

УДК 621.377.2

Информационная технология обнаружения объектов на изображении в режиме скользящего окна / Н. И. Глумов, Э. И. Коломиец, В. В. Сергеев // Научное приборостроение. — 1993. — Т. 3. — № 1: Дифракционная плоская оптика и обработка изображений. — С. 72—88.

В работе предложена информационная технология обработки изображений и обнаружения объектов в режиме скользящего окна, допускающая эффективную реализацию с точки зрения вычислительной сложности и качества обнаружения. Рассмотрены основные этапы преобразования данных: предварительная обработка изображения, рекурсивное вычисление локальных признаков, формирование поля значений дискриминантной функции, локализация объектов. Предложена процедура параметрической настройки технологии, реализуемая в форме обучения линейного классификатора. Библ. 22 назв. Ил. 1.

---

Н. И. Глумов, Э. И. Коломиец, В. В. Сергеев

(Самарский государственный аэрокосмический университет)

## ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ В РЕЖИМЕ СКОЛЬЗЯЩЕГО ОКНА

*An information technique has been proposed and aimed at data processing and object detection using a gliding window and allowing effective realization in terms of computational intricacy and detection reliability. The main stages of data transformation are as follows: preliminary image processing, recursive computation of local indications, formation of the field of the discriminant function values and localization of the object. The procedure of the technique parametric adjustment is realized as that of linear classifier learning.*

### 1. Постановка задачи

Проблема распознавания зрительных образов имеет много уровней сложности, определяемых спецификой решаемых прикладных задач: от простейшей "поэлементной" классификации точек наблюдаемого яркостного поля до построения описания сложной трехмерной сцены [1—3]. Среди этих задач важное практическое значение имеет задача анализа двумерного оптического сигнала — изображения, содержащего некоторую совокупность локальных областей ("объектов"), отличающихся по яркости или текстурным характеристикам от остальных участков ("фона").

Локальность объектов, означающая, что каждый из них может быть вписан в область малого (по сравнению с размером изображения) радиуса, упрощает процедуры обработки данных по сравнению с общим случаем, так как позволяет ограничиваться анализом поля яркости в пределах небольшого окна обработки, потенциально содержащего объект. Специфика типового процесса видеозамеров (формирование отсчетов цифрового изображения в порядке строчной развертки) и неопределенность положения объектов приводят к тому, что окно обработки целесообразно делать "скользящим", т. е. занимающим все возможные положения на плоскости изображения по ходу развертки.

Обработка данных в режиме скользящего окна характерна для известных цифровых корреляторов изображений [2, 4—6]. Она удобна тем, что при надлежащей аппаратурной поддержке позволяет ана-

лизировать изображения в реальном масштабе времени (в темпе видеоизмерений). Однако с увеличением размеров окна обработки традиционные корреляторы быстро усложняются и становятся практически нереализуемыми из-за чрезмерного объема необходимых вычислений. В этой ситуации представляется перспективным применение алгоритмов, реализующих идею рекурсивной обработки изображений в скользящем окне [6—8], которые рассматриваются в рамках предлагаемой информационной технологии.

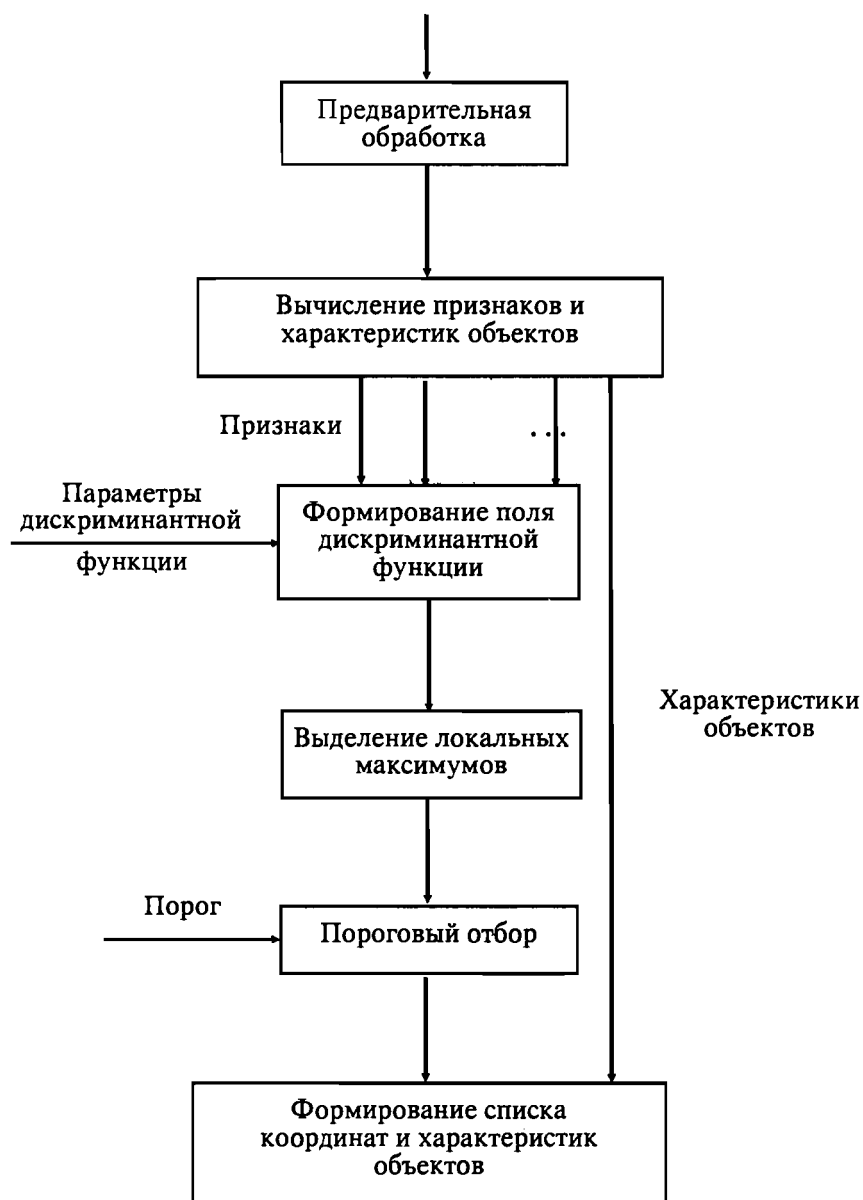
## 2. Информационная технология обработки данных

Будем считать, что имеет место следующая модель формирования и наблюдения оптического сигнала, достаточно хорошо отражающая реальные условия видеоизмерений. Фоновое изображение описывается медленно меняющимся случайным полем яркости. Обнаруживаемые объекты представляют собой двумерные области известной формы с функцией яркости, известной с точностью до контраста (коэффициента и аддитивной константы). Изображение, состоящее из объектов и фона, преобразуется линейной искажающей системой, зашумляется аддитивным стационарным шумом и подвергается равномерной дискретизации в прямоугольной системе координат.

Цель цифровой обработки изображения, содержащего объекты, состоит в их выявлении (обнаружении), локализации, т. е. указании положения на изображении, а также, возможно, оценке характеристик объектов, в частности, параметров их геометрических деформаций относительно заданных эталонов [1, 2]. Для достижения этой цели предлагается информационная технология обработки изображения, показанная на схеме. Она включает в себя несколько этапов преобразования данных.

Предварительная обработка входных изображений осуществляется, во-первых, с целью их яркостной нормализации (устранения постоянной составляющей оптического сигнала и колебаний контраста объектов) и, во-вторых, для повышения качества изображений путем фильтрации шумов и компенсации линейных (пространственно-частотных) искажений. На втором этапе для каждого положения скользящего окна обработки вычисляется некоторый вектор признаков поля яркости в пределах окна и, возможно, характеристики содержащегося в окне объекта. Далее вектор признаков преобразуется в скалярное значение дискриминантной функции, имеющей максимумы в точках наиболее правдоподобного местонахождения объектов. Следующие два этапа обработки — выделение указанных локальных максимумов и их пороговый отбор (отсевание малых значений, имеющих шумовое происхождение). После этого, на последнем шаге преобразования становится возможным составление списка координат и характеристик обнаруженных объектов, т. е. полное решение поставленной задачи.

Очевидно, что приведенная схема требует параметрической настройки. В общем случае для этой настройки предлагается использовать двухэтапную процедуру обучения по набору синтезированных тестовых изображений, содержащих объекты с известными характеристиками и искаженными в соответствии с действующей моделью наблюдения оптического сигнала. На первом этапе обучения определяются параметры дискриминантных функций для каждого класса объектов. На втором этапе при фиксированных параметрах дискриминантных функций ищется порог, обеспечивающий отбор локальных максимумов с наименьшей вероятностью ошибок распознавания.



Информационная технология обнаружения объектов на изображении.

Следует отметить важный частный случай, в котором происходит существенное упрощение как технологии обработки изображения, так и процедуры её настройки. При линейном формировании признаков и дискриминантных функций можно определять параметры последних не в результате обучения, а расчетным путем. Данную ситуацию можно проиллюстрировать примером использования линейного коррелятора, в котором признаками служат сами отсчеты изображения в пределах окна обработки, а в качестве коэффициентов линейной дискриминантной функции берутся значения яркости эталонного объекта [2, 6].

Ниже дается более детальное описание возможных вариантов реализации этапов обработки данных в рамках предлагаемой информационной технологии, а также рассматриваются вопросы расчета параметров основных алгоритмов обработки или их настройки через процедуры обучения.

### 3. Предварительная обработка изображения

Условия наблюдения и формирования изображений сильно влияют на яркостные характеристики объектов и фона. Для обеспечения инвариантности к яркостным изменениям изображения могут применяться методы его нормализации, т. е. выравнивания контраста между объектами и фоном [2, 6]. В рамках рассматриваемой технологии обработки изображений предлагается следующий алгоритм, обеспечивающий локальную нормализацию в пределах скользящего окна обработки. Для устранения медленно меняющейся (фоновой) составляющей поля яркости применяется локальное центрирование, при котором из каждого отсчета изображения  $x(n_1, n_2)$  вычитается локальное среднее  $\bar{E}(n_1, n_2)$  для соответствующего положения окна

$$y(n_1, n_2) = x(n_1, n_2) - \bar{E}(n_1, n_2). \quad (1)$$

Величина контраста между объектом и фоном нормализуется путем деления каждого центрированного отсчета  $y(n_1, n_2)$  на выборочное среднеквадратичное отклонение поля яркости в окне обработки

$$Z(n_1, n_2) = \frac{y(n_1, n_2)}{\sqrt{D}(n_1, n_2)}. \quad (2)$$

Используемые в формулах (1) и (2) величины локальных среднего и дисперсии могут быть вычислены с помощью системы рекуррентных соотношений

$$e(n_1, n_2) = e(n_1 - 1, n_2) + X(n_1 + N_1, n_2) - X(n_1 - N_1, n_2),$$

$$E(n_1, n_2) = E(n_1, n_2 - 1) + e(n_1, n_2 + N_2) - e(n_1, n_2 - N_2 - 1),$$

$$\bar{E}(n_1, n_2) = \frac{E(n_1, n_2)}{(2N_1 + 1)(2N_2 + 1)},$$

$$d(n_1, n_2) = d(n_1 - 1, n_2) + x^2(n_1 + N_1, n_2) - x^2(n_1 - N_1, n_2),$$

$$D(n_1, n_2) = d(n_1, n_2 - 1) + d(n_1, n_2 + N_2) - d(n_1, n_2 - N_2 - 1),$$

$$\bar{D}(n_1, n_2) = \frac{D(n_1, n_2)}{(2N_1 + 1)(2N_2 + 1)} - \bar{E}^2(n_1, n_2).$$

В приведенных соотношениях в качестве окна рассматривается симметричная (относительно отсчета обработанного изображения) прямоугольная область размером  $(2N_1 + 1) \times (2N_2 + 1)$  отсчетов. Данный алгоритм позволяет вычислять указанные локальные характеристики за один проход изображения. При этом вычислительная сложность (количество операций на один отсчет выходного изображения) не зависит от размеров скользящего окна. Для вычисления  $\bar{E}$  и  $\bar{D}$  требуется  $U_1 = 9$  операций сложения – вычитания и  $U_2 = 4$  операций умножения. (Здесь и далее будем пренебрегать операциями сложения – вычитания в индексах членов разностных уравнений, поскольку последовательное изменение индексов в процессе рекурсии реализуется существенно проще "полноценных" двухместных арифметических операций над произвольными числами).

Повышение качества изображения достигается с помощью фильтрации шумов наблюдения и компенсации частотно-пространственных искажений сигнала. При этом могут применяться как линейные, так и нелинейные алгоритмы обработки изображений. Линейная обработка заключается в осуществлении операции свертки входного изображения с импульсной характеристикой цифрового фильтра. Данная операция эффективно реализуется в рамках технологии скользящего окна с помощью параллельно-рекурсивных КИХ-фильтров (см. ниже). Однако возможность улучшения качества линейной обработки, как правило, ограничена, поскольку требование фильтрации шумов (ослабление высокочастотных компонент сигнала) и компенсации искажений (усиление высокочастотных компонент) вступают между собой в объективное противоречие.

Компромисс в разрешении указанного противоречия достигается применением нелинейной и, в частности, ранговой фильтрации [6, 9]. Ранговые фильтры (медианный, сигма-фильтр и др.) формируют выходной отсчет по вариационному ряду, составленному из элементов окна. Они в большей степени, чем линейные, учитывают локальные свойства изображения, что может улучшить качество его обработки. Недостатком нелинейных фильтров является сложность реализации в режиме скользящего окна: несмотря на то, что построение вариационного ряда допускает рекурсивную реализацию [6], окончательный расчет выходного отсчета в большинстве случаев необходимо выполнять независимо для каждого положения окна. Кроме того, параметры этих фильтров обычно не поддаются аналитическому расчету и должны подбираться по результатам трудоемких вычислительных экспериментов.

#### 4. Вычисление признаков объектов

В основу многих процедур вычисления признаков двумерных объектов в скользящем окне может быть положена операция линейной корреляционной обработки, которая заключается в преобразовании дискретизированного двумерного сигнала линейной инвариантной к сдвигу системой с конечной импуль-

свой характеристикой — КИХ-фильтром (см. [2, 6, 10]). Как известно, значения сигнала на выходе КИХ-фильтра являются результатом цифровой свертки входного сигнала с импульсной характеристикой фильтра и могут быть найдены взвешенным суммированием входных отсчетов в пределах окна обработки. Однако такое "прямое" вычисление свертки имеет практический смысл лишь при малых размерах окна, т.е. при короткой импульсной характеристике, поскольку объем вычислений здесь пропорционален числу ненулевых отсчетов последней. Не всегда решает проблему вычислительной сложности и применение известных алгоритмов быстрой свертки для КИХ-фильтров общего вида [6, 11]. Радикальное (на порядок и более) сокращение объема вычислений может быть обеспечено только применением алгоритмов параллельно-рекурсивной фильтрации изображений [6— 8].

Рассмотрим сначала более простой случай обработки одномерных сигналов. Преобразование КИХ-фильтром бесконечной последовательности отсчетов входного сигнала  $x(n)$  в выходную последовательность  $y(n)$  описывается сверткой

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)x(n-m) = \sum_{m=-N}^N h(m)x(n-m),$$

где  $h(m)$  — импульсная характеристика фильтра, равная нулю вне интервала  $[-N, N]$  (для простоты обозначений будем считать окно обработки симметричным). Параллельно-рекурсивный КИХ-фильтр представляется в виде  $K$  параллельных звеньев (при  $K \ll N$ ) и, следовательно,

$$h(m) = \sum_{k=0}^{K-1} a_k h_k(m), \quad (3)$$

$$y(n) = \sum_{k=0}^{K-1} a_k y_k(n), \quad (4)$$

$$y_k(n) = \sum_{m=-N}^N h_k(m)x(n-m). \quad (5)$$

Здесь  $a_k$  — коэффициенты,  $h_k(m)$  — линейно независимые базисные функции разложения  $h(m)$  в ряд (3), т.е. импульсные характеристики параллельных КИХ-звеньев,  $y_k(n)$  — сигналы на выходах звеньев ( $0 \leq k \leq K-1$ ). При этом каждое звено допускает эффективную рекурсивную реализацию, т.е. описывается достаточно простым разностным уравнением. Последнее означает, что передаточная функция ( $Z$ -преобразование импульсной характеристики) звена

$$H_k(z) = \sum_{m=-\infty}^{\infty} h_k(m)z^{-m} \quad (6)$$

записывается в дробно-рациональной форме, как отношение полиномов от комплексной переменной  $z$ , состоящих из небольшого числа слагаемых [12].

В работе [8] определен общий вид импульсных характеристик одномерных рекурсивно реализуемых КИХ-звеньев

$$h_k(m) = \sum_{l=0}^{L_k-1} S_{lk} m^{\alpha_{lk}} \beta_{lk}^{-m} [u(m - \gamma_{lk}) - u(m - \delta_{lk})], \quad (7)$$

$$\text{где } u(m) = \begin{cases} 1 & \text{при } m \geq 0, \\ 0 & \text{иначе} \end{cases}$$

— функция единичного скачка  $L_k, \alpha_{lk}, \gamma_{lk}, \delta_{lk}$  — целочисленные ( $L_k \geq 1, \alpha_{lk} \geq 0, -N \leq \gamma_{lk} \leq \delta_{lk} \leq N$ ), а  $S_{lk}, \beta_{lk}$  — произвольные вещественные или комплексные константы, а также их передаточных функций

$$H_k(z) = \sum_{l=0}^{L_k-1} S_{lk} (-1)^{\alpha_{lk}} \times \underbrace{\left\{ z \frac{d}{dz} \left[ z \frac{d}{dz} \left[ \dots z \frac{d}{dz} \left[ \frac{(\beta_{lk} z)^{-\gamma_{lk}} - \beta_{lk} z^{-\delta_{lk}}}{1 - (\beta_{lk} z)^{-1}} \right] \dots \right] \right\}}_{\alpha_{lk} \text{ раз}} \quad (8)$$

Формулы (7) и (8) могут использоваться для конструирования различных базисов разложения (3) и получения передаточных функций (и далее — разностных уравнений) для рекурсивных звеньев КИХ-фильтров. Так, при надлежащих значениях параметров из (7) следует базис комплексных дискретных экспоненциальных функций [6], базис Фурье в вещественной форме и косинусный базис [13], базис Хартли [14] и т. д. Выбор конкретных параметров в (7) и (8) следует производить из соображений простоты получающихся фильтров (так, очевидно, что  $L_k$  и  $\alpha_{lk}$  должны быть невелики) и достаточности малого числа слагаемых в разложении (3) для качественного решения задачи обнаружения объектов. Рассмотрим примеры.

Как показывают эксперименты, хороший эффект обработки при низкой вычислительной сложности обеспечивается применением в разложении (3) простейшего прямоугольного базиса [6, 7, 15]. Одномерный прямоугольный базис состоит из функций

$$h_k(m) = u(m - \gamma_k) - u(m - \delta_k) \quad (9)$$

и соответствует случаю, когда в (7)  $L_k = S_{0k} = \beta_{0k} = 1, \alpha_{0k} = 0, \gamma_{0k} = \gamma_k, \delta_{0k} = \delta_k$ . Подстановка (9) в (6) дает выражение для дробно-рациональной передаточной функции параллельного звена

$$H_k(z) = \frac{z^{-\gamma_k} - z^{-\delta_k}}{1 - z^{-1}},$$



структура которого непосредственно определяет вид разностного уравнения, описывающего звено [12]. Совершив соответствующий переход, для сигнала (5) на выходе звена получаем

$$y_k(n) = y_k(n-1) + x(n-\gamma_k) - x(n-\delta_k). \quad (10)$$

Как следует из этой формулы, для формирования одного отсчета  $y_k(n)$  нужно выполнить всего две арифметические операции: сложение и вычитание.

В качестве другого примера разложения (3) можно рассматривать разложение по степенным функциям (моментам) [1, 16]. Из общего представления (7) базисной функции, допускающей параллельно-рекурсивную реализацию, при  $L_k = S_{0k} = \beta_{0k} = 1$ ,  $\alpha_{0k} = k$ ,  $\gamma_{0k} = -N$ ,  $\delta_{0k} = N+1$  получаем одномерную базисную функцию вида

$$h_k(m) = m^k [u(m+N) - u(m-N-1)]. \quad (11)$$

По общей формуле (8) можно построить передаточные функции и разностные уравнения для одномерных линейных систем с импульсными характеристиками вида (11). Однако для данного конкретного случая можно предложить и более эффективный алгоритм рекурсивного вычисления всей совокупности величин (5) (локальных моментных характеристик сигнала) при  $0 \leq k < K$  [17]

$$y_k(n) = \sum_{i=0}^k C_k^i y_i(n-1) + (-N)^k x(n+N) - (N+1)^k x(n-N-1), \quad (12)$$

где  $C_k^i$  — биномиальные коэффициенты.

Полученные результаты могут быть обобщены на двумерный случай. Для двумерного сигнала (изображения), заданного отсчетами  $x(n_1, n_2)$  на квадратном растре, результат обработки КИХ-фильтром выражается через двумерную свертку

$$y(n_1, n_2) = \sum_{(m_1, m_2) \in D} h(m_1, m_2) x(n_1 - m_1, n_2 - m_2), \quad (13)$$

где  $h(m_1, m_2)$  — импульсная характеристика двумерного фильтра,  $D$  — конечная область ее ненулевых значений. Для быстрого параллельно-рекурсивного вычисления свертки (13) нужно, чтобы импульсная характеристика фильтра записывалась в виде суммы

$$h(m_1, m_2) = \sum_{k=0}^{K-1} a_k h_k(m_1, m_2) \quad (14)$$

при малых  $K$ , а базисные функции этого разложения соответствовали импульсным характеристикам рекурсивных КИХ-звеньев. Если ограничиться рассмотрением прямоугольных окон обработки и разделимых звеньев [6, 10]

$$h_k(m_1, m_2) = h_k^{(1)}(m_1) h_k^{(2)}(m_2), \quad (15)$$

то для каждого сомножителя в (15) можно воспользоваться "одномерной" общей формулой (7). Реализация разделимого звена заключается в последовательной (каскадной) обработке двумерного сигнала по координатам.

Опираясь на соотношения (10), (13) — (15), несложно получить, что двумерный параллельно-рекурсивный КИХ-фильтр, построенный на прямоугольном базисе и состоящий из  $K$  звеньев, требует выполнения  $U_1 = 5K - 1$  операций сложения — вычитания и  $U_2 = K$  операций умножения на каждый отсчет обрабатываемого изображения. Из формул (12) — (15) следует, что для вычисления локальных двумерных моментов до порядка  $(K_1, K_2)$  включительно требуется  $U_1$  операций сложения и  $U_2$  операций умножения на каждый отсчет обрабатываемого изображения, где

$$U_1 = 0,5 [(K_1 + 1)(K_1 + 4) + K_1(K_2 + 1)(K_2 + 4)],$$

$$U_2 = 0,5 [K_1(K_1 + 3) + K_1 K_2(K_2 + 3)],$$

а при построении КИХ-фильтра, использующего двумерный степенной базис, с учетом затрат на формирование выходного сигнала —

$$U_1 = 0,5 [(K_1 + 1)(K_1 + 4) + K_1(K_2 + 1)(K_2 + 4)] + K_1 K_2 - 1,$$

$$U_2 = 0,5 [K_1(K_1 + 3) + K_1 K_2(K_2 + 3)] + K_1 K_2.$$

Очевидно, что сложность параллельно-рекурсивных КИХ-фильтров определяется числом параллельных звеньев и не зависит от размеров окна обработки сигнала. Для сравнения, обычный КИХ-фильтр в форме прямой свертки имеет сложность, примерно пропорциональную числу отсчетов в окне. Результаты исследования показывают, что во многих практических ситуациях применение двумерных параллельно-рекурсивных КИХ-фильтров обеспечивает радикальное (на порядок и более) снижение сложности обработки изображений по сравнению с использованием КИХ-фильтров, реализуемых прямой сверткой [7, 15].

Отметим далее, что при проектировании параллельно-рекурсивного КИХ-фильтра необходимо решить три задачи: выбрать размеры окна обработки и класс базисных функций разложения (3) или (14), из полного множества базисных функций выбранного класса выделить фактически используемые в разложении  $K$  функций, рассчитать коэффициенты фильтра. Первая задача может быть решена эвристически. Вторая и третья в общем случае решаются одновременно, в ходе переборной процедуры численных расчетов.

В работах [7, 18] для быстрого выбора подмножества базисных функций предлагается использовать известный субоптимальный метод последовательного присоединения [19]. В соответствии с ним сначала выбирается единственная функция, обеспечивающая максимум некоторого показателя качества  $R$ , затем к ней присоединяется еще одна, максимизирующая показатель в паре с уже выбранной, и т.д. до получения набора из  $K$  функций. Данный метод резко сокращает вычислительные затраты по сравнению с полным перебором

при несущественной потере оптимальности формируемого подмножества базисных функций.

Как следует из сказанного выше, для каждого анализируемого подмножества базисных функций нужно рассчитывать показатель качества обработки сигналов, а для окончательного варианта подмножества — и коэффициенты разложения импульсной характеристики фильтра в ряд (3) или (14). В [7, 18] установлено, что для многих задач обработки сигналов общая схема этих расчетов оказывается одной и той же. Вектор-столбец искомым коэффициентов  $A = \{a_k\}_{k=0}^{K-1}$  всегда задается матричным соотношением вида

$$A = B^{-1} C, \quad (16)$$

а показатель качества, максимизируемый в процессе выбора базисных функций, соотношением

$$R = A^t C = C^t B^{-1} C, \quad (17)$$

где  $B = \{b_{lk}\}_{l,k=0}^{K-1}$  — невырожденная симметрическая матрица, верхний индекс  $-1$  означает обращение матрицы,  $C = \{c_k\}_{k=0}^{K-1}$  — вектор-столбец, верхний индекс  $t$  — транспонирование вектора. Специфика расчета фильтра для конкретной задачи заключается только в способе вычисления элементов  $B$  и  $C$ .

Так, пусть при линейной фильтрации сигнала на этапе его предварительной обработки (см. п. 3) искажения соответствуют следующей модели наблюдения (для краткости изложения здесь и ниже ограничимся рассмотрением одномерного случая)

$$x(n) = \sum_{m=-\infty}^{\infty} f(m) x_0(n-m) + v(n); \quad (18)$$

где  $f(m)$  — импульсная характеристика "искажающей" линейной системы,  $x_0(n)$  и  $v(n)$  — соответственно полезный сигнал и помеха, некоррелированные между собой стационарные случайные последовательности с нулевыми средними и автоковариационными функциями (АКФ)  $B_{x_0}(m)$  и  $B_v(m)$ . При расчете параллельно-рекурсивного КИХ-фильтра из условия минимальной среднеквадратичной ошибки восстановления полезного сигнала имеем [18]

$$b_{lk} = \sum_{m=-N}^N \sum_{n=-N}^N h_l(m) h_k(n) B_x(m-n), \quad c_k = \sum_{m=-N}^N h_k(m) B_{x_0 x}(-m), \quad (19)$$

где  $B_x(m)$  — АКФ наблюдаемого входного сигнала,  $B_{x_0 x}(m)$  — взаимная корреляционная функция полезного и наблюдаемого сигналов. Эти функции выражаются через введенные выше характеристики модели наблюдения (18) [12]

$$B_x(m) = \sum_{p=-\infty}^{\infty} \sum_{r=-\infty}^{\infty} f(r)f(r+p)B_{x_0}(m-p) + B_v(m),$$

$$B_{x_0 x}(m) = \sum_{p=-\infty}^{\infty} f(p)B_{x_0}(m-p).$$

Коэффициенты фильтра здесь определяются соотношением (16) с учетом (19), а достигаемая минимальная дисперсия ошибки восстановления имеет вид

$$\varepsilon_{\min}^2 = D_{x_0} - R,$$

где  $D_{x_0}$  — дисперсия полезного сигнала,  $R$  — параметр, рассчитываемый по формуле (17).

Обратимся теперь к нашей основной цели обработки — обнаружению объектов. Пусть сигнал на входе фильтра состоит либо из аддитивной смеси "объекта"  $T(n)$  известной формы и "фона" — случайной последовательности  $v(n)$

$$x(n) = T(n) + v(n), \quad (20)$$

либо только из фона

$$x(n) = v(n), \quad (21)$$

и задача состоит в том, чтобы отличить одну ситуацию от другой. Будем судить о наличии объекта по уровню сигнала  $y(n)$  на выходе фильтра, т. е. считать, что наблюдаемый входной сигнал соответствует модели (20), если

$$y(n) > \Delta, \quad (22)$$

где  $\Delta$  — некоторое пороговое значение, и модели (21) в противном случае. Данная задача интерпретируется как задача классификации входного сигнала, решаемая для каждого значения  $n$ . В рамках принятой технологии обработки изображения (см. п. 1) в качестве классификационных признаков будем использовать сигналы (5) на выходах параллельных звеньев фильтра. Совокупность этих сигналов образует вектор признаков  $Y = \{y_k(n)\}_{k=0}^{K-1}$ . Если считать, что фон  $v(n)$  стационарен, распределен по нормальному закону, имеет нулевое среднее и АКФ  $B_v(m)$ , то несложно получить, что по критерию минимума ошибок оптимальным является линейный классификатор, принимающий решение о наличии объекта при

$$C^t B^{-1} Y > 0,5 C^t B^{-1} C - d, \quad (23)$$

где  $d$  — параметр, зависящий от выбранного критерия обнаружения и априорной вероятности появления объекта в наблюдаемом сигнале [20, 21]. Сопоставление неравенств (22) и (23) с учетом формулы (4) позволяет одновременно найти вектор коэффициентов фильтра (он снова будет определяться соотношением (16)) и пороговое значение выходного сигнала

$$\Delta = 0,5 R - d.$$

Здесь величина  $R$  опять задается формулой (17) и представляет собой расстояние Махаланобиса между классами сигналов [20].

Рассмотренная выше ситуация соответствует отмеченному в п. 1 частному случаю технологии обнаружения, в котором параметры линейной дискриминантной функции рассчитываются аналитически. При модели наблюдения сигнала, более сложной, нежели (20), (21), а также при нелинейном формировании признаков (например, при конструировании из локальных моментных функций признаков, инвариантных к геометрическим деформациям объектов [16]) для определения параметров дискриминантной функции необходимо использовать описываемый ниже общий алгоритм обучения.

В заключение отметим, что нет четкой границы между упоминаемыми в п. 1 признаками и характеристиками объектов. В их качестве могут рассматриваться любые (в том числе одни и те же) количественные параметры поля яркости, вычисляемые в скользящем окне. Различие состоит только в том, что признаки используются при построении дискриминантной функции для обнаружения объектов, а характеристики просто описывают свойства объектов, интересующие получателя обработанной видеоинформации (см. схему).

## 5. Алгоритм обучения

### для определения параметров дискриминантной функции

Согласно введенной в п. 1 информационной технологии, признаки изображения, сформированные в скользящем окне обработки, далее используются для вычисления скалярной дискриминантной функции. Задачу определения параметров этой функции можно рассматривать как задачу построения классификатора на два класса. Математически данная задача формулируется следующим образом [20, 21].

Пусть при каждом положении окна измеряется  $K$  признаков, образующих вектор  $Y = (y_0, \dots, y_{K-1})$ . Обозначим  $\omega_1$  и  $\omega_2$  классы значений векторов, соответствующие наличию и отсутствию обнаруживаемого объекта в окне обработки. В этом случае пространство признаков состоит из двух областей, каждая из которых содержит точки своего класса. Граница областей определяется скалярной и однозначной функцией вектора признаков  $D(Y)$ , называемой дискриминантной. Классификатор или решающее правило, ставящее в соответствие каждому вектору признаков  $Y$  определенный номер класса, можно записать в виде

$$D(Y) \begin{cases} > 0 \\ < 0 \end{cases} \Rightarrow Y \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}. \quad (24)$$

Исчерпывающими характеристиками качества классификации являются вероятности ошибок (в нашем случае — вероятности пропуска объекта и ложного обнаружения)

$$P_1 = P(Y \in \omega_2 / Y \in \omega_1), \quad P_2 = P(Y \in \omega_1 / Y \in \omega_2). \quad (25)$$

На практике часто ограничиваются рассмотрением только линейных дискриминантных функций. Одна из причин этого заключается в том, что при высокой размерности векторов признаков только линейные или кусочно-линейные классификаторы дают приемлемый компромисс между качеством и вычислительной сложностью классификации. Другая важная причина состоит в том, что многие нелинейные дискриминантные функции могут быть интерпретированы как линейные в определенном функциональном пространстве [21]. С учетом этого ниже будем полагать, что классификация производится с помощью линейной дискриминантной функции

$$D(Y) = \sum_{k=0}^{K-1} w_k y_k + w_K, \quad (26)$$

где  $w_0, \dots, w_K$  — параметры (веса) линейного классификатора. Для удобства анализа принято вводить в вектор признаков дополнительную (фиктивную) компоненту  $y_K = 1$  и представлять соотношение (26) в векторной форме

$$D(Y) = W^t Y. \quad (27)$$

Здесь  $Y = (y_0, \dots, y_{K-1}, y_K)$  и  $W = (w_0, \dots, w_{K-1}, w_K)$  — пополненные векторы признаков и весов дискриминантной функции соответственно.

Определение компонент весового вектора в (27) может осуществляться в результате обучения классификатора [21]. Задача обучения ставится здесь следующим образом. Пусть априорная информация об анализируемом сигнале позволяет получить "обучающую выборку", т. е. множество векторов признаков, для каждого из которых известно, какому из классов  $\omega_1$  и  $\omega_2$  он принадлежит. Основываясь на обучающей выборке, требуется построить в признаковом пространстве гиперплоскость, разделяющую пространство на области, соответствующие классам. Данное разделение должно отвечать некоторому естественному принципу оптимальности, например, минимизации вероятностей (25).

В настоящее время известно большое число эффективных алгоритмов обучения классификаторов. Значительная их часть базируется на применении метода стохастической аппроксимации [20, 21]. Основное преимущество этого метода состоит в том, что он позволяет реализовать последовательное обучение в темпе поступления очередных элементов обучающей выборки. Тем самым алгоритмы стохастической аппроксимации оказываются наиболее приспособленными к решению задачи определения параметров дискриминантной функции в условиях обработки изображения скользящим окном. Ниже мы рассмотрим некоторые алгоритмы последовательного обучения линейных классификаторов.

Обучающая выборка содержит информацию о принадлежности вектора  $Y$  классам  $\omega_1$  и  $\omega_2$  и позволяет определить знаковую случайную переменную правильной классификации

$$r(Y) = \begin{cases} 1 & \text{при } Y \in \omega_1, \\ -1 & \text{при } Y \in \omega_2. \end{cases} \quad (28)$$

С другой стороны, по правилу (24), дискриминантная функция  $D(Y)$  меняет знак синхронно с переменной (28). Поэтому естественно стремиться подобрать вектор коэффициентов  $W$  в (27) так, чтобы он минимизировал уклонение функции  $D(Y)$  от  $r(Y)$ . В качестве меры уклонения можно взять математическое ожидание некоторой выпуклой функции  $F$ , зависящей от разности сравниваемых величин

$$J(W, Y) = M \{F [D(Y) - r(Y)]\} = M \{F [W^t Y - r(Y)]\}. \quad (29)$$

Очевидно, что величина  $J(W, Y)$  достигает минимума (равного нулю) при безошибочной классификации всех объектов. Обычно полагается  $F(x - y) = (x - y)^2$ , что соответствует построению классификатора по условию минимума среднеквадратичной ошибки приближения величины (28) дискриминантной функцией (27). Далее мы ограничимся именно таким вариантом показателя (29)

$$J(W, Y) = M \{[W^t Y - r(Y)]^2\}. \quad (30)$$

Для определения минимума величины (30) необходимо найти корень уравнения

$$\frac{\partial J(W, Y)}{\partial W} = -2 M \{Y [W^t Y - r(Y)]\} = 0$$

или

$$M \{Y [W^t Y - r(Y)]\} = 0. \quad (31)$$

Универсальный алгоритм поиска корня уравнения (31) дает процедура Роббинса—Монро [20, 21]. В соответствии с ней вектор коэффициентов  $W$  находится итеративно

$$W(m+1) = W(m) + \alpha_m Y(m) [r(Y(m)) - W^t(m) Y(m)], \quad (32)$$

где  $m$  — шаг итерации, начальное значение  $W(1)$  выбирается произвольным образом,  $\{\alpha_m\}$  — последовательность положительных чисел,  $r(Y(m))$  — переменная правильной классификации  $m$ -го вектора обучающей выборки  $Y(m)$ . Алгоритм (32) на каждом шаге итераций корректирует значение вектора весов  $W$  на величину, зависящую от  $\alpha_m$ . При этом для сходимости алгоритма к решению, минимизирующему показатель (30), необходимо, чтобы последовательность  $\{\alpha_m\}$  удовлетворяла следующим условиям:

$$\lim_{m \rightarrow \infty} \alpha_m = 0, \quad \sum_{m=1}^{\infty} \alpha_m = \infty, \quad \sum_{m=1}^{\infty} \alpha_m^2 < \infty,$$

(так, можно положить  $\alpha_m = 1/m$ ). Сходимость алгоритма (32) понимается в том смысле, что при  $P(|W(m) - \hat{W}| \geq \epsilon) \rightarrow 0$  при  $m \rightarrow \infty$ , где  $\hat{W}$  — значение вектора весов, минимизирующего критерий (30).

Алгоритм (32), также, как и другие подобные схемы стохастической аппроксимации, обычно обнаруживает очень медленную скорость сходимости и, как следствие, требует очень большого объема обучающей выборки. Для ускорения сходимости этого алгоритма можно предложить следующие два приема. Первый из них состоит в сохранении постоянного значения  $\alpha_m$  на тех шагах, где значения разности  $r(Y) - W^t Y$  имеют одинаковый знак [20]. Данный прием основывается на том факте, что обычно изменение знака этой разности происходит в окрестности корня уравнения (31) и обеспечивает значительные коррекции в точках, удаленных от корня.

Однако, как показывают эксперименты, сам по себе первый прием не дает существенного ускорения сходимости. Увеличить эффективность от его использования можно, если изменение  $\alpha_m$  производить только на тех шагах, на которых имеет место ошибочная классификация вектора  $Y(m)$  с помощью весового вектора  $W(m)$ . Иначе говоря, предлагается полагать  $\alpha_m = 0$  тогда, когда вектор  $W(m)$  обеспечивает правильную классификацию. Совместное использование описанных приемов дает приемлемую скорость сходимости алгоритма обучения (32) и, кроме того, повышает его вычислительную эффективность.

Применительно к задаче цифровой обработки изображений в режиме скользящего окна, сформулированной в п. 1, построение классификатора является этапом, предшествующим определению положения объектов на изображении. Результатом этого этапа (первого этапа обучения) являются параметры дискриминантной функции. Следующий шаг решения поставленной задачи связан с выделением и пороговой обработкой локальных максимумов поля значений полученной дискриминантной функции. Для определения порога необходим второй этап обучения, рассматриваемый ниже.

## 6. Локализация объектов на изображении

Применение рассматриваемой технологии обработки изображения для обнаружения объектов имеет следующую особенность. При обучении классификатора (см. п. 5) векторы признаков, вычисляемые для каждого положения скользящего окна, считаются принадлежащими классу  $\omega_1$ , соответствующему наличию объекта, в случае его полного "захвата" окном и классу  $\omega_2$  — если объект отсутствует или только частично находится в окне. При этом решение о наличии объекта по правилу (24) может быть принято для нескольких смежных положений окна обработки, в связи с чем появляется необходимость выделения для каждого объекта координат единственной точки, определяющей его наиболее вероятное положение. Естественно считать, что эта точка соответствует максимальному из значений дискриминантной функции, полученных в соседних положениях окна.

Для определения локальных максимумов предлагается использовать алгоритм обработки скалярного поля значений дискриминантной функции, названный "пиковым фильтром" [22]. Выходное поле пикового фильтра формируется следующим образом:

$$y(n_1, n_2) = \begin{cases} x(n_1, n_2) & \text{при } r_0 = N, \\ 0 & \text{при } r_0 < N, \end{cases}$$



где  $x(n_1, n_2)$  — поле на входе фильтра;  $r_0$  — ранг значения  $x(n_1, n_2)$  в вариационном ряду отсчетов входного поля в окне обработки с центром в точке  $(n_1, n_2)$ ;  $N$  — число отсчетов в окне. Для прямоугольного окна возможна параллельно-рекурсивная реализация пикового фильтра, при которой вычисляются максимумы независимо по каждому столбцу скользящего окна и далее из них определяется максимальный элемент

$$\bar{x}(n_1, n_2) = \max_{-N_1 \leq m_1 \leq N_1} x(n_1 + m_1, n_2),$$

$$y(n_1, n_2) = \max_{-N_2 \leq m_2 \leq N_2} \bar{x}(n_2, n_2 + m_2).$$

В результате пиковой фильтрации ненулевые значения сохраняются только в точках поля, соответствующих локальным максимумам и расположенных друг от друга на расстоянии, превышающем размер окна обработки. Далее они используются для окончательного решения о наличии и координатах объектов. Последний считается обнаруженным в точке с координатами  $(n_1, n_2)$ , если  $y(n_1, n_2) > R$ , где  $R$  — некоторое пороговое значение.

Порог  $R$  определяется при параметрической настройке рассматриваемой информационной технологии на втором этапе обучения. Обучающая выборка на данном этапе строится по серии тестовых изображений объектов и включает только ненулевые значения  $y(n_1, n_2)$  локальных максимумов дискриминантной функции (в отличие от первого этапа, где в качестве обучающей выборки рассматривалось множество векторов признаков, соответствующих всем положениям скользящего окна на изображении). Эта выборка содержит как элементы, соответствующие обнаруживаемым объектам (истинные цели), так и элементы, не соответствующие объектам на изображении (ложные цели). Множество элементов выборки разбивается на два подмножества (истинных и ложных целей). Для каждого подмножества строятся гистограммы, по которым легко определяется значение  $R$ , оптимальное по любому из общепринятых критериев качества процедуры обнаружения объектов [21].

## ЛИТЕРАТУРА

1. Анисимов Б. В., Курганов В. Д., Злобин В. К. Распознавание и цифровая обработка изображений. М., 1983.
2. Прэтт У. Цифровая обработка изображений: Кн. 2. М., 1982.
3. Распознавание образов при помощи цифровых вычислительных машин/ Под ред. Л. Хармана. М., 1974.
4. Баклицкий В. К., Юрьев А. Н. Корреляционно-экстремальные методы навигации. М., 1982.
5. Буймов А. Г. Корреляционно-экстремальная обработка изображений. Томск, 1987.
6. Ярославский Л. П. Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. М., 1987.
7. Сергеев В. В. // Радиотехника. 1990. N 8. С. 38—41.
8. Сергеев В. В. // Компьютерная оптика. М., 1992. Вып. 10—11. С. 186—201.

9. Чочиа П. А. // Кодирование и обработка изображений. М., 1988. С.87—98.
10. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. М., 1988.
11. Вайрадян А. С., Пчелинцев И. П., Чельшев М. М. // Зарубежная радиоэлектроника. 1982. N 3. С. 3—34.
12. Оппенгейм А. В., Шафер Р. В. Цифровая обработка сигналов. М., 1979.
13. Виткус Р. Ю., Ярославский Л. П. // Адаптивные методы обработки изображений. М., 1982. С. 6—35.
14. Бондьонадхьяй П. К. // ТИИЭР. 1988. N 10. С. 126—128.
15. Мурызин С. А., Сергеев В. В., Фролова Л. Г. // Компьютерная оптика. М., 1992. Вып. 12. С. 171—203.
16. Майтра С. // ТИИЭР. 1979. N 4. С. 297—299.
17. Glutov N. I., Krainikov N. I., Sergeev V. V., Khramov A. G. // Pattern Recognition and Image Analysis. 1991. N 4. P. 424—426.
18. Сергеев В. В., Фролова Л. Г. // Компьютерная оптика. М., 1992. Вып. 12. С. 204—230.
19. Методы, критерии и алгоритмы, используемые при преобразовании, выделении и выборе признаков в анализе данных: Сборник статей АН ЛитССР/ Под редакцией Ш. Раудиса. Вильнюс, 1988.
20. Ту Дж., Гонсалес Р. Принципы распознавания образов. М., 1978.
21. Фукунага К. Введение в статистическую теорию распознавания образов. М., 1979.
22. Сергеев В. В., Чичева М. А. // Компьютерная оптика. М., 1989. Вып. 5. С. 78—84.

Статья поступила 25.01.93