РОЗДІЛ IV. ЕНЕРГЕТИКА, ЕЛЕКТРОТЕХНІКА ТА ЕЛЕКТРОМЕХАНІКА

DOI: 10.25140/2411-5363-2023-2(32)-300-312
UDC 681.5:004.51

Volodymyr Voytenko[1], Björn Olofsson[2], Maksym Solodchuk[3], Yuriy Denisov[4]

[1]PhD in Technical Sciences, Docent, associate professor of the Electronics, Automation, Robotics and Mechatronics Department; researcher of the Department of Automatic Control
Chernihiv Polytechnic National University (Chernihiv, Ukraine); Lund University (Lund, Sweden)
E-mail: v.voytenko@stu.cn.ua. **ORCID:** http://orcid.org/0000-0003-1490-0600
ResearcherID: F-8698-2014. **Scopus Author ID**: 36167678700

[2]PhD in Automatic Control, Docent, Senior Lecturer of the Department of Automatic Control
Lund University, (Lund, Sweden)
E-mail: bjorn.olofsson@control.lth.se. **ORCID:** http://orcid.org/0000-0002-6793-7207. **Scopus Author ID**: 54793307900

[3]Head of the Scientific-Research Department
State Scientific Research Institute of Armament and Military Equipment Testing and Certification (Cherkasy, Ukraine)
E-mail: rocket15733@gmail.com. **ORCID**: https://orcid.org/0000-0002-1162-6784

[4]Doctor of Technical Sciences, Professor, head of the Electronics, Automation, Robotics and Mechatronics Department
Chernihiv Polytechnic National University (Chernihiv, Ukraine)
E-mail: den71ltd@gmail.com. **ORCID**: https://orcid.org/0000-0003-2293-7964. **ResearcherID:** G-1144-2016

COMPONENTS OF A SYSTEM FOR AUTOMATIC DETECTION OF A ZONE OF INTEREST IN IMAGES OBTAINED FROM A UAV

To reduce the load on the UAV operator during long-term search and rescue missions, an on-board automatic system that generates control signals for positioning an additional video camera with a narrow field of view is considered. The requirements for the system of automatic detection of the area of interest are defined. Various methods for detecting objects in images are considered, analyzed and compared. Software and hardware tools are discussed, which are advisable to use in the preparation and conduct of experimental studies.

Keywords: *unmanned aerial vehicle (UAV); image analysis; man-machine system; electric drive; on-board object detector; spot camera control.*
Fig.: 5. Table: 1. References: 22.

Relevance of the research. When performing search and rescue missions, UAVs are a source of important information, most of which is generated by video cameras installed on board. The number of cameras and their spectral ranges of sensitivity are selected depending on the specific task to be solved. The processing of video information itself is a cyclic procedure [1], which consists in searching for objects of interest, their detection, recognition, determination of their characteristics, and preparation of a report.

The processing speed of video information received from a UAV is critically important, since its volume is large in the case of using high-resolution cameras and long flights. Acceleration of the process of analyzing this information is possible either by parallel operation of several operators or by using appropriate automation tools.

Problem statement. Despite significant advances in computerized pattern recognition systems, in some applications, the final decision about the category of the object to be detected, as well as the subsequent actions in the search and rescue mission will remain with an operator. To improve the reliability of the classification of objects, the operator must receive all necessary and sufficient information promptly. To do this, a video camera with a varifocal lens can be placed on the UAV, which allows zooming in on the image part in the area of interest. In this case, the operator must be distracted from the direct control of the UAV and use the control of the video camera, spending precious time on this, as well as on a possible return to the original image.

Analysis of recent research and publications. Another solution [2] relies on an additional camera with a fixed narrow viewing angle (spot camera), which allows a quick overview of the object of interest at a larger scale. However, if this object turned out to be out of the direction of the optical axis of the spot camera, a positioning procedure of that camera will still be required. To reduce the time spent by the operator to perform auxiliary actions, as well as to reduce fatigue, it has been proposed to install the spot camera on a platform (or on a gimbal [3]) that can provide rotation about two coordinate axes relative to the main (navigation) video camera [2]. The image frame from the main camera can be divided into rectangular sections (zones), while the size of each of them is determined by the area of observation of the spot camera [2, Fig. 3].

An automatic detection system, in case of detecting an object of interest based on the results of the frame analysis, determines the most probable area of the location of the object, and generates control signals for appropriate positioning of the spot camera. This ultimately allows an operator to concentrate on the classification of an object of interest, as well as on making a decision based on the results from the automated system.

Isolation of previously unexplored parts of the general problem. One important specification for image processing systems using UAVs (and the direction of their development) is to have a high degree of autonomy, i.e., a concentration of functions performed directly onboard the vehicle [4]. This reduces the information load on the operator, and also reduces the required communication load between the UAV and the central control system, which expands the flight range, increases noise immunity, and reduces the power consumption of on-board radio-transmitting equipment.

Research objectives. The purpose of this article is to select components for further experimental research. The main tasks to be solved are the specification of the requirements, development of a structure, clarification of the principles and algorithms of the on-board system for automatic detection of an area of interest and positioning of the spot camera based on the results of the analysis of the image from the main video camera of the UAV.

Requirements for the system of automatic detection of the area of interest. To estimate the requirements on the computational time of the system for automatic detection of the area of interest, we use a simplified geometric construction (Fig. 1). The main difference from [2] (Fig. 1) and the corresponding mathematical relations there is that the angle α between the normal to the ground and the direction of the camera is taken into account here.

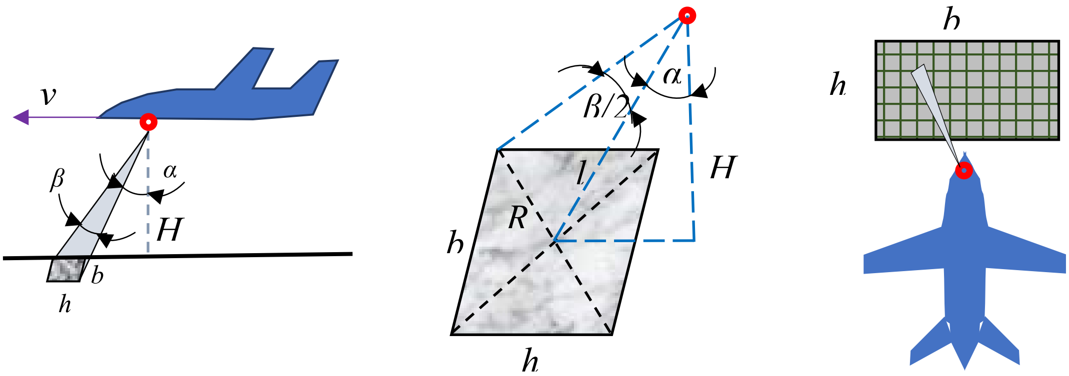

Fig. 1. Geometric parameters of the system

The UAV flies at an altitude H. In this case, the optical axis of the navigation camera lens is tilted at an angle α = 20…80°, which gives a visual three-dimensional representation of the object and the terrain [1]. Based on the definitions presented in Fig. 1, it is possible to find the size of the investigated surface in a simplified way (without considering projective distortions) depending on the camera view angle β and the UAV flight height H as follows:

$$R = l\tan\frac{\beta}{2} = H\frac{\tan\frac{\beta}{2}}{\cos\alpha}; \qquad R^2 = \frac{b^2 + h^2}{4}; \qquad b = K_f h$$

$$h = \frac{2R}{\sqrt{K_f^2 + 1}} = \frac{2H\tan\frac{\beta}{2}}{\cos\alpha\sqrt{K_f^2 + 1}}; \qquad b = \frac{2HKf\tan\frac{\beta}{2}}{\cos\alpha\sqrt{K_f^2 + 1}}$$

In the third expression, $K_f = b/h$ is the image format. On the other hand, the distance s that is covered by the vehicle which is flying at speed v during the time T_r is

$$s = v\ T_r.$$

Under the condition that $s = h$,

$$Tr = h / v = \frac{2H \ \tan\dfrac{\beta}{2}}{v \ \cos\alpha\sqrt{K_f^2 + 1}}. \tag{1}$$

Let us assume that the viewing angle of the camera is $\beta = 30°$. Moreover, if the UAV is at a height of $H = 100$ *m* with $\alpha = 58.3°$, the height of the image on the surface is $h = 50$ *m* (Fig. 1). Under the speed of $v = 72$ *km/h*, the UAV covers a distance of $s = 50$ *m*, i.e., the image is fully updated in $T_r = 2.5$ *s*.

A popular image format provided by video cameras installed on UAVs is HD $n_x \times n_y = 1920 \times 1080$ ($K_f = 16/9$). This implies that the size of one pixel for $H = 100$ *m* and $\beta = 30°$ is

$$\Delta = h/n_y = 5000/1080 \approx 4.63\ cm.$$

In this relation, n_x, n_y are the number of pixels along the horizontal and vertical direction, respectively.

Divide the entire image frame into rectangular areas (zones) with an aspect ratio corresponding to the frame format of the video camera. Considering the viewing angles of common video cameras for UAVs [5; 6], it is advisable to use the image scaling factor

$$M = 10 \ldots 32.$$

In normal navigation mode, the UAV operator uses video images with a standard frame rate of F_f. The recognition (detection) system has a time interval of T_r as margin, during which the UAV will cover a path equal to the height of the frame on the ground (h, Fig. 1). During this time, the system for automatic detection of the zone of interest must determine the numbers N_x and N_y along both coordinate axes on the captured frame and generate commands for positioning the spot camera. N_x and N_y belong to the range from 1 to M.

Table 1 shows the quantitative parameters of image zoning of the main UAV video camera operating in HD format. The flight altitude is $H = 100$ *m*, the angle of view is $\beta = 30°$, and the tilt angle of the main camera is $\alpha = 58.3°$.

Table – The number and format of split zones of the input image (the angle α is considered)

M	1	2	3	4	5	6	8	10	12	15	30
N_z	1	4	9	16	25	36	64	100	144	225	900
n_h	1920	960	640	480	384	320	240	192	160	128	64
n_v	1080	540	360	270	216	180	135	108	90	72	36
b, m	88.89	44.44	29.63	22.22	17.78	14.81	11.11	8.89	7.41	5.93	2.96
h, m	50	25	16.67	12.5	10	8.33	6.25	5	4.17	3.33	1.67
Δ, cm	4.63	2.31	1.54	1.16	0.93	0.77	0.58	0.46	0.39	0.31	0.15

The following designations are used in Table:

M – number of image zones along one of the axes (and simultaneously the scaling factor);

N_z – total number of image zones;

n_h – number of pixels in one zone of the image horizontally;

n_v –number of pixels in one zone of the image vertically;

h, b – size of the zone on the ground;

Δ – size of the pixel on the ground (or ground sample distance, GSD).

Fig. 2 makes it possible to estimate the time requirements for the system of automatic determination of the area of interest depending on the tilt angle of the main camera α.

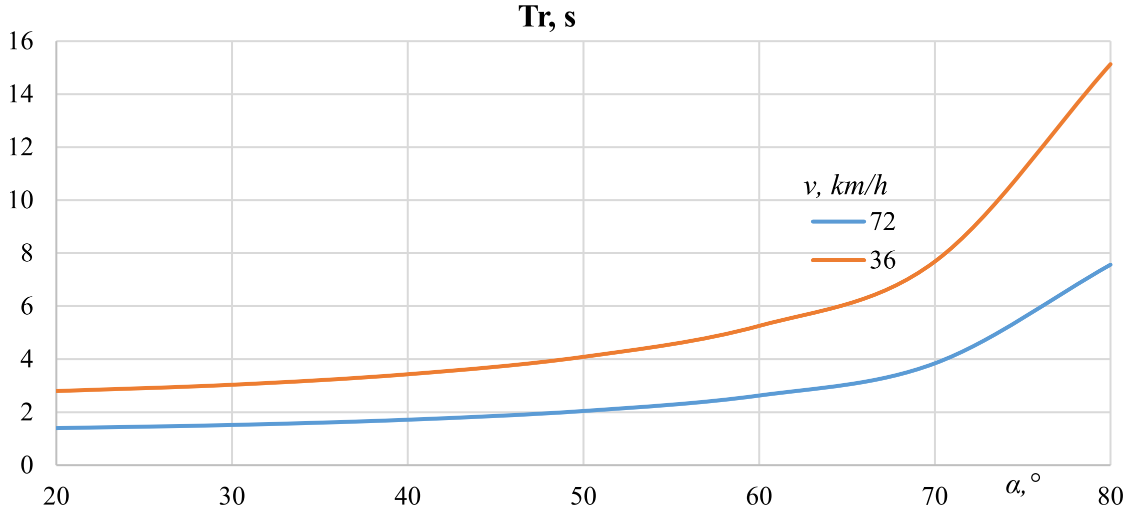

Fig. 2. Dependence of the allowable image processing time on the tilt angle of the UAV main camera (H = 100 m and β = 30°)

In Fig. 3, a three-dimensional plot of the dependence of the image update time with respect to altitude and angle at a fixed UAV flight speed is plotted.

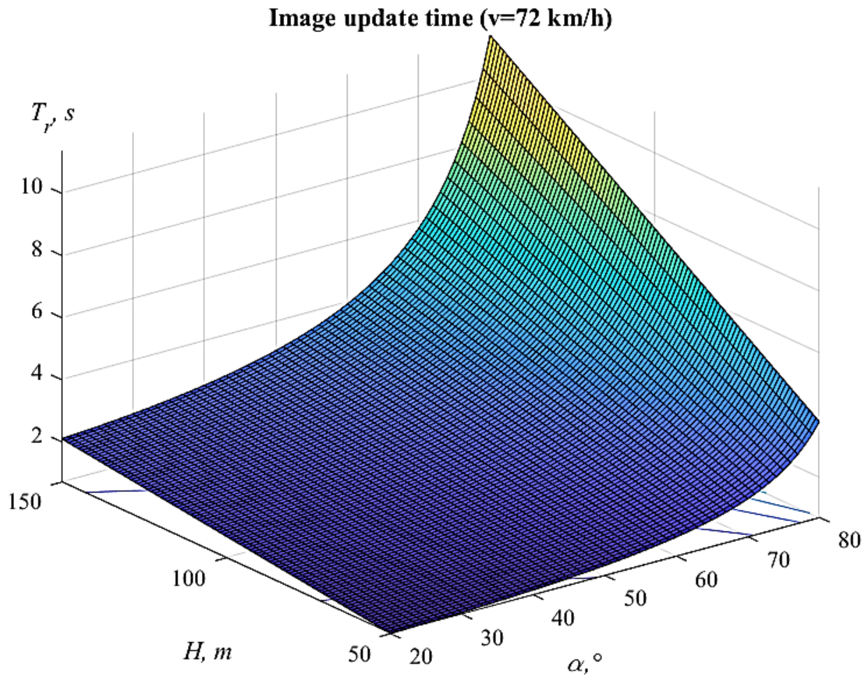

Fig. 3. The dependence of the allowable image processing time on the angle of inclination of the main camera and the flight altitude of the UAV (v = 72 km/h and β = 30°)

From Fig. 2 and Fig. 3, it follows that for realistic ranges of the geometric parameters, the allowable response time of the detection system is in the range of a few seconds. If it is possible to solve the problem of real-time detection of the zone of interest offline (on-board the UAV), this will also help to speed up the analysis of the entire accumulated array of video information by studying it after the mission is completed, which often takes many times more time than it took to complete it [1].

It could be noted that during the flight mission, the UAV commonly maintains the speed, the flight altitude, and the tilt angle of the main navigation camera at a constant level. In addition, changes in these parameters within a certain range can be considered as a result of the information acquired from the on-board sensors. As for the other requirements of the system for automatically determining the area of interest (e.g., energy consumption, dimensions, weight), despite their importance, it makes sense to clarify them in the process of solving the problem of timely detection of the area of interest. In addition, the video recording conditions (light, parallax, etc.) must also be considered at this stage.

Selection of algorithm for the system for automatic detection of the zone of interest. In fact, at this stage, the task is to determine an architecture and algorithm of the system operation that will ensure its deployment and determination of the positioning zone of a narrow-angle video camera in real-time during the UAV mission, while still causing as small computational load as possible. We are consequently faced with the task of detecting patterns in the image.

Selected parts of an image can be described using so-called descriptors. For regions, some of the simplest are [7]: the area which is defined as the number of pixels it contains; the perimeter; the compactness (the ratio of the square of the perimeter to the area); the roundness factor, which is defined as the ratio of the area of a region to the area of a circle (the most compact figure) with the same perimeter. For a global description of regions on an image plane, such topological properties as the number of holes inside the region, the number of connected components of the region, and the Euler number can be useful [7]. One important approach to describing regions is the quantitative representation of their textural features. This descriptor is a measure of local properties such as smoothness, coarseness, and regularity [7].

For UAVs, and small robots in general, it is important to reduce the amount of image information without losing key features. These features are usually scalars (e.g., area or aspect ratio) or low-dimensional vectors (e.g., object coordinates or line parameters). Image feature extraction is the stage of the information concentration, where the data transfer rate can be reduced from 10^6–10^8 *bytes/s* at the camera output to the order of dozens of features per frame [8].

Machine learning technologies have been developing for a long time and today they demonstrate the best results in solving pattern recognition problems [8-12]. The greatest advances in this area have been made with deep neural networks (DNN). Nevertheless, the very fact of the emergence of more and more new varieties of neural networks (NNs) and learning algorithms indicates their potential for further development and the presence of many unsolved problems. The most obvious problem is the computational requirements of both the DNN training process and the process of generating output from the network when the model it is deployed on an embedded device. This resource intensity can be quantified in terms of processor performance, memory usage, cost, etc.

On the other hand, the UAV has significant limitations in terms of dimensions, weight, and power consumption of devices and components that can be used on-board. These restrictions are not always compatible with the needs of real-time detection/recognition algorithms. Significant progress in reducing resource intensity can be achieved if the network learning process is implemented on another computational unit than the target processor, where the latter is placed and used on-board the UAV, i.e., on a powerful computer with a sufficiently long operation time, which is accompanied by a comparably high-power consumption. The option of generating output on a cloud device is not always feasible, and therefore the solution to the problem (smoothing the contradiction between the required and available resources) is to simplify the NN structure as much as possible while maintaining the accuracy of the output.

Given the integrated approach to solving the problem and the limited resources of the UAV, we consider only implementation of the object detection in the image from a wide-angle camera, entrusting the most important tasks of verification and recognition to a trained operator. It is shown in [9] that at present, issues related to the development of a detection and tracking system based on a neural network operating in real-time on embedded devices can be solved.

The choice of an on-board object detector. Returning to the main problem to be solved, namely, reducing the fatigue of the UAV operator, we recall that the image received from the main (navigation) video camera is divided into uniform rectangular areas (zones), and the software system for preprocessing and image analysis tries to find the zone number in which the probability of the presence of the object of interest is the highest. This makes it possible to generate electrical signals for positioning an additional (spot) camera with a narrow viewing angle to obtain an enlarged image fragment for its classification by a human operator.

To solve the problem of automatically finding the area of interest in the image from the navigation camera, it is necessary to use an object detector, the key requirement for which is the ability to work online on-board the UAV.

Object detectors which are known today can be divided into two categories [9, 10]: directly based on purely data-driven machine learning algorithms, as well as based on dedicated features computed based on the input data. The first group of object detectors basically uses both traditional neural networks and machine learning algorithms, as well as Convolutional Neural Networks (CNNs) and deep learning algorithms.

To get started with object detection using deep learning, one of the well-known approaches can be applied [10]:

1. Creating a new object detector. It allows very high-performing models in the end, but it requires significant time, computational resources, and a large amount of data to configure layers and determine the weight coefficients of a deep neural network.

2. Using a pre-trained object detector. This allows quick results, but there is a potential problem with the adequacy of the training set of images that was used when training the neural network for a specific task.

In addition, when choosing an on-board object detector, the type of convolutional neural network used should also be taken into account. In a two-stage detector, the first stage based on a Region Based Convolutional Neural Network (R-CNN) or its variants [11], is designed to determine image areas that may contain an object. The second stage classifies objects within this area. Such a detector allows for very accurate detection of an object but is usually slower than a single-stage detector [10].

An example of a single-cascade detector is YOLO [11]. Here, a single convolutional neural network analyzes the entire image and predicts the probabilities of objects of given categories within bounding boxes. Such a detector could perform faster than a two-stage detector, but at the expense of increased number of misclassifications, especially in scenes with small objects.

Some algorithms do not rely on convolutional neural networks, but still are considered machine learning algorithms. Some of these machine learning techniques are also used to detect objects in images. These are, for example:

– Aggregate Channel Features (ACF [10]). The ACF method extracts properties directly as pixel values in extended image channels without calculating rectangular sums at different locations and scales;

– classification by the method of support vector machine (SVM) using features of histograms of oriented gradient (Histograms of Oriented Gradient) [10];

– the Viola-Jones algorithm [10] for detecting the human face and upper body.

The presence of such a variety of the above and many other object detection methods is a clear indication of their non-universality. That is, each developer tries to use the method that will give the best result under the specific requirements on the specific task. Here, some general considerations can be made for applying such methods. Methods that use convolutional neural networks and deep learning allow building a high-quality object detector when a powerful GPU and a sufficient number of labeled training images are used during training. Under such conditions and if sufficient training time is available, this approach can generate the desired results. Otherwise, it can be preferable to rely on other types of machine learning algorithms. In some specific cases, when certain objects are known in advance and qualitatively displayed, it may be sufficient to completely abandon the use of resource-intensive machine learning algorithms, and instead use [10]: image segmentation and analysis of large binary fragments (blob analysis), which rely on such properties of objects as size, shape, or color; feature-based object detection, where feature extraction, pattern matching, and RANSAC are used to estimate the location of an object.

Examining the field of application of the object detector, the following can be noted.

1. The UAV has significant limitations regarding the mass, dimensions, and power consumption of any electronic devices placed on-board.

2. Intensive radio communication during the mission could be challenging. This makes cloud computing difficult to apply in this context.

The contradiction between the existing limitations and the need to perform most of the video signal processing operations on-board the UAV (which is the development trend of these devices) can be solved by combining the appropriate electronic components (multifunctional built-in productive processors) and fast object detection algorithms. From this point of view, the use of combined methods can be considered promising, e.g., that described in [12], which uses both the properties of possible objects, which are determined by textural characteristics inside the image, and a classifier configured using a machine learning algorithm.

Datasets for airborne detection of objects in UAV images. Given that universal and most successful modern object detectors are based on deep neural networks, the issue of having a suitable database for use in machine learning algorithms is very relevant. Although there are many open image data sets, we must consider the specific requirements because we need a database of images obtained from UAVs flying at known ranges of altitudes, speeds, pitch angles, and other parameters. Let us try to analyze what is most suitable for the considered case.

Popular datasets (DOTA [13], AID [14], iSAID [15], xView [16]), some of which are called "aerial", are actually created using satellites. The data set UAVVaste [17] is intended for a very specific area of use. The data set DroneDeploy [18] includes some of aerial photographs taken from a UAV. Some authors use this data set to test their methods for image processing, object detection, and tracking. Each scene in [19] has a ground sample distance (GSD) of 0.1 *m*. There is a corresponding "height" and "label" for each image. However, the database is not available for free in its entirety.

In [20], an overview of some datasets for object detection and tracking tasks in a convenient tabular form to display information is provided for comparison. Some well-known datasets are classified by the authors as a special drone-based class, and they note that these data sets have limitations in terms of scenarios. The work with the VisDrone-Dataset, which contains more than 10,000 images of the urban and rural environment of China, as well as objects from a wide range of angles, is also described in [20].

Consider the conditions when a UAV performs a search and rescue mission. Let the flight altitude $H = 100$ m; viewing angle $\beta = 30°$; tilt angle of the main video camera $\alpha = 58.3°$. Taking into account [2, Fig. 1] it is possible to obtain the main geometric parameters of the image zoning of the main UAV video camera operating in HD format:

1) the scale factor $M = 1$;

2) the number of pixels in one zone of the image $N_h = 1920$; $N_v = 1080$;

3) the size of the zone on the ground $b = 88{,}89$ *m;* $h = 50$ *m;*

4) the size of the pixel on the ground or ground sample distance (GSD) $\Delta = 4{,}63$ *cm*.

A comparative analysis of [13-20] shows that none of the data sets is created with all the input parameters in the range of those above, and none of these data sets has a sufficient number of images for the specific task. The solution can be found in two ways:

1) Using a real UAV to create new data sets. This obvious path is long and expensive, but clearly has the potential to provide good results.

2) Modeling images using computer synthesis and a monitor or projector placed at an appropriate distance.

For evaluation, we will use a regular Logitech C920 HD Pro Webcam and place it in front of the simulated image screen. The dependence of the calculated distance from the camera to the screen is shown in Fig. 4. The angle $\alpha = 0$ corresponds to the straight direction to the screen.

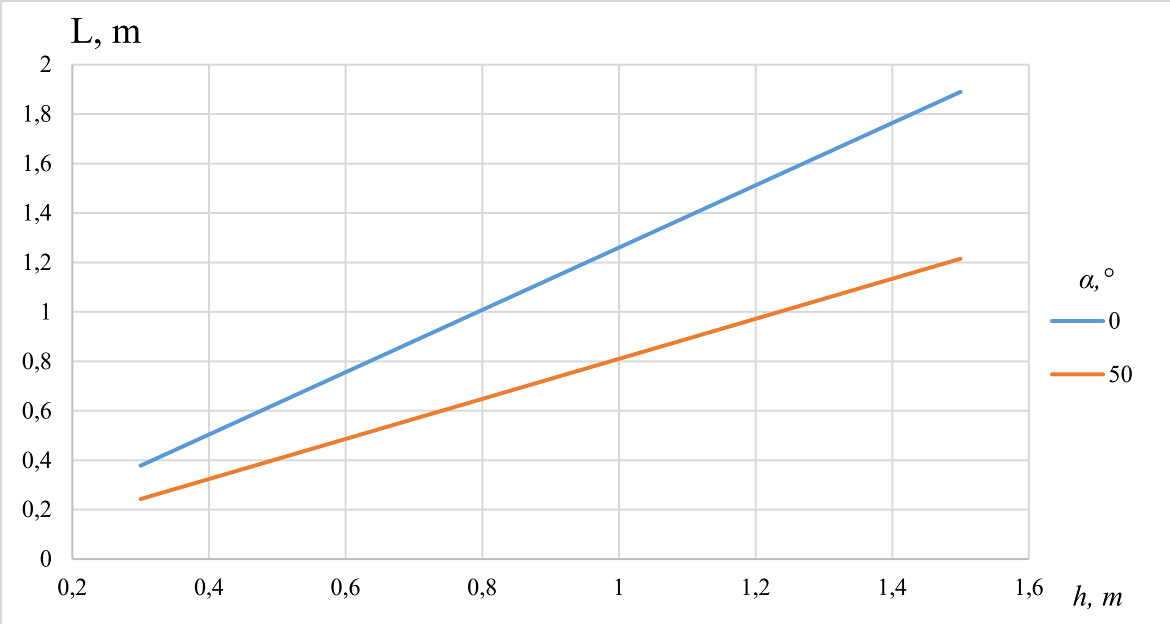

Fig. 4. Distance between camera and screen

Hardware support of the on-board system for increasing the speed of processing images received from UAVs. As discussed in [2], to solve problems associated with fatigue of the UAV operator during long missions, either the main (navigation) video camera with a varifocal lens can be used, or a spot camera with a narrow angle of view can be added. The image from the navigation video camera should be analyzed in real-time to the presence of objects of interest and, if such are detected, the required image part is automatically zoomed in on.

Using a single camera for both navigation purposes and final classification of potential candidates for the object of interest has several drawbacks.

1. The problem of documenting the mission. The availability of information from only one image (either nominal or enlarged) does not allow to conclude that no object of interest was missed, after the completion of the mission and reading the built-in storage device of the UAV. This might lead to the need to perform a repeated mission, which increases operator fatigue.

2. Purely technical problems associated with the use of a zoom lens. After all, such a lens, in addition to a complex mechatronic system of interconnected movement of several lenses, must have additional focusing subsystems implemented using separate electric drives. This leads to a deterioration in image quality (e.g., accuracy, brightness, and stability of characteristics), as well as to an increase in object classification time because of limited autofocus dynamics.

Hence, there is an advantage of using a separate spot camera located on a gimbal, which can be positioned independently of the navigation video camera for yaw and pitch angles. The structure of the electromechanical part of the overall system for increasing the speed of processing images obtained from UAVs (one of the two channels) is shown in Fig. 5. This is an extended and more detailed block diagram compared to the one in [2].

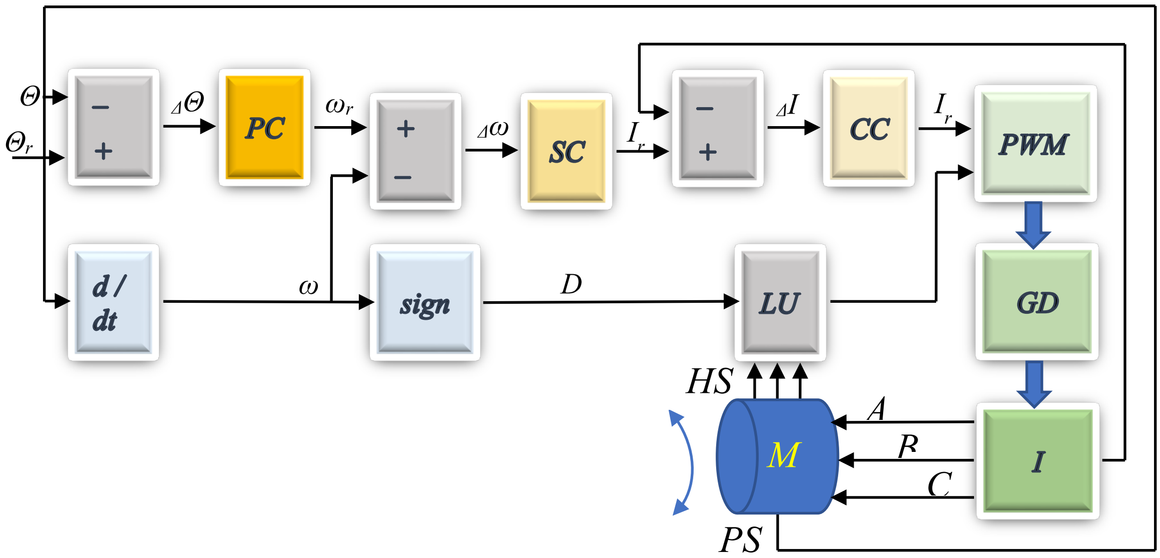

Fig. 5. Spot camera control structure (one channel)

The brushless direct-current motor *M* is the final actuating device, which, together with another similar one, can be designed for the interrelated positioning of the spot camera in one of two perpendicular directions within the image from the navigation video camera. The motors are equipped with Hall sensors (*HS*) and a position sensor (*PS*), the signals of which are used in a three-loop automatic control system that acts on the control error between reference values and actual measured values.

The index "*r*" in Fig. 5 is used for reference values. The zone number determined by the on-board object detector is recalculated into the specified angle of orientation of the spot camera Θ_r in a certain direction. Position (PC), speed (SC), and current (CC) controllers provide the required quality of the transition process of the motor rotation angle. The rotation speed is calculated by differentiating the signal from the *PS*. The *sign* block uses speed to determine the direction of rotation of the motor, which, with the help of a logic unit (*LU*), provides the appropriate alternation of *HD* signals for the pulse width modulator (*PWM*). Pulses with width modulation are supplied to the gate drivers (*GD*) of the power switches of the inverter (*I*), which ensures the required form of motor winding currents *A*, *B*, *C*.

As mentioned, a feature of UAVs is the presence of significant design and energy constraints, which fundamentally affect the choice and implementation of all components of the structure in Fig. 5. Provided that the specifications in terms of minimum dimensions, mass, and power consumption are ensured, high-performance computers must be on-board to perform both image pre-processing procedures (correction of visibility conditions, angle of view, removal of obstacles, noise, etc.) and object detection [12]. In addition to solving these resource-intensive tasks related to the implementation of the area of interest detector, the on-board computer must also control two interconnected electric drives in real-time (Fig. 5). Therefore, the selection of components for an electronic system that will provide hardware and software support for the implementation of the specified resource-intensive algorithms is a non-trivial task.

The development of electronics is accompanied by the spread of machine learning algorithms to front-end devices. An effective solution is the use of a multi-functional system on a crystal with built-in support for digital image processing procedures and the implementation of neural networks. The i.MX 8M Plus processor family [21] is designed to reliably solve tasks in machine learning and vision, multimedia, and industrial automation. Key features of this processor include: four Arm® Cortex®-A53 cores and neural processor up to 2.3 TFLOPS; dual video signal processor and two camera inputs for the video system; video codec, 3D/2D graphics accelerator, numerous audio and voice functions; real-time control using Cortex-M7.

Thus, this very large integrated circuit (VLSIC) at the hardware and the corresponding software levels support both the complex tasks related to intelligent image processing and many of the blocks in Fig. 5. It is advisable to use ready-made hardware and software solutions for conducting experimental studies at the stage of creating a prototype to evaluate the conceptual foundations. Consider the PhyBOARD Pollux AI Kit platform [22]. This kit contains a single-board computer based on the i.MX 8M Plus processor and a MIPI camera. The pre-installed software significantly speeds up the development: the OpenCV library; the GStreamer framework (video editors, streaming servers, media players and file converters, VoIP solutions). The Yocto Linux Board Support Package includes [22] NXP's eIQ machine learning software development environment. Support for pytorch, TensorFlow Lite, and the ONNX format allows implementation of machine learning algorithms.

Conclusions. To solve the problem of increasing the speed and reducing the fatigue of the UAV operator when performing long-term search and rescue missions, the choice of components of the software and hardware is discussed, which automatically on-board ensures the orientation of the spot camera in the desired direction.

The allowable computational time spent by the system for automatically determining the area of interest is estimated depending on the tilt angle of the main camera and the flight altitude of the UAV. It is shown that to solve the formulated problem, it is required to create an on-board object detector with a time of detecting the area of interest in the order of a few seconds. For this purpose, combined methods can be used that employ both the features of objects defined by textural characteristics of the image and a classifier, which is tuned using a machine learning algorithm.

Data sets for airborne object detection on UAV images are characterized and it is shown that there are limitations in terms of meeting the required input parameters and the required number of relevant images. Possible ways of solving this problem are discussed.

For experimental studies, software and hardware solutions are discussed that can be used to solve the whole range of tasks, including image processing and control of electric motors for positioning the spot camera, directly on-board the UAV under the specific operating conditions.

Acknowledgements. The authors are highly grateful to the Knut and Alice Wallenberg Foundation for providing a grant.

References

1. Sun, J., Li, B., Jiang, Y., Wen, C. (2016). Camera-based target detection and positioning UAV system for search and rescue (SAR) purposes. *Sensors, 16*(11), 1778-1802. https://doi.org/10.3390/s16111778.

2. Voytenko, V., & Solodchuk, M. (2021). Pidvyshchennia shvydkosti analizu zobrazhen, otrymanykh iz bezpilotnoho litalnoho aparatu Pidvyshchennia [Increasing the speed of analysis of images obtained from unmanned aerial vehicle]. *Tekhnichni nauky ta tekhnolohii – Technical sciences and technologies*, ((2(28)), 127-137. https://doi.org/10.25140/2411-5363-2022-2(28)-127-137.

3. Beard, R., McLain, T. (2012). *Small Unmanned Aircraft: Theory and Practice*. Princeton University Press.

4. Bashynskyi, V.G., Ragulin, V.V., Solodchuk, M.O., Fomin, A.V., & Isachenko, O.O. (2022). Obgruntuvannia neobkhidnosti obrobky videoinformatsii na bortu rozviduvalnoho BPLA [Justification of the need to process video information on board the reconnaissance UAV]. *Zbirnyk naukovykh prats Derzhavnoho naukovo-doslidnoho instytutu vyprobuvan i sertyfikatsii ozbroiennia ta viiskovoi tekhniky – Scientific works of the State Scientific Research Institute of Armament and Military Equipment Testing and Certification*, *11*(1), 105-115. https://doi.org/10.37701/dndivsovt.11.2022.12.

5. WESCAM MX-10 EO/IR. (n.d.). L3Harris. https://www.l3harris.com/all-capabilities/wescam-mx-10-air-surveillance-and-reconnaissance.

6. FCB-9500 Series – New Colour Camera Block. Sony. https://www.image-sensing-solutions.eu/FCB-New-generation.html.

7. Gonzalez, R.C., Woods, R.E. (2007). *Digital Image Processing*. 3rd ed. Pearson.

8. Corke, P. (2017). *Robotics, Vision and Control. Fundamental Algorithms in MATLAB*. Springer International Publishing AG.

9. Nousi, P., Mademlis, I., Karakostas, I., Tefas, A., & Pitas, I. (2019). Embedded UAV Real-Time Visual Object Detection and Tracking. *2019 IEEE International Conference on Real-time Computing and Robotics (RCAR)* (pp. 708-713). https://ieeexplore.ieee.org/document/9043931.

10. What Is Object Detection? 3 things you need to know. (n.d.). MathWorks. https://www.mathworks.com/discovery/object-detection.html.

11. Gandhi, R. (2018, July 9). R-CNN, Fast R-CNN, Faster R-CNN, YOLO – Object Detection Algorithms. *Towards – Data Science*. https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e.

12. Avola, D., Cinque, L., Mambro, Di, A., Diko, A., Fagioli, A., Foresti, G.L., Marini, M.R., Mecca, A., & Pannone, D. (2022). Low-Altitude Aerial Video Surveillance via One-Class SVM Anomaly Detection from Textural Features in UAV Images. *Information*, *13*(1(2)). https://www.mdpi.com/2078-2489/13/1/2.

13. Ding, J., Xue, N., Xia, G.-S., Bai, X., Yang, W., Yang, M.Y., Belongie, S., Luo, J., Datcu, M., Pelillo, M., Zhang, L. (2021, December 4). Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges. https://arxiv.org/abs/2102.12219v2.

14. Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., Zhang, L. (2016, August 18). AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification. https://arxiv.org/abs/1608.05167.

15. Zamir, S.W., Arora, A., Gupta, A., Khan, S., Sun, G., Khan, F.S., Zhu, F., Shao, L., Xia, G.-S., Bai, X. (2019, August 28). iSAID: A Large-scale Dataset for Instance Segmentation in Aerial Images. https://arxiv.org/abs/1905.12886v2.

16. Lam, D., Kuzma, R., McGee, K., Dooley, S., Laielli, M., Klaric, M., Bulatov, Y., & McCord, B. (2018, February 22). xView: Objects in Context in Overhead Imagery. https://arxiv.org/abs/1802.07856v1.

17. Kraft, M., Piechocki, M., Ptak, B., & Walas, K. (2021, March 4). Autonomous, Onboard Vision-Based Trash and Litter Detection in Low Altitude Aerial Images Collected by an Unmanned Aerial Vehicle. *Remote Sensing*, *13*(5(965)). https://doi.org/10.3390/rs13050965.

18. Full Reality Capture. Interior and exterior visual data – any altitude, any angle, all in one platform. (n.d.). *Drone Deploy*. https://www.dronedeploy.com.

19. Heffels, M. R., Vanschoren, J. (2020, December 3). *Aerial Imagery Pixel-level Segmentation*. https://arxiv.org/abs/2012.02024.

20. Zhu, P., Wen, L., Du, D., Bian, X., Fan, H., Hu, Q., & Ling, H. (2021). Detection and Tracking Meet Drones Challenge. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* (44(11)), 7380-7399. https://ieeexplore.ieee.org/document/9573394.

21. i.MX 8M Plus – Arm® Cortex®-A53, Machine Learning, Vision, Multimedia and Industrial IoT. (n.d.). *NXP*. https://www.nxp.com/products/processors-and-microcontrollers/arm-processors/i-mx-applications-processors/i-mx-8-applications-processors/i-mx-8m-plus-arm-cortex-a53-machine-learning-vision-multimedia-and-industrial-iot:IMX8MPLUS?tid=vanIMX8MPLUS.

22. phyBOARD®-Pollux AI kit. (n.d.). *NXP*. https://www.phytec.eu/en/produkte/development-kits/phyboard-pollux-ki-kit/?lang=en/.

Список використаних джерел

1. Sun J. Camera-based target detection and positioning UAV system for search and rescue (SAR) purposes / J. Sun, B. Li, Y. Jiang, C. Wen // Sensors. – 2016. – Vol. 16(11). – Pp. 1778-1802. DOI: https://doi.org/10.3390/s16111778.

2. Войтенко В. Підвищення швидкості аналізу зображень, отриманих із безпілотного літального апарату / В. Войтенко, М. Солодчук // Технічні науки та технології. – 2021. – № (2(28)). – Pp. 127-137. – Access mode : https://doi.org/10.25140/2411-5363-2022-2(28)-127-137.

3. Beard R. Small Unmanned Aircraft: Theory and Practice / R. Beard, T. McLain. – Princeton University Press, 2012.

4. Обґрунтування необхідності обробки відеоінформації на борту розвідувального БПЛА / В. Башинський, В. Рагулін, М. Солодчук, А. Фомін, О. Ісаченко // Збірник наукових праць Державного науково-дослідного інституту випробувань і сертифікації озброєння та військової техніки. – 2022. – № 11(1). С. 105-115. doi:10.37701/dndivsovt.11.2022.12.

5. WESCAM MX-10 EO/IR. (n.d.). L3Harris. https://www.l3harris.com/all-capabilities/wescam-mx-10-air-surveillance-and-reconnaissance.

6. FCB-9500 Series – New Colour Camera Block. Sony. https://www.image-sensing-solutions.eu/FCB-New-generation.html.

7. Gonzalez R. C. Digital Image Processing / R. C. Gonzalez, R. E. Woods. – 3rd ed. – Pearson. 2007.

8. Corke P. Robotics, Vision and Control. Fundamental Algorithms in MATLAB / P. Corke // Springer International Publishing AG. – 2017.

9. Embedded UAV Real-Time Visual Object Detection and Tracking [Electronic resource] / P. Nousi, I. Mademlis, I. Karakostas, A. Tefas, I. Pitas // 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR). – 2019. – Pp. 708-713. – Access mode: https://ieeexplore.ieee.org/document/9043931.

10. What Is Object Detection? 3 things you need to know. (n.d.). MathWorks [Electronic resource]. – Access mode: https://www.mathworks.com/discovery/object-detection.html.

11. Gandhi R. R-CNN, Fast R-CNN, Faster R-CNN, YOLO – Object Detection Algorithms [Electronic resource] / R. Gandhi // Towards – Data Science. (2018, July 9). – Access mode: https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e.

12. Low-Altitude Aerial Video Surveillance via One-Class SVM Anomaly Detection from Textural Features in UAV Images [Electronic resource] / D. Avola, L. Cinque, A. Di Mambro, A. Diko, A. Fagioli, G. L. Foresti, M. R. Marini, A. Mecca, D. Pannone // Information. – 2022. – № 13(1(2)). – Access mode: https://www.mdpi.com/2078-2489/13/1/2.

13. Object Detection in Aerial Images: A Large-Scale Benchmark and Challenges [Electronic resource] / J. Ding, N. Xue, G.-S. Xia, X. Bai, W. Yang, M. Y. Yang, S. Belongie, J. Luo, M. Datcu, M. Pelillo, L. Zhang; Cornell university (2021, December 4). – Access mode: https://arxiv.org/abs/2102.12219v2.

14. AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification [Electronic resource] / G.-S. Xia, J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, L. Zhang ; Cornell university (2016, August 18). – Access mode: https://arxiv.org/abs/1608.0516.

15. iSAID: A Large-scale Dataset for Instance Segmentation in Aerial Images [Electronic resource] / S. W. Zamir, A. Arora, A. Gupta, S. Khan, G. Sun, F. S. Khan, F. Zhu, L. Shao, G.-S. Xia, X. Bai; Cornell university (2019, August 28). – Access mode: https://arxiv.org/abs/1905.12886v2.

16. xView: Objects in Context in Overhead Imagery [Electronic resource] / D. Lam, R. Kuzma, K. McGee, S. Dooley, M. Laielli, M. Klaric, Y. Bulatov, B. McCord; Cornell university (2018, February 22). – Access mode: https://arxiv.org/abs/1802.07856v1.

17. Autonomous, Onboard Vision-Based Trash and Litter Detection in Low Altitude Aerial Images Collected by an Unmanned Aerial Vehicle [Electronic resource] / M. Kraft, M. Piechocki, B. Ptak, K. Walas // Remote Sensing. (2021, March 4).*13*(5(965)). – Access mode: https://doi.org/10.3390/rs13050965.

18. Full Reality Capture. Interior and exterior visual data – any altitude, any angle, all in one platform. (n.d.) [Electronic resource] // Drone Deploy. – Access mode: https://www.dronedeploy.com.

19. Heffels M. R. Aerial Imagery Pixel-level Segmentation [Electronic resource] / M. R. Heffels, J. Vanschoren; Cornell university (2020, December 3). – Access mode: https://arxiv.org/abs/2012.02024.

20. Detection and Tracking Meet Drones Challenge [Electronic resource] / P. Zhu, L. Wen, D. Du, X. Bian, H. Fan, Q. Hu, H. Ling // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – № 44 (11). – P. 7380-7399. – Access mode: https://ieeexplore.ieee.org/document/9573394.

21. i.MX 8M Plus – Arm® Cortex®-A53, Machine Learning, Vision, Multimedia and Industrial IoT. (n.d.) [Electronic resource] // NXP. – Access mode: https://www.nxp.com/products/processors-and-microcontrollers/arm-processors/i-mx-applications-processors/i-mx-8-applications-processors/i-mx-8m-plus-arm-cortex-a53-machine-learning-vision-multimedia-and-industrial-iot:IMX8MPLUS?tid=vanIMX8MPLUS.

22. phyBOARD®-Pollux AI kit. (n.d.) [Electronic resource] // NXP. – Access mode: https://www.phytec.eu/en/produkte/development-kits/phyboard-pollux-ki-kit/?lang=en.

Отримано 06.04.2023

УДК 681.5:004.51

Володимир Войтенко[1], Бьорн Олофсон[2], Максим Солодчук[3], Юрій Денисов[4]

[1]кандидат технічних наук, доцент, доцент кафедри електроніки, робототехніки, автоматики та мехатроніки; дослідник кафедри автоматичного керування
Національний університет «Чернігівська політехніка» (Чернігів, Україна); Лундський університет (Лунд, Швеція)
E-mail: v.voytenko@stu.cn.ua. **ORCID:** http://orcid.org/0000-0003-1490-0600
ResearcherID: F-8698-2014. **Scopus Author ID**: 36167678700

[2]доктор філософії з автоматичного керування, доцент, старший викладач кафедри автоматичного керування
Лундський університет (Лунд, Швеція)
E-mail: bjorn.olofsson@control.lth.se. **ORCID:** http://orcid.org/0000-0002-6793-7207. **Scopus Author ID**: 54793307900

[3]начальник науково-дослідного відділу
Державний науково-дослідний інститут випробувань і сертифікації
озброєння та військової техніки (Черкаси, Україна)
E-mail: rocket15733@gmail.com. **ORCID**: https://orcid.org/0000-0002-1162-6784

[4]доктор технічних наук, професор, завідувач кафедри електроніки, робототехніки, автоматики та мехатроніки
Національний університет «Чернігівська політехніка» (Чернігів, Україна)
E-mail: den71ltd@gmail.com. **ORCID**: https://orcid.org/0000-0003-2293-7964. **ResearcherID:** G-1144-2016

КОМПОНЕНТИ СИСТЕМИ АВТОМАТИЧНОГО ВИЯВЛЕННЯ ЗОНИ ІНТЕРЕСУ НА ЗОБРАЖЕННЯХ, ОТРИМАНИХ З БПЛА

Швидкість обробки відеоінформації, отриманої з БПЛА під час пошуково-рятувальних, моніторингових і розвідувальних місій, є параметром, який визначає успішність виконання завдання. Щоб зменшити навантаження на оператора БПЛА, безпосередньо на борту здійснюється автоматичне виявлення того фрагмента вихідного зображення поверхні, на якому може бути розташований об'єкт інтересу. Результатом роботи такої автоматичної системи є формування номера зони інтересу, генерація керуючих сигналів і відповідне позиціонування додаткової відеокамери з вузьким полем огляду (спот-камери). Остаточне рішення про виявлення об'єкта інтересу та його класифікацію здійснює оператор на підставі відеозображення з спот-камери.

Для вирішення зазначеної проблеми у статті обґрунтовано вибір компонентів програмно-апаратного комплексу, який безпосередньо на борту автоматично забезпечує орієнтацію спот-камери у потрібному напрямку.

Оцінено допустимі витрати часу для системи автоматичного визначення зони інтересу в залежності від кута нахилу основної камери та висоти польоту БПЛА. Показано, що для вирішення основної задачі потрібне створення бортового детектора об'єктів з часом виявлення зони інтересу порядку одиниць секунд. Для цього можуть бути застосовані комбіновані методи, які використовують як властивості об'єктів, що визначаються текстурними характеристиками всередині зображення, так і класифікатор, налаштований за допомогою алгоритму машинного навчання.

Охарактеризовано набори даних для бортового виявлення об'єктів на зображеннях БПЛА, і показано, що жоден із доступних вільно сьогодні не відповідає потрібним вхідним параметрам і не містить достатньої кількості релевантних зображень. Запропоновані можливі шляхи вирішення даної проблеми.

Для виконання експериментальних досліджень наведені програмно-апаратні рішення, які потенційно можуть бути використані для вирішення всього комплексу завдань, включаючи обробку зображень і керування електродвигунами позиціонування спот-камери безпосередньо на борту БПЛА в умовах численних обмежень.

Ключові слова: *безпілотний літальний апарат (БПЛА); аналіз зображень; людино-машинна система; електропривід; бортовий детектор об'єктів; керування спот-камерою.*

Рис.: 5. Табл.: 1. Бібл.: 22.

Voytenko, V., Olofsson, B., Solodchuk, M., & Denisov, Yu. (2023). Components of a system for automatic detection of a zone of interest in images obtained from a UAV. *Technical sciences and technologies,* (2(32)). Pp. 300-312.