

Метод автоматической сегментации транспортных средств на изображении

И. Г. Зубов✉

ООО "НЕКСТ", Москва, Россия

✉ZubovIG@gmail.com

Аннотация

Введение. Современные системы активной безопасности транспортных средств призваны существенно снизить количество дорожно-транспортных происшествий. Датчики на основе монокулярных камер все чаще внедряются ведущими мировыми автопроизводителями как эффективный инструмент повышения безопасности движения. Современные методы локализации и классификации в совокупности с алгоритмами семантической сегментации позволяют разделить изображение на независимые группы пикселей, соответствующие каждому объекту. Тем не менее является актуальным разработка методов сегментации, обеспечивающих улучшение качества сегментации изображений.

Цель работы. Разработка автоматического метода сегментации детектированного объекта интереса на изображении.

Методы и материалы. В статье предложен автоматический метод сегментации транспортных средств на изображении. Представленный метод позволяет провести семантическую сегментацию объекта интереса на основе априорной информации о границах прямоугольника, ограничивающего объект на изображении. Информация о границах объекта используется для преобразования изображения в полярную систему координат, где пиксели изображения выступают в роли ребер взвешенного графа. С использованием алгоритма поиска кратчайшего пути и обратного преобразования в декартову систему координат вокруг объекта интереса формируется замкнутый контур.

Результаты. Проведенные эксперименты подтвердили корректность выделения объекта интереса на основе предложенного метода. Коэффициент сходства Жаккара для открытой базы изображений Carvana составил 85 %. Предложенный метод также был успешно применен к разным классам изображений базы Pascal VOC, что доказало возможность обработки объектов различных классов.

Заключение. Основной вклад предложенного метода: 1) позволяет сегментировать объект интереса на уровне современных методов сегментации, а в отдельных случаях превосходит их; 2) предоставляется новый взгляд на способ прослеживания контура объекта.

Ключевые слова: сегментация изображений, выделение области интереса, поиск кратчайшего пути в графе, алгоритм A^* ; полярная система координат

Для цитирования: Зубов И. Г. Метод автоматической сегментации транспортных средств на изображении // Изв. вузов России. Радиоэлектроника. 2019. Т. 22, № 5. С. 6–16. doi: 10.32603/1993-8985-2019-22-5-6-16

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Благодарности. Автор выражает особую благодарность выпускнику Московского государственного технического университета им. Н. Э. Баумана Дмитрию Вадимовичу Боровому за оказанную помощь при проведении данного исследования, ценные советы и замечания.

Статья поступила в редакцию 11.07.2019; принята к публикации после рецензирования 30.09.2019; опубликована онлайн 29.11.2019

Method for Automatic Segmentation of Vehicles in Digital Images

Ilya G. Zubov✉

Ltd "Next", Moscow, Russia

✉ ZubovIG@gmail.com

Abstract

Introduction. Modern systems for active vehicle safety are designed to significantly reduce the number of road accidents. Sensors based on monocular cameras are increasingly being introduced by the world's leading automakers as an effective tool for improving traffic safety. Modern methods of localisation and classification, combined with semantic segmentation algorithms, allow for image division into independent groups of pixels corresponding to each object. However, the problem of developing segmentation algorithms ensuring improved quality of image segmentation remains to be solved.

Aim. To develop an automatic method for segmenting a given object during image analysis.

Materials and methods. An automatic method for segmenting vehicles in an image was proposed. The method presented herein allows semantic segmentation of the object of interest, based upon a priori information about the bounding boxes, which frame the objects in the image. Bounding box information is used to transform an image into a polar coordinate system where the pixels of the image act as the edges of a weighted graph. A closed contour is obtained around the object of interest by using the shortest path search algorithm and inverse transformation to the Cartesian coordinate system.

Results. The experiments confirmed the correctness of the selected area of interest based on this algorithm. Jacquard's similarity coefficient for the Carvana open database is 85 %. Furthermore, the proposed method was applied to different classes of images from the Pascal VOC database, thus demonstrating the ability to segment objects of other classes.

Conclusion. The main contribution of the proposed method was as follows: 1) segmentation of the object of interest at the level of modern methods, and in some cases in excess thereof; 2) the study presents a new look at the way of tracking object contours.

Key words: Image segmentation, selection of the region of interest; algorithm A^* ; polar coordinate system; shortest path in the graph

For citation: Zubov I. G. Method for Automatic Segmentation of Vehicles in Digital Images. Journal of the Russian Universities. Radioelectronics. 2019, vol. 22, no. 5, pp. 6–16. doi: 10.32603/1993-8985-2019-22-5-6-16

Conflict of interest. The author declares no conflict of interest.

Acknowledgements. The author expresses special gratitude to the graduate of Moscow State Technical University N. E. Bauman to Dmitry V. Borovoy for the assistance in conducting this study, valuable advice and comments.

Submitted 11.07.2019; accepted 30.09.2019; published online 29.11.2019

Введение. Внедрение систем технического зрения в повседневную жизнь становится все более популярным. Безопасность на дороге – одна из самых приоритетных задач. Датчики на основе монокулярных камер доказали свою эффективность в данной области. Например, система предотвращения столкновений Mobileye [1] позволяет уменьшить риск столкновения с пешеходом и иным транспортным средством (ТС), а также схода с полосы движения. К сожалению, указанная система оповещения работает лишь в светлое время суток.

Использование датчиков на основе монокулярной камеры позволяет решить большой спектр задач. Например, на основе данных о форме и площади ТС наряду с особенностями дизайна автомобиля возможно оценить размер ТС и расстояние до него. Также, применяя передовые алгоритмы классификации, детекции и сегментации дорожной сцены, учитывая возможные ограничения поведения ТС, основанные на теории движения автомобиля, возможно предсказать поведение ТС в следующий момент времени. Однако для

решения данных задач необходимо описание дорожной сцены как композиции объектов, имеющих форму, площадь, взаимное расположение.

Выигрыш в 2012 г. конкурса по классификации изображений сверточной нейросетью AlexNet, обученной А. Крижевским, И. Суткевером и Д. Хинтоном, заставил мировое сообщество по-новому взглянуть на методы анализа изображений. А. Крижевскому с коллегами удалось превзойти все классические методы компьютерного зрения, представленные в ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [2]. На фоне этого события, а также благодаря общедоступности цифровых камер, одним из наиболее широко развивающихся направлений в области машинного зрения стал анализ изображений на основе цветовой информации с использованием одной камеры. В настоящий момент системы, основанные на сверточных нейронных сетях, являются самыми точными в задачах классификации изображений и обнаружении объектов. С впечатляющим достижением нейронные сети были успешно применены к различным типам проблем, например [3–5].

Подавляющее большинство существующих детекторов объектов ориентировано на двумерную локализацию. Модель обнаружения 2D-объектов предоставляет информацию $(x; y; h; w)$, где $(x; y)$ – 2D-координаты центра прямоугольника, ограничивающего объект; h, w – высота и ширина объекта соответственно [3–10]. Принимая на вход изображение, детектор выдает координаты ограничивающих объекты прямоугольников и вероятности их принадлежности классам. В качестве примера на рис. 1 показаны ограничивающие прямоугольники, полученные детектором Yolo [3], вероятность нахождения транспортного средства в

которых превышает 0.5 (рис. 1, а) и 0.05 (рис. 1, б).

Для дальнейшего анализа обнаруженного объекта необходимо произвести сегментацию. Под сегментацией принято понимать разбиение изображения на множество непересекающихся связанных областей (сегментов). В процессе сегментации каждому пикселу изображения присваивается класс согласно некоторой характеристике или вычисленному свойству, например по цвету, яркости или текстуре. Результатом сегментации изображения является множество сегментов, которые вместе покрывают все изображение. Сегментация изображений позволяет перейти к описанию сцены как композиции объектов, имеющих форму, площадь, взаимное расположение, яркостные и текстурные признаки.

В настоящее время существует множество методов сегментации изображений, таких как [11–13]:

- а) выделение краев и областей: Рb-детектор краев;
- б) эвристические методы:
 - разрастания регионов (Region growing),
 - разделения и слияния регионов (Split & Merge),
 - водораздела (Watershed),
 - нормализованных разрезов графов;
- в) кластеризация по методам:
 - К-средних,
 - сдвига среднего (Mean shift) и развития;
- г) энергетические методы:
 - методы уровня,
 - турбопикселы (TurboPixels).

Метод. В настоящей статье в качестве объекта интереса рассматривается ТС. При сегментации детектированных объектов на изображении в качестве априорной информации полагаем положение прямоугольника, ограничивающего объект (рис. 2, а).

Для решения задачи сегментации ТС автором настоящей статьи разработан алгоритм обработки изображений на основе поиска кратчайшего пути



а



б

Рис. 1. Визуализация ограничивающих прямоугольников, полученных детектором объектов, при вероятности нахождения объектов в прямоугольнике более 0.5 (а) и более 0.05 (б)

Fig. 1. Visualization of bounding boxes predicted by an object detector, if the probability of finding objects in a rectangle is more than 0.5 (a) and more than 0.05 (b)



a



б

Рис. 2. Масштабирование выделенного объекта:

a – исходное изображение выделенного объекта; *б* – масштабированное изображение

Fig. 2. Scaling a selected object: *a* – the original image of the selected object; *б* – the scaled image

во взвешенном графе, представленном в полярной системе координат.

Основные шаги алгоритма.

Шаг 1. Масштабирование изображения.

Изображение масштабируется таким образом, чтобы соотношение сторон изображения составляло 1:1 (рис. 2, б). В качестве метода масштабирования используется билинейная интерполяция [14]. Масштабирование позволит центрировать полярный полюс в каждой точке (см. шаг 4).

Шаг 2. Обработка изображения оператором детектирования границ Кэнни [15]. Основные этапы этого алгоритма:

1. К изображению применяется фильтр Гаусса:

$$B = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \times A,$$

где *A* – матрица пикселей изображения.

2. Вычисляются проекции градиента изображения на координаты:

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \times A;$$

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \times A,$$

а также направление градиента:

$$\theta = \arctg(G_y/G_x).$$

3. Полученное значение направления градиента округляется до одного из четырех углов: 0, 45, 90 и 135° [15].

4. Выделение пикселей, градиент в которых является локальным максимумом относительно

соседних пикселей. Эти пиксели считаются кандидатами на формирование границы объекта.

5. Двухпороговая фильтрация, позволяющая разбить все выделенные пиксели фрагмента изображения на три множества:

- множество пикселей, значения градиента в которых превосходит верхний порог;
- множество пикселей, значение градиента в которых меньше нижнего порога;
- множество пикселей, значения градиента в которых находятся в интервале между порогами.

Пиксеты первого множества относятся к границам объекта, пиксеты второго составляют неграницные области фона или объекта. Решения в отношении пикселей третьего множества принимаются в результате дальнейшей обработки.

6. Пиксеты третьего множества относятся к границе объекта, если соседствуют с пиксетами границ. Если же эти пиксеты окружены только неграницными пиксетами, они относятся к неграницным.

7. Границы окончательно определяются в результате операции их трассировки. При этом толщина границ сводится до одного пиксела, заполняются разрывы, обрабатываются ветвления границ. Трассировка осуществляется совокупным анализом окрестности каждого пиксела границы. В результате детектором Кэнни формируется окончательное представление границ объекта исходного изображения (рис. 3).

Шаг 3. Полярное преобразование. Для описания положения точки $M(x, y)$ в полярных координатах r и φ вокруг центра прямоугольника, ограничивающего детектированный объект, используют следующее преобразование:

$$r = \sqrt{x^2 + y^2}; \quad (1)$$

$$\operatorname{tg} \varphi = y/x, \quad (2)$$



Рис. 3. Результат обработки объекта рис. 2, б детектором Кэнни

Fig. 3. The result of processing the object Fig. 2, b used Canny detector

где r – полюс (расстояние от точки M до начала координат); φ – угол, образованный лучом OM с полярной осью.

В качестве начала координат используется центр прямоугольника, ограничивающего детектированный объект, т. е. (1) и (2) имеют следующий вид:

$$r = \sqrt{(x + w/2)^2 + (y + h/2)^2}; \quad (3)$$

$$\operatorname{tg} \varphi = (2y + h)/(2x + w), \quad (4)$$

где h и w – высота и ширина ограничивающего прямоугольника соответственно.

Воспользовавшись (3) и (4), преобразуем изображение из декартовой системы координат в полярную (рис. 4). После этого преобразования контур объекта расположен в области $0 \leq \varphi < 2\pi$.

Шаг 4. Поиск кратчайшего пути во взвешенном графе. Представим изображение в виде графа, вершинами которого являются пиксели изображения в полярной системе координат. Исходя из того, что детектированный объект занимает большую часть изображения, его внешний контур будет образован пикселями, имеющими наибольшие значения радиуса. Тогда веса ребер, разделяющих два пиксела, в точке с координатами φ, r можно представить следующим образом:

$$\operatorname{weight}(\varphi, r) = \begin{cases} r_{\max}, & \text{if } I(\varphi, r) = 0; \\ r_{\max} - r, & \text{if } I(\varphi, r) \neq 0, \end{cases}$$

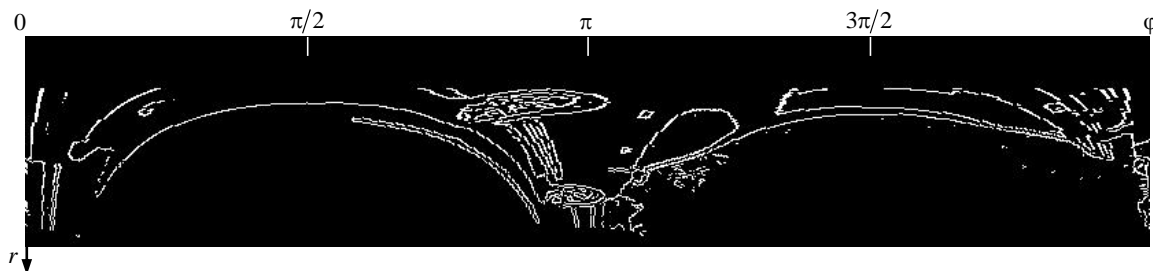


Рис. 4. Результат полярного преобразования, представленный в прямоугольной системе координат

Fig. 4. The result of a polar transformation represented in a rectangular coordinate system

где $I(\varphi, r)$ – интенсивность пиксела в точке с координатами φ, r .

Исходя из того, что внешний контур объекта расположен в области $0 \leq \varphi < 2\pi$, а веса графа зависят от значения полярного радиуса и интенсивности пикселей, задачу сегментации можно представить как поиск кратчайшего пути во взвешенном графе. Обход графа начинается с вершины $(\varphi = 0, r)$ и завершается на вершине $(\varphi = 2\pi, r)$.

Из существующего разнообразия методов, использующих теорию графов, был выбран алгоритм поиска кратчайшего пути A^* [16]. Этот алгоритм находит путь наименьшей стоимости от заданного начального до целевого узла (из одной или нескольких возможных целей). A^* следует по пути наименьших известных эвристических затрат:

$$f(v) = g(v) + j(v),$$

где v – текущая вершина; $g(v)$ – наименьшее расстояние от стартовой вершины до текущего положения; $j(v)$ – эвристическая функция (манхэттенское расстояние) приближения расстояния от текущего местоположения до конечной цели.

На рис. 5 представлен результирующий путь с наименьшим весом, рассчитанный по алгоритму A^* для изображения с рис. 4.

Шаг 5. Преобразование полученного пути в декартову систему координат с последующей заливкой. Используя обратное преобразование выражений (3), (4), переведем полученный путь в декартову систему координат. Так как кратчайший путь, полученный на шаге 4, представляет собой замкнутый контур, описывающий объект интереса, для получения маски объекта воспользуемся заливкой области, расположенной внутри данного контура. Сопоставив маску с исходным изображением, получим сегментированное изображение (рис. 6).

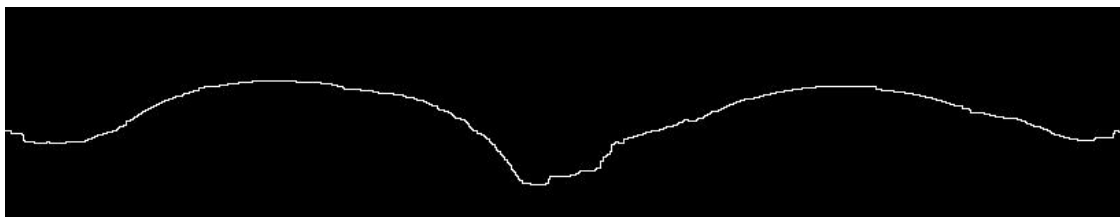


Рис. 5. Результат выделения кратчайшего пути по графу

Fig. 5. The result of selection of the shortest path through the graph



Рис. 6. Результат работы алгоритма

Fig. 6. The result of the algorithm operation

Результаты. Для оценки качества разработанного алгоритма получаемые им результаты сегментации сравнивались с результатами, даваемыми тремя распространенными методами сегментации изображения: K-Means [17], GrabCut [18] и Mask-RCNN [19].

K-Means – алгоритм кластеризации, основанный на разбиении множества элементов векторного пространства на заранее определенное число кластеров с минимизацией среднеквадратического отклонения на точках каждого кластера. На каждой итерации вычисляется центр масс для каждого кластера на основе центров, полученных на прошедшей итерации. После этого элементы

векторного пространства вновь разбиваются на кластеры в соответствии с ближайшим расстоянием до новых центров. Алгоритм завершается, если на очередной итерации центры кластеров остаются неизменными.

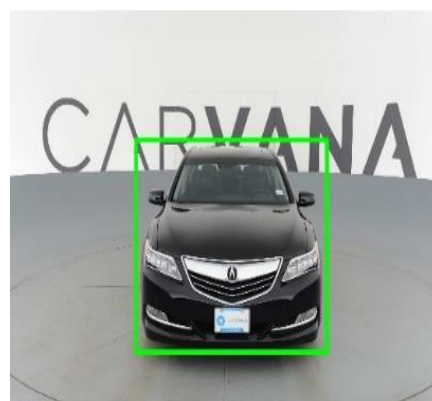
GrabCut [18] – метод сегментации изображения, основанный на алгоритме GraphCut [20]. GrabCut расширяет возможности GraphCut на обработку цветных изображений. Вначале совокупность цветов пикселей внутри и снаружи детектируемого объекта аппроксимируется смесью гауссовских величин, представляющих пиксели целевого объекта и фона. Полученная модель используется для построения марковского случайного поля с энергетической функцией, выделяющей соединенные пиксели, имеющие один и тот же класс. После чего запускается оптимизационный метод, основанный на минимальном разрезе графа.

Mask-RCNN [19] – архитектура современной нейронной сети для сегментации объектов на изображениях. Ее можно представить в виде следующих модулей:

– генератор особенностей (features extractor), формирующий трехмерную матрицу особенностей входного изображения, полученную сверточной нейронной сетью ResNet-50 [21];



а



б

Рис. 7. Предварительная обработка изображений базы Carvana:

а – исходное изображение; б – масштабированное изображение с выделенной областью объекта

Fig. 7. Preprocessing of the image of the Carava database:

а – the original image; б – the scaled image with the selected area of the object

– Region Proposal Network – сеть генерации регионов присутствия объектов;

– полносвязные слои – это сеть, которая для каждого региона вырезает из матрицы особенностей соответствующую этому региону часть и выдает класс объекта и уточненный описывающий объект прямоугольник;

– генерация бинарных масок внутри регионов присутствия объектов.

Для оценки предложенного метода и сравнения с перечисленными алгоритмами использовалась база сегментированных изображений Carvana [22], содержащая 5088 изображений ТС различных классов, а также маски для каждого изображения. Каждое изображение было масштабировано для приведения к единому количеству пикселей

500×500 (рис. 7, а). Исходя из бинарной маски изображения, выделялась прямоугольная область, в которой находится объект интереса (рис. 7, б). Изображение, заключенное внутри рамки, ограничивающей эту область, использовалось в качестве входных данных для алгоритмов сегментации.

Таблица 1. Сравнительные результаты сегментации изображений базы Carvana

Table 1. Comparative results of the segmentation of the database Carvana images

Метод сегментации/ Segmentation method	Бинарная мера сходства Жаккара/ A binary similarity measure of Jaccard, %
K-Means	0.55
GrabCut	0.68
Mask-RCNN	0.66
Предложенный метод/ Proposed method	0.85

Таблица 2. Результаты работы алгоритма сегментации детектированных транспортных средств

Table 2. The results of the algorithm for segmentation of detected vehicles







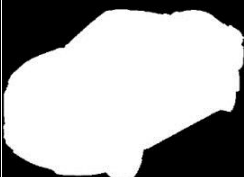

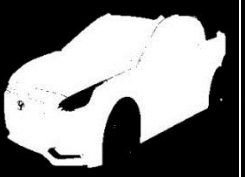
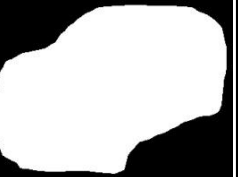

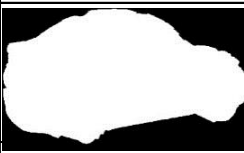

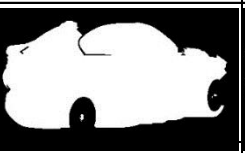
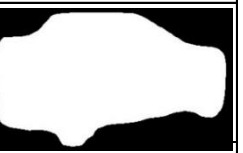

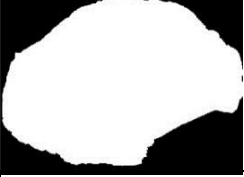

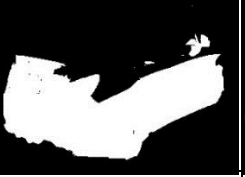

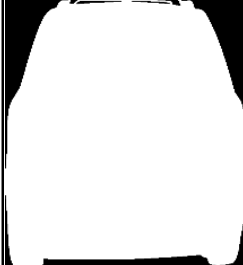
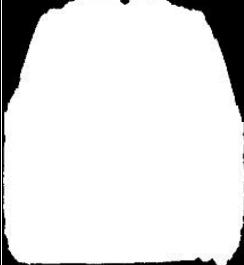


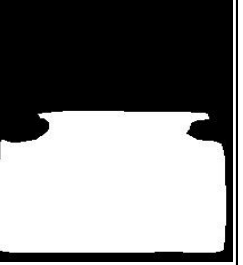

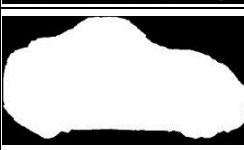
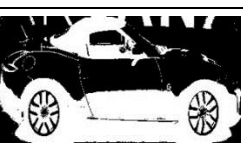
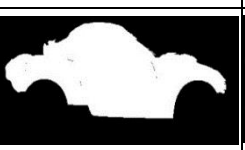
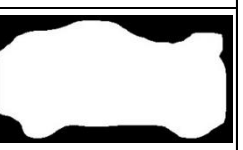
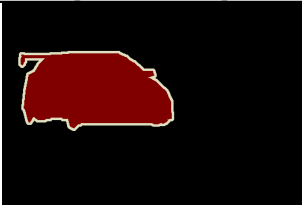





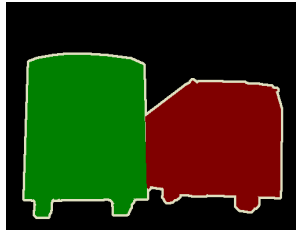




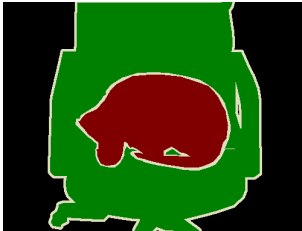
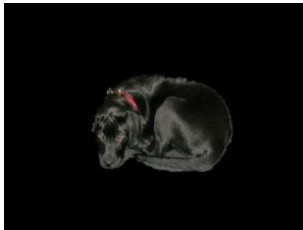

Исходная маска объекта	Предложенный метод	K-Means	GrabCut	Mask-RCNN
				
				
				
				
				
				

Таблица 3. Результат применения алгоритма к изображениям базы Pascal VOC

Table 3. The result of applying the algorithm to the images of the Pascal VOC base

Исходное изображение базы данных VOC	Сегментированное изображение	Результат работы алгоритма
		
		
		
		
		
		
		

В качестве меры сходства полученных масок сегментированных объектов использовалась бинарная мера сходства Жаккара:

$$J = |A \cap B| / |A \cup B|,$$

где A, B – бинарные маски изображения, полученного алгоритмом сегментации, и исходного

изображения соответственно.

Алгоритмы сегментации применялись ко всем изображениям базы. Полученные результаты представлены в табл. 1.

Бинарные изображения, представленные в табл. 2, показывают результаты работы алгоритмов автоматической сегментации детектированных ТС.

Предложенный метод также применялся к изображениям базы Pascal VOC [23]. Изображения в табл. 3 представляют результат применения

предложенного метода к разным классам объектов на изображениях.

Вывод. В статье представлен новый метод автоматической сегментации транспортных средств на изображении. Работоспособность и конкурентоспособность метода по отношению к известным алгоритмам сегментации K-Means, GrabCat, Mask-RCNN проверены его тестированием на базе изображений Carvana. Метод, также успешно примененный к изображениям базы Pascal VOC, показал возможность сегментации объектов разных классов.

Список литературы

1. Forward Collision Warning with a Single Camera / E. Dagan, O. Mano, G. P. Stein, A. Shashua // Proc. of the IEEE Intelligent Vehicles Symp. Parma, Italy, 14–17 June 2004. Piscataway: IEEE, 2004. P. 37–42. doi: 10.1109/IVS.2004.1336352
2. Imagenet: A Large-Scale Hierarchical Image Database / J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei // IEEE Conf. on Computer Vision and Pattern Recognition 2009. Miami, FL, USA, 20–25 June 2009. Piscataway: IEEE, 2009. P. 248–255. doi: 10.1109/CVPR.2009.5206848
3. You Only Look Once: Unified, Real-Time Object Detection / J. Redmon, S. Divvala, R. Girshick, A. Farhadi. URL: <https://arxiv.org/pdf/1506.02640.pdf> (дата обращения 29.08.2019)
4. Girshick R. Fast R-CNN // IEEE Intern. Conf. on Computer Vision (ICCV), 2015. URL: <https://arxiv.org/pdf/1504.08083.pdf> (дата обращения 29.08.2019)
5. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. URL: <https://arxiv.org/pdf/1409.1556.pdf> (дата обращения 02.09.2019)
6. Object Detection with Discriminatively Trained Part Based Models / P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan // IEEE trans. on pattern analysis and machine intelligence (PAMI). 2010. № 9. URL: <http://cs.brown.edu/people/pfelzens/papers/lsvm-pami.pdf> (дата обращения 29.08.2019)
7. Multiple Kernels for Object Detection / A. Vedaldi, V. Gulshan, M. Varma, A. Zisserman // 2009 IEEE 12th Intern. Conf. on Comp. Vision. Kyoto, Japan, 29 Sept.–2 Oct. 2009. Piscataway: IEEE, 2009. P. 606–613. doi: 10.1109/ICCV.2009.5459183
8. Viola P., Jones M. Rapid Object Detection Using a Boosted Cascade of Simple Features // CVPR, 2001. Kauai, HI, USA, 8–14 Dec. 2001. Piscataway: IEEE, 2001. URL: <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf> (дата обращения 27.08.2019)
9. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation / R. Girshick, J. Donahue, T. Darrell, J. Malik // In Proc. of the IEEE Conf. on Comp. vision and pattern recognition. Columbus, USA, 23–28 June 2014. Piscataway: IEEE, 2014. P. 580–587. doi: 10.1109/CVPR.2014.81
10. Overfeat: Integrated Recognition, Localization and Detection Using Convolutional Networks. URL: <https://arxiv.org/pdf/1312.6229v4.pdf> (дата обращения 20.08.2019)
11. A Review of Computer Vision Segmentation Algorithms. URL <https://courses.cs.washington.edu/courses/cse576/12sp/notes/remote.pdf> (дата обращения 20.08.2019)
12. Yuheng S., Yan Hao. Image Segmentation Algorithms Overview // Computer Vision and Pattern Recognition. 2017. URL: <https://arxiv.org/ftp/arxiv/papers/1707/1707.02051.pdf> (дата обращения 20.08.2019)
13. Jyotsana M., Nirvair N. A Brief Review: Super-Pixel Based Image Segmentation Methods // Imperial J. of Interdisciplinary Research. 2016. Vol. 2, iss. 9. P. 8–12. URL: <https://pdfs.semanticscholar.org/4aee/70a322c01bc4dfd51c3164e480c3984b8071.pdf> (дата обращения 20.08.2019)
14. Щерба Е. В. Анализ применимости методов интерполяции и экстраполяции для решения задачи восстановления изображения // Компьютерная оптика. 2009. Т. 33, № 3. С. 336–339. URL: <http://www.computeroptics.smr.ru/KO/PDF/KO33-3/33313.pdf> (дата обращения 20.08.2019).
15. Гонсалес Р., Вудс. Р. Цифровая обработка изображений. 3-е изд. М.: Техносфера, 2012. 834 с.
16. Dechter R., Pearl J. Generalized Best-First Search Strategies and the Optimality of A* // J. of the ACM (JACM). 1985. Vol. 32, № 3. P. 505–536. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3090&rep=rep1&type=pdf> (дата обращения 20.08.2019)
17. Bandyopadhyay S., Maulik U. An Evolutionary Technique Based on K-Means Algorithm for Optimal Clustering in RN // Information Sciences. 2002. Vol. 146, iss. 1–4. P. 221–237. doi: 10.1016/S0020-0255(02)00208-6
18. Rother C., Kolmogorov V., Blake A. “GrabCut” – Interactive Foreground Extraction Using Iterated Graph Cuts // ACM Trans. on Graphics. 2004. Vol. 23. P. 309–314.
19. Mask R-CNN / K. He, G. Gkioxari, P. Dollar, R. Girshick // Computer Vision and Pattern Recognition. URL: <https://arxiv.org/pdf/1703.06870.pdf> (дата обращения 20.08.2019)
20. Graph cut based image segmentation with connectivity priors. URL: <https://pub.ist.ac.at/~vnk/papers/connectedGC-CVPR08.pdf> (дата обращения 28.10.2019).

21. Deep Residual Learning for Image Recognition. URL: <https://arxiv.org/pdf/1512.03385.pdf> (дата обращения 20.08.2019).

22. Carvana Image Masking Challenge. URL: <https://www.kaggle.com/c/carvana-image-masking-challenge> (дата обращения 20.08.2019)

23. The PASCAL Visual Object Classes Challenge (VOC2007). URL: <http://www.pascal-network.org/challenges/VOC/voc2007/index.html> (дата обращения 20.08.2019)

Информация об авторе

Зубов Илья Геннадьевич – магистр техники и технологий (2016), программист-алгоритмист компании ООО "НЕКСТ". Автор 4 научных публикаций. Сфера научных интересов – цифровая обработка изображений; прикладные телевизионные системы.

Адрес: ООО "НЕКСТ", ул. Рочдельская, д. 15, стр. 13, Москва, 123022, Россия

E-mail: ZubovIG@gmail.com

<https://orcid.org/0000-0003-0407-5651>

References

1. Dagan E., Mano O., Stein G. P., Shashua A. Forward Collision Warning with a Single Camera. Proc. of the IEEE Intelligent Vehicles Symp. Parma, Italy. 14–17 June 2004. Piscataway, IEEE, 2004, pp. 37–42. doi: 10.1109/IVS.2004.1336352

2. Deng J., Dong W., Socher R., Li L.-J., Li K., Fei-Fei L. Imagenet: A Large-Scale Hierarchical Image Database. IEEE Conf. on Computer Vision and Pattern Recognition 2009. Miami, FL, USA, 20–25 June 2009. Piscataway, IEEE, 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848

3. Redmon J., Divvala S., Girshick R., Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. 2015. Available at: <https://arxiv.org/pdf/1506.02640.pdf> (accessed 29.08.2019)

4. Girshick R. Fast R-CNN. IEEE Intern. Conf. on Computer Vision (ICCV), 2015. Available at: <https://arxiv.org/pdf/1504.08083.pdf> (accessed 29.08.2019)

5. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. Available at: <https://arxiv.org/pdf/1409.1556.pdf> (accessed 02.09.2019)

6. Felzenszwalb P., Girshick R., McAllester D., Ramanan D. Object Detection with Discriminatively Trained Part Based Models. IEEE trans. on pattern analysis and machine intelligence (PAMI). 2010, no. 9. Available at: <http://cs.brown.edu/people/pfelzens/papers/lsvm-pami.pdf> (accessed 29.08.2019)

7. Vedaldi A., Gulshan V., Varma M., Zisserman A. Multiple Kernels for Object Detection. 2009 IEEE 12th Intern. Conf. on Comp. Vision. Kyoto, Japan, 29 Sept.–2 Oct. 2009. Piscataway, IEEE, 2009, pp. 606–613. doi: 10.1109/ICCV.2009.5459183

8. Viola P., Jones M. Rapid Object Detection Using a Boosted Cascade of Simple Features. CVPR, 2001. Kauai, HI, USA, 8–14 Dec. 2001. Piscataway, IEEE, 2001. Available at: <https://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/viola-cvpr-01.pdf> (accessed 27.08.2019)

9. Girshick R., Donahue J., Darrell T., Malik J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proc. of the IEEE Conf. on Comp. vision and

pattern recognition. Columbus, USA, 23–28 June 2014. Piscataway, IEEE, 2014, pp. 580–587. doi: 10.1109/CVPR.2014.81

10. Overfeat: Integrated Recognition, Localization and Detection Using Convolutional Networks. Available at: <https://arxiv.org/pdf/1312.6229v4.pdf> (accessed 20.08.2019)

11. A Review of Computer Vision Segmentation Algorithms. Available at: <https://courses.cs.washington.edu/courses/cse576/12sp/notes/remote.pdf> (accessed 20.08.2019)

12. Yuheng S., Yan Hao. Image Segmentation Algorithms Overview. Computer Vision and Pattern Recognition. 2017. Available at: <https://arxiv.org/ftp/arxiv/papers/1707/1707.02051.pdf> (accessed 20.08.2019)

13. Jyotsana M., Nirvair N. A Brief Review: Super-Pixel Based Image Segmentation Methods. Imperial J. of Interdisciplinary Research. 2016, vol. 2, iss. 9, pp. 8–12. Available at: <https://pdfs.semanticscholar.org/4aee/70a322c01bc4dfd51c3164e480c3984b8071.pdf> (accessed 20.08.2019)

14. Shcherba E. V. Application Analysis of Interpolation and Extrapolation Methods as Used for Image Restoration. Computer Optics. 2009, vol. 33, no. 3, pp. 336–339. Available at: <http://www.computeroptics.smr.ru/KO/PDF/KO33-3/33313.pdf> (accessed 20.08.2019). (In Russ.)

15. Gonsales R., Vuds R. *Tsifrovaya obrabotka izobrazhenii* [Digital Image Processing]. Moscow, Tekhnosfera, 2012, 834 p. (In Russ.)

16. Dechter R., Pearl J. Generalized Best-First Search Strategies and the Optimality of A*. Journal of the ACM (JACM). 1985, vol. 32, no. 3, pp. 505–536. Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.3090&rep=rep1&type=pdf> (accessed 20.08.2019)

17. Bandyopadhyay S., Maulik U. An Evolutionary Technique Based on K-Means Algorithm for Optimal Clustering in RN. Information Sciences. 2002, vol. 146, iss. 1–4, pp. 221–237. doi: 10.1016/S0020-0255(02)00208-6

18. Rother C., Kolmogorov V., Blake A. "GrabCut" - Interactive Foreground Extraction Using Iterated Graph Cuts. ACM Trans. on Graphics. 2004, vol. 23, pp. 309–314.

19. He K., Gkioxari G., Dollar P., Girshick R. Mask R-CNN. Computer Vision and Pattern Recognition. Available at: <https://arxiv.org/pdf/1703.06870.pdf> (accessed 20.08.2019)

20. Graph cut based image segmentation with connectivity priors. Available at: <https://pub.ist.ac.at/~vnk/papers/connectedGC-CVPR08.pdf> (accessed 28.10.2019)

21. Deep Residual Learning for Image Recognition. Available at: <https://arxiv.org/pdf/1512.03385.pdf> (accessed 20.08.2019)

22. Carvana Image Masking Challenge. Available at: <https://www.kaggle.com/c/carvana-image-masking-challenge> (accessed 20.08.2019)

23. The PASCAL Visual Object Classes Challenge (VOC2007). URL: <http://www.pascal-network.org/challenges/VOC/voc2007/index.html> (accessed 20.08.2019)

Information about the author

Илья Г. Zubov, Master of Engineering and Technology (2016), Ltd "Next" algorithm programmer. The author of 4 scientific publications. Area of expertise: digital image processing; applied television systems. Address: Ltd "Next", 15 Rochdelskaya st., bldg. 13, Moscow 123022, Russia
E-mail: ZubovIG@gmail.com
<https://orcid.org/0000-0003-0407-5651>



Книжные новинки

Кузнецов В. В., Москвин П. П.

МЕЖФАЗНЫЕ ВЗАИМОДЕЙСТВИЯ ПРИ ГЕТЕРОЭПИТАКСИИ ПОЛУПРОВОДНИКОВЫХ ТВЕРДЫХ РАСТВОРОВ

ISBN 978-5-8114-3809-9

СПб: Лань, 2019. 376 с. : ил.

Издательство "ЛАНЬ": lan@lanbook.ru; www.lanbook.com

Магазин электронных книг Global F5: <http://globalf5.com>

В книге на основе различных приближений теории регулярных растворов и модели диффузионного массопереноса рассмотрены особенности эпитаксии твердых растворов на основе полупроводниковых соединений A^3B^5 и A^2B^6 . Проанализировано влияние упругих деформаций на смещение фазовых равновесий в многокомпонентных системах. Изложены методики расчета равновесных и когерентных диаграмм состояния многокомпонентных систем. Дано математическое описание эффекта стабилизации периода решетки и кинетики кристаллизации многокомпонентных твердых растворов. Рассмотрены критические явления и термодинамическая устойчивость подложки в неравновесной жидкой фазе. Особое внимание уделено процессам получения изопериодических гетероструктур на основе четверных и пятерных твердых растворов, которые широко применяются в различных приборах полупроводниковой оптоэлектроники.