

ДЕТЕКТИРОВАНИЕ И КЛАССИФИКАЦИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ В ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ

Аннотация.

Актуальность и цели. Точность детектирования движущихся объектов на основе программного обеспечения с открытым кодом не превышает 60 %, это обусловлено существующей проблемой, первый аспект состоит в отсутствии алгоритмов и программного обеспечения для импортозамещения, второй аспект заключается в высокой стоимости создания отдельной системы видеонаблюдения с функциями детектирования и классификации движущихся объектов.

Материалы и методы. Обсуждается постановка и решение задачи детектирования и классификации движущихся объектов в видеопоследовательности с помощью вычисления значений статистических параметров смеси нормальных распределений. Программное обеспечение системы разработано с использованием библиотеки компьютерного зрения OPENCV.

Результаты и выводы. Предложена модель представления видеопоследовательности в задачах детектирования движения на основе трехмерного динамического информационного массива уровней яркости пикселей, отличающаяся обновляемым нулевым кадром маски движения. Разработан и реализован алгоритм детектирования движения на основе оценки плотности вероятности выпадения статичных пикселей в видеопоследовательности, отличающийся использованием смеси нормальных распределений уровней яркости элементов информационного поля. Предложен комбинированный подход к детектированию и классификации движущихся объектов в видеопоследовательности. В соответствии с результатами экспериментов точность обнаружения объектов на основе комбинированного подхода увеличилась на 15 % от показателя точности простого детектора движений и на 5 % относительно точности обычного детектора объектов на основе каскадных классификаторов.

Ключевые слова: детектирование и классификация движущихся объектов в видеопоследовательности, видеоанализ, видеонаблюдение, компьютерное зрение.

К. В. Shepelev

DETECTION AND CLASSIFICATION OF MOVING OBJECTS IN A VIDEO SEQUENCE

Abstract.

Background. The current accuracy of moving objects detection via open source software does not exceed 60 %. This problem has several aspects. The first aspect is a lack of algorithms and software for import substitution, the second is a high cost of creating a separate video surveillance system with moving object detecting and classifying functions.

Materials and methods. For the purpose of detecting and classifying moving objects in a video sequence the author uses a motion detection algorithm by calculating static pixel values over a mixture of normal distributions and an algorithm for classifying images. The software implementation of the system is performed by means of OPENCV computer vision library.

Results and conclusions. The work suggests a model of video sequence representation in motion detection problems based on a three-dimensional dynamic information array of pixel brightness levels. It is distinguished by an updated zero motion mask frame and allows increasing the accuracy of object detection by dynamically refining the static part of a video sequence using a low-pass filter. The author has developed and implemented a motion detection algorithm based on estimating the probability density of static pixels falling out in a video sequence. The algorithm is distinguished by the use of a mix of normal distributions of brightness levels of information field elements. A combined method for detecting and classifying moving objects in a video sequence is offered. The accuracy of detection of objects based on the combined method increases by 15 % compared to the accuracy of a simple motion detector and by 5 % relative to the accuracy of the Haar cascade classification

Key words: detection and classification of moving objects in a video sequences, video analysis, video surveillance, computer vision.

Введение

Одними из областей, где автоматизация осуществляется не на всех уровнях процесса и существуют неизученные вопросы внедрения компьютерных технологий, – это видеонаблюдение, системы слежения и защиты. Существующая точность детектирования движущихся объектов на основе программного обеспечения с открытым кодом не превышает 60 % [1], таким образом, существует несколько аспектов проблемы детектирования движущихся объектов и их классификации в системах видеонаблюдения. Первый аспект проблемы состоит в отсутствии алгоритмов и программного обеспечения для импортозамещения, второй аспект заключается в высокой стоимости создания отдельной системы видеонаблюдения с функциями детектирования и классификации движущихся объектов. Таким образом, создание программно-математического комплекса с алгоритмами детектирования движения и классификатора объектов в существующих системах видеонаблюдения может решить названную выше проблему детектирования и классификации движущихся объектов на видеопоследовательности.

Для реализации комбинированного подхода детектирования и классификации движущихся объектов в видеопоследовательности в виде системы необходимо решить следующие задачи:

Разработать модели представления структуры видеопоследовательности в виде трехмерного динамического информационного массива уровней яркости элементов информационного поля;

Разработать алгоритм детектирования движения;

Разработать комбинированный метод детектирования и классификации движущихся объектов на видеопоследовательности.

Анализ способов детектирования движущихся объектов

Существующие методы детектирования движущихся объектов на видеопоследовательности можно разделить на две группы: распознавание с учетом яркости и распознавание с учетом характеристик. Кроме того, можно выделить два типа детекторов, пользующихся наибольшим спросом на рынке систем подсчета посетителей – инфракрасные счетчики и системы, основан-

ные на видеонаблюдении. Сравнительный анализ существующих систем детектирования убеждает в преимуществе систем видеонаблюдения: помимо реализации статистических функций, системы обеспечивают безопасность на объекте наблюдения. При этом имеет место очевидное преимущество в обработке данных системами видеонаблюдения. Инфракрасные детекторы представляют собой узконаправленные элементарные устройства, способные выполнять только функции подсчета посетителей. На основе проведенного анализа структуры видеопоследовательности [2] можно сделать вывод, что детектирование движения на видеопоследовательности может быть зафиксировано несколькими способами, которые делятся на две основные категории: разделение переднего и заднего плана на изображении и трекинг образов на изображении.

Предметом теоретического исследования являются анализ методов детектирования движущихся объектов, оценка точности работы алгоритмов на основе сравнения математических аппаратов соответствующих методов.

Модель представления видеопоследовательности.

В процессе обработки кадров видеопоследовательности целесообразно рассматривать изображение (кадр) как двумерный массив, каждый элемент которого будет равен яркости свечения соответствующего пикселя на изображении. Для упрощения реализации предложено привести изображение формата RGB к оттенкам серого, для этого используется преобразование с использованием значений непосредственно из рекомендации ITU-R BT.601 последовательно для каждого пикселя. Конкретные значения яркости свечения каждого пикселя в оттенках серого упрощают представление пикселя, что позволяет перевести информационную модель трехмерного массива яркости свечения пикселей RGB в двумерный массив полутонового изображения, снизив пространственную размерность массива пикселей видеопоследовательности с 9 до 3 измерений.

В предлагаемой модели представления видеопоследовательности в качестве трехмерного динамического информационного массива уровней яркости элементов информационного поля рассматривается трехмерный массив, где три пространственных оси будут соответствовать ширине, высоте и номеру кадра:

$$Y_k = [y_{i,j,k}], y_{i,j,k} \in [0..255], i \in [0..w], j \in [0..h], k \in [0..n],$$

где Y_k – трехмерный динамический информационный массив уровней яркости элементов информационного поля; $y_{i,j,k}$ – значение яркости свечения пикселя $[i, j]$ в кадре k из n , который имеет ширину w пикселей и высоту h пикселей.

Уникальность модели состоит в том, что в данном динамическом массиве нулевой элемент по оси k представляет собой среднестатистический кадр неподвижного (заднего) плана $y_{i,j,0}$, который составляется при обработке видеопоследовательности выбранным методом детектирования движения. Временная ось k содержит метки, определяемые шагом дискретизации Δt в зависимости от кадровой частоты видеопоследовательности. Для каждого

кадра определяется модель переднего плана и заднего плана как совокупность соответственно сильно и слабо изменяющихся по своей яркости свечения пикселей за промежуток времени t . На основе анализа структуры видеопоследовательности было определено, что для обработки видео с точки зрения методов информационной обработки данных допустимо применение методов математической статистики.

Алгоритм детектирования движущихся объектов

Для реализации определения движущихся объектов был использован аппарат математической статистики, при этом предполагается, что выборка данных осуществляется с первых n кадров видео, когда движения объектов видеопоследовательности не происходит, т.е. на первых кадрах изображение статично. Данные кадры берутся как последовательность для предварительного анализа с целью создания модели заднего плана с помощью нормально распределенной случайной величины яркости пикселя во временной плоскости. Предварительно из массива кадров n необходимо выделить пиксели, значения яркости свечения которых на всех кадрах постоянны. Данные пиксели будут принадлежать к заднему плану, так как не изменяются и образуют подмножество элементов матрицы заднего плана.

Для пикселей с координатами $[i, j, k]$, имеющими отклонения, определяется среднее выборочное значение яркости (математическое ожидание):

$$\mu_{i,j,0} = \frac{1}{n} \sum_{k=1}^n y_{i,j,k}, \quad (1)$$

где $\mu_{i,j,0}$ – математическое ожидание (среднее значение яркости) пикселя i, j нулевого кадра; $y_{i,j,k}$ – значение яркости свечения пикселя i, j в кадре k , в битах; n – общее количество кадров для анализа.

Последовательное применение формулы (1) к каждому ряду пикселей видеопоследовательности позволяет определить разброс яркости пикселя:

$$\delta_{i,j,0}^2 = \frac{1}{n-1} \sum_{k=1}^n (y_{i,j,k} - \mu_{i,j,0})^2, \quad (2)$$

где $\delta_{i,j,0}^2$ – дисперсия яркости пикселя i, j в матрице заднего плана.

Формулы (1) и (2) дают возможность определить плотность вероятности $\rho_{i,j,k}$ принадлежности пикселя с координатами $[i, j]$ в кадре k к заднему фону в зависимости от изменяемости его значений:

$$\rho_{i,j,k}(y_{i,j,k} | \mu_{i,j,0}, \delta_{i,j,0}) = \frac{1}{\delta_{i,j,0} \sqrt{2\pi}} \exp \left(-\frac{(y_{i,j,k} - \mu_{i,j,0})^2}{2\delta_{i,j,0}^2} \right). \quad (3)$$

Для определения движения из формулы (3) необходимо установить отклонения последовательно для соответствующих значений всех пикселей изображения, определяя модель заднего плана следующим образом:

$$\frac{|\mu_{i,j,0} - y_{i,j,k}|}{\delta_{i,j,0}} \leq e, \quad (4)$$

где $k \in [1...n]$, а величина e влияет на общую чувствительность алгоритма и является адаптивным вероятностным порогом (как правило, используется $e = 3$) [3].

Определение принадлежности конкретного пикселя в конкретном кадре будет происходить по критерию, что определит наиболее вероятное значение или диапазон допустимых значений яркости текущего пикселя в результирующей матрице фона. Формула (4) определяет принадлежность пикселей к заднему или переднему плану, формируя матрицу заднего плана:

$$F(y_{i,j,k}) = \begin{cases} 1, & \text{если результат (4) удовлетворяет условию,} \\ 0, & \text{если результат (4) не удовлетворяет условию.} \end{cases}$$

Если изменение значений яркости пикселя в кадрах $1...n$ удовлетворяет условию формулы (4), то он будет принадлежать к заднему плану, т.е. по статистике и выборке можно сделать вывод, что в заданном наборе кадров данный пиксель не изменяется, и в целом элемент кадра, к которому принадлежит пиксель, неподвижен. Значения яркости пикселей нулевого кадра последовательности определяются нормальным распределением величины яркости свечения пикселя в кадрах $1...n$.

Анализ метода детектирования движения на основе смеси нормальных распределений позволил определить основные отличия от метода простого вероятностного распределения – самообучение и обновление статистических данных относительно каждого из пикселей.

Данный метод использует математическую модель смеси нормальных распределений вероятности существования статичного пикселя применимо к трехмерному массиву яркости свечения пикселей видеопоследовательности, и по поступлению нового кадра (третьей компоненты каждой ячейки пикселя) обновляет модель заднего плана:

$$\Psi_{i,j,k} = \sum_{k=1}^n \eta_{i,j,k} N(k, \mu_{i,j,0}, \delta_{i,j,0}^2), \quad (5)$$

где $\Psi_{i,j,k}$ – яркость свечения пикселя заднего плана; $\eta_{i,j,k}$ – вес события на кадре k , N – область заднего плана неподвижных пикселей на кадре k .

Параметр $\eta_{i,j,k}$ будет определять вес события k , т.е. насколько часто данное явление в отношении конкретного пикселя i, j попадало в поле зрения камеры. Параметр $\delta_{i,j}^2$ будет являться разбросом значений яркости пикселей: как правило, данный параметр задается по умолчанию и вручную, в зависимости от требуемых результатов работы алгоритма и подбирается экспериментально.

Изначально модель заднего плана будет равна первому кадру видеопоследовательности. Затем, по мере поступления следующий кадров, модель будет обновляться последовательно и попиксельно с каждым из кадров, с по-

мощью низкочастотного фильтра рекурсивного сглаживания, что также является особенностью метода:

$$\mu_{i,j,k} = (1-\varphi) \cdot y_{i,j,k-1} + \varphi \cdot \mu_{i,j,0}, \varphi \in (0;1), \quad (6)$$

где φ – коэффициент обучаемости алгоритма, прямо пропорциональный скорости обучения.

Определив формулой (6) математическое ожидание ряда пикселей, можно определить дисперсию, учитывая соответствующий коэффициент обучаемости:

$$\delta_{i,j,k} = \sqrt{(1-\varepsilon) \cdot (y_{i,j,k} - \mu_{i,j,k})^2 + \varepsilon \cdot \left(\delta_{i,j,k-1}^2 - (\mu_{i,j,k} - \mu_{i,j,0})^2 \right)}, \quad (7)$$

где ε – коэффициент обучаемости алгоритма, прямо пропорциональный скорости обучения, $\varepsilon \in (0;1)$.

Коэффициенты, определяемые формулами (6) и (7), задаются оператором системы в зависимости от оценки подвижности элементов сцены кадра, освещения и технических параметров оборудования. В дальнейшем возможна автоматизация данного процесса путем внедрения алгоритма автоматического машинного обучения, основанного на нейронных сетях. Существует также в этом методе и обновление весов процессов, происходящих в видеопоследовательности:

$$\eta_{i,j,k} = (1-\beta) \cdot \eta_{i,j,k-1} + \beta \cdot A(k), \quad (8)$$

где β – коэффициент скорости изменения веса. Параметр $A(k) = 1$, если $k = n$, и $A(k) = 0$ если $k \neq n$.

В матрицу заднего плана попадут распределения, которые удовлетворяют следующему условию:

$$B(k) = \arg \min_n \left\{ \sum_{k=1}^n \omega_k^n > T \right\}, T \in (-2,5\delta; 2,5\delta).$$

Алгоритм определения движущихся объектов в видеопоследовательности выглядит следующим образом (рис. 2).

После проведенных вычислений в соответствии с алгоритмом можно сделать вывод о принадлежности пикселя к заднему или переднему планам, если значение веса превышает значение заранее определенного порога, то пиксель не будет принадлежать движущемуся объекту, т.е. будет классифицирован как пиксель заднего плана.

Основное преимущество алгоритма, изображенного на рис. 1, заключается в том, что задний план может быть динамическим, и со временем алгоритм приспосабливается к повторяющимся движениям.

Комбинированный подход к детектированию движения и классификации объектов

Для решения задачи детектирования движения и классификации объектов обоснованно применение алгоритма, совмещающего в себе алгоритм де-

тектирования движения и каскадной классификации объектов [4]. Пусть дано некое множество объектов, для каждого из которых необходимо максимально точно определить, в какую конкретную и заранее определенную категорию он попадает, т.е. классифицировать, разбивая данное множество на подмножества – классы объектов. При данной классификации были заранее определены некоторые данные: параметры классов, их особые и уникальные свойства, позволяющие отличить данные объекты от объектов другого типа (классов), и объекты, чья принадлежность к конкретному классу неизвестна. Максимально правильное распределение объектов по классам определяет наибольшую точность подсчета.

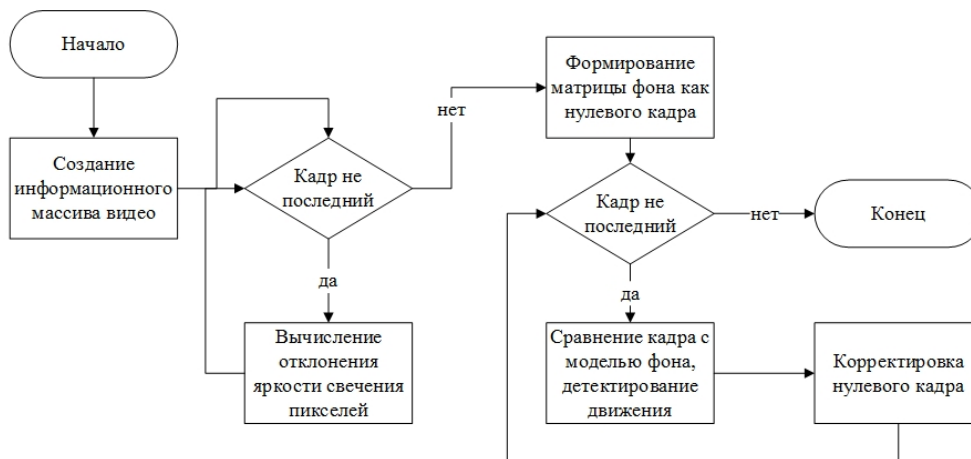


Рис. 1. Алгоритм детектирования движения на видеопоследовательности на основе смеси нормальных распределений

В качестве механизма распределения по классам использованы вэйлеты Хаара. Распознавание образов достигается за счет признаков Хаара на основе каскадного классификатора. В ходе анализа данного подхода к распознаванию образов при наличии цели распознавания посетителей было определено, что удобнее всего сосредоточиться на поиске лиц, определяя направление движения человека и точно классифицируя его как идущего к камере, и, соответственно, как входящего/выходящего. Поиск объектов на изображении в модели системы построен по принципу сканирующего окна. Для алгоритма был использован классификатор объектов, который использует деревья принятия решений, также известные как деревья классификации или деревья регрессии [5].

На основе двух алгоритмов детектирования движения и классификации объектов в видеопоследовательности был создан комбинированный подход, представленный на рис. 2.

Из рис. 2 следует, что в итоге при работе классификатора все объекты распределяются на три класса: «Не движущийся объект», «Движущийся Не человек» и «Движущийся Человек». Лишь класс «Движущийся Человек» отвечает условиям успешного детектирования и классификации посетителя.

Рассмотрены основные аспекты и методы реализации изучаемых алгоритмов. После успешной реализации программное обеспечение прошло испытание серией экспериментов в Пензенском государственном университете

и на улицах г. Пензы, решая задачу подсчета посетителей. Всего в ходе эксперимента отснято более 1000 ч видео, что доказало стабильность работы программного обеспечения [4]. В качестве единицы времени в результатах взят 20-минутный отрезок видеопоследовательности. Результаты работы детекторов были преобразованы в статистические таблицы, которые послужили материалом для анализа эффективности работы алгоритмов. Для сравнения конкретных результатов видеопоследовательности были просмотрены автором и просчитаны «вручную», давая 100 % эталонную точность. По результатам статистического сравнения эффективности работы алгоритмов были сделаны выводы о точности работы и целесообразности применения данных методов [6]. При анализе таблицы результатов работы алгоритмов выявлена средне-статистическая точность алгоритмов. Для программы на основе алгоритма детектирования движения с использованием смеси нормальных распределений точность подсчета составляет 77 %, а для алгоритма классификации объектов в видеопоследовательности точность равна 94 %. Для анализа точности детектирования и классификации движущихся объектов для каждого алгоритма были построены графики Гаусса.

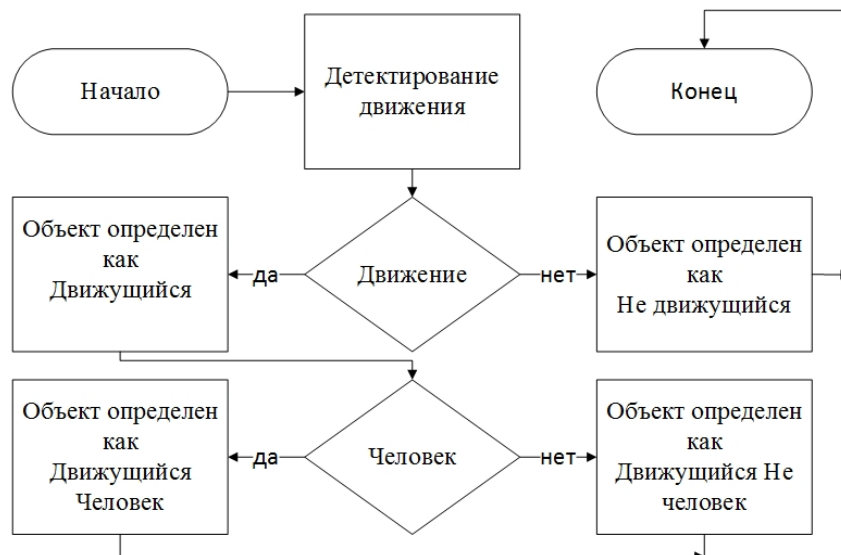


Рис. 2. Комбинированный подход к детектированию и классификации движущихся объектов

На обобщенном графике функций на рис. 3 визуально можно определить гораздо большую плотность вероятности выпадения точного значения относительно реального.

Из графика, представленного на рис. 3 следует, что гауссова функция для комбинированного метода имеет на интервале от $-2,5\delta$ до $2,5\delta$ большую плотность вероятности выпадения точного значения при работе конкретного алгоритма детектора, и в сравнительном анализе результатов исследований доказывает превосходство [7] комбинированного метода детектирования и классификации движущихся объектов над методом обычного метода детектирования [8] в отношении точности подсчета относительно реальных показателей, подсчитанных «вручную».

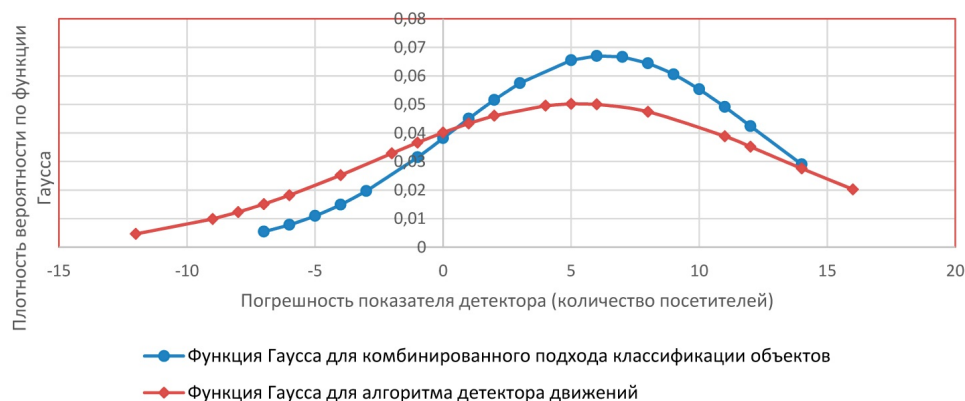


Рис. 3. График распределений Гаусса для результатов двух вариантов эксперимента

Заключение

По итогам исследования создан комбинированный подход детектирования и классификации движущихся объектов в видеопоследовательности, отличительной особенностью которого является комбинация нового алгоритма детектирования движения на основе представления структуры видеопоследовательности в виде трехмерного информационного массива яркости пикселей с нулевым обновляемым кадром и каскадного классификатора объектов, основанного на вейвлетах Хаара. Метод позволяет увеличить точность классификации движущихся объектов за счет уточнения области поиска с обнаруженным движением для классификатора, основанного на вейвлетах Хаара. Точность обнаружения объектов на основе комбинированного подхода увеличилась на 15 % [6] от показателя точности простого детектора движений и на 5 % [6] относительно точности каскадной классификации Хаара в соответствии с результатами экспериментов.

Библиографический список

1. **Шепелев, К. В.** Расширение существующих систем видеонаблюдения в системах безопасности до маркетинго-статистических систем анализа клиентопотока / К. В. Шепелев // SCIENCE TIME. Международный электронный научный журнал. – 2016. – № 4 – С. 241.
2. **Bradsky, G.** Learning OpenCV / G. Bradsky, A. Kaehler // LA, O'Reilly. – Cambridge, 2008. – 140 с.
3. **De Maesschalck, R.** The Mahalanobis distance / R. De Maesschalck, D. Jouan-Rimbaud, D. L. Massart // Chemometrics and Intelligent Laboratory Systems. – 2000. – Vol. 50 (1). – С. 1–18.
4. **Налимов, В. В.** Статистические методы планирования экстремальных экспериментов / В. В. Налимов, Н. А. Чернова. – М. : Наука, 1965. – С. 64.
5. **Шепелев, К. В.** Комбинированный метод идентификации посетителей в системах видеонаблюдения / Б. Д. Шашков, К. В. Шепелев // Естественные и технические науки. – 2017. – № 4 – С. 243.
6. **Шепелев, К. В.** Экспериментальное сравнение эффективности использования алгоритмов межкадровой разницы и нормального распределения в задачах детектирования человека / П. П. Макарычев, К. В. Шепелев, Д. С. Шепелева // OPEN INNOVATION : сб. ст. Междунар. науч.-практ. конф. – Пенза : Наука и Просвещение, 2017. – С. 298.

7. **Макарычев, П. П.** Анализ триангуляции подвижных объектов методом статистических испытаний / П. П. Макарычев // Новые информационные технологии и системы : сб. науч. ст. XII Междунар. науч.-техн. конф. – Пенза, 2015. – С. 24–26.
8. **Попов, Д. И.** Анализ цифровых систем обработки сигналов / Д. И. Попов // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2016. – № 2 (38). – С. 83–92.

References

1. Shepelev K. V. *SCIENCE TIME. Mezhdunarodnyy elektronnyy nauchnyy zhurnal* [International scientific e-journal]. 2016, no. 4, p. 241.
2. Bradsky G., Kaehler A. *LA, O'Reilly*. Cambridge, 2008, 140 p.
3. De Maesschalck R., Jouan-Rimbaud D., Massart D. L. *Chemometrics and Intelligent Laboratory Systems*. 2000, vol. 50 (1), pp. 1–18.
4. Nalimov V. V., Chernova N. A. *Statisticheskie metody planirovaniya ekstremal'nykh eksperimentov* [Statistical methods of extreme experiments planning]. Moscow: Nauka, 1965, p. 64.
5. Shepelev K. V., Shashkov B. D. *Estestvennye i tekhnicheskie nauki* [Natural and engineering sciences]. 2017, no. 4, p. 243.
6. Shepelev K. V., Makarychev P. P., Shepeleva D. S. *OPEN INNOVATION: sb. st. Mezhdunar. nauch.-prakt. konf.* [proceedings of an International scientific and practical conference]. Penza: Nauka i Prosveshchenie, 2017, p. 298.
7. Makarychev P. P. *Novye informatsionnye tekhnologii i sistemy: sb. nauch. st. XII Mezhdunar. nauch.-tekhn. konf.* [New information technologies and systems: proceedings of XII International scientific and technical conference]. Penza, 2015, pp. 24–26.
8. Popov D. I. *Izvestiya vysshikh uchebnykh zavedeniy. Povolzhskiy region. Tekhnicheskie nauki* [University proceedings. Volga region. Engineering sciences]. 2016, no. 2 (38), pp. 83–92.

Шепелев Кирилл Валерьевич
магистрант, Пензенский
государственный университет (Россия,
г. Пенза, Красная, 40)
E-mail: shepelevkir@gmail.com

Shepelev Kirill Valer'evich
Master's degree student, Penza
State University (40 Krasnaya street,
Penza, Russia)

УДК 004.932.2

Шепелев, К. В.

Детектирование и классификация движущихся объектов в видео-последовательности / К. В. Шепелев // Известия высших учебных заведений. Поволжский регион. Технические науки. – 2017. – № 4 (44). – С. 45–54. DOI 10.21685/2072-3059-2017-4-5