

COMPUTER IMAGE PROCESSING

V. A. SOIFER

This paper deals with the fundamentals of computer image processing. Computer models of images and imaging optical systems are described. The procedures of discretization and discrete image representations through 2D sequences are considered. The basic techniques of computer image processing are featured. Element-by-element image transformations, which include linear contrasting, thresholding and preparation are discussed. The methods and algorithms for extracting contours from images are reported.

В статье рассмотрены основы компьютерной обработки изображений. Построены математические модели изображений и изображающих оптических систем. Рассмотрены процедуры дискретизации и дискретные представления изображений с помощью двумерных последовательностей. Изложены основные методы компьютерной обработки изображений. Рассмотрены поэлементные преобразования изображений, которые включают линейное контрастирование, пороговую обработку и препарирование. Даны методы и алгоритмы выделения контуров на изображениях.

© Сойфер В. А., 1996

КОМПЬЮТЕРНАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

Часть 1. Математические модели

В. А. СОЙФЕР

Самарский государственный аэрокосмический университет

ВВЕДЕНИЕ

Недаром говорят: “Лучше один раз увидеть, чем сто раз услышать”. Исследования подтверждают, что информационная пропускная способность органов зрения значительно выше, чем у других каналов передачи информации, доступных человеку. В теории информации доказано, что, подбрасывая монету и наблюдая результат, мы всякий раз получаем одну двоичную единицу (бит) информации. Каждая буква в тексте несет примерно четыре бита информации. Изображение участка поверхности Земли, полученное из космоса, содержит примерно 10 миллионов бит информации! Переработать такое количество информации под силу только самому современному компьютеру. Чтобы научить машину обрабатывать изображения, требуется иметь мощный комплекс технических средств, математический аппарат, алгоритмы и большое количество программ.

Часть первая публикации посвящена построению математических моделей оптических изображений и их дискретным представлениям. Во второй части публикации рассматриваются методы и алгоритмы, а также несколько примеров решения прикладных задач.

1. МАТЕМАТИЧЕСКИЕ МОДЕЛИ ОПТИЧЕСКИХ ИЗОБРАЖЕНИЙ

1.1. Функция яркости

Необходимость построения математической модели возникает сразу же при использовании компьютера для обработки изображений. Оценивая “на глаз” расстояние между двумя предметами, мы не задумываемся о том, как это делается. Поручив это компьютеру, мы обязаны научить его выполнять подобные действия, то есть заложить в него соответствующие данные и алгоритмы. Хорошо известно, что компьютер имеет дело с массивами чисел в качестве данных. Таким образом, первой задачей компьютерной обработки изображений является перевод изображений в числовую форму. Это требует конкретизации самого понятия “изображение”.

Рассмотрим объект, освещенный источником света (рис. 1). На некотором расстоянии от объекта распределение энергии источника светового излу-

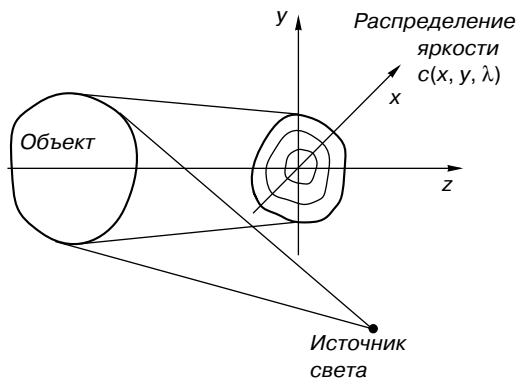


Рис. 1. Формирование изображения объекта, освещенного источником света.

нения, отраженного объектом, по пространственным координатам x , y и по длинам волн λ описывается функцией $c(x, y, \lambda)$. Эта функция является неотрицательной; ее максимальное значение в изображающих системах ограничено предельной величиной светочувствительности регистрирующих сред,

$$0 \leq c(x, y, \lambda) \leq A, \quad (1)$$

где A — максимальная яркость изображения.

Геометрические размеры изображения ограничены характеристиками формирующей системы и параметрами фоторегистрирующей среды. Будем полагать, что все изображения отличны от нуля в прямоугольной области:

$$-L_x \leq x \leq L_x, \quad -L_y \leq y \leq L_y. \quad (2)$$

Человеческое зрение и видеодатчики обладают спектральной чувствительностью, описываемой функцией $v(\lambda)$. Например, как известно, человеческий глаз обладает чувствительностью к свету в диапазоне волн от $\lambda_{\min} = 0,35$ мкм до $\lambda_{\max} = 0,78$ мкм. При этом функция спектральной чувствительности достигает своего максимума приблизительно в середине этого диапазона и спадает к его краям.

Каждый видеодатчик обладает индивидуальной характеристикой спектральной чувствительности, обусловленной физикой прибора. Имеются видеодатчики ультрафиолетового и инфракрасного диапазонов, которые широко используются, например, при проведении спектрально-аналитических съемок Земли из космоса.

Как в случае наблюдения объекта человеком, так и в случае использования видеодатчика наблюдаемое изображение является результатом усреднения функции $c(x, y, \lambda)$ по диапазону длин волн с весовой функцией $v(\lambda)$ и описывается выражением

$$f(x, y) = \int_{\lambda_{\min}}^{\lambda_{\max}} c(x, y, \lambda) v(\lambda) d\lambda, \quad (3)$$

Функцию $f(x, y)$ в дальнейшем будем называть изображением. Таким образом, изображение — это ограниченная функция двух пространственных переменных, заданная на ограниченной прямоугольной области.

1.2. Двумерные линейные системы

Из курса физики хорошо известно понятие оптической системы, осуществляющей преобразование изображений по определенным правилам, определяемым совокупностью используемых в ней оптических элементов и их взаимосвязью.

С математической точки зрения под системой будем понимать правило L , ставящее в соответствие входной функции f выходную функцию g . Различают одномерные $1D$ и двумерные $2D$ системы. Одномерные системы преобразуют функции одной переменной:

$$g(x) = L[f(x)]. \quad (4)$$

Соответственно двумерные системы преобразуют функции двух переменных:

$$g(x, y) = L[f(x, y)]. \quad (5)$$

Оптические системы по сути своей являются двумерными, но в некоторых случаях могут рассматриваться как одномерные.

Особое место среди всевозможных систем занимают линейные системы. Система называется линейной, если для нее справедлив принцип суперпозиции (наложения), который заключается в том, что отклик системы на взвешенную сумму двух входных воздействий равен взвешенной сумме откликов на каждое из воздействий, то есть

$$L[a_1 f_1(x, y) + a_2 f_2(x, y)] = a_1 L[f_1(x, y)] + a_2 L[f_2(x, y)]. \quad (6)$$

Принцип суперпозиций можно выразить в более общем виде, рассматривая произвольное число M входных воздействий:

$$L\left[\sum_{i=1}^M a_i f_i(x, y)\right] = \sum_{i=1}^M a_i L[f_i(x, y)]. \quad (7)$$

В изучении оптических систем фундаментальную роль играет понятие точечного источника света. Точечный источник света описывается дельта-функцией Дирака

$$\delta(x, y) = \begin{cases} \infty, & x = 0, y = 0, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (8)$$

Таким образом, точечный источник обладает бесконечно большой плотностью яркости в бесконечно малой пространственной области — в точке. Безусловно, это математическая абстракция, однако исключительно полезная в физике и допускающая ясную физическую трактовку: дельта-функция

может быть определена как предел обычной функции, например

$$\delta(x, y) = \lim_{a \rightarrow \infty} \{a^2 \exp[-a^2 \pi(x^2 + y^2)]\}. \quad (9)$$

Согласно выражению (9) дельта-функция может рассматриваться как бесконечно узкая колоколообразная функция (рис. 2).

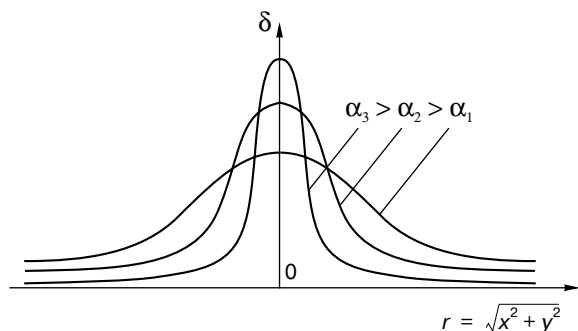


Рис. 2. Физическая трактовка дельта-функции Дирака.

Можно также ввести дельта-функцию, расположенную не в начале координат, а в произвольной точке с координатами (u, v) , по формуле

$$\delta(x - u, y - v) = \begin{cases} \infty, & u = 0, v = 0, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (10)$$

Дельта-функция обладает следующими важными свойствами:

1) Свойство нормировки:

$$\iint_D \delta(x, y) dx dy = 1. \quad (11)$$

Физически это означает, что, хотя плотность яркости точечного источника бесконечна, энергия его ограничена и равна единице.

2) Фильтрующее свойство

$$\iint_D f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta = f(x, y), \quad (12)$$

где $f(x, y)$ – произвольная функция двух переменных. Интегралы в (11) и (12) берутся по бесконечно большой пространственной области D . Доказательства свойств 1) и 2) выполняются с помощью подстановки в (11) и (12) выражения (9) и раскрытия предела.

Рассмотрим $2D$ -линейную систему, на вход которой подан сигнал в виде дельта-функции. Реакция системы на дельта-функцию будет различной для различных систем, называемых импульсным откликом и служит характеристикой $2D$ -системы. Систему называют пространственно-инвариантной,

если ее импульсный отклик зависит от разности координат входной (x, y) и выходной (ξ, η) плоскостей. Для оптической системы, показанной на рис. 3, это означает, что при перемещении точечного источника во входной (предметной) области изображение этого предмета в плоскости наблюдения будет также изменять положение, но сохранять форму.

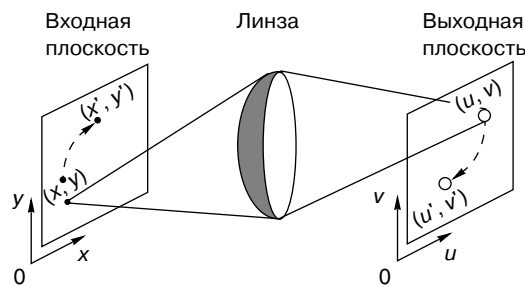


Рис. 3. Оптическая пространственно-инвариантная система.

Для пространственно-инвариантных систем импульсный отклик описывается функцией

$$h(x - u, y - v) \equiv h(\xi, \eta), \quad (13)$$

где $\xi = x - u, \eta = y - v$,

$$h(\xi, \eta) \equiv L[\delta(x, y)]. \quad (14)$$

Используя функцию импульсного отклика, можно записать уравнение, связывающее изображение на входе и выходе $2D$ -линейной оптической системы. Для этого представим входной сигнал $f(x, y)$ в виде (12) и подадим его на вход $2D$ -системы с характеристикой $h(\xi, \eta)$. Выходной сигнал запишем в виде

$$g(x, y) = L[f(x, y)] = L \left\{ \iint_D f(\xi, \eta) \delta(x - \xi, y - \eta) d\xi d\eta \right\}. \quad (15)$$

Поскольку операция L линейна и операция интегрирования в фигурных скобках (15) также линейна, их можно поменять местами и записать

$$g(x, y) = \iint_D f(\xi, \eta) L\{\delta(x - \xi, y - \eta)\} d\xi d\eta.$$

Учитывая, что по определению

$$L\{\delta(x - \xi, y - \eta)\} \equiv h(x - \xi, y - \eta),$$

окончательно получим выражение, устанавливающее связь между изображениями во входной и выходной плоскостях линейной системы:

$$g(x, y) = \iint_D f(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta. \quad (16)$$

Уравнение (16) называется интегралом свертки. Из этого уравнения следует, что, зная импульсный отклик оптической системы $h(\xi, \eta)$, можно рассчитать выходное изображение по входному.

Процесс свертки иллюстрирует рис. 4. На рис. 4а и 4б изображены функция $f(x, y)$ на входе и импульсный отклик. На рис. 4а показан импульсный отклик при обращении координат, а на рис. 4г — со сдвигом на величину x, y . На рис. 4д заштрихована область, в которой произведение $f(\xi, \eta)h(x-\xi, y-\eta)$, входящее в подынтегральное выражение (16), не равно нулю. Интегрирование по этой области дает величину $g(x, y)$ для заданных значений координат x, y . Таким образом, функция $g(x, y)$ на выходе может быть найдена сканированием входной функции скользящим “окном” — обращенным импульсным откликом, и интегрированием по области, в которой эти функции перекрываются.

1.3. Средства ввода изображений

Техническая задача, которую необходимо решить в компьютерной обработке изображений, — это ввод оптических изображений в память компьютера и вывод (визуализация) изображений. К счастью, в современных компьютерах задача визуализации решена. Для этих целей используются

высокоразрешающие цветные дисплеи и другая техника отображения информации.

Ввод изображений в память компьютера осуществляется с помощью видеодатчиков. Видеодатчик переводит оптическое распределение яркости изображения в электрические сигналы и далее в цифровые коды. Поскольку изображение является функцией двух пространственных переменных, а электрический сигнал является функцией одной переменной — времени, то для преобразования используется развертка. Например, при использовании телевизионной камеры изображение считывается по строкам: строка за строкой. При этом в пределах каждой строки зависимость яркости от пространственной координаты x преобразуется в пропорциональную зависимость амплитуды электрического сигнала от времени t . Переход от конца предыдущей строки к началу следующей осуществляется практически мгновенно. Широкое применение в качестве видеодатчиков находят также матрицы фотодиодов и матрицы приборов с зарядовой связью. При использовании матричных видеодатчиков изображение как бы наблюдается сквозь экран с множеством прозрачных ячеек. Число таких ячеек для современных видеодатчиков весьма велико и составляет величину 1024×1024 и более (см. рис. 5).

Исходное изображение, как уже отмечалось, представляет собой функцию двух непрерывных

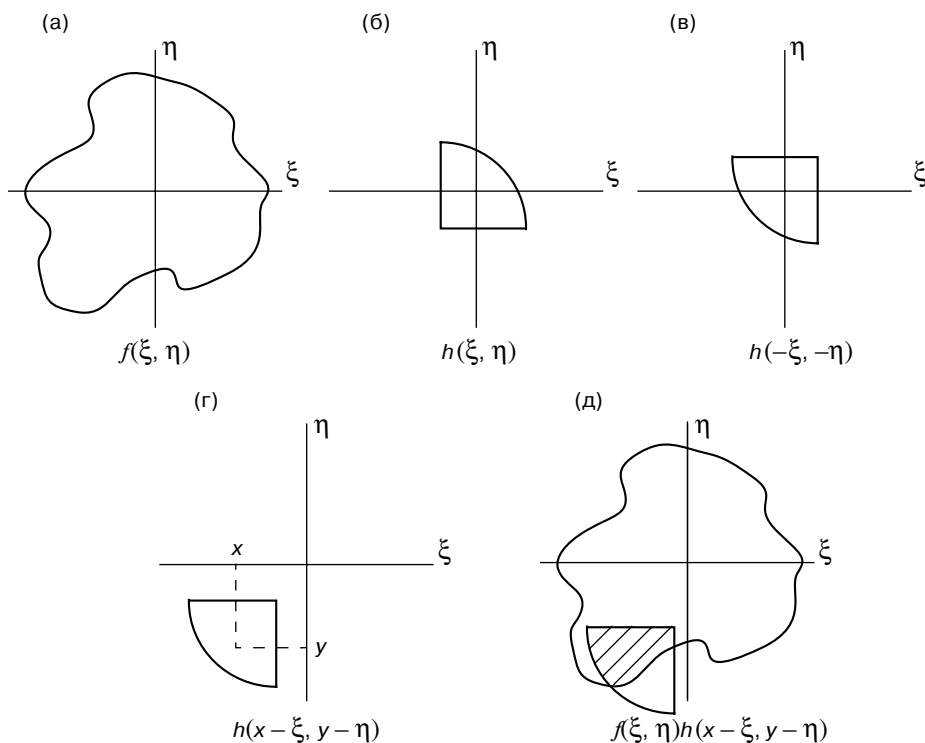


Рис. 4. Пример двумерной свертки.

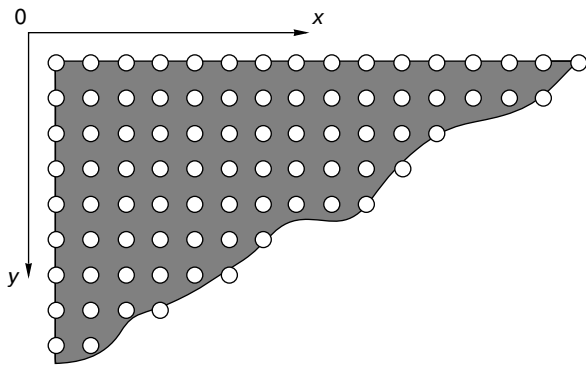


Рис. 5. Фрагмент матричного видеодатчика.

аргументов. В то же время цифровая память компьютера способна хранить только массивы данных. Поэтому ввод изображения в компьютер неизбежно связан с дискретизацией изображений по пространственным координатам и по яркости.

2. ДИСКРЕТНЫЕ ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ

2.1. Дискретизация изображений

Рассмотрим непрерывное изображение $f(x, y)$ – функцию двух пространственных переменных x и y на ограниченной прямоугольной области (рис. 6).

Введем понятие шага дискретизации T_1 по пространственной переменной x и T_2 по переменной y . Например, можно представить, что в точках, удаленных друг от друга на расстояние T_1 по оси x , расположены точечные видеодатчики. Если такие видеодатчики установить по всей прямоугольной области, то изображение окажется заданным на двумерной решетке:

$$f(n_1 T_1, n_2 T_2) = f(x, y)|_{x=n_1 T_1, y=n_2 T_2}. \quad (17)$$

Для сокращения записи обозначим

$$f(n_1 T_1, n_2 T_2) \equiv f(n_1, n_2). \quad (18)$$

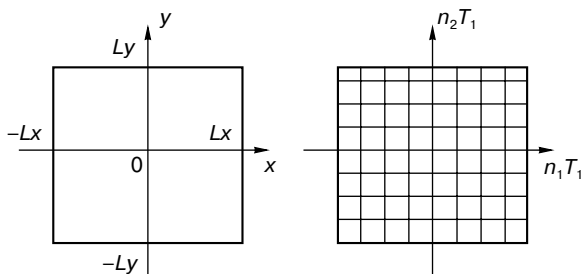


Рис. 6. Переход от непрерывного изображения к дискретному.

Функция $f(n_1, n_2)$ является функцией двух дискретных переменных и называется двумерной последовательностью. То есть дискретизация изображения по пространственным переменным переводит его в таблицу выборочных значений. Размерность таблицы (число строк и столбцов) определяется геометрическими размерами исходной прямоугольной области и выбором шага дискретизации по формуле

$$M_x = \left[\frac{2L_x}{T_1} \right], \quad M_y = \left[\frac{2L_y}{T_2} \right], \quad (19)$$

где $[\cdot]$ обозначает целую часть числа.

Если область определения непрерывного изображения – квадрат $L_x = L_y = L$ и шаг дискретизации выбран одинаковым по осям x и y ($T_1 = T_2 = T$), то

$$M_x = M_y = M \quad (20)$$

и размерность таблицы составляет M^2 .

Элемент таблицы, полученной путем дискретизации изображения, называют пиксел. Рассмотрим пиксел $f(n_1, n_2)$. Это число принимает непрерывные значения.

Память компьютера способна хранить только дискретные числа. Поэтому для записи в памяти непрерывная величина f должна быть подвергнута аналогово-цифровому преобразованию с шагом Δ (см. рис. 7).

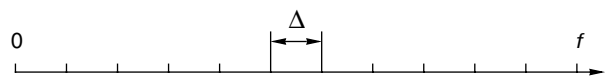


Рис. 7. Квантование непрерывной величины.

Операцию дискретизации непрерывной величины по уровням часто называют квантованием. Число уровней квантования равно

$$K = \left[\frac{A}{\Delta} \right]. \quad (21)$$

В практических задачах обработки изображений величина K варьируется в широких пределах от $K=2$ (“бинарные” (черно-белые) изображения) до $K=2^{10}$ и более (практически непрерывные значения яркости). Наиболее часто выбираются $K=2^8$, при этом пиксел изображения кодируется одним байтом информации. Из всего вышеуказанного делаем вывод, что пиксели, хранящиеся в памяти компьютера, представляют собой результат дискретизации исходного непрерывного изображения по аргументам и по уровням. Ясно, что шаги дискретизации T_1, T_2 и Δ должны выбираться достаточно малыми, для того чтобы погрешность дискретизации была незначительна и цифровое представление сохраняло основную информацию об изображении.

При этом следует помнить, что чем меньше шаг дискретизации и квантования, тем больший объем

данных об изображении должен быть записан в память компьютера. Рассмотрим в качестве иллюстрации этого утверждения изображение на слайде размером 50×50 мм, которое вводится в память с помощью цифрового измерителя оптической плотности (микроденситометра). Если при вводе линейное разрешение микроденситометра (шаг дискретизации по пространственным переменным) составляет 100 мкм, то в память записывается двумерный массив пикселей размерности $M^2 = 500 \times 500 = 25 \times 10^4$. Если же шаг уменьшить до 25 мкм, то размеры массива возрастут в 16 раз и составят $M^2 = 2000 \times 2000 = 4 \times 10^6$. Используя квантование по 256 уровням, то есть кодируя найденный пиксел байтом, получаем, что в первом случае для записи необходим объем 0,25 мегабайт памяти, а во втором случае – 4 мегабайта.

С физической точки зрения выбор шага дискретизации диктуется шириной пространственного спектра изображения. Чем больше ширина спектра Ω , тем меньше шаг дискретизации T . Практически при дискретизации стремятся удовлетворить соотношению

$$T \ll \frac{2\pi}{\Omega}. \quad (22)$$

2.2. 2D-последовательности

Рассмотрим несколько практически важных 2D-последовательностей, имеющих аналитическое выражение.

1) Цифровой единичный импульс

$$U_0(n_1, n_2) = \begin{cases} 1, & n_1 = n_2 = 0, \\ 0 & \text{при других } n_1 \text{ и } n_2. \end{cases} \quad (23)$$

Нетрудно заметить, что эта последовательность подобна дельта-функции (8). Произвольная последовательность $F(n_1, n_2)$ может быть представлена в виде

$$f(n_1, n_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f(k_1, k_2) u_0(n_1 - k_1, n_2 - k_2) \quad (8)$$

(сравним с формулой (12)).

2) Цифровой единичный скачок

$$u_{-1}(n_1, n_2) = \begin{cases} 1, & n_1, n_2 \geq 0, \\ 0, & n_1 \text{ или } n_2 < 0, \end{cases} \quad (25)$$

– функция, которая принимает единичные значения в правом верхнем квадранте координатной плоскости и нулевое значение в других квадрантах.

3) Экспоненциальная последовательность

$$e(n_1, n_2) = a_1^{n_1} a_2^{n_2}. \quad (9)$$

4) Комплексная экспонента

$$\exp(n_1, n_2) = \exp[i(\omega_1 n_1 + \omega_2 n_2)], \quad (27)$$

где ω_1, ω_2 имеют смысл пространственных частот.

2.3. 2D-системы

С математической точки зрения, 2D-система – это правило, которое ставит в соответствие 2D-входной последовательности $f(n_1, n_2)$ 2D-выходную последовательность $g(n_1, n_2)$.

Напомним, что мы рассматриваем линейные пространственно-инвариантные системы. Подавая на вход системы функцию $u_0(n_1, n_2)$, на выходе получаем функцию $h(n_1, n_2)$, которая называется импульсной реакцией системы.

Импульсная реакция позволяет записать связь между входной и выходной двумерными последовательностями системы в виде

$$g(n_1, n_2) = \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} h(m_1, m_2) f(n_1 - m_1, n_2 - m_2) \quad (28)$$

(сравним с (16)).

Формула 2D-свертки имеет большую вычислительную сложность. Для иллюстрации рассмотрим

Пример. Дана система с импульсной реакцией

$$h(n_1, n_2) = a^{n_1 n_2}, \quad -\infty \leq n_1, n_2 \leq \infty.$$

Входная последовательность имеет вид

$$f(n_1, n_2) = \begin{cases} 1, & 0 \leq n_1, n_2 \leq 2, \\ 0 & \text{при других } n_1, n_2. \end{cases}$$

Необходимо рассчитать последовательность $g(n_1, n_2)$ на выходе этой системы.

Используя формулу (28), получим

$$g(n_1, n_2) = \sum_{m_1=0}^2 \sum_{m_2=0}^2 a^{(n_1-m_1)(n_2-m_2)}.$$

Выполняя суммирование, получим

$$g(n_1, n_2) = (a^{n_1 n_2} + a^{n_1(n_2-1)} + a^{n_1(n_2-2)} + a^{(n_1-1)n_2} +) \\ + a^{(n_1-1)(n_2-1)} + a^{(n_1-1)(n_2-2)} + a^{(n_1-2)n_2} + a^{(n_1-2)(n_1-1)} + \\ + a^{(n_1-2)(n_2-2)}.$$

Сложность вычисления 2D-сверток даже в простых случаях дает представление о вычислительных трудностях, с которыми приходится сталкиваться при работе с 2D-системами.

Из рассмотренного выше примера видно, что, если входной сигнал (или импульсная реакция) имеет ограниченную протяженность, бесконечная сумма (28) в выражении двумерной свертки переходит в конечную:

$$g(n_1, n_2) = \sum_{m_1=0}^{M-1} \sum_{m_2=0}^{M-1} h(m_1, m_2) f(n_1 - m_1, n_2 - m_2). \quad (29)$$

Из формулы (29) видно, что для вычисления одного пиксела на выходе $2D$ -системы следует выполнить $\approx M^2$ арифметических операций.

В вычислительной математике разработаны так называемые алгоритмы быстрых сверток, которые позволяют сократить это число до $\approx M \log_2 M$ операций.

ЛИТЕРАТУРА

1. Прэтт У. Цифровая обработка изображений. В двух книгах. М.: Мир, 1982.

* * *

Виктор Александрович Сойфер, профессор, ректор Самарского государственного аэрокосми-

ческого университета, специалист в области обработки изображений и компьютерной оптики. Им опубликовано 310 научных работ, в том числе 3 монографии, 50 работ опубликовано в зарубежных изданиях.

В.А. Сойфер является лауреатом Государственной премии России в области науки и техники 1992 года за разработку лазерных технологий и их внедрение при создании новой авиационно-космической техники, в 1993 году награжден первой премией Германского общества содействия прикладной информатике за лучшую научную работу в области обработки изображений и распознавания образов. Член международного оптического общества SPIE.