UDC 621.391
B. Moroz, doctor of technical science, professor., A. Shcherbakov, graduate student;
Ukraine, Dnipro, University of Technology

AUTOMATED INFORMATION SYSTEM FOR VIDEO SURVEILLANCE TO PERFORM SPECIAL FUNCTIONS

Мороз Б. І., Щербаков А. Г. Автоматизована інформаційна система відеоспостереження для виконання спеціальних функцій. Представлений комплексний алгоритм створення автоматизованої системи реєстрації та відображення інформації з літальних апаратів спостереження в режимі інтерактивного управління оператором. Запропонована архітектура для шифрованої передачі потокового відео з декількох камер з літального апарату зі стабілізацією відео в польоті і проектуванням на шолом віртуальної реальності в перспективі огляду 360 градусів.
Ключові слова: FPV-система, відео реєстрація, потокове відео, алгоритм, відеокамера, алгоритм Діффі-Хеллмана, алгоритм YOLO, нейронна мережа.

Мороз Б. И., Щербаков А. Г. Автоматизированная информационная система видеонаблюдения для выполнения специальных функций. Представлен комплексный алгоритм создания автоматизированной системы регистрации и отображения информации с летательных аппаратов и наблюдения в режиме интерактивного управления оператором. Предложена архитектура для шифрованной передачи потокового видео с нескольких камер с летательного аппарата со стабилизацией видео в полете и проецированием на шлем виртуальной реальности в перспективе обзора 360 градусов.
Ключевые слова: FPV-система, видео регистрация, потоковое видео, алгоритм, видеокамера, алгоритм Диффи-Хеллмана, алгоритм YOLO, нейронная сеть.

B. Moroz., A. Shcherbakov. Automated information system for video surveillance to perform special functions. A complex algorithm for creating an automated system for recording and displaying information from aircraft and observation in interactive operator control mode was presented. An architecture for encrypted transmission of video streaming from several cameras from an aircraft with in-flight video stabilization and projection of a virtual reality helmet on a 360-degree perspective was proposed.
Keywords: FPV-system, video registration, streaming video, algorithm, video camera, Diffie-Hellman algorithm, YOLO algorithm, neural network.

Introduction and statement of the research problem. The history of the development of drones dates back to the 21st century, when the FPV-system became widespread and video recording from quadcopters became possible. Now drones are used to collect various information in large areas, aerial survey and photography, monitoring of facilities, etc. Since the technical characteristics allow the use of drones out of line of sight, improvement of the interactive system is required for comfortable control of drones.

As one of the ways to ensure comfortable interaction with the drone, an FPV system can be used. The FPV helmet will be used as the output interface, and the FPV-joystick as the input device. Compared to traditional output devices, the FPV system will change the general idea of how to control the drone. The new visual presentation will allow to control the drone from the first person with a radius of 360 degrees. Thus, the operator can fully control the environment by turning the head, without changing the position of the drone. The drone is controlled by a special joystick, and the flight path is selected by turning the head.

Purpose. To propose a complex algorithm for creating an automated system for recording and displaying information from observation aircraft in the interactive operator control mode. To describe the architecture for transmitting encrypted video stream from several cameras with video stabilization in flight from an aircraft and projecting on a helmet of a virtual reality in perspective of the review of 360 degrees.

Analysis of recent research and publications. To date, the number of scientific developments regarding the topic of research is sufficient.

K. Ratakonda [1] proposed a video stabilization technology which allows to eliminate the effects of unwanted camera movements without affecting the true image.

J. Redmon [2] revealed in the pages of his work a new way to recognize facilities with a simple neural network.

S. Gabreith [3] conducted a comparative analysis of the Diffie-Hoffman algorithm for data encryption and the possibility of improving this algorithm.

A. Protogerellis [4] considered the fundamental opportunities of an FPV system for controlling aircraft. Besides, the author covers topics such as control channel, video communication, component placement, ground station design, troubleshooting and interference. He also provided his own configuration for setting up an FPV system.

Nevertheless, despite the scale of scientific research, the issue of creating a single automated information system for performing special functions remains open and requires detailed study.

The presentation of the main material.

To create an automated video surveillance system, it is necessary to create an appropriate aircraft that will capture streaming video from several cameras and which will provide the ability to encrypt and transmit data to the operator in real time. For vision in the night spectrum, cameras must operate in the infrared. The aircraft must be able to stabilize the video in motion, because the image without stabilization may deteriorate significantly in contrast to the video, which is shot by the camera on a tripod. Other difficulties for creating an automated video surveillance system are that the aircraft must be able to recognize facilities in real time in order to follow them. In the reviewed papers there is no solution to the problem of an automated video surveillance system that would unite all the technologies of the reviewed papers. The algorithm for implementing an automated information system for video surveillance to perform special functions, which is proposed as part of this study, is shown in Figure 1.

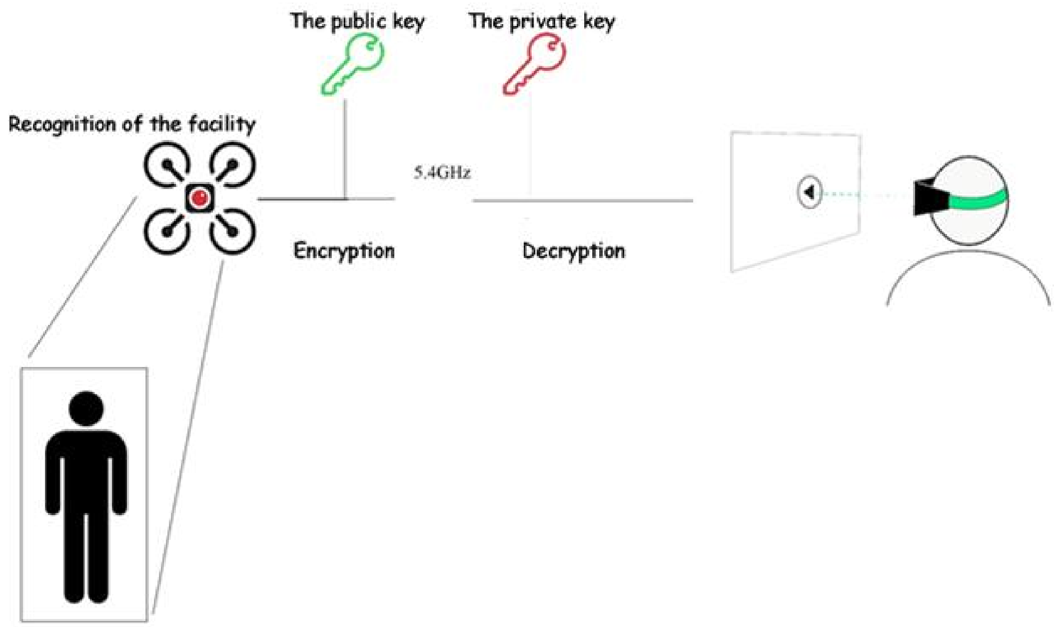

Fig. 1. Scheme of the complex algorithm of the automated system for video surveillance from aircraft

We will consider the software image stabilization principles on an aircraft. The principle of the software stabilizer, on the one hand, resembles a digital stabilizer in the camera, but there are a number of significant differences. If there is no motion at the edges of the frame, this filling of the edges works very well: it is possible to shoot a distant stationary facility with a shake even half a frame – the process of stabilization "sticks together" from this set of frames a single "panorama" and will slowly move the focus along it. Of course, this technique does not always work well, but ideally, it allows you not to reduce the frame size at all and not to lose in the resolution, minus the inevitable re-interpolation of the picture during the shift, and in the viewing angle, which is also important. Motion detection methods, as a rule, work similarly to the methods used in MPEG-like compression. That is, the frame is divided into blocks. For each of them, the most similar block in the previous frame and offset relative to it are selected. The average characteristics for the entire frame are determined from the constructed displacement map. As a rule, these are two-four values: horizontal and vertical displacement and often rotation and change of scale. In this case, it is possible to discard those blocks that move in apparent disagreement with the general direction, since they most likely correspond to the movement of individual facilities in the frame relative to the background, or are simply incorrectly identified in the previous frame. There are other ways, for example, the Fourier analysis applied in the DePan filter, but the output usually yields the same values [5]. Next, the stabilization module directly enters, which builds the optimal trajectory of the camera by smoothing the existing "chaotic" one, performs frame shifting, rotation and scaling to the corresponding values, fills edges, etc. The selection of the optimal smoothed trajectory is similar to applying a low-pass filter to the "signal" formed by displacement vectors, i.e., in fact, oscillations with frequencies higher than the cut-off frequency set by the user are removed. Thus, the operator can watch the smooth streaming video transmitted from the aircraft in real time.

At the first stage, the aircraft recognizes the facility by using the toolkit, which represents a new way to localize facilities in the image – YOLO (You Only Look Once) [6]. The localization problem is formulated as a single-stage regression problem: a single neural network accepts the entire image as input and gives the coordinates of the bounding rectangles and the probabilities of belonging to classes for them. YOLO divides the image into an S x S grid. Responsible for the detection of a facility is that grid cell, which gets its center. Each cell searches for bounding rectangles of the same class. A rectangle is characterized by five numbers - 4 coordinates and confidence in it of the neural network. Then, the confidence estimate for each detection zone is multiplied by the class probability to get the final estimate S x S x (B * 5 + C) [7]. YOLO imposes strong restrictions on the spatial location of facilities. Because of its lattice structure, the system copes well with large facilities located at a considerable distance from each other. And also, thanks to the high speed image processing, it is suitable for use in real-time systems.

To encrypt data transmitted from the aircraft to the operator, it was proposed to use the asymmetric Diffie-Hellman cipher. In an asymmetric cipher, there are two keys, k1 and k2, connected by a non-trivial relation [8]. If a message is encrypted on the key k1, it can be decrypted on the key k2. The implementation of this encryption is made due to the existence of unidirectional functions, it means such functions that can be easily calculated in one direction, but without knowing the private key, it is impossible to decrypt them. In any case, effective methods for decrypting a private key do not exist yet.

We will consider the example presented in the scheme of the complex algorithm presented in Fig. 2 between the aircraft and the operator.

Let they have common knowledge about the numbers P and G. This is the base of the field and the generator. The aircraft generates a large random number, and calculates the generator in degree a modulo R. It sends the result A to the operator. The operator takes some random number B, calculates the same generator in degree B modulo R. The generator and module G and P are open data. If we look at the real cryptographic protocols, G and P are properties of a particular group.

Then the operator calculates B and sends it to the aircraft. The drone takes B and raises to its degree, and the operator takes the number obtained from the aircraft, and raises to the degree b. The equality $B^a = A^b$ is obtained. Thus, two parties to the connection can exchange data in an encrypted communication channel.

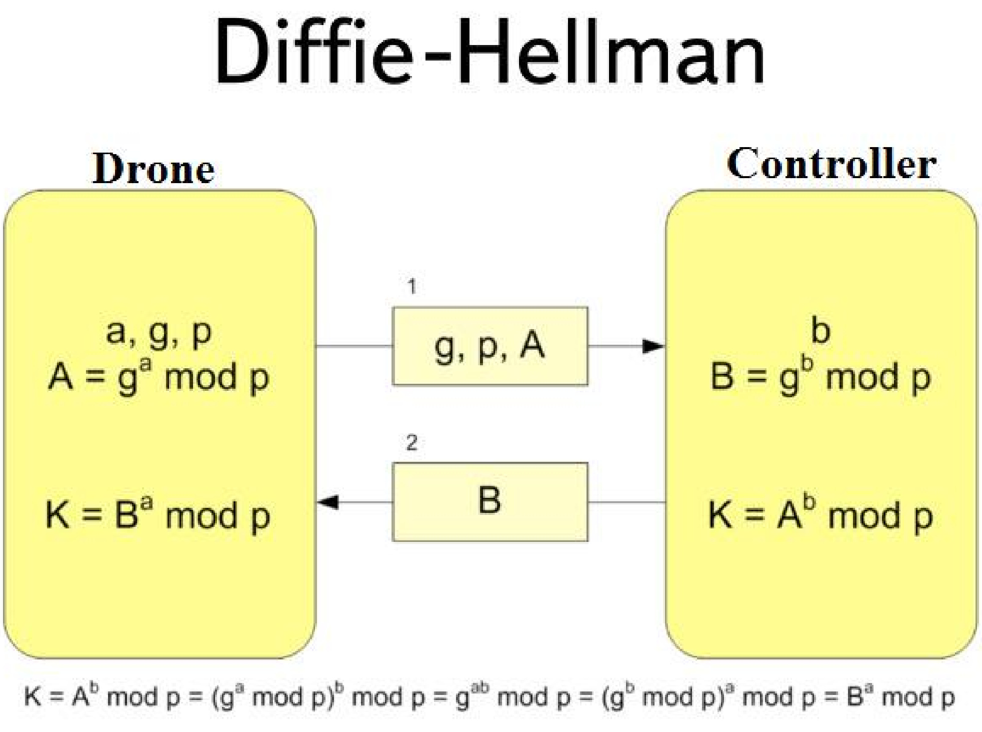

Fig. 2. The Diffie-Hellman algorithm on the example of an aircraft and an operator

The ability to control the aircraft with the ability to view the panorama of the adjacent territory through special VR-glasses appeared thanks to the FPV system. The abbreviation FPV used for flying drones appeared relatively recently. It describes the technology, the full name of which sounds like the First Person View. It appeared due to the use of cameras installed on board aircrafts. Its essence is to provide the user of the technology, in particular the operator, with the ability to view a real-time image captured by the camera. Thus, this technology allows to control the drone from the first person.

The application of this technology is almost unlimited. It can be used both for entertainment and in the professional field. Therefore, drones, in particular multi-copters and quadrocopters, with the FPV function are used in various sectors of the economy. They are used by farmers for fertilizing and tracking cattle, law enforcement agencies and rescue services associated with them also use them, but they use them in places where a person’s appearance is associated with a risk to life, and not just convenience. With their help, rescuers can inspect the facility before taking measures to eliminate the accident or apprehend armed criminals.

FPV function is not possible without an FPV system. It is multicomponent and implemented through several technologies. One of them means capturing an image, the second one - processing the received data, if necessary, complementing it with information read from sensors for implementing the OSD interface technology, the third one - by wireless transmission and reception of a signal to deliver the picture to the end user – the operator. All technologies are connected inextricably and are created from several modules. Some of them are installed on a flying drone, a quadrocopter, and some are located at the end user. At the same time, modules installed on an aircraft can be both built-in and removable. Modules that the user has, depending on the data transfer technology used, can be installed in the control equipment or they may be part of a paired mobile device, provided it is used [9].

The principle of video transmission is as follows: a camera on an aircraft – a video transmitter – a receiver in an FPV helmet – video on the helmet / glasses screen (Fig. 3). The signal is television that eliminates any delay in the transmission of the signal to the equipment. The signal is transmitted at a frequency of 5.4GHz, to prevent interference from surrounding devices and the possibility of transmitting large amounts of data over long distances. The range of stable operation of the signal using the FPV system usually does not exceed 1 km, but if we set the power to 600 mW, the aircraft will transmit video up to 3 km in line of sight [10]. The power must simultaneously increase both on the transmitter device and on the receiver device in order for the connection to remain stable throughout the entire flight of the aircraft.

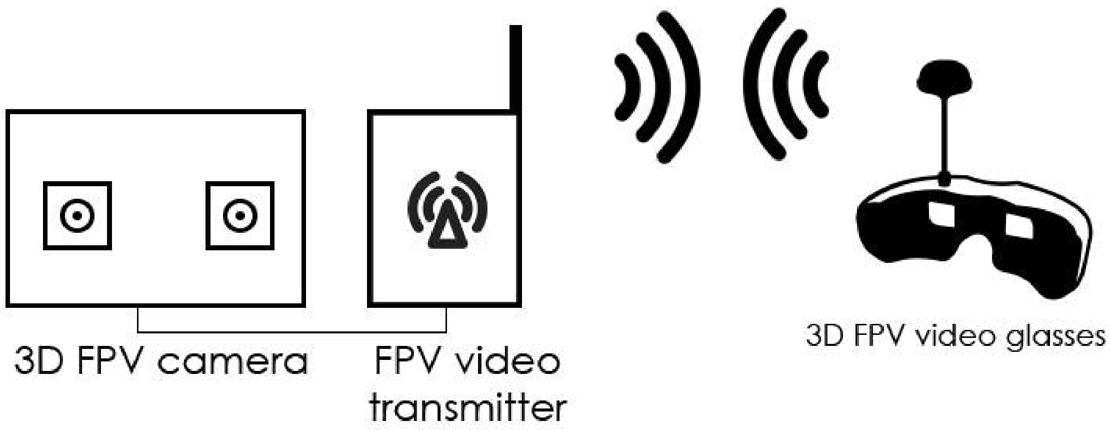

Fig. 3. Video transmission scheme in the FPV system between the camera and FPV glasses

Conclusion and prospects for further development. A new integrated algorithm for creating an automated system for recording and displaying information from observation aircraft in the interactive operator control mode is presented. An architectural solution has been proposed for encrypting the transmission of streaming video from several cameras from an aircraft with video stabilization in flight and projection on a helmet of a virtual reality in perspective of the review of 360 degrees.

It is proposed to use the YOLO algorithm for recognition of the fecility in real time with the possibility of its capture and following the drone behind it.

The algorithm works on the basis of an analog FPV system for streaming video from an aircraft to FPV glasses, a Diffie-Hellman algorithm for encrypting streaming data.

1. Ratakonda K. Real-time digital video stabilization for multi-media applications // ISCAS '98. Proceedings of the 1998

IEEE International Symposium on Circuits and Systems (Cat. No.98CH36187). – Monterey, CA: IEEE Computer Society, 1998. –P. 26-30.
2. Redmon J., Ehsani K., Bagherinezhad H., Mottaghi R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection // Conference on Computer Vision and Pattern Recognition. – Salt Lake City, UT: IEEE Computer Society, 2018. –P. 4051-4060.
3. Gabreith. S. The Diffie-Hellman Problem and Cryptographic Applications. // Key Exchange. – ACM Press, NY: IEEE Transactions on Information Theory, 2005. –P. 470-504.
4. Protogerellis A. The beginner's guide to FPV; trans. with English. – M.: McGraw-Hill Education TAB, 2016. –177 p.
5. Marius T. Digital Image Stabilization; trans. with English. – M.: InTech, 2009. - 544 p.
6. Norman Di Palo. How to add Person Tracking to a Drone using Deep Learning and NanoNets. [Electronic resource] // URL: https://medium.com/nanonets/how-i-built-a-self-flying-drone-to-track-people-in-under-50-lines-of-code-7485de7f828e.
7. Jonathan Hui. Real-time Object Detection with YOLO, YOLOv2 and now YOLOv3. [Electronic resource] // URL: https://medium.com/@jonathan_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088.
8. Steven D. Galbraith, Florian Hess, Frederik Vercauteren: Aspects of Pairing Inversion. IEEE Trans. Information Theory, 2008. –P. 5719-5728.
9. Glover J. Drone University; trans. with English. – M.: Amazon Digital Services LLC, 2014. - 134 p.
10. Yang J., Xu Y., Chen C. Gesture Interface: Modeling and Learning // IEEE International Conference on Robotics and Automation (San Diego, CA, 1994). – IEEE Computer Society, 1994. –V. 2. –P. 1747-1752.

Reviewer: Volodymyr Gnatushenko, doctor of technical science, professor, Head of the Department of computer science and IT education Oles Honchar Dnipro National University.