ARCHITECTURE OF A VIDEO ANALYTICS SYSTEM USING PARALLEL PROCESSING

A. V. Gorshkov (1), O. Ja. Kravets (2), I. A. Aksenov (3), Yu. V. Redkin (4), I. V. Atlasov (5)*

(1) Research Institute of Computing Complexes named after M.A. Kartsev, Moscow; (2) Voronezh state technical university, Voronezh; (3)Vladimir State University named after Alexander and Nikolay Stoletovs, Vladimir; (4)Admiral Ushakov Maritime State University, Novorossiysk; (5)Moscow University of Ministry of Internal Affairs of Russian Federation named by V. Ja. Kikot, Moscow Russian Federation

* Corresponding Author, e-mail: csit@bk.ru

Abstract: Modern monitoring and operational decision-making systems based on image series have become widely used for a wide variety of applications. At the stage of identification of graphic images, the affiliation of the found dynamic object to the class of objects of interest is established on the basis of a comparative analysis of its contours with a given template. The article continues the cycle of research in which the identification process is considered and the algorithm of machine graphics is given. A description of the interaction of the structural components of the video analytics system is presented.

Key words: motion vectors, image series, fast processing, analytical solution, machine graphics.

1. INTRODUCTION

Modern monitoring and operational decision-making systems based on image series have become widely used for a wide variety of applications. At the stage of identification of graphic images, the affiliation of the found dynamic object to the class of objects of interest is established on the basis of a comparative analysis of its contours with a given template. The article continues the research [1-3].

The video analytics system shown in Figure 1 includes the following structural components [1]:

- a set of video cameras,
- image capture and preprocessing module,
- dynamic object detection module,
- object recognition module,
- databases and knowledge bases,
- decision-making system, etc.

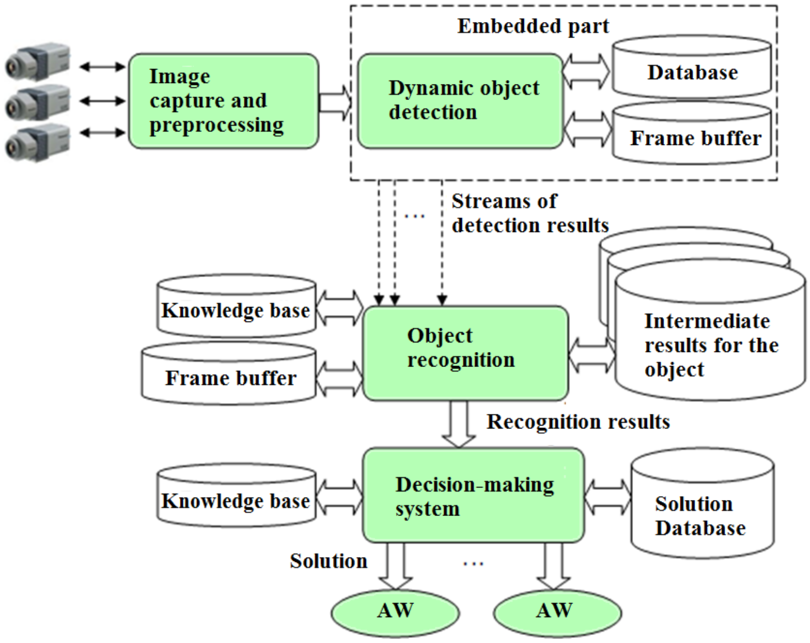

Figure 1. The video analytics system: AW - Automated workplace

Within the framework of this study, it is assumed to consider a video analytical system that includes a set of static cameras for continuous observation of a specific space of the real world. In other words, video recording is assumed within the same scene, where there are static objects marked as a common background. All changes occurring inside the scene are represented as areas of interest where the presence of a dynamic object is assumed.

The dynamic object detection module is a mathematical and image processing software embedded in a video analytics system in order to increase the efficiency of detecting dynamic objects. This component is an intermediate link for organizing interaction between software subsystems: the image capture and transformation module, discussed in more detail in Section 2, and the recognition module described in Section 3. The structure of the components of the dynamic object detection subsystem and their interaction is demonstrated in Section 4. The input data of the video analytics system is the sequence of images, as well as the settings of the detection and recognition modules. The output information is the decision received from the decision-making system in accordance with the decision-making rules based on the results obtained from the image recognition module. The decision-making system is described in Section 5.

2. IMAGE CAPTURE AND PREPROCESSING MODULE

The image capture and preprocessing module adjusts, receives and controls the video stream from CCTV cameras, as well as converts images into the required input format for further analysis.

The diagram of the structural components of the module in question, shown in Figure 2, includes the following components:

1. A computational module that includes a mathematical apparatus for calculating the resulting array of intensities and preparing input images for analysis.

In the case of a 24-bit input image, the conversion to the 8-bit Grayscale format (Grayscale) is carried out according to formula (1).

The formula for converting a 24-bit image to a grayscale format using the values of each of the channels R, G and B [4, 5]:

$$A_{ij} = \left[0{,}2989 * R_{ij} + 0{,}5870 * G_{ij} + 0{,}1140 * B_{ij} \right] \quad (1)$$

where A is the intensity matrix images; i, j – pixel coordinates.

2. The video data flow control module, which provides integration of the software environment with video surveillance cameras, interaction with the computing module, as well as sending the processed array of images.

Methods of organizing the interaction of the software environment and video surveillance cameras [6]:

- Real Time Streaming Protocol – a protocol for transmitting various streaming information in real time without pre-compression, supported by most modern IP cameras.
- Software Development Kit – a set of tools, interaction interfaces, descriptions, as well as examples of programs demonstrating integration with the hardware component provided by the manufacturer for a specific type of camera, usually machine vision.
- Open Network Video Interface Forum and Physical Security Interoperability Alliance – unified standards for working with digital video cameras, built on Web services in the WSDL language, as well as on RTSP, SOAP and compression standards H264, MJPEG, MJPEG-4 [7].
- HyperText Transfer Protocol is an application layer protocol that provides the transfer of an MJPEG video stream and interaction with an ip camera by means of Web requests and Web responses.

In video analytics systems, as a rule, machine vision video cameras or IP cameras are used. The method of integration into the system is determined depending on their type and manufacturer.

3. A video camera control module that provides the possibility of both one-way and two-way communication.

The input data of CCTV cameras are commands to control the gain settings, exposure, image transfer rate, etc. The output information is a sequence of images in various supported formats.

The block diagram of the components of the image capture and preprocessing module shows the following input and output parameters:

$I_i, i \in 1 \ldots n$ - stream of images coming from the camera *i*,

$C_i, i \in 1 \ldots n$ - a stream of camera i control commands,

$A_i, i \in 1 \ldots n$ - a parallel stream of processed images converted from input data coming from the camera *i*.

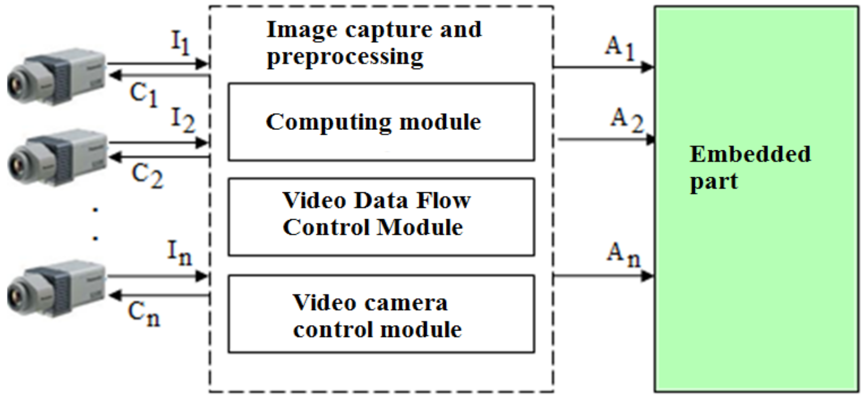

Figure 2. Block diagram of the structural components of the image capture and preprocessing module

3. OBJECT RECOGNITION MODULE

The pattern recognition module performs the function of detecting, identifying and tracking an object based on the analysis of video sequences.

The input data of the subsystem consists of a stream of images and recognition settings, including the following fields:

- static area of interest in the image;
- probability filter, which is the limit value of the probability at which an object can be considered recognized;
- scale range of the object of interest;
- recognition result;
- the period for issuing intermediate recognition results;
- the period after which, if the object has stopped appearing on the video sequence, a decision is made to lose it and issue a recognition result.;
- the angle of rotation of the image required if the camera is placed at an angle, etc.

The output data is the recognition results in a specific format depending on the specific system developer.

The recognition results include the following fields:

- the image on which the decision was made;
- image creation time stamp;
- the recognized object in the form of a set of its characteristics for unambiguous correspondence with the object of the real world;
- coordinates of the location of the object in the image, etc.

A distinctive feature of modern recognition systems is the consideration of each of the intermediate results of pattern recognition in relation to each other and making a decision after the object is lost from the visibility of the system [8].

Figure 3 shows the relationship of the databases of intermediate recognition results for each of the objects with the algorithmic part of the module. The knowledge base contains basic recognition settings, such as object shape templates and specific information from developers of a single recognition subsystem, as well as accumulated previous results.

The process of pattern recognition includes obtaining the detection result, clarifying the boundaries of the object within the area of interest, recognition and output of the result [8].

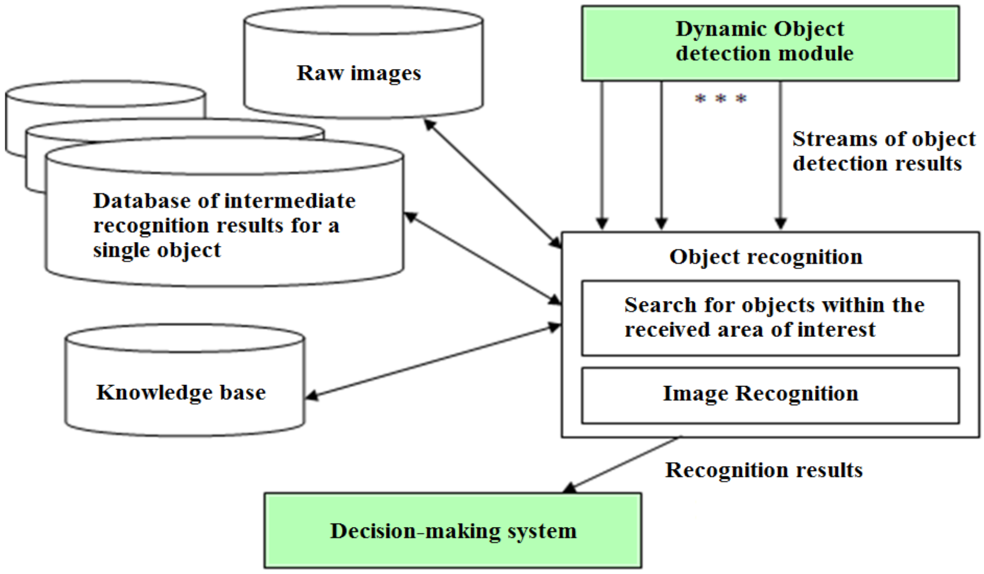

Figure 3. Data flow diagram of the pattern recognition module

The accuracy of recognition directly depends on the number of frames received of a real-world object and on the image quality.

Increasing the efficiency of the object recognition process is achieved by transferring already accumulated information about the object in the form of an image, the time of its creation and the area of interest [9].

Firstly, obtaining the location zone of the object in the image allows you to reduce the complexity of the search stage by the subsystem for recognizing the coordinates of the exact location, or abolish it altogether. Secondly, the described implementation makes it possible to take into account only the actual objects of interest of the system, avoiding false positives and increasing the accuracy of recognition.

4. DYNAMIC OBJECT DETECTION MODULE

Due to the operation of video analysis systems in real time, time constraints are imposed on the recognition subsystem. This means that the speed of image processing affects the accuracy of recognition and the number of fixes of high-speed objects.

The main purpose of the dynamic object detection module is to increase the number of recognized objects and reduce the number of false positives by transmitting information about the location of the object of interest on the frame to the recognition subsystem [1]. The detailed format of the output data is presented in Section 2 on the diagram of the relationships of program classes.

The embedded module performs parallel processing of graphic data for each of the received preprocessed video data streams. The parallelism model of graphic data processing is shown in Figure 4.

The main tasks of the detection module are, firstly, motion detection on video sequences, secondly, the selection of an area of interest and the search for an object corresponding to the shape of a predefined template on it, and thirdly, the formation and transmission of the detection result to the recognition subsystem.

The input data of the detection module are 8-bit grayscale images, as well as detection settings. The output data is the detection results received by the image recognition module.

The description of the input and output data formats is discussed in Section 2.

Figure 4 shows a model for processing a single stream of video data coming from one of the video surveillance cameras to the input of the embedded subsystem. Within a single stream, as shown in Figure 4, sequential image processing is provided up to the stage of calculating areas of interest [10].

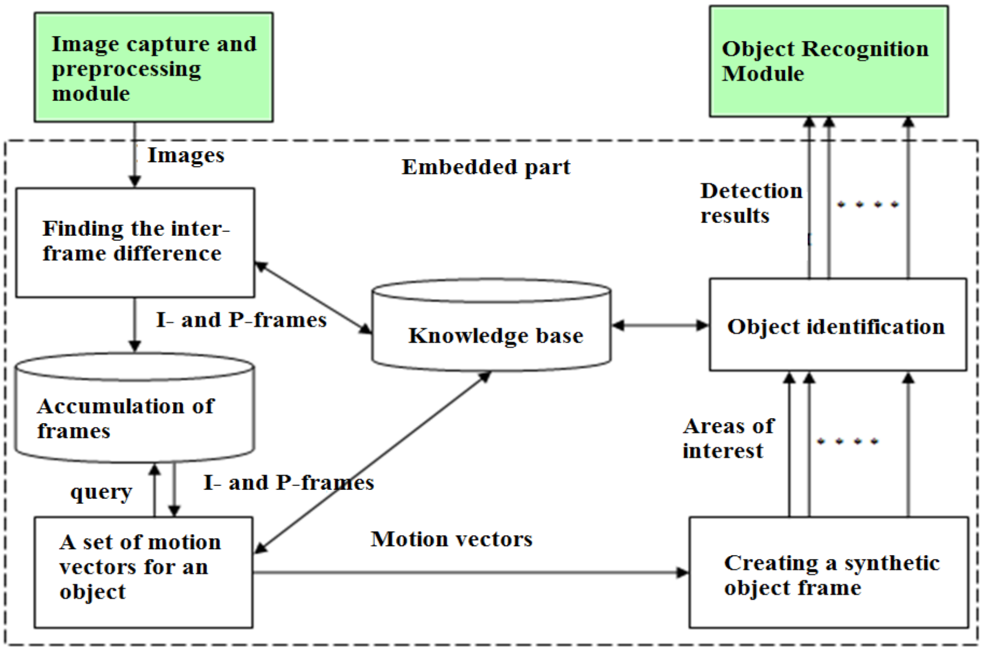

Figure 4. A model for processing a single stream of video data

Stages of processing incoming images to the input of the built-in subsystem:

1. Formation of key I-frames or intermediate P-frames at the stage of finding the inter-frame difference. The description of the algorithmic part of the stage is presented in [3].

2. Moving data to the buffer of intermediate and key frames.

3. If the number of frames is sufficient, then a set of motion vectors is formed for a separate object.

4. Calculation of areas of interest, which are synthetic frames of objects collected on the basis of sets of vectors of their movements. A description of the mathematical and algorithmic apparatus for finding vectors of block displacements and areas of interest is presented in [3].

5. Identification of the object for each of the areas of interest described in [3]. At this stage, the contours of the object are searched and compared with the standard loaded from the detection module settings. A detailed algorithm for machine graphics and image processing for recognizing the contour of an object is described in [3].

6. Output of detection results for each dynamic object found.

Since several areas of interest and dynamic objects can be recorded in one of the time intervals of observation of the real world, respectively, a mechanism for parallel data processing is provided within the framework of the subsystem under consideration. The parallelism model of data processing by the dynamic object detection module is shown in Figure 5.

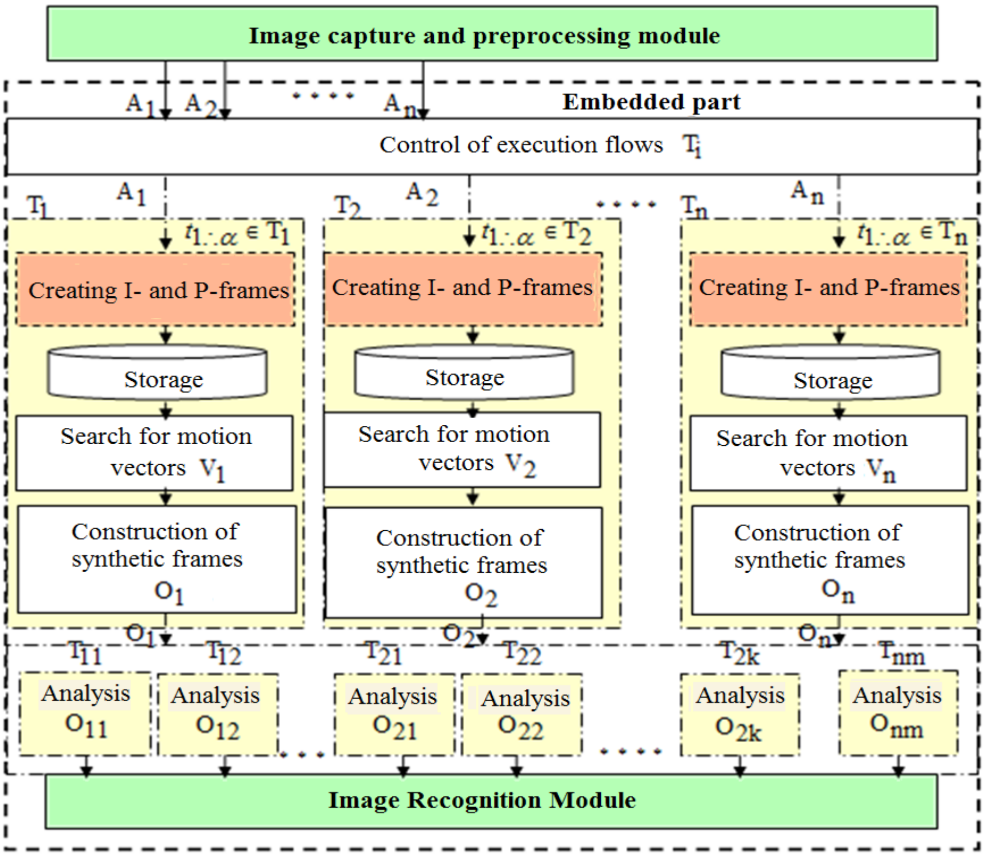

Figure 5. Multithreaded video data processing model

Parallelism is achieved at three stages of data processing. First, at the stage of receiving the video sequence. For each of the image sequences, it is assumed that a separate stream is allocated for independent computational operations. Secondly, at the stage of creating key and intermediate frames, if there is an inter-frame difference. Thirdly, the implementation of a parallel analysis of the generated synthetic frames and the output of the detection result.

The following input and output parameters are presented on the model of parallel processing of video data of the embedded subsystem:

$A_i, i \in 1 \ldots n$ - independent and parallel incoming sequence of images from the image capture and preprocessing module;

$T_i, i \in 1 \ldots n * m$ - the flow of performing image processing operations;

$V_1, i \in 1 \ldots n$ - list of motion vectors for A_i;

$O_i, i \in 1 \ldots n$ - list of synthetic frames for A_i.

The quantitative limitation of allocated streams is regulated by the operating system. If the limit is reached, the processing of operations scheduled within the new execution thread is carried out after one of the already running threads finishes its execution [11].

5. DECISION-MAKING SYSTEM

A subsystem of decision-making means such a system, the purpose of which is to help in decision-making in difficult conditions for a complete and objective analysis of the environment.

The block diagram of the decision-making subsystem is shown in Figure 6.

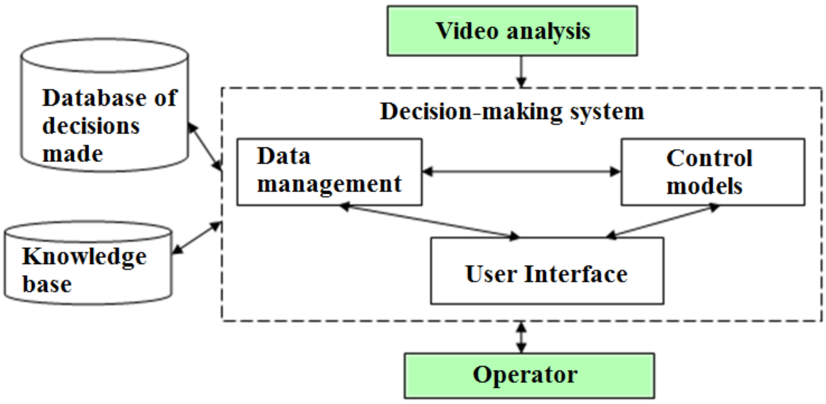

Figure 6. Block diagram of the decision support system

The input data are lists of recognition results coming from the object recognition module. The output data represents the decisions made about the need to perform any action, for example, writing to the log of logging situations or displaying an alarm notification on the screen form of the application, etc. [1].

The knowledge base contains a list of rules according to which data analysis and processing and decision-making are carried out.

The database of decisions made is also designed to support decision-making by establishing links with previous events and incidents, as well as with their conditions of occurrence.

Data management refers to the formation of new information based on available input parameters, knowledge bases and management models.

The component of the management model provides accounting and analysis of the current situation based on the available input parameters and knowledge bases and data of the decisions made.

The data management component provides the possibility of manipulating the accumulated information about incidents.

The decision-making system has interactive interaction with the system operator to ensure flexibility of configuration and decision-making.

6. CONCLUSION

1. The components of the video analytics system are considered, which are a set of video cameras, an image capture and transformation module, an object recognition module, databases and knowledge, a decision-making system, as well as an embedded part for effective detection of dynamic objects. The flow diagrams of data flows are given. The organization of the interconnection of software subsystems and the place of the embedded part in the structure of the video analytical system is shown.

2. A model of multithreaded processing of graphical data for the dynamic object detection module is demonstrated. The implementation of parallelism is carried out, firstly, at the level of receiving a video stream, providing independent processing of a single sequence of images, and secondly, at the stage of formation and transmission of areas of interest, allowing their simultaneous analysis within independent execution streams.

A more detailed structure of the embedded subsystem, as well as data flows inside the component are shown on the model of processing a single video data stream.

REFERENCES

[1] Pakhomova O.A., Kravets O.Ja. Control and preprocessing of graphic data for effective dynamic object recognition. *IOP Conference Series: Material Science and Engineering*, vol. 2019, no.537, art. No 052002. DOI: 10.1088/1757-899X/537/5/052002.

[2] Kravets O.Ja., Aksenov I.A., Rahman P.A. Algorithmization of analytical methods for finding motion vectors when processing image series. *International Journal on Information Technologies and Security*, vol. 14, no.2, 2022, pp. 39-50. http://ijits-bg.com/contents/IJITS-2022-No2/2022-N2-04.pdf.

[3] Kravets O.Ja., Aksenov I.A., Rahman P.A., Redkin Yu.V., Atlasov I.V., Zaslavskaya O.Yu. Algorithmization of image processing for identification of dynamic objects. *International Journal on Information Technologies and Security*, vol.14, no.3, 2022, pp. 47-58.

[4] Calderon F., Garnica-Carrillo A., Reyes-Zuñiga C. Binarization of images with variable lighting using adaptive windows. *Signal, Image and Video Processing*, vol. 2022, no.8, 2022. https://doi.org/10.1007/s11760-022-02150-1.

[5] Škrabánek P., Martinkova N. Tuning of grayscale computer vision systems. *Displays*, vol.74, no.2, 2022, art. No 102286. DOI: 10.1016/j.displa.2022.102286.

[6] Fedorov A., Nikolskaia K., Ivanov S., Minbaleev A. Traffic flow estimation with data from a video surveillance camera. *Journal of Big Data*, vol.6, no.1, 2019. DOI: 10.1186/s40537-019-0234-z.

[7] Ponlatha S., Sabeenian R.S. Comparison of Video Compression Standards. *International Journal of Computer and Electrical Engineering*, vol.5, no.6, 2013, pp. 549-554.

[8] Santosh Kc., Antani S.K. Recent trends in image processing and pattern recognition. *Multimedia Tools and Applications*, vol. 79, 2020, pp.47-48. DOI: 10.1007/s11042-020-10093-3.

[9] O.Ja. Kravets, I.V. Atlasov, I.A. Aksenov, A.S. Molchan, O.Yu. Frantsisko, P.A. Rahman. Increasing efficiency of routing in transient modes of computer network operation. *International Journal on Information Technologies and Security*, vol. 13, no. 2, 2021, pp. 3-14

[10] Kong L., Chen L., Wang M. An Automatic Method of Sequential Image Processing. *Journal of Physics: Conference Series*, vol. 1813, no.1, art. No 012034, 2021. DOI: 10.1088/1742-6596/1813/1/012034.

[11] Keith G. Operator Scheduling. *IIE Transactions*, vol.11, no.1, 1979, pp. 37-41.

Information about the authors:

Alexey Vladislavovich Gorshkov – Researcher, Research Institute of Computing Complexes named after M.A. Kartsev, Moscow, areas of scientific research – system analysis, optimization, simulation of complex objects

Oleg Jakovlevich Kravets – Dr. Sci (IT), professor of Voronezh state technical university, areas of scientific research – system analysis, optimization, simulation of complex objects

Ilia Antonovich Aksenov – PhD, associate professor of Vladimir State University named after Alexander and Nikolay Stoletovs, areas of scientific research – application of information technologies in economic systems

Yuriy Viktorovich Redkin, Ph. D., associate professor of Admiral Ushakov Maritime State University, Novorossiysk, areas of scientific research – control and processing of information, signal processing

Igor Victorovich Atlasov - Dr. Sci (Mathematics), professor of Moscow University of Ministry of Internal Affairs of Russian Federation named by V. Ja. Kikot, areas of scientific research – system analysis, optimization, simulation of complex objects

Manuscript received on 05 October 2022