

УДК 681.51:007.5

В. А. Гороховатский, С. В. Кузьмин

ИССЛЕДОВАНИЕ ДЕТЕКТОРОВ ТОЧЕК ИНТЕРЕСА ПРИ ПОСТРОЕНИИ ЛОКАЛЬНЫХ ПРИЗНАКОВ ИЗОБРАЖЕНИЙ

1. Введение

Построение эффективных систем локальных признаков при распознавании изображений является важнейшей проблемой современного компьютерного зрения [1–4]. Локальные признаки используются в структурных методах распознавания, которые обладают универсальностью и гибкостью применения в затрудненных условиях при действии помех и сложного фона [5, 6]. Распознавание изображений при этом сводится к сопоставлению их представлений в пространстве локальных признаков.

Под «точками интереса» в системах компьютерного зрения понимают существенные для процесса распознавания двумерные изменения сигнала [1, 4]. Реально на изображениях это соответствует углам, контрастным и изолированным точкам, концам кривых, линиям и т. д. Эти локальные признаки при геометрических преобразованиях изображений, как правило, почти не претерпевают изменений. Кроме того, множество точек интереса по мощности значительно меньше часто применяемого множества краевых (границных) точек, поэтому представляет интерес при построении эффективных методов структурного анализа и распознавания изображений [2]. Как правило, точки интереса используются в качестве отправного момента для построения в их окрестности более сложной векторной системы признаков [6], хотя имеют и самостоятельное применение [2, 4].

Целью работы является изучение свойств, особенностей, формализация, выбор параметров и компьютерное моделирование детекторов точек интереса для реальных изображений.

Задачи исследования — формализация описания признаков точек интереса, анализ вариантов их построения, применение их для решения проблемы представления изображений в признаковых пространствах, исследование статистических, инвариантных и вычислительных свойств этих признаков.

2. Формальное описание и пути построения детекторов точек интереса

Одним из путей построения признаков точек интереса является анализ автокорреляционной функции изображения, которая в значительной мере отражает локальные особенности сигнала. Автокорреляционная функция в дискретном варианте имеет следующий вид:

$$f(x, y, \Delta x, \Delta y) = \sum_{(x_k, y_k) \in W(x, y)} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2, \quad (1)$$

где $I(x, y)$ — функция яркости изображения, (x_k, y_k) — пиксели окна W с центром в точке (x, y) . Обычно окно $W(x, y)$ — это окрестность 5×5 с центром в точке (x, y) , по множеству пикселей которой вычисляется значение $f(x, y)$. С увеличением окрестности в целом растут помехозащищенность и вместе с тем вычислительные затраты, поэтому наиболее популярны на практике маски размеров 5×5 , 7×7 .

В литературе с общих позиций анализируют три варианта классификации области $W(x, y)$ фиксированных размеров. Решение принимается на основе значений автокорреляционной функции $f(x, y)$ от функции яркости $I(x, y)$, $(x, y) \in W$, заданной на этом множестве точек.

Пусть

$$\delta_1(x, y) = \min_{\Delta x, \Delta y} f(x, y, \Delta x, \Delta y),$$

$$\delta_2(x, y) = \max_{\Delta x, \Delta y} f(x, y, \Delta x, \Delta y), \quad \delta_{\max} = \max_{x, y} \delta_2(x, y).$$

Тогда область $W(x, y)$ можно классифицировать следующим образом [1]:

- Область приблизительно однородна — минимальное и максимальное по всем $(\Delta x, \Delta y)$ значения функции $f(x, y)$ незначительны (с точностью до некоторого порога δ_0) относительно глобального максимума $\delta_{\max}$ функции, определенного по всему изображению, т. е. когда выполнены условия

$$\delta_1 < \delta_0, \quad \delta_2 < \delta_0, \quad \delta_0 \ll \delta_{\max}.$$

Здесь символ $\ll$ понимается как «существенно меньше» (что в реальности обычно соответствует доли в 5...10 %).

- Краевая (границная) точка — минимальный отклик не велик, а максимальный отклик значителен относительно глобального максимума, т. е.

$$\delta_1 < \delta_0, \quad \delta_2 \approx \delta_{\max}, \quad \delta_0 \ll \delta_{\max}.$$

При этом направлении минимума и максимума, задаваемые параметрами $\Delta x, \Delta y$, должны быть перпендикулярны между собой в пределах дискретной сетки пикселей, хотя отдельные исследователи считают это условие необязательным.

- Точка интереса — минимальный и максимальный отклики по направлениям, определяемым параметрами $\Delta x, \Delta y$, значительны относительно величины глобального максимума, т. е.

$$\delta_1 \approx \delta_{\max}, \delta_2 \approx \delta_{\max} \quad (2)$$

Таким образом, точками интереса считаются пиксели изображения, полученные путем анализа множества откликов автокорреляционной функции в окрестности $W(x, y)$ по всем направлениям (в смысле 4- или 8-связности), причем оба оптимума превышают заданный порог $\delta \approx \delta_{\max}$.

В основе классификации на три типа лежит фундаментальное понятие кривизны поверхности, которую формирует функция яркости двух переменных в окрестности заданной точки [1].

Заметим, что классификация точек изображения на три указанных типа полностью определяется глобальными порогами $\delta, \delta_0, \delta_{\max}$. Реально все множество значений функции $f(x, y)$ разбивается на три области: $[0, \delta_0], (\delta_0, \delta], (\delta, \delta_{\max}]$. Пороги определяются, в основном, экспериментальным путем и зависят от контраста, освещенности и действия шумов, хотя при некоторых условиях существуют и оптимальные пути выбора порогов.

В результате анализа в соответствии с условиями (2) формируется множество точек интереса в виде

$$T = \{x, y, t\} \mid t = f(x, y, \Delta x, \Delta y), f(x, y, \Delta x, \Delta y) > \delta.$$

Затем сформированное множество T традиционно подвергается действию процедуры U отсеивания точек с целью сокращения количества пространственно группирующихся однотипных признаков [1]. Дело в том, что точки интереса на реальных изображениях имеют свойство группироваться в окрестности выделенного признака. Это связано со свойством гладкости автокорреляционной функции. В результате обработки имеем множество

$$T^* = U(T, \Omega),$$

где Ω — пространственная область, которой принадлежат точки из T , на основе анализа координат которых формируются точки множества T^* . Стандартная обработка сводится к анализу прямоугольной области изображения $\Omega = 7 \times 7$ точек с выбором максимальной по значению отклика из попавших в эту область точек. Увеличение размеров области Ω уменьшает количество точек интереса. Сформированный таким образом максимальный отклик приписывается координатам центра области Ω , а остальные точки окрестности выбрасываются из дальнейшего рассмотрения. Такая обработка получила название «немаксимального подавления значений» [4].

На базе анализа автокорреляционной функции построен ряд схем обработки [2, 3]. Например, детектор Моравица основан на поиске локального минимума в каждой точке среди откликов автокорреляционной функции по всем направлениям, опре-

деляемым вектором сдвига $(\Delta x, \Delta y)$. В дальнейшем минимумы используются для принятия решения о точке интереса с использованием заданного глобального порога. Таким образом, здесь множество точек интереса формируется в виде

$$M = \{x, y, t\} \mid t = \delta_1(x, y), \delta_1 > \delta.$$

Обычно порог δ задается как величина, пропорциональная значению максимального отклика детектора на всем множестве точек изображения. Динамика изменения откликов функции (1) зависит от контраста изображения и для области $W(x, y)$ размером 5×5 изменяется в пределах от 0 до 5×255 .

По сути, автокорреляционная функция аппроксимирует производную функции яркости по направлению, задаваемому параметрами $(\Delta x, \Delta y)$. Поэтому в детекторе Харриса [1] анализ автокорреляционной функции в точке (x, y) сводят к рассмотрению матрицы вида:

$$A(x, y) = \begin{bmatrix} \sum_{(x_k, y_k) \in W} (I_x(x_k, y_k))^2 & \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) & \sum_{(x_k, y_k) \in W} (I_y(x_k, y_k))^2 \end{bmatrix},$$

где I_x, I_y — частные производные функции изображения I в точке (x_k, y_k) .

С целью сглаживания помех суммирование значений частных производных в матрице $A(x, y)$ реализуется через свертку с матрицей гауссиана. Один из вариантов такой матрицы для сглаживания имеет вид

$$\begin{bmatrix} 0.779 & 0.882 & 0.779 \\ 0.882 & 1 & 0.882 \\ 0.769 & 0.882 & 0.779 \end{bmatrix}.$$

Кривизна автокорреляционной функции изображения определяется собственными значениями матрицы $A(x, y)$, которые в общем случае отражают понятия главной кривизны поверхности, что позволяет учитывать направления изменения значений кривизны [2].

Реализация детектора Харриса в результате сводится к анализу величины R , связанной с собственными значениями и построенной на основе взаимосвязи следа и определителя матрицы $A(x, y)$:

$$R = \text{Det} - \beta \text{Tr}^2,$$

где Tr — след (сумма элементов главной диагонали); Det — детерминант (определитель) матрицы; β — параметр, задающий свойства (классификацию) выделенных точек и фактически определяющий их количество в связи с пороговой обработкой величины R , $\beta \in [0, 1]$.

Разработано большое количество методов выделения краев (график), которые также опираются на информацию о производных от изображений. Основные принципы их построения сводятся к анализу связности пикселей или построению параметрических моделей для аппроксимации участков кривых. Для выделения краев, например, в настоящее время популярен детектор Кэнни [4], использующий на первом этапе значение и направление градиента, вычисленные с помощью оператора Робертса. Точки интереса отбрасываются детектором Кэнни ко множеству краевых точек путем сопоставления значения градиента с двумя порогами — на минимум возможного отклика и на обнаружение дополнительных точек-кандидатов, включаемых во множество пикселей формируемого края. На базе этих порогов и направлений градиентов соседних связанных точек можно формировать также и точки интереса. Однако, как показали эксперименты [1], краевые методы несколько проигрывают детектору Харриса по критерию инвариантности к группе движений.

Другое направление построения точек интереса связано с оценкой параметров аналитически заданных прямых и кривых на сформированных бинарных сигналах, полученных при обработке изображения. Его ограничением является привязка к какому-либо специфическому типу точек, например, L -, T -, X -, Y -, $\uparrow$ - углам [1].

3. Результаты компьютерных экспериментов

Целью моделирования был анализ целесообразности и эффективности применения детекторов точек интереса к решению задач распознавания реальных изображений. Распознавание в такой постановке сводится к сопоставлению множеств точек интереса некоторого объекта с соответствующими множествами для эталонов.

Одно из направлений исследования состоит в оценке влияния параметра β детектора Харриса на качество и устойчивость выделения локальных признаков.

На рис. 1 приведены исходные бинарные изображения, включающие все типы классифицируемых точек, а также величины откликов при различных β , нормализованные в целях отображения к диапазону яркостей 0...255.

Проведенные эксперименты показывают, что при изменении коэффициента β в пределах от 0 до 0,1 увеличивается величина отклика R для однородных областей, резко выделяются краевые точки, уменьшается количество выделенных точек интереса, вплоть до их полного слияния с откликами однородных областей при $\beta = 0,1$. Инверсия исходных изображений приводит к аналогичному результату.

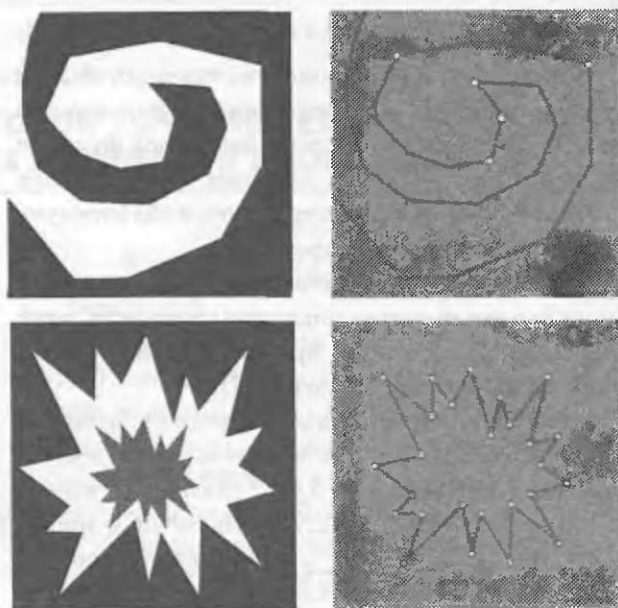


Рис. 1. Тестовые изображения и отклики на них детектора Харриса при $\beta = 0,08$

Известно, что детектор Харриса, как, впрочем, и другие детекторы, обладает инвариантностью к преобразованиям смещения (так как реализуется путем сканирования) и к повороту (поскольку анализ направлений изменения автокорреляционной функции учитывает ее повороты на 45° вместе с изображением).

С целью анализа влияния величины выделяемых углов на значение отклика R для бинарных изображений в качестве тестового использовалось изображение, которое представляет собой сектор с фиксированным значением угла (см. рис. 2). Результаты экспериментов для разных значений углов приведены в таблице 1. Согласно результатам экспериментов, наибольший отклик дают углы сектора в $30...90^\circ$, существенно меньший — $0...15^\circ$, а углы $120...165^\circ$ дают отрицательный отклик и поэтому относятся детектором к краевым точкам, а не к точкам интереса.

Приведем теперь результаты в виде гистограммы значений, где по горизонтальной оси отложены диапазоны принимаемых детектором Харриса значений, а по вертикали — количество точек в каждом

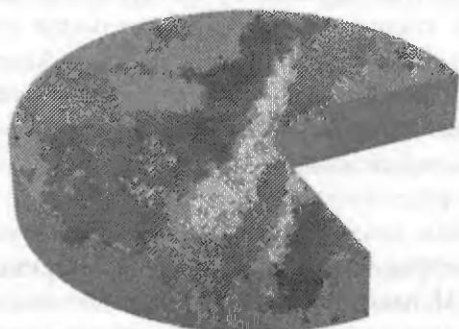


Рис. 2. Тестовое изображение сектора с фиксированным углом

Таблица 1

Значения откликов детектора Харриса для изображений углов
(диапазон 15...120°, горизонталь)
от величины угла поворота сектора рис. 2 (вертикаль)

Углы	15	30	45	60	75	90	105	120
0	2,98	12,33	12,33	13,37	11,69	9,55	3,94	-0,03
20	3,56	12,73	12,73	14,86	10,31	10,31	8,23	-1,54
40	6,31	12,30	12,30	13,44	9,02	7,57	3,89	-0,29
60	2,41	14,85	14,85	15,51	10,82	10,46	4,3	-1,09
80	1,19	12,25	12,25	13,21	9,25	6,08	4,3	-2,14
100	4,08	14,20	14,21	13,41	15,53	7,22	10,59	-1,5
120	6,65	17,42	17,42	11,65	10,89	5,73	3,7	3,69
140	1,48	13,93	13,93	15,64	8,56	8,23	3,88	5,73
160	2,49	7,96	7,96	10,08	12,55	15,58	10,94	-1,48
180	2,99	12,30	12,30	13,33	13,22	9,55	4,87	1,41
Среднее	3,41	13,02	13,02	13,22	11,18	9,03	5,86	0,27
Дисперсия	3,36	5,86	5,86	2,62	4,60	8,08	8,42	6,70

Таблица 2

Отклики лапласиана гауссоиды и детектора Харриса для угла 45°
при разных углах поворота сектора

Углы поворота	0	20	40	60	80	100	Среднее	Дисперсия
Лапласиан гауссоиды	-0,607	-3,213	-1,213	-1,82	-1,213	-1,82	-1,64767	0,79
Детектор Харриса	12,33	12,73	12,30	14,85	12,25	14,21	13,02	5,86

из них. Гистограммы (рис. 3, а) показывают, что инвариантность детектора Харриса к повороту понимается как стабильность (в некотором диапазоне) множества наибольших значений отклика. Гистограмма после удаления откликов краевых точек, соответствующих отрицательным значениям R , приобретает вид рис. 3, б. Основная масса значений здесь относится к максимальному диапазону (3...5). Поэтому обнаружение и распознавание объектов по точкам интереса можно построить путем анализа множества значений детектора, близких к максимальному значению.

Разновидностью точек интереса являются углы типа L, T, X, Y, ↑ [2, 4]. Эксперименты показали, что все они устойчиво определяются описанными детекторами.

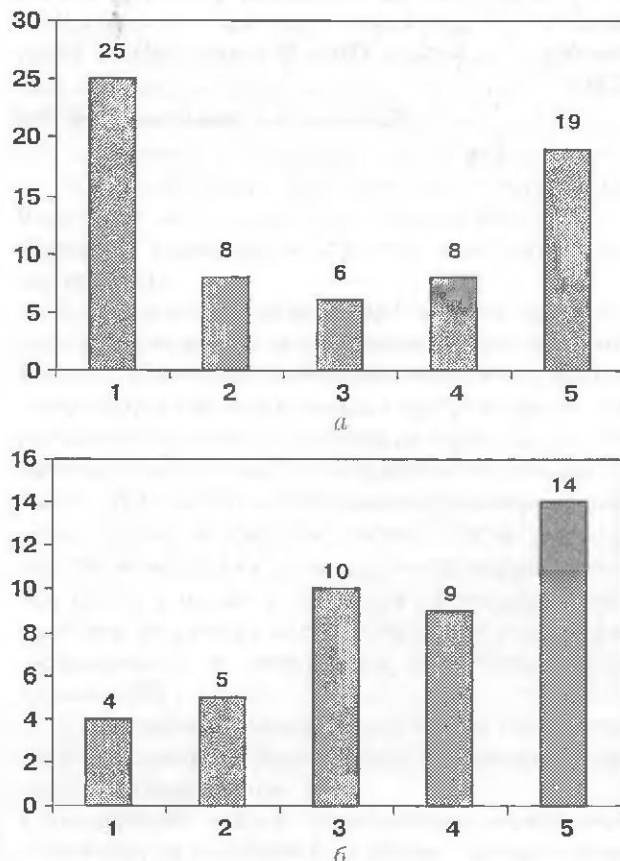


Рис. 3. Гистограммы значений детектора Харриса:

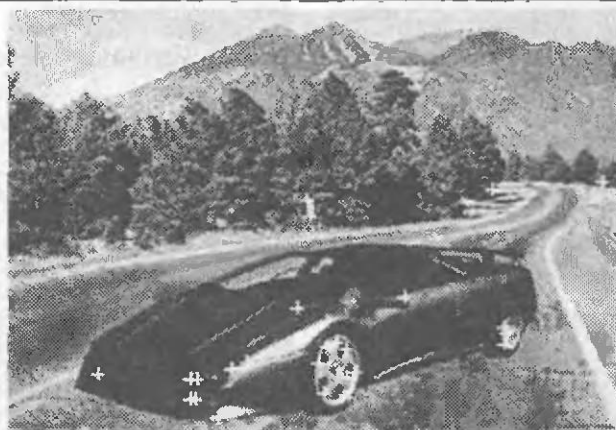
а — полная, б — после удаления откликов краевых точек

Из-за особенности построения многоэтапной обработки при реализации детектора Харриса возникают «перескоки» откликов между собой в пределах группы углов, поэтому при необходимости распознавания углов разного собой нужно использовать дополнительные признаки, позволяющие различать углы разных размеров в пределах уже выделенных групп точек. Одним из популярных признаков такого рода является лапласиан гауссоиды. Для сравнения в таблице 2 приведены значения откликов, а также среднее и дисперсия лапласиана гауссоиды и детектора Харриса для угла 45° градусов. Разброс значений для лапласиана гауссоиды, как видно из таблицы 2, несколько меньше, чем у детектора Харриса.

Эксперименты на полутоновых изображениях показали следующие результаты. На рис. 4 приведены исходное изображение (рис. 4, а), а также изображение с преобразованиями смещения и поворота на 20° на сложном фоне (рис. 4, б). Доля правильно перемещенных в процессе преобразования точек интереса составляет около 94 %. Как видим из рисунков, на фоне возникает некоторое дополнительное количество выделенных признаков.



а



б

Рис. 4. Преобразование полутонового изображения: а — выделена 31 точка; б — выделены 28 точек

4. Выводы

Новизна предложенного подхода состоит в построении, формальном описании и оценке свойств эффективной системы признаков — точек интереса для целей распознавания объектов на изображениях.

Важной особенностью детектора Харриса является то, что он может успешно справляться как с детектированием углов, так и краев (в зависимости от задаваемого параметра β). Второй особенностью при детектировании углов является чувствительность детектора к их величине — изображения углов в пределах от 30° ... 90° дают существенно больший отклик, нежели остальные.

Функция R детектора Харриса принимает положительные значения в точках интереса, отрицательные — в краевых точках и близкие к нулю значения в областях изображения, которые можно приближенно считать однородными. При выделении краев принимают $\beta \geq 1$, а для точек интереса β принимает значения в диапазоне $\beta = 0.04$... 0.06 .

Практическая значимость полученных результатов заключается в экспериментальном подтверждении таких свойств детекторов, как инвариантность, устойчивость в пределах некоторого диапазона выделяемых углов, связь с классическими детекторами краев, универсальность применения для разных видов локальных признаков.

Список литературы: 1. Schmid C., Mohr R., Bauckhage C. Evaluation of interest point detectors // *International Journal of Computer Vision*. — 2000. — 37, No. 2. — P. 151–172. 2. Harris C., Stephens M. A combined edge and corner detector // *Alvey Vision Conf.* — 1988. — 4. — P. 189–192. 3. Schmid C., Mohr R. Local Greyvalue Invariants for Image Retrieval // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. — 1997. — 19, No. 5. — P. 530–534. 4. Шапиро Л., Стокман Дж. Компьютерное зрение. — Вином. Лаборатория знаний, 2006. — 752 с. 5. Гороховатский В. А. Распознавание изображений в условиях неполной информации. — Харьков: ХНУРЭ, 2003. — 112 с. 6. Пулятин Е. П., Гороховатский В. О., Мамат О. О. Методы та алгоритми комп'ютерного зору: Навч. посібник. — Харків: ТОВ «Компанія СМІТ», 2006. — 236 с.

Поступила в редакцию 27.08.2006