

Местецкий Л.М.

Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры

Москва, Физматлит, 2009. 288 стр.

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 причёски 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 определённое противоречие между человеческим и компьютерным представлением формы объектов. Особенностью компьютерного представления видеoinформации 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. Решить 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 приёмы, 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 существенно последовательные шаги, число которых растёт 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 остаётся 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 её 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 приёмы –

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, применённые 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 всё-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 простоту 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* остаётся 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 её 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 её 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 приёмов, то в рамках непрерывного подхода появляется возможность её строгого решения эффективными алгоритмами.

Четвёртая 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 путём сравнения с эталонами, но при сравнении нужно обеспечить максимально возможное совпадение с эталоном с учётом 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 изменённого таким образом семейства окружностей, мы получаем изменённый объект, имеющий другую форму. Правда в этом новом объекте скелет уже может не совпадать с изменённым 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 автором, 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, 41-42], 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, заключённое 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 её 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 заклиниваются». 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 ещё 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 поддерживавший 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, помещённые в книгу, получены в исследованиях по проектам, выполненным по грантам Фонда 96-01-553, 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 г.г.