

ISSN 1561-8323 (print)

ИНФОРМАТИКА
INFORMATICS

УДК 004

Поступило в редакцию 12.04.2017

Received 12.04.2017

Б. А. Залесский*Объединенный институт проблем информатики Национальной академии наук Беларуси,
Минск, Республика Беларусь***ДЕТЕКТОРЫ ЭКСТРЕМАЛЬНЫХ ОСОБЫХ ТОЧЕК НА ИЗОБРАЖЕНИЯХ***(Представлено членом-корреспондентом А. В. Тузиковым)*

Аннотация. Представлены детекторы особых (ключевых, характерных) точек-экстремумов, предназначенных для описания, анализа и сравнения изображений с помощью локальных дескрипторов, которые вычисляются в окрестностях найденных точек. Отличие предлагаемых детекторов от известных состоит в том, что они находят особые точки путем поиска локальных экстремумов функции, задающей ту или иную локальную характеристику исходного изображения в каждой его точке. Большинство известных в настоящее время детекторов решают задачу поиска особых точек иным способом. Каждый такой детектор использует построенную для него функцию-характеристику изображения, значение которой в каждом пикселе сравнивается с наперед заданным числовым порогом. Если значение выбранной функции-характеристики больше заданного порогового значения, точка считается особой, в противном случае – обычной. Пороговое значение, как правило, устанавливается с помощью обучения детектора на широком классе обучающих изображений. Некоторые известные детекторы выделяют особые точки путем поиска локальных экстремумов, однако не на исходном изображении, а на его градиентных преобразованиях. Одним из недостатков использования известных детекторов (помимо использования дорогостоящего в вычислительном смысле процесса обучения) является неравномерное распределение особых точек на изображении. Нередко на изображениях возникают большие области, на которых вообще нет особых точек, что приводит к невозможности обнаружения или распознавания этих областей. Предлагаемый класс детекторов позволяет во многих случаях избежать появления областей без особых точек.

Ключевые слова: изображение, ключевые точки, детекторы особых точек, обнаружение и распознавание объектов

Для цитирования: Залесский, Б. А. Детекторы экстремальных особых точек на изображениях / Б. А. Залесский // Докл. Нац. акад. наук Беларуси. – 2017. – Т. 61, № 5. – С. 37–41.

Boris A. Zalesky*United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus***DETECTORS OF EXTREMAL KEY POINTS ON IMAGES***(Communicated by Corresponding Member Alexander V. Tuzikov)*

Abstract. Extremal key-point detectors are presented to describe, analyze and compare images by local descriptors that are determined in neighborhoods of the detected key-points. The proposed detectors select key-points, providing local extremal values of the function that characterizes local properties of the original image at each pixel. The majority of commonly used detecting algorithms are looking for key-points in another way. They mark a pixel as a key-point if the value of the function-criterion at this pixel exceeds a predetermined threshold value. The remaining known algorithms find key-points that are the local extremal values of the functions defined on gradient transforms of the image. One of the drawbacks of the known detectors (in addition to the use of learning procedures expensive in the computational sense) is a non-uniform distribution of key-points on the image. Often large image areas may be left with no key-points, making their detection or recognition impossible. The proposed extremal detectors allow one in many cases to avoid the appearance of image areas not filled with key-points.

Keywords: images, key-points, key-point detectors, object detection and recognition

For citation: Zalesky B. A. Detectors of extremal key points on images. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2017, vol. 61, no. 5, pp. 37–41 (in Russian).

Введение. В настоящее время многие задачи анализа, сравнения и распознавания изображений решаются с помощью сопоставления их локальных характеристик последовательно применяемыми алгоритмами: сначала выделяющими так называемые особые точки (key-points) – детекторами; затем задающими числовые характеристики локальных окрестностей выделенных ключевых точек – дескрипторами; и наконец устанавливающими соответствия между наборами характеристик разных изображений (matchers).

Одними из первых алгоритмов указанного типа, успешно применяемых для решения задач компьютерного зрения, были SIFT и HOG [1; 2], в первом из которых используется детектор ключевых точек, основанный на анализе гессиана разности сглаженных в разной степени копий исходного изображения, а во втором ключевыми считаются все точки регулярной решетки. Через год появился SURF [3], детектор которого использует вейвлетное приближение гессиана для выделения ключевых точек. Затем были опубликованы алгоритмы FAST [4], STAR [5], BRISK [6], KAZE [7], AKAZE [8] и другие вместе с их многочисленными модификациями и улучшениями. Результаты сравнительного анализа детекторов приведены в [9].

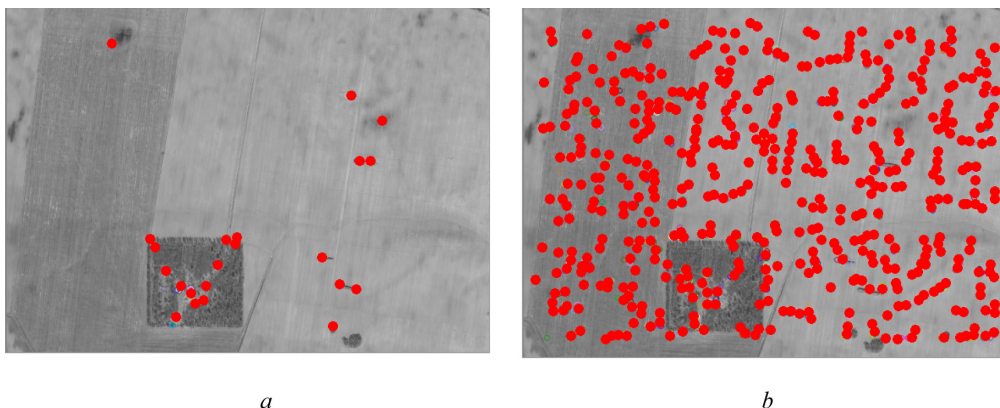
Известные детекторы особых точек обладают рядом общих недостатков: они чувствительны к изменению яркости и контраста изображений (таблица), распределение найденных особых точек на изображении часто бывает неравномерным – нередко возникают области на изображениях, не содержащие ни одной особой точки (рисунок). Такие области не могут быть обнаружены или распознаны с помощью упомянутых алгоритмов.

Предлагаются детекторы особых точек, основанные на поиске локальных экстремумов функций-характеристик непреобразованных изображений. Приведены результаты сравнительного анализа свойств известных и предложенных детекторов.

Экстремальные детекторы особых точек. Анализ известных детекторов показал, что они выделяют недостаточное количество особых точек на нерезких изображениях или изображениях с малым градиентом. На рисунке приведена часть спутникового изображения пахотных земель, на котором известные детекторы оставили большие области без особых точек. Пустые области невозможно найти, распознать или сравнить.

Предлагаемые ниже алгоритмы поиска особых точек-экстремумов находят пиксели, в которых та или иная функция, характеризующая локальные свойства изображения, достигает своего экстремума. Их отличие от известных алгоритмов состоит в использовании интегральных локальных характеристик от самого изображения, а не от его градиентных преобразований. Это позволяет находить информативные особые точки на изображении в окрестностях с малым или вообще нулевым градиентом, при этом нет необходимости использовать пороговые значения.

Рассмотрим моментные экстремальные детекторы особых точек. Несмотря на то что в настоящее время моменты широко используются при обработке изображений, до сих пор они не использовались для экстремального выделения особых точек. Обозначим через $S = \{(x, y)\}$, $x = 0, \dots, m - 1$, $y = 0, \dots, n - 1$, множество пикселей $\mathbf{p} = (x, y)$ полутонового $\mathbf{I} = \{I(\mathbf{p})\}$, $\mathbf{p} \in S$, или



Обособые точки, найденные детектором AKAZE (a) и одним из предложенных алгоритмов (b)
Key points found by detector AKAZE (a) and by one of proposed detectors (b)

цветного $\mathbf{I} = \{I_R(\mathbf{p}), I_G(\mathbf{p}), I_B(\mathbf{p})\}$, $\mathbf{p} \in S$, изображения, и пусть $O_k(\mathbf{p})$ – окрестность пиксела \mathbf{p} , которую для простоты будем считать либо кругом радиуса k , либо квадратом со стороной $2k$ с центром в пикселе \mathbf{p} . Напомним, что локальные абсолютные центрированные моменты полутонового изображения задаются соотношением $\beta_r(\mathbf{p}) = \beta_r(\mathbf{I}, \mathbf{p}) = \sum_{\mathbf{q} \in O_k(\mathbf{p})} |I(\mathbf{q}) - m(\mathbf{p})|^r$, где $m(\mathbf{p}) = k^{-1} \sum_{\mathbf{q} \in O_k(\mathbf{p})} I(\mathbf{q})$ – среднее значение. Абсолютный центрированный момент порядка r цветного изображения равен $\beta_r(\mathbf{p}) = \beta_r(\mathbf{I}, \mathbf{p}) = \beta_r(\mathbf{I}_R, \mathbf{p}) + \beta_r(\mathbf{I}_G, \mathbf{p}) + \beta_r(\mathbf{I}_B, \mathbf{p})$. Изображения являются функциями дискретного аргумента, поэтому определение локального экстремума, которое можно использовать для поиска экстремальных особых точек (ЭОТ) отличается от общепринятого. По существу, для поиска ЭОТ используются локальные экстремумы в заранее выбранной окрестности $O_\ell(\mathbf{p})$ каждого пиксела \mathbf{p} (не обязательно совпадающей с $O_k(\mathbf{p})$).

Теоретически, для фиксированного $r > 0$ экстремальные моментные особые точки \mathbf{p}_r° находятся как пиксели, обеспечивающие невырожденные локальные максимумы момента $\beta_r(\mathbf{p})$ в окрестностях $O_\ell(\mathbf{p})$ фиксированной формы и размера, в случае, если эти локальные экстремумы лежат на заданном удалении от границы окрестности. Иными словами для заданных: степени момента r , радиусов k, ℓ окрестностей O_k, O_ℓ и максимального удаления $\tau (\tau \leq \ell)$ локального экстремума от центра окрестности – ЭОТ образуют множество

$$\chi(\mathbf{I}) = \chi_{r,k,\ell,\tau}(\mathbf{I}) = \left\{ \mathbf{p}_r^\circ \in S \mid \exists \mathbf{q} \in S : \mathbf{p}_r^\circ \in O_\ell(\mathbf{q}), \left| \mathbf{p}_r^\circ - \mathbf{q} \right| \leq \tau, \forall \mathbf{u} \in O_\ell(\mathbf{q}), \beta_r(\mathbf{p}_r^\circ) \geq \beta_r(\mathbf{u}), \exists \mathbf{v} \in O_\ell(\mathbf{q}) : \beta_r(\mathbf{p}_r^\circ) > \beta_r(\mathbf{v}) \right\}. \quad (1)$$

Вторая часть условия $\exists \mathbf{v} \in O_\ell(\mathbf{q}) : \beta_r(\mathbf{p}_r^\circ) > \beta_r(\mathbf{v})$ используется для того, чтобы избежать выбора неинформативных особых точек из однородных областей изображений. Словами условие (1) может быть сформулировано следующим образом: пиксел \mathbf{p}_r° является экстремальной особой точкой, если существует окрестность $O_\ell(\mathbf{q})$ с центром в \mathbf{q} , содержащая \mathbf{p}_r° , в котором значение $\beta_r(\mathbf{p}_r^\circ)$ является локальным максимумом, при этом \mathbf{p}_r° удален от центра окрестности \mathbf{q} на расстояние, не превосходящее $\tau (\tau \leq \ell)$, и β_r не является константой в $O_\ell(\mathbf{q})$.

Практически, в некоторых случаях выделенное множество ЭОТ $\chi_r(\mathbf{I})$ приходится прореживать – оставлять из нескольких расположенных близко равноценных особых точек одну, например, случайным образом или по какому-либо дополнительному критерию. Процедура прореживания особых точек используется в большинстве известных детекторов.

Выделение ЭОТ может осуществляться с помощью моментов β_r любых порядков, однако, использование первых моментов четных степеней β_2, β_4 и модуля неабсолютного момента 3-й степени $|\alpha_3|$ наиболее предпочтительно с вычислительной точки зрения потому, что данные функции могут быть вычислены быстро с помощью интегрального изображения [10], а проведенный сравнительный анализ ЭОТ, найденных с помощью β_r , показал, что использование $r = 2, 4$ и $|\alpha_3|$ дает примерно одинаковые результаты, причем обычная дисперсия β_2 в большинстве случаев обеспечивает наилучшие наборы точек.

Результаты вычислительных экспериментов. Для исследования свойств предложенных экстремальных детекторов были использованы аэрофотоснимки и космические изображения, так как задача повышения качества дешифрирования изображений данного типа с помощью упомянутого вначале статьи подхода на основе особых точек особенно актуальна. Был проведен поиск соответствующих областей на кадрах видеопоследовательностей, полученных бортовыми камерами самолета и квадрокоптера, а также на космических снимках. Для тестирования были выбраны детекторы SIFT, SURF, ORB, AKAZE, реализованные в библиотеке OpenCV 3.1.0, и предложенный нами экстремальный детектор. Для большей объективности проводимого тестирования после выделения особых точек указанными детекторами во всех экспериментах был использован один и тот же дескриптор LATCH, который считается в настоящее время одним из лучших. Затем поиск соответствий областей изображений осуществлялся известным knn-алгоритмом.

Сравнение между собой экстремальных детекторов, основанных на использовании моментов $\beta_2, |\alpha_3|, \beta_4$, показало, что в среднем β_2 -детектор нашел на кадрах тестовых видеопоследова-

тельностью с точностью до статистической погрешности одинаковое количество ЭОТ (всего на 2,3 % больше, чем $|\alpha_3|$ -детектор, и на 1,5 % больше, чем β_4 -детектор). Количество найденных с помощью kpp-алгоритма соответствий (так называемых *good matches*) между наборами моментных особых точек сравниваемых изображений тоже оказалось практически одинаковым. При этом следует заметить, что вычисление β_2 -детектора с помощью интегральных изображений требует примерно в три раза меньше операций, чем вычисление β_4 -детектора.

Обычно при сравнении детекторов особых точек сравниваются количество найденных точек и число установленных соответствий (*good matches*) между наборами точек сравниваемых изображений. В дополнении к этим характеристикам приведем результаты проверки корректности автоматического установления соответствия между областями сравниваемых пар изображений с помощью алгоритма RANSAC на основе ЭОТ и известных алгоритмов. Результаты получены на видеопоследовательностях, снятых с борта самолета и квадрокоптера. Напомним схему экспериментов. Сначала находятся особые точки известными и разработанными алгоритмами, затем вычисляются характеристики всех найденных точек алгоритмом LATCH. После этого устанавливаются соответствия между наборами точек двух сравниваемых изображений kpp-алгоритмом и находятся проективные преобразования, совмещающие каждую пару сравниваемых изображений с алгоритмом RANSAC.

Для краткости результаты экспериментов приведены в следующей таблице для одной видеопоследовательности, снятой с борта квадрокоптера.

Результаты экспериментов

Results of Experiments

Характеристика Characteristic	ORB	SIFT	SURF	AKAZE	β_2 -детектор β_2 -detector
<i>Видео 1920 × 1080</i>					
Особые точки	3373	695	2751	163	26468
Соотв. точек	446	156	564	38	807
Соотв. кадров, %	100	100	100	100	100
<i>Видео 1920 × 1080 с гауссовым шумом 10 %</i>					
Особые точки	7434	812	3801	155	28658
Соотв. точек	640	88	428	20	479
Соотв. кадров, %	100	100	100	50	100
<i>Видео 1920 × 1080 с гауссовым сглаживанием 10 %</i>					
Особые точки	2	46	220	22	14590
Соотв. точек	0	22	77	11	796
Соотв. кадров, %	0	80	100	0	100

На видеопоследовательности, снятой с борта самолета и сглаженной гауссовым ядром, только предложенный экстремальный детектор β_2 позволил установить 100 % правильных соответствий между областями кадров. Алгоритмы ORB, AKAZE, SIFT не нашли ни одного правильного соответствия, причем первые два алгоритма не нашли ни одной особой точки.

Преимущество β_r было особенно заметно при совмещении областей слабоконтрастных спектральных космических снимков, полученных с Landsat 8. В большинстве случаев известные детекторы находили всего несколько особых точек, либо вообще не находили ни одной, в то время как β_r позволяли успешно решить задачу совмещения областей (рисунок).

Заключение. Предложенные детекторы находили на каждом из тестовых изображений большее число особых точек, чем известные. Среднее число соответствий особых точек на парах изображений, содержащих общую часть сцены, для них также оказалось большим, чем для известных детекторов. Преимущество предложенных детекторов оказалось более заметным при работе со слабоконтрастными изображениями, имеющими слабовыраженные градиентные особенности. Здесь известные детекторы нередко не находили необходимого минимального количества особых точек (или вообще не находили ни одной), которые позволили бы корректно сравнить

или распознать изображения, а также их области, в то время как предложенные – обнаруживали достаточно особенностей, позволяющих решить эту задачу (таблица). К тому же, предложенные детекторы, в отличие от большинства известных, могут быть быстро вычислены с помощью интегральных изображений. Таким образом, проведенный сравнительный анализ свойств предложенных детекторов особых точек, основанных на вычислении локальных экстремумов первых абсолютных центрированных моментов изображений, с известными детекторами, показал их применимость для решения задач обработки и распознавания изображений.

Список использованных источников

1. Lowe, D. Object recognition from local scale invariant features / D. Lowe // Proc. Int. Conf. on Computer Vision ICCV. – Corfu, 1999. – P. 1150–1157. doi.org/10.1109/iccv.1999.790410
2. Dalah, N. Histograms of Oriented Gradients for Human Detection / N. Dalah, B. Triggs // IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition CVPR'05. – San Diego, 2005. – Vol. 1. – P. 886–893. doi.org/10.1109/cvpr.2005.177
3. Bay, H. Surf: Speeded up robust features / H. Bay, T. Tuytelaars, L. Van Gool // Proc. 9th Europ. Conf. on Computer Vision ECCV. – Graz, 2006. – P. 404–417. doi.org/10.1007/11744023_32
4. Rosten, E. Faster and better: a machine learning approach to corner detection / E. Rosten, R. Porter, T. Drummond // IEEE TPAMI. – 2010. – Vol. 32, N 1. – P. 105–119. doi.org/10.1109/tpami.2008.275
5. Agrawal, M. CenSurE: Center surround extremas for realtime feature detection and matching / M. Agrawal, K. Konolige, M. R. Blas // Lecture Notes in Computer Science. – 2008. – Vol. 5305. – P. 102–115. doi.org/10.1007/978-3-540-88693-8_8
6. Leutenegger, S. BRISK: Binary Robust invariant scalable keypoints / S. Leutenegger, M. Chli, R. Y. Siegwart // Proc. IEEE Int. Conf. on Computer Vision ICCV. – Barcelona, 2011. – P. 2548–2555. doi.org/10.1109/iccv.2011.6126542
7. Alcantarilla, P. KAZE Features / P. Alcantarilla, A. Bartoli, J. Davison // Proc. Eur. Conf. on Computer Vision ECCV. – Firenze, 2012. – P. 214–227. doi.org/10.1007/978-3-642-33783-3_16
8. Alcantarilla, P. Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces / P. Alcantarilla, J. Nuevo, A. Bartoli // Proc. British Machine Vision Conference BMVC. – Bristol, 2013. doi.org/10.5244/C.27.13
9. Comparative Assessment of Point Feature Detectors and Descriptors in the Context of Robot Navigation / A. Schmidt [et al.] // J. of Automation, Mobile Robotics & Intelligent Systems. – 2013. – Vol. 7, N 1. – P. 11–20.
10. Viola, P. Robust real-time face detection / P. Viola, M. J. Jones // Int. J. of Computer Vision. – 2004. – Vol. 57, N 2. – P. 137–154. doi.org/10.1023/b:visi.0000013087.49260.fb

References

1. Lowe D. Object recognition from local scale invariant features. *Proceedings of the 7th IEEE International Conference on Computer Vision ICCV*, Corfu, 1999, pp. 1150–1157. doi.org/10.1109/iccv.1999.790410
2. Dalah N., Triggs B. Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'05*, San Diego, 2005, vol. 1, pp. 886–893. doi.org/10.1109/cvpr.2005.177
3. Bay H., Tuytelaars T., Van Gool L. Surf: Speeded up robust features. *Proceedings of the 9th European Conference on Computer Vision ECCV*, Graz, 2006, pp. 404–417. doi.org/10.1007/11744023_32
4. Rosten E., Porter R., Drummond T. Faster and better: a machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, vol. 32, no. 1, pp. 105–119. doi.org/10.1109/tpami.2008.275
5. Agrawal M., Konolige K., Blas M. R. CenSurE: Center surround extremas for realtime feature detection and matching. *Lecture Notes in Computer Science*, 2008, vol. 5305, pp. 102–115. doi.org/10.1007/978-3-540-88693-8_8
6. Leutenegger S., Chli M., Siegwart R. Y. BRISK: Binary Robust invariant scalable keypoints. *Proceedings of the 13th IEEE International Conference on Computer Vision ICCV*, Barcelona, 2011, pp. 2548–2555. doi.org/10.1109/iccv.2011.6126542
7. Alcantarilla P., Bartoli A., Davison J. KAZE Features. *Proceedings of the 12th European Conference on Computer Vision ECCV*, Firenze, 2012, pp. 214–227. doi.org/10.1007/978-3-642-33783-3_16
8. Alcantarilla P., Nuevo J., Bartoli A. Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces. *Proceedings of the 24th British Machine Vision Conference BMVC*. Bristol, 2013. doi.org/10.5244/C.27.13
9. Schmidt A., Kraft M., Fularz M., Domagała Z. Comparative Assessment of Point Feature Detectors and Descriptors in the Context of Robot Navigation. *Journal of Automation, Mobile Robotics & Intelligent Systems*, 2013, vol. 7, no. 1, pp. 11–20.
10. Viola P., Jones M. J. Robust real-time face detection. *International Journal of Computer Vision*, 2004, vol. 57, no. 2, pp. 137–154. doi.org/10.1023/b:visi.0000013087.49260.fb

Информация об авторе

Залеский Борис Андреевич – д-р физ.-мат. наук, ведущий лабораторией. Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: zalesky@newman.bas-net.by.

Information about the author

Zalesky Boris Andreevich – D. Sc. (Physics and Mathematics), Head of the Laboratory. United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: zalesky@newman.bas-net.by.