

## АПРОБАЦИЯ АЛГОРИТМА ОБНАРУЖЕНИЯ ОБЪЕКТА НА ЦИФРОВОМ ИЗОБРАЖЕНИИ

А.Т. Левинский

adam.levinskiy@yandex.ru

SPIN-код: 2301-6960

И.Д. Родионов

ir@drp.ru

SPIN-код: 2598-2061

С.В. Агальцев

sosimba-sosimbovich@mail.ru

SPIN-код: 5570-9322

Д.В. Тимофеев

dm.vl.timofeev@gmail.com

SPIN-код: 9258-7883

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

---

### Аннотация

*Предметом изучения является алгоритм обнаружения, а также исследование и модернизация алгоритма обнаружения нескольких объектов на цифровом изображении. Поставленная цель достигается благодаря применению масштабирования к исследуемому изображению с помощью нахождения точек максимума, в которых изображение имеет наивысшее сходство с эталонным изображением. Проведена работа по улучшению алгоритма распознавания объектов на изображении. Добавлена возможность дополнительно задавать в программе размеры изображений, коэффициент сходства между изображениями и коэффициент масштабирования.*

### Ключевые слова

*Алгоритм, обнаружение, объект, цифровое изображение, коэффициент, обработка изображений, фрагмент, ядра, дескриптор*

Поступила в редакцию 12.01.2018

© МГТУ им. Н.Э. Баумана, 2018

---

**Введение.** Анализ визуальных объектов в изображениях является очень важным компонентом систем компьютерного зрения, которые осуществляют распознавание объектов, поиск изображений, регистрацию изображений и многое другое. Области использования таких систем довольно разнообразны и включают в себя видеонаблюдение, видеокриминалистику, а также анализ медицинских изображений для компьютерной диагностики. В последнее время проблема распознавания объектов привлекла большое внимание по причине растущего спроса на разработку компьютерных систем.

**Цифровая обработка изображений.** Изображение можно определить как двумерную функцию  $f(x, y)$ , где  $x$  и  $y$  — координаты на плоскости, а значение функции  $f$  в любой точке, задаваемое двумя координатами  $(x, y)$ , называется интенсивностью. Если величины  $x$  и  $y$  принимают конечное число дискретных значений, то говорят о цифровом изображении. Цифровую обработку изображений выполняют с помощью вычислительных машин (компьютеров). Цифровое изображение состоит из конечного числа элементов, каждый из которых

расположен в конкретном месте и принимает определенное значение. Эти элементы называют элементами изображения, или пикселями [1].

**Обзор выбранного метода.** В данной работе в качестве дескрипторов использованы локальные ядра регрессии [2], которые очень хорошо захватывают локальную структуру данных даже при наличии значительных искажений [3]. Предложен новый подход к проблеме обнаружения объекта на изображении с помощью непараметрического классификатора ближайших соседей. Происхождение и использование этих ядер является предметом более ранней работы по адаптивной регрессии ядра для обработки и реконструкции изображений [4]. Фундаментальной составляющей метода управления ядерной регрессией является расчет локального управления ядра, при котором локальное сходство пикселя с соседями устанавливается геометрически.

Обозначив исследуемое изображение  $T$ , а эталонное  $Q$ , можно определить плотный набор локальных ядер управления от каждого из них. Эти плотно вычисленные дескрипторы высокоинформативны, но вместе они стремятся быть избыточными. Поэтому необходимо вывести функцию, которая уменьшает размерность (метод главных компонент (англ. *principal component analysis*, PCA) к этим результирующим массивам, чтобы сохранить только основные характеристики локальных управляющих ядер. Как правило,  $T$  больше, чем  $Q$ , следовательно, исследуемое изображение  $T$  необходимо разделить на набор перекрывающихся фрагментов изображения, которые имеют тот же размер, что и  $Q$ , и приписать классу каждый фрагмент  $T_i$ . В качестве примера рассмотрим рис. 1.

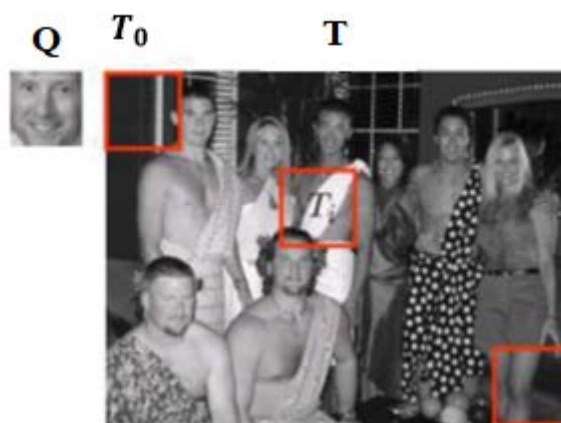


Рис. 1. Обнаружение объекта  $Q$  на изображении  $T$

Функции из форм  $Q$  и  $T$  образуют матрицы  $F_Q$  и  $F_T$ . Для поиска совпадений сравниваются характеристические матрицы  $F_T$  и  $F_Q$  от  $i$ -го фрагмента  $T$ . Если на исследуемом изображении имеется несколько объектов, осуществляется нахождение только главного максимума. Совсем недавно британские ученые Д.А Шехтман и Ирани представили систему сопоставления, основанную на дескрипторе так называемого локального самоподобия [5]. Отметим, что эта мера

локального самоподобия является частным случаем локального управления ядра и также связана с рядом других локальных адаптивных показателей данных, таких как оптимальная пространственная адаптация [6] и нелокальные средства [7], которые использовались для регрессии в обработке изображений. Хотя локальные дескрипторы смоделированы как функции простой суммы квадратов разностей между фрагментом центрального изображения и окружающими фрагментами изображений, ядра локальной регрессии имеют более сложные механизмы. Надежность метода заключается в получении локальной структуры изображений даже при наличии неопределенности данных, таких как шум и размытость.

Обзор предлагаемой системы показан на рис. 2. Первый этап состоит из вычисления весовых матриц  $W_Q$ ,  $W_T$  и получения основных матриц признаков  $F_Q$ ,  $F_T$ . На втором этапе осуществляется сравнение характеристических матриц  $F_T$  и  $F_Q$  с использованием коэффициента корреляции. Окончательный результат получают после проведения тестов значимости с последующим подавлением ложных максимумов (от англ. *a non-maxima suppression*) [8]. Ложные максимумы — пики в определенной области на изображении, которые в целом не характеризуют сходство между интересующими пользователя объектами (рис. 3).

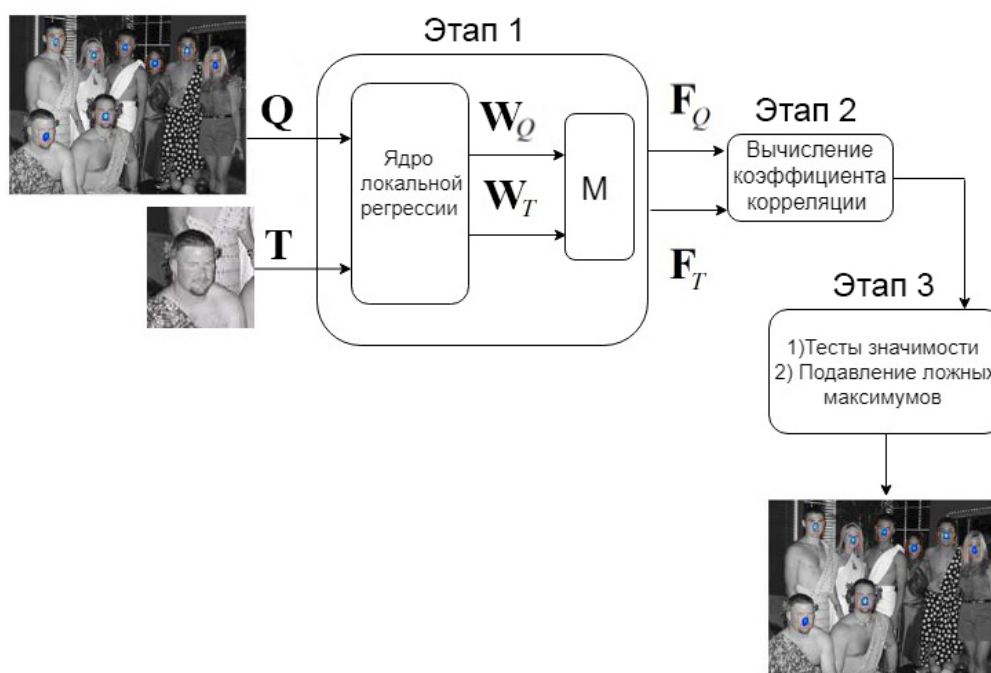


Рис. 2. Обзор системы (разделен на три этапа)

**Извлечение функций из дескрипторов локального управления ядра.** Ключевая идея локального управления ядра заключается в надежном получении локальной структуры изображений путем анализа радиометрических (пиксельных)

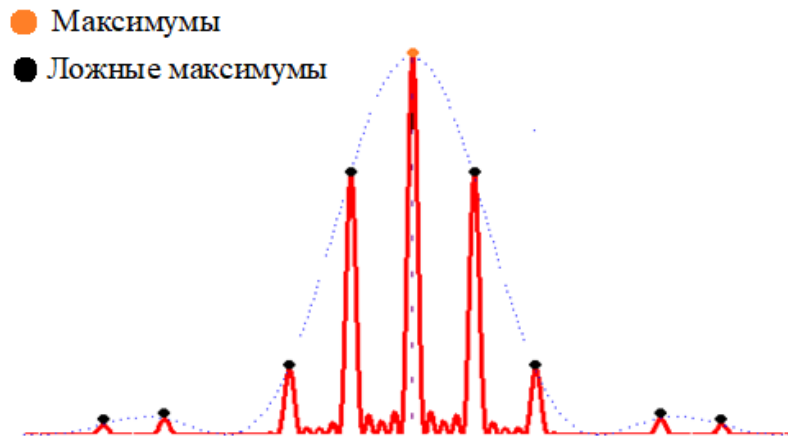


Рис. 3. Нахождение максимумов

значений) различий на основе оцененных градиентов и использования этой структурной информации для определения формы и размера канонического ядра. Локальное ядро  $K$  моделируется как радиально-симметричная функция

$$K(x_i - x; H_l) = \frac{K(H_l^{-1}(x_i - x))}{\det(H_l)}, \quad l = 1, \dots, p^2.$$

Здесь  $x_l = [x_1, x_2]^T$  — пространственные координаты;  $p^2$  — число пикселей в локальном окне ( $P \times P$ );  $H_l$  — так называемая управляющая матрица, которая определяется как

$$H_l = hC_l^{-0.5} \in R^{(2 \times 2)},$$

где  $h$  — параметр глобального сглаживания, а матрица  $C_l$  является ковариационной матрицей, оцененной из набора пространственного  $(x_1, x_2)$  градиента векторов в локальном окне анализа вокруг позиции выборки  $x$ . Управляющая матрица  $H_l$  изменяет форму и размер локального ядра способом, который кодирует локальные геометрические структуры, присутствующие в изображении [9] (рис. 4).

**Коэффициент корреляции как мера сходства.** Следующим этапом в предлагаемой структуре является правило принятия решения, основанное на измерениях «расстояния» между вычисленными характеристиками  $F_Q$ ,  $F_T$ . Корреляционные метрики хорошо подходят для задач классификации и подпространства. Основываясь на эффективности меры сходства на основе корреляции, вводим коэффициент корреляции для матричного случая и исследуем эту идею более подробно. В общем случае корреляция указывает на линейную зависимость между двумя случайными величинами. Существуют по крайней мере

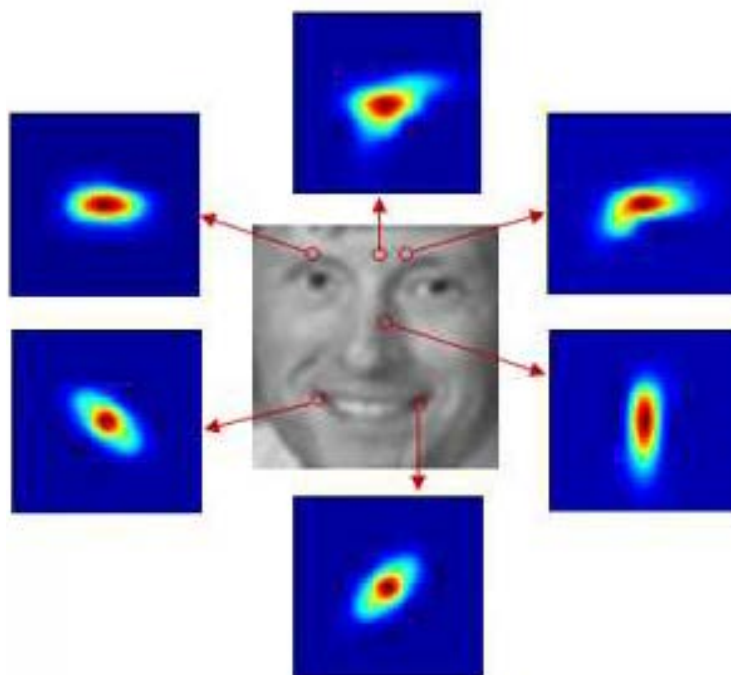


Рис. 4. Пример локального управления ядра в различных областях изображения

тринадцать различных способов использования корреляции [10]. Рассмотрим два основных типа корреляции: коэффициент корреляции Пирсона, который является известным стандартным коэффициентом корреляции, и сходство косинуса. Заметим, что сходство косинусов совпадает с корреляцией Пирсона, когда каждый вектор центрирован с нулевым средним значением. Корреляция Пирсона менее дифференцируема, чем косинусоидальное сходство, поскольку центрированные значения менее информативны. Косинусоидальное сходство определяется как скалярное произведение двух нормированных векторов:

$$p(f_Q, f_{T_i}) = \left\langle \frac{f_Q}{\|f_Q\|}, \frac{f_{T_i}}{\|f_{T_i}\|} \right\rangle = \left\langle \frac{f_Q f_{T_i}}{\|f_Q\| \|f_{T_i}\|} \right\rangle = \cos \theta_i \in [-1, 1],$$

Следующим этапом является создание «карты сходства», которая будет изображением со значениями, показывающее вероятность схожести между  $Q$  и  $T$ . Примем  $p_i^2 \in [0, 1]$ , показывающую долю дисперсии между двумя наборами признаков, в отличие от  $p_i$ , которая указывает на линейную зависимость между двумя характеристическими матрицами  $F_Q$ ,  $F_{T_i}$ . На этом этапе можно использовать  $p_i$  непосредственно как меру сходства между двумя наборами признаков. Однако общая интерпретация дисперсии  $p_i^2$  имеет ряд преимуществ. Для итоговой тестовой статистики, содержащей значения в карте сходства, исполь-

зуюем долю дисперсии  $p_i^2$  и долю «остаточной» дисперсии  $(1-p_i^2)$ . Рассмотрим более подробно вычисление  $RM$  с использованием функции отображения  $f$ :

$$RM: f(p_i) = \frac{p_i^2}{1-p_i^2}.$$

На рис. 5 приведены примеры карты сходства  $RM$  на основе  $|p_i|$  и  $f(p_i)$ . Красный цвет соответствует более полному сходству.

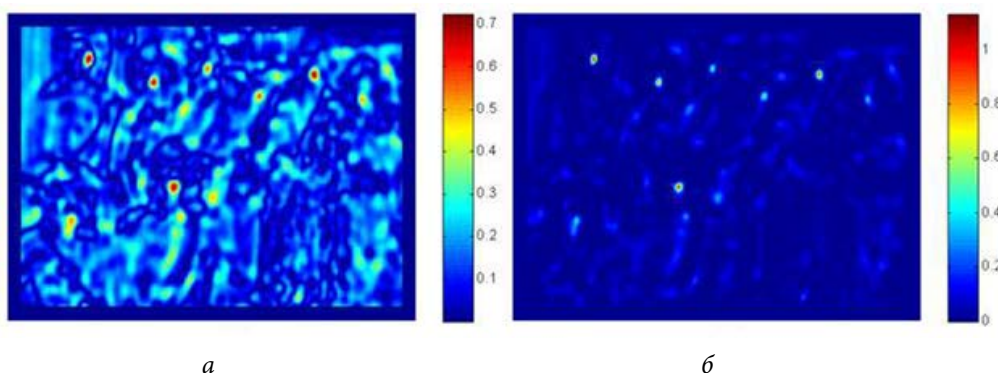


Рис. 5. Карта сходства:

$a$  — на основе  $|p_i|$ ;  $b$  — на основе  $f(p_i)$

Заметим, что  $Q$  и  $T$  — это те же самые матрицы, показанные на рис. 1. Видно, что карта сходства, созданная из  $f(p_i)$ , обеспечивает лучший контраст и динамический диапазон при условии  $f(p_i) \in [0, \infty]$ .

**Проверка значимости и подавление ложных максимумов.** Если задача состоит в том, чтобы найти наиболее подходящий фрагмент  $T_i$  исследуемого изображения  $Q$  в эталонном изображении, можно выбрать фрагмент, который даст наибольшее значение в  $RM$  (т. е.  $\max f(p_i)$ ) среди всех фрагментов, независимо от того, насколько минимальное или максимальное значение находится в диапазоне  $[0, \infty]$ . Однако это неразумно, поскольку на исследуемом изображении может отсутствовать какой-либо объект, интересующий пользователя. Рассмотрим два типа тестов значимости. Первый — это общий тест, позволяющий решить, присутствует ли в исследуемом изображении какой-либо достаточно похожий объект. Если да, необходимо узнать, сколько объектов присутствует в исследуемом изображении и где они находятся. Для этого рассмотрим два порога: общий порог  $\tau_0$  и порог  $\tau$  для обнаружения нескольких объектов, присутствующих в исследуемом изображении.

В обычном случае мы устанавливаем общий порог  $\tau_0$  равным 0,96, что составляет около 50% общей дисперсии (т. е.  $P_{\text{общ}} = 0,49$ ). Иными словами, если максимальное значение функции  $f(p_i)$  находится чуть выше значения 0,96, то существу-

ет хотя бы один интересующий нас объект. Следующим шагом является выбор  $\tau$  на основе свойств  $f(p_i)$ . Когда дело доходит до выбора  $\tau$ , нужно быть более осторожным. Если есть базовые знания о распределении  $f(p_i)$ , можно сделать прогноз о том, как эта конкретная статистика будет вести себя, и поэтому можно выбрать порог, который будет указывать сходство признаков двух изображений. Но на практике у нас нет оптимального способа смоделировать распределение  $f(p_i)$ . Поэтому вместо того чтобы предполагать тип базового распределения, используем идею непараметрического тестирования. Вычислим функцию плотности вероятности (PDF) из всех данных выборок  $f(p_i)$ , необходимо выбрать  $\tau$  таким образом, чтобы достичь, например, уровня достоверности 99% для определения того, находятся ли заданные значения в крайних хвостах распределения. Этот подход основан на предположении, что в исследовании изображении большинство фрагментов не содержат интересующий нас объект, и поэтому несколько совпадений приведут к значениям, которые находятся в хвостах распределений  $f(p_i)$  [3].

После того как выполнены два теста значимости с  $\tau_0$  и  $\tau$ , для окончательного обнаружения используется идея подавления ложных максимумов. Необходимо выбрать область с самым высоким значением  $f(p_i)$  и исключаяющий возможность повторного обнаружения какого-либо другого объекта [11]. Это позволяет избежать множественных ложных обнаружений уже обнаруженных объектов. Затем необходимо повторять этот процесс до тех пор, пока локальное максимальное значение не станет ниже порога  $\tau$ . Графическая иллюстрация тестов значимости и идею подавления ложных максимумов показана на рис. 6.

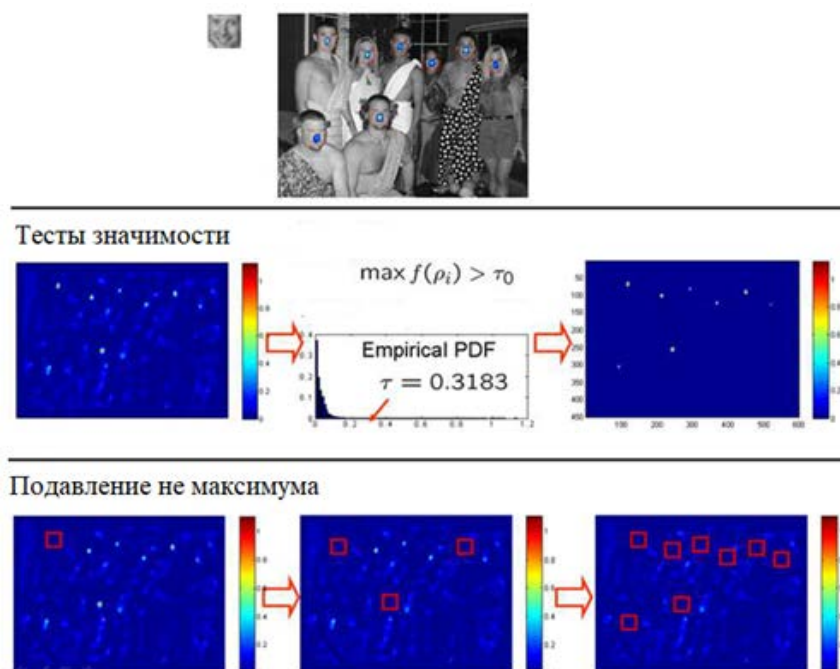


Рис. 6. Иллюстрацию тестов значимости и подавление ложных максимумов [9]

Сложные наборы реальных экспериментов с объектами продемонстрировали, что предлагаемый подход обеспечивает высокую точность обнаружения объектов, представляющих интерес даже в совершенно другом контексте и при разных условиях формирования изображения. В отличие от других современных методов обнаружения, основанных на обучении, предлагаемая структура работает с использованием единственного примера изображения, представляющего интерес для пользователя, чтобы найти подобные совпадения; не требует предварительного знания о разыскиваемых объектах и не нуждается в сегментации или предварительной обработке исследуемого изображения. Предлагаемая структура достаточно проста в использовании и может быть расширена до 3D-приложений.

**Выводы.** В работе применена новая мощная система обнаружения непараметрических объектов без обучения с использованием локальных управляемых ядер, которые хорошо фиксируют базовую структуру данных с коэффициентом корреляции в качестве меры сходства [12]. Предлагаемый метод может автоматически обнаруживать в исследуемом изображении наличие интересующих пользователя деталей. В результате работы был исследован и модернизирован алгоритм обнаружения объекта на цифровом изображении, основанный на методе локальных ядер регрессии. Фундаментальной составляющей метода является расчет локального управления ядра, который измеряет локальное сходство соседних пикселей как геометрически, так и радиометрически. Метод был реализован в среде разработки Microsoft Visual Studio на языке программирования C++. Модернизация алгоритма заключалась в нахождении на изображении нескольких схожих друг с другом объектов, а также в изменении масштаба искомого объекта на исследуемом изображении и приведении его к размерам эталонного изображения. Надежность метода гарантирует получение локальной структуры изображений даже при неопределенности данных, в частности, при наличии шума и размытости.

## Литература

- [1] Гонсалес Р., Вудс Р. *Цифровая обработка изображений*. Москва, Техносфера, 2005, 1072 с.
- [2] Seo H.J., Milanfar P. Training-free, generic object detection using locally adaptive regression kernels. *Transactions on Pattern Analysis & Machine Intelligence*, 2010, vol. 32, no. 9, pp. 1688–1704.
- [3] OpenCV шаг за шагом. Введение. URL: <http://robocraft.ru/blog/computervision/264.html> (дата обращения 20.09.2017).
- [4] Takeda H., Farsiu S., Milanfar P. Kernel regression for image processing and reconstruction. *Transactions on Image Processing*, 2007, vol. 16, no. 2, pp. 349–366.
- [5] Shechtman E., Irani M. Matching local self-similarities across images and videos. *Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
- [6] Kervrann C., Boulanger J. Optimal spatial adaptation for patch-based image denoising. *Transactions on Image Processing*, 2006, vol. 15, no. 10, pp. 2866–2878.
- [7] Buades A., Coll B., Morel J.M. Nonlocal image and movie denoising. *International Journal of Computer Vision*, 2008, vol. 76, no. 2, pp. 123–139.



- [8] Devernay F. *A non-maxima suppression method for edge detection with sub-pixel accuracy*. INRIA, 1995, 15 p.
- [9] Ke Y., Suthanakar R. PCA-SIFT: a more distinctive representation for local image descriptors. *Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 506–513.
- [10] Bay H., Ess A., Tuytelaars T., Gool L.V. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 2008, vol. 110, no. 3, pp. 346–359.
- [11] Левинский А.Т., Родионов И.Д., Агальцев С.С. Обнаружение объекта на цифровом изображении. *Политехнический молодежный журнал*, 2018, № 1. URL: <http://ptsj.ru/catalog/icec/sacip/228.html>.
- [12] Левинский А.Т., Селюто Н.М., Родионов И.Д. Отслеживание и оценка скорости движения объекта в видеопотоке. *Политехнический молодежный журнал*, 2017, № 12. URL: <http://ptsj.ru/catalog/icec/sacip/209.html>.

**Левинский Адам Тагирович** — магистрант кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

**Родионов Илья Дмитриевич** — магистрант кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

**Агальцев Станислав Сергеевич** — магистрант кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

**Тимофеев Дмитрий Владимирович** — магистрант кафедры «Информационные системы и телекоммуникации», МГТУ им. Н.Э. Баумана, Москва, Российская Федерация.

## TESTING THE OBJECT DETECTION ALGORITHM ON THE DIGITAL IMAGE

<b>A.T. Levinskiy</b>	adam.levinskiy@yandex.ru SPIN-code: 2301-6960
<b>I.D. Rodionov</b>	ir@drp.ru SPIN-code: 2598-2061
<b>S.V. Agalcev</b>	sosimba-sosimbovich@mail.ru SPIN-code: 5570-9322
<b>D.V. Timofeev</b>	dm.vl.timofeev@gmail.com SPIN-code: 9258-7883

**Bauman Moscow State Technical University, Moscow, Russian Federation**

---

### Abstract

*The subject matter of this article is a detection algorithm alongside with the investigation and modernization of the algorithm for detecting several objects on the digital image. The target goal is achieved due to the application of the studied image scaling with the aid of finding the maximum points where the image has the utmost resemblance to the master image. We have undertaken the work to improve the algorithm for detecting the objects on the image. An option has been added in the program to set the dimensions of the images, the coefficient of the similarity between the images and the scaling coefficient.*

### Keywords

*Algorithm, detection, object, digital image, coefficient, image processing, fragment, nuclei, descriptor*

© Bauman Moscow State Technical University, 2018

---

### References

- [1] Gonzalez R.C., Woods R.E. Digital image processing. Prentice Hall, 2002, 793 p. (Russ. ed.: Tsifrovaya obrabotka izobrazheniy. Moscow, Tekhnosfera publ., 2005, 1072 p.)
- [2] Seo H.J., Milanfar P. Training-free, generic object detection using locally adaptive regression kernels. *Transactions on Pattern Analysis & Machine Intelligence*, 2010, vol. 32, no. 9, pp. 1688–1704.
- [3] OpenCV shag za shagom. Vvedenie [OpenCV step by step. Introduction]. Available at: <http://robocraft.ru/blog/computervision/264.html> (accessed 20 September 2017).
- [4] Takeda H., Farsiu S., Milanfar P. Kernel regression for image processing and reconstruction. *Transactions on Image Processing*, 2007, vol. 16, no. 2, pp. 349–366.
- [5] Shechtman E., Irani M. Matching local self-similarities across images and videos. *Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
- [6] Kervrann C., Bourlanger J. Optimal spatial adaptation for patch-based image denoising. *Transactions on Image Processing*, 2006, vol. 15, no. 10, pp. 2866–2878.
- [7] Buades A., Coll B., Morel J.M. Nonlocal image and movie denoising. *International Journal of Computer Vision*, 2008, vol. 76, no. 2, pp. 123–139.
- [8] Devernay F. A non-maxima suppression method for edge detection with sub-pixel accuracy. INRIA, 1995, 15 p.
- [9] Ke Y., Suthanakar R. PCA-SIFT: a more distinctive representation for local image descriptors. *Computer Vision and Pattern Recognition*, 2004, vol. 2, pp. 506–513.

- [10] Bay H., Ess A., Tuytelaars T., Gool L.V. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 2008, vol. 110, no. 3, pp. 346–359.
- [11] Levinskiy A.T., Rodionov I.D., Agal'tsev S.S. Detecting an object in a digital image. *Politekhnicheskij molodezhnyy zhurnal* [Politechnical student journal], 2018, no. 1. Available at: <http://ptsj.ru/catalog/icec/sacip/228.html>.
- [12] Levinskiy A.T., Selyuto N.M., Rodionov I.D. Motion tracking and estimation of object speed in video stream. *Politekhnicheskij molodezhnyy zhurnal* [Politechnical student journal], 2017, no. 12. Available at: <http://ptsj.ru/catalog/icec/sacip/209.html>.

**Levinskiy A.T.** — Master's Degree student, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.

**Rodionov I.D.** — Master's Degree student, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.

**Agalcev S.V.** — Master's Degree student, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.

**Timofeev D.V.** — Master's Degree student, Department of Information Systems and Telecommunications, Bauman Moscow State Technical University, Moscow, Russian Federation.