

## Суперпиксель-сегментация на основе минимизации энергии и свертки с ядром геодезического расстояния

В.Н. Карнаухов\*, В.И. Кобер\*, М.Г. Мозеров\*, Л.В. Зимина\*\*

\*Институт проблем передачи информации, Российская академия наук, Москва, 127051, Россия

\*\*Московский политехнический университет, Москва, 107023, Россия

Поступила в редколлегию 17.04.2023

**Аннотация**—Метод минимизации энергии или максимальной апостериорной вероятности (МAB) является основой для решения многих задач компьютерного зрения, в том числе и проблемы сегментации. Однако подразумевается, что число областей при сегментации достаточно мало. В то же время в задаче суперпиксель-сегментации или иначе избыточной сегментации число таких областей превышает 1000, что делает вычислительную задачу оптимизации методом МAB практически невыполнимой. В данной работе мы предложили решение, которое сводит сегментацию с любым количеством областей к проблеме маркировки всего девяти меток. Кроме того, для усиления робастности оптимизации используется свертка с ядром геодезического расстояния. Это позволяет получать на выходе алгоритма односвязанные суперпиксели, в отличие от многих других методов, требующих дополнительной корректировки. Эффективность предложенного метода сравнивается и измеряется по критерию совпадения границ областей, а также с помощью визуальной иллюстрации.

**КЛЮЧЕВЫЕ СЛОВА:** Суперпиксель-сегментация, фильтр на основе геодезического расстояния, минимизация энергии, максимальная апостериорная вероятность (МAB).

DOI: 10.53921/18195822\_2023\_23\_2\_208

### 1. ВВЕДЕНИЕ

Целью суперпиксельной сегментации является создание согласованной группы пикселей, которая также известна как избыточная сегментация изображения [1, 2, 3]. В последнее время суперпиксельная сегментация привлекла большой интерес особенно в области компьютерного зрения, поскольку она обеспечивает удобный способ вычисления характеристик изображения и снижает сложность последующих задач обработки изображений. Для решения задачи избыточной сегментации было предложено множество алгоритмов [3, 4, 5, 6, 7, 8, 9, 10].

Ценность суперпиксельной сегментации заключается в том, что этот подход находит применение во многих областях компьютерного зрения, включая распознавание объектов [11], сегментацию изображений [12], отслеживание объектов [13], сегментацию видео [14] и классификацию [15]. Это обусловлено тем, что суперпиксель по сути является локальным структурным элементом изображения, или иначе локальным признаком, что позволяет использовать такой признак для дальнейшего анализа изображения.

В этой работе мы предлагаем использовать подход минимизации энергии, отлично зарекомендовавший себя во многих областях компьютерного зрения. Если рассматривать для примера решение проблемы сегментации, то суть подхода заключается в следующем. Предполагается, что каждый пиксель из  $P$  точек изображения может быть отнесен к одному из  $S$  дискретных значений, называемых номером сегмента. Для решения задачи необходимо оценить вероятность того, что данный пиксель принадлежит каждому из возможных номеров

сегментов. То есть на предварительном этапе мы должны получить избыточное число данных с размерностью  $P \times S$ . Затем, используя эти данные, нужно выбрать оптимальное решение, которое максимизирует совокупную вероятность, или минимизирует энергию (совокупность затрат).

В классической задаче сегментации число сегментов на задачу не превышает пятидесяти. Тогда как в суперпиксель-сегментации это число может быть очень большим и ее трудно реализовать из-за большого количества требуемых областей сегментации (для больших изображений это число превышает 2000 сегментов), таким образом делая вычислительную задачу оптимизации методом МАВ практически невыполнимой. Поэтому мы предлагаем совершенно новый подход к проблеме суперпиксель-сегментации с использованием подхода минимизации энергии. Мы предложили решение, которое сводит сегментацию с любым количеством областей к проблеме маркировки всего девяти меток, используя топологию односвязности. Кроме того, для усиления робастности оптимизации используется свертка с ядром геодезического расстояния. Это позволяет получать на выходе алгоритма односвязанные суперпиксели, в отличие от многих других методов, требующих дополнительной корректировки.

За основу мы взяли реализацию фильтра предложенную в [16, 17]. Выбор фильтра обусловлен историей развития класса фильтров с сохранением границ, так как прогресс в восстановлении зашумленных изображений и улучшении качества деталей связан, в первую очередь, с появлением билатеральных фильтров [18]. В этом случае окрестность усреднения определяется не только пространственной близостью к восстанавливаемому пикселю, но и близостью в области значений. Такие фильтры широко используются в задачах обработки изображения и компьютерного зрения, связанных с минимизацией энергии, таких как стерео [19, 20], стерео в движении [21], оценка оптического потока [22], улучшение изображений и контрастирование [23]. Однако существенным недостатком фильтров с билатеральным ядром является квадратичная зависимость вычислительной сложности от числа пикселей на изображении. И хотя были предложены быстрые методы [24, 25], они достаточно сложны в реализации, выход фильтра является неточной аппроксимацией истинного билатерального фильтра, и вычислительная сложность зависит от параметров фильтра. В данной статье мы выбрали подкласс билатеральных фильтров, который использует ядро свертки на основе геодезического расстояния. Основанием для такого выбора является то, что такое ядро свертки допускает рекурсивное вычисление, а следовательно, быструю обработку изображений.

Эффективность предложенного метода сравнивается и измеряется по критерию совпадения границ областей, а также с помощью визуальной иллюстрации. Анализ результата работы предложенного алгоритма показал, что метод показывает уровень state-of-the-art при этом скорость вычисления одна из самых быстрых для данного типа алгоритмов.

Статья организована следующим образом: первый раздел — Введение, в разделе 2 предложен новый алгоритм суперпиксель-сегментации, основанный на минимизации энергии и на свертке с ядром геодезического расстояния, в разделе 3 приведены экспериментальные результаты и, наконец, раздел Заключение суммирует наши выводы.

## 2. ПОСТАНОВКА ЗАДАЧИ

В общем виде задача сегментации изображения  $I(\mathbf{x})$  формулируется следующим образом: найти функцию  $s(\mathbf{x})$ , ставящую в соответствие каждому пикселю  $\mathbf{x}$  (вектор координаты точки на изображении) некоторое натуральное число, означающее номер сегмента. При этом пиксель на изображении должен обладать некоторыми свойствами, по которым этот пиксель можно было бы отнести к тому или иному сегменту. Обычно это некоторое расстояние в пространстве близости  $C(\mathbf{x}, s)$  или функция рассогласования (цены). Для суперпиксель-сегментации добавляется еще два важных условия: а) все пиксели одного сегмента должны образовывать

односвязанное множество; б) площадь суперпикселей должна быть примерно одинакова, а карта сегментов должна представлять из себя подобие регулярной сетки. Подход минимизации глобальной энергии заключается в нахождении функции сегментации функции  $s(\mathbf{x})$ , которая минимизирует функционал энергии в пространстве близости. Глобальная энергия, как функционал, обычно содержит два члена: член данных и член гладкости:

$$E(s(\mathbf{x})) = \sum_{\mathbf{x} \in \Omega} C(\mathbf{x}, s(\mathbf{x})) + \sum_{\mathbf{x} \in \Omega} B(s(\mathbf{x})), \quad (1)$$

где  $B$  — функция гладкости;  $\Omega$  — домен координат пикселя  $\mathbf{x} = \{x_1, x_2, \dots, x_N\}$ .

Значение веса схожести  $C(\mathbf{x}, s)$  в пространстве близости вычисляется следующим образом:

$$C(\mathbf{x}, s) = 1 - \exp\left(-\frac{|I(\mathbf{x}) - I(s)|}{2\mu^2}\right), \quad (2)$$

где  $I(\mathbf{x})$  — цветовое значения в точке  $\mathbf{x}$ ,  $I(s)$  — цветовое значение присущее данному суперпикселю (как правило максимум локальной гистограммы цветности) и  $\mu$  — параметр алгоритма. Экспоненциальная форма для оценки соответствия пикселя признаку сегмента обусловлена необходимостью нормализации этой величины в диапазоне от нуля до единицы.

Функция  $B$  гладкости в (1) задана как:

$$B(s(\mathbf{x})) = \begin{cases} \lambda s(\mathbf{x}) = s(\mathbf{x} \pm \Delta\mathbf{x}) \\ 0 \text{ elsewhere} \end{cases}, \quad (3)$$

где  $\lambda$  — параметр алгоритма. Подобная функция гладкости называется потенциалом Потса и смысл подобной функции гладкости заключается в следующем — вероятность того, что соседний пиксель принадлежит тому же сегменту намного превосходит вероятность того, что означенный пиксель принадлежит другому классу.

Решение оптимизационного уравнения:

$$s(\mathbf{x}) = \arg \min_{s(\mathbf{x})} \left\{ \sum_{\mathbf{x} \in \Omega} C(\mathbf{x}, s(\mathbf{x})) + \sum_{\mathbf{x} \in \Omega} B(s(\mathbf{x})) \right\}, \quad (4)$$

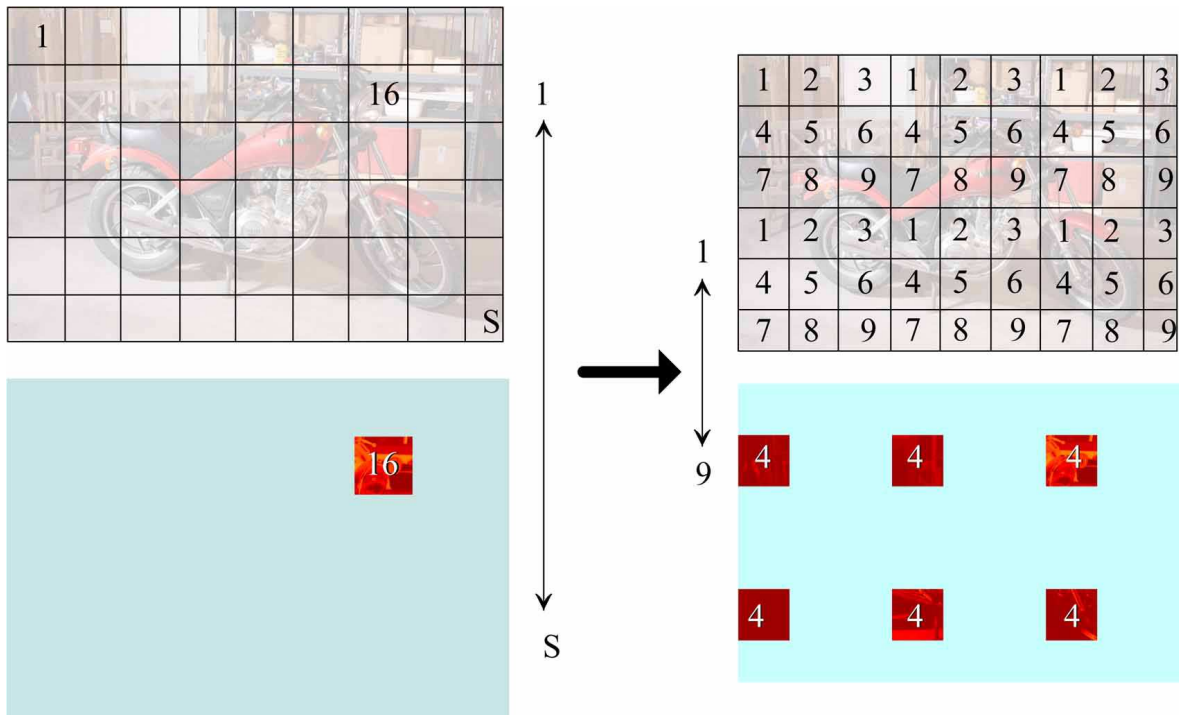
может быть получено с помощью свертки с ядром геодезического расстояния:

$$\hat{C}(\mathbf{x}, s) = \frac{\sum_{\mathbf{x} \in \Omega} w_{\mathbf{x}} C(\mathbf{x}, s)}{\sum_{\mathbf{x} \in \Omega} w_{\mathbf{x}}}, \quad (5)$$

где  $w_{\mathbf{x}}$  — веса ядра свертки в пространстве близости геодезического расстояния. Поэтому вначале рассмотрим теоретические принципы вычисления весов ядра геодезического расстояния.

$$w_{\mathbf{x}} = w_p = e^{-ad_{p,q}} \quad (6)$$

где  $p, q, k, l \in V$  — вершины графа изображения  $G$ , множество  $\varepsilon = (k, l) \in E$  — границы этого графа  $G = \{V, E\}$ . Веса ядра свёртки  $e^{-ad_{p,q}}$  в (6) определены через геодезическое расстояние



**Рис. 1.** Схема преобразования объема данных весов с произвольным числом уровней-сегментов к унифицированной системе с девятью значениями меток.

между вершинами  $(p, q)$ , которое может быть определено на изображении  $I_p$  с использованием регулярной решетки дискретизации:

$$d_{p,q} = \min_{P_{p,q}} \sum_{\varepsilon \in P_{p,q}} u_{\varepsilon}, \tag{7}$$

$$u_{\varepsilon=(k,l)} = \|I_k - I_l\| + \delta,$$

где  $P_{p,q}$  — кратчайший путь между точками  $(p, q)$ , а  $\delta$  — компонента обычного евклидова расстояния. Фильтруемый сигнал  $f_p$  в (6) может совпадать с изображением  $I_p$ , однако чаще это изображение  $I_p$  только определяет пространство близости и, в конечном итоге, тип ядра свертки.

Заметим, что параметры фильтра  $a$  и  $\delta$  в (7) приблизительно соответствуют параметрам классического билатерального фильтра следующим образом:

$$a = \frac{2}{\sigma_r^2}, \delta = \frac{\sigma_r^2}{\sigma_s^2}, \tag{8}$$

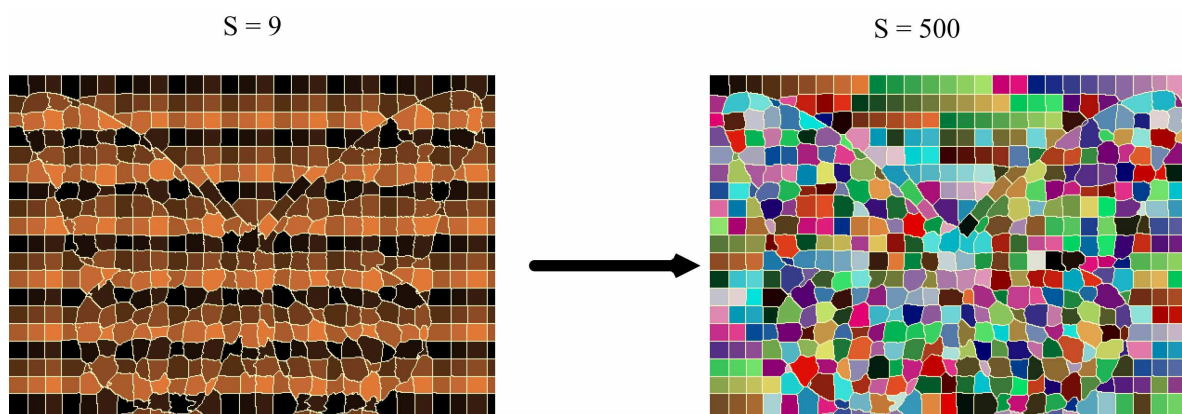
где  $\sigma_r^2$  и  $\sigma_s^2$  — дисперсия в области значений сигнала и дисперсия в пространстве изображения, соответственно. В наших экспериментах параметры фильтра:  $\sigma_r = \sigma_s = 15$ .

Итак, полный алгоритм суперпиксель-сегментации состоит из следующих шагов:

- Вычисление избыточного объема данных  $C(\mathbf{x}, s)$  с помощью (2);
- Фильтрация избыточного объема данных  $C(\mathbf{x}, s)$  с помощью (5);
- Получение окончательного решения  $s(\mathbf{x}) = \arg \min_s \hat{C}(\mathbf{x}, s(\mathbf{x}))$ .

Как видно из описания выше, объем вычислений зависит от числа сегментов. И если это число превышает 1000 возникают почти непреодолимые проблемы с оперативной памятью компьютера и временем вычисления.

В статье предлагается решение, которое сводит сегментацию с любым количеством областей к проблеме маркировки всего девяти меток. Принцип решения иллюстрируется на Рис. 1. В правой части Рис. 1 представлена классическая схема решения, когда каждому сегменту соответствует свой слой функции веса схожести. Каждый слой весов фильтруется с помощью (5). Так как веса схожести нетривиальны лишь в небольшой области всего домена изображения, то налицо высокая избыточность вычисления. Поэтому предложено свести слои весов схожести к девяти меткам, как это изображено в левой части Рис. 1, таким образом резко сократить вычислительные затраты. При этом эффективность фильтрации не теряется в силу того, что области весов схожести одной метки (например 4 на Рис. 1) никак не связаны друг с другом. Осталось решить небольшую проблему с перенумеровкой карты меток с 9 в  $S$ , как это изображено на Рис. 2. В этом случае используется стандартный алгоритм из теории графов: "поиск односвязанных областей в ширину".



**Рис. 2.** Схема преобразования сегментов в унифицированной системе с девятью значениями меток к заданному числу суперпикселей.

Разработанный метод хорошо зарекомендовал себя для большого числа суперпикселей 1000 и более. Для малого числа суперпикселей мы предлагаем пирамидальную систему фильтрации. Когда сначала изображение сегментируется с числом сегментов в  $m$  больше заданного, а затем эта карта сегментов используется для дополнительной фильтрации (усреднение весов схожести внутри сегмента). В экспериментальной части показана разница суперпиксель-сегментации с различными значениями коэффициента  $m$ .

### 3. ЭКСПЕРЕМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Экспериментальная часть была задумана таким образом, чтобы показать основные достоинства предлагаемого алгоритма. Экспериментальная секция нашей работы поделена на две части где:

- мы анализируем численные результаты сегментации на основе критериев точность-отзыв.
- часть результатов экспериментов представлена в иллюстрациях статьи, чтобы читатель мог сравнить визуально качество суперпиксель-сегментации.

В нашей работе численные оценки базируются на двух стандартных метриках  $PR$  — точность (precision) и  $RC$  — отзыв (recall), которые вычисляются по известным формулам:

$$\begin{aligned} RC(S, G) &= \frac{TP(S, G)}{TP(S, G) + FN(S, G)}, \\ PR(S, G) &= \frac{TP(S, G)}{TP(S, G) + FP(S, G)}, \end{aligned} \quad (9)$$

**Таблица 1.** Результаты сравнения по критериям отношения precision-recall нашего метода с четырьмя наиболее известными алгоритмами суперпиксель- сегментации: N-cut [1], TPS [29], SEEDS [28] и SLIC [27].

| Алгоритмы:           | N-cut [1]    | TPS [29] | SEEDS [28]   | SLIC [27]    | Наш метод    |
|----------------------|--------------|----------|--------------|--------------|--------------|
| $S = 300$ Precision  | 0.362        | 0.331    | 0.273        | <u>0.269</u> | <b>0.265</b> |
| $S = 300$ Recall     | 0.875        | 0.814    | <b>0.937</b> | 0.932        | <u>0.934</u> |
| $S = 600$ Precision  | 0.340        | 0.310    | <b>0.236</b> | 0.238        | <u>0.237</u> |
| $S = 600$ Recall     | 0.962        | 0.843    | 0.977        | <u>0.981</u> | <b>0.982</b> |
| $S = 1000$ Precision | 0.288        | 0.287    | 0.225        | <b>0.213</b> | <u>0.219</u> |
| $S = 1000$ Recall    | <u>0.989</u> | 0.981    | 0.986        | <u>0.989</u> | <b>0.991</b> |

где  $S$  — множество всех граничных пикселей, полученных в результате сегментации,  $G$  — множество пикселей истинных границ (ground truth),  $TP$  — true positives,  $FP$  — false positives,  $FN$  — false negatives.

Большинство алгоритмов суперпиксель-сегментации оцениваются на базе данных университета Berkeley [26]. Для объективной оценки мы использовали алгоритмы суперпиксель-сегментации, исходный код которых или двоичные коды находятся в открытом доступе. Мы выбрали четыре наиболее известных методов суперпиксельной сегментации: SLIC [27], SEEDS [28], TPS [29] и N-cut [1]. В Таблице 1 показаны результаты этих методов в сравнении с предлагаемым алгоритмом.

### 3.1. Эксперименты с численным анализом результатов

В этой части экспериментальной секции результаты, полученные с помощью предложенного метода, сравниваются с результатами четырех наиболее известных алгоритмов суперпиксель-сегментации: N-cut [1], TPS [29], SEEDS [28] и SLIC [27] на основе критериев precision-recall.

Анализируя результат в Таблице 1, мы можем сказать, что предложенный в статье алгоритм значительно превосходит три из четырех алгоритмов и показывает примерно равные результаты с методом SEEDS [28].

### 3.2. Результаты экспериментов для визуального анализа

В этой части экспериментальной секции на Рис. 3, Рис. 4 и Рис. 5 представлены результаты сегментации при различных параметрах алгоритма для визуального анализа. Верхняя строка каждого рисунка представляет сегментацию, полученную после прямой реализации алгоритма, средняя — при пирамидальной реализации фильтра с шагом 4, и нижняя строка — для шага 2. Из рисунков мы можем видеть, что там, где границы сегментов совпадают с границами объектов на изображении, все три реализации метода работают примерно одинаково. Отличие заметно лишь в областях с отсутствием значимых границ на изображении. Для таких областей предпочтительной является пирамидальная фильтрация с шагом 4, так как в этом случае границы сегментов ближе всего к прямой линии, а следовательно, уменьшается число ложных граничных пикселей.

## ЗАКЛЮЧЕНИЕ

В этой работе для суперпиксель-сегментации мы предлагаем использовать подход минимизации энергии, отлично зарекомендовавший себя во многих областях компьютерного зрения. В задаче суперпиксель-сегментации число областей часто превышает 1000, что делает вычислительную задачу оптимизации методом МАР практически невыполнимой. В данной



**Рис. 3.** Визуальный результат сегментации предложенного алгоритма для 300 сегментов. Верхняя строка представляет сегментацию, полученную после прямой реализации алгоритма, средняя — при пирамидальной реализации фильтра с шагом 4, и нижняя строка — для шага 2.



**Рис. 4.** Визуальный результат сегментации предложенного алгоритма для 600 сегментов. Верхняя строка представляет сегментацию, полученную после прямой реализации алгоритма, средняя — при пирамидальной реализации фильтра с шагом 4, и нижняя строка — для шага 2.





**Рис. 5.** Визуальный результат сегментации предложенного алгоритма для 300 сегментов. Верхняя строка представляет сегментацию, полученную после прямой реализации алгоритма, средняя — при пирамидальной реализации фильтра с шагом 4, и нижняя строка — для шага 2.

работе мы предложили решение, которое сводит сегментацию с любым количеством областей к проблеме маркировки всего девяти меток. Кроме того, для усиления робастности оптимизации используется свертка с ядром геодезического расстояния. Это позволяет получать на выходе алгоритма односвязанные суперпиксели, в отличие от многих других методов, требующих дополнительной корректировки. Эффективность предложенного метода сравнивается и измеряется по критерию совпадения границ областей, а также с помощью визуальной иллюстрации. Анализ результата работы предложенного алгоритма показал, что метод показывает уровень state-of-the-art при этом скорость вычисления одна из самых быстрых для данного типа алгоритмов.

#### СПИСОК ЛИТЕРАТУРЫ

1. Shi J., Malik J. Normalized cuts and image segmentation // *IEEE Transactions on pattern analysis and machine intelligence*. 2000. Vol. 22, no. 8. P. 888–905.
2. Ren X., Malik J. Learning a classification model for segmentation // *Computer Vision, IEEE International Conference on / IEEE Computer Society*. Vol. 2. 2003. P. 10–10.
3. Achanta R., Shaji A., Smith K. et al. Slic superpixels compared to state-of-the-art superpixel methods // *IEEE transactions on pattern analysis and machine intelligence*. 2012. Vol. 34, no. 11. P. 2274–2282.
4. Li Z., Chen J. Superpixel segmentation using linear spectral clustering // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. P. 1356–1363.
5. Shen J., Du Y., Wang W., Li X. Lazy random walks for superpixel segmentation // *IEEE Transactions on Image Processing*. 2014. Vol. 23, no. 4. P. 1451–1462.
6. Wang P., Zeng G., Gan R. et al. Structure-sensitive superpixels via geodesic distance // *International journal of computer vision*. 2013. Vol. 103. P. 1–21.
7. Veksler O., Boykov Y., Mehrani P. Superpixels and supervoxels in an energy optimization framework // *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5–11, 2010, Proceedings, Part V 11 / Springer*. 2010. P. 211–224.
8. Moore A. P., Prince S. J., Warrell J. “lattice cut”-constructing superpixels using layer constraints // *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition / IEEE*. 2010. P. 2117–2124.
9. Wang J., Wang X. Vcells: Simple and efficient superpixels using edge-weighted centroidal voronoi tessellations // *IEEE Transactions on pattern analysis and machine intelligence*. 2012. Vol. 34, no. 6. P. 1241–1247.
10. Papon J., Abramov A., Schoeler M., Worgotter F. Voxel cloud connectivity segmentation-supervoxels for point clouds // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013. P. 2027–2034.
11. Giordano D., Murabito F., Palazzo S., Spampinato C. Superpixel-based video object segmentation using perceptual organization and location prior // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015. P. 4814–4822.
12. Arbelaez P., Maire M., Fowlkes C., Malik J. Contour detection and hierarchical image segmentation // *IEEE transactions on pattern analysis and machine intelligence*. 2010. Vol. 33, no. 5. P. 898–916.
13. Yang F., Lu H., Yang M.-H. Robust superpixel tracking // *IEEE Transactions on Image Processing*. 2014. Vol. 23, no. 4. P. 1639–1651.
14. Tian Z., Zheng N., Xue J. et al. Video object segmentation with shape cue based on spatiotemporal superpixel neighbourhood // *IET Computer Vision*. 2014. Vol. 8, no. 1. P. 16–25.
15. Cheng J., Liu J., Xu Y. et al. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening // *IEEE transactions on medical imaging*. 2013. Vol. 32, no. 6. P. 1019–1032.

16. Karnaukhov V., Mozerov M. Restoration of noisy multispectral images with a geodetic distance filter // *Journal of Communications Technology and Electronics*. 2018. Vol. 63, no. 6. P. 612–615.
17. Karnaukhov V., Mozerov M. Restoration of multispectral images by the gradient reconstruction method and estimation of the blur parameters on the basis of the multipurpose matching model // *Journal of Communications Technology and Electronics*. 2016. Vol. 61, no. 12. P. 1426–1431.
18. Tomasi C., Manduchi R. Bilateral filtering for gray and color images // *In CVPR*. 1998. P. 839–846.
19. Mozerov M., van de Weijer J. Accurate stereo matching by two-step energy minimization // *IEEE Transactions on Image Processing*. 2015. Vol. 24, no. 3. P. 1153–1163.
20. Mozerov M. G., Van De Weijer J. One-view occlusion detection for stereo matching with a fully connected crf model // *IEEE Transactions on Image Processing*. 2019. Vol. 28, no. 6. P. 2936–2947.
21. Ershov E., Karnaukhov V., Mozerov M. Probabilistic choice between symmetric disparities in motion stereo matching for a lateral navigation system // *Optical Engineering*. 2016. Vol. 55, no. 2. P. 023101–023101.
22. Mozerov M. Constrained optical flow estimation as a matching problem // *IEEE Transactions on Image Processing*. 2013. Vol. 22, no. 5. P. 2044–2055.
23. Karnaukhov V., Kober V., Mozerov M. Improving the quality and contrast of image details using the geodesic distance filter // *Journal of Communications Technology and Electronics*. 2020. Vol. 65. P. 706–711.
24. Adams A., Baek J., Davis M. Fast high-dimensional filtering using the permutohedral lattice // *Computer Graphics Forum*. 2010. Vol. 29, no. 2. P. 753–762.
25. Mozerov M. G., van de Weijer J. Global color sparseness and a local statistics prior for fast bilateral filtering // *Image Processing, IEEE Transactions on*. 2015. Vol. 24, no. 12. P. 5842–5853.
26. Martin D., Fowlkes C., Tal D., Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics // *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001 / IEEE*. Vol. 2. 2001. P. 416–423.
27. Liu Y.-J., Yu C.-C., Yu M.-J., He Y. Manifold slic: A fast method to compute content-sensitive superpixels // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016. P. 651–659.
28. Van den Bergh M., Boix X., Roig G. et al. Seeds: Superpixels extracted via energy-driven sampling. // *ECCV (7)*. 2012. Vol. 7578. P. 13–26.
29. Tang D., Fu H., Cao X. Topology preserved regular superpixel // *2012 IEEE international conference on multimedia and expo / IEEE*. 2012. P. 765–768.

## **Superpixel-segmentation based on energy minimization and convolution with the geodesic distance kernel**

**V.N. Karnaukhov, V.I. Kober, M.G. Mozerov, L.V. Zimina**

The energy minimization or maximum a posteriori probability (MAP) method is the basis for solving many computer vision problems, including the segmentation problem. However, it is assumed that the number of regions during segmentation is quite small. At the same time, in the problem of superpixel segmentation or otherwise excessive segmentation, the number of such areas exceeds 1000, which makes the computational optimization problem by the MAP method practically impossible. In this paper, we propose a solution that reduces segmentation with any number of areas to the problem of labeling only nine labels. In addition, convolution with the geodesic distance kernel is used to enhance the robustness of optimization. This makes it possible to obtain single-linked superpixels at the output of the algorithm, unlike many other

methods that require additional adjustments. The effectiveness of the proposed method is compared and measured by the precision-recall criteria, as well as by visual illustration.

**KEYWORDS:** Superpixel segmentation, geodesic distance based filter, energy minimization, maximum a posteriori probability.