
А. Розенфельд

РАСПОЗНАВАНИЕ
И ОБРАБОТКА
ИЗОБРАЖЕНИЙ

ИЗДАТЕЛЬСТВО «МИР»

А Е

И О

А. РОЗЕНФЕЛЬД

Picture Processing by Computer

AZRIEL ROZENFELD
*Computer Science Center
University of Maryland
College Park, Maryland*

Распознавание и обработка изображений с помощью вычислительных машин

ПЕРЕВОД С АНГЛИЙСКОГО
*И. Б. Гуревича, В. М. Рудакова
и Л. П. Ярославского*

ПОД РЕДАКЦИЕЙ
*канд. техн. наук
Д. С. Лебедева*

Academic Press New York London 1969

ИЗДАТЕЛЬСТВО «МИР» МОСКВА 1972

ПРЕДИСЛОВИЕ РЕДАКТОРА

Книга представляет собой первую монографию по методам анализа и преобразований изображений при помощи вычислительных машин — этому новому перспективному направлению в науке. В ней изложены способы кодирования и аппроксимации изображений, вопросы пространственной фильтрации (в том числе с помощью оптических устройств) и ее применения в целях улучшения качества изображений, выделения объектов на изображениях, распознавания зрительных образов и составления описания изображений.

Книга предназначена для инженеров, научных сотрудников, аспирантов и студентов старших курсов, которым приходится сталкиваться с автоматизацией обработки результатов физического эксперимента и микробиологических исследований, а также с анализом аэрокосмических изображений и рентгенограмм.

Редакция литературы по новой технике

Во многих областях науки, в технике и медицине все чаще приходится сталкиваться с необходимостью извлечения информации об исследуемых объектах посредством анализа их изображений или в более общем случае связанных с ними полей. Достаточно упомянуть об изображениях, получаемых с помощью искровых или пузырьковых камер, интерферограммах, аэрокосмических снимках, рентгенограммах, изображениях микрообъектов, зарегистрированных температурных полях и многом другом. Понятно, что автоматизация процедуры анализа изображений и полей приобретает неотложную актуальность, поскольку решение этой задачи позволит повысить скорость обработки данных и точность получаемых результатов.

Большие надежды в этом отношении возлагаются на вычислительные устройства и в первую очередь на ЦВМ. Помимо автоматического анализа, ЦВМ с успехом можно применять и для решения многих других задач обработки информации, содержащейся в изображениях: преобразования изображений в целях улучшения их качества или упрощения их интерпретации человеком, моделирования методов и систем кодирования и декодирования изображений, моделирования оптических, в том числе голографических, процессов.

Надо упомянуть и о возможностях синтеза изображений с помощью ЦВМ — создании кинофильмов, визуализирующих те или иные физические процессы, синтезе голограмм и тестовых изображений.

Хотя перечисленным вопросам посвящена обширная периодическая литература, труды многочисленных симпозиумов и конференций, монографий, обобщающих накопленный опыт, до сих пор не было на мировом книжном рынке. Книга профессора А. Розенфельда «Распозна-

вание и обработка изображений с помощью вычислительных машин» — одна из первых работ подобного рода. В ней рассмотрены вопросы кодирования и аппроксимации изображений, пространственной фильтрации (в том числе с помощью оптических устройств), выделения объектов на изображениях, распознавания зрительных образов и составления описания изображений.

К сожалению, не все вопросы, связанные с анализом и обработкой изображений при помощи ЦВМ, нашли отражение в книге. Не рассмотрены, например, задачи анализа волновых полей (голограмм, интерферограмм и т. д.) и измерения параметров объектов. Полностью отсутствуют задачи синтеза изображений машинной графики и применения ЦВМ в полиграфии. Автор ограничился изучением только двумерных одноцветных изображений. Поэтому из рассмотрения исключены вопросы анализа изображений трехмерных тел (например, автоматическая стереофотограмметрия), задачи, связанные с изменением изображений во времени, и анализ цветных изображений. В книге отсутствуют теоретические обоснования выбора того или иного алгоритма обработки изображений. Правда, освещение всех этих проблем лишило бы книгу ее достоинств — небольшого объема и доступности изложения, а эти достоинства вместе с обширным списком литературы делают настоящую книгу весьма интересной для многих советских читателей и особенно ценной для студентов и инженеров, приступающих к изучению методов обработки изображений. Наряду с этим она несомненно окажется полезной и для специалистов, уже знакомых с предметом.

При переводе были дополнительно включены ссылки на иностранные монографические работы, опубликованные в 1970—1971 гг., и на важнейшие работы советских ученых. Эти ссылки отмечены звездочкой.

Книгу перевели И. Б. Гуревич (гл. 7 и 8), В. М. Рудakov (гл. 1, 9 и 10) и Л. П. Ярославский (гл. 2—6).

Переводчики и редактор выражают признательность автору книги, приславшему поправки к книге и предисловие к русскому изданию.

Д. С. Лебедев

ПРЕДИСЛОВИЕ АВТОРА К РУССКОМУ ИЗДАНИЮ

Мне доставляет удовольствие тот факт, что книга «Распознавание и обработка изображений с помощью вычислительных машин» выходит в русском переводе в Советском Союзе, где проводятся блестящие работы в этом направлении.

Выражаю надежду, что книга принесет пользу как исследователям, имеющим дело с обработкой изображений, так и студентам, специализирующимся в этой области.

Методы обработки изображений развиваются столь быстро, что сейчас следовало бы пополнить книгу многими новыми разделами, посвященными, например, вопросам оптимального квантования, преобразованию Адамара, выпуклым сфероидальным волновым функциям, теории оптимальной линейной фильтрации и оптимального обнаружения контуров, математическим результатам по связности, сужению линий и многим другим. К сожалению, сейчас нереально пытаться внести так много существенных дополнений. Тем не менее пользуюсь представившейся возможностью исправить некоторые мелкие ошибки, вкравшиеся в английское издание.

Азриэль Розенфельд

Изображения и обработка изображений

В этой главе определим «изображение» как некий математический объект и установим правомочность применения к изображениям математических операций, описанных подробно в последующих главах, доказав, что функции, представляющие изображения, «неотличимы» от функций, имеющих достаточно «хорошие» аналитические свойства. Мы также кратко познакомимся с типами задач по обработке изображений, которые будут рассмотрены в оставшейся части книги.

1.1. Изображения

Термин «изображение» обозначает плоский объект, вид которого изменяется от точки к точке. В черно-белом изображении это изменение может быть описано с помощью единственного параметра, соответствующего общему количеству света, попадающему из данной точки к наблюдателю; для цветных изображений требуется несколько параметров. Поскольку при исследовании вопросов, касающихся обработки изображений, рассматриваются почти исключительно черно-белые изображения, будем считать отныне, что учитывается лишь один параметр.

В соответствии со сделанными выше замечаниями можно формально определить изображение как действительную функцию двух действительных переменных. Поскольку реальное изображение имеет ограниченные размеры, можно предположить, что эта функция не обращается в нуль только в ограниченной области, которая имеет некоторые стандартные размеры и форму, например квадрат. В действительности разумно допустить, что существует верхняя граница возможных размеров физически реализуе-

мых изображений, так что предполагаем, что любое изображение равно нулю вне квадрата заданных размеров.

Значение функции, представляющей изображение в некоторой точке, назовем яркостью изображения в этой точке. Так как количество света от данного физического изображения, достигающее наблюдателя, конечно и неотрицательно, можно положить, что любая функция, представляющая изображение, ограничена и неотрицательна. В действительности здесь также разумно допустить, что существует верхняя граница M величины возможной яркости реального изображения, так что для любого изображения f имеем $0 \leq f(x, y) \leq M$ для всех (x, y) .

При обработке изображений иногда оказывается необходимым проведение некоторых математических операций, например применение к ним преобразования Фурье. Чтобы гарантировать проведение подобных операций, необходимо предположить, что функции, представляющие изображения, являются в некотором смысле «хорошими», т. е. обладают такими качествами, как, например, интегрируемость, непрерывность и т. д.

При обработке изображений на цифровой вычислительной машине обычно необходимо представить их дискретным массивом чисел, т. е. в виде матриц, а не функций. Действительно, любую матрицу (имеющую вещественные неотрицательные элементы) можно считать представлением кусочно-постоянного изображения. Например, пусть для каждой (i, j) , где $1 \leq i \leq n$ и $1 \leq j \leq n$, S_{ij} является квадратом, определяемым неравенствами $i - 1 \leq x < i$ и $j - 1 \leq y < j$; тогда любой матрице (a_{ij}) порядка $n \times n$ можно сопоставить изображение, значение яркости которого в квадрате S_{ij} равно постоянной величине a_{ij} для каждой пары (i, j) . Таким образом построенное изображение будет называться *дискретным изображением*¹. Часто будем говорить о матрице (a_{ij}) как о существующем (а не только определенном) дискретном изображении. Легко показать (см. ниже), что любое изоб-

¹ Выбор координат в этом примере носит иллюстративный характер; кусочно-постоянное изображение при любом выборе координат также желательно называть дискретным изображением.

ражение неотлично от дискретного изображения порядка $n \times n$ для достаточно больших значений n .

Часто оказывается желательным предположить (особенно при цифровой обработке изображения), что функция, представляющая изображение, может принимать значения только из конечного набора величин или, другими словами, что яркость квантуется. Легко показать, что любое изображение неотлично от квантованного изображения, если только допускается достаточно много уровней квантования.

Частным случаем является двоичное изображение, которое может принимать только два значения (черное и белое без промежуточных серых тонов), обычно обозначаемых как 0 и 1.

Ниже будет объяснено сделанное выше утверждение о том, что произвольная функция, представляющая изображение, неотличима от функций тех или иных специальных типов. Пусть \mathcal{S} — некоторое множество функций. Будем говорить, что функция f неотличима от функции из множества \mathcal{S} , если для любых $\delta > 0$ и $\varepsilon > 0$ существует в \mathcal{S} такой элемент g , что $|f(x, y) - g(x, y)| < \varepsilon$ для всех (x, y) , за исключением множества меры, меньшей δ .

Теперь естественно предположить, что любая функция, представляющая изображение, интегрируема, поскольку физически интегрирование соответствует измерению общего количества света, пришедшего от изображения к наблюдателю. Пусть A обозначает квадрат, на границе и вне которого имеем $f = 0$. Так как f интегрируема в A , для любых $\delta > 0$ и $\varepsilon > 0$ существует такая абсолютно непрерывная в A функция g , что $|f - g| < \varepsilon$ всюду в A , за исключением множества меры, меньшей δ ; можно также положить, что $g = 0$ вне A (см., например, Titchmarsh E. C., The Theory of Functions, Oxford Univ. Press, 1939, p. 376; обобщение на случай функций двух переменных производится непосредственно). Итак, *любая функция, представляющая изображение, неотличима от абсолютно непрерывной функции*.

Можно, следовательно, допустить «по неотличимости», что в данном случае непрерывность f означает равномерную непрерывность на замыкании A . Предположим, что квадрат A разделен на n^2 квадратов $A_{ij}^{(n)}$, определенных

аналогично S_{ij} . Тогда для любого $\varepsilon > 0$ можно найти такое n , что в каждом квадрате $A_{ij}^{(n)}$ значение функции нигде не отличается более чем на ε от ее значения в центре квадрата $A_{ij}^{(n)}$. Следовательно, если $f^{(n)}$ — дискретное изображение, имеющее это (постоянное) значение яркости в каждом квадрате $A_{ij}^{(n)}$, то имеем $|f - g| < \varepsilon$ в каждой точке квадрата A , так что *любое изображение неотлично от дискретного изображения*.

Наконец, для любого $\varepsilon > 0$ можно разделить область $[0, M]$ на интервалы $[a_0, a_1), [a_1, a_2), \dots, [a_{N-1}, a_N]$, длина каждого из которых меньше 2ε (здесь $a_0 = 0$ и $a_N = M$). По заданному изображению f определим функцию f_N , которая принимает значения a_k в точке (x, y) , где a_k — граничная точка интервала, ближайшая к $f(x, y)$. Следовательно, $|f_N - f| < \varepsilon$ в любой точке, и это доказывает, что *любое изображение неотлично от квантованного изображения*.

Упражнение. Можно ли утверждать, что любое изображение неотлично от двоичного изображения? Если нет, то почему тогда на фотографиях и растрованных изображениях (в полиграфии) возможна приемлемая передача серых тонов? (Краткое пояснение: мы видим не само изображение, а лишь его размытую копию (см. разд. 6.3); если зерна фотографической эмульсии или растровые точки достаточно малы, размазывание усредняет их по группам.)

Можно ли доказать, что любое изображение неотлично от размытого двоичного изображения?

1.2. Задачи обработки изображений

Несмотря на то что изображения были ранее определены как весьма общие математические объекты, не следует, разумеется, рассматривать любую математическую операцию, проводимую над изображением, как «обработку изображения». Например, любая квадратная матрица с действительными неотрицательными элементами представляет дискретное изображение, однако отсюда не следует, что обращение матрицы или вычисление ее собственных чисел является обработкой изображения.

То, что составляет собственно предмет обработки изображений, имеет отношение к изображениям не просто как к произвольным функциям или матрицам, а как к изображению чего-то конкретного, что должно отображать реальные явления и предметы (местность, предметное стекло микроскопа и т. д.) или абстрактный символ (например, буквенно-цифровые знаки). Именно эта возможность отображения, которая и ставит главные задачи обработки изображений, рассматривается нами в настоящей книге.

Кодирование и аппроксимация

Так как изображения не являются произвольными матрицами, они встречаются не все одинаково часто (существует, разумеется, случай, когда обрабатываемые изображения принадлежат к одному типу, например печатный материал, аэрофотографии и т. д.).

С точки зрения теории информации это значит, что количество информации в некотором изображении значительно меньше в среднем, чем в том случае, когда все возможные изображения равновероятны. Следовательно, представляет интерес измерение количества информации, содержащейся в изображении, и разработка методов кодирования для представления изображения в наиболее компактном виде; такие методы будут рассмотрены в гл. 2. Можно также исследовать возможность допустимой аппроксимации данного изображения другим изображением, содержащим меньшее количество информации, причем критерий «допустимости» может быть как объективным, так и субъективным. Большинство работ в этой области, обзор которых сделан в гл. 3, посвящено проблеме сжатия телевизионной полосы частот.

Фильтрация, восстановление и улучшение

Существует много полезных типов операций над изображениями, являющихся пространственно-инвариантными, т. е. таких, результат применения которых в данной точке не зависит от положения этой точки на изображении. Подобные операции имеют свойства, аналогичные

свойствам инвариантных во времени операций в электронике (гл. 4). Они могут быть осуществлены с помощью не только обычных вычислительных устройств, но и ряда других простых методов (гл. 5), например оптических. Операции такого типа можно использовать для «фильтрации» изображения, с тем чтобы обнаружить на нем некоторую заданную фигуру, для коррекции изображения, искаженного при применении других операций, и для «сглаживания» или «обострения» изображения в целях повышения его «качества».

Распознавание и описание изображения

При исследовании изображения наибольший интерес часто представляет получение описания того, что на нем представлено; в этом и заключается проблема распознавания зрительных образов. Это описание можно построить просто в отнесении изображения к одному из нескольких заранее указанных классов, измеряя различные параметры изображения в целом (гл. 7). С другой стороны, в это описание могут входить свойства представленных на изображении объектов и связи между ними. Для получения такого описания обычно необходимо четко выделить объекты на изображениях (см. гл. 8) и дать оценку их свойств (например, топологии, размера, формы и т. д.; см. гл. 9), а также их взаимосвязей.

Описание изображения с использованием терминов «объекты», «свойства», «взаимосвязь» может быть составлено на специальном «языке изображений» (гл. 10).

1.3. Общая литература

Следующие сборники статей неоднократно цитируются (для краткости они обозначены первыми буквами):

BPSS	«Biological Prototypes and Synthetic Systems» (Bernard E. E., Kame M. R., eds.), Plenum Press, New York, 1962.
OCR	«Optical Character Recognition» (Fischer G. L., Jr., Pollock D. K., Radack B., Stevens M. E., eds.), Spartan, New York, 1962.
OPI	«Optical Processing of Information» (Pollock D. K., Koester C. J., Tippet J. T., eds.), Spartan, New York, 1963.

OEOIP	«Optical and Electro-Optical Information Processing» (Tippet J. T., Berkowitz D. A., Clapp L. C., Koester C. J., Vanderburgh A., Jr., eds.), M.I.T. Press, Cambridge, Massachusetts, 1965.
PR	«Pattern Recognition» (Uhr L., ed), Wiley, New York, 1966.
MPSVF	«Models for the Perception of Speech and Visual Form» (Wathen-Dunn W., ed.), M.I.T. Press, Cambridge, Massachusetts, 1967.
PR ²	«Pattern Recognition» (Kanai L., ed.), Thompson, Washington, D. C., 1968.
PPR	«Pictorial Pattern Recognition» (Cheng G. C., Ledley R. S., Pollock D. K., Rosenfeld A., eds.), Thompson, Washington, D.C., 1968.
CPR	«Conference on Pattern Recognition» (Confer. Publ. № 42), Inst. Elec. Engrs., London, 1968. Proc. Eastern (Western, Fall, Spring), Joint Comput. Confer. будем сокращенно обозначать E(W, F, S)JCC.

- 1*. «Иконика», сб. статей, изд-во «Наука», М., 1968.
- 2*. «Иконика. Пространственная фильтрация изображений. Фотографические системы», сб. статей, изд-во «Наука», М., 1970.
- 3*. Ярославский Л. П., Устройства ввода — вывода изображений для цифровых вычислительных машин, изд-во «Энергия», М., 1968.
- 4*. Andrews H., Computer Techniques in Image Processing, Academic Press, New York, 1970.
- 5*. Wilkins L. C., Wintz P. A., Bibliography on Data Compression, Picture Properties, and Picture Coding, *IEEE Trans. on Inf. Theory*, IT-17, № 2 (March 1971).
- 6*. Huang T. S., Schreiber W. F., Tretjak O. J., Image Processing, *Proc. IEEE*, 59, № 11 (Nov. 1971); русский перевод: Хуанг, Шрайбер, Третьяк, Обработка изображений, *ТИИЭР*, № 11 (1971).

Кодирование изображений

2.1. Вероятность и информация

Предположим, что от некоторого источника поступает последовательность сообщений, каждое из которых может принимать одно из m возможных значений a_1, a_2, \dots, a_m . Если нельзя точно предсказать появление того или иного сообщения, то можно по крайней мере найти вероятности сообщений, определив частоту их появления в последовательности. Если в среднем в последовательности из n сообщений сообщение a_i встретится n_i раз, говорят, что оно имеет вероятность $p_i = n_i/n$. Очевидно, для каждого i $0 \leq n_i \leq n$ и $n_1 + n_2 + \dots + n_m = n$, так что $0 \leq p_i \leq 1$ и $p_1 + \dots + p_m = 1$. Заметим также, что поскольку общее число сообщений a_i и a_j в среднем в последовательности из n сообщений равно $n_i + n_j$, то вероятность появления сообщений a_i или a_j равна $(n_i + n_j)/n = p_i + p_j$. Далее, пусть F — любая функция, ставящая в соответствие каждому сообщению некоторое действительное число. Тогда среднее значение F будет равно

$$[n_1 F(a_1) + \dots + n_m F(a_m)]/n = p_1 F(a_1) + \dots + p_m F(a_m).$$

Прием сообщений уменьшает неопределенность в том, какое из возможных сообщений выбрано. Количество информации, содержащейся в данном сообщении, можно определить как меру уменьшения этой неопределенности. При таком определении высоковероятные сообщения несут малое количество информации, а маловероятные сообщения — большое. Если вероятность p_i равна 1, сообщение не несет информации, если же p_i стремится к нулю, количество переносимой i -м сообщением информации неограниченно возрастает. Простой функцией p_i ,

удовлетворяющей этим интуитивным требованиям, является функция $k \log 1/p_i$ (или $-k \log p_i$), где k — положительная константа. Именно эта величина и используется в теории информации как мера «количества информации». Выбор k (или основания логарифма) определяет единицу измерения количества информации. Если логарифм берется по основанию b , то одна единица информации содержится в сообщении, вероятность которого равна $1/b$. Обычно в качестве единицы количества информации выбирают количество информации в сообщении, вероятность которого равна $1/2$, что соответствует использованию логарифмов по основанию 2. Такая единица называется битом¹.

Пример. Пусть сообщением является одно из чисел $0, 1, \dots, 2^r - 1$, имеющих вероятность $1/2^r$. Тогда количество информации в таком числе равно $\log_2 2^r = r$ бит. Иными словами, r — значное двоичное число — содержит r бит информации. Заметим, что количество информации в сообщении можно рассматривать как число вопросов, которые нужно задать, чтобы по ответам на каждый вопрос «да» или «нет» полностью определить сообщение. Так, для r -значного числа это могут быть вопросы типа: «Равна ли первая (вторая и т. д.) двоичная цифра 1?».

Если имеется m возможных сообщений с вероятностями p_1, \dots, p_m , количество информации, приходящееся в среднем на одно сообщение, можно измерить, положив введенную выше функцию F равной количеству информации в каждом сообщении. Это среднее количество информации² равно

$$-p_1 \log_2 p_1 - \dots - p_m \log_2 p_m, \text{ или } -\sum_{i=1}^m p_i \log_2 p_i$$

Упражнение 1. Найти среднее количество информации на сообщение в следующих случаях:

а) $m = 4; p_1 = p_2 = p_3 = p_4 = \frac{1}{4};$

¹ От слов binary digit — двоичная цифра. — Прим. ред.

² Эта величина в теории информации называется энтропией. — Прим. ред.

$$\text{б) } m=4; \quad p_1=\frac{1}{2}, \quad p_2=\frac{1}{4}, \quad p_3=p_4=\frac{1}{8};$$

$$\text{в) } m=5; \quad p_1=\frac{1}{2}, \quad p_2=p_3=p_4=p_5=\frac{1}{8}.$$

Упражнение 2. Доказать, что для двух сообщений среднее количество информации на сообщение максимально, если вероятности сообщений равны $1/2$. (Указание: записать вероятности сообщений как x и $1-x$ и найти условие максимума функции $-x \log_2 x - (1-x) \log_2 (1-x)$.)

Упражнение 3. Доказать, что для m сообщений среднее количество информации на сообщение максимально, если вероятности сообщений равны $1/m$. (Указание: используя метод множителей Лагранжа, найти максимум выражения $-\sum_{i=1}^m p_i \log_2 p_i$ как функции m переменных p_1, \dots, p_m при условии $p_1 + p_2 + \dots + p_m = 1$.) Заметим, что это максимальное количество информации на сообщение равно

$$\sum_{i=1}^m (1/m) \log_2 m = \log_2 m \text{ бит.}$$

Упражнение 4. Доказать, что если p_1 близко к 1 (то есть остальные вероятности p_2, \dots, p_m близки к 0), то среднее количество информации на одно сообщение близко к 0. (Указание: $-x \log_2 x \rightarrow 0$ при $x \rightarrow 0$ или $x \rightarrow 1$.)

Введенные определения можно применить к изображениям, если рассматривать дискретные изображения как множество сообщений и считать сообщениями значения яркости элементов изображения. Если число уровней квантования яркости равно m , то количество информации в изображении, состоящем из $n \times n$ элементов (равное среднему количеству информации на элемент, умноженному на число элементов), может достигать $n^2 \log_2 m$ бит (см. упражнение 3); точное значение количества информации зависит от распределения вероятностей значений яркости. Физическая среда, служащая для регистрации изображений, может использоваться для хранения

информации с очень большой плотностью [1, 2]. Однако количество информации в изображениях, встречающихся на практике (телевизионные изображения, штриховые рисунки, печатный текст и т. п.), значительно (в два и более раз) меньше предельной величины [3, 4]. Разница между потенциально возможным и действительным количеством информации называется избыточностью.

2.2 Эффективное кодирование одиночных сообщений

Пусть имеется m возможных сообщений a_1, \dots, a_m с вероятностями p_1, \dots, p_m . Если эти сообщения представить числами от 0 до $m-1$, то для каждого сообщения потребуется по крайней мере $\log_2 m$ двоичных цифр, даже если вероятности p_i таковы, что среднее количество информации на сообщение намного меньше этой величины. Однако, используя для представления сообщений специальный код, можно уменьшить количество двоичных цифр, приходящееся в среднем на сообщение; в некоторых случаях возможно уменьшение до величины среднего количества информации на сообщение. (Можно показать, что это теоретический предел уменьшения.)

Основной принцип построения таких кодов состоит в том, что часто встречающимся сообщениям соответствуют кодовые слова с меньшим числом двоичных цифр, чем редким сообщениям. Предположим, например, что имеется четыре сообщения, которые встречаются с вероятностями, приведенными в табл. I. Если использовать код из третьего столбца таблицы, то в среднем на сообщение потребуется

$$1(0,85) + 2(0,06) + 3(0,06) + 3(0,03) = 1,24$$

двоичной цифры, т. е. меньше двух двоичных цифр, которые необходимы при простой перенумерации сообщений числами 00, 01, 10, 11. (Заметим, что код, приведенный в табл. I, не требует разделительных знаков при записи

¹ В дальнейшем для краткости основание логарифма будет опускаться.

Таблица I

Сообщение	Вероятность	Кодовое слово	Количество двоичных цифр в кодовом слове
a_1	$2^{-1/4} = 0,85$	1	1
a_2	$2^{-4} = 0,06$	01	2
a_3	$2^{-4} = 0,06$	001	3
a_4	$2^{-5} = 0,03$	000	3

последовательности сообщений.) Последовательность двоичных цифр, представляющая последовательность сообщений, может быть декодирована однозначно. Если бы этого свойства не было, код нужно было бы строить из трех символов (0, 1 и разделительного знака) и потребовалось бы значительно больше чем 1,24 двоичной цифры на сообщение.

В рассмотренном примере среднее число цифр на сообщение все еще превышает среднее количество информации, равное

$$\frac{1}{4}(0,84) + 8(0,06) + 5(0,03) = 0,84 \text{ бит.}$$

Иногда можно построить простой код, в котором число двоичных цифр на сообщение в среднем почти не будет превышать количество информации на сообщение. В этом смысле он будет близок к оптимальному.

Один из способов построения такого кода состоит в делении списка сообщений пополам на две группы, полученных групп еще раз пополам на подгруппы и так далее до тех пор, пока в каждой части не окажется только по одному сообщению, причем на каждом шаге деление необходимо производить так, чтобы суммарные вероятности получаемых частей были как можно более близки друг к другу. Сам код при этом строится таким образом. Если сообщение находится в первой половине разделяемого списка, первой цифрой кодового слова становится 0, если во второй, то 1. На втором шаге построения кода, если сообщение находится в первой подгруппе разделяемой части списка, второй цифре кодового слова присваи-

вается значение 0, если оно находится во второй, то 1 и так далее. Пример построенного по этому методу оптимального кода приведен в табл. II.

Таблица II

Сообщение	Вероятность	Кодовое слово
a_1	$\frac{1}{2}$	0
a_2	$\frac{1}{4}$	10
a_3	$\frac{1}{8}$	110
a_4	$\frac{1}{8}$	111

Для этих конкретных значений вероятностей число двоичных цифр в среднем на одно сообщение равно количеству информации на сообщение, т. е. 1,75 бит. Описанные коды известны как коды Шеннона — Фано — Хаффмана.

Упражнение 1. Обобщить разобранный в табл. II пример следующим образом. Пусть имеется m сообщений a_1, a_2, \dots, a_m с вероятностями соответственно $1/2; 1/4; 1/8, \dots, 1/2^{m-2}, 1/2^{m-1}, 1/2^{m-1}$. Построить с помощью описанной процедуры код, разделяя сообщения на подгруппы так, как показано в табл. III. (Поскольку на каждом

Таблица III

Шаг	Первая подгруппа	Вторая подгруппа
1	a_1	a_2, \dots, a_m
2	a_2	a_3, \dots, a_m
...
$m-1$	a_{m-1}	a_m

шаге первая подгруппа состоит только из одного сообщения, дальнейшему делению подлежит только вторая подгруппа.) Показать, что среднее число двоичных цифр на сообщение в полученном коде равно среднему количеству информации на сообщение, то есть $(2^m - 1)/2^{m-1}$ бит.

Упражнение 2. Показать, что если вероятности m сообщений отличаются от приведенных в упражнении 1, то любой код, построенный путем деления сообщений на подгруппы (как описано выше), требует больше $(2^m - 1)/2^{m-1}$ двоичных цифр в среднем на сообщение.

2.3. Эффективное кодирование последовательности сообщений

Если принято некоторое сообщение a_i , то вероятность того, что вслед за ним будет принято сообщение a_j , может отличаться от p_j , так как следующие друг за другом сообщения могут быть зависимыми. Обозначим условную вероятность того, что за сообщением a_i последует сообщение a_j , через p_{ij} . Если $p_{ij} = p_j$ при всех i и j , то говорят, что сообщения *независимы*. Очевидно, $p_{i1} + \dots + p_{im} = 1$ для каждого i , так как одно из всех возможных сообщений обязательно последует за a_i .

В последовательности из n пар сообщений число пар, начинающихся с сообщения a_i , в среднем равно $n_i = p_i n$. Среди них будет примерно $p_{ij} n_i = p_i p_{ij} n$ пар, в которых вторым сообщением является a_j . Таким образом, вероятность появления пары сообщений $a_i a_j$ равна $p_i p_{ij}$. Далее, число пар со вторым сообщением a_j в среднем равно

$$p_1 p_{1j} n + p_2 p_{2j} n + \dots + p_m p_{mj} n,$$

так что вероятность появления сообщения a_j на втором месте в паре сообщений равна $p_1 p_{1j} + \dots + p_m p_{mj}$. Но поскольку любое сообщение может рассматриваться как второе сообщение в паре, общее число появлений сообщения a_j приблизительно такое же, как и число появлений a_j на втором месте в парах сообщений. Отсюда следует, что в длинной последовательности сообщений для каждого j приблизительно выполняется равенство

$$p_j = p_1 p_{1j} + \dots + p_m p_{mj}.$$

Это соотношение позволяет вычислить p_j , зная p_{ij} . Например, при $m = 2$

$$p_1 = p_1 p_{11} + p_2 p_{21}$$

и, поскольку $p_2 = 1 - p_1$, а $p_{12} = 1 - p_{11}$, то

$$p_1 = \frac{p_{21}}{p_{12} + p_{21}},$$

$$p_2 = \frac{p_{12}}{p_{12} + p_{21}}.$$

Заметим, что эти рассуждения не верны, если $p_{12} = p_{21} = 0$. Но в этом случае исключается появление в паре второго сообщения, отличающегося от первого, так что имеем либо $p_1 = 1, p_2 = 0$, либо наоборот.

Упражнение 1. Выразить p_i через p_{ij} для произвольного m . В каких случаях нарушается справедливость полученного выражения?

Упражнение 2. Обобщить приведенные рассуждения на группы из k последовательных сообщений, используя условные вероятности более высоких порядков $p(a_j/a_{i_1}, \dots, a_{i_{k-1}})$, т. е. вероятности того, что в группе из k сообщений на k -м месте будет сообщение p_j , если на $k - 1$ предыдущих местах находятся сообщения $a_{i_1}, \dots, a_{i_{k-1}}$.

Рассмотрим три важных способа экономного кодирования последовательностей зависимых сообщений.

Блочное кодирование

Если следующие друг за другом сообщения независимы, то, как указано выше, вероятность любой пары сообщений (a_i, a_j) равна $p_i p_j$. Следовательно, количество информации на пару сообщений равно

$$\begin{aligned} & - p_1^2 \log p_1^2 - p_1 p_2 \log p_1 p_2 - \dots - p_1 p_m \log p_1 p_m \\ & - p_2 p_1 \log p_2 p_1 - p_2^2 \log p_2^2 - \dots - p_2 p_m \log p_2 p_m \\ & \vdots \\ & - p_m p_1 \log p_m p_1 - p_m p_2 \log p_m p_2 - \dots - p_m^2 \log p_m^2. \end{aligned}$$

Отсюда следует, что последовательность сообщений полностью определяется указанием первого сообщения и последовательностью разностей. Преимущество замены исходных сообщений разностными состоит в том, что если сообщения в исходной последовательности зависимы, то разностные сообщения будут содержать значительно меньшее количество информации на сообщение, чем исходные.

Чтобы пояснить эти соображения, рассмотрим снова случай передачи двух сообщений. Предположим, что используется простейшее правило предсказания, а именно считается, что каждое передаваемое сообщение совпадает с предыдущим. В этом случае вероятность того, что разность равна 0, равна просто вероятности появления сообщений (a_1, a_1) или (a_2, a_2) , а вероятность значения разности, равного единице, совпадает с вероятностью появления пар (a_1, a_2) или (a_2, a_1) . Используя вероятности пар из табл. IV, можно найти, что вероятности 0 и 1 в разностном сообщении равны соответственно

$$(p_{11}p_{21} + p_{12}p_{22}) / (p_{12} + p_{21})$$

и

$$2p_{12}p_{21} / (p_{12} + p_{21}).$$

Как и при анализе блочного кодирования, примем, что $p_{12} = p_{21} = p$. Тогда $p_1 = p_2 = 1/2$, а вероятности значений разностного сообщения равны $1 - p$ и p соответственно. Следовательно, количество информации на одно разностное сообщение может быть сколь угодно мало, даже если количество информации в исходных сообщениях близко к одному биту.

Упражнение 5. Обобщить этот пример.

Предположим в общем случае, что над исходной последовательностью сообщений произведена некоторая операция перестановки (например, как в предыдущем случае, операция «вычитания» соседних сообщений), причем существует обратная операция, с помощью которой исходная последовательность может быть восстановлена. Тогда можно, используя зависимость между исходными сообщениями, попытаться найти такую операцию, чтобы количество информации в полученных новых сообщениях было минимально.

Кодирование серий

В примере, приведенном в табл. IV ($p_{12} = p_{21} = p$), при малых p , т. е. когда с высокой вероятностью данное сообщение в последовательности повторяется, а не заменяется другим, среднее количество информации на парное или разностное сообщение мало. В этих случаях для последовательности сообщений характерна тенденция к сохранению длинных серий повторяющихся сообщений. Такие же серии могут появляться и в случае независимых сообщений, если вероятность одного из возможных сообщений намного больше суммарной вероятности других (как в первом примере разд. 2.2). В этом случае высоковероятное сообщение часто повторяется в последовательности, а остальные сообщения встречаются изолированно. Если эта тенденция проявляется достаточно сильно, может оказаться, что для более экономного кодирования нужно вместо кодирования каждого элемента последовательности передавать первое сообщение каждой серии и длину серии или первое сообщение серии и его позицию в последовательности сообщений. Очевидно, что эта информация полностью определяет передаваемую последовательность. Если имеется только два возможных сообщения, то достаточно даже указать первое сообщение в последовательности и размеры (или положение первых элементов) всех серий, так как серии двух сообщений должны сменять друг друга. Эти два способа кодирования называются соответственно *кодированием длин серий* и *кодированием концов серий*.

Чтобы проиллюстрировать возможность экономного кодирования последовательности сообщений с помощью кодирования серий, предположим, что имеется 2 возможных сообщения и что длина кодируемой последовательности равна n , а средняя длина серии $r > \log n$. Для полного описания кодируемой последовательности с помощью указания концов серий требуется тогда только около $n/r \log n$ двоичных цифр (а именно $\log n$ двоичных цифр для указания положения каждой из приблизительно n/r серий), т. е. меньше чем n двоичных цифр, требуемых для описания последовательности без кодирования.

Упражнение 6. Какова средняя длина серии в случае двух зависимых сообщений? В случае независимых сообщений?

Упражнение 7. Сравнить выигрыш, получаемый при блочном кодировании, кодировании с предсказанием, кодировании серий сообщений, описанных в табл. IV (где $p_{12} = p_{21} = p$), в зависимости от величины p .

2.4. Некоторые дополнительные замечания о кодировании

В принципе можно в процессе передачи изменять используемый код для согласования его со статистикой передаваемых сообщений. Пусть, например, имеется более двух возможных сообщений, которые могут быть представлены многозначными двоичными числами. Статистика значений разрядов этих чисел часто не одинакова (например, старшие разряды могут изменяться от сообщения к сообщению меньше, чем младшие). Поэтому может оказаться выгодным кодировать разряды по-разному. Этот способ кодирования называют поразрядным кодированием.

Упражнение. Построить пример уменьшения количества информации на сообщение с помощью поразрядного кодирования.

Если изображение содержит небольшое число простых по конфигурации областей с постоянными внутри каждой области значениями яркости, то оно может быть очень эффективно закодировано своим описанием, т. е. указанием геометрии расположения и яркости каждой области. Например, двоичное изображение, состоящее из $n \times n$ элементов, яркость которого в точках, составляющих некоторый прямоугольник, равна 1, а в остальных 0, можно полностью определить, задав по две координаты двух противоположных вершин прямоугольника¹. Для каждой из этих четырех координат требуется $\log n$ двоичных цифр. (Кодирование серий можно рассматривать как частный

случай такого кодирования в системах с обычной разверткой, когда «областями» с постоянным значением сообщения являются отрезки горизонтальных строк.)

Аналогичным способом часто можно компактно кодировать рисунки, в особенности если они состоят только из кривых постоянной толщины, описываемых простыми уравнениями (например, из прямых линий и конических сечений). Такой подход используется во всех системах отображения графической информации для вычислительных машин, так как он не только позволяет сократить описание изображений, но и значительно упрощает выполнение геометрических преобразований над ними.

ЛИТЕРАТУРА

1. Лебедев Д. С., Применение теории информации к фотографическим системам, *Журнал научной и прикладной фотографии*, 10, № 1, 62 (1965).
2. Jones R. C., Information capacities of radiation detectors and of light, *Appl. Opt.*, 2, 351 (Apr. 1963).
3. Deutsch S., A note on some statistics concerning typewritten or printed material, *IRE Trans. Information Theory*, IT-3, 147 (June 1957).
4. Foy W. H., Jr., Entropy of simple line drawings, *IEEE Trans. Information Theory*, IT-10, 165 (Apr. 1964).
5. Kretzmer E. R., Statistics of television signals, *Bell System Tech. J.*, 31, 751 (July 1952).
6. Schreider W. F., The measurement of third order probability distributions of television signals, *IRE Trans. Information Theory*, IT-2, 94 (Sept. 1956).
7. Wallace P. R., Real time measurement of element differences in television programs, *Proc. IEEE*, 54, 1576 (Nov. 1966); русский перевод: Уоллес, Измерение поэлементных разностей на телевизионном изображении в реальном масштабе времени, *ТИИЭР*, 54, № 11 (1966).
8. Nishikawa S., Massa R. J., Mott-Smith J. C., Area properties of television pictures, *IEEE Trans. Information Theory*, IT-11, 348 (July 1965).
9. Schwartz J. W., Barker R. C., Bit plane encoding: a technique for source encoding, *IEEE Trans. Aerospace Electronic Systems*, AES-2, 385 (July 1966).
- 10*. Лебедев Д. С., Цуккерман И. И., Телевидение и теория информации, изд-во «Энергия», 1965.

¹ Если известна ориентация прямоугольника. — Прим. ред.

Аппроксимация изображений

Во многих случаях нет необходимости воспроизводить все детали изображений. Если с точки зрения наблюдателя два разных изображения достаточно сходны друг с другом, их можно считать идентичными. Более того, часто можно заменить данное изображение сходным с ним более простым изображением, для воспроизведения которого требуется меньшее количество информации. Степень сжатия, или сокращения избыточности, достигаемая при использовании методов аппроксимации, обычно больше той, которую можно получить при статистическом кодировании. Правда, аппроксимация может привести к ухудшению качества изображения, тогда как при статистическом кодировании изображение не искажается.

В данной главе рассмотрены методы сокращения избыточности изображений с помощью их аппроксимации. К изображению можно применить любой из известных методов аппроксимации функций. Кроме того, существуют специфические методы аппроксимации, основанные на использовании особенностей зрительного восприятия изображений. Эти методы позволяют построить аппроксимацию изображений, приемлемую с точки зрения наблюдателя и в то же время существенно более экономную в информационном смысле. Работы в этом направлении связаны в основном с проблемой сокращения полосы телевизионных сигналов¹

¹ Надо указать, что в телевидении богатые возможности сжатия полосы основаны на том, что последовательные кадры, как правило, сравнительно мало отличаются друг от друга; однако подобная возможность здесь не рассматривается.

Двумя основными средствами, которыми можно воспользоваться для аппроксимации функции, являются *дискретизация* и *квантование*. Дискретизация заключается в замене непрерывной функции ее отсчетами — значениями в конечном множестве значений аргумента. При этом исходная функция аппроксимируется путем интерполяции ее значений между отсчетами с помощью простых функций. Квантование состоит в том, что значения функции в каждой точке заменяются ближайшими к ним величинами, выбираемыми из конечного множества значений (уровней квантования). Как показано в последующих разделах, положение отсчетов и (или) уровни квантования могут задаваться заранее или выбираются в зависимости от характера функции.

Другим методом аппроксимации функции, который может оказаться полезным и при аппроксимации изображений, является разложение функции в некоторый ряд, например ряд Тэйлора, Фурье (или в общем случае в ряд по данному множеству ортогональных функций), разложение матрицы на сумму диад и т. д., и использование в качестве аппроксимации суммы нескольких первых членов ряда.

3.1. Дискретизация с фиксированным положением отсчетов

Простейший класс методов дискретизации составляют методы дискретизации с фиксированным расположением отсчетов. Обычно проще всего располагать эти отсчеты через одинаковые интервалы, так чтобы они образовывали регулярную решетку. Однако иногда желательно выбирать неодинаковые интервалы между отсчетами. Например, если вследствие перспективных искажений масштаб на фотографии изменяется, удобнее брать отсчеты с переменным шагом, расположенные равномерно на исходном объекте.

Интерполяция полиномами

Для аппроксимации значений изображения по его отсчетам можно выбирать различные функции. Простейшей пригодной для этой цели функцией является функ-

ция, постоянная на данном интервале. В этом случае изображение аппроксимируется участками, содержащими один отсчет, и в пределах каждого такого участка его значение выбирается равным отсчетному значению. Заметим, что если отсчетные точки образуют квадратную решетку из $n \times n$ элементов и в качестве участков выбираются ячейки этой решетки S_{ij} (разд. 1.1), то аппроксимированное изображение представляет собой просто дискретное изображение из $n \times n$ элементов, определяемое матрицей своих значений в отсчетных точках. Таким образом, введенное в гл. 1 дискретное изображение можно рассматривать как интерполяционное приближение нулевого порядка.

В общем случае интерполяцию значений изображения между отсчетами можно производить участками плоскостей или поверхностей более высокого порядка. Например, если точки отсчетов образуют треугольную решетку, то через каждые три отсчета можно провести плоскости и таким образом построить многогранную поверхность с треугольными ячейками, проходящую через все отсчеты изображения. Если вместо плоскостей для интерполяции использовать поверхности более высокого порядка, можно потребовать, чтобы каждый участок такой интерполирующей поверхности проходил через большее число отсчетов. Для того чтобы результирующая поверхность была не только непрерывной, но и гладкой, можно построить интерполирующие поверхности так, чтобы аппроксимировались не только значения изображения, но и значения его производных вдоль кривых пересечения поверхностей.

Синусоидальная интерполяция. Теорема отсчетов

Интерполирующие функции не обязательно должны являться полиномами. Можно пользоваться также семейством трансцендентных функций, например экспонент, синусоид и т. д. Синусоидальная интерполяция представляет особый интерес ввиду того, что, согласно так называемой теореме отсчетов, при некоторых условиях этот способ интерполяции позволяет точно воспроизвести

исходную функцию по ее отсчетам (точная формулировка этой теоремы и ее доказательство для функций одной переменной даны в приложении к этой главе).

Некоторые дополнительные замечания об интерполяции

Иногда удобно разбить процедуру интерполяции значений функции между ее отсчетами на два этапа: сначала вычислить интерполированные значения для некоторого нового множества точек отсчета, а затем интерполировать функцию уже между этими новыми отсчетами. Например, так можно поступать для того, чтобы сделать более подробной первоначально грубую решетку, образованную точками отсчета, или для замены нерегулярной решетки на регулярную. Преимущество такой двухступенчатой процедуры в том, что интерполяция по регулярной решетке проще, а если эта решетка достаточно тонкая, то появляется возможность использовать интерполяционные полиномы более низкого порядка. Для вычисления интерполированных значений в новом множестве отсчетных точек обычно используют взвешенную сумму значений ближайших «старых» отсчетов, причем веса выбирают в зависимости от взаимного удаления точек.

Поскольку интерполяция функций двух переменных является достаточно сложной операцией, изображения часто аппроксимируют путем интерполяции отдельно по каждой координате. Например, в случае телевизионных изображений особенно удобно выбирать точки отсчетов на строках развертки и производить интерполяцию вдоль и поперек строк.

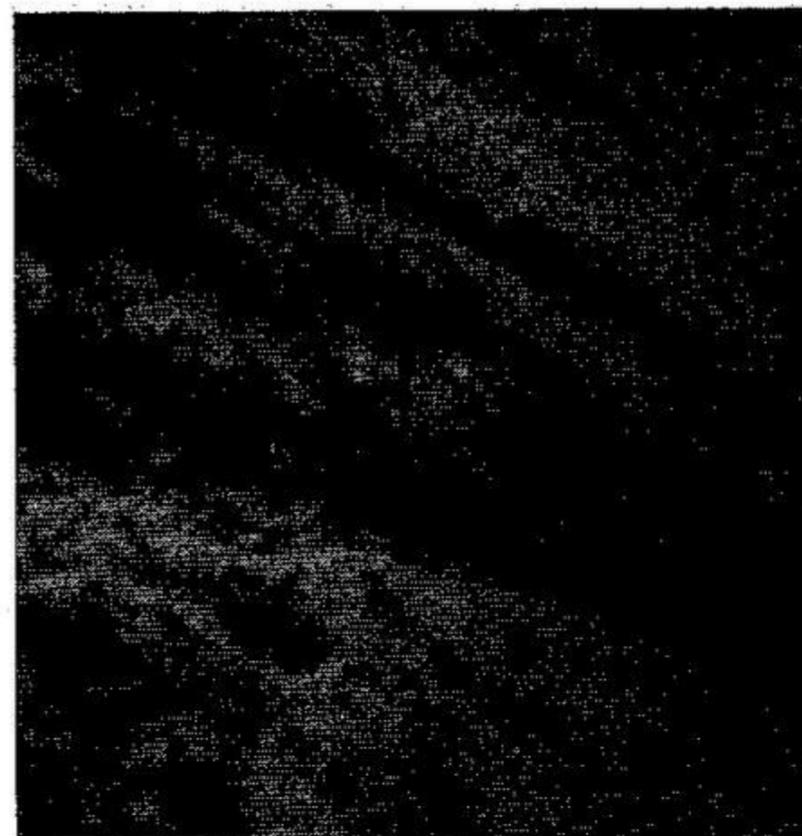
Упражнение. Пусть $f(x, y) = (r^2 - x^2 - y^2)^{1/2}$ — полусфера радиуса r с центром в начале координат. Найти максимальное и среднее значения разности между f и ее аппроксимациями по отсчетным значениям, взятым в точках с целочисленными значениями координат. Рассмотреть следующие способы интерполяции: а) интерполяцию нулевого порядка, когда аппроксимирующая функция постоянна в пределах элементарного квадрата с центром в отсчетной точке; б) интерполяцию первого порядка, когда аппроксимация производится участками

плоскостей, проходящих через каждые три отсчетные точки; в) одномерную интерполяцию сначала по x , затем по y ; г) аппроксимацию первыми k членами ряда Фурье в пределах квадрата $-r \leq x, y \leq r$.

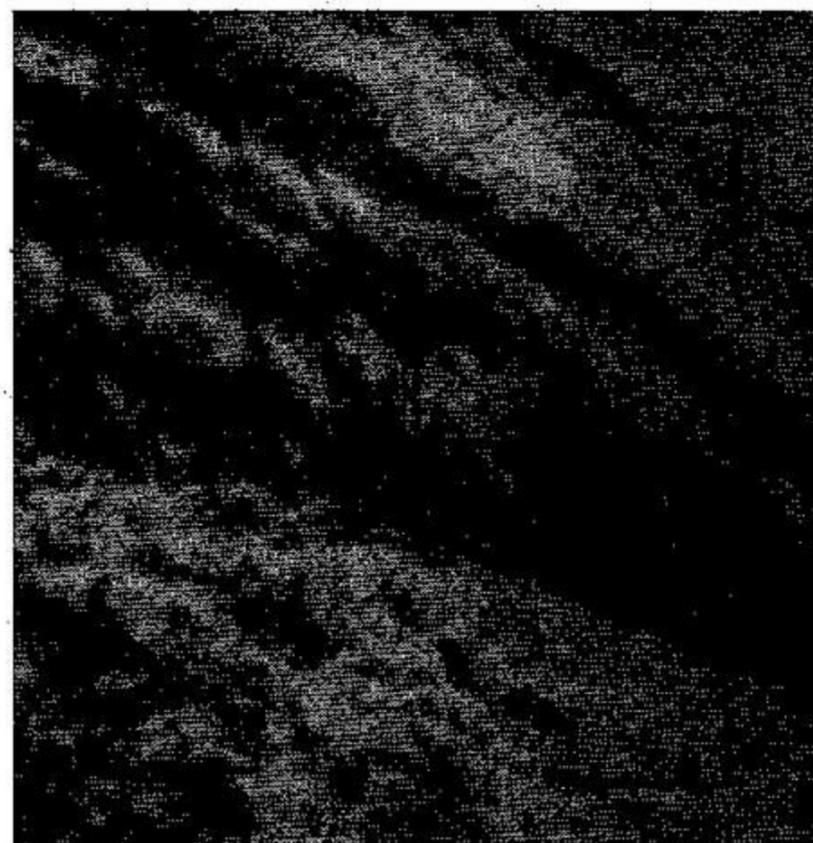
3.2 Квантование с фиксированным расположением уровней

Для того чтобы квантованное изображение имело приемлемое качество, иногда требуется около 100 уровней квантования. Такое большое число уровней квантования необходимо, в частности, если квантованию подвергаются участки изображений с плавно меняющимся значением яркости. В этом случае большим фрагментам на этих участках будут приспаны постоянные значения, а на границах между ними возникнут резкие скачки. При недостаточно большом числе уровней квантования они могут выглядеть как ложные контуры, маскирующие истинные объекты, имеющиеся на изображении и порождающие ложные. Это может сделать квантованную аппроксимацию исходного изображения неприемлемой. Влияние уменьшения числа уровней квантования на качество изображения показано на фиг. 3.1. (для каждого изображения, показанного на фигурах, уровни квантования расположены примерно в одном и том же диапазоне, и каждому элементу приспывается квантованное значение, ближайшее к тому, которое он имел при 32 уровнях квантования).

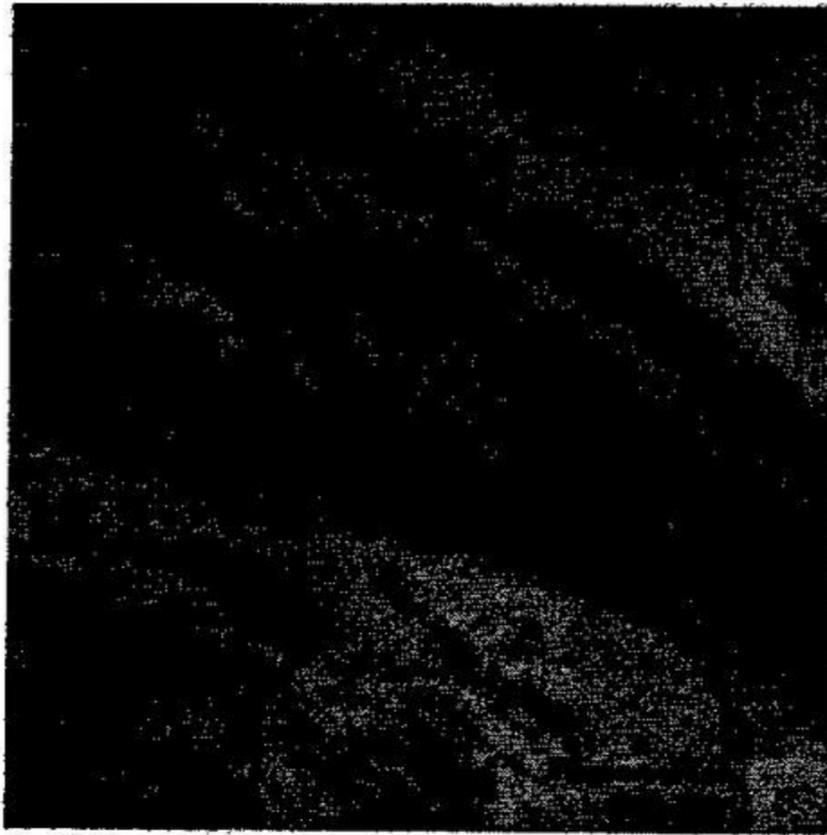
Как и в случае дискретизации по аргументу, проще всего выбрать равномерный шаг квантования. Однако иногда предпочтительнее располагать уровни квантования неравномерно. Например, предположим, что значения яркости в определенном диапазоне встречаются чаще, чем другие. В этом случае может оказаться желательным квантовать значения в этом диапазоне точнее. Это увеличивает среднюю точность квантования при том же общем числе уровней квантования. Этот метод, известный



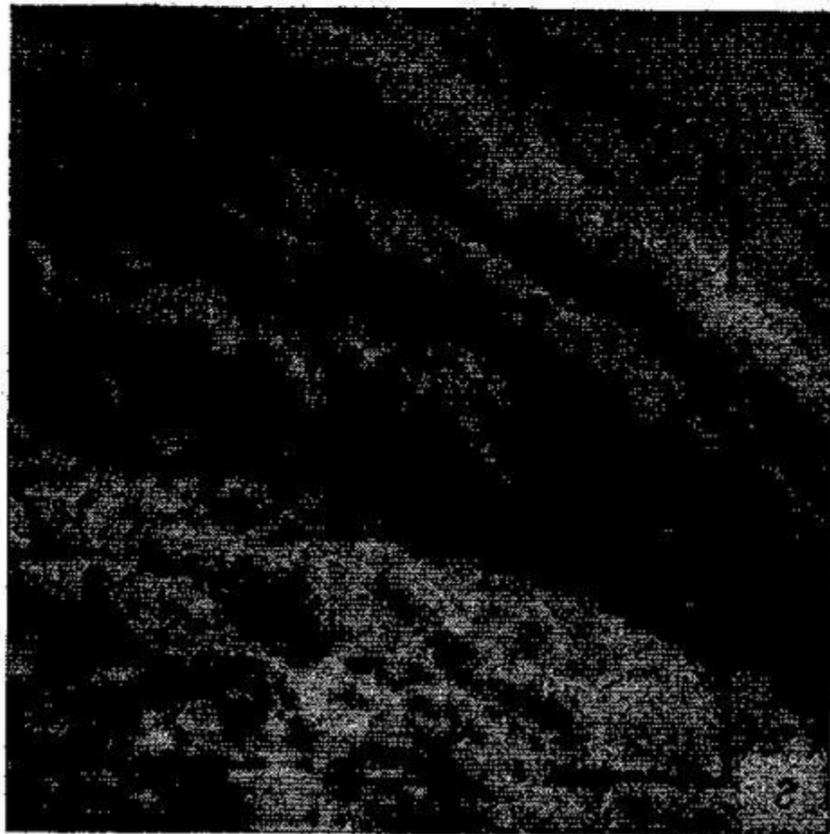
Ф и г. 3.1. Влияние уменьшения числа уровней квантования на качество изображения.
а — тридцать два уровня;



Ф и г. 3.1. б — шестнадцать уровней;



Фиг. 3.1. а — восемь уровней;



Фиг. 3.1. б — четыре уровня;



в — два уровня.

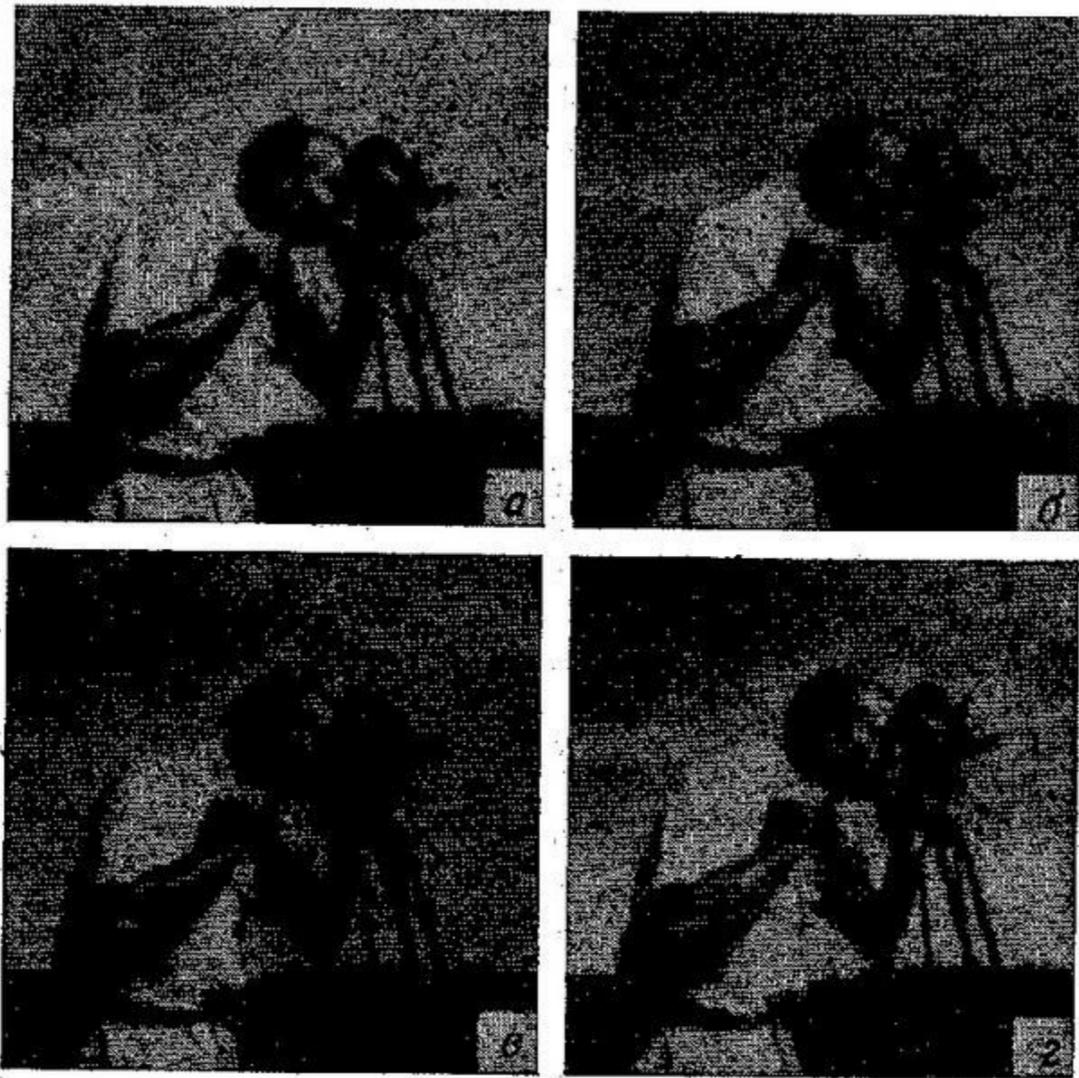
как неравномерное квантование, иллюстрируется на фиг. 3.2.

При квантовании изображения может быть даже желательным вводить в процессе квантования некоторый элемент случайности. Например, сдвигая случайным образом от элемента к элементу уровни квантования или добавляя перед квантованием в изображение случайный шум, можно разрушить ложные контуры, поскольку точки, в которых происходит смена уровня квантования, уже не будут образовывать гладких линий. На практике для этой цели лучше использовать псевдослучайный шум. При воспроизведении квантованного изображения можно вычесть из него ту же псевдослучайную последовательность¹. Этот метод иллюстрируется фиг. 3.3.

Метод квантования с псевдослучайным шумом описан в работе L. G. Roberts, Picture coding using pseudo-

¹ Псевдослучайная последовательность — это последовательность чисел, закон построения которых известен, но достаточно запутан, чтобы считать их как бы случайными. Например, в данном случае они не должны образовывать на изображении осмысленных конфигураций.

random noise, *IRE Trans. Information Theory*, IT-8, 145 (Febr. 1962). В общем случае число уровней квантования, требуемое для приемлемого воспроизведения изображения,



Ф и г. 3.2. Неравномерное квантование.

a — равномерное квантование — 16 уровней; *b* — неравномерное квантование — 16 уровней; *c* — равномерное квантование — 4 уровня; *d* — неравномерное квантование — 4 уровня.

Перепечатано из статьи: Хуанг, Третьяк, Прасада, Ямагучи, Соображения по конструированию КИМ — систем для передачи неподвижных одноцветных изображений низкой четкости, *ТИИЭР*, 55, № 3 (март 1967).

можно уменьшить путем предискажения и корректирования, т. е. соответствующего преобразования изображения перед квантованием и после него,

Упражнение. Найти среднее значение ошибки при равномерном квантовании полусферы

$$f(x, y) = (r^2 - x^2 - y^2)^{1/2}$$



Ф и г. 3.3. Квантование с использованием псевдослучайного шума (8 уровней).

Снимок из статьи: Хуанг, Третьяк, Прасада, Ямагучи, Соображения по конструированию КИМ — системы для передачи неподвижных одноцветных изображений низкой четкости, *ТИИЭР*, 55, № 3 (март 1967).

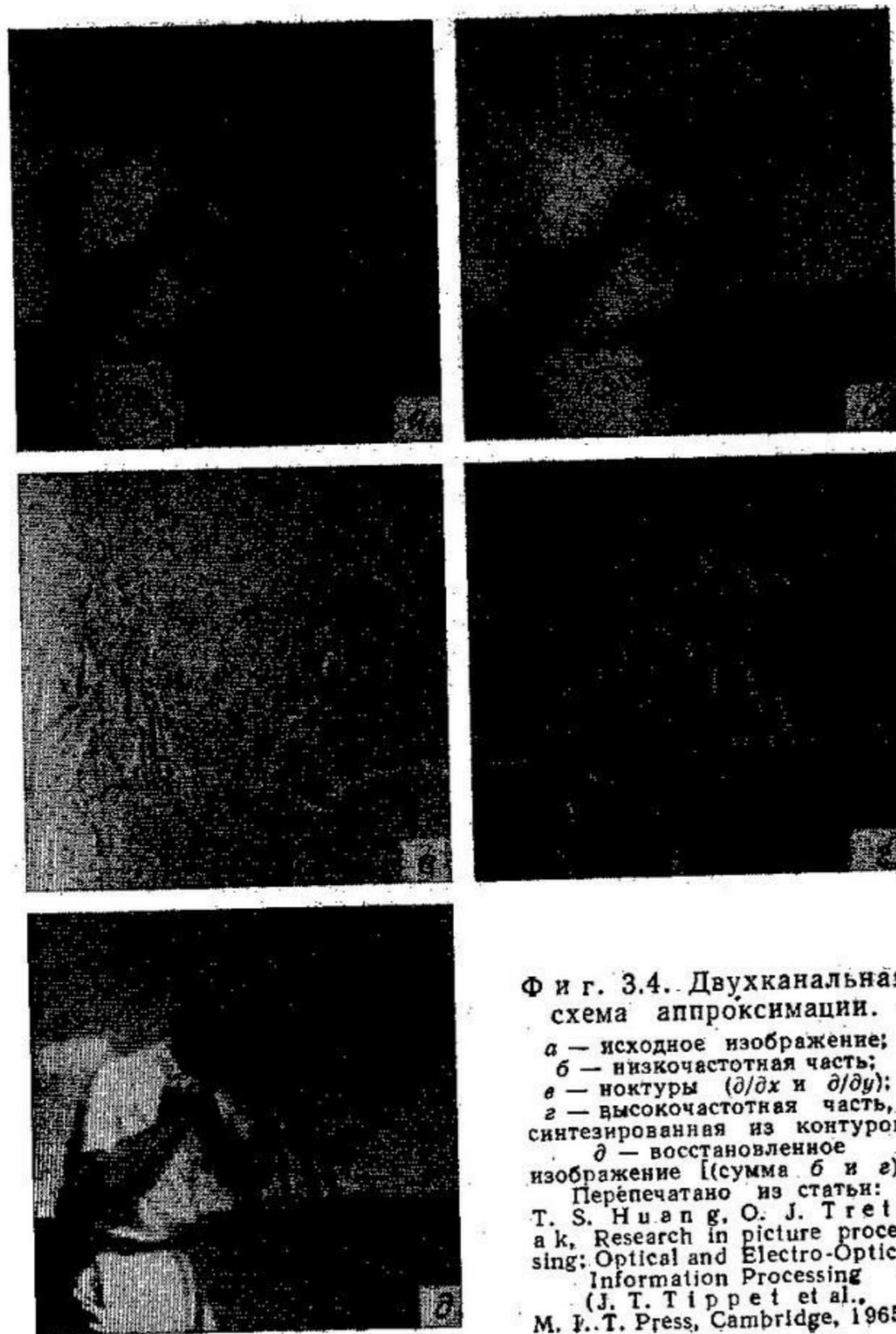
на k уровней. Как следует расположить уровни квантования, чтобы минимизировать это среднее значение?

3.3 Дискретизация и квантование с переменным шагом

Выбор отсчетных точек и уровней квантования не обязательно должен быть predetermined заранее. Его можно поставить в зависимость от характера изображения, и в частности от особенностей зрительного восприятия наблюдателем аппроксимированного изображения и его оценки сходства этого изображения с исходным. Например, глаз относительно мало чувствителен к искажениям яркости изображения вблизи резких контуров, поэтому в этих областях возможно грубое квантование. С другой стороны, на плавных участках изображения недостаточно большое число уровней квантования может

привести к появлению ложных контуров в местах переходов от одного уровня к другому. При дискретизации же, наоборот, отсчеты необходимо располагать чаще вблизи контуров для сохранения их резкости, на плавных же участках их можно брать реже. Таким образом, при дискретизации и квантовании изображения каждую точку нужно исследовать на наличие контура (резкого перепада)¹. На контурных участках следует использовать грубое квантование (конечно, участкам по обе стороны контура должны быть присвоены различные уровни), а на плавных — редкое расположение отсчетов. (Заметим, что подобные результаты в отношении квантования можно получить, используя неравномерное квантование разностей между значениями яркостей соседних элементов изображения² — грубое для больших значений разностей и точное для малых.)

В телевизионных системах низкочастотную и высокочастотную составляющие видеосигнала можно передавать отдельно, используя для них различные схемы кодирования (аппроксимации). Можно передать также только низкочастотную составляющую видеосигнала и положение (а также приблизительные значения контраста) контуров изображения и синтезировать контуры в видеосигнале на приемной стороне. Этот «двухканальный» метод передачи иллюстрируется фиг. 3.4. Другой метод дискретизации с переменным шагом состоит в том, чтобы размещать отсчеты как можно дальше друг от друга, но так, чтобы максимальное либо среднее значение разности между исходной и интерполированной функциями не превышало заданной величины³. Или, наоборот, можно разместить заданное число отсчетов таким образом, чтобы минимизировать эту разность. (Обычное квантование может рассматриваться с этой точки зрения как приближение нулевого порядка, при котором для интерполяции используется заданное множество постоянных функций —



Ф и г. 3.4. Двухканальная схема аппроксимации.
 а — исходное изображение;
 б — низкочастотная часть;
 в — контуры ($\partial/\partial x$ и $\partial/\partial y$);
 г — высокочастотная часть, синтезированная из контуров;
 д — восстановленное изображение [(сумма б и г)].
 Перепечатано из статьи:
 T. S. Huang, O. J. Tretter,
 Research in picture processing;
 Optical and Electro-Optical
 Information Processing
 (J. T. Tippett et al.,
 M. I. T. Press, Cambridge, 1965).

¹ О критериях обнаружения контуров см. разд. 6.5.

² Этот способ, конечно, предполагает наличие дискретного изображения с равномерно расположенными отсчетами.

³ Речь идет, следовательно, о минимизации числа отсчетов на изображении при заданной точности аппроксимации. — Прим. ред.

уровней квантования.) Имеется много модификаций этого метода, но, поскольку они не специфичны для изображений, они не будут здесь рассматриваться подробно.

Упражнение 1. Сколько отсчетов необходимо для аппроксимации полусферы $f(x, y) = (r^2 - x^2 - y^2)^{1/2}$ (при использовании каждой из схем аппроксимации (а) — (в), приведенных в упражнении в конце разд. 3.1) с максимальной (или средней) ошибкой r/k ?

Упражнение 2. Как следует расположить k отсчетов для наилучшей аппроксимации этой полусферы? Использовать каждую из схем интерполяции (а) — (в), приведенных в упражнении в конце разд. 3.1.

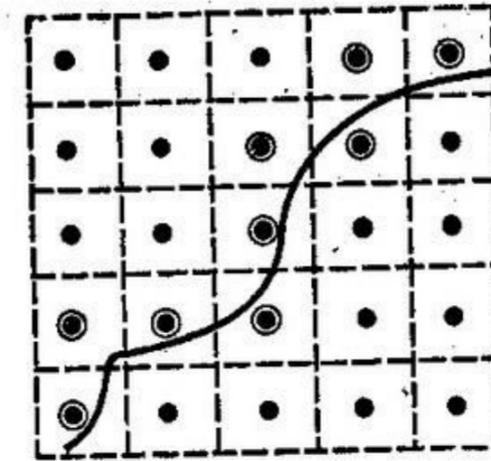
3.4. Аппроксимация рисунков

Рисунки (штриховые изображения) можно полностью описать с помощью конечного числа кривых и дуг. Если эти кривые имеют простое аналитическое описание, то, как указано в разд. 2.4, возможно очень сжатое описание такого изображения. Даже если это не так, изображение можно аппроксимировать путем замены истинных кривых отрезками достаточно простых кривых. При этом снова могут использоваться процедуры дискретизации с фиксированным или переменным расположением отсчетов. В каждом случае отрезки кривых между отсчетными точками можно аппроксимировать путем интерполяции отрезками прямых линий или кривых более высокого порядка. В первом случае получится полигональная аппроксимация, во втором — «гладкая» аппроксимация, если согласовать отрезки кривых в общих точках по наклону и кривизне.

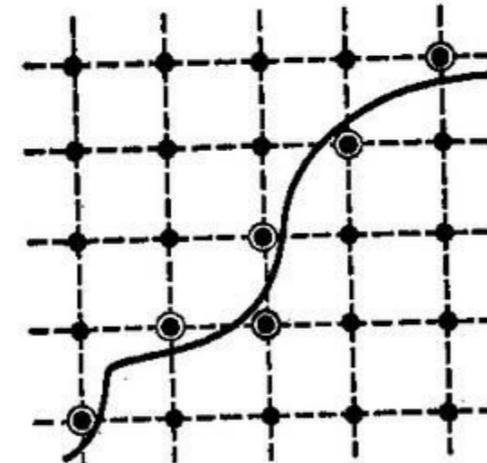
Дискретизация с фиксированным расположением отсчетов

Кривую можно аппроксимировать, используя отсчетные точки, образующие на плоскости фиксированную (например, регулярную прямоугольную) решетку. Например (фиг. 3.5), можно выбирать узлы решетки, в котором смысле ближайšie к кривой. В частности, в ка-

честве аппроксимирующих точек можно использовать узлы, ближайšie к точкам пересечений кривой и линий, соединяющих узлы решетки. В этом случае близость к



а



б

Ф и г. 3.5. Методы аппроксимации кривых на фиксированном растворе.

а — квадратная окрестность; б — измерение близости по линиям решетки.

узлам решетки определяется как расстояние до них, измеряемое вдоль этих линий. В полученной таким образом полигональной аппроксимации каждая сторона аппроксимирующего многоугольника направлена по горизонтали или вертикали и имеет единичную длину, либо по диагоналям сетки и имеет длину $\sqrt{2}$. Такая ломаная

линия полностью задана своей начальной точкой и последовательностью наклонов ее отрезков. Этот вид полигональной аппроксимации [7] известен под названием цепного кодирования. Другим возможным методом дискретизации линий является выбор отсчетных точек на кривой на заданном удалении друг от друга.

Дискретизация с переменным шагом

Как указано в предыдущем разделе, еще одним методом аппроксимации кривой является такой выбор отсчетных точек, при котором разность между исходной и интерполированной кривыми не превышает заданной величины, или такое расположение заданного числа отсчетных точек, при котором эта разность минимальна. Кроме того, в каждом случае можно потребовать, чтобы отсчетные точки лежали на кривой. Если требуется полигональная аппроксимация и число отсчетных точек ограничено, очевидно, что наилучшим размещением отсчетных точек будет их размещение вблизи тех участков кривой, где ее кривизна максимальна [8].

В связи с аппроксимацией кривых можно предложить интересный метод аппроксимации поверхностей (т. е. изображений). Он состоит в построении контурной карты поверхности и аппроксимации контурных линий [9].

Упражнение. Для окружности $x^2 + y^2 = r^2$ выполнить упражнения, аналогичные упражнениям в конце разд. 3.1 и 3.3. Какой из двух методов цепного кодирования, показанных на фиг. 3.5, обеспечивает лучшую аппроксимацию для этой окружности? (Использовать решетку точек с целочисленными значениями координат.)

ПРИЛОЖЕНИЕ

Доказательство теоремы отсчетов¹

Это приложение основано на определении и свойстве обратимости преобразования Фурье (разд. 4.3). Используется также хорошо известный факт, что если $g(t)$ яв-

¹ В советской литературе эта теорема известна как теорема Котельникова. — Прим. ред.

ляется достаточно гладкой функцией на интервале $-b \leq t \leq b$, то ее можно разложить в ряд Фурье

$$g(t) = \sum_{k=-\infty}^{\infty} a_k e^{\pi j k t / b},$$

где $j = \sqrt{-1}$,

и

$$a_k = \frac{1}{2b} \int_{-b}^b e^{-\pi j k t / b} g(t) dt,$$

где $k = 0, \pm 1, \dots$.

Сумма этого ряда является периодической функцией с периодом $2b$, которая совпадает с g на интервале $[-b, b]$.

Теорема 1. Пусть f — функция одной переменной, преобразование Фурье которой равно нулю за пределами интервала, $-b \leq \omega_t \leq b$. Тогда f можно точно восстановить из ее отсчетов, расположенных не далее чем на $1/2b$ друг от друга.

Доказательство. Пусть $F(\omega_t) = \int_{-\infty}^{\infty} e^{2\pi j \omega_t t} f(t) dt$ — преобразование Фурье функции f . По предположению

$F(\omega_t) = 0$ при $|\omega_t| > b$. Пусть $F(\omega_t) = \sum_{k=-\infty}^{\infty} a_k e^{\pi j k \omega_t / b}$ —

разложение F в ряд Фурье на интервале $-b \leq \omega_t \leq b$. Следовательно,

$$a_k = \frac{1}{2b} \int_{-b}^b e^{-\pi j k \omega_t / b} F(\omega_t) d\omega_t. \quad (1)$$

Но поскольку f есть обратное преобразование Фурье от F , то

$$f(t) = \int_{-\infty}^{\infty} e^{-2\pi j \omega_t t} F(\omega_t) d\omega_t = \int_{-b}^b e^{-2\pi j \omega_t t} F(\omega_t) d\omega_t. \quad (2)$$

Сравнив (1) и (2), получим, что $a_k = f(k/2b)/2b$, так что

$$F(\omega_t) = (1/2b) \sum_{-\infty}^{\infty} f(k/2b) e^{\pi j k \omega_t / b}, \text{ откуда}$$

$$\begin{aligned} f(t) &= (1/2b) \sum_{-\infty}^{\infty} f(k/2b) \left[\int_{-b}^b e^{\pi j \omega_t [(k/b) - 2t]} d\omega_t = \right. \\ &= \left. \sum_{-\infty}^{\infty} f(k/2b) \frac{\sin [\pi b ((k/b) - 2t)]}{\pi b ((k/b) - 2t)} \right] \end{aligned}$$

Эта формула выражает $f(t)$ через ее отсчеты $f(k/2b)$. Заметим, что если разложить F в ряд Фурье на интервале, большем чем $[-b, b]$, то в результате получим выражение для воспроизведения f из ее отсчетов, удаленных друг от друга меньше чем на $1/2b$.

Аналогично можно показать, что для функций двух переменных верна следующая теорема.

Теорема 2. Пусть f — функция двух переменных, фурье-преобразование которой F равно нулю за пределами некоторой замкнутой области S на плоскости; (a, b) , (c, d) — произвольные точки, удовлетворяющие условию, что области, получаемые при переносе всех точек S на $[m(a+c), n(b+d)]$ и $[p(a+c), q(b+d)]$, где (m, n) , (p, q) — произвольные пары целых чисел, не перекрываются. (Параллелограмм, определяемый векторами (a, b) и (c, d) , играет здесь роль интервала $[-b, b]$ в теореме 1.) Пусть также (s, t) и (u, v) таковы, что $as + bt = cu + dv = 2\pi$ и $au + bv = cs + dt = 0$. Тогда f можно точно восстановить из ее отсчетов, взятых в точках $(hs + ku, ht + kv)$, где h и k — целые числа.

Следует указать, что на практике F может быть только близка к нулю за пределами S . Вследствие этого воспроизведение функции по ее отсчетам будет приближенным.

ЛИТЕРАТУРА

1. Huang T. S., Digital picture coding, Proc. Nat. Electron. Confer., 22, 1966, p. 793.
2. Pratt W. E., A bibliography on television bandwidth reduction techniques, *IEEE Trans. Information Theory*, IT-13, 114 (Jan. 1967); см. также in A. Rosenfeld, Bandwidth reduction bibliography, *IEEE Trans. Information Theory*, IT-14, 601 (July 1968).
3. Special issue on redundancy reduction, *Proc. IEEE*, 55, 251 (March 1967); русский перевод: спец. выпуск ТИИЭР «Сокращение избыточности», ТИИЭР, 55 (март 1967).
4. Huang T. S., Tretiak O. J., Research in picture processing, OEOIP, p. 45.
5. Simpson R. S., Blackwell C. A., Frost W. O., Compendium of redundancy removal processes, *IEEE Trans. Aerospace Electronic Systems*, AES-2, 471 (July 1966).
6. Davison L. D., The theoretical analysis of data compression systems, *Proc. IEEE*, 56, 176 (Febr. 1968); Дэвиссон Л., Теоретический анализ систем сжатия информации, ТИИЭР, 56, № 2 (1968).
7. Freeman H., On the encoding of arbitrary geometric configurations, *IRE Trans. Electronic Computers*, EC-10, 260 (June 1961).
8. Attneave F., Some informational aspects of visual perception, *Psychol. Rev.*, 61, 183 (1954).
9. Boehm B. W., Tabular representations of multivariate functions — with applications to topographic modeling, Proc. ACM Nat. Confer., Aug. 1967, p. 403.
10. Petersen D. P., Middleton D., Sampling and reconstruction of wave-numberlimited functions in n -dimensional Euclidean spaces, *Information and Control*, 5, 279 (1962).
11. Prosser R. T., A multidimensional sampling theorem, *J. Math. Anal., Appl.*, 16, 574 (1966).
12. Cheng G. C., Ledley R. S., A theory of pictures digitizations and applications, PPR, p. 329.
13. Montgomery W. D., Reconstruction of pictures from scanned records, *IEEE Trans. Information Theory*, IT-11, 204 (Apr. 1965).
- 14*. Игнатьев Н. К., Оптимальная дискретизация двумерных сообщений, Известия вузов «Радиотехника», 1961, № 6.
- 15*. Поляков В. Г., Естественная дискретная модель рисунка, некоторые асимптотические ее свойства и предсказание хода следящей развертки, Проблемы передачи информации, № 3 (1970).

Теория пространственно-инвариантных операций над изображениями

В этой и двух последующих главах рассматриваются различные виды преобразований, характер которых не зависит от положения обрабатываемой области на изображении.

4.1. Пространственно-инвариантные операции

Пусть \mathcal{S} — множество функций двух действительных переменных, определенных на всей плоскости, а φ — некоторая функция, переводящая элементы множества \mathcal{S} друг в друга. Будем называть φ операцией над множеством \mathcal{S} . Пусть $T_{u,v}$ — перенос на u, v — операция «сдвига» над \mathcal{S} , определяемая для всех f из \mathcal{S} и всех x, y на плоскости соотношением¹

$$[T_{u,v}(f)](x, y) = f(x-u, y-v).$$

Операцию φ будем называть *пространственно-инвариантной*, если она коммутирует с операцией $T_{u,v}$, т. е. если

$$\varphi [T_{u,v}(f)] = T_{u,v} [\varphi(f)]$$

для всех f из \mathcal{S} и всех значений (u, v) на плоскости.

Пространственно-инвариантные операции являются двумерным аналогом инвариантных по времени операций, изучаемых в радиотехнике.

Важнейшими пространственно-инвариантными операциями являются:

¹ Здесь $[T_{u,v}(f)](x, y)$ обозначает значение функции $T_{u,v}(f)$ в точке (x, y) .

Операция сдвига

Операции сдвига являются пространственно-инвариантными операциями, так как они коммутируют друг с другом. Действительно,

$$\begin{aligned} & [T_{r,s}(T_{u,v}(f))](x, y) = \\ & = [T_{u,v}(f)](x-r, y-s) = f[(x-r)-u, (y-s)-v] = \\ & = f[(x-u)-r, (y-v)-s] = [T_{r,s}(f)](x-u, y-v) = \\ & = [T_{u,v}(T_{r,s}(f))](x, y) \end{aligned}$$

при всех r, s, u, v и f . Очевидно, что другие виды геометрических преобразований, например вращение, изменение масштаба, не являются пространственно-инвариантными.

Упражнение 1. Пусть φ — геометрическое преобразование вида $[\varphi(f)](x, y) = f[\xi(x, y), \eta(x, y)]$. Может ли оно быть пространственно-инвариантным, не являясь сдвигом?

Поэлементные операции

Предположим, что операция φ такова, что $[\varphi(f)](x, y)$ зависит только от $f(x, y)$, т. е. существует такая функция h одной действительной переменной, что $[\varphi(f)](x, y) = h[f(x, y)]$ для всех f в \mathcal{S} и всех (x, y) на плоскости. Очевидно, что поэлементные операции являются пространственно-инвариантными. Действительно, для всех (x, y) и (u, v) на плоскости и всех f из \mathcal{S} имеем $[\varphi(T_{u,v}(f))](x, y) = h([T_{u,v}(f)](x, y)) = h(f(x-u, y-v)) = [\varphi(f)](x-u, y-v) = [T_{u,v}(\varphi(f))](x, y)$.

Поэлементные операции над изображениями изменяют значения яркости каждого элемента изображения по закону, не зависящему от остальных элементов, т. е. поэлементные операции изменяют значения элементов изображения, сохраняя его геометрическую структуру. Примерами таких операций могут служить усиление и ослабление контрастности изображения (т. е. умножение зна-

чений яркости его элементов на константу), гамма-коррекция¹, квантование.

Локальные операции

Эти операции являются обобщением поэлементных операций. Пусть $[f(x, y)]$ зависит только от значений f в некоторой окрестности $N(x, y)$, причем эта окрестность неизменна для всех (x, y) на плоскости, т. е. окрестность $N(x-u, y-v)$ представляет собой множество таких точек $z-u, w-v$, что (z, w) принадлежат $N(x, y)$ для любых u, v . Такие операции Φ будем называть *локальными*. Очевидно, что они являются пространственно-инвариантными операциями. Поэлементные операции получаются как частный случай локальных операций, когда окрестность элемента состоит только из него самого. В гл. 6 будет приведено много примеров полезных локальных операций, например операции сглаживания или операции повышения резкости.

Упражнение 2. Доказать, что локальные операции являются пространственно-инвариантными.

Иногда желательно рассматривать операции Φ , которые естественным образом определяются для отдельных классов функций, но не определены для всех функций множества. Предположим, что таким классом функций является класс, замкнутый по отношению к переносу, т. е. класс функций f , для которых при всех u, v существует $T_{u,v}(f)$. К такому классу функций принадлежат, например, все интегрируемые функции, непрерывные функции, функции, представляющие изображения, класс D_n дискретных изображений, состоящих из $n \times n$ элементов при заданном n , и т. д. В таком случае на остальной части множества \mathcal{S} можно определить Φ как операцию тождественного преобразования. Очевидно, что такая обобщенная операция остается пространственно-инвариантной. Ниже описаны два полезных вида опера-

¹ Гамма-коррекцией в телевидении называют нелинейное преобразование видеосигнала, выполняемое с целью коррекции характеристик свет — сигнал и сигнал — свет видеодатчиков и устройств воспроизведения изображений. — *Прим. ред.*

ций, которые естественным образом определяются для класса функций D_n . Эти операции будут также рассмотрены в гл. 6.

Поэлементные дискретные операции

Пусть $\Phi[(a_{ij})] = [h(a_{ij})]$ — матрица, (i, j) -й элемент которой есть $h(a_{ij})$, где h — неотрицательная действительная функция одной действительной переменной. Поскольку операция Φ производится только над одним элементом матрицы, она может рассматриваться как поэлементная операция над дискретным изображением.

Дискретные локальные операции

Пусть

$$\Phi[(a_{ij})] = [H(a_{i-1, j-1}, a_{i-1, j}, a_{i-1, j+1}, \\ a_{i, j-1}, a_{ij}, a_{i, j+1}, a_{i+1, j-1}, \\ a_{i+1, j}, a_{i+1, j+1})],$$

где H — неотрицательная действительная функция девяти действительных переменных, а a_{rs} предполагаются равными нулю за пределами области $1 \leq r, s \leq n$. Такая операция может рассматриваться как локальная операция над дискретным изображением, так как она производится только над непосредственной окрестностью каждого элемента матрицы. Можно, конечно, определить такую операцию и для большей окрестности.

4.2. Пространственно-инвариантные линейные операции

Операция Φ называется *линейной* (или иногда *линейной однородной*), если

$$\Phi(af + bg) = a\Phi(f) + b\Phi(g)$$

для всех f, g из \mathcal{S} и всех действительных значений a и b .

Упражнение 1. Доказать, что: а) любая геометрическая операция (упражнение 1, разд. 4.1) является линейной; б) поэлементная операция является линейной, если

функция h имеет вид $h(z) = cz$, где c — некоторая действительная константа. Верно ли обратное?

Пусть g_0 — дискретное изображение из $n \times n$ элементов, определяемое матрицей (c_{ij}) , где $c_{11} = 1$, $c_{ij} = 0$ для остальных значений i, j . Тогда для любого дискретного изображения f из $n \times n$ элементов, заданного матрицей (a_{ij}) , получим

$$f = \sum_{1 \leq i, j \leq n} a_{ij} (T_{i-1, j-1} (g_0)),$$

где $T_{i-1, j-1}$ — оператор переноса, соответствующий сдвигу элемента S_{11} до совмещения с элементом S_{ij} . Суммирование можно производить от $-\infty$ до ∞ , поскольку по определению $a_{ij} = 0$ вне области $1 \leq i, j \leq n$.

Пусть φ — пространственно-инвариантная линейная операция. Тогда

$$\begin{aligned} \varphi(f) &= \sum_{i, j = -\infty}^{\infty} a_{ij} \varphi(T_{i-1, j-1} (g_0)) = \\ &= \sum_{i, j = -\infty}^{\infty} a_{ij} T_{i-1, j-1} (\varphi(g_0)). \end{aligned}$$

Следовательно, действие φ на произвольное дискретное изображение полностью определяется ее действием на g_0 . Кроме того, $a_{ij} = f(i-1, j-1)$ для всех i, j , и если обозначить $\varphi(g_0) = g_\varphi$, то

$$\begin{aligned} [\varphi(f)](x, y) &= \sum_{i, j = -\infty}^{\infty} g_\varphi(x-(i-1), y-(j-1)) f(i-1, j-1) = \\ &= \sum_{i, j = -\infty}^{\infty} g_\varphi(x-i, y-j) f(i, j). \end{aligned} \quad (1)$$

Упражнение 2. Доказать, что если g_φ в соотношении (1) является дискретным изображением из $m \times m$ элементов, то $\varphi(f)$ — также дискретное изображение, состоящее из $(n+m-1) \times (n+m-1)$ элементов.

Аналогом соотношения (1), соответствующим непре-

рывным изображениям и операциям, является следующее выражение:

$$[\varphi(f)](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_\varphi(x-u, y-v) f(u, v) dudv. \quad (2)$$

Ниже всегда будет предполагаться, что любая пространственно-инвариантная операция определяется выражением вида (2).

Доказательство выражения (2) с математической точки зрения не тривиально. Непрерывное изображение f можно рассматривать как предел, к которому стремится дискретное изображение из $n \times n$ элементов при возрастании n . Но существование предела для соответствующей последовательности $\varphi(f)$ не очевидно.

С другой стороны, заметим, что если операция φ определяется выражением (2), то она является пространственно-инвариантной и линейной. Таким образом, это выражение [а в дискретном случае выражение (1)] может использоваться как определение класса пространственно-инвариантных линейных операций. Пусть, например, в выражении (2) $g_\varphi(u, v) = 0$ вне некоторой окрестности N точки $(0, 0)$. Тогда $g_\varphi(x-u, y-v) = 0$ вне соответствующей (смещенной) окрестности точки (x, y) и, таким образом, $[\varphi(f)](x, y)$ зависит только от значений f в этой окрестности, являясь локальной линейной операцией.

Для аналогичного определения поэлементной операции следует предположить, что g_φ имеет вид $c\delta$, где c — константа, а δ — дельта-функция, т. е. функция, для которой

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x-u, y-v) f(u, v) dudv = f(x, y)$$

при всех x и y .

Упражнение 3. Пусть $P = (p_{ij})$ и $Q = (q_{ij})$ — матрицы, состоящие из $n \times n$ элементов, причем элементы, расположенные на любой линии, параллельной главной диагонали, равны между собой. В частности, пусть $p_{ij} =$

$= q_{ij} = 0$ при всех i, j , кроме $j = i - 1, i, i + 1$. Доказать, что операция $\varphi(A) = PAQ$ (где $A = (a_{ij})$ — произвольная матрица из $n \times n$ элементов) является дискретной локальной операцией, описанной в конце разд. 4.1, и что эта операция линейна. [О методе определения линейных локальных операций над дискретными изображениями см. Proc. Nat. Electron Confer., 18, 529 (1962); Tobler W. R., Of maps and matrices, J. Regional Sci., 7, 275 (1967).]

В заключение введем некоторые термины и обозначения. Правая часть выражения (2) называется *сверткой* функций f и g_φ и обозначается $f * g_\varphi$. Таким образом, выражение (2) можно переформулировать следующим образом: любая пространственно-инвариантная линейная операция φ является операцией свертки, т. е. имеется такая функция g (она обозначена здесь g_φ), что $\varphi(f) = f * g$ для всех f . Заметим, что при замене $x - u = u'$ и $y - v = v'$ получим

$$\begin{aligned} g * f &= \iint g(u, v) f(x - u, y - v) du dv = \\ &= \iint g(x - u', y - v') f(u', v') du' dv' = f * g. \end{aligned}$$

Пусть \bar{g} — функция, полученная из g заменой знака ее аргументов, т. е. $\bar{g}(u, v) = g(-u, -v)$ для всех u, v . Свертка $f * \bar{g}$ называется функцией *взаимной корреляции* f и g . Будем обозначать ее $f \otimes g$. Следовательно, $f \otimes g = \iint g(u - x, v - y) f(u, v) du dv$. Если положить $u' = u - x, v' = v - y$, этот интеграл перейдет в $\iint g(u', v') f(u' + x, v' + y) du' dv'$. Такая запись используется чаще. Очевидно, что $\overline{f * g} = \bar{f} * \bar{g}$, так что $g \otimes f = \overline{f \otimes g}$.

Функция взаимной корреляции f с f называется *автокорреляционной функцией* f . Заметим, что

$$\begin{aligned} [(T_{a,b}(f)) \otimes (T_{a,b}(f))](x, y) &= \\ &= \iint f(u - a, v - b) f(u - a + x, v - b + y) du dv = \\ &= \iint f(u - a, v - b) f(u - a + x, v - b + y) \times \\ &\quad \times d(u - a) d(v - b) = [f \otimes f](x, y), \end{aligned}$$

т. е. автокорреляционная функция f при сдвиге f не изменяется. То же верно и для свертки $f * f$. Заметим также, что, поскольку $\overline{f \otimes f} = f \otimes f$, автокорреляционная функция симметрична относительно начала координат.

4.3. Преобразование Фурье

Этот раздел посвящен обзору основных свойств преобразования Фурье и их применению к исследованию пространственно-инвариантных линейных операций.

Под преобразованием Фурье функции $f(x, y)$ понимается преобразование вида

$$F(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{2\pi j(\omega_x x + \omega_y y)} dx dy,$$

где j обозначает $\sqrt{-1}$. Обратным преобразованием Фурье называется преобразование

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-2\pi j(\omega_x x + \omega_y y)} F(\omega_x, \omega_y) d\omega_x d\omega_y.$$

Эта функция в силу свойства обратимости преобразования Фурье равна исходной функции $f(x, y)$ при условии, что $f(x, y)$ удовлетворяет некоторым ограничениям.

Преобразование Фурье $F(\omega_x, \omega_y)$ является комплексной функцией. Поскольку $e^{jz} = \cos z + j \sin z$, ее действительная и мнимая части равны соответственно

$$C(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos[2\pi(\omega_x x + \omega_y y)] f(x, y) dx dy$$

и

$$S(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sin[2\pi(\omega_x x + \omega_y y)] f(x, y) dx dy.$$

Эти функции называют иногда соответственно *косинусными* и *синусными* преобразованиями Фурье функции f . Квадрат модуля $F(\omega_x, \omega_y)$, т. е. произведение $F(\omega_x, \omega_y)$ на комплексно сопряженную функцию, называется энер-

гетическим спектром (а в оптике иногда винеровским спектром) f . Для одномерного преобразования Фурье можно записать аналогичные соотношения:

$$F(\omega_t) = \int_{-\infty}^{\infty} e^{2\pi j \omega_t t} f(t) dt,$$

$$f(t) = \int_{-\infty}^{\infty} e^{-2\pi j \omega_t t} F(\omega_t) d\omega_t,$$

$$C(\omega_t) = \int_{-\infty}^{\infty} \cos(2\pi \omega_t t) f(t) dt,$$

$$S(\omega_t) = \int_{-\infty}^{\infty} \sin(2\pi \omega_t t) f(t) dt.$$

В качестве простого примера одномерного преобразования Фурье рассмотрим преобразование функции $f(t) = 1$ при $-1 \leq t \leq 1$, $f(t) = 0$ при других значениях t . Тогда

$$\begin{aligned} F(\omega_t) &= \int_{-1}^1 e^{2\pi j \omega_t t} dt = e^{2\pi j \omega_t t} \Big|_{-1}^1 = \\ &= (e^{2\pi j \omega_t} - e^{-2\pi j \omega_t}) / 2\pi j \omega_t = \sin(2\pi \omega_t) / \pi \omega_t. \end{aligned}$$

Упражнение 1. Найти преобразование Фурье и энергетический спектр функций

$$f(x, y) = \begin{cases} k & \text{при } a \leq x \leq b, c \leq y \leq d, \\ 0 & \text{в противном случае.} \end{cases}$$

Преобразование Фурье, очевидно, является линейной операцией, которая определена на множестве комплексных функций двух переменных, но оно не является пространственно-инвариантной операцией. Действительно, преобразование Фурье смещенной функции $T_{r,s}(f)$ равно

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2\pi j (\omega_x x + \omega_y y)} f(x-r, y-s) dx dy.$$

Заменяя $x' = x - r$, $y' = y - s$, получим

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2\pi j [\omega_x (x'+r) + \omega_y (y'+s)]} f(x', y') dx' dy' &= \\ &= e^{2\pi j (\omega_x r + \omega_y s)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{2\pi j (\omega_x x' + \omega_y y')} \times \\ &\times f(x', y') dx' dy' = e^{2\pi j (\omega_x r + \omega_y s)} F(\omega_x, \omega_y). \end{aligned}$$

Следовательно, сдвиг функции приводит к изменению фазы ее преобразования Фурье, хотя модуль этого преобразования не меняется.

Несмотря на то что преобразование Фурье не является пространственно-инвариантным, его можно использовать для выполнения пространственно-инвариантных линейных операций. Как указано в разд. 4.2, любую такую операцию можно представить в виде свертки некоторой функции с данной функцией. Докажем следующую теорему о свертке: свертка двух функций равна обратному преобразованию Фурье от произведения преобразований Фурье этих функций. Из этой теоремы следует, что для вычисления свертки двух функций нужно найти и перемножить их преобразования Фурье, а затем взять обратное преобразование Фурье от этого произведения. Благодаря недавним успехам в разработке эффективных алгоритмов вычисления преобразования Фурье [1, 2] такой метод может ускорить вычисление свертки по сравнению с непосредственным вычислением.

Теорема. Пусть F и G — преобразования Фурье соответственно функций f и g . Тогда обратное преобразование Фурье произведения FG равно свертке $f * g$.

Доказательство. Пусть H — преобразование Фурье от $f * g$. Достаточно доказать, что $H = FG$. Имеем:

$$\begin{aligned} H(\omega_x, \omega_y) &= \iint e^{2\pi j (\omega_x x + \omega_y y)} \times \\ &\times [\iint f(x-u, y-v) g(u, v) dudv] dx dy = \\ &= \iint g(u, v) [\iint e^{2\pi j (\omega_x x + \omega_y y)} f(x-u, y-v) dx dy] dudv. \end{aligned}$$

Но интеграл в квадратных скобках, как показано выше, равен $e^{2\pi j(\omega_x u + \omega_y v)} F(\omega_x, \omega_y)$. Следовательно,

$$\begin{aligned} H(\omega_x, \omega_y) &= \iint g(u, v) e^{2\pi j(\omega_x u + \omega_y v)} F(\omega_x, \omega_y) du dv = \\ &= F(\omega_x, \omega_y) \iint e^{2\pi j(\omega_x u + \omega_y v)} g(u, v) du dv = \\ &= F(\omega_x, \omega_y) G(\omega_x, \omega_y). \end{aligned}$$

Теоремы о свертке, аналогичные доказанной выше, справедливы и для других видов интегральных преобразований, таких, как преобразование Лапласа или его дискретный аналог — Z-преобразование. Однако обычно при изучении пространственно-инвариантных операций удобнее использовать преобразования Фурье. Одна из причин этого состоит в том, что преобразование Фурье очень просто реализовать в оптических устройствах (см. разд. 5.4). Другой причиной является возможность интерпретации преобразования Фурье как спектра пространственных частот, содержащихся в исходной функции, поскольку преобразование Фурье выражается через синусоидальные функции.

Из равенства преобразования Фурье свертки $f * g$ произведению FG непосредственно следует, что преобразование Фурье от $f \otimes g = f * \bar{g}$ равно $F\bar{G}$. Но $G(\omega_x, \omega_y) = \iint e^{2\pi j[(-\omega_x)x + (-\omega_y)y]} g(x, y) dx dy = \iint e^{-2\pi j(\omega_x x + \omega_y y)} \times g(x, y) dx dy$ — это просто функция, комплексно-сопряженная с $G(\omega_x, \omega_y)$. В частности, положив $g = f$, получим, таким образом, что преобразование Фурье от автокорреляционной функции $f \otimes f$, или энергетический спектр f , равно произведению F и комплексно-сопряженной ей функции.

Упражнение 2. Пусть f симметрична относительно начала координат, т. е. $f(-x, -y) = f(x, y)$ для всех x, y . Доказать, что ее преобразование Фурье F — действительная функция.

Упражнение 3. Доказать, что энергетический спектр является симметричной относительно начала координат функцией.

4.4. Импульсные функции рассеивания и передаточные характеристики

Импульсная функция рассеивания

Функция g_ϕ в выражении (2) (разд. 4.2) называется *импульсной функцией рассеивания*. Выбор этого термина объясняется тем, что g_ϕ определяется (в случае дискретного изображения) как результат операции ϕ над точечным изображением g_0 , состоящим из одной контрастной точки на однородном фоне. В общем случае он состоит в размытии (рассеивании) одиночной точки в g_0 по некоторой окрестности. В тех случаях, когда соотношение (2) справедливо, можно говорить, что любая пространственно-инвариантная линейная операция может быть определена заданием соответствующей импульсной функции рассеивания.

В непрерывном случае соотношение (2) также означает, что любое изображение может рассматриваться как сумма множества точечных изображений. Такие изображения, конечно, физически нереализуемы. Реальные изображения могут к ним только приближаться. Импульсная функция рассеивания является двумерным аналогом импульсной реакции — понятия, используемого в теории цепей и радиотехнике и обозначающего функцию времени, описывающую реакцию цепи на поданный на ее вход импульс малой длительности.

Интеграл от импульсной функции рассеивания в пределах $(-\infty, \infty)$ является мерой действия операции ϕ на функцию, равную константе. Действительно, пусть $f(u, v) = c$ при всех u, v . Тогда функция $[\phi(f)](x, y) = c \iint g_\phi(x-u, y-v) du dv = c \iint g_\phi(x-u, y-v) d(x-u) d(y-v)$. Этот интеграл не зависит от x и y . Пусть $\iint g_\phi(u, v) du dv = k_\phi$. Тогда $\phi(c) = k_\phi c$ для всех c . Если $k_\phi > 1$, ϕ приводит к общему усилению контрастности изображения. При $k_\phi < 1$ происходит его ослабление.

Другие функции рассеивания

В некоторых случаях удобно характеризовать операцию ϕ с помощью результата ее действия не на точечную функцию, а на некоторые другие простые виды входных

функций. Часто в качестве такой функции используют функцию, равную нулю, на всей плоскости, кроме точек, лежащих на некоторой прямой линии, в которых она равна единице. Результат действия φ на такую функцию называют функцией рассеивания линии. Другой важной функцией является функция, равная нулю по одну сторону от некоторой линии и единице по другую. В этом случае результат операции φ называется функцией рассеивания границы. (Если операция φ не изотропна, эти функции рассеивания не будут, вообще говоря, инвариантны к ориентации, а будут зависеть от наклона линий.)

Чтобы вывести соотношение между импульсной функцией рассеивания и функцией рассеивания линий, примем, что h_0 — дискретное изображение, состоящее из $n \times n$ элементов и определяемое матрицей (b_{ij}) , в которой элементы b_{ij} равны 1, а остальные равны нулю. Очевидно, $h_0 = \sum_{j=0}^{n-1} T_{0j}(g)$.

Если обозначить $\varphi(h_0) = h_\varphi$, получим

$$h_\varphi(x, y) = \sum_{j=0}^{n-1} g_\varphi(x, y-j).$$

Аналогично в случае непрерывного изображения можно считать, что

$$h_\varphi(x, y) = \int_{-\infty}^{\infty} g_\varphi(x, y-v) dv = \int_{-\infty}^{\infty} g_\varphi(x, y-v) d(y-v).$$

В словесной формулировке эти соотношения означают: функция рассеивания линии может быть получена интегрированием импульсной функции рассеивания по направлению, параллельному направлению линии. Заметим, что $h_\varphi(x, y)$ не зависит от y , так что можно обозначить ее просто $h_\varphi(x)$.

Упражнение 1. Можно ли импульсную функцию рассеивания выразить через функцию рассеивания линии, если операция φ изотропна?

Упражнение 2. Как связаны между собой функции рассеивания границы и линии?

Амплитудная передаточная характеристика

Еще одной характеристикой операции φ в непрерывном случае может быть ее действие на синусоидальную функцию, например вида $a + b \cos 2\pi \omega x$. В результате операции φ эта синусоидальная функция переходит в $\iint g_\varphi(x-u, y-v) \times [a + b \cos(2\pi \omega u)] dudv = k_\varphi a + b \int h_\varphi(x-u) \cos(2\pi \omega u) du$, поскольку, как известно, $\iint g_\varphi dudv = k_\varphi$, а $\int g_\varphi dv = h_\varphi$. Положим теперь в оставшемся интеграле $z = x - u$. Тогда $du = -dz$. Но это изменение знака компенсируется переменой пределов интегрирования, так что в результате получится:

$$\begin{aligned} \int h_\varphi(x-u) \cos(2\pi \omega u) du &= \int h_\varphi(z) \cos(2\pi \omega)(x-z) dz = \\ &= \cos(2\pi \omega x) \int h_\varphi(z) \cos(2\pi \omega z) dz + \\ &+ \sin(2\pi \omega x) \int h_\varphi(z) \sin(2\pi \omega z) dz = \\ &= \cos(2\pi \omega x) C_\varphi(\omega) + \sin(2\pi \omega x) S_\varphi(\omega), \end{aligned}$$

где $C_\varphi(\omega)$ и $S_\varphi(\omega)$ — одномерные косинусное и синусное преобразования функции рассеивания линии $h_\varphi(z)$. Пусть $M_\varphi(\omega) = [C_\varphi^2(\omega) + S_\varphi^2(\omega)]^{1/2}$ и $P_\varphi(\omega) = \arctg [S_\varphi(\omega)/C_\varphi(\omega)]$, т. е. $M_\varphi(\omega)$ и $P_\varphi(\omega)$ обозначают модуль и соответственно фазу преобразования Фурье функции $h_\varphi(z)$. Тогда искомую выходную функцию можно записать

$$k_\varphi a + b M_\varphi(\omega) \cos(2\pi \omega x - P_\varphi(\omega)).$$

Таким образом, выходная функция остается синусоидальной и имеет тот же период, что и входная функция $a + b \cos 2\pi \omega_x x$, но другие амплитуду и фазу. Функция $M_\varphi(\omega)$ называется амплитудной передаточной характеристикой, $P_\varphi(\omega)$ — фазовой передаточной характеристикой. Если операция φ не изотропна, эти характеристики, конечно, зависят от пространственной ориентации входной синусоиды.

Амплитудная передаточная характеристика соответствует амплитудно-частотной характеристике, используемой в теории электрических цепей. Она показывает, как в результате операции Φ ослабляется синусоидальный входной сигнал в зависимости от его пространственной частоты.

ЛИТЕРАТУРА

1. Cochrane W. T., et al., What is the fast Fourier transform? *Proc. IEEE*, 55, 1664 (Oct. 1967); русский перевод: Кокран, Кули и др., Что такое быстрое преобразование Фурье? *ТИИЭР*, № 10 (1967).
2. Cooley J. W., Lewis P. A. W., Welch P. D., Historical notes on the fast Fourier transform, *Proc. IEEE*, 55, 1675 (Oct. 1967); русский перевод: Кули, Льюис и др., Исторические замечания о быстром преобразовании Фурье, *ТИИЭР*, 10 (1967).
3. Lamberts R. L., Application of sine-wave techniques to image-forming systems, *J. Soc. Motion Picture Television Engrs.*, 71, 635 (Sept. 1962).
4. Smith F. D., Optical image evaluation and the transfer function, *Appl. Opt.*, 2, 335 (Apr. 1963).
5. Higgins G. C., Methods for engineering photographic systems, *Appl. Opt.*, 3, 1 (Jan. 1964).
6. Linfoot E. H., *Fourier Methods in Optical Image Evaluation*, Focal Press, New York, 1964.
7. Ronchi L., van Nes F. L., Contrast transfer in the eye as a function of spatial frequency: a literature survey, *Atti Fond. G. Ronchi*, 21, 218 (March—April 1966).

Методы реализации пространственно-инвариантных операций

5.1. Цифровые методы

Цифровые методы представляются наиболее естественными для выполнения свертки изображений, заданных в дискретной форме. Однако число умножений, которое при этом требуется произвести (даже не считая те, в которых один из сомножителей равен нулю), растет как m^2n^2 , т. е. очень быстро. Существующие вычислительные машины являются машинами последовательного действия, выполняющими в каждый момент времени только одну или максимум несколько арифметических операций. Если бы можно было выполнять вычисления над большим количеством чисел параллельно (т. е. одновременно), то это дало бы значительную экономию машинного времени, затрачиваемого на выполнение свертки дискретных изображений. Эту возможность позволяют реализовать «параллельные» вычислительные машины [1—8], имеющие достаточно много периферийных процессоров, управляемых центральным процессором (например, ИЛЛИАК IV имеет 256 процессоров).

Имеются специализированные вычислительные системы для параллельной обработки изображений. Они включают специальные устройства, которые могут выполнять операции одновременно над всеми элементами матрицы чисел (т. е. элементами дискретного изображения) [9—16]. Если имеется возможность сдвигать всю матрицу в любом из четырех основных направлений и производить одновременное поэлементное сложение или умножение двух матриц, то операции по свертке двух дискретных изображений можно выполнить очень быстро. Действительно, в этом случае для изображений из $m \times m$ и $n \times n$ элементов требуется $(m + n - 1)^2$ сдвигов и умножений (и примерно такое же число сложений), а не m^2n^2 , как

при последовательной обработке. На ЦВМ ИЛЛИАК III можно таким образом обрабатывать матрицы из 36×36 элементов.

Но даже на обычных (т. е. последовательных) вычислительных машинах можно выполнять простые логические операции и сдвиги одновременно с каждой двоичной цифрой машинного слова. Это дает экономию машинного времени в число раз, равное длине слова (как правило, она составляет 30... 50 двоичных цифр). Поскольку арифметические операции можно рассматривать как совокупность нескольких логических операций, большая часть этого выигрыша может быть достигнута и для них, в особенности если разрядность обрабатываемых чисел невелика (т. е. если изображения имеют небольшое число градаций яркости). Для целей обработки изображений создано несколько специальных языков программирования ([17—20]), которые используют преимущества такого подхода.

5.2. Оптико-электронные методы

Дискретные устройства для параллельной обработки массивов с большим числом элементов являются очень сложными и содержат большое количество компонентов. Известны другие методы выполнения некоторых операций, в частности пространственно-инвариантных операций, над изображениями, представленными в виде электрических сигналов (т. е. распределения потенциалов или заряда), которые формируются непосредственно распределением света на исходном изображении. В этом разделе описаны два таких метода.

Фотопроводящие электролюминесцентные панели [21, 22]

Существуют материалы — фотопроводники, — обладающие свойством изменять свое электрическое сопротивление под действием освещения. Слой такого материала можно нанести на некоторую поверхность и создать тем самым как бы сетчатку из небольших фоточувствительных элементов. Если на такую сетчатку спроектиро-

вать изображение, то она преобразует его в соответствующее поверхностное распределение электрического тока. Эту сетчатку можно связать с электролюминесцентным экраном так, чтобы яркость свечения экрана в каждой точке управлялась разностью потенциалов в соответствующем элементе сетчатки (скажем, была пропорциональна ей). При этом изображение будет вновь восстановлено.

Предположим, что экран и сетчатка выполнены в виде «сэндвича», так что сетчатка управляет свечением экрана, а излучаемый экраном свет попадает снова на сетчатку. Таким образом, изображение на экране возбуждает сетчатку, которая в свою очередь вызывает на экране то же изображение (при необходимости усиленное). Такая фоточувствительная электролюминесцентная панель может использоваться для запоминания изображения.

Если на сетчатку проектировать несколько изображений, они будут складываться на экране. Следовательно, такое устройство позволяет легко реализовать сложение изображений. Заметим, что разность потенциалов на элементах сетчатки можно легко инвертировать, что при наличии соответствующего положительного смещения приведет к преобразованию изображения в негатив, поэтому возможно не только сложение, но и вычитание изображений. Для умножения изображения в таком устройстве на положительную или отрицательную константу достаточно просто умножить (и, соответственно, инвертировать) все разности потенциалов сетчатки. Складывая смещенные копии изображения, умноженные на соответствующие константы, можно осуществить произвольное преобразование вида

$$\sum a_i f(x + u_i, y + v_i).$$

Иными словами, можно реализовать любые пространственно-инвариантные операции над изображениями. Преобразуя нелинейно разности потенциалов сетчатки, можно производить и различные нелинейные операции над изображениями.

Электронные преобразователи изображений

При освещении некоторых фоточувствительных поверхностей возникает явление фотоэлектронной эмиссии. Число электронов, эмиттируемых некоторым участком такой поверхности, зависит от интенсивности падающего света. Если вдоль пути этих электронов сформировать соответствующее электрическое поле, их можно заставить лететь по параллельным траекториям на некоторую мишень, где они создадут распределение заряда, соответствующее распределению интенсивности света на фоточувствительном слое. Таким образом, на мишени можно запомнить изображение, спроектированное на фоточувствительный слой, а также складывать на ней изображения друг с другом. Для превращения изображений в негатив можно увеличить напряженность электрического поля, ускоряющего фотоэлектроны, до такого уровня, когда при попадании на мишень электроны станут выбивать более одного вторичного электрона. Тогда увеличение числа фотоэлектронов, вылетающих с данного участка фоточувствительного слоя, будет приводить к уменьшению заряда в соответствующей точке мишени. Изменяя напряженность ускоряющего поля, можно легко осуществлять умножение изображения на константу. Смещение изображения (отклонение) можно осуществлять с помощью поперечной составляющей поля. (Заметим, что такой способ отклонения не требует механического перемещения и поэтому отклонение может производиться очень быстро.) Таким образом, в этом устройстве также могут быть просто реализованы любые пространственно-инвариантные операции над изображением.

5.3. Оптические методы, не требующие когерентного освещения

Особенно просто пространственно-инвариантные операции над изображениями реализуются оптическими средствами. Хотя эта книга посвящена обработке изобра-

жений на вычислительных машинах, ввиду возрастающего интереса к оптическим и гибридным оптико-цифровым вычислительным устройствам желательно рассмотреть также и их возможности.

При оптической обработке изображений обычно предполагается, что изображения представлены в виде *транспарантов* и характеризуются распределением коэффициента пропускания транспаранта $f(x, y)$, т. е. отношения количества света, прошедшего через транспарант в точке (x, y) , к количеству падающего света. Заметим, что коэффициент пропускания — величина неотрицательная. Ниже будут рассмотрены способы выполнения операции вычитания и представления функций с отрицательными значениями.

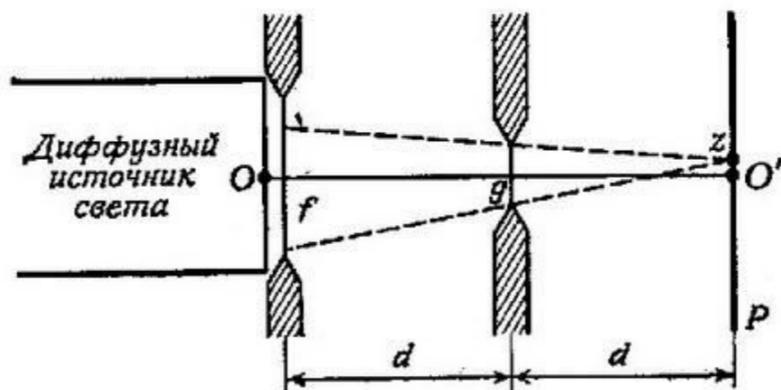
Изображение, которое хранится в памяти вычислительной машины (или в виде электрических сигналов), можно воспроизвести на транспаранте с помощью электроннолучевой трубки. Обычно в таких случаях для регистрации изображения предпочтительнее использовать такие среды, как фотохромные материалы, не требующие в отличие от обычных фотоматериалов обработки. (Другие способы воспроизведения изображений из электрических сигналов описаны в [29, 30].) Изображение на выходе оптической системы можно преобразовать в электрический сигнал с помощью телевизионной камеры или аналогичных сканирующих устройств [31—34].

Умножить одно изображение на другое можно, просто наложив их транспаранты друг на друга, поскольку коэффициент пропускания сложенных вместе транспарантов равен произведению коэффициентов пропускания каждого из них. Сложить два изображения f и g можно, совместив проекции их транспарантов. В этом случае количество света в каждой точке такой составной проекции будет равно сумме количеств света в каждой составляющей проекции и, значит, пропорционально сумме коэффициентов пропускания транспарантов в соответственных точках. Следовательно, для получения транспаранта fg можно снять контактную копию f и g , а транспарант $f + g$ можно получить, проэкспонировав регистрирующую среду сначала через транспарант f , а затем через транспарант g .

В следующем разделе описан ряд методов формирования взаимной корреляции (или, что то же самое, свертки¹) двух изображений чисто оптическими средствами. Можно, разумеется, получить взаимную корреляцию простым сдвигом транспарантов относительно друг друга, измеряя сумму получающихся при этом произведений коэффициентов пропускания транспарантов. Однако такой коррелятор должен содержать движущиеся части и потому малопроизводителен. Как будет ясно из дальнейшего, оптическими средствами можно получить взаимную корреляцию практически мгновенно без помощи каких-либо движущихся устройств [35—42].

Безлинзовый коррелятор

Исключительно простой метод получения взаимной корреляции двух функций показан на фиг. 5.1. Чтобы упростить изложение, рассмотрим этот метод для функций одной переменной. Пусть f и g равны нулю за пределами интервалов $2r$ и $2s$ соответственно. Как показано на



Ф и г. 5.1. Безлинзовый коррелятор.

фигуре, транспарант, представляющий g , имеет вдвое меньший масштаб, так что его коэффициент пропускания в точке, находящейся от оси OO' на расстоянии y , равен $g(2y)$. Масштабный коэффициент равен 2, так как g

¹ Очевидно, свертку двух функций можно получить как взаимную корреляцию одной из них с зеркальной (т. е. повернутой на 180°) копией другой.

находится на равном расстоянии d от f и плоскости P . В общем случае, если расстояния $g-f$ и $g-P$ равны соответственно d_1 и d_2 , отношение масштаба на транспаранте g к масштабу на транспаранте f равно $d_2/(d_1+d_2)$.

Лучи света от диффузного источника проходят через каждую точку транспаранта f во всех направлениях. Часть из них проходит и через g . Луч, который попадает на плоскость P в точке z и проходит через g в точке y , очевидно, проходит через f в точке $2y-z$. Следовательно, общее количество света в точке z определяется величиной $\int_{-s/2}^{s/2} g(2y) f(2y-z) dy$. Заменим в этом интервале $w = 2y$. Поскольку $g(w) = 0$ при $|w| > s$, интегрирование можно производить в пределах $(-\infty, \infty)$, т. е. этот интеграл становится равным $1/2 \int_{-\infty}^{\infty} f(w-z) g(w) dw$. Таким образом, распределение света на плоскости P как функции z пропорционально взаимной корреляции f и g .

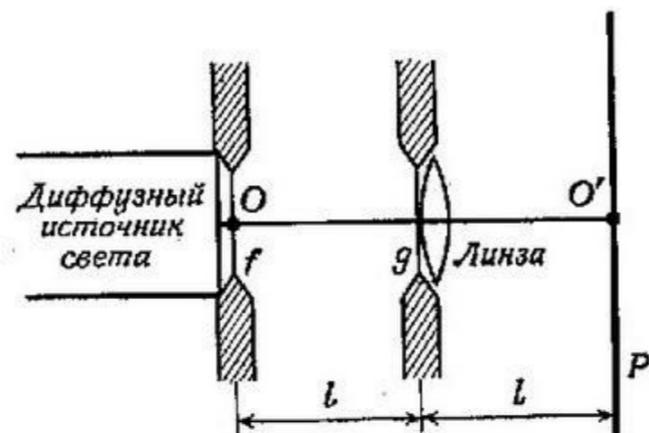
Однолинзовые корреляторы

Материал этого раздела основан на следующих свойствах идеальных линз.

Пусть L — выпуклая сферическая линза (т. е. линза, обе поверхности которой являются шаровыми сегментами). Линия, которая соединяет центры сфер, называется *оптической осью* линзы L . Если на линзу падает параллельный пучок лучей, то, пройдя ее, они сфокусируются в одной точке. Фокальные точки всех таких пучков располагаются в одной плоскости, перпендикулярной оптической оси. Эта плоскость называется *фокальной плоскостью* линзы. Положение фокальной точки на фокальной плоскости зависит от ориентации пучка относительно оптической оси и определяется точкой пересечения фокальной плоскости с линией, параллельной направлению пучка и проходящей через центр линзы. Расстояние от линзы (толщиной линзы будем пренебрегать) до фокальной плоскости называется *фокусным расстоянием* линзы. Будем обозначать его буквой l (обычно оно обозначается буквой F , но здесь F используется для обозначения фурье-преобразования функции f). Если лучи света проходят

через точку, расположенную на расстоянии l от линзы L , то они преобразуются линзой в параллельный пучок, направление которого совпадает с линией, соединяющей эту точку с центром L .

Наконец, пусть (x, y) — некоторая точка в плоскости, перпендикулярной оптической оси и отстоящей от L



Ф и г. 5.2. Однолинзовый коррелятор.

на расстояние u , причем $u > l$. Все световые лучи, проходящие через точку (x, y) , соберутся линзой в единственной точке в некоторой плоскости изображения, перпендикулярной оптической оси и расположенной с противоположной стороны линзы на расстоянии v от нее, причем v и u связаны соотношением

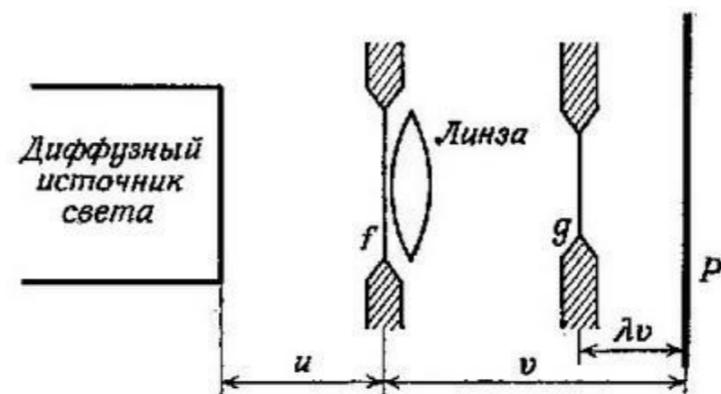
$$\frac{1}{u} + \frac{1}{v} = \frac{1}{l}.$$

Координаты этой точки в плоскости изображения равны

$$\left(-\frac{vx}{u}, -\frac{vy}{u} \right).$$

На фиг. 5.2 показана схема оптического коррелятора, в котором оба транспаранта имеют одинаковый масштаб. На этой фигуре l — фокусное расстояние линзы, так что любой пучок параллельных лучей, падающий на линзу, фокусируется в одну точку в плоскости P . Рассмотрим, в частности, луч, проходящий через точку x транспаранта f под углом θ к оси OO' . Этот луч попадает на транс-

спарант g в точке $y = x + l \operatorname{tg} \theta$ и фокусируется в точке $z = l \operatorname{tg} \theta$ в плоскости P . Общее количество света в этой точке равно поэтому $\int_{-r}^r f(x) g(x + l \operatorname{tg} \theta) dx = \int_{-r}^r f(x) g(x + z) dx$. В этом интеграле снова можно заменить конечные пределы интегрирования бесконечными и получить выражение для взаимной корреляции f



Ф и г. 5.3. Другой вариант однолинзового коррелятора.

и g . (Если расстояние между двумя транспарантами не равно l , в плоскости P все равно будет получаться их взаимная корреляция, но уже не в масштабе 1:1.)

Другая схема коррелятора, пригодная для транспарантов с различным масштабом, показана на фиг. 5.3. Здесь $1/u + 1/v = 1/l$, так что линза собирает все лучи, испускаемые из каждой точки источника, в одной точке плоскости P .

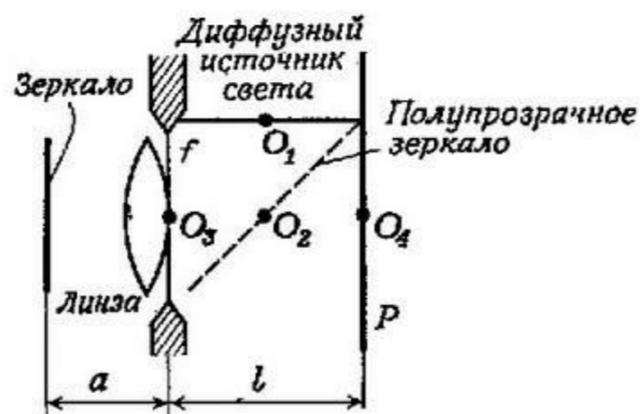
Рассмотрим лучи, испускаемые из точки с координатой x . Эти лучи, пройдя транспарант f и линзу, образуют пучок, сходящийся в точке с координатой $-vx/u$ в плоскости P . Яркость в некотором сечении этого пучка, очевидно, пропорциональна функции f , взятой в масштабе, который уменьшается по мере приближения рассматриваемого сечения к плоскости P . Пусть отношение масштабов g и f равно λ ($\lambda < 1$) и g находится на расстоянии λv от P . Световой луч, выходящий из точки x источника и проходящий через точку y транспаранта f , проходит через точку $\lambda y - (1 - \lambda) \frac{vx}{u}$ транспаранта g . Поэтому общее количество света, попадающего в точку $-vx/u$ плоскости

P , равно $\int f(y) g(\lambda [y - (1 - \lambda) \frac{ux}{\lambda u}]) dy$, т. е. в плоскости P получается в некотором масштабе взаимная корреляция g и f .

Описанная схема однолинзового коррелятора позволяет простыми средствами преобразовать коррелограмму в электрический сигнал для дальнейшей обработки. Для этого нужно заменить диффузный источник света электроннолучевой трубкой (или другим устройством развертки с бегущим лучом), в которой в каждый данный момент времени засвечивается только одна точка. При этом на коррелограмме также будет засвечиваться только одна точка, так что фотоэлемент, помещенный в плоскости изображения, будет давать выходной сигнал, соответствующий развертке коррелограммы.

Однолинзовый автокоррелятор с одним транспарантом

Для оптического формирования функции автокорреляции изображения можно просто изготовить два транспаранта и воспользоваться любой из описанных выше схем



Фиг. 5.4. Однолинзовый коррелятор с одним транспарантом.

корреляторов. Но функцию автокорреляции можно получить и с единственным транспарантом, используя схему, показанную на фиг. 5.4. Здесь $O_1O_2 + O_2O_3 = O_3O_2 + O_2O_4 = l$, где l — фокусное расстояние линзы, так что лучи света из любой точки x источника после отражения от полупрозрачного зеркала выходят из линзы в виде пучка параллельных лучей, направленных под

углом $\theta = \arctg x/l$ к оси O_3O_4 . Эти лучи попадают на зеркало и отражаются обратно на транспарант. Лучи, прошедшие через транспарант в точке y , после отражения попадают в точку $y + 2a \operatorname{tg} \theta = y + \frac{2a}{l} x$. Эти лучи после отражения по-прежнему параллельны, но направлены под углом $-\theta$ к оси O_3O_4 . Поэтому они фокусируются (после прохождения через полупрозрачное зеркало) в плоскости P в точке с координатой $z = l \operatorname{tg} (-\theta) = -x$. Общее количество света, попадающего в эту точку, равно $\int f(y) f(y + \frac{2a}{l} x) dy$, т. е. функции автокорреляции в некотором масштабе по оси x .

Все описанные схемы, основанные на использовании диффузных источников света, имеют тот недостаток, что яркость световых лучей, испускаемых источником в различных направлениях, неодинакова. Чем больше угол между лучом и перпендикуляром к поверхности источника, тем яркость меньше. Для схем на фиг. 5.2 и 5.4 это означает, что участки коррелограммы, соответствующие большим смещениям, темнее. С другой стороны, в схемах на фиг. 5.1 и 5.3 яркость лучей, приходящих в одну точку коррелограммы, неодинакова, так что общая освещенность точки только приблизительно представляется интегралом корреляции. Эти приближения, однако, достаточно точны, если углы участвующих в формировании коррелограммы лучей невелики.

Рассмотрим теперь вопрос о том, как оптическими средствами можно обрабатывать функции, которые не обязательно являются неотрицательными. Пусть g — одна из таких функций. Определим положительные и отрицательные части следующим образом:

$$g^+ = g \text{ при } g \geq 0, \quad g^+ = 0 \text{ при } g < 0,$$

$$g^- = -g \text{ при } g \leq 0, \quad g^- = 0 \text{ при } g > 0.$$

Таким образом, g^+ и g^- — неотрицательные функции, которые можно представить в виде транспаранта. Поскольку $g = g^+ - g^-$, то если бы можно было оптически вычитать изображения, можно получить и любое данное g . Далее, для любого f $g \otimes f = g^+ \otimes f - g^- \otimes f$. Поэтому любую пространственно-инвариантную операцию над f

можно выполнить путем свертки двух (неотрицательных) функций s и f и вычитания результатов. В следующем разделе описаны некоторые методы вычитания функций.

Положительное смещение. Пусть c — максимальное значение функции g , тогда $c - g$ — неотрицательная функция, которую можно представить в виде транспаранта — фотографического негатива g . Следовательно, для любого изображения f имеем $f + c - g = (f - g) + c$. Иными словами, вычтем g из f (с точностью до константы) можно, сложив f и негатив g . В более общей формулировке пусть h — произвольная функция, а c — максимальное значение h . Тогда $h + c = h^+ + (c - h^-)$ — неотрицательная функция. Следовательно, любая функция может быть с точностью до константы представлена в виде транспаранта. Так как $(h + c) \otimes f = h \otimes f + c \int f$, то с точностью до константы можно выполнить любую пространственно-инвариантную операцию над f (только константа теперь будет зависеть от f). Основным недостатком такого метода состоит в том, что добавление константы уменьшает диапазон величин, которые могут быть представлены на транспаранте. В частности, если транспарант почти полностью прозрачен, трудно передать значения $(f - g) + c$, для которых требуется еще большая прозрачность. Эта проблема становится особенно острой, если нужно выполнить последовательность пространственно-инвариантных операций, в каждой из которых вводится дополнительная константа.

Вычитание электронными методами. Разность двух изображений можно найти, вычитая электрические сигналы, полученные при синхронной развертке этих изображений. Предположим, в частности, что формируемые оптическими средствами коррелограммы f и g^+ и f и g^- сканируются синхронно бегущим лучом, а не освещаются диффузным источником света. Тогда $g \otimes f$ можно получить, вычитая получаемые при этом электрические сигналы. На самом деле нет необходимости строить два отдельных оптических коррелятора; g^+ и g^- могут быть совмещены в одном транспаранте с помощью цветоделения или методов разделения по поляризации. Предположим, например, что в схему однолинзового коррелятора (фиг. 5.3) между вторым транспарантом и плоскостью P введено

полупрозрачное зеркало, как показано на фиг. 5.5. В этом случае половина светового пучка отразится от этого зеркала на плоскость P' . Пусть H и V — пластинки поляризаторов, выделяющие соответственно вертикальную и горизонтальную составляющие поляризации, а транспарант g выполнен также из поляризаторов: в местах, где $g^+ > 0$, — вертикального, а там, где $g^- > 0$, — горизонтального. Вертикальный поляризатор прозрачен для



Фиг. 5.5. Метод разделения для вычитания изображений.

света, прошедшего также через вертикальный поляризатор, и непрозрачен для света, прошедшего через горизонтальный поляризатор, и наоборот. Следовательно, если смотреть на транспарант g с плоскости P , то там, где $g^- > 0$, он непрозрачен, его прозрачность определяется g^+ , так что в плоскости P формируется взаимная корреляция $g^+ \otimes f$. Точно так же в плоскости P' имеем $g^- \otimes f$. Если источник света заменить бегущим лучом, эти коррелограммы будут сканироваться синхронно, так что полученные при этом сигналы можно при необходимости вычитать. Подобные результаты получатся, если выполнить g не из поляризационных материалов, а из цветных фильтров с непересекающимися функциями пропускания. Каждый из таких фильтров непрозрачен для света, прошедшего через другой фильтр, и можно использовать ту же схему фиг. 5.5, заменив V и H цветными фильтрами.

Главный недостаток этих методов заключается в том, что они требуют изготовления транспаранта g в виде мозаики из двух различных материалов, что весьма за-

труднительно, если мозаика должна состоять из большого числа фрагментов. Другие методы, также использующие поляризацию, но лишенные этого недостатка, описаны в следующем разделе.

Методы фотографического обращения. Если на экспонированную, но не проявленную фотопленку попадает длинноволновое световое излучение, скрытое изображение на пленке стирается, причем степень этого стирания пропорциональна интенсивности излучения. Это явление известно как эффект Гершеля. Предположим, что пленка экспонируется (в обычном свете) транспарантом f , а затем до проявления экспонируется (на этот раз в длинноволновом свете) транспарантом g . Полученное при этом скрытое изображение представляет неотрицательную составляющую функции $af - bg$, где a и b — константы, зависящие от времени экспозиции транспарантов. Следовательно, функцию $(g \otimes f)^+$ можно получить, если дважды проэкспонировать пленку в выходной плоскости оптического коррелятора: сначала в обычном свете, сформировав на выходе коррелятора функцию $g^+ \otimes f$, а затем в длинноволновом излучении, сформировав $g^- \otimes f$. (В принципе и в этом способе можно совместить во времени обе экспозиции, если выполнить транспарант g из материала двух типов, каждый из которых непрозрачен для излучения одного диапазона и прозрачен для другого, и освещать транспарант одновременно в обоих диапазонах.)

Хотя этот метод, как и метод, описываемый ниже, позволяет получить только $(g \otimes f)^+$, для многих применений этого совершенно достаточно. Например, если g — фильтр, выделяющий контуры (разд. 6.5), то для обнаружения контуров в f достаточно положительной составляющей свертки $g \otimes f$.

Люминофоры с управляемым гашением. Существуют люминофоры, которые светятся под действием ультрафиолетового облучения и гаснут при инфракрасном свете. Если экран из такого люминофора с управляемым гашением освещать ультрафиолетовым светом, пропущенным через транспарант f , и инфракрасным светом, пропущенным через транспарант g , то его результирующее свечение будет представлять функцию $(af - bg)^+$, где константы a и b зависят от яркости соответствующих

источников излучения. Отсюда следует, что $(g \otimes f)^+$ можно получить путем оптического совмещения на этом экране изображений, представляющих в ультрафиолетовом и соответственно в инфракрасном свете функции $g^+ \otimes f$ и $g^- \otimes f$. (Это также можно выполнить в схеме с одним оптическим коррелятором, разместив в его выходной плоскости люминесцентный экран и используя транспарант g , выполненный из двух материалов, один из которых непрозрачен в ультрафиолетовой, а другой в инфракрасной области спектра.)

5.4. Когерентно-оптические методы

Для понимания этого раздела потребуются основные сведения о волновой природе света. Луч света может рассматриваться как синусоидальное колебание с некоторой длиной волны λ , определяющей ее цвет. Если эта волна распространяется со скоростью c , ее частота (число колебаний в секунду) равна c/λ . Таким образом, световые колебания в некоторой точке пространства можно представить выражением $Ae^{i(2\pi ct/\lambda + \varphi)}$. Интенсивность колебаний, как можно показать, пропорциональна A^2 . Отметим также, что при перемещении вдоль луча на расстояние z фаза колебаний изменяется на величину $2\pi z/\lambda$.

При прохождении света через объект изменяются и амплитуда и фаза световой волны, так что она превращается в колебание вида

$$K A e^{i(2\pi ct/\lambda + \varphi + \psi)}.$$

Очевидно, если параллельный пучок света проходит через объект с постоянной толщиной, все лучи изменяются одинаково: фаза каждого луча после прохождения объекта изменится на константу, и эта константа будет одна и та же для всех лучей. Можно показать также, что если параллельный пучок лучей проходит через линзу, он будет сфокусирован в точку и фаза каждого луча в фокусе будет отличаться на константу (одну и ту же для всех лучей) от его фазы в точке пересечения с произвольной плоскостью, расположенной до линзы перпендикулярно ее оптической оси. (Различие в длине пути лучей

при прохождении ими участков линзы с различной толщиной полностью компенсируется различием в длине пути лучей от линзы до точки фокуса.)

Параллельный пучок света называется когерентным, если фазы всех лучей в плоскости, перпендикулярной направлению пучка, равны. Когерентность пучка сохраняется после прохождения им объекта постоянной толщины. Если такой пучок фокусируется линзой в точку, все лучи его приходят в эту точку с одной и той же фазой.

Два световых луча одной и той же длины волны, имеющие различные фазы, интерферируют между собой. В частности, если их фазы отличаются на π , лучи могут полностью компенсировать друг друга, поскольку

$$Ae^{j(2\pi ct/\lambda + \varphi + \pi)} = e^{j\pi} Ae^{j(2\pi ct/\lambda + \varphi)},$$

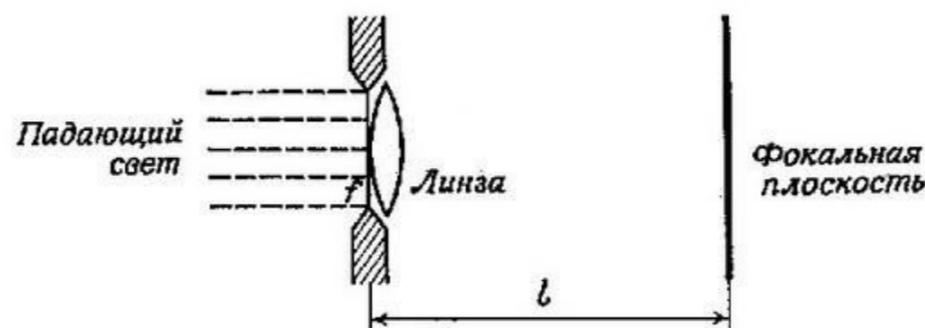
а $e^{j\pi} = -1$. Явление интерференции препятствует распространению лучей света во всех направлениях. Хотя каждую точку поперечного сечения пучка света можно рассматривать как точечный источник, излучающий сферические волны, волны, излучаемые всеми точками, суммируются, и это приводит к тому, что во всех направлениях, кроме первоначального направления пучка, они компенсируют друг друга. С другой стороны, если экранировать некоторые части пучка, эта компенсация будет происходить не во всех направлениях и часть пучка отклонится от первоначального направления — дифрагирует. Предположим, например, что пучок когерентного света проходит через синусоидальную дифракционную решетку, т. е. пластину постоянной толщины, коэффициент пропускания которой меняется по координате x по синусоидальному закону. Можно показать, что в этом случае свет дифрагирует в двух симметричных направлениях, отклоняясь от первоначального направления на угол, зависящий от частоты решетки и длины волны света.

Преобразование Фурье в фокальной плоскости

Предположим, что транспарант f (фиг. 5.6) освещается пучком когерентного света с длиной волны λ , параллельным оптической оси. Таким образом, все лучи, падающие

на плоскость f , имеют одинаковую амплитуду E_0 и фазу φ_0 . Пусть свет, проходящий через f в точке (x, y) , ослабляется в $f(x, y)$ раз и задерживается по фазе на величину $\varphi(x, y)$, так что его амплитуда и фаза после прохождения f равны соответственно $f(x, y)E_0$ и $\varphi(x, y) + \varphi_0$.

Наличие на пути лучей транспаранта f приводит к их дифракции. Рассмотрим часть света, дифрагирующую вдоль прямой, направляющие косинусы которой (по отношению к осям x и y) равны соответственно a_1 и a_2 . Эти световые лучи сфокусируются линзой в плоскости



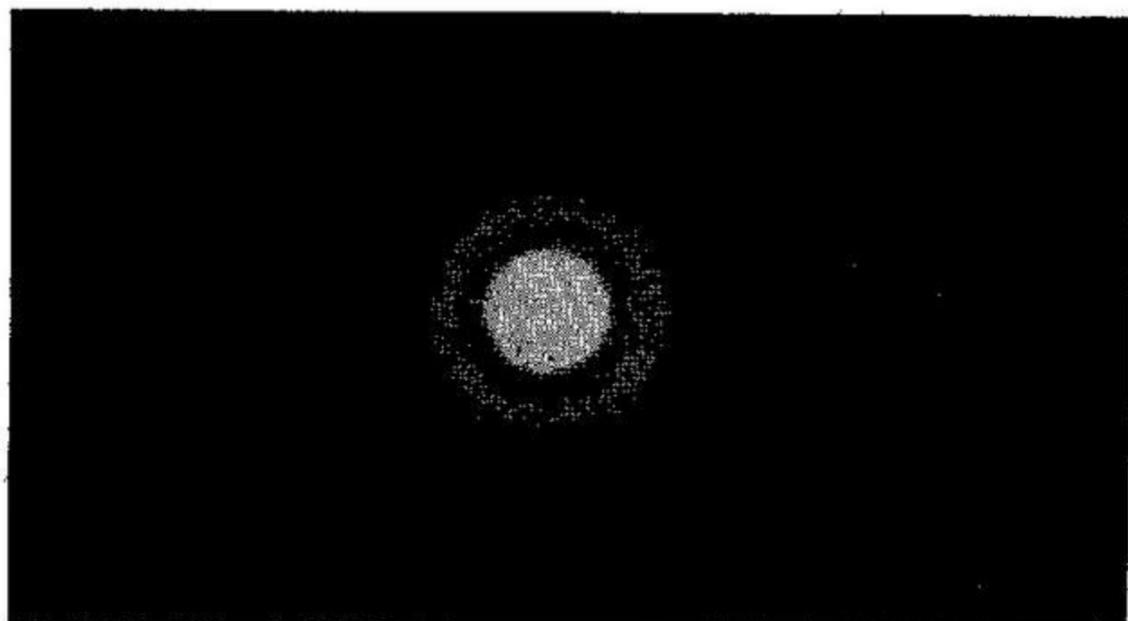
Ф и г. 5.6. Преобразование Фурье в фокальной плоскости линзы.

изображения в точку, координаты которой w_x и w_y пропорциональны a_1 и a_2 . Коэффициент пропорциональности между ними равен l/a_3 , где a_3 определяется из соотношения $a_1^2 + a_2^2 + a_3^2 = 1$, а l — фокусное расстояние линзы. Пусть P — проходящая через центр f плоскость, нормаль к которой имеет направляющие косинусы a_1 и a_2 . Тогда фаза луча, проходящего через точку (x, y) при пересечении им плоскости P , равна $\varphi_0 + \varphi(x, y) + 2\pi \frac{a_1 x + a_2 y}{\lambda}$. Но, как указано в начале этого раздела, фаза луча, проходящего через точку (x, y) , в точке (w_x, w_y) отличается на константу от его фазы в месте пересечения плоскости P . Поэтому амплитуда света в точке (w_x, w_y) пропорциональна

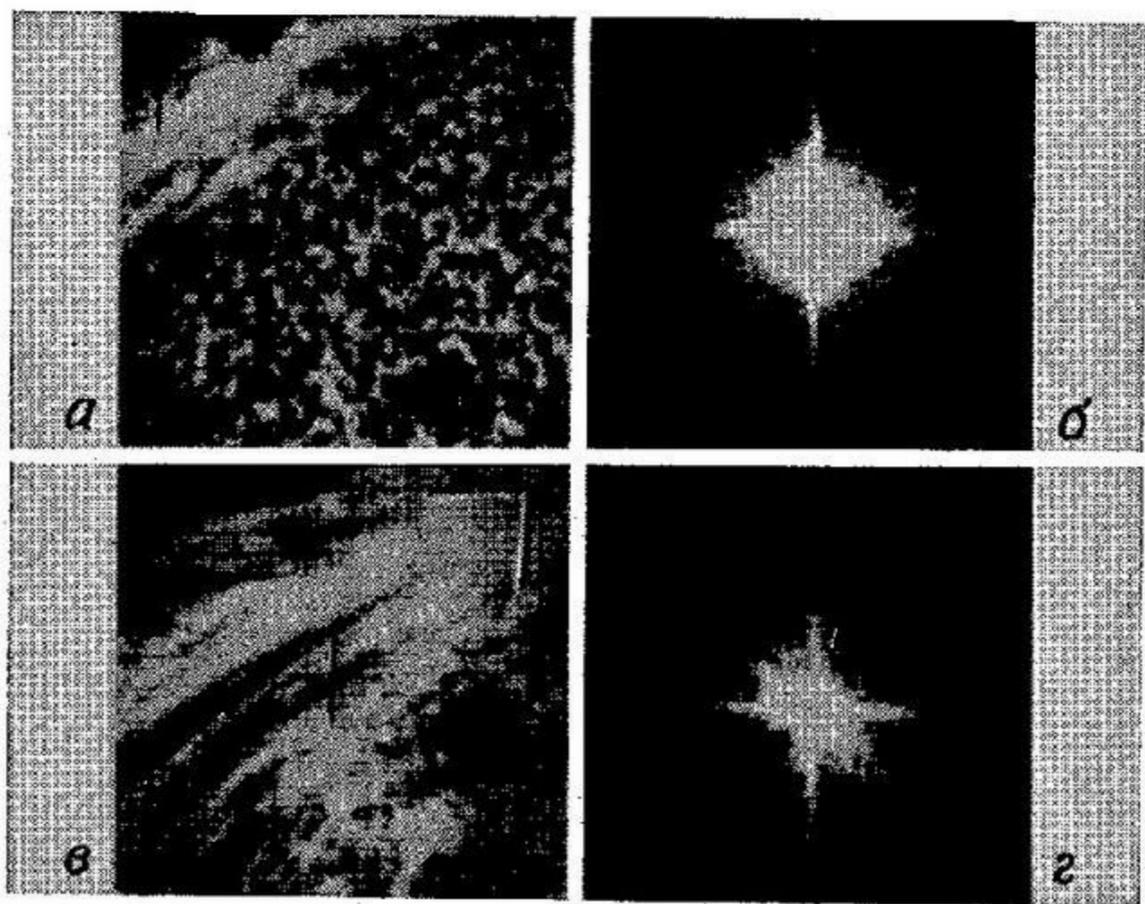
$$\iint f(x, y) e^{j[\varphi(x, y) + 2\pi(a_1 x + a_2 y)/\lambda]} dx dy.$$

Второй член в квадратных скобках в этом выражении равен $2\pi a_3 (w_x + w_y)/l\lambda$, так что интеграл переходит в

$$\iint [f(x, y) e^{j\varphi(x, y)}] e^{2\pi a_3 i (w_x x + w_y y)/\lambda l} dx dy.$$



Ф и г. 5.7. Энергетический спектр круглого отверстия [$f(x, y) = 1$ при $0 \leq (x^2 + y^2)^{1/2} \leq 0,1$ мм и $f(x, y) = 0$ при остальных значениях x, y], зафиксированный в фокальной плоскости линзы.



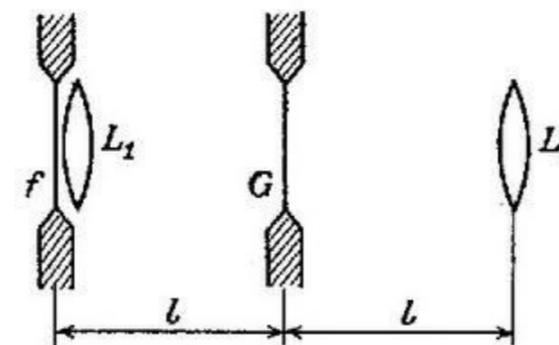
Ф и г. 5.8. Снимки облачного покрова (полученные с ИСЗ «Тирос») и их энергетический спектр.

Если изображение содержит много мелких деталей и контуров, его спектр простирается достаточно далеко от начала координат, что соответствует наличию относительно высоких пространственных частот (а и б). Если контуры на изображении ориентированы преимущественно в одном направлении, спектр оказывается существенно неизотропным (в и г).

Если дифракцией под большими углами пренебречь, можно принять, что $a_z \equiv 1$ и не зависит от (ω_x, ω_y) , так что этот интеграл является преобразованием Фурье комплексной функции $f(x, y) e^{i\varphi(x, y)}$, заданным на плоскости $\omega_x/l\lambda, \omega_y/l\lambda$. Интенсивность света в фокальной плоскости представляет, таким образом, квадрат модуля его фурье-преобразования, т. е. энергетический спектр f . Следовательно, поместив в фокальную плоскость фотопластинку, можно получить транспарант с записью энергетического спектра f . На фиг. 5.7 показан пример зарегистрированного таким способом энергетического спектра отверстия, расположенного на оптической оси системы, а на фиг. 5.8 — полученные этим же способом энергетические спектры изображений облачного покрова.

Применение когерентно-оптических методов для нахождения свертки [44—52]

Описанную схему преобразования Фурье можно использовать для выполнения пространственно-инвариантных операций над изображениями. Действительно, для



Ф и г. 5.9. Использование симметричной оптической системы для комплексной пространственной фильтрации.

свертки изображений f и g можно поместить в фокальную плоскость системы транспарант G , представляющий в соответствующем масштабе фурье-преобразование g (изготовление такого транспаранта будет описано ниже). Поскольку распределение падающего света в этой плоскости представляет собой преобразование Фурье транспаранта f , по другую сторону транспаранта G получаем

произведение FG преобразований Фурье функций f и g . Если теперь на расстоянии l от плоскости G поместить линзу L_2 , идентичную L_1 (фиг. 5.9), то в силу симметрии распределение света после этой линзы будет связано с распределением света, прошедшего через транспарант G , так же, как распределение света, падающего на G , связано с распределением света, попадающего на L_1 . Следовательно, линзой L_2 формируется обратное преобразование Фурье произведения FG , т. е. свертка функций f и g .

Реализация преобразования Фурье в виде транспаранта

В предыдущем разделе предполагалось, что G является комплексным преобразованием Фурье функции g . Как уже отмечалось, в принципе нетрудно изготовить транспарант, представляющий энергетический спектр данной функции. Значительно труднее передать на транспаранте соответствующую фазовую информацию. Ниже описаны некоторые возможные методы регистрации распределения фазы.

1. Рельеф [53]. При прохождении света через транспарант фаза световой волны изменяется на величину, пропорциональную толщине транспаранта. Следовательно, если бы можно было изготовить транспарант, прозрачность которого и толщина менялись по его площади по заданному закону, можно было бы в принципе зарегистрировать на нем модуль и фазу комплексной функции. На практике иногда достаточно даже только приблизительно передать фазовую информацию, используя несколько дискретных значений толщины.

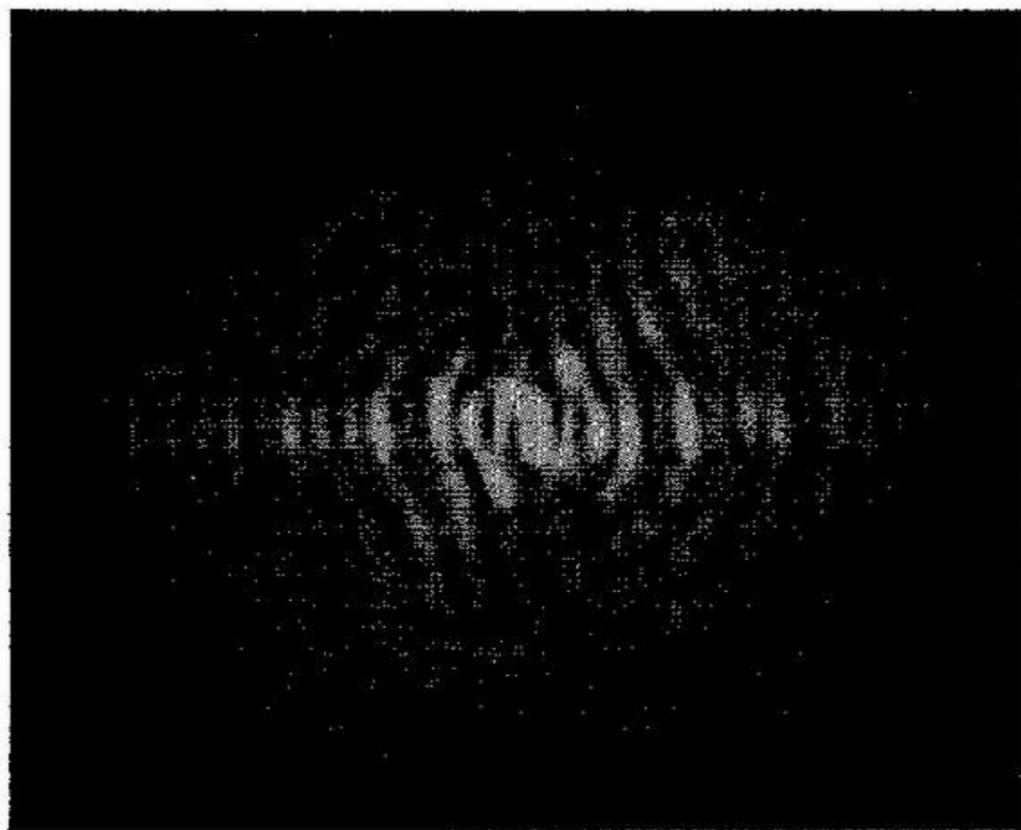
2. Разделение по поляризации [54]. Предположим, что фурье-преобразование G является действительной функцией (такое предположение оправданно, например, в тех практически важных случаях, когда функция g симметрична по отношению к началу координат). Конечно, это не исключает возможности отрицательных значений G , но в этом случае G можно разделить на положительную и отрицательную части, как описано в разд. 5.3, и воспользоваться соответствующими методами вычитания. Поскольку освещение должно быть когерентным и, следова-

тельно, монохроматическим, нельзя использовать метод цветоделения, применение же метода разделения по поляризации возможно. Наличие когерентного света позволяет применить «одноканальный» вариант этого метода. Предположим, что транспарант G изготовлен из поляризующего вещества (разд. 5.3), и пусть падающий на него пучок света поляризован под углом 45° к направлению одного из поляризаторов транспаранта, а прошедший через транспарант свет пропускается через анализатор, установленный на угол -45° . Нетрудно показать, что в этом случае после анализатора составляющие пучка, прошедшие через G^+ и G^- , будут поляризованы под углом -45° и противофазны, то есть будут вычитаться друг из друга. Можно также изготовить транспарант, не разделяя его на отдельные фрагменты G^+ и G^- , а используя двухслойный фотоматериал, известный под названием пленки Вектограф. Если на один слой такой пленки записать некоторую функцию, ее коэффициент пропускания будет зависеть от поляризации падающего света. Для некоторого направления поляризации она будет полностью прозрачна. Поэтому можно записать G^+ на один слой, полностью прозрачный для света, скажем с горизонтальной поляризацией, а G^- — на другой, прозрачный для света, поляризованного в вертикальном направлении.

3. Дифракционные решетки [55, 56]. Если заэкранировать часть пучка когерентного света, свет будет дифрагировать, причем модуль и фаза дифрагированного света будут зависеть от конфигурации экрана. Изменяя эту конфигурацию, можно в принципе реализовать любую нужную комплексную функцию не только в виде некоторого полутонового изображения, но и в виде двухградационного изображения, т. е. распределения прозрачных и непрозрачных участков. Рассмотрим в качестве простого примера одномерный случай, когда пучок света экранируется дифракционной решеткой в виде параллельных непрозрачных полосок, разделенных прозрачными промежутками. Пусть сечение пучка разделяется на полоски шириной δ . Выделим в пределах каждой полоски прозрачные щели шириной $\alpha_n \delta$, расположенные на расстоянии β_n от оси n -й полоски, причем α_n и β_n выберем таким образом, чтобы из падающего на решетку под некоторым углом

входного параллельного пучка света $Ae^{j\omega x}$ получить заданные амплитуду и фазу выходной волны $A_n e^{j\varphi_n}$. Так как комплексная амплитуда волны на выходе n -й щели равна

$$\int_{[n + \beta_n - (\alpha_n/2)] \delta}^{[n + \beta_n + (\alpha_n/2)] \delta} Ae^{j\omega x} dx = \left(\frac{2A}{\omega} \right) e^{j\omega (n + \beta_n) \delta} \sin \left(\alpha_n \frac{\omega \delta}{2} \right),$$



Ф и г. 5.10. Результат интерференции фурье-преобразования цифры 3 и опорного сигнала.

то достаточно выбрать α_n и β_n так, чтобы обеспечить выполнение условий

$$\frac{2A}{\omega} \sin \frac{\alpha_n \omega \delta}{2} = A_n$$

и

$$\omega (n + \beta_n) \delta = \varphi_n.$$

Модуляция пространственной несущей [57—69]. Если транспарант, представляющий комплексное фурье-преобразование G функции g , не нужен, а требуется только выполнить свертку функции g с другой функцией, можно использовать транспарант, на котором G смешана с некоторым опорным сигналом и который можно отделить при свертке. Действительно, предположим, что G формируется оптически в фокальной плоскости и эта плоскость освещается также параллельным пучком когерентного света R , падающего на плоскость под некоторым углом, $R = |R| \times e^{2\pi j(a\omega_x + b\omega_y)}$, где $|R|$, a и b — константы, а ω_x , ω_y — координаты в фокальной плоскости. Интенсивность суммарного светового потока, которую можно зарегистрировать на транспаранте, равна тогда $|G + R|^2 = |G|^2 + |R|^2 + GR^* + G^*R$, где звездочка обозначает комплексно-сопряженные величины. (На фиг. 5.10 показана такая запись фурье-преобразования изображения цифры 3.) Если теперь этот транспарант поместить на пути светового потока с распределением F — фурье-преобразованием функции f — и выполнить обратное преобразование, то получим

$$\begin{aligned} & \iint F (|G|^2 + |R|^2) e^{-2\pi j(\omega_x x + \omega_y y)} d\omega_x d\omega_y + \\ & + |R| \iint FG e^{-2\pi j[\omega_x (x+a) + \omega_y (y+b)]} d\omega_x d\omega_y + \\ & + |R| \iint FG^* e^{-2\pi j[\omega_x (x-a) + \omega_y (y-b)]} d\omega_x d\omega_y. \end{aligned}$$

Второй и третий члены в этом выражении представляют собой функции, смещенные в плоскости (x, y) относительно первого члена на $(-a, -b)$ и (a, b) соответственно. Следовательно, если эти функции равны нулю за пределами некоторой области, можно подобрать a и b так, чтобы все три члена в плоскости (x, y) не перекрывались. Второй член будет тогда давать смещенное изображение свертки f и g , умноженное на постоянный коэффициент $|R|$.

Интерференционные картины, с помощью которых можно воссоздать данное изображение, используются в голографии [70—73].

ЛИТЕРАТУРА

1. Slotnick D. L., Borck W. C., McReynolds R. C. The SOLOMON computer, *FJCC*, 97 (Dec. 1962).
2. Gregory J., McReynolds R., The SOLOMON computer, *IEEE Trans. Electronic Computers*, EC-12, 774 (Dec. 1963).
3. Senzig D. N., Smith R. V., Computer organization for array processing, *FJCC*, 117 (Nov. 1965).
4. Murtha J. C., Highly parallel information processing systems, *Advances Comput.*, 7, 1 (1966).
5. Lehman M., A survey of problems and preliminary results concerning parallel processing and parallel processors, *Proc. IEEE*, 54, 1889 (Dec. 1966); русский перевод: Леман М., Обзор проблем и предварительные результаты относительно параллельной обработки и параллельных процессорах, *ТНИЭР*, 54, № 12 (1966).
6. Fuller R. H., Associative parallel processing, *SJCC*, 471 (Apr. 1967).
7. Barnes G. H. et al., The ILLIAC IV computer, *IEEE Trans. Computers*, C-17, 746 (Aug. 1968).
8. Kuck D. J., ILLIAC IV software and application programming, *IEEE Trans. Computers*, C-17, 758 (Aug. 1968).
9. Unger S. H., A computer oriented toward spatial problems, *Proc. IRE*, 46, 1744 (Oct. 1958).
10. Unger S. H., Pattern detection and recognition, *Proc. IRE*, 47, 1737 (Oct. 1959).
11. Kamenskyy L. A., Pattern and character recognition systems — picture processing by nets of neuron-like elements, *WJCC*, 304 (March 1959).
12. Von Foerster H., Circuitry of clues to Platonic ideation, Aspects of the Theory of Artificial Intelligence: Proceedings, International Symposium on Biosimulation, 1st, Locarno, 1960 (C. A. Muses, ed.), Plenum Press, New York, 1962, p. 43.
13. Edelstein L. A., «Picture logic» for «Bacchus», a fourth-generation computer, *Comput. J.*, 6, 144 (July 1963).
14. McCormick B. H., The Illinois pattern recognition computer — ILLIAC III, *IEEE Trans. Electronic Computers*, EC-12, 791 (Dec. 1963).
15. Yau S. S., Yang C. C., Pattern recognition using an associative memory, *IEEE Trans. Electronic Computers*, EC-15, 944 (Dec. 1966).
16. Pease M. C., An adaptation of the Fast Fourier Transform for parallel processing, *J. ACM*, 15, 252 (Apr. 1968).
17. Narasimhan R., Labeling schemata and syntactic descriptions of pictures, *Information and Control*, 7, 151 (July 1964).
18. Narasimhan R., Syntax-directed interpretation of classes of pictures, *Comm. ACM*, 9, 166 (March 1966).
19. Pfaltz J. L., Snively J. W., Jr., Rosenfeld A., Local and global picture processing by computer, *PPR*, p. 353.

20. Moore G. A., Automatic scanning and computer processes for the quantitative analysis of micrographs and equivalent subjects, *PPR*, p. 275.
21. Hook H. O., Weinstein H., Image processing with optical panels, *Electronics*, 34, 35 (Dec. 21, 1962); русский перевод: Хук, Вайнштейн, Обработка изображений на оптических панелях, *Электроника*, 34, № 24 (1962).
22. Gray T. E., Considerations in optoelectronic logic and memory arrays, *OPI*, p. 216.
23. Hawkins J. K., Munsey C. J., An adaptive system with direct optical input, *Proc. IEEE*, 55, 1084 (June 1967); русский перевод: Хоккинс, Манси, Адаптивная система с прямым оптическим входом, *ТНИЭР*, 55, № 6 (1967).
24. Hawkins J. K., Munsey C. J., Image processing by electron-optical techniques, *J. Opt. Soc. Amer.*, 57, 914 (July 1967).
25. Hawkins J. K., Munsey C. J., Parallel logic with charge storage techniques, *IEEE Trans. Electronic Computers*, EC-16, 507 (Aug. 1967).
26. Hawkins J. K., Parallel electro-optical picture processing, *PPR*, p. 373.
27. Abraham J. M., Catchpole C. E., Goodrich G. W., Image processing with multiaperture image dissector, *J. Soc. Photo-Opt. Instr. Engrs.*, 6, 93 (Febr.— March 1968).
Описано устройство с многоэлементной (3×3 элемента) апертурой, сканирующей изображение. Локальные операции над изображением выполняются путем преобразований девяти электрических выходных сигналов элементов апертуры.
28. Russell J. K., A visual image processor, *IEEE Trans. Computers*, C-17, 635 (July 1968).
Описано устройство для запоминания изображений на двух электровакуумных трубках. Преобразования выполняются над парами соответствующих точек на этих трубках; результат хранится на третьей трубке.
29. Hoffman A. S., Electrolytic cell for use as a real-time spatial filter, *J. Opt. Soc. Amer.*, 56, 828 (June 1966).
30. Poppelbaum W. J., Adaptive on-line Fourier transform, *PPR*, p. 387.
31. Holmes W. S. et al., Optical-electronic spatial filtering for pattern recognition, *OEOIP*, p. 199.
32. Rau J. E., Real-time complex spatial modulation, *J. Opt. Soc. Amer.*, 57, 798 (June 1967).
33. Goodman J. W., Lawrence R. W., Digital image formation from electronically generated holograms, *Appl. Phys Lett.*, 11, 77 (Aug. 1, 1967).
34. Stock R. M., Deener J. J., A real-time input preprocessor for a pattern recognition computer, *Proc. IEEE Comput. Confer.*, Sept. 1967, p. 149.

35. McLachlan D., Jr., The role of optics in applying correlation functions to pattern recognition, *J. Opt. Soc. Amer.*, 52, 451 (Apr. 1962).
36. Hawkins J. K., Munsey C. J., A natural image computer, *OPI*, p. 233.
37. Hawkins J. K., Munsey C. J., A parallel computer organization and mechanizations, *IEEE Trans. Electronic Computers*, EC-12, 251 (June 1963).
38. Hawkins J. K., Munsey C. J., Automatic photo reading, *Photogrammetric Engrg.*, 29, 632 (July 1963).
39. Hawkins J. K., Munsey C. J., Eulogismographic nonlinear optical image processing for pattern recognition, *J. Opt. Soc. Amer.*, 54, 998 (Aug. 1964).
40. Hawkins J. K., Photographic techniques for extracting image shapes, *Phot. Sci. Engrg.*, 8, 329 (Nov.—Dec. 1964).
41. Jackson P. L., Correlation function spatial filtering with incoherent light, *Appl. Opt.*, 6, 1272 (July 1967).
42. De M., Lohmann A. W., Signal detection by correlation of Fresnel diffraction patterns, *Appl. Opt.*, 6, 2171 (Dec. 1967).
43. Vander Lugt A., A review of optical data-processing techniques, *Optica Acta*, 15, 1 (1968).
44. O'Neill E. L., Spatial filtering in optics, *IRE Trans. Information Theory*, IT-2, 56 (June 1956).
45. Aroyan G. F., The technique of spatial filtering, *Proc. IRE*, 47, 1561 (Sept. 1959).
46. Cutrona L. J., Leith E. N., Porcello L. J., Filtering operations using coherent optics, *Proc. Nat. Electron. Confer.*, Oct. 1959, p. 262.
47. Cutrona L. J. et al., Optical data processing and filtering systems, *IRE Trans. Information Theory*, IT-6, 386 (June 1960).
48. Marechal A., Optical filtering by double diffraction, *OPI*, p. 20.
49. Preston K., Jr., Use of the Fourier transformable properties of lenses for signal analysis, *OEOIP*, p. 59.
50. Lansraux G., Contributions of diffraction optics to optical information technology, *OEOIP*, p. 69.
51. Cutrona L. J., Recent developments in coherent optical technology, *OEOIP*, p. 83.
52. Leith E. N., Kozma A., Upatnieks J., Coherent optical systems for data processing, spatial filtering and wavefront reconstruction, *OEOIP*, p. 143.
53. Smith H. M., Photographic relief images, *J. Opt. Soc. Amer.*, 58, 533 (Apr. 1968).
54. Holladay T. M., Gallatin J. D., Phase control by polarization in coherent spatial filtering, *J. Opt. Soc. Amer.*, 56, 869 (July 1966).
55. Brown B. R., Lohmann A. W., Complex spatial filtering with binary masks, *Appl. Opt.*, 5, 967 (June 1966).
56. Lohmann A. W., Paris D. P., Computer generated spatial filters for coherent optical data processing, *Appl. Opt.*, 7, 51 (Apr. 1968).

57. Vander Lugt A., Signal detection by complex spatial filtering, *IEEE Trans. Information Theory*, IT-10, 139 (Apr. 1964).
58. Gabor D., Character recognition by holography, *Nature*, 208, 422 (Oct. 30, 1965); см. также Watrasiewicz B. M., *Nature*, 216, 302 (Oct. 21, 1967).
59. Kozma A., Kelly D. L., Spatial filtering for detection of signals submerged in noise, *Appl. Opt.*, 4, 387 (Apr. 1965).
60. Vander Lugt A., Rotz F. B., Klooster A., Jr., Character-reading by optical spatial filtering, *OEOIP*, p. 125.
61. Weaver C. S., Goodman J. W., A technique for optically convolving two functions, *Appl. Opt.*, 5, 1248 (July 1966).
62. Cathey W. T., Jr., Spatial phase modulation of wavefronts in spatial filtering and holography, *J. Opt. Soc. Amer.*, 56, 1167 (Sept. 1966).
63. Vander Lugt A., Practical considerations for the use of spatial carrier-frequency filters, *Appl. Opt.*, 5, 1760 (Nov. 1966).
64. Rau J. E., Detection of differences in real distributions, *J. Opt. Soc. Amer.*, 56, 1490 (Nov. 1966).
65. Vander Lugt A., The effects of small displacements of spatial filters, *Appl. Opt.*, 6, 1221 (July 1967).
66. Raso D. J., Simplified method to make hologram filters for target recognition, *J. Opt. Soc. Amer.*, 58, 432 (March 1968).
67. Lohmann A. W., Matched filtering with self-luminous objects, *Appl. Opt.*, 7, 561 (March 1968).
68. Binns R. A., Dickinson A., Watrasiewicz B. M., Methods of increasing discrimination in optical filtering, *Appl. Opt.*, 7, 1047 (June 1968).
69. Dickinson A., Watrasiewicz B. M., Optical filtering applied to postal code reading, *CPR*, p. 207.
70. Chambers R. P., Courtney-Pratt J. S., Bibliography on holograms, Part 1, *J. Soc. Motion Picture Television Engrs.*, 75, 373 (Apr. 1966); Part 2, 75, 759 (Aug. 1966).
71. Stroke G. W., An Introduction to Coherent Optics and Holography, Academic Press, New York, 1966; русский перевод: Строук Дж., Введение в когерентную оптику и голографию, изд-во «Мир», 1967.
72. DeVelis J. B., Reynolds G. O., Theory and Applications of Holography, Addison-Wesley, Reading, Massachusetts, 1967.
73. Goodman J. W., Introduction to Fourier Optics, McGraw-Hill, New York, 1968; русский перевод: Гудмен Дж., Введение в фурье-оптику, изд-во «Мир», 1970.
- 74*. Ленц В. Л., Матвеева А. Г., Цуккерман И. И., Пространственная фильтрация электронных изображений на мишени, *ЖТФ*, 36, № 8, 1514 (1966).

Применение пространственно-инвариантных операций над изображениями

6.1. Согласованная фильтрация

Часто необходимо определить степень «согласования» двух изображений друг с другом или выявить на одном изображении часть, согласующуюся с другим изображением. Такая необходимость возникает в основном при распознавании образов, в особенности в тех случаях, когда «изображения» в значительной мере стандартизованы (как, например, печатные знаки или изображения заведомо известных объектов), а также при навигации по картам (например, по звездным [1]), когда на наблюдаемом изображении необходимо найти область, соответствующую заданной на карте, и в автоматической стереограмметрии, где нужно идентифицировать изображения одного и того же участка на стереопаре с целью измерения параллакса этого участка (краткий исторический обзор работ в этой области дан в работе [2]).

Можно рассматривать также обратную задачу обнаружения различий между двумя изображениями, например изображениями одного и того же объекта в разные моменты времени. Если нужно обнаружить *существенные* изменения в изображении, а не изменения, вызванные шумом, разницей в экспозиции при регистрации, затенением и т. д., эта задача может оказаться не столь простой даже при точной регистрации изображений. Некоторые интересные методы обнаружения различий в двух изображениях приведены в работах [3, 4].

Существует много способов измерения отличия одной функции от другой. Например, можно просто вычислить их среднюю (абсолютную или квадратичную) разность, т. е. $\iint |f - g|$ или $\iint (f - g)^2$, деленную на площадь, по которой производится интегрирование. Наиболее часто используется метод, основанный на следующей теореме, известной как неравенство Шварца.

Теорема. Для любых интегрируемых, действительных неотрицательных функций f и g , не равных тождественно нулю,

$$\int fg \leq (\int f^2 \int g^2)^{1/2}$$

в любой области определения интегралов, причем это неравенство превращается в равенство тогда и только тогда, когда $g = cf$, где c — некоторая константа.

Доказательство. Это неравенство эквивалентно неравенству $(\int fg)^2 \leq \int f^2 \int g^2$. Рассмотрим полином второго порядка $P(z) = (\int f^2)z^2 + 2(\int fg)z + \int g^2$. Так как $F(z) = \int (fz + g)^2$, то $P(z) \geq 0$ при всех z и, следовательно, $P(z_0) = 0$ только при таком значении z_0 , при котором $fz_0 + g = 0$ во всей области определения f и g . Если такое значение z_0 существует, то $g = -z_0 f$, т. е. равно f , умноженной на константу, и при этом $(\int fg)^2 = z_0^2 \int f^2 = \int f^2 \int g^2$. В противном случае у $P(z)$ нет действительных корней и, следовательно, его дискриминант $(\int fg)^2 - \int f^2 \int g^2$ строго меньше нуля.

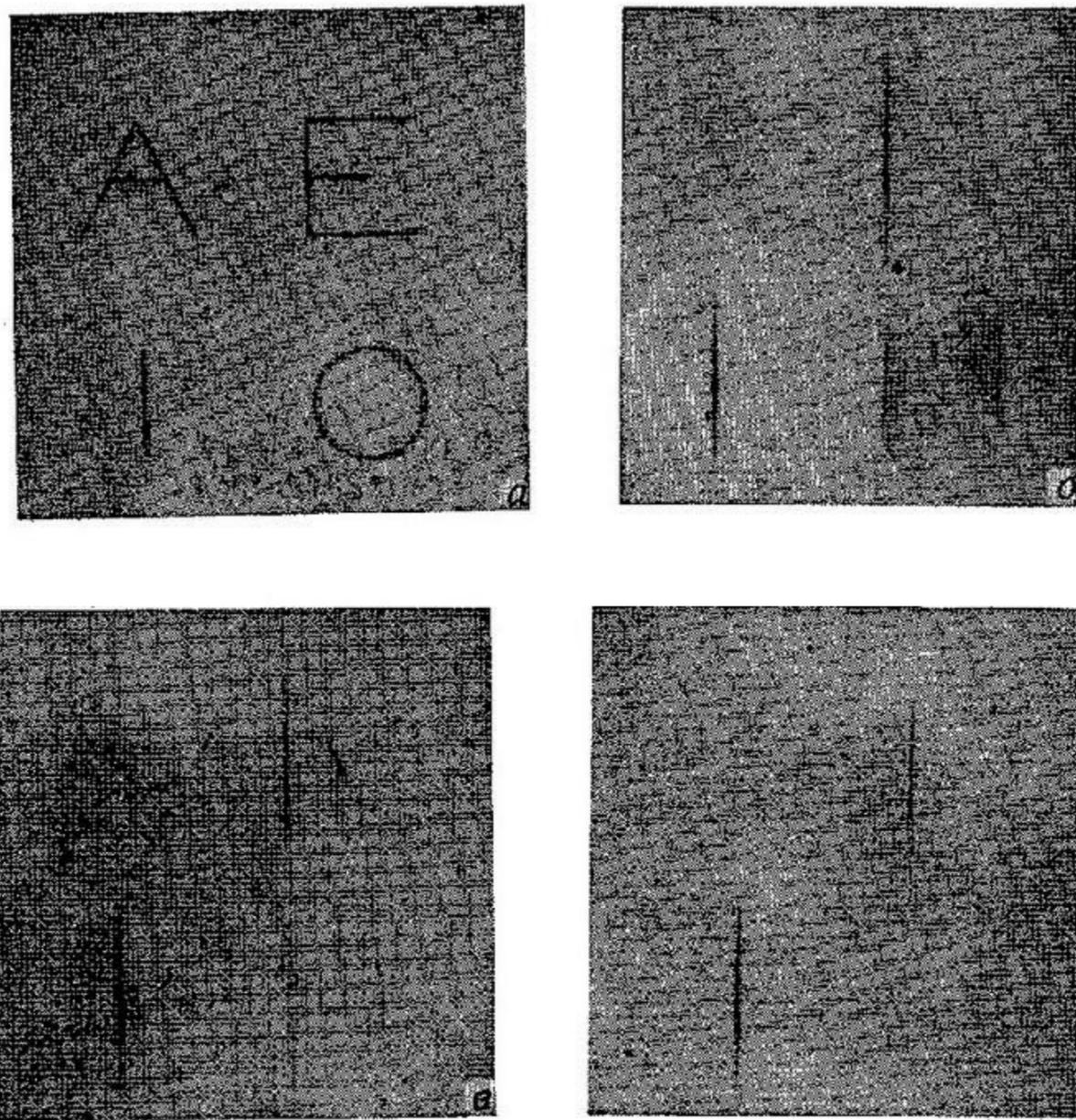
Следствие 1. $\int fg / (\int f^2 \int g^2)^{1/2} \leq 1$, причем равенство имеет место тогда и только тогда, когда $g = cf$.

Следствие 2.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v) g(u+x, v+y) dudv / \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2 \times \right. \\ \left. \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g^2 \right)^{1/2} \leq 1,$$

причем равенство имеет место тогда и только тогда, когда $g = cT_{x, y}(f)$. В словесной формулировке: нормализованная функция взаимной корреляции двух изображений f и g может принимать значение 1 только при таком смещении изображения g , при котором оно совпадает с f с точностью до мультипликативной константы.

Таким образом, вычисление взаимной корреляции f и g является простым методом определения их идентичности с точностью до сдвига и мультипликативной константы. Этот метод может оказаться очень полезным в



Ф и г. 6.1. Нормализованная взаимная корреляция.

a — двоичные дискретные изображения (255×255 элементов) букв А, Е, I, О; *б* — взаимная корреляция буквы I (заданной на растре 51×51 элемент) с изображением на фиг. 6.1, *a*; *в* — фиг. 6.1, *б* после соответствующей нормализации каждого квадранта; *г* — нелинейное масштабирование (по квадратичной шкале) изображения на фиг. 6.1, *в*.

таких задачах распознавания образов, как идентификация стилизованных незашумленных знаков стандартного размера и ориентации (но имеющих произвольное положение), когда одновременно испытывается только один знак (т. е. изображение состоит только из одного знака). В этом случае для идентификации достаточно определить

функцию взаимной корреляции данного знака с «эталонным» каждого из возможных знаков. На фиг. 6.1 показана в качестве примера функция взаимной корреляции «I» с «А», «Е», «I» и «О».

[Значение функции взаимной корреляции в максимуме часто ненамного больше значений в других точках. Но этот максимум можно подчеркнуть при помощи нелинейного масштабирования, как показано на фиг. 6.1, *г*. На практике для этой цели можно использовать нелинейность фотографического процесса.]

Если f и g являются дискретными изображениями, рассмотренные выше интегралы можно заменить суммами. Действительно, можно сформулировать следующий дискретный аналог неравенства Шварца: для любых двух множеств действительных чисел x_1, \dots, x_m и y_1, \dots, y_m , ни одно из которых не равно нулю,

$$\sum_{k=1}^m x_k y_k \leq \left(\sum_{k=1}^m x_k^2 \sum_{k=1}^m y_k^2 \right)^{1/2},$$

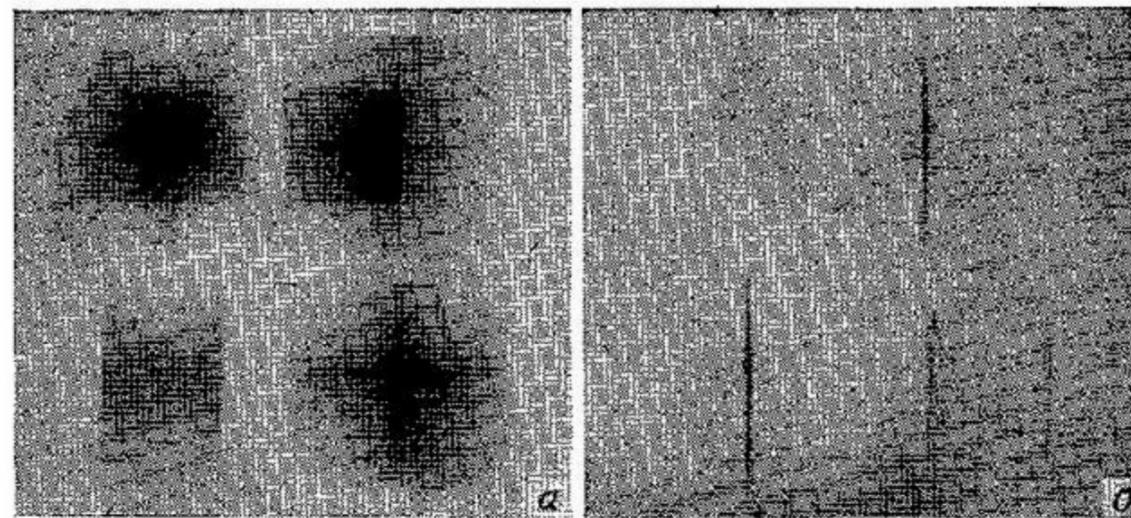
причем равенство имеет место тогда и только тогда, когда $y_k = c x_k$ при каждом k , где c — некоторая константа. Чтобы применить это соотношение к дискретному изображению из $n \times n$ элементов, нужно просто положить $m = n^2$ и считать, что x_k и y_k — значения яркости элементов изображения.

Упражнение 1. Вычислить функцию автокорреляции следующих двухградационных изображений и их попарные функции взаимной корреляции:

000	000	010	010	101	111
010	111	010	111	010	111
000	000	010	010	101	111

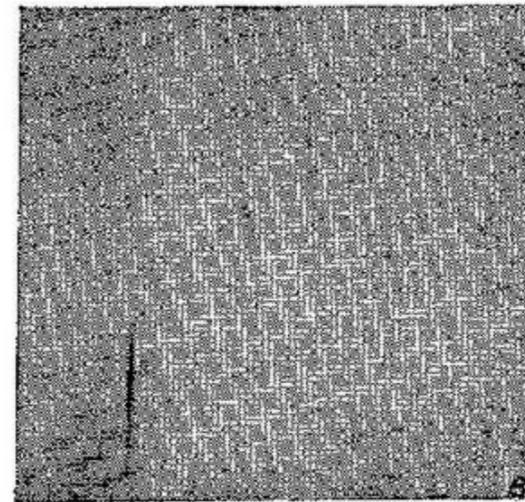
Упражнение 2. Пусть (a_{ij}) — «случайное» двухградационное дискретное изображение с вероятностью единиц p , состоящее из $m \times m$ независимых элементов.

а) Чему равно среднее число единиц в (a_{ij}) ? Иными словами, сколько единиц будет приходиться в среднем



Ф и г. 6.2. Согласованная фильтрация путем деления коррелограмм.

a — взаимная корреляция квадратной площадки размером 51×51 элемент с изображением на фиг. 6.1, *a*; *б* — результат поэлементного деления фиг. 6.1, *б* на квадратный корень значений изображения на фиг. 6.2, *a*; *в* — результат нелинейного масштабирования (по квадратичной шкале) изображения на фиг. 6.2, *б*.



на одно изображение в множестве случайных изображений с вероятностью единиц p ?

б) Пусть (b_{ij}) — другое двухградационное изображение с заданной вероятностью единиц q . Чему равно среднее число элементов изображений (a_{ij}) и (b_{ij}) , одновременно равных 1?

в) Применить результаты (а) и (б) для вычисления математического ожидания величины

$$\sum \sum a_{ij} b_{ij} / (\sum \sum a_{ij}^2 \sum \sum b_{ij}^2)^{1/2}.$$

г) Пусть (c_{ij}) — двухградационное изображение, определенное так же, как и (b_{ij}) , но с $n > m$. Чему равно математическое ожидание функции взаимной корреляции (a_{ij}) и (c_{ij}) при смещениях, не выводящих (a_{ij}) за пределы (c_{ij}) ? Чему равно математическое ожидание числа

смещений, для которых значения функции взаимной корреляции больше или равны k ($0 < k \leq 1$)?

Более общая проблема состоит в локализации положения объекта f на изображении g . При этом имеется в виду, что $g = cT_{xy}(f)$, когда u и v лежат в некоторой области S (когда u и v находятся вне этой области, $f = 0$). Если рассматривать S как область интегрирования, то в числителе левой части неравенства следствия 2 по-прежнему будет $f \otimes g$, $\iint f^2$ в знаменателе также не изменится, но $\iint g^2$ изменится и превратится в интеграл $\iint g^2(u+x, v+y) dudv$ по (u, v) , входящий в область S . Однако заметим, что этот интеграл может рассматриваться как взаимная корреляция характеристической функции S (т. е. функции, равной 1 в S и 0 вне S) с g^2 . Таким образом, объект f в g можно локализовать, вычислив отношение двух функций взаимной корреляции. Фиг. 6.2 иллюстрирует это для буквы l .

Упражнение 3. Выполнить эти вычисления, положив f равным каждому из изображений, приведенных в упражнении 1 (при этом S будет квадратом 3×3 элемента), и приняв, что g состоит из всех этих изображений, уложенных на растре из 15×15 элементов. Что произойдет, если использовать различные S , например прямоугольник шириной 3 и высотой в 1 элемент? (Проверить это, в частности, для первых трех изображений.)

Обычно желательно, чтобы согласование между f и g было четко выраженным, т. е. чтобы нормализованная функция взаимной корреляции представляла собой узкий импульс. Это требование выполняется в основном для штриховых рисунков, так как в этом случае даже небольшие смещения приводят к резкому уменьшению функции взаимной корреляции (если f построено из «широких» линий, данное требование не выполняется [5]). Наличие острого пика в коррелограмме можно обнаружить без применения поэлементного сканирования, наблюдая ее с помощью нелинейного фоточувствительного прибора, дающего более сильный выходной сигнал при неравномерном распределении по полю падающего на него света [6].

Рассмотрим теперь еще более общую постановку задачи. Предположим, что в изображении g содержится изо-

бражение $f + h$, т. е. изображение f , искаженное аддитивным шумом h . Необходимо найти положение f в g , измерив взаимную корреляцию между g и некоторым эталонным изображением f_1 . Если изображение f и шум h независимы, в частности если $\iint f_1 h = \iint f_1 \iint h$ при любом взаимном расположении f_1 и h , то, как легко показать, наилучшим эталоном f_1 будет само изображение f . Действительно, в силу независимости h и f_1 интеграл $f_1 \otimes h = \iint f_1 \iint h$ является константой для всех f_1 с одинаковой «полной прозрачностью» $\iint f_1$. Следовательно, свертка $f_1 \otimes (f + h) = f_1 \otimes f + f_1 \otimes h$ максимальна, если максимальна величина $f_1 \otimes f$, т. е. если $f_1 = f$ с точностью до сдвига и мультипликативной константы. Таким образом, наилучшим эталоном для обнаружения f в шуме (при данной «полной прозрачности») является само изображение f , т. е. наилучший фильтр — согласованный.

Разберем еще один подход к проблеме согласованной фильтрации и рассмотрим случай дискретных изображений, наблюдаемых на фоне аддитивного нормального шума. В этом случае вероятность того, что некоторое изображение (c_{ij}) получено в результате искажения шумом изображения (a_{ij}) , пропорциональна

$$\prod_{i,j=1}^n e^{-k(c_{ij} - a_{ij})^2}.$$

Вероятность того, что (c_{ij}) получено из другого незашумленного изображения (b_{ij}) , можно подсчитать по той же формуле, заменив (a_{ij}) на (b_{ij}) . Чтобы узнать, какое из исходных изображений, (a_{ij}) или (b_{ij}) , более вероятно, нужно выяснить, какое из этих произведений больше. Для этого достаточно определить знак суммы

$$\sum_{i,j=1}^n [(c_{ij} - b_{ij})^2 - (c_{ij} - a_{ij})^2],$$

которая получится, если взять логарифм отношения этих произведений и отбросить несущественные константы. Но эта сумма равна

$$\sum_{i,j=1}^n [2c_{ij}(a_{ij} - b_{ij}) + b_{ij}^2 - a_{ij}^2].$$

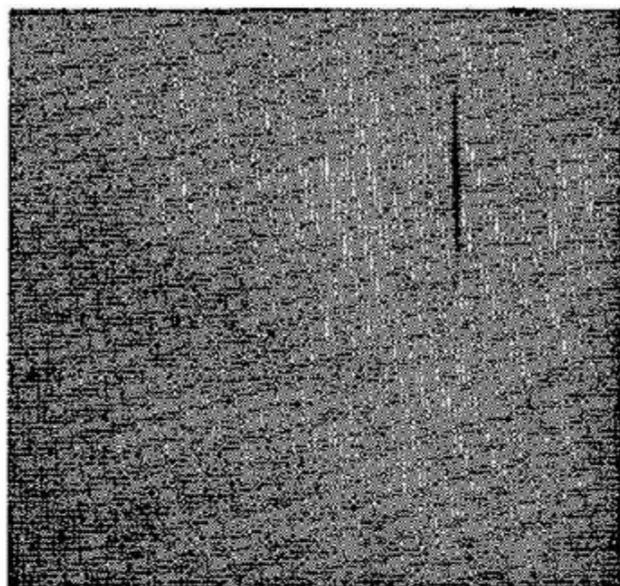
Следовательно, если $\sum a_{ij}^2 = \sum b_{ij}^2$, что аналогично условию одинаковой полной прозрачности¹, то достаточно сравнить между собой $\sum c_{ij}a_{ij}$ и $\sum c_{ij}b_{ij}$. Таким образом, в качестве фильтров следует использовать сами изображения (a_{ij}) и (b_{ij}) ([7—10]).

Во многих задачах распознавания образов можно считать, что анализируемые изображения являются двухградационными, т. е. содержат только 1 и 0 (черные и белые, прозрачные или непрозрачные участки). При решении таких задач нет необходимости использовать нормализованные функции взаимной корреляции из следствия 2, для вычисления которых требуется производить деление коррелограмм. Вместо этого можно измерить взаимную корреляцию между изображением и эталоном, а также между их негативами и затем сложить результаты. Это значительно проще в реализации (например, оптическими средствами), чем деление. Действительно, предположим, что площадь эталона f равна A , а его полная прозрачность (т. е. $\iint f$, равный для двухградационного изображения площади, на которой $f = 1$) равна T , $T \leq A$. Коррелограмма $f \otimes g$ изображения f и двухградационного изображения g будет иметь значение T только при таких смещениях g относительно f , при которых $g = 1$ везде, где $f = 1$. Однако при этом g может быть равно 1 также в точках, где $f = 0$. С другой стороны, предположим, что измеряется взаимная корреляция между негативами f' и g' , полученными соответственно из f и g заменой единиц нулями и наоборот. Полная прозрачность негатива f' эталона равна $A - T$, и $f' \otimes g'$ принимает это значение только при таких смещениях g' , при которых $g' = 1$ везде, где $f' = 1$, или, что то же самое, при которых $g = 0$ везде, где $f = 0$. Следовательно, смещение, при котором $f \otimes g + f' \otimes g' = T + (A - T) = A$, соответствует точному согласованию f и g .

Можно поступить и иначе — использовать в качестве второй коррелограммы свертку $f' \otimes g$. Она равна нулю

¹ Ранее автор назвал полной прозрачностью величину $\iint f$, дискретным аналогом которой является $\sum a_{ij}$, а не $\sum a_{ij}^2$. Только для двухградационных изображений, когда $a_{ij} = \{0, 1\}$, $\sum a_{ij} = \sum a_{ij}^2$. — Прим. ред.

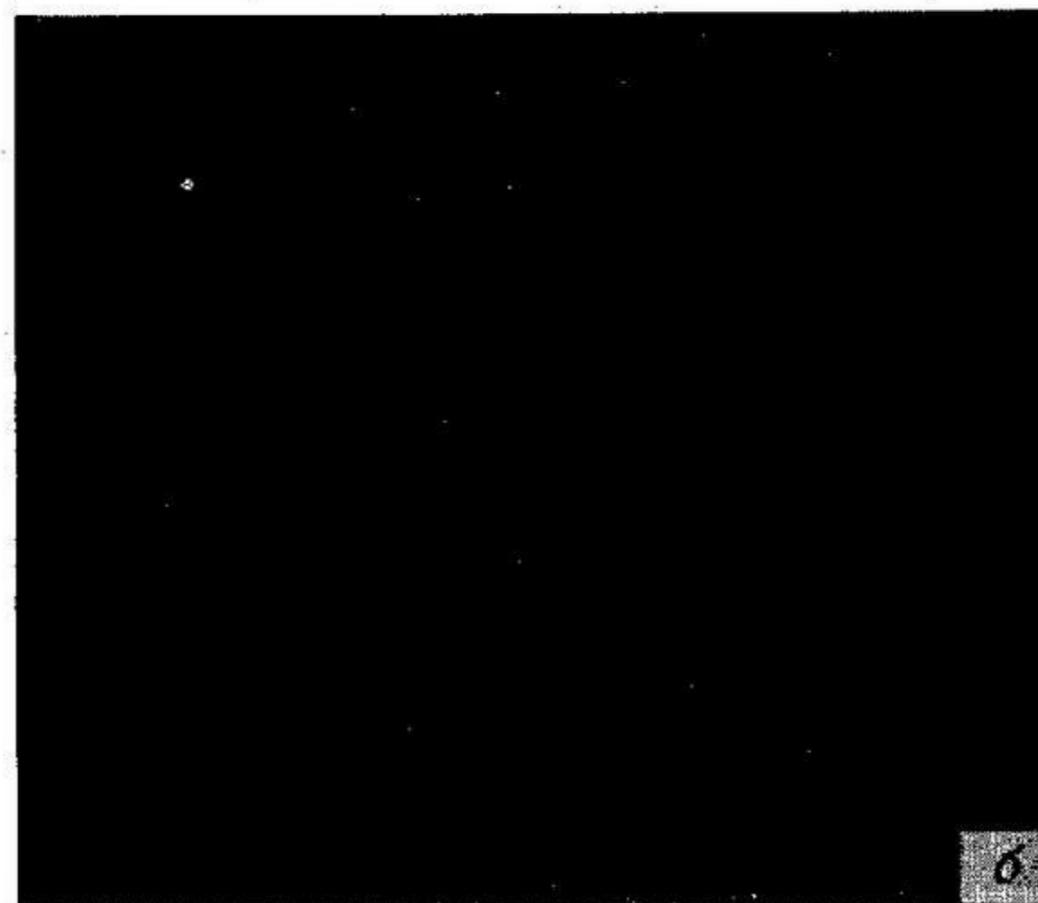
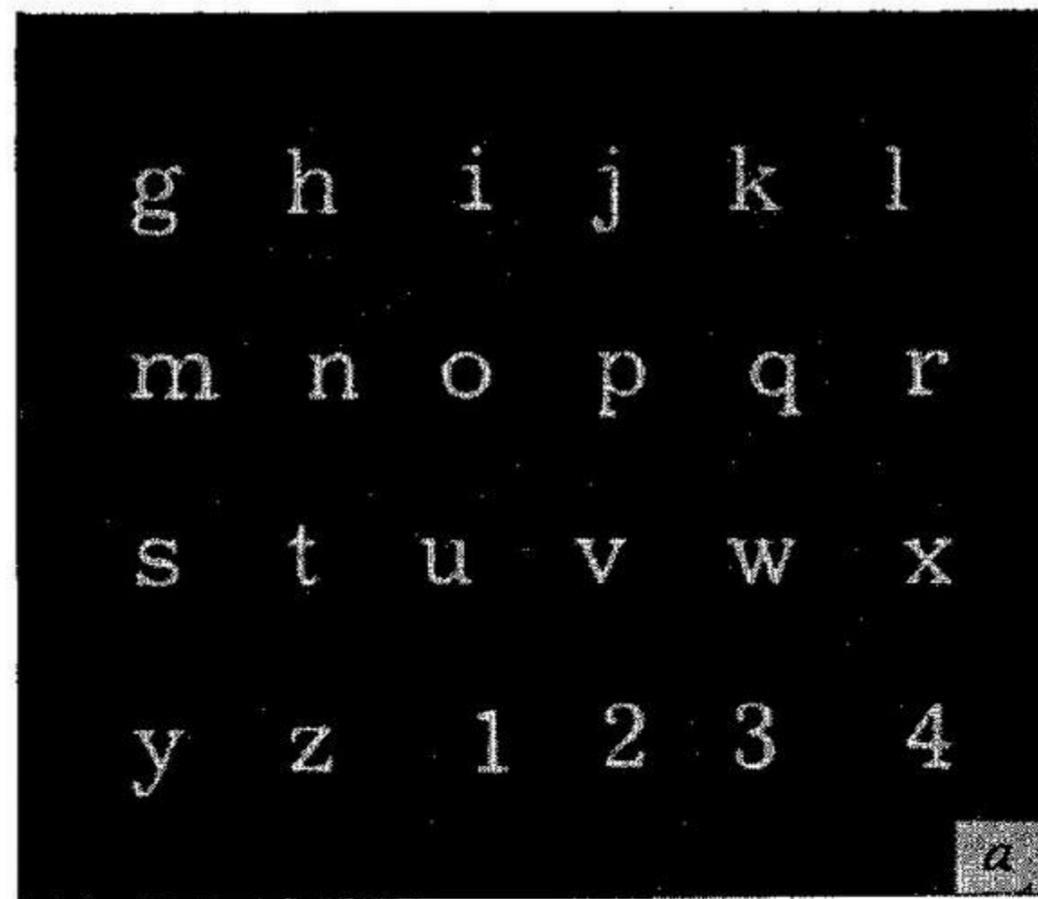
только при таких смещениях g , при которых $g = 0$ везде, где $f' = 1$, т. е. $f = 0$. Следовательно, смещение, при котором $f \otimes g - f' \otimes g = T$, соответствует точному согласованию g и f . (Это вычитание можно осуществить оптическими средствами. Кроме того, используя методы, описанные в разд. 5.3, можно, если это необходимо, совместить на одном носителе и позитив, и негатив эталона.)



Ф и г. 6.3. Согласованная фильтрация двухградационных изображений путем вычитания коррелограмм: поэлементная разность изображений на фиг. 6.1, б и 6.2., а; отрицательные значения заменены нулями.

Эти операции можно еще больше упростить, заметив, что $f \otimes g - f' \otimes g = 2f \otimes g - (f \otimes g + f' \otimes g) = 2f \otimes g - (f + f') \otimes g$, где второй член в правой части является взаимной корреляцией полностью прозрачного эталона с изображением g [11]. Описанный метод для изображения буквы l проиллюстрирован на фиг. 6.3.

Еще один метод измерения взаимной корреляции между двухградационным дискретным изображением и эталоном состоит в поэлементном умножении изображения самого на себя и на свой «негатив» со сдвигом. В самом деле, пусть эталон принимает значения 1 в элементах $(x_{i_1j_1}, y_{i_1j_1}), \dots, (x_{i_hj_h}, y_{i_hj_h})$ и 0 в элементах $(u_{i_1j_1}, v_{i_1j_1}), \dots, (u_{i_k}, v_{j_k})$. Тогда g согласуется с эталоном при



Ф и г. 6.4. Оптическая согласованная фильтрация в частотной области.

а — исходная матрица; б — обратное преобразование Фурье от произведения преобразований Фурье исходной матрицы и буквы g . Снимок из статьи: A. V a n d e r L u g t, Signal detection by complex spatial filtering, *IEEE Trans. Information Theory*, IT-10, 145 (April 1964).

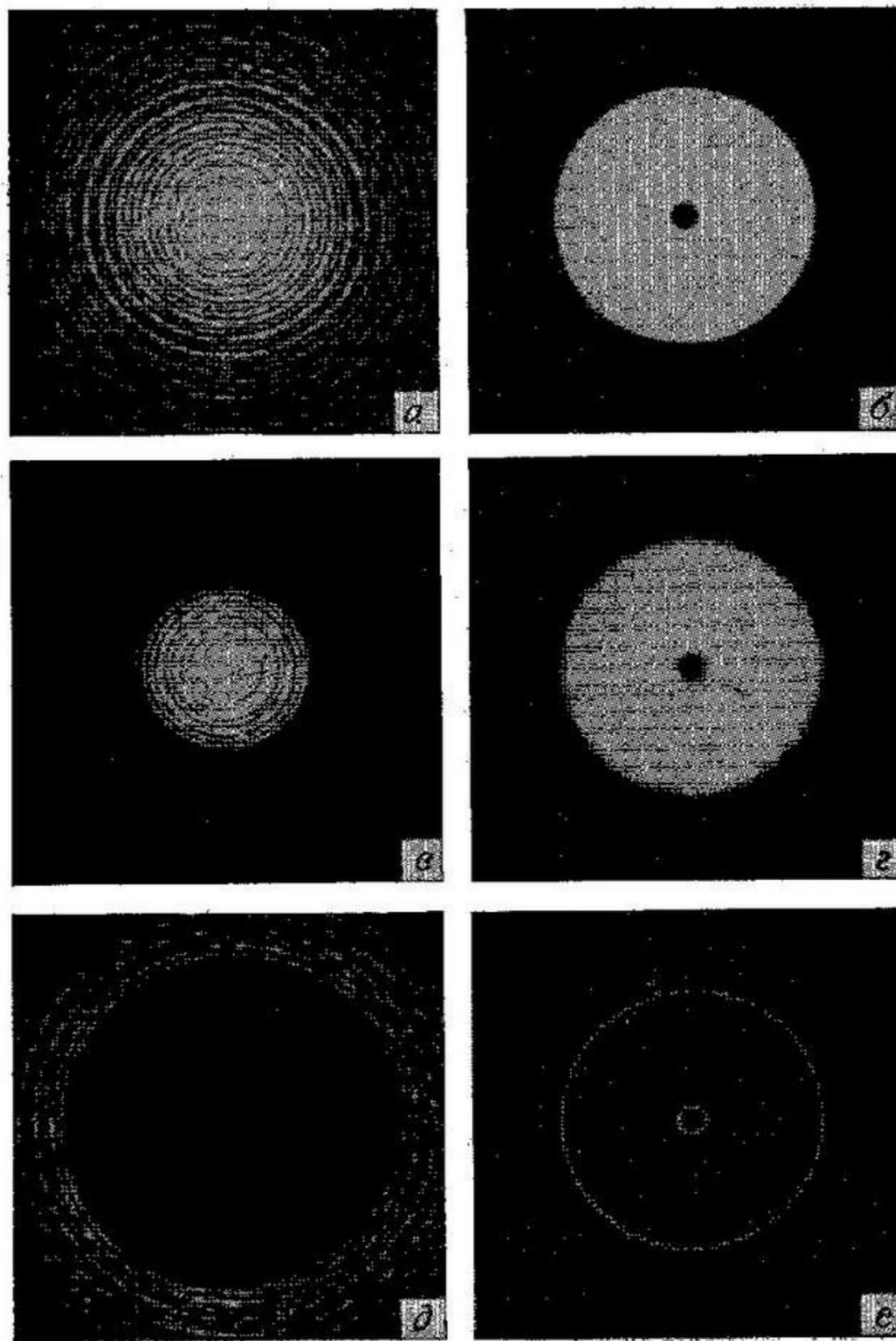
смещении (x, y) только при условии, что

$$\prod_{\alpha=1}^h g(x - x_{i\alpha}, y - y_{i\alpha}) \times \prod_{\beta=1}^k [1 - g(x - u_{i\beta}, y - v_{i\beta})] = 1.$$

6.2. Пространственная фильтрация и коррекция изображений

Как было указано в гл. 4 и 5, одним из способов вычисления свертки эталона и изображения является перемножение их пространственных спектров и вычисление обратного преобразования Фурье от полученного произведения. В этом случае согласованная фильтрация осуществляется в «частотной области», а не в «пространственной», так как операции умножения и интегрирования выполняются не над исходными изображениями, а над их преобразованиями Фурье (литература по согласованной фильтрации в частотной области приведена в разд. 5.4; см. также фиг. 6.4).

Одно из главных преимуществ работы в частотной области состоит в том, что, используя весьма простые фильтры, можно осуществлять самые разнообразные операции над изображениями. Например, если поместить в плоскость G (фиг. 5.9) непрозрачный диск (в оптической терминологии — маску), центр которого лежит на оптической оси системы, то будут подавлены низкие пространственные частоты. Это приводит к подавлению на восстановленном изображении областей с плавными изменениями яркости при сохранении резких переходов и мелких деталей. Наоборот, если в центре плоскости G поместить в качестве маски пластину с круглым отверстием, то сохранятся только низкие пространственные частоты изображения. В результате восстановленное изображение окажется размытым, расфокусированным. Действие этих фильтров показано на фиг. 6.5. Маска с длинной щелью, проходящей через центр, вызовет сглаживание на изображении резких контуров, направленных параллельно щели. (Литература по этим методам будет дана в разд. 6.4 и 6.5. О преимуществах выполнения пространственной фильтрации в частотной области изображения после логарифмического преобразования можно ознакомиться в работе [12].) Подобные операции можно, конечно, выполнить



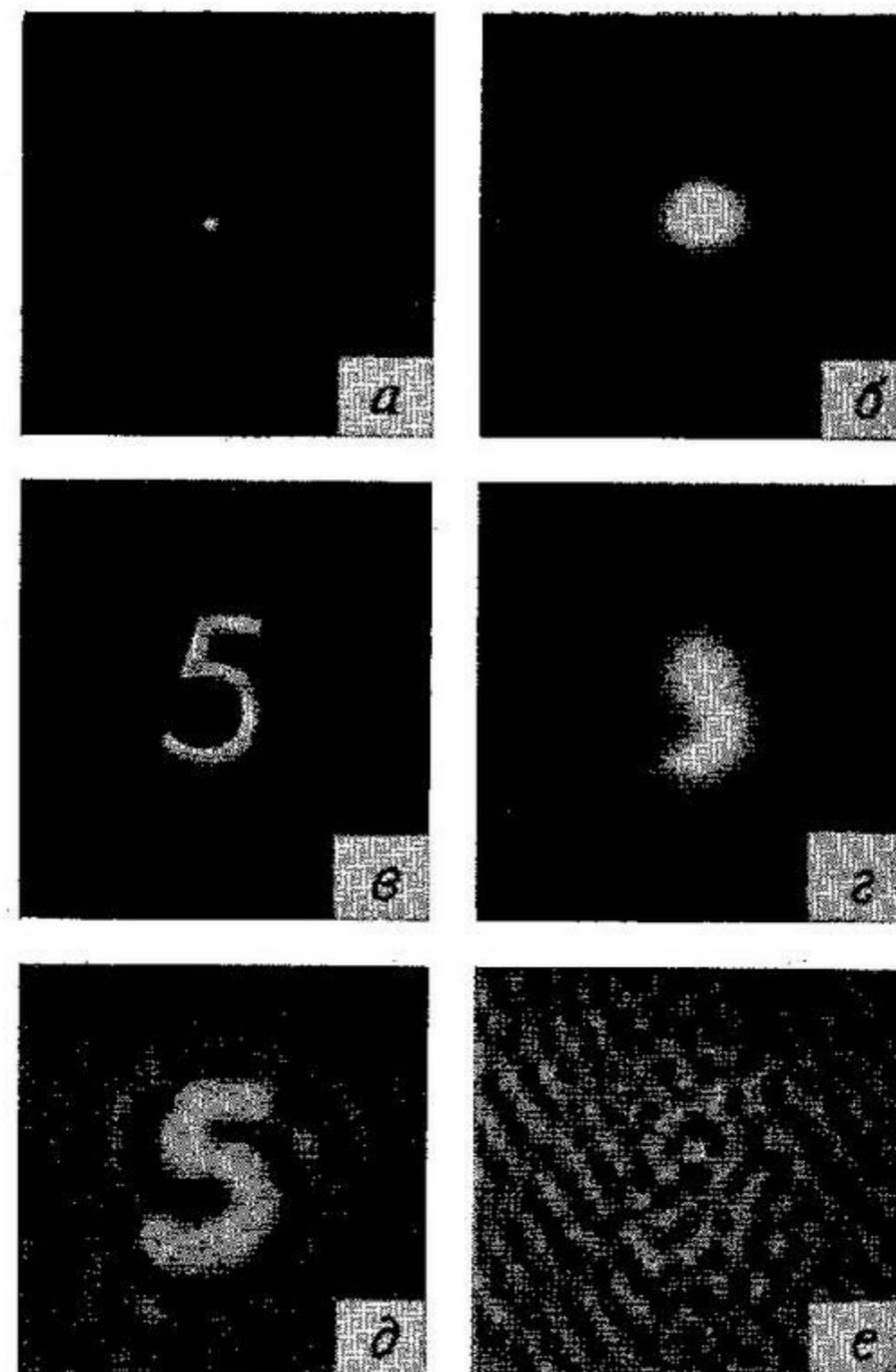
Ф и г. 6.5. Результаты оптической пространственной фильтрации изображения круглого отверстия диаметром 3 мм круглым непрозрачным экраном диаметром 0,3 мм в центре.

а — энергетический спектр, зарегистрированный в фурье-плоскости; б — обратное преобразование Фурье исходного изображения (без фильтрации); в — энергетический спектр на выходе фильтра пространственных частот; г — результат низкочастотной фильтрации; д — энергетический спектр на выходе фильтра верхних пространственных частот; е — результат высокочастотной фильтрации.

путем свертки исходного изображения с соответствующими аперттурами (а именно аперттурами, представляющими собой обратное преобразование Фурье соответствующей маски). Обратное преобразование Фурье маски (или в общем случае фильтра), являющейся действительной и симметричной относительно начала координат функцией (как, например, в случае диска, круглого отверстия, симметричной пары секторов), является также действительной функцией. Поэтому свертка (которая ввиду симметрии преобразования эквивалентна взаимной корреляции) изображения с этой аперттурой может быть выполнена оптическими средствами с помощью методов, описанных в разд. 5.3.

Важное применение пространственная фильтрация находит в задачах коррекции изображений. Предположим, что изображение формируется некоторой системой воспроизведения или некоторой системой передачи, которые его искажают. Если искажения являются обратимыми в математическом смысле, то в принципе можно компенсировать их и тем самым восстановить исходное неискаженное изображение. Предположим, в частности, что искажения можно описать с помощью пространственно-инвариантных линейных операций, скажем свертки исходного изображения f с некоторой аперттурой g , так что искаженное изображение есть $f * g$. Преобразование Фурье от $f * g$ равно произведению FG фурье-преобразований от f и g соответственно. Поделив его на G и взяв обратное преобразование Фурье, можно восстановить исходное изображение f .

По крайней мере приблизительно G можно определить как преобразование Фурье от изображения, полученного в результате искажения изображения точечного объекта. На практике искажения редко являются следствием только пространственных преобразований. Зачастую они вызваны также действием помех, а процедуры коррекции обычно очень чувствительны к помехам. Тем не менее, используя обратный фильтр (т. е. деление на G) на тех пространственных частотах, где отношение сигнал/шум велико, можно получить полезные результаты. Пример изображения, восстановленного таким способом, показан на фиг. 6.6.



Фиг. 6.6. Коррекция изображения.

a — контрастная точка на равномерном фоне (без искажений); b — то же изображение после расфокусировки; v — неискаженное изображение цифры 5; z — то же после расфокусировки; d — результат коррекции (деление выполнялось только для пространственных частот ниже 2 период/мм); e — результат деления спектральных составляющих на пространственных частотах до 3 период/мм.

Снимки из статьи: В. L. Mc G l a m e r y, Restoration of turbulence-degraded images, *J. Opt. Soc. Amer.*, 57, 295 (March 1967).

Введение в методы коррекции изображений дано в работах [13—21].

Упражнение. При каких условиях линейный дискретный локальный оператор, определяемый как произведение матриц (разд. 4.2, упражнение 3), является обратимым?

6.3. Измерение качества изображения

При любых преобразованиях изображений, например при фотокопировании, преобразованиях в оптических системах, сканировании и т. д., возникает вопрос, в какой мере в результирующем изображении сохраняется информация, содержащаяся в исходном изображении, или каково «качество» полученного изображения как копии оригинала.

Предположим, что результирующее изображение размыто по сравнению с оригиналом, т. е. является его нечеткой копией. Степень этой размытости полностью характеризуется импульсной характеристикой или передаточной функцией преобразователя, введенными в разд. 4.4. Однако в большинстве случаев желательно иметь меру размытости, выражающуюся одним числом, а не функцией. Обычно пользуются для этой цели двумя мерами — разрешающей способностью и резкостью. Первая из этих мер определяет различимость близко расположенных объектов, вторая — четкость границ. (Эти величины, вообще говоря, не всегда взаимосвязаны. Действительно, пусть φ_1 и φ_2 — операторы с «шляпообразной» и «конусообразной» функциями рассеивания соответственно. Очевидно, в изображении, подвергнутом действию оператора φ_1 , труднее разрешить близкие объекты, но контуры в нем будут резче, чем в изображении, полученном после оператора φ_2 .)

Меру разрешающей способности можно ввести, если задаться тест-объектами, которые должны обнаруживаться как отдельные, а также относительным контрастом объектов и расстоянием между ними, при которых они должны разрешаться. Объектами, которые используются для этой цели, чаще всего являются прямоугольная мира

(набор близко расположенных параллельных полосок) и кольца Ландольта (кольцо с узкой щелью).

Резкость обычно измеряется среднеквадратичной крутизной перепада (т. е. скоростью изменения яркости), деленной на общую величину этого перепада. Эту меру можно выразить в виде

$$\int_a^b (df/dx)^2 / \|f(b) - f(a)\|,$$

где a и b — точки, расположенные на противоположных краях перепада. Эксперименты показали, что эта мера хорошо коррелирует с субъективным ощущением резкости контуров (обзор методов изменения разрешающей способности и резкости приведен в [22—24]).

Упражнение 1. Пусть изображение f подвергается расфокусировке путем усреднения яркости его элементов по окрестности, ограниченной кругом радиусом r . Если в f содержится четкий контур, т. е. $f = 1$ при $x \leq 0$ и $f = 0$ при $x > 0$, то какова резкость этого контура в расфокусированном изображении? Каковы максимальное и минимальное значения яркости в расфокусированном изображении в области $0 \leq x \leq 3$, если в исходном изображении имелись две параллельные полоски с координатами $0 \leq x \leq 1$ и $2 \leq x \leq 3$ и с яркостью $f = 1$?

Преобразования изображений сопровождаются также действием различных шумов. К сожалению, в настоящее время статистические модели двумерного шума исследованы мало. Некоторые виды этого шума можно считать, по крайней мере приблизительно, аддитивными, в то время как шум зернистости, рассматриваемый ниже, не является аддитивным. (О случайных многомерных процессах см. Kendall M. G., Moran P. A. P., Geometrical Probability, Hafner, New York, 1963.)

Важным и интенсивно изучавшимся видом шума является фотографическая гранулярность. Обычную регистрирующую фотографическую среду можно рассматривать как взвесь случайно рассеянных фоточувствительных «зерен». Степень почернения некоторого участка фотографии определяется числом проявленных зерен на единицу

площади в пределах этого участка. Так как наблюдатель измеряет интегральное почернение $\iint f$ по некоторой области (в пределах некоторой апертуры конечных размеров), фотография t кажется ему полутоновой, а не двухградационной. Однако вследствие наличия зерен степень почернения фотографии любого объекта с постоянной яркостью подвержена флуктуациям. Мерой величины и пространственной протяженности этих флуктуаций является гранулярность. Ниже приведены некоторые конкретные способы измерения гранулярности [25—31].

Среднеквадратичная гранулярность, или гранулярность по Селвину. Простейшей мерой гранулярности является стандартное отклонение величины почернения

$$\sigma = [\iint (\bar{f} - f)^2]^{1/2},$$

где $\bar{f} = \iint f$ — ее среднее значение. Эта величина, конечно, зависит от размеров апертуры A , по которой производится измерение. Можно показать, что σ_f пропорционально $1/\sqrt{A}$. Следовательно, произведение $\sigma \sqrt{A}$ приблизительно постоянно и называется гранулярностью по Селвину.

Сизигетическая разностная плотность. Другой известной мерой гранулярности является сизигетическая плотность $\Delta f/\Delta x$, где Δx — расстояние на фотографии, соответствующее расстоянию между двумя соседними колбочками в фовеа глаза¹.

Автокорреляция. Так как гранулярность зависит от размеров зерен, ее можно охарактеризовать с помощью автокорреляционной функции f . Согласно неравенству Шварца (разд. 6.1), она достигает своего максимального значения $\iint f^2$ при нулевом смещении. Если бы зернистости не было, автокорреляционная функция равномерного поля была бы равна этому значению при любых смещениях. Быстрота, с которой ее значения уменьшаются при небольших смещениях, характеризует размеры зерен.

¹ О сизигетической (в русской литературе—сопряженной) плотности см. в работе [31].—Прим. ред.

Энергетический спектр. Еще одним методом измерения гранулярности является измерение пространственного спектра равномерного поля. Этот метод близко связан с методом измерения автокорреляционной функции, поскольку энергетический спектр является ее преобразованием Фурье (см. разд. 4.3).

Упражнение 2. Для двухградационного дискретного изображения, определенного в упражнении 2 разд. 6.1, найти: а) среднеквадратичную гранулярность как функцию p и числа элементов в апертуре (считая, что апертура полностью укладывается в изображение); б) гранулярность по Селвину; в) величину автокорреляционной функции для нулевого смещения (неполным перекрытием изображений при их сдвиге пренебречь).

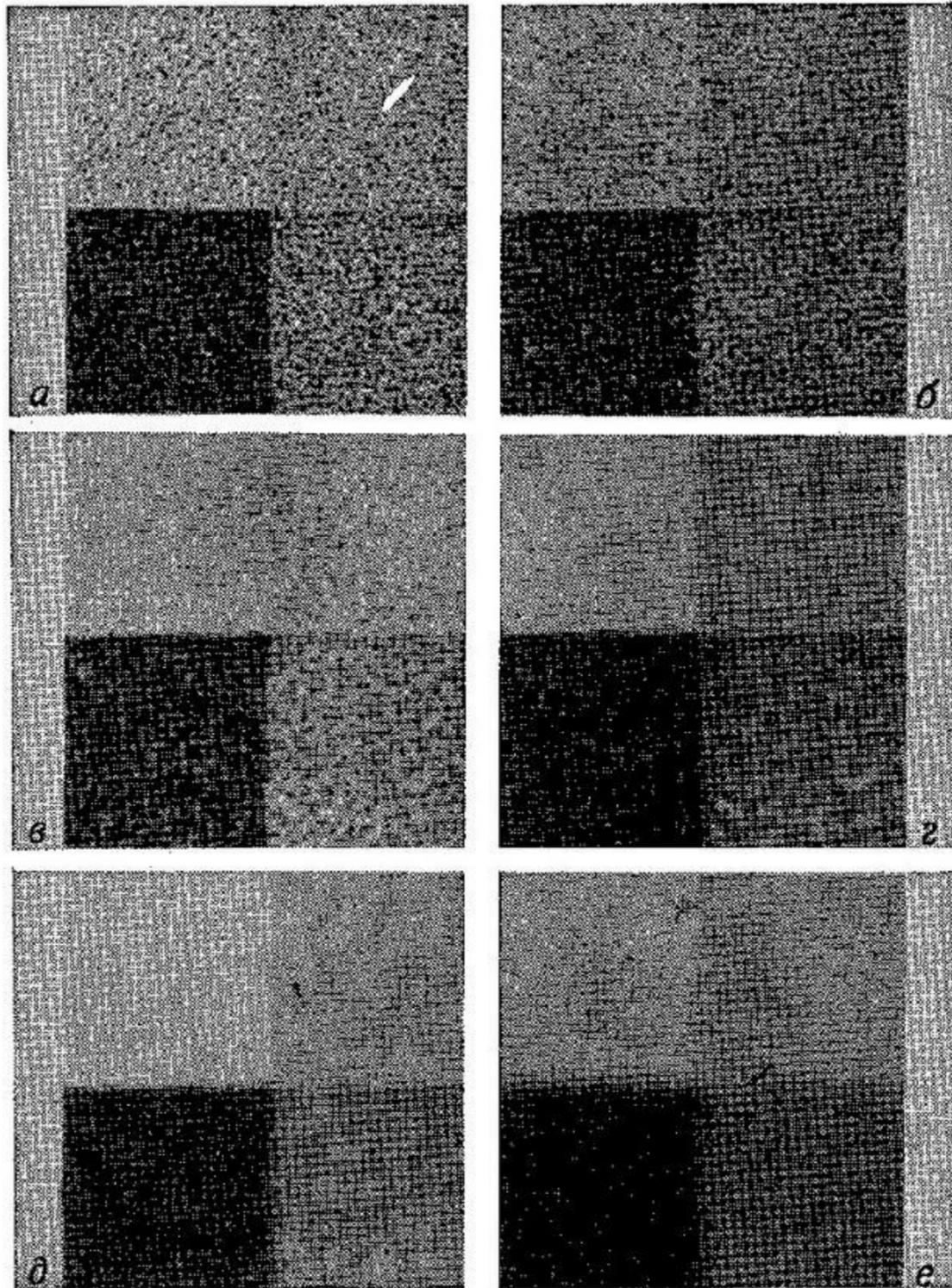
Предложены различные единые меры качества изображений, учитывающие и потерю четкости, и зашумленность. В частности, предлагалось использовать в качестве такой меры информационную емкость ([32], а также [1], разд. 2.1). Однако следует указать, что расфокусированное изображение, которое можно скорректировать, содержит такое же количество информации об исходном объекте, что и нерасфокусированное, но в общепринятом смысле качество этих изображений не одинаково.

6.4. Улучшение качества изображений. Устранение помех

В этом разделе рассматриваются пространственно-инвариантные операции над изображениями (не обязательно линейные), которые можно использовать для «сглаживания» изображений с целью подавления наложенного на них шума.

Усреднение по окрестности элементов

Простейшим способом сглаживания изображений является замена значений каждого элемента средним значением, найденным по окрестности элемента. Если для усреднения каждый раз используется окрестность одной



Ф и г. 6.7. Сглаживание путем усреднения по окрестности.

В исходном двухградационном изображении (а) вероятности черных элементов в четырех квадрантах равны соответственно 0,1; 0,2; 0,4 и 0,8. Каждое из сглаженных изображений (б — в) получено путем усреднения исходного изображения по квадратной окрестности, ориентированной под углом 45° и имеющей радиус соответственно 1, 2, 4, 8, 16 элементов.

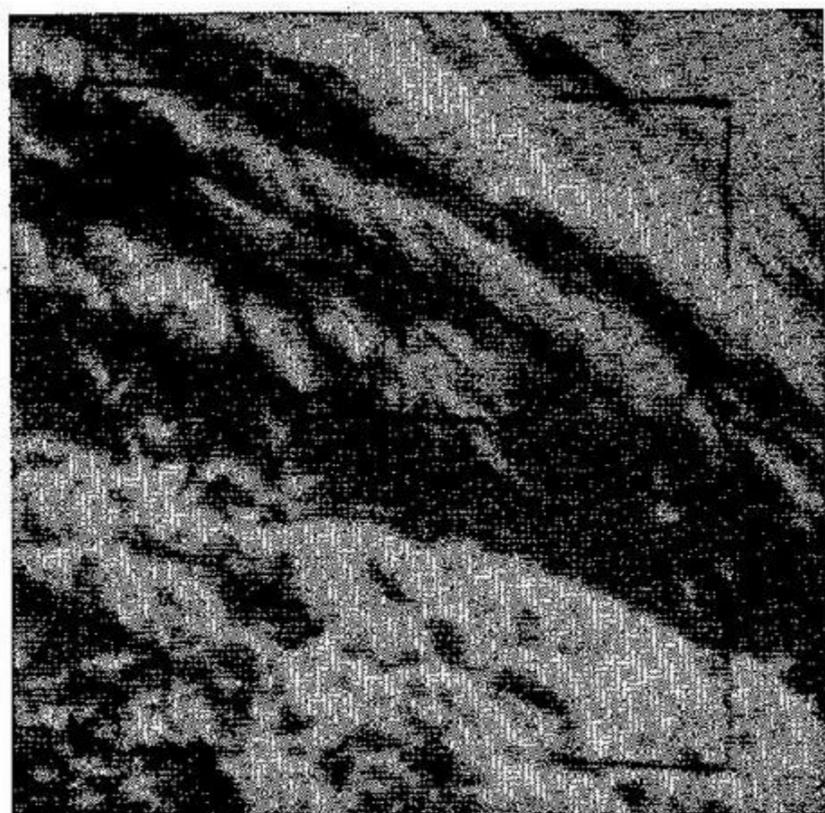
и той же конфигурации, то такая операция, очевидно, является пространственно-инвариантной. Действительно, она является сверткой изображения с функцией, равной $1/A$, в пределах этой окрестности (A — площадь окрестности) и нулю — за ее пределами. Эта операция может быть очень просто реализована оптическими средствами, например посредством расфокусирования изображения в оптической системе, или при контактном копировании изображения, если между изображением и фотоматериалом оставляют небольшой просвет. В первом случае для обеспечения хотя бы приблизительно пространственной инвариантности преобразования необходимо, чтобы края изображения находились достаточно близко к оптической оси. Примеры изображений, полученных в результате усреднения по окрестности различных размеров, показаны на фиг. 6.7. В общем случае усреднение по окрестности может производиться с некоторыми весами, например с весами, убывающими по мере удаления от центра окрестности.

Иногда выгодно использовать окрестности переменных размеров [34]. Например, можно осуществлять сглаживание двухградационного изображения, состоящего из рассеянных по полю единиц, находя в каждой точке минимальную круговую окрестность, содержащую заданное число единиц, и приписывая этой точке значение, обратно пропорциональное площади этой окрестности. Заметим, что эта операция уже не является линейной.

Для сглаживания изображения без его расфокусировки можно применять и другие нелинейные преобразования, включающие процедуру усреднения. В простейших видах таких преобразований сглаживание сочетается с применением пороговых операций.

Например, используя такое сочетание, можно отфильтровать импульсные помехи на изображении (одиночные черные точки на светлом фоне или светлые точки на темном). Для этого нужно заменять значение элемента, если оно отличается от среднего уровня по некоторой окрестности больше чем на некоторую пороговую величину, другим значением, более близким к этому среднему уровню [35]. Этот метод фильтрации проиллюстрирован на фиг. 6.8.

Еще один метод состоит в том, что в тех областях изображения, где контуры отсутствуют, производится изотропное усреднение, а при наличии контуров усреднение производится вдоль них [36] (критерии обнаружения контуров рассмотрены в следующем разделе).



Ф и г. 6.8. Применение нелинейного сглаживания изображения облачного покрова фиг. 3.1, а.

Пусть

A	B	C
D	E	F
G	H	I

— окрестность элемента E . Если $|E - 1/8(A + B + C + D + F + G + H + I)|$ (фиг. 6.11) больше $1/4$ максимально возможного значения E , то E заменяется на $1/8(A + B + C + D + F + G + H + I)$. (В противном случае E не изменяется.) Такая процедура сглаживает контрастные белые точки, а также белые черточки вблизи верхней опорной метки. Заметим, что она не приводит к расфокусировке всего изображения.

Упражнение 1. Пусть случайное изображение, определенное в упражнении 2 разд. 6.1, сглаживается путем замены нулей на единицы во всех точках, в окрестности которых больше чем m элементов из восьми равны еди-

нице. Сколько элементов изображения изменится при этом? Как это повлияет на среднеквадратичную гранулярность изображения (эффектами на границе изображения пренебречь)?

Упражнение 2. Пусть некоторое случайное изображение подвергается сглаживанию описанным выше методом нахождения окрестности, содержащей, например, заданное число k единиц. Какова среднеквадратичная гранулярность результирующего изображения как функции p и k (граничными эффектами пренебречь)?

Усреднение по нескольким копиям

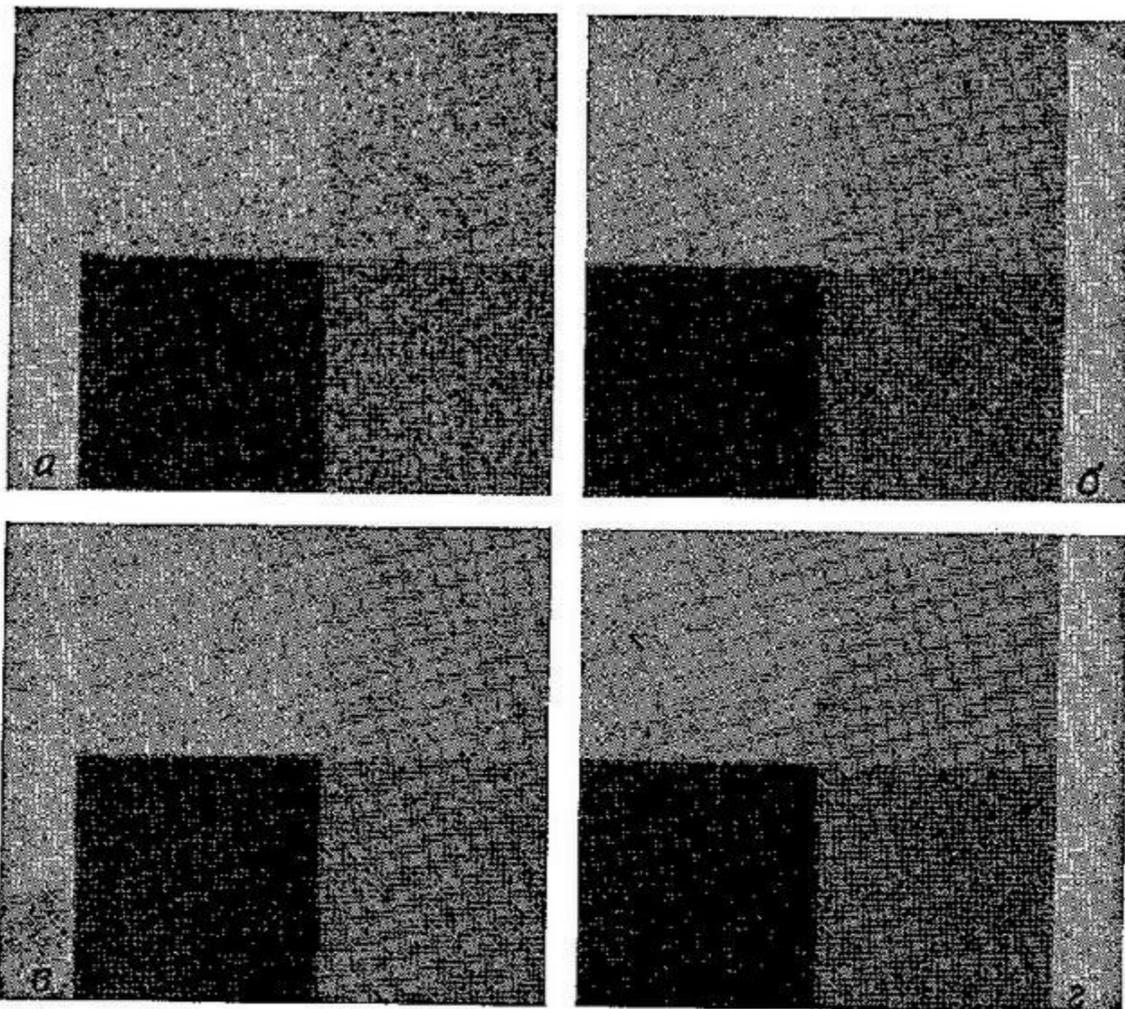
Ослабить шум на изображении, не ослабляя при этом самого изображения, можно путем усреднения нескольких независимо полученных зашумленных копий одного и того же изображения [37—39]. В принципе это можно осуществить путем оптического наложения изображений, но поскольку при этом требуется идеальная регистрирующая среда, то такой способ применим, если число копий n невелико. Пример, иллюстрирующий метод усреднения, показан на фиг. 6.9.

Чтобы понять принципы, лежащие в основе метода наложения, предположим, что изображение f зашумлено аддитивным нормальным шумом с нулевым средним значением и дисперсией σ^2 . (Конечно, распределение шума может быть только приближенно нормальным, так как f не может быть отрицательным.) Таким образом, в каждой точке (x, y) значение зашумленного изображения будет нормальной случайной величиной со средним значением $f(x, y)$ и дисперсией σ^2 .

Предположим, что можно сделать n копий исходного изображения таким образом, чтобы отсчеты шума на разных копиях были независимы. Тогда среднее значение по n копиям в точке (x, y) будет нормальной случайной величиной со средним значением $f(x, y)$ и дисперсией σ^2/n (см. Р а г з е н Е., *Modern Probability Theory and Its Applications*, Wiley, New York, 1960, p. 37).

Точно так же можно уменьшить зашумленность изображения симметричного объекта путем наложения со-

ответствующих частей изображения, например в результате поворотов, при которых объект переходит сам в себя [40, 41].



Ф и г. 6.9. Сглаживание путем усреднения по нескольким копиям. Число копий на изображениях а, б, в, г равно соответственно 2, 4, 8, 16 (исходные изображения, по которым выполнялось накопление, были вариантами изображения фиг. 6.6, а).

Мультипликативное наложение нескольких копий изображения (или многократное пропускание света через одну и ту же копию) как метод повышения контраста рассмотрено в работе [42—44].

Фильтрация пространственных частот

Как указывалось в разд. 6.2, одним из методов сглаживания изображения является подавление высших

пространственных частот в его преобразовании Фурье. Такая низкочастотная пространственная фильтрация приводит также к расфокусировке изображения и потому обычно нежелательна. Однако в некоторых случаях шум можно отфильтровать путем подавления отдельных участков пространственного спектра изображения, не затрагивая высшие пространственные частоты, чтобы не вызвать размытия контуров. В частности, если на изображение накладывается периодический шум (например, телевизионное изображение можно рассматривать как непрерывное изображение, на которое наложена сетка параллельных темных линий в промежутках между строками раstra), то он обычно может быть подавлен путем устранения соответствующего узкого участка в спектре зашумленного изображения. Точно так же растринтованное изображение можно превратить в непрерывное.

Хотя зерна в фотоэмульсии расположены нерегулярно, ослабляя пространственные частоты, соответствующие размерам зерен, можно достаточно эффективно бороться с зернистостью фотоэмульсии или других регистрирующих сред, работающих при ограниченном числе квантов [45—50].

Нанесение шума на изображение

Операции нанесения шума на изображение могут быть статистически пространственно-инвариантными в том смысле, что закон распределения шума одинаков для всех элементов изображения. Приведем два простых примера.

Аддитивный шум. Значение яркости каждого элемента изображения увеличивается или уменьшается на величину, имеющую заданный закон распределения (обычно с максимумом в нуле).

Случайное блуждание [51]. Каждая точка изображения меняется местами с другой точкой, находящейся от заданной в случайном направлении и на расстоянии, выбираемом из последовательности случайных величин с заданным законом распределения (обычно с максимумом в нуле).

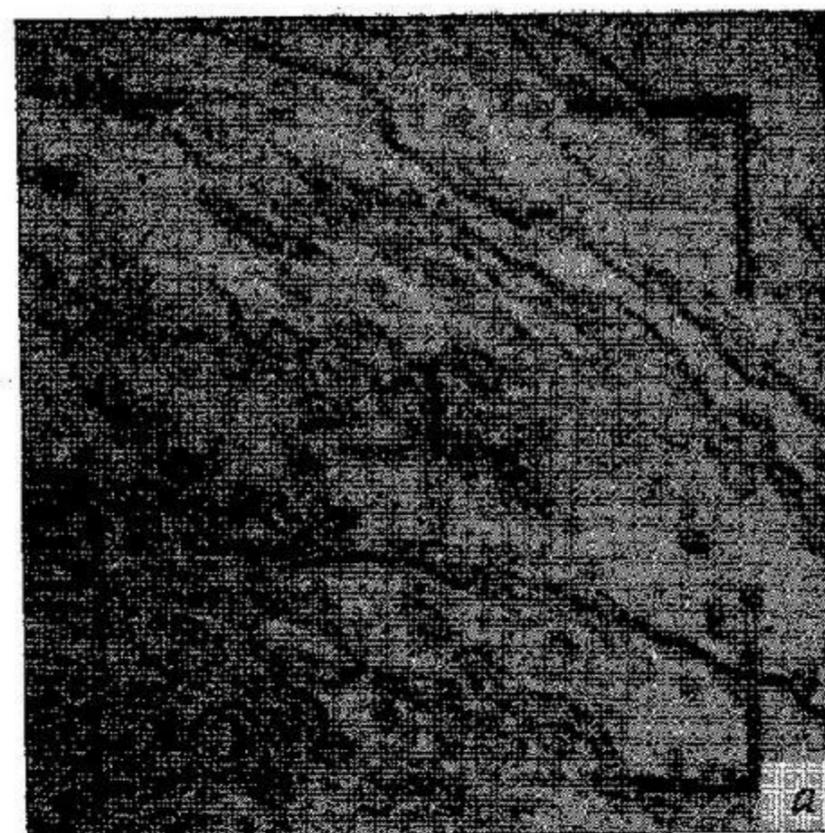
6.5. Улучшение качества изображения. Повышение резкости

Так же как многими способами можно сгладить, расфокусировать изображение, многими способами можно и скорректировать расфокусировку, повысить резкость изображения. Например, это можно сделать посредством квантования изображения («клиппирования») при воспроизведении его ксерографически или на очень контрастном фотоматериале или путем выделения высших пространственных частот [52]. В данном разделе описан ряд других методов повышения резкости.

Градиент

Поскольку расфокусировка изображения происходит в результате интегрирования (или усреднения), естественно, что для ее устранения необходимо применить некоторую дифференциальную операцию. Если коррекция должна быть выполнена только в каком-то определенном направлении, можно просто брать производную по этому направлению. В большинстве случаев, однако, требуется коррекция во всех направлениях. Ее можно выполнить, взяв в каждой точке производную в направлении градиента, т. е. в направлении наиболее быстрого изменения изображения [53]. Для гладких функций эта производная равна квадратному корню из суммы квадратов производных по любым двум взаимно ортогональным направлениям, т. е. $[(\partial f/\partial x)^2 + (\partial f/\partial y)^2]^{1/2}$. Для дискретных изображений производные нужно заменять разностями, т. е. вычислять $[(a_{i+1,j} - a_{ij})^2 + (a_{i,j+1} - a_{ij})^2]^{1/2}$. Подобные же результаты достигаются при замене квадратного корня из суммы квадратов разностей суммой их абсолютных значений. Более симметричный вариант схемы обработки получится при использовании попарных разностей значений яркости элементов, расположенных по обе стороны от элемента (i, j) . Примеры таких схем даны в работах [35, 68]. Результаты выделения градиента на дискретном изображении проиллюстрированы на фиг. 6.10.

Упражнение. а) Пренебрегая граничными эффектами, найти среднее значение дискретного градиента случай-



Ф и г. 6.10. Два варианта вычисления градиента для изображения облачного покрова (показанного на фиг. 3.1.).

$$a = \frac{((A+B+C) - (G+H+I))^2 + ((A+D+G) - (C+F+I))^2}{2}$$

$$b = |((A+B+C) - (G+H+I))| + |(A+D+G) - (C+F+I)|$$
 (обозначения, как на фиг. 6.8). Поскольку значения этого изображения редко изменяются очень быстро, значения градиента, вычисленные по разностям соседних элементов, будут, вообще говоря, меньшими.

ного изображения; б) найти среднее значение градиента на стыке двух случайных изображений с вероятностью единиц, соответственно равных p_1 и p_2 ; в) проделать аналогичные вычисления для других «разностных» величин, определенных в этом разделе (лапласиана, производной по направлению, результата «статистического дифференцирования»).

Производные по направлению (первые и более высокого порядка) можно, по крайней мере приближенно, получить оптически, выполнив свертку данного изображения с соответствующей маской ([56] гл. 5). Существуют также простые электронные средства приближенного вычисления производных по двум взаимно ортогональным направлениям или вычисления производной в направлении градиента в процессе сканирования изображения методом «изотропного сканирования». В этом методе сочетаются два взаимно ортогональных направления развертки, ориентированных под углами $\pm 45^\circ$. Так как значение градиента в некоторой точке приблизительно пропорционально разности между максимальным и минимальным значениями яркости на окружности с центром в данной точке, его также можно вычислить с помощью электронных устройств, используя быструю развертку по окружности с одновременным медленным перемещением центра окружности вдоль строк изображения и измеряя глубину амплитудной модуляции получаемого видеосигнала на частоте круговой развертки.

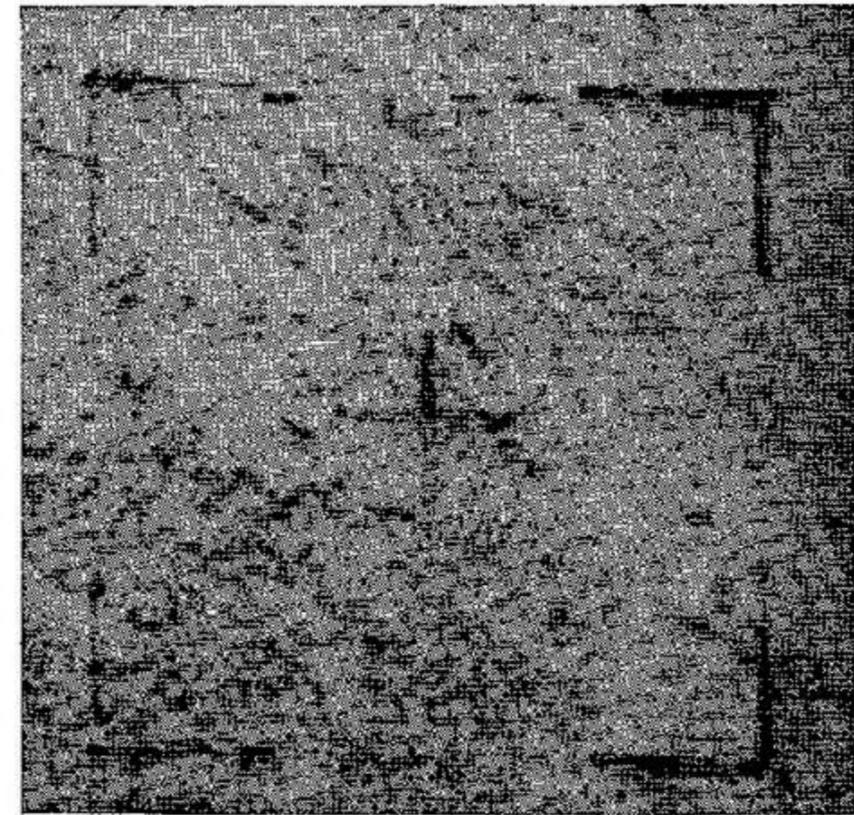
Лапласиан

Еще одним полезным дифференциальным оператором является лапласиан $\partial^2 f / \partial x^2 + \partial^2 f / \partial y^2$ [53, 54]. Для дискретного изображения он вычисляется по формуле

$$[(a_{ij} - a_{i-1, j}) - (a_{i+1, j} - a_{ij})] + [(a_{ij} - a_{i, j-1}) - (a_{i, j+1} - a_{ij})] = \\ = 4a_{ij} - (a_{i-1, j} + a_{i+1, j} + a_{i, j-1} + a_{i, j+1}).$$

Заметим, что значение лапласиана пропорционально разности между яркостью изображения в данной точке и средним значением яркости по «кольцевой» окрестности с центром в этой точке ([12], разд. 5.1, а также [55, 56]).

Приближение к лапласиану можно получить также с помощью свертки изображения с эталоном с положительным значением в центре и отрицательным по круговой периферии. Среднее значение этой маски должно быть равно нулю, т. е. положительный и отрицательный участки ее должны уравниваться. Это следует из того, что на равномерных участках изображения и на участках, где оно изменяется линейно, свертка его с такой маской



Ф и г. 6.11. Лапласиан изображения фиг. 3.1, а:

$$|E - \frac{1}{8}(A + B + C + D + F + G + H + I)|.$$

будет равна нулю, а на «контурах», где вторая производная отлична от нуля, эта свертка также будет отлична от нуля. Результат выполнения дискретного преобразования Лапласа над изображением показан на фиг. 6.11. В зрительной системе некоторых животных обнаружены группы нейронов, выполняющих над изображением на сетчатке глаза операции типа лапласиана. Эти нейроны образуют на сетчатке «рецептивные поля» с возбуждением

в центре и «торможением» по кольцевой окрестности или наоборот. В [57] приведены сведения о зрительной системе кошки, мечехвоста (*limulus*), лягушки, голубя, кролика. В некоторых случаях [58] может оказаться выгоднее выполнять преобразование Лапласа повторно.

Можно построить специализированные устройства, моделирующие обработку информации в зрительной системе мечехвоста, лягушки, голубя (подробности см. BPSS; H e r s c h e r M. B., K e l l e y T. P., Functional electronic model of the frog retina, *IEEE Trans. Military Electronics*, MIL-7, 98 (Apr. — July 1963); B e d d o e s M. P., C o n n o r D. J., M e l z a k Z. A., Simulation of a visual receptor network, *IEEE Trans. Bio-Medical Engineering*, B-ME-12, 136 (July—Oct. 1965); Cybernetic Problems in Bionics (O e s t e r r e i c h e r H. L., M o o r e D. L., eds.), Gordon and Breach, New York, 1968.

Особенно простой метод вычисления лапласиана основан на том, что значение лапласиана от изображения приблизительно эквивалентно разностям яркостей исходного изображения и его расфокусированной копии. Этот способ лежит в основе фотографического метода «нерезкого» маскирования. Он состоит в том, что регистрируется свет, прошедший через наложенные друг на друга позитивный и несколько расфокусированный негативный транспаранты. Эту операцию можно осуществить даже с одним транспарантом, используя люминесцентный экран с гашением (разд. 5.3), как показано на фиг. 6.12 [59, 60]. В этой системе при выключенном источнике инфракрасного излучения люминесцентный экран будет светиться равномерно и на неэкспонированной пленке будет зарегистрирована точная контактная копия исходного транспаранта. Если включить инфракрасный источник, его излучение, пройдя через транспарант и попав на люминофор, гасит его соответственно прозрачности транспарантов и вызывает на нем изображение, являющееся расфокусированной негативной копией транспаранта (расфокусированной потому, что между транспарантом и люминесцентным экраном имеется промежуток). Следовательно, когда свет от люминофора, пройдя через транспарант, попадает на фотопленку, регистрируется такое же изоб-

ражение, как если бы свет равномерно яркого люминесцентного экрана проходил через сложенные вместе позитивный транспарант и его расфокусированный негатив.

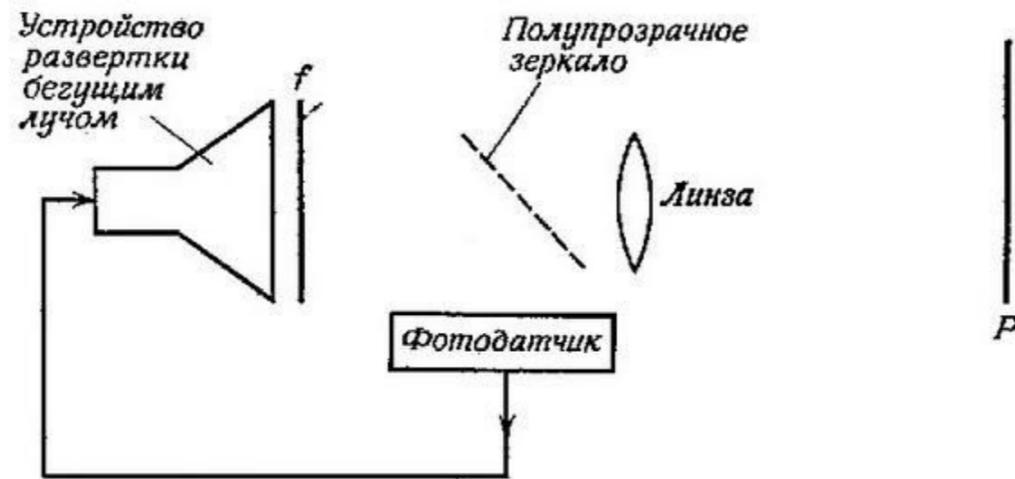
Стоит упомянуть еще несколько простых методов приближенного вычисления лапласиана при сканировании изображений [61—63]. Основная их идея состоит в одновременном сканировании двумя лучами — сфокусированным и расфокусированным — и вычитании получаемых видеосигналов. Еще проще использовать для сканирования один луч, фокусировка которого модулируется с большой частотой (намного выше частоты развертки).



Ф и г. 6.12. Метод нерезкой маски с использованием одного транспаранта.

Если вычислить разность получаемого при этом видеосигнала и видеосигнала, задержанного на полпериода частоты модуляции, то это будет эквивалентно вычитанию видеосигналов от сфокусированного и расфокусированного лучей. Существует еще более простой метод, заключающийся в том, что сканирующее пятно света проектируется на изображение через линзу с сильной хроматической аберрацией, так что для цветов на одном участке спектра он является сфокусированным, а для других цветов — расфокусированным. Если теперь использовать метод разделения цветов, описанный в разд. 5.3, получатся два канала — сфокусированного и расфокусированного пятна соответственно. Для вычисления лапласиана можно использовать также оптическую систему, показанную на фиг. 6.13. В этой системе благодаря наличию промежутка между сфокусированным пятном и транспарантом f

на фотодатчик попадает свет, прошедший через транспарант как бы от расфокусированного пятна. Сигнал фотодатчика уменьшает яркость пятна, так что f освещается как бы своим расфокусированным негативом. Таким образом, на плоскость P проектируется изображение, являющееся суммой сфокусированного оригинала и его расфокусированного негатива. (Сравнение метода нерезкого маскирования с другими методами, в которых размы-

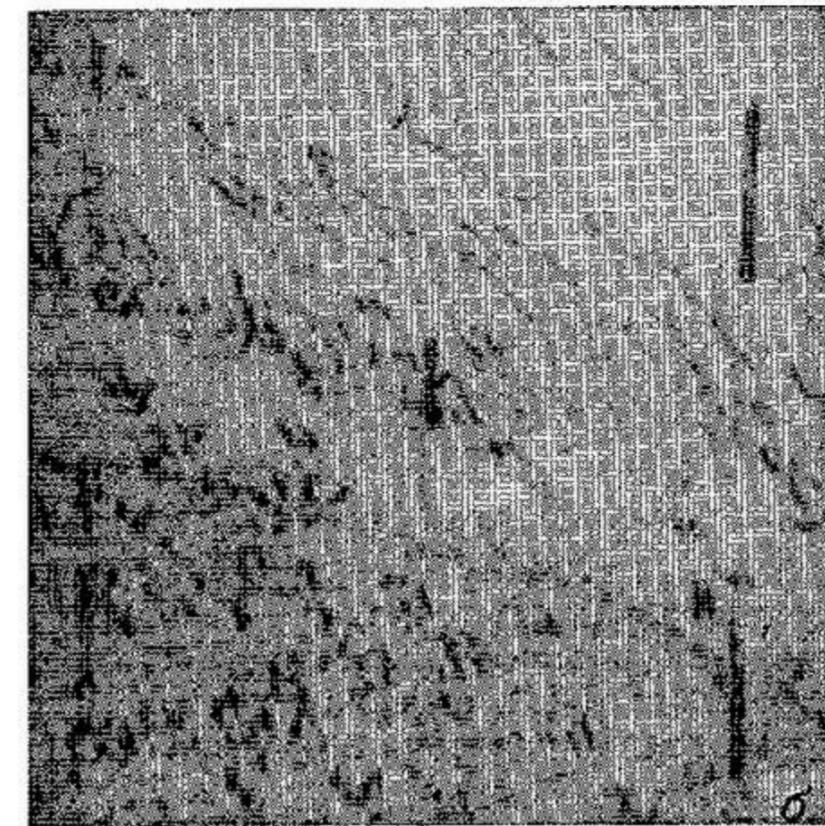
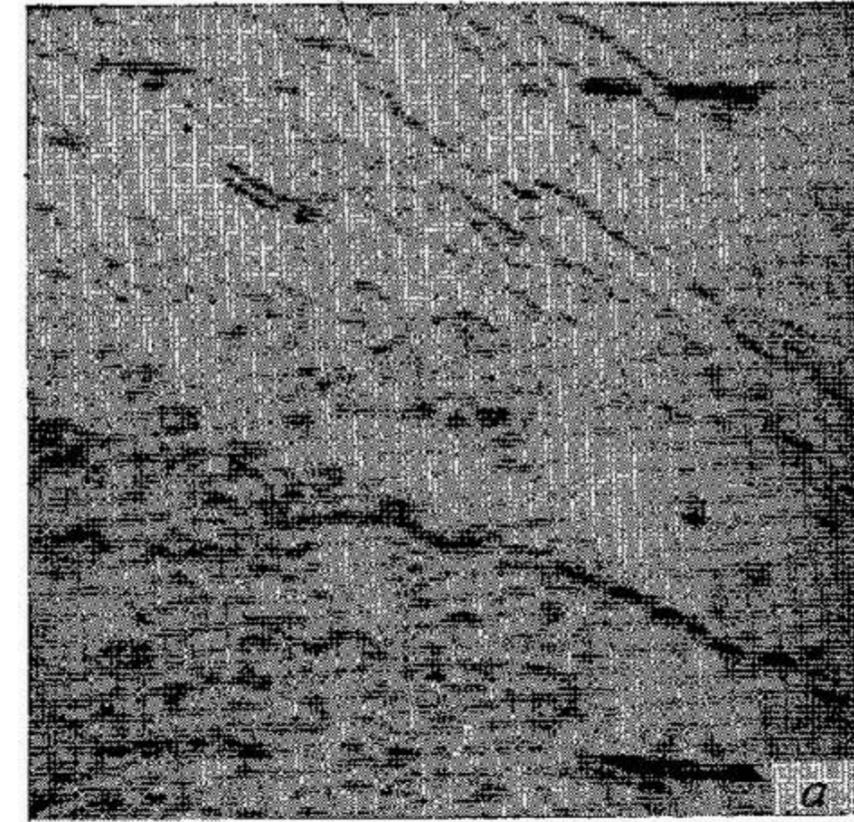


Ф и г. 6.13. Другой вариант метода нерезкой маски с использованием одного транспаранта.

тое изображение образуется посредством сдвига или размывания по кругу, а не оптической расфокусировки, приведено в работе [64].)

Производные и разности по направлению

Подчеркнуть контуры или линии, лежащие в одном направлении, и ослабить или подавить контуры и линии в других направлениях можно посредством дифференцирования или вообще дифференциальных операций, выполняемых по ортогональным направлениям. Например, для подчеркивания перепадов с наклоном m необходимо свернуть изображение с маской, которая имеет положительные значения по одну сторону от линии $y = mx$ и такие же, но отрицательные — по другую. (Результат такой операции для вертикальных и горизонтальных контуров показан на фиг. 6.14.) Точно так же для подчеркивания



Ф и г. 6.14. Результат вычисления разностей по направлениям для изображения фиг. 3.1, a .
 $a = |(A + B + C) - (G + H + I)|$; $б = |(A + D + G) - (C + F + I)|$.

линий с наклоном m следует использовать маску с положительной полоской вдоль оси $y = mx$ и отрицательными по обе стороны от нее. Подобные «рецептивные поля» из нейронов обнаружены в сетчатке глаз различных млекопитающих [57]. Заметим, что суммирование по этим полоскам, которое используется в детекторах контуров и линий, может быть выполнено путем сдвига (в направлении $y = mx$) и интегрирования. Отсюда следует метод обнаружения прямых контуров и линий, основанный на их самоконгруэнтности при сдвиге. О вероятном использовании этого метода в зрительной системе человека см. [65, 66]. В некоторых задачах предпочтительнее не аддитивные, а мультипликативные «маски» для выделения линий [67].

Статистическое дифференцирование

Вместо дифференцирования простым вычитанием средних значений яркости соседних окрестностей изображения (точки и кольца, двух соседних полосок и т. д.) можно сравнить распределение вероятностей значений яркости в этих областях. Ниже приведены два примера такого подхода.

1. Пусть \bar{f} есть результат сглаживания (т. е. усреднения) f по окрестности каждой точки, а σ — стандартное отклонение значений f по этой окрестности:

$$\sigma(x, y) = \left[\iint_N(x, y) (f - \bar{f})^2 \right]^{1/2},$$

где N — окрестность. Тогда вместо лапласиана $f - \bar{f}$ можно вычислять величину $(f - \bar{f})/\sigma$. Она будет максимальна на «изолированных точках», т. е. контрастных точках, окруженных равномерной окрестностью, и уменьшается, если окрестность содержит точки с различными значениями из всего диапазона значений. Случай, в котором окрестность является относительно большим кольцом, описан в [68, 69], а также [39] гл. 8.

2. Пусть D_1 и D_2 — распределение значений яркости для соседних областей. Тогда $|D_2 - D_1|$ может служить мерой различия между этими областями и использоваться для проведения границы между ними [70].

К видеосигналу, полученному при развертке дискретного изображения, можно применить следующий одномерный статистический критерий выделения контуров. При заданной длине серии для класса изображений, к которому принадлежит данное изображение, и статистике шума можно вычислить условные вероятности того, что соседние элементы имеют одно и то же значение при условии, что они принадлежат к одной и той же или разным сериям. По теореме Байеса можно определить условную вероятность того, что данные два соседних элемента принадлежат к различным сериям, т. е. лежат по разные стороны от контура. Если эта вероятность больше $1/2$, можно принять, что в данном месте проходит контур, даже в случае, когда эти элементы имеют одно и то же значение. Наоборот, если эта вероятность меньше $1/2$, но элементы имеют различное значение, можно предположить, что эта разница возникла в результате действия шума, и изменить их значение для устранения выброса шума.

ЛИТЕРАТУРА

1. Sowers H. W., Star pattern recognition — a survey of the literature, AD 421473 (Oct. 1963).
2. Hobrough C. L., Automation in photogrammetric instruments, *Photogrammetric Engrg.*, **31**, 595 (1965).
3. Cathey W. T., Jr., Doidge J. G., Image comparison by interference, *J. Opt. Soc. Amer.*, **56**, 1139 (Aug. 1966).
4. Rau J. E., Detection of differences in real distributions, *J. Opt. Soc. Amer.*, **56**, 1490 (Nov. 1966).
5. Diamantides N. D., Correlation measure of contrast for map matching, *J. Opt. Soc. Amer.*, **58**, 996 (July 1968).
6. Bliss J. C., Crane H. D., Relative motion and nonlinear photocells in optical image processing, OEOIP, p. 615.
7. Harris J. L., Resolving power and decision theory, *J. Opt. Soc. Amer.*, **54**, 606 (May 1964).
8. Chow C. K., An optimum character recognition system using decision function, *IRE Trans. Electronic Computers*, **EC-6**, 247 (Dec. 1957).
9. Montgomery W. D., Broome P. W., Spatial filtering, *J. Opt. Soc. Amer.*, **52**, 1259 (Nov. 1962).
10. Trabka E. A., Roetling P. G., Shape detection using incoherent illumination, *J. Opt. Soc. Amer.*, **57**, 108 (Jan. 1967).
11. Fitzmaurice J. A., Reading Russian scientific literature, OCR, p. 61.

51. White B. W., The computer as a pattern generator for perceptual research, *Behavior. Sci.*, 6, 252 (July 1961).
52. Birch K. G., A spatial frequency filter to remove zero frequency, *Optica Acta*, 5, 113 (March—Apr. 1967).
53. Kovasznaý L. S. G., Joseph H. M., Image processing, *Proc. IRE*, 43, 560 (May 1955).
54. Kulikowski J. J., Adaptive visual signal preprocessor with a finite number of states, *IEEE Trans. Systems Science and Cybernetics*, SSC-2, 96 (Dec. 1966).
55. Taylor W. K., Pattern recognition by means of automatic analogue apparatus, *Proc. IEE*, 106B, 198 (March 1959).
56. Nagy G., Preliminary investigation of techniques for automated reading of unformatted text, *Comm. ACM*, 11, 480 (July 1968).
57. Ratliff F., *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, Holden-Day, San Francisco, 1965.
58. Bell D. A., Computer aided design of image processing techniques, *CPR*, p. 282.
59. Watson A. J., The Fluoro-Dodge method for contrast control, *Photogrammetric Engrg.*, 24, 638 (Sept. 1958).
60. Clarke A. B., A photographic edge-isolation technique, *Photogrammetric Engrg.*, 28, 393 (July 1962).
61. Levine S. W., Mate H., Selected electronic techniques for image enhancement, Paper № 11, Proceedings, Image Enhancement Seminar, Soc. Phot. Instr. Engrs., Redondo Beach, California, March 1963.
62. Craig D. R., Disenhancement — a negative approach to a positive problem, Paper № V, Proceedings, Image Enhancement Seminar, Soc. Phot. Instr. Engrs., Redondo Beach, California, March 1963.
63. Hannum A. J., Techniques for electronic image enhancement, Paper № VII, Proceedings, Image Enhancement Seminar, Soc. Phot. Instr. Engrs., Redondo Beach, California, March 1963.
64. Armitage J. D., Lohmann A. W., Herrick R. B., Absolute contrast enhancement, *Appl. Opt.*, 4, 445 (Apr. 1965).
65. Platt J. R., Functional geometry and the determination of patterns in mosaic receptors, *Information Theory in Biology*, Pergamon Press, New York, 1958, p. 371.
66. Platt J. R., How a random array of cells can learn to tell whether a straight line is straight, in «Principles of Self-Organization» (H. von Foerster, G. M. Zopf, Jr., eds.), Pergamon Press, New York, 1962, p. 315.
67. Parks J. R., Elliot J. R., Cowin G., Simulation of an alphanumeric character recognition system for unsegmented low quality print, *CPR*, p. 95.
68. Holmes W. S., Leland H. R., Richmond G. E., Design of a photo interpretation automaton, *FJCC*, 27 (Dec. 1962).
69. Holmes W. S., Automatic photointerpretation and target location, *Proc. IEEE*, 54, 1679 (Dec. 1966); русский перевод: Холмс Автоматическая фотоинтерпретация и обнаружение целей, *ТИИЭР*, 54, № 12 (1966).

70. Muerle J. L., Allen D. C., Experimental evaluation of techniques for automatic segmentation of objects in a complex scene, *PPR*, p. 3.
71. Кубба М. Н., Automatic picture detail detection in the presence of random noise, *Proc. IEEE*, 51, 1518 (Nov. 1963); русский перевод: Кубба, Автоматическое обнаружение деталей изображений при наличии шума, *ТИИЭР*, 51, № 11 (1963); см. также Секки, обнаружение деталей в телевизионном сигнале, *ТИИЭР*, 53, № 1 (1965).

Свойства изображений и обработка их методами теории распознавания образов

7.1. Распознавание образов

Формальное описание изображений, как уже отмечалось в разд. 1.2, — одна из основных проблем, возникающих при использовании вычислительных машин для их обработки. Набор записанных в соответствующем языке истинностных высказываний, характеризующих изображение, может в принципе служить его формальным описанием (гл. 10). Часто, однако, требуется лишь *классифицировать* изображение, отнеся его к одной из категорий (определенных, как правило, заранее). В частности, для изображения, представляющего отдельный печатный символ, такими категориями могли бы быть буквы от A до Z , а для изображения, воспроизводящего некий вид облачного покрова, категории классификации задаются понятиями типа «перистые облака», «кучевые облака» и т. д.

Задача автоматической классификации в общем виде сводится к отысканию некоторой функции, отображающей множество образов (в данном случае изображений) во множество, элементами которого являются классы образов. В большинстве случаев процесс определения такой функции целесообразно проводить в несколько этапов.

1. *Предварительная обработка.* Заданное изображение f с помощью некоторого набора или последовательности операций преобразуется в одно или несколько новых изображений f_1, \dots, f_k (реализация такого преобразования и является «предварительной обработкой»).

2. *Выделение признаков.* Функции f_i подвергаются функциональному преобразованию $\mathfrak{F}_1, \dots, \mathfrak{F}_m$, определяющему свойства (признаки), в результате чего изображения кодируются действительными числами. Иногда этот процесс удобнее осуществлять в две ступени, отображая в

первую очередь изображения в функции одной переменной, а затем преобразуя их в действительные числа.

3. *Классификация.* В результате выполнения этапов 1, 2 появляется набор n ($n \leq mk$) действительных чисел, которые можно считать признаками исходного изображения f ; подобный набор можно рассматривать как точку в n -мерном пространстве. Если указаны области, занимаемые тем или иным классом в этом пространстве, либо на нем задана плотность вероятности для каждого класса, то, исходя из соображений геометрической близости или максимальной вероятности, данное изображение можно отнести к определенному классу. Трудность практической реализации такого подхода заключается в том, что границы областей, так же как и вероятности, обычно неизвестны.

Классификациям, выполняемым на основе оценки заданного набора признаков (свойств), посвящена обширная литература, так как эта задача не является специфической именно в применении к изображениям, в дальнейшем она обсуждаться не будет. Смысл этапа 1 — подготовить изображение к обработке таким образом, чтобы облегчить выделение на этапе 2 признаков, наилучших с точки зрения простоты или качества классификации. Методы, обеспечивающие выполнение этапа 1, были обсуждены в гл. 4—6. В этой и двух последующих главах будет рассмотрен этап 2, в частности те виды функционального определения признаков, которые хорошо «проявили себя» при решении различных задач, связанных с распознаванием изображений.

При обсуждении свойств изображений целесообразно рассматривать отдельно признаки, имеющие естественное определение на изображении произвольного вида, и признаки, возникающие только в тех случаях, когда выделяется специальное подмножество изображений (объект). Так, не имеет смысла говорить о признаках типа длины или конфигурации изображения, поскольку обычно приходится иметь дело с изображениями стандартного формата; с другой стороны, очень часто значительный интерес вызывают размеры и конфигурации изображенных объектов. Методы выделения объектов на изображениях («разбиение» изображений) будут обсуждаться в гл. 8, а геометри-

ческие свойства подмножеств изображений — в гл. 9. Данная же глава в первую очередь посвящена признакам, имеющим естественное определение даже в тех случаях, когда разбиение множества изображений на подмножества не определено.

7.2. Отбор признаков

Теоретически, имея набор признаков, можно подсчитать вклад каждого из них в классифицирующее решение [23—28], например в виде среднего количества информации о принадлежности образа некоторому классу, доставляемой признаком, если его значение определено. В тех случаях, когда подобная вычислительная процедура практически нереализуема, эффективность признаков можно оценить, реализовав или промоделировав распознающую систему [29—36]. Многие системы распознавания образов наделены способностью видоизменять используемые признаки, с тем чтобы выделить из них наиболее эффективные [37—40]. Известны другие виды «обучения» систем распознавания образов, предусматривающие не собственно замену признаков, а изменение весов, с которыми они входят в решающее (классифицирующее) правило. Мы, однако, не будем останавливаться на них, так как для распознавания изображений они не специфичны.

Выбор «хорошего» набора признаков — первоочередная проблема, непосредственно связанная с *распознаванием изображений*. Создается впечатление, что во многих известных системах распознавания выбор признаков определялся не поставленной задачей классификации, а в основном наличием подходящего математического аппарата или легкостью воспроизведения (например, возможность вычисления значения признака с помощью оптических средств). В других случаях задание признаков (как и операций предварительной подготовки) обуславливалось желанием либо имитировать физиологическую специфику зрительного тракта, либо попытаться смоделировать процесс зрительного восприятия (см. работы, помещенные в сборниках BPSS, PR и MPSVF, а также [41, 42]; характерными примерами могут служить описанные в разд. 6.5 преобразование Лапласа и «детектор края»). Обоснованием подобного подхода служит то об-

стоятельство, что во многих случаях определить границы классов оказывается не просто; при этом единственным доводом в пользу существования решения рассматриваемой задачи классификации является способность людей (или животных) правильно классифицировать данные образы. С другой стороны, целый ряд задач классификации объектов весьма просто решается с помощью операций предварительной обработки (например, включающих преобразование Фурье) или функционального выделения признаков, не имеющих почти никакого отношения к методам, используемым живыми организмами.

Наиболее логично выбирать признаки и процедуры предварительной обработки, исходя из сведений относительно «структуры» образов и определений классов. Если, например, предъявляются образы, сформированные из различных линий (в частности, рукописные и печатные буквы, треки ядерных частиц в пузырьковой камере и т. д.), то отдельные типы признаков и процедур предварительной обработки, очевидно, подходят больше, чем другие. Даже частичное описание структуры классов (буквы обычно делятся на ряд основных «штрихов» и т. д.) может оказаться очень полезным, так как систему распознавания можно «запрограммировать» на поиск специфических «локальных» образов, «фигур» с определенными очертаниями и т. д. В настоящее время разрабатываются разного рода «грамматики изображений» — формальные системы, в которых можно выразить все многообразие описаний образов и классов (гл. 10). До сих пор, однако, эта методика нашла весьма ограниченное применение.

Хотя к настоящему времени нет общей теории, позволяющей при известных определениях классов выделять «оптимальные» признаки, существует ряд особых случаев, когда оптимум известен. Допустим, например, что все принадлежащие данному классу образы можно получить, налагая шум на эталонный образ («прототип»). В таком случае (разд. 6.1) оптимальность определяется согласованностью с эталоном [43—46]. Можно рассмотреть более общую ситуацию, когда существует некоторое множество эталонных «подобразов» и каждое конкретное изображение представляет собой комбинацию этих подобразов (плюс шум).

Во многих случаях два различных образа следует отнести к одному классу, если они отличаются лишь положением, ориентацией, размерами и т. д. либо каждый из них может быть получен из другого с помощью таких простейших операций над элементами изображения, как, например, линейное преобразование яркости. Дело, однако, не всегда обстоит подобным образом; например, запятая и апостроф могут отличаться лишь положением, а буквы d , p — только ориентацией и т. д. Другими словами, во многих случаях классы оказываются *инвариантными* к преобразованиям определенных типов (параллельный перенос, поворот, увеличение, гамма-коррекция). Это обстоятельство позволяет пользоваться «сокращенными» описаниями классов, основанными на использовании свойств, инвариантных относительно одних и тех же преобразований. В следующих разделах будут рассмотрены методы определения инвариантных свойств.

1. В некоторых случаях допустимо подвергать каждое изображение всем возможным преобразованиям, выполняя их либо последовательно, либо параллельно, и находить значения признаков для преобразованных изображений. Так, например, при определении взаимной корреляции эталона и изображения вычисляется их «сходство» во всех возможных позициях. При выполнении всех возможных преобразований набор значений признаков становится, естественно, очень большим, однако инвариантные признаки в этом наборе можно выделить, используя усреднение, нахождение максимума и т. д.

2. Если набор преобразований имеет простое математическое описание, то часто оказывается возможным определить инвариантные свойства чисто аналитически. Для тех свойств фигур, которые оказываются инвариантными различным группам преобразований, можно найти соответствующую геометрическую подоплеку. В частности, расстояния между точками фигуры инвариантны относительно операции переноса и вращения, а отношения расстояний — изменению масштаба; следовательно, эти расстояния (или их отношения) можно использовать в качестве признаков при отнесении данной фигуры к одному из классов конгруэнтных (подобных) фигур (разд. 7.3, а также [47]).

3. Часто оказывается возможным ввести процедуру «нормализации» изображений; после применения такой процедуры два изображения становятся идентичными, если они отличались на некоторое преобразование, входящее в заданное множество. Следовательно, каждый признак нормализованного изображения является признаком исходного изображения, инвариантным относительно определенных преобразований. Ниже рассматриваются три соответствующих примера.

а. Изображение может быть подвергнуто такому преобразованию переноса, что его центр тяжести совместится с началом используемой координатной системы. (В этом случае мы исходим из допущения, что яркость изображения в данной точке представляет ее «массу».) Аналогичным образом можно стандартизовать ориентацию изображения, проведя, например, через его центр тяжести прямую таким образом, чтобы момент инерции изображения относительно нее был минимален, а затем поворачивая изображение до тех пор, пока эта прямая не совпадет с одной из осей координат. Следует, однако, заметить, что прямая, момент инерции относительно которой минимален, может быть не единственной. Можно считать, что глаз в какой-то степени использует нормализацию при «фиксации» («центрировании») попавшей в поле зрения фигуры, а также при адаптации к диапазону яркостей изображений.

б. Автокорреляционная и спектральная функции изображения, как показано в гл. 4, инвариантны относительно его перемещения. Так, определение сходства с эталоном с помощью автокорреляционных или спектральных функций лежит в основе позиционно-инвариантного метода обнаружения заданного изображения [48, 49]. Аналогичным образом частные значения автокорреляционной функции, соответствующие определенным сдвигам данного изображения относительно самого себя, образуют группу признаков, инвариантных относительно перемещения [50]. (Известны данные, свидетельствующие о том, что глаза некоторых насекомых наделены способностью «вычислять» значения автокорреляционной функции для своего поля зрения [51, 52].)

Следует подчеркнуть, что значения автокорреляционной функции не определяют исходное изображение «с точностью до перемещения»; две функции могут иметь одну и ту же автокорреляционную функцию даже и в том случае, если изображения не только смещены друг относительно друга. Если, например, g получено из f с помощью поворота на 180° [т. е. если $g(x, y) = f(-x, -y)$], то автокорреляционная функция $g \otimes g = \iint f(-x+u, -y+v) \cdot f(-x, -y) dx dy = \iint f(-x+u, -y+v) \times \times f(-x, -y) d(-x) d(-y) = f \otimes f$. Еще один тривиальный случай: если функции не должны быть обязательно неотрицательными, то очевидно, что $(-f) \otimes (-f) = f \otimes f$. В качестве нетривиального примера рассмотрим случай, когда $f = (a_{ij})$ и $g = (b_{ij})$ — дискретные изображения, причем $a_{11} = 9, a_{12} = 6, a_{13} = 1, b_{11} = b_{13} = 3, b_{12} = 10$, а все остальные значения a и b равны нулю; в таком случае не составляет труда убедиться в том, что $f \otimes f = g \otimes g$.

Упражнение 1. Обобщим только что рассмотренный пример следующим образом: пусть функция $f = (a_{ij})$ такова, что все элементы a_{ij} (кроме, быть может, a_{11}, a_{12} и a_{13}) — нулевые; в этом случае автокорреляционная функция $f \otimes f$ не изменится при замене трех этих элементов многочленами $\frac{1}{2} [a_{13} + (a_{13}^2 - 4a_{11}a_{13})^{1/2}]$, $a_{11} + a_{13}$ и $\frac{1}{2} [a_{12} - (a_{12}^2 - 4a_{11}a_{13})^{1/2}]$ соответственно.

Упражнение 2. В качестве f возьмем ту же функцию, что и в упражнении 1. Следует показать, что автокорреляционная функция $f \otimes f$ остается неизменной только при таких манипуляциях с элементами a_{11}, a_{12} и a_{13} , которые представляют собой комбинации следующих процедур: а) описанной в упражнении 1; б) взаимной перестановки элементов a_{11} и a_{13} ; в) изменения знаков у всех элементов a_{ij} . Получить обобщение этого результата для случая функции f , у которой $k \geq 3$ элементов могут быть ненулевыми.

Упражнение 3. Рассмотрим функции $f = (a_{ij})$ и $g = (b_{ij})$, такие, что ненулевыми являются не более чем два элемента каждой из них. Показать, что равенство $f \otimes f = g \otimes g$ выполняется только в том случае, если

функция g может быть получена из функции f посредством комбинации операций переноса, поворота на 180° и изменения знака. Другими словами, следует показать, что если этими элементами функции f являются a_{rs} и a_{uv} , то в функции g этими элементами окажутся $b_{r+h, s+k}$ и $b_{u+h, v+k}$ при определенных h и k и, более того, при этих условиях справедливы либо соотношения $a_{rs} = \pm b_{r+h, s+k}$ и $a_{uv} = \pm b_{u+h, v+k}$, либо обратные им.

Можно показать [53], что если две функции имеют одинаковые автокорреляции *всех порядков*, определяемые выражением

$$\iint f(u, v) f(u+x_1, v+y_1) \dots f(u+x_k, v+y_k) du dv \quad (k=1, 2, \dots),$$

то они отличаются лишь на перенос. В самом деле [54], для функций, либо имеющих «дискретную» природу, либо преобразование Фурье от которых является аналитической функцией, равенство автокорреляции второго (или любого четного) порядка гарантирует, что эти функции отличаются только на перенос.

Аналогичным образом можно обеспечить инвариантность относительно поворотов. Так как изображение $f(r, \theta)$, выраженное в полярных координатах, периодически с периодом 2π относительно полярного угла θ , модуль разложения этой функции в ряд Фурье (для всех r) инвариантен относительно поворота. Преобразование, инвариантное как повороту, так и изменению масштаба исходного изображения, можно определить с помощью вычисления «автокорреляции» в системе полярных координат [55]:

$$g(u, v) = \int_0^{2\pi} \int_0^\infty f(r, \theta) f(ur, \theta + v) r dr d\theta.$$

Это преобразование, очевидно, не зависит от поворота изображения, представленного функцией f , изменение же масштаба приводит только к появлению постоянного множителя, влияние которого можно исключить, скажем, введя отношение величин g . Если подобную операцию применить не к собственно функции f , а к ее автокорреляции, то в результате определяется преобразование, которое инвариантно также и переносу исходного изобра-

ражения. Известен еще один способ обеспечения инвариантности [56]: он основан на использовании спектральной функции, выраженной через полярные координаты, для сведения поворотов и изменений масштаба исходного изображения к его переносам. В работе [53] также описывается метод, обеспечивающий инвариантность к изменениям масштаба.

в. Если преобразования линейны и образуют группу относительно композиции функций, преобразовать изображение в инвариантную форму можно, применив к нему все определенные преобразования и просуммировав (или усреднив) полученные в результате изображения. Если, например, сфотографировать с выдержкой поворачиваемое изображение, то получится изображение, симметричное относительно оси вращения и не зависящее от ориентации оригинала. Это нормализованное изображение, очевидно, не представляет исходное «с точностью до поворота». Предположения о том, что мозг может обеспечивать инвариантность аналогичными способами, рассмотрены в работах [57, 58]; некоторые аспекты упомянутых проблем обсуждаются в работах [59—63].

7.3. Классы признаков

Из множества возможных видов признаков изображений для целей распознавания использовались лишь некоторые. В этом разделе рассмотрен ряд важных классов таких признаков. В данном случае под признаком изображения имеется в виду действительное число, полученное в результате применения к изображению некоторого функционала. В следующем разделе будут обсуждены признаки, имеющие функциональный характер.

Большое значение имеет двuzначный признак, способный принимать только значения 0 и 1. Такой признак может рассматриваться как предикат, являющийся для данного изображения истинным или ложным. (Имея признак \mathfrak{F} , принимающий действительные значения, можно, используя некий «порог», перейти к двuzначным признакам \mathfrak{F}_t , считая, что $\mathfrak{F}_t = 1$, если $\mathfrak{F} \geq t$, и $\mathfrak{F}_t = 0$ в противном случае, где t — действительное число.)

В случае двuzначного изображения двuzначный признак является обычной булевой функцией конечного числа аргументов.

Локальные и структурные признаки

Значение признака $\mathfrak{F}(f)$ определяется не полным множеством значений функции f , а лишь некоторым их подмножеством; другими словами, оно зависит от значений, которые функция f принимает на ограниченном участке плоскости аргументов. Если считать, что функции, представляющие изображение, определены на конечном множестве A (например, если они являются дискретными), то, очевидно, существует минимальное подмножество B , такое, что для всех функций f множества A значения признака $\mathfrak{F}(f)$ определяются сужением f на B . Это подмножество B называют *опорным множеством* функционала \mathfrak{F} . Будем считать признак \mathfrak{F} *локальным*, если диаметр множества B , или, что то же самое, количество образующих его точек, мал. Так, если, например, операция φ — локальная (разд. 4.1), то значение функции $\varphi(f)$ в любой точке представляет собой локальный признак изображения f ; так как оно определяется исключительно значениями яркости изображения в окрестности каждой конкретной точки. Крайний случай локального признака представляет так называемый «точечный признак» — функция от яркости изображения в данной точке. Можно показать [64—66], что для разделения n -различных образов достаточно учесть значения не более чем $n - 1$ точечных признаков.

Локальные признаки целесообразно использовать в основном в тех случаях, когда изображение, являющееся объектом рассмотрения, можно нормализовать, так как иначе эти признаки, вообще говоря, будут принимать разные значения даже на изображениях, отличающихся только на перенос. В работах [67—69] приведены примеры локальных признаков, которые были использованы при распознавании символов, подвергнутых позиционной нормализации. Применение локальных признаков, характеризующих спектральную функцию изображения, иллюстрируется в работах [70—71]. Установлено, что

можно сделать локальный признак в некоторой степени нечувствительным к положению изображения, уменьшив слегка его резкость перед измерением значения признака [72].

Упражнение 1. Определить набор точечных признаков, характеризующих автокорреляцию, с помощью которого можно было бы различить шесть образов, приведенных в упражнении 1 разд. 6.1, независимо от занимаемого ими положения.

Инвариантные признаки можно получить, оперируя и с ненормализованным изображением; для этого следует проанализировать *распределение вероятности* значений, принимаемых на изображении любым локальным признаком. В частности, статистические характеристики подобного распределения, такие, например, как его среднее либо доля площади изображения, на которой значения принадлежат заданному диапазону, инвариантны. Будем называть статистические характеристики локальных признаков *текстурными признаками*. Признаки такого типа использовались для автоматической классификации фотографий земной поверхности и облачного покрова, сделанных с самолетов и искусственных спутников [73—80].

Упражнение 2. Для случайного дискретного изображения (упражнение 2, разд. 6.1) подсчитать в функции от p средние значения следующих локальных признаков: а) яркости изображения; б) градиента; в) лапласиана.

Значение «визуальной структуры» изображения для его восприятия подчеркивалось Гибсоном [81]; говоря о «структуре», он, по-видимому, имел в виду средний размер элементов, образующих изображение. Следует обратить внимание на то обстоятельство, что эта величина (средний размер) обратно пропорциональна количеству границ элементов на единицу площади [82], которое для изображений с фиксированной резкостью можно рассматривать в качестве среднего значения локального признака (например, градиента или лапласиана). Одни исследователи предложили меры структуры, подобные тем, которые используются для измерения фотографической гранулярности (разд. 6.3), другие — основанные на

учете распределения условных вероятностей значений яркости элемента изображения при заданной интенсивности одного или нескольких «предшествующих» элементов [83, 84].

Линейные признаки

В разд. 4.2 было установлено, что любая позиционно-инвариантная линейная операция представляет собой свертку. Аналогично можно показать, что любой линейный признак, удовлетворяющий определенным ограничениям, является мерой согласованности с «эталоном», т. е. интегралом произведения некоторой «эталонной» функции и предъявленного изображения. Пусть, в частности, \mathcal{S} — множество всех функций двух действительных переменных, которые определены и интегрируемы на плоскости и для которых значение $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2$ конечно. В таком случае мы приходим к следующему утверждению.

Теорема. Пусть \mathfrak{F} — действительная линейная функция, определенная на множестве функций \mathcal{S} . Допустим, что существует действительное число M , такое, что для всех функций f , принадлежащих множеству \mathcal{S} , справедливо неравенство $\mathfrak{F}(f) \leq M (\iint f^2)^{1/2}$. В этом случае во множестве функций \mathcal{S} всегда найдется функция g , такая, что $\mathfrak{F}(f) = \iint gf$ для всех функций f , принадлежащих множеству функций \mathcal{S} .

Доказательство. Пусть M^* — наименьшее из чисел M , для которых посылка теоремы выполняется, и, кроме того, допустим, что функция g^* такова, что $\mathfrak{F}(g^*) = M^* (\iint g^{*2})^{1/2}$, причем $\iint g^{*2} = 0$.

Существование таких числа M^* и функции g^* следует из «полноты» множества \mathcal{S} . Эта теорема известна как теорема Рисса о представлении. (Этому вопросу посвящена, например, монография Halmos P. R., *Introduction to Hilbert Space and the Theory of Spectral Multiplicity*, Chelsea, New York, 1957, p. 31)¹. Считаем, что h выбирается таким образом, что $\mathfrak{F}(h) = 0$; тогда для любого действитель-

¹ См. также Халмош П., Гильбертово пространство в задачах / изд-во «Мир», М., 1970. — Прим. перев.

ного λ выполняются соотношения $M^{*2} \iint g^{*2} = \int (g^*)^2 = \int (g^* - \lambda h)^2 \leq M^{*2} \iint (g^* - \lambda h)^2 = M^{*2} (\iint g^{*2} - 2\lambda \iint g^* h + \lambda^2 \iint h^2)$ и, следовательно, $0 \leq \lambda^2 \iint h^2 - 2\lambda \iint g^* h$. Если $\iint h^2 \neq 0$, полагаем, что $\lambda = \iint g^* h / \iint h^2$; в таком случае выполняется неравенство $0 \leq -(\iint g^* h)^2$ и, следовательно, $\iint g^* h = 0$.

С другой стороны, если $\iint h^2 = 0$, то на основании неравенства Шварца получаем $0 \leq \iint g^* h \leq (\iint g^{*2} \times \iint h^2)^{1/2} = 0$; таким образом, мы показали, что $\iint g^* h = 0$ для любого h , при котором выполняется условие $\int (h) = 0$. Для любой функции f , однако, можно записать, что $f = [f - \int (f) g^* / M^* (\iint g^{*2})^{1/2}] + \int (f) g^* / M^* (\iint g^{*2})^{1/2}$, причем функционал \int обращает заключенное в квадратные скобки выражение в нуль. Следовательно, умножив обе части последнего выражения на функцию g^* и проинтегрировав результат, получаем равенство вида $\iint g^* f = \int (f) \iint g^{*2} / M^* (\iint g^{*2})^{1/2}$, откуда следует, что $\int (f) = M^* \iint g^* f / (\iint g^{*2})^{1/2}$; положив, что $g = M^* g^* / (\iint g^{*2})^{1/2}$, полностью доказываем теорему.

Признаки, оценивающие согласованность предъявленного изображения с эталоном, особенно удобны в тех случаях, когда изображения поддаются нормализации. Очень многие системы распознавания символов предусматривают использование именно таких признаков (см. перечень соответствующих работ, помещенный в конце разд. 7.1). Если подлежащие классификации образы в значительной степени единообразны, как, например, фотографии расположения звезд или символы шрифта определенного типа, становится целесообразным использовать в качестве признаков эталоны, представляющие образы целиком. В этом случае опознаваемый образ можно считать совпадающим с эталоном, с которым он согласуется наилучшим образом. Представляет интерес и еще один класс линейных признаков, отражающий случай, когда эталоны являются элементарными функциями, например многочленами, синусоидами и т. д.

Используя в качестве эталонов множество ортогональных функций, можно сформировать множество взаимно независимых признаков. В частности, подобное множество признаков образуют коэффициенты Фурье-функции, представляющей изображение.

Если, например, функции представляют собой одночлены вида $x^i y^j$, то получаем моменты функции, представляющей предъявленное изображение [85—93]; естественно в этом случае проводить нормализацию, помещая начало координат в центр тяжести (моменты в результате становятся «центральными»).

Упражнение 3. Вычислить значения центральных моментов для использовавшихся в упражнении 1 шести изображений при всех i и j , для которых выполняется условие $i + j \leq 2$. Определить, какой из этих моментов является наилучшим признаком для различения предъявленных образов.

Признаки, линейные относительно заданного множества признаков [94]

Признак \mathcal{E} считается линейным относительно множества признаков \mathcal{S} , если $\mathcal{E} = \sum_{a \in \mathcal{S}} a \int$, где коэффициенты a — действительные числа. Аналогичным образом, если признак \mathcal{E} является двузначным, его называют линейным пороговым признаком относительно множества \mathcal{S} при условии, что существуют некоторые действительные числа a_{\int} и действительное число t , такие, что $\mathcal{E} = 1$ в том и только в том случае, когда выполняется неравенство $\sum_{a \in \mathcal{S}} a \int \geq t$. Продолжим наш анализ, исходя из следующего допущения: все функции, представляющие изображение, являются двузначными и определены на конечном множестве A . Двузначный признак \int называется маской, если в множестве A существует такая область S , что $\int (f) = 1$ в том и только том случае, когда $f = 1$ во всех точках области S . [Заметим, что это определение эквивалентно утверждению $\int (f) = \prod_{x \in S} f(x)$.]

Утверждение. Любой двузначный признак является линейным пороговым признаком относительно множества всех масок.

Доказательство. Рассмотрим двузначную функцию \mathcal{E} ; ее аргументы y_1, \dots, y_n — двузначные переменные (ими служат значения функции f в n точках множества A). Функцию \mathcal{E} можно представить с помощью «нормальной

формы», т. е. как дизъюнкцию конъюнкций, образованных членами вида y^i и \bar{y}_i ; так как для любой заданной функции f единице может быть равна только одна из конъюнкций, то их дизъюнкция является одновременно их суммой. Если во всех конъюнкциях заменить переменные \bar{y}_i на члены вида $1 - \bar{y}_i$, перемножить их и выполнить приведение подобных членов, то функция \mathcal{G} выражается через линейную комбинацию произведений переменных y с целочисленными коэффициентами.

Под порядком функции \mathcal{G} будем понимать то наименьшее число k , при котором существует множество функций \mathcal{G} , такое, что функция \mathcal{G} является относительно этого множества линейным пороговым признаком и опорное множество любой функции \mathcal{F} , принадлежащей множеству \mathcal{G} , содержит не менее чем k элементов. Например:

1. Признак \mathcal{G} имеет порядок 0 в том и только том случае, если он является константой.

2. Порядок любой маски равен 1. Действительно, при задании любой области S на множестве A маска, соответствующая S , равна 1 в том и только том случае, если выполняется неравенство $\sum_{x \in S} \mathcal{F}_x \geq |S|$ (число элементов множества A , покрываемых областью S), где \mathcal{F}_x — маска, определяемая множеством $\{x\}$ (следовательно, маска \mathcal{F}_x имеет опорное множество, состоящее из одного элемента).

3. Можно показать, что для функций \mathcal{G} и \mathcal{H} , имеющих порядок r и s соответственно, порядок признаков $\mathcal{G} + \mathcal{H} - 2\mathcal{G}\mathcal{H}$ (\mathcal{G} или \mathcal{H} , но не одновременно) и $1 - (\mathcal{G} - \mathcal{H})^2$ тогда и только тогда, когда \mathcal{G} или \mathcal{H} не превышают величины $r, +s$. Интересно, однако, что для любого числа n найдутся такие признаки \mathcal{G} и \mathcal{H} порядка 1; что порядок признаков $\mathcal{G}\mathcal{H}$ (\mathcal{G} и \mathcal{H}) и $\mathcal{G} + \mathcal{H} - \mathcal{G}\mathcal{H}$ (\mathcal{G} или \mathcal{H}) больше n . Рассмотрим еще несколько интересных признаков \mathcal{G} различных порядков, характеризующих особенности множества S_f (набор элементов, на которых $f = 1$):

4. Признак «множество $|S_f|$ имеет нечетное число элементов» имеет порядок $|A|$.

5. Порядок признака « S_f — связное множество» возрастает с увеличением $|A|$.

6. Порядок признака «род множества S_f меньше m » не превышает 4 при любом m (определение понятия «род» и доказательство приведены в разд. 9.1).

7. Порядок признака « S_f — выпуклое множество» не превышает 3 (определение понятия «выпуклости» и доказательство приведены в разд. 8.4).

8. Из всех инвариантных относительно переноса признаков порядок, равный 1, имеют только признаки типа « $|S_f| < k$ » и « $|S_f| > k$ »¹.

Можно, кроме того, ввести признаки, являющиеся линейными пороговыми функциями относительно множества признаков с опорным множеством, ограниченным по диаметру (в отличие от ограничения на число элементов опорного множества). Легко убедиться в том, что признаку « $|S_f| \leq k$ » соответствует диаметр, равный 1; нетрудно, однако, показать, что в случае признака « $|S_f| = k$ » диаметр возрастает по мере увеличения диаметра множества A подобно тому, как это имеет место для признака « S_f — связное множество».

Может оказаться так, что, хотя признак и представляет собой линейную пороговую функцию по отношению к заданному множеству признаков, линейность эта практически нереализуема из-за недопустимо большого диапазона изменения коэффициентов. Можно, в частности, показать, что при представлении признака «множество $|S_f|$ имеет нечетное число элементов», линейной пороговой функцией, определенной на множестве всех масок, отношение наибольшего коэффициента к наименьшему экспоненциально возрастает с увеличением мощности множества $|A|$.

«Случайные» признаки

Во многих системах распознавания образов используются признаки, полученные с помощью случайной процедуры. Можно, например, определить линейный

¹ Признаки 4—8 являются в сущности предикатами, заданными на подобласти S_f области A и выражающими некоторое свойство выделенной части образа. Именно так, кстати, трактуются эти признаки в монографии: Минский М. и Пейперт С., Перцептроны, изд-во «Мир», М., 1971. — Прим. перев.

«случайный» признак, воспроизводя случайную функцию и используя ее в качестве эталона. В «простейшем персептроне» [95] каждый из «А-элементов» вычисляет значение одного из признаков такого типа (каждый элемент вычисляет «свой» признак), причем представляющие эталон случайные функции могут принимать значения из набора 0, 1, -1. Кроме того, «R-элементы» вычисляют значения признаков, являющихся линейными пороговыми по отношению к множеству признаков, «выдаваемых» отдельными А-элементами [96—99].

Упражнение 4. Бросая монету, сформировать несколько двузначных «случайных» эталонов «3 на 3»; вычислить степень согласованности полученных эталонов с шестью образцами, использовавшимися в упражнении 1.

Признаки не обязательно должны быть полностью случайными, они могут быть сформированы с учетом различных ограничений. Можно, например, воспользоваться «случайными эталонами», имеющими тенденцию к «группировке» [99, 100], или просто бессмысленной «мазней», составленной из штрихов [101—103]. Можно также создавать случайные функции с помощью процедуры «без возвращения», что обеспечивает отсутствие пересечения между опорными множествами всех функций [104, 105].

7.4. Признаки-функции

В некоторых случаях, как уже упоминалось в этой главе, целесообразно оперировать признаками изображения, являющимися действительными функциями одной переменной, а не действительными числами. Можно было бы, конечно, рассматривать признаки — функции двух (или даже большего числа!) переменных, однако обычно это не имеет смысла, так как оперировать такими объектами по меньшей мере так же сложно, как и исходными изображениями. С другой стороны, следует отметить, что признак, являющийся просто числом, может содержать столько же информации, как и исходное изображение. Так, например, двоичное изображение размером $n \times n$ полностью определяется одним числом, состоящим из

n^2 двоичных разрядов; аналогичным образом обстоит дело для многоградационных изображений. Значения признаков определяются для того, чтобы упростить классификацию изображения, но не потому, что они обязательно обеспечивают уменьшение объема информации. В данном разделе рассмотрим несколько интересных классов признаков-функций.

Частотное распределение

Как отмечалось в разд. 7.3, для любого локального признака можно определить частотное распределение значений, принимаемых им на некотором изображении. Статистические характеристики таких распределений целесообразно использовать в качестве структурных признаков.

Упражнение 1. Вычислить частотное распределение для значений градиента и лапласиана на случайном изображении.

Поперечные сечения

Во многих случаях для решения задачи распознавания достаточно информации, выделенной в результате анализа не всего изображения, а лишь его нескольких «срезов». Такой подход особенно практичен, если обрабатываемое изображение предварительно подвергнуто хоть какой-нибудь нормализации. При распознавании символов, например, разрезы можно делать на разных уровнях вдоль напечатанных или написанных строк [106—108]; результатом этих операций является набор функций вида $f(x, y_0)$.

Поперечные разрезы, ориентированные каким-либо другим способом, также могут оказаться полезными [109]. Аналогичным образом можно делать поперечные разрезы автокорреляционных функций изображения. Рассмотрим в качестве примера сечение автокорреляционной функции $g = f \otimes f$, выполненное вдоль окружности с центром в начале координат, т. е. функцию вида $g(r_0, \theta)$. Этот разрез позволяет получить значения автокорреляцион-

ной функции $f \otimes f$ для последовательности таких сдвигов, когда одна из функций движется поступательно (не вращаясь) по орбите радиуса r_0 вокруг начала координат [110]. Обобщение этой методики на случай автокорреляционных функций высших порядков рассматривается в работе [111]; в работе [112] предлагается метод, практически эквивалентный воспроизведению функции $g(r_0, \theta)$ при малых значениях радиуса r_0 .

Упражнение 2. Построить график автокорреляционной функции $f \otimes f$ при смещениях вдоль квадрата «единичного радиуса» с центром в начале координат для каждого из шести изображений, использованных в упражнении 1 разд. 6.1. Содержат ли эти графики достаточно информации для различения всех образов?

Функции, определенные на семействах подмножеств

Рассмотрим однопараметрическое семейство \mathcal{A} подмножеств A_λ , определенных на плоскости, т. е. параллельные прямые, прямые, проведенные через заданную точку, концентрические окружности и т. д. Если найти значение некоторого признака изображения на множестве A_λ , то можно считать его функцией параметра λ . Определяя, например, значение линейного признака $\iint gf$ в области A_λ , можно выполнять интегрирование не по всей плоскости, а только по множеству A_λ . Если объект A — некоторая кривая (во всех анализируемых ниже случаях дело обстоит именно так), то этот интеграл обращается в криволинейный интеграл вида $\int_{A_\lambda} gf$. В нашем рассмотрении примем, что функция g — константа, равная единице.

1. Если \mathcal{A} представлено семейством параллельных линий, например семейством вида $x = \lambda$, то интегрирование можно рассматривать как «стягивание» функции f к оси x . Для выполнения этой операции можно воспользоваться следующим несложным способом, если тонкую вертикальную щель перемещать в горизонтальном направлении поперек изображения, то общее количество света, проходящего через щель в данной позиции, приблизительно пропорционально интегралу функции, представляющей изображение, взятому вдоль оси щели [113—115]

Упражнение 3. «Стянуть» все шесть изображений упражнения 2 к оси x . Содержат ли полученные результаты количество информации, достаточное для различения всех изображений? Подвергнуть такой же операции автокорреляционные функции изображений. (Некоторые исследователи считают, что процесс стягивания изображения к осям x и y имеет определяющее значение для различения осьминогом зрительных образов [116].)

2. Рассмотрим семейство \mathcal{A} линий, имеющих общую точку; в полярных координатах это семейство представляется уравнением вида $\theta = \lambda$. Результат можно использовать для обнаружения в исходном изображении радиальных контуров или линий, так как их наличие приводит к появлению больших значений производной функции, определяющей признак. Значения этой функции можно вычислить приблизительно, используя поворачивающуюся щель или, что равнозначно, развертывающее устройство с бегущим пятном, причем это пятно имеет форму длинного штриха, который можно ориентировать в произвольном направлении [117, 118].

3. Рассмотрим семейство \mathcal{A} концентрических окружностей, т. е. семейство, определяемое в полярных координатах уравнением $r = \lambda$. Простой способ получения такой функции состоит в следующем: следует поворачивать изображение (например, с помощью вращающихся зеркал) и интегрировать (по времени) количество света, проходящего через фиксированную радиальную щель в плоскости изображения [119]. Используя отражение исходного изображения в вогнутом коническом зеркале, получаем аналогичные результаты.

В данном случае, так же как и при задании признака с помощью поперечного сечения, можно определить семейство криволинейных интегралов не на исходном изображении, а на его модификации, полученной в результате предварительной обработки первого изображения (очевидно, что в более общем смысле речь идет о функции, определяемой на семействе подмножеств). В работах [120, 121] обсуждается использование автокорреляционной функции $f \otimes f$ вместо изображения f для случая семейства окружностей вида $r = \lambda$. Если изображение является двузначным с набором масок S_r , этой функции можно при-

дать значение вероятности включения в подобласть S_i одного из концов случайно ориентированного отрезка длины λ , второй конец которого также принадлежит этой подобласти.

Упражнение 4. Построить графики сумм функций f и $f \otimes f$ на квадратах «радиусов» 0, 1, 2 с центром в начале координат для всех шести изображений упражнения 2.

ЛИТЕРАТУРА

- Смотрите работы, помещенные в сборнике OCR, а также первые десять статей сборника PR², в частности Rabinow J., The present state of the art of reading machines, p. 3.
1. National Science Foundation, «Current Research and Development in Scientific Documentation», № 14, NSF-66-17 (1966), U. S. Government Printing Office, Washington, D.C.
 2. Stevens M. E., Automatic character recognition — a state-of-the-art report, NBS TN112 (PB 161613) (May 1961).
 3. «Читающие устройства», сб. докладов на конференции по обработке информации, машинному переводу и автоматическому чтению текста; АН СССР, Институт научной информации, М., 1962.
 4. Freiburger H., Murphy E. F., Reading devices for the blind: an overview, Human Factors in Technology (Bennett E. M., Degani J., Spiegel J., eds.), McGraw-Hill, New York, 1963, p. 299; IRE Trans. Human Factors in Electronics, HFE-2, 8—19 (March 1961); Freiburger H., Murphy E. F., Reading machines for the blind, Science, 152, 679 (Apr. 29, 1966).
 5. Bijleveld W. J., Automatic Reading of Digits, Netherlands Automatic Inform. Process. Res. Ctr., Amsterdam, 1963.
 6. Lindgren N., Machine recognition of human language, Part III, Cursive script recognition, IEEE Spectrum, 2 (5), 104 (May 1965); русский перевод: Линдгрэн Н., Машинное распознавание человеческого языка, ч. 3, Зарубежная радиоэлектроника, № 5 (1966).
Эта работа в большой степени основана на использовании метода обработки информации, представляемой в естественном масштабе времени (последовательности тактов, скоростей, ускорений и т. д.), а так как это не является «обработкой изображений», данный метод в книге не рассматривается.
 7. Feidelman L. A., A survey of the character recognition field, Datamation, 12(2), 45 (Febr. 1966).
 8. Wilson R. A., Optical Page Reading Devices, Reinhold, New York, 1966; русский перевод: Уилсон Р. А., Оптические читающие устройства, изд-во «Мир», М., 1969.

9. Document Handling and Character Recognition Committee (1966), «Character Recognition», Brit. Comput. Soc., London, 1967.
10. Сб. «Читающие автоматы и распознавание образов», отв. ред. Ковалевский В. А., изд-во «Наукова думка», Киев, 1965.
11. Welford W. T., Bubble chamber optics, Appl. Opt., 2, 981 (Oct. 1963).
12. Gelernter H., Data collection and reduction for nuclear particle trace detectors, Advances Comput., 6, 229 (1965).
13. Purdue Conference on Instrumentation for High-Energy Physics, IEEE Trans. Nuclear Science, NS-12 (Aug. 1965).
14. Methods in Computational Physics (Alder B., Fernbach S., Rotenberg M., eds.), 5, Academic Press, New York, 1966.
15. Colwell R. N., The extraction of data aerial photographs by human and mechanical means, Photogrammetria, 20, 211 (1965).
16. Smillie S. F., Automatic target recognition: some considerations, IEEE Trans. Aerospace Electronic Systems, AES-2, 187 (March 1966).
17. Law Enforcement Science and Technology (Yefsky S. A., ed.), Thompson, Washington, D.C., 1967, p. 445. В сборнике помещены десять статей, посвященных автоматической классификации и распознаванию отпечатков пальцев.
18. Toiles W. E., ed., Data extraction and processing of optical images in the medical and biological sciences, Ann. N.Y. Acad. Sci., 157, 1 (March 1969).
19. Файн В. С., Автоматическое опознавание образов, Вестник АН СССР, № 10, 127 (1965).
20. Methodologies of Pattern Recognition (Watanabe M. S., ed.), Academic Press, New York, 1969.
21. Recognizing Patterns: Studies in Living and Automatic Systems (Kollers P. A., Eden M., eds.), M.I.T. Press, Cambridge, 1968; русский перевод: Колерс П., Иден И., ред., Распознавание образов. Исследование живых и автоматических распознающих систем, изд-во «Мир», М., 1970.
22. Nagy G., State of the art in pattern recognition, Proc. IEEE, 56, 836 (May 1968); русский перевод: Надь, Распознавание образов. Обзор, ТИИЭР, 56, № 5, 57 (1968).
23. Lewis P. M., II, The characteristic selection problem in recognition systems, IRE Trans. Information Theory, IT-8, 171 (Febr. 1962).
24. Marill T., Green D. M., On the effectiveness of receptors in recognition systems, IRE Trans. Information Theory, IT-9, 11 (Jan. 1963).
25. Tou J. T., Heydorn R. P., Some approaches to optimum feature extraction, Computer and Information Sciences, II (Tou J. T., ed.), Academic Press, New York, 1967, p. 57.
26. Watanabe M. S. et al., Evaluation and selection of variables in pattern recognition, Computer and Information Science, II, Academic Press, New York, 1967, p. 91; русский перевод:

- Ватанабе С., Ламберт П., Куликовский К., Бакстон Дж., Уокер Р., Оценка и отбор параметров в задачах распознавания образов, в сб. «Автоматический анализ сложных изображений, изд-во «Мир», М., 1969, стр. 276.
27. Swonger C. W., Property learning in pattern recognition systems using information content measurements, *PR²*, p. 329.
 28. Chien Y. T., Fu K. S., Selection and ordering of feature observations in a pattern recognition system, *Information and Control*, 12, 395 (May—June 1968).
 29. Greanias E. C. et al., Design of logic for recognition of printed characters by simulation, *IBM J. Res. Develop.*, 1, 8 (Jan. 1957).
 30. Evey R. J., Use of computer to design character recognition logic, *EJCC*, 205 (Dec. 1959).
 31. Stearns S. D., A method for the design of pattern recognition logic, *IRE Trans. Electronic Computers*, EC-9, 48 (March 1960);
 32. Freeman D. N., Computer synthesis of character recognition systems, *IRE Trans. Electronic Computers*, EC-10, 735 (Dec. 1961).
 33. Kamenskyy L. A., Liu C. N., Computer-automated design of multifont print recognition logic, *IBM J. Res. Develop.*, 7, 2 (Jan. 1963); русский перевод: Каментский Л. Ю., Построение с помощью ЦВМ логики для распознавания печатных шрифтов, *Зарубежная радиоэлектроника*, № 9 (1963).
 34. Liu C. N., A programmed algorithm for designing multifont character recognition logics, *IEEE Trans. Electronic Computers*, EC-13, 586 (Oct. 1964).
 35. Liu C. N., Shelton G. L., Jr., An experimental investigation of a mixed-font print recognition system, *IEEE Trans. Electronic Computers*, EC-15, 916 (Dec. 1966).
 36. Uhr L., Feature discovery and pattern description, *PR²*, p. 159.
 37. Widrow B., Generalization and information storage in networks of ADALINE «neurons»; *Self-Organizing Systems* (Yovits M. C., Jacobi G. T., Goldstein G. D., eds.), 1962, Spartan, New York, 1962, p. 435.
 38. Uhr L., Vossler C., A pattern-recognition program that generates, evaluates and adjusts its own operators, *WJCC*, p. 555 (May 1961); см. также *PR*, p. 349; русский перевод: Юр Л., Фосслер Ч., Программа распознавания образов, которая вырабатывает, оценивает и улучшает свои операторы, в сб. «Вычислительные машины и мышление», изд-во «Мир», 1967, стр. 255.
 39. Joseph R. D., Kelly P. M., Viglione S. S., An optical decision filter, *Proc. IEEE*, 51, 1098 (Aug. 1963); русский перевод: Джозеф, Келли, Вильоне. Оптический решающий фильтр, *ТИИЭР*, № 8, 1098 (1963).

40. Joseph R. D., Viglione S. S. A pattern recognition technique and its application to high resolution imagery, *SJCC*, p. 457 (Apr. 1966).
41. Uhr L., «Pattern recognition» computers as models for form perception, *Psychol. Bull.*, 60, 40 (Jan. 1963).
42. Sutherland N. S., The Methods and Findings of Experiments on the Visual Discrimination of Shape by Animals, *Exptl. Psychol. Soc. Monograph 1*. Heffer's, Cambridge, 1961.
43. Block H. D., Nilsson N. J., Duda R. O., Determination and detection of features in patterns; *Computer and Information Sciences* (Tou J. T., Wilcox R. H., eds.), Spartan, New York, 1964, p. 75.
44. Grenander U., Toward a theory of patterns, *Symp. Probability Methods Analysis*, Springer, Berlin, 1967, p. 79.
45. Ullman J. R., A simplification of the problem of choosing features, *CPR*, p. 197.
46. Richardson J. M., A rational approach to semi-adaptive pattern recognition, *CPR*, p. 220.
47. Meyer R. F., Giuliano V. E., Jones P. E., Analytic approximation and translational invariance in character recognition, *OCR*, p. 181.
48. Horwitz L. P., Shelton G. L., Jr., Pattern recognition using autocorrelation, *Proc. IRE*, 49, 175 (Jan. 1961); русский перевод: Горвиц, Шелтон, Распознавание образов с помощью автокорреляции, *ТИИЭР*, № 1, 190 (1961).
49. Armitage J. D., Lohmann A. W., Character recognition by incoherent spatial filtering, *Appl. Opt.*, 4, 461 (Apr. 1965); см. также: дискуссия, *Appl. Opt.*, 4, 1666 (Dec. 1965).
50. Kain R. Y., Autocorrelation pattern recognition, *Proc. IRE*, 49, 1085 (June 1961); русский перевод: Кайн, Распознавание знаков с помощью автокорреляции, *ТИИЭР*, № 6, 1268 (1961).
51. Reichardt W., Autocorrelation, a principle for the evaluation of sensory information by the central nervous system, *Sensory Communication* (Rosenblith W., ed.), M.I.T. Press, Cambridge, 1961, p. 303.
52. Bliss J. C., Visual information processing in the beetle *Lixus*, *OPI*, p. 124.
53. Adler R. L., Konheim A. G., A note on translation invariants, *Proc. Amer. Math. Soc.*, 13, 425 (June 1962).
54. McLaughlin J. A., Raviv J., Nth-order autocorrelations in pattern recognition, *Information and Control*, 12, 121 (Febr. 1968).
55. Doyle W., Operations useful for similarity-invariant pattern recognition, *J. ACM*, 9, 259 (Apr. 1962).
56. Brousil J. K., Smith D. R., A threshold logic network for shape invariance, *IEEE Trans. Electronic Computers*, EC-16, 818 (Dec. 1967).
57. Pitts W., McCulloch W. S., How we know universals — the perception of auditory and visual forms, *Bull. Math. Biophys.*, 9, 127 (1947).

58. Маскау D. M., Some experiments on the perception of pattern modulated at the alpha frequency, *Electroencephalog. Clin. Neurophysiol.*, 5, 559 (1953).
59. Rosenblatt F., Perceptual generalization over transformation groups; Self-Organizing Systems (Yovits M. C., Samuelson S., eds.), Pergamon Press, New York, 1960, p. 63, русский перевод: Розенблат Ф., Обобщение восприятия по группам преобразований, «Кибернетический сборник», № 4, ИЛ, 1962, стр. 150.
60. Singer J. R., Electronic analog of the human recognition system, *J. Opt. Soc. Amer.*, 51, 61 (Jan. 1961).
61. Singer J. R., A self organizing recognition system, *WJCC*, 545 (May 1961).
62. Кабрискы M., Proposed Model for Visual Information Processing in the Human Brain, Univ. of Illinois Press, Chicago, Illinois, 1966.
63. Hoffman W. C., The Lie algebra of visual perception, *J. Math. Psychol.*, 3, 65 (1966); *J. Math. Psychol.*, 4, 348 (1967).
64. Гривазкы A., Determination of redundancies in a set of patterns, *IRE Trans. Information Theory*, IT-2, 151 (Dec. 1956).
65. Gill A., Minimum-scan pattern recognition, *IRE Trans. Information Theory*, IT-5, 52 (June 1959).
66. Mayo H. B., Optimal classification of objects, Algorithm 83, *Comm. ACM*, 5, 167 (March 1963).
67. Bomba J. S., Alpha-numeric character recognition using local operations, *EJCC*, 218 (Dec. 1959).
68. Каментскы L. A., The simulation of three machines which read rows of handwritten arabic numbers, *IRE Trans. Electronic Computers*, EC-10, 489 (Sept. 1961).
69. Fischler M., Mattson R. L., Firschein O., Healy L. D., An approach to general pattern recognition, *IRE Trans. Information Theory*, IT-8, S64—S73 (Sept. 1962).
70. Lendaris G. G., Stanley G. L., An optical self-organizing recognition system, OEOIP, p. 535; русский перевод: Лендерис, Стэнли, Оптико-логическая самообучающаяся система для распознавания образов, *Зарубежная радиоэлектроника*, № 2 (1968).
71. ASENDORF R. H., The remote reconnaissance of extraterrestrial environments, PPR, p. 223.
72. CLAYDON D. O., CLOWES M. B., PARKS J. R., Letter recognition and segmentation of running text, *Information and Control*, 9, 246 (June 1966).
73. ROSENFELD A., Automatic recognition of basic terrain types from aerial photographs, *Photogrammetric Engrg.*, 28, 115 (March 1962).
74. ROSENFELD A., GOLDSTEIN A., Optical correlation for terrain type discrimination, *Photogrammetric Engrg.*, 30, 639 (June 1964).
75. STEINER D., HAEFNER H., Tone distortion for automated interpretation, *Photogrammetric Engrg.*, 31, 269 (March 1965).
76. HAWKINS J. K. et al., Automatic shape detection for prog-

- rammed terrain classification; *Filmed Data and Computers*, Paper № XVI. Soc. Photo-Opt. Instr. Engrs., Redondo Beach, California, June 1966.
77. HAWKINS J. K., ELERDING G. T., Image feature extraction for automatic terrain classification; *Computerized Imaging Techniques*, Paper № VI. Soc. Photo-Opt. Instr. Engrs., Redondo Beach, California, June 1967.
78. DARLING E. M., JR., JOSEPH R. D., Pattern recognition from satellite altitudes, *IEEE Trans. Systems Science and Cybernetics*, SSC-4, 38 (March 1968).
79. DARLING E. M., JR., JOSEPH R. D., An experimental investigation of video pattern recognition, PPR, p. 457. Автоматическая система распознавания цели, описанная в работе SWOBODA W., GERDES J. W., A system for demonstrating the effect of changing background on automatic target recognition, PPR, p. 33, также предусматривает использование набора структурных признаков.
80. PREWITT J. M. S., MENDELSON M. L., The analysis of cell images, *Ann. New York Acad. Sci.*, 128, 1035 (Jan. 1966).
81. GIBSON J. J., *The Perception of the Visual World*, Houghton, Boston, 1950.
82. CAREL W., PURDY W., LUBOW R., The visilog: a bionic approach to visual space perception and orientation, Proc. Nat. Aerospace Electron. Confer., May, 1961, p. 295.
83. JULESZ B., Visual pattern discrimination, *IRE Trans. Information Theory*, IT-8, 84 (Febr. 1962).
84. PICKETT R. M., The perception of a visual texture, *J. Exptl. Psychol.*, 68, 13 (1964).
Дополнительные сведения по поводу работ в области зрительного восприятия структуры образа можно найти в работах: JULESZ B., Some recent studies in vision relevant to form perception, MPSVF, p. 136; ROSENFELD A., On models for the perception of visual texture, MPSVF, p. 219; PICKETT R. M., The perception of random visual texture, MPSVF, p. 224.
85. SHIMBEL A., A logical program for the simulation of visual pattern recognition; *Principles of Self-Organization* (H. von Foerster and G. W. Zopf, Jr., eds.), Pergamon Press, New York, 1962, p. 521; русский перевод: Шимбел А., «Логическая программа для моделирования распознавания зрительных образов» в сб. «Принципы самоорганизации», изд-во «Мир», М., 1966, стр. 611.
86. HU M.-K., Pattern recognition by moment invariants, *Proc. IRE*, 49, 1428 (Sept. 1961); русский перевод: Ху Мин-Куэй, Опознавание фигур при помощи инвариантных соотношений между моментами, *ТНИЭР*, № 9 (1961).
87. HU M.-K., A mathematical model for visual perception, BPSS, p. 222; русский перевод: Ху Мин-Куэй, Математическая модель зрительного восприятия, в сб. «Проблемы бионики», изд-во «Мир», М., 1965, стр. 308.

88. Hu M.-K., Visual pattern recognition by moment invariants, *IRE Trans. Information Theory*, IT-8, 179 (Febr. 1962).
89. Giuliano V. E. et al. A Gestalt method of automatic pattern recognition, Proc. 3rd Intern. Congr. Cybernetics (Sept. 1961), p. 370, Association Internationale Cybernetique, Namur, Belgium, 1965.
90. Giuliano V. E., Jones P. E., Kimball G. E., Meyer R. F., Stein B. A., Automatic pattern recognition by a Gestalt method, *Information and Control*, 4, 332 (Dec. 1961).
91. Alt F. L., Digital pattern recognition by moments, OCR, p. 153; см. также *J. ACM*, 9, 240 (Apr. 1962).
92. Butler J. W., Butler M. K., Stroud A., Automatic analysis of chromosomes, I, Data Acquisition and Processing in Biology and Medicine, 3, p. 261, Pergamon Press, New York, 1963; Automatic analysis of chromosomes, II, там же, 4, p. 47, Pergamon Press, New York, 1964.
93. Moskowitz S., Terminal guidance by recognition — a new approach, *IEEE Trans. Aerospace Navigational Electronics*, ANE-11, 254 (Dec. 1964); русский перевод: Наведение на конечном участке траектории с помощью автоматического распознавания образов, *Экспресс-информация*, сер. «Техническая кибернетика», № 25, 1, 1965.
94. Minsky M. L., Papert S., Linearly unrecognizable patterns, Proc. Symp. Appl. Math., 19, 176, Amer. Math. Soc., Providence, Rhode Island, 1967; см. также Minsky M. L., Papert S., Perceptions, an Introduction to Computational Geometry, M.I.T. Press, Cambridge, 1969; русский перевод: Минский М., Пейперт С., Перцептроны, изд-во «Мир», М., 1971.
95. Rosenblatt F., Principles of Neurodynamics, Spartan, New York, 1962; русский перевод: Розенблатт Ф., Принципы нейродинамики. Перцептроны и теория механизмов мозга, изд-во «Мир», М., 1965.
96. Murray A. E., Perceptron application in photo interpretation, *Photogrammetric Engrg.*, 27, 627 (Sept. 1961).
97. Holmes W. S., Leland H. R., Muerle J. L., Recognition of mixed-font imperfect characters, OCR, p. 213.
98. Borsellino A., Gamba A., An outline of a mathematical theory of PAPA, *Nuovo Cimento, Suppl.*, 20, 221 (1961).
99. Roberts L. G., Pattern recognition with an adaptive network, *IRE Intern. Conv. Record*, Part 2, 66 (March 1960); см. также PR, p. 295; русский перевод: Робертс Л. Дж., Распознавание изображений при помощи приспособляющихся систем, «Кибернетический сборник», № 4, ИЛ, 212, 1962.
100. Bakis R., Herbst N. M., Nagy G., An experimental study of machine recognition of hand-printed numerals, *IEEE Trans. Systems Science and Cybernetics*, SSC-4, 119 (July 1968).
101. Hoffman A., The «whirling dervish», a simulation study in learning and recognition systems, *IRE Intern. Conv. Record Part 4*, 153 (March 1962).

102. Bryan J. S., Experiments in adaptive pattern recognition, *IEEE Trans. Military Electronics*, MIL-7, 174 (April-July 1963).
103. Rosenblatt F., A comparison of several Perceptron models; Self-Organizing Systems (Yovits M. C., Jacobig T., Goldstein G. D., eds.), Spartan, New York, 1962, p. 463.
104. Bledsoe W. W., Browning I., Pattern recognition and reading by machine, *EJCC*, 225 (Dec. 1959), см. также PR, p. 301.
105. Steck G. P., Stochastic model for the Browning-Bledsoe pattern recognition scheme, *IRE Trans. Electronic Computers*, EC-11, 274 (Apr. 1962).
106. Diamond T. L., Devices for reading handwritten characters, *EJCC*, 232 (Dec. 1957).
107. Harmon J. D., Handwriting reader recognizes whole words, *Electronics*, 35, 29 (Aug. 24, 1962).
108. Earnest L. D., Machine recognition of cursive writing; Information Processing 1962, North-Holland, Amsterdam, 1963, p. 462.
109. Weeks R. W., Rotating raster character recognition system, *Trans. AIEE*, 80, Part I, 353 (Sept. 1961).
110. Clowes M. B., Parks J. R., A new technique in automatic character recognition, *Comput. J.*, 4, 121 (July 1961).
111. Clowes M. B., The use of multiple auto-correlation in character recognition, OCR, p. 305.
112. Buell D. N., Chrysler optical processing scanner, *EJCC*, 353 (Dec. 1961).
113. Heasley C. C., Jr., Some communication aspects of character-sensing systems, *WJCC*, 176 (March 1959).
114. Dickinson W. E., A character-recognition study, *IBM J. Res. Develop.*, 4, 335 (July 1960).
115. Booth W. T., Miller G. M., Schleich O. A., Design considerations for stylized-font character readers, OCR, p. 115.
116. Young J. Z., A Model of the Brain, Oxford Univ. Press, London and New York, 1964.
117. Innes D. J., FILTER — a topological pattern separation computer program, *EJCC*, 25 (Dec. 1960).
118. Watts T. L., Scanning and measuring photographs of bubble chamber tracks using a computer controlled line segment («PEPR»), PRP, p. 207.
119. Uffelman M. R., Target detection, prenormalization, and learning machines, PPR, p. 503.
120. Tenery G., A pattern recognition function of integral geometry, *IEE Trans. Military Electronics*, MIL-7, 196 (April-July 1963).
121. Tenery G. R., Information flow in a Bionics image recognition system, MPSVF, p. 403.

Разбиение изображений

Часто при решении задач на распознавание зрительных образов не удается ограничиться отнесением предъявленного изображения к одному из известных классов. Необходимо получать *описание* изображения, причем число возможных описаний столь велико, что бессмысленно считать каждое определением отдельного класса. Обычно описание представляет отдельные подмножества изображения («объекты»), характеризуя их свойства. Чтобы получать описание такого рода, автоматическая система распознавания должна уметь выделять на изображении соответствующие подмножества (выполнять «разбиение» изображения). Общий метод разбиения изображения на части неизвестен; в качестве объектов в зависимости от требуемого типа описания можно выбирать подмножества самых различных видов. В разд. 8.1 будут рассмотрены методы разбиения произвольных изображений, в остальных разделах настоящей главы обсуждаются методы, позволяющие определять новые подмножества, используя уже выделенные подмножества.

8.1. Разбиение на части изображения произвольного вида

Выделение на изображении некоего подмножества равнозначно заданию его «характеристической функции», т. е. функции, равной единице в точках подмножества и нулю в остальных точках. Общепринятый метод выделения на исходном изображении некоего подмножества сводится к отысканию его характеристической функции с помощью *квантования на два уровня* исходного изображения или (в более общем случае) изображения, получен-

ного из исходного посредством соответствующей *предварительной обработки*. В частности, для любого изображения f , операции преобразования изображений Φ и действительного числа t в качестве искомого подмножества можно выбрать точечное множество $S_{\Phi, t}$, такое, что во всех образующих его точках яркость преобразованного изображения $\Phi(f)$ не меньше величины t . Можно выбирать подмножества и по-другому, например фиксировать конфигурации, на которых после применения операции преобразования выполняются соотношения $\Phi(f) \leq t$, $\Phi(f) < t$ или $\Phi(f) > t$; можно вводить два порога квантования, связанных условием $t_1 \leq t_2$, и выбирать подмножество, для которого справедливо неравенство $t_1 \leq \Phi(f) \leq t_2$; возможен и еще более общий подход — выбор подмножества, на котором после преобразования величина $\Phi(f)$ принимает значения из заданного набора. Эти приемы в сущности не более универсальны, чем простое квантование изображения $\Phi(f)$, так как, применяя операцию Φ совместно с подходящим точечным преобразованием, можно модифицировать шкалу яркостей настолько, что яркости в точках, покрываемых выбранным множеством, принимают значения, не меньшие t , не превышая эту же величину во всех остальных точках изображения.

1. *Квантование исходного изображения*. Подмножества можно выделять непосредственно на исходном изображении; для этого квантованию нужно подвергнуть яркость изображения f . Другими словами, операцию Φ можно считать тождественным преобразованием. Такой подход во многих случаях представляется наиболее естественным; например, при автоматическом распознавании символов последние являются черными, а страница — белым объектами; при распознавании типов облачного покрова облака в большинстве случаев белее земной поверхности. Когда из исходного изображения облачного покрова с 32 градациями яркости (фиг. 3.1, а) было получено двоичное изображение, мы, по существу, выбрали к качеству подмножества облака (фиг. 3.1, б). Если изображение представлено более или менее непрерывной функцией f , то при рассмотрении множества точек, интенсивность которых лежит между двумя близкими порогами, довольно часто «образуются» *точечно- или линейноподобные под-*

множества (фиг. 8.1). Конфигурации такого рода иногда называют «изофотами».

2. *Сочетание сглаживания и квантования.* Операцией Φ можно воспользоваться для того, чтобы профильтровать или очистить изображение, устранив, например, помехи, и получить в результате квантования такое подмножество, которое содержит немного точек, представляющих шум (операции сглаживания изображений рассматриваются в разд. 6.4). Для иллюстрации возможности применения этого подхода рассмотрим картину облачного покрова, содержащую участки густой и разреженной облачности, а также участки чистого неба. Если для сглаживания изображения воспользоваться усреднением и выбрать очень высокий (низкий) порог квантования, то облачный (безоблачный) участок неба можно выделить в качестве одного из подмножеств, а остальную часть изображения (редкую облачность) — в качестве второго [1]. Усовершенствованный вариант этого метода, использованный для решения другой задачи, можно найти в работе [2].

3. *Сочетание увеличения резкости и квантования.* В качестве операции Φ можно выбрать такое преобразование, при котором искомая функция принимает большие значения на границах и малые во всех остальных случаях. (Операции такого рода, обеспечивающие увеличение резкости, рассматриваются в разд. 6.5.) При квантовании $\Phi(f)$ в этом случае обнаруживается тенденция к выделению точечно- или линейноподобных подмножеств. Этот метод пригоден для выделения подмножеств, яркость которых не укладывается в фиксированный диапазон, но отличается от яркости их ближайшего окружения. Метод позволяет определить контуры таких подмножеств. Нельзя, однако, гарантировать, что эти контуры всегда будут замкнутыми кривыми: если в какой-то точке границы переход между подмножеством и окружающим фоном оказывается размытым, не всегда возможно избежать разрыва контура в этой точке. С другой стороны, непосредственное квантование яркости изображения позволяет выделять области с надпороговыми значениями яркости независимо от степени размытия ее границ.

4. *Использование согласованного фильтра в сочетании с квантованием.* Пусть операция Φ описывает прохождение

изображения через согласованный фильтр, например в виде нормированной функции взаимной корреляции $\iint fg / (\iint f^2 \iint g^2)^{1/2}$, представляющей изображение f , и



Ф и г. 8.1. Изофоты, полученные в результате квантования на два уровня картины облачного покрова.

Исходное изображение, квантованное на 32 уровня, приведено на фиг. 3.1, а. Уровни яркости с 16-го по 19-й заменены белым, остальные уровни — черным.

эталонного образа g (разд. 6.1). В этом случае квантование $\Phi(f)$ позволит выделить те точки, в которых изображение f хорошо согласуется с эталонным образом g . Отметим, что в качестве подмножества выделяется не образ g , а лишь точка или небольшая область, соответствующие множеству переносов, совмещающих изображение f с эталоном g .

Кардинальная проблема при работе с любыми методами разбиения, основанными на квантовании изображения, — выбор величины порога. Во многих случаях

нельзя сделать этот выбор заранее, так как одно и то же значение порога не для всех изображений оказывается наилучшим. Иногда, однако, удается автоматически устанавливать соответствующий порог для каждого изображения f на основе оценки распределения частот появления значений яркости изображения $\Phi(f)$. Рассмотрим два примера реализации такого подхода.

Таблица 1

Уровень яркости	Число событий	Уровень яркости	Число событий
0	406	16	1406
1	249	17	1424
2	293	18	1346
3	439	19	1342
4	784	20	1427
5	1421	21	1395
6	2014	22	1603
7	1651	23	1925
8	1667	24	2116
9	1923	25	3065
10	2105	26	4361
11	1992	27	3673
12	2033	28	2135
13	2248	29	1047
14	2051	30	443
15	1566	31	434

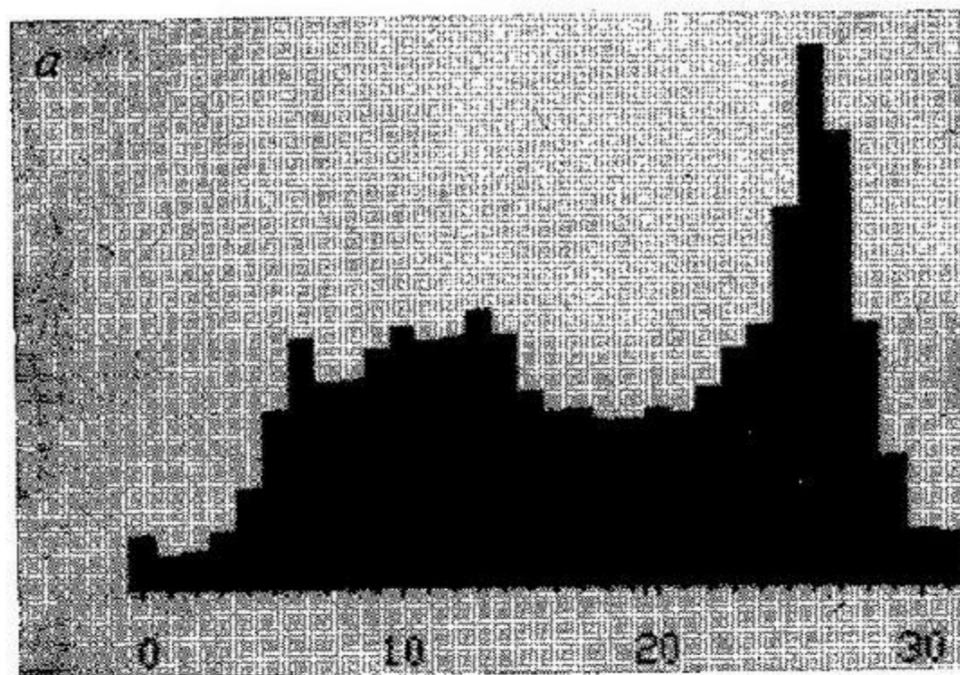
1. *Метод процентилей.* В тех случаях, когда площадь, занимаемая искомым подмножеством, приблизительно известна (например, символы занимают определенную часть площади страницы печатного текста), можно в качестве значения порога использовать соответствующий процентиль распределения; можно найти, следовательно, такой наименьший порог t , что площадь подмножества, на

котором выполняется соотношение $f(x, y) \geq t$, не превосходит заданной части площади изображения ([55, гл. 7).

2. *Метод моды.* Если диапазоны изменения яркости элементов искомого подмножества и остальной части изображения не совпадают, то в распределении частот может появиться пик (мода), соответствующий значениям яркости выделяемого подмножества; таким образом, разбиение изображения можно осуществить, проводя порог по дну «оврагов», окаймляющих этот пик ([80] гл. 7).

В качестве примера приложения изложенных методов рассмотрим приведенную на фиг. 3.1, а картину облачного покрова, распределение частот яркостей которой приведено в табл. 1. Если для назначения величины порога воспользоваться методом процентилей, например попытаться обеспечить максимально возможную близость числа элементов, расположенных по обе стороны порога, то анализ табл. 1 показывает, что лучше всего поместить порог между семнадцатым и восемнадцатым уровнями яркости. Аналогичным образом, назначая величину порога с помощью метода моды, можно с помощью таблицы (или с помощью приведенной на фиг. 8.2, а гистограммы) установить, что центр «оврага», разделяющего два пика, расположен между восемнадцатым и девятнадцатым уровнями интенсивности. На фиг. 8.2, б и в приведены изображения, полученные в результате квантования с порогами, выбранными с помощью двух описанных методов. Эти изображения следует сопоставить с приведенным на фиг. 3.1, д, которое соответствует произвольному назначению величины порога квантования в средней точке серой шкалы, т. е. между пятнадцатым и шестнадцатым уровнями интенсивности.

Другой подход к выбору величины порога состоит в последовательной проверке принадлежности точек изображения к выделяемому подмножеству. В этом случае оказывается возможным изменять порог в зависимости от точек, принадлежность которых к подмножеству уже определена. Можно, например, отнести новую точку к выделенному подмножеству только при условии, что ее яркость отличается от средней яркости точек, включенных к моменту испытания в выделяемое подмножество, не более чем на некоторую долю среднеквадратичного от-



Ф и г. 8.2.

a — гистограмма, характеризующая количество элементов изображения, приведенного на фиг. 3.1, *a*, принимающих значения яркости, соответствующие уровням квантования 0, 1, ..., 31; *b* — изображение, полученное в результате квантования на два уровня изображения, приведенного на фиг. 3.1, *a*. Порог квантования, заключенный между уровнями 17 и 18, выбран с помощью метода процентилей. Элементы, которые соответствуют уровням со значениями яркости $0 \div 17$, представлены на полученном в результате квантования изображении как белые, а уровням $18 \div 31$ — как черные; *в* — изображение, полученное в результате квантования на два уровня изображения, приведенного на фиг. 3.1, *a*. Порог квантования между уровнями 18 и 19 выбран с помощью метода моды.

клонения яркости подмножества ([70] гл. 6). Недостаток такого приема заключается в том, что результат зависит от порядка просмотра точек изображения.

8.2. Получение новых подмножеств изображения на основе заданного. 1. Связность

После выделения некоторого подмножества изображения появляется возможность совершать различные операции над последним. В частности, исходя из известного, из подмножества можно выделить новые подмножества. В настоящем разделе будут рассмотрены операции, использующие понятие связности.



Ф и г. 8.2.

Граница и внутренность множества

Рассмотрим подмножество S дискретного изображения и обозначим через \bar{S} дополнение подмножества S , т. е. множество элементов изображения, не принадлежащих подмножеству S . Элемент подмножества S с координатами (i, j) называют *изолированным*, если элементы изображения с координатами $(i-1, j)$, $(i+1, j)$, $(i, j-1)$ и $(i, j+1)$, называемые смежными, принадлежат дополнению \bar{S} ¹. Аналогичным образом элемент (i, j) называется *внутренней точкой подмножества S* , если все смежные с ним элементы входят в это подмножество. Если часть элементов, смежных с некоторым элементом (i, j) , принадлежит подмножеству S , а часть — его дополнению \bar{S} , то такой элемент называют *граничным*. В специальном случае, когда два смежных *противоположных* элемента [например, $(i-1, j)$ и $(i+1, j)$ или $(i, j-1)$ и $(i, j+1)$] принадлежат дополнению \bar{S} подмножества S , а другие два — самому подмножеству, элемент (i, j) называется *дуговым*. Если три элемента, смежных с элементом (i, j) , входят в дополнение \bar{S} , то последний можно было бы называть *дуговым конечным* элементом. Вводя в рассмотрение элементы (пары), смежные с исходным (или смежные со смежными и т. д.) по разным направлениям, можно определять еще более точные классы.

Граница и внутренность подмножества (т. е. множества соответственно граничных и внутренних элементов его) являются тривиальными примерами новых подмножеств — производных от заданного. Приведем несколько менее тривиальных примеров.

1. В разд. 6.4 было показано, что двоичное изображение можно очистить от шума, заменив единичные элементы на нулевые в тех точках, где большая часть смежных элементов — нулевые. Эту операцию можно интерпретировать как выделение в новое подмножество элементов

¹ Метод, развиваемый в данном разделе, позволяет ввести альтернативные понятия, основанные на использовании четырех смежных элементов, но на этот раз диагональных, т. е. имеющих следующие координаты: $(i-1, j-1)$, $(i-1, j+1)$, $(i+1, j-1)$ и $(i+1, j+1)$.

подмножества, имеющих много смежных элементов в подмножестве. Так, например, различные типы элементов подмножества, включающего сплошные области, дуги или кривые и точки, порожденные помехами, более или менее просто разделяются при помощи соответствующего квантования по количеству смежных элементов в подмножестве.

2. Рассмотрим изображение, над которым осуществлена операция квантования с двумя порогами $t_1 < t_2$. Обозначим через S_2 множество элементов, значения которых превышают величину t_2 . Подмножество S_2 можно считать «ядром» выделяемого и попробовать «расширить» его за счет включения элементов со значениями, превышающими порог t_1 , при условии, что они обладают соответствующим количеством смежных элементов в подмножестве S_2 .

Связные компоненты

Элементы подмножества S с координатами (i, j) и (h, k) называют *связными*, если существует такая последовательность принадлежащих подмножеству S элементов $(i, j) = (i_0, j_0)$, (i_1, j_1) , ..., $(i_m, j_m) = (h, k)$, в которой элемент (i_r, j_r) является смежным с элементом (i_{r-1}, j_{r-1}) при всех $1 \leq r \leq m$. Для всех входящих в подмножество S элементов (i, j) множество элементов (h, k) , в котором элементы (i, j) и (h, k) связаны между собой, называют связным компонентом подмножества S . Подмножество S , имеющее только один компонент, называют *связным*.

Можно рассматривать и компоненты дополнения \bar{S} подмножества S . Существует целый ряд причин [3], по которым предпочтительнее использовать понятие «связности» дополнения \bar{S} , основанное на учете восьми смежных элементов, если для подмножества S вводилось понятие, основанное на учете четырех смежных элементов, и наоборот. В дальнейшем будем считать, что любое изображение окаймлено границей, состоящей из нулевых элементов. Компонента дополнения \bar{S} , состоящая из элементов, связанных с этими нулевыми элементами, называется *внешней стороной* множества S . Если дополнение \bar{S} имеет

другие компоненты, их называют *дырами* подмножества S ; при этом говорят, что любой элемент дыры находится *внутри* подмножества S . Если подмножество S — связное и не имеет дыр, то его называют *односвязным*; при наличии дыр подмножество называется *многосвязным*.

Отслеживание границы [4—10]

Рассмотрим связное множество S и множество B граничных элементов множества S , смежные элементы которых входят в одну из компонент дополнения \bar{S} . Можно легко показать связность множества B и, более того, обосновать тот факт, что при выборе в качестве отправной точки любого элемента множества B и вполне определенном способе перехода к смежным с ним элементам обеспечивается обход множества B по контуру и возвращение в исходную точку. Ниже приводится алгоритм обхода множества B против часовой стрелки — в его основе лежит прием «сохраняйте множество S слева». Доказательство универсальности этого алгоритма приведено в работе [3].

Пусть $x_0 = (i_0, j_0)$ — исходный (произвольный) элемент множества B и $y_1^{(0)}$ представляет один из элементов $(i_0 - 1, j_0)$, $(i_0 + 1, j_0)$, $(i_0, j_0 - 1)$, $(i_0, j_0 + 1)$, входящий в дополнение \bar{S} (во всяком случае, один из них принадлежит S , так как (i_0, j_0) — граничный элемент множества S). Выберем, двигаясь против часовой стрелки, *восемь* элементов, смежных с (i_0, j_0) , причем первый из них y_1^0 ; обозначим эти элементы $y_1^{(0)}, \dots, y_8^{(0)}$. Если ни один из элементов $y_3^{(0)}, y_5^{(0)}, y_7^{(0)}$ не входит в множество S , то x_0 — единственный элемент множества S . Если такого элемента нет, обозначим через $y_{2i+1}^{(0)}$ первый принадлежащий множеству S элемент y . Далее в качестве следующего элемента x_1 множества B выбирается

Элемент $y_{2i}^{(0)} + 1$, если элемент $y_{2i}^{(0)}$ входит в дополнение \bar{S} ;

Элемент $y_{2i}^{(0)}$, если элемент $y_{2i}^{(0)}$ входит в множество S .

Отметим, что элемент x_1 в обоих случаях действительно является граничным, так как смежный с ним элемент $y_{2i}^{(0)}$ (или $y_{2i-1}^{(0)}$) принадлежит дополнению \bar{S} . Чтобы найти следующий элемент, возьмем упомянутый выше *смежный элемент* в качестве $y_1^{(1)}$ и полностью повторим процедуру первого и последующих шагов.

Упражнение 1. Показать, что при использовании в определении связной компоненты восьми смежных элементов алгоритм может быть упрощен следующим образом. Пусть элемент $y_i^{(0)}$ — первый из входящих в множество S элементов $y^{(0)}$; в таком случае можно считать, что $x_1 = y_i^{(0)}$ и $y_1^{(1)} = y_{i-1}^{(0)}$.

Упражнение 2. Пусть z_0 — произвольный элемент множества B , а z_1 — один из четырех смежных с z_0 элементов; последующие элементы z определяются следующим образом: если последний элемент z принадлежал множеству S , следует идти вправо; если этот элемент входил в дополнение \bar{S} , следует идти влево [37]. Показать, что траекторией этой процедуры является спираль, закрученная против часовой стрелки вокруг множества B (границы), при условии отсутствия в нем точек самоприкосновения.

Известны методы отслеживания границ и кривых, которые можно применить для обработки непрерывных изображений. Допустим, что, сканируя изображение, мы движемся по его поверхности по спирали; это означает, что по периметру небольшой окружности совершаются очень быстрые перемещения, в то время как центр этой окружности несколько медленнее движется поперек изображения. Пусть нас интересуют расположенные вне окружности точки, в которых производная яркости, взятая вдоль окружности, достигает больших абсолютных значений. Если центр окружности находится на кромке, существуют две такие точки, отстоящие друг от друга приблизительно на 180° . По мере перемещения центра эти точки сближаются, за исключением случая движения вдоль кромки. В сущности направление и величину смещения этих точек можно использовать для коррекции положения центра, с тем чтобы сохранить его позицию на кромке. Описанную процедуру можно модифицировать и использовать

также для отслеживания кривых; при этом следует выбирать окружность диаметра, большего, чем толщина кривой¹.

Маркировка компонент

В следующих разделах будут описаны методы, позволяющие выделять каждую компоненту заданного подмножества изображения в качестве отдельного «объекта». Рассмотрим множество S единичных элементов двоичного изображения. Введем маркировку компонент множества S для того, чтобы иметь возможность различать их. Каждому элементу множества S будет поставлено в соответствие положительное целое число таким способом, чтобы все элементы, принадлежащие определенной компоненте, характеризовались одной и той же величиной, а элементы различных компонент получали разные числовые метки.

«Распространение». Рассмотрим случай, когда элементы изображения подвергаются осмотру в любом систематическом порядке до тех пор, пока не обнаружится элемент, имеющий значение «1». Этому элементу присваивается новое значение, например «2». Затем снова производится локальная операция, в процессе которой любой элемент изображения со значением «1», смежный с элементом со значением «2», сам получает значение «2». Если в результате осуществления этой операции создается ситуация, не допускающая дальнейших замен, то, очевидно, все элементы со значением «1», принадлежащие той же компоненте, в которую входит исходный элемент, приобрели значение «2». Отметим, что необходимое количество повторений операции может оказаться соизмеримым с числом элементов изображения, например в случае, когда компонента имеет следующий характер:

```

1111111111
0000000001
1111111111
1000000000
1111111111

```

¹ О системах и методах следящей развертки см. [57*]. — Прим. ред.

На следующем этапе систематический осмотр элементов возобновляется и продолжается вплоть до обнаружения другого «единичного» элемента. Очевидно, что этот элемент должен принадлежать другой компоненте. Этому элементу присваивается новое значение, например «3», и весь процесс повторяется опять, в результате чего «единичные» элементы новой компоненты приобретают значение «3». Вся эта процедура повторяется вплоть до полного завершения осмотра изображения; к этому моменту каждая компонента единичных элементов оказывается единообразно помеченной целым числом, не меньшим 2. Проблемы, связанные с применением подобных методов для подсчета связанных компонент (разд. 9.1), рассматриваются в работах [11—13], а также [12] гл. 5.

Этот метод можно модифицировать таким образом, чтобы не требовалось применять локальную операцию к каждой обнаруживаемой компоненте отдельно. Для этого нужно начинать с замены всех «единиц» на изображении различными целыми положительными числами, скажем присваивая этим элементам последовательно значения 1, 2, 3, В таком случае можно многократно применять локальную операцию, заменяющую значение элемента a_{ij} на

$$\max \{a_{ij}, a_{i-1, j}, a_{i+1, j}, a_{i, j-1}, a_{i, j+1}\}$$

(изменение значений нулевых элементов, однако, не допускается).

Когда в результате выполнения этой процедуры возникнет ситуация, в которой дальнейшие изменения невозможны, каждому элементу определенной компоненты будет присвоено наибольшее из значений, присвоенных элементам данной компоненты первоначально.

Отслеживание границы. Допустим, что изображение осматривается строка за строкой, причем сканирование производится слева направо. В этом случае внешней границей компоненты, состоящей из единичных элементов, служит первый встретившийся при осмотре переход от нулевого к единичному элементу. Границей дыры считается соответственно первый переход от единичного значения к нулевому. Как только в процессе осмотра будет обнаружена внешняя граница, направление осмотра

изменится и далее будем следовать вдоль границы, заменяя значения ее элементов («1») положительными числами, не меньшими «3». Аналогичным образом при выходе на границу дыры осмотр продолжается вдоль нее и все «1» заменяются «2», а все значения, большие двух, сохраняются. При движении вдоль границы встречаются такие конфигурации, когда дыру от внешней стороны компоненты отделяет лишь один элемент. Когда обход границы заканчивается (как в случае компоненты, так и в случае дыры), возобновляется систематический осмотр изображения; при этом границы, отслеженные на предыдущих этапах, пересекаются и не влекут изменения направления процедуры осмотра, поскольку теперь значения их элементов не равны единице. После завершения сканирования всего изображения внешние границы каждой компоненты имеют свою маркировку, представленную целым числом, не меньшим трех. Далее можно с помощью описанного выше многократного применения к изображению локальной операции распространить «границные» значения на внутренние точки соответствующих компонент (процедура подсчета компонент, основанная на отслеживании границ, описана в работе [94] гл. 7). С другой стороны, после полного завершения отслеживания внешней границы можно предпринять систематический осмотр ее внутренней стороны: обнаружив границу дыры, можно, следуя по ней, полностью обойти ее и затем приступить к систематическому осмотру ее внутренней части; обнаружив внешнюю границу (т. е. границу новой компоненты, расположенной в дыре исходной компоненты), можно, следуя по ней, полностью обойти ее и начать затем осмотр ее внутренней части и т. д. [10, 14]. Преимуществом этого метода является то обстоятельство, что он весьма просто обеспечивает выбор такого направления сканирования, при котором последовательно осматриваются дыры, находящиеся внутри каждой компоненты, затем компоненты, лежащие в каждой из дыр, и т. д.

«Получение серой шкалы». Допустим, что изображение изучается последовательно, строка за строкой; в принципе любая строка состоит из серий, образованных из нулей, чередующихся с сериями, образованными из единиц. Первая серия из единиц, стоящая в первой строке, сохраняется

в том же виде, единицы второй серии заменяются двойками, третьей серии — тройками и т. д. Далее предпринимается исследование единичных серий второй строки. Если единичная серия второй строки не покрывает ни одну единичную серию первой строки, то она получает новую маркировку; если она покрывает одну серию, то получает метку последней; если единичная серия второй строки покрывает более одной единичной серии первой строки, то она получает наименьшую из меток покрытых серий, а в таблице эквивалентности фиксируется, что все эти метки серий первой строки принадлежат одной связанной компоненте. Сравнение серий каждой строки изображения с сериями предыдущей строки продолжается вплоть до сопоставления последней пары строк. На следующем этапе объектом изучения становится таблица эквивалентности; рассматриваются все метки и для каждой отыскивается минимальная, присвоенная той связанной компоненте, к которой принадлежит исследуемая метка. Следующий этап состоит в повторном сканировании изображения и замене каждой метки ее минимальным эквивалентом. В результате все связанные компоненты получают свои различные метки. Модификация описанного метода, предусматривающая подсчет компонент, рассматривается в работах [76, 92] гл. 7 и [26].

Внутренние и внешние элементы подмножества

Чтобы определить, находится ли элемент дополнения \bar{S} вне подмножества S или внутри него, т. е. в дыре, можно использовать описанный только что метод: маркировка элементов, имеющих значения «0», изменяется таким образом, чтобы внешние и внутренние элементы имели разные значения. Так, например, можно осматривать изображение вдоль границы и заменять все нули двойками, а затем многократно применять локальную операцию, обеспечивающую замену нулей двойками у элементов, смежных с элементами со значением «2». К моменту завершения этой процедуры все возможные замены оказываются осуществленными — все внешние элементы с нулевыми значениями превратились в «двойки», а все внутренние «нулевые» элементы такими и остались.

Подход, предусматривающий замену маркировки, целесообразно использовать, если решение «внутри или снаружи» приходится принимать на многих точках; если же таких точек лишь несколько, то можно обратиться к другим методам, например алгоритмам [16—18], позволяющим определить, находится ли точка внутри или вне многоугольника.

Стягивание и сжатие подмножества

Если в подмножество входит много небольших связанных компонент, может возникнуть желание упростить изображение, «стянув» каждую компоненту в отдельный элемент (этот прием облегчает к тому же пересчет компонент, так как подсчитываются элементы, полученные в результате стягивания [19, 20]). Если подмножество выглядит чрезвычайно удлиненным (след в пузырьковой камере, рукописное слово и т. п.), аналогичным образом может появиться желание «сжать» его в линию.

Чтобы заменить компоненты отдельными элементами, достаточно промаркировать первые и затем изъять все, за исключением одного, элементы с одинаковыми метками; к сожалению, этап назначения меток может потребовать очень много времени, особенно если число компонент велико. Ниже излагается более эффективный метод стягивания односвязных компонент единичных элементов в один «единичный» элемент. Элементы изображения исследуются последовательно и значение элемента изменяется с «1» на «0», если: 1) хотя бы один из смежных с ним элементов принадлежит подмножеству S и хотя бы один из смежных с ним элементов — его дополнению \bar{S} ; 2) все смежные с ним элементы, принадлежащие подмножеству S , образуют связное множество, являющееся подмножеством множества *восьми* смежных элементов. (Доказательство сходимости этого алгоритма в общем случае приведено в работе [3]. Известен несколько более сложный алгоритм стягивания односвязных компонент в отдельные элементы; этот алгоритм предусматривает многократную обработку изображения в параллельном режиме, т. е. он одновременно применяется ко всем точкам изображения.)

Чтобы уменьшить ширину подмножества, достаточно удалить с изображения соответствующее число последовательных слоев граничных элементов, учитывая ограничения алгоритма стягивания, обеспечивающие сохранение связности подмножества ([45] гл. 8). Несколько простых алгоритмов сжатия приведено в работах [21, 22], а также в работе [17] гл. 5. Проблема построения гладкой кривой по разрозненным точкам обсуждается в статье [23].

Упражнение 3. Показать, что при использовании определения связности, основанного на учете восьми смежных элементов, второе условие в алгоритме стягивания может быть упрощено следующим образом: множество смежных с ним элементов из подмножества S — связное.

Упражнение 4. Назовем элемент подмножества «концевым», если он имеет три смежных элемента в дополнении \bar{S} , «угловым», если его два последовательных смежных элемента принадлежат подмножеству S и в него же входит расположенный между ними диагональный элемент. Рассмотрим операцию ϕ , обеспечивающую исключение «концевых», «северо-западных» и «северо-восточных» «угловых» элементов, за исключением набранных жирным шрифтом, в следующих двух случаях (символом обозначен элемент, который может принять значение как «0», так и «1»):

$$\begin{array}{ccc} * 0 * & & * 0 0 * \\ 0 1 0 & & 0 1 1 0 \\ 0 1 0 & \text{и} & * 0 0 * \\ * 0 * & & \end{array}$$

Показать, что при многократном применении в параллельном режиме операции ϕ к каждому элементу изображения обеспечивается стягивание каждой односвязной компоненты «единичных» элементов в отдельный элемент.

8.3. Получение новых подмножеств изображения на основе заданного. 2. Расстояние

Рассмотрим множество A всех пар целых чисел и функцию d , которая парам элементов множества A ставит в соответствие неотрицательные целые числа. Эту

функцию d называют *функцией расстояния*, если она является *положительно определенной* и *симметрической*, а также удовлетворяет *неравенству треугольника*, т. е. если для всех элементов x, y и z множества A выполняются следующие соотношения:

$$d(x, y) = 0 \text{ в том и только том случае, если } x = y;$$

$$d(x, y) = d(y, x);$$

$$d(x, z) \leq d(x, y) + d(y, z).$$

Если B — подмножество множества A , то запись $d(x, B)$ означает $\min_{y \in B} d(x, y)$, т. е. расстояние между элементом x и ближайшим к нему элементом множества B .

Рассмотрим два характерных примера функции расстояния:

$$1. d_1((i, j), (h, k)) = |i - h| + |j - k|.$$

В этом случае множество точек с координатами (h, k) , отстоящих не более чем на r от точки с координатами (i, j) , ограничено ромбом со сторонами длиной $r\sqrt{2}$ и центром в точке (i, j) . В частности, точки, находящиеся на расстоянии 1 от центральной точки, являются четырьмя точками, смежными с ней по горизонтали и вертикали.

$$2. d_2((i, j), (h, k)) = \max(|i - h|, |j - k|).$$

В данном случае множество точек с координатами (h, k) , удаленных от центра (i, j) не более чем на r , ограничено квадратом (направление сторон которого совпадает с горизонтальной и вертикальной осями соответственно) со сторонами длиной $2r$ и центром в точке (i, j) ; точки, отстоящие от центра на расстояние 1, представляют собой восемь элементов, смежных с центром.

Могут быть введены и такие функции расстояния, для которых множество элементов, удаленных от определенной точки на заданное расстояние, ограничено восьмиугольником или (с некоторым приближением) шестиугольником. («Настоящие» шестиугольники можно получить преобразованием «квадратного» массива A в «шестиуголь-

ный», например, при помощи сдвига вправо строк, отстоящих на расстояние $1/2$)

Упражнение 1. Показать, что целое число $d((i, j), (h, k))$, ближайшее к значению выражения $[(i - h)^2 + (j - k)^2]^{1/2}$, не есть функция расстояния.

Функцию расстояния называют *регулярной*, если она обладает следующим свойством: для любых элементов x и z множества A в нем найдется элемент y , такой, что $d(x, y) = 1$ и $d(y, z) = d(x, z) - 1$. Очевидно, что функции расстояния d_1 и d_2 — регулярные. Используя принцип индукции, нетрудно обосновать следующий алгоритм вычисления значений регулярной функции расстояния на множестве «нулевых» элементов (которое предполагается непустым) двоичного изображения.

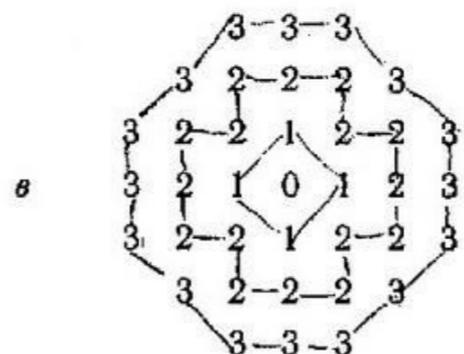
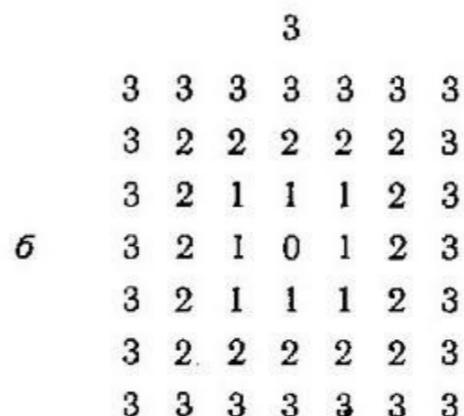
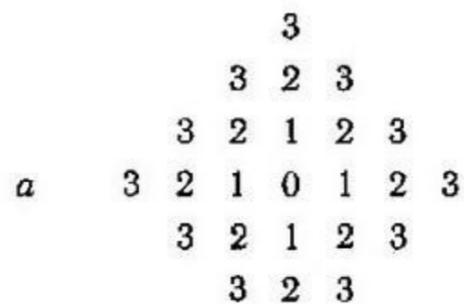
Теорема. Пусть d — регулярная функция расстояния, (a_{ij}) — двоичное изображение, состоящее из n строк и n столбцов, и S — множество точек (i, j) , для которых $a_{ij} = 0$; пусть также последовательность изображений (a_{ij}^k) с целочисленными значениями яркости определяется таким образом

$$a_{ij}^{(k)} = \min_{\substack{d((u, v), (i, j)) < 1}} a_{uv}^{(k-1)} + a_{ij}^{(0)}$$

для $k = 1, 2, \dots$, причем $a_{ij}^{(0)} = a_{ij}$. В таком случае при достаточно больших значениях k (для функции d_1 , например, можно принимать $k = 2n$, для функции d_2 достаточно брать $k = n$) для всех точек (i, j) справедливо соотношение $a_{ij}^{(k)} = d((i, j), S)$.

Действительно, легко убедиться, что данный алгоритм последовательно осуществляет замену значений всех внутренних элементов множества: «1» на «2», «2» на «3» и т. д. до тех пор, пока не будут исчерпаны все возможные замены (сравните с методом, изложенным в работе [11] гл. 8). Легко проверить, что при воспроизведении последовательности a_{ij}^k , используя попеременно через шаг алгоритмы с функциями расстояния d_1 и d_2 , контур, ограничивающий множество точек, отстоящих от выбранной на заданное расстояние, оказывается восьмиугольником. На фиг. 8.3 представлены ромбовидный, квадратный и вось-

миугольный контуры, ограничивающие точки, удаленные на определенное расстояние от отдельного «нулевого» элемента. Заметим, что данный алгоритм можно применять



Ф и г. 8.3. Подмножества точек, удаленные на определенное расстояние от отдельного «нулевого» элемента.

a — функция расстояния типа d_1 , граница подмножества имеет форму ромба; б — функция расстояния типа d_2 , граница подмножества имеет форму квадрата; в — функции расстояния типа d_1 и d_2 используются поочередно, граница подмножества имеет форму восьмиугольника.

не только к двоичным изображениям [29]; если все интенсивности исходного изображения $a_{ij}^{(0)}$ находятся в диапазоне $(0 \div 1)$, то «искомое» расстояние быстро увеличивается (практически почти столь же быстро, как и реальное расстояние) в тех областях, где интенсивность близка к

единице, и возрастает очень медленно там, где она близка к нулю.

Функции расстояния можно использовать для выделения из заданного подмножества подмножеств изображения различных типов. В следующих примерах через S^r обозначается подмножество точек (i, j) , таких, что для них выполняется неравенство $d((i, j), S) \leq r$. Говоря о d -круге радиуса r с центром в точке (i, j) , мы будем иметь в виду множество точек (h, k) , на котором справедливо неравенство $d((i, j), (h, k)) \leq r$.

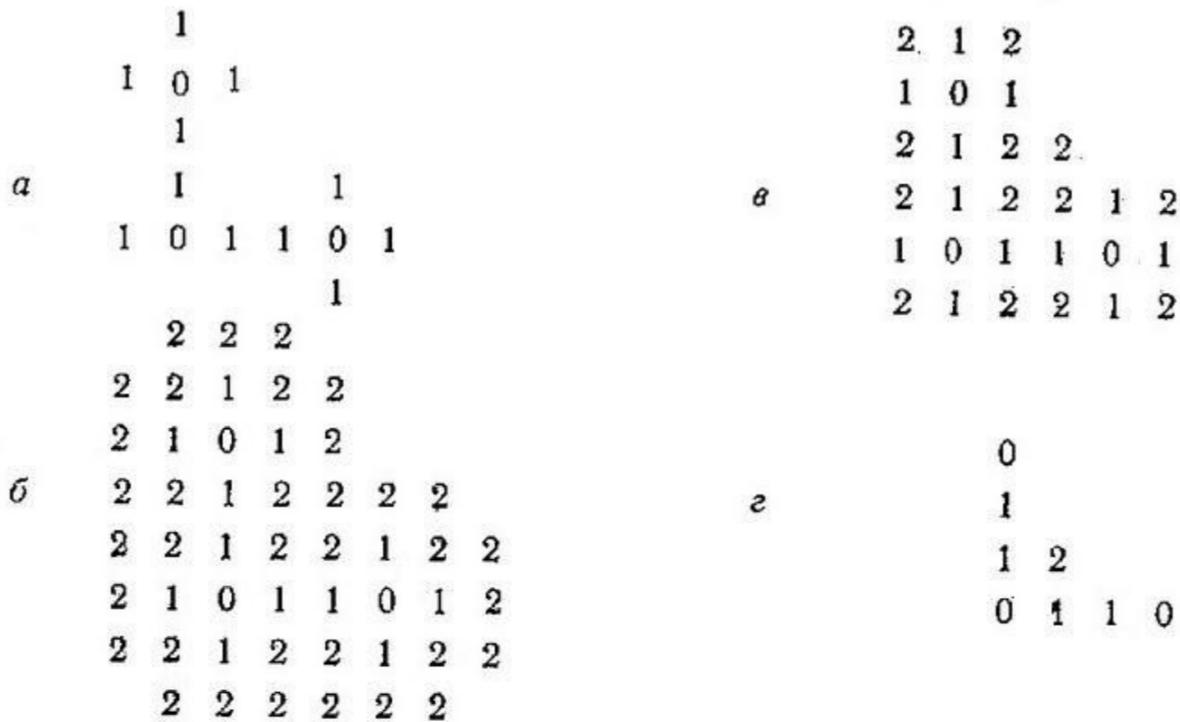
Изолированные зоны подмножества

Рассмотрим произвольную связную компоненту «С» подмножества S^r . Если площадь участка, занимаемого компонентой «С» на изображении, близка к площади d -круга радиуса r , то содержащаяся в ней часть множества S должна иметь размеры, существенно меньшие радиуса r ; так как эта часть множества S отстоит по крайней мере на расстояние $2r$ от любого другого элемента множества S , ее с полным основанием можно называть *изолированной*.

Сгущения и удлиненные зоны подмножества

При достаточно больших значениях r подмножество S^r содержит изображение полностью; в связи с чем его дополнение $\overline{S^r}$ является пустым; отсюда следует, что таким же будет $(\overline{S^r})^s$ при любых значениях s . С другой стороны, при $r = 0$ подмножество S^r совпадает с подмножеством S и, следовательно, дополнение $(\overline{S^0})^s = \overline{S^s}$ содержит дополнение \overline{S} при любых значениях s . Таким образом, для любой точки (i, j) , не принадлежащей подмножеству S , найдется такое наименьшее и явно большее нуля значение r (обозначим его r_{ij}), что точка (i, j) не входит в дополнение $(\overline{S^r})^s$. Мы будем называть этот параметр r_{ij} *радиусом слияния* S с точкой (i, j) ; если точка (i, j) при-

надлежит подмножеству S , $r_{ij} = 0$ по определению. Интуитивно понятно, что выполнение неравенства $r_{ij} \leq r$ отражает следующий факт: если подмножество S расши-



Ф и г. 8.4. Иллюстрация процесса «слияния»; подмножество точек, удаленных на фиксированное расстояние, ограничено восьмиугольником; множество S состоит из трех элементов (они обозначены через 0).

$a - S^1$ — множество элементов, отстоящих от множества S на расстояние ≤ 1 ; отметим, что все «единичные» элементы являются граничными элементами множества S^1 ; таким образом, все они отстоят на расстояние «1» от дополнения \bar{S}^1 , т. е. принадлежат дополнению $(\bar{S}^1)^1$ и, следовательно, при сжатии множества S^1 на величину одного приращения сохраняются лишь «нулевые» элементы; «слияние» не происходит; $b - S^2$ — множество элементов, ограниченное восьмиугольником и отстоящее на расстояние ≤ 2 от множества S ; e — множество элементов S^2 , из которого удалены элементы дополнения $(\bar{S}^2)^1$; e — множество элементов S^2 , из которого удалены элементы дополнения $(\bar{S}^2)^2$; «ненулевые» элементы незатронутые слиянием, представляют собой элементы дополнения \bar{S} , для которых радиус слияния равен двум.

рено так, что все точки подмножества S заменены d -кругами радиуса r , а затем снова «сжато» на такую же величину, то точка (i, j) оказывается включенной в полученное в результате этих операций подмножество. Если точка (i, j) не входит в исходное подмножество S , то единственной причиной ее отсутствия может быть слияние нескольких расширенных зон подмножества S , происшед-

шее таким образом, что поблизости не остается места, достаточного для включения точки (i, j) при сжатии к первоначальной форме; в связи с этим точка (i, j) продолжает оставаться выделенной даже после сжатия. Фиг. 8.4 иллюстрирует описанный процесс.

Упражнение 2. Определить радиусы слияния для множества точек (ki, kj) (i, j, k — целые числа), исходя из функций расстояния типа d_1 и d_2 .

Рассмотрим множество точек S_t с радиусом слияния $r_{ij} \leq t$ и связную компоненту C этого множества. Очевидно, что компонента порождается слиянием сгущения элементов подмножества S , если площадь компоненты велика по сравнению с величиной t^2 . Чем большим приходится выбирать параметр t для того, чтобы обеспечить слияние, тем более разреженным оказывается сгущение.

Рассмотрим аналогично радиус слияния r'_{ij} дополнения \bar{S} в точке (i, j) и множество точек S'_t , на котором выполняется неравенство $r'_{ij} \leq t$; пусть, кроме того, $S'_t = S - S'_t$. Если площадь связной компоненты подмножества S'_t велика по сравнению с t^2 , то, очевидно, она должна быть *удлиненной зоной* подмножества S , поскольку, несмотря на значительность площади, компонента «исчезает» при расширении и сжатии дополнения \bar{S} на r и, следовательно, ширина компоненты в любой точке не может превышать величину $2r$.

Остов подмножества

Множество точек, расстояние которых до множества \bar{S} достигает локального максимума (или близко к нему), образует своего рода остов подмножества S . (Можно показать, что эти точки вместе с соответствующими значениями расстояния полностью определяют подмножество S [26].) Множество элементов, соответствующих острым углам «контура постоянного расстояния» (т. е. геометрические места точек, отстоящих от дополнения \bar{S} на фиксированное расстояние), а также множество элементов, два или несколько ближайших элементов которого принадлежат дополнению \bar{S} , также являются своего рода

остовами. Введение понятия «остов» позволяет применить новый подход к сжатию подмножества S ; его можно использовать и для получения обобщенной информации о форме подмножества S (например, «выступам» подмно-

	0	0																			
	0	1	0																		
	0	1	1	0																	
	0	1	2	1	0																
	0	1	2	2	1	0															
	0	1	2	3	2	1	0														
<i>a</i>	0	1	2	3	3	2	1	0													
	0	1	2	3	4	3	2	1	0												
	0	1	2	3	3	3	3	2	1	0											
	0	1	2	2	2	2	2	2	2	1	0										
	0	1	1	1	1	1	1	1	1	1	1	0									
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0																			
	0	1	0																		
	0	1	1	0																	
	0	1	1	1	0																
	0	1	2	1	1	0															
<i>b</i>	0	1	2	2	1	1	0														
	0	1	2	2	2	1	1	0													
	0	1	2	3	2	2	1	1	0												
	0	1	2	3	3	2	2	1	1	0											
	0	1	2	2	2	2	2	1	1	0											
	0	1	1	1	1	1	1	1	1	1	0										
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Фиг. 8.5. Расстояния, ограниченные контурами: *a* — «ромба», *б* — «квадрата», *в* — «восьмиугольника» от элементов треугольника до окружающих его «нулевых» элементов; элементы остова выделены жирным шрифтом.

В случае *a* элемент входит в подмножество, если его значение не меньше значения не меньше значений восьми смежных с ним элементов; в случае *б* — критерий восьми элементов — если четное. Можно показать, что в квадратах или восьмиугольниках (соответственно), центры которых совпадают с центрами элементов остова, являющегося его центром,

жества S соответствуют «ветви» его остова). На фиг. 8.5 показаны остовы максимального расстояния треугольника при «ромбовидной», «квадратной» и «восьмиугольной» метриках соответственно. Остов максимального расстояния облаков, изображенных на фиг. 3.1, *д*, при «восьмиугольной» метрике приведен на фиг. 8.6.

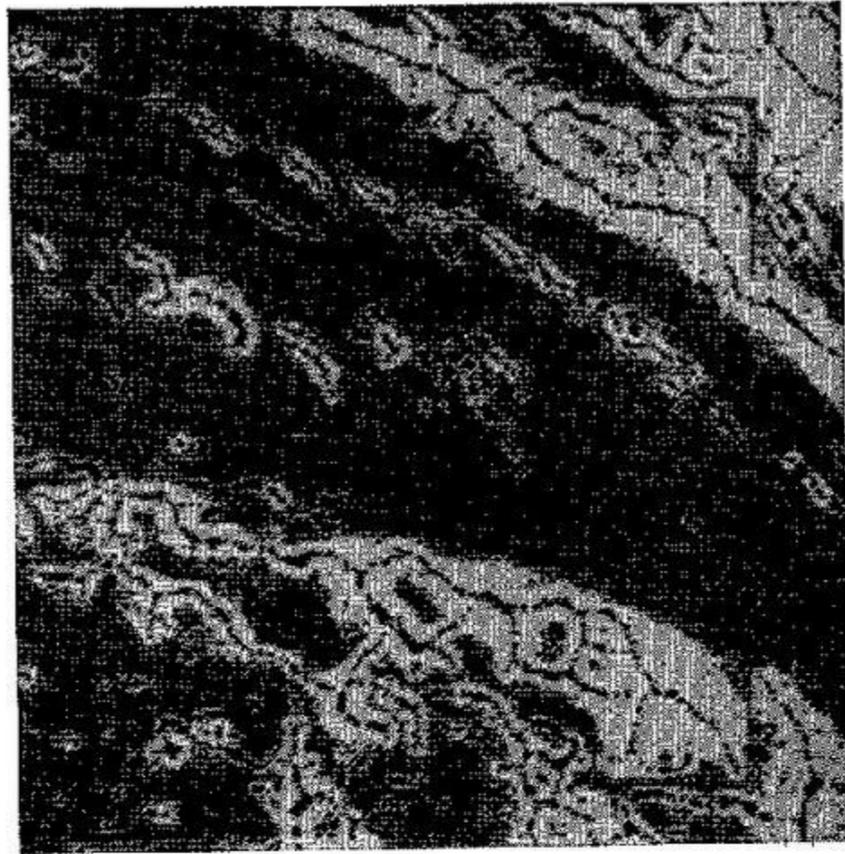
Упражнение 3. Пусть

$$c_{ij}^{(k)} = c_{i-1, j}^{(k-1)} + c_{i+1, j}^{(k-1)} + c_{i, j-1}^{(k-1)} + c_{i, j+1}^{(k-1)}, \quad k=1, 2, \dots,$$

где $c_{ij}^{(0)} = c_{ij}$. Доказать, что выражение $c_{ij}^{(k)}$ является суммой значений яркостей в точках, расположенных на концах траекторий типа «движение по границам городских кварталов» длины k , исходящих из точки (i, j) . (Замечание: траектория длины k , исходящая из точки (i, j) , состоит из шага из точки (i, j) к одной из смежных с ней точек, за которым следует движение по траектории длины $k-1$ из этой смежной точки.) Отметим, что в тех

	0	0																			
	0	1	0																		
	0	1	1	0																	
	0	1	2	1	0																
	0	1	2	2	1	0															
	0	1	2	2	2	1	0														
<i>в</i>	0	1	2	3	2	2	1	0													
	0	1	2	3	3	2	2	1	0												
	0	1	2	3	3	3	2	2	1	0											
	0	1	2	2	2	2	2	2	1	0											
	0	1	1	1	1	1	1	1	1	1	0										
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

каждом из четырех смежных с ним элементов; в случае *б* — критерий четырех элементов используется, если значение элемента во всех случаях треугольника является объединением «кругов» — ромбов, с элементами остова, причем радиус каждого круга на единицу меньше



Ф и г. 8.6. Остов подмножества, выделенный по принципу максимального расстояния элементов подмножества, ограниченного восьмиугольником (расстояние от частей изображения, свободных от облачного покрова, приведенного на фиг. 3.1, д).

Чтобы добиться большей определенности, зоне, свободной от облачного покрова, придана средняя яркость, в то же время яркость всех элементов остова превышает среднюю и соответствует своим значениям расстояния.

случаях, когда массив (c_{ij}) — двоичный, величина $c_{ij}^{(k)}$ равна числу траекторий длины k из точки (i, j) к множеству «единичных» элементов массива (c_{ij}) .

Упражнение 4. Пусть

$$h_{ij}^{(k)} = h_{i, j-1}^{(k-1)} + h_{i, j+1}^{(k-1)} - h_{ij}^{(k-2)}, \quad k=3, 4, \dots,$$

где $h_{ij}^{(0)} = c_{ij}$, $h_{ij}^{(1)} = c_{i, j-1} + c_{i, j+1}$ и

и $h_{ij}^{(2)} = h_{i, j-1}^{(1)} + h_{i, j+1}^{(1)} - 2h_{ij}^{(0)}$,

$$d_{ij}^{(k)} = d_{i-1, j}^{(k-1)} + d_{i+1, j}^{(k-1)} - d_{ij}^{(k-2)} + h_{ij}^{(k)} - h_{ij}^{(k-2)}, \quad k=2, 3, \dots,$$

где $d_{ij}^{(0)} = c_{ij}$ и $d_{ij}^{(1)} = c_{i, j-1} + c_{i, j+1} + c_{i-1, j} + c_{i+1, j}$.

Доказать, что выражение $d_{ij}^{(k)}$ является суммой яркостей элементов, отстоящих от точки (i, j) на расстояние k (расстояние определяется так же, как в предыдущем упражнении). Отметим, что в тех случаях, когда массив (c_{ij}) — двоичный, сумма $d_{ij}^{(k)}$ равна числу «единичных» элементов, находящихся на расстоянии k от точки (i, j) .

Упражнение 5. Составить алгоритмы для подсчета: а) числа кратчайших путей от точки (i, j) к множеству всех «единичных» элементов S ; б) числа «единичных» элементов, отстоящих от точки (i, j) на расстояние $d((i, j), S)$.

8.4. Получение новых подмножеств изображения на основе заданного. 3. Направление и форма

Ориентированные зоны и тени подмножества

Подмножества непрерывного или дискретного изображения можно выделять, исходя из их ориентации по отношению к заданному, точно так же как это делалось на основе оценки расстояния определенных точек от заданного подмножества. Можно, в частности, разделять те точки подмножества, у которых смежные с ними точки того же подмножества ориентированы по заданным (но разным) направлениям. Этот метод, будучи применен к рисункам, позволяет пометить каждую точку кривой округленным значением направления кривой в этой точке. В качестве другого примера подобного подхода можно привести выделение в качестве нового подмножества множества точек, заслоняемого заданным подмножеством при взгляде на него извне в определенном ракурсе. В это подмножество попадут именно те точки, которые окажутся в тени заданного подмножества при освещении его пучком параллельных лучей света, ориентированным в заданном направлении [34]. В качестве нового подмножества, вообще говоря, можно выбирать такое множество точек, что проходящий через него луч, ориентированный в заданном направлении, пересекает заданное подмножество определенное количество раз ([35, 36], а также [23] гл. 9). Площади сформированных подобным образом мно-

жеств можно использовать как признаки заданного подмножества (гл. 9).

Упражнение 1. Показать, что площади A_L и A_B , попадающие в тень подмножества при освещении его соответственно слева и снизу, можно использовать в качестве признаков при различении множеств «единичных» элементов, содержащихся в шести изображениях из упражнения 1 разд. 6.1.

Выпуклость

Подмножество изображения S называется *выпуклым*, если все точки линейного отрезка, соединяющего любые две точки, входящие в подмножество S , также принадлежат этому подмножеству. Таким образом, выпуклость является линейным пороговым признаком для множества признаков третьего порядка (разд. 8.3). Наименьшее выпуклое подмножество, содержащее любое подмножество изображения A , называется *выпуклой оболочкой* A . Проблемы, связанные с представлением многоугольников в виде объединения выпуклых многоугольников, рассматриваются в работе [29] гл. 9.

Упражнение 2. а) Показать, что эллипс является выпуклой фигурой; б) показать, что дуга является выпуклой в том и только том случае, если она образует сегмент, ограниченный прямолинейной хордой.

Разбиение границ подмножеств и кривых

Существует несколько приемлемых методов выбора «точек расчленения», т. е. тех точек, в которых можно разрывать границу подмножества изображения. Если подмножество является кривой линией, то эти методы можно применить для разбиения не только границы, но и самого подмножества.

1. *Позиционные экстремумы.* В качестве точек расчленения можно выбирать (локально) самую верхнюю, самую нижнюю, крайнюю справа или слева точки границы подмножества [37].

2. *Точки перегиба.* Границу подмножества можно разделить на последовательные выпуклые и вогнутые участки («полуострова» и «заливы»), назначая границы участков в тех точках, где касательная к границе пересекает ее [38].

3. *Точки максимальной кривизны.* Интуитивно очевидно, что оптимальными точками расчленения границы подмножества при аппроксимации его многоугольником являются те точки, в которых направление границы претерпевает резкие изменения («углы») или велика кривизна; в частности, последняя достигает локального максимума ([8] гл. 3). Кроме того, можно расчленять границу подмножества в тех точках, где значения кривизны существенно изменяются [39].

Упражнение 3. Найти позиционные экстремумы, точки перегиба и максимальной кривизны для эллипса, определяемого уравнением $x^2/a^2 + y^2/b^2 = 1$.

4. *Свойства, характеризующие форму.* Можно ввести критерии для выделения из границы подмножества таких элементов, как «прямолинейные штрихи», «заливы», «отростки» и т. д. Количество или просто наличие подобных элементов также может служить признаком заданного подмножества изображения (гл. 9). Эти же проблемы обсуждаются в работах [10] гл. 5, а также в [4, 40—43].

5. *Сопоставление границ по форме.* В принципе можно прибегнуть к поиску участков границ, согласующихся с дугами заданной формы или с избранными участками других границ [44]. Чтобы выполнить подобное согласование, совершенно не обязательно оперировать в двумерном пространстве (разд. 6.1) — участки границы могут описываться функциями одной переменной (например, с помощью их натуральных уравнений, разд. 9.4); именно эти функции и будут согласовываться. Процедура согласования может выполняться последовательно путем сравнения точка за точкой заданной и искомой функций. При этом небольшими отклонениями можно пренебрегать, если прогноз показывает, что процесс заканчивается. Последовательность процедуры позволяет использовать «регулируемые» критерии сглаживания, зависящие от предыстории подгонки участка.

6. *Вершины.* Кривую можно расчленять в вершинах, т. е. тех точках, где дуги оканчиваются или пересекаются. В результате кривая делится на ряд «ветвей», соединяющих вершины между собой [45, 46].

Упражнение 4. Составить описания прописных букв шрифта, использованного в настоящей книге, на основе элементов типа «вершины, соединенные ветвями». В частности, ввести метки для вершин каждой буквы и составить список пар меток, соответствующих связанным ветвями вершинам. Модифицировать эти описания, выделив на ветвях элементы, которые могут быть использованы как признаки, характеризующие форму, с тем чтобы каждое описание однозначно определяло соответствующую букву.

Сканирование изображения

Существует еще один способ выделения из подмножества изображения зон различных типов — следует последовательно «пройтись» по строкам изображения, входящим в подмножество. При этом естественно в качестве точек расчленения выбирать те строки, в которых последовательности точек подмножества начинаются или заканчиваются, объединяются или разветвляются либо резко изменяется их длина или положение [47—52]. (Набор длин и положений подобной последовательности можно использовать для составления упрощенных описаний формы зон подмножества, например введением таких признаков, как постоянная ширина, сходящаяся, расходящаяся, горизонтальная, вертикальная, диагональная и т. д.).

В более общем смысле это означает поиск последовательностей элементов (или концов таких последовательностей), совпадающих с кривой определенной формы или расположенных вблизи нее; в частности, можно выделять последовательности, расположенные вдоль прямых, дуг окружности и т. д. Такой метод осмотра часто используется при обработке снимков, полученных в пузырьковой камере ([11—14] в главе 7). При использовании этого метода, так же как и в случае сопоставления границ изображения по форме, можно пренебрегать девиациями и разрывами исследуемого участка подмножества, если они не

слишком велики. Критерии выделения зон и в данном случае могут изменяться в зависимости от характера предшествующих этапов процесса.

Результат выделения зон с помощью последовательного осмотра строк изображения в какой-то степени зависит от ориентации заданного подмножества относительно строк. Один из способов минимизации этого влияния состоит в нормализации ориентации подмножества. Можно, как и при нормализации изображений (разд. 7.2), стандартизовать направление главной оси подмножества (т. е. двоичного изображения, являющегося характеристической функцией исходного изображения), проведя ее, например, перпендикулярно строкам. Можно осуществить нормализацию и по-другому: в подмножестве выделяется наименьший описанный прямоугольник и его ориентация изменяется таким образом, чтобы длинная сторона прямоугольника оказалась перпендикулярной строкам [53, 54].

Упражнение 5. Выделить «зоны» на тех же, что и в упражнении 4, прописных буквах, используя горизонтальные строки.

Преобразование координат

Очень часто преобразование координат дает возможность обнаружить в рисунке присутствие дуг определенной формы. Допустим, что нас интересует наличие в изображении прямых, образованных «единицами», с заданным угловым коэффициентом m . Чтобы разрешить этот вопрос, перейдем к новой системе координат, определяемой уравнениями

$$x' = x, \quad y' = y - mx.$$

Линия, определяемая уравнением $y = mx + b$, в новой системе координат принимает вид $y' = b$. Если теперь «натянуть» преобразованное изображение на ось y' (т. е. выполнить операцию суммирования или интегрирования вдоль оси x), в результате получается функция переменной y' , характеризующаяся острыми высокими выбросами в точках $y' = b$ соответственно каждому присутствующему на исходном изображении отрезку b на оси y линии с угловым коэффициентом m . Чем больше отноше-

ние числа «единичных» элементов, содержащихся в этих линиях, к числу таких элементов в других зонах исходного изображения, тем легче обнаруживаются выбросы. Отметим, что эти линии не обязательно должны быть непрерывны; метод позволяет выделять и пунктирные линии в тех случаях, когда им принадлежит достаточно большая часть «единичных» элементов всего изображения. Аналогичные процедуры можно ввести и для выделения других «элементарных форм», например дуги окружности [55].

Представляет интерес и еще одна процедура выделения прямых [56]; в ее основу положена операция «отображения точки в линию» Φ , определенная следующим образом: каждой точке (x_i, y_i) заданного подмножества (т. е. каждому «единичному» элементу предъявленного изображения f) соответствуют «единичные» элементы во всех точках прямой $y = y_i x + x_i$ преобразованного изображения $\Phi(f)$. Легко показать, что при коллинеарности точек $(x_1, y_1), \dots, (x_n, y_n)$ соответствующие прямые имеют общую точку, т. е. все они проходят через одну и ту же точку. Если точки принадлежат линии, почти параллельной оси x , прямые становятся почти параллельными и их общая точка уходит в бесконечность. Это затруднение можно преодолеть с помощью еще одной операции отображения точки в линию: переводящей точки (x_i, y_i) в линию $y = x_i x + y_i$. Таким образом, каждый «единичный» элемент, обнаруживаемый в процессе сканирования изображения f , запоминается в виде соответствующей линии «единичных» элементов, вводимой в преобразованное изображение $\Phi(f)$ (преобразованное изображение $\Phi(f)$ можно рассматривать как матрицу, составленную из счетчиков); присутствие большого количества коллинеарных «единичных» элементов в изображении f приводит к появлению в преобразованном изображении точки, в которой функция Φ достигает большой величины. Линии изображения f , образованные «единичными» элементами, можно, следовательно, обнаружить с помощью квантования преобразованного изображения $\Phi(f)$. Если изображение f содержит много «единичных» элементов, может оказаться целесообразным воспользоваться «выборкой», например ограничиться осмотром изображения вдоль нескольких линий

различной ориентации. Если изображение f содержит линию, образованную «единицами», то по крайней мере какая-то часть сканирующих проходов ее не минует и, следовательно, для обнаружения такой линии не обязательно переводить каждый «единичный» элемент изображения f в линию преобразованного изображения $\Phi(f)$.

ЛИТЕРАТУРА

1. Rosenfeld A., Fried C., Orton J. N., Automatic cloud interpretation, *Photogrammetric Engrg*, 31, 991 (Nov. 1965).
2. Narasimhan R., Fornango J. P., Some further experiments in the parallel processing of pictures, *IEEE Trans. Electronic Computers*, EC-13, 748 (Dec. 1964).
3. Rosenfeld A., Connectivity in digital pictures, *J. ACM*, 16 (1969).
4. Greanias E. C. et al., The recognition of handwritten numerals by contour analysis, *IBM J. Res. Develop.*, 7, 14 (Jan. 1963); русский перевод: Гриниас, Мигер, Норман, Изингер, Распознавание рукописных цифр с помощью анализа контуров, *Зарубежная радиоэлектроника*, № 2, 1964.
5. Bradshaw J. A., Letter recognition using a captive scan, *IEEE Trans. Electronic Computers*, EC-12, 26 (Febr. 1963).
6. Sezaki N., Katagiri H., Character recognition by follow method, *Proc. IEEE*, 53, 510 (May 1965); русский перевод: Сэзаки, Катагирри, Распознавание знаков методом слежения, *ТИИЭР*, № 5, 589 (1965).
7. Sezaki N., Katagiri N., Kaneko T., Pattern reproduction by follow method, *Proc. IEEE*, 53, 1656 (Oct. 1965); русский перевод: Сэзаки, Катагирри, Канеко, Воспроизведение образов методом слежения, *ТИИЭР*, № 10, 1857 (1965).
8. Ledley R. S., Jacobsen J. D., Belson M., BUG-SYS: a programming system for picture processing — not for debugging, *Comm. ACM*, 9, 79 (Febr. 1966).
9. Ledley R. S. et al., Pattern recognition studies in the biomedical sciences, *SJCC*, 411 (May 1966).
10. Rintala W. M., Hsu C. C., A feature-detection program for patterns with overlapping cells, *IEEE Trans. Systems Science and Cybernetics*, SSC-4, 16 (March 1968).
11. Kirsch R. A. et al., Experiments in processing pictorial information with a digital computer, *EJCC*, 221 (Dec. 1957).
12. Weston P., Photocell field counts random objects, *Electronics*, 34, 46 (Sept. 22, 1961); русский перевод: Вестон, Подсчет числа объектов произвольной формы и расположения с помощью сетки фотоэлементов, *Электроника*, 34, № 38, 25 (1961).
13. Babcock M. L., Some physiology of automata, *WJCC*, 291 (May 1961).

14. Ledley R. S. et al., FIDAC: Film input to digital automatic computer and associated syntax-directed pattern-recognition programming system, OEOIP, p. 591.
15. Nuttal T. C., Apparatus for counting objects, пат. США. 2803406 (Aug. 20, 1957).
16. Shmirat M., Position of point relative to polygon, Algorithm 112, *Comm. ACM*, 5, 434 (Aug. 1962); Hacker R., Certification of Algorithm 112, Position of point relative to polygon, *Comm. ACM*, 5, 606 (Dec. 1962).
17. Loomis R. G., Boundary networks, *Comm. ACM*, 8, 44 (Jan. 1965).
18. Nordbeck S., Rystedt B., Computer cartography—point-in-polygon programs, *Nordisk Tidskr. Informations-Behandling (BIT)*, 7, 30 (1967).
19. Preston K., Jr., The CELLSCAN system, a leucocyte pattern analyzer, *WJCC*, 173 (May 1961).
20. Izzo N. F., Coles W., Blood cell scanner identifies rare cells, *Electronics*, 35, 52 (April 27, 1962); русский перевод: Иzzo, Коулз, Опознавание форменных элементов крови с помощью сканирующей системы, *Электроника*, 35, № 17, 21 (1962).
21. Saraga P., Woollons D. J., The design of operations for pattern processing, CPR, p. 106.
22. Hilditch C. J., Linear skeletons from square cupboards, «Machine Intelligence 4» (B. Meltzer and D. Michie, eds.), American Elsevier, New York, 1969, p. 403.
23. Sneath P. H. A., A method for curve seeking from scattered points, *Comput. J.*, 8, 383 (Jan. 1966).
24. Blum H., An associative machine for dealing with the visual field and some of its biological implications, BPSS, p. 244, русский перевод: Блум Г., Ассоциативная машина, опознающая зрительные изображения, в сб. «Проблемы бионики», изд-во «Мир», М., 1965, стр. 342.
25. Blum H., A Transformation for extracting new descriptors of shape, MPSVF, p. 362.
26. Rosenfeld A., Pfaltz J. L., Sequential operations in digital picture processing, *J. ACM*, 13, 471 (Oct. 1966).
27. Pfaltz J. L., Rosenfeld A., Computer representation of planar regions by their skeletons, *Comm. ACM*, 10, 119 (Febr. 1967).
28. Calabi L., Hartnett W. E., Shape recognition, prairie fires, convex deficiencies and skeletons, *Amer. Math. Monthly*, 75, 335 (Apr. 1968).
29. Rutowitz D., Data structures for operations on digital images, PPR, p. 105.
30. Philbrick O., Shape description with the medial axis transformation, PPR, p. 395.
31. Rosenfeld A., Pfaltz J. L., Distance functions on digital pictures, *Pattern Recognition*, 1, 33 (July 1968).
32. Kasvand T., Histogramming of nerve fiber cross-sections and water droplets by methods of pattern recognition, CPR, p. 315.

33. Montanari U., A method of obtaining skeletons using a quasi-Euclidean distance, *J. ACM*, 15, 600 (Oct. 1968).
34. Glucksmann H. A., A parapropagation pattern classifier, *IEEE Trans. Electronic Computers*, EC-14, 434 (June 1965).
35. Glucksmann H. A., Classification of mixed-font alphabets by characteristic loci, Proc. IEEE Comput. Conf. Sept. 1967, p. 138.
36. Stevens M. E., Abstract shape recognition by machine, *EJCC*, 332 (Dec. 1961).
37. Mason S. J., Clemens J. K., Character recognition in an experimental reading machine for the blind in «Recognizing Patterns: Studies in Living and Automatic Systems» (Kollers P. A., Eden M., eds.), M.I.T. Press, Cambridge, Massachusetts, 1968, p. 156; русский перевод: Мейсон С., Клеменс Дж., Распознавание букв в экспериментальной читающей машине для слепых, в сб. «Распознавание образов. Исследование живых и автоматических распознающих систем», изд-во «Мир», М., 1970.
38. Freeman H., On the classification of line-drawing data, MPSVF, p. 408.
39. Symons M., A new self-organizing pattern recognition system, CPR, p. 11.
40. Spinrad R. J., Machine recognition of hand printing, *Information and Control*, 8, 124 (Apr. 1965).
41. Munson J. H., The recognition of hand-printed text, PR², p. 115.
42. Deutsch E. S., Preprocessing for character recognition, CPR, p. 179.
43. Genchi H. et al., Recognition of handwritten numerical characters for automatic letter sorting, *Proc. IEEE*, 56, 1292 (Aug. 1968); русский перевод: Генчи и др., Распознавание рукописных букв при автоматической сортировке писем, *ТИИЭР*, 56, № 8, 34 (1968).
44. Freeman H., Garder L., A pictorial jigsaw puzzles: the computer solution of a problem in pattern recognition, *IEEE Trans. Electronic Computers*, EC-13, 118 (Apr. 1964).
45. Sherman H., A quasi-topological method for the recognition of line patterns, Proc. Intern. Confer. Inform. Process, UNESCO, Paris, 1959; UNESCO, Paris, 1960, p. 232; русский перевод: Шерман Г., Квазитопологический метод распознавания линейных изображений, в сб. «Автоматический анализ сложных изображений», изд-во «Мир», М., 1969, стр. 11.
46. Mippenham M. J., Handwritten character recognition employing topology, cross correlation, and decision theory, *IEEE Trans. Systems Science and Cybernetics*, SSC-2, 86 (Dec. 1966).
47. Grimsdale R. L. et al., A system for the automatic recognition of patterns, *Proc. IEE*, 106B, 210 (1959); reprinted in PR, p. 317.
48. Grimsdale R. L., Bullingham J. M., Character recognition by digital computer using a special flying-spot scanner *Comput. J.*, 4, 129 (July 1961).

49. Sublette I. H., Tults J., Character recognition by digital feature detection, *RCA Rev.*, 23, 60 (March 1962).
50. Uyehara G. U., A stream-following technique for use in character recognition, *IEEE Intern. Conv. Record*, Part 4, March 1963, p. 64.
51. Perotto P. G., A new method for automatic character recognition, *IEEE Trans. Electronic Computers*, EC-12, 521 (Oct. 1963).
52. Hocking K. H., Thompson J., A feature detection method for optical character recognition, *CPR*, p. 271.
53. Neurath P. W. et al., Human chromosome analysis by computer — an optical pattern recognition problem, *Ann. New York Acad. Sci.*, 128, 1013 (Jan. 1966).
54. Rutovitz D., Machines to classify chromosomes? in «Human Radiation Cytogenetics» (Evans H. J., Court Brown W. M., McLean A. S., eds.), North-Holland, Amsterdam, 1967, p. 58.
55. Bazin M. J., Benoit J. W., Off-line global approach to pattern recognition for bubble chamber pictures, *IEEE Trans. Nuclear Science*, NS-12, 291 (Aug. 1965).
56. Hough P. V. C., Method and means for recognizing complex patterns, пат. США 3069654 (Dec. 18, 1962).
- 57*. Поляков В. Г., Переверзев-Орлов В. С., Электронные системы следящей развертки, изд-во «Энергия», М., 1968.
- 58*. Горьян И. С., Цуккерман И. И., Распознавание изображений по их статистическим свойствам. *Техника кино и телевидения*, № 2, 57 (1966).
- 59*. Горьян И. С., Дубов А. С., Ильин Б. М., Цуккерман И. И., Об автоматическом распознавании форм облачности, *Метеорология и гидрология*, № 9, 22 (1970).

Геометрические свойства подмножества изображения

Как было установлено в гл. 7, некоторые признаки изображения имеют естественные определения только в том случае, если на изображении можно четко выделить обособленное подмножество, иначе говоря, некоторый объект. Выделив такое подмножество, можно вычислить значения различных признаков изображения, но только на этом подмножестве (например, положив равным нулю яркость изображения вне подмножества). Эти признаки можно рассматривать как признаки подмножества. Новый важный класс признаков подмножества, который исследуется в этой главе, связан с *геометрией* объекта, т. е. его связностью, размером, формой и т. д., и не зависит от его яркости. Под признаком обычно будет подразумеваться некоторое вещественное число, хотя кратко будут рассмотрены признаки, являющиеся функциями одной переменной.

9.1. Топологические признаки

Геометрические признаки, инвариантные относительно «эластичных» деформаций изображения, называются *топологическими* признаками. Заметим, что из этого определения следует, что «линейчатость» не является топологическим признаком, поскольку на реальном изображении линии имеют конечную ширину и поэтому могут быть превращены в «широкие» фигуры. Однако если мы рассмотрим идеализированный рисунок, линии которого могут считаться линиями в математическом смысле, или если будем применять только те деформации, которые сохраняют линейчатость, тогда такие признаки рисунков, как порядки их узлов (т. е. число линий, сходящихся в узле) (разд. 8.4), станут топологическими.

Связность и порядок связности

Наиболее известное топологическое свойство подмножества изображения — *связность*, т. е. число его связанных компонент. Очевидно, любой из методов определения компонент, описанных в разд. 8.2, можно использовать для их подсчета. Например, в методе «распространения» число единиц, появившихся в процессе сканирования, совпадает с числом компонент; аналогично в методе «обегания границы» необходимо только подсчитать, сколько раз обнаруживается внешняя граница. Специальные методы, описанные ниже, также позволяют подсчитывать количество односвязных объектов.

Пусть S — связное подмножество изображения; тогда число компонент его дополнения \bar{S} называется порядком связности S (заметим, что это число на единицу больше числа «дырок» в S ; в частности, S является односвязным тогда и только тогда, когда его порядок связности равен 1).

Род

Число компонент подмножества S минус число дырок в S называется родом (или числом Эйлера) подмножества S . Обозначим это число G_S .

Упражнение 1. Выписать в таблицу связность, порядок связности и род для каждой гласной буквы используемого в этой книге шрифта.

Теорема. Для любого подмножества S дискретного изображения имеем

$$G_S = I_S - II_S + IV_S,$$

где I_S — число элементов в S , II_S — число горизонтальных или вертикальных пар смежных элементов в S и IV_S — число входящих в S квадратов 2×2 , образованных четырьмя смежными элементами.

Упражнение 2. Доказать эту теорему индукцией по I_S .

Эта теорема в действительности аналогична более известной теореме о многогранниках $G = V - E + F$, где V , E , F — числа вершин, ребер и граней соответственно.

Отметим, что, согласно этой теореме, род является линейным пороговым признаком по отношению к набору локальных признаков порядка ≤ 4 (разд. 8.3). В действительности можно показать, что функции от G_S обладают только топологическими признаками ограниченного порядка (напомним, что сама связность имеет порядок, который растет с увеличением размера изображения).

Докажем эту теорему методом непосредственного вычисления G_S . Пусть S будет множеством единиц в двоичном изображении (a_{ij}) , и пусть

$$b_{ij} = a_{ij}(a_{i-1,j} + a_{i+1,j} + a_{i,j-1} + a_{i,j+1})/2,$$

$$c_{ij} = a_{ij}(a_{i-1,j-1}a_{i-1,j}a_{i,j-1} + a_{i-1,j}a_{i-1,j+1} \times$$

$$\times a_{i,j+1} + a_{i,j-1}a_{i+1,j-1}a_{i+1,j} + a_{i,j+1}a_{i+1,j}a_{i+1,j+1})/4.$$

Тогда, очевидно, $\sum a_{ij} = I_S$, $\sum b_{ij} = II_S$, $\sum c_{ij} = IV_S$, где суммы берутся по всем индексам (i, j) данного изображения.

Если известно, что каждая связная компонента S односвязна, то G_S оказывается равным связности S , так что, вычислив G_S , можно определить число составляющих S (разд. 8.2). Изложим два других метода подсчета односвязных компонент, основанных на методах, описанных в разд. 8.2.

1. Развертка. Подсчитаем количество серий, состоящих из единиц, в первой строке изображения. Для второй и последующих строк будем к содержимому счетчика прибавлять единицу, встретив серию, которая не перекрывается какими-либо сериями предыдущей строки, и результат счета оставляем без изменения, если серия в данной строке перекрывается одной серией предшествующей строки, и вычитаем k единиц из содержимого счетчика, если серия в данной строке перекрывается $k + 1$ сериями предшествующей строки ($k = 1, 2, \dots$). Когда все строки будут пройдены, получившееся окончательно число и будет равно числу компонент. Заметим, что, если компонента имеет дырку, серия будет расщепляться на две или более частей, которые затем опять сойдутся, после чего вычтем (хотя это и неправильно) единицу или большее число из содержимого счетчика. Если этот метод подсчета применить к изображению, в котором

компоненты не обязательно являются односвязными, то тем самым можно будет вычислить род.

Интересно отметить, что если этим методом необходимо подсчитать не только число компонент, но также и определить их площади, то можно получить правильный результат, несмотря на то что компоненты не являются односвязными. Можно поступить, например, следующим образом.

Для каждой серии единиц в первой строке будем использовать свой счетчик для подсчета длины серии. Для второй и последующих строк подключим новые счетчики для каждой серии, не перекрытой какой-либо серией предыдущей строки. Если серия перекрывается лишь одной серией предыдущей строки, то ее длина добавляется к содержимому соответствующего счетчика; если же серия перекрывается двумя или более сериями предыдущей строки, то содержимое их счетчиков суммируется, переносится в новый счетчик и к сумме добавляется длина новой серии. Когда все строки будут проверены, число счетчиков, имеющих ненулевое содержание, будет равно числу компонент, а величина, хранимая в каждом счетчике, окажется равной площади соответствующей компоненты.

2. *Сжатие*. Алгоритм сжатия, описанный в разд. 8.2, позволяет представить односвязные компоненты в виде одного элемента, в то время как компоненты, имеющие дырку, можно преобразовать лишь в замкнутую кривую.

Таким образом, если после осуществления сжатия обнаруживается, что в получившемся изображении никакие две единицы не соединяются, то все компоненты в первоначальном изображении являются односвязными, а число единиц в преобразованном изображении равно числу этих компонент.

9.2. Метрические признаки. Размеры

Рассмотрим такие признаки изображений, которые связаны с измерением площадей, расстояний, углов и т.д.; подобные признаки называются *метрическими*.

Площадь и периметр

Площадь подмножества дискретного изображения в точности равна числу элементов в этом подмножестве; это число можно получить подсчетом единиц (или суммированием всех элементов) характеристической функции подмножества. Как уже упоминалось выше, площади связанных компонент подмножества можно вычислить с помощью модифицированного метода развертки, используемого для подсчета компонент. Площади различных подмножеств, связанных с данным подмножеством (например, множество точек, находящихся на заданном расстоянии от него, выпуклая оболочка подмножества, его «тени» и т.д.), можно также отнести к характеристикам основного подмножества.

Для определения периметра подмножества можно пройти вдоль его границы и измерить общий пройденный путь. Периметр подмножества в дискретных изображениях можно вычислить, используя тот факт, что он приблизительно пропорционален площади его окаймления (или, точнее, среднему значению площадей его окаймления и окаймления дополнения). Длина дуги подмножества, образованного линиями, измеряется аналогично. Алгоритмы измерения площади, периметра и т.д. для рисунков, представленных цепным кодом, рассмотрены в работе [1].

В общем случае периметр и длина дуги непрерывных изображений не могут быть точно измерены; однако можно аппроксимировать дугу некоторым вписанным многоугольником. Очевидно, что длина дуги, вычисляемая таким способом, возрастает (или по меньшей мере не убывает) с убыванием длины сторон этого многоугольника. Таким образом, можно говорить о « ϵ -длине» дуги [2], где ϵ — длина наименьшей стороны вписанного многоугольника, используемого для аппроксимации длины дуги. Эмпирически было установлено [3], что ϵ -длину какой-либо дуги часто можно аппроксимировать показательной функцией от ϵ вида $L_\epsilon = Ke^{-d\epsilon}$, где $d \geq 0$; $d = 0$ для прямых линий, и $d \approx 0,25$ для очень «извилистых» кривых. Следовательно, d можно использовать в качестве меры прямолинейности. Для этой же цели можно применять отношение длины дуги к длине ее хорды.

Упражнение 1. Как зависит ϵ -длина окружности от ϵ ?

В следующем разделе говорится об интегрально-геометрических методах измерения площади и периметра посредством анализа пересечений семейства линий с данным подмножеством. Показывается, что периметр выпуклого подмножества равен произведению $\pi/2$ на среднюю длину проекции этого подмножества на прямую линию.

Признаки, связанные с расстоянием

Высоту или ширину подмножества изображения можно очень просто определить, вычисляя расстояния между самой высокой и самой низкой горизонтальными линиями или между самой правой и самой левой вертикальными линиями, касающимися подмножества. Аналогично можно измерить протяженность подмножества в данном направлении; использование протяженности для опознавания фотоснимков рассмотрено в работе [4].

Упражнение 2. Чему равна средняя протяженность эллипса

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1?$$

Определение длины и ширины вытянутого подмножества представляет менее тривиальную задачу. В том случае, если множество сравнительно прямое, целесообразно применять следующие методы [53, 54] гл. 8:

1. Обозначим через I главную ось подмножества; определим длину подмножества как его протяженность в направлении, параллельном оси I , а его ширину как протяженность в направлении, перпендикулярном I .

2. Обозначим через R прямоугольник наименьшей площади, который можно описать вокруг данного подмножества, и определим длину и ширину подмножества как длины большей и меньшей сторон прямоугольника R .

Более общее определение длины и ширины, которое применимо также и к искривленным вытянутым подмножествам S , можно сформулировать в следующем виде. Пусть S_d обозначает площадь множества точек S , находящихся на расстоянии d или меньшем от \bar{S} . Очевидно, что

S_d является убывающей функцией d и должна обратиться в нуль, как только d достигнет величины, равной половине диаметра S . Если S имеет почти постоянную ширину W , то S_d скачком обратится в нуль, как только d достигнет величины $W/2$. Таким образом, ширину S можно было бы определить как двойную величину d , при которой S_d скачкообразно падает до значения, близкого к нулю; длина S определится тогда делением площади на ширину.

Упражнение 3. Определить S_d для эллипса.

Точечное изображение определяется с точностью до конгруэнтных преобразований¹ перечислением расстояний между всеми парами точек; фактически, если существует более трех точек, нет необходимости для определения изображения перечислять все эти расстояния. Эквивалентная формулировка: многоугольник определяется с точностью до конгруэнтных преобразований перечислением длин его сторон и диагоналей (возможно, не всех). Если подмножество изображения аппроксимируется многоугольником, в качестве признаков подмножества можно использовать, следовательно, расстояния между точками или их комбинации [5].

Упражнение 4. Сколько диагоналей в дополнение к сторонам должно быть указано для определения n -угольника?

9.3. Метрические признаки. Форма

Признаки, определяемые семействами подмножеств

Пусть \mathcal{A} — множество подмножеств плоскости (например, множество всех линий, множество всех горизонтальных прямых, множество всех групп из трех точек и т. д.). Пусть также \mathfrak{F} — какой-либо геометрический признак. Вычислим для каждого элемента A в \mathcal{A} величину \mathfrak{F} для пересечения $S \cap A$ элемента A с данным подмножеством S . Тогда можно определить новые геометрические признаки S , используя различные статистики частот-

¹ Конгруэнтные преобразования — перенос и поворот — Прим. ред.

ного распределения величин $\mathfrak{F}(S \cap A)$. Этот фундаментальный принцип лежит в основе применения интегральной геометрии для определения геометрических свойств.

Пусть, например, \mathcal{A} — множество всех прямых, и пусть, в частности, $L(p, \theta)$ — прямая, отстоящая от начала координат на расстояние p , измеряемое вдоль перпендикуляра, расположенного под углом θ к оси x . Если вычислим среднее значение $\mathfrak{F}[S \cap L(p, \theta)]$, т. е. $\int \int \mathfrak{F}(S \cap L(p, \theta)) dp d\theta$, где $\mathfrak{F}(S \cap L(p, \theta))$ — длина пересечения $S \cap L(p, \theta)$, то, как нетрудно показать, получим площадь S , умноженную на π . Если \mathfrak{F} — связность пересечения $S \cap L(p, \theta)$, то получим периметр S . Полагая $\mathfrak{F} = 1$ (если $L(p, \theta)$ пересекает S) и $\mathfrak{F} = 0$ (если $L(p, \theta)$ не пересекает S), получим периметр выпуклой оболочки S [6, 7].

Упражнение 1. Проверить эти утверждения в случае, когда S — эллипс.

Другой полезный класс свойств получим, полагая, что \mathcal{A} — множество групп из n точек [8]. Например, при $n = 2$ положим $\mathfrak{F} = 1$ (если обе точки лежат в S) и $\mathfrak{F} = 0$ (в противном случае). Тогда среднее значение \mathfrak{F} , вычисленное для всех пар точек с заданным расстоянием друг от друга, представит как раз свойство, описанное в гл. 7.

Аналогично если рассмотреть тройки коллинеарных точек, то можно получить некоторую меру для определения степени выпуклости S , используя величину \mathfrak{F} , позволяющую установить, лежат ли две крайние точки в S ; при этом промежуточная точка не должна лежать в S .

Локальные признаки

Критерии (разд. 8.4) для разбиения границы подмножества на части — экстремальные точки, точки перегиба, максимум кривизны и т. д. — можно также использовать для определения признаков данного подмножества [9—12].

Комплексные свойства

При изучении визуального восприятия было установлено [13—15], что сложность объекта связана со следующими его свойствами: 1) числом углов (т. е. максимумов

кривизны на границе объекта) и изменением размеров этих углов; 2) числом P^2/A — квадратом периметра объекта, деленным на его площадь (в числителе берется квадрат периметра для того, чтобы частное не зависело от размеров объекта); заметим, что число P^2/A достигает наименьшего значения (равного 4π) для круга. Эти свойства не являются, очевидно, полностью независимыми, но они также и не перекрывают целиком друг друга. Если объект «звездообразный» (т. е. каждую его граничную точку можно «увидеть» из некоторой заданной внутренней точки P), то его можно полностью описать однозначной функцией $r = r(\theta)$, представляющей уравнение границы в полярных координатах с началом в точке P . В этом случае вариацию $r(\theta)$ [т. е. $\int (r(\theta) - \bar{r})^2 d\theta$, где $\bar{r} = \int r(\theta) d\theta$ — среднее значение $r(\theta)$] можно принять в качестве меры сложности [16]; заметим, что эта вариация принимает наименьшее значение (равное нулю) также для круга.

Упражнение 2. Показать, что выпуклое подмножество является звездообразным относительно любой из его точек.

Упражнение 3. Вычислить P^2/A и вариацию $r(\theta)$ для эллипса

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Применение рассмотренных выше мер сложности целесообразно в том случае, если данное подмножество является односвязным. В противоположном крайнем случае, когда подмножество — фигура, состоящая из многих разрозненных точек или линий, в качестве меры сложности изображения можно принять его случайность. Для точечных рисунков случайность можно определить, вычисляя вариацию числа точек в данной окрестности каждой точки или вариации расстояний от каждой точки до заданного числа ближайших соседних точек [17, 18].

Упражнение 4. Вычислить для случайного изображения (упр. 2 разд. 6.1): а) вариацию числа элементов в окрестности каждого элемента в виде функции от размера этой окрестности; б) вариацию расстояния от каждого элемента до ближайшего к нему соседнего элемента.

Симметрия, очевидно, также связана со сложностью. Для любого данного подмножества изображения его асимметрию по отношению к фиксированной линии или точке можно определить различными способами; например, в качестве меры асимметрии можно взять интеграл $\iint |f(x, y) - f(x', y')| dx dy$, где $dx dy$ — характеристическая функция подмножества, а (x', y') — точка, симметричная точке (x, y) относительно фиксированной линии или точки. Таким образом, можно определить ось (или центр) симметрии как линию (или точку), по отношению к которой асимметрия минимальна.

Упражнение 5. Для того же изображения (упр. 4) вычислить асимметрию относительно центра изображения и относительно вертикальной линии, проходящей через центр изображения.

Упражнение 6. Вычислить асимметрию эллипса $x^2/a^2 + y^2/b^2 = 1$ относительно прямой $y = mx$. То же сделать для множества точек (k_i, k_j) , где $1 \leq i \leq n$ и $1 \leq j \leq n$.

9.4. Функции подмножеств изображений

Уравнения

Любое подмножество изображения полностью определяется указанием его границы, которую можно представить состоящей из множества направленных кривых¹. Следовательно, задав уравнения этих кривых, можно описать это подмножество. Например (как установлено в разд. 9.3), если подмножество «звездообразно» относительно некоторой точки P , то оно может быть полностью описано перечислением расстояний от точки P до каждой точки границы подмножества — другими словами, уравнением его границы в полярной системе координат с началом в точке P . Описание подмножества можно сделать

¹ Если подмножество односвязно, то требуется указать лишь одну кривую, но если оно имеет несколько связных компонент или если оно имеет дырки, то потребуются описание добавочных кривых. Направления кривых выбираются так, чтобы при движении по ним подмножество оставалось слева от каждой кривой.

независимо как от ориентации, например с помощью разложения уравнения в ряд Фурье и последующего вычисления модуля, так и от масштаба, например разделив обе части уравнения на величину диаметра этого подмножества. Отсюда, в частности, следует, что можно сравнить два множества, сравнивая уравнения их граничных кривых. Как указано в разд. 8.3, для этого требуется одномерная взаимная корреляция, а не двумерная, необходимая для сравнения самих подмножеств. Если же уравнения нормализованы (как только что было описано), то эта задача может быть сведена к задаче простого подбора эталона.

Одним из важных методов описания границы подмножества является метод натурального уравнения, в котором кривизна представляется в виде функции от длины дуги (измеренной от начальной точки). Это уравнение полностью определяет границу с точностью до конгруэнтных преобразований. Оно также может быть приведено к виду, не зависящему от размера и от начальной точки. (Один из способов исключения зависимости уравнения от начальной точки состоит в том, что отсчет идет от некоторой характерной, легко опознаваемой точки, например от точки с наибольшей кривизной.) Кроме того, можно использовать наклон, а не кривизну граничной кривой в виде функции от длины дуги; в этом случае, разумеется, существует зависимость от ориентации. Еще одно уравнение можно вывести, определив зависимость от длины дуги угла между радиусом-вектором (т. е. прямой, идущей от начала координат) граничной точки и касательной к граничной кривой в этой точке; можно показать, что это уравнение не зависит от положения, ориентации и масштаба [19].

Упражнение 1. Какой вид имеет натуральное уравнение для эллипса, начинающегося от точки пересечения кривой с главной полуосью?

Цепное кодирование дуги в виде последовательности шагов данной длины вдоль одного из заданных наборов направлений можно рассматривать как дискретный вариант натурального уравнения, поскольку оно дает квантованный наклон в виде функции от квантованной длины

дуги. Более того, если вычислить разности между двумя последовательными наклонами, можно получить квантованную кривизну (т. е. скорость изменения наклона) в виде функции от длины дуги [20—22].

Распределение частот

Распределения частот значений признаков подмножеств дают ряд полезных функций одного параметра. Двумя примерами таких признаков являются: 1) распределение наклонов или кривизны границ подмножества [20,21] и 2) распределение расстояний между элементами подмножества (либо между элементами его границы) [23].

Функции, определяемые семействами подмножества A_λ

Пусть \mathcal{A} — однопараметрическое семейство подмножеств данной плоскости (например, множество всех линий с заданным наклоном). Тогда для любого свойства \mathfrak{F} подмножества S изображение $\mathfrak{F}(S \cap A_\lambda)$ дает нам функцию от λ для любого заданного подмножества S . Широкий класс функций такого типа рассмотрен в работах [24—29], а также в работе [29] гл. 7.

Аналогично пусть \mathcal{A} — двух- (или более) параметрическое семейство подмножеств (например, множество всех линий), так что $\mathfrak{F}(S \cap A)$ (где $A \in \mathcal{A}$) является функцией двух или более параметров; если вычислим некоторую статистическую величину по всем параметрам, кроме одного, то получим в результате функцию одного параметра. Пусть, например, \mathcal{A} — множество всех линий; усредним $\mathfrak{F}(S \cap A)$ по всем линиям, имеющим фиксированный наклон, в результате получим функцию наклона ([30], а также [34—36] гл. 8.)

Упражнение 2. Пусть \mathcal{A} — множество всех линий, пересекающих эллипс $x^2/a^2 + y^2/b^2 \leq 1$. Какова средняя длина отрезка, лежащего внутри эллипса, для всех линий из \mathcal{A} , имеющих наклон m ?

ЛИТЕРАТУРА

1. Freeman H., Techniques for the digital computer analysis of chain-encoded arbitrary plane curves, Proc. Nat. Electron. Confer., Oct. 1961, p. 421.

2. Perkal J., On the ε -length, *Bull. Acad. Polon. Sci., Classe III*, 4, 1956, p. 399.
3. Mandelbrot B., How long is the coast of Britain? Statistical self-similarity and fractional dimension, *Science*, 156, 636 (May 5, 1967).
4. Marill T., Green D. M., Statistical recognition functions and the design of pattern recognizers, *IRE Trans. Electronic Computers*, EC-9, 472 (Dec. 1960).
5. Bunge W., *Theoretical Geography*, Gleerup Publishers, Lund, Sweden, 1962; русский перевод: Бунге В., *Теоретическая география*, М., изд-во «Прогресс», 1967.
6. Steinhilber H., Length, shape and area, *Colloq. Math.*, 3, 1(1954).
7. Новиков А., «Распознавание образов с помощью интегральной геометрии» в сб. «Принципы самоорганизации», изд-во «Мир», М., 1966, стр. 428.
8. Frisch H. L., Julesz B., Figure-ground perception and random geometry, *Perception Psychophys.*, 1, 389 (1966).
9. Sprick W., Ganzhorn K., An analogous method for pattern recognition by following the boundary, Proc. Intern. Confer. Inform. Process., UNESCO, Paris, 1959, p. 238.
10. Kazmierczak H., The potential field as an aid to character recognition, Proc. Intern. Confer. Inform. Process., UNESCO, Paris, 1960, p. 244.
11. Harmon L. D., A line-drawing pattern recognizer, *WJCC*, 351 (May 1960).
12. Harmon L. D., Line-drawing pattern recognizer, *Electronics*, 33, 39 (Sept. 2, 1960).
13. Attneave F., Physical determinants of the judged complexity of shapes, *J. Exptl. Psychol.*, 53, 221 (1957).
14. Arnoult M. D., Prediction of perceptual responses from structural characteristics of the stimulus, *Perceptual Motor Skills*, 11, 261 (1960).
15. Stenson H. H., The physical structure of random forms and their judged complexity, *Perception Psychophys.*, 1, 303 (1966).
16. Boyce R. R., Clark W. A. V., The concept of shape in geography, *Geograph. Rev.*, 54, 561 (1964).
17. Dacey M. F., Tung T. H., The identification of randomness in point patterns, *J. Reg. Sci.*, 4, 83 (1962).
18. Dacey M. F., Description of line patterns, in «Quantitative Geography» (W. L. Garrison, D. F. Marble, eds), Dept. of Geography, Northwestern Univ., Evanston, Illinois, 1967, p. 279.
19. Brouillette J. W., Johnson C. W., Pattern recognition, Proc. Military Electron., 179 (June 1960).
20. Freeman H., A technique for the classification and recognition of geometric patterns, Proc. 3rd Intern. Congr. Cybernetics (Sept. 1961), Association Internationale de Cybernetique, Namur, Belgium, 1965, p. 348.
21. Freeman H., On the digital computer classification of geo-

- metrical line patterns, Proc. Nat. Electron. Confer., Oct. 1962, p. 312.
22. Sugiura T., Higashiwatoko T., A method for the recognition of Japanese hiragana characters, *IEEE Trans. Information Theory*, **IT-14**, 226 (March 1968).
 23. Deutsch J. A., A theory of shape recognition, *Brit. J. Psychol* **46**, 30 (1955); reprinted in PR, p. 117.
 24. Doyle W., Recognition of sloppy, hand-printed characters, *WJCC*, 133 (May 1960).
 25. Nadler M., Une systeme analogique-digital pour la reconnaissance de caracteres, in «Information Processing 1962» (C. M. Poplewell, ed.), North-Holland, Amsterdam, 1963, p. 456.
 26. Akers S. B., Rutter B. H., The use of threshold logic in character recognition, *Proc. IEEE*, **52**, 931 (Aug. 1964); русский перевод: Акерс, мл., Руттер, Применение пороговой логики для распознавания образов, *ТИИЭР*, № 8, 995 (Август 1964).
 27. Rubio J. E., The clustering and recognition of patterns, *Intern. J. Control*, **4**, 459 (1966).
 28. Pavlidis T., Computer recognition of figures through decomposition, *Information and Control*, **12**, 526 (May—June 1968).
 29. Pavlidis T., On the syntactic analysis of figures, Proc. ACM Nat. Confer., Aug. 1968, p. 183.
 30. Latham J. P., Methodology for an instrumented geographic analysis, *Ann. Assoc. Amer. Geographers*, **53**, 194 (June 1963); см. также Latham J. P., Witmer R. E., Comparative waveform analysis of multisensor imagery, *Photogrammetric Engrg.*, **33**, 779 (July 1967).

Описание изображений и язык изображений

Как было установлено в гл. 7, общая проблема распознавания изображений включает в себя не только задачу классификации, но также и задачу описания заданных изображений. В большинстве случаев описание изображения относится к признакам изображения (определяемым как для изображения в целом, так и для его подмножеств), именно к геометрическим признакам подмножества, и к таким отношениям между подмножествами, которые характеризуются выражениями «сверху», «слева», «вблизи», «между», «внутри», «часть чего-либо», «больше чем», «темнее чем» и т. д.

Как признак подмножества является вещественной функцией, определенной на множестве всех подмножеств, так и бинарное отношение между подмножествами является вещественной функцией, определенной на множестве всех пар подмножеств. Аналогично для тройных отношений более высокого порядка используется множество групп из трех подмножеств, однако на практике эти случаи редки. Многие важные свойства и отношения (включая и те, которые были рассмотрены выше, если предположить, что они были точно определены) являются функциями, принимающими для данных подмножеств (или пар подмножеств) лишь два значения — истина или ложь; эти функции представляют частный случай вещественных функций (признаки и отношения, в определении которых участвуют функции одного или более параметров, здесь не рассматриваются).

Если задана совокупность признаков и отношений, в принципе уже становится возможным построить словесное описание изображения в виде ответов на вопросы относительно изображения и определить, обладает ли изобра-

жение данными «конфигурациями» (т. е. комбинациями подмножеств, имеющих данные свойства и состоящих в заданных отношениях).

Если изображение, описание которого должно быть построено, достаточно простое, например если оно составлено из простых геометрических фигур, каждая из которых имеет постоянное значение яркости, то из него нетрудно извлечь всю информацию, которая обычно требуется для его описания. Для таких изображений можно построить даже «полное» описание (т. е. описание, на основании которого можно точно воссоздать изображение). Однако, за исключением этих простейших случаев, требуемая информация едва ли может быть извлечена. В действительности для желаемого описания может потребоваться информация, которая не соответствует ни одному измеряемому признаку самого изображения, но которая получается наблюдателем на основании выводов из априорных сведений о данном классе изображений или об объектах, которым они соответствуют. Для получения в таких случаях адекватного описания изображения необходима модель класса изображений (И/ИЛИ объектов); такую модель можно использовать в качестве руководства при анализе изображения и как основу для дальнейших выводов.

В этой главе мы обсудим «языки» для описания различных типов изображений и их приложения к задаче спецификации моделей частных типов изображений. Языки для операций над синтетическими изображениями, подобные тем, которые создаются для разнообразных систем машинной графики, рассматриваться не будут.

Об одной современной задаче, демонстрирующей также предельные возможности анализа изображений, см. Kulsrud H. E., A general purpose graphic language, *Comm. ACM*, 11, 247 (Apr. 1968).

10.1. Язык для описания изображений

Пусть даны наборы признаков и отношений между объектами и нужно хранить всю эту информацию в компактном, упорядоченном виде, так чтобы легко было ответить на те или иные вопросы, касающиеся изображений,

или обнаружить на них заданные «конфигурации». Для этой цели были предложены [1] различные типы «запоминающих структур». Многие из этих структур были разработаны для хранения описаний рисунков в системах машинной графики [2]. В большинстве случаев эти устройства конструируются с целью хранения ограниченного количества типов объектов (здесь имеется в виду для хранения подмножества изображений), признаков или отношений. Поскольку вопрос о запоминающих структурах не является характерным для обработки изображений, в дальнейшем он обсуждаться не будет. В этом разделе рассмотрим различные «языки», разработанные специально для описания изображений (языки, позволяющие построить полное описание, по которому можно восстановить изображение, будут рассмотрены в разд. 10.2).

Описание рисунков

Рисунки рассматриваются в большинстве работ по языку описания изображений. Такие изображения обычно относительно легко расчленяются, а получающиеся при этом подмножества (дуги и отрезки кривых) можно компактно представить с помощью целного кодирования (разд. 3.4 и 9.4).

В каждом заданном рисунке можно идентифицировать различные типы «вершин» (т. е. углы, точки перегиба, узловые точки и т. д.). Затем можно описать эту фигуру, представляя ее в виде совокупности «вершин», соединенных дугами.

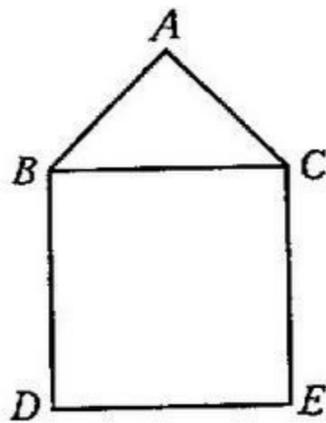
Для метрического описания самих дуг можно использовать такие признаки, как средний наклон, прямолинейность, длина и т. д.; можно даже совсем отбросить метрическую информацию и описать фигуру топологически, указывая число дуг, соединяющих каждую пару вершин [3—7].

В качестве примера такого описания рассмотрим изображение, представленное на фиг. 10.1. Здесь *A*, *D*, *E* — углы, а *B* и *C* — узлы. Если все эти пять точек считать вершинами, то можно описать топологию изображения

с помощью треугольной матрицы следующего вида:

$$\begin{array}{c|cccc}
 & A & B & C & D \\
 \hline
 B & 1 & & & \\
 C & 1 & 1 & & \\
 D & 0 & 1 & 0 & \\
 E & 0 & 0 & 1 & 1
 \end{array}$$

где 1 указывает на то, что вершина, которой соответствует столбец матрицы, соединяется простой дугой с вершиной, которой соответствует строка матрицы. Заметим, что если



Фиг. 10.1. Простой рисунок.

бы в качестве вершин рассматривать только узлы B и C , то матрица будет иметь вид

$$\begin{array}{c|c}
 B & \\
 \hline
 C & 3
 \end{array}$$

указывающий на то, что B и C соединены тремя линиями; однако в этом случае эти линии не являются отрезками прямой (в отличие от рассмотренного ранее случая, когда были использованы все пять точек в качестве вершин). Тот факт, что B и C — узлы, а A , D и E — углы, характеризуется в матрице тем, что в точках B и C сходятся три линии, а в точках A , D и E — лишь две.

Если рисунок имеет много вершин, матричное представление его топологии становится неэкономичным, поскольку в большинстве случаев лишь некоторые из всевозможных пар вершин соединяются дугами, а в матрице должно быть предусмотрено место для каждой пары. Этого можно избежать, описывая топологию рисунка другим способом. Именно для каждой вершины мы укажем другие (соседние) вершины, с которыми она соединяется дугами.

Применительно к фиг. 10.1 такой способ описания приводит к выражению вида

$$A(B, C); B(D, C, A); C(A, B, E); D(E, B); E(C, D).$$

Если бы какая-либо пара вершин соединялась более чем одной дугой, то в списке «соседей» были бы повторения обозначений соответствующей вершины. Заметим, что такое описание симметрично в том смысле, что если Y является соседом X , то X является соседом Y . Заметим также, что эту форму описания можно использовать для установления порядка расположения дуг в каждой вершине (в разобранный выше примере эти дуги указаны при обходе их против часовой стрелки); в матричном описании такой информации не содержится.

Для большинства целей определение топологии рисунка не дает адекватного описания его; помимо этого, необходима метрическая информация. Как было указано ранее, элементами такой информации являются длина и средний наклон линий. Из этой информации можно получить другие полезные метрические данные. Например, средний угол между двумя линиями определяется разностью их наклонов (здесь и далее, если дуги являются прямыми линиями, как в нашем примере, слово «средний» можно опустить). Можно указать, что два отрезка линий имеют равную длину (не указывая, какую именно), характеризуя их наклоны одним и тем же символом; этот же метод можно использовать для указания параллельности двух отрезков (их наклоны равны), перпендикулярности (их наклоны отличаются друг от друга на 90°) и т. д.

Ниже представлено множество возможных метрических данных (для фиг. 10.1); смысл обозначений ясен из контекста.

Длина $(A, B) = x$, длина $(B, C) = y$, длина $(D, E) = y$.

Длина $(A, C) = x$, длина $(C, E) = y$, длина $(B, D) = y$.

Наклон $(B, C) = \theta$, наклон $(D, B) = \theta + 90^\circ$,

наклон $(B, A) = \theta + 45^\circ$.

Наклон $(D, E) = 0$, наклон $(E, C) = \theta + 90^\circ$,

наклон $(C, A) = \theta + 135^\circ$.

Можно было бы также указать, что $x = y\sqrt{2}$ или что все «дуги» являются прямыми линиями, введя некоторый параметр, измеряющий *прямолинейность* каждой из них. Заметим, что не указаны значения y , θ и положение изображения. Таким образом, приведенное выше метрическое описание не зависит от положения, ориентации и масштаба, оно определяет рисунок только с точностью до подобия¹. Отметим, что информация, описанная выше, избыточна; например, информация о наклоне уже говорит о том, что $BCDE$ — прямоугольник и ABC — равнобедренный прямоугольный треугольник, так что для полного описания необходимо только указать на равенство любой пары смежных сторон $BCDE$. Следует обратить внимание на то, что по этому описанию можно полностью воспроизвести рисунок с точностью до подобия, а если мы оговорим величину x , то даже и до конгруэнтности. Другими словами, несмотря на то что в общем случае описания, построенные с помощью этой системы обозначений, являются лишь частичными для рисунков, составленных из отрезков прямых линий, они оказываются полными.

¹ Если бы изображение не было связным, то следовало дать какую-либо информацию об относительном положении связных его компонент для того, чтобы определить его форму с точностью до подобия. (Другой способ состоит в присоединении к изображению «невидимых дуг» и задании их длин и наклонов.)

Упражнение 1. Что можно было бы сказать о рисунке, если бы мы располагали только информацией о длине составляющих его отрезков?

Упражнение 2. Дать топологическое и метрическое описания, аналогичные приведенному выше, для каждого из следующих рисунков: 1) прямоугольный треугольник, 2) равносторонний треугольник, 3) трапеция, 4) параллелограмм, 5) прямоугольник, 6) ромб, 7) квадрат, 8) квадрат с диагоналями и 9) правильный шестиугольник.

Какие из этих описаний определяют соответствующее изображение с точностью до подобных преобразований?

Описание отношений подмножеств изображения

Другой класс языков описания изображений первоначально предназначался для описания геометрических соотношений между подмножествами изображения, а не структур обособленных подмножеств [8—11]. Обычно такие языки применяются для описания изображений, содержащих легко называемые подмножества, такие, как простые геометрические фигуры.

Очень простому частичному описанию рисунка, показанного на фиг. 10.1, можно придать форму, показывающую, что рисунок содержит два подмножества — прямоугольный треугольник и квадрат, между которыми существует отношение «примыкает сверху». (Это можно записать, например, так: «прямоугольный треугольник — примыкает сверху — квадрат».)

Более сложный пример приведен на фиг. 10.2. Здесь частичное описание может иметь следующий вид:

КВАДРАТ a , ПРЯМОУГОЛЬНИК b ,

КВАДРАТ c , ПРЯМОУГОЛЬНИК e ,

КВАДРАТ d ,

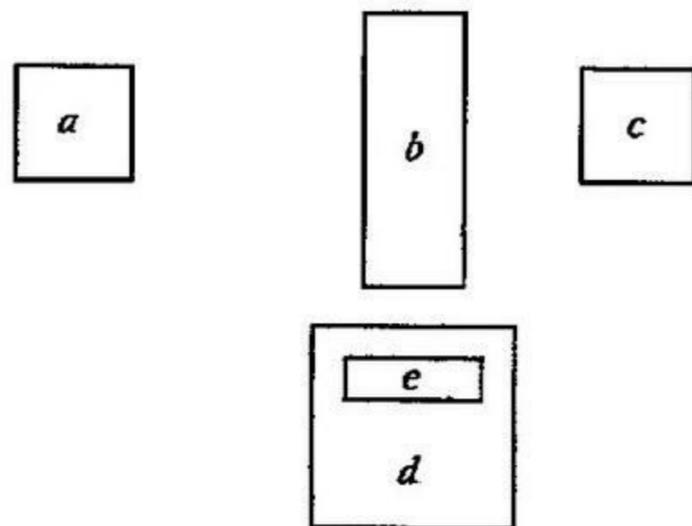
СЛЕВА a, b , КОНГРУЭНТНОСТЬ a, c , МЕЖДУ $b; a, c$,

СЛЕВА b, c , БОЛЬШЕ d, a , БЛИЖЕ $b; c, a$,

НАД b, d , ВЫШЕ b, d , БЛИЖЕ $b; d, c$,

ВНУТРИ e, d , ШИРЕ e, b ,

В идеальном случае желательно, чтобы язык описания был достаточно полным и можно было сформулировать любое утверждение (вопрос или команду), которое разумно высказать относительно изображений (или по



Ф и г. 10.2. Отношения между простыми фигурами.

меньшей мере относительно изображений некоторого заданного типа). Для описания микрофотографий нервных тканей был разработан ряд языков [12], которые в конечном счете предназначаются для реализации этой потенциальной возможности.

10.2. Язык изображений

Языком изображения будет называться такой язык описания изображения, который допускает формулирование полных описаний и по которым можно в точности воссоздать изображение. Пользуясь языком изображения, можно различными способами указать, каким образом изображения построены из частей (соответствующих определенным частям описания), так же как фразы и предложения строятся подбором последовательности слов. Части изображений, используемые как строительные блоки в языке этого типа, могут представлять собой любые геометрические образования — от отдельных элементов изображения до сложных геометрических фигур; как и следует ожидать, чем более «примитивны» элементы, тем более общий класс изображений может быть построен,

но тем более сложным является описание этих изображений по частям.

Для иллюстрации построения изображения путем подбора комбинации его частей мы опишем один способ воссоздания рисунка фиг. 10.1 из четырех блоков, показанных на фиг. 10.3, а: Здесь Г и Х могут мыслиться как «голова» и «хвост» соответственно; если мы их не будем различать, то не будет способа установить отличие двух комбинаций, показанных на фиг. 10.3, б. Если x , y являются двумя блоками, то положим, что \bar{x} («обращенный x ») обозначает x , у которого голова и хвост поменялись местами, а $x + y$ обозначает составной блок, полученный соединением головы x с хвостом y , причем хвост x становится хвостом $x + y$, а голова y становится головой $x + y$. В этих обозначениях две комбинации, изображенные на фиг. 10.3, б, можно описать

$$u + t \quad \text{и} \quad \bar{u} + t$$

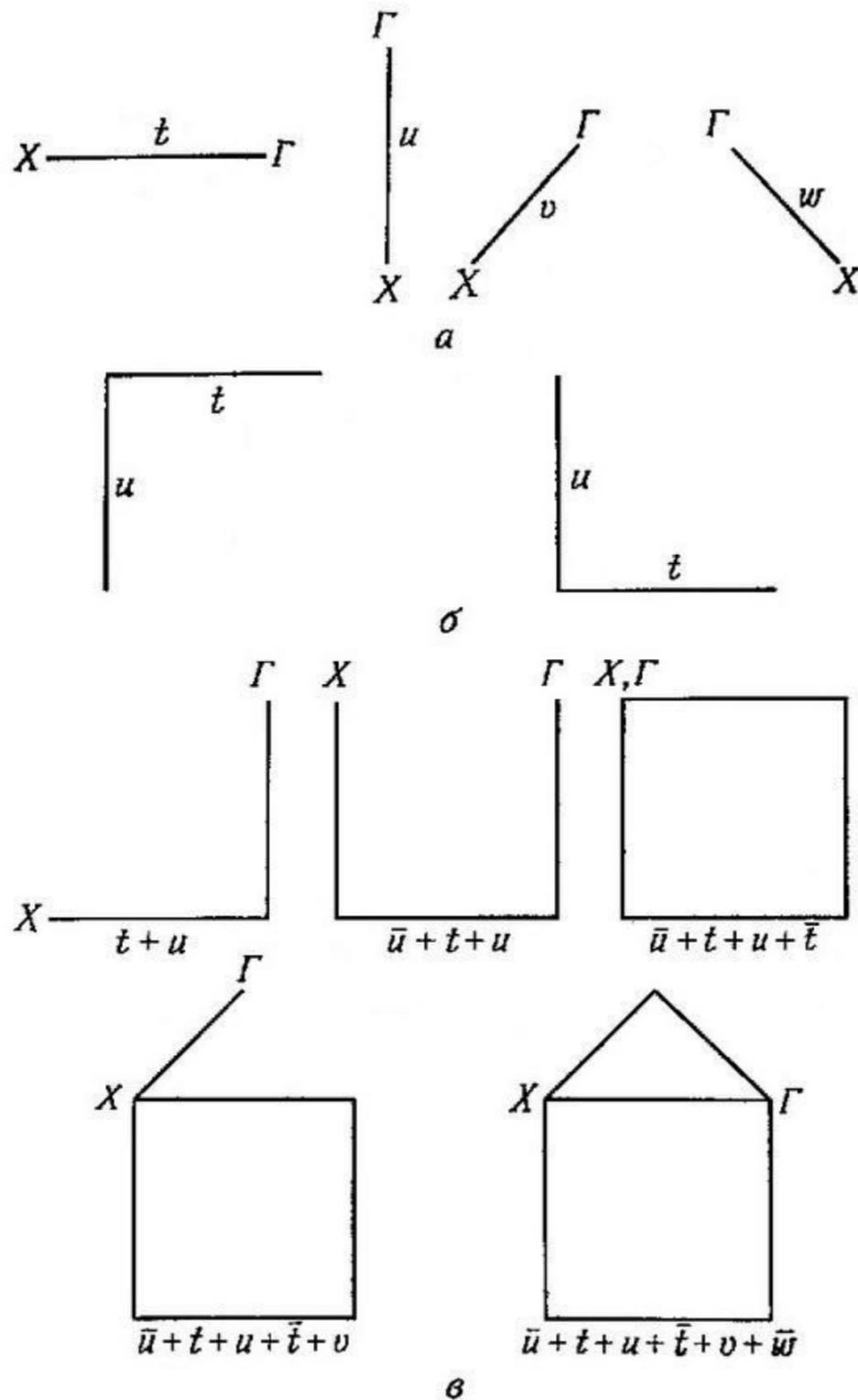
соответственно. Заметим, что операция, обозначенная знаком «+», является ассоциативной, т. е. дает ту же фигуру, что и $x + (y + z)$.

Последовательность построения фиг. 10.1 из четырех блоков с использованием операций сложения (+) и обращения изображена на фиг. 10.3, в¹.

Упражнение 1. Каким количеством различных способов можно, используя эти операции, построить квадрат из частей t и u ?

Только что описанный метод конструирования изображений может быть обобщен на случай, когда разрешено использовать еще и другие операции композиции блоков и дополнительные точки привязки (кроме головы и хвоста). Вообще операция композиции определяется указанием точек привязки элементов, которые должны быть приведены в соприкосновение, а также точек, которые должны стать точками привязки нового (составного)

¹ Когда мы соединяем $\bar{u} + t + u$ «голова к хвосту» с \bar{t} , эти части также соединяются и «хвостом к голове», но описание этого нам не показывает. Этот недостаток можно преодолеть, введя дополнительную операцию, которая четко определит, что два блока соединяются в замкнутую фигуру.



Ф и г. 10.3. Построение фигуры 10.1 из отрезков прямой.
 а — отрезки прямой; б — две комбинации отрезков; в — этапы построения.

элемента. Заметим, что таким способом можно построить из конечного множества частей бесконечное множество изображений, поскольку операции композиции можно применять произвольное число раз.

Упражнение 2. Построить конструкцию заглавных букв А, Е, F, Н, К и L, используя в качестве блоков отрезки прямых линий и а) вводя и б) не вводя точки привязки в середине отрезков.

Упражнение 3. Построить конструкцию букв В, С, D, О, Р, Q и S, используя в качестве блоков отрезки прямых линий и дуги окружностей.

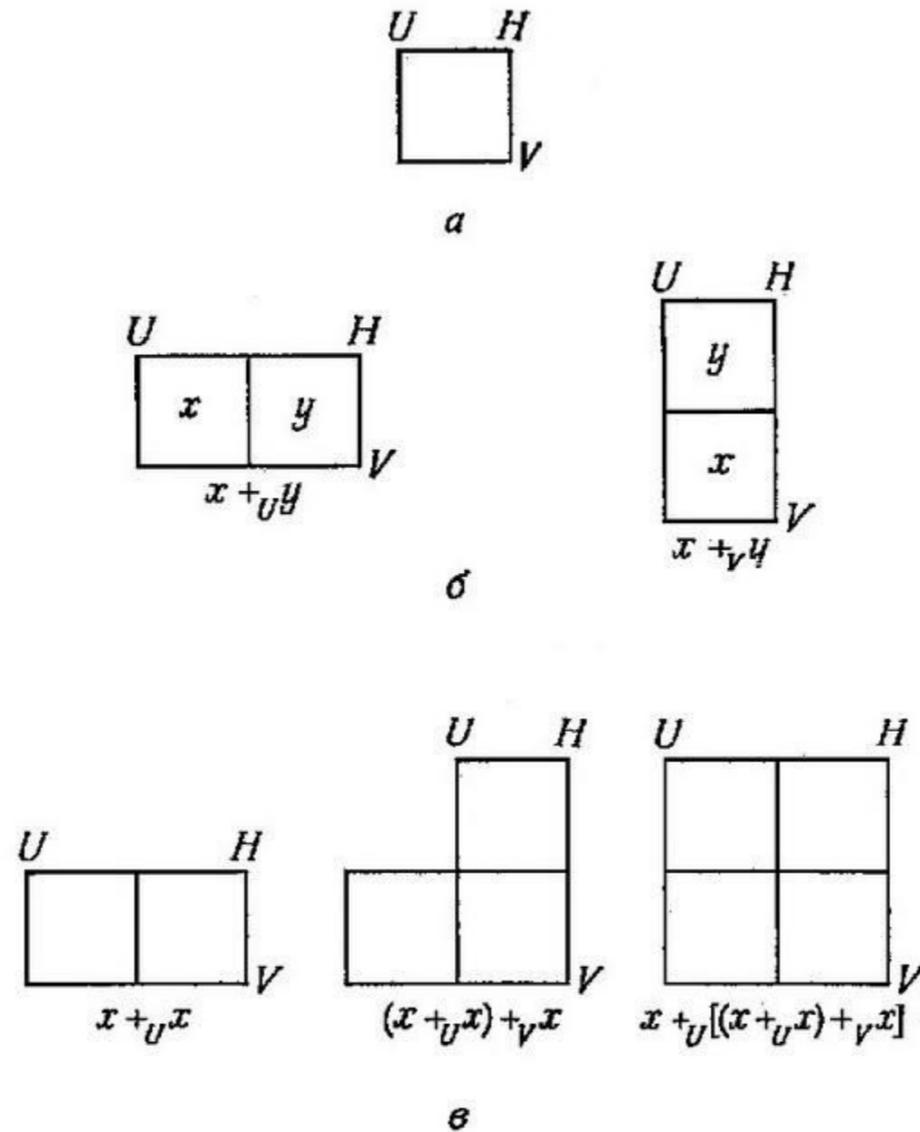
Изображения, которые можно построить подобными методами, не обязательно должны быть рисунками; точно также можно построить изображения «сплошных» объектов из «сплошных» частей. Предположим, что в качестве нашего основного блока выбран сплошной квадрат с тремя точками привязки (фиг. 10.4), и определим две операции сложения, $+_u$ и $+_v$, следующим образом. Именно $x +_u y$ означает, что мы соединяем элементы x и y , добиваясь совпадения точек H_x и U_y и принимая $H_{x+_u y} = H_y$, $U_{x+_u y} = U_x$, $V_{x+_u y} = V_y$; аналогично в операции $x +_v y$ совмещаем точки H_x и V_y и полагаем $H_{x+_v y} = H_y$, $U_{x+_v y} = U_y$ и $V_{x+_v y} = V_x$ (очевидно, что этими двумя операциями x соединяется с y слева и снизу соответственно). Используя эти операции, можно различными способами построить сплошной квадрат 2×2 , например $x +_u [(x +_u x) +_v x]$.

Упражнение 4. Сколькими способами можно проделать это построение?

Если каждый из блоков, используемых в языке изображения, имеет только по две точки привязки, то этот язык можно считать существенно одномерным, поскольку (фиг. 10.3) описание любого изображения, построенного из частей, состоит из нити элементов.

Методы конструирования изображений, описанные в этом разделе, важны для решения задачи распознавания зрительных образов. Пусть \mathfrak{F} — множество изображений, которые могут быть построены из данного множества блоков с использованием данных операций композиции.

Предположим, что дано неизвестное изображение f , принадлежащее \mathfrak{F} , тогда можно идентифицировать f , если идентифицируем его блоки и способ, которым они соеди-



Фиг. 10.4. Построение квадрата 2×2 из квадратов 1×1 .
 a — элементы; b — операции $+U$ и $+V$; c — этапы построения квадрата 2×2 .

няются. Очевидно, для того чтобы этот подход был наилучшим, блоки должны быть относительно простыми и, следовательно, более легкими для распознавания, чем сами изображения (в общем случае нельзя, конечно, использовать целые изображения в качестве блоков).

С другой стороны, если эти блоки слишком просты, то для построения изображений могут потребоваться очень сложные комбинации, так что опознавание этих комби-

наций становится сложным. Эта трудность может быть сведена к минимуму путем использования иерархического метода, в котором очень простые части комбинируются в простые «подизображения»; те в свою очередь соединяются в еще менее простые образования, и так до тех пор, пока изображение в \mathfrak{F} не будет построено.

Если интересующие нас изображения являются рисунками, то в качестве систематического метода их построения из очень простых частей, а именно из прямолинейных отрезков единичной длины, расположенных в основных направлениях (кратных 45°), можно воспользоваться методом цепного кодирования (заметим, что длины элементов, ориентированных по диагонали, в $\sqrt{2}$ больше длины элементов, расположенных горизонтально или вертикально). Естественный подход к конструированию фиг. 10.1, оперирующий вначале с такими простыми единичными отрезками, состоит в том, чтобы использовать их для построения четырех блоков t , u , v , w фиг. 10.3, a и затем — для построения фиг. 10.1, как показано на фиг. 10.3, b . В общем случае данный рисунок можно разбить на части, состоящие из вершин, соединенных дугами (разд. 10.1), а затем можно описать дуги с помощью цепного кодирования.

Упражнение 5. С помощью только что описанного метода построить конструкцию фиг. 10.1. Положить, что сторона квадрата равна шести единицам.

Упражнение 6. Как выглядит цепное кодирование отрезка прямой с наклоном 60° ?

Упражнение 7. Используя упражнение 6, дать построенную на базе цепного кодирования иерархическую конструкцию равностороннего треугольника.

Упражнение 8. Как характеризуются в цепном кодировании отрезки прямой?

В более общем случае произвольные дискретные изображения, содержащие сплошные объекты, могут быть построены с использованием в качестве элементов единичных квадратов (имеющих различные яркости), причем это построение целесообразно делать постепенно, т. е. вначале объединяя единичные квадраты в большие блоки

или фрагменты перепадов яркостей, вытянутых по основным направлениям, и затем составляя из них еще большие блоки перепада и т. д.

Упражнение 9. Построить «сплошной» дискретный вариант фиг. 10.1, используя в качестве блоков единичные квадраты. Размеры принять такими же, как и в упражнении 5. (Указание: возможно, потребуется ввести дополнительные операции композиции.)

10.3. Модели для классов изображений

Если имеется какой-либо язык для описания изображений, частичного или полного, его можно использовать для формулирования определений отдельных классов изображений. Например, как показано в разд. 10.1, язык для описания рисунков, использующий выражения «вершины» и «дуги, их соединяющие», можно применить для выделения различных простых типов геометрических фигур, например простых замкнутых многоугольников (все вершины — углы, а все дуги — отрезки прямой), параллелограммов (у которых пары несмежных сторон равны) и т. д. Аналогично (разд. 10.2, упражнения 3 и 4) определение букв алфавита может быть дано посредством соответствующего набора «штрихов» ([20—25], а также [18] гл. 5).

Во всех этих случаях изображения полностью стилизованы; определения букв описывают лишь отдельные образцы букв. Более сложная задача заключается в определении нестилизованых букв. Поскольку число различных способов написания какой-либо буквы огромно, нет смысла определять буквы путем описания каждого образца; это, разумеется, также справедливо и для более сложных; чем буквы, классов изображений.

Однако можно попытаться в таких случаях дать определения посредством перечисления ограничений, налагаемых на допустимые комбинации частей.

Например, можно попытаться определить буквы, используя комбинации единичных отрезков, образующих штрихи, и комбинации штрихов, образующих буквы.

Такой подход аналогичен подходу, применяемому в математической лингвистике, где класс рассматриваемых, т. е. «грамматических», предложений определяется указа-

нием грамматических правил, ограничивающих число способов, которыми слова могут соединяться в предложения. Ниже кратко рассмотрим некоторые основные концепции математической лингвистики и укажем, как они могут быть обобщены на языки изображений.

Рассмотрим ряд слов

a, and, another, boy, girl, hears, sees¹.

Из этих слов можно составить предложения многими способами, но только некоторые из этих комбинаций будут грамматическими предложениями в английском языке. Можно ограничить допустимые комбинации только грамматически корректными предложениями, формулируя грамматические правила следующим образом (здесь для краткости слово «или» заменено вертикальной чертой):

Артикль = a | another;

Глагол = hears | sees;

Существительное = boy | girl;

Составное существительное = артикль,
существительное;

Составная фраза = составное существительное |
составное существительное and составная фраза;

Предложение = составное существительное,
глагол, составная фраза.

В этой «грамматике» существует только одна операция композиции, а именно составление последовательности (например, артикль, существительное означает, что «существительное следует за артиклем»).

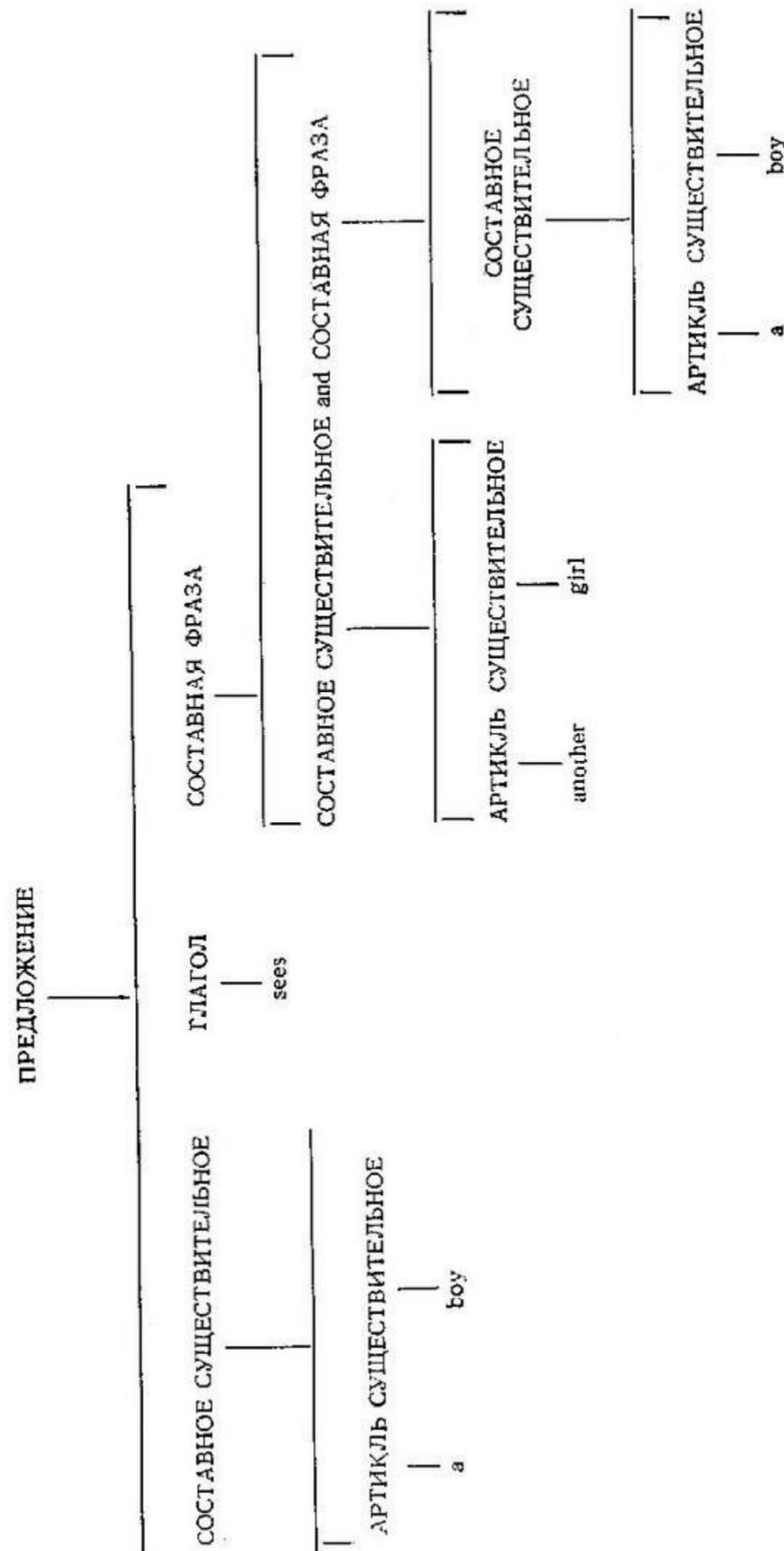
Все комбинации слов, которые можно создать, используя эти правила, будут грамматическими предложениями. Например, можно построить предложение

a boy sees another girl and a boy¹

следующим образом:

¹ a — неопределенный артикль; остальные слова означают соответственно: и, другой(ая), мальчик, девочка, слышит, видит. — Прим. перев.

² Мальчик видит другую девочку и мальчика. — Прим. перев.



Заметим, что, употребляя эту грамматику, можно построить бесконечное число предложений, поскольку правило

СОСТАВНАЯ ФРАЗА = СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ
and СОСТАВНАЯ ФРАЗА

можно использовать повторно, чтобы получить предложения произвольной длины, соединенные союзом and.

Упражнение 1. Написать, используя эту грамматику, конструкцию для предложения

another girl hears a girl and another boy and another girl¹.

Упражнение 2. Показать, что получится то же множество предложений, если заменить последние два правила данной грамматики следующими правилами:

ГЛАГОЛ, СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ = ГЛАГОЛ,
СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ and СОСТАВНОЕ
СУЩЕСТВИТЕЛЬНОЕ,

ПРЕДЛОЖЕНИЕ = СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ,
ГЛАГОЛ, СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ.

Эта перестроенная грамматика является теперь контексточувствительной в том смысле, что она позволяет нам заменять СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ на СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ and СОСТАВНОЕ СУЩЕСТВИТЕЛЬНОЕ только тогда, когда это производится в определенном контексте — точнее, когда впереди стоит глагол. Грамматика подобного типа называется контекстосвободной.

Упражнение 3. Написать правила для грамматики, использующей множество слов

a, and, another, boy, girl, hears, sees, who²

и допускающей такие предложения, как

a boy sees a girl and a boy who hears

a girl and another girl who sees a girl³.

¹ Другая девочка слышит девочку, и другого мальчика, и другую девочку. — Прим. перев.

² Который(ая). — Прим. перев.

³ Мальчик видит девочку и мальчика, который слышит девочку и другую девочку, которая видит девочку. — Прим. перев.

Существует очевидная аналогия между правилами грамматики, подобной только что описанной, и составлением изображений из блоков. Чтобы увидеть эту аналогию более ясно, запишем ряд правил — «грамматику изображений», — применимых для построения фиг. 10.1, на которой длина стороны квадрата равна шести единицам. Обозначим через a, b, \dots, g, h «единичные» отрезки в направлении $0, 45, \dots, 270$ и 315° соответственно:

$$t = a + a + a + a + a + a,$$

$$u = c + c + c + c + c + c,$$

$$\bar{t} = e + e + e + e + e + e,$$

$$\bar{u} = g + g + g + g + g + g,$$

$$v = b + b + b,$$

$$\bar{w} = h + h + h,$$

$$\text{КВАДРАТ} = \bar{u} + t + u + \bar{t},$$

$$\text{КРЫША} = v + \bar{w},$$

$$\text{ФИГУРА} = \text{КВАДРАТ} + \text{КРЫША}$$

Отметим, что эти правила допускают только одну фигуру, поскольку в них не входит союз «ИЛИ». Однако легко указать грамматику изображений, которая допускает много, даже бесконечно много фигур. Например, правило

$$\text{ГОРИЗОТОР (горизонтальный отрезок)} = a \mid \text{ГОРИЗОТОР} + a$$

дает отрезки горизонтальной линии всех возможных целочисленных длин. Точно так же, используя обозначения примера со сплошным квадратом из разд. 10.2, можно написать грамматику изображения:

$$\text{ПРЯМОУГОЛЬНИК} = x + \sqrt{x} \mid \text{ПРЯМОУГОЛЬНИК} + (x + \sqrt{x}),$$

которая дает сплошные прямоугольники с высотой в две единицы и длиной в целое число единиц. Также возможно (но более сложно) написать грамматику изображений, которая дает прямоугольники всех размеров (аналогичная задача для прямоугольных треугольников рассмотрена в разд. 10.2).

«Грамматический» подход к конструкции изображений, проиллюстрированный на предыдущих примерах, обеспечивает возможность построения описаний огромного количества изображений, используя небольшие наборы блоков и грамматические правила. Однако все приведенные примеры по-прежнему охватывают лишь простые (и, следовательно, стилизованные) геометрические фигуры. Более того, они дают слабое представление о том, как следует составлять грамматику изображений для более широких классов форм — даже для таких простых, как класс рукописных букв O , не говоря уже о таких классах, как «хромосомы» или «облака». Кроме того, до настоящего времени почти ничего не сделано в области грамматики для полутоновых изображений, включающих в себя поверхностные текстуры.

Если желательно начать с блоков относительно высокого уровня, не пытаясь строить их из более простых блоков, то можно написать грамматику для «обобщенных» изображений, составленных из данных блоков. Однако пригодность такой грамматики для распознавания изображений реальной действительности будет зависеть от нашей способности опознать эти блоки. Ниже приведен простой пример

$$\text{ПАЛОЧКА} = \text{ОСТРИЕ} \mid \text{СТОРОНА} + \text{ПАЛОЧКА} + \text{СТОРОНА}$$

$$\text{КАРАНДАШ} = \text{ПАЛОЧКА} + \text{СТИРАЛЬНАЯ РЕЗИНКА}$$

Здесь **ОСТРИЕ**, по-видимому, имеет форму V , **СТОРОНА** — отрезок прямой некоторой единичной длины и т. д. Другими примерами являются упомянутая ранее [20—25] грамматика для печатных букв, строчных рукописных, в которой используются штрихи, и грамматика ([26] и [14] гл. 8) для контуров хромосом. Этим примеры используют только одну операцию композиции типа соединения («+»); ниже приводится следующий простой пример грамматики, включающий операции композиции и представляющий другие типы отношений между частями:

$$\text{ГЛАЗ} = \text{ГЛАЗ (рядом с)} \text{ ГЛАЗОМ}$$

$$\text{ЧЕРТЫ} = \text{ГЛАЗА (над)} \text{ НОСОМ (над)} \text{ РТОМ}$$

$$\text{КОНТУР} = \text{УХО (наряду с)} \text{ ГОЛОВОЙ (наряду с)} \text{ УХОМ}$$

$$\text{ЛИЦО} = \text{ЧЕРТЫ (внутри)} \text{ КОНТУРА}$$

Другой пример (грамматика для построения изображений домов из блоков: «стена», «крыша», «фронтон», «дверь», «окно» и «труба») рассмотрен в работе [27]. В этом случае, конечно, практическая полезность такой грамматики зависит от нашей способности распознавать не только блоки, но также и отношения, представляемые операциями композиции, поскольку они также не являются точно определенными.

Если даны грамматика для класса изображений и язык (например, подмножество английского языка), в котором могут быть описаны изображения, то можно рассматривать задачу описания изображений как задачу «перевода» с языка изображений на язык описаний [12]. Подобный перевод относительно быстро приводит к цели для языка изображений «высокого уровня», в котором элементы и операции композиции имеют простое словесное описание; ситуация усложняется, если построение изображений должно производиться из отдельных растровых элементов, когда во многих случаях (например, при построении поверхностных текстур) дать словесное описание изображения очень трудно. Обычная грамматика строит предложения из слов, которые являются единицами относительно высокого уровня. Грамматика для образования сложных изображений из растровых элементов была бы аналогична «грамматике» для построения английских фраз из элементарных отсчетов речевых сигналов. Другая трудность, которая в проблеме описания изображений присоединяется к трудности перевода на естественный язык, состоит в том, что часто требуется информация, в действительности не содержащаяся в изображении, но которая может быть выведена на основании априорных знаний относительно изображаемых объектов реального мира. Общая система описания изображения должна, следовательно, не только обладать способностью к обработке изображений, но также и знать реальную модель, лежащую в основе изображения, т. е. сведения о явлениях жизни.

ЛИТЕРАТУРА

1. Gray J. C., Compound data structures for computer aided design: a survey, *Proc. ACM Nat. Confer.*, Aug. 1967, p. 355.
2. Prince M. D., Man-computer graphics for computer-aided design, *Proc. IEEE*, 54, 1698 (Dec. 1966); русский перевод: Принс, Графические методы связи человек — вычислительная машина при машинном проектировании, *ТНИЭР*, 54, № 12, 78 (1966).
3. Marill T., et al., CYCLOPS-1: a second-generation recognition system, *FJCC*, 27 (Dec. 1963).
4. Bloom B. H., Marill T., CYCLOPS-2: a computer system that learns to see, AD 624152 (Oct. 1965).
5. Marill T., Bloom B. H., Learning and perceptual processes for computers, *Ann. New York Acad. Sci.*, 128, 1029 (Jan. 1966).
6. Guzman A., Scene analysis using the concept of model, AD 652017 (Jan. 1967).
7. Guzman A., Some aspects of pattern recognition by computer, AD 656041 (Febr. 1967).
8. Simmons R. F., Answering English questions by computer: a survey, *Comm. ACM*, 8, 53 (Jan. 1965); русский перевод: Симмонс, ЭЦВМ, отвечающие на вопросы, заданные по-английски, *Зарубежная радиоэлектроника*, № 7 (1965).
9. Londe D. L., Simmons R. F., NAMER: a pattern-recognition system for generating sentences about relations between line drawings, *Proc. ACM Nat. Confer.*, Aug. 1965; p. 162.
10. Kochen M., Automatic question-answering of English-like questions about simple diagrams, *J. ACM*, 16, 26 (Jan. 1969).
11. Evans T. G., A heuristic program to solve geometric-analogy problems, *SJCC*, 327 (Apr. 1964).
12. Lipkin L. E., Watt W. C., Kirsch R. A., The analysis, synthesis and description of biological images, *Ann. New York Acad. Sci.*, 128, 984 (Jan. 1966).
13. Kirsch R. A., Computer interpretation of English text and picture patterns, *IEEE Trans. Electronic Computers*, EC-13, 363 (Aug. 1964).
14. Clowes M. B., An hierarchical model of form perception, *MPSVF*, p. 388.
15. Clowes M. B., Perception, picture processing and computers, in «Machine Intelligence 1» (N. L. Collins, D. Michie, eds.), American Elsevier, New York, 1967, p. 181.
16. Clowes M. B., Pictorial relationships — a syntactic approach, in «Machine Intelligence 4» (B. Meltzer and D. Michie, eds.), American Elsevier, New York, 1969, p. 361.
17. Clowes M. B., Transformational grammars and the organization of pictures, in «Automatic Interpretation and Classification of Images» (A. Grasselli, ed.), Academic Press, New York, 1969.
18. Feder J., Languages of encoded line patterns, *Information and Control*, 13, 230 (Sept. 1968).

19. Miller W. F., Shaw A. C., Linguistic methods in picture processing — a survey, *FJCC*, 279 (Dec. 1968).
20. Narasimhan R., On the description, generation and recognition of classes of pictures, in «Automatic Interpretation and Classification of Images» (A. Grasselli, ed.), Acad. Press, New York, 1969.
21. Eden M., Halle M., The characterization of cursive writing, *Proc. 4th London Symp. Inform. Theory*, Butterworth, London and Washington, D.C., 1961, p. 287.
22. Eden M., On the formalization of handwriting, *Proc. Symp. Appl. Math.*, 12, 1961, p. 83.
23. Eden M., Handwriting and pattern recognition, *IRE Trans. Information Theory*, IT-8, 160 (Febr. 1962).
24. Eden M., Handwriting generation and recognition, in «Recognizing Patterns» (P. A. Koles and M. Eden, eds.), M.I. T. Press, Cambridge, Massachusetts, 1968, p. 138; русский перевод: Иден М., Генерирование и распознавание рукописного текста, в сб. «Распознавание образов. Исследование живых и автоматических распознающих систем», изд-во «Мир», М., 1970, стр. 174.
25. Kpocke P. J., Wiley R. G., A linguistic approach to mechanical pattern recognition, *Proc. IEEE Comput. Confer.*, Sept. 1967, p. 142.
26. Ledley R. S., High-speed automatic analysis of biomedical pictures, *Science*, 146, 216 (Oct. 9, 1964).
27. Ledley R. S., Programming and Utilizing Digital Computers, McGraw-Hill, New York, 1962, p. 364; русский перевод: Ледли Р. С., Программирование цифровых вычислительных машин, изд-во «Мир», М., 1966.

ОГЛАВЛЕНИЕ

Предисловие редактора	5
Предисловие автора к русскому изданию	5
Глава 1. Изображения и обработка изображений	9
1.1. Изображения	9
1.2. Задачи обработки изображений	12
1.3. Общая литература	14
Глава 2. Кодирование изображений	16
2.1. Вероятность и информация	16
2.2. Эффективное кодирование одиночных сообщений	19
2.3. Эффективное кодирование последовательности сообщений	22
2.4. Некоторые дополнительные замечания о кодировании	28
Глава 3. Аппроксимация изображений	30
3.1. Дискретизация с фиксированным положением отсчетов	31
3.2. Квантование с фиксированным расположением уровней	34
3.3. Дискретизация и квантование с переменным шагом	39
3.4. Аппроксимация рисунков	42
Приложение	44
Доказательство теоремы отсчетов	44
Глава 4. Теория пространственно-инвариантных операций над изображениями	48
4.1. Пространственно-инвариантные операции	48
4.2. Пространственно-инвариантные линейные операции	51
4.3. Преобразование Фурье	55
4.4. Импульсные функции рассеивания и передаточные характеристики	59

Глава 5. Методы реализации пространственно-инвариантных операций	63
5.1. Цифровые методы	63
5.2. Оптико-электронные методы	64
5.3. Оптические методы, не требующие когерентного освещения	66
5.4. Когерентно-оптические методы	77
Глава 6. Применения пространственно-инвариантных операций над изображениями	90
6.1. Согласованная фильтрация	90
6.2. Пространственная фильтрация и коррекция изображений	100
6.3. Измерение качества изображения	104
6.4. Улучшение качества изображений. Устранение помех	107
6.5. Улучшение качества изображения. Повышение резкости	114
Глава 7. Свойства изображений и обработка их методами теории распознавания образов	128
7.1. Распознавание образов	128
7.2. Отбор признаков	130
7.3. Классы признаков	136
7.4. Признаки-функции	144
Глава 8. Разбиение изображений	156
8.1. Разбиение на части изображения произвольного вида	156
8.2. Получение новых подмножеств изображения на основе заданного. 1. Связность	162
8.3. Получение новых подмножеств изображения на основе заданного. 2. Расстояние	173
8.4. Получение новых подмножеств изображения на основе заданного. 3. Направление и форма	183
Глава 9. Геометрические свойства подмножества изображения	193
9.1. Топологические признаки	193
9.2. Метрические признаки. Размеры	196
9.3. Метрические признаки. Форма	199
9.4. Функции подмножеств изображений	202
Глава 10. Описание изображений и язык изображений	207
10.1. Язык для описания изображений	208
10.2. Язык изображений	214
10.3. Модели для классов изображений	220

А. Розенфельд

**РАСПОЗНАВАНИЕ И ОБРАБОТКА
ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ
ВЫЧИСЛИТЕЛЬНЫХ МАШИН**

Редактор Т. Б. Моисеева
Художник В. А. Медников

Художественный редактор Н. Г. Блывов
Технический редактор Е. С. Потапенкова
Корректор В. И. Киселева

Сдано в набор 6/IV 1972 г.

Подписано к печати 25/VIII 1972 г.

Бумага тип. № 2 84×108¹/₃₂ = 3,63 бум. л.

Усл. печ. л. 12,18. Уч.-изд. л. 11,51.

Изд. № 20/6351 Цена 1 р. 02 к. Зак. 251

Типография № 11 Главполиграфпрома
Комитета по печати
при Совете Министров СССР,
Москва, 88, Угрешская, 12.

ИЗДАТЕЛЬСТВО «МИР»

готовит к выпуску в 1973 году следующую книгу:

Вычислительные системы и автоматическая диагностика заболеваний сердца. Под ред. Ц. КАЦЕРСА, Л. ДРЕЙ-ФУСА, Нью-Йорк, 1970, перевод с английского

Книга написана виднейшими американскими специалистами по автоматизации медицинской диагностики. В ней рассмотрены проблемы, связанные с созданием и применением систем автоматической диагностики заболеваний сердца; характеристики оборудования для регистрации исходной информации, методы и средства ввода этой информации в вычислительные машины, принципы составления вычислительных программ для анализа электрокардиограмм и вынесения диагностического заключения, достигнутые результаты и перспективы развития систем автоматической диагностики. Большую ценность представляет описание конкретных алгоритмов автоматической диагностики по электрокардиограмме, разрабатываемых и уже функционирующих вычислительных комплексов для обслуживания больных в клинике.

Предназначена для врачей, специалистов по технической кибернетике и вычислительной технике и для всех, кто интересуется актуальными проблемами автоматизации медицинской диагностики.
План 1973 г., № 145.

Уважаемый читатель!

Заказы на эту книгу вы можете оформить в магазинах местного книготорга, торгующих научно-технической литературой, или же послать в фирменную секцию издательства «Мир» при Московском Доме книги по адресу: 121019, Москва, Г-19, проспект Калинина, 26, п/я № 42. Книга будет выслана наложенным платежом.