

UDC 004.032.2:629.7.014 (045)  
DOI: 10.18372/2306-1472.1.13651

Nataliia Kuzmenko<sup>1</sup>  
Volodymyr Kharchenko<sup>2</sup>  
Ivan Ostroumov<sup>3</sup>

## AN INVESTIGATION OF AGGREGATE CHANNEL FEATURES OBJECT DETECTOR FOR UAS APPLICATION

National Aviation University, 1, Kosmonavta Komarova Avenue, Kyiv, 03058, Ukraine  
E-mails: <sup>1</sup>nataliakuzmenko@ukr.net; <sup>2</sup>kharch@nau.edu.ua; <sup>3</sup>vany@nau.edu.ua

### Abstract

**Purpose:** The paper is aimed to point out an artificial intelligence as a key priority for research and development issues. Consideration of existing methods of object detection is one of the important tasks of the paper. The represented research results are aimed to investigate a problem of moving object detection by visual sensor data for Unmanned Aerial System application. **Methods:** Represented approach is grounded on probabilistic and statistical methods of data processing, in particular on Aggregate Channel Features approach usage for object detection. **Results:** An Aggregate Channel Features approach for object detection by video-stream at different scenarios has been practically investigated. Results of experimental investigation of ACF usage for moving vehicles detection, such as cars and trams, indicate good performance characteristics. Also, a dependence between training time of detector and amount of object positive instances has been investigated for particular case. **Discussion:** Multiple advantages of Aggregate Channel Features object detector such as universality, simplicity of realization and good compromise between computation time and detection accuracy allow to use it in the tasks of people, vehicles, artificial and natural objects detection in UAS application. Represented results can be implemented in Unmanned Aerial Systems for searching and tracking of movable objects.

**Keywords:** UAS; artificial intelligence; Aggregate Channel Features; object detection; video-stream

### 1. Introduction

An artificial intelligence (AI) has already become a part of our lives from coffeemaker up to space objects systems control.

In addition to make humans lives easier, AI helps to solve the biggest worldwide challenges, such as treating diseases, reducing traffic accidents rates, fighting climate changes or cyber security threats, etc.

Fast growing in computing power, availability of big data and progress in algorithms have made AI one of the most advanced technologies of the 21st century. Today AI defines the world we live in.

AI refers to systems that display intelligent behavior by analyzing their environment and taking actions with some degree of autonomy to achieve specific goals [1]. It is also refers to computer programs that exhibit human-like intelligence such as logical reasoning, problem solving and learning [2].

Many countries all over the world including United States, China, Japan and Canada support AI implementation in different applications and are making huge investments. Also, on April 10, 2018,

twenty five European countries signed a Declaration of cooperation on AI [3]. Moreover, European Union spend around EUR 1.1 billion into AI-related research and innovation during the period 2014-2017 under the Horizon 2020 program, including big data, health, rehabilitation, transport and space-oriented research.

According to European Union initiative on AI, the European Union including public and private sectors should aim to increase investments in AI to at least EUR 20 billion by the end of 2020 that proves topicality of AI for mankind.

An object detection and recognition by Unmanned Aerial Vehicle (UAV) video-stream is a good example of AI-related tasks important for modern life. Its application can be found not only in military but also in civilian purposes. The task of object detection consists in determining object position and its size relative to the frame. The primary criteria in task solving for Unmanned Aerial System (UAS) application are accuracy and speed of object detection.

## 2. Analysis of object detection and recognition methods

Nowadays, there are a lot of methods for solving the task of object detection and recognition. However, each of them has particular application and a set of constraints.

Cascade methods and neural networks are particularly notable approaches.

Cascade methods are widespread due to their high detection speed, allowing them to be used in real-time applications. But at the same time, comparing them with modern solutions, cascade methods carry out object detection with a large percentage of false detections, decreasing accuracy in comparison with other methods [4].

Cascade algorithms, also known as classifiers with a sliding window, are widely used. The first method of this type was proposed by Paul Viola and Michael Jones for the task of finding faces in photos. They used Haar features that are rectangular whose appearance is similar to Haar wavelets. Later, usage of local binary pattern features became popular due to their faster computation speed. Object detection by cascade methods using these features is based on calculation of the feature convolution with an image.

The main advantage of this algorithm is speed of detection, compromising to accuracy. According to authors statements, the detector works with a speed equal to 15 frames per second, reducing the level of false detections by a half [4]. This is achieved by using the integral representation of the image instead of pixel image, that allows the feature to be calculated very fast. In addition, the choice of the optimal feature of all possible types is carried out by modifying AdaBoost learning algorithm, that improves accuracy of detection. The AdaBoost is an algorithm that amplifies the classifiers that showed poor results at the previous stages, for further combining them into one final classifier [5]. This algorithm is called cascade due to the use of a cascade of classifiers, which during training focuses on the relevant area with the desired object. This approach contributes to the early rejecting of regions where the object could not be located [6].

Neural network is another method that has become popular to combine with other approaches. The study of neural networks has a vast history from the 20th century. Their analysis is connected with the development of artificial intelligence. Classic neural networks are a directed graph with nodes as neurons, which are arranged in layers in such a way that neurons of one layer are connected with neurons of the next layer.

Neural networks, show high accuracy of detection, greatly reducing the level of false detections. Disadvantage of neural networks is low detection speed, that make them worse in comparison with other approaches despite the network size and layers connectivity. In addition, in paper [7] was proved that usage of neural networks is not allowed in task of pedestrians detection in real applications due to low computation speed. At the same time, the approaches that combine both speed and accuracy remain very topical in real applications. These methods include hybrid approaches, combining cascade methods with deep neural networks, have become popular today. This decision to combine the best of each method demonstrates good results.

Bayesian networks (BN) based on probabilistic models are well known and very effective for data analysis and pattern recognition in domains with uncertainty [8,9].

The VeryFast method is a cascading sliding-window algorithm that uses Histogram of Oriented Gradient (HOG) features. This is the first method that has combined cascades with HOG features [10].

HOG features, were proposed by Navneet Dalal and Bill Triggs in 2005 for the task of pedestrians detection on the INRIA database [11]. They showed good results adapting these features for the Support Vector Machine learning algorithm. This method focuses on detection speed, reaching 50 frames per second [12]. This is achieved by combining ChnFtrs and FPDW (Fastest Pedestrian Detector in the West) highly efficient methods.

ChnFtrs method is the basis of the VeryFast algorithm. The ChnFtrs detector was proposed by Piotr Dollar in 2009, and is based on the idea of Integral Channel Features [12]. These are rectangular features that determine response of the filter on the considered image area. In the original article, 6 quantized orientations, 1 gradient value, and 3 LUV color channels were used, which are sufficient for obtaining competitive results. According to the ChnFtrs algorithm, after such a representation a two-level decision tree is built, according to which a strong classifier is created as a linear combination of weights. The training of such trees and their weights is carried out by the discrete AdaBoost algorithm.

FPDW method is a modification of the ChnFtrs algorithm to a multiscale detector developed by Piotr Dollar in 2010 [13]. If classic detectors working on multiple scales are trained in a single model for a single pre-selected scale, followed by changing the

image size  $N$  times, the idea of the FPDW detector is in resizing images by  $N/K$  times, also learning from a single model on a single scale. Each modified image in such a way participates in a convolution with a feature, which are later used in a convolution together with the remaining  $N-N/K$  scales.  $K$  coefficient was chosen by authors in an empirical way with the purpose to reduce the number of changes in the size of images and features calculating, and is equal about to 10.

### 3. Research tasks

The research tasks are:

- Analysis of object detection and recognition methods;
- Study of Aggregate Channel Features approach for object detection;
- Practical investigation of Aggregate Channel Features approach for object detection by video-stream at different scenarios.

### 4. Object detection by Aggregate Channel Features approach

Aggregate Channel Features object detector has been proposed in [14].

ACF detector computes several channels  $C = \Omega(I)$  on an input image  $I$ , sums every block of pixels in  $C$ , and smooths the resulting lower resolution channels.

Features are single pixel lookups in the aggregated channels. Boosting is used to train and combine decision trees over these features (pixels) to distinguish object from background and a multiscale sliding-window approach is employed. With the appropriate choice of channels and careful attention to design, ACF achieves great performance in pedestrian detection.

ACF uses the following channels: normalized gradient magnitude, histogram of oriented gradients (6 channels), and LUV color channels

LUV color space contains 3 channels,  $L$  channel describes the lightness of the object,  $U$  channel and  $V$  channel represent the chromaticity of the object. Compared to RGB space, LUV space is able to partially invariant to illumination change. So the proposed detector can work under different light conditions. Images can be converted to LUV space by using a specific transformation.

A normalized gradient magnitude is used to measure the edge strength. Gradient magnitude  $M(x,y)$  at location  $(x,y)$  is computed by the equation:

$$M^2(x,y) = I_x^2 + I_y^2,$$

where  $I_x$  and  $I_y$  are first intensity derivatives along the x-axis and y-axis, respectively. Since the gradient magnitude is computed on 3 LUV channels independently, only the maximum response is used as the gradient magnitude channel.

A histogram of oriented gradients is a weighted histogram where bin index is determined by gradient orientation and weighted by gradient magnitude [8]. The histogram of oriented gradients at location  $(x,y)$  is computed by the following equation:

$$M(x,y) \cdot i[\Theta(x,y) = \theta],$$

where  $i$  – the indicator function,  $M(x,y)$  and  $\Theta(x,y)$  are the gradient magnitude and discrete gradient orientation, respectively. ACF quantizes the orientation space to 6 orientations and compute one gradient histogram channel for each orientation.

Comparison of miss rates of pedestrian detection on four datasets using ACF and leading approaches show its high accuracy.

### 5. ACF for UAS application

Due to several advantages of ACF object detector in comparison with other methods, its usage for UAS application is approved. ACF object detector is a universal mean of detection, because it can be trained for any detection tasks including object detection in video-stream. In comparison with neural networks, ACF approach requires less computation time and is much easier in its realization. Also, it automatically collects negative instances from the images during training and provides confidence scores for detected objects, in contrast to Cascade object detector. ACF implementation for video data processing allows to detect any type of objects such as people, vehicles, artificial and nature objects, etc.

A typical Unmanned Aerial System consists of UAV, Ground Control Station (GCS) and communication data link (Fig.1). A camera is a basic UAV payload for object detection and recognition. Angles of video capturing are maintained by gimbaled system in UAV. Angles of camera orientation are set at the GCS by remote pilot. UAV positioning and guidance is provided by autopilot unit, receivers of Global Navigation Satellite System signals and numerous on-board sensors. Planned UAV trajectory is prepared in GCS and loaded to autopilot by communication units. Received UAV video data in GCS is decoding to one of the suitable formats such as .mp4 or .avi. Video data is represented as a set of frames for ACF object detector. Moreover, ACF detector is pre-trained for the desired object of

detection using its positive samples. As a result of detection, detected objects and their confidence scores are saved in a data base available for remote pilot to make appropriate decision.

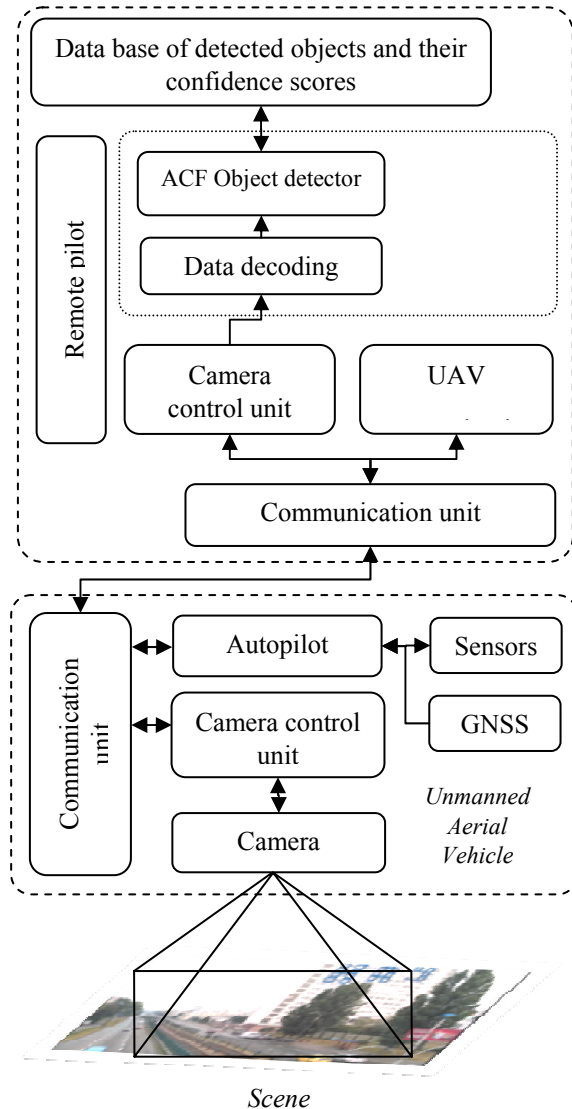


Fig.1. Block-scheme of a typical UAS video data processing

### 5. Simulation

For ACF detector verification, video-stream data from Nikon D3200 camera was used. Video duration was equal to 30 sec. Video frame size was equal to 720x1280 pixels with frame rate 29.8 fps. ACF detector was trained to detect trams in congested road traffic. For training 45 tram’s positive samples and Negative samples factor equal to nine was used in four-stage training. The number of negative samples to use at each stage was equal to negative samples factor multiplied by the number of positive

samples used at each stage. In order to create a ground truth table, the Image Labeler was used.

Also, video data processing was performed at GCS with Processor Pentium Dual-Core CPU 2.30 GHz.

Detected object is marked by bounding box defined by rectangular form with coordinates of the left bottom corner with height and width. In addition, each bounding box (detected object) is complemented with its confidence scores. Larger score values indicate higher confidence in the detection.

The task of simulation was a practical investigation of ACF approach for different vehicle types detection (trams and cars) at congested road traffic at a randomly selected frames.

Results of tram detection with confidence scores using trained ACF detector (acfObjectDetector) are represented on Fig.2 (frame 262) and Fig.3 (frame 242). On fig.2 confidence score of tram detection is equal to 86.2, on Fig.3 – 79.5.

Results of cars detection with confidence scores using pretrained vehicle detector using ACF (vehicleDetectorACF) that uses 295 positive vehicles samples are represented on Fig.4 (frame 342) and Fig.5 (frame 647). On Fig.4 ten cars were detected with scores from 8.5 to 32.4; on Fig.5 – five cars were detected with scores from 9.8 to 22.8.



Fig. 2. Detected tram and detection scores for frame 262



Fig. 3. Detected tram and detection scores for frame 242



Fig. 4. Detected cars and detection scores for frame 342



Fig. 5. Detected cars and detection scores for frame 647

In order to achieve high performance of object detection and recognition in terms of accuracy and computation time, it is necessary to investigate a dilemma between the amount of positive instances and required time for training. Positive instances includes images of the detecting object given in the training data table. It is evident, that bigger number of positive instances will guarantee higher accuracy of detection, however it simultaneously increases time of detector training. Result of dependence estimation between training time and amount of positive object instances is represented on Fig.6.

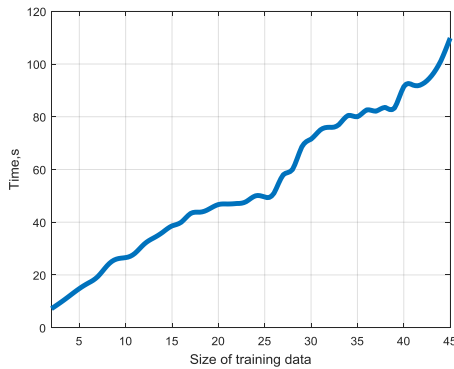


Fig. 6. Dependence between training time and amount of positive instances

It is necessary to point that training time also depends on quality and size of positive instances, however the dependence still has continuously increasing behavior.

ACF object detector was investigated for car detection ( $N=10$ ) within the same frame with different Negative Samples Factor(NSF) (Table). Results of ACF object detection include total number of detected BLOBs(Binary Large Objects) that contains number of cars correct detection; omitted cars and false detected objects.

Table

Results of ACF object detection

NSF	2	3	4	5	6	7	8	9
Total BLOBs	15	16	23	25	24	10	8	8
Object Detection	4	5	8	9	7	5	4	5
Target omission	6	5	2	1	3	5	6	5
False detection	11	11	15	16	17	5	4	3

One of the possible ways to estimate ACF detector performance is usage of statistical method with frequency estimation of event occurrence in probabilistic representation. Probability of object correct detection ( $P_{11}$ ), probability of target omission ( $P_{01}$ ), probability of false alarm ( $P_{10}$ ), and probability of correct omission ( $P_{00}$ ) have been estimated (Fig 7).

$$P_{11} = N_{OD} / N,$$

where  $N_{OD}$  – number of detected cars on the frame;  $N$  – number of existing cars on the frame.

$$P_{10} = (N_{TB} - N_{OD}) / N_{TB},$$

where  $N_{TB}$  – number of detected BLOBs on the frame.

Moreover  $P_{11}$  creates total probability with  $P_{01}$ ;  $P_{10}$  creates total probability with  $P_{00}$ . Therefore,

$$P_{01} = 1 - P_{11}; P_{00} = 1 - P_{10}.$$

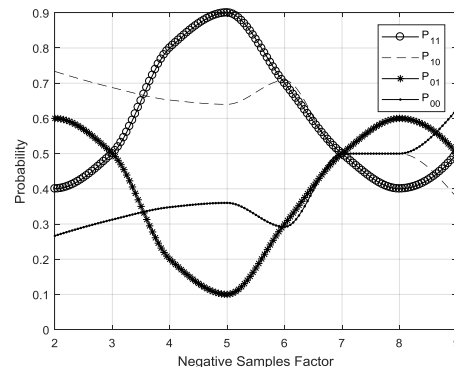


Fig. 7. ACF detector performance according to value of negative samples factor

Results of ACF detector performance estimation show that the best characteristics such as the maximum probability of correct object detection and minimum probability of target omission is reached for NSF equal to five.

## 8. Conclusions

Due to sophisticated algorithms, high computational power and huge amount of data, technologies such as computer vision achieved amazing results. AI is becoming more topical because it makes the quality of human lives better. However, AI should not be considered as a substitute of a mankind, it should be a complement to humans.

Usage of ACF object detector for UAS application is approved due to a set of its advantages in comparison with other existing methods. Results of experimental investigation of ACF usage for moving vehicle detection by video data processing indicate good performance characteristics. Also, an investigation proved continuously increasing behavior between training time of detector and amount of object positive instances.

## References

[1] European Commission (2018) Communication from the Commission to the European Parliament, the European Council, the Council, the European economic and social committee and the committee of the regions. Artificial Intelligence for Europe.– Brussels, 19 p.

[2] IATA (2018) AI in aviation. Exploring the fundamentals, threats and opportunities of artificial intelligence (AI) in the aviation industry. White Paper, 20 p.

[3] European Commission (2018) Declaration of cooperation on Artificial Intelligence, Brussels. Available at:

<https://ec.europa.eu/jrc/communities/sites/jrccties/files/2018aideclarationatdigitaldaydocxpdf.pdf>

[4] Viola P., Jones M. (2001) Rapid object detection using a boosted cascade of simple features.

IEEE Conference on Computer Vision and Pattern Recognition, 9 p.

[5] Freund, Y. and Schapire, R.E. (1997) A decision-theoretic generalization of on-line learning and an application to boosting. Journal of computer and system sciences, 55(1). – P. 119–139.

[6] Jones, M.J. and Viola, P., (2001) Robust real-time object detection. In Workshop on statistical and computational theories of vision, Vol. 266. – 56 p.

[7] Angelova, A., Krizhevsky, A. and Vanhoucke, V. (2015) Pedestrian detection with a large-field-of-view deep network. In Robotics and Automation (ICRA), IEEE International Conference.– P. 704–711.

[8] Pérez, A., Larrañaga, P. and Inza, I., (2009) Bayesian classifiers based on kernel density estimation: Flexible classifiers. International Journal of Approximate Reasoning. – 50(2). – 341 p.

[9] Kharchenko, V., Kukush, A., Kuzmenko, N. and Ostroumov, I., (2017). Probabilistic Approach to Object Detection and Recognition for Videostream Processing. Proceedings of the NAU, № 2(71). – P. 8–14.

[10] Benenson, R., Mathias, M., Timofte, R. and Van Gool, L., (2012) Pedestrian detection at 100 frames per second. In Computer Vision and Pattern Recognition (CVPR), IEEE Conference. – P. 2903–2910.

[11] Dalal, N. and Triggs, B., (2005) Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition. CVPR 2005. IEEE Computer Society Conference on Vol. 1.– P. 886–893.

[12] Dollar P., Tu Z., Perona P., Belongie S. (2009) Integral Channel Features. British Machine Vision Conference. – 11p.

[13] Dollár, P., Belongie, S.J. and Perona, P., (2010) The fastest pedestrian detector in the west. In Bmvc, Vol. 2, No. 3.–7 p.

[14] Dollár, P., Appel, R., Belongie, S. and Perona, P., (2014) Fast feature pyramids for object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(8). – P.1532–1545.

**Н. С. Кузьменко<sup>1</sup>, В. П. Харченко<sup>2</sup>, І. В. Остроумов<sup>3</sup>**

**Дослідження методу виявлення об'єктів за узагальненими каналними характеристиками кольору для застосування у БАС**

Національний авіаційний університет, просп. Космонавта Комарова, 1, Київ, Україна, 03058

E-mail: <sup>1</sup>nataliakuzmenko@ukr.net; <sup>2</sup>kharch@nau.edu.ua; <sup>3</sup>vany@nau.edu.ua

**Мета:** Робота спрямована на те, щоб визначити штучний інтелект як ключовий пріоритет у питаннях досліджень та розробок. Розгляд існуючих методів виявлення об'єктів є одним з важливих завдань статті. Представлені результати досліджень спрямовані на дослідження проблеми виявлення рухомих об'єктів за допомогою даних візуальних датчиків для застосування у безпілотній авіаційній системі.

**Методи:** представлений підхід ґрунтується на імовірнісних та статистичних методах обробки даних, зокрема використання підходу узагальнених каналних характеристик для виявлення об'єктів.

**Результати:** Метод пошуку об'єктів за узагальненими каналними характеристиками за використанням відеопотоку з різними сценаріями був досліджений практичним способом. Результати експериментального дослідження використання узагальнених каналних характеристик для виявлення рухомих засобів транспорту, таких як автомобілів та трамваїв, свідчать про високі характеристики методу. Також досліджено залежність між часом тренування детектора та об'ємом позитивних екземплярів у конкретному випадку. **Обговорення:** Численні переваги методу виявлення об'єктів за узагальненими каналними характеристиками, такі як універсальність, простота реалізації та компроміс між часом обчислення та точністю виявлення, дозволяють використовувати його у завданнях виявлення людей, транспортних засобів, штучних та природних об'єктів для застосуванні у БАС. Представлені результати можуть бути впроваджені в безпілотні авіаційні системи для пошуку та моніторингу рухомих об'єктів.

**Ключові слова:** БАС; штучний інтелект; узагальнені каналні характеристики; виявлення об'єктів; відеопотік

**Н. С. Кузьменко<sup>1</sup>, В. П. Харченко<sup>2</sup>, И. В. Остроумов<sup>3</sup>**

**Метод обнаружения объектов по обобщенным каналными характеристиками цвета для применения в БАС**

Национальный авиационный университет, ул. Космонавта Комарова, 1, Киев, Украина, 03058

E-mail: <sup>1</sup>nataliikuzmenko@ukr.net; <sup>2</sup>kharch@nau.edu.ua; <sup>3</sup>vany@nau.edu.ua

**Цель:** Статья направлена на то, чтобы определить искусственный интеллект как ключевой приоритет у вопросах исследований и разработок. Рассмотрение существующих методов обнаружения объектов является одной из важных задач статьи. Представлены результаты исследований направлены на исследование проблемы обнаружения движущихся объектов с помощью данных визуальных датчиков для применения в беспилотной авиационной системе. **Методы:** представлен подход основывается на вероятностных и статистических методах обработки данных, в том числе использование подхода обобщенных каналных характеристик для обнаружения объектов.

**Результаты:** Метод поиска объектов по обобщенным каналными характеристиками за использованием видеопотока с разным сценариям был исследован практическим способом. Результаты экспериментального исследования использования обобщенных каналных характеристик для обнаружения подвижных средств транспорта, таких как автомобилей и трамваев, свидетельствуют о высоких характеристиках метода. Также исследована зависимость между временем тренировки детектора и объемом позитивных экземпляров в конкретном случае.

**Обсуждение:** Многочисленные преимущества метода обнаружения объектов по обобщенным каналным характеристикам, такие как универсальность, простота реализации и компромисс, между временем вычисления и точностью обнаружения позволяют использовать его в заданиях обнаружения людей, транспортных средств, искусственных и естественных объектов для применения в БАС. Представленные результаты могут быть внедрены в беспилотные авиационные системы для поиска и мониторинга подвижных объектов.

**Ключевые слова:** БАС; искусственный интеллект; обобщенные каналные характеристики; обнаружение объектов; видеопоток

**Nataliia Kuzmenko** (1990). Candidate of Engineering. Senior researcher.

National Aviation University.

Education: National Aviation University, Kyiv, Ukraine (2013).

Research area: navigation and control of dynamic systems, artificial intelligence.

Publications: 39.

E-mail: nataliikuzmenko@ukr.net

**Volodymyr Kharchenko.** Doctor of Engineering. Professor.  
Vice-Rector on Scientific Work, National Aviation University, Kyiv, Ukraine.  
Editor-in-Chief of the scientific journal Proceedings of the National Aviation University.  
Winner of the State Prize of Ukraine in Science and Technology, Honored Worker of Science and Technology of Ukraine.  
Education: Kyiv Institute of Civil Aviation Engineers, Kyiv, Ukraine.  
Research area: management of complex socio-technical systems, air navigation systems and automatic decision-making systems aimed at avoidance conflict situations, space information technology design, air navigation services in Ukraine provided by CNS/ATM systems.  
Publications: 540.  
E-mail: knarch@nau.edu.ua

**Ivan Ostroumov** (1983). Candidate of Engineering.  
Associate Professor. National Aviation University.  
Education: National Aviation University, Kyiv, Ukraine (2005).  
Research area: Navigation, Alternative Position, Navigation and Timing.  
Publications: 118.  
E-mail: vany@nau.edu.ua