



<http://dx.doi.org/10.35596/1729-7648-2020-18-2-96-104>

*Оригинальная статья*  
*Original paper*

УДК 623.618

## СПОСОБ ОБРАБОТКИ ИЗОБРАЖЕНИЙ В ЗАДАЧЕ ОБНАРУЖЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ В ОПТИКО-ЭЛЕКТРОННЫХ СИСТЕМАХ НАБЛЮДЕНИЯ ТЕПЛОВИЗИОННОГО ТИПА

МИХНЁНОК Е.И.

*Военная академия Республики Беларусь (г. Минск, Республика Беларусь)*

*Поступила в редакцию 13 марта 2020*

© Белорусский государственный университет информатики и радиоэлектроники, 2020

**Аннотация.** В статье рассмотрен предлагаемый автором способ обработки изображений применительно к задаче автоматического обнаружения движущихся объектов в оптико-электронных системах тепловизионного типа. Исследованию подлежат движущиеся объекты на наблюдаемой сцене, поэтому для решения задачи обнаружения целесообразно применение алгоритмов, основанных на методах вычитания фона. Однако среди наблюдаемых объектов могут находиться как объекты интереса (человек, транспортное средство), так и другие объекты и элементы фона, повышающие шумовую составляющую наблюдаемой обстановки. Также на повышение шумовой составляющей огромное влияние оказывает ложная сегментация в передний план областей обрабатываемых изображений при переносе поля зрения датчика оптико-электронной системы наблюдения. Целью настоящей статьи является доказательство снижения вероятности ложной тревоги автоматического обнаружителя за счет предлагаемых автором подходов к обработке изображений. Для исследования используется математический аппарат теории вероятностей и имитационное моделирование с последующей статистической обработкой данных. В статье показано, что вероятность ложной тревоги автоматического обнаружителя, построенного на методе вычитания фона, существенно возрастает после переноса поля зрения датчика оптико-электронной системы наблюдения и уменьшается после прекращения движения по мере автоматической сегментации ложно выделенных в передний план областей обрабатываемого изображения. Проведенное моделирование показало: предлагаемые автором подходы позволяют повысить пиковое отношение сигнал/шум обрабатываемых изображений и снизить вероятность ложной тревоги автоматического обнаружителя объектов интереса. Полученные результаты показывают реализуемость адаптации алгоритмов обнаружения, построенных на методах вычитания фона, для работы в сканирующих оптико-электронных системах наблюдения.

**Ключевые слова:** способ обработки, движущиеся объекты, оптико-электронная система.

**Конфликт интересов.** Автор заявляет об отсутствии конфликта интересов.

**Для цитирования.** Михнёнок Е.И. Способ обработки изображений в задаче обнаружения движущихся объектов в оптико-электронных системах наблюдения тепловизионного типа. Доклады БГУИР. 2020; 18(2): 96-104.

## METHOD OF IMAGE PROCESSING IN THE PROBLEM OF DETECTING MOVING OBJECTS IN OPTICAL-ELECTRONIC SURVEILLANCE SYSTEMS OF THERMAL IMAGING TYPE

EVGENY I. MIKHNIONOK

*Military Academy of the Republic of Belarus (Minsk, Republic of Belarus)*

*Submitted 13 March 2020*

© Belarusian State University of Informatics and Radioelectronics, 2020

**Abstract.** The article considers the method of image processing proposed by the author in relation to the problem of automatic detection of moving objects in optoelectronic thermal imaging systems. Moving objects on the observed scene are subject to investigation, so it is advisable to use algorithms based on background subtraction methods to solve the detection problem. However, the observed objects may include objects of interest (a person, a vehicle), as well as other objects and background elements that increase the noise component of the observed situation. Also, the increase in the noise component is greatly influenced by false segmentation in the foreground of the areas of processed images when transferring the field of view of the sensor of the optical-electronic surveillance system. The purpose of this article is to prove the reduction of the probability of false alarm of an automatic detector due to the author's proposed approaches to image processing. The research uses the mathematical apparatus of probability theory and simulation with subsequent statistical processing of data. The article shows that the probability of a false alarm of an automatic detector based on the background subtraction method increases significantly after the transfer of the field of view of the sensor of the optical-electronic surveillance system and decreases after the movement stops as the areas of the processed image that are falsely highlighted in the foreground are automatically segmented. The simulation showed that the approaches proposed by the author can increase the peak signal-to-noise ratio of processed images and reduce the probability of a false alarm of the automatic detector of objects of interest. The results obtained show the feasibility of adapting detection algorithms based on background subtraction methods to work in scanning optoelectronic surveillance systems.

**Keywords:** processing method, moving objects, optical-electronic system.

**Conflict of interests.** The author claims no conflict of interest.

**For citation.** Mikhniionok E.I. Method of image processing in the problem of detecting moving objects in optical-electronic surveillance systems of thermal imaging type. Doklady BGUIR. 2020; 18(2): 96-104.

### Введение

В настоящее время системы видеонаблюдения активно внедряются в различные сферы человеческой деятельности: в геологических и гидрологических исследованиях, лесном и сельском хозяйстве, охране окружающей среды, исследованиях космоса, в образовательных, военных и других целях. Для решения задачи обнаружения малококонтрастных, малоразмерных объектов актуально применение многоканальных сканирующих оптико-электронных систем наблюдения, функционирующих в любое время суток (день, ночь), в любых погодных условиях (дождь, туман и др.). Однако наличие каналов наблюдения различных частотных спектров вносит свои особенности в процесс функционирования данных систем. Так, например, при использовании датчиков ИК-диапазона в дневное время суток, особенно в летний период, местные предметы наблюдаемой сцены (столбы опор электропередачи, участки земли и др.) сильно нагреваются и на общем фоне имеют повышенный тепловой контраст. Это приводит к их ложному автоматическому обнаружению. Также дребезжание листвы, качание крон деревьев, наличие других дестабилизирующих факторов приводит к увеличению шумовой составляющей. Исключение влияния данных факторов на процесс обнаружения позволяет значительно повысить эффективность работы алгоритмов обнаружения и расширить их область применения.

В рассматриваемой задаче объектом интереса выступают человек и транспортное средство, а датчиком – оптико-электронная система с тепловизионным каналом, функционирующая в сложной фоноцелевой обстановке.

Учитывая тот факт, что объекты интереса характеризуются преимущественным движением на анализируемых участках местности, для реализации алгоритмов обнаружения с учетом особенностей наблюдаемой фоноцелевой обстановки целесообразно применение методов, основанных на вычитании фона.

Однако использование данных методов в «чистом виде», без дополнительных этапов обработки результатов их работы, не всегда в полном объеме позволяет минимизировать дестабилизирующие факторы в анализируемой сцене, а также шумовые воздействия, возникающие при переносе поля зрения датчика системы наблюдения.

### Основная часть

В качестве исходного метода для реализации автоматического обнаружителя использован метод вычитания фона, в котором модель фона строится на основе улучшенной адаптивной смеси гауссовых распределений (MOG2) [1]. Достоинствами данного метода является его адаптированность к резким изменениям освещения сцены, что достигается автоматическим выбором количества компонент смеси, описывающих пиксель изображения.

Принцип работы обнаружителя заключается в следующем. При построении модели фона с использованием смеси гауссовых распределений производится накопление статистики изменения интенсивности каждого пикселя  $(x, y)$  изображения  $I_k$  на определенном числе кадров  $\{X_1, X_2, \dots, X_k\} = \{I_j(x, y), j = \overline{1, k}\}$ .

Соответственно, наблюдение значения  $X_k(x, y)$  может определяться вероятностью [2]:

$$P(X_k) = \sum_{j=1}^S \omega_j^k N(X_k | \mu_j^k, \Sigma_j^k), \quad (1)$$

где  $\omega_j^k$  – вес  $j$ -го распределения кадра  $k$ ;  $\mu_j^k$  – математическое ожидание  $j$ -го распределения;  $\Sigma_j^k$  – среднеквадратичное отклонение  $j$ -го распределения;  $N(X_k | \mu_j^k, \Sigma_j^k)$  – функция плотности нормального распределения;  $S$  – количество распределений в смеси.

Таким образом, для каждого пикселя обрабатываемого кадра задается распределение значений его интенсивности. Для обновления параметров модели фона используется новое значение интенсивности, выраженное одной из компонент смеси распределений.

Сортировка распределений в порядке уменьшения величины (2) используется для классификации пикселей (фона или переднего плана). Распределение с наименьшей дисперсией и максимальным весом соответствует пикселю фона [2].

$$r_j^k = \frac{\omega_j^k}{\sigma_j^k}. \quad (2)$$

Определим, что распределения, удовлетворяющие условию (3), относятся к фоновым пикселям фона.

$$B^k = \operatorname{argmin}_b \left\{ \sum_{j=1}^b \omega_j^k > T_{\min} \right\}, \quad (3)$$

где  $T$  – минимальный объем данных, описывающих распределение пикселей фона.

Для каждого пикселя кадра  $I_{k+1}$  производится определение соответствия распределению из группы  $B^k$ :

$$\sqrt{(X_{k+1} - \mu_j^k)^T (\sigma_j^k)^2 (X_{k+1} - \mu_j^k)} < 2,5\sigma_j^k. \quad (4)$$

При наличии соответствия пиксель относится к группе фоновых. При отсутствии соответствующего распределения принимается решение о соответствии пикселя объекту.

В зависимости от соответствия пикселя распределению из группы  $B^k$  производится обновление параметров распределений для обработки кадра  $I_{k+1}$ .

При нахождении соответствия весовые коэффициенты и параметры распределений пересчитываются согласно формул [2]:

$$\begin{aligned} \omega_j^{k+1} &= \omega_j^k + \alpha(1 - \omega_j^k) - \alpha c_T, \\ \mu_j^{k+1} &= (1 - \rho)\mu_j^k + \rho X_{k+1}, \\ (\sigma_j^{k+1})^2 &= (1 - \rho)(\sigma_j^k)^2 + \rho(X_{k+1} - \mu_j^{k+1})(X_{k+1} - \mu_j^{k+1})^T, \end{aligned} \quad (5)$$

где  $\alpha$  – заданная константа;  $\rho = \alpha N(x_k | \mu_j^k, \Sigma_j^k)$ ;  $c_T = c/T$ ,  $c$  – параметр модели.

Распределениям, которым не соответствует  $X_{k+1}$ , пересчитываются веса согласно (6):

$$\omega_j^{k+1} = (1 - \alpha)\omega_j^k. \quad (6)$$

При отсутствии соответствия последнее распределение замещается новым со следующими параметрами:  $\mu_s^{k+1} = X_{k+1}$ ,  $(\sigma_s^{k+1})^2$  максимальная,  $\omega_s^{k+1}$  минимальным.

Таким образом, обнаруживаются все движущиеся объекты на наблюдаемой сцене. Однако вместе с объектами интереса могут ложно обнаружиться местные объекты с повышенным тепловым контрастом, а также ложно детектироваться пиксели изображения в объекты переднего плана после переноса поля зрения датчика. Поэтому дальнейшие этапы направлены на исключение данных объектов.

Обнаруженные объекты поступают на этап пространственной фильтрации. Данный этап заключается в последовательном применении операций математической морфологии, оконтуривания и фильтрации объектов, не удовлетворяющих условиям минимальных размеров.

Термин математической морфологии относится к описанию свойств формы областей на изображении. Основными операциями являются расширение (дилатация), сжатие (эрозия), замыкание и размыкание [3].

Операции расширение (7) и сжатие (8) определяются следующим образом:

$$\delta_B(S) = S \oplus B = \{s + b \mid s \in S; b \in B\}, \quad (7)$$

$$\varepsilon_B(S) = S - B = \{B_h \subseteq S \mid h \in E\}, \quad (8)$$

где  $\delta_B(S)$  – расширение подмножества  $S$  со структурным элементом  $B$ ;  $\varepsilon_B(S)$  – сжатие подмножества  $S$  со структурным элементом  $B$ ;  $E$  – дискретная плоскость;  $h$  – вектор, определяющий каждую позицию  $B$  на  $E$ ;  $B_h$  – перенос  $B$  относительно  $h$ ;  $S$  – совокупность пикселей изображения объекта интереса;  $B$  – структурный элемент.

Результат обработки существенно зависит от конфигурации и размеров выбранного структурного элемента. Определение его конфигурации является эмпирическим процессом, а выбор размера представляет собой оптимизационную задачу, решение которой зависит от определенных факторов. Во-первых, необходимо сохранить топологию интересующего объекта, во-вторых, максимально использовать возможности методов математической морфологии для устранения дефектов изображения, в-третьих, необходимо минимизировать время обработки.

Последовательное выполнения операции сжатия и расширения определяют операцию размыкания (9):

$$A \circ B = (A(-)B) \oplus B, \quad (9)$$

где  $A$  – обрабатываемое изображение,  $B$  – структурный элемент.

В данном случае операция размыкания позволяет удалить объекты, размеры которых меньше размеров структурного элемента, а также отлично удаляет линии, толщина которых меньше, чем диаметр структурного элемента.

В свою очередь последовательное выполнение операций расширения и сжатия определяют операцию замыкания (10).

$$A \bullet B = (A \oplus B)(-)B. \quad (10)$$

Применение операции замыкания позволяет убрать малые дыры и щели в изображениях объектов. Параллельно с применением операций математической морфологии производится оконтуривание объектов. Целью данного этапа является выделение объектов, имеющих замкнутый контур. Вследствие того, что движение на наблюдаемых участках сцены могут осуществлять не только объекты интереса, но и другие объекты, например, дикие животные, то возникает необходимость определения минимального значения порога размеров обрабатываемых объектов.

Так как объектами интереса с наименьшими размерами выступает человек, то именно для данного типа объектов необходимо определить минимальный порог. В целях определения значения данной величины было проведено экспериментальное исследование. Объектом исследования выступал взрослый человек среднего роста (165 см), средством наблюдения являлся тепловизионный канал оптико-электронной системы наблюдения.

Для получения значения с доверительной вероятностью 0,9 и доверительным интервалом 5 % было проведено 2180 измерений. Результаты измерений показали, что при наблюдении в фас фигура человека занимает в среднем 28 пикселей, а при наблюдении в профиль – 24 пикселя. Поэтому для определения минимального порога размера объекта принимался случай при наблюдении его в профиль. Среднее значение составило 24 пикселя, а среднеквадратическое отклонение – 1,4 пикселя. Следовательно, значение минимального порога размера объектов интереса составило  $S_{\min} = 20$  пикселей.

Конечным является этап временной фильтрации объектов, заключающийся в выполнении условия

$$k / n, \quad (11)$$

где  $k$  – требуемое количество последовательно обрабатываемых кадров, на которых наблюдался объект интереса;  $n$  – количество обрабатываемых кадров.

С целью устранения эффекта ложного детектирования областей изображения и отнесения их к группе объектов интереса в случаях переноса поля зрения датчика применяется отключение работы обнаружителя при движении оптико-электронной системы наблюдения.

Блок-схема алгоритма разработанного способа обработки изображений в задаче автоматического обнаружения движущихся объектов представлена на рис. 1.

Оценку эффективности применения разработанного способа проведем для показателей, характеризующих качество обработки изображения и качество решения задачи обнаружения. В качестве примера используем видеозапись движения группы нарушителей по участку Государственной границы. Одним из показателей, характеризующих качество обработки изображений, является пиковое отношение сигнал/шум [4]. В качестве критерия используем максимизацию его значения. Пиковое отношение сигнал/шум (PSNR) представляет собой отношение максимального значения сигнала к среднеквадратичному уровню шума, которое определяется разницей между свойствами пикселей исходного и обработанного кадров [4]:

$$PSNR = 10 \log \left( \frac{I_{\max}^2}{MSE} \right), \quad (12)$$

где  $I_{\max}$  – максимальное значение яркости пикселя изображения;  $MSE = \frac{1}{n} \sum_{i=1}^n (I_{1,i} - I_{2,i})^2$ , – среднеквадратическая ошибка между двумя изображениями;  $n$  – количество пикселей изображения;  $I_1$  – яркости пикселей обработанного изображения;  $I_2$  – яркости пикселей исходного изображения.

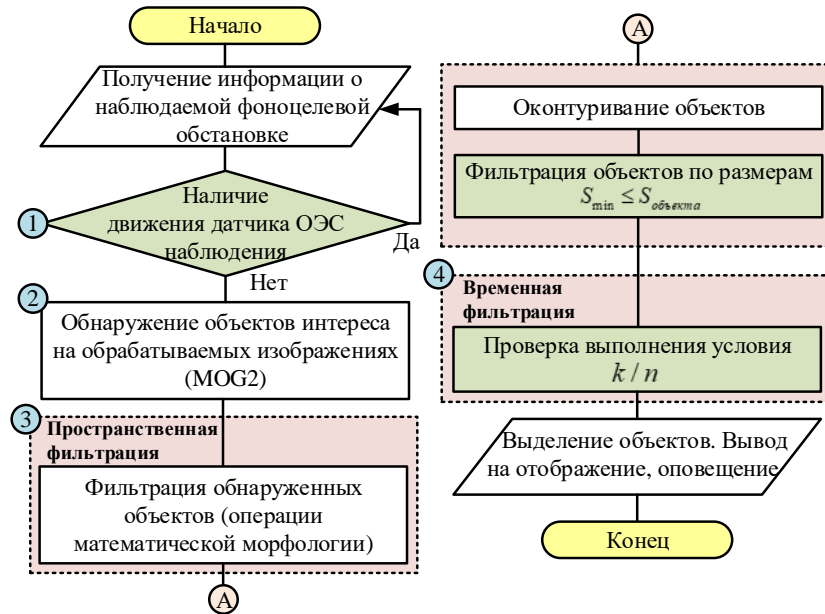


Рис. 1. Блок-схема алгоритма разработанного способа обработки изображений  
Fig. 1. Block diagram of the algorithm of the developed image processing method

Следует отметить, что критерий измерения пикового отношения сигнал/шум основывается на сравнении пары изображений до и после их обработки. В качестве изображения до обработки примем результат работы алгоритма обнаружения, основанного на методе вычитания фона, в котором модель фона представляется смесью гауссовых распределений. Изображением после обработки выступит результат работы разработанного способа обработки изображений и метода извлечения визуального фона (ViVe) [5].

Для характеристики качества решения задачи обнаружения рассчитаем вероятность ложной тревоги. Следует отметить, что на выходе обнаружителя, основанного на методе вычитания фона, формируется изображение, где пиксели представляют решение о наличии или отсутствии объекта интереса. Учитывая, что координаты искомого объекта известны, можно вычислить количество правильно и ложно обнаруженных объектов.

Рассматривая итоговое изображение как набор принятых независимых решений и исходя из формул частоты события и полной вероятности, вероятности ложной тревоги можно вычислить по формуле [6]

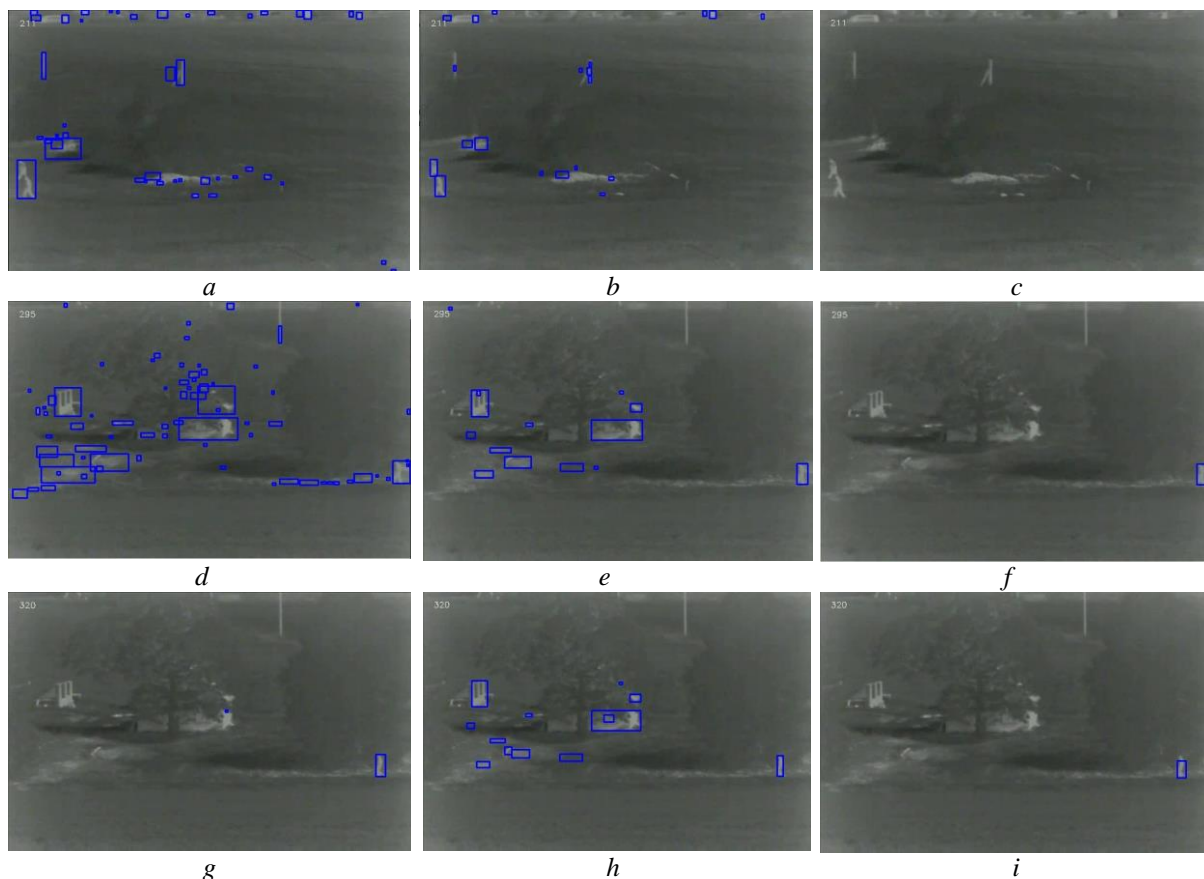
$$F_k = \frac{N_{\text{лож}_i}}{N_k}, \quad (13)$$

где  $N_{\text{лож}_i}$  – количество пикселей  $i$  ложно обнаруженных объектов;  $N_k$  – количество пикселей изображения  $k$ .

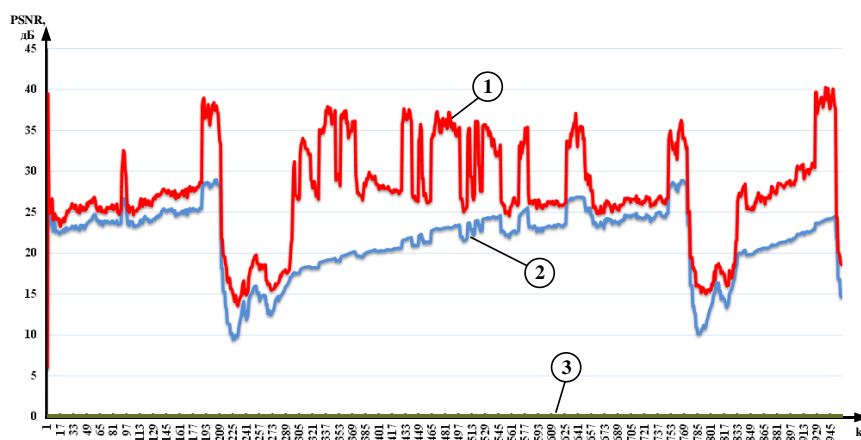
На рис. 2 представлены примеры обнаружения объектов интереса исследуемыми методами. Из рисунка видно, что при переносе поля зрения датчика (211-294 кадр) происходит неправильная сегментация пикселей обрабатываемых изображений (MOG2, Vibe), что, в свою очередь, приводит к уменьшению пикового отношения сигнал/шум (рис. 3) и повышению вероятности ложной тревоги (рис. 4).

Из рис. 3 видно, что для данного примера значение  $PSNR$  разработанного способа превышает значения  $PSNR$  метода MOG2 в среднем на 27 дБ, а  $PSNR$  метода Vibe – на 21 дБ.

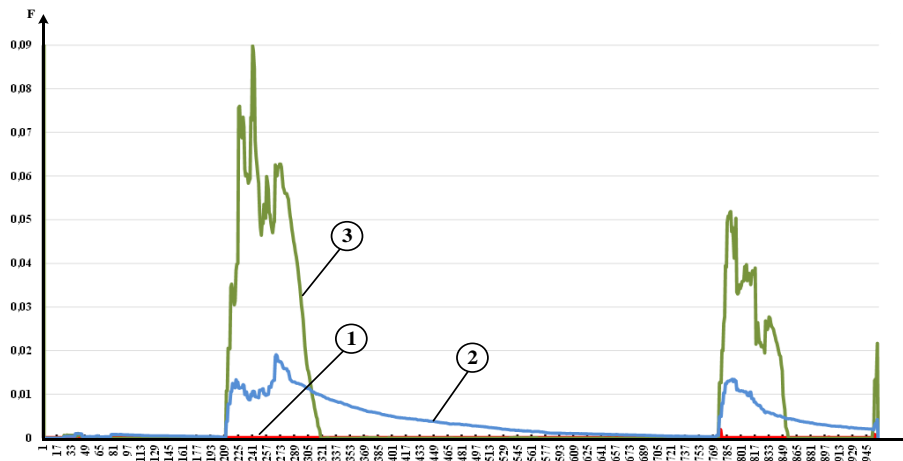
Из рис. 4 видно, что для данной фоноцелевой обстановки значение вероятности ложной тревоги при переносе поля зрения датчика опико-электронной системы наблюдения возрастает на несколько порядков (MOG2, Vibe) и уменьшается после прекращения движения по мере автоматической сегментации ложно выделенных в передний план областей обрабатываемого изображения. Среднее значение вероятности ложной тревоги для данного примера составляет: разработанный способ –  $1 \times 10^{-6}$ , Vibe –  $1 \times 10^{-3}$ , MOG2 –  $2 \times 10^{-3}$ .



**Рис. 2.** Примеры работы исследуемых методов: 211 кадр (*a* – MOG2, *b* – Vibe, *c* – разработанный способ); 295 кадр (*d* – MOG2, *e* – Vibe, *f* – разработанный способ); 320 кадр (*g* – MOG2, *h* – Vibe, *i* – разработанный способ)  
**Fig. 2.** Examples of the methods under study: 211 frame (*a* – MOG2, *b* – Vibe, *c* – developed method); 295 frame (*d* – MOG2, *e* – Vibe, *f* – developed method); 320 frame (*g* – MOG2, *h* – Vibe, *i* – developed method)



**Рис. 3.** Значения пикового отношения сигнал/шум (*PSNR*) на каждом кадре исследуемой видеопоследовательности (1 – разработанный способ, 2 – Vibe, 3 – MOG2)  
**Fig. 3.** Values of the peak signal- to- noise ratio (*PSNR*) on each frame of the video sequence under study (1 – developed method, 2 – Vibe, 3 – MOG2)



**Рис. 4.** Значения вероятности ложной тревоги на каждом кадре исследуемой видеопоследовательности (1 – разработанный способ, 2 – Vibe, 3 – MOG2)

**Fig. 4.** False alarm probability values for each frame of the video sequence under study (1 – developed method, 2 – Vibe, 3 – MOG2)

### Заключение

Таким образом, представленный способ обработки изображений применительно к задаче обнаружения движущихся объектов в оптико-электронных системах тепловизионного типа позволяет адаптировать к работе в сканирующих системах наблюдения алгоритмы обнаружения, построенные на методах вычитания фона, повысить не менее чем в 10 раз в зависимости от наблюдаемой фоноцелевой обстановки пиковое отношение сигнал/шум обрабатываемых изображений, тем самым уменьшить вероятность ложной тревоги не менее чем на один порядок.

### Список литературы

1. Zivkovic Z. Improved adaptive gaussian mixture model for background subtraction. *IEEE Proceedings of the 17<sup>th</sup> International Conference*. 2004;28-31.
2. Заливин А.Н., Балабанова Н.С. Обнаружение движущихся объектов методом вычитания фона с использованием смеси Гауссовых распределений. *Автоматизированные технологии и производства*. 2016;3:45-48.
3. Dougherty E.R. The dual representation of gray-scale morphological filters. *IEEE Trans. PAMI*. 1989.
4. Конюхов А.Л., Костевич А.Г., Курячий М.И. Критерии оценки отношения сигнал/шум в активно-импульсных телевизионно-вычислительных системах. *Доклады ТУРСУРа*. 2012;2:111-115.
5. Barnich O. ViBe: A universal background subtraction algorithm for video sequences. *IEEE Transactions on Image Processing*. 2011;20(6):1709-1724.
6. Вентцель Е.С., Овчаров Л.А. *Теория вероятностей и ее инженерные приложения*. Москва: Высшая школа; 2000:480.

### References

1. Zivkovic Z. Improved adaptive gaussian mixture model for background subtraction. *IEEE Proceedings of the 17th International Conference*. 2004;28-31.
2. Zalivin A.N., Balabanova N.S. [Detecting moving objects by subtracting the background using a mixture of Gaussian distributions]. *Avtomatizirovannyye tekhnologii i proizvodstva = Automated Technologies and Production*. 2016;3:45-48. (In Russ.)
3. Dougherty E.R. [The dual representation of gray-scale morphological filters]. *IEEE Trans. PAMI*. 1989.
4. Konyukhov A.L., Kostevich A.G., Kuryachiy M.I. [Criteria for evaluating the signal-to-noise ratio in active-pulse television and computer systems]. *Doklady TURSURA = Proceedings of Tomsk State University of Control Systems and Radioelectronics*. 2012;2:111-115. (In Russ.)



5. Barnich O. ViBe: A universal background subtraction algorithm for video sequences. *IEEE Transactions on Image Processing*. 2011;20(6):1709–1724.
6. Venttsel' Ye.S., Ovcharov L.A. [*Probability theory and its engineering applications*]. Moscow: Vysshaja Shkola; 2000:480. (In Russ.)

#### **Сведения об авторах**

Михнёнок Е.И., младший научный сотрудник НИЛ кафедры автоматизированных систем управления войсками Военной академии Республики Беларусь.

#### **Information about the authors**

Mikhnionok E.I., Junior Researcher of Research laboratory automated control system of troops of Military Academy of the Republic of Belarus.

#### **Адрес для корреспонденции**

220057, Республика Беларусь,  
г. Минск, пр. Независимости 220,  
Военная академия Республики Беларусь  
тел. +375-29-291-27-36  
e-mail: actosum@gmail.com  
Михнёнок Евгений Игоревич

#### **Address for correspondence**

220057, Republic of Belarus,  
Minsk, Nezavisimosti ave., 220,  
Military Academy of the Republic of Belarus  
tel. +375-29-291-27-36  
e-mail: actosum@gmail.com  
Mikhnionok Evgeny Igorevich