

Обучаемый детектор особых точек изображения

А.В. Веричев^а

^а Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Московское шоссе, 34, Самара, Россия

Аннотация

В работе представлен алгоритм построения обучаемого детектора особых точек изображения. В качестве локальных признаков используются моменты и моментные инварианты изображения. Для сокращения числа требуемых арифметических операций применяется метод рекурсивного вычисления степенных моментов с помощью целочисленных факториальных полиномов. Использование методов машинного обучения позволяет обеспечить настройку детектора к требуемым откликам. Результаты экспериментальных исследований демонстрируют возможность настройки детектора под задачу, повторяемость результатов, инвариантность к преобразованиям подобия изображений и возможность использования детектора на различных масштабах без построения пирамиды разрешений.

Ключевые слова: особые точки изображений; детекторы особых точек; обобщённые моменты изображений; моментные инварианты; машинное обучение

1. Введение

Особыми называются такие точки изображения, локальные окрестности которых обладают некоторыми отличительными особенностями в сравнении с окрестностями остальных точек изображения. Особые точки применяются при решении ряда задач, таких как создание панорамных снимков, стереотождествление, отслеживание движущихся объектов, распознавание и поиск объектов [1]. Решение упомянутых задач возможно благодаря тому, что особые точки обладают следующими свойствами [2]:

- *Повторяемость (repeatability)* - особая точка находится в одном и том же месте сцены или объекта изображения, несмотря на изменение угла обзора или освещения наблюдаемой сцены.
- *Информативность (informativeness)* – окрестности особых точек значительно отличаются друг от друга.
- *Локальность (locality)* – особая точка занимает небольшую область изображения.
- *Количество (quantity)* – число особых точек достаточно велико для того, чтобы их хватило для обнаружения даже небольших объектов.
- *Точность (accuracy)* – обнаруженные особые точки точно локализируются.
- *Эффективность (efficiency)* – особые точки изображения обнаруживаются за приемлемое время в критичных по времени приложениях.

Совокупность алгоритмов и процедур нахождения особых точек на изображении называется детектором особых точек. В литературе описано большое количество детекторов, подавляющее большинство которых основано на некотором тесте - эвристике, неявно определяющей понятие особая точка. Часто используемые эвристические критерии можно условно разделить на следующие группы [3]:

- Критерии, основанные на интенсивности изображения: значения яркости элементов изображения используются для построения некоторой функции, определяющей меру «особенности» точки. Примерами детекторов, основанных на таком подходе, являются Förstner [4], Harris [5], Shi-Tomasi [6].
- Критерии, основанные на шаблоне: шаблон определяет локальную окрестность точки, значения яркостей элементов которой участвуют в проверке выполнения определённого условия. На критериях по шаблону основаны детекторы SUSAN [7], FAST [8], AGAST [9].
- Критерии, основанные на контурах объектов: особая точка определяется как изгиб или пересечение границ двух областей. Примеры подобных детекторов: DoG-curve [10], ANDD [11].

Формулирование эвристических правил требует однозначного понимания того, что именно является особой точкой в контексте решаемой задачи, что в свою очередь подразумевает знание предметной области. Сформулированный критерий настраивает детектор на конкретную задачу, что повышает качество его работы, однако вместе с тем снижает гибкость и затрудняет возможность применения настроенного детектора в других задачах.

В данной работе предпринята попытка отказаться от явного определения понятия особая точка с помощью эвристического критерия и сконцентрироваться на требуемых свойствах особых точек. С этой целью мы прибегаем к методам машинного обучения. В качестве признаков используются некоторые локальные характеристики изображения, в том числе моменты и моментные инварианты. Детектор обучается решению конкретной задачи в заданных условиях, что фактически эквивалентно заданию формального критерия.

Предлагаемый метод детально представлен в разделе 2, где описывается признаковое пространство, методы формирования обучающей выборки и используемый метод обучения. Критерии оценки качества обучения детектора, качества детектирования особых точек настроенным детектором, а также результаты экспериментальных исследований приведены в разделе 3. Выводы и итоги подведены в заключении.

2. Предлагаемый метод

Предлагаемый алгоритм построения детектора особых точек изображения основывается на предложенной в [12] идее построения алгоритма детектирования по прецедентам и заключается в конструировании искомого преобразования - детектора - на специальном образом сформированной обучающей выборке, содержащей примеры требуемых откликов.

2.1. Построение признакового пространства

Первым этапом построения искомого преобразования является формирование признакового пространства. Признаковое пространство представляет собой векторное пространство \mathbb{R}^{15} . Каждому элементу изображения соответствует вектор из \mathbb{R}^{15} . Извлечение признаков является локальной операцией, выполняемой в окне размером 9×9 , центрированном на обрабатываемом элементе изображения. В этой локальной области вычисляется ряд характеристик, описанных ниже, которые образуют элементы вектора признакового пространства. Таким образом, извлечение признаков есть отображение $P^{9 \times 9} \rightarrow \mathbb{R}^{15}$, где $P = \{n: 0 \leq n < 256\}$ - множество значений яркости полутонового изображения.

Первыми двумя признаками элемента изображения $I[x, y]$ являются выборочное среднее квадратичное отклонение функции яркости, нормированной в окне 9×9 , а также нормированное значение с.к.о.:

$$\begin{aligned}\phi_1 &= \sqrt{\frac{1}{80} \sum_{i=-4}^4 \sum_{j=-4}^4 \frac{1}{n^2} (I[x+i, y+j] - \bar{I})^2}, \\ \phi_2 &= \frac{\phi_1}{n},\end{aligned}\tag{1}$$

где норма n и среднее \bar{I} определены:

$$\begin{aligned}\bar{I} &= \frac{1}{81} \sum_{i=-4}^4 \sum_{j=-4}^4 I[x+i, y+j], \\ n &= \sqrt{\sum_{i=-4}^4 \sum_{j=-4}^4 (I[x+i, y+j])^2}.\end{aligned}$$

Использование этой пары признаков призвано облегчить различение монотонных и текстурированных областей.

Следующие четыре признака являются центральными моментами локальной области изображения: $\phi_{t+3} = \mu_{tt}$, $0 \leq t \leq 3$. Центральные моменты вычисляются следующим образом [13]:

$$\mu_{ij} = \sum_{k=-4}^4 \sum_{l=-4}^4 k^i \cdot l^j \cdot \frac{1}{n} I[x+k, y+l].\tag{2}$$

Признаками, отвечающими за инвариантность к преобразованию поворота, являются инвариантные моменты X_y [13, 14]:

$$\begin{aligned}\phi_7 &= \mu_{20} + \mu_{02}, \\ \phi_8 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\ \phi_9 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\ \phi_{10} &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \\ \phi_{11} &= \frac{(\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]}{(\mu_{21} + \mu_{03})^2}, \\ \phi_{12} &= (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}), \\ \phi_{13} &= \frac{(3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]}{(\mu_{21} + \mu_{03})^2}, \\ \phi_{14} &= \mu_{11}[(\mu_{30} + \mu_{12})^2 - (\mu_{03} + \mu_{21})^2] - (\mu_{20} - \mu_{02})(\mu_{30} + \mu_{12})(\mu_{03} + \mu_{21}).\end{aligned}\tag{3}$$

Нахождение значений центральных моментов изображения является вычислительно сложной задачей. С целью снизить количество выполняемых операций используется метод рекурсивного вычисления степенных моментов с помощью целочисленных факториальных полиномов [15].

Дополнительный признак, характеризующий удалённость центра масс фрагмента от центра локальной области, определяется формулой:

$$\phi_{15} = \sqrt{(x_c - x)^2 + (y_c - y)^2}, \quad (4)$$

где $x_c = \mu_{10}/\mu_{00}$ и $y_c = \mu_{01}/\mu_{00}$.

Упорядоченный набор признаков ϕ_i , $1 \leq i \leq 15$, определённых формулами (1) - (4), образует вектор признаков из \mathbb{R}^{15} , соответствующий элементу изображения с координатами (x, y) .

2.2. Построение детектора

2.2.1. Формирование обучающей выборки

Для настройки детектора требуется обучающая выборка, образуемая размеченными векторами признаков, извлекаемыми из изображения. В зависимости от решаемой задачи, разметку экземпляров обучающей выборки можно осуществлять следующими способами:

- вручную, привлекая специалистов в соответствующей предметной области;
- используя хорошо зарекомендовавшие себя детекторы, такие как Harris для нахождения углов или Canny для нахождения точек, образующих рёбра;
- комбинируя предыдущие методы.

В случае использования первого или третьего способов, неизбежно возникает т. н. шум учителя – неточности разметки данных [16]. Помимо этого, в типичных сценариях количество особых точек мало по сравнению с не особыми точками. Для сглаживания негативных эффектов этих явлений следует считать особыми не только выбранные точки, но и их ближайших соседей.

В случае предъявления повышенных требований устойчивости к определённым искажениям, например, преобразованиям подобия, размеченную обучающую выборку можно дополнить т. н. виртуальными примерами [17]. С этой целью каждое изображение из обучающей выборки подвергается некоторому искажению, параметры которого известны, что позволяет устанавливать соответствия между элементами исходного и искажённого изображений. Далее, из искажённого изображения извлекаются векторы-признаки и им назначаются метки соответствующих элементов исходного изображения. Новые векторы признаков представляют собой виртуальные примеры. Такой подход позволяет дополнить обучающую выборку данными, несущими информацию об эффектах некоторых преобразований на векторы признакового пространства.

2.2.2. Настройка детектора

Наличие размеченной обучающей выборки позволяет приступить к этапу настройки детектора. Так как объём обучающей выборки даже для небольших изображений сравнительно велик, для построения показателя, характеризующего выраженность особенности точки, предлагается использовать оценку плотности вероятностей для соответствующего ей вектора признаков.

Обозначим обучающую выборку $D = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$, где \mathbf{x}_i – вектор признаков, y_i – метка класса, принимающая значение C_1 для особых точек и C_2 для не особых точек. Тогда оценки условных плотностей нового вектора \mathbf{x} определяются соотношением:

$$\hat{p}(\mathbf{x}|C_i) \propto \sum_{j=1}^N [y_j = C_i] K\left(\frac{\|\mathbf{x} - \mathbf{x}_j\|}{h}\right), \quad (5)$$

где K – некоторое ядро, h - его параметр, контролирующий ширину окна. Из теоремы Байеса следует:

$$\hat{p}(C_i|\mathbf{x}) \propto \hat{p}(\mathbf{x}|C_i) \cdot \hat{\pi}_i, \quad (6)$$

где $\hat{\pi}_i$ – оценка априорной вероятности i -го класса:

$$\hat{\pi}_i = \frac{1}{N} \sum_{j=1}^N [y_j = C_i]. \quad (7)$$

Определим функцию особенности точки $l(\mathbf{x})$:

$$l(\mathbf{x}) = \ln(\hat{p}(C_1|\mathbf{x})) - \ln(\hat{p}(C_2|\mathbf{x})). \quad (8)$$



Рис. 1. Примеры изображений одной сцены с разных ракурсов.

Для сглаживания отклика предлагаемого детектора функцию особенности точки $l(\mathbf{x})$ следует предварительно подвергнуть локальной пиковой фильтрации. Пиковый фильтр подавляет не максимальные значения функции $l(\mathbf{x})$ и определяется следующим образом:

$$\tilde{l}(\mathbf{x}) = \begin{cases} l(\mathbf{x}), & l(\mathbf{x}) > l(\mathbf{g}) + \delta \quad \forall \mathbf{g} \in W \setminus \{\mathbf{x}\}, \\ 0, & \text{иначе} \end{cases} \quad (9)$$

где W – множество векторов признакового пространства, соответствующих элементам изображения, содержащимся в окне размером 3×3 , центрованном на элементе изображения, которому соответствует вектор признаков \mathbf{x} ; δ – некоторое пороговое значение, выбираемое экспериментальным путём.

Используя отфильтрованное значение функции особенности точки, приходим к простому решающему правилу:

$$y(\mathbf{x}) = \begin{cases} C_1, & \tilde{l}(\mathbf{x}) > t = \ln\left(\frac{\hat{\pi}_2}{\hat{\pi}_1}\right), \\ C_2, & \text{иначе} \end{cases} \quad (10)$$

3. Экспериментальные исследования

3.1. Условия экспериментов

Для проведения экспериментальных исследований была собрана база изображений, состоящая из 10 снимков 6 разнородных сцен, снятых с различных ракурсов - всего 60 изображений. Примеры различных изображений одной сцены приведены на рис. 1. Съёмка осуществлялась фотоаппаратом Panasonic Lumix DMC-TZ10. База была разделена на обучающую и контрольную выборки в отношении 80/20 независимо для каждой сцены. Выборки размечались с помощью детектора Harris. Для тестирования повторяемости настроенного детектора выборки были дополнены виртуальными примерами, как описано в разделе 2.2.1. Используемые при этом преобразования описаны ниже, в разделе 3.3.

3.2. Оценка качества настройки детектора

Пусть $V = \{(\mathbf{x}_i, y_i)\}_{i=1}^N$ – обучающая или контрольная выборка. Основным показателем качества работы детектора на выборке V является точность детектора, (*accuracy*):

$$A(V) = \frac{1}{N} \sum_{i=1}^N [y(\mathbf{x}_i) = y_i]. \quad (11)$$

Помимо точности детектора, для оценки качества работы детектора используется две дополнительные характеристики: точность отклика детектора (*precision*), P , и полнота отклика детектора (*recall*), R [18]. Точностью отклика называют отношение верных откликов к общему числу положительных откликов детектора, полнотой называют отношение верных откликов к общему числу положительных примеров в обучающей выборке.

Обозначим ложное срабатывание, или ошибку первого рода, FP , пропуск события, или ошибку второго рода, FN , и верное срабатывание детектора TP . Эти величины определяются следующим образом:

$$\begin{aligned} FP(V) &= \sum_{i=1}^N [y(\mathbf{x}_i) = C_1] \cdot [y_i = C_2], \\ FN(V) &= \sum_{i=1}^N [y(\mathbf{x}_i) = C_2] \cdot [y_i = C_1], \\ TP(V) &= \sum_{i=1}^N [y(\mathbf{x}_i) = y_i]. \end{aligned} \quad (12)$$

Таблица 1. Характеристики качества настройки детектора

Заголовок колонки	$A(D)$	$P(D)$	$R(D)$
Обучающая выборка, D	0.997	0.905	0.960
Контрольная выборка, C	0.9766	0.730	0.580

Тогда точность и полнота отклика вычисляются с помощью формул:

$$P(V) = \frac{TP(V)}{TP(V)+FP(V)}, \quad (13)$$

$$R(V) = \frac{TP(V)}{TP(V)+FN(V)}.$$

Качество работы настроенного на обучающей выборке D детектора было оценено на обучающей выборке D и контрольной выборке C , результаты представлены в таблице 1. Учитывая большой объём и характер формирования обучающей и контрольной выборок, данные свидетельствуют о хорошем качестве обучения.

3.3. Оценка повторяемости детектора

Как описано во введении, важнейшей характеристикой особых точек является повторяемость. К тому же это одна из немногих характеристик, допускающих объективную количественную оценку. Поэтому для оценки качества работы детектора используется повторяемость отыскиваемых особых точек.

Методика тестирования повторяемости:

- Особые точки детектируются на некотором исходном изображении и сохраняются в список P_o .
- Исходное изображение подвергается одному из тестовых преобразований подобия (см. ниже).
- Особые точки детектируются на преобразованном изображении и сохраняются в список P_t .
- Так как параметры преобразования известны, координаты найденных на исходном изображении точек P_o отображаются на преобразованное изображение и сохраняются в список отображённых точек P_m .
- Наборы P_m и P_t сравниваются. Две точки $a \in P_m$ и $b \in P_t$ считаются совпавшими, если $a \in V_\varepsilon(b)$, $\varepsilon = 2.0$.
- В результате сравнения получаем три набора точек: P_{TP} – точки, найденные в обоих списках, P_{FP} – новые точки, появившиеся на преобразованном изображении, P_{FN} – потерянные точки, не воспроизведённые на преобразованном изображении. Количества точек в списках равны, соответственно, величинам TP , FP и FN настроенного детектора. С помощью этих величин вычисляются значения характеристик точности и полноты отклика детектора.

В оценке повторяемости особых точек используются следующие преобразования изображений:

- поворот на угол α , изменяющийся в диапазоне $-45^\circ \leq \alpha \leq 45^\circ$ с шагом 3° ;
- сдвиг с субпиксельной точностью, приводящий к интерполяции недостающих значений яркости, с параметром t , изменяющимся в диапазоне $0.25 \leq t \leq 0.75$ с шагом 0.05 (сдвиг производится по диагонали изображения, т.е. по строкам и столбцам одновременно);
- масштабирование с коэффициентом s , изменяющимся в диапазоне $0.5 \leq s < 1.5$ с шагом 0.1 .

Результаты экспериментального исследования повторяемости особых точек, находимых настроенным на обучающей выборке D детектором, представлены на рис. 2. Детектор демонстрирует умеренную повторяемость при поворотах на угол α , $-9^\circ < \alpha < 9^\circ$, обнадёживающую повторяемость при масштабировании с коэффициентом s , $0.8 \leq s \leq 1.2$, а также высокую повторяемость при сдвигах с параметром t на всём тестовом диапазоне.

4. Заключение

В данной работе осуществлена попытка построения настраиваемого детектора особых точек изображения. В отличие от стандартного подхода к решению этой задачи, особые точки не определяются путём задания эвристического критерия. Понятие особая точка выводится индуктивно с помощью методов машинного обучения, что позволяет настраивать детектор особых точек на решение конкретной задачи. Результаты экспериментальных исследований демонстрируют возможность настройки детектора к требуемым откликам, приемлемую повторяемость при преобразованиях поворота и масштабирования, высокую повторяемость при преобразованиях субпиксельного сдвига, а также возможность использования настроенного детектора на изображениях различного масштаба без построения пирамиды разрешений.

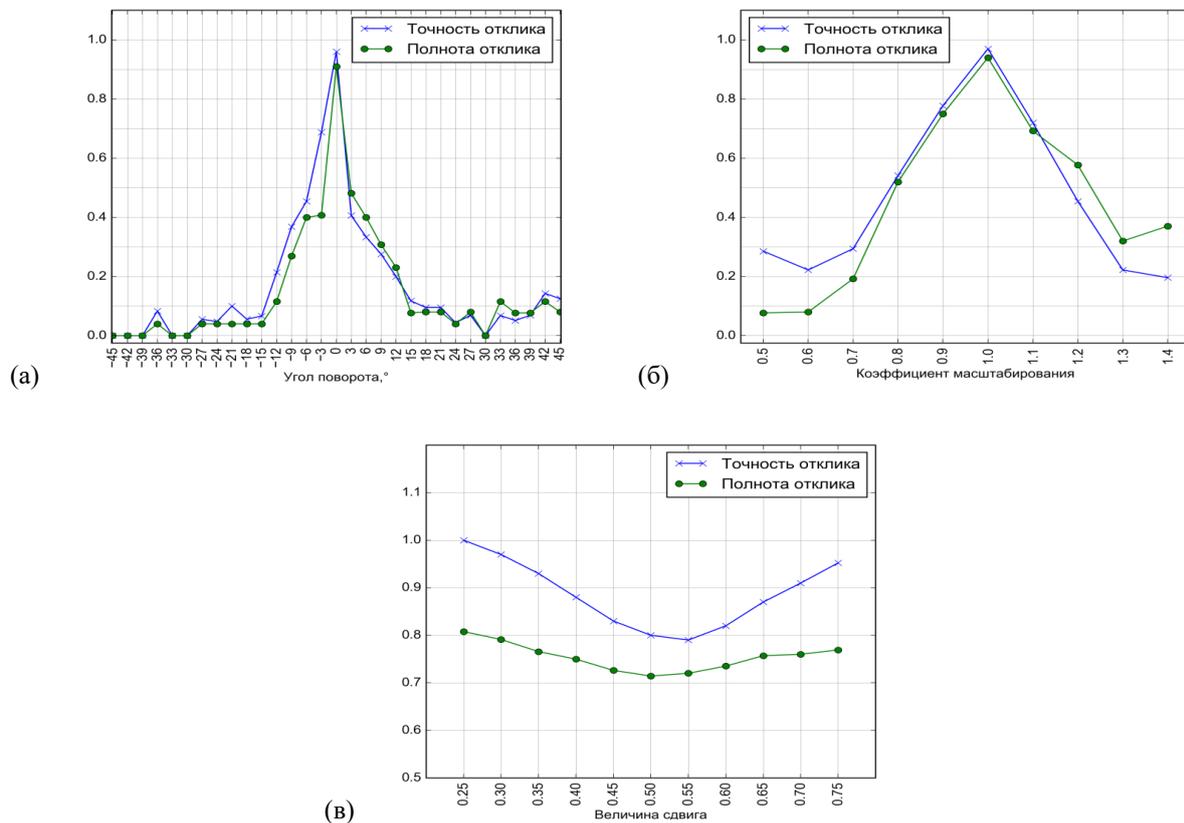


Рис. 2. Повторяемость детектора при преобразованиях: (а) поворота, (б) масштабирования, (в) сдвига.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-29-03190-офи.

Литература

- [1] Szeliski, R. Computer Vision: Algorithms and Applications / R. Szeliski – London: Springer, 2011. – 812 p.
- [2] Tuytelaars, T. Local invariant feature detectors: a survey / T. Tuytelaars, R. Mikolajczyk // Foundations and trends® in computer graphics and vision - 2008. - Vol. 3(3). - P. 177-280. DOI: 10.1561/0600000017.
- [3] Li, Y. A survey of recent advances in visual feature detection / Y. Li, S. Wang, Q. Tian, X. Ding // Neurocomputing. - 2015. - Vol. 149. - P. 736-751. DOI: 10.1016/j.neucom.2014.08.003.
- [4] Förstner, W. A fast operator for detection and precise location of distinct points, corners and centres of circular features / W. Förstner, E. Gülch // Proc. ISPRS intercommission conference on fast processing of photogrammetric data - 1998. - P. 281-305.
- [5] Harris, C. A combined corner and edge detector / C. Harris, M. Stephens // In Alvey vision conference - 1988. - Vol. 15(50). - P. 147-151.
- [6] Shi, J. Good features to track / J. Shi, C. Tomasi // Proc. Intl Conf. on Comp. Vis. and Pat. Recog (CVPR) - 1994. - P. 593-600.
- [7] Smith, S.M. SUSAN—A new approach to low level image processing / S.M. Smith, J.M. Brady // International journal of computer vision - 1997. Vol. 23(1). - P. 45-78. DOI: 10.1023/A:1007963824710.
- [8] Rosten, E. Machine learning for high-speed corner detection / E. Rosten, T. Drummond // In European conference on computer vision - 2006. - P. 430-443. DOI: 10.1007/11744023_34.
- [9] Mair, E. Adaptive and generic corner detection based on the accelerated segment test / E. Mair, G.D. Hager, D. Burschka, M. Suppa, G. Hirzinger // In European conference on Computer vision - 2010. - P. 183-196. DOI: 10.1007/978-3-642-15552-9_14.
- [10] Zhang, X. Corner detection based on gradient correlation matrices of planar curves / X. Zhang H. Wang, A.W.B. Smith, X. Ling, B.C. Lovell, D. Yang // Pattern recognition - 2010. - Vol. 43(4). - P. 1207-1223. DOI: 10.1016/j.patcog.2009.10.017.
- [11] Shui, P.L. Corner detection and classification using anisotropic directional derivative representations / P.L. Shui, W.C. Zhang // IEEE Transactions on Image Processing - 2013. - Vol. 22(8). - P. 3204-3218. DOI: 10.1109/TIP.2013.2259834.
- [12] Chernov, A.V. Fast Method for Local Image Processing and Analysis / A.V. Chernov, V.V. Myasnikov, V.V. Sergeev // Pattern Recognition and Image Analysis - 1999. - Vol. 9(2). - P. 237-238.
- [13] Flusser, J. Pattern recognition by affine moment invariants / J. Flusser, T. Suk // Pattern recognition - 1993. Vol. 26(1). - P. 167-174. DOI: 10.1016/0031-3203(93)90098-H.
- [14] Hu, M.K. Visual pattern recognition by moment invariants / M.K Hu // IRE transactions on information theory - 1962. - Vol. 8(2). - P. 179-187. DOI: 10.1109/TIT.1962.1057692.
- [15] Мясников, В.В. Рекурсивный алгоритм вычисления свертки изображения с неразделимым двумерным полиномиальным ких-фильтром / В.В. Мясников // Компьютерная оптика. - 2004. - № 26. - С. 81-84.
- [16] Theodoridis, S. Machine learning: a Bayesian and optimization perspective / S. Theodoridis - San Diego: Academic Press, 2015. - 1062 p.
- [17] Alpaydin, E. Introduction to machine learning / E. Alpaydin - Cambridge: MIT press, 2014 - 584 p.
- [18] Hastie, T. Elements of statistical learning: data mining, inference, and prediction / T. Hastie, R. Tibshirani, J. Frieman - London: Springer, 2011. - 745 p.