Л.М. Местецкий

# НЕПРЕРЫВНАЯ МОРФОЛОГИЯ БИНАРНЫХ ИЗОБРАЖЕНИЙ

ФИГУРЫ
СКЕЛЕТЫ
ЦИРКУЛЯРЫ

Местецкий Л.М.

# Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры

МОСКВА
ФИЗМАТЛИТ ®

УДК 519.72
ББК 22.19; 32.97
М 53

*Издание осуществлено при поддержке Российского фонда фундаментальных исследований по проекту 08-07-07017*

Местецкий Л.М. **Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры.** — М.: ФИЗМАТЛИТ, 2009. — 288 с. — ISBN 978-5-9221-1050-1.

Книга посвящена исследованию способов описания формы объектов в цифровых бинарных изображениях с помощью непрерывных моделей. Использование непрерывных моделей существенно упрощает решение многих задач анализа, распознавания и преобразования изображений. В книге в качестве универсальной непрерывной модели формы используется понятие фигуры — замкнутой области, граница которой состоит из конечного числа непересекающихся жордановых кривых. Рассматриваются три взаимосвязанных способа представления фигур: многоугольными границами, скелетами, семействами кругов (циркулярами). Задача построения непрерывной модели для бинарного изображения состоит в аппроксимации его фигурами. В книге описываются разработанные автором методы решения этой задачи и их практические приложения.

Книга рассчитана на научных работников и инженеров, профессионально занимающихся вопросами математического обеспечения цифровой обработки и анализа изображений. Она также может быть полезна аспирантам и студентам соответствующих специальностей.

Оформление обложки:
Анна Аренштейн

В оформлении обложки использовано изображение
антропозооморфного идола с птичьими крыльями,
сильно выступающими в стороны, и широким головным убором.
I–III вв. Из раскопок Н.Н. Новокрещеных 1896–1897 гг.
на Гляденовском костище (Пермской обл.). Бронза, литье. $14,5 \times 9$ см.

Изображение опубликовано в книге:
*В. А. Оборин, Г. Н. Чагин.* Чудские древности Рифея.
Пермский звериный стиль. Серия «Искусство Прикамья». — Пермь:
Пермское книжное издательство, 1988.

ISBN 978-5-9221-1050-1

# ОГЛАВЛЕНИЕ

ЧАСТЬ III

## ГРАНИЦЫ И СКЕЛЕТЫ

ЧАСТЬ IV
## СКЕЛЕТЫ И ЦИРКУЛЯРЫ



ЧАСТЬ V
## ПРИКЛАДНЫЕ ЗАДАЧИ

*Маме Евгении Львовне и жене Ирине*
*с благодарностью за вдохновение и поддержку*

# ПРЕДИСЛОВИЕ

Тема этой книги — апология непрерывного подхода к описанию формы объектов в цифровых бинарных изображениях, названного здесь непрерывной морфологией. Предметом нашего внимания является понятие формы, а также способы описания формы, удобные для анализа, преобразования и распознавания изображений.

С одной стороны, понятие формы предмета является простым и ясным для любого человека. Маленький ребенок демонстрирует такое понимание, когда рисует домики в виде прямоугольников с треугольными крышами и солнышко в виде кружочка. По форме прически юные меломаны классифицируют друг друга. Любому дачнику ясно, какое ведро имеет цилиндрическую форму, а какое — коническую. Спортсмены хорошо знают дельтовидную, т. е. треугольную, и трапециевидную мышцы. Психологи умеют определять характер человека на основе его отношения к треугольникам, квадратикам и кружочкам. Но с другой стороны, если попросить их дать общее определение, что же такое форма предмета, вряд ли удастся услышать что-либо вразумительное. Форма является одним из тех фундаментальных понятий, смысл которых понимают все, но объяснить строго могут немногие.

В повседневной жизни люди не нуждаются в строгом определении понятия формы. Однако специалисты по работе с цифровыми изображениями должны волей-неволей такое определение иметь. Стремительное проникновение компьютерных технологий во все сферы жизни современного общества, их использование в различных видах человеческой деятельности естественным образом привело к разработке «видящих» машин, способных выполнять такие специфические человеческие функции, как зрение и анализ увиденных изображений. Для создания алгоритмов анализа и распознавания тех сцен, которые компьютер «увидел» через фото- и видеокамеры, сканеры или какие-нибудь другие устройства, способные регистрировать изображения, требуются строгие математические модели формы объектов реального мира.

Разработка систем машинного понимания изображений ставит перед их создателями сложную задачу — научить машину выделять форму объектов в изображениях, сравнивать объекты по их форме и определять сходные и различные формы. Сложность состоит в том, что механизмы работы человеческого интеллекта при решении этих задач не вполне понятны. Сами люди, постоянно и многократно решая подобные задачи в повседневной жизни, оперируют интуитивными понятиями формы. Поэтому простые вопросы, на которые легко ответит даже маленький ребенок, например, одногорбый или двугорбый верблюд нарисован в книжке или кто за кем вытягивает репку — где внучка, где Жучка, а где кошка — превращаются в самостоятельную сложную задачу для компьютера.

При разработке алгоритмов, связанных с пониманием изображений, проявляется определенное противоречие между человеческим и компьютерным представлением формы объектов. Особенностью компьютерного представления видеоинформации является дискретная или растровая структура изображений, размещенных в памяти компьютера. Изображение представляется прямоугольной матрицей точек, обладающих определенным цветом и яркостью. Именно в этой матрице нужно увидеть какие-то формы, выделить их и сравнить между собой или с некоторыми эталонами, чтобы произвести их классификацию и обеспечить понимание компьютером сцены, представленной в изображении.

А человеческий глаз не замечает растровую природу цифрового изображения в компьютере. Высокая разрешающая способность современных видеокамер и сканеров приводит к тому, что точки в матрице изображений имеют очень малые размеры и расположены с высокой плотностью. Поэтому в глазах человека эти точки сливаются и сами изображения воспринимаются им как непрерывные «сплошные» объекты. Воспринимая форму таким образом, человек легко формулирует характеристики и признаки объектов в терминах геометрических фигур или непрерывных функций, но не в терминах дискретных растровых изображений, представленных в виде матриц точек.

Таким образом, перед создателями систем обработки изображений возникает проблема таких преобразований формы изображений, при которой «человеческие», интуитивно понятные геометрические способы анализа, сравнения, распознавания и преобразования формы можно было бы применить к изображениям, представляемым в компьютере в виде матриц. Решать эту проблему можно, сближая разными путями дискретные компьютерные и непрерывные человеческие представления формы объектов. Можно пытаться переформулировать человеческие представления о форме, как о непрерывном «сплошном» объекте, в терминах дискретных растровых изображений. Здесь возникают непростые задачи описания таких существенно непрерывных понятий как связность, гладкость, кривизна применительно к матрицам точек. В противоположность такому подходу, состоящему в «дискретизации» непрерывных понятий, можно попытаться продвинуться с другой стороны: построить непрерывные модели для объектов, выделенных в дискретной матрице изображений, а далее формулировать алгоритмы в терминах непрерывных моделей, адекватных человеческому восприятию понятия формы. Эти два пути, которые можно условно охарактеризовать как «дискретный» и «непрерывный» подходы, представляют собой крайние варианты, между которыми возможны компромиссные решения, комбинирующие в различной степени дискретные и непрерывные представления формы.

В этой книге предпринята попытка разработки и реализации в наиболее полном виде непрерывного подхода к описанию формы объектов в цифровых изображениях — начиная с бинарных растровых изображений, полученных в результате сегментации исходных цветных либо полутоновых изображений. Задача ставится следующим образом. Для объектов, представленных на изображении в виде матриц точек, нужно построить непрерывные модели формы в виде геометрических фигур на евклидовой плоскости.

В качестве универсальной непрерывной модели формы мы используем понятие замкнутой ограниченной области. Для представления таких областей предложены так называемые граничное, скелетное и циркулярное описания. Эти представления формы являются непрерывными. Задача построения непрерывной модели для цифрового бинарного изображения состоит в том, чтобы превратить исходное дискретное изображение, представленное матрицей точек, в непрерывные геометрические объекты: границы, скелеты, циркуляры. Соответствующие преобразования называются в книге дискретно-непрерывными преобразованиями изображений.

Предлагаемый непрерывный подход не является традиционным. Обычно при анализе формы изображений используются уже упомянутые дискретные модели формы, представляющие собой изображения границ и скелетов в виде «линий на растре». Дискретные методы, в основе которых лежат, как правило, эвристические соображения и приемы, получили чрезвычайно широкое распространение, поскольку они достаточно просто описываются и допускают простую программную реализацию. Такой подход вполне имеет право на существование и оправдан во многих случаях. Однако недостатки, присущие дискретным методам, сдерживают возможности создания современных технических систем машинного зрения. Несоответствие дискретного описания изображений непрерывному человеческому представлению является, по-видимому, одним из главных факторов такого сдерживания. Другим существенным ограничением является относительно невысокая скорость работы дискретных алгоритмов с изображениями большой размерности. Это ограничение проявляется в системах машинного зрения, работающих в реальном времени. Не внушает больших надежд на ускорение дискретных алгоритмов возможность использования параллельных вычислений. В дискретных алгоритмах имеются строго последовательные шаги, число которых растет с ростом размерности и сложности изображений. А размерность и сложность, в свою очередь, неизбежно возрастают по мере повышения разрешающей способности видеокамер и сканеров.

Таким образом, преодолеть указанные недостатки в рамках дискретных методов вряд ли возможно. Непрерывный подход, напротив, открывает широкие возможности для решения возникающих проблем.

В обработке и анализе изображений принято называть инструменты и методы, связанные с извлечением компонент изображения, относящихся к форме объектов, *математической морфологией*. Этот термин заимствован из биологии, где под морфологией обычно понимается исследование формы и строения животных и растений. Математическая морфология представляет собой вполне сложившийся набор инструментов для представления и описания границ, скелетов, выпуклых оболочек, а также способов предобработки, фильтрации, утончения, усечения изображений. При этом все эти задачи решаются в рамках дискретного подхода, методы работают непосредственно с дискретными изображениями, оперируя над матрицами яркости, в том числе и бинарными. Поскольку целью предлагаемых непрерывных методов остается решение тех же самых задач: построение границ, скелетов, дескрипторов формы, мы будем использовать термин *непрерывная морфология* для обозначения непрерывного подхода.

В рамках этой книги мы хотим продемонстрировать важные преимущества непрерывного подхода по сравнению с дискретным:

— математическая корректность;

— адекватность человеческому представлению о форме и ее преобразованиях;

— широкие возможности для преобразования и сравнения форм;

— высокая вычислительная эффективность.

Платой за эти преимущества является существенное усложнение алгоритмов с точки зрения их математического содержания и программной реализации. Но, как отмечал академик Н. Н. Моисеев, «...инженерам не нужны примитивные математические приемы — ими инженер владеет сам не хуже математиков. Трудные технические задачи требуют настоящего математического творчества» [17].

Начальный импульс к работе над непрерывными методами для цифровых бинарных изображений автор получил, когда обнаружил, что красивый и эффективный алгоритм скелетизации простого многоугольника, опубликованный Дэвидом Ли в 1982 году [39], не используется в обработке изображений, а вместо него применяются дискретные алгоритмы скелетизации, основанные на морфологическом утончении, дистанционной карте и им подобные. Получаемые при этом новые бинарные изображения по внешнему виду хоть и напоминают скелеты и даже называются скелетами, но при этом не имеют строгого математического определения, а их построение требует больших вычислительных затрат. Отсутствие строгой математической модели приводит к тому, что разные методы дискретной скелетизации, примененные к одному и тому же изображению, дают разные результаты. А представление скелета в виде нового бинарного изображения, а не в виде графа и семейства окружностей, существенно ограничивает возможности классификации и преобразования изображений.

Задавшись целью все-таки применить к построению скелетов изображений мощные и элегантные алгоритмы вычислительной геометрии, автор столкнулся с необходимостью решения нескольких дополнительных задач.

Во-первых, эффективные алгоритмы скелетизации в вычислительной геометрии разработаны только лишь для областей с кусочно-линейной границей. В частности, алгоритм Ли работает только с простыми многоугольниками. Поэтому для работы с цифровыми изображениями вначале нужно решить *задачу аппроксимации присутствующих в них объектов простыми многоугольниками*, т. е. не имеющими пересекающихся сторон. Сложность этой задачи состоит именно в обеспечении простоты получаемых многоугольников. В частности, многочисленные алгоритмы, строящие многоугольники, вершинами которых являются все граничные точки дискретного объекта, не гарантируют простоты этих многоугольников, поскольку их стороны могут иметь самопересечения. Более того, в случае, когда сложное изображение содержит несколько объектов или объекты не являются односвязными (имеют «дыры»), такие граничные многоугольники часто «слипаются», т. е. имеют пересекающиеся границы.

Во-вторых, несмотря на то, что алгоритм Ли скелетизации многоугольника с $n$ вершинами имеет сложность $O(n\log n)$ в худшем случае, все известные попытки его обобщения на случай многосвязной многоугольной фигуры (многоугольника с многоугольными дырами) приводили к сложности $O(nk + n\log n)$, где $k$ — это количество внутренних многоугольных дыр в объекте. При большом числе дыр, что часто имеет место в реальных изображениях, такие затраты времени оказываются неприемлемыми. Таким образом, *задача разработки эффективных алгоритмов скелетизации многосвязной многоугольной фигуры* остается актуальной.

Третья задача — регуляризация скелета — связана с высокой чувствительностью скелета к шумовым эффектам, неизбежно присутствующим в изображениях. Суть задачи в следующем. Поскольку эффективные алгоритмы скелетизации известны только для многоугольников, в случае произвольной замкнутой области приходится сначала аппроксимировать многоугольниками ее границу, а затем строить скелет полученной многоугольной фигуры. Многоугольная аппроксимация в этом случае не является однозначной, однако при выборе другой аппроксимирующей многоугольной фигуры (в пределах той же точности) получается совсем другой скелет. Это свойство задачи скелетизации показывает ее некорректность в терминологии А. Н. Тихонова: *решение задачи неустойчиво и нуждается в регуляризации*. Данная проблема хорошо известна и в дискретных методах скелетизации. Однако, если там она решается на основе эвристических приемов, то в рамках непрерывного подхода появляется возможность ее строгого решения эффективными алгоритмами.

Четвертая задача возникает из стремления использовать скелеты не только для анализа изображений, но и для их преобразования. Такие задачи возникают в компьютерной графике, например, при создании анимационных клипов. Также подобная задача появляется в распознавании изображений, когда классификация объектов осуществляется путем сравнения с эталонами, но при сравнении нужно обеспечить максимально возможное совпадение с эталоном с учетом допустимых деформаций объекта. Осуществить подобные операции с использованием дискретных скелетов невозможно в принципе. В рамках же непрерывного подхода это может быть сделано на основе преобразования скелетного графа и связанного с ним семейства окружностей. Построив огибающую измененного таким образом семейства окружностей, мы получаем измененный объект, имеющий другую форму. Правда в этом новом объекте скелет уже может не совпадать с измененным скелетным графом исходного объекта. Для того чтобы сделать корректными такие преобразования, нужно каким-то образом обобщить понятие скелета. Решение этой задачи привело к концепции описания формы объекта в виде так называемой циркулярной фигуры. Таким образом, возникла *задача аппроксимации бинарных изображений циркулярными фигурами* и последующей деформации объектов через преобразование циркулярных фигур.

Решение всех перечисленных задач составляет содержание этой книги.

Нужно отметить, что автор не первый, кто предпринял попытку развития непрерывной морфологии в плане практического использования для обработки изображений. В работах Уго Монтанари [44, 45] были сформулированы основные принципы подобного непрерывного подхода: многоугольная аппроксимация связных объектов бинарного изображения и построение непрерывных скелетов многоугольников. Однако реализация этого подхода и эксперименты, проведенные в этих работах, видимо, не вызвали большого оптимизма. Во всяком случае, развития этих работ не последовало. В то же время, успехи дискретной морфологии того времени явились приемлемой альтернативой непрерывным математическим моделям.

Описанный в книге непрерывный подход к анализу формы изображений разрабатывался автором с начала 1990-х годов. Основные его принципы, методы и алгоритмы публиковались в научной периодике, докладывались на конференциях, частично размещались в Интернете [12–16, 42, 43], однако в собранном виде публикуются впервые. Собирая материал в книгу, автор стремился выдержать единство терминологии и обозначений. Если это не везде удалось, автор просит прощения у придирчивого читателя. Все описанные в книге методы и алгоритмы были реализованы автором в различных программных комплексах и прошли практическую проверку.

Несмотря на отсутствие в книге программных кодов, она адресована программистам, той их части, кого принято называть математиками-программистами, основным содержанием деятельности которых является разработка алгоритмов при создании программной системы и для кого одной из главных «радостей ремесла» является «...очарование, заключенное в самом процессе создания сложных загадочных объектов, состоящих из взаимосвязанных непостоянных частей, и наблюдения за тем, как они работают в запутанных циклах, сохраняя верность принципам, заложенным в них с самого начала» [3].

Все — и заказчики, и пользователи, и разработчики программного обеспечения — понимают, что алгоритмы должны работать правильно и быстро. Но именно алгоритмист должен строго сформулировать в каждом конкретном программном проекте, что есть правильная работа алгоритма и каковы критерии оценки этой правильности. Применительно к рассматриваемым в книге задачам анализа формы изображений это означает, что нужно, в первую очередь, дать математически строгое определение формы изображений и способов ее вычисления и сравнения. Именно наличие строгой математической модели является главной основой для создания «правильных» алгоритмов, которые дают корректный результат и хорошо работают на практике, т. е. «не падают и не зацикливаются». А второй заботой математика-программиста является высокая скорость работы алгоритма. Среди алгоритмов, работающих правильно, хочется выбрать такой, который работает быстрее всех остальных. Но задача ускорения алгоритма становится осмысленной лишь в случае, когда понятие правильности работы алгоритма опирается на строгую математическую модель. Если же такой модели нет, то нет возможности проверки адекватности алгоритма. В этом случае алгоритм можно разгонять до любого предела: для плохого, но быстрого алгоритма всегда можно придумать алгоритм, работающий еще быстрее, но так же плохо.

Появлению этой книги способствовали многие люди, которым автор приносит свою искреннюю признательность. В первую очередь это академик РАН Ю. И. Журавлёв, с самого начала поддержавший данное научное направление и оказывающий ему внимание на протяжении многих лет. Это член-корреспондент РАН К. В. Рудаков, в сотрудничестве с которым проводилась значительная часть теоретических и прикладных исследований, описанных в книге. Важным источником идей и методов, вошедших в книгу, стали прикладные задачи из различных областей. Автор считает своим приятным долгом поблагодарить «постановщиков» этих задач: В. Н. Елшанского, Э. Л. Якупова, Т. У. Седерберга, С. В. Клименко, О. М. Черномордика, А. Г. Халтурина. Автор также благодарен своим ученикам И. А. Рейеру, А. Б. Семёнову, А. Г. Нефёдову, А. А. Масаловичу, А. К. Цискаридзе, совместная работа с которыми способствовала развитию идей и методов, описанных в этой книге.

Особую признательность автор приносит Российскому Фонду фундаментальных исследований за неоценимую поддержку работ по методам анализа и распознавания формы изображений начиная с 1996 года по настоящее время. Все основные результаты, помещенные в книге, получены в исследованиях по проектам, выполненным по грантам Фонда 96-01-00553, 99-01-00829, 02-01-00667, 05-01-00542, 08-01-00670. Также автор благодарен фонду CRDF (Civilian Research and Development Foundation), поддержавшему исследования по данной тематике в рамках проектов RM2-133 и RM2-2245 в 1996–2000 гг.

ГЛАВА 1

# ВВЕДЕНИЕ

1.1. *Бинарные изображения.* 1.2. *Принципы непрерывного подхода.* 1.3. *История вопроса.* 1.4. *Структура книги.*

## 1.1. Бинарные изображения

Бинарное изображение — это двухцветная картинка, на которой представлены один или несколько объектов одного цвета на фоне, имеющем другой цвет. Такие изображения представляются в компьютере в виде матрицы точек, каждая из которых может иметь лишь одно из двух значений цвета, условно обозначаемых через 0 или 1. Точки также называются пикселями (*pixel* — picture's element). Не нарушая общности, будем считать бинарное изображение «черно-белым»: точки объекта являются черными (значение цвета равно 1), а точки фона — белыми (значение цвета равно 0). Такая раскраска соответствует привычному способу рисования черной тушью на белой бумаге. Бинарные изображения требуют минимальных размеров памяти для хранения (достаточно одного бита на пиксел) и позволяют минимизировать расход других вычислительных ресурсов для решения задач обработки и анализа.

Бинарные изображения в системах машинного зрения формируются, как правило, в результате сегментации более сложных — полутоновых монохромных, либо цветных изображений, которые, в свою очередь, получаются в результате сканирования картин, документов, фотографий, а также ввода видеоинформации непосредственно с цифровых фото- и видеокамер (рис. 1.1).

В современных системах машинного зрения бинарные изображения играют очень важную роль. Они позволяют выделить объект интереса в изображении и являются основой для анализа и распознавания формы объектов в изображении [6, 26, 49].

Процесс преобразования цветных и полутоновых изображений в бинарные называется бинаризацией. Множество преобразований, используемых при бинаризации, весьма многочисленно и разнообразно [6, 49]. К ним относятся всевозможные методы точечных преобразований, свертки, усиления краев, выделения низкочастотных и высокочастотных компонент изображения и т. д. Результатом таких преобразований, в частности, является определение каких-то свойств точек изображения, по которым можно судить о том, принадлежат ли точки объекту или их следует отнести к фону. На основе проверки

Рис. 1.1. Примеры полутоновых изображений (вверху) и полученных на основе их обработки бинарных изображений (внизу)

выполнения этих свойств для всех точек изображения и создается, в конечном счете, бинарное двухцветное растровое изображение, в котором черные точки соответствуют объекту, а белые — фону. Заметим, что в явном виде построение такого бинарного изображения может и не осуществляться, но неявно оно обязательно присутствует и, в принципе, выделенный таким образом набор свойств объекта может быть визуализирован в виде черно-белого изображения.

Способы построения бинарных изображений выходят за рамки этой книги. Для нас существенно, что задача анализа формы объекта в изображении возникает после того, как этот объект выделен, т. е. определены те точки цифрового изображения, которые относятся к объекту. Поэтому мы рассматриваем бинарное изображение в качестве исходного материала для анализа формы.

## 1.2. Принципы непрерывного подхода

Несмотря на то, что изображение, представленное в виде матрицы точек, является чрезвычайно удобным для ввода, запоминания, обработки в компьютере, человеку привычнее и проще при описании формы объектов оперировать геометрическими понятиями. Простые геометрические фигуры, такие как многоугольники, круги, а также более сложные фигуры, например, ограниченные сплайновыми кривыми, однозначно понимаются всеми, имеют ясное математическое описание и могут быть легко использованы для представления и преобразования формы объектов. В то же время, описать какую-либо форму, например, круг или многоугольник, в терминах матрицы точек представляется намного более трудной задачей. Поэтому использование непрерывных моделей существенно упрощает создание алгоритмов анализа, классификации, преобразования формы объектов. Непрерывное описание формы соответствует интуитивному представлению человека о «сплошном» объекте, например, нарисованном краской на бумаге с помощью кисти. Изображение может быть при этом достаточно сложным в том смысле, что оно состоит из большого числа отдельных пятен, каждое из которых может иметь, в свою очередь, внутренние отверстия.

Такое различие в человеческом и машинном описании формы ставит разработчика алгоритмов анализа изображений перед дилеммой. Можно использовать дискртеное описание формы объекта в виде матрицы точек, рассчитывая на простоту компьютерной обработки таких данных. Но при этом нужно потратить большие усилия на то, чтобы переформулировать человеческие геометрические правила и алгоритмы описания, сравнения, преобразования формы объектов в операции над такими матрицами. Сложность такой алгоритмизации может оказаться значительной и с лихвой перекрыть выигрыш от простоты операций над матричными изображениями.

Другой путь состоит в преобразовании формы объектов из дискретного представления в виде матрицы точек в какую-либо непрерывную модель с тем, чтобы выполнять операции с формами в терминах, адекватных пониманию и опыту человека. В этом случае затраты на построение непрерывной модели достаточно высоки, но зато существенно упрощается создание алгоритмов анализа, сравнения и преобразования формы объектов, представленных на изображении.

Эти два похода, условно названные нами дискретным и непрерывным, являют собой крайние позиции, между которыми могут существовать смешанные варианты: та или иная часть обработки выполняется над матричным изображением, а затем можно осуществить переход к непрерывным формам.

Подход, который развивается в этой книге, занимает крайнюю «непрерывную» позицию в этом диапазоне. Он основан на идее как можно более раннего перехода от дискретного изображения формы к непрерывному аналогу. Предполагается, что на дискретном «матричном» уровне обработки должна быть решена лишь задача собственно бинаризации — выделения объекта интереса на изображении. Результатом такого выделения является бинарное изображение, в котором точки объекта имеют один цвет, а точки фона — другой (для определенности черный и белый соответственно). После этого предлагается по бинарному изображению немедленно построить непрерывное описание формы выделенных объектов. Такой переход от матричного представления к непрерывной модели будем называть дискретно-непрерывным преобразованием формы изображения.

Дискретно-непрерывное преобразование строит для бинарного изображения, представленного матрицей точек двух цветов, некоторое множество «непрерывных» примитивов, приближенно описывающих это изображение. Непрерывные примитивы — это геометрические фигуры, которые допускают простое аналитическое описание на плоскости — в двумерном евклидовом пространстве. Такое «непрерывное изображение» должно быть подобрано достаточно тонко, чтобы адекватно отразить особенности исходного дискретного изображения. Таким образом, в рамках нашего подхода дискретно-непрерывное преобразование представляет собой главный этап обработки изображения, который призван перевести изображение из формата, удобного для компьютерного ввода и хранения, в формат, удобный для человеческого понимания.

С точки зрения математики дискретно-непрерывное преобразование представляет собой аппроксимацию некоторого объекта одного класса (заданного черно-белой матрицей) объектом другого класса, обладающего свойством «непрерывности». Иногда подобную аппроксимацию называют сглаживанием. Задача сглаживания функции, имеющей низкую степень регулярности (например, разрывной или не дифференцируемой), состоит в аппроксимации ее другой, более регулярной функцией. При желании можно усмотреть в подборе непрерывного аналога для дискретного изображения именно задачу сглаживания некоторой функции.

Важную роль в реализации непрерывного подхода играет выбор класса непрерывных примитивов для аппроксимации формы дискретных объектов. В этой книге в качестве универсальной математической модели для описания формы объектов в изображении мы будем использовать понятие фигуры — замкнутой ограниченной области, граница которой состоит из конечного числа непересекающихся жордановых кривых. Такое описание вполне соответствует вербальному определению понятия формы из словаря русского языка: «внешние очертания, наружный вид предмета» [23]. Действительно, замкнутая ограниченная область имеет четко выраженные внешние очертания в виде границы, имеет внутреннюю часть, состоящую из всех внутренних точек области. Эта внутренняя часть отделена границей от внешнего окружающего пространства. И при этом «внешние очертания» могут иметь как очень простой, так и самый замысловатый и сложный вид.

Способность компьютера быстро оперировать с объектами, описываемыми замкнутыми областями, определяется тем, насколько удобно представлены эти области в числовом виде. Именно от формата представления объектов зависит скорость обработки информации в компьютере. А от скорости обработки, в свою очередь, зависит, можно ли решать задачу в реальном масштабе времени, обрабатывая, например, в секунду 5–10 кадров, снятых видеокамерой, либо на обработку одного кадра понадобится несколько минут. Соответственно, область применимости алгоритмов в системах анализа изображений принципиально меняется. Поэтому способам представления объектов в виде непрерывных моделей в книге уделено особое внимание. Предлагаемый подход состоит в использовании трех различных способов непрерывного представления фигур: граничного, скелетного и циркулярного (рис. 1.2).

Граничное представление фигуры состоит в описании границы в виде замкнутых кривых линий конечной длины. Граница однозначно определяет замкнутую область и, вместе с тем, хорошо характеризует ее внешние очертания, т. е. форму. Важную роль в таком представлении играет выбор способа численного описания границ области. Для

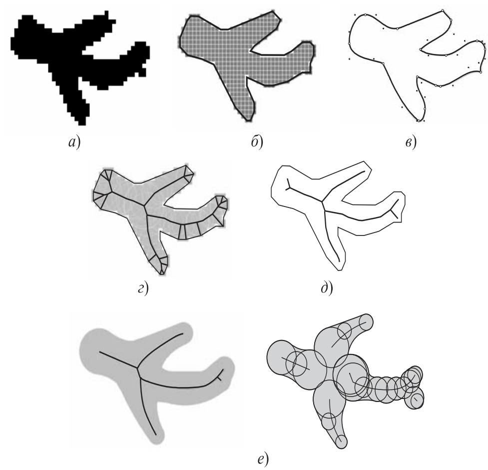

Рис. 1.2. Непрерывный подход к описанию формы бинарного изображения: дискретное бинарное изображение (*а*), полигональная аппроксимация границы (*б*), аппроксимация границы сплайнами (*в*), скелет многоугольной фигуры (*г*), регуляризация скелета (*д*), циркулярная фигура (*е*)

этого мы используем многоугольники, а также замкнутые сплайновые линии — составные кривые Безье.

Скелет замкнутой области, составляющий основу скелетного представления фигуры, — это множество точек, являющихся центрами максимальных окружностей, вписанных в эту область. Скелет фигуры представляет собой плоский граф, структура которого очень хорошо отражает особенности формы объекта. Дополнительную информацию для анализа формы дает радиальная функция скелета, которая ставит в соответствие каждой его точке радиус максимальной вписанной окружности с центром в этой точке. Описание скелета и радиальной функции составляет скелетное представление фигуры.

Циркулярное представление фигуры основывается на использовании однопараметрических семейств окружностей. При этом центры окружностей заполняют гладкие линии конечной длины, а радиус окружности является гладкой функцией параметра семейства. Циркулярное представление замкнутой области является развитием понятия скелета и дает возможность не только анализировать форму объекта, но и осуществлять ее преобразования. Потребность в таких преобразованиях возникает в задачах сравнения формы гибких объектов и в задачах синтеза новых форм на основе преобразования имеющихся эталонов.

## 1.3. История вопроса

Содержательно понятия границы и скелета объекта давно и широко используются в качестве базовых инструментов анализа формы изображений.

Граница, как основной дескриптор формы, является наиболее понятным и естественным инструментом для ее анализа. Вычисление границ в бинарном изображении никаких особых трудностей не вызывает. Обычно это множество точек объекта, имеющих соседние точки, не принадлежащие объекту. Соединив эти граничные точки, получаем многоугольники, описывающие контуры границы. Зная границу, можно вычислить периметр, площадь, выпуклую оболочку фигуры, ее топологическую структуру (количество дыр), определить характер края (сильно или слабо извилистый) и т.п.

Скелет является более сложной конструкцией, получение и использование которой требует решения гораздо более сложных задач. На протяжении сорока лет с момента выхода в свет первой работы по скелетам [28], показывающей широкие возможности их использования при анализе и распознавании изображений, опубликованы сотни статей по методам построения скелетов и их применениям в алгоритмах обработки изображений. Однако сам факт такого большого числа публикаций свидетельствует о том, что получить совершенные алгоритмы скелетизации, пригодные для любых реальных цифровых изображений, пока не удается.

Изначально определение скелета было сформулировано в непрерывном виде для замкнутых областей с непрерывной границей на евклидовой плоскости [28]. Естественно, что сразу же возник вопрос об адаптации этого понятия к растровым изображениям. Однако такая адаптация столкнулась с серьезными трудностями.

Первая фундаментальная проблема состоит в том, что для непосредственного использования понятий непрерывной геометрии в анализе дискретных изображений требуется корректно решить задачу аппроксимации границы дискретного объекта непрерывными линиями. Однако, несмотря на кажущуюся простоту, сделать это не всегда удается. В частности, упомянутый выше традиционный «наивный» способ представления границы бинарного изображения многоугольниками с вершинами во всех граничных точках не является корректным, поскольку полученная замкнутая линия не обязательно будет простым многоугольником. Оказывается, вообще невозможно математически корректно определить понятие границы изображения в терминах дискретной матрицы таким образом, чтобы сохранить адекватность с непрерывным аналогом. Причина этого состоит в невыполнении теоремы Жордана для линий на растре. Если считать, что в дискретной матрице, описывающей изображение, линия представляет собой цепочку смежных пикселей, то оказывается, что замкнутая растровая линия, вообще говоря, не всегда делит дискретную плоскость на два связных множества, для которых она является границей. Поэтому наивный подход к непрерывной аппроксимации границы изображений не дает хорошего геометрического описания формы. В частности, полученная таким образом граница может иметь самопересечения. Появление самопересечений сильно усложняет интерпретацию моделей формы, свидетельствует об их неадекватности.

Вторая проблема состоит в том, что не для всяких замкнутых областей, ограниченных жордановыми кривыми, существуют эффективные алгоритмы построения скелетов. Вычислительная геометрия предоставляет эффективные алгоритмы построения скелетов только для фигур, ограниченных простыми многоугольниками [37, 39]. Скелетизация фигур с нелинейной границей требует слишком больших затрат времени. Возникает необходимость сведения сложных многосвязных объектов с криволинейными границами к более простым фигурам с кусочно-линейными границами, для которых существуют эффективные алгоритмы скелетизации. Необходимо построить такую полигональную аппроксимацию объектов бинарного изображения, чтобы все многоугольники были простыми и не имели пересечений друг с другом. Выполнить это требование в сложных изображениях не так-то просто. Следует также отметить, что, несмотря на доказанное существование эффективных алгоритмов построения скелета сложной многосвязной фигуры, практическая разработка таких алгоритмов также является непростой задачей.

Наконец, существует еще одна проблема, связанная с интерпретацией полученного скелета аппроксимирующей многоугольной фигуры. Чем точнее аппроксимирует многоугольник границу фигуры, тем больше у него вершин и, соответственно, тем больше в его скелете шумовых ребер, не несущих никакой информации о форме объекта.

По-видимому, эти трудности привели к тому, что большинство известных решений задачи скелетизации бинарных изображений не использует мощных и элегантных методов вычислительной геометрии. Несмотря на большое число работ по скелетизации, до настоящего времени нам не известны другие примеры успешной реализации чисто непрерывного подхода, являющегося наиболее приближенным к геометрическому решению (непрерывная граница + непрерывный скелет).

Нерешенные проблемы непрерывного подхода привели к тому, что наибольшее распространение получили прагматические методы, основанные на чисто дискретном подходе (дискретная граница + дискретный скелет) [6, 29, 31, 49]. Дискретные методы не претендуют на полную адекватность с непрерывным аналогом, хотя и стремятся эту аналогию поддержать. В рамках дискретного подхода, как само собой разумеющееся, граница изображения определяется как множество точек объекта, имеющих соседние точки фона. А для скелета строгих определений вообще не предлагается. Обычно под скелетом понимается результат некоторого процесса преобразования изображения, например, так называемого утончения за счет перекрашивания граничных пикселей.

Всякий дискретный аналог корректной непрерывной модели хорошо работает «в пределе», т. е. в тех случаях, когда шаг дискретизации очень мал. Применительно к изображениям шаг дискретизации выражается в размерах пикселей. В изображениях с высоким разрешением, где размеры объектов в сотни раз превосходят размер пикселя, дискретные границы и скелеты часто оказываются вполне пригодными для анализа формы. Но в изображениях с низким разрешением их использование оказывается неприемлемым.

Недостатки дискретного подхода, служащие постоянным мотивом для его развития, отмечаются во многих работах. Дискретная граница в ряде случаев не позволяет определить топологическую структуру объекта — количество и вложенность граничных контуров, особенно для изображений с низким разрешением. Дискретные скелеты, полученные методом утончения, не являются серединными осями в евклидовой метрике, а скелеты, полученные с помощью так называемой дистанционной карты, оказываются не гомотопными исходному изображению, поскольку имеют другую связность.

Кроме этих недостатков, отметим еще одно важное обстоятельство. Дискретные границы и скелеты описываются в форме нового дискретного изображения, являющегося результатом преобразования исходного изображения. Для того чтобы решать задачи анализа, распознавания, сравнения и преобразования формы, требуется создание новых, часто нетривиальных алгоритмов обработки полученных изображений границ и скелетов. Например, для дискретных границ нужно «вытянуть» граничные пиксели в цепочки, т. е. упорядочить пиксели и образовать из них растровые линии. Нужно определить, какие линии задают внешние, а какие внутренние границы объектов. В скелете нужно выделить графовую структуру, т. е. найти пиксели — вершины графа и найти растровые линии, описывающие его ребра и т.п.

Кроме «чистого» дискретного подхода известны также варианты комбинированного промежуточного решения (дискретная граница + непрерывный скелет) [9, 46]. Метод [46] состоит в построении диаграммы Вороного для множества граничных точек изображения. Непосредственно в качестве скелета рассматривается некоторый подграф диаграммы Вороного, получаемый на основе процедуры регуляризации. Достоинство метода состоит в том, что полученный скелет является непрерывным, соответствует евклидовой метрике, имеет радиальную функцию, что определяет его полезность для анализа формы изображений. Второе важное достоинство этого метода состоит в его высокой скорости за счет использования эффективных алгоритмов вычислительной геометрии для построения диаграмм Вороного большой размерности. Недостатками метода являются отсутствие строгого определения понятия скелета, а также нерегулярность, зигзагообразность ветвей скелета. В методе [9] получается скелет в неевклидовой метрике и с нарушениями связности.

Следует отметить еще одну проблему, которую нужно решать при анализе формы изображений. Она связана с высокой чувствительностью границ и скелетов к шумовым эффектам в изображении. Небольшие изменения границы области могут привести к существенным изменениям топологической структуры скелета. Это приводит к тому, что для объектов, имеющих на взгляд человека очень высокую степень сходства, компьютерные алгоритмы строят весьма отличающиеся скелеты. Для того чтобы выделить в форме устойчивые свойства, инвариантные к шумовым эффектам, разрабатываются различные методы регуляризации границ [44, 52] и скелетов [27] для улучшения их свойств. Нужно сказать, что регуляризация необходима для всех скелетов, как непрерывных, так и дискретных. Вопрос состоит в том, насколько корректно ставится задача регуляризации и насколько удобно регуляризировать скелеты. Анализ работ по регуляризации дискретных скелетов показывает, что применяемые здесь методы, как правило, используют эвристические приемы и не имеют строгого математического обоснования.

## 1.4. Структура книги

Основной нашей целью является обоснование того, что полномасштабная реализация чисто непрерывного подхода к описанию формы изображений возможна и достижима. В книге описывается вариант полномасштабной реализации непрерывного подхода, включая построение корректных непрерывных границ и скелетов, строгих критериев их регуляризации. Кроме этого, предлагается циркулярный способ описания формы, близкий к скелетному графу, но ориентированный не только на анализ формы, но и на ее преобразование. Для обоснования практической реализуемости и полезности непрерывного подхода в книгу включены примеры решения практических задач с помощью описанных методов. Практическая реализация позволила также оценить скорость работы непрерывных алгоритмов и показать, что она на порядок и более превосходит быстродействие дискретных алгоритмов даже в варианте их параллельной реализации на графическом процессоре.

В соответствии с этим построена структура книги.

Первая часть «Формы и фигуры» посвящена анализу методов представления формы в изображениях. В ней описываются дискретные и непрерывные модели формы изображений. Дискретной моделью формы является дискретная фигура — связное множество точек растровой решетки. Связность определяется при этом в соответствии с выбранной структурой соседства пикселей в изображении. Непрерывная модель формы — это непрерывная фигура, представляющая собой замкнутую область с границей, состоящей из жордановых кривых. Демонстрируются возможности дискретного подхода и его недостатки.

Во второй части «Фигуры и границы» рассматривается задача получения непрерывных границ дискретных фигур. В основе решения лежит модель, включающая понятия граничного коридора дискретной фигуры и аппроксимирующего многоугольника, лежащего внутри граничного коридора. В качестве аппроксимирующего многоугольника выбрана замкнутая линия, разделяющая точки объекта и точки фона и имеющая минимальную длину, — так называемый *минимальный* *разделяющий многоугольник*. Показана единственность этого объекта для заданной структуры соседста точек квадратичной решетки. Описываются методы построения граничных коридоров и аппроксимирующих многоугольников для бинарных изображений любой сложности. Описан метод построения границы в виде замкнутых составных кривых Безье на основе аппроксимации многоугольников.

Третья часть «Границы и скелеты» описывает методы построения скелетов для многосвязных многоугольных фигур. Предложен эффективный алгоритм, основанный на построении графа смежности сторон и вершин многоугольников, составляющих границу фигуры. Этот граф, имеющий прямые аналогии с триангуляцией Делоне, назван *графом смежности многоугольной фигуры*. Смежными сторонами и вершинами многоугольной фигуры считаются такие, которые имеют общую касательную окружность, вписанную в фигуру. Задача скелетизации многосвязной фигуры, имеющей несколько граничных многоугольников, сводится к известной задаче скелетизации простого многоугольника за счет построения дерева смежности граничных контуров. Здесь же описывается метод регуляризации полученного скелета, основанный на строгом критерии сходства фигур в хаусдорфовой метрике.

Четвертая часть «Скелеты и циркуляры» посвящена описанию нового представления формы в виде так называемой циркулярной фигуры, для краткости, циркуляра. Циркуляр представляет собой объединение конечного числа жирных линий, которые, в свою очередь, определяются как семейства окружностей с центрами на гладких кривых и гладкой радиальной функцией. Циркуляр представляет собой обобщение понятия скелета, однако обладает дополнительными свойствами, дающими возможность преобразования формы изображений. Предложен метод построения циркуляров на основе скелетов, а также методы преобразования формы дискретных фигур на основе их циркулярного представления.

Наконец, пятая часть «Прикладные задачи» описывает приложения непрерывного подхода к решению практических задач анализа, сравнения и преобразования формы изображений. Рассматриваются задачи разработки компьютерных шрифтов, распознавания сканированных текстов, биометрической идентификации личности, компьютерной графики, геоинформационных технологий. Здесь же приводятся экспериментальные данные по результатам сравнения предлагаемых методов с известными решениями с точки зрения вычислительной эффективности.

# ЧАСТЬ I
# ФОРМЫ И ФИГУРЫ

## ГЛАВА 2
## НЕПРЕРЫВНЫЕ МОДЕЛИ ФОРМЫ

2.1. *Содержательное понятие формы.* 2.2. *Фигура как модель формы.* 2.3. *Граничное представление фигуры.* 2.4. *Скелетное представление фигуры.* 2.5. *Характеристические свойства скелета.* 2.6. *Вычисление и регуляризация скелета.*

### 2.1. Содержательное понятие формы

Восприятие и последующая обработка электронных волн зрительной системой — базовая функция в деятельности живых существ в их взаимодействии с окружающей средой. В результате длительной эволюции развились зрительные системы, которые могут регистрировать резкие изменения яркости и цвета в световых потоках, характерные для границ, отделяющих объекты от окружающего пространства и других объектов. Разнообразная геометрия граничных линий объектов, видимо, привела к человеческому понятию формы.

Точного однозначного определения понятия формы нет. Это понятие из тех, которые являются базовыми, легко понимаются всеми, но формулируются немногими и с большим трудом. Обычно формулировки имеют вид метафор через сходство (круглый как луна, извилистый как змея, кольцеобразный, в форме верблюда и т. д.) Для человеческого общения точная формулировка понятия формы не очень востребована, но для алгоритмов анализа и распознавания изображений строгие определения всех понятий необходимы. Поэтому определение формы должно быть сформулировано в виде общих математических моделей. Адекватность этих моделей, определяемая как непротиворечивость практическому опыту человека, служит подтверждением правильности сформулированного понятия в тех границах, в которых этот опыт имеется.

Абстрагируясь от конкретных физических свойств объекта, человек обычно представляет его форму в виде сплошного пространственного тела в трехмерном пространстве либо в виде плоского объекта с непрерывной границей в двумерном пространстве. Поскольку картины трехмерного мира воспринимаются через двумерные проекции на сетчатку глаза, да и системы машинного зрения также регистрируют изображения в виде двумерных картин, первичной является именно двумерная модель формы. Трехмерные модели реконструируются впоследствии человеческим либо машинным интеллектом на основе двумерных проекций.

Среди многочисленных содержательных понятий слова «форма» в словаре русского языка приводится одно, наиболее подходящее для использования его при выборе математической модели. *Форма* — это внешние очертания, наружный вид предмета [23].

В этом же словаре приводятся значения близкого по смыслу термина «фигура». *Фигура* — внешнее очертание, форма чего-либо. Там же для слова фигура указан еще один вариант с пометкой «математический термин» — это часть плоскости, ограниченная замкнутой линией. Эти значения слов фигура и форма вполне могут быть положены в основу математической модели понятия формы для использования в системах обработки и анализа изображений.

## 2.2. Фигура как модель формы

Наиболее распространенные и доступные математические модели формы объектов реального мира дает евклидова геометрия. Геометрические фигуры на плоскости и тела в пространстве, изученные еще в школе, предоставляют общий язык для описания формы окружающих объектов и инженерам, и художникам, и домохозяйкам. Сама по себе потребность людей использовать некоторый понятийный аппарат для определения формы предметов окружающей действительности явилась, по-видимому, одним из важных источников для создания и развития геометрии в древности. Однако строгие геометрические очертания многоугольников, окружностей, параллелограммов или трапеций не исчерпывают всего многообразия форм, которые выделяет и классифицирует человеческий глаз. Силуэт человека, кленовые или дубовые листья, военный или пассажирский самолеты — эти и другие изображения легко воспринимаются человеком как объекты определенной формы, присущей именно им и отличающей их от других объектов.

Для описания всего множества плоских «сплошных» форм требуется более широкое общее определение. С этой точки зрения приемлемой общей математической моделью представляется понятие фигуры. Понятие фигуры, ключевое для описания моделей формы, мы сформулируем более расширительно по сравнению с упоминавшимся определением из словаря русского языка. Все определения даются применительно к евклидовой плоскости $R^2$.

**Определение 2.1.** *Жордановой кривой* называется образ окружности при непрерывном инъективном ее отображении в евклидову плоскость [20]. Отображение инъективно, если любые две различные точки прообраза отображаются в две различные точки образа.

**Определение 2.2.** *Фигурой* называется связная замкнутая область на плоскости, ограниченная конечным числом непересекающихся жордановых кривых.

Будем использовать следующие обозначения. Пусть $R^2$ — евклидова плоскость с расстоянием $d(p,q)$ между точками $p$, $q \in R^2$.

Фигура $A$ — это некоторое точечное множество на плоскости $R^2$. Граница фигуры определяется как множество точек

$$\partial A = \big\{p\colon\ p \in R^2, \forall\, r > 0,\ S_r(p) \cap A \neq \varnothing \text{ и } S_r(p) \cap \overline{A} \neq \varnothing\big\},$$

где $S_r(p) = \big\{q\colon\ q \in R^2,\ d(p,q) < r\big\}$ — открытый круг радиуса $r$ с центром в точке $p$, $\overline{A} = R^2 \backslash A$ — дополнение фигуры $A$ в пространстве $R^2$.

Внутренностью фигуры называется множество $\widehat{A} = A \backslash \partial A$, состоящее из внутренних точек фигуры.

Таким образом, фигура в нашем понимании — это связная часть плоскости, ограниченная одной или несколькими замкнутыми линиями. Под это определение подходят сплошные объекты с конечным числом дыр. С одной стороны, разнообразие этих объектов неисчерпаемо, а с другой, существует несколько способов их конструктивного и ясного математического описания. Поэтому математическую модель фигуры мы выберем в качестве базовой для представления понятия формы.

Фигура представляет собой замкнутую ограниченную область. Как известно, *областью* на евклидовой плоскости называется непустое связное открытое множество точек. *Замкнутой областью* называется замыкание области, т. е. минимальное замкнутое множество, содержащее область. Другими словами замкнутая область — это объединение некоторой области (открытой) и всех ее предельных точек. Область называется *ограниченной*, если она содержится целиком в некотором прямоугольнике.

Не всякая ограниченная замкнутая область является фигурой в смысле определения 2.2. В примере на рис. 2.1 изображены замкнутые области, представляющие собой квадрат с треугольным отверстием.

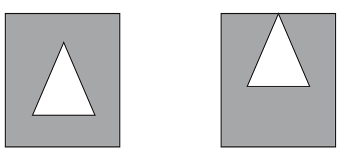

Рис. 2.1. Замкнутые области: слева — фигура, справа — нет

На левом рисунке треугольник целиком лежит внутри квадрата, а на правом одна из вершин треугольника лежит на границе квадрата. Замкнутая область, изображенная слева, является фигурой, поскольку имеет границу, состоящую из двух непересекающихся жордановых кривых. А замкнутая область справа не является фигурой. Если