

УДК 519.6+004.9

ОПТИМАЛЬНЫЕ РЕАЛИЗАЦИИ МЕТОДОВ ОБНАРУЖЕНИЯ ПРОСТРАНСТВЕННО РАЗДЕЛЕННЫХ ОБЪЕКТОВ В БИНАРНЫХ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЯХ

Е.А. КРАСНОБАЕВ

(Витебский государственный университет им. П.М. Машиерова)

Рассматриваются методы маркирования связанных компонент цифровых бинарных изображений. На их основе в результате исследования разработаны оптимальные программные алгоритмы выделения пространственно разделенных объектов в бинарных изображениях, полученных при сегментации движущихся объектов в видеопоследовательностях. Данные программные алгоритмы позволяют определять координаты центра масс объектов, выделять их прямоугольными рамками и выполнять их распознавание. Важным этапом исследования явилась оценка времени работы алгоритмов с целью выполнения требований работы в режиме реального времени. Установлено, что время обработки одного кадра видеоряда составило 25...28 мс, а точность распознавания силуэта человека – 72 %.

Введение. Наряду с полноцветными и полутоновыми изображениями при решении широкого круга задач компьютерного зрения используются бинарные (двухградационные) изображения. Использование бинарных изображений дает преимущество в условиях ограничения вычислительных ресурсов, а также в тех случаях, когда полноцветные изображения несут избыточную информацию. Переход к бинарным изображениям упрощает выделение силуэтов, контуров объектов в изображениях, а также их обнаружение и распознавание [1].

Одна из широкоизвестных областей компьютерного зрения – анализ движения в видеопоследовательностях. В случае, когда необходимо определять пространственное расположение движущихся объектов, используются методы сегментации движения. Результатом сегментации при этом является бинарное изображение, определяющее наличие или отсутствие движения в пикселе изображения. Вследствие чего возникает задача обнаружения пространственно разделенных объектов в таких бинарных изображениях, а также определения их характеристик: координат, площади и др. Для решения этой задачи нами предлагается использовать так называемые алгоритмы «маркирования связанных компонент» изображения. Данные алгоритмы приводят к выделению групп пикселей, принадлежащих каждому объекту, по принципу взаимной пространственной расположенности составляющих их пикселей.

В связи с ресурсоемкостью задачи обработки изображений в реальном времени становится актуальной разработка оптимальных по быстродействию алгоритмов обнаружения пространственно разделенных объектов в бинарных видеопоследовательностях с целью определения их характеристик: координат центра, координат описанных вокруг них прямоугольников.

Маркирование связанных компонент бинарных изображений. Пиксели в бинарном изображении B имеют две градации яркости – 1 (переднеплановые пиксели) и 0 (фоновые пиксели). Для определения объектов в бинарных изображениях необходимо выделить группы переднеплановых пикселей, принадлежащих каждому объекту на изображении. Пиксели в каждой группе при этом будут объединены по принципу пространственной расположенности или соседства (связанности). Различают две разновидности связанности пикселей: 4-связанность и 8-связанность. В первом случае у каждого пикселя $B(i, j)$ в качестве соседей рассматриваются четыре пикселя: $B(i-1, j)$, $B(i+1, j)$, $B(i, j-1)$, $B(i, j+1)$ (рис. 1, а). Во втором случае – соседними считаются восемь пикселей $B(i-1, j)$, $B(i+1, j)$, $B(i, j-1)$, $B(i, j+1)$, $B(i-1, j-1)$, $B(i+1, j-1)$, $B(i-1, j+1)$, $B(i+1, j+1)$ (рис. 1, б).

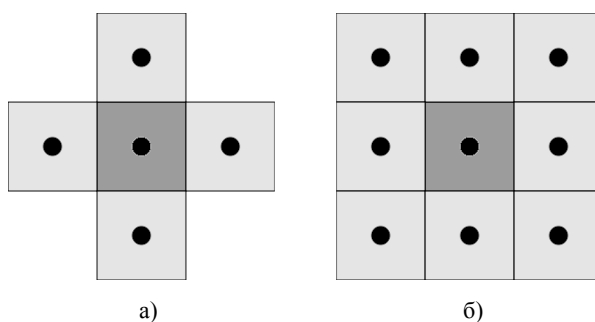


Рис. 1. Связанность компонент бинарного изображения:
а – четыре соседа; б – восемь соседей

В соответствии с [2], если B бинарное изображение и $B\langle c \rangle = B\langle c' \rangle = v$, где $v = 0$ или $v = 1$, то пиксель $\langle c \rangle$ является связанным с пикселем $\langle c' \rangle$, если есть последовательность пикселей $r, c = r_0, c_0, r_1, c_1, \dots, r_n, c_n = r', c'$, в которой $B\langle c_i \rangle = v, i = 1..n$ и $\langle c_i \rangle$ соседи $\langle c_{i-1} \rangle$ для каждого $i = 1..n$. Последовательность пикселей $\langle c_0 \rangle, \langle c_1 \rangle, \dots, \langle c_n \rangle$ соединяет между собой пиксели $\langle c \rangle$ и $\langle c' \rangle$. Связанные компоненты являются множеством пикселей S , равных v , таких, что каждая пара пикселей в множестве является соседями и их значения равны v .

Таким образом, маркирование связанных компонент бинарного изображения B представляет собой получение маркированного изображения LB , в котором значение яркости каждого пикселя равно метке группы связанных компонент (пикселей).

Алгоритмы и методы маркирования связанных компонент бинарных изображений. Для быстрого маркирования связанных компонент бинарных изображений в основном используются два подхода: рекурсивный подход и «двухпроходной метод выделения эквивалентных областей». Первый метод основан на принципе рекурсивного обхода соседей каждого пикселя бинарного изображения. Проверка начинается с первого произвольного переднепланового пикселя. Если хотя бы один из его соседей оказывается также переднеплановым пикселем, то проверяются и его соседи и т.д.

Построим рекурсивный алгоритм маркирования пикселей бинарного изображения с целью определения координат центров объектов и координат описанных вокруг них прямоугольников:

- 1) для каждого пикселя $B\langle j \rangle$ бинарного изображения и для каждого пикселя $LB\langle j \rangle$ маркированного изображения;
 - 2) вызов рекурсивной подпрограммы:
 - а) проверка возможного выхода за пределы изображения;
 - б) проверка принадлежности пикселя $B\langle j \rangle$ к переднему плану. В случае если значение пикселя равно 1, продолжаем алгоритм; в обратном случае – присваиваем значение текущему пикселю $LB\langle j \rangle$, равное 0, и переходим к пункту 3;
 - в) увеличиваем счетчик пикселей k_m текущего объекта m на 1;
 - г) определяем, являются ли координаты текущего пикселя $B\langle j \rangle$ максимальными или минимальными в массиве пикселей m -го объекта, заносим их в x_{\min}^m, y_{\min}^m и x_{\max}^m, y_{\max}^m ;
 - д) присваиваем значение текущему пикселю $B\langle j \rangle$, равное нулю (с целью исключения повторной его проверки);
 - е) присваиваем значение текущему пикселю $LB\langle j \rangle$, равное значению текущей метки m ;
 - ж) вызываем рекурсивную подпрограмму для каждого соседнего пикселя $B\langle -1, j \rangle, B\langle +1, j \rangle, B\langle j-1 \rangle, B\langle j+1 \rangle$ – пункт 2;
 - 3) увеличиваем значение m на 1.

Результатом работы является маркированное изображение LB , значение яркости пикселей которого будет равно метке группы связанных компонент. При этом определяется площадь каждого m -го выделенного объекта k_m , а также левая верхняя и правая нижняя координаты x_{\min}^m, y_{\min}^m и x_{\max}^m, y_{\max}^m прямоугольника, описанного вокруг объекта. Условный центр объекта при этом определяется следующим образом:

$$\begin{aligned} x_c^m &= x_{\min}^m + \frac{1}{2} (x_{\max}^m - x_{\min}^m) ; \\ y_c^m &= y_{\min}^m + \frac{1}{2} (y_{\max}^m - y_{\min}^m) . \end{aligned} \quad (1)$$

Данный алгоритм является рекурсивным, и для исключения переполнения стека следует избегать большой глубины рекурсии или выполнять его с использованием параллельных вычислений.

Второй алгоритм, в отличие от первого, состоит из двух проходов: при первом проходе каждому пикселю изображения назначается некоторая метка, на втором проходе эквивалентные метки объединяются в разделенные области. Построим рекурсивный алгоритм маркирования пикселей бинарного изображения с целью определения координат центров объектов и координат описанных вокруг них прямоугольников:

Первый проход бинарного изображения:

- 1) для каждого пикселя $B\langle j \rangle$ бинарного изображения и для каждого пикселя $LB\langle j \rangle$ маркированного изображения;
- 2) проверка принадлежности пикселя $B\langle j \rangle$ к переднему плану. В случае если значение пикселя равно 1, продолжаем алгоритм; в обратном случае – присваиваем значение текущему пикселю $LB\langle j \rangle$, равное 0, и переходим к пункту 1;

3) проверка соседних пикселей:

а) если верхний соседний пиксель $LB \llbracket -1, j \rrbracket$ и левый соседний пиксель $LB \llbracket j-1 \rrbracket$ являются переднеплановыми и имеют метку m , метке $LB \llbracket j \rrbracket$ присваивается значение m ;

б) если хотя бы один верхний соседний пиксель $LB \llbracket -1, j \rrbracket$ или левый соседний пиксель $LB \llbracket j-1 \rrbracket$ является переднеплановым и имеет метку m , метке пикселя $LB \llbracket j \rrbracket$ присваивается значение m ;

в) если верхний соседний пиксель $LB \llbracket -1, j \rrbracket$ имеет метку m и левый соседний пиксель $LB \llbracket j-1 \rrbracket$ – другую метку n , метке пикселя $LB \llbracket j \rrbracket$ присваивается значение m . Вносим метки m и n в эквивалентную таблицу E ;

г) если верхний соседний пиксель $LB \llbracket -1, j \rrbracket$ и левый соседний пиксель $LB \llbracket j-1 \rrbracket$ являются фоновыми, метке пикселя $LB \llbracket j \rrbracket$ присваивается новое значение l .

Второй проход маркированного изображения $LB \llbracket j \rrbracket$:

1) объединение эквивалентных меток в эквивалентной таблице E . Если m и n две эквивалентные метки в E , то переназначают n в m ; если $m < n$ – наоборот;

2) определяем, являются ли координаты текущего пикселя $LB \llbracket j \rrbracket$ максимальными или минимальными в массиве пикселей m -го объекта, заносим их в $\llbracket x_{\min}^m, y_{\min}^m \rrbracket$ и $\llbracket x_{\max}^m, y_{\max}^m \rrbracket$. Увеличиваем счетчик пикселей k_m текущего объекта m на 1.

Распознавание бинарных изображений. Для распознавания бинарных изображений, полученных на предыдущем этапе, удобно использовать моментные характеристики, а также рассчитываемые на основе них моментные инварианты [3]. Моментные характеристики представляют собой некоторые взвешенные суммы пикселей изображения, характеризующие такие свойства изображения, как площадь, центр тяжести, ориентация и др. Основным достоинством моментных инвариантов является устойчивость к поворотам изображений, сдвигу, масштабу и другим преобразованиям.

Для цифровых изображений, представленных дискретной функцией $I \llbracket x, y \rrbracket$, выражение для вычисления простых моментов имеет вид:

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y).$$

Теорема единственности утверждает, что если $I \llbracket x, y \rrbracket$ кусочно-непрерывна и имеет ненулевые значения только в конечной части плоскости, моменты всех порядков существуют и однозначно определяются функцией $I \llbracket x, y \rrbracket$. На практике вычисляются только некоторые моменты нижнего порядка. Простые моменты изображения определяют площадь (для бинарного изображения) M_{00} , координаты центра $\llbracket x_c, y_c \rrbracket \equiv \llbracket M_{10}/M_{00}, M_{01}/M_{00} \rrbracket$.

Чаще при распознавании изображений применяют центральные моменты, обладающие инвариантностью к сдвигу изображения:

$$\mu_{pq} = \sum_x \sum_y (x - x_c)^p (y - y_c)^q I(x, y).$$

Центральные моменты второго порядка позволяют получать информацию об ориентации изображения и могут быть нормированы. В результате этого они будут инвариантны к изменению масштаба и параллельному переносу:

$$\eta_{ij} = \frac{\mu_{ij}}{\binom{i+j}{2} \mu_{00}}.$$

С помощью центральных моментов определяются характеристики изображения, инвариантные к многим преобразованиям. Семь функций, основанных на центральных моментах, инвариантных к параллельному переносу, изменению масштаба и вращению, ввел М. Ху [4]:

$$\phi_1 = \eta_{20} + \eta_{02};$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2;$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2;$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2;$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right];$$

$$\phi_6 = (\eta_{20} - \eta_{02}) \left[\eta_{30} + \eta_{12}^2 - \eta_{21} + \eta_{03}^2 \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03});$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \left[\eta_{30} + \eta_{12}^2 - 3\eta_{21} + \eta_{03}^2 \right] + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \left[3\eta_{30} + \eta_{12}^2 - \eta_{21} + \eta_{03}^2 \right].$$

Данные моментные инварианты предлагается рассматривать как признаки распознаваемых объектов.

Определим диапазон значений моментных инвариантов для некоторого заданного класса эталонных объектов. В качестве распознаваемого класса выберем наиболее часто анализируемый объект – силуэт идущего человека, полученный в результате сегментации движения в видеопоследовательности (рис. 2). Значения первых четырех моментов X_u для данной последовательности приведены на рисунке 3.

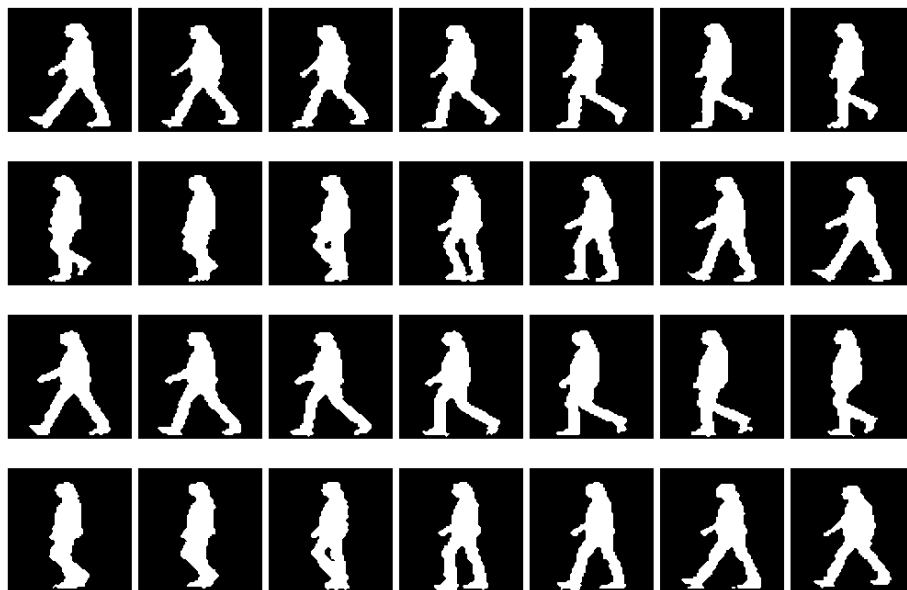


Рис. 2. Последовательность бинарных изображений силуэта идущего человека

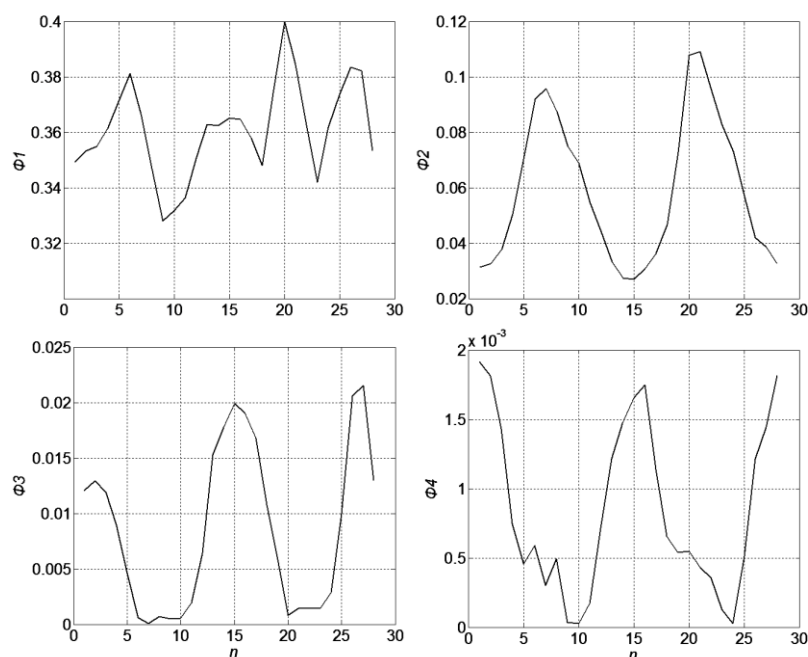


Рис. 3. Зависимости значения моментов X_u от номера кадра

Анализируя графики, видно, что диапазон значений моментов лежит в строго заданном числовом интервале. В итоге распознавание объекта бинарного изображения может осуществляться путем проверки на принадлежность значений его первых моментов X_0 диапазону значений эталонного класса образов.

В результате исследования установлено, что точность распознавания силуэта человека данным способом составила 72 %. Необходимо отметить, что для повышения точности распознавания, следует либо повышать количество инвариантов изображений, либо максимально полно задавать ряды эталонных изображений.

Закключение. В результате данного исследования разработаны оптимальные программные алгоритмы маркирования связанных компонент бинарных изображений, полученных при сегментации движущихся объектов в видеоизображениях, с последующим их распознаванием. Важным этапом исследования является оценка времени работы алгоритмов с целью выполнения требований работы в режиме реального времени. Известно, что стандартная видеопоследовательность поступает с видеокамеры в режиме реального времени со скоростью порядка 20...25 кадров/с. Следовательно, временная задержка между кадрами составляет около 40...50 мс, именно этим временем и должны ограничиваться алгоритмы обработки кадра видеопоследовательности.

Проведена оценка времени работы алгоритмов на тестовых изображениях (таблица) с различным разрешением на компьютере с процессором AMD Athlon X2 4400+, объемом ОЗУ 2 Гб, среда разработки – Microsoft Visual C# 2005 Express.

Время работы алгоритмов с тестовыми изображениями

Вид алгоритма	Время (мс)		
	320 × 240	640 × 480	800 × 600
Рекурсивный	11	21	28
Двухпроходной	9	17	25

Как видно из таблицы, даже для изображения с разрешением 800 × 600 все алгоритмы вкладываются в заданный временной интервал.

Помимо разрешения на время работы алгоритмов влияет также количество переднеплановых пикселей (выражающие размеры и количество объектов) в бинарном изображении. Зависимости процента переднеплановых пикселей от общего объема и времени работы алгоритмов для бинарного изображения с разрешением 320 × 240 приведены в графиках на рисунке 4.

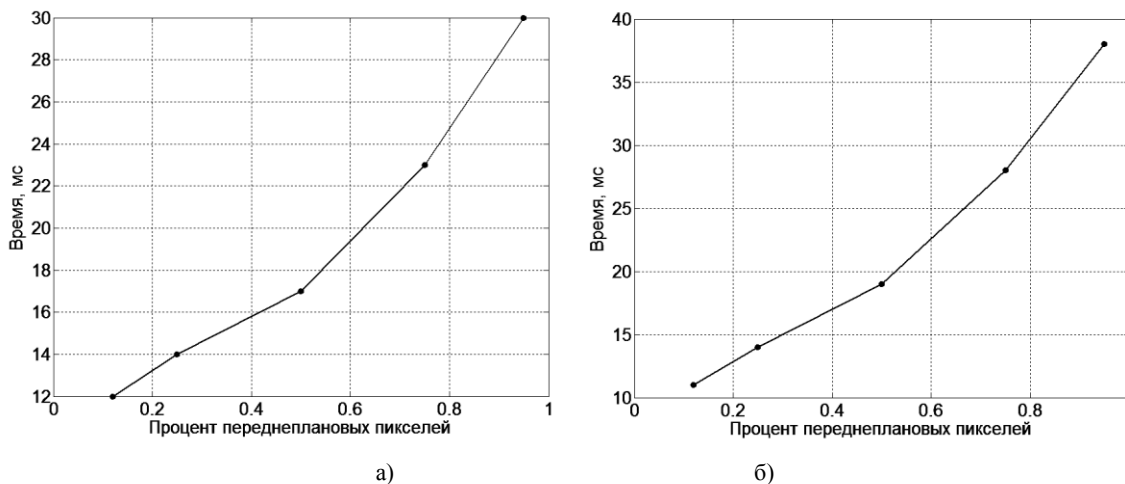


Рис. 4. Зависимости времени выполнения алгоритмов от процента переднеплановых пикселей в изображении с разрешением 320 × 240 пикселей: а – рекурсивный алгоритм; б – двухпроходной алгоритм

Анализируя рисунок 4 можно отметить, что рост времени выполнения практически одинаков для двух рассматриваемых алгоритмов.

На рисунке 5 приведен результат обработки бинарного изображения, которое является результатом сегментации движения тестовой видеопоследовательности. Определены и выведены координаты центров объектов, а также описанные вокруг них прямоугольники.



Рис. 5. Исходное изображение (а) и отсегментированное бинарное изображение с выделением объектов, определением координат их центров и описанных прямоугольников (б)

В качестве тестовых последовательностей использовались видеоролики камеры видеонаблюдения School of Informatics of The University of Edinburgh [5]. Данные алгоритмы реализованы в разрабатываемом программном комплексе детектирования и сопровождения движущихся объектов в видеоизображениях в режиме реального времени, а также могут использоваться для решения широкого класса задач компьютерного зрения.

ЛИТЕРАТУРА

1. Фурман, Я.А. Цифровые методы обработки и распознавания бинарных изображений / Я.А. Фурман. – Красноярск: Изд-во Краснояр. ун-та, 1992. – 248 с.
2. Shapiro, L. Computer vision / L. Shapiro, G. Stockman. – Prentice Hall, 2000. – 617 p.
3. Flusser, J. Rotation Moment Invariants for Recognition of Symmetric Objects / J. Flusser, N. Suk // IEEE Trans. Image Proc. – 2006. – Vol. 15. – P. 3784 – 3790.
4. Hu, M. Visual Pattern Recognition by Moment Invariants / M. Hu // IRE Trans. Info. Theory. – 1962. – Vol. IT-8. – P. 179 – 187.
5. Caviar test case scenarios [Electronic resource] / School of Informatics of The University of Edinburgh. – 2004. – Mode of access: <http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1>. – Date of access: 06.10.2009.

Поступила 23.01.2010

OPTIMAL IMPLEMENTATION OF METHODS FOR THE DETECTION OF SPATIALLY SEPARATED OBJECTS IN BINARY VIDEO SEQUENCES

E. KRASNOBAEV

The methods for labeling connected components of digital binary images are considered. Based on them, developed optimal software algorithms of allocation of spatially separated objects in binary images obtained by the segmentation of moving objects in video sequences. These algorithms allow to determine the coordinates of the mass center of objects, allocate their rectangles and carry out the recognition. An important step in the investigation was to evaluate the speed of algorithm for real time. Found that the processing time of one frame video sequence was 25...28 ms, and the recognition accuracy of the silhouette man – 72 %.