

СИСТЕМЫ АНАЛИЗА
И СИНТЕЗА ИЗОБРАЖЕНИЙ И СИГНАЛОВ

УДК 619.24

В. А. ГОРОХОВАТСКИЙ, Ю. Н. ЕРЕСЬКО, Е. П. ПУТЯТИН,
В. И. СТРЕЛЬЧЕНКО
(Харьков)

ЛОКАЛИЗАЦИЯ ОБЪЕКТОВ
НА ИЗОБРАЖЕНИЯХ ВИЗУАЛЬНЫХ СЦЕН

Введение. При построении систем технического зрения интеллектуальных роботов возникает задача создания эффективных алгоритмов для выделения отдельных объектов из сцен и оценки параметров объектов. Сложность построения таких алгоритмов состоит в необходимости обработки больших объемов информации, а также в том, что объекты могут претерпевать преобразования относительно их эталонных описаний. Другим важным моментом, усложняющим эффективный анализ, является неравномерное освещение объектов и наличие в поле зрения фоновых объектов [1], воспринимаемых как помехи. Под объектом в задачах анализа изображений понимают множество точек изображения, характеризующихся общими признаками [2].

Применение корреляционных алгоритмов к решению задачи локализации приводит к значительному объему вычислений. Наличие фоновых образований снижает надежность этих алгоритмов.

Такие методы локализации, как анализ контурных признаков, классические структурные методы, метод моментных инвариантов [2], требуют предварительного выделения точек объектов из фона, что является самостоятельной сложной задачей.

В [1, 3, 4] развит иерархический подход, заключающийся в том, что вначале производят анализ и формирование признаков отдельных фрагментов объекта, по которым впоследствии из структурных пространственных соображений формируются признаки объекта. Локализация осуществляется по этим признакам. Такой подход, безусловно, перспективен в силу его универсальности для широкого разнообразия объектов и визуальных сцен в целом, а также из-за простой обработки на уровнях формирования признаков объектов и принятия решения об их местоположении на изображении. Кроме того, при переходе от одного уровня иерархии к другому сжимается анализируемый объем информации. При практической реализации алгоритмов, основанных на иерархическом подходе, появляется возможность распараллеливания обработки и решения задач в реальном времени.

Настоящая работа посвящена созданию быстродействующих алгоритмов локализации объектов на изображениях при наличии сложного фона на основе использования иерархического подхода.

Формализация и обоснование алгоритма. Пусть $B(x, y) \in W_1(D)$ — изображение, а

$$W_1(D) = \{B(x, y), (x, y) \in D \subset R^2 | 0 \leq B(x, y) \leq I\}.$$

Здесь D — поле зрения; I — максимальная яркость изображений. Пусть

© 1990 Гороховатский В. А., Ересько Ю. Н., Путятин Е. П., Стрельченко В. И.

$x \in X, y \in Y$, где $X = [-a, a]$, $Y = [-b, b]$, и, следовательно, $D = X \times Y$, где \times — декартово произведение. Под объектом будем понимать множество

$$A = \{(x, y) \in D \mid Q(x, y) = 1\},$$

где Q — правило, устанавливающее соответствие точек поля зрения объекту: если $Q(x, y) = 0$, то $(x, y) \notin A$.

В предположении широкополосности регистрирующей системы изображение $B(x, y)$ можно представить в виде

$$B(x, y) = G(x, y) + [1 - Q(G(x, y))] \Phi(x, y) + \xi(x, y), \quad (1)$$

где $G(x, y)$ — финитная функция яркости объекта; $\Phi(x, y)$ — фоновое изображение; $\xi(x, y)$ — флуктуационный шум. В таком случае A — носитель $G(x, y)$, т. е. $A = \text{supp } G(x, y)$, а правило Q соответствует характеристической функции.

Создание алгоритма сводится к построению такого отображения $\Omega = \Omega_1 \circ \Omega_2 \circ \dots \circ \Omega_v$, которое осуществляет переход $W_1(D) \rightarrow D$, где v — число уровней иерархии.

Предлагаемый алгоритм имеет пять уровней иерархии, на каждом из которых решается своя подзадача, связанная с подзадачами других уровней. Представим действие каждого уровня отображением Ω_i . Суть этих отображений состоит в следующем:

$$1. \Omega_1: W_1(D) \begin{cases} \rightarrow W_2(X) \\ \rightarrow W_2(Y) \end{cases},$$

где Ω_1 — оператор развертки по направлениям ox и oy ; $W_2(\cdot)$ — множество одномерных функций, каждый элемент которого получен дискретизацией $B(x, y) \in W_1(D)$ по выбранным ортогональным направлениям:

$$W_2(X) = \{\{\zeta_i(x) = B(x, -b + ih_y), \quad i = 0, 1, \dots, n_1\}\};$$

$$W_2(Y) = \{\{\eta_i(y) = B(-a + ih_x, y), \quad i = 0, 1, \dots, n_2\}\};$$

$n_1 + 1, n_2 + 1$ — число строк и столбцов соответственно; $h_x = 2a/n_1, h_y = 2b/n_2$.

Переход к анализу одномерных функций значительно упрощает и дает возможность распараллелить дальнейшую обработку.

$$2. \Omega_2: W_2(X) \rightarrow W_3^x, \quad \Omega_2: W_2(Y) \rightarrow W_3^y,$$

где $W_3^{x(y)}$ — конечномерное векторное множество локальных признаков.

Отображение Ω_2 осуществляет сжатие информации, так как преобразует непрерывные функции в набор W_3 векторов небольшой размерности:

$$W_3^x = \{\{q_i^{kx} = \{q_{i1}^{kx}, q_{i2}^{kx}, \dots, q_{in_1}^{kx}\}; \quad i = 0, 1, \dots, n_1; \quad k = 1, 2, \dots, r_i\}\};$$

$$W_3^y = \{\{q_i^{ky} = \{q_{i1}^{ky}, q_{i2}^{ky}, \dots, q_{in_2}^{ky}\}; \quad i = 0, 1, \dots, n_2; \quad k = 1, 2, \dots, s_i\}\},$$

где i — номер строки (столбца); k — номер фрагмента функции $\zeta_i(x) \times \eta_i(y)$ в i -й строке (столбце), определяемого вектором $q_i^{kx} (q_i^{ky})$; $r_i (s_i)$ — количество фрагментов в i -й строке (столбце); n_3 — размерность векторов q_i^{kx} . Компоненты векторов q_i^{kx} содержат информацию о координате центра фрагмента, его длине, полярности относительно фона и об интегральных характеристиках яркости выделенного фрагмента.

Априорной информацией, используемой для выделения фрагментов строк (столбцов), являются наличие скачка первой производной в граничных точках объекта, диапазон изменения длительностей одномерных «срезов» объекта по каждому направлению развертки, а также некоторые другие характеристики, такие, как яркость и контраст изображения объекта. Для выделения одномерных сигналов существует набор известных алгоритмов [5], в основу которых положены как оптимальные, так

и эвристические подходы. Большинство из них неприменимы в нашем случае из-за необходимости наличия априорных сведений о распределении вероятностей, описывающих фоновые изображения, и ограничений, накладываемых требованиями работы в реальном времени.

$$3. \Omega_3: W_3^x \rightarrow W_4^x, \Omega_3: W_3^y \rightarrow W_4^y,$$

где W_4^x, W_4^y — множества признаков объекта, элементами которых являются векторы v_i^x, v_i^y размерности n_4 .

Введем отношение упорядоченности Λ на множествах векторов из W_3^x, W_3^y . Все элементы, удовлетворяющие отношению Λ , т. е. когда $q_i^{kx} \wedge q_j^{ky}$ — истина, будем называть связными. Предполагается, что одному объекту соответствует одна или несколько групп связных элементов из множеств W_3^x и W_3^y . Ω_3 по заданному отношению упорядоченности определяет связные элементы и для каждой группы таких элементов вычисляет компоненты векторов $v_i^x \in W_4^x$ и $v_i^y \in W_4^y$.

При формировании векторов v_i^x, v_i^y учитываются априорные сведения о возможных размерах объектов по осям координат, в результате чего из дальнейшего анализа частично исключаются ложные объекты.

Таким образом, каждый набор векторов $\{q_i^{kx} \wedge q_j^{ky}\}$ отображается в единственный вектор $v_i^x = \{v_{i1}^x, v_{i2}^x, \dots, v_{in_4}^x\}$, а набор векторов $\{q_i^{ky} \wedge q_j^{lx}\}$ — в вектор v_i^y . Векторы v_i^x, v_i^y содержат информацию о пространственном положении связного набора в области D , контрасте, геометрических размерах, интегральных свойствах функций яркости объектов в поле зрения.

4. Применим отображение $\Omega_4: W_4^x \rightarrow D, \Omega_4: W_4^y \rightarrow D$, получив координаты объекта независимо по каждому из множеств W_4^x и W_4^y . Отображение Ω_4 осуществляется путем оптимизации сходства векторов из W_4^x, W_4^y с заданными эталонными наборами векторов $e^x = \{e_i^x\}, e^y = \{e_i^y\}$.

При этом эталонный объект может описываться как одним, так и несколькими векторами, что дает возможность локализации широкого разнообразия форм объектов, а также объектов, отдельные части которых имеют различный контраст по отношению к фону. В таком случае функция сходства должна учитывать структурные связи векторов эталона. Правило принятия решения можно представить как

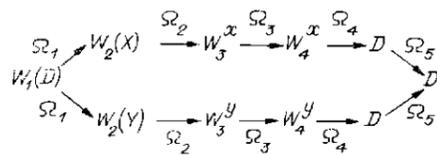
$$\{v_i^{x*} \infty e^x \mid \arg [\text{opt } \rho(v_i^x, e^x)] = v_i^{x*}; P(\rho(v_i^{x*}, e^x), \delta) = 1\}, \quad (2)$$

где ∞ — знак эквивалентности; ρ — мера сходства; P — предикат, связанный с пороговым ограничением меры сходства; δ — порог. Введение предиката вызвано необходимостью снижения вероятности определения ложных объектов. Однако при этом повышается и вероятность отказа в принятии решения по одному из множеств W_4^x, W_4^y .

Анализируем векторы, полученные по правилу (2), и определяем соответствующие им координаты.

5. Отображением $\Omega_5: D \rightarrow D$ сопоставляются координаты, соответствующие векторам v_i^{x*}, v_i^{y*} , и принимается окончательное решение о координатах искомого объектов. Процедура принятия таких решений основана на пространственных структурных особенностях анализируемых объектов.

Порядок обработки изображения алгоритмом схематично показан на диаграмме



Как видим, алгоритм обладает следующими особенностями:

1. Иерархичность, позволяющая распараллелить и упростить обработку на отдельных уровнях, а также сократить размерность задачи от уровня к уровню.

2. Идентичная обработка по двум ортогональным направлениям, что дает возможность анализировать широкий класс форм объектов и объекты с различной полярностью их частей относительно фона.

3. Согласованная обработка (учет априорной информации об объектах, сопоставление с эталонным описанием) на этапах Ω_2, Ω_4 , что обеспечивает помехозащищенность алгоритма.

Реализация алгоритма. Конкретный вид алгоритма определяется отображениями $\Omega_1, \Omega_2, \dots, \Omega_5$. В [6] нами предложен вариант алгоритма, где в качестве Ω_2 используется одномерный интегродифференциальный фильтр с импульсной характеристикой вида

$$H_m(x) = -h_m(x+l) + 2h_m(x) - h_m(x-l), \quad (3)$$

где

$$h_m(x) = \underbrace{(\Pi_l * \Pi_l * \dots * \Pi_l)}_m(x). \quad (4)$$

* — символ свертки; m — кратность свертки; $\Pi_l(x)$ — функция прямоугольного импульса длительностью l .

Функция h_m аппроксимирует гауссиану вида

$$\varphi(x) = C \exp(-6x^2/ml^3),$$

C — константа. Фильтр с импульсной функцией h_m является квазиоптимальным [5] для длительностей сигналов $l_0^* = \sqrt{ml^3/3}$, а фильтр с функцией отклика (3) — центрирующим для длительностей сигналов $l_0 \leq 2l$. Сверточная аппроксимация вида (4) с амплитудой импульсов Π_l , равной 1, обеспечивает реализацию фильтра без операции умножения, что приводит к высокому быстродействию фильтрации.

Многочисленные эксперименты, проведенные на ЭВМ, показали, что для практической реализации достаточно взять $m=3$, так как дальнейшее увеличение m практически не улучшает соотношения сигнал/шум на выходе фильтра, но зато увеличивает время фильтрации. Эксперименты по оценке помехозащищенности алгоритма фильтрации (3) показали его высокую эффективность [6]. Так, например, при входном соотношении сигнал/шум $\mu=3$ обеспечивается вероятность правильного определения координат центров импульсных сигналов не ниже 0,97. После пороговой обработки результатов фильтрации формируются векторы q_i^k , компонентами которых являются полярность, амплитуда отклика фильтра, координата центра импульса.

Отображение Ω_3 реализовано следующим образом. Отношение Λ задавалось предикатом, устанавливающим совпадение полярностей, близость координат и амплитуд откликов фильтра между соседними строками (столбцами). В результате сформирован набор векторов v_i , компонентами которых являются полярность, координаты центров связанных групп, число связанных элементов (размер), сумма амплитуд этих элементов.

Отображение Ω_4 реализовано как селекция тех векторов из полученных наборов, которые близки к эталонным векторам без учета координат, но с учетом пространственных структурных соотношений между компонентами эталона. В качестве меры близости ρ выбрано евклидово расстояние.

Эксперименты. Эффективность и работоспособность алгоритма были подтверждены экспериментами на реальных изображениях, характерных для задач робототехники, таких, например, как распознавание элементов низковольтной аппаратуры, выделение отдельных элементов печатных плат. Эксперименты проводились на многомашинном комплексе обработки изображений, включающем микропроцессорную телевизионную систему ввода — вывода изображений, ЭВМ СМ-4, ЕС-1050.

Для оценки устойчивости предложенного алгоритма к помехам типа «световое пятно» и аддитивному гауссовому шуму на ЭВМ синтезировались эти помехи. Моделирование работы алгоритма проводилось при изменении отношения сигнал/шум μ от 2 до 15. Эксперименты показали, что для широкого класса объектов вероятность правильного определения их местоположения была не меньше 0,95 при $\mu \geq 5$. Работоспособность алгоритма практически не зависит от наличия помех типа «световое пятно», от изменения параметров геометрических преобразований в пределах 10 %, а также от изменения контрастности объектов относительно фона в пределах 30 %. Эффективность алгоритма зависит от близости геометрической формы искомого и ложных объектов.

В результате экспериментов установлено, что разработанный алгоритм позволяет с высокими показателями достоверности и быстродействия решать практические задачи анализа изображений.

СПИСОК ЛИТЕРАТУРЫ

1. Гороховатский В. А., Кацалап С. Ф., Пуятин Е. П. Анализ изображений в условиях локальных искажений // Автометрия.— 1986.— № 6.
2. Прэтт У. Цифровая обработка изображений.— М.: Мир, 1982.— Кн. 1, 2.
3. Павлидис Т. Алгоритмы машинной графики и обработки изображений — М.: Радио и связь, 1986.
4. Распознавание образов и обработки изображений // ТИИЭР.— 1979.— 67, № 5.
5. Гуткин Л. С. Теория оптимальных методов радиоприема при флуктуационных помехах.— М.: Сов. радио, 1972.
6. Пуятин Е. П., Гороховатский В. А., Ереско Ю. Н., Стрельченко В. И. Метод экстремально-логической фильтрации в задачах автоматической сегментации изображений // АСУ и приборы автоматики.— Харьков: Вища шк., 1987.— Вып. 83.

Поступила в редакцию 2 июля 1987 г.

УДК 621.397

Н. Н. КРАСИЛЬНИКОВ

(Ленинград)

ОБОБЩЕННАЯ ФУНКЦИОНАЛЬНАЯ МОДЕЛЬ ЗРЕНИЯ И ЕЕ ПРИМЕНЕНИЕ В СИСТЕМАХ ОБРАБОТКИ И ПЕРЕДАЧИ ИЗОБРАЖЕНИЙ

При проектировании и оптимизации параметров систем обработки и передачи изображений серьезные трудности вызывает отсутствие формализованного критерия качества воспроизведения изображений, функционально связанного с электрооптическими характеристиками самой системы.

В настоящее время для оценки качества воспроизведения изображения телевизионной системой используется либо метод экспертных оценок, либо набор так называемых частных критериев, к числу которых относятся: разрешающая способность, отношение сигнал/шум, число воспроизводимых градаций яркости, погрешности воспроизведения цвета и т. д. Их значения определяются параметрами системы и могут быть вычислены. Недостатком первого метода является отсутствие математических выражений, устанавливающих связь между оценкой и электрооптическими характеристиками аппаратуры, недостатком второго — невозможность сравнивать системы, которые различаются между собой по разным частным критериям. Так, например, если система *A* обеспечивает большую разрешающую способность, но меньшее отношение сигнал/шум, чем система *B*, то, не опираясь на опыт, невозможно сказать, какая из них лучше. Из этого следует, что метод экспертных оценок не может быть использован при аналитическом решении оптимизационных задач,

© 1990 Красильников Н. Н.