

УДК 681.323 : 523.8

П. Г. Попов

(Новосибирск)

### ДИНАМИЧЕСКАЯ МЕРА БЛИЗОСТИ ИЗОБРАЖЕНИЙ. Ч. I. ВЫДЕЛЕНИЕ ДВИЖЕНИЯ

Рассматривается один из вариантов межкадровой обработки изображений в следящих системах. Вводится понятие динамической меры близости изображений, описанных локальными анизотропными признаками. Приведены экспериментальные результаты по селекции динамических объектов.

В последнее время проводится много исследований методов межкадровой обработки изображений с целью выделения подвижных объектов и устранения дефектов изображений, вызванных движением. Межкадровая обработка изображений в системах технического зрения позволяет не только улучшить характеристики этих систем, но и придать им дополнительные функциональные возможности [1, 10].

Типичная ситуация, где приходится сталкиваться с межкадровой обработкой подвижных изображений, — это телевизионные системы, и в частности следящие системы. Как правило, значения элементов изображения в соседних кадрах изображения (с фиксированными координатами) существенно отличаются друг от друга только в зонах относительного перемещения. Далее будем предполагать, что наблюдаемый объект всегда находится в поле зрения камеры (т. е. в пределах кадра), но может смещаться за счет прямолинейного движения, вращения и других видов движения.

Рассмотрим работу телевизионной следящей системы в случае, когда характер обрабатываемой видеоинформации и специфика решаемой задачи не позволяют оперативно селектировать наблюдаемый объект или использовать в качестве эталонов набор заранее выбранных изображений. В этом случае предполагается использование в системе априорной информации: поиск объекта производится оператором, после чего включается режим слежения. Эталоном при этом является фрагмент (строб) телевизионного кадра, определенный оператором в качестве объекта. Дальнейшая смена эталона производится автоматически при смене кадров: за эталон принимается фрагмент телевизионного изображения, имеющий наибольшее сходство с текущим эталоном. Общая структурная схема такой системы показана на рис. 1.

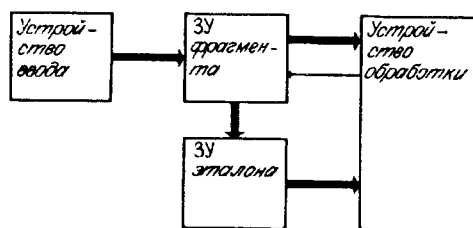


Рис. 1

Дальнейшее изложение построено на следующих предположениях:

1. Следящая система является чисто пассивной системой.

2. Алгоритмы обработки изображений не являются синтаксическими.

3. Объект, даже в случае сложных изображений, обладает несколькими признаками, отличающими его от фона. Предполагается, что изображение объекта обладает внутренней текстурой.

4. Класс трансформаций объекта не ограничен.

5. Эталонное изображение обновляется от кадра к кадру.

Для повышения надежности слежения в подобных системах предпринимались попытки (например, [1]) ввести определенные правила на перезапись элементов эталона. Для этого производилось сравнение элементов старого и нового эталонов. Такое поэлементное сравнение при обычном яркостном описании изображений представляет собой сложную задачу, если учесть, что значения яркости элементов искажены шумами. Для того чтобы сделать вывод об идентичности элементов, необходимо применение каких-либо пороговых критериев. Если при этом учесть, что в случае динамического объекта могут происходить яркостные изменения эталона и изменяться условия наблюдения, то такая задача становится практически неразрешимой при сложных форме объектов и фоне.

Такого рода трудности не возникают в случае, когда следящая система работает с изображениями, предварительно описанными локальными анизотропными признаками (ЛАП) [2, 11]. Смысл локальных анизотропных признаков заключается в следующем. Каждый элемент изображения описывается условным числом (кодом), соответствующим общей направленности его окрестности. Под направленностью окрестности понимается направление, перпендикулярное направлению, соответствующему наибольшей мощности фурье-спектра данной окрестности. Эти условные числа и выбираются в качестве признаков.

Окрестности всех точек одинаковы и для удобства вычислений чаще всего имеют форму квадрата. Для адекватного представления структуры изображения линейный размер окрестности не должен превышать величину интервала автокорреляции исходного изображения.

Для хорошего описания сложных изображений достаточно пяти классов признаков. Четыре из них соответствуют преобладающим корреляционным связям в одном из четырех направлений, составляющих углы  $0, 45, 90, 135^\circ$  с вертикалью. Пятый класс составляют окрестности с изотропной структурой, при которой имеются две пары направлений с одинаковыми по величине корреляционными связями: пара вертикаль — горизонталь и пара диагональных направлений. При этом корреляции различных пар могут отличаться по величине друг от друга. На практике класс окрестности определяется по наибольшей мощности элементов энергетического спектра окрестности, лежащих на выбранных осях. Разделение окрестностей на изотропные и анизотропные производится по расстоянию Бхаттачария для классов «изотропность — анизотропность». Подробно эти вопросы рассмотрены в [2].

Выбранные пять классов обозначаются условными числами, т. е. каждая точка изображения кодируется одним из пяти чисел. В частном случае, когда окрестности разделяются только на два класса (изотропные и анизотропные), исходное изображение преобразуется в бинарный контурный препарат. На рис. 2 приведено реальное изображение, а на рис. 3 то же изображение представлено в закодированном виде. Для наглядности на рис. 3 ЛАП представлены не в виде чисел, а в виде направлений. Изображения имели размер  $32 \times 32$  элемента, кодирование производилось по окрестностям размером  $4 \times 4$  элемента.

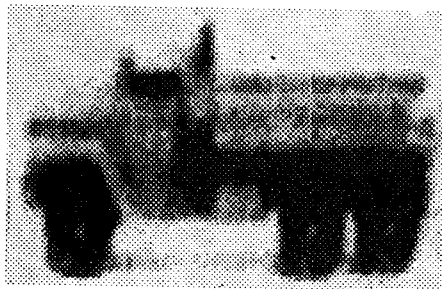


Рис. 2

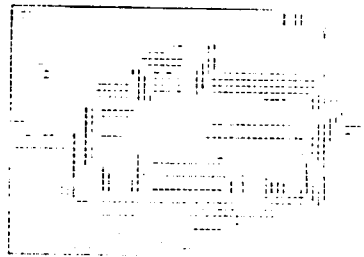


Рис. 3

При построении следящей системы для изображений, описанных ЛАП, вводятся специальные меры близости в пространстве признаков. Наиболее простой из предложенных [2] мерой близости является величина

$$S_1 = \sum_j \sum_k \delta [B_1(j, k) - B_2(j, k)],$$

где  $B_1(j, k)$  и  $B_2(j, k)$  — элементы сравниваемых изображений, а

$$\delta(x) = \begin{cases} 1, & \text{если } x = 0; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Иными словами,  $S_1$  дает количество совпадающих элементов изображений при их совмещении. При перемещении эталонного изображения по анализируемому фрагменту образуется матрица меры близости (аналог корреляционной матрицы), по максимальному значению которой осуществляется координатная привязка эталона.

Вторая мера близости имеет вид

$$S_2 = \sum_j \sum_k \delta_2 [B_1(j, k) - B_2(j, k)],$$

где

$$\delta_2(x) = \begin{cases} 1, & \text{если } x = 0; \\ \tau \bmod_2 |x| & \text{в остальных случаях.} \end{cases}$$

Эта мера близости учитывает не только совпадающие элементы, но и элементы с близкими направлениями. Вклад элементов с близкими направлениями определяется весовым коэффициентом  $\tau$ .

Если рассматривать изображение в расширенном пространстве признаков, т. е. описывать каждую точку признаком ее локальной окрестности и в то же время яркостью, то можно применять комбинированные меры близости изображений. Смысл таких мер близости заключается в том, что используются обычные меры близости при поточечном сравнении значений яркости изображений, но это сравнение допускается только в том случае, когда признаки сравниваемых точек совпадают. Например, комбинированная корреляционная мера близости будет иметь следующий вид:

$$S_k(m, n) = \sum_j \sum_k \delta_2 [B_1(j, k) - B_2(j, k)] b_1(j, k) b_2(j, k),$$

где  $b_1, b_2$  — яркости изображений. Для примера строки нормированных матриц мер близости  $S_1(a)$ ,  $S_2(b)$  и  $S_k(c)$  приведены на рис. 4 при привязке двух кадров изображения, показанного на рис. 2. Дальнейшее изложение будет основываться на применении меры близости  $S_1$ .

При сравнении элементов двух эталонных изображений не требуется применения пороговых методов, а речь идет лишь о совпадении элементов.

Следуя [3], используем модель изображения для случая одного объекта:

$$a(\bar{i}; kT) = f(\bar{i}; kT) \lambda(\bar{i}; kT) + [1 - \lambda(\bar{i}; kT)] b(\bar{i}; kT),$$

где 3-мерные функции определены как:  $a(\bar{i}; kT) \triangleq$  закодированное изображение;  $b(\bar{i}; kT) \triangleq$  закодированная функция фона в точке  $\bar{i}$  и время  $kT$ ;  $f(\bar{i}; kT) \triangleq$  закодированная функция объекта в точке  $\bar{i}$  и время  $kT$ , при этом

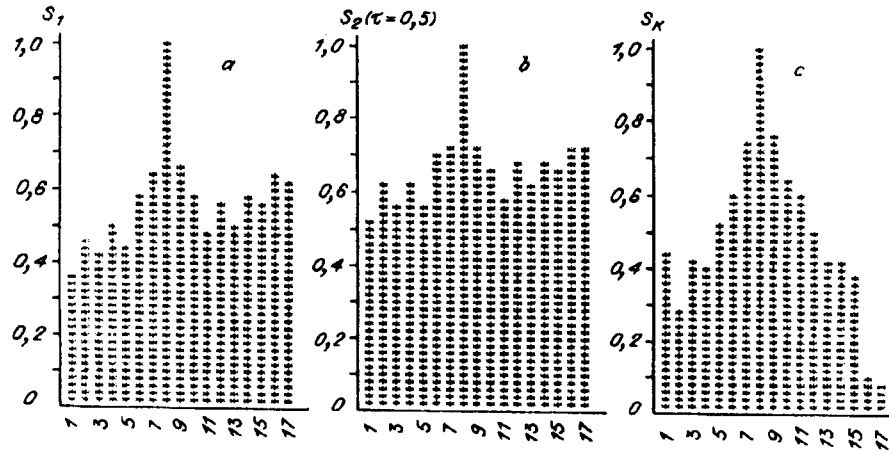


Рис. 4

$\bar{i} = \begin{bmatrix} i \\ j \end{bmatrix} \Delta$ ;  $i, j \in I$ ;  $\Delta$  — интервал пространственной выборки;  $T$  — интервал временной выборки и  $\lambda(\bar{i}; kT) \triangleq$  функция положения, т. е.

$$\lambda(\bar{i}; kT) = \begin{cases} 1 & \forall \bar{i}, \text{ которые принадлежат объекту;} \\ 0 & \text{для всех остальных } \bar{i}. \end{cases}$$

Таким образом,  $\lambda(\bar{i}; kT)$  является функцией «бинарной маски», которая, «подключая» и «отключая» объект и фон, позволяет получить контур объекта. Используя эту модель и описание изображения ЛАП, можно получить оценку  $\bar{\lambda}$  функции  $\lambda$  для динамического объекта. Для этого в процессе перезаписи производится поэлементное сравнение старого  $[t_1 = (k-1)T]$  и нового  $[t_2 = kT]$  эталонов:

$$\bar{\lambda}(\bar{i}; kT) = \begin{cases} 1, & \text{если } a(\bar{i}; kT) = a(\bar{i}; (k-1)T); \\ 0 & \text{в противном случае.} \end{cases}$$

Делая следующий шаг, введем функцию  $D(\bar{i}; kT)$ , для которой зададим начальное значение  $D(\bar{i}; 0) = 1 \forall \bar{i}$ , определив ее как

$$D(\bar{i}; kT) = \begin{cases} D(\bar{i}; (k-1)T) + \tau, & \text{если } a(\bar{i}; kT) = a(\bar{i}; (k-1)T); \\ 1 & \text{в противном случае.} \end{cases} \quad (1)$$

Здесь  $\tau > 0$  — постоянная величина. В данном случае число  $k-1$  определяет количество перезаписей эталона. Как следует из (1), матрица  $D$  по размеру равна эталону. Будем называть матрицу  $D$  матрицей сравнений (МС).

Полученную таким образом МС можно использовать при вычислении меры близости изображений. Дадим новое определение меры близости  $S$ :

$$S_D = \sum_{\bar{i}} \delta(i) D(\bar{i}; kT). \quad (2)$$

Меру близости, определенную выражением (2), назовем динамической мерой близости изображений.

Смысл МС состоит в том, что элементы эталона, не изменяющие своего значения от кадра к кадру, будут иметь больший «вес» по сравнению с осталь-

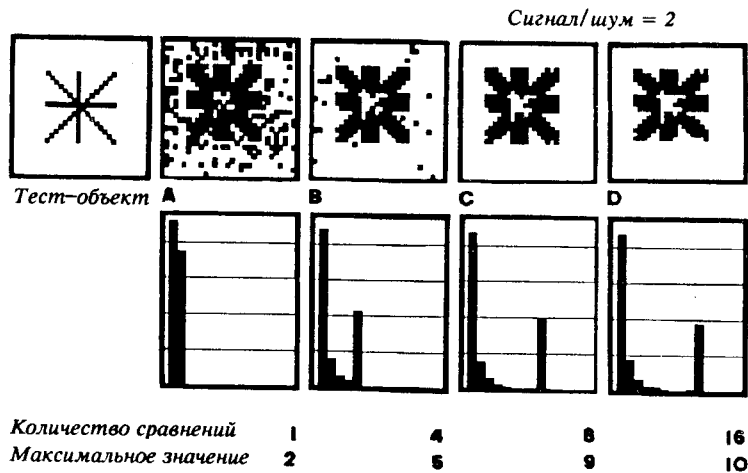


Рис. 5

ными. При устойчивом слежении объект неподвижен относительно эталона, в то время как фон может перемещаться. Это обстоятельство приводит к тому, что в случае динамического объекта элементы МС, соответствующие объекту, будут иметь большие значения, нежели фоновые элементы. Это хорошо видно на рис. 5 и 6.

Тестовое изображение, показанное на этих рисунках, перемещалось относительно сложного фона и при этом зашумлялось. Выделение признаков на изображениях проводилось на окрестностях размером  $4 \times 4$  элемента, и рассматривались четыре класса признаков (исключая изотропный случай). На изображениях МС черные элементы соответствуют максимальным значениям этой матрицы. Под изображением каждой матрицы приведена гистограмма ее значений. На приведенных примерах  $\tau = 1$ .

Из-за априорного незнания формы объекта эталонное изображение обычно выбирается в форме прямоугольника или квадрата. Введение МС приводит к адаптации эталона к форме объекта. Иначе говоря, на счет динамической меры близости идет в основном по элементам изображения, принадлежащим объекту. При этом чем более устойчивым признаком описывается элемент изображения, тем больший вклад он вносит в динамическую меру близости.

Адаптация эталона к наиболее устойчивым признакам объекта приводит к возрастанию пика матрицы динамической меры близости и в конечном итоге

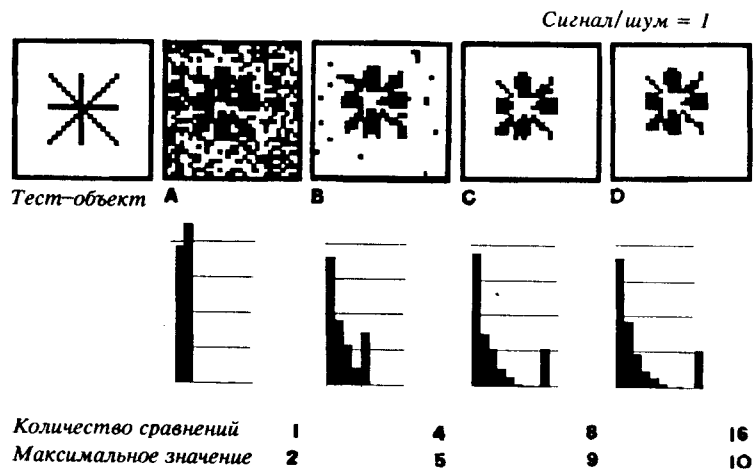


Рис. 6

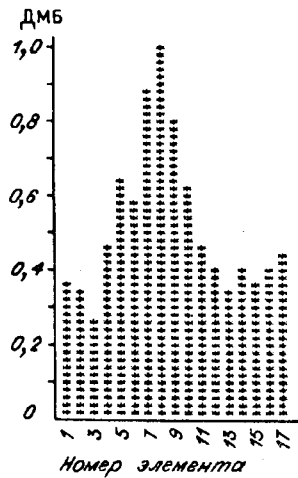


Рис. 7

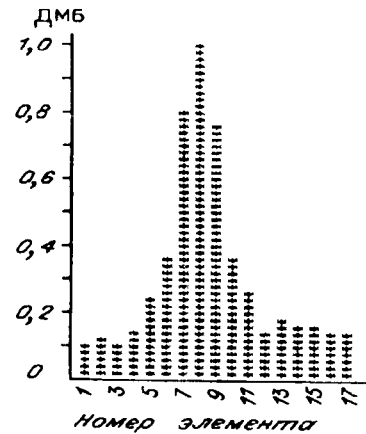


Рис. 8

к повышению надежности слежения. На рис. 7 и 8 приведены графики пиков динамической меры близости (ДМБ) для реального изображения, показанного на рис. 2 для 2-го и 9-го кадров после начала слежения, т. е. после одного и восьми сравнений эталона. При этом эталон имел размер  $16 \times 16$  элементов разложения, а поиск производился на фрагменте размером  $32 \times 32$  элемента.

Рассмотрим вопрос о том, как изменяются значения МС при росте количества перезаписей эталона. Для этого оценим вероятность появления какого-либо значения элемента, не принадлежащего объекту, в случае движения объекта относительно фона.

Будем считать, что  $p$  — вероятность совпадения признаков элемента при перезаписи,  $q$  — вероятность несовпадения ( $p + q = 1$ ),  $p$  и  $q$  независимы для всех кадров (испытания Бернулли). Положим при этом  $\tau = 1$ . При таких условиях процесс роста значения элемента МС описывается цепью Маркова, связанной с рекуррентными событиями [4]. Рассмотрим цепь с набором состояний  $E_1, E_2, \dots$  и переходными вероятностями

$$P = \begin{bmatrix} q & p & 0 & 0 & \dots \\ q & 0 & p & 0 & \dots \\ q & 0 & 0 & p & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \end{bmatrix}.$$

По достижении элементом МС значения  $x$  процесс возрастания с вероятностью  $p$  продолжается, а с вероятностью  $q$  начинается заново — с единичного значения. Последовательные прохождения через состояние  $E_1$  представляют собой рекуррентное событие, и вероятность того, что процесс вернется в состояние  $E_1$  из состояния  $E_x$ , равна  $p^{x-1}q$ . Легко показать, что появление значений элемента МС будет при этом задаваться матрицей

$$P_3 = \begin{bmatrix} q & p & 0 & 0 & \dots \\ q & pq & p^2 & 0 & \dots \\ q & pq & p^2q & p^3 & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \\ \cdot & \cdot & \cdot & \cdot & \dots \end{bmatrix}, \quad (3)$$

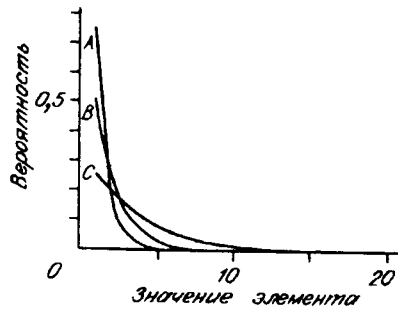


Рис. 9

где номер строки соответствует количеству сравнений, а номер столбца — значению элемента.

На рис. 9 приведены расчетные графики вероятностей появления значений элементов для случая  $p = 0,25$  (кривая А),  $p = 0,50$  (кривая В),  $p = 0,75$  (кривая С).

Значения элементов МС могут неограниченно возрастать при росте количества перезаписей эталона. Чтобы избежать этого, достаточно ограничить рост этих значений, т. е. ввести максимальное допустимое значение  $x_m$ . Зная вероятность появления больших значений элементов и задавая числом  $\epsilon > 0$ , легко определить коли-

чество сравнений  $k_0$ , выше которого вероятность  $p(x > x_m | k > k_0) < \epsilon$ , где  $x$  — значение элемента. Число  $k_0$  служит отправной точкой для задания максимального значения МС.

Из выражения (3) видно, что, ограничив значения МС числом  $x_m$ , при количестве сравнений  $k > k_0$  вероятности появления каких-либо значений элементов имеют постоянные значения независимо от дальнейшего количества циклов перезаписи. Будем называть такую матрицу установившейся.

Зная вероятности появления значений каждого элемента установившейся матрицы, можно оценить вероятность того, сколько элементов с заданным значением будет иметь МС. Эта вероятность определяется биномиальным распределением

$$P_B(m; n, p_y) = \binom{n}{m} p_y^m (1 - p_y)^{n-m}, \quad (4)$$

где  $m$  — количество элементов, имеющих заданное значение,  $n$  — количество элементов МС,  $p_y$  — вероятность появления заданного значения элемента установившейся матрицы. Распределение (4) имеет максимум, который соответствует наиболее вероятному числу  $m = m_B$ .

На рис. 10 приведена расчетная кривая, отражающая вероятность появления числа элементов  $m = m_B$  для каждого из значений МС при  $p = 0,50$  и  $\tau = 1$ . На этом же рисунке приведены экспериментальные результаты, полученные по нескольким установившимся матрицам для случая, показанного на рис. 5, и  $x_m = 10$ .

Матрицу сравнений саму по себе можно рассматривать как изображение. Поэтому к этой матрице применимы все методы обработки и анализа, которые существуют для изображений. Если анализировать только максимальные значения этой матрицы, то ее можно рассмат-

ривать как бинарное изображение. Алгоритмы обработки бинарных изображений хорошо известны, и обработка изображений такого типа не вызывает принципиальных трудностей [5—7].

В случае обработки бинарной матрицы сравнений можно выделить три основные цели: 1) подавление случайных выбросов и выделение связанных областей, принадлежащих объекту; 2) выравнивание краев объекта; 3) компенсирование уширения контуров для приведения выделенной части МС к истинным размерам объекта (поскольку при выделении признаков используются локальные окрестности конечных размеров, то происходит уширение контуров; при размере

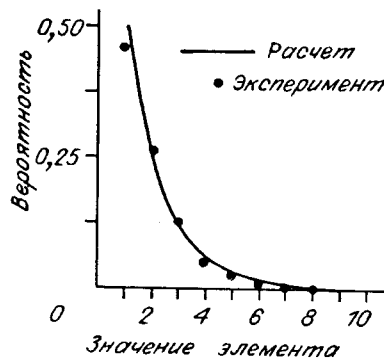


Рис. 10

окрестности  $l \times l$  элементов границы объекта могут расширяться на  $l - 1$  элемент по каждой координате относительно истинных границ).

Обработанная и приведенная к истинным размерам бинарная МС может рассматриваться как оценка функции  $\lambda$  для динамического объекта. Значение функции  $\lambda$  позволяет сегментировать изображение и как следствие оценивать аффинные трансформации объекта в плоскости изображения (перемещение, поворот, изменение масштаба) [8—10].

Важным следствием введения МС и динамической меры близости является возможность продолжать слежение в том случае, когда видна лишь часть объекта (например, при заезде за препятствие, затенении и т. д.).

Применение подобных методов дает дополнительные возможности при построении систем технического зрения: на их основе возможно создание систем селекции и обнаружения подвижных объектов, систем контроля за срывом слежения, центрирование объектов и т. д. Более подробно эти вопросы рассматриваются во второй части статьи.

#### СПИСОК ЛИТЕРАТУРЫ

1. Pat. 4272783 US. Correlation follower for tracking an object /R. Jonsson, G. Ludvigsson, L.-A. Warnstam.—Publ. 9.06.81.
2. Попов П. Г., Порядина С. А. Кодирование изображений локальными анизотропными признаками // Оптические и оптико-электронные средства обработки информации.—Л.: ФТИ, 1989.
3. Nahi N., Logez-Mora. Estimation-detection of object boundaries in noisy images // IEEE Trans. Automat. Contr.—1978.—АС-В.—Р. 834.
4. Феллер В. Введение в теорию вероятностей и ее приложение.—М.: Наука, 1967.
5. Прэтт У. Цифровая обработка изображений.—М.: Мир, 1982.
6. Rosenfeld A., Kak A. C. Digital picture processing.—N. Y.: Academic Press, 1976.
7. Престон К. и др. Основы клеточной логики с приложениями к обработке изображений в медицине // ТИИЭР.—1979.—67, № 5.
8. Shalkoff R. J., Mc'Veigh E. S. A model and tracking algorithm for a class of video targets // IEEE Trans. Pattern Anal. Machine Intell.—1982.—PAMI-4, N 1.—Р. 2.
9. Cafforio C., Rocca F. Methods for measuring small displacements of television images // IEEE Trans. Inform. Theory.—1976.—IT-22, N 5.—Р. 573.
10. Netravali A. M., Robbins J. D. Motion compensated television coding: Part 1 // Bell Syst. Tech. J.—1979.—58.—Р. 631.
11. Попов П. Г., Порядина С. А. Построение анизотропных фильтров для коррекции строчных геометрических шумов // Автометрия.—1991.—№ 2.

*Поступила в редакцию 9 июня 1992 г.*