

СИСТЕМЫ И МЕТОДЫ АНАЛИЗА ИЗОБРАЖЕНИЙ

УДК 517.518.85

В. М. Ефимов, А. Н. Колесников

(Новосибирск)

СИНУСНОЕ ПРЕОБРАЗОВАНИЕ СИГНАЛА
С ПРЕДВАРИТЕЛЬНОЙ ИНТЕРПОЛЯЦИЕЙ

Изложен метод сжатия одномерных и двумерных сигналов на основе дискретного синусного преобразования с предварительной интерполяцией, позволяющий ослабить фрагментарный шум. Приводятся соотношения для интерполяции изотропных и анизотропных изображений по значениям на границе фрагмента. Предложенный метод проверен на тестовых изображениях.

Введение. В ряде систем сжатия изображений с преобразованием для декорреляции сигнала $f(x, y)$ ($-l \leq x, y \leq l$) используют косинусное преобразование [1—4]. Это преобразование асимптотически оптимально в одномерном случае для сигнала с корреляционной функцией типа

$$\rho(x) \cong 1 - \alpha |x| \quad (1)$$

при $\rho(x) \rightarrow 1$ [5].

Недостаток косинусного преобразования — появление неприятного для зрительного восприятия фрагментарного шума при сильном сжатии сигнала, несмотря на то что общее качество изображения можно считать приемлемым. Предпринимался ряд попыток борьбы с этим явлением путем фильтрации изображения на границах фрагмента либо применением дискретного косинусного преобразования (ДКП) ко всему изображению [4].

Теоретические и экспериментальные исследования авторов показали, что при том же самом уровне сжатия можно ослабить эффект фрагментарности изображения, если использовать дискретное синусное преобразование (ДСП) изображения после его предварительной интерполяции, а также общие границы смежных фрагментов. Суть предлагаемого метода заключается в следующем: все изображение разбивается на частично перекрывающиеся фрагменты с общей границей (рис. 1), по значениям функции интенсивности на границах фрагментов строится интерполирующая функция, к разностной функции применяется дискретное синусное преобразование, полученные трансформанты квантуются и кодируются.

В разд. 1 и 2 рассмотрены вопросы построения интерполирующих полиномов в одномерном и двумерном случаях соответственно. В разд. 3 изложен алгоритм кодирования изображения. В разд. 4 приведены результаты

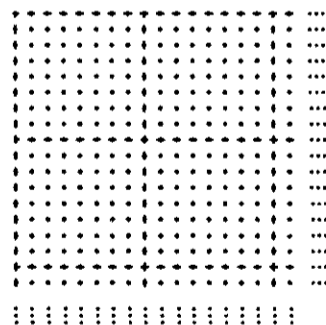


Рис. 1. Схема разбиения массива на частично перекрывающиеся (один пиксел) фрагменты ($N = 9$)

проверки алгоритма на тестовых изображениях.

1. Преобразование на прямой. Из классического анализа Фурье известно, что для применения синусного преобразования необходимо, чтобы функция (сигнал) равнялась нулю на краях интервала разложения.

Памятуя о связи дисперсии ошибки интерполяции и ее энтропии, естественно также потребовать, чтобы дисперсия ошибки интерполяции была минимальной. Асимптотически оптимальной интерполяционной формулой, удовлетворяющей этим двум условиям, в одномерном случае является [6]

$$f^*(x, y) = (1/2)f(-l)(1 - x/l) + (1/2)f(l)(1 + x/l). \quad (2)$$

Сравним в одномерном случае дисперсии трансформант косинусного и синусного преобразований с предварительной интерполяцией по формуле (2), когда корреляционная функция сигнала определяется соотношением (1), т. е. при $\alpha l < 1$. Вычисления показывают, что при косинусном преобразовании сигнала с этой корреляционной функцией дисперсия k -й трансформанты

$$\lambda_k = 4\alpha l / (k\pi)^2 \quad (k > 0). \quad (3)$$

Этим же соотношением определяются дисперсии трансформант разностного сигнала

$$\overset{\circ}{f}(x) = f(x) - f^*(x) \quad (4)$$

с корреляционной функцией

$$\overset{\circ}{\rho}(x, \xi) = -\alpha |x - \xi| + \alpha(l - x\xi/l). \quad (5)$$

Отметим, что синусное преобразование является решением интегрального уравнения

$$\int_{-l}^l \overset{\circ}{\rho}(x, \xi) \varphi_k(\xi) d\xi = \lambda_k \varphi_k(x), \quad (6)$$

если корреляционная функция сигнала определяется соотношением (5). Если $\overset{\circ}{\rho}(x, \xi) = 1 - \alpha |x - \xi|$, то асимптотически оптимальное решение уравнения (6) — косинусное преобразование.

Таким образом, если рассматривать достаточно длинную последовательность прилегающих друг к другу интервалов разложения, то различие между косинусным и синусным преобразованиями в непрерывном случае сводится к различию между дисперсией нулевой составляющей косинусного преобразования и дисперсией значения сигнала на границе между интервалами разложения.

2. Преобразование на плоскости. В двумерном случае использование синусного преобразования связано с интерполированием функций двух переменных внутри квадрата. При этом интерполирующая функция на границе квадрата ($-l \leq x, y \leq l$) должна совпадать с сигналом. Такую функцию можно получить, последовательно применяя к сигналу $f(x, y)$ формулу (2) сначала вдоль одной, а затем вдоль другой оси. При этом (рис. 2)

$$f^*(x, y) = f_1^*(x, y) + f_2^*(x, y) - f_3^*(x, y), \quad (7)$$

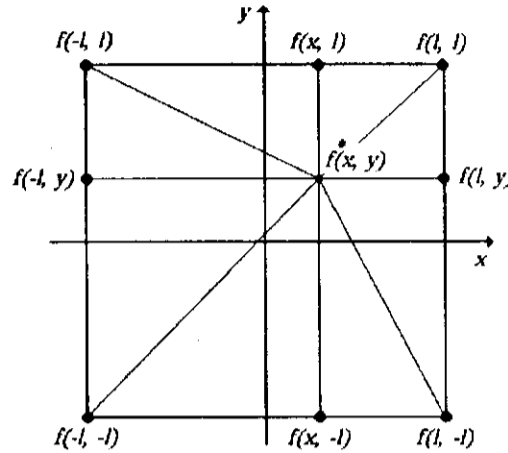
где

$$f_1^*(x, y) = (1/2)(f(x, -l)(1 - y/l) + f(x, l)(1 + y/l)), \quad (7.1)$$

$$f_2^*(x, y) = (1/2)(f(-l, y)(1 - x/l) + f(l, y)(1 + x/l)), \quad (7.2)$$

Рис. 2. Схема интерполяции по восьми точкам (для анизотропного сигнала)

$$\begin{aligned}
 f_3^*(x, y) = & (1/4)(f(-l, -l) \times \\
 & \times (1 - y/l)(1 - x/l) + \\
 & + f(-l, l)(1 + y/l)(1 - x/l) + \\
 & + f(l, -l)(1 - y/l)(1 + x/l) + \\
 & + f(l, l)(1 + y/l)(1 + x/l)).
 \end{aligned}
 \tag{7.3}$$



Отметим, что в случае сигнала с корреляционной функцией

$$\rho(x, y) = (1 - \alpha|x|)(1 - \alpha|y|) \tag{8}$$

интерполяционная формула (7), как показывает анализ, обеспечивает минимум среднего квадрата ошибки интерполяции, что при нормальном распределении сигнала минимизирует его энтропию. Поиск точного решения для интерполяционной функции, когда сигнал изотропен, приводит к линейной системе четырех интегральных уравнений для весовых функций на сторонах квадрата (см. приложение). Анализ этой системы показывает, что для изотропного сигнала $\rho(x, y) = \rho(\sqrt{x^2 + y^2})$ более близкой к оптимальной, чем (7), оказывается интерполяционная формула (рис. 3):

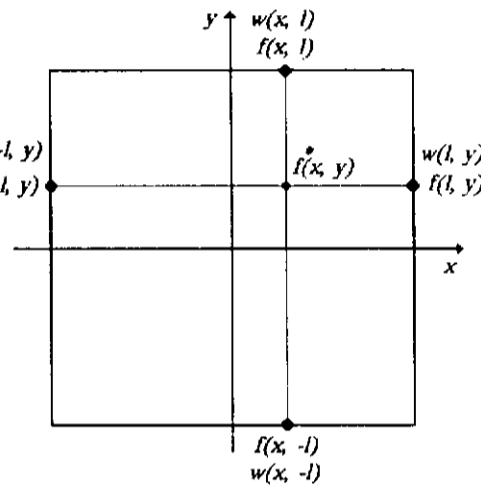


Рис. 3. Схема интерполяции по четырем точкам

$$f^*(x, y) = (1/2 + \mu)f_1^*(x, y) + (1/2 - \mu)f_2^*(x, y), \tag{9}$$

где коэффициент μ определяется из условия минимума среднего квадрата ошибки интерполяции и равен

$$\mu = \frac{\langle f(x, y)(f_1^*(x, y) - f_2^*(x, y)) \rangle - (1/2)(\langle (f_1^*(x, y))^2 \rangle - \langle (f_2^*(x, y))^2 \rangle)}{\langle (f_1^*(x, y) - f_2^*(x, y))^2 \rangle}, \tag{10}$$

знаком $\langle \rangle$ обозначено математическое ожидание.

Для корреляционной функции

$$\rho(x, y) \cong 1 - \alpha\sqrt{x^2 + y^2} \tag{11}$$

в асимптотике ($\rho(x, y) \rightarrow 1$)

$$\begin{aligned}
 \mu = & [(l+x)(l-x) - (l+y)(l-y)] / [\sqrt{(l+y)^2 + (l+x)^2}(l-y)(l-x)/l + \\
 & + \sqrt{(l+y)^2 + (l-x)^2}(l-y)(l+x)/l + \sqrt{(l-y)^2 + (l+x)^2}(l+y)(l-x)/l + \\
 & + \sqrt{(l-y)^2 + (l-x)^2}(l+y)(l+x)/l]
 \end{aligned}$$

$$+ \sqrt{(l-y)^2 + (l-x)^2} (l+y)(l+x) / l - 2(l+y)(l-y) - 2(l+x)(l-x)]. \quad (12)$$

Использование (9) вместо (7), если корреляционная функция определяется соотношением (11), снижает максимум среднего квадрата ошибки примерно в 1,4 раза.

3. Алгоритм сжатия изображения. Процедура сжатия изображения включает в себя три этапа: предварительную интерполяцию, дискретное синусное преобразование и кодирование трансформант.

Предварительная интерполяция. А. Все изображение делится на частично перекрывающиеся фрагменты размером $N \times N$ с общей границей (см. рис. 1); в случае необходимости изображение дополняется соответствующим числом столбцов (строк) путем повтора крайнего столбца (строки).

В. Для функции интенсивности $b(j)$ на границе фрагмента по значениям функции интенсивности в вершинах строится интерполирующая $b_i(j)$ и разностная $b_D(j)$ функции:

$$b_D(j) = b(j) - b_i(j), \quad j = \overline{0, N-1}.$$

С. Для функции интенсивности $f(j, k)$ внутри фрагмента по значениям функции интенсивности на границах фрагмента строится интерполирующая $f_i(j, k)$ и разностная $f_D(j, k)$ функции (см. рис. 3):

$$f_D(j, k) = f(j, k) - f_i(j, k) \quad (j, k = \overline{0, N-1}).$$

Дискретное синусное преобразование. А. Для разностной функции $b_D(j)$ (принимаяющей нулевые значения на краях интервала длиной N) находятся трансформанты ДСП $B_D(k)$:

$$B_D(k) = \sqrt{2/(N-1)} \sum_{j=1}^{N-2} b_D(j) \sin\left(\pi \frac{jk}{N-1}\right), \quad k = \overline{1, N-2}. \quad (13)$$

В. Для разностной функции $f_D(j, k)$ (принимаяющей нулевые значения на границах фрагмента) находятся трансформанты двумерного ДСП $F_D(m, n)$ ($m, n = \overline{1, N-2}$); двумерное ДСП сводится к двукратному применению одномерного ДСП (например, сначала по столбцам, а затем по строкам).

Кодирование. Полученные на предыдущем этапе данные (значения функции интенсивности в вершинах фрагментов и трансформанты ДСП функций $B_D(k)$ и $F_D(m, n)$) квантуются с некоторыми шагами q_b и q_f соответственно, регулирующими степень сжатия и уровень ошибки восстановления, и кодируются.

Учитывая, что каждая вершинная точка фрагмента принадлежит четырем смежным фрагментам, а каждая граничная точка — двум фрагментам (см. рис. 1), получаем, что фрагмент полностью описывается

$$1 + 2(N-2) + (N-2)(N-2) = (N-1)(N-1)$$

отсчетом (без учета краевых эффектов).

Восстановление сигнала на приемном конце включает в себя также три этапа: декодирование, обратное ДСП и интерполяцию.

Декодирование. На приемном конце полученные трансформанты $B_D(k)$ и $F_D(m, n)$ декодируются и домножаются на коэффициенты q_b и q_f .

Обратное ДСП. А. С помощью обратного синусного преобразования из трансформант $B_D(k)$ восстанавливаются значения функции интенсивности $b_D(j)$ на границах фрагмента; как следует из (13), обратное ДСП в данном случае тождественно прямому, что упрощает его программную и аппаратную реализацию:

$$b_D(j) = \sqrt{2/(N-1)} \sum_{k=1}^{N-2} B_D(k) \sin\left(\pi \frac{jk}{N-1}\right), \quad j = \overline{1, N-2}.$$

В. С помощью обратного двумерного синусного преобразования из трансформант $F_D(m, n)$ восстанавливаются значения функции интенсивности внутри фрагмента $f_D(j, k)$.

Интерполяция. А. По полученным значениям в углах фрагмента строится интерполирующая функция $b_I(j)$.

В. По интерполирующей $b_I(j)$ и разностной $b_D(j)$ функциям восстанавливается функция интенсивности на границе фрагмента:

$$b(j) = b_D(j) + b_I(j).$$

С. По функции $b(j)$ строится интерполирующая функция $f_I(j, k)$ внутри фрагмента (см. рис. 3);

Д. По интерполирующей $f_I(j, k)$ и разностной $f_D(j, k)$ функциям восстанавливается функция интенсивности на фрагменте.

Для вычисления одного отсчета разностной функции $b_D(j)$ требуется два умножения и два сложения; для вычисления одного отсчета функции $f_D(j, k)$ требуется четыре умножения и четыре сложения; при этом полагается, что коэффициенты интерполяционного многочлена вычисляются заранее и хранятся в виде таблицы. В общей сложности на один фрагмент размером $N \times N$ требуется $4(N-2)(N-1)$ умножений и столько же сложений.

Характерной особенностью разложения по тригонометрическим функциям является наличие быстрых алгоритмов их вычисления [7—9]. На рис. 4 представлена блок-схема быстрого ДСП, разработанная для предлагаемого метода кодирования при $N = 9$. Стоит отметить, что в данном случае весь алгоритм ДСП определяется только тремя коэффициентами: S_1, S_2, S_3 , где $S_k = \sin(k\pi/8)$, $k = 1, 2, 3$ (в связи с тем, что в данном случае прямое и обратное ДСП тождественны, эту блок-схему можно привести к симметричному виду).

Как следует из представленной схемы, для вычисления ДСП одной строки для $N = 9$ требуется шесть операций умножения и 18 — сложения; для сравнения укажем, что прямая реализация ДСП требует 36 умножений и 44 сло-

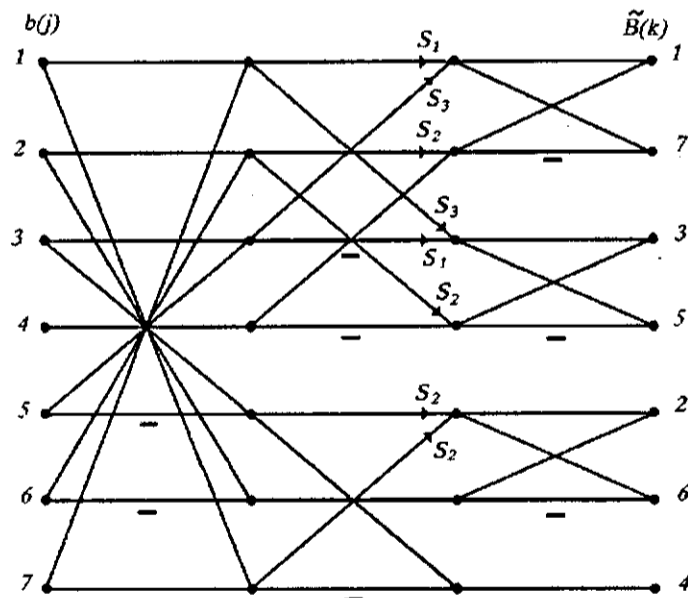


Рис. 4. Блок-схема быстрого дискретного синусного преобразования



Рис. 5. Исходное изображение:
 $H_0 = 7,49$ бит/отсчет



Рис. 6. Восстановленное изображение:
ДСП; $q_b = 16$; $q_f = 32$; $H_1 = 0,78$ бит/отсчет;
С/Ш = 34,9 дБ

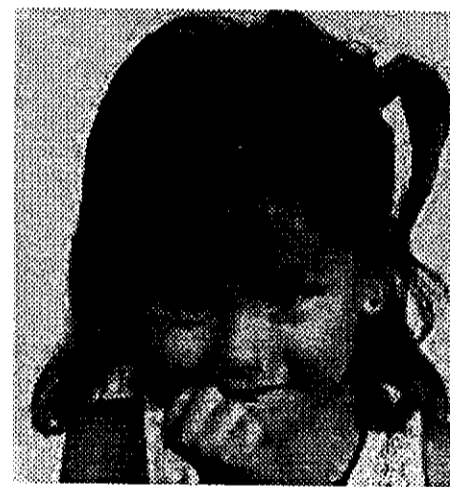


Рис. 7. Восстановленное изображение:
ДСП; $q_b = 64$; $q_f = 128$; $H_1 = 0,25$ бит/отсчет;
С/Ш = 29,0 дБ



Рис. 8. Восстановленное изображение:
ДСП; $q_b = \infty$; $q_f = \infty$; $H_1 = 0,11$ бит/отсчет;
С/Ш = 23,3 дБ



Рис. 9. Восстановленное изображение:
ДСП; $q_b = 1$; $q_f = \infty$; $H_1 = 1,17$ бит/отсчет;
С/Ш = 26,8 дБ



Рис. 10. Восстановленное изображение:
ДКП; $q = \infty$; $H_1 = 0,11$ бит/отсчет; С/Ш = 20,5 дБ

жения. Строго говоря, на представленной схеме фигурирует восемь умножений, но, как нетрудно видеть, два умножения на S_2 можно исключить.

Таким образом, для синусного преобразования одного фрагмента, включающего в себя обработку 8 строк и 8 столбцов, быстрый алгоритм требует 96 умножений и 288 сложений; прямая реализация потребует 576 умножений и 704 сложения соответственно (здесь также учтено, что некоторые коэффициенты равны нулю либо единице).

Если сравнивать предложенный алгоритм с дискретным косинусным преобразованием, то в рассматриваемом случае быстрое дискретное синусное преобразование фрагмента 9×9 само по себе требует гораздо меньше операций, чем быстрое ДКП соответствующего фрагмента 8×8 (192 умножения, 292 сложения [8]), но затраты на предварительную интерполяцию приводят к тому, что по общему числу операций предлагаемый метод на 50 % уступает кодированию на основе ДКП.

4. Результаты проверки. Предлагаемый метод кодирования был проверен на тестовых изображениях, одно из которых приведено на рис. 5—10. Коэффициент сжатия определялся энтропией H_1 квантованных трансформант; для восстановленных изображений вычислялось отношение сигнал/шум, которое является некоторой мерой близости восстановленного изображения к исходному, хотя, как известно, не является критерием визуального качества.

На рис. 5 представлено тестовое изображение "Girl" (формат 303×372), на рис. 6—10 — восстановленные изображения с разной степенью сжатия.

При малых шагах квантования восстановленное изображение, естественно, практически не отличается от исходного. При довольно больших значениях шага квантования восстановленное изображение все еще мало отличается от исходного (см. рис. 6). При увеличении шага квантования ошибка восстановления проявляется в виде ореола на тех фрагментах, которые характеризуются резким перепадом интенсивности, участки с плавными изменениями интенсивности передаются неплохо (см. рис. 7).

В предельном случае, когда $q_b = \infty$ и $q_f = \infty$, все подлежащие кодированию трансформанты обращаются в нуль, восстановленное изображение полностью определяется значениями в вершинах фрагментов, а его энтропия — энтропией вершинных точек: $H_1 = 0, 11$ бит (см. рис. 8). Для сравнения на рис. 9 представлено изображение, у которого обращены в нуль только трансформанты $F_D(m, n)$ ($q_f = \infty$), но точно переданы трансформанты $B_D(k)$ ($q_b = 1$), т. е. все изображение описывается интерполирующей функцией $f_i(j, k)$; энтропия в этом случае определяется энтропией вершинных отсчетов и трансформант $B_D(k)$.

Что же касается кодирования с ДКП, то в предельном случае, когда все трансформанты зануляются ($q = \infty$), восстановленное изображение представляет собой мозаику из квадратов, интенсивность которых определяется средней интенсивностью по фрагменту (см. рис. 10), а коэффициент сжатия — энтропией этих средних значений.

Заключение. Результаты моделирования показали, что предложенный метод кодирования на основе дискретного синусного преобразования с предварительной интерполяцией на фрагментах с частичным перекрытием позволяет избежать резких скачков интенсивности между фрагментами на участках изображений с плавно меняющейся интенсивностью. По своим характеристикам предложенный метод примерно соответствует известным методам на основе ДКП.

ПРИЛОЖЕНИЕ

Необходимо выбором весовых функций минимизировать средний квадрат ошибки интерполяции

$$\min_{\omega} \langle (f(x_0, y_0) - f^*(x_0, y_0))^2 \rangle, \quad (П1)$$

где интерполирующая функция

$$f^*(x_0, y_0) = \int_{-l}^l f(x, -l)w(x, -l)dx + \int_{-l}^l f(l, y)w(l, y)dy + \\ + \int_{-l}^l f(x, l)w(x, l)dx + \int_{-l}^l f(-l, y)w(-l, y)dy. \quad (\text{П2})$$

После дифференцирования (П1) по весовым функциям и вычисления математического ожидания получим линейную систему четырех интегральных уравнений:

$$\rho(x - x_0, l + y_0) = \int_{-l}^l \rho(\xi - x, 0)w(\xi, -l)d\xi + \int_{-l}^l \rho(l - x, l + \xi)w(l, \xi)d\xi + \\ + \int_{-l}^l \rho(\xi - x, 2l)w(\xi, l)d\xi + \int_{-l}^l \rho(l + x, l + \xi)w(-l, \xi)d\xi; \quad (\text{П3.1})$$

$$\rho(x - x_0, l - y_0) = \int_{-l}^l \rho(\xi - x, 2l)w(\xi, -l)d\xi + \int_{-l}^l \rho(l - x, l - \xi)w(l, \xi)d\xi + \\ + \int_{-l}^l \rho(\xi - x, 0)w(\xi, l)d\xi + \int_{-l}^l \rho(l + x, l - \xi)w(-l, \xi)d\xi; \quad (\text{П3.2})$$

$$\rho(l - x_0, y - y_0) = \int_{-l}^l \rho(l - \xi, l + y)w(\xi, -l)d\xi + \int_{-l}^l \rho(0, \xi - y)w(l, \xi)d\xi + \\ + \int_{-l}^l \rho(l - \xi, l - y)w(\xi, l)d\xi + \int_{-l}^l \rho(2l, \xi - y)w(-l, \xi)d\xi; \quad (\text{П3.3})$$

$$\rho(l + x_0, y - y_0) = \int_{-l}^l \rho(l + \xi, l + y)w(\xi, -l)d\xi + \int_{-l}^l \rho(2l, \xi - y)w(l, \xi)d\xi + \\ + \int_{-l}^l \rho(l + \xi, l - y)w(\xi, l)d\xi + \int_{-l}^l \rho(0, \xi - y)w(-l, \xi)d\xi. \quad (\text{П3.4})$$

Непосредственной проверкой можно убедиться, что решением системы для анизотропного поля $\rho(x, y) = (1 - \alpha|x|)(1 - \alpha|y|)$ является соотношение (7).

СПИСОК ЛИТЕРАТУРЫ

1. Ahmed N., Natarajan T., Rao K. R. Discrete cosine transform // IEEE Trans. Comput.—1974.—С-23.—Р. 90.
2. Андерсон Дж., Фралик С. К., Хамилтон Э. и др. Кодек для передачи цветной видеоинформации по цифровым телефонным линиям // Электроника.—1983.—№ 22.
3. Wallace G. K. The JPEG still picture compression standard // Commun. of the ACM.—1991.—34, N 47.
4. Ho В. К. Т., Chan К. К., Ishimitsu Y. et al. Expandable image compression system, a modular approach // SPIE (Medical Imaging).—1987.—767.—Р. 286.
5. Ефимов В. М., Золотухин Ю. Н., Резник А. Л. Асимптотически оптимальная декорреляция стационарной последовательности равноотстоящих отсчетов сигнала // Автометрия.—1991.—№ 5.

6. Ефимов В. М. Асимптотически оптимальные интерполяционные соотношения // Автоматика.—1992.—№ 4.
7. Chen W., Smith C. H., Fralick S. A fast computational algorithm for the discrete cosine transformation // IEEE Trans. Commun.—1974.—COM-25.—P. 1004.
8. Haque M. A. A two-dimensional fast cosine transform // IEEE Trans. on Acoustics, Speech and Signal Proces.—1983.—ASSP-33, N 6.
9. Yip P., Rao K. R. A fast computational algorithm for the discrete sine transform // IEEE Trans. on Commun.—1980.—COM-28, N 2.

Поступила в редакцию 25 июня 1992 г.

УДК 621.396

Е. П. Путятин, А. А. Добрынин, О. В. Сытник
(Харьков)

АЛГОРИТМ ВТОРИЧНОЙ ОБРАБОТКИ ДИНАМИЧНЫХ СЦЕН

Предложена процедура фильтрации при анализе динамических сцен. Обработка сцены строится на основе модели марковского процесса, параметры которой оцениваются в процессе анализа изображения. Процедура предназначена для сглаживания последовательности оценок параметров изображений в динамических сценах.

При анализе сцен с помощью алгоритмов, основанных на преобразовании Хо [1], возникает проблема обработки сильно зашумленных изображений, когда большая часть контура обнаруживаемого объекта разрушена вследствие воздействия флуктуационной помехи. При этом используются многоступенчатые пошаговые системы обработки с нелинейной в общем случае передаточной характеристикой [2]. На рис. 1 изображена обобщенная схема получения оценок в таких системах. Исходное изображение 1 подвергается предварительной обработке в блоке 2, где оно приводится к бинарному путем пороговой обработки. Флуктуационная помеха 3 условно показана на рис. 1 как аддитивный бинарный шум. Бинарное изображение 4 подвергается преобразованию Хо 5, в результате которого получаем спектр 6 параметров наблюдаемой сцены. В результате анализа спектра 6 и оценки максимальных спектральных компонент в блоке 7 вычисляется оценка 9. Ошибки оценивания, порождаемые воздействием флуктуационной помехи 3 на нелинейную процедуру обработки 5—7, показаны в виде аддитивной компоненты 8. При этом свойства результирующей оценки определяются как статистическими свойствами исходных (наблюдаемых) процессов, так и величинами выбранных порогов. Если анализируется последовательность сцен, то на выходе системы обработки будет последовательность оценок, отражающая не только поведение исследуемого объекта во времени и пространстве, но и состояние и способ обработки исходного сигнала. Из этого следует, что полученная последовательность оценок является негауссовской и нестационарной.

Негауссовские процессы для исчерпывающего вероятностного описания требуют такого количества статистической информации, которое трудно получить из физических измерений по причинам ограниченного времени анализа, аппаратных сложностей или ввиду ограничений, предусмотренных условиями самой задачи. Поэтому приходится отказываться от полного ве-

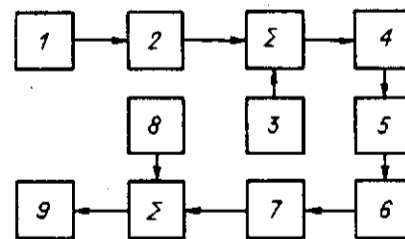


Рис. 1