

КОМПЬЮТЕРНЫЕ СИСТЕМЫ
АНАЛИЗА И СИНТЕЗА ИЗОБРАЖЕНИЙ

УДК 528.8 : 512.642

Л. В. КАСПЕРОВИЧ, М. И. КОЛЕСНИК, Д. А. УСИКОВ
(Москва)

БЫСТРЫЙ АЛГОРИТМ ОКОНТУРИВАНИЯ ОБЪЕКТОВ
НА ИЗОБРАЖЕНИЯХ

Введение. Этап выделения контуров является ключевым в решении большинства задач анализа изображений. Разработке эффективных алгоритмов выделения контуров посвящено большое число публикаций (см., например, [1, 2]). Характерное время работы таких алгоритмов определяется размером входного изображения и быстродействием процессора. Время обработки страницы изображения размером 256×256 элементов простейшими алгоритмами составляет на мини-ЭВМ класса СМ-4 порядка 10 мин.

С появлением современных цифровых дисплейных систем (ЦДС), или, иначе, видеопроцессоров, открылась возможность существенно ускорить решение этой задачи. Использование аппаратных возможностей ЦДС позволяет добиться снижения времени расчетов в 100 раз, что дает заметный экономический эффект с учетом важности этапа выделения контуров в алгоритмах распознавания образов. Следует отметить, что задачи классификации такие, как выделение стыков, пересечений линий и т. п., тоже успешно решаются методами, описываемыми ниже.

В настоящей работе сделана также попытка формализовать, используя аппарат векторной алгебры, описание алгоритмов, ориентированных на архитектуру видеопроцессоров, с целью достижения максимума удобства и краткости в изложении.

Общее описание алгоритма выделения контуров. Выделение контуров на реальных полутоновых изображениях осложнено существованием большого числа градаций яркости (обычно до 256 градаций). В этом случае исходное изображение преобразуют к такому виду, когда яркостей всего две (так называемая бинаризация изображения). Однако выделить объекты на изображении, бинаризованном по следующему правилу: яркости, не превышающие некоторую фиксированную величину, переводятся в 0, остальные — в 1 — обычно невозможно, так как объект на исходном изображении может не иметь ярко выраженных границ. Эффективной оказывается процедура бинаризации с программируемым порогом в скользящем окне, которая в простейшем случае заключается в следующем. Пусть F_{ij} — значение яркости элемента изображения с координатами (i, j) . Вычисляется средняя яркость \bar{F}_{ij} в окне $n \times n$ (n нечетное) с центром, имеющим координаты (i, j) , по формуле

$$\bar{F}_{ij} = \frac{1}{n^2} \sum_{k=i-(n-1)/2}^{i+(n-1)/2} \sum_{l=j-(n-1)/2}^{j+(n-1)/2} F_{kl}. \quad (1)$$

Положим $\tilde{F}_{ij} = F_{ij} - \bar{F}_{ij}$. Тогда значение яркости B_{ij} бинаризованного

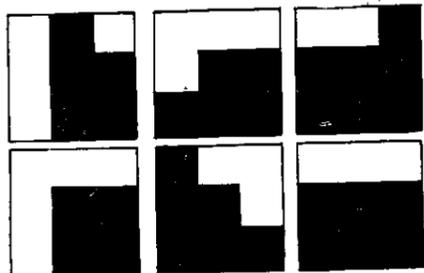


Рис. 1. Основные шаблоны границ в окне размером 3×3 элемента. Существует еще 30 шаблонов, которые получаются из основных поворотами на 90° и зеркальными отражениями

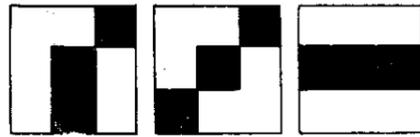


Рис. 2. Основные шаблоны линий в окне размером 3×3 элемента. Остальные 9 шаблонов — результат поворотов на 90° и зеркальных отражений

изображения в точке (i, j) определяется с помощью пороговой операции

$$B_{ij} = \begin{cases} 1, & \text{если } \hat{F}_{ij} \geq 0; \\ 0, & \text{если } \hat{F}_{ij} < 0. \end{cases}$$

На полученном бинарном изображении производится поиск заданных геометрических объектов в скользящем окне 3×3 . Для этого формируется система элементарных объектов (шаблонов), с которыми сравнивается бинаризованное изображение. На рис. 1 приведены 36 шаблонов для определения граничных точек на изображении, а для распознавания того, принадлежит ли граничная точка некоторой линии, требуется еще 12 шаблонов (рис. 2). Бинаризованное изображение просматривается скользящим окном, и определяются расстояния Хэмминга между фрагментом изображения в окне и каждым из шаблонов сформированной системы. Если минимальное из полученных расстояний не превышает некоторой заданной величины, то полагается, что данный фрагмент содержит элементарный объект-шаблон.

Устранение разрывов в сформированных таким способом контурах производится аналогичным сравнением с шаблонами линий.

На рис. 3 приведен обработанный алгоритмом оконтуривания снимок Памира, полученный с космического корабля «Союз-22».

Алгебраический подход к описанию архитектуры видеопроцессора. Видеопроцессор типа СВИТ [3], I^2S [4] и другие является конвейерным (pipe-line) процессором систолического типа, состоящим из нескольких параллельно работающих структурных единиц (рис. 4). Обычно для кодирования яркости пиксела страницы памяти ЦДС требуется 8 бит, однако для простоты записи операций в окне 3×3 условимся считать, что каждый пиксел памяти видеопроцессора кодируется 9-ю битами. Переложение алгоритмов с 9-битовой памяти на 8-битовую носит чисто технический характер.

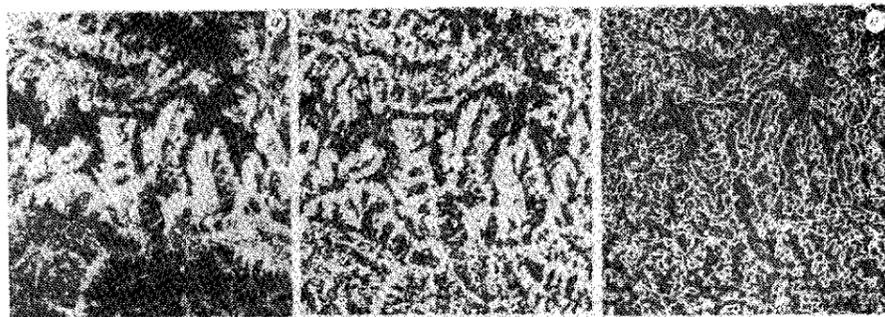


Рис. 3. Исходный снимок (а), бинаризованное изображение (б), оконтуренное изображение (в)

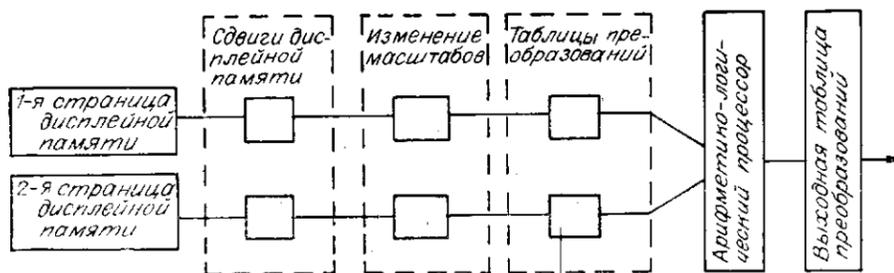


Рис. 4. Блок-схема структурной единицы видеопроцессора

Для удобства записи алгоритма введем операторные обозначения аппаратных единиц видеопроцессора (здесь и далее $[a/b]$ равняется целой части a/b , $\{a/b\} = a - b * [a/b]$).

Определение 1. Назовем картой $F = \langle F_0, \dots, F_9 \rangle$ кортеж матриц $F_i \in M_{256}(Z_2)$, $0 \leq i \leq 9$. Фрагментом (циклическим) карты F в окне размером $n \times m$, имеющим координаты левого верхнего угла $(i_0, j_0) \times \times (1 \leq i_0, j_0, n, m \leq 256)$, будем называть карту G такую, что если $F_i = \{f_{kl}^i\}$; $G_i = \{g_{kl}^i\}$ ($0 \leq i \leq 9$), то

$$g_{kl}^i = \begin{cases} f_{kl}^i, & \text{если } k = \{(i_0 + s - 2)/256\} + 1, 1 \leq s \leq n, \\ & l = \{(j_0 + t - 2)/256\} + 1, 1 \leq t \leq m; \\ 0 & \text{в остальных случаях.} \end{cases}$$

В данной работе будет использоваться лишь случай $n = m = 3$.

Понятие карты объединяет в себе представление о содержимом страницы памяти ЦДС, с одной стороны, как об изображении реальных объектов, а с другой — об информации, полученной в процессе обработки и не имеющей естественной природы.

Матрицы F_i в определении 1 играют роль битовых слоев страницы памяти видеопроцессора.

Множество всех карт как прямое произведение векторных пространств $M_{256}(Z_2)$ наделяется естественной структурой векторного пространства относительно операций, индуцированных операциями матричного сложения и умножения на скаляр, а именно

$$F + G = \langle F_0 + G_0, \dots, F_9 + G_9 \rangle; \lambda F = \langle \lambda F_0, \dots, \lambda F_9 \rangle, \lambda \in Z_2.$$

Определение 2. Назовем отображение Δ дополнением карт, если $\Delta: \langle F_0, \dots, F_9 \rangle \rightarrow \langle 0, F_1 + 1, \dots, F_9 + 1 \rangle$, 1 — матрица из пространства $M_{256}(Z_2)$, все элементы которой равны 1.

Определение 3. Фиксируем вектор $u = (u_1, \dots, u_9) \in Z_2^9$. Отображение Π назовем маской вектора u , если $\Pi: \langle F_0, \dots, F_9 \rangle \rightarrow \langle 0, u_1, F_1, \dots, u_9, F_9 \rangle$.

Определение 4. Пусть $a, b, k, l \in Z$, $1 \leq k, l \leq 256$. Пусть также $q_a(\cdot)$ — некоторая целочисленная функция со множеством значений $[1, 256]$, зависящая от параметра a . Отображение Θ назовем координатным преобразованием карт на вектор (a, b) , если $\Theta: \langle F_0, \dots, F_9 \rangle \rightarrow \langle QF_0, \dots, QF_9 \rangle$, где Q — оператор, действующий на пространстве $M_{256}(Z_2)$ следующим образом:

$$(QA)_{k,l} = A_{q_a(k), q_b(l)}.$$

$$\text{Если } q_a(k) = \begin{cases} 256 - \{(a - k)/256\}, & \text{если } a \geq k; \\ 1 + \{(k - a - 1)/256\}, & \text{если } a < k, \end{cases}$$

то координатное преобразование назовем циклическим сдвигом карт на вектор (a, b) . Если же $a, b \geq 0$ и $q_a(k) = [(k - 1)/a] + 1$, то такое координатное преобразование назовем зуммированием карт на вектор (a, b) .

Определение 5. Пусть L — некоторый оператор на Z_2^{10} . Отображение Λ назовем таблицей преобразования яркости (ТПЯ) или «look-up»

table», если $\Lambda: \langle F_0, \dots, F_9 \rangle \rightarrow \langle G_0, \dots, G_9 \rangle$, где $G_i = [g_{kl}^i]$ ($1 \leq i \leq 9$) определяются как

$$(0, g_{kl}^1, \dots, g_{kl}^9) = L(f_{kl}^0, \dots, f_{kl}^9).$$

Оператор L назовем характеристическим оператором отображения Λ .

Определение 6. Внешним произведением карт назовем билинейное отображение $P = F \otimes G$, где P_i ($0 \leq i \leq 9$) является результатом почечного умножения матриц $P_i = F_i \otimes G_i$.

Определение 7. Внешним объединением карт назовем отображение $P = F \odot G$ такое, что $P = F \otimes G + F + G$.

Определение 8. Внешней суммой карт назовем отображение $P = F \oplus G$, действующее на картах F и G ($F_0 = G_0 = 0$) следующим образом:

$$P_9 = F_9 + G_9;$$

$$P_i = F_i + G_i + (F_{i+1} + G_{i+1}) \underset{0 \leq i \leq 8}{\odot} (I + P_{i+1}) + F_{i+1} \odot G_{i+1}.$$

Определение 9. Внешней разностью карт назовем отображение $P = F \ominus G$ такое, что $P = F \oplus \Delta G$.

В приведенных определениях внешнее произведение описывает реализованную в ЦДС операцию логического умножения, внешнее объединение — логическое сложение, а оператор дополнения — операцию отрицания.

Поскольку при фиксированных k и l вектор $f = (f_{kl}^0, \dots, f_{kl}^9) \in Z_2^{10}$ суть двоичное представление яркости пиксела изображения с координатами $(k-1, l-1)$, то определим функцию $\text{dec}(f) = \sum_{i=0}^9 f_{kl}^i * 2^{9-i}$, которая устанавливает соответствие между ними.

В соответствии с введенными обозначениями элементарное действие структурной единицы видеопроцессора можно представить формулой

$$R = \Lambda(AP \circ BG), \quad (2)$$

где P и G — карты из входного потока ($F|F_0 = 0$); R — результирующая карта; A и B — операторы из множества $\{\Lambda, \Theta, \Pi, \Delta\}$; \circ — одна из внешних операций.

Поскольку любое действие структурной единицы видеопроцессора является композицией элементарных, то можно считать, что формула (2) полностью описывает работу ЦДС. Например, так называемый виток обратной связи, реализованный в ЦДС, записывается в виде $R = \Pi_u P \odot \odot \Pi_v G$. Координаты вектора \bar{u} удовлетворяют равенству $\bar{u}_i = u_i + 1$ ($1 \leq i \leq 9$).

Реализация алгоритма на ЦДС. Перейдем к детальному описанию алгоритма выделения контуров, используя введенную терминологию.

Бинаризация. Усреднение исходной карты F по формуле (1) производится рекурсивной процедурой ($\bar{F} = R[n^2]$):

$$P[j] = \Lambda_r P[j-1] \oplus Q_j F \quad (2 \leq j \leq n^2), \quad P[1] = F;$$

$$R[j] = \Lambda_r P[j] \oplus R[j-1] \quad (2 \leq j \leq n^2), \quad R[1] = 0,$$

в которой характеристические операторы ТПЯ Λ_r и Λ_p определяются следующим образом:

$$L_r v = v_0 \text{dec}^{-1} \circ c(v), \quad L_p v = (1 - v_0)v + v_0 \text{dec}^{-1} \circ d(v),$$

где $c(v) = [\text{dec}(v)/n^2]$; $d(v) = \{\text{dec}(v)/n^2\}$.

Сдвиги Θ_j выбирались так, чтобы обход окна $n \times n$ осуществлялся по раскручивающейся от центрального элемента спирали. Координаты вектора сдвига (a_j, b_j) на шаге j ($j = (2t-1)^2 + s$, $1 \leq s \leq 8t$, $t = 1, 2, \dots, (n-1)/2$) вычисляются по формулам

$$a_j = (-1)^{\lfloor (s/(2t)+1)/2 \rfloor} * (t - \{s/(2t)\}) * \text{sign}(\{([s/(2t)]+1)/2\}); \quad (3)$$

$$b_j = (-1)^{\lfloor s/(4t) \rfloor} * (t - \{s/(2t)\}) * \text{sign}(\{[s/(2t)]/2\}). \quad (4)$$

Формула (2) для построения бинарной карты преобразуется к виду $B = \Lambda(F \ominus F)$. ТПЯ Λ определяется своим характеристическим оператором $L(v_0, \dots, v_9) = (0, \dots, 0, v_0)$.

Упаковка. Элементы бинарной карты из окна 3×3 помещаются в битовые слои элемента упакованной карты с координатами центра окна. Процедура упаковки записывается в виде

$$P[j] = \Pi_{u[i] \ominus \Lambda_i B} \ominus \Pi_{v[i]} P[j-1] \quad (2 \leq j \leq 9);$$

$$P[1] = \Pi_{u[1]} \Lambda_1 B.$$

Вектор $u[j]$ имеет единственную отличную от нуля координату j , $\bar{u}_i[j] = u_i[j] + 1$, $1 \leq i \leq 9$. Координаты вектора сдвига (a_j, b_j) на шаге j вычисляются по формулам (3), (4) при $n=3$. ТПЯ Λ_j на этом шаге имеет характеристический оператор L такой, что

$$(Lv)_i = v_9(1 - \text{sign}(|i-j|)), \quad 0 \leq i \leq 9.$$

Поиск шаблонов. Расстоянием Хэмминга между векторами v и $w \in Z_2^r$ называется целочисленная функция ρ такая, что

$$\rho(v, w) = \sum_{i=1}^r \{(v_i + w_i)/2\}.$$

Расстояние между картами F и G определяется следующим образом:

$$\rho(F, G) = \sum_{i=0}^9 \rho(F_i, G_i) \quad (r = 256^2).$$

Задача классификации бинарной карты B состоит в получении карты R минимальных расстояний от фрагментов карты B до всех заданных шаблонов и карты P номеров соответствующих шаблонов, на которых такое расстояние достигается. Этот процесс реализуется на ЦДС следующим образом. Каждый шаблон системы упаковывается соответственно в вектор $b[k]$ (k — номер шаблона), и строятся карты $R = \Lambda_r P[9]$, $P = \Lambda_p P[9]$.

Характеристические операторы ТПЯ Λ_r и Λ_p определяются формулами

$$L_r v = \text{dec}^{-1} \circ l(v), \quad L_p v = \text{dec}^{-1} \circ m(v),$$

где

$$l(v) = \min_k \rho(v, b[k]);$$

$$m(v) = \min \{m_1, \dots, m_k | \rho(v, b[m_i]) = l(v), 1 \leq i \leq k\}.$$

Визуализация классифицированных структур. Заключительный этап состоит в построении изображения классифицированных структур, т. е. в том, как воссоздать совокупность элементарных объектов. Фрагментами подобной карты являются объекты, расстояния от которых до соответствующих фрагментов бинаризованной исходной карты не превышают некоторой заданной величины, причем если таких объектов несколько, то выбирается элементарный объект с минимальным номером из тех, на которых достигается минимальное расстояние.

Зафиксируем n и построим карту G , которая будет содержать номера шаблонов, расстояние до которых от фрагментов карты B не превышает n :

$$G = P \otimes \Lambda R.$$

ТПЯ Λ определяется своим характеристическим оператором L :

$$(Lv)_i = 1 - \text{sign}(\text{dec}(v) + |\text{dec}(v) - n| - n), \quad 1 \leq i \leq 9.$$

Карта классифицированных структур получается рекурсивной процедурой, обратной процедуре упаковки:

$$R[j] = \Lambda_r G \ominus \bar{\Lambda}_r [j-1] \quad (2 \leq j \leq 9), \quad R[1] = \Lambda_r G,$$

в которой на шаге j ($1 \leq j \leq 9$) ТПЯ A_j определяется характеристическим оператором $L_j v = (0, \dots, 0, b_j[\text{dec}(v)])$, а координаты вектора сдвига такие же, как в процедуре упаковки.

В заключение отметим возможные модификации алгоритма. Во-первых, простое изменение элементарных признаков (см. рис. 1 и 2) позволяет с той же скоростью находить любые бинарные геометрические структуры в скользящем окне 3×3 (точки, стыки, пересечения и др.). Во-вторых, поиски структур в окнах большего размера можно производить многоступенчатой классификацией в окнах 3×3 . Таким образом, открывается возможность для поиска сложных геометрических структур на аэрокосмических изображениях.

СПИСОК ЛИТЕРАТУРЫ

1. *Proceedings of the 7-th Conference on Pattern Recognition.*—North Holland Press, 1984. — P. 42.
2. Прэтт У. Цифровая обработка изображений.— М.: Мир, 1982.
3. Чесалов Л. С., Лубман С. В., Борисенко В. И. и др. Самостоятельный видеотерминальный терминал СВИТ.— М., 1982.— (Препр. ИКИ АН СССР; 724).
4. Adams I., Wallis R. New concepts in display technology // *Computer.*— 1977.— 10, N 8.— P. 61.

Поступила в редакцию 20 июня 1988 г.

УДК 778.38 : 621.373

А. И. ЗАВАЛИН, В. Ф. ЛАМЕКИН, В. Л. СМЕРНОВ
(Москва)

ИССЛЕДОВАНИЕ ПИКОСЕКУНДНЫХ ДИНАМИЧЕСКИХ ГОЛОГРАММ В ПОЛУПРОВОДНИКОВЫХ ЭНИТАКСИАЛЬНЫХ СТРУКТУРАХ НА ОСНОВЕ СОЕДИНЕНИЙ $A^{III}B^V$

Для сверхскоростной обработки информации, представленной в виде двумерных изображений, перспективным является применение динамических голограмм, сформированных на основе светондуцированных неравновесных носителей заряда (ННЗ) в полупроводниках. В [1] исследовалась запись динамических голограмм двумерных изображений на светондуцированных ННЗ в кристаллическом кремнии, однако используемая длительность записывающих лазерных импульсов (7 нс) и большая величина коэффициента амбиполярной диффузии кремния (~ 10 см²/с) существенно ограничивали разрешающую способность регистрирующей среды (до 100 мм⁻¹) и частоту циклов записи до 20 МГц. В то же время, как показывают результаты записи динамических дифракционных решеток с помощью ультракоротких импульсов (УКИ) [2, 3], в полупроводниках состава GaAs_xP_{1-x} диффузия вносит малый вклад в распад решеток и слабо ограничивает разрешающую способность ($\leq 2-3 \times 10^3$ мм⁻¹). Кроме обычных динамических голограмм, перспективными являются поляризационные динамические голограммы, позволяющие записывать и восстанавливать поляризацию предметной волны. В [4, 5] показана возможность записи поляризационных динамических дифракционных решеток в кристаллическом кремнии с помощью УКИ и в сульфиде кадмия с помощью наносекундных импульсов.

Цель данной работы — исследование возможности записи динамических (в том числе поляризационных) голограмм двумерных изображений с помощью индуцированных пикосекундными световыми импульса-