2+b^2}\mu_2+\sqrt{a^2+b^2}\mu_3\right)$$

has the explicit solution:

$$\omega=\frac{r}{3r_2I+I_1\left(2a^2+2b^2\right)}\left(\sqrt{4b^2+a^2}\int_0^t\mu_1dt+\sqrt{4a^2+b^2}\int_0^t\mu_2dt+\sqrt{a^2+b^2}\int_0^t\mu_3dt\right)$$

The moments necessary for the movement on the set trajectory $\tilde{z}(t)$:

$$\vec{\mu}=-r\left(M^T\right)^{-1}\left(\left(A+\frac{I_1}{r^2}M^TM\right)\dot{\vec{z}}+\Gamma\omega\vec{z}\right)$$

For the free movement $\left(\vec{\mu}\left(t\right)=0\right)$ the first integrals are $\omega=const_1$ and

$$\left(m+\frac{I_1}{r^2}\left(M^TM\right)_{11}\right)v_x^{\,2}+2\frac{I_1}{r^2}\left(M^TM\right)_{12}v_xv_y+\left(m+\frac{I_1}{r^2}\left(M^TM\right)_{22}\right)v_y^{\,2}=const_2.$$

Let us notice that a matrix M^TM is symmetric one and for given arrangement of wheels

$$\left(M^TM\right)_{13}=\left(M^TM\right)_{23}=0.$$

For the constant moments, let us investigate stationary movements.

For constant angular speed, ω_0 the following condition is required

$$\left(\sqrt{4b^2+a^2}\int_0^t\mu_1dt+\sqrt{4a^2+b^2}\int_0^t\mu_2dt+\sqrt{a^2+b^2}\int_0^t\mu_3dt\right)$$

otherwise ϖ = ϖ (t).

Let us rewrite dynamics equations:

$$A'\begin{pmatrix}\dot{v}_x\\ \dot{v}_y\end{pmatrix}+B'\begin{pmatrix}v_x\\ v_y\end{pmatrix}=\begin{pmatrix}u_1\\ u_2\end{pmatrix},$$

where

$$A' = \begin{pmatrix} m + \frac{I_1}{r^2}\left(M^T M\right)_{11} & \frac{I_1}{r^2}\left(M^T M\right)_{12} \\ \frac{I_1}{r^2}\left(M^T M\right)_{21} & m + \frac{I_1}{r^2}\left(M^T M\right)_{22} \end{pmatrix}, \quad B' = \begin{pmatrix} 0 & -m\omega_0 \\ m\omega_0 & 0 \end{pmatrix},$$

u_1; u_2 – the record of the right parts of the equations of dynamics reduced for convenience (u_1 and u_2 – are constant). Stationary solutions:

$$v_x = \frac{u_2}{m\omega_0}, \qquad v_y = -\frac{u_1}{m\omega_0} ..$$

4. Movement on a trajectory

Thanks to a design of wheels, the robot can move along the set trajectory in various ways. Two the most interesting is a forward (translational) movement and the movement on a tangent to a curve.

At first, we will consider the movement of the robot if he does not rotate around an axis O_1z. Therefore, $\varpi = 0$. Let us assume that the center of the robot already is in a certain curve point and axis O_1x makes some angle β with axis $O\xi_1$. According to requirements of a task, angle β has to remain constant. Then

$$v_x = \xi_1 \cos\beta + \xi_2 \sin\beta, \quad v_y = -\xi_1 \sin\beta + \xi_2 \cos\beta \quad .$$

In these expressions it is necessary to substitute the trajectory equation in a parametrical form. From (2) necessary moments for realization of this movement are obtained

Now we will consider the movement of the robot along a curve provided that in each time an axis O_1x is directed on a tangent to a trajectory. Then $\varpi = kv$, where k – path curvature.

Let us assume that at the initial moment the robot already is located so that an axis O_1x and a tangent to a trajectory coincide. Then

$$v_x = \dot{\xi}_1, \quad v_y = \dot{\xi}_2, \quad \omega = k\sqrt{\dot{\xi}_1^2 + \dot{\xi}_2^2} \quad .$$

The moments for movements of this kind are also expressed from (2).

5. Control

To verify model and to consider the impact of possible slipping the prototype of the robot on three omni-wheels with electric motors was created (see Fig. 2).

Main goal of prototyping is synthesis of control for the movement along various trajectories and comparison of results with results of mathematical modeling. Use of omni-wheels simplifies the kinematic scheme of the apparatus because there is no need for difficult steering devices, at the same time there is a possibility of controlling of the curvilinear movement. Any arrangement of wheels allows to generalize the theory of robots of the omnidirectional movement.

The hardware and program architecture of a control system of the mobile robot was developed based on the STM32F4 microcontroller. MATLAB Simulink was chosen as the environment of modeling and programming of a control system. For modeling of kinematics and dynamics of the robot the program complex "Universal Mechanism" was used The universal system of low-level control for electric motors with feedback on sensors for assessment of possible wheel slip and to sensors of angular speeds of wheels on the basis of Hall's effect was developed for the multipurpose mobile platform in the form of flowcharts for multidomain modeling and model design in MATLAB Simulink. For feedback for correction of positioning the onboard navigation, system is used. Navigation is carried out by means of beacons based on ultrasonic sensors of distance.

6. Logistics system on basis of triangular robots

For transportation of non-standard or heavy freights the basic robots described above have to unite in group. It is supposed that robots unite side to sides of the cases without gaps so that completely to cover a freight projection to the plane parallel to the transportation plane. Mathematically it is a problem of a covering or the task of tiling (a task about parquets). In work, its solution for a problem of tiling of a projection of the transported freight is discussed. The connected robots keep in the connected state or by means of special control of their movement, or with electric locks on sides. Modeling and prototyping of a system is executed, results are effective and presented in the report.

The idea of a system is shown below in Fig. 2 and fig. 3. Fig.2 – the triangular omni-robot, a view from the part of running wheels. Fig. 3 is a freight projection covering triangular robots; the additional robots not obligatory in a covering are highlighted with color.

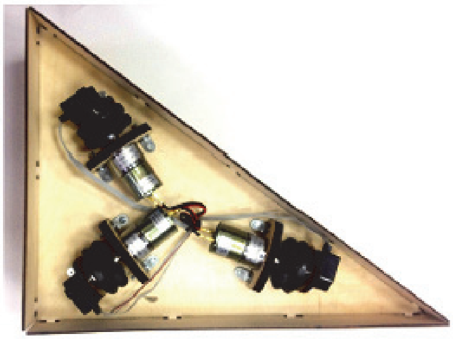


Figure 2 – Prototype of the robot

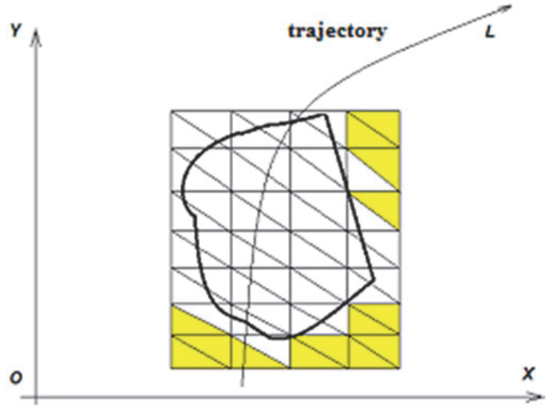


Figure 3 – Example of transportation of bulky goods

7. Conclusion

In article the kinematics, dynamics and control of the triangular device on omni-wheels are described. The explicit formulas of the moments necessary for the movement along the set any curve are given in work. Testing of such control systems was carried out on a prototype. On the basis of the described robots the reconstructed logistics system intended for transportation of non-standard freights is designed. Modeling confirmed its efficiency.

References

1. Lin L. and Shih H. Modeling and Adaptive Control of an Omni-Mecanum-Wheeled Robot. Intelligent Control and Automation. v.4. 2013. N.2. pp. 166–179.
2. Taheri H., Qiao B. and Ghaeminezhad N. Kinematic Model of a Four Mecanum Wheeled Mobile Robot. International Journal of Computer Applications. v.113. 2015. N.3. pp. 6–9.
3. Williams R.L., Carter B.E., Gallina P., Rosati G. Dynamic model with slip for wheeled omnidirectional robots. IEEE Transactions on Robotics and Automation. Volume 18. Issue 3. 2002. pp. 285 – 293.
4. Ashmore M., Barnes N. Omni-drive Robot Motion on Curved Paths: The Fastest Path between Two Points Is Not a Straight-Line. AI 2002: Advances in Artificial Intelligence, 15th Australian Joint Conference on Artificial Intelligence Canberra. Australia. December 2–6. 2002 Proceedings. pp. 225–236.
5. Balakrishna R., Ghosal A. Modeling of slip for wheeled mobile robots. IEEE Transactions on Robotics and Automation. v. 11. Issue 1. 1995. pp. 126 – 132.
6. Gfrerrer A. Geometry and kinematics of the Mecanum wheel. Comput. Aided Geom. Design 25. 2008. N.9. pp. 784–791.
7. Zobova A.A., Tatarinov Ya.V. The dynamics of an omni-mobile vehicle. Journal of Applied Mathematics and Mechanics. 2009. Vol. 73. № 1. pp. 8-15. (In Russ.).
8. Borisov A. V., Kilin A.A., Mamaev I.S. An omni-wheel vehicle on a plane and a sphere. Russian Journal of Nonlinear Dynamics. 2011. Vol 7, No 4. pp.785–801. (In Russ.).

S.M. Sokolov, A.A. Boguslavsky, N.D. Beklemishev

IMPLEMENTATION OF INTERPRETIVE NAVIGATION USING VISION SYSTEM MODULES

M.V. Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences (KIAM), Moscow, Russia
sokolsm@list.ru, anbg74@mail.ru, n_bekl@mail.ru

Abstract

The article considers the implementation of the interpretive navigation method using an information system based on a vision system. The interpretive navigation approach makes it possible to effectively solve problems in ensuring targeted movement of mobile robots which has received the abbreviated name of SLAM (simultaneous localization and mapping). Interpretive navigation excludes the collection and processing of large three-dimensional data arrays and enables the concentration of resources on identifying and processing features – landmarks in the environment. Combined with a wide range of unified modular software/hardware architecture of real-time vision systems, interpretive navigation makes it possible to provide the necessary information for mobile devices with a high autonomy and fully autonomous ones. The paper describes the solution of a navigation problem in various conditions when a robot moving into a room with known landmarks.

Keywords: mobile robots navigation, interpretive navigation, vision systems, markers, landmarks.

Acknowledgments

This work is partially supported by the RFBR grant no. 19-07-01113.

Introduction

The interpretive navigation method (IN) is being developed in the Keldysh Institute of Applied Mathematics RAS since 1970s. The essence of this method is that the position of a mobile robot on the ground is determined not in the Cartesian coordinate system (CCS), but on the basis of the descriptions of the environment in the language of the visible environmental features – landmarks and their changes during movement [1-5]. In this case, an analogue of the CCS quantitative model is a different, qualitative model called the information equivalence graph. In such description, the vertices of the graph correspond to connected areas of the terrain with the same informational and visual content – information equivalence areas (IEA), and the edges correspond to changes in these descriptions when going from one local area to another. In this case, the organization of a robot movement is based on a fixed set of recognition and movement actions. For ground and air robots, the most complete information support system forms the so-called "navigation cross", including: global satellite system; conventional navigation system consisting of an inertial navigation, dead reckoning and landmarks adjustment system; interpretive (or quality) navigation and operator. The integrated use of all components of this cross provides a solution to the navigation problem in most practical cases but requires a variety of infrastructural and sensory support. Vision systems of different ranges represent the most complete and reliable information for identifying and determining the relative objects location in the environment and robot. IN offers one of the ways to implement the SLAM concept [6, 7] – a common methodology for solving the two tasks: 1) providing environment knowledge (a map); 2) determining a robot trajectory on this map.

We focus our efforts on building an effective (in terms of price/quality ratio) unified modular hardware and software architecture of the vision system (VS) for solving a wide range of tasks of movements of vehicles with a high autonomy and fully autonomous ones in difficult operating conditions when it is not possible to use the whole variety of navigation tools. The article describes studies in which three types of VS modules (monocular, stereo and omnidirectional) are involved, solving navigation problems in a different formulation based on the method of interpretive navigation. A set of artificial markers is used as landmarks which makes it possible not to pay much attention to the issues of identifying and recognizing landmarks in the environment at this stage of research. We consider the directions of further work on studying the concept of interpretive navigation based on a unified modular VS.

The architecture of an onboard vision system

For mobile robots with an enhanced autonomy, we have developed a unified modular real-time software and hardware architecture of VS. The hardware part of the onboard VS architecture is formed by reconfigurable combinations (network) of a data registering units (DRU) and a processing and control units (PCU), and the software part – by a large-scale framework of real-time VS software [8-10]. Fig. 1a shows the general structural diagram of this framework, Fig. 1b shows the core frame structure of real-time VS software.

Algorithmic software

The theoretical foundations of the IN are formed by the introduction of the mathematical structure – terrain [11, 12]. This space provides a good basis for the study of the information and motor behavior of mobile robots and distributed mobile systems. To solve applied problems, mathematical foundations allow optimizing the onboard knowledge/data base replenishing it as required with permissible volume portions, rationally composing and using various data acquisition modules to highlight/define reference points in the environment and process the collected data taking into account the existing knowledge. Analytically obtained ratios can optimize the algorithmic support of data collection and processing when solving specific problems. For ground and air robots, the most complete information support system forms the so-called "navigation cross", including: global satellite system; conventional navigation system consisting of an inertial navigation system, dead reckoning and landmarks adjustment system; interpretive (or quality) navigation and operator. The complexing of all components of this cross provides a solution to the navigation problem in most practical cases but requires a variety of infrastructural and sensory support. Vision systems provide the most complete and reliable information about the identification and relative location of objects in the environment and the robot.

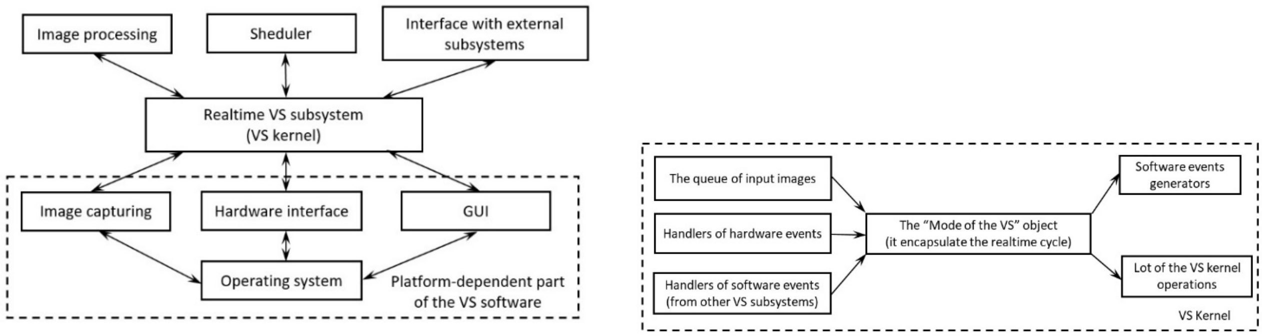


Figure 1 – (a) The general diagram of real-time VS software. (b) The VS kernel structure

One of the main requirements/trends in the creation of modern mobile robots with enhanced autonomy is the unification of their software. Software unification improves the reliability of operation, reduces the cost of creating robot options, increases the economic viability of use. The issues of algorithmic software implementation are one of the main directions of our research. As already mentioned, the basis for constructing the entire software of the information management system of the robot is the framework of the real-time VS software. In terms of knowledge representation, task formation and organization of information and movement actions – technologies for working with graphs. For the preparation of an onboard database and the generation/filling of tasks structure, well-known libraries and environments for working with graphs are used [13-17]. In the onboard control system, its information support uses a special implementation of the DBMS. It actively uses polymorphism, a technique known in object-oriented programming: the same mechanism is used both to represent the entire information equivalence graph (IEG) – the robot mission map, and to describe specific information equivalence areas (IEA) and individual landmarks (a set of related features).

Examples of the implementation of navigation information support in various conditions

The selection of landmarks in the environment is one of the key objectives of ensuring targeted movements. At the described stage of research, we simplified the task of identifying landmarks and used special tags – markers. We analyzed the known systems of markers [18-20] and decided to prepare our own set, reliable and able to make improvements. Fig. 2 shows a set of markers located for system adjustment and calibration.

The proposed markers set was investigated from the point of view of the possibility of reliable recognition in laboratory conditions: providing observation from different distances and at different angles, subject to restrictions on the size of the markers themselves (ease of manufacture and the possibility of easy placement in the room). As a result of these studies, reliable recognition results were obtained: for a video camera with a field of view of 60° and a resolution of 2464 × 2056 pixels, the observation distance is 1.5-4 meters, the angles of the main optical axis to the normal to the marker plane are ± 45°; for an omnidirectional camera with a field of view of 360° × 180° the resolution of 2592 × 1920 pixels, the observation distance is 1.5-3m, angles are ±30°.

Using this marker set we investigated two modes of information support for targeted movements. First: to observe landmarks with known absolute coordinates, the trajectory of the moving robot is determined in a given coordinate system. To solve this problem, a VS with one field of view is used (monocular DRU with a viewing angle of about 60° or an omnidirectional one with a viewing angle of about 180°). Second: the absolute coordinates of the landmarks are unknown in advance. In this case, a VS with a stereo module is used. This allows us to determine the 3D coordinates of the landmarks in the coordinate system associated with the initial position of the robot. Further, the robot trajectory in the selected relative coordinate system is constructed similarly to the first mode using only one camera. When entering data on the initial position of the robot or specifying the absolute coordinates of three of the landmarks, it is possible to make an absolute binding of all landmarks and transfer the data to a traditional GIS.

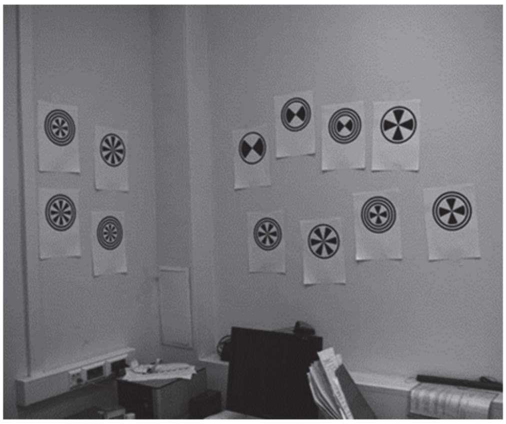


Figure 2 – A set of markers located on the walls of the room
(on the left – the image in the left camera of the stereo module,
on the right – in the omnidirectional module)

Figure 3 – Laboratory rooms for experiments with a robot.
Above – room with number 219, below – 220

Fig. 3 shows a panorama of the laboratory rooms where the movement of the mobile robot equipped with the described VS took place. Fig. 4a shows an IEG describing a map of the building second-floor where the laboratory rooms are located. Fig. 4b presents a graph describing the IEA – the second of the laboratory rooms.

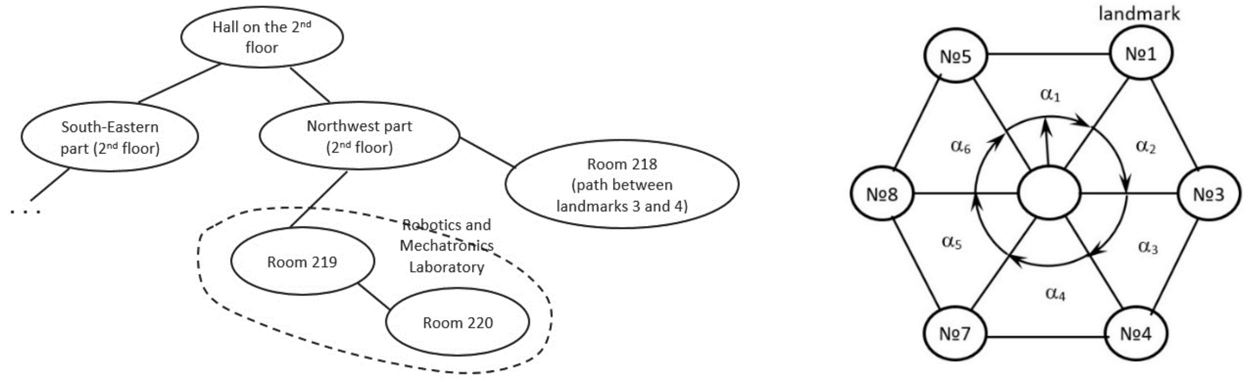


Figure 4 – (a) An example of an IEG for the area of a mobile robot functioning.
(b) IEA graph for the description of laboratory room 220

The sequence of processing visual data obtained from a single field of view (graph of the marker recognition).

[Rings determination]: Contour selection (Canny operator) and their vector description. ⇒ Selecting curves (ellipses). ⇒ Selecting ellipses having a common center. ⇒ Selecting objects having from 2 to 6 contours of ellipses having a common center. ⇒ Checking the relationship of the selected object ellipses sides. ⇒ If there is no general relationship, such an object is discarded.

[Counting "spokes" – black sectors]: Only image areas within the selected boundaries of candidate objects are processed. The interior of the candidate object in the source frame is converted into a circle by the affine transformation in accordance with the ratio of the ellipse axes. The intensities of the pixels of this circle are summed in a one-dimensional array of length $N = 128$ in accordance with what angle is formed by the pixel radius vector relative to the center point with the horizontal (in gradations $2\pi/N$).

The resulting array is transformed by fast Fourier transform (FFT), and the number of sectors is located at the position of the maximum modulus point of the Fourier transformed array.

This method contributes to a reliable recognition of tags of this type at angles between the plane of the marker and the normal to the line of sight on the center of this marker to 45°, as shown in the figure below. In this case, the imaging resolution, the distance to the tags and the size of the images must be such that the thickness of the black rings on the frame was at least 3 pixels.

In parallel with the recognition of markers, angles between the plane of the marker and the normal to the line of sight at the center of this marker for each of them are determined in this algorithm. Fig. 5 shows a photo of markers in the field of view of a stereo module left camera with an indication of a certain angle between the plane of the marker and the normal to the line of sight at the center of this marker.

Figure 5 – Angles between the plane of the marker and the normal to the line of sight at the center of this marker for recognized markers found in the recognition algorithm based on data obtained from one field of view of the forward-looking camera

Using the stereo module (base 0.56 m), the three-dimensional coordinates of each of the markers in the coordinate system associated with the left camera of the stereo system were found. This coordinate system was used to calculate the track shown in Fig. 6. The movement of the VS in this experiment was implemented by the manual movement, which determined the “bounce” in the trajectory. To get the track, only the left camera record was used. Marker recognition was performed on each frame. If, among the recognized markers, there were at least 4 markers with known 3D coordinates calculated by the stereo pair, the coordinates and angles determining the camera position for this frame were found by solving the inverse photogrammetric problem (PnP problem) [21, 22].

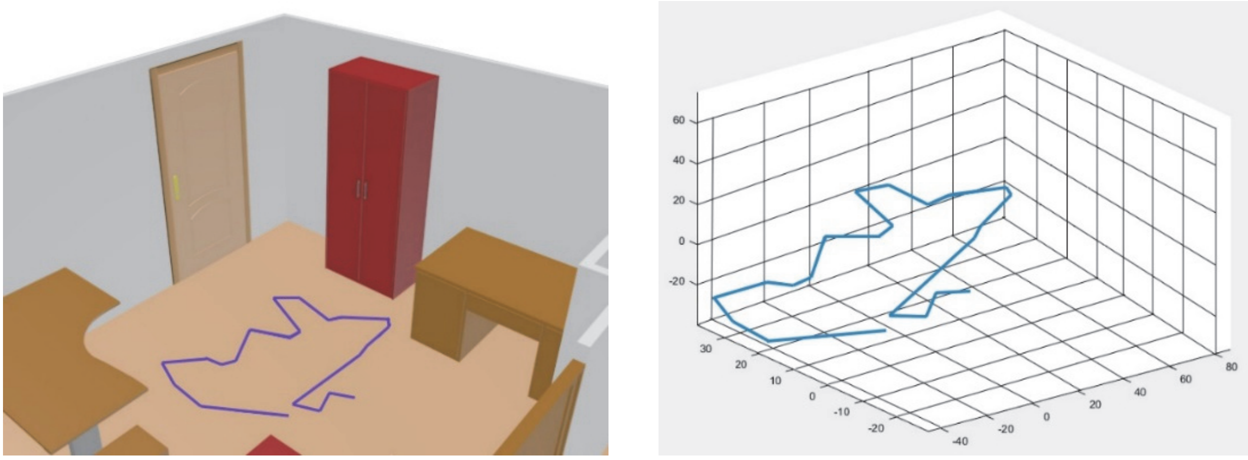


Figure 6 – The robot trajectory estimated by VS. On the left – in the projection on the laboratory room floor plane. On the right – in the 3D coordinate system.
A laboratory room layout was used as a GIS map

When solving the PnP problem, residual tolerance was set to 10 pixels. If the solution found for all markers did not fit into the tolerances, then one of the markers was discarded in the cycle of the markers recognized on the frame, and the problem was solved by the remaining markers. In the experiments, only one of the markers presented from 464 frames processed was incorrectly identified.

Quantitative estimates – the frequency of obtaining navigation data using the laboratory model of the VS is as follows. The accuracy of determining the position of the mobile robot on which the VS modules are installed is ±1 cm (this accuracy was estimated using manual measurements of individual static VS positions with a laser distance meter), the frequency of data acquisition was 10 Hz (limited by the frame rate of high-resolution video cameras with the GigE interface). At the described stage of research, optimization of the used implementations of algorithmic support was not performed. The above results were obtained after combining individual algorithms in the framework of the real-time VS software.

Conclusion

As the example of solving a model problem shows, the use of computer vision systems in combination with interpretive navigation approaches demonstrates good results. The interpretive navigation method provide the possibility of rational acquisition and processing of visual data and solving navigation problems by onboard computational tools in real time. The modular arrangement of a VS on the framework of the real-time VS software achieves the unification of software tools, offers a basis for comparing and quantifying the algorithmic support of navigation tasks. Operational maneuvering in the selection of information support tools for specifically targeted movements allows generating reliable and economically feasible information support systems based on real-time VS software.

More success can be expected when using this approach in the natural environment. Here, the algorithmic support for the selection of landmarks comes to prominence. Such studies are already being actively conducted, as well as by the authors of this paper. In the near future, the application of the method of interpretive navigation will be implemented for autonomous movements in the natural environment without adding special tags.

References

1. Platonov A.K., Kirilchenko A.A., Kugushev E.I. Using local landmarks to determine the position of a mobile robot (in Russian). // “Problems of machine vision in robotics”, M.: KIAM USSR, 1981, p. 31-47.

2. Kirilchenko A.A. Interpretation of local relative medium descriptions by a mobile robot (in Russian). / Preprint of the KIAM USSR AS, 1983, No. 149, 28 p.
3. Sokolov S.M. The use of photometric information in the complex of an integrated locomotion robot (in Russian). – In the book of Robotic systems control and sensing. M.: Nauka, 1983, p. 150-163.
4. Kirilchenko A.A., Platonov A.K., Sokolv S.M. Theoretical aspects of the organization of mobile robot interpretive navigation. (in Russian). Preprint of the KIAM RAS No.19, Moscow, 2008.
5. Kirilchenko A.A., Zueva E.Yu., Platonov A.K., Sokolov S.M. Formal approaches to the design of algorithms for information support of mobile systems (in Russian). Preprint of the KIAM RAS No.5, Moscow, 2002.
6. Mur-Artal Raúl, Tardós Juan D. ORB-SLAM: Tracking and mapping recognizable features // MVIGRO Workshop at Robotics Science and Systems (RSS), Berkeley, USA. – 2014.
7. SLAM in realistic environments. http://www.nada.kth.se/utbildning/forsk.utb/avhandlingar/lic/
8. Sokolov S.M., Boguslavsky A.A. Intellectual Images Processing for a Realtime Recognition Problem. // Proc. The 2nd Intern. Multi-Conf. on Complexity, Informatics and Cybernetics (IMCIC2011), Orlando, Florida, USA, March 27th-30th, 2011, Orlando, Florida, USA, Vol. II, pp. 406-411.
9. Sokolov S.M., Boguslavsky A.A. Methodological aspects for the development of information systems of unmanned mobile vehicles. // Proc. of the 13th Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO2016) July 29-31, 2016, vol.2 pp. 492-498.
10. Sokolov S.M., Boguslavsky A.A., Podtelkina O.A., Dmitriev A.A., Ikonnikov M.A., Petrov A.A. Real-time computer vision system software framework (in Russian). // Certificate of state registration of computer software RU2017613311, March 15, 2017, Federal Service for Intellectual Property.
11. Zeeman E.C. The topology of the brain and visual perception. // The topology of 3-Manifolds, M.K. Fort (ed.), 1962, pp. 240-256.
12. Aleksandrova A.A., Akhterov A.V. et al. Fundamentals of theoretical robotics. Theory of tolerant spaces (review).(in Russian) / Preprint of the KIAM RAS No.45, 2009, 25 pg.
13. Nechepurenko M.I., Popkov V.K., Mainagashev S.M. et al. Algorithms and programs for solving problems on graphs and networks (in Russian). – Novosibirsk, Наука (Siberian Branch), 1990.
14. Mehlhorn K., Naher St. The LEDA Platform of Combinatorial and Geometric Computing. – Cambridge University Press, 1999.
15. Gansner E. R., North S.C. An open graph visualization system and its applications to software engineering // Software-Practice and Experience – 1999.
16. Newick Tree Format graph description language. URL: http://evolution.genetics.washington.edu/phylip/newicktree.html
17. Protégé URL community official website: https://protege.stanford.edu/
18. QR Code company site https://www.qrcode.com/en/index.html
19. GOST R ISO / IEC 18004-2015 Information technologies. Automatic identification and data collection technologies. QR Code Bar Code Symbols Specification
20. DeChant Consulting Services DCS Inc. iWitness Target Accessories. URL: https://iwitnessphoto.com/
21. Lobanov A.N. Photogrammetry (in Russian). M., Nedra, 1984, 543 pg.
22. Lepetit V., Moreno-Noguer F., Fua P. EPnP: An Accurate O(n) Solution to the PnP Problem. / International Journal of Computer Vision, vol. 81, pp. 155-166.

P.S. Baranov[1], A.S. Kurnikov[1], A.A. Mantsvetov[1], V.V. Pyatkov[2]

MULTI-PULSE ACTIVE CCD TELEVISION SYSTEM MODEL FOR 3D IMAGING

[1] ETU "LETI", Saint Petersburg, Russia, psbaranov@etu.ru,kurnikov93@inbox.ru, spmtv@yandex.ru
[2] Television Scientific Research Institute, Saint Petersburg, Russia, pyatkov@niitv.ru

Abstract

The existing 3D imaging methods using active systems are analyzed and the closest one using the pulse method is selected. On its basis, a new method is proposed, based on the use a CCD in the mode of ultra-small integration time. To test the performance of the proposed method, a model has been developed.

Keywords: LIDAR, 3D imaging, multi-pulse active television system.

Introduction

Currently, the direction of active television systems to build 3D images of the object is being developed. They find their application in a variety of industries, such as design, construction, autonomous transport, and also space television. There are a large number of methods for constructing a full or partial 3D image of the object. These include systems using structured light, a mechanical scanning system, systems with a phase or pulse range measurement method.

Each of the methods has its advantages and disadvantages, which determine possible areas of application. When implementing a specific method, it is necessary to take into account which system parameters are most significant and which have lower priority. The following characteristics can be distinguished that describe a system for constructing a 3D image:

- spatial resolution;
- depth resolution;
- weight and size;
- power consumption;
- resistance to vibration and shock loads;
- 3D imaging frame rate;
- maximum range.

Pulse range determination system based on FLASH LIDAR

To build a moving object 3D image, for example, in autonomous transport, the FLASH LIDAR system associated with the pulse method is most suitable. One of the implementations of such a system is given in [1]. From it, the following parameters should be highlighted:

- CMOS sensor built on vertical avalanche photodiodes;
- frame rate - 60 Hz;
- spatial resolution - 688x384 pix;
- IR laser with a peak power of 1.2 kW,
- laser pulse width less than 50 ns;
- laser pulse repetition rate 6 kHz;
- depth resolution 10 cm;
- working distances range to the object from 1 to 250 m.

In this system, the laser emits pulse in a certain area of space in which the object is located, the reflected signal from the object is first fixed by the photodiode and the reflected signal is fixed on the photodetector of the television camera, it carries information about the distance to individual points of the object. The resulting frames are processed, then a 3D image of the object is restored based on them.

The FLASH LIDAR system block diagram is shown in fig. 1.

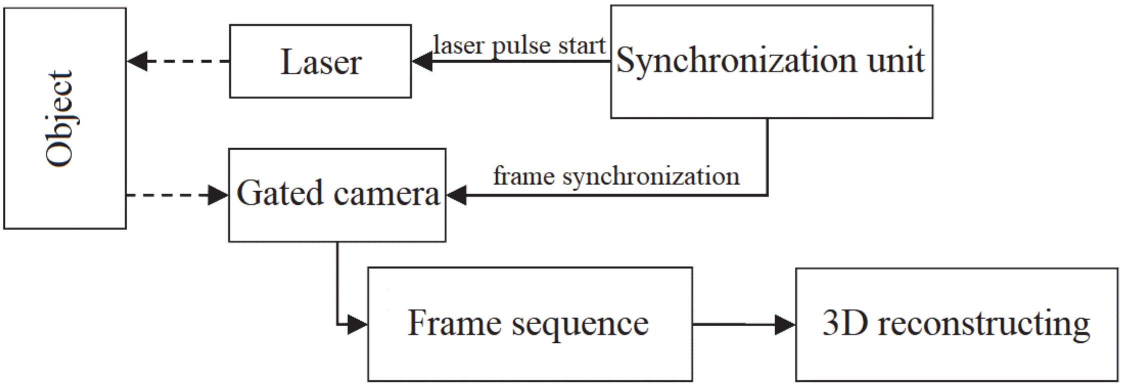


Figure 1 – FLASH LIDAR system block diagram

Multi-pulse active television CCD system with the image intensifier

This solution has such advantages as: the absence of mechanical moving parts, small dimensions and high scanning speed in comparison with systems with mechanical scanning. The disadvantages include the low depth resolution associated with the duration of the emitted pulse, and the high power of the laser radiation. Nevertheless, this solution is the most promising for building 3D imaging system without scanning. In several NASA missions, systems with a maximum range of about 3000 meters were used [2–4]. They used specially designed photodetectors of resolution 128 × 128 pixels.

In [5, 6], an original method was proposed for forming a 3D image using an active-pulse system. The difference of the proposed method from the classical active-pulse system is the use of multi-pulse illumination of the object. Due to this, it is possible to significantly increase the depth resolution to 10-50 mm [5], as well as to reduce the peak radiation power of the laser. In fig. 2 shows the principle of forming a 3D profile of an object by this method.

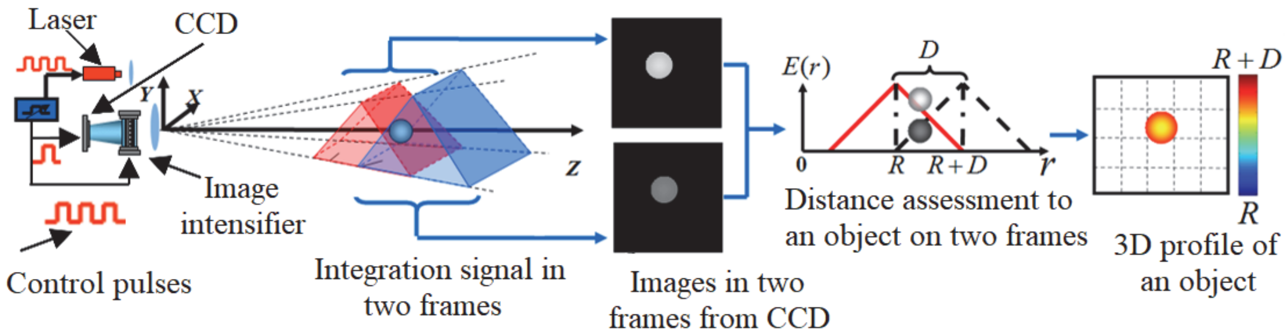


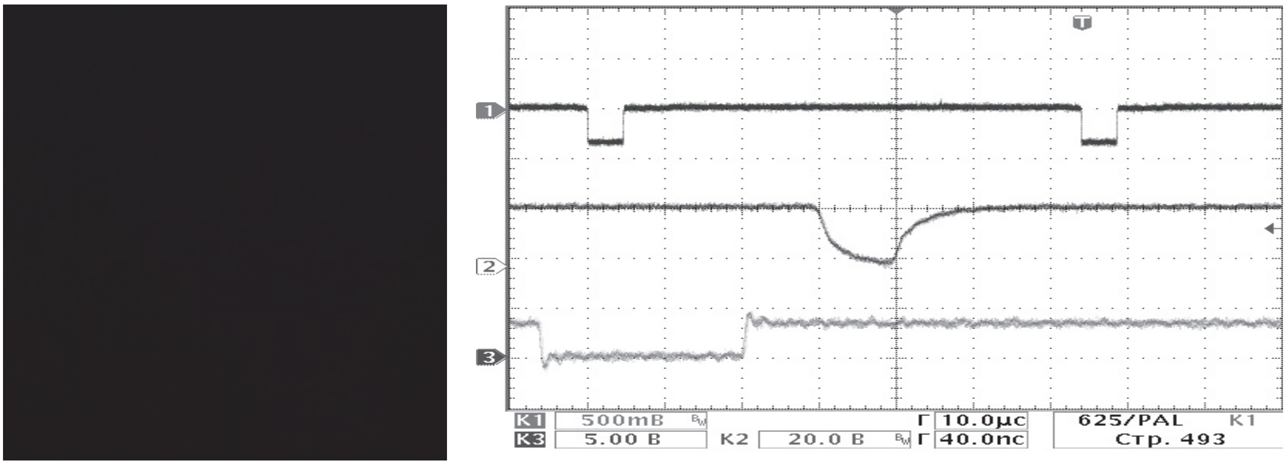
Figure 2 – Principle of forming a 3D profile of an object by the method [5]

In this method, for the realization of short integration time (10–200 ns), an image intensifier is used. One of the significant disadvantage of the image intensifier is the small resource of work and low resistance to light overloads.

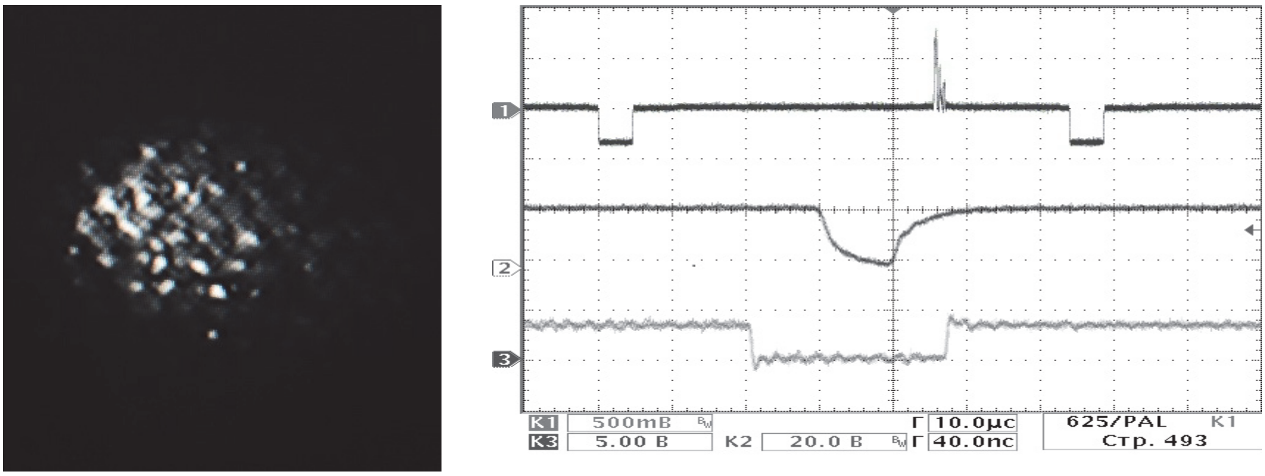
Multi-pulse active television CCD system in the mode of ultra-small integration time

To get rid of the image intensifier, used in the previous method, line transfer CCD system that implements an ultra-small integration time is proposed. In the experimental studies described in [7], a minimum integration time of about 40 ns was obtained, which is comparable to the time of gating of an image intensifier. The results of [7] are shown in fig. 3.

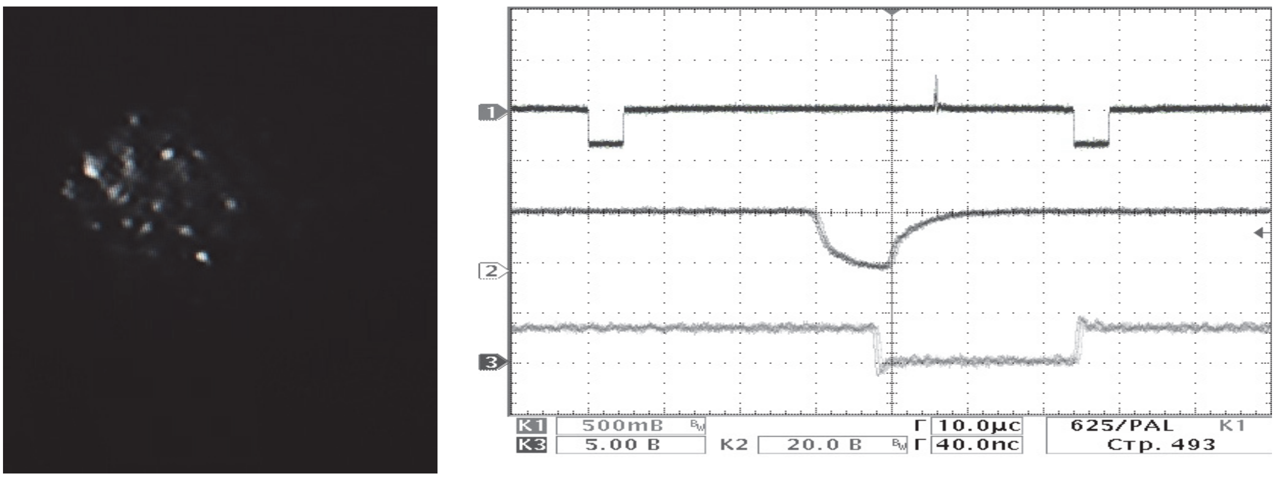
In [8], the idea [5, 6] of multi-pulse backlighting is used to increase the energy of the optical signal, and also CCD is used in the mode of ultra-small integration time (up to 200 ns). In this paper, the task was not to construct a 3D profile, but only to construct a gated television system for observing objects in nuddy environments — fog, haze, rain, and so on. In [8], the experimental studies results are presented, which indirectly confirms the correctness of the proposed method for constructing a system, based on CCD.



a



b



c

Figure 3 – Images of the reflector taken when the reflected light pulse does not overlap with the integration interval – a, with full overlap – b with partial overlap – c; video signal waveform – upper, integration interval – medium, laser pulse – lower [7]

Frame processing and the further depth map construction in this system is assumed to be similar to the method described in [5].

In fig. 4 shows the CCD control signals timing diagrams and the laser for the implementation of the proposed method. The moment the laser pulse starts is the start of time reference. The pulse repetition period of the laser T_L is constant and is determined by the particular laser technical parameters and the data generation required frequency.

The laser pulse duration τ_L and the integration pulse τ_{int} doesn't need to be equal. The cases when $\tau_L \ll \tau_{int}$ are considered in [4, 5]. In this implementation, it is proposed to use the option $\tau_{int}=\tau_L$, which allows to obtain a triangular shape of the resulting integration pulse.

To register the reflected optical signal from each laser pulse, a separate integration pulse is formed. Time shift between pulses is determined as follows:

$$t_{L\_INT} = t_d + (i-1)\Delta\tau + (j-1)\Delta\tau_F, \quad (1)$$

where t_d – main time delay between the laser pulse and the integration time pulse, determined by the distance to the object, $\Delta\tau$ – step shift delay between the laser pulse and the integration time pulse in each pair of pulses, $\Delta\tau_F$ – shift of the time delay between the laser pulse and the integration time pulse in the second frame, i – laser pulse number, j – frame number.

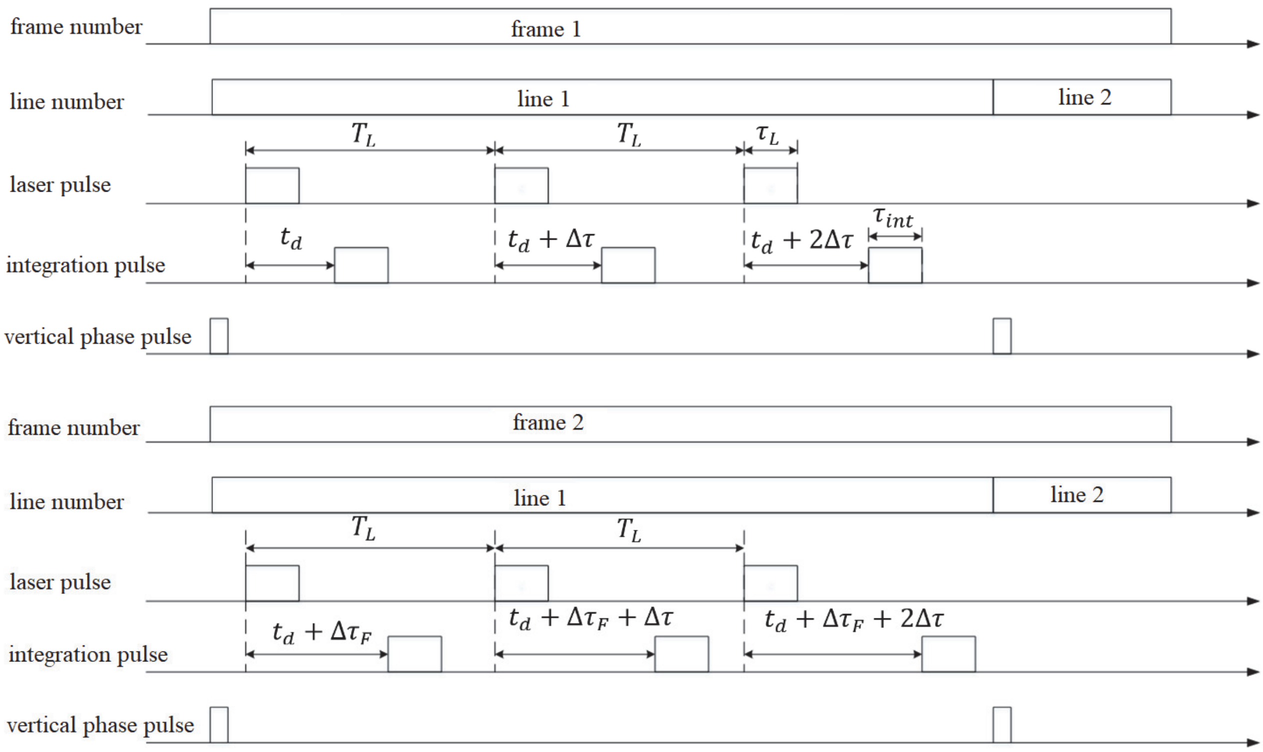


Figure 4 – CCD control signals timing diagrams and a laser for the 3D profile formation according to the method [5]

The standard adjusting integration time method in CCD is the use of electronic shutter pulses during the reverse time on the line or frame. The minimum integration time for the vast majority of CCD television system is more than 5-8 µs. The complexity of the shorter integration time implementation is associated with different capacitive loading of vertical phases electrodes and electronic shutter. This leads to different rates of rise/fall of control pulses fronts. For example, vertical phase electrode capacitances V_1 and the electronic gate electrode capacitances V_{sub} for KAI-01050 are 6 nF and 0.8 nF, respectively. In order to realize ultra-small integration time, it is necessary that integration time be determined only by electronic shutter pulse, since in this case it is possible to provide a steeper control pulse front.

At the time when the optical signal reflected from the object is registered, it is necessary to form pulse sequence on vertical phase electrode V_1 as shown in fig. 5. Pulses duration is usually about 1 µs. Pulse duration τ_{sub} must be greater than τ_{V1}. Moreover, electronic shutter pulse consists of three parts. The first pulse part (marked in light gray) is necessary to reset the charge integrated during the background frame from the photodiode, which is not informative. The second electronic shutter section (marked in dark gray) overlaps in time with charge transfer moment to the vertical register. CCD pixel architecture is organized in such a way that the electrons discharge into the substrate is first carried out, and in the second place the accumulated charge transfer into the vertical register. Second section duration is chosen so that the pulse V_1 transition process is over.

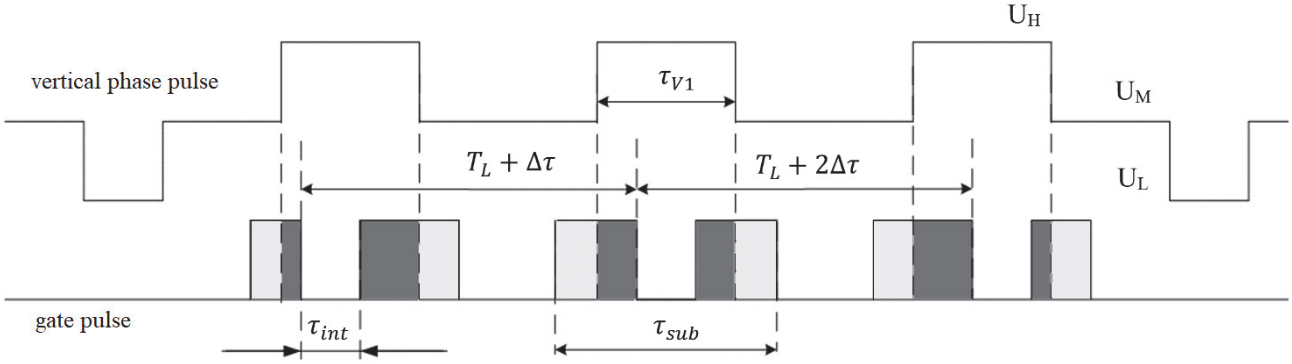


Figure 5 – CCD control signals timing diagram for implementation of the proposed method

During the third section of the V_{sub}, integrated charge reset is forbidden. At the same time, charge transfer to the vertical register has already been resolved. In this case, all electrons formed in the photodiode are transferred to the memory section in a vertical register.

To receive the next laser pulse, only the third portion of V_{sub} is temporarily shifted, while all other temporal ratios are preserved.

The described method of CCD control allowed realizing a short integration time (up to 40 ns) [7].

System model

A new method of forming a 3D image is proposed to use, using the principle of multi-pulse backlighting [5, 6], but taking into account the use line transfer CCD in the mode of ultra-small integration time [7]. For this purpose, a model was developed consisting of television system based on a Sony ICX429ALL CCD sensor with a resolution of 752x582 pixels and a frame frequency of 50 Hz in interlace decomposition, as well as a SPL LL90_3 laser diode with a minimum pulse duration of 40 ns and a pulse repetition rate of 1 kHz. The appearance of the model is shown in Fig.6.

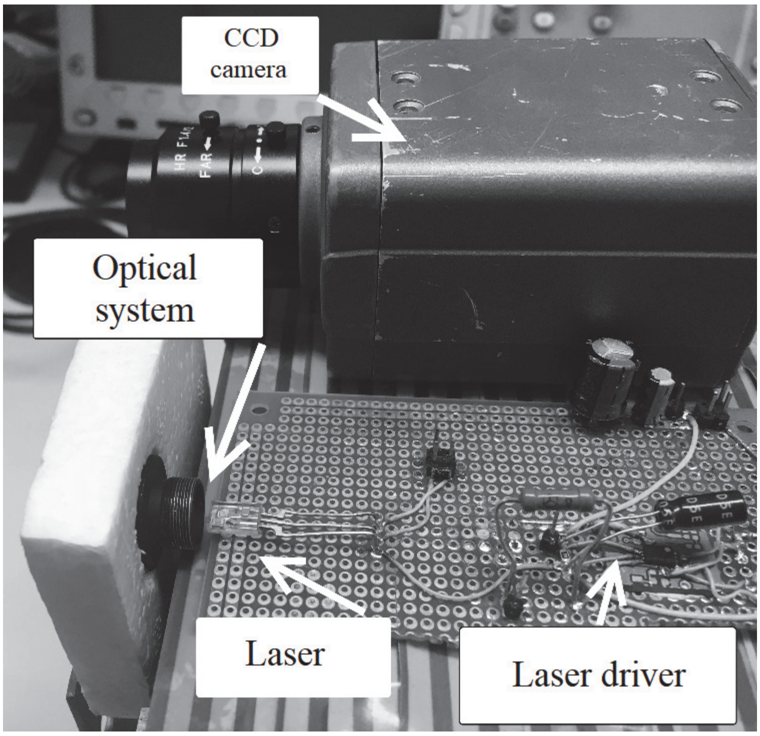
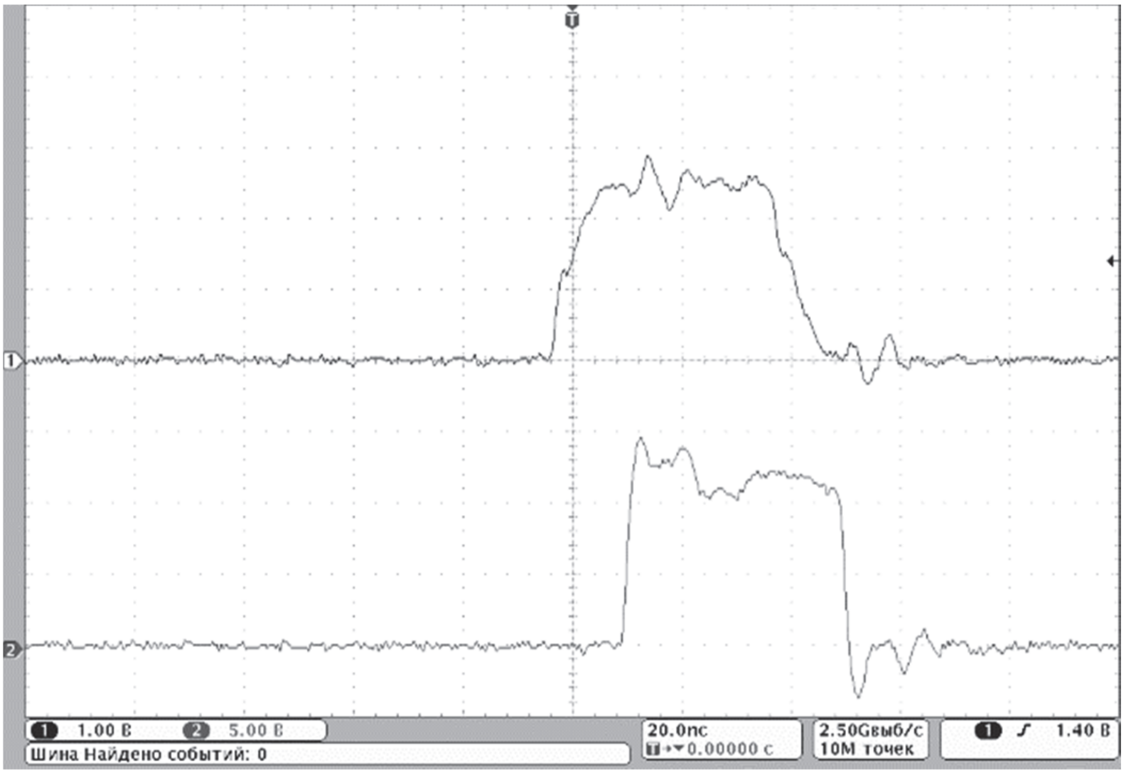


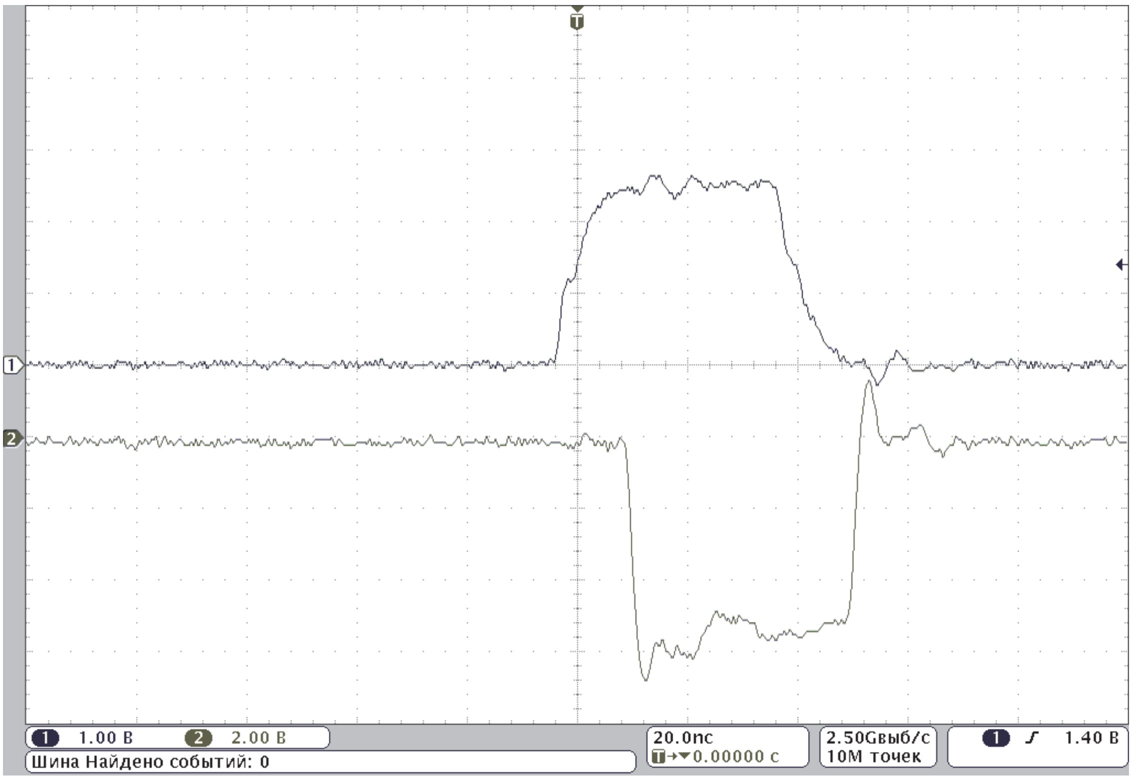
Figure 6 – Television system model

Laser control pulses and integration control pulses waveforms are shown in Fig. 7. In the figure, control pulses fronts are tightened and the shape is slightly different from rectangular, this is due to signal flow through the wires connecting camera and the laser driver (Fig. 7 (a)), which leads to a signals shape distortion that are so short. However, this disadvantage doesn't affect the model functionality; in the future, this disadvantage will be eliminated by combining the system into a single structure.

The laser is controlled and synchronized using the FPGA installed in the camera, its performance is sufficient to ensure synchronous operation of camera and laser. Frames processing received using this system can be done in FPGA, and also in PC using a video capture card.



a



b

Figure 7 – Laser control pulse from camera waveform – upper (a),
laser pulse waveform – lower (a), integration pulse from camera waveform – upper (b),
integration pulse at CCD input – lower (b)

Conclusion

Existing methods of active television systems for building a three-dimensional image of the object are considered, a new method based on the use CCD in the mode of ultra-small integration time is proposed. An active system model is developed based on the proposed method, a Sony ICX429ALL IPSD is used as a photodetector, for which an integration time of 40 ns is realized, and the laser pulses is controlled with an accuracy of 10 ns. In the future, the model will be used to obtain practical results of the existing and new method. After the results are obtained, the two methods will be compared.

References

1. Yutaka Hirose, Shinzo Koyama, Motonori IshiiA / 250 m Direct Time-of-Flight Ranging System Based on a Synthesis of Sub-Ranging Images and a Vertical Avalanche Photo-Diodes (VAPD) CMOS Image Sensor // MDPI Sensors, 2018
2. Farzin Amzajerdian, Michael Vanek / Utilization of 3D imaging flash lidar technology for autonomous safe landing on planetary bodies // Proceedings Volume 7608, Quantum Sensing and Nanophotonic Devices VII; 760828. 2010.
3. Vincent E. Roback1, Farzin Amzajerdian / 3-D Flash Lidar Performance in Flight Testing on the Morpheus Autonomous, Rocket-Propelled Lander to a Lunar-Like Hazard Field // Proceedings Volume 9832, Laser Radar Technology and Applications XXI; 983209. 2016.
4. Farzin Amzajerdian, Vincent E. Roback, Alexander Bulyshev, Paul F. Brewster, and Glenn D. Hines. "Imaging Flash Lidar for Autonomous Safe Landing and Spacecraft Proximity Operation", AIAA SPACE 2016, AIAA SPACE Forum, (AIAA 2016-5591).
5. Xinweil W., Youfu L., Yan Z. Multi-pulse time delay integration method for flexible 3D super-resolution range-gated imaging // Opt. Express. 2015. N 23(6). P. 7820—7831.
6. Xinwei Wang; Yinan Cao; Wei Cui; Xiaoquan Liu; Songtao Fan; Yan Zhou; Youfu Li. Three-dimensional range-gated flash LIDAR for land surface remote sensing // Proceedings Volume 9260, Land Surface Remote Sensing II; 92604L. 2014.
7. Umbitaliev A. A., Tsytsulin A. K., Mantsvetov A. A., Kozlov V.V., Rychazhnikov A. E., Baranov P. S., Ivanova A. V. Solid-state photodetectors integration mode control // Optical journal. 2012
8. Golitsyn A. A., Seyfi N. A. Active-pulse observation method using CCD photodetector with interline transfer // Journal of Instrument Engineering. 2017. Vol.60, N 11. P. 1040-1047.

N.V. Bykov, N.S. Vlasova, M.Yu. Gubanov

A WALL-CLIMBING ROBOT WITH A MAGNETIC-TAPE ADHESION MECHANISM

Bauman Moscow State Technical University, Moscow, Russia
bykov@bmstu.ru

Abstract

This study develops a new combined hybrid adhesion method with the operating surface for the wall-climbing robot, based on a combination of passive magnetic and glue adhesion. To implement this method, a metal tape with an adhesive layer on one side is used. Experimental studies of the magnetic adhesion were performed to measure the holding force of the robot on a vertical surface. The proposed hybrid method allows the robot to move along vertical surfaces of various types.

Keywords: wall-climbing robots, mobile robots, locomotion mechanisms, adhesion methods, magnetic adhesion, glue adhesion, hybrid adhesion methods.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (Grant No. 16-29-09596 ofi-m).

Introduction

There are lot of wall-climbing robots (WCR) prototypes, operating on various combinations of locomotion mechanisms and adhesion methods. Among the WCRs locomotion mechanisms, there are wheeled, tracked, walking, sliding frame, hybrid mechanisms that combine several of these types, as well as specific mechanisms that cannot be attributed to any of the listed groups [1–9]. Adhesion methods are based on friction forces, magnetic forces, air pressure, electrostatic adhesion, molecular forces, rheological properties of a fluids, and some combinations of these methods [1-5, 9-15]. The adhesion mechanisms that technically implement these methods can be active or passive, depending on whether energy is expended to create a holding force. However, for most WCRs, adhesion methods and locomotion mechanisms limit the possible types of operating surfaces. This paper proposes an alternative hybrid adhesion method for WCR to provide a more universal adhesion to various materials.

1. The hybrid WCR principle of operation

To ensure the greatest versatility of adhesion to surfaces of various types, it is convenient to use glue adhesion methods, the main disadvantage of which is contamination of the adhesive contact surface of the robot. A solution to this problem is a replaceable adhesive surface, and various types of adhesive tapes satisfy this requirement. However, the effective locomotion of the robot is difficult to realize, when using glue adhesion of the robot itself to the tape. In this paper, we consider a passive magnetic way of the WCR adhesion to the surface of the tape.

The magnetic method implies the presence of a metal surface, the role of which in this case is the tape itself. The use of a tape from a thin metal sheet with an adhesive layer put on one of the sides makes it possible to ensure a simple mechanism for the adhesion of the WCR to the tape by placing passive magnets on the tracks. Thus, the basis of the magnetic-tape locomotion principle is based on two components: the interaction of the tape with the operating surface by means of the glue layer and the interaction of the tape with the robot platform by means of magnetic force. A general view of a WCR prototype is shown in Fig. 1.

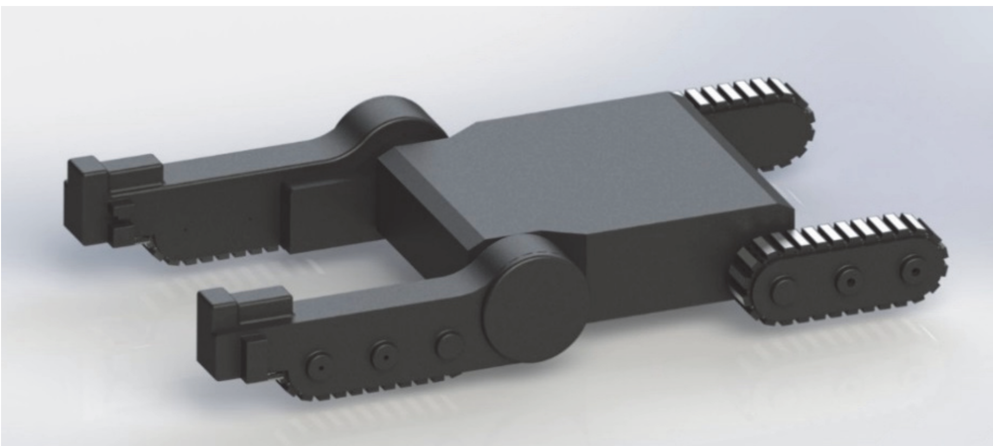


Figure 1 – 3D model the WCR

To ensure autonomous operation in the process of climbing in a surface, the robot sticks a magnetic tape on the surface in its front part. At the same time, the WCR is attached to the tape with utilizing permanent magnets located in the caterpillar tracks. The feed of the tape is carried out by a mechanism that spins the tape cassette, placing the initial section of the tape on the caterpillar. Interacting with the track magnets, the tape moves along a guide. The element at the end of the guide is separating the protective film of the tape's glue layer. Then tape contacts with the surface of the wall by fixing itself on it with a glue layer. The rear track modules of the WCR prototype can make a full turn around the axis of attachment in the robot platform. This allows one to create additional force for gluing magnetic tape at the initial stage of robot's movement.

2. Experimental studies of the adhesion method

There were three stages of experimental studies to determine the characteristics of the new adhesion method. At the first stage, the properties of passive magnets and magnetic tapes were investigated. At the second stage, the adhesion force of the platform to the surface was measured. At the third stage, studies of the WCR locomotion on a vertical surface were carried out.

2.1 Adhesion properties of magnets and magnetic tapes

Neodymium magnets in the form of a parallelepiped with dimensions of 20 × 5 × 5 mm and a mass of 3.7 g each were used as magnetic elements. To compare the adhesion forces between the magnetic elements and the tape, studies were conducted on two types of tapes, the characteristics of which are listed in Table 1.

Table 1. Characteristics of magnetic tapes

Tape characteristics	**Tape 1**	**Tape 2**
Magnetic interaction layer material	Magnetic vinyl	Soft iron
Material density, g/cm^3	4.375	7.48
Tape thickness, mm	1.5	0.4
Tape width, mm	25.4	20

In these experiments, the magnetic tape was glued to the surface, and several magnets were fixed on the frame which imitated robot's platform. The magnets on the frames and the frames themselves on a magnetic tape were held by magnetic forces. A flexible element was attached to one of the magnets, to which the load was suspended. The maximum load mass at which the contact between the magnet and the tape was maintained determined the holding force.

Schemes of experimental studies are shown in Fig. 2, and the results of the experiments in Table 2 and Fig. 3. In the experiments, the type of tape (according to Table 1), the type of frame (rigid (R) and flexible (F)), the number of magnetic elements in the frame and the point of force application varied.

The holding force directly depends on the frame type on which the magnets are fixed. It can be explained by the interaction of the magnets with each other. When using a rigid frame, this interaction is minimal, since the magnets cannot change their position. When using a flexible frame, the magnets, interacting with each other, deform the frame. As a result, the contact area of the magnets with a tape decreases as does the holding force. It can also be concluded that the holding force is essentially dependent on where the load is applied.

Thus, tape 2 is clearly superior to tape 1 by the holding force. In addition, tape 2 has a thickness significantly less than tape 2 (0.4 mm vs. 1.5 mm). Based on this, tape 2 was chosen for further testing.

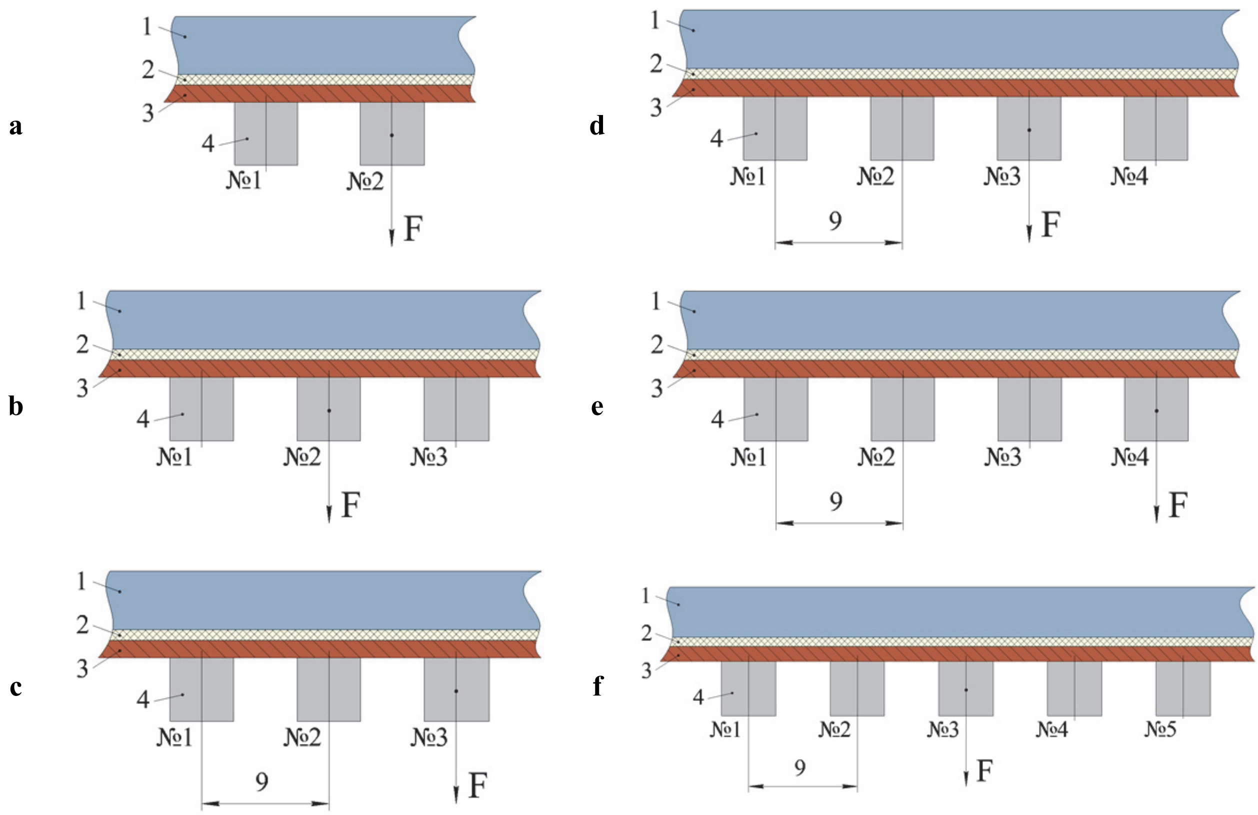


Figure 2 – Schemes of experimental studies:
1 - operating surface, 2 - glue layer, 3 - magnetic tape, 4 - magnets (№1 ... №5), F - force,
a ... f - the designation of the experiment (see Table 2)

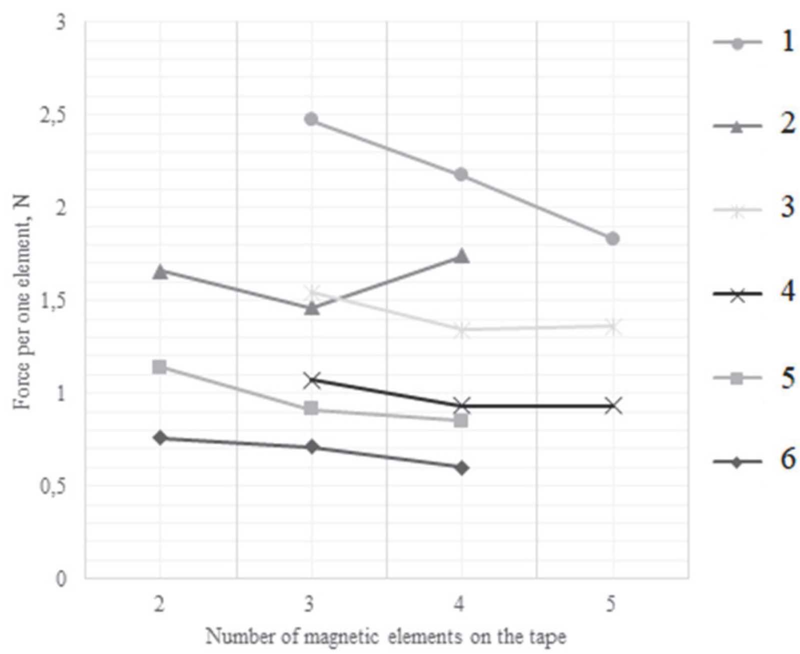


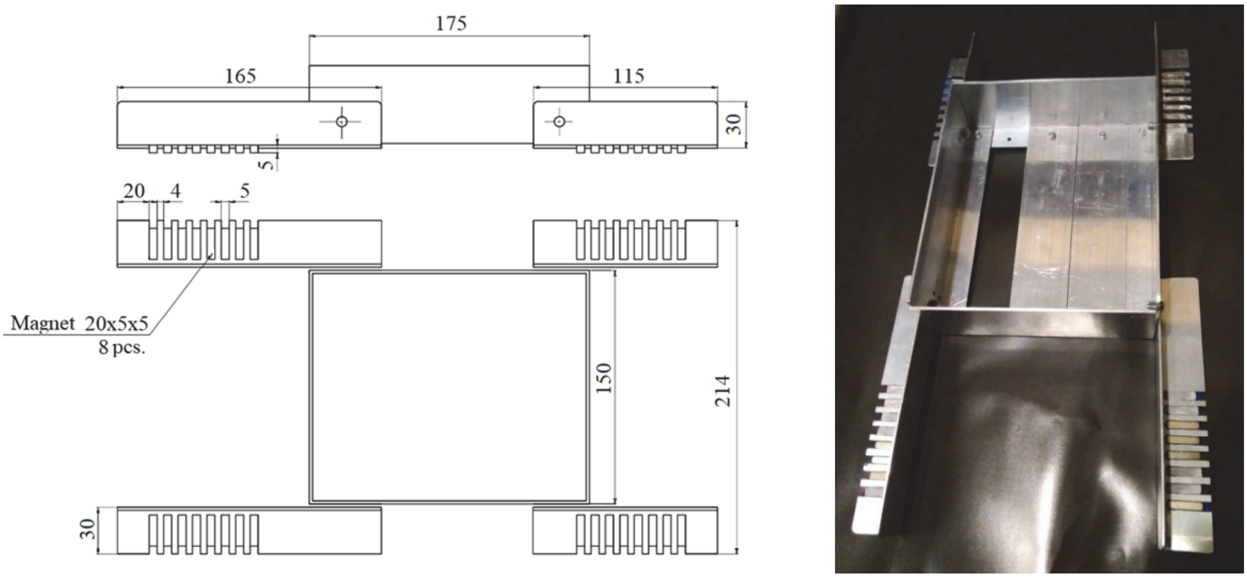
Figure 3 – Dependence of adhesion force per one magnetic element, on the number of magnets in the frame and the point of force application for two types of tapes

Table 2. Results of experimental study of adhesion forces between magnetic elements and tapes

№	The number of magnetic elements	Tape type	Frame type	Point of force application	Average holding force, N	Force per element, N
a	2	1	F	2	1.51	0.76
		2	F	2	2.27	1.14
		2	R	2	3.31	1.66
b	3	1	F	2	3.21	1.07
		2	F	2	4.63	1.54
		2	R	2	7.42	2.47
c		1	F	3	2.12	0.71
		2	F	3	2.73	0.91
		2	R	3	4.38	1.46
d	4	1	F	3	3.70	0.93
		2	F	3	5.35	1.34
		2	R	3	8.69	2.17
e		1	F	4	2.41	0.60
		2	F	4	3.39	0.85
		2	R	4	6.94	1.74
f	5	1	F	3	4.65	0.93
		2	F	3	6.79	1.36
		2	R	3	9.14	1.83

2.2 The adhesion of the platform to the operating surface

Fig. 4 shows the layout of the locomotion mechanism. The body of the layout has dimensions 175×150×50 mm. Elements that imitate the caterpillar tracks are 165×30×30 mm for the front track and 115×30×30 mm for the rear one. These elements are fixed on the layout and can change their angle relative to it. On each of the parts of the layout, imitating the caterpillars of the robot, 8 neodymium magnets, which were used in the previous experiment, are fixed at 9 mm between the centers of the adjacent magnets.



a *b*

Figure 4 – Scheme (a) and view (b) of the layout for the study of adhesive characteristics

In the first experiment, the situation of fixing the robot on the ceiling was simulated and the actual value of the magnetic holding force for a robot in a horizontal position was determined (Fig. 5, a). The experiment involved all 32 magnets located on the layout. The load was placed inside of layout. Next, the layout was placed on a horizontal surface on which two strips of magnetic tape were placed with a width of 20 mm. The maximum holding force was determined by the maximum mass of the load.

Then the measurement of the holding force was carried out when the layout was placed on a vertical wall. Cases with rubber lining and without, as well as the case without sticking imitators of rear tracks, were considered. Rubber linings were used to increase the holding force by increasing the friction between the contacting surfaces. The measurement results are shown in Table 3.

Table 3. The results of measurements of the adhesion of the prototype corps

Type of experiment	**Number of magnetic elements**	**Average holding force, N**
On the ceiling (Fig. 5, a)	32	35.81
On the wall (Fig. 5, b)		
Without rubber linings (Fig. 5, b)	32	24.82
Rubber linings on the front "tracks" (Fig. 5, b)	32	46.35
Rubber linings on the front and rear "tracks" (Fig. 5, b)	32	58.95
Only front "tracks" with rubber linings (Fig. 5, c)	16	29.95

The experimental results are shown that without the rubber linings on the magnets, a robot that has dimensional parameters like the layout can have a mass of not more than 2.5 kg to be held on a vertical surface. And when using rubber linings on magnets in case of contact with a vertical surface with only two front “tracks”, the robot can have a total mass of about 3 kg.

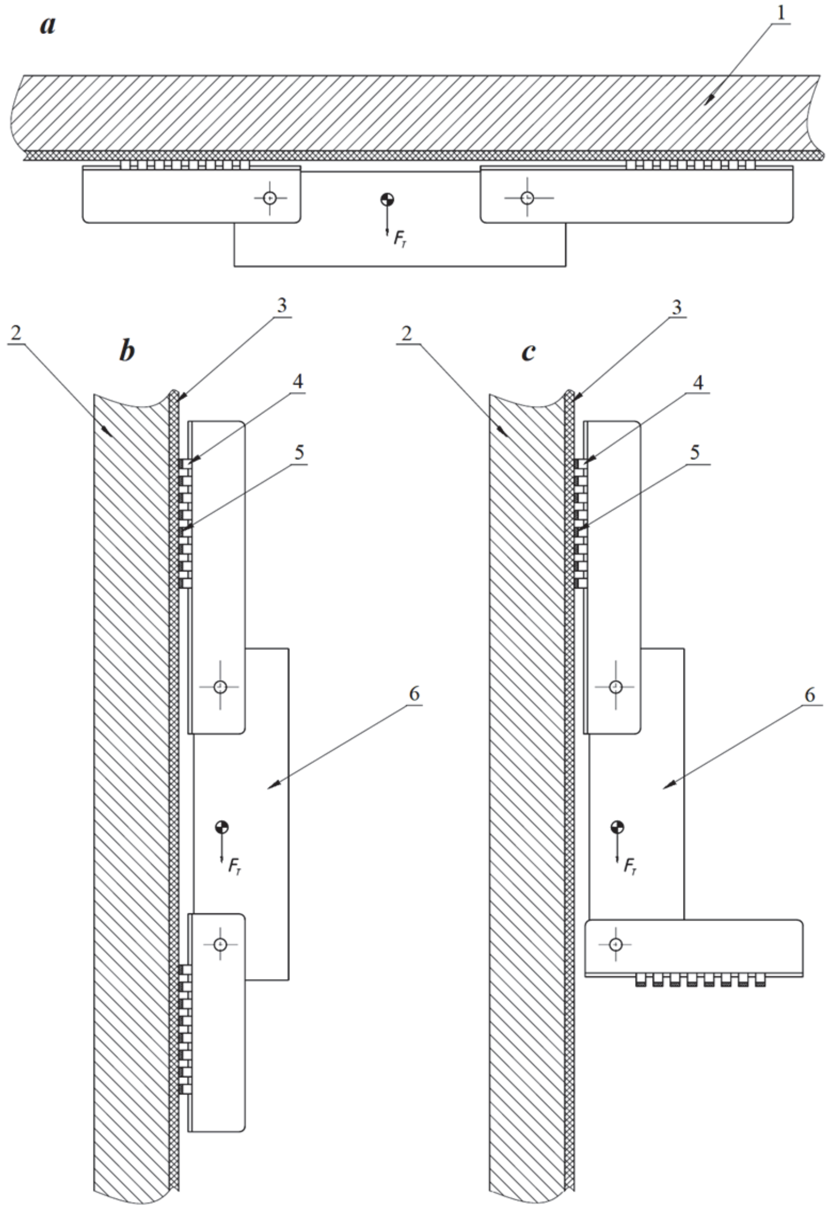


Figure 5 – Schemes of experiments on the study of the adhesion of the layout

2.3 WCR locomotion

The third series of experimental studies was carried out to test the WCR possibility of locomotion on vertical surface. In this case, a magnetic tape was pre-glued to a vertical surface, and the WCR prototype was placed on it before the start of the test. Movement of the WCR prototype was performed by tracked drives with four Maxon DC motors EC-max 22 283837 with a rated power of 12 W and planetary gearboxes GP 22 C 143997. Fig. 6 shows the video frames of the experimental studies. Tests have shown that the average speed of the robot on a vertical surface, at which the contact of the WCR prototype with a magnetic tape was not lost, is approximately 2 cm/s.

Conclusion

A new adhesion method of wall-climbing robot with a vertical surface, based on a combination of passive magnetic adhesion with magnetic tape and tape with an operating surface due to the glue layer, is proposed. It is noteworthy that without a tape, the robot can be used on metal surfaces. The concept of constructing a locomotion mechanism and the design of a prototype platform for vertical movement have been developed. Experimental studies of the magnetic properties of the proposed locomotion mechanism has showed its efficiency.

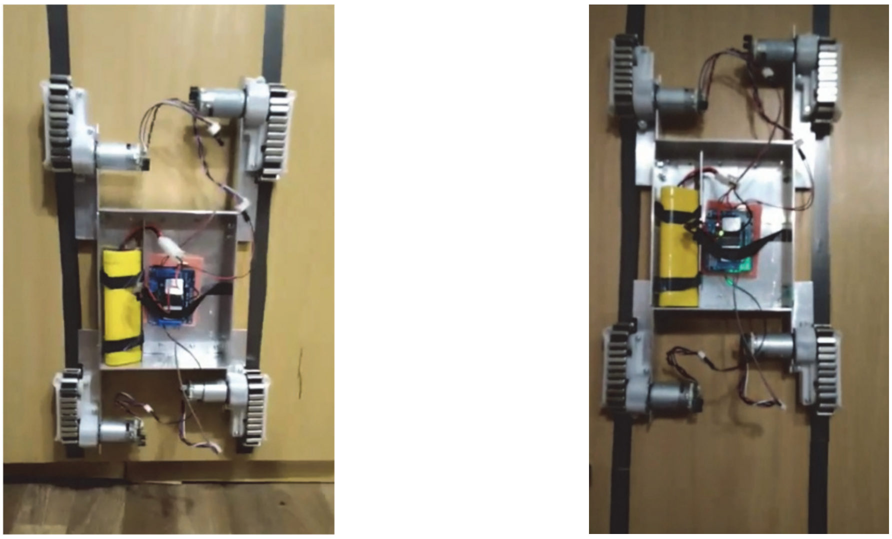


Figure 6 – Video frames of the locomotion mechanism testing

References

1. A survey of climbing robots: locomotion and adhesion / B. Chu, K. Jung, C.-S. Han, D. Hong // International Journal of Precision Engineering and Manufacturing. 2010. Vol. 11 (4). P. 633-647.
2. Bisht R.S., Alexander S.J. Mobile robots for periodic maintenance and inspection of civil infrastructure: a review // Proceedings of the 1st International and 16th National Conference on Machines and Mechanisms (iNaCoMM 2013). Roorkee, India, 2013. P. 1050-1057.
3. Silva M.F., Machado J.A.T., Tar J.K. A survey of technologies for climbing robots adhesion to surfaces // Proceedings of the 6th International Conference on Computational Cybernetics (ICCC 2008). Stara Lesná, Slovakia, 2008. P. 127-132.
4. Schmidt D, Berns K. Climbing robots for maintenance and inspections of vertical structures - A survey of design aspects and technologies // Robotics and Autonomous Systems. 2013. Vol. 61 (12). P. 1288-1305.
5. Brusell A., Andrikopoulos G., Nikolakopoulos G. A survey on pneumatic wall-climbing robots for inspection // Proceedings of the 24th Mediterranean Conference on Control and Automation (MED). Athens, Greece, 2016. P. 220-225.
6. Series of multilinked caterpillar track-type climbing robots / G. Lee, H. Kim, K. Seo, J. Kim, M. Sitti, T.W. Seo // Journal of Field Robotics. 2016. Vol. 33 (6). P. 737-750.
7. Fu Y., Li Z., Wang S. A wheel-leg hybrid wall climbing robot with multi-surface locomotion ability // Proceedings of the IEEE International Conference on Mechatronics and Automation. Takamatsu, Japan, 2008. P. 627-632.
8. Development of a wall-climbing robot with biped-wheel hybrid locomotion mechanism / W. Dong, H. Wang, Z. Li, Y. Jiang, J. Xiao // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013). Tokyo, Japan, 2013. P. 2333-2338.
9. Chan B., Balmforth N.J., Hosoi A.E. Building a better snail: lubrication and adhesive locomotion // Physics of Fluids. 2005. Vol. 17 (11). P. 113101.
10. Nansai S., Mohan R.E. A survey of wall climbing robots: recent advances and challenges // Robotics. 2016. Vol. 5 (3). P. 5030014.
11. Wet adhesion inspired bionic climbing robot / B. He, Z. Wang, M. Li, K. Wang, R. Shen, S. Hu // IEEE/ASME Transactions on Mechatronics. 2014. Vol. 19 (1). P. 312-320.
12. Osswald M., Iida F. Design and control of a climbing robot based on hot melt adhesion // Robotics and Autonomous Systems. 2013. Vol. 61 (6). P. 616-625.
13. Wang L., Graber L., Iida F. Large-payload climbing in complex vertical environments using thermoplastic adhesive bonds // IEEE Transactions on Robotics. 2013. Vol. 29 (4). P. 863-874.
14. Wiltsie N., Lanzetta M., Iagnemma K. A controllably adhesive climbing robot using magnetorheological fluid // Proceedings of the IEEE International Conference on Technologies for Practical Robot Applications (TePRA). Woburn, USA, 2012, P. 91-96.
15. Koh K.H., Sreekumar M., Ponnambalam S.G. Hybrid electrostatic and elastomer adhesion mechanism for wall climbing robot // Mechatronics. 2016. Vol. 35. P. 122-135.

V.K. Abrosimov, V.V. Eliseev

CURRENT STATE AND DEVELOPMENT POTENTIAL FOR THE NATIONAL AGRICULTURAL ROBOTICS

Scientific and Technical Center "RoboPROB", Ltd., Moscow, Russia
avk787@yandex.ru

Abstract

The issues of making robots in the Russian Federation of various types and designation for the agriculture needs are reviewed. The primary focus is on agricultural robots intended for crop growing. The requirements for different types and areas of agricultural robots from different points of view expressed by agricultural producers are formulated. Along with traditional irrigation and differentiated fertilization, the priority tasks cover the following new tasks of robotic farming: soil sampling, weed destruction, plant diseases recognition, crops pests control, etc. The reasons for the gaps with the world trends in robotics development and precision farming are systematized. The necessity to improve intelligence of agricultural robot control systems with migration to autonomous modes of operation is justified. The main directions for research and technical innovations use to create domestic agricultural robots are determined.

Keywords: robot, agriculture, intelligence, task, innovation, precision farming

1. Modern challenges for agriculture in the Russian Federation

Agriculture is an essential sector of economy in the Russian Federation. But Russian agricultural producers faced several major challenges in the last 10-15 years.

The main ***political challenge*** is the need to ensure food security of the country. Except for the clear issues of domestic seed and genetic material and crop protection agents development, import substitution of agricultural machinery becomes a major issue. The percentage of Russian machinery used in the fields is extremely low. According to the figures provided in 2016, 52% of the tractors produced in Russia were assembled from the tractor units imported by the foreign brands.

The global unresolved issue is the yield reduction. The yield is an extremely complex, essentially non-linear and sometimes discontinuous function of soil agrophysics, fertilizers applied, weather conditions, and field history. The idea of natural fertility has been put into practice for a long time in the Russian Federation, resulted in significant land depletion. Therefore, the main ***economic challenge*** has been and remains to increase efficiency of agricultural production on depleted lands.

The main ***technical*** challenge is the need to significantly reduce the share of manual labor in agriculture. Partially this issue is resolved in farm animal production, where many production processes can be and are already largely automated. But there is certain contradiction in respect to crop growing: it appears to be more profitable to pay for manual unskilled or even seasonal non-professional labor than to invest in agricultural machinery and equipment.

These challenges to the agricultural sector in the Russian Federation, acting in an integrated manner, resulted in significant dependence of the agriculture on modern technological innovations and solutions. In the victorious reports issued by the Ministry of Agriculture of the Russian Federation "... Russian AIC is a successful sector feeding the country and conquering international markets ...". However, this statement is poorly applicable to the use of modern information and communication technologies in this sector. Indeed, famous paradox of the ancient Greek philosopher Zeno says that the swift-footed Achilles will never catch up with the leisurely turtle. But "Achilles", the bright representatives of which are the USA and Europe with actively developing technologies in precision farming, is not initially behind, but ahead of the Russian "turtle", with its dream and wish to catch up and overtake "Achilles", at least by 2035 (as resulted from the numerous concepts of agriculture digitalization and in general still primitive informatization of agricultural enterprises). The "turtle", of course, applies different technical innovations left by Achilles on its way (GPS navigators, GIS systems, parallel driving technologies, etc.), but "Achilles" "... takes 100 steps forward but the turtle only ten ones."

The above listed challenges require both scientific and technical responsive operational decisions.

2. Current situation in the field of agricultural robots

Worldwide active development of robotics leads to intelligent robots penetration in such a historically conservative area as agriculture. According to Robotrends.ru portal, agriculture robotechnics is targeted at increasing efficiency and reducing the cost of agricultural production, improving the quality of monitoring and predicting plant growth, onset of plant diseases, pests detection, environmental load reduction and agricultural production safety improvement. Making of intelligent agricultural robots can be an effective result for modern achievements use in technology trends such as digitalization, unmanned vehicles, the Internet of things, etc. According to some forecasts, the market of agricultural robots will exceed $ 25 billion by 2024. Agricultural robots production in the world is expected to increase tenfold in the coming years.

The main areas of agricultural robotics include robotic tools for crop growing (sowing, irrigation, unmanned tractors, monitoring of agricultural lands using unmanned aerial vehicles, harvesting robots, robots for plant-protecting agents application, fertilizers, pest control, etc.) and robots to improve efficiency of farm animal production (primarily milking, cleaning), etc.

The highest practical results were achieved in the field of farm animal production robotics on dairy and meat livestock farms. Milking robots, automated feeding systems, feed adjusters and movers, cleaners, automated poultry farms, etc. refer to well-known solutions. As a rule, such robots are electric-driven mobile autonomous trolley having standard programmed simple traffic routes. However, livestock robots also have problems: a robotic carousel cannot deal with "non-standard cows" (in Russia, with our high genetic variation of breeds at least 10% - 20% of animals refer to this category), some cows do not wish to join milking circle due to their psychotype, etc.

In the field of crop growing, agricultural robots are entrusted with the tasks of drawing up digital field maps with the shape and boundaries determination, purposeful movement along specified or selected field routes without harm to the plants and soil, information receipt about field properties (heterogeneity, temperature, humidity, wind speed and direction), soil sampling, weed and pest control, dosed fertilizers application in real time and other tasks [1].

It shall be stressed on incommensurability of the pace of modern robotics development and improvement of farming technologies. Development of ground and airborne robots control systems and mathematical methods for information collection, processing and analysis anticipate classical and highly conservative processes development in agro-industrial production. Developers propose to introduce the following required for effective work in the fields, but so far, unfortunately, high-cost elements of robotics: advanced computer visual facilities, ultra-precise navigation, modern image recognition methods, special software, etc. Understanding of the term "big agricultural data" is being formed. The need to increase autonomy of agrotechnics operation in difficult field conditions and use of intelligent information processing and decision-making methods is also recognized. But the views of even large Russian agricultural producers on the capabilities of modern agricultural robots are rather cautious.

The study of foreign literature shows that required statistics is actively accumulated abroad and specialized software is developed. Foreign robots collect fruits one by one, persistently search for weeds, recognize them, trample, drive into the ground, cut off with mills, spoil weed leaves with laser, introduce herbicides exactly under the plant root, etc. [2, 3, etc.]. It shall be noted, however, that all of above listed is done with varying success and at the level of individual experiments.

The so-called robot-oriented farming is becoming a completely new innovative direction. It means transition in agricultural production to crops that are easier processed by robots, including crop collection and storage. It requires special conditions for fields design, planting and growing, but at the same time it significantly simplifies robots making, since their most complex operations become simple and standard.

In domestic literature, unfortunately, more general descriptions of the relevance of such tasks can be found, scattered in mass media, specialized scientific journals and numerous informational messages from Internet websites. However, statistics required for intelligent methods use to solve agricultural issues is usually unavailable, or such bases are only created. Currently, the share of national funds attributed to agricultural robots (understanding that the term itself is still interpreted in many ways) does not exceed 10% in Russia. The rest of robotic tools, where they are available, are of foreign manufacture. The gap in domestic agricultural robotics with the world level is observed in many parameters: reliability, material intensity, power consumption, efficiency, speed and accuracy. For industrial robots developed to solve agricultural transport issues, the ratio of the payload weight (for example, portable elements) to the robot mass is very low. Especially, the gap is noticed in the element base of control systems and software.

In the context of on-going political and economic anti-Russian sanctions, it is clear that it is not possible to rely on massive purchase of foreign machinery in the future, which is, moreover, accessible only to certain large agricultural holdings due to their prices. Therefore, over the next years (no later than 2025–2030), it is necessary to "reverse" the dangerous tendency of the Russian backwardness in the field of robotic devices making for the needs of agriculture. The most important component of the entire innovation work in the field of robotics referring to agriculture should be import substitution and orientation towards national agricultural robots making.

3. The main areas of concern for the national agricultural robots development

A significant global trend in agriculture during the last decade is a precision (coordinate) farming [4]. The basis of scientific concept for precision farming lies in objective and proven by practice fact of substantial heterogeneities within one field. The latest technologies are used to assess and recognize such heterogeneities: global positioning systems (GPS,GLONASS, Galileo), special sensors, aerial photographs, satellite images, special programs for agricultural management based on geoinformation systems. The most important issues in precision farming include soil cover knowledge, inner-field variability of the plant's habitat and targeted fertilizers application.

However, the use of current technical facilities to solve the issues of precision farming is objectively limited. Modern solutions are aimed at working on large field areas. It is not possible to get rid of a lot of manual operations. Large tractors have a negative impact on the soil and require special operating conditions. Small agricultural aviation, for obvious reasons, is practically unavailable. Soil samples are taken mainly by hand. Field transmitters and sensors have high costs and cannot be placed in a variety of places needed. Diseases monitoring and pests search can be performed only by an agronomist being personally present on the site. But required efficiency in decision-making is extremely high; specific biology of many pests has not yet been studied. So, the rate of some pests development from eggs to caterpillars is no more than a week (for example, meadow moth).

Technological achievements over the recent years in the areas of global positioning, machine vision, laser technology, mechatronics and unmanned vehicles allow to develop and implement a variety of specialized intelligent and robotic systems, and to promote intelligent technologies for agriculture. Analysts refer to them: a) Earth remote sensing satellites; b) unmanned ground transport robots; c) unmanned aerial vehicles (UAVs); d) field transmitters and sensors integrated into complex and interconnected networks. The main areas of concern in this area should be considered based on technical requirements for the agricultural robots, which, in turn, are still being formed to meet agrotechnical requirements for the quality of agricultural works, in particular, in terms of duration and quality of works, material consumption, acceptable losses in products, etc. It shall be also stressed on invariance of the principle in obtaining maximum yield with high quality processes performed by the agricultural robots, increasing soil fertility and observing environmental requirements.

It is clear that ***agricultural robots efficiency*** depends on many factors: power loading, movement speed through the fields, performance of attachable equipment, etc. A priori, it is clear that agricultural robots efficiency will exceed human efficiency working in the fields. The effect will depend both on seed cultures and on the agricultural robot functionality. In some examples (for example, soil sampling), it has already been shown that processes efficiency can be increased several times with a significant improvement in the quality of field works [5]. In some situations, due to agricultural robots capabilities implementation, even the entire business processes will fundamentally change; presumably this will refer to the plant diseases definition, agricultural pests search, etc., information support and automation of which are not currently started yet in the Russian Federation.

An important component of efficiency factor is the movement speed between the given points of their route. When calculating the required speeds, it is advisable to consider the time from complete stop in working operations performance until the movement speed increase by the agricultural robot. As a rule, maximum movement speed up to 15-20 km/h is quite sufficient to solve the key tasks of precision farming. The time intervals between the stop and subsequent movement, based on the practice, shall be set within the range from a few seconds to 10-15 seconds.

In terms of work quality performed, the agricultural robots are guaranteed to be out of competition, but only for those processes that comply with described patterns. The quality of agricultural works performed by agricultural robots depends on the quality indicators entered according to their individual functions. In decision-making processes, agricultural robots perform a significant, but still supporting role.

When making agricultural robots, it will be necessary to solve a number of technical issues associated with the work in the fields of relatively heavy unmanned vehicles, primarily on the issues of maneuverability and the effect of the running gear on the soil. Running gear compacts the soil, which negatively affects its fertility and crop yields. Currently, according to some standards, the impact of agricultural machines and vehicles on the soil is allowed no more than 45 kPa for the tracked vehicles and 110 kPa for the wheeled vehicles. It seems that the very figures will be laid in the relevant requirements. The issues of manoeuvrability are also associated with special conditions of agricultural robots operation. So, for tilled crops, the agricultural robot manoeuvrability between the rows, which can be significantly different, is especially important. As a rule, inter-row spacing is 45, 60, 70, 90 cm, which suggests the possibility to create solutions close to universal for robot lorries.

The main requirement for technical facilities is to ensure precision farming tasks solution, which directly relate to ground-based agricultural robots - high planning-altitude field accuracy with annual repeatability of results (in the plan - up to 1 cm and in height up to 5 cm). As a rule, to set requirements for positioning accuracy no worse than 10–20 cm is both reasonable and sufficient for solving basic tasks of precision farming. Currently, such characteristics are already achievable.

An important and, in some cases, defining characteristic of agricultural robots efficiency is readiness to work in difficult weather conditions. Agricultural tasks are closely related to the natural features and regional factors, they often require high efficiency, resolution in different climatic conditions and during prolonged rains, fog, twilight, etc. It is advisable to lay down requirement for agricultural robots working capacity in difficult weather conditions with details (rain, mud, thaw, twilight, etc.), ideally in the format of 24*7*365. However, this requirement can sometimes significantly increase the cost of agricultural robot, since it will require from the developers to install more complex and expensive equipment, in particular video cameras. Here it shall be found mutually acceptable solutions; practical experience of the customers and relevance of agricultural robot ordering particularly for specific conditions and crops should be decisive.

In the future, it is advisable to seek to the maximum autonomy of agricultural robots; but in the short term, one can hardly expect here essentially innovative solutions. The actions of agricultural robots will be controlled by the Operator trained according to the special programs. The relevant data transmission channel will be arranged for communication between the agricultural robot and the Operator. Now agricultural robot exchanges two types of data-information with the Operator in a significant extent (video data) and, if necessary, commands to perform actions. Obviously, these channels shall be designed as wireless and implemented on the basis of well-known wireless data standards. Gradually, the amount of data transmitted should be reduced to the level of individual major commands.

The issue of “turnkey” precision farming is a promising direction. To do this, organization of agricultural robots control systems linking with other agricultural machines and so-called cyber-physical production systems of agro-industrial complexes, containing information for agricultural operations planning, routes and transport loading, crop yields, flow charts, data on resources, supplies, sales, and other information systems of agricultural enterprises is crucially important. To implement above mentioned is possible in the future on the basis of service-oriented architectures [6], in which robot functionality will be described as a service with the relevant meta-information and interfaces for automatic access and control objects involvement. Preliminary analysis does not reveal any technically unsoluble problems here; moreover, both the volumes of the data transmitted and the frequency of the data exchange to solve the problem of such linking should not become significant.

Separate requirements will be imposed for the information processing. Decision-making processes should and will involve both agricultural robots control systems and various mathematical models of agrocynosis. Therefore, the level of agricultural robot intelligence should be sufficiently high. Most of these problems should be solved directly on board of agricultural robot, which may require significant computation capacity. But data processing in the Internet clouds can be also useful. Online interaction channels should not be wide. Therefore, perhaps the promising LoraWAN technology will be effective here.

Summarizing requirements in terms of agricultural practice, the main technical requirements for the promising domestic agricultural robots shall be formulated.

– To have computer vision system (parallel driving for standard tasks) for movement in difficult geophysical conditions, including over uneven ground.

– To have wireless connection with the Operator; in the absence of communication, operate autonomously until communication is restored.

– To form traffic route according to the task received and adjust it involving Operator (or independently in autonomous operation).

– To implement the tasks in recognizing situations occurring in the habitat of the plants, assigning the situations to a certain class and preparing recommendations for the agronomist on the actions in this situation.

– To have an ability to install a variety of attachable equipment for solving the issues of precision farming (differential fertilization, recognition of plant diseases, weed and pests control, etc.).

– To have fuel reserve for no less than 6-8 hours of operation.

– To have a significant time between failures - hundreds of hours.

– To be easy maintainable.

− To be served with a minimum number of specialists with average qualification.

The analysis of the technical requirements listed essentially determines the main directions for the national agricultural robotics making.

4. Agricultural robot mental power

It is assumed that the agricultural robot will function in the general information and communication environment of the agricultural enterprise. In this case, most functions of the agricultural robot will be associated with the statements and methods for the intellectual control tasks solving, which in the future should be solved in autonomous mode (the Operator is only entrusted with the tasks of agricultural robot functions control).

The term "artificial intelligence", being underdetermined, was used for various agricultural controlled objects in different senses. In some cases, it is associated with the term "intelligent". As a result, it is difficult to understand what the control system for "intelligent attachable equipment", tractor "with the elements of artificial intelligence" and other agricultural machines shall really be. The authors do not share the general euphoria about the multi-valued interpreted word "smart" attached to various agricultural terms ("smart" tractor, "smart" field, "smart" greenhouse, etc.). Obviously, this is deeply marketing and journalistic approach to promote some, often simply automated technologies, into the sphere of agriculture, which obscures the true essence of changes made.

Indeed, in a general sense, the concept of "intelligence" means that the controlled objects are adaptable to the maximum extent to changes in external environment, equipped with existing high-performance communication, navigation and computer vision systems, have significant possibilities for manoeuvring in space and, most importantly, are independent in choosing behaviour strategies. In the medium term, one of the innovative agricultural robot functions will refer to independent Plan for Problem Solution development, described by the Customer in a general way and its implementation mainly offline. Above stated shall form intellectual component of the robot, which makes it possible to consider it as an intelligent agent for this purpose [7]. This means that agricultural robot shall possess the following properties: ability to adequately percept environment and rapidly respond to its change (reactivity), ability to both replenish databases and use them (awareness), ability to perform various tasks within its functionality (mobility), learnability, etc. These functions implementation is quite complicated and time consuming process. In this context, agricultural robots can be considered "intelligent" to the extent that they are able to display the above properties. The concept of the "control system with artificial intelligence" is more rigid, as it assumes other functions associated with modern achievements implementation in the field of artificial intelligence.

Let`s consider possible methods application of the so-called "weak" artificial intelligence as referred to the agricultural robots making.

Knowledge-based agricultural robot systems development

Models and methods for knowledge presentation, extraction and structuring, and knowledge databases creation are developed in this area of artificial intelligence. The above is intended for solving unstructured or weakly structured problems.

Knowledge necessary for effective agricultural robots operation is acquired from various sources: a) own technical facilities of the agricultural robot, which record situation in the environment accessible to its means of observation; b) information received from other agricultural robots and c) information from external sources. In this case, the tasks of farming, farm animal production, etc., which require robotics, despite of their numerous specific features, cannot be considered weakly structured, since the elements and links between them are quite easily observed.

In this context, knowledge required by the agricultural robot to perform its tasks is advisable to accumulate within the framework of such a resource as "field situational awareness" [8]. In principle, this resource can and should be built as a distributed knowledge database, in which the output data of agrocenosis models, current data on the agricultural robot operation and data from information systems of agricultural enterprises will be logically related. As a result, a regularly updated and filled with all necessary information knowledge database with estimates of the likelihood / possibility of various typical and/or random events occurrence in the course of field works is built up. The logical mechanism of successive development of the situations based on a set of events occurrence, situations recognition and algorithm for situation consequences on crop yields prediction, taking into account retrospectives for a given geographic region and agricultural crop, will be determined by the recommended procedure for actions in this situation.

The class of intelligent knowledge-based systems includes expert systems. Signs of the artificial intelligence in the expert systems historically refer to the knowledge database with a set of rules for a certain range of tasks and software and hardware tools that allow formulating recommendations for actions in the current Agrorobotics based on the data, which may become an effective technical element in the overall expertise process with a focus on correctness of decisions made, choice of crops, sequence of various technologies use, choice of time intervals for planning crops, irrigation, fertilizers application taking into account autopilots setting up, automated monitoring systems introduction, cloud services development for the field history, etc. So, agricultural robots will contribute to the expert systems development for the agricultural enterprises through the information contribution to appropriate maps development, information accumulation and updating in the field of situational awareness and robotic processes implementation.

Visual information processing by agricultural robots

The tasks for images receipt, processing, analysis and synthesis are almost the main for agricultural robots. They are solved in computer vision systems when agricultural robot moves to the place of its work, in programs for weeds and plant diseases recognition, pests search, etc., based on the required video sequence, etc. As a result of processing, the original images are converted into another type data, for example, into the commands of agricultural robot control systems, attachable equipment sensors, etc.

Agricultural robots training and self-training

This actively developing area of artificial intelligence includes models, methods and algorithms development implementing automatic knowledge accumulation and generation using knowledge analysis and synthesis procedures. This area refers to the Data-mining, Knowledge Discovery in databases, etc.

Automatic information accumulation can be organized for field situational awareness data. The use of Knowledge Discovery in situational awareness data will contribute to all necessary for robotic farming regularities generation. This possibility is based on the fundamental similarity of the situations on the field; and general signs of situations are usually the same for selected cultures and differ in clear details. As a new idea, it is possible to propose agrocinosis models development on the neural networks and presentation of the situations as neuron ensembles with common signs of situations as a “core” of the ensemble and external conditions affecting the situation as a “fringe” for the ensemble [9].

Pattern recognition as a task for special agricultural robots systems. The tasks of patterns recognition directly include the tasks of plant diseases recognition, pests habitats and places of weeds growth search and recognition. Currently, this scientific direction is the most interesting in terms of intellectual methods application and the most undeveloped. It is required to develop intelligent autonomous weeding systems for weeds recognition and removal using very precisely controlled impacts, sets of multispectral images allowing you to quickly respond to changes in the state of crops, signs of deficiency recognition in the substances required (nitrogen, phosphorus), which shall be replenished in the form of fertilizers and others. Convolutional neural networks are considered to be very efficient mathematical tools.

Separately, it is necessary to distinguish such direction of artificial intelligence as speech recognition. In principle, we seek autonomy of agricultural robots, but it is still too early to eliminate the controlling function of the Operator. If the voice of the Operator can be converted into commands for the agricultural robot, such functions will be certainly in demand.

Well-behaved applications in those areas of artificial intelligence associated with machine translation, development of natural language interfaces and new computer architectures in relation to agricultural robots in the future are not observed.

Essentially, all the algorithms developed shall be programmed initially in the algorithmic language and then implemented in electronic boards of agricultural robot control systems in a programmatic manner. All tools for intelligent systems development can be used, including special programming languages, orientation to the symbolic information processing (LISP, SMALLTALK, REFAL), logic programming languages (PROLOG), knowledge representation languages (OPS5, KRL, FRL), software environment integrated, expert system shells, etc.

So, the following can be recorded.

If agricultural robot software has automated navigation and manoeuvring functions in the field space, routing when performing tasks, and a number of processes are partially automated (for example, automatic soil sampling by samplers, automatic differential fertilizer application which doses are calculated by man, etc.), and agricultural robot is equipped with a computer vision system, required for work in difficult conditions, high-performance communication systems and intelligent attachable equipment, then such controlled object can be considered as **intelligent agricultural robot**.

If agricultural robot software has the functions of its own mission development for the general description of agricultural task, autonomous movement across the field with obstacles avoidance and high-precision navigation, processing in the on-line photo and video information format solving situation recognition issues (for example, plants diseases recognition with acceptable accuracy and weeds and plant pests detection), recommendations development for decision-making in a particular situation and consequences prediction caused by solutions taken, and agricultural robot learns the rules for working in the field in the course of its development and corrects them subject to its operation and responds to human voice commands, such controlled object can be considered as **highly intelligent agricultural robot with the elements of "weak" artificial intelligence.**

5. Prospects for the national agricultural robotics development

It can be predicted the following perspective.

It shall be developed three main classes of agricultural robots in the Russian Federation.

Class A. **Large ground-based robots** (weighing over 1 ton) based on unmanned tractors with parallel and self-driving technologies. Drivers-controllers will be provided at the initial stage in the tractor cabin to simplify the cost of large agricultural robot. Tractors will vary by subclass, most likely by total weight and functionality. The tractors will be equipped with self-driving systems and satellite positioning receivers, lidar-based software and hardware computer vision systems, position correction modules considering relief, etc. The relevant attachable equipment will be installed on the tractor robots. This class of agricultural robots will be used primarily for agricultural works on large areas possibly preliminary prepared for efficient robotic farming. Such robots will have to implement robotic farming technologies.

The efforts of developers will be focused on remote and autonomous agricultural robot control technology and master-slave or follow me technology. In practice, transition to such robots will be gradual, from "one driver - two tractors" principle and then telecontrol and autonomous tractors with self-driving function operating under a given program.

At the time of this publication, it is known about three Russian conceptual solutions and about ten foreign solutions for this class of agricultural robots.

Class B. **Small ground-based robots** (weighing up to 1 ton) based on robot lorries equipped with intelligent attachable equipment and specialized robotic manipulators to solve agricultural tasks, mainly precision farming. Such robots will be a remotely controlled by radio channel (semi-autonomous, autonomous) intelligent robotic on wheel or track platform. Computer vision system will be installed on the agricultural robot for safe movement and obstacles avoidance, with significant ranges in azimuth and angle of elevation. Optionally, such agricultural robots will be equipped with photo and video recording system to obtain photo and video images of the situations in the specified field areas (delivered and approaching to research site by manipulator). The platform has radio control antenna, attachable equipment for solving specific tasks and special type manipulator installed.

In general, in medium term, a gradual transition from large, heavy and man-controlled agricultural machineries to a variety of small, light, inexpensive and specialized autonomous robots is predicted. This class of agricultural robots will be used primarily to solve precision farming issues, in particular, soil sampling, variable fertilization, weeds and pests control, etc.

Class B. **Small air** agricultural robots based on Unmanned Aerial Vehicles of aircraft- or quadcopter-type. As a rule, they will be equipped with high-resolution video and photo cameras with payload weight (video camera and multispectral camera) up to 5-10 kg. GPS or GLONASS 3G/4G modems will provide agricultural robot navigation. This class of agricultural robots will be used to solve monitoring issues (primary purpose), determine plant diseases and perform simple operations on the field.

It can be assumed that for all types of robots the so-called "service" model of interaction with the Customer will become most effective, in which all functions of the robot will be presented as software-defined information and active services implemented by technical means. The task for agricultural robot will be generated through a special cloud service in a dialogue with the Customer, where it will be able to select any task from wide menu or formally describe the task. After task processing with resources availability checking, terms feasibility, etc., the Mission Plan will be agreed upon and operational flow charts will be drawn up for the respective types of field works, agricultural machineries and agricultural robots. Contrary to the modern type of charts, they will contain not only the working conditions and agrotechnical requirements for the operations performed, but also routes for agricultural robots movement through the fields, characteristics of robotic operations, etc. Agricultural robots will be controlled either by the Operator over the radio channel or (if communication is lost) autonomously according to the program stored in the robot pilot memory of the agricultural robot control system.

Agricultural robots will be an important element for field situational awareness replenishment and large agricultural data generation provided by agricultural enterprise. This data will be located in the Internet clouds. Agricultural robots servicing is likely to require sufficiently qualified specialists.

Robot-oriented farming will lead to the new businesses that will be supported by specialized service companies.

Conclusion

Agricultural robots will be used in solving precision farming issues making a significant competitive advantage for the agricultural enterprises in the next 15-20 years. Currently, negative trend in national robotics gap with foreign developments, which are still at the stage of experimental and pilot projects, is being formed.

Domestic developments from large agricultural machine engineers in the field of agricultural robotics are essentially oriented on innovations actively introduced abroad; the main one is a technology of automatic parallel driving.

Special unmanned aerial vehicles are also among the promising agricultural robots. Although very useful but limited functionality will be developed on their basis, aimed mainly at solving informational but not production issues.

Due to objective reasons, robotic farming technologies in the Russian Federation will be exotic for a long time.

It is justified that there is a certain technical and methodological “niche” within which it is possible to obtain not only import substituting, but also export-oriented solutions. This “niche” includes: a) in the methodological terms - development of the models for plant diseases recognition, pests and weeds detection aggregated with agrocytosis models and b) in the technical terms - making small agricultural robots with intelligent attachable equipment, including specialized manipulators for solving precision farming issues.

References

1. Z.A. Godzhaev, A.P. Grishin, A.A. Grishin, Prospects for the robotic technologies development in crop growing // Tractors and agricultural machinery. - 2015. - No.12. - pp. 42-45.
2. Bechar, A.; Vigneault, C. Agricultural robots for field operations: Concepts and components. Biosyst. Eng. 2016, 149, 94–111.
3. Wable A., Bhongal M.B., Yeole K. Mukund, Bedke A.A. An Autonomous Polyhouse Robot for Plant Health Indication and Detection of Plant Disease using Image Processing, International Journal for Scientific Research & Development| Vol. 3, Issue 12, 2016- pp. 355-358
4. V.V. Yakushev Precision farming: theory and practice-SPb.: FSBI API, 2016, - 364 p.
5. V.K. Abrosimov, V.V. Eliseev, Small intellectual robots for solving precision farming issues: issues and solution, Robotics and Technical Cybernetics, 2019, No.4 (21), pp. 14-19.
6. Aiello M., Johnsen E.B., Dustdar S., Georgievski I. (eds.) Service-Oriented and Cloud Computing, ESOCC, Springer, 2016-266 p.
7. V.I. Gorodetskiy, M.S. Grushynskiy, A.V. Khabalov, Multi-agent systems (review) [Electronic resource]. Available at: http://spkurdyumov.ru/networks/mnogoagentnye-sistemy-obzor
8. V.K. Abrosimov, V.V. Eliseev, An approach to solving precision farming issues in the Internet paradigm of agricultural “things”, Agricultural machinery and technologies, 2019, No. 3, in press.
9. E.M. Kussul, Associative neural structures. - Naukova Dumka. Kiev. 1992. - 144 p.

O.B. Shagniev, S.F. Burdakov

THE ROBOT VIBRATION CONTROL UNDER EXTREME LOADS DURING MACHINING

Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia
shagniev_ob@spbstu.ru

Abstract

The problem of the occurrence and rapid suppression of unstable self-excited vibrations arising in the process of turning is considered. It is assumed that tool is connected with manipulator by an elastic suspension, which is used for force sensation. The tool moves evenly along the work surface with a given pressure on it. Pressing of the tool provides the necessary axial depth of cut. Uniform movement along the work surface provides the required tool feed. Unstable self-excited vibrations or chattering is a deterrent to increase productivity. In this paper we consider the possibility of promptly detecting the onset of unstable auto-oscillations from the amplitude spectrum of the sensor readings of the horizontal forces of interaction between the instrument and the working surface. The amplitude spectrum is obtained using the fast Fourier transform, which allows to determine the beginning of unstable processes in system. Timely change of the axial depth of cut allows to transfer the turning process into the stable zone.

Keywords: robot; chattering; adaptive control; adaptive control; position-force control; math modeling

Introduction

Metal machining is one of the main methods of parts manufacturing in mechanical engineering. Due to the desire to expand the scope and increase productivity, optimization and intensification of the machining process are actual. The great diversity and complexity of physicomechanical processes, including thermodynamic processes in the tool-surface contact area, makes it necessary, in addition to the force sensation using the elastic tool suspension, to use various elements of artificial intelligence. This is necessary not only because of the complexity of the models, but also with a priori uncertainty in a wide range of modes of contact interaction.

In practice the main limiting factor in improving machining performance is the possible loss of dynamic stability caused by tool vibrations. Vibrations cause tool breakdowns, premature wear of the cutting edges, reduced quality and precision of machining. In practice, to minimize the probability of system stability losing, the parameters of the machining mode are consciously underestimated, which naturally leads to performance degradation.

According to the existing hypotheses [1], the cause of self-excited oscillations is the formation of a self-sustaining oscillatory mechanism in the process of chip formation. The limited stiffness of the tool causes the relative movements of the tool and the workpiece caused by the interaction forces, which in turn can lead to the formation of a wavy cutting surface with each new cutter pass. In this case, the wavy surface left on the previous turn of the workpiece is removed during the next pass [2]. This mechanism leads to the formation of waves on both sides of the chip, the thickness of which depends on the phase shift between them. As a result, cutting forces can increase without limit.

For turning process, Budak E. and Altintas Y. developed methods for determining the dynamic stability conditions of the system, allowing to relate the value of the axial depth of cut, feed and spindle speed, corresponding to a stable area. Further research by many authors has allowed us to develop a number of practical ways to reduce vibrations during milling. Today the most common ways of dealing with vibrations are offline (for example speed modulators). The developing of an adaptive control system for the turning process, which is capable to adjust the machining parameters in the online mode, remains actual.

Many studies focuses is on the design features of CNC machines. This seriously limits the ability to quickly adjust the parameters of the machining in the online mode. From the point of view of the control system flexibility and the possibility of spatial processing of the arbitrary profile parts, the using of multi-axis robots with a tool installed in an elastic suspension is perspective. Elastic suspension provides a force sensation of the robot in at least three axes. This configuration allows to use the standard robot manipulator in the hybrid position-power control mode, in which robot moves taking into account the contact forces of interaction between the tool and the working surface. Installation the tool in an elastic suspension provides additional opportunities for organizing adaptation contours that can predict the possible loss of machining process stability and provide the necessary ratio of cutting parameters.

In this paper, we investigate the possibility of the axial depth of cut automatic correction when signs of dynamic stability loss appear. In this case, the corresponding time moment is detected using the amplitude spectrum of the horizontal cutting force. Amplitude spectrum is obtained using the fast Fourier transform (FFT). The results of computer simulation of the adaptation contour are presented.

Mathematical model of regenerative self-excited vibrations

The robot equipped with a cutter, in accordance with the technological task, moves at a certain speed along the workpiece with a given pressing to it. To ensure the required level of the tool to the surface force pressing, a position-force control algorithm [3] is built based on the formation of a corrected task for a regular positional control system of the robot based on the vertical force sensor readings feedback. The horizontal movement control of the robot is based on the positional control algorithm based on the velocity feedback and provides the required tool feed. Calculation model of the robot in contact mode is shown in Figure 1.

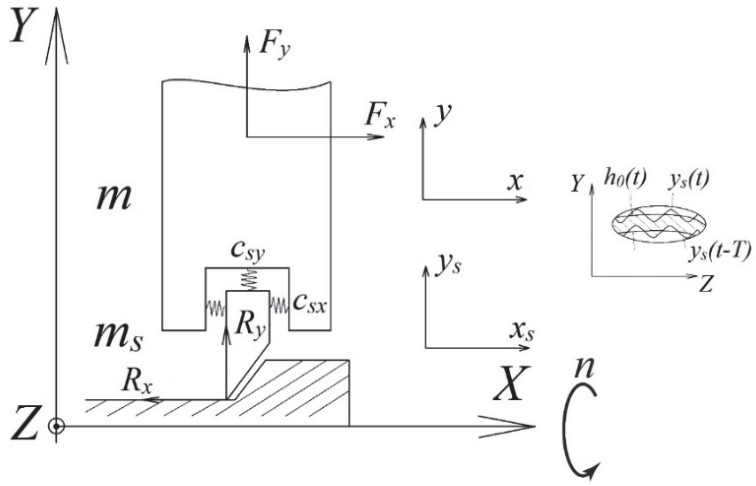


Figure 1 – Calculation model of the robot

In Figure 1, the following notation is introduced: m – equivalent mass of robot arm; m_s – equivalent mass of the tool; c_{sx}, c_{sy}– tool suspension stiffness; y – robot arm vertical coordinate; y_s – tool vertical coordinate; x – robot arm horizontal coordinate; x_s– tool horizontal coordinate; $y_{sf}(x)$ – surface equation; R_x, R_y – cutting forces acting on the tool from the machining surface. It is assumed that the coordinates of the robot and the tool take into account their structural dimensions. Task $\widetilde{y^d}$ is formed using the PID controller based on the vertical interaction force feedback $F_{sy} = c_{sy}(y - y_s)$.

In the contact mode the mathematical model of the sensible robot as a control object is

$$m\ddot{y} + b_y\dot{y} + c_{sy}(y - y_s) = F_y - mg,$$

$$m_s\ddot{y}_s + b_{sy}\dot{y}_s + c_{sy}(y_s - y) = R_y,$$

$$F_y = \left(k_p + k_i\frac{1}{p} + k_d\frac{N_y}{1 + N_y\frac{1}{p}} \right)\left(\widetilde{y^d} - y\right),$$

$$\widetilde{y^d} = \left(\theta_p + \theta_i\frac{1}{p} + \theta_d\frac{N_f}{1 + N_f\frac{1}{p}} \right)\left(F_{sy}^d - F_{sy}\right),$$

$$m\ddot{x} + b_x\dot{x} + c_{sx}(x - x_s) = F_x,$$

$$m_s\ddot{x}_s + b_{sx}\dot{x}_s + c_{sx}(x_s - x) = R_x,$$

$$F_x = \left(k_{px} + k_{ix}\frac{1}{p} + k_{dx}\frac{N_x}{1 + N_x\frac{1}{p}} \right)\left(\dot{x}^d - \dot{x}\right).$$

The chip thickness can be expressed as

$$h(t) = h_0 - \big(y(t) - y(t - T)\big),$$

where h_0 – constant component of chip thickness, which is equal to the feed per revolution, $y(t) - y(t-T)$ – dynamic component of chip thickness. Cutting force can be written as

$$R_y = K_f a h(t) = K_f a\left(h_0 - \left(y(t) - y(t-T)\right)\right)$$

where K_f - cutting constant in feed direction, a - depth of cut.

Analysis of the stability of the system [4, 5] allows to write the expression for the maximum depth of cut

$$a_{\text{кр}} = -\frac{1}{2K_f G(\omega_c)},$$

$$T = \frac{2\pi k + \epsilon}{2\pi} = \frac{60}{n}.$$

where $G(\omega_c)$ – the real part of the transfer function of the system from the cutting force to the tool coordinate; ω_c – frequency of regenerative self-excited oscillations; ϵ – phase shift between external and internal modulation; n – workpiece rotation speed; T – rotation period of the workpiece.

Thus, in the general case, the system of equations describing the milling process is a system of nonlinear differential equations with delay. Analysis of the system stability according to Altintas Y. allows us to build a stability lobe diagram in the space of the axial depth of cut and the spindle speed. In computer simulation of the milling process, the following parameter values were used:

$$m = 2\ kg; \quad m_s = 0.2\ kg; \quad K_f = 10^8 \frac{\text{N}}{\text{m}^2};$$

$$c_{sx} = c_{sx} = 10^4\ \frac{\text{N}}{\text{m}};\ b_s = 20\frac{N \cdot s}{\text{m}};\ b = 50\frac{\text{N} \cdot \text{s}}{\text{m}}.$$

Fig. 2 shows the stability lobe diagram of the turning process.

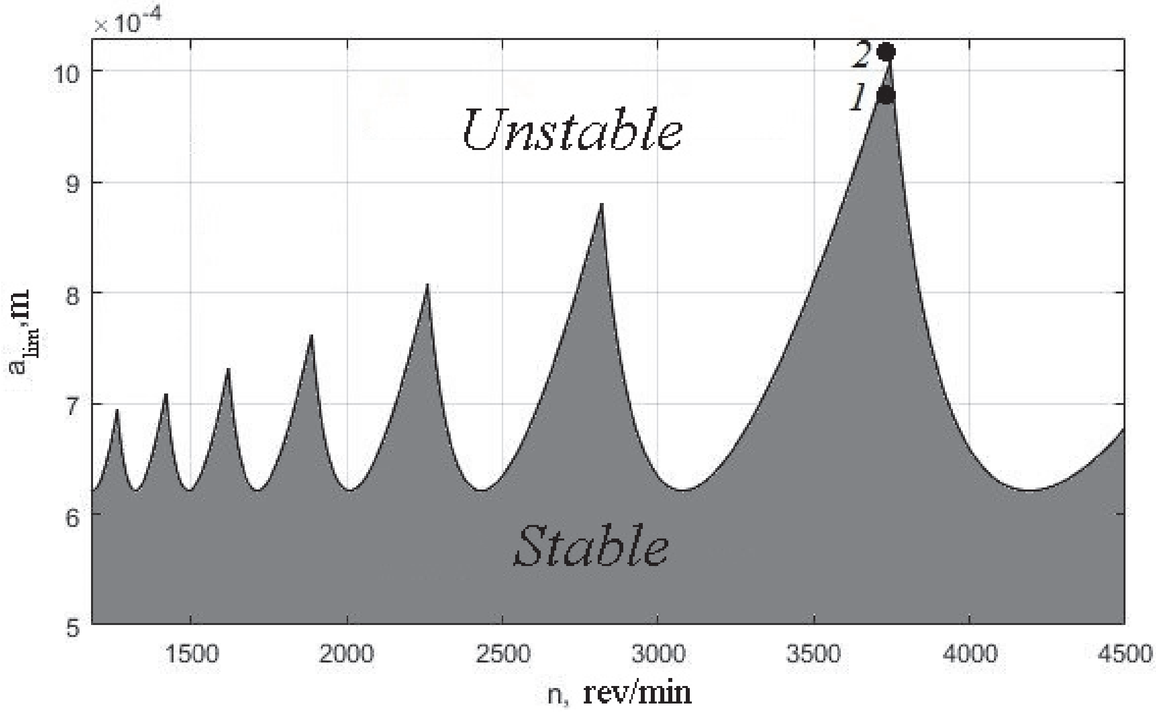


Figure 2 – Stability lobe diagram

In Figure 2 point 1 corresponds to stable cutting and point 2 to unstable. Figures 3 and 4 present the time history of the tool coordinate, the vertical force sensor readings and its fast Fourier transform (FFT) for points 1 and 2.

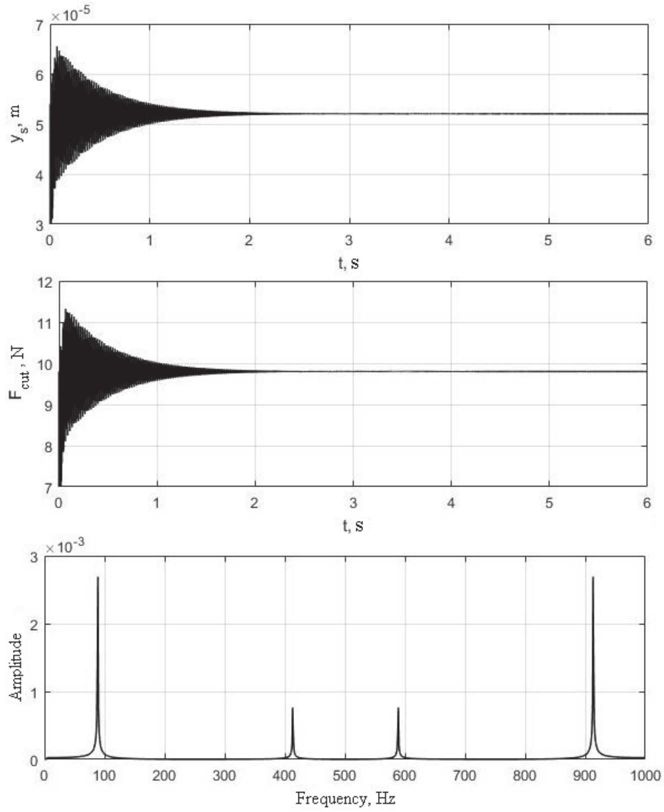


Figure 3 – Time history of the tool coordinate, the vertical force sensor readings and its FFT for point 1

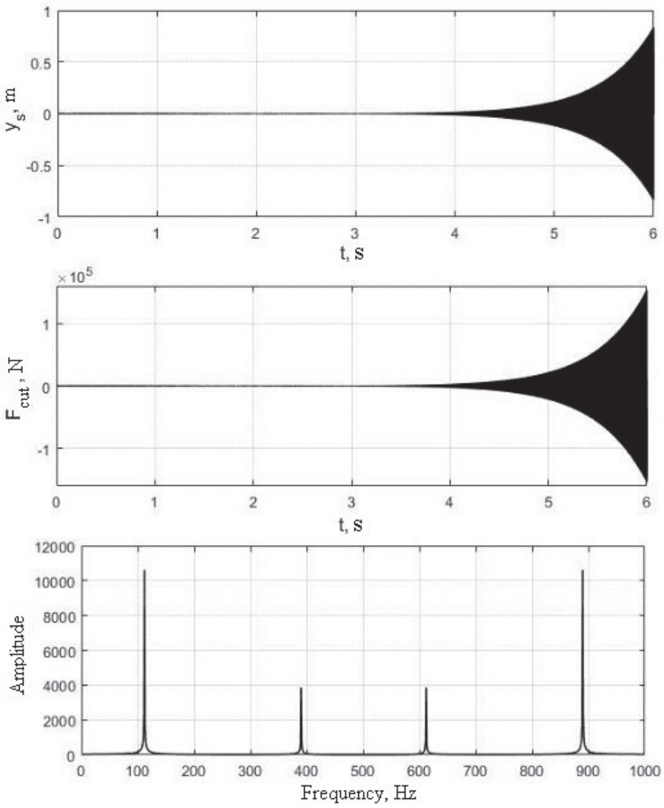


Figure 4 – Time history of the tool coordinate, the vertical force sensor readings and its FFT for point 2

Adaptive stabilization of unstable self-excited vibrations by changing the axial depth of cut

Significant for the stability of the turning process system parameter, which can be controlled using the position-force algorithm, is the depth of cut, which is strictly related to the vertical coordinate of the tool y_s.

Based on the measurements of force sensors, in accordance with expert evaluation (EE), which includes data from the stability diagram, a modified task along the vertical coordinate is formed in the logic block. To predict the instability of the system, the FFT of the cutting force is used. At the moment of the beginning of the increase of the peak associated with the frequency of self-excited oscillations, the control system reduces the depth of cut. The block diagram of the control system is shown in Figure 5.

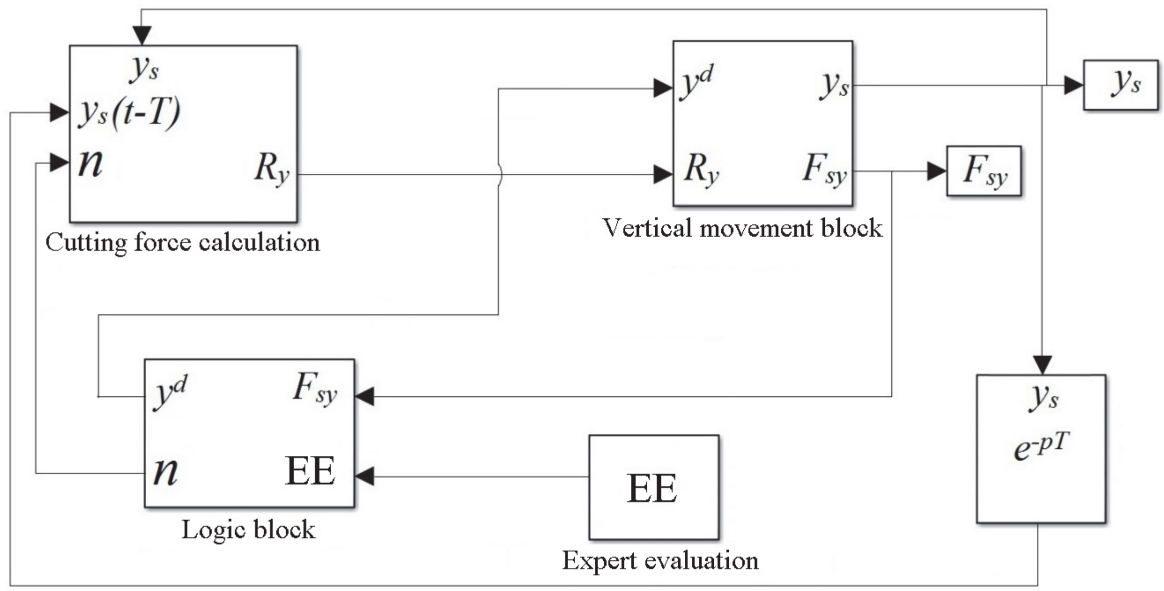


Figure 5 – Block diagram of the control system

Figure 6 shows the time histories of the tool coordinate and vertical force sensor readings for the system transition mode from an unstable (point 2 in Fig. 3) to a stable (point 1 in Fig. 3) state.

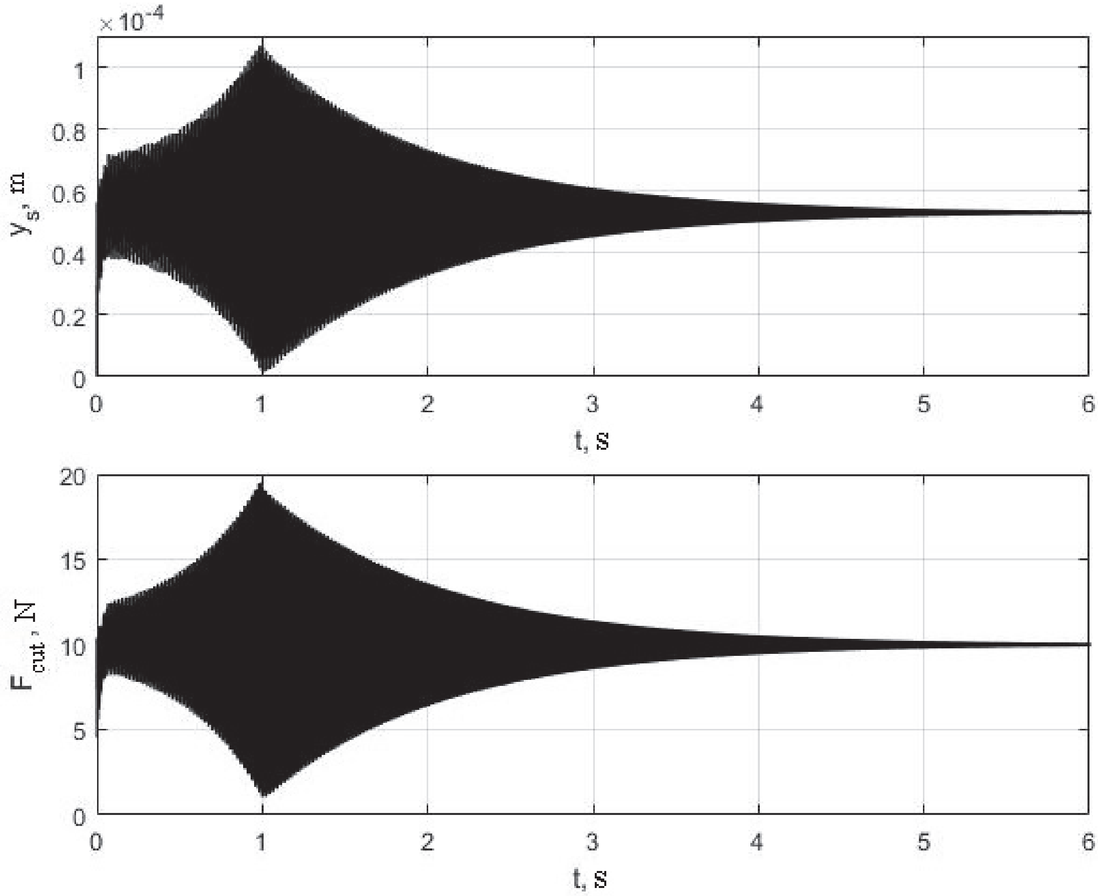


Figure 6 – Time histories of the tool coordinate and vertical force sensor readings

The construction of additional adaptation contours allows to prevent the occurrence of undesirable regimes during mechanical processing. In terms of the use of robotic systems in mass production, the task of intellectualizing adaptation contours with the help of trained artificial neural networks is an urgent task. [6]. The block diagram of the adaptive control system is shown in Figure 7.

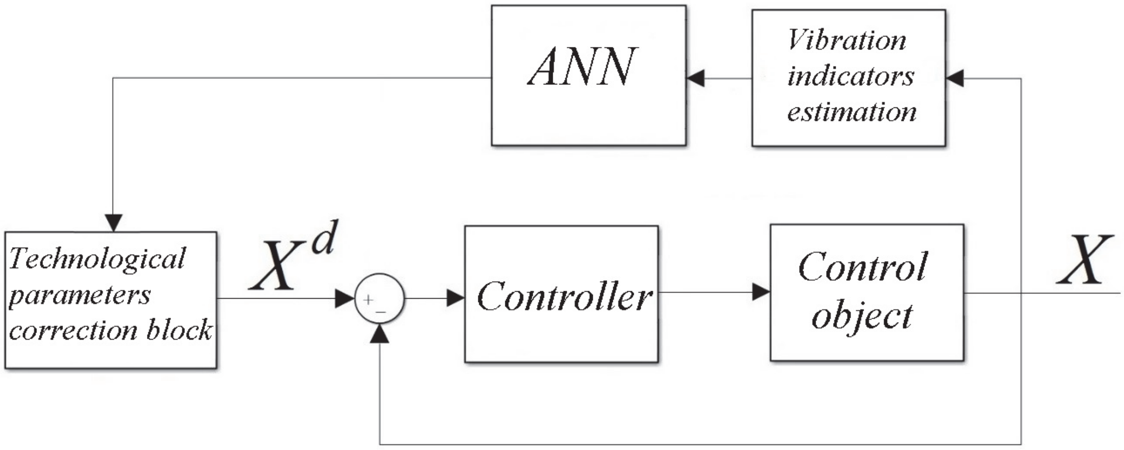


Figure 7 – Block diagram of the adaptive control system

Conclusion

The introduction of such systems allows large batches to seriously reduce the scrap rate and increase productivity. Artificial neural network (ANN) in terms of the vibration state of the system [7], receiving with the help of force sensors, outstanding commands to change the technological parameters that affect the accuracy and performance of the turning process. If any signs of loss of stability or unsatisfactory quality of machining occur, the adaptation contour adjusts the task for the position-force control system, which provides the new required machining parameters.

References

1. Siddhpura M., Paurobally R. A review of chatter vibration research in turning // International Journal of Machine Tools and Manufacture. 2012. Vol. 61. P. 27-47.
2. Filippov A.V., Rubtsov V.E., Tarasov S.Yu., Podgornykh O. A., Shamarin N. N. Detecting transition to chatter mode in peakless tool turning by monitoring vibration and acoustic emission signals // The International Journal of Advanced Manufacturing Technology. 2018. Vol. 95. P. 157-169.
3. Burdakov S.F., Shagniyev O.B. Modeli mekhaniki v zadache upravleniya silovym vzaimodeystviyem robota s poverkhnostyu neopredelennogo profilya [Mechanics models in the control problem of the force interaction between a robot and a free-formed surface]// Nauchno-tekhnicheskiye vedomosti SPbGPU. Informatika. Telekommunikatsii. Upravleniye. 2015. Iss. 4. P. 68-79.
4. Budak E. Maximizing Chatter Free Material Removal Rate in Milling through Optimal Selection of Axial and Radial Depth of Cut Pairs// CIRP Annals – Manufacturing Technology. 2005. Vol. 54. Iss. 1. P. 353 – 356.
5. Altintas Y. Metal cutting mechanics, machine tool vibrations, and CNC design. 2nd-ed. Cambridge University press. 2012. 382 p.
6. Chuangwen X., Jianming D., Yuzhen C., Huaiyuan L., Zhicheng S., Jing X. The relationships between cutting parameters, tool wear, cutting force and vibration// Advances in Mechanical Engineering. 2018. Vol. 10(1). P. 1–14.
7. Lamraoui1 M., Barakat M., Thomas M., El Badaoui M. Chatter detection in milling machines by neural network classification and feature selection// Journal of Vibration and Control. 2015. Vol. 21(7). P. 1251–1266.

V.P. Andreev, V.L. Kim

MODULAR ARCHITECTURE OF A MOBILE ROBOT TRANSPORT PLATFORM FOR A MOTION TASK ON A ROUGH TERRAIN

MSTU "STANKIN", IINET RSUH, Moscow, Russia
andreevvipa@yandex.ru, top7733@gmail.com

Abstract

This paper deals with a design of a mobile robot transport platform with modular architecture build as distributed system. The necessity of transport module's structure decomposition into submodules is due to computational complexity of navigation tasks in case of a mobile robot motion in nondeterministic environments. It is shown that solving of a trajectory tasks using potential fields for motion on perfectly even surface is possible when computational control process is distributed between two microcontrollers of small computational power. The transition to consideration of scenarios in which modular robot operates in an environment with complex terrain causes the complication of navigation algorithms. To implement these algorithms in real time, it is proposed to consider the transport module as a modular architecture which should include unified nodes-submodules. A set of these submodules will allow rapid reconfiguration of a transport module structure in accordance with the desired goal.

Keywords: mobile robot, modular architecture, reconfigurable robots, distributed computing systems, computer network, trajectory tracking.

Acknowledgments

Research is supported by the Russian Foundation for Basic Research: Grant 19-07-00892a.

Introduction

The successful completion of the majority of mobile robot's (MR) practical tasks requires from an onboard computer execution of a great number of mathematical computations with significant amount of data. The computational load is quite large even when a mobile robots act in conditions of well-developed urban infrastructure, office and laboratory buildings. The complexity of algorithms for supervisory or autonomous navigation increases even more in environments where human presence is prohibited due to health or life risk.

In significant part of known works mobile robot motion planning is considered only on flat surfaces that can be found not very often even in office compartments. An autonomous motion on uneven terrain requires different algorithms for path planning and tracking with integration of more sensors. This leads to significant increasing of mobile robot's onboard computer computational load.

The distributed structure of modular robot with full-featured modules

A distribution of computational load may be performed by robot structure decomposition onto unified full-featured modules [1-2]. The research was supported by the RFBR: Grant 16-07-00811a. According to the proposed concept every module is functionally independent and solves strictly defined set of tasks. The full-featured modules united into one robotic system must interact between each other using special protocol [3] that standardizes modules command system, transmission data format, message priority etc. These modules are responsible only for particular functions so they require less computational resources compared with a whole robot. This is a key advantage of a modular architecture along with reconfiguration possibility and rapid development of new robotic devices.

Nevertheless some of the robot modules must perform significant amount of work despite present function distribution between mobile robot nodes. For example the transport module is responsible for navigation in environment, obstacle avoidance, path following and low-level control of the actuators. The transport module performs all these tasks in cooperation with a sensor module that in turn is in charge of systematization of the data obtained from an environment (mapping, obstacle shape determination, etc.).

Realization of the transport module motion algorithm using potential fields method

The algorithm of motion along splines with use of potential fields was proposed for the transport module [4]. This algorithm allows the modular mobile robot with differential drive to move smoothly along splines.

The proposed algorithm is a particular case of navigation task – trajectory tracking, the only difference is that robot trajectory is defined implicitly. The motion of the modular robot along spline like trajectories (B-splines) was achieved using a set of control points on the plane and potential fields linked with them.

The algorithm was implemented on the modular mobile robot laboratory prototype (fig. 1). The transport module is a wheeled platform with a differential drive. The control system hardware consists of two parts: high level and low level units. The module is also divided physically in two parts. At the core of the high level unit the microcontroller board Arduino Due is situated, at the core of the low level unit we use Arduino Mega 2560. Two geared DC motors with Hall effect sensors are used as electric drives, voltage regulation is performed by motor drivers L298N.

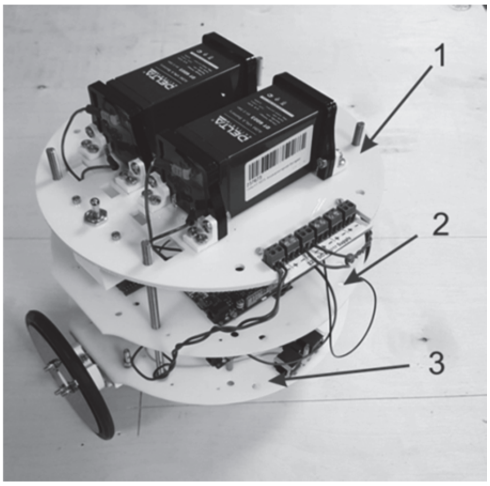


Figure 1 – Modular mobile robot laboratory prototype: 1 – the power module, 2 – the transport module high level part, 3 – the transport module low level part

Arduino Due is a system main controller and it's responsible for proposed motion algorithm calculation. Arduino Due is a microcontroller board based on AT91SAM3X8E ARM Cortex-M3 CPU (32-bit ARM core) with 84 MHz clock. According to algorithm Arduino Due computes desired linear V and angular ω robot velocities that are reference-input signals for Arduino Mega board. Arduino Mega performs low-level control: actuators control, encoder data processing and odometry. The board converts linear and angular velocities from Arduino Due to desired angular velocities of module's wheels. The wheels speed control was performed using integer PID controller with filtration and anti-windup technique [5].

Two boards communicate via I2C bus (fig. 2). The choice of I2C bus is conditioned by the next reasons: simplicity of electrical connection, high transmission speed for desirable traffic volume and integrated software libraries for programs.

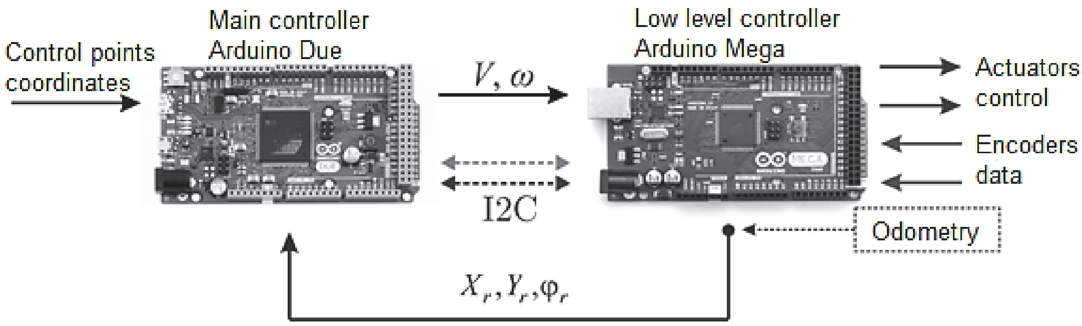


Figure 2 – Scheme of interaction of high and low levels boards of the transport module control system hardware part

Working frequency of internal electrical drive control loop is 100 Hz. Working frequency of external closed-loop control system with robot position feedback is also 100 Hz (it's a I2C bus transmission frequency). This frequency is enough for mobile robot motion with maximum linear velocity 0.5 m/s.

The utilization of two microcontroller boards is due to situation that proposed algorithm could not be implemented on the single board (Arduino Mega 2560) even with a low velocity of the robot (less than 0.3 m/s). At the same time it was planned that every full-featured module must be based on a low performance microcontroller [6]. The proposed design could not be implemented because of several main reasons:

– low computing speed of the board especially when evaluating floating point numbers;

– necessity for simultaneous solving of a large number of tasks: low-level control of actuators, odometry, motion algorithm calculating;

– the motion algorithm is quite demanding of a module controller: almost all evaluations include operations with trigonometric functions, also the potential fields described by two-dimensional Gaussian functions (computation of exponents).

Note that a computational load of the module onboard computer is large even such important parts of navigation task are absent: path planning (control points evaluation), obstacle avoidance, sensory data fusion etc.

Distribution of the tasks between two boards connected via I2C bus allowed to successfully perform experiments using proposed motion method.

In all experiments path of the modular mobile robot was compared with B-splines which control points are coincident with algorithm control points.

Experiment 1. In the first experiment an open uniform second order B-spline with control points (0, 0), (0, 0.75), (1.5, 0.75) and (1.5, 0) was set (all values in meters). The vector of knots is $u = (0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1)$. Figure 3 shows the next curves:

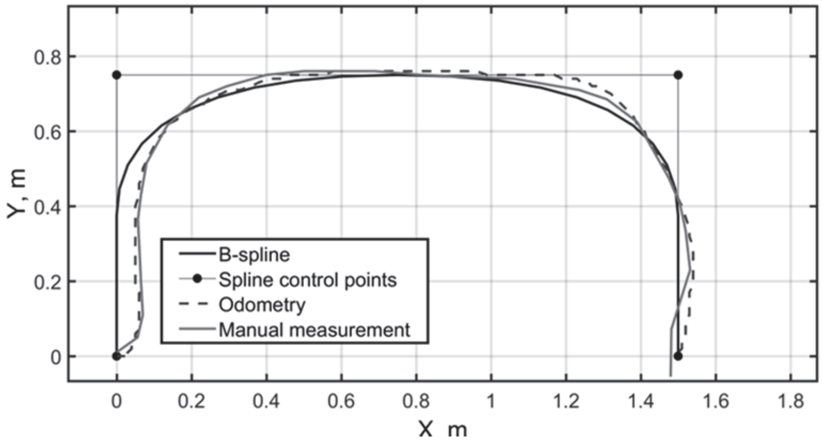


Figure 3 – B-spline and robot trajectories with control points (0, 0), (0, 0.75), (1.5, 0.75) and (1.5, 0)

The maximum deviation of the robot from target curve (B-spline) is not exceed 0.069 m (69 mm with a total path length 2.69 m) moreover it occurs at the initial moment of motion, when the robot start to approach to desired trajectory defined by potential fields.

The robot deviation from goal position (1.5, 0) is about 0.02 meters along X axis and -0.055 meters along Y axis; position error is amount to 0.058 meters (58 mm). From the plot on the fig. 3 one can see that this error is a maximum deviation of a practically measured position from a position obtained using odometry.

Figure 4 shows linear and angular velocities of the mobile robot during the motion.

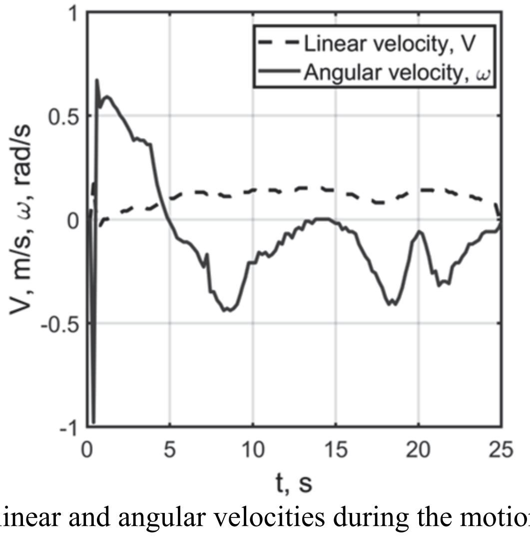


Figure 4 – Robot linear and angular velocities during the motion along the spline with control points (0, 0), (0, 0.75), (1.5, 0.75) and (1.5, 0)

From the plot on the fig. 4 it is apparent that in steady-state control points tracking by the robot angular velocity doesn't exceed 0.59 rad/s and linear velocity doesn't exceed 0.15 m/s. One can see the smooth change of velocities except that initial moment (robot switch on by the tumbler). This result corresponds to computer simulations [4].

Experiment 2. In the second experiment an open uniform second-order B-spline with control points (0, 0), (0.5, 0.75), (1.5, 0.75) and (1.5, 0.75) was set; other B-spline parameters stayed the same. Figure 5 shows the B-spline and robot trajectory obtained using odometry.

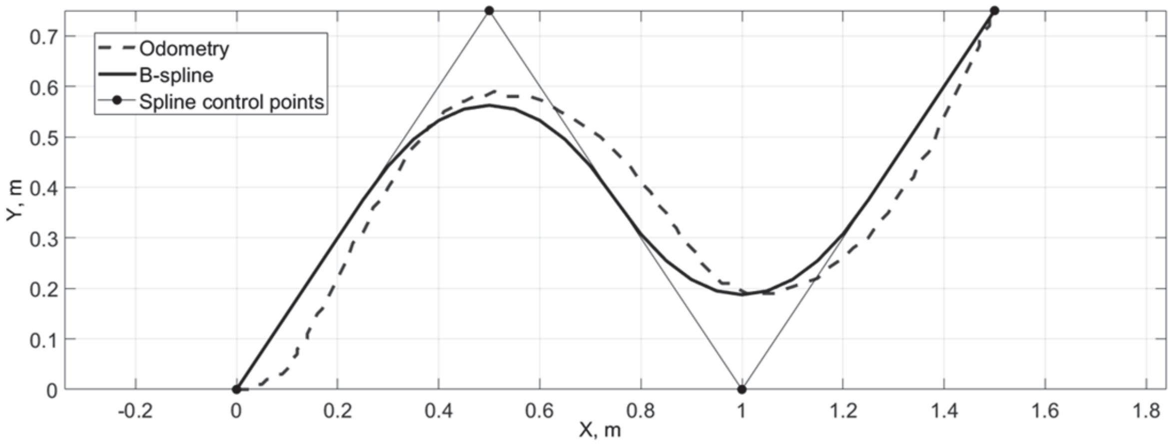


Figure 5 – B-spline and robot trajectory with control points
(0, 0), (0.5, 0.75), (1.5, 0.75) and (1.5, 0.75)

The maximum deviation of the robot from desired B-spline doesn't exceed 0.065 m (65 mm with total path length 2.31 m). At the initial moment there is a position error (about 0.064 m) when the robot motion settling on the spline.

Figure 6 shows mobile robot linear and angular velocities during the motion.

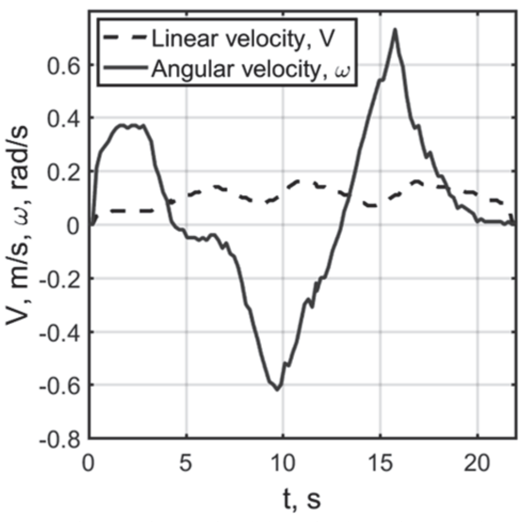


Figure 6 – Robot linear and angular velocities during the motion along the spline with
control points (0, 0), (0.5, 0.75), (1.5, 0.75) and (1.5, 0.75)

As in previous case velocities change smoothly; moreover as it seen from plot when the angular velocity increases the linear velocity decreases i.e. robot slows down on turns and accelerates on straight paths. Linear velocity changes slightly and doesn't exceed 0.16 m/s. The maximum angular velocity is 0.73 rad/s.

The results of performed experiments show efficiency of proposed motion algorithm of the mobile robot using potential fields. The motion trajectories match provided B-splines but some position errors are present which is due to the following main reasons:

1) an odometry inaccuracy resulting from position error accumulation during robot motion;

2) no functional relation between algorithm parameters (for example Gaussian function width) and geometrical parameters of B-spline (spline degree for example);

3) absence of robot dynamics estimation in the proposed algorithm.

The further transport module functional possibilities enhancement will lead to addition of new algorithms in the high level unit control system. Moreover robot motion was considered on a perfectly even surface. It is expected that the transport module must allow robot movement not only in environments with level floor but on surfaces with different properties and complex terrain.

In accordance to increasing computational needs of onboard computers mobile robots can be designed based on three most commonly used schemes [7]:

1. **Fully centralized control system** – one central processor fully responsible for robot's behavior, reading sensors, processing data, low-level actuators control etc.
2. **Distributed hardware (mechatronics) and centralized control** (partially distributed control system) – system is designed using central embedded processor and microcontrollers which regulate actuators and process data of a robot sensors.
3. **Distributed hardware (mechatronics) and distributed control** (fully distributed control system) – system is based on microcontrollers in such a way that central processor is not needed (but it can be set up optionally).

The advantage of the fully centralized control system is a simple hardware and software design because there is no need in a setup of communication between distinct nodes (all communication goes through a CPU). But with a number of sensors and actuators increasing this topology becomes not quite appropriate for several reasons [8]:

1) a common bus may be overloaded with big data flows which in turn causes robot reaction delay to external stimuli;

2) an addition of new devices to the system may be difficult because all signal and electrical cables have to be wired through the whole robot construction;

3) the number of devices that can be connected to the central processor is limited to the number of I/O ports (for example limited PWM channels for motors control).

The main disadvantage of centralized systems is their reliance on a one processor; a system scalability is bounded and reconfiguration possibility is missing.

Distributed systems have next advantages:

1) a computational load is distributed between all system nodes which allow to increase robotic device processing speed while performing complex tasks;

2) a possibility of new devices addition – more expandability of the system;

3) a possibility of a robot reconfiguration.

The main disadvantage – a distributed system is generally less reliable because include large number of microcontrollers; regulation of information interaction between these microcontrollers gives rise to a second problem – choice of communication medium (bus), message format, command system, etc., which ensure all robot nodes working in real-time.

According to the concerned concept mobile robot has distributed modular architecture with full-featured nodes. But as was described above the transport module solves many tasks which require performing a large number of math operations especially when robot motion is considered on a rough terrain. As a consequence the transport module itself must be built as a *modular* architecture with distributed nodes – *submodules*.

The transport module distributed architecture for motion on a rough terrain

The distributed architecture of the transport module is based on the classical hierarchical structure of an intelligent mobile robots control systems (fig. 7) proposed by Popov [9]. The basis of the structure consists of four levels: intellectual, strategic, tactical and executive.

The intelligent level in the transport module is not considered, since it is assumed that the module-supervisor of the mobile robot is responsible for it [1]. All other levels can be represented by **submodules** with limited functions.

The main property of the original concept – full functionality of the device is transferred to submodules, i.e. each submodule may not have complete information about the structure of other submodules, except for the most common parameters and properties.

This architecture should allow distributing the computational load of different functional significance between submodules, especially in cases where the number of actuators and sensors may increase or their type may change. This may be true for mobile robots off-road type. In addition, it is possible to combine submodules, which provides reconfigurability of the transport module.

The structure of the distributed architecture of the transport module is shown in fig. 8.

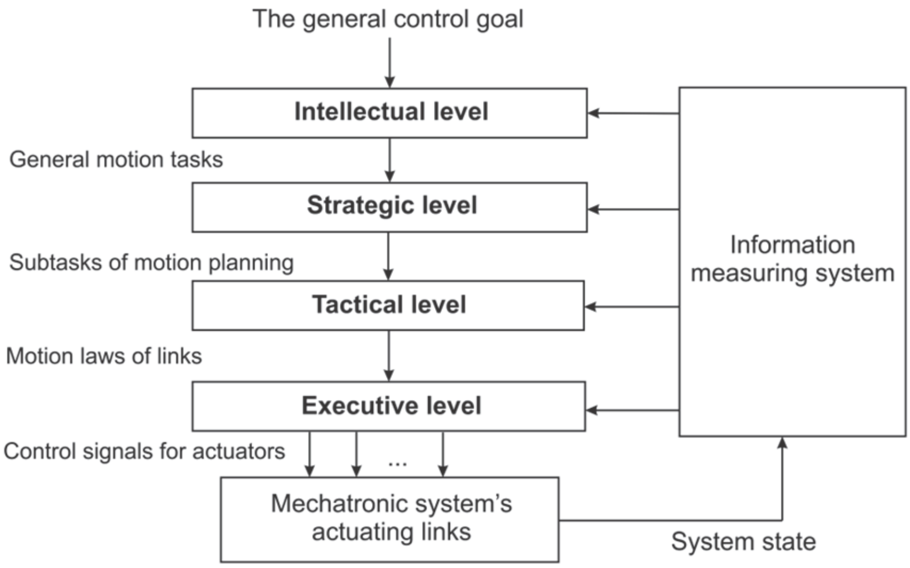


Figure 7 – The classical hierarchical structure of an intelligent mobile robots control systems

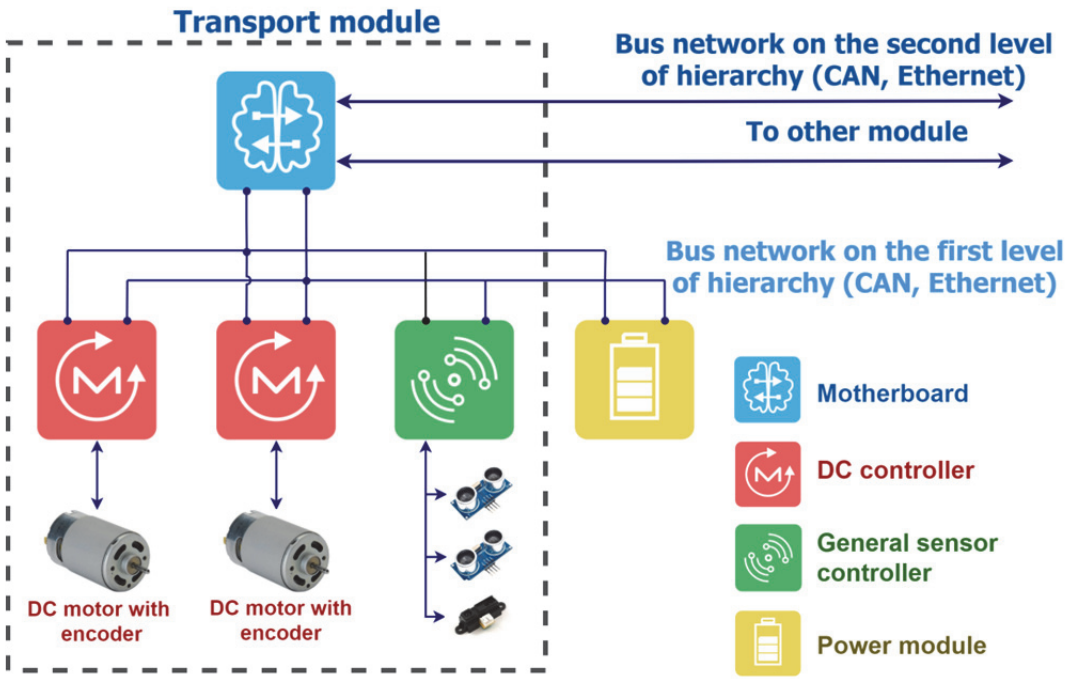


Figure 8 – The distributed architecture of the mobile robot's transport module

In this hierarchical structure the interaction of the transport module with the other full-featured nodes of the robot takes place on the bus network at the second level of the hierarchy (CAN, Ethernet, USB, etc.). It is assumed that this interaction can be carried out in a "soft" real-time (the sampling rate of the modules is not less than 10 Hz, but in general is determined by the required speed of the robot itself).

The transport module structure is based on three main types of submodules:

1) **the motherboard submodule** that responsible for control and setup of other submodules,

2) **the actuator submodule** (fig. 8 shows partial case – submodule for DC control) that control attached electromotor,

3) **the general sensor submodule** which can be connected to a variety of distance sensors, gyroscopes, accelerometers, etc.

The presence of a generalized sensor submodule in the transport module is optional; it can be used in situations where the sensor module (at the top level of the hierarchy) is not present in the mobile robot and there is a risk of collision with unexpected obstacles.

Figure 8 also shows the power module that provides power to all connected devices. The power module is not a part of the transport module and is considered as a more specific node of the system, since it communicates with the modules (the second level of the hierarchy), but provides power supply to *submodules*. So in fig. 8 the power module is schematically shown at the lower level – the submodule level.

The interaction between all submodules must be performed in a "hard" real-time (the sampling rate is not less than 100 Hz) on a bus network at the first level of the hierarchy (CAN, I2C).

Let's take a closer look at the structure of submodules.

The submodules of the modular robots transport platforms

The motherboard – is a computing unit corresponding to strategic and tactical levels of control. The main functions of submodule are determined by user which programs device based on known set of desired platform functions. In a general case the motherboard may be responsible for:

1) path planning from current to goal position assigned by a module-supervisor (top level of the hierarchy);
2) calculation of a platform linear and angular velocities according to a planned path;
3) getting data about robot real position and its comparison with goal position;
4) getting data about emergency situations and planning further actions in case of their occurrence;
5) determination of the feasibility of the task set by the module-supervisor;
6) calculation of the path taken by the robot since the beginning of the task.

Besides of mentioned functions the motherboard performs a task of a mediator (*bridge*) between developer's PC and other submodules. This is necessary for these submodules setup: setting PID controller coefficients of actuator submodule if needed, setting filtration and interpolation parameters for sensor, etc.

The submodule should include a microcontroller (for example, the STM32 family) or a single-board computer, hardware interface connectors, an electronic "harness", a controller for communication, if necessary, a power supply regulator.

The actuator submodule – is an electronic device (controller) designed to control various electric motors: brushed and brushless DC motors, stepper motors, servomotors. In the transport module the actuators actually regulate various locomotors mounted on motors shafts: wheels, legs, tracks, etc.

The main submodule functions:

1) motor's shaft speed control (if speed control),
2) motor's shaft position control (if position control),
3) actuator locomotor position determination with respect to platform coordinate system,
4) calculation of the path taken by locomotor.

Also in certain cases some additional functions may be required:

1) fixing the fact of a wheel (leg) slipping when moving;
2) determination of a force and torque vectors applied at the contact point of a link with ground.

The submodule should include the following components:

- microcontroller,
- motor driver (H-bridge for example),
- electronic "harness",
- hardware interface connectors,
- power supply regulator,
- communication controller.

According to the proposed concept actuator submodules depending on a controlled electromotor type may be of three main kinds:

1) for a brushed DC motor control (DC controller),
2) for a brushless DC motor control (BLDC controller).
3) for a stepper motor control (Stepper controller),
4) for a servomotor control (Servo controller).

Each of these types should also vary depending on the maximum permissible power of the electromotor to be connected.

It's supposed that a actuator submodule should not be mechanically connected with a certain motor which it controls because even in a one mobile robot motors with different size and design may be used.

The general sensor submodule – is an electronic device designed for attaching of different sensors, mainly range finders and inertial measuring devices (gyroscopes, accelerometers). It's supposed that submodule must be compliant with the most common typical sensors.

The main submodule functions:

1) collecting data from connected sensors,

2) sensor data processing (interpolation, filtration etc.),

3) sensor data representation in a form defined by a intermodule communication protocol used in the system.

The submodule should include the following components:

– microcontroller,
– electronic “harness”,
– hardware interface connectors,
– connectors for sensors attaching,
– power supply regulator,
– communication controller.

An important feature of the general sensor submodule is the ability to connect to it different models of sensors of the same type, for example, ultrasonic rangefinders. The principle of operation of ultrasonic rangefinders is the same in different models; differ only technical parameters that determine the further processing of sensor readings (linear interpolation coefficients, filter parameters, etc.). It is assumed that the user must configure the submodule before installation: select the model of the connected sensor, the interpolation method, filtering, etc.; this configuration should be done using the motherboard submodule that connects to the user's computer.

Figure 9 shows the computer models of the submodules considered above. Each submodule is a complete electronic device to which different controlled peripherals (motors, sensors, buttons, etc.) must be connected. All submodules are equipped with connectors for connection to the CAN network (various network topology options are possible).

Figure 9 – Computer models of submodules: DC controller, general sensor submodule and the motherboard

Most of the different versions of the transport module require the motherboard and the general sensor submodule, as well as a several actuator submodules, a number of which depends on a number of controlled links of the mobile platform.

Figure 10 shows the block diagram of the architecture of a four-wheel transport module, which consists of four actuator submodules, that control geared DC motors, the motherboard submodule and the general sensor submodule.

When the robot moves on an uneven surface, the task of following a given trajectory can be set. Then one needs to know the information about the distances traveled by each wheel - $S_1, S_2,$ etc. The motherboard can fusion data collected from actuator submodules. If additional sensors are present the actuator submodule must determine its wheel position.

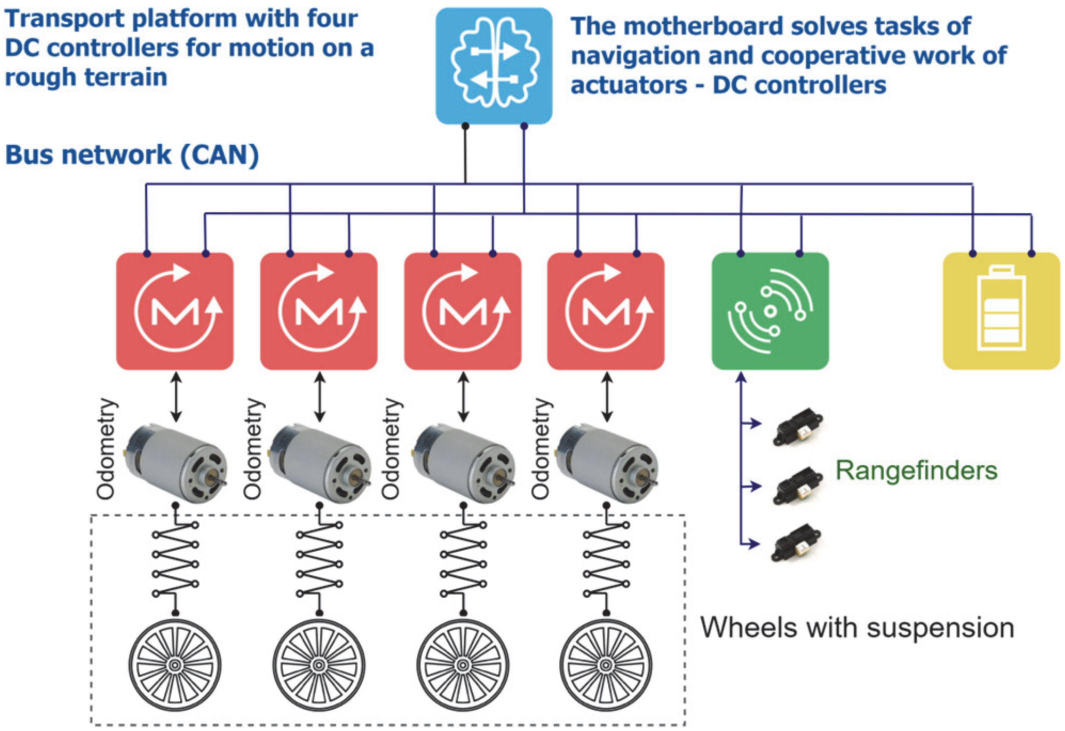


Figure 10 – The transport module architecture for motion on a rough terrain

Depending on the specific application for which the transport module is assembled, the chassis design and the suspension mechanism will vary, but the hardware of the control system (submodules) must remain unchanged. The use of submodules should ensure the reconfigurability of the transport module: instead of wheels, one can use walking mechanisms, then the composition of submodules will not change, but their number may increase and the overall algorithm may change. At the moment, the following laboratory prototypes of modules and submodules are made: DC controller, the motherboard and the power module. Figure 11 shows the photos of the manufactured devices.


Figure 11 – From left to right: DC controller (35 V and 4 A maximum), the power module (voltage: 10 – 16.8 V, current 5 A), the motherboard

Conclusion

The considered approach to the design of a full-featured transport module as a distributed system consisting of independent computing nodes (submodules) should provide a solution to the problems of navigation of a mobile robot in an environment with complex terrain. This can be achieved by separating functions between submodules belonging to different levels of the hierarchy, which is based on the classical structure of intelligent robot control systems. It is shown that the transport module can be represented by three main submodules corresponding to the levels of the hierarchical structure of the mobile robots control systems. On the laboratory model of the transport module, the structure of the control system which currently contains only two levels, was demonstrated the efficiency of the algorithm for motion along splines using potential fields. It is assumed that the expansion of the transport module architecture should ensure the movement of the mobile robot along a given trajectory on a complex terrain.

Future work

The further task of the study is to determine a sufficient set of unified submodules, their manufacture and testing of the protocol of intermodule interaction. It is necessary to conduct experiments with modules and submodules to study the possibility of using the hierarchical topology of the transport module control system in the navigation tasks of mobile robots in an environment with complex terrain.

References

1. Andreev V.P. Network–based design of heterogeneous modular mobile robotic systems / Andreev V.P., Kim V.L., Poduraev Yu.V. // Robotics and technical cybernetics. – Saint-Petersburg: Russian State Scientific Center for Robotics and Technical Cybernetics (RTC), 2016. – №3(12). – P. 23-29.
2. Andreev V.P. Hardware & software solution for rapid reconfiguration of heterogeneous robots / Andreev V.P., Kim V.L., Pletenev P.F. // Mekhatronika, Avtomatizatsiya, Upravlenie. – M.: Publishing house «New technologies», ISSN: 1684-6427. – 2018. – №6, Vol. 19. – P.387-395. DOI: 10.17587/mau.19.387-395.
3. Pletenev P.F. etc. 1/ PMMV - Protokol vzaimodeystviya v geterogennom modulnom mobilnom robote. URL: https://asmfreak.github.io/modular_robots_rfc/1/ПММВ/ (access data: 27.04.2019).
4. Andreev V.P., Kim V.L. motion control method for modular mobile robot using two-dimensional vector fields // Robotics and technical cybernetics. – Saint-Petersburg: Russian State Scientific Center for Robotics and Technical Cybernetics (RTC), ISSN 2310-5305. – 2017. – №4(17). – P. 22-27.
5. Astrom K.J., Murray R.M. Feedback systems: an introduction for scientists and engineers – New Jersey: Princeton University Press, 2008. – 396 P.
6. Andreev V.P., Pletenev P.F. Method of information interaction for distributed control systems of robots with modular architecture // Tr. SPIIRAN. ISSN 2078-9181 (print.), ISSN 2078-9599 (online). – Saint-Petersburg: SPIIRAN. – 2018. – №2(57). – P.134-160.
7. ASEBA: a modular architecture for event-based control of complex robots / S. Magnenat, P. Retornaz, M. Bonani, V. Longchamp, F. Mondada // IEEE/ASME Transactions on mechatronics. – 2011. – Vol. 16, №2. – P. 321-329.
8. Taira T. Design and implementation of reconfigurable modular humanoid robot architecture / Taira T., Kamata N., Yamasaki N. // IEEE/RSJ International Conference on Intelligent Robots and Systems. – Edmonton, 2005. – P. 3566-3571.
9. Popov Ye.P. Pismennyy G.V. Osnovy robototekhniki: vvedeniye v spetsialnost. - M .: Vysshaya shkola, 1990.– 224 P.

V.P. Andreev, P.F. Pletenev

A STUDY OF THE APPLICABILITY OF DIFFERENT NETWORKS AND TOPOLOGIES IN A MODULAR ROBOT WITH A PYRAMIDAL STRUCTURE OF CONTROL SYSTEM

MSTU "STANKIN", IINET RSUH, IL "Sensorika", Moscow, Russia
andreevvipa@yandex.ru, cpp.create@gmail.com

Abstract

The features of mobile modular robots with a pyramidal (multi-level) structure of information-measuring and control system (IMCS) from the point of view of inter-module network inter-level and intra-level interaction are considered. The modular architecture allows for rapid reconfiguration of robotic systems. The use of hierarchical topology for the construction of IMCS robot, when each module and submodule has its own IMCS with a separate computer, allows one to increase the performance of the system by distributing the computational load between the computing devices of the modules. The implementation of distributed computing in a multiprocessor system with a hierarchical topology imposes a number of restrictions on the organization of intermodule information interaction. Based on the analysis of logical intermodule connections and evaluation of possible information flows in a system with distributed computing, the requirements for both the types of networks themselves and their topology are formulated. In the context of these requirements, the existing networks and protocols of information exchange are evaluated, their brief description is given.

Keywords: modular robot, mobile robot, network technologies, informational interaction, reconfiguration.

Acknowledgments

Research is supported by the Russian Foundation for Basic Research: Grant 19-07-00892a.

Introduction

The expansion of the use of mobile robots (MR), which is currently observed, leads to an increase in the algorithmic complexity of the processes of controlling the functioning of robotic systems (RTS). This problem can be solved using high performance computers in the information-measuring and control system (IMCS) of RTS, which is known to be not always possible. Another way is to create a RTS with a modular architecture [1], in which each module is equipped with its own IMCS, built on the basis of a low-performance systems – *embedded computing devices* (microcontrollers and microcomputers). Then the implementation of the computational process of system-wide control and decision-making RTS is solved on the basis of distributed computing methods performed on a set of processors.

The second problem is the need for reconfigurable RTS. This is due to the obvious need to use MR where human stay is fraught with danger to their health or life, for example, for work in extreme conditions of the Arctic and Antarctic, space and in solving the problems of the Ministry of emergency situations. In these conditions assembly of robots of necessary functionality shall be carried out on the place of the carried-out works and during them. Therefore, the assembly process should be simplified as much as possible, and the reconfiguration or scaling of the RTS should be performed in the "plug and play" mode, which is associated with the need to organize the automatic configuration of the IMCS. As shown in [2], and this problem is solved by creating robots with modular architecture of IMCS.

On the basis of the principle of full functionality [3] it is proposed to consider the structure of the MR control system (in the minimum version) as a synergetic union of full-featured modules (Fig.1): transport module – TM, power module – PM, sensor system consisting of short-range module – SRSM and long-range module – LRSM, active interaction modules with the external environment – EIM (manipulators, grippers, etc.) and intelligent control module – ICM. In this structure, each module is responsible for only one function of the robotic system. The principle of full functionality of modules is formulated as follows: *each robot module should be able to perform its goal function in any convenient way, using only its own means to execute commands from an external control system.*

Research carried out within the framework of RFBR grant 16-07-00811a "Development of functional-modular principle of construction of hardware and software of intelligent mobile robotic systems" [3] showed that the two-level hierarchy of the modular architecture (see Fig. 1) is unable to provide real-time control process on embedded computing devices, because functional modules IMCS algorithms are quite complex. The amount of information processed in modules is still large, and the computing power of embedded systems is not always sufficient to provide real-time performance. Especially expensive is the process of processing and integration of sensory information and decision-making at the level of intelligent system-wide control module. As the complexity of the tasks to be solved by MR both in stand-alone and in supervisory mode grow, so the algorithms of the IIUS modules will become more complex, and the use of microcomputers with greater performance will still lead to limited functionality of the module.

A number of experiments conducted with the transport module showed the possibility of further development of the MR modular architecture in division the full-featured modules into submodules and creating a multi-level hierarchical (pyramidal) network topology of IMCS, which ensures the use of embedded computing systems at each level of the hierarchy [5].

1. Multi-level subordination modules and submodules

Within the framework of the given concept, it is supposed to introduce a multi-level subordination of modules and submodules so that the resulting system is (a) implemented on embedded devices and (b) the computational load overall robot control is distributed among a variety of different computers. Let us consider this approach on the examples of the main functional modules of MR (see Fig. 1): function blocks of the transport module [6], the short-range and long-range sensor module [7], and the power module. The proposed schemes of the modules are shown in fig. 2, 3 and 4. At least 3 levels of interaction and subordination can be distinguished in the modules.

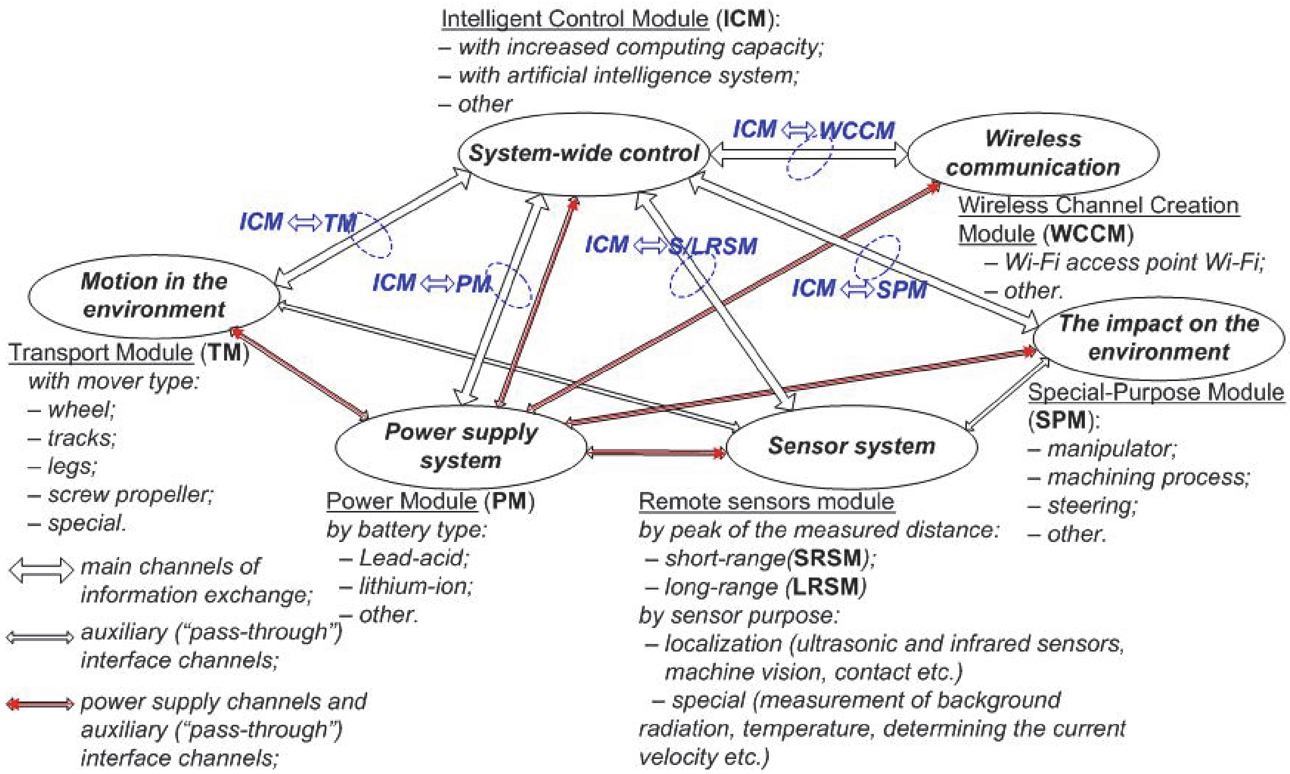


Figure 1 – Functional-modular architecture of information-measuring and control system of the mobile robot

– 0^{th} (zeroth) level of interaction (and corresponding “networks” of zeroth level, see fig. 2-4):

– primary analysis of sensors data, filtering and conversion to structural representation of information;

– creating closed-loop drives and other control systems.

– 1^{st} level of interaction (and corresponding First Level Network, see fig. 2,3) – level of full-functional submodules. The following is implemented here:

– secondary data analysis, aggregation of readings from several submodules;

– development and distribution of tactical tasks for this module as a whole.

– 2^{nd} level of interaction (and corresponding Second Level Network, see. fig. 2-4) – level of full-functional modules:

– tertiary data analysis, aggregation of data from same and different types of modules;

– development and distribution of strategic tasks for the robot as a whole.

It is assumed that the interaction of modules with an external supervisor will also occur at the 2 level of interaction.

Another, specialized method of interaction is also possible – through the power line. Since it is assumed (see [5]) that all modules will be powered from **one shared power line**, it is possible to create a communication system between all modules and submodules included in the robot. However, it should be borne in mind that such a communication channel will have low bandwidth.

As one can see in figures 2-4, each of the modules and submodules is, in fact, a **specialized border router**, which converts network traffic from a higher level to network traffic from a lower level, except when all modules and submodules are connected to one common bus. Also, figure 4 shows the case when the network of the 2nd level interacts with the networks of the 0-th level, i.e. the 1st level "collapses" inside one submodule, and the submodule of the generalized control of the power supply performs the tasks of two levels at once, which will require separation in time.

This concept is complicated by the multiplicity of possible technical solutions. In principle, there are no universal solutions for zeroth level networks – each actuator or sensor has its own unique ways of connection and control. Many sensors and actuators can be grouped together, and the ways they interact at the 0-th level of the network are the details of the implementation of each specific module and should not be important in the organization of intermodule interaction.

First and Second Level Networks are also a part of the implementation of each module, but they generally need to interact with each other, as well as with all possible modules according to general principles, and obey the same set of commands. This is only possible if these networks are standardized.

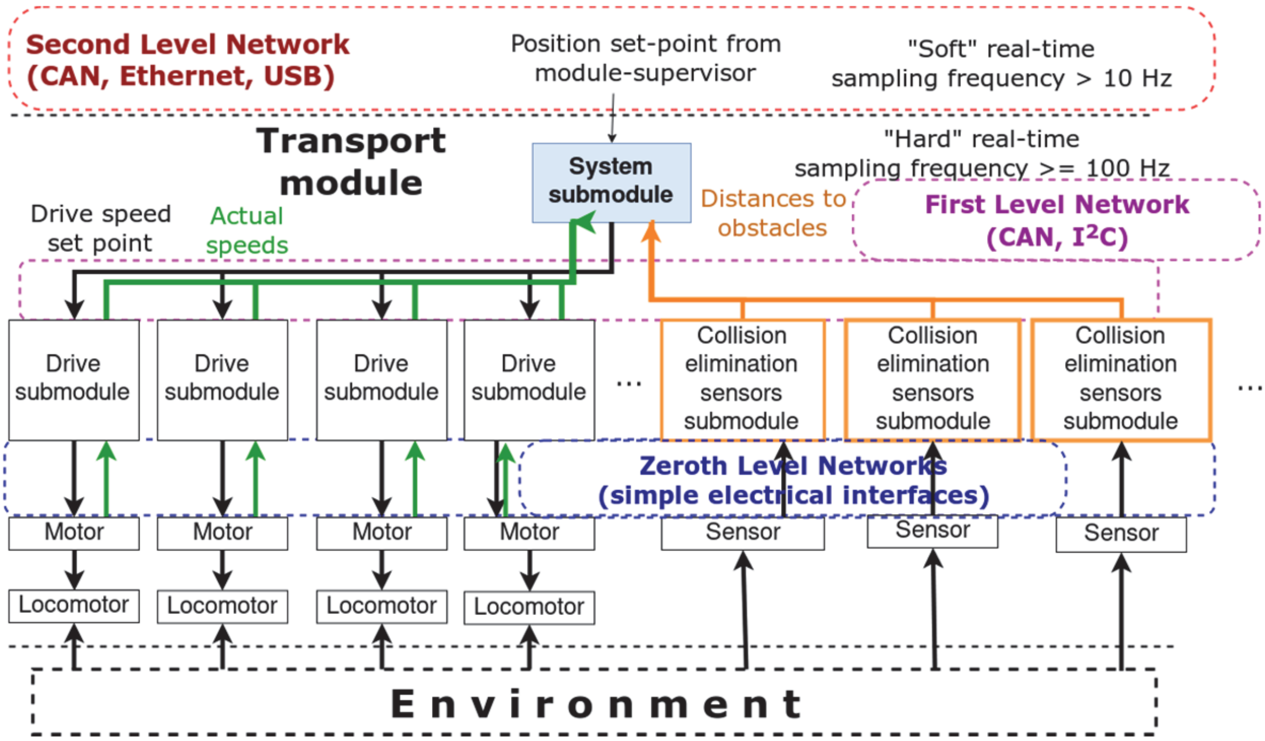


Figure 2 – Transport module's structure [6]

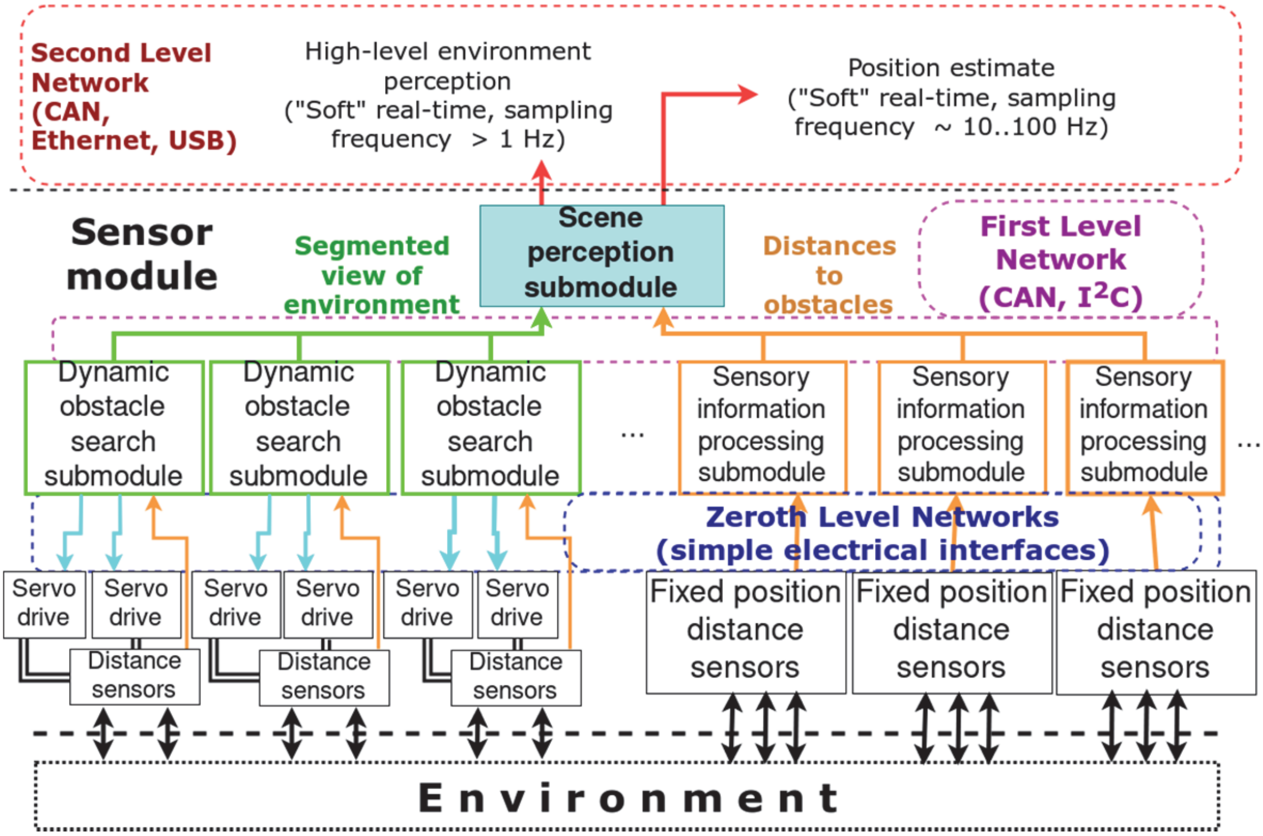


Figure 3 – Sensor module's structure

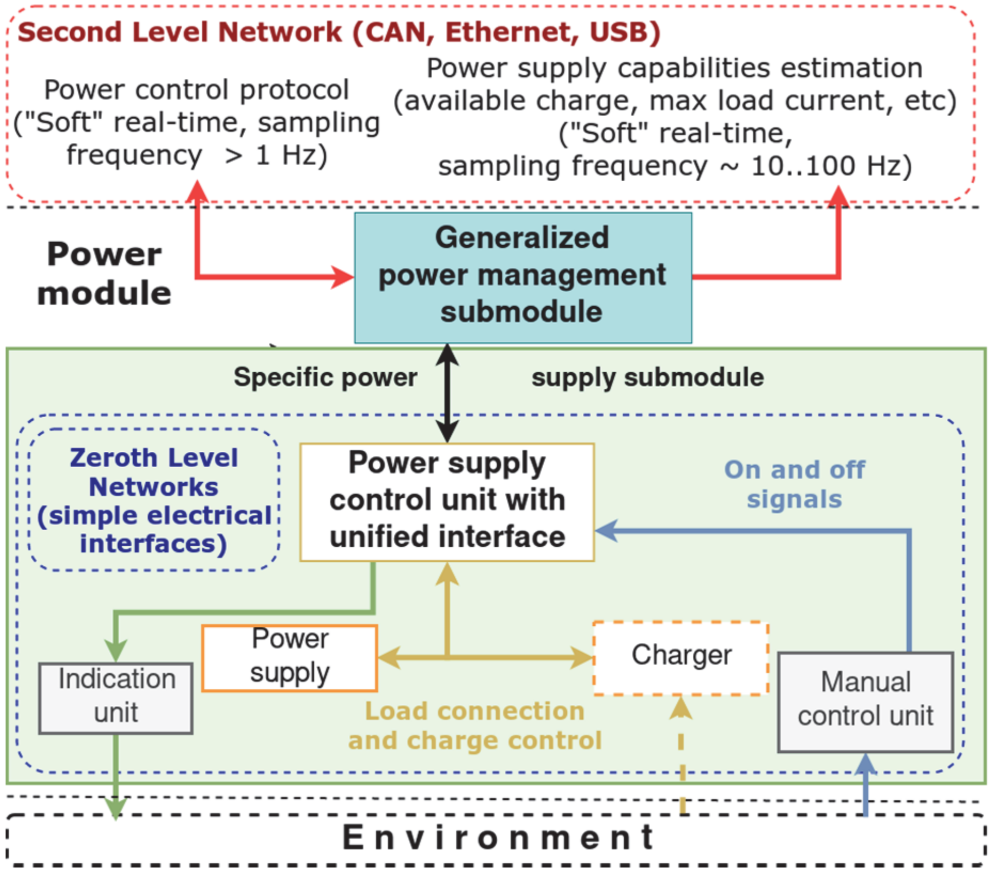


Figure 4 – Power module's structure

2. The use of frameworks and standard libraries

Existing libraries and frameworks offered by different researchers in [8-11] do not allow dynamic reconfiguration. Of particular interest is the approach implemented in the framework of EmsBoT [12], which proposes the creation of an abstraction layer over the real-time OS. This layer would allow the same interaction between different networks, providing the same software interface for communication via CAN, Ethernet, USB and other networks, using the same software code and a set of tools for creating agents. Unfortunately, the implementation of this approach presented in the article does not imply dynamic reconfiguration.

From this point of view, we consider the previous work – specification 1/PIMI [13] (Protocol of intermodule interaction) and the method of information interaction for distributed control systems in robots with modular architecture [14]. In these works, the structure of the modular robot IMCS is used, shown in Fig. 1. The following was assumed:

– All modules of one type within one robot are available in a single instance and fully perform their own functional purpose.

– All properties of the module and its functional characteristics are given in its meta-information.

The current approach – a modular robot with a hierarchical IMCS topology – is that any type of module can be represented in multiple instances, each of which can consist of several functional submodules. In this case, each of the submodules can be further divided into smaller, possibly incomplete functionally blocks – specific sensors and actuators (drives).

Due to the fact that changing the structure of a modular robot changes different characteristics (mass, moment of inertia, etc.), it is necessary to rebuild and adjust the control system of modules and submodules of robots. Therefore, for the correct operation of the mechanisms for automatic reconfiguration of modules and submodules at all levels, it is necessary to automatically update the knowledge of modules about the relative location of modules and submodules, their size, weight, etc. Therefore, First and Second Level Networks need to provide information about the topology of the modules – information on where exactly a module is attached to or detached from.

The previous specification [13] referred only to Ethernet. This network has shown its not the highest efficiency in terms of module size and complexity of implementation of all levels of abstraction for protocols using TCP/IP stack. This means that further development of the specification needs to be carried out taking into account old and new requirements and limitations. Since the existing communication networks, which are used in various modular robots, differ in physical methods of data transmission and protocols, it is necessary to use the formalism of the OSI model [15] to compare the protocol stacks of communication networks.

1. A communication network **must** provide data transmission at a speed sufficient for the modules and, where necessary, ensure deterministic delivery of messages.

2. A communication network **must not** have an explicit master node for the data transmission, i.e., a communication network **must** ensure the creation of horizontal links.

3. A communication network **must be** able to provide the node with information about the topology of the connection of other nodes in the network.

4. If one selects multiple networks for one layer, there **must be** one defined way to convert message of one network type to message of another network type.

5. Standard connectors for network communications, if they are specified by the physical layer of standard of a network, **must be** as small as possible.

6. As many levels of communication network as possible **must be** implemented at the hardware level in embedded systems.

7. For most embedded devices, there **must be** a library (or multiple libraries) to work with a communication network.

8. The physical and/or link layer of the communication network **must be** noise-proof and/or have error detection and correction mechanisms for transmitting information at distances specific to the modules.

It should be noted that requirements 2 and 3, in fact, contradict each other. Therefore, any choice of network will be a compromise between the ability to create horizontal links and obtain information about the topology and connection points of the nodes. This list of requirements may not be complete, and some items may not be required. The solution of this problem requires further both theoretical and practical research.

3. Existing communication networks

At the moment, many robots with modular architecture have been created: homogeneous [16-19] and heterogeneous [20-23]. They use many different types of networks. Each type of network can be implemented in different topologies and include only a part of the standard model levels. To assess the applicability of each network, we estimate the volume of messages between the modules and the frequency of their transmission for the worst case. A special mathematical representation is often used for such estimates – network calculus, which is based on min-plus algebra.

In network calculus, all network nodes are represented as black boxes with multiple inputs and outputs. Each input is characterized by the so-called arrival function (arrival curve), and each output – departure function (departure curve). Both of these functions characterize the total amount of information received or sent by a particular host. Some of the networks are more amenable to such formalism than others. For example, when using networks with priority of delivery such as CAN, must make extra efforts to analyze traffic due to the fact that the message with higher priority to interrupt the transmission of a message with lower priority [24]. To evaluate the performance of the network as a whole, you must first evaluate the parameters of the arrival curves for each transmitter in each module and submodule. Let us use the Heaviside function (1) to set the step curve (2), which is the arrival curve for the submodule transmitter (PM), where f_{sm} , Hz, is the frequency, a V_{sm} , bytes, is the payload volume of the transmitted messages.

$$h(t) = \begin{cases} 0, x < 0 \\ 1, x \geq 0 \end{cases}. \tag{1}$$

$$A_{sm}(t) = \sum_{n=0}^{\infty} V_{sm} \cdot h(\mathrm{t} - \frac{n}{f_{sm}}). \tag{2}$$

Let's evaluate these parameters for various modules and submodules:

1. For transport module (see fig. 2):

a) For drive submodules: $f_{sm} >= 100\,Hz$, new speed set point $V_{sm} = 4\,bytes$, передача readings on the current speed and position of the output shaft $V_{sm} = 8\,bytes$. There can be several such submodules.

b) For collision elimination sensors submodules: $f_{sm} >= 100\,Hz$, sensors can be two kinds: (1) bumper with touch sensors $V_{sm} = 1\,byte$, (2) non-contact distance sensor $V_{sm} = 4\,byte$. There can be several such submodules.

c) For the motion control submodule: $f_{sm} >= 100\ Hz$, the submodule is given the desired linear and angular velocities $V_{sm} = 8\,bytes$, the actual position of the module relative to a certain coordinate system $V_{sm} = 4*6 = 24\,bytes$.

2. For sensor module (see fig. 3):

a) For a dynamic search obstacles submodule: $f_{sm} \sim 1\,Hz$, this submodule can transmit the segmented view of environment in either of n line segments in the plane $V_{sm} = n \cdot 16\ bytes$ or n line segments in space $V_{sm} = n \cdot 16 \cdot 6 = n \cdot 48\,bytes$. There can be several such submodules.

b) For sensory information processing submodule: $f_{sm} \sim 1\,Hz$, this sub-module processes signals from various sensors and generates the average estimate of the distance $V_{sm} = 8\ bytes$.

3. For power module (see fig. 4):

a) For the generalized power management submodule: $f_{sm} \sim 1\,Hz$, this submodule periodically publishes information about the current state of the battery, the maximum load current, etc $V_{sm} = 1..4 \cdot 1..4\ bytes$.

In the above data, all frequency parameters f_{sm} depend on the speed of the robot. For other modules, the pattern will be approximately the same. Actuating modules, such as transport module, manipulator module or process module (capture module), have high frequencies of interaction inside the module with small sizes of transmitted messages. Sensor modules or data processor modules, on the other hand, have a much lower frequency of message transmission, which is compensated by the large volume of these messages. A separate case is the power module – on the one hand, this module should not very often publish sensor readings, and on the other hand – should quickly convey a message about the need to start a safe shutdown to all modules.

Thus, it is important that messages in the communication network are transmitted with the least overhead (overhead), and that all packets have a roughly equal chance of being transmitted despite their size, i.e. that long packets do not occupy the common bus for too long.

It can also be noted that different topologies may be preferred for different module types. Topology "bus" is preferable to the modules of the actuators, especially if they are the creation of horizontal linkages. The star topology is preferable in the case of sensor modules, as it allows data to be transmitted directly to the supervisor. The definition of requirements for the use of certain topologies for different types of networks will be formulated in further research.

4. Conclusion

In the future, it is necessary to perform a detailed analysis (within the indicated problems) of such networks as I2C, SPI, RS-232, RS-485, CAN, USB, Ethernet, EtherCAT. It is necessary to create a new, universal specification for the method of intermodule interaction, which could work with any coding method and in any communication network.

References

1. Andreev V.P., Poduraev Yu.V. Functional-modular design of heterogeneous mobile robotic systems // Extreme robotics. Proceedings of the International Scientific and Technological Conference. 2016. – Sankt-Peterburg: OOO «AP4Print». 2016. pp.44-49.
2. Andreev V. The principle of full functionality – the basis for rapid reconfiguration in heterogeneous modular mobile robots / Andreev V., Kim V., Pletenev P. // Annals of DAAAM and Proceedings of the 28th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2017; Zadar; Croatia,B. Katalinic (Ed.), Published by DAAAM International, ISBN 9781510853270, ISSN 1726-9679, Vienna, Austria. Curran Associates, Inc. (Feb 2018). – pp.0023-0028.
3. Andreev V.P., Kim V.L. Development of functional unitsof a heterogeneous modular mobile robot // Extreme robotics. Proceedings of the International Scientific and Technological Conference. 2016. – Sankt-Peterburg: OOO «A4Print». 2016. pp.359-369.
4. Andreev V.P. Hardware & Software Solution for Rapid Reconfiguration of Heterogeneous Robots. / Andreev V.P., Kim V.L., Pletenev P.F. // *Mekhatronika, Avtomatizatsiya, Upravlenie.* – M.: Publishing house «New technologies», ISSN: 1684-6427. – 2018. – №6, Vol. 19. – pp.387-395.
5. Andreev V.P. The concept of using the theory of multi-agent systems to design control systems for mobile robots with modular architecture. (proceedings of ER-2019).
6. Andreev V.P., Kim V.L. Modular architecture of a mobile robot transport platform for a motion task on a rough terrain. (proceedings of ER-2019).
7. Andreev V.P., Tarasova V.E. Determination of the Form of Obstacles by a Mobile Robot Using Scanning Angular Movements of Ultrasonic Sensor. Mekhatronika, Avtomatizatsiya, Upravlenie. 2017;18(11):759-763. (In Russ.).
8. The shift in the robotics paradigm — the Hardware Robot Operating System (H-ROS); an infrastructure to create interoperable robot components / V. Mayoral, A.Hernandez, R.Kojcev, I. Muguruza et al. // NASA/ESA Conference on Adaptive Hardware and Systems (AHS), Pasadena, CA. – 2017. – pp.229-236.
9. AMiRo: a modular & customizable open-source mini robot platform / S. Herbrechtsmeier, T. Korthals, T. Schopping, U. Ruckert // 20th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia. – 2016. – pp.687-692.
10. R2P: An open source hardware and software modular approach to robot prototyping / A. Bonarini, M. Matteucci, M. Migliavacca, D. Rizzi // Robotics and Autonomous Systems. – 2014. – No.62. – pp.1073-1084.
11. Distributed and modular CAN-based architecture for hardware control and sensor data integration / D. P. Losada, J. L. Fernández, E. Paz, Rafael Sanz // Sensors. – 2017. – No.17. – pp.1013-1030.
12. EmSBoT: A lightweight modular software framework for networked robotic systems / L. Peng, F. Guan, L. Perneel, H. Fayyad-Kazan and M. Timmerman // 2016 3rd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA), Beirut, 2016, pp.216-221. doi: 10.1109/ACTEA.2016.7560142.
13. Pletenev P. F. 1/PIMI – Protocol of intermodular interaction [1/ PMMV – Protokol vzaimodejstvija v geterogennom modul'nom mobil'nom robote] Available at: https://asmfreak.github.io/modular_robots_rfc/1/ПММВ/ (accessed: 20.01.2017). (In Russ.).

14. Andreev V.P., Pletenev P.F. Method of information interaction for distributed control systems of robots with modular architecture. // SPIIRAS Proceedings. 2018. № 57 (2). P.134-160.
15. GOST R ISO/MEK 7498-1-99. — « Information technology. Open systems interconnection. Basic reference model. Part 1. The basic model». — OKS: 35.100.70. — In use since 01.01.2000. — 62p.
16. M-TRAN: self-reconfigurable modular robotic system / S.Murata, E.Yoshida, A.Kamimura, H.Kurokawa, K.Tomita & S.Kokaji // IEEE/ASME Transactions on Mechatronics.2002. no. 7(4). pp.432-441.
17. Design of the ATRON lattice-based self-reconfigurable robot / E.H.Østergaard, K.Kassow, R.Beck & H.H.Lund // Autonomous Robots. 2006. no. 21(2). pp.165-183.
18. Design of Transmote: a Modular Self-Reconfigurable Robot with Versatile Transformation Capabilities / Guifang Qiao, Guangming Song, Jun Zhang, Hongtao Sun, Weiguo Wang & Aiguo Song // Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics. 2012. pp.1331-1336.
19. Concept of cellular robotic system (CEBOT) and basic strategies for its realization / Toshio Fukuda, Tsuyoshi Ueyama, Yoshio Kawauchi, Fumihito Arai // Computers Elect Engng. 1987 vol.18. no.1. pp.11-39.
20. Baca J. A heterogeneous modular robotic design for fast response to a diversity of tasks / Baca J., Ferre M., Aracil R. // Robotics and Autonomous Systems. 2012. vol. 60. no. 4. pp. 522-531.
21. On sub-modularization and morphological heterogeneity in modular robotics / A.H.Lyder, K.Stoy, R. F.Mendoza-Garcia, J.C.Larsen & P.Hermansen // Intelligent Autonomous Systems of Advances in Intelligent Systems and Computing. Springer Berlin Heidelberg. 2013. vol.193. no. 12. pp. 649-661.
22. Hancher M.D., Hornby G.S. A modular robotic system with applications to space exploration // 2nd IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT'06). Pasadena, CA: Publisher «IEEE». 2006. pp.132-140.
23. Reusable Electronics and Adaptable Communication as Implemented in the Odin Modular Robot / Ricardo Franco Mendoza Garcia, Andreas Lyder, David Johan Christensen and Kasper Stoy // IEEE International Conference on Robotics and Automation. 2009. pp.1152-1158.
24. "Delay Bounds for CAN Communication in Automotive Applications," / U. Klehmet, T. Herpel, K. Hielscher and R. German // 14th GI/ITG Conference - Measurement, Modelling and Evalutation of Computer and Communication Systems, Dortmund, Germany, 2008, pp. 1-15.

E.A. Abrosimov[1,2], V.A. Dyacheko[2], A. V. Bakhshiev[1,2], E.K. Ignatiadi[1], A.A. Shavlikov[1]

TECHNOLOGIES OF ARTIFICIAL INTELLIGENCE IN THE PROBLEM OF ANALYSIS OF ROAD SITUATION BY AUTONOMOUS VEHICLE

[1] *Russian State Scientific Center for Robotics and Technical Cybernetics, St. Petersburg, Russia*
[2] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*
e.abrosimov@rtc.ru

Abstract

Nowadays the automotive industry is experiencing a revolution, the emergence of unmanned vehicles. The main stages solved by the control system of an unmanned vehicle are the recognition of the road environment, route planning and vehicle movement control. It is necessary to solve the problem of analyzing the traffic situation for the successful development of unmanned vehicles. The main challenge of this task is caused by the fact that the vehicle needs to function in a dynamic, non-deterministic environment, affected by external factors (weather conditions, other road users, pedestrians, etc.).

Thus, the analysis of road situation is an intellectual task and requires the application of artificial intelligence methods for its solution. However, the solution of some problems of image analysis can be carried out using certain algorithmic methods of image analysis, without using of artificial intelligence technologies.

Artificial intelligence methods as fuzzy systems, neural networks and hybrid systems, as well as some algorithmic methods are considered in this work. It is indicated what methods should be used in certain tasks of the road situation analysis.

Keywords: autonomous driving, fuzzy systems, neural networks, hybrid systems, algorithmic methods of image analysis.

Acknowledgments

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 “Development and study of new architectures of reconfigurable growing neural networks, methods and algorithms for their learning”.

1. Introduction

A modern unmanned vehicle is a mobile robot that operates in the same environment as a human. This environment cannot be deterministic by definition. Consequently, a completely different approach to navigation is needed for the further development of the unmanned vehicles industry It is necessary to educate an unmanned vehicle to move in a dynamic, changeable environment. Such navigation cannot be based only on statically described classical maps used, for example, in car navigator. Unmanned vehicle must understand what is happening around it at the moment. This can be achieved only by combining various methods of artificial intelligence and algorithmic methods of image analysis.

2. Relevance

The unmanned car industry has evolved rapidly over the past few years. The following Table 1 presents data showing the average number of kilometers traveled in unmanned mode prior to driver intervention [1].

Table 1. Comparative characteristics of the leading companies in the field of unmanned driving by the number of kilometers traveled without driver intervention

Company	Kilometers before human interrupting, km			
	2015	2016	2017	2018
Waymo	2 002	8252	9 005	17 951
GM Cruise	–	82	2018	8 376
Zoox	–	–	455	3 094
Nissan	22	227	335	337
Baidu	–	–	67	331

The leaders are Waymo and General Motors with the Cruise project. Cars of these companies have sensors that create detailed digital maps. Waymo's has with five lidars, four radars and an omnidirectional camera. Cruise has five lidars, 14 cameras and 21 radars.

3. Artificial intelligence technologies

Creating accurate digital maps is a complex task, since processing large data arrays requires large computational resources. In addition, such maps should be constantly updated in real time in order to be always relevant for navigation of unmanned vehicles. Therefore, it is necessary to train unmanned vehicles to move in natural environment, which cannot be fully determined. To achieve this goal, various technologies of artificial intelligence are used (see Fig. 1).

Fuzzy, neural and hybrid systems are most interesting for analyzing road conditions. Reinforcement learning can be used to improve systems based on the listed technologies.

In addition, algorithmic approaches for image analysis are also used. Geometrical clustering, analysis of disparity maps and image segmentation on stixels are related to these methods. Below are the main areas of application of these methods in the task of analyzing the road conditions.

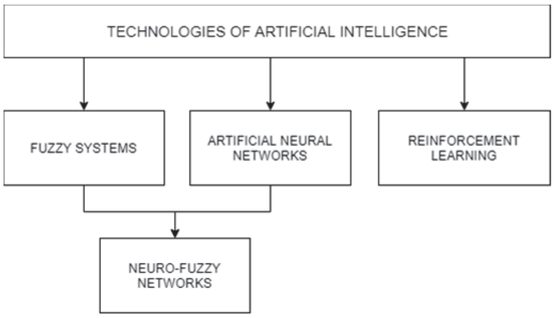


Figure 1 – Artificial intelligence technologies classification

3.1. Fuzzy systems applying

Fuzzy logic is applied with incomplete knowledge of the system [2]. Mostly fuzzy systems are used to create hierarchical vehicle control systems. Examples of such systems are given in [3, 4, 5]. However, in [6], an example of a fuzzy logic controller with a PD-like behavior for tracking the line of motion is presented. It has two entrances and one exit. The input and output variable terms are defined by triangular membership functions (see Fig. 2). The first input variable is the difference between the reference line and the measured line position in pixels relative to the center of the image (see Fig. 2a). The second input variable is a derivative of this error (see Fig. 2b). The output of the controller is the absolute steering angle in degrees to correct this error (see Fig. 2c).

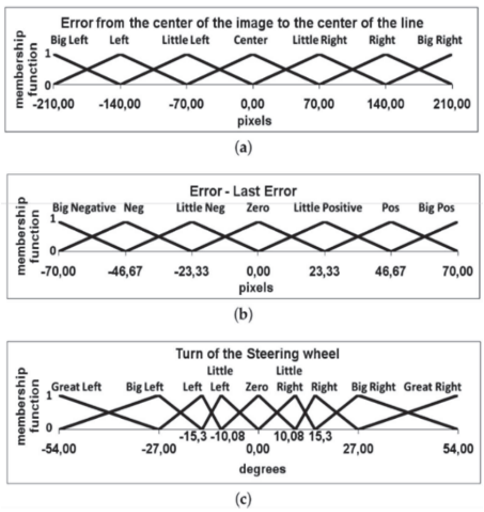


Figure 2 – Definition of two inputs and outputs of a fuzzy controller: (a) – the first input variable of a fuzzy controller: the difference between the reference line and the measured position of the line relative to the image center in pixels, (b) – the second input variable of the fuzzy controller: the difference between the last error and actual in pixels, (c) – fuzzy controller output variable: steering wheel angle, in degrees

In the task of analyzing the road conditions, fuzzy systems can be used to create fuzzy ontologies — descriptions of objects and their relationships in the environment [7]. Fuzzy ontologies are one of the varieties of semantic maps. Such maps not only describe the environment comprehensively, but also make it possible to predict the behavior of objects using knowledge and a huge data set. An example of such a map built using a Waymo's car (see Fig. 3).

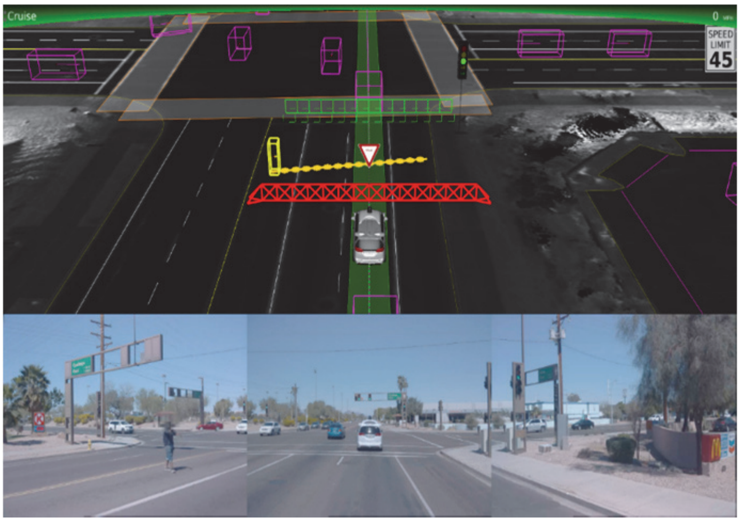


Figure 3 – Semantic map example

3.2. Neural networks applying

The last few years there has been significant progress in the field of neural network traffic analysis methods. This is due to the appearance on the market of relatively inexpensive and high-performance computers (GPU), the development of effective neural network architectures and approaches to learning. Neural network methods allow to solve problems of classification and detection of objects, image segmentation. This solves a wide range of road analysis problems.

Convolutional neural networks are used for image recognition. The input layer accepts not a vector of values, but a matrix (which is the image in digital format, with one channel in the case of a black and white image and with three in the case of color image). Convolution networks perform on the image a series of convolutions with filters (which are also matrices, usually 1x1, 3x3, 5x5 and 7x7, with an arbitrary number of channels), the weights of which are tuned during the training process. After convolutional layers, there may be several fully connected ones for displaying the probability that the result belongs to a particular class in classification task. In detection task, often fully connected layers are absent, only convolutional layers exist. The result of such networks is selected areas of the image containing objects of interest. All modern architectures are deep, i.e. contain a few dozens of layers.

The main advantage of neural networks is universality (the same architecture can be used to detect different classes of objects through training on appropriate data samples). However, due to the complexity of the neural network and the large number of tunable parameters, it is difficult to analyze its work and learning process. This leads to the fact that learning does not always give the expected result. Currently, new approaches to learning deep networks (deep learning), their analysis and visualization are being actively developed.

In [8], an empirical assessment of existing deep learning methods was carried out in application to autonomous driving tasks. The authors implemented a system for detecting road markings and vehicles by means of technical vision and deep learning. The result of the system is shown in Fig. 4 [8].

A set of labeled video data containing marked markings and vehicles in various weather conditions are used in this work. As an architecture, a straightforward implementation of Overfeat [9] was chosen. It’s a scalable architecture that simulates a detector with a sliding window in one direct pass through the network by effectively reusing convolutional results on each layer. As a result, the use of a neural network has improved the detection accuracy of vehicles at a distance of about 100 m from 10% using only radar to 80% using the developed system [8].

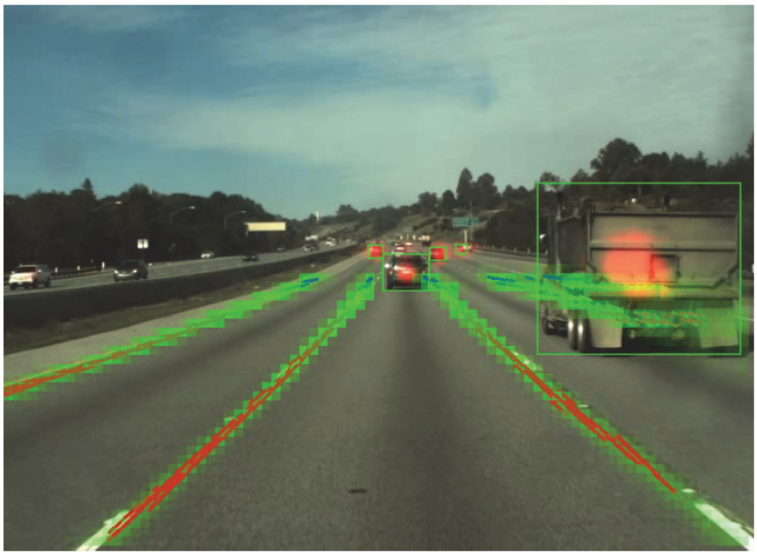


Figure 4 – Output of markup detection system and vehicles
based on neural network methods

Neural networks for detecting objects are used to detect road infrastructure elements and obstacles. As a rule, they are able to work with dozens of classes. The most modern representatives are Faster R-CNN, SSD, YOLOv2. However, for their successful application it is necessary to create a database of images and videos for training and testing, as close as possible to the actual conditions of use of the system.

Another area of application for neural networks is the segmentation of an image or a cloud of points obtained from stereo cameras or lidars. Segmentation allows not only to divide objects into classes, but also to determine the areas of patency for the vehicle. In [10], the SegNet network architecture is used to segment the image into a rather large set of classes (see Fig. 5).

Figure 5 – The result of the SegNet neural network for image segmentation

Simple methods of localization can be applied to classified areas to select individual objects. However, neural networks for image segmentation are still rather slow, making it difficult to use them in real time, especially when driving at high speed.

3.3. Hybrid systems applying

A hybrid network is a system in which conclusions are drawn on the basis of fuzzy logic, but the corresponding membership functions are adjusted using neural network learning algorithms (for example, the backpropagation algorithm). These systems not only use a prior information, but can acquire new knowledge and are logically transparent to the user [11]. In other words, hybrid systems combine the advantages of neural networks and classical fuzzy systems. Unlike neural networks, hybrid systems are characterized by a clear view of the knowledge contained in the fuzzy rules. The parameters of the membership functions of judgments and conclusions of fuzzy rules are most often subjected to training. The listed advantages of hybrid systems caused their wide use for solving problems of modeling, approximation and classification.

The most common integrated model is the ANFIS (Adaptive-Network- Fuzzy Inference System). ANFIS Implements the Takagi-Sugeno-Kang fuzzy inference system and is a five-layer forward propagation neural network shown in Fig. 6.

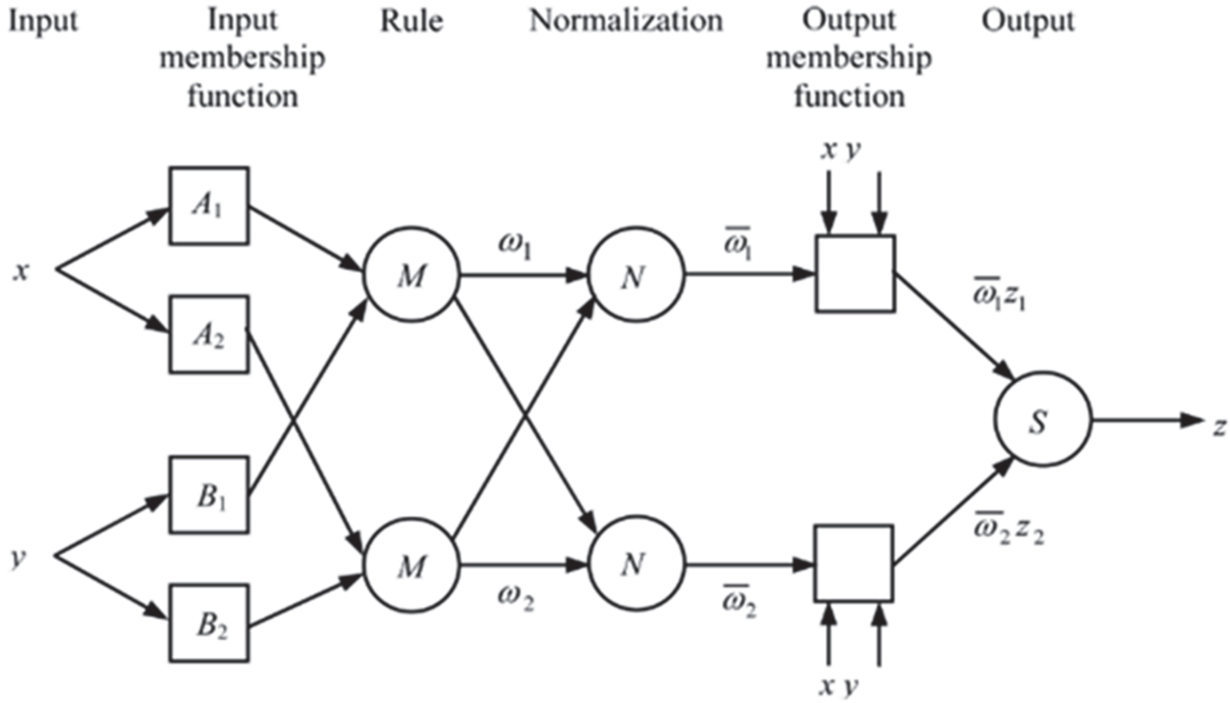


Figure 6 – ANFIS structure

Hybrid network along with fuzzy systems used in control tasks [11, 12]. They are mainly used to improve the characteristics of fuzzy controllers. This is due to the fact that systems based only on fuzzy logic are not capable of self-learning in the process. Hybrid systems, while maintaining the transparency of the conclusion and the simplicity of the formation of the rule base, allow the system to be trained on the knowledge of experts. In the road analysis system, hybrid networks can be used to predict the movement of objects [13]. Such information will be useful in compiling a semantic roadmap.

4. Algorithmic methods of image analysis

The algorithmic methods of image analysis include: geometric clustering, analysis of disparity maps when working with stereo systems, pixel segmentation, and others. Such methods are used for relatively simple tasks that do not require the creation of a large database of objects. Often, these methods work faster than neural network analogues in the same applications.

4.1. Geometric clustering

An algorithm that determines the criterion of geometric clustering of points on obstacles, based on their location in space and on the physical characteristics of the controlled vehicle, is given in [14]. An example of the result of this algorithm is shown in Fig. 7.

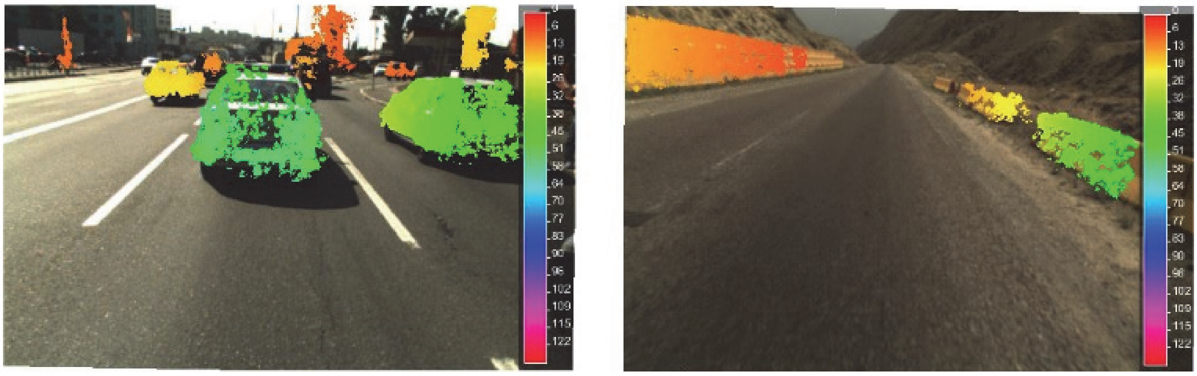


Figure 7 – Geometric clustering of obstacles

4.2. Disparity maps analysis

Stereo systems are used in the task of road segmentation on the basis of v- and u-disparity. According to [15], the V-disparity reflects a discrepancy histogram for each row of the image (coordinate v), and U-disparity does the same, but for each column (coordinate u). In other words, U-disparity is constructed by summing the pixels of each column with the same coordinates (u, d), and V-disparity by adding the pixels of each row with the same coordinates (v, d).

Information on the environment can be obtained from these disparity maps (u-disparity map and v-disparity map). For example, on the u-disparity map, perpendicular obstacles in front of the vehicle appear as horizontal lines, the intensity of which pixels reflect the height of these obstacles. In the case of v-disparity, perpendicular obstacles appear as vertical lines with pixel intensity proportional to the width of the obstacles. Another interesting feature is that the road profile in front of the vehicle appears as an inclined line [16] (see Fig. 8).

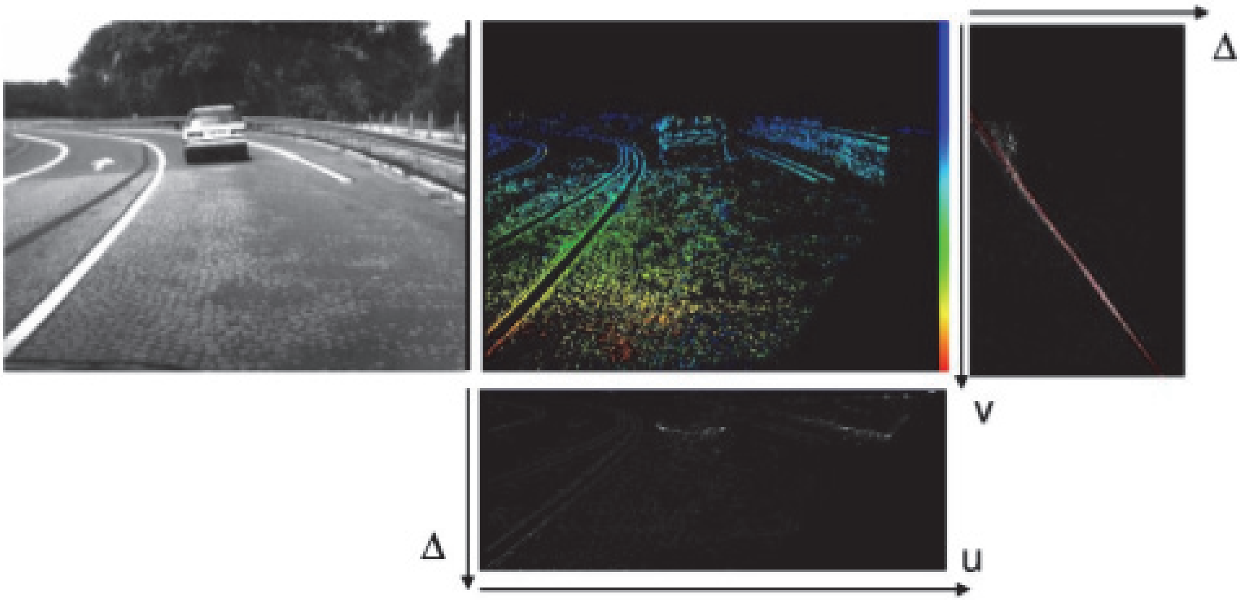


Figure 8 – Application of the method of mapping u- and v-disparity for the segmentation of the road

In addition to the segmentation task, the method of building disparity maps is also used for the tasks of detecting and classifying objects. A system for detecting obstacles in front of a vehicle and classifying them as pedestrians and other objects is described in [17]. To obtain a dense map, the SORT-SGM stereo-design method is used. The road plane is calculated using the v-disparity map. The obstacles are then determined by analyzing the u-disparity map. Objects are then classified. The result of this algorithm is shown in Fig. 9 [17].

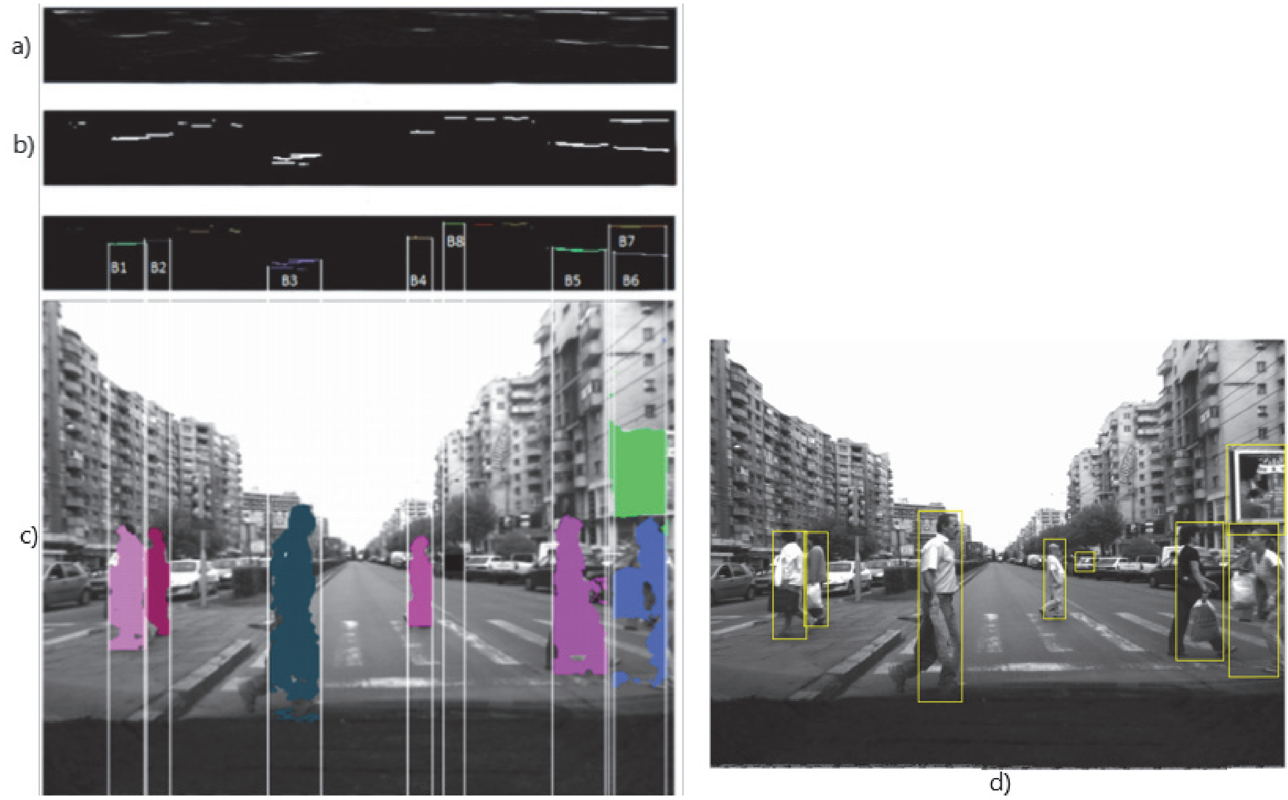


Figure 9 – Using the method of building maps of *u-* and *v-disparity* for the classification of objects:
(a) – the result of accumulation of *u-disparity*,
(b) – threshold filtering,
(c) – detection of horizontal borders of objects,
(d) – detection of areas of interest in the space of *uv-disparities*

4.3. Digital elevation map

Stereo cameras are also used in the task of constructing terrain models based on two-dimensional, three-dimensional, and the so-called 2.5-dimensional maps (digital elevation maps).

Construction of three-dimensional maps on a cloud of points gives the most complete description of the area, but is a very resource-intensive task. Two-dimensional maps are stored only points of average height in the cell, so they cannot be distinguished inclined walkable surface and a vertical obstacle.

Compromise in this case is the method of constructing digital height maps, described in [18]. This method makes it possible to detect and classify the pavement, walkable and not walkable obstacles based on their height. 3D data obtained from dense stereo is converted to a rectangular digital elevation map (DEM). Unlike the raw cloud of three-dimensional points, DEM provides a compact representation in the form of a flat grid with an explicit connection between adjacent three-dimensional positions.

In other words, for several three-dimensional points, only one height is stored that falls in the DEM. This allows three to five times to reduce the amount of calculations performed with them and to perform processing in real time. Based on the density of three-dimensional points, each DEM cell is classified as a road, obstacle, or road island [18] (see Fig. 10). Such maps can be used to create a semantic map of the area.

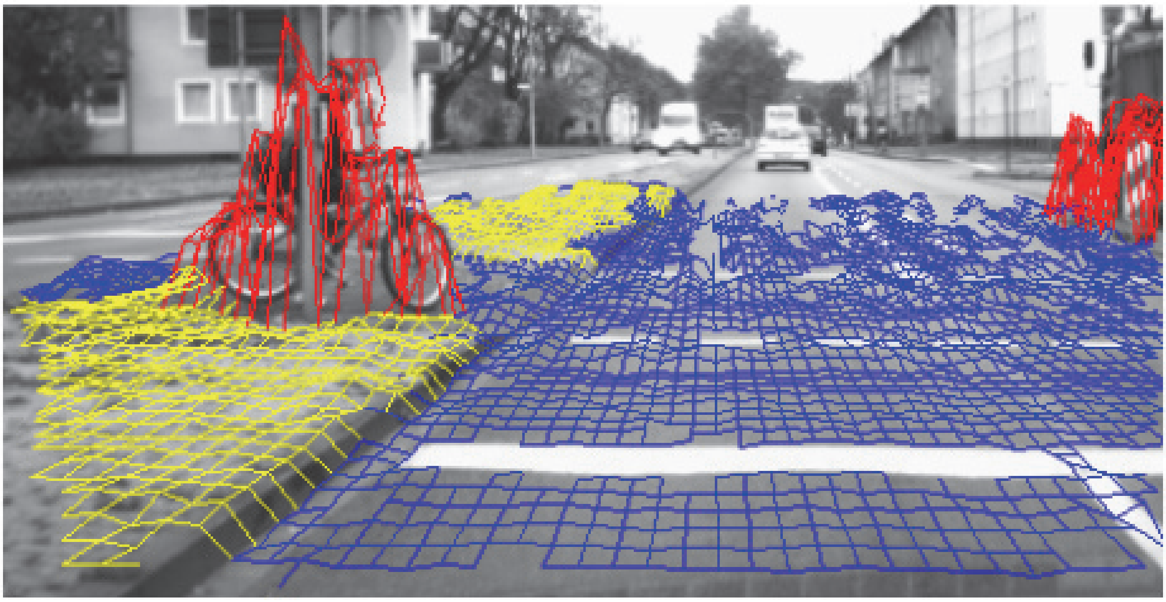


Figure 10 – Digital elevation map

4.4. Stixel segmentation

In [19] considered another segmentation option to build terrain model based on the so-called stixel. As in the urban environment is dominated by the vertical and horizontal surfaces, a dense cloud of three-dimensional points obtained from the stereo system or a scanning lidar approximated by a set of thin flat rectangles – stixels. The stixel representation is fed to the input of further processing steps and significantly reduces the amount of computation needed for them, which allows for real-time calculations.

For the first time, the concept of approximation of the three-dimensional environment of a vehicle using stixels appeared in [20]. The result of approximation by stixels turns out to be a reliable and very compact representation of the road situation, describing both the free space in front of the car, and static and moving objects. The result of the construction of the permeability model based on the pixel representation of the environment is shown in Fig.11 [20].

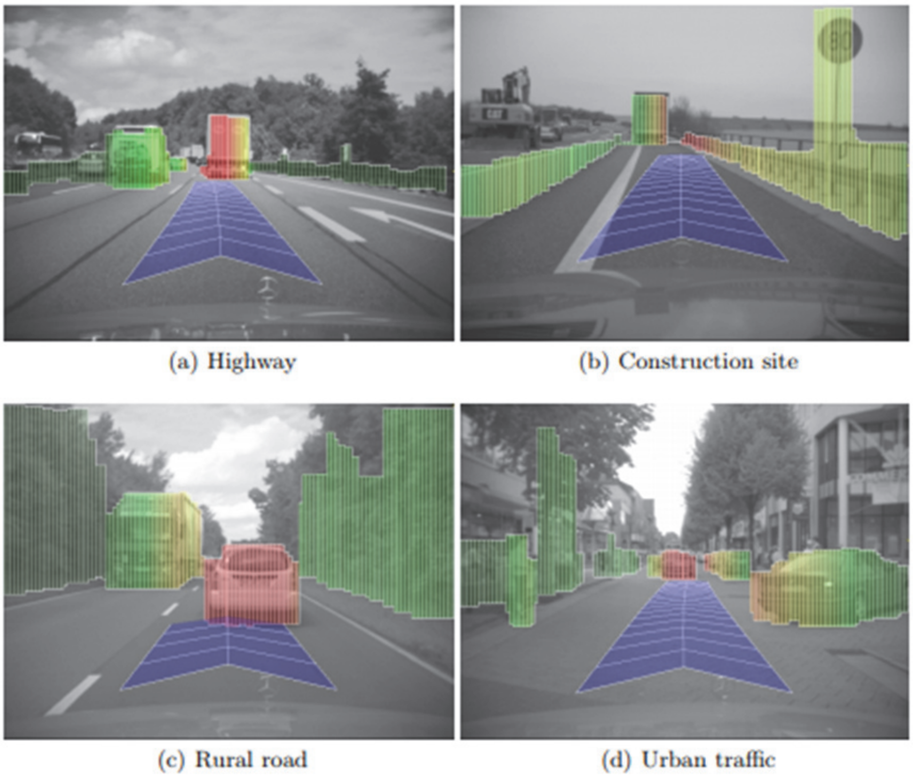


Figure 11 – Using stixel segmentation to build a terrain model in various environmental conditions

5. Road conditions analysis algorithm

To achieve the best results in the task of analyzing the traffic situation is possible only by combining the technologies of artificial intelligence and algorithmic methods discussed above. No technology is a universal solution that provides comprehensive information about the traffic situation around the vehicle. In [21], a classification of algorithms in unmanned vehicle vision systems is presented. According to it and the considered technologies of artificial intelligence, a scheme of a system for analyzing road conditions is developed (see Fig. 12).

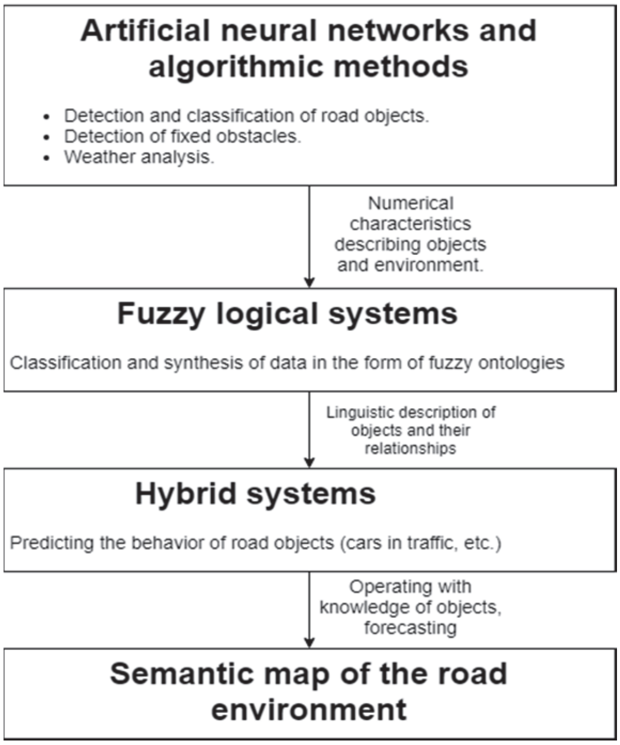


Figure 12 – The scheme of road conditions analysis algorithm

In this scheme moving from top to bottom increases the abstraction level concepts and increases the degree of working with knowledge than the data. For the purposes of the definition of road markings and construction of cross-country map it is preferable to use an algorithmic approach. For recognition and detection of road infrastructure (such as road signs, other cars and pedestrians) used neural network techniques. Then the detection results are analyzed by a fuzzy system, which reduces the number of classes required for learning neural networks. In the next step, the data enters the hybrid network, which analyzes from the knowledge put into the system and modified during the adaptation process.

The result of the analysis of the road situation is a semantic map that qualitatively describes the road situation around the vehicle, the objects and the relationship between all participants in the movement.

6. Conclusion

In this paper presents the classification of artificial intelligence technologies. From the point of view of the task of analyzing the road situation around an unmanned vehicle, the technologies of fuzzy production systems, artificial neural networks and hybrid systems are considered. In addition, algorithmic methods for image analysis, such as geometric clustering, disparity analysis, and pixel segmentation, are considered. It is shown in which areas of the analysis of road conditions these or those methods are most applicable. The scheme of the algorithm, the purpose of which is to build a semantic map of the environment, is presented.

The semantic map allows you to operate not only with data, but also with knowledge. It is based on the ontological description of objects and the relationship between them. The presence of a semantic map will improve the quality of navigation unmanned vehicles.

References

1. Herger, M.: Changes in Disengagements Over the Years, https://thelastdriverlicenseholder.com/2019/03/09/changes-in-disengagements-over-the-years/comment-page-1/, last accessed 2019/07/17.

2. Stankevich, L.: Intellectual information and control systems. Publishing house of the Polytechnic University, St. Petersburg (2011).
3. Widaa, A.H.A., Talha, W.A.: Design of Fuzzy-based autonomous car control system. In: International Conference on Communication Control Computing and Electronics Engineering (ICCCCEE), pp. 1-7. IEEE, Khartoum, Sudan (2017).
4. Armagan, E., Kumbasar, T.: A fuzzy logic based autonomous vehicle control system design in the torcs environment. In: 10th Electrical and Electronics Engineering (ELECO), pp. 737-741. IEEE, Bursa, Turkey (2017).
5. Naranjo, J.E., Gonzalez, C., Garcia, R., Pedro, T.: Using fuzzy logic in automated vehicle control. IEEE Intell. Syst., vol. 22 (1), pp. 36-45. IEEE (2007).
6. Olivares-Mendez, M. A., Sanchez-Lopez, J.L., Jimenez, F.: Vision-Based Steering Control, Speed Assistance and Localization for Inner-City Vehicles. In: Sensors (Basel), vol. 16(3): 362, (2016).
7. Buoncompagni, L., Mastrogiovanni, F., Saffiotti, A. Scene learning, recognition and similarity detection in a fuzzy ontology via human examples, https://www.researchgate.net/publication/320075283_Scene_learning_recognition_and_similarity_detection_in_a_fuzzy_ontology_via_human_examples, last accessed 2019/07/17.
8. Mandellos, N. A.: A background subtraction algorithm for detecting and tracking vehicles. Expert Systems with Applications, vol. 38, pp. 1619-1631. Pergamon Press, Inc. Tarrytown, NY, USA (2011).
9. Dkhil, M. B., Wali, A., Alimi, A.M.: Towards a Real Time Road Moving Object Detection and Tracking System. Journal of Information Assurance and Security, vol. 11, pp. 039-047. MIR Labs, USA (2016).
10. Badrinarayanan, V., Kendall, A., Cipolla., R.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.39, pp 2481 – 2495. IEEE (2017).
11. Demin, A.: Model of an adaptive control system and its application for controlling the movement of a virtual robot. Young Scientist 11 (46), 114 – 119 (2012).
12. Al-Mayyahi A., Wang W., Birch P.: Adaptive Neuro-Fuzzy Technique for Autonomous Ground Vehicle Navigation. Robotics 3(4), 349-370 (2014).
13. Khodayari, A., Ghaffari, A., Kazemi, R., Manavizadeh, N.: ANFIS Based Modeling and Prediction Car Following Behavior in Real Traffic Flow Based on Instantaneous Reaction Delay. In: 13th International IEEE Annual Conference on Intelligent Transportation, pp 599-604. IEEE, Funchal, Portugal (2010).
14. Broggi, A., Buzzoni, M., Felisa, M., Zani, P.: Stereo obstacle detection in challenging environments: the VIAC experience. 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1599–1604. San Francisco, CA, USA (2011).
15. Musleh, B., Escalera, A., Armingol, J.M.: U-V Disparity Analysis in Urban Environments. Computer Aided Systems Theory – EUROCAST, pp. 426–432. Springer, Heidelberg (2011).
16. Soquet, N., Aubert, D., Hautiere, N.: Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation. In: IEEE Intelligent Vehicles Symposium, pp. 160–165. IEEE, Istanbul, Turkey (2007).
17. Iloie, A., Giosan, I., Nedevschi, S.: UV-disparity based obstacle detection and pedestrian classification in urban traffic scenarios. In: IEEE Int. Conf Intelligent Computer Communication Processing, pp. 119-125. IEEE, Cluj Napoca, Romania (2014).
18. Oniga F., Nedevschi S. Processing dense stereo data using elevation maps: Road surface, traffic isle, and obstacle detection. In: IEEE Transactions on Vehicular Technology. 2010. vol. 59, №. 3. pp. 1172-1182. IEEE (2010).
19. Pfeiffer, D., Franke, U.: Towards a Global Optimal Multi-Layer Stixel Representation of Dense 3D Data. In: The 22nd British Machine Vision Conference, pp. 1-12. BMVA Press, Dundee, UK (2011).
20. Badino, H., Franke, U., Pfeiffer, D.: The stixel world – a compact medium level representation of the 3d-world. In: Pattern Recognition, 31st DAGM Symposium, pp. 51-69. Springer, Heidelberg (2009).
21. Bakhshiev, A., Orlova, S., Komarov, A., Stepanov, D.: Classification of scenarios and algorithms in the technical vision systems of unmanned vehicles. EXTREME ROBOTICS 1(1), 400-409, (2018).

S. Orlova[1], T. Isakov[2]

USING OF DEEP NEURAL NETWORKS FOR SEGMENTATION OF DRIVING ENVIRONMENT IMAGES

[1] The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
[2] Peter the Great St.Petersburg Polytechnic University, Saint Petersburg, Russia
s.orlova@rtc.ru

Abstract

An increased interest in the development of technologies for vehicular automation is the reason that the computer vision problems associated with the recognition of driving environment are highly relevant now. Neural networks application in such tasks allows to achieve better performance than the use of classical methods of technical vision. Currently, a lot of development of neural network methods is underway, including for image segmentation, which gradually push the state of the art.

However, despite the large number of works devoted to the development of neural network segmentation methods, there are a number of problems associated with the practical solution of the segmentation problem in the development of computer vision system. One of the main problems is the lack of experimental studies comparing the performance of existing methods, which would be a support when choosing the most appropriate method for the task.

The report discusses the use of deep convolutional neural networks to solve the problem of segmentation of images of the road situation, compares modern algorithms and formulates recommendations on the choice of specific algorithms depending on the task.

Keywords: image semantic segmentation, neural network, computer vision, driving environment.

Acknowledgements

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 “Cloud services for automatic synthesis and validation of datasets for training deep neural networks in pattern recognition tasks”.

1. Introduction

Currently, there is a very high interest in the field of transport automation and autonomous driving technologies. Some results have already been achieved, but these areas are still on the path of active development. An important part of an automated vehicle or mobile robot is a computer vision system, the tasks of which include perception and low-level analysis of the surrounding scene. In the past few years, the ability to solve such problems has increased significantly due to neural networks. Classification and segmentation of images, objects detection in images and some other tasks - for all these problems, the solutions with the best quality and speed-quality ratio are generally achieved using neural networks.

Segmentation of road images allows to get the needed data both for direct use in control – for example, separation of roadway areas and sidewalks, lawns and other surfaces that cannot be paved, as well as data useful as auxiliary for other computer vision algorithms - for example, as an additional condition for object detection algorithms: pedestrians and cars cannot be located in the image areas corresponding to the sky, trees, the upper parts of the facades of buildings. It is the segmentation of the driving environment images that reflects the simplified structure and composition of the scene.

However, when developing and implementing image segmentation algorithms, there are difficulties with the choice of the algorithm that corresponds to the task, or the choice of direction (group of methods), which should be taken as the basis for research and development of new algorithms. There is a significant amount of scientific literature with proposals for new methods, but the authors do not present a comparison with all modern methods – which is why it is difficult to form a complete picture. In addition, the authors do not always use the original implementations or the best weights for the neural network models with which they are compared. Not always in literature there is a fairly complete description of such models, which allows us to understand what kind of implementation of the algorithm the authors used. This article addresses this problem. Let us try to bring some clarity to the current situation in the field of segmentation methods, taking the task of segmentation of road images.

2. Image Segmentation in Computer Vision

2.1. Problem of Image Segmentation

Image segmentation is the process of dividing an image into parts that are homogeneous by some criterion, usually corresponding to objects, surfaces, etc., and the assignment of classes or identifiers to these parts. We can distinguish several segmentation problems that involve obtaining various outputs (Fig. 1).

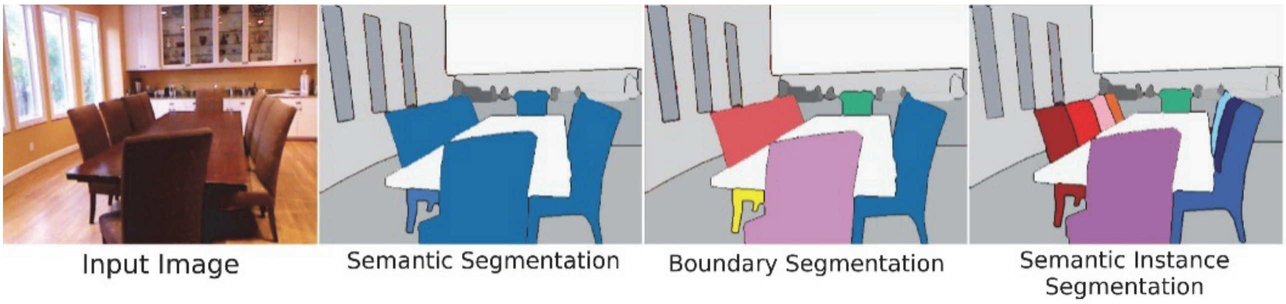


Figure 1 – Types of image segmentation

Semantic segmentation is a per-pixel classification, i.e. each pixel of an image is assigned a class. Border segmentation is the selection of the borders of objects in an image. Appearance of such a segmentation task seems to be related to the relative simplicity of determining edges and borders in an image by classical methods of computer vision. Instance semantic segmentation extends the task of semantic segmentation to the separation of individual objects (instances of classes). Usually, the result of the algorithm is presented in the form of mask images, on which the value of each pixel represents the class or object identifier for the corresponding pixel in the original image.

The most common tasks or applications of segmentation are the segmentation of biomedical images, the segmentation of images of road and city scenes for the computer vision systems of mobile robots, automated vehicles, for security systems. Segmentation is also used to process aerial or satellite images to select objects (equipment, buildings, roads) and terrain types (forest, meadow, water).

2.2. Image Segmentation Techniques

We did not find in the literature any generally accepted classification of segmentation methods, but we can still separate seven common groups of methods [1]:

1. Threshold methods – image pixels are divided into parts based on their intensity level, i.e. threshold filtering is performed.
2. Edges based methods – often at the edges of areas there is a strong difference in brightness, however, to select an object on the image, you need closed area borders.
3. Regions based methods – the image is divided into regions with similar characteristics, usually using one of the two most common techniques: region growing methods or region splitting and merging methods.
4. Clustering based methods – image pixels are divided into clusters so that the elements inside the cluster are similar to each other according to certain characteristics more than the elements of other clusters. There are two main groups of clustering methods: hierarchical, based on the idea of trees, and partition methods that use optimization algorithms to iteratively minimize the objective function.
5. Watershed methods – the absolute value of the image gradient is considered as a topographic surface, i.e. pixels on a gradient image with greater brightness (large changes in brightness – edges, borders) are “mountains”, and pixels with lower brightness are “valleys”. Lines of bright pixels - "ridges" are the lines of watersheds – the borders of objects.
6. Partial differential equation based methods – the result of the work is blurred edges and boundaries that can be shifted with the help of closing operators; second-order methods are used to better detect edges and borders, and the fourth methods are used to reduce the noise from the image.
7. Methods based on neural networks – artificial neural networks (usually convolutional) are used.

In the past few years, various competitions have become popular in which teams compete in solving a given task. The organizers provide a set of annotated data that can be used to train or customize the parameters of the algorithm, and a set of non-annotated data for which teams get the result, and then, using the metrics set by the organizers, the results of various teams are compared to determine the winners. Many competitions related to image segmentation can be cited: “Ultrasound Nerve Segmentation” 2016 [2], “Carvana Image Masking Challenge” 2017 [3], “KONICA MINOLTA - Pathological Image Segmentation Challenge” 2017 [4], “Robotic Instrument Segmentation Sub-Challenge” 2017 [5], “CVPR 2018 WAD Video Segmentation

Challenge" 2018 [6], "TGS Salt Identification Challenge" 2018 [7] and others. In all these competitions, the teams that won prizes used neural network methods. In addition, on the websites of some public datasets, there are tools for evaluating the results and a constantly updated list of rated methods is presented. All of these methods also use neural networks.

Thus, it becomes clear that the most promising group of methods for solving the segmentation problem in the general case are neural network methods.

2.3. Neural networks for semantic segmentation

The modern approach to the construction of neural network algorithms for solving the segmentation problem dates back to 2015, when neural network architectures FCN [8], UNet [9] and DeepLab [10, 11, 12] were proposed.

The FCN (Fully Convolutional Network) architecture, built using only convolutional layers. The paper considers a neural network for classification, taken without the last fully connected layers that convert the resulting feature map to a class number, as an encoder. The encoder converts the input image into feature maps of various levels - on the initial layers, low level features (points, edges, spots) are identified, on the final, a high level feature search (complex shapes, parts of objects, objects) is found. Note that such feature maps are significantly smaller than the input image. In order to further obtain a segmentation mask from the features, it is proposed to use a neural network decoder, which accepts feature maps from the encoder and increases their dimension to the size of the input image using deconvolutional layers (using the deconvolution operation). The accuracy of the results obtained on popular datasets turned out to be quite high and surpassed the results of existing methods. All subsequent modern architectures used the principle of the encoder-decoder proposed by the authors of FCN.

UNet got its name because of the U-shaped architecture (Fig. 2) and in 2015 won in several competitions on the segmentation of medical images. The left descending part is the encoder decreasing the dimension of the output feature maps, and the right ascending part is the decoder increasing the dimension. The encoder is represented as a sequence of four blocks. Each block consists of a pair of 3x3 convolutions with the ReLU activation function and a 2x2 subsample layer in increments of 2. In each block, the number of channels of feature maps is twice as large as in the previous one. After the encoder and before the decoder, in the middle of the network, there are two 3x3 convolutional layers with the ReLU activation function. The decoder also consists of four blocks, each block contains a deconvolutional (or up-sampling) layer, which increases the resolution of the feature map, and then two 3x3 convolutional layers with the ReLU activation function. There are additional links between the encoder and the decoder: the output feature map from each encoder block is input to the corresponding decoder block, i.e. the first block of the encoder is connected to the last block of the decoder, the second block of the encoder is connected to the third block of the decoder, etc. After the encoder and before the decoder, in the middle of the network, there are two convolutional layers with the ReLU activation function. The neural network architecture ends with a 1x1 convolutional layer with the number of channels equal to the number of classes, where for each pixel there are values corresponding to a class. Also, in order for the output image (mask) size to be equal to the size of the input image, and not to be reduced due to convolution operations, the authors propose an addition of the input image at the edges with a reflection. A feature of UNet is a large number of channels in the decoder part, which allows to store more information in the last layers, as well as the use of elastic deformations during augmentation of the training base, which is often found in biomedical images. Although the authors assumed the use of UNet for the segmentation of biomedical images, this neural network has become very popular, and is currently showing good results in various tasks. Often the original version of the encoder is replaced with another classifier - usually VGG-16, ResNet.

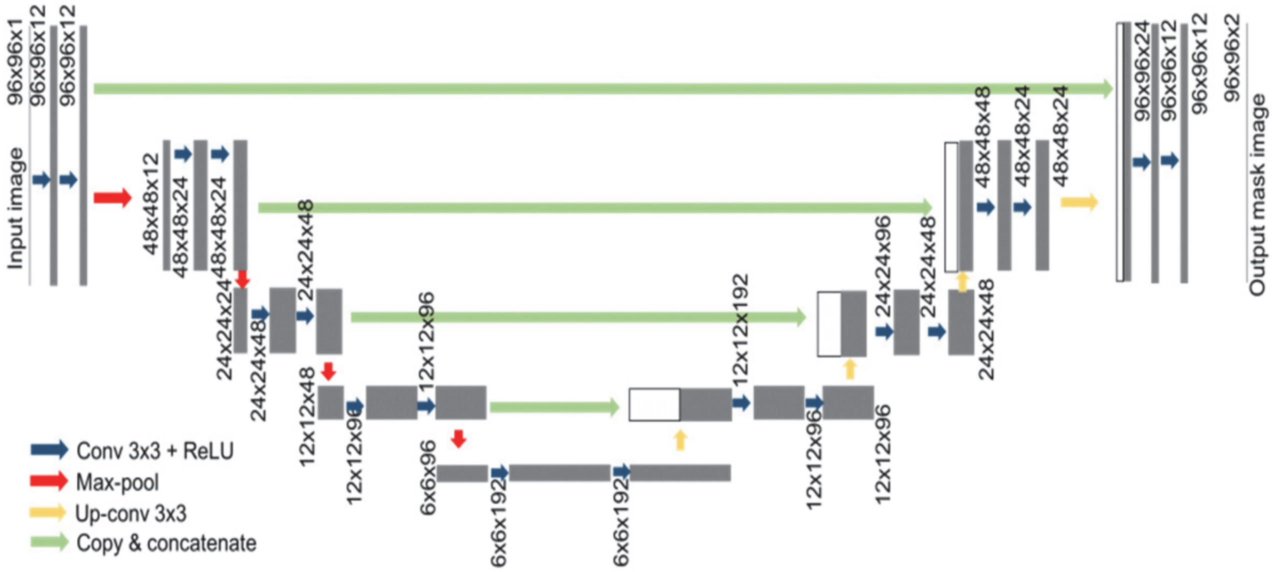


Figure 2 – Neural network UNet

The DeepLab algorithm uses a different approach to decoder implementation — using CRF (Conditional Random Fields). DeepLab treats each pixel at the output of a convolutional neural classifier network as a CRF node, and uses the CRF output directly to optimize the loss function controlled by a convolutional encoder network. Such an algorithm makes it possible to obtain more accurate contours of objects on the output segmentation mask, in contrast to the full-screwed methods that redundantly smooth the contours, thereby missing details (Fig. 3, below).

The principle of the algorithm is shown in Figure 3, above. The coarse mask obtained at the output of the deep convolutional network is increased in size by bilinear interpolation, and then fully connected CRFs are applied to it, specifying the result, resulting in a detailed, precise segmentation mask. The authors, experimenting with algorithm parameters, used atrous (or dilated, convolution with holes) convolutional layers, which significantly increased the accuracy and reduced the number of parameters to be trained, and the connection between the initial and final layers of the feature extractor network for better segmentation of various-sized objects, which is also positive affected the quality of work, but to a much lesser extent than the use of CRF and atrous convolutional layers.

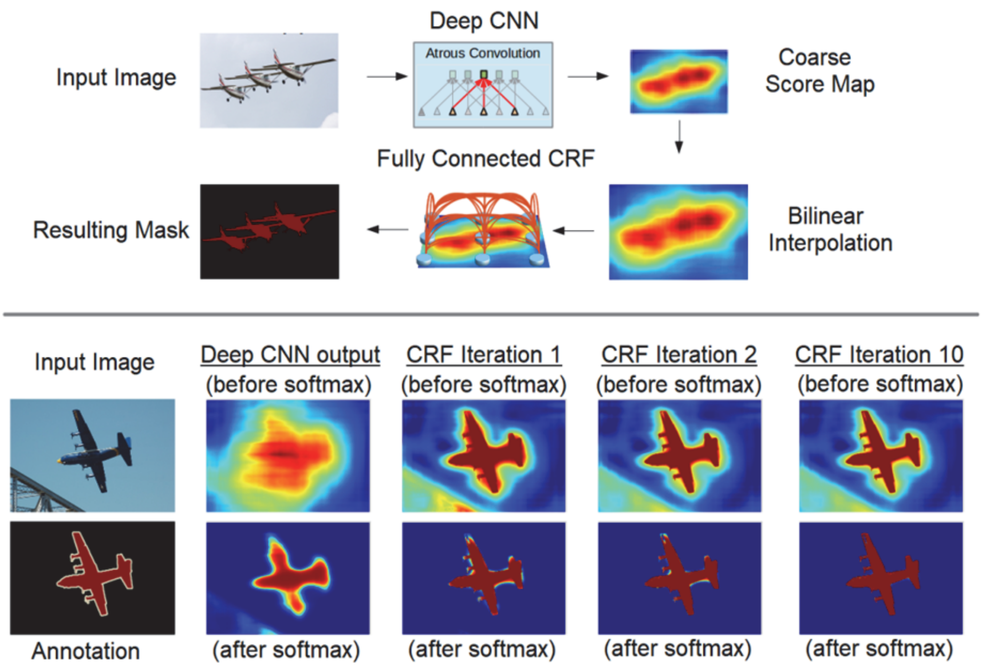


Figure 3 – Algorithm DeepLab: principle of operation (top) and the effect of CRF on the final result (bottom)

Then in 2016, the authors of DeepLab-CRF presented the second version - DeepLab v2 – and an article describing in detail the work of this method [13]. In addition to various small changes in the algorithm following the principles of the first version, in the second version the atrous convolutional layers are combined with the idea of connections between non-neighboring layers, i.e. parallel branches of the architecture, forming ASPP – atrous spatial pyramid pooling, allowing high-quality segmentation of multi-scale objects. This allowed DeepLab v2 to demonstrate better performance than the first version, and become one of the most effective neural network algorithms at that time.

In addition, in 2016, SegNet [14, 15], ENet [16] and PSPNet [17] were introduced.

The SegNet neural network uses the first 13 layers of the VGG-16 classifier as a classifier part (encoder). The role of the decoder is to expand the resulting low resolution feature map from the encoder output into the final network output, which has the same resolution as the input data. That is, the principle of FCN is used. The authors assumed the use of this architecture in the tasks of understanding scenes and conducted experiments with images of the driving environment and indoor scenes. The network structure, like UNet, is presented in the form of blocks with links between the encoder and the decoder, however, the blocks contain a different number of convolutional layers (2 or 3). Also the way to increase resolution is different, in SegNet it is upsampling layers. Upsampling in SegNet is performed by storing the indexes of the maximum elements remaining after the pooling layer. For this, 2 bits are enough for each 2x2 subsample window. Thus, the way to increase the dimension is the main feature of SegNet.

ENet (from “Efficient neural network”) was created for tasks that require high speed, and its main feature is the network structure and the use of the PReLU activation function. The network architecture is shown in Figure 4.

ENet arcitecture with input size of 512 x 512

Name	Type	Output size
initial		$16 \times 256 \times 256$
bottleneck1.0	downsampling	$64 \times 128 \times 128$
4× bottleneck1.x		$64 \times 128 \times 128$
bottleneck2.0	downsampling	$128 \times 64 \times 64$
bottleneck2.1		$128 \times 64 \times 64$
bottleneck2.2	dilated 2	$128 \times 64 \times 64$
bottleneck2.3	asymmetric 5	$128 \times 64 \times 64$
bottleneck2.4	dilated 4	$128 \times 64 \times 64$
bottleneck2.5		$128 \times 64 \times 64$
bottleneck2.6	dilated 8	$128 \times 64 \times 64$
bottleneck2.7	asymmetric 5	$128 \times 64 \times 64$
bottleneck2.8	dilated 16	$128 \times 64 \times 64$
Repeat section 2, without bottleneck2.0		
bottleneck4.0	upsampling	$64 \times 128 \times 128$
bottleneck4.1		$64 \times 128 \times 128$
bottleneck4.2		$64 \times 128 \times 128$
bottleneck5.0	upsampling	$16 \times 256 \times 256$
bottleneck5.1		$16 \times 256 \times 256$
fullconv		$C \times 512 \times 512$

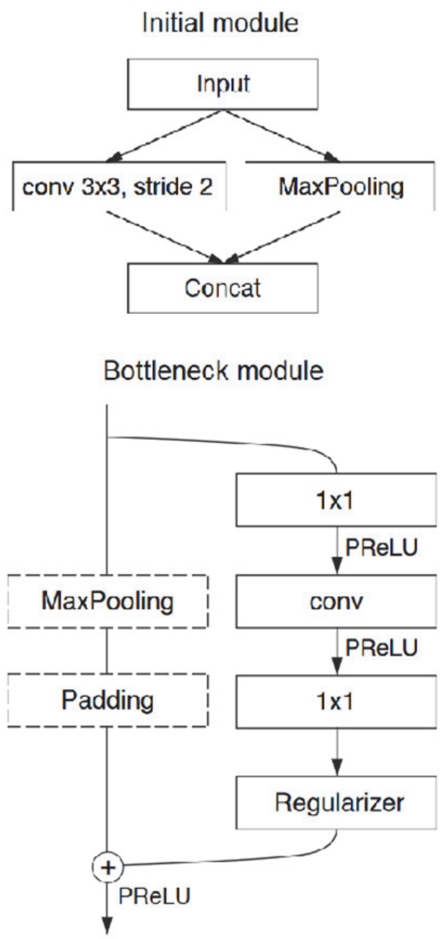


Figure 4 – The main modules of the ENet neural network: the initial module and the bottleneck module

The initial network module consists of two parallel layers - 3x3 convolutions with offset of 2 and pooling, the outputs from which are then concatenated. The convolutional layer has 13 channels, and after concatenation with the output of the pooling layer, a feature map with 16 channels is obtained. This is followed by 19 bottleneck modules. The bottleneck module consists of two branches, one of which contains MaxPooling and Padding layers, and the other contains a sequence of 1x1 convolution layers, an adjustable convolution (denoted as conv), a convolution 1x1, and a regularizer layer. Between all layers of convolution, batch normalization and the activation function PReLU are also present. The two branches are then connected. Adjustable convolution can be a regular convolutional layer, a deconvolutional layer or an atrous 3x3 convolution, or a 5x5 convolution divided into two asymmetric ones.

The authors [16] also compared ENet with SegNet, the results of which are shown in Figure 5 on the Cityscapes dataset [18]. It can be seen that, in terms of accuracy, ENet is on par with SegNet, and in terms of speed, ENet is many times more productive and less demanding on computing resources.

Speed comparison

Computer	NVIDIA TX1						NVIDIA Titan X					
Input size	480×320		640×360		1280×720		640×360		1280×720		1920×1080	
	ms	fps	ms	fps	ms	fps	ms	fps	ms	fps	ms	fps
SegNet	757	1.3	1251	0.8	-	-	69	14.6	289	3.5	637	1.6
ENet	47	21.1	69	14.6	262	3.8	7	135.4	21	46.8	46	21.6

Comparison of size, trainable parameters number and required computer performance for input size of 640 x 360

Model	GFLOPs	Par. number	Size (fp16)
SegNet	286.03	29.46M	56.2 MB
ENet	3.83	0.37M	0.7 MB

Results on Cityscapes dataset

Model	Class IoU	Class iIoU	Category IoU	Category iIoU
SegNet	56.1	34.2	79.8	**66.4**
ENet	**58.3**	**34.4**	**80.4**	64.0

Figure 5 – Comparison of the performance of ENet and SegNet

In 2016, the PSPNet segmentation neural network [17] won first place in the ImageNet Scene Parsing Challenge (INSPC) competition, the ASCAL VOC 2012 Benchmark and Cityscapes Benchmark. The authors of the architecture paid great attention to the problem of not using global information segmentation algorithms in the image. Examples of such a problem: the neural network mistakenly classifies a boat on the background of a river as a car, gets a “ripped” mask of a skyscraper compared to other buildings, confusing it with the “building” class (confusing categories), doesn’t distinguish a pillow from a bed sheet (inconspicuous categories), although usually there is a pillow on each bed.

PSPNet uses the ResNet classifier with dilated convolutional layers, after which it receives an output matrix 8 times smaller than the original image. After that, the output matrix is subjected to a pooling operation in parallel along four branches. Then, the number of channels in the matrices obtained after the subsample is reduced N times with the help of the 1x1 convolutional layer, where N is the number of pooling branches, i.e. in this case, four times. After that, the resulting matrices are brought to the size of the initial feature map obtained from the output of the encoder. All these matrices are concatenated and pass through the convolutional layer to obtain a final prediction (segmentation mask).

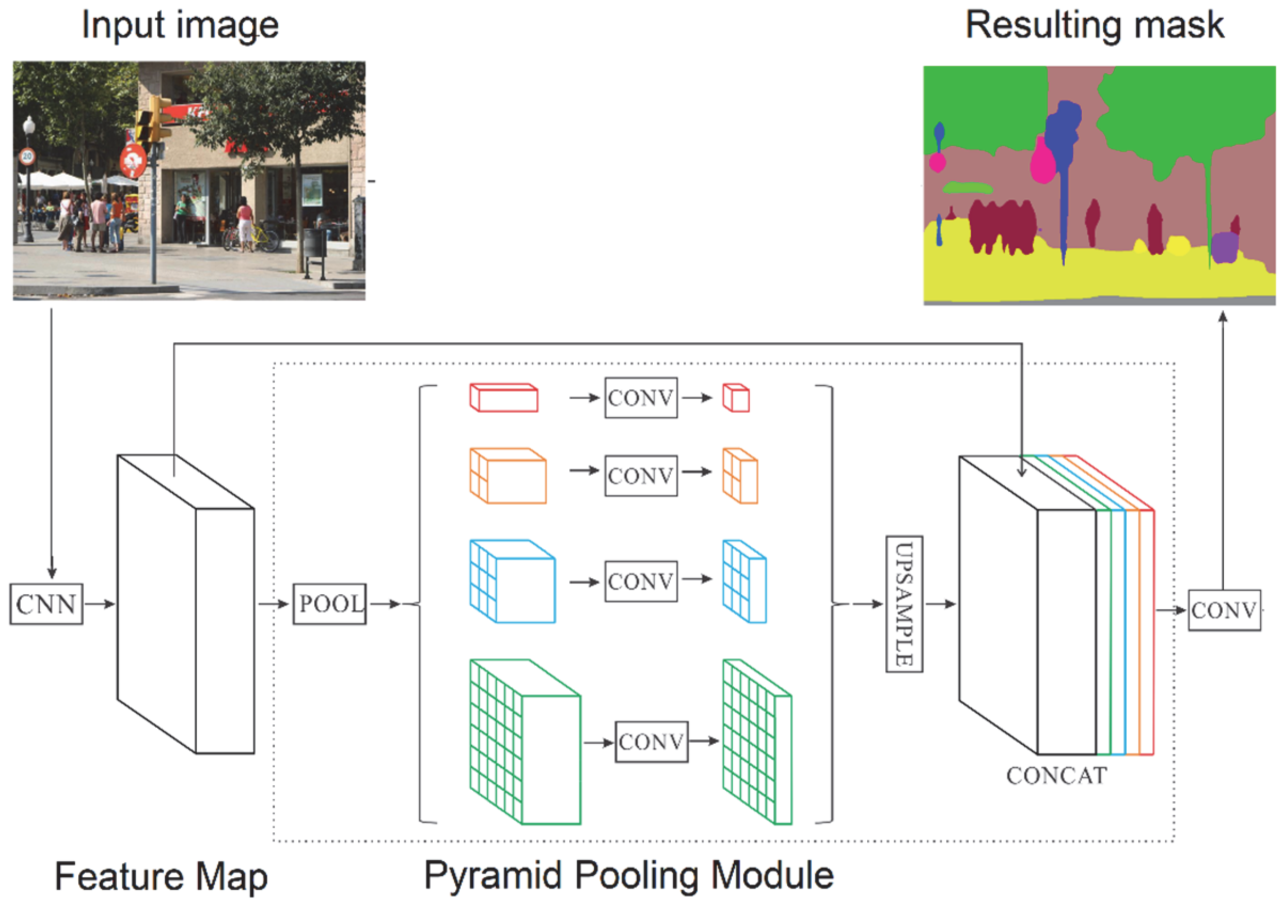


Figure 6 – PSPNet segmentation neural network

In 2017, the most popular and significant models were DeepLab v3 [19, 20] and ICNet [21].

The authors of the third version of DeepLab deepened even more at the possibility of atrous convolution, which allows to effectively control the size of the output feature map and the area of vision of the convolution filter. As a result, the network architecture is similar to the PSPNet architecture, consisting of a feature encoder as a classifier, then a pooling pyramid and one final convolutional layer. The feature of DeepLab v3 is that in the pooling pyramid it does not contain the stage of reducing the number of channels; also, atrous convolutions are used in the pyramid, replacing the subsample by region (by cell), and only one branch containing the global pooling layer. Network accuracy is higher than PSPNet performance.

ICNet (from "Image Cascade Network") was developed as an instance of a high-speed algorithm suitable in practice in tasks that require high speed image processing. The architecture is based on a cascade of branches, in which the image is processed with different resolutions. The input image is reduced by 2 and 4 times, forming a cascade input for three branches - low, medium and high (original) resolution. Next, the image in each branch passes a series of convolutional layers (Fig. 7).

Under each of the input images and under each feature map in parentheses indicate the ratio of their size to the size of the original image. Some operations are used only during training, some - only during inference (testing and prediction). CFF (Cascade Feature Fusion) blocks denote operations of connecting features of different scale, obtained in different branches of the network. Low and medium resolution branches use shared weights and operations.

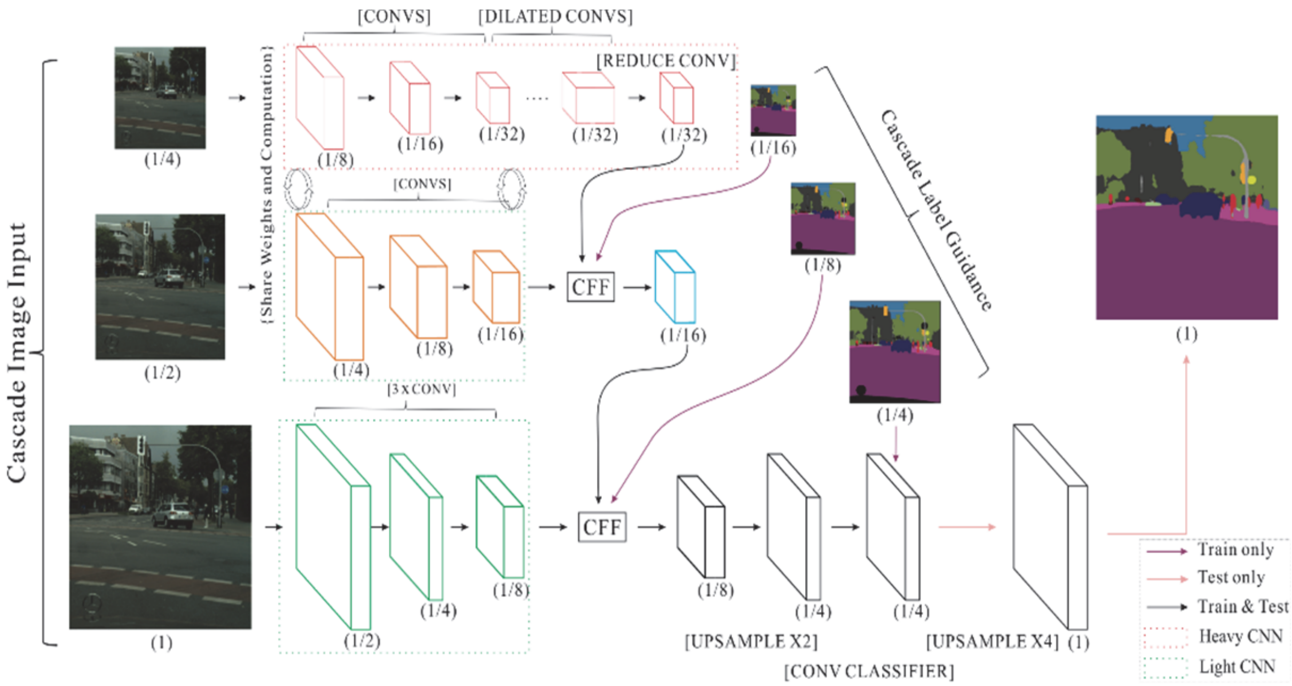


Figure 7 – ICNet neural network architecture

Branching reduces the computational complexity of the algorithm. Thus, the image reduced by 4 times is fed to the input of the upper branch (low resolution branch) containing a complete segmentation neural network, the output of which is a coarse mask, reduced by 32 times the original image size. Due to the small input resolution, the calculations do not take much time. The resulting mask does not contain fine details and is not accurate, but displays the main semantic areas in the image, which can significantly reduce the number of trainable parameters in the middle and lower branches. Further, in the branches of medium and high resolution, the mask is improved and supplemented with details, forming the resulting segmentation mask having the same size as the input image. When training a network, scaling annotations is used.

Thus, the main feature of ICNet lies in the idea of using a post-convolutional neural network to obtain a basic, coarse mask, and the sequential refinement of the mask by convolutional networks with a small number of layers. Due to the fact that the base mask is calculated for the image in low resolution and therefore does not lead to long calculations, and convolutional networks in the branches of medium and high resolution due to their small size also do not require much time for calculations, ICNet achieves a very favorable balance of accuracy and speed. In terms of quality, it is on par with PSPNet, but it takes several times less time for inference.

3. Implementation of Image Segmentation Algorithms using Neural Networks

3.1. Software tools for implementation of neural network models

Since the objective of this study is to compare modern segmentation algorithms, it is obvious that all implementations of the selected algorithms should be executed and tested on a single software base. And since such a comparison is necessary for formulating recommendations on the use of certain algorithms for solving applied problems, it is also necessary to choose the software base taking into account the possibility of embedding algorithms in other software.

At the moment there are several popular libraries and frameworks for machine learning: DeepLearnToolbox for Matlab [22], Torch [23] and PyTorch [24], Caffe [25], TensorFlow [26] and Keras [27]. The DeepLearnToolbox library is only suitable for research purposes, which does not meet our requirements.

Torch has many tools for working with deep learning, has the ability to parallelize calculations using CUDA, but it has an interface for the not so popular Lua language and at the moment does not have a very large audience of users. Nevertheless, there is a more popular framework, PyTorch, created on the basis of Torch, which has a Python programming interface and several modules for working with neural networks of different levels of abstraction. Currently supported by Facebook.

Caffe is one of the first popular open source libraries for machine learning. It was developed in 2013 at the University of Berkeley by Yangqing Jia in the process of preparing his thesis and is still supported by the Berkeley Vision and Learning Center. Implemented in the C++ programming language and supports the interface in C++, Python and Matlab. Supported operating systems: Windows, Linux and Mac OS X. The library has a significant set of tools: layers of various types, activation and loss functions, various teaching methods.

The framework for machine learning TensorFlow was originally developed by Google for internal use and appeared in open access at the end of 2015. The computational part is written in C++, the main API is implemented in Python, and there are also implementations for C++, Haskell, Java and Go. TensorFlow is available for 64-bit Linux, MacOS, Windows, and mobile computing platforms, including Android and iOS. There is a built-in TensorBoard tool that can be used to visualize the computational graph, weights and maps of neural network layers and other data, as well as automatically log errors, accuracy and any other metrics during training or testing and then represent them as charts. It has a wide range of tools for machine learning and neural networks, the ability to implement custom operations and layers, and also has various tools for embedding trained models into software systems. That is, this library is designed as a universal tool for researching, implementing and embedding machine learning algorithms.

Keras is an open library for working with deep learning, written in Python. It is built on top of the TensorFlow or Theano frameworks and provide higher abstraction level. However, the work on Theano was stopped two years ago, support – a year ago, therefore Keras is further considered on the basis of TensorFlow. It is aimed at simplifying work with neural networks, while being designed to be compact, modular and expandable. The main author and supporter is Francois Chollet, Google engineer. Keras supports almost all TensorFlow functionality related to neural networks, since it is more like an interface than an independent machine learning library. Contains numerous implementations of widely used building blocks of neural networks, such as layers, loss and activation functions, optimizers. Provides a high level of abstractions and a lot of auxiliary and additional tools for working with data. In the case of using TensorFlow as a low-level library when building a neural network from modules, Keras automatically builds a computational graph, which can be used later as a regular TensorFlow graph (without Keras), and all TensorFlow tools will be compatible with such a graph, which is very convenient for later embedding of resulting model.

In [28], the author aims to find out which library or framework is growing in popularity more quickly than others, and also compares recent activity related to these libraries - new articles, activity on GitHub, requirements in vacancies. TensorFlow was both the most popular at the moment and the fastest growing, the second place was taken by the PyTorch framework, the third - by Keras. Thus, any of these three libraries is a worthy choice.

In this study, TensorFlow and Keras were chosen, mainly due to the possibility of rapid development using Keras and the possibility of simple embedding using TensorFlow. Also, one of the reasons is that the authors have working experience with these libraries.

3.2. Datasets

For training and testing segmentation algorithms, suitable datasets are needed. Datasets should, first of all, contain annotations. Annotations for the segmentation of the driving environment are usually images in which each pixel is already assigned to a particular class. Classes are the main objects and areas that are represented on the images (road, pavement, signs, pedestrians, cars, trees, etc.).

Cityscapes [18]. This is a large dataset that contains a diverse set of stereo video sequences recorded on the streets of 50 different cities, with high-quality (fine) annotations-masks for 5,000 frames in addition to a large set of 20,000 coarse annotated frames. There are 33 semantic classes in the dataset, however not all of them are suitable for training, therefore usually 19 classes are used in this set. The Cityscapes dataset is intended for training and evaluating the effectiveness of algorithms for the two main tasks of a semantic understanding of road conditions: semantic segmentation at the pixel level and at the instance level.

Berkeley DeepDrive (BDD100K) [29]. BDD100K is by far the largest publicly available set of video data. It contains 100,000 videos representing over 1000 driving hours with over 100 million frames. Videos include GPS location, IMU data and timestamps. BDD100K allows solving such tasks as detection of road objects, driving zone segmentation and semantic segmentation, represented by 40 classes, but not all of them are suitable for learning, therefore, usually in this set they use 19 classes - the same as in Cityscapes. For the first two tasks, 70,000 training and 10,000 evaluation images are provided. For the third task, 7000 training and 1000 evaluation images were provided.

SYNTHIA [30]. SYNTHIA is a collection of synthetic images and annotations (200,000 images from video streams and 20,000 independent frames), which is a set of data created for solving semantic segmentation problems and related problems of understanding the scene in the context of driving scenarios. The data set consists of a collection of photorealistic frames rendered in a virtual city and contains precise semantic annotations at the pixel level, represented by 23 classes. There is a wide variety of scenes in the database: a European-style city, a modern city, a highway and a rural area, different seasons, days and different weather conditions.

CamVid [31]. The Cambridge CamVid database is the first collection of videos with semantic annotations of classes of objects, supplemented by metadata. The database contains more than ten minutes of high-quality video with a frame rate of 30 Hz, with corresponding annotations with a frequency of 1 Hz (in some places, 15 Hz). The annotation assigns each pixel one of 32 semantic classes. The database contains 700 annotated images.

MS-COCO [32]. COCO is a large-scale set of data for detection and segmentation, consisting of 300 thousand images, 200 thousand of which have annotations. The set contains layouts in 80 categories, among which there are categories suitable for the task of segmentation of road images (car, bus, traffic light, train, truck, motorcycle, bicycle, person), but only some of the images taken from the cameras located on the front panel car, so this set is not very suitable for a computer vision system of an autonomous vehicle. There is a Python API for working with this dataset.

PASCAL [33]. The PASCAL VOC dataset was created to solve the problems of classification, detection, and segmentation. From 2005 to 2012, the dataset was updated annually in accordance with the tasks set by the developers. The latest version of PASCAL VOC 2012 contains 9993 images for segmentation. Segmentation is performed on 20 classes.

3.3. Metrics

Evaluation of segmentation methods can be done using various metrics. The most obvious is pixel accuracy (global pixel accuracy) or pixel error, which is the ratio of true or falsely predicted pixels to the total number of pixels in an image or total number of pixels in mask corresponding to a given class (Expressions 1, 2). Such a metric can also be calculated over a whole set of images - then it will be the ratio of true or falsely predicted pixels on all images to the total number of pixels of all images.

$$\text{Pixel accuracy} = N_{\text{true}} / N_{\text{total}} \quad (1)$$

$$\text{Pixel error} = N_{\text{false}} / N_{\text{total}} \quad (2)$$

Pixel accuracy allows a fairly qualitative assessment when different semantic classes have the same significance. However, such a metric does not provide information on errors by class, although such information would be very useful. Often the methods have a high error on the classes of small and complex objects - signs, poles, pedestrians, while on large classes (sky, road, sidewalk, facades) the error is small. With this information, you can adjust the penalty coefficients by class and use other training methods with unbalanced data. There are also cases when a dataset contains too few members of a class, and the neural network is not well trained with this class. Assessing the error of the trained model by class, such a problem could be noticed.

Average pixel accuracy by class shows how accurately segmentation is performed on average for each of the classes. To begin with, pixel accuracy is calculated for each of the classes: the number of correctly predicted pixels is divided by the number of pixels corresponding to the correct annotation. The resulting values are summed, the amount is divided by the number of classes. This metric is less susceptible to class imbalances in frequency of occurrence than global metrics. However, the background (non-annotated areas) absorbs all the false predictions without affecting the accuracy of the class. This metric is more suitable for datasets without a background class, or it is necessary to perform calculations in another way to remove this problem.

IoU (Intersection over Union) solves this problem. It is based on the Jaccard similarity coefficient (or Jaccard index) and measures the ratio of the intersection to the union of the labeled segments for each class (Expression 3). In this method, the pixels on the resulting segmentation mask are divided into 3 groups:

– True positive (TP) – pixels correctly assigned to this class, being the intersection of the predicted mask and annotation mask;

– False positive (FP) – pixels falsely assigned to this class, being pixels of the class in the predicted mask, but belonging to another class in the annotation mask;

– False negative (FN) – pixels, falsely assigned to another class, being pixels of this class, which are falsely missing in the predicted mask.

To calculate the Jaccard index, the number of true positive pixels is divided by the sum of pixels from all three groups. It can be noted that such a sum is the union of the class mask in the prediction and annotation, therefore the metric is called “intersection over unoin”, i.e. intersection divided by union. Having the values of the IoU metric for each class (Expression 4: C is the number of classes, i is the class number), mean IoU (mIoU) can be calculated. Thus, the Jaccard index takes into account both false predictions and missing areas for each class.

$$IoU = TP / (TP + FP + FN) \tag{3}$$

$$mIoU = (\sum_i^C IoU_i) / C \tag{4}$$

However, sometimes there are tasks in which the quality of segmentation of the contours of objects is important. There are many contour evaluation metrics. The most qualitative and promising are the BF metric [34] based on the F1-measure and the BJ metric [35], which combines the BF and the Jaccard index. When calculating the BF metrics, the precision Pc and recall Rc are calculated for class C based on the boundary B_{gt}^c of annotation S_{gt}^c for class C and the boundary B_{ps}^c of the predicted area for a given threshold:

$$P^c = \frac{1}{|B_{ps}^c|} \sum_{x \epsilon B_{ps}^c} \left[\left[d(x, B_{gt}^c) < \theta\right]\right] \tag{5}$$

$$R^c = \frac{1}{|B_{gt}^c|} \sum_{x \epsilon B_{gt}^c} \left[\left[d(x, B_{ps}^c) < \theta\right]\right] \tag{6}$$

The measure F_1^c for class C is defined as:

$$BF^c = F_1^c = \frac{2 \cdot P^c \cdot R^c}{P^c + R^c} \tag{7}$$

The BJ metric is based on the Jacquard index, the coefficients for which are calculated taking into account the distance from the annotation border to the predicted region ($TP_{B_{gt}}^c$, FN^c) and the distance from the border of the predicted region to annotation ($TP_{B_{ps}}^c$, FP^c). The total number of true predictions is defined as:

$$TP^c = TP_{B_{gt}}^c + TP_{B_{ps}}^c \tag{8}$$

False predictions are calculated using the formulas:

$$FN^c = |B_{gt}^c| - TP_{B_{gt}}^c \tag{9}$$

$$FP^c = |B_{ps}^c| - TP_{B_{ps}}^c \tag{10}$$

Accuracy over the entire set of images is calculated similarly to the Jaccard index. This metric, in contrast to BF, takes into account the segmentation content beyond the threshold distance θ, under which the boundaries are matched.

There is also an approach to assessing the quality of segmentation contours, when areas close to the contours of class masks are given more weight when calculating accuracy. Also, the calculation of accuracy can be made separately for areas near the contours.

3.4. Implementation of neural networks models

After the programming language and libraries are chosen - Python and the TensorFlow and Keras libraries, it is necessary to select the implementations of the algorithms. Since the development of a neural network, even according to the description, requires a lot of time, including debugging, we decided to use publicly available solutions. A prerequisite is the ability to use the original weights obtained by the authors of the algorithm (can be converted from one framework to another), or weights successfully trained on a set of images of the driving environment. Also we did not find any implementation of UNet with trained weights, but still included it in the study due to the high popularity of UNet architecture.

Taking into account these conditions, five implementations were chosen:

1. The implementation of ENet [36], not the original (not from the authors of the algorithm). Implemented using TensorFlow. The model was trained on CamVid dataset with an input size of 480 × 360 pixels, 12 classes. Contains software tools for training, testing (evaluation) and running the model. The input image size for running the model is unlimited.

2. PSPNet implementation [37], not original. Implemented using TensorFlow. Converted original weights are available, trained on the Cityscapes data set with an input size of 713 × 713 pixels, 19 classes. The input size to run the model in any mode must be at least 713 pixels in width and height.

3. The implementation of DeepLab v3 [38], the original. Implemented using TensorFlow. Includes various configurations of models and various weights. There is no documentation, only examples of use for training, evaluation, output, so we needed a detailed study of the implementation. Has 4 models trained on Cityscapes dataset, 19 classes. The first model uses the MobileNet v2 neural network as a feature extractor and does not have a decoder, which is not suitable for the task. The second and third models use the Xception 65 neural network as a feature extractor, the fourth uses Xception 71. The fourth model's accuracy is 82%, which is 2% higher than the second and third accuracy. There is no significant difference in the size of the models, so it is sensible to choose the fourth model. The training took place on fragments of images of 513 × 513 pixels, but the input size of the images for launching the model is not limited. The training was done sequentially on ImageNet, MS-COCO, a coarse-annotated set of Cityscapes and, finally, a fine-annotated set of Cityscapes.

4. ICNet implementation [39], not original. Implemented using TensorFlow. Converted original weights are available, trained on the Cityscapes data set with an input size of 1025 × 2049 pixels, 19 classes. The model can be configured with and without batch normalization layers. The input size must be at least 97 × 97 pixels in width and height for a model with batch normalization, and at least 161 × 161 for a model without batch normalization.

5. UNet implementation, not original, without trained weights. The architecture implementation is taken from [40]. Software tools for training, testing (evaluation) and running models are implemented by us. The model input was set to 104x512 pixels. First, the model was trained from scratch on sets of Cityscapes with coarse annotations — original, augmented with geometric and elastic transformations, with color augmentation — a total of 185 epochs of 3,658 batches of size 6, i.e. 4 million examples. Then, the resulting model was trained on the original and all augmented Cityscapes sets with fine annotations and Berkeley Deep Drive dataset with augmentation - 100 epochs of 3658 batches of size 6, i.e. just under 2.2 million examples.

Despite the fact that all the implementations, except for UNet, use the same library for machine learning, working with neural network models in them is organized differently, and collecting them into one project does not seem to be a reasonable task. Therefore, to simplify the experiments and possible subsequent embedding, it was decided to use the TensorFlow tool to "freeze" the trained graph (in this case, the neural network) and save it in the protobuf format (protocol buffers, ".pb" file extension). An example of use can be seen in [41, 42]. It is impossible to train such a "frozen" model, however, it allows to abstract from the details of the implementation of a specific model and run it without having any code to build the neural network architecture, since all this is stored in the protobuf file.

Only the input size of the model should be known, the name of the input node where the image should be sent, and the name of the output node where the resulting mask should be taken from. Thus, the model can be run with both Python and C ++. In addition, since the "freezing" cuts off unnecessary operations related to training, the model can increase in productivity.

TensorFlow has another tool for working with trained models - TensorFlow Serving [43], which provides a flexible server architecture for deploying and maintaining models, organizing the exchange of information between models and software applications that use them. This is very useful for embedding models, but for the purposes of this study, it is sufficient to simply convert models to the protobuf format.

All models were "frozen" in such a way that their inputs and outputs had a uniform format. Thus, the input of any model has the form [H, W, 3], that is, it receives a color (three-channel) image of W × H pixels, and the output is [1, H, W, C], where C is the number of classes. Together with the model, the files were saved in json format, which contained meta-information — the names of the input and output nodes, the name of the model's architecture, the size of the images on which the model was trained, the size of the images to which the frozen model was set up, the name of the dataset, on which the model is trained, class names, and some other parameters. To run the model, only the names of the input and output nodes and the input size are needed, but other data simplify the work during the experiments.

The functionality for "freezing" the model in protobuf format was added separately to each of the selected implementations. Thus, it is supposed to work with the model (training, changing the configuration and parameters) in its own software project, and for evaluation and inference it is supposed to use models in the protobuf format, and the code is universal for all models.

4. Experiments

To compare the performance of the models, different input sizes were set, that is, the input images and annotations were scaled to these sizes, after which the resulting predicted masks in the same size were compared with annotations-masks. If the aspect ratio of the image is significantly different from the aspect ratio of the input network size, the image is padded in width or height to get the aspect ratio as the input network size. Also, the prediction execution time was measured without preprocessing. Metrics mIoU and PA (Pixel accuracy) were used.

The main part of the experiments was carried out on the following equipment: Intel Core i7-3770 processor (3.4 GHz × 8), GeForce GTX 1070 graphics card (8 GB), 32 GB of RAM.

Evaluation was performed on the Cityscapes and Berkeley Deep Drive databases in 19 classes. These data sets have the same set of classes.

4.1. ENet

Table 1 shows the performance of the ENet neural network model. The input size corresponding to the one with which the model was trained is shown in bold. It can be seen that the model No. 8 is the best option in terms of speed and accuracy, however, the maximum accuracy of 44% and 66% for mIoU and PA metrics is rather low today. As the size of the input increases, the speed drops, but the accuracy does not increase. It should be noted that a large contribution to the error is made by the fact that the model is trained on the CamVid data set with 12 classes, and the evaluation was performed on datasets with 19 classes. Test datasets were converted to CamVid format, but a complete match is not possible. Some CamVid classes include several classes from the Cityscapes and BDD sets at once. Annotation masks of such classes are combined.

However, the model has a small number of trainable parameters and a very small size; It is possible that training on the Cityscapes and Berkeley Deep Drive databases will significantly improve the accuracy of the neural network.

Model 8, ENet_8 was chosen for further evaluation.

Table 1. ENet performance

Name	Input size	Cityscapes		Berkeley Deep Drive		FPS
		mIoU, %	PA, %	mIoU, %	PA, %	
ENet_1	**360×480**	28.26	39.66	33.94	50.96	41.94
ENet_2	480×480	24.31	31.71	30.03	46.85	34.07
ENet_3	720×960	34.71	46.98	31.48	47.32	11.63
ENet_4	900×1200	32.51	46.17	28.97	45.50	7.33
ENet_5	360×720	43.78	65.50	38.73	59.05	30.85
ENet_6	720×1280	37.70	55.79	33.20	52.08	8.86
ENet_7	720×1440	37.19	55.44	32.55	51.56	7.92
ENet_8	360×640	**44.03**	**65.65**	**39.05**	**59.20**	**34.24**

4.2. PSPNet

Table 2 shows the performance of the PSPNet neural network. It can be seen that the accuracy performance on the basis of Berkeley Deep Drive is lower than that on the basis of Cityscapes, which is expected, since the models were trained only on the basis of Cityscapes. To increase the model's generalizing ability, it is necessary to increase the representativeness of the training dataset, that is, the neural network should be trained on several more datasets besides the Cityscapes, for example, PASCAL, MS-COCO, Berkeley Deep Drive. Model No. 6 is the best option for accuracy on Cityscapes, Model No. 1 is the best option for speed, Model No. 4 can be considered balanced, while it is the best in Berkeley Deep Drive. For further evaluations, models No. 1, 4, 6 were chosen (PSPNet_1, PSPNet_4, PSPNet_6).

Table 2. PSPNet performance

Name	Input size	Cityscapes		Berkeley Deep Drive		FPS
		mIoU, %	PA, %	mIoU, %	PA, %	
PSPNet_1	**713x713**	76.16	92.58	51.76	81.99	3.48
PSPNet_2	720x720	75.02	92.05	49.79	81.38	3.48
PSPNet_3	1024x1024	79.68	93.95	49.62	80.90	1.87
PSPNet_4	713x1426	82.37	95.37	49.42	81.27	1.83
PSPNet_5	720x1440	82.93	95.3	49.11	80.07	1.82
PSPNet_6	1024x2048	**84.15**	**95.78**	42.53	73.81	0.96
PSPNet_7	720x1280	81.79	95.09	49.09	81.37	2.08

4.3. DeepLab v3

Table 3 shows the performance of the DeepLab v3 neural network. When training these models, several datasets were used, but Berkeley Deep Drive was not among them, which is the reason for the lower accuracy on this set.

For further evaluation, models with not too low accuracy and speed were chosen - models No. 2, 6, 8 (DeepLab_2, DeepLab_6, DeepLab_8).

Table 3. DeepLab v3 performance

Name	Input size	Cityscapes		Berkeley Deep Drive		FPS
		mIoU, %	PA, %	mIoU, %	PA, %	
DeepLab_1	256x256	44.98	82.16	32.33	59.68	39.21
DeepLab_2	**513x513**	69.00	91.45	49.25	**73.21**	12.67
DeepLab_3	768x768	74.85	92.71	50.29	**73.48**	6.02
DeepLab_4	1024x1024	79.50	93.88	50.99	72.78	3.46
DeepLab_5	256x512	65.28	88.92	46.18	72.08	25.41
DeepLab_6	512x1024	78.81	93.13	**51.19**	**73.46**	6.84
DeepLab_7	640x1280	81.41	93.95	50.69	72.05	4.42
DeepLab_8	720x1440	**82.68**	**94.32**	49.65	70.59	3.56
DeepLab_9	1024x2048	**84.89**	**94.88**	46.62	67.02	1.77
DeepLab_10	720x1280	81.77	94.08	50.34	71.36	3.93

4.4. ICNet

Table 4 shows the performance of the ICNet neural network. It can be seen that batch normalization improves accuracy and speed to a small extent. As with the other models, the accuracy on the Berkeley Deep Drive data set is lower than on the Cityscapes.

Table 4. ICNet performance

Name	Input size	Cityscapes		Berkeley Deep Drive		FPS
		mIoU, %	PA, %	mIoU, %	PA, %	
ICNet_1	**1025x2049**	87.34	96.50	43.13	78.01	16.84
ICNet_2	720x1280	77.69	93.56	42.03	78.35	32.90
ICNet_3	780x1560	83.49	95.34	**44.12**	**78.02**	25.23
ICNet_4	512x1024	72.11	92.29	37.42	74.24	**54.88**
ICNet_BN_1	1025x2049	**87.73**	**96.57**	43.24	77.45	15.93
ICNet_BN_2	720x1280	78.47	93.83	41.91	78.03	31.48
ICNet_BN_3	780x1560	83.61	95.39	43.91	77.65	23.87
ICNet_BN_4	512x1024	73.15	92.55	37.33	73.81	51.02

One of the best options in terms of accuracy is Model 5 (ICNet_BN_1), while the model's speed is minimal and is approximately 16 frames per second, which, relative to other neural network architectures, is nevertheless a fairly high speed. The fastest model No. 4 (ICNet_4) at a speed of almost 55 frames per second also has relatively high accuracy on Cityscapes, which are sufficient for many applications, but quite low accuracy on the Berkeley Deep Drive dataset. As solutions with the best balance of accuracy-speed, models No. 3 (ICNet_3) and No. 7 (ICNet_BN_3) can be taken. Thus, models 3, 4, 5, 7 (ICNet_3, ICNet_4, ICNet_BN_1, ICNet_BN_3) were taken for further evaluation.

4.5. Best models comparison

Tables 5 and 6 show the performance of the best neural network models. Table 5 contains the accuracy values for the best models. In comparison, it is clear that the PSPNet and DeepLab models have a better generalizing ability, since they obtained the highest accuracy values for the dataset, on which the models were never trained. However, the highest accuracy on the test set of Cityscapes is obtained for the ICNet neural network (with a significant margin), which indicates a high learning ability. UNet has the best PA result on Berkeley Deep Drive dataset among all models. However, the model was trained on this dataset, and the big difference between PA and mIoU results requires further study.

Table 5. Accuracy comparison

Name	Input size	Cityscapes		Berkeley Deep Drive	
		mIoU, %	PA, %	mIoU, %	PA, %
UNet	512x1024	32.97	89.19	31.46	**89.36**
ENet_8	360×640	44.03	65.65	39.05	59.20
PSPNet_1	713x713	76.16	92.58	**51.76**	**81.99**
PSPNet_4	713x1426	82.37	95.37	49.42	**81.27**
PSPNet_6	1024x2048	84.15	**95.78**	42.53	73.81
DeepLab_2	513x513	69.00	91.45	49.25	73.21
DeepLab_6	512x1024	78.81	93.13	**51.19**	73.46
DeepLab_8	720x1440	82.68	94.32	49.65	70.59
ICNet_3	780x1560	83.49	95.34	44.12	78.02

	Name	Input size	Cityscapes		Berkeley Deep Drive	
			mIoU, %	PA, %	mIoU, %	PA, %
0	ICNet_4	512x1024	72.11	92.29	37.42	74.24
1	ICNet_BN_1	1025x2049	**87.73**	**96.57**	43.24	77.45
2	ICNet_BN_3	780x1560	83.61	95.39	43.91	77.65

Table 6 shows the frame rate values and sizes for the models measured on various sets of equipment:

– Set 1: CPU Intel Core i7-3770 (3.4 GHz × 8), GPU GeForce GTX 1070 (8 GB), 32 GB RAM.

– Set 2: CPU Intel Core i7-6700 (3.4 GHz × 8), GPU GeForce GTX 1080 Ti (11 GB), 16 GB RAM.

PSPNet and DeepLab v3 models have a higher generalizing ability than ENet and ICNet, however, their operation speed is much lower, and the model size is several times larger. ICNet with sufficiently high accuracy has a high speed, and the size of its model allows it to be used on equipment with limited resources. It allows to perform image segmentation in almost real time even on medium power equipment (set 1), which is rarely seen among neural networks for segmentation at the present time. UNet and ENet with a sufficiently high FPS have low accuracy. We assume that it is possible to increase their accuracy by additional training, so that they can be applied in some tasks.

Table 6. Speed and size comparison

	Name	Size, MB	FPS	
			Set 1	Set 2
	UNet	139	31.42	47.16
	ENet_8	1.8	34.24	53.90
	PSPNet_1	264	3.48	6.51
	PSPNet_4	264	1.83	3.50
	PSPNet_6	264	0.96	1.79
	DeepLab_2	168	12.67	22.68
	DeepLab_6	168	6.84	11.97
	DeepLab_8	168	3.56	6.56
	ICNet_3	27	25.23	39.26
0	ICNet_4	27	54.88	75.99
1	ICNet_BN_1	27	15.93	23.36
2	ICNet_BN_3	27	23.87	37.35

4.6. Analysis of accuracy results by category

Table 7 presents the IoU accuracy by category. Figure 8 shows the histograms of the distribution of images by category.

It can be noted that in most classes a certain tendency is maintained regardless of the model, this is due to the frequency of the appearance of the class in dataset. Frequently found classes "road", "building", "car", "vegetation", "person" give high accuracy for any model, while the rare classes "wall", "truck", "train", "pole" have low accuracy regardless of model. It should be noted that the PSPNet and DeepLab B3 models are less affected by the class imbalance, and the UNet model seems to be able to learn only a few classes, however, with very high quality.

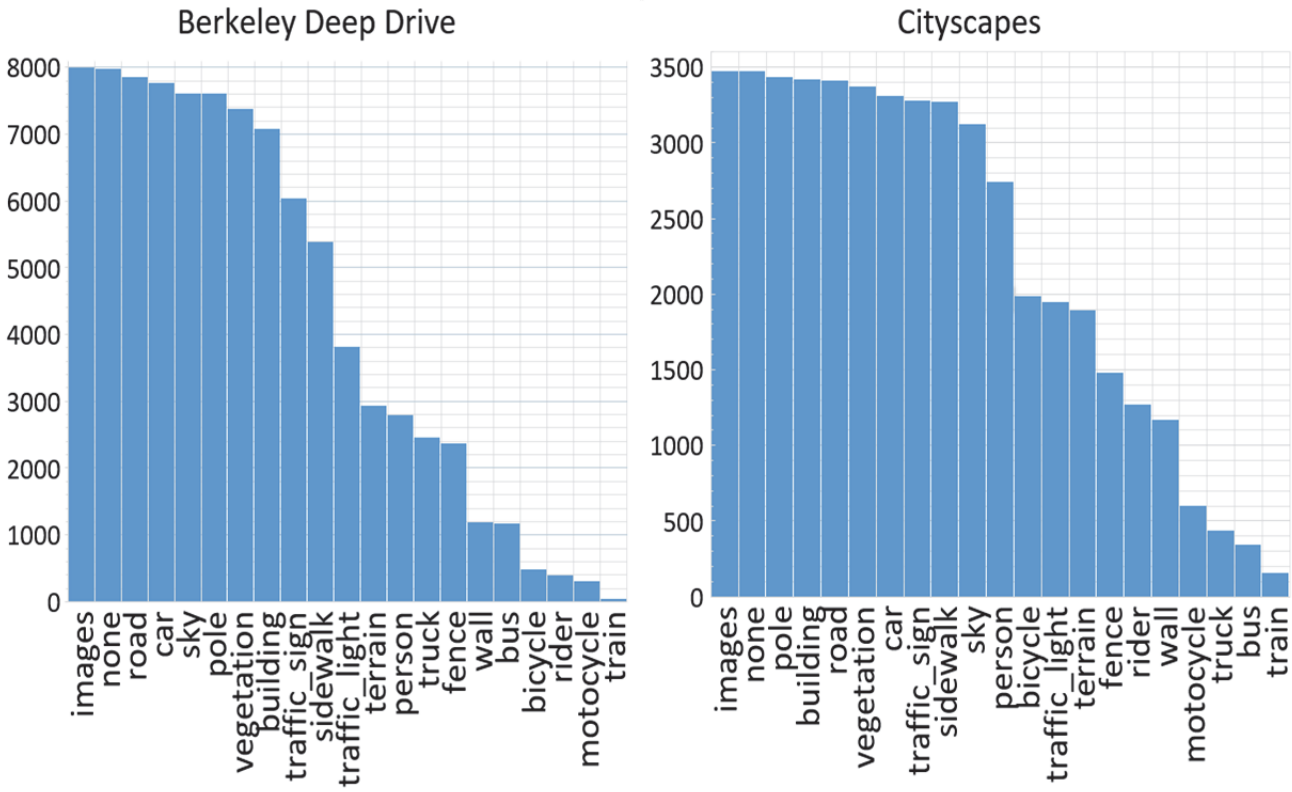


Figure 8 – Histograms of the distribution of images by category

Table 7. IoU accuracy by category

	Class	IoU, %			
		UNet	PSPNet	DeepLab	ICNet
	road	97.34	88.87	86.24	76.39
	sidewalk	73.27	38.87	41.58	48.78
	building	93.71	84.80	87.68	83.91
	wall	00.00	15.24	15.38	14.40
	fence	00.00	46.63	50.52	15.49
	pole	00.00	26.87	40.12	29.55
	traffic light	00.00	42.17	26.95	34.10
	traffic sign	00.00	33.63	44.80	23.63
	vegetation	91.88	90.38	87.67	84.84
0	terrain	48.76	34.62	18.23	38.35
1	sky	96.84	77.25	38.35	80.33

	Class	IoU, %			
		UNet	PSPNet	DeepLab	ICNet
2	person	00.00	69.86	65.16	58.26
3	rider	00.00	60.21	42.59	27.22
4	car	95.90	95.13	91.85	93.85
5	truck	00.00	24.41	28.14	11.24
6	bus	00.00	21.82	50.71	24.12
7	train	00.00	00.23	00.00	00.00
8	motorcycle	00.00	67.10	46.44	32.71
9	bicycle	00.00	65.41	73.48	42.40
	mIoU	31.46	51.76	49.26	43.14

Figure 9, 10 shows examples of predictions.

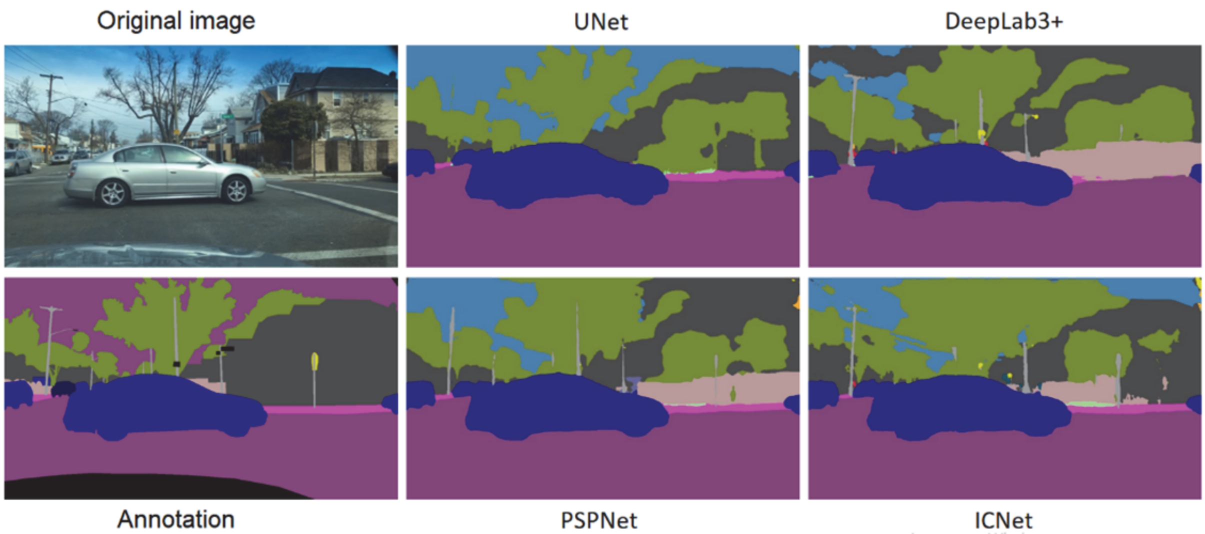


Figure 9 – UNet ignores small objects: signs, poles. DeepLab v3, PSPNet and ICNet correctly predict fence, although it is missing from the annotation

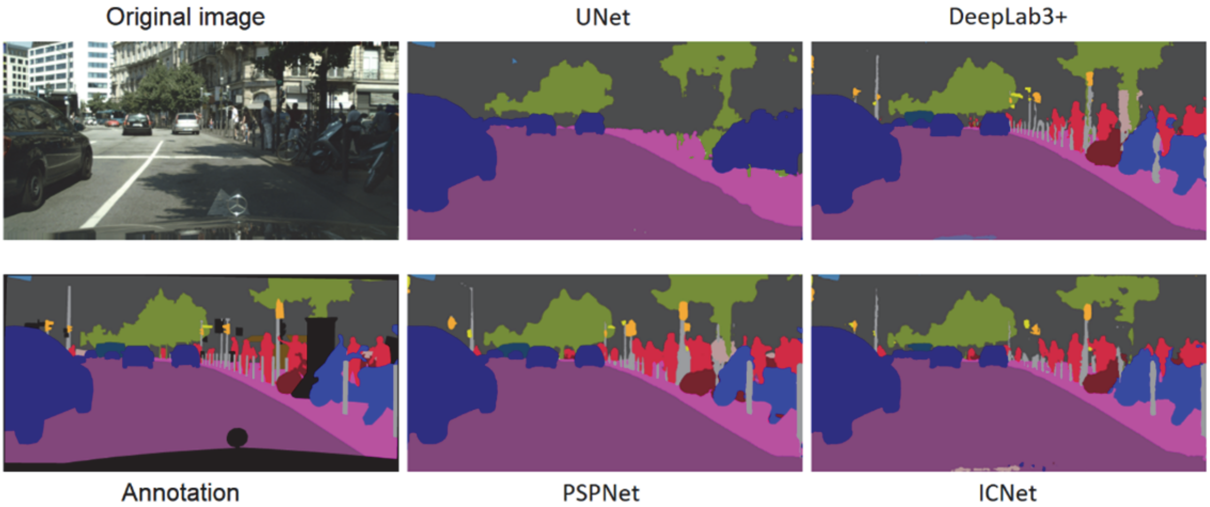


Figure 10 – UNet ignores small categories including people, the most visually accurate result belongs to DeepLab v3 model

5. Conclusion

In the paper, the current level of image segmentation methods is studied on the example of the task of driving environment segmentation. A review of segmentation methods was performed, which showed that the most promising group of segmentation methods for today is neural network algorithms. A review of neural networks for segmentation, libraries and frameworks for working with neural networks and datasets for training and testing was conducted. This paper also addresses the problem of lack of experimental studies showing the performance of neural segmentation with respect to each other. This complicates the selection of a neural network for a certain task and does not allow to see the real situation in the field of neural network segmentation methods.

In the course of the work, several implementations of popular neural network architectures (models) were chosen - Unet, Enet, PSPNet, DeepLab v3, ICNet. All of these models were brought to a uniform format for the purity of the experiment and evaluated with different parameters.

The best balance of speed-accuracy has the architecture ICNet. While for other architectures, performance measures are a bargain between speed and segmentation quality, ICNet allows both high accuracy and high speed. The advantages of this model are also its small size and the ability to use on high-resolution images. However, this architecture is not very high generalizing ability, so the dataset for training should be as representative as possible.

Neural networks DeepLab v3 and PSPNet have a higher generalizing ability, but these are “heavy” architectures - they require a sufficiently large amount of memory and have a low speed. However, in tasks that require high accuracy in terms of incomplete representativeness of training data, these models will show themselves better than ICNet. It should also be noted a feature of the entire DeepLab line of architectures, including the third version: high-quality selection of the contours of objects, while full-convolutional networks often get a fuzzy, smoothed silhouette.

The ENet neural network was developed as “light”, i.e. lightweight, not occupying a large amount of memory neural network with high speed. This architecture can actually be run on weak hardware and at the same time have relatively high speed, but the accuracy of ENet is much lower than the other models considered. However, this network can be successfully applied in tasks that require segmentation of large areas that do not have small details (roadway, sidewalks, sky, facades, trees), if there is a limit on resources.

UNet showed inability to learn many classes, but also showed high accuracy in a small number of learned classes. This model requires a fairly large amount of memory, but also has a fairly high speed. UNet is suitable for segmentation problems of a small number of classes with simple smooth contours that require high accuracy.

In further work, it is advisable to train the considered neural networks on a group of datasets and conduct an evaluation again. There is a possibility that ICNet will show accuracy on the same level with PSPNet DeepLab v3. It also makes sense to conduct an evaluation with contour metrics.

References

1. Dilpreet Kaur, Yadwinder Kaur: Various Image Segmentation Techniques: A Review. IJCSMC, 3(5), 809–814 (2014).

2. Ultrasound Nerve Segmentation. Identify nerve structures in ultrasound images of the neck – Kaggle: Your Home for Data Science, https://www.kaggle.com/c/ultrasound-nerve-segmentation, last accessed 2019/06/20.
3. Carvana Image Masking Challenge. Automatically identify the boundaries of the car in an image – Kaggle: Your Home for Data Science, https://www.kaggle.com/c/carvana-image-masking-challenge, last accessed 2019/06/20.
4. KONICA MINOLTA. Pathological Image Segmentation Challenge – topcoder: leading in crowdsourcing, https://www.topcoder.com/KonicaMinoltaChallenge, last accessed 2019/06/20.
5. Robotic Instrument Segmentation Sub-Challenge (Part of the Endoscopic Vision Challenge) – Grand-challenges: Endoscopic Vision Challenge, https://endovissub2017-roboticinstrumentsegmentation.grand-challenge.org, last accessed 2019/06/20.
6. CVPR 2018 WAD Video Segmentation Challenge. Can you segment each objects within image frames captured by vehicles? – Kaggle: Your Home for Data Science, https://www.kaggle.com/c/cvpr-2018-autonomous-driving/, last accessed 2019/06/20.
7. TGS Salt Identification Challenge. Segment salt deposits beneath the Earth's surface – Kaggle: Your Home for Data Science, https://www.kaggle.com/c/tgs-salt-identification-challenge, last accessed 2019/06/20.
8. J. Long, E. Shelhamer, T. Darrell: Fully convolutional networks for semantic segmentation, In: Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 3431-3440 (2015).
9. Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS, Vol.9351, 234--241 (2015).
10. Chen L.-Ch., Papandreou G., Kokkinos I., Murphy K., Yuille A.L. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. ICLR (2015), http://arxiv.org/abs/1412.7062, last accessed 2019/06/20.
11. George Papandreou, Liang-Chieh Chen, Kevin Murphy, Alan L Yuille: Weakly- and Semi-Supervised Learning of a DCNN for Semantic Image Segmentation. ICCV (2015).
12. DeepLab System Public Repository, https://bitbucket.org/deeplab/deeplab-public, last accessed 2019/06/20.
13. Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L Yuille: DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs (2016), http://arxiv.org/abs/1606.00915, last accessed 2019/06/20.
14. Vijay Badrinarayanan, Alex Kendall and Roberto Cipolla: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(12), 2017.
15. SegNet project page, http://mi.eng.cam.ac.uk/projects/segnet/#publication, last accessed 2019/06/20.
16. Adam Paszke, Abhishek Chaurasia, Sangpil Kim, Eugenio Culurciello: ENet: A Deep Neural Network Architecture forReal-Time Semantic Segmentation, https://arxiv.org/pdf/1606.02147.pdf, last accessed 2019/06/20.
17. Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, Jiaya Jia: Pyramid Scene Parsing Network. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2881-2890 (2017).
18. Cityscapes Dataset – Semantic Understanding of Urban Street Scenes, https://www.cityscapes-dataset.com, last accessed 2019/06/20.
19. Liang-Chieh Chen, George Papandreou, Florian Schroff, Hartwig Adam. Rethinking Atrous Convolution for Semantic Image Segmentation, https://arxiv.org/pdf/1706.05587.pdf, last accessed 2019/06/20.
20. Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: The European Conference on Computer Vision (ECCV), 801-818 (2018).
21. Hengshuang Zhao, Xiaojuan Qi, Xiaoyong Shen, Jianping Shi, Jiaya Jia: ICNet for Real-Time Semantic Segmentationon High-Resolution Images, arxiv.org/pdf/1706.05587.pdf, last accessed 2019/06/20.
22. R.B. Palm. DeepLearningTolbox – Rasmus Berg Palm personal page on GitHub.com, https://github.com/rasmusbergpalm/DeepLearnToolbox, last accessed 2019/06/20.
23. Torch: a scientific computing framework for lua, http://torch.ch/, last accessed 2019/06/20.
24. PyTorch web-site, https://pytorch.org, last accessed 2019/06/20.

25. E.S. Yangqing Jia. Caffe – Berkeley AI Research Lab, http://caffe.berkeleyvision.org/, last accessed 2019/06/20.
26. TensorFlow: Google's Machine Intelligence research organization, https://www.tensorflow.org/, last accessed 2019/06/20.
27. Keras: The Python Deep Learning library, https://keras.io/, last accessed 2019/06/20.
28. Jeff Hale. Which Deep Learning Framework is Growing Fastest? (2019) – Towards Data Science. Sharing concepts, ideas and codes, https://towardsdatascience.com/which-deep-learning-framework-is-growing-fastest-3f77f14aa318, last accessed 2019/06/20.
29. Berkeley DeepDrive, https://bdd-data.berkeley.edu/, last accessed 2019/06/20.
30. SYNTHIA Dataset, http://synthia-dataset.net/, last accessed 2019/06/20.
31. Motion-based Segmentation and Recognition Dataset, University of Cambridge, http://mi.eng.cam.ac.uk/research/projects/VideoRec/CamVid/, last accessed 2019/06/20.
32. COCO – Common Objects in Context, http://cocodataset.org/#home, last accessed 2019/06/20.
33. Visual Object Classes Challenge 2012 (VOC2012), http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html, last accessed 2019/06/20.
34. G. Csurka, D. Larlus, F. Perronnin, F. Meylan: What is a good evaluation measure for semantic segmentation?. In: BMVC, 27, 32.1-32.11 (2013).
35. Eduardo Fernandez-Moral, Renato Martins, Denis Wolf, Patrick Rives: A new metric for evaluating semantic segmentation: leveraging global and contour accuracy. Workshop on Planning, Perception and Navigation for Intelligent Vehicles, PPNIV17, Vancouver, Canada (2017).
36. TensorFlow implementation of ENet – Kwotsin GitHub page, https://github.com/kwotsin/TensorFlow-ENet, last accessed 2019/06/20.
37. An implementation of PSPNet in tensorflow – Hellochick GitHub page, https://github.com/hellochick/PSPNet-tensorflow, last accessed 2019/06/20.
38. DeepLab: Deep Labelling for Semantic Image Segmentation – TensorFlow Research Models GitHub page, https://github.com/tensorflow/models/tree/master/research/deeplab, last accessed 2019/06/20.
39. TensorFlow-based implementation of "ICNet for Real-Time Semantic Segmentation on High-Resolution Images" – Hellochick GitHub page, https://github.com/hellochick/ICNet-tensorflow, last accessed 2019/06/20.
40. Kaggle Carvana Image Masking Challenge solution with Keras tensorflow – Petros Giannakopoulos GitHub page, https://github.com/petrosgk/Kaggle-Carvana-Image-Masking-Challenge, last accessed 2019/06/20.
41. A Tool Developer's Guide to TensorFlow Model Files – TensorFlow website, https://www.tensorflow.org/guide/extend/model_files, last accessed 2019/06/20.
42. Morgan. TensorFlow: How to freeze a model and serve it with a python API – MetaFlow.ai blog, https://blog.metaflow.fr/tensorflow-how-to-freeze-a-model-and-serve-it-with-a-python-api-d4f3596b3adc, last accessed 2019/06/20.
43. Serving Models – TensorFlow website, https://www.tensorflow.org/tfx/guide/serving, last accessed 2019/06/20.

L.Yu. Vorochaeva[1], A.S. Yatsun[1], S.I. Savin[2], A.V. Repkin[1]

GAITS OF A SEARCH TWO-LINK CRAWLING ROBOT

[1] *Southwest State University, Kursk, Russia*
[2] *Innopolis University, Innopolis, Russia*
mila180888@yandex.ru

Abstract

The present paper deals with the consideration of a gait of a two-link crawling robot designed to inspect buildings on a horizontal rough surface. The interaction with the latter is carried out with the help of three supporting elements with a controlled friction coefficient, which allows the robot links to be fixed periodically on the surface. A classification of possible gaits of the device is given, two basic types are considered in detail: controlled and combined. For each gait, the sequence of stages, the conditions for the beginning and end of each of them are determined, and the diagrams of angular velocities and friction coefficients in the supports are constructed.

Keywords: two-link crawling robot, friction coefficient control, gaits, controlled gaits, combined gaits.

Acknowledgments

The work was carried out within the framework of the Presidential Grant MK-200.2019.1.

Introduction

Nowadays, with increasing frequency, in the event of natural and man-inflicted disasters, such as earthquakes, fires, floods, acts of terrorism, robotic means are used to carry out rescue, search and reconnaissance operations in completely or partially destroyed buildings. As the main areas of their application, searching for people under the rubble, detecting fires, analyzing the status of floors for safe rescue operations, monitoring the state of explosive substances in buildings, searching for passages to certain rooms, determining the presence and concentration of toxic substances in the air, etc. can be distinguished.

Currently, scientists from different countries are conducting research in the field of creating snake-like robots for the above given purposes. Indeed, articles [1-3] deal with issues related to the development of the gaits of snake-like robots, as well as the issues of movement on surfaces with obstacles. For example, in [1], the development of such an algorithm (gait) for the movement of a crawling snake-like robot over an obstructed surface is described, so that the trajectory of the robot movement is smooth. In [2], a method for designing two gaits of a snake-like robot moving through pipes and over rough terrain is presented, in which the required configuration of a robot is formed by connecting the curve segments. The spatial movement of a snake robot in the presence of obstacles is considered in [3]; overcoming obstacles is provided by vertical and horizontal deformations of the robot's body in the presence of anisotropic friction between its links and the surface. Works [4-6] present issues on the motion of snake-like robots in limited spaces, including pipes and along cables.

The choice of snake-like robots as reconnaissance robots is conditioned owing to their higher maneuverability and terrain crossing capacity over a non-deterministic surface in the presence of obstacles, as well as the ability to penetrate through narrow restricted spaces (slots, pipes, ventilation shafts), which are most often observed after natural and man-inflicted disasters.

In this paper, it was decided to focus on the consideration of a two-link crawling robot, which moves along a horizontal rough surface. The design feature of the device is the presence of three supporting elements with a controlled coefficient of friction between them and the surface. This allows the robot to implement various types of gaits, which is the subject of this study.

Description of the crawling robot

We will consider a two-link crawling robot moving along the horizontal plane *Oxy*, the links of which 1 and 2 are interconnected by a rotating hinge O_2 (Fig. 1).

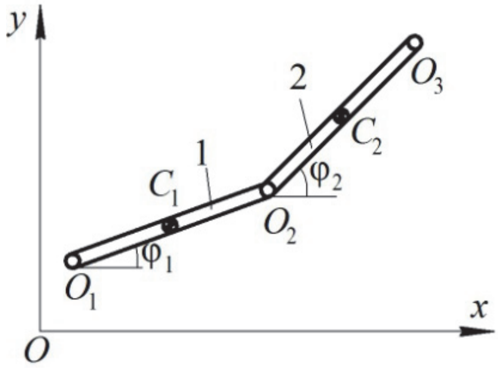


Figure 1 − Computational scheme of the two-link crawling robot

The links of the robot will be envisaged in the form of absolutely solid rods with the lengths l_i, $i = 1, 2$, and the masses $m_i = m$, the centers of which coincide with the centers of symmetry of the rods - the points C_i. At the ends of the links - in the points O_i, $i = 1-3$ - there are supports with controlled friction, similar to those described in the works [7-9]. Supports of this type allow for the increase of the coefficient of friction between the support element of the robot and the surface (the value of the friction coefficient is denoted as $f_{\max}$) to fix the link on the latter and reduce it for the link movement (the value of the friction coefficient $f_{\min}$) owing to the change of supporting elements. The position of the robot on the plane is described by the vector of generalized coordinates

$$\mathbf{q} = (x_{C1}, y_{C1}, \varphi_1, \varphi_2)^{\mathrm{T}}, \tag{1}$$

where x_{C1}, y_{C1} − stand for the mass center coordinates of link 1, φ_1, φ_2 − link 1 and 2 rotation angles relative to the axis *Ox* counterclock-wise.

In the hinge O_2 a rotary motion drive is installed, ensuring the rotation of one link of the two-link mechanism relative to another.

Classification of gaits and paths of motion of the crawling robot

The gaits implemented by this robot can be classified depending on the presence or absence of the friction control in the supports, as shown in Fig. 2, into three types: inertial, combined and controlled.

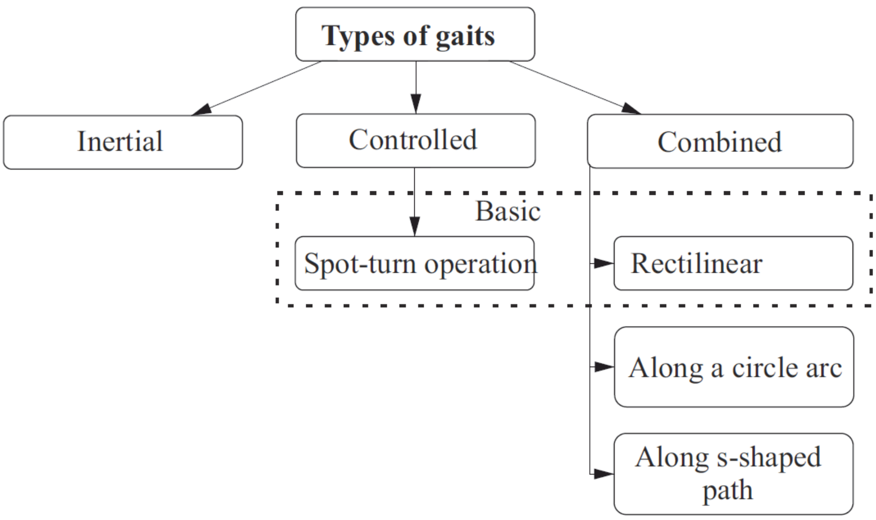
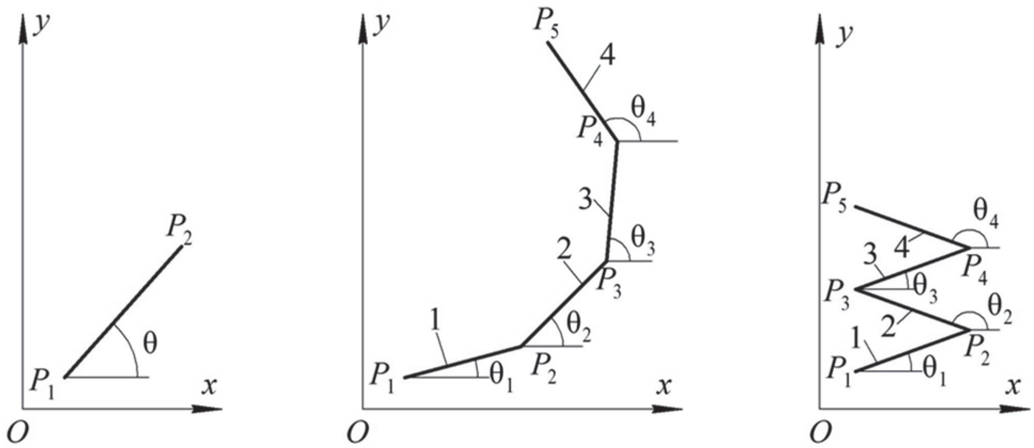


Figure 2 − Gait classification of the two-link crawling robot

Inertial gaits are considered in detail in the works [10-12] and are observed in the absence of friction control in the supporting elements, in this work we will not study them. Combined gaits are possible with a combination of stages, during which inertial movements occur, with stages that are possible while controlling friction in supports. Controlled gaits are carried out only if there is friction control in the supports. We assume that during inertial movements the coefficient of friction of the robot's support on the surface is small − $f_{\min}$ −

and allows the support to slide, and the friction control leads to the increase of the friction coefficient up to the value $f_{\max}$ and fixing of the support on the surface.

Among the gaits of the robot it is possible to single out two basic ones: rectilinear motion and spot-turn operation. Using these two gaits, it is possible to realize the movement of a robot along various complex trajectories by dividing them into components. The most common complex trajectories are the arc of a circle and the s-shaped trajectory, therefore, for these types of trajectories, the corresponding combined gaits are considered. By the trajectory of the robot, we will understand the trajectory of the point O_2 $y_{O2}(x_{O2})$. The most common types of trajectories along which the robot must move are shown in Fig. 3.



a b c

Figure 3 – Types of trajectories (paths): a – straight, b – arc of a circle, c – s-shaped trajectory, 1-4 – trajectory sections

The trajectory in the form of a straight line is shown in Fig. 3, and is described by the equation

$$y = kx + b, \tag{2}$$

where k and b − denote the coefficients, and the angle of inclination of the straight line to the axis Ox is equal to

$$\theta = \operatorname{arctg}(k). \tag{3}$$

The arc of the circle will be described by a piecewise continuous function.

$$y = \begin{cases} k_1 x + b_1,\ x \in [x_{P1}, x_{P2}], \\ k_2 x + b_2,\ x \in (x_{P2}, x_{P3}], \\ \ldots \\ k_n x + b_n,\ x \in (x_{Pn}, x_{Pn+1}]. \end{cases} \tag{4}$$

We will assume that the given trajectory is formed by the n-th number of straight line segments of the same length $l_1 = l_2 = \ldots = l_n = l$, the relative angles between which are equal to each other $\theta_{21} = \theta_{32} = \ldots = \theta_{n,n-1}$, where $\theta_{n,n-1} = \theta_n - \theta_{n-1}$, θ_n, θ_{n-1} − the angles of inclination of the corresponding straight sections to the axis Ox.

The s-shaped trajectory can be described by a piecewise continuous function

$$y = \begin{cases} k_1 x + b_1,\ x \in [x_{P1}, x_{P2}], \\ k_2 x + b_2,\ x \in (x_{P2}, x_{P3}], \\ \ldots \\ k_n x + b_n,\ x \in (x_{Pn}, x_{Pn+1}]. \end{cases} \tag{5}$$

The lengths of straight-line sections are equal $l_1 = l_2 = \ldots = l_n = l$, and its main distinction from the circular arc is that the odd straight-line sections are parallel to each other, whereas the even ones are parallel to each other, i.e.

$$\theta_1 = \theta_3 = \theta_5 = \dots, \tag{6}$$

$$\theta_2 = \theta_4 = \theta_6 = \dots \tag{7}$$

Gaits of the crawling robot

Let us consider the gaits of a two-link crawling robot.

Rectilinear motion

The rectilinear motion consists of three stages, shown in Fig. 4. In this and the subsequent figures on the left, the initial positions of the stages are shown, and on the right, the final fixed supports are painted black, and the fixed link is gray, the dashed line corresponds to the specified trajectory of movement. Let in the initial position the links of the robot be located relative to the trajectory of motion at angles $\varphi_{10} = \varphi_{10}^{(n)}$, $\varphi_{20} = \varphi_{20}^{(n)}$, where φ_{10}, φ_{20} − the angles of links 1 and 2 of the robot relative to the trajectory, $-\varphi_{10}^{(n)} = \varphi_{20}^{(n)} = \varphi_0^{(n)}$ − some given angle value at the moment of the motion start along a given trajectory, the friction coefficients in all the supports are equal to $f_{\min}$.

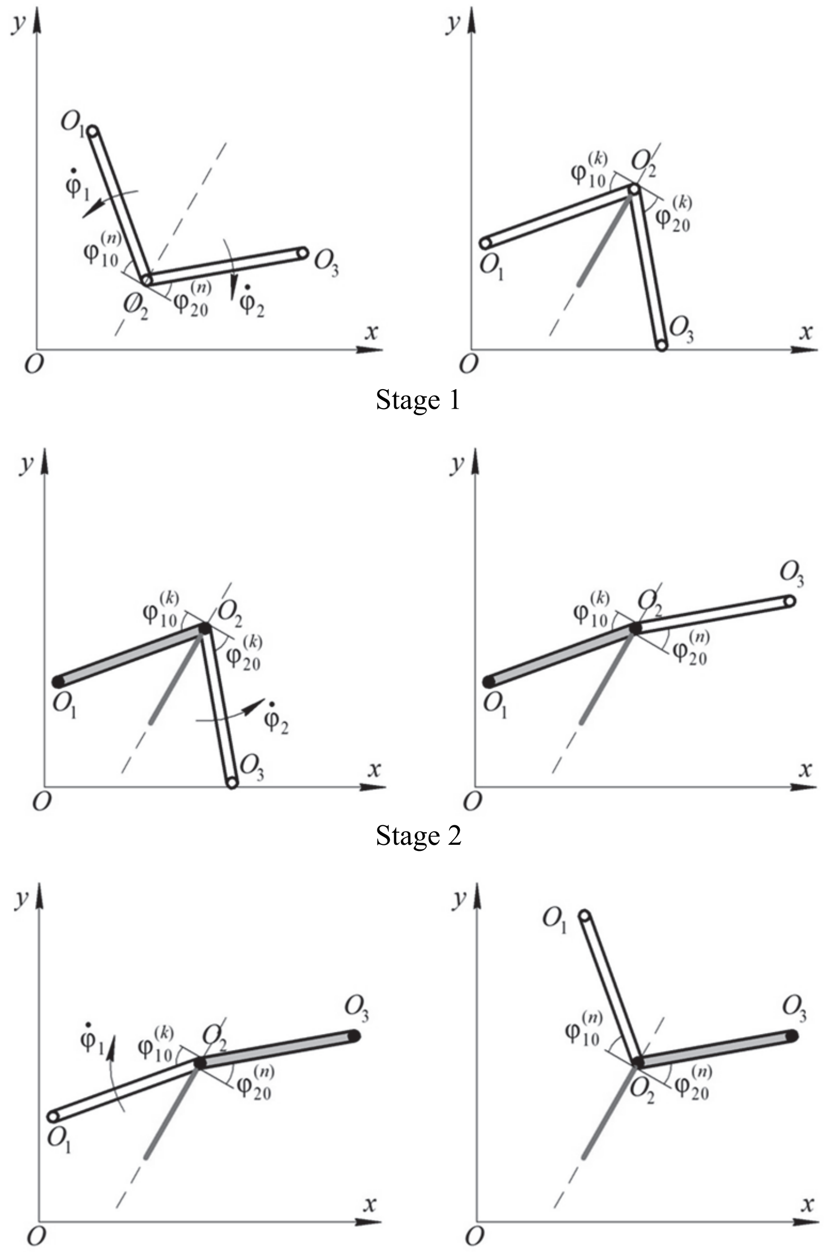
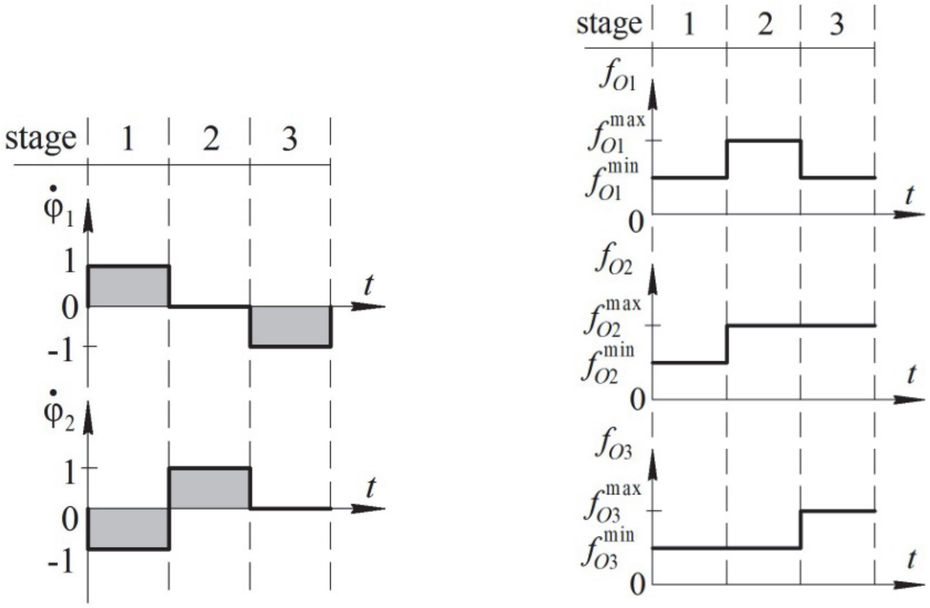


Figure 4 − Pictograms of the stages of movement of a two-link robot with rectilinear motion, − the path covered

During the first stage (inertial), links 1 and 2 rotate in opposite directions (in Fig. 4 it is shown that link 1 rotates counterclockwise, and link 2 rotates clockwise), while the point O_2 moves along the motion path. This stage is terminated when the links reach the angles $\varphi_{10} = \varphi_{10}^{(k)}$, $\varphi_{20} = \varphi_{20}^{(k)}$, where $\varphi_{10}^{(k)} = -\varphi_{20}^{(k)} = \varphi_0^{(k)}$ − some given angle value corresponding to the completion of the first stage. The second stage (controlled) consists in rotation of link 2 with respect to fixed link 1 until it reaches the angle $\varphi_{20} = \varphi_{20}^{(n)}$, for fixing on the surface of link 1 the coefficient of friction in the supports O_1 and O_2 is increased. The third stage is also controlled and consists in the rotation of link 1 relative to the fixed link 2 until the condition $\varphi_{10} = \varphi_{10}^{(n)}$ is met, for this the coefficient of friction in the support O_1 decreases and in the support O_3 increases. At the time of termination of the third stage, the robot is in the same position relative to the trajectory of movement, in which it was at the beginning of the first stage. In fig. 5, diagrams of angular velocities of the links are shown, where ± 1 means the presence of angular velocity and its direction ("+" counterclockwise, "-" clockwise), and 0 is the stationary state of the link. In fig. 5, b diagrams of changes in the friction coefficients in the supports are given.



a b

Figure 5 − Diagrams: a − $\dot{\varphi}_1(t)$, $\dot{\varphi}_2(t)$, b − $f_{Oi}(t)$ in the rectilinear motion

For visual clarity of the description of the rectilinear motion, all its main characteristics are summarized in Table. 1, the initial gait conditions are: $\varphi_{10} = \varphi_{10}^{(n)} = -\varphi_0^{(n)}, \varphi_{20} = \varphi_{20}^{(n)} = \varphi_0^{(n)}$.

Table 1. Characteristics of the rectilinear motion

Stage	**Type of motion**	**Generalized coordinates**	**Termination conditions**	**Fixed link**	**Friction coefficients**		
					f_{O1}	f_{O2}	f_{O3}
1	inertial	$x_{C1}, y_{C1}, \varphi_1, \varphi_2$	$\varphi_{10} = \varphi_{10}^{(k)} = \varphi_0^{(k)}$ $\varphi_{20} = \varphi_{20}^{(k)} = -\varphi_0^{(k)}$	-	min	min	min
2	controlled	φ_2	$\varphi_{20} = \varphi_{20}^{(n)} = \varphi_0^{(n)}$	1	max	max	min
3	controlled	φ_1	$\varphi_{10} = \varphi_{10}^{(n)} = -\varphi_0^{(n)}$	2	min	max	max

Spot-turn operation

Let us proceed to the consideration of the controlled gait of the robot - spot-turn operation. This movement is necessary if there is a change in the direction of movement of the robot, for example, from trajectory 1 to trajectory 2, as shown in Fig. 6.

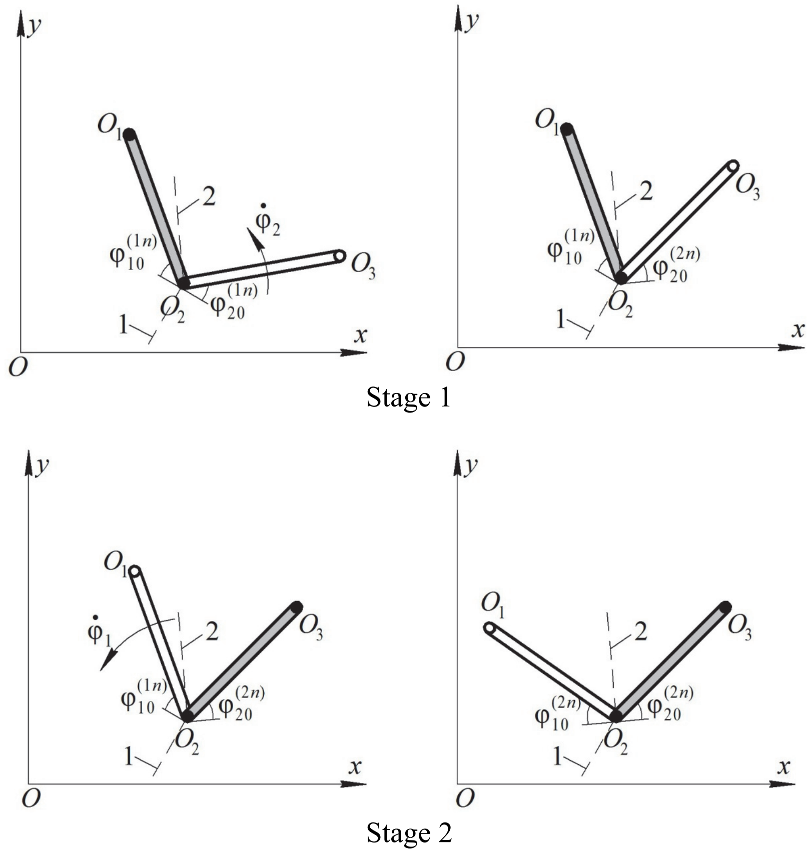
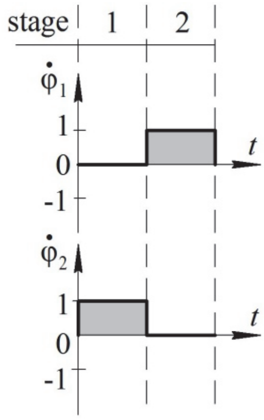
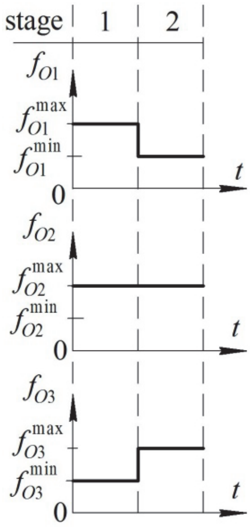


Figure 6 − Pictograms of the stages of movement of a two-link robot while spot-turn operation

The gait consists of two stages, in the initial position the robot is located relative to the trajectory 1 at the angles $\varphi_{10}^{(1)} = \varphi_{10}^{(1n)} = -\varphi_0^{(n)}$, $\varphi_{20}^{(1)} = \varphi_{20}^{(1n)} = \varphi_0^{(n)}$ (Hereinafter, the superscript in brackets corresponds to the number of the straight line segment). During the first stage, link 1 is fixed on the surface, and link 2 is rotated relative to it clockwise or counterclockwise, depending on the relative position of the trajectories (in the example shown in Fig. 6, it is turned counterclockwise) until the condition $\varphi_{20}^{(2)} = \varphi_{20}^{(2n)} = \varphi_0^{(n)}$ is met. Following that, during the second stage, link 2 is fixed on the surface, and link 1 is rotated relative to it in the considered case also counterclockwise until the condition $\varphi_{10}^{(2)} = \varphi_{10}^{(2n)} = -\varphi_0^{(n)}$ is met, that is illustrated in the diagrams of Fig. 7.



a



b

Figure 7 − Diagrams: a − $\dot{\varphi}_1(t)$, $\dot{\varphi}_2(t)$, b − $f_{Oi}(t)$ while spot-turn operation

Characteristics of the spot rotation of the robot are given in Tab. 2, the initial conditions of the gait are: $\varphi_{10}^{(1)} = \varphi_{10}^{(1n)} = -\varphi_0^{(n)}$, $\varphi_{20}^{(1)} = \varphi_{20}^{(1n)} = \varphi_0^{(n)}$.

Table 2. Characteristics of the spot-turn operation

Stage	**Type of motion**	**Generalized coordinates**	**Termination condition**	**Fixed link**	**Friction coefficients**		
					f_{O1}	f_{O2}	f_{O3}
1	controlled	φ_2	$\varphi_{20}^{(2)} = \varphi_{20}^{(2n)} = \varphi_0^{(n)}$	1	max	max	min
2	controlled	φ_1	$\varphi_{10}^{(2)} = \varphi_{10}^{(2n)} = -\varphi_0^{(n)}$	2	min	max	max

Motion along the arc of a circle

As mentioned above, the use of two basic gaits allows for the realization of the movement of the robot along different trajectories, the most common among them are movement along a circle and along an s-shaped curve. Each of these gaits is a sequential combination of rectilinear movement with the robot's spot-turn operation in order to provide movement to the next rectilinear segment. Motion along a circle arc can be represented as consisting of three stages, diagrams of changes in angular velocities and friction coefficients on which correspond to diagrams in straight-line motion, shown in Fig. 5. The characteristics of this movement are given in table. 3. This gait can be used to overcome turns by the robot.

Table 3. Characteristics of motion along the circle arc

Stage	**Type of motion**	**Generalized coordinates**	**Termination condition**	**Fixed link**	**Friction coefficients**		
					f_{O1}	f_{O2}	f_{O3}
1	inertial	x_{C1}, y_{C1}, φ_1, φ_2	$\varphi_{10} = \varphi_{10}^{(1k)} = \varphi_0^{(k)}$ $\varphi_{20} = \varphi_{20}^{(1k)} = -\varphi_0^{(k)}$	-	min	min	min
2	controlled	φ_2	$\varphi_{20} = \varphi_{20}^{(2n)} = \varphi_0^{(n)}$	1	max	max	min
3	controlled	φ_1	$\varphi_{10} = \varphi_{10}^{(2n)} = -\varphi_0^{(n)}$	2	min	max	max

Movement on the s-shaped trajectory

The movement along the s-shaped trajectory is formed by six stages (Fig. 8, Tab. 4). The first three of them describe how the robot overcomes the first rectilinear section of the trajectory and its rotation in place for positioning on the second section of the trajectory; In the future, the stages of movement are repeated, because the odd trajectories are parallel to each other, and the even ones to each other. This gait can be used for a detailed survey of a certain area of limited space, for example, a corridor; in this case, each section of the s-shaped trajectory allows the robot to move from one corridor wall to another, inspecting the area. Alternatively, it is possible to bypass obstacles encountered on the way of the robot.

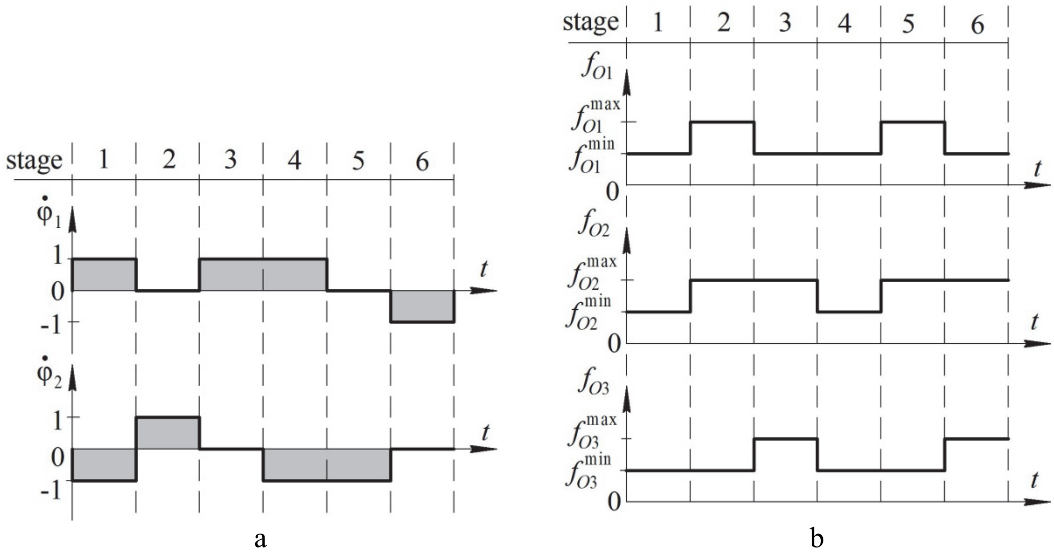


Figure 8 − Diagrams: a − $\dot{\varphi}_1(t)$, $\dot{\varphi}_2(t)$, b − $f_{Oi}(t)$ while motion along the s-shaped trajectory

Table 4. Characteristics of movement along the s-shaped trajectory

Stage	**Type of motion**	**Generalized coordinates**	**Termination condition**	**Fixed link**	**Friction coefficients**		
					f_{O1}	f_{O2}	f_{O3}
1	inertial	$x_{C1}, y_{C1}, \varphi_1, \varphi_2$	$\varphi_{10} = \varphi_{10}^{(1k)} = \varphi_0^{(k)}$ $\varphi_{20} = \varphi_{20}^{(1k)} = -\varphi_0^{(k)}$	-	min	min	min
2	controlled	φ_2	$\varphi_{20} = \varphi_{20}^{(2n)} = \varphi_0^{(n)}$	1	max	max	min
3	controlled	φ_1	$\varphi_{10} = \varphi_{10}^{(2n)} = -\varphi_0^{(n)}$	2	min	max	max
4	inertial	$x_{C1}, y_{C1}, \varphi_1, \varphi_2$	$\varphi_{10} = \varphi_{10}^{(2k)} = \varphi_0^{(k)}$ $\varphi_{20} = \varphi_{20}^{(2k)} = -\varphi_0^{(k)}$	-	min	min	min
5	controlled	φ_2	$\varphi_{20} = \varphi_{20}^{(3n)} = \varphi_0^{(n)}$	1	max	max	min
6	controlled	φ_1	$\varphi_{10} = \varphi_{10}^{(3n)} = -\varphi_0^{(n)}$	2	min	max	max

The use of all the gaits considered in this work allows the crawling robot to move in limited spaces in the presence of obstacles. In Fig. 9 an example of surpassing the corridor by a robot is given, the obstacles are painted gray.

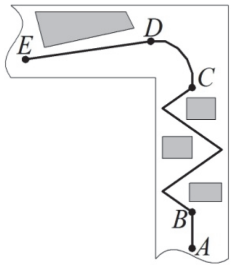


Figure 9 – Movement of the crawling robot along the corridor

The trajectory of the device motion is divided into four sections: *A-B* – rectilinear motion, *B-C* – motion along the s-shaped path to avoid obstacles, *C-D* – motion along the arc of the circle to overcome a turn, *D-E* – rectilinear motion while by-passing obstacles. It should be noted that at the points *A-E* of connecting the sections of movement, the robot performs spot-turn operation for positioning on each of the sections of the trajectory.

Conclusion

The present article touches upon the gait of a two-link crawling robot moving inside buildings for reconnaissance and prospecting. A feature of this device is the presence of controllable supporting elements at the ends of its links, which allow interacting with the surface by supporting surfaces with two different friction coefficients: the support can move when the contact has a low friction coefficient, and the support is stationary with a large friction coefficient. This allows the robot to implement two types of gaits: controlled (one of the links is always fixed on the surface) and combined (inertial movements combined with controlled ones). In the paper, the most common gaits of the robot are considered in detail: a rectilinear motion, a spot-turn operation, moving along an arc of a circle and along an s-shaped path; it is shown how using the gait data it is possible to realize the device movement in a limited space with obstacles.

References

1. Conkur E.S., Gurbuz R. Path Planning Algorithm For Snake-Like Robots // Information Technology and Control. – 2008. – Vol. 37. – № 2. – P. 159-162.
2. Takemori T., Tanaka M., Matsuno F. Gait Design for a Snake Robot by Connecting Curve Segments and Experimental Demonstration // IEEE Transactions on Robotics. – 2018. – Vol. 34. – № 5. – P. 1384-1391.
3. Fu Q., Mitchel T., Yi N., Gart S., Li C. Snake robot's poor 3-D obstacle traversal reveals snake's better stability mechanisms // Bulletin of the American Physical Society. – 2018. – Vol. 63. – №. 1.
4. Lounis D., Spinello D., Gueaieb W., Sarfraz H. Planar kinematics analysis of a snake-like robot // Robotica. – 2014. – Vol. 32. – № 5. – P. 659-675.
5. Wu W., Jun X. Y., Wei H. L., Ri S. M., Chun X. C., Zhen Y. H., Zhong L. Structure Design of Climbing Snake-Like Robot for Detection of Cable-Stayed Bridge // Applied Mechanics and Materials. –2014. – Vol. 598. – P. 610-618.
6. Huang C.W., Huang C.H., Hung Y.H., Chang C.Y. Sensing pipes of a nuclear power mechanism using low-cost snake robot // Advances in Mechanical Engineering. – 2018. – Vol. 10. – №. 6.
7. Vorochaeva L.Yu., Panovko G.Ya., Savin S.I., Yatsun A.S. Movement Simulation of a Five-Link Crawling Robot with Controlled Friction Forces // J. of Machinery Manufacture and Reliability. – 2017. – Vol. 46. – № 6. – P. 527–535.
8. Vorochaeva L.Yu., A.S. Yatsun, S.F. Yatsun Simulation of the motion of a five-link crawling robot with controlled friction on a surface having obstacles // J. of Computer and Systems Sciences International. – 2017. – V. 56. – № 3. – P. 527–552.
9. Vorochaeva L.Yu., Naumov G.S., Yatsun S.F. Simulation of Motion of a Three-Link Robot with Controlled Friction Forces on a Horizontal Rough Surface // J. of Computer and Systems Sciences International. – 2015. – Vol. 54. – № 1. – P. 151–164.
10. Chernousko F.L. O dvizhenii trekhzvennika po gorizontalnoy ploskosti // PMM. – 2001. – Vol. 65. – № 1. – P. 15-20.
11. Chernousko F.L. Upravlyayemyye dvizheniya dvuzvennika po gorizontalnoy ploskosti // PMM. – 2001. – Vol. 65. – № 4. – P. 578-591.
12. Chernousko F.L. Dvizheniye ploskogo mnogozvennika po sherokhovatoy gorizontalnoy ploskosti // Dokl. RAN. – 2000. – Vol. 370. – № 2. – P. 186-189.

A.V. Mal'chikov, L.Yu. Vorochaeva, A.V. Repkin

IMPLEMENTATION OF THE SET OF MEASURING TOOLS OF A WHEELED JUMPING ROBOT FOR THE TASKS OF AUTONOMOUS OVERCOMING OF OBSTACLES

Southwest State University, Kursk, Russia
mila180888@yandex.ru

Abstract

In this paper a mobile wheeled jumping robot is considered. The problem of creating an autonomous control system that allows to overcome obstacles is investigated. For the implementation of this system, a set of measuring tools has been developed, which allows to determine the presence and size of obstacles. The paper also analyzes the existing methods of circumventing obstacles and finding the shortest path, taking into account the different cost of moving through the wheel and jump propulsion.

Keywords: wheeled robot, jumping robot, sensing, overcoming of obstacles.

Acknowledgments

The work was carried out within the RFBR project № 18-31-00075.

Introduction

In recent decades, more and more mechanisms can be found that combine the principles of a wheeled propulsion and a jumping mechanism. This is due to the wide possibilities to overcome difficult areas on rough terrain opening up in front of a mobile robot, moving with a separation from the supporting surface [1-11]. Wheeled robots can be widely used for patrol missions, search work in difficult conditions, including among the rubble, i.e. in conditions where human movement may be difficult or unsafe.

As a rule, mobile platforms are equipped with a system of cameras, additional sensors, information from which is transmitted to the operator's post, which controls the trajectory of the device [1-11]. The disadvantages of this approach are obvious, therefore, within the framework of this project, it is proposed to develop a measuring system that allows the implementation of an autonomous mode of movement of a wheeled jumping robot.

To do this, it is necessary to develop a sensing system, a system for preprocessing and data preparation, analyze the possibility of using existing SLAM-algorithms for solving the problem of constructing a map, review the path finding algorithms if the terrain map is known. As part of the work, one of the possible hardware implementations of the measuring set of an autonomous mobile wheeled jumping robot is shown.

1. Description of the robot and justification of the measuring system

One of the possible layouts of the wheeled jumping robot is shown in Fig.1.

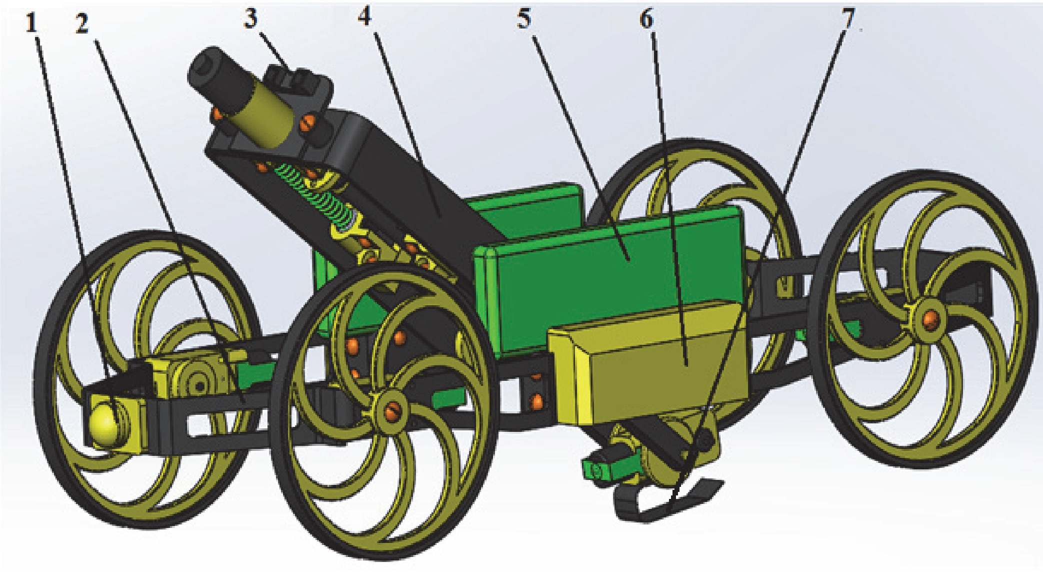


Figure 1 – Wheeled jumping robot: 1 – camera, 2 – power frame, 3 – Infrared rangefinder, 4 – jumping mechanism, 5 – batteries, 6 – control system, 7 – jumping mechanism support

Fig. 1 shows a three-dimensional solid model of a prototype jumping wheeled robot. The design of the robot includes four driving wheels set in motion by DC drives. The robot has a rotary jump module, on which, in addition to the adjustable elastic support, there is an infrared distance sensor (rangefinder), which allows determining the distance to the obstacle, as well as its height. In front of the power frame, a video camera is installed. The mobile robot is equipped with a pair of lithium-polymer batteries and a control unit. The wheels are made of elastic plastic, which avoids damage when landing after a jump. Optionally, the robot can be equipped with a swiveling lidar, which is fixed on the power frame in front of the robot (not shown in Fig. 1). In more detail the design and operation of the mechanisms described in [12-15].

As is already clear from the presented figure, to obtain information about the surrounding space, it is proposed to use a system that includes an infrared rangefinder and a low-resolution video camera. The choice of this combination of sensors due to the wide possibilities that open with the simultaneous use of various types of sensors. Let us consider in more detail the methods for determining the presence and parameters of obstacles.

2. Building a two-dimensional map of the environment

When the robot moves in the autonomous mode, it is necessary to determine the presence of obstacles, to assess the possibility of overcoming them using the wheel propulsion or with the help of jumps. In case the obstacle is insurmountable for jumping by the robot, it is necessary to establish the shortest way to bypass it. If the environment map is known in advance, it is possible to apply path finding techniques. A detailed analysis of the applicability of various algorithms will be discussed below. If the environment map is unknown, then the mobile robot needs to search for the path simultaneously with the construction of the map. This task is solved using the SLAM-method.

The essence of SLAM is to build a map of a previously unknown environment while simultaneously monitoring the current location relative to the starting point [16-18]. To date, the most popular methods are the “particle filter” and the advanced Kalman filter. The need to use complex filters due to the low accuracy of devices that measure the distance to the object. As part of this work, we will pay special attention to the selection of a set of measuring tools that allow implementing various SLAM-methods.

To construct the map, it is necessary to accurately position the robot relative to the starting point [16-18]. To do this, the mobile robot must be equipped with angle sensors on each of the wheels. Note that the evaluation of the angular velocities of all four driving wheels and the measurement of the orientation of the robot using a magnetometer allow us to estimate both the trajectory of the robot and the amount of slippage. To determine the orientation of the device in space, it is convenient to use IMU-modules, for example, CMPS11 (Fig. 2, a). As a rule, these devices contain acceleration sensors, a gyroscope, an electronic magnetometer, a barometer, etc. Information from the sensors is filtered and multiplexed using various filters (Kalman, Majvik, etc.).

To estimate the distance to the obstacles, laser rangefinders are usually used, often mounted on a movable base. As such a device, the lidar URG-04LX-UG01can be used (Fig. 2, b). However, the use of lidars requires a high-performance computing unit and is not always justified [19]. Mapping the environment, its analysis, finding the optimal path can be useful when using the map for several times. For example, in the process of searching, a robot can map an environment, moving from one section to another, remembering the best routes between sections, and in the case of repeated movement along them, use the maps and trajectories already received. But for the task of a one-time movement in an unknown area, mapping may be required only in order not to come to a dead end again in search of the passage. If the probability of hitting the “dead end” is low, for example, when driving on open rough terrain, then the presence of a complex system of orientation is not necessary. On the contrary, to ensure energy saving, the obstacle detection system should be simple and effective.

To detect the presence of an obstacle and determine the distance to it, it is proposed to use the infrared rangefinder SHARP-GP2Y0A02YK0F (Fig. 2, c). When the robot moves, the control system constantly polls the distance sensor. At the time of the appearance of an obstacle, the robot reduces speed and moves to the required distance to the object. Next, an assessment is made of the shape and size of the obstacle using a video camera.

To identify obstacles and complement a two-dimensional environment map, a digital camera with VC0706 data processing module is used (Fig. 2, d). The main advantages of the solution are shooting speeds of up to 10 frames per second, resolution of 640 by 480 pixels, relatively small size; USART or SPI interface. The processor has a built-in JPEG codec that allows to take pictures in automatic mode. For this module there are libraries that greatly simplify the processing of data from the camera.

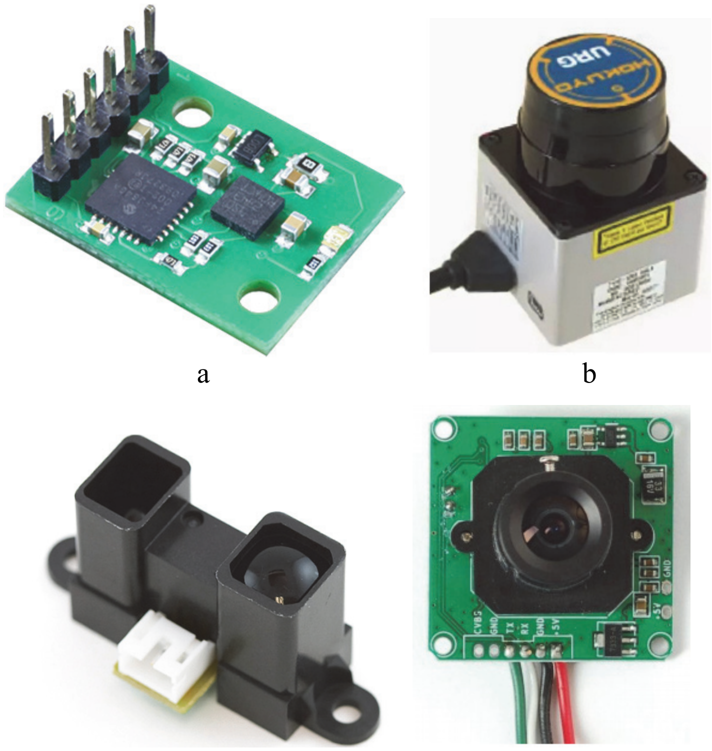


Figure 2 − Components of the measuring system: a – IMU-module CMPS11, b – lidar URG-04LX-UG01, c – infrared rangefinder SHARP, d – digital camera with processor VC0706

Since the distance to the object is known, the system only needs a separate image of the space in front of the robot to determine the possibility of its detour or jump over. If the object blocking the path to the robot has an extended length and its detour without additional information is considered impossible, the possibility of making a jump on (behind) the obstacle is estimated. The height of the object can be estimated both from the image from the camera and from the infrared rangefinder, which is attached to the turning mechanism of the jump. Thus, using the rotation of the accelerating module relative to the body, the rangefinder is aimed at an obstacle (Fig. 3).

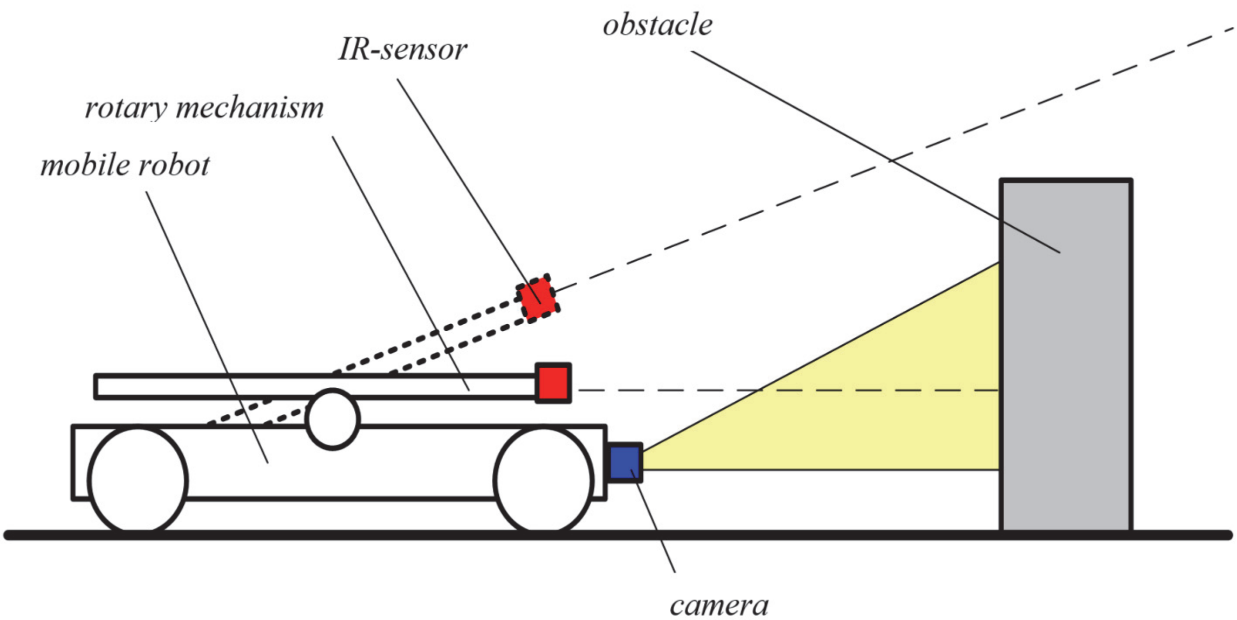


Figure 3 – Diagram of the use of the IR-sensor on the rotary mechanism of the jump to determine the height of the obstacles

If the set height of the object exceeds the capabilities of the jumping mechanism, then the search for a way around the obstacle is carried out. Schematically, the algorithm for passing an obstacle is shown in Fig. 4.

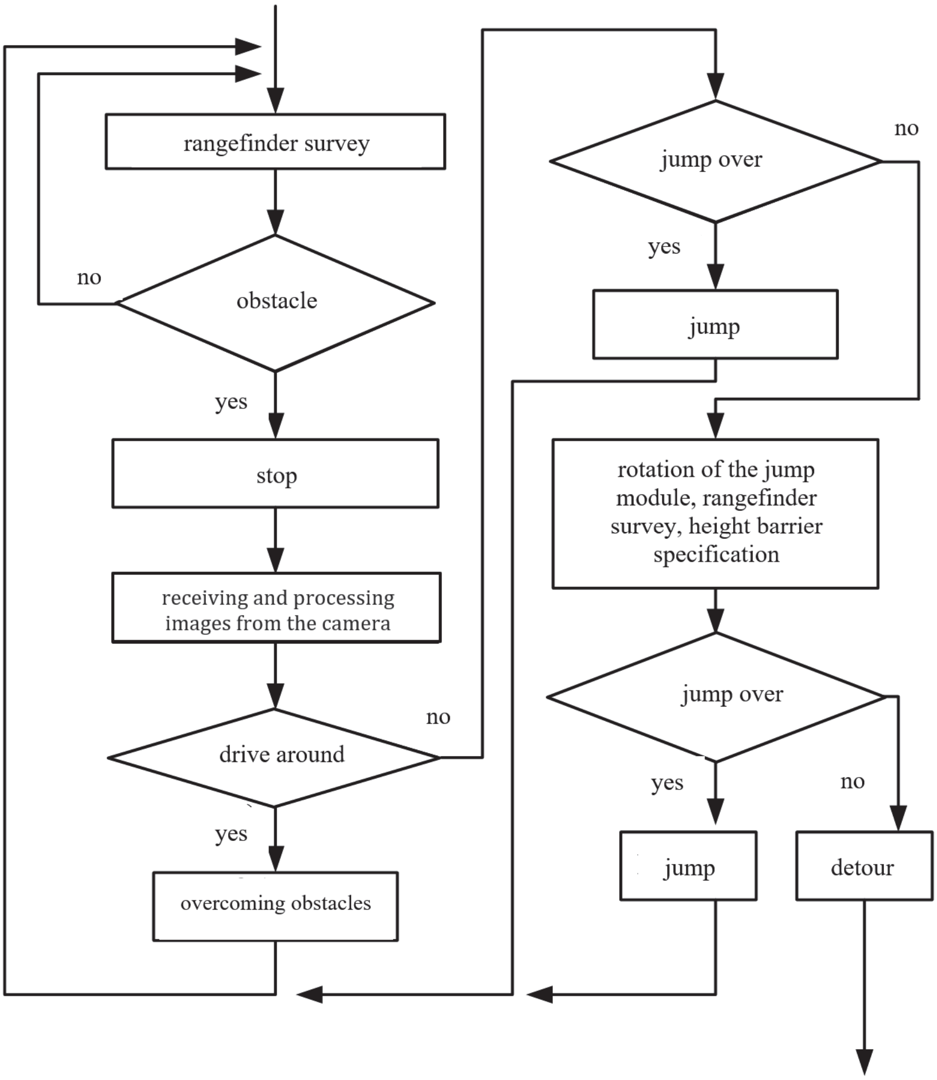


Figure 4 − Algorithm for overcoming obstacles detected by the rangefinder and camera

The method of implementation of movement along the blocking object may be different. When using a single distance sensor and a camera, the algorithm will be based on movement along the obstacle, periodically turning on the spot and re-analyzing the possibility of overcoming it by a detour or by jumping over.

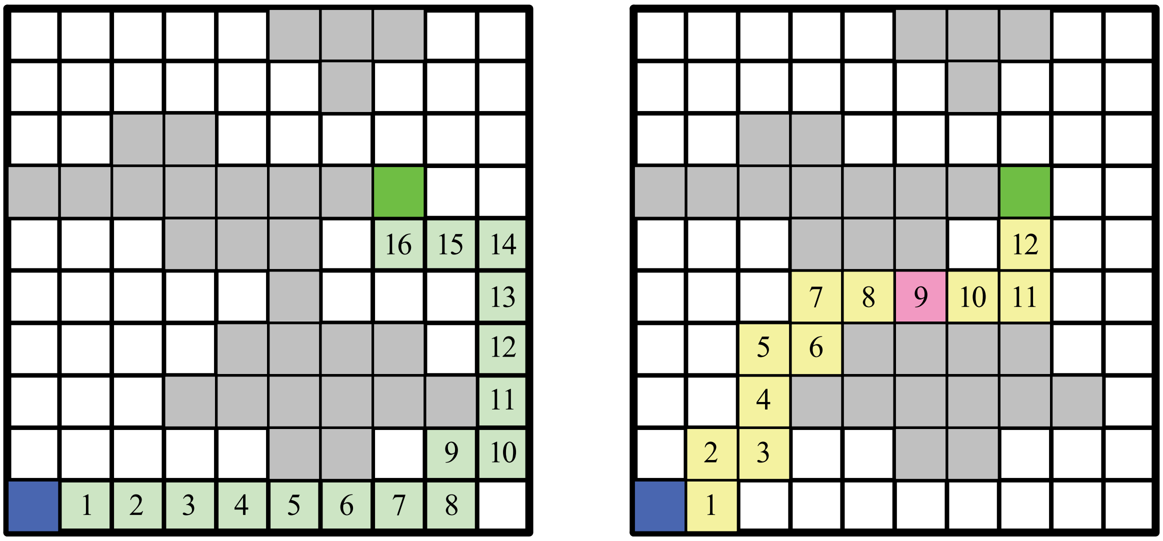
Two types of sensing system implementations allow building an environmental map with varying accuracy and speed. However, the search for the optimal (shortest) path with a known map is also an interesting task.

3. Determination of the optimal trajectory of overcoming obstacles

If there is a well-known environment map, the robot must determine the optimal, shortest path to the object. The classic way to search for a path is the wide search algorithm [20]. This algorithm performs research in all directions and can be useful not only for finding a path, but also for mapping when generating test polygons, etc. The disadvantage of the search algorithm in width is the impossibility of finding the shortest path at different cost of moving.

In this example, the robot can move not only using the wheel propulsion, but also by means of jumps. The cost of such a move is different. So from the point of view of speed, the robot will spend time planning the trajectory of the jump, cocking the jump mechanism, etc. From the point of view of energy consumption, the jump also has a higher cost, since the consumption of the jumping mechanism is much higher than the drives of the wheel propulsion. However, detachment from the surface allows to move with greater speed and shorten the path by overcoming small obstacles.

Schematically, two different approaches to the implementation of movement on the map with obstacles from the starting point, located in the lower left corner of the map, to the end point, located near the upper right corner of the map, are shown in fig. 5. Movement using the wheel module is shown in Fig. 5, a, the number of steps in this method of movement is equal to 16. The use of the combined method of movement - due to the wheel module and jumps - is presented in Fig. 5, b, in this case the number of steps is 12.



a b

Figure 5 – Scheme of finding the shortest path: a – when driving only on wheels, b – when using a jumping mechanism

To find the way, taking into account the cost, use different methods. One of the most common is the Dijkstra algorithm [21]. This method allows you to make a path with different cost of movement, which fits the described task. The disadvantage of the methods shown is a uniform calculation in all directions, which increases the number of necessary operations in the search process. To improve the search efficiency, various heuristic search methods are used [21, 22].

To solve such a problem, it is convenient to use the methods of "greedy search", for example, the method of "greedy search by first best match." In contrast to the “search in width”, the closest point to the target will be used first. This method allows you to determine the path to the goal much faster, however, this path will not necessarily be the shortest.

There are also more complex algorithms, including the wave algorithm, route algorithms, methods of the navigation grid, hierarchical obstacle avoidance algorithms, the "divide and conquer" method.

The simplest and at the same time allowing to determine the best way, taking into account the cost of movements in various ways, is the A-Star – A* method [21]. There are ready-made libraries that implement such an algorithm, and its implementation does not require a high-performance computing module.

Conclusion

This paper discusses possible options for implementing a measuring system that allows you to implement automatic and semi-automatic modes of movement of a wheeled jumping robot. To detect obstacles, it is proposed to use an infrared distance sensor. To assess the shape and size of the obstacles are encouraged to use images taken from the camera. Additionally, you can estimate the height of the obstacle using also the IR sensor, but only when turning the hopping mechanism to aim the sensor. The paper provides a brief overview of the existing methods for finding a path and methods for constructing environmental maps using the selected sensor system.

References

1. Ackerman E. Boston dynamics sand flea robot demonstrates astonishing jumping skills // IEEE Spectrum Robotics Blog. − 2012. – Vol. 2(1).
2. Armour R., Paskins K., Bowyer A., Vincent J., Megill W. Jumping robots: a biomimetic solution to locomotion across rough terrain // J. Bioinspiration and Biomimetics − 2007. – Vol. 2. – P. 65–82.
3. Kovac M. Bioinspired jumping locomotion for miniature robotics: Ph.D. dissertation // Ecole Polytechnique F´ed´erale de Lausanne. − 2010. – 194 p.

4. Sato A., Buehler M. A planar hopping robot with one actuator: design, simulation, and experimental results // Intelligent Robots and Systems: Proc. IEEE/RSJ Intern. Conf., Pasadena. − 2004. − Vol. 4. − P. 3540−3545.
5. Yatsun S.F., Lupekhina I.V., Rukavitsyn A.N. Issledovaniye upravlyayemogo dvizheniya prygayushchego minirobota // Izvestiya vysshikh uchebnykh zavedeniy. Severo-Kavkazskiy region. Tekhnicheskiye nauki. – 2011. – №. 2. – S. 10-15.
6. Stoeter S., Papanikolopoulos N. Kinematic Motion Model for Jumping Scout Robots // Transactions on Robotics and Automation: Proc. IEEE Intern. Conf., Orlando. − 2006. − Vol. 22(2). − P. 398−403.
7. Yatsun S. F., Volkova L. YU., Vorochayev A. V. Issledovaniye rezhimov razgona chetyrekhzvennogo prygayushchego apparata // Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta. – 2013. – T. 19. – №. 24. – S. 86-92.
8. Umehara A., Yamamoto Y., Nishi H., Takanishi A., Lim H.O. Jumping Pattern Generation for One-legged Jumping Robot // Control, Automation and Systems (ICCAS): Proc. 17th IEEE Intern. Conf., Jeju, South Korea. − 2017. – P. 1396−1400.
9. Yatsun S.F., Volkova L.YU. Modelirovaniye dvizheniya mnogozvennogo prygayushchego robota i issledovaniye yego kharakteristik // Izvestiya RAN. TiSU. – 2013. – №. 4. – S. 137-137.
10. Yatsun S.F., Vorochayeva L.YU. Matematicheskoye modelirovaniye upravlyayemogo dvizheniya kolesnogo pyatizvennogo prygayushchego robota // Izvestiya RAN. TiSU. – 2015. – №. 4. – S. 68-68.
11. Zhang Z., Zhao J., Chen H., Chen D. A Survey of Bioinspired Jumping Robot: Takeoff, Air Posture Adjustment, and Landing Buffer // Applied bionics and biomechanics. – 2017. – Vol. 2017. – P. 1−22.
12. Yatsun S.F., Vorochayeva L.YU., Savin S.I. Issledovaniye voprosov upravleniya oriyentatsiyey kolesnogo prygayushchego robota v polete // Mekhatronika, avtomatizatsiya, upravleniye. – 2019. – T. 20. – №. 4. – S. 236-243.
13. Vorochayeva L.YU., Mal'chikov A.V., Savin S.I. Konstruktivnyye osobennosti i klassifikatsiya prygayushchikh robotov // Cloud of science. – 2018. – T. 5. – №. 3. – S. 473-497.
14. Vorochayeva L.YU., Mal'chikov A.V., Savin S.I. Opredeleniye diapazonov dopustimykh znacheniy geometricheskikh parametrov kolesnogo prygayushchego robota // Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. – 2018. – T. 22. – №. 2. – S. 76-84.
15. Mal'chikov A.V., Vorochayeva L.YU., Savin S.I. Konstruktsiya i sistema upravleniya kolesnogo prygayushchego robota // MIKMUS-2018: sbornik trudov konferentsii, Moskva. − 2019. − S. 479−482.
16. Durrant-Whyte H., Bailey T. Simultaneous localization and mapping: part I // IEEE Robotics & Automation Magazine. – 2006. – Vol. 13. – №. 2. – P. 99-110.
17. Montemerlo M., Thrun S., Koller D., Wegbreit, B. FastSLAM: A factored solution to the simultaneous localization and mapping problem // Aaai/iaai. – 2002. – P. 593-598.
18. Dissanayake M. W. M. G., Newman P., Clark S., Durrant-Whyte H. F., Csorba M. A solution to the simultaneous localization and map building (SLAM) problem // IEEE Transactions on robotics and automation. – 2001. – Vol. 17. – №. 3. – P. 229-241.
19. Schwarz B. LIDAR: Mapping the world in 3D // Nature Photonics. – 2010. – Vol. 4. – №. 7. – P. 429-430.
20. Moore E. F. The shortest path through a maze // Proc. Int. Symp. Switching Theory. – 1959. – P. 285-292.
21. Delling D., Sanders P., Schultes D., Wagner D. Engineering route planning algorithms // Algorithmics of large and complex networks. – Springer, Berlin, Heidelberg. – 2009. – P. 117-139.
22. Geisberger R., Sanders P., Schultes D., Delling D. Contraction hierarchies: Faster and simpler hierarchical routing in road networks // Intern.l Workshop on Experimental and Efficient Algorithms. – Springer, Berlin, Heidelberg. – 2008. – P. 319-333.

B.S. Lapin[1], I.L. Ermolov[2], S.A. Sobolnikov[1]

THE SIMPLY INTEGRATED APPROACH FOR SURFACE PARAMETERS DETECTION BY UGV

[1] MSTU "STANKIN", Moscow, Russia
[2] Ishlinsky Institute for Problems in Mechanics RAS, Moscow, Russia
ermolov@ipmnet.ru

Abstract

Autonomy of UGV's should be supported by efficient models of environment. Such models provide the implementation of various control tasks including navigation, trajectory planning, transportation over rough terrain with poor environment visibility etc [1]. One of the challenges is terrain identification for further motion planning [2,3]. This task is greatly complicated by the fact of significant limitation of energy and computational resources provided by mobile robots.

The paper presents a new approach for terrain identification by UGV with known position and limited computational resources. It considers a combined method with direct and remote approach for terrain identification.

The terrain data obtained from the identification system allows to effectively solve tasks of motion planning, SLAM and odometry correction [3].

Keywords: robotics, UGV, terrain identification, soil parameters.

Acknowledgments

This work is supported by RFBR Grant №16-29-04199 ofi_m.

Part of this research was supported by Program "Advanced Topics of Robotic Systems" of the Presidium of the Russian Academy of Sciences.

Introduction

Contemporary practice of using autonomous UGV on rugged terrain faces difficulties in solving auto move task. Classical navigation and mapping systems take into account only geometrical and color parameters of the environment but do not consider other important factors, such as friction, that strongly affecting the movement.

It is important to identify the surfaces around the moving robot, for example in order to avoid slippery areas, weak soils, etc. Efficient path planning and environmental mapping is impossible without proper information about terrain properties.

There are many papers devoted to the terrain and traversable regions identification. In these works, two main approaches are considered: remote and direct. The following section presents these methods in more detail.

Direct and remote approaches

The direct approach consists in detecting surface (or terrain) parameters directly during the interaction between robot chassis and surface. The most promising implementation of this approach involves measurement of the interaction forces. In the paper [4] a method for determining the underlying surface characteristics using reactive forces analysis on the wheels obtained from the load cell integrated into the suspension elements was considered. In this work, the chassis performs standard movements for various types of underlying surface, accumulating information for training the neural network.

Common direct approach disadvantages are:

– direct interaction between robot chassis and surface;

– using additional devices (force sensors, penetrometers, etc).

The main advantage of this approach is ability of the friction forces direct measurement between robot chassis and surface.

The remote approach is based on using computer-vision system for surface parameters estimation away from the robot. In paper [5], a system for surface type remote identification using integrated information obtained from video cameras and a scanning laser range finder is presented.

Common remote approach disadvantages are:

– indirectly surface parameters measurement,

– pre-training for surface type and parameters detection.

The main advantage of the remote approach is ability for surface parameters estimation far from a robot.

Combined Surface Identification Method

Authors propose a new method, based on combining two previously mentioned approaches [4,5] for surface parameters detection. In this case it is possible to simplify each of the approaches separately, which allows to solve the identification problem with small computational resources. In figure 1 a variant of the algorithm that implements proposed method is shown.

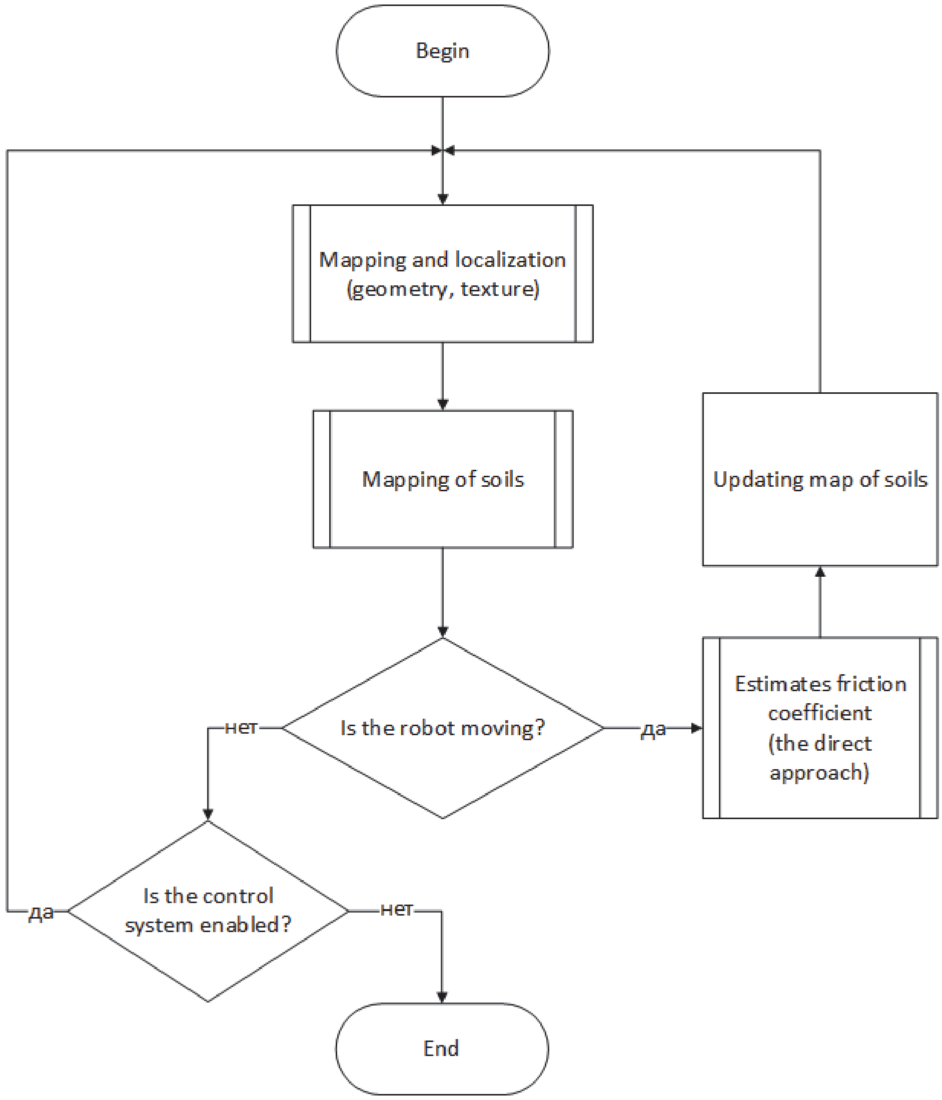


Figure 1 – Simple combining algorithm

The algorithm presented in figure 1 shows the tasks distribution between direct and remote parts in surface parameters identification. The main task of the remote approach is mapping, while general goal of the direct approach is adding reliable information (friction parameters) on the map.

The disadvantage of the algorithm is a long time needed for terrain exploration and surface map building by mobile robot. At the system starts the robot will have reliable information only about the surface on which it stands. The surface even at small distance from the robot will be less known to it.

Direct identification

One of the tasks involved in the development of the direct identification system was simplification of integration into existing robots' models, thus only a standard set of UGV sensors was considered. In this work we used only wheel torque feedback (or current, in the case of indirect measurement), and localization system. The localization system consists of wheels odometry and remote 6DOF tracking system by AprilTag 3 [6]. The remote tracking system includes AprilTag markers installed on the robot (see figure 3) and RGB camera installed remotely to track the markers.

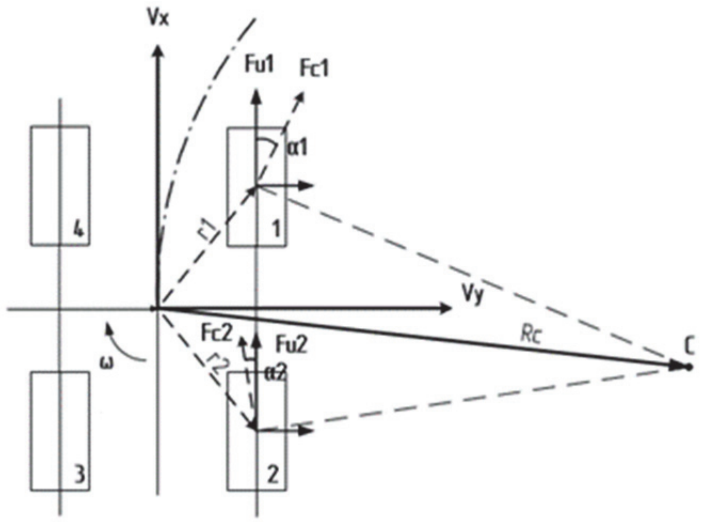


Figure 2 – Wheeled robot chassis

In this work we used skid-steer robot with torque feedback for each wheel (fig. 2).

We assume the task of the robot position estimation solving by localization system. The lower limit of sliding friction coefficient between each wheel and the surface can be estimated using a simple maximum search for the torque measured from the wheels:

$$\mu_{min} = \frac{F_{ci}}{N_i} = \frac{\max\left(\frac{M_i(t)}{r_{\kappa i}}\right)}{\cos(\alpha_i)\cdot N_i}, \tag{1}$$

where $M_i(t)$ – torque on wheel, $r_{\kappa i}$ – wheel radius, α_i – angle between F_{ci} and F_{ui}, N_i – reaction force.

Expression (1) is valid only if the wheel has direct contact with the surface. To estimate real value of the friction coefficient, it is necessary to consider only values obtained when slippage between the wheel and the ground occurs. The slippage can be determined using data from the robot's localization system and based on the torque information measured at the wheel. The second case is more complicated.

Remote identification

We tried to simplify the remote identification system integration of into UGV, so we used standard robot's vision system hardware. Such system may consist of different sets of sensors that provide two types of data: a points cloud, containing distances to obstacles, and color image of the robot working area.

The main disadvantage of the remote surface identification systems is pre-training for different terrains. It is difficult because of training set creation complexity. In addition, these systems allow only to identify the type of surface without its parameters. Typically, these solutions are based on classifying the data by surface type.

In this paper, we propose to use the remote identification system to solve the problem of clustering, rather than classification. This greatly simplifies implementation of such system, since this task does not require its training.

The clustering task involves dividing the entire surface into regions with similar characteristics. From the robot vision system, at least two surface parameters can be obtained: the dispersion of the uneven heights D and the color information, which can be conveniently represented as three components: the color tone H, the saturation S, and the brightness V. Thus, the distance function characterizing similarity of surface parameters depends on four components: D, H, S, V.

There are many types of distance functions used in cluster analysis to compare data: Euclidean, Minkowski, cosine, Mahalanobis distance, etc. [7]. There are also a large number of clustering methods that can be used to solve the problem: DBSCAN [8], Mean-shift [9], methods based on the hierarchical approach [10]. The particular method choice depends on computational capabilities of certain UGV model.

The focus of this work is on the direct part of the algorithm that provide identification of the surface parameters. This part is discussed in detail in the next section.

Direct identification algorithm

The KUKA Youbot researching platform, with omni-wheels chassis (see fig. 3), was selected for experiments.

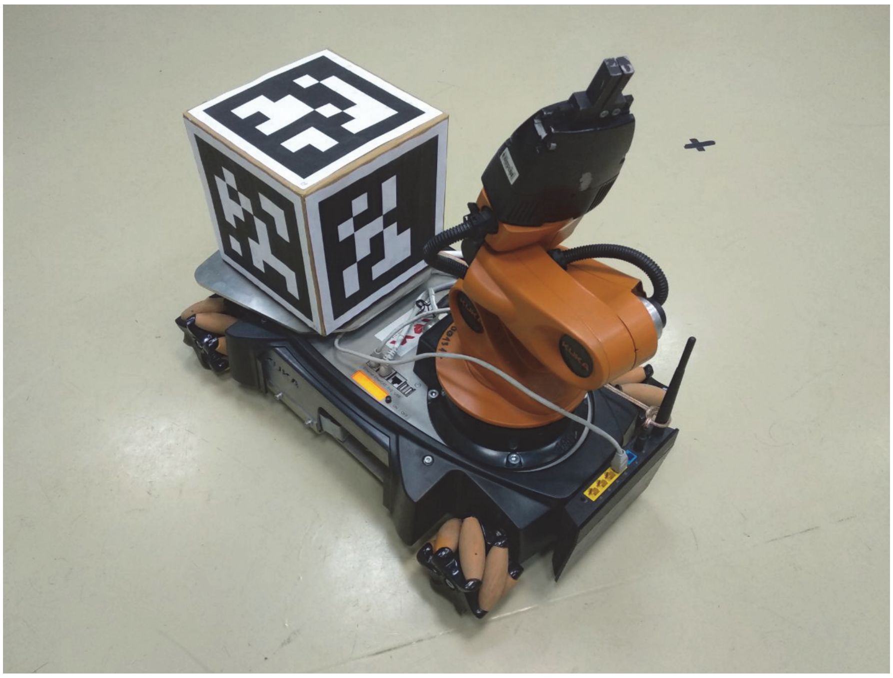


Figure 3 – Kuka Youbot robotics researching platform

This work not considering omni-wheel chassis, that is why other wheels with low friction force (see fig. 4) was developed for replacement.

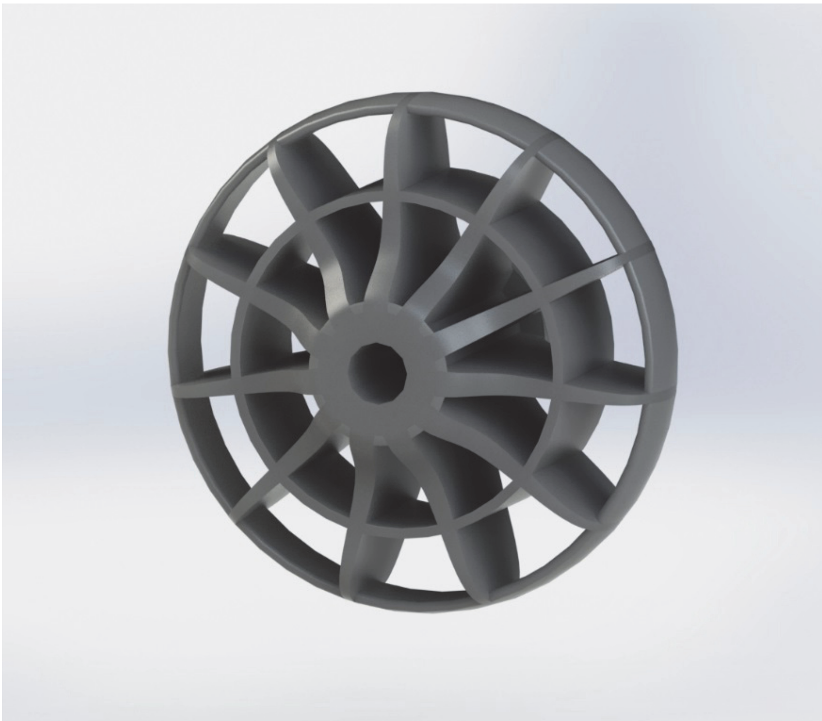


Figure 4 – Wheel with low friction force for Kuka Youbot

For validating the approach, a special software for friction coefficient detection between wheels and surface was developed. The algorithm that describes software operation process presented in Figure 5.

Determination of slippage was performed by the operator. The lower limit of the friction coefficient was estimated using formula (1). The measured values of the wheel torque were filtered by Kalman filter.

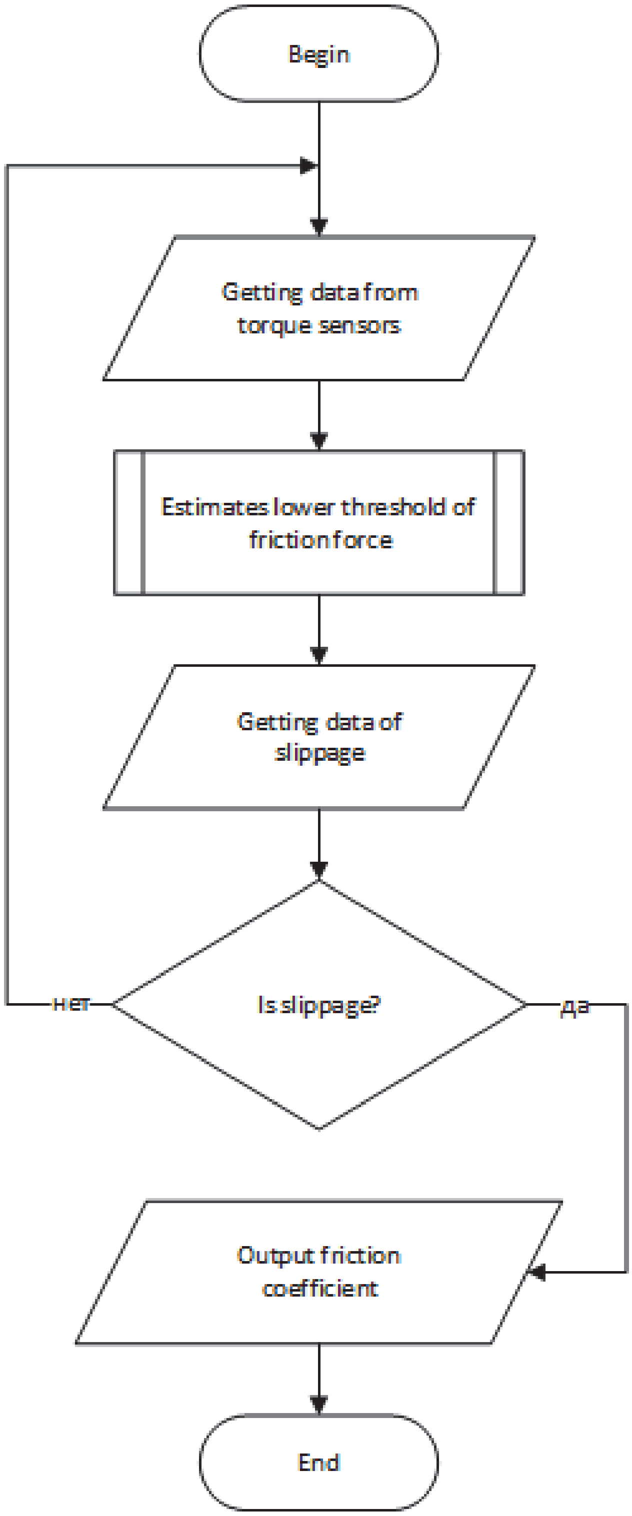


Figure 5 – Algorithm of the direct part

The developed software, which corresponds to the algorithm described above, was used to determine the friction coefficient of the robot wheels (see Fig. 4) with Stankin TP-15 laboratory floor. In Figure 6 the result of this experiment is presented.

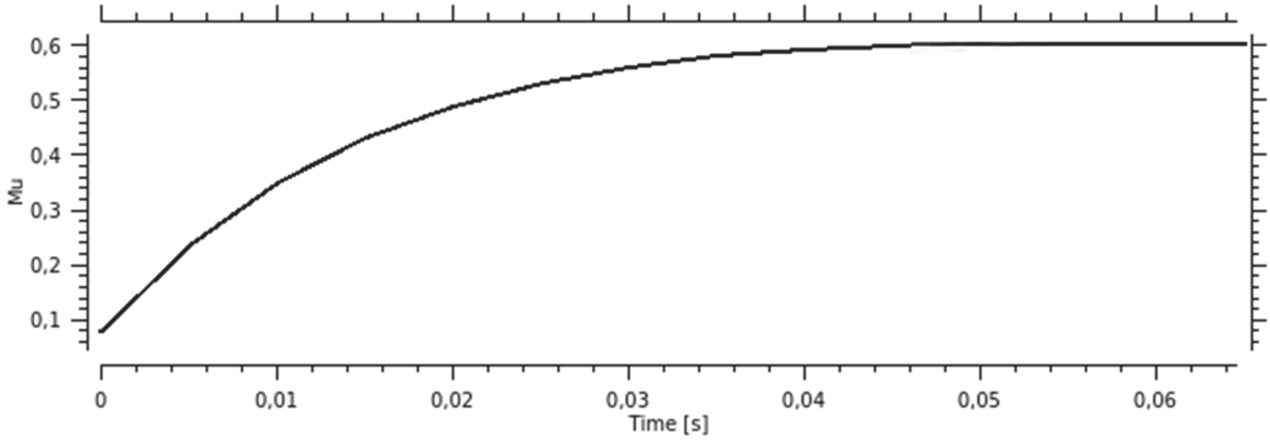


Figure 6 – Friction coefficient

The experiment was conducted with maximum robot speed - 0.8 m/s. Measurement plots are presented in Figure 7.

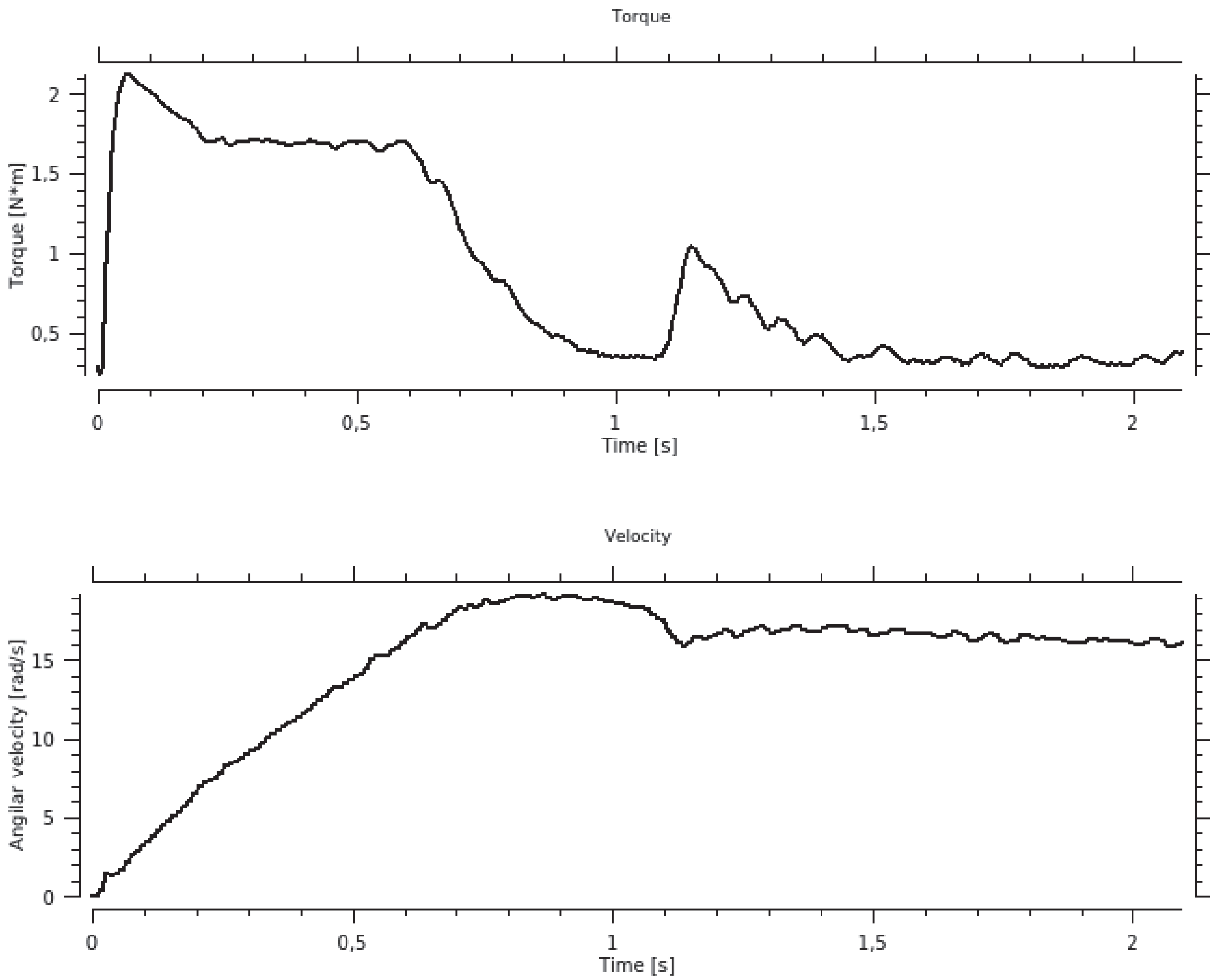


Figure 7 – Wheel torque (top), wheel rotation speed (bottom)

Figure 7 shows all robot acceleration stages. The movement begins with a sharp rise of torque, it is sufficient for slippage effect appearance. The fact that slippage has begun can be traced by the subsequent preservation of the torque, and at the same time an almost linear increase in the wheel rotation speed. The following part of speed curve is explained by overshoot in the wheel drive control system.

Meanwhile, the decline in the torque is explained by the frictional force approaching the boundary. Wheels rotation speed and real robot speed match each other and at the same time slippage decreases. This is followed by a small step in torque, as well as a drawdown in the wheel rotation speed, indicating the final overcoming of the friction boundary force and the end of slippage. In confirmation, Figure 8 shows similar plots, but for lower speed when slip did not occur.

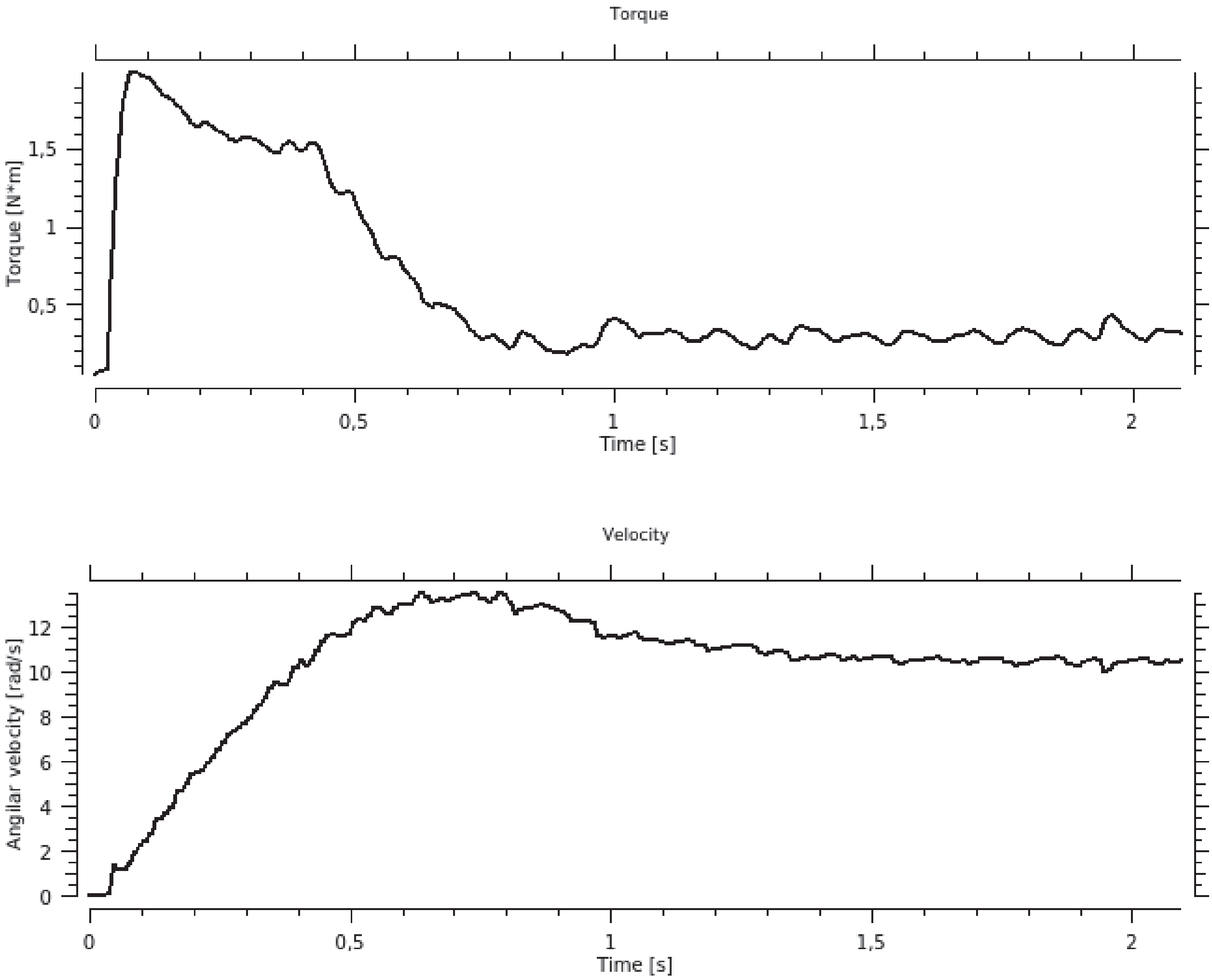


Figure 8 – Wheel torque (top), wheel rotation speed (bottom) without slippage

As one can see, Figure 8 shows no torque stabilization with subsequent decline and jump.

These features inherent in the slip process can be used for slippage detection, controlling a propulsive device with variable passability [11], building of accurate odometry calculation models and multi-UGV control on rugged terrains [12].

References

1. Ermolov I.L., O faktorah, vliyayushchih na uroven' avtonomnosti v prostranstve transportnyh shassi nazemnyh mobil'nyh robotov (*On the factors affecting the level of autonomy of transport chassis of ground mobile robots*) “Izvestiya YuFU. Technical Sciences”, №1, pp 210-218, 2016. *In Russian*
2. V.G. Gradetsky, I.L. Ermolov, M.M. Knyazkov, E.A. Semenov, A.N. Sukhanov, Silovoe vzaimodejstvie mobil'nogo nagruzhennogo robota s gruntom (*The force interaction of loaded mobile robot with the ground*). Mechatronics, automation, control, (12):819–824, 2017. *In Russian*
3. Kudryashov V.B., Lapshov V.S., Noskov V.P., Rubtsov I.V., Robotics problems in UGV, “Izvestiya YuFU. Technical Sciences”, pp. 42-57, 2014. *In Russian*
4. Mashkov K.Y., Naumov V.N., Rubtsov V.I., Sistema avtomaticheskogo opredeleniya harakteristik grunta pri dinamicheskom vzaimodeistvii dvighitelya MRK s opornoi poverhnost’yu (*The system of automatic soil characteristics detection with dynamical interaction between the robot movement base and bearing*

surface) Materials of Eight All-Russia scientific and practical conference “Perspective systems and control tasks”. – Taganrog: TTI SFU, pp 87-95, 2013. *In Russian*
5. A.V.Vazaev, V.P.Noskov, I.V. Rubtsov, S.G. Tsarichenko, *Object detection and ground type classification with combined computer vision system*. “Izvestiya YuFU. Technical Sciences”, №2, pp 127-139, 2013. *In Russian*
6. J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, pp. 4193-4198, 2016. doi: 10.1109/IROS.2016.7759617
7. Xu, D. & Tian, Y. Ann, A Comprehensive Survey of Clustering Algorithms // Data. Sci. (2015) 2: 165. https://doi.org/10.1007/s40745-015-0040-1
8. Sander, Jörg; Ester, Martin; Kriegel, Hans-Peter; Xu, Xiaowei. "Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications". Data Mining and Knowledge Discovery. Berlin: Springer-Verlag. 2 (2): pp 169–194. 1998. doi:10.1023/A:1009745219419.
9. Hartigan, J. A.; Wong, M. A.. "Algorithm AS 136: A k-Means Clustering Algorithm". Journal of the Royal Statistical Society, Series C. 28 (1): pp 100–108, 1979. JSTOR 2346830.
10. Müllner, Daniel, Modern hierarchical, agglomerative clustering algorithms, 1109cc, 2011. arXiv:1109.2378, Bibcode:2011arXiv1109.2378M.
11. Gradetsky V., Ermolov I., Knyazkov M., Lapin B., Semenov E., Sobolnikov S., and Sukhanov A.. Highly Passable Propulsive Device for UGVs on Rugged Terrain. 13th International Scientific-Technical Conference on Electromechanics and Robotics “Zavalishin’s Readings”, 161(03013):1–5, 2018.
12. V. Gradetsky, I. Ermolov, M. Knyazkov et al., Implementation of a joint transport task by a group of robots / Smart Electromechanical Systems. — Vol. 174 of Studies in Systems, Decision and Control. — Springer International Publishing, 2019. — P. 203–214.

V.G. Vhashchukhin

ORIENTATION SYSTEM OF THE AERODYNAMICALLY ADHESIVE WALL CLIMBING ROBOT

Ishlinsky Institute for Problems in Mechanics RAS, Moscow, Russia
ketlk@mail.ru

Abstract

The system of orientation of the wall climbing robot with aerodynamic pressing to the surface is considered. The design of the robot is described. The problem of determining the orientation of the robot on the basis of data obtained from the accelerometer is solved. Using a moving average filter in real time, implemented on the basis of the onboard microcontroller is proposed. The filter allows to significantly reduce the noise and make the data obtained from the accelerometer suitable for determining the orientation of the robot.

Keywords: wall climbing robot, moving average filter.

Acknowledgments

The work was performed on the topic of the state task № AAAA-A17-117021310384-9 and was supported by the Russian Foundation for Basic Research (grant No. 18-01-00650 A) and Russian Academy of Sciences program “Actual problems of robotic systems”.

Itroduction

A promising method of autonomous wall climbing robot’s fixation on surface is aerodynamicaly adhesion due to the creation of underpressure in the vacuum chamber located under the bottom of the robot [1]. A schematic image of the robot is depicted on Figure 1. The vacuum chamber is formed by the bottom of the robot body (1) and the displacement surface. The gap between the robot’s body and the surface should be minimized to minimize the required air flow by means of a sliding seal (5). The underpressure in the vacuum chamber is created by a fan (2) driven by an electric motor. Inadequate underpressure in the vacuum chamber can lead to the separation of the robot from the surface, or to its tilting or sliding, that ultimately will lead to a fall. Excessive underpressure will lead to increased energy consumption of the on-board power source and reduced battery life. One of the drawbacks of this type of fixation on the surface is the increased vibration of the robot’s body, due to the high frequency of rotation of the fan impeller, affecting the readings of sensors installed on board. It is necessary to know robot’s orientation on the plane and the angle of inclination of the plane to the horizon to determine the necessary equilibrium conditions on the displacement surface and control the movement of the robot [2]. The robot is equipped with two wheels (4) with differential drive and a turning wheel to ensure balance on the surface. Wheels driven by electric motors (3). Drive wheels fitted tires with increased friction coefficient.

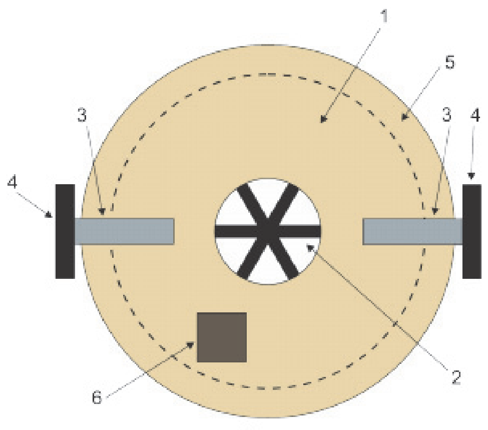


Figure 1 — Schematic illustration of the robot

Since the movements of the robot are rather slow, we can neglect the acceleration of the robot when moving and consider its movements quasi-static. To determine the orientation angles, it is proposed to use a triaxial accelerometer IMU-10 (6) mounted on the body of the robot. In the quasistatic mode, the corresponding component of the gravitational acceleration is measured for each axis of the accelerometer. The effect of vibrations on the accelerometer readings was studied at various orientations of the robot and at different robot's rotation speeds of the fan impeller. A method is proposed for processing data from an accelerometer by using a moving average filter, which is widely used in various fields. [3-5]. The algorithm of the method is implemented on an onboard microcontroller operating in real time. The effect of deformation of tires on accelerometer readings is identified. A method for obtaining the orientation angles of a robot based on accelerometer readings is described. The error of determining the orientation of the robot is estimated at different robot's orientation on the plane and the angles of plane's inclination to the horizon at different speeds of rotation of the fan impeller. It is revealed that the error of orientation measurement essentially depends on the speed of rotation of the impeller. The use of a moving average filter allows to reduce the noise of the accelerometer readings to the required level in real-time calculations. As the onboard controller making calculations used Arduino uno board. Thus, the data coming from the accelerometer become suitable for use in movement control and calculating the conditions for ensuring the equilibrium of the robot.

Determining the orientation of the robot

Let be g_x, g_y, g_z – the values of the accelerometer readings, which correspond to the projections of the gravitational acceleration vector on the axis of the coordinate system rigidly connected with the accelerometer.

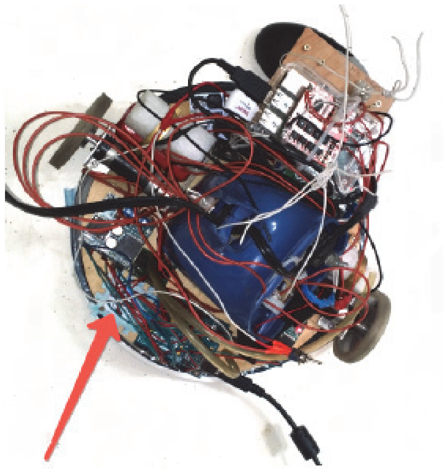


Figure 2 — Physical prototype of the robot, the arrow shows the location of the accelerometer

The angle α of inclination of the robot`s body tj the horizontal plane and the angle β of the robot`s rotation relative to the horizontal can be expressed as follows:

$$\alpha = arccos\left(\frac{-g_z}{\sqrt{g_x^2 + g_y^2 + g_z^2}}\right)$$

$$\beta = arccos\left(\frac{g_y}{\sqrt{g_x^2 + g_y^2}}\right)$$

In Figures 3, 4, 5 shows the data obtained during the experiment and the result of data processing with a moving average filter. In fig. 3 and 4 show the dependence of the inclination angle of the robot on time. It can be seen that the first 30 seconds before turning on the fan motor, the angle values are almost constant. After turning on the engine, the data spread increases significantly. In Figure 3 shows the result of the experiment when the robot rests on the wheels.

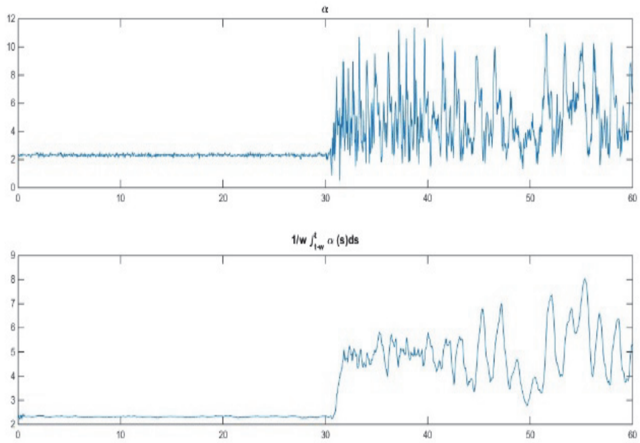


Figure 3 — Data on the angle of inclination, obtained on the basis of the accelerometer readings before turning on the fan and after turning it on at full capacity. The robot relies on wheels

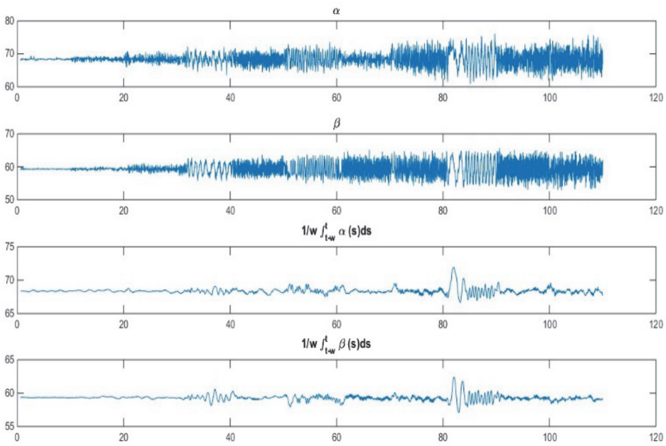


Figure 4 — Data on the angle of inclination, obtained on the basis of the accelerometer readings before turning on the fan and after turning it on at full capacity. The robot rests on the stand, so the wheels do not rest on the surface

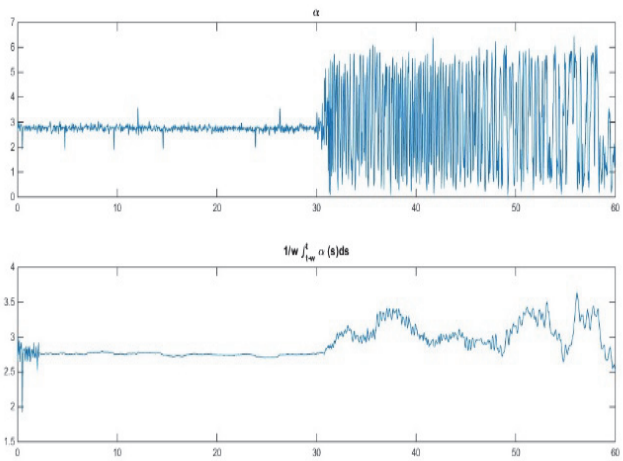


Figure 5 — Applying a moving average filter to inclination and orientation data.
Fan speed gradually increases

The figure shows that after turning on the fan, the inclination angle increases. This is due to the deformation of tires on which the robot rests. In fig. 4 shows the result of the experiment when the effect of tire deformation is excluded. The robot is on a stand so that the wheels do not touch the surface.

To reduce the noise of the accelerometer readings, a moving average filter was used using the following averaging algorithm:

$$\hat{\alpha}(t_i) = \frac{1}{w_i} \sum_{j=i-N_w}^{i} (t_j - t_{j-1})\alpha(t_j)$$

$$\hat{\beta}(t_i) = \frac{1}{w_i} \sum_{j=i-N_w}^{i} (t_j - t_{j-1})\beta(t_j)$$

The parameter $w_i = t_i - t_{i-N_w-1}$ depends on time, averaging occurs on the interval $[t_{i-N_w-1}, t_i]$. A value $N_w = 20$ was determined experimentally, which corresponds to an interval of approximately 800 ms..

In Figure 5 shows an example of applying a filter for processing data about the angles of inclination and orientation of the robot. The fan speed gradually increased every ten seconds by 1/10 of the maximum speed. The graphs are for the case when the inclination angle of the robot is 67.5 degrees, and the orientation angle is 60 degrees. From the two upper graphs it can be seen that the error in determining the angles is about 5 degrees for the highest fan rotation speed. The bottom two graphs show the angles after the data was processed by a moving average filter. The error in determining the angles is no more than 1 degree. The algorithm of the method does not require large computing power, which allowed it to be implemented on the Arduino UNO microcontroller, which controls the operation of both the fan and the drive wheels of the robot. Filtering the signal from the accelerometer occurs in real time.

Conclusion

The influence of tire wheel deformation on accelerometer readings has been revealed. A method for obtaining the orientation angles of a robot based on accelerometer readings is described. The estimated error of determining the orientation of the robot at different positions of the robot on the plane and the angles of inclination of the plane to the horizon at different speeds of rotation of the fan impeller. It is revealed that the error of orientation measurement essentially depends on the speed of rotation of the impeller. The use of a moving average filter allows to reduce the noise of the accelerometer readings to the required level in real-time calculations. The Arduino uno board was used as the onboard controller for the computation. Thus, the data coming from the accelerometer becomes suitable for use in controlling the movement and calculating the conditions for ensuring the equilibrium of the robot.

References

1. D. Longo and G. Muscato, ”The Alicia 3 Climbing Robot: a Three-module Robot for Automatic Wall Inspection,” IEEE Robotics & Automation Magazine, vol. 13, no. 1, pp. 42–50, 2006.
2. A.M. Nunuparov, V.G. Chashchukhin The control system of an autonomous wall climbing robot with aerodynamic adhesion// Proc. of the CLAWAR 2017, Porto, Portugal, September 11 – 13, 2017, pp. 118-126
3. L. Xiong, F. Zhuo, X. Liu, F. Wang, and Y. Chen, ”Optimal design of moving average filter and its application in distorted grid synchronization,” in Proceedings of the 2015 IEEE Energy Conversion Congress and Exposition (ECCE), Montreal, Canada, 2015, pp. 3449–3454.
4. R. Irani, K. Nasrollahi and T. B. Moeslund, ”Improved pulse detection from head motions using DCT,” in Proceedings of the 2014 International Conference on Computer Vision Theory and Applications (VISAPP), Lisbon, Portugal, 2014, pp. 118–124.
5. A. Garfinkel, Y.-H. Kim, O. Voroshilovsky, Zh. Qu, J.R. Kil, M.-H. Lee, H. S. Karagueuzian, J. N. Weiss, and P.-Sh. Chen, ”Preventing ventricularfibrillation by flattening cardiac restitution,” Proceedings of the National Academy of Sciences USA, vol. 97, no. 11, pp. 6061–6066, 2000.

A.S. Kreusova, A.N. Yusupov

THE DEVELOPMENT OF A DISCRETE MODEL OF MECHATRONIC MODULE

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
a.kreusova@rtc.ru, a.n.yusupov@gmail.com

Abstract

The paper describes the software simulation model of mechatronic module. Mechatronic module consists of a DC motor, a planetary and wave gear and a torque sensor. The simulation model represents the electrical and mechanical part of the mechatronic module. Simulation model provides an opportunity to develop and configure control systems, security systems of the manipulator before the possibility of full-fledged experiments. The model implements an algorithm for calculating the friction torque, which differs from the common one by the presence of a hypothetical minimum speed at which the mechanism can continue proper motion. The model was prototyped in Matlab/Simulink package, implemented in C programming language, tested as part of the operating software of the manipulator.

Keywords: simulation model, mechatronic module, friction.

Acknowledgments

The results were obtained in the framework of the state task of the Ministry of education and science of Russia № 075-00924-19-00.

Now the RTC is actively developing manipulators. A full-fledged experiment is often not possible to conduct during robot development. For setting up experiments and debug software, we set a goal to develop a discrete model of the robot manipulator, including a discrete model of the mechatronic module.

Problem could be solved using the mechatronic module discrete model:

- Test and debug control software for mechatronic module
- Test the control system with variation of motor parameters and the environment in a wide range
- Emergency situation simulation
- Modeling, explanation and systematization of emergency situations
- Capitalization of knowledge and technology.
- Automated search for dangerous conditions

In our research, we consider manipulators with 6-8 mechatronic modules, which include a DC motor, a planetary and wave gears, an elastic shaft and a torque sensor. The mechatronic module also includes nonlinear elements – nonlinear electromagnetic effects in the motor, distributed friction in the motor and in the wave gear, elasticity in the wave gear. All these effects had to be taken into account in the discrete model of the mechatronic module.

In our work, the mathematical model of friction contains a hypothetical minimum speed at which the mechanism can continue proper motion. This hypothesis is based on the assumption that at some low speed the kinetic energy is not enough to overcome the next "potential pit" and the kinetic energy is either immediately dissipated into heat or passes into the energy of damped elastic oscillations, which can be observed when stopping various mechanisms.

System describing friction (on the example of motor friction)

$$\begin{cases} M_{fr} = \begin{cases} M_m - M_{load\,m} + M_{to\text{дв}}, & |\omega_m| \le \Delta\omega_m \text{ и } M_m - M_{load\,m} + M_{\text{то}\,m} \le M_{\text{fr stat}\,m} \\ M_{\text{fr stat}\,m}, & |\omega_m| \le \Delta\omega_m \text{ и } M_m - M_{load\,m} + M_{to\,m} > M_{\text{fr stat}\,m} \\ sign(\omega) M_{\text{fr coul}\,m} + M_{\text{fr vis}\,m}, & |\omega_m| > \Delta\omega_m \end{cases} \\ M_{to\,m} = K_{to}\omega_m \\ M_{\text{fr vis}\,m} = K_{\text{fr vis}\,m}\omega_m \end{cases} \quad (1)$$

M_{to} – torque providing a complete stop of the rotor and simulating its "sticking" at low speeds when the static friction begins to start

$M_{\text{to}} = K_{to} \cdot \omega_m$

$K_{\text{то}}$ – viscous friction coefficient of M_{to}

$M_{\text{fr stat}}$ – static friction torque,

$M_{\text{fr coul}}$– coulomb friction torque;

$M_{\text{fr vis}}$ – viscous friction torque,

M_{load} – load torque,

ω_m – motor speed,
$\Delta\omega_m$ – minimum speed of the motor at which the static friction stop function.

The structure of the mechatronic module includes an elastic link – wave gear. We suspect that the dynamic equilibrium (redistribution of forces) in the system of elasticity is established faster than the period of integration of the discrete model (50 μs) and the whole system behaves as a second-order system from an external observer point of view ("load on a spring").

After research made, we stopped at the model of mechatronic module, which conditionally divides it into two units so that the links of elements in the composition of each node are relatively rigid. The first unit includes the model DCM, planetary gear and wave generator, the second unit – the driven part of the wave gear and the shaft between the wave gear and the torque sensor. Each of the units has its own friction model and reduced torque of inertia.

Discrete model functionality depends on a set of modes, produced by the state of nonlinear elements. Each unit of the mechatronic module model can be in one of the states:

1) State of rest - the static friction force keeps the node in a stationary state
2) The state of “breakaway” - the node comes into motion, but static friction force still exists
3) The node is in motion and a typical model of dry and viscous friction exists

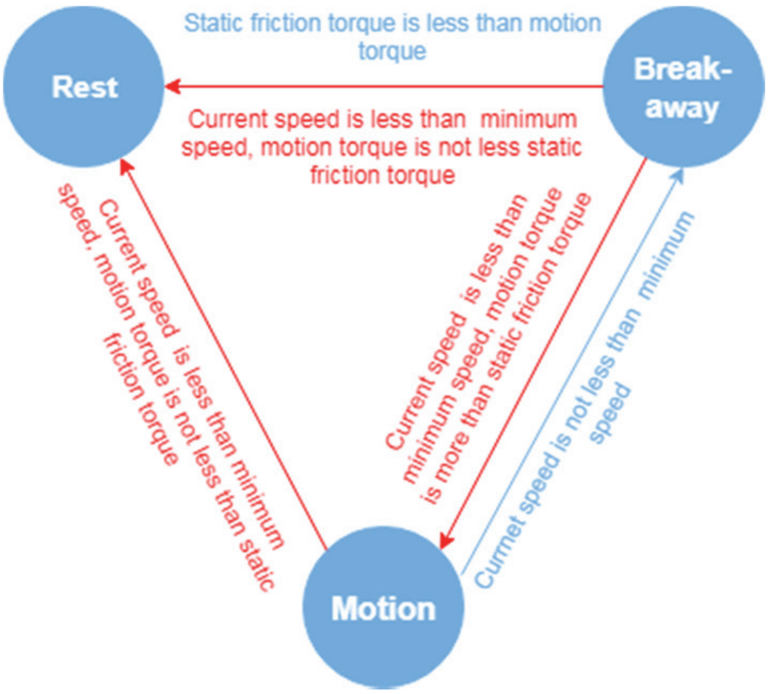


Figure 1 – Graph of transitions between working modes of the mechatronic module depending on friction

The mechatronic module model has 9 modes – for each of the combinations of node States. In General, when all the mechanisms are set in motion and the friction forces no longer cause nonlinear distortions (we believe that the static friction on is constant and is considered as a static load), the model of the mechatronic module dynamics is presented as follows:

$$\begin{cases} \dfrac{dI}{dt} = \dfrac{-RI - K_v\omega_m + U}{L} \\ \dfrac{d\omega_m}{dt} = \dfrac{K_m I}{J_m} - \dfrac{K_{\text{elastic}}}{J_m Z_{plan\ gear}^2 Z_{wave\ gear}^2}\varphi_m - \dfrac{K_{\text{elastic}}}{J_m Z_{plan\ gear} Z_{wave\ gear}}\varphi_{wave\ gear} \\ \dfrac{d\varphi_m}{dt} = \omega_m \\ \dfrac{d\omega_{wave\ gear}}{dt} = \dfrac{K_{\text{elastic}}}{J_{wave\ gear} Z_{plan\ gear} Z_{wave\ gear}}\varphi_m - \dfrac{K_{\text{elastic}}}{J_{wave\ gear}}\varphi_{wave\ gear} - \dfrac{M_{load}}{J_{wave\ gear}} \\ \dfrac{d\varphi_{wave\ gear}}{dt} = \omega_{wave\ gear} \end{cases} \quad (2)$$

R, L – resistance and inductance of rotor winding,
Km, Kv – motor constants,
U, I – voltage and current of rotor winding,
ω_m, φ_m – motor speed and angle position,
J_m,$J_{wave\ gear}$ – moment of inertia of motor shaft and wave gear shaft
$Z_{plan\ gear}$, $Z_{wave\ gear}$ – planetary and wave gears ratios,
$K_{elastic}$ – elasticity coefficient of wave gear,
M_{load} – load torque.

Or in vector-matrix form:

$$\dot{X} = AX + BU, \tag{3}$$

$$A = \begin{bmatrix} -\frac{R}{L} & -\frac{K_v}{L} & 0 & 0 & 0 \\ \frac{K_м}{J_m} & 0 & \frac{K_{\text{elastic}}}{J_m Z_{plan\ gear}^2 Z_{wave\ gear}^2} & 0 & \frac{K_{\text{elastic}}}{J_m Z_{plan\ gear} Z_{wave\ gear}} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{K_{\text{elastic}}}{J_{wave\ gear} Z_{plan\ gear} Z_{wave\ gear}} & 0 & \frac{K_{\text{elastic}}}{J_{wave\ gear}} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4}$$

$$X = \begin{bmatrix} I \\ \omega_m \\ \varphi_m \\ \omega_{wave\ gear} \\ \varphi_{wave\ gear} \end{bmatrix}$$

The matrix A from the system presented earlier has no inverse. So we need to change the basic vector of the system. New basis includes current, motor speed, the speed of the output shaft of the wave gear and the position difference given to the output shaft of the wave gear.

$$\begin{cases} \dfrac{dI}{dt} = \dfrac{-RI - K_v\omega_m + U}{L} \\ \dfrac{d\omega_m}{dt} = \dfrac{K_м I}{J_m} - \dfrac{K_{\text{elastic}}}{J_m Z_{plan\ gear} Z_{wave\ gear}} \left(\dfrac{\varphi_m}{Z_{plan\ gear} Z_{wave\ gear}} - \varphi_{wave\ gear}\right) \\ \dfrac{d\omega_{wave\ gear}}{dt} = \dfrac{K_{\text{elastic}}}{J_{wave\ gear}} \left(\dfrac{\varphi_m}{Z_{plan\ gear} Z_{wave\ gear}} - \varphi_{wave\ gear}\right) - \dfrac{M_{load}}{J_{wave\ gear}} \\ \dfrac{d\left(\dfrac{\varphi_m}{Z_{plan\ gear} Z_{wave\ gear}} - \varphi_{wave\ gear}\right)}{dt} = \dfrac{\omega_m}{Z_{plan\ gear} Z_{wave\ gear}} - \omega_{wave\ gear} \end{cases}$$

$$A = \begin{bmatrix} -\frac{R}{L} & -\frac{K_v}{L} & 0 & 0 \\ \frac{K_м}{J_m} & 0 & 0 & \frac{K_{\text{elastic}}}{J_m Z_{plan\ gear} Z_{wave\ gear}} \\ 0 & 0 & 0 & \frac{K_{\text{elastic}}}{J_{wave\ gear}} \\ 0 & \frac{1}{Z_{plan\ gear} Z_{wave\ gear}} & -1 & 0 \end{bmatrix} \tag{7}$$

$$X = \begin{bmatrix} I \\ \omega_m \\ \omega_{wave\ gear} \\ \frac{\varphi_{дв}}{Z_{plan\ gear} Z_{wave\ gear}} - \varphi_{wave\ gear} \end{bmatrix} \tag{8}$$

$$B = \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & 0 \\ 0 & -\frac{1}{J_{wave\ gear}} \\ 0 & 0 \end{bmatrix} \tag{9}$$

$$U = \begin{bmatrix} U \\ M_{load} \end{bmatrix} \tag{10}$$

The solution of the system will be obtained by the matrix exponent method. We introduce a time discretization.

$$X_{k+1} = X_{max} - e^{At}(X_{max} - X_k) \tag{11}$$

$$X_{max} = -A^{-1}BU \tag{12}$$

It is necessary to calculate the position of the motor and the output shaft of the wave gear, since only their difference is calculated, given to the output shaft of the wave gear. The position of the motor is the integral of the motor speed, so we need to find the antiderivative of $\dot{X}$.

$$\dot{Y} = F\dot{X} \tag{13}$$

The matrix F need to receive only the integral of speed, so we fill diagonal only 2 and 3 line

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{14}$$

$$X(t) = X_{max} - e^{At}(X_{max} - X_0) \tag{15}$$

$$Y = F\big(X_{max}t + A^{-1}(e^{At} - E)(X_0 - X_{max})\big) + Y_0 \tag{16}$$

Each mode has its own system of linear differential equations with constant coefficients. The algorithm of the discrete model is as follows:

The calculation of new moments of motor load and wave gear ($M_{load_{k+1}}$)

Selected model mode:

$$r_{k+1} = S(\mathrm{X}, M_{load_{k+1}}) \tag{17}$$

r_{k+1} – mode,
S – mode calculation function

The friction torque is determined according to the current mode. From the specified voltage, the friction torque and the load torque on the motor is formed by the input vector U_{k+1}.

Calculated the increment of phase coordinates

$$\Delta X = A_{r_{k+1}}X_k + B_{r_{k+1}}U_k \tag{18}$$

Calculated the values of phase coordinates

$$X_{k+1} = X_k + \Delta X \tag{19}$$

Calculated the rotation angles

$$Y_{k+1} = FX_{k+1} + Y_k \tag{20}$$

To verify the obtained discrete model in the MATLAB Simulink system, a prototype of the model was developed, which implements a complete mathematical model of the mechatronic module taking into account all elastic and nonlinear interactions and its discrete model with a switchable structure.

The model and the prototype showed similar results under the same effects. Speed misalignment of the output shaft of the wave gear $5{\cdot}10^{-5}$. The calculation time of 1 second of the model took 0.85 seconds with a sampling time of 250 μs.

Next, the verified discrete model of the drive dynamics is translated into the "C" language. This makes to integrate the dynamics model into the simulation model. The calculations are significantly accelerated. Thus, all other things being equal, the calculation time decreased from 0.85 to 0.016 seconds. The simulation model can be easily integrated into top-level control systems. We replaced the real drive with a simulation model, leaving all the software. This allows us to develop and explore software in near-real-world environments.

To verify the simulation model, an experiment was carried out in which the same input signal was applied to the real mechatronic module and to the simulation model. We gave a triangular signal to the input (see Fig. 2). The simulation model and the real module worked out the position and speed on the output shaft of the wave gear with negligible error.

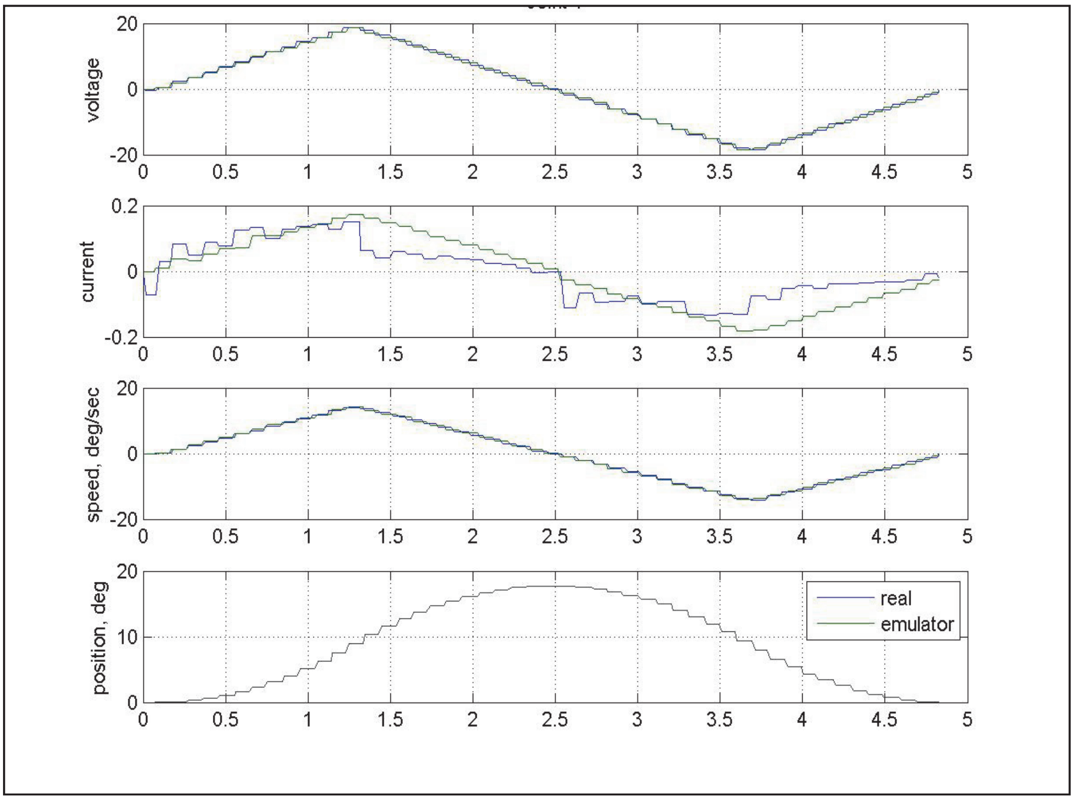


Figure 2 – Compare of simulation model with real mechatronic module

As a result the model of mechatronic module compatible with real software was developed, which simulates the operation of the real module with tolerable accuracy. This makes possible to configure the control system, check its performance with various parameters, check the reaction of the mechatronic module in emergency situations. Ability of simulation model to work in real time allows to use it for complex modeling and testing software of robotic systems before full-fledged experiments.

I.N. Bubnikov[1,2], A.N. Yusupov[2]

SAFETY OF ROBOTIC SYSTEMS IN EXTREME CONDITIONS

[1] Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia
[2] The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
i.bubnikov@gmail.com, a.n.yusupov@gmail.com

Abstract

The robotic system operates in extreme conditions. This means that the output of certain system parameters is at the limits of permissible values, up to the violation of technical requirements. Under these conditions, system reliability indicators are no longer relevant. A relevant indicator is the safety of the system. Procedures for detecting hazardous situations are carried out and methods for their prevention are defined to determine the specific safety requirements. Each hazardous situation is associated with the violation of safety requirements imposed on a system element or on a step of an operation. For robotic manipulator operating in an extreme mode, the identified warning methods will be ineffective from the operator's point of view. There are a lot of indicators that must be monitored to prevent a hazardous situation (tens or even hundreds), and the time to make a decision is severely limited (hundreds of milliseconds). However, as a result of a series of procedures, it was possible to obtain specific safety requirements and descriptions of the conditions under which such violations become possible. The obtained data is proposed to be used to calculate the safety index. The safety index is an integrated evaluation of the situation that is displayed to the operator. The value of the index is the result of numerous simulations of the system elements work with certain variations of significant parameters. This evaluation will allow the operator to quickly make a decision when working in an extreme mode. Thus, the key goal of the work is to determine the possibility of implementing this approach.

Keywords: robotics, extreme conditions, safety.

Acknowledgments

This research is supported by the Ministry of Science and Higher Education of the Russian Federation. State assignment at the Russian State Scientific Center for Robotics and Technical Cybernetics № 075-00924-19-00.

Introduction

A number of tasks solved by robotic manipulators are associated with work in extreme modes. In addition to providing the basic functionality, it is required to ensure the safety of the robot and its environment. For this, it is necessary to formulate the safety requirements for the robot and explore ways to ensure them. The robot safety issues relate to the interaction of the robot and the person acting as an object of the robot's activity (for example, a medical robot manipulator and a patient). The safety of robots in extreme conditions remains an open question. This work is aimed at determining the safety concept of a robotic arm in extreme conditions, developing an approach to formulating security requirements and maintenance of it.

Features of the extreme modes

Extreme modes of robotic manipulators operation have their own specific features. Firstly, the damage from the failure to complete a given task is comparable or even higher than the damage from the loss of the robot. Examples of such tasks are an accident liquidation, people rescue, tasks for repairing space stations, on which the safety of astronauts depends, tasks of moving dangerous loads. The robot may fail, but it must solve the problem. Secondly, the deliberate use of the robot at the boundary of operating conditions. Work in extreme conditions involves the output of the robot parameters beyond the limits specified in the documentation – increased moments, long cycles of continuous operation, work in conditions of high/low temperatures, the excess level of exposure by the harmful environmental factors (radiation, the chemical composition of the environment, humidity).

In such modes, the reliability of the robot is not an indicator on which the operator can rely when choosing an operation algorithm. The operator's goal is a safety assurance of the robotic system. The probability of the successful implementation of tasks assigned to the robot depends on how long the safety will be maintained. By safety we mean the following:

1) a robot safety for the environment and the working area;
2) a robot safety for itself.

The first is determined by the preservation of the objects and the health of people in close proximity to the robot, the lack of movements that are not obvious to the operator, which may cause him to react inadequately to what is happening. For example, work on the boundary of the working area, when for a small displacement of the output link it may be necessary to move several links with high speed to change the configuration of the robot.

The second includes the structural integrity of the robot (no mechanical damage) and its electrical part – drives, power converters (no faults due to exceeding the permissible limits of currents or voltages).

The time the operator performs the main task and the results of its implementation depends on the time during which the goal is still being achieved. The failure of the robot, which does not allow the operator to complete the task, is unacceptable. At the same time, the robot works on the boundary of the operating conditions. These factors determine the extremely high, and sometimes incompatible with the human capabilities, requirements for the operator in terms of decision making speed and quality of a decision. In addition to performing the main task, the operator must continuously ensure the safety of the robot.

Safety requirements and HAZOP procedure

To present safety requirements for the characteristics of various parameters of the robot, it is necessary to determine which states of the robot should be considered hazardous. For this, a hazard and operability study (HAZOP) is conducted [1, 2]. HAZOP – is a process of detailed and structured identification of hazards for individual technological systems (sectors, nodes). This method is preferable at the completion stage of the project development when the main constructive and technological solutions are worked out. At present, a number of manipulators with a similar structure have already been developed, which allows us to consider the robotic manipulation system as a finished product [3, 4].

During the HAZOP procedure, the manipulation system is divided into separate components, the deviations of the characteristics of which must be considered using control words. Such division into the elements is presented in fig. 1.

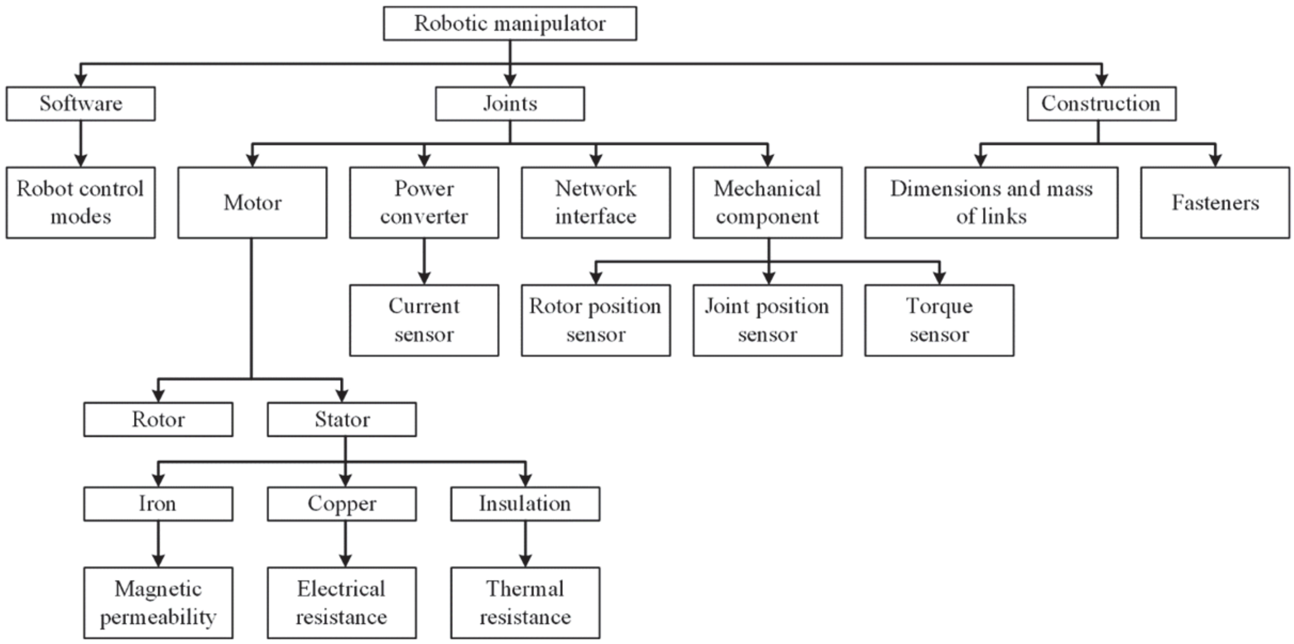


Figure 1 – Division of the system of the robotic manipulator into the elements

For example, one of the most common causes of a robot hazardous condition is chosen – an overheating of the stator winding insulation. If the robot is in such a hazardous state, the result will be either a failure of the power converter or failure of the motor. Both cases deprive the robot of the main functionality and make it impossible to continue working with it. A fragment of the HAZOP worksheet for this hazardous condition is presented in tabl. 1.

During the procedure, it was found that there is no way to prevent an increase in the insulation temperature. The HAZOP group proposed the following possible solutions:

1) an installation a temperature sensor on the insulation;
2) a power consumption measurement for determining the thermal mode of a motor;
3) a use of the temperature observer for motor winding.

Installing an insulation temperature sensor reduces reliability, increases system complexity, adds another element that requires research. In addition, the motor often does not provide for such modifications, and installing a sensor outside the insulation will introduce a delay, which will make the detection system useless.

Table 1. A fragment of the HAZOP worksheet

Control word	HIGHER
Element	The temperature of the stator winding insulation
Error	Exceeded insulation temperature
Possible reasons	Violation of the motor thermal mode by the operator
Effects	– Short circuit in the power converter – Motor failure – Electric shock to persons
Existing safety measures	Do not exist
Assessment of the situation	Unacceptable
Action required	Install an insulation temperature sensor or use a motor temperature observer or an algorithm for determining the thermal mode of a motor based on the power consumption

Measurement of the motor power consumption allows evaluating the motor thermal mode, in particular, the temperature of the motor windings, without changes in design. However, in case of a three-phase motor, this option may not be very informative – the current can be distributed between the phases evenly and can flow through one, leading to local overheating.

The use of a motor phase temperature observer allows estimating the temperature of each phase based on the current that flows in it. In this case, additional structural elements are not required, and the achieved result is comparable with the installation of the temperature sensors on each phase. This option is considered the most acceptable.

If one of the proposed options is implemented, the temperature of the motor windings of the joint will be displayed to the operator who can determine the order of further actions based on the requirements for maximum temperature (for example, change the configuration of the robot so that the heated joint becomes less loaded).

However, it should be noted that the robotic arm usually consists of 6-8 joints. This means that the operator needs to monitor the temperatures of each winding of all motors, this is 18-24 parameters, in parallel performing of the main task. When working in extreme mode, the temperature of most joints will be at the boundary of the maximum allowance. In this case, the operator will not be able to determine which configuration of the robot to choose in order to accomplish the task. The presence of a temperature indication will no longer allow him to determine adequately the order of further actions.

Safety index as an integrated evaluation

An approach to determine the integrated evaluation of the situation (the safety index) is proposed. The main tool for determining the safety index is the simulation model. During the development of robotic manipulators, sufficient experience has been gained in working with simulation models; therefore, they were chosen for the safety task.

The main idea of the approach is as follows: the current parameters of the robot are determined in real-time. Then the system is simulated for the near future in the most likely states. For example, it can be the reaction time of the operator. After that, the number of safety requirements violations is determined. The conditional frequency, which determines the ratio of unsuccessful outcomes to all the simulated variants under given initial conditions, is called the safety index.

The proposed approach can be described by the following algorithm:

1) an identification of condition/goal combination;

2) a statistics set – multiple simulation modeling of a robotic system with appropriate variations of operating conditions for a given period of time;

3) an approximation of the function of the conditional frequency of safety requirements violations;

4) determining the safety index – the ratio of the number of negative outcomes to all the results of the simulation model.

A condition is a set of current characteristics of the robot parameters, a goal is the state of the robot defined by the operator. The choice and variation ranges of parameters under given conditions and goals are determined by a group of experts by brainstorming.

Consider the application of the proposed approach by an example. Suppose the simulation model is a thermal model of the stator windings. The initial parameters of the model are:

1) the temperature at the beginning of calculations;

2) thermal mode of the motor at the beginning of calculations;

3) reference values of thermal resistances.

The task of the simulation model is to calculate the temperature of the stator winding after a fixed period of time with the variation of thermal resistance and thermal mode of the motor within the possible range. After that, the calculated temperature is compared with the allowable maximum for insulation of the stator winding and the number of outcomes leading to motor failure is determined.

The expert group determines the parameters, the variation of which may affect the characteristic to which the safety requirements are imposed. To determine the parameters, the following situation was considered: a six-link robotic arm maintains the position of the gripping device under the action of high torque. The configuration of the robot is shown in fig. 2.

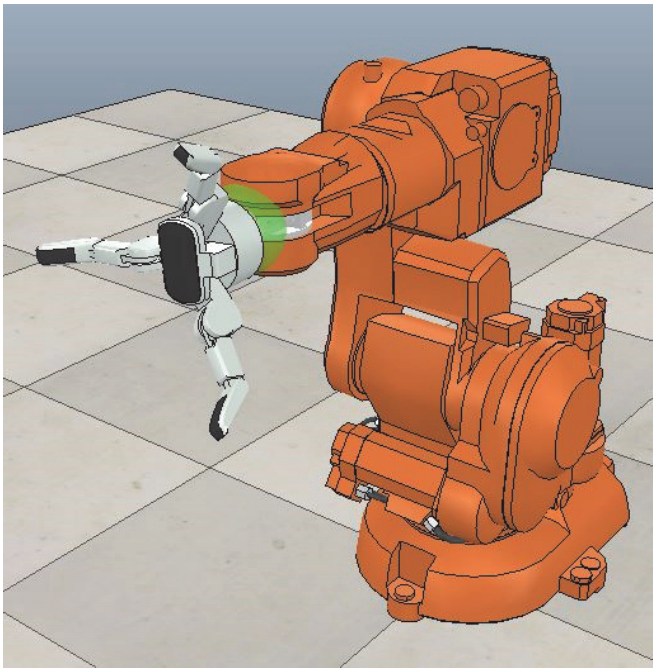


Figure 2 – The position of the robot for which the calculation was made:
the goal is to hold the position of the gripper under the external moment
(the direction of the action is shown by the arrow)

If in this case, the operator keeps the links of the manipulator stationary, then in the motors of revolute joints, compensating the moment, the current will constantly flow through one winding. This can lead to a rapid increase in the temperature of this winding and overheating of the insulation. It was suggested that if the coaxial links rotate with equal speed in opposite directions (see fig. 3), the heating of the windings will be longer and more uniform. Simulation has shown that the assumption is true. Fig. 4 shows the heating of the windings at different motors speeds.

The speed parameter is one of those, which variations affect the characteristic of the parameter to which the safety requirements are imposed. In addition to speed, the temperature of the windings will be influenced by the change in current, since it is the flow of current through the winding that causes its heating. A group of experts during a brainstorming determined the ranges of the most probable variations of parameters: current – from 20 A to 30 A, speed – from -0.5 rad/s to +0.5 rad/s.

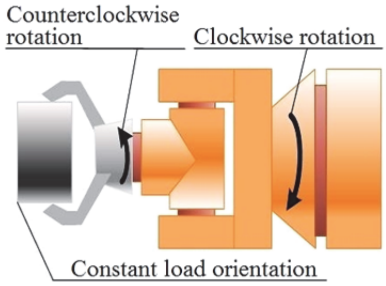


Figure 3 – The choice of variable parameters

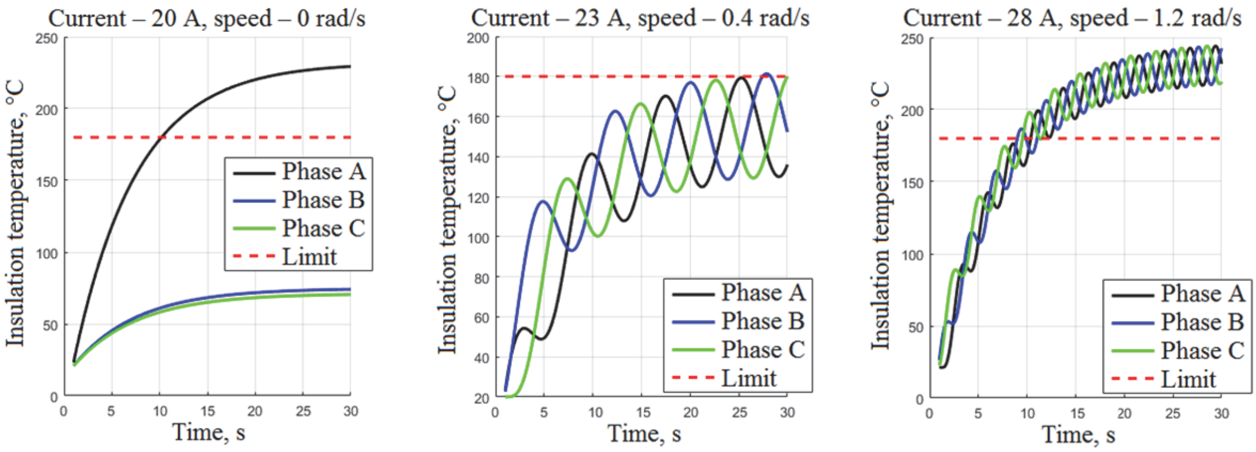


Figure 4 – The effect of speed variation on the phase insulation temperature

On a certain space of parameter variations, a simulation was carried out with a subsequent approximation of the conditional frequency function and a safety index was determined. The results are presented in fig. 5. Calculation parameters were as follows:

1) the calculation was made for one joint of the robot (temperatures of three windings) for the case of rotation with a speed of 0.4 rad / s and a current of 23 A;

2) time range of the calculation was 25 s;

3) variable parameters for a set of statistics were the speed and the current in the joint motor phases. The range of parameters variation: speed – from -0.5 rad/s to +0.5 rad/s, current – from 20 A to 30 A;

4) simulation time span was 300 ms (approximate reaction time of an experienced operator).

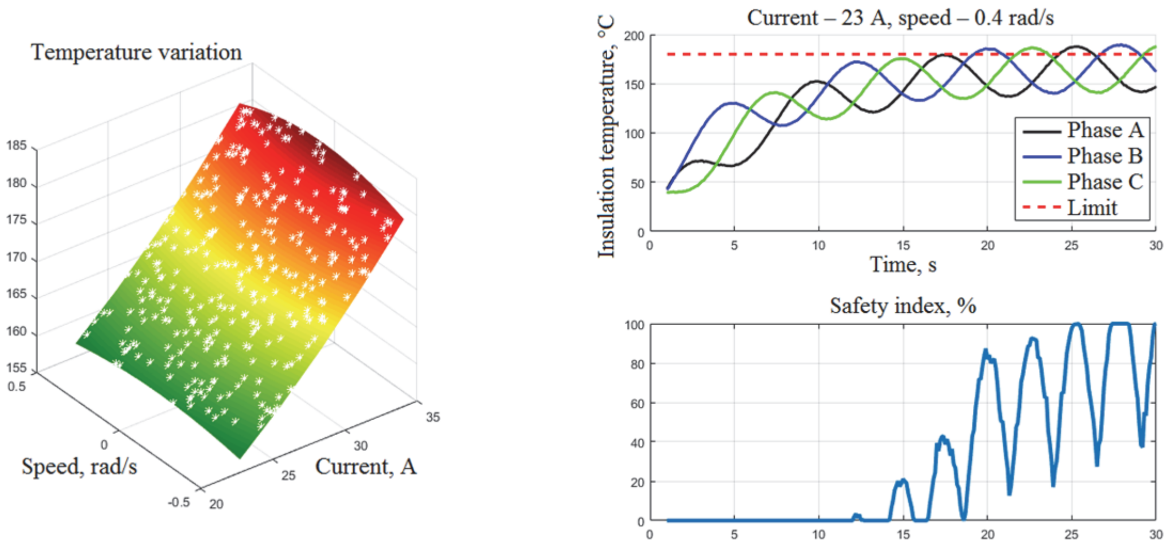


Figure 5 – Approximated winding temperature function
with selected parameter variation range and the calculated safety index

In addition to the elements of the robot system, to which the safety requirements can be applied, various operations performed by the operator can also lead to adverse consequences for the environment and for the robot itself and the operator. Therefore, it is necessary to identify possible hazardous conditions and put forward safety requirements for specific phases of the operation.

Perform a HAZOP procedure for the operation. To do this, it is necessary to divide it into elements, choose the parameters of the elements and determine the characteristics of these parameters. After that, for each of the characteristics, apply control words to determine possible hazardous states. As an operation, moving of a gripping device of the robot with a given speed was chosen. For the case of operator control, this type of movement is one of the most common. The decomposition of the robot movement process is shown in fig. 6.

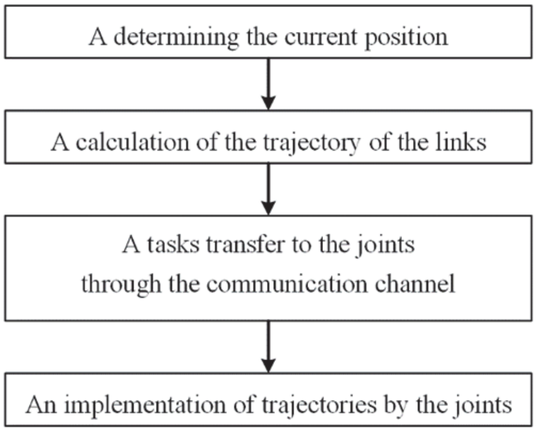


Figure 6 – Elements of the operation «Moving a gripping device of the robot»

Partially carried out the HAZOP procedure. The result is presented in tabl. 2.

When working on the boundaries of the working area, the movements of the manipulator are not always obvious to the operator, so it is not always possible to visually determine the possibility of an accident.

Table 2. HAZOP procedure for the element «Determining the current position»

Control word	NOT LIKE AS
Element	Determining the current position
Error	The actual position does not match measured
Possible reasons	Plastic deformation of links under load, thermal expansion
Effects	When working on the border of the working area, surrounding damage – jamming the robot, the failure of the joint
Existing safety measures	Choice of the working area, taking into account possible deformations of the links, protection against over-torque of the link (under conditions of extreme operation, these measures are often neglected)
Assessment of the situation	Unacceptable on the boundary of the working area
Action required	Does not exist

Thus, the following should be taken as a safety requirement: deviations of the real position from that determined by the control system should not lead to a collision of the manipulator with those objects which positions are known to the control system (including collision with itself).

After the safety requirements are presented, the algorithm proposed earlier is executed.

Firstly, conditions and goal are identified. Conditions are the current positions of the links of the manipulator located on the boundary of the working area. The goal is to move the gripper with a given speed.

Then, a group of experts determines a set of parameters which variations may affect the characteristic of the element and the setting of the range of these parameters variation. The dimensions of the links are considered as such a parameter. Characteristics of this parameter are length, width, height of links (in meters), moments of inertia of links (in kilograms per square meter). For these characteristics, possible variations are chosen. The simulation model for this case is a geometric model of the robot.

As a result, it becomes possible to implement the proposed approach in two different cases. The approach allows calculating a common safety index for the case of several hazardous states, using several simulation models. The calculated safety index should be displayed to the operator as a single indicator that determines the level of hazard at the current time, taking into account the permanence of the goal. It can be converted into various risk assessments or can be applied by the operator for a comprehensive assessment of the need to continue the work of the robot, taking into account the factors known to him.

The aim of further research is to check the possibility of implementing the approach for a larger number of special cases, checking the adequacy of combining the results and implementing the approach on a real robot.

Conclusion

The possibility of implementing an approach to ensuring the safety of the robotic arm has been determined. The proposed approach allows calculating a single safety index for the case of several hazardous conditions, using several simulation models. The display of the safety index allows the operator to pre-detect hazardous situations. The purpose of further research is to verify the feasibility of the proposed approach for a larger number of special cases, to verify the adequacy of the merging of the results, and to implement the approach on a real robot.

References

1. GOST R 51901.11-2005 «Risk management. Hazard and operability studies. Application guide».
2. Ibadulaev V., Stepanov I., Turusov S. Experience of creating decision support systems in emergency situations // Monitoring. Science and Technologies. 2014. № 2. Pp. 14-31.
3. Shardyko I.V., Yusupov A.N. Implementation of stiff and compliant control for joint of space manipulation systems // Robotics and Technical Cybernetics. 2018. № 4. Pp.60-67.
4. Shardyko I.V., Titov V.V. A closed-form solution of IK task for a 6-DOF manipulator with pitch axes offset and a technique of fast joint space trajectory computation // Proceedings of the International Scientific and Technological Conference «Extreme Robotics». November 2-3, 2017, Saint Petersburg. Pp. 23-29.

A.N. Kosenko, D.M. Korolev, O.A. Shmakov

METHODS OF USING MODULAR CIRCUIT UNITS FOR MOBILE ROBOTIC SYSTEMS DESIGN

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
kosenko.ank@yandex.ru, d.korolev@rtc.ru, shmakov@rtc.ru

Abstract

The object of the research is mobile robotic systems built using modular units. The aim of the work is to substantiate the effectiveness of building a control system for robots for various environments based on a modular basis.

Keywords: modular unit, robot, control system.

Acknowledgments

This research is supported by the Ministry of Science and Higher Education of the Russian Federation. State assignment at the Russian State Scientific Center for Robotics and Technical Cybernetics № 075-00924-19-00.

Introduction

One of the important features in the development of mobile robotic systems is the possibility of simplifying the development or unification of the systems. Unification means that in different systems there are common interchangeable blocks. In addition, it should be possible to expand or change the functionality of the system depending on the tasks assigned to it. The possible solution is the use of modular units as part of a robotic system. Various modules can be interchangeable to perform specialized tasks, as well as common within a robotic platform and control system. The aim of the paper is to develop the structure of robotic systems for various environments (ground, air, and sea) using a modular approach.

Modular approach in the control system development

The use of modular units is one of the approaches in the development of a control system for robots with the possibility of application in different home-based environments – ground, air, and underwater. Modularity provides a single common approach in the design, which greatly simplifies the workflow [1].

When designing a device, several basic points can be highlighted. Firstly, it is necessary to form a technical task that describes in detail the required functionality of the device. Secondly, it is necessary to develop a functional diagram of the device. Thirdly, it is necessary to develop a schematic diagram of the device. The next steps are the development of printed circuit boards, the purchase of components, installation, assembly, debugging, and so on.

The first two steps give a general understanding of what the developed device is. Directly at these steps, the common nodes that are present in individual parts of the device are already visible. Accordingly, at the step of the schematic diagram design, there are two possible approaches. First, it is possible to combine all nodes into groups and to develop a separate printed circuit board with exclusively required functionality for each. This approach has several advantages and disadvantages. The advantages of this approach are found during the development of large or medium-sized batches of devices – because the order of large batches of boards and their components is cheaper than purchasing a single device.

The second approach is to use of modular construction of the device. The essence of this approach is to use separate modules – nodes that are common in the development of various devices. In particular, in the development of mobile robotic systems, such common components can be distinguished: a computer, an autonomous power supply, an electric drive controller, and others.

On the one hand, several drawbacks can be identified in this approach. First, a large number of modules in one device requires a large number of cable connections, which can lead to errors when installing the product. Secondly, such modules should be unified, which is both an advantage and a disadvantage. This means that the board should contain the maximum possible functionality that may be required to solve various tasks. For example, a module that must switch between the interfaces UART and RS485, RS232, CAN, etc. The board must contain all types of converters, there must be input connectors for connecting to all these interfaces, but at the same time, there are no guarantees that under the conditions of this task all interfaces will be needed. Thus, the board will have dimensions that are larger than required and have extra components on it. However, the modular approach has one major advantage (especially if the disadvantages of this approach can be reduced). The use of modularity can significantly simplify the development process of the device and significantly reduce the time spent on development. In the ideal case, the time for the development of printed assemblies is zero (all the required modules have already been developed in advance). In addition, the use of the distributed modular structure of the control system will make it possible to reduce the price and expand the possibilities of upgrading such system by replacing existing modules or adding new ones that increase the functionality of the robot.

The development of a robotic system for different basing environments

The most obvious is the modular approach in the development of robotic systems for various environments – ground, air, and sea. Directly in each robotic system of its home environment, there will be common elements – a computer, an autonomous power supply, an electric drive controller and others. Further, each of these modules is considered in more detail.

The computer is a component of the control system that provides centralization of computational operations in the process of control a separate robot. The basis of this device is the motherboard of the robotic system, the key node of which is a microcontroller.

The tasks of this unit include:

– implementation of centralized control of individual nodes of the robotic system,

– ensuring and monitoring the status of a power supply of individual components of the robotic system,

– formation of a data bus, switching signal lines between the nodes of the robotic system.

To perform these tasks, the device is equipped with the following components:

– connectors for connecting an autonomous power supply, functional blocks and microprocessor debugging,

– LEDs indicating the normal operation of the connected functional blocks,

– smart protection switches on the power of the connected blocks,

– interface converter, which forms the internal data transmission bus,

– programmer based on a microcontroller.

The autonomous power supply provides electricity for all functional units of a mobile robot, ensuring a given autonomy duration, high current output to enable simultaneous activation of the maximum number of functional blocks, as well as protection of the power supply in terms of voltage, current, and temperature. When choosing a microcircuit that charges an autonomous power supply, it should be guided by such parameters as the supported number of cells, cell types and the presence of built-in cell protections. When choosing a microcircuit that monitors and controls the battery pack, it is necessary to provide a possibility of independent measurement of the battery charge level, along with the process of charging and discharging, to protect from emergency situations.

Also, an important component of the robotic system is the motor controller. At present, the overwhelming majority of electric drive modules presupposes the use of brushless electric motors due to their power parameters (from units of watts to hundreds of watts) and the possibility of implementing flexible control modes. The absence of the collector reduces the weight and size of this design, and also reduces interference. The design of the brushless motor allows its use in water and aggressive environment [2]. Accordingly, a single controller module allows widespread use of brushless motors. This module contains the control part of the circuit (microcontroller, power bridges, power sources) and is a universal platform for controlling drives, equipped with the interfaces shown below:

– two SPI channels,

– two incremental encoders,

– absolute position sensor

– Hall sensor,

– CAN,

– UART,

– analog inputs/outputs of general-purpose,

– digital inputs/outputs of general-purpose.

Another necessary node of the robotic system is a unit of audio-video surveillance equipment. This unit allows monitoring the surroundings the robot. The unit itself can be divided into separate modules that perform their tasks. It is supposed to use the following elements as modules of the audio-video surveillance equipment unit:

– microphone module,

– a standard video camera module

– a zoom camera module,
– a thermal imaging camera module,
– switching module.

The microphone module allows audio monitoring with support for hardware compression of the sound range. A standard video camera module includes a video camera and a filtering and control board. This module is designed for general monitoring of the environment in conditions of sufficient illumination. For example, camera BHV1000 can be used as a standard video camera. The zoom camera module includes a zoom camera and filtering and control board. A camera WONWOO EM-363S [3] can be used as a zoom camera. A feature of this camera is the ability to remotely adjust the focus and zoom. The thermal imaging camera module includes the thermal imaging camera itself and the filtering and control board. A camera IRIDIUM-640 [4] can be used as a thermal imaging camera. The above cameras allows full audio and visual monitoring of the surrounding area, regardless of the light and time of day.

It is also necessary to provide a single module to control the audio and video information received from several video cameras for further transmission to the radio channel – the switching module. For this, a video processing board is used, which additionally improves the quality of the video signal and the noise immunity of the video paths. In addition, this module is equipped with individual power supplies for each type of video paths and the physical separation of video lands from digital ones using an in-phase choke.

The GPS beacon module is designed to track the position of the robotic system in case of an emergency situation. GPS beacon is a receiver of the global positioning system, coupled with a redundant data channel.

As transceivers for the control channel and the video transmission channel, it is mostly efficient to use the purchased system, for example, «Cordon», JSC «SET-1» [5]. The radio modem from the Cordon system is intended for two-way data transmission (remote control commands, telemetry, device position, etc.) via a digital radio channel as part of the control equipment of the robotic systems. This radio channel supports the necessary functionality from the point of view of data transmission and provides a sufficient reserve of the control of the robot. An effective solution is to use it in conjunction with the video transmission channel if the second integrated control channel is in place to increase system reliability and provide redundancy.

For transmitting video and telemetry, it is proposed to use a transceiver from the «Luch» system [6].

Inside the robotic system, it is supposed to install video cameras with an analog video output that are switched on the video switching board and then transmitted in analog form to the Luch device. The microphone module is connected to a separate video transmitter connector. Then the analog video signal is digitized and already in digital form is transmitted to the receiving part of the Luch device, where it enters the recording device and is displayed on the display, as well as to the headphones. The resulting video signal resolution will be in the region of 450-500 television lines.

As noted above, one of the most obvious methods for developing a robotic system using a modular approach is the development of robots for various home-based environments – ground, air, and water. Figures 1-3 show the developed structural diagrams for the design of the corresponding robotic systems.

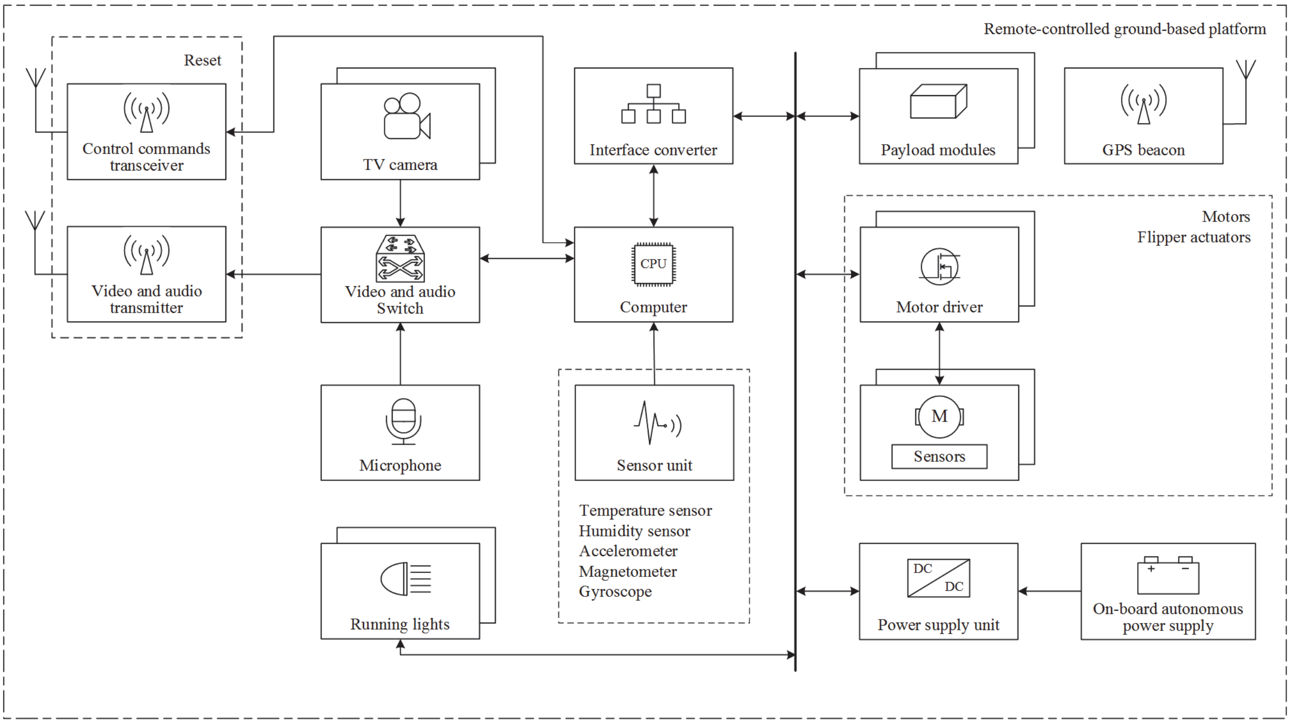


Figure 1 – Structural diagram of the control system of the ground-based robot

The control system of the ground-based robot includes the following modular elements:

– computer,
– autonomous power supply,
– lighting unit,
– audio-video surveillance unit (course cameras, microphone),
– motor controllers,
– GPS beacon,
– radio transmitter.

In addition, the robotic system is equipped with a number of non-modular solutions in terms of sensing and drives. If necessary, it is allowed the integration of various modules of the payload. To implement the motion control of the robot, it is supposed to use two groups of drive systems: the motors control system and the flippers control system to overcome obstacles.

The drives of these systems consist of modular controllers, brushless electric motors, gearboxes and auxiliary electronic devices, such as encoders, Hall sensors, absolute position sensors.

To ensure the correct operation of the onboard electronics of the robotic system, microclimate control (temperature and humidity) is necessary – microclimate sensors are provided for this function. In addition, inside the robotic system, there should be a unit that performs navigation functions, as well as a warning about undesirable states, such as capsizing. Inertial sensors are used for this task. The two groups of sensors described above can be combined into one separate module – a sensor unit, the main task of which is to detail the state of the system and provide this information to the central computing device.

To extend the functionality of a robot, payload modules are used. In particular, the payload modules of the ground-based robotic system can perform the following functions:

– delivery of payload,
– moving objects,
– audio-video surveillance,
– the implementation of technological operations (drilling, cutting, etc.).

The control of these modules is also carried out using a computer. Audio-video data can be transmitted directly to the switching device with the appropriate configuration of the module. As a payload module, directional microphones, cameras of various spectra, electromechanical locks and grippers, manipulators, a set of interchangeable instruments, etc. can be used.

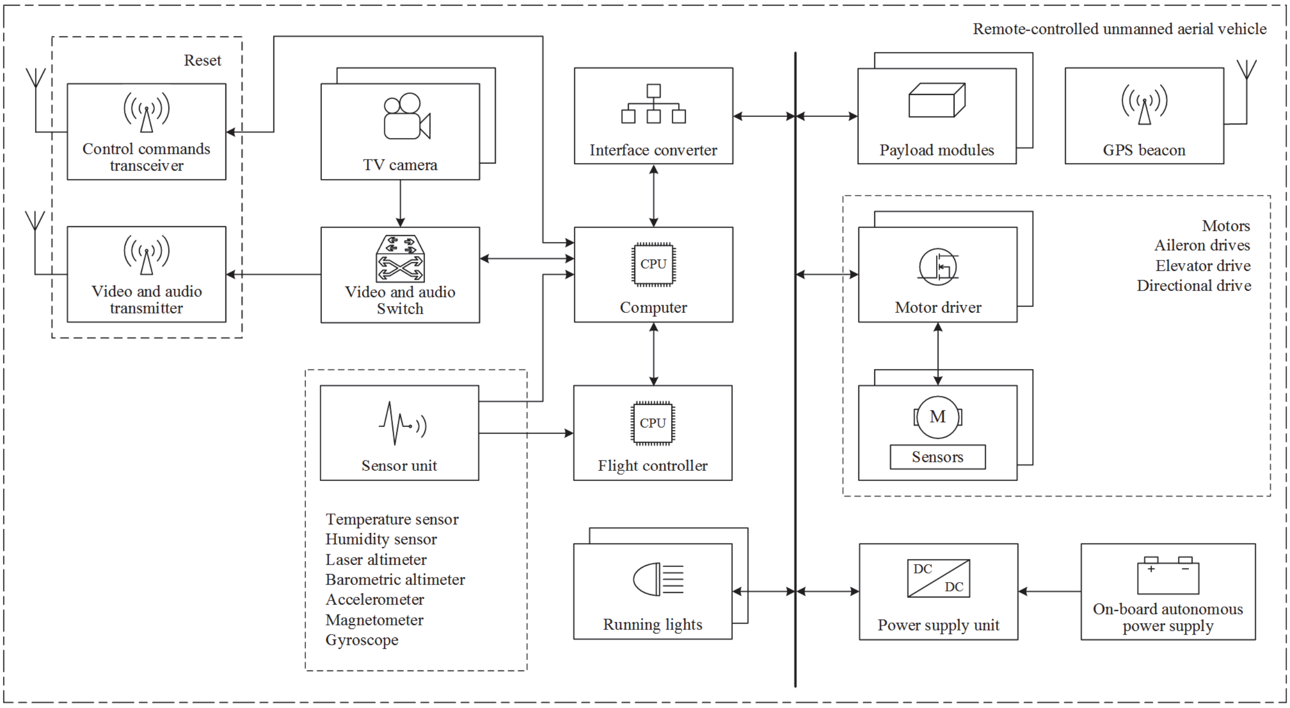


Figure 2 – Structural diagram of the control system of the air-based robot

The air-based robotic system is equipped with a number of non-modular solutions in terms of sensing and drives, as well as a specialized flight controller. In particular, the tasks of this controller include monitoring and retaining a flight altitude, which requires a continuous workload of computing resources. In this regard, the control of the flight component of the movement is carried out by a specialized flight controller, processing the flight data from the sensor unit and issuing corrective commands to the motor controllers in order to maintain a given mode of motion. The module of the sensor unit must be expanded – it must include an altimeter unit. If necessary, it is allowed to integrate various modules of the payload. In particular, the modules of the payload of the air-based system can perform the following functions:

– delivery of payload (including discharge),
– remote sensing,
– audio-video surveillance.

The remaining modules are similar to ones used in the ground-based robotic system.

The general structure of building an underwater robotic system is practically the same as the above-described ground- and air-based robots. The underwater robotic system is equipped with a number of non-modular solutions in terms of sensing and drives, as well as a specialized controller of a remote-controlled unmanned underwater vehicle. If necessary, it is allowed to integrate various payload modules to this platform. For example, a manipulator for cargo delivery. If it is necessary to operate equipment with a narrow zone of permissible microclimate, a system for ensuring thermal conditions (cooling, dehumidification, heating, etc.) is being installed.

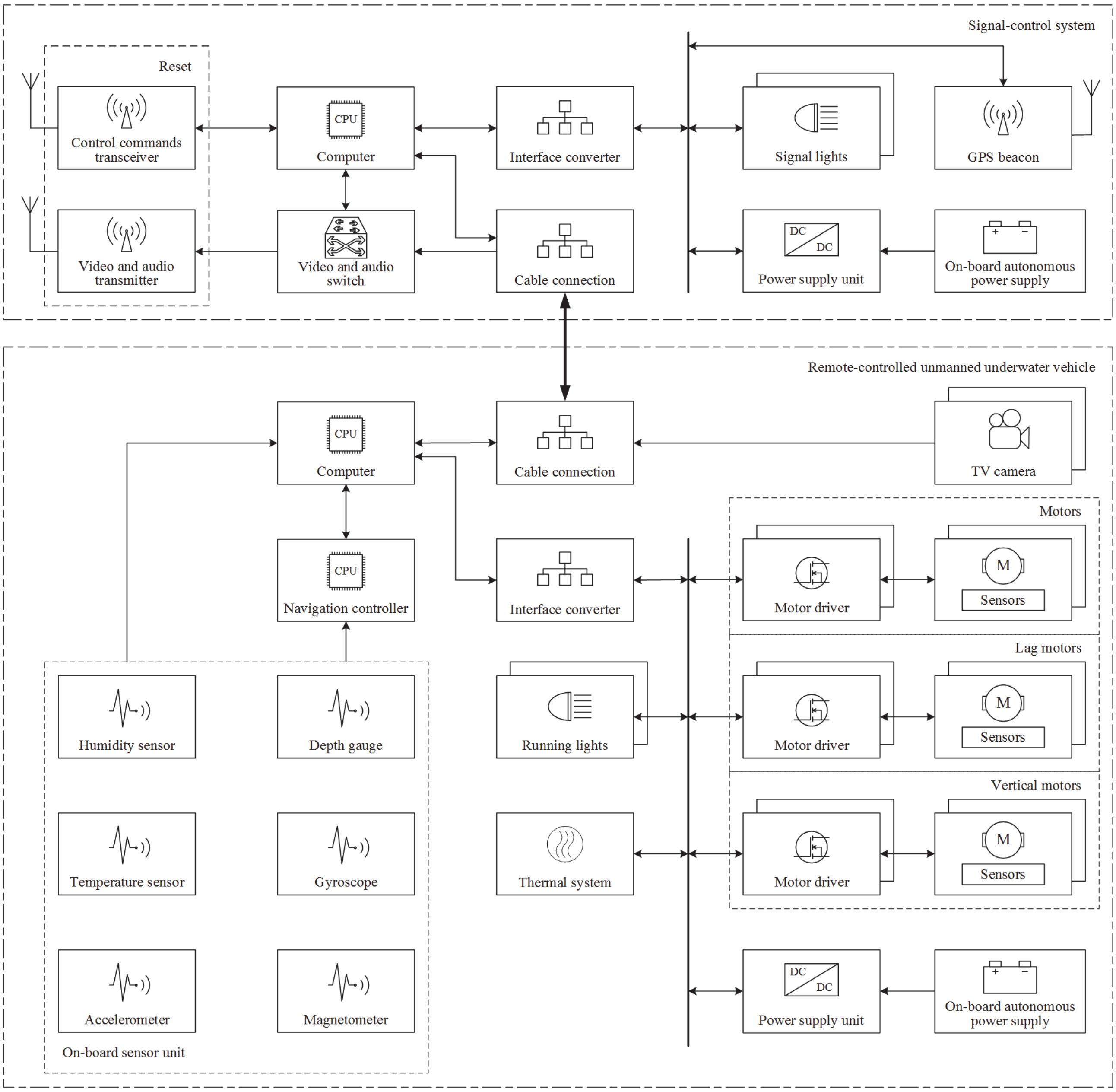


Figure 3 – Structural diagram of the control system of the underwater-based robot

As can be seen from the above description of robotic systems of various home environments, these systems can be developed using specialized modules common to each system (for example, a computer, a motor driver module, etc.). The use of these modules greatly simplifies the process of developing a robotic system in terms of electronics, it takes much less time compared to the approach, when all boards are designed individually for each node. In addition, it reduces the cost of printed circuit assemblies of the robotic system, since the purchase of a single set of components and single boards will be cheaper than the purchase of separate various boards with separate sets of components for each of them.

In addition, the use of a modular approach allows to significantly expand the functionality of the robotic system by adding payload modules. In particular, any audio-video surveillance module can be installed on a robotic system depending on the situation and the required functions – a standard camera, zoom or thermal imaging camera. The manipulator, sensors, and other modules that are required to perform the tasks set in this situation can also be installed. Thus, all the previously described shortcomings of the modular approach, such as excessive functionality and a large number of connecting wires, are leveled by the advantages that the use of modules provides.

Conclusion

The use of a modular approach seems to be the most practical and correct solution for the development of mobile robotic systems. It allows to simplify the process of the development, to avoid design errors due to the use of proven modular solutions, and also to expand the functionality of the system through the use of specialized modules for solving specific problems. The paper shows the advantages of using a modular approach by the example of developing the structure of a robotic system for different basing environments (ground, air, and sea).

References

1. The concept of modular device design. URL: http://kmpu.ru/theory/index.html.
2. Brushless DC Motors. URL: http://www.avislab.com/blog/brushless01/.
3. WONWOO official site. URL: http://www.wonwoo.com/.
4. Thermal iridium modules with optics. URL: http://astrohn.ru/thermal_modules/iridium_modules_with_optics/iridium-640/50.
5. «Kordon» video receivers. URL: https://www.set-1.ru/products/videosistemy/cordon-videopriyemniki/.
6. «Luch» radio modem. URL: https://www.set-1.ru/products/dopolnitelnoe-oborudovanie-dlya-bla-i-rtk/.

M.A. Nogin[1,2], A.L. Korotkov[2], O.A. Shmakov[2]

METHODOLOGY FOR THE QUALITY ASSESSING OF THE OBSTACLE OVERCOMING BY MOBILE ROBOTS

[1] Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia
[2] The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
m.nogin@rtc.ru, a.korotkov@rtc.ru, shmakov@rtc.ru

Abstract

The paper discusses the methodology for the quality assessment of the obstacle overcoming by mobile robots of the ultra-light and light classes based on the RTC testing ground. This methodology is proposed for using during the tests to obtain an assessment of traversability, which reflects not only the ability of robots to overcome specific obstacles but also the completeness and speed of overcoming. Such an assessment can be applied to compare the traversability of robots with various design features, which can help identify the strengths and weaknesses of typical robots designs. As a quality assessment of the obstacle overcoming, the term of traversability class is proposed.

Keywords: robot, tests, methodology, mobile robotic system, ground robot, obstacles, quality assessment, class of traversability.

Acknowledgments

This research is supported by the Ministry of Science and Higher Education of the Russian Federation. State assignment at the Russian State Scientific Center for Robotics and Technical Cybernetics № 075-00924-19-00.

Introduction

It is important to test the technical characteristics of mobile robotic systems for companies engaged in the development of them. Long-term operation of robots in real conditions is necessary to fully understand how the mobile platform will function in such conditions. Testing in real conditions often entails many complicating factors that impede rapid research, for example, due to repairs. Testing at a specialized ground simulating operating conditions can simplify the testing process greatly. However, conducting full-length research and obtaining an objective assessment of the technical characteristics of the developed platforms require a methodological research base.

At present, Russian standards for robots traversability testing are being developed in the form of projects, for example, methods for quantitative assessment of the obstacle overcoming [1], which are analogs of foreign methods [2]. In this paper, we propose the use of a qualitative assessment technique, as an extension to quantitative methods applied to test. The described methodology is being developed as part of the test procedure at the testing ground of the Russian State Scientific Center for Robotics and Technical Cybernetics (RTC) [3]. The use of qualitative assessment methods in the context of a flexible test procedure that adapts to the purpose, requirements and actual capabilities of the tested robots can improve the quality and information content of the assessment obtained by the robot during the tests at this ground. The further development of the methodology for testing can make it applicable to any other testing ground.

Test procedure

The test procedure includes a preliminary passage of the robot across all obstacles of the RTC testing ground in the order shown in Figure 1. If the robot does not have a manipulation system, items 7 and 8 are skipped. For robots equipped with manipulation systems, at the step of the reconfigurable room module this item is expanded. The extension is shown in Figure 2. Figure 3 shows the relative position of obstacles within the testing ground.

The actual values of the wheelbase W and the radius of the wheels R of the robot are measured before testing. If the robot has pairs of wheels of different radius, only the larger one is taken into account. Also, for robots with sealed enclosures, the height h of the robot body elements is measured for assessing the traversability of the basin module. The body elements include any elements of the robot shell that may affect the water from entering into the housing and its direct influence on the actuators. Body elements do not include the protruding parts such as antennas, video cameras, lighting devices and manipulation systems with the exception of the root joint.

Based on the purpose of the test robot and the requirements for it, different methodologies for conducting and evaluating the tests are built. After receiving the results of the preliminary passage, on the basis of a preliminary assessment of the actual capabilities of the robot, the methodology is improved and, according to it, the main tests are conducted.

During the preliminary passage, the robot is sent for recharging if the power reserve is more than 75 % unless the robot has already completed the overcoming of the module of the reconfigurable room.

The operator has no more than three attempts to overcome each obstacle during the preliminary passage. At the first unsuccessful obstacle overcoming of the testing ground during the preliminary passage, the operator can use the two remaining attempts, if all attempts fail, the obstacle is considered to be impassable. The operator can also move to the next obstacle after the first unsuccessful attempt. An attempt is considered unsuccessful if the robot loses the possibility of further movement or if there is no progress in overcoming the obstacle within 15 seconds. Before the start of each attempt, the robot's power reserve should be at least 25 % of the maximum.

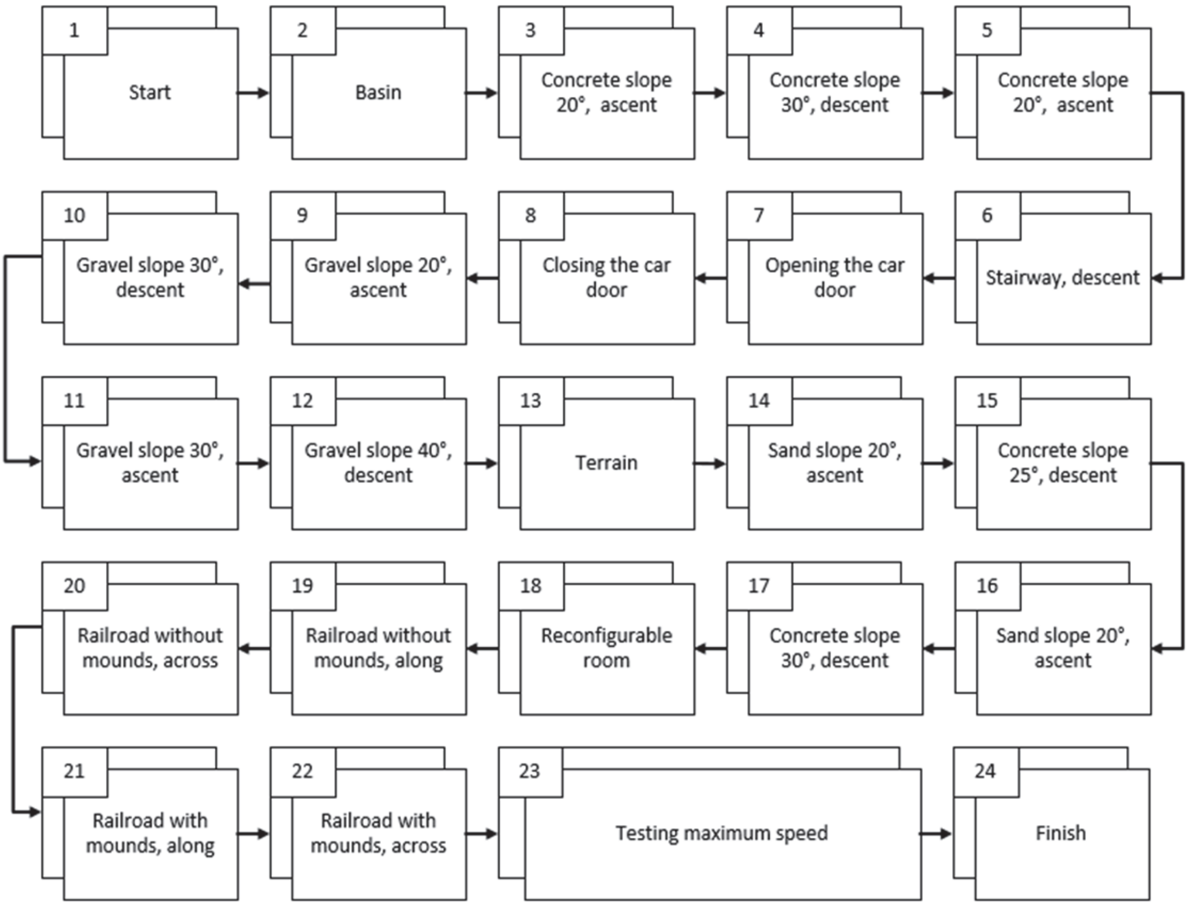


Figure 1 – The order of the obstacles at the preliminary overcoming of the testing ground

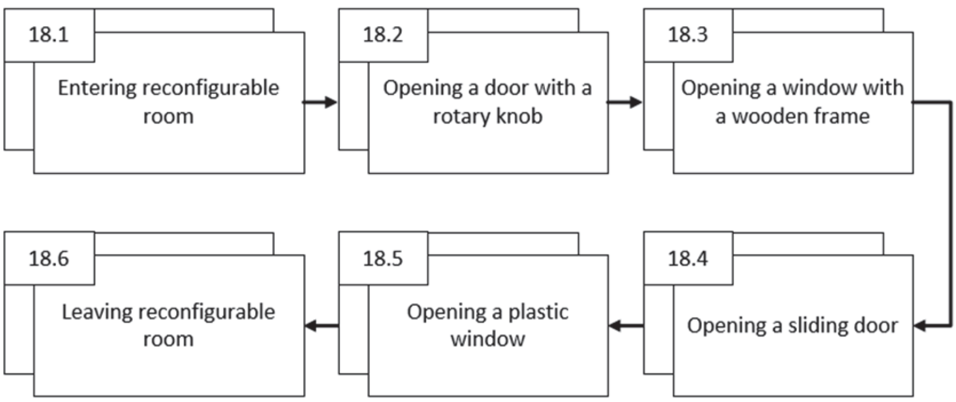


Figure 2 – Expansion of the obstacles sequence for robots with manipulation systems

During the preliminary passage, the average maximum velocity V of the robot movement is measured. It is later used to qualitatively assessment of the speed of obstacle overcoming.

The quality assessment of the obstacle overcoming

To determine the quality of the obstacle overcoming, the concept of traversability class is introduced. It is denoted by the letters TC and the two numbers following them, the first of which X indicates the extent to which the robot overcomes the obstacle, and the second shows how fast the robot overcomes it. If the obstacle is impassable for the robot, then the class TC-00 is assigned to such an obstacle.

The traversability class for each obstacle is calculated individually, the maximum traversability class for each obstacle is TC-44. After assigning the traversability class for each obstacle, the robot is assigned a general traversability class of the testing ground, which is calculated from the arithmetic average rounded up to an integer.

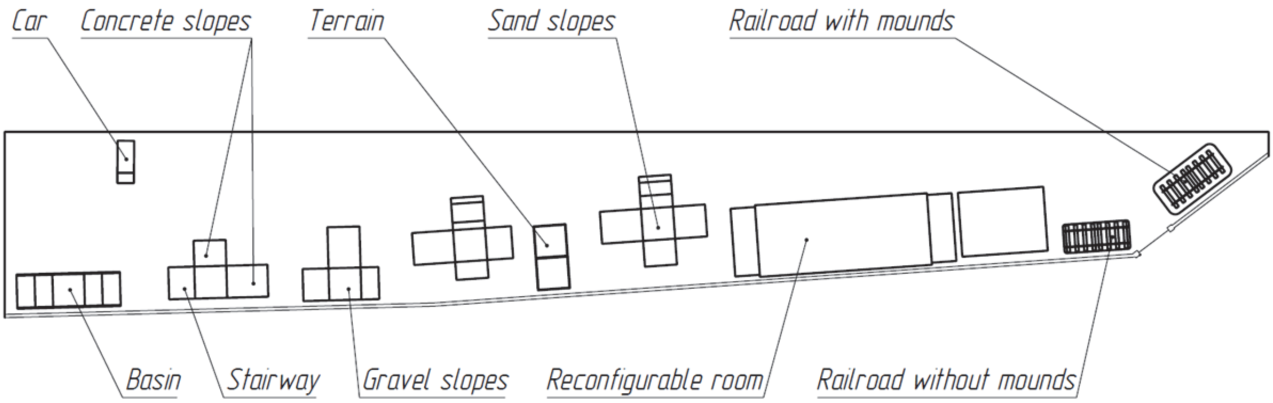


Figure 3 – Location of the obstacles within the testing ground

The traversability class for the basin module is calculated from the maximum relative depth of water in the basin at which the robot overcomes the pool. The relative depth is calculated based on the radius R of the robot wheels, as well as the height h of the robot body elements. Table 1 shows the values of the traversability class of the basin module for different depths.

Table 1. Selection of the X traversability class value for the basin module

Water depth	**Traversability assessment**	**Traversability class**
Up to 0.5R	Elementary	TC-1(Y)
More than 0.5R to R	Partial	TC-2(Y)
More than R to h	High	TC-3(Y)
More than h	Complete	TC-4(Y)
More than 600 mm	Complete	TC-4(Y)

The traversability class for the flight of stairs is calculated from the possibility of moving up the flight of stairs of the LM-27.12.14-4 type in accordance with GOST 9818-2015. If it is possible to move up, but if the robot cannot move from the lower horizontal position to the position parallel to the slope of the stairs, the attempt starts from this position. In this case, the robot drives by itself to the top of the slope and stairway module along an inclined surface with 20° and goes backward in reverse until it touches the lower stair. If the robot cannot move up the flight of stairs, the possibility of the downward movement is taken into account. In this case, only the attempt, in which the robot goes down from a horizontal position on the top platform to a horizontal position on the ground, is counted. When the robot is flipped, the attempt is considered failed. For further calculation of the speed of the obstacle overcoming, the speed class is reduced by 1 if the robot moves up not from a horizontal position and by 2 if the robot moves down. If, as a result, the speed class drops below 1, the traversability class for the flight of stairs is assigned to TC-00. This is true for all such cases of checking the speed of the obstacle overcoming. Table 2 shows the values of the traversability class of the stairway module.

Table 2. Selection of the X traversability class value for the stairway module

Movement	**Traversability assessment**	**Traversability class**
Downward	Elementary	TC-1(Y-2)
Upward, from the slope	Partial	TC-2(Y-1)
Upward, diagonally	High	TC-3(Y)
Upward, straight	Complete	TC-4(Y)

When calculating the traversability class for the slope modules, the middle traversability class for each type of surface is considered as an estimate of the traversability class. To define the average value, the traversability class is calculated for each sloping surface and the arithmetic average is calculated, rounded up to the integer. Tables 3 to 5 show the traversability class values for the slope modules. If the robot can not pass up a slope, on slopes with concrete and gravel surface, a required condition for obtaining the first traversability class is the possibility of stopping on an inclined surface of the slope. The ability to stop is checked separately from the attempt at which the overcoming speed is measured. When measuring the overcoming speed for TC-1 (Y), the value of Y is reduced by 1 for the slope module with sand covering, by 2 for the slope module with gravel covering, and by 3 for the slope module with concrete covering.

Table 3. Selection of the X traversability class value for the slope module with concrete covering

Movement	**Traversability assessment**	**Traversability class**
Downward	Elementary	TC-1(Y-3)
Upward, diagonally	Partial	TC-2(Y)
Upward, straight with stops	High	TC-3(Y)
Upward, straight without stops	Complete	TC-4(Y)

Table 4. Selection of the X traversability class value for the slope module with gravel covering

Movement	**Traversability assessment**	**Traversability class**
Downward	Elementary	TC-1(Y-2)
Upward, diagonally	Partial	TC-2(Y)
Upward, straight with stops	High	TC-3(Y)
Upward, straight without stops	Complete	TC-4(Y)

Table 5. Selection of the X traversability class value for the slope module with sand covering

Movement	**Traversability assessment**	**Traversability class**
Downward	Elementary	TC-1(Y-1)
Upward, diagonally	Partial	TC-2(Y)
Upward, straight with stops	High	TC-3(Y)
Upward, straight without stops	Complete	TC-4(Y)

To calculate the traversability class for the terrain module, the possibilities of a getting on the module, a passage across the module and a turning are taken into account. A getting on the module is defined as a robot possibility to take a horizontal position within the module, while the time is not measured and if successful, the class TC-11 is assigned. A turning is considered as the possibility of a crossing one section of a module and out into another section at an angle of more than 45° from the edge of the module. For the passage with a turn, the traversability class for the passage time increases by 1. Table 6 shows the values for defining the traversability class during the passage of the terrain module.

Table 6. Selection of the X traversability class value for the terrain module

Overcoming	**Traversability assessment**	**Traversability class**
Getting on the module	Elementary	TC-11
Passage across the module	High	TC-3(Y)
Passage with a turn	Complete	TC-4(Y+1)

When calculating the traversability of the module of the reconfigurable room, the possibilities of maneuvering in spaces with limited volume and in conditions of limited visibility are checked. With the overcoming of the module of the reconfigurable room, the operator takes a position diagonally from the corner of the module at a distance of ten meters and cannot change it during the attempt. The lack of progress in the passage of this test is the preservation of the position of the robot within one section. The possibilities of the doorways passing and the doors (that are not fixed by locking mechanisms) opening are estimated. During the passage through a doorway with the door opening against the direction of movement, the traversability class (in terms of the speed) increases by 2. Table 7 shows the traversability class values for the module of the reconfigurable room.

Table 7. Selection of the X traversability class value for the module of the reconfigurable room

Doors and openings	**Traversability assessment**	**Traversability class**
Without passing doorways	Elementary	TC-1(Y)
With passing doorways, without opening the doors	Common	TC-2(Y)
With the door opening in the direction of movement	High	TC-4(Y)
With the door opening against the direction of movement	Complete	TC-4(Y+2)

To calculate the traversability class of the railroad modules, the possibility of movement along the rails and across the rails with and without mounds is checked. If the robot cannot to overcome the rails across but can transfer the front wheel axle through the rail, this is considered partly traversable. In this case, the attempt time is not measured, and the robot receives the TC-22 traversability class. Table 8 shows the values for defining the traversability class during the passage of the railroad modules.

Table 8. Selection of the X traversability class value for the railroad module

Movement	**Traversability assessment**	**Traversability class**
Along the rails	Elementary	TC-1(Y)
Getting on the rail	Partial	TC-22
Across the rails without mounds	High	TC-3(Y)
Across the rails with mounds	Complete	TC-4(Y)

The absolute value of the maximum velocity of the robot in this method does not affect the traversability class for each obstacle; a subjective estimate of speed is used. When measuring the average maximum velocity, the traversability class is also assigned, depending on the length of the robot. The traversability class for speed is calculated depending on the ratio of the speed (in meter per second) to the length of the wheelbase (in meters). Table 9 shows the X values for the traversability class when measuring the maximum velocity of the robot.

Table 9. Selection of the X traversability class value when measuring the speed

The value of the ratio V/W	**Speed assessment**	**Traversability class**
Up to 0.3	Extremely low	TC-11
Over 0.3 to 1	Low	TC-22
Over 1 to 3	Common	TC-33
Over 3	High	TC-44

After defining the value of the obstacle traversability class, the positions of the origin O and the ending E points of the obstacle overcoming are determined to calculate the class of the overcoming speed. The position of the origin point of the obstacle is always fixed for each obstacle, the position of the end point is chosen based on the traversability class. Figures 4-6 show the positions of points O and E for various obstacles and traversability classes. For classes TC-1 (Y) during the overcoming of the slope and stairway modules the points E and O change places with each other.

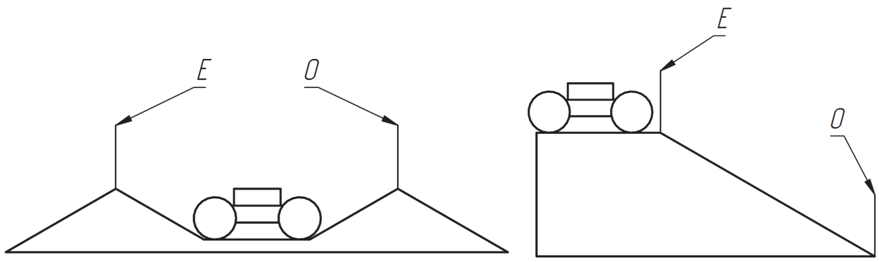


Figure 4 – Positions of the points O and E for the basin module (on the left) and for the slope module for classes other than TC-1(Y) (on the right)

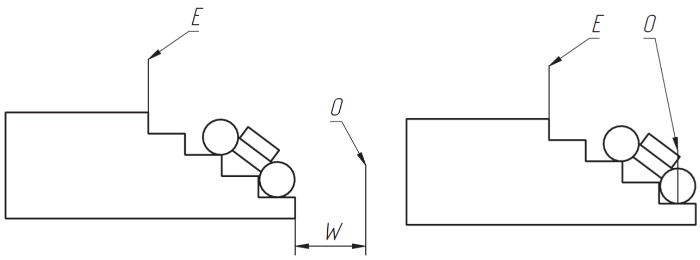


Figure 5 – Positions of the points O and E for the stairway module for classes other than TC-1(Y) (on the left) and for the TC-1(Y) class (on the right)

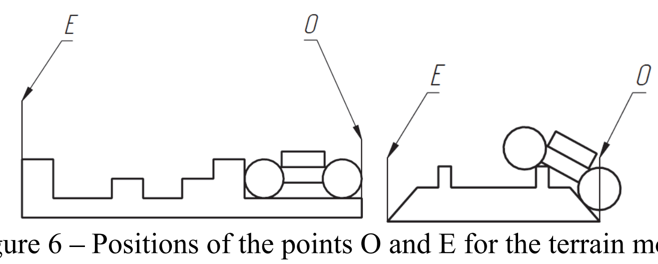


Figure 6 – Positions of the points O and E for the terrain module for the TC-3(Y) and TC-4(Y) classes (on the left) and for the railroad module for the TC-3(Y) and TC-4(Y) classes (on the right)

To calculate the speed class of the obstacle overcoming, the average speed U of the obstacle overcoming is measured. The ratio of the average speed of the obstacle overcoming to the average maximum velocity of the robot determines the second digit in the traversability class in accordance with Table 10. When measuring the speed of the obstacle overcoming, the starting and the ending points of the time count correspond to the moments of an intersection of the origin and the ending points of the obstacle respectively by the axis of the rear pair of wheels of the robot.

Table 10. Selection of the Y traversability class value

The value of the ratio U/V	**Traversability assessment**	**Speed class of the obstacle overcoming**
Up to 0.1	Extremely difficult	TC-(X)1
Over 0.1 to 0.5	Difficult	TC-(X)2
Over 0.5 to 0.8	Common	TC-(X)3
Over 0.8	Complete	TC-(X)4

Conclusion

A methodology for a qualitative assessment of the capabilities of ground-based mobile robotic systems in terms of the obstacle overcoming has been developed. This methodology implies assigning a quality assessment to the robot in the form of traversability class value based on the results of passing the testing ground. The methodology also implies the possibility of further development and expansion for use it at other testing grounds.

References

1. The project of the GOST R 60.6.3.2 «Robots and robotic devices. Test methods for robots in extreme conditions. Traversability. Cracks overcoming».
2. ASTM E2801-11 «Standard Test Method for Evaluating Emergency Response Robot Capabilities: Mobility: Confined Area Obstacles: Gaps».
3. Nogin M.A., Korotkov A.L., Rogov A.V., Shmakov O.A., Lopota A.V. RTC proving ground for mobile robotic complexes // Proceedings of the International Scientific and Technological Conference «Extreme Robotics and Conversion Tendencies», June 7-8, 2018, Saint Petersburg, pp. 352-360.

D.S. Popov, O.A. Shmakov

REQUIREMENTS FOR REMOTE CONTROL SYSTEMS FOR GROUND-BASED MOBILE ROBOTS

The Russian State Scientific Center for Robotics and Technical Cybernetics,
Saint Petersburg, Russia
d.popov@rtc.ru, shmakov@rtc.ru

Abstract

The article provides a brief overview of remote control systems (RCS) that are parts of ground-based mobile robots and an analysis of the main characteristics and requirements for the RCS. Due to the wide variety of work performed by the robotic systems, any mobile robotic system should be divided into the main functional parts: basic robotic platform, hinged equipment and a control station. Based on this division, the basic requirements for the RCS are grouped and analyzed, and examples of existing systems are given. Key areas for improving the effectiveness of the RCS are highlighted.

Keywords: remote control system, mobile robot, robotic platform, attachments, control panel.

Acknowledgments

This research is supported by the Ministry of Science and Higher Education of the Russian Federation. State assignment at the Russian State Scientific Center for Robotics and Technical Cybernetics № 075-00924-19-00.

Introduction

At present, the service robotics market is rapidly expanding. The most common category is ground mobile robots. The vast majority of actually used mobile robots does not yet have full autonomy and require operator control. Only a few models have the ability to perform certain tasks without constant operator intervention. Therefore, here we consider the robot control systems as precisely the systems with remote control by a human operator. Initially, it is useful to evaluate the existing approaches to the development of remote control systems. It is also necessary to analyze the current level of requirements for the products that are similar in terms of size, mass, and functional characteristics.

Ground-based mobile robotic systems

Due to the wide variety of work performed and the modular construction principle of robots, it is reasonable to divide any mobile robotic system (MRS) into the main functional parts:

– basic robotic platform, on the basis of which the system is developed;
– hinged equipment, directly located at the RP and performing a functionally-executive function;
– control station, which includes a remote control panel for control of the system (see figure 1).

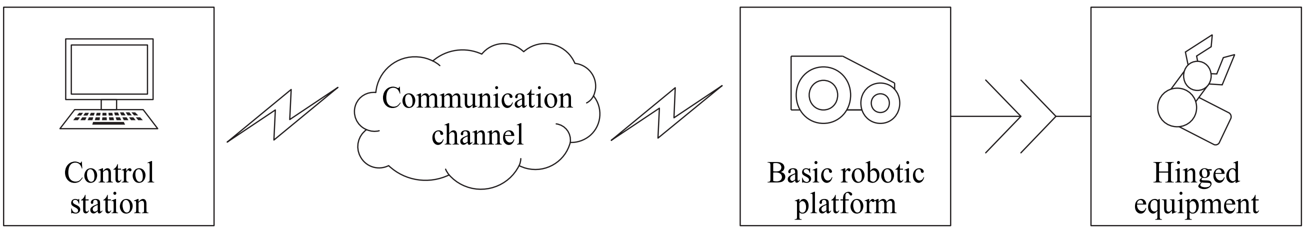


Figure 1 – **S**tructural diagram of the MRS

For the functional division of the MRS, it is reasonable to rely on mass and size characteristics of the robotic platform and, consequently, the type of hinged equipment that is included in its composition. In [1], an extended, classification was proposed and justified (see table 1) comparing with GOST R 22.9.22-2014 (“Safety in emergency situations. Emergency rescue devices. Classification”).

The robotic platform can have the form of a supporting structure and mounted on it drive units, on-board electronics, and peripheral systems. Hinged equipment is connected to the platform through a connecting device. Reconfiguring of the basic platform allows to adapt it maximally to the operating conditions.

To expand the functionality of the MRS, it should be equipped with hinged equipment. To reduce MRS setup and reconfiguration time, equipment should be quickly replaceable. For this purpose, fast-locking connections are often used [2].

It is efficient to distinguish two levels in the architecture of remote control systems:

– a high level, which determines the general logic of work, performs such tasks as controlling manipulators by solving direct and inverse problems of kinematics and dynamics, navigation, technical vision, etc. (in the conventional hierarchy of control systems such level corresponds to a combination of intelligent, strategic and tactical levels);

– a low level that provides motors control, sensors data merging and primary data processing, communication between MRS and control panel, an interface of various payloads, etc. in hard real-time (executive level).

Table 1 – MRS classification

Class, subclass	**Total mass, kg**	**Possible equipment**
Ultralight first (UL1)	up to 1 kg	Passive video/audio surveillance systems
Ultralight second (UL2)	from 1 to 5 kg	Driven video/audio surveillance systems, micromanipulators
Ultralight third (UL3)	from 5 to 15 kg	Driven video/audio surveillance systems, manipulators
Ultralight fourth (UL4)	from 15 to 50 kg	
Ultralight fifth (UL5)	from 50 to 100 kg	
Light first (L1)	from 100 to 300 kg	Driven video/audio surveillance systems, manipulators, technological equipment
Light second (L2)	from 300 to 1000 kg	
Middle first (M1)	from 1000 to 5000 kg	
Middle second (M2)	from 5000 to 20000 kg	
Heavy (H)	over 20000 kg	

Architecture of the remote control system

The most important principles laid down in control systems are reliability, flexibility, and simplicity.

From the position of architecture, reliability is ensured by dividing the remote control system into separate subsystems according to their functional purpose and parameters. For example, in the remote control system of the MRS of the third and fourth ultralight classes are often distinguished:

– a separate hard real-time subsystem responsible for low-level control of the robotic platform, i.e. motors, sensors, power supply system;

– a hard real-time subsystem responsible for the low-level control of removable hinged equipment that is critical to the response time, for example, a manipulator;

– a soft real-time subsystem that controls such elements of the hinged equipment as cameras, backlight modules, etc.

A flexibility entails the need to implement the system based on individual modules common for MRS of various classes.

The architecture is also significantly influenced by the class of the robot, even with a similar set of tasks performed and similarelements of the remote control system. So, for UL1, the entire system is usually performed with the maximum level of integration on one or two printed circuit boards, for UL4 and higher, as separate units in their own shielded enclosures.

The composition of the remote control system

Following the MRS, the remote control system can also be structurally divided into main parts:

– an onboard control system, including control systems of the robotic platform and hinged equipment;

– remote control station;

– control and data transmission channel.

Onboard control system

An onboard control system generally can be considered as a complex consisting of the following structural elements:

– a control system of the robotic platform;
– a control system of the hinged equipment;
– a system of audio/video devices;
– a system of receiving and transmitting information;
– a power supply system;
– a self-diagnostic system.

An example of the block diagram of the onboard control system is presented in figure 2.

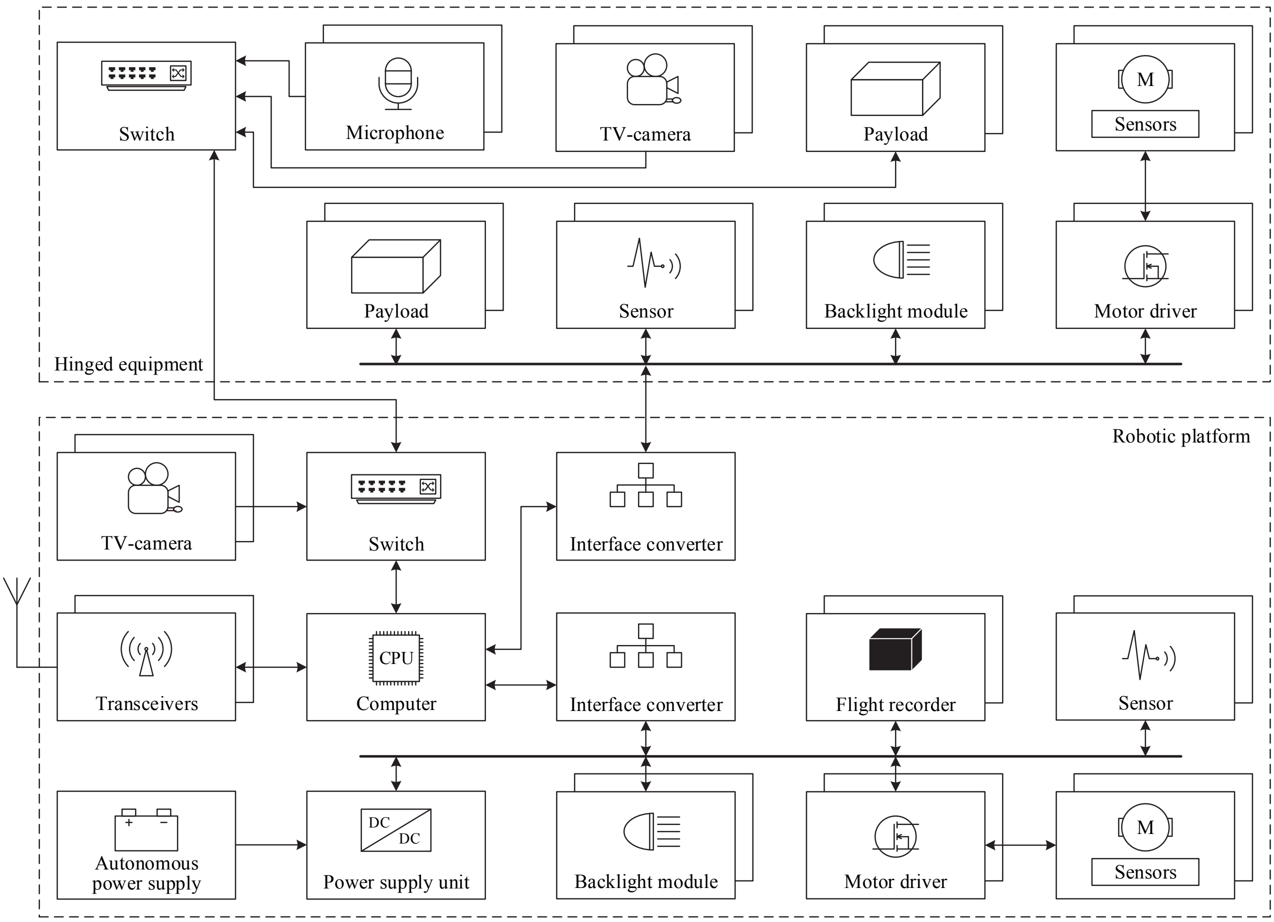


Figure 2 – An example of the block diagram of the onboard control system

Control system of the robotic platform

The central element of the control system of the robotic platform is a computer, designed for the general control of the mobile part of the MRS and ensuring interaction between the main subsystems. For the MRS of the first and second ultralight classes, this computer is based on a microcontroller ensuring the performance of all general control tasks. For larger MRS, such computer can be made on the basis of an industrial computer. In the case of the most sophisticated systems, the computer consists of several blocks and may include:

– an industrial computer solving high-level control problems;
– a specialized computer that processes video from cameras – a video stream compressing for transmission to a remote control panel, technical vision problems, etc.;
– a microcontroller that provides low-level control of all motors and peripherals in hard real-time.

Such decomposition of a complex computing unit allows not only to reduce labor costs for developing and debugging software but also to increase the reliability of the system significantly. The microcontroller is ready to work after a reboot, caused, for example, by voltage surges, other external influences or "freezes" due to software errors and inaccuracies. Unlike a full computer, which can take up to several minutes to load an operating system the microcontroller is ready in milliseconds. It allows to keep controllability of the MRS in emergency situations.

All units of the onboard control system are interconnected and linked to the computer via a local network, which is closed inside the robot and has not direct output beyond the system. The local network can also be implemented in different ways, depending on the class of the MRS. In the simplest case, for example, for UL1, the local network may be completely absent or be represented by several independent serial interfaces on a single control board. In more complex systems, the local network is built from several independent segments:

– a network of low-level control of motors and peripheral systems of the robotic platform based on the "bus" topology;

– a network of low-level control of hinged equipment, also based on the "bus" topology;

– a system-wide onboard network of the MRS based on the "star" topology.

For the first two segments, interfaces such as RS-485, CAN, MIL-STD-1553B, EtherCAT and similar are used. EtherNet is usually used for the latter segment. Segment separation is performed in order to increase the overall reliability of the system.

Control system of the hinged equipment

Hinged equipment is connected to the robotic platform through external connectors and controlled by a central computer using local networks. In some cases, system equipment may have its local controller. Then it also connects to the network and only receives high-level commands from the central computer.

System of audio/video devices

The system of audio/video devices is designed to merge, process and transmit audio and video information from the mobile robot to the remote control panel. Since in most cases the bandwidth of the video transmission channel from the robot to the control panel is limited, the functions of the quad, picture-in-picture, etc. are widely used to provide the operator with the maximum amount of information.

System of receiving and transmitting information

The system of receiving and transmitting information is intended for:

– receiving control commands from the remote control panel;

– transmission telemetry to the remote control panel;

– audio and video transmission to the remote control panel.

The system consists of the separate transceiver modules, the number, and functionality of which is determined by the type of communication and control channel used in a particular MRS. To ensure internal electromagnetic compatibility and reduce the interference of the mobile robot components, the elements of the system are almost always executed as separate shielded modules, even in the case of small robotic platforms of the UL1 class, in which much attention is paid to the integration of components.

Power supply system

The power supply of the mobile robot can be carried out:

– using rechargeable batteries;

– using an internal combustion engine and a generator;

– remotely by cable;

– via a combination of the above items.

The remote control system provides power control, remaining battery charge calculation, and diagnostics.

Self-diagnostic system

The self-diagnostic system is designed to monitor the state of the remote control system during debugging and operation. Usually, it is represented by software modules as part of the computer, redundant lines of communication and control, as well as flight recorders engaged in logging traffic in the local networks of the MRS.

Control and data transmission channel

The control and data transmission channel can be implemented both on the basis of a radio channel, and on the basis of wired communication.

Currently, in the development of MRS two approaches to the organization of a radio channel for information sharing are used.

In the first case, a digital low-speed transceiver is responsible for receiving commands and transmitting telemetry data, while audio and video information is transmitted to the remote control panel via a second broadband transmitter. In some cases, it is possible to transmit telemetry data with a video signal; then the digital transceiver operates in simplex mode, which provides better electromagnetic compatibility and, as a result, improves the quality of communication.

The second approach is based on the use of one broadband transceiver. This can be either a specialized radio channel or industrial Wi-Fi systems (for example, the MikroTik Metal series). To improve reliability, the broadband digital transceiver can be supplemented with a narrow-band low-speed radio modem (as previously) used for control in case of temporary loss of communication.

It is also worth mentioning the possibility of implementation on the basis of some radio channels of signal relay systems. In this case, one mobile robot can perform the functions of a repeater, transmitting information from the remote control panel to another robot, when the remote control panel does not have direct radio communication.

Wired communication is used in cases when the use of a radio channel is difficult or undesirable:

- a use of MRS in radio silence mode;
- a need to protect from strong external electromagnetic interference, natural or artificial;
- increased requirements for the reliability of the control channel;
- high bandwidth requirements.

Sometimes a wired communication is used in conjunction with a remote power supply system, which allows to extend the time of MRS operation significantly.

Remote control station

The control station generally includes, depending on the composition of the MRS, a remote control panel, coils of cable communication, an outlying antenna post with transceivers and an antenna tripod, etc.

The remote control panel is equipped with visual, audio, telemetry and control information output systems, as well as with control nodes sufficient for interaction with all executive nodes of the robot.

An example of the block diagram of the remote control panel is shown in figure 3.

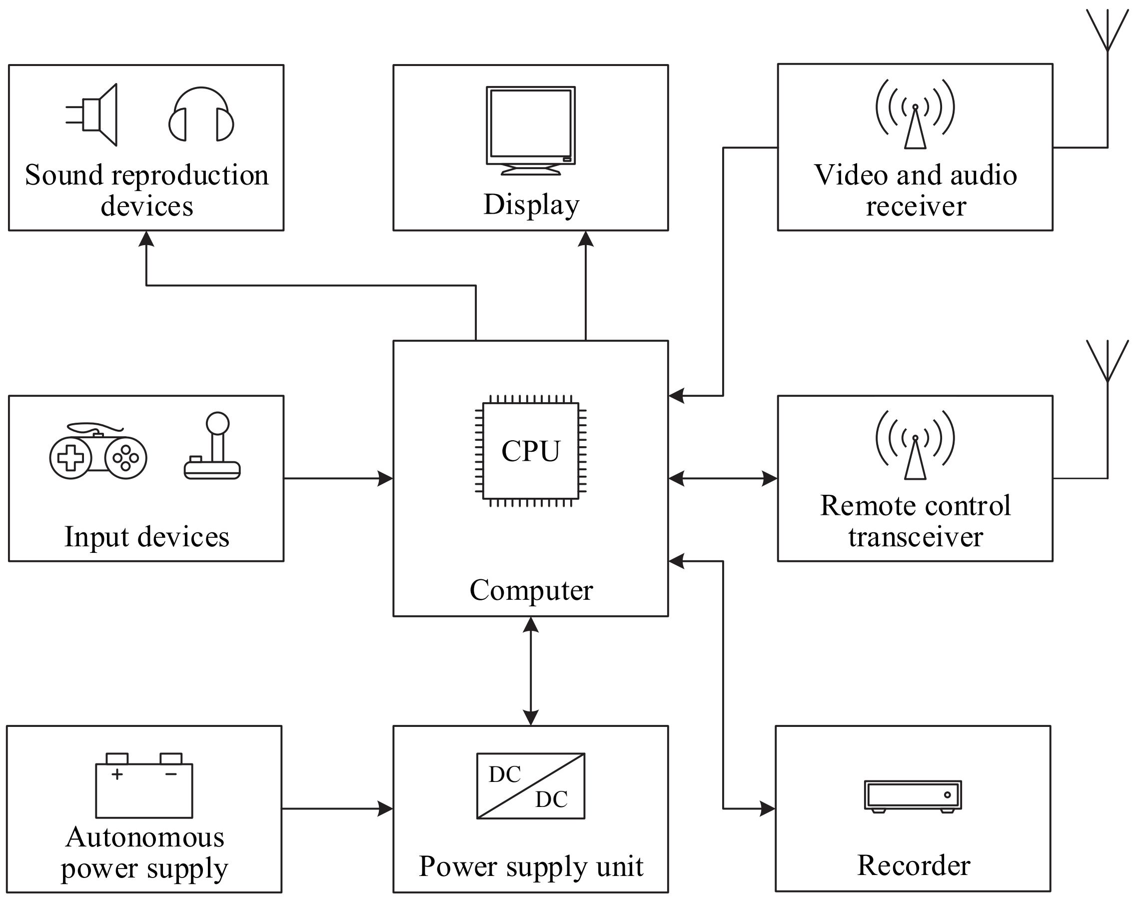


Figure 3 – An example of the block diagram of the remote control panel

One of the main requirements for the remote control panel is to provide the human operator with the maximum possible visual information about the environment and indicative information for predicting actions in solving operational problems. This requirement is fulfilled by equipping the remote control panel with a human-machine multimedia interface, the peculiarity of which is to provide information in the forms most effective for human perception, including methods of structuring the display of information so that it attracts the operator's attention to the most important information elements.

Based on the requirements for operating modes, modern MRS can be equipped with three types of remote control panels:

– a command remote control panel as part of a mobile command control station, designed to display the maximum possible visual information about the environment (video data from a controlled mobile robot, a terrain map with markers of the location of a group of robots, telemetric information about the state of the MRS, video information about the surrounding setting);

– a stationary remote control panel designed to control the mobile robot on the ground at a distance from the command point or in its absence;

– a portable remote control panel designed to control the mobile robot from various types of shelters, which can not be reached with a command or stationary remote control panel.

Examples of remote control panels are shown in figure 4.

a)

b)

c)

Figure 4 – Examples of remote control panels:
a) a command remote control panel developed at the RTC,
b) a stationary remote control panel MRS-27M,
c) a portable remote control panel MRS «Scarabaeus»

Modern requirements to the remote control systems

The increasingly stringent requirements for modern MRS in general, are largely related to remote control systems. Based on the above architectural principles, it is possible to group and analyze the basic requirements for the remote control systems.

Common key requirements for modern remote control systems are:

− modular principle of system design;

− high reliability;

− resistance to external factors.

The modular principle of system design provides the ability to quickly reconfigure the system, greatly simplifies debugging and tuning, and allows the use of previously developed blocks and nodes when building new systems without a significant re-engineering.

Many modern MRS are composed from a wide range of attachments, sets of accessories, individual nodes, and modules. For example, the MRS family of the Telerob company (Germany) tEODor and telemax are equipped with replaceable tools, sensors, additional batteries, coils of cable communication, radio repeaters, additional sets of wheels and tracks, turning cameras, etc. [3].

The overall *reliability* of the remote control systems is determined both by the system architecture and the reliability of individual mechanical parts, electronic components, and software. The main parameters are the operating time of the failure and the service life of the product. Maintainability, also related to reliability parameters, is regulated by state standards (for example, GOST RV 20.39.309 «Military hardware, instruments, devices, and equipment. Structural and technical requirements»).

In most cases, the permissible ambient temperature is the main requirement for the remote control system in terms of *external factors*. Since the mobile robot is often controlled from shelters, the requirements for the control system are usually lower than those for the robot.

Common possibilities of modern MRS are (see table 2):

− from minus 20 to plus 60 °C for MP;

− from 0 to plus 40 °C for the remote control panel.

In rare cases, manufacturers claim lower operating temperatures. Examples include domestic MRS produced by special design and technology bureau of applied robotics (LLC «SDTB AR»), BMSTU and RTC [4-6].

Requirements for the onboard remote control system

Requirements for the robotic platform are imposed on the operation time of the system in terms of energy supply and the level of autonomy of the mobile robot.

The operating time requirements are determined, first of all, by the purpose of the system and can be from few minutes for small robots or «disposable» systems designed to perform a narrow range of operations, and up to ten hours for light and heavy multifunctional MRS. The operation time depends on the battery capacity, the use mode of the MRS, the possibility of recharging or full power supply of the mobile robot from the control station by cable.

The average operation time is (see table 2):

− from one to four hours for mobile robots of the UL1-UL3 class;

− from four to eight hours for mobile robots of the UL4-L2 class.

Recently, more and more attention has been paid to increasing the autonomy of MRS, which is necessary for:

− reduce the requirements for quality and reliability of the communication channel;

− increase the reliability and survivability of the system;

− implementation of new scenarios for the use of a mobile robot (following the object, along a predetermined route, etc.).

For this, the remote control system includes the corresponding navigation subsystems, including computers with specialized software, rangefinders, GPS / GLONASS receivers, cameras. Examples are MRS aunav.NEXT and aunav.MEGA of the Spanish company Proytecsa [7, 8] and Uran-9 produced by JSC «766 UPTK» [6].

In a simpler case, it is sufficient to use the so-called «auto-return module», which provides an automatic movement of the mobile robot along the traversed route in the opposite direction in case of loss of communication, for example, when entering the skip zone. Such functionality can be implemented even on small robots, such as, for example, Nerva LG manufactured by Nexter (France) [9].

Requirements for the remote control system in terms of *payloads* are determined mainly by their connection to a mobile robot. Many modern robots have quick-tightening joints or Picatinny rail on the cover, which allows fixing a wide range of interchangeable payloads. A connection is made via unified interfaces, most often RS-485, CAN and EtherNet.

Requirements for the control station

The basic requirements for the control station are determined by ergonomics, operation time, delay in displaying video information, and the reaction time of the mobile robots to commands.

Ergonomics is generally regulated and evaluated by state standards and relevant guidelines.

The operating time of the control station from an autonomous power supply must not be less than the operating time of a mobile robot.

The delay in displaying video information on the monitor of the control panel and *the response time* of the mobile robot to the commands are important characteristics of the remote control system, determining not only the comfort of the operator in the control process but also the capabilities of the MRS.

Taking into consideration that the average response time of a person is 200 ms, a significant excess of this value due to delays and reaction time can significantly degrade the effectiveness of the remote control system, if no additional measures are taken (for example, described in [10]).

Requirements for the control and data transfer channel

The requirements for the control and data transmission channel are following:

– control range;

– the ability of several mobile robots to operate in the same work area;

– safety and information security.

The control range of the mobile robot via the radio channel is determined by the type of transceivers and antenna-feeder devices. Most often, relatively small antennas of a circular directivity are installed on the mobile robots and portable remote control panel. The structure of the control stations using the command and stationary remote control panels usually includes remote tripods with directional antennas.

It should be noted that the range in direct visibility conditions is usually two to three times longer than the control range in urban areas and similar obstacles.

The maximum range of control is (see table 2):

– for mobile robots of the UL1-UL2 class with a portable remote control panel – up to 200 m;

– for mobile robots of the UL3 class – from 300 to 400 m;

– for mobile robots of the UL4-UL5 class – up to 1000 m;

– for heavier mobile robots – from 1000 m and more.

A large part of the MRS of the UL4 class and higher is equipped, in addition to the radio channel, with a cable line. The maximum distance is determined by the type and diameter of the cable (twisted pair, coaxial, optical fiber), the number of coils (usually one for the robot and one for control station), permissible overall dimensions, the presence of a cable layer and ranges from 100 to 300 meters.

The ability of several mobile robots to operate in the same work area is important when using a radio channel and is determined by the capabilities of transceivers. Modern communication systems allow using from 6 to 10 same type robots in the same work area without creating radio interference with each other.

The protection of the radio channel is necessary to prevent unauthorized connections in order to obtain information or intercept control of the mobile robot. The optimal solution is to use equipment that includes software and hardware cryptographic modules to provide radio communications. An example is the «Cordon» transceiver series used in the MRS «Scarabaeus» and «Captain» [11].

Table 2 presents the characteristics of some mobile robotic systems of Russian and foreign production.

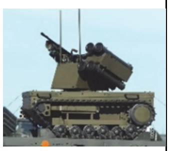
Table 2. Characteristics of the mobile robotic systems

MRS	**Image**	**Developer**	**Year**	**Weight**	**Operating time, h**	**Communi-cation range**	**Control channel**	**Operating temperature**	**Payloads and features**
Mobile robotic systems of the ultralight class									
Iris [12]		Robo-team (Israel)	2012	1,85 kg	1-3 h	200 m	n/a	-20…+60 °C	TV cameras, microphone, IR lights. Removable attachments on Picatinny rail, Ethernet interfaces, RS232.
110 FirstLook [13]		iRobot (USA)	2010	2,4 kg	6 h	200 m	2,4/4,9 GHz meshnet	-20…+55 °C	TV cameras, microphone, IR lights. Ability to work in repeater mode for other MRS.
StoneMarten [14]		Novatiq (Switzerland)	2013	4,5 kg	3 h	200 m	n/a	-20…+40 °C	TV cameras, microphone, IR lights.
Nerva LG [9]		Nexter (France)	2012	4,5 kg	2 h	300-1000 m	2,4 GHz	n/a	Auto return, navigation module.
Scarabaeus [15]		CJSC «SET-1» (Russia)	2014	4,7 kg	1 h	300 m	1,3-1,4 GHz 917 MHz	-20…+40 °C	TV cameras, microphone, IR lights.
Avatar [16]		RoboteX (USA)	2013	11 kg	4 h	300 m	2,4 GHz	n/a	TV cameras, microphone, IR lights. Removable attachments, manipulators, sensors.
Scorpio [17]		CJSC «SET-1» (Russia)	2014	17 kg	4 h	1000 m	1110-1230 MHz	-20…+50 °C	Pan-tilt TV cameras, microphone, IR lights.
MRS Captain [18]		RTC (Russia)	2018	35 kg	4 h	300 m (cable) 500-1000 m (radio)	1,3-1,4 GHz 917 MHz	-30…+40 °C	Removable attachments. Cameras, microphone, lights.

Table 2. Characteristics of the mobile robotic systems

MRS	**Image**	**Developer**	**Year**	**Weight**	**Operating time, h**	**Communi-cation range**	**Control channel**	**Operating temperature**	**Payloads and features**
MRS-15M [4]		LLC «SDTB AR» (Russia)	2017	80 kg	4 h	135 m (cable) 1000 m (radio)	n/a	-30…+40 °C	Removable attachments. Cameras, microphone, lights.
Mobile robotic systems of the light class									
telemax PLUS [19]		Telerob (Germany)	2018	113 kg	2,5-5 h (NiMh) 3-12 h (Li-ion)	n/a	n/a	-20…+60 °C	Large assortment of interchangeable attachments. Cameras, microphone, lights.
RMI-9XD [20]		Pedsco Ltd. (Canada)	2014	176 kg	5-6 h	200 m (cable) 1000 m (radio)	2,4 GHz 700 MHz	-20…+50 °C	Removable attachments. Cameras, microphone, lights.
MPK-27M [4]		LLC «SDTB AR» (Russia)	2013	210 kg	4 h	200 m (cable) 1000 m (radio)	n/a	-30…+40 °C	Removable attachments. Cameras, microphone, lights.
MPK-28 [4]		LLC «SDTB AR» (Russia)	2012	270 kg	5 h	200 m (cable) 60 m (cable layer) 1000 m (radio)	n/a	-30…+40 °C	Removable attachments. Cameras, microphone, lights.
Sentinel [21]		AB Precision (Poole) Ltd. (UK)	2014	270 kg	n/a	150 m (cable) 1000 m (radio)	1,2-1,4 GHz 400-450 MHz	n/a	Removable attachments. Cameras, microphone, lights. GPS, navigation module.
tEODor EVO [19]		Telerob (Germany)	2018	383 kg	3-4 h	n/a	n/a meshnet	-20…+60 °C	Large assortment of interchangeable attachments. Cameras, microphone, lights.

Table 2. Characteristics of the mobile robotic systems

MRS	**Image**	**Developer**	**Year**	**Weight**	**Operating time, h**	**Communi-cation range**	**Control channel**	**Operating temperature**	**Payloads and features**
aunav.NEXT [7]		Proytecsa (Spain)	2013	495 kg	5 h	700-3000 m	n/a	-10…+50 °C	Navigation system (following the object, on a certain route, bypass obstacles, auto return).
Platform-M [22]		CJSC «SRTI «Progress» (Russia)	2014	800 kg	n/a	3000 m	n/a	n/a	Cameras, weapons.
Mobile robotic systems of the middle class									
Nerekhta [23]		Degtyaryov Plant (Russia)	2014	1000 kg	n/a	1500 m	n/a	n/a	Cameras, weapons.
MRS High-scaler [5]		BMSTU (Russia)	2013	1400 kg	8 h	300 m (cable) 500-1000 m (radio)	n/a	-40…+40 °C	Removable attachments. Cameras, microphone, lights.
aunav.MEGA [8]		Proytecsa (Spain)	2013	8090 kg	8 h	700-3000 m	n/a	-10…+50 °C	Navigation system (following the object, on a certain route, bypass obstacles, auto return).
Uran-9 [6]		JSC «766 UPTK» (Russia)	2016	12000 kg	6 h (diesel)	3000 m	meshnet	n/a	Cameras, weapons. Navigation system (following the object, on a certain route, bypass obstacles, auto return).

Conclusion

A brief overview allows assessing the current capabilities of the remote control systems and the trends in customer and consumer requirements. Over the past few years, the capabilities of the «low level» of the remote control systems have slightly changed in terms of battery life, cable or radio control range, and operating temperature. Several major changes affected the «high level».

Taking into account the above, it is possible to identify the main areas of further work in improving the efficiency of remote control systems for mobile robotic systems.

1. The use of modern element base in the development and modernization of systems. The capabilities of the radio-electronic industry are growing rapidly, together with them the possibilities for integrating and reducing the weight and size characteristics of the remote control systems are growing.

2. Universalization of both individual elements of the system, and the remote control system as a whole. Development of common universal interfaces and information exchange protocols, which allow:

– using universal control panels to interact with different types and classes of new MRS;

– using interchangeable modules in the construction of robotic systems;

– developing universal attachments that connect to a wide range of robotic platforms.

3. Intellectualization of the remote control systems. An adding element of technical artificial intelligence into the remote control system allows:

– increasing the ease of use and ergonomics of the control unit;

– improving the reliability of data transmission over a highly noisy radio channel;

– increasing the autonomy of MRS, significantly reducing the requirements for the control channel.

4. Development of approaches and methods that provide the possibility of comparing, evaluating and structural-parametric synthesis of remote control systems. For this, it is necessary to formalize, analytically express the requirements for the system and determine the basic, numerically expressed parameters for assessing its quality. Such parameters can be based on numerical methods for assessing reliability, manufacturability, level of unification, etc.

References

1. Pryamitsin I.B., Rogov A.V., Shmakov O.A., at al. Ultralight ground-based mobile robotic platforms // Materials of the 10th Russian Multiconference. 2017. Pp. 48-50.
2. Pryamitsin I.B., Nogin M.A., Rachitskiy A.V., at al. Modular hinged equipment // Materials of the 10th Russian Multiconference. 2017. Pp. 45-47.
3. Telerob official site. URL: https://www.telerob.com/en/.
4. LLC «SDTB AR» official site. URL: http://sktbpr.ru/content/produkciya-sktb-pr/.
5. BMSTU official site. URL: http://robotmobot.ru.
6. JSC «766 UPTK» official site. URL: http://766uptk.ru/index.php?do=static&page=bmrk.
7. Aunav-NEXT. URL: https://www.proytecsa.net/en/aunav-NEXT.php.
8. Aunav-MEGA. URL: https://www.proytecsa.net/en/aunav-MEGA.php.
9. Robotiq de defense et securite. URL: https://www.nexter-group.fr/images/stories/fichiers_pdf/Robotics/Robotics_FR.pdf/.
10. Zaborovskiy V.S., Guk M.Yu., Mulyukha V.A., Silinenko A.V., Volkov O.N., Belyaev M.Yu. Teleoperation of on-ground robots from onboard the ISS in frame of «Kontur-2» space experiment // Proceedings of international scientific and practice conference «Scientific exploration and experiments on the ISS». 2015, Moscow. Pp. 64-65.
11. KORDON radio receivers. URL: https://www.set-1.ru/products/videosistemy/cordon-videoperedatchiki/.
12. Iris. URL: http://www.robo-team.com/products/iris/.
13. iRobot 110 FirstLook. URL: http://www.darley.com/documents/guides/robotics/spec_sheets/FirstLook_Specs.pdf/.
14. Ground-based robots: from drop systems to unmanned traffic convoys. URL: http://oko-planet.su/science/scienceday/268283-nazemnye-roboty-ot-zabrasyvaemyh-sistem-do-bezlyudnyh-transportnyh-kolonn.html.
15. Scarabaeus. URL: https://www.set-1.ru/products/robototekhnicheskie-kompleksy-/skarabey/.
16. RoboteX official site. URL: http://robotex.com.
17. Scorpio. URL: https://www.set-1.ru/products/robototekhnicheskie-kompleksy-/scorpion-robot-razvedka-razminirovanie/.
18. MRS Captain. URL: http://rtc.ru/ru/nazemnaya-robototekhnika/robot-kapitan/.
19. Telerob official site. URL: https://www.telerob.com/en/.
20. RMI-9XD. URL: http://www.pedsco.com/products-rmi9xd.php.
21. AB Precision (Poole) Ltd. Official site. URL: http://www.abprecision.co.uk/explosive-ordnance-disposal/remotely-operated-vehicles/.
22. Platform-M. URL: http://www.arms-expo.ru/news/perspektivnye-razrabotki/platforma-m-robotizirovannyy-kompleks-shirokikh-vozmozhnostey/.
23. Nerekhta. URL: https://topwar.ru/84742-proekt-robototehnicheskogo-kompleksa-nerehta.html.

A.A. Vlasenko, A.L. Korotkov, O.A. Shmakov

MODULAR MANIPULATION DEVICE WITH REMOVABLE OPERATIONAL EQUIPMENT

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
a.vlasenko@rtc.ru, a.korotkov@rtc.ru, shmakov@rtc.ru

Abstract

The aim of the work is to develop a manipulator with the ability to use interchangeable tool modules on a small-sized robotic platform.

An analytical review and comparison of existing manipulators and gripping devices were carried out and their advantages and disadvantages were revealed in the article. Based on the requirements for the performing task and the review, the requirements for the manipulator and the gripping device were formed. A preliminary calculation was carried out, on its basis was designed the manipulator and gripping device.

The simulation of technological operations was implemented to ensure efficiency. Relevant results were obtained for manipulator control. The design of the manipulator and the gripper was developed and described.

Keywords: manipulator, robotic platform, gripping device, drilling, technological operation.

Acknowledgments

This research is supported by the Ministry of Science and Higher Education of the Russian Federation. State assignment at the Russian State Scientific Center for Robotics and Technical Cybernetics № 075-00924-19-00.

Introduction

Mobile robotic systems are used for various tasks, including the search and inspection of the objects that can pose a threat to human life and health.

Existing manipulators for mobile robotic systems weigh from 5 to 15 kg and have limited functionality, consisting usually of the movement of objects due to a two-finger gripper, which can optionally be equipped with a video surveillance camera. The using of interchangeable tool modules can expand the functionality of the manipulation device.

The using of tool modules at the manipulator is promising, as it expands its functionality. The tool may be placed on the gripper, but this excludes the possibility of its rapid change or the possibility of the object capturing by the manipulator. It is also possible to use a sophisticated gripping device that combines several tools and gripping fingers to provide greater functionality, but the mass of such module is unreasonably large, which increases the requirements for the manipulation device itself.

Recent developments of manipulators for mobile robotic platforms

Manipulator «Servosila Robotic Arm»

Manipulator Servosila Robotic Arm (figure 1) is a removable module for robotic platforms. It has water and dust protection, so it can be used in rainy and snowy weather. Servosila Robotic Arm is able to withstand shocks and collisions with obstacles.

The manipulator has 5 degrees of freedom and is able to lift an object weighing up to 6 kg on the border of the working area, with its own mass of 8 kg. Motors with a wave gear and integrated bearing unit are placed inside the joints.

The modular design, consisting of motors and links connecting them, allows varying the configuration of the manipulator. Individual modules can be replaced to change the parameters of the manipulator, such as expanding or reducing the radius of the working area.

Figure 1 – Manipulator «Servosila Robotic Arm»

An optional vision module, which includes various cameras, can be installed at the gripper of the manipulator as interchangeable equipment.

Manipulator «DM4-A2»

DM4-A2 (figure 2) is a modular manipulator for robotic platforms. It consists of links providing 4 degrees of freedom and is equipped with a two-finger gripper. The manipulator is manufactured according to the standards of Resquared Robotics. The control is implemented with a joystick.

The manipulator has 4 degrees of freedom and is able to lift an object weighing up to 4.5 kg near the robot, with a manipulator mass of 2 kg. Motors with planetary gears transmitting the rotational moment to the links through the bearing units are placed inside the joints.

Optionally, an additional camera can be installed on the link in front of the gripper to monitor the capture of an object visually.

Figure 2 – Manipulator «DM4-A2» produced by Resquared Robotics, USA

Manipulator of the robot «SUGV»

The manipulator is designed to be used on the SUGV mobile platform (figure 3) for capturing and moving objects, opening doors.

The manipulator has 4 degrees of freedom and is able to lift an object weighing 5.4 kg at the border of the working area, with its own mass of 3.7 kg.

The manipulator is equipped only with rotary joints driven by electric motors through a worm gear. The rotation of the gripper around the longitudinal axis is implemented through a cylindrical gear.

Figure 1 – Manipulator of the robot «SUGV» produced by Endeavor Robotics, USA

Requirements for the manipulator

The considered manipulators are equipped only with rotational kinematic pairs, as a result, all of them have a spherical working area bounded below by the structural geometry of the mobile platform and the location of the manipulator. In the joints, mainly planetary and worm gearboxes are used, since they provide the required positioning accuracy and the necessary gear ratio, with lower economic cost and manufacturing complexity, comparing to wave gearboxes. Optionally, additional cameras and other attachments can be mounted on the manipulators. The considered manipulators consist of a root link, two or three intermediate links and a gripper. All links are connected by rotational kinematic pairs, however, some of the manipulators do not have a degree of freedom in the root link.

The disadvantages of the considered manipulators include the following:

– external wiring;
– part-turn joints;
– a limited set of additional replacement equipment;
– the using of low-efficiency mechanical gears (worm gear).

Based on the review and analysis of the considered manipulators, the following requirements are imposed on the manipulator:

– the using of rotational kinematic pairs;
– the using of mechanical transmission with high efficiency;
– the possibility of quick change of attachments;
– the load capacity of 3 kg at the border of the working area;
– length – up to 900 mm;
– mass of the manipulator – up to 3 kg.

The ability to change equipment without returning the robot to the place of service significantly reduces runtime of the mission, expands the functionality of the mobile robotic platform and saves the battery charge of the robot that is needed to move to the place of service and back. The analysis of manipulators for mobile platforms weighing from 5 to 15 kg allows concluding that the development of a manipulator with the ability to quickly change attachments is a relevant and promising task.

Recent developments of gripping devices for mobile and stationary robots

Gripping device «Robotiq2F-140», Robotiq

The gripping device Robotiq2F-140 (figure 4) is intended for use on industrial robots and mobile platforms. It is a simple and reliable scheme with two fingers that move in parallel. The gripper is mainly intended for use on industrial manipulators, however, installation on mobile platforms is also possible. The standard configuration of the gripping fingers allows capturing and moving objects of various shapes, the dimensions of which correspond to the requirements. However, using the tool with this gripping device is impossible.

A specific feature of this gripper is the presence of multiple sensors that help to control the capture of an object fully. It is also optional to install a camera in the form of a small add-on to control the capture of an object using vision systems.

Figure 4 – Gripper «Robotiq2F-140»

Gripping device of the robot «SUGV»

The considered gripping device (figure 5) was designed and intended for the manipulator of the mobile platform SUGV. Structurally, the gripper is a continuation of the manipulator and does not have its own degree of freedom. The drive part is located at the end link of the manipulator, which makes the gripping device compact. The gripper has two fingers that move plane-parallelly. The shape and structure of the fingers allow not only to grab and hold objects but also to rotate them, for example, to open door handles. Optionally, additional lights and a camera can be mounted at the gripping device.

The specificity of this gripping device, as well as the location of the drive part at the end link of the manipulator, allow it to be used only on the given manipulator.

Gripper of the manipulator «Servosila Robotic Arm»

The gripper of the Servosila Robotic Arm (figure 6) is designed to work with objects of various shapes and dimensions. The gripper is equipped with a quick-release sealed mount, which allows its use only with a suitable manipulator.

The fingers of the gripper are moving in parallel, the drive part is located at the base of the gripper. Dustproof and waterproof construction allows it to operate in different environments and weather conditions.

Figure 5 – Gripper «SUGV»

Figure 6 – Gripper «Servosila Robotic Arm»

Requirements for the gripping device

There is no possibility of quick installation and using a tool that extends the functionality of the considered grippers.

In the considered gripping devices of mobile robots, electric motors are used, due to the compactness of the system and ease of operation. In this case, the fingers move plane-parallelly, which allows capturing objects of various shapes and sizes, within the working area of manipulation device. In addition, the fingers can rotate about the axis of the gripper, which allows to perform simple technological operations, unscrewing or twisting. In this case, the driving part of the fingers movement is located at the base of the gripping device or in the previous link of the manipulator. The movement of the fingers is carried out via a worm gear or a gear with a rack, which provides the necessary compression force.

Based on the review and analysis of gripping devices, and also taking into account the specifics of the tasks performed, the following requirements can be formulated for a gripping device:

- work with interchangeable tool modules;
- two fingers;
- resistance to workloads;
- plane-parallelly movement of the fingers;
- coordination of working surfaces with a captured mating part of the modules;

– an additional degree of freedom in the gripper (rotation of the fingers around the axis of the gripper);
– holding an object weighing 3 kg.

For the full operation of the manipulator and the gripping device, it is necessary to carry out modeling and research the arising loads in the joints and the necessary torque of the motors. Further, a drilling simulation is considered as the most loaded. It is necessary to check the performance of the manipulator when performing this technological operation.

Modeling of technological operation (drilling)

For the manipulator, the four most common technological operations were selected (figure 7), such as gripping and moving objects, drilling, cutting wires, and unscrewing.

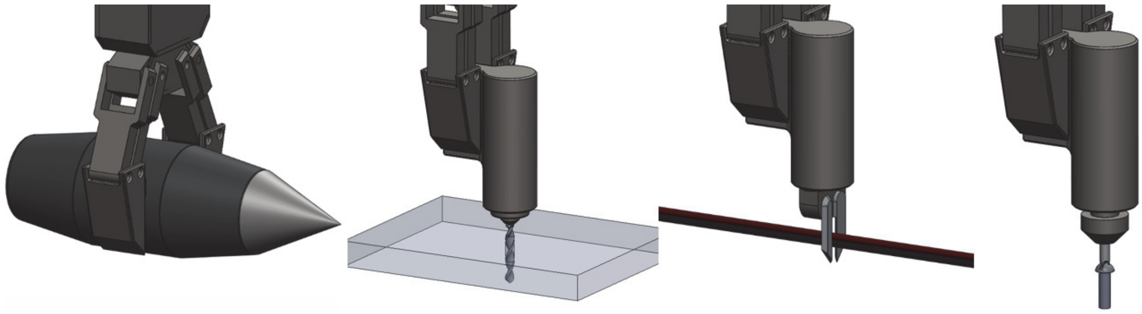


Figure 7 – Possible technological operations

For simulations of the technological operation, a model (figure 8) with a mass distributed over the links was created to analyze torque moments in the drives of motors and reaction forces in the joints of the manipulator.

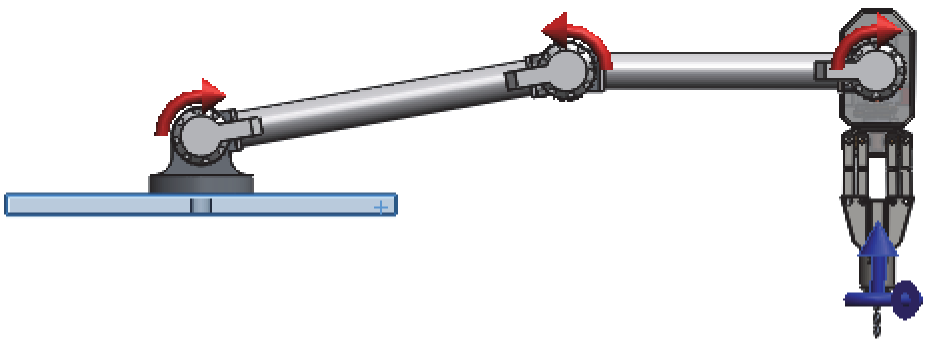


Figure 8 – Model for loads analyzing

The following mass distribution was used in the simulation:
– the first link, m2 – 1000 g;
– the second link, m3 – 800 g;
– the third link with the gripper, m4 – 550 g.

Link lengths:
– the first link, l2 – 375 mm;
– the second link звено, l3 – 325 mm;
– the third link with the gripper, l4 – 185 mm.

The motion of the manipulator was set as follows:
– from 0 to 5 seconds – setting the manipulator to the operating position;
– from 5 to 7 seconds – drilling with the force of pressing at the drill;
– from 7 to 9 seconds – exit from drilling, stopping the process;
– from 9 to 15 seconds – setting the manipulator to its original position.

The torque of an average BOSCH GSR 1440 screwdriver is taken as the maximum torque of the drill, equal to 30 N·m.

The rotational moment of drilling (figure 9) is set by an oscillatory function with an oscillation amplitude of 5 N·m and an average value of 25 N·m. This gives a variation of the rotational moment from 20 to 30 N·m.

The calculation uses a drill with a diameter of 6 mm. Such diameter is the maximum available for drilling. The pitch of the screw edges of the drill is approximately equal to two diameters, so it is taken equal to 12 mm.

Theoretical measurements show that 80 % of the total cutting moment is accounted for by the main cutting edges, 8 % – by the transverse edge and 12% – by the friction of the chips on the drill and ribbons on the machined surface. The power spent on the longitudinal flow is small and does not exceed 1.0 - 1.5 % of the total power spent on drilling. Therefore, for the maximum axial feed force, we take the maximum loading capacity of the manipulator of 3 kg or 30 N.

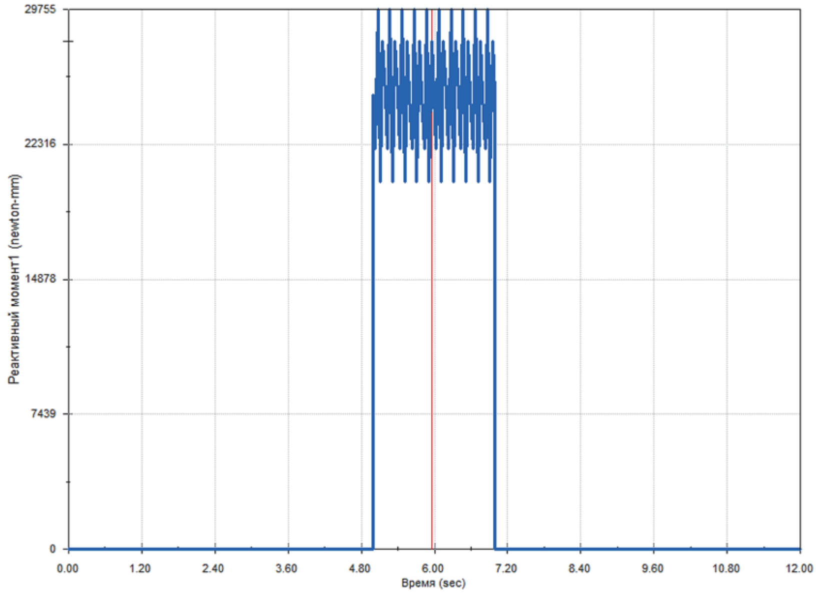


Figure 9 – Drill torque plot

Analysis of the joints

Consider the required torque of the motors and the reaction forces arising in the joints (figure 10).

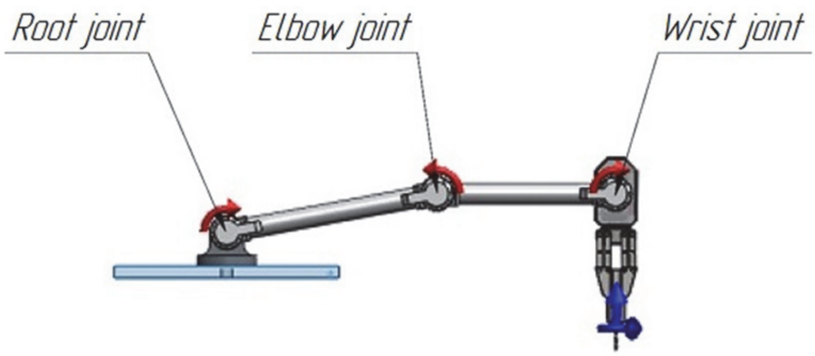


Figure 10 – Design model of the manipulator

The root joint is the most loaded in the manipulator. According to figure 11, when the manipulator moves and at the beginning of drilling, the torque transmission occurs, from lifting the gripper together with the tool to pressing into the drilled material. The motor needs to produce a torque of 30 N·m.

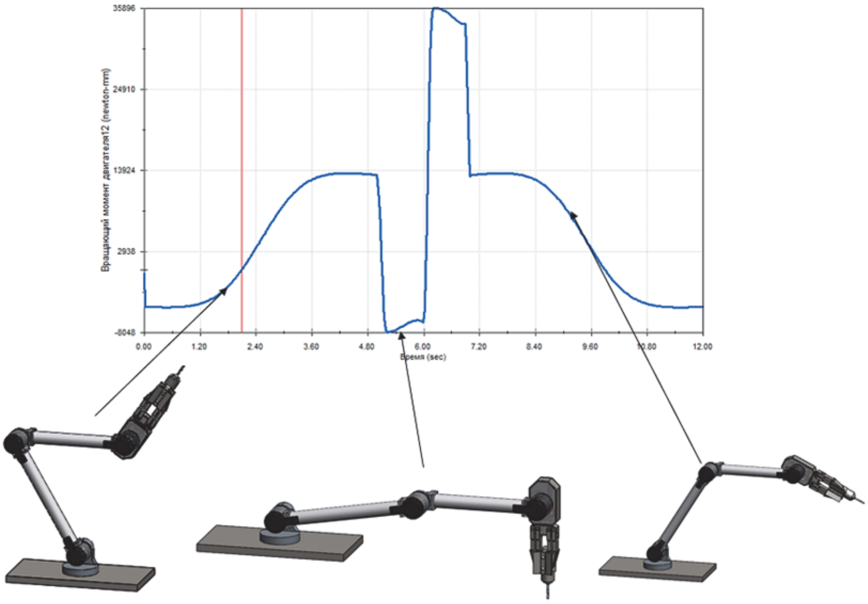


Figure 11 – Motor torque plot of the root joint

Based on the results, shown in figure 12, the reaction forces arising in the joint do not exceed 51 N. The resulting reaction forces are radial and do not cause axial displacement.

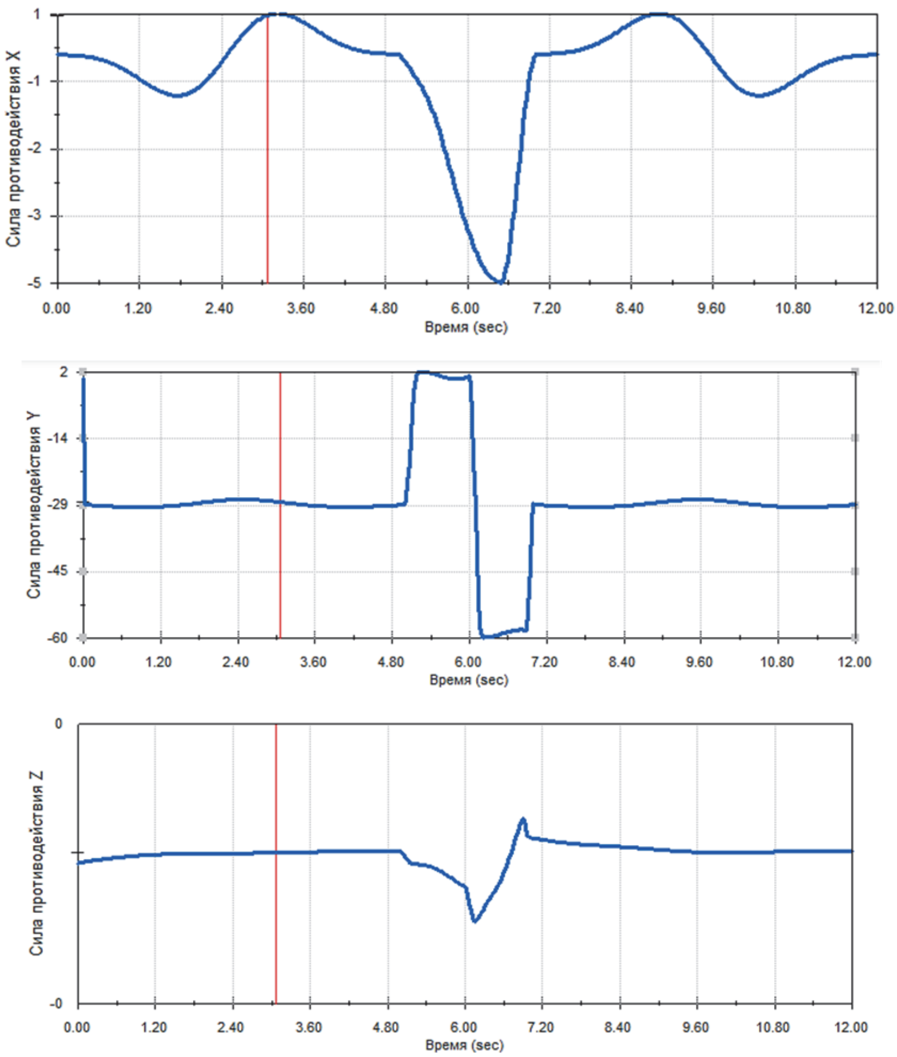


Figure 12 – Reaction forces plot of the root joint along the axes X, Y, and Z

The elbow joint is the second most loaded joint in the manipulator. According to figure 13, when the manipulator moves and at the beginning of drilling, the torque transmission occurs, from lifting the gripper together with the tool to pressing into the drilled material. The motor needs to produce a torque of 12.5 N·m.

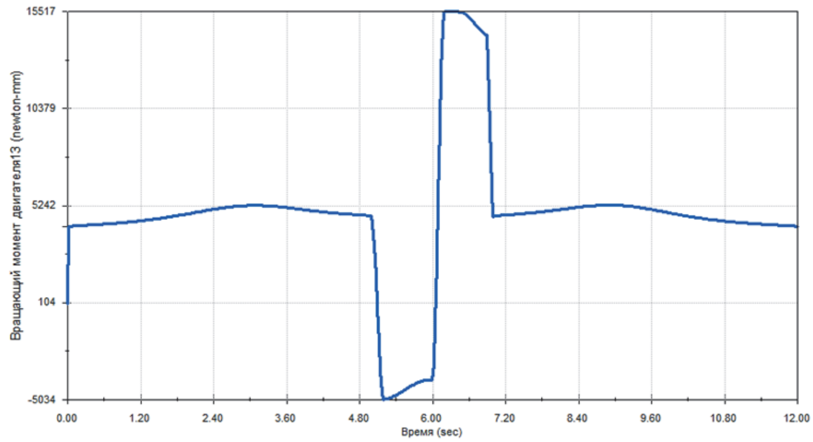


Figure 13 – Motor torque plot of the elbow joint

Based on the results, shown in figure 14, the reaction forces arising in the joint do not exceed 41 N. The resulting reaction forces are radial and do not cause axial displacement.

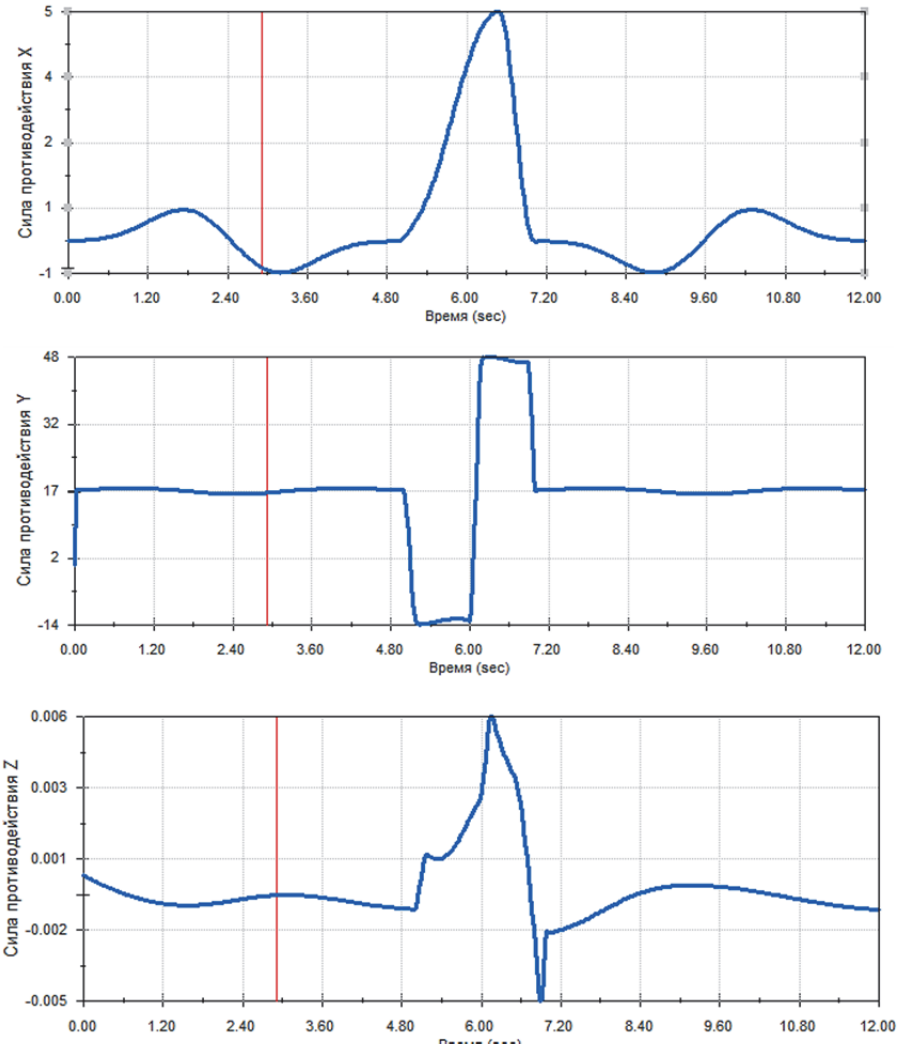


Figure 14 – Reaction forces plot of the elbow joint along the axes X, Y, and Z

The wrist joint is the least loaded joint in the manipulator. According to figure 15, when the manipulator moves and at the beginning of drilling, the torque transmission occurs, from lifting the gripper together with the tool to pressing into the drilled material. The motor needs to produce a torque of 1.2 N·m.

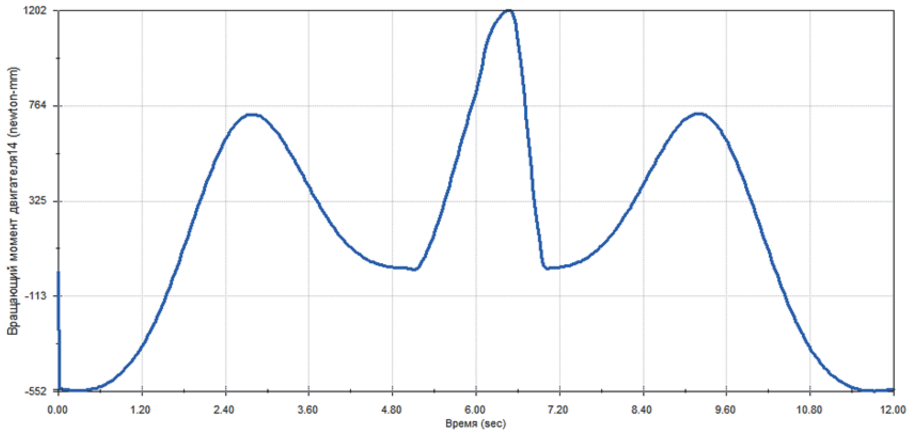


Figure 15 – Motor torque plot of the wrist joint

Based on the results, shown in figure 16, the reaction forces arising in the joint do not exceed 34 N. The resulting reaction forces are radial and do not cause axial displacement.

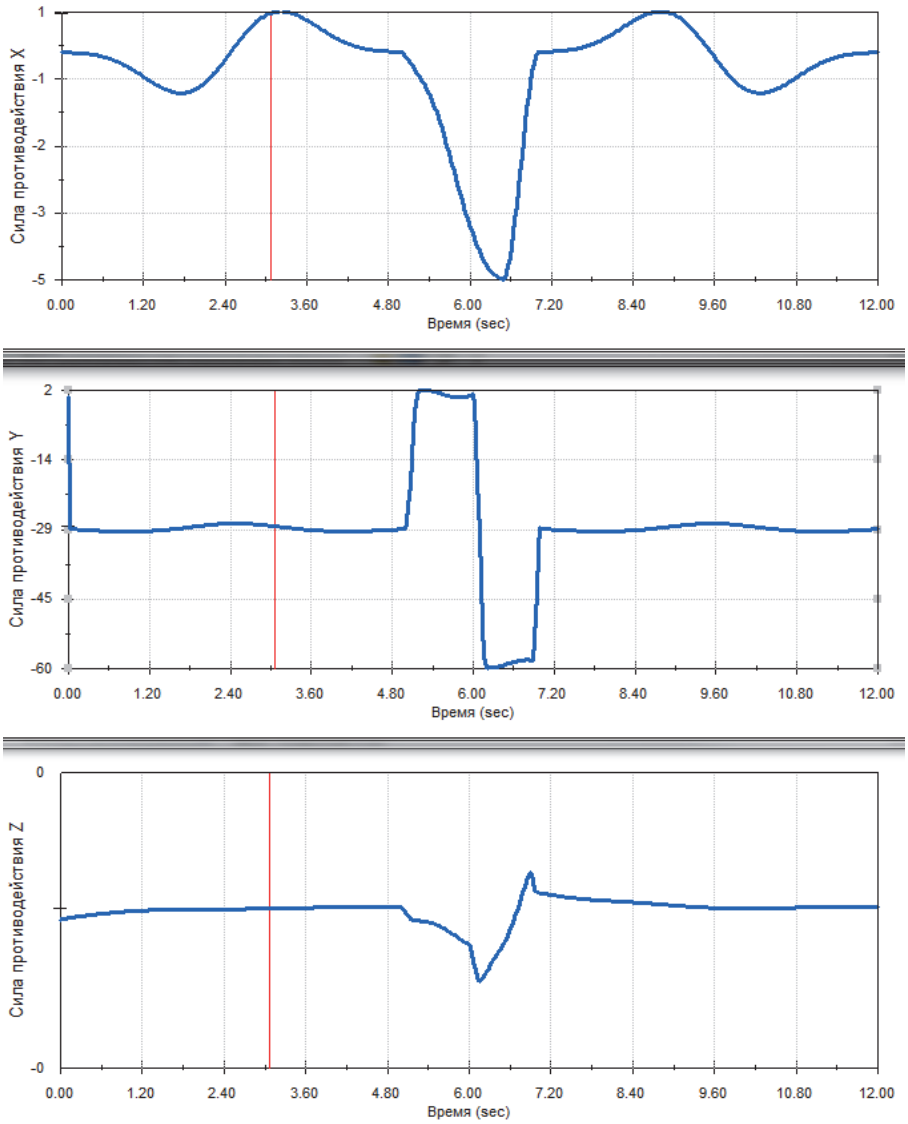


Figure 16 – Reaction forces plot of the wrist joint along the axes X, Y, and Z

In general, the simulation results show that the joints need to be strengthened to ensure the performance under these loads.

Design of the manipulator and the gripping device

The gripping device (figure 17) is driven by two electric motors. The first electric motor moves the gripping fingers. Through the coupling, the movement is transmitted to the intermediate shaft with the spur gear. Then the next gear drives the worm wheel shaft and the worm, which transmits the movement to the fingers rigidly connected to the worm wheel. The second electric motor rotates the fingers around the gripper axis. The gear motor transmits movement to the block with fingers through cylindrical gears, one of which is mounted on the output shaft of the gear motor, and the second on the block with fingers in size with it.

Interchangeable tool modules are captured by the fingers of the gripper and held in them using a special adapter (figure 18), which also provides power.

Overall dimensions of the manipulator in the folded state are 463 x 182 x 97 mm (figure 19), with a mass of 2.9 kg.

For unification, the manipulator drives have similar structure and design. The drive is an open-frame motor with a developed planetary gearbox, located inside the aluminum shell (figure 20). The ILM50x14 motor is used in the root and elbow drives, but the elbow drive uses two steps in the planetary gearbox, and the root drive uses three steps. The wrist joint uses an ILM50x8 motor and two steps of a planetary gearbox.

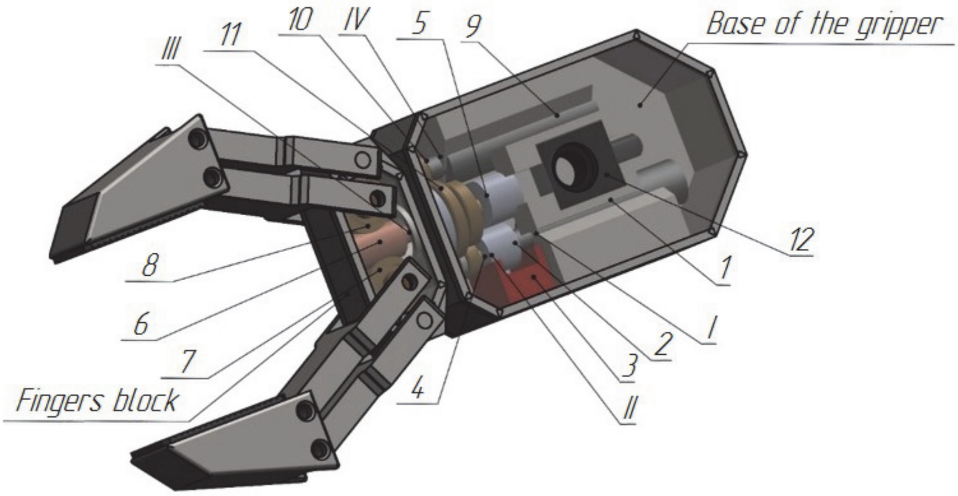


Figure 17 – 3D model of the gripping device:
pos. 1, 9 – electric motors, pos. 2 – coupling, pos. 3 – brake, pos. 4-5, 10-11 – spur gears,
pos. 6 – worm, pos. 7-8 – worm wheels, pos. 12 – three-component force sensor,
pos. I, IV – motor output shafts, pos. II – intermediate shaft, pos. III – worm wheel shaft

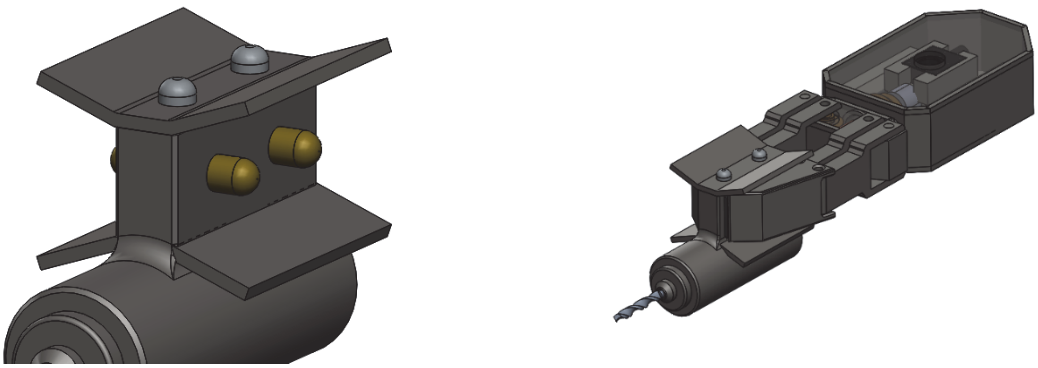


Figure 18 – Adapter for the gripping device (on left),
and the gripping tool in the gripper (on right)

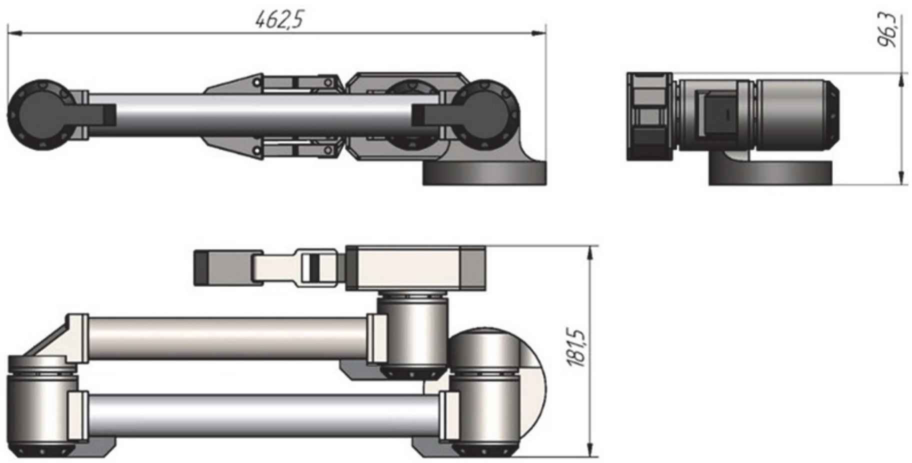


Figure 19 – Developed manipulator

The first link of the manipulator is located on the base and includes two drives connected by a shoulder, representing a tube with welded flanges for fastening to the drives (figure 21). The second link of the manipulator is attached to the first with special fastening. At the second end of the link, there is a drive carrying the gripping device (figure 21).

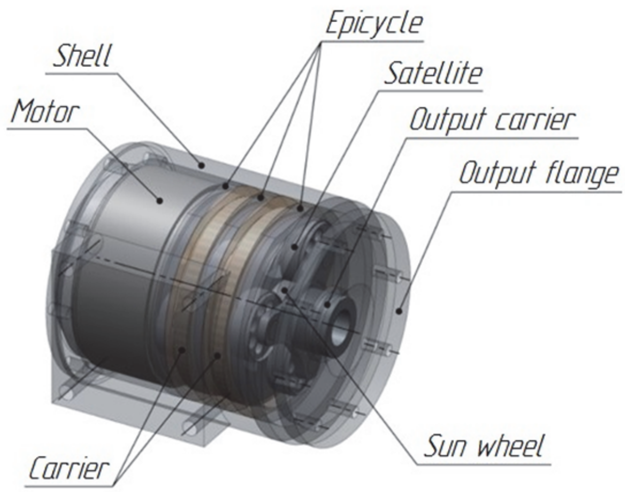


Figure 20 – Drive of the manipulator joint (some parts are shown transparently for clarity)

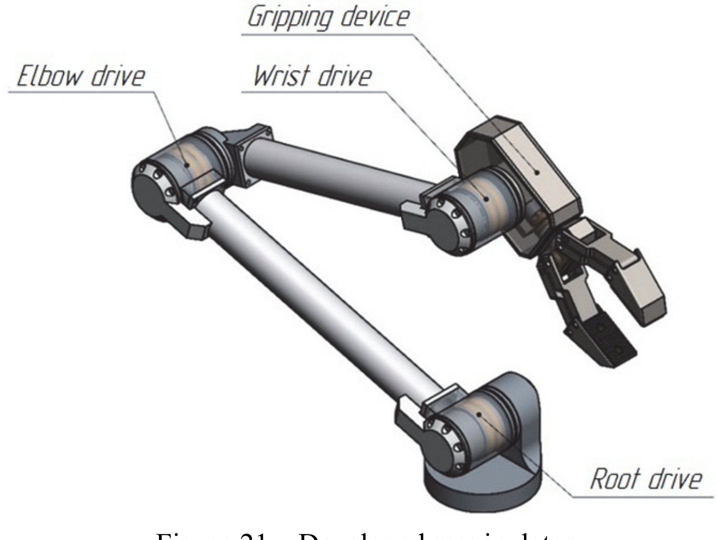


Figure 21 – Developed manipulator

Conclusion

An analytical review was conducted, the advantages and disadvantages of the existing structures of manipulators and gripping devices for mobile platforms weighing from 5 to 15 kg were revealed. The configuration and geometrical dimensions of the links of the manipulator were determined. The type of gearboxes in the joints of the manipulator was selected. The gripping device structure was selected.

The proposed manipulator can be used on various mobile robotic systems with a mass from 5 to 15 kg, responding to the formulated requirements. Interchangeable tool modules will significantly expand the functionality of the robots, and equipment replacement without additional intervention will significantly reduce the time of performing a number of tasks.

References

1. Vasiliev A.S., Schukin P.O. Kinematicheskie shemy [Kinematic circuits]. Textbook for students of engineering colleges. Petrozavodsk: PetrSU. 2013. URL: http://forest.petrsu.ru/courses/dm/kinematic_schemes.pdf.
2. Servosila Robotic Arms. URL: https://www.servosila.com/en/roboticarms/
3. reSquared robotics. URL: http://www.resquared.com/ products/manipulators/
4. EndeavorRobotics. URL: http://endeavorrobotics.com/products#310-sugv
5. Robotiq 2F-140. URL: https://robotiq.com/products/2f85-140-adaptive-robot-gripper?ref=nav_product_new_button
6. Shadow lit. URL: https://www.roscomponents.com/en/robotic-hands/116-shadow-lite-hand.html#/shadow_lite_lite_fingers-extra_lite_2_fingers_y_1_thumb
7. Shahinpur M., Kurs robototehniki [Robotics course]. Moscow: Mir Publ. 1990. p. 527.

V.I. Petrenko, F.B. Tebueva, V.O. Antonov, V.B. Sychkov, M.M. Gurchinsky

ANTHROPOMORPHIC MANIPULATOR MOTION PLANNING FOR COPYING CONTROL

North-Caucasus Federal University, Stavropol, Russia
gurcmikhail@yandex.ru

Abstract

One of the main tasks of robotics is to replace a person when working in conditions dangerous to his life and health. The functionality and effectiveness of robotic systems is determined by the control method used. To solve complex problems in a non-deterministic environment, the use of anthropomorphic robots with copying control is promising. As a master for copying control, systems based on various principles can be used. The master device in the form of an exoskeleton combines mobility and the possibility of realizing moment-force sensation. Unlike the copying manipulator system, where the master and slave are proportional to each other, in the general case, the exoskeleton and executive manipulator are not proportional or even have a different kinematic structure. One of the tasks that arise when implementing copying control using a master in the form of an exoskeleton is planning the movement of an anthropomorphic manipulator based on known information about the position of the operator's hand. The purpose of the article is to calculate the generalized coordinates of the anthropomorphic manipulator with a copying control that satisfy the conditions for the effectiveness of copying control. As conditions for the effectiveness of copying control, we used the conditions for maximizing the involved operating space of the executive manipulator, coordination of the orientation of the wrist link of the manipulator and the palm of the operator, similarity to the orientation of the plane formed by the wrist, elbow and shoulder joints of the manipulator, and similar operator joints. To calculate the generalized coordinates of the manipulator, an analytical calculation technique was proposed based on the Denavit-Hartenberg representation and Euler angles. The novelty elements of the presented solution is an analytical calculation of the generalized coordinates of the manipulator that satisfy complex requirements. The practical significance of the analytical calculation of the generalized coordinates of the manipulator lies in the low computational complexity, which determines the relevance of using real-time copying control systems.

Keywords: anthropomorphic manipulator, anthropomorphic robot, copying control.

Introduction

One of the main tasks of robotics is to replace a person when working in conditions dangerous to his life and health. One of the factors determining the functionality and efficiency of robotic systems is the control method used. Often, the robotic complex is tasked with performing heterogeneous targeted operations in a complex non-deterministic environment. For example, in [1] the task of robotic servicing of a target satellite is described, which generally does not meet the standards of a service robot, which still remains an open research area, which is associated with many technical problems. One of the biggest problems is ensuring safe and reliable maintenance of the target spacecraft with the help of a service one, or capturing a target for its stabilization and subsequent maintenance. The task is further complicated if the target has unknown motion characteristics. It is obvious that the current level of automatic control systems is not able to ensure the implementation of this operation with high reliability. The presence of a large number of classes of tasks that only a person can handle determines the prospects of using anthropomorphic robots with a copy control method. Using the copy control allows you to realize the virtual presence of the operator in the body of the robot. In [2], a hardware-software complex is presented, designed to perform basic technological operations at the International Space Station. The composition of the complex includes a copy type master and an anthropomorphic second-generation SAR-401 robot, as well as a control system. Testing of the developed complex was carried out at the research and testing center for cosmonaut training in 2013 and 2015, the results of which showed the possibility of using the copy type control to perform typical space operations.

Copy control can be implemented using various technologies. In [3], a technique is proposed in which the manipulator is controlled using a Kinect sensor. To form the movement of the manipulator that copies the movement of the operator's hand, an algorithm is used to solve the inverse kinematics problem. The movement of a human hand in three-dimensional space is captured, processed and replicated by a robotic arm. In [4], the concept of robot teleoperation was presented, which tracks the position of the operator using the Kinect motion capture device and the use of algorithms for solving the inverse kinematics problem. The methods were demonstrated on the Mitsubishi Movemaster EX with five degrees of freedom. The disadvantages of non-contact motion capture are the inability to implement force moment feedback and the need to collect and synchronize visual information from several angles to ensure reliable motion capture.

An alternative and the basis of the first copy control systems is the use of lever systems as master devices. A known system of the copy manipulator [5], used for remote work with radioactive substances. The master and executive manipulators are proportional to each other, which makes the formation of control laws simple. The value of the generalized coordinates of the executive manipulator fully corresponded to the similar values of the master manipulator.

An alternative to stationary copy control systems are mobile copy control systems using a master in the form of an exoskeleton. In [6–7], a system for copying movements using an exoskeleton is considered. An analysis of the performance of a motion copying system based on information about the speed of a robot with two degrees of mobility is presented in [6]. In [7], a system for copying movements is proposed. The control system is constructed in such a way that the desired movement of the master body, specified by the human operator, is completely copied up to a scale factor by the executive body, for which purpose tracking systems are used that use feedback on the efforts or moments that arise during the operation of the executive body.

In [8–9], the problem of the accuracy of positioning of the actuator during copy control of the robot is investigated. In work [8], a parallel manipulator for minimally invasive surgery is presented, the remote center of which is capable of generating control actions with four independently controlled degrees of mobility. The geometry and decomposability of the manipulator motion are investigated, and the problem of its inverse kinematics is solved. The domestic development of an anthropomorphic gripping manipulator with high accuracy of copying the movement of a human hand is presented in [9]. The application of the method of planning the trajectory of movement allows you to pre-calculate the parameters of movement of the joints. In [10–11], the solution of problems of dynamics and kinematics is considered to determine the spatial position of the nodal points of the manipulator. In [10], the kinematic model of the Fanuc AM100iB robot was developed and tested. The test consisted in the robot working in the mode of controlling the values of natural angles at selected points in its workspace and reading the readings of the coordinates of the point in the global coordinate system of the robot on the operator panel.

One of the tasks to be solved when implementing copying control is to determine the rotation angles of the operator's hand, based on the measured values of the generalized coordinates of the exoskeleton. The article [11] describes the development of an exoskeleton with an excessive degree of mobility for restoration of the upper limbs. The technique of analytical solution of the inverse problem of the kinematics of exoskeleton movement is presented to ensure synchronous movement with patients and ensure the natural interaction of humans and robots. For the mathematical description of excessive mobility, the angle of rotation of the elbow is introduced as an additional parameter for determining the position of the human hand with a predetermined location of the wrist. A kinematic criterion is proposed for determining the angle of rotation by simulating the natural reflexive reaction of a human hand. Experimental results show that the kinematic criterion for determining the angle of rotation is suitable for describing free motion without additional restrictions. Moreover, with the presented rotation angles, the root-mean-square errors between the actual and calculated hinge angles were not more than 8°. In [12], a complex technique for calculating the rotation angles of an operator's hand based on the rotation angles of a master device in the form of an exoskeleton is presented.

Another important task is to determine the rotation angles of the actuator. Unlike the copy manipulator, the operator's hand and the executive manipulator are not proportional to each other. In addition, the task is complicated by the presence of two executive manipulators, cooperatively performing target operations. If the mismatch of the movement of one manipulator with the movement of the operator's hand can be compensated for by the inherent proprioception, then with two manipulators some problems cannot be solved in this way. For example, even with equal lengths of the links of the manipulator and parts of the operator's arm, if the «shoulder width» of the robot, that is, the distance between the shoulder joints, is greater than the width of the operator's shoulders, the situation shown in Figure 1 is possible when the manipulator copies the rotation angles in the joints of the operator's arm.

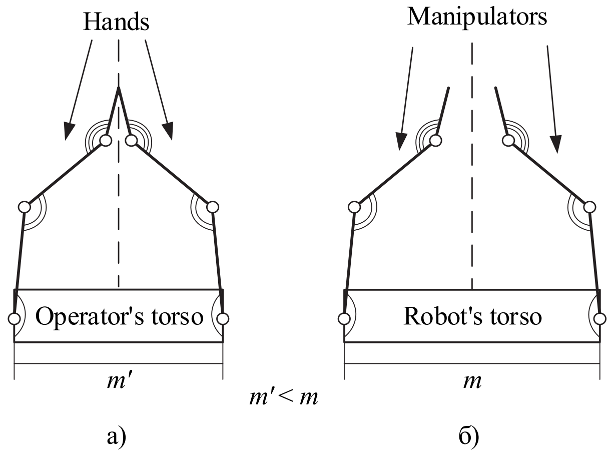


Figure 1 – Problematic situation of copying rotation angles

Due to the fact that the «shoulder width» of the robot is greater, when the operator's hands are brought down palm by palm, there will be a gap between the wrist links of the manipulator. Therefore, the urgent task of calculating the rotation angles of the anthropomorphic manipulator when copying control with a master in the form of an exoskeleton.

Anthropomorphic manipulator motion planning

The kinematic diagram of the manipulator under consideration is shown in Figure 2.

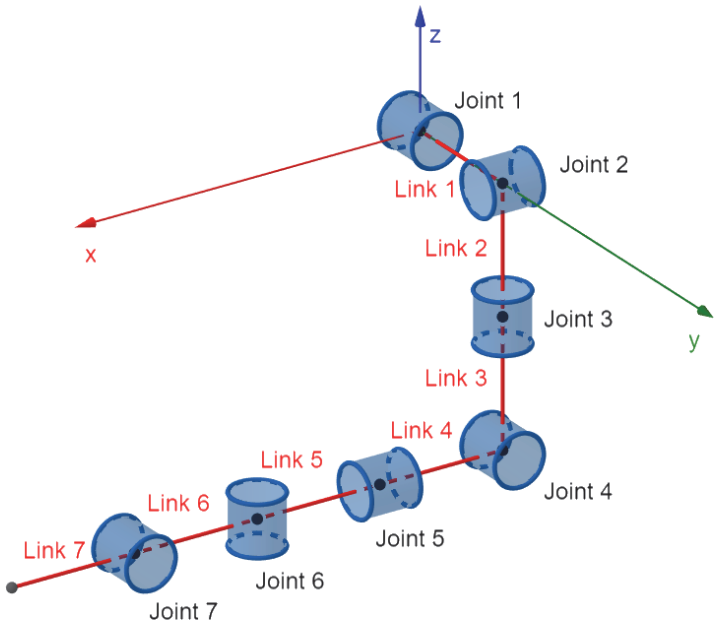


Figure 2 – The manipulator kinematic diagram

The number of rotational degrees of mobility $n = 7$. Joints 1-3 match the shoulder joint, joint 4 – elbow joint, joints 5-7 – wrist joint. Let us designate the part of the manipulator between the shoulder and elbow joints as the shoulder link, between the elbow and wrist joints as the forearm link, between the wrist joint and the working end as the carpal link.

To describe the parameters of the kinematic scheme, we use the Denavit-Hartenberg representation. In accordance with the Denavit-Hartenberg representation, each i-th link is described by a set of four parameters: a, d, α, θ one of which characterizes the movement of the manipulator in i-th jopint and is the generalized coordinate of the manipulator. The system of seven links is respectively described by parameter vectors $\mathbf{a}, \mathbf{d}, \boldsymbol{\alpha}, \boldsymbol{\theta}$. Since all the kinematic pairs of the manipulator are rotational, the vector of generalized coordinates that uniquely describes the position of the manipulator is $\boldsymbol{\theta}$.

In accordance with the Denavit-Hartenberg representation, each link of the manipulator is associated with a related coordinate system according to a certain rule. The direction and location of the associated coordinate systems of the manipulator under consideration is shown in Figure 3.

The homogeneous transformation matrix from the i -th coordinate system to the j -th can be found by the formulas [13]:

$$ {}^{i}\boldsymbol{T}_{j} = \prod_{k=i+1}^{j} {}^{i-1}\boldsymbol{A}_{i},\ i < j, \tag{1} $$

$$ {}^{i-1}\boldsymbol{A}_{i} = \boldsymbol{T}_{z,\theta}\left(\theta_{i}\right)\boldsymbol{T}_{z,d}\left(d_{i}\right)\boldsymbol{T}_{x.a}\left(a_{i}\right)\boldsymbol{T}_{x,\alpha}\left(\alpha_{i}\right), \tag{2} $$

$$ \boldsymbol{T}_{z,\theta}\left(\theta\right) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3} $$

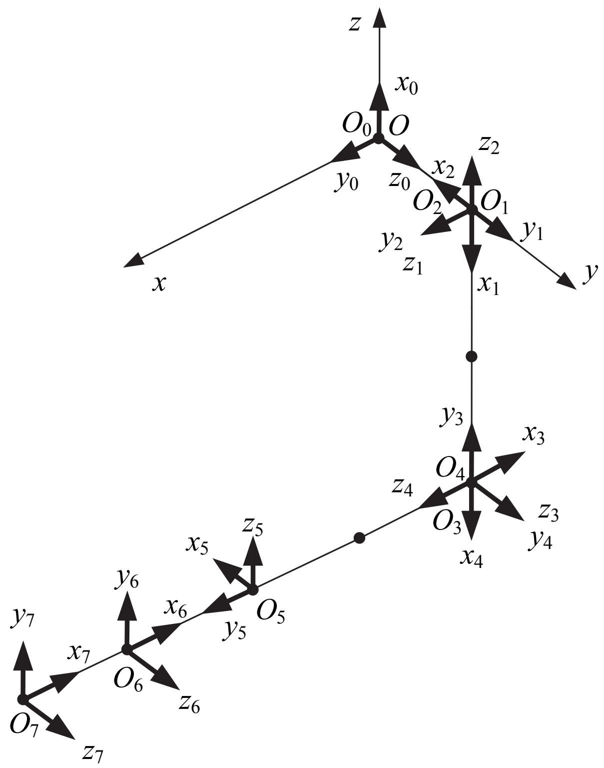


Figure 3 – Coordinate systems associated with the links of the manipulator

$$ \boldsymbol{T}_{z,d}\left(d\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{4} $$

$$ \boldsymbol{T}_{x.a}\left(a\right) = \begin{bmatrix} 1 & 0 & 0 & a \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5} $$

$$\boldsymbol{T}_{x,\alpha}(\alpha)=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(\alpha) & -sin(\alpha) & 0 \\ 0 & sin(\alpha) & cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

where ${}^{i}\boldsymbol{T}_{j}$ – transformation matrix from j -th to i -th coordinate system;

${}^{\mathbf{i-1}}\mathbf{A}_{\mathbf{i}}$ – homogeneous matrix of complex transformation between adjacent coordinate systems;

$\mathbf{T}_{\mathbf{z},\boldsymbol{\theta}}(\theta)$ – homogeneous matrix of elementary rotation about the z axis by an angle θ;

$\mathbf{T}_{\mathbf{z},\mathbf{d}}(d)$ – homogeneous displacement matrix along the z axis by d;

$\mathbf{T}_{\mathbf{x}.\mathbf{a}}(a)$ – homogeneous displacement matrix along the x axis by a;

$\mathbf{T}_{\mathbf{x},\boldsymbol{\alpha}}(\alpha)$ – homogeneous matrix of elementary rotation about the x axis by an angle α;

$\mathbf{a},\mathbf{d},\boldsymbol{\alpha},\boldsymbol{\theta}$ – Denavit-Hartenberg parameter vector columns describing the kinematic structure of the manipulator.

The desired position of the manipulator in the space of generalized coordinates must satisfy the following conditions. The condition for maximizing the use of the operating space of the manipulator has the form:

$$\frac{|\mathbf{O}_7|}{r_m}=\frac{|\mathbf{O}'_7|}{r'_m}, \tag{7}$$

where $\mathbf{O}_7$ – radius-vector of the working end of the manipulator;

r_m – maximum outstretched manipulator lenght;

$\mathbf{O}'_7$ – radius vector of the center of the palm of the operator;

r'_m – maximum outstretched arm length.

The condition for coordinating the orientation of the operator’s brush and the manipulator’s hand link:

$$\begin{cases} \varphi_h=\varphi'_h, \\ \theta_h=\theta'_h, \\ \psi_h=\psi'_h, \end{cases} \tag{8}$$

where φ_h, θ_h, ψ_h – Euler angles of the manipulator hand;

φ'_h, θ'_h, ψ'_h – Euler angles operator hand.

The condition for the similarity of the orientation of the plane formed by the wrist, elbow and shoulder joints of the manipulator, and similar joints of the operator is:

$$\begin{cases} \dfrac{|\mathbf{O}_7|}{r_m}=\dfrac{|\mathbf{O}'_7|}{r'_m}, \\ \dfrac{|\mathbf{o}_7|}{r_m}=\dfrac{|\mathbf{o}'_7|}{r'_m}, \\ \theta_2=\theta'_2, \end{cases} \tag{9}$$

where $\mathbf{o}_7$ – vector connecting the shoulder joint of the second manipulator, and the working end of the first;

$\mathbf{o}'_7$ – vector connecting the shoulder joint of the second hand of the operator, and the center of the palm of the first;

θ_2 and θ_2' – angle of rotation in the joint 2 of the manipulator and a similar angle of rotation in the shoulder joint of the operator's hand.

The purpose of the article is to calculate the angle of rotation of the manipulator $\boldsymbol{\theta}$, satisfying the conditions for maximizing the use of the operating space of the manipulator, matching the orientation of the operator's hand and the wrist link of the manipulator, similarity to the orientation of the plane formed by the wrist, elbow and shoulder joints of the manipulator, and similar operator joints based on the following initial data:

– manipulator's constant parameters of the Denavit-Hartenberg $\mathbf{a}, \mathbf{d}, \boldsymbol{\alpha}$;

– the distance between the shoulder joints of the manipulator m;

– radius vector of the center of the palm of the operator $\mathbf{O}_7'$;

– vectors connecting the shoulder joint of the second hand of the operator, and the center of the palm of the first $\mathbf{o}_7'$;

– Euler angles operator hand φ_h', θ_h', ψ_h';

– rotation angle θ_2' in the shoulder joint describing the adduction/abduction of the operator's arm in the coronary plane.

To solve this problem, the following calculation method is proposed.

The calculation of the maximum length r_m of the extended arm:

$$r_m = l_1 + l_2 + l_3, \tag{10}$$

where l_1, l_2, l_3 – the length of the shoulder, forearm and carpal links of the manipulator.

Based on the maximum length r_m of the extended arm, the distance from the center of the palm of the hand of the operator to the center of the shoulder joint is recalculated based on the condition for maximizing the involved operating space of the manipulator:

$$\frac{|\mathbf{O}_7|}{r_m} = \frac{|\mathbf{O}_7'|}{r_m'}, \tag{11}$$

$$|\mathbf{O}_7| = r_m \frac{|\mathbf{O}_7'|}{r_m'}. \tag{12}$$

Similarly, the distance from the center of the grip to the origin of the second manipulator is calculated:

$$|\mathbf{o}_7| = r_m \frac{|\mathbf{o}_7'|}{r_m'}. \tag{13}$$

Vectors $\mathbf{O}_7$, $\mathbf{o}_7$ and the line connecting the shoulder joints of both manipulators of the anthropomorphic robot form a triangle, for which the cosine theorem is valid:

$$\mathbf{o}_7{}^2 = \mathbf{O}_7^2 + m^2 - 2|\mathbf{O}_7| m \cos\varphi, \tag{14}$$

where φ – the angle between the line passing through the shoulder joints of both manipulators of the anthropomorphic robot and the vector $\mathbf{O}_7$.

Angle φ can be found by the formula:

$$\varphi = \arccos\left(\frac{\mathbf{O}_7^2 + m^2 - \mathbf{o}_7{}^2}{2|\mathbf{O}_7| m}\right), \tag{15}$$

and the coordinates of the working end on the plane in question are expressed in terms of the angle φ:

$$O_{7x} = |\mathbf{O}_7| \cdot \cos\varphi, \tag{16}$$

$$O_{7y} = |\mathbf{O}_7| \cdot \sin\varphi. \tag{17}$$

Projection O_{7z} of vector $\mathbf{O}_7$ on axis z can be found with scaling factor:

$$O_{7z} = O'_{7z}\frac{r_m}{r'_m}. \tag{18}$$

Based on the known coordinates of the center of the tong, a homogeneous transformation matrix can be found from the coordinate system associated with the tong into the global coordinate system:

$$\mathbf{T}_7 = \begin{pmatrix} C\varphi_h C\theta_h & C\varphi_h S\theta_h S\psi_h - S\varphi_h C\psi_h & C\varphi_h S\theta_h C\psi_h + S\varphi_h S\psi_h & O_{7x} \\ S\varphi_h C\theta_h & S\varphi_h S\theta_h S\psi_h + C\varphi_h C\psi_h & S\varphi_h S\theta_h C\psi_h - C\varphi_h S\psi_h & O_{7y} \\ -S\theta_h & C\theta_h S\psi_h & C\theta_h C\psi_h & O_{7z} \\ & 0 \quad 0 \quad 0 & & 1 \end{pmatrix}, \tag{19}$$

where $C\varphi_h$, $C\theta_h$ and $C\psi_h$ – cosines of angles φ_h, θ_h and ψ_h respectively;

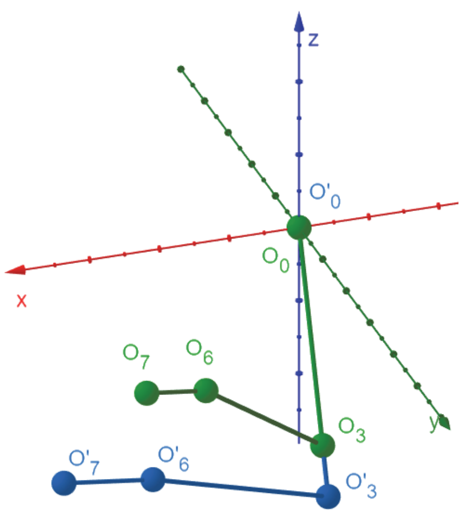
$S\varphi_h$, $S\theta_h$ and $S\psi_h$ – sinuses of angles φ_h, θ_h and ψ_h respectively.

Based on a known matrix $\mathbf{T}_7$ the position in the Cartesian coordinates of the wrist joint of the manipulator can be found:

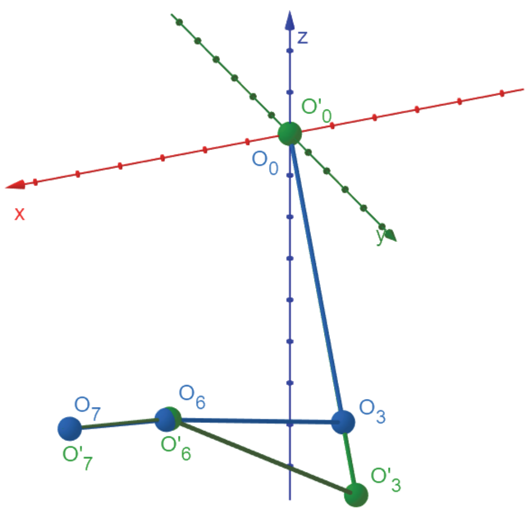
$$\mathbf{O}_6 = \mathbf{T}_7 \cdot \begin{pmatrix} l_3 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{20}$$

Based on the known positions of the joints in Cartesian coordinates, subject to restrictions $\theta'_2 = \theta_2$, imposed by a condition of similar orientation of the plane formed by the wrist, elbow and shoulder joints of the manipulator, and similar operator joints, the rotation angles of the manipulator can be calculated using the method for solving the inverse kinematics of the excess manipulator described in [12].

Visualization of the results of calculating the angle of rotation of the manipulator using the proposed mathematical apparatus is shown in Figure 4. As can be seen from Figure 4b, despite the disproportionality of the manipulator and the operator’s hand, the wrist link of the manipulator is oriented just like the operator’s brush, and the plane formed by the shoulder, elbow and wrist the joints of the manipulator coincides with the same plane of the operator’s hand.



a)



b)

Figure 4 - Example of calculation results: a) absolute scale; b) relative scale

Conclusion

The article proposes an analytical technique for planning the movement of an anthropomorphic manipulator with copying control using a master in the form of an exoskeleton. The proposed technique allows us to calculate the angle of rotation of the manipulator, satisfying the conditions for maximizing the use of the operating space of the manipulator, matching the orientation of the operator's hand and the wrist link of the manipulator, similar to the orientation of the plane formed by the wrist, elbow and shoulder joints of the manipulator, and similar operator joints. Low computational complexity and accounting for complex requirements for copying control determine the prospect of applying the proposed methodology in existing systems of copying control.

References

1. Wu, Q.-C., Wang, X.-S., Du, F.-P. Analytical Inverse Kinematic Resolution of a Redundant Exoskeleton for Upper-Limb Rehabilitation. (2016) International Journal of Humanoid Robotics, 13 (3), article № 1550042.
2. Rose, C.G., Kann, C.K., Deshpande, A.D., O'Malley, M.K. Estimating anatomical wrist joint motion with a robotic exoskeleton. (2017) IEEE International Conference on Rehabilitation Robotics, статья № 8009450, pp. 1437-1442.
3. Wang, M., Luo, J., Yuan, J., Walter, U. Detumbling strategy and coordination control of kinematically redundant space robot after capturing a tumbling target. (2018) Nonlinear Dynamics, pp. 1-21.
4. Shi, Z., Huang, X., Hu, T. Human motion capture similarity control for space teleoperation. (2018) Lecture Notes in Electrical Engineering, 445, pp. 263-279.
5. Yurevich, E.I. Upravlenie robotami i robototekhnicheskimi sistemami [Management of robots and robotic systems]. SPb: SPbSTU, 2000. 171 p.
6. Yu, W., Rosen, J., Li, X. PID admittance control for an upper limb exoskeleton // PID admittance control for an upper limb exoskeleton. 2011. Pages 1124-1129.
7. Margineanu, D., Lovasz, E.-C., Gruescu, C.M., Ciupe, V., Tatar, S. 5 DoF haptic exoskeleton for space telerobotics – Shoulder module (2018) Mechanisms and Machine Science, 52, pp. 111-120.
8. Yajima, S., Katsura, S. Multi-DOF motion reproduction using motion-copying system with velocity constraint. – (2014) IEEE Transactions on Industrial Electronics, 61 (7), статья № 6636073, pp. 3765-3775.
9. Igarashi, K., Katsura, S. Motion-data processing and reproduction based on motion-copying system (2015) IEEJ Journal of Industry Applications, 4 (5), pp. 543-549.
10. Jia, Q., Zhang, Q., Gao, X., Chen, G., Song, J. Dynamic obstacle avoidance algorithm for redundant robots with pre-selected minimum distance index. (2013) Jiqiren/Robot, 35 (1), pp. 17-22.
11. Zhou, S., Wu, X., Qi, Y., Gong, W. Visual location method of indoor mobile robot based on beeline detection. (2016) Huazhong Keji Daxue Xuebao (Ziran Kexue Ban)/Journal of Huazhong University of Science and Technology (Natural Science Edition), 44 (10), pp. 93-97.
12. Petrenko V.I., Tebueva F.B., Sychkov V.B., Gurchinsky M.M., Antonov V.O. Calculating rotation angles of the operator's arms based on generalized coordinates of the master device with following anthropomorphic manipulator in real time // International Journal of Mechanical Engineering and Technology (IJMET). 2018. Vol. 9. № 7. pp. 447-461.
13. Fu, K., Gonzalez, P., Lee, K. Robototekhnika [Robotics]. Moscow: Mir, 1989. 621 p.

V.V. Epaneshnikova, P.V. Koroletskiy, V.E. Pryanichnikov, E.A. Prysev, O.V. Punenkov

DESIGNING UNINTERRUPTED WIRELESS COVERAGE NETWORK ENSURING THE SERVICE ROBOT CONTROL

M.V. Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences (KIAM), Moscow, Russia,
IINET RSUH, STANKIN, Moscow, Russia
v.e.pr@yandex.ru

Abstract

While building a supervisor robot control, even in extreme conditions, Wi-Fi wireless communications are allowed (for example, we implemented multi-stream video transmission for the Brokk-type EMERCOM robots). However, in this case there are two problems - it is necessary to foresee that the attenuation of the signal upon detection of coverage areas and the provision of switching control using the automatic return algorithm. In our work, a number of approaches to assessing the quality and quality of work were proposed. An experimental study of this parameter was carried out in robotized systems (IINET RGGU – Keldysh Institute of Applied Mathematics RAS), consisting of many rooms in which the control of mobile service robots Amur-307, Amur-105, Robotino and others.

To control the service robot, seamless roaming (handover) technology was implemented - providing Wi-Fi coverage, with access to several access points (clients), switching between zones, switching from one point to another without losing the signal [1]. When designing such a seamless wireless network, one of the most important conditions affecting the quality of the built network infrastructure is to determine the degree of attenuation of the wireless signal.

With the help of specialized software, a survey of parts of the building was conducted. Wall types and the actual scale of the plan were identified, and inconsistencies were corrected. D-Link DWL-3200AP provides the optimal location of access points to the most important signal transmission zones, taking into account the existing wired infrastructure. To reduce the level of interference, it was decided to use non-overlapping channels (1, 6, 11) at 2.4 GHz. The implementation of the project to cover the premises with uninterrupted Wi-Fi signal coverage allowed not only to manage service robots, but also to create a public Wi-Fi infrastructure for visitors without the need for reconfiguration.

Keywords: mobile service robots, WiFi seamless coverage for supervisory management.

The first step in building a wireless network is to design it, taking into account the radio survey of all rooms. The plan of the rooms was drawn up, optimal places were chosen to install of access points, taking into account the absorbing capabilities of walls, doors and other objects.

The project of a wireless Wi-Fi network should always include a radio inspection of the object at the design stage and prior to the installation of the equipment. Radio testing was carried out at the design stage to determine the configuration and optimal installation sites of equipment to ensure uninterrupted and high-quality network operation. There are two main types of radio survey:

1. Passive examination. It’s conducted with a test access points, that are selected to build the network. The level of the received signal is measured, and the signals of all access points in the study zone are recorded in the 2.4 GHz and 5 GHz bands.

2. Active examination. It implies the additional use of metering devices that imitate a real subscriber device (mobile phone, tablet, laptop, barcode scanner) communicating with the access point. Measurement of real data: connection speed, packet loss, switching between points, etc., diagnostics of client devices.

In addition to surveys using actual measurements at the site, a radio survey can be used to design still non-deployed Wi-Fi networks. This type of survey is called “predictive” or “virtual”, since Wi-Fi characteristics are calculated for a virtual model of the environment created by the user. The process of creating and configuring a virtual environment, choosing the location of virtual access points (APs) and analyzing a simulated Wi-Fi network is usually called “RF-planning” [2].

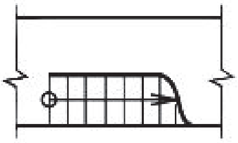
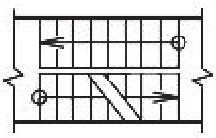
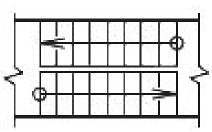
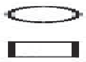
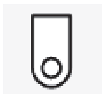
Prior to the beginning of the radio survey, a plan of the robotarium, a department, was received, on which a Wi-Fi network is expected. A study was conducted of the premises, the accuracy of the plan, the thickness and height of the walls were determined, as well as columns and fire doors were marked. The plan (floor) is a schematic of the object, presented as a drawing with special designations (for example, a “BTI” plan). The designations used when creating floor plans of an object are called graphical symbol (GS) and are governed by various State Standarts:

– GOST 21.201-2011 Project documentation system for construction. Conventional graphic images of elements of buildings, structures and structures.

– GOST 21.205-93 SPDS. Symbols of elements of sanitary and technical systems.

Based on the information obtained during the analysis of these standards, a table of symbols used in terms of premises was created (Table 1).

Table 1. GOSTs, used in terms of premises

Name of the object	**GS facility on the plan**	**Standard**	**Note**
Ladder (lower march)		GOST 21.201-2011, p. 4.6	On the plans of the stairs the arrow indicates the direction of raising the march.
Ladder (intermediate marches)		GOST 21.201-2011, p. 4.6	On the plans of the stairs the arrow indicates the direction of raising the march.
Ladder/stairway (Upper March)		GOST 21.201-2011, p. 4.6	On the plans of the stairs the arrow indicates the direction of raising the march.
Column (support)		GOST 21.201-2011, p. 4.2	
Sink		GOST 21.205-93, tab. 2	
Toilet		GOST 21.205-93, tab. 2	
Window opening		GOST 21.201-2011, Clause4.4	
Room number and its area	$\frac{3}{11,5}$	Examples of plans	

When designing a seamless wireless network, it is worth considering one of the most important conditions affecting the quality of the built network infrastructure - determining the degree of attenuation of the wireless signal. In order to take into account this important factor, a model of radio propagation in rooms is being developed. When building this model, the following factors should be considered:

– Signal interference with various electrical devices has a significant effect on signal propagation.

– Attenuation of a radio signal or its complete loss may occur as a result of passing through various obstacles or reflections from them.

Obstacles include walls, beams, furniture, different types of doors, columns. Such objects, being on the path of radio signal propagation, partially / significantly (depending on the obstacle) absorb or reflect it, which leads to a deterioration in the signal quality. Each obstacle in the zone of propagation of a signal reduces its power. The more obstacles are there, the worse the signal becomes. It is also worth noting, that the Wi-Fi signal not only tries to bend around an obstacle, but also passes through it, which leads to additional reflection and absorption of a part of the original signal. Quality affects not only the number of walls located on the signal propagation path, but also their thickness. There are materials with a different signal absorption coefficient. For example, wood, plastic, ordinary glass, drywall refer to materials with low absorption. Tinted glass, water (large aquarium), bricks, plaster are materials with medium absorption. Materials with a high absorption coefficient, which have a strong negative effect on the signal, include metal (iron doors, aluminum and steel beams), concrete (inside which there is a reinforcing lattice), and ceramics. Indoors, mirrors can also cause radio signal interference (strongly reflect the signal) and tinted windows. The table below shows the loss of Wi-Fi signal efficiency, when passing through various environment. The values (not absolute, but approximate) are given for a wireless network operating in the 2.4 GHz frequency band (Table 2).

Table 2. Wi-Fi signal efficiency losses

Obstacle	**Additional loss, dB**	**Effective distance**
Open space	0	100%
Window without tinting (no metallized coating)	3	70%
Window tinted (metallized coating)	5-8	50%
Wooden wall	10	30%
Interroom wall (15.2 cm)	15-20	15%
Bearing wall (30.5 cm)	20-25	10%
Concrete floor / ceiling	15-25	10-15%
Monolithic reinforced concrete floor	20-25	10%

Effective distance means how much the range of a Wi-Fi signal will decrease after passing the corresponding obstacle compared to open space [3]. To obtain a more extensive and accurate model of radio signal propagation, four main types of walls were identified that were used when working with the floor plan:

1. Type 1 - thin partitions up to 10 cm thick (this type also includes windows inside the cabinets).
2. Type 2 - interior walls with a thickness of 10 cm to 20 cm, consisting of drywall.
3. Type 3 - interior walls with a thickness of 10 cm to 20 cm, consisting of concrete, sheathed with plasterboard.
4. Type 4 - concrete and brick walls with a thickness of 20 cm.

Considering these four views, the corresponding marks were made on the plans, showing where the type of walls is and their width. Formulated network requirements include:

– Number of users per point at least 10

– Acceptable signal level, which is enough for both remote control of robots and Internet surfing (up to -60 dBm)

– Seamless roaming

In order to check the quality of the coating, a passive survey was conducted with test access points. Measurements were made of the received signal level in the study zone in the 2.4 GHz band. The access points (AP) D-Link DWL-3200AP available at the customer participated in the modeling of the coverage. Based on the fact that there were 4 access points, their optimal location was determined in terms of covering the most important signal propagation zones (Fig. 1.). Also the existing wired infrastructure was taken into account to minimize the cost of installation work. To reduce the level of interference, it was decided to use non-overlapping channels (1, 6, 11) at 2.4 GHz.

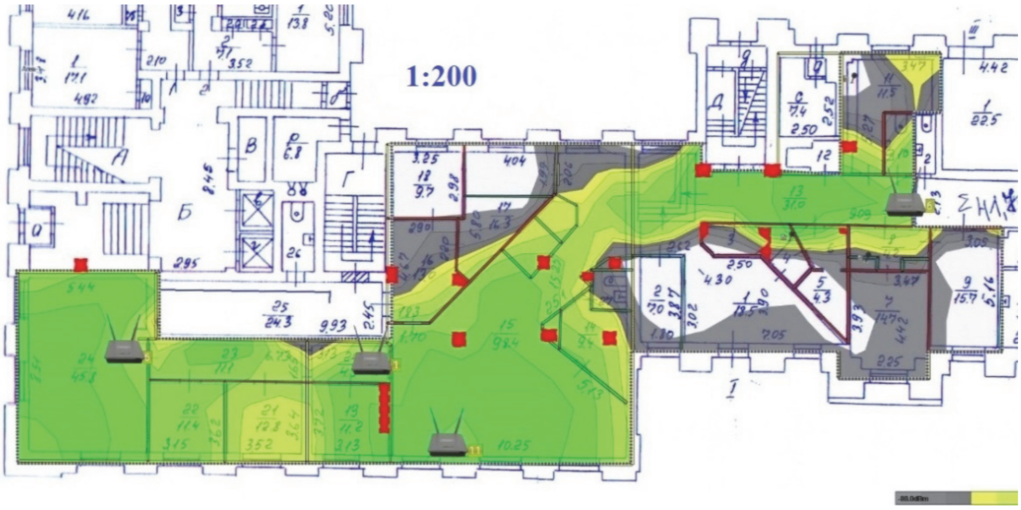


Figure 1 – The level of signal propagation using four APs

Further analysis of the telecommunication equipment market was carried out in order to select the AP models most suitable for the tasks assigned and the following APs were considered: Ubiquiti UniFi AP (UAP), Ubiquiti UniFi AP LR (UAP-AC-LR), Ubiquiti UniFi AP Pro (UAP-AC -PRO), Ubiquiti UniFi AP AC Lite (UAP-AC-LITE). Table 3 shows the price range for these AP models, and Table 4 compares the characteristics of these APs.

Table 3. Estimated cost of equipment on the market

AP Model	**Cost of 1 pcs., RUB**	**The cost of a set of 3 pieces., RUB**	**The cost of a set of 5 pcs., RUB**
Ubiquiti UniFi AP	4520	11800 (megabitcomp.ru), 13400 (Beru.ru)	-
Ubiquiti UniFi AP LR	7200	17379 (wifimag.ru)	32210 (its-wifi.ru)
Ubiquiti UniFi AP PRO	9830	28050 (its-wifi.ru), 29350 (wifimag.ru)	46300 (its-wifi.ru), 47716 (wifimag.ru)
Ubiquiti UniFi AP AC Lite	5800	-	25650 (its-wifi.ru), 26708 (wifimag.ru)

Table 4. Equipment specifications

Specifications	**UAP**	**UAP-AC-LITE**	**UAP-AC-PRO**	**UAP-AC-LR**
Wi-Fi standard	802.11n (802.11 b / g / n)	802.11a / b / g / n / ac	802.11a / b / g / n / ac	802.11a / b / g / n / ac
Frequency range Wi-Fi devices	2.4 GHz	2.4 / 5 GHz	2.4 / 5 GHz	2.4 / 5 GHz
Max. Speed wireless connections	300 Mbit / s	1167 Mbit / s	1750 Mbit / s	1300 Mbit / s
Simultaneous work in two ranges	none	available	available	available
Radius actions outpremises	no data	122 m	no data	183 m
Power transmitter	20 dBm	20 dBm	22 dBm	24 dBm (22 dBm at 5GHz)
Speed ports	100 Mbit / s	1000 Mb / s	1000 Mb / s	1000 Mb / s
amount Ethernet ports	1	1	2	1
amount internalantennas	2	2	3	1

Specifications	UAP	UAP-AC-LITE	UAP-AC-PRO	UAP-AC-LR
Coefficient gaininternal antennas	8 dBi	3 dBi	3 dBi	6 dBi
MIMO	1x1	2x2	3x3	3x3 (2x2 for 5 GHz)
IEEE 802.1Q support(VLAN)	available	available	available	available
Security	WEP, WPA-PSK, WPA-TKIP, WPA2 AES, 802.11i	WEP, WPA-PSK, WPA-Enterprise (WPA / WPA2, TKIP / AES)	WEP, WPA-PSK, WPA-Enterprise (WPA / WPA2, TKIP / AES)	WEP, WPA-PSK, WPA-Enterprise (WPA / WPA2, TKIP / AES)
Nutrition	Passive Power over Ethernet 12-24 V, 24V 0.5A	802.3af / A PoE 24V Passive PoE (Pairs 4, 5+; 7, 8 Return)	Passive Power over Ethernet (48V), 802.3af / 802.3at Supported (Supported Voltage Range: 44 to 57VDC)	802.3af / A PoE 24V Passive PoE (Pairs 4, 5+; 7, 8 Return)
Number of simultaneous customers	Up to 100	More than 250	More than 250	More than 250
Width	200 mm	160 mm	176 mm	176 mm
Height	200 mm	32 mm	43 mm	43 mm
Depth	36 mm	160 mm	176 mm	176 mm
Weight	290 g	170g	350 g	240 g

Let us consider the signal level by using various options. Hereinafter, the name of the pictures in brackets will indicate the signal level. For example, “AC LITE x 5 (-60 dBm)” means that the signal level of -60 dBm ... -50 dBm is highlighted in gray, the signal of the best level is colored in other color. Analyzing the data obtained, the customer was given the following recommendations on the choice of equipment:

– It does not make sense to consider UAP AP, since it does not know how to work in the 5 GHz band, and in the future, if there is a need to connect, besides robots, also visitors, it is better to “unload” the 2.4 GHz band for smooth work with robots. On the other hand, for 20 tr. one can take 5 of these access points.

– UAP AC LR and UAP AC PRO show almost identical results when using three points, and when using five. Therefore, it makes no sense to overpay for the PRO version.

– When choosing between UAP AC Lite and UAP AC LR, you should rely on the budget and the number of points.

In this work, a survey was conducted of a rather large room of a robotarium at the Minot of the RSUH with the subsequent installation of access points and a passive radio survey that allowed testing the simulated coverage. Thus, confirmation was obtained of the effectiveness of the proposed method of building room models for designing sustainable supervisory control of a group of mobile robots for various purposes - service, for telemedicine, for surveying premises in extreme situations, including network deployment using the repeater robots themselves. Studies are also based on other works of our team - see [4-7].

References

1. "Zyxel. Access points ”[Electronic resource] - Access mode: https://kb.zyxel.ru/hc/ru/articles/115002584514- Description of the modes of operation- wireless- point-access- Wi Fi, free. Reference date: 07/20/2018
2. "TP-Link. Building a WLAN ”[Electronic resource] - Access Mode: https://itnan.ru/post.php?c=1&p=335674, free. Reference date: 07/20/2018
3. "Zyxel. Wi-Fi ”[Electronic resource] - Access mode: https://kb.zyxel.ru/hc/ru/articles/115002572093-Ratio-attenuation ratios- Signal- Wi-Fi- through passage- through- various- environments, free. Reference date: 07/25/2018

4. Pryanichnikov V.E., Aryskin A.A., Eprikov S.R., Kirsanov K.B., Khelemendik R.V., Ksenzenko A.Ya., Prysev E.A., Travushkin A.S. (2017). Processing of Consumables, Proceedings of the 28th DAAAM International Symposium, pp. 1202-1207, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-11 -2, ISSN 1726-9679, Vienna, Austria. DOI: 10.2507 / 28th.daaam.proceedings.167
5. Ksenzenko A.Ya., Marzanov Yu.S., Prysev E.A., Pryanichnikov V.E., Chernyshev V.V. Prototyping of contactless data exchange and energy supply of a group of underwater satellite robots with a base station walking along the bottom. Extreme Robotics. // Collection of theses of the International Scientific and Technical Conference Extreme robotics. - 2017.- 272 p., P. 268-269. ISBN 978-5-85875-522-7. URL: http://er.rtc.ru/images/docs/Sbornik_tezisov_ER_2017.pdf
6. Pryanichnikov V.E., Ksenzenko A.Ya., Kuvshinov S.V., Poduraev Yu. V., Prysev E.A., Khelemendik R.V., Eprikov S. (2016). Intelligent robotics, Proceedings of the 27th DAAAM International Symposium, pp.0225-0229, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-08-2, ISSN 1726 -9679, Vienna, Austria DOI: 10.2507 / 27th.daaam.proceedings.033
7. Bogdanovich A.V., Kirsanov K.B., Pryanichnikov V.E., Helemendik R.V. Software and hardware components of intelligent service mobile robots / Information-measuring and control systems (Issue. Intelligent adaptive robots, V. 14, No. 1-2, 2019), M .: Radio Engineering, 2018, V. 16, No. 12 ,. p.33-39.

L.D. Smirnaya[1,2], E.S. Briskin[1,2]

THE INTERACTION OF THE FOOT WALKING PROPULSION OF MOBILE UNDERWATER ROBOT WITH THE BOTTOM SOIL

[1]*Volgograd State Technical University, Volgograd, Russia*
[2]*Center for Technology Components of Robotics and Mechatronics, Innopolis University, Innopolis, Russia*
dtm@vstu.ru

Abstract

The control of normal reactions of walking propulsion devices underwater robots in the implementation of the traction force and the separation from the ground in the paper considered.

Keywords: propulsion device, underwater walking robot, pulling force, the force of resistance to motion.

Acknowledgments

The work was supported by the Russian Science Foundation (grant No. 18-71-10069).

Introduction

One of the main tasks in the development of underwater mobile robots to move around the bottom surface is to ensure proper profile and reference maneuverability. Its solution may be based on the use of walking propulsion devices [1]. Improving the reference maneuverability is an urgent task for both ground vehicles and mobile underwater robots. The main factor affecting the supporting permeability is the magnitude of the normal reaction exerted by the propulsion device on the ground [2, 3, 4]. For underwater robots, this effect is more significant due to the “compression effect” [5]. A characteristic feature of walking thrusters is that with an increase in normal reactions linear traction properties increase linearly, but at the same time proportional to the square of the normal reaction increases the resistance forces to movement [6, 7]. Another feature is that the normal reactions depend on the movement of the body relative to the fixed supports propulsion device. However, these reactions can be controlled both by deliberately changing the stiffness of the supports, and by changing the buoyancy for pontoon-type robotic systems [8, 9] and by vertical micro-movement of the foot of the walking mechanism, for robots with orthogonal-type propulsion devices (Fig. 1) [10]. For mobile robots moving along the bottom of reservoirs, due to a certain specificity of movement, manifested in such well-known phenomena as: arising pushing force, possible presence of flow, difficult separation of the propulsion device foot from the bottom surface caused by the “compression” effect [11, 12, 13] optimal control problems were not set. So for the implementation of the transfer of the propulsion device to a new position, it is required to detach the foot from the ground, and this, due to the “compression” effect, requires much more effort. This is confirmed experimentally, but both the mathematical model of the process and the regularities established in accordance with it, which allow accounting and purposefully controlling the lifting of the foot in order to ensure the permissible drive forces, are absent.

Figure 1 – Walking robot «Ortonog»

Statement of problem

1. An absolutely rigid body, supported on N supports, as propulsion devices, periodically interacting with a supporting surface (Fig. 2) is taken as the design scheme of an underwater mobile robot. The center of mass of the robot and the center of pressure do not coincide and are at a distance b from each other. Each of the supports is located at a distance a_j from the center of mass, and in the process of movement, these coordinates change by the amount of horizontal movement of the robot S before the transfer of one of the thrusters to a new position.

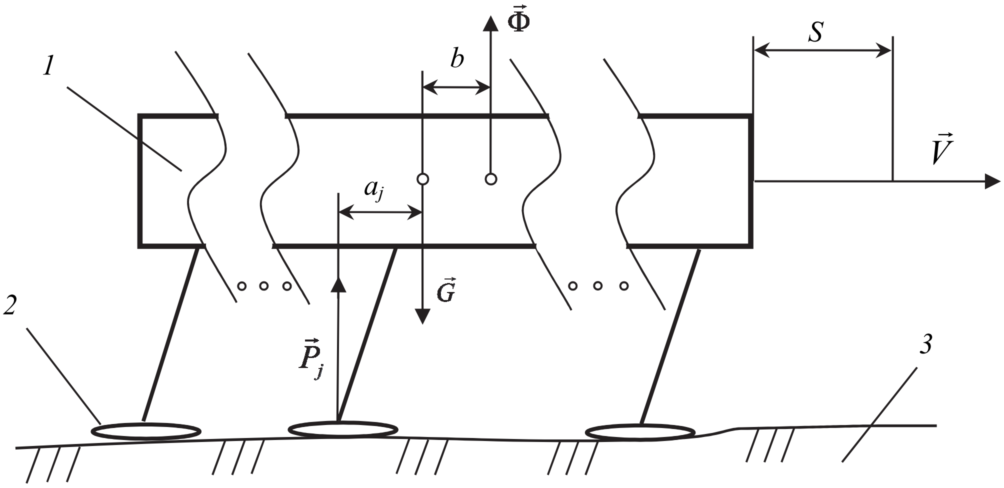


Figure 2 – The design scheme of the robot: *1* – the body of the robot; *2* – propulsion foot; *3* – ground

The control efficiency is evaluated by criterion I, consisting of two indicators: the amount of traction and coupling properties I_1 and the force of resistance to movement I_2

$$I_1 = f\left(G - \Phi\right),$$
$$I_2 = -\frac{1}{\lambda}\sum_{j=1}^{N} P_j^2; \quad \lambda = 2c_n l, \tag{1}$$

where G is the weight of the robot; f – coefficient of adhesion; c_n is the stiffness of the support-soil system; l is the step length.

Moreover, the relationship between them k_1, k_2 is determined by the developer in each specific situation with manual control and control system for robots capable of self-controlling [14].

The controlled parameters are the normal reactions P_j. This is possible, since when studying the translational motion of the robot body there are only two equations for determining the normal reactions.

$$\sum_{j=1}^{N} P_j - G + \Phi = 0,$$
$$-\sum_{j=1}^{N} P_j\left(a_j + S\right) + \Phi b = 0. \tag{2}$$

Therefore, if there are more than two support reactions, then they can be controlled. The simplest relations for their determination together with equations (2) are

$$\sum_{j=1}^{N} P_j \gamma_{jn} = D_n \quad \left(n = 1, \ldots N-2\right), \tag{3}$$

where γ_{jn}, D_n – parameters to be determined.

The minimum number of thrusters, ensuring steady movement of the considered mobile robot, allowing the transfer of thrusters and reaction management is three.

Therefore, to test the control method, a robot based on three movers is considered. In this case, instead of (3) one equation holds, for example

$$P_3=\gamma P_2, \quad D=0. \tag{4}$$

The physical meaning of the coefficients introduced is the control parameters.

The task is to determine the control parameter γ of the maximum performance criterion I

$$I=k_1 I_1+k_2 I_2. \tag{5}$$

2. To solve the problem of foot separation, a horizontal plate of mass m and area S is progressively moving along the vertical axis in a liquid medium with a known permeability coefficient k and dynamic viscosity μ [15] as a design scheme for the foot of a walking propulsion device interacting with the ground of the reservoir. At the stage of the traction mode, the plate (Fig. 3) is at rest, and at the stage of lifting it is affected by forces: mg is the weight of the plate, T is the force developed by the propulsion foot lifting drive, p_0S, pS are the upper and lower pressure forces respectively plate surface. In the equilibrium position $p = p_*$

$$p_* S-p_0 S=F, \tag{6}$$

where F is the buoyancy force.

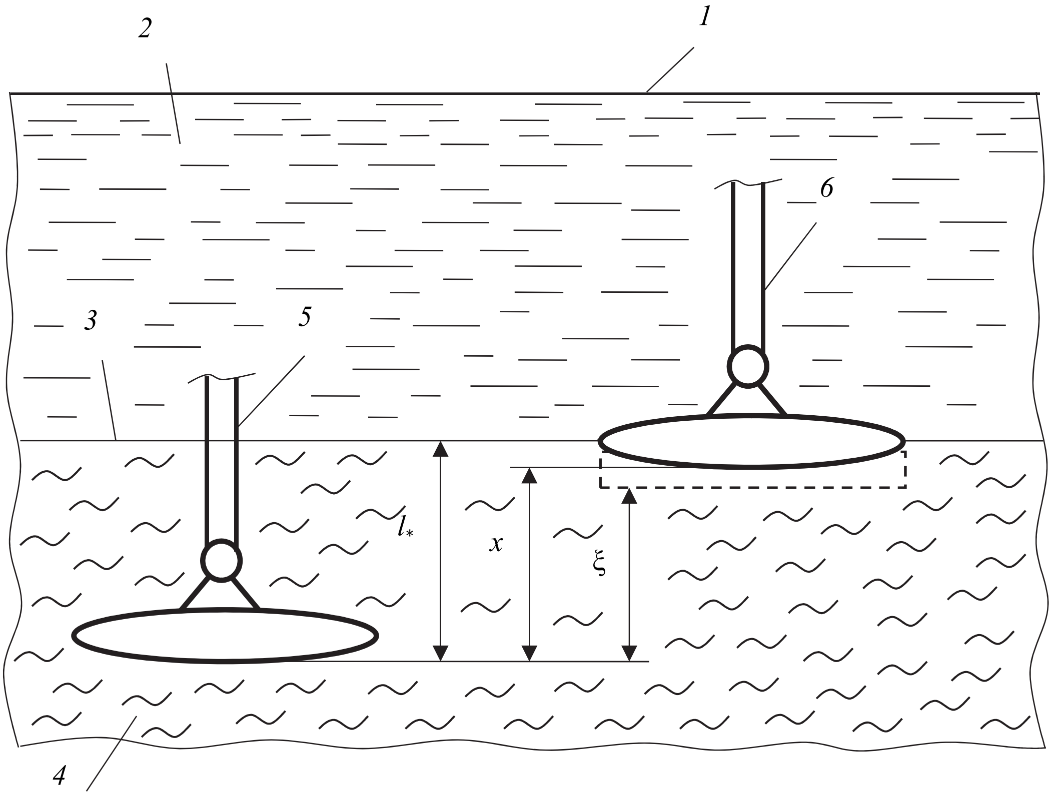


Figure 3 – The design scheme of the engine immersed in the ground:
1 - the surface of the reservoir, *2* - the reservoir, *3* - the boundary of the reservoir and bottom soil,
4 - bottom soil, *5* - propulsion support, *6* - propulsion stop

When the plate moves submerged upward, a vacuum forms under it, due to the expansion of the volume of dissolved air in the liquid in accordance with Henry's law [16]. In the design scheme, this is taken into account by the difference of coordinates between the bottom surface of the plate x and the surface of the liquid ξ

$$\Delta = x - \xi \tag{7}$$

In the equilibrium position of the foot at $t = 0$

$$\xi_0=-\Delta_0, \quad x_0=0, \quad \dot{x}_0=0, \quad \dot{\xi}_0=0 \tag{8}$$

The rise of the foot is carried out with an initial zero speed under the action of a force T developed by a lifting drive. When the foot reaches the height $x = l_*$, the foot breaks away from the ground and the force caused by "compression" under its lower base ceases to act on it.

In this case, one of the tasks is to determine the drive force for a given law of change in the speed of the foot.

Solution method

1. The solution of the first task is

$$\begin{aligned} P_1 &= \mu_{11} G + \mu_{12} \Phi, \\ P_2 &= \mu_{21} G + \mu_{22} \Phi, \\ P_3 &= \mu_{31} G + \mu_{32} \Phi, \end{aligned} \tag{9}$$

where

$$\begin{aligned} &\mu_{11} = \frac{a_2 + \alpha a_3}{(a_2 - a_1) + \gamma(a_3 - a_1)}, \qquad \mu_{12} = \frac{b(1+\gamma) - (a_2 - \gamma a_3)}{(a_2 - a_1) + \gamma(a_3 - a_1)}, \\ &\mu_{21} = -\frac{a_1}{(a_2 - a_1) + \gamma(a_3 - a_1)}, \quad \mu_{22} = \frac{a_1 - b}{(a_2 - a_1) + \gamma(a_3 - a_1)}, \\ &\mu_{31} = \gamma \mu_{21}, \qquad \mu_{32} = \gamma \mu_{22}, \\ &a_1 = a_{10} + V_0 t, \quad a_2 = a_{20} + V_0 t, \quad a_3 = a_{30} + V_0 t, \end{aligned} \tag{10}$$

where V_0 is the speed of the robot, a_{n0} is the initial coordinates

The indicator I_1 does not depend on the support reactions P_j, and the indicator I_2 has the form

$$I_2 = \frac{1}{\lambda}\left[\left(\mu_{11}^2 + \mu_{21}^2 + \mu_{31}^2\right) G^2 + 2G\Phi\left(\mu_{11}\mu_{12} + \mu_{21}\mu_{22} + \mu_{31}\mu_{32}\right) + \left(\mu_{12}^2 + \mu_{22}^2 + \mu_{32}^2\right)\Phi^2\right]. \tag{11}$$

The graph (Fig. 4) shows the dependence of the indicator I_2 on time with the initial coordinates a_{10}= -4, a_{30}= 4, with different coordinates a_{20}.

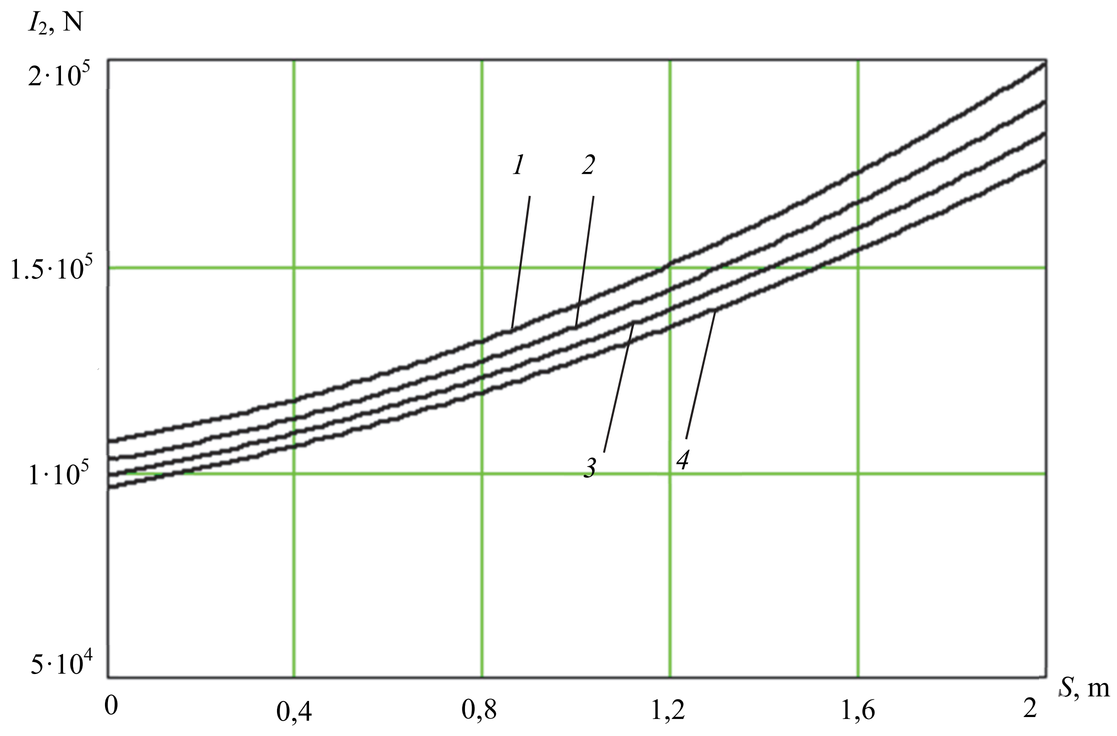


Figure 4 - Dependence of the indicator I_2 on the movement of the robot S with a different location of one of the propulsion devices a_{20} at $\gamma = 0,5$:

1 – a_{20} = -0,5; *2* – a_{20} = 0; *3* – a_{20} = 0,5; *4* – a_{20} = 1

The results of the calculations show that the location of the propulsion devices in particular of one central one influences the force of resistance to movement, and in the process of moving the force of resistance changes, in particular, in the considered model problem by 100%.

2. The solution of the second task is based on Henry's law on the concentration of gas in a liquid medium, the law of its expansion for the polytropic process when lifting the foot, the Darcy law of fluid filtration and the theorem on the movement of the center of mass of the foot as a solid, when the foot begins to separate from the ground.

As a result, the equation

$$\frac{m\Delta_0 S}{\alpha\left(p_0 S+F\right)}\ddot{V}+m\dot{V}+\frac{S}{\alpha}V+mg-F=T \tag{12}$$

where m, S are, respectively, the foot mass of the walking mover and the area of its interaction with the ground, $\alpha = k/\mu l_{*n}$ is the proportionality coefficient between the fluid pressure in different horizontal sections of the soil and the filtration rate V, k is the soil permeability coefficient, μ is the dynamic viscosity liquids, l_*– conditional height of the bottom soil column through which filtration occurs.

The solution of equation (12) can be obtained if set the speed of movement of the foot

$$\dot{x}=u\left(1-e^{-\delta t}\right) \tag{13}$$

Then

$$x=ut-\frac{u}{\delta}\left(1-e^{-\delta t}\right). \tag{14}$$

Now, it is possible to determine both the filtration rate of the fluid and its derivatives.

$$\begin{aligned}
\dot{\xi}&=V=u\left[-\frac{\lambda}{\lambda-\delta}e^{-\delta t}+\frac{\lambda}{\lambda-\delta}e^{-\lambda t}\right]+u,\\
\ddot{\xi}&=\dot{V}=u\left[\frac{\lambda\delta}{\lambda-\delta}e^{-\delta t}-\frac{\lambda\delta}{\lambda-\delta}e^{-\lambda t}\right],\\
\dddot{\xi}&=\ddot{V}=u\left[-\frac{\lambda\delta^2}{\lambda-\delta}e^{-\delta t}+\frac{\delta\lambda^2}{\lambda-\delta}e^{-\lambda t}\right].
\end{aligned} \tag{15}$$

The final determination of the force $T(t)$ is carried out using the equation

$$T\left(t\right)=mg-F+\frac{m}{\lambda}\ddot{V}+m\dot{V}+\frac{S}{\alpha}V. \tag{16}$$

It is considered raising the foot to a height of 0,2 *m* before leaving the ground

Thus, if the law of variation of the foot speed is specified, the force T increases smoothly and the more, the smaller the permeability coefficient k and the greater the depth of the foot immersed in the ground l_*. When the foot reaches the position corresponding to its exit from the ground, the force is reduced to the magnitude of the corresponding force without taking into account the “compression” effect. In the calculations of the model problem, it was assumed that the depth of the foot immersion in the bottom soil is l_* = 0.2 m.

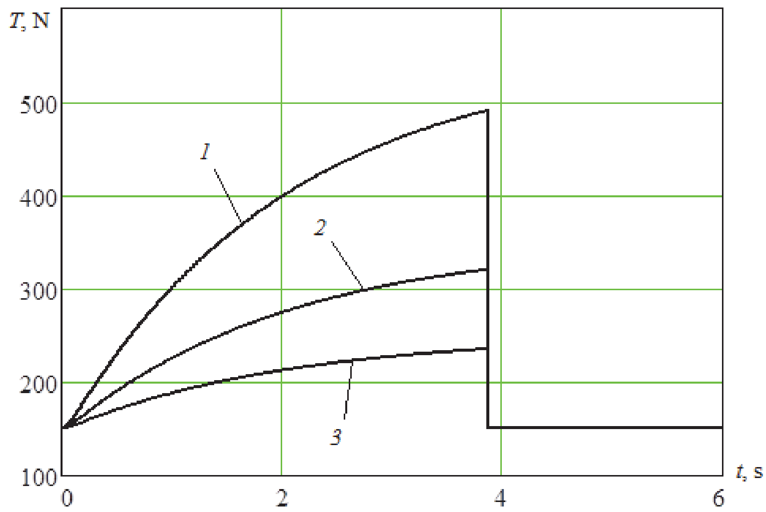


Figure 5 – Dependence of the force T developed by the drive lifting the foot on time:
1 – $\alpha = 2{,}5\cdot 10^{-8}$ m^2/ kg·s; *2* – $\alpha = 5\cdot 10^{-8}$ m^2/ kg·s; *3* – $\alpha = 10\cdot 10^{-8}$ m^2/ kg·s

The dependences of the displacement $x(t)$ and velocity $V(t)$ are presented in the graphs (Fig. 6, 7)

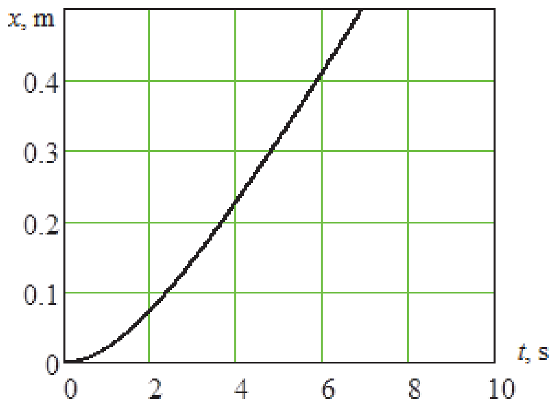


Figure 6 – Dependence of the foot displacement x on time

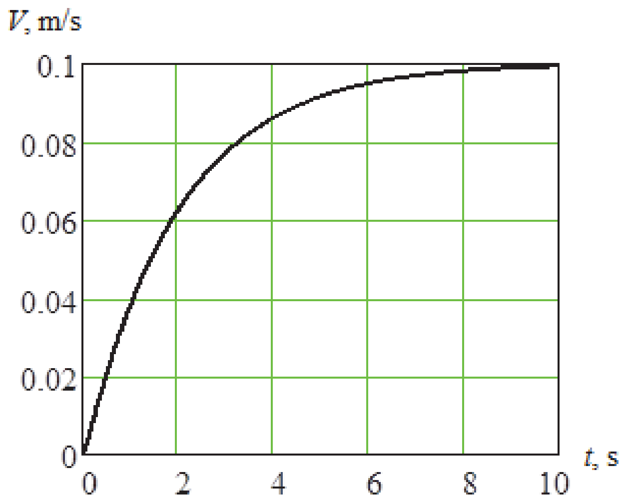


Figure 7 – Dependence of the foot velocity V on time

References

1. Chernyshev V.V., Arykancev V.V., Gavrilov A.E. Upravlenie dvizheniem podvodnyh shagayushchih apparatov peredvigayushchihsya po dnu // Izvestiya YUFU. Tekhnicheskie nauki. 2016. № 1 (174). S. 141-155.

2. Traktory: Teoriya: / V. V. Gus'kov [i dr.]; pod obshch. red. V. V. Gus'kova. – M.: Mashinostroenie, 1988. – 375 s.
3. Smirnov G.A. Teoriya dvizheniya kolesnyh mashin / G. A. Smirnov. – 2-e izd., dop. i pererab. – M.: Mashinostroenie, 1990. – 352 s.
4. Briskin E.S., Chernyshev V.V., Zhoga V.V., Maloletov A.V., Sharonov N.G., Frolova N.E. Koncepciya proektirovaniya, dinamika i upravlenie dvizheniem shagayushchih mashin. CH. 3. Algoritmy upravleniya dvizheniem shagayushchih mashin serii "Vos'minog" i eksperimental'nye issledovaniya // Mekhatronika, avtomatizaciya, upravlenie. – 2005. – № 7. – S. 13–18.
5. Arykancev V.V., Chernyshev V.V. Issledovanie "kompressionnogo effekta", voznikayushchego pri smene stop shagayushchego dvizhitelya na podvodnyh gruntah // V sbornike: Sovremennye metody i sredstva okeanologicheskih issledovanij Materialy XV Vserossijskoj nauchno-tekhnicheskoj konferencii (MSOI-2017). Federal'noe agentstvo nauchnyh organizacij, Institut okeanologii im. P.P. SHirshova RAN, Rossijskij fond fundamental'nyh issledovanij i dr. 2017. S. 178-182.
6. Algoritmy upravleniya robotami i manipulyatorami / Ignat'ev M. B., Kulakov F. N., Pokrovskij A. M. i dr. – L.: Mashinostroenie. [Leningr. otd-nie], 1972. – 247 s.
7. Briskin E.S., Sobolev V.M. Tyagovaya dinamika shagayushchih mashin s ortogonal'nymi dvizhitelyami // Problemy mashinostroeniya. № 3. – M., 1990. – S.28–34.
8. Shnejder A.Yu. Upravlenie opornymi reakciyami shagayushchego apparata pri dvizhenii po gruntam s razlichnymi nesushchimi svojstvami / A. Yu. Shnejder, D. M. Gorinevskij // Preprint instituta problem peredachi informacii AN SSSR. – 1986.– 72 s.
9. Briskin E.S., Sharonov N.G., Serov V.A., Pen'shin I.S. Upravlenie dvizheniem podvodnogo mobil'nogo robota s yakorno-trosovymi dvizhitelyami // Robototekhnika i tekhnicheskaya kibernetika. 2018. № 2 (19). – S. 39–45.
10. Briskin E.S., Kalinin YA.V., Maloletov A.V., Serov V.A., Ustinov S.A. Ob upravlenii adaptaciej ortogonal'nyh shagayushchih dvizhitelej k opornoj poverhnosti // Izvestiya Rossijskoj akademii nauk. Teoriya i sistemy upravleniya. 2017. № 3. S. 184-190.
11. Arykancev V.V., Goncharov A.A., Chernyshev V.V. Modelirovanie kontaktnogo vzaimodejstviya opornyh elementov (stop) shagayushchego dvizhitelya s gruntom v usloviyah slozhnogo nagruzheniya // Ekstremal'naya robototekhnika. 2018. T. 1. № 1. S. 258-265.
12. Arykancev V.V., Chernyshev V.V. Model'nye ocenki vliyaniya "kompressionnoj" sily na dinamiku glubokovodnogo shagayushchego apparata // V sbornike: Nelinejnaya dinamika mashin - School-NDM 2017 sbornik IV Mezhdunarodnoj SHkoly-konferencii molodyh uchenyh. 2017. S. 113-119.
13. Ksenzenko A.Y., Prysev E.A., Pryanichnikov V.E., Chernyshev V.V. Design the contactless charger and contactless data transfer between underwater robotsatellits and underwater 6-legged vehicle // V sbornike: Annals of DAAAM and Proceedings of the International DAAAM Symposium 28. Ser. "Proceedings of the 28th International DAAAM Symposium "Intelligent Manufacturing and Automation", DAAAM 2017" 2017. S. 1197-1201.
14. Briskin E. S. Walking robot «character» as element of intelligent system / E. S. Briskin, A. V. Maloletov, N. G. Sharonov, Ya. V. Kalinin, A. V. Leonard, V. A. Serov, V. A. Shurygin // V sbornike: Advances in Cooperative Robotics: Proceedings of the 19th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, (CLAWAR 2016 19th.) 2016. – S. 386–394.
15. Zhuzhikov V.A. Fil'trovanie. Teoriya i praktika razdeleniya suspenzij. M., Izd-vo – Himiya. 1971. 440 s.
16. Ramm V. M. Adsorbciya gazov. Izd. 2-e, pererab. i dop. M., Izd-vo – Himiya. 1976. 656 s.

A.V. Klekovkin[1], Yu.L. Karavaev[1], A.A. Kilin[2], I.S. Mamaev[1]

CONTROL SCREWLESS FISH-LIKE ROBOT WITH INTERNAL ROTOR

[1] *Kalashnikov Izhevsk State Technical University, Izhevsk, Russia, klanvlad@mail.ru*
[2] *Udmurt State University, Izhevsk, Russia*

Abstract

The article presents the screwless overwater fish-like mobile robot, driven by an internal rotating rotor. The robot has a rigid case which doesn't be transform while moving. Mathematical model of robot movement was developed; differential equations were written. Few series of experimental researches with different control actions were carried out. Results of experimental researches were compared with modelling results.

Keywords: mobile robot, swimming robot, screwless overwater robot, fish-like robot.

Acknowledgments

The work was carried out within the framework of the state assignment 1.2405.2017/PCh and was supported by the RFBR grant No 18-08-00995-a.

Introduction

Many floating robotic vehicles move by rotating of propeller screws. Also mechanisms which copy organisms moving are popular. There are methods of moving in water by jet reaction drive, moving by transforming of body shape, moving by action of internal mechanisms.

Using moving by action of internal mechanisms all driving elements are in body and don't associate with fluid. As a result, a construction of these robots are simple because contact movable elements with water is missing. First theoretical researches were presented at the beginning 2000 years [1, 2]. The practical implementation of such a movement has become more complicated due to the difficulty of implementing the task of moving internal masses along an arbitrary trajectory. For this reason, for the movement due to the action of internal mechanisms, technical systems with a simpler implementation are used: the movement of masses along rectilinear guides [3], the change in the center of mass position due to rotating eccentrics [4], the change in the gyrostatic moment due to the rotation of the internal rotors [5], etc.

This paper presents a screwless overwater fish-like robot which has the shape of a wing profile NACA 0040 [6]. The robot is moved by the rotation of the internal rotor. The robot design combines two principles of movement of floating robotic devices: a form close to the shape of fish, and movement due to the action of internal mechanisms. Most of the movement of existing fish-like robots is carried out by the movement of the fins, that is, changes in the shape of the body [7, 8]. In this paper, we consider a robot that retains the principles of motion, but does not change its shape.

Mathematical model

Consider the motion of a fish-like robot on a plane. To describe the motion, we introduce two coordinate systems: a fixed coordinate system Oxy and a moving coordinate system Cx_1x_2, attached to the body, located on the axis of symmetry of the body. Axis Cx_1 is directed from tail to nose of the robot (see Fig. 1).

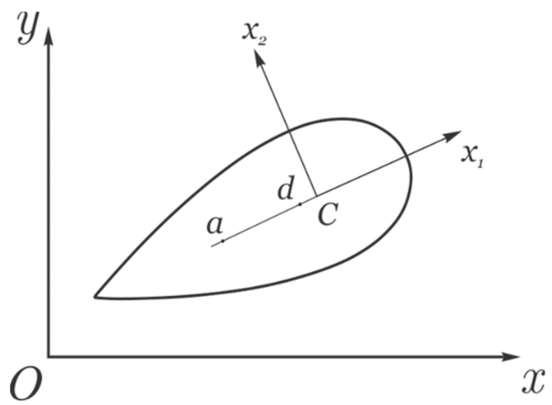


Figure 1 – The robot geometry. Point *a* is a center of mass; point *d* is a rotor position

The kinetic energy of the entire system can be represented as:

$$T = T_f + T_b + T_r$$

where T_f – the kinetic energy of the fluid, T_b – the kinetic energy of the body, T_r – the kinetic energy of the rotor.

Let's describe each of the energies:

$$T_f = \frac{1}{2}(\lambda_1 v_1^2 + \lambda_2 v_2^2 + \lambda_3 w^2); T_b = \frac{1}{2}m(v_1^2 + (v_2 + aw)^2) + \frac{1}{2}I_b w^2;$$

$$T_r = \frac{1}{2}m_r\left(v_1^2 + (v_2 + \mathrm{d}w^2)\right) + \frac{1}{2}I_r(w + \Omega)^2 ;$$

where λ_1, λ_2 are the added masses and λ_3 is the added moment of inertia; v_1, v_2 are the components of the velocity vector; w is the angular velocity of the body; I_b, I_r are the moment of inertia of the body and moment of inertia of the rotor respectively; a is the displacement of the center of mass relative to the moving coordinate system; d is the displacement of the rotor relative to the moving coordinate system; m, m_r are the mass of the robot and the mass of the rotor respectively.

After simplification it is possible to write expressions of the momentums and the angular momentum of the system:

$$p_1 = \frac{\partial T}{\partial v_1} = Av_1; \quad p_2 = \frac{\partial T}{\partial v_2} = Bv_2 + c_1 w; \ M = \frac{\partial T}{\partial w} = Iw + c_1 v_2 + k,$$

where $A = m + m_r + \lambda_1$; $B = m + m_r + \lambda_2$; $c_1 = ma + m_r d$; $I = I_b + I_r + \lambda_3 + ma^2 + m_r d^2$; $k = I_r\Omega$; Ω is the rotor angular velocity that is the control action.

Let us define the equations of motion taking into account the viscous resistance and circulation:

$$\dot{p}_1 = p_2 w - \Gamma v_2 - \mu_1 v_1|v_1|; \dot{p}_2 = -p_1 w + \Gamma v_1 - \mu_2 v_2|v_2|;$$

$$\dot{M} = p_1 v_2 - p_2 v_1 - \mu_3 w|w|;$$

where $\Gamma = \alpha v_2 + \beta w$ is the circulation; μ_1, μ_2, μ_3 are the coefficients of viscous resistance. More detailed derivation of the equations of motion, as well as their analysis can be found in the works [9, 10]

The parameters α and β we can determine from Kutta-Chaplygin condition. To do this, we replace our profile on the Zhukovskii foil, which are almost identical. For the Zhukovskii foil the same sizes we get $\alpha = 4.18$; $\beta = 0.698$. The coefficients μ_1, μ_2, μ_3 depend on the mode of motion and are determined experimentally. For modeling, we accept $\mu_1 = 0.5$; $\mu_2 = 40$; $\mu_3 = 5$. The added masses were calculated based on general theory [11]. $\lambda_1 = 0.0925$; $\lambda_2 = 0.4615$; $\lambda_3 = 0.0006$. For example, we take the control action in the simulation $\Omega(t) = \frac{\pi}{8}\sin(8x)$ (see Fig 2, a).

Solving the system of differential equations, we find the trajectory of the robot (see Fig. 2, b).

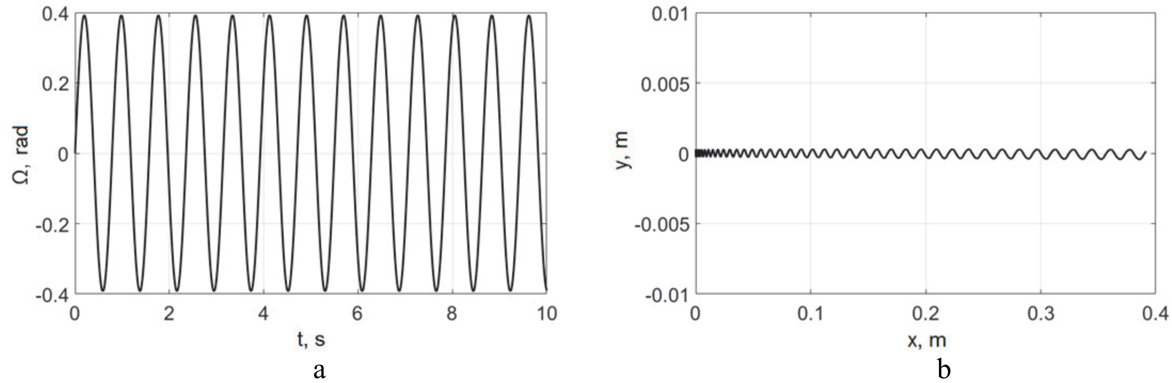


Figure 2 – The theoretical trajectory of the robot (b) with the control action $\Omega(t) = \frac{\pi}{8}sin(8x)$ (a)

Description of the robot design

The robot is a hollow object, which has the shape of a wing profile NACA 0040 (see Fig. 3). Robot length is 340 mm, width is 134 mm, height is 80 mm. The body is made on a 3D printer out of PLA plastic, the wall thickness of 2 mm. The rotor with the motor fixed inside the body so that the center of mass of the system is as close as possible to the bottom face of the robot. Pololu gear motor with encoder was used as the drive motor. Characteristics of the motor: rated supply voltage – 6 V, gear ratio – 47:1, stall torque – 0.459 Nm, maximum speed – 120 rpm. A pair of gears with a gear ratio of 3.5:1 were used to transfer the rotation from the motor to the rotor. Also there are a battery and a board with a microcontroller model STM32F303K8, controlling the rotation of the DC motor inside the robot. To control the motor, the driver of the DC motor VNH3SP30 by STMicroelectronics is used. The encoder is used to determine the position of the rotor during the experiments. To obtain the angular velocity of the rotor, we need to differentiate the data from the encoder.

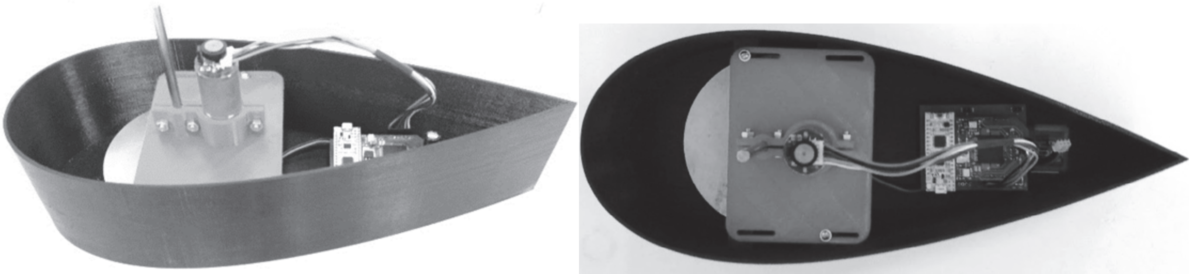


Figure 3 – The screwless fish-like robot

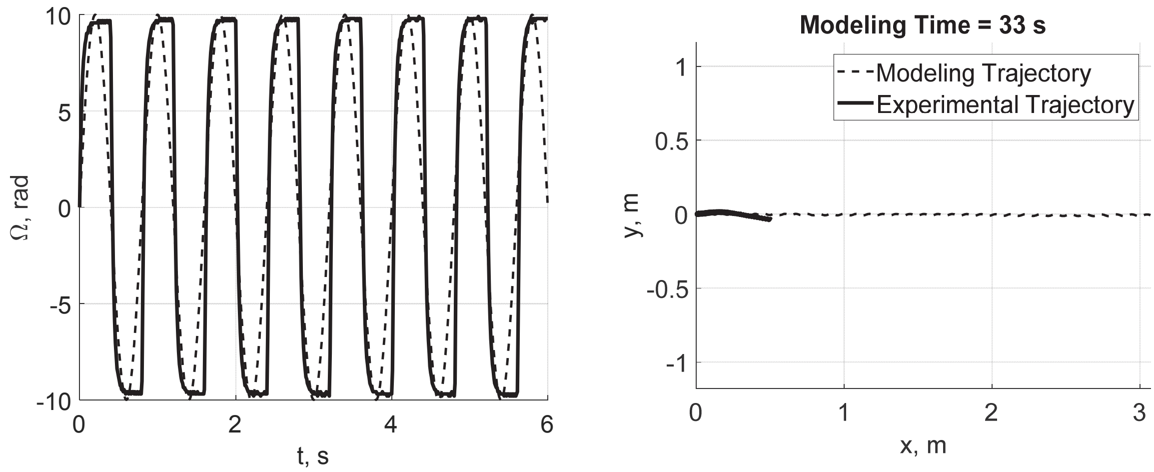
The real model of the robot has the following characteristics: m = 0,905 kg; I_0 = 0.00844 kg*m^2; the rotor is made of aluminum and has an external diameter of 110 mm, a height of 12 mm. Rotor mass m_r = 0,327 kg; rotor moment of inertia I_r = 0,00058 kg*m^2. The design of the robot allows you to shift the center of rotation of the rotor.

Experimental investigations

It was carried out four series of five experiments in a circular pool with a diameter of 0.8 meters with the same initial conditions to verify the motion theoretical model of the screwless overwater fish-like robot. The control was carried out in such a way that the rotor changed the direction of rotation at equal intervals of time, the speed and acceleration were maximum (limited by the capabilities of the DC motor). Thus, the control actions in different series of experiments differed in the frequency of changing the direction of rotation of the rotor.

The maximum value of the robot movement was 0.5 meters because the robot movement was limited by the size of the pool. The simulation time was equal to the time for which the robot with this control action floated a distance of 0.5 meters in the experiment.

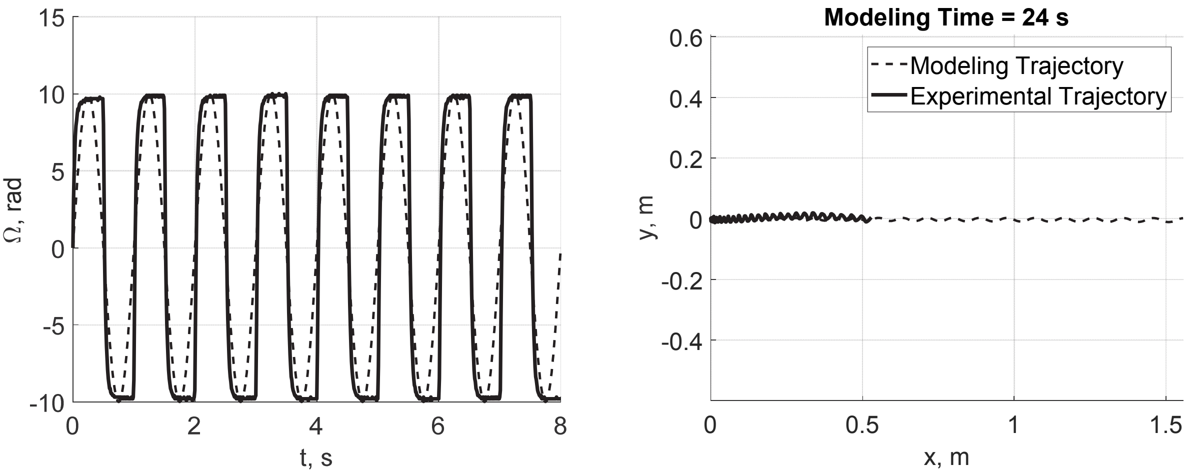
In the first series of experiments, the frequency of change in the rotor rotation direction is 1.25 Hz. Respectively, the period is 0.8 s. The time to overcome the distance of 0.5 m is 33 seconds. Figure 4, *a* shows the average dependence of the rotor speed on time in a real experiment (solid line) and the dependence of the rotor speed on time used in the simulation (dashed line). Figure 4, *b* shows the experimental trajectory (solid line) and the trajectory obtained as a result of modeling (dashed line). In this experiment, the average speed of the robot was 0.015 m/s.



a b

Figure 4 – Results of comparison of experimental and theoretical data on the motion of a screwless fish-like robot with the frequency of change in the rotor rotation direction is 1.25 Hz

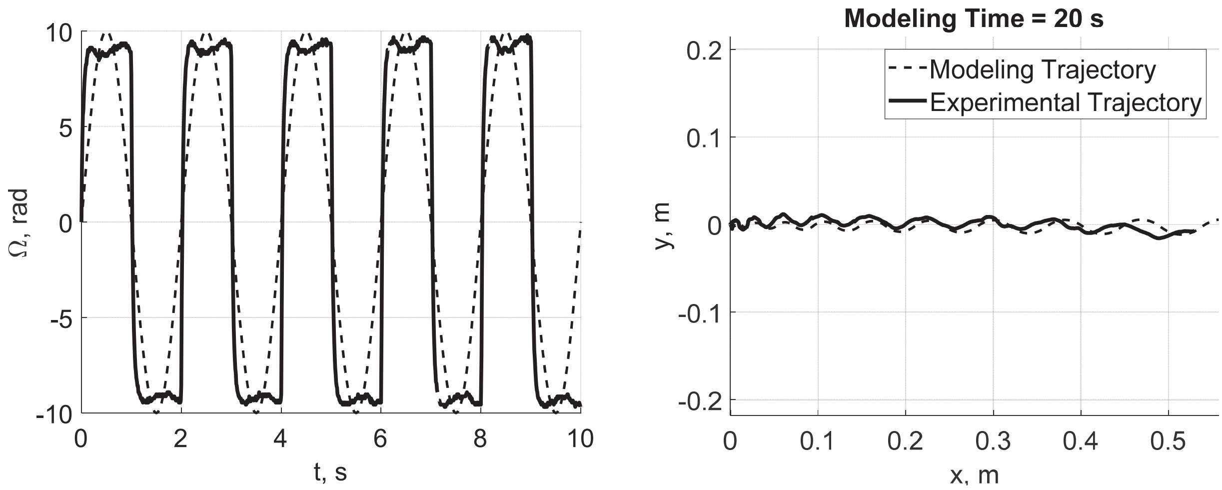
In the second series of experiments, the frequency of change in the rotor rotation direction is 1 Hz. Respectively, the period is 1 s. The time to overcome the distance of 0.5 m is 24 seconds. Figure 5, *a* shows the average dependence of the rotor speed on time in a real experiment (solid line) and the dependence of the rotor speed on time used in the simulation (dashed line). Figure 5, *b* shows the experimental trajectory (solid line) and the trajectory obtained as a result of modeling (dashed line). In this experiment, the average speed of the robot was 0.021 m/s.



a b

Figure 5 – Results of comparison of experimental and theoretical data on the motion of a screwless fish-like robot with the frequency of change in the rotor rotation direction is 1 Hz

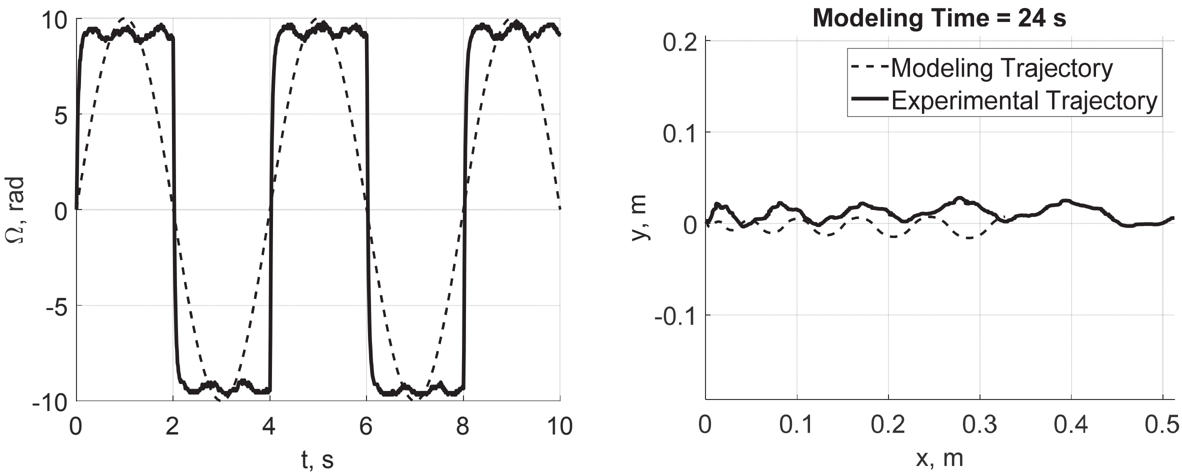
In the third series of experiments, the frequency of change in the rotor rotation direction is 0.5 Hz. Respectively, the period is 2 s. The time to overcome the distance of 0.5 m is 20 seconds. Figure 6, *a* shows the average dependence of the rotor speed on time in a real experiment (solid line) and the dependence of the rotor speed on time used in the simulation (dashed line). Figure 6, *b* shows the experimental trajectory (solid line) and the trajectory obtained as a result of modeling (dashed line). In this experiment, the average speed of the robot was 0.025 m/s.



a b

Figure 6 – Results of comparison of experimental and theoretical data on the motion of a screwless fish-like robot with the frequency of change in the rotor rotation direction is 0.5 Hz

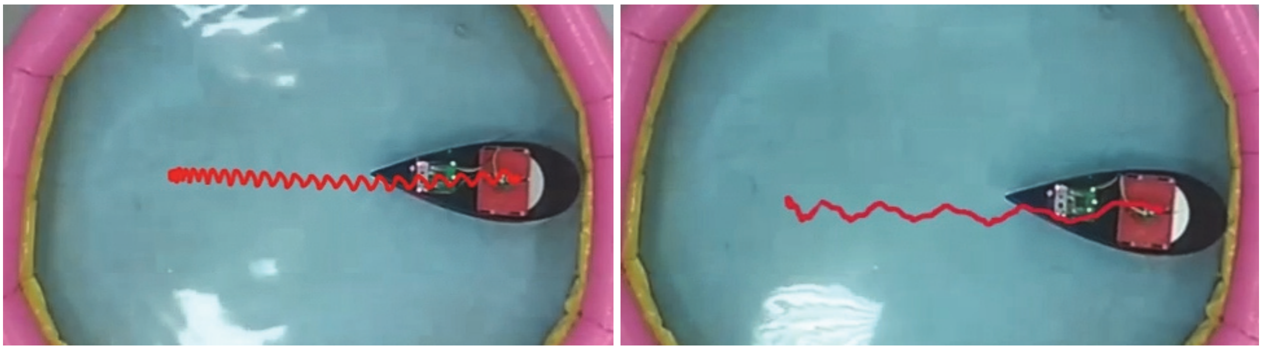
In the fourth series of experiments, the frequency of change in the rotor rotation direction is 0.25 Hz. Respectively, the period is 4 s. The time to overcome the distance of 0.5 m is 24 seconds. Figure 7, *a* shows the average dependence of the rotor speed on time in a real experiment (solid line) and the dependence of the rotor speed on time used in the simulation (dashed line). Figure 7, *b* shows the experimental trajectory (solid line) and the trajectory obtained as a result of modeling (dashed line). In this experiment, the average speed of the robot was 0.021 m/s.



a б

Figure 7 – Results of comparison of experimental and theoretical data on the motion of a screwless fish-like robot with the frequency of change in the rotor rotation direction is 0.25 Hz

For example, figure 8 presents video frames from experiments with the frequency of change in the rotor rotation direction – 1 Hz (Fig. 8, a) and 0.25 Hz (Fig. 8, b).



a b

Figure 8 – Frames from experiments with the screwless overwater fish-like robot

Conclusion

The obtained results confirm the possibility of movement of screwless fish-like robots due to periodic rotation of the internal rotor. It is shown that the nature and trajectory of motion depend significantly on the frequency of controls. For the considered design, the results of numerical simulation are consistent with the experimental data only for control actions with a frequency of 0.5 Hz. The quantitative discrepancy of the simulation results is possible due to the fact that the fluid resistance and circulation coefficients are not constant, but depend on the speed of the robot in the liquid. In the future, it is planned to develop a mathematical model of motion, taking into account these dependencies. Also the waves reflected from the walls affects on the results of the experiments, which make a contribution to the trajectory of the robot.

References

1. Kozlov V. V. Ramodanov, S. M. The Motion of a Variable Body in an Ideal Fluid// J. Appl.Math.Mech., 2001, vol. 65, no. 4, pp. 579–587.
2. Kozlov, V. V., Ramodanov, S. M. On the Motion of a Body with a Rigid Hull and Changing Geometry of Masses in an Ideal Fluid// Dokl. Phys., 2002, vol. 47, no. 2, pp. 132–135.
3. Vorochaeva (Volkova) L. Y., Jatsun S. F. Control of the three-mass robot moving in the liquid environment, Rus. J. Nonlin. Dyn., 2011, Vol. 7, No. 4, pp. 845-857.

4. Klenov A.I., Kilin A.A. Influence of vortex structures on the controlled motion of an above-water screwless robot // Regular and Chaotic Dynamics. 2016. Vol. 21. No. 7–8. pp. 927–938.
5. Karavaev Y. L., Kilin A. A., Klekovkin A. V., Experimental Investigations of the Controlled Motion of a Screwless Underwater Robot, Regular and Chaotic Dynamics, 2016, Vol. 21, No. 7-8, pp. 918-926.
6. Loftin Jr L. K., Smith H. A. Aerodynamic Characteristics of 15 NACA Airfoil Sections at Seven Reynolds Numbers from 0.7 x 10 (exp 6) to 9.0 x 10 (exp 6). – 1949.
7. Jatsun S., Lusnikov B., Politov E., Knyazev S. Underwater floating robot-fish: a comparative analysis of the results of mathematical modelling and full-scale tests of the prototype // MATEC Web of Conferences. – EDP Sciences, 2017. – Vol. 113, 02014, P. 1-5.
8. Jatsun S.F., Lushnikov B.V., Kazaryan K.G., Vorochaeva L. Yu., Vorochaev A.V. Design features of the bionic robot fish // Proceedings of South-West State University. Series Technics and Technologies. 2017. Vol. 7, No 2(23). pp. 94-102.
9. Borisov A. V., Mamaev I. S., Vetchanin E. V. Dynamics of a Smooth Profile in a Medium with Friction in the Presence of Parametric Excitation, Regular and Chaotic Dynamics, 2018, vol. 23, no. 4, pp. 480-502.
10. Mamaev I. S., Vetchanin E. V., The Self-propulsion of a Foil with a Sharp Edge in a Viscous Fluid Under the Action of a Periodically Oscillating Rotor, Regular and Chaotic Dynamics, 2018, vol. 23, no. 7-8, pp. 875-886.
11. Korotkin A. I.Added Masses of Ship Structures,Fluid Mech. Appl., vol. 88, Dordrecht: Springer, 2009.

A.A. Boreiko, A.A. Kushnerik, D.N. Mikhailov, A.F. Scherbatyuk

CURRENT EXPERIENCE FOR USAGE OF SOME AUV DEVELOPED IN IMTP FEB RAS

Institute for Marine Technology Problems FEB RAS, Vladivostok, Russia
alex-scherba@yandex.r

Abstract

The unification of autonomous underwater vehicles /AUV/ developed in IMTP FEB RAS is important part of work. The base model of unified AUV is created with open mechanical and information architecture that is lightly modified for different tasks. The specific feature of this approach is wide range of equipment and sensors installed on board of AUV. The AUV MMT 3000 series intended for operation on depth up to 3000 meters is one of examples of this approach. There were developed four AUV with different configuration that were used the same base of this series. Other examples of AUV series is MARK and later X200 that are intended for operation on depth up to 200 meters. Some details about mentioned AUV and fulfilled operations with MMT 3000 and MARC usage are described in the report.

Keywords: autonomous underwater vehicles, marine operations, unification.

Currently the different types of unmanned underwater vehicles are widely used for scientific and industrial experiments in the sea. More often autonomous underwater vehicles have used for different ecological monitoring and bottom objects inspection. IMTP FEB RAS has extended experience for execution of scientific researches and industrial missions in the sea with usage of different AUV developed by own. AUV MMT 3000 [1] and MARC [2] are the examples of IMTP production.

Some operations in Arctic seas

The vehicles of MMT 3000 series were used widely for operation in Arctic seas. In 2012 the work was intended for pollution task decision. The operation in 2016 was connected with investigation of the processes connected with bottom gas hydrate instability.

Operation in Okhotsk sea

In 2017 the AUV MMT-3000 with a modernized navigation system was used to carry out deep-sea operations for surveying of several sections of the seabed with a total length of more than 250 km [3-4]. The photo with AUV MMT 3000 on the supporting ship is shown on Figure 1.

Figure 1 – AUV MMT 3000 on the supporting ship

In accordance with the customer requirements each section should be traversed by two parallel lines at a distance of 50 m from each other. Thus, the total path length exceeded 500 km. The specified altitude of the AUV relative seabed during survey is 20 m, the specified speed of the AUV is 1 m/s. The maximum depth in the work area was not exceeded 2000 m. An average AUV path length during a diving was about 20 km. In the longest dive the AUV surveyed a seabed section with a length about 37900 m in 10.5 hours. During each dive the AUV was accompanied by a supporting ship. The supporting ship moved aside from the AUV at a horizontal distance of about 200-400 m.

The need to use the AUV for long deep-sea routes required the modernization of the present navigation system. It was required to provide a bottom survey on the routes with a total length of more than 250 km in 15-20 days. The working conditions were characterized by complex relief with inclines of up to 40 degrees, and sea current up to 1 m/s. A change in the magnetic declination up to 20 degrees was expected on some sections of the route.

The modernization of the AUV MMT-3000 navigation system included the realization of single mobile beacon positioning and introduction of additional equipment on board the underwater vehicle and the supporting ship:

– the AUV was equipped with a digital hydroacoustic modem EvoLogics S2C M 18/34 and a fiber-optic angular velocity sensor ВГО35КД to reduce the effect of magnetic anomalies and magnetic declination variations;

– the combined system of digital hydroacoustic communication and navigation with ultra-short base-line (USBL) EvoLogics USBL S2C 18/34, as well as the system of positioning and determining of spatial orientation Applanix POS MV V5 were used to avoid reinstallation of the transponder-beacons in the LBL APS.

Figure 2 shows the scheme which explains the operation of the modernized navigation system. The upgraded navigation and communication system works as follows. The supporting ship initiates the exchange on the hydroacoustic channel by periodically sending to the AUV of a special format messages. When the hydroacoustic modem of AUV receives this message, it sends a response package backward to the ship with the data on current state of the AUV.

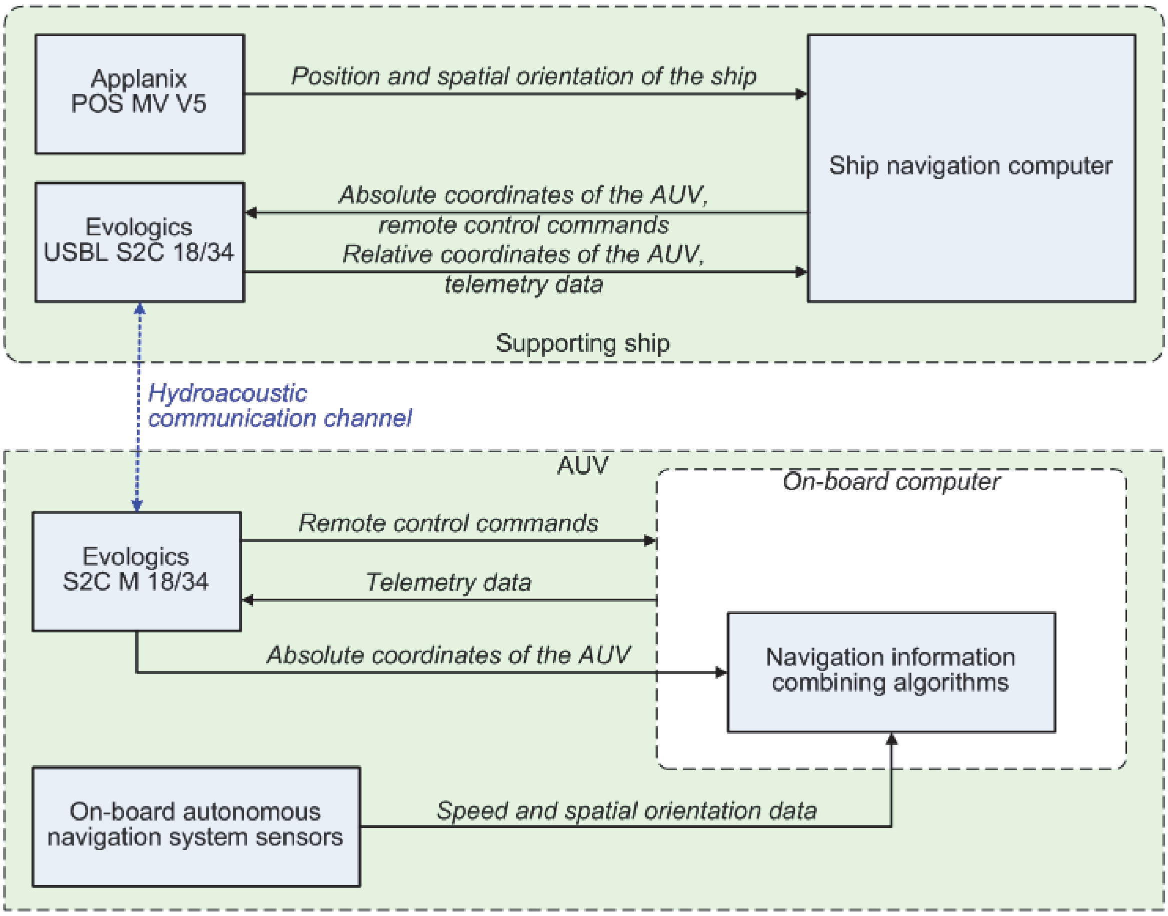


Figure 2 – Scheme of the modernized navigation system of AUV MMT-3000

When the ship modem receives this reply message, it transmits the telemetry data to the ship navigation computer for processing and for analysis by the operator. In addition, the position of the AUV relative to the ship's receiving hydroacoustic antenna is determined, and the absolute coordinates of the AUV are calculated by using the current position of the ship and its spatial orientation (the heading, pitch and roll angles from Applanix device).

These coordinates are transmitted in the next message from the ship to the AUV. Also the operator in case of need can send a command of remote control to the AUV (the corresponding data fields are provided in the transmitting message for that). It is important to note that the hydroacoustic communication system transmits data in the digital form and uses error correcting codes, which ensures high reliability of the received data.

Ecological Investigation in Zolotoy Rog Bay

Last years the researches connected with study of ecological conditions and influence of human activity for water areas became more and more wide. The processes of different substances distribution caused by the domestic and industrial activities attract special attention. One of the examples for water pollution is connected with plumes that arise as the result of industrial screenings from tubes or near the confluence of river with industrial effluents. Localization of water area with increased level of concentration for polluting substance or pollution source position detection are the tasks for investigation.

The traditional sampling methods based on research ship exploitation are expansive and do not allow the detailed covering of demanded water areas. Obviously, that the platforms like the autonomous underwater vehicles /AUV/ are mobile and able to investigate adaptively the processes during long time and for considerable distances. It supposed that AUV is equipped with the navigation system, intended for position detection with demanded accuracy, and with sensors for dissolved matter measuring. When the purpose for investigation is connected with detection of position for the pollution source or with localization of water area with increased level of concentration for polluting substance, the detailed covering of whole water area can take unacceptable long time. In this case the AUV allows applying the adaptive algorithms for trajectory planning.

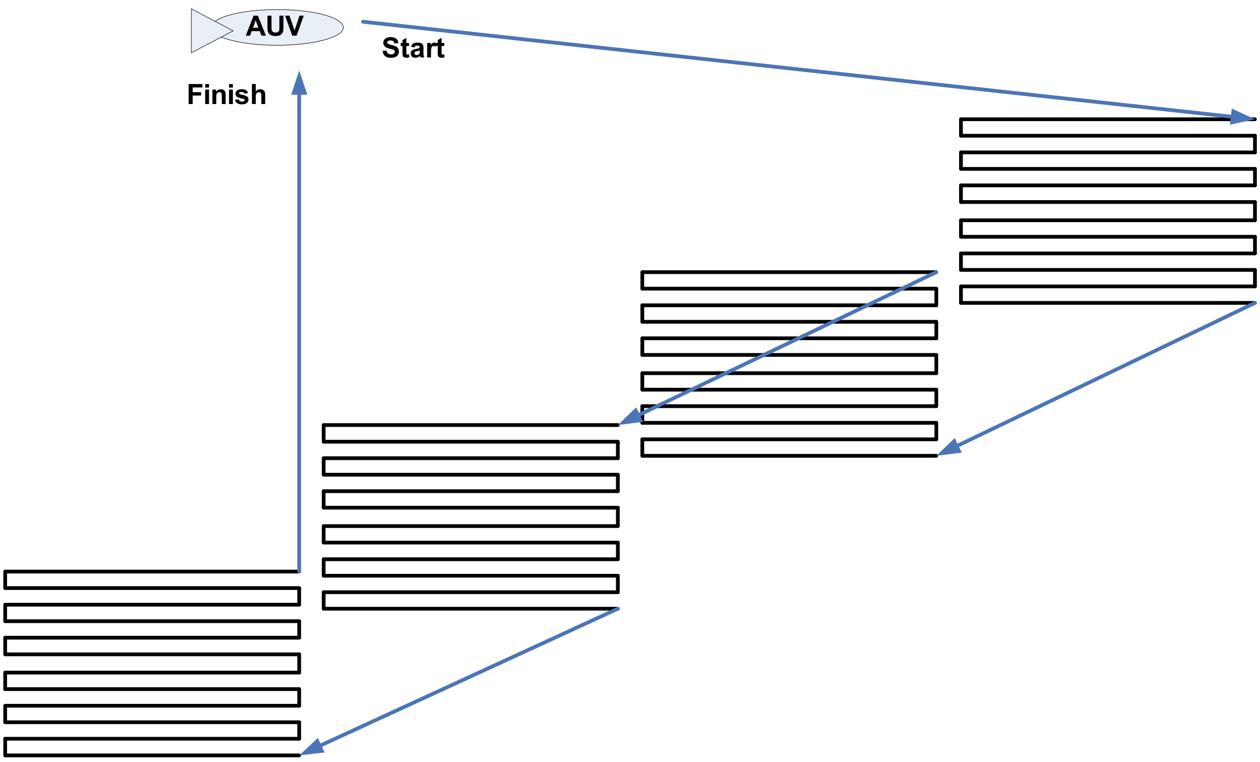


Figure 3 - The scheme of demanded water area covering by uniform horizontal net of vertical cuts

Let us consider the task of the bay ecological state estimation using the autonomous underwater vehicles. It is supposed that AUV is fulfilling the water parameters sampling in the demanded area by covering the water column with uniform horizontal net of the vertical cuts (Figure 3). In the marine experiments the length of tacks may be 100-200 meters for the current cut and vertical distances between the neighbor tacks may be 2-5 meters. For allowing the AUV motion along the program trajectory it is necessary to detect on board of the vehicle its position with high precision and to correct AUV trajectory when it deflects from the predefined one, for example, because of existence of current in the water area. It is supposed that AUV is equipped with navigation system intended for the own position estimation with predetermined accuracy and with the sensor of the concentration for the dissolved substance.

In June 2017 IMTP FEB RAS together with Far Eastern Federal University fulfilled the research intended for estimation of the ecological state for Zolotoy Rog Bay (Golden Horn) in Vladivostok city [5]. The purpose of the work was data obtaining for 3D information receiving about physical and chemical parameters of water in the vicinity of the influx of the Obyasneniya river into the bay. AUV of type MMT 3000 (Figure 4), equipped with a regular autonomous navigation system (which includes GPS, depth sensor, Doppler log, compass and echosounding system) was used to fulfill the task. The AUV also contains hydroacoustic link and navigation facilities, which allow efficiently monitoring and controlling the motion of the robot. The robot was additionally equipped with FLCDRT-926 and FLNTU-665 fluorometers installed on the outside of the hull. The meter CTD-NV-2406 was used for temperature and salinity values measuring.

Figure 4 – AUV of type MMT 3000 during the operation in Zolotoy Rog Bay
with the environmental sensors on the top part of the AUV

The AUV performed several parallel transects across the bay to obtain the raw data. The bay state was estimated on the basis of the analysis of water medium parameters. They were measured along AUV motion path and linked with navigational information about AUV’s position and orientation. The vehicle enabled measurement and logging of such parameters as concentration of dissolved organic matters (CDOM) and chlorophyll, conductivity, temperature and turbidity of water. The accompanying navigation information included coordinates, yaw, pitch, roll, depth and speed of the motion.

To obtain the pattern for distribution of the parameters being measured in terms of depth, each transect included three forward and reverse passes at the depth of 2, 4 and 6 meters. Hovering motion mode was used for transition between the working horizons to minimize maneuvers within the water area. The robot was used in such a way, that vessels navigation in the bay didn’t stop during the operations. As far as a real AUV motion cannot coincide with rectangular grid (is not rectilinear), the methods of interpolation on curvilinear grid was used for environment parameters mapping in the specified water area.

Collected data allowed reconstructing a three-dimensional pattern of the distribution of the measured parameters in the water column and the profile of the bottom surface along the moving trajectory. Figure 5 shows the general picture of the experiment, with the AUV's motion profiles and measured density of CDOM along the track.

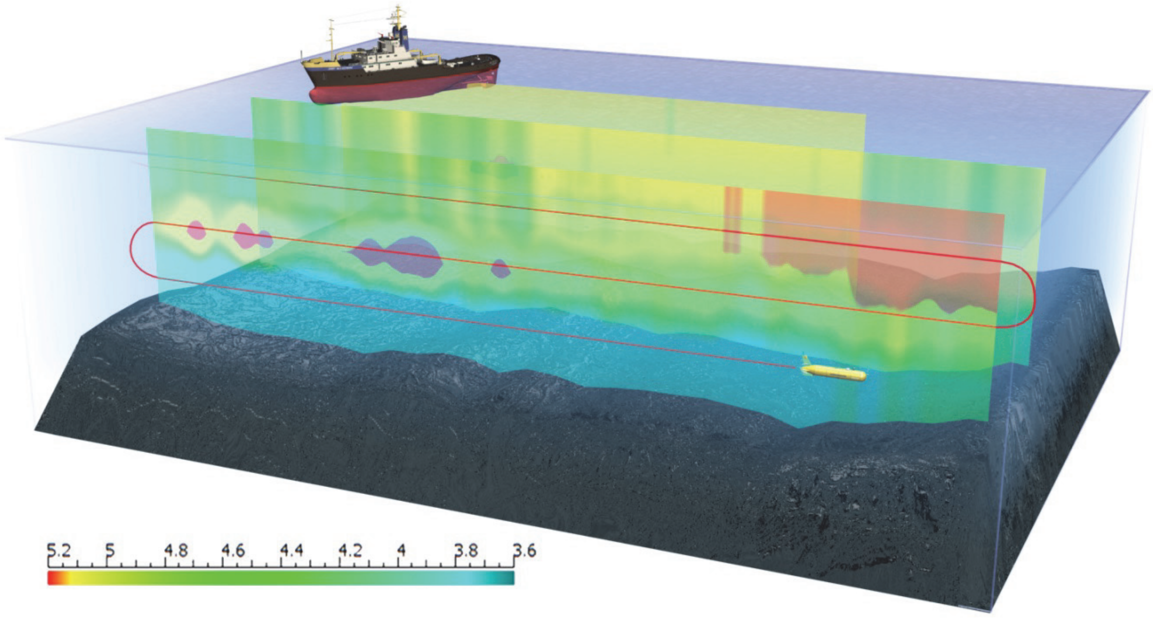


Figure 5 – General picture of the experiment, with the AUV motion profiles and measured density of CDOM [ppb] along the track

The obtained data made it possible to determine the picture of the impact of organic matter emissions produced by the Obyasneniya river on the state of the aquatic environment in the adjacent part of the Golden Horn Bay. The results of the operations also confirmed the effectiveness of AUV use for automated study of the parameters of water area medium and mapping with high spatial resolution.

Acoustical experiment on the shelf

In 2016 and 2018, IMTP together with POI FEB RAS were performed a test experiment in Vityaz Bay, the Sea of Japan, for studying of hydroacoustic oscillations and waves transformation on border of system "hydrosphere - lithosphere" in natural conditions according to the scheme, described in [6-7]. AUV MMT 3000 or MARC were used for hydro acoustical field sensor transportation.

The AUV hydrophone registered pressure variations of the hydroacoustic field and the obtained data were memorized onboard AUV. At the same time the system of coastal laser strainmeters recorded the seismoacoustic waves resulting from transformation of hydroacoustic waves on border "water-bottom".

The method supposes:

– the mechanism investigation of hydroacoustic waves transformation into superficial and volume waves of the earth's crust;

– the space-time wave and power structure definition of the hydroacoustic waves extending on the wedge-shaped shelf;

– studying of physics for different multi-scale oscillations and waves of the atmosphere, hydrosphere and lithosphere on space-time structure of hydroacoustic fields of the inspected water areas.

In 2016 AUV MARC was used in the marine experiment. It was equipped for experiment with hydroacoustic receiving devices: hydrophone 8104 of the Bruel & Kjaer Company, preamplifier, analog-to-digital converter, device of record and data storage. This equipment was installed inside pressure volume in nose part of the vehicle (figure 6). The Bruel & Kjaer hydrophone has the following technical characteristics: the frequency range of measurements from 0.1 Hz to 120 kHz with sensitivity – 205 dB 1 V / mkPa. In addition to a low-frequency hydroacoustic oscillator 33 Hz, the equipment includes low-frequency hydroacoustic oscillators 22 and 24.5 Hz.

Figure 6 – Additional module in the nose part of AUV MARC
with hydro acoustical receiving system

The method of experiment includes some steps:

− supporting ship was fixed at the shallow water area on the laser strainmeters line;

− low-frequency hydroacoustic oscillator (radiator) was installed from research vessel /RV/ on the 15 meters depth and operated continuously;

− the AUV was moving along the laser strainmeters line from supporting ship to coast and back and was measuring the level of acoustical signal; AUV used meander like trajectories in vertical plane with definite step on the constant heights above the bottom or at the constant depths;

− the laser strainmeters measured simultaneously the level of acoustical signal on the land;

− the level of transmitted acoustical signal was measured by the hydrophone installed near the radiator;

− all measuring and transmitting devices were synchronized using the system of exact time.

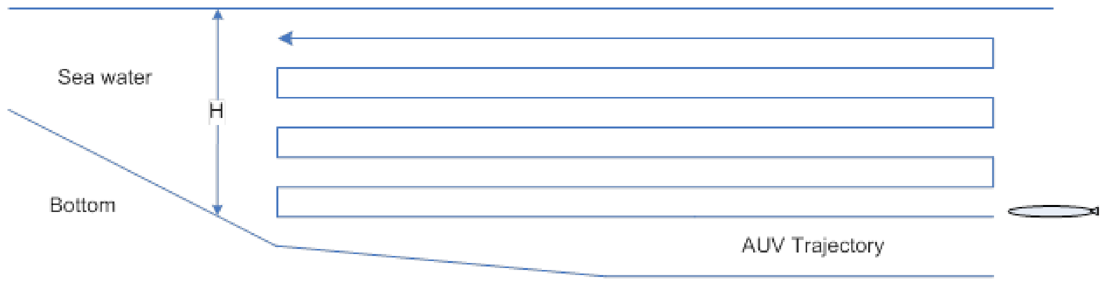
The main aim of the experiment in 2018 concludes in more precise measurements fulfilling. It was used AUV MMT 3000 in the trials that allows more precise navigation and following along the program trajectories. AUV MMT 3000 tethered the hydroacoustic field sensor by link of 7 meters length that allows to decrease significantly the own AUV acoustical noises in the received data.

Two laser strainmeters of classical type with working shoulder of 52.5 and 17.5 m, one 52-meter laser strainmeter of pendulum type [12], AUV MMT 3000 (fig. 7) [13] with hydroacoustic receiving system, a low-frequency hydroacoustic oscillator (radiator) at 33 Hz and system of exact time were used in the marine experiment. Laser strainmeters are intended for measurement shifts of crust with an accuracy of 0.1 nanometers in the frequency range from about zero (conditionally) to 1000 Hz.

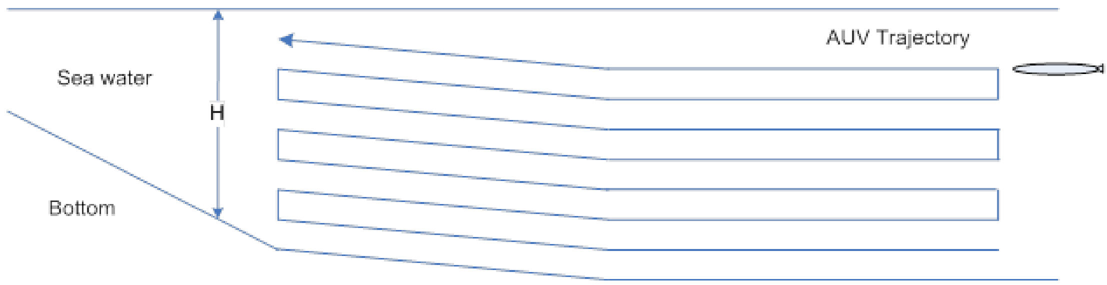
The operation of AUV MMT 3000 positioning system is based on modem communicational link. It allows synchronous exchanging by parcel with navigation data between AUV and supporting ship and simultaneously measuring of the acoustical signal propagation time between them.

AUV and supporting ship exchange by navigation data in turn manner during operation. The underwater vehicle detects own position and supporting ship tracks the AUV motion trajectory on the base of this information. Information parcel from supporting ship includes the current coordinates of ship beacon calculated using data from ship GPS. The AUV coordinates are calculated on board of the vehicle using ship beacon position information and measured distances data received during several cycles of synthesized LBL acoustic positioning system /APS/ operation. Derived AUV position together with measured AUV speed, heading and depth are transmitted to the supporting ship in answer informational parcel.

The AUV MMT-3000 positioning equipment includes the AUV digital hydroacoustic modem EvoLogics S2C M 18/34 and a fiber-optic angular velocity sensor ВГО35КД to reduce the effect of magnetic anomalies and magnetic declination variations.



a) AUV meander like trajectory at the constant depths,



b) AUV meander like trajectory on the constant heights above the bottom

Figure 7 – AUV trajectories in vertical plane

The combined system of digital hydroacoustic communication and navigation with ultra-short base-line (USBL) EvoLogics USBL S2C 18/34, as well as the system of positioning and determination of spatial orientation Applanix POS MV V5 were used on the supporting ship.

References

1. V.E.Gornak, A.V. Inzartsev, O.Yu.Lvov, Yu.V. Matvienko, A.Ph. Scherbatyuk. MMT-3000 – Small AUV of New Series of IMTP FEB RAS. Proceedings of the OCEANS 2006 MTS/IEEE Conference, September 18-21, 2006, Boston, USA, ISBN CD-ROM: 1-4244-0115-1.
2. F.S. Dubrovin, A.F Scherbatyuk, Vaulin Yu. V., Kushnerik A.A., Tuphanov I.E.. Small size AUV MAPC of new generation for group operation execution. // Mechatronics, automation, control, №6, 2012, p. 59-65.(in Russian).
3. F.S. Dubrovin, A.F Scherbatyuk, Yu. V. Vaulin. Some results of operation for the AUV MMT 3000 mobile navigation system on long and deep water trajectories. Proceedings of the OCEANS 2018 MTS/IEEE Conference, Kobe, Japan. ISBN: 978-1-5386-1653-6.
4. Mikhailov D.N., Dubrovin F.S., Boreyko A.A. and others. Application of AUV for hydrographic research in Okhotsk sea. // Underwater investigation and robotics, №2, 2017, p. 4-13. (in Russian).
5. Inzartsev A.V., Boreyko A.A., Borovik A.I., Vaulin Yu.V., Matvienko Yu.V., Scherbatyuk A.F and others. Experience for AUV MT 2010 usage for ecological investigation Zolotoy Rog Bay // Ecological systems and devices, 2018, № 12, p. 38-45. (in Russian).
6. G.I. Dolgikh, V.A. Chupin and A.F. Scherbatyuk. A Method for the Space-Time Distribution Studying of the Hydroacoustic Fields Near the Bottom Using AUV. Proceedings of the IEEE OES International Symposium on Underwater Technology, February 21-24, 2017, Busan, Korea.
7. G.I. Dolgikh, V.A. Chupin, Boreiko A.A., Mikhailov D.N., M.S. Sporyshev and A.F. Scherbatyuk. Improved Marine Experiment for Studying of the Hydroacoustic Fields Space-Time Distribution Near the Bottom Using AUV. Proceedings of the IEEE OES International Symposium on Underwater Technology, April 16-19, 2019, Kaoshiung, Taiwan.

V.B. Schneider, I.P. Janayt, I.A. Shavyrin

DESIGNING A HYDROACOUSTIC RANGE SENSOR OF THE UNDERWATER ROBOT NAVIGATION COMPLEX

Moscow Aviation Institute (National Research University), Moscow, Russia
vschndr@gmail.com

Abstract

The analysis and justification of the requirements for the hydroacoustic distance-measuring sensor of the underwater robot navigation system was performed. The specificity of the application is given and the features of its design are considered. The questions of modeling, testing the components of the sensor, the manufacture of a prototype are considered, its characteristics are given.

Keywords: uninhabited underwater vehicle, hydroacoustic distance-measuring sensor, altimeter, obstacle detection, integrated navigation system.

1. Introduction

The need for accurate navigation for underwater apparatus was considered in [1,2,3] and is determined by the tasks and assignments of underwater vehicles. An accurate navigation system is needed for localization, positioning, tracking, guidance and control. Typical navigation system of uninhabited underwater vehicles (UUV) includes a magnetic compass, angle sensors (roll, trim), angular velocity sensors (yaw, roll, trim) and depth sensor [4]. However, an important feature of the navigation system is the use of sonar sensors that are, in particular, locators designed to measure distances to objects.

Directed down locator is designed to measure the distance to the seabed. It is used to stabilize the vertical deflection of the apparatus from the ground, both when inspecting bottom objects and when inspecting ship hulls, quay walls, etc.

Front locators - designed to stabilize the horizontal distance of the device to the object being inspected. Based on data from these locators, the angular position of remotely controlled unmanned underwater vehicles (RCUUV) in the horizon relative to the vertical surface being examined, as well as the distance to it, is calculated.

The locator directed up is designed to measure the distance upwards and is used to organize the movement of the apparatus under the ice or in cavities [4].

The issue of preventing collisions of autonomous vehicles with obstacles (pier, ship hull, examined objects, bottom) is actual and is considered in [2]. To solve this problem, the use of different types of acoustic systems is considered: with a narrow directional beam and scanning systems (electronically or mechanically).

Scanning systems have several disadvantages:

− Systems with electronic scanning have high energy consumption, high cost and complexity of integration into the navigation and control system of the device, which requires ensuring a high speed of data exchange.

− Systems with mechanical scanning have a short range and low scan rate, which significantly complicates their use in the underwater vehicle control circuit.

The use of a sonar distance measuring sensor for solving collision avoidance problems allows detecting an obstacle in a narrow beam with high accuracy and a high scan rate. The possibility of simultaneous operation of several such sensors (each in its own direction) makes it possible not only to hold the device at a given distance from the bottom, but also to detect an obstacle in the direction of moving or to provide parking to the docking station.

Single-channel acoustic systems with a narrow directivity pattern have traditionally been used to solve problems of navigation, as an altimeter, and to prevent collisions with obstacles, as a measurer of distance to obstacles. Table 1 shows the most widely used in the world systems of single-channel altimeter distance meters.

Table 1. Technical characteristics of altimeters for underwater vehicles

Name	ГА-10	ГА-30	PA200	PA500	Model 864	Model 863
Country Developer-firm	Russia “Akvazond” LLC		Great Britain «Tritech International Limited»		Canada «IMAGENEX TECHNOLOGY CORP»	

Name	ГА-10	ГА-30	PA200	PA500	Model 864	Model 863
Working frequency, kHz	100	300	200	500	330	330
Weight, kg in air	≤ 4	≤ 1	1,15	1,15	1,2	–
Weight, kg in water	–	–	0,8	0,8	–	–
Overall dimensions, mm			Ø47 x160	Ø47 x160	Ø89 x125	
Supply voltage	from 22 V to 30 V		from 12 V to 24 V		from 22 V to 32 V	
Consumed Power, W	≤ 20	≤ 15	2	2	2,2	2,2
Beam width	12°	10°	20°	6°	10°	10°
Range of detectable distances	0,4m to 1000m	0,2m to 150m	0,7 to 100	0,3 to 50	from 0,1 to *	from 0,1 to *
Instrumental error, m	≤ 0,007		≤ 0,001	≤ 0,001	≤ 0,02	
Maximum immersion depth	6000 m		700 m polyacetal 4000 m aluminium alloy 4000 m stainless steel 6800 m titanium alloy 6Al-4V		100 m aluminium alloy 6061-T6	
Interface	Ethernet or RS-232		RS232 or RS485		RS232 or RS485	

In Russia, the manufacturer of altimeters for underwater vehicles is LLC “Akvazond” (Taganrog) [5]. Two types of deepwater long-range sensors are presented: GA-10 – 1000 m and GA-30 – 150 m. They are intended, mainly, for operation on heavy deep-water apparatuses. A significant drawback of these devices is their high power consumption - 15...20 watts. For light inspection RCUUV and light AUUV, these altimeters are not optimal [6]. Unfortunately, most foreign systems are not available on the Russian market for foreign policy reasons.

2. Analysis of the technical requirements for distance measuring system

The main technical requirements [6,7,8] for acoustic range measurements by altimeters are:

1. The maximum range for the altimeter caused by the required maximum height of the device holding above the bottom, as well as the smallest detection range of obstacles, required to perform a maneuver or stop the device. These values are determined by the class of the apparatus. In the case of the most common light class, the required distance is from 50 to 100 meters.
2. Requirements for the minimum distance are determined by the ability to move safely the underwater vehicle near the ground or object and lies within the range from 0.1 to 0.7 m.
3. Requirements for the maximum depth of immersion are determined by this characteristic of underwater vehicles. Devices with the immersion depth of 300 ... 500 m, 600 ... 1000 m are most demanded. Machines with extreme immersion depths are single, and their supply does not exceed 1-5 units per year.
4. Requirements for the types of power source and the limit on power consumption. The most widespread UUV’s have instrument power of 24...30 V DC. The use of altimeters on light RCUUV of inspection class and light autonomous uninhabited underwater vehicles (AUUV) requires the device to consume not more than 2...5 watts.
5. Technical limitations on size and weight. The use of altimeters on light RCUUV of inspection class and light AUUV requires that the device (in the monoblock variant) have a mass in air of not more than 1.2 kg and a mass in water of not more than 0.8...0.9 kg. The altimeter in the monoblock version should have the following overall dimensions: diameter not more than 50.0 mm, length not more than 160...250 mm.
6. Acoustic sensor compatibility with other sonars
7. Low cost

3. Features of the projected acoustic distance measuring sensor

The design of a sonar distance-measuring sensor is considered in the paper. This sensor solves navigation tasks as part of the navigation system of underwater vehicles, as well as tasks for preventing collisions of underwater vehicles with obstacles. Characteristics of the developed system are presented in table 2.

The acoustic distance measuring sensor consists of an antenna and an electronic module placed in a sealed case (the altimeter is structurally designed as a monoblock). The block diagram of the hydroacoustic distance measuring sensor is shown in Figure 1.

Table 2. Characteristics of hydro-acoustic distance measuring sensor

Parameter	**Value**
Working frequency, kHz	500
Antenna beam width at -3 dB level, degrees	±3
Operating frequency bandwidth, kHz	50
Minimum measurable range, m	0,1
Maximum measured distance, m	100
Range resolution, mm	1,5
Data update rate, Hz	1 - 10
Working depth, m	600
Input power supply voltage	9-36V
Overall dimensions, mm	50x250

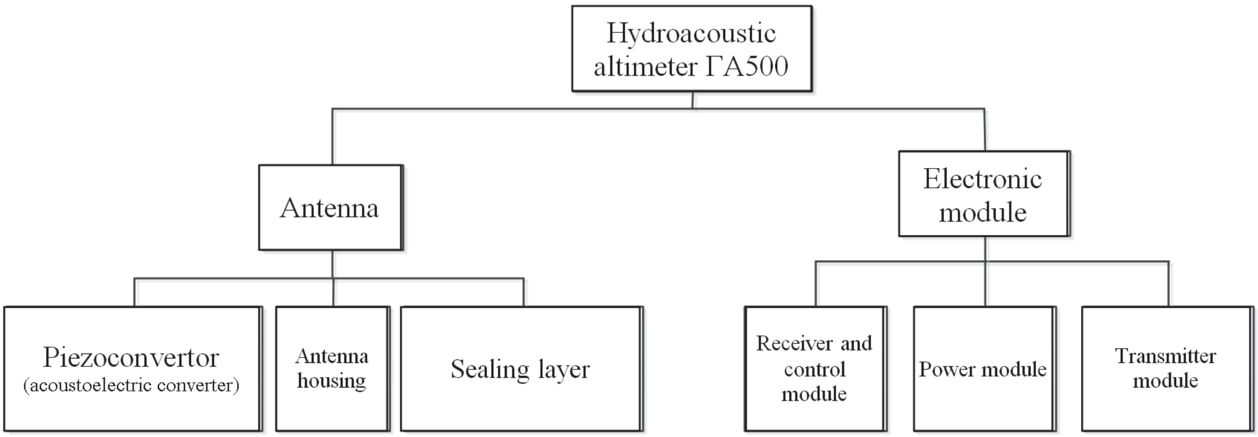


Figure 1 – Block diagram of the hydroacoustic distance measuring sensor

Antenna

Directivity characteristic

As an acoustoelectric converter of the antenna module, a piezoceramic disk was chosen, the dimensions of which provide a predetermined width of the radiation pattern of ± 3 degrees. The working frequency band of this type of converters is up to 10%, which, at an operating frequency of 500 kHz, will allow a signal bandwidth of 50 kHz, which in turn will provide a resolution of 1 mm.

Materials

Studies have been conducted on antennas characteristics made from different types of materials. An antenna made of soft piezoceramic material CTS-19(ЦТС-19) showed the best sensitivity values for reception and radiation at a static pressure of 6 MPa. The appearance of the antenna is shown in Figure 2.

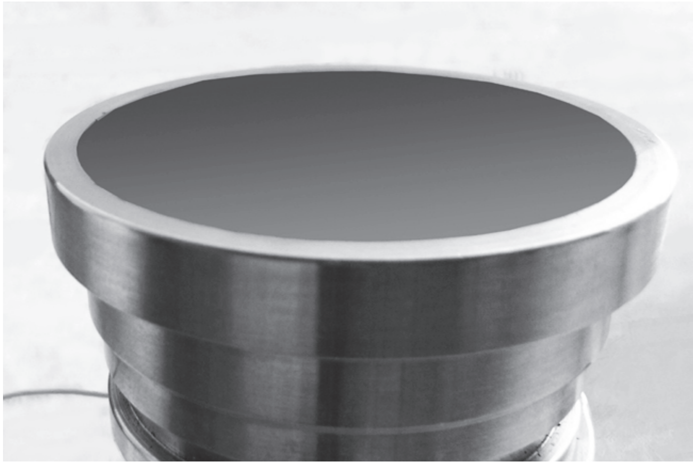


Figure 2 – Antenna of acoustic distance measuring sensor

The measured antenna parameters - radiation directivity pattern and sensitivity for reception and emission are shown in Figures 3, 4 and 5.

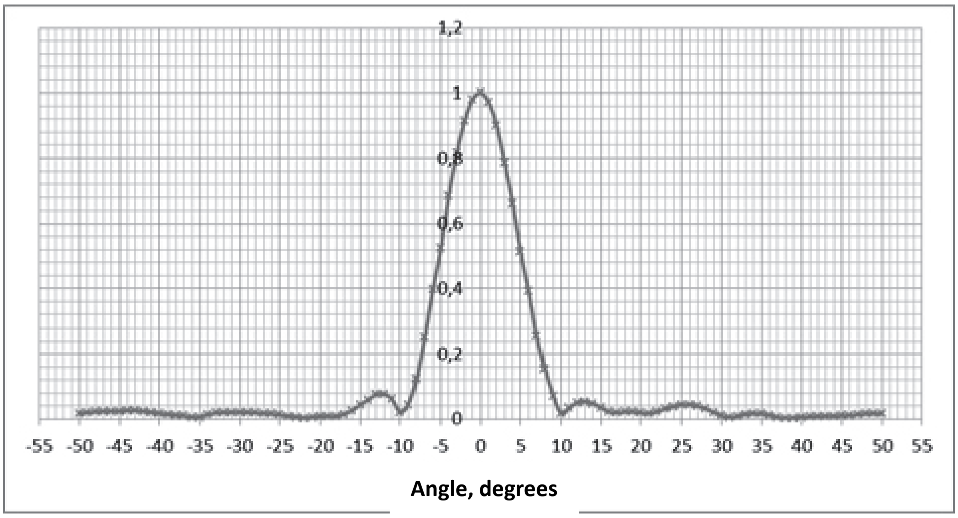


Figure 3 – Antenna directivity pattern

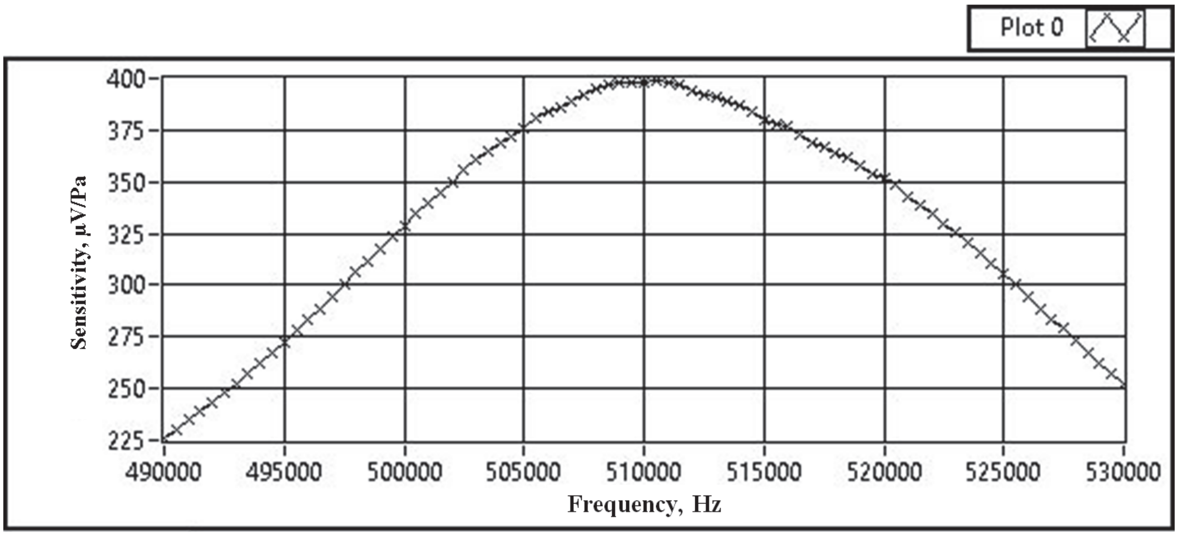


Figure 4 – Antenna reception sensitivity

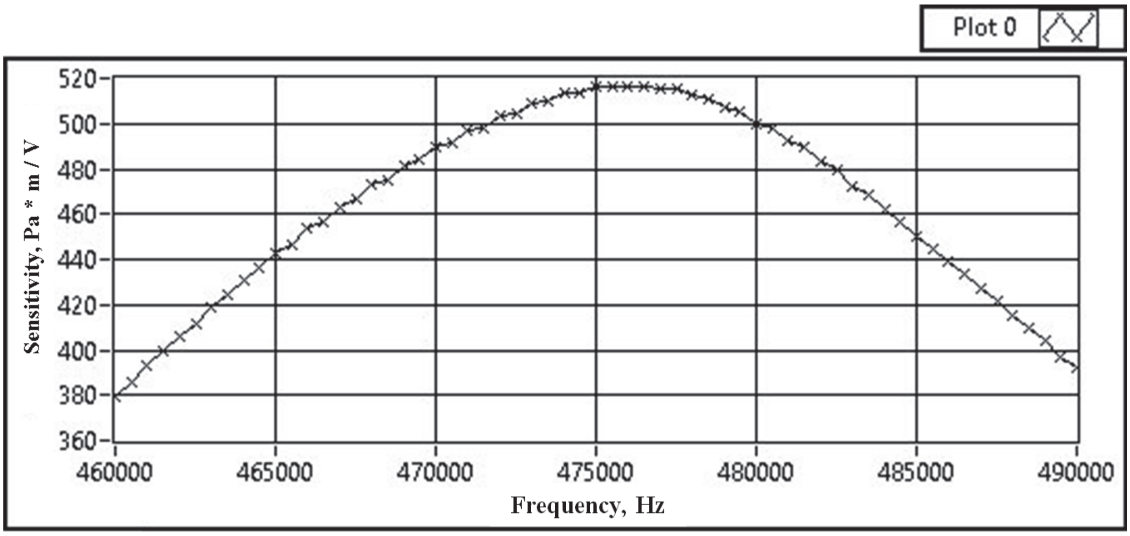


Figure 5 – Antenna radiation sensitivity

Electronic module

The electronic module of the sensor implements the classical optimal scheme for receiving a deterministic signal. The functional diagram is presented on Figure 6.

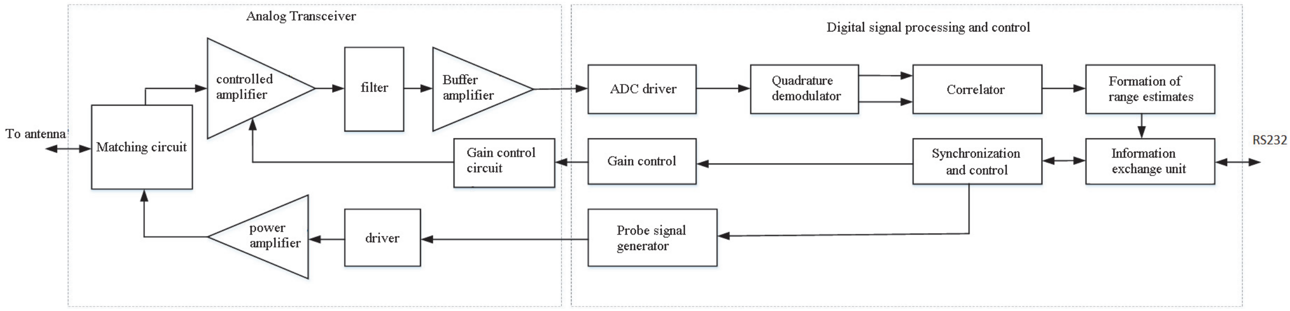


Figure 6 – Functional diagram of the electronic module

The scheme shown in Figure 6 can be divided into two parts: an analog transceiver and a digital signal processing and control unit. In turn, the first part of the electronic module can be divided into a receiver and a transmitter, connected to the antenna through the matching circuit.

The transmitter consists of a power amplifier that generates an electrical signal at the antenna input, which is converted to an acoustic pressure of 150 kPa at distance of 1 m. A driver is connected to the input of the power amplifier, which converts the signal from the driver output to the required level for the amplifier.

The receiver consists of a gain-controlled input amplifier, an anti-alias filter and a buffer amplifier. To increase the dynamic range of the receiver, gain control in time is used, which allows to provide a dynamic range of received signals up to 160 dB. In this case, the spectral density of the receiver noise, reduced to the input, is 3 nV/√Hz.

Digital signal processing is implemented by Kotelnikov's receiver [9]. A feature of this scheme is the selection of threshold values at different times, which is adaptive to the noise level. This solution allows to set the false alarm probability. It is known that correlation reception is effective when using complex signals with a base of more than 20 [10]. However, since many complex signals are very sensitive to the Doppler effect, the generation and processing of signals from an acoustic range sensor was simulated.

As a result, a probe signal with frequency modulation according to a hyperbolic law was chosen, the correlation function of which is relatively stable to the Doppler effect [11].

The system has a table method for signals generating. At the same time, the size of the signal representation table was chosen on the basis of ensuring the level of quantization noise [12], minus 80 dB, which is due to the dynamic range of the signal representation in the processing system.

It should be noted that the implementation of a table-like representation of a signal makes it possible, with simultaneous operation of several sensors, to emit orthogonal signals, the values of mutual correlation functions of which are substantially (about 40-50 dB) less than the values of their auto-correlation functions.

Housing

In the design process, a study for cases of various materials was made. The use of titanium alloy materials OT-4 and stainless steel 12X18H10T was considered as a material for the case. A simulation was performed for loading parts of an acoustic range sensor made from different materials in the static analysis module of the Catia program.

Hydrostatic pressure of 6 MPa and distributed loads corresponding to loads from interaction with other parts of the carcass were taken as factors affecting the details. Figure 7 shows the simulation results for the hull made from OT-4 material.

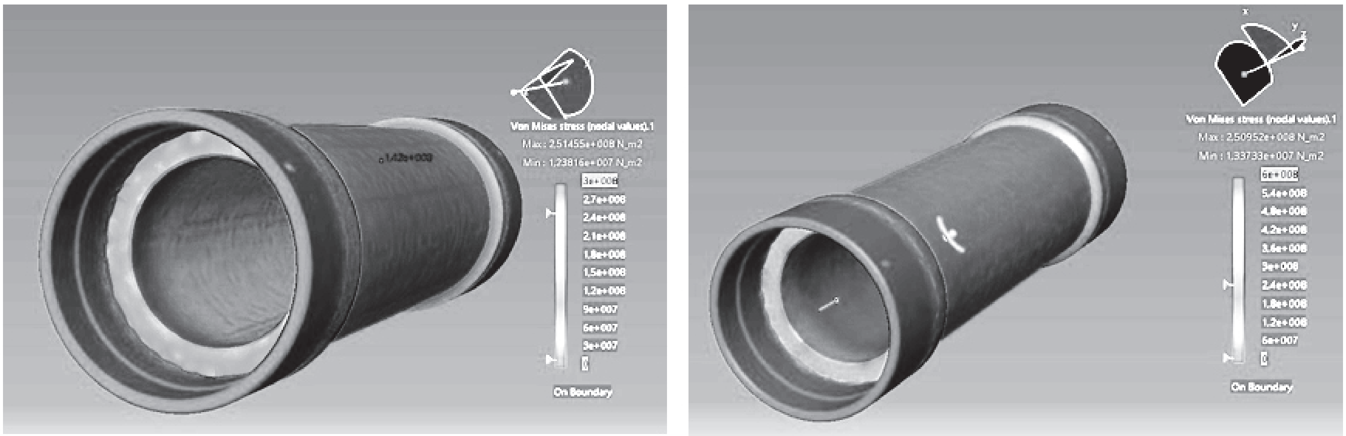


Figure 7 – Housing stress distribution

The simulation showed the following results:

– Stainless steel case model. Maximum stress: - σ_{Max} = 251 MPa. Reserve factor: σ_T / σ_{Max} = 1.19

– Titanium case model. Maximum stress: - σ_{Max} = 251 MPa. Reserve factor: σ_T / σ_{Max} = 2.39.

The simulation made it possible to optimize the design in terms of mass and dimensional parameters while ensuring resistance to the effects of pressure arising at a depth of 600 meters.

The housing photo is shown in Figure 8.

Figure 8 - Photo of the manufactured housing for acoustic range sensor

The measured parameters of the manufactured sensor modules allow the calculation of the maximum range, taking into account the real system sensitivity and the achieved signal compression parameters. The calculated values of the range depending on the noise at the input are shown in Figure 9. The reflection coefficient from the bottom is assumed to be -18 dB, the pulse duration is 10 ms.

Calculations show that the acoustic range sensor has an input noise margin for operation in a wide range of underwater vehicle motion modes.

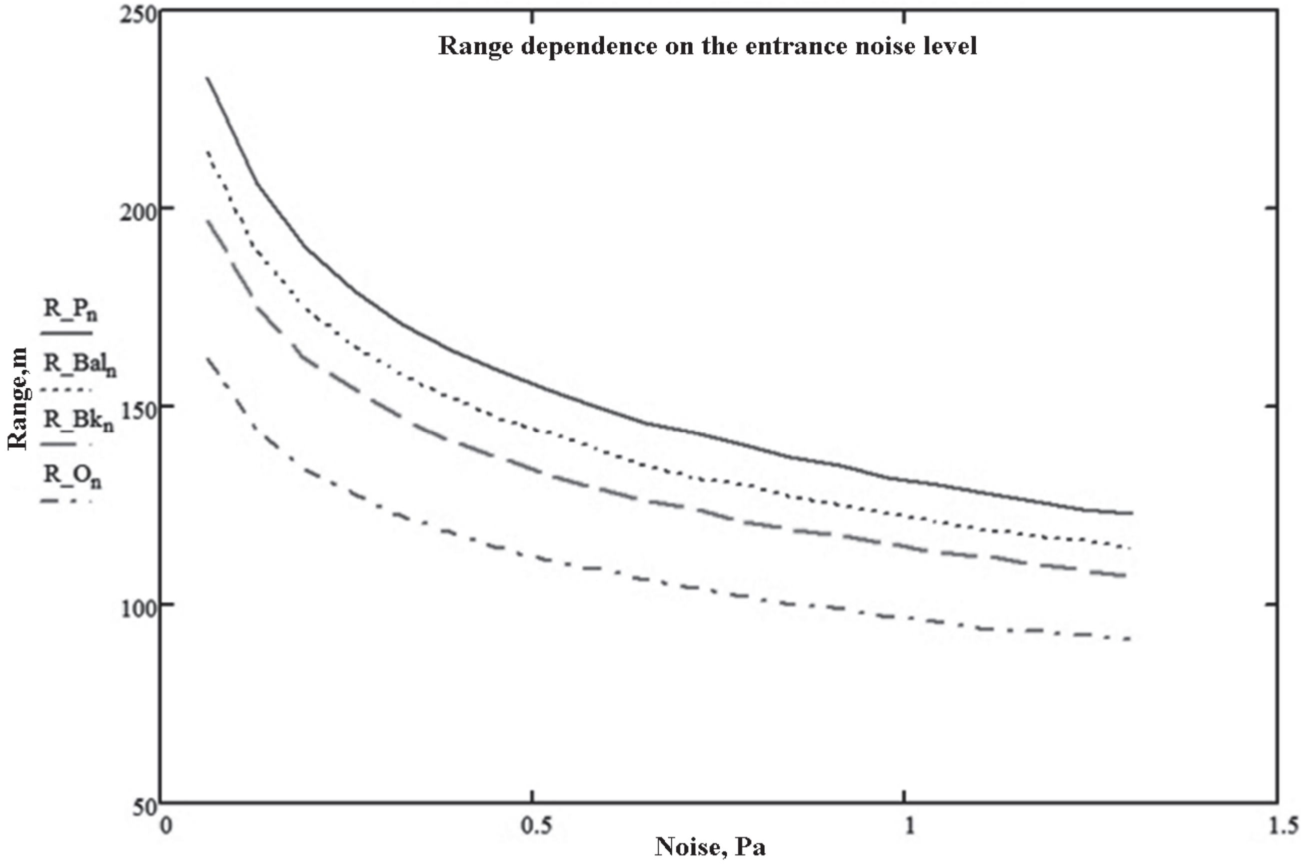


Figure 9 – Range depending on the noise at the entrance for the environment of different salinity.
Designations: R_P – fresh water; R_Bal – Baltic; R_Bk – Black Sea; R_O – World Ocean

It should be noted that the acoustic noise of the aquatic environment (in the absence of movement) is about 30-60 MPa. The major sources of noise are the noises of the flow around the antenna as the vehicle moves and the noise of the engines [13].

For typical conditions of use without additional external noise, the expected working range varies from 260 meters for fresh water to 180 m in the ocean.

Conclusion

The validation of the requirements for the acoustic distance-measuring sensor has been fulfilled. The specificity of the application is given and the features of its design are considered.

At present, the modeling has been completed, all the sensor components have been tested, a prototype has been manufactured, and its field tests are being prepared.

References

1. Achieving High Navigation Accuracy Using Inertial Navigation Systems in Autonomous Underwater Vehicles Robert Panish and Mikell Taylor Bluefin Robotics Corporation 553 South Street Quincy, MA 02169 USA.
2. Polar Grid Navigation Algorithm for Unmanned Underwater Vehicles Zheping Yan, Lu Wang,* Wei Zhang, Jiajia Zhou, and Man Wang.
3. Organization of the navigational support of maritime robotic complexes of military appointment in operational-important areas of the world ocean galkin extreme robotics // Collection of theses of the International Scientific and Technical Conference. - SPb: Gangut publishing and printing complex, 2017 - 272p.
4. Navigational and algorithmic support of rcuuv for effective solution of tasks identification of bottom purposes and inspection of marine objects. U.V. Vaulin, V.V. Kostenko, A.M. Pavin.
5. Deep-water altimeters LLC Akvazond, Russia, Taganrog http://aquazond.ru/gal-series.html.
6. Autonoums Echosounders for Sediment Monitoring in shallow waters. Dr. Doowon Choi EofE Ultrasonics Co., Ltd. Prof. Guan-Hong Lee Inha University, Korea 2012. 10. 18 https://geo-matching.com/uploads/default/m/i/migrationhzjnus.pdf.

7. A bottom-following problem approach using an altimeter José Melo, Aníbal Matos INESC TEC (formerly INESC Porto) and Faculty of Engineering University of Porto Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal {jose.melo, anibal}@fe.up.pt
8. Obstacle detection and avoidance for AUV: problem analysis and first results (REDERMOR) Alain HETET (1),Isabelle QUIDU (2), Yann DUPAS(1)(1) Groupe d'Etudes Sous-Marines de l'Atlantique, DGA/DET/GESMA, BP 42 – 29240 Brest Armées, France, {alain.hetet, yann.dupas} @dga.defense.gouv.fr (2) Laboratoire E3I2 EA3876, Ecole Nationale des Etudes et Techniques de l'Armement – 2, rue François Verny, 29806 Brest Cedex 9, France {isabelle.quidu@ensieta.fr}.
9. Sklyar B. Digital communication. Theoretical foundations and practical application.
10. Cook C., Bernfeld M. Radar signals / Trans. from English to ed. V.S. Kelzon - M.: Soviet Radio, 1971 - 568p.
11. Correction of side lobes of the autocorrelation function of radar and sonar signals of complex shape, Baskakov A.I., Boldinov R.O., Radio-technical and telecommunication systems. 2015. № 1(17). Pp. 21-25.
12. Quantization noise in digital signal synthesizers and methods for reducing them. N.P. Kandyrin. Kharkov University of the Air Force named after Ivan Kozhedub, Kharkov.
13. Principles of Underwater Sound - Robert J. Urick.

V.S. Taradonov, A.P. Blinkov, I.V. Kozhemyakin,
V.A. Ryzhov, D.N. Shamanov, D.A. Dmitriev

THE CONCEPTUAL SHAPE OF THE ROBOTIC UNDERWATER – SURFACE VEHICLE OF THE INCREASED AUTONOMY WITH CHANGEABLE GEOMETRY OF THE HULL FOR THE SYSTEM OF ROBOTIZED UNDERWATER SEISMIC EXPLORATION IN SUBGLACIAL WATER AREAS

Saint-Petersburg State Marine Technical University, Russia
office@smtu.ru

Abstract

The purpose of researches was the development of a new underwater-surface vehicle of high autonomy based on the analysis of advanced developments of leading foreign states in the creating marine robotic complexes.

The main results of initiative research engineering on development of conceptual shape of the perspective robotic underwater-surface unmanned vehicle of the increased autonomy with changeable geometry of the hull are presented; options of its use at the solution of a wide range of tasks are given. Particularly there are presented any variants of its use in the system of robotic underwater seismic exploration in the subglacial waters, and in the information-measuring network during oceanographic studies. The proposed conceptual image of a promising robotic tool is developed based on the analysis of the advanced developments of the leading foreign states in the creation of surface, underwater and surface-underwater marine robotic complexes; advantages of the prospective robotic underwater- surface unmanned vehicle with increased autonomy in comparison with the best foreign analogs prototypes are given. One of undoubted advantages of the vehicle is the large volume reserved for payload, allowing installing onboard more precision, but power-intensive detectors and sensors. The variable configuration of the hull and the underwater-surface version of the vehicle minimize the contradiction between the need for high-speed deployment in the mission area, stability during searching and research operation, independence from external conditions (under water), and also ensure the operation of high-tech marine robotic tools. On the basis of calculations of hydrodynamic and power characteristics and also the analysis of different types of rechargeable batteries, available in the world market, preliminary configuration of the power station is made; the reserve of diesel fuel and rechargeable batteries and also ballast tanks is defined.

As a result of development of the concept of the robotic underwater-surface vehicle of the increased autonomy with changeable geometry of the hull we obtained the multipurpose universal robotic platform which combines all advantages autonomous unmanned surface and underwater vehicles and can be used at the solution of a wide range of tasks both in military, and in civil areas.

Keywords: robotic complex, autonomous unmanned surface vehicle, autonomous unmanned underwater vehicle, robotic underwater-surface vehicle, multifunctional universal platform, conceptual image, variable geometry of hull.

Introduction

Carried out recently developments of space, air, sea and ground-based robotic systems for monitoring of the World Ocean, show significant technical progress in the field and speak about emergence of innovative approaches in this sphere.

Extensive networks of the drifting and anchored buoys, autonomous measuring stations, gliders and also satellites of remote sensing of Earth enter the systems of monitoring today, except vessels and submersibles.

Undoubtedly that the priority by efficiency and coverage in monitoring of the World Ocean belongs to facilities of space sounding of Earth, but ocean depths are inaccessible to them. Here are yet out of competition the research vessels (RV), especially their role considerably are increased with application of remotely-controlled and autonomous underwater vehicles. They are capable to carry out the long, limited only by vessel's autonomy, researches in a defined sea area, however the size of the World Ocean are so great (its area is 71% of the area of all Earth, that is 360 million sq.km) that the most high-speed vessel will need many decades to visit all areas of the ocean. In that time the condition of its waters significantly changes, as a result only the fragmentary picture turns out, distorted because of observations are extended through time [1-2].

The traditional autonomous drifting and anchor buoys and also gliders and AUVs are much cheaper, but they are capable to register only some data (temperature, salinity, pressure, etc.) and cannot compete fully yet with the research vessels which are carrying out complex researches of sea.

The interest shown by various departments to marine robotic systems and also research and development in the field of modern technologies (new materials, nanotechnologies in power, IT) is led to rapid upgrading of MRS and their technical qualities necessary for the solution of oceanography problems.

Thus, in the short term oceanographic researches have to be conducted in a complex with use both traditional (research vessel, drifting and the anchored buoy-based station), and innovative research means (spacecrafts, gliders, AUUV). Their combined use is capable to lower considerably expenses and to compensate separate shortcomings of each of research means.

Development of conceptual shape of the robotic underwater-surface unmanned vehicle of the increased autonomy with changeable geometry of the hull

The St. Petersburg State Marine Technical University (SMTU) in the cooperation with the Main research test center of robotics of the Ministry of Defence of the Russian Federation in 2017-2018 carried out initiative researches of the name "Shadow" [3] and "Shadow-2" [4]. Research supervisor of both researches: Taradonov V.S. is candidate of technical sciences, head of University laboratory.

During this work it was carried out the analysis of the advanced developments of the leading foreign states in the field of creation of surface, underwater-surface, and underwater marine robotic complexes. In particular, it was studied materials on the autonomous semisubmersible vehicle of the SAIC company (USA) [5-10], the autonomous unmanned surface vehicle (AUSV) "Sea hunter" of trimaran type (USA) [11-14], the autonomous unmanned anti-submarine USV System "Seagull" (Israel) [15-19], the unmanned surface vehicle "Sentry", and on its prototype - the manned submersible craft "GHOST" (USA) [20-23].

Their characteristics, merits and demerits are analyzed.

The main shortcomings of all foreign complexes are:

– dependence on hydro-meteorologic conditions (HMC);
– impossibility of their operation in ice conditions;
– insufficient endurance.

On the basis of made analysis, specialists of St. Petersburg SMTU developed conceptual shape of the perspective robotic surface-underwater unmanned vehicle of the increased autonomy (RSUUV IA) with changeable geometry of the hull.

Its concept allows to combine the high speed at deployment to the region of mission (working) and independence from HMC in the underwater mode.

Calculations of its main hydrodynamic and power characteristics were made in accord withthe proven techniques, modern manufacturing techniques of energy-power plants and their actual parameters [24-32].

In figure 1 results of calculation of the actual spent power of one of options of full-scale RSUUV IA "Shadow-2" in various modes of its movement depending on speed are presented.

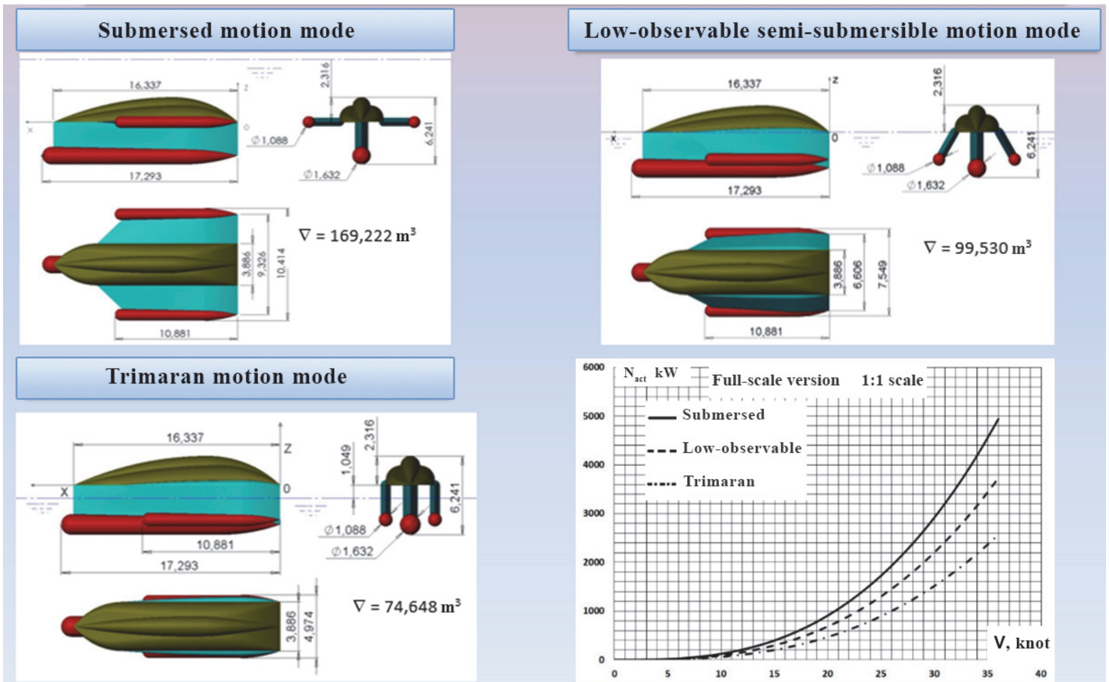


Figure 1 – Results of calculation of the actual spent power of RSUUV IA in various modes of its movement depending on speed

The conducted analytical studies showed advantage of design characteristics RSUUV IA in comparison with the best foreign analogs-prototypes.

Comparative performance characteristics of unmanned trimaran "Sea Hunter" (USA) and RSUUV IA are presented in Table 1 at the underwater mode, in the "unobtrusive" mode, and in high-speed "trimaran" mode [4]. From comparison it is visible what with the same combined capacity of engines, RSUUV IA "Shadow-2" has smaller mass-dimensional characteristics and much higher rates of endurance and range.

Table 1. Comparative performance characteristics unmanned trimaran "Sea Hunter" and RSUUV IA "Shadow-2"

Parameter	Unmanned trimaran "Sea Hunter"* (USA)	"Shadow-2" underwater mode	$\frac{(3)}{(2)}$	"Shadow-2" semisubmersible mode*	$\frac{(5)}{(2)}$	"Shadow-2" "trimaran" mode *	$\frac{(7)}{(2)}$
1	2	3	4	5	6	7	8
Displacement, tons	138,5	169,22	1,22	99,53	0,72	74,65	0,54
Maximum length, m	43,2	17,3	0,40	17,3	0,40	17,3	0,40
Maximum width, m	12,2	10,4	0,85	7,6	0,62	5,0	0,41
Maximum height, m	12,2	6,24	0,51	6,24	0,51	6,24	0,51
Maximum speed, knot	27	26,43	0,98	29,9	1,11	34,15	1,26
Patrol speed, knot	6	6	1,0	6	1,0	6	1,0
Number of diesel-generator sets	-	2	-	2	-	2	-
Total power of engines, kW	2134	2134	1	2134	1	2134	1
Mass of a reserve diesel fuel, tons	40	15,1	0,38	15,1	0,38	15,1	0,38
Mass of rechargeable batteries, tons	-	32,2	-	32,2	-	32,2	-
Relative weight of diesel fuel **	0,29 (29%)	0,09 (9%)	0,31	0,15 (15%)	0,51	0,2 (20%)	0,7
Relative weight of rechargeable batteries **	-	0,19 (19%)	-	0,32 (32%)	-	0,43 (43%)	-
Endurance at V = 6 knot, day	70	88,7	1,27	121,2	1,73	176,15	2,52
Range at V = 6 knot, miles	10080	12770	1,28	17451,3	1,75	25365,6	2,54
* Still water or small sea state **In comparison with displacement							

The offered conceptual shape of the robotic underwater-surface vehicle of the increased autonomy with changeable geometry of the hull "Shadow" in comparison with foreign analogs – prototypes has the following main advantages:

1. The modes of the movement "Shadow" at the same time include the modes of movement of all presented earlier foreign analogs – prototypes.

1.1. On still water and at small sea state RSUUV IA can travel in surface "trimaran" and "unobtrusive" modes and also under water close to a surface with a retractable or folding snorkel mast (semisubmersible mode).

1.2. At hard sea state, storm weather or in the difficult ice conditions, the "Shadow-2" can move under water at a necessary depth.

2. At the movement in a surface or underwater mode at failure in the functioning of any from three propulsion devices, RSUUV IA can by means of the remained operating propulsion device keeping a direct exchange or be able to steering in the horizontal and vertical planes.

2.1. If at the surface or underwater modes one of side propulsion devices failed, then it is necessary to disconnect the second working side propulsion device and to move forward during the work of an average propulsion device in a torpedo-shaped hull, or to provide necessary maneuvering in the horizontal and vertical planes by means of rudders.

2.2. At failure in the functioning of the average propeller the similar movement and maneuvering of "Shadow-2" is provided due to operation of two side propellers and rudders.

3. In all three torpedo-shaped hulls and on the deck of the top hull it is possible to install necessary research or other special equipment.

4. RSUUV IA "Shadow-2" can be deployed to the region of the mission without assistance and does not demand additional carriers.

In Fig. 2 it is presented preliminary configuration (based on calculated data) the energy unit [4, 27-30], a reserve of diesel fuel [4, 28-29] and rechargeable batteries [4, 31-32] and also ballast tanks.

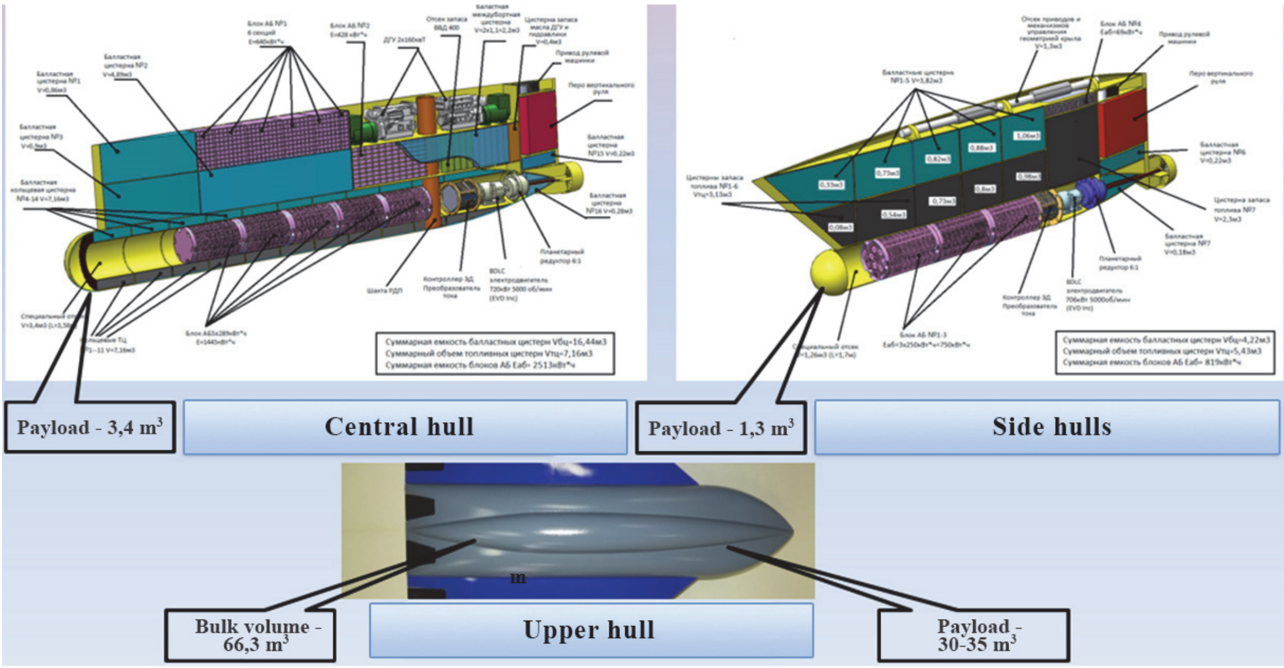


Figure 2 – Preliminary configuration the energy unit, reserve of diesel fuel and rechargeable batteries, and also ballast tanks

The top hull having the form of a wing, of RSUUV IA with the volume displacement of 66.4 m^3 remains almost completely free, and payload can be placed in it: research equipment and additional sources of energy supply [4].

In general as a result of development of the RSUUV IA concept, the multipurpose universal platform turned out which can be used at the solution of a wide range of tasks both in military, and in civil areas.

The system of robotic underwater seismic exploration in subglacial waters

RSUUV IA "Shadow" [3] was used as a prototype of the heavy hybrid underwater-surface robot - seismosource - the carrier of autonomous underwater vehicles group – seismographs when developing "The concept of robotic underwater seismic exploration in subglacial waters" [4]. The concept got the first award at a contest of the scientific, scientific and technical and innovative developments directed to studying of the Arctic and the continental shelf.

Due to the complexity of carrying out full-scale surveying by means of traditional methods (with using of vessels and towed seismographs) in the difficult ice conditions and storm activity, which typical for the waters interest of the Russian Federation, the problem of creation the surveying system, which having a minimum need for surface components, such as seismological vessels and the towed surface antennas is urgent. Thus, the distributed system of carrying out seismological survey consisting of the following components is offered:

– universal hybrid surface-underwater robot (carrier);

– small-sized underwater autonomous vehicles - carriers of seismometers with a possibility of group control;

– underwater autonomous vehicles with a possibility of sampling and analysis of samples of soil, water, and functioning as benthonic observation stations (or ensuring deployment and assembling such stations).

The vehicle - carrier has an opportunity to move both in underwater, and in the surface mode (including a complicated ice conditions) and has sufficient available power and the sizes to serve as a source, the carrier, and a command and navigation center for group of small-sized autonomous crafts performing functions of seismic detectors. This carrier is supplied with compressor installation therefore at the movement on a surface it can use the standard procedure of application of seismosources (pneumatic seismoguns). Under water the system of chemical gas generation can act as a source of gas of high pressure for work of pneumatic installations – the similar systems of domestic production provide necessary parameters of the gas environment and the required productivity.

As carrier of seismic detector and benthonic station it is supposed to use an underwater glider - the autonomous underwater vehicle set in motion by hydrodynamic forces due to change of buoyancy [33]. Generally it is the multiple-use autonomous underwater vehicle which moves in water space like the glider - with the minimum energy consumption and according to program. This craft has no "propeller" or "the traditional engine", and the movement is provided due to change of its buoyancy. It can be used as the universal platform - carrier of various tools for a research of any water areas of the World Ocean, including the Arctic zone where it can be used in the "subglacial" mode on "super big" autonomous time intervals with a possibility of "landing" on sea floor and works as the ground seismometer, and in hybrid option – the group movement in the generated swarm (network) of seismometers in the thickness of water. In addition it can carry out all these operations, being stationed at heavy hybrid underwater-surface robot - a seismosource - the carrier of autonomous underwater vehicles group – seismometers, both leaving it for carrying out measurements, and coming back for recharge and transfer of the accumulated data for processing.

Conclusion

In general the conducted researches showed that performance characteristics (including endurance and range) in an underwater, semisubmersible and surface mode of the vehicle "Shadow-2" can significantly exceed the same parameters of the best known foreign analogs - prototypes. In particular, considering that the volume displacement of the vehicle "Shadow 2" is 22% more, than displacement of American robotic trimaran "Sea hunter" of 138.5 tons, and at the equal capacity of power stations, the maximum speed "Shadow-2" in the underwater mode is 26.43 knots (98% of the maximum speed "Sea hunter"), at the surface "unobtrusive" mode or semisubmersible mode – 29.9 knots (110%) and at the "trimaran" mode - 34.15 knots (126%). Range of the surface- underwater vehicle "Shadow-2" in comparison with "Sea hunter" at the movement with a speed of 6 knots in an underwater mode approximately by 1.3 times, in the surface "unobtrusive" mode – by 1.74 times, and in surface "trimaran" mode – by 2.53 times more.

The multifunctionality of the surface-underwater vehicle with changed geometry of the hull is that it can perform all functions and tasks which are carried out and solved separately surface, semisubmersible and underwater unmanned civil and military vehicles now. RSUUV IA "Shadow-2" organically fits into the concept of multiagent information and measuring network as the carrier of small-sized AUUV with limited endurance or for acceleration of low-speed gliders's delivery to the region of the mission. It can be also used as docking station for AUUV, and considerable increase the accuracy of their positioning during long actions under water without a possibility of updates of their location on the GLONASS/JPS system including under ices. Application principles of glayder type marine robotic system were investigated in the concept of multiagent information and measuring network of dual purpose for illumination of a situation in the sea of Arctic zone, developed by group of authors of St. Petersburg State Marine Technical University under the general guide of professor V.A. Ryzhov [33].

The offered surface-underwater autonomous unmanned vehicle is capable to work effectively, practically, under any weather conditions. On still water and at small sea state it can move in surface and semisubmersible mode and also in the mode of operation of the diesel under a periscope. At heavy sea or in a difficult ice conditions this vehicle will move under water.

It should be noted that the given analytical calculations have approximate character and confirm only a basic possibility of creation such surface-underwater vehicle with changeable hull geometry, which efficiency considerably exceeding the best foreign analogs. Further development of the concept requires conducting of research for technology development of creation of the robotic surface-underwater unmanned vehicle of the increased autonomy with changeable hull geometry and to confirmation of its design characteristics, and design and development work on creation of a RSUUV IA prototype.

References

1. S.B. Kuzmin, A.U. Ipatov. Modern devices and technologies of observation of hydrological conditions in the Arctic Ocean. Oceanography and sea ice. A contribution of Russia in the international polar year 2007/08. pp. 7-22
2. G.V. Antsev, V.V. Kobylyansky. Research fleet of Russia. Whether there is a future? Sea management information systems. 2013 No. 2(3) pp. 40-54
3. "Research and development of conceptual shape of the perspective Robotic Underwater-Surface Unmanned Vehicle of the Increased Autonomy (RUSUV IA) with changeable geometry of the hull for search and tracking underwater objects in a long-range sea zone" (Code "Shadow"). Saint-Petersburg State Marine Technical University, GNIITS RT of the Ministry of Defence of the Russian Federation, Report on research, 2018.
4. Gaykovich B.A., Zanin V.Yu., Taradonov V.S., Kozhemyakin I.V., Tokarev M.Yu., Biryukov E.A. The concept of robotic underwater seismic exploration in subglacial water areas. The collection of works of winners of the international competition of the scientific, scientific and technical and innovative developments aimed at the development and development of the Arctic and the continental shelf, 2018, pp. 64-86.
5. ASW Continuous Trail Unmanned Vessel (ACTUV), Phase I, DARPA-BAA-10-43, 31 pp., http://www.fbo.gov, 2010.
6. Till Geoffrey. The automatic semisubmersible vehicle for ensuring anti-submarine operations of US Navy, Unmanned Vehicles, June/July, pp. 40-41, 42, 2010; Seapower, July, 2010, pp.18-19, 2010.
7. Pilotless anti-submarine complex of tracking ACTUV, http://bmpd.livejornal.com/, 2011.
8. American designers began to develop an unmanned underwater hunter SSK http://topwar.ru/, 2012.
9. Unmanned surface means of different function, Digest of Federal State Unitary Enterprise Krylov State Scientific Center, Issue 72, pp. 90-91, 2014
10. Project antisubmarine anti-submarine defense boat ACTUV (USA), http://topwar.ru/, 2015.
11. The American "submarines hunter" again in the sea, http://vpk.name/, 2016.
12. Tests robot ship "Sea Hunter" in the United States will last until autumn 2017, http://topwar.ru/, 2017.
13. ACTUV – Catalog of surface robotic vehicles, https://robotrends.ru, 7.02.2018.
14. Research and analysis of the main characteristics of the high-speed three-hull ships and vessels perspective projects and developments, the code "Trimarans", the scientific and technical report of Company Limited «Marine innovative technologies» under the St. Petersburg State Marine Technical University, 194 p., 2017.
15. Seagull Multi-Mission USV System, http://www.elbitsystems.com.
16. Elbit Systems unveils Seagull unmanned naval vessels, http://www.globes.co.il
17. Israel showed the robotic submarines hunter, http://flot.com/2016/Израиль4/
18. The Israeli Elbit Systems represents military multi-purpose robotic USV Seagull, http://robotrends.ru/pub/1606
19. B Israel showed the robotic boat for hunting for submarines, http://tech.onliner.by/2016/02/09/seagull
20. SENTRY: Reconfigurable USV SWATH. Patents Pending, http://www.julietmarine.com
21. Americans were engaged in arms of the supercavitating boat, http://www.membrana.ru, 18.01.2012.
22. Meet the underwater ghost ship: Stealth military machine can travel at high speed on water - then submerge in a 'supercavitation bubble' to hit similar speeds under the sea, http://www.daily-mail.co.uk/, 20.10.2016.
23. High speed surface craft and submersible craft: Patent US № 8683937 B2, опубл. 01.04.2014.
24. The reference book on the theory of the ship, under the editorship of V.F. Droblenkov, M., Military publishing house, 590 p., 1984.
25. Egorov S.K., Research of hull extension impact on performance of torpedoes and submersibles, Collection of scientific works "Fundamental and applied hydrophysics", No. 2(8), pp. 58-67, 2010.
26. Dronov B.F., Pyalov V.N., Introduction to architecture of submarines, St. Petersburg State Marine Technical University, St. Petersburg, 2014

27. Kostenko V.V., Mikhaylov D.N., Determination of parameters of the energy-power plant of the autonomous underwater vehicle on the assigned speed //IZVESTIYA SFedU. ENGINEERING SCIENCES . - 2013-№3 (140) pp. 70-73.
28. Brighenti, Parametric analysis of the configuration of autonomous underwater vehicles, IEEE. 1 Oceanic Eng. vol. 15. pp.179-188, 1990.
29. Diesel-generator Tontekpower installations. [Electronic resource] URL: http://www.tontekpower.com/diesel-generators (date of address 27.02.2017).
30. Reducers of offshore Tontekpower power plants. (Marine gearbox. [Electronic resource]) URL: http://www.tontekpower.com/marine-gearbox (date of address 27.02.2017).
31. YASA Motors. [Electronic resource] URL http://www.yasamotors.com/products/yasa-750 (date of address 20.02.2017).
32. iXBIT Live. Complex testing of various accumulators. 18650, 16650, 18500, 26650, AA, AAA. [Электронный ресурс] URL: http://www.ixbt.com/live/kirich/kompleksnoe-testirovanie-razlichnyh-akkumulyatorov-18650-16650-18500-26650-aa-aaa.html (date of address 20.02.2017).
33. Specifications of LIPO of the AA Portable Power Corp accumulators. (Category: LiFePO4/LiFeMnPO4 Batteries.) [Electronic resource] URL: http://www.batteryspace.com/LiFePO4/LiFeMnPO4-Batteries.aspx (date of address 27.02.2017).
34. Kozhemyakin I.V., Rozdestvenskiy K.V., Ryzhov V. A. Development of the technical platform of a global sea information and measuring system on the basis of glider tipe autonomous unmanned vehicle . The Russian innovative technologies for development of hydrocarbon resources of the continental shelf, 2016, pp. 91-108.

A.S. Shustov, A.E. Kutsko, S.V. Belov

COMPACT POSITIONING, DATE TRANSMISSION AND VOICE COMMUNICATION SYSTEM FOR UNDERWATER APPLICATIONS

Saint-Petersburg State Marine Technical University (SMTU), Saint-Petersburg, Russia
semte@semte.ru, akutsko@mail.ru

Abstract

Underwater system being designed by SMTU for positioning robots and divers, voice communications and remote control, digital data transmission is demonstrated. The system consists of multifunctional surface and diver stations, compact beacon transponders and high-speed digital modems. One of the key features of the system is the ability to underwater searching with surface and divers stations, independently of each other. The structural diagrams of major system components and description of used algorithms are presented. The analysis of the factors affecting positioning accuracy are made. The results of prototype field tests are shown in condition of shallow water.

Keywords: hydroacoustic positioning, uninhabited underwater vehicle, hydroacoustic communication, high-speed hydroacoustic modem, beacon responder, emergency.

Hydroacoustic equipment is used by divers and underwater robots (autonomous underwater vehicle – AUV) for technical and rescue operations must provide:

– voice (telephone) or signal (telegraph) communication between divers and support vessels;

– positioning of divers (AUV) not only “from above” (with support vessel), but also with other divers “from below”;

– determining the location of the emergency, both by the diver and without him, by remotely turning on his emergency beacon;

– digital information communication with transfer of various kinds of messages, commands, information from sensors, photos, etc .;

– efficiency and reliability management by involved underwater participants;

– continuous recording of communication on hydroacoustic channel with the playback for analyzing actions.

The key point of this system is to ensure maximum safety of divers.

At present, there is no equipment that provides these functions as part of a single set on the international market. Practically at all known diving underwater communication stations [1], [2] neither emergency beacons nor means of underwater navigation and telemetry are provided. In positioning systems [3], including those that combining positioning with data transmission [4], there are no voice communication channels and no interaction with divers is expected during mission performance. The development of a unified positioning system, communication, telemetry, digital data transmission minimizes consumer costs, reduces size, power consumption, while improving usability, safety of work and therefore is, in our opinion, relevant.

From the beginning of presented work, we came to the conclusion, that it is not possible to use technical devices available on the world market for building a system for several reasons:

– the inability to make changes to the software;

– constructive inconsistency with the overall design of the development;

– high price of products and their after-sales service.

For this reason, it was decided to independently develop all the components of the system. When developing the above multifunctional underwater system (referred as “Poisk-01”), the following tasks were set and solved:

– providing hydroacoustic voice and signal communication divers among themselves and with the surface part of the system;

– positioning in automatic mode on the display of the boat station of the location of the AUV (more precisely, the beacons installed on them), divers equipped with diving hydroacoustic stations, in coordinates associated with the boat and on the tablet - on the map (output geographic coordinate), with the possibility of transmitting information from divers in the form of a diver number (AUV) and an emergency signal initiated by a diver from the diving station's control panel;

– remote activation of any emergency the diver’s beacon, both from the remote control of the “Poisk-01” surface part and other divers;

– determining the location of the emergency beacon by a diver;

– digital data transmission via underwater communication channel.

The composition of the developed system is shown in Figure 1.

The system consists of surface and underwater parts.

The surface part is located on the support vessel (or on the pier, platform, etc.) and consists of a communication console, control, display, a lowering device with an antenna unit (inside the unit is an electronic module of preliminary signal processing), a tablet, a personal computer necessary (Fig. 1). The underwater part consists of multifunctional diving underwater stations and responder beacons. The underwater station is designed for use with full-face masks AGA type and consists of an electronic unit, a communication headset with a control panel and an antenna base (Fig. 2). The stations are controlled from a two-button remote, combined with a microphone node, which is inserted into a special slot of the AGA mask. Hydroacoustic antennas are placed on the diver's breathing apparatus.

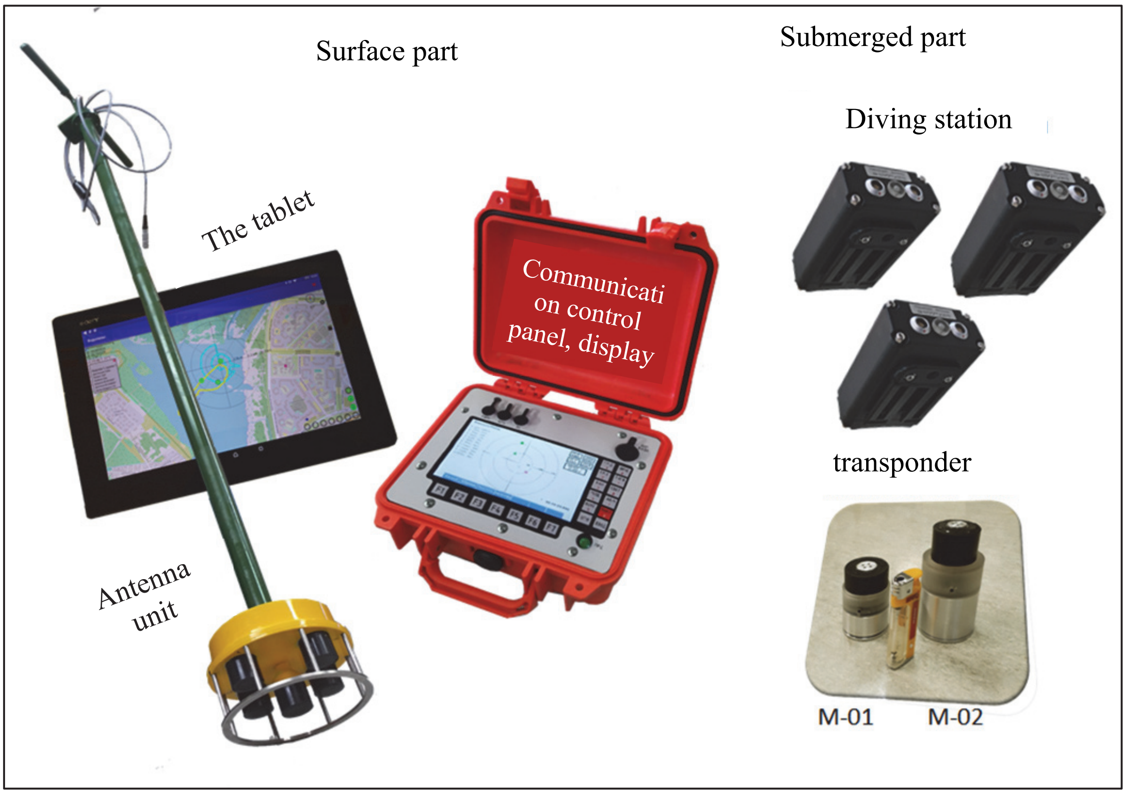


Figure 1- The composition of the prototype of the system "Poisk-01"

Responder beacons are equipped with depth sensors, have a battery or external power supply and are located on the AUV. Beacon can be combined with underwater modem

The generalized block diagram of the surface part of the system (located on the support vessel) is shown in Figure 3 and works as follows.

In the positioning mode diving stations or responder beacons:

– From the control panel, the numbers of the stations or the military formation (the "subscribers") are set.

– The processor forms the teams of sequential switching on (hereinafter referred to as the "request") of subscribers. The request signal through the switch goes to the power amplifier and from its output to the radiating antenna. After the request signal is emitted, the switch disconnects the from the radiating antenna, which is also used to receive / emit voice communication and digital communication signals.

The subscriber's response signal is sent to the antenna unit and from it to the processor where the preprocessing is performed, the distance between request and response is determined, the heading angle is based on the measured signal delays on the antennas, taking into account the roll / pitch sensor data. By the known course angle and distance, the display shows the location of the subscriber relative to the antenna unit. Taking into account the GPS receiver (compass) data, the processor calculates the subscriber's position in global coordinates and transmits them via Bluetooth protocol to the tablet, where it indicates a point on the map of the work area.

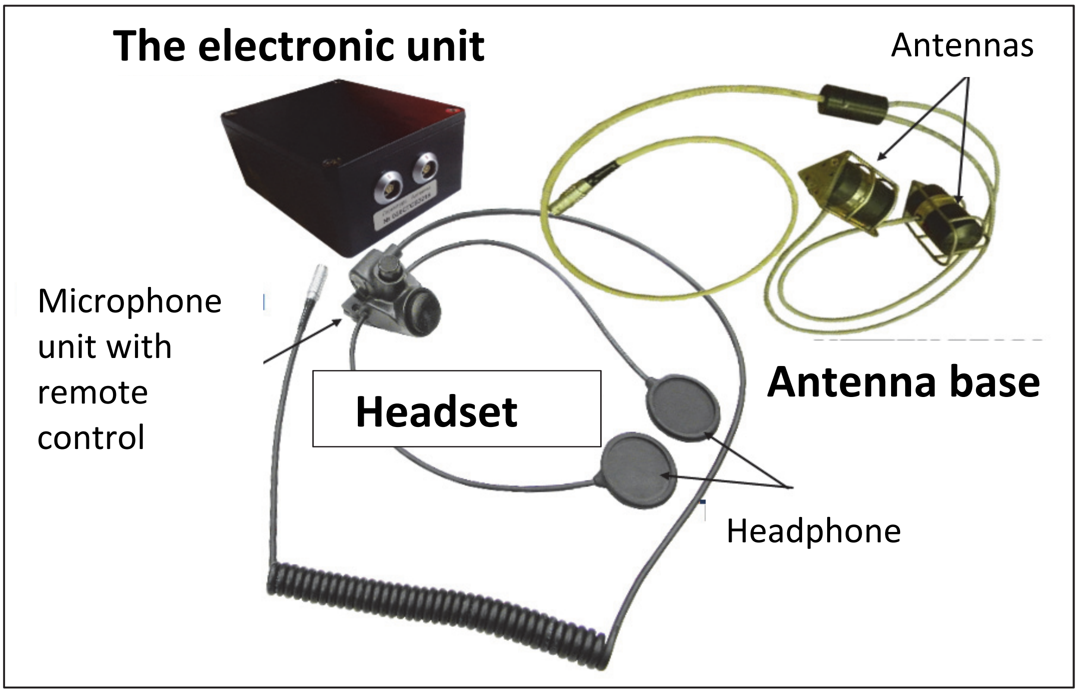


Figure 2 – The composition of the diving station hydroacustic

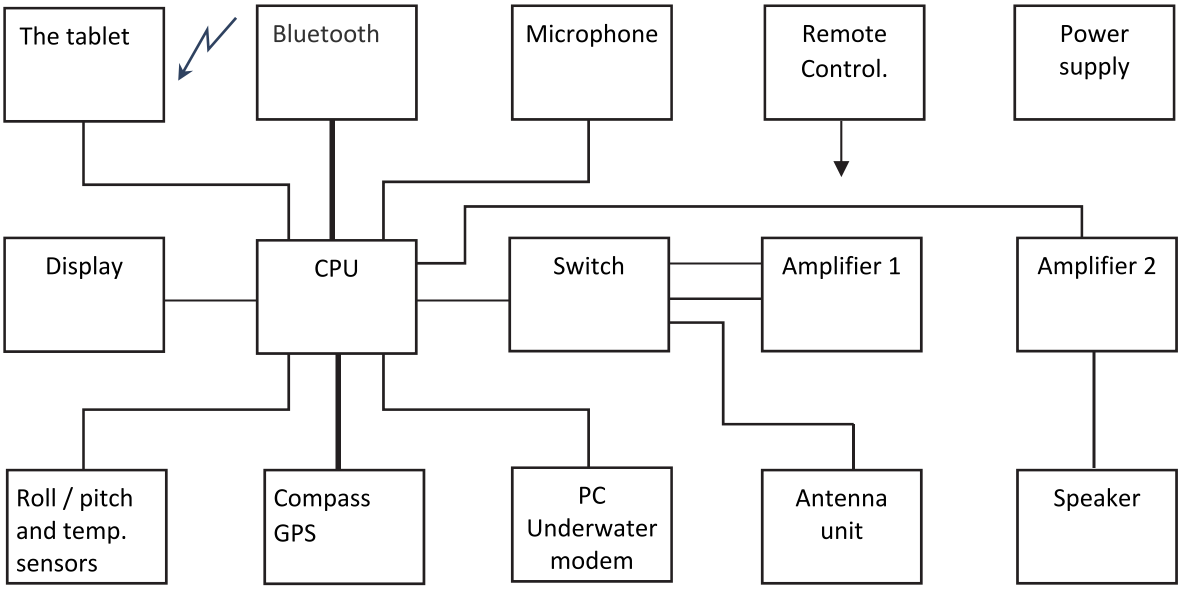


Figure 3 – Block diagram of the surface part of the system

Obtained on the display or tablet points of each subscriber are interconnected, forming a trajectory of its movement. The underwater beacon forms a response parcel with information about the subscriber number, the depth of its location, and the presence of an emergency.

In the voice communication mode: the voice transmission mode is initiated by a button or voice. The speech signal from the microfilm microphone enters the processor, where it is converted into a single-sided (SSB) signal, which is transmitted through the amplifier to the receiving-emitting antenna and transmitted to water. When receiving a voice information signals from the antenna unit through the switch enter the processor, where they are decoded (converted into voice signals) and through the power amplifier arrive at the speaker.

In the digital communication mode: the modem software installed on the PC generates digital communication signals that go through the processor to the switch, then to the power amplifier and, from its output, to the radiating antenna of the antenna unit. After the transmission signal is emitted, the switch disconnects the amplifier from the radiating antenna, the signals from the answering modem from it through the switch and the processor arrive at the PC (modem).

The peculiarity of the diver station is its multifunctionality and a large number of commands calling certain functions. The stations provide: telephone and telegraph communication (2 commands), remote control by beacons with determination of distance and bearing (15 commands), voice activated transmission mode (2 commands), direction finder control (2 commands), control of its own emergency beacon (2 commands), transmission power control (3 commands), volume control (5 commands), noise reduction control (2 commands). The control is performed from the 2-button remote control by combinations of the number of button presses, which requires memorization of the combinations and leads to the presence of control errors. To improve the reliability of control, a speaker-independent speech recognition system with a limited vocabulary (about 50 commands) based on chalk-core analysis and hidden Markov models is built into the aircraft. At present, the algorithm provides a probability of correct recognition in air of about 95% and about 80% in a full-face mask, regardless of the type of breathing apparatus. As the system asks for the correctness of the command, control errors are practically eliminated.

The development of a modem is based on OFDM modulation and was carried out practically without a real prototype, since the developers do not disclose the technological and algorithmic characteristics of the underwater modems available on the market, and the well-known radio communication solutions cannot be applied due to significantly different characteristics of the electromagnetic hydroacoustic fields [5].

The block diagram of the transmitter of the OFDM modem is presented in Fig. 4. The input data is converted into separate frames containing the data itself, the service information and the code (CRC) to verify the integrity of the information in the receiver. In the scrambler, bitwise summation of frame data with a pseudo-random sequence is performed to reduce the peak-factor of transmission and provide information cryptographic protection. Convolutional coding (CC) is intended to correct errors in the receiver. To combat the error packing in fading and impulse noise in the communication channel, bit interleaving is used according to a known pseudo-random law.

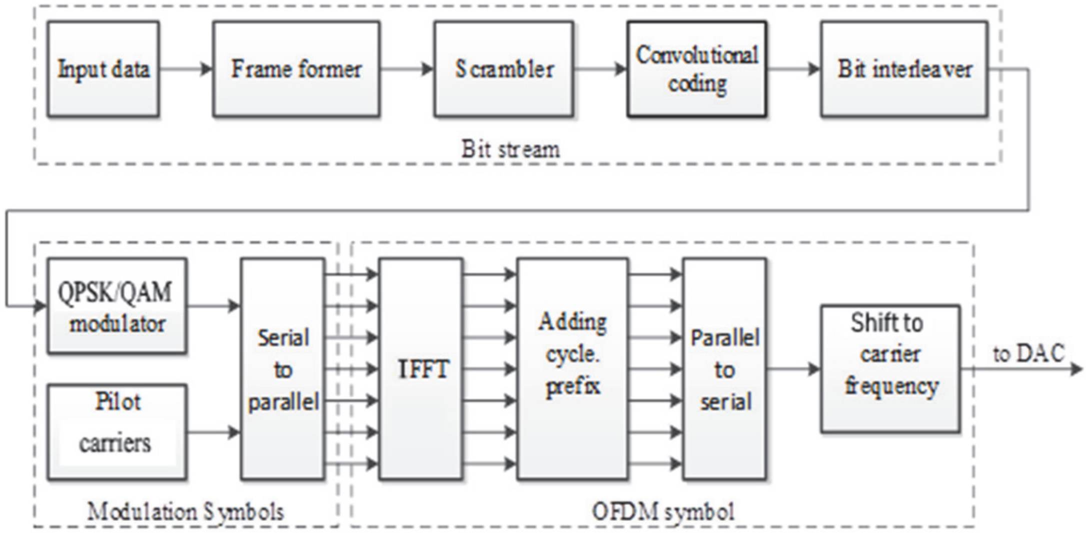


Figure 4 – Transmitter block diagram

In the QPSK / QAM modulator block, quadrature phase shift keying (QPSK) or 16-position quadrature amplitude shift keying (QAM-16) is performed. The quadrature information symbols are distributed over the subcarriers and are complemented by well-known pilot symbols, by which the amplitude-frequency characteristic in the receiver is corrected. The inverse Fourier transform in the IFFT block translates symbols from the frequency domain into the time domain, after which its end is written at the beginning of each OFDM symbol to create a cyclic prefix. The obtained time samples in a complex form at zero frequency are converted into real samples at a given carrier frequency using an interpolator and a quadrature local oscillator. The generated data is transmitted via a digital-to-analog converter (DAC), power amplifier to the underwater antenna (not shown in Fig. 4).

The block diagram of the receiving part of the modem is presented in Fig. 5. The signal from the antenna through the receiving amplifier and the ADC (not shown in Fig. 5) goes to the transfer unit at zero frequency, from where the complex samples go to the symbol synchronization block, where the start of the data frame transmission is determined by the synchronization symbol (SS) Y, which is in the time domain threefold repetition of one pseudo-random sequence $Y = [\ Y_{lp}\ ;\ Y_{cp}\ ;\ Y_{rp}\ ]$,, where Y_{lp}, Y_{cp}, Y_{rp} - left, central and right parts of the SS For each complex reference, scalar values of cross-correlation are calculated: $J_{lp} = Y_{lp} * Y_{cp}'$ and $J_{rp} = Y_{cp} * Y_{rp}'$, after which the metric J is determined using the expression:

$$J = \frac{J_{lp}+J_{rp}}{Y*Y'}, \quad (1)$$

where Y_{cp}', Y_{rp}', Y' are the values of the complex conjugation of the readings of the left side, the right side, and the full SS, respectively. Based on the comparison of the J metric with a threshold, a decision is made to detect SS.

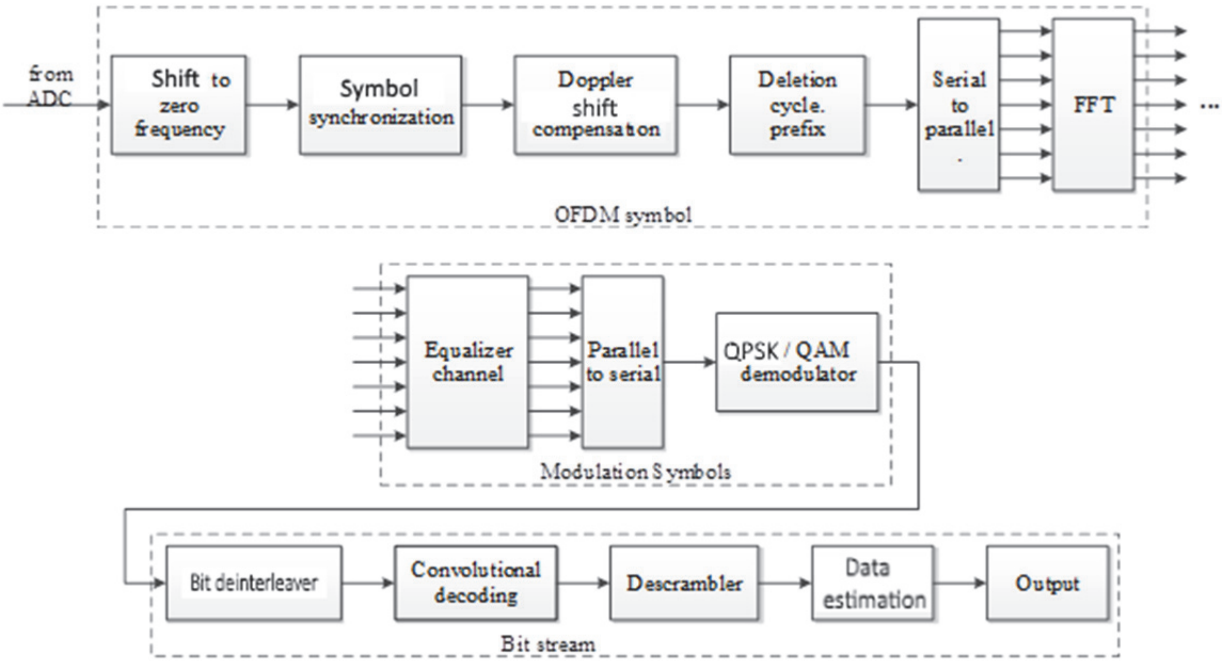


Figure 5 – The structural scheme of the receiver

In the Doppler shift compensation unit, frame-by-frame (over the measured SS period) and symbol-by-cell (over subcarriers with zero energy) estimates of the magnitude of the frequency offset (due to the Doppler effect and frequency drift at the receiver) are performed. The method used is presented in [6]. Next, the cyclic prefix is removed from the frame and is performed using the direct FFT to transfer it from the time domain to the frequency domain. The channel equalizer block compensates for distortions that occur when a signal passes through a multipath communication channel. The channel response estimate is calculated using the MMSE algorithm for known pilot symbols and interpolated into information symbols [7].

After equalizing the channel frequency response, the data symbols are converted by a QPSK / QAM demodulator into a sequence of soft bits quantized into 8 levels. The bit deinterleaver performs the reverse permutation of soft bits, after which the convolutional decoding block performs decoding of the obtained sequence using the soft decision Viterbi algorithm according to the maximum likelihood criterion [8]. The descrambler block performs the functions inverse to the transmitting part of the scrambler. In the block for estimating the received data, the CRC is checked and the numbers of neighboring frames are compared, on the basis of which a decision is made about the correct reception of the frame.

The frame structure of the transmitted information is shown in Fig. 6

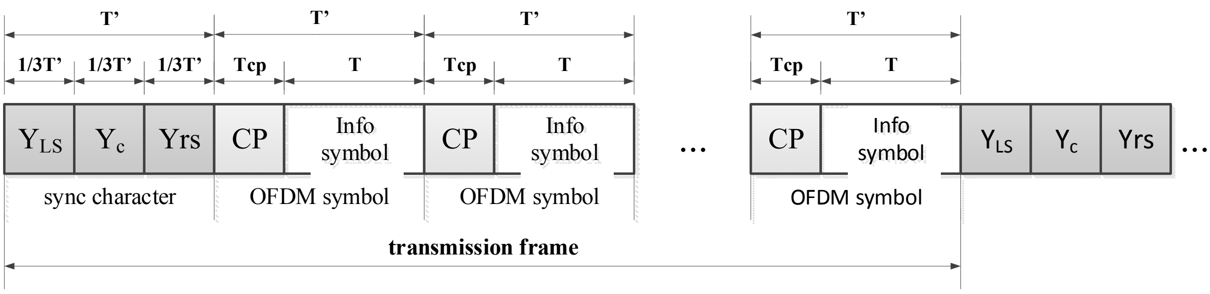


Figure 6 – Transmitter block diagram

The modem provides 3 modes (A, B, C) for use in various interfering situations. Mode A is intended for transmission channels having frequency and time selectivity with propagation delays; mode B - similar to mode A, but with an increased Doppler effect; mode B is similar to mode B, but with significant propagation delays and a significant Doppler effect. The main signal parameters of the modem in these modes are presented in Table. 1.

Table 1. The main parameters of the modem

Mode	**Number of subcarriers**	**Number of pilots**	**Distance between subcarriers Δf = 1 / T, Hz**	**Duration of information symbol T, ms**	**The duration of the CP Tcp, ms**	**Total character duration *T'*, ms**	**The number of characters in the frame N_s**
A	854	106	11,7	85,3	21,3	106,6	15
B	584	97	17,1	58,6	21,3	79,9	20
C	373	93	26,8	37,3	29,3	66,6	24

The estimated modem transmission rate for various modes and types of modulation is presented in Table 2.

Table 2. Modem data rates

Mode	**Modulation**	**Speed, kbps**		
		CC 1/3	CC 1/2	CCs 2/3
A	QPSK	3,7	5,5	7,3
	QAM-16	7,3	11,0	14,7
B	QPSK	3,0	4,6	6,1
	QAM-16	6,1	9,2	12,2
C	QPSK	2,1	3,2	4,3
	QAM-16	4,3	6,4	8,5

Comparative tests of the developed OFDM modem with the Evologics S2C 42/65 modem were carried out in two water areas with different hydrological parameters: Lakhta Spill with a depth range of 4-6 m (St. Petersburg) and the Gulf of Finland with a depth range of 10-15 m (near Primorsk city). The operation of the OFDM modem was provided using software running on a PC: the physical and channel levels were implemented in the C / C ++ programming language, the transport level in Matlab R2017a. The modem operation parameters are transmission modes (A, B, C), convolutional coding rate (1/3, 1/2, 2/3) and modulation type (QPSK, QAM-16) were determined by the software automatically based on the evaluation test packet. Digital-analog and analog-digital conversions of received and emitted signals were performed using an external sound card Creative EMU 0204 connected to a PC via USB interface. The appearance of a set of OFDM modems is shown in Fig. 7

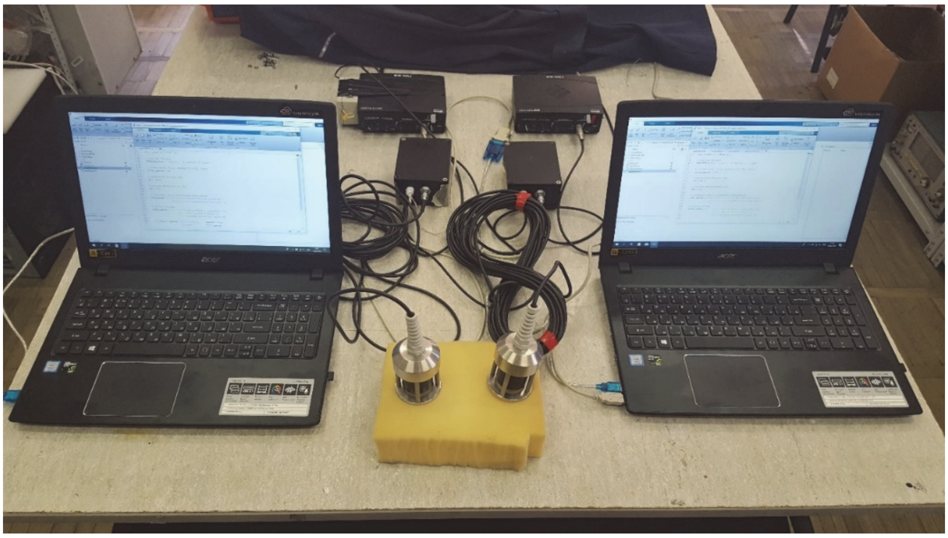


Figure 7 – The appearance of a set of OFDM modems

Typical changes of impulse response (IR) for real communication channel are presented in fig. 8.

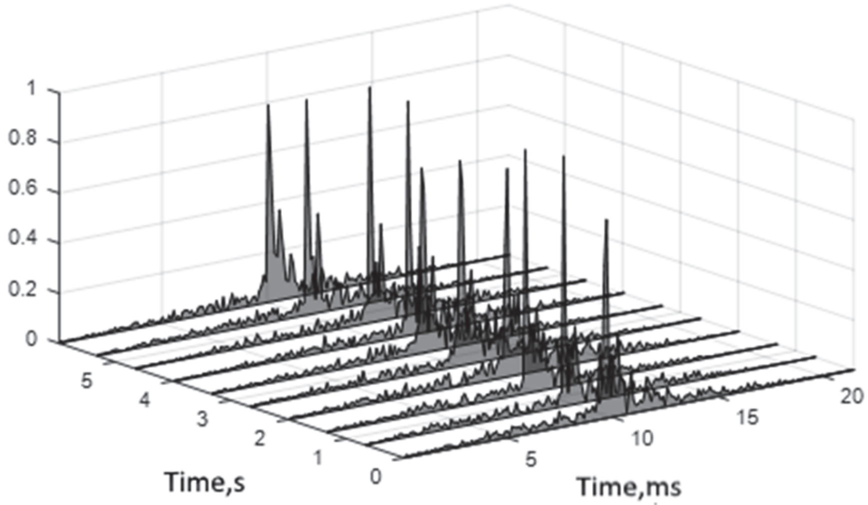


Figure 8 – Typical IR of underwater channel

Table 3 presents the values of the measured average data transfer rates (including all repeated requests) in the Lakhta spill, and Table 4 presents transfer rates in the Gulf of Finland area.

Table 3. Field test results, Lakhta Spill

Distance	**Average transfer rate, bit / s**				**Notes**
	Modem OFDM		**Modem S2CR 42/65**		S2CR 42/65 - 23 kHz bandwidth; OFDM - 10 kHz bandwidth.
Direction	1-2	2-1	1-2	2-1	
200 m	2960	2270	2920	3040	
600 m	3140	5820	3900	1750	
800 m	Not tested	3180	No connection	No connection	S2CR 42/65 - directional antenna ± 50^0; OFDM - non-directional antenna.
1100 m	270	302	No connection	No connection	

Table 4. Field test results, Gulf of Finland

Distance	Average transfer rate, bit / s				Notes
	Modem OFDM		Modem S2CR 42/65		
Direction	1-2	2-1	1-2	2-1	S2CR 42/65 - 23 kHz bandwidth;
100 m	2126	2048	2735	2624	OFDM - 10 kHz bandwidth.
200 m	8829	8790	643	No connection	
600 m	8835	4826	No connection	No connection	S2CR 42/65 - directional antenna ± 50^0
1000 m	No connection	No connection	No connection	No connection	OFDM - non-directional antenna

The underwater OFDM modem was realized as a C / C ++ program for the BeagleBone Black hardware platform [9]. The board is equipped with an ARM processor with a clock frequency of 1 GHz (2000 DMIPS). Board dimensions are 86x53x20 mm and average power consumption is 1 W.

Despite the identified advantages of the OFDM modem we will continue field tests to determine the characteristics of the modem in different areas to collect statistics.

At present, the «Poisk-01» system has the following characteristics:

– communication / positioning range — not less than 1 km;
– the error in determining the bearing - not more than 2 degrees;
– the error in determining the distance - not more than 3% of the distance;
– the number of frequency communication channels - 8;
– frequency range of voice communication from 24,700 to 33,200 Hz;
– voice modulation type - SSB (single sideband, with suppression of one sideband and carrier frequency);
– number of simultaneously serviced divers (UAV) - not less than 10;
– speed of relative movement of divers and “Poisk-01” - up to 3 m / s;
– continuous operation time - at least 8 hours (at the rate of the transmission 1 time per 10 s);
– display - on the surface part console and the tablet with Android OS;
– communication protocol with the tablet – Bluetooth;
– range of communication with the tablet – at least 10 m;
– the number of antennas in the “Poisk-01” - 5 antenna unit;
– type of positioning signals – phased manipulated pseudo-random sequences of 255 chips;
– bandwidth of positioning signals from 12 to 36 kHz;
– bandwidth of the digital communication channel from 19 to 29 kHz;
– data transfer rate - up to 14 kbps;
– acoustic transmission power - up to 10 W;
– number of HA antennas - 5;
– dimensions - no more than 500x300x150 mm.

The diving stations included in the system have the following characteristics:

– communication range - at least 1 km;
– working depth - up to 50 m;
– the number of frequency communication channels -8;
– modulation type - SSB;
– type of telemetry signals - PM PN;
– the error of determining the distance - no more than 3%;
– the error in determining the bearing - no more than 5 degrees;
– acoustic transmission power - up to 10 W;

– number of HA antennas - 2;
– dimensions of electronic block - no more than 250x120x40 mm.

The compact hydroacoustic beacon (without a modem) has the following characteristics:

– range of action - not less than 1 km;
– working depth - up to 50 m;
– type of signals - PM PN;
– time in the "request-response" mode - at least 6 hours;
– working time in standby mode - at least 24 hours;
– working time in sleep mode – up to 24 days;
– dimensions D = 60mm, H = 100 mm.

Positive results of field tests will allow developing a more technically advanced hydroacoustic system for positioning of underwater vehicles and divers, control, voice communications, digital data transmission in the near future. The system will have the following general characteristics:

– range - 5 km;
– bearing error - 1 degree;
– working depths - 500 m;
– number of positioned UAV – 15;
– data transfer rate - 5 kbps;
– vocabulary voice command - 32 words;
– probability of command recognition in the air - 95%;
– probability of command recognition in a full face mask - 90%.

References

1. https://www.oceantechnologysystems.com/product-category/wireless/.
2. http://oceanreefgroup.mybigcommerce.com/gsm-dc-wireless-communication-system-w-nacs/.
3. https://www.ixblue.com/products/usbl-positioning-systems.
4. https://evologics.de/acoustic-modem/42-65/r-serie.
5. Bakulin M.G., Kreindelin V.B., Shlyuma A.M., Shumov A.P., 2015, The technology of OFDM, Moscow, Telecom, 351 pp.
6. Li B., Zhou S., Stojanovic M., 2007, Non-uniform Doppler compensation for zero-padded OFDM over fast-varying underwater acoustic channels, OCEANS, pp. 1-6.
7. Cho Y., Kim J., Yang W., Kang G., 2010, MIMO-OFDM Wireless Communications with MATLAB, Wiley, 544 pp.
8. Morelos-Zaragoza R., 2006, The art of error correcting codes. Methods, algorithms, usage. Moscow, Technospere, 320 pp.
9. BeagleBone Black, https://www.beagleboard.org.

S.Yu. Pribylov, V.V. Sergeev, V.N. Karpov, V.A. Sokolov

FEATURES OF CONSTRUCTION OF ACTIVE VISION SYSTEMS FOR AUTONOMOUS UNDERWATER VEHICLES

JSC "Research Institute of television", Saint-Petersburg, Russia
npk-62ypr@niitv.ru

Abstract

The article discusses the features of the construction of an active underwater vision system for Autonomous unmanned underwater vehicles, provides recommendations for the selection of the light-signal Converter and illumination systems and their placement to reduce the influence of backscatter interference and increase the visibility range under water. The block diagram and the main technical characteristics of APV for ANPA are given.

Keywords: autonomous and remote-controlled unmanned underwater vehicles, underwater television, solid-state imaging devices, underwater visibility, led illumination system.

Equipment of Autonomous or remote-controlled unmanned underwater vehicles to operate at different depths of the dive underwater vision systems allows to expand the range of scientific and applied problems of the development of the Oceans, to successfully perform search and rescue operations, to solve problems of defense value. Therefore, at present, in all countries that occupy a leading position in the field of marine technology, special attention is paid to the creation of Autonomous unmanned underwater vehicles (ANPA), while the direction of robotics using ANPA is rapidly progressing.

In the foreign literature of ANPA generally klassificeret to limit the working depth of the dive and the weight and size characteristics – portable AUV (weight up to 20 kg), mini AUV (20-100 kg), small AUV (100-500 kg), average AUV (500-2000 kg) and large AUV (>2000 kg) [1].

When designing the underwater vision system for the ANPA, depending on the class of the ANPA and the tasks to be solved, respectively, there are requirements for the range of visibility, resolution and quality of the video images, as well as requirements for weight and size characteristics, for the working depth of immersion and power consumption, since it is necessary to provide a predetermined battery life, while the capacity of the batteries in the ANPA is limited.

Consider the features of the underwater vision system for the ANPA with these requirements.

Due to the lack or complete absence of natural light at working depths, the underwater vision system in addition to the TV camera should have an external illumination system that provides the necessary level of illumination of the object of observation to obtain the necessary quality of photo and video materials. Given this fact, a system of underwater vision are further referred to as active system for underwater vision (ASPV).

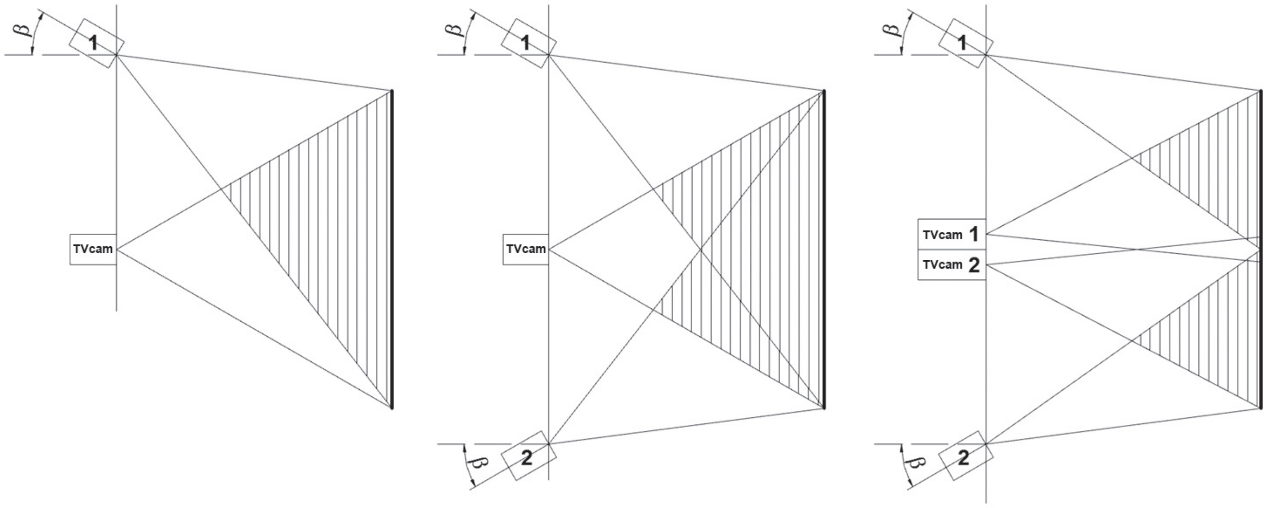
In General, any object that is in the water is perceived by vision or television system only if it is different in brightness or color from the background.

As the observation distance increases, the contrast of the object decreases not only due to the weakening effect of water, but also due to the phenomenon of "smearing" the boundaries associated with the scattering of light rays reflected from the surface of the object, as well as due to the effect of volumetric energy scattering during the passage of the light beam through the water medium. Volumetric scattering is perceived by the optical receiver as the glow of the water itself (light haze) and is called backscattering interference (PORE), which greatly reduces the brightness contrast of the object of observation [2,3].

Thus, in order to increase the range of observation in sea water, it is necessary to overcome both the attenuation of radiation in water and to reduce PORES in the construction of underwater lighting schemes. Note that an increase in the power of light sources to compensate for the attenuation of radiation in the water is necessary, but this does not guarantee an increase in the range of visibility, since with an increase in the power of radiation, the veiling brightness of the PORES simultaneously increases.

Spatial separation by moving the light source away from the TV camera at a certain angle (β) relative to the optical axis of the lens, depending on the shooting distance, is the most accessible way to reduce the illuminated volume of water (this area is shaded in figure 1).

It is for this reason that the examples of diversity are widely used in the ANPA. It should be noted that the size of the separation also depends on the layout of the equipment in the ANPA, so the real angle of separation in the design may differ from the optimal. With the use of two or more light sources, you can get more illumination and more uniform illumination of the subject, if there is a constructive possibility of placing several light sources on the ANPA and sufficient battery capacity. It is possible to implement, for example, in ASPV for medium and large ANPA.



a) one light source b) 2 light sources c) effect reduction method THEN, using multiple TV cameras

Figure 1 – Examples of light source and TV camera separation to reduce PORES

In addition to the known methods of separation, it is possible to propose a method to reduce the effect of PORES using several TV cameras with narrow angles of view and with consistent angles of light sources (see figure 1, b), and the observation of the overall image from all cameras to provide crosslinking in split-screen mode. This method can be implemented in ASPV, for example, for medium and large ANPA and thus increase the range of visibility.

Let us now consider the light sources to implement ASPV for AUV.

Analysis of the main lighting characteristics of modern light emitters shows that the most promising are led light sources [4,5].

JSC "research Institute of television" has experience in the creation of underwater lighting devices using led emitters for inhabited underwater vehicles for various purposes. For example, developed and widely used device underwater light TM-1245 [6] on the led matrix with a power of 100 W, the axial luminous intensity of 4000 CD, the radiation angle in the water of 60 degrees, working depth – 3000 m. The protective glass of the light device is made in the form of a lens, which allowed to concentrate the light flux within a given angle of radiation, and by reducing the size of the protective glass – lens, including its thickness, provided the necessary strength under the influence of external hydrostatic pressure, characteristic of the operating conditions of the underwater light device.

Taking into account the above mentioned advantages and on the basis of the experience available in JSC "research Institute of television", in the near future it is advisable to use led light sources in the design of ASPV for ANPA.

As mentioned above, the range of observation and, accordingly, the quality of the images obtained in underwater vision systems is significantly affected by the optical characteristics of the aqueous medium, where due to the processes of scattering and selective absorption of light in the water, the quality of the images obtained deteriorates. This is manifested in the weakening of the brightness and reducing the contrast of the image, distortion of its spatial structure, loss of information about the color of the object, the relative increase in the noise component of the image and the limitation of the visibility range [2].

Thus, from the point of view of increasing the range of visibility in the aquatic environment, the photodetector of the underwater vision system must have a high contrast sensitivity, and for the recognition of small objects and high resolution.

Within the linear characteristic of the light-signal transformation, the estimation of the threshold contrast of the imaging system is determined by the following formula [7]:

$$K_{\text{пор}} = \frac{\Psi_{\text{пор}}}{\sqrt{N_{\text{э}}}} \tag{1}$$

That is, the potential value of the threshold contrast is limited by the signal-to-noise ratio and the cumulative capability of the photodetector element (its capacity or the limit number of Ne photoelectrons).

Real CCD and CMOS photodetectors have their own noise, which worsens the potential sensitivity and signal/noise of the imaging system, which should be taken into account when choosing a photoconverter for operation at low light levels.

The paper [3] provides recommendations for the selection of photosensitive matrices for underwater vision systems based on the method of objective comparison using the standard EMVA1288 (European Machine Vision Association).

Absolute threshold of sensitivity (the number of photons that are required to achieve the noise level. The lower the threshold, the more sensitive the camera is and the less illumination is required to generate a useful signal) depends on both the signal level (e-), equal to the product of the number of photons (NF) per unit area of the photosensitive element (photon/μm2), the area of the photosensitive element (μm2) and quantum efficiency, and the magnitude of the intrinsic noise of the matrix (e-). The number of photons NF is determined by the energy collected by the unit area of the photosensitive element during the exposure, divided by the photon energy, that is, the number of photons NF includes the exposure time [8].

Based on the recommendations of [3], it is shown that for underwater vision systems it is necessary to choose a more sensitive photodetector with a lower absolute sensitivity threshold, a better signal/noise ratio, a large dynamic range and a large number of elements. This will allow for the same accumulation time to provide a system of underwater vision longer range of vision and distinguish between lower contrast and small objects with better quality.

Note that the isolation of the useful signal from the noise by increasing the accumulation time, as a rule, is not possible for underwater vision systems installed on moving underwater vehicles due to the appearance of "blur" images.

Among the photodetector matrices considered in [3], the best parameters are IMX249 matrices, the element size is 5.86×5.86 μm, the number of elements is 1920×1200 and IMX265, the element size is 3.45×3.45 μm, the number of elements is 2048×1536, which is preferable for use in underwater vision systems. Using research data, TV cameras on the matrix IMX249 and IMX265 [9,10], the best relationship of s/W has a TV camera on the matrix IMX249 - 45 dB, and the matrix IMX265 - 40 dB, it is possible to get the best results when observing low contrast objects, and it was confirmed by our test cameras at extreme levels of illumination on the object.

The disadvantage of TV cameras on the imx249 matrix is that the matrix has an optical format 1/1.2", so the size of the lens is large enough, in addition, the lens with this format from manufacturers is small and it will be difficult to choose a lens with the desired focal length.

For underwater vision systems of small-sized ANPA, where in addition to high lighting characteristics, small dimensions of TV cameras are required, it is advisable to use matrices with a smaller optical format, for which the required lenses will have smaller dimensions and weight. In particular, the TV camera on the IMX265 matrix slightly loses to the TV camera on the IMX249 matrix in relation to C/W [9,10], but has an optical format 1/1.8". With this optical format, there is a wide range of lenses and, in our opinion, for small-sized ANPA it is advisable to use TV cameras with the imx265 matrix.

Note that the matrix IMX249 and IMX265 are available in color format. Color matrices have a lower contrast sensitivity compared to black-and-white matrices [11], so the range of visibility when using color cameras is less. Color TV cameras are used in cases where information about the color of the object of observation is required, and in order to obtain a good color reproduction, the object of observation should be as close as possible to the optical receiver, since water differently absorbs color components in the radiation spectrum, and the degree of absorption depends on the distance of observation.

In General, when choosing the observation distance for underwater vision systems, it is necessary to take into account that when forming an image, light radiation passes a double distance - from the light source to the object and from the object to the radiation receiver.

In conclusion, we present the structural scheme (figure 2) and the algorithm of ASPV with one illuminator, developed in collaboration with JSC "Concern "NGO "Aurora" to create a small-sized ANPA with photo-video.

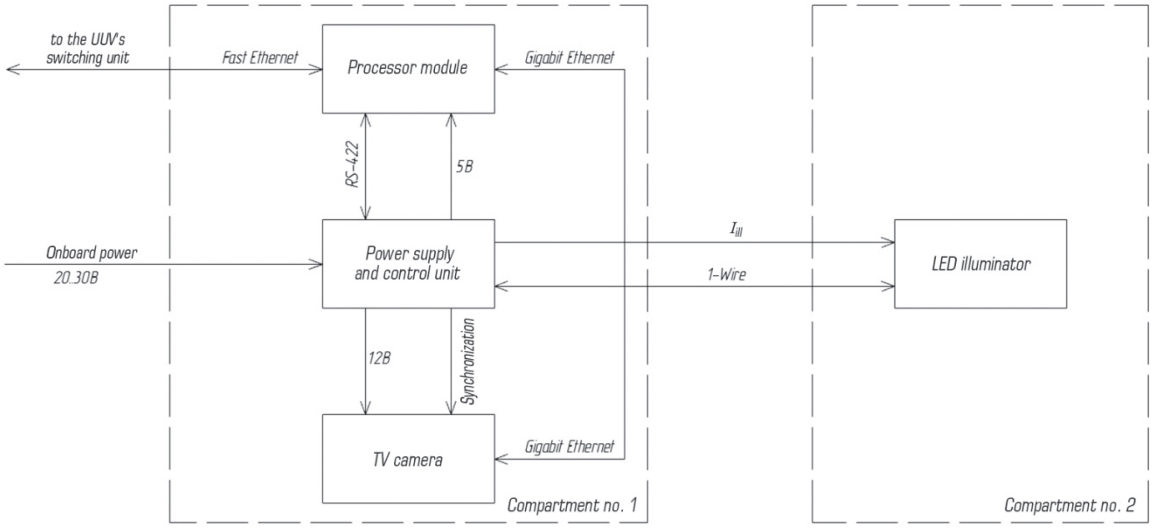


Figure 2 – Block diagram of APV with one illuminator

Characteristics of ANPA:

– maximum working depth of immersion - 1000 m;
– speed during photo and video shooting - 1.5-3 knots;
– supply voltage range - 20...30 Volts;
– outer diameter of the compartments - 200 mm;
– internal (light) diameter - 160 mm.

ASPV is located in two adjacent compartments of the ANPA (figure 3).

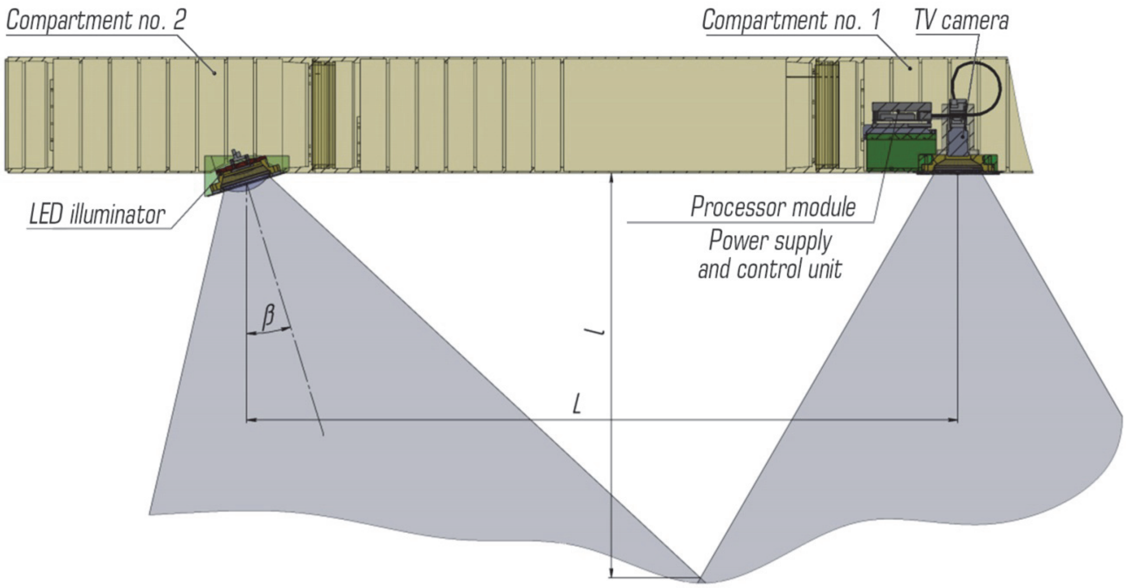


Figure 3 – Layout ASPV in AUV

In the compartment №1 there is a TV camera, power Supply and control unit, Processor module. In the compartment №2 - led Illuminator.

The main technical characteristics of the components of ASPV are given in table 1.

The algorithm works ASPV composed of ANPA:

Software control unit (BPU) ANPA transmits to the Processor module (PM) ASPV interface FastEthernet command mode - "Photography", "Video" and service information - "date and Time", "speed ANPA", "Distance to the bottom", "Geographical coordinates".

Software control unit (BPU) ANPA transmits to the Processor module (PM) ASPV interface FastEthernet command mode - "Photography", "Video" and service information - "date and Time", "speed ANPA", "Distance to the bottom", "Geographical coordinates".

In the photo shooting mode, PM and the controller of the power supply and control unit ASPV generates current pulses for the lamp and pulses control the accumulation time of the TV camera, where the duration and frequency of the pulses of the lamp depends on the speed of the ANPA, the amplitude (determines the radiation power) adapts to the transparency of the water environment and to the shooting distance, and the pulses control the accumulation time of the TV camera are consistent with the pulse duration of the illuminator.

Image frames from the TV camera are transmitted over the GigabitEthernet network to the APV PM and recorded together with service information on the Flash drive. AVI file type, JPEG compression format.

In the "video Shooting" mode, the video file is recorded with the frame rate of the TV camera (25-30 Hz) in the continuous backlight mode, the radiation power can also adapt to the transparency of the water environment and to the shooting distance.

At the completion command received from the ANPA BPU, the ASPV system goes into standby mode to save power consumption.

Transfer of photo and video files recorded in the PM ASPV, carried out by standard means of the operating system ANPA network FastEthernet.

Table 1. Main technical characteristics of ASPV

TV camera (monochrome or color)	**Name/ parameter**
Monochrome/Color CMOS photosensitive matrix	IMX265LLR-C/IMX265LQR-C
Optical format	1/1,8″
Resolution, PCs	2048×1536 (3,15Mpx)
Pixel size, μm	3.45×3.45
Focal length of the lens, mm	6
Angle of view diagonally (in water), deg	52
Minimum illumination (on site), Lux	0,1
Shooting distance range, m	2...5
Interface type	GigabitEthernet
Power, V	+12
Power consumption, VA, no more	3
Maximum operating temperature, deg. After	+50
Overall dimensions (without lens), mm	60×29×29
Processor module	
Processor	Intel Atom E3845(1,91 ГГц, 4-core)
RAM/Flash memory	4 GB RAM/16 GB Nand
Flash-disk	C-Fast (SATA), 512 GB
Interfaces	2 x Ethernet 1Gb, 4xRS-422
Operating mode	Photography and video
Maximum video frame rate, Hz	30
Recording images shooting with a frequency of 3 GHz, the hour	50

TV camera (monochrome or color)	Name/ parameter
Video controller (integrated)	VGA, DP
Supply voltage, V	+5
Power consumption, VA	15
Power and control unit	
Input supply voltage range, V	20-30
The mode of formation of the current illuminator	Continuous/Pulse
The duration of the current pulses of the illuminator	Adaptive to ANPA travel speed
The pulses control the accumulation time TV camera	Consistent with the duration of the current pulses of the illuminator
The frequency of photography	1-5 Hz, adaptive to the speed of the ANAPA
Minimum current pulse, μs	Hundred
The amplitude of the illuminator current (determines the radiation power)	Adaptive to the transparency of the water environment and distance shooting
The range of smooth variation of the amplitude of the current of the illuminator, And	0,1-3,5
Exchange interface with processor module	RS-422
Exchange Protocol with a program control unit AUV	fast Ethernet
Telemetry data reception interface from the illuminator	1-wire
Output voltage, V	+5, +12
Led illuminator	
Light source	Led matrix of SvL-23iP100
Axial light intensity, CD	Four thousand
Angle of view (in water)	Sixty
Telemetry data	The temperature of the substrate matrix, humidity, time between
Maximum power consumption in continuous mode of radiation, VA	One hundred
Minimum power consumption in pulsed mode of radiation, VA	Two

Conclusion

The field of robotics using Autonomous unmanned underwater vehicles equipped with underwater vision systems is important for solving a wide range of scientific, economic and defense tasks.

The given recommendations on the choice of light-signal converters of TV cameras using the EMVA technique allow to choose the optimal matrix for solving specific problems of selection of objects in the aquatic environment, where due to the processes of scattering and selective absorption of light, brightness and contrast are reduced, the spatial structure is distorted, the noise component is increased, and, accordingly, the visibility range and the quality of the images is reduced.

For the construction of ASPV for ANPA it is advisable to use led light sources, which have advantages over other light sources.

These recommendations for the separation of the TV camera and the illuminator in the body of the ANPA allow to reduce PORES, improve the quality of the images and to damage the range of visibility of the ASPV in the water environment.

On the basis of the considered recommendations, the block diagram, layout and main technical characteristics of ASPV for use in ANPA of different classes and purposes are given.

References

1. Bocharov L. "unmanned underwater vehicles: current Status and General trends of development". Electronics: Science. Technology. Business 7/2009.
2. Dolin L. S., Levin I. M. "Handbook of underwater vision theory". – L.: Gidrometeoizdat, 1991. – 229s.
3. Voytov A. A., Sergeev V. V., Sokolov V. A., Fursov B. Yu, "the Formation of images in the active system of underwater vision." Questions of radio electronics, ser. Television engineering, 2015, vol. 5. P. 21-31.
4. Bugrov V. E., Vinogradova K. A. "Optoelectronics of LEDs. Textbook.» – SPb: ITMO, 2013. -174 S.
5. Radomsky N. In. "Comparative analysis of the products of the leading manufacturers of white LEDs". Semiconductor lighting engineering, 2010, issue. 4.
6. Pribylov Yu. S., Sergeev V. V., Kosyanchuk S. I., Karpov V. N. "Systems of illumination of objects of observation in television complexes of underwater vehicles", the collection of reports of MNTK "a Look into the future-2018", CDB MT "Rubin", St. Petersburg, 2018, P. 435-441.
7. Berezin, V. V. Solid-state revolution in television: Television system based on charge coupled devices, systems on chip and systems on a chip [Text] / V. V. Berezin, A. A. Umbitaliev, S. Fahmy, A. K., Cytolin, N. N. Shipilov; Under the editorship of A. A. and A. K. Umbetalieva Tsytsulina. – Moscow: Radio and communication, 2006. – 312 p., Il. – ISBN 5-256-01814-0.
8. EMVA 1288 A Standard For Characterising Cameras/ F. Dierks, Basler Vision Technologies.
9. [Electronic resource.] – URL: https://www.baslerweb.com/fp-1489067453/media/downloads/documents/emva_data/BD00094001_Basler_acA1920-40gm_EMVA_Standard_1288.pdf (date accessed: 01.03.2019)
10. [Electronic resource.] – URL: https://www.baslerweb.com/fp-1529914631/media/downloads/documents/emva_data/BD00100301_Basler_acA2040-35gm_EMVA_Standard_1288.pdf (date accessed: 01.03.2019)
11. Baranov S. P., Kozlov V. V., Mansueto A. A. "Sensitivity-matrix color television cameras". Questions of radio electronics, ser. Television engineering, 2008, vol. 1, pp. 1-14.

S. Polovko, V. Tseluyko, A. Popov, D. Stepanov

A COMPUTER VISION SYSTEM FOR DETERMINATION OF AUV POSITION IN THE PROBLEM OF COOPERATIVE DOCKING

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
dnstepanov@rtc.ru

Abstract

With development of the autonomous underwater vehicles the task of automatic rendezvous and docking with an underwater carrier or a stationary dock become actual. The article describes a solution of the problem of determining the position of an underwater vehicle relative to the docking module using a computer vision system. Various existing solutions are considered, the structure of the vision system with based on active light landmarks is presented and two new methods of position determination are described. The applicability of the developed solution is confirmed by numerical experiments and marine field tests.

Keywords: computer vision, AUV, docking, navigation, underwater vehicle, mobile robot, PnP.

Acknowledgments

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 “Investigation of ways to create a multifunctional modular reconfigurable hyper-redundant unmanned underwater vehicle for integration into a robotic complex of three home environments”.

1. Introduction

Determining the relative position of underwater vehicles (AUV) is an important task when approaching or maneuvering at close distances [1]. For underwater vehicles, the most urgent problem requiring high-precision position determination is docking. In most cases the AUV docking is performed under water due to the greater stability of the environment compared to the surface.

The task of docking with an underwater carrier is one of the most difficult, since it requires close interaction of dynamic objects in a mobile environment. The developed AUV requires bringing to the carrier from a long distance, accurate positioning and precise vertical movement until the docking with simultaneous holding of the position in the horizontal plane. At the same time, it is necessary to ensure the possibility of the system operation in a case of a partial failure of its components.

2. Related works

The are several known approaches to the implementation of the underwater docking [2]. Existing solutions for docking and holding torpedo-shaped AUV and remote-controlled underwater vehicles (ROV) traditionally based on the use of special connecting funnel-shaped frames. A typical solution of this kind [3] is shown in figure 1.

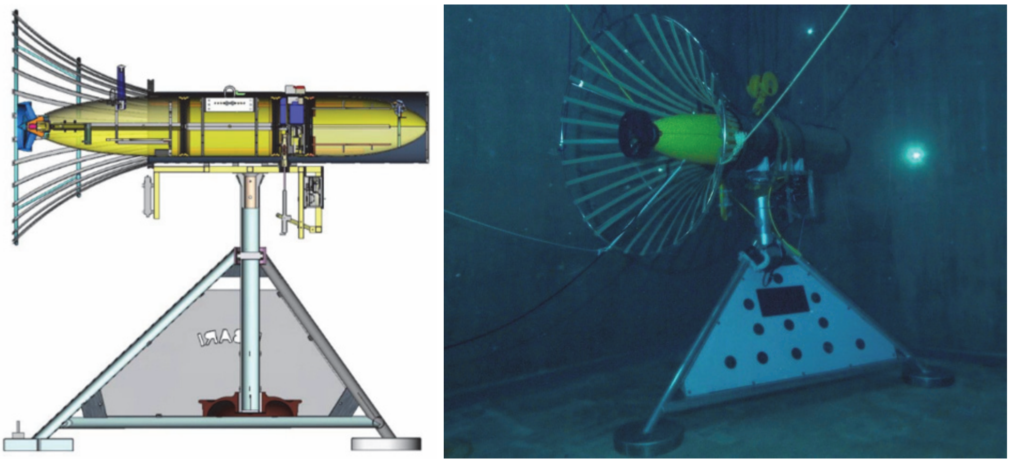


Figure 1 – A funnel shaped frame for a Dorado class AUV

This design involves a horizontal approach of the underwater vehicle using main engines, which allows it to be used for vehicles that cannot hang in place. For underwater vehicles, which have separate engines for lag and vertical displacement, can also be used schemes with a vertical landing in the module. In the docked state, the module provides holding of the AUV, charging of battery modules, high-speed information exchange for obtaining the results of the mission, setting new missions, hardware diagnostics, software integrity check, etc.

Other possible solutions are based on the use of docking pins and frames of various designs, captured by a special device from the underwater vehicle, as well as active grips from the docking modules [1]. However, such solutions are less popular because of the limited opportunities for interaction of the module with the underwater vehicle.

To perform the rendezvous and docking, the docking module is equipped with hydroacoustic and, if necessary, optical means for determining the position of the underwater vehicle. Hydroacoustic systems based on a short-range scheme (with the location of the antennas on the carrier) are used to determine the position over long distances. Ultra-short-range scheme with the location of the antennas on the underwater vehicle allows to determine the position at short distances.

There are also single experimental studies showing the applicability to the guidance problem of scanning sonars (together with a computer model of the docking device), as well as electromagnetic sensors.

At the final stages of docking the use of optical systems providing high accuracy and continuity of measurements is recognized as the most efficient. The main limiting factor in the use of such systems is the range limited by the visibility distance in the water. We can distinguish approaches based on the observation of active light landmarks, passive patterns, as well as systems working on arbitrary objects. Among these approaches, the use of active light landmarks considered in the article is the most efficient, since it provides the maximum range and reliability of measurements, and is the least expensive in terms of computing resources. An example of an image of such a complex system using active light and passive (code) landmarks [4] is shown in figure 2.

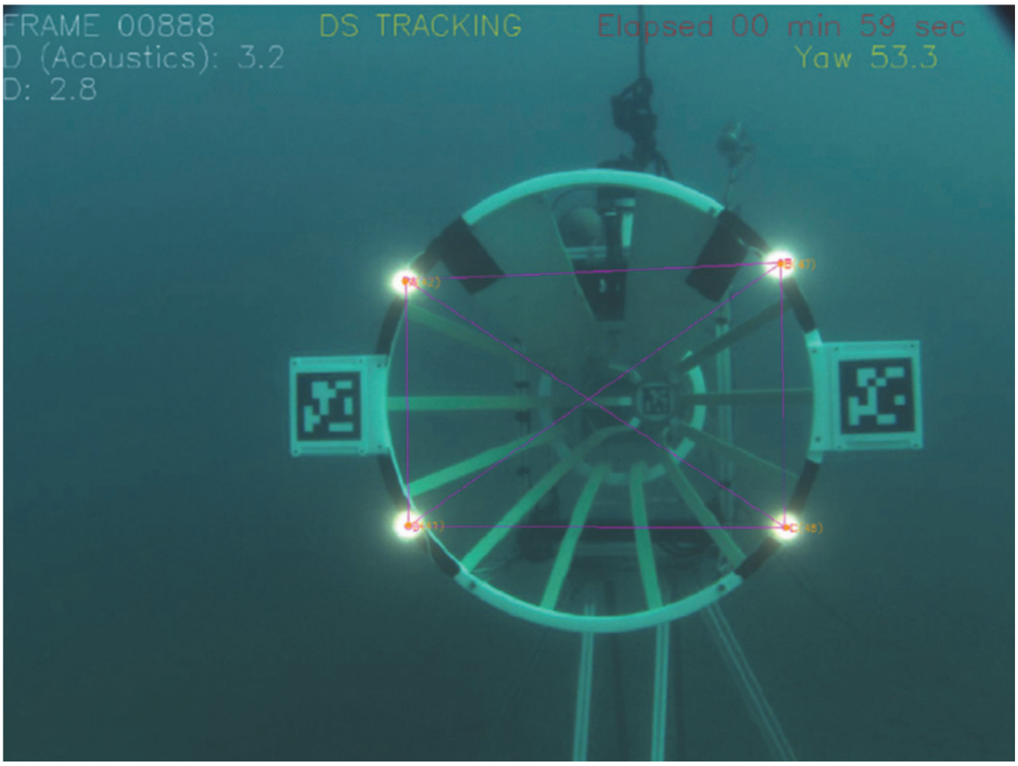


Figure 2 – An example of using active light landmarks together with passive QR-like codes for precise docking of AUV Sparus II in a transparent water

The development of solutions for a particular AUV is always closely related to the features of the vehicle itself (geometry, location of the propulsion and steering modules), with the planned placement of the docking station, with the tasks to be solved. Horizontal approach is convenient for vehicles that cannot provide controllability at low speeds. However, the placement of such a docking module, for example, on an underwater carrier, presents a number of technological challenges and significantly limits access to the AUV for maintenance.

Another factor influencing the choice of technical solutions is the optical properties of water, in which the AUV should work. Existing solutions focus on greater water transparency and, often, the illumination of the passive landmark system by scattered light (from the Sun). This significantly limits the applicability of existing solutions when creating a complex designed for operation at different depths and in conditions of different water transparency.

3. The computer vision system

3.1. The light landmarks and the cameras

The proposed computer vision system (CVS) for determining the parameters of the AUV position during docking consists of a landmarks system located in the docking module, two TV cameras of the AUV directed downwards (to the docking module), an on-board computer and a special software.

As described above, the most effective scheme for optical guidance is the use of special active light landmarks. This solution provides a maximum range of guidance from 5 to 20 m depending on the water transparency. The landmark system consists of 4 light emitting landmarks: two main ones with masks forming 4 points on each landmark, and two additional without masks. TV cameras provide images of landmarks, and on-board computer uses the special software for joint analysis of images from both cameras.

The light landmarks should fall into the field of view of the television camera and be clearly visible on the frame both at the far border of the operating range of the system (about 10-20 meters, depending on the transparency of the water) and at the near (about 0.5 meters).

In most cases the AUV have data from the inertial navigation system, giving the angles of roll and trim with high accuracy. Taking into account these data, it is possible to solve the problem of finding the mutual position of the devices by only two landmarks observed by one television camera. And in the presence of four or more distinguishable landmarks in the field of view of the camera, the problem can be solved without knowing the angles of roll and trim for each vessel [5, 6].

The light landmarks should be in the field of view of the television camera both on the far border of the operating range of the system and on the near one. The width of the camera field of view for different distances (at an angle of view of 45 degrees) is shown in figure 3.

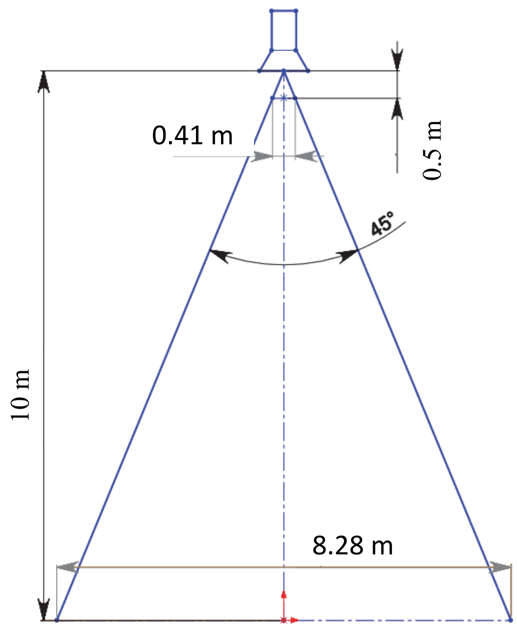


Figure 3 – The width of the camera field of view (angle of view is 45°)

As can be seen from the figure 3, at a distance of 10 meters, the width of the field of view is 8.28 meters, and 0.5 meters it is only 0.41 meters. In order not to lose the accuracy and stability of the system it is necessary to have several sets of landmarks for different distances. This approach is used in other existing systems too [4].

To solve the problem for each camera on the far edge of the operating range, you must have 2 landmarks. It is desirable that when approaching along the axis of the camera, they do not go out of its field of view to about 3 meters, which will give some margin for possible lateral displacements when approaching. In this case, each light landmark may consist of several spaced light spots (or, for example, one large lamp with a mask superimposed on it). This would be guaranteed to give at a long distance one clearly distinguishable spot, and several spots at the near (figure 4).

Since data from two television cameras will be used to rectify the angle of the mutual course of underwater vehicles, it is desirable that the cameras are as far apart as possible. An example of the configuration of the system of light landmarks and cameras is shown in figure 5.

Figure 4 – Light landmark at a) – distance 10 m, b) – distance 0.5 m

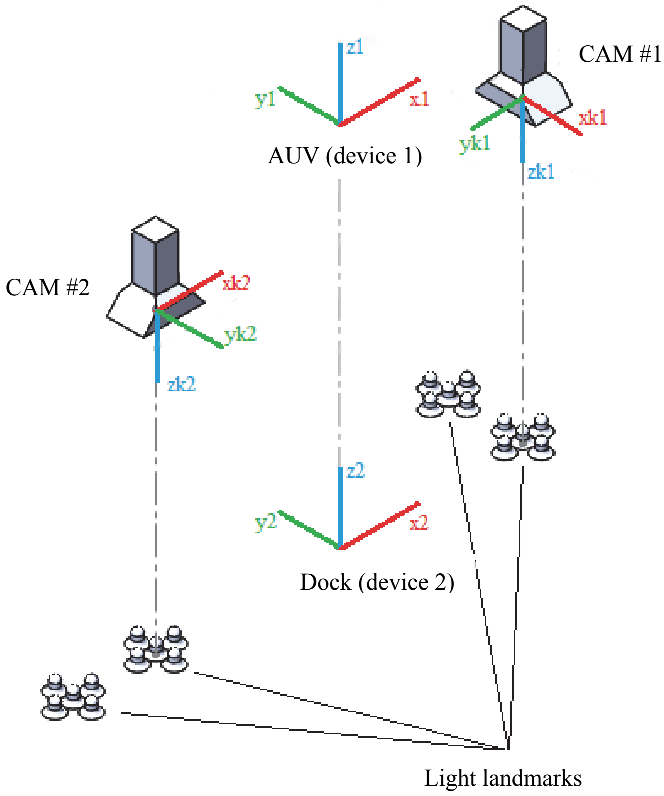


Figure 5 – An example of positioning of light landmarks and cameras of the AUV

3.2. Automatic landmarks detection and their grouping

After receiving the television frames (images) it is necessary to find the light landmarks, and then to number them in accordance with the order defined in the system. It may also be necessary to group several light spots into one and find the center of the landmark.

To find landmarks, binarization and subsequent selection of contours [7] with automatically adjustable parameters are used.

After finding the landmarks is the numbering of the found points. If the approximate angles of the course of the vehicles are known, the points are numbered using the knowledge of these angles, or simply by the part of the image in which they are located.

If point grouping is required, the DBSCAN clustering method is used [8].

3.3. A method of determining the relative position by the two points on the image and the known angles of the roll and trim

At known angles of roll and trim for each of the two underwater vehicles (AUV and dock), it's possible to calculate the angle of their mutual roll by the image. To do this, it is necessary that in the field of view of the television camera were at least two points with known spatial coordinates. In our case the points are the light landmarks.

After that, it will be possible to calculate the rotation matrix between the coordinate systems of the devices. And, knowing how the devices are rotated relative to each other, by the same image with two points, we will be able to calculate all three relative displacements.

Some PnP methods are able to find the position of an object by three points [9, 10]. But this raises an ambiguity, and choosing the right solution is paired with the computational and algorithmic complexity [11]. The proposed method, as already mentioned, finds a solution for only two points.

a) **Finding the rotation matrix between the reference frames of AUV and the dock.** Let the roll and trim angles for the both devices are know: $a_{x1}, a_{x2}, a_{y1}, a_{y2}$.

The figure 6 shows the XY-planes of the devices and the horizontal axes when the mutual roll is zero. Lines P11 и P12 are parallel to the plane 1, while the P21 и P22 are parallel to the plane 2. All these lines lie in a vertical plane perpendicular to the horizontal plane.

Knowing the orts of these lines it's possible to construct the equations of the planes 1 and 2 relative to the horizontal reference frame. While now we consider the mutual course equals zero to make the explanation more clearly, one of the planes will finally be rotated by the necessary angle.

Axes X of the devices coincide with the plane x0, so the angles al_{x1}, al_{x2} equal to a_{x1}, a_{x2}. The angles al_{y1}, al_{y2} don't equal to a_{y1}, a_{y2} and need to be calculated (figure 7).

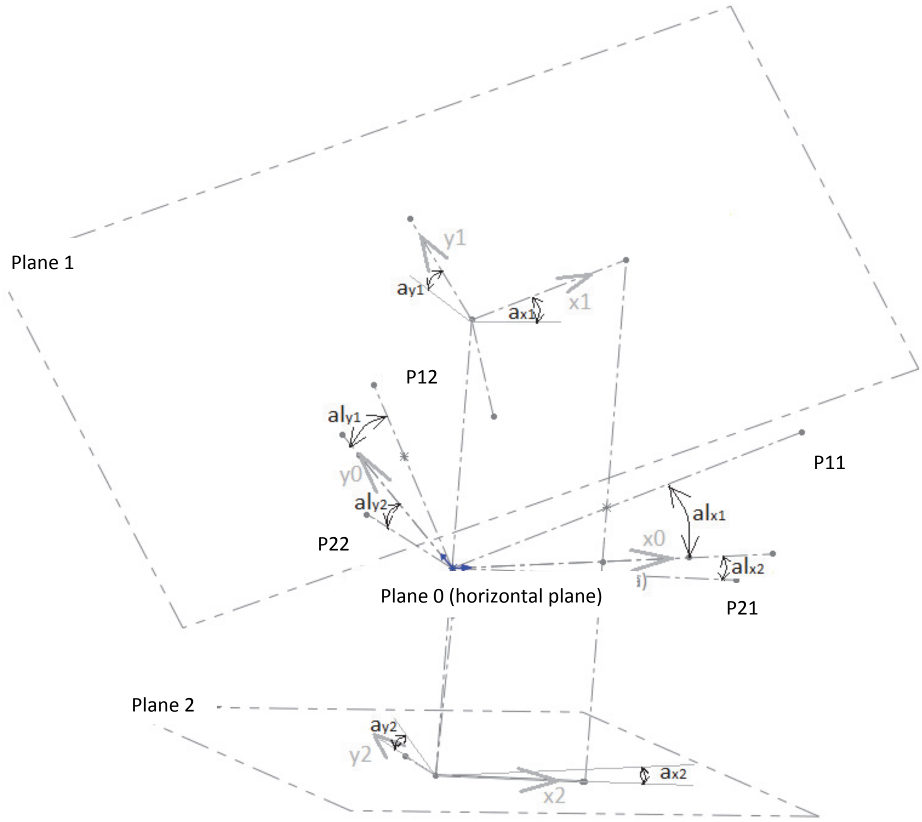


Figure 6 – Reference frames and planes of the devices

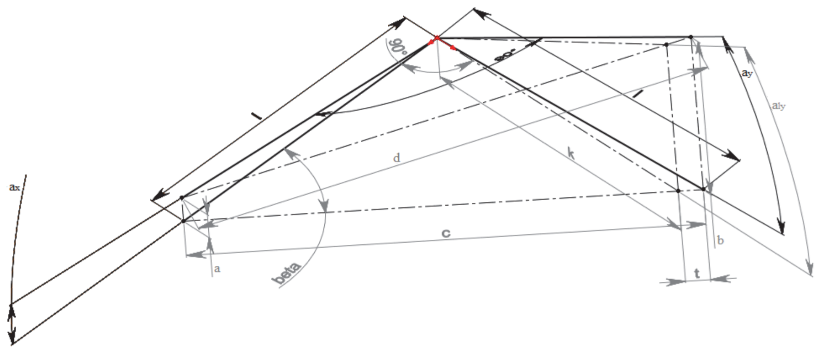


Figure 7 – Calculation of the al_y angle

b) **Calculation of the auxiliary angle al_y.** From the $abcd$ trapezoid:

$$c = l * \sqrt{\frac{1}{\cos(a_x)} + \frac{1}{\cos(a_y)} - (\mathrm{tg}(a_x) - \mathrm{tg}(a_y))^2} \tag{1}$$

$$d = l * \sqrt{\frac{1}{\cos(a_x)} + \frac{1}{\cos(a_y)}} \tag{2}$$

Knowing c, one can calculate the angle β:

$$\beta = \frac{(\pi - \arccos(\frac{-(c)^2 + 2 * l^2}{2 * l^2}))}{2} \tag{3}$$

According to the 1 – 3 the angle al_y becomes:

$$al_y = \mathrm{arctg}\left(\mathrm{tg}(\alpha_y) + \left(c - l * \sqrt{\mathrm{tg}(\beta)^2 + 1}\right) * \mathrm{tg}\left(\arcsin\left(\frac{l * \left(\mathrm{tg}(\alpha_x) - \mathrm{tg}(\alpha_y)\right)}{d} \right)\right)\right) \tag{4}$$

c) **Calculation of the rotation angles.** Knowing $al_{x1}, al_{x2}, al_{y2}, al_{y2}$, one can construct the equations for the planes 1 and 2 and find their normal vectors (which are the orts of the z axes):

$$Z_1 = n_1 = \begin{pmatrix} -\mathrm{tg}(al_{x1}) \\ -\mathrm{tg}(al_{y1}) \\ 1 \end{pmatrix} \qquad Z_1 = n_2 = \begin{pmatrix} -\mathrm{tg}(al_{x2}) \\ -\mathrm{tg}(al_{y2}) \\ 1 \end{pmatrix} \tag{5}$$

Let's also find the orts of the x axes of the planes:

$$X_1 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{tg}(al_{x1}) \end{pmatrix} \qquad X_2 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{tg}(al_{x2}) \end{pmatrix} \tag{6}$$

And the y axes are determined as following:

$$Y_1 = Z_1 * X_1 \qquad Y_2 = Z_2 * X_2 \tag{7}$$

After the vectors normalization one can construct the rotation matrixes between the reference frames of the planes and the horizontal plane:

$$M_{01} = \begin{pmatrix} X_1 & Y_1 & Z_1 \end{pmatrix} \qquad M_{02} = \begin{pmatrix} X_2 & Y_2 & Z_2 \end{pmatrix} \tag{8}$$

Next you need to rotate the coordinate system of one of the planes (for example the second one) around the z0 axis of the horizontal plane. The rotation can be described using the quaternion constructed from the rotation angle (mutual course angle k_ang) and the direction vector $z0 = (0,0,1)^T$.

$$quat = \begin{pmatrix} 0 \\ 0 \\ 1 \\ k\_ang \end{pmatrix} \tag{9}$$

After the rotation we get the rotation matrix M_{02r} .

Now the rotation matrix between the two devices reference frames can be determined:

$$M_{12} = M_{01} * M_{02r} \tag{10}$$

Knowing the rotation matrices between the AUV reference frame and the camera reference frame M_k one can determine the rotation matrix of the camera relative to the second device (dock) frame:

$$M_{k2} = M_k * M_{12} \tag{11}$$

d) **Determination of the offsets.** Now, having a matrix of camera rotation relative to the second device and the screen coordinates of the two landmarks on the image, one can calculate the displacement and get the full transition matrix.

Let one of the points have spatial coordinates x and y equal to zero. If this is not the case, then shift the coordinate system of the second device, get the result and do the opposite shift.

Consider the picture in the ZX-plane of the camera, since the ZY-plane will be the same (figure 8).

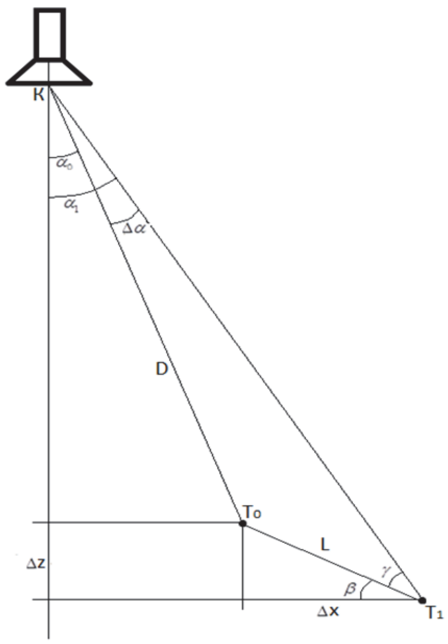


Figure 8 – Points in one of the camera planes

Using this matrix, we obtain the spatial coordinates of the points in the reference frame of the camera taking into account the rotation:

$$\begin{aligned} T_0 &= M_{k2} * T_0' \\ T_1 &= M_{k2} * T_1' \end{aligned} \tag{12}$$

where T_0', T_1' – spatial coordinates of the points in the second device reference frame.

Knowing the internal camera parameters one can calculate the visible angle of the points:

$$\alpha_{0x} = \text{arctg}(t_0.x - x_c / f_x)$$
$$\alpha_{1x} = \text{arctg}(t_1.x - x_c / f_x) \tag{13}$$

where t_0, t_1 – image points coordinates;

x_c – principal point x coordinate;

f_x – focal length.

Let’s calculate the line L length:

$$L_x = \sqrt{(T_1.x - T_0.x)^2 + (T_1.z - T_0.z)^2} \tag{14}$$

Angle β:

$$\beta_x = \text{arctg}(\frac{T_1.z - T_0.z}{T_1.x - T_0.x}) \tag{15}$$

Knowing that the angle at vertex T_1 equals $\alpha + \beta$, having considered triangles, KT_0T_1 и $KT_1\alpha_1$, we can calculate angle γ:

$$\gamma_x = \frac{\pi}{2} - \alpha_{1x} - \beta_x \tag{16}$$

Knowing the γ angle and using the sine theorem one can calculate the length of the D line:

$$D_x = \frac{L_x * \sin(\gamma_x)}{\sin(\beta_x)} \tag{17}$$

And knowing D and angles to the points, one can calculate all the three displacements of the cameras relative to the second device reference frame:

$$Z = \cos(\alpha_{0x}) * D$$
$$X = \sin(\alpha_{0x}) * D \tag{18}$$
$$Y = Z * \text{tg}(\alpha_{0y})$$

Thus, we fully know the position of the camera of the first device (dock) relative to the second (AUV).

e) **Finding the angle of the mutual course.** It remains to figure out how to initially determine the angle of the mutual course k_ang.

This is done using the above model. At one point, the assumption is made about the X and Y displacements of the camera coordinate system relative to the coordinate system of the second device. After that, the projection of two points on the frame with a zero angle of the mutual course and these displacements is constructed. The angle between the line between the projected points on the frame and the x-axis of the camera is considered, and the misalignment of this angle with the same angle of the screen points is considered. After that, by iteratively changing the value of the course angle, this mismatch is minimized.

3.4. A method for determining the relative position of a set of points on an image

Having in one image 4 or more points with known spatial coordinates relative to the dock reference frame one can state and solve the PnP-problem [9]. There are methods for solving the problem for a certain number of points [6,12], and methods that allow to solve the problem for an arbitrary number of points (more than four) with an arbitrary geometry [5,13].

The system uses the method proposed in [14], because it shows high accuracy, sufficient speed and robustness.

3.5. Rectification of the angle of the mutual course according to data from several television cameras

The angles of roll and pitch with high accuracy is taken from the testimony of the inertial system or of the inclinometers and the angle of relative roll is determined directly according to the testimony of a television camera and is dependent on errors of location of the landmarks on the frame.

After the mutual position of the devices with the use of several cameras is obtained, it becomes possible to clarify the angle of the mutual roll. And for a method that uses two points for calculation, after finding the angle of mutual roll, it is possible to refine the entire solution by restarting the method with the calculated angle.

Consider the case where there are two cameras. More cameras can always be reduced to this situation by taking them in pairs.

After the calculation there are to transition matrices from the reference frame of the dock (figure 3) to the coordinate systems of the cameras of the AUV M_{k12}, M_{k22}.

Taking the coordinates of any of the points visible by the first camera P_{1t2} and any point visible by the second camera P_{2t2} we can get their coordinates in the camera coordinate systems:

$$\begin{aligned} P_{2t1k} &= M_{k12} * P_{2t2} \\ P_{1t2k} &= M_{k22} * P_{1t2} \end{aligned} \tag{19}$$

The same coordinates can be obtained in a different way knowing the transition matrix between the cameras:

$$M_{k1k2} = M_{k1}^{-1} * M_{k2} \tag{20}$$

where M_{k1}, M_{k2} – transition matrices from the AUV reference frame to their cameras reference frames.

The resulting formula is:

$$\begin{aligned} P'_{12k} &= M_{k1k2}^{-1} * M_{k12} * P_{1t2} \\ P'_{21k} &= M_{k1k2} * M_{k22} * P_{2t2} \end{aligned} \tag{21}$$

So, we have two points whose coordinates are calculated in two different ways: (P_{1t2k} and P'_{12k}) relative to the first camera reference frame and (P_{2t1k} and P'_{21k}) to the second camera reference frame.

The distance between the points can be written as follows:

$$\begin{aligned} D_{12} &= P_{1t2k} - P'_{1t2k} \\ D_{21} &= P_{2t1k} - P'_{2t1k} \end{aligned} \tag{22}$$

The smallest of these distances will show which camera has a more accurate calculation of the relative position.

Further, for this camera, by adjusting the angle of the mutual course of the devices, the minimum value D is found, at which the course will be adjusted optimally.

4. Computer model for the study of the methods

A computer model was developed to study the system and methods for determining the relative position of underwater vehicles using television cameras and light landmarks. The model is written in C++ using the Qt library.

The graphical interface for working with the model is shown in figure 9.

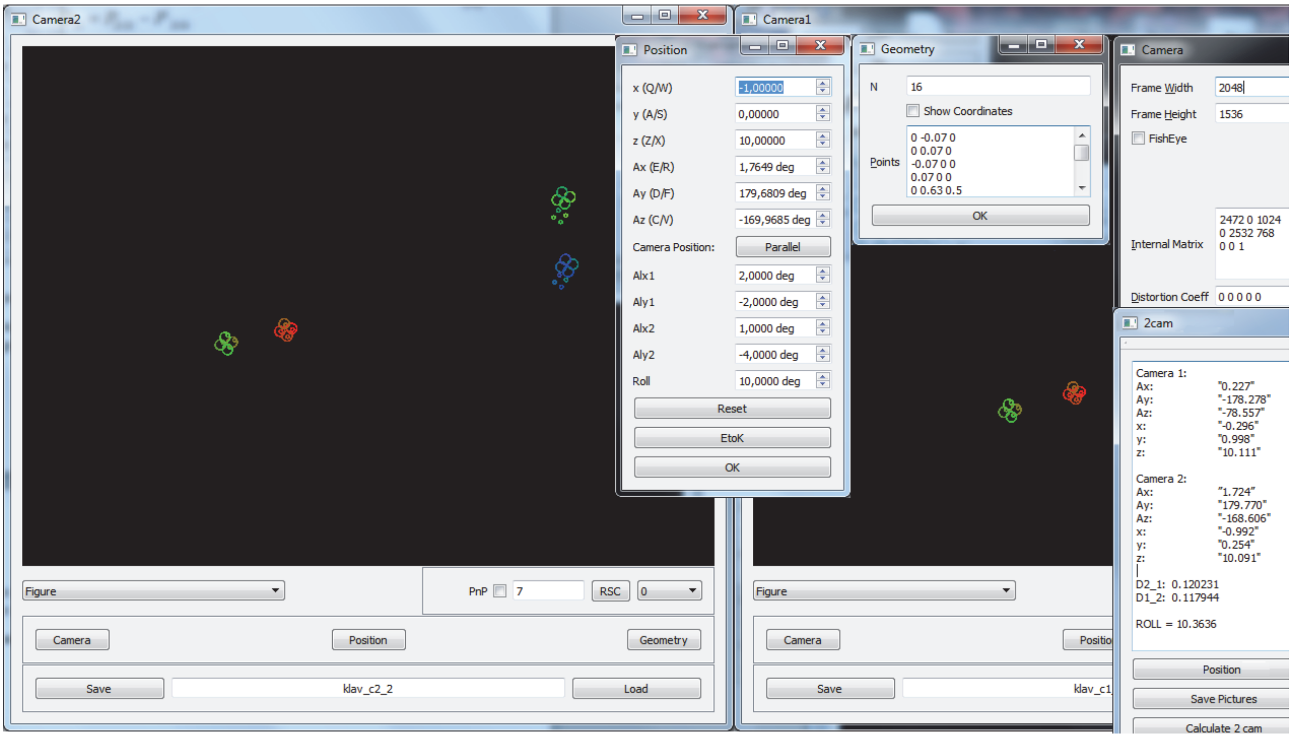


Figure 9 – The graphical interface for working with the model

The computer model allows one to set different landmark configurations, camera locations and parameters. It visually simulates frames from cameras with landmarks and other objects of the scene. The model allows one to introduce different bias in the original data (pixel coordinates of points, the roll and pitch, etc.). Allows one to test methods, both manually and automatically.

5. Accuracy characteristics of the methods

Accuracy characteristics, speed and stability of PnP-methods, including those used in the presented system, are studied in detail in [14]. Therefore, for the two methods used in the system, tables of dependence of the accuracy of the method on the pixel error of the screen coordinates of landmarks at different distances are given (tables 1-4).

Table 1. Method 1, 10 meters

RMS error, pixels	x, m	y, m	z, m	Ax,	Ay,	Az,
0	0,000	0,000	0,000	0,000	0,000	0,002
1	0,007	0,007	0,046	0,000	0,000	0,335
2	0,015	0,013	0,091	0,000	0,000	0,463
3	0,026	0,026	0,146	0,000	0,000	1,386
4	0,048	0,033	0,235	0,000	0,000	1,456

Table 2. Method 1, 3 meters

RMS error, pixels	x, m	y, m	z, m	Ax, °	Ay, °	Az, °
0	0,000	0,000	0,000	0,000	0,000	0,002
1	0,001	0,002	0,006	0,000	0,000	0,090
2	0,001	0,004	0,008	0,000	0,000	0,209
3	0,004	0,006	0,017	0,000	0,000	0,203
4	0,005	0,007	0,018	0,000	0,000	0,410

Table 3. Method 2, 10 meters

RMS error, pixels	x, m	y, m	z, m	Ax, °	Ay, °	Az, °
0	0,000	0,000	0,000	0,000	0,000	0,000
1	0,019	0,018	0,136	1,371	1,343	1,003
2	0,036	0,029	0,259	2,533	2,758	1,808
3	0,054	0,052	0,386	4,797	3,923	3,309
4	0,072	0,053	0,464	5,353	5,190	3,905

Table 4. Method 2, 3 meters

RMS error, pixels	x, m	y, m	z, m	Ax, °	Ay, °	Az, °
0	0,000	0,000	0,000	0,000	0,000	0,000
1	0,001	0,001	0,006	0,398	0,383	0,287
2	0,001	0,001	0,013	0,926	0,734	0,629
3	0,002	0,002	0,021	1,330	1,114	0,954
4	0,003	0,003	0,025	1,661	1,354	1,129

Method 1 is a method described in the article that works on 2 points and on the given angles of roll and trim, Method 2 is a PnP method that works on at least 4 points.From these tables it can be seen that the method 1 gives smaller errors than the method 2.

6. Conclusion

The presented system makes it possible to determine with high accuracy the mutual position of underwater vehicles at distances at which special light landmarks are visible for the television cameras.

The method of determining the relative position by two points on the television frame and two known angles of orientation of the vehicle (obtained from the navigation system with high accuracy) showed better results in accuracy than the classical PnP-methods. Due to the need to find only two points on the television frame, it can be used in various systems to find the position of a device relative to a scene with a small number of features with known geometry.

Refinement of the angle of the mutual course using data from several television cameras showed good results in improving the accuracy of the studied methods. If one has multiple cameras without specialized methods of finding the position it can help to refine the solution obtained by the cameras.

The computer model developed for the study and testing of this system has turned out to be quite universal and can be used in the future to test other similar methods or systems.

The conducted offshore field tests of the system showed high accuracy of measurements of the coordinates of the AUV relative to the carrier (sufficient to perform the automatic docking). During the tests, for the first time in the Russian Federation, an AUV was landed on an underwater carrier in an automatic mode.

References

1. Bakhshiev, A.V. and others. Proposals development for the design and technical implementation of the determining parameters relative motion system based by video processing. - Technical report. - / RTC. – Saint-Petersburg (2010).
2. Podder, T., Sibenac, M., Bellingham, J. Applications and Challenges of AUV Docking Systems Deployed for Long-term Science Missions (2019).
3. McEwen, R. S., Hobson, B. W., McBride, L., Bellingham, J. G., Docking Control System for a 54-cm-Diameter (21-in) AUV. In: IEEE Journal of Oceanic Engineering, vol. 33, no. 4, pp. 550-562, Oct. 2008.
4. Palomeras, N., Vallicrosa, G., Mallios, A., Bosch, J., Vidal, E., Hurtos, N., Carreras, M., Ridao, P., AUV homing and docking for remote operations. In: Ocean Engineering, Volume 154, Pages 106-120 (2018).

5. Moreno-Noguer F., Lepetit Y., Fua P. Accurate noniterative o(n) solution to the PnP-problem // Proceeding of the IEEE International Conference on Computer Vision / Rio de Janeiro, Brazil, October, 2007. – Rio de Janeiro (2007).
6. Horaud R., Conio B. An Analytic Solution for the Perspective 4-Point Problem // Computer Vision, Graphics and Image Processing. – Vol. 47. – P. 33-44 (1989).
7. Suzuki, S. and others. Topological structural analysis of digitized binary images by border following // Computer Vision, Graphics, and Image Processing. –Vol. 30. – P. 32–46 (1985).
8. Ester, M., Kriegel, H., Sander, J., Xiaowei X. A density-based algorithm for discovering clusters in large spatial databases with noise // Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. / AAAI Press. – P. 226–231 (1996).
9. Fischler, M., Bolles, R. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography // Communications of the ACM. – Vol. 24(6). – P. 381-395 (1981).
10. Gao, X. S., Hou, X. R. Complete Solution Classification for the Perspective-Three-Point Problem // IEEE Transactions on Pattern Analysis and Machine Intelligence. – Vol. 25. – P. 930-934 (2003).
11. Carceroni, R., Brown, C. Numerical Methods for Model-Based Pose Recovery. – Rochester – (Technical Report 659 / Computer Science Dept. / U. Rochester) (1997).
12. Bujnak, M., Kukelova, Z., Pajdla, T. New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion // Proceedings of the Asian Conference on Computer Vision 2010 / Queenstown, NZ, November 8-12, 2010. – Queenstown (2010).
13. Hesch, J.A., Roumeliotis, S.I. A Direct Least-Squares (DLS) method for PnP // Proceedings of the IEEE Conference on Computer Vision (2011).
14. Kirpan, N.A. The algorithm for determining the position and orientation of an object with a reference target by a flat image recorded at a space docking: Master degree thesis: 220400.68.01: 18.06.13 – Saint-Petersburg (2013).

I.V. Pashkevich, A.V. Grinenkov, G.V. Konyukhov, L.A. Martynova, A.O.Pronin, G.A.Podshivalov, V.V.Prokopovich, N.I.Gorbachev

FEATURES OF THE IMPLEMENTATION OF AUV EMERGENCY SUBSYSTEM DURING THE USE OF MULTI-AGENT TECHNOLOGY IN ITS CONTROL SYSTEM

Concern Central Research Institute Electropribor, St. Petersburg, Russia
iv@bk.ru, grin_a_v@mail.ru, kongv1@yandex.ru, martynowa999@bk.ru, pronin.ao@gmail.com, 930730@mail.ru, wm.prokopowich@yandex.ru, gorbachev_ni @ elprib. ru

Abstract

The causes of accidents on autonomous underwater vehicles when they overcome super-long crossings are considered. It was found that the causes of accidents may be the apparatus's getting into dangerous zones: increased shipping zones, fishing net setting zones, shallow sea zones with complicated uneven relief, ice drift zones, garbage "island" zones, current zones and underwater volcanoes. To avoid accidents, it is necessary to identify and classify the apparatus in hazardous areas, which was not previously considered.

An analysis of the causes of emergencies in hazardous areas has been carried out, signs have been identified by which emergency situations can be detected, and means capable of detecting these signs. As a result of the analysis, it was revealed that the sources of information about the apparatus's entry into the danger zone are: means of lighting the situation that can detect icebergs, garbage "islands" and fishing nets; the navigation subsystem, which allows to detect strong currents and eddies; and the subsystem of the main propulsion and steering complex that compares the parameters of the position of the rudders and the rotation of the propellers with the navigation data and the results of predicting the behavior of the device using a mathematical dynamic model.

Identified the signs by which it is possible to carry out the classification of pre-emergency and emergency states of the device.

To implement in-depth analysis to identify signs of emergency situations, it was proposed to use the AUV hardware control subsystem due to its relative computational underload and direct control of emergency vehicles.

The results of the analysis make it possible to foresee in the algorithms for processing information arriving at the inputs of the apparatus's subsystems, identifying and classifying previously unresolved features: fishing nets, garbage islands, currents, etc., and also taking them into account, along with traditional ice detection, additional information on the area of increased shipping, etc. All this contributes to the trouble-free navigation of the AUV, including over long distances.

Keywords: autonomous uninhabited underwater vehicle, accident, danger zone, detection, classification.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (Projects No. 17-08-00666, No. 19-08-00253).

Introduction

The increasing complexity of tasks with the use of autonomous underwater vehicles (AUV) [1-2] led to the use of modern technologies and energy sources [3], which are often unsafe in the event of emergency situations on AUV. The occurrence of an emergency can lead not only to economic losses caused by the flooding of the AUV and its subsequent search and recovery, but also to the leakage of chemically hazardous substances and even to an explosion. This, in turn, may entail both environmental consequences associated with marine pollution and human casualties. In this regard, the safety issues of AUV are relevant.

A feature of the emergency subsystem is the flow of information from external and internal subsystems and sensors into the AUV control system, in which the integrated assessment of the possibility of an accident occurs, its classification and issuing control commands to prevent or eliminate the accident [4].

With the centralized organization of the control system, information from the systems and sensors about the pre-emergency and emergency states goes to the center, which, in turn, according to the results of the analysis of the received information, issues commands to emergency means of the AUV.

A more modern approach to building a control system is a hierarchical approach, in which a control and alarm system is formed. At the top level of the hierarchy, decisions are made based on the development scenarios provided by the mid-level control and alarm system, based on data supplied by sensors and lower-level systems [4].

However, today one of the promising approaches to building a control system for complex AUV is a multi-agent approach, in which each subsystem of AUV has its own agent [5-7]. Agents perform the functions assigned to them and within these functions independently make decisions. Each agent communicates only with those agents from whom he receives the information necessary for decision-making and with whom he must issue the decisions or commands made by him.

A feature of the functioning of the emergency subsystem is the collection of data on the state of the hardware of all AUV subsystems, data from sensors of its internal state and the state of the environment, their processing and transmission of control commands to the AUV subsystems. Therefore, in order to determine the approach to the implementation of the emergency subsystem in the multi-agent control system of AUV, we will analyze possible options for the occurrence of AUV emergency conditions.

As a result, to ensure trouble-free operation of the AUV, there is a contradiction between the need to collect data from all subsystems of the AUV, on the one hand to conduct an integrated assessment of the state of the AUV, and the distributed functioning of the subsystems-agents of the AUV multi-agent control system, on the other hand.

The goal was to form the work of the emergency subsystem of the AUV in the control system of the AUV based on multi-agent technology.

To achieve the goal in work the following tasks were solved:

– analysis of the causes of accidents on the AUV;

– analysis of information from sources about the detection of signs of pre-emergency and emergency situations;

– the analysis of the influence of the pre-emergency state of the AUV on the work of its subsystems;

– a list of tasks assigned to the emergency subsystem;

– an analysis of the features of the multi-agent control system of AUV was conducted;

– proposed an approach to the implementation of the emergency subsystem in the multi-agent control system of the AUV with the characteristics of its operation.

1. Causes of AUV emergency situations

The most severe consequences of accidents include: flooding of AUV with the lack of information about the site of flooding, the release of environmentally hazardous substances, loss / damage to equipment, damage to the building of AUV and its subsequent flooding, equipment displacement, fires, etc. There is a reasonably close causal relationship between “hazard factors — causes — effects”. The degree of accuracy of forecasting the development of events in the event of a threat of emergency situations depends on the completeness of information about causal effects during the operation of the AUV. The need for such a forecast is due to the need to find a rational solution to anticipate the occurrence of emergency situations or minimize the consequences. Consider the causes of emergency situations.

On the AUV they can be both external and internal.

External causes are caused by ingress of the AUV in areas of potential danger to the AUV. These zones are [8-10]:

– areas of increased shipping;

– zone setting fishing nets;

– zones of the shallow sea with a complex uneven relief;

– ice drift zones;

– zones of garbage islands;

– zones of currents and volcanoes.

A prerequisite for avoiding the occurrence of emergencies is the construction of a route to bypass areas that represent a potential danger to the AUV. However, during the movement of the AUV according to the route assignment, a deviation from the specified route is possible (accumulation of positioning errors, the impossibility of conducting a timely observation), the AUV may be in the danger zone. In addition, the boundaries of the danger zone may move over time (during the mission) or untimely marked on the sea map.

As a result, the AUV may be in the zone of increased risk of an emergency situation on the AUV.

Thus, the movement of AUV in the zone of increased shipping can lead to a collision with other vessels, resulting in possible damage to the AUV skin and violation of the integrity of the hull, its depressurization, displacement of the AUH internal equipment, fire, fire, explosion, flooding of the AUV.

Fishing nets, which are particularly abundant along the coastline, can be a serious source of accidents for the AUV [9]. Moreover, the fishing net can be wound up on the propulsion unit, so the AUV itself can also get into it [11]. The reason for entering the fishing net may be a low speed of the AUV and the low power of the sonar that can detect fishing nets only at small distances, since any sound impulse that falls on it dissipates. In addition, even if the sonar detects the fishing net ahead of the course, the low speed of the AUV (4-5 knots) does not allow an accurate maneuver to evade the detected objects, especially if maneuvering occurs against the current. In addition, the current may catch fishing nets and other debris on the AUV.

Another source of an accident after fishing nets may be garbage "islands", which can grow due to the characteristics of currents. The danger of getting into such a region is due to the movement of AUV in a heterogeneous medium with a high density, some fragments of which can lead to entanglement and sticking of the body of AUV, as a result of which it can become uncontrollable.

An emergency situation can occur in the ice region: the presence of ice crumb can damage the hull and propeller blades, slow down the course of the AUV, and the presence of drifting icebergs can lead to a collision with them.

It is possible for the AUV to enter the zone of whirlpools and currents that move in a certain direction and direction [9]. Another cause of accidents can be underwater volcanoes: when a volcano erupts, methane is released, resulting in a huge gas bubble rising to the surface of the water, capable of provoking the failure of the AUV to the beyond depth [9].

In the absence of designation on the map of derricks installed on the seabed, or deviation of the AUV from the route, collision with them is also possible.

A separate cause of an emergency may be the attack of giant squids or sharks, which has already happened in practice [9].

But even in the absence of serious external causes, in the course of the mission, technical malfunctions of the AUV hardware may arise, or natural phenomena may provoke or aggravate them. It is not so important whether there was a technical failure due to external or internal reasons. Therefore, along with identifying external causes, constant monitoring of the technical state of the hardware of all subsystems of the AUV is necessary. This is done by an automated diagnostic system (ADS).

The description of possible situations that could lead to an accident allows us to analyze which sensors and with what information we can prevent the occurrence of an AUV accident.

2. Sources of information to prevent the occurrence of an accident

Since an emergency situation is characterized by deviations in the functioning of the AUV from normal operation, in order to prevent an accident, it is necessary to monitor the state of the AUV hardware [12–22]. In order to know what exactly to control in terms of the causes of the accident (breakdown of the hull, etc.), their occurrence and in terms of consequences, it is necessary to consider the sources of information by which the causes of accidents of the AUV can be classified.

These include:

– means providing the emergency system with environmental data (presence of icebergs, garbage “islands”, currents, fishing nets, etc.) that could lead to an accident;

– means providing the emergency system with data on the internal state of the AUV, which may change both as a result of the operation of the AUV hardware and as a result of the appearance of external causes leading to the accident.

Sources of environmental information can be obtained primarily from the agent of the situation lighting subsystem (SLS).

SLS, which may include, for example, an active-passive sonar, ice thickness meter, provides:

– target detection - mobile and immobile;

– classification of detected targets;

– periodic diagnostics of the subsystem hardware and software.

SLS provides detection of various obstacles by detecting the presence of other vessels within a radius of about 4 km with a reliability of 99.9%, and their identification [23], as well as more complex obstacles such as fishing nets, drilling rigs and other bottom structures. As a rule, sonars are installed in the forward part of the AUV, while they scan the space ahead of the course. With a passing current, the aft part can be in the dead zone of the sonar, and then the network can be thrown over the AUV. In this regard, it is advisable to use antennas that provide all-round visibility.

Additionally, information on the location of civil courts, formed by Automatic Identification System (AIS), can be used.

Another important source of information for detecting an emergency is the navigation subsystem agent. Navigation subsystem (NS) is designed to develop its own coordinates and motion parameters of AUV (course, speed, depth, angles of heel and trim) and forecast the current deviation from a given route. The composition of the NS may include, for example, inertial navigation system (INS), which forms the coordinates of the AUV as a result of the reckoning of the path; hydroacoustic (absolute) lag (GAL), measuring the components of the speed of the AUV; relative lag, measuring the speed of the AUV relative to the flow; depth gauge; echo sounder, measuring the depth under the keel.

The agent of the navigation subsystem allows to detect the flow. The presence and rate of flow can be measured by calculating the difference of the data obtained from the absolute and relative lags, as well as the number and observational position of the AUV. Determining the presence of a current, its direction and width makes it possible to assess the approach of the AUV to dangerous zones and take timely measures to bypass these zones, as well as to take into account the choice of movement parameters along a given route.

Another source of information is the agent of the subsystem of the marching propulsion complex (MPC) in the event that the response to the commands given to the propulsion team do not correspond to the required movement parameters. Data MPC can be used in conjunction with navigation data and the results of predicting the behavior of the AUV when submitting certain control commands. To assess the adequacy of the MPC response to the commands given, it is advisable to use a mathematical dynamic model, as was done, for example, when diagnosing a distributed computer network [23,24].

The mathematical dynamic model takes into account that when the ANP moves in an infinite fluid, gravitational, mass, and hydrodynamic forces act on it, and it is possible to simulate all the main maneuvering modes of the AUV during normal operation:

– movement at constant depth (balancing);
– movement at constant depth with varying untrimming;
– depth change maneuvers with automatic control of the bow horizontal rudders and stern rudders;
– depth change maneuvers with automatic control of vertical thrusters;
– maneuvers to change the depth when using equalizing-differential system;
– maneuvers of change of course by transferring the feeding rudders;
– course change maneuvers using lag thrusters;
– maneuvers change course during acceleration, reverse and free coasting;
– movement lag under lag thrusters;
– joint control of marching and lag propulsion to hold at a given course and in given coordinates;
– hang at the point.

When using a mathematical dynamic model of AUV behavior in the event of an accident, one of the signs is the mismatch of the simulation results with the current values of the AUV motion parameters - the course, the components of the AUV motion speed, its roll and trim, immersion depth, angular velocities of roll and trim. If the analysis revealed a deviation of the current parameters from the calculated ones, then a decision is made on the inadequate functioning of the AUV.

As an emergency can be considered situations involving:

– with the failure of the stern rudders or the bow horizontal rudders (separately),
– with the appearance of negative buoyancy and the need for its compensation.

To prevent accidents can be used:

– control of equalizing-differential tanks;
– movement control with the help of bow horizontal rudders and stern rudders;
– control of the speed of the course of the AUV;
– AUV motion control using vertical and lag thrusters.

Sources of information on the internal state of the AUV (ADS)

Information about the state of the AUV hardware is transmitted as a result of ADS operation. ADS is designed for automatic and automated diagnostics of AUV at the operational stage, which guarantees the operability and reliability of AUV during the mission.

ADS AUV solves the following tasks:

– monitoring the functioning of the components and the AUV as a whole;
– control of the technical state of the AUV hardware and software;
– automated search of places of failures;
– troubleshooting for recovery;
– reconfiguration control;
– control and accounting of the operating time of devices, individual units.

The control of the technical condition (TS) consists in checking the compliance of the values of the parameters of the object of diagnosis with the parameters given in the technical documentation, and determining on this basis one of the specified types of technical condition (good, efficient, inoperable) of the object of diagnosis at a given time. When assessing the type of vehicle used criteria and indicators of technical failure AUV. Checked the flow of data to ADS from instrumentation to determine the emergency and localize it. ADS has a built-in hardware self-diagnostics subsystem for all AUV systems. The self-diagnostic subsystem is based on regular data acquisition from sensors. In the event of termination of the receipt of data, it is assumed that there was a failure of the sensor or communication with the sensor.

Since the diagnostics are ongoing, the emergency subsystem is constantly running.

According to the aggregate data SLS, NS and MPC can be a classification of an emergency. According to the results of the classification, control commands are developed, for example, to perform a maneuver to circumvent an obstacle or to change the buoyancy of the AUV - to prevent an accident. In the event of an accident, control commands are developed for the use of emergency and other technical means, which are transmitted to the technical means control subsystem (TMCS). The TMCS that are not included in other systems.

Such systems and devices, for example, may be:

– equalizing system,
– lifting and mast device;
– device of return of emergency ballast;
– power distribution device;
– emergency facilities (hydroacoustic beacon-responder, light-signal subsystem).

The task of classification by the apparatus of the emergency and emergency situations will be considered in more detail.

3. Classification of objects - potential sources of emergency

The classification of objects - potential sources of an emergency is made according to a number of signs formed using information obtained from the subsystems-agents of the AUV.

Classification of entanglement in the fishing net can be carried out on the combination of the following characteristics:

– detection by the SLS agent of the fishing net (usually at short distances);
– the discrepancy between the current position of the AUV and the position determined by a previously developed mathematical dynamic model;
– the absence of a change in the position of the AUV (according to the navigation subsystem) when the propulsion unit is running
– the lack of compliance of the real movement of the AUV with the results of dynamic modeling of the motion of the AUV with the same control commands and the parameters of the propulsive-steering complex.

The classification of an AUV into a “garbage island” can be performed by detecting a garbage “island” by means of SLS, since the density of the garbage island differs from the density of water and has characteristic signs of signal scattering by heterogeneous fragments. Obviously, after the discovery and classification of the garbage island, the SLS will look for ways to bypass it. If, nevertheless, the AUV is in the “environment” of this garbage “island”, then there will be an obvious discrepancy between the commands sent to the marching propulsion and steering complex and the AUV motion parameters, as can be seen by comparing the actual motion parameters of AUV with simultaneous mathematical modeling of dynamic behavior AUV using mathematical dynamic model.

Getting into the trash can of the AUV “island” can cause deformation of the AUV body or damage to it. As a result, a leak may occur, water will begin to flow into the AUV, as a result of which a change in the AUV trim may occur, due to an increase in buoyancy, the depth of the AUV immersion will change. All this will receive data from pressure and humidity sensors, which will show the deviation of indicators from the norm within the AUV. The movement of the AUV will slow down to a full stop of the AUV with the propeller running. There will also be a clear discrepancy between the current parameters of the movement of the AUV and the results of mathematical modeling of the dynamic behavior of the AUV.

Scrapping fragments on a screw or handlebars can be recognized:

– according to the discrepancy between the speed of rotation of the propeller and the speed of movement of the AUV;
– propulsion overheating (in case of winding on the screw or its jamming);
– changes in buoyancy in case of hanging garbage on the rudders and inadequate reaction of the AUV to the shifting of the rudders.

Inadequate behavior of the AUV is estimated using additional data on the parameters of the MPC propellers and steering wheels.

When the AUV moves in ice conditions, the ice crumb hits the propeller blades and, as a result, its efficiency can be reduced by comparing the behavior of AUV with a dynamic model. Damage to the hull can be identified according to the internal sensors in the AUV chassis: changes in pressure, temperature, humidity.

As a result of these factors, an emergency situation may occur on the AUV. Data about it will come from sensors and subsystems.

If an accident does occur, it is necessary to classify it in order to perform certain actions in accordance with the prescribed instructions.

An accident in the marching propulsion and steering complex subsystem can be classified as a result of analyzing the root cause of an accident, and its influence on the functioning of the AUV - based on a comparison of the actual motion parameters of the AUV and the parameters obtained from the results of mathematical modeling of the AUV dynamic behavior.

In addition to the listed signs of external character, a combination of external and internal characters is possible.

The failure of the energy supply subsystem is determined by internal sensors and parameters of devices and units of the power supply subsystem.

The failure of the main propulsion and steering complex can be generated by both external factors and internal ones. The external factors include the ingress of the AUV in an unfavorable environment - in an ice area, in the "trash island", in the fishing net. The internal factors include the abnormal work of the propulsion unit and the rudders, the lack of execution of the specified commands, or the impossibility of stabilizing the set position. At the same time, internal causes may be the result, including external causes, but - not necessarily.

Failure of the central on-board computer can be caused, most likely, by internal causes of the electrical network or overflow of resources of the central on-board computer - lack of memory, speed, etc. In addition, a malfunction of the central on-board computer can also be caused by increased temperature and humidity in the corresponding compartment, which should be clear from the data of these sensors.

The deviation of the AUV from the route trajectory to an unacceptable distance can be caused primarily by an increase in the error when calculating coordinates in the INS, the impossibility of conducting a timely observation, and also - the influence of external factors - drift by a current, falling into a large funnel, a forced deviation from the route trajectory (for example, to avoid obstacles) and the inability to return to it;

The ingress of the AUV into the fishing nets may be due to the inadequate detection range of the fishing net due to a low-power sonar or a net falling into the dead zone of the sonar, as well as the weak maneuverable characteristics of the AUV. For example, an attempt to bypass detected fishing nets may fail due to:

− weak maneuverability characteristics of the AUV;

− currents whose speed is comparable to the speed of the AUV;

The failure of equipment related to the payload depends, of course, on the nature of the payload, so in general, there can be a wide variety of causes, both internal and external.

Based on the specified features of the work of the emergency subsystem, let us proceed to consider the organization of its work as part of the management system of the AUV, built on a multi-agent basis.

4. Feature of multi-agent control system

The representation of the AUV control system as a multi-agent subsystem [5-7] is based on the allocation of subsystems in the control system, the main ones of which are (Figure 1):

− mission planning subsystem;

− subsystem lighting situation;

− navigation subsystem;

− subsystem hydroacoustic communication;

− radio communication subsystem;

− control subsystem of the propulsion and steering complex;

− power supply subsystem;

− subsystem control of technical means.

Each subsystem of the AUV control unit is an agent (Figure 1).

When using multi-agent technology in the control system of the AUV, the agents are the subsystems of the AUV control system (CS), which are inherently rational components whose characteristic features are:

− autonomy, i.e. the ability to act independently, controlling their actions and internal state;

– activity, i.e. the desire to achieve their goals;
– reactivity, i.e. adaptive behavior as a reaction to external influences;
– social behavior, i.e. interaction with other agents to reach agreed decisions;
– ability to self-study.

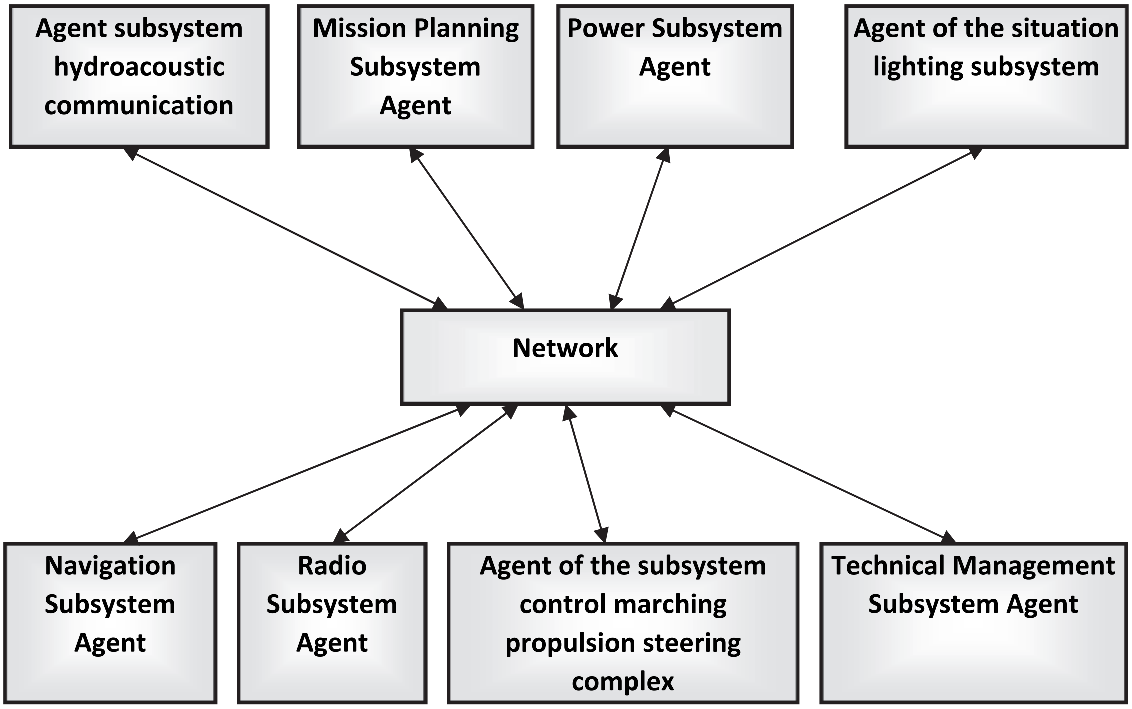


Figure 1 – Scheme of multi-agent control system

The multi-agent approach consists in the representation of each subsystem as an independent agent, making certain decisions depending on the current state and the conditions that have developed by the present moment. Each agent independently makes a decision regarding its actions, based on the information coming to it and the algorithms in place. The flow of information occurs as a result of mutual exchange between agents.

The operation of the multi-agent control system of AUV in the process of accomplishing the mission of moving AUV from one point to another is carried out as follows.

During the movement of AUV according to the data of motion sensors, the agent of the navigation subsystem constantly evaluates the numerical path, which it then recounts in absolute coordinates of AUV. AUV coordinates are transmitted to the route agent to estimate the AUV deviation from the route path. At the same time, the agent of the “power supply subsystem” constantly undergoes an assessment of the sufficiency of the remaining energy resource to fulfill the mission, taking into account the current energy consumption at the current speed limit of the AUV. The agent of the “power supply subsystem” compares the required volume of energy with the remaining one. If the remainder is close to its minimum, the agent of the “energy supply subsystem” calculates alternative movement options (reduction of electricity consumption by the AUV subsystems, reduction of the speed of movement, reduction of the route of movement) and gives the most suitable option to the mission planning agent. If one of the proposed options is adopted, then commands are issued to the relevant subsystems, for example, to reduce the propulsion speed - to the propulsive-steering subsystem, to reduce the radiation power - to the subsystem of situation lighting, etc.

From the above description, it follows that the subsystems-agents during the operation of the AUV communicate with each other, and each agent receives exactly the information that is necessary for it to function and make decisions. Accordingly, each subsystem-agent transfers the generated solution only to those agents who need it.

5. Implementation of the emergency subsystem

To organize the work of the emergency subsystem in a multi-agent control system, we consider three alternatives:

– formation of an additional agent;
– distribution of functions of the emergency subsystem by the existing subsystems-agents;

– placement of the emergency subsystem on one of the agent subsystem.

The approach associated with the creation of an additional agent of the emergency subsystem is inexpedient, since this specially created agent, in essence, will duplicate the functions of other subsystems in collecting and transmitting information.

The approach associated with the distribution of the function of the emergency subsystem among other subsystems of the AUV agents, will not allow to form an integral assessment of the emergency situation, based on the entire set of data obtained from all AUV systems. The third approach is to assign the functions of the emergency subsystem to one of the existing AUV subsystems, but this approach requires a weighted decision. This agent will always be in the active state, since the diagnostics of the state of the hardware is constant, even if the external environment does not create the prerequisites for the occurrence of an accident.

The approach associated with the placement of the emergency subsystem on one of the agent subsystem requires additional analysis of the AUV subsystem in which the placement of the emergency subsystem is advisable. The use of SLS is impractical because SLS is overloaded with data processing from hundreds of nose antenna sensors. The use of NS is also impractical, since the NS conducts a temporary processing of incoming data, their filtering, smoothing, integration of information coming from dissimilar in physical means.

It is inexpedient to use MPC, since it constantly tunes the parameters of the automated control system.

TMCS remains, and the use of TMCS seems most appropriate.

Of all the considered systems, only the TMCS has the by technical means control commands without in-depth analysis of their control and information processing. This is done by other systems - SLS, NS. In the TMCS, information from the sensors is not processed, except for the mast lift angle, etc. The TMCS is, in fact, an assistive system that is not burdened with complex analysis and decision making. In addition, the TMCS is engaged in the control of technical equipment that is not used in other subsystems.

Thus, the rise of the mast device occurs, for example, to ensure radio communications, a change in the water level in the leveling trim tanks occurs, for example, to change the buoyancy of the AUV, the light signaling system is turned on - to highlight the AUV and facilitate its visual tracking. Thus, the PUTS is most suitable for implementing an emergency subsystem in it.

When ADS is implemented in ITU, the status of all hardware of the AUV agent subsystems will be transferred to TMCS for ADS.

In addition, to obtain an integrated assessment of the emergency and emergency situations of the AUV, it is necessary to carry out the integration of external information from the subsystem agents.

Interaction with other agents-subsystems with which interaction was not previously provided, can be organized according to the existing protocols of information and technical interaction with subsystems with which other agents communicate - subsystems.

Thus, if all the information pertaining to pre-emergency and emergency situations is accumulated in the TMCS, then before the TMCS, along with the traditional control of technical means, additional nontrivial tasks will arise:

– receiving data from state sensors inside the AUV enclosure (temperature, pressure, humidity, etc.) in the AUV compartments;

– obtaining the current parameters of the AUV based on the results of modeling the dynamic position of the AUV;

– data acquisition from agents SLS, NS, MPC;

– integration of all received information to identify signs of a pre-emergency and emergency situation;

– classification of pre-emergency and emergency situations (fishing nets, garbage "islands", icebergs, volcanoes, whirlpools) and the state of the external environment (ice conditions).

As a result, the work of the emergency subsystem in the composition of the TMCS is as follows.

When data is received from the sensors of the internal state of the AUV that do not meet the standards, the state of the AUV is classified as abnormal.

When receiving signs of malfunction of individual devices from the subsystem-agents, the emergency subsystem of the TMCS first solves the problem of criticality of the failure of the device or element. If the failure is not critical, then the emergency is not registered, if critical, then the possibility of reserving and reallocating resources is assessed.

If this did not help or is impossible, or if all possible reserves are exhausted, and:

– AUV is not able to move;

– the central computer failed;

– AUV sinks;
– AUV turns over (overturns);
– AUV is not controlled;
– a fire or explosive situation has occurred;
– one of the elements of MPC has failed;
– the external environment is defined as unfavorable (ice, fishing nets, etc.),

then the PUTS solves the classification problem, identifying the type and source of the accident.

According to the results of the classification, a decision is made in accordance with ready-made scenarios of actions performed. The entire analysis of the pre-emergency is aimed at trying to eliminate them, and if it fails, then to allocate, for example, six types of emergency situations, each of which has clear instructions on how to turn on the relevant emergency equipment:

1) failure of the power supply subsystem;
2) failure of a marching propulsion and steering complex;
3) failure of the central on-board computer;
4) deviation from the route trajectory at an unacceptable distance;
5) getting the AUV in fishing nets;
6) failure of equipment related to payload.

What to do in each of these situations is prescribed in advance in emergency scenarios. There is a development of control commands that are transmitted to the relevant technical subsystems in the pre-emergency situation. For example, if, nevertheless, the AUV gets entangled in the fishing net, then with the help of special algorithms and devices, he independently tries to free himself and continue the mission. The elimination of collision or grazing of these structures with the hull is possible due to a sharp AUV maneuver to the opposite course with a decrease in the speed of movement.

Conclusion

Based on the analysis of the causes of accidents at the AUV and sources of detection of the occurrence of accidents in order to take timely measures, the analysis of the features of the multi-agent control system of the AUV - the implementation of the emergency subsystem as part of the technical control subsystem is proposed.

This allowed, in turn, to resolve the contradiction between the collection of information from all AUV subsystems and the peculiarity of the implementation of the multi-agent approach, in which the functioning of AUV subsystems is distributed.

The organization of the work of the emergency subsystem and the list of tasks assigned to it are defined. Such an organization of the emergency subsystem as part of a multi-agent control system provides:

– rejection of the formation of a special agent allows you to use the existing contacts with other subsystems.

– the ability to transfer data to other subsystems-agents in accordance with the same protocols as other subsystems for functional purposes, that is, nothing new needs to be developed.

References

1. Ageev, MD and others. Autonomous underwater robots. Systems and technologies. - M .: Science, 2005. 400 p.
2. Illarionov G.Yu., Sidenko KS, Bocharov L.Yu. Threat from the depths: XXI century. - Khabarovsk: KGUP "Khabarovsk Regional Printing House", 2011, 304 p.
3. Appolonov E.M., Bachurin A.A., Gorokhov A.I., Ponomarev L.O. On the possibility and necessity of creating an ultra-large uninhabited underwater vehicle // Collection of materials of the XIII All-Russian Scientific and Practical Conference "Perspective Systems and Control Problems". Rostov-on-Don - Taganrog, SFU. 2018. pp.34-42.
4. Matvienko Yu.V., Inzartsev A.V., Kiselev L.V., Scherbatyuk A.F. Prospects for increasing the efficiency of autonomous underwater robots // News of SFU. №1. Engineering science, 2016., p.123 - 141.
5. Innocenti B. A motion control. Ph.D. dissertation –Universitat de Girona, 2009, pp.147.
6. Martynova L.A. The use of multi-agent technology in the control systems of an autonomous uninhabited underwater vehicle // Promising information technology works of the International Scientific and Technical Conference. 2016. pp. 292-296.
7. Martynova LA, Mashoshin A.I. Building a control system for autonomous, uninhabited underwater vehicles based on multi-agent technology. Izvestia SFU. Technical science. 2016. № 2 (175). Pp. 38-48

8. Illarionov G.Yu., Laptev K.Z., Matvienko A.P. Additional requirements for autonomous uninhabited long-range underwater vehicles // Proceedings of the 7th All-Russian Conference "Technical Problems of the Development of the World Ocean", October 2-6. - Vladivostok. - 2017. - pp. 25-33.
9. Laptev K.Z., Illarionov G.Yu. What can prevent the autonomous uninhabited underwater vehicle from underwater navigation: // Collection of materials of the XIII All-Russian scientific-practical conference "Perspective systems and control tasks". Rostov-on-Don: Southern Federal University. 2017 p. 138 - 146.
10. Naumov, LA, Illarionov, G.Yu., Laptev, KZ, Babak, AV, On the Question of Planning Principles and Features of the Formation of Global Routes of Autonomous Underwater Robots // Izvestiya TSU. Technical science. Issue 11: at 2 pm. Part 2. Tula: Publishing House of TSU. 2015. 219c.
11. Why do submarines get into fishing nets? // Internet resource https://inosmi.ru/world/20050811/221492.html. The date of the appeal is 06.02.2019.
12. Faustova O. G. Development of a method of integrated assessment and risk control of emergency situations to improve the safety of ships. Diss. on the competition Uch.st.kand.tehn.nauk. - Kaliningrad. - 2016. 200 p.
13. Rodimova R.I., Shinkevich Yu.G. Automated systems for technical diagnostics of modern hydroacoustic complexes. // Hydroacoustics. 2015. №3. pp.75-87.
14. Topalov V.P. Risks in shipping / V.P. Topalov, V.G. Torsky. - Odessa: Astroprint, 2007. - 368 p.
15. Casualty statistics and investigations. (Very serious and serious casualties for the 2009). - London: International Maritime Organization, 2010. – P. 32-45.
16. Donald Waters. Supply chain management. Vulnerability and Resilience in Logistics. - London and Philadelpia. - Kogan Page Limited, 2007. - 264 p.
17. International Safety Management. Code and Guidelines on Implementation of the ISM Code. — London: IMO, 2012.
18. Yin J. Quantitative Risk Assessment for Maritime Safety Management / PhD thesis. Hong Kong Polytechnic University. — 2011. <http://repository.lib.polyu.edu.hk/jspui/bitstream/10397/4317/2/b24415613_ir.pdf>.
19. Karahalios, H. A risk appraisal system regarding the implementation of maritime regulations by a ship operator / H. Karahalios, Z. L. Yang, J. Wang // Maritime Policy & Management. — 2015. — Vol. 42. — No. 4. — P. 389–413.
20. Kobyliński L. Risk analysis and human factor in prevention of CRG casualties / L. Kobyliński // TransNav. —2009. — № 3 (4). — P. 443–448.
21. Kristiansen Svein. Maritime Transportation: Safety Management and Risk Analysis, 2005. – 528 p.
22. Underwater robots will patrol the world ocean // Internet resource http://zoom.cnews.ru/rnd/news/line/podvodnye_roboty_budut_patrulirovat_mirovoj_okean/print. The date of the appeal is 06.02.2019.
23. Kolesov N.V., Gruzlikov A.M., Lukoyanov E.V. Using fuzzy interacting observers for fault diagnosis in systems with parametric uncertainty // Procedia Computer Science – 2017. Vol. 103, pp. 499–504.
24. Gruzlikov A.M., Kolesov N.V., Lukoyanov E.V. Active diagnosis of a discrete event systems // International Workshop Navigation and Motion Control 2017. Proceedings - 2018. - pp. 47-55.

A.C. Golubev, O.V. Litvinov, A.V. Bakshiev, I.A. Vasilyev

RESEARCH OF THE APPLICATION REINFORCEMENT LEARNING METHODS IN THE TASK OF CONTROLLING THE REDUNDANT AUTONOMOUS UNDERWATER VEHICLE

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
alexab@rtc.ru

Abstract

The aim of this work is to analyse the limits of applicability of various deep reinforcement learning algorithms and study the methodology of object models integration into deep reinforcement learning framework in the task of designing autonomous underwater vehicle control systems. This work is concerned with existing approaches of designing autonomous mobile robot control systems, analyses deep reinforcement learning methods and frameworks that implement these algorithms.

Keywords: deep learning, neural networks, reinforcement learning, AUV, autonomous robot, robot control.

Acknowledgements

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 "Development and study of new architectures of reconfigurable growing neural networks, methods and algorithms for their learning".

1. Introduction

Despite the tremendous advances in the field of automation, there are still spheres of human activity, the automation of which is either extremely inefficient or not at all possible when seemingly very simple operations (which any unqualified personnel can easily cope with) cannot be automated at all. Many of these “automation-free niches” could be filled with highly adaptable and autonomous mobile robots (MR).

“Intellectualization” of MR is one of the most important direction of development of MR. This term here means increasing the level of MR adaptability to complex rapidly changing external conditions or increasing the degree of independence (autonomy) of the process of MR functioning from a human operator. With increasing degree of autonomy of the MR, the control of the robot is simplified, the negative influence of the human factor decreases, and the overall effectiveness of the use of MR increases. The main difficulties in this are to create algorithmic software that allows you to automatically control the movement of robots.

Taking into account the success of deep learning algorithms in various poorly formalized tasks, it was decided to analyze and compare different approaches (both on the basis of neural networks and classical non-neural networks) for designing of algorithms of autonomous MR control.

Thus, the urgency of the task of creating MR control system algorithms based on deep neural networks, on the one hand, is determined by the demand for autonomous MR, on the other hand, by the lack of efficient algorithms capable of solving a poorly formalized autonomous control problem for autonomous MR in various environments.

2. Analysis of approaches of autonomous robotic systems design

The most relevant approaches for designing autonomous robots control algorithms are the following:

– evolutionary algorithms [1];
– fuzzy logic [2];
– formal approach;
– reinforcement learning [3].

The mentioned approaches are considered in order to identify the best for solving an autonomous MR control problem by the following criteria:

– formalization complexity;
– ability to work in an unknown environment;
– application complexity;
– demands on computing resources.

Formalization simplicity is the most important criterion, since the main task is to simplify the existing methods for creating control algorithms. Work in a previously unknown environment is provided by the generalizing ability of the algorithm. The criterion of complexity of the application determines the complexity of the reproduction of the algorithm for new initial conditions and configurations of the robot. The criterion of requirements for computing resources directly affects the cost of applying the chosen approach. The results of the comparison of the approaches are presented in table 1.

Table 1. Comparison of approaches to the creation of control systems

Criterion	**Fuzzy logic**	**Formal approach**	**Reinforcement learning**	**Evolutionary algorithms**
Formalization complexity	High	High	Low	Medium
Generalizing ability	No	No	Yes	Yes
Labor costs	Medium	High	Low	Medium
Requirements for computing resources	Low	Low	Medium	High

In the evolutionary approach and reinforcement learning, takes place abstraction from the physical characteristics of the agents, because of which the labor costs of creating control algorithms for complex systems, i.e. on formalization, are reduced. Also, due to the fact that these methods do not use the description of a specific task, but only a function that contains general aspects of the algorithm, the algorithm itself has a large generalizing ability. For the work of evolutionary methods requires the simulation of a variety of agents, which greatly increases the requirements for computing resources. Describing the objective function in evolutionary methods and in learning with reinforcement is a simpler task than describing the dynamics and kinematics of a system. However, if we compare evolutionary algorithms and training with reinforcement, encoding the agent's genome in evolutionary algorithms is incomparably harder. In connection with the foregoing, the best of the considered approaches on the selected criteria is reinforcement learning.

3. Reinforcement learning

Reinforcement learning is one of the methods of machine learning, during which an agent (robot) learns by interacting with a certain environment. The agent actions a_t transfer the environment to the new state x_t and the agent receives from the environment some reward r_t or punishment in accordance with the reward function. The reward function defines the goal in the reinforcement learning process and is essentially a correspondence between environmental states and a number, reward showing the desirability, value of the condition. As a function of reward, there may be, for example, the distance from the robot to a certain point in space; the height of the top point of the walking robot (indirectly reporting that the robot did not fall); battery charge onboard battery. In this class of methods, much attention is paid to encouraging or punishing not only current actions that directly led to positive or negative reinforcement, but also those that preceded them. Therefore, the objective function is the sum of reinforcements over a certain interval of space, for example, total reinforcement for traversing a certain length. While the reinforcement function determines the direct, characteristic desirability of the state of the environment, the objective function detects the long-term desirability of the states after taking into account the states that follow the current one, and the reinforcements corresponding to these states. For example, a condition may entail low direct reinforcement, but it will have a very positive effect on the total score, because it is regularly followed by other conditions that bring high reinforcement. The only goal of the agent is to maximize the total reinforcement (objective function) that he receives in the course of long work.

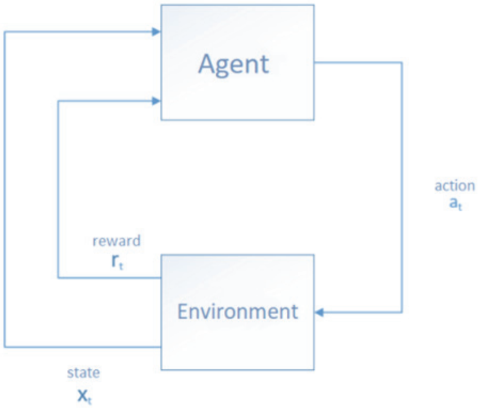


Figure 1 – Scheme of reinforcement learning

4. Comparison of reinforcement learning algorithms

Following reinforcement learning algorithms with analog control were selected for conducting an experiment:

– DDPG [4];
– PPO [5];
– TRPO [6].

And also 4 algorithms with discrete control:

– CEM [7];
– DQN [8];
– Dueling DQN [9];
– SARSA [10].

Experiments were conducted to compare reinforcement learning algorithms. Experiments for discrete and analog controls were separated. This separation due to existing different methods for 2 types of control. Task of pendulum balance was solved for discrete control and task of double pendulum balance were solved for analog control. Both type of algorithms were compared by similar criterion. Speed of learning determine by best episode with maximum reward. The less this value, the more learning speed. Best result determines quality of algorithm. The higher value of this criterion, the better quality. Stability criterion determine overfitting degree. In case value is «Yes» then quality of algorithm didn't decrease after reaching best value in experiment. Amount of learnings parameters affect on learning speed. The more parameters, the longer algorithm will learn. Table 2 was made by conducted experiments.

Table 2. Comparison of reinforcement learning algorithms

Algorithm	**Learning speed**	**Best result**	**Stability**	**Amount of learning parameters**
CEM	2500	35	Yes	658
DQN	80	200	Yes	658
Dueling DQN	140	200	Yes	658
SARSA	450	160	No	658
DDPG	1500	2800	No	9000
PPO	4200	7500	Yes	4485
TRPO	3000	6200	Yes	4485

According to the results of the experiment, it was established that DQN and Dueling DQN, which showed similar results, are the best reinforced learning algorithms for discrete control. The DQN algorithm learned quickly enough - it took him 60 episodes less than the Dueling DQN, so we can conclude that the DQN algorithm is the best in this experiment. For analog control, the DDPG algorithm showed the highest learning rate, but its maximum result was 2.2 times less than the TRPO result, and 2.7 times less than the PPO result. The PPO and TRPO algorithms were not overfitting in 4500 episodes, while the DDPG began to overfitting immediately after its maximum value, after 1500 episodes. The best algorithms for analog control are the PPO and TRPO algorithms, if the major criterion is quality, and if the major criterion is learning speed, then DDPG better.

5. Framework selecting for agent model and environment realization

Reinforcement learning involves an agent which locate in an environment. Each step the agent takes action which return a reward and new state of the environment. The agent performs a certain number of actions from the starting position (which in general is random), that make up one episode of the training. The natural goal of reinforcement learning algorithms is to maximize the total reward that an agent receives during interaction with the environment during an episode. Validation of algorithms can occur during test runs, when the parameters of the algorithms are fixed, and the environment conditions are as close as possible to the required ones. The methodology described above is implemented in frameworks and libraries for different programming languages. This paper will discuss about the most popular ones for the Python language.

The selection of the toolkit will be based on the following criteria:

- complexity of use for a new user;
- a variety of built-on learning reinforcement learning algorithms;
- a variety of different environments;
- performance of implemented algorithms.

The following frameworks are considered for agent and environment modeling:

- OpenAI Gym + keras-rl;
- RLlib;
- TensorForce;
- Intel's Coach;

The results of the comparison are presented in table 3.

Table 3. Comparison of reinforcement learning frameworks

Criterion	**OpenAI Gym + keras-rl**	**RLlib**	**TensorForce**	**Intel's Coach**
Complexity	Low	High	Middle	High
Variety of algorithm	Middle	High	Very high	Very high
Variety of environments and agents	High	High	Very high	High
Performance	High	Very high	Middle	Middle

The complexity of use was the highest priority criterion due to limited time. The complexity of use is made up of the volume of documentation, the number of different functions and their complexity, the presence of various examples in the network. RLlib and Intel's Coach due to various optimization of computations are difficult to use in this work. OpenAI Gym in the keras-rl bundle has an advantage over TensorForce by this criterion due to the large number of examples in the network and the larger community.

A variety of algorithms made possible to test different approaches to solving the same problem, but this is not a key criterion, since the objectives of this paper included testing only the basic, most popular algorithms.

A variety of environments and agents made it possible not only to identify the features of the algorithms, but also to study more examples before implementing own model. At the moment, all of the considered frameworks have a large set of supported environments and agents, however, extensive research will be preferable to TensorForce.

The performance of the algorithms also strongly affected on the choice of the framework, since even the simplest agents, such as the inverse pendulum, learn a significant amount of time. Therefore, the possibility of using unstable algorithms in this work was rejected. As for the possibility of parallel computations in the RLlib library, this approach seems to be unreasonably difficult in this task.

According to the results of the analysis, the best of the considered frameworks within the framework of this work is OpenAI Gym in conjunction with keras-rl.

6. AUV model realization

In this work, as an example, the MR is an autonomous underwater vehicle (AUV). The AUV mathematical model described below has five degrees of freedom. To simplify the model, we assume that the AUV has a torpedo shape. Device has five thrusters:

– sustainer thruster (p. 1 fig. 2);
– horizontal thruster (p. 2, 3 fig. 2);
– vertical thruster (p. 4, 5 fig. 2).

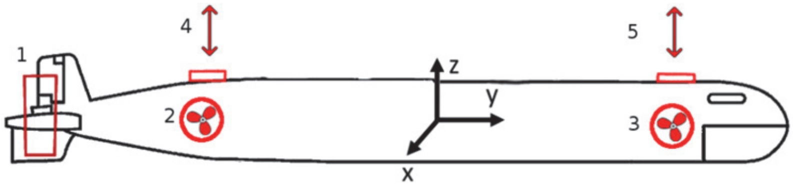


Figure 2 – Schematic image of the AUV

Local coordinate system rigidly connected with the AUV was introduced. The center of mass is the geometric center (see Fig. 3) and coincides with the center of the axes of coordinates. We denote all the forces acting on the AUV. All thrusters create traction forces (see Fig. 3). The traction forces of the rear propulsive thruster are applied to a point separated from the center of coordinates by l/2 in the negative direction and located on the y axis. Where l is the length of the AUV. The traction forces of the forward thrusters are applied to a point located on the Oy axis and spaced l / 2 in the positive direction. The resistance force Fc (1) is also applied to this point.

$$F_c = C_x S(\theta) * \rho * v^2 \tag{1}$$

where C – dimensionless forward resistance coefficient;

S(θ) –Midels cross-section area;

ρ – density of water;

V – velocity in S orthogonal direction.

The gravity force mg and the force of Archimedes Fa (2) are applied to the center of gravity.

$$F_a = \rho g V \tag{2}$$

where ρ – density of water;

g – acceleration of gravity;

V – AUV value.

Also moments from these forces act to this system.

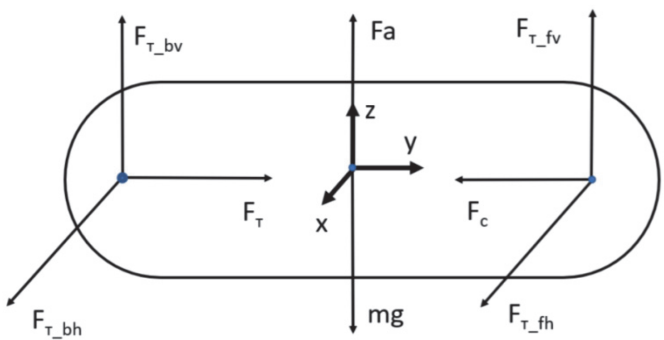


Figure 3 – Schematic image of the AUV with forces acting on it

Make equations for forces and moments in vector form. (3, 4):

$$\vec{F} = \vec{F}_{t\_fv} + \vec{F}_{t\_fh} + \vec{F}_{t\_bv} + \vec{F}_{t\_bh} + \vec{F}_t + m\vec{g} + \vec{F}_a + \vec{F}_c \tag{3}$$

$$\vec{M} = \vec{M}_{t\_fv} + \vec{M}_{t\_fh} + \vec{M}_{t\_bv} + \vec{M}_{t\_bh} + \vec{M}_t + \vec{M}_c \tag{4}$$

Make projections of the received equations on axes of coordinates (5, 6):

$$\begin{cases} F_x = F_{t\_fh} + F_{t\_bh} \\ F_y = F_t - F_c \\ F_z = F_{t\_fv} + F_{t\_bv} \end{cases} \tag{5}$$

$$\begin{cases} M_x = {}^{l}/_{2} * (F_{t\_fh} - F_{t\_bv}) \\ M_y = 0 \\ M_z = {}^{l}/_{2} * (F_{t\_fv} + F_{t\_bh}) \end{cases} \tag{6}$$

To determine the coordinates in the global system, it is necessary to translate forces using the rotation matrix R (α, β, γ), where α is the trim angle, β is the roll angle, and γ is the course angle.

Acceleration can be obtained through the resultant force, and angular accelerations can be obtained through the resultant moment. Speed, angular velocity, position and angles in the global coordinate system can be obtained further differentiation.

To use the AUV model, it is necessary to determine, in addition to the mathematical description of the laws of motion, the observation function, the action function, the reward function, and the reset function. To do this, you need to create a class inherited from the class «gym.Env». In the new class the mentioned functions should be redefined:

7. Conclusion

In that paper we discover and compare approaches to the development of control systems and reinforcement learning algorithms. We analyzed existing frameworks for creating environments and agents and reinforcement learning algorithms libraries. An AUV model was created for use in the OpenAI Gym framework.

References

1. L.Silva, A.Silva An Evolutionary Algorithm for Autonomous Robot Navigation // The International Conference on Computational Science. – 2013.
2. L.Zade The concept of a linguistic variable and its application to approximate reasoning Mir, – 166 p, – 1996.
3. C.Florensa, D.Held Automatic Goal Generation for Reinforcement Learning Agents // arXiv:1705.06366. – 2018.
4. S.O.Nikolenko Deep learning. Deep learning. Immersion in the world of neural networks. Piter. – 480 p. – 2018.
5. J.Schulman, F.Wolski, P.Dhariwal, A.Radford, O.Klimov Proximal Policy Optimization Algorithms // arXiv:1707.06347v2. – 2017.
6. J.Schulman, S.Levine Trust Region Policy Optimization // arXiv:1502.05477v5. – 2017.
7. D.P.Kroese Cross-Entropy Method // arXiv:1503.01842v1. – 2015.
8. V. Mnih Human-Level Control through Deep Reinforcement Learning // journal Nature, 2015.
9. Z.Wang T.Schaul Reinforcement Learning : An Introduction // arxiv.org:1511.06581. – 2016.
10. R.Sutton, A.Barto Dueling Network Architectures for Deep Reinforcement Learning // The MIT Press. – 2017.
11. Roboschool - URL: https://github.com/openai/roboschool
12. Yuval Tassa MuJoCo: A physics engine for model-based control, 2012 IEEE/RSJ International Conference

D.A. Frolov, D.A. Gromoshinskiy, A.M. Korsakov, E.Yu. Smirnova, A.V. Popov

DETECTION OF UNDERWATER METAL-CONTAINING OBJECTS WITH FUSION OF FLUXGATE SENSORS WITH NAVIGATIONAL DATA

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
d.frolov@rtc.ru

Abstract

With the underwater human activity actively developing and the rise of the autonomous unmanned underwater vehicles (AUV) industry, researchers and engineers face the task of maintenance and repair of communications. For this purpose AUVs can be used, with e.g. the task of inspecting pipelines. The article describes the processing algorithm for the signal from passive fluxgate sensors mounted on an AUV carrier used to search for metal-containing objects at the sea bottom. A scheme for such a measurement is proposed - the installation of two sensors at the opposite ends of the carrier. This allows to measure the gradient of magnetic field between the sensors. The characteristic form of such a signal and the dependence of the signal on the motion parameters of the vehicle and external factors are determined. To eliminate false positives, filters are used based on the readings of the position, speed and orientation sensors of the navigation system. Using data on the motion parameters of the device allows to generate a reference signal, which is used to validate the detection of an object using the cross-correlation method. The use of data on orientation angles makes it possible to compensate for the influence of the orientation of the device in the Earth's magnetic field.

Keywords: AUV, UUV, signal processing.

Acknowledgments

This work was done under the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 "Investigation of the methods for creating a multifunctional modular reconfigurable hyper-redundant uninhabited underwater vehicle for the integration into the robotic complex of three environments".

1. Introduction

In the course of repairing or replacing undersea infrastructure objects passing along the sea bottom, there are problems of their detection and assessment of their length. To solve the problems of detecting metal-containing objects (gas pipelines, etc.) there exist two main types of electromagnetic sensors: active and passive. For an active detection method, additional energy is required for irradiation of the scanning signal into the surrounding space. At the same time, the active detection principle is characterized by resistance to external sources of magnetic fields (MF), both environmental or created by the carrier of the magnetic sensor. On the experimental AUV electromagnetic sensors of the passive type (fluxgate sensors) were installed. In contrast to the active, the passive sensors require compensation for the effects of external magnetic fields, which can have a significant impact on the readings of the probes. To search for metal-containing objects, it is proposed to use two passive three-component ferromagnetic probes mounted on the bow and stern of the vehicle. The presence of two probes separated by a known distance from each other allows to calculate the gradient of the magnetic field between the probes by simply subtracting the readings of the sensors. Calculation of the gradient compensates the influence of external magnetic fields, the impact of which on each of the sensors is the same. When crossing a metal-containing object, it is firstly registered by the bow sensor, and then on the stern one. In this case, a characteristic pattern is observed in the gradient signal — the alternation of the maximum and minimum (minimum and maximum) of the gradient. The presence of such a pair of extremes is a necessary condition for the detection.

2.Signal processing pipeline

2.1. Initial signal processing

The algorithm for detecting an extended metal-containing object is based on the detection of the intersection of a metal-containing object by the vehicle. For this, the pre-filtered and corrected signal is compared with the reference signal using the cross-correlation method. To reduce the computational load, the signal before the cross-correlation is pre-tested for compliance with the necessary criteria. The scheme of the algorithm is shown on Fig.1.

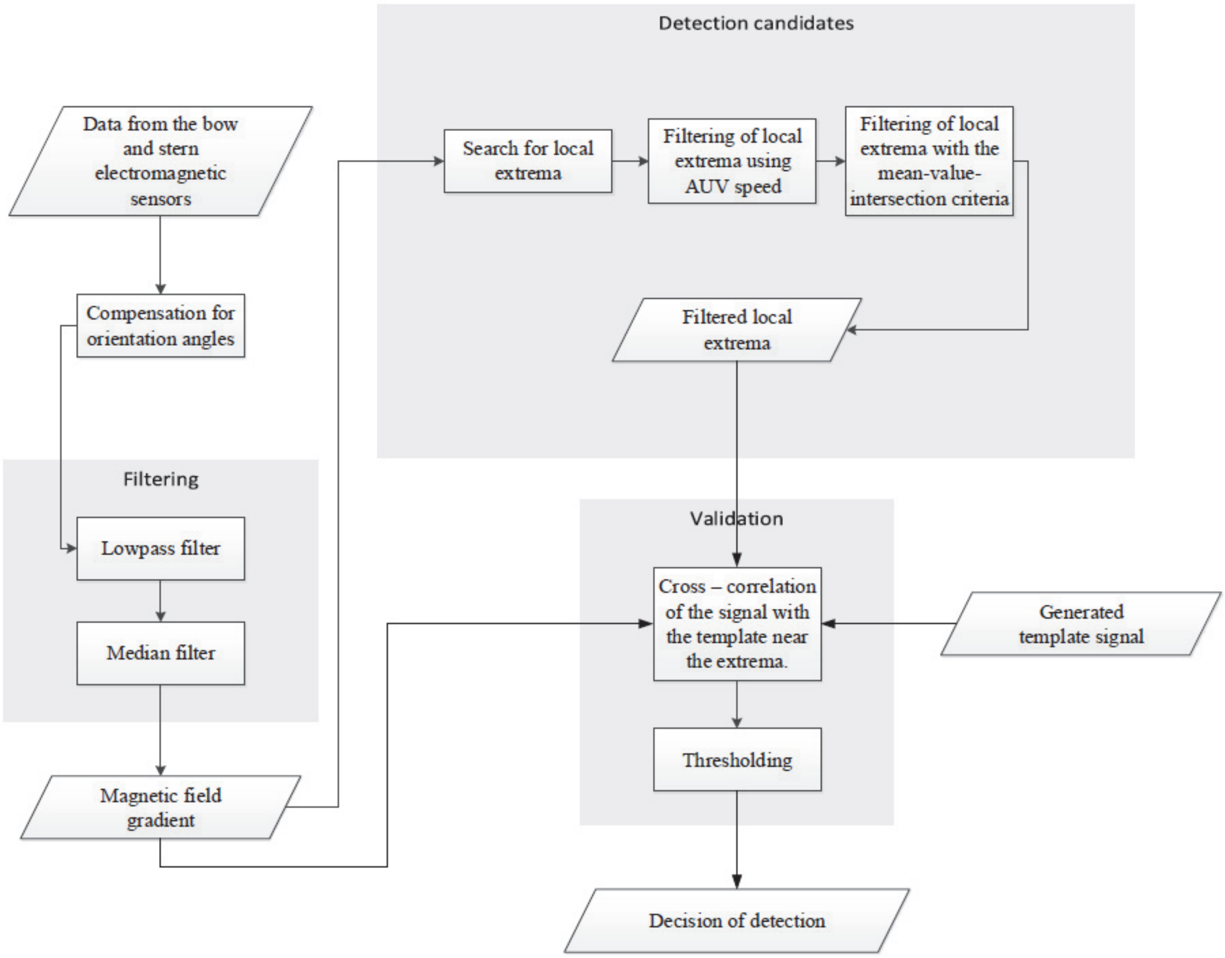


Figure 1 – Structure of the signal processing algorithm

To compensate for the influence of the orientation of the apparatus in the Earth's magnetic field, the coefficients of dependence of sensor readings on course, roll and trim angles were determined using the linear regression method. The dependence on heel and trim was linear, and on the heading it was harmonic. In the bare signal from a passive ferromagnetic sensor with a sampling frequency of 50 Hz, a constant harmonic component with a frequency of 4 Hz was also detected, for filtering of which a low-pass filter is used. To smooth the signal and eliminate sharp outliers, a median filter is applied to the signal. On Fig. 2 the filtered signal is shown. After orientation compensation, all values are given in arbitrary units.

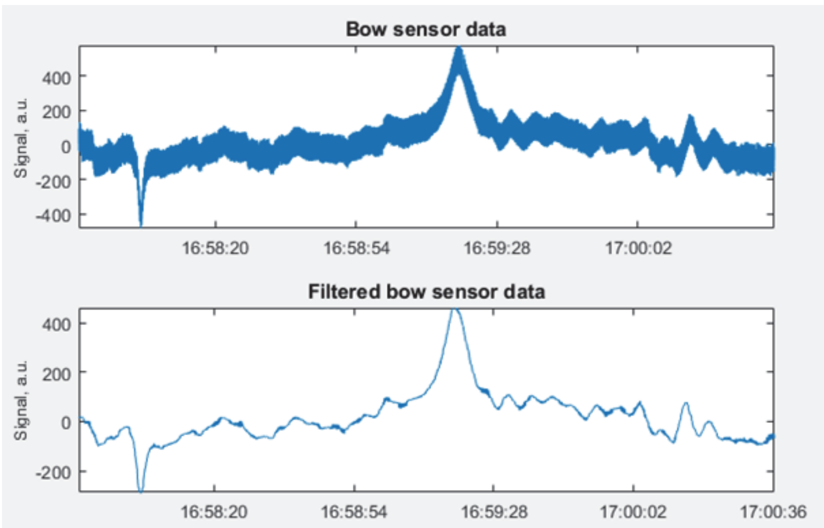


Figure 2 – Fluxgate sensor signal before and after lowpass and median filters

Then, calculating the difference between the readings of the bow and stern sensor, we obtain a signal that has the meaning of the gradient of the magnitude of the magnetic field along the vehicle. The signal of the gradient on passing over a metal-containing object has a characteristic shape - two alternating extrema. After that the local extrema of the received signal are searched for. Local extrema are filtered by the width of the peaks that they form. Fig. 3 shows the result of a search for the local extrema of the magnetic field gradient along the vehicle. The local maxima are marked with the red cicles and the local minima with blue. To filter such a pair of extrema, it is assumed that the time between them should be approximately equal to the speed of the AUV (it is assumed that it is constant) multiplied by the distance between the bow and stern fluxgate sensor.

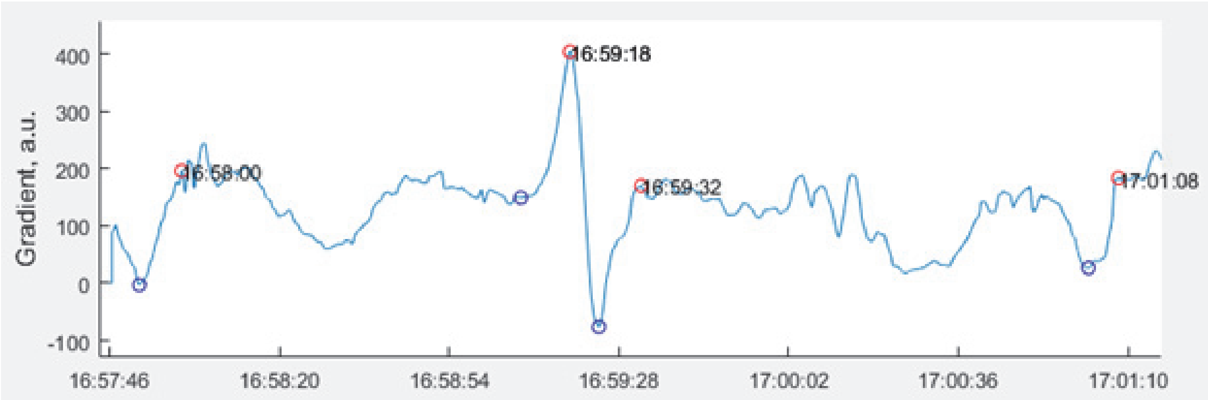


Figure 3 – Local extrema search result example

To filter by the intersection of the mean value criteria, a similar pair of points described earlier is taken. For each such pair of points, a check is made on the intersection of the mean gradient value by the line connecting this pair of points. A pair of points for which this condition is not fulfilled is eliminated. Then, for each pair of extrema found, validation is performed.

2.2. Detection validation

The task of searching for a signal of a known form in time series arises in many areas of knowledge. In medicine such is the task of identifying complexes in signals of various vital indicators, for example, a QRS complex in a cardiogram [1], or patterns in an electroencephalogram [2], and in modern physics this problem was solved in experiments that recorded gravitational wave signals [3]. This problem can be solved in different ways: using filter combinations [4], using wavelet transform [5], using machine learning methods and neural networks [6]. All these methods are briefly reviewed in [1]. We have chosen a matched filter for our purpose for its incomplexity and fast implementation.

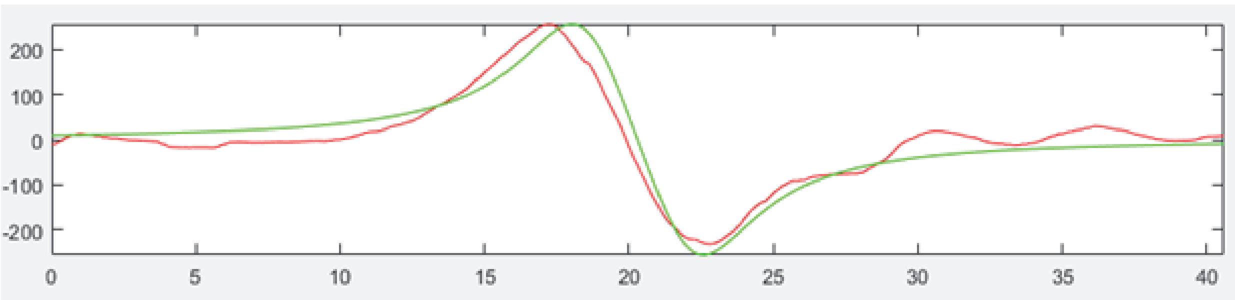


Figure 4 – A template signal (green) and a real one (red)

The matched filter is used to search for a known sequence (template) in the signal by correlating the template with the signal being studied. The maximum of the cross-correlation function then marks the point in time that corresponds to the appearance of the template signal in the analyzed one. With the appropriate normalization, it is not necessary for the sample and template signal to have equal magnitude, since the cross-correlation function will always have values in the -1: 1 range, but it is necessary that the pattern and the signal under study have the same sampling frequency. In this case, false detections can be filtered by setting the threshold of the cross-correlation function at the maximum, which is determined experimentally.

As a result of analyzing the experimental data from the device, a hypothesis was put forward about the shape of the sample signal: the signal from each sensor must be proportional to 1 / r, where r is the distance to the object, and the template signal is the difference of such functions. The template is defined by two dynamic parameters — the speed of movement and the impact parameter (height above the sea bottom), and one static parameter — the distance between the sensors. An example of such a generated signal is shown in Fig. 4, in this case, the vehicle speed was 0.62 m/s and the distance to the sea bottom was 1.34 m. At the time of receiving the information about the speed and position of the device from the navigation system, the metal-containing object detection system changes the template signal accordingly, which allows filtering out external noise that does not fit the shape.

3. Experiments

3.1. System performance

We performed 23 experiments in which the vehicle moved uniformly and rectilinearly, with an average speed of about 0.5 m/s and at a distance of 1.0 -– 2.5 meters above the sea bottom. The metal pipe and the other metal objects were placed on the sea bottom. The sea bottom was also filmed by a video camera, which made possible the markup of the data to count the number of true detections, the number of false positives and the number of false negatives. The presence of a cross-correlation maximum with a value greater than the threshold value 0.7 was considered the detection for the algorithm with validation. The choice of threshold value was performed heuristically, based partly on the uncertainties in the readings of the navigational system. The detection without validation was considered to be the result of the work of the candidate selection algorithm in Fig.1. Since the output of the algorithm without validation is a pair of extrema and several pairs can be found on the true signal, then each pair of extrema was taken as a detection, and not the fact of the presence of at least one pair. This allowed to correctly account for the false detected pairs of extrema.

Table 1. Experimental results

Alg. Type	True pos.	False pos.	False neg.	Precision
1	67	85	0	44.1
2	27	9	5	65.9
3	63	60	0	51.2
4	31	5	2	81.6

The comparison involved 4 different combinations of candidate selection and analysis algorithms, the results are presented in Table 1. Numbers in the first column correspond to: algorithm without orientation correction and form validation (1), algorithm without orientation correction with form validation (2), algorithm with orientation correction without form validation (3), and algorithm with orientation correction and form validation (4).

3.2. Impact of the vehicle control on system performance

We have also conducted experiments to study the influence of electric engine control and performance on the precision of our system. Non-stationary magnetic field generated by the unstable engine rotation speed may account for the noise and false positives in detection system, especially when the modulation introduced is of a temporal scale similar to that of the template signal. On Fig.5 and Fig.6, histograms of engines rotation speed (radians per second) are plotted with FFT spectra of rotation speed signal and the template. The data for this analysis was selected only from the linear parts of the vehicle trajectory. Engine performance data collected during experiments from Table.1 is shown on Fig.5. It is clearly seen, that stable performance of all engines may not impose any noise and false positive detections, since its Fourier image has negligible intersection with that of the template.

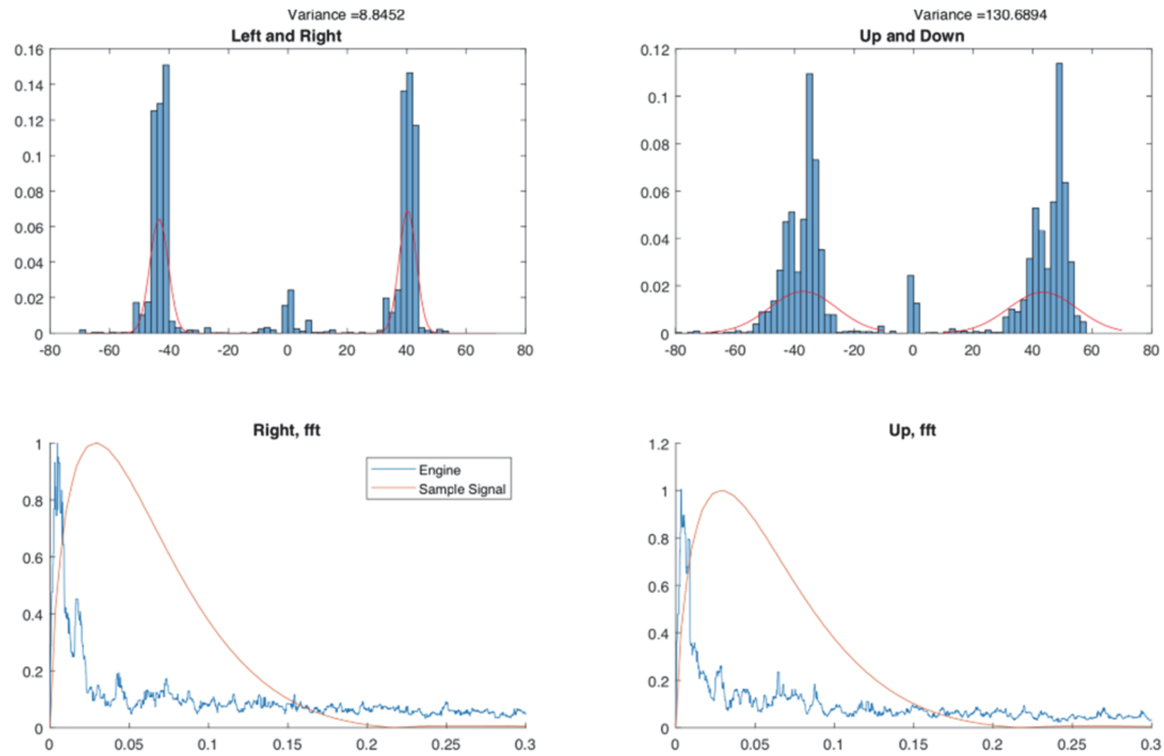


Figure 5 – Engine performance in experiments from Table 1

Because the computation of cross-correlation involves multiplication of signals` Fourier images, one might expect that introducing signals with spectra similar to the template will impact the system`s precision. And really, our experimets have shown that introducing the sinusoidal component to the vehicle control system results in dramatic drop of the system precision performance from 81.6% to 38%, and at these conditions the whole system becomes unreliable. Thus we have determined that providing the uniform constant motion of the vehicle is crucial for the performance of the system.

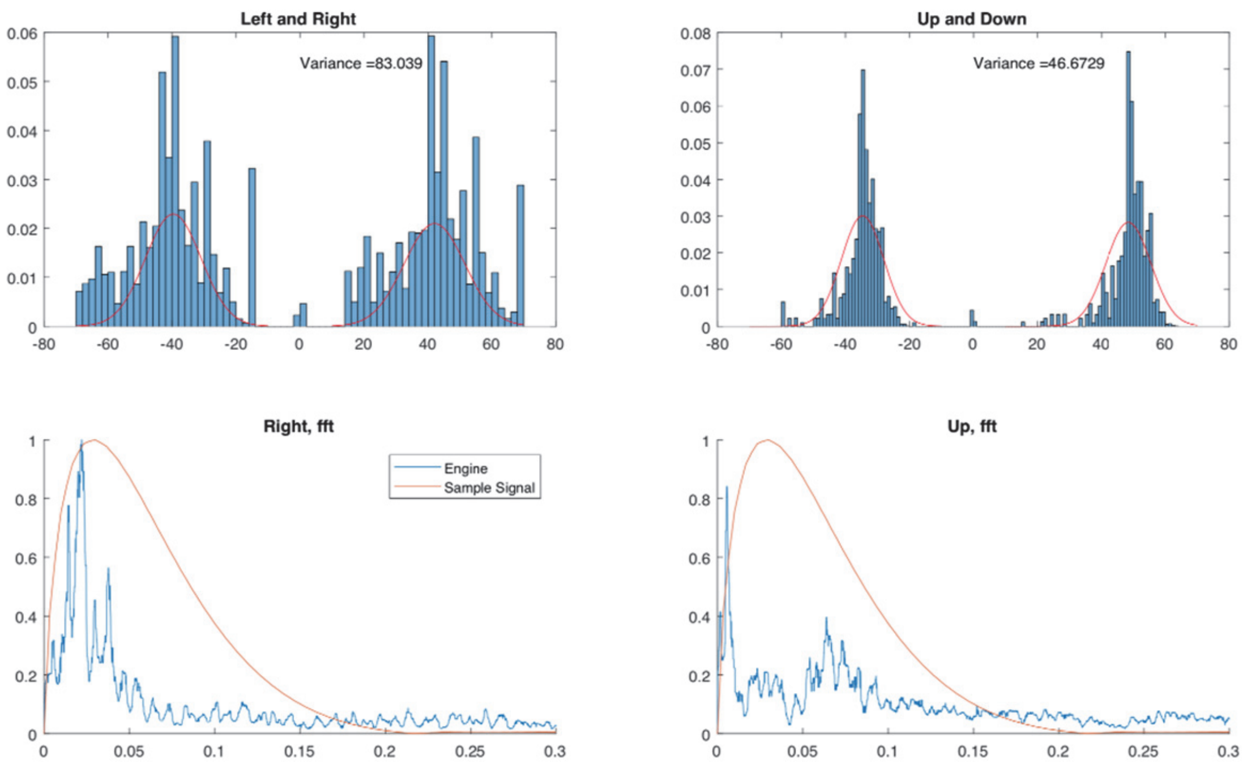


Figure 6 – Engine performance with sinusoidal component introduced

4. Conclusions

The paper proposes a measurement scheme and a signal processing pipeline for determining the presence of metal-containing objects at the sea bottom. A description of the characteristic features of the signal to identify it in the data of fluxgate sensors is given. According to the results of the experiments, one can conclude that the algorithm for finding pairs of extremas and filtering them by distance and amplitude is not sufficient for the qualitative detection of an object, and for determining the time of its intersection by the vehicle. Using a matched filter allows to determine the point in time at which the test and the signal under study coincide as much as possible, as well as significantly reduce the number of false detections. The paper also discusses the influence of vehicle control quality on the performance of the detection system. It is shown that introduction of periodic fluctuations in the control system has a negative effect on the precision of the detections, lowering it from 82 to 38 percent, which is an undesirable and unsatisfactory result. In this regard, the authors emphasize that ensuring a straight and uniform movement of AUVs is crucial in the search mode for a metal-containing object using the gradient method with fluxgate sensors.

References

1. Kohler, B. U., Hennig, C., & Orglmeister, R. (2002). "The principles of software QRS detection" *IEEE Engineering in Medicine and Biology Magazine*, **21(1)**, 42-57.
2. S. Bulárka and A. Gontean, "EEG pattern recognition techniques review," *2015 IEEE 21st International Symposium for Design and Technology in Electronic Packaging (SIITME)*, Brasov, 2015, pp. 273-276.
3. Abbott, Benjamin P., et al. "Observation of gravitational waves from a binary black hole merger." *Physical review letters* **116.6** (2016): 061102.
4. S. Suppappola and Y. Sun, "Nonlinear transforms of ECG signals for digital QRS detection: A quantitative analysis," IEEE Trans. Biomed. Eng., vol. **41**, pp. 397-400, 1994.
5. Kadambe, S., Murray, R., & Boudreaux-Bartels, G. F. (1999). Wavelet transform-based QRS complex detector. *IEEE Transactions on biomedical Engineering*, ***46*(7)**, 838-848.
6. Neocleous, A., Azzopardi, G., Kuitems, M., Scifo, A., & Dee, M. (2019). Trainable filters for the identification of anomalies in cosmogenic isotope data. *IEEE Access*., **7**

I.A. Vasiliev[1], A.A. Nikiforov[2]

MODELING OF THE UNDERWATER APPARATUS WITH A VARIABLE THRUST VECTOR, EQUIPPED WITH BALLAST TANKS

[1] The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia, vas@rtc.ru
[2] Peter the Great SPbPU, Saint Petersburg, Russia, pilrause@gmail.com

Abstract
The article describes the principles for controlling a mobile robot of underwater use moving in an environment and having redundancy associated with different control capabilities: variable thrust vector of the cruise propulsion, four ballast tanks and a set of thrusters for accurate maneuvering in the narrow areas.
Keywords: mobile robot, control, inverse dynamic task.

Acknowledgments
The work was carried out within the framework of the fulfillment of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 "Investigation of ways to create a multifunctional modular reconfigurable hyper-redundant uninhabited underwater vehicle for integration into the robotic complex of three home-based environments".

Introduction
Over the past few decades, interest in the use of unmanned robots for underwater operations has been steadily increasing. Currently, throughout the world, the autonomous underwater vehicle (AUV) two kinematic schemes are used [1]. The first is the already established classic scheme of underwater robots: the propeller and thruster units (TU). Vertical thrusters are used to control the ascent and the dive, and horizontal launchers are used for maneuvering along the course. Such a scheme has noticeable difficulties [1].

The second is the use of rotary launchers without a propulsion unit and tanks. This scheme is used mainly for vehicles that do not need to move over considerable distances, and only need to maneuver almost in place.

In this paper, an apparatus having a single propeller (sustainer propulsion — SP), which has a variable stop vector, that is, can be rotated around the vertical and horizontal axes to create moment on the course and the trim of the apparatus, is considered and modeled. Also, for controlling the AUV, the unit is equipped with four ballast tanks (BT) for trim and roll: in the bow to the left and right of the longitudinal axis and, respectively, in the stern, also to the left and right of the apparatus axis (see Fig. 1) and vertical thrusters. To control the course additionally applied horizontal TU.

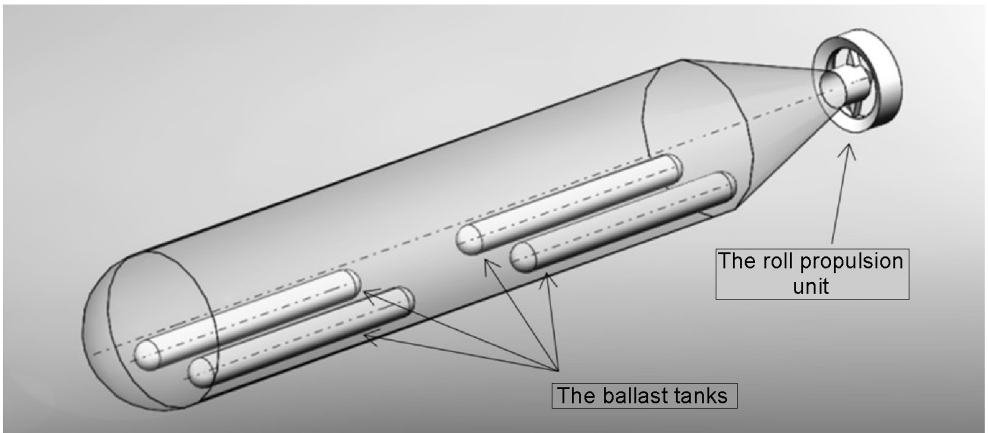


Figure 1 - Kinematic scheme of the AUV (TU not shown)

1. The idea of a control model for the general case
Let there be a control object (CO), described by the following differential equation:

$$U(t) = F(t, y, y^{(1)}, y^{(2)} \dots, y^{(n-1)}) \quad (1)$$

Where:
$U(t)$ – is a control depending (generally speaking) on time;

t – is time;

$F(*)$ – is a control function;

$y = y(t)$ – is a controlled value;

$y^{(i)} = y^{(i)}(t)$ – is the *i*-th derivative with respect to time ($i \leq n - 1$).

By controlled variable we mean three parameters: position, speed (derived from position) and acceleration (derived from speed). Therefore, in the system of equations (1), the *i*-th derivative may be negative — that is, we will assume that $y^{(-1)}$ – is an integral of y(t), and $y^{(-2)}$ - is an integral of integral, etc. That is, if the controlled quantity is a position. That is, if the controlled quantity is a position, then

$$i \geq 0$$

if controlled by speed –

$$-1 \leq i \geq +1$$

By controlling the acceleration, we have

$$i \leq 0.$$

The task is classic: to create such a control so that the mismatch *e* between the task y_g and the current position (working off) y_c is minimal. Consider some current state of the CO, which we call the initial:

$$e(0) = e_0$$
$$\ldots \tag{2}$$
$$e^{(i)}(0) = e_{i0}$$

Where:

e, e_0 – is the mismatch and the current mismatch of the system;

$e^{(i)}, e_{i0}$ – is the *i*-th derivative of the time error.

Naturally, the presence of task constants does not change the magnitude of the derivatives; therefore, it is logical to rewrite the system of states (2) in the following form:

$$e = y_g - y_c$$
$$\ldots \tag{3}$$
$$e^{(i)} = y^{(i)}(t)\Big|_{t\,=\,0}$$

To solve the problem, it is logical to try to reduce all current error statements to zero. Here you can offer a parameterized function $y(t) = f(t)$, which will give the path of the CO to the target in the phase space, and the number of parameters is equal to twice the order of the differential equation (1) plus 2, that is, 2*n*. To determine these parameters, it is necessary to solve the system of equations naturally arising from system (3). But in the system (3) there are only *n* equations, therefore it is required to expand the system (3) by the conditions at the end of the regulation cycle (RC):

$$e(t_1) = 0$$
$$\ldots \tag{4}$$
$$e^{(i)}(t_1) = 0$$

That is, to determine the parameters of the function $f(t)$, it is required to solve a system of 2*n* equations describing the states of the OU at the initial and final moment of time, i.e. in systems (3) and (4), substitute the corresponding derivatives of *f*.

Having thus obtained the function of the controlled parameter on time, $y = f(t)$, we substitute it into equation (1). We get the required control.

There is an additional complexity. The point is that the rate of change of control *U(t)* is limited. Therefore, the only control cycle, most likely, will not be able to fully extinguish the mismatch. Hence the additional parameter t_1 – the time to perform the operation. It can be determined from the expression

$$|\dot{U}(\tau_{iэ})| \leq U_{1max},$$
$$\tau_{iэ} : \ddot{U}(\tau_{iэ}) = 0$$

Where U_{1max}, is, naturally, the maximum control growth rate. And yet: the management of $U(t)$ itself is limited in magnitude. That is, we have another expression to determine the control phase trajectory

$$|U(t_{iэ})| \le U_{max},$$
$$t_{iэ} : \dot{U}(t_{iэ}) = 0$$

However, this cannot be done, that is, not to find time to perform the operation t_1. The fact is that the real drive device of the control object already has both a certain certain overclocking characteristic and a certain maximum value, and not exceed them. Moreover, as shown below, even the knowledge of these characteristics is not strictly required.

Now we will consider a real digital control system with a control cycle, which we denote here by Δt. Suppose that at each control cycle we believe that we are being controlled from scratch. That is, the function $y(t)$ is "decoupled" from the time parameter t and made constant at each clock cycle, where the time for the RC is the time for a single clock.

Let also as the phase trajectory of the controlled variable choose a polynomial function

$$y(t) = at^5 + bt^4 + ct^3 + dt^2 + et + f, \tag{5}$$

because the polynomial curve of odd degree can be adjusted not only for the required values, but also for the required derivatives at the ends of the segment.

Next, we compose two systems of equations (3) and (4) and, after simplifications, we obtain the result

$$y(t) = \frac{t^2 a_0}{2} + e_0 - \frac{6t^5 e_0}{t_1^5} + \frac{15t^4 e_0}{t_1^4} - \frac{t^5 a_0}{2t_1^3} - \frac{10t^3 e_0}{t_1^3} + \frac{3t^4 a_0}{2t_1^2} - \frac{3t^3 a_0}{2t_1} + tv_0 - \frac{3t^5 v_0}{t_1^4} + \frac{8t^4 v_0}{t_1^3} - \frac{6t^3 v_0}{t_1^2} \tag{6}$$

Where e_0 – is a mismatch of the controlled variable;

v_0 – is a current derived variable;

a_0 – is a current derivative v_0.

There are no contradictions in considering the current time as the initial (i.e., zero), the time at the next moment is equal to the time between control cycles $t = \Delta t$, and the time for the entire control is considered to be double tact $t_1 = 2\Delta t$.

Under these assumptions, we obtain compact formulas for magnitude, velocity, and acceleration:

$$\begin{aligned} q &= \frac{a_0\,\Delta t^2 + 8e_0 + 5\,\Delta t\, v_0}{16} \\ q' &= -\frac{a_0\,\Delta t^2 + 15\, e_0 + 7\,\Delta t\, v_0}{16\,\Delta t} \\ q'' &= -\frac{a_0\,\Delta t + 3\ v_0}{4\,\Delta t} \end{aligned} \tag{7}$$

Here, the designations q, q' and q'' are specially introduced in order not to be confused with derivatives. To illustrate, consider simple examples.

2. Control model for linear system

Consider a simple case - a linear link of the 2nd order, described by the equation

$$U(t) = A\,\ddot{y} + B\,\dot{y} + C\,y, \tag{8}$$

Where *A, B, C* – are constants;

$y = y(t)$ – is a controlled value.

As the control law, we choose the polynomial function (1.5).

Substituting (6) into equation (8) and simplifying we get a cumbersome expression to control:

$$\begin{aligned} U(t) = e_0 &\left(B\mathrm{v}1 - \frac{180At^5}{t_1^7} + \frac{300At^4}{t_1^6} + \frac{30Ct^5}{t_1^6} - \frac{120At^3}{t_1^5} - \frac{60Ct^4}{t_1^5} + \frac{30Ct^3}{t_1^4}\right) + \\ v_0 &\left(B\mathrm{v}1 - \frac{60At^5}{t_1^6} + \frac{96At^4}{t_1^5} + \frac{12Ct^5}{t_1^5} - \frac{36At^3}{t_1^4} - \frac{24Ct^4}{t_1^4} + \frac{12Ct^3}{t_1^3}\right) + \\ a_0 &\left(B\mathrm{v}1 - \frac{6At^5}{t_1^5} + \frac{9At^4}{t_1^4} + \frac{3Ct^5}{2t_1^4} - \frac{3At^3}{t_1^3} - \frac{3Ct^4}{t_1^3} + \frac{3Ct^3}{2t_1^2}\right), \end{aligned} \tag{9}$$

Determine the value of t_1, as mentioned above, can be the magnitude of the change control. It should be noted here that the maximum in absolute value derivative of function (8) becomes only at the ends of the segment.

Consider the formula (9). It specifically grouped values at e_0, v_0 и a_0. Since the time parameter $t = \Delta t$ is always equal to the duration of the control cycle (i.e., each time is controlled as the first time), and the control duration parameter is $t_1 = 2\Delta t$, then the constants are in brackets for e_0, v_0 and a_0 . That is, the control law, in general, is as follows:

$$U(t) = const1 * e_0 + const2 * v_0 + const3 * a_0, \tag{10}$$

Recalling that v_0 s a derivative of e_0 etc. we get the analytical output of the PID controller.

That is, formulas (9) and (10) - proportional-integral-differential controller (PID-controller) - differential controller of the 2nd order. Moreover, it should be noted that this regulator is adaptive, since the coefficients are calculated not once and for all, as in the classical regulator, but on each control cycle.

In general, the differential controller of the *n*-th order is as follows:

$$U(t) = \sum_{i=0}^{n} C_i \, e^{(i)}$$

In practice, the regulator above the 2nd order is problematic to use.

3. Control model for nonlinear system

Suppose there is an underwater vehicle, one degree of mobility which is described by a nonlinear differential equation

$$U(t) = A\,\dot{v}(t) - C_x\, v^2(t), \tag{11}$$

Where A и C_x – are constants,

$v(t)$ – is a speed of AUV.

The first component of the right side is Newton's law ($F = ma$), the second is the law of hydrodynamic resistance, which is proportional to the square of the velocity. The difference from the first example is that the dependence is non-linear.

For this object, I want to minimize the position, i.e. integral of the controlled variable - speed. Therefore, as a wish, we consider the fifth degree polynomial, but for the position:

$$s(t) = at^5 + bt^4 + ct^3 + dt^2 + et + f, \tag{12}$$

Acting in a way completely analogous to the first example, we get the control dependence:

$$\begin{aligned} U(t) = {} & \frac{900t^{10}c_x e_0^2}{t_1^{12}} - \frac{3600t^9 c_x e_0^2}{t_1^{11}} + \frac{5400t^8 c_x e_0^2}{t_1^{10}} - \frac{3600t^7 c_x e_0^2}{t_1^9} + \frac{900t^6 c_x e_0^2}{t_1^8} \\ & + \frac{27t^8 a_0^2 c_x}{2t_1^6} + \frac{9t^{10} a_0^2 c_x}{4t_1^8} - \frac{9t^7 a_0^2 c_x}{t_1^5} + \frac{9t^6 a_0^2 c_x}{4t_1^4} \\ & + \frac{144t^{10} c_x v_0^2}{t_1^{10}} - \frac{576t^9 c_x v_0^2}{t_1^9} + \frac{864t^8 c_x v_0^2}{t_1^8} - \frac{576t^7 c_x v_0^2}{t_1^7} + \frac{144t^6 c_x v_0^2}{t_1^6} \\ & - \frac{180At^5 e_0}{t_1^7} + \frac{300At^4 e_0}{t_1^6} - \frac{120At^3 e_0}{t_1^5} \\ & - \frac{60At^5 v_0}{t_1^6} + \frac{96At^4 v_0}{t_1^5} - \frac{36At^3 v_0}{t_1^4} \\ & - \frac{6At^5 a_0}{t_1^5} + \frac{9At^4 a_0}{t_1^4} - \frac{3At^3 a_0}{t_1^3} \\ & + \frac{90t^{10} a_0 c_x e_0}{t_1^{10}} - \frac{360t^9 a_0 c_x e_0}{t_1^9} + \frac{540t^8 a_0 c_x e_0}{t_1^8} - \frac{360t^7 a_0 c_x e_0}{t_1^7} + \frac{90t^6 a_0 c_x e_0}{t_1^6} \\ & + \frac{720t^{10} c_x e_0 v_0}{t_1^{11}} - \frac{2880t^9 c_x e_0 v_0}{t_1^{10}} - \frac{2880t^7 c_x e_0 v_0}{t_1^8} + \frac{4320t^8 c_x e_0 v_0}{t_1^9} + \frac{720t^6 c_x e_0 v_0}{t_1^7} \\ & + \frac{36t^{10} a_0 c_x v_0}{t_1^9} - \frac{144t^9 a_0 c_x v_0}{t_1^8} + \frac{216t^8 a_0 c_x v_0}{t_1^7} - \frac{144t^7 a_0 c_x v_0}{t_1^6} + \frac{36t^6 a_0 c_x v_0}{t_1^5} \end{aligned} \tag{13}$$

Here, due to the nonlinearity of the dynamic equation (11), we obtained a nonlinear PID controller.

4. Experimental studies on the model

For the one-dimensional case, we obtain the results of the refinement — a displacement of 10 meters. The results are shown in Figure 2.

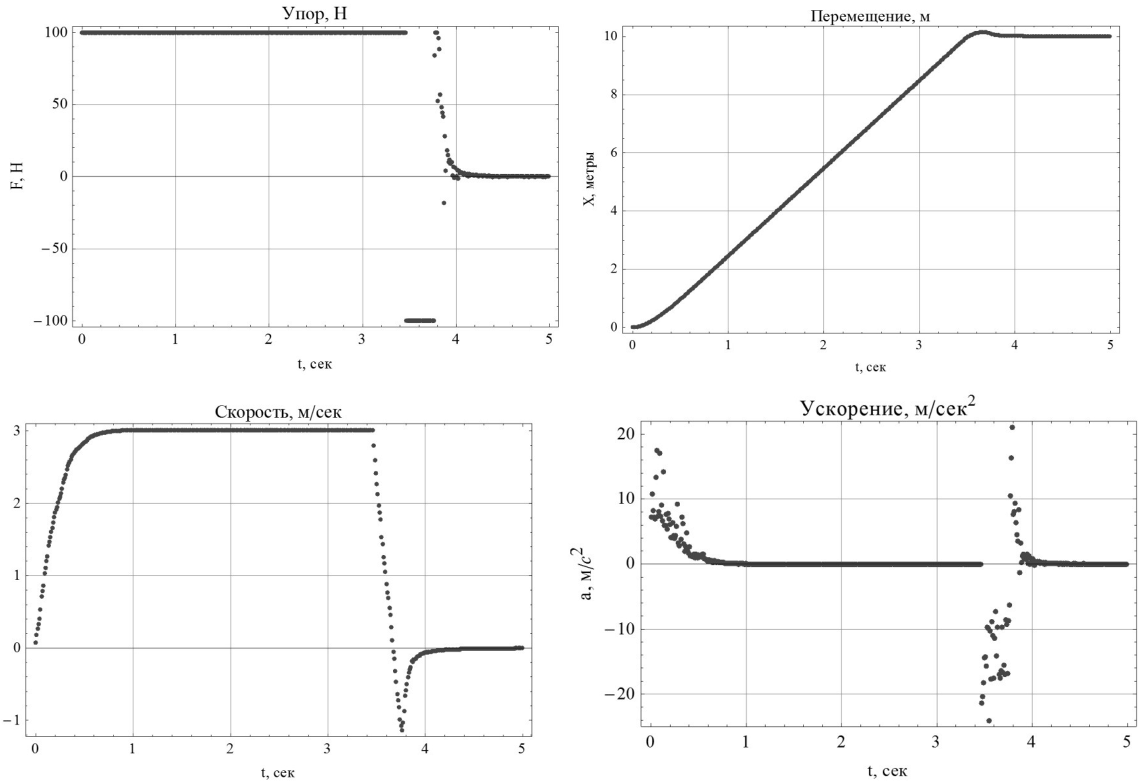


Figure 2 – Schedules of testing the movement of 10 meters

Two-dimensional case: moving by TU along the Y axis and marching on the X axis. Results - Figure 3.

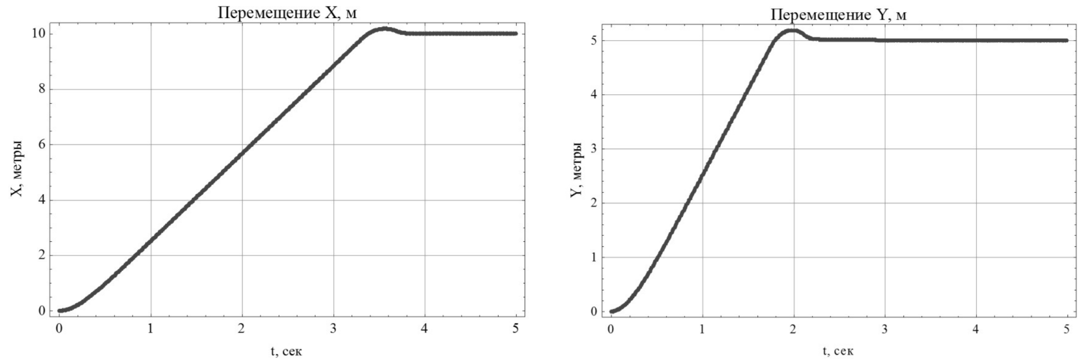


Figure 3 – Results of refinement for two-dimensional case

The graphs show that mining is proceeding satisfactorily with an overshoot of less than 3%. On the other hand, both control channels do not work synchronously and the DU according to the coordinates does not reach the goals.

Conclusion

Mathematical-computer model [2] AUV performed in the system Wolfram Mathematica.

Modeling of the device without TU but with tanks showed:

1) Convenience of ballancting and balancing. These two parameters can be changed here during the execution of the AUV work, to compensate for disturbances caused by exposure to the external environment;

2) Reliability of retention of the required depth, which is a significant problem for devices of other kinematic schemes;

3) Easy trim control. To perform depth retention for classic AUVs, in the case of non-zero buoyancy, it is required to hold non-zero and trim, since PUs at cruising speeds are not applicable. This may be unacceptable for some technological operations. Here, the depth and trim control channels are unleashed.;

4) Ease of management roll. To compensate for the roll moment developed by the main engine, it is necessary to do some (precisely calculated in advance!) Imbalance in the ballast center on the right and left sides;

5) High maneuverability at the rate, since the change in the vector of the stop MD directly determines the circulation radius of the AUV.

Of course, there are some downsides to this scheme:

1) There is no possibility to produce lag movement. Here the question immediately arises: how often are the movements “sideways” necessary?

2) AUV cannot turn "on the spot".

Both of these minuses can be removed if the front and rear TU are added to the AUV scheme, the inclusion of which will only be episodic to perform lag movements and turns “on the spot”.

In general, the work showed the promise of such studies. The authors hope to continue research in the future and, if possible, continue them on the current AUV model.

References

1. A.B. Nikolaev. Naval military robotics: state and prospects. Robotics and technical cybernetics, №1 - 2017.
2. I.A. Vasilyev, D.A. Vokhmintsev, S.A. Polovko. The emergence of forces of resistance to the movement of underwater robots and other sea-based objects when the thruster units work. Robotics and technical cybernetics, №1 - 2017.
3. S.I. Antonenko, V.P. Makarychev. The study of the dynamics of the underwater vehicle propulsion. Robotics and technical cybernetics, №1 - 2018.
4. T.I Fossen. Guidance and Control of Ocean Vehicles. – New York: Wiley, 1994.
5. I.A. Vasilyev. Control a multi-module underwater mobile robot. Robotics and technical cybernetics, №2 - 2019.
6. V.A. Leontyev, B.A. Smolnikov. Computer simulation of the dynamics of the deep-water cable with a load at the end. Robotics and Technical Cybernetics, №3 - 2018.
7. D.A. Yuhimets, V.F. Filaretov. Contour control systems for an autonomous underwater vehicle. Robotics and Technical Cybernetics, №2 - 2015.
8. A.V. Timofeev, R.M. Yusupov. Intellectualization of robotic control and navigation processes. Robotics and Technical Cybernetics, №3 - 2014.
9. V.V. Kozlov, V.P. Makarychev, A.V. Timofeev, E.I. Yurevich. Dynamics of control robots. M., "Science", 1984.
10. Autonomous underwater robots: systems and technologies / Ed. Acad. M.D. Ageev. Institute of Marine Technology Problems the Russian Academy of Sciences, M., "Science", 2005

V.S. Verba, V.I. Merkulov

OPTIMIZATION PROBLEM FOR A GROUP OF UAVS OF JOINT CONTROL, ENSURING THEIR DESIRED SPATIAL TOPOLOGY

JSC "Concern "Vega", Moscow, Russia
from_fn@mail.ru

Abstract

Coordinated group management of unmanned aerial vehicles (UAVs), providing, along with the solution of the route problem, the desired spatial topology of the participants, is a complex task.

The purpose of the report is to propose simplified versions of the synthesis of group control of a UAV with a long-term preservation of a given topology of participants with reduced requirements for computing performance. The problem was solved in two stages.

At the first stage, on the basis of the dynamic programming method, a strict solution of the linear quadratic-Gaussian problem was obtained using the quality functional, in which besides the typical terms there was a quadratic form of the weighted arrangement of the UAV. The specificity of the obtained control law is the need to solve a high-dimensional two-point boundary value problem in the reverse time, which requires large computational costs.

At the second stage, by simplifying the previous solution, the task was reduced to local optimization, which significantly simplified obtaining the control law, which ensures not only the group's flight along the desired trajectories, but also the prevention of collisions of the UAV within it.

The examples of the synthesis of control of a group of UAVs are considered, the results of studies of options for constructing the desired UAV topology on a plane when solving various routing problems are presented.

Keywords: group management optimization, network communication.

The solution to a number of economic and military tasks is possible only when using sufficiently large groups of jointly controlled aircraft [1]. Such tasks include: mobile monitoring of large volumes of airspace and the earth's surface; liquidation of consequences of various kinds of disasters; overcoming enemy air defenses, clearing large areas of the territory, etc.

It should be noted that the effectiveness of solving such problems largely depends not only on the chosen direction of movement of the group, but also on the relative position of the participants. However, in such situations, it becomes necessary to solve the problem of synthesizing coordinated control by group members, which is much more complicated than controlling a single object [2-4]. General approaches to solving the group management problem are considered in [4,5]. The increase in complexity is due to several reasons [6]:

− the complexity of the description of group actions with the formulation of collective interest, which each of the group members must realize;

− increasing the complexity of the formation of the control signal, which at the same time ensures the fulfillment of the intended purpose, and ensuring the desired position of the participants in the general topology of the group;

− a significant increase in the dimension of the solved problem of simultaneous management of all members of the group;

− an increase in the complexity of information support for the procedure for generating control signals [7].

In this regard, the synthesis of group management, in which the listed disadvantages will be reduced, is very popular.

The purpose of the report is to develop simplified options for the synthesis of group management with long-term preservation of the given topology of participants with reduced requirements for computing performance.

The dimensionality of the problem to be solved can be reduced due to its decomposition based on the synthesis of control for an individual participant taking into account the spatial position of other participants.

Statement and solution of the problem

The solution to the problem of synthesis of control of individual objects, taking into account the state of all participants, will be performed on the basis of the mathematical apparatus of the statistical theory of optimal control in a modified version of the solution of the Letov-Kalman problem, which will reduce the requirements on the computing performance of the control system.

Moreover, we will assume that the following assumptions are fulfilled:

– each object independently forms its own control for the current situation on the basis of information about the goal facing the group, about its state and the condition of other objects;

– as optimal, such control of each object in the current situation is understood that makes the maximum possible contribution to the achievement of a common goal, i.e. provides the maximum increment of the general functional during the transition of the group from the current state to the final [8-10];

– in the optimized functional, both the requirements of the group's purpose and the requirements for ensuring the given topology of the participants should be taken into account;

– information exchange between participants is carried out as part of a local network on the principle of "each with each" [5].

In mathematical terms, the problem is formulated as follows.

For a group consisting of N objects of the same type, each of which is determined by a state model

$$\dot{\mathbf{x}}_i(t) = \mathbf{F}_i(t)\mathbf{x}_i(t) + \mathbf{B}_i\mathbf{u}_i(t) + \xi_{xi}(t),\ i = \overline{1,N}, \tag{1}$$

where

$$\mathbf{x}_i = \left[\mathbf{x}_{di}^T \mathbf{x}_{ci}^T\right]^T - \tag{2}$$

is a composite vector in which the n-dimensional vector

$$\dot{\mathbf{x}}_{di}(t) = \mathbf{F}_{di}(t)\mathbf{x}_{di}(t) + \xi_{di}(t) \tag{3}$$

determines the desired state coordinates (program of actions), and the n-dimensional vector

$$\dot{\mathbf{x}}_{ci}(t) = \mathbf{F}_{ci}\mathbf{x}_i(t) + \mathbf{B}_{ci}\mathbf{u}_i(t) + \xi_{ci}(t) \tag{4}$$

displays the current (controlled) state of the object, in the presence of measurements

$$\mathbf{z}_i(t) = \mathbf{H}_i\mathbf{x}_i(t) + \xi_{mi}(t) \tag{5}$$

it is necessary to find the r-dimensional ($r \le n$) vector $\mathbf{u}_i$ of control signals that are optimal for the minimum functional

$$I = M_c\left\{\int_0^{t_\kappa}\left[\mathbf{x}_i^T(t)\mathbf{L}_{i1}\mathbf{x}_i(t) + \sum_{j=1, j\ne i}^{N}\Delta\mathbf{x}_{ij}^T(t)\mathbf{M}_{j1}\Delta\mathbf{x}_{ij}(t) + \mathbf{u}_i^T(t)\mathbf{K}_i\mathbf{u}_i(t)\right]dt + \right.$$
$$\left. + \mathbf{x}_i^T(t_\kappa)\mathbf{Q}_{i1}\mathbf{x}_i(t_\kappa) + \sum_{j=1, j\ne i}^{N}\Delta\mathbf{x}_{ij}^T(t_\kappa)\mathbf{G}_{j1}\Delta\mathbf{x}_{ij}(t_\kappa)\right\}, \tag{6}$$

where $\Delta\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$, $j = \overline{1,N}$, $j \ne i$.

In (1)-(6) $\mathbf{F}_{di}$ and $\mathbf{F}_{ci}$ in the general case are non-stationary dynamic matrices that take into account the internal relationships of processes (3) and (4), while $\mathbf{F}_{di}$ in the general case is a non-stationary matrix that determines the program of movement of the group; $\mathbf{B}_{ci}$ is the matrix of the effectiveness of control signals $\mathbf{u}_i$; $\mathbf{z}$ is the m-dimensional ($m \le 2n$) vector of measurements; $\mathbf{H}_i$ is the communication matrix for $\mathbf{x}_i$ and $\mathbf{z}_i$; ξ_{di}, ξ_{ci} and ξ_{mi} are the centered Gaussian noises of state and measurements with known spectral density matrices; $\mathbf{x}_j$, $j = \overline{1,N}$, $j \ne i$ are state vectors for other objects of the group; $\mathbf{K}_i$ is a positive definite matrix with the $r \times r$ size of fines for the magnitude of the control signals; M_c is a sign of conditional expectation;

$$\mathbf{F}_i=\begin{bmatrix}\mathbf{F}_{di} & \mathbf{O}_1\\ \mathbf{O}_2 & \mathbf{F}_{ci}\end{bmatrix};\ \mathbf{B}_i=\begin{bmatrix}\mathbf{O}_3\\ \mathbf{B}_{ci}\end{bmatrix};\ \xi_{xi}=\begin{bmatrix}\xi_{di}\\ \xi_{ci}\end{bmatrix};\ \mathbf{L}_{i1}=\begin{bmatrix}\mathbf{L} & \text{-}\mathbf{L}\\ \text{-}\mathbf{L} & \mathbf{L}\end{bmatrix};$$
$$\mathbf{M}_{j1}=\begin{bmatrix}\mathbf{M} & \text{-}\mathbf{M}\\ \text{-}\mathbf{M} & \mathbf{M}\end{bmatrix};\ \mathbf{Q}_{i1}=\begin{bmatrix}\mathbf{Q} & \text{-}\mathbf{Q}\\ \text{-}\mathbf{Q} & \mathbf{Q}\end{bmatrix};\ \mathbf{G}_{j1}=\begin{bmatrix}\mathbf{G} & \text{-}\mathbf{G}\\ \text{-}\mathbf{G} & \mathbf{G}\end{bmatrix}, \tag{7}$$

where $\mathbf{O}_1$, $\mathbf{O}_2$, $\mathbf{O}_3$ are zero matrices of appropriate sizes; $\mathbf{L}$ and $\mathbf{Q}$ are non-negatively determined $n\times n$ matrices with penalties for the accuracy of approaching $\mathbf{x}_{ci}$ to $\mathbf{x}_{di}$ at the current time t and end time t_e; $\mathbf{M}$ and $\mathbf{G}$ are non-negatively defined $n\times n$ matrices with penalties for violating the desired relationship between $\mathbf{x}_i$ and $\mathbf{x}_j$ at the current and final times.

The specificity of functional (6) is that, in the framework of solving the general problem, it takes into account the requirements for the formation of the desired phase trajectory and the requirements for ensuring a given group topology.

Since the initial models (1)-(5) are linear, the quality functional is quadratic, and the perturbations are Gaussian, then, in accordance with the separation theorem [11], control synthesis and filtering problems can be solved separately. In this case, control synthesis can be performed in a deterministic setting $\left(\xi_{xi}=\mathbf{0}\right)$, provided that in the resulting algorithm the state coordinates $\mathbf{x}_i$ and $\mathbf{x}_j$ are replaced by their optimal estimates $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j$.

In such conditions (LKG problem), Bellman's equation [12] can be used to find the control:

$$\frac{-\partial S\left(\mathbf{x}_i(t),t\right)}{\partial t}=\min_{\{\mathbf{u}_i\}}\left\{\Phi_d\left[\mathbf{x}_i(t),\mathbf{x}_j(t),\mathbf{u}_i(t),t\right]+\dot{\mathbf{x}}_i^T(t)\frac{\partial S\left(\mathbf{x}_i(t),t\right)}{\partial \mathbf{x}_i(t)}\right\}, \tag{8}$$

$$S\left[\mathbf{x}_e(t),t_e\right]=\Phi_e\left[\mathbf{x}_i(t_e),\mathbf{x}_j(t_e),\mathbf{u}_i(t_e),t_e\right]. \tag{9}$$

Here S is the Bellman function; $\Phi_d\left[\bullet\right]$ is the integrand of the functional used (6); $\dot{\mathbf{x}}_i$ is determined by model (1), $\Phi_e\left[\bullet\right]$ is the terminal part of the functional that defines the boundary conditions for S.

Using (1) and (6) in (8), (9), we obtain:

$$-\frac{\partial S}{dt}=\min_{\{\mathbf{u}_i\}}\left\{\mathbf{x}_i^T\mathbf{L}_i\mathbf{x}_i+\sum_{j=1,j\neq i}^{N}\Delta\mathbf{x}_{ij}^T\mathbf{M}_j\Delta\mathbf{x}_{ij}+\mathbf{u}_i^T\mathbf{K}_i\mathbf{u}_i+\left[\mathbf{x}_i^T\mathbf{F}_i^T+\mathbf{u}_i^T\mathbf{B}_i^T\right]\frac{\partial S}{\partial \mathbf{x}_i}\right\}=$$
$$=\mathbf{x}_i^T\mathbf{L}_i\mathbf{x}_i+\sum_{j=1,j\neq i}^{N}\Delta\mathbf{x}_{ij}^T\mathbf{M}_j\Delta\mathbf{x}_{ij}+\mathbf{x}_i^T\mathbf{F}_i^T\frac{\partial S}{\partial \mathbf{x}_i}+\min_{\{\mathbf{u}_i\}}\left\{\mathbf{u}_i^T\mathbf{K}_i\mathbf{u}+\mathbf{u}_i^T\mathbf{B}_i^T\frac{\partial S}{\partial \mathbf{x}_i}\right\}, \tag{10}$$

$$S\left[\mathbf{x}(t_e),t_e\right]=\mathbf{x}_i^T(t_e)\mathbf{Q}_{i1}\mathbf{x}_i(t_e)+\sum_{j=1,j\neq i}^{N}\Delta\mathbf{x}_{ij}^T(t_e)\mathbf{G}_{j1}\Delta\mathbf{x}_{ij}(t_e). \tag{11}$$

In (10) and (11), to simplify the notation, the dependence of vectors on time was omitted and all terms independent of $\mathbf{u}_i$, were taken out of the sign of the minimum operation.

A control $\mathbf{u}_i$ for each object $\left(i=\overline{1,N}\right)$ that minimizes (6) can be found by equating in (10) to zero the result of differentiation by $\mathbf{u}_i^T$ terms in curly brackets:

$$2\mathbf{K}_i\mathbf{u}_i+\mathbf{B}_i^T\frac{\partial S}{\partial \mathbf{x}_i}=\mathbf{0};$$

$$\mathbf{u}_i=-0{,}5\mathbf{K}_i^{-1}\mathbf{B}_i^T\frac{\partial S}{\partial \mathbf{x}_i}. \tag{12}$$

Substituting (12) in (10), we obtain:

$$-\frac{\partial S}{\partial t}=\mathbf{x}_i^T\mathbf{L}_i\mathbf{x}_i+\sum_{j=1,j\neq i}^{N}\Delta\mathbf{x}_{ij}^T\mathbf{M}_j\Delta\mathbf{x}_{ij}+\mathbf{x}_i^T\mathbf{F}_i^T\frac{\partial S}{\partial\mathbf{x}_i}-0,25\left(\frac{\partial S}{\partial\mathbf{x}_i}\right)^T\mathbf{B}_i\mathbf{K}_i^{-1}\mathbf{B}_i^T\frac{\partial S}{\partial\mathbf{x}_i}. \tag{13}$$

A solution to this partial differential equation will be sought in the class of quadratic forms

$$S=\mathbf{x}_i^T\mathbf{R}_{i1}\mathbf{x}_i+\sum_{i=1,j\neq i}^{N}\Delta\mathbf{x}_{ij}^T\mathbf{P}_{j1}\Delta\mathbf{x}_{ij}; \tag{14}$$

$$\frac{\partial S}{\partial\mathbf{x}_i^T}=2\mathbf{R}_{i1}\mathbf{x}_i+2\sum_{i=1,j\neq i}^{N}\mathbf{P}_{j1}\Delta\mathbf{x}_{ij}; \tag{15}$$

$$\frac{\partial S}{\partial t}=\mathbf{x}_i^T\dot{\mathbf{R}}_{i1}\mathbf{x}_i+\sum_{i=1,j\neq i}^{N}\mathbf{x}_i\dot{\mathbf{P}}_{j1}\Delta\mathbf{x}_{ij}. \tag{16}$$

Using (15) in (12), we have

$$\mathbf{u}_i=-\mathbf{K}_i^{-1}\mathbf{B}_i^T\left(\mathbf{R}_{i1}\mathbf{x}_i+\sum_{i=1,j\neq i}^{N}\mathbf{P}_{j1}\Delta\mathbf{x}_{ij}\right). \tag{17}$$

Using (15) and (16) in (13), we can obtain

$$\dot{\mathbf{R}}_{i1}=-\mathbf{L}_{i1}-\mathbf{F}_i^T\mathbf{R}_{i1}-\mathbf{R}_{i1}\mathbf{F}_i^T+\mathbf{R}_{i1}\mathbf{B}_i\mathbf{K}_i^{-1}\mathbf{B}_i^T\mathbf{R}_{i1}, \tag{18}$$

$$\dot{\mathbf{P}}_{j1}=-\mathbf{M}_{j1}+\sum_{j=1,j\neq i}^{N}\mathbf{P}_{j1}\mathbf{B}_i\mathbf{K}_i^{-1}\mathbf{B}_i^T\mathbf{P}_{j1}. \tag{19}$$

The boundary conditions for (18) and (19) are determined by comparing the results of (11) and (14) with $t=t_e$:

$$\mathbf{R}_{i1}(t_e)=\mathbf{Q}_{i1},\ \mathbf{P}_{j1}(t_e)=\mathbf{G}_{j1}. \tag{20}$$

Upon receipt of (16), it was taken into account that the Bellman function does not depend on the state coordinates [12].

Since the LKG problem was considered, on the basis of the principle of statistical equivalence [11] it can be argued that the statistical control law will be adequate to the deterministic one provided that the state coordinates $\mathbf{x}_i$, $\mathbf{x}_j$, and $\Delta\mathbf{x}_{ij}$ are replaced in them by their optimal estimates $\hat{\mathbf{x}}_i$, $\hat{\mathbf{x}}_j$ and $\Delta\hat{\mathbf{x}}_{ij}$, i.e.

$$\mathbf{u}_i=-\mathbf{K}_i^{-1}\mathbf{B}_i^T\left(\mathbf{R}_{i1}\hat{\mathbf{x}}_i+\sum_{i=1,j\neq i}^{N}\mathbf{P}_{j1}\Delta\hat{\mathbf{x}}_{ij}\right). \tag{21}$$

Analysis (18)-(21) allows us to draw the following conclusions.

In the general case, the collective control system with network information exchange should include: optimal controllers that calculate control signals $\mathbf{u}_i$ at each object; optimal filters forming estimates $\hat{\mathbf{x}}_i$ and $\Delta\hat{\mathbf{x}}_{ij}$, $j=\overline{1,N}$, $j\neq i$, and a system for exchanging estimates of the coordinates of the state of each object.

The control signal generated at each object depends on its parameters $\mathbf{F}_{ci}$ and $\mathbf{B}_{ci}$, the desired control law $\mathbf{F}_{di}$, the state of the object $\hat{\mathbf{x}}_i$ itself, and the state of other objects that determine the errors $\Delta\hat{\mathbf{x}}_{ij}$.

The coefficients of the matrices $\mathbf{R}_{i1}$ and $\mathbf{P}_{j1}$ calculated according to the rules (18) and (19) together take into account in (21) the parameters of the object $\mathbf{F}_i$ and $\mathbf{B}_i$, the requirements for the accuracy $\mathbf{L}_{i1}$ and the economy of the control $\mathbf{K}_i$, and the degree of compliance with the desired states of the *i*-th and *j*-th objects defined by the coefficients of the matrix $\mathbf{M}_{j1}$.

Using (21) allows us to ensure not only the movement of each object along the desired trajectory due to accounting for $\mathbf{x}_i$, but also to prevent collisions between them in the process of joint movement due to accounting $\Delta\mathbf{x}_{ij}$.

The specificity of the obtained algorithm is the need to solve a two-point boundary value problem, the essence of which is that the matrices $\mathbf{R}_{i1}$ and $\mathbf{P}_{j1}$ are calculated in the inverse time from t_e to t, while the control (21) for the object is formed in the direct direction from t to t_e, which predetermines increased requirements for computational performance.

The collective management procedure can be significantly simplified if optimization of the local quality functional is minimized [13], in which in (6) $\mathbf{L}_{i1} = \mathbf{0}$, $\mathbf{M}_{j1} = \mathbf{0}$, and, every moment of time is considered as the moment of the possible termination of management, i.e. $t_e = t$. Then from (6) and (20) it follows:

$$I = M_c \left\{ \mathbf{x}_i^T(t)\mathbf{Q}_{i1}\mathbf{x}_i(t) + \sum_{j=1, j\neq i}^{N} \Delta\mathbf{x}_{ij}^T(t)\mathbf{G}_{j1}\Delta\mathbf{x}_{ij}(t) + \int_0^t \mathbf{u}_i^T(t)\mathbf{K}\mathbf{u}_i(t)dt \right\}, \tag{22}$$

$$\mathbf{R}_{i1}(t) = \mathbf{Q}_{i1}, \mathbf{P}_{j1}(t) = \mathbf{G}_{j1}. \tag{23}$$

Using (22), (2), (7) in (21), we obtain:

$$\begin{aligned} \mathbf{u}_i &= -\mathbf{K}_i \left[\mathbf{O}_3 \mathbf{B}_i^T\right] \left\{ \begin{bmatrix} \mathbf{Q} & \text{-}\mathbf{Q} \\ \text{-}\mathbf{Q} & \mathbf{Q} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_{di} \\ \hat{\mathbf{x}}_{ci} \end{bmatrix} + \sum_{j=1, j\neq i}^{N} \begin{bmatrix} \mathbf{G} & \text{-}\mathbf{G} \\ \text{-}\mathbf{G} & \mathbf{G} \end{bmatrix} \begin{bmatrix} \Delta\hat{\mathbf{x}}_{dij} \\ \Delta\hat{\mathbf{x}}_{cij} \end{bmatrix} \right\} = \\ &= \mathbf{K}^{-1}{}_i \mathbf{B}_i^T \left\{ \mathbf{Q}\left(\hat{\mathbf{x}}_{di} - \hat{\mathbf{x}}_{ci}\right) + \sum_{j=1, j\neq i}^{N} \mathbf{G}\left(\Delta\hat{\mathbf{x}}_{dij} - \Delta\hat{\mathbf{x}}_{cij}\right) \right\}, \end{aligned} \tag{24}$$

where $\Delta\hat{\mathbf{x}}_{dij} = \hat{\mathbf{x}}_{di} - \hat{\mathbf{x}}_{dj}$, $\Delta\hat{\mathbf{x}}_{cij} = \hat{\mathbf{x}}_{ci} - \hat{\mathbf{x}}_{cj}$, $j = \overline{1,N}$, $j \neq i$.

An undoubted advantage of (24) is the possibility of generating a group control signal without solving in high time a high-dimensional two-point boundary-value problem using equations (18)-(20).

In addition, the resulting collective control system is characterized by the presence of negative feedbacks on all coordinates in each object. This indicates its high stability and low sensitivity to the accuracy of the parameters. Moreover, the control signal in it depends on the control errors $\hat{\mathbf{x}}_{di} - \hat{\mathbf{x}}_{ci}$ and $\Delta\hat{\mathbf{x}}_{dij} - \Delta\hat{\mathbf{x}}_{cij}$. This feature allows you to halve the requirements for transmission capacity of transmission lines, since it is enough to exchange only errors $\Delta\hat{\mathbf{x}}_{dij}$ and $\Delta\hat{\mathbf{x}}_{cij}$.

Moreover, from (24) its ability to ensure not only the movement of each object along the desired path due to accounting for $\hat{\mathbf{x}}_{di} - \hat{\mathbf{x}}_{ci}$, but also to prevent their mutual collisions due to accounting for $\Delta\hat{\mathbf{x}}_{dij} - \Delta\hat{\mathbf{x}}_{cij}$ is clearly traced. This feature makes it possible to further simplify the group management procedure and its information support due to the transition to management "with a leader" [5], which can be obtained as a special case of management (24). In the framework of this approach, the control for the leader forming the flight path is determined only by the first term, while the relative position of other elements of the group is determined by using only the second term.

Thus, the proposed network management option allows for the joint operation of objects, ensuring movement along predetermined paths and preventing collisions between them with the minimum computing performance requirements of a digital computer and to obtain simpler group control options.

Below are the results of a study of the effectiveness of the considered method of synthesis of group management.

Example

The performance check of the proposed method for optimizing group control was carried out on the example of the synthesis of joint control of a group of three UAVs in the process of their joint flight at predetermined intervals [6] along both straight and complex, curved paths, provided that the control for them is formed according to the law (24), and by law:

$$\mathbf{u}_i = \mathbf{K}^{-1}{}_i \mathbf{B}_i^T \mathbf{Q}\left(\hat{\mathbf{x}}_{di} - \hat{\mathbf{x}}_{ci}\right), \tag{25}$$

which does not take into account the influence of other members of the group.

The initial position of the UAV was taken: for the first – *z*=0, *x*=0, for the second – *z*=50 m, *x*=0 and for the third – *z*=1900 m, *x*=0. The desired flight paths are three parallel lines at an angle $\psi_{di} = 45^\circ$ to the OZ axis of a rectangular coordinate system located at a distance of 550 m from each other along which the UAVs move at speed $V_i = 42\, м/c$. Initial UAV courses: for the first $\psi_1 = 30^\circ$, for the second $\psi_2 = 60^\circ$ and third $\psi_3 = 45^\circ$ (see Fig. 1).

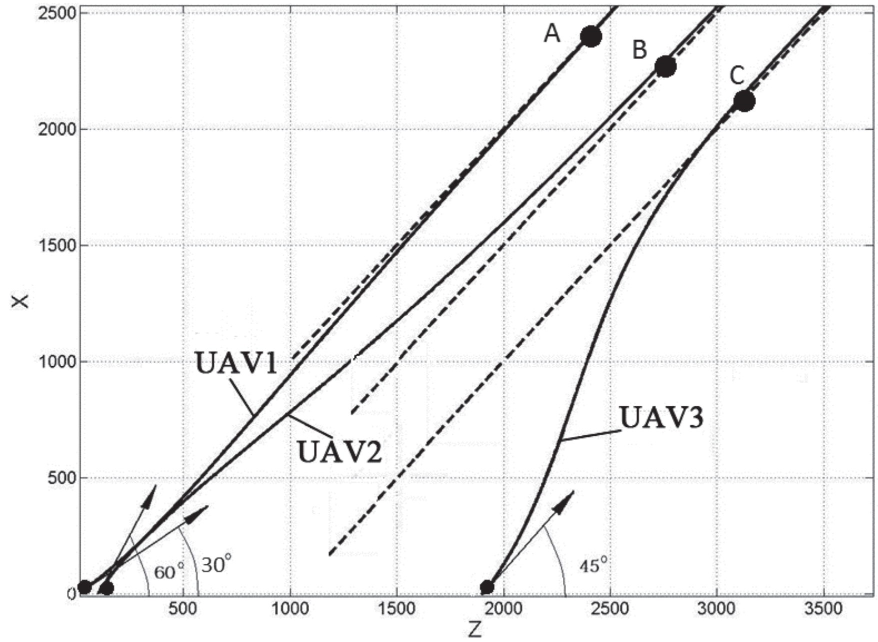


Figure 1 – UAV flight paths with individual control

It is necessary to check the ability of the group to reach the desired points in space A, B, C with the desired course, keeping the desired intervals in the horizontal plane, regardless of the initial courses of each UAV and the initial distances between them.

The following indicators were used as efficiency indicators: linear errors of UAV output to the desired points, control errors in angular coordinates, as well as regulation time and realized overloads.

It should be noted that the UAV control calculation was carried out in the polar coordinate system, and the construction of trajectories in a rectangular one.

To calculate control signals, it is necessary to have state models (3), (4), which are appropriate to use kinematic equations [6]:

$$\begin{aligned} &\dot{\varphi}_{gi} = \omega_{gi} + \frac{j_{gi}}{\dot{D}_i}, && \varphi_{gi}(0) = \varphi_{di}; \\ &\dot{\omega}_{gi} = -\frac{2\dot{D}_i}{D_i}\omega_{gi} - \frac{j_{gi}}{D_i}, && \omega_{gi}(0) = \omega_{g_0}; \end{aligned} \tag{26}$$

provided that

$$\begin{aligned} &\dot{\varphi}_{di} = 0, \quad \varphi_{di}(0) = 30^\circ, 60^\circ, 45^\circ \\ &\dot{\omega}_{di} = 0, \omega_{di}(0) = 0, \end{aligned} \tag{27}$$

where φ_{di} and φ_{gi} are desired and current bearings of the desired points A, B, C of the UAV exit in the horizontal plane; ω_{di} and ω_{gi} are the desired and current angular velocities of the lines of sight of these points from the UAV (Fig. 1); j_{gi} is instantaneous lateral acceleration (control signal), D_i is range to desired point, and $\dot{D}_i$ is rate of its change.

The relations (3), (4), (7) and (26), (27) correspond to the vector-matrix representations:

$$\mathbf{x}_{di}=\begin{bmatrix}\varphi_{di}\\0\end{bmatrix},\mathbf{x}_{ci}=\begin{bmatrix}\varphi_{gi}\\\omega_{gi}\end{bmatrix},\mathbf{B}_{ci}=\begin{bmatrix}1/\dot{D}_i\\-1/D_i\end{bmatrix},\mathbf{Q}=\begin{bmatrix}q_\varphi & 0\\0 & q_\omega\end{bmatrix},\mathbf{G}=\begin{bmatrix}g_\varphi & 0\\0 & g_\omega\end{bmatrix},$$
$$\mathbf{K}=k_i,\mathbf{u}=j_{gi}. \tag{28}$$

Using (28) in (24), we obtain control laws for each UAV:

$$\begin{aligned}
j_{gd1}&=\left[\frac{q_{\varphi1}}{k_1\hat{\dot{D}}_1}\left(\hat{\varphi}_{d1}-\hat{\varphi}_{c1}\right)+\frac{q_{\omega1}}{k_1\hat{D}_1}\hat{\omega}_1\right]+\left\{\frac{g_{\varphi2}}{k_2\hat{\dot{D}}_2}\left(\Delta\hat{\varphi}_{d12}-\Delta\hat{\varphi}_{c12}\right)+\frac{g_{\omega2}}{k_2\hat{D}_2}\hat{\omega}_2+\right.\\
&\left.+\frac{g_{\varphi3}}{k_3\hat{\dot{D}}_3}\left(\Delta\hat{\varphi}_{d13}-\Delta\hat{\varphi}_{c13}\right)+\frac{g_{\omega3}}{k_3\hat{D}_3}\hat{\omega}_3\right\};\\
j_{gd2}&=\left[\frac{q_{\varphi2}}{k_2\hat{\dot{D}}_2}\left(\hat{\varphi}_{d2}-\hat{\varphi}_{c2}\right)+\frac{q_{\omega2}}{k_2\hat{D}_2}\hat{\omega}_2\right]+\left\{\frac{g_{\varphi1}}{k_1\hat{\dot{D}}_1}\left(\Delta\hat{\varphi}_{d21}-\Delta\hat{\varphi}_{c21}\right)+\frac{g_{\omega1}}{k_1\hat{D}_1}\hat{\omega}_1+\right.\\
&\left.+\frac{g_{\varphi3}}{k_3\hat{\dot{D}}_3}\left(\Delta\hat{\varphi}_{d23}-\Delta\hat{\varphi}_{c23}\right)+\frac{g_{\omega3}}{k_3\hat{D}_3}\hat{\omega}_3\right\};\\
j_{gd3}&=\left[\frac{q_{\varphi3}}{k_3\hat{\dot{D}}_3}\left(\hat{\varphi}_{d3}-\hat{\varphi}_{c3}\right)+\frac{q_{\omega3}}{k_3\hat{D}_3}\hat{\omega}_3\right]+\left\{\frac{g_{\varphi1}}{k_1\hat{\dot{D}}_1}\left(\Delta\hat{\varphi}_{d31}-\Delta\hat{\varphi}_{c31}\right)+\frac{g_{\omega1}}{k_1\hat{D}_1}\hat{\omega}_1+\right.\\
&\left.+\frac{g_{\varphi2}}{k_2\hat{\dot{D}}_2}\left(\Delta\hat{\varphi}_{d32}-\Delta\hat{\varphi}_{c32}\right)+\frac{g_{\omega2}}{k_2\hat{D}_2}\hat{\omega}_2\right\}.
\end{aligned} \tag{29}$$

Rational values f the ratios $q_{\varphi i}/k_i$, $q_{\omega i}/k_i$ and $g_{\varphi i}/k_i$, $g_{\omega i}/k_i$ of the coefficients of the matrices $\mathbf{Q}$, $\mathbf{G}$ and $\mathbf{K}$, which are responsible for the accuracy and efficiency of the control, are selected according to the well-known rules [14], taking into account restrictions on the magnitude of the control signal, the transient time and the admissible control error in the steady state mode. In addition, a simpler so-called “principle of equal strength” [15] can be used to calculate them, the essence of which is the assumption that the products of all weighted error coefficients by the squares (variances) of these errors) are assumed to be the same for all state coordinates. However, this approach is approximate and requires refinement due to the variation of error coefficients according to the results of modeling and calculation of specific values of the quality functionals.

We study the control case according to the law (25), in which changes in the states of other objects are not taken into account. In relations (29), the terms in square brackets correspond to them. It should be noted that such a law can be used to control the leading UAV when using control with a leader.

The flight paths for this control are shown in Fig. 1, where the dashed line denotes the desired trajectories, and A, B, C are the end points of control. It can be seen from the figure that all three UAVs can reach the desired course, however, the flight paths of the first and second UAVs intersect in the initial section, which indicates their collision.

For comparison, in Fig. 2 shows the flight paths of three UAVs under the same conditions obtained using collective control (29).

It can be seen from the figure that the regulation time characterizing the time the group entered the desired courses has slightly increased, however, the flight is provided without crossing paths, which eliminates their collision. Moreover, the group seeks to maintain the desired distance between UAVs throughout the flight.

To evaluate the linear control errors, we use the relation determining the linear missile drone UAV in the horizontal plane [14]:

$$h_{gd}\approx\frac{D_i^2\omega_i}{V_i},$$

where D_i is current range to the desired points A, B, C (Fig. 2); ω_i is angular velocity of the line of sight of these points with a UAV; V_i is UAV speed at a given time.

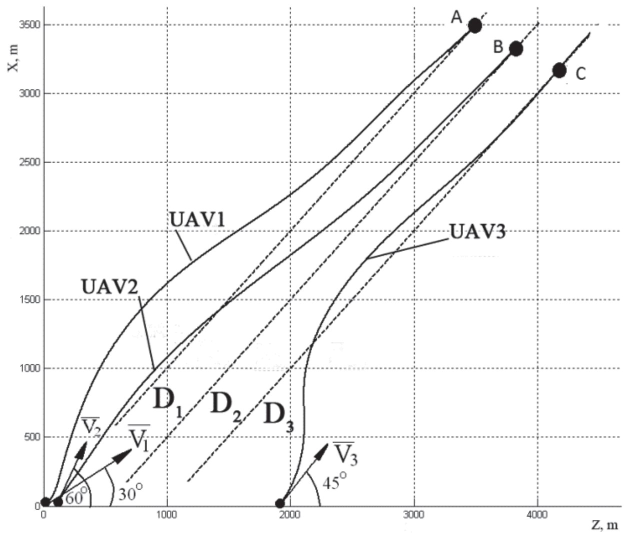


Figure 2 – Flight paths of UAVs with collective control

Figure 3 shows the dependence of the current linear misses of each UAV on time.

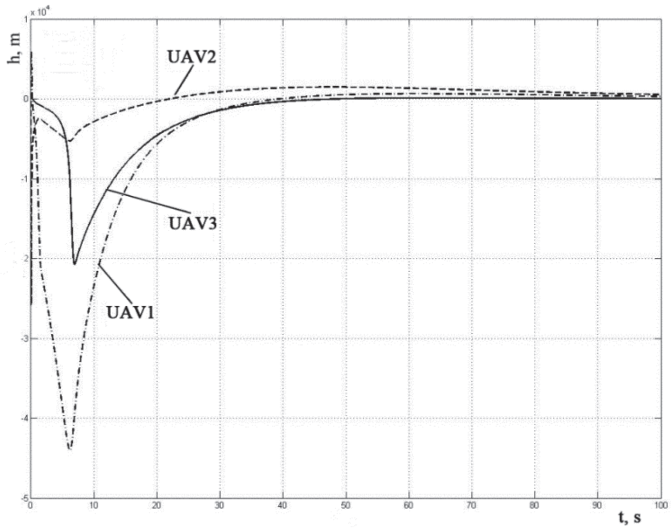


Figure 3 – Dependence of the current miss on time

Figure 4 shows the time dependences of the angular errors $\varphi_{di} - \varphi_{ci}$ of the control. From figures 3 and 4 it follows that collective control according to the proposed law (24) implements the elimination of both linear and angular errors. Moreover, the time for eliminating errors in linear and angular coordinates will depend on the initial conditions: on the presence of intersecting courses between UAVs, the distances between UAVs and the distances between the desired exit points of the group.

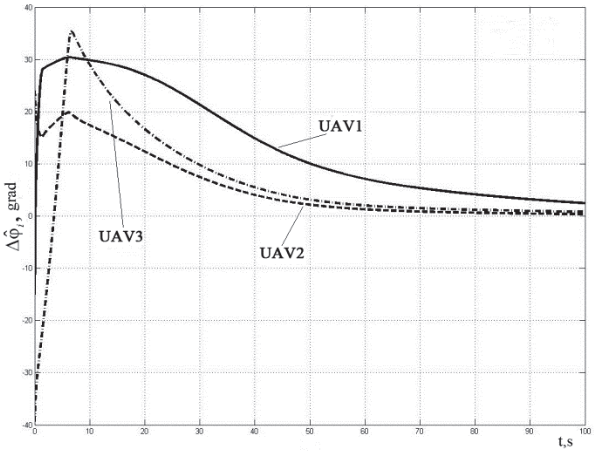


Figure 4 – The dependence of errors in angular coordinates on time

The results of the study of the group control law (29) when moving along more complex paths are shown in Figures 5, 6. Figure 5 shows the flight paths of the group while bypassing the danger zone, and Figure 6 shows the transverse acceleration modules corresponding to this situation.

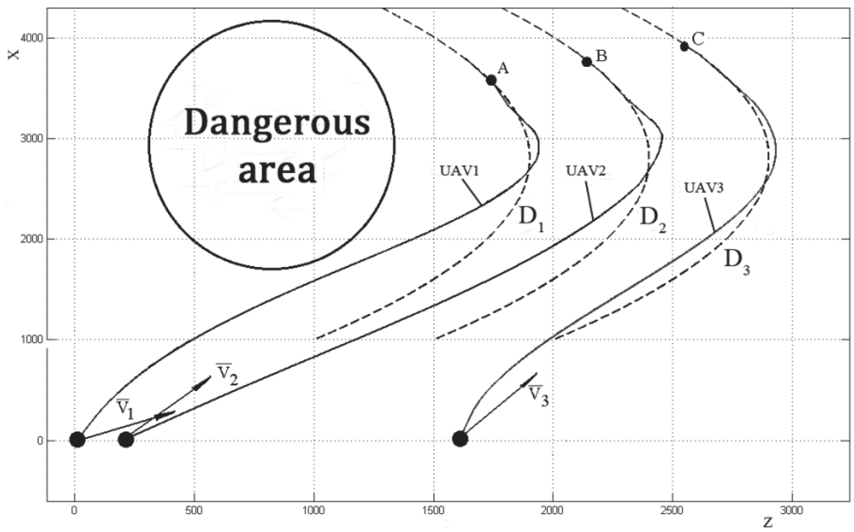


Figure 5 – UAV group flight along curved paths

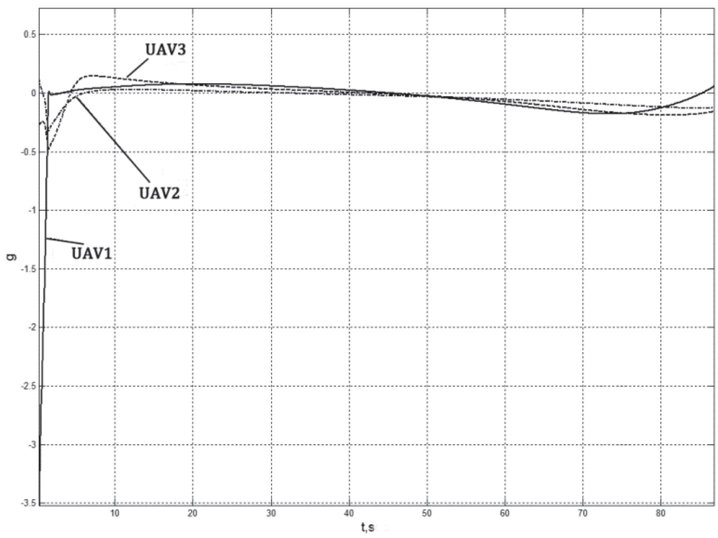


Figure 6 – Time dependence of transverse UAV overloads

Figure 6 shows that the maximum value of overloads acting on UAVs does not exceed 2.5g, which indicates the absence of fundamental restrictions on the possibility of implementing the proposed control law.

Conclusion

Based on the results of the work performed, the following conclusions can be drawn:

– in the general case, a group management system with network information exchange should include: optimal controllers that implement control signals for each object; optimal filters, forming estimates $\hat{\mathbf{x}}_i$, $\hat{\mathbf{x}}_j$ and $\Delta\hat{\mathbf{x}}_{ij}$, $j=\overline{1,N}$, $j \neq i$, and a system for exchanging estimates of the coordinates of the state of each object;

– the synthesized collective control algorithm provides adequate control of the group both in linear and in angular coordinates with minimal computational costs, realizing high safety of joint movement, including when using complex trajectories;

– the exchange of errors $\Delta\hat{\mathbf{x}}_{dij}$, $\Delta\hat{\mathbf{x}}_{cij}$ between objects allows you to save the desired topology throughout the flight.

It should be noted that the proposed version of group management based on (24), (29) is highly versatile. On its basis, as special cases, other types of group management can be implemented [5]. Under centralized control, it suffices to use terms in square brackets for each object from (29). When using control with a leader, it is enough for him to use terms in square brackets, while for all others, terms in curly brackets.

In general, the conducted studies suggest that the considered algorithm does not impose fundamental restrictions on the possibility of its implementation.

References

1. Verba V.S. Radioelektronnije sistemi upravlenija novogo pokolenija. Teoreticheskije problem razrabotki. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2015. №5. pp. 3-9.
2. Verba V.S., Kapustyan S.G., Merkulov V.I., Kharkov V.P. Optimizatsija radioelektronnih system upravlenija. Metodi i algoritmi sinteza optimal'nogo upravlenija. Part 1. Klassifikatsija metedov reshenija zadachi optimal'nogo upravlenija. Metod dinamicheskogo programmirovanija. Printsip maksimuma Pontrjagina. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2012. №12. pp. 5-16.
3. Verba V.S., Kapustyan S.G., Merkulov V.I., Kharkov V.P. Optimizatsija radioelektronnih system upravlenija. Metodi i algoritmi sinteza optimal'nogo upravlenija. Part 2. Prikladnije metodi i algoritmi teorii optimal'nogo upravlenija. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2013. №3. pp. 3-18.
4. Verba V.S., Kapustyan S.G., Merkulov V.I., Kharkov V.P. Optimizatsija radioelektronnih system upravlenija. Metodi i algoritmi sinteza optimal'nogo upravlenija. Part 3. Metodi i algoritmi obespechenija zadannih svojstv upravljajemih dinamicheskih ob'ektov i sistem. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2013. № 11. pp. 3-20.
5. Kalyaev I.A., Gaiduk A.R., Kapustyan S.G. Modeli i algoritmi kollektivnogo upravlenija v gruppah robotov. Moscow. Fizmatlit. 2009.
6. Verba V.S., Merkulov V.I., Milyakov D.A. Problemi upravlenija bol'shimi plotnimi gruppami bespilotnih letatel'nih apparatov. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2018. №6. pp. 3-13.
7. Verba V.S., Polivanov S.S. Organizatsija informatsionnogo obmena v setetsentricheskih bojevih operatsijah. Radiotehnika. 2009. №8. pp. 57-62.
8. Merkulov V.I., Kharkov V.P., Shamarov N.N. Optimizatsija kollektivnogo upravlenija gruppoj bespilotnih letatel'nih apparatov. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2012. №7. pp.3.
9. Kharkov V.P., Merkulov V.I. Sintez algoritma ijerarhicheskogo upravlenija gruppoj bespilotnih letatel'nih apparatov. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2012. №8. pp. 61-67.
10. Gaiduk A.R., Kapustyan S.G. Kontseptsija postroenija system kollektivnogo upravlenija bespilotnimi letatel'nimi apparatami. Informatsionno-izmeritel'nije i upravljajushije sistemi. 2012. №7. pp. 8-16.
11. Chernousko F.A., Kolmanovsky V.B. Optimal'noje upravlenije pri sluchajnih vozmushenijah. Moscow. Nauka. 1978.
12. Roytenberg Y. N. Avtomaticheskoje upravlenije. Moscow. Nauka. 1992.
13. Kompleksi s bespilotnimi letatel'nimi apparatami. Book 1. Printsipi postrojenija i osobennosti primenenija kompleksov s BLA. Pod. red. V.S. Verbi i B.G. Tatarskogo. Moscow. Radiotehnika. 2016.
14. Merkulov V.I., Verba V.S., Ilchuk A.R. Avtomaticheskoje soprovozhdenije tselej v RLS integrirovannih aviatsionnih kompleksov. T.1. Teoreticheskije osnovi. RLS v sostave integrirovannogo aviatsionnogo kompleksa. / Pod. red. V.S. Verbi. Moscow. Radiotehnika. 2018.
15. Krutko P.D., Maksimov A.I., Skvortsov L.M. Algoritmi I programmi projektirovenija avtomaticheskih sistem. Moscow. Radio i svijaz'. 1988.

E. Lyapustin, R. Meshcheryakov, M. Mamchenko

IDENTIFICATION OF UNCLEANED AIRCRAFT ON SOUND PRINTING FOR DETECTION, RECOGNITION AND DETERMINATION OF THE POSITION FOR PREVENTION OF COLLISIONS IN AIRSPACE

Infosystems Jet, Moscow, Russia, es.lyapustin@mail.ru
Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia, meshcheryakov.roman@gmail.com
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia

Abstract

The article proposes a method for solving the problem of sound recognition, produced by the motor and blades of the UAV, as well as the detection of its location, to adjust the trajectory of motion in order to avoid a collision. The methods of analysis and the system of coefficients of the audio stream are considered. Algorithms for preliminary signal processing and selection of criteria are given.

Keywords: audio analysis, UAV identification, noise elimination, sound filtering, UAV recognition, filtering algorithm, algorithm for eliminating quiet signal areas, audio signal, cepstral coefficients.

Unmanned aerial vehicles (UAV) are a rapidly developing technology that is currently widely used in the private, commercial and public sectors. However, there are no safety measures that facilitate the safe operation of these devices in populated uncontrolled airspace without potential danger to other manned or unmanned aerial vehicles. Conventional collision avoidance systems are only required for passenger aircraft weighing more than 5,700 kg and must be present on both in order to be operational [1].

To accomplish the task, it is necessary to create an individual system on each UAV to facilitate the detection and acceleration of the reaction rate, which will help save another approaching UAV. It is believed that acoustic sounding can be used to achieve this goal

The most common form involves the use of sensors, known as microphones, that detect pressure fluctuations of sound waves generated during UAV piloting. Acoustic sensing has many potential advantages over more traditional non-cooperative technologies such as electro-optical (EO), infrared (IR) and radar. Since the sensors are typically one-way, can be achieved in a full spherical coverage of the sensors. This is a very important feature because most air collisions occur from behind, from the side, above or below; locations that are usually out of sight for most other sensing technologies [2]. Sensor systems are usually very small and light, as they consist of several microphones and a recording/data processing unit. Requirements for the collection and processing of data is also much less than for EO or IR due to the decrease in data transfer speed from sensors. With the simultaneous use of multiple spatially spaced microphones in the array configuration, it is possible to detect, locate and track a sound source such as an aircraft [3-8]. In some cases, analysis of the Doppler frequency shift of the source signal over a period of time can also determine the speed and direction of the sound source [9-13].

To implement the UAV collision avoidance system, it is necessary to solve the following tasks:

1. To identify the sound of the blades and the engine blah BLAH from other parasitic noises of the environment.
2. Determine the location relative to another UAV.
3. Perform a collision avoidance maneuver

1. Identification of unmanned aerial vehicles

Analyzing the physical features of the sounds produced by the UAV, the UAV's audio data stream was recorded. Through the Fourier transform, the sound was decomposed into sound components [14].

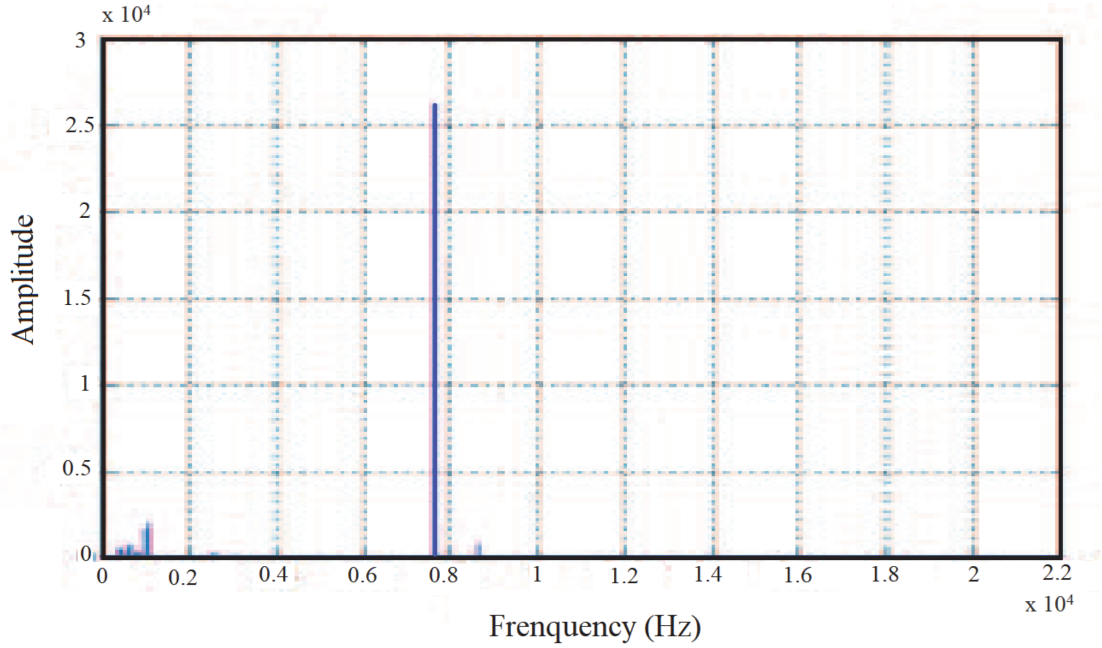


Figure 1 – The sound of one engine with a UAV blade

In Figure 1, a value between 7 and 8 kHz, which corresponds to the sound of one motor with UAV blades. However, in order to recognize a UAV group, it is necessary to analyze the sound stream with at least four blades.

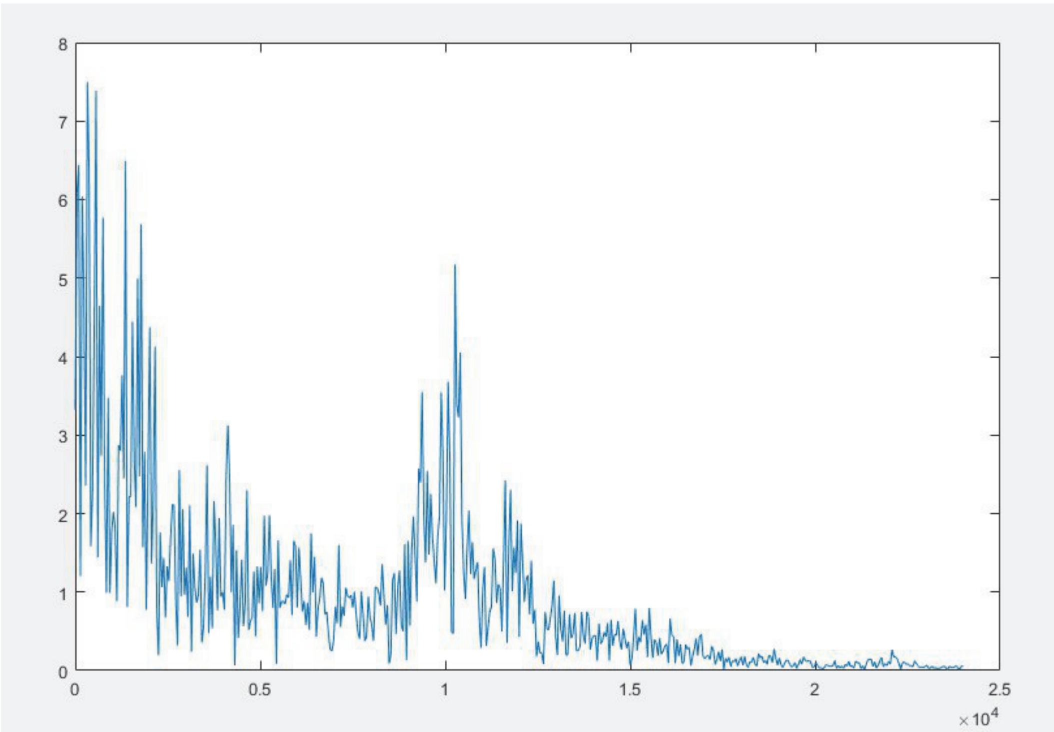


Figure 2 – The sound of several motors with UAV blades

The unsteady noise spectrum depends on the rotation speed of each engine and blades, which varies with time [15]. In this case, it is not possible to find one or two separate frequency parts. Nevertheless, you can find a reference point where you can start analyzing the sound stream. Most useful frequencies range from 5 kHz.

The main problem of processing the received audio data from the data acquisition and transmission module is the choice of functions for constructing an accurate identification system. The audio functions must be efficient, reliable and physically interpreted in order to accurately differentiate from the environmental audio data the key properties of the audio signal produced by the UAV. In general, environmental sounds can generate a similar spectrum of UAV signals, with different sources in different contexts, and no assumptions can be made about their spectral and temporal structure. [16]

The received audio signal is not monotonic, which complicates the identification of UAVs, so the signal can be considered locally stationary only for a short time interval of 10-30 ms [17]. It is understood that the time and spectral behavior of the audio signal can be considered almost uniform in the time range of several milliseconds. Therefore, to capture heterogeneous audio information in the phase of analysis of environmental sound, two different time scales are considered:

1. Short audio segments are processed in a short period of time of 20 ms to determine which set of time and frequency characteristics is effective in the task of voice identification.
2. A medium-term time analysis is performed for 200 ms in order to differentiate environmental noise, and subsequently, by filtering, process a discrete audio signal to identify the UAV.

After the acoustic signals were received in digital form, various processing steps were implemented to enable source detection and spatial localization.

Any designed system will function at the first stage receiving the sound emitted by the UAV, then the module that processes this sound, compares it with the UAV identifier-standard and using triangulation in terms of sound strength, provide information on the instant location of the device. For this, the detection system was divided into five separate modules:

1. Sound detection;
2. Signal conditioning;
3. Collection and transmission of data;
4. Processing;
5. Data output.

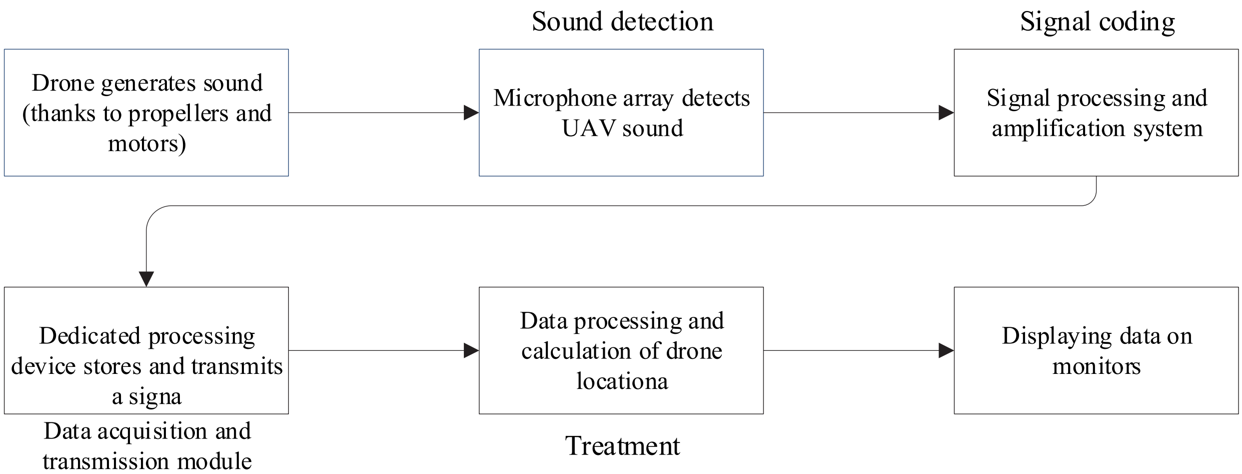


Figure 3 – Logical identification scheme of UAVs by sound fingerprint

The short-term analysis, denoted by x(n), is a normalized discrete-time audio packet, the audio frame s(n) of each subframe is determined by the equation.

$$s(n) = x(n) \cdot w(m-n) \tag{1}$$

$$w(n) = 0.54 - \cos\left(\frac{2\pi(m-n)}{L-1}\right) n \in [0, L) \tag{2}$$

where $w(n)$ is the Hamming window of length L, and m is its time shift. To calculate the raw features and locally characterize the corresponding audio waveform and spectrum shape, each subframe $s\ (n)$ is processed by a set of special filters based on certain features described in detail below.

To identify the UAV, we will consider determining the following unknowns, based on the physical properties of the audio signal:

1. Short-term energy;
2. Zero crossing level;
3. Time centroid of gravity for the time domain;
4. Spectral centroid;
5. Spectral decay, Mel-cepstral frequency coefficients, for the frequency domain.

Short-term energy is calculated in accordance with the expression and provides a measure of changes in the environmental sound energy over time.

$$STE = \frac{1}{L}\sum_{i=0}^{L-1}|s(i)|^2 \tag{3}$$

The time centroid is defined as the temporal balancing point of the amplitude distribution of the audio signal. It is expressed as:

$$\mathrm{C} = \frac{\sum_{h=1}^{L} h \cdot s(i)}{\sum_{h=1}^{L} s(i)} \tag{4}$$

The zero crossing level determines the average number of times the sound signal changes its sign in a short-term window. This feature is especially useful for identifying a voiced subframe.

$$ZCR = \frac{1}{2(L-1)} \cdot \sum_{i=0}^{L-1}\left|sgn\big(s(i)\big) - sgn(s(i-1))\right| \tag{5}$$

where

$$sgn[s(n)] = \begin{cases} -1, s(n) \geq 0, \\ \ \ 1, s(n) < 0. \end{cases} \tag{6}$$

The spectral centroid determines the balancing point of the sound spectrum *G* (*f*), indicates whether to contain lower or higher frequencies in the spectrum.

$$SC = \frac{\sum_f f \cdot G(f)}{\sum_f p(f)} \tag{7}$$

$$p(f) = \left|\sum_{i=0}^{L-1} s(i) \cdot e^{-j2\pi f/L}\right|^2 \tag{8}$$

The spectral decay determines the frequency below which a certain amount β of spectral energy is concentrated (8). This paper assumes that β = 0.9

$$SRO = \arg\max_m \sum_{f=1}^{m} p(f) \leq \beta \cdot \sum_{f=1}^{F} p(f) \tag{9}$$

Key to many speech recognition systems is the approach to calculating speech signal parameters based on kepstral analysis [18].

Mel-kepstral coefficients (MFCC) are discrete cosine transforms in the scale log power spectrum G (f).

Mel-kepstral coefficients are used to display the power spectrum of a *G*(*f*) signal. In order to decompose the resulting spectrum on the mel scale, we need to create a "comb" of filters. In fact, each mel-filter is a triangular window function, which allows you to sum up the amount of energy at a certain frequency range and thereby get a mel-coefficient. Knowing the number of chalk coefficients and the analyzed frequency range, we can build a set of such filters:

$$f_{Mel} = 1127{,}01048ln(1 + \frac{f}{700}) \tag{10}$$

– The energies *Em* are calculated at the output of the mth filter and then transferred to a logarithmic scale. Suppose that *Cm* is the relative value of the frame.

– The larger the index number of the chalk coefficient, the wider the filter base. This is due to the fact that the frequency range s (n) of interest to us into the ranges processed by the filters occurs on the chalk scale

$$c_i = \sum_{k=1}^{M} C_k \cdot \cos(\frac{\pi(2i+1)k}{2M}) \tag{11}$$

To solve the problem of identifying UAVs by sound, it is necessary to extract 12–13 MFCCs (Fig. 4), since it has been established that they contain identifying information for higher frequencies.

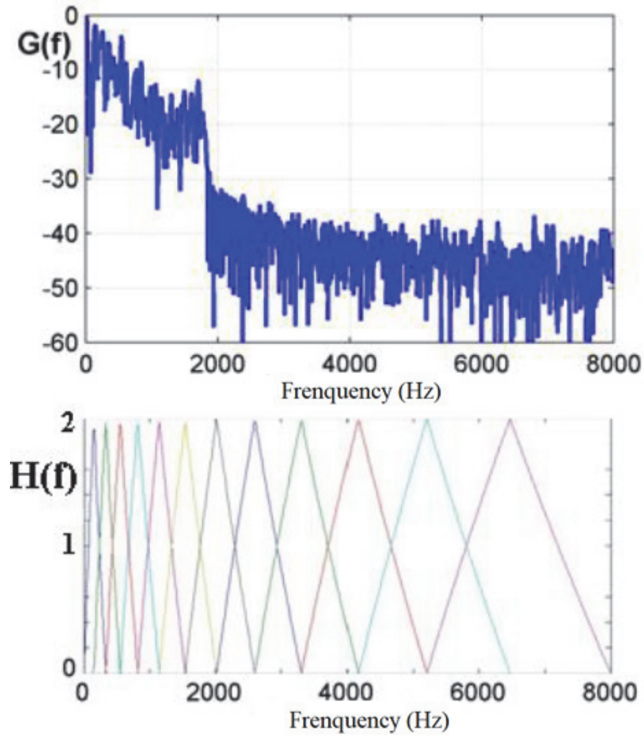


Figure 4 – 12 MFCC filters

2. UAV positioning

To build an acoustic system, the detection distances must be large enough to facilitate the evasion maneuver. There are currently no formal requirements, but it is generally accepted that 150 meters is the absolute minimum that must be constantly maintained between aircraft [1]

$$d_0 \approx 5.6\, V_{rel}\sqrt{\frac{\pi}{2} - \varphi} \quad (12)$$

where d_0 is the minimum distance at which a maneuver to avoid a collision should begin, V_{rel} is the relative speed of approach of two UAVs, and φ is the angle of heel (heel) initiated by the sensitive aircraft upon detection.

To limit the size of the microphone system, the maximum distance between any two microphones should be less than half the wavelength of the maximum frequency used during processing, in the case of this device 8000 Hz.

This limitation is required when the signal is periodic. The phase difference between the signal received by microphone 1 at time t1 and the other signal received by microphone 2 at time t2, at $t2 + kT$, will be the same for any $k \in \mathbb{Z}$, where T is the period of the signal.

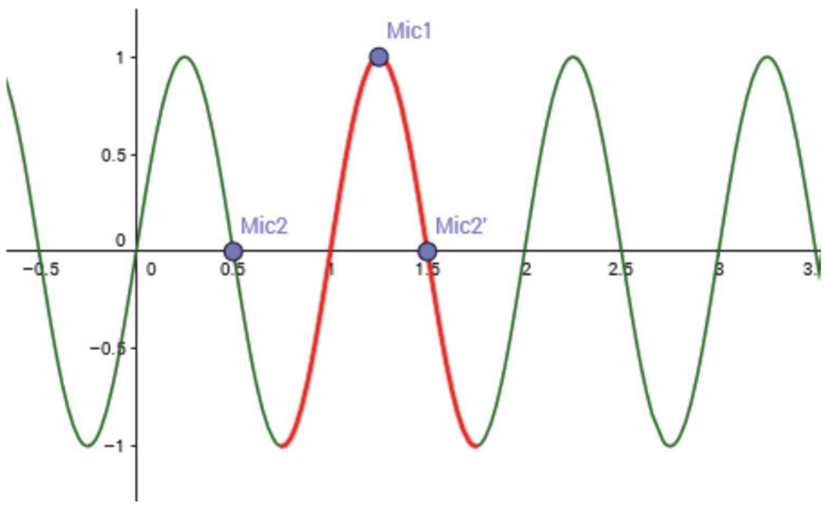


Figure 5 – Possible temporary positions taking into account a certain phase difference

As shown in fig. 5, when a sinusoidal signal with a frequency of 1 Hz is received, microphone 2 selects two samples, at t = 0.5 s and t = 1.5 s, microphone 1 will not be able to distinguish between two samples. because they have the same phase difference. This illustrates how a sample taken by microphone 2 can be mistakenly interpreted as the leading sample of microphone 1 when it is actually behind. Therefore, imposing a physical restriction, the signal of the microphone 2 can lie only in the red region ($t1 \pm T\ /\ 2$), eliminating any uncertainty (each phase difference will have a unique solution). Therefore, taking into account the maximum frequency under consideration, the maximum distance should be

$$d_{mic} = \frac{0.5*C}{f_{max}} \tag{13}$$

Where C is the speed of sound in vacuum, and f_{max} is the maximum sound frequency of the device.

Below are two diagrams illustrating the location and size of the microphone with microphone 1 selected as the reference microphone (located at the origin).

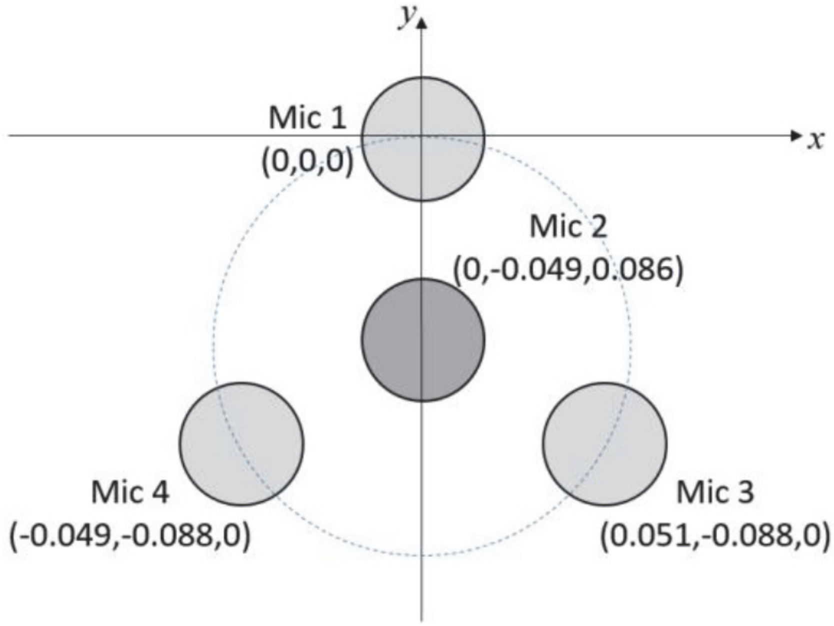


Figure 6 – Microphone system (Top view) (left) and side view (top)

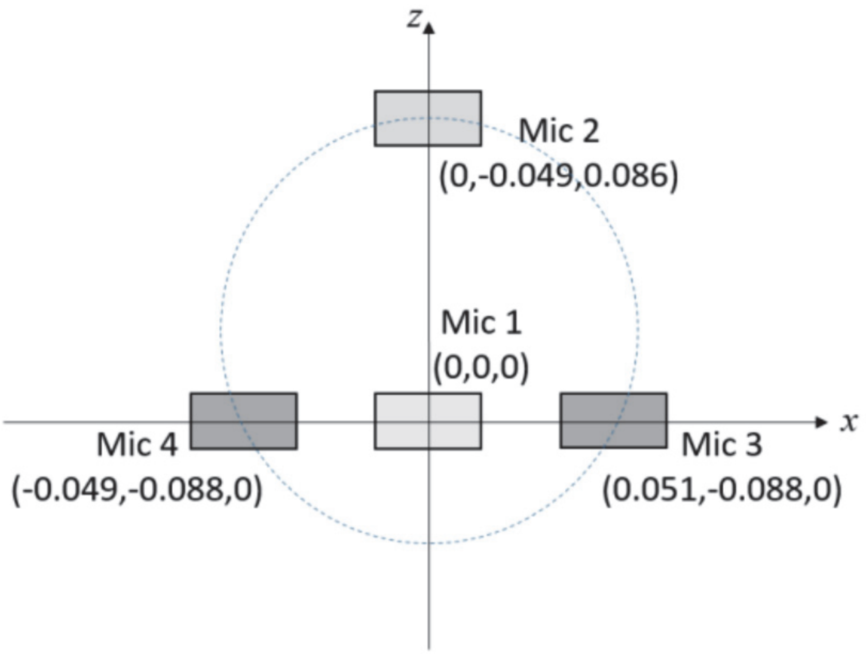


Figure 7 – Microphone system (side view)

For each angle of sound arrival to the structure of the tetrahedral microphone, there is a unique set of delays by which the audio signal reaches various microphones. Having found these delays between the signals, we can determine the angle of sound arrival.

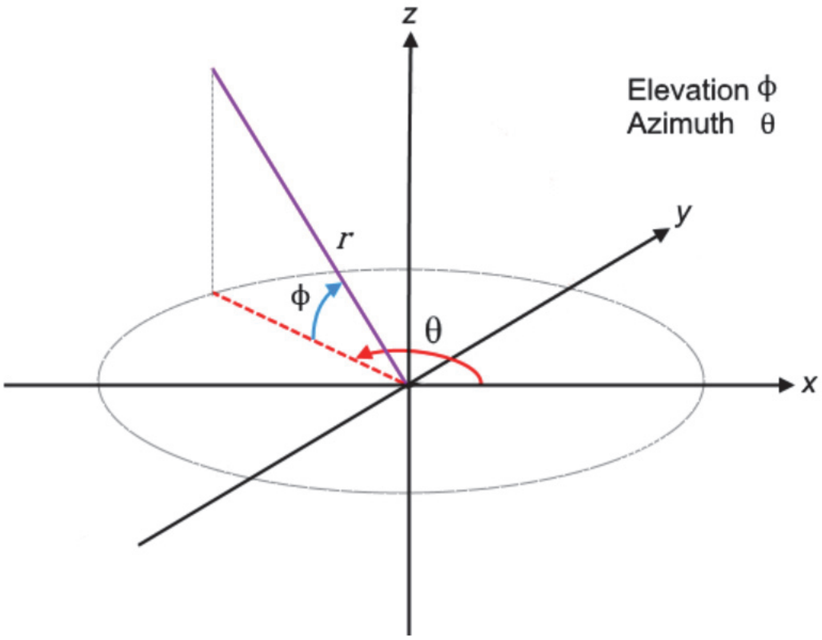


Figure 8 – Determination of the azimuth of the object

Step one is to set one of the microphones as the default reference microphone. When a sound wave is received by an array, other microphones will receive a signal with a negative or positive time delay compared to a reference microphone. Since the array is designed so that the maximum distance between the microphones is less than half the wavelength of the highest frequency, there will be a unique delay for each microphone and reference microphone for each direction in the sky. In a 2D plane:

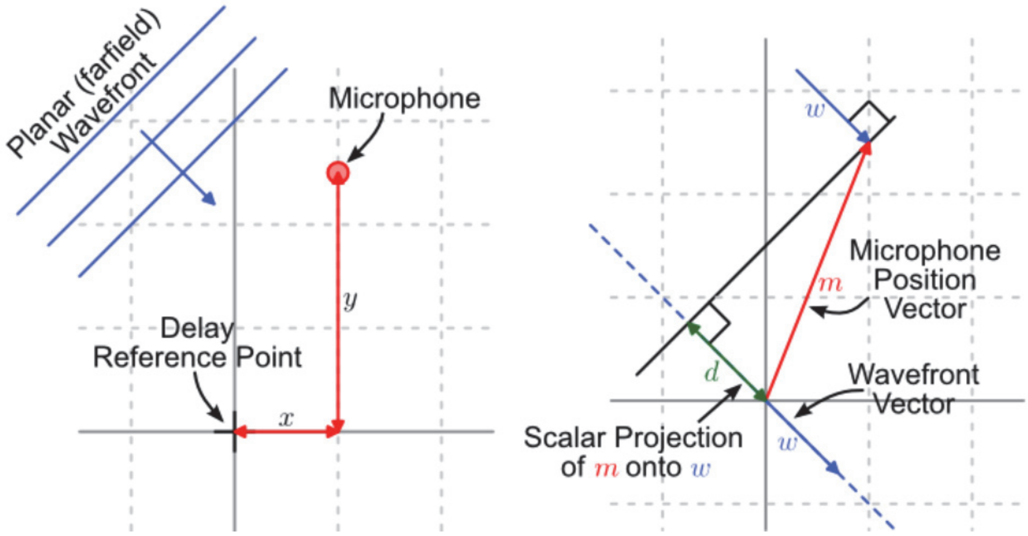


Figure 9 – Diagrams depicting the position of a 2D microphone (left) and calculating the delay of a point product (right)

The following relationship, obtained from the figure above, is then used to determine the position of the sound source (drone):

$$d_{dron} = \mathrm{C} \times x_0 = m \times w \tag{14}$$

Where x0 is the delay of the sound stream, C is the speed of sound, **w** is the unit vector indicating the direction of propagation of the wavefront of the sound signal, and **m** is the microphone position vector.

The result of the determination in two-dimensional space can be translated into three-dimensional, to determine the location of the UAV in a three-dimensional coordinate system.

From expression 14, it turns out:

$$x_0 = \frac{m \times w}{\mathrm{C}} \tag{15}$$

$$x_0 = \frac{m_x w_x + m_y w_y + m_z w_z}{\mathrm{C}} \tag{16}$$

There is a unique delay between each microphone and the reference microphone, so this result can be superimposed on each pair of microphones, which includes the reference microphone. The result is three equations with three unknowns wx, wy and wz, since three microphones other than the one selected were chosen as the reference. This system of equations can then be solved using the Gaussian elimination method or other mathematical methods to determine the components of the direction vector of the wavefront of the UAV sound signal and by expanding the direction of UAV movement.

The last step is to use the trigonometric ratios of the direction vector components to calculate the azimuth pair and the UAV elevation angle relative to the device (or, more precisely, the reference microphone) that provides the UAV location in terms of elevation and azimuth.

$$wx = -\cos\theta\cos\phi;$$
$$wy = -\cos\theta\sin\phi;$$
$$wz =- \sin\phi.$$

Where θ is the azimuth, and ϕ is the elevation angle from the source to the sound source. The negative sign is due to the fact that the UAV position vector has a direction opposite to the wavefront vector.

Summary

It is possible to detect and identify UAVS to prevent collisions [19]. Acoustic sensing can be a viable technology for creating an Autonomous collision avoidance system. However, the results regarding the system's ability to localize and track detected targets still need to be worked out to reduce the number of false positives. You can also use multiple arrays to further extend the identity system. For example, the modernization of the array of microphones electro-optical (EO), infrared and radar detection systems. Also, since the direction of the motor noise remains constant with respect to the array of microphones, this information can be used to further improve the localization characteristics of the source, by suppressing the noise of its own blades and motors is an interesting topic for future research.

References

1. Geyer, C.; Singh, S.; Chamberlain, L. Avoiding Collisions between Aircraft: State of the Art and Requirements for UAVs Operating in Civilian Airspace; Technique Reports; CMU-RI-TR-08-03; Robotics Institute, Carnegie Mellon University: Pittsburgh, PA, USA, 2008.
2. Finn, A.; Franklin, S. Acoustic sense & avoid for UAV's. In Proceedings of the 7th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), Adelaide, Australia, 6–9 December 2011.
3. Zelnio, A.M.; Case, E.E.; Rigling, B.D. A low-cost acoustic array for detecting and tracking small RC aircraft. In Proceedings of the 2009 IEEE 13th Digital Signal ProcessingWorkshop and 5th IEEE Signal Processing Education Workshop, Marco Island, FL, USA, 4–7 January 2009.
4. Case, E.E.; Zelnio, A.M.; Rigling, B.D. Low-cost acoustic array for small UAV detection and tracking. In Proceedings of the 2008 IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 16–18 July 2008.
5. Zelnio, A.M. Detection of Small Aircraft Using an Acoustic Array;Wright State University: Fairborn, OH, USA, 2009.
6. Sutin, A.; Salloum, H.; Sedunov, A.; Sedunov, N. Acoustic detection, tracking and classification of low flying aircraft. In Proceedings of the 2013 IEEE International Conference on Technologies for Homeland Security (HST),Waltham, MA, USA, 12–14 November 2013.
7. Salloum, H.; Sedunov, A.; Sedunov, N.; Sutin, A.; Masters, D. Acoustic system for low flying aircraft detection. In Proceedings of the 2015 IEEE International Symposium on Technologies for Homeland Security (HST), Waltham, MA, USA, 14–16 April 2015.
8. Nielsen, R.O. Acoustic detection of low flying aircraft. In Proceedings of the IEEE Conference on Technologies for Homeland Security, Boston, MA, USA, 11–12 May 2009.

9. Ferguson, B.G. A ground-based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller-driven aircraft. J. Acoust. Soc. Am. 1992, 92, 1403–1407. [CrossRef]
10. Sadasivan, S.; Gurubasavaraj, M.; Sekar, S. Acoustic signature of an unmanned air vehicle exploitation for aircraft localisation and parameter estimation. Def. Sci. J. 2002, 51, 279. [CrossRef]
11. Tong, J.; Xie, W.; Hu, Y.; Bao, M.; Li, X.; He, W. Estimation of low-altitude moving target trajectory using single acoustic array. J. Acoust. Soc. Am. 2016, 139, 1848–1858. [CrossRef] [PubMed]
12. Ferguson, B.G.; Lo, K.W. Turbo-prop and rotary-wing aircraft flight parameter estimation using both narrow-band and broadband passive acoustic signal processing methods. J. Acoust. Soc. Am. 2000, 108, 1763–1771. [CrossRef] [PubMed]
13. Reiff, C.; Pham, T.; Scanlon, M.; Noble, J.; Landuyt, A.V.; Petek, J.; Ratches, J. Acoustic Detection from an Aerial Balloon Platform; US Army Research Laboratory: Adelphi, MD, USA, 2004.
14. P. Misra, A. A. Kumar, P. Mohapatra, and P. Balamuralidhar, "Aerial drones with location-sensitive ears," IEEE Communications Mag. vol. 56, no. 7, pp. 154-160, Jul. 2018.
15. S.Allen Broughton, Kurt M. Bryan-Discrete Fourier Analysis and Wavelets,Wiley, 2008, ISBN: 978-0-470-29466-6.
16. G. Sinibaldi and L. Marino, "Experimental analysis on the noise of propellers for small UAV," Appl. Acoust., vol. 74, no. 1, pp. 79–88, Jan. 2017.
17. Rabiner L.R. Hidden Markov models and their application in selected applications for speech recognition: Overview // IEEE, vol.77, no. 2 February 1989 — pp. 86–120.
18. Lie N. V. Speech pre-processing for speech recognition system / N. V. Lie, D. P. Panchenko // Young scientist. — 2011. — no. 5. vol. 1. — pp. 74-76.
19. Deller, R. Jr., Hansen, H. L., Proakis, G. Discrete -Time Processing of Speech Signals / John R. Deller, Jr., John H.L. Hansen, John G. Proakis. – Wiley-IEEE Press, 1999. – 936p.
20. Meshcheryakov R.V., Bondarenko V.P. Dialogue as a basis for construction of speech systems // Cybernetics and Systems Analysis 44 (2), 175-184

A.E. Ananenkov, D.V. Marin, V.M. Nuzhdin., V.B. Schneider

INTERFEROMETRIC RSA FOR THE ICE SITUATION MONITORING

Moscow Aviation Institute (National Research University), Moscow, Russia
vschndr@gmail.com

Abstract

The goals, tasks and ways of constructing a miniature RSA suitable for installation on small-sized unmanned aerial vehicles are considered. The estimates of the technical and some tactical characteristics of the radar being created are given. The results of flight experiments with radar prototype equipment are discussed.

Keywords: ice conditions, monitoring, emergency oil spills, miniature radar, airborne radar system, unmanned aerial vehicle, all-weather round-the-clock surveillance, radar image, antenna aperture synthesis, interferometric signal processing.

Acknowledgments

This report is based on the results of applied research and experimental work (PNIER) on the topic: “Creating a scientific and technical reserve in the field of building a unified miniature onboard radar target load of small-sized unmanned aerial vehicles for monitoring ice conditions in the construction and operation of oil and gas platforms”. Agreement No. 14.577.21.0226. The unique identifier of PNIER: RFMEFI57716X0226.

Introduction

To conduct successful economic activities in the Arctic and solve a number of scientific problems, it is necessary to monitor the ice situation (MIS) – regular observations of the distribution, characteristics, and dynamics of sea ice. Depending on the frequency and area of simultaneous surveys, the MIS systems are divided into three main categories: strategic, operational and tactical.

From the beginning of the twentieth century, the main means of tactical ice reconnaissance was the visual observation of sea ice from aircraft carriers. However, the difficult weather conditions in the region required the use of meteo-independent observation tools - radar. In 1966, since the launch of the Cosmos-122 satellite, the use of satellite imagery for strategic monitoring and mapping of ice formations began. At the same time, the main drawback of orbital carriers became obvious - the lack of efficiency in obtaining data associated with both the parameters of the orbit and the process of receiving and transmitting the MIS data. Therefore, the orbital MIS system was supplemented since 1968 by side-view aviation radar stations (without aperture synthesizing) Toros («Торос») and Thread («Нить»), which made it possible to proceed to operational radar ice observations. Later, the aviation complex of the two-frequency radar station of the side view “Iceberg - cut” (JSC "Scientific-Production Enterprise" Radar MMS ") [1] was created.

In the high-latitude regions, it is still difficult to ensure the necessary regularity and efficiency of radar surveying of ice formations by a group of spacecraft (despite the launch of the Russian “Severjanin-M” satellite radar for ice reconnaissance developed by NII TP [2]), and difficult weather conditions do not allow to ensure the regularity of observations by optical means. Regular ice monitoring is necessary to increase the speed of vessels on the Northern Sea Route, to ensure the safety of operations on the shelf drilling platforms, when establishing communication with islands away from the mainland, in emergency situations, in the interests of special ministries and departments. In addition, the operational ice monitoring complex will significantly simplify navigation in the mouths of northern rivers, the Bay of Ob river and similar regions of the North.

Until now, the main source of tactical ice information is visual ice reconnaissance using manned helicopters based on icebreakers. Visual aviation inspection is also used to ensure safety during the operation of drilling platforms. This monitoring technology is associated with high costs for the operation of aircraft, the need to maintain appropriate infrastructure and the inevitable risk to the lives of pilots and observers in the difficult weather conditions of the Arctic.

The need for the widespread introduction of unmanned aerial MIS vehicles in this area is due to the ability to reduce the risk of loss of human life and the cost of monitoring, taking into account the lower cost of flight hours, to eliminate the need to create and maintain readiness of the landing field or helicopter platforms. The most promising in terms of reducing the costs are small-class UAVs, the explosive growth of interest in which is observed all over the world today.

International experience shows the desire to create radar monitoring tools for unmanned carriers and, if possible, reduce their weight and size characteristics, since UAVs with low take-off weight do not allow the use of radar samples developed for traditional aircraft. As a result, the creation of compact radars with synthesized antenna aperture (RSA), which allows reducing the size and weight of the antenna system qualitatively, requires the development of specialized solutions. Practical work on their creation is carried out by experts in Germany, USA and Russia. MicroSAR and NanoSAR [3, 4] can be mentioned as the most well-known examples of successfully implemented projects.

An additional advantage in expanding the scope of such RSAs usage can be obtained by implementing the restoration of the heights of the underlying surface in a single pass by interferometry methods [5]. The most obvious areas of application for interferometric small RSAs include:

– operational monitoring of the ice situation to search for divorces that pose a threat to icebergs, both in the interests of escorting ships and ensuring the safety of oil-producing platforms in the northern regions of the world ocean;

– operational monitoring of the situation in flood zones, ice jams on rivers, avalanche zones, landslides and mudflows;

– operational monitoring of oil pipelines to detect fistulas (oil leaks) and critical deformations caused by the soil subsidence.

Such a tool will be especially effective in obtaining comparative images of the same area of terrain periodically with a small time interval for observing the situation dynamics.

MAI(NRU) technical background

At present, the MAI (NRU) has developed and tested a compact multi-mode on-board radar system (MMBRL) of the Ku-wavelength range for remote sensing of the Earth from a UAV. The key features of the MMBRL are the ability to monitor the earth's surface in various modes, and the transition from mode to mode is carried out by the operator in real time, without requiring landing of the UAV carrier. Its appearance is shown in Figure 1, the characteristics are shown in Table 1.

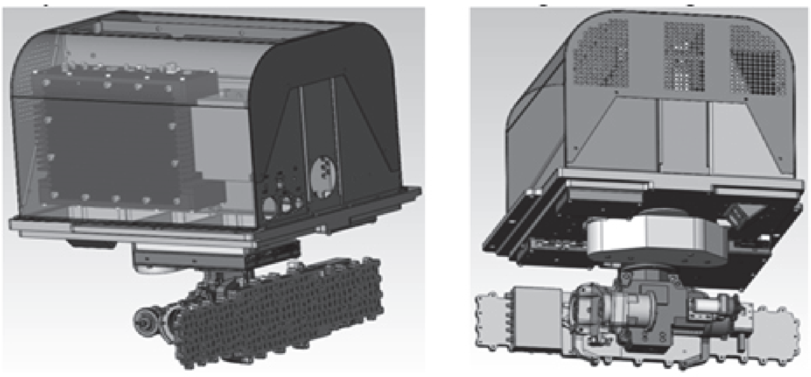


Figure 1 – 3D MMBRL layout model

Table 1. Main MMBRL technical characteristics

Parameter	**Value**
Aircraft speed (V), km / h	80…220
Flight height (H), m	30…7000
The angle of the antenna survey in azimuth, degree	-85…+85
Signal wavelength, cm	2
Transmitter impulse power (Pi), W	200
Surface mapping range, km	40
Maximum geometric resolution, m	0,5
Spectrum width, MHz	640
Weight, kg	40

Analysis of the characteristics given in Table 1 shows that the mass of the MMBRL is too large to be installed on small-class UAVs, so it was decided, using the experience of previous developments, to create an interferometric radar with an aperture synthesizing and not exceeding 10 kg in weight.

Characteristics and functional diagram of the radar prototype being created

The prototype of the radar being created represents a panoramic short-range front-end radar sensor with a chirp probing signal and a homodyne reception method [6, 7]. A structural diagram of the prototype [6] is shown in Figure 2.

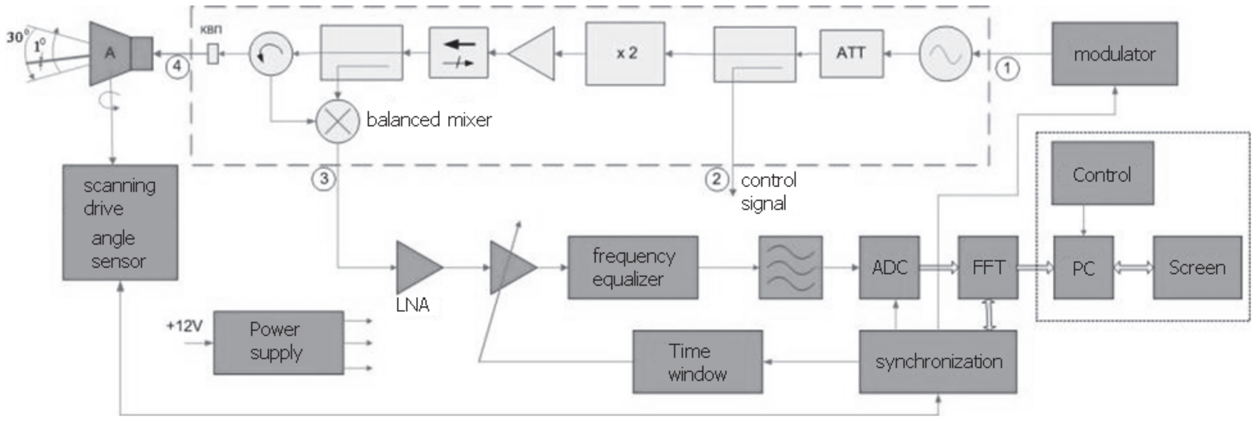


Figure 2 – A generalized block diagram of the created radar prototype equipment

The receiving-transmitting antenna is made in the form of a narrowly directed traveling-wave antenna. To control the scan in the azimuth plane, the antenna position sensor is installed. The microwave unit consists of three separate modules: a transmitter module, an attenuator module and a balanced mixer module, which are desirable to be integrated into one for mass production. The transmitter module contains a high frequency generator (HFG) with an amplifier. In the transmitter unit, frequency modulation of the probing signal is provided using the “Chirp Modulator”. A periodic ramp signal is used as the modulating signal. The balanced mixer module is designed to multiply the input signal reflected from the target with the reference signal previously obtained from the directional coupler. The digital signal processing unit (DSP) consists of three main modules: an analog-to-digital converter (ADC), a clock pulse generator (CPG), and an interface with a digital computer.

The clock pulse generator (CPG) provides the formation of all service and control signals and frequencies. The COS modulator is used to form the time window with which the beat signal is multiplied. Using the time window allows to suppress the parasitic amplitude modulation of the transmitter and reduce the level of the beat signal spectrum side lobes.

An amplifier with quadratic frequency response is used to compensate the dependence of the beat signal amplitude on the distance to the target from which the reflected signal comes. Using this amplifier allows to reduce the dynamic range of the signal applied to the ADC input.

The power supply provides the possibility of long-term operation from both AC 220V/50Hz and the aircraft’s on-board electrical system.

The on-board digital computer calculates the spectrum of the beat signal for each reflected signal, and thereby forms one column in range of the radar image being generated. Accumulating separate columns of radar image in the process of antenna scanning in the azimuthal plane, computer forms the primary radar image in azimuth-distance coordinates and processes it for presentation on the indicator. The problem of creating a wideband amplifying device with a wide dynamic range can be circumvented in our case by using homodyne reception (multiply reducing the band of received signals) and frequency-spreading the received signals, which allows a significant narrowing of the dynamic range of the amplified signals at the output of the last stages of the amplifying device.

Construction of interferometric RSA

To create such a RSA, homodyne radar equipment with a fixed antenna located along the flight line must be supplemented with a second identical receiving path and a second antenna. The antennas of the receiving paths should be spaced vertically on the base, ensuring the unambiguity of the height measurements at the chosen distance. RSA being created, should form radar images (RI) of the underlying surface with a resolution in range and azimuth of the order of 1 m in a 3 km viewing range from a height of about 1 km. The width of the microwave signal spectrum necessary to ensure the required resolution in range (at the near edge of the capture band), taking into account the loss on apodization, is 300 MHz. Taking into account the difficult weather conditions of the Arctic region, we will choose an emission wavelength of - 3 cm, which has a low attenuation.

The evaluation of the RSA energy when mapping the underlying surface, carried out according to the radar formula [5], shows that with the required signal-to-noise ratio q=20 dB at the maximum working distance, antenna gain G=26.4 dB, UEPR of the underlying surface σ = -20 dB, the power of continuous radiation with the selected modulation should be P = 15 W. The main technical characteristics of the interferometric MIS RSA are summarized in Table 2.

Table 2. Main technical characteristics of RSA for MIS

Parameter	**Value**
Aircraft speed (V), km / h	30…50
Flight height (H), m	500…1500
The width of the antenna radiation pattern in azimuth, degree	4
The width of the radiation pattern of the antenna on the slope, degree	23
Radiation wavelength, cm	3
Transmitter power, W	15
Surface mapping range, km	4
Maximum geometric resolution, m	1
Spectrum width, MHz	300
Weight, kg	less than 10

The requirements for the RSA antenna system are contradictory: on the one hand, it is necessary to provide sufficient suppression of the side principal maximums of the synthesized radiation pattern and the greatest possible gain of the useful signal (which requires an increase in antenna area); from the other hand, real antenna beam should be wide enough to minimize losses from random rolls and fluctuations of the course angle during the movement of a light carrier in air flows at low altitude.

To increase the RSA energy, we choose an azimuthal opening of 0.5 m. The width of the beam in azimuth will be about 4 degrees. Such a wide beam in azimuth allows the use of incoherent accumulation of up to 4 frames of radar images to reduce the level of speckle noise. In this case, the requirements for the accuracy of maintaining the course angle will be less than ±1.5 degrees. To ensure the required capture bandwidth, the antenna beam width should be at least 23.5 degrees, which is provided in the X-range of wavelengths with an aperture of 9 sm.

In our case, the problem of creating a broadband amplifying device with a wide dynamic range can be circumvented by using homodyne reception (multiply reducing the band of received signals) and frequency-spreading received signals according to the technology described in [6]. This allows the output of the last stages of the amplifying device to achieve a significant narrowing of the dynamic range of the signals [7, 8].

Let us estimate the beat signal band Fb to be amplified, digitized and processed:

$$F_b = \frac{2\Delta F F_{\text{M}} R}{c}; \tag{1}$$

where:

ΔF – is the microwave signal band that provides the necessary resolution by the delay (300 MHz),

F_{M} – is the modulation frequency of the radiated signal (1 kHz),

R – is the maximum distance to the observed objects (3740 m),

c – is the speed of light.

Thus, the required band Fb for amplifying and processing the received signals will be 7.5 MHz, which allows the use of cheap, light and low-consuming equipment for amplifying, digitizing and processing radar signals.

Let us consider an algorithm for forming an estimate of the relief height, assuming that the on-board computer has two associated matrices of complex amplitudes of signals received at the antennas A1 (i, j) and A2 (i, j). The indices i and j correspond to discrete readings of the range and azimuth, respectively.

The computer must implement the procedure of multiplying the conjugate complex amplitudes, as a result – we get the matrix of the following form:

$$\dot{A}_1(i,j)\cdot A^*(i,j)=\left|a(i,j)\right|^2\cdot e^{j\cdot[\phi_1(i,j)-\phi_2(i,j)]};\qquad(2)$$

where $\left|a(i,j)\right|^2$ – power factor proportional to the intensity of signals received on two antennas, or the radio brightness field,

$\Delta\phi(i,j)=\phi_1(i,j)-\phi_2(i,j)$ – phase distribution or phase difference field.

Let us consider the dependence of the phase difference of the signals for some fixed value of the azimuthal coordinate j = const and variable slant range R = i ΔR. The analysis of this dependence is carried out at different angles of inclination of the antennas aperture base - θo, relative to the direction of the local horizontal. The geometrical relationships used are explained in Figure 3.

It is easy to show that the phase difference $\Delta\phi(i, j=const)$ will be determined by the following expression:

$$\Delta\phi(R, j=const)=kD\cdot\left[\frac{H}{R}\cdot\cos(\Theta_0)-\sqrt{1-\left(\frac{H}{R}\right)^2}\cdot\sin(\Theta_0)\right];\qquad(3)$$

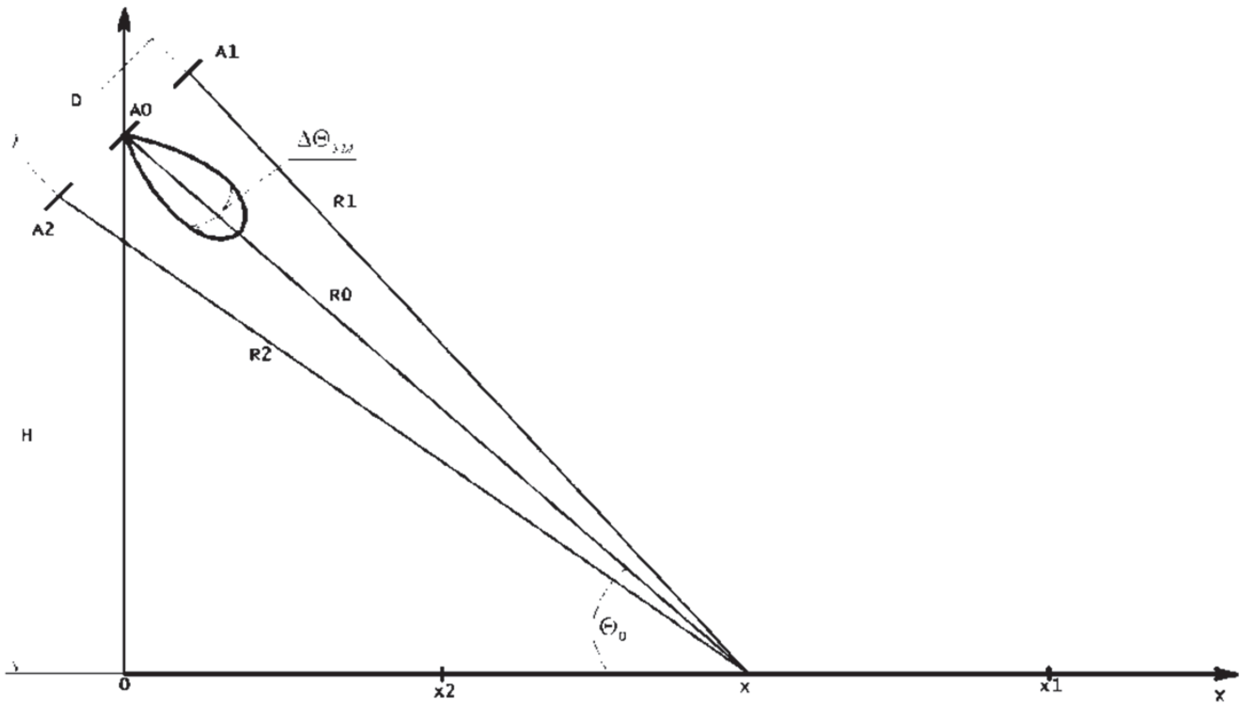


Figure 3 – Geometric relations in the interferometer

where: k = 2π / λ is the wave coefficient, D is the separation base of receiving antennas, H is the antenna system phase center height above the reflecting surface.

The graphs of the normalized phase difference, with different angles of base inclination - θo, are shown in Figure 4. These graphs were obtained in the absence of uneven terrain (flat horizontal reflecting surface) and constant altitude of air carrier (H = 1000 m).

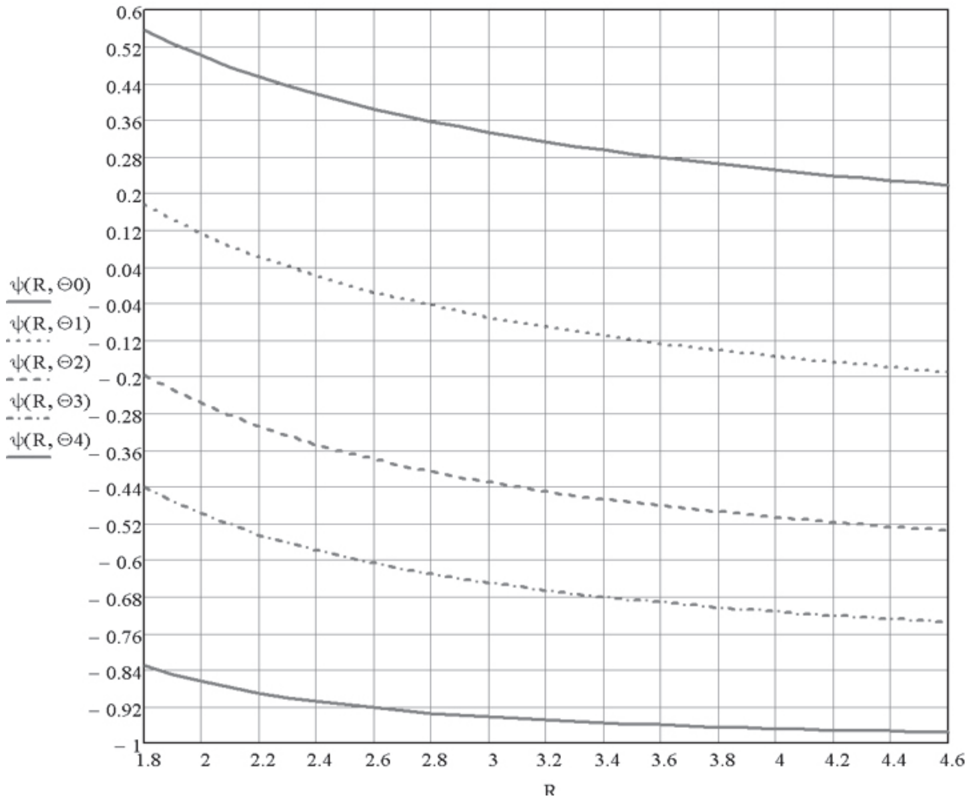


Figure 4 – The normalized dependence of the phase difference (in radians) at a constant height H and a variable slant range R in kilometers (for setting angles θo = 0; 23.5; 45; 60 and 90 degrees) at the ratio d/λ = 20

To obtain the relationship of the phase difference with the height of the irregularity h (discriminatory characteristic), we differentiate the function of the phase difference with respect to the variable H and write the product of the derivative by the variable h, which will give us the desired characteristic.

$$\Delta\psi(h) = kD \cdot \frac{H}{R} \cdot \left[\cos(\Theta_0) + \frac{H/R}{\sqrt{1-\left(\frac{H}{R}\right)^2}} \cdot \sin(\Theta_0) \right] \cdot \frac{h(R)}{H}; \quad (4)$$

Expression analysis shows that the maximum slope of the discriminatory characteristic is realized at θo = 0, increasing the angle of inclination to θo = 90 ° leads to a decrease in the slope and, accordingly, to a decrease in the sensitivity of the interferometer. The slope decreases inversely with the range at which the estimated elevation of the relief above the average level is located. An example of a discriminatory characteristic for installation angles θo = 0; 23.5; 45; 60 and 90 degrees, D / λ = 20 and for an average range of R = 3600 m, and H = 1000 m, are shown in Figure 5.

A change in the angle of inclination in the range of values θo = 0 ° ÷ 60 ° does not reduce the interferometer sensitivity, therefore, the angle θ ° = 60 ° can be considered as the working angle of the antenna base inclination.

Since the conducted studies showed the possibility of implementing equipment with the required characteristics, the manufacture of an experimental sample of an interferometric RSA MIS was started. The appearance and characteristics of the antenna module directivity pattern are shown in Figure 6. In parallel, the development and testing of software and algorithmic tools for obtaining radar images, including the reconstructed information about the height of the underlying surface, was carried out. At the first stage, models of background-target environment were used to test the software, and at the final stage, real radar signals registered in flight from a prototype radar installed in a flying laboratory were used. Figure 7 shows a

RSA image of a terrain area (brightness corresponds to the intensity of reflection) and the same terrain area with interferometric processing (brightness corresponds to height).

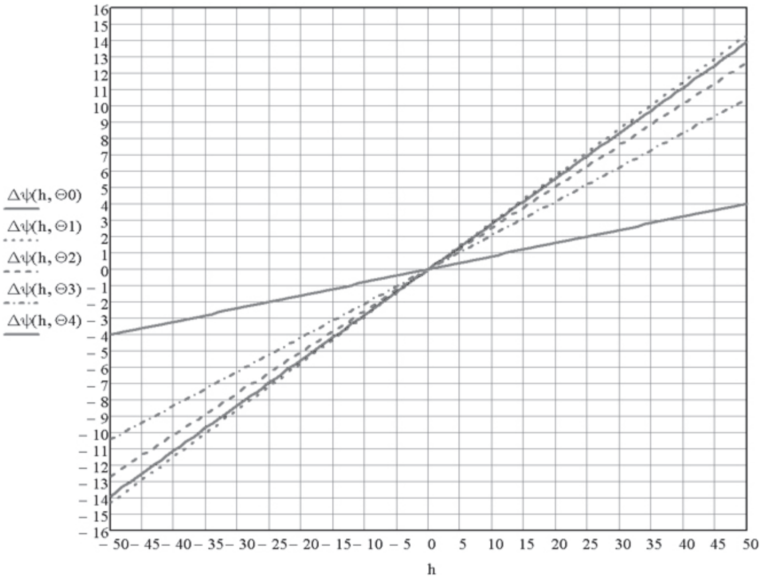


Figure 5 – Discriminatory characteristic (dependence of the phase difference in radians on the surface height in meters) for angles of 0; 23.5; 45; 60 and 90 ° with the ratio d / λ = 20

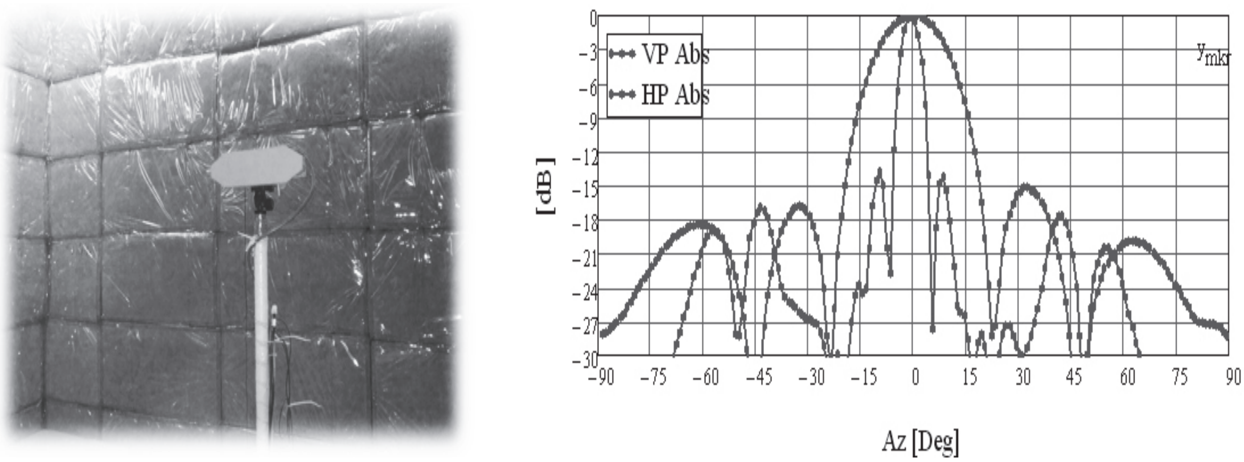


Figure 6 – The photo of the antenna element during measurements of its parameters (left) and the radiation pattern in the elevation and azimuthal planes (right)

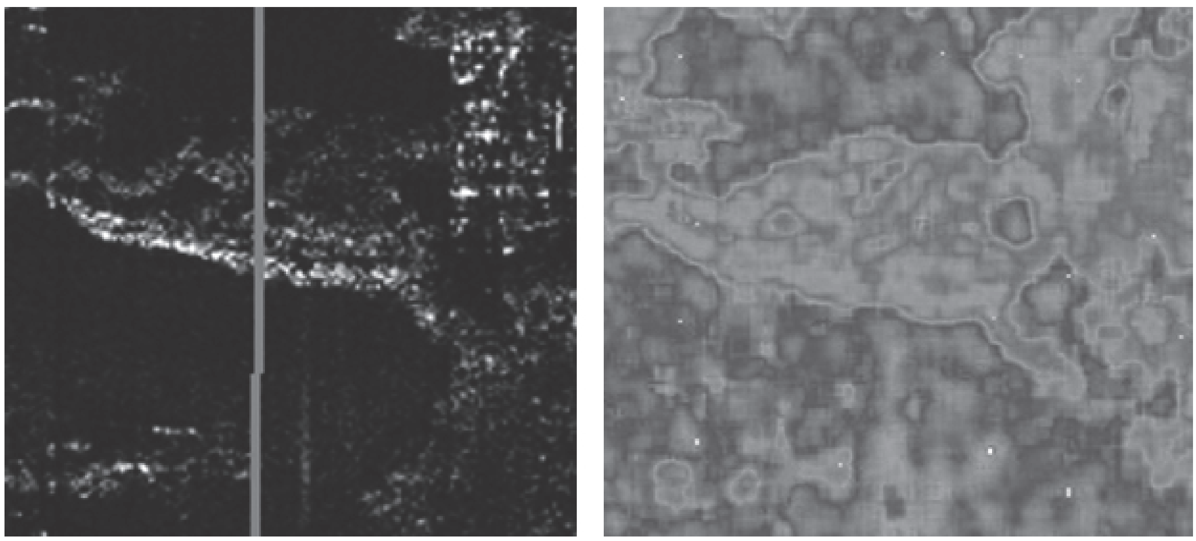


Figure 7 – RSA image of the terrain (brightness corresponds to the intensity of reflection) and the same terrain with interferometric processing (brightness corresponds to the height)

Conclusion

The appearance, main technical characteristics and proposals for using a miniature interferometric radar with antenna aperture synthesizing are formulated to ensure all-weather operational monitoring of the ice situation to search for divorces and icebergs that pose a threat to navigation, both in the interests of vessel navigation and monitoring of the safety of oil production platforms in northern areas of the oceans.

Eliminating the risk to human lives, together with the possibility of using small class UAVs as carriers with low cost of the complex and flight hour, will dramatically improve the information support of ice management systems. The use of radar equipment with interferometric techniques for measuring the height of the underlying surface and synthesizing the antenna aperture will allow creating highly detailed maps of ice formations, significantly expanding the possibilities of their classification, and allowing timely detecting icebergs and assessing their degree of danger.

The implementation of real-time processing on board of the UAV significantly reduces the flow of transmitted data, automation and interpretation of the received images, provides the ability to change interactively the flight task during the course laying from the icebreaker vessel.

The implementation of the proposed solutions allows:

– to ensure the possibility of producing a radar complex of operational ice monitoring on the basis of a small UAV with a minimum life cycle cost and increased reliability;

– an increase in the reliability of the MIS system by an order of magnitude, an increase in the speed and safety of vessel navigation, to ensure the safe and uninterrupted operation of the production platforms on the shelf;

– to reduce operating costs, ensure the possibility of long-term operation of the MIS system with maintenance according to the current state, thanks to diagnostics using built-in means of functional control;

– to eliminate the use of scarce foreign radio components under “sanctions”, increase the technological security of the state.

Therefore, the development of this topic seems relevant and complies with the “Fundamentals of the state policy of the Russian Federation in the Arctic for the period up to 2020 and beyond” in the development of the information and telecommunication environment in the Arctic.

References

1. http://www.radar-mms.com
2. http://www.niitp.ru.
3. The features of characteristics aviation radar and its application- Neiman P.I., Geomatics. 2011. №3.
4. www.sandia.gov/radar/
5. Radar Handbook, Third Edition. Editor Merrill I. Skolnik, Copyright 2008 by the McGraw-Hill Companies, ISBN 978-0-07-148547-0.
6. The features of of the FMCW radar design - Ananenkov A.E., Konovaltsev A.V., Kukharev A.V., Nuzhdin V.M., Rastorguev V.V., Skosyrev V.N./ 2-nd Russian Scientific and Technical Conference «Radiovysometryia- 2007», Kamensk - Uralsky, October 2007г.
7. Patent for invention: Homodyne radar registration number No. 2626405 of 27.07.2017
8. Highly informative short-range radars - Ananenkov A.E, Nuzhdin V.M , Rastorguev V.V., Skosyrev V.N /MAI, Moscow, 2018.

A.B. Belskiy

TASKS OF CREATION OF ROBOTIC HELICOPTER COMPLEXES

JSC Moscow Helicopter Plant M.L. Mile, Lyubertsy, R.-P. Tomilino, Russia
abelskiy@mi-helicopter.ru

Abstract

The analysis of the state and directions of development of robotic helicopter complexes. The basic principles and tasks of building robotic helicopter complexes are defined.

Keywords: helicopter, onboard systems, automation, airline artificial intelligence, decision support system.

In modern conditions and expanding the conditions for the use of helicopters in various geographic regions and climatic zones, the processes of information-management interaction in the helicopter-crew system in a real situation are becoming increasingly complicated.

One of the ways to improve piloting efficiency and perform tasks is automation of control processes and intellectualization the functions of onboard system set for decision making by the helicopter crew during individual, group and network-centric use of helicopter complexes and under natural and artificial interference.

The gradual robotization the functions of onboard systems should be directed primarily to the helicopter performing tasks in an automatic mode: piloting safety (take-off, flight, landing, hovering), flight along a given route, use of special and search and rescue facilities, etc.

Recently, helicopters have introduced automatic monitoring of the technical condition and diagnostics of the main units and power equipment, as well as functional systems of the onboard systems.

The further process of automating the onboard systems functions of upgraded and new types of helicopters will include solving problems in terms of state recognition, communications, information support, offshore, search and rescue, and other special tasks.

For optimal implementation of the whole complex of tasks in the direction of the robotization of the onboard systems, structural-functional restructuring and integration of avionics, systems of onboard radio-electronic equipment (hereinafter - avionics), optical-electronic, radar, navigation and other systems is required. Intellectualization of onboard systems, automation of piloting and performance of special tasks, operation and efficiency of decision making are the basic criteria for the gradual robotization of the onboard systems and helicopter complexes as a whole.

The main problem of the timing of technical implementation in the field of robotization of the helicopter complexes is the current gap between modern scientific theoretical achievements in the field of intellectualization of complex systems and applied solutions for their implementation in the onboard helicopter complexes.

The introduction of elements of artificial intelligence in onboard systems the function of a helicopter should significantly increase the level and ability of helicopters to perform piloting and special tasks (by reducing the time to make a decision (excluding the “human factor” when detecting objects, maneuvering, landing, special work, etc.). the crew remains with only the control functions and, if necessary, the operational adjustment of the decision support system.

The tasks of assessing situational awareness, monitoring the condition of helicopter onboard systems, optimal development and decision-making should be primarily automated in the process of controlling helicopters. In modern high-speed conditions, when solving special and search and rescue tasks, crews are affected by significant information and intellectual overloads; therefore, to fulfill an ever-increasing spectrum of external factors, it is necessary and expedient to automate and intellectualize decision-making processes (with minimal pilot involvement) and create sufficient conditions for transition to the optionally manned (robotization of the onboard systems) and unmanned (fully robotized) helicopter complexes.

When is such helicopter complexes developing, it is necessary first of all to “robotize” the onboard systems, the means realizing the principles of artificial intelligence which solving the tasks of safe piloting and the effective accomplishment of tasks in different operating conditions and helping to make the best decision on the implementation of the flight task.

At present day, scientific and technical groundwork has been created in scientific centers of the Russian Academy of Sciences and industry-specific research institutes for the implementation of phased robotization (general theory of intelligent systems, theory of situational modeling, theory of expert systems, etc.) of helicopter complexes.

Development of functional intellectualization of the onboard helicopter systems, i.e. the ability of the onboard helicopter systems to function optimally in changing conditions based on adaptive or heuristic algorithms should ensure the following tasks:

– automation of the control of the onboard systems, including search and rescue, offshore, special tasks;

– automatic execution of all stages of helicopter piloting (take-off, flight, landing);

– highly intellectual decision support by the crew;

– guaranteed safety of automatic flight.

The level of "intelligence" of the onboard systems should be achieved by using appropriate algorithms and programs that ensure the effectiveness of the decision support system.

In the near future, it is advisable to develop and implement an intelligent decision support system on the basic helicopter complexes platforms, which form the basis for the “robotization” of the CCD functions of helicopter complexes.

Decision support system functionality should be provided by the use of:

– stronger algorithms; one optimal solution is provided with known input parameters (“hard” algorithm);

– technologies of artificial intelligence using the methods of expert systems, fuzzy logic, neural networks, bionics (“creative” algorithm).

The feasibility of an “intelligent onboard systems” as part of helicopter complexes should be ensured by using a promising methodology for constructing a structural-functional onboard systems based on a top-level network architecture system (or information management system).

The top-level onboard systems should ensure the integration of sub-systems, algorithmic problem solving, the implementation of decision support system based on artificial intelligence technologies and the corresponding software and hardware (super-productive processors, "highly developed" algorithms, high-speed exchange buses of large bandwidth) and their optimal and effective operation.

Intellectualization algorithms of the onboard systems will ensure the expansion of the information field of structural and functional characteristics and an increase in the available computing power of the onboard systems (taking into account the use of the domestic element component base).

The main areas of robotization helicopter complexes (in terms of the implementation of intelligent algorithms) include:

– implementation of effective mathematical methods for modeling processes and operations (based on theoretical and applied mechanisms for numerical modeling of physical processes of systems functioning);

– systematization, collection and analysis of data on the state of the systems (in the process of solving various problems, allowing to obtain patterns and data on the functioning of the systems);

– development of technologies and tools for operational modeling of complex technical systems.

Robotics helicopter complexes are a complex process associated with the introduction of advanced scientific and technical results in the field of modeling and with the achieved level of technology and technological solutions in the field of complex information and control dynamic systems.

Based on the analysis of modern scientific and technical achievements and the level of the production and technological base of developers of onboard systems, step-by-step robotization of helicopter complexes can be implemented in 3 stages involving the gradual «adaptation» of a manned helicopter to a robotic (unmanned) version.

At the same time, helicopter complexes can be considered as an element of the automated control system (within the network-centric model) and as an autonomous unit capable of independently (or as part of a mixed LA group) performing a given amount of tasks (exploration, search and rescue, other special tasks).

In the basis of this approach, using the example of foreign development programs, it is advisable to apply the ideology associated with the evolution (division) of robotic complexes according to the principles:

– «man in the control system» (human-in-the-loop);

– «man over the control system» (human-on-the-loop);

– «man outside the control system» (human-out-of-the-loop).

In the near future – a manned helicopter can be equipped with a system of intelligent decision making (elements) that will provide support to the pilot during the flight.

In this case, helicopters will be equipped with special software and hardware complexes (without interfering with the standard design of the onboard systems).

At this stage, piloting should mainly be carried out in a semi-automatic mode, trainings and training of decision support system crews are performed with the development of automated control algorithms (at the main stages of the helicopter flight), the modernization of the helicopter complexes begins on the basis of a unified Computer-Assisted Action Information System.

In the medium term, the intelligent control system is fully implemented on helicopters. All tasks of the CCD VC should be performed in semi-automatic mode, but under the control of the pilot (with the transition to semi-automatic or manual control mode, if necessary).

In this case, all onboard systems helicopter complexes operate in automatic mode, the interaction of the helicopter complexes with the automatic control system (hereinafter - ACS), ground control points, ensuring accurate exit, and landing are being worked out. Perform special functions of communication with other helicopter complexes and aircraft.

Within the framework of the second stage the following technologies should be implemented:

– automatic control of group use;

– decision support system of intellectual support of crews when performing basic flight, tactical and special tasks.

In the long term, the helicopter complexes should perform piloting and solve problems autonomously without the participation of the pilot, fully in an automated mode under the control of the decision support system (intelligent system), including reconnaissance and special ones in the group of mixed-type aircraft.

Helicopter complexes of the third stage can be optionally manned (or unmanned), having robotic functions, including:

– the ability to «think» and «make independent» decisions (under the control of a person) with significantly better parameters (in terms of speed and accuracy, quality of the decision taken);

– ensuring full safety of the flight in an automated mode, including as part of a group of helicopters and other manned and unmanned aircraft;

– integration into a unified network-centric control system for aircraft and vehicles (an integral part of the interspecific system), ensuring interaction with unmanned flying vehicles (UAVs), ground control points taking into account the analysis of preferred application scenarios, information exchange options and management strategies.

References

1. Zheltov Yu.S. Vizilter Yu.V. Prince V.A. Vygolov O.V. Obstacle detection in vehicle control system. Report at the 3rd International Conference "Digital Information Processing and Emergency Management" Minsk, 2002
2. Makarenko N.S. Military robotic complexes - the current state and development prospects. UDC 007.52.
3. Sizov V.Yu. What combat robots are needed by Russia // Military Review (electronic resource) 03/07/2016 - URL: https: //topwar.ru/91912-kakie-boevye-roboty-nuzhny-rossii.html.
4. Mchenry R. ACTUV ASW Continuous Trail Unmanned Vessel Industry Day - USA, DARPA, 2010.
5. The Navy Unmanned Undersea Vehicle (UUV) Master Plan - USA, Department of the Navy, 2004.
6. Unmanned aerial vehicle ELBIT Hermes 900 (electronic resource). URL: https: //www.ejwiki.org./wiki

V.I. Merkulov, D.I. Milyakov, A.S.Plyashechnik

SYNTHESIS OF PHASED ANTENNA ARRAYS FOR LONG-RANGE MOBILE RADARS BASED ON QUADCOPTERS

JSC "Concern "Vega", Moscow, Russia
from_fn@mail.ru

Abstract

Strengthening the role of unmanned aerial vehicles (UAVs) for various purposes due to their group use allows obtaining a number of advantages in solving a wide class of tasks. At the same time, the new advantages of group use of UAVs are due to: the difficulty of separate observation of the group members and, accordingly, the difficulties of tracking and target distribution; the inability to serve the entire large group with the number of participants exceeding the capacity of the information control system of the opposing side; the increase in the behavioral complexity of the UAV in solving various problems through the use of artificial intelligence; random change of the spatial position of individual UAVs within the group, preventing their detection and selection of virtually all types of information systems.

The noted advantages of the UAVs groups are especially pronounced in the implementation of such a new task as the formation and use of temporary phased antenna arrays (PAA) of large sizes based on multicopter for the implementation of long-range radar systems.

In this regard, the purpose of the report is to present a variant of the algorithm for the formation and functioning of such a PAA. On the example of solving the task of maintaining an air object, the features of the operation of a radar with a PAA based on a group of UAVs are illustrated.

Keywords: phased antenna arrays, mobile radar, quadcopter.

The problem of the formation and functioning of a phased antenna array (PAA) based on unmanned aerial vehicles (UAVs) was solved in two stages.

The first was devoted directly to the removal of all UAVs from the initial points of their ground location to points with coordinates set for each UAV in the air and the formation of a stabilized fixed UAV group in this way.

This stage in the work was presented by the results of modeling an algorithm consisting of the following steps:

1) construction of a reference trajectory for each UAV;
2) calculation of derivatives of the trajectory up to 4th order inclusive;
3) substitution of the derived derivatives into the laws of control of each UAV;
4) recalculation of the status of each UAV.

At the first step, when constructing the UAV reference trajectory, mathematical relations were used for models of its movement in the horizontal plane [1]

$$\begin{cases}\dot{\mathbf{p}}_h = \mathbf{v}_h, \\ \dot{\mathbf{v}}_h = -g\mathbf{A}_\Psi \mathbf{\Theta}_h, \end{cases} \quad . \tag{1}$$
$$\mathbf{p}_h = \begin{bmatrix} p_x \\ p_y \end{bmatrix}, \mathbf{A}_\psi = \mathbf{R}_\psi \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \mathbf{R}_\psi = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix}, \mathbf{\Theta}_h = \begin{bmatrix} \varphi \\ \theta \end{bmatrix},$$

where g is the acceleration of gravity, φ is the roll angle, θ is the pitch angle, ψ is the yaw angle, in the vertical plane [1]

$$\begin{cases}\dot{p}_z = v_z; \\ \dot{v}_z = g - \dfrac{f}{m}, \end{cases} \tag{2}$$

where f is the traction force of the engines, m is the mass of the UAV, and angular motion [1]

$$\begin{cases}\dot{\mathbf{\Theta}} = \boldsymbol{\omega}; \\ \mathbf{J}\dot{\boldsymbol{\omega}} = \boldsymbol{\tau}, \end{cases} \tag{3}$$

where $\mathbf{\Theta}$ is the vector of Euler angles, $\boldsymbol{\omega}$ is the vector of angular velocity of rotation of the UAV, $\mathbf{J}$ is the matrix of inertia moments of the UAV, $\boldsymbol{\tau}$ is the total torque of the UAV screws.

The second step associated with the calculation of derivatives can be represented by the following relationships [2]:

$$
\begin{aligned}
&\mathbf{R}_d = \begin{bmatrix}\mathbf{b}_{1d} & \mathbf{b}_{3d}\times\mathbf{b}_{1d} & \mathbf{b}_{3d}\end{bmatrix} \in \mathrm{SO}(3);\\
&\mathbf{b}_{1d} \equiv \Psi_d;\, \mathbf{b}_{3d} = \mathbf{R}_d\mathbf{e}_3 = -\frac{-k_x\mathbf{e}_x - k_v\mathbf{e}_v - mg\mathbf{e}_3 + m\ddot{\mathbf{x}}_d}{\left\|-k_x\mathbf{e}_x - k_v\mathbf{e}_v - mg\mathbf{e}_3 + m\ddot{\mathbf{x}}_d\right\|};\\
&\mathbf{\Omega} \equiv \boldsymbol{\omega};\, \mathbf{e}_x = \mathbf{x} - \mathbf{x}_d;\, \mathbf{e}_v = \mathbf{v} - \mathbf{v}_d;\, \mathbf{e}_R = \frac{1}{2}\left(\mathbf{R}_d^T\mathbf{R} - \mathbf{R}^T\mathbf{R}_d\right)^{\vee};\ \mathbf{e}_\Omega = \Omega - \mathbf{R}^T\mathbf{R}_d\Omega_d; \qquad (4)\\
&\dot{\mathbf{R}} = \mathbf{R}\cdot\hat{\mathbf{\Omega}} \Rightarrow \dot{\mathbf{R}}_d = \mathbf{R}_d\hat{\mathbf{\Omega}}_d \rightarrow \hat{\mathbf{\Omega}}_d = \mathbf{R}_d^T\dot{\mathbf{R}}_d \rightarrow \dot{\hat{\mathbf{\Omega}}}_d = \dot{\mathbf{R}}_d^T\dot{\mathbf{R}}_d + \mathbf{R}_d^T\ddot{\mathbf{R}}_d \rightarrow \begin{bmatrix}\ddot{\Psi}_d = \dots\\ \ddddot{x}_d = \dots\end{bmatrix},\\
&\hat{\mathbf{\Omega}}_d = \mathbf{R}_d^T\dot{\mathbf{R}}_d \rightarrow *^{\vee} \rightarrow \mathbf{\Omega}_d = \left(\mathbf{R}_d^T\dot{\mathbf{R}}_d\right)^{\vee},
\end{aligned}
$$

where $\mathbf{R}$ and $\mathbf{R}_d$ are the current and desired UAV orientation matrix in space, respectively $\mathrm{SO(3)}$ is group of orthogonal rotation matrices; ($\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$) is triple of vectors corresponding to the direction cosines associated with the UAV, in this case $\mathbf{b}_{1d}$ is determined by the desired UAV orientation around its axis OZ when moving along the trajectory as the law of change in time of rotation of the axis OX around the axis OZ of the UAV; $\mathbf{b}_{3d}\times\mathbf{b}_{1d}$ is determined by the vector product of the vectors $\mathbf{b}_{3d}$ and $\mathbf{b}_{1d}$; $\mathbf{e}_3$ is the unit vector of the OZ axis; k_x, k_v are some positive constants; $\mathbf{e}_x$, $\mathbf{e}_v$, $\mathbf{e}_R$, $\mathbf{e}_\Omega$ are UAV following errors along the desired trajectory; $\mathbf{\Omega}$ and $\mathbf{\Omega}_d$ are the vector and its desired value of the angular velocity of rotation in the coordinate system associated with the UAV; $*^{\vee}$ is the sign of the operation of converting the matrix into a vector.

As a result of performing the second step from (4), we obtain derivatives $\ddot{\Psi}_d$ and $\ddddot{x}$ of following along the desired trajectory.

At the third step, the propeller thrust $\mathbf{f}$ and torque $\boldsymbol{\tau}$ as UAV control signals were generated according to the laws:

$$
\begin{aligned}
&\begin{bmatrix} f\\ \tau_x\\ \tau_y\\ \tau_z\end{bmatrix} = \begin{bmatrix} c_T & c_T & c_T & c_T\\ 0 & -dc_T & 0 & dc_T\\ dc_T & 0 & -dc_T & 0\\ c_M & -c_M & c_M & -c_M\end{bmatrix}\begin{bmatrix}\omega_1^2\\ \omega_2^2\\ \omega_3^2\\ \omega_4^2\end{bmatrix};\\
&\mathbf{f} = -\left(-k_x\mathbf{e}_x - k_v\mathbf{e}_v - mg\mathbf{e}_3 + m\ddot{\mathbf{x}}_d\right)\mathbf{R}\mathbf{e}_3; \qquad (5)\\
&\boldsymbol{\tau} = -k_R\mathbf{e}_R - k_\Omega\mathbf{e}_\Omega + \mathbf{\Omega}\times\mathbf{J}\mathbf{\Omega} - \mathbf{J}\left(\hat{\mathbf{\Omega}}\mathbf{R}^T\mathbf{R}_d\mathbf{\Omega}_d - \mathbf{R}^T\mathbf{R}_d\dot{\mathbf{\Omega}}_d\right);\\
&\boldsymbol{\tau} = \begin{bmatrix}\tau_x\\ \tau_y\\ \tau_z\end{bmatrix}.
\end{aligned}
$$

Here $\mathbf{f}$ and f is the thrust vector directed along the Z axis of the UAV, and its module; $\boldsymbol{\tau}$ is torque of UAV screws with projections (τ_x, τ_y, τ_z) on the coordinate axis; c_T and c_M are dimensionless coefficients corresponding to the thrust and torque of the screws, respectively; d is the distance between the center of the UAV and the center of its *i*-th engine; ω_i is angular rotation speed of the screws of the *i*-th UAV engine; k_R, k_Ω are some positive constants.

For the UAV configuration considered in this work in the form of an X-shaped quadrocopter with perpendicular guides, the control signals in (5) can be converted into the following form [1]:

$$\begin{aligned}
& f = c_T\left(\omega_1^2+\omega_2^2+\omega_3^2+\omega_4^2\right) \\
& \tau_x = dc_T\left(\frac{\sqrt{2}}{2}\omega_1^2-\frac{\sqrt{2}}{2}\omega_2^2-\frac{\sqrt{2}}{2}\omega_3^2+\frac{\sqrt{2}}{2}\omega_4^2\right); \\
& \tau_y = dc_T\left(\frac{\sqrt{2}}{2}\omega_1^2+\frac{\sqrt{2}}{2}\omega_2^2-\frac{\sqrt{2}}{2}\omega_3^2-\frac{\sqrt{2}}{2}\omega_4^2\right); \\
& \tau_z = c_M\left(\omega_1^2-\omega_2^2-\omega_3^2-\omega_4^2\right).
\end{aligned} \tag{5.1}$$

At the fourth step, the UAV state was calculated according to the model [1],

$$\begin{cases}
{}^e\dot{\mathbf{p}} = {}^e\mathbf{v}; \\
{}^e\dot{\mathbf{v}} = g\mathbf{e}_3 - \dfrac{f}{m}\mathbf{R}\mathbf{e}_3; \\
\dot{\mathbf{R}} = \mathbf{R}\left[{}^b\boldsymbol{\omega}\right]_\times; \\
J\cdot{}^b\dot{\boldsymbol{\omega}} = -{}^b\boldsymbol{\omega}\times\left(J\cdot{}^b\boldsymbol{\omega}\right)+\mathbf{G}_a+\boldsymbol{\tau}.
\end{cases} . \tag{6}$$

$$\left[{}^b\boldsymbol{\omega}\right]_\times = \begin{bmatrix} 0 & -{}^b\omega_z & {}^b\omega_y \\ {}^b\omega_z & 0 & -{}^b\omega_x \\ -{}^b\omega_y & {}^b\omega_x & 0 \end{bmatrix}. \tag{7}$$

In (6) ${}^e*$ means the indication of * in the Earth's coordinate system, and ${}^b*$ – in the onboard coordinate system associated with the UAV; $\mathbf{p}$ and $\dot{\mathbf{p}}$ are vectors of coordinates and rates of change for the center of gravity of the UAV, respectively; the operation $[*]_\times$ in (7) means constructing a matrix by a vector *

The UAV trajectory was drawn according to the model [1]

$$\begin{cases}
{}^e\dot{\mathbf{p}} = {}^e\mathbf{v}; \\
{}^e\dot{\mathbf{v}} = g\mathbf{e}_3 - \dfrac{f}{m}\mathbf{R}\mathbf{e}_3; \\
\dot{\boldsymbol{\theta}} = \mathbf{W}\cdot{}^b\boldsymbol{\omega}; \\
J\cdot{}^b\dot{\boldsymbol{\omega}} = -{}^b\boldsymbol{\omega}\times\left(J\cdot{}^b\boldsymbol{\omega}\right)+\mathbf{G}_a+\boldsymbol{\tau},
\end{cases} . \tag{8}$$

whose kinematics is associated with the representation of the rotation matrix $\mathbf{R}$ from (6) in the form of a matrix $\mathbf{W}$ of Euler angles.

Thus, we obtained the results of modeling the output of a group of 9 UAVs at given points in space, illustrated in Figure 1.

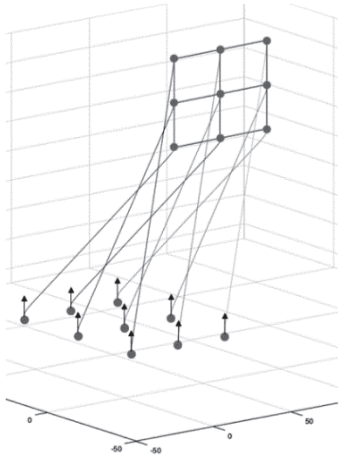


Figure 1 – An example of the formation of phased array UAV

The second stage of solving the problem was devoted to the formation of the radiation pattern of obtained PAA based on the UAV and the tracking of the air object (AO) during its maneuverable movement.

The research results of this step are illustrated in Figures 2-4.

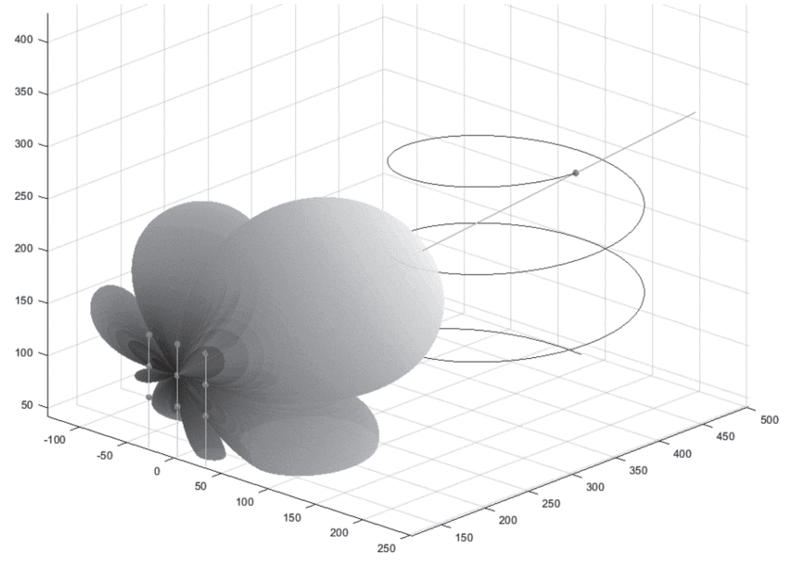
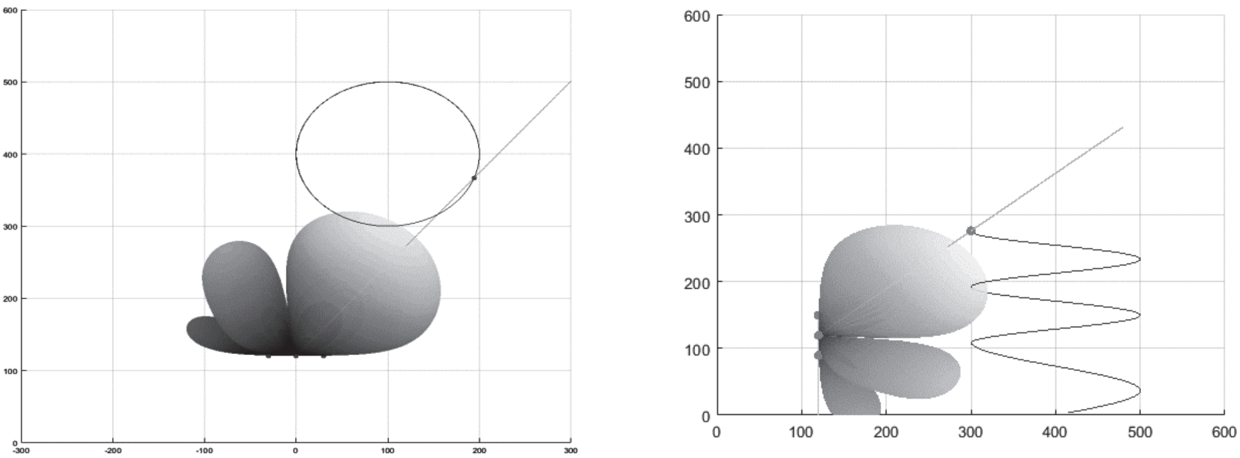


Figure 2 – Maintenance of maneuvering AO using PAA based on the UAV group



a) b)

Figure 3 – Maintenance of maneuvering AO using PAA based on the UAV group (a – top view, b – side view)

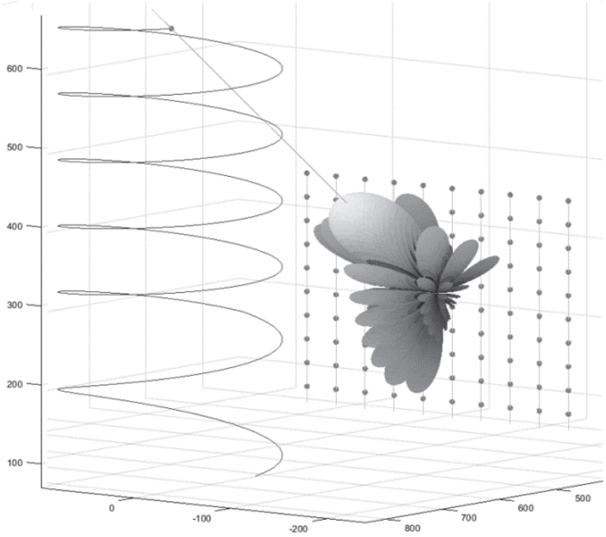


Figure 4 – The appearance of the radiation pattern of
the PAA when changing its dimension

In the work, a study was made of the antenna array (see Fig. 5), consisting of 50 × 50 carriers (with an emitter and a transceiver module (ETM)).

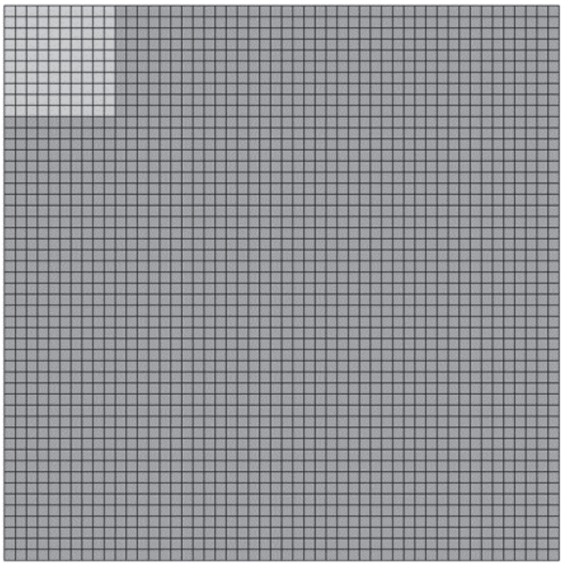


Figure 5 – An example of a formed PAA

A cell (10 × 10) corresponds to a carrier (with an emitter and an ETM) operating both in transmission and reception. The remaining cells correspond to the carrier (with emitter and ETM), which works only on reception. The transmitting area can be moved along the aperture for uniform consumption of energy carriers. The size of the transmission region can be increased to obtain a given beam width and level of the side lobes.

One of the options for the emitter of the array is a log-periodic antenna (see Fig. 6).

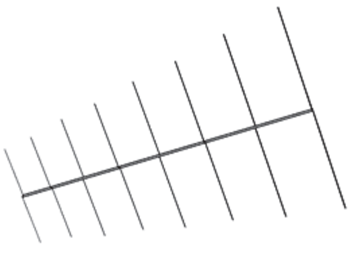


Figure 6 – Log-periodic antenna

Its main advantages are simplicity of design, low weight, high gain and relatively low level of back radiation without a metal screen. The radiation pattern of such an antenna is shown in Fig. 7.

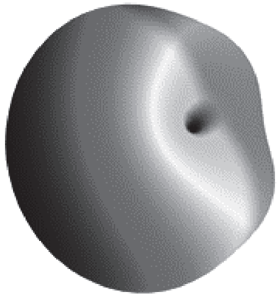


Figure 7 – The radiation pattern of a log-periodic antenna

Also, studies were carried out on reconfiguration of the PAA in the case of nodes diverging from 0.25λ to 0.5λ and from 0.25λ to 0.9λ (here λ is the wavelength of the emitted signal), the results of which are illustrated in Figures 8 and 9, respectively.

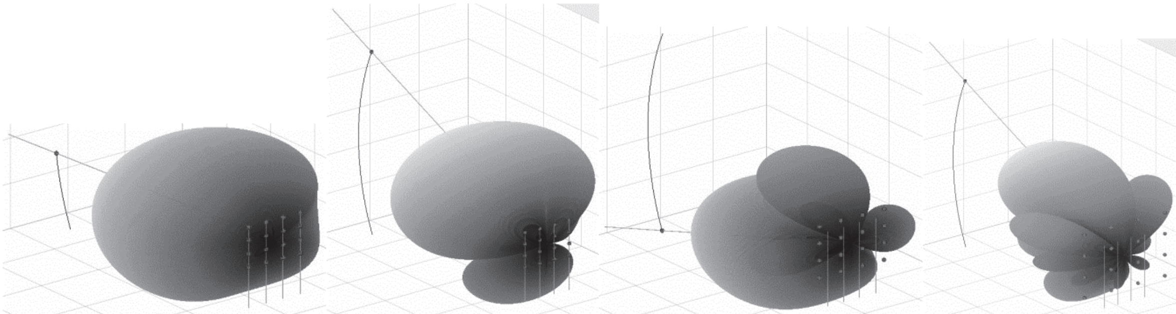


Figure 8 – Reconfiguration of the PAA in case
of divergence of its nodes from 0.25 λ to 0.5 λ

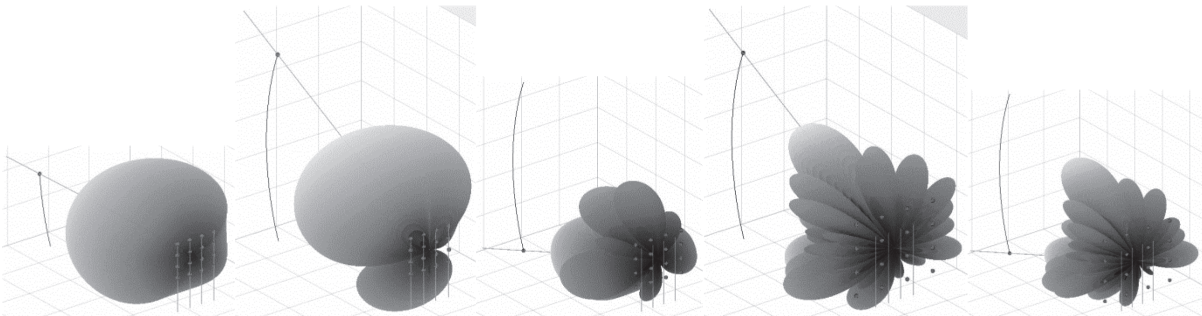


Figure 9 – Reconfiguration of the PAA in case
of divergence of its nodes from 0.25λ to 0.9λ

The research results confirmed the emergence of standard difficulties for the processing of multi-beam patterns for antenna systems. It should be noted that there are various ways to overcome these difficulties that determine the various properties of the resulting systems.

References

1. Quan Quan. Introduction to Multicopter Design and Control. Springer. 2017. DOI 10.1007/978-981-10-3382-7.
2. Taeyoung Lee, Melvin Leok, and N. Harris McClamroch. Geometric Tracking Control of a Quadrotor UAV on SE(3). arXiv:1003.2005v1 [math.OC] 10 Mar 2010.

S.S. Tataurshchikov

PERSPECTIVE DEVELOPMENTS OF PHOTODETECTORS BY JSC "NRI "ELECTRON"

JSC "NRI "Electron", St. Petersburg, Russia

Abstract

Future development of photoelectronic devices of JSC "NRI "Electron", produced on their own production and technological base (vacuum and solid-state), are discussed. The devices feature with a high level of basic characteristics providing leadership in the market. It is reported about new directions of development of new solid-state and hybrid photoelectronic devices, which have the required potential in the creation of modern opto-electronic devices and systems.

Keywords: CCD, photoelectronic devices for IR, VIS and UV, low light level hybrid TV device photodetectors, EMCCD on chip, wide-format photodetector, CMOS.

The company develops and manufactures photoelectronic devices and combined systems based on them for various spectral ranges (from UV to near IR) and different fields of applications. Company own technological base (vacuum and solid-state) creates a unique opportunity to develop hybrid photodetectors, combining the advantages of vacuum and solid-state technologies, as well as the ability to develop fundamentally new technologies on an existing base, such as radiation resistant, or with back-side illumination, which ensures market leadership in photoelectronic devices [1].

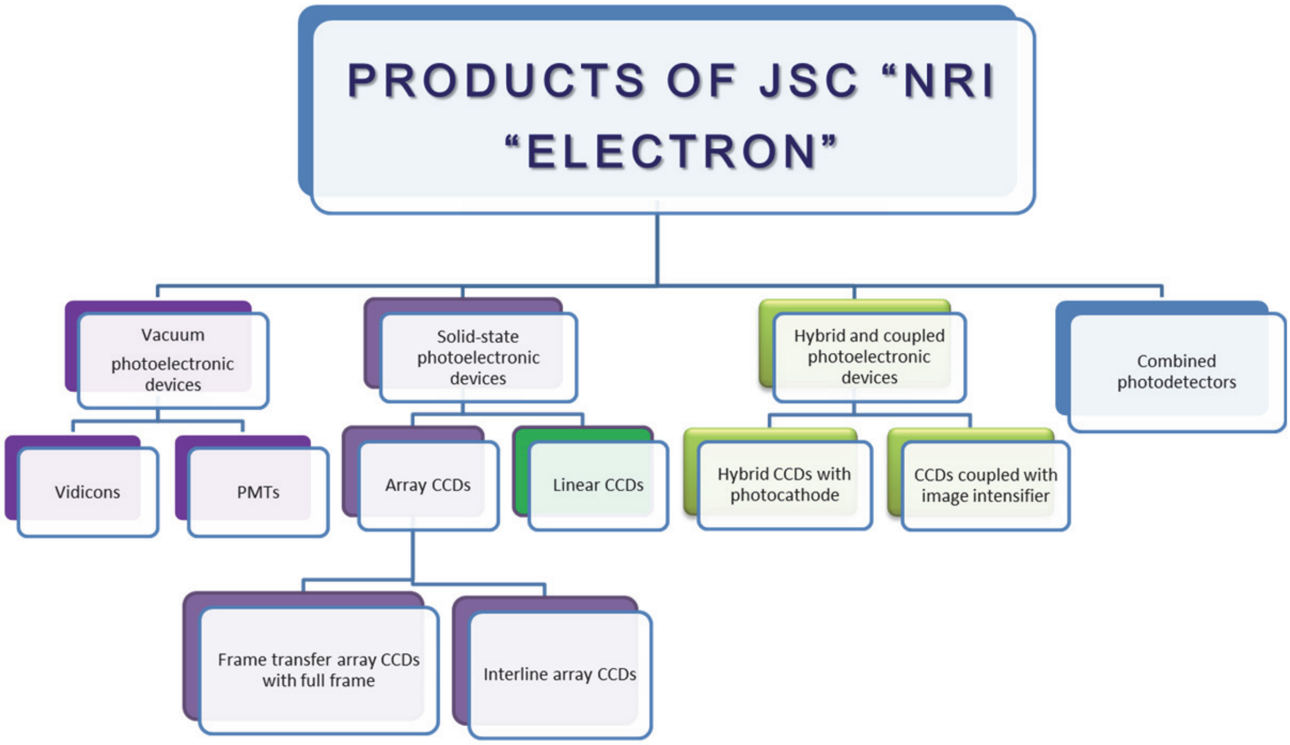


Figure 1 – Nomenclature

Based on its achievements, JSC "NRI "Electron" plans in the next 5 years to conduct the following new developments of photoelectronic devices:

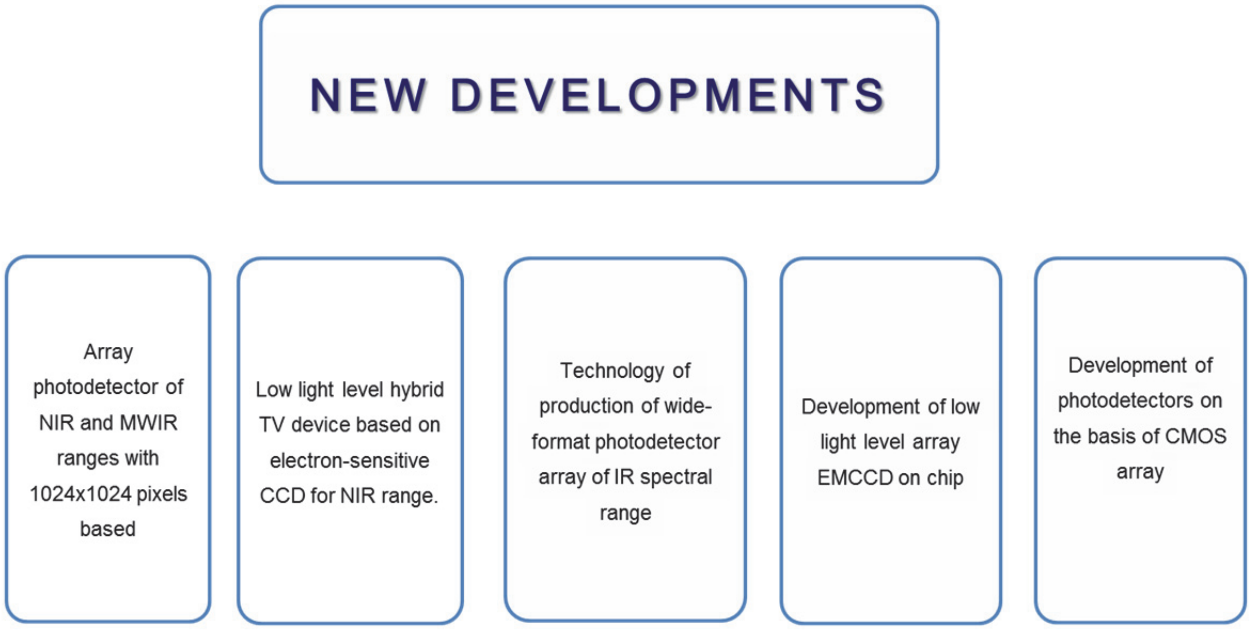


Figure 2 – New developments

1.Array photodetector of near- and mid- IR ranges with 1024×1024 pixels based on photosensitive Schottky diodes

Currently IR array photodetectors for ground and airborne optoelectronic devices are created on the base of hybrid photodetectors, have a small number of elements and high levels on the cooling system, as well as complex digital signal processing devices due to the instability of sensitivity parameters of photodetectors. The available wide-format IR photodetectors with electron-beam reading system are limited in the frame rate of the output information. The application of the proposed for the development MFPU will significantly improve the information and operational characteristics, reduce the thermal load of cooling devices, reduce the mass and energy consumption of ground, air and space-based equipment.

High sensitivity and resolution, low level of structural interference and the possibility of its full compensation with interframe substraction due to a hard digital raster in the developed MFPU will improve the spatial resolution and sensitivity of ground and airborne OES.

The use of the proposed MFPU will also simplify the calibration of the OES during operation (due to the long-term stability of the planned photodetectors), and, as a consequence, reduce the required computing resources of the OES. This will reduce the cost of OES creation, as well as its operation. In addition, the use of a monolithic silicon photodetector based of Schottky diodes will ensure its stable operation for about 30,000 hours, which significantly exceeds the lifetime of photodetectors based on MKT.

2. Low light level hybrid TV device based on electron-sensitive CCD for near IR range

Currently, television systems used to solve military and civilian tasks are being developed on the basis of image intensifiers of 2+ and 3 generation. The sensitivity of such intensifiers is close to theoretically possible, which will not allow a substantial increase in the further range of action of television detection and recognition systems. The installation of the CCD array inside the vacuum volume of the image intensifier dramatically improves the signal-to-noise ratio of the photoelectronic device, since the array unlike the microchannel plate, completely accepts the signal flow of electrons. Such devices are characterized by very high speed.

Development of hybrid television module with low light level solar-blind photocathode for the UV range will provide both increased sensitivity and improved “solar-blindness”.

A significant (several times) increase in the range of action is possible due to the transition to the infrared diapason of the spectrum up to 1.6 µm. At night, in this diapason, the irradiance of objects is in 10-20 times higher than in the range of 0.5-0.9 µm, the contrast of objects increases, and atmospheric losses decrease. This allows to increase the range of action of passive systems in 2-2.5 times. In addition, the use of such system with laser illumination will provide observation of the image with high resolution, as well as the detection and identification of the target at a distance of up to 20 km. This is in 3-3.5 times more than the best surveying systems of 3-5 and 10-12 micron ranges.

A device containing a TE-photocathode and electron-sensitive CCD array provides 10 times more sensitivity than a solid-state analog.

3. Technology of production of wide-format photodetector array of IR spectral range

At present, photodetectors based on epitaxial structures of A3B5 are used to obtain a television image in the infrared region of the spectrum. Such photodetectors have high sensitivity and resolution, but their disadvantage is that in order to provide the necessary photoelectric characteristics, photodetectors require cooling to temperatures close to 80 K. This leads to an increase in the weight and size characteristics of the equipment, reduces its reliability. Reduction of the weight and size characteristics of the equipment and increase of its reliability can be achieved by moving from a cryogenic level of cooling to a thermoelectric level. In this regard, the development of photodetectors, characterized by a working temperature sufficient for transfer to a thermoelectric cooling level (150-170K) seems relevant.

4. Development of low light level array EMCCD on chip

The aim of the work is creation of a large format array CCD with 1024×1024 pixels, improved sensitivity to detection of single photon signals and threshold exposure (1-2)×10-6 lx due to electron multiplication in CCD additional output register and increase of signal charge up to 1000 times. Such photodetectors, comparable in sensitivity to image intensifiers of 2+ generation, are the most important basis for the construction of promising high-precision optoelectronic systems of target detection in low light level conditions.

5. Development of photodetectors on the basis of CMOS array

The aim of the work is to create a large format CMOS array with 1024×1024 elements of higher sensitivity 3,2×E8e-/((W/m2)·s), expand significantly the range of photodetector sensitivity (0.2-1.0 μm) and increase the maximum of quantum efficiency up to 90% at back-side illumination, that is the most important basis for building promising on-board high-precision optoelectronic target detection systems.

The considered CMOS array will ensure the registration of weak light signals at the level of single photons, and also optimize the processes of the array operation control and output digital signal processing, which will improve in several times the basic technical characteristics of the onboard optoelectronic complexes – dimensional-mass in 1.5-2 times, accuracy and dynamic in 3-5 times.This situation is caused by the fact that in modern high-precision optoelectronic complexes a higher frequency of information output up to 100 MHz and more is required, which is an unattainable value for currently used photodetectors based on CCD, and the use of imported CMOS analogs is associated with sanctions restrictions.

Airborne systems and complexes of target detection, tracking and aiming, which are in service with the Ministry of Defense of the Russian Federation, are oriented by modern optoelectronic devices. There is need to improve the accuracy of determining the coordinates, as by increasing the photosensitive array of the photodetector up to 1024×1024 elements, so by improving the sensitivity and expanding the spectral sensitivity range at a sufficiently high information processing speed. Therefore, the development of optoelectronic devices for onboard equipment of the specified purpose for the promising systems is limited by the capabilities of the existing photodetectors based on CCDs.

References

1. Stepanov R.M. *Televisionnye photoelectronnye pribory* [TV photoelectronic devices]. St.-Petersburg, SPbETU “LETI” Publ., 2014.

J.S. Bodrova, G.F. Karabadzhak, K.G. Raykunov

SPACE ROBOTICS MOBILE VEHICLE PLATFORMS, THEIR PRIORITY TASKS AND POTENTIAL USAGE SCENARIOS TO SUPPORT RUSSIAN MANNED MOON EXPLORATION PROGRAM

State Space Corporation Roscosmos, Central Research Institute of Machine Building, Korolev, Russia
RaykunovKG@tsniimash.ru

Abstract

In accordance with the "Fundamentals of the State Policy of the Russian Federation in the field of space activities for the period till 2030 and beyond" the key development thrust of Russian manned cosmonautics is being towards to the Moon exploration. Two key Russian Moon Exploration Program events shall be: the landing of Russian cosmonauts on the Moon in 2030 and the start of manned flights to the Moon on a regular basis from 2032. Up-to-date, man activities at LEO are clearly considered to be of applicable nature and this activity shifts from "exploration" stage towards "utilization" one. Spacefaring nations are expecting to get notable practical results, including commercial benefits and ability to export "space" technologies to domestic market, as well as to reach a new "continent" - the Moon.

It is expected that the profound exploration of the lunar territories will become the main task of the world cosmonautics of the 21st century. The restored interest in the exploration of the Moon is largely caused by the discovery of the unique regions in the immediate vicinity of the lunar poles that have highly favorable conditions for the future deployment of lunar bases. There is hardly ever any sunset in the given regions, as well as in the very vicinity thereof there is high probability of sizable lunar ice deposits to be located.

If any favorable outcomes of these regions investigation are obtained the competition to obtain access to the regions under consideration and their resources will unfold, particularly taking into account new development trends in international space law. In this very case space robotics gets a new rational field of application, as well as its new development vectors are emerging to address the tasks on effective Moon exploration and manned missions support.

In this paper a methodological approach to form a basic set of scenarios and determine the required range of spacecraft for the implementation of a comprehensive long-term program of research and exploration of the Moon is considered. Baseline scenarios are the input data for the requirements definition to spacecraft, including space robotics. An evaluation of the relevance of various baseline platforms of automatic spacecraft to accomplish the suggested scientific tasks to support manned missions has been conducted based on the analysis of the Russian Academy of Sciences priority tasks in the domain of lunar studies. Requirements for the proposed platforms are set within the scope of the baseline scenarios.

Keywords: the Moon, manned Moon exploration, lunar missions, space robotics, Russian Moon Exploration Program, rovers, lunar rovers, service lunar rover.

1. Approaches to program development on manned research and exploration of the Moon and its main stages

The Moon has suitable conditions and is an accessible testing ground for:

– development and trying out of new technologies;
– outer space exploration;
– scientific research in various science areas;
– development of not only cosmonautics but also industry and international cooperation.

Manned Moon exploration can fall into the three following main stages:

The First Exploration Stage – exploration of the Moon by automatic spacecraft (SC), a number of manned flights to the Moon orbit and its surface, development of Space Transportation System (STS) to maintain people and cargo deliveries, trying out of critical technologies. On this stage the possibility and feasibility of lunar base development either on the Moon surface or on its orbit is investigated.

The Second Stage – construction of manned lunar base of entry-level configuration and development of infrastructure ensuring manufacturing of life support system components and resources to make continuous presence of people on the Moon possible, as well as construction of scientific, test, and production facilities. On this stage the evaluation of lunar base expansion feasibility is investigated.

The Third Stage – expansion of the lunar base and development of a closed-loop life support system that uses lunar resources to maintain its operation; setup of propellant components production and use of this propellant for the transport system; extraction of oxygen, water, metals, construction materials and other elements from the lunar resources.

The Moon research and exploration program shall aim at delivery of a long-term integrated and interconnected scientific and technological program rather than at sole delivery of a demonstration program.

It is obvious that the up-to-date spacecraft being under development and state of their readiness is clearly considered to be insufficient for fulfillment of the objectives set.

To develop the long-term integrated program a list of baseline scenarios aimed to achieve strategic goals is being developed. These scenarios are to be implemented at the initial stage with the use of existing spacecraft and spacecraft under current development taking into account their development time and state of readiness. The scenarios under consideration are the input data for defining the requirements to conceptual designs of spacecraft, including space robotics.

The mission scenarios shall include a more extended list of spacecraft and specify requirements to their functionality and availability dates based on the needs of successful accomplishment of both scientific and technological tasks. Scientific and applied purposes and tasks shall be distributed on time and Moon exploration stages basis in the course of scenarios development process, having regard to launch vehicle capabilities, as well as to capabilities to developed required spacecraft. Spacecraft shall be subdivided into basic platforms both based on dimensional characteristics (light-, heavy- and super heavy classes) to correspond to LV dimensions and on spacecraft classification (landing platforms, rovers, landing and launching platforms). The proposed approach will allow for reduce in the dimension of the task of optimization of required range of spacecraft, functional and technical requirements for them and their development sequence, and focus on optimizing the range of “baseline spacecraft” (BS) and “baseline mission scenarios” (BMS) to ensure the implementation of the manned cosmonautics development strategy adopted.

Development of off-the-shelf basic platforms of different dimensions could allow not only to decrease expenditures on development of new spacecraft due to the continuity of technologies and technical solutions, but also to adapt these platforms flexibly to solve various specific tasks and to develop gradually new technological solutions during the missions to be performed.

Basic architecture of spacecraft and mission scenarios is used for specification of functional and technical requirements to advanced manned systems and complexes to develop potential pre-design concepts of spacecraft, to define the list of the required technologies, as well as to evaluate the feasibility of the manned spacecraft development concept adopted.

Such approach (see Fig.1) ensures generation of interconnected activities (projects) list; their proper performance in an appropriate timeline will ensure implementation of the adopted development strategy.

A significant practical effect of the given approach is a direct association of the technologies list to the mission scenarios and the spacecraft involved.

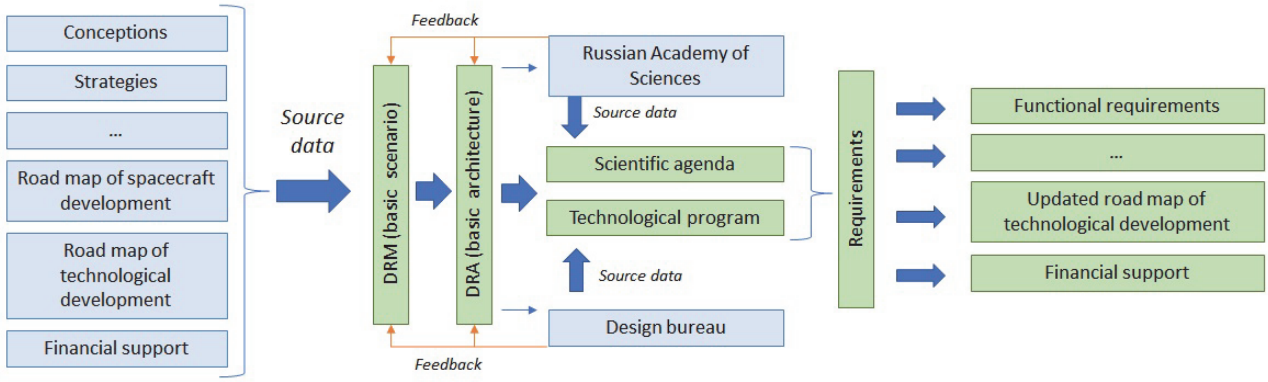


Figure 1 – Development algorithm for Moon research and exploration program, list of spacecraft and requirements to them

2. Priority robotic spacecraft to use in the Moon surface exploration program

To develop a scientific program on the Moon research and exploration the Council of the Russian Academy of Sciences for space has drafted proposals to be included in the “Program on the Moon Research and Exploration for the period till 2030 and beyond” among which the following list of high-priority research areas can be emphasized:

– selection of location and operation principles for the lunar base (water and volatiles deposits exploration, exploration of other minerals, in-situ technologies trying out, regolith layer depth and cavities sizes determination, hazards analysis, biomedical research to be performed on the Moon, lunar lava tubes exploration, etc.)

– clarification of the Moon origin and evolution data (regolith analysis down to 30-meter depth, seismology, gravity measurements, thermal monitoring, research of the Moon using magnetic methods, etc.)

– search for interstellar substance (of comet origin) on the Moon (interstellar and interplanetary dust research, comets impact locations research – search for water, minerals, compounds, etc.)

– research of the Sun and the Earth (the solar corona and solar wind research, Moon-based Earth observation, etc.)

– outer space research (radio astronomic surveys performed from the far side of the Moon, cosmic rays spectrometric measurements, etc.)

For a number of experiments/projects/tasks suggested by the Council of the Russian Academy of Sciences for space, deployment of scientific equipment on the Moon surface will allow for overcoming a range of significant constraints, inherent to research performed on the Earth and orbiters, and obtain completely new results not available to obtain in the frame of the on-going projects. However, taking into account complexity of not only installation, but also setting of the scientific equipment, required for the implementation of the given projects, cooperative operation of both robotic spacecraft and cosmonauts is required. To perform a number of operations requiring high level of accuracy and related to fast-evolving consequences that may occur due to the operation error or inaccuracy, the cosmonaut may require direct control of the robotic spacecraft ensuring minimal time lags. It is expected that within the manned flights there will be an opportunity to service, repair and replace individual units of the automatic scientific stations on the lunar surface to extend their operation lifetime. One of the main tasks can also be the repair and retrofitting of lunar research rovers.

As was stated above, the initial stage of the Moon research and exploration is characterized by the requirement to solve a wide range of interrelated fundamental and applied problems. Consequently, the main scope of work of the first manned flights will include implementation of a scientific research and prospecting and exploration program. Taking into account limited resources required for the implementation of manned flights, cosmonauts' professional activity to be performed on the Moon shall focus on the solution of such tasks that cannot be performed without a man in-the-loop.

Based on the analysis of Council of the Russian Academy of Sciences for space proposals to be included in the "Program on the Moon Research and Exploration for the period till 2030 and beyond", an evaluation of the relevance of various baseline platforms of automatic spacecraft to accomplish the suggested scientific tasks has been conducted. The evaluation outcomes are presented in Figure 2. In the framework of the given proposals 17 lines of research have been distinguished (some of the experiments have been combined in one group). The column "Portion of scientific tasks to be solved" represents percentage of the scientific tasks that can be solved using the given spacecraft to the total number of tasks. In-between parenthesis, in the "Portion of scientific tasks to be solved" column the values of this parameter are specified taking into account the requirement to deliver lunar rovers of the corresponding dimensions to the surface of the Moon.

Based on the information in Figure 2, we can suggest that the most highly-demanded mobile vehicle platform to perform the tasks under consideration is heavy-class lunar rover platform (mass – 1-2 t) that can be used to perform 70.5% of the most relevant Moon research and exploration program scientific tasks. The light-class lunar rover platform with mass of ~ 500 kg is to be used for performance of 47% of the tasks given, while its lighter modifications with mass of ~ 300 kg can ensure performance of only 23.5% of the tasks. Significant decrease in percentage of the tasks solved with the decrease in the mass of the light-class lunar rover is caused by a significant decrease in the designed capacity of the platform payload. For the experiments requiring installation of small-scale scientific equipment as mobile platforms payload, based on the weight-and-dimensional characteristics analysis results it was demonstrated that to perform most of the experiments under consideration light-class lunar rover would be enough to use. However implementation of a heavy-class lunar rover can ensure successful performance of combination of tasks to be performed by several light-class lunar rovers within a single mission, thus this very rover is considered to be more operationally efficient.

Medium-class Reusable Automatic Landing and Launching Platform (MBVPK) can be used to perform 52.9% of the scientific tasks under consideration provided it would have a manipulating device with interchangeable equipment and drilling capability.

Baseline platform	Class	Payload mass	Scientific tasks 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Portion of scientific tasks to be solved, %
Landing platform	Light	300-500 kg	OR	OR	OR	OR	OR							AND	OR		AND			53,0% (58,8%)
	Heavy	1-2 t								OR					OR			AND		17,6% (82,3%)
	Super-heavy	6 t													OR				AND	11,7% (23,5%)
Reusable Automatic Landing and Launching Platform (MBVPK)	Light	~ 20 kg													OR					5,8%
	Medium	100 kg; On lunar orbit: 50 kg	OR	OR	OR	OR	OR	AND		AND	AND		AND		OR					52,9%
Lunar rover	Lunar rover for lava tubes	Up to 40 kg						AND												5,8%
	Light, m ~ 300 ± 50 kg	50-60 kg			OR						OR				OR		OR			23,5%
	Light, m ~ 500 kg	50 - 100 kg	OR	OR	OR	OR	OR				OR				OR		OR			47,0%
	Heavy	200 kg	OR	OR	OR	OR	OR	AND			OR		AND	AND	OR		OR			70,5%
	Super-heavy	Geologist: SE ~5 t; MLC: ~500 kg								OR					OR					11,7%
	Super-heavy	Construction rover: SE ~300 kg ; MLC: ~ 4 t													OR			AND	AND	17,6%
Orbital platform	-						AND								AND					17,6%

– Scientific task can be accomplished with the given spacecraft.
– Scientific task can be partly accomplished with the given spacecraft.
– Landing platform is used for the delivery of other spacecraft.
AND – Given spacecraft is necessary to accomplish the scientific task.
OR – Scientific task can be accomplished with either of the given spacecraft.
SE – Scientific equipment.
MLC – Manipulator lifting capacity.

Figure 2 – Evaluation of relevance of various baseline platforms of automatic spacecraft required to accomplish scientific tasks from among Council of the Russian Academy of Sciences for space proposals to be included in the "Program on the Moon Research and Exploration for the period till 2030 and beyond"

3. Scenarios with scientific heavy-class rover platform to support manned missions to the surface of the Moon

After the completion of a number of systems and subsystems trying outs onboard the automatic spacecraft, as well as conduction of the Moon surface remote sensing, exploration of the Moon surface in near-polar area including cryogenic drilling and delivery of lunar regolith samples to the Earth and specification of the landing site for the first manned missions using Luna 25 – Luna 28 spacecraft, the delivery of a heavy-class lunar rover is reasonable to be performed in 2029 to conduct deeper geological surveys and a unified heavy-class landing platform (Luna 29 spacecraft) trying out. At the initial stage, the spacecraft can be used to perform geological and fundamental surveys, collect a significant number of lunar regolith samples and drilling samples on its way and deliver them to the first manned mission landing site. The results obtained from scientific heavy-class lunar rover can help to specify the cosmonauts' potential landing site and an area for the future deployment of the visited base on the Moon surface.

After the first manned mission cosmonauts meet the Luna 29 scientific heavy-class lunar rover it can serve cosmonauts as a mobile scientific laboratory and accompany them within the expedition period, transport collected samples, supply power to the scientific equipment and transport Automatic Scientific Stations (ASS) from the Lunar launch and landing complex (LVPK) or MBVPK to the location of installation to be performed by the cosmonauts or lunar rover manipulating device, perform in-situ analysis of regolith composition and physical and chemical characteristics (depending on the selected payload), drill the moon regolith upon the cosmonaut's command, perform express analysis of water and hydrogen presence in the regolith and perform other functions. The lunar rovers payload allows for determination of the principal geological composition of the area under investigation using the contact investigation methods that could not be replaced by noncontact ones. Based on the results obtained, evaluation of mineral reserves and their mode of occurrence can be performed to develop advanced mining technologies.

An option of scenario for lunar rover and MBVPK interaction is presented in Figure 3.

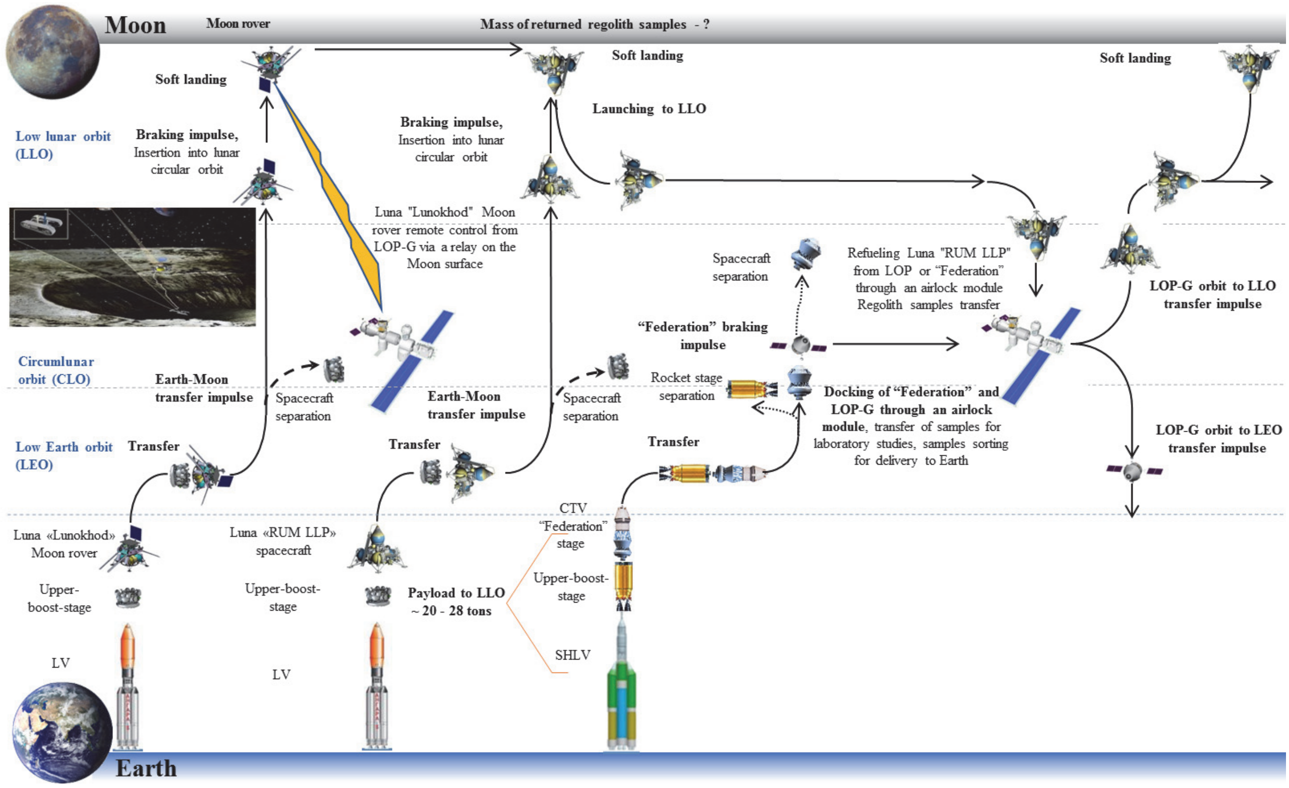


Figure 3 – Scenario for lunar rover and MBVPK interaction

Scientific heavy-class lunar rover shall perform the following tasks:

– delivery and transportation of scientific equipment to the surface of the Moon to be used to conduct surveys, operation with the scientific equipment, power supply of the scientific equipment;

– core drilling along the pre-defined route and collecting of regolith stratified core samples with subsequent well logging;

– cryogenic drilling to perform regolith sampling;

– collecting and loading of fragments using the manipulating device;

– regolith samples collecting and logging (individual cell per each sample, each sample has a description specifying sampling time, location and coordinates, as well as physical conditions and photo);

– enroute explosives deployment and detonation in accordance with the pre-defined program to conduct active seismic experiment;

– area photographing and objects photomacrographing;

– evaluation of mineral reserves;

– radiation environment measurements;

– lava tubes exploration possibility;

– loading/unloading and deployment of container-type Automatic Scientific Stations (ASS) of long-term monitoring;

– loading of regolith and experiment samples to the MBVPK.

Taking into account the tasks mentioned above, we can develop initial requirements to the scientific heavy-class lunar rover to be used for geological survey of the surface of the Moon:

1) the lunar rover shall be equipped with the manipulation system capable of operation of the scientific equipment, collecting material fragments samples, loading and unloading them to the return vehicle to deliver to the Earth;

2) the lunar rover maximum travel range shall be not less than 400 km;

3) the lunar rover shall be equipped with drilling machine to perform well and core drilling;

4) the lunar rover shall perform multiple drilling (not less than 5 wells) throughout the entire route;

5) based on the modular configuration design, the lunar rover shall comprise two independent assembly units: chassis and body;

6) the lunar rover shall be designed with the all-terrain wheel chassis capable of reliable movement in loose sandy soil and crossing individual rock obstacles and crater wall slopes;

7) the lunar rover shall ensure operation in continuous shaded areas at the bottom of craters where the temperature can go down to -230°C.

One of the design concepts of scientific heavy-class lunar rover is a lunar rover project – “Robot-Geologist” being developed by Russian State Scientific Center for Robotics and Technical Cybernetics and Vernadsky Institute of Geochemistry and Analytical Chemistry of the Russian Academy of Sciences and TSNIIMASH, which is presented in Figure 4; and a scientific heavy-class lunar rover developed by Lavochkin Research and Production Association and TSNIIMASH, which is presented in Figure 5.

The heavy-class lunar rover shall perform the tasks to solve in manual, semi-automatic and automatic modes, i.e. under the control of:

– cosmonaut from the Lunar Orbital Station, e.g. from LOP-G;
– ground station;
– fully automatic control system.

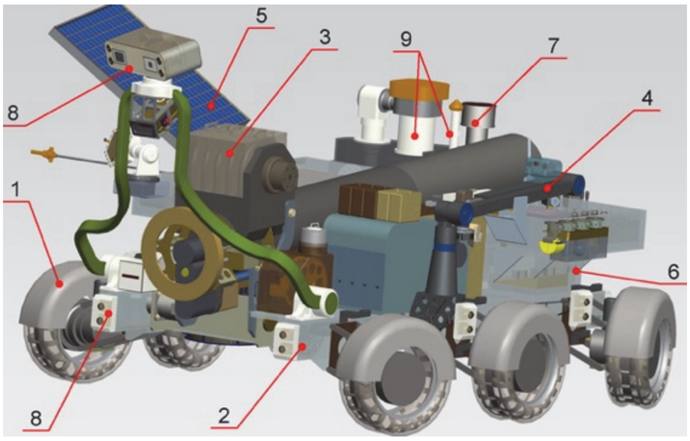


Figure 4 – “Robot-Geologist” scientific heavy-class lunar rover developed by Russian State Scientific Center for Robotics and Technical Cybernetics, Vernadsky Institute of Geochemistry and Analytical Chemistry of the Russian Academy of Sciences and FGUP TSNIIMASH

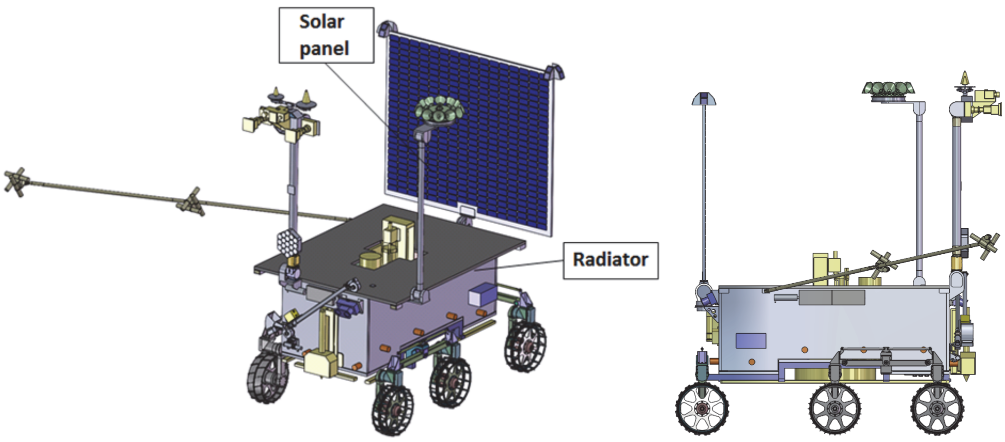


Figure 5 – Scientific heavy-class lunar rover developed by Lavochkin Research and Production Association and FGUP TSNIIMASH

4. Scenarios with manned unpressurized lunar rover based on heavy-class rover platform to support manned missions

At the stage of regular manned flights to the Moon by 2032, starting from regular-basis cosmonauts’ interaction with the automatic systems for manned missions support, use of manned unpressurized lunar rovers would become efficient, they would allow for significant move away from the LVPK landing site during the mission implementation and for significant extension of the area to be explored by the cosmonauts with identical spacesuit life-support system resources, as well as for increase in the transported equipment mass, use of loading/unloading manipulating device of a larger size comparing to a scientific heavy-class rover. Unpressurized manned heavy-class rover can be designed based on the chassis of the scientific heavy-class rover platform with different body – cabin and payload (Figure 6), it can be delivered using LVPK. In this regard, unification of heavy-class rovers platforms would ensure interchangeability of a number of their chassis elements that can be changed by a cosmonaut or service lunar rover if required. A potential configuration of a manned unpressurized lunar rover, being developed by Russian State Scientific Center for Robotics and Technical Cybernetics, is presented in Figure 6.

For the first time a manned unpressurized rover has been tested by the American astronauts in the frame of manned missions and ensured obtaining of highly-important results that would not be obtained without using it.

Figure 6 – A potential configuration of an unpressurized rover being developed by Russian State Scientific Center for Robotics and Technical Cybernetics

5. Scenarios with service-engineering lunar rover to support manned missions

A service rover to support infrastructure on the surface of the Moon can be designed based on the light- and heavy-class rovers chassis. A service-engineering rover can serve as a mobile engineering laboratory used by cosmonauts. This platform can accommodate processing equipment to perform experiments on lunar regolith baking, various additive technologies trying out (including 3D printing using lunar regolith) and others. Upon the completion of experiments program a cosmonaut can take results/samples back to the Earth.

Not only experiments can be performed but repair works as well, including those using remotely controlled manipulating device with interchangeable equipment. Rover of such class could ensure not only man-visited infrastructure maintenance (protection coating, replacement of consumable elements, repairing of module skin damage, replacement and rerouting of cables on the surface, etc.), but transportation of small-scale payload, changing of higher heavy-class rovers equipment and their partial repair and maintenance as well.

Maintenance of lunar infrastructure objects by service lunar rovers is considered to be a crucial task, since it can ensure a significant extension of operation lifetime of the objects under maintenance, thus reducing the need to manufacture and set into operation of new objects (to replace unserviceable ones). This can allow for decrease in expenses to be used for development and maintenance of the lunar infrastructure.

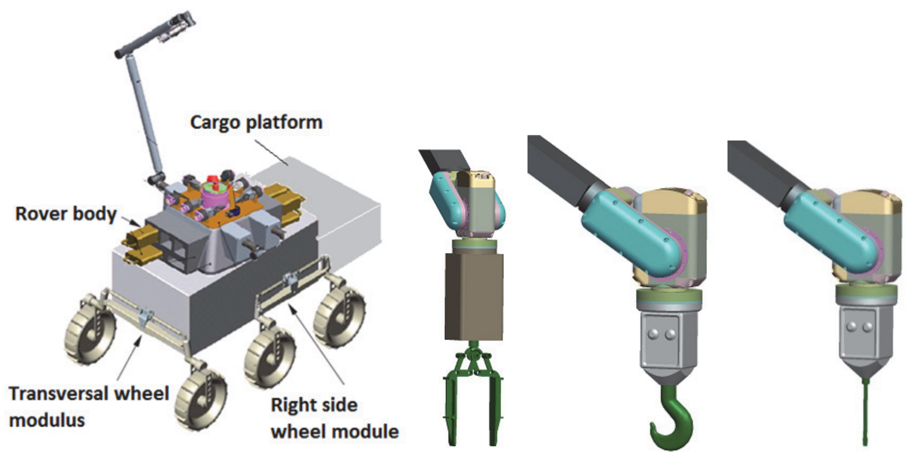


Figure 7 – A potential configuration of a service lunar rover based on the lunar rover platform developed by Lavochkin Research and Production Association using robotic means of Russian State Scientific Center for Robotics and Technical Cybernetics

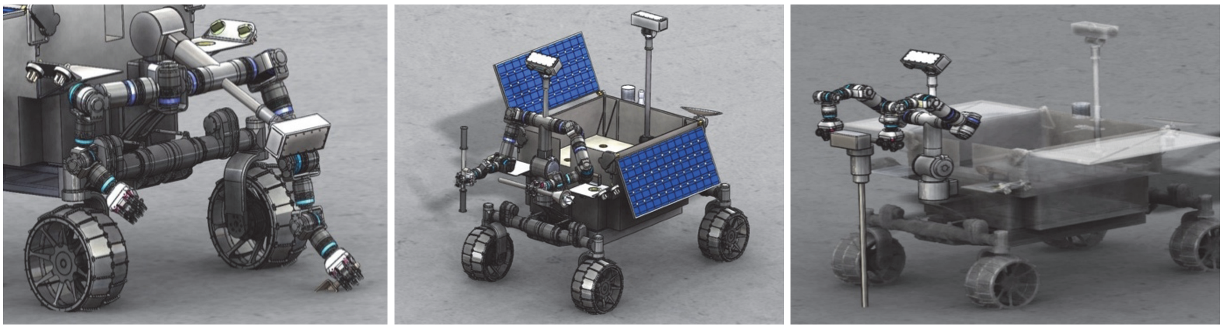


Figure 8 – A potential configuration of a service lunar rover developed by Russian State Scientific Center for Robotics and Technical Cybernetics

6. Scenarios with super heavy-class lunar construction rover

Beginning with the stage of a man-visited lunar base construction using a modular architecture design there will emerge a demand for special-purpose construction machinery capable of performing building and construction works under control of a man. Each individual base module will be delivered to the surface of the Moon separately. To prevent collisions and potential damages modules of a man-visited infrastructure will be delivered to the surface of the Moon a certain distance apart from each other. Upon the completion of soft landing the modules shall be unloaded from the landing platform and be transported to each other to be coupled by the lunar transport-loading manipulating device. Transport-loading manipulating device can be accommodated on the super heavy-class mobile vehicle platform.

Development of a basic super heavy-class rover platform (approximately 5-6 t) with remote control and AI elements will significantly contribute not only to the Moon exploration but to the Moon research as well. Availability of even one rover of the type will ensure the use of interchangeable equipment to perform a wide range of work (a crane, a bucket, a drilling machine to drill down to 15 meters).

Thus, the platform under consideration can ensure deployment of not only initial man-visited lunar base elements but deployment of lunar observatories and installation of heavy scientific equipment as well (experiments suggested by the Russian Academy of Sciences include the tasks that require accommodation of equipment with mass of ~ 4 tons on the surface of the Moon which is comparable with the mass of man-visited infrastructure modules), and set up of experiments that are impossible to conduct with the help of an automatic spacecraft of a lighter class or the manned missions crew.

After the landing on the surface of the Moon the lunar rover shall ensure:

– performance of construction works related to the installation of manned infrastructure elements in the surface of the Moon;

– performance of landing site clearing to support arrival of cargo spacecraft and manned lunar launch and landing complexes;

– conduct of experiments on regolith excavation;

– conduct of experimental digging of ditches and trenches;

– conduct of experimental regolith transportation at various distances;

– testing of regolith application as a construction material (banking and mounding of habitable modules);

– transportation of large-scale scientific, industrial and processing equipment;

– performance of lifting and rigging works using a crane;

– towing of trolleys with heavy processing equipment;

– performance of work with the manipulating device;

– power supply of the industrial processing equipment;

– area photographing and objects photomacrographing;

– performance of experimental works on extraction of volatile compounds from regolith (hydrogen, oxygen, water, helium-3, etc.);

– performance of experimental works on extraction of chemically bonded elements from regolith (oxygen, silica, metals);

– performance of experiments on thermal processing of regolith (baking, smelting);

– performance of trying out works on regolith excavating and construction and installation machinery;

– performance of experiments results/samples loading on MBVPK;

– operation under control of a cosmonaut from the orbital station, e.g. from LOP-G;

– operation under control of ground station.

Preliminary works are proactively performed by Lavochkin Research and Production Association in collaboration with FGUP TSNIIMASH. At a later stage such platform can become a prototype for development of a chassis of a manned pressurized rover, while the trying out of its significant part can be performed using an automatic prototype. The manned pressurized rover compartment can be arranged on the loading platform. Furthermore, using of super heavy-class lunar rover will allow for deployment of a fully automatic lunar observatory and perform its maintenance in the future. A potential configuration of the super heavy-class construction rover being developed by Lavochkin Research and Production Association is presented in Figure 9.

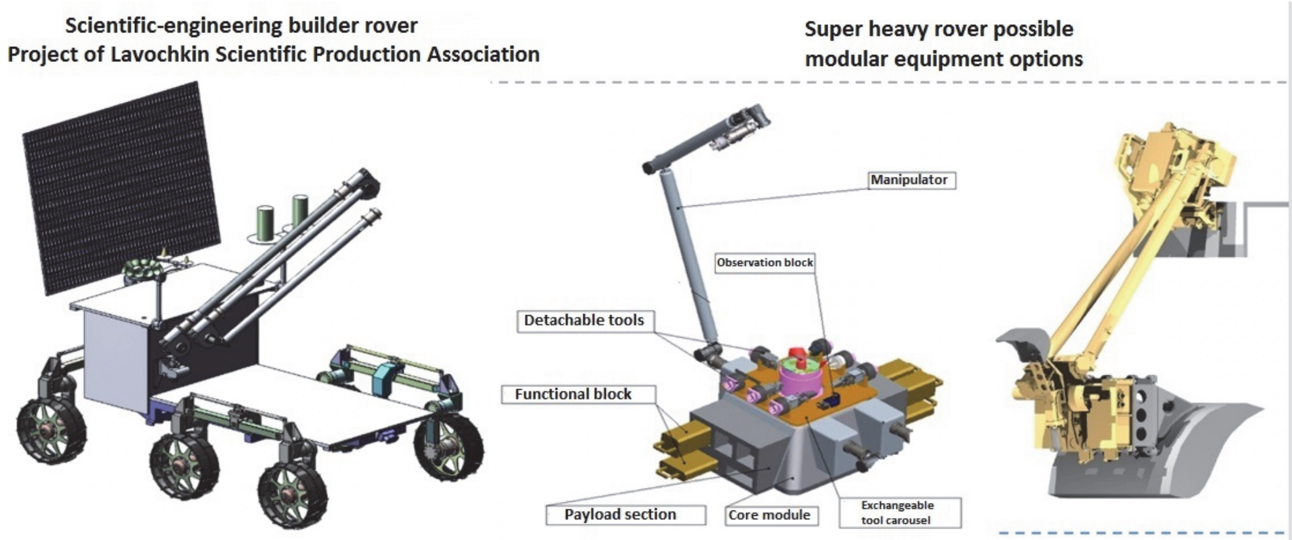


Figure 9 – Potential configuration of a construction rover being developed by Lavochkin Research and Production Association and FGUP TSNIIMSH with possible using of robotic means of Russian State Scientific Center for Robotics and Technical Cybernetics

7. Conclusion

To perform research and exploration of the lunar surface it is reasonable to develop a heavy-class rover platform being more efficient than a light-class rover platform in terms of its functionality and prospects of its use at the subsequent exploration stages. The use of a light-class rover is efficient only at the early lunar exploration stage to perform tasks on trying out of technologies, systems and subsystems. After the development of a heavy-class mobile platform, the relevance of the new light-class rovers development will significantly decrease.

Heavy-class lunar rover platform can be used for trying out of radioisotope power source technology that can significantly expand potential platform functionality.

Development of a unified heavy-class mobile platform will ensure development of various specific-purpose heavy-class rovers on its basis using tried out technical solutions and at minimum expense:

– scientific-engineering rover designed for minerals exploration, implementation of cartographic and other applied tasks, including conduct of fundamental scientific experiments using interchangeable equipment, as well as processing experiments (e.g. regolith baking, additive technologies, etc.), can serve as a mobile laboratory accompanying cosmonauts during the mission on the lunar surface;

– service rover for the maintenance of the first infrastructure elements can be used for installation of navigation beacons, changing of interchangeable equipment and replacement of consumable elements of the man-visited modules, etc.;

– manned unpressurized rover to extend exploration area in the vicinity of the landing site.

As a result of the mentioned-above spacecraft development the exploration area on the lunar surface and list of tasks to be implemented by cosmonauts could be extended by a several-fold factor. Furthermore, spacecraft developed on the basis of a heavy-class rover platform will be in a high demand at the subsequent stages of research and exploration of the Moon as well.

During the Moon exploration spacecraft development the technologies that can be applied in related technological areas on the Earth and in space could be developed and tried out, including:

– radiation- and electromagnetic irradiation-resistant ECB;

– compact computing equipment generating minimum heat;

– artificial intelligence systems, robotic systems and spacecraft automatic control technologies under composite topography conditions;

– technologies and systems of self-contained medical and life support systems;

– additive technologies used to develop large-scale structures.

The development of the mentioned-above spacecraft is a precondition for deployment of a serviced automatic test site to perform applied scientific experiments and semi-industrial manufacturing works, including:

– transport system with manned and cargo spacecraft, landing and launch modules;

– uninterrupted power supply facilities;

– crew accommodation modules;

– key elements of support systems (retransmission units, medical supplies, robotic equipment to perform work on the testing site);

– equipment to perform scientific research;

– experimental complexes for utilization of the lunar resources to perform construction and base operation.

O.V. Rudakova

TESTING-OUT ROBOTICS CONTROL TECHNOLOGIES FOR MOON EXPLORATION

Federal state unitary enterprise Central Research Institute for General Machine Building, Korolev, Russia
EmeldyashevaOV@tsniimash.ru

Abstract

Existing experience in development and utilization of space-purpose robotic systems shows that such systems present a powerful and unique instrument in solving tasks related to outer space research activities, development of space systems and exploitation of such systems. The main development vector of Russian national human space exploration activities is Moon research and exploration. To achieve such a goal it is required to start, as in today, developing key technologies in space robotics and manufacturing control means for on-planet robotic systems. It would be prudent to use ISS resources to run advanced technologies and space systems to be able to successfully run interplanetary spaceflights and to be able, to build on-planet bases as well as to simulate manned space flight to other objects in our Solar System because ISS is a reliable, safe and well-equipped platform for implementation of scientific and practical research and experimental activities.

Keywords: space experiment, Moon, Moon rovers, International Space Station, on-planet robotics, orbital station, testing out of technologies, robotic systems, telecommand, human-machine interface.

1. Existing scientific and technological capacity in frame of space robotics design and advanced robotic facilities for Moon research and exploration

According to the Manned Space Exploration Strategy until 2035, the following tasks are critical [1]:

– Further exploitation of Low-Earth Orbits (LEO) for the purpose of testing out advanced systems and technologies as well as switch to serviceable space on LEO;

– Moon research and exploration;

– Generating technological capacity for further human expansion into deep space (towards Mars and asteroids).

To achieve such goals, we need to develop new technologies and create advanced crew support systems of on-planet and space base, particularly a wide variety of mission-critical Space Robotics Systems (SRS).

SRS introduction into space environment will allow cosmonauts to increase the percentage of their time which will be used for scientific-technological and research activities resulting from all the time saved on the decreased amount of routine work that crews of space crafts (SC) are ought to do in frame of their extravehicular activities and activities in pressurized modules. SRS introduction will make it possible to increase the safety of conducted space activities, diversify the list of robotic operations aimed at localizing and eliminating off-nominal and emergency situations as well as support space crew activities, both extravehicular and in pressurized modules and, in the long run, even in frame of research and exploration activities on surface of the Moon and other planets.

Main areas of space robotics implementation are:

– Status control of maintainable objects;

– Re-docking of space station modules and spacecrafts;

– Mounting and dismantling activities;

– Preventing emergency situations;

– Repair work activities;

– Maintaining payloads and service systems of space stations;

– Building protective structures for on-planet manned bases;

– Technical maintenance of on-planet base elements;

– Execution support of intra and extravehicular activities of cosmonauts;

– Taking soil probes from the surface of Solar System objects;

– Assembling large-scale space systems.

Currently, national space sector managed to gather a formidable amount of scientific-technological capacity in frame of space-purposed SRS design. Particularly, regarding Federal Space Program related activities, there are certain achievements made and further research and R&D being fulfilled, aimed at designing cargo handling dollies, humanoid robot-assistants, technological mobile rectangular units for extravehicular works on Manned Space Complexes (MSC), crew info-support robots etc. But there are other issues that exist in this area which have been looked into on a surface level only, that are the things related to on-orbit technical maintenance, assembly-related activities in space environment and space debris disposal. These are very pressing issues and they require much more attention from the government.

Today, the further development vector of the whole space industry is Moon research and exploration.

Moon exploration may comprise the following stages:

1. «Preparation» - 2022 – 2029.
2. «Outing» - 2030 – 2032.
3. «Outpost» - 2032 – 2035.
4. «Base» - 2035 onwards.

Each stage has a set of unique tasks to solve. These tasks define baseline scenarios of missions (manned and unmanned) and SRS baseline layout required to execute them.

The following on-planet based SRS should be designed to support manned spaceflights aimed at cislunar space and space crew landings on the Moon surface.

“Preparation” stage is planned to incorporate development tests and demonstration of manned lunar and siclunar flight key technologies. This stage requires maximum utilization of ISS RS and automatic space crafts (ASC) to run development tests of manned lunar missions’ support systems and subsystems. Particularly, it is necessary to run development tests for key technologies used to design moon rovers, control systems and interaction means between SRS and cosmonauts-operators and ground-based operators.

“Outing” stage is planned to incorporate short-term (duration of 7 to 14 days) manned missions with 4-man crews. This stage implies Moon research via automatic and manned facilities, Moon surface exploration activities should be implemented as well as development tests of on-planes technologies and preparation work for Moon surface infrastructure establishment. The result of the analysis showed that “Outing stage” requires construction of two types of moon rovers: light class (300-500 kg mass) and heavy class (up to 2 t mass). At that, utilization of light-class lunar rovers is only feasible at early-stage research – to run development tests for systems and subsystems. Design and utilization of heavy-class lunar rovers is more efficient from the angle of functional capabilities and their utilization perspectives on further stages of Moon exploration. Heavy-class lunar rovers might solve the following issues:

– Selenological exploration capable of probing for mineral resources using drill fixture;

– Running fundamental scientific research with the help of a wide variety of scientific equipment onboard;

– Development tests of construction and maintenance technologies in Moon environment (subsoil baking, additive technologies, clearing sites etc.);

– Crew and cargo transport to Moon surface.

“Outpost” stage is planned to run recurrent crewed expeditions on Moon (duration of up to 30 days, 4-man crew). The results of “Outpost” stage research activities should lead to a decision over feasibility of permanent manned Moon base establishment, allocation of permanent manned Moon Base (MB) elements, MB baseline infrastructure should be established as well (transport, power supply and communication). This stage implies design of super-heavy class lunar rovers (on top of already utilized light- and heavy-class lunar rovers) with a mass up to 10 t. Such a rover, using interchangeable instruments and equipment (e.g. crane unit, bulldozer unit etc.) will be able to solve the following tasks:

– Transport and assembly of man-tended modules;

– Clearing landing site for cargo spacecrafts and manned lunar takeoff and landing sites;

– Construction work execution (excavation, trenching and digging ditches, banking habitable modules, drilling down to 15 m, regolith extraction etc.);

– Execution of transport activities aimed to move oversized cargo and equipment;

– Running scientific and technological research on the Moon (including establishment of an observatory).

Aside from builder-rover, there is a need to design a super-heavy class pressurized manned rover required to enlarge the investigation area for crews during their long-stay missions.

“Base” stage is planned to run full-fledged expeditions (duration of up to 60 days, 4-man crew), building different structures, itemization of such structures with corresponding instruments and equipment, establishment of corresponding infrastructure and, further down the line, establishment of proper pilot production on the Moon surface with long-term expeditions (duration of up to 180 days, 2 expeditions per year, 4-man crew). This stage implies rapidly growing demand for super-heavy class builder-rover and heavy-class rovers required to solve maintenance-related tasks.

Availability of such robotics utilized in maintaining crewed missions allows for a drastic increase in a crew's capabilities at all stages of Moon exploration. Moreover, it allows us to run development tests for key technologies required in further human expansion into deep space.

2. Space robotics facilities design technologies

SRS development, to a great part, depends on progress made in contemporary baseline technologies, such as technologies of accurate positioning, machine vision, biomechanical equipment, and artificial intelligence and so on. One should note that designing such advanced robotics requires continuing development in the following key robotics industry sectors:

– Mechatronic units (rockers), multi-joint manipulator design technologies which allow continued operation in extreme conditions, including space environment;

– Mechatronic-modular system design technologies with the ability to modify their own structures;

– Advanced sensor system and machine vision design technologies;

– Remote control technologies for space robotics complexes, with signal distribution delay (in communication lines);

– Adaptive control technologies (for robotics systems) with intelligent technics: expert systems, neural networks, fuzzy logic etc.;

– Gang control technologies for robotics systems;

– Advanced human-machine interface design technologies;

– Advanced small-scale power saving systems design technologies for robotics complexes;

– Virtual simulation and ground development technologies for robotics systems;

– Information processing technologies and software for integrated control systems and robotics systems control and identification;

– Design technologies and prototypes of dedicated equipment and interchangeable rigging with corresponding mounting elements.

3. Main principles of on-planet robotics facilities control

Solving complex technological tasks and running scientific research using Moon rovers looks feasible in supervisory and teleoperator-mode from ground control site, from the board of Lunar Orbital Station (LOS) and, in prospect, from on-planet habitable modules. Furthermore, rover control form the ground (Earth) implies time delay in communication channels – more than 2.5 sec. Taking into account the time required by a human-operator to process the signal, cumulative time delay doesn't allow full-fledged telecommand control (that includes force feedback) from the ground (Earth). If we look at rover control from LOS, delay is minimal (depending on the height of LOS flight, median delay time is 60 milliseconds) and is compared to human reaction delay which makes telecommand control feasible. It will allow minimization of command dispatching signal delay, as well as rover feedback delay minimization, making it harder for unintended consequences to occur.

Currently, front-end engineering design of Lunar Orbital Platform-Gateway (LOP-G) is running, together with negotiations over Russian participation in this project. Taking that into account, the issue of development and testing control technologies over on-planet robotics from the LOS board looks very concerning. LOS is not the only topic in the list of plans related to cislunar infrastructure build-up. Cislunar orbit will see relay satellites fostering navigation and communication between the Moon, cislunar infrastructure and the Earth. When human activity on the Moon surface will begin, orbital constellation will be reinforced by relay satellites facilitating stable broad-band communication channels.

Analysis of existing international technological capacity in frame of rover control showed that, currently, space objects' (on-planet, orbital) control methods and systems based on Earth surface are the most well-developed. Communication establishment and control (between rovers and Earth) experience was gained during the following missions: «Lunohod» (1970 – 1973), Mars Pathfinder (1996 – 1997, Sojourner rover), Mars Exploration Rover (beginning from 2003 to the present day, Spirit and Opportunity rovers), Mars Science Laboratory (beginning from 2011 to the present day, Curiosity rover).

To establish efficient control over on-planet SRS, one must develop telecommand technologies which will allow for establishment of advanced human-machine interface giving operator a full immersion capability into the environment of an object under his control and giving operator comprehensive feedback information (visual and force-torque). Currently, virtual reality based interfaces are experiencing rapid development. On the other hand, employment of virtual reality technologies in frame of complex drivers working in space environment is not studied sufficiently yet.

Creating efficient human-machine interface on must run a plethora of preliminary studies, such as:

– Investigate microgravity effects over robotics control processes employed by human operators using different drivers and feedback (kinesthetic effect over operator);

– Define a set of drivers which can be used to maintain sound and efficient control over on-planet robotics and estimate applicability and convenience of use for such drivers;

– Run a number of investigations into psychological and psych-physiological status of a human-operator working in an environment of sensor disintegration and long-term microgravity exposure;

– Run ergonomic analysis of an operator workstation design (used for on-planet mission control purposes);

– Run development tests of on-planet robotics' control means using different communication channels (with different capacity) and with present delay in communication channels;

– Define and test-out reference cases for on-planet robotics' operation as well as control methods over robotics systems in frame of solving targeted tasks in undetermined conditions of lunar terrain;

– Run research of virtual reality implementation into processes of on-planet robotics telecommanding taking into account integration of information received from SRS machine vision means.

Abovementioned factors affect telecommand efficiency, reliability and responsiveness.

4. Testing out key elements and technologies of space robotics control on ISS

Is would be prudent to use ISS RS resources to run mentioned research and experimental activities as ISS is a reliable, safe and well-equipped testing stand, well-suited for development tests of advanced technologies and space facilities aimed at interplanetary flights and establishment of on-planet bases as well as human flight simulation with respect to other objects of Solar System. It is assumed that ISS fill continue is operation till the year 2024.

SRS development tests in space environment onboard of an MSC allows for the most accurate consideration of space flight – related extreme conditions (overloads, zero gravity, communication delay etc.) in frame of SRS control processes and different SRS operations.

Currently, ISS RS hosts a considerable number of research activities focused on different scientific and applied areas. The number of development tests focusing on different technologies related to Moon and Mars expeditions is growing. In that regard, the following ISS RS resources might be employed:

– Power supply;
– Information exchange;
– Generic interfaces;
– Universal workstations;
– Containment and external surface of modules;
– Cargo traffic;
– Crew time.

Currently, there are some changes being implemented onboard of ISS RS in relation to ISS experiments and research scheduling and execution for the purposes of improving efficiency of ISS designated use, solving general scientific and applied tasks as well as streamlining ISS RS related operational procedures [2]. Particularly, these changes are related to:

– Scientific and applied research program scheduling in the form of Long-term Program of Applied Activities (LPAA);
– ISS RS applied activities preparation and execution algorithm;
– Applied activities' results analysis execution algorithm.

ISS RS has already seen space Experiments (SE) related to development and development tests of SRS remote control. Such activities, for the most, include international projects like «ROKVISS», «Kontur» and «Kontur-2» (in partnership with German aerospace center DLR and RSSC RTC). «ROKVISS» experiment managed to demonstrate control capabilities over double-joint robotic arm installed on the external surface of ISS RS. S-band communication channel used allowed for simultaneous transmission of commands and telemetric data between the robot and ground automated workstation (AWS) of an operator which included force-torque joystick. This made it possible to simulate telepresence operating mode and test-out force-torque sensing technology.

Joint Russian-German «Kontur» experiment data transfer network infrastructure and «ROKVISS» experiment scenarios but, in this case operator's AWS was installed in RSSC RTC and access to S-band channel was maintained via Internet. A control environment with significant delay in communication channel was tested-out during this experiment.

In frame of «Kontur-2» experiment, cosmonaut-operator's AWS was installed onboard ISS RS and robots under control – on Earth, in DLR and RSSC RTC. The aim of «Kontur-2» experiment was to test-out force-torque telecommanding technologies operating on-planet rovers from the board of manned orbital complex.

The following tasks were set for this experiment:

– Design telecommand system for on-planet robots to be operated from ISS RS by a driver with a force-torque feedback and using communication channels with limited capacity and existing time delay;

– Design automated workstation onboard ISS RS which will allow for telecommanding over robots on Earth;

– Research particularities of human-machine interface with force-torque feedback used for remote control in micro gravitational environment.

At the moment, there are a number of space experiments being prepared which are aimed at running space SRS development tests onboard of ISS RS. Among them, there are SEs (currently, at ground preparation stage) focused on investigating operational capabilities of remote-controlled anthropomorphic robots within next-gen advanced Manned Cargo Ships (SE «Ispitatel'», SE "Telegroid"). At that, these experiments will incorporate testing out operation modes of anthropomorphic SRSs placed onboard of ISS RS or MCS "Federation" (in case of SE «Ispitatel'»), from Earth and ISS RS in automatic and "duplicating" modes.

5. "Kontur-3" joint Russian-German experiment overview

There is a very promising (from the on-planet-objects-control-from-space development tests point of view) new "Kontur-3" joint project being planned for ISS RS by DLR, RSSC RTC and FSUE TSNIIMASH. This project is at the stage of program integration at the moment. There is substantial technological contribution foreseen from both Russian and German party, particularly, interactive utilization of developed (by respective parties) human-machine interface devices as well as development of new space station-based telecommand solutions for robotics systems.

For the purpose of testing out telecommand systems exercising control over on-planet robotics objects on the Moon surface, particularly, lunar rovers, it is proposed to establish, in frame of this experiment, a simulation of lunar rover telecommand system by via creation of (physical) closed-loop control system, which includes:

– Cosmonaut-operator;

– Human-machine interface with a variety of control means;

– S- and Ku-band space communication channels;

– Lunar rover prototype (planned for manufacturing if frame of the project);

– Site (stand) simulating lunar terrain and illumination properties;

– Ground-based control consoles (AWS of ground operators);

– Corresponding software.

The main features of «Kontur-3» project are:

– Control subjects used are doing to be lunar rover prototypes close in their technical properties to real rovers purposefully made for lunar missions;

– Development and utilization of a number of human-machine interface drivers allowing for efficient and intuitive telecommanding;

– Utilization of cosmonaut activity research development tools in a world picture distorting micro gravitational environment;

– Multiplexing S- and Ku-band space communication channels (basing it on «Luch'» SC).

The following thins are meant to be used as human-machine interface drivers:

– «RJo» single-degree-of-freedom joystick with force feedback which was verified in frame of «Kontur-2» SE and placed onboard of ISS RS;

– «Delta» Six-degree-of-freedom driver (developed by RSSC RTC) with force feedback;

– Virtual reality headset with integrated SW (developed by RSSC RTC);

– «DataSuit» and «DataGlove» drivers (developed by DLR);

– «VibroTac» vibrotactile sensing wristband (developed by DLR).

According to RSSC RTC calculations, telecommunication infrastructure will incorporate S- and Ku-band communication channels. It is planned to utilize existing S-band network infrastructure which was created during previous experiments and to update it to be able to use Ku-band. Multiplexing these two channels will give us the opportunity to combine operational features of S-band direct radio contact (low signal delay, low capacity, short period of radio contact) with Ku-band indirect contact using «Luch'» satellite relay SC (significant signal delay, high capacity, long period of radio contact). This way we can execute SRS telecommanding operations with high longevity.

«Kontur-3» experiment envisages establishment and testing out technical feasibility of channel switching and their simultaneous use.

Figure 1 shows baseline interaction pattern in frame of «Kontur-3» project fulfillment (based on RSSC RTC data).

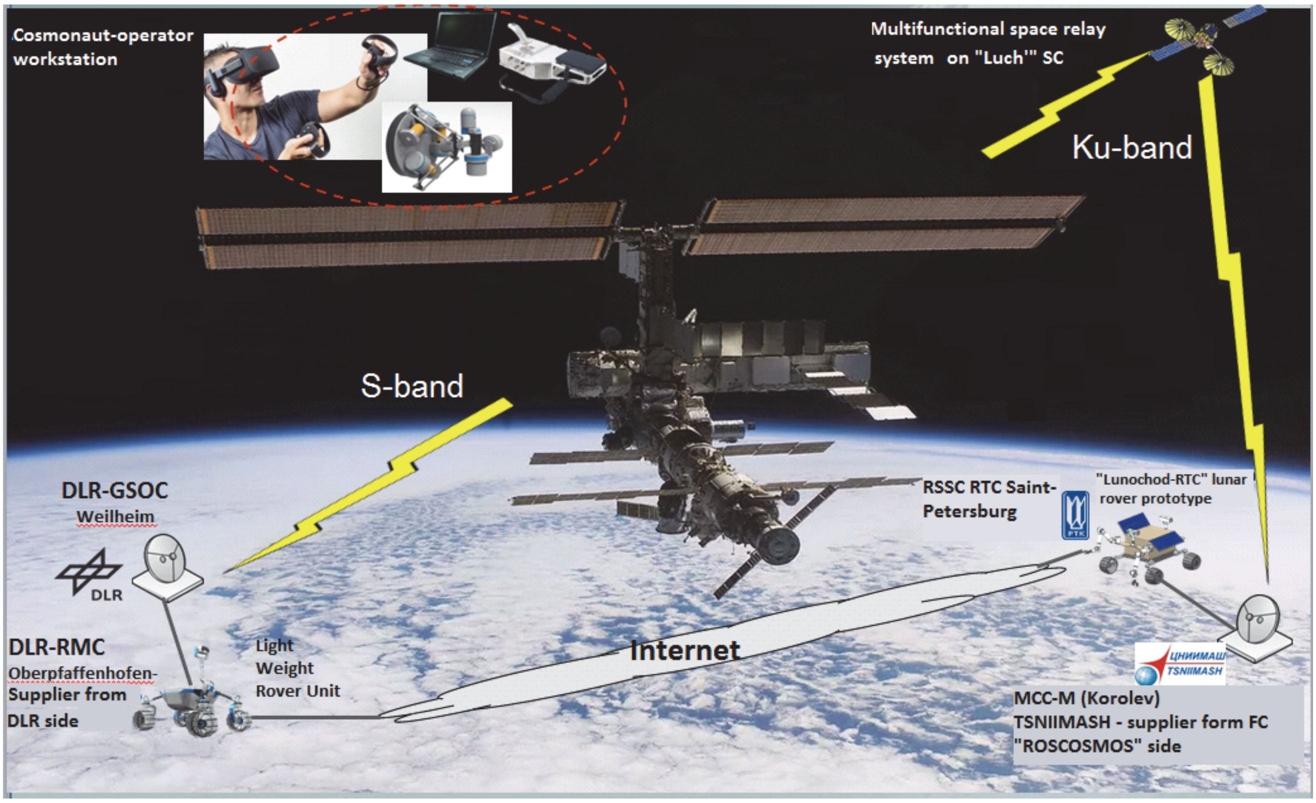


Figure 1 – Baseline interaction pattern in frame of «Kontur-3» project fulfillment

FSUE TSNIIMASH is a technical supplier in frame of this project as well as an organization responsible for the lunar rover prototype design. RSSC RTC here serves the role of scientific supplier.

In frame of "Konur-3" project, the following key technologies are planned for testing-out (based on RSSC RTC data):

– Telecommanding technologies (working from the board of an orbital station) focused on control over the movement of a mobile robotics system (lunar rover) operating on the Moon;

– Telecommanding technologies focused on control over manipulator system and grappling arms with force-torque feedback from board of an orbital station;

– Automated (supervisor mode) control technologies focused on control over the movement of a mobile robotics system (lunar rover) operating on the Moon;

– Lunar rover mobility system design technologies;

– Movement control technologies for multistage robotic self-propelled platform focused on cases of extreme obstacles crossing while operating on the Moon;

– Design technologies for manipulator system and grappling arms purposed for work on the Moon;

– Research technologies for the Moon surface, using mobile robotics system equipped with a robotic arm with a force-torque feedback.

The data acquired during«Kontur-3» experiment might be used in:

– Design and development of advances MSCs;

– Placing ergonomic workstations onboard MSCs – such workstations will provide cosmonaut-operators (controlling remote technical complexes, incl. on-planet SRS) with effective working conditions;

– Planning and execution of future missions prompting use of on-planet SRS, e.g. rovers, in frame of research missions on the Moon and other Solar System objects;

– Organizing space communication systems purposed for on-planet SRS control;

– Developing scientific and applied engineering disciplines for designers of complex technical systems using human-machine interaction in micro gravitational environment as well as other space environment factors.

Conclusion

The design of robotic facilities aimed at manned missions support is essential for Moon research and exploration. Particularly, it is necessary to design three classes of rovers: light class (300-500 kg), heavy class (1-2 tons) and super-heavy class (up to 10 tons). Design of such advanced facilities is closely interconnected with development of corresponding baseline technologies. One of the main technology development vectors he is remote control technology over robotic facilities (incl. on-planet ones).

At the moment, national space industry has managed to gather a substantial amount of experience in frame of space robotics systems design, there is a number of research and experimental activities being carried out in a real space flight environment (using ISS resources) to test-out key elements of space robotics and robotic facilities control technologies.

"Kontur-3" joint Russian-German experiment is being prepared at the moment onboard of ISS RS in order to facilitate establishment of technological capacity in the field of on-planet robotic facilities control for the time of carrying-out future Moon research and exploration missions. In frame of this experiment it is planned to simulate telecommanding process over lunar rover prototypes from the board of ISS RS as well as carry out research focused of cosmonaut-operator while he is executing control tasks.

References

1. «Russian manned spacecraft strategy till the year 2035 », 2015.
2. «Provision №1 about planning and execution procedures for target-oriented activities on International Space Station », 2018.

V.N. Dmitriev, B.V. Burdin, V.A. Dovzhenko, Yu.S. Chebotarev

APPLICATION OF SPACE ROBOTIC SYSTEMS TO SUPPORT COSMONAUTS' ACTIVITY FOR THE IMPLEMENTATION OF EXISTING AND FUTURE SPACE PROGRAMS

FSBO "Yu.A.Gagarin Research&Test Cosmonaut Training Center", Star City, Russia
V.Dmitriev@gctc.ru, B.Burdin@gctc.ru, V.Dovzhenko@gctc.ru, Y.Chebotarev@gctc.ru

Abstract

The paper discusses issues of the application of space-dedicated robotic systems (RS) and shows the role and place of anthropomorphic robotic systems as high-tech service systems used to support cosmonauts' activity in the course of future space missions. The principles for creating a multipurpose computer-based stand of anthropomorphic RS in order to carry out ergonomic and psychophysiological research and to train cosmonauts to operate various robotic systems using VR-technology and hardware for controlling motor and behavioral activity of the operator are reviewed as well.

Keywords: space-based robotic systems, robotic manipulator, anthropomorphic robot, copy type master device, remote control mode, multipurpose computer-based stand, virtual 3D models, human-computer interface.

Overview of robotic manipulators used in manned space programs

Future projects in the sphere of manned space exploration presuppose the expansion in the application of space robotic systems (SRSs) as a result of the increased complexity of crews' activities aboard manned space complexes (MSCs) and volume of labour-intensive assembly-recovery works and routine operations performed by cosmonauts aboard MSCs as well as the heightened risks of dangerous situations for human life and health during near and deep space missions. In this regard, the use of SRSs as a high-tech toolkit of a cosmonaut when performing intra- and extravehicular activity as well as on-planet operations during future lunar and Martian missions is of particular interest.

A number of strategies for designing and using the RS in manned space exploration, developed by leading foreign and domestic companies are well known today. An example of the successful creation and use of a robotic manipulator is the Soviet manipulator "Lyappa". "Lyappa"-type manipulators were installed on 4 modules on the Mir orbital complex and were used to assemble the station. [1].

The second vivid example is the Canadian "Canadarm" manipulator (Figure 1), which was mounted on the Space Shuttles (reusable low Earth orbital spacecraft system) and used from November 1981 through July 2011. The main purpose of the "Canadarm" was to transfer payloads from the cargo compartment of the MTKK to a certain place in the working area with the required orientation. It was used to unload more than 200 tons of the ISS elements and also to capture the Hubble telescope to repair it in the orbit.

The "Canadarm" has evolved into the "Canadarm-2" (Space Station Remote Manipulator System – SSRMS) (Figure 2), designed to perform in-orbit operations with big-size and heavy-weight cargoes [1-8].

Among other manipulators, the Japanese Experiment Module Remote Manipulator System – JEMRMS (Figure 3) [2-7] and the European Robotic Arm – ERA should be noted [9, 10].

Figure 1 – The Canadian "Canadarm" manipulator

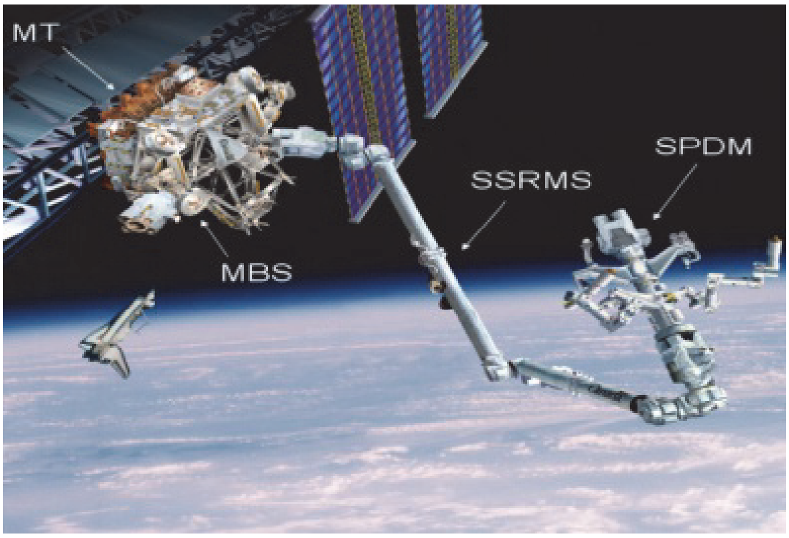


Figure 2 – Mobile Service System (MSS)


Figure 3 – JEMRMS manipulator system on the Japanese experimental module of the ISS

All the above-mentioned manipulator systems are the result of high-tech, expensive innovative projects requiring the building of technological stands which are necessary for the experimental testing of samples and the development of simulators for cosmonaut training.

At present, along with the design of manipulator robots, intended to perform assembly, installation and various operations on the outer surface of the ISS RS modules, special attention is paid to the building and use of anthropomorphic robotic systems (ARTS).

Just as in manipulator systems, the kinematics of the ARTS includes the shoulder, elbow and wrist parts. The difference lies in the number of manipulators. A torso-type robot has two arms hinged to its torso part. Due to its hands and fingers the ARTS can use standard human tools.

Initially, the ARTS is planned to be used in the master-slave mode for performing maintenance operations on the ISS [11]. It should be noted that there are various ways to build a human-computer interface for remote-controlled robots with autonomous operation capabilities. So, as an alternative to the contact data retrieval methods, where a human operator interacts with the robot controllers at the physical level, non-contact methods of generating control signals using motion recognition technology, motion capture system and oculographs can be applied. [12-15].

Participation of Yu.A Gagarin Reasearch&Test CTC in the building, experimental development and scientific support of the space-dedicated ARTS

Currently the issues with evaluating the utilization efficiency of the ARTS, the possibility of performing flight operations using anthropomorphic robots, optimizing the interaction between humans and robots are open for discussion and require special studies. The problem of matching up the elements of the “man-ARTS-environments” ergatic system (ES) comes to the fore, since the limitations of the human factor and the ways to overcome them are not well understood, especially in cases where this interaction is undertaken in a remote control mode. This problem encompasses the difficulty of prompt data exchange through communication channels in the course of deep-space missions because of long signal delays (for example between the Earth and the Moon or between the Earth and Mars).

Since 2010, Yu. A. Gagarin Research&Test CTC (hereinafter the Center) has been taking part in the development and scientific - technical support of the space-based robotic systems. A number of ergonomic studies and tests of various space-based robotic systems are carried out in conjunction with FSUE TsNIIMash.

From July 2011 to the present the anthropomorphic robotic systems, developed by NPO “Android Technology” (Figure 4) are being tested. [16]. A robot-crew assistant, developed by Neurobotics LLC (Figure 5) and a robotic EVA complex, developed by TSNII RTK (St-Pb.) were examined in 2014.

During testing, the ARTS SAR-401, designed by NPO “Android Technology” performed “on-board operations” copying the motions of the human operator via a special exoskeleton. The exoskeleton is a master device which is placed on the operator’s shoulders. [16]. Operator’s movements were copied by the robot in a real-time mode. [12-17].

Figure 4 – ARTS replaces the elements of the “Biorisk” scientific equipment in a master-slave mode

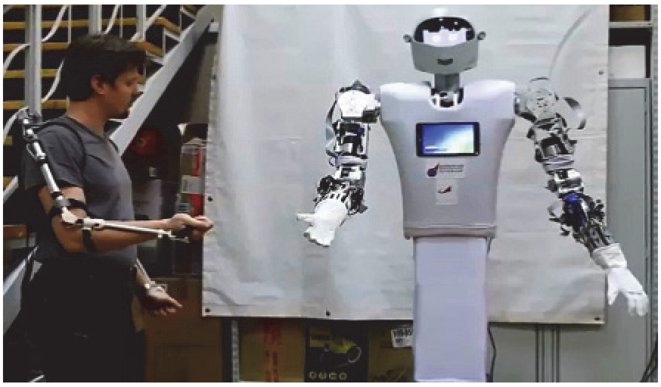


Figure 5 – ARTS developed by Neirobotics LLC in the process of data acquisition on the psychophysiology and emotional state of the cosmonaut-operator

In the course of the “Kontur-2” experiment a cosmonaut-operator being aboard the ISS operated the robot located on the Earth via combined communication channels with limited bandwidth and random delays in signal transmission (Figure 6).

The objectives of the experiment were:

– to develop a robot remote control system using a master manipulator with force-moment feedback via communication links with limited bandwidth;

– to develop a computer workstation providing remote control of ground-based robots from the board of the ISS;

– to study the functionality of the human-computer interface with force-moment feedback in microgravity conditions.

Taking into account a high cost of building and high risks of using space robotic systems, the development of the RTS computer stands suitable for experimental ergonomic study of multipurpose “cosmonaut-RS-environment” system is of practical significance. The stands are also used to train cosmonauts to control robotic systems during manned space flights thereby reducing the potential risks for cosmonauts in the course of their work with new systems and equipment.

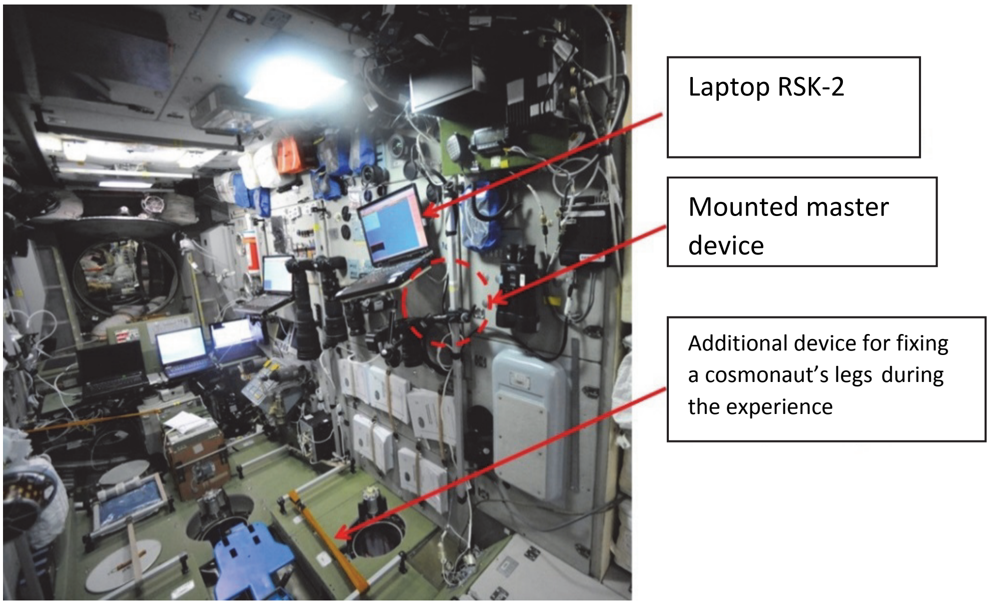


Figure 6 – The workstation aboard the ISS for carrying out the “Kontur-2” experiment

Composition, purpose, prospects for the use of the Multipurpose Computer Stand of Space Robotic Systems (VCS SRSs)

In the early stages of designing the “ARSs-cosmonaut-environment” ergatic system the Center created and used the research computer-based stand of robotic systems as the most acceptable solution for the experimental testing of real anthropomorphic robotic systems (ARSs).

In this case, the main requirements for the stand were the openness of architecture and versatility. This allows to simulate the "RSs-operator-activity environment" system and interface tools of various types on the basis of a unified set of software and hardware tools and visualization (virtual interactive 3D-models of robots and their external environment).

Since the stand belongs to high-tech innovative projects, its grounds and development have required the participation and coordination of efforts of a number of leading Russian organizations in the field of robotics with experience in working out technologies and ergonomic support for creating prototypes of robots and master devices: FSUE TsNIImash, JSC "Scientific and Production Association Android Technology", Scientific Research Institute of System Studies of RAS, Moscow State University named after Lomonosov.

VCS SRSs is a complex (Figure 7) including the following main elements: four computer workstations, two servers, a shared screen with a projector, a copy-type master device (CTMD) to control a physical robot or its 3-D model in a virtual environment, and a virtual reality glasses.

In order to control the RSs, special mathematical software has been created in the VCS SRSs, which makes it possible to investigate and pre-plan safe trajectories of movement of the kinematic links of robots and to simulate the performance of flight technological operations of cosmonauts in a virtual environment, and then execute it using a robot. This significantly reduces the risk of errors caused by the "human factor" and improves crew functioning.

VCS SRSs includes physical (full-scale) models, as well as software and hardware that imitate the operation of the RSs of various purposes and design in the “crew – manned spacecraft - environment” ergatic system with elements of virtual reality.

VCS SRSs is designated:

– to work out experimentally in-flight operations performed by the RSs with the participation of cosmonauts, in order to form well-grounded requirements for programs and methods of cosmonauts’ training for work with the RSs and assessing the quality of operator activity;

Figure 7 – Multipurpose Computer Stand of Robotic Systems

– to familiarize cosmonauts with the latest progress in the field of space robotics not only at the level of theoretical knowledge, but also at the level of forming eye-mindedness concepts and basic skills of interaction and work with space robots, as well as controlling the RSs in various operating conditions (forming a general operator culture);

– to study the principles and methods of monitoring the "behavioral" activity of robots to ensure flight safety;

– to assess the limits of the cosmonauts' functional capabilities for performing operations when working with the RSs of various technical designs and purposes;

– to study possible ways of building and using multimodal interfaces with "sensitive" robots, RSs control systems;

– to carry out with the participation of cosmonauts and specialists of the Center ergonomic examination of new technical solutions when creating RSs of various purposes in the interests of increasing the efficiency and safety of the activities of manned spacecraft crews;

– to form the baseline data for carrying out full-scale space experiments in terms of cosmonauts' interaction with RSs samples planned for delivery to the space station after preliminary studies on the stand, etc.

The stand allows to create and accumulate "databases" when performing technological operations, detail them by elements, and then execute them under operator control or autonomously. This may lead to subsequent control of the robots not only from the spacecraft, but also remotely, for example from Earth. Such robots will become indispensable human assistants in the exploration of the Moon and Mars and the construction of lunar (Martian) bases.

Automated workplaces (AWP) are intended for the RSs operator and specialists of the experimental support team. For the convenience of interaction of participants of the experiments, all workstations are located on the same line in close proximity to each other.

The RTS operator's workplace interface provides:

1) display of a three-dimensional scene including the RS, the working area and objects of interaction and surrounding interior;

2) scalability on screens of different resolutions.

The interface of the workplace of operational research control specialists provides:

1) display of the three-dimensional scene, identical to the image on the operator's helmet on the monitor and the shared screen;

2) display of current values of state vectors of mathematical models in the form of numerical data and in the form of graphs.

The communication interface with the copy- type master device provides:

1) transfer of command information;

2) data exchange with the master device;

3) conversion of the output values of the mathematical model of controls into values for issuing to the master device.

The VCS RS rests on the following principles:

1) openness and scalability. The modular software architecture allows for phased development and implementation, as well as allows for further expansion of both integrated data and information exchange objects and RS models;

2) data integration and consolidation. Integration of disparate data from various systems of the RS and its controls into an orderly unified database structure, which is a subject-oriented, integrated, chronological data set, constantly updated with new reliable information. This data set is a single source of noncontradictory and consistent data for all VCS SRS;

3) continuity. If necessary, it is possible to add and expand the range of functions, technical, informational and program components without affecting its functioning.

In January 2019 the Center conducted experimental studies to work out a number of tasks involving cosmonauts, when a cosmonaut controlled the robot remotely in a copy control mode (Figure 8). Analysis of their results has allowed putting forward proposals for improving the design and technical performance indicators of the software-hardware interface and virtual interactive 3-D model of the ARSs. The results of experimental studies allow us to accumulate proposals for updating the physical prototype of the ARSs and creating a higher level physical model of the ARSs suited for use onboard the ISS in the future.

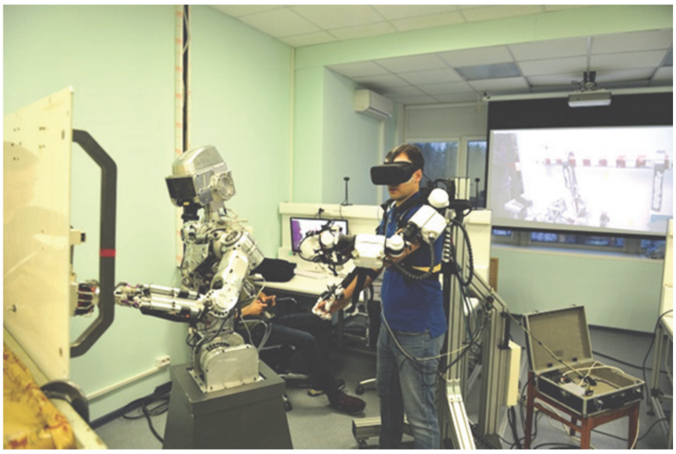


Figure 8 – Cosmonaut controls the RS in a copy-type mode on VCS SRS

The next stages of the VCS SRS improvement should involve studies on perfection of the design and technical and operational indicators of the software-hardware interface, the virtual interactive 3-D model of the ARSs and physical prototype of the ARS. It is also important to determine if 3-D model of the ARSs and the physical prototype of the ARSs are capable to perform the enlarged list of on-board flight operations. VCS STS allows to work out the interaction of cosmonauts with various space-dedicated RSs and to conduct ergonomic research both in a virtual environment and with physical prototypes of the space-dedicated RS. The use of VCS SRS is possible at all stages of on-ground training of cosmonauts.

At present, a research program is being formed aimed at meeting the requests of space robot designers regarding the cosmonaut's psycho-physiological capabilities during an on-planet activity and limitations due to the presence of the human factor.

Conclusions

Experience in the implementation of manned space programs shows that the development of technologies for the creation and application of robotic systems (such as: US – "Canadarm", "Canadarm-2", Japanese – JEMRMS, European – ERA) in order to carry out EVA and on-planet operations by cosmonauts should be considered as one of the promising directions of an innovative development in the field of manned space exploration. The transport-manipulation robots have found practical application in carrying out assembly and mounting works on the ISS, technological operations on recovery of the Hubble telescope, examination of ceramic tiles on the outer surface of Shuttles, and a number of other works.

Along with the use of transport-manipulation robots there are projects to create space anthropomorphous robotic systems designed for the future lunar missions. However, at the same time, to ensure an efficient and safe interaction of cosmonauts with the ARSs it is required to solve problems of the “advance” ergonomic designing of space robotics.

The designing and application of the computer-based stands of robotic systems of various purposes are very important practically for the experimental maturing of real ARSs in the early stages of designing the “ARSs – cosmonaut – environment” ergatic system, given the high cost of the creation and risks of the use of space-dedicated robots. The multipurpose computer-based stand of robotic systems at the Center is based on the virtual reality technologies and means of controlling motor and behavioral activity of a cosmonaut-operator and allows training cosmonauts for controlling various RSs when implementing anticipated scripts of promising manned space programs.

References

1. Kryuchkov B.I., Usov V.M., Yaropolov V.I., Sosyurka Yu.B., Troitskiy S.S., Dolgov P.P. On the features of professional activity of cosmonauts when implementing Lunar missions // Scientific journal “Manned Spaceflight” 2016. No 2 (19). pp. 35-58
2. NASA // NASA: official website. Available at: http://www.nasa.gov/, accessed 12.11.2016.
3. CSA // CSA: official website. Available at: http://www.asc-csa.gc.ca/eng/default.asp, accessed 12.11.2016.
4. E. Kaupp, E. Bains, R. Flores, G. Jorgensen, Y.M. Kuo, H. White Shuttle Robotic Arm // Engineering Innovations / P. 286 -- 301. Available at: http://www.nasa.gov/centers/johnson/pdf/584734main_Wings-ch4h-pgs286-301. pdf , accessed 12.11.2016.
5. B. Stockman, J. Boyle, J. Bacon International Space Station Systems Engineering Case Studyю Available at: http://spacese.spacegrant.org/uploads/images/ISS/ISS%20SE%20Case%20Study.pdf , accessed 12.11.2016.
6. The role of dexterous robotics in ongoing maintenance of the ISS / Lyndsey Poynter, P. Andrew Keenan // IAC-12,B3,4-B6.5,6,x16014.
7. Flexible robot manipulators: modelling, simulation and control. - (IET control series) – ISBN 978-0-86341-448-0
8. Kryuchkov B.I., Krikalyov S.K., Salaev A.M., Usov V.M. A man and a robot aboard a manned spacecraft // // Collection of scientific articles based on reports of the First Russian-German Seminar on Space Robotic, Germany, Stuttgart, 2012, non-legible.
9. Alpatov A.P., Belonozhko P.A., Belonozhko P.P., Kuzmina L.K., Tarasov S.V., Fokov A.A. Study of the dynamics of flexible space robots and prospect of their using // *Tekhnicheskayea mekhanika* [Engineering mechanics]. --2012. -- No 1. -- pp. 82 -- 93.
10. Yaskevich A.V., Ostroukhov L.N., Egorov S.N., Chernyshev I.E. Practicing the semi-realistic docking of the Russian module to the ISS using the remotely controlled SSRMS simulator // *Robotics and technical cybernetics.* 2013. No 1. pp. 53 58.
11. Karpov A.A., Kryuchkov B.I., Rogatkin D.A., Usov V.M. [Conceptual strategy for using service robots: common problems of implementation (On the example of manned space exploration and high-tech)]. *Biotekhnosfera – [Biotechnosphere]*. 2013. No.6. pp. 48-59
12. Alpatov A.P., Belonozhko P.A., Belonozhko P.P., GrigoryevS.V., Tarasov S.V., Fokov A.A. Modeling the dynamics of space manipulators on a movable platform // *Robotics and technical cybernetics*. 2013. No 1. pp. 59-65.
13. Lonchakov Yu.V., Sivolap V.A., Sokhin I.G., Sorokin I.G., Burdin B.V. Ergonomic research of interaction between cosmonauts and robot-assistants. // Ideas of K.E. Tsiolkovskiy in scientific and technical innovations / Proceedings of the 51st Tsiolkovskiy Scientific Readings – Kaluga. 2016 – 496p. ISBN 978-5-905849-46-6.
14. Burdin B.V., Dovzhenko V.A., The development of stands with VR-elements for ergonomic research of the «Operator – RS - Professional Environment» system // The ideas of K.E. Tsiolkovskiy in scientific and technical innovations / Proceeding of the 51 st Tsiolkovsky Scientific Readings – Kaluga. 2016 – 496 p
15. Sokhin I.G., Dovzhenko V.A., Burdin B.V., Grebenschikov A.V., Solvyova I.B. and others. Experimental ergonomic studies of an anthropomorphic robotic system controlled by cosmonauts while maintaining spacecraft and lunar infrastructure facilities. // Proceedings of the XI International Scientific and Practical Conference “Manned Space Flights”, November 10–12, 2015. – Gagarin Test& Research CTC, Star City, Moscow Region, 2015. pp. 31-33.
16. http://express-news.ru/index.php/news/science/item/2943-predstavlen-rossijskij-robot-kosmonavt-sar-401.
17. Burdin B.V., Mihayluk M.V., Sokhin I.G., Torgashev M.A. The use of virtual 3D models for experimental testing of the flight operations performed by means of anthropomorphic robots // Robotics and Technical Cybernetics, No 1. 2013, ISSN 2310-5305 – pp. 42-46.

M.V. Mikhailyuk [1], B.I. Kryuchkov [2], V.M. Usov [2,3]

VIRTUAL REALITY TOOLS FOR COMPUTER MODELING OF A COSMONAUT'S INTERACTION WITH A GROUP OF AUTONOMOUS MOBILE ROBOTS ON THE LUNAR SURFACE

[1] Federal State Institution "Scientific Research Institute for System Analysis of the Russian Academy of Sciences", Moscow, Russia, mix@niisi.ras.ru
[2] Federal State Budgetary Institution "Gagarin Research&Test Cosmonaut Training Center", Moscow region, Star City, Russia, bik43@mail.ru
[3] State Research Center of the Russian Federation - Institute for Biomedical Problems of the Russian Academy of Sciences, Moscow, Russia, khoper.1946@gmail.com

Abstract

The paper considers the long-range directions of creating the research simulating complexes (RSCs) using virtual reality technology (VR-technology). As applied to the lunar exploration projects, it is highly required to ensure the situational awareness (SA) of a human operator (HO) when controlling the movements of a mixed group of autonomous mobile robots (AMRs) on the lunar surface.

In the framework of the anthropocentric approach to making the human-robot interface (HRI) when organizing the feedback, it is proposed to use the combination of several types of visualization, as part of a human-robot interface (HRI) in order to improve the SA. They are: panoramic representation of movements of the AMR group; on-the-fly updated electronic maps (EMs) of lunar terrains; character-digital data from the simulated virtual sensors of on-board AMR systems; the images, obtained from the opto-electronic sensors of on-board ARM devices, etc. To this end, the choice of the basic RSC structure is justified, which is supplemented by VR-tools for building the AMR "virtual similitude" or "digital twins" (proxy-agents) in the virtual intellectual environment (VIE) and virtual sensors of AMR on-board systems. In order to support the cosmonauts in making a decision on potential AMR collisions, the graph representation of the task structure performed by AMRs is used, which allows the HO to manage the execution of tasks in the group, making choose between "initiation", "suspension" and/or "termination" in case of detection of the non-nominal situations.

Keywords: human operator (HO), group of autonomous mobile robots (AMR), human-robot interface (HRI), situational awareness (SA), research simulation complex (RSC), proxy-agents, virtual intelligent environment (VIE).

Introduction

The computer-based RSCs are widely used for ergonomic design and training of the HOs. The modern RSC has a complex multi-module structure and is constructed using VR-technologies. At first, the main idea of using the VR-tools for cosmonaut training was to visualize the interior of spacecraft and external environment, but as they improved, new areas of application were found. The VR-tools provide the HO with new opportunities to control objects (e.g. robots) in virtual reality while being in real physical world. More importantly, that for today many control devices and information display systems (IDSs) included in cosmonaut workplaces are transformed into the virtual interactive forms as part of RSCs. Typical configurations of modern RSCs, used VR-tools for modeling HRI, and the purpose of various modules are considered in [1].

The originally developed RSC scheme is shown in fig. 1.

The current state and development directions of RSCs for robotics groups

Experience in RSC building and the purpose of its primary subsystems

As it follows from numerous publications, RSCs are widely used for ergonomic designing the HRIs. Methods of "immersing a person” in the digital virtual interactive environment give a HO the opportunities to control artificial digital objects while staying in real physical world. In practical aspects, the virtual prototyping methods in the field of space robotics follow the experience in building space simulators for training cosmonauts to control spacecraft. This inheritance can be traced when constructing the control loops of dynamic objects and implementing the multiple interactive real-time feedback using the synthesized visual picture of the environment to improve the HO's SA.

Examples of such approaches to modeling and lab testing the HRIs are given in [2-4].

The bench for simulating the control of an anthropomorphic robot in a copying mode (or a "master-slave” mode) is presented in [2].

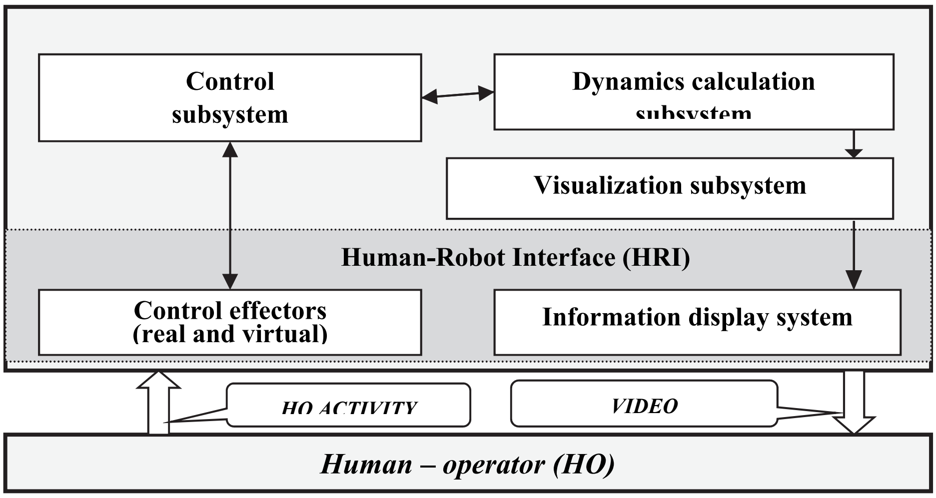


Figure 1 – The general scheme of the RSC

Testing of the telecontrolling the space robots are conducted during space experiments (SE) aboard the ISS. In particular, the "Kontur-2" SE is designed to study the remote control of ground-based robots from the board of the ISS [3].

The VR-tools for developing the prototype of a robot-manipulator for EVA are proposed in [4].

At the same time, it can be stated, that the interaction of the HO with a mixed group of lunar AMRs when the HO performs functions of the work coordinator and the "dispatcher" to prevent collisions of AMRs on the lunar surface is not enough investigated.

Consider the initial RSC scheme, shown in fig. 1, which was adapted to the task of modeling the control of an anthropomorphic robot [1, 2].

The initial RSC configuration consists of several subsystems [1]:

– "Virtual scene simulation".
– "Control".
– "Dynamics".
– "Visualization".

The "Virtual scene" is created in the 3DS MAX three-dimensional modeling system of the: 1) standard primitives (rectangular parallelepipeds, spheres, cones, etc.), available in the 3DS MAX system, and 2) special objects, added to this system (wheels, engines, particle systems, etc.), containing dynamic parameters (masses, inertia tensors, etc.). The "Modeling system" and added objects form the "Virtual scene designer".

The "Control subsystem" consists of virtual control panels and functional diagrams to calculate control signals. It receives at the input the digitalized actions of the HO to virtual control elements and generates control outputs that are transmitted to the "Dynamics subsystem". "Dynamics subsystem" loads the virtual model and calculates new positions and orientations of its objects according to their dynamic parameters in a period of the simulation time. This data are transmitted to the "Visualization subsystem", which synthesizes the virtual scene in the Information Display System (IDS) for the HO.

In the current version of RSC, developed by the Federal State Institution Scientific Center of the NIISI RAS (Moscow) using the original "GLView" software, the entire modeling cycle takes no more than 40 msec, what ensures operation of the RSC in real time. The fundamentally important feature of this modeling method is the use of the highest priority for transmitting the digitalized control actions of the HO to the "Control subsystem".

As a result, the adequacy and efficiency of the VR-methods have increased the accuracy of the robots' response to the HO's control actions. The same priority level can be assigned to the technical failure signal of AMR's onboard systems due to the impossibility of executing the HO's commands.

In most cases, using the visual feedback the HO can monitor the current situation and makes sure that the robot has processed the transmitted commands with success.

For mixed groups of AMRs, the situation changes dramatically, and the need to monitor the activity of the whole AMRs group in real time is accounted for the high level of uncertainty relative to the prediction of attendant risks when the AMRs move.

Opportunities for improving the HO's SA when interacting with the group of AMRs on the lunar surface

The experience in performing EVA on the ISS shows that the monitoring of the current situation performed by the crew members provides the possibility of timely intervention to correct the operation in the event of unscheduled EVA. Following this principle, it is very important to monitor joint activities of the cosmonaut crew and the mixed group of AMRs for ensuring the EVA safety. Thus, at the risk of collisions the necessity to improve the HO’s SA is recognized, since the joint execution of tasks with robots requires cosmonauts to know operational information and certain regulations of exchanging information about the tasks performing by AMRs [5].

When setting the task of monitoring the environment, it is assumed that crew members have a priori knowledge about the tasks to be performed during EVA on the lunar surface using the group of AMRs, about the priorities of the tasks to be performed by robots and the specific tasks for robots under the current planning of operations, in order to decide on resolving the "conflict of interest" in the group of AMRs in the access to common resource.

One of the possible ways to improve the HO’s SA consists in generating the virtual intellectual environment (VIE), as part of the RSC, which uses the Multi-agent model to simulate the movements of the group of AMRs and to transmit and receive data from active agents. Using virtual agents (so-called proxy-agents or "virtual understudies"), the behavior and interactions between the AMRs in a group is simulated not only in 3D visual representation of their movements but also in a graph form what allows monitoring the requests for access to shared resources from proxy-agent. This gives the opportunity to display for the HO the progress of the going tasks. On the basis of graph representations, the HO controls the group of AMRs at the level of understanding of their joint goals and assigned tasks in accordance with current priorities and available resources. Using the data from proxy-agents the HO makes the decision on resolving potential conflicts between AMRs.

The stages of designing the modes and RSC operation scheme are shown in fig. 2, which presents a fragment of ontology, reflecting the specific character of the HRI configuration when monitoring activity of the group of AMRs in terms of preventing collisions.

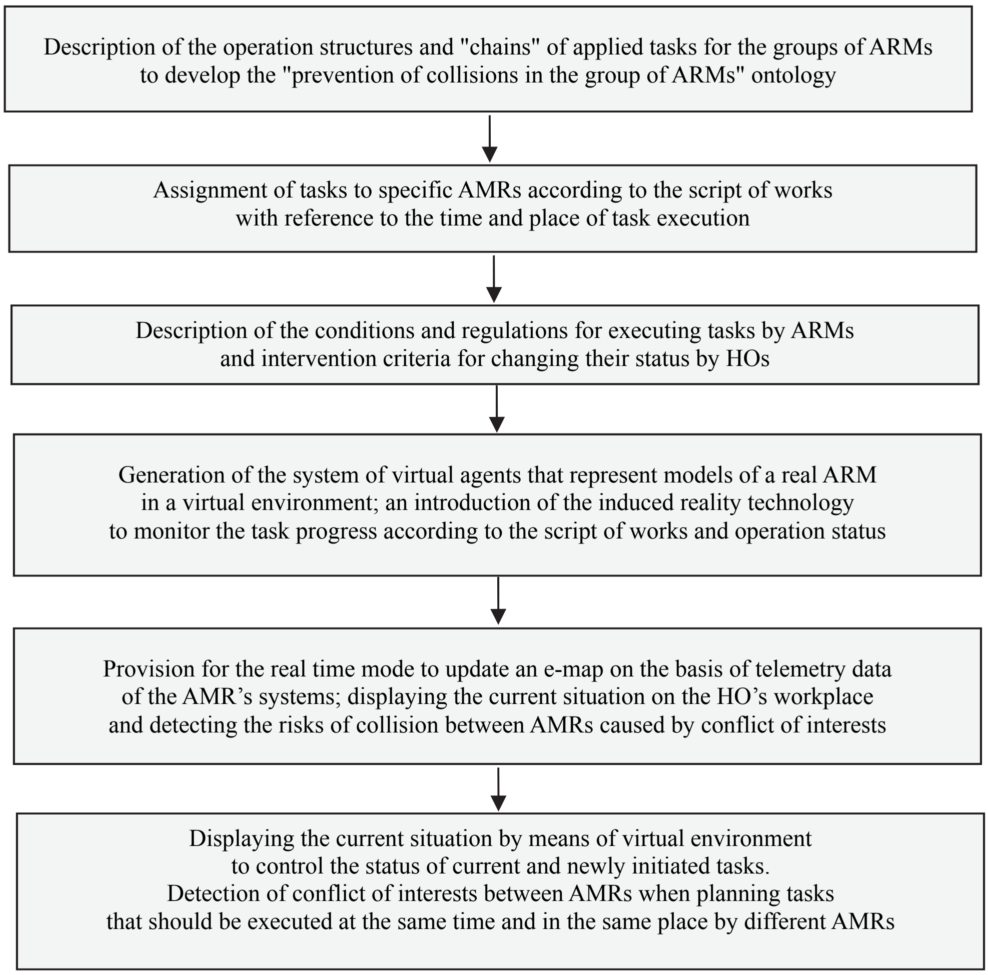


Figure 2 – Configuration of the HRI when monitoring the group of AMRs

In the general case, it is assumed that an AMR periodically moves along a closed route and transmits using proxy-agents the information about the current task, selected route, its own coordinates in the local coordinate system (fixed for areas of the electronic map) and the request for using the nearest territorial zone along the route considered in this case as a common resource that can be claimed by another AMR.

As far as the organization of the computing process is concerned, the above said assumptions require for transmitting signals on the independent operations of the AMR to the input of the "Control subsystem” to generate the control signals that are transmitted to the "Dynamics subsystem”.

It is assumed that AMRs are equipped with collision prevention systems when moving on the lunar surface. Embedded sensors scan the nearest area along the line of sight and, in the event of collision, interrupt the movement, and proxy-agents wait for command from the HO to continue an operation or to complete it.

In this case, the signals from the HO are transmitted to the input of the "Control subsystem" to control the execution of tasks (fig. 3).

Telemetry information from the AMR's onboard systems, images from onboard robot vision systems are also required as part of the monitoring system of the robotic group.

The RSC in the proposed extended configuration consists of a number of new subsystems (fig. 3).

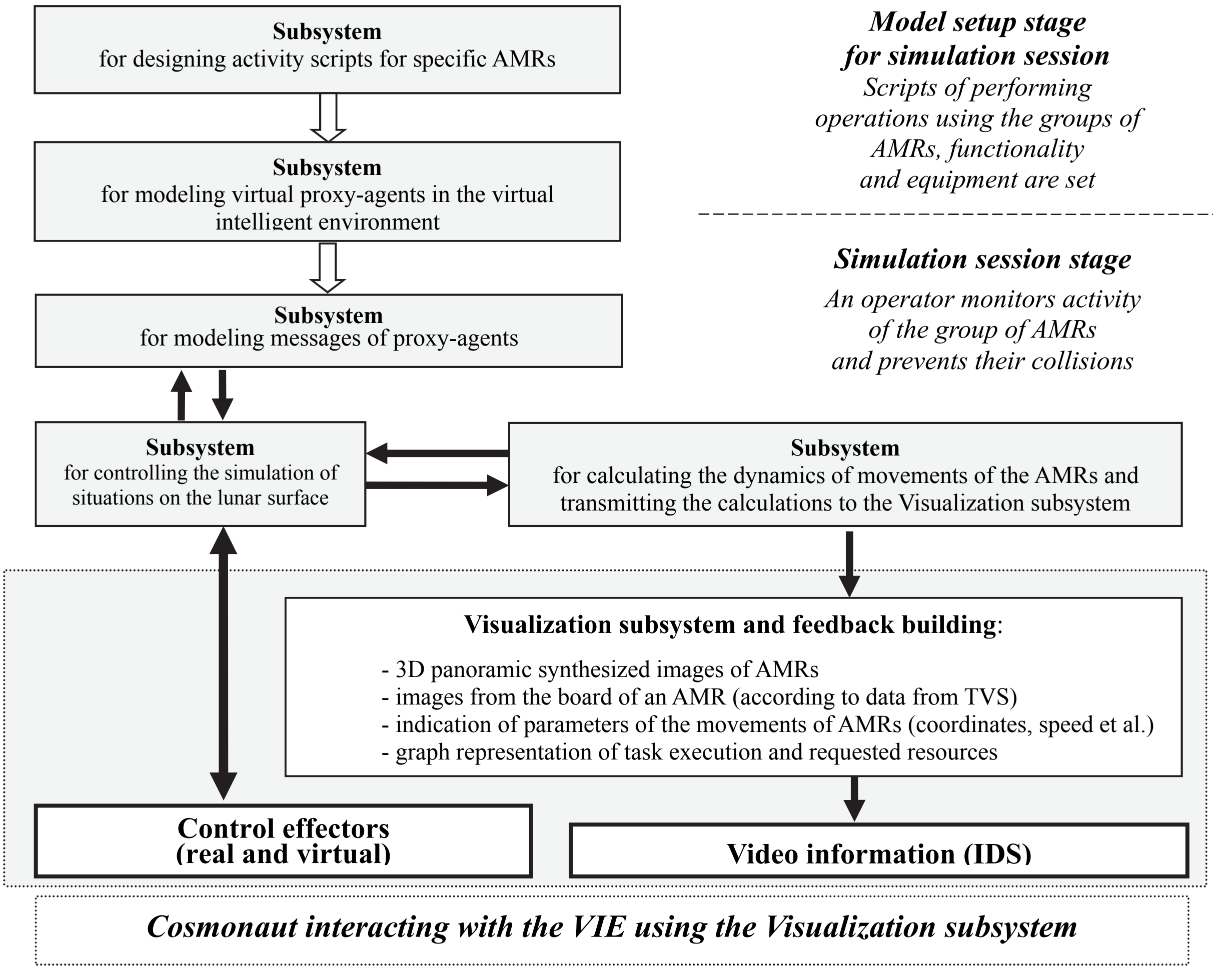


Figure 3 – Functional diagram for designing the modes and operation of the RSC

The subsystem for developing scenarios of the specific operations of AMRs includes defining their autonomous activities, taking the following into account:

1) The scope of tasks to be performed (in accordance with the purpose of AMR);
2) The Electronic Map (EM) of the areas on the lunar surface they move on;
3) The navigational landmarks marked on the EM;
4) The joint resources designed to perform specific tasks that may cause conflicts between AMRs on some routes.

Information about the route and movement parameters of the specific AMR is a priori contained in the mission plan and the routes are indicated on the EM. The nearest route segment is considered by the virtual agent as the resource for which the request is submitted to VIE. A conflict of interest which occurs when different virtual agents request the same resources should be resolved by the HO. In such situations while monitoring execution of the tasks, the HO can use the graph representation generated using data received from proxy-agents in the VIE. Detection of a conflict of interest at the level of requests for resources from proxy-agents gives grounds to suspend and activate the execution of tasks by the AMR, in this case the HO acts as a "Coordinator" when planning tasks and as a "Manager" when monitoring their execution [5].

In connection with the need for monitoring, additional modules are proposed for modeling the system of virtual sensors (VS) associated with AMRs and their proxy-agents.

Virtual sensors of the complex and their role in designing proxy-agents and VIE

In real physical control systems of complex dynamic objects there is a feedback, which is provided using sensors. The most commonly used sensors are: position sensors, angular and linear accelerations sensors, magnetometers, rangefinders, height sensors, and others. Each sensor transmits its readings to a real physical control system at a certain frequency; those data are used partly as information ones and partly for generating control actions.

This paper describes the approach of creating virtual sensors (VS) the readings of which are calculated by the simulating complex.

Sensors combined on the principle of being a part of the specific proxy-agent are added to the "Dynamics subsystem" and "Visualization subsystem" of the initial RSC structure. At every simulation stage, each sensor transmits its readings to the control system. These values are used in function diagrams to calculate control actions.

The paper deals with the following virtual sensors that are necessary for virtual representation of the AMRs in the VIE using a proxy-agent, they are: position sensor, orientation sensor, angular velocity sensor, accelerometer, touch sensor, force-torque sensor, rangefinder.

These virtual sensors are designed for the following:

– "Position sensor" determines the coordinates of the object's center in the local geographic coordinate system, which is set for the electronic map of an area;

– "Orientation sensor" calculates the angles of rotation of the Euler around the axes of the coordinate system;

– "Angular velocity sensor" determines the angular velocities of rotation of an object in its local coordinate system;

– "Accelerometer" calculates the acceleration of an object along its local coordinate axes, i.e. determines the acceleration vector of an object in its local coordinate system;

– "Touch sensor" detects the collision of an end-effectors with other objects of the virtual environment;

– "Force-torque sensor" measures the force and moment that are applied by any object of the environment;

– "Rangefinder" determines the distance to the nearest virtual object along the line of sight.

For each AMR, a set of sensors is configured so that the proxy-agent can promptly transmit to the VIE information about its location and the nearest surrounding objects within the sensor's tolerance for resolution, as well as warning signals if the sensors detect an unacceptable approach to another object.

In the proposed approach to building the RSC, the HO plays the key role in decision-making by coordinating the activity of robots. He initiates and interrupts the execution of current tasks, taking into account the priorities for executing predetermined plans of operations defined by a cosmonaut acting as a dispatcher.

This interpretation brings the interaction of an HO with a group of robots to a higher level of generalization in four main areas of building visual feedback:

1) The computer synthesis of a panoramic 3D visual pattern of movements of AMRs for monitoring the supervised area from the standpoint of the "external observer";

2) The presentation of an updated electronic map considering current indications received from virtual sensors of on-board AMR systems.

3) The graph representation of current tasks and possible conflicts of interest on the requested shared resources.

4) The images from the opto-electronic devices of a specific robot (and positioning of an observer in the coordinate system associated with this AMR, if necessary).

Thus, in ergonomics terms, an improvement of the SA can be achieved by combined methods of constructing DIS, including a panoramic display of the synthesized virtual scenes, by constructing and updating the EM with the indication of moving AMRs and stationary objects(as reference points of an information and navigation model), markers on overview displays, optical-television images, etc.

In the computer model, the panoramic representation of virtual scenes from the view of an external observer should be provided by several "virtual camera" (to follow the information technology of "external observer") with the multiport display on the monitors of IDS.

Conclusion

The problem of using robots during manned flights, associated with human factors, is largely determined by the quality of the HRI. Due to the complexity of the verification tests of the HDI under full-scale conditions, the paper proposes the method of designing and testing the HDI in lab experiments using existing RSCs.

From the methodological point of view, the possible solution of emerging interdisciplinary problems consists in a combination of several complementary concepts:

a) The robot-centric approach at the stage of designing a separately considered AMR in order to obtain the required functionality in its adaptability to extreme environmental factors and to the interaction in mixed "men-robots" groups;

b) The theory of multi-agent systems for VR-modeling the group of AMRs to ensure the coordinated movements of robots and to prevent conflict situations;

c) The anthropocentric approach to the interaction between operators and robots in the course of planned operations, implying the priority of human decision-making in the interests of entire mission safety.

The present-day stage of designing advanced robotic systems for use in extreme environments is characterized by the increased interest both to the improvement of the intelligence of individual autonomous robots and to the building of the VIE which ensures communication between intelligent agents, acting as "virtual counterparts" of real robotic products (existing in physical reality).

As a result of the study, the configuration and purpose of various types of sensors that generate data about the status of real physical environment to model activities of artificial agents were determined.

The proposed method and algorithms are implemented and tested in the VirSim complex, developed at the Federal State Institution Scientific Center of the NIISI RAS (Moscow). Obtained results together with the previously developed GLView software package and stored in the memory of the lunar surface electronic maps, allow visualizing the movement routes of the observed group in real time and displaying the motion parameters on the monitors of the information display system.

Thus, one can conclude that there is a possibility to scale the structure of the previously used computing platform and original VR-tools for creating the RSC in relation to the ergonomic design of the HRI for the “cosmonaut – mixed group of AMRs” team what is the continuation of previously obtaining solutions.

References

1. Mikhailyuk M.V., Bragin V.I. Virtual reality technologies in simulator-training complexes for the training of cosmonauts // Piloted spaceflights. 2013. № 2 (7). Pp. 82-93.
2. Sokhin I.G., Kuritsyn A.A., Usov V.M. Problems of interaction of the crews of promising space missions with anthropomorphic helper robots // The human factor in complex technical systems and environments: ERGO 2018. Coll. works. III International conf. ERGO 2018 St. Petersburg. July 4-7, 2018 - St. Petersburg: LETI Saint-Petersburg Electrotechnical University. (789 p.) P. 782-289.
3. Zaborovsky V.S., Kondratiev A.S., Mulyukha V.A., Silinenko A.V., Ilyashenko A.S., Filippov M.S. The remote control of robotic objects in space experiments of the “Contour” series // Scientific and technical statements of SPb GPU. Computer science. Telecommunications. Control. 2012. № 6 (162). C.23-32.
4. Sergeev A.V., Guk M.Yu. Mobile space robot control with the use of virtual reality and power moment sensation. // Piloted spaceflights. 2018. №4. P.44-52.
5. Kryuchkov BI, Usov V.M., Ivanko D.V. Prospects for the use of intelligent spaces for the information support of the operator in the remote monitoring of a group of mobile robots on the lunar surface. // Proceedings of the conference "Information Technologies in Management" (ITU-2018). St. Petersburg: Concern Central Research Institute Elektropribor, 2018. (742 p.) P. 678-687.

N.S. Slobodzyan

METHODS OF IMPROVING HEXAPOD' LINEAR ACTUATORS ACCURACY IN SPACE APPLICATION

Baltic State Technical University «VOENMEH» named after D.F. Ustinov, St. Petersburg, Russia
ja-nikita@mail.ru

Abstract

The article presents analysis of specific functioning features of hexapod-type mechanisms with parallel kinematics under space conditions. Comparative characteristics of the most widespread linear movement sensor types are given. Schematic structure of hexapod' linear actuator with indirect position feedback is presented. The principal sources of errors of linear actuator rod movement are described. Methods of error compensation for ball screw and linear thermal expansion of the actuator elements' size.

Keywords: hexapod, Gough-Stewart Platform, linear, actuator, compensation, ball screw, gear, linear, thermal, expansion, calibration.

Acknowledgments

This work was carried out in accordance with the decree of the Government of the Russian Federation dated 09.04.2010 No. 218. Research, development and technological work is carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation (contract dated 01.12.2015 No. 02.G25.31.0160). The work was carried out in the BSTU «VOENMEH» named after D. F. Ustinov [10].

Robotic drive systems for kinematic arrangement are divided into two large groups characterized by serial or parallel chains connection.

Serial connection mechanisms have open-loop kinematic chain, and each chain link bears weight of the subsequent part including the payload weight (processing tool, object, etc.) thus subjecting the chains to high bending torques effect. To maintain accuracy of the discussed mechanisms functioning, improvement of the structure stiffness or introduction of additional load relief elements is required, which inevitably results in the mechanisms' mass and size increase.

On the contrary, in parallel structure mechanisms (PSM) load is distributed between all the chains, so high stiffness, precision and load capacity are assigned to the PSM' advantages. However, PSM have some drawbacks, such as limited workspace, as well as complexity and – in general case – ambiguity of analytical solutions for kinematic and dynamic problems at control algorithm synthesis [1, 2, etc.].

There are parallel structure mechanisms having various number of degrees of freedom of mobile platform. Of special interest are mechanisms of hexapod type providing six degrees of freedom: three rotational (Euler angles) and three translational (cartesian coordinates) ones.

The initially proposed mechanism of the said type was V. Gough Platform for testing of a car' wheel-hub assembly at any angle to the surface [3]; then, D. Stewart article [4] was published describing a pilot trainer mechanism. Later, the said mechanism, also called Gough-Stewart Platform, was studied by many other researchers.

Hexapod (See Figure 1) consists of the fixed baseplate, movable platform and connecting linear actuators (LA). The latter are fixed to the platform by joints.

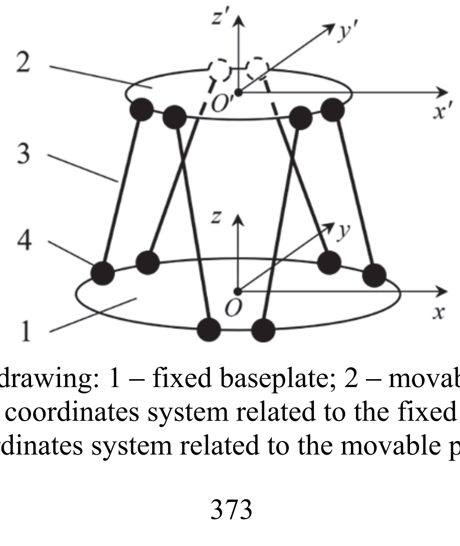


Figure 1 – Hexapod schematic drawing: 1 – fixed baseplate; 2 – movable platform; 3 – linear actuator; 4 – joint; *Oxyz* – coordinates system related to the fixed baseplate center; *O'x'y'z'* – coordinates system related to the movable platform center

Hexapod ensures transfer of an object situated on its movable platform according to six degrees of freedom relative to its baseplate by changing linear sizes of its 'legs' with the help of linear actuators. At that, depending on the version of hexapod' kinematic diagram plotting, joints and linear actuators having their own various degrees of freedom can be used [5].

Hexapods are used for solving the task of objects positioning in various fields: machinery elements' test systems, precision measuring machines, machine-robots and manipulators of various applications, as well as guidance and orientation systems for board devices placed on the non-stabilized baseplate.

An advanced application of PSM is its use in space technology to solve the task of precision devices' guidance and orientation, as well as their stabilization and vibration isolation [6]. Russian specialists are considering possibility of developing a multi-sectional manipulator to ensure navigation management of the Millimetron Orbital Astrophysical Observatory [7].

At that, use of general-purpose industrial grade hexapod in space technology is considered to be impossible due to abnormal operating conditions. Requirements to space application devices include ability to operate in wide temperature range and vacuum environment. Besides, there are special requirements to the components related to resistance to the integrated radiation and total energy of heavy ions [8].

In the open sources, hexapods of the dedicated space application are not presented. For instance, hexapods manufactured by Symetrie intended for operation in vacuum and fabricated from materials with low thermal expansion coefficients [9] cannot function in space since their components are of general-purpose industrial grade and are not resistant to ionizing radiation. So, development of a parallel structure mechanism intended for space application became the vital objective and was solved within the range of Research, Development & Technology Program [10].

In view of special features of the discussed hexapod' space application, problems related to the mechanism components choice and to the possibility of ensuring its functioning in the specific environment are being faced, such as wide range of working temperature levels, ionizing radiation, vacuum, etc. [11]. As to the theory of the parallel structure mechanisms' positioning, trajectory and contour control, it has been rather completely described in literature including studies performed within the range of Research, Development & Technology Program [1, 2, 12, etc.].

The main principle of designing high quality system of an object' positioning and orientation is its position feedback control. For the discussed guidance devices, it includes three angular and three linear coordinates.

In the systems of guiding radio antennae, optical devices and other monitoring instruments, the object tracking principle is used wherein automatic adjustment of the guidance device' positioning and orientation based on the received information (radio frequency signal power, optical image).

However, the above method is inadequate for monitoring such complex objects as stars, black holes, galaxies, etc., and the task of the guidance device' positioning and orientation tends to revolve around the task of positioning and orientation of the hexapod' movable platform. Below, possible scenarios of creating the platform position' measuring system are discussed.

For spacecraft orientation, the principle of point sources (stars and other radiating objects) sighting. Orientation parameters are calculated by comparison of the observed image of the sky of stars area vs the data of the star catalogue stored in the on-board computer memory. In spite of high-precision determination of the order of angular seconds' units and fractions for angular attitude [13], use of such sensors for measuring linear translations within the range of the discussed PSM movable platform does not seem essentially possible. The said sensor sizes can exceed the object of guidance' dimensions, and error measurements of linear translations make up kilometers.

To measure the platform' movements, it is possible to use combined spatial system of inertial sensors, viz. gyroscopes and accelerometers. However, even modern microelectromechanical measuring systems are not able of ensuring high-precision control of the movable platform and are subject to substantial drifts in time and temperature readings [14].

It is possible to design the hexapod' measuring system as a series connection of linear and angular sensors in analogy with the design of coordinate-measuring machines with serial kinematics. However, considering the serial design, the said machines accumulate significant errors that by several times exceed requirements to accuracy of the hexapod movable platform' positioning and orientation. Moreover, serial mechanism' kinematic scheme making possible measuring any combination of six angular and linear coordinates of the movable platform will be non-rigid and rather cumbersome.

As an alternative version, the hexapod' measuring system can be designed according to six separate linear translation sensors mounted parallel to each other between the baseplate and the platform. In this case, the so called ‘spatial sensor of position’ is created [15]. A measuring system’ kinematic structure similar to the most multi-stage mechanism is the optimum version for multi-chain parallel structure mechanism.

In this case, the highest control precision can be achieved at the expense of supplying the hexapod with the opto-electronic system of measuring the platform’ spatial position relative to the baseplate. At that, precision of the platform position measurement is improved at the expense of separating the hexapod’ power and measuring subsystems. Based on the measured data, calculation of the hexapod legs’ actual length is performed by way of solving the inverse kinematic problem, and the produced data are entered into the hexapod control system as feedback signals. In the above scheme, measuring sensors can be either strictly bound to the actuator system (built into the linear actuators), or be randomly positioned relative to the linear actuator (LA), but the said linear sensors number should be equal to the number of the mechanism’ degrees of freedom.

Spacecraft instrumentation and devices operate in the extremely aggressive environmental conditions, which imposes certain restrictions on the components used. First and foremost, there are strict functioning requirements applied to the sensors including resistance to ionizing radiation, vacuum, wide range of temperature levels, as well as their reliable functioning. Moreover, a spacecraft and its payload should be of minimum mass and size aimed at cost saving for placing payload into orbit by launch vehicle.

Sensors of various functional principles can be used as linear sensors of position. Potentiometers ensure information retrieval by the LA control system (typically, analog-to-digital conversion is used), but such essential faults as low precision and poor reliability are typical for them. Along with high reliability and resistance to external factors effect, inductive and capacitance sensors demonstrate poor accuracy. Interferometric sensors allow measuring position with accuracy to micrometer fractions, though have a complex structure with expensive components and are completed with the specialized opto-electronic system for information processing. However, the above types of sensors having the precision accuracy required for space application PSM are produced worldwide in single quantities, are very expensive and require, at least, stable temperature environment.

It should be noted that translation accuracy of the object fixed onto the hexapod' movable platform depends both on manufacturing errors of the mechanism and on the adequate consideration of all the kinematic parameters entered into the mechanism control for solving the inverse kinematic problem. Moreover, measurement readings for six linear translations are not enough for accurate defining of the hexapod platform position since linear actuators, movable platform itself, and hexapod’ baseplate are subject to thermal deformation. A temperature change results in displacement of fixing point of all the joints including fine joints in the measuring system. To improve the hexapod accuracy, it is possible to fit its basement and platform with additional temperature sensors, thus making possible to carry out measurements and further check of thermal deformation.

Still it is accurate functioning of linear actuators that influences precise translation of the output link, i.e. movable platform with the object.

Considering disadvantages of the above-described approaches to designing the measuring system of the movable platform’ position and orientation and modular principle applied for the discussed PSM measuring system creation, the most optimum strategy is arrangement of the linear actuators functioning as closed-loop actuation devices, or servo drive. Set of requirements to the essential ranges of positioning and orientation of the hexapod’ movable platform, as well as requirements to tens of micrometers and tens of angular seconds accuracy of the platform’ translation and rotational movement, result in the need for the linear actuator rod translation with ±2 micrometers accuracy at travel range making up to 200 mm [16].

There are various kinematic schemes and design concepts on the hexapod’ linear actuator arrangement. The actuating part of the discussed hexapod’ linear actuator is arranged according to the typical scheme: stepper motor – gearbox – ball screw. Possible scenarios of the actuating part of the hexapod’ linear actuator arrangement are as follows.

The principal requirement to the linear position actuator is the minimal absolute error of measurement. We assume that to achieve precise travel of the linear actuator rod, error in linear position measurement should not exceed 50% of the acceptable error for the LA rod travel, i.e. ±1 μm. The desired range of translations measured by the sensor equals the required linear actuator travel and makes 200 mm.

Just as in case with separate measuring system of the movable platform’ position and orientation, extreme environmental conditions of a spacecraft equipment operation impose a number of restrictions on the linear actuator components as well.

Strict functioning requirements are applied to the LA information-measuring system, such as ionizing radiation resistance, ability to function in vacuum and at wide temperature levels range (±80 °C), resistance to electromagnetic disturbances and severe mechanical disturbances at payload placing in orbit.

The elementary sensor and its processing system should not be expensive and should be of minimum mass and size for the sake of cost saving of the spacecraft placing into orbit. Depending on the elementary sensor type, signal processing and its useful component extraction can be realized either in rather simple way (filters on passive elements, operating amplifiers, analogue-digital converters), or by using high frequency schemes of extraction and conversion, optical modulators, standard measures, and other specialized, complex and expensive devices.

Based on the above-mentioned requirements Table 1 sums up specifications [17-19, etc.] of the most common types of linear travel sensors. The sensors' typical absolute error in measurement is given for measurements range of 200 mm on condition of the elementary sensors and processing system operation at normal temperature level with negligible permissible deviation range.

Table 1 Specifications of the most common types of linear travel sensors

Type of linear position sensor	Typical absolute error, μm	Measured travel range, mm	Simple scheme of measurement	Availability of ionizing radiation resistance solutions	Insensitivity to electromagnetic disturbances	Insensitivity to mechanical disturbances	Small size	Low cost
eddy current	±100	10						
capacitance	±10	10	+	+			+	+
inductive	±30	> 200	+	+	+	+	+	+
interferometric	±0.2	> 200			+			
magnetoresistive	±400	> 200						+
magnetostrictive	±20	> 200			+			
optical	±0.6	> 200			+			
potentiometric	±500	> 200	+	+	+		+	+
strain gauge	±400	> 200	+	+	+		+	+
ultrasound	±1600	> 200			+	+		
based on Hall effect	±20	> 200				+	+	

Based on the Table, it is evident that the desired absolute error in measurement is ensured by the interferometric and optical sensors, however no solutions for their space application on commercial scale are available. Thus, comparative analysis results reveal the principal challenge of PSM linear actuators' precise control of in space technology: in the absence of possibility to use linear position sensor as the principal feedback element arrangement of precise LA rod linear travel is required.

Let us discuss possible implementation of a linear actuator with position indirect feedback [11] (See Figure 2). The linear actuator' drive part includes a motor, gearbox, ball screw (BS), as well as a rod fixed on the transmission' movable element.

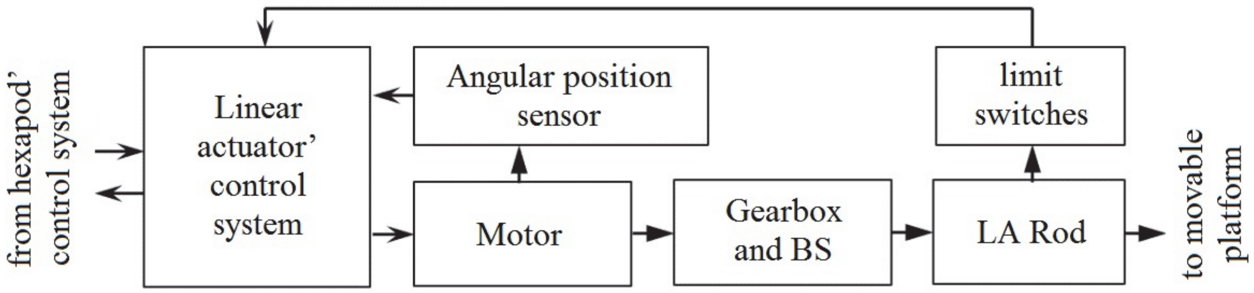


Figure 2 – Schematic diagram of the linear actuator with position indirect feedback

The LA travel measuring part is composed by the motor shaft' angular position sensor (APS), as well as by the limit switches of the LA rod' linear position. So, indirect sensor of linear position measuring is used in the discussed LA. APS ensures getting the local feedback generating information on angular position of the motor shaft within one rotation, the said information being converted into the rod linear position l by way of recalculation.

Similar to the case with the linear position sensors (See Table 1), various types of sensors can be used as APS. However, angular position inductive sensors converting nonelectric value measured in the induction EMF have gained the most widespread use for actuators operating in extreme environment. Among the said sensors, there is a rotary transformer (resolver) intended for conversion of angle of rotation into two electric signals, their amplitude being the angle' sine and cosine function, respectively. Joint processing of the said two signals allows getting absolute value of the angle of rotation. Resolvers can be of both one pair of windings and dual pairs of windings types. Availability of dual pairs of windings improves accuracy of angular position determination. Owing to reliability, resistance to electromagnetic disturbances, mechanical effect and ionizing radiation, resolvers are widely used in aerospace, defense and petrochemical industries. Both home-produced and foreign-manufactured resolvers are available for space application. As limit switches, use of either inductive sensors or specialized optocouplers is possible.

At feedback control, accuracy of the linear actuator rod travel is influenced by measuring error of the applied measurement system and by the actuator' total sensitivity to control signals. So, at designing, measuring and actuating elements of the actuator are selected so that the preset accuracy of the rod travel be ensured in the course of operation according to the feedback signal. As mentioned above, development of the space-application LA with the rod position measuring system is associated with significant technical, technological and economic challenges. That is why despite of all the obvious advantages of the above measuring system, factors influencing travel precision of the rod in linear actuator without its position feedback, as well as methods of compensation of the most important error sources shall be discussed.

In the linear actuator in quasistatic mode the following sources of the rod travel accumulative error are defined: geometrical errors; temperature errors; errors caused by mechanical loads; control errors; other errors.

Geometric errors mean interrelation of the LA mechanism' angles and their relative movement caused by manufacturing and/or assembling faults.

Temperature errors occur due to thermal expansion, shrinkage or distortion of LA elements owing to the effect of heat sources, such as environment, electric motor, friction in the ball screw and bearings, etc.

Errors caused by mechanical loads include mechanism structure distortion under the effect of static load in the Earth' gravitational field (at operating in earth-bound environment), as well as under the effect of inertial forces at the acceleration motion of the mechanism' elements.

Control errors occur at performance of measurements by sensors or at assessment of position in the sensor-free system, at recalculation of the measured or assessed values by the control device, at control signals generation by the device, at signal transmission, etc.

Other errors are those related to instability of the LA fabrication materials geometry, to friction forces change, as well as the ones caused by external disturbances (vibrations, etc.).

The most important error sources as per their share in LA rod travel accumulated error are such as ball screw errors (mismatching of the screw' angular position and the transmission nut' linear travel) and linear thermal changes of the LA elements size.

In general, there are two ways of improving accuracy at actuating mechanisms designing, namely: elimination of error sources and compensation of errors occurring at functioning. The first way envisages taking special actions in the process of designing and manufacturing aimed at complete elimination or minimization of errors. However, it is impossible to get rid of errors using this approach, so compensation measures are to be additionally undertaken. Below, ways of compensation of the above-mentioned most important error sources are discussed.

Figure 3 demonstrates the chart of LA rod travel error relative to its current linear position [11] caused by inaccurate functioning of the ball screw that, in turn, is the result of limited manufacture precision.

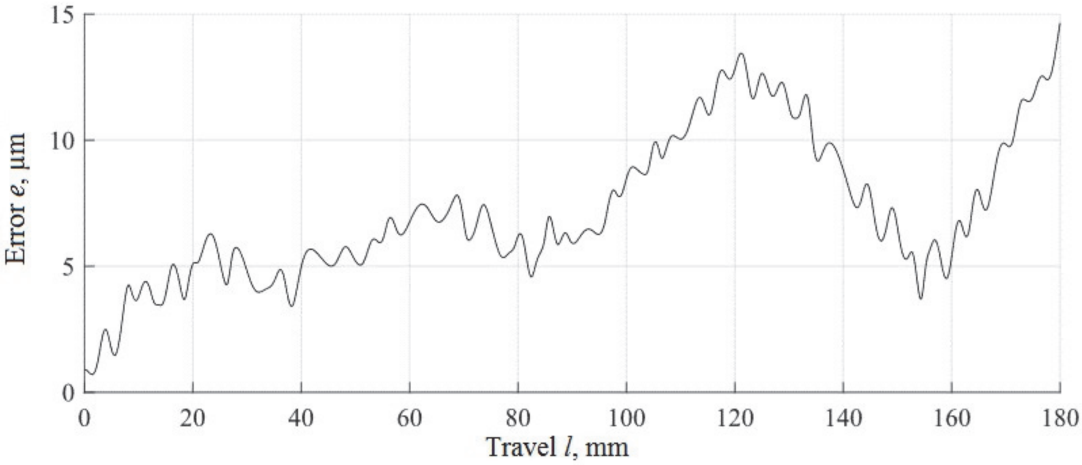


Figure 3 – Value of error in LA rod travel relative to its current linear position

According to the chart, error e amounts to 15 μm and more, which markedly exceeds maximum permissible value of error, i.e. 2 μm. Compensation of the said error can be achieved at the control signals generation preceded by the linear actuator calibration [11].

Based on Figure 3, it is evident that dependency of the LA rod travel error e relative to the current linear position l is a non-linear one. At discrete travel with 1 μm step and 200 mm slide, it is necessary to store 200,000 correction values in memory, which might be not feasible in case of implementation of the LA control system on the basis of microcontroller with limited memory capacity.

The most optimum scenario is to store incomplete correction data array, i.e. only some discrete set of n nodal points $\{e_{d1}, e_{d2}, \dots, e_{d(n-1)}, e_{dn}\}$ making possible finding intermediate correction values. In mathematics, the task is called interpolation, and has a great number of solutions for the general case.

It seems to be impossible to describe the above dependency $e(l)$ by one common function (i.e. to perform global interpolation), so local piecewise interpolation should be considered. At that, a separate interpolation function is specified for each interval $\left[e_{d(i)}, e_{d(i+1)}\right]$.

Out of the multiple variants of interpolation (piecewise-constant, piecewise-linear, piecewise-quadratic, spline-interpolation, etc.) use of the piecewise-linear interpolation is believed to be the most optimum solution for correlation of requirements to the memory capacity and to the computational resources.

At this type of interpolation, the nodal points are connected by line segments, i.e. is linear polynomial $e(l) = a_0 + a_1 l$ is plotted at each two points interval $\left(l_{d(i)}, e_{d(i)}\right)$ and $\left(l_{d(i+1)}, e_{d(i+1)}\right)$, where $l_{d(i)} \le l \le l_{d(i+1)}$. Coefficients are defined according to the equations

$$a_0 = e\left(l_{d(i-1)}\right) - a_1 l_{d(i-1)}, \tag{1}$$

$$a_1 = \frac{e\left(l_{d(i)}\right) - e\left(l_{d(i-1)}\right)}{l_{d(i)} - l_{d(i-1)}}. \tag{2}$$

Figures 4 and 5 give charts of, respectively, corrections and errors of the piecewise-linear interpolation with 0.5 mm step. According to Figure 5, it is evident that maximum error value occurring at the piecewise-linear interpolation is substantially lower than the permissible rod linear travel error, while using the above described method to compensate the error source, i.e. ball screw allows increasing accuracy of the linear actuator operation.

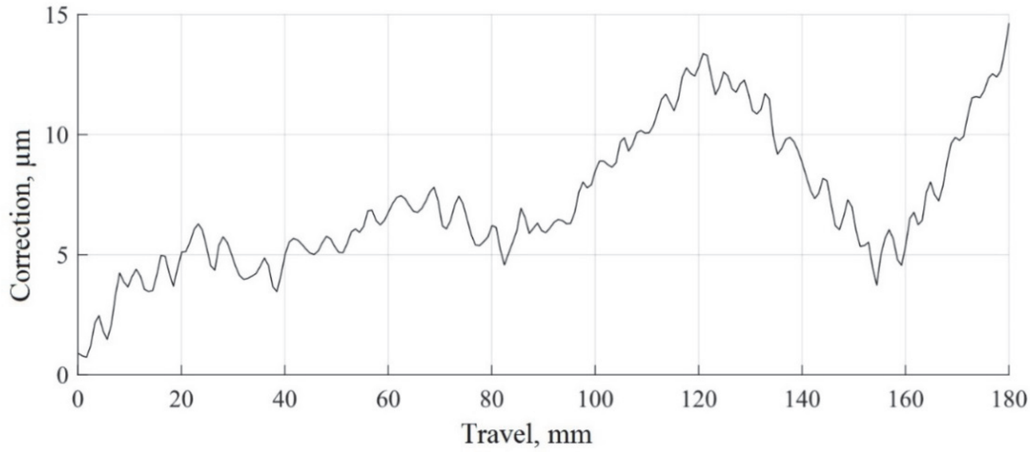


Figure 4 – Correction values at piecewise-linear interpolation

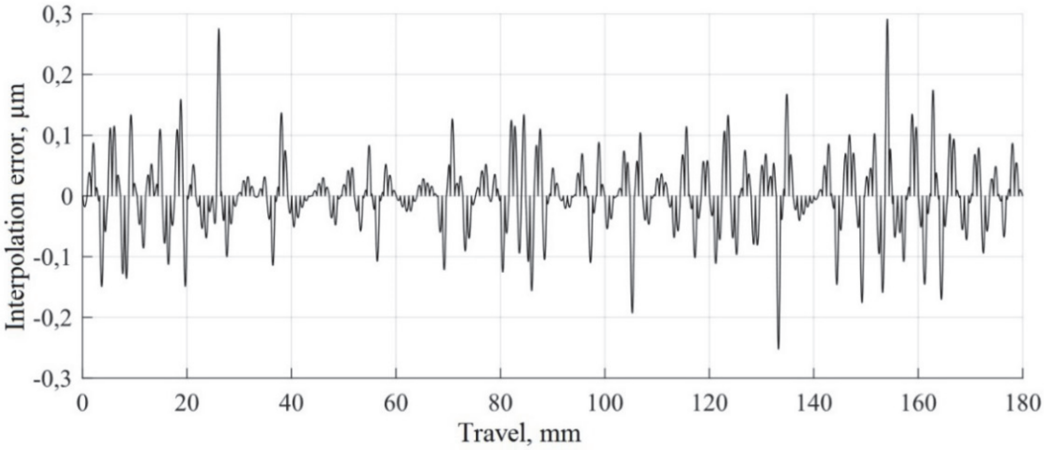


Figure 5 – Interpolation error values at piecewise-linear interpolation

In the presence of linear actuator' wide-range temperature fluctuations, temperature changes of the rod position can reach hundreds micrometers [11], which is inadmissible at the preset maximum accuracy of 2 µm. Let us discuss method of this error source compensation.

A body length Δl change at the expense of linear thermal expansion is known to be expressed by the formula

$$\Delta l = \Delta T \cdot \alpha \cdot l, \tag{3}$$

where α is a body' coefficient of linear thermal expansion (CLTE); l is a body' initial length at T_0; ΔT is a temperature level change relative to T_0.

A linear actuator consists of n components, each having nominal length in the axial direction l_i defined for $T_0 = 22\,°C$, as well as coefficient of linear thermal expansion α_i. The linear position' total thermal change Δl_T shall be described by the expression

$$\Delta l_T = \Delta T \cdot \sum_{i=1}^{n} \alpha_i l_i \quad . \tag{4}$$

On the screw, BS' slide nut is in the position l_{BS} that is changed depending on the preset linear travel of the rod. For the discussed LA scheme, length l_i of the rest linear actuator' components are not changed in the course of operation. So, the expression (4) can be converted as follows:

$$\Delta l_T = \Delta T\left(\alpha_{BS} l_{BS} + \sum_{i=1}^{n-1} \alpha_i l_i\right) = \Delta T\left(\alpha_{BS} l_{BS} + K_{LA}\right), \tag{5}$$

where α_{BS} is the coefficient of BS' linear thermal expansion; $K_{LA} = \sum_{i=1}^{n-1} \alpha_i l_i$ is coefficient accounting for the linear thermal change of size of all linear actuator elements with the exception of BS.

It is possible to calculate such BS nut' linear position l_{BS} wherein the resulting linear position l_T of the LA joint' external plane relative to its position at $T_{LA} = 22\ °C$ and $l_{BS} = 0$ mm will be equal (considering the thermal change of sizes) to the preset one: $l_T = L$.

Dependency between the above said values will be expressed as follows:

$$L = l_T = \Delta l_T + l_{BS}. \tag{6}$$

Based on the above and considering (5), the following expression is produced: $L = \Delta T\left(\alpha_{BS} l_{BS} + K_{LA}\right) + l_{BS}$.

Then, $L = l_{BS}\left(1 + \Delta T \cdot \alpha_{BS}\right) + \Delta T \cdot K_{LA}$, hence:

$$l_{BS} = \frac{L - \Delta T \cdot K_{LA}}{1 + \Delta T \cdot \alpha_{BS}}. \tag{7}$$

It should be noted that in the expression (7) constant coefficients α_{BS} и K_{LA} are used. They can be defined experimentally, which would ensure higher precision as compared to the calculation method.

For this, at the known and constant temperature level T_{LA} (though different from the temperature of defining LA' nominal sizes $T_0 = 22\ °C$), error values Δl_T are defined at two extreme positions of the BS nut $l_{BS0} = 0$ мм и $l_{BS200} = 200$ мм ($\Delta T = T_{LA} - T_0$):

$$\Delta l_{T,l_{BS0}} = \Delta T\left(\alpha_{BS} l_{BS0} + K_{LA}\right), \tag{8}$$

$$\Delta l_{T,l_{BS200}} = \Delta T\left(\alpha_{BS} l_{BS200} + K_{LA}\right), \tag{9}$$

By subtracting (8) from (9) and expressing α_{BS}, we get:

$$\alpha_{BS} = \frac{\Delta l_{T,l_{BS\,200}} - \Delta l_{T,l_{BS\,0}}}{\Delta T\left(l_{BS200} - l_{BS0}\right)}. \tag{10}$$

Similarly, it is possible to experimentally define value K_{LA}. For this, at the known and constant temperature level T_{LA} (though different from the temperature of defining LA' nominal sizes $T_0 = 22\ °C$) error value Δl_T is defined Δl_T at BS nut position $l_{BS} = 0$ mm. Then,

$$\Delta l_T = \Delta T\left(\alpha_{BS} l_{BS0} + K_{LA}\right) = \Delta T \cdot K_{LA}, \tag{11}$$

$$K_{LA} = \frac{\Delta l_T}{\Delta T}. \tag{12}$$

Thus, compensation of the sizes' thermal change at the linear actuator control is performed. Combination of two methods of errors compensation presented in the article allows substantially improve the linear actuator precise operation without using precision linear sensor.

References

1. Lebret G., Liu K., Lewis F. Dynamic analysis and control of a Stewart platform manipulator // Journal of Robotic System. 1993. Vol. 10. Issue 5. Pp. 629-655. https://doi.org/10.1002/rob.4620100506
2. Zhukov Yu. A., Lychagin Yu. V., Slobodzyan N.S. Hexapod kinematics real-time problem solving // Intellektual'nye sistemy, upravlenie i mekhatronika – 2017. Materialy III Vserossijskoj nauchno-

tekhnicheskoj konferencii molodyh uchenyh, aspirantov i studentov (Russia). Science editor A.T. Barabanov. 2017. Pp. 87-91. (In Russ.)
3. Gough V.E. Contribution to discussion of papers on research in automobile stability, control and tyre performance // Proceedings of the Institution of mechanical engineers (Automotive Division). 1956-1957. Pp. 392-394.
4. Stewart D. A platform with six degrees of freedom // Proceedings of the Institution of mechanical engineers. 1965. Vol. 180. Pt. 1. No. 15. Pp. 371-385. https://doi.org/10.1243/PIME_PROC_1965_180_029_02.
5. Weck M., Staimer D. Parallel Kinematic Machine Tools – Current State and Future Potentials // Cirp Annals-manufacturing Technology. 2002. Vol. 51. Issue 2. Pp. 671-683. https://doi.org/10.1016/S0007-8506(07)61706-5
6. Liang D., Lu A., Wu and R. Dong. Active vibration isolation for stewart platforms using adaptive backstepping sliding mode control // 37th Chinese Control Conference (CCC). Wuhan. 2018. Pp. 990-994. https://doi.org/10.23919/ChiCC.2018.8483664K.
7. Artemenko Y.N., Karpenko A.P., Volkomorov S.V. Capability of Parallel Mechanism Used for Attitude Control of "Millimetron" Space Telescope // Science and Education: Scientific Publication (Russia). 2014. Vol. 11. Pp. 357-370. (In Russ.) https://doi.org/10.7463/1114.0740118.
8. International Space Station Researcher's Guide. Available at: https://www.nasa.gov/sites/default/files/files/NP-2015-03-015-JSC_Space_Environment-ISS-Mini-Book-2015-508.pdf/ (accessed 17.07.2019).
9. ZONDA hexapod. Available at: https://www.symetrie.fr/en/ products/positioning-hexapods/zonda/ (accessed 17.07.2019).
10. Information about R&D No AAAA-A16-116052710015-3. Available at: https://rosrid.ru/nioktr/IJJUCSNSQASBJTR70JILYZHY/ (accessed 17.07.2019). (In Russ.)
11. Slobodzyan N.S. Evaluation of open-loop linear drive accuracy achieved by calibration and linear thermal expansion compensation // Radio industry (Russia). Vol. 29 (2). Pp. 54-61. https://doi.org/10.21778/2413-9599-2019-29-2-54-61.
12. Navvabi H., Markazi A. H. D. Position control of Stewart manipulator using a new extended adaptive fuzzy sliding mode controller and observer (E-AFSMCO) // Journal of the Franklin Institute. 2018. Vol. 355. Issue 5. Pp. 2583-2609. https://doi.org/10.1016/j.jfranklin.2018.01.032.
13. Lipatov A.N., Lyash A.N., Makarov V.S., Antonenko S.A., Zakharkin G.V. Star sensor for nanosatellite // Proceedings of the 1st All-Russian Scientific and Technological Conference "Contemporary Problems of Spacecraft Attitude Determination and Control" // Space Research Institute of the Russian Academy of Sciences (Russia). 2009. Pp. 66-77. (In Russ.)
14. Guerard J., Larroque M., Lizin G., Verstraeten L., Delavoipiere G. MEMS gyroscope demonstration for space application, using a DPC // 6th International Workshop on Analogue and Mixed-Signal Integrated Circuits for Space Applications (AMICSA 2016). 2016. Pp.15-21. https://hal.archives-ouvertes.fr/hal-01994045.
15. Zhukov Yu. A., Korotkov E.B., Slobodzyan N.S. Management of a high-precision positioning and orientation system for a space purpose based on a hexapod with a "spatial position sensor" // Extreme Robotics (Russia). 2017. Vol. 1. No 1. Pp. 256-265. (In Russ.)
16. Zhukov Yu. A., Korotkov E.B., Moroz A.V., Slobodzyan N.S. Evaluation of the sensitivity of a precision hexapod to design parameters, technological and measuring errors // Desyataya vserossiiskaya mul'tikonferentsiya po problemam upravleniya (MKPU-2017). Materialy 10 Vserossiiskoi mul'tikonferentsii v 3 tomakh. (Russia). Editor-in-chief: I.A. Kalyaev. 2017. Pp. 35-37. (In Russ.)
17. Absolute Linear Transducer Type 3711. Available at: https://www.zettlex.com/wp-content/uploads/Linear-Transducer-3711-OEM-Product-Guide.pdf (accessed 17.07.2019).
18. SMART Position Sensor. Available at: https://sensing.honeywell.com/honeywell-sensing-smart-position-sensor-35-mm-75-mm-225-mm-linear-configurations-productsheet-000674-6-en.pdf2.pdf (accessed 17.07.2019).
19. XL-80 laser system. Available at: https://www.renishaw.ru/ru/xl-80-laser-system--8268 (accessed 17.07.2019).

A.V. Vasiliev, A.V. Sergeev

DEVELOPMENT OF REQUIREMENTS FOR A GROUND TESTBED FOR MODELING AND RESEARCH OF REMOTE CONTROL TECHNOLOGIES FOR A SMALL LUNAR EXPLORATION ROVER

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
andrey@rtc.ru, i.shardyko@rtc.ru

Abstract

The issues of technical support of the planned space experiment Kontur-3 are discussed in terms of the defining the foundation for the construction of a terrestrial testground for communication sessions with a ground-based prototype Lunokhod-RTC as a control object. The conditions for the locomotion of moon rovers on the surface of the Moon are analyzed. Formed requirements for the composition of the objects of the testground. A possible option for constructing such a testground is given.

Keywords: Moon, lunar rovers, planetary rovers, testground, control technologies, ground testing.

Introduction

The possible schedule of the Moon exploration was repeatedly discussed by domestic and foreign experts. To date, the overall structure of these phases is mostly clear, although there are some differences in the composition, sequence, and projected timing of their implementation.

An essential part of the first stages of the Moon's exploration in the next 10-15 years will be to exploit mobile robots in solving a wide range of tasks on the Moon's surface - from research to construction. At present, an outline of the possible range of required mobile robotic systems (MRS) has already been formed, including: a heavy research rover for geological exploration on the lunar surface, a heavy unpressurized lunar rover with piloting capability for transportation, loading/unloading and service tasks, heavy sealed manned lunar rover, heavy moon rover-builder, light exploration and service lunar rovers [1-8].

An important scientific and technological challenge on the way of celestial bodies exploration by means of robotics is the creation and development of a technology for their remote control. Despite the significant success achieved in recent years in the field of autonomous control systems for MRS, the role of humans in the control loop of planetary robots in the near future will remain significant. Therefore, the development and improvement of remote control technologies (teleoperation) for objects located at a considerable distance from the control centers is a vitally important scientific and technical task.

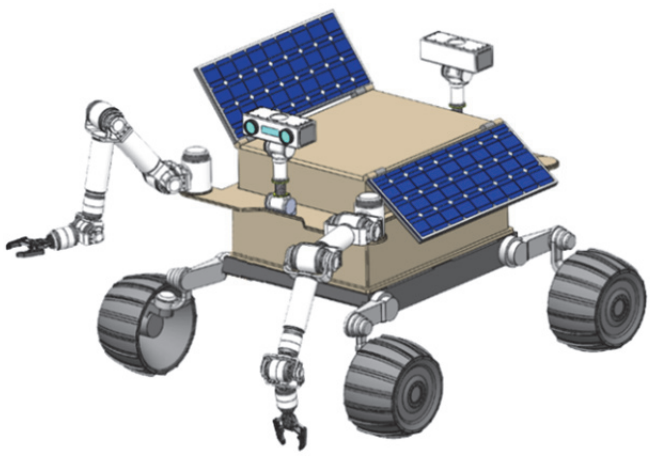


Figure 1 – The design concept of the robot «Lunokhod-RTC» for SE «Kontur-3»

A series of joint Russian-German space experiments (SE) Kontur are aimed at solving this task. The first two SEs in the series were carried out in 2008-2016 [9-12]. The purpose of the SEs Kontur is to study how hostile to humans weightlessness factors affect the operator located on board the manned space complex (MSC), and also to work out the control technologies of the robotic complexes taking into account delays in the communication channel, working out the ergonomic issues of building places of the operator in zero gravity, working out scenarios for the functioning of moon rovers during future expeditions on the moon.

Within the framework of the currently planned SE Kontur-3, from the Russian side, it is proposed to use a small model (weighing about 150 ... 200 kg) of the lunar rover, intended both for scientific exploration and a number of service tasks during maintenance of objects of scientific and industrial importance on the surface of the Moon [13]. Figure 1 shows the preliminary technical concept of the ground rover prototype for development and testing of the different scenarios of perspective exploration and exploitation missions on the Moon.

An important condition for conducting the planned SE Kontur-3 is to create the testground (polygon) in order to perform control sessions with the ground rover prototype simulating scenarios of future lunar missions. The focus of this work is to define the requirements for such a polygon and also to suggest possible polygon implementations

Tasks of the service and exploration lunar rover

Based on the analysis of the well-known publications [1-2, 6-7] and scientific reviews carried out at the Central Research & Development Institute (RTC) [3-5, 8] on the tasks potentially faced by moon rovers in future missions, the considered object can be seen as a prototype of a small service/exploration rover, whose tasks include:

– conducting research and exploration to study the composition and distribution as well as to mark useful resources on map, for example, loosely coupled and frozen volatiles in the polar regions of the moon;

– assisting astronauts on the surface of the moon when they perform maintenance tasks on lunar infrastructure or transport small objects, etc.:

– tasks that can be performed in the telepresence mode, when using the virtual and augmented reality technologies, the operator has the impression of being on the surface of the moon. The latter tasks may include: transporting and placing on the lunar surface small instruments for scientific or industrial use, assembling spatial structures, connecting individual components together, working with connectors.

Analysis of the motion conditions of the lunar rover on the Moon surface

If we analyze information about the surface of the Moon, which is known from the results of previous studies, we can get the following picture [14-16].

General information about the Moon and its surface is given in Table 1.

Table 1. The main physical characteristics of the moon and its surface

Parameter	Value
Distance from the Earth, km, average	384 400
Period of circulation, earth day	27,3
Radio transmission time to Earth, s	1,3
Gravity at the surface, m/s^2	1,62
Atmosphere presence	No
The temperature at the surface, K (°C), according to [16]	from 120 to 400 (from minus 150 to plus 130)
Solar radiation on the surface, W/m^2	1400
Surface Albedo	0,075
Magnetic field	Nearly absent
Space radiation on the surface	Approximately 50 % from outer space
Characteristic processes on the surface	Meteorite bombardment

The surface of the Moon is made of craters and hills, saturated with stony inclusions. The constant bombardment of the moon by meteorites is the reason that its entire surface is covered with a layer of fine crushed matter, the regolith. Regolith is a good thermo-insulating material, therefore even at a depth of several centimeters a constant temperature is maintained.

Regolith is a fairly homogeneous class of soil in terms of particle size – silty sand with a noticeable amount of stones and stone blocks, which can be considered separate protruding obstacles when analyzing traversability.

In comparison with terrestrial analogs, the lunar soil is easily compressed (compacted) from a density of 1,2 to 1,9 g/cm^3, highly tends to electrization, has anomalous adhesion and low thermal conductivity. Despite noticeable stickiness, it is unstable to vibrations. Being able to easily hold a vertical wall, but it has a slope angle of approximately 45° in the case of free pouring.

As a result of research by Soviet and American apparatuses, it was found that the difference in the geomorphological situation in the areas of their work had little effect on the physical and mechanical properties of the soil [14]. Therefore, it is possible to consider the average, most typical soil parameters in natural occurrence, for fairly large areas of Moon.

Table 2 shows the main mechanical properties of the surface layer of the lunar soil according to the results of using the Lunokhod-1 research using contact methods, which are of greatest interest from the point of view of ensuring movement on the lunar surface and modeling the appropriate conditions on Earth.

The uppermost surface layer along the Lunokhod-1 route was a pulverized, easily deformable material. The fine-grained soil layer generally had a depth of at least 5 cm and was fairly uniform. In the case of straight movement of the Lunokhod-1, the depth of the rut varied from 1 to 5 cm. When moving on steep slopes, the rut depth increased due to skidding. Severe destruction of the soil occurred when turning - the depth of the track reached 10 cm.

Table 2. The main mechanical properties of the surface layer of the lunar soil

Parameter	**Value**
The average particle size, microns	10-100
Bulk density, g/sm^3	1-1,8
Porosity, %	40-50
Cohesion, kPa	0,5-5,0
Angle of internal friction, °	10-30
Carrying capacity, kPa	10-100

The most difficult parts of the lunar surface from the point of view of ensuring the movement of the vehicle are deposits of loose soil in combination with large slopes on the slopes of craters.

Measurements of the carrying capacity of the soil showed a significant heterogeneity of mechanical properties along the route of the Lunokhod-1 from 0,2 to 1,0 kg/cm^2 (20-100 kPa). For relatively even areas and intercrater spaces, the most frequent (60% probability) areas of soil with carrying capacity from 25 to 55 kPa were encountered. Along the entire route of the Lunokhod-1 movement, the most frequently (46 %) carrying capacity was equal to 0,34 kg/cm^2 (34 kPa) [14].

Carrying capacity of less than 25 kPa was observed on crater rims and on slopes with steepness of more than 10°. Lower carrying capacity was found in areas with pronounced traces of the processes of crater formation and other forms of relief in combination with significant slopes [16].

The resistance to rotational shear at the explored sites ranged from 0,02 to 0,09 kg/cm^2 (2-9 kPa). The most common value of the resistance of the rotational shear was 0,048 kg/cm^2 (4,8 kPa) [14].

The lunar soil along the routes of the Soviet lunar rovers has good compactibility, while with the compaction of the soil the values of its carrying capacity and resistance to rotational shear increase. The carrying capacity of the uppermost layer of soil (from 1 to 5 cm depth) was estimated by calculation from the depth of the track, the load on the ninth wheel and its dimensions. It turned out to be very small: from 0,02 to 0,03 kg/cm^2 (2-3 kPa). At a depth of more than 10 cm, the soil has the best mechanical properties: the carrying capacity is 1 kg/cm^2 (100 kPa), the resistance to rotational shear is 0,06 kg/cm^2 (6 kPa).

The study of the coupling properties of the Lunokhod-1 chassis revealed that the movement resistance coefficient ranged from 0,15 to 0,25. Specific free thrust (ratio of tractive force to weight) reached values from 0,2 to 0,41. Traction characteristics ensured steady movement of the moon rover on the slopes of steepness up to 20°. The parameters of interaction with the ground of the Lunokhod chassis were very close to the values obtained during testing the self-propelled chassis on Earth. Slippage of the drive wheels ranged from 0 to 15 % in horizontal areas and up to 30 % on slopes [14].

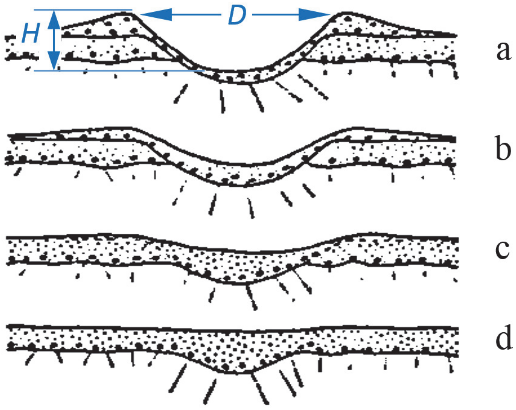
Experts of VNIItransmash used grinded volcanic pumice stone and dry cohesionless quartz sand [17] as analogs of lunar regolith on Earth for physical modeling and running trials of experimental chasses prototypes, because these materials have similar to regolith values of the movement resistance coefficient f, and cohesion φ (see table 3).

Table 3. Earth analogs of the lunar regolith [17]

Soil Type	***f***	***φ***	**Description**
Lunar regolith	0,15…0,25	0,7…0,8	–
«Soil-1»	0,12	0,67	Grinded volcanic pumice
«Soil -2»	0,17	0,61	Cohesionless quartz sand

Craters, typical cross-sections of which are shown in Figure 2, represent a specific and distinctive form of lunar relief formations. Freshly formed small craters usually have a nearly spherical shape. About 95 % of lunar craters are of this shape. Large craters, as a rule, are flat-bottomed, but there are craters of conical, cylindrical, and other forms. Older craters are smoothed and deformed due to erosion.

According to Figure 2, craters are classified into three main morphological classes — A, B, and C by their geological and morphological properties. Additional classes AB, BC are also distinguished. The youngest class A craters are characterized by a maximum steepness of the walls, a pronounced rim and a large number of stones both inside the crater and in the concentric area around it [15].



a – class A crater; b – class B crater; c – class C crater;
d – intercrater space (fully buried crater)

Figure 2 – Presumptive cross-section of typical geological and morphological formations [15]

As can be seen from Table 4, the vast majority of craters in the considered marine areas - up to 97 % - are represented by class C craters (wall slope angle from 8 to 10°), BC (from 10 to 15 °) and B (from 20 to 25°). The share of "young" craters - A (from 35 to 45°) and AB (from 30 to 40°) does not exceed 3 %. The most common craters of classes C and BC have wall slope angles that does not exceed 20° and are characterized by a very small number of fragmentation inclusions (stones), often in a partially buried position in the regolith.

Table 4. The percentage of craters of various classes on the lunar seas and on the route of the Lunokhod-1

Crater Class	**H/D**	**Crater's Wall Slope Angle, °**	**Total Percentage on Moon Seas, %**	**Total Percentage on the Lunokhod-1 Route, %**
A	1/4…1/5	35 - 45	0,5	1
AB	1/6…1/7	30 - 40	2,5	2
B	1/8…1/10	20 - 25	20	17
BC	1/11…1/12	10 - 15	30	32
C	$\leq$ 1/14	8 - 10	47	48

The other characteristic elements of the lunar surface are stones and stone ridges. Stones are most often found near craters. In this case, the larger the crater, the more stones there are around it.

All the morphological diversity of the primary stones in emissions from fresh craters can be reduced to four main types, presented in Figure 3: stones of irregular (I), pyramidal (II), prismatic (III) and flattened (slab) shape (IV). In terms of morphology, as shown in Figure 3, the stones of the selected types can be further divided by degree of destruction into three age classes: primary angular (1), angular-rounded (2), rounded (3).

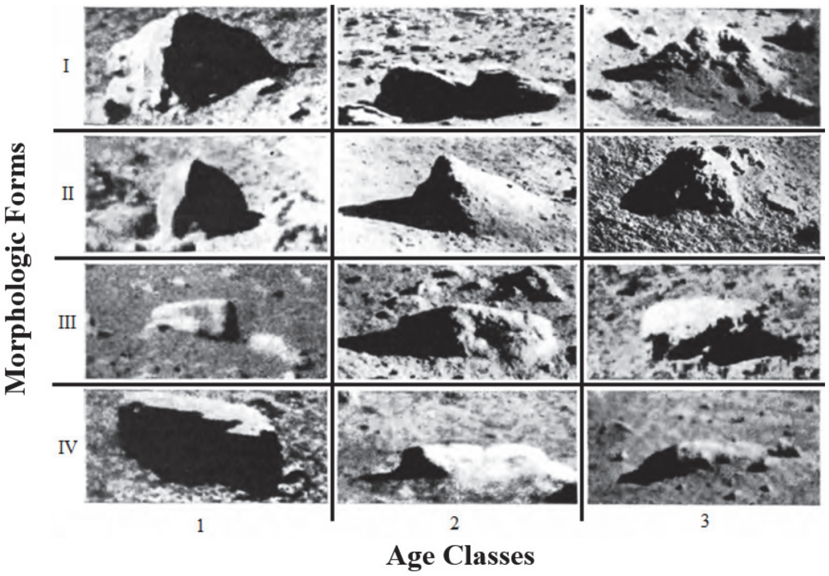


Figure 3 – The morphology of the stones in the Lunokhod-1 exploration area [14, 15]

Class 1 includes stones of a clear, initially angular shape, associated mainly with class A craters; Class 2 – angular-rounded stones, most often related to class B craters; class 3 consists of rounded stones, distributed mainly in the craters of class C and in the intercrater space.

For large stones of class 3 (more than 20 cm), deeply buried location is widespread, as well as the presence of a plume of small clastic material around them.

The distribution of stones of various size along the Lunokhod-1 route shows that the surface of impact craters belonging to a special morphological group as well as of class A craters is the most rocky, while the surface of craters C and intercrater space are the least rocky areas. At the same time, the largest number of encountered stones have diameter from 2 to 15 cm. The probabilistic number of stones with diameter from 10 to 15 cm: up to 40 pieces per 100 m^2. Stones larger than 15 cm occured much rarer. Therefore, stones larger than 15...20 cm can be considered as single obstacles in the path of the moon rover, which can be surmountable or insurmountable.

Thus, it is necessary to consider the following elements as distinctive features for modeling the conditions of movement on the Moon:

– sandy coverage of the surface to imitate lunar regolith;
– sandy slopes that are typical for large craters of classes C and BC;
– small craters with a diameter of not greater than 1 m;
– protruding large stones (up to 200...300 mm in diameter);
– placers of small stones.

Terrestrial testground requirements

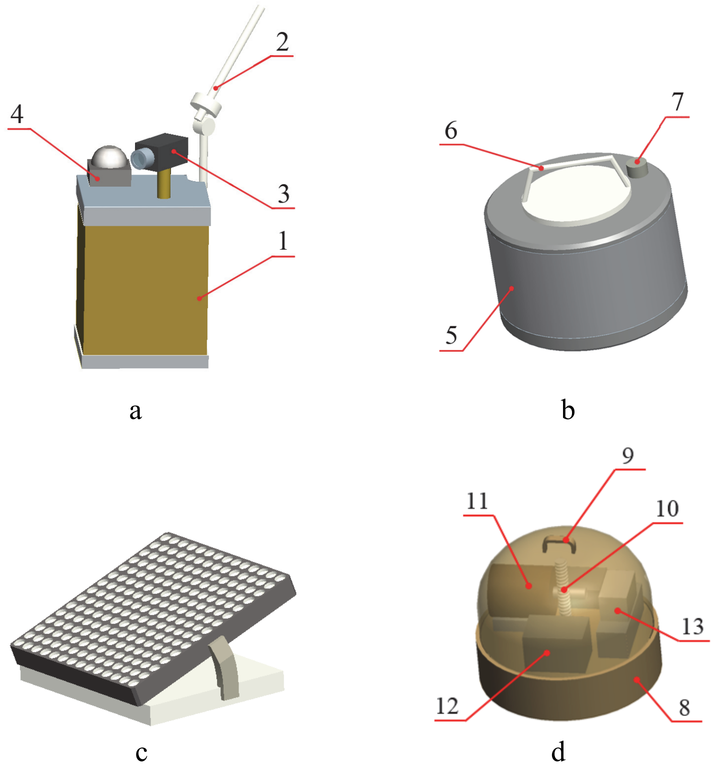
In order to develop the skills of an astronaut operator concerning control of a prototype of a lunar rover from spacecraft, as well as to develop and test the technologies for such control, it is necessary to investigate both locomotion and manipulation tasks within scenarios simulating conditions of lunar missions. To do this, it is necessary to equip a room in which the testground protected from external weather conditions as well as the operators workstations would be placed to provide technical support for the sessions.

In order to implement the SE Kontur-3 sessions, in accordance with the above information about the nature of the lunar surface, the objects listed in Table 5 should be available at the testing ground simulating lunar conditions.

Table 6 shows the infrastructure elements of the testground (active and auxiliary) that are necessary for implementing the SE scenarios. To study manipulation tasks of various degrees of complexity, it is proposed to use mock-ups of scientific instruments similar to those considered in the scenarios developed for a heavy research lunar rover «Robot Geologist» in the framework of the corresponding research [3-5]. An approximate design of the scientific instruments taken as the basis is presented in Figure 4.

Table 5. Elements of the testground relief

№	Object	Studied Tasks	Description
1	Sand coverage of the entire testground area	All tasks associated with the movement of the moon rover	Dry cohesionless quartz sand is considered as a model of the soil
2	Small craters with a diameter of not greater than 1...3 m	Control of a moon rover with moving around or over local obstacles	Craters of classes C and BC with wall slope angles from 5° to 15°
3	Placers of small stones		Small stones imitating the fields of grinded material on the surface of the moon. Size of stones is from 20 to 150 mm
4	Large stones		Stones from 200 to 300 mm in height, partially buried in sand
5	Slopes	Control of the moon rover locomotion with partial skidding of propellers. Actions in emergency situations	Imitation of hilly terrain or long slopes inside large craters. Slope angles from 5 to 10 °. The long slope (hill) structurally serves as a preparatory relief element for the organization of a large crater (pit)
6	Large crater		Large crater of intermediate class BC / B with the wall slope angle up to 15°...20°



a – central unit (CU) of the automated scientific station; b – seismometer; c – laser reflectometer; d – gas analyzer (cap is transparent only to show the inner elements)

1 – CU housing; 2 – antenna of radio channel; 3 – TV-spectrometer camera; 4 – detector of space rays; 5 – seismometer housing; 6 – holder; 7 – electrical connector; 8 – housing; 9 – holder; 10 – small-size drill; 11 – drive; 12 – electronics unit; 13 – mass-spectrometer

Figure 4 – Scientific equipment prototypes for an exploration rover [5]

Table 6. Testground infrastructure

№	Object	Study Tasks	Description
1	A mock-up of the landing platform with ramps for the lunar rover descent and compartment for installing the container with the gained samples	Descending from the landing platform along the ramps. Delivery of the container with the gained samples to the platform. Scenarios of possible emergency situations during the descent of the lunar rover (optional)	The most simplified and lightweight design, the main elements of which are the ramps
2	Imitator of the central unit of an automatic scientific station	Manipulation. Maintenance of scientific equipment (connection to power supply), visual inspection,	Simplified mock-up allowing to perform maintenance operations (cable connection, replacement of functional modules)
3	Seismometer imitator	Manipulation. Removal from the moon rover and installation of scientific instruments on the surface of the moon	Simplified mock-up of a cylindrical device with a handle for grip by a manipulator
4	Imitator of a laser corner reflector	Manipulation. Installation on the ground, setting a given orientation, bringing into position (unfolding)	Simplified mock-up imitating a panel of laser corner reflectors
5	Imitator of a gas analyzer situated on the ground	Manipulation. Removal from the moon rover and installation of scientific instruments on the surface of the moon.	Simplified mock-up of a cylindrical device with a dome-shaped lid and a holder for grip by a manipulator
6	Imitator of energy module with solar panel	Manipulation tasks. Cable connection Visual inspection. Cleaning the working surfaces of the equipment.	Simplified mock-up in the form of a "box" connected by a cable to the mock-up of the solar array with the possibility of changing the orientation
7	Cable	Manipulation tasks. Running the cable in a given way. Working with connectors	A mock-up of cable with connectors. Option of composite cable is possible (connection of several segments)
8	Telecameras	Auxiliary elements for the preparation and maintenance of SE sessions	Cameras (up to 8 psc) for SE sessions visual support and recording, mounted on racks around the perimeter of the testground
9	Fencing		Testground perimeter fencing with technological passages for staff
10	Stairs		Technological stairs for climbing the elevations in case of emergency

Possible implementation of the testground

A possible variant for constructing a testground for developing and testing the functional scenarios of a lunar rover ground prototype during the SE Kontur-3 is presented in Figure 5.

The testground has dimensions of 20×20 m. Its central part is occupied by an elevation of about 1 m in height (indicated by H1000 in the figure), in which a bowl pit is made imitating a large crater of class B/BC with wall slope angles from 10 to 20°. The diameter of the crater is about 10 m. A ten-degree slope leads to the crater rim from outside.

The rest of the testground, marked H150, is a relatively flat surface filled with sand. In this part there are small irregularities such as small hills, craters, placers of small stones and imitators of large stones as insurmountable obstacles for the lunar rover.

The testground has a fence around the perimeter. There are video cameras for recording and tracking the rover on the elevation at the perimeter.

The elevation with a large crater has a vertical wall with a fence along its perimeter. For quick access to the top of the hill in case of any emergency situations, technological ladders for staff are provided.

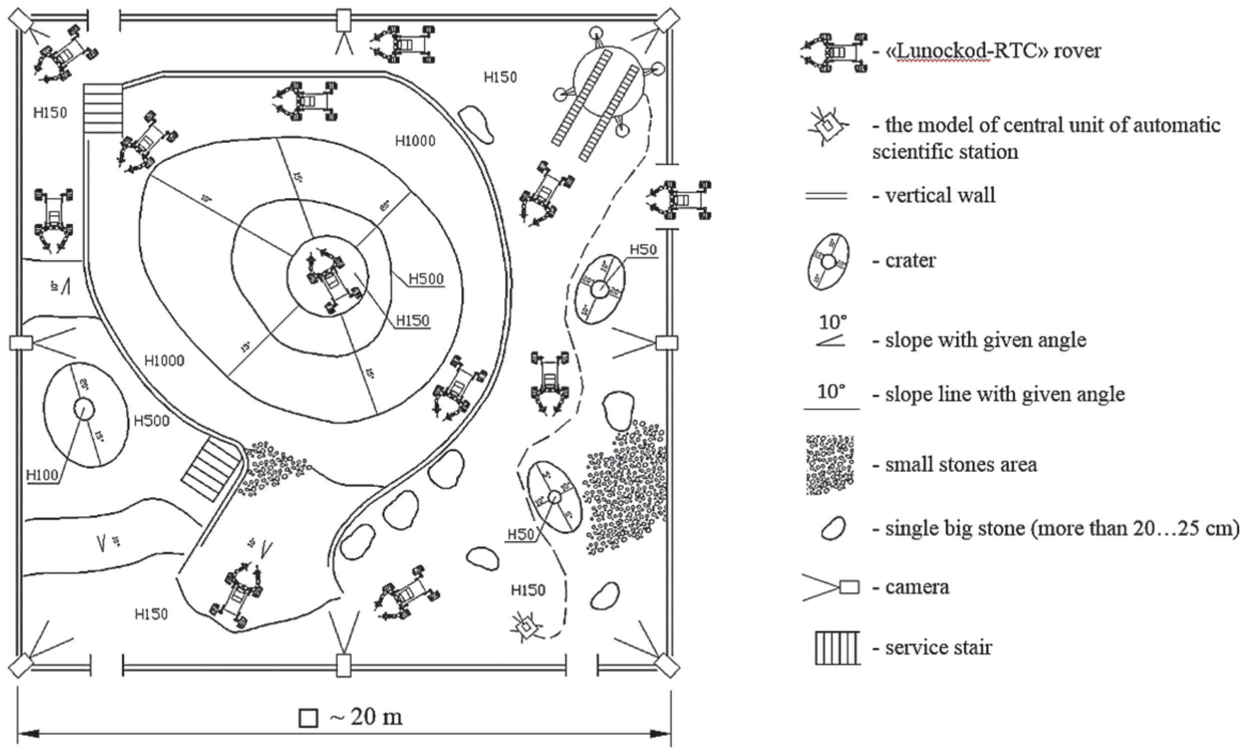


Figure 5 – Diagram of a possible testground layout

Conclusion

In the course of the currently planned SE Kontur-3, from the Russian side, it is proposed to use a small model (weighing about 150...200 kg) of the lunar rover, intended both for scientific research and a number of service tasks during maintenance of scientific and industrial facilities on the Moon surface.

Requirements for both the Lunokhod-RTC ground prototype, which will be used as a control object in the SE Kontur-3, and the testground for the SE sessions, are preliminarily determined.

Further work is related to the detailed elaboration of scenarios for conducting SE sessions and the refinement of the design of the control object and the testground site.

References

1. Luna – shag k tehnologijam osvoenija Solnechnoj sistemy / pod nauchn. red. V.P. Legostaeva i V.A. Lopoty. – M: RKK «Energiya», 2011. – 584 s.
2. Bodrova J.S., Karabadzhak G.F., Raykunov K.G. Space robotics mobile vehicle platforms, their priority tasks and potential usage scenarios to support Russian manned Moon exploration program // Abstracts of the 30th International Scientific and Technological Conference «Extreme Robotics». – SPb: OOO «Izdatel'sko-poligraficheskii kompleks «Gangut», 2019. – pp. 318-319.

3. Vasiliev A.V. Obosnovanie trebovanij k mobil'noj robototehnicheskoj sisteme dlja geologicheskoj razvedki na poverhnosti Luny // Jekstremal'naja robototehnika: Trudy mezhdunarodnoj nauchno-tehnicheskoj konferencii. – SPb: Izd-vo «Politehnika-servis», 2015. – S.186-191.
4. Vasiliev A.V., Kondratiev A.S., Gradovtsev A.A., Dalyaev I.Yu. Issledovanie i razrabotka proektnogo oblika mobil'noj robototehnicheskoj sistemy dlja provedenija geologicheskoj raz-vedki na poverhnosti Luny // Trudy SPIIRAN. 2016. Vyp. 2(45). – URL: http://proceedings.spiiras.nw.ru/ojs/index.php/sp/article/view/3266 (access date 18.07.2019). – DOI: 10.15622/sp.45.9
5. Proektno-poiskovye issledovanija po opredeleniju tehnicheskogo oblika i taktiko-tehnicheskih harakteristik mobil'noj robototehnicheskoj sistemy dlja provedenija geologicheskoj razvedki na poverhnosti Luny. Razrabotka predlozhenij v proekt TZ na OKR po sozdaniju mobil'noj robototehnicheskoj sistemy: otchjot o NIR (zakljuchit.)/ RTC; ruk. A.S. Kondratiev; ispoln.: A.A. Gradovtsev [i dr.]. – SPb., 2015. – 303 s.
6. Predlozhenija po kosmicheskomu apparatu s lunohodom [Electronic resource] / V.A. Vorontsov, A.M. Kraynov, E.V. Vlasenkov [i dr.] // Inzhenernyj zhurnal: nauka i innovacii. – 2016. – №5. – 31 s. – URL: http://engjournal.ru/catalog/arse/itae/1492.html (access date 18.07.2019). – DOI: 10.18698/2308-6033-2016-5-1492
7. Novye proektno-komponovochnye reshenija dlja povyshenija podvizhnosti i funkcional'nyh vozmozhnostej planetohodov / M.I. Malenkov, A.N. Bogachev, V.A. Volov [i dr.] // Izvestija JuFU. Tehnicheskie nauki. – 2017. – №1-2. – S. 42-54.
8. Sozdanie i issledovatel'skie ispytanija jeksperimental'nogo obrazca mobil'noj transportno-manipuljacionnoj robototehnicheskoj sistemy s avtonomnoj sistemoj navigacii i orientacii v prostranstve: Vybor napravlenija issledovanij. Teoreticheskij issledovanija: otchjot o NIR (promezhut.) / RTC; ruk. I.Yu. Dalyaev; ispoln.: A.V. Vasiliev [i dr.]. – SPb., 2018. – 286 s.
9. Udaljonnoe upravlenie robototehnicheskimi ob#ektami v kosmicheskih jeksperimentah serii «Kontur» / V.S. Zaborovskij, A.S. Kondrat'ev, A.V. Silinenko [i dr.] // Nauchno-tehnicheskie vedomosti SPbGPU. Informatika. Telekommunikacii. Upravlenie. – 2012. – № 6 (162). – S. 23-32.
10. Guk M.Yu., Silinenko A.V. «Kontur-2» – sovmestnyj rossijsko-germanskij kosmicheskij jeksperiment po silomomentnomu upravleniju nazemnymi robotami s borta MKS // Trudy mezhdunarodnoj nauchno-tehnicheskoj konferencii «Jekstremal'naja robototehnika», 1-2 oktjabrja 2014. – SPb.: Izd-vo «Politehnika-servis», 2014. – S. 59-64.
11. Kontur-2 Mission: the DLR Force Feedback Joystick for Space Telemanipulation from the ISS / C. Riecke, J. Artigas, R. Balachandran [et al.] // Proceedings of International Symposium on Artificial Intelligence, Robotics and Automation in Space (i-SAIRAS), At Beijing, China. – 2016.
12. Rezul'taty kosmicheskogo jeksperimenta «Kontur-2» po otrabotke tehnologij udaljonnogo upravlenija naplanetnymi robototehnicheskimi ob#ektami / V.A. Muljuha, V.S. Zaborovskij, M.Ju. Guk, A.V. Silinenko // Izvestija JuFU. Tehnicheskie nauki. – 2017. – №9. – S.153-169. – DOI: 10.23683/2311-3103-2017-9-153-169
13. Vasiliev A.V., Dalyaev I.Yu. Razrabotki CNII RTK v oblasti robototehniki dlja obespechenija budushhih orbital'nyh i naplanetnyh missij // Jekstremal'naja robototehnika i konversionnye tehnologii: Sbornik tezisov Mezhdunarodnoj nauchno-tehnicheskoj konferencii. – S.43-48.
14. Peredvizhnaja laboratorija na Lune Lunohod-1 / pod red. A.P. Vinogradova. – M.: Nauka, 1971. – T.1. – 128 s.
15. Peredvizhnaja laboratorija na Lune Lunohod-1 / pod red. V. L. Barsukova. – M.: Наука, 1978. – T. 2. – 184 s.
16. Planetohody / A.L. Kemurdzhian, V.V. Gromov, I.F. Kazhukalo [i dr.]; pod red. A.L. Kemurdzhiana. – 2-e izd., pererab. i dop. – M.: Mashinostroenie, 1993. – 400 s.
17. Dinamika planetohoda / A.L. Kemurdzhian [i dr.]. – M: Nauka, Glavnaja redakcija fiziko-matematicheskoj literatury, 1979. – 440 s.

A.I. Bykov[1], A.V. Artemev[1], A.N. Sova[2]

RESULTS OF ANALYSIS OF EXPERIMENTAL GROUND TESTING METHODS OF PLANETARY ROVERS

[1]Affiliate of Lavochkin Association, Kaluga, Russia, bykovartem1994@yandex.ru, arav@laspace.ru
[2]Moscow Automobile and Road State Technical University (MARI), Moscow, Russia, slsova@mail.ru

Abstract

The article presents the results of analysis of experimental ground testing methods of planetary rovers. An abnormal situations analysis that arise during planetary rovers operation, shows the need for ground tests to improve the completeness and accuracy of planetary rovers' functioning evaluation in normal mode. Main approaches of planetary rovers ground tests conduct were analyzed. The semi-natural test methods of planetary rover chassis are considered. The structure of the soil for conducting planetary rovers chassis semi-natural tests is determined. Approach to study and modeling of the longitudinal slip six-wheeled planetary rover climbing along an inclined plane is considered. The methodology for testing fault-tolerant planetary rover control systems is given. Analyzed the telemetry testing method with the "ExoMars" rover program. The results of analysis of experimental ground testing methods of planetary rovers is the determination of the merits and shortcomings of the approaches to testing, as well as formation of main provisions integrated test methodology to improve completeness and reliability the assessment rovers functioning in normal mode and in abnormal situations case.

Keywords: planetary rovers, wheel chassis, rovers' experimental ground tests, fault-tolerant control system, telemetry system testing methodology.

Introduction

Self-propelled vehicle "Lunokhod-1" was designed and assembled by Lavochkin Association. "Lunokhod-1" successfully completed scientific task on the surface of the moon. The planetary rover traveled a distance of 10,540 meters, during this time 200 telephotometric panoramas and about 20 thousand images were transmitted to Earth. During the shooting, the most interesting features of the relief stereoscopic images were obtained, allowing for a detailed study of their structure. Ground tests were very important to make this research mission successful.

Space objects studies using planetary rovers are also being carried out at the present time. Consider the basic composition of the rover used for research missions. In general, the rover is a mobile robotic system, which is equipped with a set of instruments and sensors, as well as a manipulation mechanism. The rover must solve research problems in a gravitational effect, different from the earth. In addition, when solving scientific research tasks, a planetary rover needs to move around difficult terrain. Planetary rover's control is possible in manual and automatic modes. In manual mode, the control is carried out by the operator. Operator gets data from cameras and sensor systems. Delays in telemetry sessions when data is exchanged between the rover and the control center on Earth can cause an emergency situations. An emergency situation in this article means rover's mobility loss, as a result it cannot continue to carry out its scientific and other functional tasks.

When controlling a rover in automatic mode, the data for the control system also comes from sensors and cameras. Errors of sensor readings, as well as the visual assessment complexity of the soil carrying capacity using the vision system, can significantly affect a planetary rover control system operation can lead to an emergency situations. Therefore, in order to control the planet rover operation, both in normal mode and in the event of an emergency situations, manual and automatic control modes combined is necessary. But even operator and the automatic control system joint work do not guarantee that there will be no an emergency situations during the operation of the planet rover.

Consider the main emergency situations that may arise during a planet rover operation, for example, the rover Opportunity [1]. Over the entire operation period, Opportunity has been in emergency situations several times. After a successful landing in the Eagle crater, the rover first attempts to go in the direction of crater walls greatest rise angle the ended in failure. It was possible to leave the crater only moving across the slope.

A year later, the rover's front right wheel steering wheel was frozen at the 8° position to the longitudinal axis. To preserve mobility, it was decided to deploy the rover and move in reverse. Later the rover was stuck in a sand dune about 0.35 m high. The lifting angle was no more than 12°. Front in the course movement wheels completely have been dug into the ground . It was decided to go back, that is, to move again the main forward course. A little more than a month later, as a result of numerous local movements, the rover managed to get out onto solid ground.

Also, the Opportunity rover fell into a sand trap during operation and could lose mobility several times as a result of strong skidding.

The considered an emergency situations encountered by the Opportunity rover prove the need for conducting complex planetary rovers' ground tests. At the same time, in the complex ground tests, a planetary rovers' work should be considered both in normal mode and in an emergency situations event.

The work purpose is to analyze the methods of experimental ground testing of planetary rovers in terms their work assessing the completeness and reliability in normal mode and in the event of emergency situations.

Proper function wheel chassis, motion control systems with a sensors set and telemetry systems ensure make a rover functioning without losing mobility. Therefore, in this paper we analyze experimental ground tests techniques this systems.

Methods of testing wheeled chassis planetary rovers

At the moment, wheeled chassis are the main mobile platform for planetary rovers. The main tasks of testing a wheeled chassis are maneuverability evaluation and rover stability assessment during normal operation.

Depending on physical and mechanical properties of a a soil and relief, there is a distinction between supporting and profile maneuverability. At the same time, when evaluating these types of terrain for ground-based landing gear, ground-based experimental tests are carried out [2]. These tests can be carried out both on the natural landscape and on special landfills. As a natural landscape, a complex relief is usually chosen on volcanic sands or in deserts (Figure 1). When testing at special test sites, quartz sand or a soil analogue is used [1, 3].

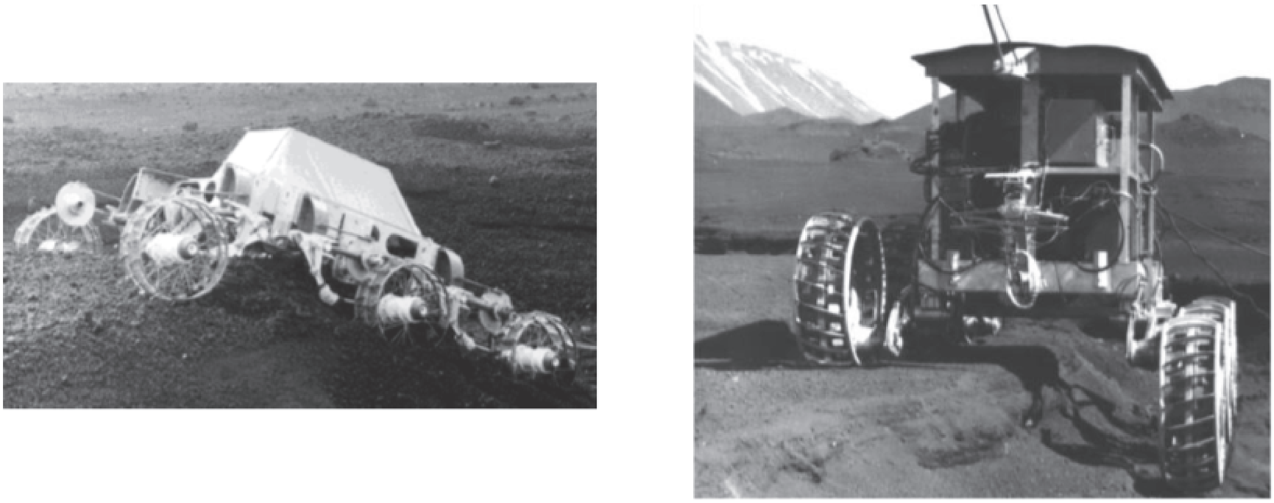


Figure 1 – Fragments of planetary rovers' tests

Soil and terrain choice depends on the space object on which the rover will operate. At the moment, research using planetary rovers are taking place on the Moon and Mars. Both the lunar and the Martian surface represent a complex sandy relief with a large craters number. Therefore, special chassis planetary rovers ground-based experimental tests are required on loosely coupled soils. In the works [4] and [5], the parameters of the soil-analogue for planet rovers ground tests are given. Soil analogue parameters were obtained as a result of the images analysis of the planetary rovers wheels, which got stuck in the ground during operation on Mars.

When testing a planetary rovers' running gear, the main characteristic is the time to move from a starting point to a final point of a route. In addition, the estimated movement average speed over rough terrain. Also tracked are the parameters of the obstacles that the planetary rover can overcome without maneuvering. These parameters include: steps and stones height, terrain elevation angle, loose soils carrying capacity [6].

Also for testing the chassis used rovers' virtual models. Thus, in [7], the problem of a six-wheeled planetary rover longitudinal slip when climbing along an inclined plane is considered. Longitudinal glide occurs when a planetary rover rises uphill on loosely bound ground. An example of such a situation is a planet rover departure from the crater. Longitudinal gliding can lead to a situation where the planetary rover cannot get out of the crater.

Also for testing the chassis used rovers' virtual models. Thus, in [7], the problem of a six-wheeled planetary rover longitudinal slip when climbing along an inclined plane is considered. Longitudinal glide occurs when a planetary rover rises uphill on loosely bound ground. An example of such a situation is a planet rover leaving from the crater. Longitudinal gliding can lead to a situation where the planetary rover cannot get out of the crater.

The authors of [7] collected experimental data to simulate the rover rise on an inclined surface for only one rover's wheel. For this purpose, a special bench equipment was developed (Figure 2), imitating the dynamic loads on the rover's wheel when climbing on an inclined surface with loosely coupled soil. Further modeling of the six-wheeled planetary rover was carried out uphill. The input data for the simulation of each rover's wheel were received from the stand equipment. As a modeling result, an assessment of longitudinal slip effect on the rover's mobility is given.

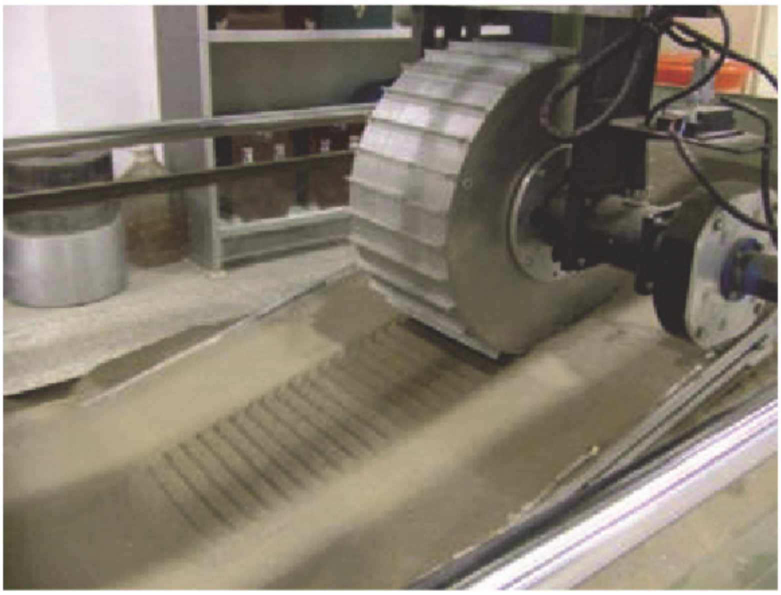


Figure 2 – Bench equipment for experimental data collection

A rover's stability loss is associated with a change in its position in space or on a plane. Therefore, it is necessary for the planetary rover to distinguish between rollover resistance and skidding or sliding resistance. In both cases, two stability types can be considered - static and dynamic. Static stability is usually understood to mean a planetary rover's ability to maintain its original position while at rest or in motion, and dynamic - a rover's ability of to withstand a inclination critical angle achievement its hull under the influence of disturbing factors [8].

In [8], the experimental stability assessment method is presented. When testing used special layouts. Due to the complexity of the gravitational attraction imitation, which is different from the earthly, and also because of the equations describing validity a rover's stability for any gravitational field, the rover's mock-up's tests are conducted under the earth's gravity conditions. Further, an analytical assessment of the planet rover's stability was carried out. Table 1 shows the layouts parameters on which the tests were carried out, and, for comparison, the parameters of Lunokhod-1 are presented. Experimental studies were carried out in cases where mock-up No. 1 and No. 2 were moving down by a stepped obstacle, as well as when mock-up No. 3 ran into an insurmountable obstacle by two wheels simultaneously and braked.

The stepped obstacles were made of planks whose height varied from 0.3 to 0.55 m. A stability and subsequent study a rover deceleration when they collided with an insurmountable obstacle was carried out on a special platform, the angle of which could change. The blow against an insurmountable obstacle, which was a wall with a height of 0.35 m and a width of 0.25 m, was carried out at a speed of 0.555 m / s. The wall was installed at the slope end strictly perpendicular to the layout longitudinal axis. A rover stability investigation when hitting an insurmountable obstacle was carried out at tilt angles from 25 to 27 degrees, and when braking – from 17 to 21 degrees [8].

When testing on mock-ups, the devices recorded angular displacements in a vertical-longitudinal plane, vertical accelerations and deformations of suspensions.

As the tests result, it was obtained that the calculated values of the rover's stability are very close to the experimental ones. A comparison was also made of the calculated stability values for Lunokhod-1 and the actual data obtained during normal operation, the results of which led to the conclusion that the calculated and actual data were in agreement.

Test method for a planetary rover's fault tolerant motion control

In most planetary rovers' models, control can be carried out both in manual and automatic modes [9]. The main systems for a planetary rover's implementation control process are: a technical vision system, a radio communication system, a navigation system, and a mechanical position sensor system [6]. As in all autonomous mobile robots, automatic control systems for planetary rovers can be divided into a strategic control level and a mechanics' functional control level. At a strategic management level, the rover determines the trajectories of movement on an another planet's surface. After determining the trajectory of movement at a mechanics' functional control level, it is determined: which drives should turn on, for what period of time and with what power.

Table 1. Mock-up's and "Lunokhod-1"'s technical characteristics

Parameters	**Dimension**	**Mock-up**			**«Lunokhod-1»**
		№1	№2	№3	
Mass	kg	179	217	255	780
Inertia moment about x axis	kgm^2	310	230	450	290
Gravity height center	m	0,835	0,75	0,98	0,835
Coefficient of accounting for rotating masses	-	-	-	3,8	1,96
Double stiffness extreme suspensions	N/m	16600	1520	1660	1660
Double stiffness of medium suspensions	N/m	7000	8800	7000	7000
Wheel stiffness	N/m	$25 \cdot 10^4$	$25 \cdot 10^4$	$25 \cdot 10^4$	$25 \cdot 10^4$
Double brake stiffness	N·s/m	-	-	2000	2000
I-th wheel damping ratio	N·s/m	125	125	125	125

A planetary rover's strategic level chassis control system is based on neural networks. The input to the planetary rover's neural network comes from a technical vision system. In an ambiguity event in the interpretation of input data, the rover can query the operator to adjust the movement in manual mode. At a mechanics' functional control level systems, besides a task of moving , it is important to prevent slipping. To do this, drive control should be carried out according to the torque [10].

The functioning of a planet rover's motion control systems requires an integrated approach application based on the data on the physical layout and on the results of computational experiments based on a rover's digital model. Consider a control system's test method for example scientific work [11]. This paper discusses the fault-tolerant planetary rover management system. The authors consider the possibility of using this control system in a malfunctions event in a rover's electro-mechanical system. The method task of control is to restore the performance of a faulty planetary rover to a level acceptable for making further scientific tasks [11].

To obtain the initial data, the rover mock-up's tests within the ExoMars program framework were carried out (Figure 3). The rover's model moved straight through a special area with sand. During the tests, the sand was filled up at such a level as to prevent the mock-up from sticking in it. In order to collect a sufficient amount of initial data, 26 experiments were carried out and malfunctions were simulated in a mechanical layout system.

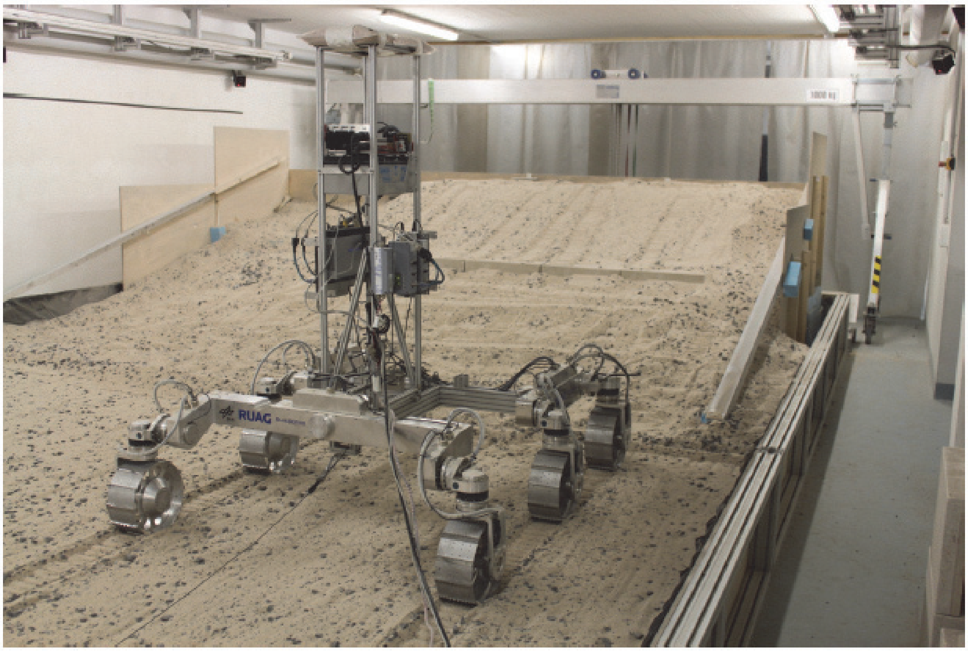


Figure 3 – Moke-up for working out fault-tolerant control system

The use of a fault-tolerant planetary rover control system allows the restoration of a given movement trajectory in the faults presence in the mechanical system.

A planet rover telemetry system testing methodology

A telemetry system fuction is very important when operating the planet rover in manual mode. Communication with the planet rover for direct management can take place with a long time delay. To a planet rover's test in manual mode with delays in telemetry sessions, the European Space Agency (ESA) conducted Mars Rover moke-up test on the ExoMars program [12]. ESA scientists conducted tests on a special field measuring 80 by 50 meters, which simulated the surface structure of Mars. The purpose of the test was to check how the rover could move off the landing module after landing on Mars (Figure 4).

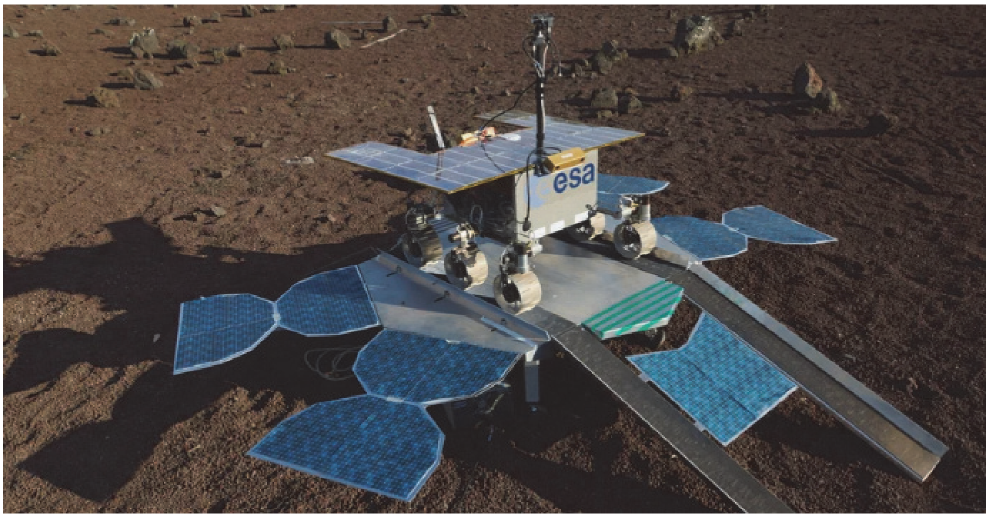


Figure 4 – The rover's moke-up of on the landing module

The rover was controlled by a team located in the Netherlands. At this time, the rover itself was in France. With each new test, the team operating the rover did not know where the rover was located in the test site. The only information they owned came from cameras and sensors installed on the rover itself and the landing module. The remote control complexity was added by the fact that it was not carried out in real time. The team could receive telemetry data at specific time intervals, after which it sent the necessary commands to the rover for execution. Telemetry sessions were held once an hour. Each new task was initially modeled as a virtual model of the rover environment (Figure 5), and then compared with a panoramic image obtained from various cameras installed on the rover [12].

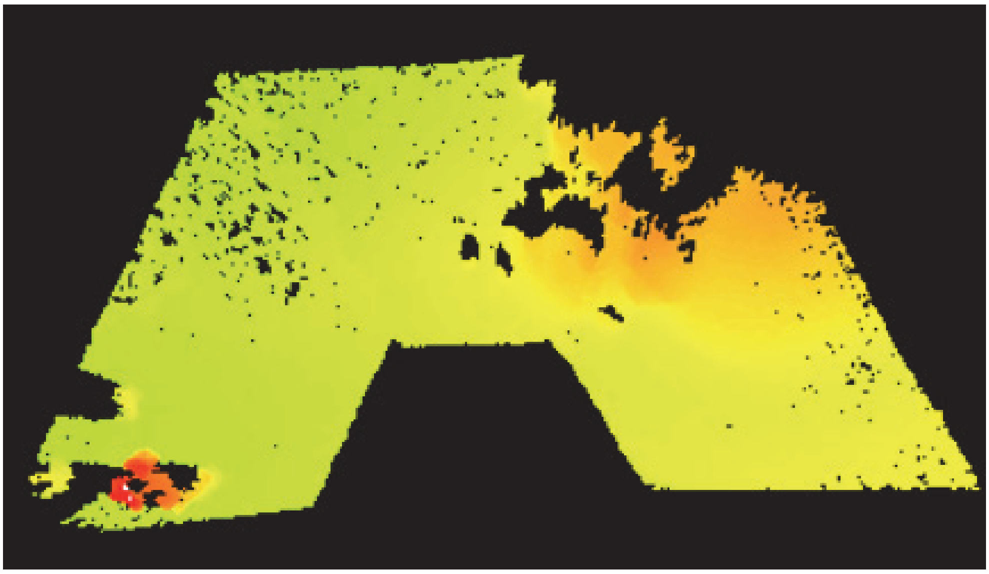


Figure 5 – Digital model environment of the rover

Four of the five tests were successful. Only in one of the five tests, the rover fell from the ramp, on which it needs to move down from the landing module.

Conclusion

A planetary rovers' ground-based experimental tests analyzed in the article are effective in studying a specific subsystems capabilities. In this case, each approach disadvantage is the absence of accounting for the gravitational effect, which is different from the earthly one. Despite the fact that in [8] it is stated that the equations for estimating the stability of a planetary rover are valid for any gravitational field, the authors also note the need to clarify data during ta rover normal operation. In [7] and [11], the initial data for the models were obtained as a result of experimental testing with assumptions that do not fully reflect the real operating conditions of the rover. Testing the method of telemetry, which was conducted by ESA, considers only the exit of the Mars rover from the landing module and does not consider movement on the planet.

As a result of the planetary rovers' ground tests considered methods analysis, it can be concluded that the ground tests of a planetary rovers' individual subsystems cannot give a complete assessment of a planetary rover work under normal conditions. To clarify the parameters of stability and throughput, it is necessary to develop a comprehensive methodology for ground tests, which takes into account the interaction of a rovers' subsystems and the gravitational effect, which differs from the earthly one.

To imitate a gravitational effect, different from the earthly one, one should apply the method of de-suspension. In addition, it is necessary to analyze and improve the existing technological equipment to simulate the gravity of other planets [3].

The tests' results can be the basis for a rover's detailed digital model development. The effectiveness of digital models use for complex electromechanical systems is shown in [13], [14], [15].

Digital models are necessary for the study of critical situations in a rover normal operation. In addition, a planetary rover's digital model can be used to simulate its movement in order to substantiate and develop proposals for making a motion control decision. In the future, the digital model of the planet rover can be upgraded in order to provide the ability to predict the failures of the planet rover systems during normal operation on another planet.

References

1. M.I. Malenkov, V.A. Volov, N.K. Guseva, Ye.A. Lazarev Analiz podvizhnosti marsokhodov dlya razrabotki sistem peredvizheniya i algoritmov upravleniya planetokhodami novogo pokoleniya // Izvestiya YUFU. Tekhnicheskiye nauki. 2015. №1 (162). S. 82-95.
2. M.I. Malenkov, V.A. Sravnitel'nyy analiz komponentov khodovoy chasti samokhodnykh shassi planetokhodov // Izvestiya YUFU. Tekhnicheskiye nauki. 2016. №1 (174). S. 169-185.
3. A.F. Batanov, V.A. Vorontsov, YU.A. Khakhanov Sozdaniye innovatsionnykh kosmicheskikh apparatov dlya issledovaniy v oblasti fundamental'nykh i prikladnykh issledovaniy. problemy sozdaniya sluzhebnykh i nauchnykh sistem »: Trudy konferentsii. (Anapa, Krasnodarskiy kray, 04-09 sentyabrya 2017 g.) M: Izdatel'stvo: Aktsionernoye obshchestvo «Nauchno-proizvodstvennoye ob"yedineniye im. S.A. Lavochkina », 2017 - S. 141-150.

4. YU.A. Khakhanov Rezul'taty vizual'nykh issledovaniy sloyev marsianskogo grunta metodom sravneniya ikh s instrumental'nymi ispytaniyami analogov na Zemle // KHL Akademicheskiye chteniya po kosmonavtike, posvyashchennyye pamyati akademika S.P.Koroleva. Moskva, 26-29.01.2016 g.
5. A.F. Batanov, YU.A. Khakhanov Issledovaniye poverkhnostnogo sloya Luny s pomoshch'yu devyatogo kolesa i razvitiye etogo elementa // KHL Akademicheskiye chteniya po kosmonavtike, posvyashchennyye pamyati akademika S.P.Koroleva. Moskva, 26-29.01. 2016 g.
6. M.I. Malenkov, A.N. Bogachev, V.A. Volov, N.K. Guseva, A.G. Konkolovich, D.N. Kuz'menko, V.M. Kurdzyuk, Ye.A. Lazarev, A.B. Fedorushkov, D.B. Fedorushkov Novyye proyektno-komponovochnyye resheniya dlya povysheniya podvizhnosti i funktsional'nykh vozmozhnostey planetokhodov. // Izvestiya YUFU. Tekhnicheskiye nauki. 2017. №1 (186).
7. L. Chzhendzha, V. Yan. Nadezhnoye adaptivnoye nechetkoye upravleniye dlya planetarnykh roverov pri pod"yeme po deformiruyemym sklonam s prodol'nym uklonom // Hindawi Publishing Corporation.
8. Ye.V. Avotin, I.S. Bolkhovitinov, A.L. Kemurdzhian, M.I. Malenkov, F.P. Shpak Dinamika planetokhoda. // M .: Nauka, 1979 g. 440 s.
9. M.V. Mikhaylyuk, Ye.V. Strashnov, L.A. Finagin Sistema upravleniya virtual'noy model'yu marsokhoda // Trudy nauchno-issledovatel'skogo instituta sistemnykh issledovaniy rossiyskoy akademii nauk. 2018. №4. S. 74-79.
10. V.N. Naumov, K.YU. Mashkov, D.A. Chizhov Matematicheskoye modelirovaniye dinamiki pryamolineynogo dvizheniya robotizirovannogo transportnogo sredstva po deformiruyemomu gruntu // Transportnoye i energeticheskoye mashinostroyeniye. №2, 2012 g. - s. 19-24.
11. A.C. Leite, B. Schafer, M.L. Strategiya upravleniya otkazoustoychivost'yu Souza dlya upravleniya povorotami kolesnykh planetarnykh roverov // Hindawi Publishing Corporation, Journal of Robotics Volume 2012, ID stat'i 694673, 15 stranits.
12. URL:https://www.esa.int/Our_Activities/Space_Engineering_Technology/ExoMars_software_passes_ESA_Mars_Yard_driving_test (data obrashcheniya: 18.03.2019).
13. N.A. Yeremin, L.N. Yeremin Tsifrovoy dvoynik v neftegazovom proizvodstve // Neft'. Gaz. Novatsii. №12, 2018 - s. 14-17 let.
14. A.V. Gur'yanov D.A. Zakoldayev, A.V. Shukalov, I.O. Zharinov, M.O. Kostishin Organizatsiya tsifrovykh proizvodstv Industrii 4. 0 na osnove kiberfizicheskikh sistem i ontologiy // Nauchno-tekhnicheskiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki. 2018. №2.
15. YU.P. Vozmozhnosti proyektirovaniya izdeliy s maloy veroyatnost'yu otkazov v usloviyakh industrii 4.0 // Ontologiya proyektirovaniya. 2019. №1 (31).

I.P. Nanjageev, V.V. Titov

A SYSTEM FOR PAYLOAD INERTIA PARAMETER ESTIMATION BASED ON AN INDUSTRIAL MANIPULATOR WITH A6 DOF FORCE/TORQUE SENSOR: DESIGN AND APPLICATION

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
i.nanjageev@rtc.ru, vtitov@rtc.ru

Abstract

The paper describes the design and application experience of a system for payload inertia parameter estimation based on an industrial manipulator equipped with 6 DoF force/torque sensor. The methods for mass, center of gravity and inertia tensors identification are presented. Error analysis is conducted for the mass and center of gravity identification methods. Theoretical analysis of the errors propagation is confirmed by and simulation and experiment.

Keywords: center of gravity, mass, inertia tensor, identification, industrial manipulator, force/torque sensor.

Introduction

In fast and accurate motion dynamics expected from industrial manipulators the precise knowledge of robotic system dynamic parameters is of great importance. Both the manipulator and its payload parameters are required for the overall model to be valid. Several papers are devoted to the manipulator dynamic parameters identification [1,2] which can be used to identify the manipulator parameter. These parameters are considered constant and usually supplied by the manufacturer. The methods are too complicated for the payload identification alone that typically deemed as a rigid body with constant parameters.

There are several ways to perform inertia parameters identification [3]. In [4] all the methods are classified into two big categories:

1. Static
2. Dynamic

In static methods, the main principle is kineto-static relationships that allow to estimate only the mass and the center of gravity (CoG) of the payload. In the dynamic methods the motion of the payload is utilized to estimate the inertia tensor (typically relative to the body fixed frame). For example in [5], from motion analysis of multi cable suspension the inertia tensor is estimated. The same estimation is done via eigen-motion analysis of the object in [6]. Its noteworthy that in most cases the static parameters (mass and CoG) and the inertia tensor are estimated separately. Typically, the former are used in process of the later estimation.

Industrial manipulators are designed to both accurately follow special trajectories and provide static positioning that makes it convenient tool for its own payload identification. This is quite acute since payload type change or change of payload parameters in the manufacturing process can occur frequently in modern automated manufacturing while estimation of the payload parameters with application of special tools and/or methods is time consuming, technologically cumbersome or pricy (equipment cost).

In the paper a system based on an industrial manipulator equipped with a force/torque sensor is used in two steps inertia parameters identification process:

1) estimation of mass and CoG;
2) inertia tensor estimation.

The first step assumes multiple measurements of the force/torque at the sensor mounted at the end effector of the manipulator before the payload in the kinematic chain. The force/torques are measured at different payload orientations to ensure linear independence and robustness of estimation. The estimation is performed in the least square sense.

In the second step the data obtain in the first is utilized to generation rotation about the CoG of the payload to estimate the inertia tensor of the later. The force/torque measurements are accumulated in motion and post processed to get the required parameters also in the least square sense.

System description

The kinematic structure of the system is presented in the fig. 1.

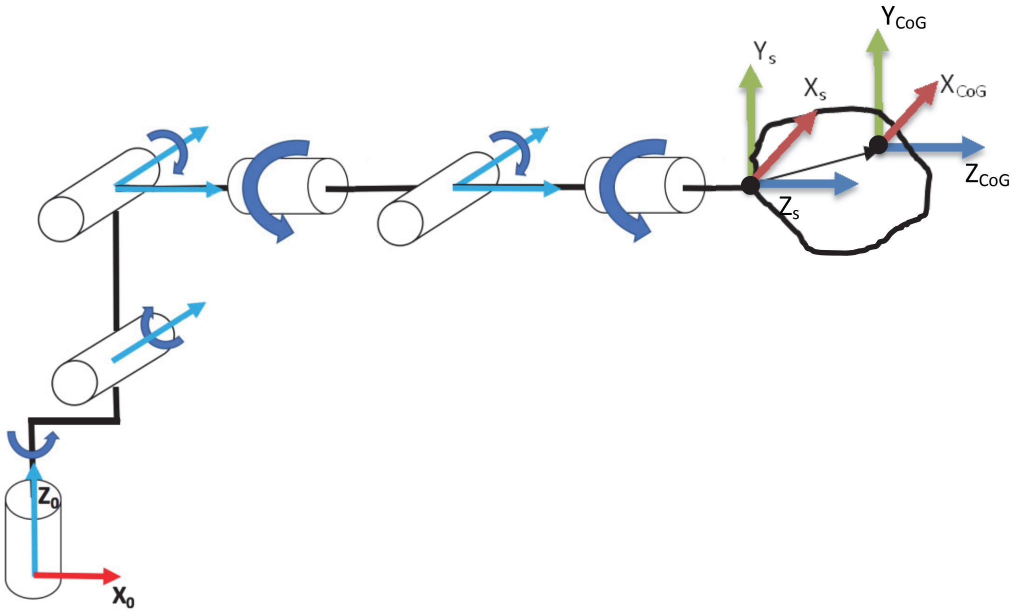


Figure 1 – Kinematics of the system for inertia parameter estimation

In the picture:
XsYsZs – coordinate system (CS) associated with the force/torque sensor (FTS);
CG – objects center of gravity;
g – gravity vector;
r_{cg}^{S} – objects center of gravity vector in FTS coordinate system.

The force/torque acting on the FTS at the mounting point S represented in $X_sY_sZ_s$ CS are gives as (1) [2]:

$$\hat{F}^{s} = m \cdot g^{S}$$
$$\hat{T}^{s} = m \cdot \vec{r}_{CG}^{\,S} \times g^{S} \tag{1}$$

All real FTSs have zero biases of the measured force/torque components. For this reason eq. (1) should be augmented to account for it. The modifications leads to (2):

$$\hat{F}^{s} = m \cdot g^{S} + b_F$$
$$\hat{T}^{s} = m \cdot \vec{r}_{CG}^{\,S} \times g^{S} + b_T \tag{2}$$

with

$\hat{F}^{s} = [F_x^S, F_y^S, F_z^S]^T$ – FTS measurements (forces);

$\hat{T}^{s} = [T_x^S, T_y^S, T_z^S]^T$ – FTS measurements (torques);

$g^S = R_b^S \cdot g^b$ – gravity vector in FTS CS;

$g^b = [0,0,-9,81]^T$ – gravity vector in base frame $X_0Y_0Z_0$;

R_b^S – rotation matrix from base frame into $X_SY_SZ_S$;

$b_f = [b_{fx}, b_{fy}, b_{fz}]^T$ - FTS force components biases;

$b_t = [b_{tx}, b_{ty}, b_{tz}]^T$ - FTS torque components biases.

In general motion the torque at the center of gravity S is calculated according to the Euler motion equation which can be written in the axes of CS $X_sY_sZ_s$ as (3) [3]:

$$\hat{T}^{s} = \vec{r}_{CG}^{\,S} \times F^{s} + I_{CG} \cdot \dot{\omega} + \omega \times (I_{CG} \cdot \omega) + b_T \tag{3}$$

where F_s– force vector at FTS (biases subtracted);

$\omega = [\omega_x, \omega_y, \omega_z]^T$ – angular velocity of the object transformed into $X_{CoG}Y_{CoG}Z_{CoG}$ frame;

$\dot{\omega} = [\dot{\omega}_x, \dot{\omega}_y, \dot{\omega}_z]^T$ – angular acceleration of the object transformed into $X_{CoG}Y_{CoG}Z_{CoG}$ frame;

$I_{CG} = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{12} & I_{21} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{bmatrix}$ – inertia tensor at $X_{CoG}Y_{CoG}Z_{CoG}$ frame.

Step 1: Mass and CoG estimation

The object mass and center of gravity estimation is achieved through measurement of the forces and torques at the object attachment point to FTS when gravity force is applied in different direction. Changing the direction of gravity force is done by positioning only two joint of the industrial manipulator. The manipulator is posed in initial configuration such that the last to joints closest to the attachment point form polar coordinate system. The measurements are taken according to the algorithm:

1) m tangent joint positions are equally spaced in the interval [-90°;90°];

2) for each k-th tangent joint position n yaw joint positions are chosen equally spaced in the interval [-180°; 180°];

3) at each k-th tangent and j-th yaw joint positions the manipulator is paused for 2-3 seconds to mitigate the dynamic effects in the compliant elements of the sensor and the manipulator;

4) after the pause the FTS measurements are recorded.

The accumulated nxm pairs of the force/torque measurements and angles (α_{5k}, α_{6j}) are stored and processed as follows.

For each k-th tangent and j-th yaw poses qq. (2) can be expressed as (4)

$$\hat{F}_i^S = U_{Fi}(\hat{\alpha}_5, \hat{\alpha}_6) \cdot \Psi_F$$
$$\hat{T}_i^S = U_{Ti}(\hat{\alpha}_5, \hat{\alpha}_6) \cdot \Psi_T \tag{4}$$

where α_{5k} – angular position of the k-th tangent measurement;

α_{6j} – angular position of the j-the yaw measurement at the k-th tangent measurement;

$$U_{Fi}(\alpha_5, \alpha_6) = \left[({}^S R_b(\alpha_{5i}, \alpha_{6j}))^T \cdot g_b, I_{3x3} \right]$$

$$\Psi_F = \left[m, b_1, b_2, b_3 \right]^T$$

$$U_{Ti}(\alpha_5, \alpha_6) = \left[m \cdot S(\left[{}^S R_b(\alpha_{5i}, \alpha_{6j}) \right]^T \cdot g_b), I_{3x3} \right]$$

$$S(g^s) = \begin{pmatrix} 0 & -g_z^s & g_y^s \\ g_z^s & 0 & -g_x^s \\ -g_y^s & g_x^s & 0 \end{pmatrix}$$

$${}^S R_b(\alpha_{5i}, \alpha_{6j}) = {}^S R_b(\alpha_{5i}) \cdot {}^S R_b(\alpha_{6j}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sin(\alpha_{5i}) & \cos(\alpha_{5i}) \\ 0 & -\cos(\alpha_{5i}) & -\sin(\alpha_{5i}) \end{bmatrix} \cdot \begin{bmatrix} \sin(\alpha_{6j}) & -\cos(\alpha_{6j}) & 0 \\ \cos(\alpha_{6j}) & \sin(\alpha_{6j}) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Psi_T = \left[r_{CGx}, r_{CGy}, r_{CGz}, b_4, b_5, b_6 \right]^T$$

Index *i* just indicates the number of the measurement in the total sequence.

Stacking all the measurement vectors in one big vector one obtains (5)

$$\hat{f}^S = \Re_{56F} \cdot \Psi_F$$
$$\hat{t}^S = \Re_{56T} \cdot \Psi_T \tag{5}$$

with $\Re_{56F} = \begin{bmatrix} U_{F1} \\ \vdots \\ U_{Fn} \end{bmatrix}, \Re_{56T} = \begin{bmatrix} U_{T1} \\ \vdots \\ U_{Tn} \end{bmatrix}, \hat{f}^S = \begin{bmatrix} \hat{F}_1^S \\ \vdots \\ \hat{F}_n^S \end{bmatrix}, \hat{t}^S = \begin{bmatrix} \hat{T}_1^S \\ \vdots \\ \hat{T}_n^S \end{bmatrix}$

Thus the required parameters are obtain from (5) using Moore-Penrose pseudo-inversion as (6):

$$\begin{aligned} \Psi_F &= \Re_{56F}^{+} \cdot \hat{f}^S \\ \Psi_T &= \Re_{56T}^{+} \cdot \hat{t}^S \end{aligned} \tag{6}$$

$\Re_{56F}^{+}, \Re_{56T}^{+}$ – pseudo-inverse of $\Re_{56F}, \Re_{56T}$.

Step 2: Inertia tensor estimation

The inertia tensor is estimated by sequentially rotating the object about its CoG (estimated in the previous step) about three orthogonal axes. The angle about an axis has sinusoidal time law.

The inertia tensor is calculated from the force/torque measurements and the angular velocities/accelerations recorded during each axis rotation.

Eq. (3) can be rearranged as(7) to form the regressor matrix [3]

$$\widehat{M}_i^S - b_T = \Omega_i \cdot I^* + \Pi_i \cdot \vec{r}_{CG}^S \tag{7}$$

$$\Omega_i = \begin{bmatrix} \dot{\omega}_{xi} & (\dot{\omega}_{yi} - \omega_{xi} \cdot \omega_{zi}) & (\dot{\omega}_{zi} + \omega_{xi} \cdot \omega_{yi}) & (-\omega_{yi} \cdot \omega_{zi}) & \left(\omega_{yi}^2 - \omega_{zi}^2\right) & \omega_{yi} \cdot \omega_{zi} \\ \omega_{xi} \cdot \omega_{zi} & (\dot{\omega}_{xi} + \omega_{yi} \cdot \omega_{zi}) & \omega_{zi}^2 - \omega_{xi}^2 & \dot{\omega}_{yi} & \dot{\omega}_{zi} - \omega_{xi} \cdot \omega_{yi} & -\omega_{xi} \cdot \omega_{zi} \\ -\omega_{xi} \cdot \omega_{yi} & \omega_{xi}^2 - \omega_{yi}^2 & \dot{\omega}_{xi} - \omega_{yi} \cdot \omega_{zi} & \omega_{xi} \cdot \omega_{yi} & \dot{\omega}_{yi} + \omega_{xi} \cdot \omega_{zi} & \dot{\omega}_{zi} \end{bmatrix};$$

$$I^* = \begin{bmatrix} I_{11} & I_{12} & I_{13} & I_{22} & I_{23} & I_{33} \end{bmatrix}^T;$$

I_{ii} – diagonal elements of the inertia tensor;
I_{ij} – off-diagonal elements of the inertia tensor;

$$\Pi_i = \begin{bmatrix} 0 & F_z - b_3 & -F_y + b_2 \\ -F_z + b_3 & 0 & F_x - b_1 \\ F_y - b_2 & -F_x + b_1 & 0 \end{bmatrix}$$

Again stacking all the measurement in one be vector one can obtain (8)

$$\widehat{M}^S = \widehat{\Omega} \cdot I^* + \widehat{\Pi} \tag{8}$$

where

$$\begin{aligned} \widehat{M}^S &= [\widehat{M}_1^S - b_T \cdots \widehat{M}_n^S - b_T]^T \\ \widehat{\Omega} &= [\Omega_1 \cdot \vec{r}_{CG}^S \cdots \Omega_n \cdot \vec{r}_{CG}^S]^T \\ \widehat{\Pi} &= [\Pi_1 \cdots \Pi_n]^T \end{aligned}$$

From eq. (7) and (8) using pseudo-inversion the inertia tensor components are expressed as (9)

$$I^* = \left[\widehat{\Omega}^T \cdot \widehat{\Omega}\right]^{-1} \cdot \left[\widehat{\Omega}^T \cdot \left(\widehat{M}^S - \widehat{\Pi}\right)\right] \tag{9}$$

Error sources analysis

According to (6) and (9) the errors in the inertia parameters estimated are mostly due to two underlying measurements errors of the kinematic motion parameters and the force/torque measurements.

The first source is due to trajectory following errors and finite angle sensor resolution and noise as well as other mechanical uncertainties.

The second source is FTS resolution combined with FTS calibration errors and non-linearities.

Analysis of influence of the force/torque measurement errors onto the mass and CoG estimation

The model of the force/torque measurements is depicted in fig.2.

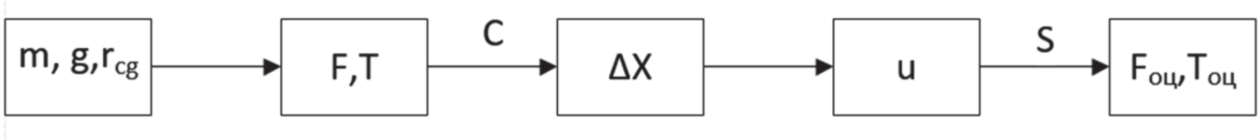


Figure 2 – Force/torque measurement model

m – object mass;
g – gravity vector;
rCG – CoG radius from FTS frame;
F – vector of actual forces acting at the output flange of the FTS at XsYsZs frame;
T – vector of actual torques acting at the mounting point of the FTS at XsYsZs frame;
C – stiffness matrix;
ΔX – vector of FTS mechanical deformations;
u – voltages at Whitstone bridges of the tenso-elements;
S – calibration matrix;
Fоц,Tоц – force/torque estimates at the at the mounting point of the FTS.

This model give inside into the causality of physical and information processes of FTS measurements with attached payload of mass m, center of gravity at r_{CG}:

The payload act on FTS with the force F ant torque M causing the deformations ΔX of the tenso-elemets. Customly this deformation are considered linear in the applied force/torque and related via the stiffness matrix C. However, linear approximation does not account for non-linear effect such as material hysteresis, temperature deformations and other physical factor changing the resistance of the tenso-elements.

The deformation of the tenso-elements are transduced into electrical signals in Whitstone briges and preamplified before the analog-to-digital conversion. The parameters of preamplifiers and passive elements in electrical circuits are temperature dependent and generally different for each force/torque component. The electrical signals and the force/torque estimates are related through the calibration matrix S.

According to the description above the following notes about the possible error sources in the force/torque measurements seems:

1) the assumption about the relationship between the acting force/torque and the mechanical deformation to be linear and constituted by the stiffness C is only the first order approximation;

2) the calibration matrix S relating the electrical equivalents of the deformations to the estimates of the force/torque may be determined not exactly. Furthermore, the relationship itself is linear only.

Let

F_k – actual force vector at k-th measurement;
T_k – actual torque vector at k-th measurement;
ΔF_k – force error vector at k-th measurement;
ΔT_k – torque error vector at k-th measurement.

Thus the FTS measurements vectors $\widehat{F}_k, \widehat{T}_k$ are (10)

$$\begin{aligned} \widehat{F}_k &= F_k + \Delta F_k \\ \widehat{T}_k &= T_k + \Delta T_k \end{aligned} \tag{10}$$

Eq. (4) takes the from (11)

$$\begin{aligned} F_k + \Delta F_k &= U_{Fk}(\alpha_5, \alpha_6) \cdot \Psi_F \\ T_k + \Delta T_k &= U_{Tk}(\alpha_5, \alpha_6) \cdot \Psi_T \end{aligned} \tag{11}$$

Staking the measurements in one big vector like in (5) equations (12) are easily obtained:

$$F + \Delta F = \Re_{56F} \cdot \Psi_F \tag{12}$$

$$T + \Delta T = \Re_{56T} \cdot \Psi_T$$

From eq. (12) the required inertia parameters are expressed as (13)

$$\Psi_F = \Re^+_{56F} \cdot F + \Re^+_{56F} \cdot \Delta F$$
$$\Psi_T = \Re^+_{56T} \cdot T + \Re^+_{56F} \cdot \Delta T \tag{13}$$

From (13) it can be deduced that the influence of the FTS measurement errors on the inertia parameters estimates is linear and in general can not be compensated by some acquisition scheme. However, for the large enough data set this influence reduces to zero provided zero mean assumption for ΔF_k, ΔT_k.

Analysis of influence of the positioning errors onto the mass and CoG estimation

The error due to kinematic uncertainties for the mass and CoG estimation represents itself in the orientation errors. The orientation errors modeled as the uncertainties in the tangent and yaw joints angular positions. The actual angles of the joints are (14)

$$\widehat{\alpha}_{5i} = \alpha_{5i} + \Delta\alpha_5$$
$$\widehat{\alpha}_{6i} = \alpha_{6i} + \Delta\alpha_6 \tag{14}$$

where

$\Delta\alpha_5, \Delta\alpha_6$ – joint position errors;

α_5, α_6 – required (or expected) angle values;

$\widehat{\alpha}_{5i}, \widehat{\alpha}_{6i}$ – actual angles

Using (14) eq. (4) can be written as (15)

$$\hat{F}_i^S = U_{Fi}(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6) \cdot \Psi_F = \left[({}^S R_b(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6))^T \cdot g_b, I_{3x3} \right] \cdot \Psi_F$$
$$\hat{T}_i^S = U_{Ti}(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6) \cdot \Psi_T = \left[m \cdot S(({}^S R_b(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6))^T \cdot g_b), I_{3x3} \right] \cdot \Psi_T \tag{15}$$

Considering that combine rotation about the same axes can be presented as a sequential multiplication of rotation matrices and also assuming that $\Delta\alpha_5, \Delta\alpha_6$ are small enough to replace all trigonometric functions with their linear approximations, matrix ${}^S R_b(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6)$ can be approximated as (16)

$${}^S R_b(\alpha_{5i} + \Delta\alpha_5, \alpha_{6i} + \Delta\alpha_6) = {}^S R_b(\alpha_{5k}) \cdot {}^S R_b(\alpha_{5k}) + {}^S R_b(\alpha_{5k}) \cdot \Delta R_{56} \cdot {}^S R_b(\alpha_{6j}) \tag{16}$$

$$\Delta 56 = \begin{bmatrix} -1 + \Delta\alpha_6 & -1 - \frac{(\Delta\alpha_6)^2}{2} & 0 \\ 1 + \frac{(\Delta\alpha_6)^2}{2} & -2 - \Delta\alpha_5 + \Delta\alpha_6 & 1 + \frac{(\Delta\alpha_5)^2}{2} \\ 0 & -1 - \frac{(\Delta\alpha_5)^2}{2} & -1 - \Delta\alpha_5 \end{bmatrix}$$

Remember the property of a skew-symmetric matrix S (17)

$$S(R \cdot a) = R \cdot S(a) \cdot R^T \tag{17}$$

Under aforementioned assumption and using (16), (17) eq. (15) is expressed as (18)

$$\widehat{f}_i^S = \widehat{\mathbb{R}}_{Fi} \cdot \Psi_F + \Delta\widehat{\mathbb{R}}_{Fi} \cdot \Psi_F = \widehat{\mathbb{R}}_{Fi} \cdot \Psi_F + \Delta\Re_{Fi} \cdot m \tag{18}$$

$$\hat{t}_i^S = \widehat{\mathbb{R}}_{Ti} \cdot \tilde{\Psi}_T + \Delta\widehat{\mathbb{R}}_{Ti} \cdot \tilde{\Psi}_T = \widehat{\mathbb{R}}_{Ti} \cdot \tilde{\Psi}_T + \Delta\Re_{Ti} \cdot \vec{r}_{CG}^{\,S}$$

$$\widehat{f}_i^S = \left({}^S R_b(\alpha_{6k})\right)^T \cdot \hat{F}_i^S$$

$$\widehat{\mathbb{R}}_{Fi} = \left[\left({}^S R_b(\alpha_{5k})\right)^T \cdot g_b, {}^S R_b(\alpha_{6k})\right]$$

$$\Delta\Re_{Fi} = \left[\Delta\Re_{Fi}, 0_{3x3}\right] = \left[\left(\Delta R_{56}\right)^T \cdot \left({}^S R_b(\alpha_{5k}) \cdot\right)^T \cdot g_b, 0_{3x3}\right]$$

$$\widehat{\mathbb{R}}_{Ti} = \left[m \cdot S\left[\left({}^S R_b(\alpha_{5k})\right)^T \cdot g_b\right] \cdot {}^S R_b(\alpha_{6k}), \left({}^S R_b(\alpha_{6k})\right)^T\right]$$

$$\Delta\Re_{Ti} = \left[\Delta\Re_{Ti}, 0_{3x3}\right] = \left[m \cdot S\left[\left(\Delta R_{56} \cdot {}^S R_b(\alpha_{6j})\right)^T \cdot g_b\right] \cdot {}^S R_b(\alpha_{6k}), 0_{3x3}\right]$$

$$\hat{t}_i^S = {}^S R_b(\alpha_{6k})^T \cdot \widehat{T}_i^S$$

i – measurement index.

Stacking again all the measurement vectors in one big vector gives (19)

$$\widehat{f}^S = \widehat{\mathbb{R}}_F \cdot \Psi_F + \Delta\widehat{\mathbb{R}}_F \cdot \Psi_F = \widehat{\mathbb{R}}_F \cdot \Psi_F + \Delta\Re_F \cdot m$$
$$\hat{t}^S = \widehat{\mathbb{R}}_T \cdot \tilde{\Psi}_T + \Delta\widehat{\mathbb{R}}_T \cdot \tilde{\Psi}_T = \widehat{\mathbb{R}}_{Ti} \cdot \tilde{\Psi}_T + \Delta\Re_T \cdot \vec{r}_{CG}^{\,S} \tag{19}$$

with $\widehat{\mathbb{R}}_F = \begin{bmatrix} \widehat{\mathbb{R}}_{F1} \\ \vdots \\ \widehat{\mathbb{R}}_{Fn} \end{bmatrix}, \widehat{\mathbb{R}}_T = \begin{bmatrix} \widehat{\mathbb{R}}_{T1} \\ \vdots \\ \widehat{\mathbb{R}}_{Tn} \end{bmatrix}, \Delta\Re_T = \begin{bmatrix} \Delta\Re_{T1} \\ \vdots \\ \Delta\Re_{Tn} \end{bmatrix}, \Delta\Re_F = \begin{bmatrix} \Delta\Re_{F1} \\ \vdots \\ \Delta\Re_{Fn} \end{bmatrix}$

Applying pseudo-inversion in (19) the values for the mass and the CoG vector components as well as FTS biases are obtained as (20)

$$\widehat{\Psi}_F = \Psi_F + \Delta\Psi_F = \Psi_F + (\widehat{\mathbb{R}}_F^T \cdot \widehat{\mathbb{R}}_F)^{-1} \cdot \widehat{\mathbb{R}}_F \cdot \Delta\Re_F \cdot m$$
$$\widehat{\Psi}_T = \Psi_T + \Delta\Psi_T = \Psi_T + (\widehat{\mathbb{R}}_T^T \cdot \widehat{\mathbb{R}}_T)^{-1} \cdot \widehat{\mathbb{R}}_T \cdot \Delta\Re_T \cdot \vec{r}_{CG}^{\,S} \tag{20}$$

where

$\widehat{\Psi}_F$ – vector containing the estimates of the mass and the FTS biases for the force components;

Ψ_F – vector containing exact mass and the FTS biases for the force components;

$\Delta\Psi_F$ – errors in estimates of the mass and the FTS biases for the force components;

$\widehat{\Psi}_T$ – vector containing the estimates of the CoG vector components and the FTS biases for the torque components;

Ψ_F – vector containing exact the CoG components and the FTS biases for the torque components;

$\Delta\Psi_F$ – error vector of the CoG components and the FTS biases for the torque components estimates;

From eq. (20) following observation are immediately available:

1) the errors in the estimates of the FTS biases for the force/torque components does not depend on their actual values;

2) the errors in the estimates of the the mass and CoG vector are proportional to their actual values;

3) it can be shown by expanding the expression for $(\widehat{\mathbb{R}}_F^T \cdot \widehat{\mathbb{R}}_F)^{-1} \cdot \widehat{\mathbb{R}}_F \cdot \Delta\Re_F$ and $(\widehat{\mathbb{R}}_T^T \cdot \widehat{\mathbb{R}}_T)^{-1} \cdot \widehat{\mathbb{R}}_T \cdot \Delta\Re_T$ in (20), that the minimum of the CoG and FTS biases errors is attained when the values set of α_5, α_6 is symmetric with respect to zero. This means that the points of the manipulator end-effector are placed uniformly and symmetrically on a sphere/hemisphere.

Experimental evaluation of the proposed method for the mass and CoG estimation

This section is devoted to the experimental verification of the theoretical analysis presented in the previous sections. The verification is performed in two ways:

1) Computer simulation of the error propagation;

2) Experiments with the hardware.

Simulation

The errors introduced by the uncertainties in the orientation angles are evaluated by numerically calculating the differences of corresponding values at particular point $\widehat{\alpha_{5i}}, \widehat{\alpha_{6i}}$ from the same values at the exact positions α_5, α_6 given the equally spaced network of the joint position errors $\Delta\alpha_5, \Delta\alpha_6$.

The typical result are shown in the figures 3-6. The figures shows qualitative

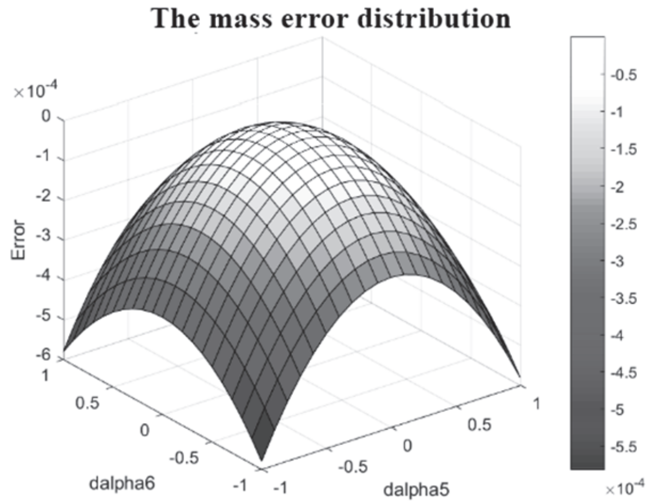


Figure 3 – The mass error distribution against the errors in tangent and yaw joints positions

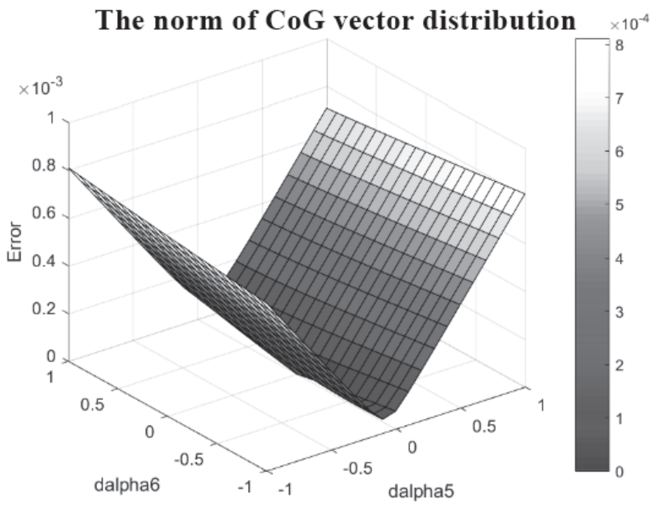


Figure 4 – The norm of CoG vector distribution against the errors in tangent and yaw joints positions

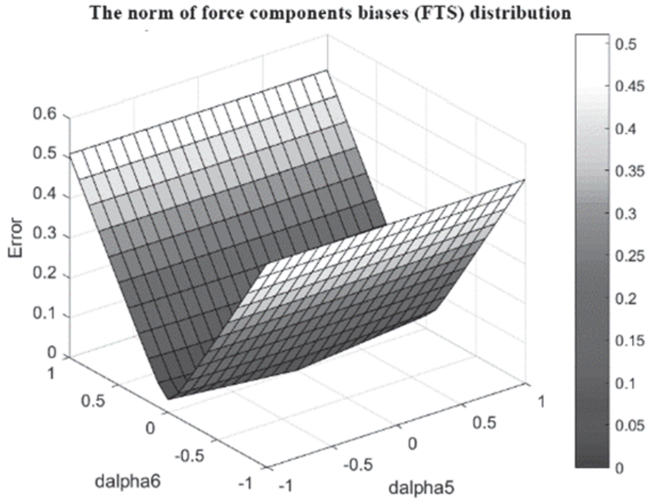


Figure 5 – The norm of force components biases (FTS) distribution against the errors in tangent and yaw joints positions

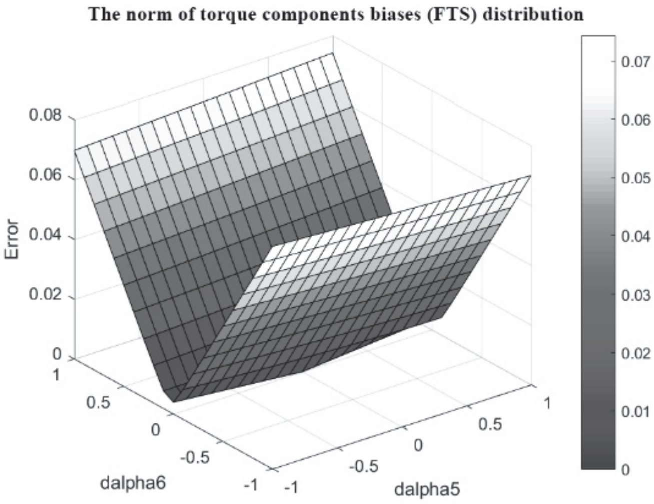


Figure 6 – The norm of torque components biases (FTS) distribution against the errors in tangent and yaw joints positions

Analyzing fig. 3-6 the following observation can be made:

1) Fig. 3 shows that the errors in tangent and yaw joints positions have similar influence on the mass estimation;

2) Fig. 4 reveals that tangent joint position error gives greater add-up to the CoG estimation error than that of the yaw joint;

3) Figures 5-6 illustrates the opposite situation: the yaw joint error gives greater rise in force/torque biases than that of the tangent joint.

The dependence of the errors on the value of the parameters are shown in fig. 7-10.

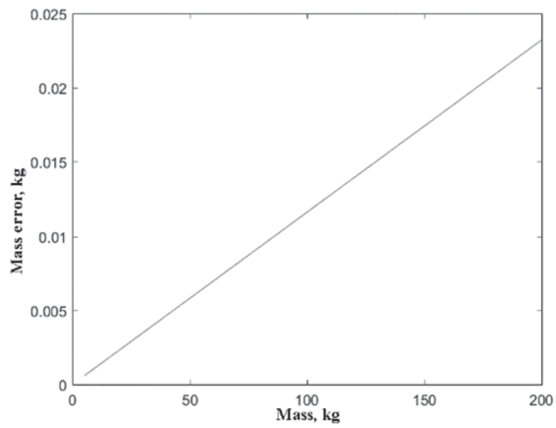


Figure 7 – The error in mass determination against the object mass

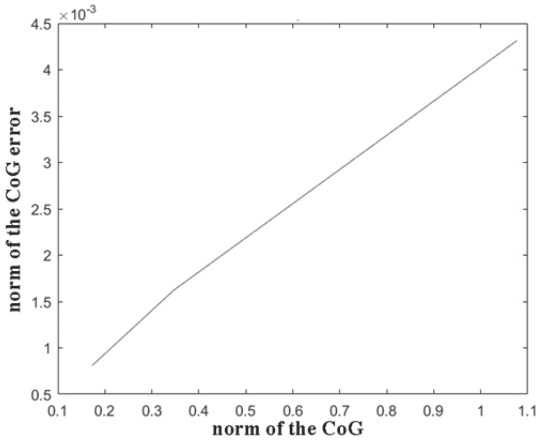


Figure 8 – The norm of the CoG vector error against the norm of the CoG vector

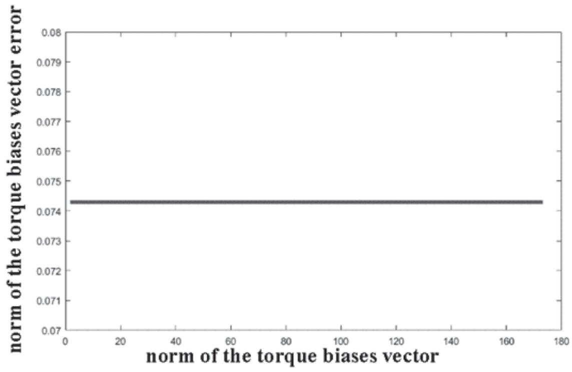


Figure 9 – The norm of the torque biases vector error against the torque biases vector

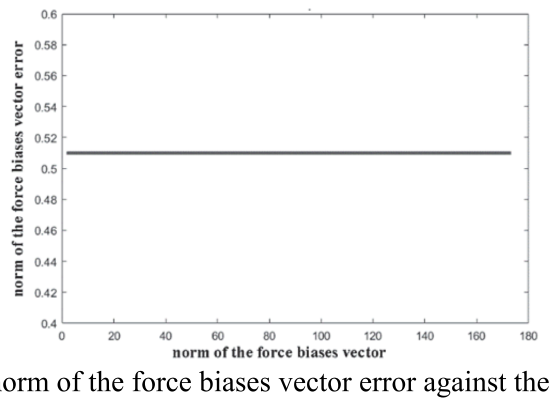


Figure 10 – The norm of the force biases vector error against the force biases vector

Brief analysis of the fig. 7-10 confirms the theoretical result in the previous sections.

Simulation also confirmed that any systematic error in the force/torque measurements will be accumulated in the FTS biases. The random errors in this measurements can't be mitigated by any particular orientation set but the increasing the number of points has shown its steady decline.

Experimental estimation of mass and center of gravity

Experimental setup is shown in fig.11. It consists of the industrial manipulator KukaKR300R2500 Ultra equipped with specialized flange with orthogonal rods. Each axes has screw for mounting a specialized weights at particular distances measured with engraved rulers. All coordinate systems are given in fig. 11. Z axis is collinear with the yaw joint axis and the mounting rod axis. Y and X axes are parallel to the corresponding rod axes and lay in the appropriate plane of the FTS measurement axes.

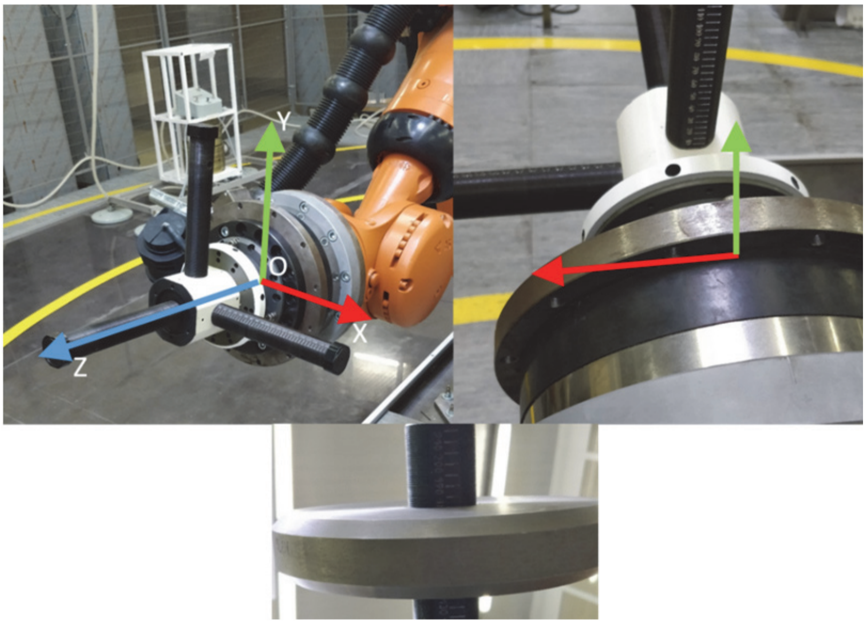


Figure 11 – Experimental setup

The first series of experiment consist in differential evaluation of the CoG estimation accuracy. The goal is to estimate the method accuracy by shifting the CoG vector by the known value and comparing the shift vector with the calculation results. This allows to refrain from actual knowledge of the CoG vector and work only with its relative shifts.

The measured object was placed at three different positions on the X axis at given distances. The results are presented in table1.

Table 1. First experiments results

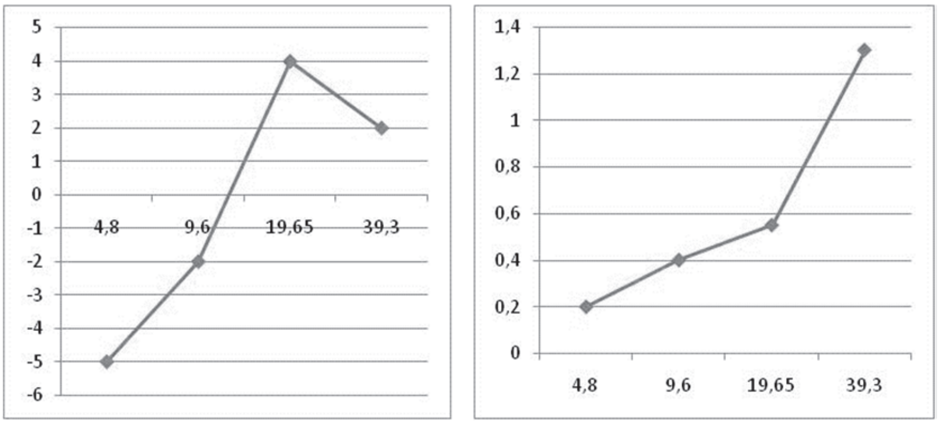
	Expected values				**Actual values**			
№	X,mm	Y,mm	Z,mm	M,kg	X,mm	Y,mm	Z,mm	M,kg
1	245	0	160	4,8	242	2	162	5
Δx	55				55			
2	190	0	160	4,8	187	1	162	5
Δx	60				64			
3	120	0	160	4,8	123	0	163	5

Table 1 shows the maximum deflection of the CoG vector from the expected value to be 3 mm that can be explained by difficulties in the payload positioning and its final pose estimation by the rulers. The mass error is explained by the actual FTS resolution being about the error magnitude. The FTS hysteresis also takes place.

The second series of experiment is devoted to evaluate the influence of the true mass value on the accuracy of the mass and CoG estimation. In the experiment the CoG vector was kept constant while the mass was variable. The results are given in table 2 and fig. 12.

Table 2. Second experiment result

	Expected values				Actual values				
№	X,mm	Y,mm	Z,mm	M,kg	X,mm	Y,mm	Z,mm	M,kg	ΔM
1	0	0	370	4,8	-2	-2	375	5	0,2
2	0	0	370	9,6	-0	-1	372	10,0	0,4
3	0	0	370	19,65	0	0	366	20,2	0,55
4	0	0	370	39,3	0	0	368	40,6	1,3



a) b)

Figure 11 – Table 1 graphical representation: a) Z axis component of CoG vector error; b) mass error. Both a) and b) are againt the actual mass in kg

The experiment confirmed the theoretical prognosis about dependence of the mass estimation error on the actual mass.

Conclusion

The paper presents the analysis of the error sources in the mass and center of gravity estimation procedures. The method of estimation is based on utilizing an industrial manipulator equipped with a force/torque sensor. Theoretical analysis gives qualitative relationships between the parameters of the system and the expected errors. Theoretical results are verified by computer simulation and experiment with the system prototype.

References

1. Hanssen S., Hovland G. E., Brogardh T. Verification of Physical Parameters in a Rigid Manipulator Wrist Model //the 3rd Imacs Symposium on Mathematical Modelling. – P. 849-855.
2. Gaz C., Flacco F., De Luca A. Extracting feasible robot parameters from dynamic coefficients using nonlinear optimization methods //2016 IEEE international conference on robotics and automation (ICRA). – IEEE, 2016. – P. 2075-2081.
3. Khalil W., Gautier M., Lemoine P. Identification of the payload inertial parameters of industrial manipulators //Proceedings 2007 IEEE International Conference on Robotics and Automation. – IEEE, 2007. – P. 4943-4948.
4. Schedlinski C., Link M. A survey of current inertia parameter identification methods //Mechanical systems and signal processing. – 2001. – V. 15. – №. 1. – P. 189-211.
5. Gobbi M., Mastinu G., Previati G. A method for measuring the inertia properties of rigid bodies //Mechanical Systems and Signal Processing. – 2011. – V. 25. – №. 1. – P. 305-318.
6. Jamisola R. S., Dadios E. P. Identifying moments of inertia parameters for rigid-body manipulators //Int. Conf. Mechatron. Technol., Cebu City, Philippines. – 2009.
7. Sivukhin D.V. The General Course of Physics. Volume 1. Mechanics // Uch. pos. Fizmatlit. - 1989.
8. Barreto J. P., Muñoz L. E. Low uncertainty method for inertia tensor identification //Mechanical Systems and Signal Processing. – 2016. – V. 68. – P. 207-216.

Ji Sup Yoon, Youngsoo Choi, Kyung-Min Jeong, Jongwon Park

RESEARCH WORKS OF EMERGENCY RESPONSIVE ROBOTS AT KAERI

Nuclear Robotics Lab, KAERI, Daejeon, Korea
jsyoon@kaeri.re.kr, yschoi1@kaeri.re.kr,
kmjeong@kaeri.re.kr, jwpark@kaeri.re.kr

Abstract

In 2012, the Nuclear Robotics Lab (NRL) of Korea Atomic Energy Research Institute (KAERI) initiated research for an unmanned emergency-response robotics system. To date, it has developed various types of robotics systems based on the NPP (Nuclear Power Plant) maintenance experiences. In this paper, its recent work is introduced, including an all-terrain vehicle (ATV) and a remote controlled forklift, a radiation monitoring drone, a range-gated image (RGI) camera system, and a heavy duty mobile manipulator.

Keywords: emergency, response, robot, KAERI

In 2012, the Nuclear Robotics Lab (NRL) of Korea Atomic Energy Research Institute (KAERI) initiated research for an unmanned emergency-response robotics system. To date, it has developed various types of robotics systems based on the NPP (Nuclear Power Plant) maintenance experiences. In this paper, its recent work is introduced, including an all-terrain vehicle (ATV) and a remote controlled forklift, a radiation monitoring drone, a range-gated image (RGI) camera system, and a heavy duty mobile manipulator.

Radiation monitoring system

The NRL of KAERI has proposed a rapid radiation monitoring system [1]. The system comprises an ATV, a drone, and two radiation detectors (Fig. 1). The ATV has been modified for remote control and has excellent speed and adaptability to various types of terrains. Moreover, it can be used to carry the drone to a site to be monitored. The drone is equipped with two radiation detectors, and can measure radiation and send radiation data to a control station or smart phone. The two radiation detectors were designed to have light weight so that they can be carried by the drone. Each detector incorporates both low and high range Geiger Muller counters for wide range measurement. The measuring range of each detector is 10 uSv/h ~ 100 Sv/h and 1 uSv/h ~ 100 Sv/h.

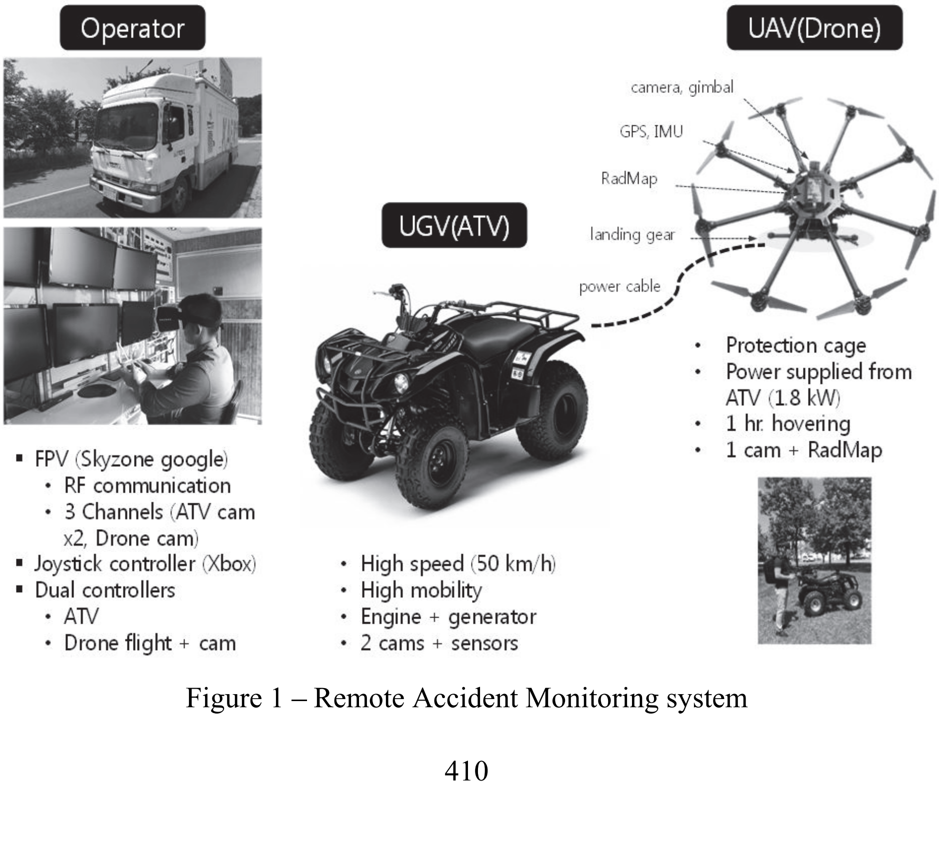


Figure 1 – Remote Accident Monitoring system

Remotely operated forklift

The forklift (Doosan Pro 5) has been modified for remote operation and has a maximum payload of 1.5 tons [2]. Moreover, it has been fitted with seven motors for control of steering, brakes, acceleration, forward/reverse gears, lift, and tilt (Fig. 2). In addition, it has been fitted with two laser range finders, one on each side of the forklift. The forklift can be used to transport heavy objects or to remove debris.

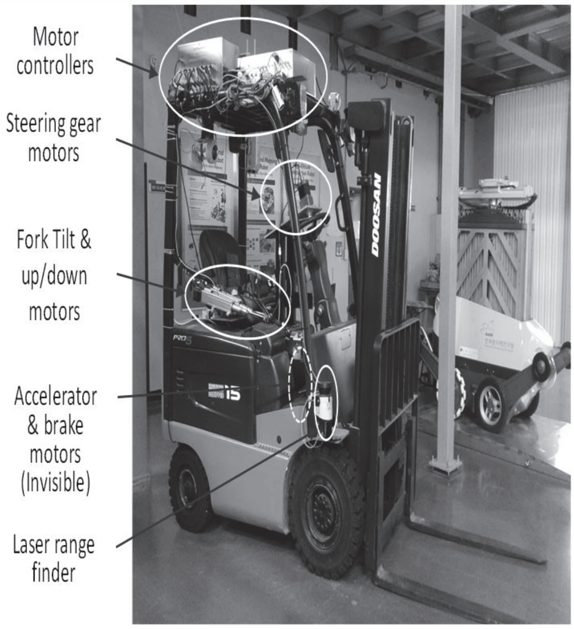
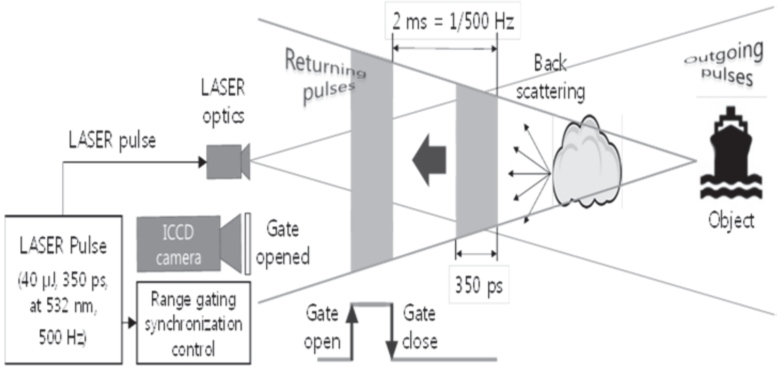


Figure 2 – Remotely operated forklift 1

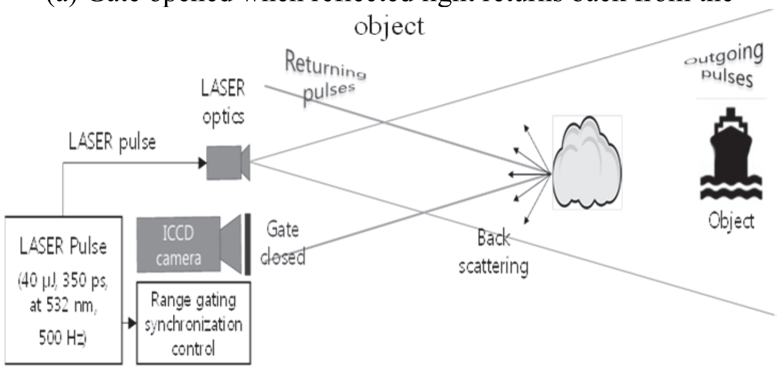
Range-Gated Imaging Camera

In the event of a loss-of-coolant accident at a nuclear power plant (NPP), visibility inside the reactor containment building of the NPP is likely to be poor due to vaporized water. Visibility in a severe accident scenario is likely to be only in the range of 0.4 ~ 4 m [3].

In this study, we sought to utilize RGI technology to obtain visual information in a poor visibility environment. As shown in Fig. 3, the developed RGI camera system uses an ultrashort pulse laser as an illuminator and an intensified CCD (ICCD) camera. Laser light is irradiated to the 'object to be observed', and the ICCD camera receives only that light which is returned from the 'object to be observed': all other light, such as that reflected from particles or objects unrelated to the 'object to be observed', is filtered out. This is achieved by the gate control of the developed RGI camera system, which regulates the opening of the shutter of the ICCD camera at the instant where the returned laser light from the 'object to be observed' reaches the camera. Received light is amplified using the on-board electronics.



(a) Gate opened when reflected light returns back from the object

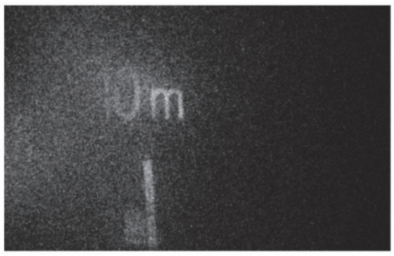


(b) Gate closed when reflected light returns back from the particles located between camera and object

Figure 3 – Operation principle of RGI system

A test result of RGI system is shown in Fig. 4. Fig. 4(a) depicts a CCD camera image of the chess board is not visible, even at a distance of 3 m from the camera. On the other hand, a chess board located 10 m ahead from the ICCD is visible (Fig. 4(b)).

(a) CCD camera



(b) ICCD camera

Figure 4 – Image of CCD(a) and ICCD(b) with fog

References

1. J. Park and Y. S. Choi, "An Aerial and Ground Monitoring System for Nuclear Accidents," Transactions of the Korean Nuclear Society Spring Meeting, Jeju, Korea, 2017.
2. J. Y. Park, et al., Advanced system establishment for nuclear system integrity, KAERI/TR-6750/2017, 2017.
3. J. W. Cho, Y. S. Choi, and K. M. Jeong, "Verification of Range-Gated Imaging Technology under Dense Aerosol Environments," Journal of Institute of Control, Robotics and Systems, Vol. 23, No. 7, pp. 606-617, 2017.

Jongwon Park,Young Soo Choi

HEAVY DUTY DUAL ARM ROBOT FOR DISASTER RESPONSE

Korea Atomic Energy Research Institute, Daejeon, Korea
jwpark@kaeri.re.kr

Abstract

An accident in a nuclear facility causes a great social cost. To prevent an unexpected nuclear accident from spreading to the catastrophic disaster, emergency response action in early stage is required. A small sized high power robotic manipulator can be an appropriate candidate to deal with a wide spectrum of tasks in an emergency situation. In this paper, we discuss about the design of a high power robotic manipulator, which is capable of handling high payloads for an initial response action to the nuclear facility accident.

Keywords: heavy duty, dual arm, robot, disaster, response

Acknowledgments

This paper was supported by the Ministry of Science, ICT & Future Planning.

1. Introduction

An accident in a nuclear facility causes a great social cost. To prevent an unexpected nuclear accident from spreading to the catastrophic disaster, emergency response action in early stage is required. However, high radiation environment has been proved as a challenging obstacle for human workers to access to the accident site and take an action in previous accident cases. Therefore, emergency response robotic technology to be used in a nuclear accident site instead of human workers are actively conducted in domestically and internationally.

Robots in an accident situation are required to carry out a variety of tasks depend on the types and patterns of accidents. An emergency response usually includes removing of debris, make an access road to a certain place and handling valves. These tasks normally involve high payload handling. However existing human-sized robotic manipulators are not appropriate to deal with heavy duty tasks due to the limit of payloads. Thus, a small sized high power robotic manipulator can be an appropriate candidate to deal with a wide spectrum of tasks in an emergency situation.

In this paper, we discuss about the design of a high power robotic manipulator, which is capable of handling high payloads for an initial response action to the nuclear facility accident.

2. Heavy Duty Dual Arm Robot

The size of the robotic arm is an important factor to be utilized in a disaster situation such as nuclear accident. Because most man-made structures are built for human, human-sized robots are ideal for maneuvering freely inside and outside the accident site and handling the tasks performed by workers. In this paper, we propose a robot manipulator with a length of 1 m that simulates the linkage structure of a human arm in order to carry out a disaster response tasks.

The robotic arm should be able to handle heavy loads for various operations required in case of an accident, such as door opening and closing, valve operation, radioactive contamination treatment, and debris removal. Therefore, the robotic arm should be able to produce high power and large torque in a limited size. In order to implement the size and output conditions mentioned above, we designed the robot based on the hydraulic actuators.

The hydraulic actuators can produce approximately 10 times more output compared to electric motors. In addition, the hydraulic actuators can generate large force and torque without a gear reduction, therefore the structure can be simple. Hydraulic actuators are strong against external forces; thus they are suitable for work requiring large load.

In this study, we designed an eight degree of freedom dual arm robot capable of handling objects over 100 kg.

Figure 1 – Heavy duty dual arm robot, ARMstrong

The linkage structure of the ARMstong robot was designed by simulating the human body. For mobility, the caterpillar was adopted. The small mobile hydraulic power pack was installed for the operation of the robot. Fig. 1 shows the concept design of ARMstrong (Accident Response Robot).

3. Simulation

The V-REP robot simulator was used to evaluate the characteristics of the designed robot. In the simulator, the robot's work space, joint torque, and control performance were accessed (Fig. 2). In addition, we constructed a virtual nuclear accident environment and conducted the detail work such as door opening and closing, valve operation, debris removal, hull movement, and transfer of radioactive materials.

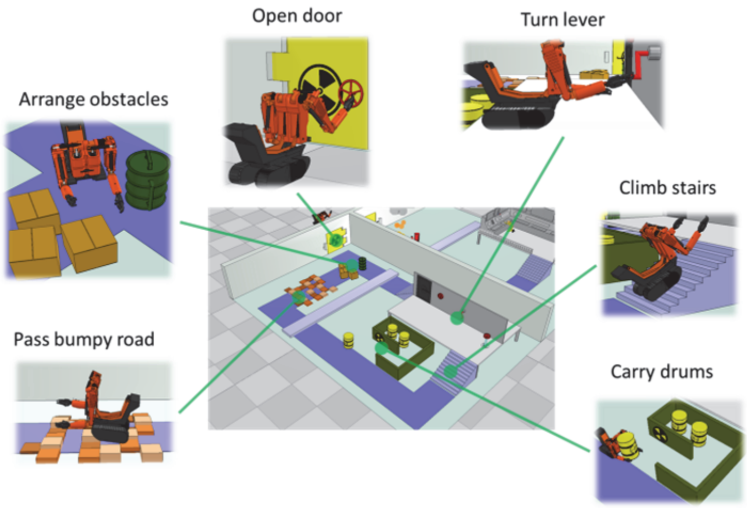


Figure 2 – Armstrong robot simulation in a virtual accident environment

4. Conclusion

In this paper, we discussed about the design and simulation of heavy duty dual arm robot for disaster response in case of a nuclear accident. In the future, detailed design, manufacturing and control system development of robot will be expected.

References

1. Design of a High Power Robotic Manipulator for Emergency Response to the Nuclear Accidents, Jongwon Park, Yeong-Geol Bae, Myoung Ho Kim and Young Soo Choi, Proceedings of the KNS 2016 Autumn Meeting

***Jianghai Li*[1], *V. Promyslov*[2], *K. Semenkov*[2]**

CYBER-PHYSICAL ASSESSMENT OF USING ROBOTS FOR SAFETY OPERATIONS OF A NUCLEAR POWER PLANT

[1]Institute of Nuclear Energy Technology (INET), Tsinghua University, Beijing, China, lijianghai@tsinghua.edu.cn
[2]V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russia, v1925@mail.ru, semenkovk@mail.ru

Abstract
The paper considers the problem of assessment of digital threats for cyber-physical objects of high operational risk – nuclear power plants (NPPs). An assessment method combining an information model with a simulator of physical processes is suggested. The method supports the integration of smart mobile digital assets (e.g., robots). The method scope matches with the standard template of cybersecurity assessment developed by IAEA.
Keywords: cybersecurity, nuclear power plant, simulator, model, discretionary model, cyber, physical.

Introduction

The implication of information technologies into the operation of nuclear power plants (NPPs) is steadily growing. The implementation of digital and computer-based instrumentation and controls (I&C) systems had been proceeding slower than in other technological sectors, for example, petrochemistry and transport because of the conservatism, which is inherent to the atomic industry.

However, recently, the process has gathered pace and found acceptance [1,2,3]. A set of factors contributed to it: increased confidence in digital technologies based on the operation experience in the industry; the toughening competition with other power generation systems and the need to implement more complex control algorithms that are difficult to implement with analog technologies; NPP accidents (like Chernobyl and Fukushima). The latter has highlighted the need for cybernetic autonomous and automated devices in NPP critical events when a human cannot be directly on the object.

However, digital technologies, being indispensable for some critical cases or significantly increasing the operational efficiency, have also a negative impact of operational safety of NPPs. The effect appears in new vulnerabilities of NPP— cyber threats. The threats are mostly similar to the informational threats known to the IT-industry, but they have a peculiarity distinguishing it among the pure informational threats. The peculiarity is: a cyberthreat unfolds to the full extent only by influencing the processes that pass directly in the control object, and digital medium is just the carrier to transfer the malicious influence.

To assess threats of that nature and resist them, one has to use not only computer security methods but also some methods developed for nuclear (technological) safety provisioning for NPP. The paper suggests combined usage of information models [4], which describe NPP cybersecurity architecture at different stages of the NPP's life cycle, and a simulator of physical processes within the NPP.

The integrated security model of the NPP digital I&C system includes the information model [4], and the physical simulator model. The simulator model is under development in a frame work of IAEA project [5].

1. Description of the information model

The component DM is described by the discretional cybersecurity model, which is a simplified information model derived from the model [4].

$DIM =< G^*, OP >$, where $G^* =< \{G_i | i = 1, N\} >$ are all possible system states characterized by security graph G_i with assets $\{O, S\}$ as vertices. O means objects as "passive" actives and S means "active" actives.

Some mobile assets (for instance, robots) that interact with the object either physically (direct access) or via information channels can act as subjects.

The edges represent the binary relation of directed information transfer between assets in the frame of system operation, OP is the set of security graph transformation corresponding to the model (we call it allowed transformations) $G_i =< A, \{\mapsto\} >$.

The entry $a \mapsto b$ means information or right (physical access) transfer from the asset a to the asset b.

2. Description of the simulator

The goal of the simulator is to imitate the I&C systems in NPP, providing a realistic platform for the threat assessment and security control verification. The specialty of I&C systems from IT system should be implemented in the platform. Not only the digital control components can be compromised, but also the control object, the physical process of NPP, can be affected in the simulator.

To meet the above requirements, the simulator could be designed in this architecture, in Fig.1. The simulator is consisting of three major parts, a process simulator of NPP, control components, and a Human-Machine Interface (HMI). The process simulator will simulate the continuous processes of a NPP, offering measurements to the control components, and receiving manipulated variables from the control components. Control components generate manipulated variables from the input measurement based on the control algorithms, set points, and commands from the HMI. The HMI of the simulator will be served as the interfaces of the part of control components, as well as of the part of process simulator. It can display the “original” process states from the process simulator. It can also display the “reported” states from the control components. In this way, the consequences of cyber-attacks on the digital control components, together with the indirect consequences of the physical process through the digital components, can be illustrated on the HMI.

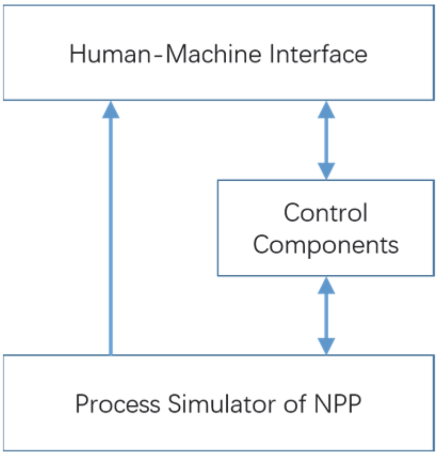


Figure 1 – Architecture of Asherah Simulator

Part of the Asherah Simulator developed by the IAEA Coordinated Research Project (CRP) J02008 [6] can satisfy the above requirements and architecture. The participants of the CRP are 17 institutes from 13 countries. The Asherah NPP Simulator developed by University of São Paulo (USP) is based on the Matlab/Simulink model. It includes primary, secondary & tertiary cooling loops of a hypothetical Pressurized Water Reactor (PWR) named “Asherah”. Control components of the Asherah NPP are hardware of programmable logic controller (PLC) developed by Austrian Institute of Technology (AIT), Framatome, Canadian Nuclear Laboratories (CNL), Korea Atomic Energy Research Institute (KAERI), and Otto von Guericke University Magdeburg (OvGU). The HMI is developed by Tsinghua University connecting with USP’s Simulink model and PLCs via OPC UA protocols.

The reconfigurable feature of Asherah Simulator can support the zoning by IEC 62645 [1].

3. A method of risk assessment with the combined approach

The technique of risk assessment consists in the application of information model for the generation of a tree of attacks on the object with the use of a digital asset followed by “playing” the attack distribution process on the simulator.

The transition between the models may be performed either by an expert or automatically (in case of matching of model interfaces). A standard template developed by IAEA for the project [5] suits for the description of an attack scenario. We provide a tailored and simplified template form (Table 1), where all the fields unrelated to the work have been removed.

Table 1. A standard template for a cybersecurity assessment scenario

Field	Guidance	Model used
ID	*scenario's identifier*	-
Threat Actor		*DIM*
Effect(s)	*Cyber effects on active*	*DIM*

Field	Guidance	Model used
1st order	*Specify effect on targeted component*	*Simulator*
2nd order	*Specify system-level consequence(s)*	*Simulator*
		DIM
3rd order	*Specify enterprise-level consequence(s)*	*Simulator*
		DIM
Exploited vulnerability	*Specify a vulnerability identifier*	*DIM*
Vulnerability reference(s)	*Provide URL reference(s) to identified vulnerability (if applicable)*	*DIM*
Entry point	*Starting point for attack (the root of attack tree)*	*DIM*
Targeted component	*targeting component of active (if applicable)*	*DIM*
NPP resource(s) impacted	*Identify NPP functions or resources affected*	*Simulator*
		DIM
Indicator(s) of compromise	*Identify technical or behavioral conditions associated with the attack*	*Simulator*
		DIM
Risk	*Select from enumerated*	-
Mitigation(s)	*Specify mitigating security control(s)*	*Simulator*
		DIM

Conclusions

The paper considers the problem of assessment of digital threats for cyber-physical objects of high operational risk (like modern NPPs). We suggest a method based on the consistent application of an informational model combined with a simulator of physical processes in the NPP. The advantages of the method are its comprehensive approach which considers both the attack medium (digital content) and the attack object (the real equipment of NPP) and also transparent integration of smart mobile digital assets (for example, robots) into the model. Existing formal security models and physical simulators can be used as the tools of the method. The scope of every model matches with the standard template of cybersecurity assessment developed by IAEA.

References

1. Computer Security at Nuclear Facilities Technical Guidance Reference Manual IAEA Nuclear Security Series. – 2011. – N 17.
2. IEC 62645 Nuclear power plants - Instrumentation and control systems - Requirements for security programmes for computer-based systems, Edition 1, IEC. 2014.
3. Babaev, D.I., Poletykin, A.G., Promyslov, V.G., Timofeev, M.Yu. (2018) Managing cyber security safety of APCS of nuclear power plants. Control Sciences, vol. 3, p. 47-55. (in Russian).
4. Promyslov V., Consideration for formal security models in I&C system design . / Proceedings of the 9th International Conference on Application of Information and Communication Technologies (AICT2015, Rostov on Don). Rostov on Don: IEEE, 2015. C. 188-190. 2015.
5. IAEA, Enhancing Computer Security Incident Analysis at Nuclear Facilities https://www.iaea.org/ru/projects/crp/j02008.
6. Michael T. Rowland1, Mislav Findrik, and Paul Smith. Computer Security Incident Response and Analysis for Nuclear Facilities, https://www.energypact.org/wp-content/uploads/2018/03/FindrikMislavAndSmithPaulCaseStudy.pdf.

A.V. Zhukov[1], V.V. Prikhodko[1], V.V. Svetukhin[2], A.A. Sobolev[1], E.M. Chavkin[1], A.N. Fomin[1],P.E. Kapustin[1], V.E. Kiryukhin[1], V.V. Levshchanov[1]

A ROBOTIC COMPLEX FOR HOT CELLS AND A TRAINING SIMULATOR

[1] S.P. Kapitsa Technological Research Institute of Ulyanovsk State University, Ulyanovsk, Russia, vp@kapitsa.tech
[2] SMC «Technological Center», Zelenograd, Moscow, Russia, tc@tcen.ru.

Abstract
The results of developing a radiation-hardened robotic complex to work in hot cells at nuclear industry facilities as well as a training simulator which is a virtual copy of the production site, are presented. The robotic complex includes the following original constituting elements: a robotic manipulator, a control device with force feedback, a control server with software. The training simulator is a hardware and software complex, the software part of which is identical to that used for the robotic arm control. The software has been developed on a modular basis and contains both original and the state-of-the-art components. The simulator hardware consists of an original control device with feedback, VR equipment and a control server.

Keywords: robotic complex, manipulator, control device, force feedback, joystick, radiation-protected chamber, hot cell, training simulator, virtual reality, VR, control software.

Acknowledgments
The work was supported by the Ministry of Science and Higher Education of the Russian Federation, project RFMEFI57417X0173.

1. Robotic complex

Mechanical master-slave manipulators (MSM) are widely used at nuclear facilities for performing technological operations in hot cells. The undeniable advantages of such manipulators are their high reliability and relative simplicity of design. At the same time manipulators of this type have a limited set of operations like capture, retention, and movement of an object in the horizontal and vertical planes. At the moment, in Russia, this area is practically not covered by robotization which could expand the capabilities of MSMs by adding comfortable remote control, possibility to work with heavy loads, automation of routine tasks. There exist robotic manipulators for the nuclear industry on the market such as A1000, Telbot by HWM [1] but their use in Russia is limited due to their relatively high price.

The need to expand the technological capabilities of MSMs while maintaining high radiation resistance and moderate cost allowed us to formulate the task of developing a hardware-software complex that includes a robotic manipulator, a control device with force-moment feedback and control software. The generalized diagram of the robotic complex is shown in Figure 1.

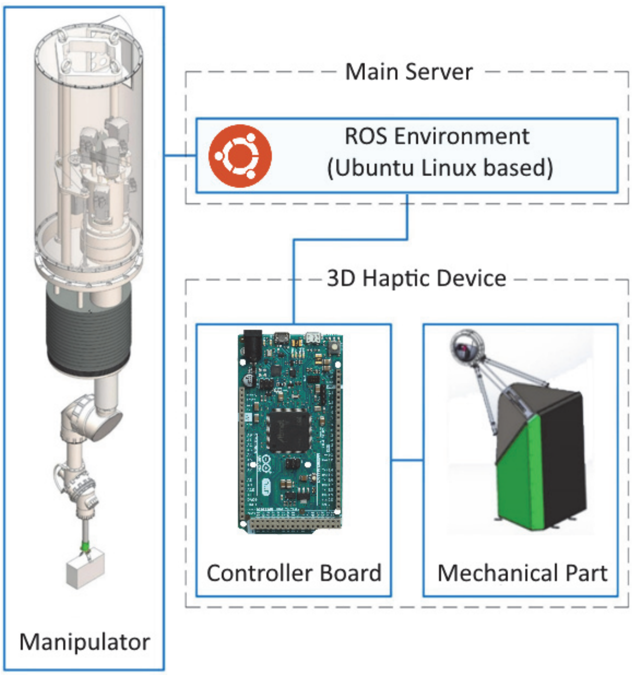


Figure 1 – The generalized diagram of the robotic complex

The control scheme of the robot contains three levels:

a) a robotic arm with a control device, sensors and actuators;

b) a control server;

c) the EtherCAT real-time computer network.

The robotic complex consists of the following elements: a robotic manipulator, a control device with force feedback; a control panel with hardware buttons and a touch screen; a server with control software; a servo control cabinet; a power cabinet.

The manipulator is a 6-link 6-DOF robotic arm with a 2-finger gripper having the following features:

1) To ensure high resistance to ionizing radiation, the drives are placed in the base of the robot arm outside the hot cell and are made in the form of separate quick-detachable units.

2) Mechanical transmission of rotational motion from the drive unit to the wave gearbox is implemented through coaxial gear shafts and backlash-free bevel gears to reduce the positioning errors of the robotic arm.

3) Resolvers are used as feedback sensors in drive units. They are less susceptible to the damaging effects of ionizing radiation in comparison with semiconductor optical encoders.

4) The object moving distance is the area inside the cubic hot cell with dimensions of 1m x 1m x 1m.

At the design stage of the robotic arm, a series of numerical experiments were performed in the ANSYS software. With the help of the ANSYS Rigid Body Dynamics module, a comprehensive simulation analysis of the kinematic scheme of the structure was carried out, its stability was evaluated and the occurrence of mechanically forbidden states was excluded. Strict requirements for technological equipment designed to operate at high levels of ionizing radiation impose serious restrictions on the choice of structural materials [2]. Since the body of the manipulator and most of the elements of its kinematic scheme are made of stainless steel, the problem of maximally reducing the mass of the structure without deteriorating its strength characteristics was solved. The results of modeling in the ANSYS mechanical environment can be presented in the form of pictures of stress-strain states of elements of the kinematic scheme under operating conditions and maximum load. As an example, Fig. 2 shows a picture of the stress-strain state of the manipulator body under a load of 6kg on the gripper.

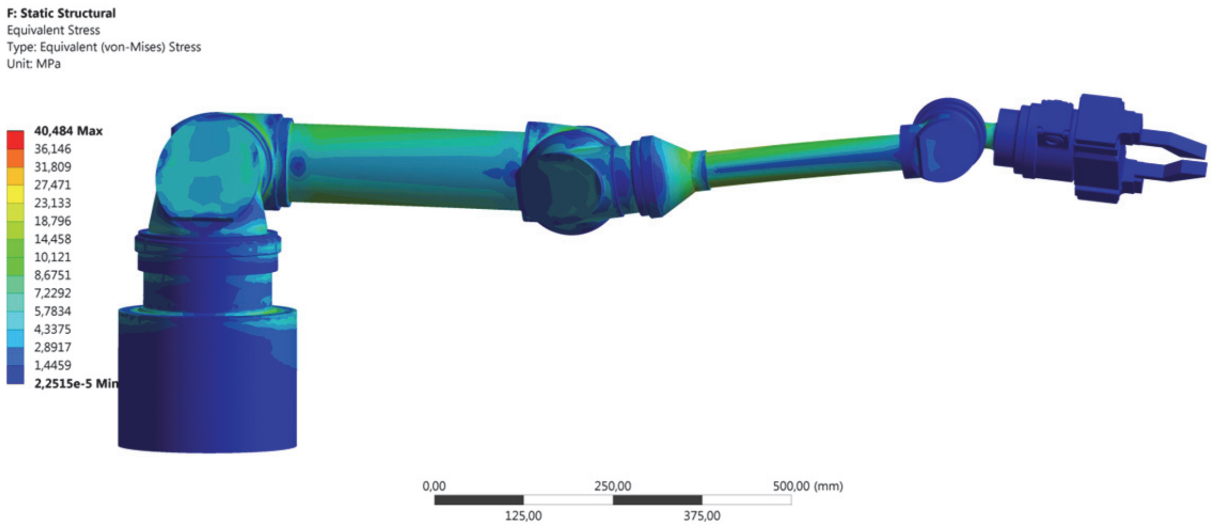
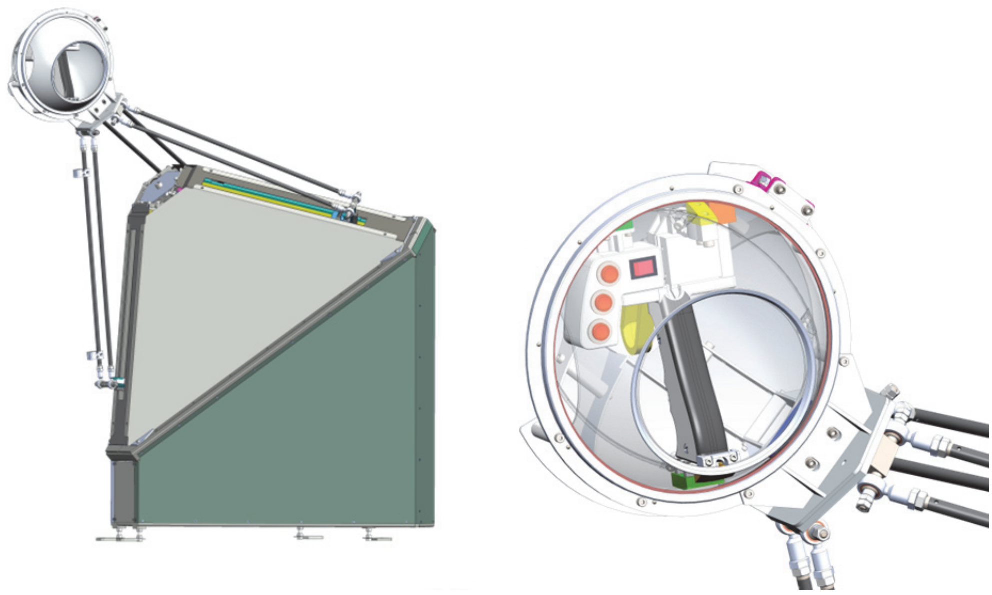


Figure 2 – A picture of the stress-strain state of the manipulator body in the frontal projection

The control device has been designed in such a way that, on the one hand, to provide ergonomics close to that of MSMs, and on the other hand, to expand the functionality of MSMs. The platform with the joystick unit has six degrees of freedom and provides force feedback. The control device prototype with 6 degrees of freedom is a modified version of the Stewart platform [3, 4] with three electric linear actuators (Fig. 3a). The design is based on the principle of parallel kinematics due to which the closed kinematic chain provides high rigidity of the system and accuracy of geometric movements with a relatively small mass of moving parts. The platform provides three degrees of freedom. Additional three degrees of freedom are provided by the joystick unit (Fig. 3b). The mechanism of translational motion is represented by ball-screw gears driven by servomotors. In addition to the movement, the control unit provides imitation of resistance to movement when the specified boundaries of the working area are reached, as well as accelerated or slowed movement of the platform depending on the force applied to the load cells applied by the operator.



(a) (b)

Figure 3 – Electromechanical control device for the robot manipulator (3D-model):
(a) – general view; (b) – joystick unit inside a rotating sphere

The control server was assigned to implement the following tasks:

a) collecting data from the control device and providing feedback to the operator in real time;
b) collecting data about the current state of the manipulator and errors;
c) controlling the actuators of the robot in real time;
d) providing a graphical operator interface.

The control software for the robotic complex was designed on a modular basis and employed both original and the state-of-the-art program components. Ubuntu 16.04 OS with Linux-RT 4.4.0 real-time kernel was installed on the server. The direct and inverse kinematics problem, dynamics problem and collision problem were solved using ROS Kinetic middleware and its ROS-MoveIt and ROS-Industrial extensions.

The control software for the robotic complex was designed on a modular basis and employed both original and the state-of-the-art program components. The functional diagram of the motion control software modules is shown in Figure 4. The dashed line in Figure 4 highlights the subroutines that are part of the developed ROS node. The node receives information from the controller board about the position of the moving elements of the control device and computes the current position in Cartesian coordinates. Then it checks whether the joystick is in the workspace. In case of moving beyond this region, the algorithm forms resistance (feedback) to the movement of the operator's hand. The debug module allows for testing and calibration procedures.

Information about the spatial position of the platform with a joystick installed on it, analog signals from load cells and control commands are transmitted to an Atmel SAM3X8E microcontroller with ARM Cortex-M3 architecture. Original algorithms for processing the information were developed that solve the direct and inverse kinematic problems of the tripod device. In addition to the low-level task of managing servomotors, the microcontroller provided two-way communication channel with the ROS middleware using the UART protocol.

The Qt / PyQt library was the main tool for implementing the operator's GUI. The program code of the robot control process was implemented in C / C ++ / Python. Among the third-party open-source software modules used in the development of the RTK, the following can be highlighted: MoveIt for solving the inverse kinematic problem, Gazebo simulator with its powerful physical solver, and RViz as the visualization tool. These allowed us to carry out a preliminary simulation of the movements of the manipulator along the optimal trajectory in a limited space of the hot cell. The modular approach to managing a robotic arm allowed to provide a high level of movement safety a minimum cost of the solution and to automate a number of service functions, such as returning to the initial position after completion of a technological operation or safe movement of an object to a predetermined position.

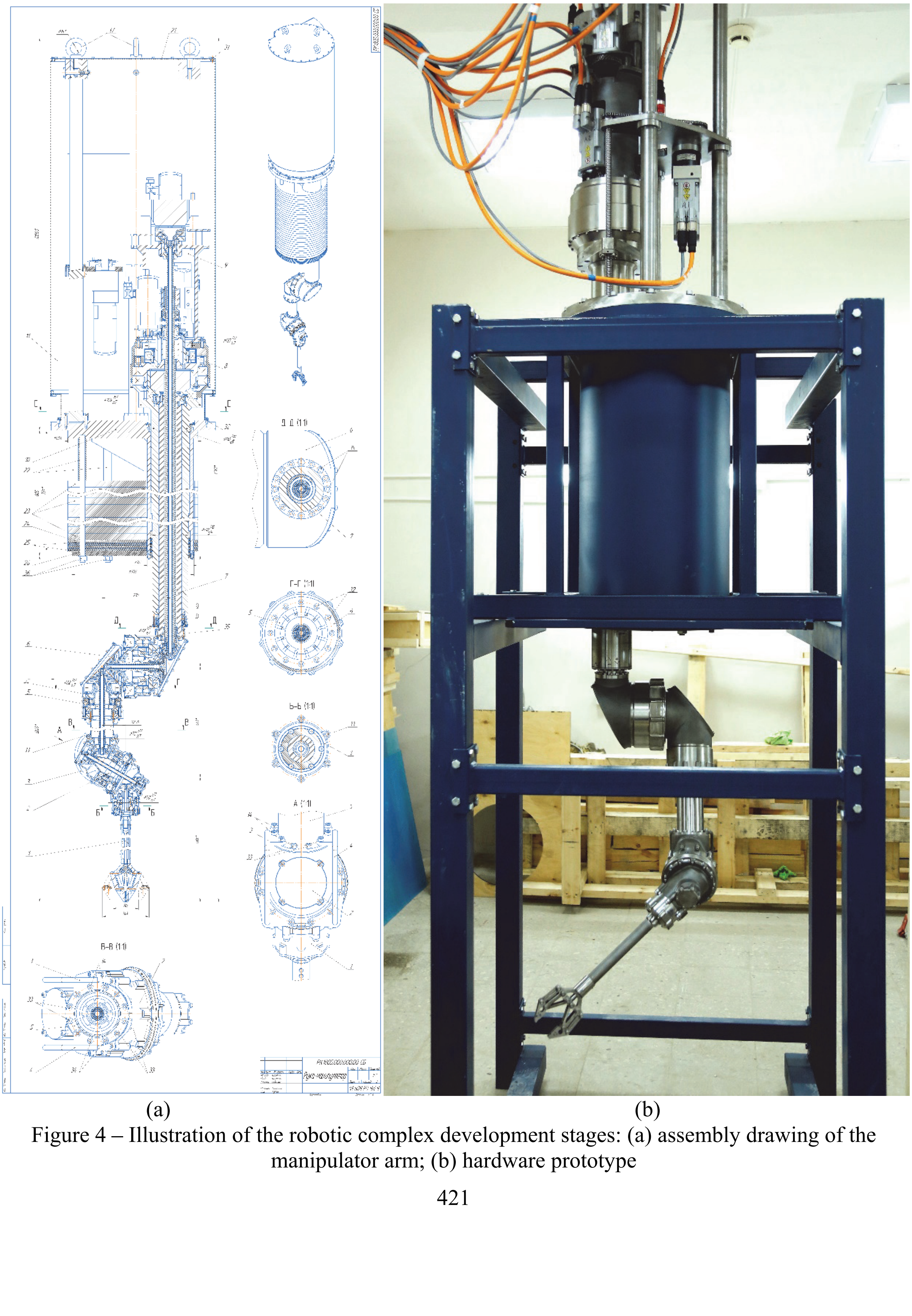
Informational interaction of the elements of the robotic complex is performed through real-time communications. The industrial EtherCAT network protocol was chosen as the real-time communication technology. Festo EMMS-AS-70-S-LS-RRB servos were used together with Festo CMMP-AS-C2-3A-M3 controllers. The distance between the controller and servomotors is up to 15 meters. The control computer can be moved away from the controllers up to 100 meters away, which would entail the need for video surveillance of the robot.

The control process suggests two main modes:

a) manual (interactive) control mode, in which the control of the robot is performed directly by the operator through the control device;

b) an automated control mode in which the control of the robot is performed by the control computer.

In accordance with the above requirements, the technical documentation has been developed and a prototype of the robotic complex has been manufactured (see Fig. 4).



(a) (b)

Figure 4 – Illustration of the robotic complex development stages: (a) assembly drawing of the manipulator arm; (b) hardware prototype

2. Training simulator

Currently, the market of training simulators is actively developing. According to the Global Market Insights report [5], the market of virtual simulators for training operators in 2018 amounted to 8.5 billion US dollars and will grow by an average of 13% per year, with the simulator market size predicted to exceed 20 billion USD by 2025. This is connected with the growing introduction of robotic technologies and the existence of a number of industries which face a shortage of professionals and feel the need to train personnel to work with equipment, technological processes and study safety to minimize errors.

In particular, in the space and defense industries, as well as in civil aviation, full flight simulators (FFS) are used, providing movement and visual means for training professional pilots in flight control, aerodynamics and control of ground support of the aircraft.

In the nuclear industry, due to the greater conservatism, the introduction of new technologies for personnel training is going on at a slower pace, and in this regard, the solution proposed in this work has important relevance.

The presence of an exact simulation model of the robotic complex, obtained from CAD software at the design stage, as well as control software developed, made it possible to quickly solve the problem of developing an appropriate interactive simulator that implements the following functionality:

– implementation of an accurate three-dimensional model of the robotic manipulator, hot cell and the corresponding equipment in accordance with the technical documentation;

– the ability to view the model of the hot cell and equipment using a VR headset and/or a monitor;

– implementation of the interaction of various objects of the virtual production site;

– the ability to perform technological operations with a virtual manipulator using a joystick with the required number of degrees of freedom, and other specialized control devices;

– implementation of feedback in the control device: the representation of collisions, the effort when interacting with the equipment, the weight of objects;

– output of additional (auxiliary) information messages on the monitor screen and in a VR headset, providing auxiliary information audio signals.

The following main stages in the development of the simulator can be distinguished:

1. Development / digitization of the technical documentation for the robotic complex and equipment. As part of the project, the corresponding technical documentation was already prepared in CAD software, and by the time the simulator development started, there had been a 3D model of the objects, which allowed to save resources and proceed to the next steps of implementation. In case of the absence of a 3D model at this stage, it is supposed to develop a geometric model based on a paper or electronic version of the technical documentation. The optimum CAD system at this stage is SolidWorks.

2. Manufacturing hardware elements of the simulator. The stage of manufacturing feedback control devices and other hardware elements in this project was part of the development of the robotic manipulator. However, in case the simulator is an independent product that requires a specialized control device, it is necessary to carry out the design work and a full cycle of development and manufacture of the product.

3. Preparation of URDF models of objects that are part of the simulated complex. URDF (Unified Robot Description Format) is a universal robot description format that is based on the XML markup language and contains the following information:

– description of the visual model;

– description of the geometric model of collisions;

– mass-inertial characteristics of the model.

URDF preparation is facilitated by the ability to import a visual model and its mass-inertia characteristics from SolidWorks CAD using the SW2URDF converter.

4. Development of control software.

At this stage, the following tasks are solved:

– collision avoidance;

– solution of the inverse kinematic problem;

– calculation of forces, interactions between the objects of the simulator;

– implementation of the necessary control and feedback modes.

The first two tasks are solved through the use of the ROS middleware. For the task of calculating forces and interactions, the Gazebo software was used (simulation of physical reality). To interact with the control device and provide feedback, an original software has been developed.

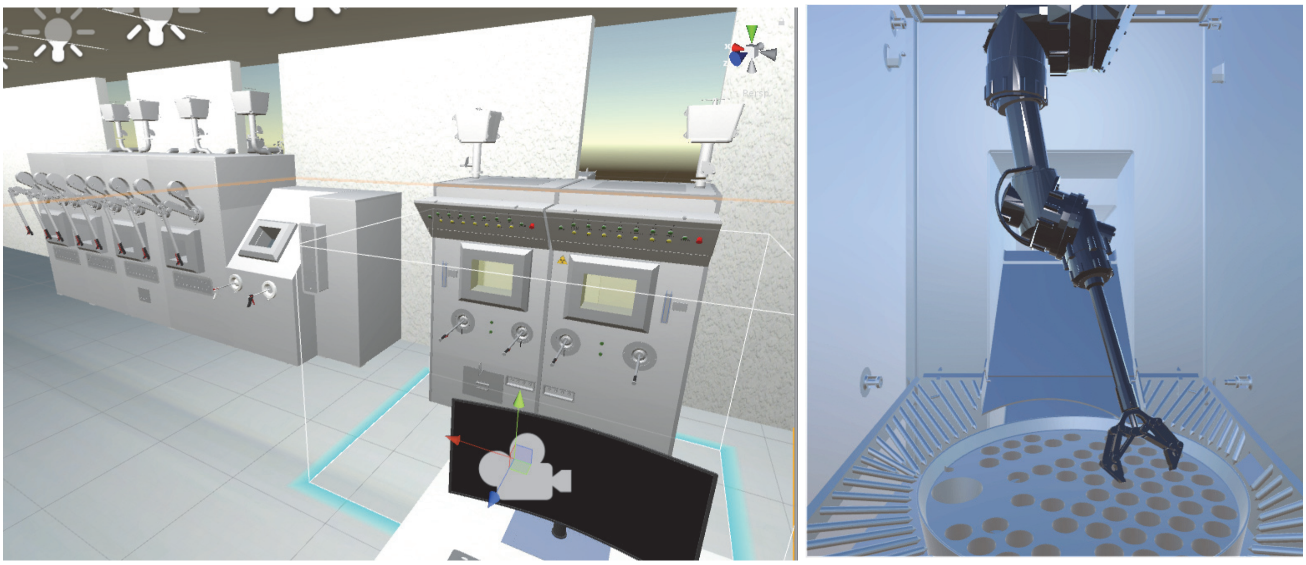
5. Creation of a realistic, optimized 3D model.

At this stage, the following tasks are solved:

1) conversion of graphic formats;
2) optimization of the geometric model:
– removal of unimportant details;
– reduction in the total number of polygons;
3) setting up the scene (see the example in Fig. 5a);
4) creation of visual effects (lighting, gleaming, refraction, etc.);
5) design of the user interface / elements of augmented reality.

The main software used to visualize the optimized model was Unity, which, by means of the rosbridge_suite and ROS# utilities, is able to interact with the ROS nodes to calculate the positions of the joints of the manipulator arm. The Blender software was also used to prepare the 3D model.

As a result of the implementation of the above steps, a training simulator of the robotic complex has been created aimed at gaining skills in performing technological operations in hot cells (see a fragment of the simulator image in Fig. 5b). The simulator is focused on the use of a VR headset in order to ensure the effect of complete immersion of the operator in the process being studied.



(a) (b)

Figure 5 – Illustrations of various modes of operation of the simulator: (a) setting up a three-dimensional scene with hot cells in the Unity software; (b) a three-dimensional image of the robotic arm in a hot cell

The optimization of the model, carried out at Stage 5, makes it possible to launch the simulator on a personal computer with a mid-level Geforce video adapter. Recommended requirements for the simulator hardware are given in table 1.

Table 1. Recommended requirements for the simulator hardware

№	Equipment
1	Computer / laptop including: Intel Core i7 CPU, 32GB RAM, Geforce GTX 1080Ti video adapter, 2TB hard drive
2	VR kit: - a HTC Vive-level headset– 1 unit., - HTC Vive controller – 2 units., - base station – 2 units.
3	Electromechanical control device with feedback

Fulfilling the recommended hardware requirements ensures a comfortable work with the simulator, eliminating such undesirable effects as abrupt movement of objects during a quick turn of the scene or a delay in response to user actions. As a minimum configuration, one can consider a laptop with an i7 8750H CPU, 16GB RAM and a Geforce GTX 1060 video adapter. In some scenes the above effects may be observed to a small extent, however, in case of using a laptop, an additional advantage appears: the mobility of the software and hardware complex.

In conclusion, the report presents the results of the development of a robotic complex, which includes original components: a robotic arm-manipulator, a feedback control device and software. The robotic arm has 6 degrees of freedom, a gripper with two fingers and is characterized by high radiation resistance, similar to that of mechanical master-slave manipulators. The prototype of the robotic complex has demonstrated the ability to perform basic technological operations in a demo hot cell. A number of service functions have been automated.

Software and hardware solutions that formed the basis of the robotic complex allowed, in a short time, to develop a training simulator of the robotic arm for practicing the operators' skills of performing technological operations in hot cells.

References

1. Manipulators and robots for the nuclear, oil and gas industries URL: https://www.hwm.com.
2. Dementyev B.A. *Nuclear power reactors* [In Russian – Yadernye energeticheskiye reaktory]. M.: Energoatomizdat, 1990, pp. 21-22.
3. Stewart D. 1965 A platform with six degrees of freedom *Proc. Institution of mechanical engineers* (London) vol 180 pp 371–86.
4. Lebret G., Liu K., Lewis F. L. 1993 Dynamic analysis and control of a Stewart platform manipulator *J. Robotic Systems* **10** 629–55.
5. Operator Training Simulator Market Size By Component (Hardware, Software {Control Simulation, Process Simulation, Immersive Simulation}, Service {Consulting, Installation & Deployment, Maintenance & Support}), By Simulation Environment (Console Operator Training, Field Operator Training), By End-Use (Aerospace & Defense, Chemical, Energy & Power, Healthcare, Industrial, Oil & Gas), Industry Analysis Report, Regional Outlook (U.S., Canada, UK, Germany, France, Italy, Spain, Russia, China, India, Japan, South Korea, Australia, Brazil, Mexico, GCC, South Africa), Growth Potential, Competitive Market Share & Forecast, 2019 – 2025. Published Date: Mar 2019, 220 Pages, Report ID: GMI1285, Authors: Ankita Bhutani, Pallavi Bhardwa. URL https://www.gminsights.com/industry-analysis/operator-training-simulator-market.

M.V. Nosikov

CONTROL SYSTEM SYNTHESIS OF RADIATION-PROOF MANIPULATOR MR-48 FOR CHAMBERS

Miass branch of South Ural State University (National Research University), Miass, Russia
nosikovmv@susu.ru

Abstract

This work covers the architecture, functional specifications and structure of the control system and operating modes of a 6-degree-of-freedom radiation-proof manipulators designed for the replacement of an outdated human-driven copying-type manipulator. The proposed robotic systems provide manual, semi-automatic and automatic modes, including the definition of workspace, collision avoidance, and automatic grasping of tare and materials. The key feature of such manipulator systems is presence of human in control loop, by reason of variety and complexity of manipulations and, in some cases, non-deterministic environment. Due to these factors, it is very desirable to form operator sustainable skills to control such manipulators, including learning process using training systems. A relatively large distance between control stations and manipulators, located in sealed chambers, imperfection of visibility from control zone form necessity of vision systems usage. Chamber videocameras are used for general area observation, while cameras installed on manipulator end-effector allow local work zone observation. High level of radiation in chambers requires the cameras to comply with limitations regarding their functionality in radiation fields, lifecycle and so on. Special technical solutions such as additional radiation protection and algorithms of manipulator control allow increasing the lifetime of the vision system.

Keywords: robot-manipulator, control system, ROS, training system, vision system.

Intriduction

Modern multi-axis robotic systems (RS) and manipulators with electric actuators are widely used in many industries, and in most cases are highly autonomous systems operating in automatic mode. Their typical modes of operation are: movements along a given trajectory with the required speed without external environment non-determinism; adaptive execution of movements along the required trajectory with the analysis of external environmental conditions (for example, when using a technical vision system).

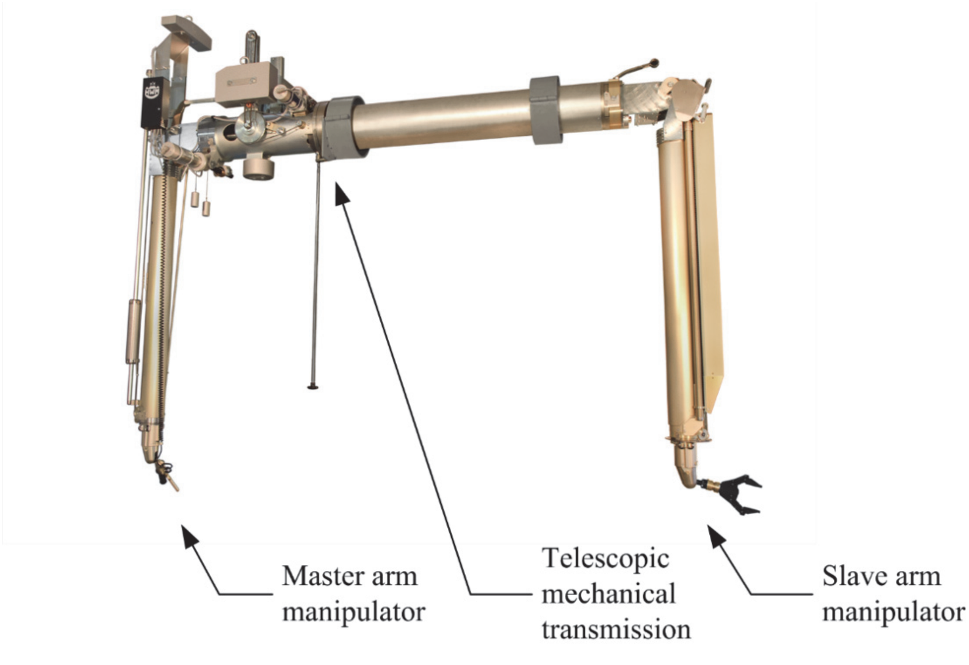


Figure 1 – Example of «Master arm – Slave arm» manipulator system

A separate class of robotic systems, mobile and stationary manipulating systems is systems where a human is included directly in the control loop. RSs of this type are used in areas with non-stationary and / or non-deterministic environmental conditions. One of the most important applications of this RS type is nuclear industry where a number of hazardous factors (radiation fields, chemically aggressive environment, etc.) and the high variability of operations performed are combined with the need to protect people from hazardous factors. Developed in the 1960-1970s manipulators were built copying the position of the master organ with an actuator (manipulator) with mechanical gears or an electromechanical synchro system [1]. At present, a number of Russian and foreign enterprises still produce electromechanical copying manipulators, and the number of RSs of this type at the enterprises of the Russian nuclear industry reaches several hundreds. Figure 1 shows a typical lightweight electromechanical manipulator with a lifting capacity of up to 15 kg. Slave arms are usually installed on the ceiling of a thick-walled pressurized chamber, which guarantees the protection of the operator and maintaining staff from the influence of dangerous factors. The operator performs the required technological operations by moving the master arm, kinematically similar to the slave arm. The control is provided by visual observation through the thick-walled glass of the chamber front wall. Figure 2 shows a section (side view) of a typical sealed chamber with the installed manipulator slave arm and the master arm.

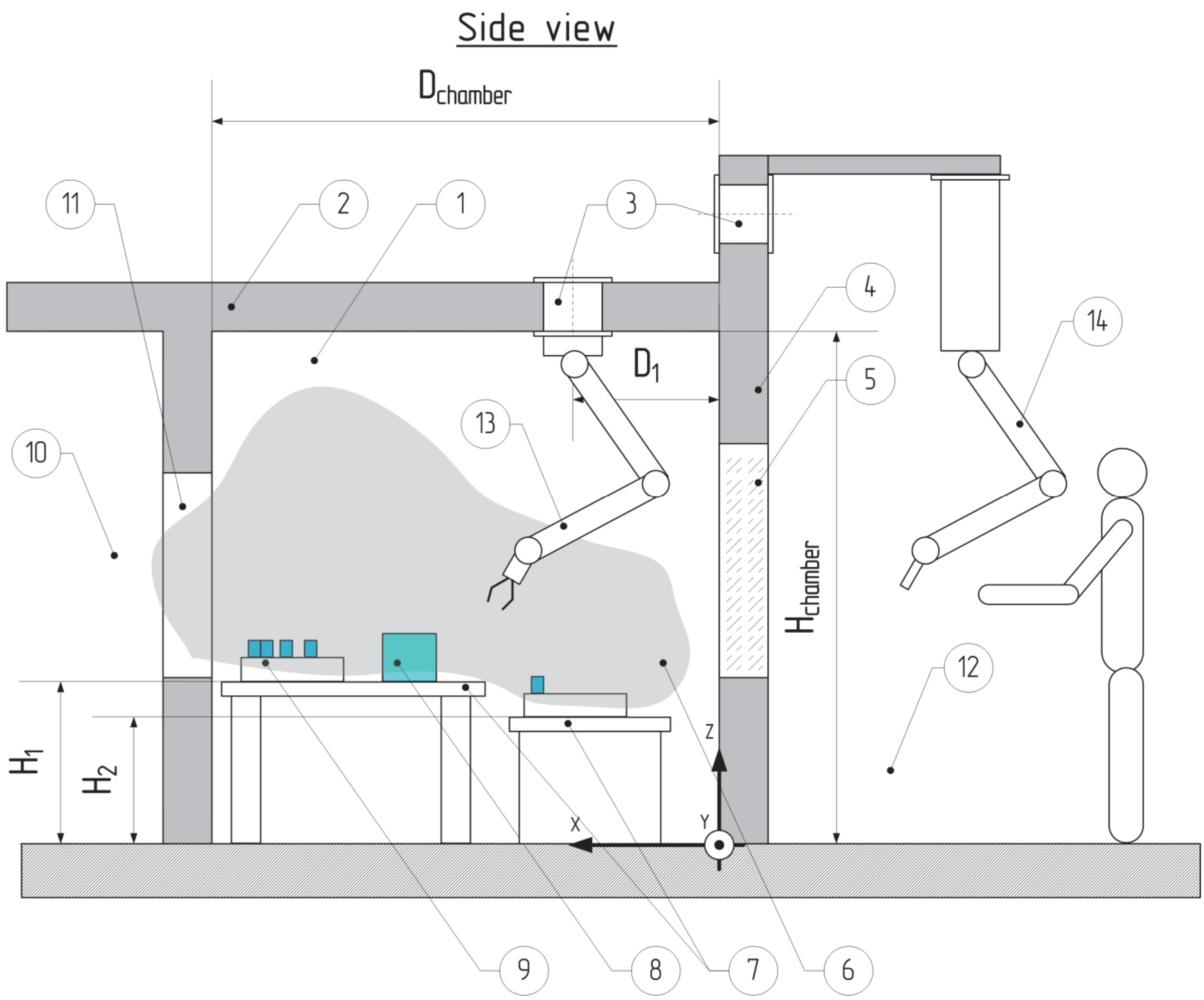


Figure 2 – Sealed chamber (side view)

Positions on Figure 2:
1 – intrachamber volume;
2 – chamber ceiling;
3 – manipulator and equipment openings (canals);
4 – chamber front wall;
5 – protective glass;
6 – manipulator and equipment workspace at typical operations execution;
7 – main and auxiliary surfaces for parts and materials positioning;
8 – intrachamber equipment;
9 – tare in a container;
10 – transportation corridor;
11 – chamber to corridor zone;

12 – operator zone;

13 – intrachamber manipulator (slave arm);

14 – kinematically identical master arm.

Despite the obvious advantages of kinematically identical systems in terms of the identity of kinematic configurations of the master and slave arms, this type of manipulators has several disadvantages:

1. High mechanical loads on the operator's hands, high fatigue during continuous technological operations.
2. Backlashes in the kinematic pairs of both the manipulator and the master arm, resulting in a positioning accuracy loss, and, in some cases, in making the required operations impossible.
3. Fixed gear ratios (1:1) between the master arm and the manipulator, which causes certain difficulties for operations requiring the high accuracy of positioning the gripper.
4. The impossibility of automation operations performed by manipulators of this type.

In general, the tasks of RS control systems with a human operator in the control loop include:

1. Ensuring technological manipulations with the required accuracy and speed. Providing the possibility of manual, automated and automatic modes.
2. Providing the operator with means of influence/control (master arms) to choose and control the magnitude and direction of controlling impact.
3. Forming and visualizing with technical means the system state vector necessary and sufficient for its unambiguous perception.
4. Providing control and security while manipulating.
5. Ensuring interface with higher-level systems to integrate separate technological zones into a single information space of modern productions (technologies like , “Industry 4.0”, etc.).

Hardware and software of the control system with radiation-proof robotic system

Commissioned by one of the leading companies in the nuclear industry, the robotics laboratory in Miass branch of South Ural State University designed and manufactured a prototype of the intra-chamber robotic arm MR-48. In accordance with the technical specification, the product includes a 6-degree manipulator arm equipped with a gripper with smoothly adjustable gripping force, and an operator control station with an integrated control system. The manipulator has a 950 mm spherical workspace and provides a load capacity of 15 kg. The specification also defines the design of the wrist mechanical tool interface for a number of standard tools used in the industry, and formulates the requirements for the geometric parameters of movable segments, maximum rotation angles and the required angular rates in rotational joints.

This RS control system is based on widely used digital computing devices and software algorithms, electromechanical sensors indicating the segments position and velocity, and devices for interface with an object. A significant proportion of the control system functioning is assigned to a program code. The functional diagram of the control system is shown in Figure 3

The core of the control system is the computer of the operator’s console interfaced with controls, sensors and slave arms via input-output modules being also hardware and software components of the control system. Input and intermediate signals of the manipulator control system are:

1. Operator actions, applied to controls (sticks, buttons, etc.) $\boldsymbol{M_i}$;
2. Manipulator joint angles $\boldsymbol{q_i}$;
3. Manipulator joint angular rates $\boldsymbol{\omega_i}$;
4. Engine currents $\boldsymbol{I_i}$.
5. Required angles and rates $\boldsymbol{q_{i\,SP}}$, $\boldsymbol{\omega_{i\,SP}}$, calculated in control loops.

The output signals of the control system are PWM voltage pulses Ui (i = 1 ... 6) generated at each moment of time so that the manipulator performs the required movement pattern. The functional diagram of the electromechanical drive of the manipulator’s i-th segment is shown in fig. 4.

In general, the control system is discrete, the control actions generation of which is 20 ms. Due to the specific features of the control object (manipulator) and its operation modes, the control system is made as several nested control loops implemented at the hardware, firmware/software levels.

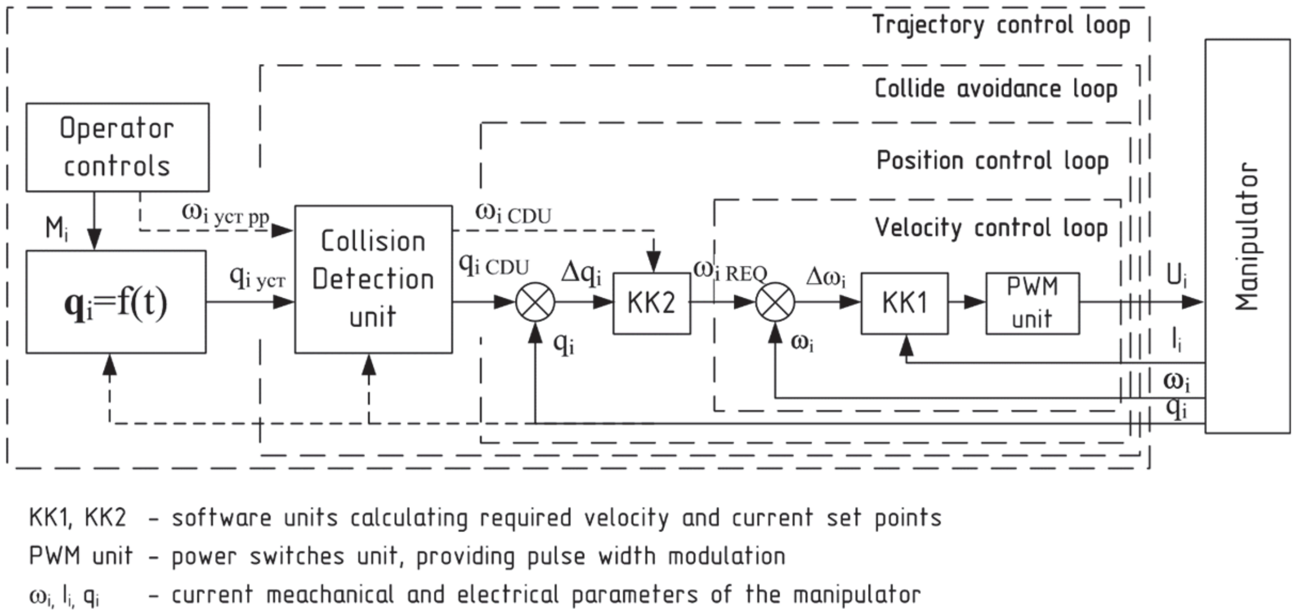


Figure 3 – Functional diagram of manipulator control system

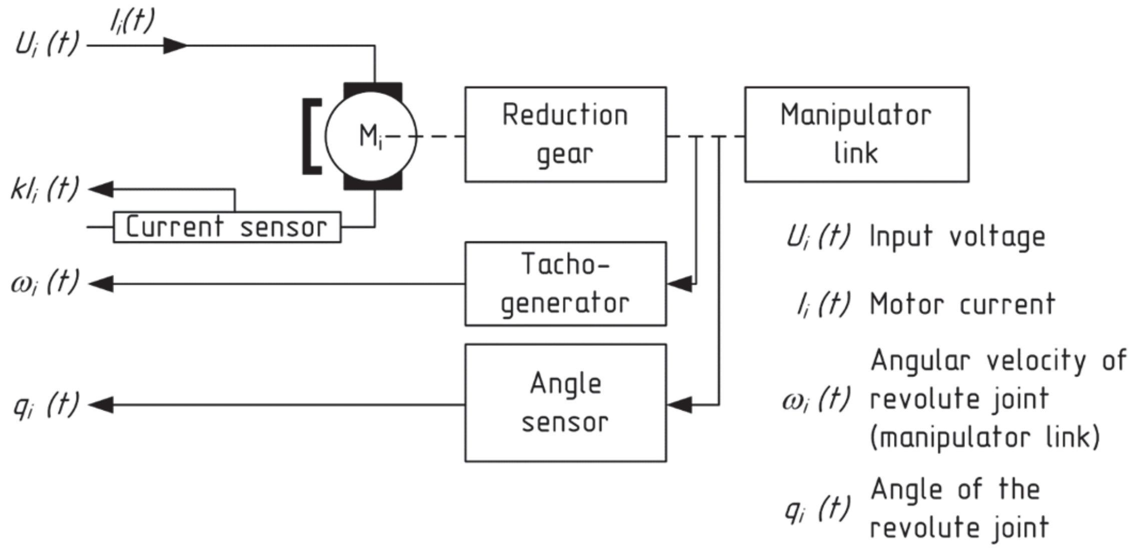


Figure 4 – Electromechanical structure of the «MR-48» revolute joint

As a human-machine interface (HMI) in terms of operator controls, it is possible to use a number of modern devices: mouse-type control devices; dual- and multi-axis joysticks; operator controls with three linear and three angular degrees of freedom (DoF), including wireless ones based on inertial sensors; exoskeletons, etc. Multi-axis joysticks are considered the most suitable for the production area in terms of reliability and ergonomic controls. At the MP-48 control station there are two multi-axis joysticks with proportional and discrete control outputs, whose static characteristics (angular deviation - output code) can be adapted to the specific control mode of the manipulator in order to provide more precise control (Figures 5-7). The most frequently the following methods are introduced into static characteristics: dead zones, saturation zones; change in gain ratio; relay characteristics formation; arbitrary characteristic formation (for example, by a piecewise linear method).

Figure 5 – Manipulator «MR-48» and its operator control station

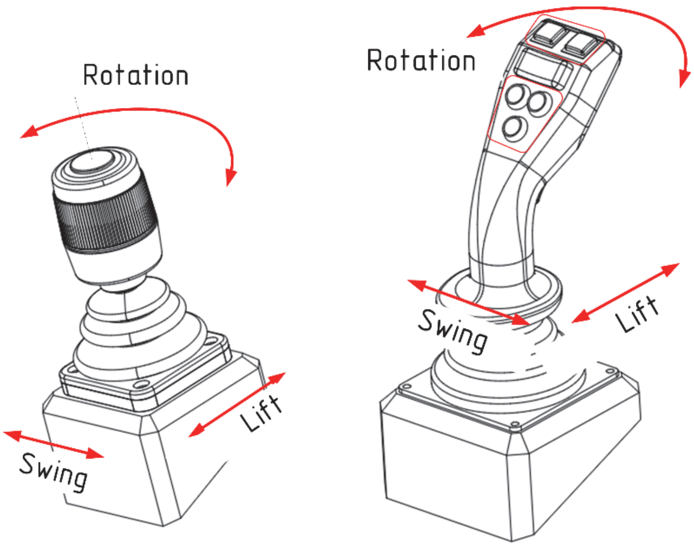
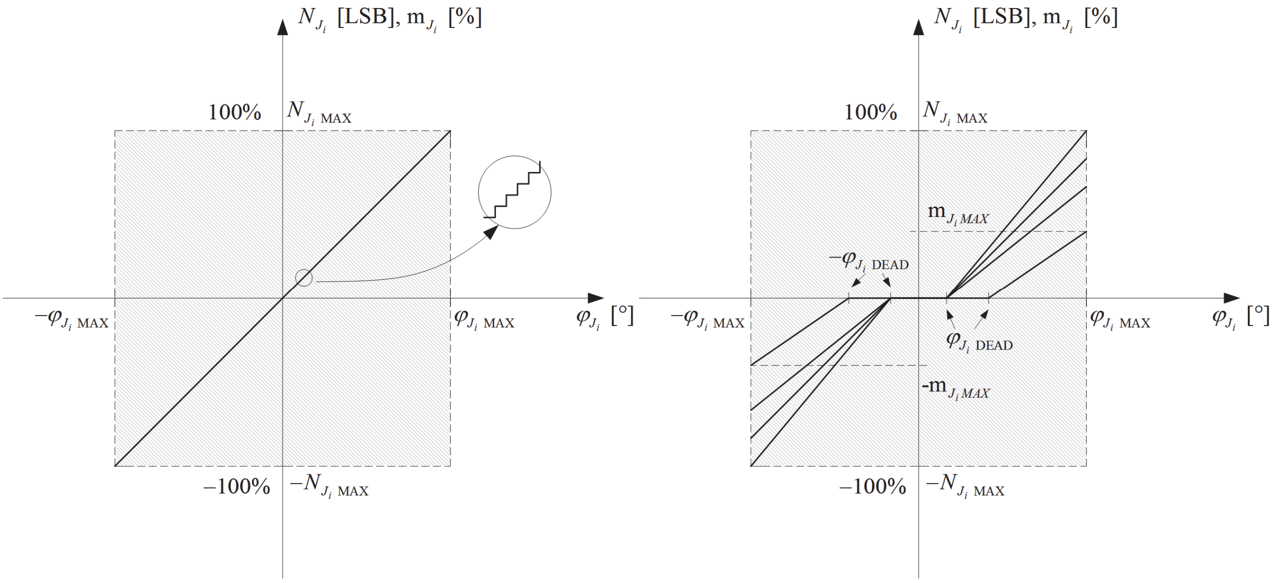


Figure 6 – Manipulator «MR-48» control joysticks



a) b)

Figure 7 – Nominal (a) and adapted (b) joystick static characteristics

Items on Figure 7:

φ_{J_i} – mechanical angle (inclination) of ith joystick rotation axis;

$\varphi_{J_i MAX}$ – maximum mechanical angle (inclination) of ith joystick rotation axis;

$\varphi_{J_i DEAD}$ – "dead zone" of ith joystick rotation axis;

N_{J_i}, m_{J_i} – absolute and relative code output of ith joystick rotation axis.

When implementing software modules of the control system, it is necessary to pay attention to the hardware structure of the control system, the functionality of its computing devices, requirements for speed and accuracy. Today there are a number of approaches and software platforms used for implementing control systems, one of the most widely used is specialized software library sets ("frameworks", algorithm sets, functions and program classes). One of the most popular solutions for the problems of mobile and stationary robotics is the Robotic Operating System (ROS, ROS2, ROS-Industrial) framework [2, 3]. ROS is an add-on for the operating system (OS; the base is the Linux class OS). ROS includes the OS core which provides arbitration, synchronization, interprocess communication, a large number of specialized software libraries and modules that implement algorithms for receiving and processing data of primary meters, solving manipulator kinematics and dynamics problems, localization and path and trajectory planning. ROS software architecture is supported by a large number of research projects, leading manufacturers of industrial robotic equipment, such as FANUC, ABB, Universal Robots. The developers of RS control systems are provided with a wide range of development and debugging tools for the user software modules integrated into the common ROS software infrastructure ("Module 1", "Module N", Figure 8).

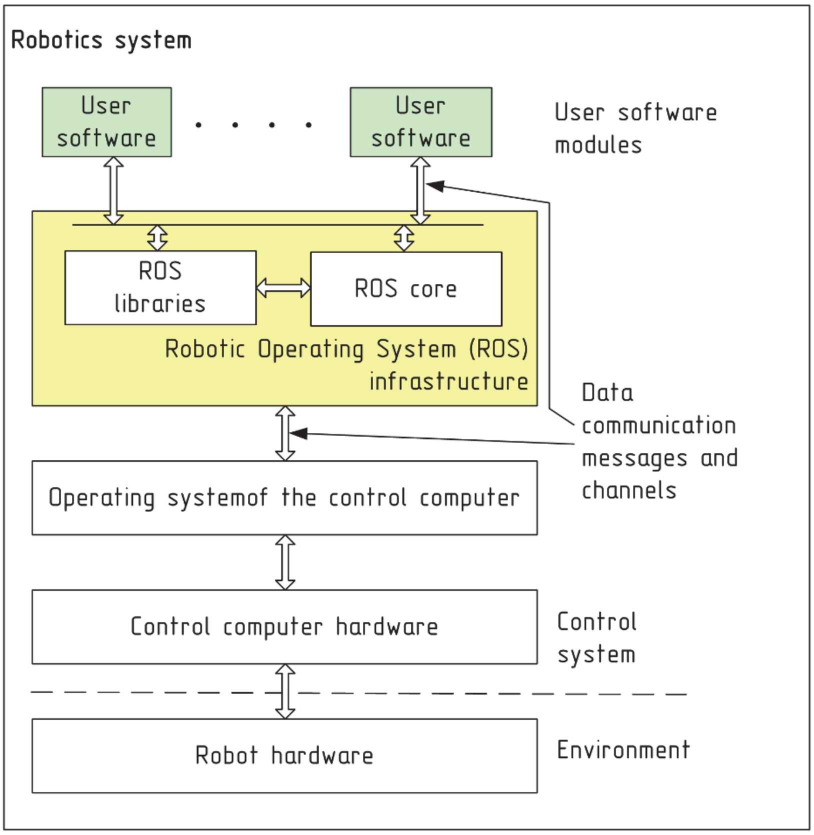


Figure 8 – Robotic operating system (ROS) and its place in the manipulator control system software/hardware

One of the ROS key features is the subsystem of messaging between the software components of the control system, which ensures the guaranteed delivery of information from the sending to the receiving module. The subsystem allows exchanging typical data structures (whole and fractional numbers of various formats, string messages, messages from standard I / O devices), specialized messages (homogeneous matrices, manipulator configuration state vector), and complicated user message structures. Algorithms for launching and initializing custom software modules (called nodes in ROS) allow the ROS kernel to send a message list that the node requires from other components and nodes of the system, as well as a message list generated by this node when performing periodic or aperiodic actions. Thus, a table of information interaction of software nodes is formed, an example of which is graphically prepared by the standard ROS utility *rqt_graph*, shown in Figure 9. When a message of a certain type is received, the program node calls the message processing function, performs the necessary math and/or logic operations, and generates an answer message.

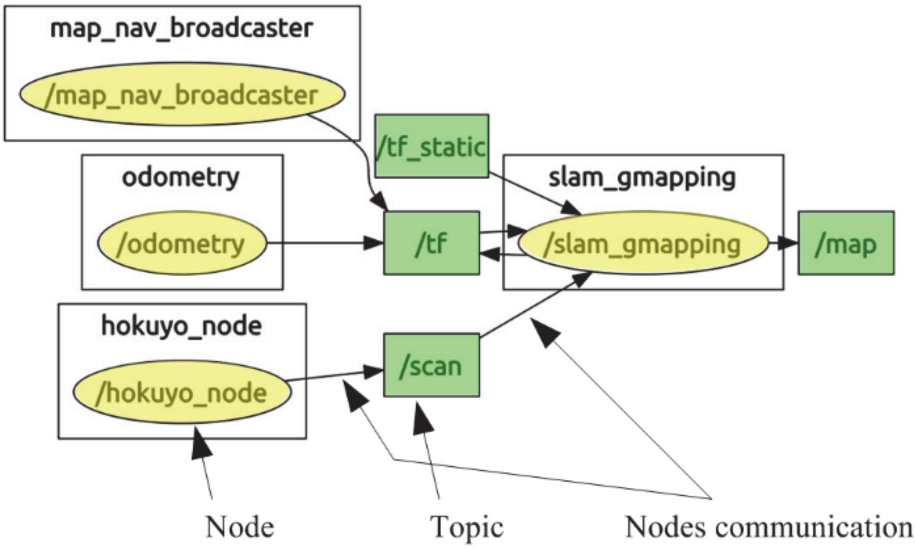


Figure 9 – Typical ROS software structure, built using «rqt_graph» utility

ROS-based MR-48 control system includes the following ROS nodes (Table 1):

Table 1. Main ROS nodes of MR-48 control system

ROS node	Functional description
/mr_algorithms	Main program unit of control system
/mr_hrdw_interface	Middle and low-level protocol communication unit. Receiving information from manipulator sensors and sending control set points
/trainee_ctrls /instructor_ctrls	Program units, performing data acquisition from proportional and discrete data channels of trainee and instructor joysticks
/training_task	Training session task generation module
/storage	Motion and action parameters file and database storage module
/usb_cam	USB videocameras image capture module
/speech_database /sound_play	Sound and speech generation modules

In the process of trial operation of the MP-48 manipulator, registration and analysis of the operator's actions on the operator controls were carried out, as well as a general analysis of the operators readiness to use remotely operated RSs with joystick-type operator controls. In order to improve the operator's skills both under normal operating conditions and in emergency, a simulator training, monitoring and analysis of operator's actions was proposed as part of the MP-48 software system.

Training system architecture

Taking into account the equipment and materials cost, limited chamber workspace, measurement and technological equipment inside chamber, precision nature of the operations and individual perception and tactile characteristics of the personnel, it can be say, that operator training system and different training procedures are mandatory to initiate and improve operator skills.

Based on principle “from simple to complex”, the training procedures (course) can be built on the following steps:

Theoretical part:

– manipulator characteristics, workspace, construction, tools;
– control station structure;
– manipulator start-up and power-down;
– general manipulator operations;
– gripper operations.

Practical part:

– screen operations (screen tabs and views, button pressing, view rotation, etc.);
– basic joint movements;
– operations in Cartesian end-effector frame;
– operations in Cartesian base frame;
– object taking and releasing;
– trajectory record, save and replay;
– parking operations;
– emergency procedures.

In theoretical part it is possible to make intermediate tests or exam, and in practical part the decision of trainee’s skill is based on the instructor expert opinion, analysis of recorded control system data and so on. The approaches to practical realization of the training system are proposed below.

When designing a training system (practical part) it is necessary to ensure the implementation of the following tasks:

1. Forming a set of training and control practical tasks (motion trajectories, initial and final manipulator configurations, etc.) for various scenarios of technological operations.

2. Guaranteed recording on the data carrier of all control vectors, state vectors, modes and tuning parameters of the system. The solution of this task allows you to:

– analyze the operator’s impact on the controls throughout the entire training stage or implementing a technological operation;

– analyze the adequacy of the operator's actions, based on the system state (the manipulator kinematic configuration);

– reproduce repeatedly the operator's actions and manipulator’s motion parameters including the parameters with a different time scale. The repeated reproduction and analysis of the system state can be performed both on the manipulator and its software mathematical model.

3. Involving the instructor into the process of training in order to promptly correct the actions of the trained operator. This approach assumes the instructor's workplace with controls identical to the operator’s control system, and the instructor’s controls must have a higher priority than the operator’s ones.

Before the training stage, the instructor makes the current assignment or selects it from the list of typical operations, taking into account the qualifications of the trained operator. Setting up the task is performed at the instructor's workstation (Figure 10) and may include: the initial kinematic manipulator configuration, the number of manipulation objects, the obstacles location configuration, etc.

To register (save) inter-module ROS communication, the *rosbag* service is used - a set of functional modules and procedures to record the required messages into the file system using the specified criteria. While recording, the message sequence number is automatically generated, and it is strictly bound to the system time. The recorded messages can be visualized in graphical and tabular form with internal ROS tools or with external software. In the training system, the recorded information can be used to analyze actions in the “delayed time” mode, and be replayed to evaluate actions in the “real time” mode. Table 2 shows the types of messages stored in the MR-48 control and training systems, the frequency of their receipt and the estimated amount of stored information.

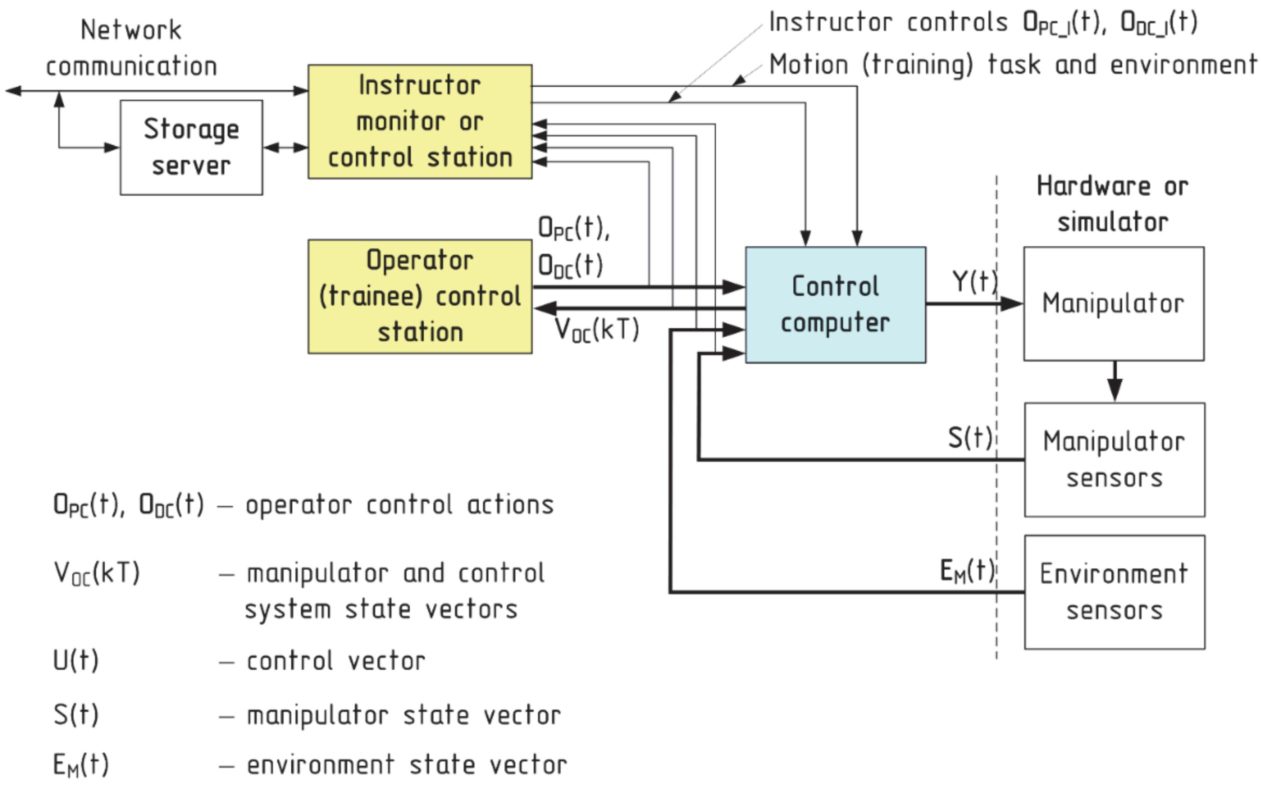


Figure 10 – Training system structure

Table 2. Types and sizes of manipulator and joysticks registration data

Message type	**Description**	**Sample rate, Hz**	**Data stream size, MB/min**
/tr_ctrls /instr_ctrls	Trainee and instructor control actions	1-50	0.1-1
/joint states	Current kinematic configuration and motion state of the manipulator (joints angles and angular rates)	50-200	0.2-0.8
/tf	Homogeneous matrices of manipulator kinematic configuration	50-200	0.3-1.2
/training_events	Text and numeric messages of current training session	aperiodic	0.1
/usb_cam/image_raw	Data stream from video camera	10-50	10-100

The “Database engine” software module provides storing the information about the training process including the identification information about the operator, instructor, training tasks, training sessions, and operator’s performance of a particular task. The “Database engine” module integrates this information with the information registered in the *rosbag* file in a single relational SQL database (MariaDB 5.0). A simplified structure of database tables and their relationships is shown in Figure 11.

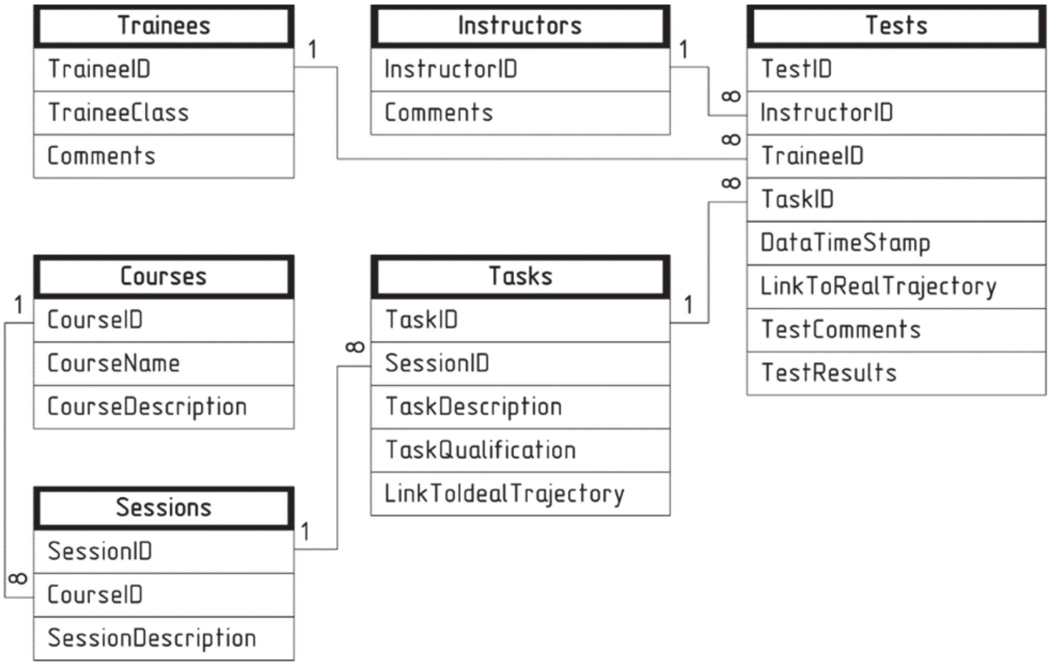


Table	Table description
Trainees	Trainee personal infromation table
Instructors	Instructor personal infromation table
Courses	Trainee courcses (programs) table
Sessions	Trainee sessions data table
Tasks	Performing / performed tasks data table
Tests	Results data table

Figure 11 – Training system databases tables structure

The training system of MR-48 has been built using following information technologies/components (Figure 12):

– ros-bridge system, sending particular list of ROS messages over TCP connection using WebSocket and JSON technology;

– Web-browser with WebSocket technology support;

– HTML5 language;

– JavaScript (ECMAScript 2018);

– JSON5 communication protocol.

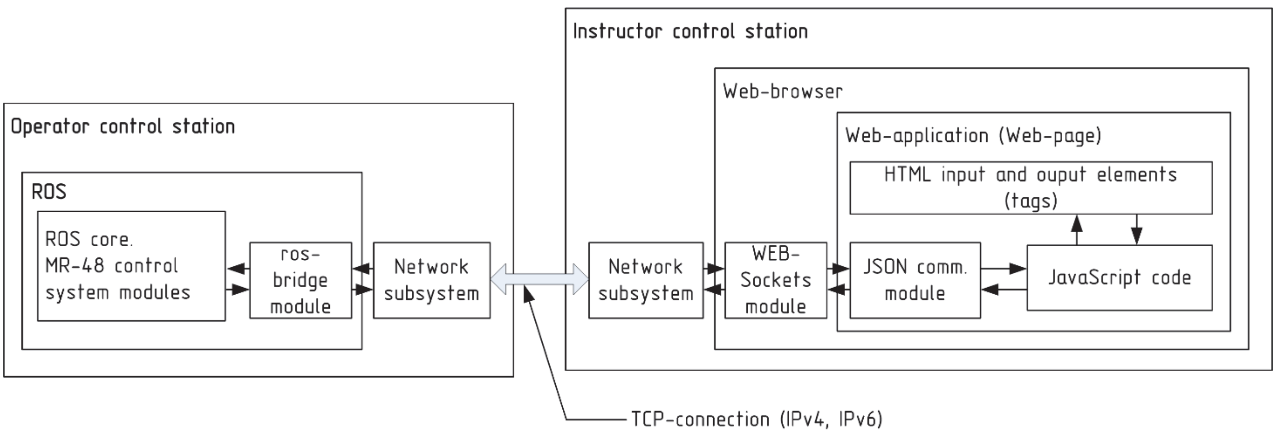


Figure 12 – Web-browser and control system communication structure

Figure 13 demonstrates the instructor’s web-console screen functioning in the operator’s action real-time monitoring mode.

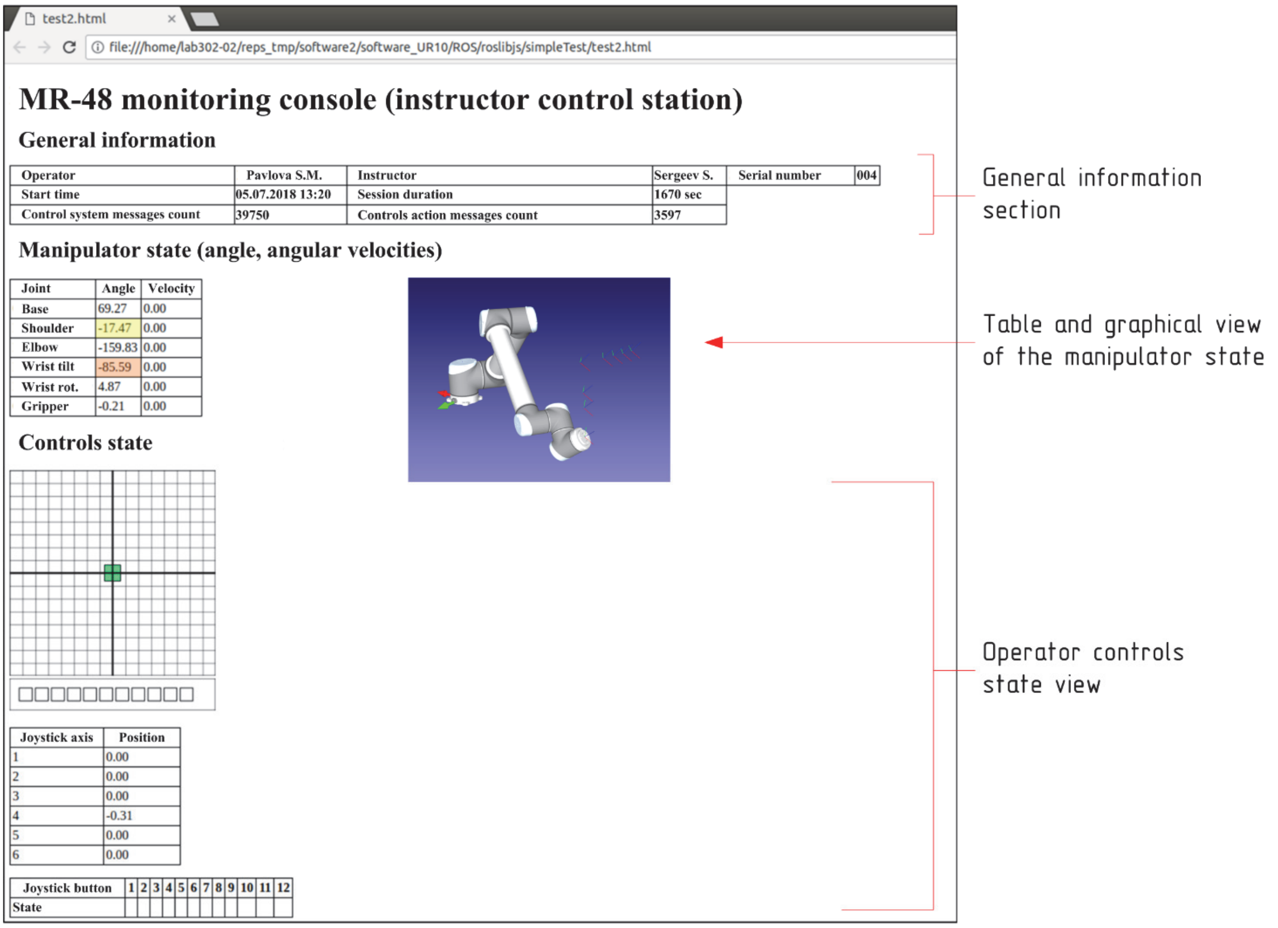


Figure 13 – Instructor’s control station Web-view of the operator’s actions and manipulator state

Computer vision system architecture

Computer vision systems (CVS) are currently one of the main means of developing motion object control systems in conditions when the amount of prior information is insufficient to make a decision on further actions and real-time environment perception and analysis is needed. The most common tasks of vision systems are:

1. Organizing the channel for transmitting video information to the operator(s) of the processing equipment. This CVS structure can be quite complicated and include several video signal sources located in different points of the studied space, cameras optical axis orientation drives along one or more angular or linear coordinates, focusing devices, multiplexing and data transmission channels.

2. Organizing an additional data transmission channel based on image processing algorithms in order to identify key information required for operator and dispatch personnel’s control systems. This CVS task can also include search and recognition of individual objects, synthesis of messages about the nature of the objects’ relative position, parameters of their movement. As a rule, these tasks require real time solutions.

3. Registering data flows listed in paragraphs 1 and 2 to data storage devices in order to further information transfer to the consumer, its reproduction and analysis in the “delayed time” mode.

The operating conditions (high levels of ionizing radiation, chemically aggressive environment) of stationary radiation-resistant manipulators described in [4-6], operating in sealed chambers, impose a number of restrictions on the CVS hardware composition, its spatial configuration and protection from external impact. One of the features of operating this equipment in sealed chambers is the chamber environment influence on the elements and systems that are not practically removed from the chamber working zone so they cannot be protected by distance and time. Thus, the only possible way is protection by shielding (shown in Figure 14 as “Protection cover”). It should be noted that the electronic components and units of the control system (CS) and computer vision systems must be placed in a radiation-safe zone.

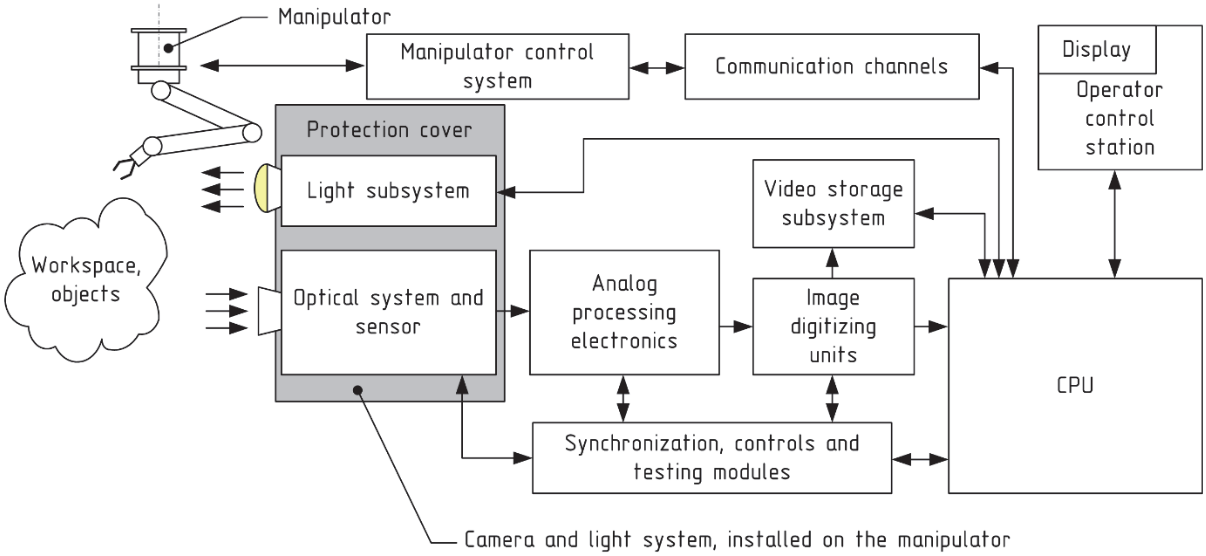


Figure 14 – Robotics vision system structure

Depending on the tasks solved by the RS, the following CVS options are possible: a system without surveillance cameras with one or two cameras in the manipulator's gripper zone, a system with additional one or more surveillance cameras mounted on a fixed or rotary base. The optical axes of the short range working chambers in the manipulator's gripper zone can either be coaxial with the gripper’s geometric axis or separated (Figure 15).

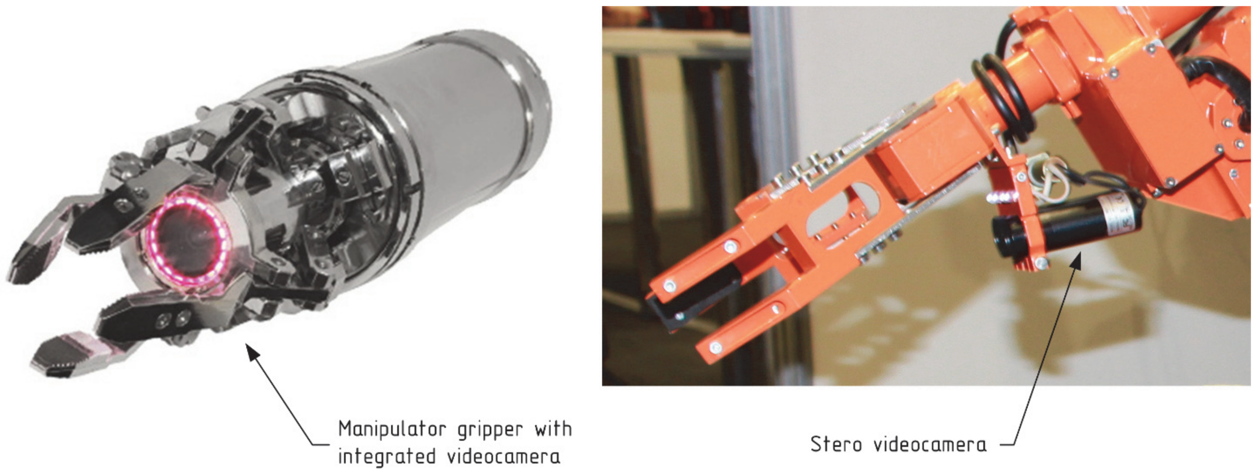


Figure 15 – Axial and side positioning of the gripper video cameras

There are the following stages of operation with a digitized image (video stream frame) received from the CVS transformation modules to its computing device: image segmentation, image analysis (calculation of qualitative and quantitative attributes) and description by the internal structures and fields of these CVS algorithms. The final stage of CVS operation is transferring data packets of the main and auxiliary information to the RS control system which, depending on the mode, uses this information to control the manipulator.

In order to improve technical and economic indicators, modern production (technological) processes are faced with the task of optimal information support. The integration of technological and production data into a single enterprise information environment enables to assess the current state of the production cycle, correlation with the plan, creating intermediate and final reports, implementing analytical and forecast functions. The key factor in solving these tasks is the identification of objects and means of production, accounting for the operations execution time. Used since the early 1970s identification systems on the basis of bar codes allow to partially or completely solve the accounting problem. One of the modern widely used barcode options is a two-dimensional QR code (Figure 16). Barcodes (correction codes) ensures guaranteed decoding primary information with the loss of up to 30% of visual information from the data zone [7].

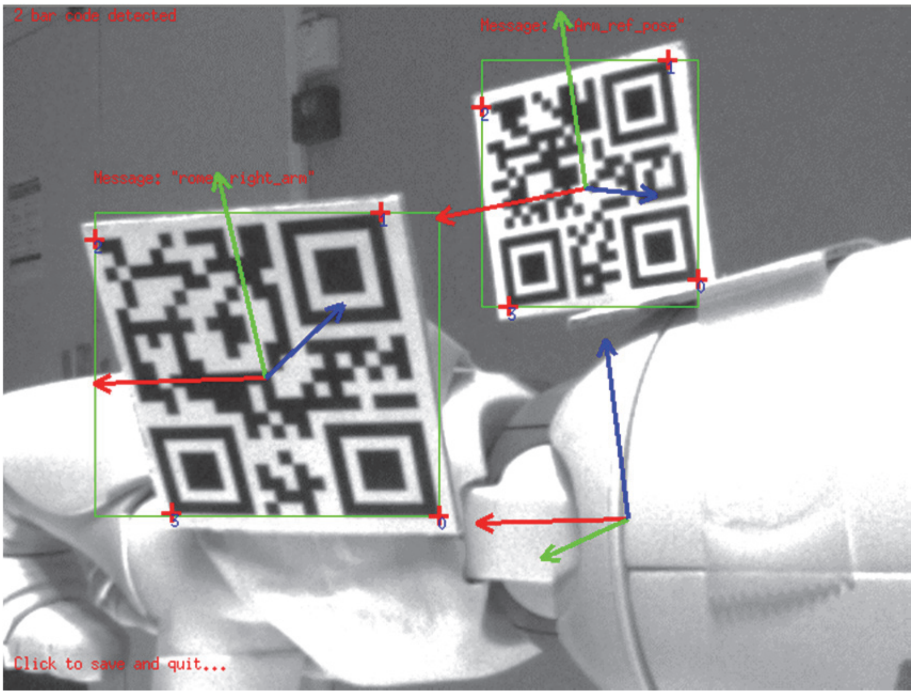


Figure 16 – Vision system camera view with identified QR-coded objects

The software modules of the MP-48 manipulator control system, built on the basis of the modified (using the real-time kernel) Ubuntu operating system and the ROS infrastructure software (framework) [2, 3], includes the QR_TRACKER software module (C ++) providing the interaction interface with the ROS core, the Video4Linux API v2 subsystem, the ZBar image recognition library and implementing the selection and "tracking" of several QR codes.

The CVS and QR codes identification system allow point the manipulator's gripper at an identified object in an automatic (semi-automatic) mode, i.e. to form the manipulator kinematic configuration for a certain time interval so that the coordinate system associated with the video camera is positioned in a certain way and oriented relative to the coordinate system associated with the object (Fig. 17 before and after pointing at the object). Taking into account prior information about the type of an object or information obtained during the QR code analysis, the task of the guidance subsystem is to bring the manipulator's gripper to a given point in the space S, whose coordinates are formed in the $O_{IMi}X_{IMi}Y_{IMi}Z_{IMi}$ object coordinate system. In the particular case, point S may have zero coordinates X_{IMi}, Y_{IMi} and a positive coordinate Z_{IMi}.

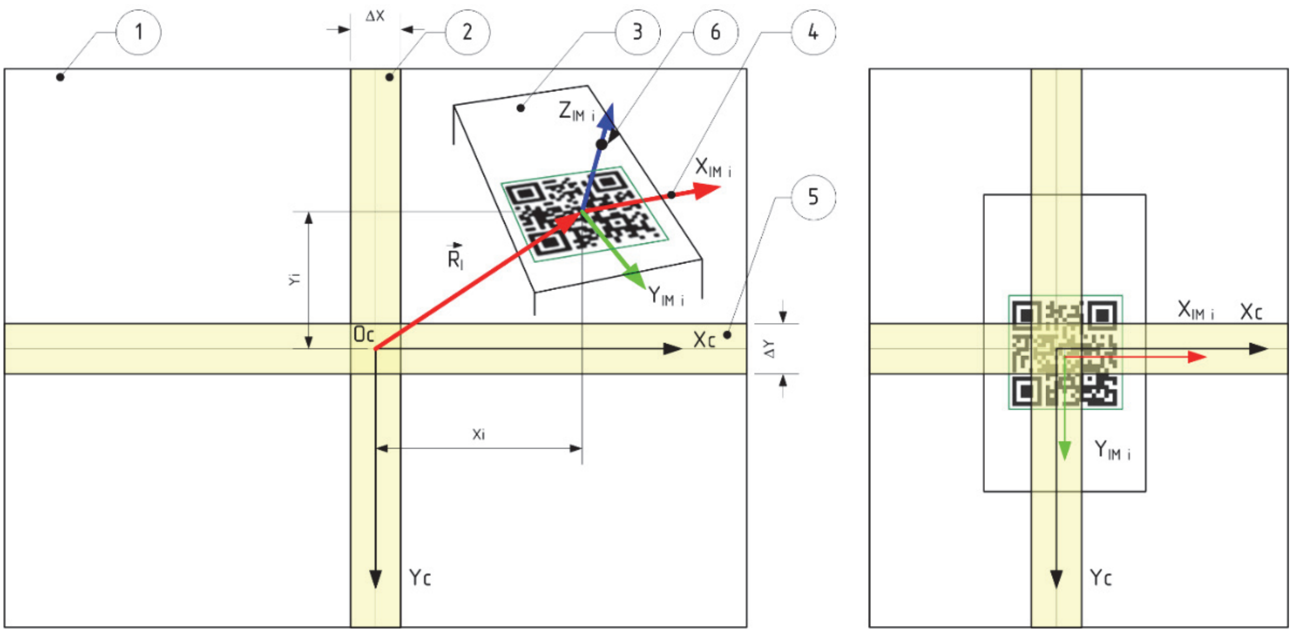


Figure 17 – Identified QR-coded object before (left) and after (right) manipulator/gripper positioning

Positions of Fig. 17:
1 – camera field of view (FoV);
2 – gripper (camera) positioning tolerance relative to the center line of camera FoV (X-axis);
3 – identified object;
4 – object coordinate frame;
5 – gripper (camera) positioning tolerance relative to the center line of camera FoV (Y-axis);
6 – gripper target positioning/orientating point.

The algorithm for the gripper positioning and orienting can be implemented using the following two approaches:

1. The gripper is guided by calculating the control actions on the drives at each clock cycle (frame) of the CVS received information.
2. The gripper is guided by a single calculation of the gripper's required position and orientation, with its further automatic guidance.

The first approach is preferable for targeting the objects having a high probability of movement during the guidance phase, and it requires increased computational resources. Algorithms of the second approach are less resource intensive, since the guidance point is calculated once by the operator's command and the CVS image is not analyzed during the guidance stage.

Along with the manual gripper's velocity control method (with proportional joysticks deviation) proposed in [5] and implemented in the manipulator prototype, it is also proposed to implement the algorithms of automatic positioning in the form of velocity vector control:

$$\dot{\mathbf{x}} = f(\vec{\mathbf{R}}_i, \Lambda_i) = f(R_{xi}, R_{yi}, R_{zi}, \lambda_{0i}, \lambda_{1i}, \lambda_{2i}, \lambda_{3i})$$
$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x} & \dot{y} & \dot{z} & \dot{\varphi} & \dot{\psi} & \dot{\theta} \end{bmatrix}^T$$
$$\dot{\mathbf{q}} = \mathbf{J}^{-1}\dot{\mathbf{x}},$$

where

$\dot{\mathbf{x}}$ – end-effector state derivative (linear velocities and angular rates);

$\vec{\mathbf{R}}_i$ – position vector of ith object coordinate frame origin;

Λ_i – quaternion, describing ith object coordinate frame orientation relative to camera coordinate frame;

$\dot{\mathbf{q}}$ – angular rates of manipulator joints, calculated using inverse Jacobian $\mathbf{J}^{-1}$

Figure 18 shows variants of phase trajectories of a linear gripper velocity $\dot{x}$ using an example of component R_{Xi}.

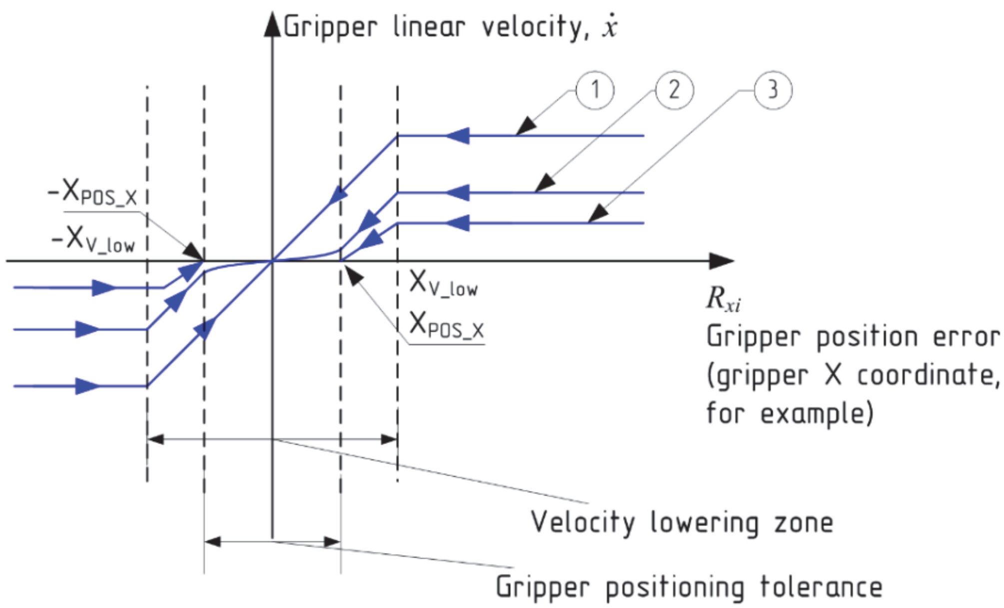


Figure 18 – Possible scenarios of gripper positioning (gripper cartesian X coordinate)

Positions of Fig. 18:

1, 2, 3 – variants of trajectories on phase plane (gripper velocity vs positioning error);

X_{V_low} – switching boundary (gripper velocity slowdown);

X_{POS_X} – positioning tolerance.

The paper suggests following set and structures of information messages between computer vision system, control system and operator regarding object optical identification:

1. Computer vision system status.
2. QR-code recognition mode status.
3. Number of QR-coded objects within camera view.
4. Object(s) identification number.
5. Object(s) position (linear coordinates).
6. Object(s) orientation (Euler angles, quaternions).
7. Selected object for following tracking.
8. Tracking mode status.
9. Kinematic gripper (manipulator) motion parameters during positioning.

Figure 19 shows an image from the RS gripper's camera of a vial (transport tube) with a QR code applied, detected and decoded by a 17-digit identification code.

Figure 19 – Operator's station screen view of a lab tare (vial) with the decoded information

Conclusion

In order to replace obsolete manipulators designed for technological operations at the nuclear industry enterprises, within the framework of the completed package of works a scientific and technical reserve was created and the MR-48 robotic complex prototype was developed and manufactured.

When performing this research and development work, there was proposed and implemented a basic control system [4], currently operating in the simulator training mode. At the present moment, the vision system algorithms are being perfected.

MP-48 robot-manipulator has the following advantages over the MEM-10 manipulator currently used:

1. Broader functionality allowing three control modes, including the automatic "Trajectory" mode.
2. MP-48 power consumption is 1 kW, which is significantly lower than the power consumed by MEM-10 (6 kW).
3. MP-48 slave arm tightness prevents the ingress of a chemically active medium which increases the wear resistance of kinematic pairs and positioning accuracy. In addition, the increased MP-48 maintainability is due to the modular design of the slave arm, in particular, the use of planetary cycloid gears, each of which is made in a separate package, unlike the long shafts with bevel gear pairs in MEM-10.

At the moment, the MP-48 manipulator prototype has an operation time of about 1,400 hours in various modes and operating conditions. Individual electromechanical components have been successfully tested for a significant absorbed dose.

References

1. E.I. Yurevich Robotics basics. - Saint-Petersburg: BHV-Petersburg, 2005. - 368 pp.
2. Anis Koubaa. Robot Operationg System (ROS). The Complete Reference (Volume 3)., Springer, 2019, 605 pp. DOI: 10.1007/978-3-319-91590-6.
3. Lentin Joseph. Mastering ROS for Robotics Programming. Packt Publishing, 2015, 481 pp. ISBN 978-1-78355-179-8.
4. I.V. Voynov, A.M. Kazantsev, B.A. Morozov, M.V. Nosikov Control system of the robotmanipulator with use of neural network algorithms of restriction of work area of the gripper, Bulletin of the South Ural State University. Series «Computer Technologies, Automatic Control & Radioelectronics» Vol 17, No 4, 2017. 29-36pp., DOI: 10.14529/ctcr170404
5. I.V. Voinov, I.F. Kruglov, B.A. Morozov, A.M. Kazantsev, M.V. Nosikov. Manipulator MR-48 for nuclear industry: patent. №172431 RU, 2016., F42D5/04.
6. I.V. Voinov ; M.V. Nosikov. Automatic and Manual Control Algorithms of Radiation-Proof Manipulators, Proceedings of 2018 Global Smart Industry Conference, IEEE Xplore., DOI:10.1109/GloSIC.2018.8570161
7. Hicham Tribak, Youssef Zaz. QR Code Patterns Localization based on Hu Invariant Moments. International Journal of Advanced Computer Science and Applications, Vol. 8, No. 9, 2017.

M.A. Akbarova, V.M. Bitnyi-Shliakhto, E.U. Smirnova, A.V. Popov

INTEGRATED SAFETY AND LABOR PROTECTION SYSTEM FOR HAZARDOUS INDUSTRIES

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
m.akbarova@rtc.ru, v_bitny@rtc.ru, eus@rtc.ru, apopov@rtc.ru

Abstract

Modern digital production (smart factory) is a cyber-physical production system (CFPS), which combines four main components: intelligent sensors, communications; cloud applications, terminal devices of the new generation. The traditional data sources for the security system are stationary sources of alarm and emergency information (detectors, alarms, TV cameras, etc.) and mobile means of monitoring the territory (ground and air). At the same time, factors related to the operation of mobile equipment, which is a source of increased danger to personnel, including in hazardous industries, remain outside the control. The advent of the Internet of things technology makes it possible to significantly expand the functionality of the security system by using information about the status of all units of equipment. The article introduces the concept of integrated industrial safety system (ISPS) and proposes the concept of working in a smart factory.

Keywords: digital production, smart factory industrial safety, labor protection, hazardous production.

Acknowledgments

This work was done as the part of the state task of the Ministry of Education and Science of Russia No 075-00924-19-00 "Methods of automatic synthesis of optimal control of the behavior of a group of robots on the basis of situational analysis with the use of semantic technologies».

Every year around the world there are hundreds of industrial accidents, most of which are the result of non-compliance with safety requirements and violation of the rules of work with the equipment. In hazardous industries, accidents are primarily caused by collisions with vehicles and moving machinery.

The main causes of accidents and deaths are:

– violation of safety requirements during operation of the equipment;

– violation of the organization of works on maintenance of the equipment;

– unsatisfactory organization of work during loading and unloading operations.

Modern digital production (smart factory) is a cyber-physical production system (CFPS), which combines four main components [1]:

– mobile smart sensors equipped with built-in technology to interact with each other or with the environment;

– means of communication, including the integration of point devices using peripheral computing technology;

– cloud application layer that provides semantic links between production data based on the ontological production model;

– terminal devices of new generation (smart phones, smart watches, etc. gadgets).

Safety and trouble-free operation are important tasks in any production and are currently provided by a set of organizational and technical measures, United by the concept of "industrial safety".

All security systems can be divided into:

– preventive, whose main task is the forecasting of emerging emergency situations to prevent;

– warning, are a "softening" layer of protection – their main purpose is to reduce the severity of the consequences of an emergency.

The traditional approach includes warning security systems - different types of sensors, alarms and simple automated workstations to which the signal is supplied, or video surveillance is conducted. In addition to sensors, the traditional approach includes route maps, workflows, and safety rules that serve to track operations and prevent emergencies.

Currently, industrial safety systems based on the principles and algorithms of machine vision are being widely implemented to provide forecasting and continuous monitoring of the situation. Along with the functions of measuring, processing, forecasting and documenting the parameters of the situation, the complex provides:

– rapid calculation of the boundaries of hazardous areas;

– the display of the hazardous areas on the TV picture of the monitored space or object;

– the display of the hazardous areas on the working surfaces using laser pointers;

– control of personnel access to hazardous areas with changing borders and planning of safe movement in these areas.

Preventive systems are extremely complex and require the use of intelligent equipment. At the moment, there are practically no examples of a preventive industrial safety system that allows avoiding emergency situations, assessing potential accidents.

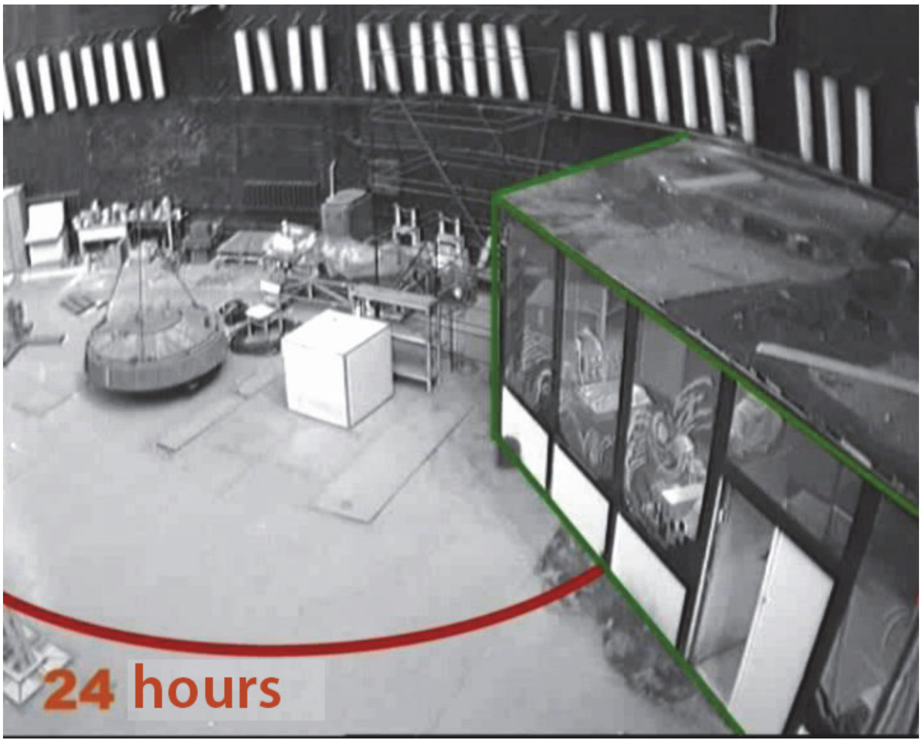


Figure 1 – The new generation complex for monitoring the situation and ensuring the safety of personnel based on the integration of sensors of ionizing radiation and STZ

A preventive security system can be created in a smart production environment and must be part of it. Objective forecasting of the situation is possible only with the use of intelligent sensors and vision systems.

Table 1. The structure of an integrated system of industrial safety

Subsystem	**Functional purpose**
Network of stationary sensors (detectors, TV cameras, etc.)	– Equipment status (on / off / canvas) – Detection of abnormal, alarming, dangerous, emergency situations
Mobile robotic means of territory monitoring	– Detection of abnormal, alarming, dangerous, emergency situations in areas not covered by stationary sensors
Navigation and communication	– Monitoring of movements of personnel and mobile equipment based on vision and radio frequency identification (RFID) technologies).
Cloud computing	– Formation of the operational map of danger zones in relation to the digital spatial model of the area and the enterprise – Multi-criteria analysis of the current situation – Situation forecasting – Support of decisions regarding industrial safety and labor protection
Terminal device	– Smartphones – Integrated means of informing staff functions operational display of danger zones, flow alarm when approaching dangerous equipment, select a safe route of travel

One of the key characteristics of the CFPS is proactivity [2], that is, the ability to form advanced solutions, including the identification of anomalies and alarming situations. In terms of occupational health and safety, the CFS is an environment with constantly changing areas of greater or lesser risk to personnel. In view of these features, ensuring trouble-free and safe operation of the CFS requires the introduction of appropriate means of industrial safety, integrated into an integrated system (ISPS). The most popular ISPS will be both at hazardous industrial facilities and at industrial sites, where many accidents are the result of collision of moving vehicles or mobile equipment [3].

ISPB should have the following basic properties.

1) Traditional stationary sources of alarm and emergency information (detectors, alarms, TV cameras, etc.) should be supplemented with mobile robotic means of territory monitoring (ground and air). Since the status of each piece of equipment (on, off, idle, etc.) is constantly known in the CFS, this data must be continuously processed in the ISPS to form an operational map of the danger zones. Means of monitoring the use of personal protective equipment based on technical vision systems should be introduced;

2) Important from a security point of view is the ability to localize personnel, as well as monitor compliance with the use of PPE. Navigation and communication facilities should provide continuous monitoring of movements of personnel and mobile equipment based on technology of technical vision and radio frequency identification (RFID) [4].

RFID system consists of antenna, receiver, transmitter and memory for data storage.

Currently, RFID method is actively used in logistics and warehouses. However, its broad capabilities and ease of operation can be used to control the movement of personnel in the workplace, as well as when working with various equipment.

The advantages of RFID technology are:

– the lack of requirements for line of sight;

– readability at a large distance and at high speed. For example, when using labels operating at ultra-high frequencies, reading is possible at a distance of up to 10 m at a speed of 128 kbps. Active microwave labels are used in real-time systems for determining the location of objects.

– resistance to moisture and dirt;

– long service life;

– multipurpose use;

– high degree of security due to unique identifier and data encryption algorithms.

The main disadvantages of RFID technology are:

– sensitivity to mechanical damage;

– high cost;

– the complexity of self-production;

– interference at electromagnetic impact;

– a small amount of technical solutions.

3) At the level of cloud applications should be provided with the formation of an operational map of danger zones in relation to the digital spatial model of the enterprise, multi-criteria analysis of the current situation, forecast and support solutions in terms of industrial safety.

Currently, a standard on ontology for robotics and automation has been developed [5] with the presentation of the fundamental concept. This standard also defines the methodology of ontological engineering used to construct ontologies. It is expected that this standard will simplify the programming of robotic solutions for the automation of technological processes, as well as expand the capabilities of robots in the field of information processing, self-learning and the construction of logical chains. Ultimately, this standard should facilitate the interaction between robots and humans, as well as between the machines themselves.

The purpose of this standard is to provide a methodology for the presentation of knowledge and reasoning in robotics and automation. The standard provides a unified way of representing knowledge and provides a common definition of terms, allowing unambiguous transfer of knowledge among any group of people, robots and other artificial systems.

With the growth of robotics production, the need for a standard and well-defined representation of knowledge is becoming increasingly apparent. The standard methodology of knowledge representation will allow:

– to define more precisely the concept of;

– to ensure mutual understanding between members of the community;

– will facilitate the integration of data and transfer of information among robotic systems.

The intended audience for this standard are robot manufacturers, system integrators, robot end users (parts manufacturers, automotive industry, construction industry, service and solution providers, etc.), robotic equipment suppliers, robot software developers and researchers/developers. Based on this approach, a number of projects have been implemented [6-8], and we plan to use this approach in ST. PETERSBURG.

4) At the level of terminal devices, special means of informing personnel should be developed. The concept of the device "smart helmet" (one of the options proposed in [9]) is proposed, which, among other things, will provide an operational display of danger zones of various kinds, the supply of warning signals when the employee approaches the working equipment, the choice of the safest route of movement.

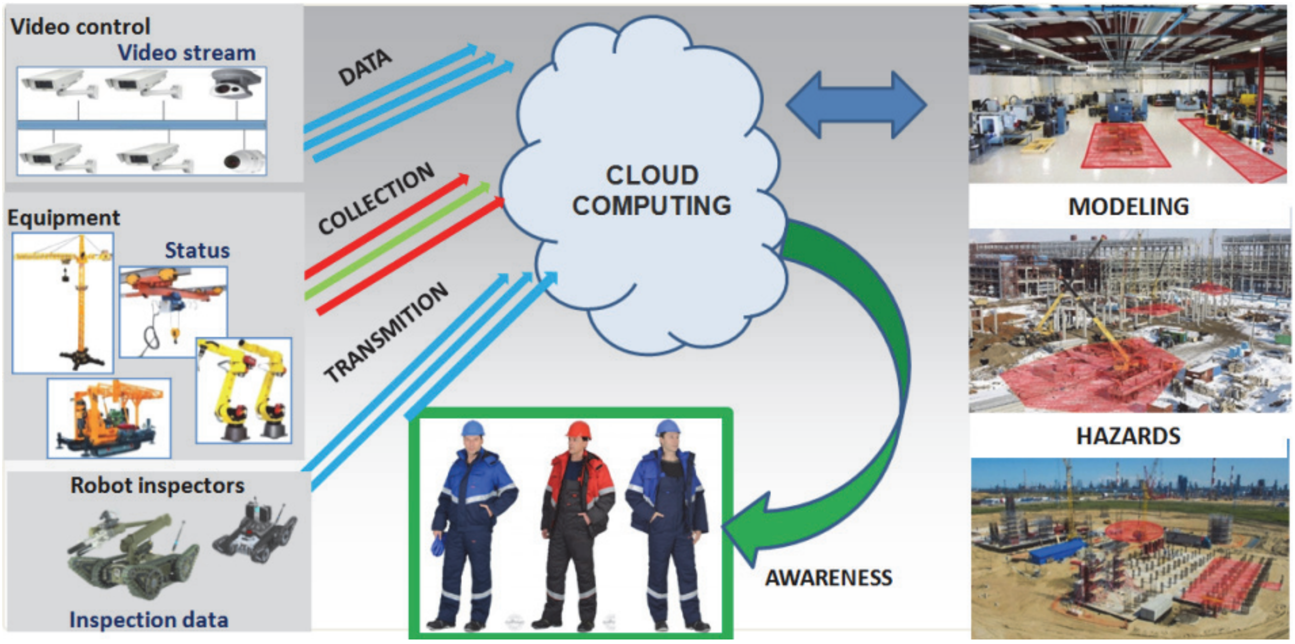


Figure 2 – Integrated security system based on CFPS

Thus, in order to increase the level of safety in production and reduce the level of injury, it is necessary to introduce an ISPB system based on intelligent sensors and vision systems and the use of concepts proposed in the ontology standard for robotics and automation to ensure the formation of an operational map of danger zones at the level of cloud applications.

References

1. B. Chen, J. Wan, L. Shu, P. Li. M. Mukherjee, and B. Yin, “Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges” IEEE Access, volume 6, pp.6505-6518, 2018.
2. R. Burke, A. Mussomeli, S. Laaper, M. Hartigan, B. Sniderman, “The Smart factory – responsive, adaptive, connected manufacturing” Deloitte University Press, 2017.
3. GOST 12.0.003-2015
4. S. Lu, C. Xu, R.Y. Zhong, L.Wang, “A RFID-enabled positioning system in automated guided vehicle for smart factories” Journal of Manufacturing Systems 44, pp 179–190, 2017.
5. IEEE 1872-2015 – IEEE Standard Ontologies foe Robotics and Automation
6. T. Haidegger, M. Barreto, P. Goncalves, M.K. Habib, S.K.V. Ragavan, H. Li, A. Vaccarella, R. Perrone, E. Prestes “ Applied ontologies and standards for service robots” Robotics and Autonomous Systems 61, pp 1215–1223, 2013.
7. S.R. Fiorini, J.L. Carbonera, P. Goncalves, V.A.M. Jorge, V.F. Rey, T. Haidegger, M. Abel, S.A. Redfield, S. Balakirsky, V.Ragavan, H. Li, C. Schelenoff, E. Prestes. “Extensions to the core ontology for robotics and automation” Roboticsand Computer-Integrated Manufacturing 33, pp. 3–11, 2015.
8. C. Schlenoff, A. Pietromartire, Z. Kootbally, S. Balakirsky, S. Foufou “Ontology-based state representations for intention recognition in human–robot collaborative environments” Robotics and Autonomous Systems 61, pp. 1224–1234, 2013.
9. https://softline.com/ru/smart_business/article/new-challenges-of-production-from-needs-to-solutions-softline

A.N. Afonin, E.L. Smovdarenko

NEURAL-CONTROL INTERFACE IN ROBOTICS

Belgorod State University, Belgorod, Russia
afonin@bsu.edu.ru

Abstract

The article describes the principle of operation, advantages and disadvantages of the most common neurointerfaces of robotic devices. Invasive and non-invasive devices operating on the basis of electrical (EMG, EEG, etc.) and chemical (fMRI, fNIRS, etc.) activity of the nervous system are described. The prospects for the use of neural-control interface in robotics are analyzed. It is concluded that the use of combined non-invasive neural-control interface is promising.

Keywords: brain-computer interface, robot, neurotechnology, control system, cyborg.

Acknowledgments

The study was carried out with the financial support of the Russian Foundation for Basic Research and the Administration of the Belgorod Region as part of a research project 18-48-310028 p_a.

Neural-control interfaces (brain-computer interfaces) are devices designed to directly exchange information between the nervous system and an electronic device. Currently, they are increasingly used in robotics. Since any cognitive activity of a person is accompanied by the activation of neuron ensembles corresponding to it, the principle of operation of most modern neural-control interface is based on brain mapping and identifying areas of its activity. Successful identification of activation of the desired brain zone allows interpreting it as a mental command to perform a particular action. By stimulating the corresponding neuron ensembles, it is also possible to implement the reverse transmission of information directly to the brain.

It is known that cognitive activity is carried out through the exchange of electrical impulses of the brain's neurons. At the same time, the electrical activity of neurons arises due to the chemical processes occurring in them.

Despite the fact that neurotechnologies have gained significant development only in recent years, existing neural-control interface are distinguished by diversity [1, 2, 3, 4]. The most important signs of neurointerface classification for robotic devices are their location in the body (invasive and non-invasive; in the central or peripheral nervous system) and the principle of action (Fig. 1).

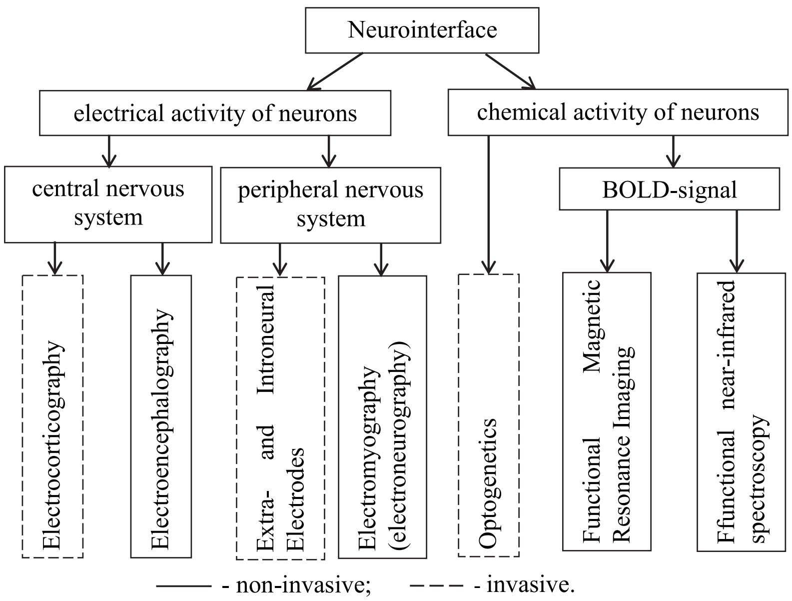


Figure 1 – Classification of the most common neural-control interface in robotics by location and principle of operation

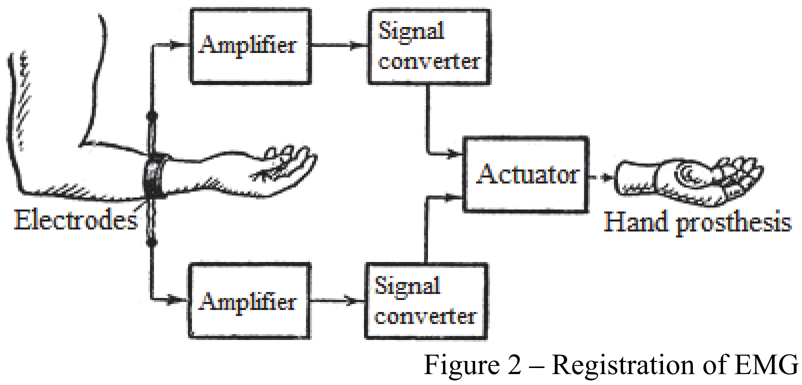
Invasive neurointerfaces are located inside the human body through surgical intervention. They make it possible to sufficiently successfully detect the activity of individual ensembles of neurons. When using invasive neural-control interface, the operator can usually give immediate mental commands to the control system to perform the required actions. In addition, only invasive interfaces today allow for feedback, passing information directly to the brain. However, installing them on the human brain poses a significant health hazard. Also, such interfaces overgrow with connective tissue, the neurons in contact with them die, which impairs their work and leads to the need for repeated surgery. In this regard, the use of invasive neural-control interface in healthy people is unacceptable for ethical reasons.

Non-invasive neural-control interface are installed on the skin surface and do not pose a health hazard. However, their sensitivity is several orders of magnitude worse than that of invasive interfaces, and therefore, using them, the activity of individual ensembles of neurons is extremely difficult. Commands to the control system when using non-invasive neural-control interface usually have to be given indirectly, generating clearly recognizable patterns in the brain, for example, mentally representing movements with hands and other parts of the body, each of which serves as a conditional signal for certain actions to be performed by a robotic device.

By the principle of operation, neural-control interface can be divided into those based on the analysis of electrical and chemical activity of the nervous system.

The oldest and most common are neural-control interface based on an analysis of the electrical activity of the nervous system. The most important advantage of these neural-control interface is a high reaction rate. The main disadvantage is the difficulty of identifying and recognizing signals. Depending on the location of the sensor electrodes, the neural-control interface based on the analysis of the electrical activity of the nervous system can be divided into the neural-control interface of the central and peripheral nervous systems.

Neurointerfaces based on the analysis of the electrical activity of the peripheral nervous system include electromyographic sensors and extraneous electrodes. Non-invasive electromyographic (EMG) sensors are located on the skin at any point of the body, except for the upper part of the skull. They capture, as a rule, electrical impulses from the work of muscles, but they can also remove much weaker electrical signals from the peripheral nervous system (in this case, the method is called electroneurography). EMG-sensors can be performed as in the form of separate disposable electrodes, and in the form of reusable bracelets with a variety of sensors. Since the signals perceived by them mainly reflect information about motor activity, EMG sensors today have found wide use for controlling bionic limb prostheses using residual stump muscles [5, 6, etc.] (Fig. 2). The first works in the field of similar prostheses were carried out in the USSR and Yugoslavia in the 1950s [6, 7]. At present, research in this field is successfully continuing both abroad [5, 8, and others], and in the Russian Federation [9, 10, and others]. There are a large number of industrially manufactured constructions of bionic limb prostheses with control from EMG sensors (Bebionic, Michelangelo, i-Limb, Stradivarius, etc.). Feedback in such prosthetic control systems can be implemented, for example, by applying electrical impulses to the skin of a stump.



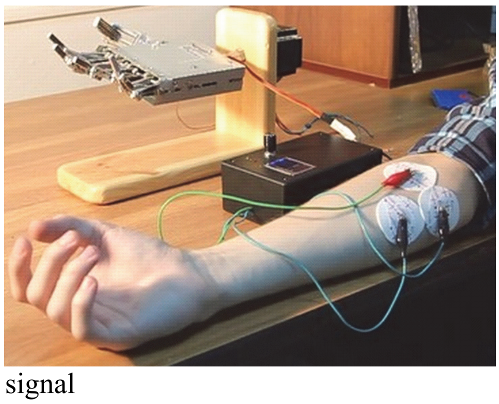


Figure 2 – Registration of EMG signal

Since bionic limb prostheses undoubtedly belong to anthropomorphic robotic manipulators, electromyography is the most successful example of the use of neural-control interface in robotics. However, its possibilities are limited only to the currently occupied niche. Its application for other purposes is impossible (rehabilitation of the paralyzed) or impractical (manipulator control).

The main advantage of invasive extra- and intraneural electrodes [8, 11, 12, etc.] is the possibility of implementing a more effective feedback with bionic prostheses. The difference between extra-and intraneural electrodes is that the former are attached to the nerve from the outside, and the latter are inserted into it. The design of the electrodes may be different. Due to the above disadvantages inherent to all invasive neural-control interface, control systems on extra- and introneural electrodes have not yet found wide practical application in robotic devices. At the same time, there is experience of successful use of these neural-control interface to control the behavior of beetles, dragonflies, butterflies, etc. in order to create "Cyborg Insects" [12, 13, etc.] (Fig. 3). Due to the supply of electric impulses into the nervous system of an insect by means of extra- or introneural electrodes, it is ensured that it moves along a trajectory set by a remote control or program. Cyborg insects are a very promising alternative to mechanical mobile microrobots in agriculture, military affairs and other fields, but work in this area has not yet gone beyond the scope of experiments.

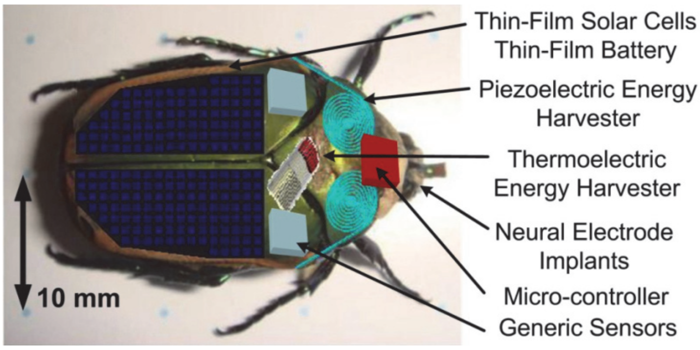


Figure 3 – Cyborg beetle with a neural-control interface based on extraneural electrodes

Neurointerfaces based on an analysis of the electrical activity of the central nervous system include electroencephalography and electrocorticography. In electroencephalography (EEG), electrodes are mounted on the skin of the upper part of the skull and pick up electrical signals from brain neurons. The location of the electrodes on the head is usually carried out in accordance with the so-called international system "10-20" (Fig. 4).

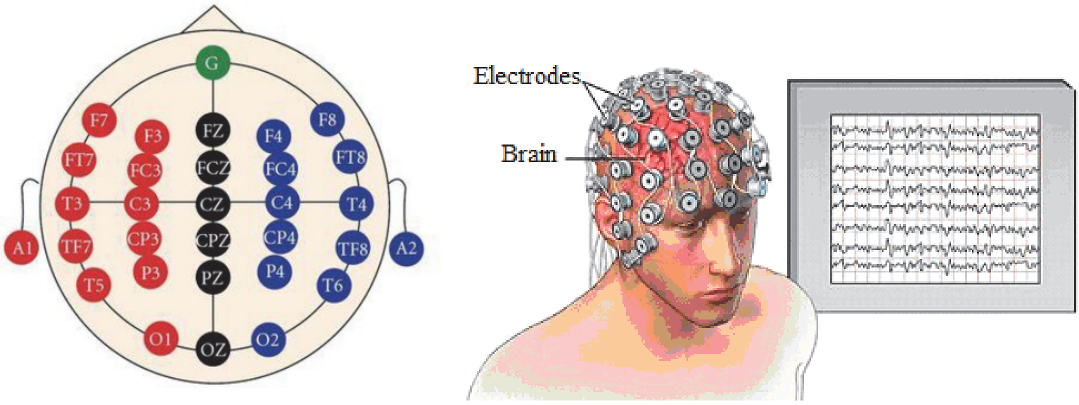


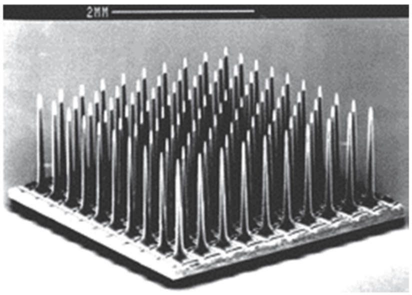
Figure 4 – System for installing EEG electrodes

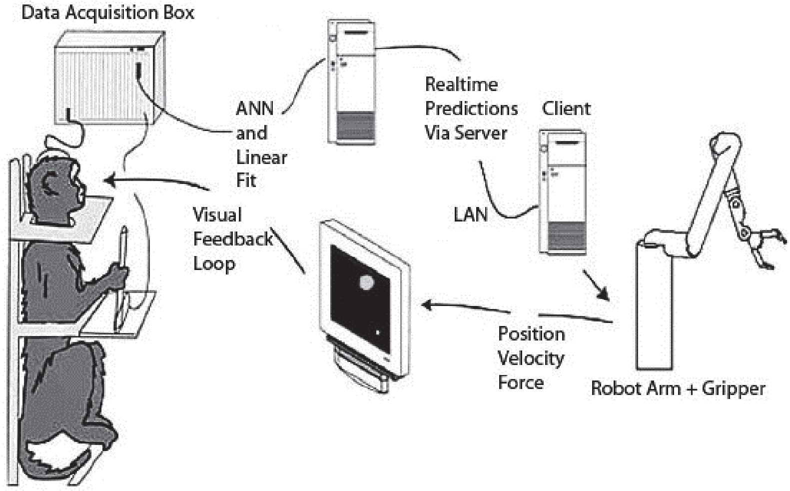
EEG is one of the oldest methods of brain research (known since the 1920s [14]). The attempts to use it as a neural-control interface are also very old. However, in relation to robotics, they should be considered unsuccessful. Since electrical impulses propagate at the speed of light and are significantly distorted when passing through the skin and bones of the skull, it is almost impossible to detect the activity of individual neuron ensembles using EEG sensors. In this regard, when creating neural-control interface based on EEG, as a rule, they are guided by integral indicators of the electrical activity of the brain, in particular, by different electrical rhythms, such as, for example, α-rhythm. Commands to the control system, as a rule, are not mental desires to perform a given action, but a certain complex set of more simply recognizable mental signals, for example, alternation of relaxation-concentrations in a predetermined manner.

The most frequent attempts were made to use EEG to control bionic prosthetic limbs and exoskeletons in order to rehabilitate people with disabilities [1, 2, 3, 8, 15, etc.]. Experiments were also carried out on control using neural-control interface based on EEG manipulators [16], mobile robots [17] and wheelchairs [18], multikopter [19], etc. For recognition of control commands, various methods were used, including artificial neural networks [20, 21, etc.]. However, the error of recognition of control commands when using EEG usually exceeds 50%. Such a percentage of errors is unacceptable for controlling moving objects. In this regard, neurointerfaces based on EEG, despite the large amount of work in this area, have found practical application only in cases where a large percentage of control system errors are uncritical: in devices for communicating paralyzed people using typing "Neurochat" and in numerous games and simulators.

In addition to the EEG, a method of magnetic encephalography (MEG) is also known, based on the measurement of the magnetic fields of the brain. However, since the measurement of magnetic fields is more difficult than the measurement of electric, it has not found wide application for neural-control interface.

Electrocorticography, an invasive neurointerface, installed directly on the cerebral cortex, shows high efficiency [1, 12]. Electrocorticography makes it possible to accurately register the activity of individual neuron ensembles. The most commonly used to register the electrical activity of neural ensembles are matrices of metal wire electrodes (Fig. 5, a). Electrodes in such a matrix can be rigidly fixed or moveable. Such matrices can be used to record the activity of the cerebral cortex and subcortical nuclei. In addition to wire electrodes, other constructions are also used, for example, an implant based on silicon.





a b

Figure 5 – Experiment on manipulator control using electrocorticographic neurointerface

There have been successful experiments in controlling the hands-manipulators with signals directly from the brain using electrocorticographic interfaces, conducted on monkeys since the 90s of the 20th century [22, 23, 24]. In the studies conducted, the monkey using a manipulator controlled by an electrode mounted on its cerebral cortex, to bring pieces of food to its mouth (Fig. 5, b). Similar studies were conducted on humans [25]: silicon-based implants were installed in fully paralized patients in the motor cortex of the brain, which recorded the electrical activity of several dozen neurons. Patients have learned to mentally control the cursor on a computer screen and a robotic arm. Using a manipulator, in particular, they could bring a cup of coffee to their mouths, and later even control various devices with the help of buttons. However, electrocorticographic interfaces have all the disadvantages of the invasive neural-control interface described above. In this regard, today they are used only in animal experiments and in individual experimental attempts at the rehabilitation of completely paralyzed people.

Neurointerfaces based on the chemical activity of the nervous system, allow you to determine the change in the concentration of certain substances that affect the activity of neurons, and some and controle it. Their main advantage is the ability to detect the activity of individual neuron ensembles even when using non-invasive interfaces. The main disadvantage of these neural-control interface is the low reaction rate, since the speed of chemical processes is much lower than the speed of electrical ones. As a rule, these neurointerfaces are based on the peculiarities of the interaction of various chemical compounds with radiation fluxes.

Since the activation of neurons in the brain begin to absorb several times more oxygen than in an inactive state, most modern non-invasive neural-control interface based on an analysis of the chemical activity of the nervous system are based on the BOLD principle (blood-oxygenation-level-dependent). The BOLD principle is based on detecting hemoglobin concentrations (oxygenated and deoxygenated) in oxygenated areas that bring oxygen to neurons.

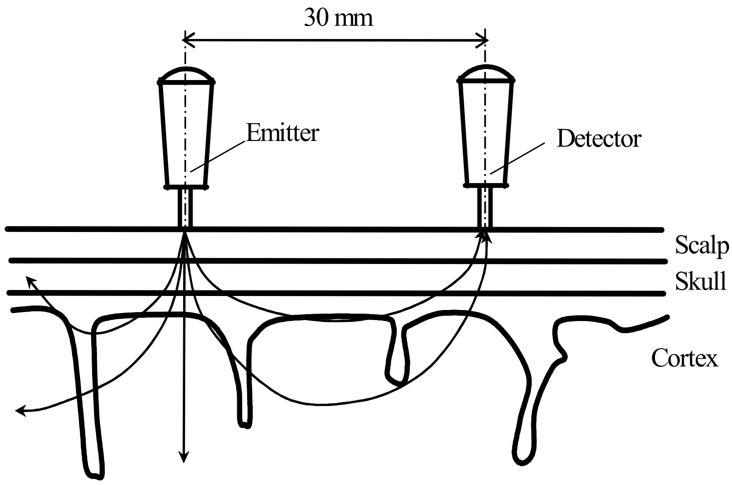
The most powerful non-invasive neurointerface based on the BOLD-principle today is functional magnetic resonance imaging (fMRI) [26, 27]. The principle of fMRI is based on measuring the electromagnetic response of atomic nuclei in a strong magnetic field. fMRI allows to make volumetric maps of neuron activity with a resolution of up to 1 mm throughout the whole brain volume. Successful control experiments were performed using a fMRI-based neurointerface with android robots [28] and manipulators

[29]. However, fMRI when taking measurements requires immobility of the patient. In addition, equipment for fMRI very cumbersome, energy-intensive and has a high cost. Thus, the use of fMRI directly as a neural interface for robotic devices today is not rational. fMRI can only be used when setting up other neural-control interface under laboratory conditions [30].

The lack of fMRI is deprived of another neural interface based on the BOLD-principle - functional near-infrared spectroscopy (fNIRS) [31, 32, 33, 34]. fNIRS is based on fNIRS in the near-infrared range and uses near-infrared radiation to measure the optical absorption spectrum of hemoglobin. fNIRS as a tool for monitoring hemodynamics of the brain appeared more than 30 years ago. The fNIRS scheme is presented in fig. 6. Infrared rays, passing from the emitter through the skin, the bones of the skull and the cerebral cortex, fall on the sensor. Depending on the number of rays absorbed in the body, the device calculates the concentration of hydroxy and deoxyhemoglobin. Photodiodes are commonly used as emitters and sensors. The emitters and sensors on the head are located, as a rule, similarly to the EEG in accordance with the "10-20" system.

The main advantages of fNIRS compared to fMRI include portability and the relatively low cost of equipment and the absence of serious limitations to the operator's physical activity. The greatest disadvantages of fNIRS y are its spatial resolution (not more than 3 cm deep), as well as a time delay of 3-5 seconds in identifying areas of brain activity associated with the inertia of the processes of inflow and outflow of blood.

There is a relatively successful experience in using fNIRS to control bionic limb prostheses [35, 36, etc.], android robots [37, 38], wheelchairs [18], etc. An example of the successful application of fNIRS in the gaming industry is the device "Optorhythmograph", which allows you to control the characters of computer games using the fNIRS interface. However, the low reaction rate of fNIRS makes it difficult to use it to control moving robotic devices.



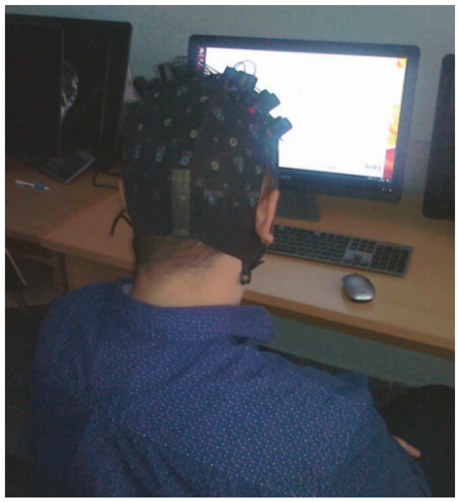


Figure 6 – Scheme of fNIRS

The invasive neurointerfaces based on the control of the chemical activity of the nervous system include methods of optogenetics [32, 39]. The essence of these methods lies in the fact that using genetic engineering, nerve cells make them sensitive to electromagnetic radiation of a certain range. The method is one of the newest, it appeared in the mid-2000s. To make neurons sensitive to visible light, they are injected with proteins sensitive to it, such as rhodopsins. For the introduction of photosensitive proteins, various methods of genetic engineering can be applied, for example, the creation of transgenic organisms in which this protein is incorporated from birth or a viral transduction method in which this protein is introduced into neurons using genetically modified viruses. To provide light radiation to the necessary groups of neurons are fed beams of thin optical fibers. Due to the irradiation of photosensitive neurons with light, chemical reactions occur in them, changing their electrical potential, similar to those that occur during their natural activation. Compared with the activation of neurons by electric current using invasive electrodes, this method is more selective and less traumatic. To remove information about the activity of neurons, along with optometers, they introduce electrodes-sensors. A type of optogenetics is a method of thermogenetics developed by Russian scientists [40], in which neurons are made sensitive to infrared radiation.

Due to the low level of knowledge and potential danger, the methods of optogenetics are currently at the stage of animal experiments. Their use on humans in the foreseeable future will not find wide distribution. From the point of view of robotics, they can be interesting so far only as a means of creating cyborg animals.

Greater selectivity compared with other invasive neurointerfaces will allow the use of optogenetics to control highly developed mammals, such as laboratory mice and rats [41].

Based on the analysis performed, it can be concluded that existing invasive neurointerfaces cannot yet be widely used for use in humans due to health hazards, and non-invasive ones do not allow full control of moving objects. It follows from this that the most promising in the near future is the use of combined non-invasive neurointerfaces in robotics, for example, based on a combination of EEG and fNIRS. In these neural-control interface, the disadvantages of one method are compensated for by the advantages of another. There are successful examples of the implementation of such combined neural-control interface to control bionic prostheses [42], quadrocopters [43] and other robotic devices.

Thus, we can draw the following conclusions:

1. For invasive neural-control interface, the main area of application in the near future will be animal experiments, including the creation of cyborg animals. The use of invasive neural-control interface on a person, including for the rehabilitation of disabled people, will be widely used only after the emergence of fundamentally new, secure neural-control interface of this kind. At the same time, since work in the field of invasive neural-control interface is sometimes highly controversial from an ethical point of view, they should be carried out under strict public control.

2. For bionic prostheses, exoskeletons, wheelchairs and other robotic devices intended for the rehabilitation of disabled people, combined non-invasive neurointerfaces, in particular neurointerfaces based on a combination of EEG and fNIRS, will be widely used.

3. Robot-based games and simulators with neural-control interface based on EMG, EEG and fNIRS will find more and more widespread use.

4. The use of neural-control interface for other robotic devices (mobile robots, UAVs, manipulators, etc.) will have only scientific value until new safe invasive neural-control interface appear, since non-invasive neural-control interface have no significant advantages over traditional control systems for healthy people.

References

1. Lebedev M.A. Neurocomputer interfaces for expanding brain functions // Science and innovation in medicine. - 2016. - №3. - Pp. 12 – 27 (in Russian).
2. Hassanien A.E., Azar A.T. Brain-Computer Interfaces. Current Trends and Applications. - Springer. - 2015. -422 p.
3. Wolpaw J.R., Wolpaw E.W. Brain-Computer Interfaces: Principles and Practice. - Oxford: Oxford University Press. - 2012. - 419 p.
4. Brain-Computer Interfaces Handbook: Technological and Theoretical Advances / Editet by C.S. Nam, A. Nijholt, F. Lotte - Boca Raton: CRC Press, 2018. - 814 p.
5. Muzumdar A. Powered Upper Limb Prostheses: Control, Implementation and Clinical Application. - Berlin: Springer-Verlag. - 2004. - 220 p.
6. Bioelectric control / V.S. Gurfinkel, V.B. Malkin, M.L. Tsetlin, A.Yu. Schneider - M .: Science. —1972. - 245 p. (in Russian).
7. Tomovic R. The hand of man as a feedback system. - M .: Publishing House of the Academy of Sciences, 1969. - 13 p. (in Russian).
8. Navarro X., Krueger T.B., Lago N. A critical review of the radio system and the hybrid bionic systems // Journal of the Peripheral Nervous System, 2005, No. 10 - Pp. 229–258.
9. Issues of constructing neuro-controlled prostheses / D.R. Safin, I.S. Sawers, M.A. Urakseev, R.M. Migranova // Medical equipment. - 2009. No. 4. - Pp. 16–21 (in russian).
10. Development and implementation of a mock-up of a bionic prosthetic hand / A.N. Afonin, A.Yu. Aleinikov, A.R.Gladyshev, A.V. Popova // Robotics and Technical Cybernetics, 2016, No. 3 (12). - Pp. 68-71 (in Russian).
11. Implantable Bioelectronics / Edited by Evgeny Katz. - N.-Y .: Wiley, 2014. - 472 p.
12. Thompson C.H., Zoratti M.J., Langhals N.B. and Purcell E.K. Regenerative Electrode Interfaces for Neural Prostheses // Tissue Engineering, 2016, Part B, Vol. 22, No. 2. - Pp. 125 - 135.
13. Sato H. Maharbiz M.M. Frontiers in Neuroscience, 2010, Volume 4, Article 199.
14. Haas L.F. Hans Berger (1873-1941), Richard Caton (1842-1926), and Electroencephalography // Journal of Neurology, Neurosurgery & Psychiatry, 2003, Vol. 74 (1). - Pp. 9.
15. Control of a robotic exoskeleton based on the technology of brain-computer interface of a motor-imaginary type / S.Yu. Gordleeva, M.V. Lukoyanov, S.A. Mineev, A.Ya. Kaplan et al. // Modern Technologies in Medicine, 2017. —T. 9, No. 3. - Pp. 31-38 (in Russian).

16. Salazar-Gomez, A.F., DelPreto, J., Gil, S., et al. Correcting robot mistakes in real time using EEG signals // International Conference on Robotics and Automation (ICRA), 2017.
17. Milan Jose del R., Renkens F., Mourino J., Gerstner W. Non-violent Brain-Actuated Control of a Mobile Robot by Human EEG // Transactions on Biomedical Engineering, 2004, Vol. 51, No. 6. - Pp. 1026-1033.
18. Smart Wheelchairs and Brain-computer Interfaces / Edited by Pablo Diez. - Academic Press, 2018. – 473 p.
19. LaFleur K., Cassady K., Doud A. et al. Quadcopter control in a three-dimensional space using a noninvasive motor imagery-based brain-computer interface // Journal of Neural Engineering, 2013. Vol. 10, № 4.
20. Classification of electroencephalographic patterns of imaginary finger movements for the development of the brain-computer interface / L.A.Stankevich, K.M. Sonkin, J.V. Nagornova, Yu.G. Khomenko, N.V. Shemyakin // Proceedings of SPIIRAN, 2015, №3. - Pp. 163-182 (in Russian).
21. The use of artificial neural networks for the analysis of human brain activity during imaginary movements / S.A. Kurkin, V.Yu. Musatov, A.E. Runnova, A.E. Khramov // Design of machines, robots and mechatronic systems. Sat mater vseros. scientific methodological conf. - Orel: O. Turgenev OGU, 2017. - Pp. 10 – 11 (in Russian).
22. Levitskaya, OS, Lebedev, MA The brain-computer interface: the future in the present // Bulletin of the Russian State Medical University, 2016, №2. - Pp. 4 – 16 (in Russian).
23. Carmena, J.M., Lebedev M.A., Crist R.E. et al., Learning to control the brain by the machine interface // PLoS Biol, 2003, Vol. 1 (2). - Pp. 193 - 208.
24. Velliste, M., Perel S., Spalding M.C. et al., Cortical control of a prosthetic arm for self-feeding // Nature, 2008. 453 (7198). - Pp. 1098-1101.
25. Hochberg L.R., Donoghue J.P. Sensors for braincomputer interfaces // IEEE Eng Med Biol Mag. - 2006, vol. 25 (5). - Pp. 32 - 38.
26. Kremneva E.I., Konovalov R.N., Krotenkova M.V. Functional magnetic resonance imaging // Annals of Clinical and Experimental Neurology, 2011, № 1, Volume 5. - Pp. 30-34 (in Russian).
27. fMRI: Spin to Brain Functions / Edited by K. Uludağ, K. U K.urbil, L. Berliner. - N.-Y .: Springer, 2015. - 929 p.
28. Cohen O., Druon S., Lengagne S. et al. fMRI Robotic Embodiment: A Pilot Study // IEEE RAS / EMBS International Conference on Biomedical Robotics and Biomechatronics, 2012. - Pp. 314-319.
29. Lee J.-H., Ryub J., Jolesza F.A. et al. Brain – machine interface via real-time fMRI: Preliminary study on thought-controlled robotic arm // Neuroscience Letters, 2009 .; Vol. 450 (1). - Pp. 1–6.
30. Barrios L.J., del Castillo M.D., Serrano J.I., Pons J.L. A review of fMRI as a tool for enhancing EEG-based brain-machine interfaces // Applied Bionics and Biomechanics, 2012, vol.9. - Pp. 125–133.
31. Ferrari M., Quaresima V. A brief review on the history of human functional and infrared spectroscopy (fNIRS) development and fields of application // Neuroimage, 2012, Vol. 63. - Pp. 921–935.
32. Chen Y., Kateb B. Neurophotonics and brain mapping. - Boca Raton: CRC Press, 2017. - 587 p.
33. Near infrared tomography and spectroscopy in the study of brain activity / K.V. Kvashneva, V.A. Ilyukhina, E.V. Kryzhanovsky, A.V. Chistov // Biotechnosfera, 2013, №2 (26). - Pp. 33 – 37 (in Russian).
34. Sitnikova, MA, Nyurk, G.Kh. Functional optical tomography: a reliable method for measuring brain activation in the process of solving various mathematical problems // Modern problems of science and education, 2016. - №3 (in Russian).
35. Bianchi T., Croitoru N.I., Frenz M. et al. NIRS monitoring of muscle contraction to control a prosthetic device // Proceedings of SPIE - The International Society for Optical Engineering, 1998, Vol. 3570. - Pp. 157-163.
36. Afonin A.N., Asadullaev R.G., Sitnikova M.A. Data analysis of the fNIRS tomograph for the control of limb prostheses using the brain-computer interface // Scientific and Technical Bulletin of the Volga region, 2018, №11. - Pp. 182 – 185 (in Russian).
37. Batula, A.M., Kim, Y.E., Ayaz.H., Motor-Imagery-Optical Brain-Computer Interface, Virtual and Actual Humanoid Robot Control with Four-Class // BioMed Research International, Vol. 2017, Article ID 1463512.
38. Matsuyama Y., Ochiai N., Hatakeyama T., Noguchi K. Multimodal human-humanoid interaction using motions, brain NIRS and spike trains // Proceedings from the 5th ACM / IEEE International Conference on Human-Robot Interaction, 2010. - Pp. 173 -174.

39. Implantable devices for optogenetic research and stimulation of excitable tissues / M.V. Matveev, A.I. Erofeev, S.G. Terekhin, P.V. Plotnikova, K.V. Vorobev, O.L. Vlasova // Scientific and technical statements SPbGPU. Physics and Mathematics, 2015, No. 3 (225). - Pp. 75 – 85 (in Russian).
40. Roshchin M., Ermakova Y.G., Lanin A.A. et al. Thermogenetic stimulation of single neocortical pyramidal neurons transfected with TRPV1-L channels // Neuroscience Letters, 2018, vol. 687. - Pp. 153 - 157.
41. Park S.-G., Jeong Y.-C., Kim D.-G. et al. Medial preoptic circuit and prey // Nature Neuroscience, 2018, vol. 21 (3). - Pp. 364-372.
42. Yinfeng F., Nalinda H., Dalin Z. et al. Multi-Modal Sensing Techniques for Interfacing Hand Prostheses // IEEE Sensors Journal, 2015, Vol. 15, Iss. 11.
43. Khan M.J., Hong K.-S. Hybrid EEG – fNIRS-Based Eight-Command Decoding for BCI: Application to Quadcopter Control // Frontiers in Neurorobotics, 2017, Vol.11.

M.A. Chumichev [1], D.A. Gribkov [2], V.E. Pavlovsky [2], I.A. Orlov [2]

A MODEL OF THE PNEUMATIC ARTIFICIAL MUSCLE

[1]Lomonosov MSU, Moscow, Russia, chym4@mail.ru
[2]Keldysh Institute of Applied Mathematics, Moscow, Russia,
vlpavl@mail.ru, legovas@gmail.com, i.orlov@keldysh.ru

Abstract

This paper gives a mathematical description of the pneumatic artificial muscle. The fundamental of such a mechanism is conversion the energy of expanding gas to elastic elements tension and subsequently to the mechanical work purposing to move objects. In this article we will consider the work of a single pneumatic muscle set in motion by a compressor and research its movements, as well as various objects to the muscle both under the effect of gravity without it.

Keywords: artificial muscle, mathematical model, exoskeleton.

Acknowledgments

The research was made with the financial support of RFBR within scientific projects no. 18-31-20068 и № 18-08-01441.

Introduction

In the modern world various helping systems are increasingly developing. They are created to help a person to perform hard work, which require great physical strength or cannot be produced without certain equipment or can cause harm to physical health. These devices are divided into various types: personal protective equipment, mechanical devices, automated systems, robotic systems and autonomous robots, which completely exclude human participation.

The separate direction stands out among the robotic systems - exoskeletons. Today, they not only help to carry out difficult and dangerous work but for the rehabilitation of patients with locomotor impairment.

Exoskeletons, according to the principle of work, can be active or passive. Active devices independently help the person for a given scenario. Passive exoskeletons use a sensory system to determine the action that a person is trying to perform in order to help him.

Exoskeletons divide by type of construction into two main types - rigid and flexible. Rigid exoskeletons are analogous to the external skeleton, which is represented in the animal world. But in practice, these designs are cumbersome and inconvenient to wear, as well as expensive for mass implementation in the production process. Therefore, scientists decided to replace rigid structures with soft and elastic ones. A human can put on separate parts of the body like clothes. In this way, devices appeared that became known as flexible exoskeletons.

The development of exoskeleton is of great scientific interest to specialists in the field of mechanics. To create such a device, it is necessary to simulate the action of a person, to describe the equations of the limb movement trajectories and etc. Control, stabilization, and device positioning systems are created based on these equations. Next, you need to choose the types of drives to perform the task. To do this, it is necessary to solve a variety of problems: theoretical and mechanical, problems of the rigidity, and strength of structures

One of the main problems in the creation of flexible exoskeletons today is the selection and development of new types of drives. Among them may be elastic and stretchy cables, harnesses, and artificial muscles. Artificial muscle works by reducing elastic elements, similar to the contractions of real human muscles. This type of device is similar to a real human muscle and can replace the actions of limbs, working in a complex of several muscles or independently. These devices can be used to move loads, reduce the load on the locomotor system, or perform the role of an active controlled spring in stabilizing the movement of mechanical systems.

For the first time, the design of an artificial pneumatic muscle was proposed by the British doctor Jerry McKibben for orthopedic purposes in the 60s of the last century. At the same time a modern name was given to this device due to its similarity with a real muscle (there is a similarity both visual (Fig. 1) and, in principle of creation the tractive effort by reducing fibers). A modern artificial muscle is a thin-walled elastic tube - 1, reinforced with a strong, elastic braid - 2, consisting of fibers laid at a certain angle crosswise.The edges of the braid are firmly fixed to the ends of the tube. Attached to the ends of the tube are devices (loops) for attaching other muscles or objects. The supply from the compressor is attached to one of the ends. When the air pressure inside the muscle increases, the muscle changes its volume by stretching the tube. Due to the elastic properties of the braid muscle increases in diameter and reduces in length, creating a pulling force at the ends of the muscle.

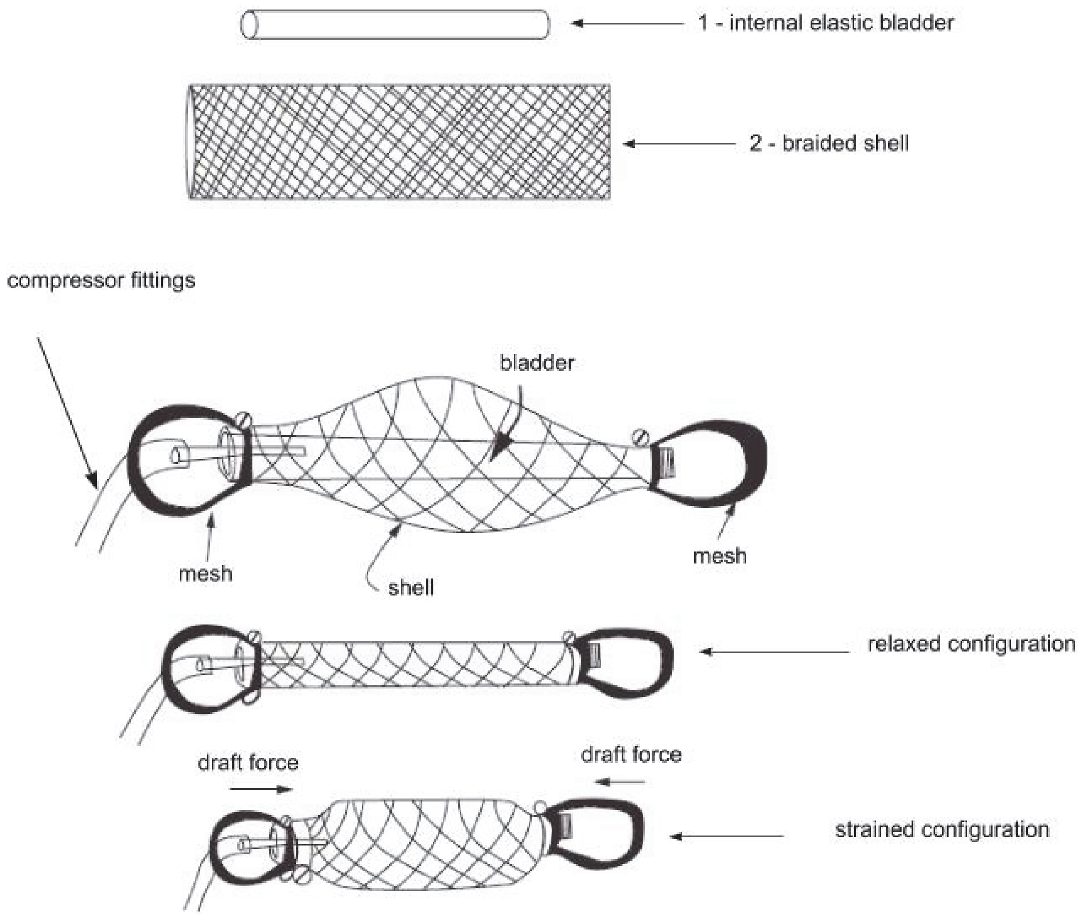


Figure 1 – Artificial muscle device

Figure 2 – Artificial muscle from the company Festo

The first structured mechanical description and study of pneumatic muscle was presented in [1]. It is believed that the shell of the tube itself is infinitely thin, and when pressure and volume change, the muscle remains cylindrical. Further studies of scientists were aimed at numerical analysis and engineering of sheaths of various materials, to achieve greater accuracy and reduce the error in muscle contraction, as well as to achieve maximum reliability.

Unlike other pneumatic actuators, such as piston devices and pneumatic cylinders, the work of a pneumatic muscle depends to a much lesser extent on external temperature, humidity and other factors. In the case when the ends are sealed, the use of muscle is possible even in water and at depth. Air is pumped through the compressor through the muscle, without special filtration and injection of oil impurities, so there is no need for special maintenance, and there will be no stagnation during operation.

Most of the existing research is aimed at compiling static formulas and staging static experiments to determine the dependence of the traction force at the ends of a muscle depending on the applied pressure and various braids. For example, in [1] and [3] are presented formulas for calculating the tractive force, depending on pressure, with the setting of further experiments. In these studies, a hysteresis is observed on the numerical graphs of the force due to compression, and therefore the instability of the work and the error are obtained.

In accordance with all of the above and the results, it can be concluded that by now most of the research is focused on the engineering aspect of creating devices (developing various types of braids, creating sensory systems of artificial muscles, for determining deformations and stresses on the surface, and building a numerical model). The purpose of this work is to develop a rigorous mathematical model and write the kinematic and dynamic equations of muscle movement, as well as the subsequent assessment of the compliance of this model with a real experiment.

General formulation of the problem

The subject of this research is to study the behavior of the pneumatic muscle in the active mode. From [1] and [3], the principle of building a muscle model and its tractive effort at the end due to the deformation of the braid is taken.

The work will be considered a single muscle, one end of which is fixed. The role of the braid is played by an elastic thread, wound on an elastic tube several turns at a certain angle. When air is supplied through the compressor, the muscle will begin to increase in volume, and due to the elastic braid, increasing in diameter, muscle contraction in longitudinal dimensions will occur. Since the air velocity in the muscle is not high (compared with the speeds at which uneven pressure is observed and various dynamic effects during gas movement), as well as with a working pressure up to 5-8 bar, we can assume that we are dealing with an ideal incompressible gas.

First it is necessary to investigate the movement of the free end. Further, expressing the quantities of interest (in particular, force) as a function of time, we write the equations of dynamics for a material point fixed at the end of the muscle. After that, it will be possible to find out from the obtained results: how quickly the contraction process occurs in the case of a free muscle, and in the case when it is under load; dependence of the obtained values (maximum deformation, time of contraction, traction effort) on the main parameters of the muscle, such as the initial length, diameter and length of the wound braid.

To solve the problem described above, we construct a mathematical model. Having obtained the necessary kinematic equations, we will determine the main parameters of interest and construct graphs of solutions of specific equations depending on these parameters.

The experimental part of the work is to create a graphic illustration of the process of functioning of a pneumatic muscle in various modes of operation. After conducting virtual experiments, it is necessary to assemble an experimental setup and compare the data of virtual and real experiments. The end result of the experimental part will be a conclusion about the compliance of the mathematical model with the experiment and the identification of optimal modes of muscle operation (the greatest reduction with the best accuracy).

Construction of a mathematical model

Consider an artificial pneumatic muscle, in which the left end is fixed. We assume that the force of gravity acting on the shell is negligible. The muscle is a thin-walled cylinder on which a thread of length b is wound, n is the number of turns of the thread. The initial length and radius of the muscle are x_0 and y_0, respectively. θ - the angle between the wound thread and the axis Ox. By virtue of circular symmetry, we will consider the plane problem of moving a rectangle.

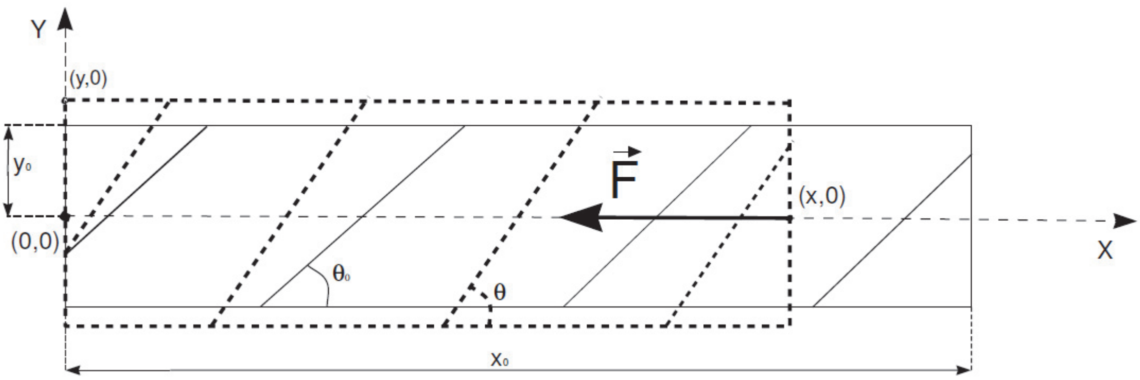


Figure 3 – Artificial muscle

In addition to the assumption that the muscle walls are negligibly thin, we assume that the muscle remains cylindrical when contracted (in fact, the experiment shows that except for the ends on which the braid is fixed, the muscle remains cylindrical).

The hose from the compressor is brought to the left end of the muscle. The diameter of the hose (inlet) — d_0. Characteristics of the compressor is consumption — $Q \ [m^3/s]$.

$$Q = \pi d_0^2 W \tag{1}$$

$W \ [m/s]$ — air velocity at compressor outlet.

Connection of geometrical parameters of the system:

$$\begin{cases} x = b\cos\theta, \\ 2\pi n y = b\sin\theta, \\ 0 < \theta < \frac{\pi}{2}, \\ V = \pi y^2 x. \end{cases} \tag{2}$$

V — muscle volume.

By virtue of the fact that the number of winding turns remains constant, it is possible to get rid of this parameter *n* by expressing it through x_0 and y_0.

$$n = \frac{b}{2\pi y_0}\sqrt{1 - (\frac{x_0}{b})^2} \tag{3}$$

Next, we consider the static problem: Denote by p the pressure difference created in the muscle and atmospheric. Then the work of an ideal gas will be equal to $A = pdV$. At the same time, this work is transformed into work on moving the right end of the muscle $A = (\vec{F}, \vec{dr}_x) = Fdx$, where F — this is a pull on the end of the muscle, dr_x — the displacement vector of the rightmost point, F and dx are the projections of these vectors to the axis Ox. Equating these equalities and substituting in them the geometric expressions of their system 2 and equality 3 we get:

$$F = \frac{p\pi y_0^2}{1 - (\frac{x_0}{b})^2}(1 - 3cos^2\theta) \tag{4}$$

From this formula it can be seen that there is a critical value of the angle at which the projection on the axis Ox of the pulling force changes sign: $\theta_{extr} = arccos\frac{1}{\sqrt{3}}$

Write down new restrictions on muscle movements, taking into account the obtained critical angle:

$$\begin{cases} \frac{b}{\sqrt{3}} = x_{extr} < x < x_0, \\ y_0 < y < y_{extr} = y_0 b\frac{\sqrt{\frac{2}{3}}}{\sqrt{b^2 - x_0^2}}, \\ 0 < \cos\theta < \frac{1}{\sqrt{3}}. \end{cases} \tag{5}$$

Rewrite the expression for strength as follows:

$$F = \frac{p\pi y_0^2}{1 - (\frac{x_0}{b})^2}(1 - 3(\frac{x}{b})^2) \tag{6}$$

Thus, we obtain the dependence of the traction force, which is directed in the direction of decreasing *x* coordinate, on the length of the muscle, with constant pressure inside the muscle.

Comparing the time required to fill a gas with a constant volume obtained in [2] and the time obtained for an ideal gas using the Mendeleev-Clapeyron law, we introduce the assumption that the pressure equalization time in the muscle is negligible compared to the contraction time. If we turn to the experiment and observe the work of the artificial muscle, it is clear that this assumption is quite reasonable.

Proceeding from all of the above, using all the above assumptions, as well as the equalities 1, 5, 6 and the relation $pdV = Fdx$, we get the complete mathematical model of our problem.

Derivation of the kinematic equations of the free end

Next, we derive the kinematic equations of muscle movement. The volume of air supplied by the compressor during dt is $Qdt = \pi d_0^2 W$. At the same time, the muscle volume will change to$dV = d(\pi y^2 x)$.

The main task of this section is to obtain expressions for the position of the free end - *x* and the diameter of the muscle - *y* as functions of time.

To begin with, from system 2 and equality 3 we obtain the following expression for y:

$$y = \frac{y_0}{\sqrt{1-(\frac{x_0}{b})^2}}\sqrt{1-(\frac{x}{b})^2} \quad (7)$$

Neglecting the time pressure equalization in the muscle, we can write the following equality:

$$Qdt = dV \quad (8)$$

Which can be rewritten by substituting the expression for y in the following form:

$$Qdt = \frac{\pi y_0^2}{(1-(\frac{x_0}{b})^2)}(1 - 3(\frac{x}{b})^2)dx \quad (9)$$

Integrating the obtained equality and substituting the initial conditions ($x(0) = x_0$), we obtain the implicit dependence of the coordinate $x(t)$:

$$x^3(t) - b^2x(t) + \frac{b^2}{\pi y_0^2}(1 - (\frac{x_0}{b})^2)(Qt + \pi y_0^2 x_0) = 0 \quad (10)$$

Ultimately, muscle contraction is described by a system of two kinematic equations, one of which describes muscle contraction, and the second - the change in its diameter.

$$\begin{cases} x^3(t) - b^2x(t) + \frac{b^2}{\pi y_0^2}(1 - (\frac{x_0}{b})^2)(Qt + \pi y_0^2 x_0) = 0, \\ y(t) = \frac{y_0}{1-(\frac{x_0}{b})^2}\sqrt{1-(\frac{x(t)}{b})^2}, \\ \frac{b}{\sqrt{3}} < x(t) \leq x_0, \\ y_0 \leq y(t) < y_0 b\frac{\sqrt{\frac{2}{3}}}{\sqrt{b^2-x_0^2}}. \end{cases} \quad (11)$$

The limitations written in the form of inequalities for x and y are due to the fact that muscle contraction cannot be greater than a certain value, due to the structure of the structure. As was shown above, at a certain angle of braiding the traction force changes its sign, we will not consider this case in the framework of this work, because the reduction is planned to be done by forcing air, and the relaxation is performed by pumping.

Derivation of dynamic equations

In this section we will consider the same system, only now the material point of mass m will be fixed at its right end, with radius-vector $\boldsymbol{dr_m}$ and coordinates $(x(t),0)$.

Let us write Newton's second law for this point, in the case when only the muscle's pull force acts on it:

$$m\frac{\boldsymbol{d^2r_m}}{dt^2} = \boldsymbol{F}, \quad (12)$$

Given the ratio of 6 and the initial conditions we obtain the Cauchy problem for the differential equation:

$$\begin{cases} \frac{d^2x(t)}{dt^2} + \frac{3p\pi y_0^2}{mb^2(1-(\frac{x_0}{b})^2)}x(t)^2 = \frac{p\pi y_0^2}{m(1-(\frac{x_0}{b})^2)}, \\ x(0) = x_0, \\ \dot{x}(0) = v_0. \end{cases} \quad (13)$$

v_0 — the initial speed of the free end (for a simpler study of the solution, while we consider $v_0 = 0$).

The general solution of this ordinary inhomogeneous nonlinear differential equation of the second order can be written in the following form, using the Weierstrass function:

$$x(t)=\frac{\sqrt[6]{3}\left(\frac{\sqrt[3]{\frac{y_0^2p}{2m(x_0^2-b^2)}}\sqrt[3]{\pi}(t+c_1)}{\sqrt[6]{3}};2\sqrt[3]{2}\sqrt[6]{3}b^2\left(\frac{y_0^2p}{m(x_0^2-b^2)}\right)^{2/3}\pi^{2/3},c_2\right)}{\sqrt[3]{\pi}\sqrt[3]{\frac{y_0^2p}{2m(x_0^2-b^2)}}}$$

But in the future, it is planned to receive and study solutions by numerical methods in software packages. Also, later we will check how close the solutions of the kinematic and dynamic equations are if we take the trial mass in the second.

In the case when an additional force acts on a muscle, the equation will not fundamentally change, except that additional non-uniformity will be added on the right, and equation 13 will take the following form:

$$\frac{d^2x(t)}{dt^2}+\frac{3p\pi y_0^2}{mb^2(1-(\frac{x_0}{b})^2)}x(t)^2=\frac{p\pi y_0^2}{m(1-(\frac{x_0}{b})^2)}+R, \tag{14}$$

where R — projection of additional force on the axis Ox.

Graphic representation in Wolfram Mathematica and analysis of solutions of equations

In this chapter, we will consider numerical solutions of the above equations, on the basis of which it is expected to obtain some sort of muscle selection algorithm with specific characteristics for a specific task. In the future we plan to check the results of numerical solutions for compliance with the real experiment.

The graph of the solution of the kinematic equation will be as follows:

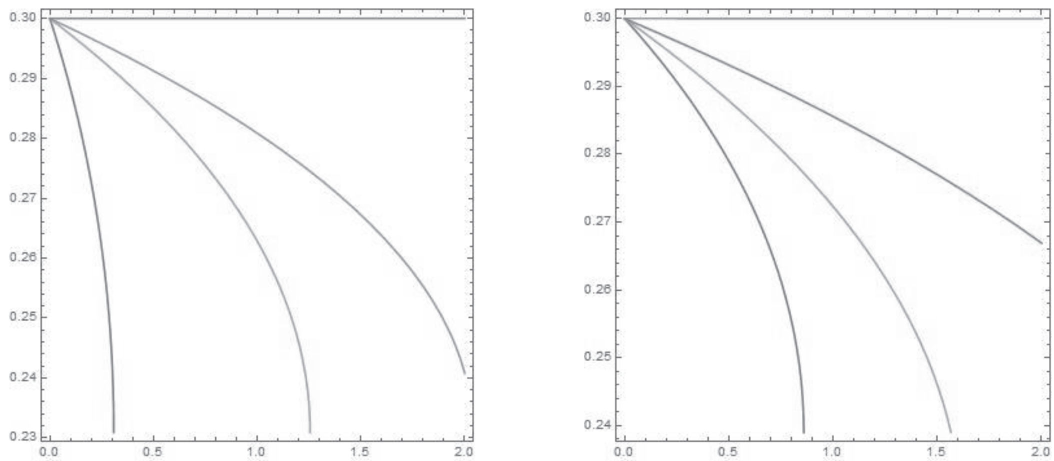


Figure 4 – A diagram of the kinematic equation solution for various values of parameters

For clarity, the graphs of solutions in the first 2 seconds are given for a specific pneumatic muscle: the initial length is $x_0 = 0.3\ m$; the initial diameter is $y_0 = 0.01\ m$. The left figure shows several curves for different values of compressor flow - Q, and in the right figure the value b of the winding length varies at a constant flow rate.

According to these solutions, it can be concluded that trivial solutions exist (at zero flow, as with a winding length equal to the initial muscle length - the length is constant). In addition, the graphs show that the rate of reduction increases with increasing consumption.

If the task does not require large cuts from the muscle, then the sheath can be wound at a large angle. And when winding at the lowest possible angle - maximum cuts are achieved. Based on this, the engineering task of creating such a braid that would be wound at the minimum angle, but would not contradict the model and the principle of muscle operation, is put.

Next, we present a graphical solution to equation 14:

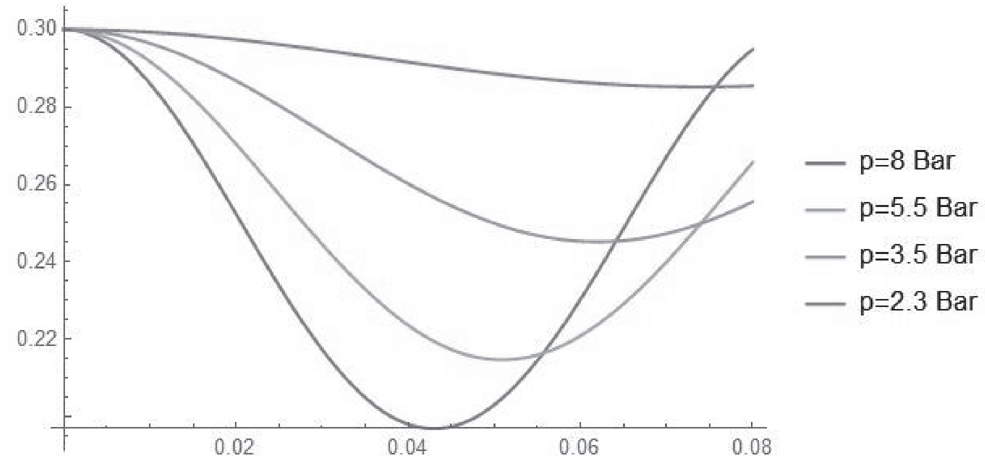


Figure 5 – A diagram of the differential equation solution for various values of pressure

This solution is given at a constant value of the external tensile force $R = 100\ N(10\ kgF)$ acting on the mass point $m = 1\ kg$ fixed at the free end of the muscle with the same parameters (initial length — $x_0 = 0.3\ m$; initial diameter — $y_0 = 0.01\ m$). The curves are given for four different pressure values: 2.3 *Bar*, 3.5 *Bar* 5.5 *Bar* and 8 *Bar* (from top to bottom, respectively).

Due to the fact that the elastic force changes sign when the critical angle is reached (generally speaking, it is predicted that the muscle behaves like an elastic spring), each curve is considered only up to its inflection point. Further movement does not fall under the case considered at the moment. In the case of muscle work, when the system reaches an equilibrium position, the system shuts off the compressor and thus the necessary contraction is achieved.

The expected result is observed from the analysis of the solution of the dynamic equation: with greater pressure, a larger reduction occurs with greater speed, and also with the balancing of forces, a stationary solution will be performed.

The following muscle selection algorithm was developed by processing the obtained solutions using the parameter variation method and proceeding from the conditions of a specific task:

1. It is necessary to determine the dimensions of the muscles allowed by the task, as well as the maximum and minimum working contraction;
2. The corresponding parameters are selected from the solution of the kinematic equation: length, diameter, braid;
3. Knowing the workload applied to the muscle, it is necessary to choose a compressor (it must inject the necessary pressure). In accordance with the applied load, the geometrical parameters are coordinated (since in this work we do not consider the complex problems of the theory of elasticity and want to avoid destruction and wear of the muscle) it is necessary to select parameters with a reasonable margin.

The solution of the model problem

Consider the following task to show an example of the use of a pneumatic muscle and the selection of its parameters, as well as compressor parameters (for ease of calculation, we will perform intermediate calculations, with an accuracy of 1 millimeter). Consider a two-link consisting of two rods of the same length. The upper link is rigidly fixed vertically, and the lower one can freely rotate around it (the mass of the lower link is 5 Kg). One of the ends of the pneumatic muscle is fixed to the lower link in its center of mass, and the other end is rigidly fixed on the fixed rod. The challenge is to raise the bottom bar from position $\phi = 150^0$ to position $\phi = 120^0$. It is easy to express the length of the muscle through an angle ϕ: $x^2 = 0.2 + 0.16cos\phi$, and also to determine that the initial length of the muscle should be 0.582 m, and the final one 0.529 m. The initial radius is 0.01 m (this choice is due to small loads and small dimensions compared to those offered in finished products). The projection of gravity on the muscle axis is:

$$mg_x = mg\frac{0.175 + 0.2cos\phi}{\sqrt{0.2 + 0.16cos\phi}}$$

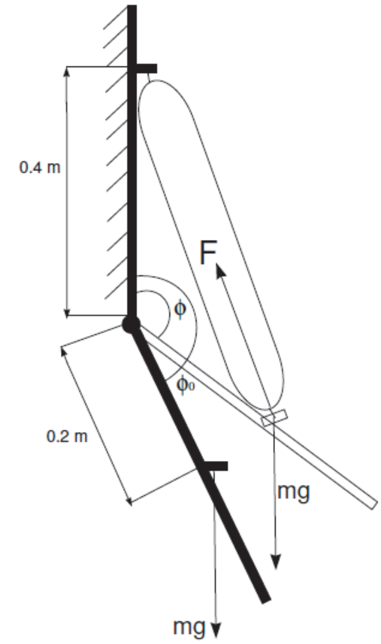


Figure 6 – Figure to model problem

Consider a muscle, with an initial braid angle of 45^0 ($b = 0.823\ m$). From system 11, which describes the contraction of a free muscle, we get that this angle allows for deformations equal to 20%, and in the problem, the deformation is 9%. The expression for the traction force, through the angle ϕ will take the following form:

$$F = p(-0.00067 + 0.00442cos\phi)$$

From where we can get the expression for the pressure required to balance the rod:

$$p = \frac{8.584+9.81}{(0.00067-0.00442cos\phi)\sqrt{0.2+0.16cos\phi}}$$

This expression is necessary to control the process of reduction and create a tracking or stabilizing system that provides the necessary reduction, as well as eliminating the impact (achieving the necessary reduction with high speed). From the expression obtained, we find the minimum operating pressure of the compressor so that the muscle begins to contract — $9282\ Pa$. And the pressure necessary for the muscle to keep the rod in balance at the required angle is equal to — $16653\ Pa$. Based on the obtained values, you can choose a compressor with the necessary parameters.

Conclusion

This paper presents a study of the mechanical drive - an artificial pneumatic muscle. According to the results of the review of existing devices and their applicability, it can be concluded that the use of pneumatic muscles can greatly facilitate the production process when replacing rigid structures with flexible ones, and they can also be used in cases when it is impossible to use other types of drives.

In the conducted research a rigorous mathematical model was constructed, and kinematic and dynamic equations were obtained, which was the main goal of the work. This was not done in the review articles reviewed. On the basis of the obtained solutions, a method for determining muscle parameters for solving specific problems has been proposed. It can be argued that the main tasks considered in the work were performed.

In addition to the fact that it has already been noted, it is worth mentioning that the achieved result does not fully satisfy the initial expectations. It was not possible to obtain a simpler explicit analytical solution for the equations of kinematics and dynamics than those that were written out in the paper. In this regard, there may be difficulties with traffic control or the implementation of certain laws of motion and trajectories specified analytically.

Nevertheless, the main task was carried out, making it possible to explore complex mechanical systems consisting of several pneumatic muscles that perform complex movements.

In the future we plan to continue studying the work of pneumatic artificial muscle. Within the framework of this scientific work it was not possible to conduct a properly experimental part. To do this, it is necessary to assemble an experimental setup and check the experimental results for compliance with the mathematical model constructed in this paper and the derived equations. After conducting the necessary tests and obtaining an objective assessment of the effectiveness of the use of this type of drive, it will be possible to investigate more complex mechanical systems, as well as issues of managing systems containing pneumatic muscles.

References

1. Chou C.-P., Hannaford B., "Measurement and modeling of McKibben pneumatic artificial muscles" , IEEE Transactions on Robotics and Automation, vol. 12, n. 1, february 1996.
2. Hertz E.I., Kreinin G.V., "Design of pneumatic mechanisms" , 1975.
3. Frank Daerden, Dirk Lefeber, "Pneumatic Artificial Muscles: actuators for robotics and automation" , 2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics.
4. Terenziano Raparelli, Francesco Durante and Pierluigi Beomonte Zobel, "Numerical Modeling and Experimental Validation of a Pneumatic Muscle Actuator" , 2016.
5. Festo, "PneumoMuscle MAS — Specification" , 2019.
6. H.F. Shulte, “The characteristics of the McKibben artificial muscle” , Proc. Application of external power in prosthetics and orthotics, National Academy of Science-National Research Council, Washington D.C., 1961.
7. Jackson Wirekoh, Yong-Lae Park, "Design of flat pneumatic artificial muscles" , 2017.

I. Kagirov[1], A. Karpov[1], I. Kipyatkova[1], K. Klyuzhev[2], A. Kudryavcev[2], I. Kudryavcev[2], D. Ryumin[1]

DESIGN OF AN INTELLECTUAL INTERFACE IN THE CONTEXT OF THE MEDICAL EXOSKELETON CONTROL TASK

[1] St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences SPIIRAS, St. Petersburg, Russia
[2] Volga State University of Technology (Volga Tech), Yoshkar-Ola, Russia
karpov@iias.spb.su

Abstract
The paper presents an intelligent human-machine interface designed for controlling the medical robotic lower limb exoskeleton (MRLLE) “Remotion”. The developed intelligent bi-modal interface combines tools of voice control and sensor-based control. The use of intelligent ways of user-exoskeleton interaction increases the level of ergonomics of the product and promotes its use in medical rehabilitation practice due to the intuitive and natural way of human-machine interaction and control.
A somewhat more detailed description of the exoskeleton can be found in the paper: Kagirov I.A., Karpov A.A., Kipyatkova I.S., Klyuzhev K.S., Kudryavcev A.I., Kudryavcev I.A., Ryumin D.A. Intelligent interface for controlling the robotic medical lower limb exoskeleton “Remotion” // Aviakosmicheskaya i ekologicheskaya medicina. №5. 2019 (in print).
Keywords: medical exoskeleton, lower limbs exoskeleton, intelligent control, voice interface, bi-modal interface, assistive technologies.

Acknowledgments
This work was performed within the framework of the complex project designed to develop a high-tech production “Creating a high-tech production of multifunctional robotic exoskeleton for medical purposes (MRLLE)” No. 2017-218-09-1807, and as a part of state research No. 0073-2019-0005.

This article describes an intelligent human-machine interface for managing a medical robotic lower limb exoskeleton “Remotion” [1], which was developed in 2017-2019 as part of a joint project. The project participants were Volzhsky Electromechanical Plant, Volga State University of Technology, and St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS).

The term ‘intellectual human-machine interface’ implies methods of interaction or control of a computer, system or robot, involving the use of artificial intelligence technologies [2, 3]. The intelligent user interface, on the one hand, allows increasing the level of autonomy of the control object, on the other hand, it contributes to naturalness and improved ergonomics of human-machine interaction, allowing the user to use convenient and / or natural ways of interacting with a machine for certain situations. If the first aspect, namely, increasing the level of autonomy of the system, is important for complex robotic systems that operate in conditions of informational ambiguity of the environment, then the second aspect - improving ergonomics and naturalness of the interface - is important in cases where the robot system is designed to interact with humans (social robotics) or human interaction is a necessary condition for its functioning (medical robots, helper robots, exoskeletons, etc.) [4, 5].

One of the components of the intellectual control system is a natural-language (in the particular case of speech) or a multimodal interface of human-machine interaction. This article is exactly dedicated to the management of the medical exoskeleton.

The article is structured as follows: after the present Introduction, in which the meaning of the term “intellectual control” was revealed, brief information on exoskeleton and their classification are given. The following section provides the technical specifications of the developed exoskeleton. The following describes the concept of intelligent control, implemented for the interface of this device, and reveals its main components: touch and voice control. In the Conclusion, the main features of the developed interface are presented and the advantages of its use are formulated.

Overview of Robotic Exoskeletons

The term “exoskeleton” refers to devices designed to increase the physical characteristics of a person with the use of an external frame. Exoskeletons are not autonomous; however, they are usually classified into the service robot’s domain [6, 7].

Despite the prevalence of exoskeletons in various branches of human activity [8], no widely accepted classification of exoskeletons has been elaborated. Among the most common aspects for grouping of exoskeletons can be noted [9-11]:

a) Energy source: active and passive - powered exoskeletons use external power banks, or use cable connections to run sensors and actuators, while passive exoskeletons depend on the kinetic energy of the user;

b) localization: upper / lower limb exoskeletons, suit exoskeletons;

c) scope: military, medical, industrial, space, etc;

d) weight of the structure: from ultralight (up to 2 kg) to heavy (over 30 kg);

e) user mobility: mobile and stationary.

Nowadays, exoskeletons are being actively developed in many countries of the world [12, 13, 14]. Examples of lower limb exoskeletons include ReWalk (by ARGO Medical Technologies, Israel) [15] and eLEGS (Berkeley Bionics, USA: https://bleex.me.berkeley.edu/research/exoskeleton/). Lots of exoskeletons have been created for military purposes. The well-known examples are: the HULC exoskeleton by Lockheed Martin (US) [16] and the XOS exoskeleton (Sarcos, USA: http://www.army-technology.com/projects/raytheon-xos-2-exoskeleton-us/). A detailed review of trends in the development of exoskeletons is given in [17]. This work is especially notable due the fact that it classifies current systems by assistive strategies of application.

The paper [18] provides a list of research priorities in the exoskeletons industry. These include:

− studies on the kinematic and biomechanical properties of current devices in order to formulate of the principles of their application;

− development of new methods to determine parameters of exoskeletons and control their functioning, which would provide researchers with tools for quick and systematic evaluation of exoskeletons of various types according to selected criteria;

− computer analysis of virtual topographic-anatomical environments when designing biomechanical systems;

− creation and improvement of materials and basic parts of exoskeleton, ensuring their effective work.

Medical Robotic Exoskeleton

The developed multifunctional medical robotic lower limb exoskeleton “Remotion” (hereinafter referred to as MRLLE) can be used to mitigate or liquidate the effects of the central and peripheral nervous system damage, consequences of injuries and diseases of the musculoskeletal system, accompanied by dysfunction of the lower extremities.

Currently there exists four modifications of the MRLLE:

− REM-B – the basic version with basic sensors and control system, designed for rehabilitation in medical institutions.

− REM-E – similar to REM-B, with the addition of electromyography (EMG) channels, designed for rehabilitation in medical institutions.

− REM-F - an enhanced version of REM-E with a system of functional electrical stimulation of the patient’s muscles (FES), designed for rehabilitation in medical institutions.

− REM-D – the basic version with basic sensors and control system, designed for rehabilitation of children in medical institutions.

MRLLE is mounted of an external metal frame equipped with four servos and fixed on the patient's body via soft cuffs, belts and fasteners. One of the features of the exoskeleton RME is the modular design, and options to program various walking patterns, as well as the availability of electromyography and functional electrical muscle stimulation. The parts of the MRLLE are: a lumbar module, electrically driven hip modules, electrically driven lower-leg modules, foot modules, handles, a fixation system, and a power supplying system (battery), a control panel, and a charger. In addition, the exoskeleton is equipped with crutches or walkers. The overall 3D design of the exoskeleton is set out in Fig. 1, and in Figures 2A and 2B is a diagram of the equipped exoskeleton and a photo of the product.

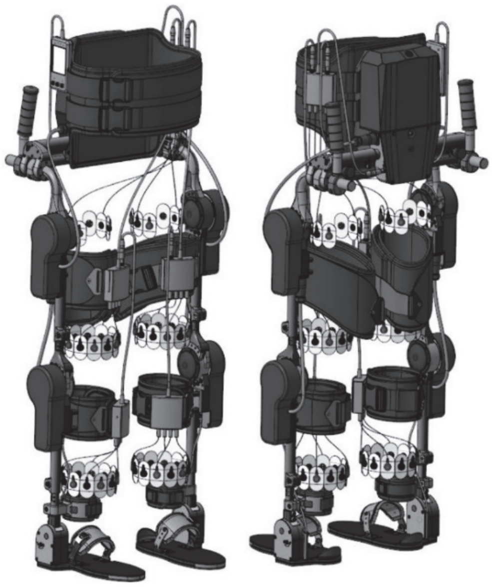


Figure 1 –3D design of the exoskeleton (frontal and rear views)

The lumbar module, modules of the hips and legs can be adjusted in width, depth and length depending on size of the patient. For the same purpose, the foot module is equipped with an adjustable hinge, and the system of belts, cuffs and buckles is adjustable.

The servos from the hip and lower leg module are motor reducers. The exoskeleton control system software allows various motion programs depending on medical purposes.

Figure 2 – MRLLE general view: A) a sketch of the equipped exoskeleton
(1 – plastic protective case; 2 – waist belt; 3 – user; 4 – walkers; 5 – confirmation button);
Б) a photo of the equipped MRLLE

MRLLE allows rehabilitation of the user in the following modes of operation:

– get up / sit down;
– go / stop;
– turning;
– mark time;
– asymmetric user goniograms;
– diagnostic mode "synchronous electromyography"
– mode "synchronous functional electrostimulation"

In order to improve the effectiveness of rehabilitation sessions, MRLLE is equipped with FES and EMG systems. This uses one of the modes of operation: reading EMG signals from certain muscles of the patient, or muscle stimulation at a point in time synchronized with step periods.

The activation and control in FES and EMG modes is carried out remotely from a PC via the RemotionTool software utility. The connection between a PC and the exoskeleton is set up automatically via a wireless Wi-Fi network. RemotionTool graphical user interface allows the medic to maintain several medical files, to create and update databases of rehabilitation techniques, to analyze the data from each goniography session.

Exoskeleton Control System

The control of the exoskeleton is possible using touch or voice interfaces, which are implemented in a bimodal control system. Touch control is carried out using an application with a graphical user interface (GUI) that is installed in the smartphone. Voice control is aimed at automatic recognition of commands that the user pronounces in Russian. Thus, the control system of the exoskeleton is aimed at helping the patient during the processes of activation, interaction and deactivation of the exoskeleton. The developed devices and software are designed in such a way that they allow you to export a profile with the necessary rehabilitation program from a personal computer to a mobile device, which leads to time savings. During the initial phase of the rehabilitation course, patient users rarely can perform a certain movement on their own. In this case, automatic mechanisms for controlling the exoskeleton are the only ways to solve the described problem. Figure 3 shows some examples of interfaces that have graphical shells, due to which patient users can easily perform the necessary operations on the exoskeleton.

Movement pattern choice	Fine-tuning characteristics of moves	Telemetric data

Figure 3 – Graphical interfaces as displayed on the touch screen of the control panel

Voice interface for contactless control of exoskeleton

The efficiency and naturalness of exoskeleton control is significantly enhanced by the voice interface and well-thought-out individual commands. Intellectual voice input contains a small number of narrowly directed commands that are available for performing, which leads to a significant improvement not only in speed, but also in ease of interaction (a certain command is more often easier to pronounce than to press a touch button in the graphical user interface). It was also revealed that during the course of rehabilitation courses, patients are distracted by a mobile device for sensory interaction, which negatively affects the overall concentration. Entering an additional modality in the form of voice control improves ergonomics. In the future, this technology will allow the use of exoskeletons not only under the careful supervision of doctors in medical institutions, but also in everyday life.

Currently, a number of domestic and foreign developments are aimed at implementing voice controls for intelligent methods of exoskeleton control. For example, the ExoAtlet exoskeleton from the Russian company OOO Exoatlet (https://www.exoatlet.com/ru/node/84/) and the exoskeleton of the lower limbs ARKE (https://exoslelerereport.com/product/arke/) from Canadian company Bionic Laboratories Corp. The last example is particularly interesting in that for speech recognition the creators used Alexa automatic speech recognition (APP) technology from the American company Amazon, integrating it into the exoskeleton control system. A similar approach was chosen by the authors when creating their own voice control system for the exoskeleton.

For automatic speech recognition in the exoskeleton control system, features are used embedded in the open source software interface of the Android operating system developer from Google, which allows you to convert the audio signal to a textual representation on mobile devices running Android (https://developer.android.com/reference/ android / speech / SpeechRecognizer).

The voice interface module software was developed using a high-level Java programming language and the professional development environment Android Studio. The logical structure of voice interface modules for intelligent control of SEM is shown in Figure 4.The main functions of the developed contactless voice interface to control the exoskeleton MRLLE are:

1) automatic conversion of isolated Russian voice commands into a textual representation;

2) voice feedback by repeating the recognized voice command via the headphones or the smartphone (female or male voice at the user's choice);

3) sending the code of the recognized speech command to the electronic control system of the MRLLE before the further actuation;

4) duplication of the progress of the voice control process using the graphic interface of the smartphone.

The complete list of voice commands currently available to the user is presented in the Table below. Before each voice command, the user must either press a virtual button on the screen, which turns on voice input mode, or say an activation keyword, which reduces the probability of false alarms of the voice control module. The user can choose one of the three keywords of his/her choice: (1) “Robot”, (2) “Command”, and (3) “Execute” (‘Robot’, ‘Komanda’, and ‘Vypolni’ in Russian).

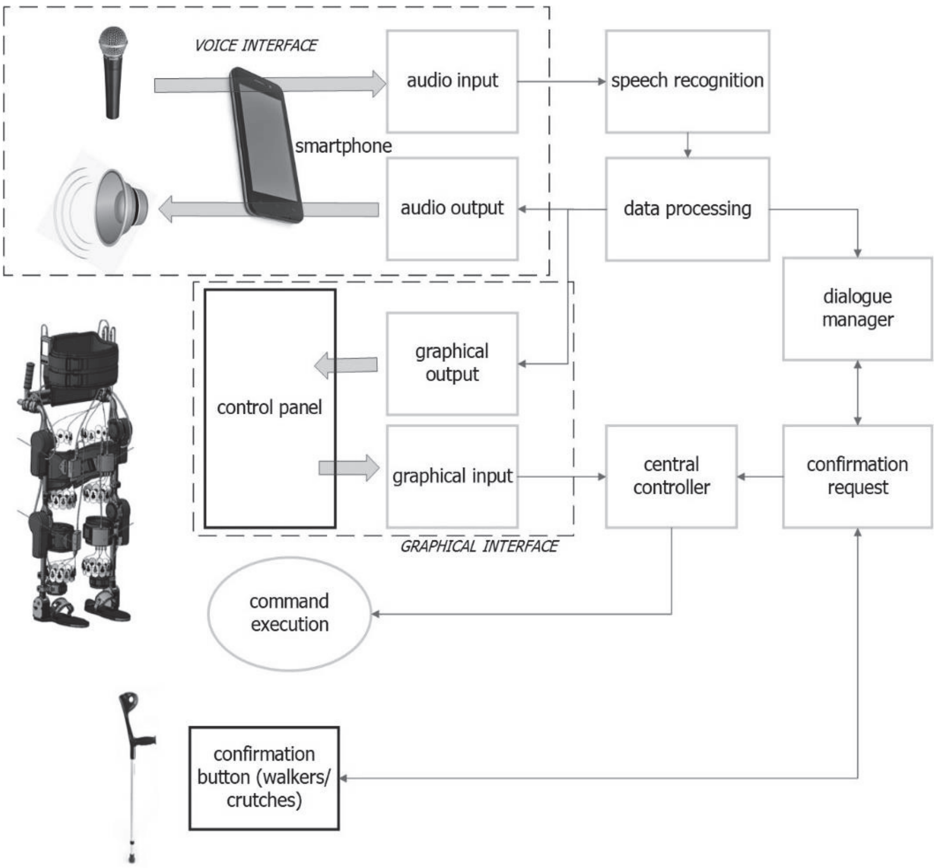
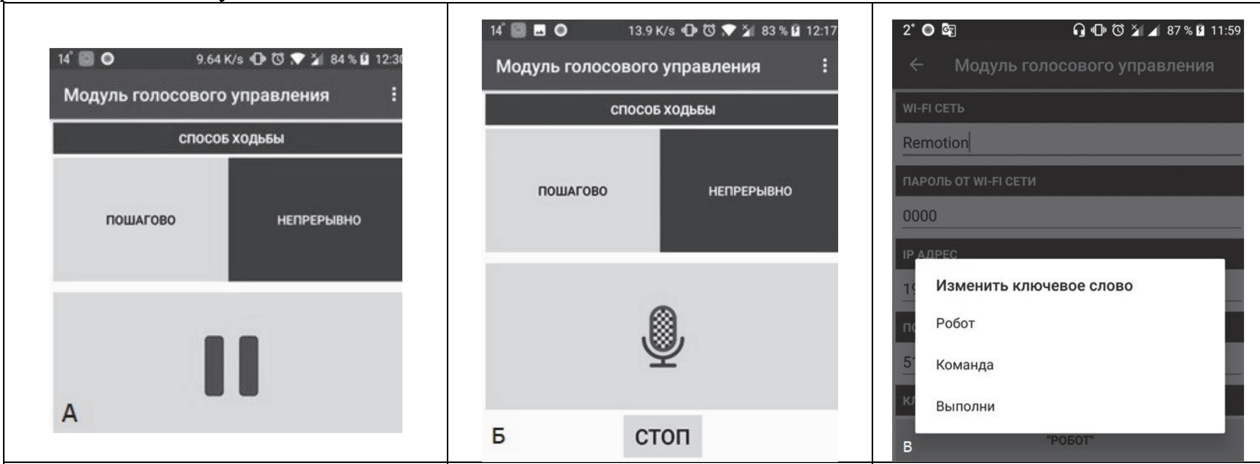


Figure 4 – The logical structure of the MRLLE voice interface

Table 1. Voice command list and the corresponding actuations by the exoskeleton.

Voice command (in Russian)	**Actuation performed by MRLLE**
‘Take a step’	One step forward
‘Go’	Continuous steps forward
‘Stop’	Keeping a standing position
‘Get up’	Getting up from a sitting position
‘Sit down’	Take a seated position
‘Turn left’	Turning left 90 degrees
‘Turn right’	Turning right 90 degrees
‘Abort’	Cessation of the last active actuation

In order to reduce power consumption, the audio stream captured through the smartphone’s microphone is checked for any speech-like audio signal, and, in case of successful detection, the mode of command detection is activated. If the activation command matches the preferred keyword, the speech signal following the activation command is considered as the control command. The software delivers the code of the recognized command to the IP address specified in the settings in accordance with the developed protocol. The recognized command is then displayed in the graphic interface of the system is doubled by the voice via smartphone speakers. Figure 5 shows the graphical interface of the voice control system: a dialog box during the capture of a speech signal, a window with a recognized voice command (“Stop”) and a window with the option to select a keyword.



Audio signal capture for voice command.	A recognized speech command.	Key word choice.

Figure 5 – the graphical interface of the voice control system: dialog boxes

If the input speech command is not recognized by the system correctly, the user will be requested to repeat the attempt. In order to increase safety level and foolproofness, actuations of the exoskeleton are not executed immediately after the recognition process, but by active user confirmation: the execution of the action takes place after a certain manipulation by the user (pressing a button on the crutch).

Conclusion

The article describes the intelligent human-machine interface for controlling a robotic exoskeleton for medical use by the Exoskeleton “Remotion” and provides specific technical details of the exoskeleton architecture. Based on the results of these developments and studies, the following conclusions can be drawn:

1) the created robotic medical exoskeleton Remotion differs from a number of similar devices by the complete modularity of the device, which greatly facilitates the maintenance of the product and leaves room for flexible adjustment of the exoskeleton depending on the specific conditions of use;

2) despite conceptually similar devices in the domestic market, the Exoskeleton Remotion differs from them, first of all, by the presence of an intelligent user interface, which significantly improves the ergonomics of the exoskeleton and the possibility of contactless control;

3) the combination of intelligent control methods through voice commands, touch control, and the presence of voice and visual cues about the state of the exoskeleton allows to increase the security level of the user when using the exoskeleton;

4) the availability of electromyography, functional muscle electrostimulation and the possibility of medical control through the developed program utility RemotionTool improves the efficiency of rehabilitation, medical treatment or mitigating the effects of the user-patient disease;

5) a prototype of the exoskeleton neurocomputer control module (neurointerface) is currently at the development stage, the existing similar developments are relevant in the development of assistive exoskeleton for paralyzed people [19, 20].

References

1. Kapustin A.V., Loskutov Yu.V., Skvortsov D.V., Nasybullin A.R., Klyuzhev K.S., Kudryavtsev A.I. Circuit solutions for the management of a rehabilitation exoskeleton for medical purposes // Vestnik Povolzhskogo gosudarstvennogo tekhnologicheskogo universiteta. 2018. № 2 (38). P. 77–86.
2. An Introduction to Intelligent and Autonomous Control. Antsaklis P.J., Passino K.M. (eds.). Kluwer Academic Publishers, 1993.
3. Shcherbatov I.A. Intellectual control of robotic systems in conditions of uncertainty // Vestnik Astrakhanskogo gosudarstvennogo tekhnicheskogo universiteta. 2010. № 1. P. 73–77.

4. Karpov A.A., Yusupov R.M. Multimodal Interfaces of Human-Computer Interaction // Herald of the Russian Academy of Sciences. 2018. V. 88. № 1. P. 67-74.
5. Ushakov I.B., Karpov A.A., Kryuchkov B.I., Polyakov A.V., Usov V.M. Promising solutions in the field of medical robotics to support crew life and reduce medical risks in space flight // Aviakosmicheskaya i ekologicheskaya meditsina. 2015. V. 49. № 6. P. 76–83.
6. World Robotics – Service robots 2017: Statistics, Market Analysis, Forecasts and Case Studies. Frankfurt-am-Main: VDMA Verlag, 2017.
7. Ermolov I.L., Knyaz'kov M.M., Kryukova A.A., Sukhanov A.N., Kryuchkov B.I., Usov V.M. Method of Controlling an Exoskeleton Device Using the System of Recognition of Arm Movements on basis of Biosignals from the Skeletal Muscles of a Human Operator's Arms. // Pilotiruemye polety v kosmos. 2015. № 4 (17). P. 80–93.
8. Ferris D. The exoskeletons are here // Journal of neuroengineering and rehabilitation. 2009. V. 6 (17).
9. Vorob'ev A.A., Andryushchenko F.A., Ponomareva O.A., Solov'eva I.O., Krivonozhkina P.S. Controversial terminology and classification of exoskeletons. (Analytical review, own data, clarifications, suggestions.) // Volgogradskii nauchno-meditsinskii zhurnal. 2015. № 3 (47). P. 14–20.
10. Herr H. Exoskeletons and orthoses: Classification, design challenges and future directions // Journal of neuroengineering and rehabilitation. 2009. V. 6 (21).
11. Gorgey A.S. Robotic exoskeletons: The current pros and cons // World Journal of Orthopedics. 2018. Vol. 9 (9). P. 112–119.
12. Vorob'ev A.A., Zasypkina O.A., Krivonozhkina P.S., Petrukhin A.V., Pozdnyakov A.M. Exoskeleton - the state of the problem and the prospects for the introduction of the system of habilitation and rehabilitation of persons with disabilities (analytical review) // Vestnik Volgogradskogo gosudarstvennogo meditsinskogo universiteta. 2015. № 2 (54). P. 9–17.
13. Banala S.K., Agrawal S.K., Kim S.H., Scholz J.P. Novel gait adaptation and neuromotor training results using an active leg exoskeleton // IEEE/ASME Transactions on Mechatronics. 2010. V. 15 (2). P. 216–225.
14. Banala S.K., Kim S.H, Agrawal S.K., Scholz J.P. Robot assisted gait training with Active Leg Exoskeleton (ALEX) // IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2009. V. 17 (1). p. 2-8
15. Talaty M., Esquenazi A., Briceno J.E. Differentiating ability in users of the ReWalk(TM) powered exoskeleton: an analysis of walking kinematics // IEEE Int Conf Rehabil Robot 2013.
16. Bednyak S.G., Eremina O.S. HAL Robotic Exoskeletons (Feel like a HALc) // Sworld. 2014. Vol. 2. № 1. P. 49–51.
17. Yan T., Cempini M., Oddo C.M., Vitiello N. Review of assistive strategies in powered lower-limb orthoses and exoskeletons // Robotics and Autonomous System. 2015. V. 64. P. 120-136.
18. Ergasheva B.I. Lower limb exoskeletons: brief review. Scientific and Technical Journal of Information Technologies, Mechanics and Optics. 2017. V. 17. №. 6. P. 1153–1158.
19. He Y., Eguren D., M Azorín J., Grossman R., Luu, T. Ph., Contreras-Vidal J. Brain–machine interfaces for controlling lower-limb powered robotic systems // Journal of Neural Engineering. 2018. № 15.
20. Rosen M. Mind to Motion: Brain-computer interfaces promise new freedom for the paralyzed and immobile // Science News. 2013. V. 184 (10). P. 22–24.

V. Vlasenko[1], S. Orlova[1], A. Bakhshiev[2]

REVIEW OF MODERN METHODS OF SEGMENTATION OF MEDICAL IMAGES

[1] The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
v.vlasenko@rtc.ru, s.orlova@rtc.ru, alexab@rtc.ru

Abstract

The paper analyzes the main methods of image segmentation. The advantages and disadvantages of the methods in the task of segmentation of organs and cell tissues in medical images are considered. A review of datasets used to study prostate segmentation methods is given. Considered in detail the U-net convolutional network and its modifications. A comparison of their effectiveness with other neural network solutions in various problems of segmentation of organs and tissues is given.

Keywords: deep learning, image segmentation, medical images, U-net.

Acknowledgments

This work was done as the part of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00 “Cloud services for automatic synthesis and validation of datasets for training deep neural networks in pattern recognition tasks”.

1. Introduction

There are a number of tasks for the segmentation of medical images (MRI, CT, ultrasound and X-ray), for example, the allocation of organs, tissues and tumors in the images, often requiring high precision segmentation. There are many images per patient. Processing, including manual segmentation of images, is a time-consuming process. In view of the growing computing capabilities of the technique, it became possible to create algorithms that can significantly accelerate the process of segmentation of medical images and have good quality work. Although such algorithms already exist, they continue to develop actively, raising the level of segmentation quality and expanding the scope, and new methods and experimental results are appearing.

The aim of this work is to analyze methods and systems of technical vision for the segmentation of medical images.

2. Image Segmentation Techniques

There are many methods of image segmentation; most of them can be divided into seven popular groups [1, 2, 3], which are shown in Figure 1.

Threshold methods. Within this method, threshold values are selected, and pixels, depending on their intensity value, are assigned to certain classes. There are three main types of threshold segmentation: global threshold segmentation, variable threshold segmentation, and multiple threshold segmentation.

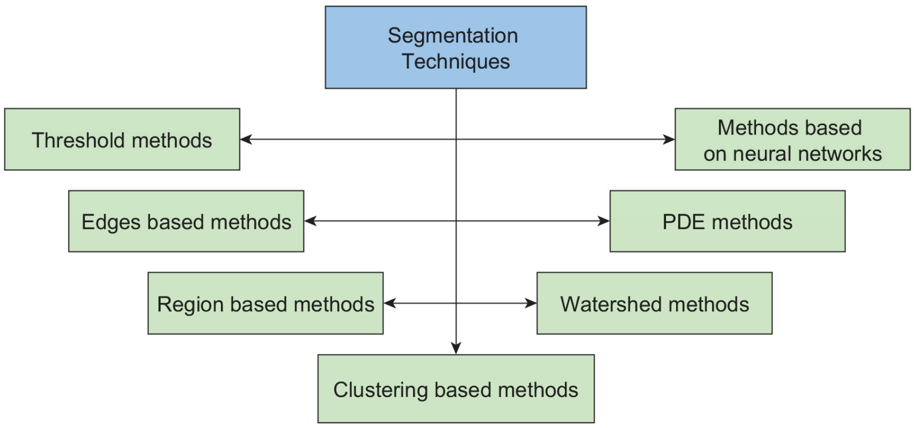


Figure 1 – Image segmentation Techniques

1. Edges based methods. The intensity value alone does not provide sufficient accuracy. This method is based on a sharp change in the intensity value in the image. Edges can be found in places where either the first derivative of the intensity is greater than a certain value, or the second derivative has an intersection with 0. Pixels that are not separated by boundaries are assigned to the same class.

2. Regions based methods. The method of segmentation based on regions is based on the selection of segments that have similar characteristics. There are two main types: region growing methods or region

splitting and merging methods. Region growing methods segment the image into different regions based on the initial pixels (seeds), which can be selected manually (based on existing knowledge) or automatically (depending on the specific application), and regions gradually grow out of them, using a measure of similarity - if the intensity If an unoccupied neighboring pixel matches sufficiently with the average intensity of the region, then the pixel is included in the region. In region splitting and merging methods, the image is iteratively divided into areas with similar characteristics, then similar areas are joined together.

3. Clustering based methods (Fig. 2). This method is based on clustering pixels with similar characteristics. Clustering-based segmentation consists of dividing data elements into clusters in such a way that the elements in one cluster are more similar to each other than to others. There are two main categories of clustering methods: the hierarchical method and the partitioning method. Hierarchical methods are based on the concept of trees. In this case, the root of the tree represents the entire database, and the internal nodes are clusters. On the other hand, the partition-based method uses optimization methods iteratively to minimize the objective function. Between these two methods, there are various algorithms for searching for clusters. There are two main types of clustering:

– Hard clustering is a simple clustering method that divides an image into multiple clusters so that one pixel can belong to only one cluster. These methods use a membership function that takes values 1 or 0 i.e. one specific pixel can either belong to a specific group or not belong. An example of such a clustering method is the K-means method. In this method, the centers are calculated first, and then each pixel refers to the nearest center.

– Soft clustering is a more natural form of clustering, because in real life accurate separation is not possible due to the presence of noise. Thus, such clustering methods are most useful for image segmentation, in which the separation is not strict. An example of this type is fuzzy C-means clustering. In this method, pixels are divided into clusters based on partial membership, i.e. one pixel can belong to more than one cluster. Each pixel is assigned a degree of belonging to clusters. This method is more flexible than other methods.

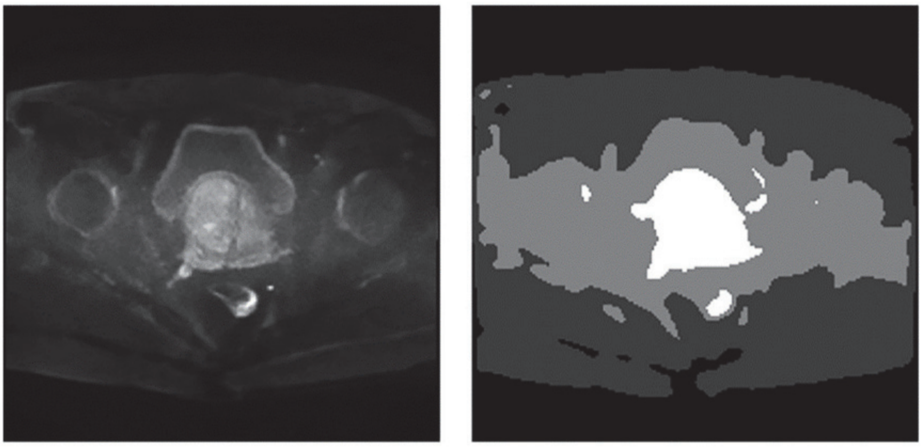


Figure 2 – Clustering based method K-means

4. Watershed methods (Fig. 3). In segmentation by watersheds, the image is represented as a function of z = I(x, y) intensity or the modulus of the gradient of the pixel coordinates. The function is plotted along the OZ axis. Ridges are formed in places of sharp changes in intensity, and valleys are formed in places of uniform intensity. Next, a threshold is selected along the OZ axis to which the image is filled with “water”. The remaining ridges are the contours of the objects, and the spaces separated by them, filled with water, are separate clusters. This algorithm is well suited for images with a small number of local minima.

5. Partial differential equation (PDE) based methods. Partial differential equations methods are fast segmentation methods. They are suitable for real-time applications. There are two main PDE methods: a nonlinear isotropic diffusion filter (used to reinforce the ribs) and restoration of convex non-quadratic variations (used to eliminate noise). The result of the PDE method is blurry edges and borders that can be shifted using closing operators. The fourth-order PDE method is used to reduce image noise, and the second-order PDE method is used to better detect boundaries [4].

6. Methods based on neural networks. These methods use artificial neural networks, usually convolutional, to solve the segmentation problem. Currently widely used in the problems of segmentation of medical images, road images, separation of the target object from the background.

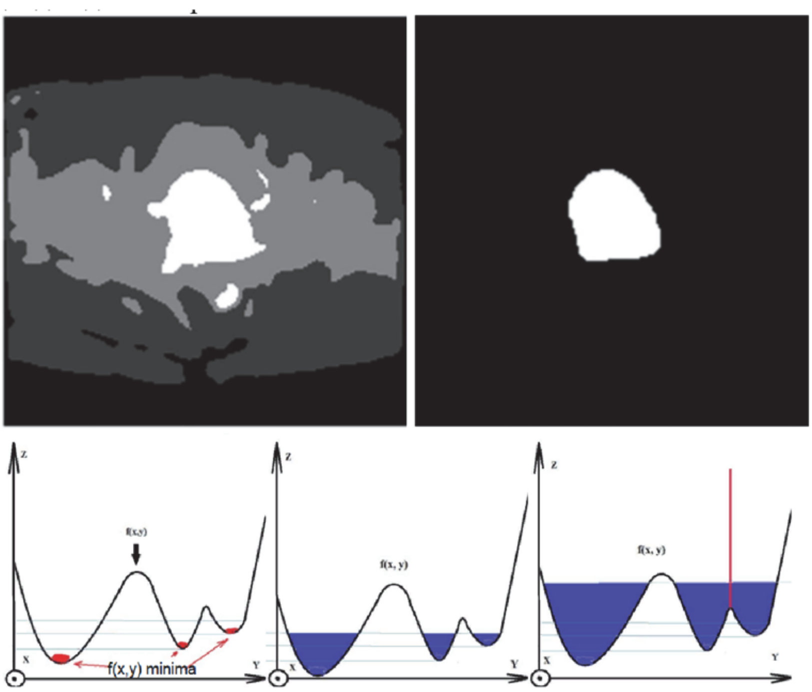


Figure 3 – Watershed method

Also not so popular among general purpose segmentation methods, but quite effective is a method for segmenting biomedical images — the atlas-based method [4, 5, 6]. Atlases are collections of images with manual markings by experts. This method may have the following algorithm [6]. At the first stage, atlases are selected with which the input image will be matched, for which you need to get markup (segmentation mask). At the second stage, the atlases are compared with the input image using special software for the analysis of medical images. At the final stage, to combine the marking of the atlases in such a way as to obtain a new markup corresponding to the input image, the weights are calculated based on the measure of similarity of the input image to the atlases and then merge the markings by weighted voting. The algorithm works quite effectively on large objects, but is not suitable for working with small or thin objects. The operating time, depending on the number of atlases and other parameters, varies from 40 seconds to several minutes.

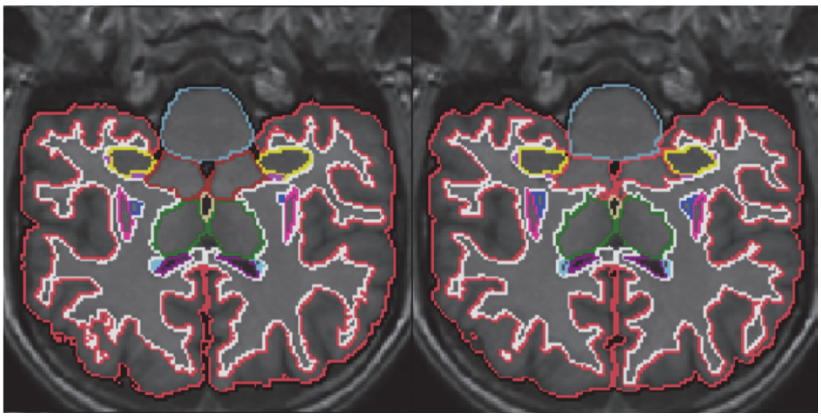


Figure 4 – The atlas-based algorithm [6]: on the left is expert markup, on the right is the result of the algorithm

The resulting comparison of methods is given in Table 1.

The threshold segmentation method is not suitable for solving the problem of high-quality segmentation of medical images, because it has very specific conditions for its work: the object stands out against an almost uniform background. The edge method may fail, because biomedical images often contain a lot of noise and a large number of borders, and sometimes, on the contrary, the borders of the target objects are too blurred. The same applies to watershed segmentation methods.

Methods based on regions, clustering, and based on PDE can be suitable for pre- and post-processing, they can also be tested as independent methods of segmentation and to evaluate their performance compared to other methods.

Among the considered methods, methods based on neural networks make it possible to distinguish and classify objects and areas in images not only on the basis of pixel brightness or some of their individual characteristics, but also on the basis of complex features that are formed in the learning process. This implies a drawback of neural network methods - not very high speed (or high requirements for computing resources) and the need to prepare a representative training base of a large volume, however, this particular group of methods is universal and allows to get the highest quality result in segmentation tasks when it is difficult to identify features segment based on pixel values only.

In recent years, competitions have become widespread, in which teams solve the problem, having a set of annotated data from the organizers, which can be used for training or selecting algorithm parameters. Algorithms are evaluated based on a set of non-annotated (for participants) data, for which teams receive results. Organizers have markup for such a test suite. Several competitions related to image segmentation can be mentioned: “Ultrasound Nerve Segmentation” 2016 [7], “KONICA MINOLTA - Pathological Image Segmentation Challenge” 2017 [8], “Digital Pathology: Segmentation of Nuclei in Images” 2018 [9], “PAIP 2019: Liver Cancer Segmentation” 2019 [10]. In all these competitions, the winning teams used neural network methods.

According to a review of segmentation methods, the most promising general-purpose method should be considered segmentation methods based on convolutional neural networks.

Table 1. Segmentation Method Groups

Method group	Description	Advantages	Disadvantages
Threshold methods	The threshold value is determined by the histogram of the image	do not need a priori information; simple methods in implementation, undemanding to resources	depend on abnormal values; spatial details are not taken into account
Edges based	based on detection of intensity spikes	suitable for pictures that have good contrast between objects	not suitable for cases when there are a lot of edges
Regions based	based on the partition of the image into homogeneous regions	more resistant to noise; useful when it is easy to define similarity criteria	very expensive methods in time and memory
Clustering based	based on division into homogeneous clusters	the use of partial membership is therefore more useful for a real problem	complex definition of membership function
Watershed methods	based on topological interpretation	results are more stable, detected boundaries are continuous	sophisticated gradient calculation
PDE based	based on differential equations	high speed	high computational complexity
Atlas-based	based on image matching with atlases and markup transfer	do not require a lot of data to implement	not a flexible algorithm; does not work well on small objects; low speed
based on neural networks	based on modeling the learning process for decision making	universality of algorithms	large time costs for training; must have training datasets

3. Existing solutions for the segmentation of medical images based on convolutional neural networks

Convolutional networks have been around for a long time. One of these networks, containing 8 layers and millions of customizable parameters, made a breakthrough in the classification of the ImageNet dataset with more than 1 million training images [11]. Since then, a large number of different neural network architectures of even greater complexity have been developed.

A typical use of convolutional networks is a classification task when a class label from a set is matched to an image. However, in most tasks of the analysis of biomedical images, segmentation is required, i.e. match the class label to each pixel in the image. Moreover, the number and volumes of publicly available labeled datasets that can be used to train neural network algorithms are relatively small.

A successful attempt to solve this problem was made at the 2012 ISBI Championship. In the winning solution [12], the neural network processed fragments of the image using a sliding window, predicting the label for each pixel and taking as input some local area around this pixel (Fig. 5). This network made it possible to localize the result and, since the neural network worked directly only with image fragments, and not with the entire images, there were much more training examples than the number of training images.

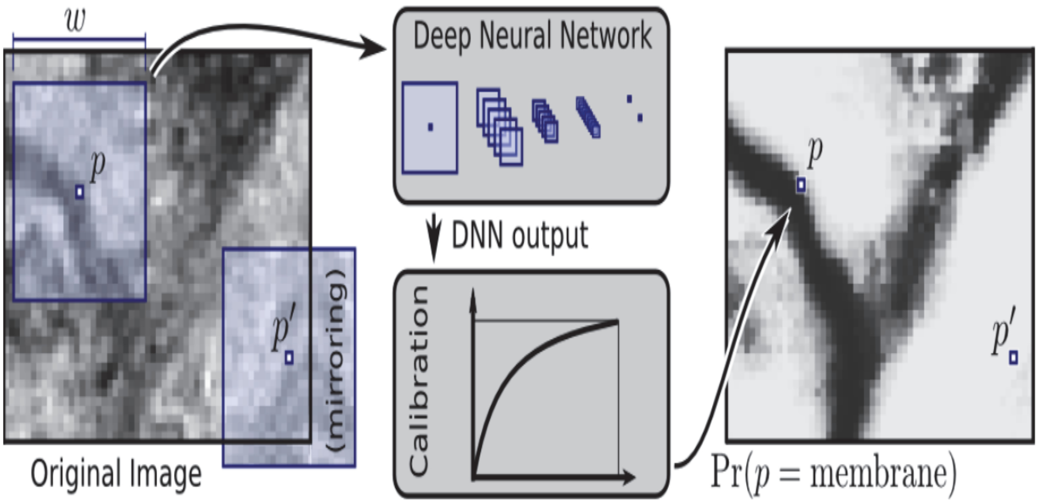


Figure 5 – The principle of the algorithm [12]

But this approach has two drawbacks. Firstly, low speed, because the network must separately process each area of the image, and there is a lot of redundancy due to the overlap of these areas. Secondly, when using this approach, it is necessary to find a balance between the magnitude of the input of the neural network and the use of context, i.e. between localization accuracy and classification accuracy: larger input image areas that work with a large volume of context require more pooling layers, which reduce localization accuracy, while small areas can lead to a decrease in classification accuracy.

2015 can be considered the year of the emergence of modern neural segmentation networks when the neural network architectures FCN [13], U-net [14] and DeepLab [15] were introduced. And while FCN and DeepLab were positioned as general-purpose segmentation methods, the U-net architecture won several competitions in medical image segmentation. U-net uses the principle proposed by the authors of FCN - a full-convolutional architecture (consisting entirely of convolutional layers), which can be divided into two functional blocks: an encoder (part of downsampling) and a decoder (part of upsampling). The encoder is built as a neural network classifier, which receives an input image and receives a class attribute map at the output, and then a decoder follows that converts the attribute map into a segmentation mask. For more efficient use of information, the decoder receives not only the final encoder feature map, but also intermediate feature maps (Figure 5). The architecture looks symmetrical, U-shaped, which determines the name of the neural network.

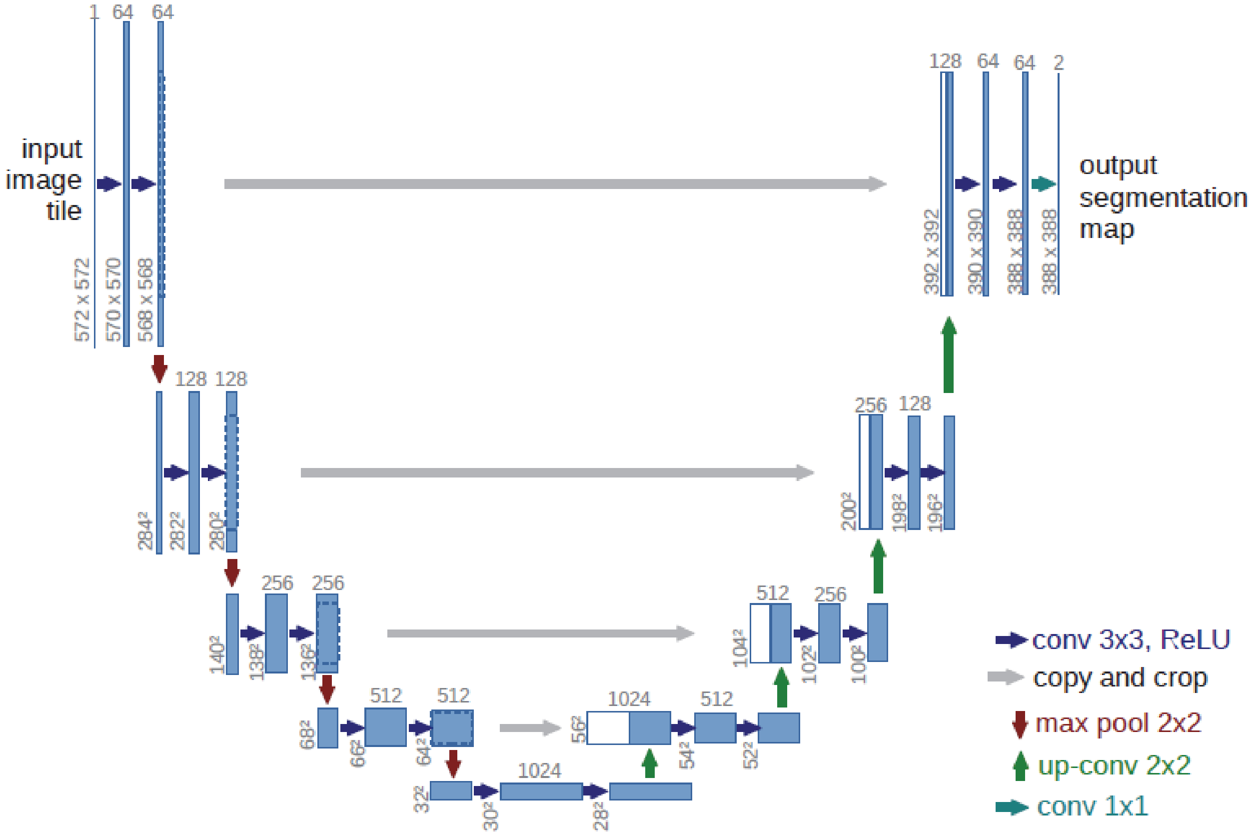


Figure 6 – U-net architecture

A feature of U-net is a large number of channels in the decoder part, which allows you to save more information in the last layers, as well as the use of elastic deformations when augmenting the training base, which is often found in biomedical images. The problem in many tasks of segmentation of medical images of cells is the separation of contiguous objects of the same class (Figure 7). To this end, the authors propose the use of a weighted loss function, in which the separation of background labels between objects in contact has a greater weight. Although the authors suggested the use of U-net for segmenting biomedical images, this neural network has become very popular and still shows good results in various tasks and competitions. Moreover, often the original version of the encoder is replaced with another classifier without the last layers - VGG-16, ResNet, etc.

U-net surpassed the neural network [12] when evaluating on the basis of the ISBI 2012 competition without any preliminary or further post-processing (Table 2). The training data for this competition is a set of 30 images 512x512 pixels in size. For each image there is a mask on which cells are highlighted in white and the membrane is black. Nevertheless, according to the table, it can be noted that manual labelling at that time exceeded the result of the best algorithm by accuracy tenfold.

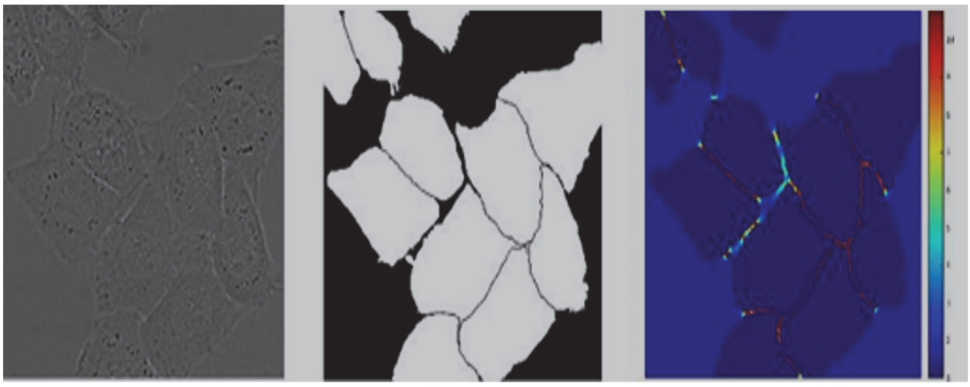


Figure 7 – HeLa cells on glass are fixed by differential interference contrast.
On the left is the original image, in the middle is the mask obtained by segmentation (white - cells, black - background), on the right - a map with pixel weights for the loss function

Table 2. ISBI 2012 Leaderboard, sorted by warping error (March 6th, 2015)

Name of algorithm or team	Warping Error	Rand Error	Pixel Error
Hand labelling	0.000005	0.0021	0.0010
U-net	0.000353	0.0382	0.0611
DIVE-SCI	0.000355	0.0305	0.0584
IDSIA	0.000420	0.0504	0.0613
DIVE	0.000430	0.0545	0.0582

U-net architecture managed to win the ISBI 2015 competition (Table 3) on the PhC-U373 and DIC-HeLa datasets (Figure 8). It took the authors 10 hours to train U-net on an NVidia Titan GPU (6 GB), segmenting a single 512x512 pixel image took less than one second.

Table 3. Accuracy Algorithm Results at ISBI 2015

Name	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
HOUS-US (2014)	0.5323	-
second-best (2015)	0.83	0.46
U-net (2015)	0.9203	0.7756

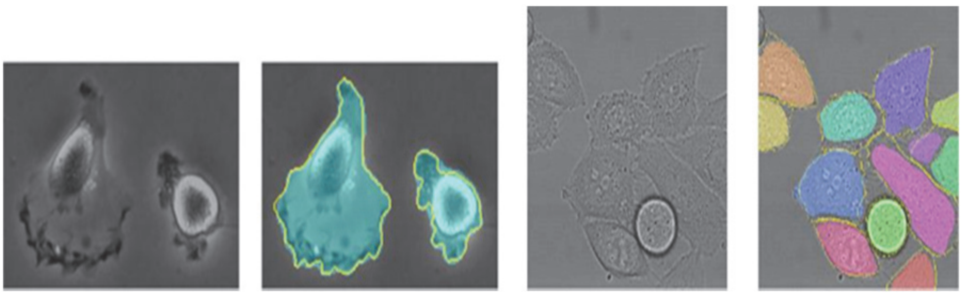


Figure 8 – Examples of images and annotations of datasets used at the ISBI 2015 competition. On the left is the “PhC-U373” dataset, on the right is “DIC-HeLa”

In [16], the authors identified three problems of the classical U-net architecture. First, U-net, as a popular model for segmenting medical images, is difficult to train when convolutional layers increase, but a deeper network usually has better generalization due to more parameters. Secondly, exponential ReLU (ELU), as an alternative to ReLU, is not much different from ReLU when the network of interest is deepened. Thirdly, the Dice coefficient, as one of the widespread loss functions for segmenting a medical image, is ineffective when the forecast is close to true and will cause fluctuations during training. To solve the three above problems, the authors propose a deeper network that can be trained on medical image data sets that are usually small. They also offer a new loss function to speed up the learning process and a combination of various activation functions to increase network performance.

In the article, the authors propose a bridge architecture between two U-nets (Figure 9). They connect each decoder block of the first U-net with the corresponding encoder block of the second U-net, which directly introduces features from previous levels to the last levels. This process should reduce the cost of training and show better performance than one U-net.

As an improvement to U-net, stacked U-net is used. Such a network improves network performance by using the first U-net to find rough signs and using the second U-net to get an accurate result. However, this increases the complexity of the network. Unlike the previous U-net with accumulation, which requires a large amount of training data, connecting two U-net can reduce the cost of training and make the network suitable for medical use, where training data is usually not enough. Table 4 shows the effect of various connection methods on quality.

The authors also presented the results of comparison with other teams in the leaderboard of the MICCAI PROMISE12 competition (Table 5).

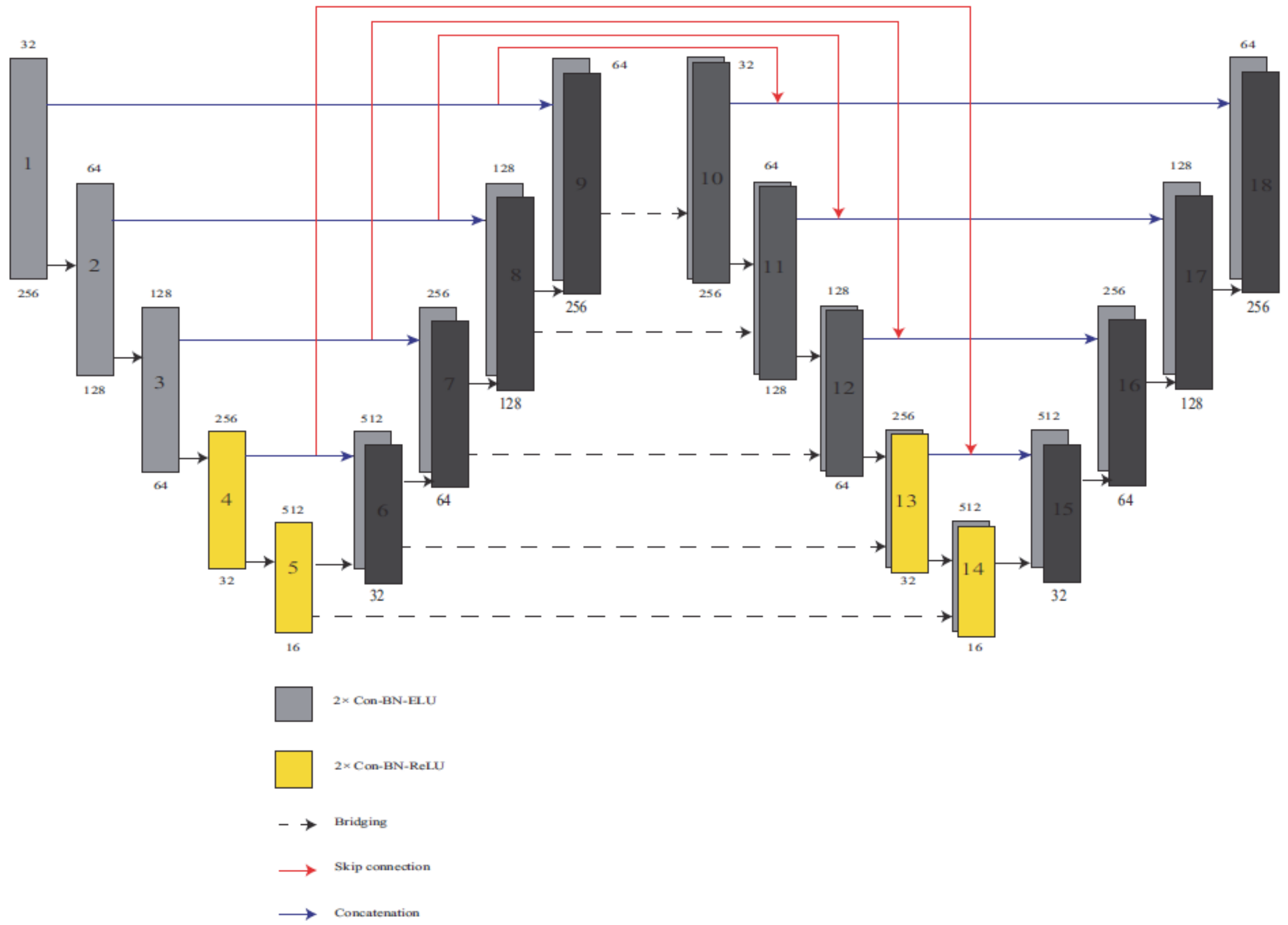


Figure 9 – Bridged U-net

Table 4. Influence of network bridging and skip connection. vDSC is the abbre-viation of volumetric Dice Similarity Coefficient

Method	Bridging method	Skip connection	Mean vDSC, %
U-net	None	None	86.73
Stacked U-net	None	None	85.57
Stacked U-net	Addition	None	86.99
Stacked U-net	Concatenation	None	87.85
Stacked U-net	Concatenation	Concatenation	86.02
Bridged U-net	Concatenation	Addition	88.12

Table 5. Quantitative comparison between the proposed method with othermethods on testing data

Team	DSC, %
Bridged U-net	89.96
DenseNet	88.98
U-net	88.06
ResNet	87.42
Stacked U-net	87.15

In [17], two new U-net-based architectures are proposed - Recurrent U-net and Recurrent Residual U-net (R2Net), Fig. 10.

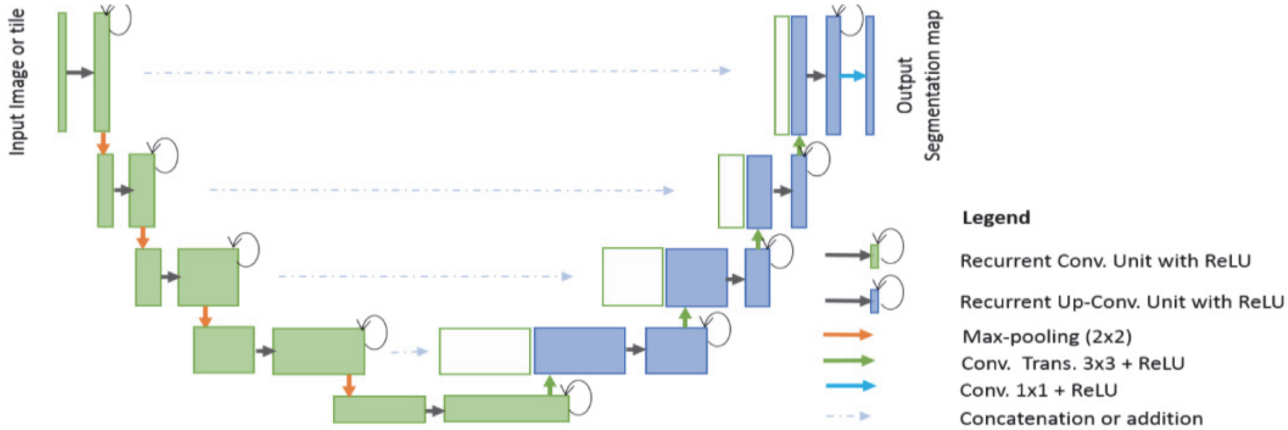


Figure 10 – Recurrent U-net architecture

In the case of R2Net, the recurrent convolutional block with ReLU and the recurrent unwrapping block with ReLU are replaced by the recurrent residual block with convolution and ReLU and the recurrent residual block with sweeps and ReLU. The blocks are shown in Figure 11.

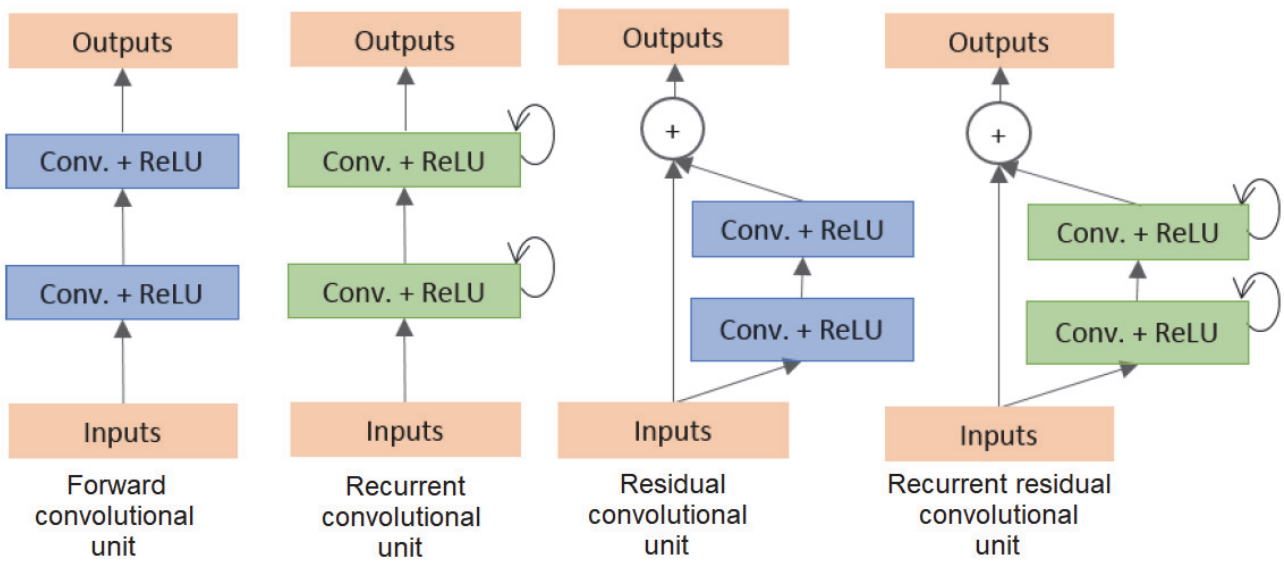


Figure 11 – Blocks used in the Recurrent U-net and Recurrent Residual U-net architectures

Recurrent layers allow to accumulate features, which increases accuracy, and residual layers contribute to better learning of a deep neural network. This was confirmed during the experiments. On all datasets, the proposed methods showed better results than the original U-net. Figure 12 shows a typical accuracy result obtained when testing neural networks on a dataset for lung segmentation. The signature t = 2 means that in the model, after each simple convolutional layer, two successive recurrent convolutional layers follow. Figure 13 shows examples of Recurrent Residual U-net output segmentation masks. The authors noticed that at the borders of the masks, the confidence of the network is higher (the edges are yellow).

Methods	Year	SE	SP	JSC	F1-Score	AC	AUC
U-Net (t=2)	2018	0.9696	0.9872	0.9858	0.9658	0.9828	0.9784
ResU-Net(t=2)	2018	0.9555	0.9945	0.9850	0.9690	0.9849	0.9750
RU-Net (t=2)	2018	0.9734	0.9866	0.9836	0.9638	0.9836	0.9800
R2U-Net (t=2)	2018	0.9826	0.9918	0.9897	0.9780	0.9897	0.9872
R2U-Net (t=3)	**2018**	**0.9832**	**0.9944**	**0.9918**	**0.9823**	**0.9918**	**0.9889**

Figure 12 – The accuracy of the proposed models compared to the original U-net

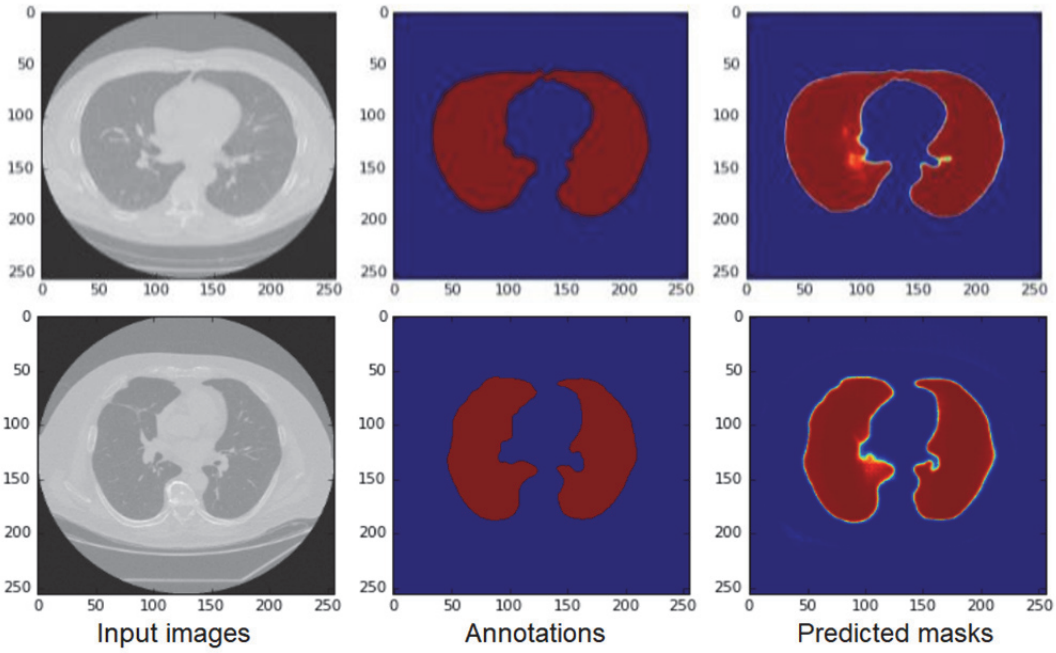


Figure 13 – Recurrent Residual U-net network prediction example

In [18], the authors propose the Sequential SegNet architecture (SeqSegNet), based on the SegNet segmentation architecture. SeqSegNet allows to qualitatively segment various organs in the image, i.e. solves the problem of multiclass segmentation. The neural network contains a sequential feature extractor, which receives a set of intra-slice feature maps obtained from a sequence of sliced images of computed tomography (CT) images and generates cross-feature features maps (in the case of a sequence of sliced images, intra-slice features are distinguished, obtained from the image of the slice, and inter-slice obtained from the slices relations), which is shown in Figure 14. The set of seven input feature maps is divided into a direct sequence (marked in blue), the reverse sequence (marked in green) and the target indication map (red). The feature extractor contains LSTM (Long Short-Term Memory) recursive blocks, and first the feature maps that are most unlike the target are processed, while the most similar ones are sent to the input of convolutional recurrent blocks last.

Figure 15 illustrates the general architecture of the neural network and the method of combining intra-slice and inter-slice feature maps. It can be seen that instead of copying the output feature maps at the outputs of the encoder blocks and subsequent merging at the inputs of the corresponding decoder blocks, the feature maps are processed by a sequential feature extractor.

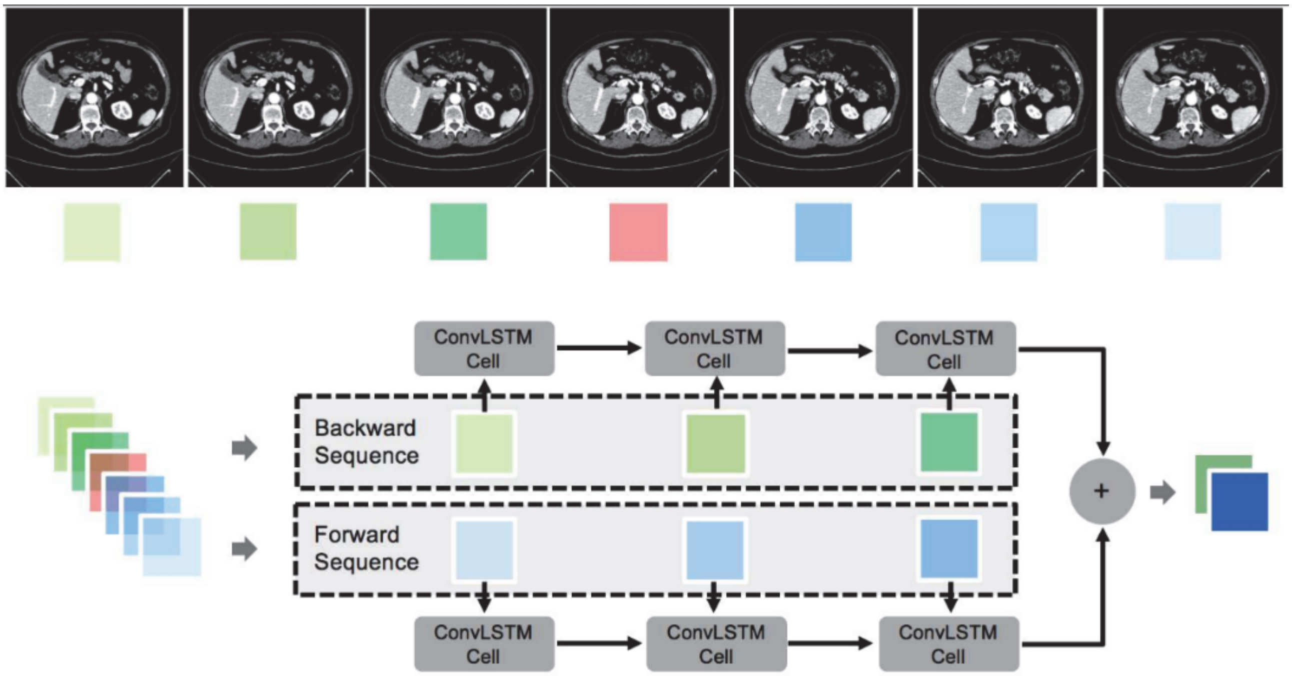


Figure 14 – Sequential Feature Extractor

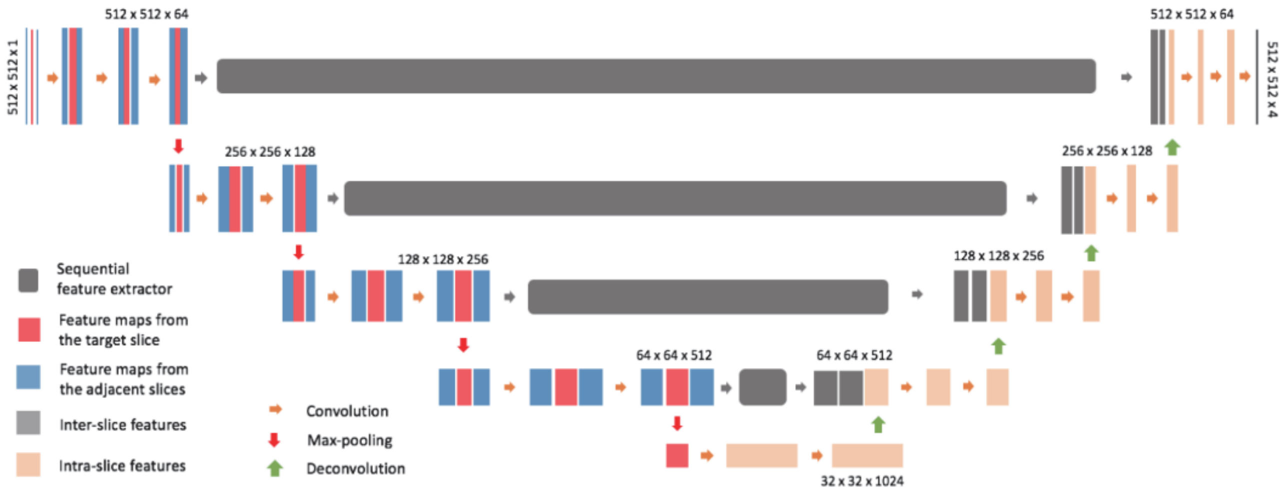


Figure 15 – SequentialSegNet Architecture

The SequentialSegNet architecture turned out to be more accurate than U-net in both the single-class segmentation problem and the multi-class segmentation problem, as expected (Table 6).

Table 6. SequentialSegNet (SSN) and U-net Evaluation Results

Methods	Dice Similarity Index Per Case, %		
	Liver	Spleen	Gallbladder
Single-organ U-net	90.45 ± 4.74	80.45 ± 13.15	59.15 ± 22.92
Multi-organ U-net	92.74 ± 4.82	91.17 ± 5.77	71.27 ± 18.48
Single-organ SSN	93.60 ± 4.19	89.94 ± 8.85	69.31 ± 21.79
Multi-organ SSN	**95.58 ± 1.57**	**91.41 ± 4.38**	**74.72 ± 13.83**

Two-dimensional methods tend to have limited segmentation performance, since large volumes of spatial information about an organ are discarded during the segmentation process in parts. The use of three-dimensional methods can also improve the quality of segmentation, since they use isotropic kernels to perform three-dimensional convolutions, while most MRI images of the prostate have anisotropic spatial properties. In addition, fully convolutional structural methods achieve good performance for localization problems, but neglect the classification of voxels for segmentation problems.

In [19], the authors proposed a 3D global convolutional adversarial network (3D GCA-Net) for solving the problems of segmentation of MRI images of the prostate (Figure 16). First, they designed the 3D ResNet encoder to extract three-dimensional features from sections of the prostate gland, and then developed a decoder that consists of a multi-scale three-dimensional global convolutional block and a three-dimensional border refinement unit to simultaneously solve classification and localization problems for volume segmentation. In addition, the authors combined an encoder-decoder segmentation network with an adversarial network in the training phase to ensure continuity of spatial predictions. Throughout the proposed model, they use anisotropic convolution processing to better study the features of MRI scans of the prostate gland.

This method was tested on two data sets: MICCAI PROMISE 12 (Table 7) and ASPS 13 (Table 8). Also, several segmented images are shown in Figure 17.

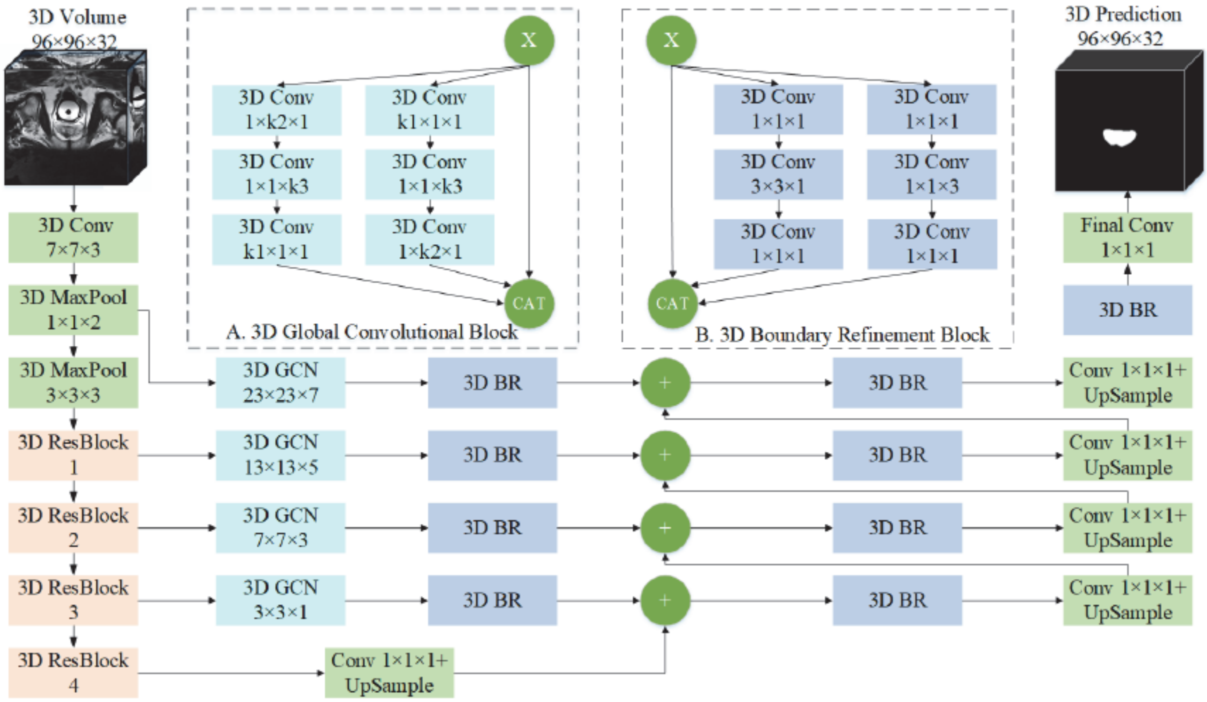


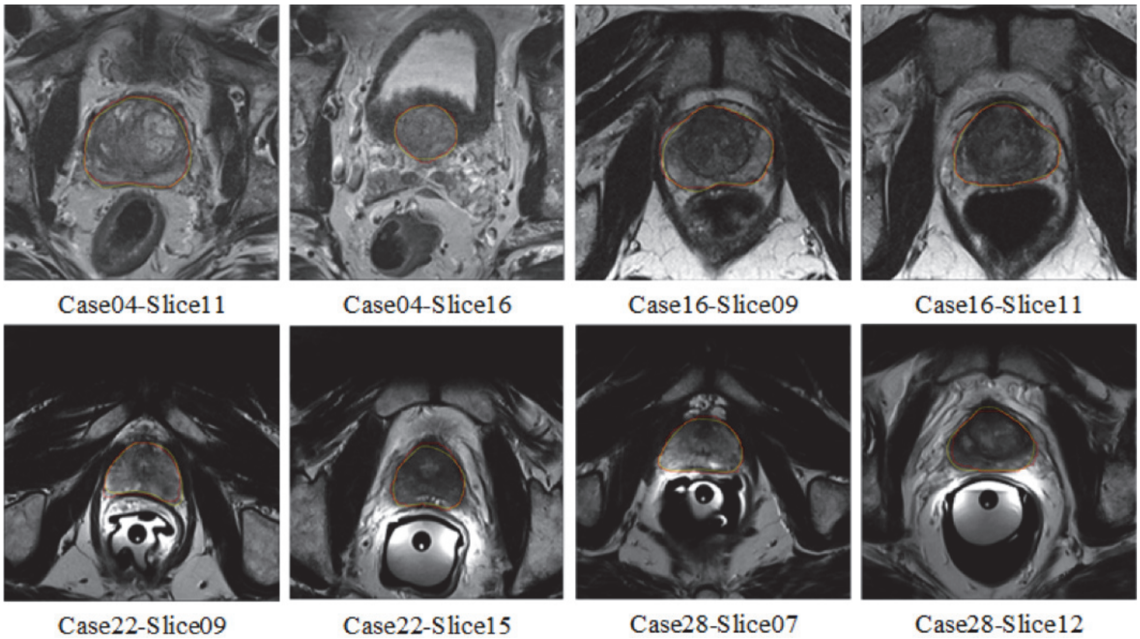
Figure 16 – Three-dimensional network encoder-decoder architecture for segmentation. A and B are a three-dimensional global convolutional block and a three-dimensional boundary refinement block, respectively.

Table 7. Comparison with several variations of convolutional encoder-decoder networks on the PROMISE12 dataset

Method	**Type**	**Dice coefficient**		
		Whole	**Base**	**Apex**
CAMP-TUM2	3D	0.869	0.843	0.844
UdeM 2D	2D	0.874	0.849	0.842
MBIOS	2D	0.881	0.850	0.847
BDSLab	3D	0.883	0.876	0.798
3D GCA-Net	3D	0.889	0.877	0.861
CREATIS	2D/3D	0.893	0.866	0.868
CUMED	3D	0.894	0.864	0.860

Table 8. Comparison with various methods on the NCI-ISBI 13 dataset

Method	Conv. layers	Parameters	Dice coefficient
V-net	31	65,191,134	0.841
3D U-net	23	33,854,722	0.862
ResNet-50	53	23,507,904	-
3D Encoder-decoder	141	29,601,094	0.878
3D GCA-Net	148	33,540,327	0.880



Figrкy 17 – The result of 3D GCA-Net on the PROMISE 12 dataset

Conclusion

A review of segmentation methods was carried out, groups of methods and their advantages and disadvantages were identified. The atlas segmentation method shows good results on large objects, but at the same time on small objects its accuracy drops significantly, which makes it not a universal solution. Neural network methods have good speed and quality. This group of methods is most effective for image segmentation. Classical methods can also be implemented for comparison with the results of methods based on deep learning or as a pre- and post-processing of the results. Thus, methods based on neural networks are defined as the most promising group of methods.

Modern neural network algorithms for solving the problem of medical images are considered. The most popular development base in this area is the U-net neural network, proposed in 2015. At the moment, its modifications have good performance and occupy high places in competitions.

There are several trends in the development of neural network algorithms for the segmentation of biomedical images. This is the use of residual connections to improve the quality and convergence of neural network training, the use of recurrent blocks to improve the quality of segmentation through the accumulation of features, as well as work with sequences of image slices through the use of 3D convolutions or recurrent blocks.

References

1. Dilpreet Kaur, Yadwinder Kaur: Various Image Segmentation Techniques: A Review. IJCSMC, 3(5), 809–814 (2014).

2. S. Kannan, Vairaprakash Gurusamy, G. Nalini: Review on image segmentation techniques. RTRICS, At Podi, 2014, https://www.researchgate.net/publication/273127438_REVIEW_ON_IMAGE_SEGMENTATION_TECHNIQUES, last accessed 2019/03/20.
3. S. Inderpal and K. Dinesh: A Review on Different Image Segmentation Techniques, IJAR, 4, (2014).
4. S. D. Pirozzi, A. S. Nelson, J. W. Piper: Atlas-based segmentation: comparison of multiple segmentation approaches for lymph level targets and normal structures in head and neck cancer. International Journal of Radiation Oncology, Biology, Physics (IJROBP), 81(2), 3294 (2011), https://www.redjournal.org/article/S0360-3016(11)02282-6/abstract, last accessed 2019/05/20.
5. Arnau Oliver, M. Cuadra, Mariano Cabezas, Xavier Lladó, Jordi Freixenet: A review of atlas-based segmentation for magnetic resonance brain images. Computer methods and programs in biomedicine, 104(3), 2011, https://www.researchgate.net/publication/51599738_A_review_of_atlas-based_segmentation_for_magnetic_resonance_brain_images, last accessed 2019/05/20.
6. A.Y. Zubov, O.V. Senyukova:. Segmentation of magnetic resonance images of human brain by registration with multiple atlases. In: GraphiCon 2015 Proceedings, 56-60 (2015), http://www.graphicon.ru/html/2015/Proceedings.pdf, last accessed 2019/06/20.
7. Ultrasound Nerve Segmentation. Identify nerve structures in ultrasound images of the neck – Kaggle: Your Home for Data Science, 2016, https://www.kaggle.com/c/ultrasound-nerve-segmentation, last accessed 2019/05/20.
8. KONICA MINOLTA. Pathological Image Segmentation Challenge – topcoder: leading in crowdsourcing, 2017, https://www.topcoder.com/KonicaMinoltaChallenge, last accessed 2019/05/20.
9. Digital Pathology: Segmentation of Nuclei in Images Challenge – MICCAI 2018 Computational Precision Medicine, http://miccai.cloudapp.net/competitions/83, last accessed 2019/05/20.
10. PAIP 2019 Challenge – MICCAI 2019 Grand Challenge for Pathology, https://paip2019.grand-challenge.org/, last accessed 2019/05/20.
11. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: NIPS. pp. 1106-1114 (2012)
12. Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: NIPS. pp. 2852-2860 (2012)
13. J. Long, E. Shelhamer, T. Darrell: Fully convolutional networks for semantic segmentation. Proc. IEEE Conf. Comput. Vis. Pattern Recognit., (3431-3440) 2015, https://people.eecs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf, last accessed 2019/05/20.
14. Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention. MICCAI, 9351, 234-241 (2015), https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/, last accessed 2019/05/20.
15. Chen L.-Ch., Papandreou G., Kokkinos I., Murphy K., Yuille A.L.: Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. ICLR (2014), http://arxiv.org/abs/1412.7062, last accessed 2019/05/20.
16. Wanli Chen, Yue Zhang, Junjun He, Yu Qiao, Yifan Chen, Hongjian Shi, Xiaoying Tang: Prostate Segmentation using 2D Bridged U-net. https://arxiv.org/abs/1807.04459, last accessed 2019/05/20.
17. Alom Md. Zahangir, Hasan Mahmudul, Yakopcic Chris, M. Taha Tarek, Asari Vijayan: Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation. Journal of medical imaging, 6(1), 2019, https://www.ncbi.nlm.nih.gov/pubmed/30944843, last accessed 2019/05/20.
18. Yao Zhang, Xuan Jiang, Cheng Zhong, Yang Zhang, Zhongchao Shi, Zhensheng Li, Zhiqiang He: SequentialSegNet: Combination with SequentialFeature for Multi-organ Segmentation. In: 24th International Conference on Pattern Recognition (ICPR). Beijing, China, 2018, https://dblp.org/pers/hd/l/Li:Zhensheng, last accessed 2019/05/20.
19. Haozhe Jia, Yang Song, Donghao Zhang, Heng Huang, Dagan Feng, Michael Fulham, Yong Xia, Weidong Cai: 3D Global Convolutional Adversarial Network for Prostate MR Volume Segmentation. https://arxiv.org/abs/1807.06742, last accessed 2019/05/20.

A.V. Kapustin, Y.V. Loskutov, I.A. Kudryavtsev

PROVIDING VERTICAL SUPPORT OF A MEDICAL EXOSKELETON. PROBLEMS AND TECHNICAL SOLUTIONS

Volga State University of Technology, Yoshkar-Ola, Russia
KapustinAV@volgatech.net; LoskutovYV@volgatech.net; KudryavtsevIA@volgatech.net

Abstract

Based on the review of Russian and foreign scientific papers and analysis of unstable positions, the present work suggests possible ways to solve the problem of stability of a medical exoskeleton. The paper contains design and technical solutions to provide support; ways based on the controlling processor algorithm, models based on equilibrium principle of an unstable inverted pendulum with one or two links. To provide different ways for the implementation of balance maintenance system, there have been identified basic criteria which maintain the balance. Examples which illustrate different ways to provide balance are also given. There have been introduced applied design solutions and principles which will prevent the human-exoskeleton system from turning over. The problems having difficult solutions which are still to be found are stated.

Keywords: exoskeleton, providing exoskeleton's balance, human locomotions, balance stability

Nowadays, there are frequently used medical exoskeletons designed for rehabilitation purposes of those people who have lost their ability to walk. Their application has proved the effectiveness of the method of the outer movement reproduction to restore muscular activity.

One of the problems while using and designing medical exoskeletons is considered to be the problem of providing a stable vertical support of the "human-exoskeleton" mechanical system in the moving process. The specific feature of these exoskeletons, unlike the other exoskeletons, is that a person (user, operator, patient) is unable (or only partially able) to hold upright position while moving. That is why the task of providing vertical support of the "human-exoskeleton" system rests with the design of an exoskeleton, or with a user and care-taking personnel.

When designing the exoskeleton control system, it is necessary to set the laws of motion and eliminate unnecessary ("extra") movements. An exoskeleton must have appropriate sensors and mechanisms which would respond to the movement and position of the body in space taking into account right locomotions and system balance stability. The developed algorithms to control an exoskeleton must be able to "feel" and "predict" human behavior, provide action timing which would eliminate serious injuries for a user [1, 2].

After a series of the experiments performed with "REMotion" exoskeleton prototypes, there have been found some approaches to solve the task to provide the stability of an exoskeleton with a patient inside it.

To perform movements, a person and an exoskeleton act within the available degrees of freedom of a lower extremities kinematic chain. For a person it is more difficult. If we take into consideration Hanavan model, hip and ankle joints can be represented in the form of spherical kinematic pairs each having three degrees of freedom. A knee joint is considered as a turning kinematic pair. Thus, to provide stability a person uses 7+7 (right leg plus left leg) degrees of freedom which make 14 controlled movements (in reality there are much more of them), see Figure 1.

The degrees of freedom of exoskeleton kinematic chains usually do not exceed this number and vary between 2 and 10. However, in an exoskeleton design a distinction should be drawn between controlled (active) degrees of freedom and uncontrolled (passive) degrees of freedom. REMotion exoskeleton has 2+2 active degrees of freedom in the hip and knee joints and 1+1 passive degrees of freedom in an ankle joint. Such design is quite common and is due to the maximum simplicity and performance. It should be noted that all six degrees of freedom of a REMotion exoskeleton operate in one saggital plane and have no opportunity to move in the frontal plane.

A "user-exoskeleton" system is considered to be quite unstable in case if a person suffers from locomotor disorders. Nowadays, there are a few mathematical models of such a system each of which is necessary and sufficient to solve a specific balance support task.

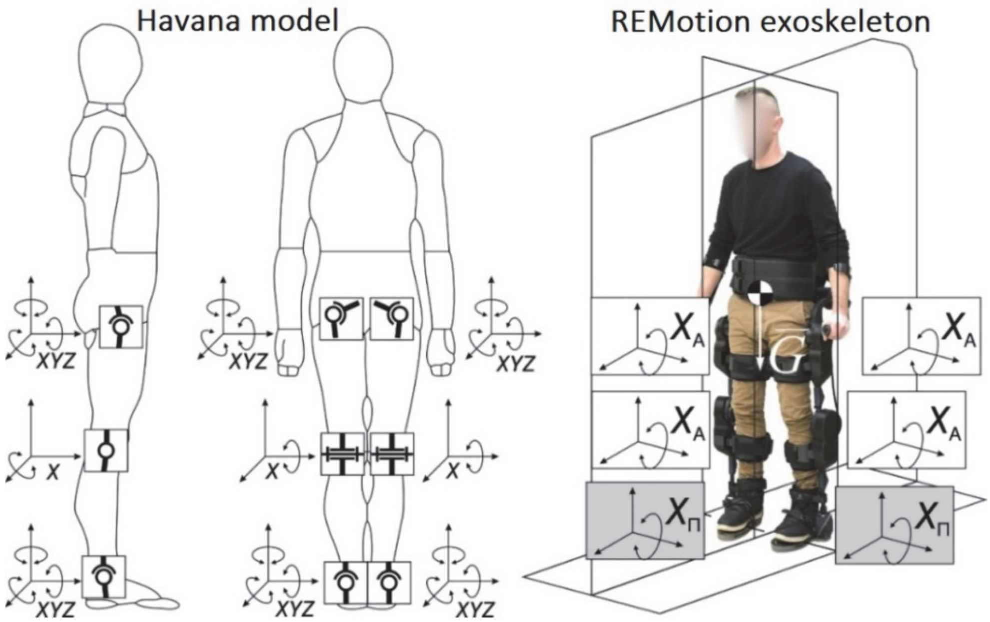


Figure 1 – The number of degrees of freedom of lower extremities for a person and REMotion exoskeleton

The design model for a "human-exoskeleton" system stability can be represented as a spherical pendulum or inverted plane pendulum [3, 4, 5]. Based on the chosen criteria, initial and boundary conditions, the pendulum can have one, two or many chains. The more chains (and consequently degrees of freedom) there are in a model, the more complex it is to practically solve the balance support task, see Figure 2.

The differential equation for one-chain pendulum motion is represented as [3]:

$$mr^2\ddot{\beta} - mgb\sin\beta = L, \tag{1}$$

where m is a pendulum mass, b is the distance between point O and centre of mass C, r is the radius of pendulum gyration with the respect to point O, β is the deflection angle, g is gravitational acceleration.

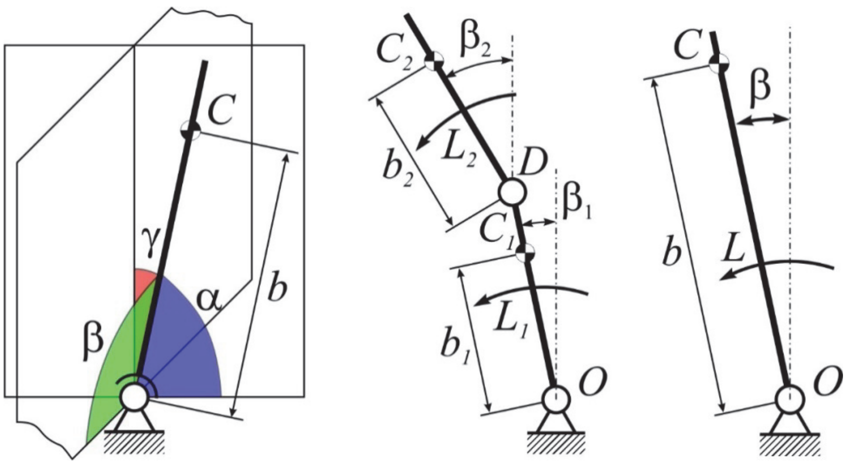


Figure 2 – Possible design models to set the exoskeleton balance stability task

If we consider an exoskeleton itself as a mechanical system with two degrees of freedom, the mathematical model can be represented as a two-chain inverted pendulum. The equation for a two-chain inverted pendulum is much more complex. It can be represented in matrix form [3]:

$$[A]\{\ddot{\phi}\} + [F]\{\dot{\phi}^2\} + [B]\{\sin\phi\} = L\{c^{(i)}\}. \tag{2}$$

Here matrixes and column-vectors are represented as:

$$[A]=\begin{Vmatrix} a_{11} & a_{12}\cos(\phi_2-\phi_1) \\ a_{12}\cos(\phi_2-\phi_1) & a_{22} \end{Vmatrix}, \tag{3}$$

$$[F]=\begin{Vmatrix} 0 & -a_{12}\sin(\phi_2-\phi_1) \\ a_{12}\sin(\phi_2-\phi_1) & 0 \end{Vmatrix}, \tag{4}$$

$$[B]=\begin{Vmatrix} -b_1 & 0 \\ 0 & -b_2 \end{Vmatrix}, \tag{5}$$

$$\{\ddot{\phi}\}=\begin{Vmatrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \end{Vmatrix}, \{\dot{\phi}^2\}=\begin{Vmatrix} \dot{\phi}_1^2 \\ \dot{\phi}_2^2 \end{Vmatrix}, \{\sin\phi\}=\begin{Vmatrix} \sin\phi_1 \\ \sin\phi_2 \end{Vmatrix}, \{c^{(1)}\}=\begin{Vmatrix} 1 \\ 0 \end{Vmatrix}, \quad \{c^{(2)}\}=\begin{Vmatrix} -1 \\ 1 \end{Vmatrix}. \tag{6}$$

The corresponding coefficients are equal to:

$$a_{11}=I_1+m_2l^2=I_{1O}+m_2\left|OD\right|^2, a_{22}=I_2=I_{2D}, a_{12}=m_2r_2l=m_2\left|DC_2\right|\cdot\left|OD\right|,$$
$$b_1=\left(m_1r_1+m_2l\right)g=\left(m_1\left|OC_1\right|+m_2\left|OD\right|\right)g, b_2=m_2r_2g=m_2\left|DC_2\right|g,$$

where I_1 and I_2 are moments of inertia of the first and second chains with the respect to joints O and D respectively, m_1 and m_2 are masses of the first and second chains, r_1 and r_2 are the distances from joints O and D to centers of mass C_1 and C_2 of the first and second chains respectively, l is the length of the first chain OD, g is gravitational acceleration. It should be assumed that the chain centers of mass do not match with the joints. If the control torque L is applied to the pendulum support O, then i=1. If the control torque is applied in a joint between chains, then i=2.

Further increase of chains of the inverted pendulum leads to the sophistication of the model and causes difficulties while looking for possible solutions. The same opinion can be found in the work [6]: the solution of the obtained system of inequalities in analytical form is lengthy and it is impossible to draw a 4D-area which would correspond to the solution of these inequalities.

It is necessary to mark the basic problems to provide stable position and movements of a person in an exoskeleton. The first one is the complexity of the mathematical model used to implement the operation algorithm of the stability support system. If we increase degrees of freedom of the design model, it becomes impossible controller's processor to solve it.

The second problem is the absence of degrees of freedom in the frontal plane. To simplify the design and reduce the costs of a medical exoskeleton, we intentionally eliminate the controlled degree of freedom in the frontal plane. Thus due to its design, it is impossible to create a motion which would prevent falling.

The third problem deals with the first one because of the natural motion of a person inside an exoskeleton. It is observed an uncontrolled shift of the common gravity center of the whole system. It is impossible to predict the position of the common mass centre of a "human-exoskeleton" system. The only thing which we can do in this case is to measure it somehow and extrapolate with the help of motion sensors.

The fourth problem is significant power consumption and sophistication of the design to create a balancing moment. The problem deals with some critical conditions upon the reach of which it would be useless to balance the system due to high power waste. Perhaps, we should avoid the fact that the power consumption for stability maintenance would not exceed the power consumption required for movement.

The fifth problem, which is considered to be main one, deals with the security. If the manufacturer states that its exoskeleton is equipped with the balance support system, it must totally eliminate the possibility to fall. Any falling is regarded as an accident and it is the manufacturer who takes the responsibility for it. Nowadays, there are no medical exoskeletons which allow motion without support.

There are a lot of ways to stabilize the objects when the point of load application does not correspond with the center of mass. For instance, based on the inverted pendulum we can point out the following ways to stabilize the system.

Method 1. The use of additional outer supports (links).

This method is considered to be the simplest one as based on the motion of a usual pendulum near the stable equilibrium position. Outer support can be created with the help of support tools. See Figure 3.

In case if a patient uses crutches, the task to provide support becomes rather simple. The support area increases significantly and it provides better stability of the "four-point gait". The support area is a polygon with the vertexes in the exoskeleton's feet and points where the crutches touch the floor. However, for a locomotor patient it is difficult to choose and set the most stable position of a crutch. This fact causes a problem of balance stability when a patient uses crutches.

In this case it is possible to provide support by the correlation of the step length and the position of a crutch which was experimentally proved [7]. It was found out that if the step length is fixed, different position of the crutches can cause different influence on the balance and gait stability. Having analyzed possible forms of support planes, it is possible to create the conditions at which the position of a "human-exoskeleton" system can be always stable.

Figure 3 – Use of additional support tools

In reality the "four-point gait" can be performed in the following way. First, a person sets a crutch as he moves and shifts his weight to this crutch. The data obtained from the pressure and position sensors placed in all four support points let us adjust the step length to provide the maximum stability at the end of the step. Thus, balance support is due to prediction of the final leg position.

Apart from crutches, walking frames having four legs can be used as support tools while walking in an exoskeleton. A walking frame provides extra support area in the form of a tetragon. One of the drawbacks of such a support tool is that it is difficult to use it if the surface is uneven, e.g. in case if there are stairs, soft ground, some rough spots on the road. So, we offered a walking frame which design would allow to adjust it to the rough surface. The walking frame can automatically change its leg length to provide horizontal position of the handle as it is shown in Figure 4.

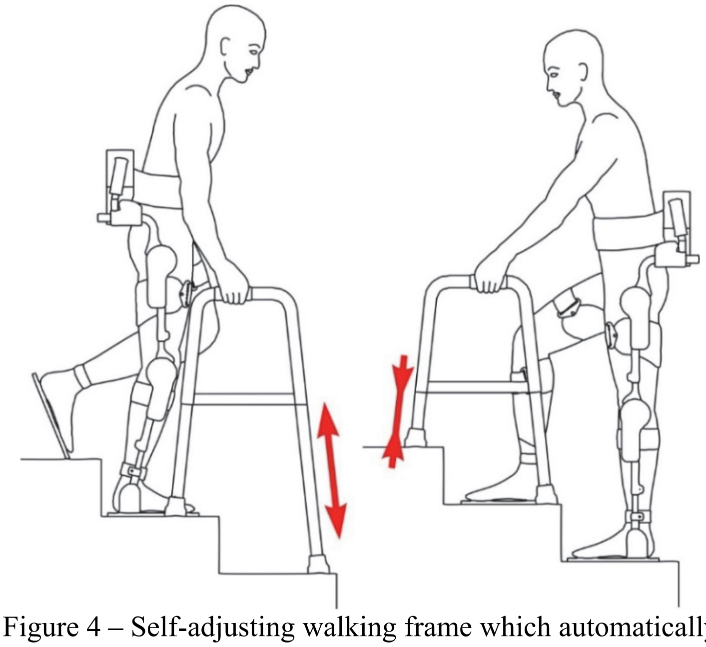


Figure 4 – Self-adjusting walking frame which automatically provide horizontal position of the handle

Method 2. Feedback between position control system and motion algorithm.

This method requires the development of a sensitive feedback system and an algorithm which would return the pendulum to the equilibrium position taking into account the prior movement. The task becomes even more difficult if we deal with a multi-chain pendulum. This method is widespread nowadays.

There was offered an exoskeleton design equipped with the hybrid drive [8] which uses the combination of pneumatic muscles and electric motors to create a required controlling moment to be able to keep the balance without crutches. That research focused its analysis only on the robotic system; a patient was not taken into account. There was offered and tested a robotic exoskeleton with a balance stabilization mechanism. The project needs further research to control the mass center shift of a body in the frontal plane. The developers of Mindwalker [9] exoskeleton designed the control methods to provide active gait in the saggital and frontal planes. Its stable gait without crutches was shown for able-bodied people.

The criteria of balance stability of a human and exoskeleton based on the idea of catching and extrapolating the common centre of masses of the system were used for automated control and balance recovering against moving disturbances [10]. That research also dealt with able-bodied people having sense of equilibrium and full mobility. There was offered a flexible zero dynamics control to provide automated gait. Lower part of a human body and exoskeleton were considered as solid body without any disturbance caused by legs. They also offered an improved model of an exoskeleton which provided balance with the help of a spring-mass-damper system in each joint having the parameters obtained by extrapolation and optimization of the data received during the experiments.

The condition of the constant balance in a static situation can be expressed in the following way: a vertical projection of a mass centre must be in the support footing. For small movements within the support area there was developed a model [4] which considers a body as an inverted pendulum the motion of which can be described with the help of equation (1). The solution of the task is aimed at a person whose support area includes his feet (exoskeleton's feet) and space between them.

The authors of the paper [7] developed an exoskeleton for people suffering from minor muscle-skeleton disorders. Such an instruction was also given in a paper [11]. An exoskeleton has two main functions: assistive walking and maintaining walking stability. A hip joint drive is equipped with suitable sensors and a control system which allows the speed of a moving in an exoskeleton person up to 0.8 m/s (2.88 km/h). The specifications include the following parameters: if a person weighs 80 kg, each drive must have a rated load torque equal to 40 Nm which can be increased up to 80 Nm; maximum angular rate of a hip joint equals 150 degrees/s. The exoskeleton structure has 4 degrees of freedom, two drives for each hip joint in saggital and frontal planes.

The developers offer two systems to prevent falling: passive and active. The passive system has a constructive character and consists of stops which prevent the motion of chains in critical positions. The active security system has a complex design and consists of software which controls the exoskeleton's drives based on the data obtained from the position and rotation velocity sensors of each drive, current, temperature and turning moment. If any of these indicators go beyond the safety limits, the software stops the motor and produces an alarm signal. The mathematical model for a walking stabilization is based on the principles of an inverted pendulum and is described in work [4]. The authors note that the correction of the motion is due to the extrapolation of the mass centre.

Paper [12] describes the functionality of the stability support by an exoskeleton's controller. The controller allows to change the length of the step cycle, the length of the step and to count the time of the walk during the walking process. It is pointed out that for the rehabilitation of patients it is very important both to repeat the natural locomotions and try to feel the state of imbalance. It helps a patient to automatically choose a suitable step length and cycle duration. In this paper the stability support is provided due to the angle change cyclogram adjustment to the ideal state. The exoskeleton has 6 degrees of freedom, two of which are structural couplings in the ankle joints.

In paper [13] the authors use the model of an inverted pendulum mentioned in Hof's works [4, 14]. The control system corrects the position using three aspects: mass centre shift, the effort created by the exoskeleton's leg due to the accelerated motion, step length correction. Due to the use of an alternating force in the support point relative with respect to the mass centre point, the system brings the exoskeleton back to the stable position.

Method 3. The use of the law of conservation of momentum. Gyroscopic effect.

This is the way where the stability support is due to the inertia. This way uses a model of two-chain inverted pendulum described by Formalsky A.M. [3].

The balance stabilization is due to the inertia of the upper chain and a controlling moment in the medium joint. The value of the moment is set by the controller after it processes the data of the velocity and position sensors.

The mass centre of the upper chain can be both in the rotation point of the medium joint and can be shifted, see Figure 5. The way is similar to what a rope-walker does keeping his balance on the rope with the help of long stick or heavy clubs.

The two-chain pendulum which motion can be described with the help of equations (2) apart from the position of stable balance with two hanging down chains and unstable balance with two inverted chains can have two more balanced states. If we take into consideration a state when the first chain hangs down and the second is inverted, we can have an opportunity to control the balance in the frontal plane of a "human-exoskeleton" system (Fig. 6). If we linearize the equation (2) near the state of equilibrium $\phi_1 = 0,\ \ \phi_2 = \pi,\ \ \dot{\phi}_1 = \dot{\phi}_2 = 0$, we get the system

$$\left[A_o\right]\left\{\ddot{\phi}\right\} + \left[B\right]\left\{\phi\right\} = \left\{\,L\right\}, \tag{7}$$

where

$$\left[B\right] = \left\| \begin{matrix} -b_1 & 0 \\ 0 & -b_2 \end{matrix} \right\|,\ \left\{\ddot{\phi}\right\} = \left\| \begin{matrix} \ddot{\phi}_1 \\ \ddot{\phi}_2 \end{matrix} \right\|,\ \left\{\phi\right\} = \left\| \begin{matrix} \phi_1 \\ \phi_2 \end{matrix} \right\|,\ \left\{\,L\right\} = \left\| \begin{matrix} -L \\ L \end{matrix} \right\|.$$

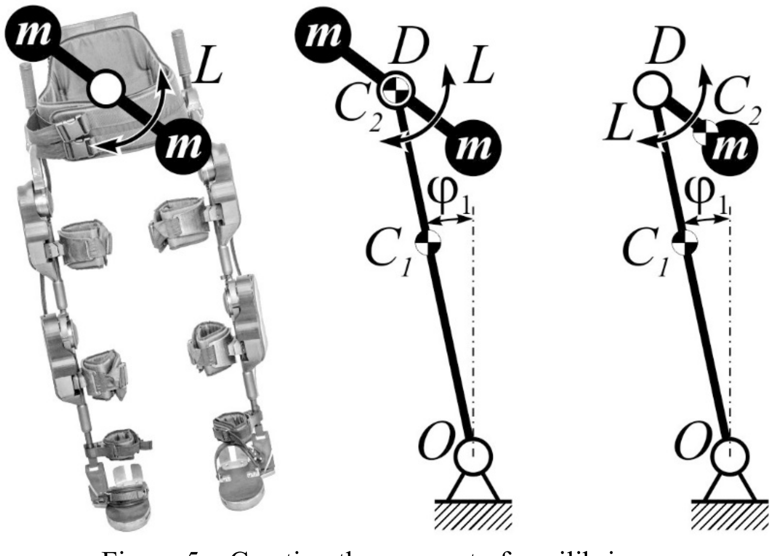


Figure 5 – Creating the moment of equilibrium

Acceptable control moments act in the medium joint and are sectionally continuous functions *L(t)* which are limited by a constant value in a module L_0: $L \le L_0$. At the beginning and during the whole stabilization process the pendulum is near the marked position of equilibrium. The drift of the angle is considered to be within 8 or 10 degrees. Circular motion of the whole pendulum or of one of its chains is not taken into consideration, i.e. the stabilization task is considered locally.

In this case the equation (7) has only one positive proper value in the right semi-plane of the whole plane, i.e. the imbalance degree of such a system equals 1 apart from the pendulum with two inverted chains and the imbalance degree which equals 2. It significantly simplifies the calculation. If we have a linear (with saturation) feedback in one unstable variable, the domain of attraction corresponds with the whole controlled domain and is considered to be a maximum possible one. The solution of this task can be found in paper [13]. The balance stabilization is done with the help of inertia of the upper chain and controlling moment in the medium joint. The value of the moment is set by the controller after it processes the data obtained from the velocity and position sensors located in the exoskeleton control system.

Using the same law of the conservation of the angular momentum, it is possible to stabilize the vertical position with the help of two gyroscopes (gyrodynes) located in the frontal and saggital planes.

In conclusion, we would like to point out that a person being an erect-walking creature is constantly in a state of unsteady balance. The human being learnt how to use all the ways of dynamic maneuvering to keep stable vertical position. Improving our machine, we look at a human being and try to create a harmonious combination of motion and strength. The aim of a medical exoskeleton is to reconstruct those simple basic movements to launch the perfect balancing system a person has.

References

1. Kapustin, A.V., Loskutov, Y.V., Skvortsov D.V., Nasibullin, A.R., Klyuzhev K.S., Kudryavtsev, A.I. Design models of rehabilitation exoskeleton control system // Annals of Volga state university of technology / Radio technology and IT systems. 2018. № 2 (38). P. 77-86. DOI: 10.15350/2306-2819.2018.2.77
2. Loskutov, Y., Kapustin, A., Kudryavtsev, A., Nasibullin, A., Lebedeva, A. (2018). Synthesis of exoskeleton control algorithms based on kinematic analysis of locomotions and human gait modelling. Journal of Applied Engineering Science, 16(4), 583-591. doi:10.5937/jaes16-17230
3. Formalsky, A.M. Control of the motion of unstable objects. – M.: Fizmalit, 2014. – 232 p. – ISBN 978-5-9221-1460-8.
4. Hof A.L., Gazendam M.G.J., Sinke W.E., The condition for dynamic stability // Journal of Biomechanics., vol. 38, no. 1, pp. 1–8, 2005.
5. Vallery, H., Bögel, A., O'Brien, C., Riener, R. (2012). Cooperative control design for robot-assisted balance during gait. At-Automatisierungstechnik, 60(11), 715-720. doi:10.1524/auto.2012.1041
6. Borisov, A.V., Konchina L.V., Abrosov Y.A. The algorithm to control the vertical stability of an exoskeleton with a person inside while moving // Problem of social security in Russia. – M.: Russian university of transport, 2016. - №2 – P. 184-201.
7. Chen, C., S. Zhang, C. Wang, G. Wu, and X. Wu. 2017. Dynamic step length planning method based on stable threshold analysis for exoskeleton. Yi Qi Yi Biao XueBao/Chinese Journal of Scientific Instrument 38, no. 3: 523-529, www.scopus.com.
8. Mummolo C., Peng W., Agarwal Sh., Griffin R., Neuhaus Peter D., Kim J.H. Stability of Mina v2 for Robot-Assisted Balance and Locomotion // Frontiers in Neurorobotics. – October 2018. – V.12., A.62. – 16 p. DOI: 10.3389/fnbot.2018.00062
9. Barbareschi G, Richards R, Holloway C, Carlson T, Thornton M. Statically vs dynamically balanced gait: Analysis of a robotic exoskeleton compared with a human // 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), August 2015, 2015. – 4 p. DOI: 10.1109/EMBC.2015.7319937.
10. Yatsun S.F., Savin S.I., Malchikov A.V. The study of the controlled movement of an exoskeleton in the frontal plane in the balance recovering mode // Extreme robotics. – 2016. T. 1. № -1. P. 236-245.
11. Vallery, H., Bögel, A., O'Brien, C., &Riener, R. (2012). Cooperative control design for robot-assisted balance during gait. At-Automatisierungstechnik, 60(11), 715-720. doi:10.1524/auto.2012.1041
12. Andres Martinez, Brian Lawson, Michael Goldfarb. A Controller for Guiding Leg Movement during Overground Walking with a Lower Limb Exoskeleton // IEEE Transactions on robotics, 2017. DOI: 10.1109/TRO.2017.2768035
13. Huynh, V., C. Bidard, and C. Chevallereau. Balance Control for an Active Leg Exoskeleton Based on Human Balance Strategies, vol. 48, 2018. www.scopus.com, doi:10.1007/978-3-319-59972-4_15
14. Hof, A. L. "The Equations of Motion for a Standing Human Reveal Three Mechanisms for Balance." Journal of Biomechanics, vol. 40, no. 2, 2007, pp. 451-457. www.scopus.com, doi:10.1016/j.jbiomech.2005.12.016.

Yu.V. Loskutov, A.V. Kapustin, I.A. Kudryavtsev

CALCULATED JUSTIFICATION OF ENERGY EQUIVALENT OF SERVICE LIFE TESTS OF A MEDICAL EXOSKELETON

Volga State University of Technology, Yoshkar-Ola, Russia
LoskutovYV@volgatech.net;KapustinAV@volgatech.net; KudryavtsevIA@volgatech.net

Abstract
In the work, a calculated justification of equivalence in terms of the energy indices of the modes of operation of the exoskeleton with a weight dummy was performed. Modes of operation are provided: cycling on a stationary bike, walking, sitting down and standing up. Criteria's for the equivalence of movements in the hinge are accepted: the rms torque, equivalent power, the average speed turns between the links. Comparison of the values of the main parameters of locomotion gives small discrepancies of the largest equivalent torques, average speeds of revolutions and equivalent powers in the drives. According to the energy indices of the operating modes of the exoskeleton with the locomotion weight dummy (walking, standing up, sitting down and cycle-pedaling) may well be equivalent in the right choice of movement parameters. It is shown that the use of an exercise bike may well be acceptable to conduct laboratory reliability tests of the exoskeleton.
Keywords: exoskeleton, service life tests, human locomotion, energy equivalence, walking, cycling.

Implementing a complex project called “Creation of high-tech manufacturing for a multifunctional robotic medical exoskeleton (“REM”)” (code 2017-218-09-1807 according to the governmental decree dated April 9, 2010 N 218) resulted in the development of a multifunctional robotic medical exoskeleton REMotion designed by Volga state university of technology. The exoskeleton is to be produced at Volzhsky Electromechanical plant [1, 2]. This exoskeleton is to be used to help patients suffering from muscle-skeleton disorders.

During the launch of the manufacturing process of a product or during the refining process it is necessary to perform laboratory tests to check the reliability, i.e. to perform the service life tests. They are performed to define or evaluate the technical resource of the whole product or of some of its parts. These tests can be done both during sampling tests or acceptance tests depending on the purpose of a product. The tests check parameters of an exoskeleton specified by the developers, find out some drawbacks and manufacturing errors of some links, algorithm errors which occur in the electromechanical part and in the processor.

The service life tests specifications included the cycle which corresponds to the rehabilitation period of a patient having:

1. Walking along a straight line at a speed equal to 1.5 km/h.
2. Five cycles of standing up and sitting down. The duration of each one is about 30 sec.

The duration of a rehabilitation cycle is one hour and last in a non-stop mode. During the tests we took into account the parameters of an average person (patient): weight – 75 kg, height – 175 cm.

The service life tests were performed using a weight dummy (its weight was 75 kg) and included 160 hours of walking on the running machine and 3500 cycles of “standing up and sitting down”. In total it required 10 days of work of two laboratory assistants (see Fig. 1).

The performance of the tests in the regular operation mode (walking, sitting down) is dangerous for an operator, it requires a lot of space, time and operators and assistants looking after the balance stability and failure-free performance of a human-exoskeleton system (or weight dummy-exoskeleton system). So the relevance of the research is due to the necessity to develop some fixed-site ways to tests an exoskeleton which would allow to create various operation cycles including those performed in the forced conditions eliminating spatial motion.

The aim of the paper is to evaluate the opportunity and explain the tests of electromechanical parts of a medical exoskeleton drive performed on an exercise bike.

During the work we planned to determine the energy indicators for exoskeleton operation modes when walking, sitting down, standing up, cycling, as well as to choose the criteria of equivalence and give the justifying calculations concerning the energy indicators for chosen modes while working with a weigh dummy.

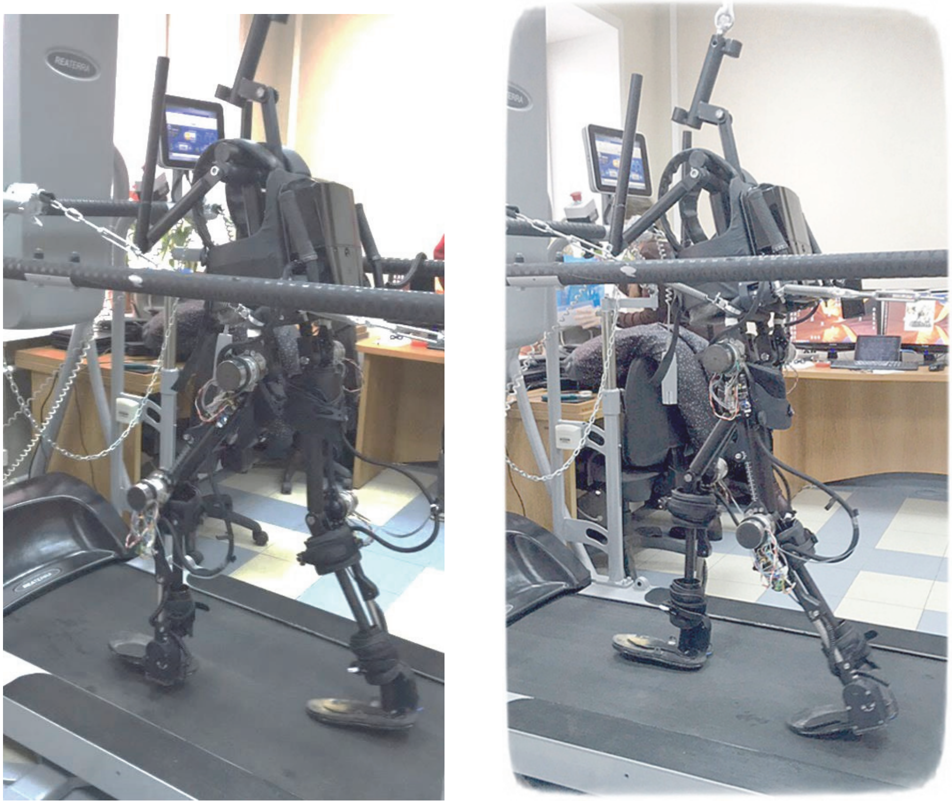


Figure 1 – Service life tests of “REMotion” exoskeleton performed on the running machine

Functional motion analogies

To evaluate the possibilities of the service life tests of an exoskeleton on an exercise bike, we should consider the parameters of cycling: frequency and step. These parameters have analogues in other types of functional motions, for example, while walking along the horizontal surface [3, 4]. The analysis of the walking shows that specific work $\tilde{A}$ (per unit of the distance), developed during the walking power P and ideal step length $L_{\text{опт}}$ can be counted using the following formulae [5]:

$$\tilde{A} = \left(\frac{mgL}{4h}\right) + \left\{4\mu\frac{mgh}{L}(Fr)^2\right\}, \tag{1}$$

$$\tilde{A}_{\text{опт}} = 2 \cdot m \cdot g \cdot (Fr) \cdot \sqrt{\mu}\,, \tag{2}$$

$$P = \left(\frac{mgL}{4h}\text{v}\right) + \left\{4\mu\frac{m}{L}\text{v}^3\right\}, \tag{3}$$

$$L_{\text{опт}} = 4 \cdot \text{v} \cdot \sqrt{\frac{\mu \cdot h}{g}}\,, \tag{4}$$

where m is a person’s mass; g is a gravitational acceleration; L is a step length; h is a vertical point of the mass center position; $\mu = \frac{m_{leg}}{m_{body}}$ is a ration of leg mass to the body mass; $Fr = \frac{\text{v}}{\sqrt{g \cdot h}}$ is the Froude number; v is a speed of motion.

The first elements of the equations (1) and (3) placed in brackets define work and power which are spent to support the body, the second elements (placed in braces) define work and power which are spent to move the leg. Numerical solutions of the equations (1-4) [5] showed that when $g = 9.8$ *m/s*^{2}, $h = 1$ *m*, $\mu = 0{,}2$ and v = 1.25 *m/s* (or 4.5 *km/h*), the ideal step length while walking is $L_{\text{опт}} = 0{,}7\,m$ if the step frequency is $f_{\text{опт}} = 105.3$ steps a min . Then, the power is $P_{\min} = 150\,W$.

It is noted that [5] the ideal step while walking equals approximately the step of cycling, i.e. it equals four-time the length of a bicycle crank. The length of a racing bike crank according to the existing standard equals 0.171 m (6.75") or 0.178 m (7.0"). Thus, the cycling step is about 0.684 or 0.712 m. If the power is the same (P=150 W) cycling with the frequency f=105.3 rpm can pick up speed equal to 24 km/h.

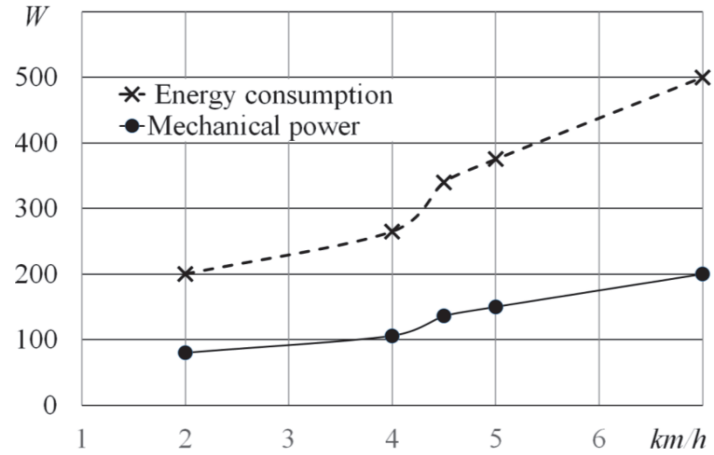


Figure 2 – Energy consumption and mechanical power of a person while walking

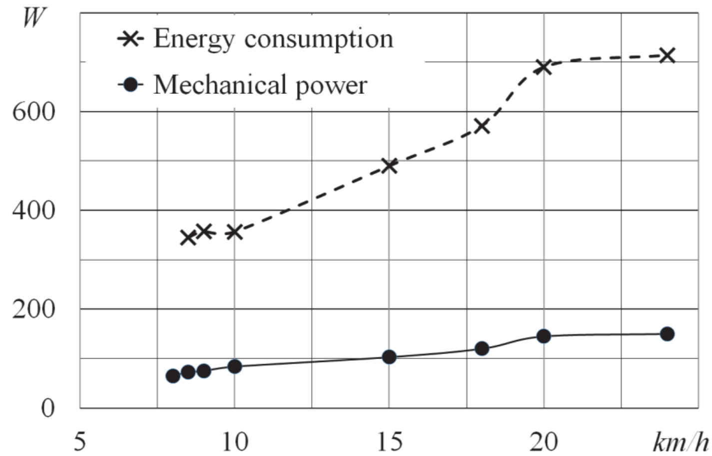


Figure 3 – Energy consumption and mechanical power of a person while cycling

Such a coincidence is not accidental. The logic is based on the intention of designers and sportsmen to provide maximum efficient functional motion. In the theory of walking and cycling the researchers and designers followed different ways but obtained the same results: ideal in terms of power walking step and cycling step are approximately equal if the functional motion frequency is the same.

Let us consider the dependences of energy consumption and gained power on the speed of motion for walking and cycling (see Fig. 2 and 3). It should be noted that the gained power is the same if you walk at a speed of 4.5 km/h or cycling at a speed of 24 km/h.

Based on the researches [3-6] and the results of our own experiments, we obtained the graphs which show the changes of cycle frequencies, torques in the hip and knee drives for walking, cycling and sitting down-standing up cycle (see Fig. 4-5) depending on the cycle phase T, % for various rates of motion and system mass. The graphs are built for a person whose weight is 75 kg and who walks at a speed of 4.5 km/h or cycling at a speed of 24 km/h.

Energy equivalence of the motions

The criteria of equivalence of the motions can be the rms torque [7] which occurs in a drive:

$$M_{экв} = \sqrt{\frac{\sum t_i M_i^2}{\sum t_i}}\text{, } Nm \tag{5}$$

equivalent power [7] in the drive

$$P_{экв} = \sqrt{\frac{\sum t_i P_i^2}{\sum t_i}}\text{, } W \tag{6}$$

and average rotation frequency between the links

$$n_{cp} = \frac{\sum t_i n_i}{\sum t_i}, \, rpm \tag{7}$$

Based on the analysis of the motion graphs and approximation of the obtained results, we got basic parameters of the motion for various modes (see Table 1).

Conclusions:

The analysis of the results allows us to make the following conclusions:

1. The comparison of the values of the basic locomotion parameters in various modes has given us some discrepancies of the highest equivalent torques, average interlinks rotation cycles and equivalent power in the joints.
2. As for energy consumption of the operation modes of an exoskeleton with a weight dummy inside, the locomotions (walking, standing up, sitting down, cycling) can be quite equivalent if the parameters are chosen in a right way.
3. The use of such a way of service life tests has the following advantages:

– a fixed site of the tests of an “exoskeleton-weight dummy” system; there is no need for much space; higher security of a fixed system.

– it is simpler to switch the exoskeleton to the measuring equipment.

– it is possible to perform the tests in the forced mode.

– it is possible to perform various loading modes in various time periods including those which are critical for an exoskeleton.

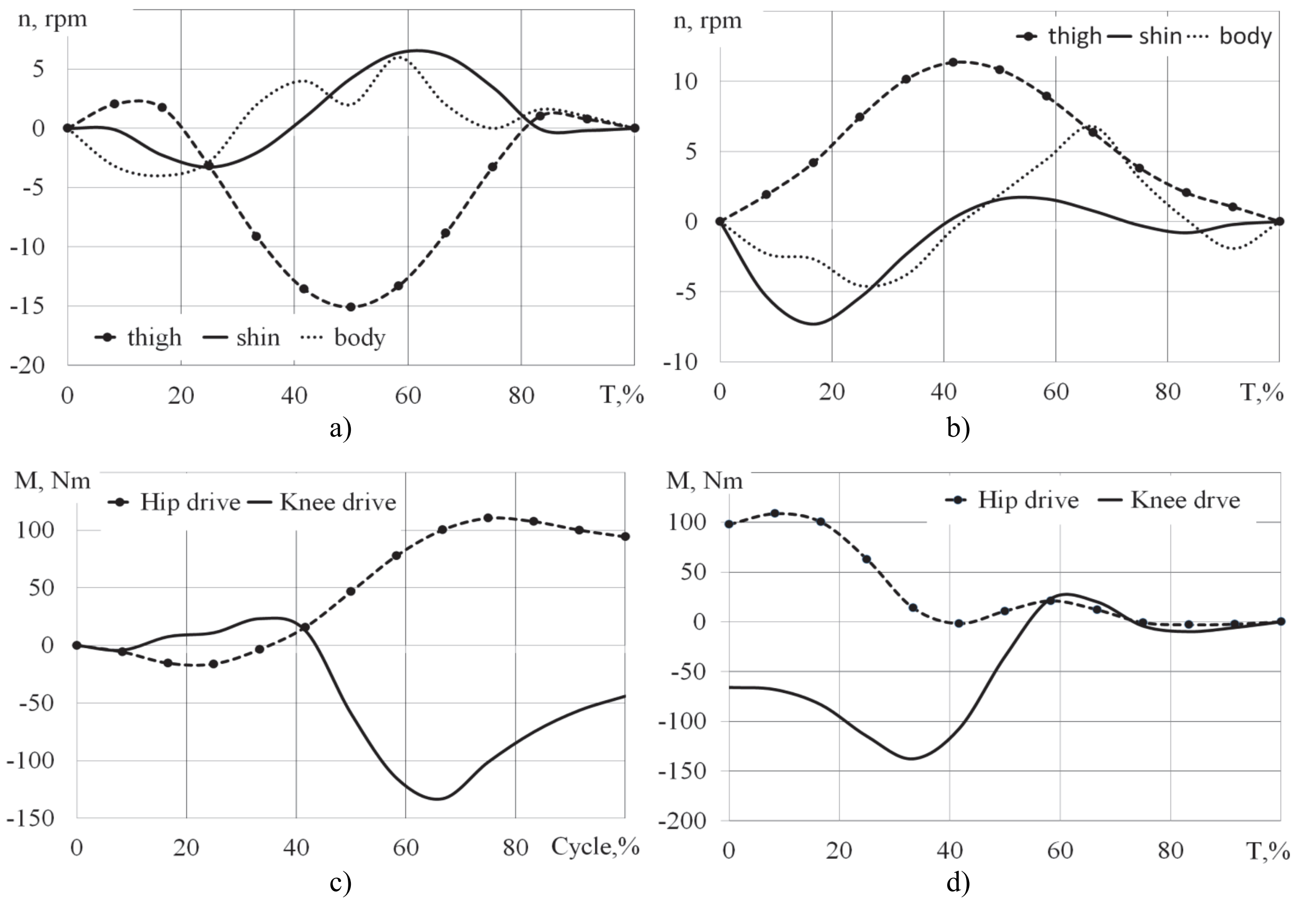


Figure 4 – Rotation frequencies and torques in “stranding up-sitting down” cycle:
a) rotation frequencies when standing up
b) rotation frequencies when sitting down
c) torques when standing up
d) torques when sitting down

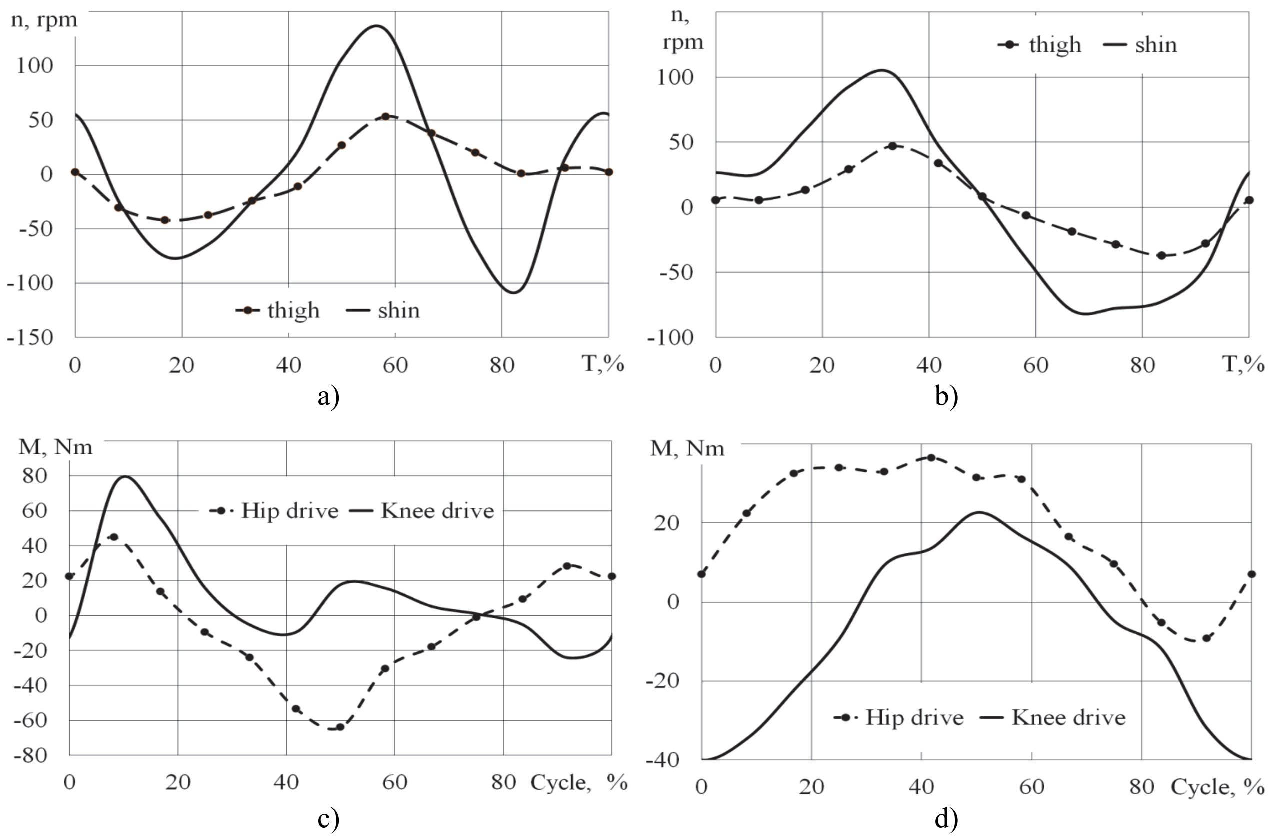


Figure 5 – Rotation frequencies and torques in locomotions:
a) rotation frequencies when walking along the straight line at a speed of 4.5 km/h
b) rotation frequencies when cycling at a speed of 24 km/h
c) torques when walking along the straight line at a speed of 4.5 km/h
d) torques when cycling at a speed of 24 km/h

Table 1. Comparison of various motions according to the chosen criteria

Type of motion	**Joint**	**Torque $M_{экв}$**	**Rotation frequency $n_{ср}$**	**Power $P_{экв}$**	
				in isolation	**in total**
		Nm	**rpm**	**W**	
Walking	Speed 4.5 km/h, rate 105.3 steps/min, step length 0.7 m, power $P_{operating}$ 150 W				
	hip	30.5	24.24	73.29	165.68
	knee	25.38	60.54	92.39	
	Speed 1.5 km/h, rate 65 steps/min, step length 0.7 m, power $P_{operating}$ 55.23 W				
	hip	30.5	8.08	24.43	55.23
	knee	25.38	20.18	30.8	
Standing up	hip	68.52	5.07	103.25	253.81
	knee	66.24	1.09	150.56	

<table>
<tr><th rowspan="3">Type of motion</th><th rowspan="3">Joint</th><th rowspan="2">Torque $M_{экв}$</th><th rowspan="2">Rotation frequency $n_{ср}$</th><th colspan="2">Power $P_{экв}$</th></tr>
<tr><th>in isolation</th><th>in total</th></tr>
<tr><th>Nm</th><th>rpm</th><th colspan="2">W</th></tr>
<tr><td rowspan="2">Sitting down</td><td>hip</td><td>50.24</td><td>5.67</td><td>37.13</td><td rowspan="2">152.00</td></tr>
<tr><td>knee</td><td>68.79</td><td>1.45</td><td>114.87</td></tr>
<tr><td rowspan="12">Cycling</td><td colspan="5">Speed 24 km/h, rate 105.3 rpm, step 0.684 – 0.712 m,
power $P_{operating}$ 150 W</td></tr>
<tr><td>hip</td><td>24.73</td><td>21.74</td><td>70.07</td><td rowspan="2">125.43</td></tr>
<tr><td>knee</td><td>20.99</td><td>56.43</td><td>55.36</td></tr>
<tr><td colspan="5">Mode 1 corresponding to walking at 1.5 km/h. Crank rotation frequency
39.20 rpm, power P_{output} 69 W</td></tr>
<tr><td>hip</td><td>30.55</td><td>8.09</td><td>86.55</td><td rowspan="2">68.96</td></tr>
<tr><td>knee</td><td>25.93</td><td>21.01</td><td>68.39</td></tr>
<tr><td colspan="5">Mode 2 for standing up-sitting down. Crank rotation frequency 27.5 rpm,
power P_{output} 129 W</td></tr>
<tr><td>hip</td><td>81.11</td><td>5.68</td><td>229.77</td><td rowspan="2">128.44</td></tr>
<tr><td>knee</td><td>68.84</td><td>14.74</td><td>181.56</td></tr>
</table>

– it is possible to model short-term overloads including those which occur during equipment misuse, control failures, emergency situations, huge overloads (when shifting the weight from one leg to the other, load occurred on the foot, etc) with a load coefficient equal to 1.2-1.5.

4. Based on the all mentioned above it is possible to use an exercise bike for service life tests of an exoskeleton.

References

1. Kapustin, A.V., Loskutov, Y.V., Skvortsov D.V., Nasibullin, A.R., Klyuzhev K.S., Kudryavtsev, A.I. Design models of rehabilitation exoskeleton control system // Annals of Volga state university of technology / Radio technology and IT systems. 2018. № 2 (38). P. 77-86. DOI: 10.15350/2306-2819.2018.2.77
2. Loskutov, Y., Kapustin, A., Kudryavtsev, A., Nasibullin, A., Lebedeva, A. (2018). Synthesisofexoskeletoncontrolalgorithmsbased on kinematic analysis of locomotions and human gait modelling. Journal of Applied Engineering Science, 16(4), 583-591. doi:10.5937/jaes16-17230
3. Winter, David A., Biomechanics and motor control of human movement / David A. Winter. – JOHN WILEY & SONS, INC., 2009 – 370 p.
4. Beletsky, V.V. Biped gait: model tasks of dynamics and control. – M.: Nauka, 1984. – 288 p.
5. Lyubovitsky, V.P. racing bicycles. – L.: Mashinostroenie. Leningradskoe otdelenie (Machine building. Leningrad Department), 1989. – 319 p.
6. Turlapov, R.N. Models and algorithms to control an exoskeleton's movements to keep a person in upright position and expand his opportunities // a PhD thesis manuscript. – Kursk: FSBEI of HE "Southwest state university". – 2015. – 172 p.
7. Masandilov, L.B., Sergievsky Y.N., Kozyrev S.K. Electric drive. Hydro and vibration drives // Machine building. 40-volume encyclopedia. Frolov K.V. (editor-in-chief). Volume IV-2. Book 1. Electric drive. – Moscow: Mashinostroenie (Machine building), 2012. – 520 p.

M.D. Solovyova

AN EXOSKELETON WITH A PARALLEL STRUCTURE FOR PATIENTS SUFFERING FROM MUSCLE-SKELETON DISORDERS OF LOWER EXTREMITIES

Volga State University of Technology, Yoshkar-Ola, Russia
masha_ru94@mail.ru

Abstract
The article deals with the analysis of the existing medical exoskeletons and their systems which provide stability. It also considers a solution to the problem of keeping balance while moving without support.
Keywords: exoskeleton, keeping balance, parallel structures, support tools.

Nowadays, exoskeletons have become widespread. An exoskeleton (Greek ἔξω "outer" and σκελετος "skeleton") is a device designed to enhance the force of a person using an outer frame. There are different kinds of exoskeletons: industrial or professional used by people to perform activity which involve great weights, military to enhance physical abilities of a person (stamina, speed, weight lifting ability), medical designed for rehabilitation and social adaptation of people suffering from muscle-skeleton disorders of lower extremities.

According to experts, the number of people who use medical exoskeletons exceeds 60 million including disabled people, elderly, children and people who got an injury and are in the process of their recovery [1].

The government of the Russian Federation has adopted the Strategy for industrial development of products aimed at rehabilitation for the period till 2025.

The development of medical exoskeletons is considered to be an innovative area. Robotic exoskeletons can make the everyday life of many musculoskeletal patients better [2]. The use of such exoskeletons for lower extremities as ReWalk, Ekso, Indego, ExoAtlet and REMotion allows to make a significant progress in the mobilization of people who had spinal cord traumas, strokes and suffer from muscle weakness [3]. Unlike exoskeletons designed to facilitate the movements or enhance physical abilities, a medical exoskeleton must fully replace human locomotions.

The rehabilitation is due to the anthropomorphic structures fixed on the legs and body to produce locomotions such as walking, sitting down, going up and down the stairs. Medical exoskeletons are required to have very precise and reliable nodes and drives which are able to bear the human weight.

One of the main tasks which the developers face nowadays is to provide stable position of a person while moving in an exoskeleton without using extra support tools such as crutches and walking frames.

Analysis of existing medical exoskeletons

Honda, a Japanese company, developed a device called WalkingAssistDevice (see Fig. 1) which helps people suffering from muscle weakness in lower extremities. The development began in 1999. The device weighs 2.5 kg. It is fixed around the waist and with the help of belts draws hip upwards and set the walking pace.

Figure 1 – WalkingAssistDevice exoskeleton

The device has a simple design, small number of drives and simple control system. All this resulted in a relatively low cost of the product. WalkingAssistDevice is an exoskeleton which includes hard belt, knee caps and bars which a located along the outer part of a hip.

The disadvantage of this device is that there are no drives in knee joints which minimizes its use for rehabilitation purposes.

Cyberdyne, a Japanese robotics and technology company, developed an experimentalexoskeleton called HAL (HybridAssistiveLimb). So far there have been developed two prototypes: HAL 3 (see Fig. 2) to restore leg mobility and HAL 5 used to restore the mobility of arms, legs and body. HAL exoskeletons are used in hospitals in Japan, Europe and the USA for rehabilitation of patients suffering from chronic diseases of nervous and muscular systems.

Servomotors are driven by electrical pulses produced by muscles and caught by the fixed on the skin electrodes. These pulses go to the built-in computer which analysis the load and triggers necessary servomotors of an exoskeleton.

A standard exoskeleton weighs 15 kg and is fixed to waist and legs. The run time (under the conditions of maximum load) is 2.5 hours. Its price is 4 200 US dollars. There is also a model for one leg only. The exoskeleton is aimed at enhancing the mobility of a patient. HAL 3 does not provide a vertical position of a patient.

Figure 2 – HAL 3 exoskeleton

Figure 3 – EksoBionics GT exoskeleton

EksoBionics, an American company, develops and produces smart exoskeletons for various purposes including industrial exoskeletons to perform hard work. EksoBionics GT (see Fig. 3) is a medical exoskeleton aimed at rehabilitation of patients suffering from muscle-skeleton disorders of lower extremities. It weighs 21.4 kg and can be used by patients whose weight is up to 100 kg. the price of such an exoskeleton is 100 000 US dollars.

Its design allows patients to use crutches and walking frames. One of the features of this exoskeleton is a patented weight load, it means that the structure of the exoskeleton is able to bear its own weight and keep itself in the vertical position without the load on a patient.

ParkerHannifin Company, USA, developed a robotic medical exoskeleton called Indego (see fig. 4). The weight of the device is only 12 kg. besides, it has a modular design and is easy to be taken apart. The exoskeleton has small dimensions, slender profile, and it is not equipped with a heavy backpack with batteries and a computer. All this allows a patient to move in a wheelchair wearing the exoskeleton.

Figure 4 – Indego exoskeleton

With the help of Indego exoskeleton it is possible to move along different surfaces including ramps and stairs. Special software provides smooth work repeating natural human locomotions and providing constant control of a body position. Also, Indego is the first wearable device which is equipped with effective and tested through practice treatment technology of functional electrical stimulation (FES). This technology enhances the muscle strength of patient suffering from paresis, improves the blood flow, prevents the loss of bone density and helps to avoid serious muscular atrophy.

With the help of Indego it is possible to stand up from the wheel chair and make some steps without help. The exoskeleton provides full support and helps to keep upright position of people who have a minimum control of their leg muscles. Patient who cannot walk at all can use crutches while moving in an exoskeleton. However, the price of the device is very high and reaches 80 thousand US dollars.

An American company of robotic and technology, SuitX, developed an exoskeleton called Phoenix (see Fig. 5). It is also designed for people suffering from muscle-skeleton disorders of lower extremities.

Phoenix exoskeleton consists of a hip joint module, two knee joints modules and leg modules. Patients are free to choose a module depending on their individual features. Besides, patients can put on and take off each component without help. Each component can be easily adjusted depending on different user parameters.

The exoskeleton weighs 7 kg and is considered to be the lightest exoskeleton in the world. Its price is about 40 thousand US dollars.

Figure 5 – Phoenix exoskeleton

Figure 6 – ReWalk 6.0 exoskeleton

Israel company ReWalkRobotics Ltd developed an exoskeleton called ReWalk (see. Fig. 6) which helps patients suffering from paralysis stand, walk and move up and down the stairs. The system is driven by a battery which is in a backpack. The weight of an exoskeleton is 23.3 kg, including the battery which weighs about 2.3 kg. The run time is three hours of continuous work. The exoskeleton allows to walk at a speed of up to 2.6 km/h. It is controlled by a remote device which looks like a watch. It is worn on the wrist and recognizes the movements of a patient. The system has three operation modes: walking, sitting, standing.

The price of the ReWalk 6.0 is about 77 thousand US dollars. The exoskeleton was subjected to clinical tests in a rehabilitation centre called MossRehab in Philadelphia, the USA. ReWalk was approved for the use in US hospitals by the US Food and Drug Administration (FDA) in 2011 [5]. In June 2014 the exoskeleton was approved by the FDA to be used at home and public places [6].

REX Bionix company from New Zealand created an exoskeleton called REX P (see Fig. 7) for patients suffering from the paralysis of lower extremities. The exoskeleton is able to stand up on its own, move along different surfaces, turn, and go up and down the stairs if the height of the footsteps does not exceed 18 cm without extra support tools. The power system can feed the exoskeleton for an hour. The control is done via a joystick or a tablet.

Figure 7 – REX P exoskeleton

The exoskeleton weighs 38 kg. Big weight of the device and its high price (about 150 thousand USD) makes it unaffordable for a large number of people [7].

ExoAtlet, a Russian company, in cooperation with Moscow State University and with the support of "Skolkovo" developed an exoskeleton for rehabilitation of unable to move patients. It is called ExoAtlet (see fig. 8). So far the exoskeleton has been testes and is available on the market. Its price is over 3.6 million rubles (57 thousand USD).

With the help of ExoAtlet patients get an opportunity to stand up, sit down, walk, go up and down the stairs on their own. The frame with sensors is fixed to the lower and upper parts of legs. A computerized control system and a battery, which can run for 6-8 hours in the motion mode, are placed behind. The whole device weighs about 20 kg. The speed of motion is 1 km/h. The balance when making a step as well as the control of the device is done with the help of crutches. The control buttons are on the crutch handles. The exoskeleton has an extra option which is muscle stimulation with the help of electric pulses synchronized with the motion of the exoskeleton [8].

Figure 8 – ExoAtlet exoskeleton

Figure 9 – REMotion exoskeleton

Since 2016 the researchers from Mechatronic systems laboratory, Volga state university of technology (Yoshkar-Ola), together with Volzhsky Electromechanical plant have been developing a medical exoskeleton called REMotion (see Fig. 9). The exoskeleton performs a very important task in rehabilitation of patients suffering from muscle-skeleton disorders of lower extremities. There have been developed four models of an exoskeleton: a basic model for rehabilitation of adults weighing up to 100 kg, a basic model with electromyography sensors, a basic model with electromyography sensors and functional electrical stimulation of muscles, and also a model for children [9].

The speed of motion is up to 2.8 km/h. REMotion exoskeleton has such features as fully modular structure, electromyography sensors and functional electrical stimulation of muscles, ability to create any walking algorithms.

The literature review shows that the task to provide balance stability while moving of a patient in an exoskeleton has not been fully solved and requires further research.

The aim of the present work is to develop and analyze an effective way to solve the task which would provide unassisted stability of an exoskeleton while performing locomotions.

Existing solutions to provide stability of an exoskeleton

The simplest solution for this problem is the use of extra support tools by a patient. Those include crutches, walking frames which a patient uses as a support while walking.

When using crutches (see Fig. 10) the task of balance support can be considered as a simple one. The support area expands significantly and it leads to a better balance. However, for locomotor patients it is difficult to choose a more stable position for a crutch. This causes a problem with balance support while using crutches.

The use of a walking frame (see Fig. 11) provides a better balance of a patient while moving in an exoskeleton. However, this system also has some drawbacks. For instance, when going up and down the stairs or along the rough surface, the walking frame cannot provide any stable support [10].

Figure 10 – Patient moving in an exoskeleton with the help of crutches

Figure 11 – Patient moving in an exoskeleton with the help of walking frame

These ways provide a good mobility for a patient but are not ideal solutions of the problem. When using extra support tools, the speed of walking drops. Moreover, a patient uses arms and shoulders to move the tools and keep the balance; it means he must be physically fit. Furthermore, even the use of crutches and walking frames means that a patient moving in an exoskeleton must have enough experience to keep balance. While training the assistance of a doctor or a physician is obligatory to stabilize the patient's position and providing some security means in case of emergency. They are also necessary to set the control commands in the exoskeleton control system. Those factors result in fatigue of a patient and his assistant.

Exoskeleton with parallel structures

There has been offered a new method to provide stability of a user while moving in an exoskeleton. It excludes the use of extra support tools such as crutches, walking frames, etc.

Design and mode of operation

Described technical result is achieved by the fact that an exoskeleton has extra right and left support devices which shift forward one-by-one when making a step in an exoskeleton. Each of these devices consists of two links connected with each other. The first link has an angular rotation gear in its base which is fixed on the exoskeleton on the hip joint level. The second link is located on the same level as a knee joint and has in its base a linear or angular rotation drive.

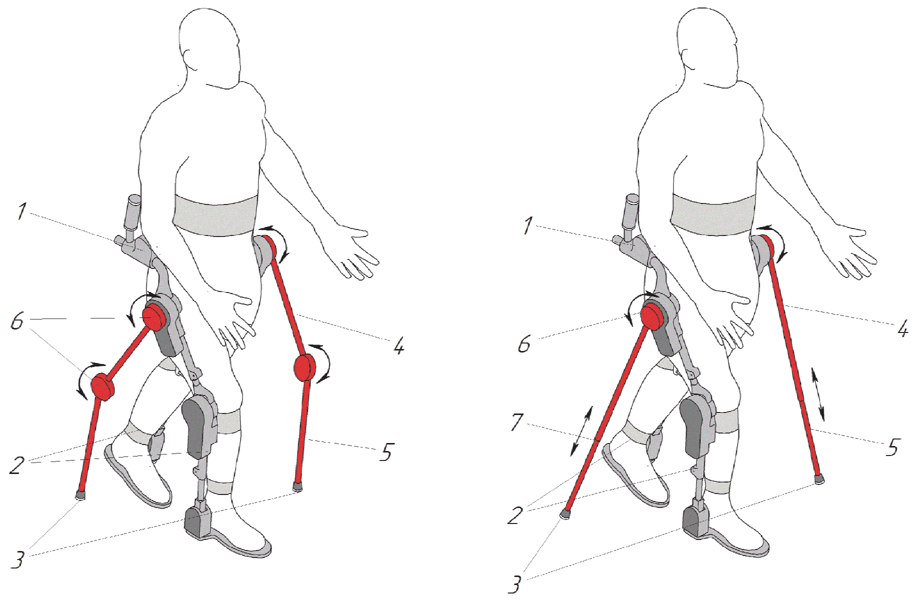


Figure 12– Diagram shows an exoskeleton with parallel structure: a) with an angular rotation drive in the base of the second segment; b) with a linear rotation drive in the base of the second segment

The shown method (see Fig. 12) includes an exoskeleton 1, which contains legs 2 and support devices 3. The support devices consist of the first link 4, second link 5 with angular rotation drives 6 and linear rotation drives 7 in the base.

An exoskeleton with a parallel structure operates in the following mode. When the exoskeleton makes the first step, support device 3 fixed on the left side of the hip joint moves at the same time too. The same motion is performed by the support device fixed on the right leg when making the left step. With the help of angular rotation drive 6 placed in the base of the first link 4, exoskeleton leg 2 performs a synchronous angular movement of the support device 3. With the help of angular rotation drive 6 or linear drive 7 placed in the base of the second link 5, support device 3 presses the surface.

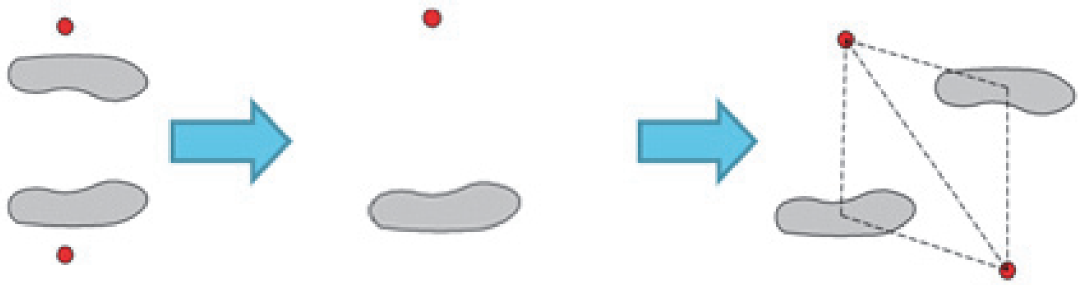


Figure 13 – Motion algorithm for an exoskeleton with parallel structure

The described motion algorithm of an exoskeleton with parallel structure (see Fig. 13) illustrates that the system has two support points as minimum when moving a leg while making a step and four support points during the pause before making the next step.

Kinematic analysis and explanation of energy efficiency of support devices

It is possible to provide a motion method for support devices if on the support part of the path, the drive 1 is locked (see Fig. 14). In this case the motion looks like a rolling of a polygon along the hard surface. The energy consumption is defined by the work required to lift the system when making each step [11].

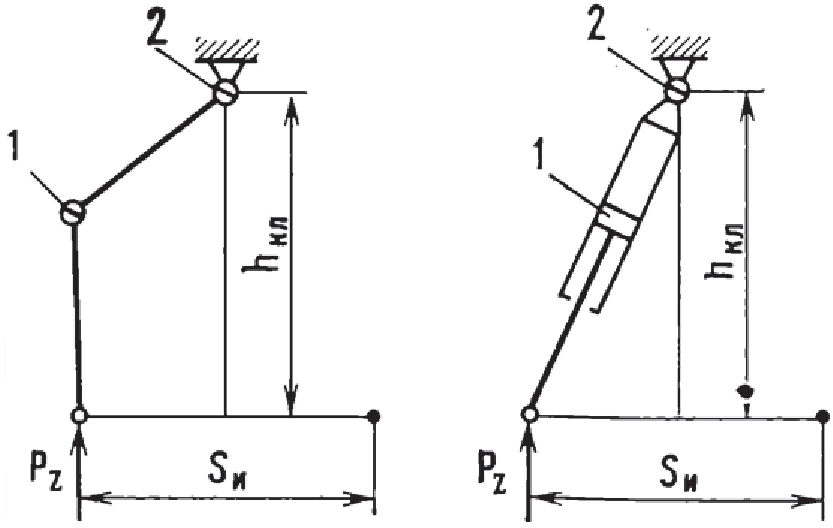


Figure 14 – Kinematic diagrams of support devices;
1 and 2 – actuation mechanism of a drive

In case if there are angular rotation drives used as actuation mechanisms (see Fig. 14), energy consumption depending on the pushing stage has a more complex equation:

$$f_0 = \frac{k_{пх} \cdot h_{кл}}{(1-s_{ю})}\left(\sqrt{0{,}25 + \frac{h_{кл}^2}{S_и^2}} - \frac{h_{кл}}{S_И}\right) \tag{1}$$

where $S_И$ is a step; $h_{кл}$ is a clearance of the mechanism; $s_{ю}$ is sliding (when moving along the straight line $s_{ю}$=0).

The gait coefficient is defined by the equation:

$$k_{пх} = \sum_{i=1}^{n_0} P_{zi} \cdot \frac{z_0}{n_0} \cdot G \tag{2}$$

where P_{zi} is a normal reaction of a support;z_0 is the number of supports; G is the weight of the system.

Energy consumption if there is one linear drive in the base of the actuation mechanism (see Fig. 14b) has the same physical meaning and can be expressed by the equation:

$$f_0 = \frac{k_{пх} \cdot h_{кл}}{(1-s_{ю})S_и} \ln\left(1 + \frac{S_и^2}{4 \cdot h_{кл}^2}\right) \tag{3}$$

Such pushing path can be considered as energy efficient because potential energy of mass centre lifting converts into kinematic energy of translator motion of the support devices.

How to prevent turning over of the "human-exoskeleton" system

When using an exoskeleton there can occur emergency situations when the "human-exoskeleton" system loses its balance, for example, when there is some rough surface or obstacles. The described method can help prevent falling by putting forward both support devices at the same time in saggital plane in the turning-over direction. After that they push it back to restore the balance (see Fig. 15).

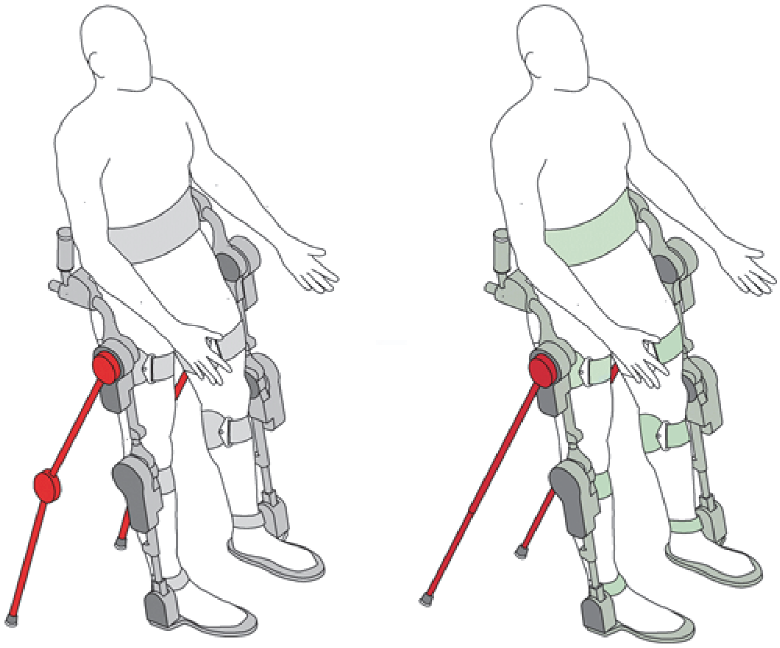


Figure 15 – Prevention of turning over of the “human-exoskeleton” system with the help of support devices

The advantages of an exoskeleton with parallel structures

The describes parallel structure of an exoskeleton has some advantages compared with the classical approaches used to solve the set task: it improves the balance of an exoskeleton; it eliminates the use of extra support tools when a patient moves in an exoskeleton; it makes the motion of a patient safer as it uses the support devices to prevent turning over; it reduces the pressure on the basic drives of an exoskeleton due to the pushing path of the support devices; it reduces the load on the patient and medical assistant.

References

1. Recommendation of the Disability Services Advisory Council in the Council of the Federation of the Federal Assembly of the RF [electronic source] // Official website of the Council of the Federation – Access mode: http://council.gov.ru/media/files/fDqG0zASJNAALmOjQoiE96H3APp4ivlh.pdf
2. Biomechatronic complex of neurorehabilitation – concept, structure, models and management / V. E. Pavlovsky [and others.] // Preprinty IPM named after M.V. Keldysh. 2014. № 111. 19 p.
3. Stability of Mina v2 for Robot-Assisted Balance and Locomotion [Электронныйресурс] / C. Mummolo, W. Peng, Sh. Agarwal, R. Griffin, Peter D. Neuhaus, J. H. Kim // Frontiers in Neurorobotics. October 2018. V. 12, A. 62. 16 p. DOI: 10.3389/fnbot.2018.00062. – Access mode: URL: https://www.researchgate.net/publication/328288680_Stability_of_Mina_v2_for_Robot-Assisted_Balance_and_Locomotion (access date 15.04.2019).
4. Moreno J., Turowska E., Wearable Lower Limb and Full-Body Robots // ArantesWearable Robots: Biomechatronic Exoskeletons , 2008, pp. 283–321.
5. Bionic legs get FDA approval, NewsComAu. [electronic source] – Access mode: URL: https://www.news.com.au/technology/rewalk-bionic-legs-get-fda-approval/news-story/60aad387a54a6897e758e1c341560220 (access date 29.10.2018).
6. FDA approves Israeli ReWalk robotic exoskeleton technology | Israel Company Reports - Search Israeli Companies. [electronic source] – Access mode: URL: www.israelbizreg.com. (access date: 29.04.2019).
7. Meghan Rosen Mind to motion: Brain-computer interfaces promise new freedom for the paralyzed and immobile// Science News Volume 184, Issue 10, 16 November 2013, pp. 22–24.
8. For clinics [electronic source] // ExoAtlet is a Russian exoskeleton – Access mode: URL: https://www.exoatlet.com/for-clinics (access date 03.05.2019)
9. A medical exoskeleton is getting ready for acceptance tests [electronic source] // Volga state university of technology – Access mode: URL: https://www.volgatech.net/news/Novosti_nauki/394563/ access date: 01.05.2019
10. Kapustin, A.V., Loskutov Y.V.,Kudryavtsev I.A., Belogusev V.N., Ways to provide stable position of a medical exoskeleton designed for rehabilitation purposes while walking // Bulletin of Volga state university of technology. Section: Materials. Constructions. Technology. 2018. № 3 (8). P. 44-54.
11. Planet research vehicles / A.L. Kemurdzhian. – M.: Mashinostroenie (Machine building). 1982. – 319 p.

A.A. Meldo [1,2], L.V. Utkin [2]

RADIOMICS AND THE MULTIDISCIPLINARY APPROACH IN THE DEVELOPMENT OF CAD SYSTEM IN LUNG CANCER DIAGNOSTICS

[1] *Saint-Petersburg Clinical Research Center of specialized types of medical care (oncological), Saint-Petersburg, Russia, anna.meldo@yandex.ru*
[2] *Peter the Great Saint-Petersburg Polytechnic University, Saint-Petersburg, Russia, lev.utkin@mail.ru*

Abstract

The interpretation of visual information obtained through various modalities is a reflection of properties of the pathology analyzed by the doctor. From the point of oncological view, this is not only reflection of the phenotypic characteristics of the tumor, but also a certain system of the disease staging, and predicting of the disease, as well as planning a particular treatment tactics. Standardization of data interpreted by a doctor and their transformation into a quantitative format are necessary for reducing subjectivity in making decisions. The main idea of Radiomics is extraction quantitative features from diagnostic images, which can then be supplemented by other information, including anamnestic and clinical data of patients. Most of technologies of the artificial intelligence systems development in medicine are directly related by this approach. The article reflects the concept of radiomics in the development of CAD system of diagnostics of lung cancer, which are a new feature extraction of peripheral carcinomas, not previously used. The article also reflects the role of a radiomics in the transformation of radiology. All this contributes not only for the technical development of diagnostics in general, but also for a new approach of the organizational and educational process.

Acknowledgments

The project is supported by grant of Russian scientific Foundation (project No. 18-11-00078).

Radiomics: current state

The concept of Radiomics involves technologies and methods of converting the digital medical images in quantitative data for establishing of viewing biomarkers. The main idea of radiomics is that some features of visualization can be useful for predicting the course of diseases [1]. Before the description of the issue of radiomics we would like to describe the it's prototypes. In oncology, as in most other medical specialties, the information generated by radiation modalities, is usually limited to the visual interpretation of the phenotypic characteristics of the tumor such as shape, size, contours, structure, the contrast enchacement, etc., in a word assessment of tumor characteristics and properties [2]. The role of radiation imaging methods is due to the fact that more than 90% of patients with cancer are evaluated by these methods [3]. Image techniques such as ultrasound diagnostic (USD), computed tomography (CT), magnetic resonance imaging (MRI) shows the anatomical changes in organs and systems. Another group of image methods, such as functional MRI, positron emission tomography (PET) and single-photon emission computed tomography (SPECT), as well as hybrid methods (PET-CT, PET-MRI) are used to assess the metabolic activity of the tumor [4, 5].

The traditional activities of the radiologist are the choice of scanning Protocol, image acquisition control, making radiological reports [6]. Image phenotypes are often separated according to specific classifications, which helps to separate patients into subgroups or subpopulations for treatment selection or prognosis evaluation. Thus RADS criteria developed by the American x-ray society, are accepted for tumor evaluation systems, they include some standardized categories for characterizing the probability of tumor malignancy: TI-RADS for the evaluation of thyroid tumors [7], BI-RADS – for the characterization and description of formations in the breast [8], PI-RADS similarly for the diagnosis of prostate cancer [9], LI-RADS for the evaluation of hepatocellular cancer [10]. Lung-RDS is a system of standardized description of lung nodules in the CT, used in screening, it is also applicable for the description of lung nodules in first examined patient. It is a fact that 8 – 51% of patients from the high-risk group for lung cancer have at least one focus in the lung [11, 12], and 95% of them are usually benign [13,14]. Clear criteria of the malignancy characteristics determine the follow up or the decision about the operation. So RADS-criteria are prototypes of radiomics in aspect of transformation visual features into quantitative ones which can be used in making decisions. Another prototype of radiomics is size coefficient proposed by Japanese scientists in 2003 which is the ratio of the axial maximum size to the maximum vertical size of the object. The large value of this parameter indicates that the formation is flat and benign [15, 16, 17]. That is, attempts to formalize the data of radiation studies and convert them into a quantitative format were made before the beginning of active development of mashine learning in medicine. This is necessary to reduce the subjectivity in the interpretation of the pathology by the doctor. Technologies of the development of AI systems in medicine involve solving some tasks related to the extraction of quantitative features from diagnostic images, which can then be supplemented by other data, including anamnestic and clinical information of patients. This is the paradigm of radiomics [2]. It means that the radiologist chooses most significant signs for making decision, which can be rendered quantitatively. Thus, radiologist can improve his skills by choosing radiomical criteria of the pathology for feature engineering by data-scientist. Therefore, the radiologist becomes the participant of CAD development for it's quality improvement [6].

In the paper [18] quantitative image characteristics relating to the size/shape, a volume histogram of the intensity (function of the first order); the functions of texture functions (second order) and the function transformation analysis presents as a radiomical features in the CAD of prostate cancer.

In some cases, the scan parameters can be different, however, the quantitative expression of features in radiomical analysis can standardize and average the information obtained, improve the quality of the developed CAD through the unification of classification algorithms [19].

Thus, radiomics is not only the method of creating CAD. In our opinion the selection of significant signs for CSD development may be limited without the participation of a doctor with his clinical and radiological knowledge. As a result, the quality of the developed system will not be satisfactory. This aspect leads to the creating and development of the new methodology in the extraction of diagnostic criteria of diseases for the effective using of AI in medicine.

Multidisciplinary collaboration in the development of CAD Doctor AIzimov

Currently, most of the developed CAD are based on the principle of data processing without taking into account clinical and radiological classifications adopted in medicine. But in this case, even with a good result, there is no acceptance of AI ideology by the medical community because of it's conservatism. Therefore, the approach of the AI algorithm to the "logic of the doctor" has the greatest prospects for its development and implementation in practice. In the development of CAD radiologist needs additional skills and basic knowledge of machine learning in contrast to traditional activities. So, activities of radiologist become close to data-scientist, and he becomes the main person in the interaction of "AI-doctor".

CAD Doctor AIzimov was developed by laboratory of neural network technologies and artificial intelligence as a part of collaboration between Saint-Petersburg clinical research center of specialized types of medical care (oncological) and Saint-Petersburg Peter the Great Polytechnic university. This laboratory integrates activities of data-scientists of Polytechnic university and radiologists of oncocenter. Main steps and the part of radiologist in development CAD system are described below:

1) Data collection. The data must be homogeneous for processing. In the developed system, CT-scans with a slice thickness of not more than 2.5 cm were included in the database. Despite the fact that contrast enhancement was not the inclusion criterion, most studies were performed with contrast infusion (in the arterial phase). The heterogeneity of the data can lead to errors in classification.

2) Analysis and preliminary conclusion. The traditional function of radiologist was expanded – after making radiological protocol data was saved on the hospital server and special registry pending the morphology results for creating the dataset.

3) Control of morphological confirmation. To minimize inaccuracies, all cases used in the database were morphologically verified. If surgical treatment was absent, confirmation of the disease was based on clinical, laboratory or other data (for example, with dynamic observation, the lack of growth of the nodulus for 6 months indicates benign nodulus). After confirmation, the data were entered into the database and the register of confirmed cases.

4) Marking of data. The marking can be made by different approaches in different radiological modalities. It depend on the task. As an example the marking of pathological objects was made with special made software MAIA (Medical Artificial Intelligence Assistant). It is characterized by the use of multiplanar reconstructions of CT series, allowing to outline the pathology 2-10 times faster (depending on the size of the nodule) [21]. The figure 1 shows the example of outlining of lung nodules with MAIA software. It is realized in Python language. The graphical shell is implemented using the PyQT library, as well as standard image processing tools. To process images of computed tomography pydicom module was used, for three-dimensional visualization of the program – VTK (visual tool kit).

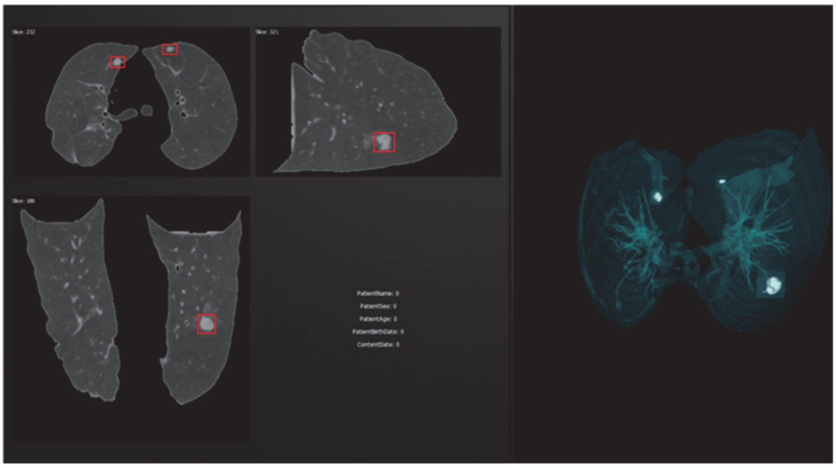


Figure 1 – A screenshot of the program, delineation of pathology in the lung using multi-planar reconstructions of a series of CT scans of MAIA (Medical Artificial Intelligence Assistant)

5) Feature selection. In accordance with the principles of interpretation of the pathology by the radiologist main signs of lung cancer were chosen for machine learning. They are the shape of the object, the internal structure of the node, the external structure (perifocal lung tissue). It is a fact that malignant tumors are more characterized by a form approaching the spherical, fuzzy, uneven contours, spicules caused by local lymphangitis, as well as an inhomogeneous structure associated with central necrosis. And vice versa pathognomonic signs for benign tumors are smooth clear contours and the calcium in the structure. After feature selection by the radiologist next step is feature extraction and feature engineering for machine learning. The main idea of multidisciplinary approach is in a constant collaboration between the data-scientist and the radiologist for feature selection and feature extraction correction.

6) The development of the dataset.

New approaches for the feature representation in the implementation of CAD Doctor AIzimov

As said above the "doctor's logic" in interpretation of chest CT of the patient with lung cancer is based on such phenotypical signs as a shape, margins and structure of the nodule and the structure of perifocal lung tissue. For the feature representation according to the shape criteria we used the method of cords [20]. The idea of this approach is in the construction of a set of segments (chords) which connect random points on the surface of the object. The lengths of the chords are normalized by the longest of them and a histogram of the lengths of the chords is constructed. The histogram is invariant to the size of the tumor, its movement or rotation, as well as stable to noise and small distortions. Therefore, it is a new representation of signs of a smaller dimension of the tumor for further classification (Fig.2). The number of chords and the size of the histogram are adjustable parameters of the classification system.

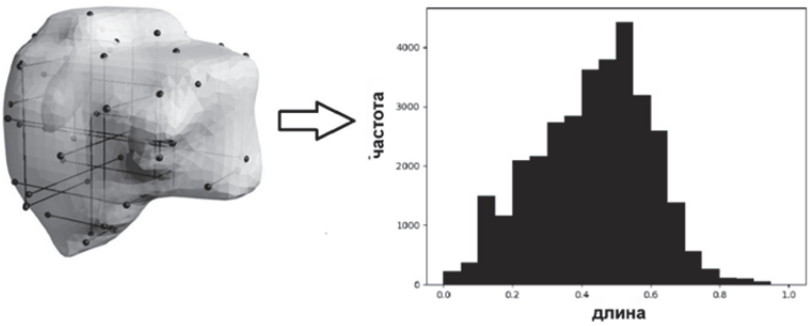


Figure 2 – The model of the nodule with a chords and the histogram of lengths of chord which is characterized the shape of the object

To confirm and clarify the form feature, we used following method: a virtual cube was built around the object, the size of one was greater than the maximum size of the tumor. Next, the chords connecting the surfaces of the cube and the points on the surface of the tumor were constructed, as shown in figure 3. These chords are perpendicular to the faces of the cube.

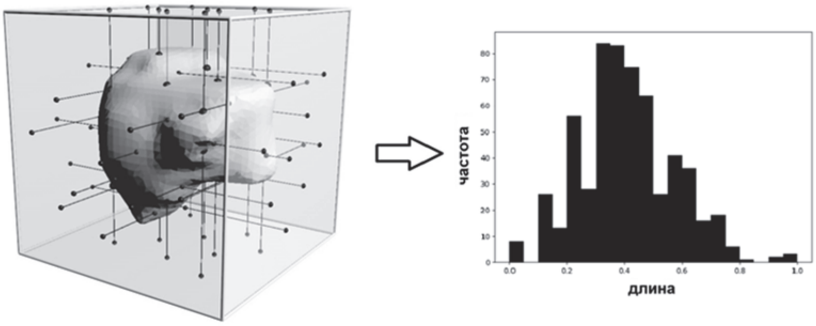


Figure 3 – The model of the nodule with virtual cube and chords which connect faces of the cube and the surface of the object, and the histogram of external chords.

Thus, the first and second histograms are characterize the surface of the tumor.

The histogram of densities is a new representation of features by the criterion of the internal structure of the object. To express this feature on each chord randomly selected points, which are determined by the density to build a histogram of densities. Example of neoplasm and corresponding histogram of densities are shown in Fig. 4.

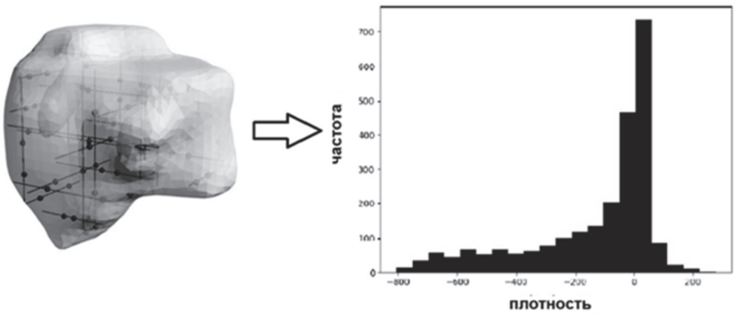


Figure 4 – The model of the node with points for measuring densities on the chords and the histogram of densities which characterize the internal structure of the object

Despite the fact that the third histogram characterizes the internal structure of the tumor, it reflects some average values of densities without taking into account their changes. The internal structure of the tumor is usually not homogeneous, it is characterized by the central zone of necrosis, and therefore, there is a central-peripheral gradient of the density of formation. To represent the characteristic of the central-peripheral density gradient, we propose to use a histogram which characterizes the change in density from the center of the tumor to its surface. To build this histogram, the approximate center of the tumor is determined and spheres with an increasing radius from the center of the tumor are built. The last sphere covers the most distant points of the tumor. Radii from the center to the surface are randomly constructed and density values are determined at the points of intersection of radii with spheres. The differences between the densities at the neighboring points of each radius (for neighboring spheres) are the basis for the histogram. Thus, the histogram contains the distribution of density differences, as shown in Fig. 5.

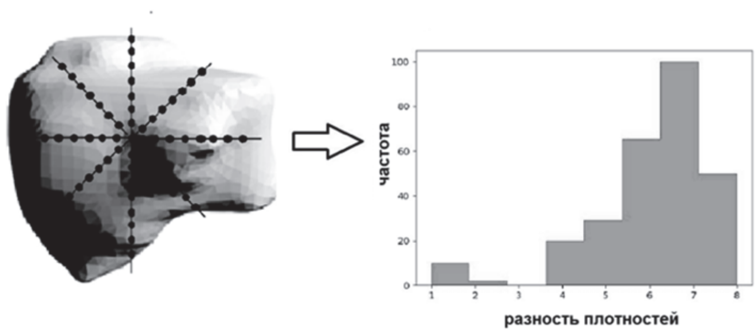


Figure 5 – The model of the node and the distribution of densities in accordance with the central-peripheral gradient, histogram which reflect the difference of densities

In some cases, the perifocal lung tissue characterizes the one or the other disease. As an example, bronchogenic perifocal lesions are characterize the tuberculosis, radial spicules which are demonstrate the local lymphangitis are lung cancer signs, in a benign tumor case the perifocal lung tissue stay normal. For the represent features which characterize perifocal lung tissue we used histogram of external densities based on the model depicted in Fig.2. This is a histogram of densities on the outer chords (Fig. 6). Objects between the virtual cube and the node can affect the change in the histogram.

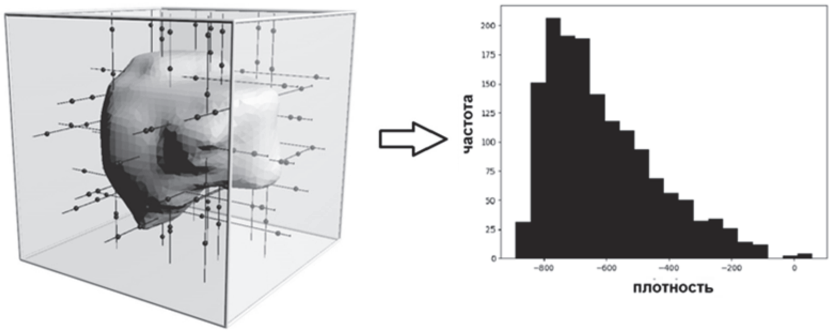


Figure 6 – A model of nodal formation enclosed in a virtual cube with constructed external chords and points for measuring of densities, and a histogram of external densities which is a representation of the external environment of the tumor (perifocal lung tissue)

Thus, we reduce the dimension of analyzed data by representing the features in the form of five histograms. This approach has some advantages: histograms fully reflect the main signs of lung cancer, which are analyzed by the doctor. They can be interpreted for explaining the classification results. Using the feature selection procedure, can determine which features are more significant (densities, and their gradients, size, form etc.), it can show how each of them affects the classification results of a particular tumor, which combinations of features have the greatest impact on the results. In addition, simple classifiers other than deep neural networks, such as random forest, support vector method, AdaBoost, etc., can be used to classify histograms due to their small dimension.

The ideology of creating of the database

The process of database making is also the result of interaction between the doctor and the data-scientist. The role of the radiologist is to create the ideology of the dataset architecture, the collection of confirmed material, the data-scientist makes a structure, software, forms learning and testing subsets for machine learning. In the project Doctor Azimov we used as open databases LIDC-IDRI, LUNA-16, and our own replenished database LIRA (Lung Image Resource Annotated). The analysis of public data sets showed that the morphological verification of diseases will minimize false positive results. LIRA includes more than 450 fully anonymized chest CT studies with a slice thickness of no more than 2.5 mm. RadiAnt DICOM Viewer have been chosen for the preparation of the dataset. Archiving of the prepared data was carried out on the server of the oncological center before the morphological diagnosis. Then data was anonymized with DicomCleaner™, that is the identification of CT-scans was changed by rename them with special code in according with internal hospital register. One more peculiarity of LIRA is it's structure – dataset is separated in subsets in according with the type of morphologically confirmed lung pathology. The 3 years trial of lung cancer visualization showed that not all cases have clear pathognomonic signs. It is about 9% cases which are difficult to recognize because of the similarity of their CT-portrait with other diseases. So we decided to separate our dataset into 3 subsets: "typical lung cancer" – is the subset of cases which were interpreted as a cancer by 3 radioligists and were morphologically confirmed as a cancer as well; "not typical lung cancer" – the subset of cases which are similar with other diseases on CT scans, were interpreted as "not cancer" at least by 1 from 3 radiologists and morphologically identified as lung cancer; "not lung cancer" – subset of cases which are not lung cancer by morphology.

Due to the necessity and relevance improvement of the machine learning algorithm to differential diagnosis of lung cancer, the LIRA data were supplemented with the marking of each case in accordance with the international classification of diseases (ICD). The improved LIRA2019 database includes 300 cases of lung cancer, 60 cases without pathology of the pulmonary system and 92 cases with other diseases which are similar to lung cancer - from inflammatory changes and benign tumors to metastases of other localization tumors. There are 70 cases in data base LIRA which are not morphologically verified. That CT of patients who refused surgery or biopsy was impossible due to contraindications. That data was included in the training sample after conclusion of 3 radiologists with the mark “not confirmed”. In order to take into account unverified data, we apply an approach which is based on the fact that the weight of each feature vector representing training instance in the loss function is replaced with a new weight. In particular, if the corresponding instance is from the a verified part of the training set, then its weight becomes equal to

$$w1 = 1 / (N [1-p + qp]),$$

where w is the replaced weight, N is the number of cases in the training set, p is the fraction of the unverified cases, q is the probability of the doctor’s correct decisions.

If the instance is unverified, then its weight becomes equal to

$$w2 = q / (N [1-p + qp]).$$

To calculate q, we analyzed the dataset analysis results provided by five independent radiologists who evaluated a part of the verified sample in accordance with the criterion “cancer” / “not cancer”.

Thus, in conclusion: multidisciplinary collaboration results in the new quality of developing CADs. In one hand – increasing of functional tasks of radiologist make machine learning algorithms be closer to doctor’s logic that is, it allows to develop an explainable AI system. An explainability of AI system is an important condition for using it in medical practice. In other hand this approach take developer the ability to make CAD on the base of clinical and radiological classifications. Besides this way makes radiologist to improve his professional level and skills, because the control of morphological verification and data collection enhance the feedback between the final diagnosis and preliminary conclusion. Implementation of an extended set of tasks of radiologist is possible only when he has basic machine learning and AI knowledge.

References

1. Yip, Stephen S F; Aerts, Hugo J W L (7 July 2016). Applications and limitations of radiomics // Physics in Medicine and Biology, 61(13):150–166, 2016.
2. Ognerubov N.A., Shatov N.A., Shatov A.V. Radiogenomics and radiomics in diagnostics of malignant tumors: review // Vestnik TGU, 22(6):1-9, 2017 (in Rus.).
3. Goyen M. Radiogenomic imaging-linking diagnostic imaging and molecular diagnostics // World. J. Radiol, 6(8):519-522, 2014.
4. O’Connor J.P.B., Rose C.J., Waterton J.C., Carano R.A., Parker G.J., Jackson A. Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome // Clin. Cancer Res, 21(2):249-257, 2015.
5. O’Connor J.P.B, Aboagye E.O., Adams J.E., Aerts H.J. et al. Imaging biomarker roadmap for cancer studies // Nat. Rev. Clin. Oncol, 14(3):169-186, 2017.
6. Meldo A.A., Utkin L.V. Radiomics as a basis for transformation of radiologists skills and partnership. International Conference "Emerging Trends in Applied and Computational Physics 2019" (ETACP-2019), Saint-Petersburg, Russian Federation, 21–22 March 2019, URL: https://iopscience.iop.org/journal/1742-6596.
7. Timofeeva L.A., Alyoshina T.N. TI-RADS system application in differential diagnostics of thyroid cancer // Kazansky meditsinsky jurnal. 98(4):632 – 636, 2017 (in Rus.).
8. Hamy A.S., Giacchetti S., Albiter M. et al. BI-RADS categorisation of 2,708 consecutive nonpalpable breast lesions in patients referred to a dedicated breast care unit. Eur.Radiol. 22(1): 9–17, 2012.
9. Mishchenko A.V., Rubtsova A.V., Alekseev B.Ya., Petrov S.B., Belyaev A.M., Kaprin A.D. The unified approach to interpretation of MRI of the prostate under the guide PI-RADSv2 // Oncourology. 12(1):81 – 89, 2016 (in Rus.).
10. Tumanov U. N., Karmazanovskii G. G., Shchegolev A. I. the System of LI-RADS with computed tomographic diagnosis gepatocelullar cancer // Medical imaging. 6:44, 2014 (in Rus.).
11. Aberle D., Duan F., Apgar C.K. et al. Comments on National Coverage Analysis (NCA) for Lung Cancer Screening with Low Dose Computed Tomography (CAG-00439N) Provided by members of the National Lung Screening Trial Research Team. Available at: Accessed October 13, 2014.
12. Khoruzhik S. A., Bogushevich, E. V., Sprindzhuk M. V., Snezhko E. V., Kovalev V. A., A. V. Tuzikov Computer-assisted diagnosis of nodules in the lungs. Oncology issues. 57(1):25-35, 2011 (in Rus.).
13. Tyurin I. E. Single foci in the lungs: criteria for differential diagnosis. Russian electronic journal of radiology. 3(3):50-52, 2013 (in Rus.).
14. Brandman S., Ko J.P. Pulmonary nodule detection, characterization, and management with multidetector computed tomography. J. Thorac. Imaging. 26:90–105, 2011.

15. Takashima S., Sone S., Li F. Maruyama Y., Hasegawa M., Kadoya M. Indeterminate Solitary Pulmonary Nodules Revealed at Population-Based CT Screening of the Lung: Using First FollowUp Diagnostic CT to Differentiate Benign and Malignant Lesions. AJR. 180:1255-1263, 2003.
16. Takashima S., Sone S., Li F., Maruyama Y., Hasegawa M., Matsushita T., Takayama F., Kadoya M. Small Solitary Pulmonary Nodules (1 cm) Detected at Population-Based CT Screening for Lung Cancer: Reliable High-Resolution CT Features of Benign Lesions. AJR. 180:955-964, 2003.
17. Moiseenko V. M., Meldo A. A., Utkin L. V., Prokhorov I. Yu., Ryabinin M. A., Bogdanov A. A. Automated system of detection of volume formations in the lungs as a stage of development of artificial intelligence in the diagnosis of lung cancer. Radiation diagnosis and therapy. 3:62-68, 2018, URL: https://doi.org/10.22328/2079-5343-2018-9-3-62-68 (in Rus.)
18. Lambin P., Rios-Velazquez E., Leijenaar R. et al. Radiomics: extracting more information from medical images using advanced feature analysis // Eur. J. Cancer. 2012. V. 48. No 4. P. 441-446. ; Kumar V., Gu Y., Basu S. et al. Radiomics: the process and the challenges. Magn. Reson. Imaging. 30(9):1234-1248, 2012.
19. Gillies, Robert J; Kinahan, Paul E; Hricak, Hedvig (2016). "Radiomics: Images Are More than Pictures, They Are Data". Radiology. 278 (2):563–577, 2015, URL: doi:10.1148/radiol.2015151169.
20. Smith S.P., Jain A.K. Chord distribution for shape matching // Computer vision, graphics, and image processing, 20(3):259-271, 1982.
21. Meldo A., Utkin L . Intelligent Data Processing: Theory and Applications. Book of abstracts of the 12th International Conference (Moscow: TORUS PRESS). 35, 2018, URL: http://mmro.ru/en/2018/05/01/2018-idp-12.

S.A. Matyunin

FIBER-OPTIC SENSORS FOR ANTHROPOMORPHIC ROBOT GRIPPERS

Samara National Research University, Department of Power Plant Automatic Systems, Samara, Russia
S.A.Matyunin@yandex.ru

Abstract

Currently, technologically advanced countries are conducting intensive research in the development of new types of sensors, especially for robotics. Fiber-optic sensors (FOS) with the closed optical channel with insensitive to electromagnetic pickups, operating from cryogenic (minus 200 ºC) to high temperatures (+400 ºC) and in a principally explosion-proof design are of particular interest [1-6].

Samara University has developed theory and made a report that discussed the features of the implementation of Tactile Force (FOS-TF) and the angular position (FOS-AP) of phalange grips of an anthropomorphic robot based on optical fiber macro bends [6-10].

Experimental studies of Tactile Force FOS prototypes and the angular position of the phalanges of the grips prototypes were carried out (with the participation of "Android Technique", Moscow) and the following characteristics were achieved [10]:

– controlled tactile force is not less than 0 ... 10 N;
– controlled angular positions of phalanges are 0 ... 60 degrees;
– the basic error of tactile force control is not worse than 1.0%;
– the main error in the control of the angular position is not worse than 0.5%;
– size of the contact patch of tactile force is at least 3x3 mm;
– resolution of the electronic transceiver is at least 10 bits;
– operating temperature is from minus 80 to plus 80 °C;
– relative air humidity is up to 100%;
– supply voltage of an electronic transceiver is 22-32 V, 0.1 A.

Keywords: anthropomorphic robot grip, fiber optic tactile force sensor, fiber optic angle position sensor, electronic transceiver, mathematical model, experimental studies.

Acknowledgments

This work was supported by the Ministry of Education and Science of the Russian Federation. Unique identifier: Applied research and experimental work RFMEF157816X0209.

Introduction

Modern systems for collecting and transmitting of measurement and control information based on fiber-optic sensors are the most rational elements for receiving and transmitting due to their high resistance to electromagnetic interference, destructive factors (chemical, radiation, temperature, etc.). They also provide the possibility of ultrafast transfer of measurement results and control commands. According to numerous estimates [1-6] in the next 10-15 years, such FOSs will cover 80 ... 90% of the robotics needs, and will also be actively introduced in such areas as nuclear power, hazardous chemical and oil, and gas processing industries, aerospace instrumentation, etc.

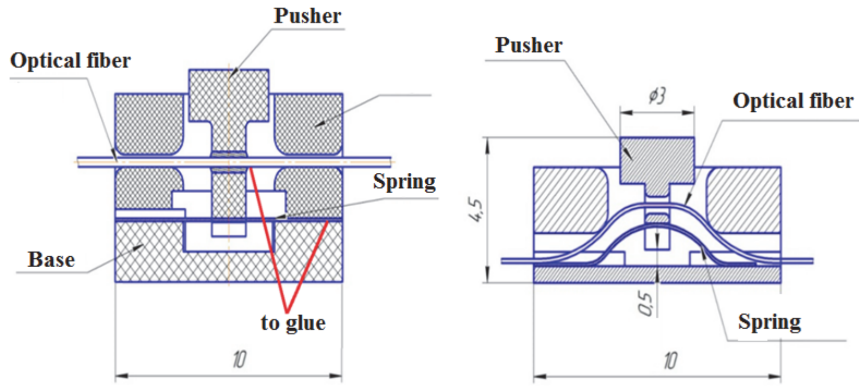
In this paper, we present the results of the development and research of sensitive elements (SE) of FOS-TF and FOS-AP, realized on the effect of changing optical losses in an optical fiber during its macrobending.

Sensors of tactile force based on macro bending losses in optical fiber

The following options for the implementation of SE FOS-TF, differing in ways to ensure the necessary values of tactile force (the highest measured tactile force 10N) were considered [7, 10]:

– FOS-TF with a flat curved spring;
– FOS-TF with a flat straight spring;
– FOS-TF with a helical spring.

SE FOS-TF with a flat spring was designed in two versions - with a flat curved spring (along the radius) and flat straight spring. The modeling and calculation of the elastic element were carried out in the finite-element software package ANSYS-14. The first variant of the SE, by the means of spring bend, made it possible to form a compact space for the stroke of the pusher due to the deformation of the spring. As a result, the geometrical dimensions, in particular, the height, the sensitive element were significantly decreased. However, this design has its own disadvantages that were emerged from the test results. Figures 1 a) and b) show the layout of the structures, for the sensitive element of the tactile force sensor with a flat straight and curved springs, respectively.



a) b)

Figure 1 – Layout scheme of the SE FOS-TF with a flat straight and curved springs

Experimental studies have shown that SE with straight and curved springs have low initial sensitivity and begin to work with a force of 4-5N, which can be explained by the lack of initial bending of the optical fiber.

An increase of FOS-TF sensitivity in the initial part of the conversion function is possible by creating a cantilever loop, the top of which will be fixed in the pusher. During assembling of such sensitive elements, it was noticed that the fiber during bending, under the action of internal elastic forces, rotates around its axis at a certain angle. At the same time, even a small influence on a similar element, a significant change in the optical signal was caused. It was decided to "artificially" create an additional loop slope at an angle of 30-45°. To do this, a variant of the SE in which grooves for the optical fiber laid at the required angle, while preserving the bending of the fiber with a radius of 5-6 mm was designed. The scheme of creating an optical fiber bend at an angle is shown in Figure 2.

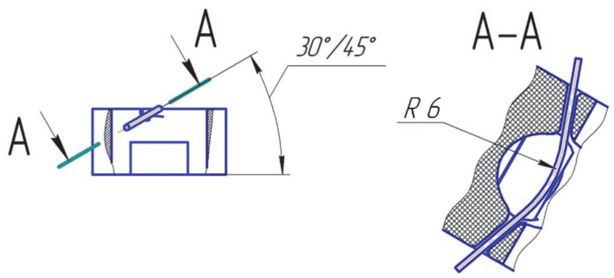


Figure 2 – Diagram of creating an optical fiber bend at an angle

The test results of the SE with a fiber angle of 30 ° and 45 ° showed that their conversion functions (Figure 3) are almost identical. From the obtained results, we can conclude that a prototype SE of FOS-TF with flat curved spring and an optical fiber inclined at an angle of 45 ° is the most acceptable for use.

SE FOS-TF with a helical (spiral) cylindrical spring is characterized by a significantly increased vertical size (more than 2.5 times) and therefore is not considered further.

Angular position sensor based on macro bending loss in optical fiber

A sketch of the mounting of SE FOS-AP on the grip phalanges is shown in Figure 4. Here the design of the SE FOS-AP is made in the form of loops of optical fiber with a radius of 5-7 mm, formed around the axis of the grip rotation. SE FOS-AP components include special supports of optical fiber with fixation 1 and free stroke optical fibers 2.

The installation of the FOS-AP unit on the third rotary gripping unit is shown in Figure 5. The test results of the SE of the FOS-AP unit are shown in Figure 6. As can be seen from the experimental results, the characteristic of the transfer function of the sensitive element has three areas of uncertainty (angles 20, 30 and 45 degrees). This picture can be explained by the fact that the fiber, under the action of internal elastic forces from its deformation, begins to additionally change its curvature at the attachment points in the fixed supports. This leads to a redistribution of the mode content of the optical radiation and to a corresponding change in the optical signal.

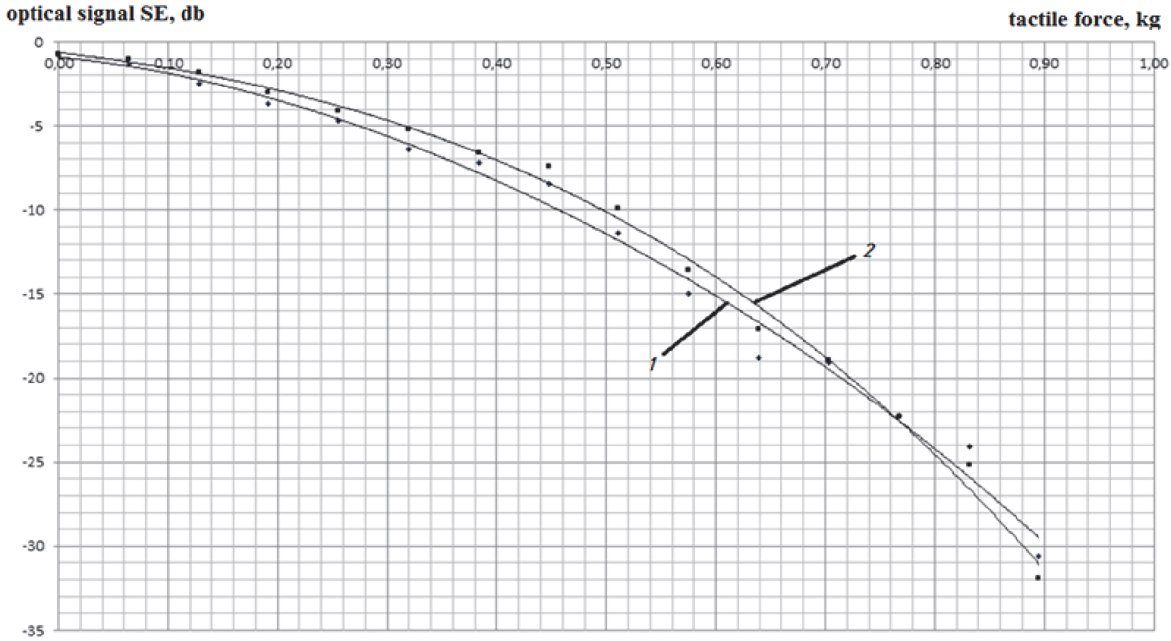


Figure 3 – Conversion function for the SE FOS-AP with a fiber tilt angle of 30° (1) and 45° (2)

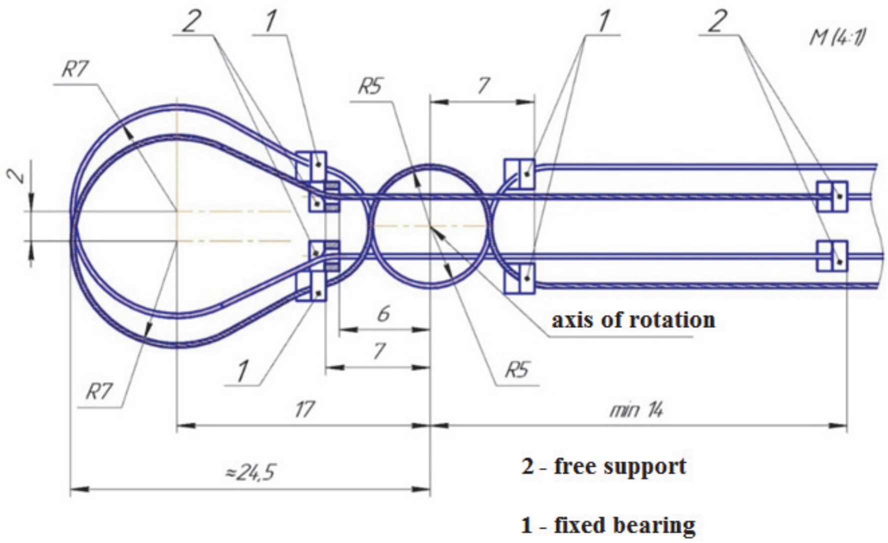


Figure 4 – Sketch of mounting SE FOS-AP on the phalange of capture

One of the options for eliminating this effect is considered and implemented as an application for the invention of the "Fiber optic sensor of the rotation angle". Here, the SE and the measuring system of the rotation angle of the phalanges of the robot's finger grip consist of an electronic transceiver, fiber-optic SE and fiber-optic communication lines connecting the SE with an electronic transceiver. In this case, the sensitive element is made in the form of two optical fiber segments rolled into rings of different radius. Moreover, the ends of the rings are fixed on the phalanges of the fingers, and the rings are arranged in such a way that the radii of the rings increase / decrease synchronously with the angular rotation of the phalange. During operation, the electronic transceiver measures the optical signals of both rings of the optical fiber SE. According to values of these signals, the only result of measuring counts when a segment of the optical fiber (ring) has the conversion function which does not have ambiguity in this area of characteristics.

In some variants of the gripper designs, there is no possibility of using fiber-optic sensing elements in the form of full rings due to their relatively large dimensions. In such grips, ring-shaped segments of optical fibers can be made in the form of half-rings. This will eliminate at the same time a possible change in the position of the plane of the ring that occurs during its bending.

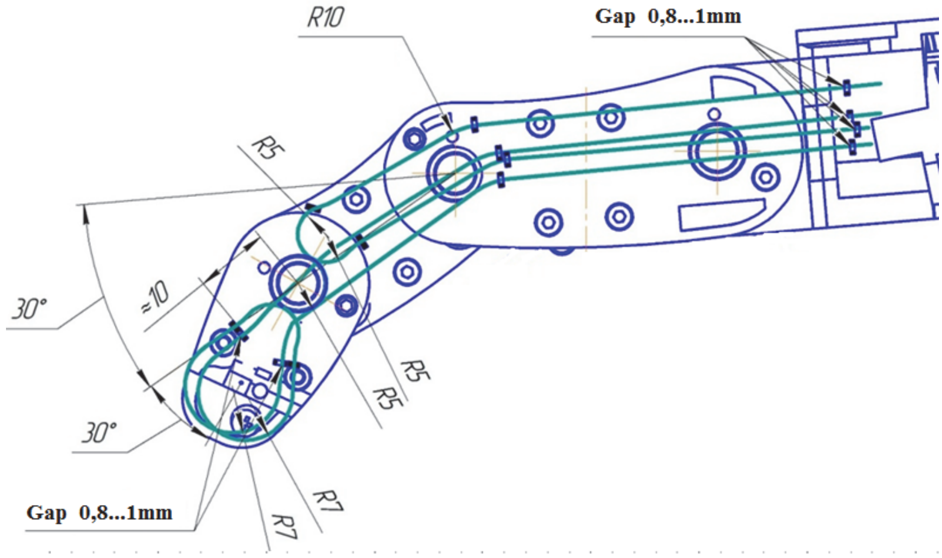


Figure 5 – Scheme of installation of the SE FOS-AP on the third rotary gripping unit

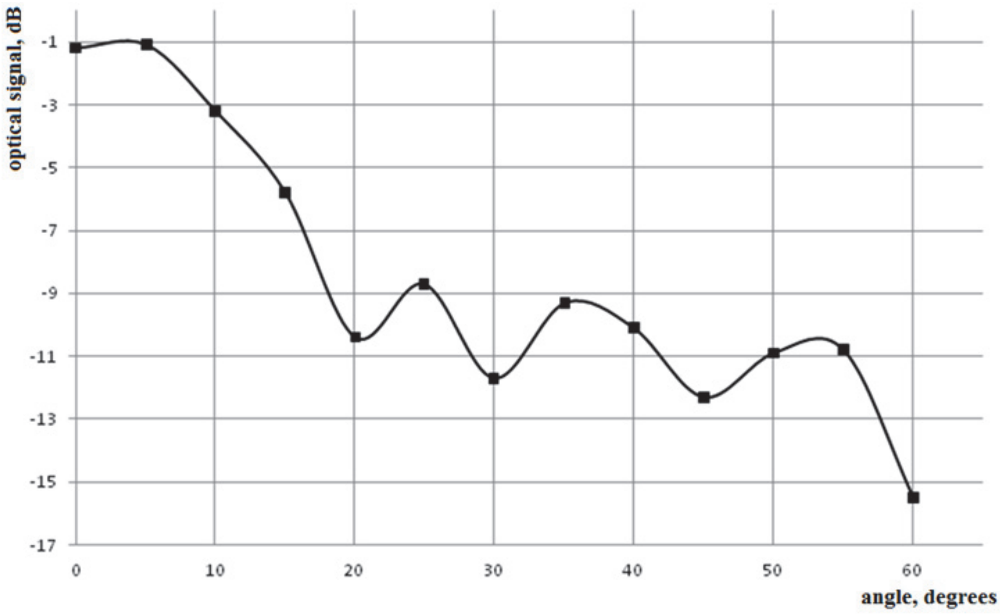


Figure 6 – Conversion function for SE FOS-AP

Both presented options can be used as sensors of the angle of rotation for phalanges of anthropomorphic grippers. However, with further development of the sensitive elements variants, one design variant was revealed. This variant uses the same physical principle, but differs in smaller dimensions, with the absence of the need for preliminary tuning, the absence of ambiguity zones and greater stability of the characteristics during operation. This embodiment was created during the rotation of the phalange and has an artificial preload of the optical fiber. Since at the same time, when the phalange is rotated, a change in the length of the optical fiber occurs, the use of rigid fixation, in this case, is impossible. Figure 7 shows the graph of the conversion function of the SE of FOS-AP on macro bends, that has no sections of the ambiguity of the conversion function.

Analysis of the mutual influence of the phalange grip rotation on the neighboring SE FOS-AP

Figures 8 a) ... 8 d) show the location of the SE FOS-AP installed in the grips of the anthropomorphic robot finger phalange. Here 1-3 are phalanges of the fingers; 4 - an element of the forearm; 5, 6, 7 - sensitive elements of FOS-AP installed in phalange 1, 2, 3, respectively; 8, 9, 10 - the position of the same SE when bending the phalanges. From figure 8 it can be seen that only the rotation of phalange 3 influences SE 7, the turns of phalanges 1, 2 do not affect it. SE 6 is affected by the rotation of the phalange 3 and the phalange 2, the rotation of the phalange 1 does not affect it. SE 5 is affected by the turns of all three phalanges.

It is also seen from Fig. 8 that the bending radii of the sensitive optical fiber are significantly different in the places where the SE is installed and in “transit” places, where the optical fiber passes through the other phalanges. The bending radii of the “transit” optical fiber are significantly greater than the bending radii of the optical fiber of SE.

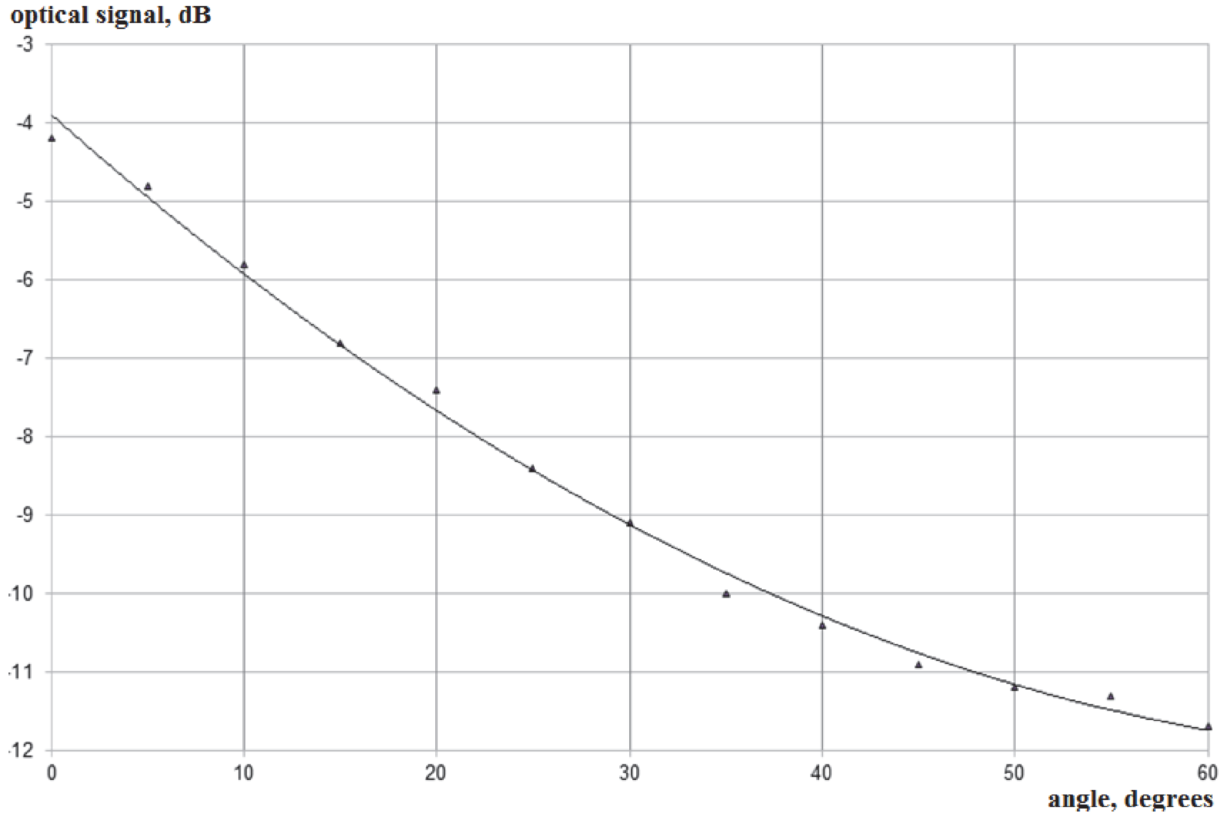


Figure 7 – The conversion function of SE FOS-AP on a macro bend

Figure 9 shows the experimental characterization of the SE transfer function as a function of the bend radius of the optical fiber: curve a) is the transfer coefficient expressed in dB; curve b) is the transfer coefficient, expressed in relative units. As can be seen from the figure, the dependence of the transformation function, expressed in relative units, is almost linear, which makes it possible to approximate it with a linear function.

Figure 10 shows the dependence of the SE transformation function on the initial bending diameter of the optical fiber. This dependence is well approximated by a second-degree polynomial.

Figure 11 a) shows a diagram of the change in the bend radius of optical fiber from the angle of rotation of the phalange. The value of the bending radius of an optical fiber can be determined from the expression:

$$R(\varphi) \approx \frac{\pi}{\pi+\varphi} R_0 \tag{1}$$

where $R(\varphi),\ R_0$ – are the current and initial bending radii of the optical fiber, respectively; φ – is the angle of rotation; $\pi = 3{,}14156$.

The radius of the initial bending is determined by the fixing features of the optical fiber in the phalanges and for the case shown in Figure 11 a) is determined from the expression:

$$R_0 \approx \frac{H}{4} \tag{2}$$

where H – the distance between optical fiber supports. In particular, when $H = 40\ mm,\ R_0 = 10\ mm, \varphi = 1{,}047\ rad. (60\ deg.)$ we will get $R(\varphi) = 7{,}5\ mm$, which corresponds to the diameter $D(\varphi) = 15\ mm$.

In accordance with the graph in Figure 11, the SE transmission coefficient changes by 1.3 times.

For the optical fiber passing through the phalange in "transit", when $H = 40\ mm,\ \ \varphi = 1{,}047\ rad. (60\ \ deg.)$ we get $R(\varphi) = 38{,}2\ mm.$. In accordance with the graph in Figure 11, the SE transformation function changes by only 0.5%.

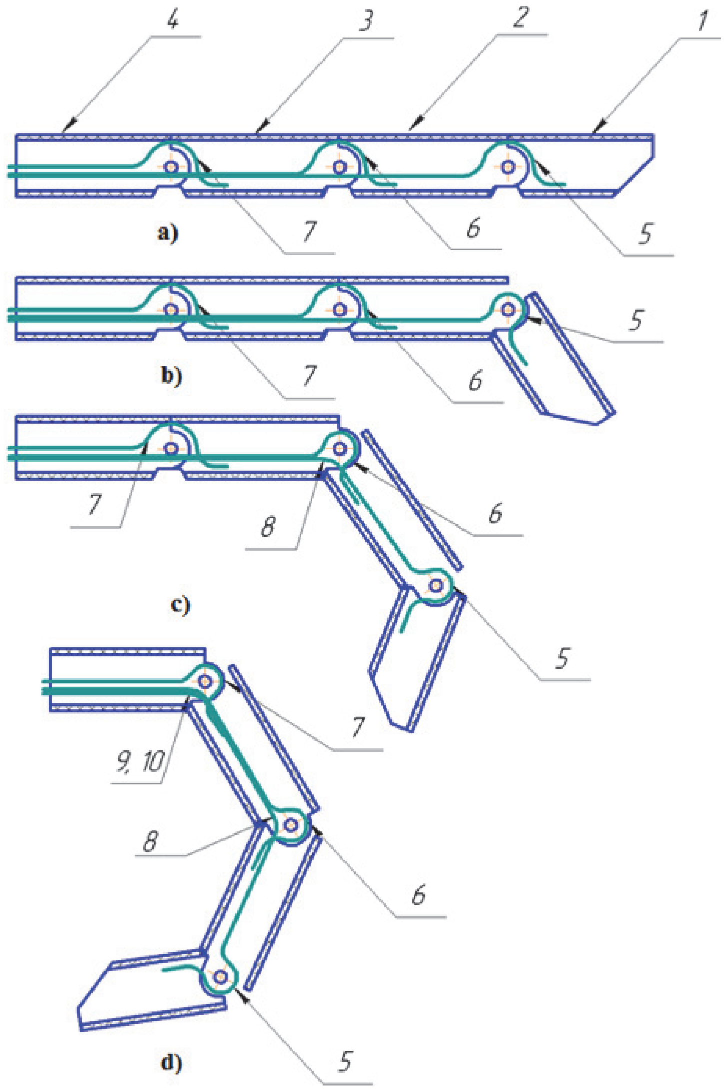
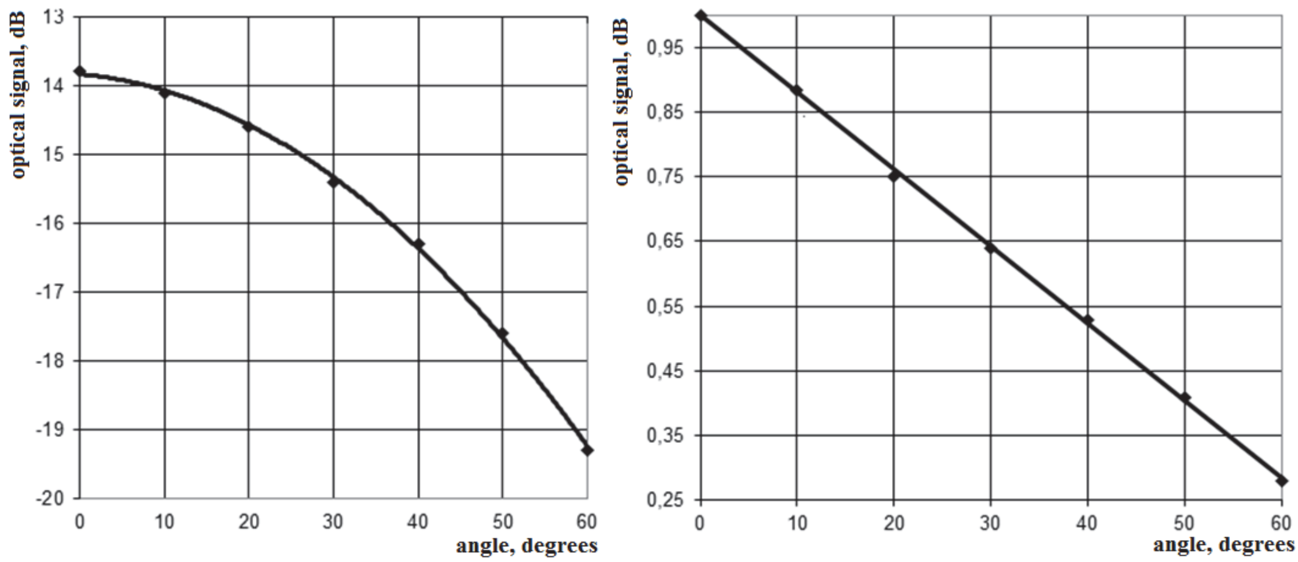


Figure 8 – Location and effect of the phalange rotation on SE FOS-AP



a) b)

Figure 9 – Experimental dependence of the transfer function of SE FOS-AP from the angle of rotation of the phalange

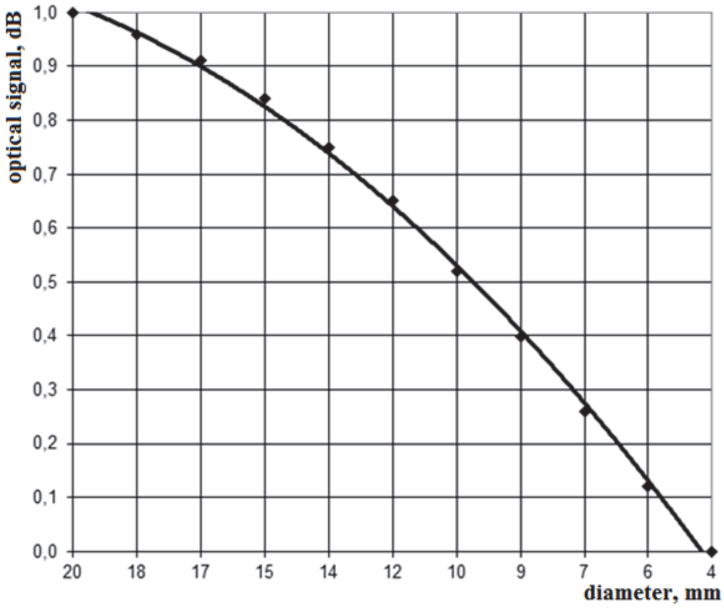


Figure 10 – Experimental dependence of the transfer function of SE FOS-AP from the initial diameter of the bend of the optical fiber

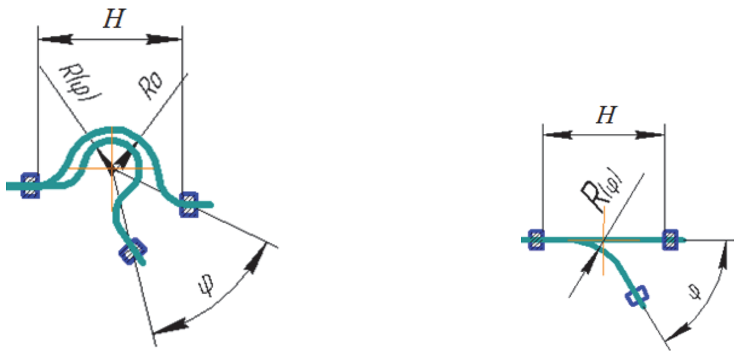


Figure 11 – Change of the bend radius of the optical fiber from the angle of rotation of the phalange

In light of the above, optical signals $Y(\varphi_5), Y(\varphi_6), Y(\varphi_7)$ SE of phalanges 1, 2, 3 can be written in the form:

$$\begin{aligned} Y(\varphi_5) &= a_5\varphi_5(1-b_7\varphi_7)(1-b_6\varphi_6) \\ Y(\varphi_6) &= a_6\varphi_6(1-b_7\varphi_7) \\ Y(\varphi_7) &= a_7\varphi_7 \end{aligned} \tag{3}$$

where: $\varphi_5, \varphi_6, \varphi_7$ – are current angles of rotation of the corresponding phalanges 1, 2, 3; a_5, a_6, a_7 – are SE sensitivity of the corresponding phalanges to bending; b_6, b_7 – influence on the transformation function (sensitivity) of bending of nearby phalanges.

With this:

$$\begin{aligned} \varphi_5 &= \Delta\varphi_5 + \varphi_5^{\text{H}} \\ \varphi_6 &= \Delta\varphi_6 + \varphi_6^{\text{H}} \\ \varphi_7 &= \Delta\varphi_7 + \varphi_7^{\text{H}} \end{aligned} \tag{4}$$

where $\varphi_5^{\text{H}}, \varphi_6^{\text{H}}, \varphi_7^{\text{H}}$ – are the angles of SE initial installation; $\Delta\varphi_5$, $\Delta\,\varphi_6$, $\Delta\,\varphi_7$ - the amount of deviation of the corresponding angle from the initial setting during phalanges bending.

In accordance with expressions (3-4), the method of compensating for the influence of the rotation of the nearby phalanges on the signals of the SE FOS-AP signals can be the following:

1) measure the value of the third phalange signal $Y(\varphi_7)$;

2) from the expressions (3-4) determine the phalange 3 angle of rotation the value:

$$\Delta\varphi_7 = \frac{Y(\varphi_7)}{a_7} - \varphi_7^{\mathrm{H}} \tag{5}$$

3) measure the value of the second phalange signal$Y(\varphi_6)$;

4) from expressions (3-4), taking into account expression (5), determine the phalange 2 angle of rotation value, thereby compensating for the effect of bending of the phalange 2 on the measurement result:

$$\Delta\varphi_6 = \frac{Y(\varphi_6)}{a_6\left[1-b_7\left(\frac{Y(\varphi_7)}{a_7} - \varphi_7^{\mathrm{H}}\right)\right]} - \varphi_6^{\mathrm{H}} \tag{6}$$

5) measure the value of the optical signal of the third phalange $Y(\varphi_5)$;

6) from expressions (3-4) determine the phalange 1 angle of rotation value, thereby compensating for the effect of bending of the phalanges 1, 2 on the measurement results:

$$\varphi_5 = \frac{Y(\varphi_5)}{a_5\left[1-b_7\left(\frac{Y(\varphi_7)}{a_7} - \varphi_7^{\mathrm{H}}\right)\right]\left[1-b_6\left(\frac{Y(\varphi_6)}{a_6\left[1-b_7\left(\frac{Y(\varphi_7)}{a_7} - \varphi_7^{\mathrm{H}}\right)\right]} + \varphi_6^{\mathrm{H}}\right)\right]} - \varphi_5^{\mathrm{H}} \tag{7}$$

Since the dependence of the signal variation of the SE FOS-TF signal in the first approximation is also linear (Figure 12), the method for rotation influence compensation of nearby phalanges on the measured tactile force is similar to the method for compensation of changing measured angular position.

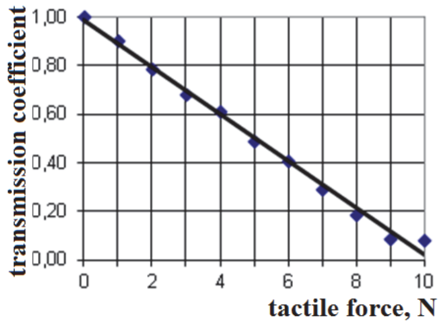
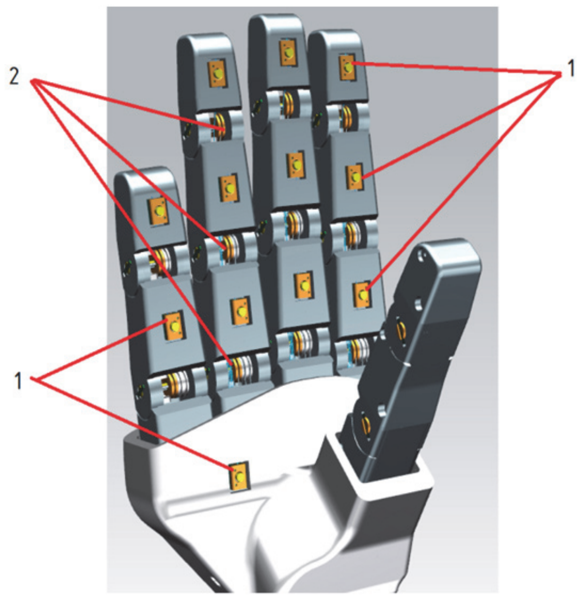


Figure 12 – Experimental characteristics of the conversion function SE FOS-TF

The layout scheme of SE FOS-AP and SE FOS-TF on anthropomorphic grip is shown in Figure 13.



1 - SE FOS-TF, 2 - SE FOS-AP

Figure 13 – Layout scheme of SE FOS-AP and SE FOS-TF on anthropomorphic grip

Electronic transceiver

The block diagrams of the FOS-AP and FOS-TF electronic transceivers (ET) are identical. Simulation and experimental study of several variants of ET were performed.

Figure 14 shows a block diagram of a single-channel variant of an ET with a local stabilization circuit of a laser diode (LD).

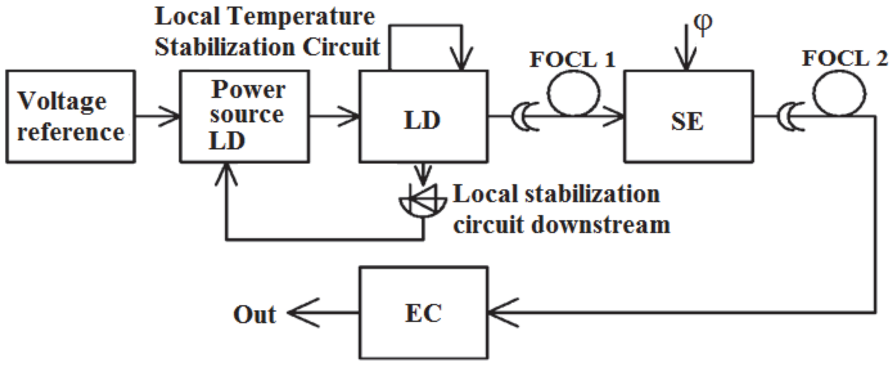


Figure 14 – Structural scheme of single-channel FOS with a local stabilization circuit LD radiation

Stabilization of LD optical radiation this variant is carried out by local stabilization circuits embedded in the LD housing - the feedback circuit for the optical flow and the feedback circuit for the temperature of the LD crystal. The output signal of the SE is detected by the photodetector, and after amplification (electronic converter controller unit - EC) is fed to the output of ET. Such variant of ET, despite its simplicity, provides high stability of the light flow at the output of the LD. However, this option does not allow to compensate changes, incl. temperature, occurring in the fiber-optic communication line (FOCL 1, FOCL 2) and in SE FOS.

To reduce the additive and multiplicative errors, the output signal of the ET is formed in the EC from the signal Y_1 of the SE sensor, for example, FOS-AP (for phalange 1 rotation angle), according to the expression:

$$Y_{\text{H}}(\varphi) = 4096(\frac{Y_1(\varphi_5)-Y_{min}(\varphi_5)}{Y_{max}(\varphi_5)-Y_{min}(\varphi_5)}), \tag{8}$$

where $Y_{max}(\varphi_5), Y_{min}(\varphi_5)$ - are maximum and minimum values of the SE transformation function.

Since the temperature coefficient of the LD radiation power, not covered by the local stabilization chain is 2.5… 4% / ºC, when introducing the local stabilization chain via the optical channel, the LD temperature coefficient is determined by the temperature coefficient of the feedback circuit, i.e. by the photodiode, and amounts to 0.2 ... 0.3%/ºC. Maintaining of LD temperature by the means of local stabilization circuit of the LD - thermistor - Peltier element can stabilize the LD temperature with an error of ± 1 ºC. Then the additional temperature instability of the LD radiation will be ± 0.2 ... 0.3%.

Since the rest of the FOS elements (FOCL-SE-FOCL-EC) are not covered by feedback, the resulting additional temperature error of the FOS will be about 105% in the temperature range minus 80 ... +80 ºC. As can be seen, the additional temperature errors of the sensors at the macro bends (nondistinctive inclusion option) reach significant values. Therefore, in order to reduce the value of the additional temperature error, differential inclusion schemes were investigated.

Figure 15 shows a block diagram of the differential variant of the ET with a local channel for stabilization of the LD radiation. Two “identical” measurement channels are used here. Stabilization of LD optical radiation in this variant is carried out by local stabilization circuits embedded in the housing – by the means of a feedback circuit for the optical flow and feedback circuit for the temperature of the LD crystal.

The output signal of the ET is formed in EC from the signals Y_1, Y_2 of two sensor channels, for example, FOS AP for the phalange 1 rotation angle, will be calculated according to the expression:

$$Y(\varphi_5) = 2048(\frac{Y_1(\varphi_5)-Y_2(\varphi_5)}{Y_1(\varphi_5)+Y_2(\varphi_5)} + 1) \tag{9}$$

When implementing EC expression (9), compensation is performed both for the additive measurement error and for the multiplicative error. In this case, for the input/output of radiation to/from the CE, a fiber optic line consisting of four optical fibers is required.

This option allows you to compensate the changes, incl. temperature, occurring in the optical communication channel (FOCL 1, FOCL 2, FOCL 3, FOCL 4) and in SE1, SE2 only in the case of the identity of these channels and the conditions of their operation. In addition, as noted in [8-9], the use of expressions (8-9) is effective only in the case when $Y_1 + Y_2 \approx const$. For larger changes in Y_1, Y_2 , the condition is not satisfied and the effectiveness of stabilization decreases sharply. In addition, there is a sharp increase in the nonlinearity of the transformation function. Experimental studies confirm the above conclusions.

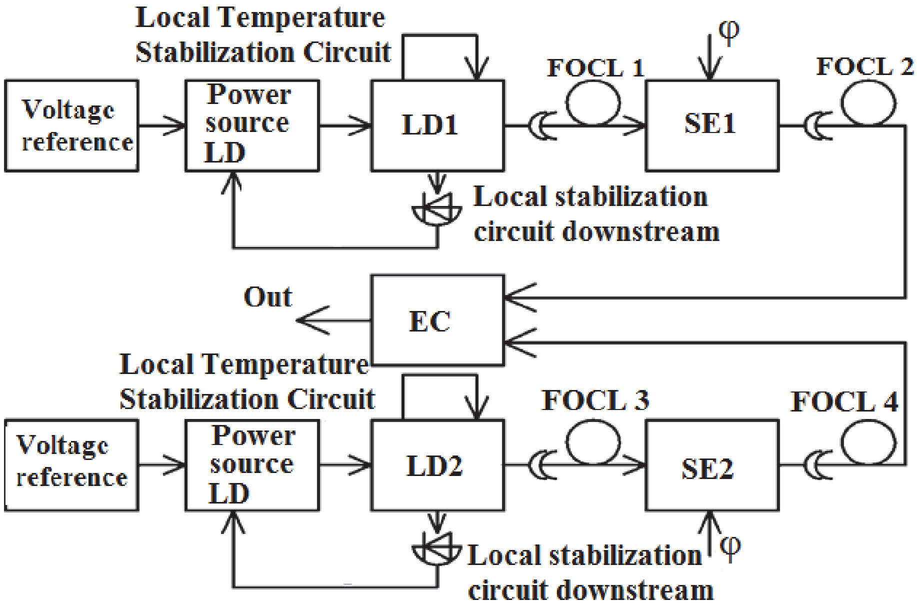


Figure 15 – Block diagram of a differential FOS with local channels stabilization

As in the previous version, when introducing a local stabilization circuit via the optical channel, the temperature coefficient LD is determined by the temperature coefficient of the feedback circuit, i.e. by the photodiode, and amounts are in range 0.2 ... 0.3%/ºC. As in the previous version, while maintaining the LD temperature, the local stabilization circuit of the LD thermistor-Peltier element manages with stabilization of the LD temperature with an error of ± 1 ºC. Then the additional temperature instability of the LD radiation will be ± 0.2 ... 0.3%. The following additional temperature error values were obtained experimentally for 15 specimens of sensitive element sensors (without temperature compensation): specimens 1-3 - 13.7%; specimens 4-6 - 16.2%; specimens 6-9 - 10.7%; specimens 9-12 - 20.1%; specimens 12-15 - 10.3%.

Hence, for the differential structure under consideration, the final additional temperature error of the FOS should be about 20.1% in the temperature range minus 80… + 80 ºC.

Figure 16 shows a block diagram of a differential FOS with a common channel of stabilization along the ES circuit, which has no indicated disadvantages. The optical signal LD by the means of an optical radiation divider is divided into two parts, which through identical FOCL-1, FOCL-2 each transfer at its own SE: SE1, SE2 combined in one design. The output optical signals SE through identical FOCL-1, FOCL-2 enter the EC, which calculates the corrective action and controls the operation of the LD. In this case, for the input/output of radiation to/from the SE, a fiber optic line consisting of four optical fibers is required. The stabilization of the LD parameters in this variant is carried out by the means of a digital-to-analog converter (DAC) with a common stabilization circuit covering SE1, SE2, FOCL-1-1, FOCL-1-2, FOCL-2-1, FOCL-2-2, LD, and optical radiation divider. This option effectively provides high stability of optical radiation at the output of LD and transmission coefficients SE1, SE2, FOCL-1-1, FOCL-1-2, FOCL-2-1, FOCL-2-2, optical radiation divider, however, requires the identity of the optical radiation division channels.

Analysis of the ET structure (Figure 16) allows determining the sensitivity coefficients of the basic elements of the FOS-TF (LD; photodiode – PD, operational amplifier OA, optical splitter OS, FOCL, DAC, ADC) to temperature and additional temperature error in measuring the tactile force (Table 1)

Table 2 shows the experimental results of the additional temperature error when measuring the angular position of the phalanges of the differential design of the FOS-AP for several angular positions of the SE.

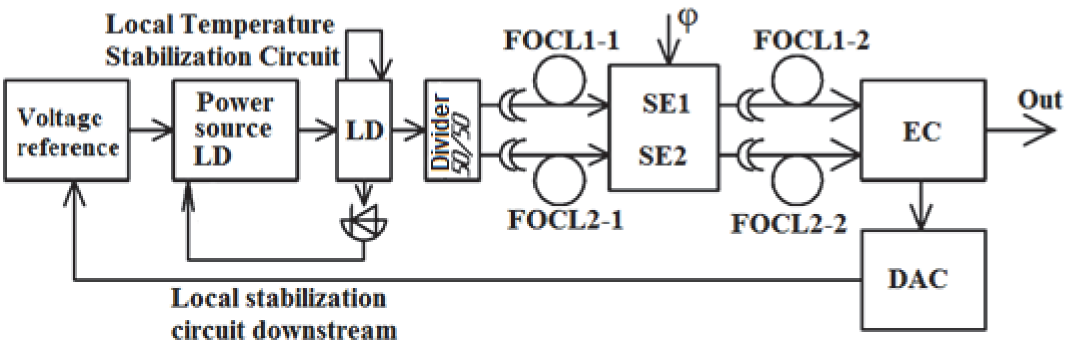


Figure 16 – Block diagram of differential FOS with local and common stabilization channels

Table 1. Estimation of the components of the main and additional reduced errors of FOS-TF

FOS-TF element	**Transfer ratio**	**Components of the main error**	**Components of the additional error**		
		Deviation from nominal	**Temperature, %/°C**	**Temperature range, °C**	**Deviation from nominal**
LD	5 mW / mA	±2,5 mW	−2,5	±1	−2,5 mW
PD	0,2 mA / mW	$1\cdot10^{-4}$ mA	−0,2	±1	−0,2 mA
OA	200 rel. units	2	$1\cdot10^{-4}$	−40…+80	0,12 %
OS	0,5 rel. units	0,025	0,01	−40…+80	−1,2 %
FOCL	0,85 rel. units	0,02	-	−80…+80	-
ADC	248 bit / V	0,001 bit / V	-	−40…+80	0,01 %
DAC	0,00488 V / bit	$4{,}9\cdot10^{-6\ V/bit}$	-	−40…+80	0,005 %
The basic reduced measurement error of tactile force					0,86 %
The additional reduced measurement error of tactile force					0,54 %

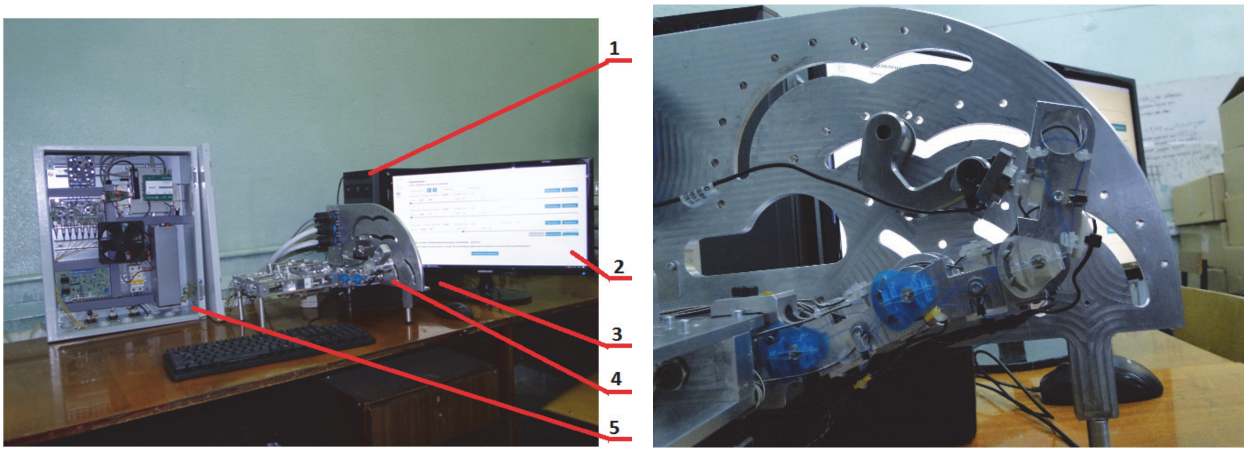
Table 2. Additional temperature error of differential FOS-AP

Temperature, °C	**Additional temperature error, %**		
	0 deg.	30 deg.	60 deg.
-80	0,233	0,139	0,085
+80	0,291	0,233	0,022

Experimental research facility

Figure 17 shows a photograph of a specialized facility for investigation of the conversion function of FOS-AP and FOS-TF. Figure 18 contains a photograph of its electromechanical unit with an installed gripper and sensors. The stand (Figure 17) consists of a personal computer 1 and a monitor 2, an uninterruptible power supply unit 3; an electromechanical unit with an executive gripping group (EGG) 4, the control and management unit 5. The EGG is an anthropomorphic finger, consisting of three phalanges, each of which has its own independent drive located in the electromechanical unit. The computer contains a control program in which the settings of the movement of the EGG are set. Settings of movement are the angle of inclination of each phalange in the range of 0 ... 60 degrees and the amount of tactile force developed by each phalange. The computer program generates a test report in which the experimental dependence of the output signals of FOS-AP and FOS-TF on successively specified angular displacements and tactile loads is presented in tabular form. The monitor displays the program settings and the experimental data obtained. The control and management unit generates control electrical signals for the electromechanical unit in accordance with the specified settings and receives signals from control (precision) sensors that track the position of each phalange and the developed tactile force.

The main technical characteristics of the research facility are given in Table 3. The RF is equipped with control sensors of angular position, tactile force, and temperature. Appearance of the anthropomorphic robot AR-600 produced by PJSC “Android Technique” with a modified capture option, in which the sensitive elements of fiber-optic sensors of the phalanges of the arms (15 pieces) and sensitive elements of the fiber-optic sensors of the tactile force of the phalanges of the arms (16 pieces) are shown in Figure 19.

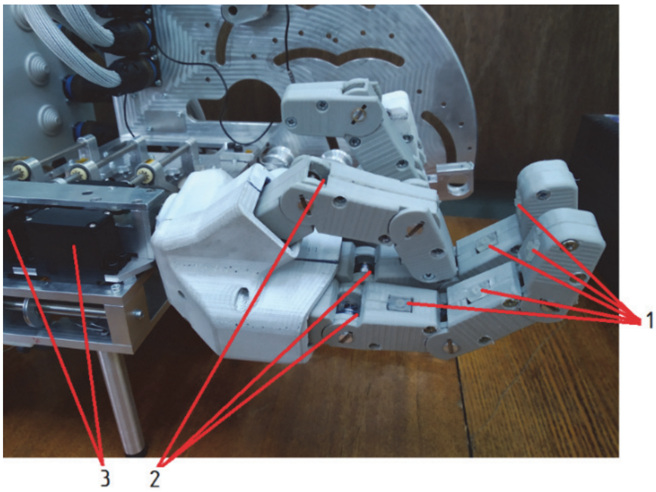


a) b)

a) 1 – PC; 2 – monitor; 3 – uninterruptible power supply unit; 4 – electromechanical unit with an executive gripping group; 5 – control and management unit; b) an executive capture group on an enlarged scale

Figure 17 – Photograph of the research facility

KTXB-64-M heat and the cold chamber was used for the investigation of temperature effect on the positional and tactile characteristics of the FOS. Chamber has the following characteristics: temperature maintenance range from minus 80 to + 80 ^{O}C, temperature uniformity over the chamber volume ± 0.2 ^{O}C, time to reach the limit values of 50-60 minutes.



1 – tactile force sensor (16 pieces); 2 – angle position sensor (15 pieces); 3 – electric drive grip

Figure 18 – Photograph of the electromechanical unit with the set grip

Table 3. Main technical characteristics of the research facility

Parameter	**Value**
Number of grip fingers, pcs	1
Number of finger phalanges, pcs	3
The range of angular positions of the fingers phalanges, degrees	0-60
The way to set the angular positions of the phalanges	independent
The main error of positioning by angle, %	0,025
Tactile force setting range, N	10
The main error in setting the tactile force, %	0,2

Parameter	Value
The size of the contact patch, at which tactile force is set, mm	3x3
The range of measured phalanges temperatures, ^{O}C	-80…+80
The basic error of temperature measurement, ^{O}C	0,2
Operation temperature for the block of electronics RF, ^{O}C	laboratory
Operating temperature of the electromechanical unit, ^{O}C	-40…+80

Figure 19 – Photo of the anthropomorphic robot AR-600
with sensors FOS-AP and FOS-TF

References

1. Koyama, Yu. Multi-channel measurement for hetero-core optical fiber sensor by using CMOS camera [Текст] / Yuya Koyama, Michiko Nishiyama, Kazuhiro Watanabe // Proc. of SPIE, Fifth Asia-Pacific Optical Sensors Conference, 2015. – Vol. 9655. – № 965525-4.
2. Silva, A.S. Design and characterization of a wearable macro bending fiber optic sensor for human joint angle determination [Текст] / Ana S. Silva, André Catarino, Miguel V. Correia, Orlando Frazão // SPIE, Optical Engineering, 2013. – 52(12), № 126106.
3. Martellucci, S. Optical Sensors and Microsystems [Текст] / S. Martellucci, A.N. Chester, A.G. Mignan // Boston: Kluwer academic publishers, 2002. -318 pp.
4. Garmash, V.B. Opportunities, tasks and prospects of fiber-optic measuring systems in modern instrument-making [Text] / V.B. Garmash, F.A. Egorov, L.N. Kolomiets, A.P. Neugodnikov, V.I. Pospelov // Special edition "Foton-Express", 2005. - №6. - p. 128 - 140.
5. Khlybov, A.V. Fiber-optic polarimetric sensors of physical quantities. [Text]: dissertation for the degree of Ph.D.: 04/01/03 / AV Khlybov. - SPb., Petersburg, 2004. - 215 p.
6. Tatmyshevsky K.V., Makarova N.Yu., Pavlov D.D. Focused and distributed (tactile) sensors based on the phenomenon of mechano-luminescence for instruments for measuring and measuring pulse pressure [Text]: R & D report, State contract No. 3 of November 11, 2009 / State Educational Institution of Higher Professional Education "Vladimir State University", 07.10.2011. - Inv. No. 01200965317.
7. Matyunin, S.A., Babaev, O.G. Fiber-optic Sensor of Tactile Force for Anthropomorphic Robot Grips. IOP Conference Series: Materials Science and Engineering 302(1),012040
8. Babaev, O.G., Matyunin, S.A., Paranin, V.D. Linearization of Positional Response Curve of a Fiber-optic Displacement Sensor. IOP Conference Series: Materials Science and Engineering 302(1),012051
9. Matyunin, S.A., Stepanov, M.V., Babaev, O.G. Simulation of the Characteristics of a Magneto-Optical Displacement Transducer. Measurement Techniques 59(8), c. 832-837 DOI:10.1007/s11018-016-1053-7
10. Matyunin, S.A. Research on Characteristics of Fiber Optic Sensors for Anthropomorphous Robots. Procedia Engineering 176, c. 128-136 DOI:10.1016/j.proeng.2017.02.280

V.P. Andreev

THE CONCEPT OF USING THE THEORY OF MULTI-AGENT SYSTEMS TO DESIGN CONTROL SYSTEMS FOR MOBILE ROBOTS WITH MODULAR ARCHITECTURE

MSTU "STANKIN", IINET RSUH, IL "Sensorika", Moscow, Russia
andreevvipa@yandex.ru

Abstract

The approach to the solving of the actual fundamental and applied scientific problem – development of principles and methods of robotic complex and systems design with modular architecture based on pyramidal (hierarchical) build-up topology of their computational and control systems is considered. Modular architecture makes it possible to execute fast reconfiguration of robotic systems. The use of hierarchical topology for constructing the robot IMCS, when each module and submodule has its own IMCS with a separate processing unit, allows increasing the computation speed in the system by distributing the computational load between the modules computing devices. It begs an analogy with such directions of scientific research in the field of information technologies as distributed information systems, computer networks, methods of artificial intelligence. Achievements in these areas are integrated in research, united by the common name "multi-agent systems". In this paper, we consider an approach that allows us to describe the hierarchical topology of the design of the robot IMCS with modular architecture in terms of multi-agent systems, which will allow us to use modern achievements in this scientific field.

Keywords: modular robot, mobile robot, modular architecture, multi-agent system, hierarchical topology, distributed computing systems, computer network, information interaction, hardware and software system.

Acknowledgments

Research is supported by the Russian Foundation for Basic Research: Grant 19-07-00892a.

Introduction

Currently, the control system (CS) of mobile robots (MR) are being designed mainly as a centralized system, i.e. the entire information-measuring and control system (IMCS) of robot has a single computational center responsible for control and processing of sensory information. Centralized control requires a high performance computer so that the robot can solve its tasks in real time. With the increase in the functionality of MR, as well as with the expansion of the conditions of their use (for example, for tasks in extreme conditions of the Arctic and Antarctic, space and in solving the problems of the Emergencies Ministry), the complexity of robotic system control tasks increases. Consequently, the algorithmic complexity of executable control programs is increasing, which places extremely high requirement on the system CPU. The complication of robot IMCS leads to the need to find methods for implementing the computational control process and decision-making based on distributed calculations performed by a multiple processors. This will significantly reduce the requirements for computing power of multiprocessor system and increase the system performance as a whole. Using this approach to implement the IMCS of a mobile robot is *to design and create robots with a modular architecture*.

A precise definition of the term of "Robot with modular architecture" has not yet been formed. Both robots with homogeneous [1-4] and heterogeneous [5-8] design can be attributed to this area of research. Recently, a new area in this research has separated, which is rather a search for design methods aimed at providing developers and manufacturers of MR with similar modular architectures. Using the proposed ready-made architectures, developers can create their own robotic and mechatronic devices. Thus, some developers offer not a specific architecture of any robot, but a whole infrastructure – a set of hardware and software tools for design new robots.

One example is the system H-ROS [9], based on the ROS framework. This system allows on to design new robots by assembling them from components that are compatible with H-ROS. The user only needs to program the robot components responsible for the perception ("brain" of the robot) and independently develop program applications for specific tasks, without solving the integration problem of different technologies and interfaces. The H-ROS infrastructure is built on a standard industrial bus and is compatible with unified software frameworks such as ROS-2. At the same time, for the use of H-ROS, it is necessary on each component to have a special microprocessor required to run ROS-2, which imposes a number of restrictions on the use of this system.

The modular architecture of a miniature mobile platform AMiRo is considered in the paper [10]. Features of the architecture are: powerful hardware with small dimensions, open source software, distinct division of functions and tasks between modules, the ability to work with external applications and the ability to connect other devices (modules). The disadvantages include the fact that the inclusion of third-party modules can be difficult due to the complex multi-level software architecture, as well as the small size of the robot.

A common approach is to create modular architectures for combining different modules – actuators, sensors, controllers, etc. In [11], the authors propose a software framework R2P (Rapid Robot Prototyping) to create robots based on standardized hardware modules with appropriate software and developed interaction protocol Real-time CAN (RTCAN). R2P provides easy integration of mostly simple modules, without taking into account more complex devices (for example, entire mobile platforms).

The article [12] discusses a similar modular architecture RoboCAN, based on the CAN bus and embedded in a Robotic Integrated Development Environment (RIDE) or other similar environments. Similar to R2P, the RoboCAN framework allows the integration of various actuators, sensors and control devices into mobile robot architectures (mostly simple components). The software of the modules is independent of the microcontrollers used; this allows one to create complex devices without modifying the existing approach in the implemented solution.

The analysis of articles in the field of modular robots architectures leads us to the conclusion that these studies focused mainly on the creation of standard software and hardware interfaces between the modules to ensure the robots reconfigurability. In these studies, there are no fundamental solutions to the robot's IMCS architecture that can solve not only the problem of reconfigurability, but also the problem of providing real-time mode for the ever-increasing complexity of the robot functionality if *embedded computing devices* (relatively low-power microprocessors or single-board computers) are used in modules.

Functional-modular principle of design of the robots IMCS architecture

In [13, 14] it is proposed to create a distributed control system, in which the general task is divided into subtasks according to the functional-modular principle, and then these sub-tasks are distributed among the modules. This approach provides a significant reduction in the computational load on the computing device of each module, which allows use of the *embedded computing devices* in IMCS of modules.

Research is supported by the Russian Foundation for Basic Research: Grant 19-07-00892a. The basis of the functional-modular architecture is the ***principle of full functionality of modules*** [15, 16], which is formulated as follows: *each robot module should be able to perform its goal function in any convenient way, using only its own means to execute commands from an external control system*.

In the mentioned articles, the principles and methods of construction of the hardware and software parts of the MR control system with a modular architecture are considered. The proposed structure is quite simple and it can be automatically reorganized in the "plug and play" mode.

Based on the principle of full functionality, it is proposed to consider the structure of the MR control system (in the minimum version) as a synergetic union of full-featured modules (fig.1) [17, 18]: Transport Module – TM, Power Module – PM, the sensory system, consisting of Short-Range Sensor Module – SRSM and Long Range Sensor Module – LRSM, modules of the impact on the environment – SPM (manipulators, machining process, grippers etc.) and Intelligent Control Module – ICM. In this structure, each module is responsible for only one function of the robotic system.

System-wide control. The function is implemented in the Intelligent Control Module (ICM), which is a supervisor in relation to all other robot modules. In our case, ICM forms the control goal and controls only the result of its achievement by the slave module, and does not control the process of task execution by this module. This functional completeness is the main difference between our approach and others, in which the module-supervisor controls the operation of all modules at the executive level.

The transport function is implemented by the transport module (TM) – the module allows the robot to move in the environment. The module can be equipped with different locomotors types: wheels, tracks, legs, screw propeller, etc. The module task is to move the modular robot from the current location to the goal specified by ICM. Naturally, the motion control algorithms for TM with different locomotors type will be different and implemented by own IMCS of the module. If necessary, the TM may directly address a request to the robot sensor system, thus reducing the load on the ICM processor and the data channel. Therefore, it is necessary to have a direct interaction (in the figure "pass-through") between the TM and the SRSM.

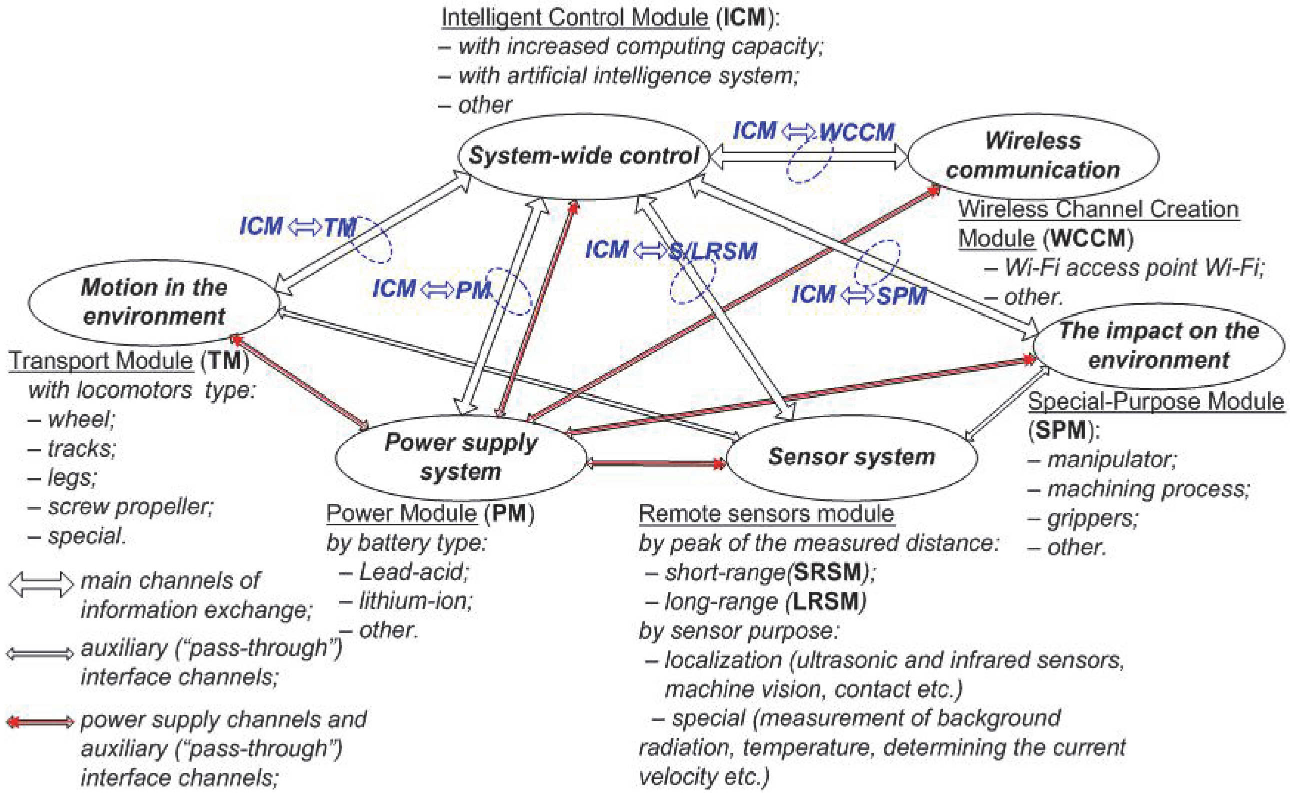


Figure 1 – Functional-modular architecture of an information-measuring and control system of a mobile robot

The information function is implemented by a remote sensor system consisting of several sensor modules (SM): the short-range sensors module (SRSM) and the long-range sensors module (LRSM). Modules detect obstacles and the manipulation objects and can provide sensory information to other modules by "pass-through" channels (bypassing ICM). In addition, these modules can give them a command to suspend execution of their functions in special cases, for example, to prevent a possible collision of MR with an obstacle. In this case, the ICM can prohibit such communication if the collision is part of the necessary actions. One should consider that each module has its own sensor system consisting of internal state sensors when implementing direct information interaction between full-featured modules.

The technological function (and partially of the *transport function*) – the impact on the environment. Special Purpose Modules (SPM) implements the function. These modules are manipulators-actuators (MAM) and the gripping modules (GM), which provide the ability to use different tools for moving the manipulation objects within the robot workspace. The *technological function* is also implemented by the mechanical operations modules (MOM) – technological equipment for the manipulator module, which allows making various technological impact – drilling, milling, etc.

The communication function is implemented by means of the network organization of information interaction between modules (on Fig.1 shows bi-directional arrows). A special Wireless Channel Creation Module (WCCM) creates communication of a mobile robot with an external supervisor – a human operator of a robotic system. Figure 1 shows the network topology of a star type, but a bus and mixed topology is possible. The choice of network topology is likely to depend on a number of factors that have yet to be determined.

The power function is implemented in the power module (PM) – the module provides power supply to the electronic and electromechanical components of the MR modules and their safe switching on and off: the correct shutdown of the program blocks, saving the file system state, monitoring and indication of the batteries state, etc. Disconnection of the power supply of some module is performed only at the command of the ICM – disconnection of all modules (including itself) in the case of robot shutdown, or those modules which functions are not needed now (for example, to save energy). The structure of distributed stabilized power supply, implemented using the PoE principle, is proposed [14]. The total unstabilized power supply of the robot, including electric drives, is carried out from the PM containing the battery. Each Module is supplied with the own Voltage Stabilizer (MVS) to produce the desired number of stable voltages, which supply with stable voltages only to the module electronic devices, e.g., microprocessors. The voltage to the stabilizer is supplied directly from the battery. When there are oscillations of the input voltage (caused, for example, by starting currents of electric motors), the presence of such voltage stabilizer eliminates the possible interruption of the program code execution by computing devices of modules. This distributed stabilized power supply does not impose strict restrictions on the required number of stabilized voltages in the new module design – the developer can make their own decision. In addition, in case of exceeding the total power consumption, it is enough only to install additional batteries in the PM or replace the module itself by installing a module with more battery capacity. In addition, the heat generated in the process of voltage stabilization is distributed between the modules, and is not concentrated in a single power supply unit, which would require installing radiators. The use of distributed stabilized power supply reduces the number of inter-module electrical connectors; simultaneously a power supply lines can be used to pass rare commands and small information.

It should be noted that the given functional-modular architecture of the mobile robot IMCS reflects only the general approach. Research has shown the need for further development of this architecture. The results of research supported by the RFBR Grant 16-07-00811a "Development of functional-modular principle of construction software and hardware of intelligent mobile robotic systems" are published in [19-25]. The main results are as follows:

1. *The functional-modular* principle of construction of hardware and software of reconfigurable mobile robotic systems (MRS) is developed and reasoned.

2. *The principle of the modules full functionality* is formulated and efficiency of its use for the implement of distributed control in MRS of robotic systems is shown. This approach allows us to divide the computational process of implementing the objective function by the robot into functional subtasks and distribute them between the computing devices of individual modules.

3. It is shown that the implementation of a full-featured distributed control requires such inter-module information interaction, in which *each full-featured module should be able to exchange information directly with any other module*.

4. The method of inter-module information interaction is proposed, which is based on the system of geographically distributed control of mechatronic devices developed by the authors, in which *the minimum object of control is any electronic or mechatronic device* – an onboard computer, a microcontroller, a sensor, a manipulator, or other executive mechanism.

5. *The network organization of the general structure of the MRS control system* is proposed, which makes it possible to transfer such network properties as scalability and reconfigurability to the robot modular structure.

6. The mechanism of automatic reorganization (in "plug and play" mode) of the IMCS general structure of robot with modular architecture was developed. The mechanism is based on the use of appropriate drivers – a set of software management instructions and network protocols for the application-programming interface (API).

7. A specification has been developed that is designed to create such drivers and a language for inter-module information exchange, which provide a "plug and play" mode to enable modules of third-party manufacturers. *The specification* is based on the principles of the ROS framework, but *allows to implement software on embedded computing devices*.

8. *The principle of distributed stabilized power supply* is proposed for MR modules.

9. The efficiency of using multi-criteria Pareto optimization to select a modules computing platform is shown.

10. A new *method for determining the obstacle shape according to ultrasonic distance sensors data* is proposed. The method is based on the micro-scanning of space by ultrasonic sensors (analogy with the human eyes drift).

11. The possibility of determining the obstacle shape (for simple shapes) by a sensor system, which is a ring of simultaneously launched ultrasonic sensors is experimentally shown, which allows performing a quick analysis of the environment.

For experimental verification of the proposed methods and decisions the laboratory models of mobile robots with a modular architecture were developed and manufactured (Fig.2), consisting of 4 variants of the transport module with different locomotors type (two wheel TM of different design, TM with omnidirectional wheels and walking TM), power module (PM), ultrasonic sensor module of parallel action (SRSM) and intelligent control module (ICM). Inter-module information interaction is implemented by combining the computing devices of the modules into a local area network (LAN). The models were made using modern digital technologies implemented on 3D printers, laser cutting machines and digital engraving machines. The work was carried out on the equipment and with the participation of specialists of the Institute of new educational technologies of the Russian state humanitarian University (INET RSUH).

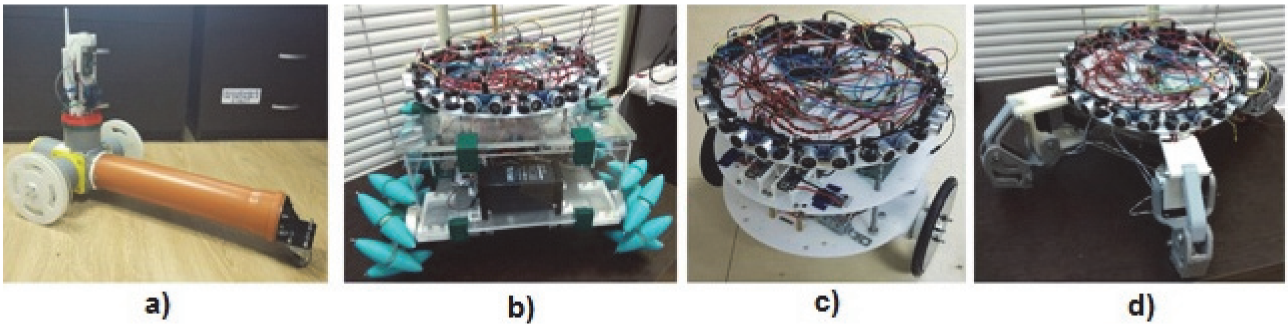


Figure 2 – Laboratory models of mobile robots with modular architecture:
a – wheel robot (option 1), b – omnidirectional robot, c – wheel robot (option 2), d – walking robot

Experimental studies of the developed methods and solutions on the MR models showed that *the proposed architecture with two-level hierarchy is not capable of implementing the control process on embedded computing devices in real-time mode*, since the functional modules control algorithms are still quite complex. The amount of data processed in modules is still large enough and the embedded systems computing power is not always sufficient to provide real-time mode. The procedure of analysis and sensory data fusion (integration) and decision-making at the system-wide control level by intelligent control module (ICM) is especially expensive. *It is necessary to find a fundamental solution to this problem*, because the modules control algorithms will become more complicated due to the tasks complexity to be solved by the MR both in autonomous and in supervisory mode and the use of better performance micro-computer will still lead to the limited functionality of the module.

A number experiment with the transport modules showed the possibility of further development of the MR modular architecture by dividing the full-featured modules into submodules and creating *a multi-level hierarchical (pyramidal) network topology of the robot IMCS*, which ensures the use of low performance microprocessors – embedded computing systems at each level of the hierarchy. We believe that the design principles of such a pyramidal structure can be based on *the multi-agent systems theory*, which will make it possible to use the achievements in this scientific field for the implementation of distributed computing in the information-measuring and control system of mobile robots.

Modular architecture and multi-agent systems

The use of multi-level pyramidal (hierarchical) topology in the build-up of information measuring and control systems of modular robotic systems is a promising area of research. If each module is equipped with its own IMCS, implemented on a separate computing device, one can distribute the computing tasks according to the modules functions. For example, the MR transport module consists of several actuator-modules that are included in the TM control system. Each actuator-module can also be designed as a separate full-featured module with its own IMCS. The microprocessor of the actuator-module control system should be responsible only for the motor operation, the parameters of which are set by the IMCS of TM: the position of its output link, the speed vector, and giving information about its current state to other modules. The sensor modules, for example, process the readings of groups of identical sensors, filter and convert a data into the required format for its transmission to the next level hierarchy modules, which perform the information fusion. The upper level hierarchy modules are responsible for complex, behavioral tasks: the mobile platform moves the robot to a given point in space and independently performs obstacle avoidance, the remote sensors module builds a distances map, etc.

Studies [22] have shown that our approach to the design of robots with modular architecture has a number of analogies with multi-agent systems, the use of which in the modular robotics field has not yet been widely reflected in the well-known publications. The principles of functional partitioning are yet to be developed, the requirements for the topology of the hierarchical network structure and interfaces of intermodule information interaction are not defined, the corresponding network protocols are not developed. Conditions and mechanisms of combination of intra-level and inter-level inter-module information interaction have not been developed at all. It is also necessary to develop requirements and conditions (including network protocols) for the network structure hierarchical organization of the IMCS, which provide the possibility to promptly reconfigure and scale of the robot modular structure depending on its purpose. Methods of remote dynamic reprogramming of robot modules computing devices are not developed.

As noted in [26] "Intelligent multi-agent systems – one of the new promising areas of artificial intelligence, which was formed on the basis of research results in the field of distributed computer systems, network technologies for solving problems and parallel computing". The proposed approach of a hierarchical organization of the control system network structure fully fits into the above definition. It has already become clear that the computational process of implementing complex interactions of an autonomous robot with the environment in real time can be implemented only through distributed computing – no, even promising, mono-computer is not able to implement such an algorithm. Therefore, we believe that the solution of this problem is possible with the use of distributed computer systems.

In the considered functional-modular architecture of the MR control system at its representation in the form of hierarchical structure an analogy with such areas of research in the information technology field as distributed information systems, computer networks, methods of artificial intelligence is clearly seen. Achievements in these areas are integrated in research, united by the common name "multi-agent systems".

Figure 3 shows a possible variant of the mobile robot IMCS architecture (shown in Fig.1) in the form of a multi-agent system built on a hierarchical principle.

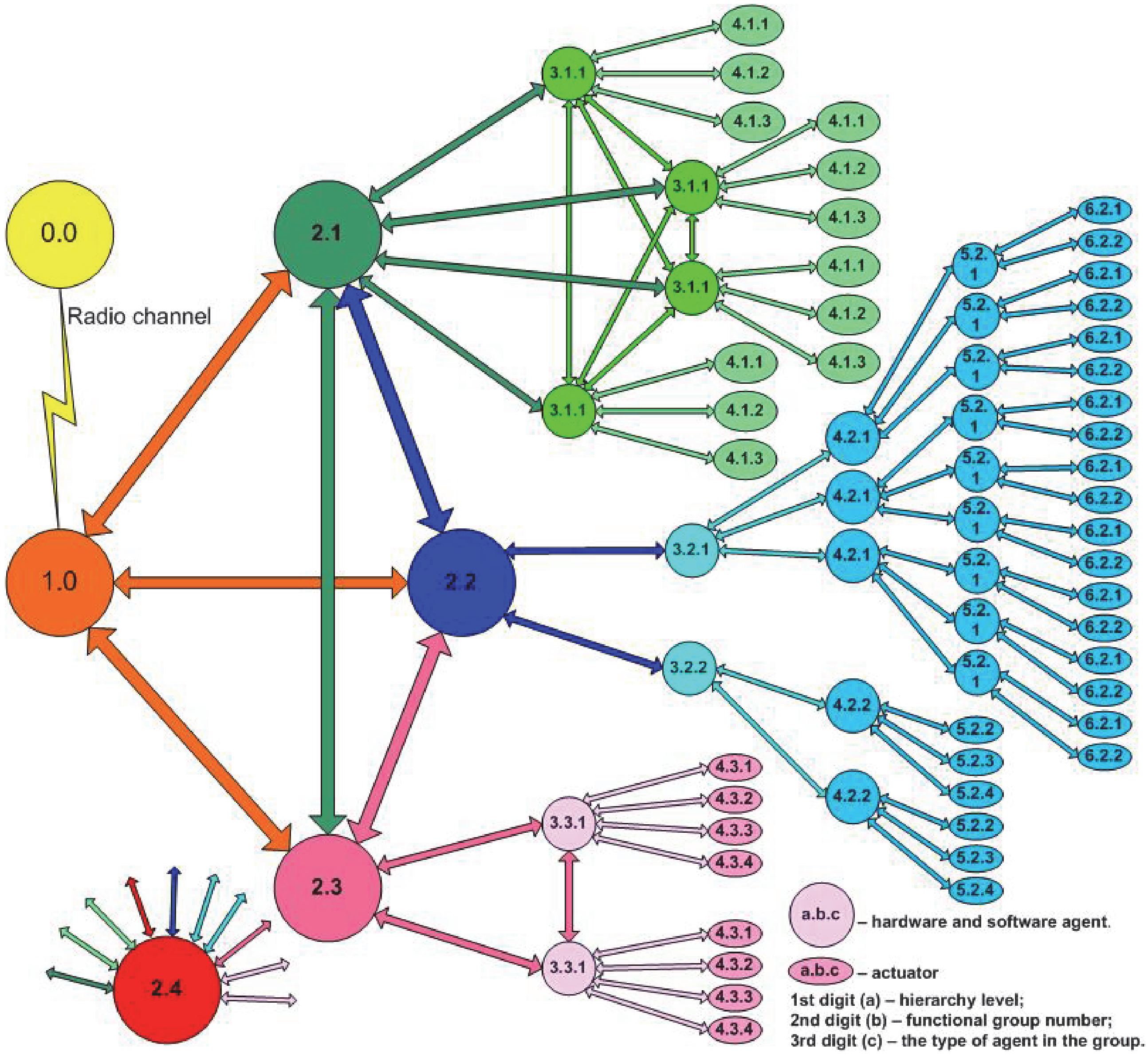


Figure 3 – Multi-agent system – information measuring and control system of mobile robot

The agent with index 0.0 is a MR control panel (for the human operator). It has information interaction via a radio channel with the agent with the index 1.0, which is a module-agent ICM (legend in Fig.1) and is located at the first level of the hierarchy. The agent ICM, in accordance with its functional purpose, generates tasks for the agent-module TM (agent with index 2.1), agent-module SM (agent with index 2.2), agent-module

SPM (agent with index 2.3) and agent-module PM (agent with index 2.4), which are located at the 2nd level of the hierarchy. All agents of the hierarchy second level must have channels of information interaction with each other. Information interaction between modules of the same level of the hierarchy, if necessary, should be carried out directly, and not via the top-level module, which allow reducing the computational load on this module and reduces traffic between hierarchy levels. Such information interaction, for example, between TM and SRSM will allow the transport module to move in a dynamic environment without the participation of the ICM, excluding possible collisions with obstacles in relatively simple cases.

Agents with index 3.1.1 form the third level of the hierarchy and represent the software and hardware modules of the locomotors control systems (wheels, legs, etc.), and the agents with index 4.1.1, which are located at the 4th level of the hierarchy are actuators equipped with appropriate internal state sensors. All agents 3.1.1 have communication channels for information interaction with each other, because in the process of TM's motion each agent needs to know the state of other agents at each moment of time, especially in the case of moving on the rough terrain. The presence of horizontal links violates the hierarchical principle, but it can be justified in the specific technical implementation of the interactions logic. The number of agents 3.3.1 can vary, for example, from 3 to 8, depending on the TM design: three-wheeled – 8-wheeled; 4-legs – 8-legs, the combination of legs-wheels, etc. In addition, the functionality of this agent can be changed – master (with actuator) or passive (sensors only).

Agents with indices 3.2.1 and 3.2.2 belong to the 3rd level of the hierarchy and represent the SRSM and LRSM (according to Fig.1). The SRSM (agent with index 3.2.1) can be implemented on ultrasonic distance sensors, which are located along the MR perimeter and form a distance map. An example of such a sensor module is given in [24, 25]. Due to the high computational complexity of simultaneous signals processing from a large number of ultrasonic sensors (agents with index 6.2.1), such processing is parallelized using a similar pyramidal architecture – agents of the 4th and 5th levels of the hierarchy (with indices 4.2.1 and 5.2.1). The SRSM can also have a line of infrared (IR) distance sensors (elements with index 6.2.2), for example, duplicating the ultrasonic sensors readings. The presence of six levels of hierarchy here may not be necessary – there may be more or less; this depends on the sensors total number, the algorithm complexity for data fusion and the microprocessor performance. In general, this signal processing hierarchical principle can be constructed in the same way as described in [27], which follows that the functional agents 3.2.1, 4.2.1 and 5.2.1, most likely, should be the same. In this case, there is no inter-module information interaction between agents of the same hierarchy level. This is explained by the sequential, level-to-level process of information processing in the absence of mutual influence of localized information; here the segmentation process is performed – the restoration of areas connectivity. However, in some cases, such interaction may be necessary, for example, in the implementation of the lateral inhibition algorithm.

The agent with index 3.2.2, i.e. LRSM, can be a stereoscopic computer vision, performing video information fusion from the left and right TV-cameras (indices 5.2.2 – 5.2.4), video data from which (for example, from three zones of each view field) are fused in agents with index 4.2.2. This hierarchical tree branch can also be extended by connecting other remote sensors, such as thermal cameras or laser scanners. It can be assumed that due to the huge flow of video information on the input of the computer vision, the hierarchy tree will consist of a much larger number of levels than shown in the figure.

Agent with an index of 2.3, which corresponds to the SPM (Fig.1), acts as a coordinator of two manipulators, controlled by its own IMCS, implemented as two identical agents with the index 3.3.1. These agents actually implement the IMCS of the multi-link manipulator (in Fig.3 – three-link manipulator with an agent-actuator of 4.3.1–4.3.3), equipped with a gripper (agent-actuator 4.3.4). Here, similar to agents with index 3.1.1, an information interaction channel between agents 3.3.1 is needed.

Agent with index 2.4, which corresponds to the PM (Fig.1), has information interaction with all modules and submodules of the hierarchy, since its function is to monitor and control the state of their power supply during the system operation, as described earlier. The main purpose of this module is to provide power to all electronic systems of modules. Therefore, this module must have direct electrical contact with each module and submodule, for example, as daisy-chain connection schema. Therefore, the information interaction of this module with the others can be technically implemented directly on the same wires as the modules power supply, bypassing the agent-modules chain. The bandwidth requirements of such a communication channel are extremely low due to the small frequency of information exchange and its small amount.

Each higher-level agent (with a less value for the index first digit) is the supervisor for the agents at the hierarchy next level.

The inter-module communication channels in this hierarchical structure of the MR control system can be based on networks of different topologies: "bus", "star" or a combination of them. The use conditions of networks certain types, topologies and network protocols have yet to be determined.

Analogies to multi-agent systems

As in multi-agent systems (MAS), in our case, knowledge and resources are distributed among sufficiently "independent" agents – modules (according to the principle of full functionality), while there is a software module of general command control – the base agent ("resident" in the terminology of multi-agent systems). Each agent-module generates and executes scripts on their own. Agents-modules react to unpredictable events by means of their sensor system, and the reaction can be independent, or be carried out in interaction with the resident [28] (for example, the human operator of a mobile robot).

"An agent is an entity that resides in some environment from which it receives data and that reflects events occurring in the environment, interprets them, and executes commands that affect the environment. The agent can contain software and hardware components… The lack of a clear definition of the agents world and the presence of a large number of attributes associated with it, as well as the availability of a wide variety of agents examples suggests that agents are a fairly common technology that accumulates several different areas» [29]. In our case, "agent" is not only a program; it is a hardware and software solution of module IMCS, which *provides its full functionality*.

According to the agent's theory, there are two definitions of intellectual agent – "weak" and "strong" [29].

Intellectual agent (IA) in a weak definition has the following properties [30]:

– ***autonomy*** – the ability of IA to function without human intervention and at the same time implement self-control over its actions and internal state;

– ***social ability*** — the ability to function in a community with other agents, exchanging messages with them using some generally understood communication language;

– ***reactivity*** — the ability to perceive the state of the environment and promptly respond to the changes that occur in it;

– ***pro-activity*** — the ability of an agent to take the initiative, i.e. the ability to generate goals and act rationally to achieve them, and not only to respond to external events.

According to above definition, the agent-module in our structure has *the autonomy property*; this follows from the previously formulated principle of module full functionality. In our modular architecture (see Fig.1) each agent-module necessarily has an information interaction channel with other agents-modules, therefore, has *the social ability property*. The only difference is that the interaction is not implemented on the principle of "all with all"; interaction is implemented on a hierarchical principle, as well as between agents of the hierarchy same level.

By means of internal state sensors, as well as remote sensors of the sensor module, with which there is information interaction, each agent-module is able to perceive the state of the external environment and promptly respond to the changes that occur in it; therefore, our hierarchical architecture has *the property of reactivity*. As for *the pro-activity property*, in our case, not all agents-modules can/should possess this property; one of the research tasks is to determine the conditions under which the agent-module can or should possess this property.

Intellectual agent in a strong definition is characterized by the presence of "at least some subset" of additional properties, called "*mental properties*" or, otherwise, *intensional concepts* [30]:

– ***goals*** — a specific set of final and intermediate states, the achievement of which the agent has adopted as the current strategy of its behavior;

– ***knowledge*** — a constant part of the agent's knowledge about itself, the environment and other agents, i.e. the part that does not change during its operation;

– ***beliefs*** — the agent's knowledge of the environment, in particular, about other agents; this is the knowledge that can change over time and become incorrect, but the agent may not have information about it and continue to remain in the belief that it is possible to base their conclusions on them;

– ***desires*** — states or situations, the achievement of which for various reasons is desirable for the agent, but they can be contradictory and therefore the agent does not expect that all of them will be achieved;

– ***intentions*** — what an agent is either obliged to do by virtue of his obligations to other agents (he is "assigned this task" and has taken on this), or what follows from his desires (i.e. a noncontradictory subset of desires, chosen for one reason or another, and which is compatible with the obligations assumed);

– ***commitments*** (obligations towards other agents) — tasks that the agent undertakes at the request (assignment) of other agents within the cooperative goals or goals of individual agents within the boundaries of cooperation.

In our hierarchical architecture, agents-modules should be referred to as "strong" IA due to their inherent mental properties. The hard distribution of modules functions indicates that the agents-modules necessarily have a specific *goal*.

In addition, agents-modules must have *knowledge about themselves and about other agents-modules* (*knowledge* — a constant part of the agent's knowledge), because only the whole system is able to solve a common task that none of the agents-modules cannot perform independently. Only in the aggregate of interactions of fully functional agents-modules the system as a whole achieves the set goal. The presence of this property in this case is also required because it is necessary to provide the "plug and play" mode to implement the dynamic reconfiguration – a promptly replacement of the agent module with a similar or modified one (for example, having a large number of mental properties). Moreover, in the reconfiguration process, each newly connected agent-module must inform all other agents-modules not only about its presence in the system, but also about its mental properties, as well as receive knowledge about other agents-modules. For example, when replacing conventional wheels in the transport module with omnidirectional wheels or "legs" (locomotors), the mental properties of the agent-TM changes (the motion control algorithm changes significantly), but the goal for the agent-TM remains the same – to achieve a given position in space. In this case, it is possible to change the algorithm of interaction of agents-locomotors. Therefore, agents-actuators should be equipped with information interaction channels with each other, and with the higher-level agent, which is a supervisor in relation to agents-actuators.

The agent-module should have the *benevolence* property – the readiness of agents to help each other to achieve a common goal and the readiness of the agent to solve exactly the tasks that the agent-supervisor assigns to him, which implies that the agent does not have *conflicting goals*. For example, when a TM moves on an uneven surface, each of the module-actuators may have at any time different conditions for movement, but for the movement of the transport module in accordance with a given speed vector, each agent-actuator needs to know the state of the other agents-actuators in order not to interfere with their behavior.

It is obvious that agents-modules have the beliefs property (knowledge about the environment and about other agents that can change over time), because we consider mobile robots that work in ever-changing environment so the knowledge about this environment is non-deterministic. This knowledge may become incorrect, for example, if any sensors or even individual agents-modules are damaged, but the agent may not have information about it and continue to be convinced that everything works fine. *Beliefs* should be based on previous experience (memory) and be probabilistic. To minimize errors, the system state should be under constant monitoring, and the failure of any system element should be accompanied by an appropriate signal – "pain". Apparently, this property can be realized with the use of such mathematical apparatus as ***pentalogy*** [31].

Due to the software and hardware implementation of agents-modules and their rigid functionality, such agent property as *mobility* is excluded, i.e. the agent's ability to migrate through the network in search of the necessary information for solving its tasks.

Conclusion

The considered functional-modular approach to the design of information-measuring and control system of mobile robots, based on the hierarchical principle of distributed computing implementation, can be presented in the form of a specific multi-agent system. The modular architecture of mobile robot control system with such properties as full functionality, promptly reconfigurability and scalability, the ability to implement the "plug and play" mode and the modules feasibility on embedded systems can be considered as a multi-agent system.

In accordance with the agents classification [26] according to the development degree of their internal understanding of the outside world and of the decision-making method, at each hierarchy level, hardware and software agents of different complexity can be used from a simple reflex agent (at the level of the transport module actuators – agent 4.1.1) to a learning agent (neural network at the level of the agent-ICM – agent 1.0). The agents use conditions of different complexity levels in this hierarchical multi-agent system will need to be determined in further research.

Summary

Proposed hierarchical topology (see Fig.3) has more similarities to a graph than a tree due to the presence of horizontal links between agents of the hierarchy same level (between agents 2.1, 2.2, 2.3; agents 3.1.1 and

agents 3.3.1). The necessity of such interaction follows from the *agents mental properties*. However, when implementing this logical structure, a number of algorithmic difficulties may arise, and such interaction will have to be organized through a higher agent – agent-supervisor in relation to the considered group of agents. This, in turn, can cause an increase in the load on the appropriate communication channels, as well as require large computing power for the agent-supervisor. Future research in this direction will allow finding compromise solutions.

Using the theory of multi-agent systems, it will be necessary to find formal models of mental properties and rules of manipulation with them to describe the pyramidal structure of distributed computing in relation to the IMCS of mobile robot. In addition, it is necessary to find a representation of the dynamic aspects of the functioning of both the individual agent-module and the community of agent-modules in mobile robots with a hierarchical modular architecture.

References

1. M-TRAN: selfreconfigurable modular robotic system / S. Murata, E. Yoshida, A. Kamimura, H. Kurokawa, K. Tomita & S. Kokaji // IEEE/ASME Transactions on Mechatronics. – 2002. – No.7(4). – pp. 432-441.
2. Design of the ATRON lattice-based self-reconfigurable robot / E.H. Østergaard, K. Kassow, R. Beck & H.H. Lund // Autonomous Robots. – 2006. – No.21(2). – pp.165-183.
3. Design of Transmote: a Modular Self-Reconfigurable Robot with Versatile Transformation Capabilities / Guifang Qiao, Guangming Song, Jun Zhang, Hongtao Sun, Weiguo Wang & Aiguo Song // Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics, 2012. – pp.1331-1336.
4. Reusable Electronics and Adaptable Communication as Implemented in the Odin Modular Robot / Ricardo Franco Mendoza Garcia, Andreas Lyder, David Johan Christensen & Kasper Stoy // IEEE International Conference on Robotics and Automation, 2009. – pp.1152-1158.
5. Concept of cellular robotic system (CEBOT) and basic strategies for its realization / Toshio Fukuda, Tsuyoshi Ueyama, Yoshio Kawauchi, Fumihito Arai // Computers Elect Engng. 1987. vol.18, no.1. – pp.11-39.
6. Baca J. A heterogeneous modular robotic design for fast response to a diversity of tasks / Baca J., Ferre M., Aracil R. // Robotics and Autonomous Systems. 2012. Vol.60. No.4. – pp.522-531.
7. On sub-modularization and morphological heterogeneity in modular robotics / A.H. Lyder, K. Stoy, R.F. Mendoza-Garcia, J.C. Larsen & P. Hermansen // Intelligent Autonomous Systems of Advances in Intelligent Systems and Computing. Springer Berlin Heidelberg. 2013. Vol.193. No.12. – pp.649-661.
8. Hancher M.D., Hornby G.S. A modular robotic system with applications to space exploration // 2nd IEEE International Conference on Space Mission Challenges for Information Technology (SMC-IT'06). Pasadena, CA: Publisher «IEEE». 2006. – pp.132-140.
9. The shift in the robotics paradigm — the Hardware Robot Operating System (H-ROS); an infrastructure to create interoperable robot components / V. Mayoral, A. Hernandez, R. Kojcev, I. Muguruza et al. // NASA/ESA Conference on Adaptive Hardware and Systems (AHS), Pasadena, CA. – 2017. – pp.229-236.
10. AMiRo: a modular & customizable open-source mini robot platform / S. Herbrechtsmeier, T. Korthals, T. Schopping, U. Ruckert // 20th International Conference on System Theory, Control and Computing (ICSTCC), Sinaia. – 2016. – pp.687-692.
11. R2P: An open source hardware and software modular approach to robot prototyping / A. Bonarini, M. Matteucci, M. Migliavacca, D. Rizzi // Robotics and Autonomous Systems. – 2014. – No.62. – pp.1073-1084.
12. Distributed and modular CAN-based architecture for hardware control and sensor data integration / D.P. Losada, J. L. Fernández, E. Paz, Rafael Sanz // Sensors. – 2017. – No.17. – pp.1013-1030.
13. Andreev V.P., Poduraev Yu.V. Functional–Modular Design of Heterogeneous Mobile Robotic Systems // Extreme robotics (ER-2016). Proceedings of the International Scientific and Technological Conference. – Saint-Petersburg: «AP4Print», 2016. – pp.39-44.
14. Andreev V., Poduraev Y. Network–based Design of Heterogeneous Modular Mobile Robotic Systems // 27th International DAAAM Symposium on Intelligent Manufacturing and Automation 2016, Proceedings of a meeting held 26-29 October 2016, Mostar, Bosnia and Herzegovina, B. Katalinic (Ed.), ISBN 978-1-5108-3300-5. Curran Associates, Inc., NY 12571 (Jan 2017). – pp.0004-0009.
15. Andreev V. The principle of full functionality – the basis for rapid reconfiguration in heterogeneous modular mobile robots / Andreev V., Kim V., Pletenev P. // Annals of DAAAM and Proceedings of the 28th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2017; Zadar; Croatia, B. Katalinic (Ed.), Published by DAAAM International, ISBN 9781510853270, ISSN 1726-9679, Vienna, Austria. Curran Associates, Inc. (Feb 2018). – pp.0023-0028.
16. Andreev V.P. The principle of the full functionality of modules in heterogeneous modular mobile robots / Andreev V.P., Kim V.L., Pletenev P.F. // Extreme robotics (ER-2017). Proceedings of the International Scientific and Technological Conference. – Saint-Petersburg: «AP4Print», 2017. – pp.81-86.

17. Andreev V., Kim V. Control system and design of the motion module of a heterogeneous modular mobile robot // Proceedings of the 27th DAAAM International Symposium, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-08-2, ISSN 1726-9679, Vienna, Austria, 2016. – pp.0586-0594.
18. Andreev V.P., Kim V.L. Development of Functional Units of a Heterogeneous Modular Mobile Robot // Extreme robotics (ER-2016). Proceedings of the International Scientific and Technological Conference. – Saint-Petersburg: «AP4Print», 2016. – pp.359-363.
19. Andreev V., Kim V. The application of artificial vector fields for motion control of a heterogeneous modular mobile robot // Annals of DAAAM and Proceedings of the 28th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2017; Zadar; Croatia, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-1-510853-27-0, ISSN 1726-9679, Vienna, Austria. Curran Associates, Inc. (Feb 2018). – pp.0635-0644.
20. Andreev V.P., Pletenev P.F. Method of Information Interaction for Distributed Control Systems of Robots with Modular Architecture. SPIIRAS Proceedings. Saint-Petersburg - 2018. Issue 2(57). ISSN 2078-9181 (print), ISSN 2078-9599 (online). – pp.134-160.
21. Andreev V., Pletenev P. Organizing Intermodular Communication for Heterogeneous Modular Mobile Robot. Annals of DAAAM and Proceedings of the 28th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2017; Zadar; Croatia, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-1-510853-27-0, ISSN 1726-9679, Vienna, Austria. Curran Associates, Inc. (Feb 2018). – pp.0474-0480.
22. Andreev V. Hardware & Software Solution for Rapid Reconfiguration of Heterogeneous Robots / Andreev V., Kim V., Pletenev P. // Mekhatronika, Avtomatizatsiya, Upravlenie, 2018, no.6, v.19, – pp.387-395.
23. Andreev V., Kim V., Pletenev P. Using Pareto Optimum to Choose Module's Computing Platforms of Mobile Robot with Modular Architecture. Annals of DAAAM and Proceedings of the 29th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2018; Zadar; Croatia, B. Katalinic (Ed.). Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-9679, Vienna, Austria, v.29(1), 2018. – pp.0559-0565.
24. Andreev V.P., Tarasova V.E. Determination of the Form of Obstacles by a Mobile Robot Using Scanning Angular Movements of Ultrasonic Sensor // Mekhatronika, Avtomatizatsiya, Upravlenie, 2017, no.11, v.18, – pp.759-763.
25. Andreev V., Tarasova V. Identification of the Obstacle Shape Using the Ultrasonic Sensors Module of Modular Mobile Robot. Annals of DAAAM and Proceedings of the 29th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2018; Zadar; Croatia, B. Katalinic (Ed.). Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-9679, Vienna, Austria, v.29(1), 2018. – pp.1001-1009.
26. https://intellect.icu/11-multiagentnye-sistemy-5354. Multi-agent systems. Basic concepts of the theory of agents (Rus). Accessed on: 2019-03-18.
27. Andreev V.P. On the possibility of constructing a technical visual system based on a parallel computing environment // Collection of articles "Methods and models for controlling robots and manipulators in production and research". Moscow House of scientific and technical propaganda. F. E. Dzerzhinsky, 1979. – pp.76-82.
28. http://www.intuit.ru/studies/courses/11068/1102 /lecture/17391. NOU INTUIT. Multi-agent technologies (Rus). Accessed on: 2019-03-18.
29. Wooldridge M., Jennings N.R. Agent Theories, Architectures, and Languages: A Survey. In: Intelligent Agents. ECAI-94 Workshop on Agent Theories, Architecture and Languages. Amsterdam, The Netherlands, August 8-9, 1994, (Eds. M.J.Wooldridge and N.R.Jennings). Proceedings. Springer Verlag, 1994. – pp.3-39.
30. https://refdb.ru/look/2448008-pall.html. Gorodetsky V.I., Grushinsky M.S., Khabalov A.V. Multi-agent system (review) (Rus). Accessed on: 2019-03-18.
31. Technology of Multi-Agent Control for Industrial Automation with Logical Processing of Contradictions / Pryanichnikov V.E., Aryskin A.A., Eprikov S.R., Kirsanov K.B., Khelemendik R.V., et al. // Annals of DAAAM and Proceedings of the 28th International DAAAM Symposium on Intelligent Manufacturing and Automation, DAAAM 2017; Zadar; Croatia, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-1-510853-27-0, ISSN 1726-9679, Vienna, Austria. Curran Associates, Inc. (Feb 2018). – pp.1202-1207.

V.Ya. Vilisov, B.Yu. Murashkin, A.I. Kulikov

SIMULATION MODEL OF TWO-ROBOT COOPERATION IN COMMON OPERATING ENVIRONMENT

University of Technology, Korolyov, Moscow region, Russia
vvib@yandex.ru

Abstract

The article considers a simulation modelling problem related to the chess game process occurring between two three-tier manipulators. The objective of the game construction lies in developing the procedure of effective control of the autonomous manipulator robots located in a common operating environment. The simulation model is a preliminary stage of building a natural complex that would provide cooperation of several manipulator robots within a common operating environment. The article addresses issues of training and research.

Keywords: manipulator robot, chess, operating environment, simulation model.

Introduction

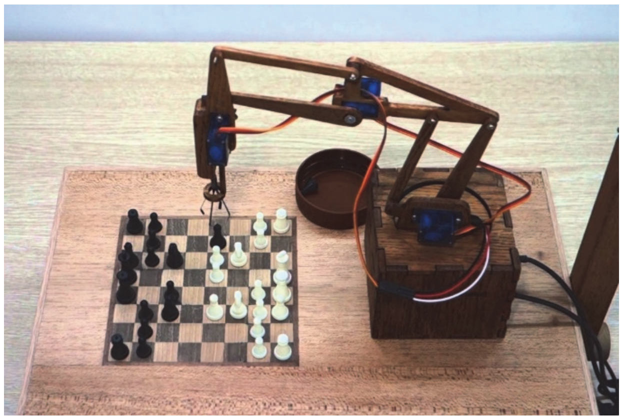
Modern robotics has a wide range of tasks where robots are to perform various manipulations with objects. These applications include: products' assembly, room cleaning, cargo moving etc. [1, 2].

As a rule, university students of robotics study mostly construction of separate robotic units and their control. Thus, little attention is paid to the issues of robots' cooperation in groups. The main focus of the work is the construction of models and robotic complexes that would allow obtaining necessary research and development skills related to the robot cooperation in groups and their coordinated actions to solve the problems that require participation of several robots and/or heterogeneous "robot-human" groups.

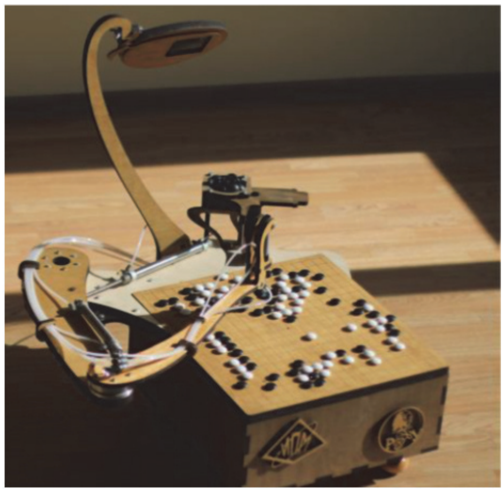
The important subset of such problems is comprised of the operations that are executed by several robots or in robot-human groups. They are also called "collaborative robots" [3, 4]. The main requirement set for such robot-robot groups or robot-human groups is taking into account other group members or coordination of their engagement. Thus, in various assembly operations, robots must adhere to a certain sequence of actions. Therefore, the chess game between two manipulator robots may become a useful model for developing the robot operation algorithms, their sequence of actions etc.. Moreover, using the framework of the model, it is also possible to work over various types of grips and/or to optimize the control system used for grabbing different objects etc. [2]. There are many publications that are devoted to the above-mentioned research [1-4].

Usually, the first stage of solving such problems would include mathematical modelling [4], in particular – simulation modelling. Here we have demonstrated a simulation model of cooperation between two manipulator robots within the chess game environment. The simulation is based on the geometrical dimensions of the robots' tiers, chess game logic and time required to make a move by each robot competitor. Statistical time characteristics pertaining to the moves of each robot may be changed as a part of the simulation model.

Currently, there are several implementation types of mechatronic devices [5-12] that can move chess pieces on the board and choose the move (see Fig. 1). Industrial robots are also used when moving the chess pieces in a robot-human game (see Fig. 1c) or in other games like Go [5] (see Fig. 1d). However, the work is focused on the chess game between two manipulator robots. This problem is a part of a wider research on a robot-robot and robot-human cooperation as well as the robot training.



a) b)



c) d)

Figure 1 – Mechatronic players

In order to simulate the chess game between two manipulator robots, we have developed a program that implements and demonstrates the process as a sequence of 3D stages.

Structure and Combination of Simulation Model

The program has been developed in *C#* (C Sharp) using ***Unity***, a cross-platform environment for the computer game development [13]. The structure of main program elements is provided on Fig. 2.

Main program file (*Main.cs*) contains settings, interface elements, decoder, separator and program's logic control elements. The decoder transforms chess game text file *Game.txt* (Fig. 3) where the moves are recorded in the way that is customary for users, into the numerical form (file *Game_Symbols.txt*).

The separator splits the *Game_Symbols.txt* file for each manipulator. Thus, *PlayerWhite.txt* file contains all odd array elements (moves of the player with the white chess pieces) that should be executed by the manipulator robot 1 (MR1), while *PlayerBlack.txt* file is created for MR2, containing all even array elements, i.e. the moves of the player with the black chess pieces.

Child files of *Main.cs* script, i.e. scripts *PlayerWhite.cs* and *PlayerBlack.cs* read the data from the files *PlayerWhite.txt* and *PlayerBlack.txt* accordingly, receiving the "next move" message from the script *Main.cs* and sending data regarding the players' moves together with the "next move" request to *BattleField.cs* script.

Game object *BattleField* contains script file *BattleField.cs* together with its child game objects, where each of them contains an additional script file *KubeScript.cs* that corresponds to a specific game field square and possesses a square-specific name. Moreover, we have created scripts for the game objects *Manipulator1* and *Manipulator2,* acting as foundations for their child objects *Detal1, Detal2, Detal3* and *Detal11, Detal22, Detal33.* Script files *Manyapul1.cs* and *Manyapul2.cs* were created specifically for these objects. Afterwards we have created game objects for the chess pieces that contain script file *FigureScript.cs* and which are child game objects of *WhitePlayer* and *BlackPlayer*.

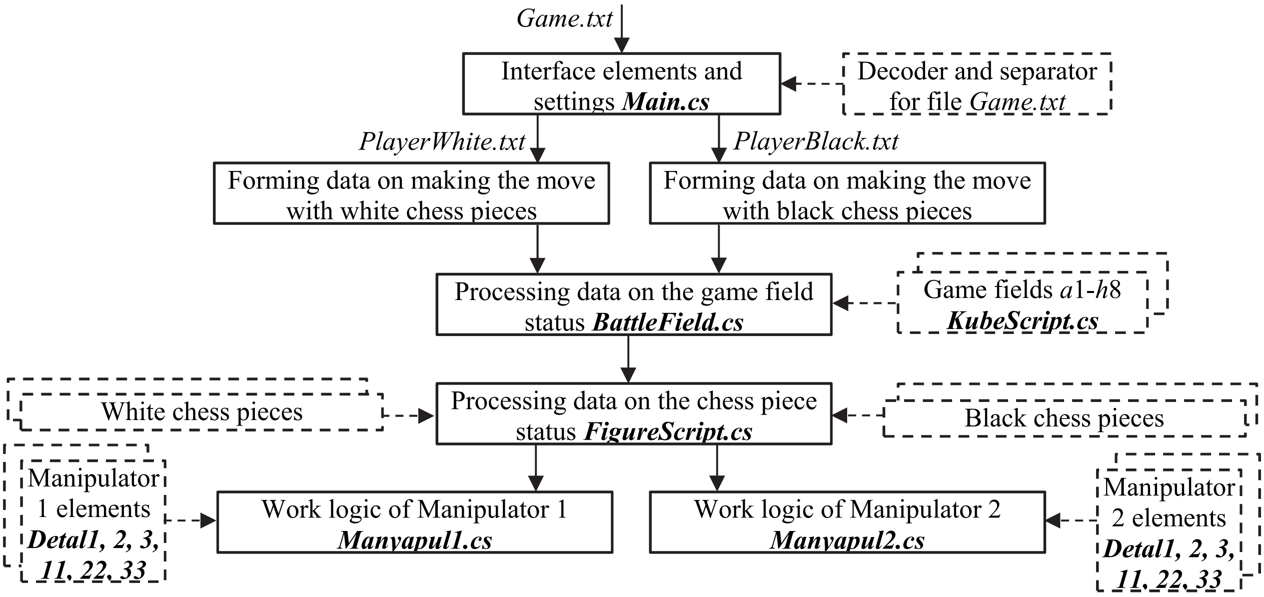


Figure 2 – Simulation program object structure

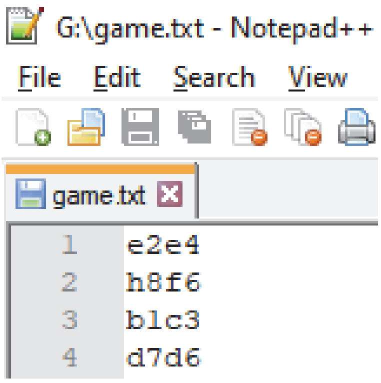


Figure 3 – Part of a game's simulation file

Algorithm and Program Implementation

Main program work is set and controlled by script *BattleField.cs* that receives requests to make the move and to use the value of variables that contain information about the move. Depending on who is going to make the move (which object sent the request), the script provides the manipulator robots MR1 and MR2 with values of variables and the request to make a move.

Depending on the current code, the scripts *Manyapul1.cs* and *Manyapul2.cs* receive the request to make a move together with current values of variables. Based on the values of variables, the scripts provide the manipulator's elements with the values which shall be used to open its tiers in order to move the grip to the required square. Then the script launches the turn function *PovorotFrom*, which is responsible for the manipulator's turn, further informing the chess piece's script about its move and after that launching the function *PovorotTo*.

PovorotTo function provides manipulator's elements with the values, according to which they should be turned in order to move the grip towards the target square, then sending a moving message to the chess piece's script file and putting the manipulator into motion. Afterwards the program delays the manipulator's movement for the time that can be set by the vertical sliders located on the screen; then *DeFolt* function is launched. *DeFolt* function provides the manipulator elements with the source location values, performing the manipulator's turn and sending the move-end message to the script file *Main.cs*.

Program scripts contain several service functions, whose values we are not going to describe here due to a limited size of the publication.

After launching the program we can see a configuration window, where we can make changes in the current settings of graphic interface and control elements. The settings are saved in the configuration file. The game window is opened after pressing Play button (see Fig. 4).

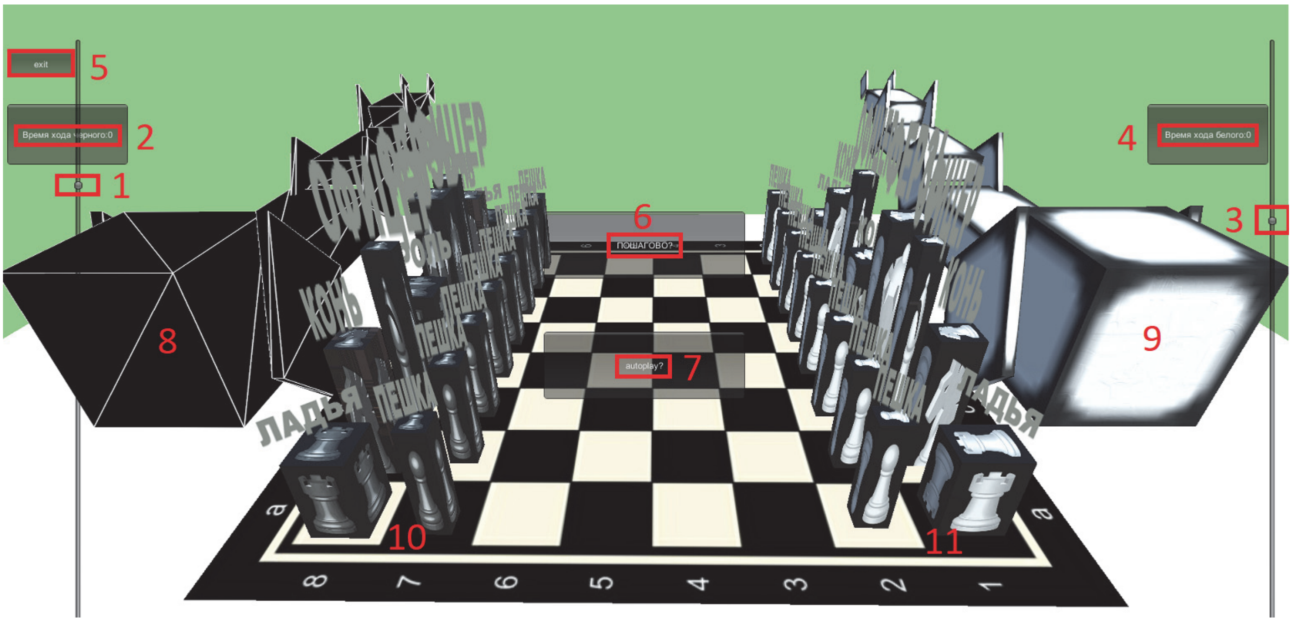


Figure 4 – Game window. Source location of chess pieces and control elements

In the current version of the simulation program we can only replay the sequence of moves that were recorded previously in the file *Game.txt*. Should the user wish to replay his/her own game, he/she should create his/her own file *Game.txt* and place it into the folder that contains the executable file *chessgame.exe*.

In the game window the user can access the following control elements together with the data display fields (see Fig. 4):

– The slider that changes movement speed of the black manipulator (see red mark 1); the higher the slider is, the faster the manipulator moves (black - mark 8, white - mark 9).

– The slider that changes movement speed of the white manipulator (mark 3); the higher the slider is, the faster the manipulator moves.

– *Exit* button (mark 5), pressing which the user stops the application.

– POSHAGOVO? (STEP BY STEP?) button (mark 6) launches the game in a step-by-step mode, i.e. the manipulators shall stop after each step.

– *Autoplay?* button (mark 7) launches the game in an automatic mode, where the manipulators shall make moves one after another until the game is over.

– Khodim? (Go?) button (see mark 1 on Fig. 5) appears only when the step-by-step mode has been selected; if the move has already been made, the next move shall be made only when pressing the button.

– Nachnem? (Start?) button appears after the launch game option was chosen, the first move is made after pressing it.

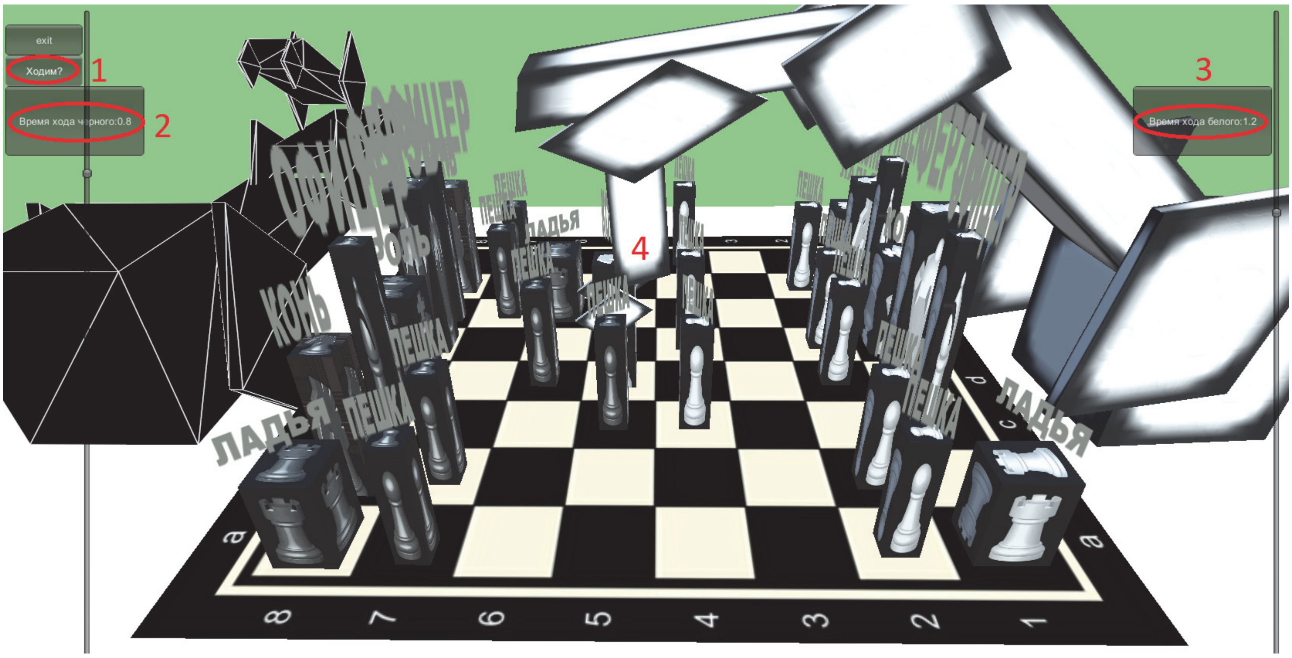
– Text fields "Black Manipulator's Move Time: 0.8" (mark 2 on Fig. 5) and "White Manipulator's Move Time: 1.2" (mark 3 on Pic. 5) display duration of the last move, made accordingly by black and white MRs.

– In the beginning of the game the chess pieces are in their source locations, i.e. the black are on the left (mark 10) and the white are on the right (mark 11).

Picture 5 shows game position after the manipulators have made several moves. The manipulator with white chess pieces (mark 4) is completing its move.

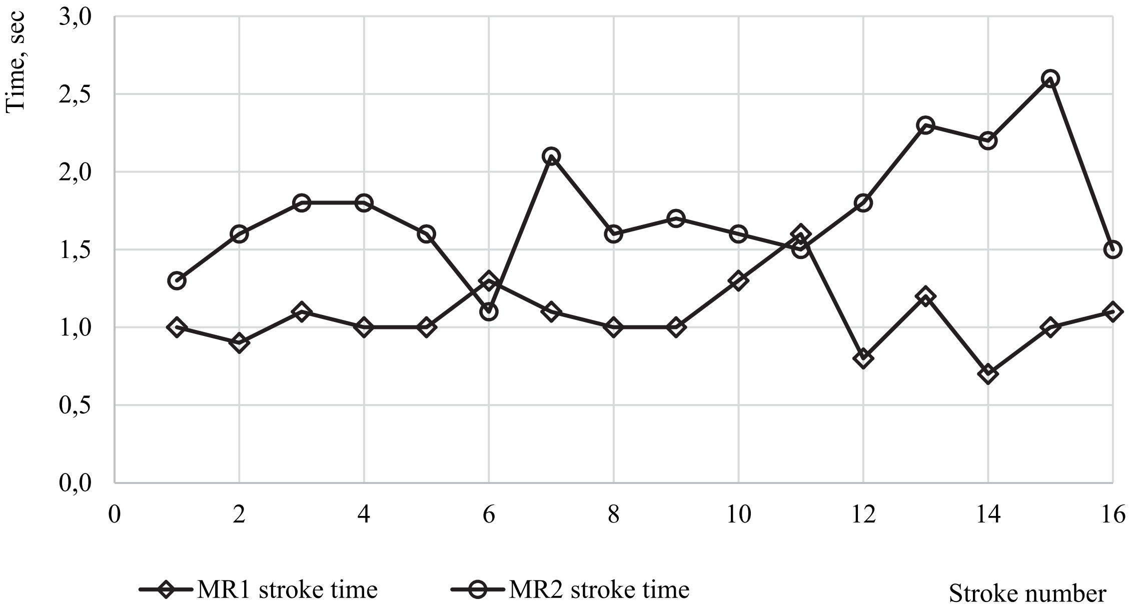
The program provides accumulation and output of measurable data per each simulated game. In particular, it is possible to measure the length of the grip trajectory as well as the time required for each move.

These indicators can be used to optimize the movement of manipulators, for example in relation to the operating speed and/or power consumption. Picture 6 shows registered duration of each move made by the manipulators.



Picture 5 – Intermediate position of the game

Statistical characteristics of this data are the following: average value of MR1 making its move is 1.07 sec, MR2 - 1.76 sec, while the correlation coefficient value of these two datasets constitutes $r_{12} = -033$. These three indicators can be used when solving the problem of manipulator optimization, i.e. tier size, location of their bases, choosing the type of a servo drill etc.. It is also necessary to mention that this model can be launched in a virtual mode (without screen image but with measurement of all parameters) thus giving a possibility to apply optimization search algorithms.



Picture 6 – Move duration per each manipulator

Results and Discussion

Provided simulation game model is only a part of the simulation & nature robotic complex, which shall be used for further research, working over the construction elements as well as the work algorithms.

Nowadays, the work of autonomous manipulator robots or android robots in a non-determined environment is a highly sought after and necessary topic [1, 3, 4]. It can be applied at the space station when assembling its construction elements, at other planets or on the ground when building various objects etc.. At that, the operating range of MR can vary greatly: from the static set of homogeneous objects to a large variety of diverse dynamic objects [4]. Here, the variety of possible operating environments is also determined by the fact that MRs can be both stationary or moving, located in a stable and spatially-linked group or as a totality of independently moving MRs.

The chess context allows working over structure and combination of algorithmic and program elements, their cooperation as well as optimizing the parameters.

At the next stage we expect to add a free chess engine to the current simulation model [14]. It is obvious that a full-function simulation model of MR playing chess does not allow working over some situations that may occur in real conditions of the above-mentioned spheres. Thus, the model does not include sensory capabilities which are required when manipulating the objects in a real time. Therefore, the natural part of the model complex shall include stage surveillance cameras, corresponding situation (condition, scene) identification algorithms, as well as other means of controlling the work area.

Apart from that, the environment of the observed simulation & nature complex shall include various grip options for various objects (chess pieces): two- and three-finger grips (also the collet type) as well as a five-finger anthropomorphic hand [15-18]. It would also be interesting to perform a research on the grip construction, including the feedback on strengthening the grip with installed strain indicators and actuators, installed in the human's operating gear.

Apart from the autonomous work of the group of (now only two) MRs, there should also exist a possibility to implement a human-machine mode where one MR is to be controlled by the human operator using various interface options, including somatosensory glove (Exoskeleton), gyroscope sensors, neurobionical helmet and neurosensors on the hand. Here, the computing environment shall include elements of Arduino family [19] (connected to the personal computer) and the robot development environment EZ-Robot [20]. These options allow implementing control algorithms practically of any complexity.

Human-machine mode would also give a possibility of working over some collaborative work modes of MRs [3].

The simulation environment shall provide a possibility of working on adaptive training algorithms where MR would perform actions based on observation (registration) of the human operator's actions, for instance using a somatosensory glove [19] or video cameras.

The above-mentioned functional capabilities of the created simulation & nature robotic complex are mostly focused on the training objectives applied from college level to the doctoral studies. However, its modular nature together with the possibility to replace parts of real elements with the simulation modules allows performing research that would also prove useful to the application in practice.

Conclusions

1. The developed totality of the chess game simulation modelling proved the possibility of building the cooperation model for manipulator robots in a chess game. Visual demonstration of the game course shows the cooperation of the manipulators, while the control elements provide a possibility of showing the results as it is desired, when changing some modelling parameters.

2. The modelling program is built in such a way that it allows performing different kinds of research of the manipulator robot cooperation, in particular, changing time characteristics of each MR's actions.

References

1. Capable of working with heavy materials at construction sites (HRP-5p). URL: https://youtu.be/fMwiZXxo9Qg.
2. Autonomous Tidying-up Robot System. URL: https://projects.preferred.jp/tidying-up-robot/en/
3. A.S. Yushchenko, K.V. Ermishin. Collaborative Mobile Robots - a New Stage of Robotics // Works of International Scientific and Technological Conference "Extreme Robotics". Saint Petersburg: AP4Print LLP. 2016. p. 480. pp. 455-459.
4. I.A. Kalyayev, A.R. Gaiduk, S.G. Kapustyan. Models and Algorithms of Collective Control in Robotic Groups. - Moscow: Fizmatlit. - 2009, p. 280
5. V.Ye. Pavlovsky, A.V. Podoprosvetov, V.S. Smolin, I.A. Orlov. ManGo Manipulator with Intelligent Technologies // Intellectual Systems and Technologies: Modern Condition and Prospects. Collection of Scientific Papers of IV International Summer Seminar School on Artificial Intelligence (Saint Petersburg: June 30 - July 30, 2017). Saint Petersburg: Politeknika-Servis. 2017. p. 224. pp. 128-137.
6. ChessBot - wooden chess-playing robot. URL: https://youtu.be/tmG-FJrXAj8
7. Chess. Human and robot. URL: https://youtu.be/V9xFqpHNypM
8. Robot «Kuka» - Robot «Chesska» Game 1. URL: https://youtu.be/TFeXUjQ87Ps
9. Lego Mindstorms NXT 2.0 - Chess playing robot - Charlie. URL: https://vk.com/video-100375733_171533035

10. Dvorkovich played chess with a robot. URL: https://youtu.be/FYSOlEoc25s
11. REEM-A, humanoid robot. Playing chess URL: https://youtu.be/2wzT4vafXOA
12. 7Bot Desktop Robot Arm playing chess with human. URL: https://youtu.be/OHazT3y0WpI
13. Unity, a cross-platform environment for the computer game development. URL: https://unity.com
14. Chess engines. URL: http://chess-boom.online/shahmatnye-dvizhki/
15. Prosthetic devices (Rocket and Space Corporation Energia). URL: https://www.energia.ru/ru/conversion/prosthetic/pvk/pvk-02.html#
16. Arm Dynamics. URL: https://www.armdynamics.com/
17. Motorica company. URL: https://motorica.org/
18. BeBionic. URL: http://bebionic.com/
19. Arduino. URL: http://arduino.ru
20. EZ-Robot. URL: http://www.ez-robot.com

V.I. Shiryaev, D.P. Klepach, D.O. Malyugina, A.A. Romanova

ABOUT GUARANTEED ESTIMATION OF THE LINEAR DYNAMICAL SYSTEM STATE VECTOR IN THE CONDITIONS OF UNKNOWN INPUT

Federal Autonomous Educational Institution of Higher Professional Education South Ural State University (National Research University), Chelyabinsk, Russia,
shiriaevvi@susu.ru, klepachd@mail.ru

Abstract

We consider the construction of guaranteed estimates of the dynamical system state vector in the conditions of indeterminacy, when initial state statistics, process noise, measurement noise is absent and only the sets of their possible values are known. The algorithm for constructing the approximation of information sets is described, when the sets of possible values of the initial state, the process noise, and the measurement noise are polyhedron set by the systems of linear inequalities. The operation of the algorithm is demonstrated by the example of a two-dimensional model. A comparison is made between the approximations of information sets and the Kalman filter estimates.

Keywords: approximation, guaranteed estimates, dynamical systems, information sets, linear inequalities, minimax filters, the conditions of unknown input.

Introduction

The task of estimating the state of dynamic systems arises in various technical applications, such as aircraft and spacecraft control systems, tracking and target detection systems, automated process control systems, etc. [8,13-20,22]. There are two different approaches to estimating the state vector of linear dynamic systems: probabilistic approach, for example, Kalman filter [21,23], and guaranteed approach, for example, minimax filter [2,4-7,9-11]. The Kalman filter is based on the assumption that the statistical characteristics of the process noise and the measurement noise influencing the system are known. However, in real conditions, statistical information about process noise may be absent and reduces either to specifying the corresponding areas of their change or to setting a whole class of permissible distribution functions that determine process noise realizations if the latter are of a statistical nature, therefore the use of the Kalman filter may not be justified [3] . Then the estimation problem is considered in guaranteeing or minimax setting [2,4-6,9-11].

This paper describes a procedure of constructing an approximation of information sets from above. A comparison of the estimates of the Kalman filter and the minimax filter for different implementations of the process is given. Work continues the researches [7,9,10].

Minimax filter

Let us consider the problem of estimating the state of a dynamic system, when statistical information about the process noise and the measurement noise influencing the system is missing, but there are many known possible sets of their values. The equations of motion and measurements in a linear approximation are set

$$x_{k+1} = Ax_k + Bu_k + \Gamma w_k, k = 0,1, \dots, N-1, \tag{1}$$

$$y_{k+1} = Gx_{k+1} + Hv_{k+1}, \tag{2}$$

where $x_k \in R^n, w_k \in R^{n_w}, y_{k+1} \in R^m, v_{k+1} \in R^{m_v}$ – are the vectors of state, process noise, measurement, measurement noise, respectively, A, B, Γ, G, H – are known matrices of corresponding dimensions, u_k – given control.

The initial state x_0, the process noise w_k, the measurement noise v_k are known to be able to take at any k-th instant of time any value from the sets

$$x_0 \in X_0, w_k \in W, v_{k+1} \in V, k = 0,1, \dots, N-1, \tag{3}$$

which are given in the form of convex polyhedron, which are known.

It is known [2] that the result of the guaranteed estimation of the state vector x_{k+1} of systems (1) - (3) is the information set $\bar{X}_{k+1}$. Based on the previous information set $\bar{X}_k$ prediction locus state vector x_{k+1} system is calculated

$$X_{k+1/k} = A\bar{X}_{k+1} + Bu_k + \Gamma W. \tag{4}$$

According to the measurement results y_{k+1} set of states compatible with measurement are calculated

$$X[y_{k+1}] = \{x \in R^n | Gx + Hv = y_{k+1} \; \forall \, v \in V\}. \tag{5}$$

Then we find the information set as the intersection area of the prediction locus and the set that is compatible with the measurements

$$\bar{X}_{k+1} = X_{k+1/k} \cap X[y_{k+1}]. \tag{6}$$

The operations in (4) - (6) are performed on sets: a linear transformation, the sum of sets in the Minkowski sense, and the intersection of sets. Let us consider an example to clarify operations (4) - (6) in the case when the sets X_0, W, V are given as polyhedron.

Example 1

The matrices in the system (1), (2) have the value:

$$A = \begin{pmatrix} 1 & 0.01 \\ 0 & 1 \end{pmatrix}, \Gamma = (0 \quad 2 \cdot 10^{-4})', B = (0 \quad 0.01)', \; G = I_{2x2}, \qquad H = I_{2x2}.$$

The sets X_0, W, V and the corresponding values of the variables x_0, w_k, v_k are given (fig. 1). An example of building an information set is given $\bar{X}_1$ (fig. 2), for u_0 accepted $u_0 = 0$.

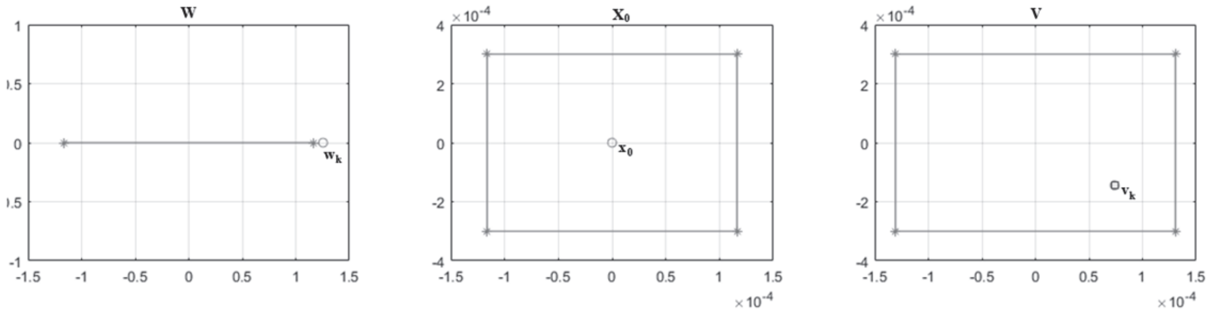


Figure 1 – Sets X_0, W, V

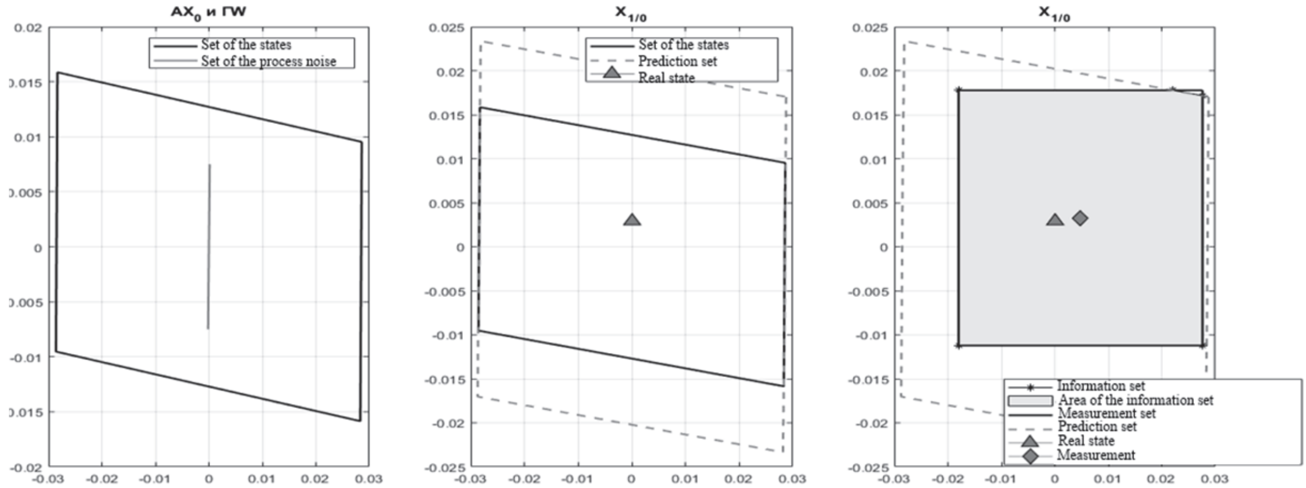


Figure 2 – An example of building an information set

As a rule, the form of information sets $\bar{X}_{k+1}, k = 0,1, \ldots$ is quite complex, that is, the set can contain a large number of vertices and faces, and operations on sets are computationally complex for systems of large dimensionality in real time . Instead of the exact construction of information sets $\bar{X}_{k+1}$ use [7] the approximate value $\bar{X}_{a\,k+1} \supseteq \bar{X}_{k+1}$, i.e. approximation from above.

Approximation of information sets

We present the procedure of approximation of the information set without constructing the sum and the intersection of the sets. It is proposed to build an estimate $\bar{X}_{a\,k+1}$ of the information set $\bar{X}_{k+1}$ in the form of a convex polyhedron, obtained by approximating from above the "exact", but implicitly given by the system of linear inequalities of the information set obtained for the system (1) - (3). When the initial constraints (3) on x_0, w_k, v_k , $k = 0,1, \ldots$ are polyhedrons, they can be set by systems of linear inequalities

$$x_0 \in X_0\colon A_{x_0} x_0 \le b_{x_0}, w_k \in W\colon A_w w_k \le b_w, v_k \in V\colon A_v v_k \le b_v, k = 0,1, \ldots, N-1. \tag{7}$$

The equations of the system (1), (2) together with the system of linear inequalities (7) describe the state of the dynamical system at the k-th step. In this case, the approximate information set $\bar{X}_{a\,k+1}$ at the k-th step will be set by the system of linear inequalities

$$\bar{X}_{a\,k+1} = \{x_k \in R^{n_x} \colon A_{x_k} x_k \le b_{x_k}\},\ x_k \in \bar{X}_{a\,k+1}, \tag{8}$$

Then the task of building an information set is reduced to solving a system of linear equations (1), (2) and inequalities (7) with respect to the estimated variable x_{k+1}, the problem of estimation in this case is reduced to a linear programming problem.

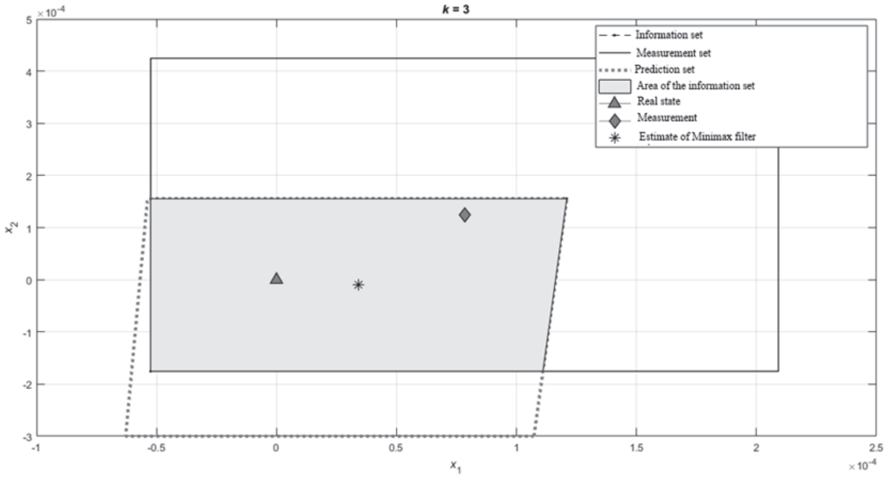


Figure 3 – Approximation of the information set

Let us consider an algorithm describing the approximation of the information set $\bar{X}_{k+1}$ based on the current measurement of y_{k+1} and the known value of the control u_k. The vectors $x_{k+1}, x_k, w_k, v_{k+1}$ are unknown. The equations of the model (1), (2) the unknown variables will be moved to the one part, and the known - to another. We get a system of linear equations describing the model

$$\begin{cases} x_{k+1} - Ax_k - \Gamma w_k = Bu_k, \\ Gx_{k+1} + Hv_{k+1} = y_{k+1}. \end{cases} \tag{9}$$

In matrix form, the system of equations (9) can be represented in the following way

$$\begin{pmatrix} I & -A & -\Gamma & 0 \\ G & 0 & 0 & H \end{pmatrix} \begin{pmatrix} x_{k+1} \\ x_k \\ w_k \\ v_{k+1} \end{pmatrix} = \begin{pmatrix} Bu_k \\ y_{k+1} \end{pmatrix}. \tag{10}$$

We get the system of linear inequalities from the constraints (7)

$$\begin{cases} A_{x_k} x_k \le b_{x_k}; \\ A_w w_k \le b_w; \\ A_v v_k \le b_v. \end{cases} \tag{11}$$

In matrix form

$$\begin{pmatrix} 0 & A_{x_k} & 0 & 0 \\ 0 & 0 & A_w & 0 \\ 0 & 0 & 0 & A_v \end{pmatrix} \begin{pmatrix} x_{k+1} \\ x_k \\ w_k \\ v_{k+1} \end{pmatrix} \le \begin{pmatrix} b_{x_k} \\ b_w \\ b_v \end{pmatrix}. \tag{12}$$

We approximate the information set $\bar{X}_{k+1}$ by the polyhedron $\bar{X}_{a\,k+1}$ with a set of faces, a vector of normals forming the rows of the matrix $A_{x_{k+1}}$, i.e. we obtain the polyhedron (8). For each direction a_i, (the i-th row of the matrix $A_{x_{k+1}}$) we solve the linear programming problem

$$x_{k+1}^* = argmax\langle a_i, x_{k+1}\rangle \tag{13}$$

under constraints (10), (12), where $\langle a_i, x_{k+1}\rangle$ is the scalar product of vectors a_i and x_{k+1}. Then from the equation of the hyperplane passing through the point x_{k+1}^* it follows that the i −th coordinate of the vector $b_{x_{k+1}}$ is equal to

$$b_{x_{k+1}}(i) = \langle a_i, x^*_{k+1} \rangle. \tag{14}$$

Next, we proceed to the calculation in the next step.

The described algorithm allows one to construct an approximation of an information set in the form of a polyhedron of any shape without performing computationally expensive operations of the Minkowski sum and intersection of sets. Since the shape of the approximating set $\bar{X}_{a_{k+1}}$ is unknown, the directions A_{x_k} should be chosen based on the requirements of the problem. The closer to the true information set the shape of the approximating polyhedron is given, the more accurate the approximation will be get and the less the estimation error will accumulate.

Example 2

Let us compare the approximations of the information set on a two-dimensional model of the angular motion of a spacecraft (SC). The state vector consists of the rotation angle φ and the angular velocity $\dot{\varphi}$.

The model of lateral motion of a spacecraft is described by a system of differential equations

$$\begin{cases} \dot{x}_1 = x_2; \\ \dot{x}_2 = u_0(1 + \Delta u), \end{cases} \qquad u_0 = \begin{cases} 0.02 \\ 0 \\ -0.02 \end{cases}, \tag{15}$$

where $x_1 = \varphi$ – spacecraft rotation angle, $u_0 -$ given control, $\Delta u -$ process noise arising from the operation of the propulsion system.

The state space model and the measurement equation are

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + \Gamma w_k, \\ y_{k+1} = x_{k+1} + v_{k+1}, \end{cases} \tag{16}$$

where $A = \begin{pmatrix} 1 & 0.01 \\ 0 & 1 \end{pmatrix}, \Gamma = (0 \quad 2 \cdot 10^{-4})', B = (0 \quad 0.01)', x_0 = (0 \quad 0)'.$

The set X_0 is set by a rectangle.

$X_0 = \{x \in R^2 | -1.162755 \cdot 10^{-4}\, rad \le x_1 \le 1.162755 \cdot 10^{-4}\, rad, -3 \cdot 10^{-4}\, rad/s \le x_2 \le 3 \cdot 10^{-4}\, rad/s\}.$

The set W is given by a segment

$W = \{w \in R^1 | -1.16 \le w \le 1.16\}.$

The set V is set by a rectangle.

$V = \{v \in R^2 | -1.30767 \cdot 10^{-4} rad, \le v_1 \le 1.30767 \cdot 10^{-4}\, rad, -3 \cdot 10^{-4} rad/s \le v_2 \le 3 \cdot 10^{-4} rad/s\}.$

To approximate the information set by polyhedral, we define the sets $\bar{X}_0, W, V$ (fig.1) in the form of systems of linear inequalities

$$\bar{X}_0 : \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} x_0 \le \begin{pmatrix} 1.16 \\ 3 \\ 1.16 \\ 3 \end{pmatrix} \cdot 10^{-4}, W : \begin{pmatrix} 1 \\ -1 \end{pmatrix} w \le \begin{pmatrix} 1.16 \\ 1.16 \end{pmatrix}, V : \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} v \le \begin{pmatrix} 1.31 \\ 3 \\ 1.31 \\ 3 \end{pmatrix} \cdot 10^{-4}. \tag{17}$$

The simulation results are presented in figures 4 and 5.

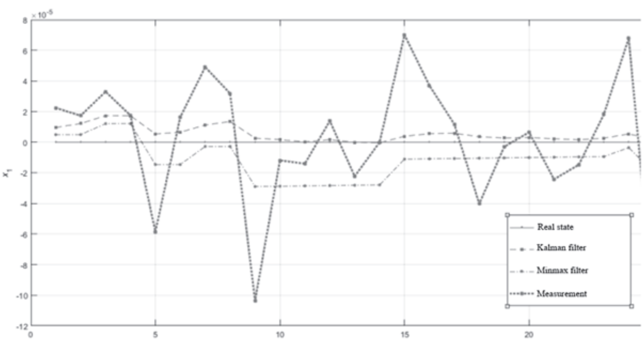


Figure 4 – Results of modeling the system of equations (16) and inequalities (17)

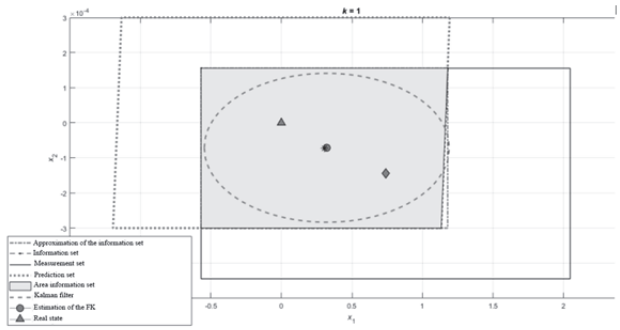


Figure 5 – The result of the approximation of the information set, the comparison with the Kalman filter

Comparison of estimates for the Kalman filter and minimax filter

The Kalman filter is used when it is assumed [21,23] that the process noise w_k and measurement noise v_k, acting on the system are random variables with known distribution functions (fig. 5). The following Kalman Filter equations are used to calculate the state vector x_k estimate $\hat{x}_k$

$$\hat{x}_{k+1} = A\hat{x}_k + K_{k+1}(y_{k+1} - G\hat{x}_k) \tag{18}$$

$$K_k = (AP_{k-1}A' + \Gamma Q_{k-1}\Gamma')((AP_{k-1}A' + \Gamma Q_{k-1}\Gamma') + HR_kH')^{-1} \tag{19}$$

$$P_{k+1} = (\mathrm{I} - K_k)(AP_{k-1}\mathrm{A}' + \Gamma Q_{k-1}\Gamma'), k = 1, \dots, N-1. \tag{20}$$

where the initial state vector $x_0 \sim N(0, P_0)$, the process noise $w_k \sim N(0, Q)$, the measurement noise $v_k \sim N(0, R)$, P_0, Q, R – are given.

Covariance matrix P_0, Q, R set in such a way that the random variables x_0, w_k, v_k at the level 3σ fall into the sets X_0, W, V.

The actual value of the state vector of the system falls into the set with a certain probability.

$$\tilde{x}_k: (X_k - \bar{X}_k)'^{p^{-1}}(X_k - \bar{X}_k) \leq l^2, \tag{21}$$

where $\tilde{x}_k -$ boundary of the confidence ellipse. The parameter l is selected depending on the dimension of the vector x and the confidence level. For example, the probability of finding a vector inside a two-dimensional ellipse with $l = 3$ is 0.989. Compare the confidence areas constructed using the Kalman filter and the information using the minimax filter.

Consider the situation when the process noise w_k and measurement noise v_k are not random variables, but change, for example, over the boundaries of the sets W and V, respectively. The Kalman filter in this case does not work correctly, because the true value of x_k does not fall into the FK evaluation region (fig. 6 upper) or the confidence ellipse region is too large in comparison with the information set. The minimax filter gives the most accurate estimate, and the estimate of the information set $\bar{X}_{a\,k+1}$¯gradually tightens up to a point (fig. 6 below), i.e. get true value.

If the process noise and measurement noise are random variables, with an unknown distribution law, then the use of the Kalman filter will not be justified [3]. In this case, it is possible that the true value of the state vector x_k goes beyond the limits of the confidence ellipse (fig. 6b).

If necessary, increase the accuracy of solving the problem of estimating the angular motion of a spacecraft, which is determined by the information set $\bar{X}_{k+1}$, and it depends (4) on the set W. The sensor, which would measure the angular acceleration [12], can be included in the sensors of the onboard spacecraft control complex. This will lead to a decrease in the set W (3) and, as a consequence, an increase in the accuracy of estimation.

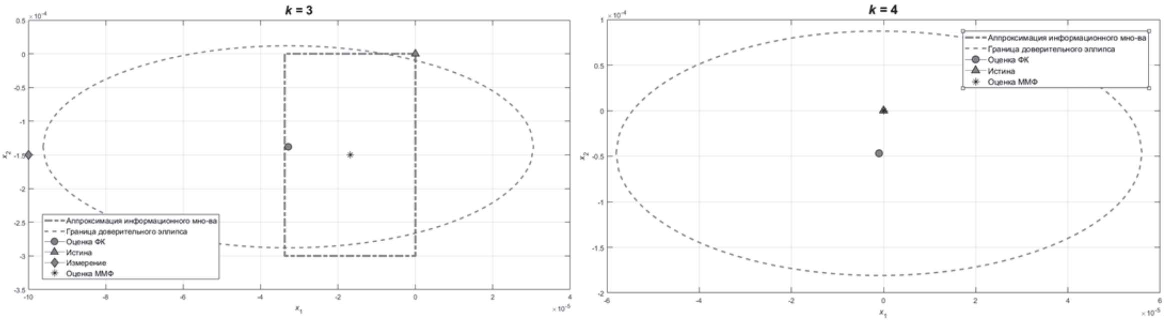


Figure 6 – Simulation result when the process noise w_k and measurement noise v_k along the set boundaries

Multistep procedure for approximation of information set

Example 2 presents an algorithm for approximation of the information set $\bar{X}_{k+1}$ based on one current measurement y_{k+1}. If this estimate is used to obtain an estimate of the information set $\bar{X}_{k+1}$ in the next step, then the inaccuracies resulting from the approximation in the previous step will accumulate with each step. Therefore, we will consider processing the measurement information for a given length L of the measurement sequence, i.e. over the entire observation period or the last few steps.

Let us consider an approximation algorithm that takes into account information about measurements y_{k+1}, process noise w_k, and measurement noise v_{k+1}, obtained not from one previous step but from several previous steps. Now it is necessary to increase the dimension of the system (10), (12) by including the variables $x_k, x_{k-1}, \dots, x_{k-L}, w_{k-1}, \dots w_{k-L}, v_k, \dots, v_{k-L}$, where L is the number of previous steps. Using this algorithm, we find a more accurate estimate $X_{a\,k}$ of the information set $\bar{X}_k$.

Set the number of previous steps L, measurement values $y_{k}, y_{k-1}, \dots, y_{k-L}$ and controls $u_{k-1}, u_{k-2}, \dots, u_{k-L}$ will be used to calculate the current estimate. Create a system of linear inequalities describing system (1) - (3) in the last L steps

$$\begin{pmatrix} I & -A & 0 & \cdots & -\Gamma & 0 & 0 & 0 & \cdots & 0 \\ 0 & I & -A & \cdots & 0 & -\Gamma & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & I & -A & \cdots & 0 & -\Gamma & \cdots & \cdots & 0 \\ G & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & G & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & H \end{pmatrix} \begin{pmatrix} x_k \\ x_{k-1} \\ \cdots \\ x_{k-L} \\ w_{k-1} \\ \cdots \\ w_{k-L} \\ v_k \\ \cdots \\ v_{k-L} \end{pmatrix} = \begin{pmatrix} Bu_{k-1} \\ \cdots \\ Bu_{k-L} \\ y_k \\ \cdots \\ y_{k-L} \end{pmatrix}. \quad (22)$$

Create a system of linear inequalities describing the sets X_{k-L}, W, V in the last L steps

$$\begin{pmatrix} 0 & \cdots & A_{x_{k-L}} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & A_w & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & A_w & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & A_v & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & A_v \end{pmatrix} \begin{pmatrix} x_k \\ x_{k-1} \\ \cdots \\ x_{k-L} \\ w_{k-1} \\ \cdots \\ w_{k-L} \\ v_k \\ \cdots \\ v_{k-L} \end{pmatrix} = \begin{pmatrix} b_{x_{k-L}} \\ b_w \\ \cdots \\ b_w \\ b_v \\ \cdots \\ b_v \end{pmatrix}. \quad (23)$$

We construct an approximation of the information set X_k on the basis of systems (22), (23)

$$X_k = \{x_k | A_{x_k} x_k \le b_{x_k}\}, \ \ \bar{X}_k \subseteq X_k \quad (24)$$

For each direction a_i, (the i-th row of the matrix A_{x_k}) we solve the linear programming problem

$$x_k^* = argmax\langle a_i, x_k \rangle \quad (25)$$

with constraints (22), (23), where $\langle a_i, x_k \rangle$ is the scalar product of vectors a_i and x_k. Then about the i-th coordinate of the vector b_{x_k} is equal to

$b_{x_{k+1}}(i) = \langle a_i, x_k^* \rangle$. (26)

Next, go to the next step $k = k + 1$.

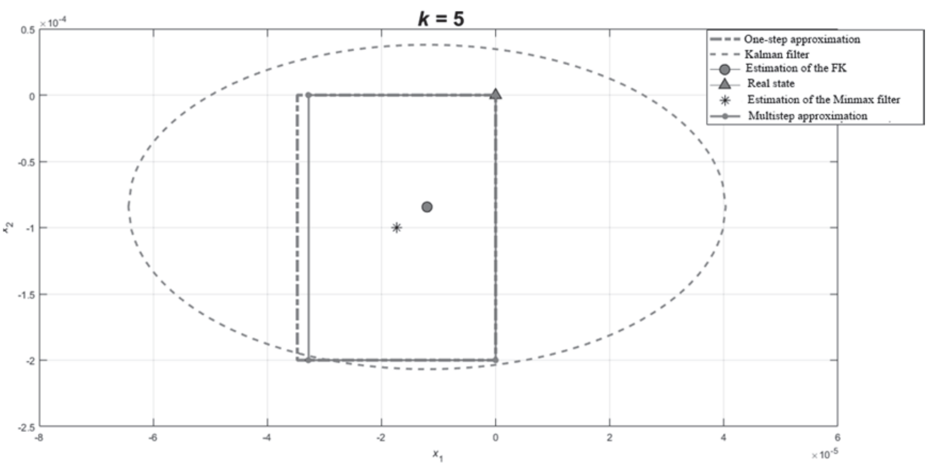


Figure 7 – Results of solving a system of linear inequalities

Let us compare the approximations of the information set for system (16), (17) of a one-step procedure with a multistep process, when the process noise w_k and measurement noise v_k vary along the boundaries of the sets W and V, respectively. According to the results of modeling (fig. 7), the information set obtained using a multi-step procedure is 0,04 times less. In comparison with the Kalman filter, the information set is 0,58 times smaller than the confidence ellipse.

The neural network (NN) algorithm for solving the system of linear inequalities

Traditional numerical methods for solving a system (7) do not allow achieving the necessary accuracy or requiring more time to be processed. Thus, the appliance of the neural network approach is of particular interest.

Let us consider the neural network (NN) algorithm for solving the system of linear inequalities $Ax \leq b,$ where $x \in R^n,\ b \in R^M -$ vectors, $A -$ matrix $n \times M$ with the use of analogy error module minimization criterion [1]. The structure of NN, realizing the converting of input signal into an output one, adequate to solving the system of linear inequalities, will be the following (fig.8) with the function of activation

$$y = f(Ax - b), \tag{27}$$

$$f(g_m) = \begin{cases} 0, & g_m \leq 0, \\ \neq 0, & g_m > 0, \end{cases} \qquad m = 1, \dots, M, \tag{28}$$

where $f \epsilon R^M$ – continuously differentiable vector function, A – weighting coefficient matrix (input data), b – displacement vector, M – the number of limitations, equal to the number of inequalities.

Error signal is defined as the difference between the received y and the desired $\bar{y}$ signals. As the required value for an output signal is $\bar{y} = 0$, then error can be found in the following way

$$e = y - \bar{y} = y. \tag{29}$$

The criterion for the quality of the system will be the achievement of a minimum of functional depending on the error. Let us consider the sum of modules of the error component vector as a F functional

$$F = |e| = \sum_{i=1}^{M} |y_i| = \sum_{i=1}^{M} \left(f\left(\sum_{j=1}^{N} a_{ij} x_j - b_i \right) \right), \tag{30}$$

where a_{ij} – A matrix elements, b_i – b matrix elements, x_j – x vector elements

The process of solving finding will be in the search of such an output signal *x* in which the functional will be minimal

$$F = \sum_{i=1}^{M} |y_i| \to min. \tag{31}$$

Adjustment will be done with the use of the iterative gradient method

$$x_{k+1} = x_k - \Delta x_k,\ \Delta x_k = H \left. \frac{\partial F}{\partial x} \right|_{x = x_k}, \tag{32}$$

where H – gain factor.

Let us find criterion gradient by differentiating (32), we will have

$$\frac{\partial F}{\partial x} = y^T A,\ \frac{\partial F^T}{\partial x} = y A^T. \tag{33}$$

Then the adjustment block equation will be the following

$$x_{k+1} = x_k - H A^T y_k. \tag{34}$$

Thus, the full algorithm equation will be

$$\begin{aligned} & x_0 = 0, \\ & y_k = f(A x_k - b), \\ & x_{k+1} = x_k - H A^T y_k. \end{aligned} \tag{35}$$

The scheme of this algorithm in the form of closed system is shown in figure 8.

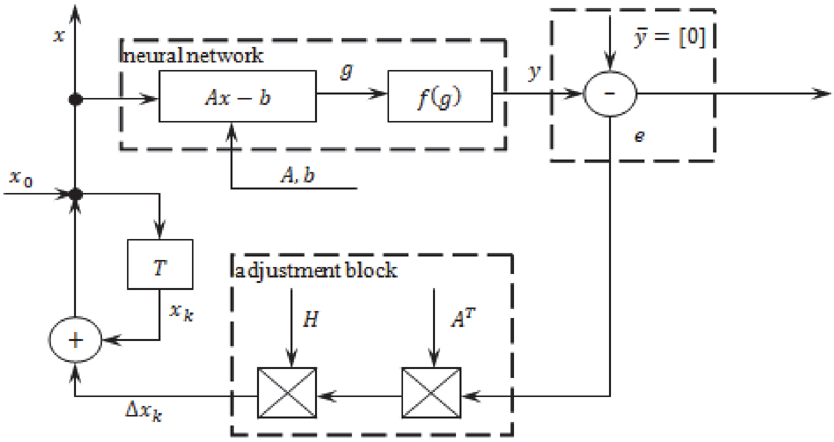


Figure 8 – The structure of the neural network algorithm for solving systems of linear inequalities in case of error minimization

Example 3

Let's compare the time costs of solving a system of linear inequalities using the neural network approach with solving a linear programming problem. Consider a system of linear inequalities

$$\begin{cases} -x_1 - x_2 \leq 1; \\ -2x_1 + x_2 \leq 4; \\ 4x_1 - x_2 \leq 2. \end{cases}$$

The results are presented in figure 9. The time spent on solving the linear programming problem is $t = 0{,}549\ s$, and the time spent on the neural network algorithm is $t = 0{,}105\ s$.

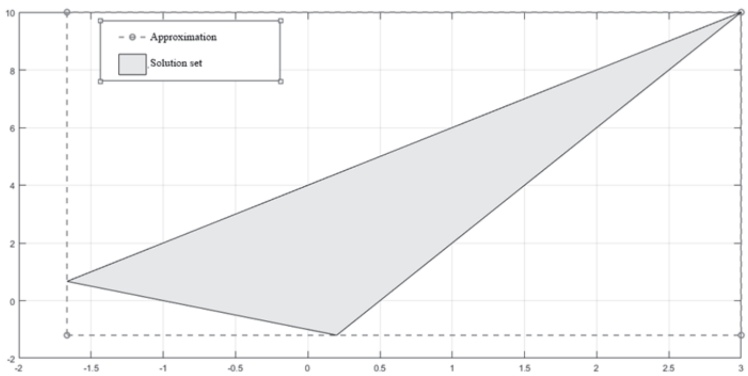


Figure 9 – Results of solving a system of linear inequalities

Conclusion

Algorithms for estimating the state of a dynamic system using linear programming methods are obtained. The results of the comparison of the guaranteed estimation with the Kalman filter are given, where the guaranteed estimation shows more accurate results in the case when there is no statistical information about changes in process noise and measurement noise. A multi-step procedure for the approximation of information sets is described, which makes it possible to increase the accuracy of estimation. One of the possible methods for accelerating the operation of the algorithm is their neural network implementation.

References

1. Ageev, A.D. Neuromathematics. Book 6 / A. D. Ageev, A. N. Balukhto, A. V. Bychkov, et al. / General ed. A.I. Galushkina. // M.: Radio Engineering. – 2002. – p. 448. (in Russian)
2. Kats, I.YA., Minimax Multistep Filtration in Statistically Undefined Situations. Avtomatica i telemekhanica, 1978, no. 11, pp. 79–87. (in Russian)
3. Kalman, R.E. Identification of Noisy Systems. Russian Mathematical Surveys, 1985, no. 40 (4), pp. 25–42.

4. Kurzhanski, A.B Ellipsoidal techniques for reachability analysis: Part I: External approximations. Part II: Internal approximations, boxvalued constraints /A.B. Kurzhanski, P. Varaiya // Optimization. Methods and Software. – 2002. – Vol. 17. – pp. 177–237. (in Russian)
5. Kuntsevich, V.M. Accuracy of construction of approximating models under bounded measurement noises / V. M. Kuntsevich // Automation and Remote Control. – 2005. – №5 – pp. 125-133.
6. Matasov, A.I. Method of Guaranteed Estimation / A.I. Matasov // Mosk. Gos. Univ.. – 2009. – 100 c. (in Russian)
7. Podivilova, E.O. On the approach of dynamic system state estimation as solving linear inequalities system / E.O. Podivilova, V.I. Shiryaev // Bulletin of the South Ural State University. Series "Computer technology, automatic control, radio electronics". – 2013. – Vol. 13, № 3(13). – pp. 133–136. (in Russian)
8. Ryabogin, N. V. Inertial Measurement Units and Star Tracker Measurements Fusion on the Basis of the Rodrigues Parameters Estimation/ N. V. Ryabogin , V. N. Sokolov, N. M. Zadorozhnaya // Mekhatronika, Avtomatizatsiya, Upravlenie. – 2017. – Vol. 18, № 5 – pp. 351-357. (in Russian)
9. Shiryaev, V.I. Algorithm of Dynamic System Control Under Condition Of Uncertainty. Mekhatronika, 2001, no. 8, pp. 2–5. (in Russian). (in Russian)
10. Shiryaev, V.I. Synthesis of control of linear systems in incomplete information / V.I. Shiryaev // Journal of Computer and Systems Sciences International. – 1994. – No. 3. – pp. 229–237. (in Russian)
11. Chernousko, F.L. State Estimation for Dynamic Systems. Boca Raton, CRC Press, 1994. 304 p. (in Russian)
12. Borisov, M.I. Angular accelerometer for bench microvibration testing of high-precision guidance and stabilization system for line-of-sight of science hardware / Borisov M.I., Vladykin S.A., Zhartovsky G.S. and ect. // The Space Engineering and Technology magazine. – 2016. – №2(13). pp. 62-69. (in Russian)
13. Galkin, D.I. Algorithm of estimation parameters of spacecraft's angular attitude with Kalman filter / D.I. Galkin // Engineering Journal: Science and Innovation. – 2013. – №2(14).
14. Ivanov, D.S. Kalman filter algorithm for satellite attitude parameters estimation / D.S. Ivanov, S.O. Karpenko, M.Yu. Ovchinnikov // Keldysh Institute preprints. – 2009. – T.9.
15. Kirilin, A.N Kosmicheskoe apparatostroenie: nauchno-tekhnicheskie issledovaniya i prakticheskie razrabotki GNPRKTs "TsSKB-Progress" (Space Engineering: Scientific and Technical Research and Practical Developments of TsKB-Progress) / A.N. Kirilin, G.P. Anshakov, R.N. Akhmetov, D.A.. Storozh // Samara: AGNI. – 2011. – 280 p.
16. Vandersteen, J. Observation and Estimation for Space Applications / J. Vandersteen // KU Lueven. – 2012.
17. Lefferts, E.J. Kalman filtering for spacecraft attitude estimation / E.J. Lefferts, F.L. Markley, M.D. Shuster, // AIAA Journal. – 1983. – Vol. 1 № 8. – pp. 135-149.
18. Lisakov, M.M.Operation of the Spektr-R Orientation System/ MM. Lisakov, S.M. Woynakov, A. S. Syrov and ect. // Cosmic Research.– 2014. – Vol.52. №5. –pp. 365-372.
19. Markley, F.L. How to Estimate Attitude from Vector Observations / F.L. Markley, D. Moratri // AAS/AIAA Astrodynamics Conference: Proceedings: Advances in the Astronautical Sciences. – 1999.
20. . Reshetnev, M. F. Control and Navigation of Satellites on Near-CircularOrbits / M. F. Reshetnev, A. A. Lebedev, V. A. Bartenev, et al. // Mashinostroenie – 1988.
21. Stepanov, O.A. Fundamentals of the Estimation Theory with Applications to the Problems of Navigation Information Processing. Part 1. Introduction to the Estimation / O.A. Stepanov // Concern CSRI Elektropribor. – 2009. – 509 p.
22. Steyn, W.H. Full Satellite State Determination from Vector Obdervations // Automatic Control in Aerospace – 1994. A Postprint Volume from the IFAC Symposium. – 1995. – pp. 195-200.
23. Kalman, R.E. A New Approach to Linear Filtering and Prediction Problems/ R.E. Kalman // J. Basic Engineering. – 1960. – pp. 35-45.
24. Aleshin, B.S. Orientation and navigation of mobile objects: modern information technologies / B.S. Alyoshin (Ed) // M.: Physmathlit. – 2006. – 424 p.

V.M. Kopylov, I.V. Shardyko, A.A. Truts, K.A. Volnyakov, D.S. Bobkov

DEVELOPMENT OF MECHATRONIC UNIT WITH MODULAR DESIGN AND INCREASED TORQUE MEASUREMENT RELIABILITY

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint Petersburg, Russia
v.kopylov@rtc.ru

Abstract

A robotic joint containing a flexible element approach is considered. The methods of increasing the torque measurement accuracy, reducing the joint weight and its manufacturing costs are presented.

Keywords: mechatronics, modular design, flexible joint, compliance, safety, reliability, additive manufacturing.

Introduction

Among many trends in modern mechatronic units design for robotics there are the development of the modular principle and the torque control systems improvement. The modular principle is well known in technology, and in robots it is applied at several levels. The upper level is represented by large subsystems such as mobile platforms or robot arms; for example, the same manipulators are used as the arms of Justin[1] and TORO[2] robots developed by DLR. Further, robotic joints can be modules themselves, and it is common practice that a manipulator often includes a limited number of joint sizes [3] or even only one size [4]. At the lower level, inside the joint, modularity is manifested in interchangeable functional elements – drives, gearboxes, electronics units, sensors, etc. In general, the idea of modularity is associated with two application strategies: interchangeability of components and extensibility of the system. In the latter case, the use of modularity opens up the possibility to create a family of joints that differ in configuration, for example, with(without) an electromagnetic brake, or with (without) a link position sensor, which leads to increased flexibility in development and reduces the average development costs.

It is quite difficult to separate the torque control problem in manipulation systems from the compliance control problem. Currently in torque control design, there are three main technologies for joint output torque measuring, i.e. indirect torque measurement, micro- and macro-measurement. Indirect, or sensorless, torque measurement [5] is based on the information about the motor current, as well as on models connecting motor current with joint torque, primarily, the friction model. Torque micro-measurement involves the measurement of micro-deformations, for example, by means of strain gauges [6-8]. On the contrary, macro-measurement is associated with integrating the elements with significant compliance into the joints and measurement of their deformation by position sensors, such as resolvers or optical position sensors [9,10]. Considering compliance, the first two cases, we are talking about the virtual elasticity provided by software, while in the latter case, the compliance is natural, mechanical. In practice, this means that in the case of external shock loads, such as impacts, the mechanical elasticity will react instantly, with a time constant determined by the physical characteristics of the element, with the deformation primarily subjected to the elastic element, while the gearbox and the drive will perceive only a long-time impact. Thus, the elastic mechanical element acts as an external load filter and contributes to the mechatronics unit reliability.

Reliability is also improved by strain gauges elimination, because they are less reliable than position sensors, while requiring a large number of signal wires. At the same time, the problem of temperature sensitivity is reduced, since the only temperature-dependent parameters are elastic (young's modulus) and thermal (linear expansion coefficient) properties of the flexible element, and the thermal effects are weak in the typical temperature range. Indirect torque measurement (by drive current) significantly loses to other methods in accuracy. Thus, the modular approach and the inclusion of a mechanically flexible element into the joint structure are of interest for the design of a robotic arm joint with increased torque measurement reliability.

Design description

The following design requirements were laid down:

– quick replacement of individual functional elements of the joint (motor, gearbox, brake, etc.);

– interchangeability of any functional element with any similar element in another version with other characteristics;

– the complete set of the joint should include: the output torque sensor, the rotor position sensor and the link position sensor;

– the shaft should be hollow, providing internal cabling;

– joint weight and dimensions should be minimal, taking into account other requirements;
– joint cost should be minimal, considering the pilot and small-scale production;
– output torque should be up to 200N•m, speed up to 14 rpm.

To meet these requirements, the joint was divided into functional modules. This solution always leads to the deterioration of weight and size characteristics due to own housing of each module and increased number of fasteners. To reduce the weight and dimension growth as a result of modular design use, a joint structural analysis was carried out with the partial material replacement from aluminum to a lighter polycarbonate plastic. Since the strength is reduced by almost 5 times (tensile strength of PC plastic is 61 MPa, bending strength – 103 MPa) [11], detailed strength analysis is required. In particular, as will be shown below, it is important to take into account the bolted joints tightening forces and contact stresses.

The use of plastic compatible with 3D printers operating on FDM technology can significantly reduce the parts manufacturing costs. This raises the technological problem of ensuring the required surface treatment quality, accuracy and dimensional stability. Thus, for each plastic joint part, a modified 3D-model was created with dimensions different from the nominal, taking into account the technological features of printing. The initial data for the models modification were obtained by analyzing the test samples geometry, as a result of which the shrinkage coefficients, the processing allowances values and allowances "in the middle of the tolerance field" were obtained.

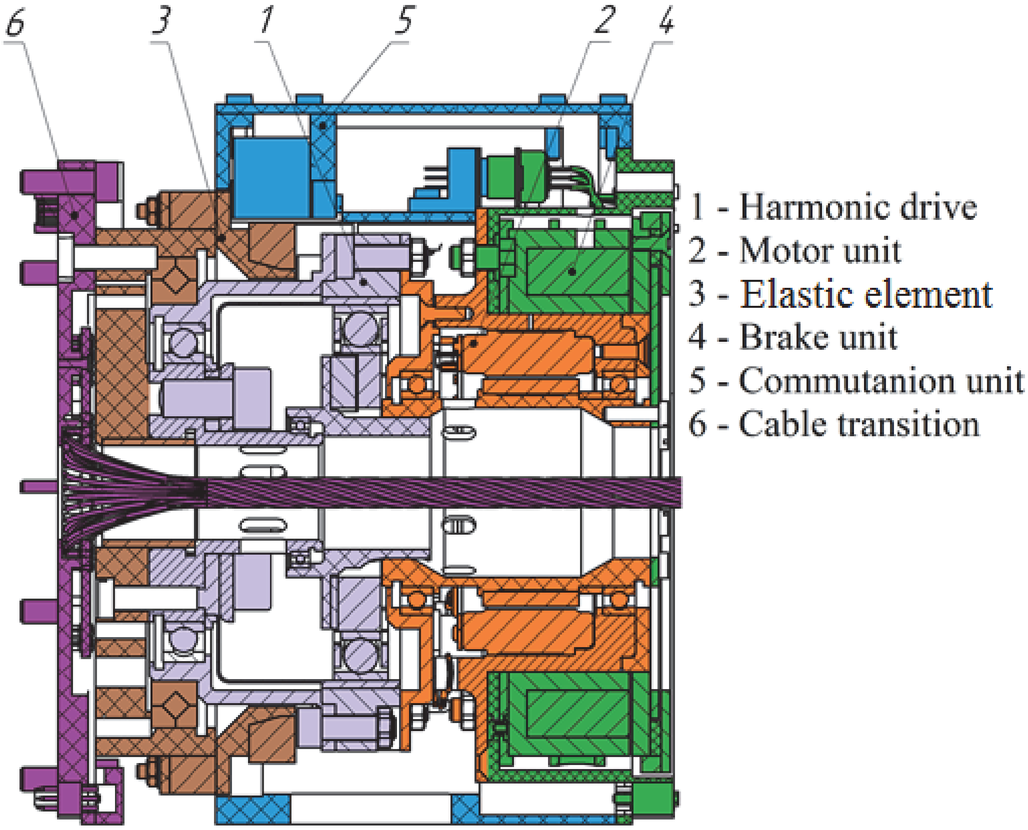


Figure 1 – Joint design in cross-section

Free volumes in the radial direction are used to preserve the joint dimensions provided in case the modular design is used. Thus, it was possible to place the brake module around the drive module, almost without increasing the axial dimension (not counting the thickness of the brake disc), and the elastic element is located around the gearbox (harmonic drive) module on 70% of its axial dimension. Module connections are made either by screws in a thread cut in the housing of another module (cable transition and the elastic element, the elastic element and the harmonic drive, the commutation unit and the brake), or with the nuts on the free ends of the bolts fixed in the unit hull (drive with brake and gearbox flange connection).

The electric circuit runs from the brake flange, which external diameter has four RPMM1 connectors. These connectors provide one of the best ratios of transmitted current to the connector area among the connectors for volume mounting. The use of PCB installation connectors, having smaller dimensions, is difficult because of the need to organize additional guides in order to protect the contacts from the loads perceived by the joint, first of all, bending moment. At the same time, the RPMM1 connectors design provides centering bushings and studs, and the contacts are made floating, which provides a simple articulation and protects against lateral loads.

In the absence of a brake in the joint configuration, a simple transition flange similar to the cable transition flange is used to attach to the manipulator's previous link. The connectors and mounting points for the brake flanges and the cable transition are fully consistent with each other.

Both the brake and the motor comprise one internal connector for docking with the commutation module. D-SUB and RPS connectors are used. The choice of the RPS connector is due to its small size and weight, which increases resistance to vibration. The choice of the D-SUB connector is due to the convenience of angular mounting on the PCB, the housing rigidity and the connector compactness. The main purpose of the commutation unit is the housing of the link angular position sensor, signal transmission from the drive rotor position sensor (RPS), as well as power transmission to the brake and to the stator. In this case, the control electronics and the power supply are located separately from the joint. This solution allows to minimize the joint dimensions, without requiring the use of microassemblies or dense installation of components on the control boards. If necessary, the commutation module hull can be made in the variant involving installation of the joint control and the power drive boards.

The hall sensor installed on the stator current drive board [12] is used as a RPS. The relatively low accuracy of the sensor is compensated by a high gear ratio. For stator cooling the air channels are provided inside the rotor shaft, as well as the grooves in the lid of the motor block, and the grooves in the housing of the switching unit. Thus, it is possible to overload the motor if the forced ventilation is provided to the hollow shaft. The air cooling line resistance is ~15 Pa, which allows the use of an axial fan with a diffuser for cooling.

The harmonic drive is installed according to the classical scheme on two supports, one of which corresponds to the fast shaft (rotor). The mechanical power supply to the gearbox is realized via Oldham clutch that ensures correct wave generator alignment [13].

Flexible flange

The use of a flexible flange as an additional (elastic) element in a joint allows to solve the following problems:

– when operating in position mode, the flange deformation allows to eliminate positioning errors;

– the macro-scale flange deformation measurement improves the torque measurement accuracy, and the measurement itself is made in a contactless way;

– the flange compliance allows the collision with an obstacle to stop the movement in a timely manner, preventing damage to the joint and/or obstacles.

In addition, in the presence of a sufficiently high structural damping of the flexible flange material, it is possible to avoid self-oscillating processes after frequent changes in the joint rotation direction. Based on the advantages that are expected to be achieved by installing a flexible element, one can choose the material for its manufacturing. Possible flexible flange materials are: spring steel type G92600, aged high-strength aluminum alloy type 2024, magnesium alloys type AZ80A, plastics. The main characteristics of these materials are given in table 1. Materials strength and elasticity are obtained from [11], hysteresis loss coefficients are obtained from [14].

Table 1. Flexible flange materials

Material	Strength, MPa	Elastic modulus, GPa	Strength/ elasticity relation	Hysteresis loss ratio
Spring steel	1375	212	$6{,}5\cdot10^{-3}$	0,01
Hardened aluminum alloy	320	72	$4{,}5\cdot10^{-3}$	0,02
Magnesium alloy	330	42	$7{,}8\cdot10^{-3}$	0,02
PLA plastic	55...57	2,3...3,3	$17...24\cdot10^{-3}$	0,1
PC plastic	61...103	2,1	$32...49\cdot10^{-3}$	0,1

Table 1 shows that the best material that provides both high torque sensitivity and high structural damping coefficient is polycarbonate plastic. However, the limited plastic strength imposes restrictions on the elastic element dimensions from below, and also requires precautions against bending loads. Thus, the elastic element is a flange, designed to receive the torque output from the outer diameter corresponding to the maximum size of the joint. For the perception of lateral loads the outer rim of the flange rests on a crossed-roller bearing [15]. Since bending of the flange flexible elements can lead to the change of the external diameter, the direct placement of the position sensor scale ring on the flexible flange is impossible. To install the position sensor ring, a slot connection is provided with centering on the slot sides and sufficient clearance on the slot inner diameter.

Increasing the flange compliance increases torque measurement accuracy, and also reduces the speed requirements to the electric motor, increasing the time constant of the flexible flange. At the same time, too high compliance of the flange leads to undesirable resonances in the low frequency band, and to increasing of joint settling time in the case of a fast load change. Thus, there is some optimal value of the flange torsional stiffness. However, since the lower limit for the flange natural frequency depends on the manipulator configuration, and the dynamic requirements are determined by specific tasks of the robotic system, it is impossible to specify the optimal value of the joint stiffness in advance. In this regard, it is assumed to manufacture a number of various flexible flanges having different stiffness values. Analysis of scientific and technical information on elastic elements [16-19] allows the following classification:

1) a continuous disk with a groove to a minimum cross-section [16];
2) disc with a circular array of holes;
3) perforated disc;
4) wheel with rim: straight spokes, tangent to the inner sleeve;
5) wheel with rim, spiral spokes, tangent to inner and outer bushings;
6) wheel with rim, spokes of variable thickness, normal to the circumferences of bushings [18];
7) wheel with spiral multi-coiled spring [19];
8) wheel with T-shaped spring spokes [17].

As a criterion evaluating the flange geometry we will use the complex value

$$K = \frac{\alpha E}{T[\sigma]n},$$

where T is the applied torque, n is the minimum safety factor, E is the modulus of elasticity, $[\sigma]$ is the tensile strength corresponding to the selected destruction model, α is the flange twist angle. This criterion makes it possible to compare flanges made of different materials and designed for different loads. Figure 2 shows the above classification with indications of the approximate value of K.

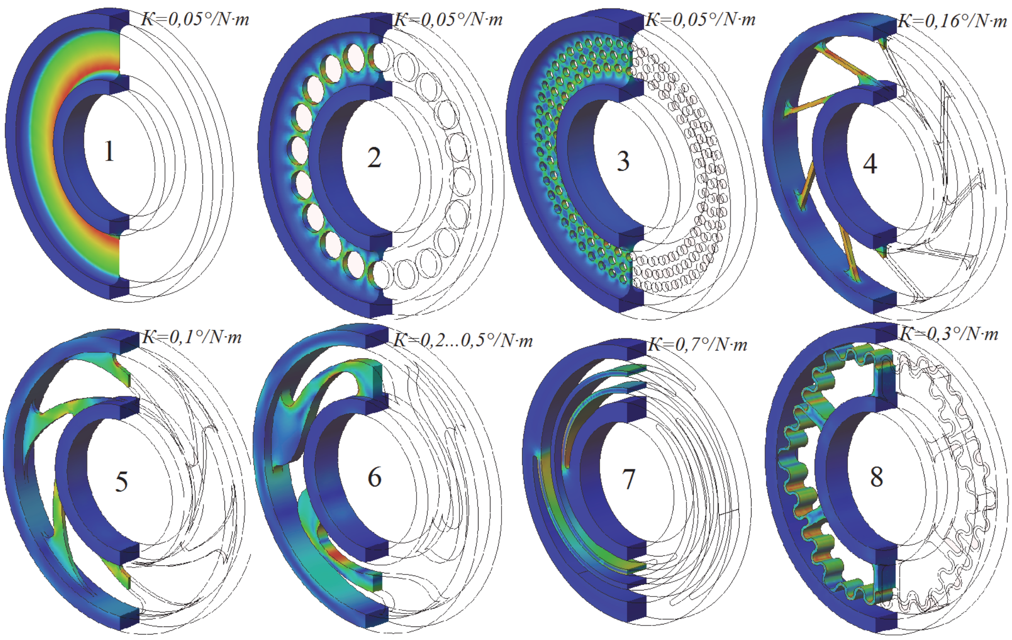


Figure 2 – Stress distribution and compliance in flanges of different type

Type 7, a spiral multi-coiled spring, is represented by a multiple spiral. At the same time, when implementing a single spiral, the K complex can take values much greater than 1, depending on the number of revolutions of the spiral. Thus, the considered classification allows to choose the desired value of the flange stiffness within a wide range (from values close to zero to the values of K equal 7 to 10 when using a multiple spring) for any selected material. However, it should be kept in mind that the single helix distributes the load unevenly, does not support the concentricity of the load diameters, and is most susceptible to changes in the outer diameter under deformation. The double spiral does not cause negative effects associated with the outer diameter non-concentricity, and is much more rigid in the radial direction, but its application requires high rigidity of the supporting bearing due to the extremely low bending stiffness of the flange. All other considered designs can have an arbitrary number of spokes, due to the desired uniformity of the load distribution, the exact value of the stiffness and K-value, and geometric constraints in the design.

Figure 3 shows the type of elastic flange developed for the joint in question. As the main scheme was taken the type 6, modified to take into account the chosen material and the requirements for bearing capacity and stiffness. Polycarbonate plastic was chosen as the material for the elastic flange. As for the joint prototype PLA plastic is used, which is more simple at printing, the strength analysis was carried out for lower values of critical stress. Due to the brittleness of the plastic, the fracture criterion is adopted by the Mohr-Coulomb model [20]. The flange compliance is 12,5 °/kN·m, the K-value for this type is 0,5 °/N·m. The size of the gaps between the flange spokes ,is selected in such a way that when exceeding the maximum torque the spokes come into contact with each other.Thus,if short-term impacts cause cracks, due to the grooves in the contact area they occur in the area above the groove, which does not bear load. Thus, the protection of the harmonic drive from high short-term loads is provided.

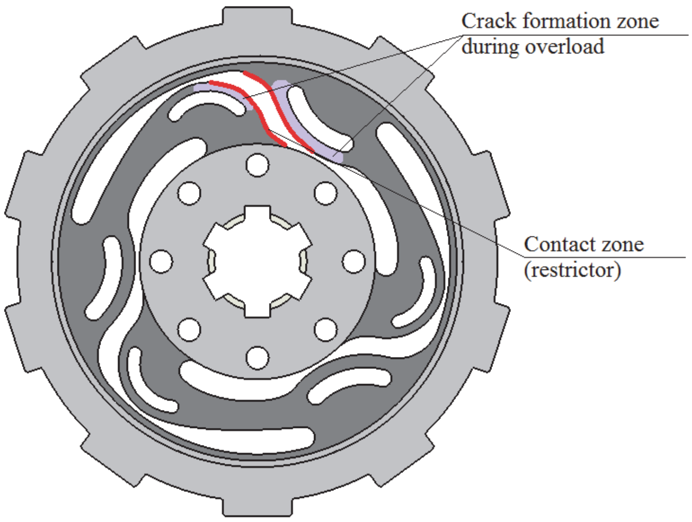


Figure 3 – Flexible flange design used in developed joint

Torque measurement accuracy estimation

Torque signal is built upon the readings of the RPS and the link position sensor. The discretization of the Hall sensor, acting as RPS, is 12 bits per revolution, which corresponds to a resolution of 316 angular seconds. A scale ring with a readhead Resolute by Renishaw[21] are used as position sensor. With a 115mm scale ring diameter, the accuracy of this system is 2,44 angular seconds. Since the position of the output link is determined by the position of the rotor with a reduction of 1:160, the effective accuracy of the Hall sensor is 1,98 angular seconds. The error introduced by the harmonic drive is up to 5 angular seconds [13]. Thus, the total measurement error (its upper limit) of the output flange twist is 8,14 angular seconds. Given the stiffness of the elastic flange, this corresponds to an accuracy of 0,09 % of the maximum measured torque. The level of torque measurement systems accuracy currently corresponds to 0,1-0,3 % of the measured torque [22].

Using the Renishaw measurement system error graph, the accuracy of the latter can be increased to ~1 angular second, so the torque measurement accuracy will reach 0,05 %. In this case, the measurement linearity is provided by the loading of the material in the proportionality area, which in fragile polycarbonate plastic practically coincides with the allowable stress area. The increase in the compliance (K-value) of the flange reduces the influence of the error of the harmonic drive backlash. Thus, at K = 1 °/N·m, the backlash effect is comparable to the errors of position sensors. However, the joint dynamic performance deterioratea as a result of the increase in compliance.

Additive manufacturing usage

To reduce the manufacturing costs and improve the joint parts manufacturability its parts are made of polycarbonate plastic by 3D printing. Piece molding was performed by the 3D-printer Picaso Designer Pro 250 [23]. The main fabrication limitations production on this printer are related to parts dimensions (no more than 210x200x200 mm), manufacturing accuracy (passport accuracy 11 microns), operating loads and local stresses. The test printing revealed the following distinctive features of manufacturing:

– shrinkage of the material is 0,5 – 1 %, and has good repeatability at similar points;

– shrinkage of the material at different points of the platform differs to 0,6 %;

– dimensional accuracy instead of the passport ±11 microns is actually ±24 microns, because thin printing slices lead to local defects due to the uneven layout of the material;

– printing accuracy of the first 4 layers is significantly lower than the following.

To eliminate these shortcomings, the part is scaled before sending to the preprocessor in accordance with a pre-defined map of shrinkage (see Figure 4), and a technology allowance of 50 microns in the particularly important places is left. If it is necessary to obtain higher accuracy, the allowance is removed by a mechanical lathe or grinding tool.

The contact stresses in the tightening zone of bolted joints are shown in Figure 5. It's worth noting that the stresses caused by operating loads are close to zero in this case. In all cases, allowing the use of washers, the contact area was increased, which made possible the use of plastic parts instead of metal. In Figure 6, details made of plastic are highlighted in color. The replacement of the material resulted in weight saving of 0,77 kg (13 %), and the cost savings for the parts manufacturing amounted to almost 60 % of the original costs (taking into account the features of pilot and small-scale production).

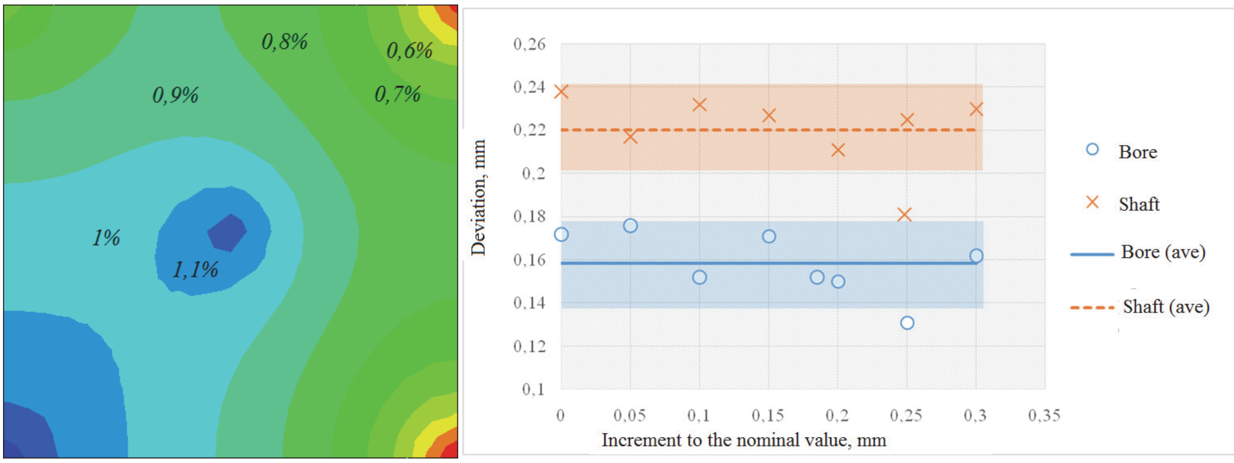


Figure 4 – Map of shrinkage ratios and tolearnce fields of printed parts

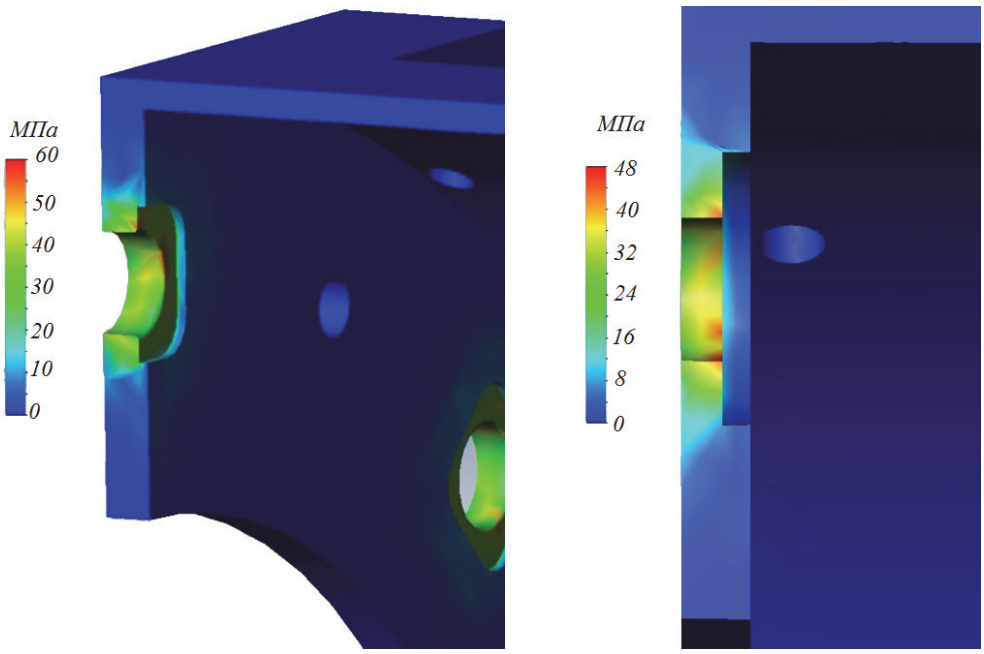


Figure 5 – Contact stress in brake hull without and with a washer

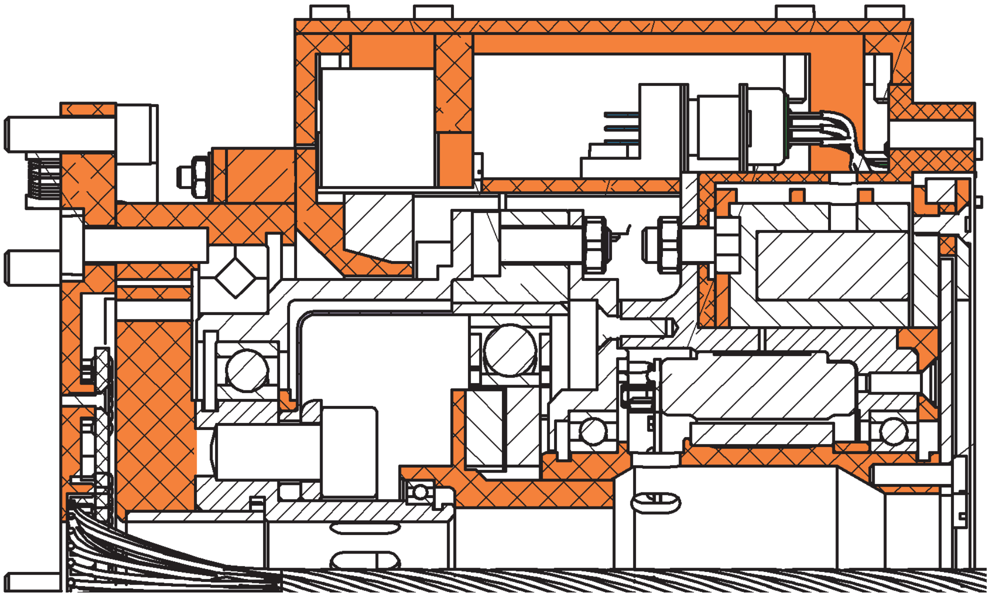


Figure 6 – Parts replaced with polycarbonate ones

Conclusion

As a result of installing an elastic element as one of the modules of the robotic joint, it is expected to that the torque measurement accuracy will significantly increase, while the contactless measurement removes restrictions on the joint rotation angle.

The proposed design solution is a compromise, combining compactness, low weight and the advantages of modularity, such as quick replacement of the main functional elements of the joint and flexible configuration. At the same time, the modular principle could not be fully implemented, since if some elements (for example, brakes) are excluded, some "passive" modules that perform secondary functions should be installed in their place. In particular, when the electromagnetic brake modules are excluded, an empty hull with electrical connectors must be installed in its place to preserve the integrity of the power supply circuit.

Manufacturing of parts with FDM technology printing has significantly reduced the joint manufacturing cost. This substitution was made possible only the calibration of the 3D printer based on the results obtained during testing printing. When analyzing the possibility of manufacturing parts using additive technology, it is necessary to conduct a strength analysis for each part, including analysis of contact stresses arising during assembly - from threaded connections tightening, rivet installation, etc.

Currently a simplified laboratory prototype of a 3-DoF manipulator is designed and being manufactured based on the developed joints. It will be used to test manipulator control algorithms that consider the joints elasticity.

References

1. C. Borst, T. Wimböck, F. Schmidt, M. Fuchs, B. Brunner, F. Zacharias, P.R. Giordano, R. Konietschke, W. Sepp, S. Fuchs, Rollin'Justin – mobile platform with variable base, in IEEE International Conference on Robotics and Automation, 2009, pp. 1597–1598
2. J. Englsberger et al., "Overview of the torque-controlled humanoid robot TORO," 2014 IEEE-RAS International Conference on Humanoid Robots, Madrid, 2014, pp. 916-923.
3. Bischoff R. et al. The KUKA-DLR Lightweight Robot arm-a new reference platform for robotics research and manufacturing //ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics). – VDE, 2010. – C. 1-8.
4. Jaekel S. et al. Design and operational elements of the robotic subsystem for the e. deorbit debris removal mission //Frontiers in Robotics and AI. – 2018. – T. 5. – C. 100.
5. De Luca A., Mattone R. Sensorless robot collision detection and hybrid force/motion control //Proceedings of the 2005 IEEE international conference on robotics and automation. – IEEE, 2005. – C. 999-1004.
6. Kanemoto Y. et al. Compact and High Performance Torque-Controlled Actuators and its Implementation to Disaster Response Robot //2018 IEEE International Conference on Robotics and Automation (ICRA). – IEEE, 2018. – C. 1-7.

7. Kim I. M., Kim H. S., Song J. B. Design of joint torque sensor and joint structure of a robot arm to minimize crosstalk and torque ripple //2012 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). – IEEE, 2012. – C. 404-407.
8. Titov V., Shardyko I., Isaenko S. Force-torque control implementation for 2 DoF manipulator //Procedia Engineering. – 2014. – T. 69. – C. 1232-1241.
9. Pratt G. A., Williamson M. M. Series elastic actuators //Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots. – IEEE, 1995. – T. 1. – C. 399-406.
10. Bodie K., Bellicoso C. D., Hutter M. ANYpulator: Design and control of a safe robotic arm //2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). – IEEE, 2016. – C. 1119-1125.
11. MatWeb: Material Property Data. [Electronic source]. URL: http://www.matweb.com
12. TQ-Drives Motor Characteristics ILM70x18 Rev100 [Electronic source]. URL: https://www.tq-group.com/filedownloads/files/products/robodrive/data-sheets/en/DRVA_DB_Servo-Kits_ILM_EN_Rev402_Web_01.pdf
13. CobaltLine-CPM/CPH/CPS Units. Engineering Data. [Electronic source]. URL: https://harmonicdrive.de/fileadmin/user_upload/ED_CobaltLine-CP_E_1019640_01.2016_V01.pdf
14. Henrik Sönnerlind. Theory and mechanisms of damping in structural mechanics [Electronic source]. URL: https://www.comsol.ru/blogs/damping-in-structural-dynamics-theory-and-sources/
15. THK. Bearings with cross rollers. Catalogue. [Electronic source]. URL: https://tech.thk.com/upload/catalog_claim/pdf/382-1R_rossRollerRing.pdf
16. Sergi F., Lee M. M., O'Malley M. K. Design of a series elastic actuator for a compliant parallel wrist rehabilitation robot //2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR). – IEEE, 2013. – C. 1-6.
17. Martins L. T. et al. Design of a modular series elastic upgrade to a robotics actuator //Robot Soccer World Cup. – Springer, Cham, 2014. – C. 701-708.
18. Liu H. et al. Design and vibration suppression control of a modular elastic joint //Sensors. – 2018. – T. 18. – №. 6. – C. 1869.
19. Lagoda C. et al. Design of an electric series elastic actuated joint for robotic gait rehabilitation training //2010 3rd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics. – IEEE, 2010. – C. 21-26.
20. Mohr-Coulomb Stress Criterion. Help system article. [Electronic source]. URL: https://help.solidworks.com/2017/English/SolidWorks/cworks/r_Mohr-Coulomb_Stress_Criterion.htm
21. RESOLUTE™ FS absolute optical encoder with Siemens DRIVE-CLiQ serial communications. [Электронный ресурс] URL: https://resources.renishaw.com/en/details/--87330
22. Torque Sensors for Static and Dynamic Applications [Electronic source] URL: https://www.te.com/usa-en/products/sensors/torque-sensors.html?tab=pgp-story
23. Designer Pro 250 printer. Technical specification. [Electronic source] URL: https://picaso-3d.com/ru/products/printers/designer-pro-250/

A. Vasiliev, I. Shardyko

ANALYSIS, DETECTION, REACTION AND PREVENTION OF POTENTIAL CRITICAL SITUATIONS FOR LIGHT-WEIGHT MOBILE ROBOTS

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg
andrey@rtc.ru, i.shardyko@rtc.ru

Abstract

Mobile robots represent an extensive trend in robotics and are widely used in the modern world. Due to the fact that the failure of the mobile robot to perform a specified task often has a high price, questions of reliability and robustness are of utmost urgency. This paper gives a brief description of modern mobile robots and analysis of the tasks they solve, also factors that could lead to the failure of the robot or the failure of the mission (risk factors) were identified. Risk factors may have an objective (environmental factors and inertial forces) or abnormal nature (failure of aggregates). In this study, only objective factors are considered. Based on the analysis of tasks and risk factors, the main types of critical situations that lightweight mobile robots can face are defined. An analytical review on the detection and countering of critical situations has been carried out, on the basis of which underexplored areas that deserve further research are established.

Keywords: mobile robots, mobile manipulators, lightweight robotic complexes, kinematic model, dynamic model, detection of critical situations, robustness, tip-over, sliding, slippage.

Acknowledgments

The results of the article were obtained in the framework of the fulfillment of the state task of the Ministry of Education and Science of Russia No. 075-00924-19-00.

Introduction

Mobile robots have received widespread employment in the modern world is such fields as inspection, intelligence, disaster response in conditions hazardous to humans, planetary research, etc. In general, mobile robots are determined by presence of two main subsystems: a mobile platform and a manipulation complex. There are a variety of mobile robots created to date that primarily differ in weight, size and design features.

Various classifications have been proposed to set the structure of research and development in the topic of mobile robots. For example, in [1] a classification of mobile robots by weight and size was suggested. According to the method of locomotion, there is a well-established division [2] of mobile platforms into wheeled [3, 4], tracked [5, 6], legged [7, 8] and hybrid [9, 10], examples of which are shown in Figure 1.

Mobile robots applications are largely associated with non-deterministic environmental conditions, which naturally increases the likelihood of mission failure due to the loss of robot locomotion ability.

In [11] mobility of a transport-technological vehicle is defined as an integral property of its exploitation which determines the ability of this vehicle to perform the specified task with optimal adaptability to the operating conditions and technical state of the vehicle itself. There are two well-established types of mobility, i.e. operational and structural mobility. Operational mobility determines the ability of a vehicle to perform a task under specified operating conditions, including both environmental features and loading modes of the vehicle. Structural mobility is determined by the ability to resist changes in the vehicle structure during its functioning (wear of components, systems and units, destruction of structures due to the influence of aggressive environmental factors, impacts, etc.). Maintaining mobility is closely linked with ensuring the dynamic adaptability of the vehicle to changing conditions.

Wheeled Tracked Legged Hybrid

Figure 1 – Mobile robots with different methods of locomotion

With regard to mobile robots, which are unmanned autonomously operating vehicles, the solution of the adaptability problem is directly related to the ability of their control systems to detect (predict) critical situations (CS) in due time and take effective measures to respond to them or prevent them if possible.

The purpose of this article is to identify risk factors that pertain to the movement of a mobile robot and a list of possible CSs that may lead to emergencies, as well as to analyze the current level of research concerning CS detection and emergencies prevention and identify promising areas of research in the field of mobile robot adaptive control.

To be specific, in the early stages of this study the authors confine themselves to lightweight mobile robots (light category and lower, i.e. weighing less than 150 kg [1]). This restriction is rather arbitrary and is caused, first of all, by the fact that the problem of providing mobility for small robots is more critical [2]. Particularly the problem considered in this paper is relevant for small-scale mini-and microclass robots, which, due to their small sizes (tens of centimeters), have to move in an environment with macro-obstacles, i.e. obstacles comparable or superior to the size of the robot [12].

Analysis of risk factors and critical situations

Study of factors that can adversely affect the operation of mobile robots should begin with a task analysis. The tasks of mobile robots can be considered at two levels: operational, i.e. in terms of the robot applications; and tactical, regarding elementary motion and manipulation tasks.

The list of operational tasks can comprise the next entries:

- CBRN (Chemical, Biological, Radiological and Nuclear) activities [3, 6];
- disaster response [5, 6, 9, 10];
- planetary research [13, 14, 15];
- service applications (logistics, agriculture etc.).

These tasks assume the execution of some scenarios built from elementary (tactical) actions, the latter comprising:

- locomotion on firm plane surfaces with normal cohesion (concrete, asphalt);
- locomotion on soft or compliant plane surface with cohesive soil (grass, loam);
- locomotion on plane surfaces with low-cohesive soil (sand, lunar regolith);
- locomotion on firm slippery surfaces (ice, thin dust on firm surface);
- locomotion on a slope with normal cohesion;
- locomotion on a slope with low cohesion;
- locomotion upstairs and downstairs;
- traversing trenches;
- traversing cliffs;
- in-tube locomotion;
- traversing/avoidance static obstacles;
- traversing/avoidance moving obstacles;
- manipulation.

Considering the control system of a mobile robot, which lies in the focus of the present study, the solution of the listed tasks is primarily the responsibility of the mobile platform. In turn, the manipulation complex may both complicate the tasks associated with the robot motion and assist the mobile platform, e.g. by translating the center of mass of the robot in a given direction or providing additional support for the mobile platform to surpass obstacles [16]. Manipulation tasks are not considered thoroughly in the context of the proposed research, since they present a vast topic themselves and they are not directly related to the problem of CSs response and prevention. At the same time, the motion as well as the current configuration of the manipulation complex should certainly be taken into account, since they can play a significant role regarding CS arising and response to it. Thereby, a number of elementary actions are taken as manipulation tasks in the current study, such as: free motion of a manipulator, interaction with moving and stationary environmental objects, and interaction with the ground.

Speaking of non-deterministic environments, real-time determination of environmental parameters (geometrical and supporting) as well as robot own actual parameters, is essential to perform the motion of a mobile robot. The solution of this problem lies in the responsibility of the sensing system of a mobile robot.

The risk factors that can lead to critical situations include both external (environmental) factors and internal factors (related to the robot itself). External risk factors are represented by three groups (see Figure 2):

1) geometric characteristics of the environment (local irregularities, various obstacles);

2) specification of wheels/tracks interaction with the ground (slippery surfaces, low-cohesive soils);
3) gravity.

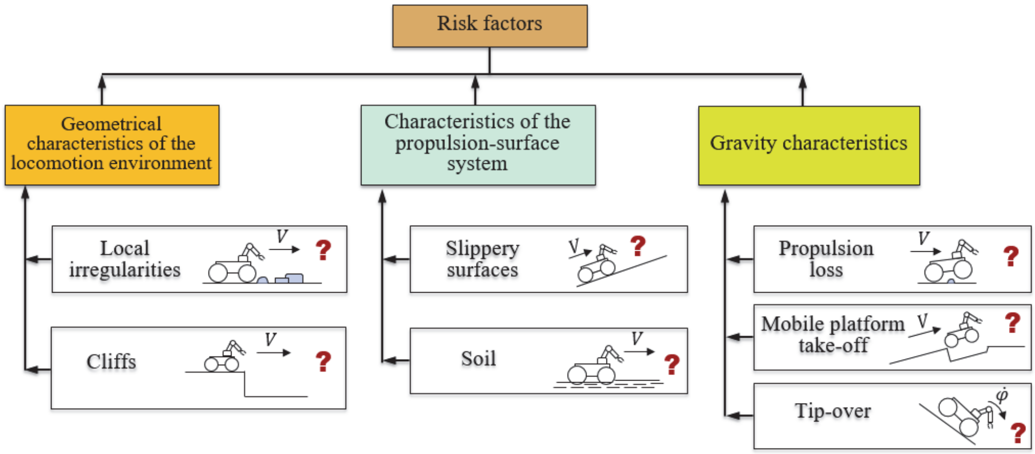


Figure 2 – Risk factors that can cause critical situations

Internal risk factors include normal, that refer to inertial forces affecting the platform and the arms, and abnormal, reflecting the possible failure of the units and hardware of the robot. Abnormal factors present a subject of a separate study, therefore not considered in this paper.

CSs are essentially processes concerning a mobile robot, leading either to the termination of the robot activity, or to deviations from the nominal execution of the mission, requiring additional actions from the operator, including nontrivial ones. An example of the latter situation is the tip-over of the robot without significant damage, after which the operator, using the robotic arm(s) and platform mobility, must find a way to return the robot to its nominal position.

The ultimate goal of this study is to build a mobile robot control system capable of detecting a CS and responding to it, as well as to prevent CSs if possible.

Potential critical situations that can be caused by the previously mentioned risk factors are presented in Table 1.

Table 1. Critical situations and risk factors causing them

Critical situation	**Risk factors**
Traction loss (longitudinal skidding)	Slippery surfaces Uneven ground (slopes) Inertial forces
Spinning	Low-cohesive soil Slippery surfaces
Deviation from trajectory (lateral skidding)	Slippery soil Low-cohesive soil Uneven ground (slopes)
Embedding/entrapment	Low-cohesive soil
Tip-over	Uneven ground (slopes) Gravity Inertial forces Contacts with environment

Critical situation	Risk factors
Fall	Cliffs Gravity Inertial forces
Propulsion loss (one or more wheels/tracks break from the ground)	Uneven ground (slopes) Obstacles
Full surface loss (take-off)	Local (surface) irregularities Inertial forces

The action of such factors as slippery surfaces (low friction at traction points), low-cohesion soils, uneven ground (mostly slopes), as well as inertial forces may cause the propulsion loss of the mobile platform, resulting in slippage of some kind.

Slippage is the process of simultaneous translation of all track/wheel (further collectively referred as traction units) points that are in contact with the surface [17] (i.e. traction points). Slippage occurs when longitudinal or lateral forces affecting a traction unit exceed cohesive forces between the traction unit and the surface. Two types of slippage are typical for vehicle motion, that are not usually separated in literature, being referred to as just slippage. Still, both the cause and the effects of these phenomena are rather different, thus it is reasonable to consider them separately.

The first one is the longitudinal slippage of traction unit, or spinning, the direction of which coincides with the direction of tangential velocity of traction points, so that linear velocity of the traction unit axis connected with the vehicle body becomes less than the velocity of that axis in case there is no slippage [17]. Spinning is caused by excess of tractive forces over cohesive forces. The danger of spinning lies in the possibility of embedding of traction units into soil, which may further lead to absolute loss of vehicle mobility. Skidding (longitudinal skidding) is the opposite phenomenon, when the longitudinal slippage of traction unit comes in the direction, opposite to the tangential velocity of traction points, so that linear velocity of the traction unit axis connected with the vehicle body becomes greater than the velocity of that axis in the no-slippage case [17]. There may also be lateral slippage which also has the skidding, or sliding, nature. Sliding here assumes skidding in the case when longitudinal velocity of the platform is zero.

Tip-over happens when non-zero moment acts on the robot about one of the axes (axes of tip-over), bounding the support polygon of the robot on the surface. This moment may be caused by gravity, inertial forces, or contacts with external objects.

Falling off a cliff is dangerous by itself as a shock load, however, landing in a nominal position results in minimal damage for the mobile robot, whereas landing on a robot side or upside-down will have a high probability of destruction or failure of onboard equipment and robotic arms.

If one or more traction units lose contact with the ground (take-off from the surface occurs) for a short time, it causes controllability loss of a mobile robot which in the best case leads to more or less substantial deviation from the specified trajectory, while in the worst case – to more dangerous situations up to an accident or a crash (falling, tip-over).

Detection, as well as response and prevention of CSs, requires tracking some metrics that represent the distance of current robot state from the CS boundary. Calculation of metrics, in turn, requires knowledge of corresponding mathematical models of robot itself and of its interaction with the environment, particularly, dynamic models.

Specific features of mobile robots modeling

It’s important to make a note that there are three types of mathematic models depending on their purpose: natural model, simulation model and control model. Natural model is represented by the system of equations comprising physical laws and geometric relations, which are continuous and consequently valid at any moment of time [18]. Simulation models are always discrete by their nature, because they are designed for implementation on computers. Such models should provide coincidence of their output variables with natural models or real-world objects with desired accuracy given the same input. Finally, control model is a system of equations, which is designed to compute control signals for some plant that is controlled. This model is discrete too, but it need not be as accurate as simulation model, while the main criterion of its quality is the performance of the plant. Besides, it should be fast enough to be implemented in real-time software, whereas simulation models are usually employed offline taking a relatively long time (days or even weeks).

Detection and prevention of CSs directly during task execution is also the objective of the control model, so this model can exclude some interactions that actually exist. Required accuracy of the control model depends on the application of the robot, so mobile robots of different kind will also account for different set of relations.

Simulation models also may be employed to detect and prevent critical situations but on the robot development stage in order to assess the performance of a mobile robot, or while developing a specific mission to estimate the probability of its success and to reveal potential vulnerabilities.

As mentioned before, mobile robots are considered as complex structures that integrate the mobile platform and the manipulation complex. Undoubtedly both components influence each other, however, their interconnection can be described by a limited parameter set, while the control models could be designed separately, making it possible to combine different platforms and different manipulators (or payloads).

Dynamic modelling of a mobile robot in general involves a number of interactions (submodels):

1. traction unit – surface («wheel-soil» model);
2. motors – traction unit («drive and transmission» model taking into account elasticities and nonlinearities);
3. traction units – mobile platform («suspension» model including active elements of the chassis geometry changing, if necessary);
4. mobile platform – robotic arms («platform - manipulator» models);
5. robotic arm («manipulator» model);
6. arm – tool/payload;
7. tool/payload – external objects;
8. mobile platform – environment;
9. mobile robot – external («robot body - external objects»).

The issue is to find out which interactions are necessary and sufficient to model for successful detection and prevention of critical situations.

Spinning analysis

Spinning of traction unit implies that its actual linear velocity is less than the theoretical value, which is based upon the actuator speed value [19, 20]. As an estimate of spinning a dimensionless factor s (slippage) is usually used, which can be expressed more conveniently via actual and rated traveled distance [21]:

$$s = 1 - S_r/S_c \,, \tag{1}$$

where S_r – actual distance, S_c – rated distance, calculated upon the data from the speed sensor of the traction unit (wheel encoder).

Knowing spinning factor makes it possible to adapt the robot motion in order to execute the desired trajectory and prevent embedding. Importance of s for modelling and control of mobile robots induced a significant number of articles in this field, including search for means of its estimation. Basically, the knowledge of actual platform velocity and actuators speed is sufficient to find s. The latter is usually measured by motor encoders, while there is a wide list of methods to estimate the platform velocity, including:

1. passive wheel in front of or behind the robot [22];
2. sensor fusion from wheel encoders and accelerometers, mounted on the mobile platform [23];
3. measurement by the camera, looking down the robot (visual odometry, VO, which, according to [24] works well up to velocities of 1 m/s);
4. 3D-odometry, i.e. sensor fusion of inertial measurement unit (IMU) with VO [25];
5. fusion of wheel encoders, IMU and GPS [26];
6. active and passive beacons (require special infrastructure) [27];
7. machine learning based on sensor fusion data, mostly proprioception-based [28];
8. visual observation of wheel tracks [29, 30].

The study on quality of spinning estimation is beyond the scope of this paper, one of the reasons for which is the fact that the choice of the method is strongly connected with the architecture of the sensor suite, and this can be done only on the design stage. Speaking of modelling, s can be set expertly, with possible addition of reasonable amount of noise. To enhance the trajectory tracking performance, it is necessary not only to know the spinning value, but also to compensate it, that can be done with two groups of methods [21]:

– dynamic, based on motor current or torque;

– kinematic, based on speed.
Also, these methods can be combined to some degree.

Tip-over analysis

Tip-over detection is based on the notion of tip-over stability. All variety of methods that estimate tip-over stability can be divided into static and dynamic ones. The difference here is that the methods of first group do not account for dynamic effects like manipulator motion or platform acceleration. With respect to direction, tip-over can also be forward, backward, or sideways, i.e. rollover [31].

Static methods include the method of normalized energy (NE) [32], geometry method [33], and system centroid position (SMP) method [31].

Dynamic methods, in turn, include:
– Zero-moment point (ZMP) method [7, 10, 34, 35];
– Force-angle (FA) method [36, 37];
– Moment-height stability (MHS) method [38];
– Energy Stability Level (ESL) method [39];
– tip-over stability on stairs [40],
– tip-over stability on slopes [41].

ZMP, FA and MHS methods exploit the notion of support polygon. It is a polygon in a horizontal plane, and if some reference point lies within it, than robot is considered as stable. The reference point in ZMP method is ZMP itself, i.e. the point, the sum of horizontal moments of reaction forces about which equals zero, while FA and MHS exploit the dynamical center of mass as reference point. Also, FA and MHS consider each side of the support polygon as potential tip-over axis (PTA), which is illustrated in Figure 3, where $\mathbf{p_c}$ is directed toward the dynamical center of mass, while PTA axes are shown as orts $\hat{a}_i$. The study [42] also features FA method, introducing such a metric as time to fall, also defined with respect to each PTA. MHS differs from FA in accounting for the robot's moment of inertia with respect to each PTA, i.e. the point model is substituted with a solid-body model.

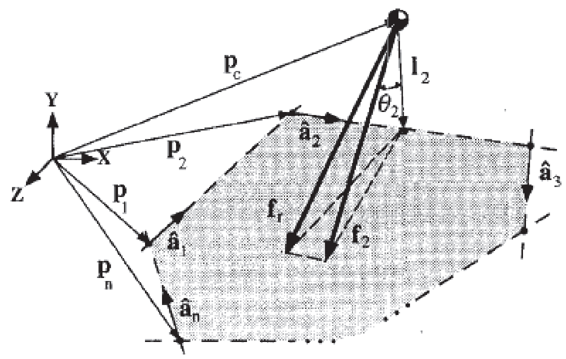


Figure 3 – FA method graphical description

ESL method also has clear physical meaning, however, requires substantial computational resources even in comparison with other dynamical methods.

Methods presented in [40] and [41] are based on an elaborate dynamic model, still the stability criteria are chosen with respect to a particular situation: motion on stairs and on slope respectively.

The main advantage of static methods is lesser computational requirements, however, with increased performance of modern controllers this advantage loses its significance. On the other hand, multi-DoF robotic arms with high dynamical capabilities that are often employed in mobile robots nowadays makes the suitability of static methods questionable. Nevertheless, in order to get reasonable verdict on suitability of any methods one would require thorough research that should consider various test scenarios for robots with different structure. An example of such research could be found in [43] and [44], however, the authors confined themselves with a limited number of situations, studying a specific robot in a specific configuration, which cannot be sufficient to present a sound comparison of methods. Moreover, it is necessary to take into account the influence of inaccurate measurement of quantities on base of which the stability metrics are constructed. These inaccuracies are caused by the errors in the sensor and control hardware and may drastically decrease the performance of methods.

The methods estimating tip-over stability has one interesting feature, i.e the tip-over stability metric can be used to counteract tip-over and also to prevent it. The most common approach here is to set a stability potential function which is employed to find the compensated motion of a robotic arm [31, 34, 41, 45] or active elements of suspension [14]. It’s worth to mention that [34, 45] consider an arm with only 2 DoF, while [41] and [31] consider 4 DoF arms, however in the latter case the arm motion is possible only in one plane. Finally, compensating motion is executed by slippers of a tracked vehicle in [32, 33, 40].

Summarizing, it should be noted that none articles on tip-over response and prevention were found that consider mobile robots with usual to date robotic arms with 5 to 7 DoF. Thus, the comparison of various proposed methods of tip-over detection, response and prevention for mobile robots with complex structures present another topic of utmost importance.

Conclusion

In this paper an analysis of tasks for mobile robots of different kind has been conducted, based on which a list of basic critical situations that can happen to a mobile robot on mission has been defined. Abnormal situations caused by failures in robot hardware and mechanics require special analysis and remained beyond the scope of the article. Publication review showed that some critical situations, i.e. spinning, embedding and tip-over, present a well-studied topic, while the other, including traction loss and deviation from trajectory, received less attention. Furthermore, none publications have been found that consider free fall or take-off.

The review showed that while spinning and embedding problems present a great interest for planet rovers designers, the tip-over problem is perceived by designers of any kind of mobile robots. There have been proposed a number of metrics and algorithms to estimate the chance of tip-over, however, a thorough analysis of them with such criteria as reliability, time to response and applicability for concrete type of robot, is missing. Also, there have been not found any research of the influence of the mobile platform velocity on its stability towards tip-over, spinning and skidding as well as studies of motion on compliant surfaces or surpassing compliant obstacles. All these topics remain to future work.

References

1. Vasiliev A.V., Lopota A.V. Utochnenie tiporazmernyh grupp nazemnyh distancionno upravljaemyh mashin dlja primenenija v opasnyh dlja cheloveka uslovijah // Nauchno-tehnicheskie vedomosti SPbGPU. – 2015. – №1 (214). – S.226-234.
2. Vasiliev A.V. Issledovanie i klassifikacija strukturno-kinematicheskih shem shassi mobil'nyh robotov // Perspektivnye sistemy i zadachi upravlenija: Materialy Devjatoj vserossijskoj nauchno-prakticheskoj konferencii. – Taganrog: Izd-vo YuFU, 2014. – S.115-128.
3. RESCUER: Development of a Modular Chemical, Biological, Radiological and Nuclear Robot for Intervention, Sampling and Situation Awareness / R. Guzman, R. Navarro, J. Ferre, M. Moreno // Journal of Field Robotics. – Wiley Periodicals, Inc., 2016. – Volume 33, Issue 7. – pp.931-945. – DOI: 10.1002/rob.201588
4. Deployment of an Autonomous Mobile Manipulator at MBZIRC / J. Carius, M. Wermelinger, B. Rajasekaran [et al.] // Journal of Field Robotics. – Wiley Periodicals, Inc., 2018. – Volume 35, Issue 8. – pp.1342-1357. – DOI: 10.1002/rob.21825
5. ResQuake: A Tele-Operative Rescue Robot / A.A. Moosavin, A. Kalantari, H. Semsarilar // Journal of Mechanical Design. – ASME, 2009. – Volume 131. – DOI: 10.1115/1.3179117
6. Integration of the Fido Explosives Detector onto the PackBot EOD UGV // Pavlo Rudakevych P., Clark S., Wallace J. Integration of the Fido Explosives Detector onto the PackBot EOD UGV // Proceedings of SPIE 6561, Unmanned Systems Technology IX, 656125 (Defense and Security Symposium, Orlando, Florida, USA, 2 May 2007). – SPIE, 2007. – Volume 6561. – DOI: 10.1117/12.720025
7. A Highly Mobile and Dynamic Quadrupedal Robot / M. Hutter, C. Gehring, D. Jud [et al.] // Proceedings of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). – IEEE, 2016. – DOI: 10.1109/IROS.2016.7758092
8. MIT Cheetah 3: Design and Control of a Robust, Dynamic Quadruped Robot / Bledt G., Powell M.J., Katz B., Di Carlo J. // Proceedings of IEEE International Conference of Intelligent Robots (IROS) (Madrid, Spain, October 2018). – IEEE, 2018. – DOI: 10.1109/IROS.2018.8593885
9. CENTAURO: A Hybrid Locomotion and High-Power Resilient Manipulation Platform / N. Kashiri, L. Baccelliere, L. Muratore, A. Laurenz [et al.] // IEEE Robotics and Automation Letters. – 2019. – Volume 4, Issue 2. – DOI: 10.1109/LRA.2019.2896758
10. Mobile Manipulation and Mobility as Manipulation - Design and Algorithms of Robosimian / P. Hebert, M. Bajracharya, J. Ma [et al.] // Journal of Field Robotics. – 2015. – Volume 32, Issue 2. – pp.255-274. – DOI: 10.1002/rob.21566

11. Koncepcija podvizhnosti nazemnyh transportno-tehnologicheskih mashin / V.V. Beljakov, A.M. Beljaev, M.E. Bushueva [et al.] // Trudy Nizhegorodskogo gosudarstvennogo tehnicheskogo universiteta im. R.E. Alekseeva. – 2013. – №3(100).– S.145-174.
12. Vasiliev A.V. Metodika sinteza strukturno-kinematicheskih shem shassi malogabaritnyh mobil'nyh robotov // Perspektivnye sistemy i zadachi upravlenija: Materialy Desjatoj vserossijskoj nauchno-prakticheskoj konferencii: v 2-h t. – Rostov-na-Donu: Izd-vo JuFU, 2015. – T.II. – S.146-157.
13. Mars Science Laboratory Mission and Science Investigation / J.P. Grotzinger, J. Crisp, A.R. Vasavada [et al.] // Space Science Reviews. – 2012. – Volume 170, Issue 1–4. – pp.5–56. – DOI 10.1007/s11214-012-9892-2
14. Control of Robotic Vehicles with Actively Articulated Suspensions in Rough Terrain / K. Iagnemma, A. Rzepniewski, S. Dubowsky, P.S. Schenker // Autonomous Robots. – 2003. – Volume 14, Issue 1. – pp.5-16. – DOI: 10.1023/A:1020962718637
15. Vasiliev A., Dalyaev, I., Slyuta E. Design Concept of Lunar Rover for the Moon Geological Exploration / B. Katalinic (Ed.) // Proceedings of the 28th DAAAM International Symposium (Vienna, Austria, 2017). – Published by DAAAM International, 2017. – pp.0780-0786. – ISBN 978-3-902734-11-2, ISSN 1726-9679, DOI: 10.2507/28th.daaam.proceedings.110
16. Shinov S.N. Ispol'zovanie manipuljatora mobil'noj robototehnicheskoj sistemy dlja preodolenija prepjatstvij // Aktual'nye problemy zashhity i bezopasnosti: Tr. desjatoj Vseross. nauch.-prakt. konf. – SPb.: NPO Special'nyh materialov, 2007. – T.5: Jekstremal'naja robototehnika. – pp.194-201.
17. GOST 17697-72. Avtomobili. Kachenie kolesa. Terminy i opredelenija ; Vved. s 01.07.1973. – M.: Izd-vo standartov, 1972. – 26 s.
18. Y. Liu, G. Liu. Modeling of tracked mobile manipulators with consideration of track terrain and vehicle manipulator interactions // Robotics and Autonomous Systems. – 2009. – Volume 57, Issue 11. – pp.1065-1074. – DOI: 10.1016/j.robot.2009.07.007
19. Vong J. Teorija nazemnyh transportnyh sredstv: per. s angl. – M.: Mashinostroenie, 1982. – 284 s.
20. Planetohody / A.L. Kemurdzhian, V.V. Gromov, I.F. Kazhukalo, M.I. Malenkov, V.K. Mishkinjuk, V.N. Petriga, I.I. Rozenceejg; pod red. A.L. Kemurdzhiana. – 2-e izd., pererab. i dop. – M.: Mashinostroenie, 1993. – 400 s.
21. Iagnemma K., Gonzalez R. Slippage Estimation and Compensation for Planetary Exploration Rovers. State of the Art and Future Challenges // Journal of Field Robotics. – 2018. – Volume 35, Issue 4. – pp.564-577. – DOI: 10.1002/rob.21761
22. Peredvizhnaja laboratorija na Lune Lunohod-1 / pod red. A.P. Vinogradova. – M.: Nauka, 1971. – T.1. – 128 s.
23. Ward C.C., Iagnemma K. Classification-Based Wheel Slip Detection and Detector Fusion for Outdoor Mobile Robots. Proceedings of 2007 IEEE International Conference on Robotics and Automation (Rome, Italy). – IEEE, 2007. – DOI: 10.1109/ROBOT.2007.363878
24. Combined visual odometry and visual compass for off-road mobile robots localization / R. Gonzalez, F. Rodriguez, J.L. Guzman [et al.] // Robotica. – 2012. – Volume 30(6), Issue 6. – pp.865-878. – DOI: 10.1017/S026357471100110X
25. Lamon P, Siegwart R. 3D-Odometry for Rough Terrain – Towards Real 3D Navigation // Proceedings of 2003 IEEE International Conference on Robotics and Automation (ICRA-2003) (Taipei, Taiwan). – IEEE, 2003. – Volume 1. – DOI: 10.1109/ROBOT.2003.1241634
26. Ward C.C., Iagnemma K. A dynamic-model-based wheel slip detector for mobile robots on outdoor terrain // IEEE Transactions on Robotics, August 2008. – pp.821-831. – DOI: 10.1109/TRO.2008.924945
27. Olson E. AprilTag: A Robust and Flexible Visual Fiducial System // Proceedings of 2011 IEEE International Conference on Robotics and Automation (ICRA) (Shanghai, China, 9-13 May 2011). – IEEE, 2011. – pp.3400-3407. – DOI: 10.1109/ICRA.2011.5979561
28. Brooks C., Iagnemma K. Self-supervised terrain classification for planetary surface exploration rovers // Journal of Field Robotics. – Wiley Periodicals, Inc., 2012. – Volume 29, Issue 3. – pp. 445-468. – DOI: 10.1002/rob.21408
29. Odometry correction using visual slip angle estimation for planetary exploration rovers / G. Reina, G. Ishigami, K. Nagatani, K. Yoshida // Advanced Robotics. – 2010. – Volume 24, Issue 3. – pp.359-385. – DOI: 10.1163/016918609X12619993300548
30. K. Skonieczny, S.J. Moreland, D.S. Wettergreen. A Grouser Spacing Equation for determining appropriate geometry of planetary rover wheels // Proceedings of 2012 IEEE/RSJ International

Conference on Intelligent Robots and Systems (IROS), Vilamoura, Portugal, 7-12 Oct. 2012. – IEEE, 2012. – pp.5065-5070. – DOI: 10.1109/IROS.2012.6386203

31. H. Zhang, A. Song. System Centroid Position-based Tipover Stability Enhancement Method for a Tracked Search and Rescue Robot // Advanced Robotics. – 2014. – Volume 28, Issue 23. – pp.1571-1585. – DOI: 10.1080/01691864.2014.976654
32. Rollover Avoidance using a Stability Margin for a Tracked Vehicle with Sub-tracks / K. Ohno, E. Takeuchi, V. Chun // Proceedings of 2009 IEEE International Workshop on Safety, Security & Rescue Robotics (SSRR 2009), Denver, CO, USA, 3-6 Nov. 2009. – IEEE, 2009. – DOI: 10.1109/SSRR.2009.5424149
33. R. Yajima, K. Nagatani. Investigation of the Tip-Over Condition and Motion Strategy for a Tracked Vehicle with Sub-Tracks Climbing over an Obstacle on a Slope // Proceedings of 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), Philadelphia, PA, USA, 6-8 Aug. 2018. – IEEE, 2018. – DOI: 10.1109/SSRR.2018.8468638
34. Q. Huang, S. Sugano, I. Kato. Stability Control for a Mobile Manipulator using a Potential Method // Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'94), Munich, Germany, 12-16 Sept. 1994. – IEEE, 1994. – pp.839-846. – DOI: 10.1109/IROS.1994.407542
35. S. Sugano, Q. Huang, I. Kato. Stability Criteria in Controlling Mobile Robotic Systems // Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'93), Yokohama, Japan, 26-30 July 1993. – IEEE, 1993. – pp.832-838. – DOI: 10.1109/IROS.1993.583186
36. E.G. Papadopoulos, D.A. Rey. A New Measure of Tipover Stability Margin for Mobile Manipulators // Proceedings of IEEE International Conference on Robotics and Automation (ICRA'96), Minneapolis, USA, 22-28 April 1996. – IEEE, 1996. – pp.3111-3116. – DOI: 10.1109/ROBOT.1996.509185
37. E. Papadopoulos, D.A. Rey. The Force-Angle Measure of Tipover Stability Margin for Mobile Manipulators // Vehicle System Dynamics. – 2000. – Volume 33, Issue 1. – pp.29-48. – DOI: 10.1076/0042-3114(200001)33:1;1-5;FT029
38. S. Ali, A. Moosavian, K. Alipour. Stability Evaluation of Mobile Robotic Systems using Moment-Height Measure // Proceedings of 2006 IEEE Conference on Robotics, Automation and Mechatronics, Bangkok, Thailand, 1-3 June 2006. – IEEE, 2006. – DOI: 10.1109/RAMECH.2006.252730
39. A. Ghasempoor, N. Sepehri. A Measure of Machine Stability for Moving Base Manipulators // Proceedings of 1995 IEEE International Conference on Robotics and Automation, Nagoya, Japan, 21-27 May 1995. – IEEE, 1995. – DOI: 10.1109/ROBOT.1995.525596
40. Y. Liu, G. Liu. Track-Stair Interaction Analysis and Online Tipover Prediction for a Self-Reconfigurable Tracked Mobile Robot Climbing Stairs // IEEE/ASME Transactions on Mechatronics. – 2009. – Volume 14, Issue 5. – pp. 528-538. – DOI: 10.1109/TMECH.2009.2005635
41. Y. Liu, G. Liu. Interaction Analysis and Online Tip-Over Avoidance for a Reconfigurable Tracked Mobile Modular Manipulator Negotiating Slopes // IEEE/ASME Transactions on Mechatronics. – 2010. – Volume 15, Issue 4. – pp. 623-635. – DOI: 10.1109/TMECH.2009.2031174
42. D. Rey, E. Papadopoulos. On-line Automatic Tipover Prevention for Mobile Manipulators // Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '97). – IEEE, 1997. – pp.1273-1278. – DOI: 10.1109/IROS.1997.656414
43. 2010_Real-World Validation of Three Tipover Algorithms for Mobile Robots / Philip R. Roan, Aaron Burmeister, A. Rahimi [et al.] // Proceedings of 2010 IEEE International Conference on Robotics and Automation (ICRA'10), Anchorage, USA, 3-7 May 2010. pp.4431-4436. – IEEE, 2010. – DOI: 10.1109/ROBOT.2010.5509506
44. Y.H. Liu, X.C. Meng, M.L. Zhang. Research on Mobile Manipulator Tip-over Stability and Compensation // Proceedings of the 8th WSEAS International Conference on Robotics, Control and Manufacturing Technology (Hangzhou, China, April 6-8, 2008). – WSEAS Press, 2008. – pp.114-120.
45. Real-time ZMP Compensation Method using Null Motion for Mobile Manipulators / J. Kim, W.K. Chung, Y. Youm, B.H. Lee // Proceedings 2002 IEEE International Conference on Robotics and Automation (Washington, USA, 11-15 May 2002). – IEEE, 2002. – DOI: 10.1109/ROBOT.2002.1014829

Yin Shuai, A.S. Yuschenko

COLLABORATIVE ROBOT - SURGEON ASSISTANT

Bauman Moscow State Technical University, Moscow, Russia
shuai.yin@yandex.ru, arkadyus@mail.ru

Abstract

In this paper, the collaborative robot is seen as a surgeon's assistant who accompanies the operation, providing the necessary tools and performing other auxiliary actions. Such a robot should be mobile, have a manipulator, means of visual communication, a system of autonomous navigation in the operating room, and an interactive system of interaction with the surgeon-operator. The paper deals with the organization of the speech dialogue of the surgeon with the robot, issues of image processing and control of the collaborative robot. The main means of solving these problems are convolutional neural networks. In the ROS environment, a simulation of the behavior of a collaborative robot was performed. The main results were tested when operating the KUKA LBR iiwa 14 R82 manipulator.

Keywords: dialogue system, image processing, convolutional neural network, speech recognition.

1. Introduction

Medical science is one of the most promising applications of cooperative robotics. At present, one of the most successful projects in this area is the Da Vinci Robot System, which has been widely used in small-invasive surgery in major medical centers around the world, including Russia. Note that similar surgical robots are being developed in China and in Russia, and their performance is the same as that of the system [1]. The Da Vinci system can be attributed to the remote control robotic system rather than to collaborative one, since the operator directly, although remotely, controls the movements of the robot. Along with the remote control robotic systems collaborative robots may be used as surgeon assistants performing the operation in the traditional way.

Robot actions consist of three phases. Firstly, by voice request of the surgeon, robot determines which tool is required by the surgeon. At second phase, the robot finds the right tool on the desktop and grabs it. Finally, at third phase, it gives the instrument to the surgeon.

To solve the first problem, a system of object-oriented dialog control of a robot has been proposed. It based on the theory of finite state machines using a deep neural network. The second problem has been solved using a double convolutional neural network. Robot control procedure by dialogue has already been considered in a number of papers. In [2] fuzzy logic was used to solve it, and in [3] theory of finite automata and Petri nets were used. Speech dialogue allows to adjust both the results of speech recognition of the surgeon and the corresponding actions of the robot. Feature of the proposed recognition system, which uses a double convolutional network, is to combine dialogue and control the actions of robot. Initially, the network is trained using the vision system to recognize the main surgical tools. Grip position and tool orientation conditions necessary for planning the movement of manipulator are also determined. A trained neural network should ensure the pace of performing all the necessary actions close to the rate of performing the same actions by a human assistant, which is important for surgical operation.

To solve the problem of transferring the tool to the surgeon, it is necessary not only to have on board the robot a navigation system that determines the robot's own position relative to the operating table, but also a spatial vision system that recognizes the position and orientation in the space of the surgeon's hand. The last problem may be solved by various technical devices and have already been considered in relation to the control of robots using gestures [4], as well as in the task of exchanging information using the finger "alphabet of the deaf" [5].

Note that the robot - assistant of surgeon is only one of the tasks of robotic surgery. Such a robot (or several robots for various purposes) can solve other auxiliary tasks in the operating room, including monitoring the patient's condition and the operation of other devices used during the operation. The system of interaction between humans and robots could facilitate the work of the surgeon, relieving him of auxiliary functions and increasing the accuracy and reliability of their implementation (Fig. 1).

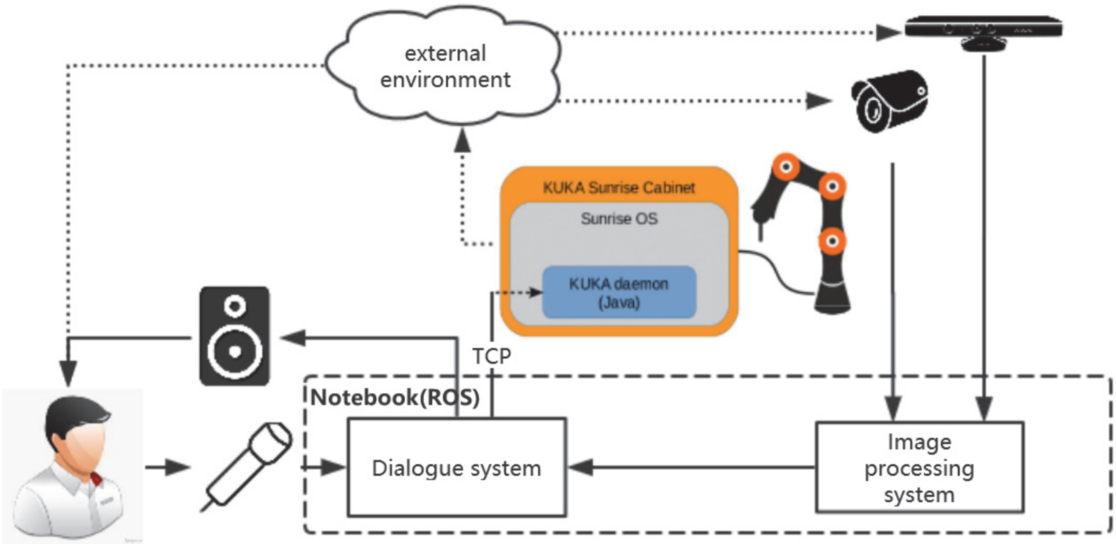


Figure 1 – The system of interaction between humans and robots

In Figure 1 operator sends a voice request to the dialogue system. The dialogue system translates the voice signal into text information, analyzes it and converts it into robot actions. The system is able to clarify the task in the process of dialogue with the operator. The robot control system performs the necessary actions on a voice request using the image processing subsystem. At the same time, the image processing system returns the execution result to an interactive system that is capable of implementing voice feedback for the operator.

2. Organization of speech dialogue using neural networks

The dialogue system is a means of exchanging information between humans and robots using natural language. Such a system allows robots to be controlled by operators who do not have special training, which is important for medical robotics. The dialogue system includes several subsystems, including an automatic speech recognition unit, a speech understand unit, a dialogue control unit, response generation units and voice synthesis units. The speech recognition unit converts a voice message into text. The task of the speech understand unit is correctly interpret this message regardless of the form of its expression. The dialogue control unit receives information from the robot control system of the machine and controls the dialogue process in accordance with the dialogue scenario. The text of the synthesized answers is converted into a voice response in the “speech synthesis” block.

The dialogue involves the recognition not only of separate words, but of individual phrases and sentences. Therefore, in this case, for speech recognition, it is advisable to use a Hidden Markov Model (HMM). In HMM the observable probabilities $b_i(k) = p(o_t = v_t \mid s_t = i)$, are first calculated using the Forward-Backward Algorithm (FBA) [6]. Next, the model is trained using the Baum-Welch algorithm (BWA) [7], as a result of which the parameters {A, π, B} are determined, with known observable vectors v_k. The calculation of the most probable state, with known parameters {A, π, B}, and observable vectors v_k, is carried out using the Viterbi algorithm [8].

To solve the problems of voice control in robotics, it is advisable to use Deep Neural Networks (DNN), which have much greater capabilities than ordinary neural networks, from which they differ in the number of layers [9]. In DNN model there are at least 5 hidden layers. However, neural networks are not able to model voice signals directly. Therefore, in order to use the ability of neural networks to classify, it is advisable to consider the combined model DNN-HMM (Figure 2).

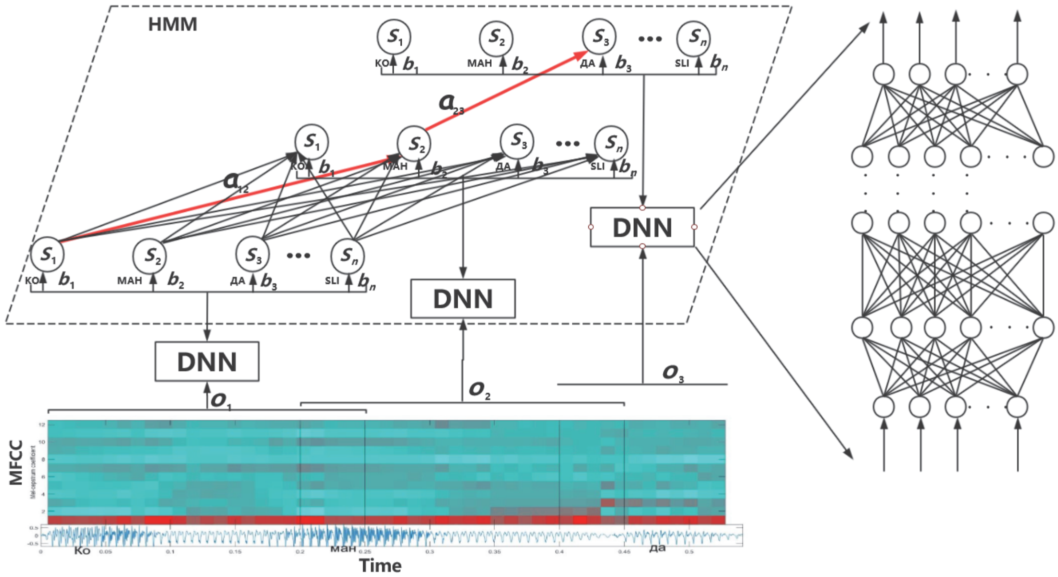


Figure 2 – Speech recognition scheme using deep neural network

The DNN-HMM model, shown in Figure 2, consists of deep neural networks DNN and a hidden Markov model HMM; it used to describe the dynamics of voice signals. The neural networks calculated probabilities of the observed vectors. Probability of the initial state assumed known beforehand. For given characteristics of voice signals, a priori probabilities of state are to be obtained at the outputs of neural networks. Figure 2 indicates : a_{12}, a_{23} - the probability of transition between states; $b_1, b_2, b_3, ..., b_n$ - observable probabilities; shown in Figure 2 DNN network has 6 hidden layers of 2048 neurons, 440 input and 1024 output signals of the network.

Number of output signals of DNN network is determined by the Bayes formula:

$$P(o \mid s) = \frac{P(s \mid o) P(o)}{P(s)} \tag{1}$$

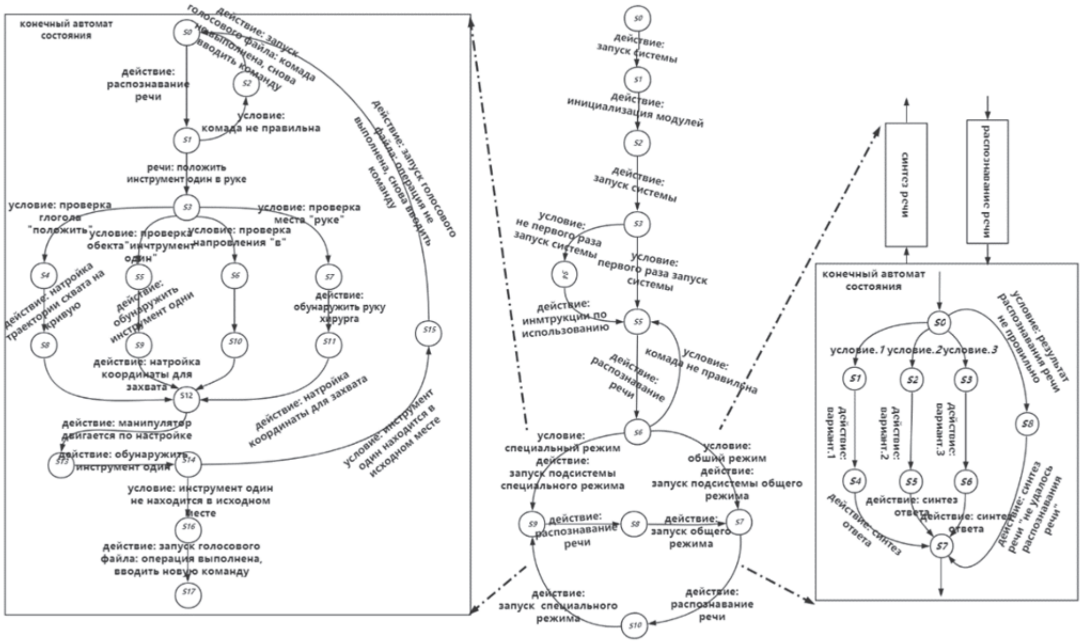
where $P(s)$ - the a priory probability of the states; $P(o)$ - does not depend on the sequence of words and can be taken equal to 1.

The neural network is trained using the Backward Error Propagation Algorithm (BPA). Unlike the DTW (Dynamic Time Warping) model, or the GMM-HMM combined method, here GMM (Gaussian Mixture Models [9] [10], the DNN-HMM model is capable of recognizing natural speech, which is especially important when developing a collaborative-type robotic system, oriented to work with a user who does not have special training.

3. Dialog control with state finite automata

Dialog management directly influences how the natural and intellectual dialogue of the system is perceived by users. The task of managing the dialogue is to determine the next action of the system, taking into account the previous course of the dialogue. The conversational process consists of referrals between the subject making the request and the subject responding to the request, which depend on the context of the discourse. In this case, one of the subjects of the dialogue is a man, and the other is a robot. The dialogue process can be viewed as the exchange of information, in which the initiative can belong to both the user and the technical system.

Below we consider the possibility of using an object-oriented dialog system for controlling a manipulation robot in a special mode and in a general mode (Figure 3).



a) b) c)

Figure 3 – Diagram of dialogue control system:
a) special mode;
b) general scheme of dialogue management;
c) graph of general mode

For example, in the special mode, the statement "to put the tool one (number 1) in hand" is converted into a sequence of commands: "move the manipulator from position B (obtained by executing the previous command "take the tool one") to position A. Note that in more complex cases, it is necessary to solve the problem of automatic action planning. For example, if in the above command on the tool one is item C, which must first be removed [12]. Using the technical vision of the robot, one can determine the coordinates of the tool at position B and the coordinates of the position A specified by the operator, and then transfer these coordinates to the manipulator control system. For complicated objects recognition it is possible to use a convolutional neural network [13]. After the predetermined operation was completed, supposed a real state check whether the block really is in position A. If the answer is positive, then the task is considered completed.

The general mode extends capabilities of operator, allowing him to obtain additional information in real time. The easiest way to implement a common mode is to connect the interactive system to the Internet, which contains the necessary Yandex service for speech synthesis.

4. Recognition of objects using convolutional neural network

The complete process of object recognition in our robotic system includes the detection of recognition areas, the recognition of objects, the determination of the orientation of objects, and the transformation of object coordinates. The task of detecting areas of recognition is realized as a search for areas where objects may exist. For example, methods for separating an object from the environment according to the properties of objects [13], [14], [15], 16], etc. The object recognition task as the classification of object recognition areas using the key points of SIFT [17], Hog [18] and the dual classifier SVM [19].

Determining of the objects orientation involves the analysis of the spatial orientation of recognizable objects using 3D models [20] and geometric approaches [22] in the case when the shape of the object is known in advance. It is necessary also to transform coordinates of the object relative to the camera to the coordinates of the object relative to the manipulator. If the coordinates of camera do not change, the homogeneous transformation matrix is also keep constant. If the coordinates of the camera relative to the manipulator are changed, then each time, in order to perform the transformation of the coordinates of the object, it is necesary to recalculate the uniform transformation matrix.

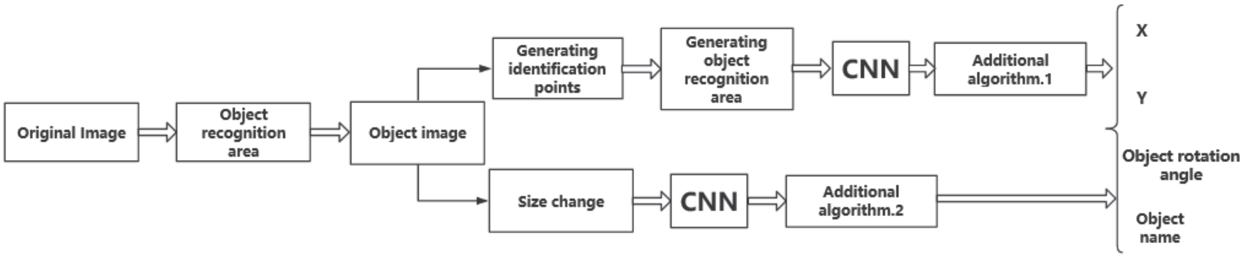


Figure 4 – Functional diagram of convolutional neural network

Fig. 4 shows the scheme of a convolutional neural network. The upper convolutional neural network should classify the recognition areas of each medical tool. The network also divides the recognition areas into three groups: the head of the tool, the tail of the tool and the rest of the tool. This network include an input matrix, three convolutional layers, two fully connected layers and forms three output signals. Size of input matrix is 56 * 56 * 3. The inferior convolutional neural network (CNN) classifies the recognition regions and determines the names of medical tools. Its size is the same as the upper convolutional neural network. The size of the input matrix is 224 * 224 * 3; the network has six outputs. Additional algorithm No. 1 calculates the position and orientation of objects on the desktop plane, as well as the characteristic features of objects. Using the obtained average value, the angle of rotation of medical tools is calculated by formula:

$$\delta_{угол\ поворота} = \tan^{-1}\left(\frac{\overline{y_1} - \overline{y_2}}{\overline{x_1} - \overline{x_2}}\right) \tag{2}$$

where $(\overline{x_1}, \overline{y_1})$ - the average value of the head parts; $(\overline{x_2}, \overline{y_2})$ - average value of tail parts.

The additional algorithm No. 2 determines the name of medical tools according to the formula:

$$a = \mathrm{MAX}(p_i)\, i = 1 \ldots 6 \tag{3}$$

where a is a number of the name of medical tools; p_i - probability of each output of the lower convolutional neural network.

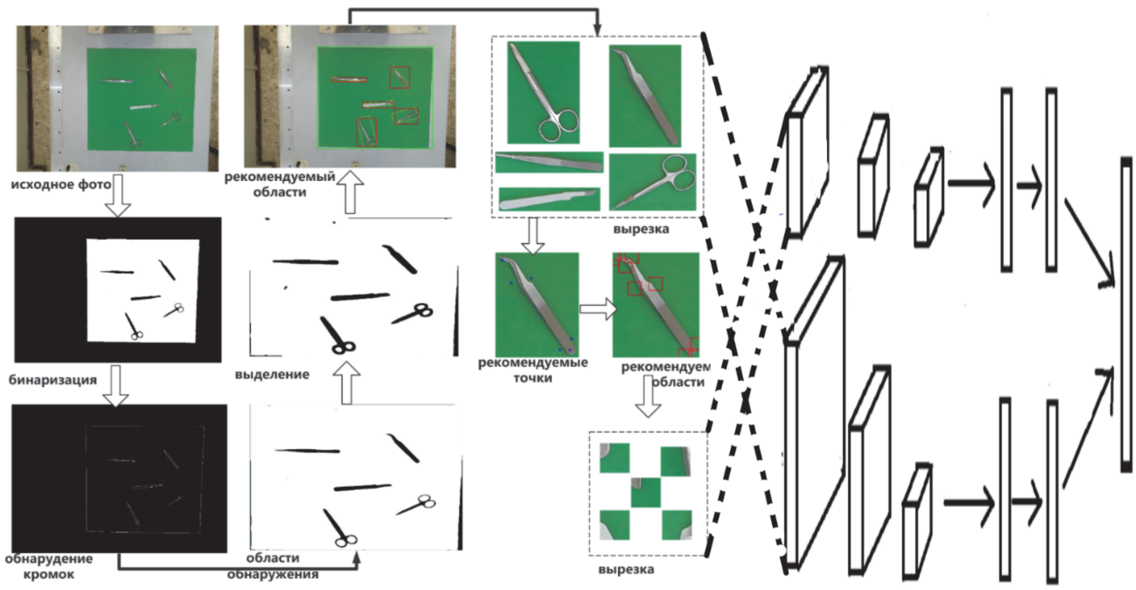


Figure 5 – Object recognition process

Figure 5 shows the image preprocessing procedure. Using binarization procedure, one can distinguish green and other colors from the results of edge detection and determine the boundaries of the identification area shown in green. After this, identification areas and objects are distinguished. Note the identification areas and consider them as input for the lower convolutional neural network at Fig.4. To determine the orientation of tools, it is necessary to create recognition areas (size 56 * 56) for each tool separately. Recognition points are generated based on changes in the curvature of the edges of each tool. The recognition area is the input to the upper convolutional neural network (Fig.4).

In experiments was used the usual RGB color camera (photo size 3968 * 2976), mounted on the capture arm. According to the known coordinates of the head and tail parts of the tools, it is possible to calculate approximately the center of gravity of objects after formula:

$$\begin{cases} x_{\text{з}} = k \cdot x_{\text{г}} + (1-k) \cdot x_{\text{х}} \\ y_{\text{з}} = k \cdot y_{\text{г}} + (1-k) \cdot y_{\text{х}} \end{cases} \tag{4}$$

where $(x_{\text{з}}, y_{\text{з}})$ - the capture coordinate; $(x_{\text{г}}, y_{\text{г}})$ - coordinate of the head; $(x_{\text{х}}, y_{\text{х}})$ - tail coordinate; k - coefficient of the center of gravity of objects.

After obtaining the capture coordinates from the images, it is necessary to perform a uniform transformation of the coordinates of the objects relative to the camera to their coordinates relative to the manipulator:

$$[x_0 \quad y_0 \quad z_0]^T = T \cdot [x_1 \quad y_1 \quad z_1]^T \qquad T = \begin{bmatrix} R & P \\ 000 & 1 \end{bmatrix} \tag{5}$$

where R - 3x3 rotation matrix; P - 3x1 displacement matrix; $\begin{bmatrix} x_0 & y_0 & z_0 \end{bmatrix}^{\mathrm{T}}$ - coordinates of objects relative to the camera; $\begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix}^{\mathrm{T}}$ - coordinates of objects relative to the manipulator;

5. Determination of the spatial coordinates of the operator's hand using the KINECT device

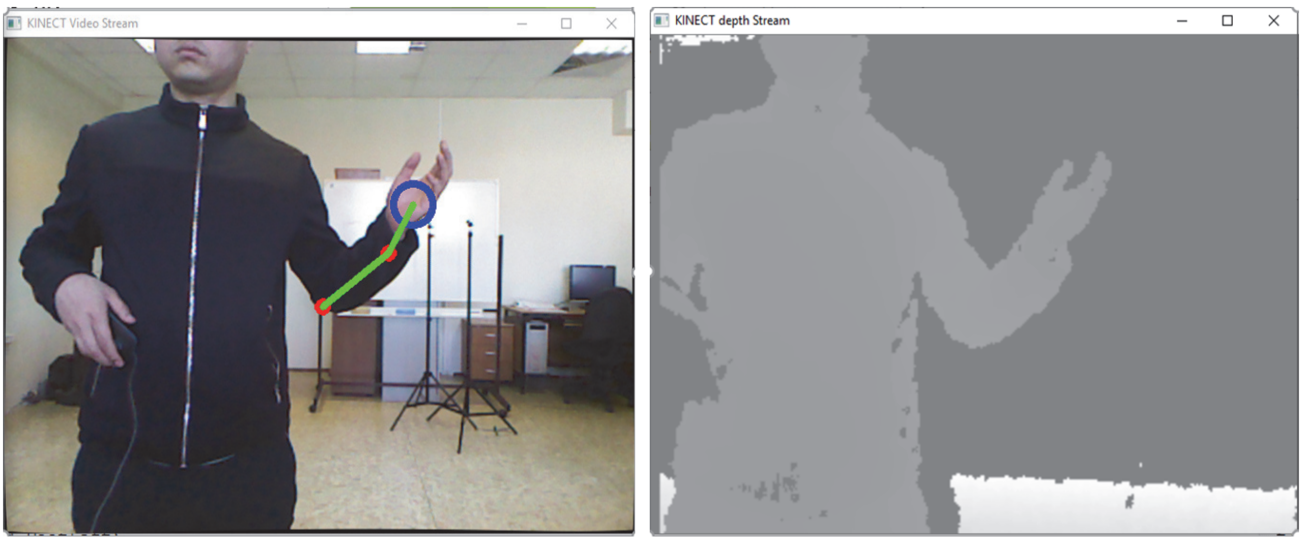
The last task for the collaborative robotic system is to give the prescribed medical instrument to the surgeon. Therefore, it is necessary to determine the position of the surgeon's hand in previous restricted region. We intend to use for this task a pinhole camera model with internal parameters f_x, f_y, c_x, c_y corresponding to the focal length and the optical center. This allows you to calculate the position of the point (i, j) in the resulting image at a known position of the hand $(x, y, z)^{\mathrm{T}}$ according to the formula:

$$\pi(x, y, z) = (\frac{f_x x}{z} + c_x, \frac{f_y y}{z} + c_y) \tag{6}$$

Now we can restore the three-dimensional point corresponding to the pixel $(i, j)^T \in \mathbb{R}^2$ with depth using the formula:

$$p(i, j, z) = (\frac{(i - c_x)z}{f_x}, \frac{(j - c_y)z}{f_y}, z)^{\mathrm{T}} \tag{7}$$

To calculate the internal parameters f_x, f_y, c_x, c_y, the camera calibration tests is used.



a) b)

Figure 6 – Output information of KINECT: a) RGB color image, b) the image of depth D

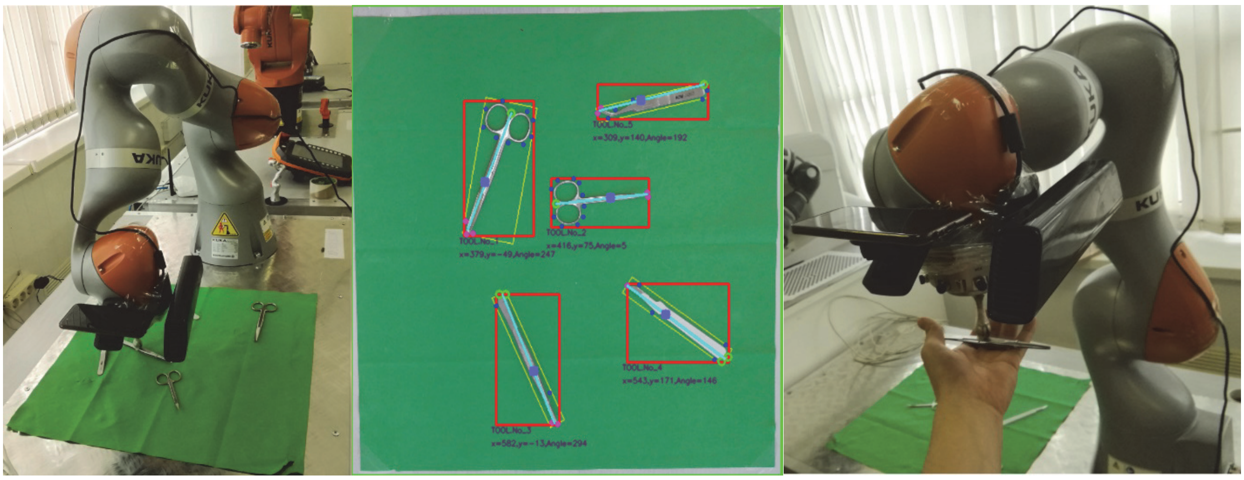
The output of the KINECT device includes a color image (RGB) and image depth (D) (Figure 6). The task of the KINECT camera is included in that getting the depth of the hand by the depth of the image (D). Pixels $(\boldsymbol{i},\boldsymbol{j})^T \in \mathbb{R}^2$ of the hand image were calculated using a convolutional neural network (see pt. 4). Through (7), we can calculate the spatial coordinates of the surgeon’s hand.

Note that partially we also solve the problem of security i.e. the absence of a robot, or a tool collision with the operator's hand. However, security task is not limited to this, since other objects may appear in the robot's workspace, the operator himself may fall into this area, etc. The simplest way to solve this issue is to require the manipulator to stop moving at the risk of a collision and to require the operator to clear the work area. After remove the obstacle, the manipulator continues to move in accordance with the original instructions. However, this requires continuous additional analysis of the working area and the solution of the problem of predicting further movements of objects in the working area. These issues are to be considered in our further work.

6. Experimental studies.

For experiments the KUKA LBR iiwa 14 robot was used (Figure. 7 a)). The control system of the collaborative robot as a whole was developed in the PYTHON programming environment in the ROS system.

To solve the problem of speech recognition, the DNN-HMM acoustic model was applied using the KALDI software. Training data was taken from the open site VOXFORGE. When controlling the dialogue, a state machine was created that can independently detect errors, including speech recognition errors, errors of control commands, robot actions, etc. For recognition of specific objects, training data were mainly created manually: 200 samples for each object and 1200 samples for objects shown in Figure 7 b). These are usual surgical tools. Object recognition algorithms were implemented using the TENSORFLOW and OPENCV software.



a) b) c)

Figure 7 – Experimental study of the dialogue system

Figure 7 b) the results of recognition of surgical tools placed randomly on the desktop are shown. The desktop was green for contrast with metallic instruments. The rectangles (red) are the areas of recognition selected as results of information processing. The minimum area of objects are shown with yellow rectangles. The head and tail parts of objects are also distinguished. The big blue dot is the center of gravity of objects found using the procedure described in pt.4. After these points were determined the tool was captured by robot using standard control programs.

The workflow of the collaborative robotic system begins after the voice command of the surgeon. For example, "put tool one in hand." Note that the tool may also be called by surgeon directly (tweezers, scalpel, etc.). Robot recognizes the voice command and confirms the correctness of the surgeon’s voice command using the state machine described above. In the case when the voice command is incorrect, the interactive system will ask the surgeon to re-enter the voice command, or refine it. If the voice command is correct, the robot performs the task using visual communication tools (see clause 5) and transmits the tool to the surgeon (Figure 7 c)).

Conclusion

The results of the experiments show that the principles of collaborative management proposed in the paper and the algorithms developed are realizable and quite effective. However, further research is necessary before the results may be recommended for practical use. First of all they should include the issues of safety and reliability of using the collaborative system “surgeon - robot assistant” in practice.

References

1. Jobannes Bodner, Florian Augustin, Heinz Wykypiel, Jobn Fisb, Gilbert Mueblmann, Gerold Wetscber, Tbomas Scbmid, The da Vinci robotic system for general surgical applications: acritical interim appraisal. SWISS MED WKLY, 2005, p.p. 674–679
2. Yuschenko A.S., Dialog control robots based on fuzzy logic. - Proceedings of the International Scientific and Technical Conference «Extreme Robotics», St. Petersburg, 2012, p.p. 29-36.
3. Yuschenko A.S., Morozov D.N., Zhonin A.A. Speech control for mobile Robotic systems: Proc.of 4th International Conference «Mechatronic Systems and Materials» MSM-2008, Byalostok, Poland, July, 2008, p.p. 14 –17.
4. Jakub Kanis., Dmitry Ryumin., Zdenek Krnoul. Improvements in 3D Hand Pose Estimation Using Synthetic Data, Interactive Collaborative Robotics, Germany, 2018, p.p. 105– 115
5. Ivan Gruber., Dmitry Ryumin., Marek Hruz., Alexey Karpov. Sign Langugage Numeral Gestures Recognition Using Convolutional Neural Network, Interactive Collaborative Robotics, Germany, 2018, p.p. 71– 77
6. Yu Z. S., Kobayashi H. An Efficient Forward-Backward Algorithm for an Explicit-Duration Hidden Markov Model//IEEE Signal Processing Letters, 2003, p.p. 11 – 14.
7. Tu S. Derivation of Baum-Welch Algorithm for Hidden Markov Models. URL: https://people.eecs.berkeley.edu/~stephentu/writeups/hmm-baum-welch-derivation.pdf
8. Tao C. A generalization of discrete hidden Markov model and of Viterbi algorithm. Department of Computer Science, 1992, p.p. 1381–1387.
9. Arisoy E., Sainath T., Kingsbury B., Ramabhadean B. Deep neural network language model. In Proceedings of the Joint Human Language Technology Conference and the North American Chapter of the Association of Computational Linguistics Workshop, 2012, p.p. 20 – 28.
10. Rabiner L., Juang B.H. Fundamentals of Speech Recognition. Prentice-Hall, Upper Saddle River, 1993, p.p. 321 – 386.
11. Yushchenko A.S. Intellektual planning in robot’s operation. Mekhatronika, avtomatizatsiya, upravlenie. 2005. No.3. Pp.5 – 18.
12. Huang J., Rathod V., Sun C., Zhu M. l., Korattikara A. Speed/accuracy trade-offs for modern convolutional object detectors. Computer Vision and Pattern Recognition, 2017.
13. Alexe B., Deselaers T., Ferrari V. Measuring the objectness of image windows. *TPAMI*, 2012.
14. Uijlings J., van de Sande K., Gevers T., Smeulders A. Selective search for object recognition. *IJCV*, 2013.
15. Endres I., Hoiem D. Category independent object proposals. In *ECCV*, 2010.
16. Carreira J., Sminchisescu C. CPMC: Automatic object segmentation using constrained parametric min-cuts. *TPAMI*, 2012.
17. Lowe D. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.
18. Dalal N., Triggs B. Histograms of oriented gradients for human detection. In *CVPR*, 2005.
19. Cortes C., Vapnik V, “Support vector machine,” *Machine Learning*, vol. 20, no. 3, p.p. 273–297, 1995.
20. Zhou H., Huang T. S. Tracking articulated hand motion with Eigen dynamics analysis. Proc. of the Intern. conf. on computer vision, Nice (France), 14–16 Oct. 2003. Washington DC: IEEE Computer Soc., 2003. V. 2. p.p 1102–1109.
21. Yoruk E., Konukoglu E., Sankur B., Darbon J. Shape-based hand recognition. IEEE Transactions on Image Processing, 2006, p.p. 1803 – 1815.

A. Nikolaev

PROSPECTS OF DEVELOPMENT OF NEW CYBERPHYSICAL SYSTEM BASED DIGITAL PRODUCTION PARADIGM - "INDUSTRY 4.0" IN THE WORLD AND RUSSIA. EXPERT OPINIONS

The Russian State Scientific Center for Robotics and Technical Cybernetics, Saint-Petersburg, Russia
a.nikolaev@rtc.ru

Abstract

Based on general publications, the brief analysis of main trends of the future industrial production in the "Industry 4.0" paradigm. Analysis demonstrates that the basis of the forthcoming economic order will consist of cyberphysical systems which will be simultaneously considered both as the intellectual machines (robots) and information technology concept.

It assesses the up-to-date economic situation in Russia focusing on the machine building industry as to the "Industry 4.0" development. It reveals the main reason for reduction in machine building product output in Russia – downturn of internal consumption, both the investment and end consumption. It emphasizes the difference in approaches to the economy digitalization in Russia and leading global industrial states.

The conclusions list the main features of the global macroeconomic development directly influencing the production sphere, new fields of production development and transformation effects in approach to the new concept "Industry 4.0".

The main result of the analysis is the substantiation of the Russia's current unreadiness for active participation in creation of new production and technology order based on the domestic economy digitalization.

Keywords: Industry 4.0, cyberphysical systems, robots, industrial production, economy digitalization, Internet of Things, product customization, digital production.

Acknowledgments

The results were obtained within the fulfilment of the state assignment of the Russian Ministry of education and Science No. 075-00924-19-00.

Cyberphysical systems are the next stage of the technological development of the society

Formation of production infrastructure in the specific time period is determined by the economic demands of the society and achieved level of technology development.

The economic demands at the modern stage is characterized by two main trends: the first one consists in the efficient, from the perspective of profit-making, satisfaction of the society with the customized products with the mass production industrial equipment. The second trend is related to the more increasing need in unique products and equipment requiring record accuracy, purity characteristics, operation parameters, etc., for aerospace, medicine, precision weapons, nuclear and other extreme applications.

The foregoing trends occurred in the conditions of appearance and rapid development of information technology supported by the progress in creation of micro and nano-electronic components and devices.

A particular effect on transformation of the industry functioning is expected from the Internet of Things (IoT) that has been already partially used in specific industry and service sectors.

The most comprehensive definition of IoT integrating the definitions and features that are the most common in multiple publications is given in [1]: "IoT is the dynamic global network infrastructure with self-sustained set of functions based on standard and compatible communication protocols where the physical and virtual things have identifiers, physical attributes and virtual personalities, use the intelligent interfaces and are easily integrated into the information network."

The requirement of the cost efficient satisfaction of the society's need in the product customization with the use of industrial equipment intended for mass production leads to formation of a new view to production as the "smart" system characterized by decentralization, self-sufficiency, re-configuration, and distinguished by continuous information exchange between subsystems and products in the framework of the performed production and logical processes.

Therefore, the IoT must serve as an information connecting "skeleton" of new industrial order.

In 2012, General Electric (GE) published its vision of IoT development in the industry - the "industrial Internet" [2, 3] where the physical objects, the "intelligent machines", are interconnected into a network at the component level. Within these networks, the "intelligent machines" can communicate with each other and people irrespective of the time and place they are. In other words, the "intelligent machines" are the particular combinations of industrial equipment and IT where the boundaries between people (consciousness) and machines are indistinct.

The National Institute of Standards and Technology (NIST), instead of the term "intelligent machines", used the term "cyber-physical systems (CPS) to be defined as follows [2, 3]: "smart systems covering the computational (i.e.) and efficiently integrated physical components closely interacting in order to feel the changes in the real world." As the examples of cyber-physical systems, the NIST considers the robots, intelligent buildings, medical implants. Self-controlled vehicles, unmanned aerial vehicles (UAV) and so on. Like the GE, the NIST treats the joint operation of such systems and people as one of the most import tasks in the CPS design, development and management. The key role here belongs to determination and simulation of "situational understanding" (circumstances of the system) that can have crucial significance in decision-making. According to the NIST vision, the cyberphysial systems form the basis for innovative production development.

In [4], the CPS is defined as the "information technology concept supposing the computing resources integration into physical processes. In such systems, the sensors, equipment and information systems are connected in a single cost creation chain moving beyond a company or business. These systems interact with the help of standard Internet protocols for forecasting, self-setting and adaptation to changes."

As the above definition states, the CPS is a large number of interconnected hardware and software components simultaneously functioning in different spatial and time conditions that are changed online, i.e. the distributed system, that requires for new approaches and techniques in the control system.

According to the author [5], when using the distributed systems, such widely used techniques as hierarchic, matrix, situational and network-centric control do not provide for sufficient efficiency: "This problem is enhanced by the problem of the "big data" setting large volumes of data and large number of links increasing in geometrical progression", and then "only the intelligent control appears to be efficient for complex distributed systems. "And then: "... The cyberphysical systems are the distributed systems with the function of intelligent processing and re-configuration of streams due to intellectual control." "The efficiency of use of the cyberphysical systems can be compared with the efficiency of application of parallel computing systems. Parallel computations are efficient for large volumes of information with the function of data stream parallelization. In simple tasks, the parallel techniques are not efficient, but the linear computing techniques are. Just in a similar way, the cyberphysical systems are efficient for the complex distributed system control and complex task solving. For simple tasks, the CPSs are not efficient like the parallel computing systems."

By their architecture, the modern robotic systems are CPS with the integrated (control) system, mainly. Moreover, vice versa, that is even more conventional, CPSs mean the robots most often today.

However, as the functionality of mass consumption of the customized products and demand in the life cycle technologies develops, the CPS will be treated as the information technology concept of the distributed system with a function of intelligent processing and re-configuration of streams due to intelligent control."

"Industry 4.0" as the cyberphysical system based organizational and technical production architecture. Robots

The CPS information technology concept covering the industrial production process is called as "Industry 4.0" that, in general, is characterized as global multilayer organizational and technical system based on the integration of physical operations and concurrent processes into a single information space [5].

In the "Industry 4.0" concept, the conventional data sources give way to CPS and IoT by the VVV parameters (a complex of defining characteristics of the Big Data: Volume - physical data volume, Velocity - data volume growth rate , Variety - variety and simultaneous processing of different data types). The CPS is the basis for industrial IoT where the PLM (Product Lifecycle Management - application for product/item life cycle control) structured and controlled information for the Big data is created. As a result, the individual production stages/operations are connected from the product design and production resource planning to the field testing of the mechanisms. The most important element of the organizational and technical concept "Industry 4.0" is the functional compatibility (interoperability) of its elements [6].

According to one of the leading specialists in the Internet of future, Prof. Wolfgang Wahlster, Chief Executive Officer, as well as Director for Science and Technology of the German Research Center for Artificial Intelligence (DFKI) in Kaiserslautern, Saarbrücken, Bremen and Berlin; member of the R&D Union of the Federal government; Chairman of the EU Higher Advisory Group for the Internet of future (FI-PPP), [7], the CPS change radically the traditional production logics, namely: the working object will individually define what operation should be done. Production equipment and products will become active system components that control their production and logistics processes. They will differ from the existing mechatronic systems by the opportunity to interact with the environment, by planning and adapting their behaviour as to the surrounding conditions, by being capable of self-optimization. This will make it possible to ensure efficient release of minimum product batches in a number of options with quick changing of the final products. Applying the actuating mechanism integrated sensors, ensuring the inter-machine data exchange and using the active semantic memory will lead to appearance of new optimization techniques aimed at resource preserving in the production environment.

Successful, in terms of its concision, definition of “Industry 4.0” was given by the Russian representatives of Bosch Rexroth LLC:

“Industry 4.0” shall mean highly self-sustained decentralized re-configured production that is distinguished by continuous data exchange between its subsystems and production objects in the frameworks of the performed production and logistics processes.”

According to the authors [8], the “Industry 4.0” industrial concept is a multilayer organizational and technical system combining six subsystems:

– PLM – Product Lifecycle Management: domain-specific application software packages intended for structuring the array and automation of the physical and information process control during the whole product lifecycle, including production cooperation,

– Big Data: a complex of approaches, instruments and methods of bulk data processing and large variety of data intensive production systems that is intended for receiving the results efficient in the avalanche-like data stream by a human or computer system.

– SMART Factory – a well-thought out factory (a mnemonic abbreviation used in management to define the goals and set the tasks): a concept of seamless combination of individual production stages (operations) are connected from the product design and production resource planning to the field testing of the actuating mechanisms. The Smart Factory is based on the concept of Digital Manufacturing that characterizes the organizational and technical system of the adequate production modelling based on the advanced CAD (Computer-Aided Design) simulators.

With the development of adaptive cognitive II based systems, the Smart Factory will significantly change. If in the “Industry 4.0” the key role of a human is in the algorithm development and teaching of machines by computer programming method, the self-teaching of machine, copying the actions of a human or other robots and automatic optimization of production algorithms will be the basis for new production order, “Industry 4.0”.

– CPS – cyberphysical systems: can be both in the frameworks of a single company, and dynamic business model consisting of several companies. The IoT cannot exist without cyberphysical systems as they appear with its infrastructure,

– IoT: links the Internet connected things and ensures their joint work controlled by cloud computing systems; ensures interaction of physical production operations and concurrent processes.

– Interoperability – functional compatibility: the “Industry 4.0” serviceable integrated system cannot be created without functional compatibility.

It should be noted that currently “Industry 4.0” has not been implemented anywhere in the world on an industrial scale, and only its individual elements are used. In the developed countries, intensive works are being in progress to prepare for the companies’ large-scale transition to the cyberphysical production principles. Many countries invest a lot of money in the corresponding developments. For example, the UK government started the £5 mln. IoT project implementation, the German government invests about €200 mln. in this sphere. Yet in 2008-2009, Japan launched the “u-Japan” and “i-Japan” strategies to prepare for IoT implementation into daily life; in the EU, the European Research Cluster on the Internet of Things (IERC), FP7, offered a number of projects for the Internet of Things and created the International IoT Forum for development of common strategy and technical vision of the IoT use in Europe. In the USA in 2012, the non-profit Coalition of Smart Production Leaders was created. It included the industrial manufacturers, suppliers, IT companies, state authorities, universities and laboratories. The purpose of organization is to create a smart platform for industrial IT applications. In 2014, General Electric, AT&T, Cisco, IBM and Intel set up the Industrial Internet Consortium currently combining 170 organizations. The purpose of non-profit union is to remove the barriers between different technologies in order to ensure maximum access to the big data and improve integration of physical and digital environment. The similar programs have been launched in the Netherlands, France, UK, Italy, Belgium and other technologically advanced countries.

Success of “Industry 4.0” is in many aspects dependent on standardization that must provide for interoperability, compatibility, reliability and efficiency of the integrated devices globally. Due to the development of the common standards, the designers and consumers can widely use the IoT applications and services while saving costs for the IoT service and expansion on a long-term horizon. The International Intercommunications Union, International Electrotechnical Commission, International Organization for Standardization, Institute of Electrical and Electronics Engineers, European Committee for Electrotechnical standardization, Chinese Institute for Electronic Standards, and American National Standards Institute are engaged in development of different standards for the Internet of Things. This is complicated by a need in agreeing standards of different organizations with the international standards, as well as national and regional organizations for standardization [1].

The German specialists forecast the appearance of the first companies complying with the “Industry 4.0” principles not earlier than by 2020, and complete transition to the wired industry is planned for 2030.

It was noted above that “Industry 4.0” is the CPS information technology concept covering the industrial production process and is a global multilayer organizational and technical system based on the integration of physical operations and concurrent processes into a single information space.

This concept/system implementation is not possible without its constituent subsystems (Smart Factory, IoT, CPS) and, as a result, their material and technology basis - robotics that is characterized by a comprehensive application nature. In addition, the robots as smart machines represent the CPS by themselves, but like in the “shortened” version with the integrated control system (including the II based) and data exchange with the environment.

Therefore, the robots are like puzzles which will be the basis for future “picture” of the industrial Internet (or “Industry 4.0”), but now already having the widest and variable application in different spheres of industry and services.

At the modern industrial production stage with the prevailing of the comparatively simple tasks of product manufacture and process organization localized in the limited space, the most efficient way is to use the CPS in the “shortened” form - robots, robot packages or systems. As the needs of the society in changing customized products which modification also depends on the geography of its consumption develop, the CPS concept in the “Industry 4.0” architecture with complex tasks of the distributed system control is becoming more efficient.

Economic situation of Russia as to the “Industry 4.0” development

The general assessment of the modern economic situation of Russia in the general publications is rather contradictory.

For example, recently, the domestic mass media widely informed of the Russia’s second place (after Malaysia) in the new rating of comparative indicator of the current situation and economic perspectives of the developing countries (markets) published by Bloomberg. See, for example [9]. It was stated that during the rating preparation a number of parameters had been taken into account, among which were the GDP growth forecast for 2018-2019, state of the current transaction account, sovereign credit rating, stock and bond market, foreign exchange reserves, etc. By the way, it is not clear why China takes the third place in this rating although it supersedes Russia by its economic power by multiple times.

Simultaneously, a number of articles appeared representing the Russian Federation economic situation as not very optimistic.

Particularly, [10] the results of survey “Barometer of Confidence of the Companies” conducted by the audit and consulting group EY (the research concurred with the round of the American sanctions against the large Russian business in April; 60 managers of the Russian companies from 14 economic sectors took part in the survey) that show that 84% of small and medium business owners in Russia are ready to refuse from some part of assets due to their inefficiency or any revealed risks, and only 32% of investors are ready to invest in new projects. According to 55% of respondents, the main risk threatening the business development is the geopolitical tension: The research points out that: “The crisis has weakened the small and medium business suffering from restrictions introduced regarding Russia, as well as from the regulators’ uncertainty and requirement strengthening. And they are not supported by the state like the large business.”

The expectations of business regarding the situation improvement are related to the return of direct investment funds as the main assets buyers. In June 2018, the Head of the Russian Direct Investment Fund (RDIF), Kirill Dmitriev, promised to invest more than 7 tln.RUB in the infrastructure and high-tech projects in Russia. For these purposes, in the end of 2017, a new investment platform “RDIF Technology” was created.

The main risks for market development are still the political uncertainty in economy due to the rouble volatility, dependence on the oil and gas sector, increased VAT and decreased consumer activity [10].

The negative effect on the Russian economy is also due to the capital outflow initiated by toughening the money and credit policy of the US Federal Reserve System, strengthening the foreign trade restrictions, economic crisis in Turkey, sanction police of the West as to the Russia and dependence of the Russian economy on the situation in the oil and gas markets. According to the Central Bank, based on the seven month results (January - July) 2018, the net capital exports by the Russian private sector increased by $21.5 bln., while for 2017 in general the similar indicator was $27.3 bln. ($18.5 bln. – in 2016). As a result, the Russian Ministry of Economy and Development reviewed the forecast for capital outflow in 2018 - by $41 bln. instead of $18 bln. estimate [11].

High poverty rate of Russians should be also noted. It does not promote for development of the mass domestic demand in goods and services. According to the Manager of the Accounts Chamber, Aleksey Kudrin [12], 19.3 mln. of Russians, as of 01.01.2018, had income lower than the minimum of subsistence which by the beginning of 2018 averaged to RUB 10,328 per head (Order of the RF Government dd 08.12.17 No. 1490). As a result, the consumption structure of the ordinary Russian citizen is shifted towards the food products. If in the USA the food products make about 6% of the consumption structure, and 12% considering the public catering, the most of the domestic households show more than 50%.

The general economic situation is in many aspects determined by the situation in machine building the "Industry 4.0" development is directly related to.

Here we use the materials provided in the analytical report [13] and data from description of the sphere of implementation of the subprogram "Production Development of Production Means in the RF" of the State Program of the Russian Federation "Industry Development and Competitive Growth" (Order of the RF Government dd 15.04.2014 No. 328 (rev. on 31.03.2017)).

According to referenced sources, the machine building includes the following fields: machine tool industry and investment machine building (power engineering, electrical engineering and cable engineering, oil and gas and chemical engineering, heavy engineering), car building, transport and special-purpose machine building, machine tool production, radio-electronics and instrument making industry, aviation industry and ship-building industry.

During five years (since 2014), the Russian machine building industry showed negative output dynamics. In 2015, according to the RIA rating, the decrease made 8.9 %, and in 2016 – 0.9 %. The machine building products export, irrespective of devaluation of the rouble, has reduced by 4.3 %. i.e. $24.3 bln., according to the RF Federal Customs Service (FCS).

The main reason for reduction in machine building product output in Russia is downturn of internal consumption, both the investment and end consumption. Due to the output negative dynamics and high volatility, there appears a risk of restored degradation of production capacities of the machine building enterprises, and machine tool fleet, first of all. Degradation of production assets appears both in reduction of the total tool fleet, and increased share of the worn-out equipment. That is the reduced number of equipment is not compensated by the increase in effectiveness of the remaining ones. Now, the average age of the production equipment in the domestic machine building industry exceeds 20 years that determines the high level of wear: appr. 15% of the machine tool fleet and other equipment are worn and must be decommissioned. Great underrun is observed for the CNC machinery: if in Japan this class comprises more than 90% of machinery, in Germany and USA - more than 70%, and in China - about 30%, in 2016 in Russia, the CNC machinery share made less than 10%.

The reduction in quantity and degradation of quality of the production capacities in the Russian machine building industry are progressing faster than the staff reduction. As a result, while at the Russian enterprises, 1 piece of machinery falls on 4.7 workers of the machine building industry, in the EU - 0.8 worker, and the labour capacity in the EU machine building industry exceeds the Russian one by 6 times, and at this, the expensive investment machine building products prevail in the Russian machine building product output structure.

The consequence of the above mentioned indicators are the low competitive ability of the manufactured products both by price and quality. As a result, only a small share of the existing demand is available for the Russian machine building enterprises in the global market, and it is restricted by: firstly, share in the chain of new cost creation (among the world manufacturers), secondly, low share of the value-added cost in the products by the Russian enterprises, as such. This is mainly the assembly and production of inexpensive components and materials.

In the defense and investment machine building market, the share of the products manufactured by the Russian enterprises is considerably higher. But there is the other problem - the lesser volume of these markets as compared to the mass ones, their high volatility, as well as high share of imported components (mainly, from China).

Low competitive ability makes it impossible for the domestic machine building enterprises to recover investments in the fixed assets and production processes as the result of the restricted sales market of Russia is the low level of utilization of production capacities as to the work hours - about 20% that is absolutely non-acceptable considering the world competitors with the indicator exceeding 90%.

The low utilization of production capacities leads to inefficient automation without changing the production organization and product sales principles. It is obvious that the investments in the means of time reduction for facilities preparation and utilization with the most optimal route become unreasonable if there is no sufficient facilities for production equipment to ensure its optimal utilization. In other words, in the situation when the utilization of the expensive production equipment makes 20-25% of the work hours, it is not reasonable to invest in accelerating the preparation of facilities for machinery and optimizing their repairs, as there is 75-80% of free time for these operations, and the situation will remain the same if nothing is done to increase the flow of production facilities [13].

Therefore, the small size of the available (with the existing prime cost and quantity) sales markets is the primary reason for crisis in the Russian machine building industry, but not the low marginality and expensive loans [13].

The additional negative factor is that the partially upgraded production assets are scattered among separate, non-related production facilities. As a result, there is no possibility, even in particular production facilities, to build a through automated production chain (and this is the approach providing for utilizing the CNC machinery and processing centres in full) not speaking about building of such chains with the use of facilities of various enterprises. The most production facilities provide only for the domestic needs and are optimized only for own developed products; the contract production is practically absent.

All the mentioned above mean that if no efficient actions for the domestic machine building development will be implemented in the nearest future, than the retirement of machinery without increased utilization of the remaining ones will lead to decrease in the machine building output irrespective of the market situation. Moreover, the enterprises will have to reduce the excess staff. This may result in that 2-3 mln. people will lose their jobs without any opportunity to find a new ones, as the staff reduction in the industry is not compensated by the increased demand in specialists in the other sectors of economy.

Such degradation due to transition of the largest world machine building manufacturers to the PSS model (Product-Service Systems are single systems integrating the physical product as such and all processed related to its production and operation) will be destructive not only for the material production but for the sphere of services which share exceeds 55% of the national GDP.

In Russia, more than a half volume of the service market is directly related to re-sale and servicing of the imported equipment, that is why the transition of the largest world manufacturers to the PSS model when they gain direct access to the end consumer means that the entire chain of domestic mediators loses such access and goes under the manufacturer's supervision. That is why the absence of own competitive production in Russia in the PSS paradigm will mean the absence of the national sphere of services related to the machine building products sale and servicing.

The currently existing in Russia concepts of ways out of the situation with the help of digitalization of fields are solely focused on creation of new kinds of services based on collection and analysis of data from different physical objects and do not cover the issues of fundamental change in approaches to the lifecycle of these objects.

As distinguished from the Russian approach, according to the world industrial states (USA, Germany, Japan), the "digital economy" means the processes of creation and use of the PSS that is the products which are initially designed as a single system integrating the physical product as such and all processed related to its production and operation. Thus, the appearance of the IoT service is the consequence of such approach but not a self-sufficient field of economic activity (like Big Data and other services).

Similarly to the total automation of design, engineering, production and operation of such systems, this is only the way of the PSS concept implementation allowing, fully or partially, to eliminate the negative effect of human factor on consumption properties of the product service system and make them completely controlled. The program determinability ("cyberphysical systems" in the "Industry 4.0" concept) and modularization of the "product" part of the PSS are the single way for economically efficient creation of such systems [13]. Such approach is implemented in the form of brand new business models, for example, Performance-Based Lifecycle Product Support – the lifecycle contract for the machine building products, Circular Economy - production of the products modified at the operation stage and having no disposal stage.

Conclusions

The main features of the global macroeconomic development that directly influence the production sphere are currently as follows:

– increasing customized demand that leads to transition from large volumes of small nomenclature to low volumes of large nomenclature of the products manufactured and forms a demand for expansion of the production facilities and proposition of services,

– reduction of the lifecycle (reduced time of use) of the mass products with its simultaneous modification that requires flexible automation of production and flexible efficient technology in the sphere of services,

– global competition requiring for constant increase in the production efficiency,

– required achievement of record-rate quality characteristics and/or consumption properties of individual products in high-tech industries, such as precision weapons, medicine, aerospace, nuclear energy, and so on,

– increasing salary being average for industries and related with objective trend in improvement of the living standards of society,

– aging of population,

– rapid development of information technology, first of all, IoT in the background of success in creation of micro- and nano-electronic components and devices,

– formation of new paradigm of the industrial production as the most efficient in the conditions of the above mentioned factors, this is "Industry 4.0", cyberphysical system based organizational and technical production architecture.

Appearance of new production concept (both for products and services) creates new fields in development, namely [14]:

– decentralization and more flexible scale control for cost saving; transition to through automation processes;

– providing the things with AI functions conversion of each thing simultaneously into the consumption object (end or intermediate - for further processing) and into information source; participation of such smart things in own structuring and repair;

– conversion of the industry of services into the field controlled by interaction of client and service end II with the use of the dig data as the information source for forecasting and planning;

– reduced participation of a human in the interaction of things;

– creation of infrastructure of the alternate reality and its communication protocols with smart things and devices;

– expansion of "passive entrepreneurship" of population due to creation and development of electronic trade systems and other resources, first of all, in utility services;

– blockchain technology development;

– development of alternative networks, like Internet, and their integration into the alternate reality infrastructure.

As to the machine building industry, the transformation effects of the "Industry 4.0" implementation can be presented by three groups [13]:

– the appearance of the manufactured machine-building products changes, they become more program-determined and modular:

1) the external form of products is simplified;

2) constant modification of products due to upgrading of their software and replacement of individual modules becomes possible (platform-based principle).

– approaches to product development change:

1) "the human factor" is reduced to zero;

2) maximum reduction of time and resources at the production stage (robotization, automatic control) and operation stages (remote automatic monitoring) is provided.

– approaches to production change: organization of "hybrid" model for utilization of production capacities becomes possible. Using such model, the enterprise having production facilities use them not only for own production, but for individual production operations ordered by third-party companies, thus being a part of the distributed production chain combined by a single automatic control. The third-party companies preserve full control over the quality of their "remote" operations. This makes it possible to maintain high, near 100% utilization of production facilities.

Today Russia is not ready actively to participate in creation of new production and technology order. This is due to the increased compatibility of the products, terms of order fulfilment and production quality stability required by the main domestic machine building customers, as well as achievement of the service level to ensure for transition to the PBL model (Performance Based Lifecycle Product Support) in the relations with the consumers are possible only with the implementation of the cyberphysical systems and IoT into domestic machine building industry. Transition to PBL will require for absolute review of design, production and operation approaches.

[13] states that the R&D analysis conducted by the leading universities of the industrially developed countries shows their activity in development of scientific and practical approaches to PSS task-solving according to the market requirements and development of business models to monetize the PSS advantages. The technology part, particularly, in information technology (big data, II, etc.), is not so much considered as it is thought not to be the main restraining factor preventing the transition to product service model. In Russia, the studies of problems of the PPS creation and monetization are not conducted, as a result, all latest initiatives in the economy digitalization looking like groundbreaking have, in fact, no key success factors, that is understanding the appearance of the competitive product and business models of its monetization.

According to the authors [13], this the underlying reasons of the "programmed failure" of the currently made efforts on the Russian economy digitalization as they are.

References:

1. Li Da Siy. Internet of Things in the Industry: Review of Key Technologies and Trends [Electronic resource] / Li Da Siy, Wu Hem Sianchan Li; trasl. A. Osotov // PT Electronics: [Web-site]. – Access mode: https://ptelectronics.ru/stati/ (applied on: 03/12/2018).
2. Spimes, Cyberphysical Systems and Industry 4.0 [Electronic resource] // Internet of Things.ru: [Web-site]. – Access mode: http://internetofthings.ru/issledovaniya/33-spimes-kiber-fizicheskie-sistemy-i-promyshlennost-4-0 (applied on: 03/12/2018).
3. Ben van Lier. Spimes, Cyber Physical Systems and Industrie 4.0 [Electronic resource] // Centric: [site]. – 2013. – URL: https://www.centric.eu/NL/Default/Themas/Blogs/2013/06/27/Spimes-Cyber-Physical-Systems-and-Industrie-40-[1]- (applied on: 03/12/2018).
4. Cyberphysical system [Electronic resource] // Wiki 2: [site]. – 05/11/2018. – Access mode: https://wiki2.org/ru/ (applied on: 03/12/2018).
5. Tsvetkov V.Ya. Cyberphysical system // International Journal for Applied and Fundamental Studies. – 2017. — No. 6-1. – P. 64-65. – Access mode: https://applied-research.ru/ru/article/view?id=11623 (applied on: 10/12/2018).
6. Automated production control systems in the organizational and technical concept "Industry 4.0" [Electronic resource] // MPLM IG™ Industrial Engineering: [e-journal]. – 27/02/2017. – Access mode: http://www.plm.pw/2017/02/PLM-in-Industry-4-0.html (applied on: 05/12/2018).
7. Industry 4.0: Production process of future [Electronic resource] // production control: [Internet portal]. – Access mode: http://www.up-pro.ru/library/opinion/industriya-4.0.html (applied on: 22/08/2017).
8. 6 Components of Industry 4.0 [Electronic resource] // MPLM IG™ Industrial Engineering: [e-journal]. – 29/09/2016. – Access mode: http://www.plm.pw/2016/09/The-6-Factors-of-Industry-4.0.html (applied on: 05/12/2018).
9. Russia took the second place in the rating of developing countries by Bloomberg [Electronic resource] // Kommersant: [site]. – 11/28/2018. – Access mode: https://www.kommersant.ru/doc/3813786 (applied on: 10/12/2018).
10. Ageeva O. More than 80% Russian companies declared about their readiness to sell assets [Electronic resource] / O. Ageeva // RBK: [site]. – 04/09/2018. – Access mode: https://www.rbc.ru/economics/04/09/2018/5b8d33f39a7947867fceb7d6?utm_source=yxnews&utm_medium=desktop (applied on: 10/12/2018).
11. Kalyukov E. Ministry of Economic Developement increased the capital outflow forecast by $23 bln. [Electronic resource] / E. Kalyukov // RBK: [site]. – 05/09/2018. – Access mode: https://www.rbc.ru/economics/05/09/2018/5b8ffd9b9a794735f25e7a20 (applied on: 20/11/2018).
12. Markelov R. Golikave responded to Kudrikov concerning poverty rate [Electronic resource] / R. Markelov // Rossiyskaya Gazeta: [site]. – 28/09/2018. – Access mode: – https://rg.ru/2018/09/28/golikova-otvetila-na-slova-kudrina-ob-urovne-bednosti.html?utm_source=yxnews&utm_medium=desktop (applied on: 21/11/2018).
13. Market review Economic effects of digitalization and IoT implementation in the Russian machine building industry [Electronic resource] // JSON.TV. Technology. Investments. Telecommunications: [site]. – 17/08/2018 – Access mode: http://json.tv/ict_telecom_analytics_view/ekonomicheskie-effekty-ot-tsifrovizatsii-i-vnedreniya-iot-v-mashinostroenii-v-rossii-20180817013305 (applied on: 21/11/2018).
14. Shpurov I. Industry 4.0 article [Electronic resource] / I. Shpurov // Expert.ru: [site]. – 03/102016. – Access mode: http://expert.ru/expert/2016/40/industriya-4_0/ (applied on: 10/12/2018).

Proceedings

of the International Scientific and Technological Conference

EXTREME ROBOTICS

June 13-15, 2019, Saint-Petersburg, Russia

Труды

Международной научно-технической конференции

ЭКСТРЕМАЛЬНАЯ РОБОТОТЕХНИКА

13-15 июня 2019 года, Санкт-Петербург

Подписано в печать 01.10.2019
Формат А4. Печать – цифровая.
Тираж 600 экз. Бумага офсетная.
Отпечатано в
ООО "Издательско-полиграфический комплекс "Гангут"
с оригинал-макета заказчика