ПОИСК ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ С ИСПОЛЬЗОВАНИЕМ СТРУКТУРНОГО ДЕСКРИПТОРА НА ОСНОВЕ ГРАФОВ

А.А. Захаров [1], А.Е. Баринов [1], А.Л. Жизняков [1], В.С. Титов [2]

[1] Муромский институт (филиал) ФГБОУ ВО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Муром, Россия,

[2] ФГБОУ ВО «Юго-Западный государственный университет», Курск, Россия

Аннотация

В работе рассматривается разработка структурного дескриптора для поиска объектов на изображениях. Дескриптор построен на основе графа, вершинами которого являются центры масс сегментов особенностей. Для вложения графа в векторное пространство используется преобразование Юнга–Хаусхолдера. Предложенный способ вложения графа в векторное пространство базируется на методах дифференциальной геометрии. Для описания связи между точками используются составные кривые. Граф изображения описывается матрицей параметров кривизны. С помощью метрики Хаусдорфа вычисляется матрица расстояний для графов объекта-кандидата и объекта-эталона. Для представления результатов используется метод многомерного шкалирования. Для исследования разработанного подхода были использованы изображения тестовых объектов и изображения лиц людей. При обнаружении головы человека на изображениях сравнение разработанного дескриптора осуществлялось с методом Виолы–Джонса. Достоинством разработанного подхода является инвариантность к повороту изображения на плоскости при поиске объектов. Также дескриптор позволяет обнаруживать объекты с углом поворота в пространстве до 50 градусов. Использование центров масс сегментов особенностей в качестве вершин графа значительно повышает устойчивость подхода при изменении ракурса съемки по сравнению с подходом, в котором вершинами графа являются особые точки изображения.

Ключевые слова: анализ изображений, поиск объектов, структурный дескриптор, вложение графов, компьютерное зрение.

Цитирование: **Захаров, А.А.** Поиск объектов на изображениях с использованием структурного дескриптора на основе графов / А.А. Захаров, А.Е. Баринов, А.Л. Жизняков, В.С. Титов // Компьютерная оптика. – 2018. – Т. 42, № 2. – С. 283-290. – DOI: 10.18287/2412-6179-2018-42-2-283-290.

Введение

Поиск объектов на изображениях является актуальной задачей в различных областях компьютерного зрения: человеко-машинные интерфейсы, биометрия, дистанционное зондирование Земли, базы данных изображений, системы контроля технологических процессов и т.д.

Часто поиск объектов осуществляется на основе сопоставления с использованием дескрипторов. Дескриптор представляет собой метод, который идентифицирует некоторую область изображения на основе набора признаков. Выделяют следующие группы дескрипторов двумерных изображений [1, 2]: локальные двоичные дескрипторы, дескрипторы на основе спектрального представления, дескрипторы на основе базисных функций, дескрипторы формы. Следует отметить, что некоторые методы по своим характеристикам можно отнести к различным группам одновременно.

Локальные двоичные дескрипторы представляют собой описание небольшой области изображения в виде бинарных векторов. Наиболее известными локальными дескрипторами являются локальные двоичные шаблоны и их модификации (Local Binary Patterns – LBP) [3]. Также к этой группе относятся BRIEF (Binary Robust Independent Elementary Features) [4], ORB (Oriented BRIEF) [5], BRISK (Binary Robust Invariant Scalable Keypoints) [6] и т.д.

Дескрипторы на основе спектрального представления используют различные величины для идентификации областей: интенсивность, цвет, градиенты, статистические характеристики и т.д. По сравнению с двоичными дескрипторами, дескрипторы на основе спектрального представления требуют более интенсивных вычислений. К таким методам относятся следующие дескрипторы и их модификации: SIFT (Scale Invariant Feature Transform) [7], SURF (Speeded Up Robust Features) [8], DAISY [9], HoG (Histogram of Gradients) [10], корреляционные шаблоны (Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Normalized Cross Correlation (NCC)) [11], Local Gradient Pattern (LGP) [12], код Фримана (Chain Code Histograms (CCH)) [13], признаки Хаара (HAAR Features) [14] и др.

Дескрипторы на основе базисных функций представляют описание изображения в заданных пространствах. Наиболее известным дескриптором на основе базисных функций является дескриптор Фурье [15]. К этой группе дескрипторов можно также отнести методы разреженного кодирования (Sparse Coding) [1]. Примером дескриптора из этой группы является метод «мешок слов» (Bag of Words) [16]. В дескрипторах на основе разреженного кодирования вместо базисных функций используется набор кодов для идентификации объектов.

Дескрипторы формы (Polygon Shape Descriptors) позволяют выполнять поиск объектов на основе та-

ких характеристик, как площадь, параметры контура, моменты областей, центр тяжести полигона, коэффициенты прямоугольности и округлости, количество дыр и т.д. К наиболее известным подходам относятся MSER (*Maximally Stable Extremal Regions*) [17], код Фримана [13], дескриптор Фурье [15], контекст формы (*Shape Context*) [18], дескриптор на основе площади дискового покрытия [19], морфологические дескрипторы формы [20], дескриптор на основе ограничивающих областей [21], моменты региона [22], кривизна границ (*curvature scale-space*) [23], скелетон (*Shock graphs*) [24] и др.

Недостатком многих подходов является то, что в системах компьютерного зрения часто затруднительно выделить границы анализируемого объекта из-за эффектов резкой смены освещения, взаимных перекрытий, сложного текстурированного фона, изменения ракурса и т.д.

В работе предлагается структурный дескриптор, который идентифицирует область, состоящую из сотен и тысяч пикселей. Такая область сопоставима с изображениями объектов, поиск которых ведется в системах компьютерного зрения. Для построения дескриптора предлагается использовать граф, вершинами которого являются особенности изображения. В предыдущих работах рассматривались вопросы нахождения соответствий [25], кластеризации особенностей [26, 27] и обнаружения объектов на основе спектральной теории графов [28]. В представленной работе новизной является следующее: рассматривается использование дескриптора для обнаружения лиц на изображениях, проводится исследование точности обнаружения лиц от угла поворота головы человека в пространстве.

1. Разработка структурного дескриптора на основе графов

Преимуществом структурных подходов является то, что они позволяют анализировать большое множество элементов на основе малого количества простых составляющих и правил формирования графической модели. Также структурные методы позволяют описать те характеристики объекта, которые исключают его отнесение к другому классу, что повышает надежность распознавания.

Предлагаемый структурный дескриптор включает в себя построение графов изображения-эталона и текущего изображения, вложение графов в векторное пространство и классификацию.

Построение графа на основе изображения

На вход дескриптора поступают текущее изображение и изображение-эталон, полученное на этапе инициализации. Следует отметить, что при поворотах объектов в пространстве некоторые особенности на различных снимках будут пропадать. Поэтому для повышения надежности поиска предлагается отслеживать центры масс сегментов особенностей. При повороте объекта набор особых точек в таких областях может измениться, но сама область будет присутствовать на изображении (рис. 1). Особенности выделяются с использованием детектора SURF [8].

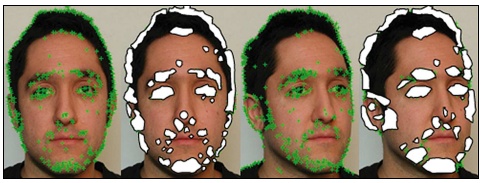

Рис. 1. Кластеризация особенностей, выделенных на изображении лица человека детектором SURF

Проводится кластеризация особых точек. В ходе данного процесса все особые точки, выделенные на изображении, группируются на основе метода связных компонент. В этом случае строится граф на основе триангуляции Делоне. Задается параметр W, и в графе удаляются все ребра, длина которых больше W. Для подбора параметра W строится гистограмма распределений попарных расстояний между вершинами. Для изображений с хорошо выраженной кластерной структурой на гистограмме будет два пика, один из которых соответствует внутрикластерным расстояниям, второй – межкластерным расстояния. Параметр W подбирается из зоны минимума между этими пиками. Соединенными остаются только наиболее близкие вершины. Если в каком-либо полученном сегменте оказывается мало особенностей (меньше 3), то он не рассматривается далее. Таким образом, формируются сегменты, соответствующие наиболее характерным областям изображения.

После кластеризации осуществляется построение графа изображения. В качестве вершин выбираются центры масс полученных сегментов. Граф строится на основе триангуляции Делоне.

Нормализованная матрица Лапласа графа рассчитывается на основе выражения (1):

$$L_n = \begin{cases} 1, \text{ если } u = v \text{ и } d_v \neq 0; \\ -\dfrac{1}{\sqrt{d_u d_v}}, \text{ если } A(u,v) = 1; \\ 0, \quad \text{в другом случае,} \end{cases} \tag{1}$$

где d_u, d_v – степени вершин u и v соответственно.

Вложение графов в векторное пространство и классификация объектов на изображениях

Для поиска объектов предлагается выполнять вложение графов в векторное пространство. Это позволит представить изображения в виде векторов числовых характеристик, что даст возможность приближенного сравнения структур, не требующего точного сопоставления графов. Достоинством разрабатываемого подхода является инвариантность к повороту изображения на плоскости, так как спектральные характеристики графа не зависят от маркировки его вершин [29].

Для сравнения изображения-кандидата с эталоном предлагается использовать спектральную теорию графов. Для решения поставленной задачи выполняется вложение графов, построенных по сравниваемым изображениям, в векторное пространство [30]. В этом

случае графы сравниваемых изображений преобразуются в вектор числовых характеристик, на основе которого выполняется сравнение.

На основе декомпозиции нормализованной матрицы Лапласа вычисляются спектральные характеристики графа:

$$L_n = \Phi\Lambda\Phi^T, \tag{2}$$

где Λ – диагональная матрица собственных значений $\lambda_1, \lambda_2, \ldots, \lambda_{|V|}$;

Φ – матрица собственных векторов $\varphi_1, \varphi_2, \ldots, \varphi_{|V|}$.

Вложение графа основано на решении термодинамического уравнения:

$$\frac{\partial H_t}{\partial t} = -L_n H_t, \tag{3}$$

где t – время изменения состояния графа; H_t – тепловое ядро.

Обычно значение времени подбирается опытным путем [30]. Однако кластеризация особых точек позволяет снизить зависимость результата от значения t. В рассматриваемых далее примерах начальное значение $t = 0{,}01$.

Тепловое ядро является решением уравнения (3) и вычисляется с помощью собственных значений и собственных векторов:

$$H_t(u,v) = \sum_{i=1}^{|V|} e^{-\lambda_i t} \varphi_i(u) \varphi_i(v), \tag{4}$$

где λ_i – собственные значения матрицы Лапласа;

φ_i – собственные векторы матрицы Лапласа.

При проецировании графа в векторное пространство с помощью спектральных характеристик используются значения теплового ядра. Выполняется декомпозиция Юнга–Хаусхолдера:

$$H_t = Y^T Y, \tag{5}$$

где $Y = (y_1 \ldots y_u \ldots y_{|V|})$ – матрица координат размером $|V| \times |V|$, в которой каждый столбец является вектором координат соответствующей вершины.

Раскрывая это выражение, получаем

$$Y = e^{-(\Lambda t/2)} \Phi^T. \tag{6}$$

Следовательно, для вершины u координатный вектор рассчитывается следующим образом:

$$y_u = \left(e^{-\frac{\lambda_1 t}{2}} \varphi_1(u),\ e^{-\frac{\lambda_2 t}{2}} \varphi_2(u), \ldots,\ e^{-\frac{\lambda_{|V|} t}{2}} \varphi_{|V|}(u) \right)^T. \tag{7}$$

При вложении графа в векторное пространство используется дифференциальная геометрия, в которой для описания связи между элементами матрицы координат используются составные кривые. По теореме Гаусса–Боне часть такой кривой может быть аппроксимирована дугой окружности. Такие кривые характеризуются параметрами кривизны, значения которой могут быть получены по значениям кратчайшего пути по окружности (длина дуги) и евклидовому расстоянию (длина хорды). Далее длину дуги будем обозначать d_G, длину хорды – d_E (рис. 2).

Квадратичное евклидово расстояние в этом случае имеет следующий вид:

$$d_E(u,v)^2 = (y_u - y_v)^T (y_u - y_v) = \\ = \sum_{i=1}^{|V|} e^{-\lambda_i t} \left[\varphi_i(u) - \varphi_i(v) \right]^2. \tag{8}$$

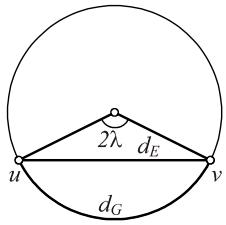

Рис. 2. Представление кривой с помощью дуги окружности

Составные кривые характеризуются параметрами кривизны. Для u и v значение кривизны рассчитывается следующим образом:

$$k(u, v) = \frac{1}{R(u, v)}, \tag{9}$$

где $R(u, v)$ – радиус окружности.

Представим кривую дугой окружности с радиусом $R(u, v)$. В таком случае длина дуги, то есть кратчайший путь по окружности между точками u и v, вычисляется следующим образом:

$$d_G = 2\alpha R(u,v), \tag{10}$$

где α – угол дуги.

При вложении графов кратчайшее расстояние характеризуется весом ребра. Следовательно, для невзвешенного графа можно принять $d_G = 1$. В таком случае угол α вычисляется следующим образом:

$$\alpha = 1/2R(u, v). \tag{11}$$

Евклидово расстояние равно длине хорды в рассматриваемой окружности между точками u и v:

$$d_E = 2R(u, v)\sin(\alpha). \tag{12}$$

В дальнейших расчетах необходимо избавиться от вычисления функции синуса неизвестного угла. Для этого предлагается разложить функцию $\sin(x)$ в ряд Маклорена, то есть в ряд Тейлора при $x = 0$:

$$d_E = 2R(u,v)\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \ldots \right). \tag{13}$$

Принимается, что в расчетах используется точность до двух членов последовательности ряда Маклорена. Таким образом, имеем следующее:

$$d_E = 1 - \frac{1}{24R(u,v)^2}. \tag{14}$$

Решая это уравнение в поисках $R(u, v)$, получим следующее значение кривизны для точек u и v:

$$k(u, v) = \sqrt{24(1 - d_E)}. \tag{15}$$

Таким образом, каждый граф можно описать матрицей параметров кривизны. С помощью метрики Ха-

усдорфа рассчитывается подобие между матрицами графов [30]. Имея матрицу подобия, с помощью метода многомерного шкалирования MDS (Multi Dimensional Scaling) можно определить, насколько соответствуют друг другу объект-кандидат и объект-эталон. Метод MDS используется для уменьшения размерности набора данных [31].

Пусть имеются два графа $G_1=(V_1,E_1,k_1)$ и $G_2=(V_2,E_2,k_2)$, где V_1, V_2 – набор вершин; E_1, E_2 – набор ребер; k_1, k_2 – матрицы кривизны. Таким образом, расстояния между графами можно описать с помощью метрики Хаусдорфа:

$$HD(G_1,G_2)=\max_{i\in V_1}\max_{j\in V_1}\min_{I\in V_2}\min_{J\in V_2}\left\|k_2(I,J)-k_1(i,j)\right\|. \quad (16)$$

С помощью метрики Хаусдорфа вычисляется матрица расстояний для графов объекта-кандидата и объекта-эталона. Для визуализации результатов используется метод MDS. Результатом применения метода MDS к матрице расстояний является представление, в котором характеристики графов описываются точками в евклидовом пространстве. Точки кластеризуются с помощью метода связных компонент для отнесения изображений объектов к какому-либо эталону.

Алгоритм реализации структурного дескриптора на основе графов

Алгоритм реализации структурного дескриптора на основе графов состоит из следующих шагов.

Шаг 1. Выделяются особенности изображения-кандидата и изображения-эталона.

Шаг 2. Проводится кластеризация особенностей двух изображений.

Шаг 3. На основе центров масс сегментов строятся графы Делоне. Рассчитываются нормализованные матрицы Лапласа (1) и тепловые ядра (4).

Шаг 4. С помощью преобразования Юнга – Хаусхолдера происходит проецирование координат вершин графов в векторное пространство (7).

Шаг 5. Рассчитываются значения кривых, связывающих точки. Строятся матрицы кривизны (15).

Шаг 6. Вычисляется матрица расстояний между характеристиками графа изображения-эталона и графа анализируемого изображения (16).

Шаг 7. Для визуализации данных применяется метод многомерного шкалирования, позволяющий отобразить степень близости изображения-кандидата к какому-либо объекту из базы эталонов.

Шаг 8. Происходит кластеризация точек в пространстве меньшей размерности на основе метода связных компонент. Вхождение точек в кластер свидетельствует об отнесении изображений объектов к определенному изображению-эталону.

2. Исследование структурного дескриптора на основе графов

В ходе исследования были использованы следующие базы данных изображений: Carnegie Mellon University (CMU) Image Database, база данных Массачусетского технологического института MIT-CBCL Face Recognition Database. На используемых изображениях меняется угол поворота объекта в пространстве. В данном случае обнаружение лица осуществляется на портретных снимках. При использовании разработанного дескриптора на изображениях реальных сцен происходит предварительное выделение особых точек, относящихся к области лица на изображении [26, 27].

Было проведено исследование разработанного подхода. Выполнялось сравнение с алгоритмом вложения графа, в основе которого лежит использование всех особых точек сцены [30].

На рис. 3*а* вершинами графа являются все особые точки изображения, а на рис. 3*б* вершинами графа являются только центры масс сегментов особенностей.

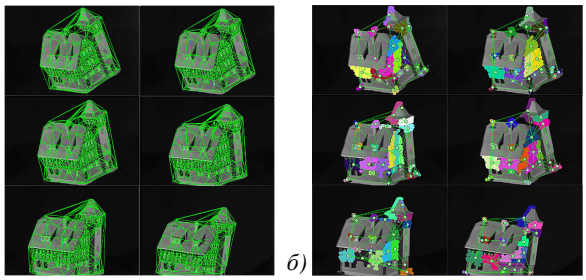

Рис. 3. Построение графа объекта по тестовым изображениям: а) вершинами графа являются все особенности изображения; б) вершинами графа являются центры масс сегментов особенностей

Были получены следующие результаты вложения графов с использованием метода многомерного шкалирования (рис. 4).

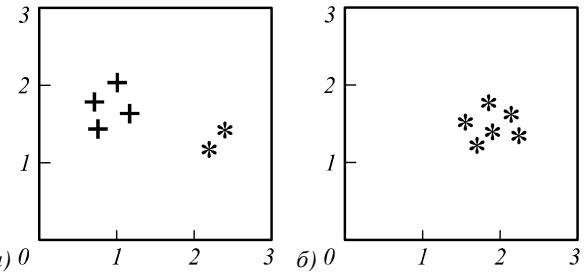

Рис. 4. Результаты вложения графов, изображенных на рис. 3, в векторное пространство: вершинами графа являются все особые точки изображения (а); вершинами графа являются центры масс сегментов особенностей (б)

Несмотря на то, что используются изображения одного объекта, в ситуациях, когда вершинами графа являются все особые точки изображения, алгоритм выделяет две разные группы. Напротив, разработанный алгоритм, рассматривая в качестве вершин графа центры масс сегментов особенностей, отнес объекты к одной группе. На рис. 5 показано построение графов по изображениям лиц.

На рис. 6 показаны результаты вложения графов изображений лиц, представленные с помощью метода многомерного шкалирования. Разработанный дескриптор разделил на группы изображения лиц разных людей.

Одна из функций дескриптора – обнаружение головы человека на изображении. В данном случае сравнение работоспособности осуществлялось также с мето-

дом Виолы–Джонса [32], так как он является некоторым стандартом в области компьютерного зрения и широко применяется в задачах поиска лиц на изображениях. Рассматривалась реализация метода Виолы–-Джонса, когда при обучении лицо человека расположено строго анфас. Вычисление характеристик изображения на основе разработанного дескриптора выполнялось в аналогичных условиях.

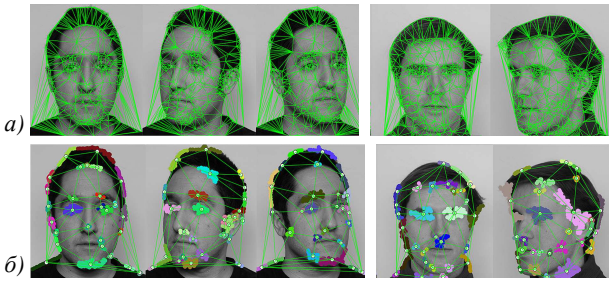

Рис. 5. Построение графов по изображениям лиц: вершинами графа являются все особенности изображения (а); вершинами графа являются центры масс сегментов особенностей (б)

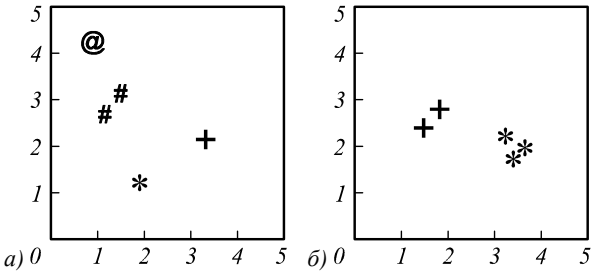

Рис. 6. Результаты вложения графов, изображенных на рис. 5, в векторное пространство: вершинами графа являются все особенности изображения (а); вершинами графа являются центры масс сегментов особенностей (б)

Было выполнено тестирование разработанного алгоритма на изображениях лиц из базы данных MIT-CBCL. В данном случае диапазон кивка, наклона и поворота головы составлял от 0 до 90°. При тестировании алгоритма для каждого лица из базы данных была составлена выборка изображений, на которых голова человека имела различную ориентацию в пространстве. Были получены следующие результаты (рис. 7). Главной особенностью разработанного подхода является то, что при любом значении угла поворота в плоскости *XOY* правильное обнаружение лица было не ниже 90 %. Это связано с тем, что спектральные характеристики графа не зависят от маркировки его вершин.

При использовании метода Виолы–Джонса с увеличением углов поворота, кивка и наклона снижается точность правильного обнаружения лица человека. Напротив, разработанный алгоритм обладает большей точностью. В случаях вращения объекта в плоскости *XOY* точность обнаружения остается неизменной.

Заключение

При исследовании разработанного подхода выявлено, что дескриптор обладает инвариантностью к повороту изображения на плоскости, а также способностью обнаруживать объекты с углом поворота в пространстве до 50°. Также выявлено, что использование центров масс сегментов особенностей в качестве вершин графа значительно повышает устойчивость подхода при изменении ракурса съемки объекта.

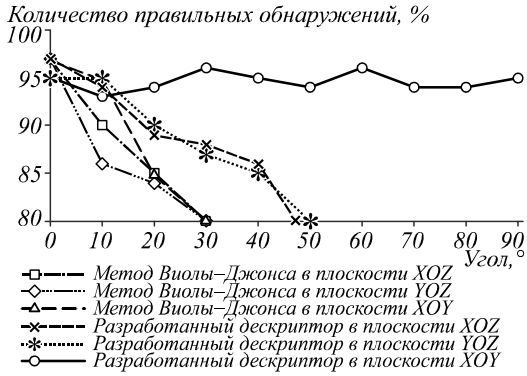

Рис. 7. Графики зависимости количества правильных обнаружений головы от угла поворота в плоскостях XOY, XOZ и YOZ для метода Виолы–Джонса и разработанного алгоритма

Благодарности

Работа выполнена при финансовой поддержке государственного задания Минобрнауки России (проект № 2.1950.2017/ПЧ), РФФИ (проект № 16-37-00235).

Литература

1. **Krig, S.** Computer vision metrics: Survey, taxonomy, and analysis / S. Krig. – Berkeley, CA: Apress Media, 2014. – 498 p. – ISBN: 978-1-4302-5929-9.
2. **Jain, M.** A survey on CBIR on the basis of different feature descriptor / M. Jain, D. Singh // British Journal of Mathematics & Computer Science. – 2016. – Vol. 14, Issue 6. – P. 1-13. – DOI: 10.9734/BJMCS/2016/24000.
3. **Ojala, T.** A comparative study of texture measures with classification based on feature distributions / T. Ojala, M. Pietikäinen, D. Hardwood // Pattern Recognition. – 1996. – Vol. 29, Issue 1. – P. 51-59. – DOI: 10.1016/0031-3203(95)00067-4.
4. **Calonder, M.** BRIEF-binary robust independent elementary features / M. Calonder, V. Lepetit, C. Strecha, P. Fua // European Conference on Computer Vision. – 2010. – Part IV. – P. 778-792. – DOI: 10.1007/978-3-642-15561-1_56.
5. **Rublee, E.** ORB: An efficient alternative to SIFT or SURF / E. Rublee, V. Rabaud, K. Konolige, G. Bradski // IEEE International Conference on Computer Vision (ICCV). – 2011. – P. 2564-2571. – DOI: 10.1109/ICCV.2011.6126544.
6. **Leutenegger, S.** BRISK: Binary Robust invariant scalable keypoints / S. Leutenegger, M. Chli, R. Siegwart // IEEE International Conference on Computer Vision (ICCV'11). – 2011. – P. 2548-2555. – DOI: 10.1109/ICCV.2011.6126542.
7. **Lowe, D.G.** Distinctive image features from scale-invariant keypoints / D.G. Lowe // International Journal of Computer Vision. – 2004. – Vol. 60, Issue 2. – P. 91-110. – DOI: 10.1023/B:VISI.0000029664.99615.94.
8. **Bay, H.** SURF: Speeded up robust features / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // Computer Vision and Image Understanding. – 2008. – Vol. 110, Issue 3. – P. 346-359. – DOI: 10.1016/j.cviu.2007.09.014.
9. **Tola, E.** DAISY: An efficient dense descriptor applied to wide-baseline stereo / E. Tola, V. Lepetit, P. Fua // IEEE Transactions on Pattern Analysis and Machine Intelli-

gence. – 2010. – Vol. 32, Issue 5. – P. 815-830. – DOI: 10.1109/TPAMI.2009.77.
10. **Dalal, N.** Histograms of oriented gradients for human detection / N. Dalal, B. Triggs // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005). – 2005. – Vol. 1. – P. 886-893. – DOI: 10.1109/CVPR.2005.177.
11. **Scharstein, D.** A taxonomy and evaluation of dense two-frame stereo correspondence algorithms / D. Scharstein, R. Szeliski // International Journal of Computer Vision. – 2002. – Vol. 47, Issue 1-3. – P. 7-42. – DOI: 10.1023/A:1014573219977.
12. **Jun, B.** Robust face detection using local gradient patterns and evidence accumulation / B. Jun, D. Kim // Pattern Recognition. – 2012. – Vol. 45, Issue 9. – P. 3304-3316. – DOI: 10.1016/j.patcog.2012.02.031.
13. **Freeman, H.** On the encoding of arbitrary geometric configurations / H. Freeman // IRE Transactions on Electronic Computers. – 1961. – Vol. EC-10, Issue 2. – P. 260-268. – DOI: 10.1109/TEC.1961.5219197.
14. **Gonzalez, R.** Digital image processing / R. Gonzalez, R. Woods. – 3rd ed. – Upper Saddle River, NJ: Prentice-Hall, 2007. – 976 p. – ISBN: 978-0-13-168728-8.
15. **Bracewell, R.** The Fourier transform and its applications / R. Bracewell. – 3rd ed. – New York: McGraw-Hill Science, 1999. – 640 p. – ISBN: 978-0-07-303938-1.
16. **Fei-Fei, L.** Recognizing and learning object categories / L. Fei-Fei, R. Fergus, A. Torralba // Conference on Computer Vision and Pattern Recognition. – 2007.
17. **Matas, J.** Robust widebaseline stereo from maximally stable extremal regions / J. Matas, O. Chum, M. Urban, T. Pajdla // British Machine Vision Conference. – 2002. – P. 384-393.
18. **Belongie, S.** Shape matching and object recognition using shape contexts / S. Belongie, J. Malik, J. Puzicha // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24, Issue 4. – P. 509-522. – DOI: 10.1109/34.993558.
19. **Ломов, Н.А.** Площадь дискового покрытия – дескриптор формы изображения / Н.А. Ломов, Л.М. Местецкий // Компьютерная оптика. – 2016. – Т. 40, № 4. – С. 516-525. – DOI: 10.18287/2412-6179-2016-40-4-516-525.
20. **Сидякин, С.В.** Морфологические дескрипторы формы бинарных изображений на основе эллиптических структурирующих элементов / С.В. Сидякин, Ю.В. Визильтер // Компьютерная оптика. – 2014. – Т. 38, № 3. – С. 511-520. – DOI: 10.18287/0134-2452-2014-38-3-511-520.
21. **Bauckhage, C.** Bounding box splitting for robust shape classification / C. Bauckhage, J.K. Tsotsos // IEEE International Conference on Image Processing (ICIP 2005). – 2005. – P. 478-481. – DOI: 10.1109/ICIP.2005.1530096.
22. **Sonka, M.** Image processing, analysis and machine vision / M. Sonka, V. Hlavac, R. Boyle. – London: Chapman and Hall, 1993. – 872 p. – ISBN: 978-0-412-45570-4.
23. **Abbasi, S.** Enhancing CSS-based shape retrieval for objects with shallow concavities / S. Abbasi, F. Mokhtarian, J. Kittler // Image and Vision Computing. – 2000. – Vol. 18, Issue 3. – P. 199-211. – DOI: 10.1016/S0262-8856(99)00019-0.
24. **Siddiqi, K.** A shock grammar for recognition / K. Siddiqi, B. Kimia // Proceedings of the 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '96). – 1996. – P. 507-513. – DOI: 10.1109/CVPR.1996.517119.
25. **Zakharov, A.A.** Finding correspondences between images using descriptors and graphs / A.A. Zakharov, A.Yu. Tuzhilkin, A.L. Zhiznyakov // Procedia Engineering. – 2015. – Vol. 129. – P. 391-396. – DOI: 10.1016/j.proeng.2015.12.131.
26. **Barinov, A.E.** Clustering using a random walk on graph for head pose estimation / A.E. Barinov, A.A. Zakharov // 2015 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS). – 2015. – DOI: 10.1109/MEACS.2015.7414876.
27. **Баринов, А.Е.** Алгоритм спектральной кластеризации с ограничениями для выделения лица человека на изображениях / А.Е. Баринов, А.А. Захаров, А.Л. Жизняков // Динамика систем, механизмов и машин. – 2016. – Т. 2, № 1. – С. 222-228.
28. **Zakharov, A.A.** Recognition of human pose from images based on graph spectra / A.A. Zakharov, A.E. Barinov, A.L. Zhyznyakov // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. – 2015. – Vol. XL-5/W6. – P. 9-12. – DOI: 10.5194/isprsarchives-XL-5-W6-9-2015.
29. **Chung, F.R.K.** Spectral graph theory / F.R.K. Chung. – Provides, Rhode Island: AMS, 1997. – 207 p. – ISBN: 0-8218-0315-8.
30. **ElGhawalby, H.** Measuring graph similarity using spectral geometry / H. ElGhawalby, E.R. Hancock. – In book: Image Analysis and Recognition. ICIAR 2008 / ed. by A. Campilho, M. Kamel. – Berlin, Heidelberg: Springer, 2008. – P. 517-526. – DOI: 10.1007/978-3-540-69812-8_51.
31. **Borg, I.** Modern multidimensional scalling: Theory and applications / I. Borg, P. Groenen. – 2nd ed. – New York, NY: Springer-Verlag, 2005. – 614 p. – P. 207-212. – ISBN: 0-387-25150-2.
32. **Viola, P.** Rapid Object detection using a boosted cascade of simple features / P. Viola, M. Jones // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). – 2001. – P. 511-518. – DOI: 10.1109/CVPR.2001.990517.

Сведения об авторах

Захаров Алексей Александрович, 1978 года рождения, к.т.н., доцент, ведущий научный сотрудник кафедры САПР МИ ВлГУ. Область научных интересов: компьютерное зрение, САПР. E-mail: *aa-zaharov@ya.ru* .

Баринов Алексей Евгеньевич, 1990 года рождения, лаборант-исследователь кафедры САПР МИ ВлГУ. Область научных интересов: компьютерное зрение, САПР. E-mail: *alexey.barinov.murom@ya.ru* .

Жизняков Аркадий Львович, 1975 года рождения, д.т.н., профессор, заведующий кафедрой САПР МИ ВлГУ. Область научных интересов: компьютерное зрение, цифровая обработка изображений, САПР. E-mail: *lvovich1975@mail.ru* .

Титов Виталий Семёнович, 1943 года рождения, д.т.н., профессор, заведующий кафедрой вычислительной техники Юго-Западного государственного университета. Область научных интересов: теоретические и методологические основы построения адаптивных оптико-электронных систем, применяемых при автоматизации технологических процессов и производств различного назначения. E-mail: tas_06@mail.ru .

ГРНТИ: 28.23.15.
Поступила в редакцию 21 ноября 2017 г. Окончательный вариант – 29 января 2018 г.

OBJECT DETECTION IN IMAGES WITH A STRUCTURAL DESCRIPTOR BASED ON GRAPHS

A.A. Zakharov [1], A.E. Barinov [1], A.L. Zhiznyakov [1], V.S. Titov [2]
[1] Murom Institute (branch), Vladimir State University named after Alexander and Nikolay Stoletovs, Murom, Russia,
[2] Southwest State University, Kursk, Russia

Abstract

We discuss the development of a structural descriptor for object detection in images. The descriptor is based on a graph, whose vertices are the centers of mass of segment features. The embedding of the graph in a vector space is implemented using a Young-Householder decomposition and based on differential geometry. Compound curves are used to describe the relationship between the points. The image graph is described by a matrix of curvature parameters. The distance matrix for the graphs of the candidate object and the reference object is calculated using the Hausdorff metric. A multidimensional scaling method is used to represent the results. Images of test objects and images of human faces are used to study the developed approach. A comparison of the developed descriptor with the Viola-Jones method is performed when detecting a human head in the image. The advantage of the developed approach is the image rotational invariance in the plane while searching for objects. The descriptor can detect objects rotated in space by angles of up to 50 degrees. Using the mass centers of segments of features as the graph vertices makes the approach more robust to changes in image acquisition angles in comparison with the approach that uses image features as the graph vertices.

Keywords: image analysis, objects detection, structural descriptor, graph embedding, computer vision.

Citation: Zakharov AA, Barinov AE, Zhiznyakov AL, Titov VS. Object detection in images with a structural descriptor based on graphs. Computer Optics 2018; 42(2): 283-290. – DOI: 10.18287/2412-6179-2018-42-2-283-290.

Acknowledgements: This work was partly funded by the RF Ministry of Education and Science under a state contract (project 2.1950.2017/ПЧ) and the Russian Foundation for Basic Research (RFBR grant No. 16-37-00235).

References

[1] Krig S. Computer vision metrics. Survey, taxonomy, and analysis. Apress, Berkeley, CA; 2014. ISBN: 978-1-4302-5929-9

[2] Jain M, Singh D. A survey on CBIR on the basis of different feature descriptor. British Journal of Mathematics & Computer Science 2016; 14(6): 1-13. DOI: 10.9734/BJMCS/2016/24000.

[3] Ojala T, Pietikainen M, Hardwood D. A comparative study of texture measures with classification based on feature distributions. Pattern Recognition 1996; 29(1): 51-59. DOI: 10.1016/0031-3203(95)00067-4.

[4] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF-Binary robust independent elementary features. ECCV 2010; IV: 778-792. DOI: 10.1007/978-3-642-15561-1_56.

[5] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: an efficient alternative to SIFT or SURF. ICCV 2011: 2564-2571. DOI: 10.1109/ICCV.2011.6126544.

[6] Leutenegger S, Chli M, Siegwart R. BRISK: Binary robust invariant scalable keypoints. ICCV 2011: 2548-2555. DOI: 10.1109/ICCV.2011.6126542.

[7] Lowe DG. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 2004; 60(2): 91-110. DOI: 10.1023/B:VISI.0000029664.99615.94.

[8] Bay H, Ess A, Tuytelaars T, Van Gool L. SURF: Speeded up robust features. Computer Vision and Image Understanding 2008; 110(3): 346-359. DOI: 10.1016/j.cviu.2007.09.014.

[9] Tola E, Lepetit V, Fua P. DAISY: An efficient dense descriptor applied to wide-baseline stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence 2010; 32(5): 815-830. DOI: 10.1109/TPAMI.2009.77.

[10] Dalal N, Triggs B. Histograms of oriented gradients for human detection. CVPR 2005; 1: 886-893. DOI: 10.1109/CVPR.2005.177.

[11] Scharstein D, Szeliski R. Taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 2002; 47(1-3): 7-42. DOI: 10.1023/A:1014573219977.

[12] Jun B, Kim D. Robust face detection using local gradient patterns and evidence accumulation. Pattern Recognition 2012; 45(9): 3304-3316. DOI: 10.1016/j.patcog.2012.02.031.

[13] Freeman H. On the encoding of arbitrary geometric configurations. IRE Transactions on Electronic Computers 1961; EC-10(2): 260-268. DOI: 10.1109/TEC.1961.5219197.

[14] Gonzalez R, Woods R. Digital image processing. 3rd ed. Upper Saddle River, NJ: Prentice-Hall; 2007.

[15] Bracewell R. The Fourier transform and its applications. 3rd ed. New York: McGraw-Hill Science; 1999. ISBN: 978-0-07-303938-1.

[16] Fei-Fei L, Fergus R, Torralba A. Recognizing and learning object categories. Conference on Computer Vision and Pattern Recognition 2007.
[17] Matas J, Chum O, Urban M, Pajdla T. Robust widebaseline stereo from maximally stable extremal regions. British Machine Vision Conference 2002: 384-393.
[18] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence 2002; 24(4): 509-522. DOI: 10.1109/34.993558.
[19] Lomov NA, Mestetskiy LM. Area of the disk cover as an image shape descriptor [In Russian]. Computer Optics 2016; 40(4): 516-525. DOI: 10.18287/2412-6179-2016-40-4-516-525.
[20] Sidyakin SV, Vizilter YV. Morphological shape descriptors of binary images based on elliptical structuring elements. Computer Optics 2014; 38(3): 511-520. DOI: 10.18287/0134-2452-2014-38-3-511-520.
[21] Bauckhage C, Tsotsos JK. Bounding box splitting for robust shape classification. IEEE International Conference on Image Processing 2005: 478-481. DOI: 10.1109/ICIP.2005.1530096.
[22] Sonka M, Hlavac V, Boyle R. Image processing, analysis and machine vision. London: Chapman and Hall; 1993. ISBN: 978-0-412-45570-4.
[23] Abbasi S, Mokhtarian F, Kittler J. Enhancing CSS-based shape retrieval for objects with shallow concavities. Image and Vision Computing 2000; 18(3): 199-211. DOI: 10.1016/S0262-8856(99)00019-0.
[24] Siddiqi K, Kimia B. A shock grammar for recognition. CVPR '96 1996: 507-513. DOI: 10.1109/CVPR.1996.517119.
[25] Zakharov AA, Tuzhilkin AYu, Zhiznyakov AL. Finding correspondences between images using descriptors and graphs. Procedia Engineering 2015; 129: 391-396. DOI: 10.1016/j.proeng.2015.12.131.
[26] Barinov AE, Zakharov AA. Clustering using a random walk on graph for head pose estimation // 2015 International Conference on Mechanical Engineering, Automation and Control Systems (MEACS) 2015. DOI: 10.1109/MEACS.2015.7414876.
[27] Barinov AE, Zakharov AA, Zhyznyakov AL. The algorithm of spectral clustering with restrictions for the selection of a person's face in images [In Russian]. Dynamics of systems, mechanisms and machines 2016; 2(1): 222-228.
[28] Zakharov AA, Barinov AE, Zhyznyakov AL. Recognition of human pose from images based on graph spectra. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2014; XL-5/W6: 9-12. DOI: 10.5194/isprsarchives-XL-5-W6-9-2015.
[29] Chung FRK. Spectral graph theory. Provides, Rhode Island: AMS; 1997. ISBN: 0-8218-0315-8.
[30] ElGhawalby H, Hancock ER. Measuring graph similarity using spectral geometry. In book: Campilho A, Kamel M, eds. Image Analysis and Recognition: ICIAR '08 2008; 5112: 517-526. DOI: 10.1007/978-3-540-69812-8_51.
[31] Borg I, Groenen P. Modern multidimensional scalling: theory and applications. New York, NY: Springer-Verlag; 2005: 207-212. ISBN: 0-387-25150-2.
[32] Viola P, Jones M. Rapid Object detection using a boosted cascade of simple features. CVPR 2001: 511-518. DOI: 10.1109/CVPR.2001.990517.

Author's information

Alexei Alexandrovich Zakharov (b. 1978), PhD, assistant professor, leading researcher of CAD department of MI VlSU. Research interests: computer vision, CAD. E-mail: *aa-zaharov@ya.ru* .

Alexei Evgenyevich Barinov (b. 1990), laboratory assistant researcher of CAD department of MI VlSU. Research interests: computer vision, CAD. E-mail: *alexey.barinov.murom@ya.ru* .

Arkady Lvovich Zhiznyakov (b. 1975), doctor of engineering, professor, head of CAD department of MI VlSU. Research interests: computer vision, image processing, CAD. E-mail: *lvovich1975@mail.ru* .

Vitaliy Semenovich Titov (b. 1943), doctor of engineering, professor, head of Computer Technology department of SWSU. Research interests: theoretical and methodological foundations of the construction of adaptive optoelectronic systems used in the automation of technological processes and industries for various purposes. E-mail: *tas_06@mail.ru* .

Received November 21, 2017. The final version – January 29, 2018.