

Московский государственный университет имени М. В. Ломоносова

Ломов

На правах рукописи

Ломов Никита Александрович

**Морфологические дескрипторы объектов переменной
ширины на цифровых изображениях**

Специальность 05.13.17 —
«Теоретические основы информатики»

Диссертация на соискание учёной степени
кандидата физико-математических наук

Научный руководитель:
доктор технических наук, профессор
Местецкий Леонид Моисеевич

Москва — 2020

Оглавление

	Стр.
Введение	4
Глава 1. Описание ширины на основе непрерывного скелета	17
1.1 Описание ширины средствами математической морфологии	17
1.2 Непрерывные модели для описания ширины	24
1.3 Непрерывная морфология бинарных изображений	27
1.4 Скелет многоугольной фигуры и бициклы	31
1.5 Выводы к главе 1	38
Глава 2. Метод построения дескрипторов формы на основе дискового покрытия	40
2.1 Дисковое покрытие и морфологическая ширина	40
2.2 Собственные области и внешние сектора	43
2.3 Усечённые бициклы	45
2.4 Пересечения бициклов	49
2.4.1 Общая структура пересечений	50
2.4.2 Поиск смежных бициклов	65
2.4.3 Оценки числа смежных бициклов	70
2.4.4 Лакуны дискового покрытия	78
2.5 Структура алгоритма и анализ вычислительной сложности	86
2.6 Эксперименты с гибкими объектами	88
2.7 Выводы к главе 2	92
Глава 3. Производные дескрипторы ширины	94
3.1 Частная морфологическая ширина	95
3.2 Морфологическая ширина полутонового изображения	110
3.3 Карты морфологических признаков	120
3.3.1 Расчёт карты расстояний	121
3.3.2 Расчёт карты толщин	127
3.4 Морфологические моменты многоугольных фигур	134
3.4.1 Задача пространственного описания ширины	134

3.4.2	Концепция морфологических моментов	137
3.4.3	Инварианты морфологических моментов	139
3.4.4	Непрерывный алгоритм вычисления моментов	144
3.4.5	Морфологические моменты примитивных областей	145
3.5	Выводы к главе 3	152
Глава 4. Практические приложения алгоритмов вычисления признаков ширины		154
4.1	Распознавание компьютерных шрифтов	154
4.1.1	Описание задачи	154
4.1.2	Сравнение шрифтов на основе диаграмм ширины	155
4.1.3	Эксперименты	159
4.2	Компьютерная морфометрия планарий	163
4.2.1	Описание задачи	163
4.2.2	Метод решения	164
4.2.3	Эксперименты	168
4.3	Выводы к главе 4	170
Заключение		172
Список литературы		174

Введение

Диссертационная работа посвящена исследованию и разработке математических моделей универсального описания ширины объектов на бинарных изображениях. Предложенные модели способны отражать во всей полноте распределение ширины в тех случаях, когда ширина объекта варьируется в различных его частях, а форма объекта может изменяться путём гибких деформаций. Разработаны высокоэффективные вычислительные процедуры для расчёта дескрипторов, позволяющих сравнивать такие объекты по ширине и классифицировать их.

Актуальность темы. В настоящее время активно развиваются методы компьютерного зрения, которые находят своё применение в системах видеонаблюдения, медицине, геоинформатике и других областях. При этом одной из наиболее популярных задач в этой сфере является выделение и классификация объектов: людей, животных, транспортных средств и т. д.

Для решения указанной задачи зачастую используются методы, работающие либо с отдельными полноцветными изображениями, либо с их последовательностями: в частности, входные данные такого типа являются стандартными для нейросетевых методов. Можно заметить, что такое представление в некотором смысле является избыточным: и в отсутствие цветовой информации человек способен определить класс объекта, основываясь на знании только о его форме. Это подтверждают исследования в области психологии [1; 2], причём меньше всего от отсутствия информации о цвете страдают изображения живых существ — животных или людей. Ещё одним аргументом в пользу сведения этой задачи к распознаванию формы служит тот факт, что в обиход входят приборы, позволяющие напрямую получить удобное представление формы: 3D-сканеры, сенсоры для получения дальностных изображений, трекеры скелетов.

Входные данные для этой задачи могут быть достаточно разнообразными не только в силу различий между самими объектами, но и в силу того, что один и тот же объект может находиться в разных позах, быть зафиксирован с различных ракурсов, образовывать самоокклюзии и т. д. С точки зрения описания формы это означает, что мы имеем дело с так называемыми гибкими артикуляциями. При всём разнообразии гибких артикуляций интуитивно ясно, что

есть свойства формы, стабильные относительно воздействия таких артикуляций: это свойства, связанные с шириной отдельных частей объекта. К примеру, при движении человеческих рук ширина самих рук остаётся неизменной.

Человек руководствуется информацией о ширине в тех случаях, когда форма состоит из одних и тех же семантических составляющих: к примеру, можно отличать породы собак по пропорциям их отдельных частей тела: туловища, лап, хвоста. По ширине листа можно определить, какому растению принадлежит этот лист, хотя форма листьев может быть в целом достаточно похожей. Тем не менее человек оперирует понятием ширины не только в случае гибких объектов: так, мы различаем обычное и полужирное начертания одного и того же шрифта, судя по ширине его штрихов, конструкторские детали сложной формы могут иметь переменную ширину, отличие силуэтов самолётов или иных транспортных средств, снятых с подходящего ракурса, также может быть обнаружено по ширине их составляющих частей.

Это позволяет сделать вывод, что признаки, связанные с шириной объекта, сообщают нам важную информацию о его форме, поэтому разработка эффективных математических моделей ширины объекта выглядит актуальной задачей.

Задача описания ширины объекта затрагивает целый спектр областей математики: от теории множеств и численных методов до алгоритмов вычислительной геометрии. Первую проблему создаёт неопределённость самого понятия ширины. Толковый словарь русского языка под редакцией С. И. Ожегова и Н. Ю. Шведовой [3, стр. 896] даёт неожиданно проницательное с точки зрения математики определение ширины — это «величина, протяжённость чего-нибудь в том направлении, в котором две крайние точки плоскости, тела лежат, в отличие от длины, на наименьшем расстоянии друг от друга». Хотя с точки зрения этого определения ширина объекта является одним из его габаритов и описывается единственным числом, в определении точно отмечены сущностные особенности ширины — это расстояние между крайними точками фигуры (то есть лежащими на её границе), а направление выбрано таким образом, что расстояние оказывается минимальным. Такой взгляд на ширину погружает задачу её описания в круг методов, которые описывают форму объекта с помощью распределения расстояний между точками его границы или точками, полученными из них с помощью геометрических построений.

В методе описания формы контура с помощью *функции высоты дуги* [4] рассматриваются пары точек, между которыми расположены части контура фиксированной длины. Эти точки связываются отрезками, к которым проводятся серединные перпендикуляры. Значением функции высоты считается расстояние от середины отрезка до ближайшего пересечения серединного перпендикуляра с отсечённой частью контура (расстояние может быть и отрицательным в зависимости от полуплоскости, в которой расположена точка пересечения). Большие значения функции высоты детектируют значительные изгибы формы (выпуклые или вогнутые в зависимости от знака). Варьируя длину части контура, заключённой между точками, можно получить набор функций высоты, который используется для описания формы контура.

В *методе высотных функций* [5] на границе фигуры равномерно отмечаются точки, к каждой из них проводится касательная, и вычисляются расстояния от каждой точки до каждой касательной. Таким образом, длина дескриптора оказывается равной квадрату числа выбранных точек. Другой возможный способ описания формы контура — *расстояние от точки до центра масс фигуры* при движении по контуру с постоянной скоростью [6]. Так как этот дескриптор сам по себе является достаточно простым, он, как правило, используется на этапе предобработки изображения для извлечения более сложных признаков, например, при помощи преобразования Фурье [7]. Преобразование Фурье можно применить и к исходным координатам контура объекта, сэмплированным с равномерным шагом и рассматриваемым как комплексные числа [8]. В этом случае форму объекта можно воспринимать как совокупность сжатий, растяжений и утолщений в отдельных её частях.

Наиболее общий подход к построению дескрипторов формы такого рода был предложен в работе [9] под названием *распределений формы*: он анализирует распределения различных характеристик (углов, расстояний, площадей) при случайном выборе точек на границе или внутри формы, причём рассматривает такие распределения с точки зрения теории вероятностей. Для некоторых простых форм — например, для треугольника или круга — оказывается возможным вывести необходимые функции распределения аналитически. Также ряд рассмотренных выше дескрипторов формы приводится к общему знаменателю в работе [10], где рассматриваются способы описания формы в виде функций, полученных в результате анализа контура плоского объекта. Кроме того, предлагается ряд функций ширины, выраждающих зависимость некоторых

характеристик формы — например, длины проекции или сечения, вычисленных относительно опорной прямой, — от направляющего угла прямой.

Контекст формы был предложен в работе [11] как метод сравнения контуров объектов, основанный на поиске соответствия между точками двух контуров. Каждая точка описывается гистограммой распределения расстояний от этой точки до остальных точек контура в полярной системе координат, причём для радиусов используется разбиение гистограммы на ячейки в логарифмической шкале. В качестве меры сходства распределений между точками используется критерий χ^2 . Расстоянием между дескрипторами двух изображений является сумма таких статистик при сопоставлении точек «один-к-одному». Контекст формы является весьма эффективным дескриптором в том случае, если форма меняется незначительно. В том случае, когда форма претерпевает гибкие артикуляции, дескриптор в его изначальном виде является довольно неустойчивым. Однако форма частей контура должна сохраняться с точностью до поворота, поэтому было предложено вычислять *контексты для характерных частей контура* [12], равномерно выбирая на них фиксированное число точек и предварительно поворачивая фрагмент контура в соответствии с углом наклона касательной в его центре. Другой возможный подход к преодолению сложностей, связанных с гибкими деформациями, — рассматривать пути, лежащие полностью внутри фигуры. В *методе внутренних расстояний* [13] расстояние между точками определяется как длина кратчайшего пути такого рода. Авторы предполагают возможность замены традиционного евклидова расстояния на внутреннее в дескрипторах формы с повышением качества работы систем распознавания, что демонстрируется на примере контекста формы.

Разнообразие изложенных выше методов описания формы через набор расстояний обеспечивается различным способом задания расстояний и выбором точек интереса. Все они, однако, позволяют составить представление о ширине фигуры в разных её частях. Важно отметить, что почти все упомянутые методы рассматривают только точки контура и считают их равноправными, используя равномерное сэмплирование по контуру. Наиболее эффективный в практических приложениях метод — контекст формы — анализирует не только распределение расстояний, но и распределение направлений в выбранных точках, что в ещё большей степени соответствует словарному определению.

Ряд техник анализа формы объектов на изображениях опирается на теорию математической морфологии Серра [14], в рамках которой вводятся опе-

рации преобразования изображения с помощью некоторого примитива формы, называемого структурирующим элементом. Общий принцип задания форморазмерных дескрипторов был разработан Матероном и известен под названием *гранулометрического анализа* [15]. Для описания гранулометрии можно воспользоваться метафорой просеивания. Гранулометрией называется семейство фильтров, параметризованных неотрицательными значениями, для которых выполняются следующие свойства: антиэкстенсивность (просеивание уменьшает количество зёрен), свойство возрастания (просеивание подмножества вложено в просеивание всего множества) и стабильность (результат просеивания через два последовательных сита соответствует просеиванию через более крупное). Можно считать, что при очередном просеивании удаляются компоненты или их части, имеющие недостаточную ширину. Наиболее распространённой морфологической операцией при определении гранулометрий является операция открытия, заключающаяся в последовательном сужении и расширении объекта относительно структурирующего элемента.

Гранулометрии получили широкое использование в практических приложениях, в частности, при анализе медицинских изображений. Отметим, что входными данными для методов этого типа могут служить и полутоновые изображения, поэтому гранулометрии можно рассматривать также как метод анализа текстур. Разнообразие гранулометрий достигается путём выбора различных структурирующих элементов, а также способа подсчёта числа оставшихся после просеивания элементов. К примеру, в работе [16] предлагается оставлять компоненту целиком, если хотя бы одна её точка выдержала очередное просеивание. В терминах ширины это означает, что ширина компоненты во всех точках считается одинаковой и соответствует размеру её наибольшей части. Примечательно, что хотя теоретически гранулометрии могут быть определены для множеств на пространстве \mathbb{R}^2 , в практических приложениях речь, как правило, идёт о множествах на дискретной решётке. Как будет показано далее, это приводит не только к погрешностям измерения и подсчёта, но и делает сами гранулометрические операции некорректными, так как нарушается свойство стабильности.

Наиболее известный способ построения гранулометрического дескриптора как функции распределения размеров фигуры в различных её точках был предложен Марагосом [21]. Благодаря работам Ю. В. Визильтера [22–24], он получил в русскоязычной литературе название *морфологического спектра*.

Помимо этого, известен ряд модификаций традиционного морфологического спектра, основанный на различиях в способе выбора точек интереса, определения значения размерной функции и группировки этих значений. Несколько таких модификаций предложены в работах Шекара. В частности, в одной из его работ [25] в качестве точек интереса рассматриваются только точки скелета, а размером считается значение радиальной функции. В другой работе [26] значением размера или ширины считается радиус максимального круга, центр которого расположен в данной точке и который вписан в блоки изображения, полученные после разбиения. В третьей работе [27] рассматриваются только точки на границе объекта, а значением ширины является расстояние от точки до охватывающего фигуру круга, центр масс которого совпадает с центром масс фигуры.

Важная связь между морфологическими операциями и одним из важнейших понятий морфологического анализа изображений, скелетом [17], была отмечена ещё на ранних этапах развития математической морфологии [18; 19]. Скелет — это способ описания фигуры в виде множества её серединных осей, по сути являющихся некой основой, каркасом формы. При этом не только сам скелет можно получить в результате морфологических операций, но и, зная, на каком этапе какая часть скелета получена, можно восстановить результат операции открытия. Это приводит нас к идее не просто хранить скелет в виде набора точек, но и связать с каждой точкой некоторый размерный параметр, то есть задать функцию на множестве точек скелета. Способ описания формы с помощью скелета и размерной (радиальной) функции называется медиальным представлением [20]. Радиальная функция каждой точке скелета ставит в соответствие радиус максимального вписанного в фигуру круга с центром в данной точке. Это значение можно рассматривать как ширину объекта в окрестности точки скелета.

На такой способ описания формы следует обратить особое внимание, поскольку он остаётся устойчивым при гибких артикуляциях, если только они не приводят к самоокклюзиям. Действительно, при подобных деформациях топологическая структура скелета, длины его ветвей и значения радиальной функции на них меняются незначительно. Таким образом, медиальное представление является весьма удобным описанием формы фигуры для дальнейшей работы с её шириной. С этой точки зрения любой метод, который оперирует

медиальным представлением, учитывая радиальную функцию, можно считать способом описания ширины.

Существует класс сложных дескрипторов формы, извлекающих информацию о ширине из медиального представления, которые нашли применение при анализе формы объектов, сильно варьирующихся в позе и артикуляции своих частей, а также способных образовывать самоокклюзии. К ним относятся *мешок скелетных путей* [28], который записывает последовательность значений радиальной функции при движении от одной терминальной точки скелета к другой, *мешок ассоциированных со скелетом контурных фрагментов* [29], в котором извлекаются контурные фрагменты, и каждой точке контура соответствует набор признаков: расстояние до центра фрагмента, угол поворота и длина спицы, проходящей через данную точку (а по сути ширина фигуры в точке контура). *Скелетный контекст* [30] похож на контекст формы, но рассматриваются точки не контура, а скелета, для которых строятся те же гистограммы распределения расстояний и ориентаций. Примечательно, что данные методы извлекают набор локальных характеристик ширины, но не дают её интегрального описания, а отсутствие учёта «полной картины» компенсируется использованием сложных техник машинного обучения.

Таким образом, несмотря на разнообразие подходов к описанию ширины и тем более конкретных методов построения дескрипторов, само понятие ширины практически нигде в явном виде не формализуется и остаётся достаточно расплывчатым. Поэтому имеет смысл дать этому понятию строгое математическое наполнение уже на первых этапах исследования.

В результате создание математически элегантных моделей ширины фигуры в сочетании с эффективной процедурой их получения и разработка мощных дескрипторов на их основе составляют во многом открытую проблему, и этот факт определяет актуальность данного исследования.

Целью данной работы является разработка математической модели формы объекта на бинарном изображении, удобной для интегрального описания ширины этого объекта, и создание на базе данной модели методов и алгоритмов вычисления ширины, обладающих высокой вычислительной эффективностью. В качестве «отправной точки» рассматривается медиальное представление формы многоугольной фигуры, полученное в результате скелетизации на основе диаграммы Вороного.

Научные задачи работы заключаются в следующем:

1. Формализация понятия ширины объекта сложной формы с точки зрения как локального (ширина фигуры в точке), так и глобально-го (функция распределения ширины объекта) описания. Разработка морфологических дескрипторов ширины объектов на бинарных изображениях, основанных на введённой формализации.
2. Исследование и разработка математических моделей, позволяющих вычислять предложенные дескрипторы, разработка соответствующих алгоритмов вычисления, обладающих высокой вычислительной эффективностью и точностью. Доказательство корректности разработанных алгоритмов.
3. Адаптация разработанных моделей и алгоритмов для вычисления дескрипторов ширины с особыми условиями: в случае полутоновых входных данных, наличия у объекта сегментации, необходимости учёта пространственного расположения объекта и с целью построения мультимодальных дескрипторов формы.
4. Практическая реализация разработанных методов построения моделей и оценки ширины объектов, проведение вычислительных экспериментов для проверки корректности полученных результатов.
5. Проверка применимости разработанных методов оценки ширины для описания гибких объектов, оценка устойчивости методов к гибким артикуляциям.
6. Применение разработанных методов для специализированных прикладных задач компьютерного зрения, связанных с распознаванием плоских объектов на изображениях по ширине.

Научная новизна. В диссертации предложены новые математические модели, позволяющие анализировать форму объектов переменной ширины на бинарном изображении. Разработаны новые алгоритмы вычисления признакового описания формы, характеризующего ширину объекта, опирающиеся на данные модели. Рассмотрены вариации разработанных моделей и дескрипторов для случаев различных типов входных данных и особенностей постановки задачи. Предложены алгоритмы для расчёта дистанционного преобразования фигуры и вычисления ширины фигуры в точке на основе медиального представления формы и решения оптимизационной задачи. В силу этого примеры применения предложенных методов к решению практических задач также являются новыми.

Разработанные методы значительно превосходят по вычислительной эффективности существующие аналоги. В отличие от большинства существующих методов работа ведётся с непрерывной, а не растровой моделью формы. Предложенные методы способны работать с бинарными изображениями любой сложности, а источником погрешности в измерении ширины выступает только аппроксимация многоугольными фигурами формы объектов на растровых изображениях. При этом для самих многоугольных фигур ширина определяется с точностью, ограниченной лишь точностью компьютерных вычислений.

Научная значимость работы состоит в разработке методов вычисления дескрипторов ширины для сравнения и классификации объектов по форме. Предложен подход, позволяющий сравнивать объекты между собой по ширине, используя их медиальное представление. Такой подход не зависит от сложности и особенностей этого представления. Изложенная в работе методика предоставляет математический аппарат для конструирования новых дескрипторов формы изображения, использующих данные о ширине, и методов их сравнения.

Практическая значимость состоит в разработке детальной программной реализации всех предложенных методов, а также их приложений к задаче распознавания шрифтов по изображениям текста и оценке морфогенеза плоских червей. Для решения практических задач разработан соответствующий программный комплекс.

Методология и методы исследования. В работе использованы методы вычислительной геометрии, математической морфологии, теории графов, методов оптимизации, обработки и анализа изображений, теории сложности алгоритмов. Работа носит экспериментально-теоретический характер. Проведены эксперименты на модельных данных, открытых базах изображений и изображениях из реального мира, полученных в результате фотосъёмки или сканирования. Также исследованы приложения предлагаемых подходов к практическим задачам (распознавание шрифтов и оценка скорости регенерации плоских червей), которые подробно описаны в заключительной главе.

Основные положения, выносимые на защиту:

1. Морфологические дескрипторы объектов на цифровых изображениях, основанные на непрерывном медиальном представлении формы: площадь и моменты дискового покрытия многоугольных фигур, описывающих форму объектов.

2. Метод вычисления площади дискового покрытия многоугольной фигуры и функции морфологической ширины как интегрального описания ширины фигуры.
3. Метод вычисления геометрических моментов дискового покрытия многоугольной фигуры и функции морфологических моментов как интегрального описания пространственного распределения ширины фигуры.
4. Методы вычисления производных дескрипторов ширины — площади покрытия сегментированной фигуры и совокупной интенсивности системы дисковых покрытий полутонового изображения.
5. Методы вычисления точечных дескрипторов ширины бинарного изображения — карты расстояний (distance transform) и карты толщин — на базе непрерывного медиального представления формы объектов изображения.
6. Методы идентификации шрифта по изображению текста и оценки скорости регенерации плоских червей-планарий на основе разработанных дескрипторов ширины.

Достоверность полученных результатов обеспечивается их теоретическим обоснованием, а также подтверждается результатами вычислительных экспериментов, проведённых при помощи программного обеспечения, реализующего предложенные методы.

Апробация работы. Результаты работы докладывались на следующих конференциях:

- 24-я международная конференция по компьютерной графике, обработке изображений и машинному зрению «Графикон-2014» (Ростов-на-Дону, 2014) [31];
- 26-я международная конференция по компьютерной графике, обработке изображений и машинному зрению «Графикон-2016» (Нижний Новгород, 2016) [32];
- 11-я международная конференция «Интеллектуализация обработки информации» ИОИ-2016 (Барселона, 2016) [33; 34];
- 8-я научно-техническая конференция «Техническое зрение в системах управления» ТЗСУ-2017 (Москва, 2017) [35];

- Международная конференция «Техники фотограмметрии и компьютерного зрения для задач видеонаблюдения, биометрии и биомедицины». PSBB-2017 (Москва, 2017) [36];
- 12-я международная конференция по теории и приложениям компьютерного зрения VISAPP 2017 (Порту, 2017) [37];
- 27-я международная конференция компьютерной графике, обработке изображений и машинному зрению «ГрафиКон-2017» (Пермь, 2017) [38];
- 18-я Всероссийская конференция с международным участием «Математические методы распознавания образов» ММРО-2017 (Таганрог, 2017) [39].

Автор принимал активное участие в работе научного семинара Л. М. Местецкого «Непрерывные морфологические модели и алгоритмы» (факультет вычислительной математики и кибернетики МГУ). Результаты научного-исследовательской работы были представлены на московском семинаре «Морфологический анализ данных» под руководством Ю. П. Пытьева (физический факультет МГУ), международном семинаре «Распознавание RGB-D изображений в системе компьютерного зрения» и в рамках совместного российско-индийского научного проекта (доклады в университетах городов Хайдарабад, Майсур и Мангалор, декабрь 2017).

Материалы диссертации использовались в работах по грантам РФФИ №№ 14-01-00716, 15-07-01323 и 17-01-00917.

Личный вклад. Все результаты, выносимые на защиту, получены автором самостоятельно. Постановка задачи была выполнена совместно с научным руководителем. В совместных публикациях в трудах конференции «ГрафиКон-2014» [31] и «Приволжском научном журнале» [40] автор внёс вклад в теоретическое обоснование и практическую реализацию метода, в публикации в трудах конференции «ГрафиКон-2017» [38] автором разработаны методы, использующие описание ширины, и проведены эксперименты с ними.

Публикации. Основные результаты по теме исследования изложены в 12 печатных изданиях, 3 из которых изданы в журналах, рекомендованных ВАК [40—42], 2 опубликованы на английском языке в изданиях, индексирующихся системами Scopus и Web of Science [36; 37], 3 — в сборниках трудов конференций [31; 32; 38], 4 — в сборниках тезисов докладов [33—35; 39].

Предлагаемый подход к построению дескрипторов формы, чувствительных к изменению ширины и инвариантных к изгибу объекта, основан на идеях использования медиального представления формы в виде множества срединных осей (скелета) и радиальной функции, заданной на этих осях [43]. На основе медиального представления исследуется модель дискового покрытия многоугольной фигуры и площадь дискового покрытия как базовый дескриптор ширины. Далее рассматриваются вариации этого дескриптора в случаях особой постановки задачи: задания области интереса или множества особых точек, необходимости учесть пространственную информацию или скомбинировать информацию о ширине с признаками других типов. Практическая применимость разработанных методов демонстрируется на примере реальных задач, кроме того, описание каждого метода снабжено вычислительными экспериментами, демонстрирующими их надежность и вычислительную эффективность.

Изложение вышеперечисленных этапов исследования определяет **структуру** диссертации.

Первая глава является вводной, в ней описан общий подход к интегральному описанию ширины фигуры с помощью математической морфологии, изложены базовые принципы и понятия непрерывной морфологии бинарных изображений и рассмотрены основные методы построения дескрипторов ширины на базе контурно-скелетного представления формы.

Во второй главе вводится понятие дискового покрытия многоугольной фигуры, анализируется его форма при заданном размере круга, описывается способ его последовательного преобразования при увеличении круга и метод расчёта дескриптора ширины — площади дискового покрытия.

В третьей главе рассматриваются производные дескрипторы ширины, опирающиеся на модель дискового покрытия. Описан метод вычисления дескриптора ширины для сегментированной фигуры, предложено обобщение дискового покрытия в виде системы покрытий для полутонового изображения, а также введено обобщение понятия морфологической ширины в виде концепции морфологических моментов. Кроме того, разработаны методы расчёта растровых карт признаков ширины и предложен общий принцип построения мультимодальных дескрипторов формы, учитывающих информацию о ширине.

Четвёртая глава посвящена практическим приложениям разработанных методов: распознаванию шрифтов из обширной базы и оценке скорости

регенерации плоских червей-планарий. Проведены эксперименты по оценке скорости работы реализованных методов и её сравнению с имеющимися аналогами.

В заключении подводятся итоги работы.

Объем и структура работы. Диссертация состоит из введения, четырёх глав и заключения. Полный объём работы составляет 185 страниц, включая 80 рисунков и 6 таблиц. Список литературы содержит 115 наименований.

Глава 1. Описание ширины на основе непрерывного скелета

Ширина объектов является важной характеристикой формы изображений. Для объектов сложной формы, у которых разные части имеют разную ширину и протяжённость, эта характеристика не может быть хорошо описана скалярной величиной, например, «средней» шириной. Поэтому в качестве дескриптора ширины требуется описание «распределения» ширины, которое характеризует весь спектр её значений.

Данная глава посвящена обзору математических средств — теоретических оснований, способов представления и преобразования формы, алгоритмических приёмов — предназначенных для определения подобных дескрипторов и их эффективного вычисления. Далее приведена структура главы. Раздел 1.1 содержит основные положения математической морфологии и общий принцип построения «функций распределения» ширины, заданный на их основе. В разделе 1.2 выделены основные направления разработки вычислительно эффективных процедур для реализации требуемых морфологических преобразований, отмечены перспективы использования непрерывных представлений формы. Способ перехода от дискретной бинарной сцены к эквивалентной ей непрерывной описан в разделе 1.3. В разделе 1.4 представлены непрерывные морфологические модели, позволяющие извлекать информацию о ширине, а также проверяется соответствие известных способов описания ширины на базе непрерывных моделей требованиям классической математической морфологии. Завершает главу раздел 1.5 с выводами.

1.1 Описание ширины средствами математической морфологии

Математическая морфология, разработанная Серра [14] и Матероном [15], представляет собой математический аппарат для анализа и обработки геометрических структур, основанный на теории множеств. Пусть дано произвольное множество E , на множестве подмножеств которого $\mathcal{P}(E)$ определено отношение включения (\subseteq), а также операции объединения (\cup) и пересечения (\cap). Рассмотрим некоторое преобразование $\Psi : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ (оператор Ψ).

Пусть $X, Y \in \mathcal{P}(E)$. Оператор Ψ называется:

- *возрастающим*, если $(X \subset Y) \Rightarrow \Psi(X) \subseteq \Psi(Y)$;
- *убывающим*, если $(X \subset Y) \Rightarrow \Psi(Y) \subseteq \Psi(X)$;

- *экстенсивным*, если $\psi(X) \supseteq X$;
- *антиэкстенсивным*, если $\psi(X) \subseteq X$;
- *усиливающим*, если $\psi(\psi(X)) \supseteq \psi(X)$;
- *ослабляющим*, если $\psi(\psi(X)) \subseteq \psi(X)$;
- *идемпотентным*, если $\psi(\psi(X)) = \psi(X)$.

Оператор ψ называется *дилатацией* (*расширением*), если он сохраняет объединение:

$$\psi(\cup X_i) = \cup\psi(X_i), \quad \forall X_i \in \mathcal{P}(E),$$

и *эрзией* (*сжатием*), если он сохраняет пересечение:

$$\psi(\cap X_i) = \cap\psi(X_i), \quad \forall X_i \in \mathcal{P}(E).$$

Также оператор ψ называется *закрытием*, если он является экстенсивным, возрастающим и идемпотентным, и называется *открытием*, если он является антиэкстенсивным, возрастающим и идемпотентным. Если операторы φ и ψ являются дилатацией и эрзией соответственно, то операторы $\psi\varphi$ и $\varphi\psi$ могут служить примерами, соответственно, операторов закрытия и открытия.

Общий принцип построения форморазмерных описаний множества X известен под названием *гранулометрии* [44].

Определение 1.1. Гранулометрией называется семейство операторов открытия $\Psi = \{\psi_\lambda\}$, параметризованных неотрицательной величиной λ (размером), обладающее свойством стабильности:

$$\psi_\lambda\psi_\mu(X) = \psi_\mu\psi_\lambda(X) = \psi_{\max(\lambda,\mu)}(X), \quad \lambda, \mu \geq 0.$$

Определение 1.2. Рассмотрим преобразования множества X гранулометрией Ψ . Гранулометрической функцией называется отображение $G_\Psi : X \rightarrow \mathbb{R}_+$, $\mathbb{R}_+ = \{a \in \mathbb{R} \mid a \geq 0\}$, определяемое как

$$G_\Psi(x) = \max\{\lambda \mid x \in \psi_\lambda(X)\}. \quad (1.1)$$

Отметим, что следствием антиэкстенсивности и стабильности операторов из семейства $\{\psi_\lambda\}$ является их монотонность по λ :

$$\psi_\mu(X) = \psi_\mu\psi_\lambda(X) = \psi_\mu(\psi_\lambda(X)) \subseteq \psi_\lambda(X), \quad \text{если } \mu \geq \lambda.$$

Из этого также следует, что

$$\begin{cases} x \in \psi_\lambda(X), & \text{если } \lambda \leq G_\Psi(x), \\ x \notin \psi_\lambda(X), & \text{если } \lambda > G_\Psi(x). \end{cases}$$

Определение 1.3. Гранулометрической кривой называется зависимость некоторой меры m множества $\psi_\lambda(X)$ от размерного параметра:

$$F_\Psi(\lambda) = m(\psi_\lambda(X)). \quad (1.2)$$

Очевидно, что из монотонности Ψ по λ следует, что $F(\lambda)$ не возрастает.

Представим, что X — фигура, образованная множеством точек из E , а ψ_λ выделяет в фигуре X достаточно широкую часть (область заданной ширины λ). Тогда требования, предъявляемые к гранулометрии, получаются весьма логичными и довольно ясно интерпретируемыми:

- антиэкстенсивность ($\psi(X) \subseteq X$) означает, что любой оператор в гранулометрии действительно выделяет *часть* X ;
- возрастание ($(X \subset Y) \Rightarrow \psi(X) \subseteq \psi(Y)$) означает, что ширина большей фигуры должна быть больше ширины меньшей;
- стабильность означает, что последовательное выделение областей заданной ширины эквивалентно выделению самой широкой области.

В такой интерпретации можно считать, что гранулометрическая функция $G_\Psi(x)$ определяет локальную ширину фигуры в точке $x \in X$, а область заданной ширины $\psi_\lambda(X)$ состоит из точек, в которых локальная ширина не меньше λ . Гранулометрическую кривую же можно рассматривать как функцию распределения ширины по точкам фигуры.

Для задач распознавания изображения особый интерес представляют операции математической морфологии, определённые для подмножеств плоскости \mathbb{R}^2 или решётки \mathbb{Z}^2 с помощью структурирующих элементов. Далее мы будем вести речь только о таких операциях.

Более формально, пусть X — двумерная фигура на плоскости \mathbb{R}^2 , а B — структурирующий элемент с центром в начале координат $(0; 0)$ этой плоскости. Для определения операций морфологии Серра нам также понадобится *перенос множества $X \subset \mathbb{R}^2$ на вектор z* : $X_z = \{y \mid x \in X, y = x + z\}$. Этую операцию можно считать частным случаем *суммы Минковского*, которая образует множество C , состоящее из сумм всевозможных векторов из A и B :

$$C = \{c \mid c = a + b, a \in A, b \in B\}.$$

Базовые операции морфологии Серра со структурирующим элементом определяются следующим образом.

Эрозия множества $X : X \ominus B = \{z : B_z \subset X\}$. На рис. 1.1 X — вся фигура, B — круг, $X \ominus B$ — закрашенная область.

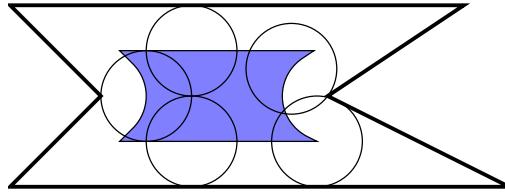


Рисунок 1.1 — Эрозия множества X

Дилатация множества $X : X \oplus B = \{x + b : x \in X, b \in B\}$. На рис. 1.2 X — закрашенная область, B — круг, $X \oplus B$ — вся фигура.

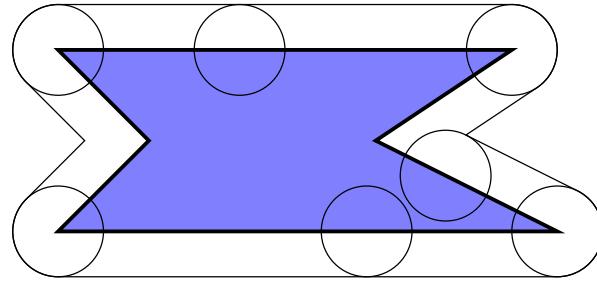


Рисунок 1.2 — Дилатация множества X

Операции открытия и закрытия в морфологии Серра получаются как производные от базовых.

Открытие множества $X : X \circ B = (X \ominus B) \oplus B$. На рис. 1.3 граница X обозначена жирной линией, граница $X \ominus B$ — пунктиром, $X \circ B$ — закрашенная область.

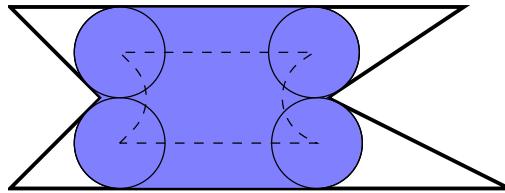
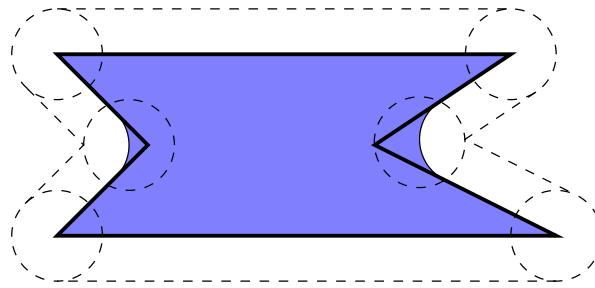


Рисунок 1.3 — Открытие множества X

Закрытие множества $X : X \bullet B = (X \oplus B) \ominus B$. На рис. 1.4 граница обозначена жирной линией, граница $(X \oplus B)$ — пунктиром, $X \bullet B$ — закрашенная область.

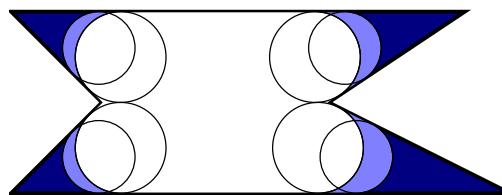
Пусть структурирующий элемент B масштабирован по параметру $r \geq 0$: $B(r) = \{rb \mid b \in B\}$. Тогда на основе операции открытия множества определяется также *морфологический спектр Марагоса* (pattern spectrum) множества X :

Рисунок 1.4 — Закрытие множества X

$$\begin{cases} PS(r) = -\partial S(X \circ B(r))/\partial r, & r \geq 0, \\ PS(-r) = \partial S(X \bullet B(r))/\partial r, & r > 0, \end{cases} \quad (1.3)$$

где $S(X)$ — площадь фигуры X , $PS(r)$ — спектр положительной части оси r (спектр объектов изображения), $PS(-r)$ — спектр отрицательной части оси r (спектр фона изображения). Можно считать, что $S(X \circ B(r))$ — это количественная мера элементов $B(r)$ в X . Таким образом, морфологический спектр представляет собой морфологическую технику, которая описывает изменение площади открытия при небольших изменениях размера примитива, то есть вклад примитивов каждого размера в форму изображения. Особую роль при этом играют дисковые структурирующие элементы $B(r) = \{b \mid \|b\| \leq r\}$, так как они обеспечивают инвариантность спектра к повороту.

На рис. 1.5 X — вся фигура. Тёмно-синим закрашено множество $X \setminus (X \circ B(r))$, светло-синим — множество $(X \circ B(r)) \setminus (X \circ B(r + \Delta r))$ для некоторых $r > 0$ и $\Delta r > 0$.

Рисунок 1.5 — Открытие множества X с использованием примитивов разного радиуса

Из книги [15] известно, что семейство открытий с масштабированным по r структурирующим элементом B определяет гранулометрию, если B — выпуклое множество. Соответственно, спектр Марагоса в этом случае представляет собой производную гранулометрической кривой со знаком минус. Можно также рассматривать гранулометрическую кривую и спектр Марагоса как функцию и плотность распределения ширины в точках фигуры X . Таким образом,

математическая морфология предоставляет мощный аппарат для описания распределения переменной ширины объектов.

Ключевую роль для нашей работы играет понятие скелета, представляющего собой множество срединных осей фигуры. Пока определим его конструктивно на основе морфологических операций, в виде формулы, приведённой в [18]:

$$\mathcal{S}(X) = \bigcup_{r \geq 0} \mathcal{S}_r(X), \quad \mathcal{S}_r(X) = \bigcap_{q > 0} [(X \ominus B(r)) - (X \ominus B(r)) \circ B(q)]. \quad (1.4)$$

Здесь $B(r)$ — диск радиуса r , взятый с границей, а X — замкнутое множество¹. Пример скелета показан красным на рис. 1.6.

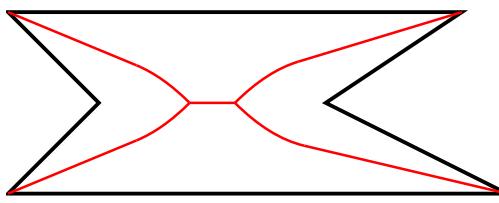


Рисунок 1.6 — Скелет множества X

Важнейшим свойством скелета является его способность восстановить само изображение и весь набор результатов операции открытия $\{X \circ B(r)\}, r \geq 0$:

$$X \circ B(r) = \bigcup_{q \geq r} \mathcal{S}_q(X) \oplus B(q). \quad (1.5)$$

Это свойство используется для сжатия и восстановления объектов на бинарных изображениях и результатов морфологических операций с ними [19; 45].

Тем не менее в практических приложениях изображения редко хранятся и анализируются в «векторном» виде, как множество фигур на \mathbb{R}^2 — как правило, работа ведётся с матрицами точек на дискретной сетке \mathbb{Z}^2 . Поэтому приходится иметь дело с дискретизацией семейства структурирующих элементов $\{B(r)\}$ и дискретной процедурой вычисления спектра:

$$\begin{cases} PS(i) = -S(\mathbf{X} \circ \mathbf{B}(i)) - S(\mathbf{X} \circ \mathbf{B}(i+1)), & i \geq 0 \\ PS(-i) = S(\mathbf{X} \bullet \mathbf{B}(i)) - S(\mathbf{X} \bullet \mathbf{B}(i-1)), & i > 0 \end{cases} \quad (1.6)$$

где \mathbf{X} и $\mathbf{B}(i)$ — соответствующие дискретные множества, площадь которых S определяется как число их точек. Рассматривают только целые значения размеров i , а дисковые структурирующие элементы обычно также составляют из

¹Построение скелета относительно произвольных выпуклых структурирующих элементов также возможно.

точек, удалённых от центра на расстояние, не превышающее i , но эти точки теперь имеют только целые координаты.

Несмотря на то что такая дискретизация кажется естественной, она может приводить к довольно неприятным эффектам. На рис. 1.7 изображены дискретные дисковые структурирующие элементы размеров 1 и 3, которые мы обозначим $B(1)$ и $B(3)$, а также результат открытия большего круга меньшим. Нетрудно видеть, что $B(3) \circ B(3) = B(3)$, но $B(3) \circ B(1) \neq B(3)$, и следовательно, $S(B(3) \circ B(1)) < S(B(3) \circ B(3))$. В результате в спектре фигуры $B(3)$ появляются отрицательные значения, и его больше нельзя интерпретировать как оценку вклада примитивов заданного размера в образование формы.

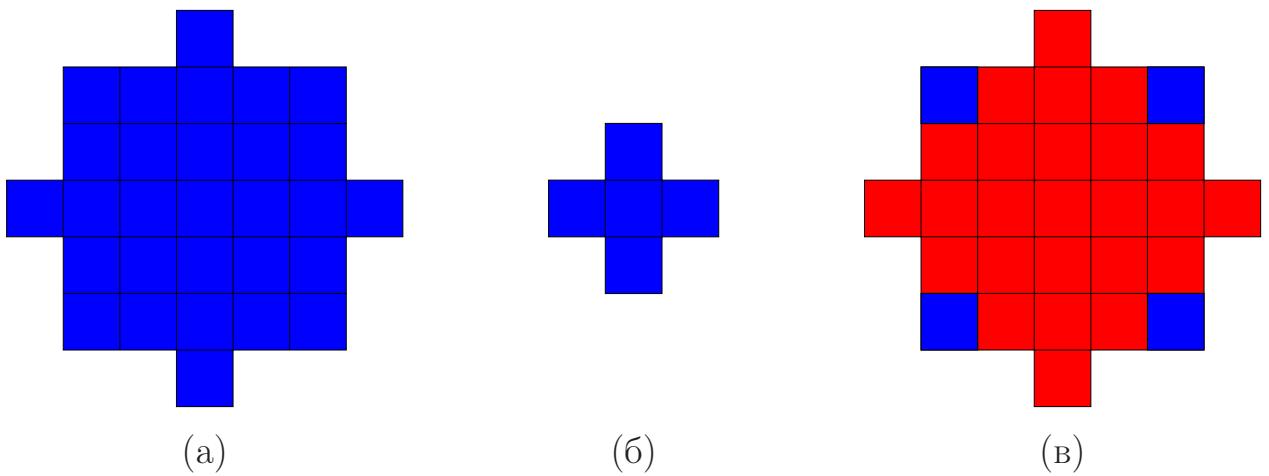


Рисунок 1.7 — (а-б) Дискретные структурирующие элементы, (в) открытие большего элемента меньшим (красные пиксели)

Строго говоря, можно постановить, что открытие с таким семейством элементов не определяет гранулометрию, так как семейство открытий не обладает свойством стабильности:

$$(B(3) \circ B(3)) \circ B(1) \neq B(3) \circ B(3) = B(3) \circ B(\max(1,3)).$$

Для борьбы с этим эффектом дисковый структурирующий элемент может быть аппроксимирован многоугольником, что обеспечивает монотонность результата² и допускает более эффективную программную реализацию [46], однако сам результат оказывается вычислен с существенной погрешностью. Идея метода заключается в возможности разложить центральносимметричный многоугольник в сумму Минковского своих сторон (рис. 1.8) и далее произвести

²К примеру, такой подход используется в системе MATLAB. Примечательно, что в документации к системе (<https://www.mathworks.com/help/images/ref/strel.html>) прямо сказано, что структурирующие элементы, не использующие аппроксимацию, не подходят для вычисления гранулометрий.

последовательность операций эрозии и дилатации с линейными структурирующими элементами (рис. 1.9). Отметим, что в практической реализации все элементы по-прежнему остаются растровыми.

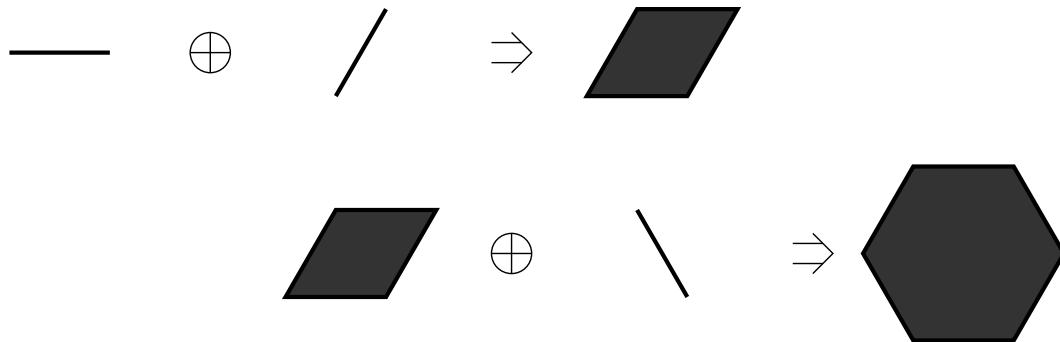


Рисунок 1.8 — Разложение многоугольника в сумму Минковского своих сторон

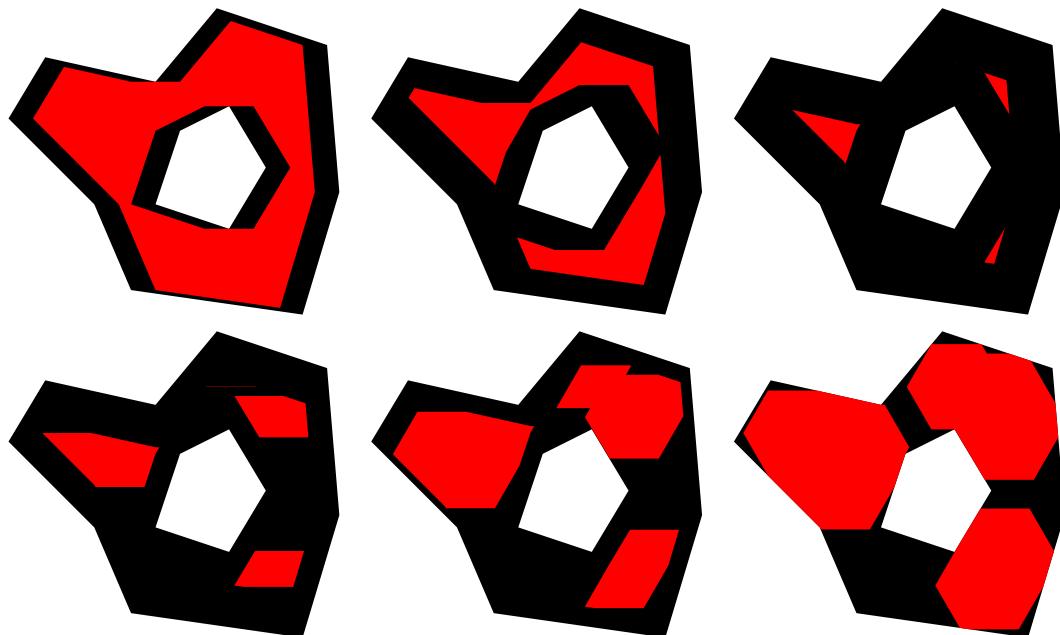


Рисунок 1.9 — Реализация операции открытия с многоугольником через последовательность эрозий (верхний ряд) и дилатаций (нижний ряд) с линейными элементами

Таким образом, задача точного, быстрого и математически корректного вычисления морфологического спектра по-прежнему остаётся актуальной.

1.2 Непрерывные модели для описания ширины

Морфологический спектр является популярной техникой для извлечения информации о ширине объектов на изображении. Дескриптором ширины объекта может служить диаграмма морфологического спектра, которая строится на основе операции морфологического открытия с использованием дискового

структурирующему элемента. Пример использования этого подхода для решения задачи распознавания личности по форме ладони описан в [47]. Также известны примеры использования морфологического спектра для распознавания личности по форме губ [48] и обнаружения человеческих фигур на видеопоследовательностях [49], которые, однако, используют прямоугольные структурирующие элементы, что требует фиксированного положения камеры.

Как правило, нужный дескриптор вычисляется напрямую — путём последовательного применения к растровому изображению морфологических операций, что приводит к простой программной реализации. В силу этого перспективы ускорения процедуры вычисления спектра традиционно связывались с оптимизацией базовых морфологических операций. Лучший из дискретных алгоритмов описан в [50], он основан на разбиении структурирующего элемента на хорды — отрезки из идущих подряд пикселей в строке или столбце, принадлежащих элементу (рис. 1.10) — и использовании специальной структуры данных для хранения результата преобразования с каждой из хорд. Дальнейшие улучшения алгоритма связаны в основном с аппаратными подходами: адаптацией для графического [51] и векторного процессора, распараллеливанием, реализацией на программируемых логических интегральных схемах [52]. Это свидетельствует о том, что эффективность базового алгоритма достигла определённого предела, но по-прежнему является недостаточной, особенно при работе с большими изображениями высокого разрешения.

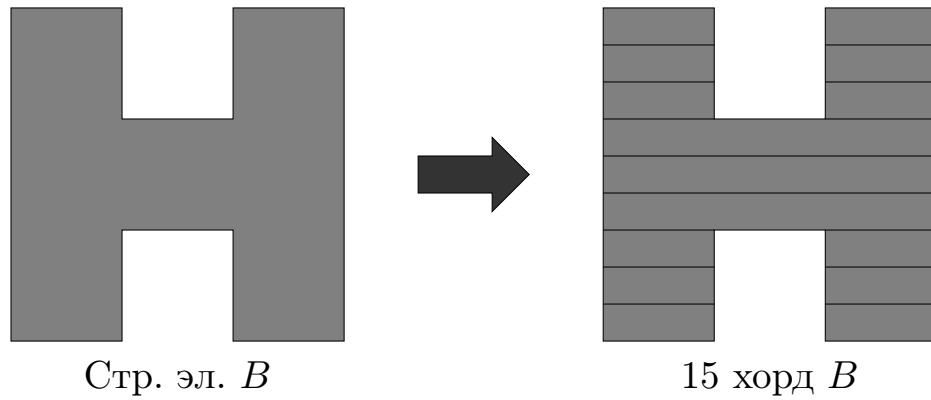


Рисунок 1.10 — Разбиение структурирующего элемента на хорды

Упомянутый ранее подход с разложением структурирующего элемента в сумму Минковского представляет особый интерес, поскольку может быть адаптирован для «векторного» случая при приближении формы многоугольными фигурами — достаточно модифицировать широко известный алгоритм Бентли–Оттманна [53], вместо события пересечения пары отрезков отслеживая

другое событие — интервал между соседними отрезками в фигуре (её дыре) стал равен длине стороны многоугольника. Это означает, что либо фигура стала в этом месте достаточно узкой, и интервал полностью устранился эрозией, либо дыра фигуры стала достаточно узкой, и интервал полностью заполняется дилатацией. Таким образом, этот алгоритм «прокладывает мостик» от дискретной математической морфологии к непрерывной.

Теоретические принципы морфологических преобразований фигур, заданных непрерывными граничными описаниями, разработаны достаточно давно. В работе [54] показано, что базовые морфологические операции можно осуществить путём преобразования контурного представления формы — сдвига точек контура по направлению внешней нормали для дилатации и в противоположном направлении для эрозии. Отдельную сложность представляет обработка самопересечений в результирующей кривой. Достаточно развиты алгоритмы, реализующие морфологические операции с многоугольниками, при этом, так как оперировать выпуклыми многоугольниками значительно легче, был разработан аппарат *отрицательных форм* [55], который позволяет представить невыпуклый многоугольник в виде объединения выпуклых многоугольников разных знаков, а также предложены эффективные алгоритмы для требуемой декомпозиции [56]. Однако использование подобных методов затрудняется высокой сложностью программной реализации, кроме того, необходимость приближать многоугольниками структурирующие элементы сильно ограничивает точность. Лишь совсем недавно появился пример использования в практическом приложении морфологических операций с непрерывными моделями [57], который опирается на концепцию прямолинейного скелета [58]. Принцип работы метода заключается в сдвиге вершин исходного многоугольника в направлении либо смежных сторон вершины, либо их бисектора. Тем не менее, так как, во-первых, прямолинейный скелет является некоторым приближением «точного» скелета, во-вторых, сами операции также производятся с некоторым приближением, вычисление спектра ширины на базе этого метода может привести к значительным ошибкам.

Другое направление развития спектров ширины связано с непрерывной морфологией бинарных изображений, разработанной Л. М. Местецким [43], в рамках которой был создан алгоритм построения точного непрерывного скелета многоугольной фигуры. Эта модель используется в работах [22; 59], однако после построения скелета его рёбра и вписанные круги снова растеризуются, так

что данный подход правильнее назвать дискретно-непрерывным. Его применение позволило существенно сократить время вычислений, однако не настолько, чтобы задачу можно было решать в реальном времени работы систем компьютерного зрения. Стоит также отметить модификацию этого комбинированного подхода для расчёта спектров с эллиптическими элементами [60], принцип которого заключается в аффинном преобразовании исходной фигуры, построении для неё скелета, а затем возвращении к исходной системе координат, в результате чего вписанные круги превращаются во вписанные эллипсы.

Полностью непрерывный подход был исследован в работе [61], где были получены аналитические формулы для спектров бициклов — областей, связанных с отдельными ребрами скелета. Однако исследование показало, что спектр всей фигуры нельзя представить как сумму спектров составляющих фигуру бициклов, поскольку открытия бициклов пересекаются между собой. Таким образом, для точного вычисления морфологического спектра необходим анализ этих пересечений, который был проведён лишь для простейших случаев. Этот анализ осложнялся тем, что в работе рассматривались не сами открытия бициклов, а их изменения при бесконечно малом изменении радиуса примитива.

В работах [40; 62] была предпринята попытка построения так называемой медиальной ширины, также основанной на непрерывной модели формы. Этот подход обеспечил требуемую вычислительную эффективность за счёт упрощения вида областей заданной ширины и возможности представить общую площадь области как сумму вкладов отдельных бициклов. Однако диаграмма медиальной ширины оказалась слишком чувствительной к результату промежуточной операции стрижки скелета. Отметим также, что операция извлечения области медиальной ширины, строго говоря, не является операцией открытия.

В итоге, хотя модель непрерывного скелета многоугольной фигуры выглядит весьма перспективной для использования в задаче описания ширины, проблема создания точного, вычислительно эффективного и устойчивого алгоритма вычисления ширины до сих пор окончательно не решена.

1.3 Непрерывная морфология бинарных изображений

В памяти компьютера изображения, как правило, хранятся в растровом формате, определяющем прямоугольную сетку точек (пикселей), обладающих определённым цветом или интенсивностью. Растровое изображение в оттенках серого можно рассматривать как двумерный массив (матрицу), изображение с

несколькими цветовыми каналами (например, использующее цветовую модель RGB или HSV) — как трёхмерный массив. Однако о наличии объектов на изображении, их положении и особенностях формы можно вести речь только в том случае, когда решена задача сегментации — для всех пикселей изображения известно, являются ли они пикселями объекта или пикселями фона. Таким образом, отправной точкой для задачи распознавания формы объектов в нашем случае являются бинарные изображения, в которых все пиксели принимают значения 0 — пиксель принадлежит фону, или 1 — пиксель принадлежит объекту.³ Объектом на бинарном изображении можно считать компоненту связности в четырёх- или восьмисвязной структуре соседства. Однако, как мы отметили выше, основной моделью формы для нас будет служить многоугольная фигура — многоугольник с многоугольными дырами. Поэтому возникает задача аппроксимации границ объектов на бинарном изображении многоугольниками.

Естественно потребовать, чтобы в результате аппроксимации получалась непрерывная сцена, эквивалентная дискретной: содержащая то же число объектов и сохраняющая их взаимное расположение, включая возможную вложенность одного объекта в другой, и общие очертания формы. Используемый способ построения эквивалентной непрерывной сцены описан в [43] и базируется на следующих принципах:

- граф смежности пикселей является триангуляцией, которая включает в себя 4-связную структуру соседства. Это означает, что граф разбивает область изображения на треугольники, причём две его стороны соединяют соседние в строке или столбце точки, а третья сторона проведена по диагонали;

- два смежных по диагональной стороне треугольника образуют ячейку.

Ячейка считается граничной, если она содержит точки разного цвета. Граничная ячейка считается внешней, если относится к нечётному уровню вложенности граничных контуров, и внутренней — если к чётному. Граничная ячейка имеет приоритет точек определённого цвета, если её диагональ соединяет пару точек этого цвета, либо, если такой пары нет — пару разноцветных точек. Все внешние граничные ячейки имеют приоритет чёрного цвета, а все внутренние — приоритет белого цвета;

³Далее для удобства визуализации сегментированных изображений пиксели объекта мы будем обозначать чёрным цветом, хотя это расходится с традиционным способом вывода на экран бинарных изображений, при котором чёрным цветом отображаются пиксели, имеющие значение 0.

- внутренние чёрные (не имеющие соседних белых) пиксели оказываются внутренними точками объекта, внутренние белые — внутренними точками фона;
- граничные чёрные и белые точки могут оказаться либо внутренними для объекта и фона соответственно, либо лежать на границе объекта и фона;
- для аппроксимации используются разделяющие многоугольные контуры минимального периметра.

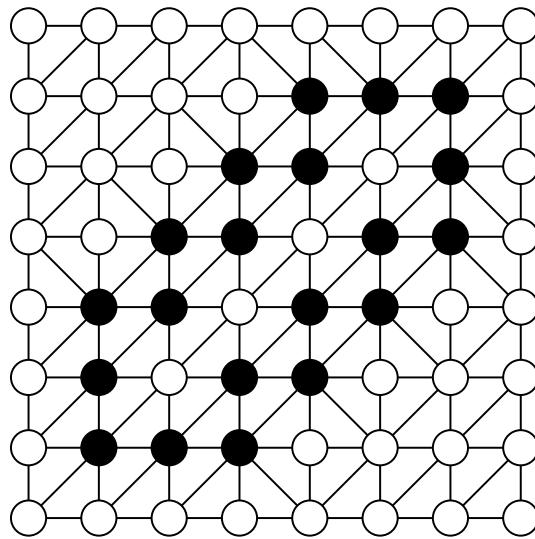


Рисунок 1.11 — Связные компоненты при компонентной смежности

Пример построения компонентной структуры смежности для заданного изображения приведён на рис. 1.11. Такая нетривиальная структура зависит от конкретного изображения, позволяет считать единым целым группы пикселей соприкасающиеся только по диагонали (то есть 8-связные, но не 4-связные) как для самих объектов, так и для дыр в них, и избегать дробления компонент.

Процедура поиска и прослеживания границ состоит из следующих этапов:

1. Обнаружение граничного коридора — связного множества треугольников в структуре соседства, соответствующего отдельному замкнутому контуру границы.
2. Прослеживание граничного коридора — построение последовательности из чёрных и белых граничных точек, определяющих форму коридора.
3. Построение минимального разделяющего многоугольника — многоугольника минимальной длины, лежащего внутри граничного коридора.

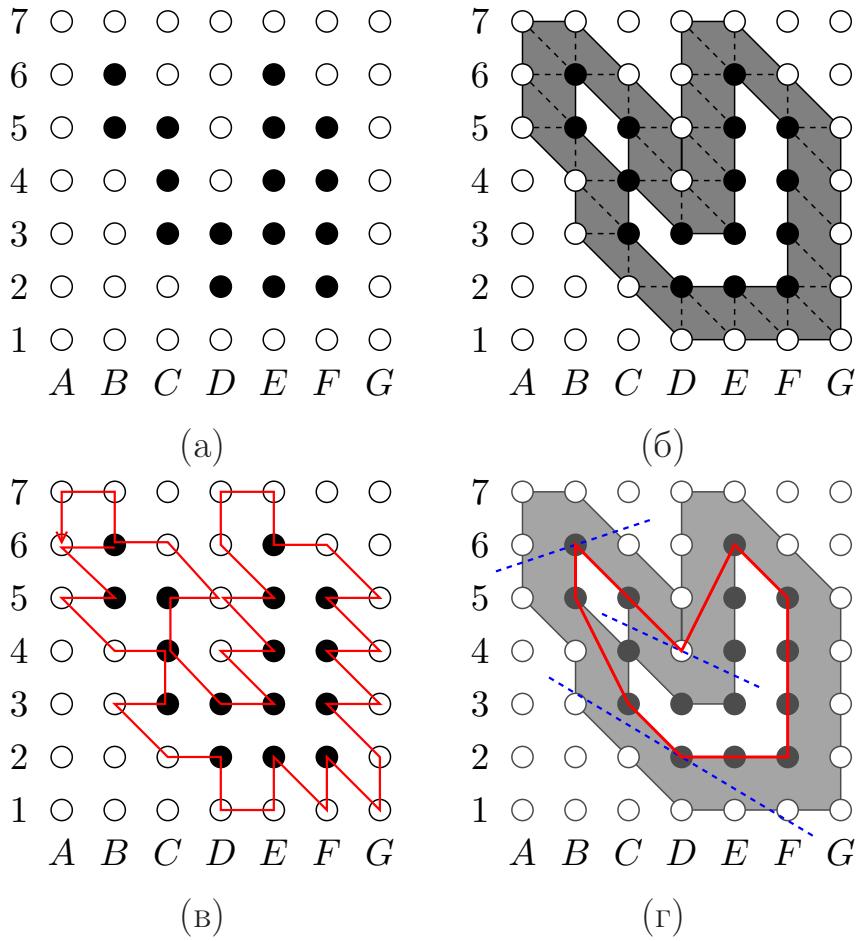


Рисунок 1.12 — Связные компоненты при компонентной смежности

Эта процедура проиллюстрирована рис. 1.12, на котором показаны исходное изображение 1.12а, его единственный граничный коридор 1.12б, след трассировки при прослеживании 1.12в и итоговый минимальный разделяющий прямоугольник 1.12г.

Любой многоугольник, с границей, проложенной внутри граничного коридора, будет разделяющим в том смысле, что все внутренние пиксели бинарного изображения сохранят свой цвет. Поэтому возможны и другие способы многоугольного представления контуров, приводящие к непрерывным сценам, эквивалентным дискретным. Достоинством многоугольников минимального периметра является их относительная простота: они приводят к границам, состоящим из минимально возможного числа элементов, что обеспечивает выигрыш в скорости при обработке непрерывных сцен алгоритмами вычислительной геометрии. Это достигается путём выбора не оптимальной по точности аппроксимирующей многоугольной сцены, однако ширина граничного коридора в пикселях в любом случае не превышает $\sqrt{2}$, а такой точности вполне доста-

точно, если размер объектов на изображении исчисляется не в пикселях, а хотя бы в десятках пикселей.

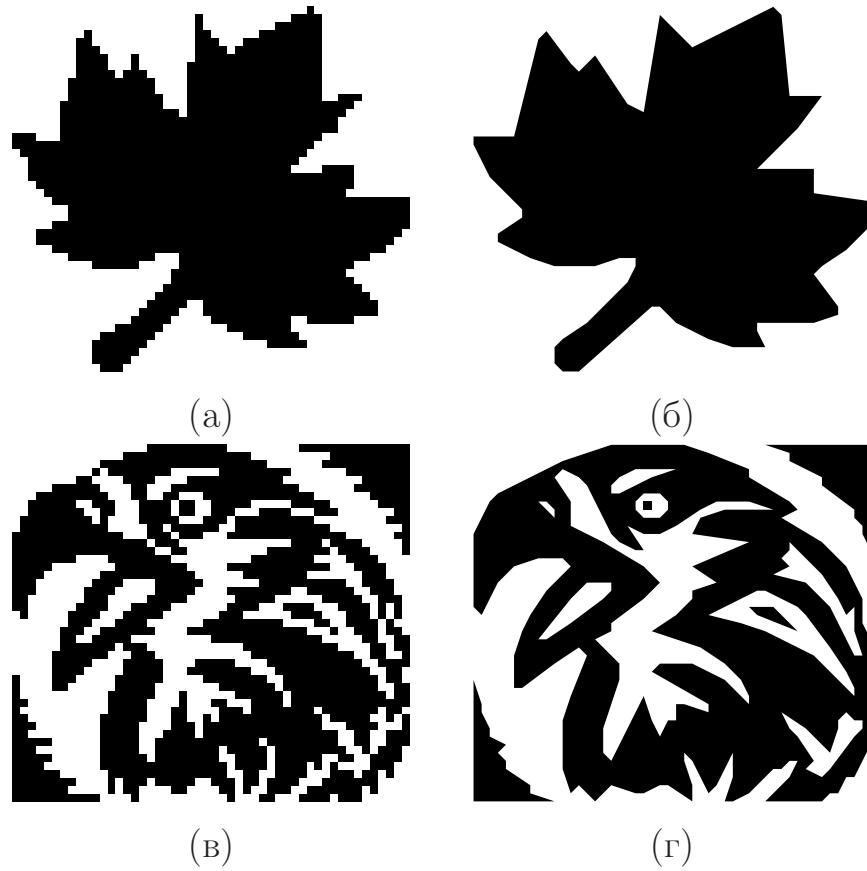


Рисунок 1.13 — Построение бинарных векторных изображений, эквивалентных растровым

Примеры аппроксимации границ для более реалистичных изображений приведены на рис. 1.13. Можно считать, что в результате аппроксимации мы получаем векторные изображения (рис. 1.13б,г), и они оказываются убедительными не только с точки зрения формальной близости к исходным растровым (рис. 1.13а,в), но и визуально.

Таким образом, существует универсальный способ перехода от дискретной сцены к эквивалентной ей непрерывной, подходящий для бинарных изображений любой сложности. Поэтому в качестве входных данных всех последующих алгоритмов можно рассматривать именно непрерывные модели формы — многоугольные фигуры.

1.4 Скелет многоугольной фигуры и бициклы

Определение 1.4. *Фигурой называется замкнутая область на плоскости, ограниченная конечным числом непересекающихся замкнутых жордановых кривых.*

Определение 1.5. Пустым кругом называется круг, целиком лежащий в фигуре.

Множество вписанных кругов тесно связано с медиальным представлением фигуры в виде скелета и радиальной функции.

Определение 1.6. Вписаным кругом фигуры называется пустой круг, который является максимальным, т.е. не содержит ни в каком другом пустом круге.

Определение 1.7. Скелетом фигуры называется множество точек-центров всех вписанных кругов⁴.

Такое определение эквивалентно определению на основе морфологических операций согласно формуле 1.4.

Определение 1.8. Радиальная функция определена в точках скелета и ставит в соответствие точке скелета радиус вписанного круга с центром в этой точке.

Определение 1.9. Многоугольной фигурой называется фигура, у которой граница состоит из замкнутых ломаных линий.

Границу многоугольной фигуры можно представить как объединение конечного числа подмножеств, называемых сайтами: сайтов-точек (вершины фигуры) и сайтов-сегментов (стороны фигуры без концевых точек).

Скелет многоугольной фигуры (рис. 1.14) имеет вид геометрического графа, рёбрами которого являются отрезки прямых линий и квадратичных парабол, а вершинами — концевые точки рёбер. Каждое ребро представляет собой связное множество точек-центров вписанных кругов, касающихся одной и той же пары сайтов, называемых образующими сайтами ребра. Если это пара однотипных сайтов (два сайта-точки или два сайта-сегмента), то ребро имеет форму прямолинейного отрезка. Если же сайты разнотипные (сайт-точка и сайт-сегмент), то ребро представляет собой отрезок квадратичной параболы.

Построение скелета многоугольной фигуры осуществляется с помощью высокоеффективных алгоритмов вычислительной геометрии [43]. Ряд методов вычисления дескрипторов ширины использует особые свойства скелета

⁴Отдельно отметим, что определения пустого и вписанного кругов распространяются и на круги нулевого радиуса. Таким образом, скелет — это замкнутое множество.

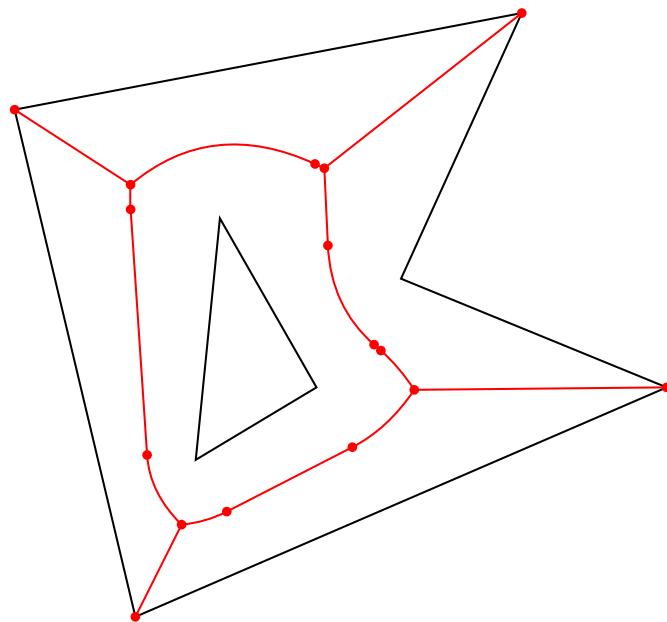


Рисунок 1.14 — Многоугольная фигура и её скелет

многоугольной фигуры [22; 61; 62] и основан на декомпозиции фигуры на составляющие элементы — бицикли.

Определение 1.10. *Бициклом называется множество всех вписанных кругов с центрами на одном ребре скелета. Линия ребра называется осью бицикла.*

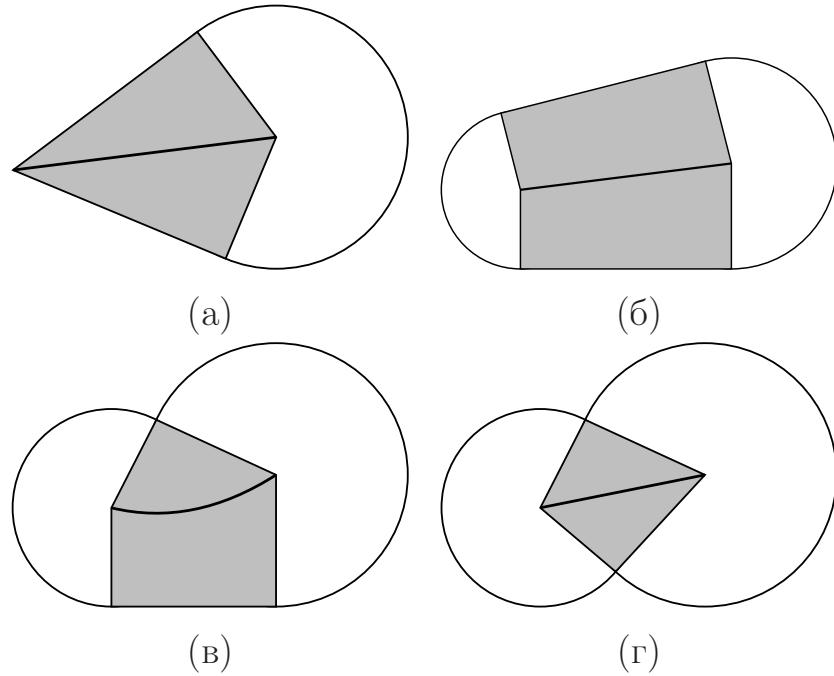


Рисунок 1.15 — Бицикли: оси, собственные области, внешние сектора концевых кругов

Различаются три типа бициклов в зависимости от пары образующих сайтов его ребра: линейный (два сайта-сегмента — рис. 1.15а-б), параболический (сайт-сегмент и сайт-точка — рис. 1.15в) и гиперболический (два сайта-точки — рис. 1.15г). Такая терминология определяется характером зависимости радиальной функции от положения точки на оси бицикла.

Определение 1.11. Круги с центрами в вершинах скелета называются концевыми кругами бицикла.

Граница бицикла представляет собой огибающую семейства входящих в него кругов. Границы линейных и параболических бициклов включают полностью либо частично образующие сайты-сегменты (рис. 1.15а-в). Кроме этого, границы бициклов всех типов состоят из дуг концевых кругов.

Определение 1.12. Внешним сектором бицикла называется сектор концевого круга, опирающийся на дугу в границе бицикла.

Определение 1.13. Спицей называется отрезок прямой, соединяющий точку скелета с ближайшей точкой границы фигуры.

Определение 1.14. Собственной областью бицикла называется объединение всех спиц, инцидентных точкам оси бицикла.

Бицикл представляет собой объединение собственной области и пары внешних секторов. Форма собственной области зависит от типа бицикла (рис. 1.15). Для линейного бицикла она складывается из двух треугольников (рис. 1.15а) либо двух трапеций (рис. 1.15б). В параболическом бицикле — это фигура «домик», которую можно рассматривать как объединение трапеции и треугольника (рис. 1.15в), в гиперболическом — объединение двух треугольников (рис. 1.15г).

Собственные области бициклов образуют покрытие всей многоугольной фигуры, совпадающее с объединением всех спиц, причём пересекаться области могут только по своим граничным спицам.

Определение 1.15. Бицикл называется монотонным, если радиальная функция монотонно убывает или возрастает вдоль его оси.

Очевидно, что линейный бицикл является монотонным, поскольку радиальная функция линейно зависит от положения точки на оси. Линейный бицикл

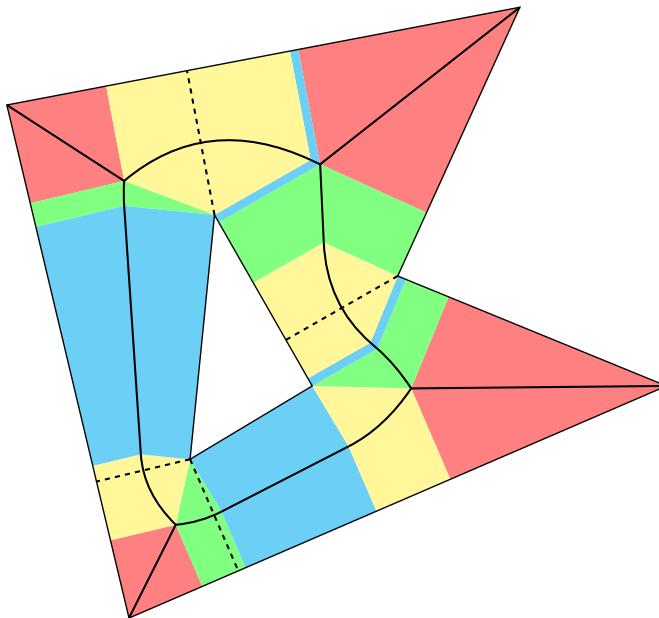


Рисунок 1.16 — Покрытие многоугольной фигуры собственными областями бициклов

постоянной ширины будем считать монотонным по определению. В параболическом бицикле в случае, когда вершина параболы является внутренней точкой оси бицикла, при прохождении через вершину характер изменения радиальной функции меняется с убывающего на возрастающий (рис. 1.15в). Вершина параболы является точкой локального минимума радиальной функции, и бицикл при этом не является монотонным. В остальных случаях, когда вершина параболы лежит вне оси либо совпадает с концевой точкой оси, параболический бицикл является монотонным.

В гиперболическом бицикле монотонность определяется положением центров концевых кругов относительно линии сайтов (прямой, проходящей через сайты-точки). Если центры лежат по разные стороны от этой линии, то точка пересечения её с осью бицикла находится внутри оси и в ней достигается минимум радиальной функции — бицикл не монотонный (рис. 1.16г). В остальных случаях гиперболический бицикл является монотонным.

По опыту работы [62] вычисление ширины для монотонных бициклов представляет собой более простую задачу, чем для немонотонных. Поэтому целесообразно представить фигуру в виде объединения только монотонных бициклов. Для этого каждый немонотонный бицикл может быть заменен парой монотонных бициклов. В немонотонном бицикле всегда существует лишь один минимум радиальной функции, поэтому его ось можно разбить на два участка монотонности. Разбиение немонотонных бициклов на монотонные па-

ры представляет собой коррекцию скелетного графа путём добавления в него вершин в минимальных точках бициклов и разбиения соответствующих рёбер на две части. В примере (рис. 1.16) четыре экстремальных бицикла разбиты на монотонные пары. Пунктиром показаны новые участки границ собственных областей после разбиения.

Ряд способов описания ширины на основе медиального представления формы был обобщён в работе [40]. Эти способы основаны на явном задании в точках фигуры X функции локальной ширины $T(x)$ и построении дескриптора ширины как функции зависимости площади области заданной ширины $\Psi_r(X) = \{x \mid T(x) \geq r\}$ от размерного параметра. Конкретный способ определения ширины фигуры в точке обуславливает вид областей заданной ширины, возможны следующие варианты:

- морфологическая ширина (значение в точке — радиус максимального покрывающего фигуру круга);
- медиальная ширина (значение в точке — длина проходящей через точку спицы);
- дистанционная ширина (значение в точке — расстояние до ближайшей точки границы).

Области заданной ширины при различных способах её определения показаны на рис. 1.17.

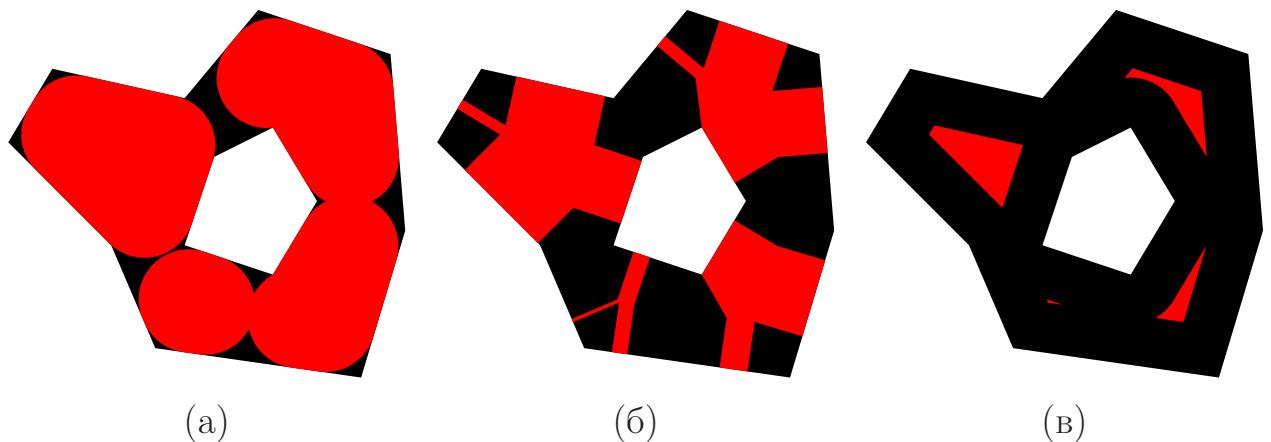


Рисунок 1.17 — Области заданной ширины: (а) морфологической (б) медиальной (в) дистанционной

При работе в рамках непрерывной морфологии области заданной ширины описываются с помощью непрерывных примитивов — бициклов, участков границы, спиц. При этом задачи вычисления медиальной и дистанционной ширины можно считать относительно простыми, так как они сводятся к анализу

спиц: в случае медиальной r -ширины необходимо оставить только спицы длины не меньше r , а в случае дистанционной r -ширины — сделать все спицы короче на r (спицы, которые изначально короче r , пропадают полностью).

Заметим, что область $\psi_r(X)$ в каждом случае является невозрастающей по r , что совпадает с одним из свойств гранулометрии. Проверим, действительно ли семейство операторов $\Psi = \{\psi_r\}_{r \geq 0}$ определяет гранулометрию. Напомним, что гранулометрия задаётся множеством операторов открытия (то есть антиэкстенсивных, возрастающих и идемпотентных), обладающих свойством стабильности (усиленным свойством идемпотентности). Заметим, что все три типа определения ширины обладают свойством антиэкстенсивности, так как область заданной ширины является подмножеством исходной фигуры.

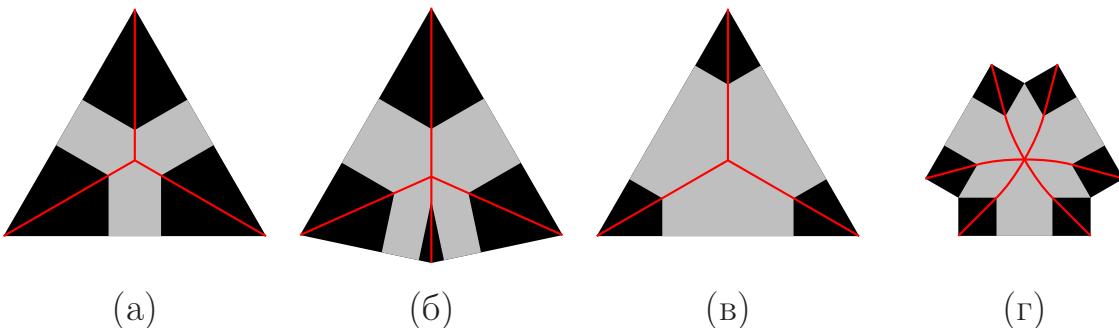


Рисунок 1.18 — Нарушение свойств возрастания (а)-(б) и идемпотентности (в)-(г) в случае медиальной ширины. Чёрным отмечена исходная фигура, красным — её скелет, серым — область медиальной ширины.

Помимо этого,

- морфологическая ширина обладает всеми необходимыми свойствами, так как задаётся операцией открытия с выпуклым структурирующим элементом — диском;
- медиальная ширина не обеспечивает свойства возрастания и идемпотентности. Действительно, при одном и том же значении ширины область меньшей фигуры (рис. 1.18а) может не быть вложена в область большей (рис. 1.18б), а последовательное построение области, вообще говоря, приводит к сужению фигуры (рис. 1.18в,г);
- дистанционная ширина обеспечивает свойство возрастания (очевидно, при расширении фигуры расстояние до ближайшей точки границы может либо увеличиться, либо остаться прежним), но не обеспечивает свойство идемпотентности (последовательное построение области ненулевой ширины снова приводит к сужению фигуры).

Таким образом, несмотря на то что её области имеют более сложную структуру, именно морфологическая ширина сочетает математическую строгость и возможность анализа средствами медиального представления формы, что обосновывает наше желание разработать точные и высокоскоростные алгоритмы её вычисления.

1.5 Выводы к главе 1

Математическая морфология предлагает мощный аппарат для анализа формы плоских объектов. Морфологические операции с выпуклыми структурирующими элементами позволяют анализировать вклад примитива в образование формы объекта при разных масштабах детализации. Становится возможным определить локальную ширину фигуры в точке и описать переменную ширину как функцию распределения значений этой характеристики в точках фигуры. При этом на практике базовые морфологические операции вычисляются для форм в дискретном представлении, что делает эти операции, вообще говоря, математически некорректными, а кроме того, вычислительно затратными.

В большинстве случаев непрерывное представление формы имеет меньшую сложность (число элементов), нежели дискретное, что даёт надежду на создание более эффективных вычислительных процедур. Однако алгоритмы вычисления морфологических операций на базе непрерывных представлений формы недостаточно разработаны и слишком сложны в реализации. Помимо этого, они, как правило, также требуют аппроксимации формы примитива — в данном случае многоугольной фигурой — что приводит к невысокой точности измерения форморазмерных характеристик, связанных с шириной. Тем не менее существуют предпосылки для разработки эффективных процедур вычисления морфологических дескрипторов на базе непрерывной морфологии бинарных изображений.

Подход на базе непрерывной морфологии позволяет представить форму объектов непрерывными фигурами — многоугольниками с многоугольными дырами. При этом сохраняется число, взаимное расположение объектов и цвет точек, соответствующих точкам растровой решётки. Медиальное представление многоугольных фигур в виде скелета и радиальной функции даёт возможность описать фигуру и результаты её морфологических преобразований через геометрические примитивы, связанные с отдельными рёбрами скелета.

Среди предложенных методов описания ширины с помощью непрерывного медиального представления только морфологическая ширина соответствует требованиям классической математической морфологии к форморазмерным дескрипторам. При этом в отличие от других методов вычисление морфологической ширины не сводится к анализу отдельных скелетных примитивов. Разработка точного алгоритма вычисления по-прежнему остаётся открытой задачей, решению которой посвящена следующая глава.

Глава 2. Метод построения дескрипторов формы на основе дискового покрытия

2.1 Дисковое покрытие и морфологическая ширина

Один из возможных подходов к описанию переменной ширины объекта на изображении — задать функцию локальной ширины, определённую в каждой точке фигуры. Локальное описание ширины может быть основано на следующем интуитивном представлении: чем большую ширину имеет объект в какой-либо из своих частей, тем большего размера примитив можно разместить в этой части. Если же вписать в объект примитивы заданного размера, например, *диски определённого диаметра*, то часть объекта, покрытую примитивами, можно рассматривать как *область заданной ширины*. Тогда функцию, описывающую зависимость площади этой области от размера примитива, можно считать интегральным описанием ширины объекта.

В данной главе предлагается подход к построению дескриптора ширины изображения, который основан на площади дискового покрытия объекта (рис. 2.1). Так как дескриптор определяется на основе морфологических преобразований объекта и предназначен для описания ширины, он получил название «морфологической ширины» объекта. Выбор диска в качестве примитива позволяет, во-первых, сделать дескриптор ширины инвариантным к сдвигу, повороту, а с дополнительной нормировкой — и к масштабированию изображения. Во-вторых, он даёт возможность описать области заданной ширины с помощью медиального представления формы — скелета и радиальной функции.

Выбор многоугольной фигуры (многоугольника с многоугольными дырами) в качестве модели формы объекта обусловлен двумя причинами. С одной стороны, с помощью многоугольных фигур можно с высокой точностью аппроксимировать границы сложных объектов, представленных как нелинейными кривыми, так и дискретными растровыми изображениями. С другой стороны, для многоугольной фигуры медиальное представление может быть получено с помощью высокоэффективных алгоритмов вычислительной геометрии.

Дескриптор ширины объекта представляет собой диаграмму зависимости площади покрытия от размера покрывающих дисков (рис. 2.2).

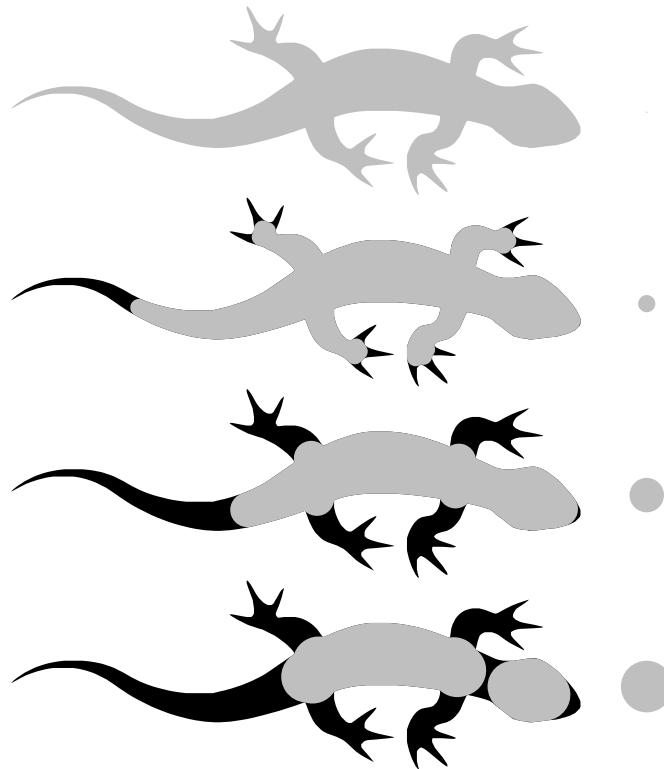


Рисунок 2.1 — Дисковые покрытия фигуры «ящерица» (справа изображены образцы покрывающих дисков)

Предлагаемый в главе метод аналитического вычисления площади дискового покрытия для многоугольных фигур включает следующие этапы:

1. Переход к медиальному представлению многоугольной фигуры в виде скелета и радиальной функции на основе диаграммы Вороного линейных сегментов, составляющих границу фигуры.
2. Представление многоугольной фигуры сложной формы в виде объединения элементарных геометрических фигур, называемых бициклами. Бицикл — это объединение вписанных в фигуру кругов с центрами на ребре скелета.
3. Представление дискового покрытия фигуры в виде объединения подмножества бициклов и вычисление площади дискового покрытия на основе площадей бициклов.
4. Построение функции распределения площади дискового покрытия как функции от размера диска.

Предложенный метод расчёта морфологической ширины многоугольной фигуры на основе площади дискового покрытия имеет высокую точность и вычислительную эффективность.

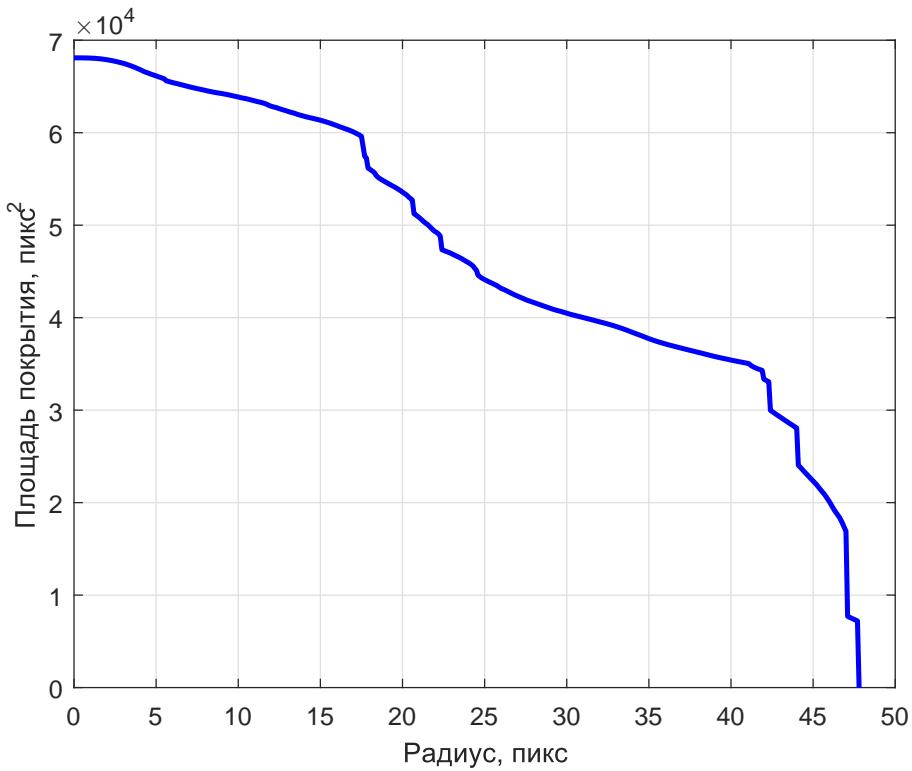


Рисунок 2.2 — Диаграмма зависимости площади дискового покрытия фигуры «ящерица» от размера (радиуса) покрывающих дисков

Определение 2.1. *Дисковое r -покрытие фигуры X (обозначается $X(r)$) — объединение всех пустых (целиком лежащих в фигуре) кругов радиуса r . Примеры дискового r -покрытия для разных значений r представлены на рис. 2.1.*

Определение 2.2. *r -площадь фигуры — площадь её дискового r -покрытия.*

Согласно этому определению, площадь всей фигуры — это её 0-площадь.

Определение 2.3. *Морфологическая ширина фигуры $\mathfrak{F}(r)$ — это r -площадь фигуры как функция от r .*

Очевидно, что каждый пустой круг радиуса больше r может быть представлен как объединение пустых кругов радиуса r . Таким образом, дисковое r -покрытие не расширяется с ростом r , то есть $X(r') \subseteq X(r)$ при $r' \geq r$, и морфологическая ширина является невозрастающей функцией параметра r .

Нетрудно также заметить, что множество вписанных в фигуру примитивов B является результатом операции открытия фигуры X с B в качестве структурирующего элемента:

$$\begin{aligned}
X \circ B = (X \ominus B) \oplus B &= \{z + b : z \in (X \ominus B), b \in B\} = \\
&= \{z + b : z \in \{c : B_c \subset X\}, b \in B\} = \bigcup_{\{z : B_z \subset X\}} B_z. \quad (2.1)
\end{aligned}$$

Таким образом, дисковое r -покрытие фигуры X является её открытием с дисковым структурирующим элементом $B(r) = \{z : \|z\| \leq r\}$, а производная его площади — морфологической ширины — с точностью до знака даст нам морфологический спектр с этим элементом¹

Из рассуждений выше следует, что любой вписанный круг с радиусом r или больше r содержитя в дисковом r -покрытии. Следовательно, дисковое r -покрытие фигуры совпадает с объединением всех вписанных кругов радиуса не меньше r . Центры таких вписанных кругов составляют подмножество точек скелета. Таким образом, для вычисления морфологической ширины фигуры достаточно рассматривать только круги, центры которых лежат на скелете. Задача состоит в том, чтобы для заданных значений аргумента r получить соответствующие значения r -площади фигуры. Для этого нужно на основе медиального представления фигуры выделить множество всех вписанных кругов радиуса r и больше и найти площадь объединения этих кругов. Решение этой задачи для многоугольных фигур будет получено в явном виде. Итак, сформулируем постановку основной задачи главы.

Задача 2.1. *Дана многоугольная фигура X и набор значений радиуса $\{r_i\}_{i=1}^n$, $r_i \geq 0$ и $r_i < r_j$ при $i < j$. Вычислить $\mathfrak{F}(r_i)$, $i = 1, \dots, n$.*

2.2 Собственные области и внешние сектора

Пусть X — многоугольная фигура, $X'(r)$ — подмножество фигуры, образованное объединением всех спиц длины r и более. Очевидно, что множество $X'(r)$ целиком лежит в r -покрытии, иначе говоря, область медиальной ширины r вложена в область морфологической ширины r . Поэтому анализ r -покрытия мы начнём с рассмотрения множества спиц длины r и более внутри каждого из бициклов.

¹Так как морфологическая ширина — монотонная функция, она дифференцируема почти всюду, везде имеет левую и правую производные и может иметь точки разрыва только первого рода. При отсутствии скачка в точке r можно принять $PS(r) = -\frac{\mathfrak{F}'(r-0)+\mathfrak{F}'(r+0)}{2}$, а при его наличии использовать дельта-функцию Дирака: $PS(r) = (\lim_{x \rightarrow r+0} \mathfrak{F}(x) - \lim_{x \rightarrow r-0} \mathfrak{F}(x))\delta(r)$, как это сделано в [21].

На рис. 2.3 представлены монотонные бициклы всех трёх типов. Здесь r и R — радиусы меньшего и большего концевых кругов, l — расстояние между их центрами. Если бицикл является линейным или параболическим, то он имеет образующий сайт-сегмент, и тогда t — длина проекции оси бицикла на этот сайт:

$$t = \sqrt{l^2 - (R - r)^2}.$$

В параболическом бицикле p — это фокальный параметр параболы, то есть расстояние между сайтом-точкой и прямой линией, содержащей сайт-сегмент. В гиперболическом бицикле q — расстояние между сайтами-точками. Для линейного бицикла (рис. 2.3а) площадь собственной области определяется как сумма площадей двух трапеций, имеющих основания r и R и высоту t :

$$S_{lin} = 2 \cdot ((R + r)/2) \cdot t = (R + r) \cdot t. \quad (2.2)$$

Угловой размер внешнего сектора меньшего концевого круга:

$$\varphi_{lin} = 2\alpha = 2 \cdot \arcsin(t/l). \quad (2.3)$$

Для параболического бицикла (рис. 2.3б) площадь собственной области складывается из площади такой же трапеции и площади треугольника с вершинами в центрах концевых кругов и в сайте-точке. Площадь треугольника вычисляется по формуле Герона:

$$S_{par} = ((R + r)/2) \cdot t + \sqrt{P(P - R)(P - r)(P - l)}, \quad (2.4)$$

где $P = (R + r + l)/2$.

Угловой размер внешнего сектора меньшего концевого круга параболического бицикла есть

$$\varphi_{par} = (\pi/2) + \alpha = (\pi/2) + \arcsin((p - r)/r). \quad (2.5)$$

Площадь собственной области гиперболического бицикла (рис. 2.3в) складывается из площадей двух треугольников, вычисляемых по формуле Герона:

$$S_{hyp} = 2\sqrt{P(P - R)(P - r)(P - l)}. \quad (2.6)$$

Угловой размер внешнего сектора меньшего концевого круга

$$\varphi_{hyp} = 2 \arcsin(q/2r). \quad (2.7)$$

В итоге получены формулы, по которым вычисляются площади собственных областей и внешних секторов концевых кругов для всех трёх типов бициклов.

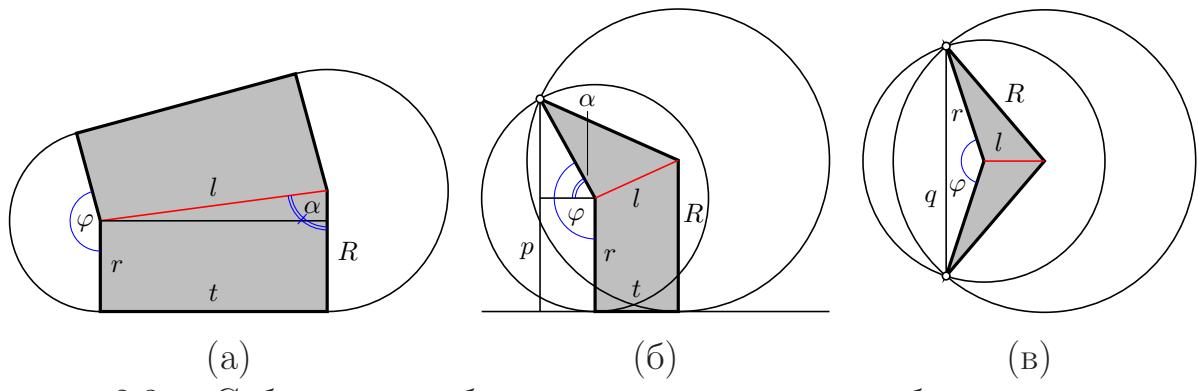


Рисунок 2.3 — Собственные области и внешние сектора бициклов: линейного (а), параболического (б), гиперболического (в)

2.3 Усечённые бициклы

Дисковое r -покрытие многоугольной фигуры при $r = 0$ совпадает с самой многоугольной фигурой. По мере роста r покрытие сжимается: часть фигуры, покрытая дисками, уменьшается (рис. 2.4). Само покрытие при этом представляет собой фигуру, граница которой состоит из линейных сегментов и дуг окружностей.

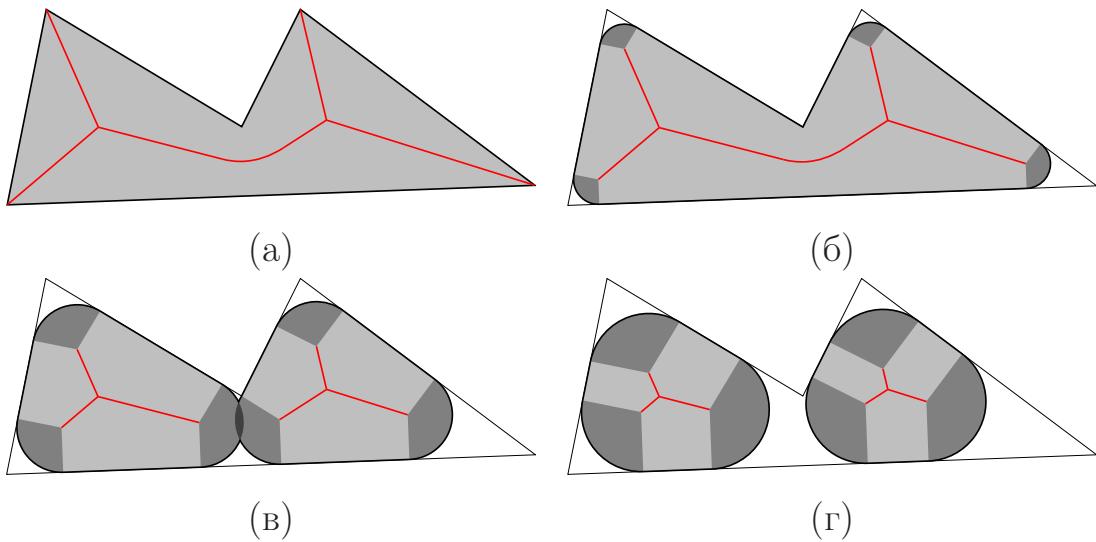


Рисунок 2.4 — Изменение дискового r -покрытия с ростом радиуса дисков

Дисковое r -покрытие есть объединение вписанных в многоугольную фигуру кругов с радиусом больше или равным r . Будем называть множество центров этих кругов *осью* r -покрытия. Очевидно, ось r -покрытия представляет собой подмножество скелета многоугольной фигуры. Это подмножество при малых значениях r является связным, но с ростом r оно может разбиться на несколько связных компонент (рис. 2.4).

Таким образом, скелет многоугольной фигуры делится на две части: ось r -покрытия — это подмножество с радиальной функцией r или больше, и осталь-

ная часть — подмножество, в котором радиальная функция меньше r . Оба эти подмножества можно рассматривать как геометрические графы.

Для каждого $r > 0$ все бициклы многоугольной фигуры разбиваются на три группы: широкие (входят в r -покрытие), узкие (не входят в r -покрытие) и усечённые (входят в r -покрытие частично).

Пусть в монотонном бицикле R_1 и R_2 — радиусы минимального и максимального кругов. В широком бицикле $R_1 \geq r$, в узком $R_2 < r$. Если в бицикле $R_1 < r \leq R_2$, то в r -покрытие входят лишь те его круги, у которых радиус не меньше r . Определим операцию сечения (коррекции) такого бицикла, которая состоит в удалении кругов с радиусом меньше r . Полученный новый бицикл будем называть усечённым. Усечённый бицикл имеет те же образующие сайты и тот же тип (линейный, параболический или гиперболический), что и исходный бицикл. Минимальным кругом усечённого бицикла становится круг радиуса r , а максимальным остаётся круг с радиусом R_2 .

Обозначим C_1, C_2 — центры меньшего и большего концевых кругов. Для полного описания усечённого бицикла и расчёта его площади достаточно вычислить новое положение меньшего концевого круга. Пусть точка C — это искомый центр круга с радиусом r (рис. 2.5).

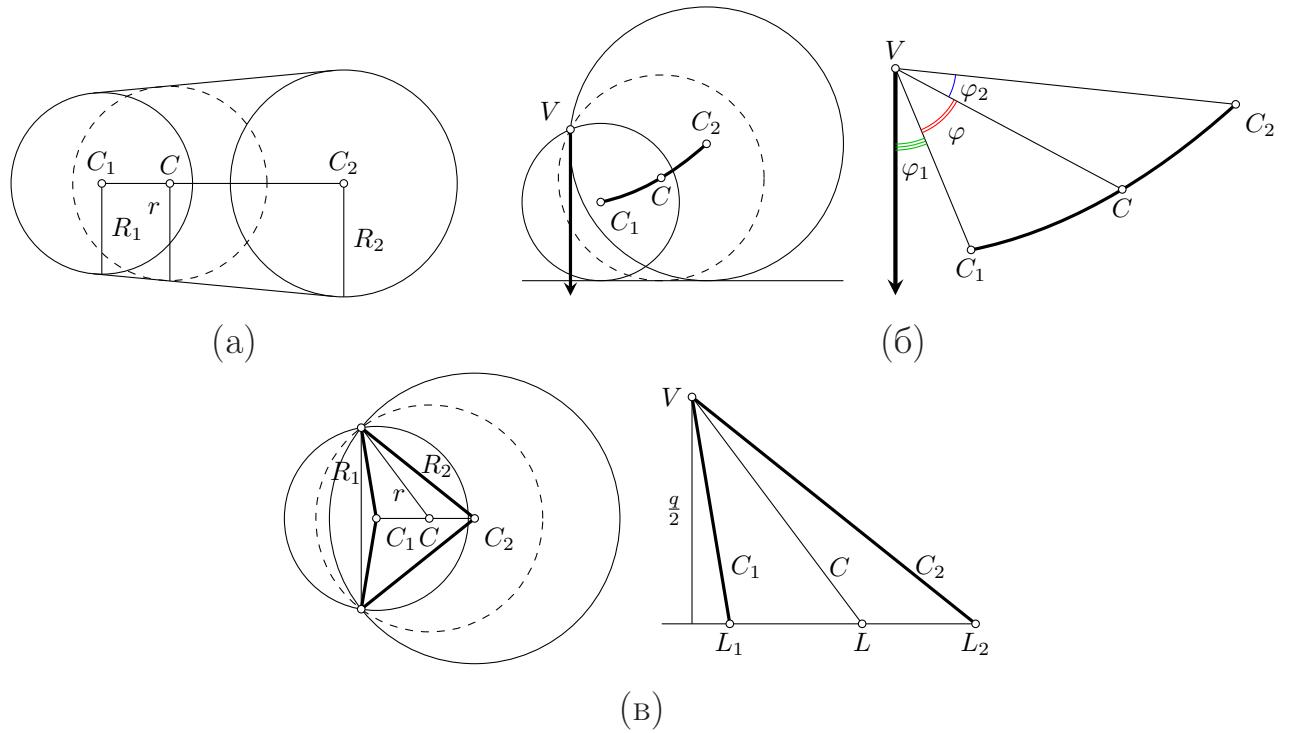


Рисунок 2.5 — Коррекция усечённых бициклов

Для линейного бицикла (рис. 2.5а) имеем $C = C_1 + \lambda \cdot (C_2 - C_1)$, где $\lambda = (r - R_1)/(R_2 - R_1)$. В частном случае, когда $R_2 = R_1$, полагаем $\lambda = 0$.

Для параболического бицикла (рис. 2.5б) выберем полярную систему координат (r, φ) с началом в сайте-точке V бицикла и осью, ортогональной сайту-сегменту. Уравнение параболы в этих координатах имеет вид $\rho = p/(1 + \cos(\varphi))$, где p — фокальный параметр параболы. Центры концевых дисков имеют координаты $C_1(R_1, \varphi_1)$ и $C_2(R_2, \varphi_2)$, где $\varphi_1 = \arccos(p/R_1 - 1)$, $\varphi_2 = \arccos(p/R_2 - 1)$. Искомая точка $C(R, \varphi)$, $\varphi = \arccos(p/R - 1)$. Не нарушая общности, будем считать $\varphi_1 < \varphi_2$. Вектор \overline{VC} получается из $\overline{VC_1}$ поворотом на угол $\theta = \varphi - \varphi_1$ и умножением на коэффициент r/R_1 . Тогда искомый центр круга есть $C = V + (r/R_1) \cdot \mathbf{G} \cdot \overline{VC}$, где \mathbf{G} — матрица поворота на угол θ :

$$\mathbf{G} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

В гиперболическом бицикле (рис. 2.5в) точка C лежит между C_1 и C_2 . Пусть q — расстояние между сайтами-точками. Если V — сайт-точка, то проекции векторов \overline{VC} , $\overline{VC_1}$, $\overline{VC_2}$ на ось бицикла имеют длину

$$\begin{aligned} L &= |\overline{VC}| = \sqrt{r^2 - \left(\frac{q}{2}\right)^2}, \\ L_1 &= |\overline{VC_1}| = \sqrt{R_1^2 - \left(\frac{q}{2}\right)^2}, \\ L_2 &= |\overline{VC_2}| = \sqrt{R_2^2 - \left(\frac{q}{2}\right)^2}. \end{aligned}$$

Тогда $C = C_1 + \lambda \cdot (C_2 - C_1)$, где $\lambda = (L - L_1)/(L_2 - L_1)$.

Полученные формулы позволяют найти новое положение меньшего концевого круга, после чего вычисление площади собственной области бицикла и угловых размеров внешних секторов осуществляется по тем же формулам (2.2)–(2.7), что и для широких бициклов.

Таким образом, дисковое r -покрытие представляет собой объединение двух множеств бициклов: полных бициклов, у которых $R_1 \geq r$, и усечённых бициклов, у которых $R_1 < r \leq R_2$.

Напомним, что каждый бицикл состоит из собственной области и пары внешних секторов концевых кругов. При этом собственные области бициклов пересекаются только по общим границам, имеющим нулевую площадь, поэтому общая площадь объединения собственных областей равна сумме площадей собственных областей бициклов. Легко показать, что для анализа не покрытой

собственными областями части дискового покрытия достаточно из всех внешних секторов рассмотреть только внешние сектора меньших кругов усечённых бициклов.

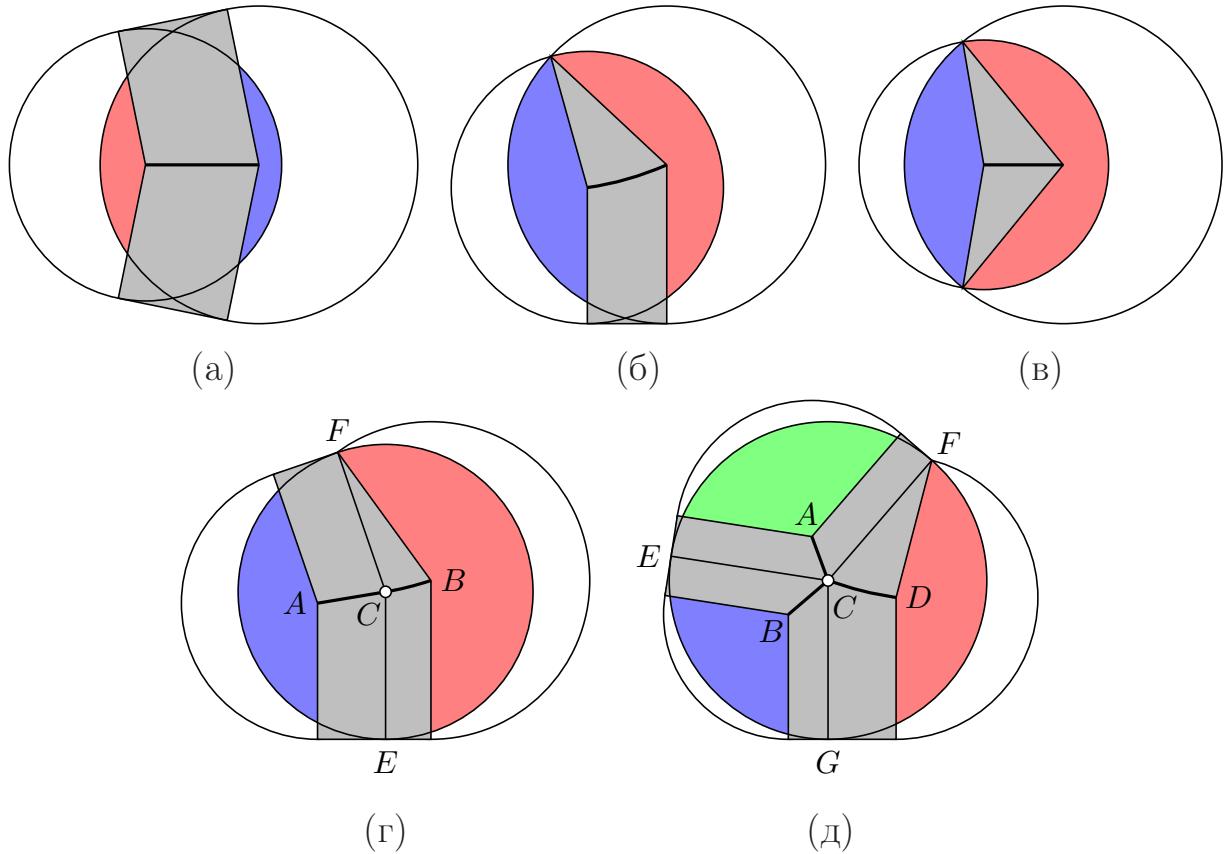


Рисунок 2.6 — К доказательству леммы 2.1

Лемма 2.1. *Часть дискового покрытия, не входящая в объединение собственных областей полных и усечённых бициклов, покрывается внешними секторами меньших концевых кругов усечённых бициклов.*

Доказательство. Заметим, что при любом типе бицикла часть внутреннего (дополнительного к внешнему) сектора его концевого круга, не покрытая собственной областью, оказывается вложенной во внешний сектор другого концевого круга. На рис. 2.6а–в такие сектора выделены светло-синим и светло-красным светом. При соединении соседних бициклов оказывается, что для каждого бицикла внешний сектор концевого круга с центром в общей вершине C является либо внутренним сектором второго бицикла, если вершина имеет степень 2 (рис. 2.6г), либо объединением внутренних секторов остальных бициклов, если вершина имеет степень 3 и более (рис. 2.6д). К примеру, на рис. 2.6д внешний сектор бицикла CFE с осью AC складывается из внутренних секторов CFG и CGE бициклов с осями CD и CB соответственно. Но для

каждого из внутренних секторов его часть, не покрытая собственной областью, вложена во внешний сектор другого концевого круга бицикла, то есть сектор круга с центром в соседней вершине скелета. Повторив подобное рассуждение несколько раз, для каждого внешнего сектора или его части можно построить цепочку вложений, которая заканчивается внешним сектором одного из кругов с центром в терминальной вершине оси покрытия.

Таким образом, объединение группы бициклов можно представить как объединение их собственных областей и внешних секторов в терминальных вершинах подграфа скелета, образованного осями бициклов. Необходимо лишь, чтобы сумма внутренних секторов в нетерминальных вершинах была равна 2π , т.е. чтобы нетерминальные вершины либо сохраняли все свои рёбра, либо исключались из подграфа. Множество полных и усечённых бициклов удовлетворяет этому требованию, так как вершина с кругом радиуса ρ при $r \leq \rho$ сохранит в r -покрытии все свои рёбра, некоторые, возможно, в усечённом виде, а при $r > \rho$ будет исключена из подграфа. Терминальными вершинами подграфа, образованного осями бициклов, будут центры меньших концевых кругов усечённых бициклов.

В результате приходим к выводу: покрытие формируется из собственных областей всех этих бициклов, а также из внешних секторов меньших кругов усечённых бициклов. ■

Такая структура дискового покрытия позволяет свести задачу вычисления r -площади многоугольной фигуры к суммированию площадей собственных областей и объединения внешних секторов бициклов. На рис. 2.4 собственные области выделены светлым, а внешние сектора — тёмным серым цветом.

Концевые круги усечённых бициклов в r -покрытии имеют радиус r . Площадь внешнего сектора с углом φ есть $S_{sec} = (\varphi/2) \cdot r^2$. Но сектора могут иметь нетривиальные пересечения. Пересечения возникают, когда происходят «разрывы» оси r -покрытия (рис. 2.4в). Кроме того, если сектора «задевают» собственные области, часть площади секторов изначально включается в совокупную площадь собственных областей. В таких случаях вычисление общей площади внешних секторов представляет собой более сложную задачу.

2.4 Пересечения бициклов

В предыдущих разделах мы рассмотрели общую структуру дискового покрытия и пришли к выводу, что покрытие можно представить в виде объ-

единения двух множеств A и B , где первое объединяет собственные области полных и усечённых бициклов, а второе — внешние сектора меньших концевых кругов усечённых бициклов. Так как множества A и B , вообще говоря, пересекаются, для площади r -покрытия справедлива формула:

$$\mathfrak{F}(r) = S(A) + S(B \setminus A). \quad (2.8)$$

Множество A по сути является объединением спиц длины r и более, то есть областью медиальной ширины $X'(r)$, площадь которой можно найти как сумму площадей собственных областей, так как они пересекаются только по своим границам. Эта задача решена в разделах 2.2–2.3, поэтому вычисление площади покрытия сводится к вычислению второго слагаемого в формуле 2.8.

Задача 2.2. *Дана многоугольная фигура X и радиус покрывающего круга r , для которого определена область медиальной ширины $X'(r)$ и множество внешних секторов меньших концевых кругов усечённых бициклов $\{S_i\}_{i=1}^k$. Вычислить площадь множества $\left(\bigcup_{i=1}^k S_i\right) \setminus X'(r)$.*

2.4.1 Общая структура пересечений

Решение задачи 2.2 осложняется возможным наличием пересечений между секторами усечённых бициклов. Для того чтобы определить бициклы, пересечения которых необходимо учесть при расчёте площади r -покрытия, исследуем общую структуру пересечения вписанных кругов фигуры. Пусть вписанные круги, не обязательно одинакового радиуса, с центрами в точках A и B пересекаются. Их пересечение представляет собой «линзу», фигуру, граница которой состоит из двух дуг окружностей. Имеет место следующее свойство.

Лемма 2.2. *Между центрами двух вписанных кругов, имеющих непустое пересечение, существует единственный простой (не проходящий дважды через одну и ту же точку) путь в скелете, такой, что все вписанные круги с центрами в точках пути содержат это пересечение.*

Доказательство. Для начала докажем, что таких путей не может быть больше одного. Альтернатива при выборе путей возможна, если скелет содержит циклы, и различные пути проходят по разным их участкам. Рассмотрим два пути-кандидата. Так как начала и концы путей фиксированы, найдётся пара точек, для которых участки, заключённые между ними в выбранных путях,

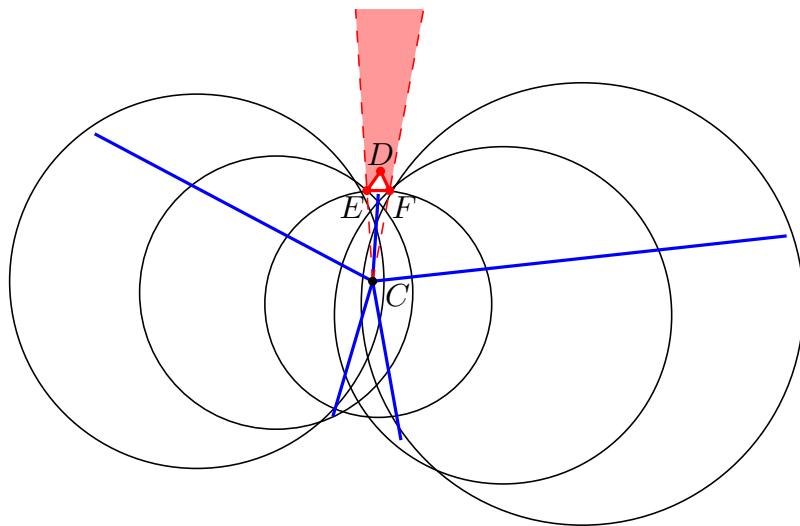


Рисунок 2.7 – К доказательству леммы 2.2. Все точки кругов должны быть видны из общей точки C (синие отрезки лежат внутри кругов). При наличии дыры DEF имеется зона, в которой не могут располагаться круги (закрашена красным). В результате объединение кругов не может охватывать дыру, что ведёт к противоречию.

не имеют других общих точек. Объединение этих участков образует цикл. Если все круги так на первом, так и на втором участке охватывают линзу, значит, все круги в цикле имеют общую точку C . Но тогда, в силу выпуклости кругов, все точки объединения должны быть видны из точки C , т.е. отрезки, проведённые из точки C , должны полностью лежать внутри объединения. Это невозможно, так как объединение кругов в цикле представляет собой фигуру с дырой, ограничивающей видимость. Ситуацию иллюстрирует рис. 2.7.

Для доказательства существования искомого пути определим его конструктивно. Рассмотрим две пересекающихся окружности с центрами в точках A и B . Каждая из них касается границы всей фигуры как минимум в двух точках, которые расположены вне дуги, ограничивающей линзу. Возьмём две последовательных точки касания, между которыми заключена такая дуга, обозначим их A_1 и A_2 для первого круга, B_1 и B_2 для второго (рис. 2.8а). Тогда объединение кругов имеет условную «верхнюю» и «нижнюю» границы, лежащие по разные стороны от линзы, на рис. 2.8а они обозначены синим и красным цветом. Построим такой путь между точками A и B , чтобы верхняя огибающая его кругов была не ниже верхней границы, а нижняя — не выше нижней границы. Точка A лежит на ребре, образующие сайты которого содержат точки A_1 и A_2 . Будем называть их верхним и нижним сайтом, а процедуру построения

пути представим как процесс прослеживания сайтов. Начнём движение по этому ребру внутрь сектора, граница которого содержит граничную дугу линзы. Возможны следующие случаи:

1. Мы достигли точки B .
2. Мы достигли вершины скелета степени 2. Тогда движение продолжается по соседнему ребру. При этом один из образующих сайтов сохраняется, а другой меняется на соседний (сайт-точка на сайт-сегмент или наоборот).
3. Мы достигли вершины скелета степени 3². С этой вершиной связаны три образующих сайта, два из которых — те же, что и у ребра, по которому мы пришли в вершину. Если третий сайт расположен сверху, значит, нужно обновить верхний сайт, если снизу — нижний сайт, и далее продолжить движение по ребру-бисектору верхнего и нижнего сайтов.

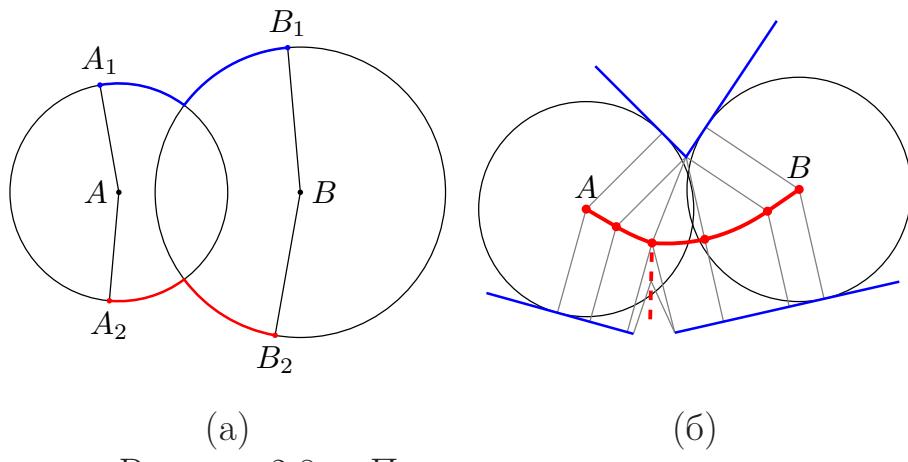


Рисунок 2.8 — Построение опорного пути

В итоге мы достигнем точки B за конечное число шагов, так как число рёбер в скелете конечно, а зацикливание невозможно, ведь тогда одна из огибающих будет лежать внутри другой.

По построению каждый круг на пути имеет спицы, пересекающие как нижнюю, так и верхнюю границы объединения пары исходных кругов. Пример

²Несмотря на то что число инцидентных вершин скелета рёбер формально может быть больше трёх, не ограничивая общности, можно считать, что все «развилки» имеют степень 3, так как в противном случае группу смежных рёбер можно разделить на тройки, добавив фиктивные вершины с теми же координатами и радиусом вписанного круга и фиктивные рёбра нулевой длины. Более того, именно такая структура с ограничением на число инцидентных рёбер используется в программном комплексе для построения гранично-скелетного представления бинарного изображения.

построения искомого пути показан на рис. 2.8б: путь не идёт по ребру, отмеченному пунктиром, так как оба сайта ребра — нижние. Примечательно также, что найденный путь не лежит полностью внутри объединения кругов. Однако можно установить, что все промежуточные круги содержат интересующую нас линзу. Форма любой линзы, не обязательно симметричной, однозначно определяется своими радиусами кривизны и величиной граничной дуги с меньшим радиусом. Отметим свойства монотонности линзы как множества: она сужается с ростом большего радиуса при фиксированной дуге с меньшим радиусом (рис. 2.9а) и расширяется с ростом величины этой дуги, если зафиксировать радиусы и точку пересечения дуг (рис. 2.9б).

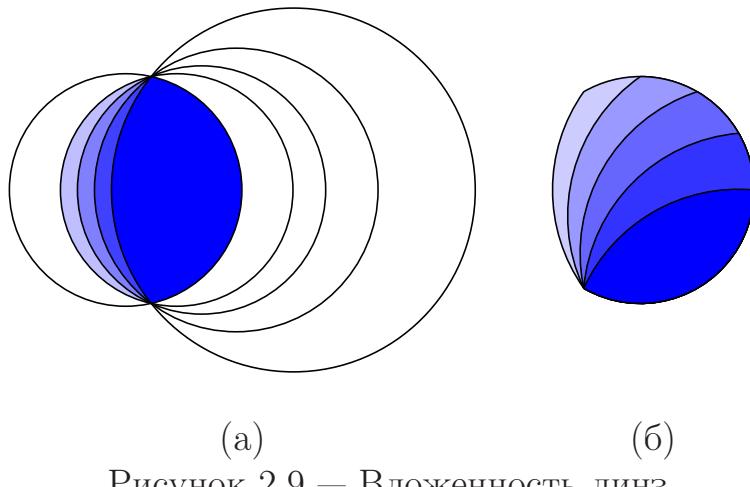


Рисунок 2.9 — Вложенность линз

Не ограничивая общности, будем считать, что круг с центром в точке A имеет радиус r , круг с центром в точке B — радиус R , и $r \leq R$, а линзу в пересечении кругов будем называть EF по её концам. Рассмотрим множество возможных положений центра промежуточного круга. Оно зависит от радиуса этого круга z . Пусть $z \leq r$. Тогда центр круга должен лежать на расстоянии не больше z как от верхней, так и от нижней границы (рис. 2.10а). В результате он удалён не более чем на z от концов линзы EF , а значит, расположен в другой линзе, с концами C и D , радиусы кривизны которой равны z . Нужно убедиться, что пересечение всех кругов радиуса z с центрами внутри второй линзы покрывает исходную. В силу выпуклости кругов-примитивов в пересечении достаточно проверить только точки её границы. Причём, так как все круги с центрами на каждой из граничных дуг опираются на общую точку (верхняя дуга — на F , а нижняя — на E), мы имеем ситуацию, изображённую на рис. 2.9б. Поэтому чем сильнее удалены точки на дуге, тем меньше линза в их пересечении. Таким образом, достаточно проверить концы дуги C и D , но они, в

свою очередь, образуют линзу с той же высотой, что и первая, но меньшими радиусами кривизны, и потому охватывают её (тип вложенности линз, показанный на рис. 2.9а). Следует отметить, что ситуация $z \leq r$ возможна только, если четырёхугольник $AEBF$ — выпуклый, поэтому угловые величины дуг линзы EF меньше π . В результате при любом положении промежуточного круга радиуса z он содержит исходное пересечение.

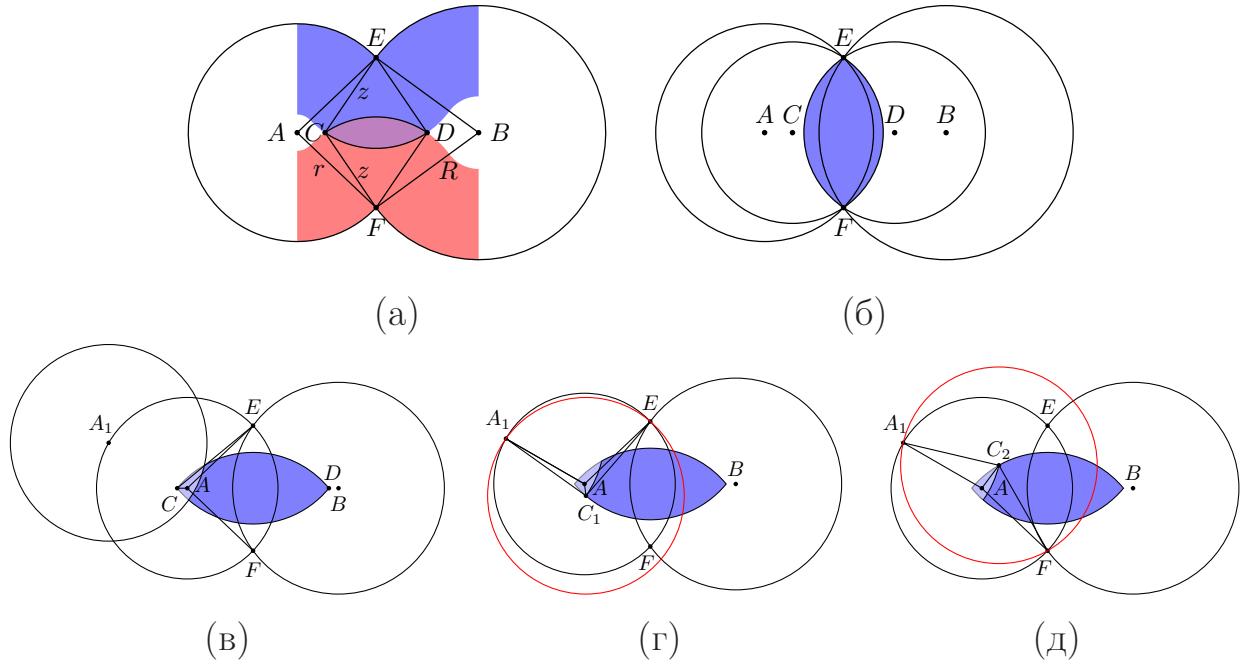


Рисунок 2.10 — Возможные положения центра промежуточного круга

Теперь рассмотрим случай $r < z \leq R$. Обратим внимание на то, что на внешней дуге круга с центром в точке A есть точка границы фигуры A_1 , которая не может быть покрыта промежуточным кругом (иначе круг не будет вписанным), так что его центр должен быть удалён от A_1 на расстояние не менее z . Все точки внешнего сектора расположены от точки C на расстоянии не более z , поэтому возможные положения центра ограничены не линзой, а криволинейным треугольником — линзой с вырезанной областью, состоящей из точек, расположенных слишком близко к точке A_1 (рис. 2.10в). По доказанному ранее, достаточно рассмотреть вершины этого треугольника. Пусть точка пересечения круга радиуса z с центром в точке A_1 пересекается с нижней дугой линзы CB в точке C_1 , а с верхней дугой — в точке C_2 . Рассмотрим линзу EA_1 в пересечении кругов с радиусами AA_1 и C_1A_1 (рис. 2.10г). Одна из её граничных дуг (EFA_1 , перечисляя по часовой стрелке) содержит правую дугу линзы EF , а радиус кривизны второй дуги не превышает соответствующий радиус исходной линзы (z против R). Таким образом, линзу EA_1 можно получить из линзы

EF , сначала увеличив величину дуги с меньшим радиусом, как на рис. 2.9б, а затем понизив радиус большего круга, как на рис. 2.9а. Поэтому круг радиуса z с центром в точке C_1 содержит линзу EF . Аналогичные рассуждения можно привести для точки C_2 (рис. 2.10д). Наконец, круг радиуса z с центром в точке D при пересечении с кругом, соответствующим точке A , образует линзу, в которой дуга с меньшим радиусом кривизны совпадает с аналогичной дугой в линзе EF , а больший радиус не превышает аналогичный радиус в EF (z против R ; тип вложенности линз с рис. 2.9а). В результате линза EF покрывается при любом возможном положении промежуточного круга.

Наконец, случай $z > R$ получается модификацией второго: достаточно рассмотреть точки A_1 и B_1 на внешней дуге каждого из исходных кругов и удалить из линзы, ограничивающей возможные положения промежуточного круга, не одну область, а две. Лемма доказана. ■

Определение 2.4. Путь между центрами пересекающихся кругов, в котором каждый промежуточный круг содержит пересечение конечных, называется опорным.

Понятие опорного пути понадобится нам для теоретического обоснования метода учёта пересечений между усечёнными бициклами при расчёте площади дискового покрытия, а также непосредственной организации процедуры поиска пересечений. Очевидно, что проверять надо не все возможные пары бициклов, так как не во всех парах сектора образуют непустое пересечение. Однако рассмотрение всех пар пересекающихся секторов также является избыточным. Мы будем искать пары секторов, пересечения которых являются максимальными по включению.

Определение 2.5. Два пересекающихся усечённых бицикла в r -покрытии называются смежными, если радиальная функция во всех точках опорного пути между центрами его меньших концевых кругов меньше r .

Пересекаться могут и внешние сектора бициклов, не являющихся смежными. На рис. 2.11а опорный путь между точками A и B проходит через точки C и D , радиальная функция в которых равна r , поэтому бициклы смежными не являются. Тем не менее, бицикл, внешний сектор концевого круга которого пересекается с другим сектором, имеет смежный: достаточно построить опорный путь между центрами кругов, взять первую встретившуюся на пути точку с

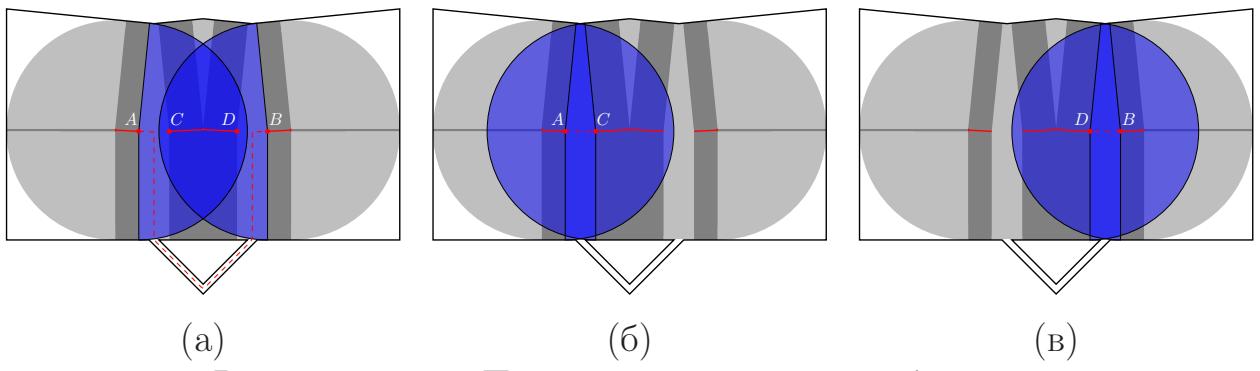


Рисунок 2.11 — Пересечение несмежных бициклов

радиусом r и определить соответствующий ей бицикл. К примеру, пары бициклов с центрами кругов в точках A и B , а также C и D , являются смежными (рис. 2.11б,в). При этом требование достаточно низких значений радиальной функции именно на опорном пути является существенным: из A в B существует путь, обозначенный на рис. 2.11а пунктиром, радиальная функция во всех точках которого меньше r , но он не является опорным. Отметим также, что по определению опорного пути пересечение концевых кругов несмежных бициклов оказывается вложенным в пересечение концевых кругов смежных бициклов, лежащих на их опорном пути. Таким образом, несмежные бициклы не несут необходимой информации об устройстве покрытия, что приводит нас к идею исключить их из рассмотрения.

Рассмотрим пару пересекающихся усечённых бициклов. Нас будут интересовать только внешние сектора меньших концевых кругов бициклов. В монотонном бицикле угловые размеры внешнего сектора меньшего концевого круга $\varphi \leq \pi$.

Внешний сектор усечённого бицикла может иметь пересечение не только с внешним сектором другого бицикла, но и с его собственной областью. При вычислении вклада внешних секторов двух пересекающихся усечённых бициклов в r -площадь фигуры необходимо найти площадь только той части каждого сектора, которая не входит в собственные области. На рис. 2.12 представлены примеры возможных взаимных расположений внешних секторов двух усечённых бициклов. В первом случае (рис. 2.12а) в пересечении двух секторов образуется фигура «линза», граница которой состоит из двух одинаковых дуг окружностей. Во втором случае (рис. 2.12б) пересечение секторов представляет собой более сложную фигуру, граница которой включает прямолинейные отрезки спиц и дуги окружностей. Серым цветом на рис. 2.12 выделены области, образованные объединением внешних секторов за вычетом пересечения

с внутренними секторами. Такие области будем называть *внешними зонами* пары бициклов.

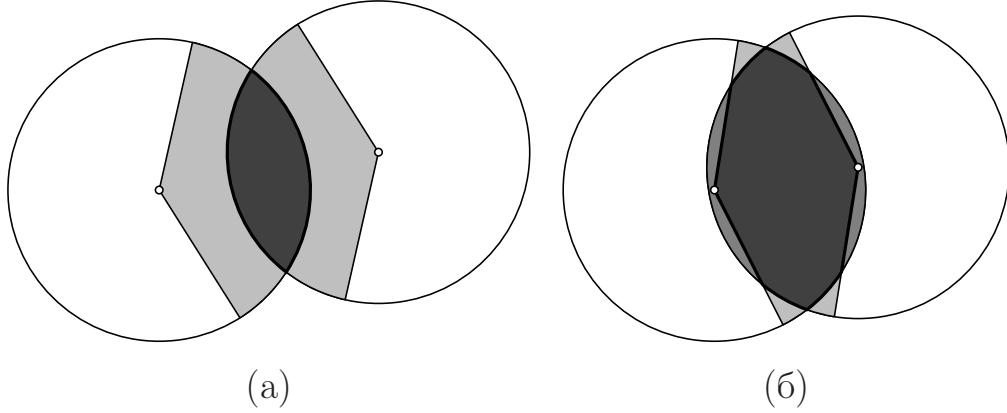


Рисунок 2.12 — Взаимное положение пары пересекающихся внешних секторов усечённых бициклов

Обозначим:

S_{disk} — площадь концевых кругов бициклов,

S_{lens} — площадь линзы, образованной пересечением концевых кругов,

$S_1^{(ext)}$, $S_2^{(ext)}$ — площади внешних секторов бициклов,

$S_1^{(int)}$, $S_2^{(int)}$ — площади внутренних секторов концевых кругов.

Внутренний сектор — это дополнение внешнего сектора в концевом круге. Внутренние сектора смежных усечённых бициклов не имеют пересечений между собой. Поскольку угловые размеры внешних секторов не превосходят π , получается, что размеры внутренних секторов не меньше π .

Лемма 2.3. *Площадь внешней зоны, образованной парой внешних секторов двух пересекающихся усечённых бициклов, равна сумме площадей этих секторов за вычетом площади линзы, образованной пересечением концевых кругов:*

$$S = S_1^{(ext)} + S_2^{(ext)} - S_{lens}. \quad (2.9)$$

Доказательство. Общая площадь объединения двух пересекающихся концевых кругов равна $2S_{disk} - S_{lens}$.

Поскольку внутренние сектора кругов не пересекаются, получаем:

$$S = (2S_{disk} - S_{lens}) - (S_1^{(int)} + S_2^{(int)}).$$

Очевидно, $S_{disk} = S_1^{(ext)} + S_1^{(int)} = S_2^{(ext)} + S_2^{(int)}$.

С учетом этого получаем искомое соотношение для площади внешней зоны

$$S = (S_{disk} - S_1^{(int)}) + (S_{disk} - S_2^{(int)}) - S_{lens} = S_1^{(ext)} + S_2^{(ext)} - S_{lens}.$$

Лемма доказана. ■

Пусть φ_1, φ_2 — угловые размеры двух пересекающихся внешних секторов. Тогда

$$S_1^{(ext)} = (\varphi_1 \cdot r^2)/2, \quad S_2^{(ext)} = (\varphi_2 \cdot r^2)/2.$$

Угловые размеры линзы, образованной двумя окружностями радиуса r , с центрами, расположенными на расстоянии $h < 2r$ друг от друга, составляют $\theta = \arccos(h/2r)$.

Площадь этой линзы

$$S_{lens} = r^2(\theta - \sin \theta).$$

Таким образом, площадь (2.9) внешней зоны пары пересекающихся бициклов равна

$$S = (\varphi_1 \cdot r^2)/2 + (\varphi_2 \cdot r^2)/2 - r^2(\theta - \sin \theta). \quad (2.10)$$

Далее мы докажем, что для расчёта площади покрытия достаточно рассматривать пары «ближайших» усечённых бициклов, между которыми нет других усечённых бициклов.

Лемма 2.4. *Любая точка дискового r -покрытия:*

- либо принадлежит собственной области полного или усечённого бицикла;
- либо покрывается внешним сектором концевого круга усечённого бицикла, не имеющего смежных;
- либо входит во внешнюю зону пары смежных усечённых бициклов.

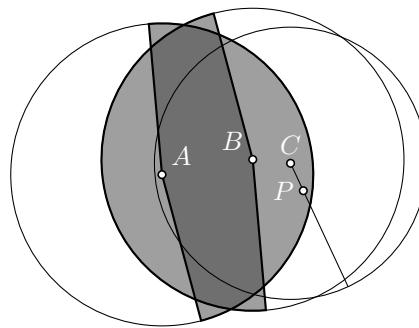


Рисунок 2.13 — К доказательству леммы 2.4

Доказательство. Покажем, что при исключении первых двух случаев вступает в силу третий. Пусть точка P не принадлежит собственной области

или изолированному внешнему сектору. Тогда, так как точка входит в покрытие, она принадлежит хотя бы одному из внешних секторов концевых кругов усечённых бициклов, причём все такие бициклы имеют смежные. Спица, которой принадлежит точка P , не входит в собственные области, значит, её длина, а следовательно, и радиус вписанного круга с центром в основании спицы меньше r . Построим опорный путь из основания спицы C до центра меньшего концевого круга одного из бициклов, внешний сектор которого покрывает точку, и возьмём первую точку на этом пути со значением радиальной функции, равным r (обозначим её A). Вписанный круг с центром в A — промежуточный на пути и покрывает точку P . При этом, так как путь проходит по отсечённой части ребра с радиальной функцией меньше r , а эта часть расположена со стороны внешнего сектора, точка P попадает именно во внешний сектор. Рассмотрим бицикл, смежный к найденному, центр его меньшего концевого круга обозначим B . Чтобы точка попадала во внешний сектор, но не попадала во внешнюю зону, необходимо, чтобы она попадала во внутренний сектор смежного бицикла, как показано на рис. 2.13. Но тогда спицы, выпущенные из точки B , пересекают как верхнюю, так и нижнюю границу объединения кругов с центрами в точках A и C , поэтому опорный путь между ними проходит через B . Это противоречит тому условию, что A — первая точка радиуса r на этом пути. Поэтому точка C не может лежать вне внешней зоны. Лемма доказана. ■

С помощью рассуждений, аналогичных доказательству леммы, можно прийти к выводу, что внешний сектор, не имеющий смежных, полностью лежит вне собственных областей: иначе длина спицы с точкой P не меньше r , и на пути от A к C найдётся точка с радиальной функцией, равной r , что противоречит отсутствию смежных бициклов.

Более сложным представляется случай пересечения трёх и более внешних секторов. Возможные варианты пересечения трёх одинаковых кругов представлены на рис. 2.14. Однако, как будет показано ниже, при пересечении трёх усечённых бициклов варианты, изображенные на рис. 2.14а,б, невозможны.

Рассмотрим сначала случай, когда линза в пересечении двух кругов целиком лежит внутри третьего круга (рис. 2.14а). В примере линза кругов A и C накрывается кругом B .

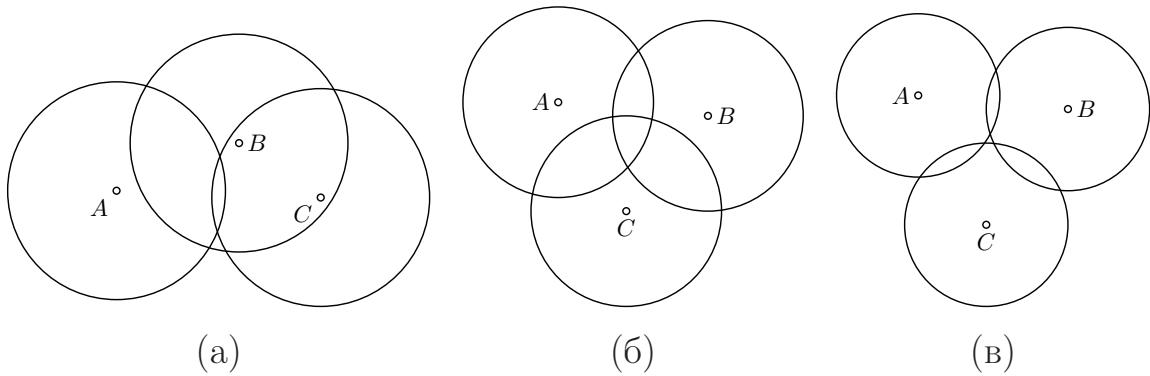


Рисунок 2.14 — Пересечения трех концевых кругов усечённых бициклов

Лемма 2.5. Пусть A, B, C — одинаковые круги и линза в пересечении кругов пары A и C лежит внутри круга A (рис. 2.15а). Тогда сумма дуг круга B , накрытых кругами A и C , больше π .

Доказательство. Обозначим α, β угловые размеры линз, лежащих в пересечении пар кругов $A \cap B$ и $B \cap C$. Изменим положение круга C на C' так, чтобы в $\triangle ABC'$ было $BC' = BC$, $\angle B = (\alpha + \beta)/2$ (рис. 2.15б). Треугольник $\triangle ABC'$ получается из $\triangle ABC$ путем поворота стороны BC так, чтобы в точке D совместились углы линз пар кругов $A \cap B$ и $B \cap C$. Построим касательную GH для круга B в точке D . Построим также хорды DE и DF , проходящие через угловые точки линз. Поскольку в $\triangle ABC'$ сторона $AC' \leq AC$, это означает, что круг C' лежит не дальше от круга A , чем C , и поэтому тоже имеет пересечение с A . Из простых геометрических рассуждений получаем:

$$\angle GDE = \alpha/2, \angle FDH = \beta/2.$$

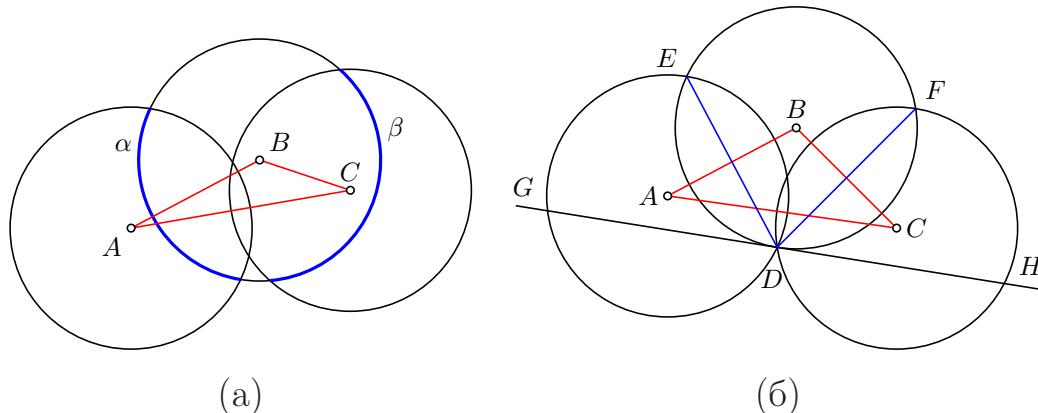


Рисунок 2.15 — К доказательству леммы 2.5

Условие пересечения линз состоит в следующем:

$$2 \cdot \angle GDE + 2 \cdot \angle FDH > \pi,$$

что доказывает утверждение леммы $(\alpha + \beta) > \pi$. ■

Следствием леммы 2.5 является вывод, что круг B не может быть меньшим кругом усечённого бицикла, смежного с обоими остальными, поскольку его внешний сектор не меньше $\alpha + \beta$ и в этом случае круг не является минимальным в бицикле. Таким образом, бицикл с концевым кругом B не может быть смежным к бициклам с кругами A и C одновременно. Тем не менее возможна ситуация, когда одна из дуг α и β принадлежит внешнему сектору, а другая — внутреннему. В этом случае бициклы, соответствующие кругам A и C , всё равно не могут быть смежными, так как опорный путь между ними должен проходить через точку B . Наконец, если обе дуги принадлежат внутреннему сектору концевого круга, опорные пути из A в B и из C в B имеют общий участок DB . Но радиальная функция возрастает в начале движения из B в D , так как путь проходит через внутренний сектор бицикла. Пусть в этом случае B_1 — первая точка с радиальной функцией, равной r , на пути из A в B . Точка B_1 отлична от B , но если B_1 лежит на участке DB , можно повторить те же рассуждения, заменив точку B на B_1 , что приводит нас к противоречию. Поэтому радиальная функция в точке D не меньше r и бициклы при точках A и C в этом случае не могут быть смежными.

Рассмотрим теперь второй случай, представленный на рис. 2.14б.

Лемма 2.6. *Если в дисковом r -покрытии концевые круги трёх бициклов имеют непустое пересечение, то в объединении этих кругов существует пустой круг с радиусом r .*

Доказательство. Пусть три круга с центрами в точках A, B, C и радиусом r имеют непустое пересечение (рис. 2.16). Обозначим α_{AB}, α_{BC} угловые размеры линз, лежащих в пересечении пар кругов $A \cap B$ и $B \cap C$ соответственно. Угловые размеры дуг внешних секторов не превышают π , поэтому $\alpha_{AB} \leq \pi$ и $\alpha_{BC} \leq \pi$.

Область пересечения трёх кругов представляет собой круговой треугольник $\triangle D_1E_1F_1$, сторонами которого являются дуги $\gamma_A, \gamma_B, \gamma_C$ кругов A, B, C . Треугольник $\triangle DEF$ лежит в объединении кругов A, B, C . Для доказательства утверждения леммы достаточно показать, что радиус описанной окружности

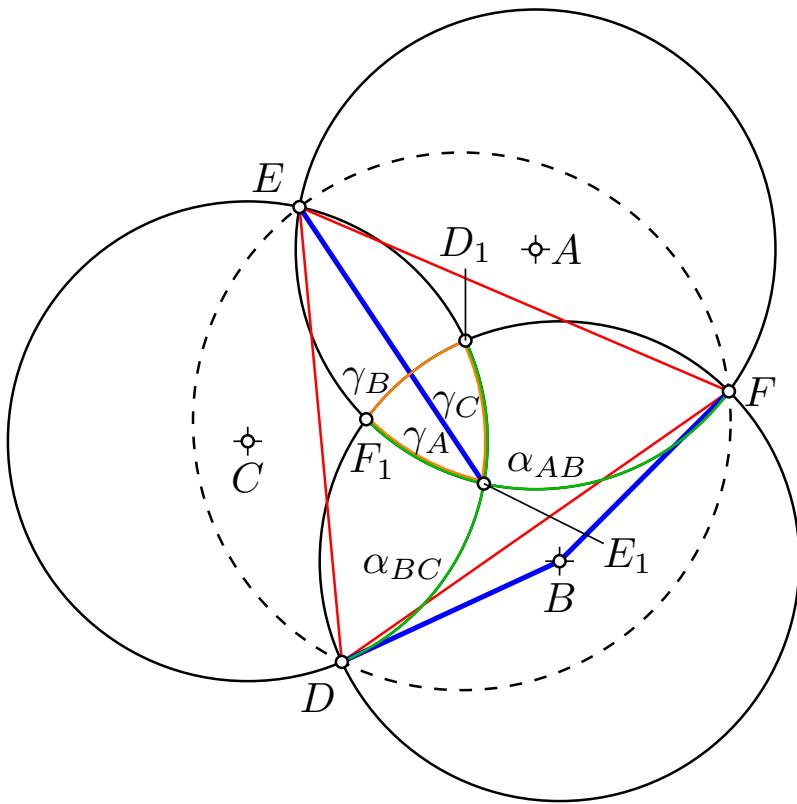


Рисунок 2.16 — К доказательству леммы 2.6

$\triangle DEF$ больше или равен r . Имеем:

$$\begin{aligned}\angle DEF &= \angle DEE_1 + \angle FEE_1, \\ \angle DEE_1 &= (\alpha_{BC} - \gamma_C)/2, \\ \angle FEE_1 &= (\alpha_{AB} - \gamma_A)/2, \\ \angle DEF_1 &= ((\alpha_{BC} + \alpha_{AB}) - (\gamma_A + \gamma_C))/2, \\ \angle DBF &= \alpha_{BC} + \alpha_{AB} - \gamma_B.\end{aligned}$$

В круговом треугольнике $\triangle D_1E_1F_1$ выполняется неравенство треугольника $\gamma_B \leq \gamma_A + \gamma_C$. Действительно, круговой треугольник со сторонами, составленными из дуг окружностей одинакового радиуса, можно рассматривать как плоскую развертку сферического треугольника на сфере такого же радиуса. В этом случае окружности равны большим кругам этой сферы и длина сторон сферического треугольника совпадает с длиной дуг кругового треугольника. Поскольку для сферического треугольника, как известно, неравенство треугольника выполняется, оно выполняется и для кругового треугольника тоже.

Отсюда следует, что

$$\angle DBF \geq 2 \cdot \angle DEF. \quad (2.11)$$

В $\angle DBF$ имеем

$$DF = 2r \cdot \sin(\angle DBF/2).$$

А в $\triangle DEF$ согласно теореме синусов:

$$DF / \sin(\angle DBF/2) = 2R^*, \quad (2.12)$$

где R^* — радиус описанного круга $\triangle DEF$, который, в свою очередь, является вписанным в объединение кругов A, B, C . Поскольку углы $\angle DBF < \pi$, а $\angle DEF < \pi/2$, получаем из 2.11

$$\sin(\angle DBF/2) \geq \sin(\angle DEF).$$

Отсюда из равенства 2.12 имеем

$$R^* = \frac{DF}{2 \cdot \sin(\angle DBF/2)} = \frac{2r \cdot \sin(\angle DBF/2)}{2 \cdot \sin(\angle DEF)} \geq r.$$

Лемма доказана. ■

Лемма 2.7. *Если концевые круги трёх бициклов имеют непустое пересечение, то бициклы образуют не более одной смежной пары.*

Доказательство. При типе пересечения кругов, соответствующем рис. 2.14а, из леммы 2.5 следует, что бициклы с концевыми кругами A и C — несмежные и не могут быть смежными к бициклу с кругом B одновременно. Таким образом мы исключаем две из трёх возможных пар смежных бициклов.

При типе пересечения кругов, соответствующем рис. 2.14б, согласно лемме 2.6 существует круг радиуса r или больше, лежащий в объединении трёх концевых кругов бициклов. Поскольку концевые круги являются пустыми, то и этот круг пустой. Тогда существует максимальный пустой круг, его содержащий и имеющий радиус не меньше r . Центр этого максимального круга является точкой скелета, которая лежит на опорном пути между центрами концевых кругов бициклов. Следовательно, условие смежности бициклов не выполняется и бициклы смежными не являются, что и требовалось доказать. ■

Это показывает, что варианты пересечения пары усечённых бициклов, смежных к третьему (рис. 2.14а,б), невозможны. Единственный возможный вариант пересечения тройки усечённых бициклов — это попарные пересечения, как в примере на рис. 2.14в. Однако, чтобы избежать «двойного счёта»,

необходимо правильно обработать ситуацию, при которой усечённый бицикл имеет сразу несколько смежных, а значит, одна и та же область попадает в несколько различных внешних зон. Эта проблема решается довольно просто, так как можно сформулировать аналог леммы 2.3 сразу для всех внешних секторов усечённых бициклов, учитывая возможность одного бицикла иметь сразу несколько смежных.

Лемма 2.8. *Площадь области, покрытой внешними секторами концевых кругов и лежащей вне собственных областей, равна сумме площадей этих секторов за вычетом площадей линз, образованных пересечением концевых кругов смежных бициклов.*

Доказательство. Для бицикла, имеющего несколько смежных, тройные пересечения секторов невозможны, поэтому дуги сектора, накрытые другими секторами, отделимы друг от друга. Но тогда каждый сектор можно разделить на части таким образом, что каждая часть будет пересекаться только с одним сектором (рис. 2.17). Внешние зоны частичных секторов не будут иметь существенных пересечений друг с другом и будут полностью лежать вне собственных областей, поэтому для них также справедлива формула из леммы 2.3. Но из площадей частичных секторов собирается сумма площадей полных секторов, а каждое пересечение частичных секторов равно пересечению соответствующих полных секторов, поэтому из площадей пересечения частей складывается общая сумма площадей пересечений, что и доказывает утверждение леммы. ■

В итоге приходим к утверждению основной теоремы.

Теорема 2.1. *Площадь дискового покрытия многоугольной фигуры равна сумме площадей собственных областей всех бициклов и внешних секторов усечённых бициклов за вычетом площадей парных пересечений смежных усечённых бициклов.*

Доказательство. Часть дискового покрытия — это подмножество фигуры, образованное объединением всех спиц длины r . Площадь этой части равна сумме площадей всех широких и усечённых бициклов. Остальная часть дискового покрытия состоит из внешних секторов усечённых бициклов. Как следует из лемм 2.5–2.7, возможны лишь попарные пересечения внешних секторов усечённых бициклов. Суммарная площадь этих внешних секторов согласно леммам 2.3 и 2.8 равна сумме площадей этих секторов за вычетом общей площади линз,

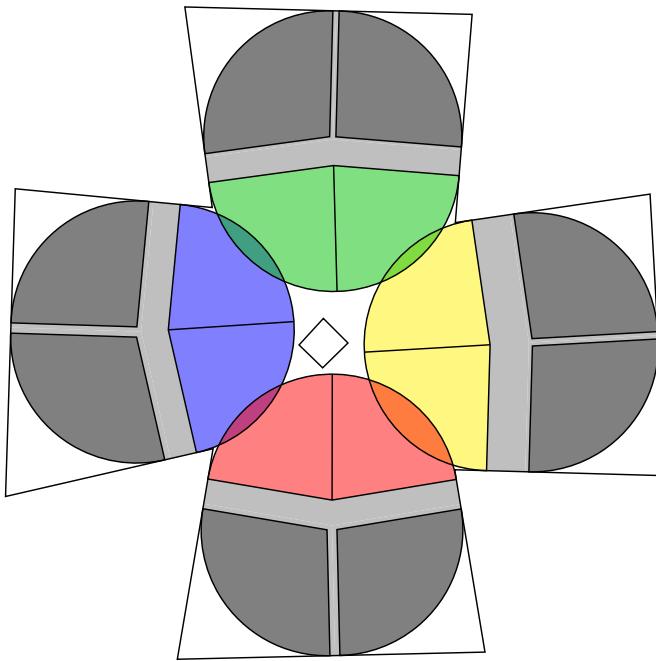


Рисунок 2.17 — Пересечение группы бициклов

лежащих в пересечениях концевых кругов смежных усечённых бициклов. Теорема доказана. ■

В итоге для решения задачи 2.2 остаётся решить задачу поиска смежных бициклов и обновления графа их смежности.

2.4.2 Поиск смежных бициклов

Задача 2.3. Пусть $\mathcal{B} = \{\mathcal{B}\}_{i=1}^k$ — множество усечённых бициклов в дисковом r -покрытии многоугольной фигуры X . Найти множество пар смежных бициклов \mathcal{E} .

Такое определение множеств \mathcal{B} и \mathcal{E} позволяет говорить о *графе смежных бициклов* $\mathcal{G} = (\mathcal{B}, \mathcal{E})$.

Очевидно, процедура прямого построения опорного пути для каждой пары бициклов с целью проверки смежности является слишком затратной с точки зрения объёма вычислений. Однако можно организовать процедуру поиска таким образом, что граф смежности на каждом шаге перестраивается локально: мы дополняем имеющиеся пути очередной вершиной и проверяем, не слишком ли отдалились круги на концах путей. Итоговый граф смежности в дисковом покрытии при заданном r будет результатом ряда последовательных перестроений при меньших значениях радиуса. Для организации обхода скелета при поиске смежных бициклов мы придадим рёбрам направление: концом ребра

будем считать вершину с большим радиусом, соответственно, другая вершина будет началом ребра.

Определение 2.6. Скелет, в котором рёбра монотонные и ориентированы по направлению возрастания радиальной функции, будем называть ориентированным скелетом.

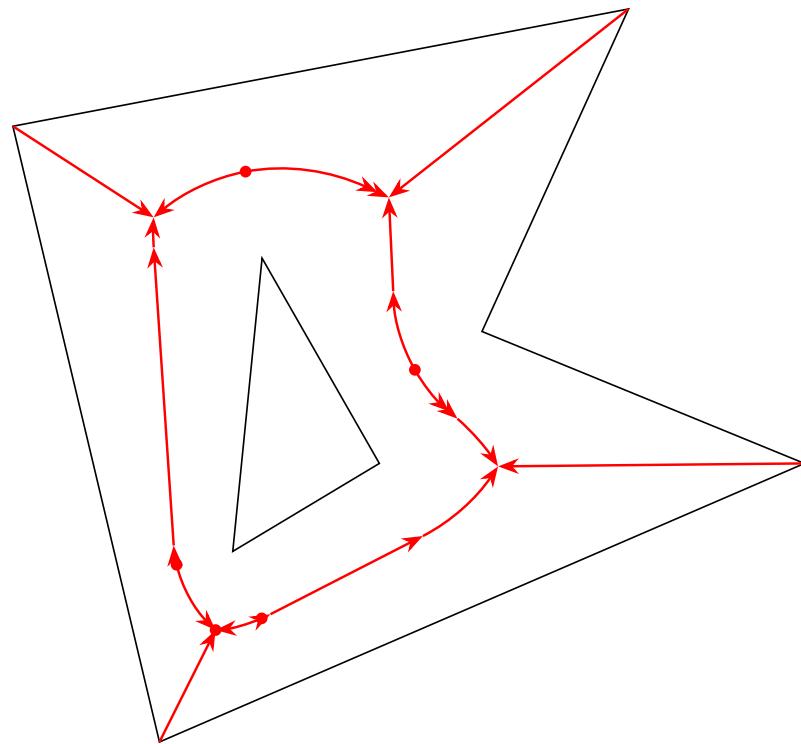


Рисунок 2.18 — Ориентированный скелет

Пример ориентированного скелета показан на рис. 2.18. Определим возможные полустепени захода и исхода такого скелета. Для каждого ребра угловая величина внутреннего сектора меньшего концевого круга не меньше π , а величина большего сектора — не больше π . При этом сумма внутренних секторов нетерминальных вершин скелета равна 2π , а радиальная функция в терминальной вершине равна нулю. Это означает, что возможны следующие сочетания числа входящих и исходящих рёбер (мы по-прежнему рассматриваем скелет общего положения, в котором все вершины имеют степень не более 3):

- одно исходящее ребро, если вершина терминальная (рис. 2.19 α);
- два исходящих ребра (рис. 2.19 β);
- одно исходящее ребро и одно входящее (рис. 2.19 γ);
- одно исходящее ребро и два входящих (рис. 2.19 δ);
- три исходящих ребра (рис. 2.19 ϵ).

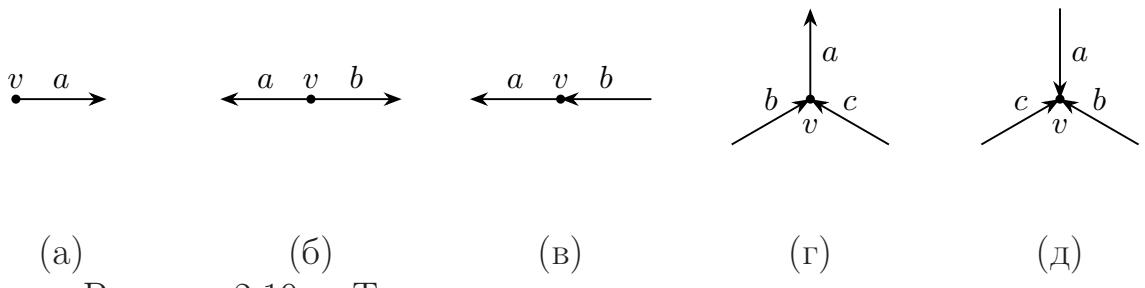


Рисунок 2.19 — Типы вершин в ориентированном скелете

Определение 2.7. Вершину скелета с m входящими и n исходящими рёбрами будем называть вершиной типа (m,n) .

При возрастании радиуса дискового покрытия структура смежности усечённых бициклов сохраняется, если сохраняется само множество усечённых бициклов — возможно лишь, что концы усечённых бициклов достаточно отдалились друг от друга, и линза в их пересечении пропала. Таким образом, признаком изменения структуры смежности является изменение статуса бицикла — с полного на усечённый и с усечённого на узкий, что происходит в том случае, когда радиус покрытия превышает радиус вписанного круга с центром в одной из вершин скелета. Статус меняется для бициклов, инцидентных вершине, а структура смежности перестраивается в окрестности вершины. Вершины обрабатываются в порядке возрастания их радиуса, а способ обработки вершины зависит от её типа.

1. Терминальные бициклы не могут иметь смежные, поэтому терминальные вершины игнорируются.
2. Вершина с двумя исходящими рёбрами порождает пересечение соответствующих бициклов.
3. Вершина имеет исходящее ребро a и входящее b . Тогда бициклы, которые были смежными с b , могут стать смежными с a . Бицикл передаётся от b к a , если его крайняя точка отдалась от обрабатываемого круга на расстояние не более $2r$ — его удвоенного радиуса, и пересечение не исчезло. Отметим также, что процедуре усечения смежного бицикла всегда можно произвести, так как радиус его большего концевого круга не меньше радиуса обрабатываемого круга.
4. Вершина с исходящим ребром a и входящими рёбрами b и c . В этом случае бициклы, смежные с b и c , также могут быть переданы a . При этом новые пересечения между смежными бициклами образоваться не могут: по свойствам опорного пути они должны пересекаться внутри

рассматриваемого круга, причём согласно лемме 2.5 сумма дуг, покрытых концевыми кругами смежных бициклов, окажется больше π , но круг является меньшим концевым для бицикла a , поэтому покрыт может быть только его внешний сектор, который не больше π .

5. Вершина имеет три входящих ребра a , b и c . Исходящих рёбер нет, поэтому бициклы, смежные ко входящим, не передаются, зато могут образовать новые пересечения между собой. При этом опорные пути снова должны проходить через вершину, а бициклы пересекаться внутри её круга. Заметим, что при радиусе покрытия, равном r , оси бициклов a , b и c вырождаются в точки, а внешние сектора их концевых кругов складываются во весь круг³. В силу отсутствия тройных пересечений дуги, покрытые смежными бициклами, попарно пересекаются не более чем по одной точке. Сумма угловых величин всех дуг не превышает 2π , а пересечение может образоваться только в том случае, если сумма величин пары дуг превышает π . Поэтому большая дуга в пересечении должна быть больше $\frac{\pi}{2}$, но таких дуг не может быть больше трёх. Таким образом, достаточно проверять только пересечения, в которых задействована одна из трёх наибольших дуг (факт наличия пересечения зависит не только от величины дуг, но и от их взаимного положения), и мы имеем не квадратичный по числу бициклов перебор пар, а линейный. Пусть $\alpha_1, \alpha_2, \dots, \alpha_k$ — угловые величины покрытых дуг, упорядоченные по убыванию. Заметим, что $\alpha_2 + \alpha_4 \leq \pi$, так как иначе $\alpha_1 + \alpha_3 > \pi$ и $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 > 2\pi$, что невозможно. Поэтому бициклы со второй и четвёртой наибольшими дугами не пересекаются. Аналогично можно прийти к выводу, что первый и четвёртый, и второй и третий бициклы не могут пересекаться одновременно. Таким образом, все возможные пересечения либо включают в себя только бициклы из первой тройки, либо во всех из них участвует первый бицикл. Поэтому число образовавшихся пересечений не превышает числа смежных бициклов, то есть числа ликвидированных пересечений.

³Это свойство может нарушаться, если один из бициклов — линейный с одинаковыми радиусами концевых кругов. Однако в таком случае сектор обрабатываемого круга, который может быть покрыт со стороны этого бицикла, меньше сектора, о котором идёт речь в рассуждении, поэтому рассуждение остаётся справедливым.

Описанную выше процедуру поиска смежных бициклов назовём *инкрементным прослеживанием*.

Теорема 2.2. *Граф, полученный в результате инкрементного прослеживания, является графом смежных бициклов.*

Доказательство. Так как существует взаимно однозначное соответствие между точками скелета и вписанными кругами, будем для удобства говорить об опорном пути между кругами и считать, что опорный путь состоит из множества кругов, а не точек. Заметим для начала, что если пересечение кругов A и B вложено во все круги на некотором пути из A в B , являющиеся концевыми кругами бициклов, то оно вложено вообще во все вписанные круги на этом пути. В этом легко убедиться с помощью рис. 2.5: для бицикла любого типа пересечение его концевых кругов вложено в любой из промежуточных.

Утверждение теоремы докажем по индукции. Инкрементное прослеживание начинает работу с пустого графа, что корректно для 0-покрытия, в котором нет смежных бициклов. Два исходящих ребра может быть только у вершины, полученной в результате разбиения немонотонного ребра на пару монотонных. Очевидно, что опорный путь между точками таких рёбер лежит через их общий концевой круг C , и это единственный концевой круг на опорном пути. Таким образом, процедура способна поддерживать все опорные пути длины не больше 1. Покажем, что сохраняются все опорные пути длины k , если сохраняются все опорные пути меньшей длины. Пусть опорный путь между смежными бициклами содержит концевые круги $C_1C_2\dots C_k$, а C_t – круг с максимальным радиусом среди них. Если $t = k$ (случай $t = 1$ рассматривается аналогично), последовательность кругов должна получиться в результате шага третьего или четвёртого типа инкрементной процедуры из последовательности $C_1C_2\dots C_{k-1}$. Если круг с максимальным радиусом находится в середине последовательности (то есть $1 < t < k$) последовательность получается в результате шага пятого типа из последовательностей $C_1C_2\dots C_{t-1}$ и $C_{t+1}C_{t+2}\dots C_k$. Так как часть опорного пути по построению также представляет собой опорный путь, любой опорный путь в итоге сводится к последовательностям длины 1, то есть к базе индукции. Таким образом, ни один опорный путь не теряется.

Лишние рёбра в графе смежных бициклов не возникают, так как мы на каждом шаге можем в явном виде получить последовательность концевых

кругов на потенциальном опорном пути и проверить вложенность линзы-пересечения в каждый из кругов. Теорема доказана. ■

Отметим, что последовательность кругов на опорном пути можно не хранить в явном виде. Если последовательность кругов $C_1C_2\dots C_k$ получена добавлением круга C_k , достаточно проверить вложенность только в круги C_1 , C_{k-1} и C_k , так как $C_1 \cap C_k \in C_i, 1 < i < k$ по определению опорного пути. Аналогично, в случае образования последовательности из двух частей $C_1C_2\dots C_{t-1}$ и $C_{t+1}C_{t+2}\dots C_k$ достаточно проверить круги $C_1, C_{t-1}, C_t, C_{t-1}$ и C_k . Но, за исключением круга C_t , оставшиеся в обоих случаях круги будут меньшими концевыми кругами бициклов в «унаследованных» пересечениях. Таким образом, при обработке вершины в ходе инкрементного прослеживания для каждой пары возможных смежных бициклов достаточно проверить вложенность линзы в круг с центром в вершине, меньшие концевые круги этих бициклов и меньшие концевые круги их смежных бициклов на предыдущем шаге.

Алгоритм инкрементного прослеживания даёт нам первое решение задачи 2.3.

2.4.3 Оценки числа смежных бициклов

Согласно алгоритму построения графа смежных бициклов, рост числа пересечений возможен только при обработке вершины, имеющей два исходящих ребра, а после обработки вершины другого типа число пересечений останется в лучшем случае прежним. Вершины с двумя исходящими рёбрами образуются при разбиении немонотонного ребра на пару монотонных, поэтому число смежных пар бициклов не превышает числа немонотонных рёбер в исходном скелете.

Можно также оценить число смежных пар, исходя из общего числа вершин скелета.

Лемма 2.9. Число пар смежных бициклов в r -покрытии не превышает $\frac{3n-12}{5}$, где n — число вершин скелета. Предполагается, что скелет не содержит вершин степени больше 3.

Доказательство. Обозначим число вершин каждого типа по их полустепеням захода и исхода как $n_{01}, n_{02}, n_{11}, n_{21}$ и n_{30} . Нам нужно оценить величину n_{02} . Так как общее число исходящих рёбер по всем вершинам равно общему

числу входящих, имеем соотношения:

$$\begin{aligned} n_{11} + 2n_{21} + 3n_{30} &= n_{01} + 2n_{02} + n_{11} + n_{21}, \\ 2n_{02} &= 3n_{30} + n_{21} - n_{01}, \\ 2n_{02} &= 3(n - n_{01} - n_{02} - n_{11} - n_{21}) + n_{21} - n_{01}, \\ 5n_{02} &= 3n - 4n_{01} - 3n_{02} - 3n_{11} - 2n_{21} \leq 3n - 4n_{01}. \end{aligned}$$

Но число терминальных вершин n_{01} равно количеству выпуклых вершин внешней границы и вогнутых вершин дыр, поэтому таких вершин не меньше трёх. В результате получаем заявленную оценку. ■

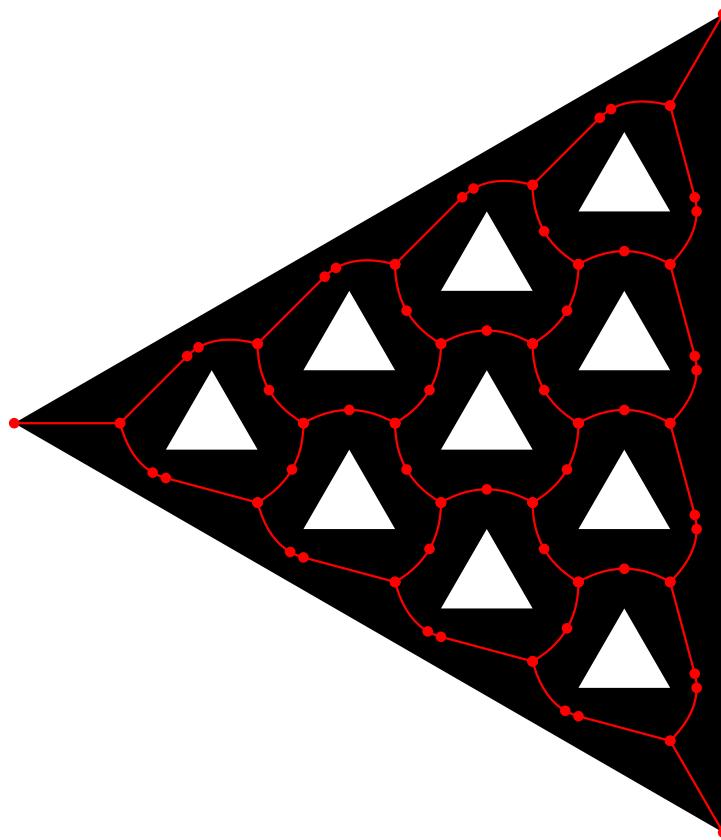


Рисунок 2.20 — Скелет с наибольшим удельным числом минимальных точек

Оценка доли минимальных точек в $\frac{3}{5}$ достигается асимптотически на бесконечной фигуре, в которой повторяется определённый шаблон, пример которого приведён на 2.20. Посчитаем общее число вершин и число минимальных вершин в фигуре, содержащей бесконечное число треугольников. При добавлении k -го ряда из k треугольников в скелете возникают $4(k - 1)$ рёбер между новыми треугольниками и треугольниками из предыдущего ряда, а также $2(k - 1)$ рёбер между парами новых треугольников. При этом число вершин, не имеющих среди образующих сайты сайты внешнего контура, возрастает на $5(k - 1) - 1$, из которых $3(k - 1)$ точек минимума. Часть границы, опирающаяся

на внешний контур, состоит из $9k + 3$ вершин, из которых $3k$ — точки минимума. В итоге для фигуры с n рядами треугольников число вершин V_n и число минимальных вершин V'_n определяются по формулам:

$$\begin{aligned} V_n &= \left(\sum_{k=2}^n 5(k-1) - 1 \right) + 9n + 3 = \\ &= \left(\frac{5(n+2)(n-1)}{2} - n + 1 \right) + 9n + 3 = \frac{5}{2}n^2 + \frac{21}{2}n - 1, \\ V'_n &= \left(\sum_{k=2}^n 3(k-1) \right) + 3n = \frac{3(n+2)(n-1)}{2} + 3n = \frac{3}{2}n^2 + \frac{9}{2}n - 3. \end{aligned}$$

Таким образом,

$$\lim_{n \rightarrow \infty} \frac{V'_n}{V_n} = \lim_{n \rightarrow \infty} \frac{\frac{3}{2}n^2 + \frac{9}{2}n - 3}{\frac{5}{2}n^2 + \frac{21}{2}n - 2} = \frac{3}{5}.$$

Очевидно также, что число минимальных точек составляет не больше половины от числа рёбер, поскольку каждой минимальной точке соответствует пара смежных рёбер, для каждой точки различных. Число рёбер во внешнем контуре, включая терминальные, равно $9k + 3$. Поэтому общее число рёбер E_n равно

$$E_n = \left(\sum_{k=2}^n 6(k-1) \right) + 9n + 3 = 3(n+2)(n-1) + 9n + 3 = 3n^2 + 12n - 3,$$

так что

$$\lim_{n \rightarrow \infty} \frac{V'_n}{E_n} = \lim_{n \rightarrow \infty} \frac{\frac{3}{2}n^2 + \frac{9}{2}n - 3}{3n^2 + 12n - 3} = \frac{1}{2}.$$

В итоге приведённая фигура является асимптотически оптимальной с точки зрения доли минимальных вершин в обоих смыслах. Интересно, что значения радиальной функции во всех внутренних минимальных точках будут одинаковыми, так что при некотором радиусе покрытия доля смежных бициклов также будет стремиться к оптимальной.

Тем не менее можно заметить, что сайты границы приведённой фигуры не являются сайтами общего положения, так как существуют точки скелета, равноудалённые более чем от трёх сайтов. В частности, каждый тройник фигуры, расположенный между внутренними треугольниками, на самом деле равноудалён сразу от шести сайтов — трёх сайтов-точек и трёх смежных к

ним сайтов-сегментов: бисектором смежных сайта-точки и сайта-сегмента является перпендикуляр, проведённый к сайту-сегменту из сайта-точки, и сразу три таких перпендикуляра пересекаются в тройнике. Поэтому при организации структуры данных каждый тройник будет «расщеплён» на четыре вершины, координаты и значения радиальной функции в которых будут совпадать, но образующие сайты — отличаться. Также между дубликатами тройника возникнут три дополнительных фиктивных ребра. В результате с точки зрения практической реализации алгоритма оценка предельной доли минимальных точек в $\frac{3}{5}$ оказывается завышенной. Однако её можно уточнить, пользуясь особенностями скелета как подграфа диаграммы Вороного.

Лемма 2.10. *Доля минимальных точек в скелете с монотонными рёбрами для множества граничных сайтов общего положения не превышает $\frac{1}{3}$ по отношению как к числу рёбер, так и числу вершин скелета.*

Доказательство. Начнём с того, что скелет — планарный граф, поэтому для него выполняется теорема Эйлера:

$$V - E + F = 2,$$

где V — число вершин, E — число рёбер, а F — число граней (циклов) скелета, включая внешнюю.

Общее число вершин и общее число рёбер можно выразить через число вершин каждого типа:

$$\begin{aligned} V &= n_{01} + n_{02} + n_{11} + n_{21} + n_{30}, \\ E &= n_{01} + 2n_{02} + n_{11} + n_{21}. \end{aligned}$$

Заметим, что каждая грань в скелете, в том числе и внешняя, состоит из бисекторов, имеющих общий граничный контур. Число сайтов в каждом из контуров не менее шести (контур имеет хотя бы три сайта-точки и три сайта-сегмента). При обходе цикла против часовой стрелки образующие сайты, принадлежащие этому контуру, расположены слева и также меняются в цикле: происходит переход с сайта-точки на смежный сайт-сегмент, потом на следующий сайт-точку и так далее. Такие переходы могут происходить не в каждой вершине цикла, однако общее их число равно числу сайтов в контуре, то есть не менее шести. При этом терминальное ребро скелета прослеживается при обходе одного и того же цикла дважды, и можно считать, что в терминальной вершине

скелета происходит двойной переход, так как левый сайт-сегмент меняется не на следующую за ним в списке сайт-точку, а сразу на сайт-сегмент, расположенный в списке через одну позицию. Переход не может происходить в тройнике, так как в случае общего положения граничных сайтов это точка, равноудалённая от трёх попарно несмежных сайтов. Также переход не может происходить в точке минимума, поскольку в ней образующие сайты не меняются. Поэтому переходы случаются только в вершинах типа 0–1 и 1–1. Для внутреннего контура терминальная вершина может обеспечивать переходы только между сторонами при вогнутой вершине многоугольника-дыры, но у него есть ещё как минимум 3 выпуклых вершины и 6 соответствующих сайтов, которые, таким образом, требуют нужного количества точек типа 1–1. Поэтому

$$\begin{aligned} 6(F - 1) &\leq n_{11}, \\ 6(2 + n_{02} - n_{30} - 1) &\leq n_{11}, \\ 6n_{02} - 6n_{30} - n_{11} &\leq -6. \end{aligned}$$

Пусть $p_{ij} = \frac{n_{ij}}{\sum_{t,s} n_{ts}}$. Тогда ограничение на долю минимальных вершин можно найти как решение классической задачи линейного программирования — максимизации p_{02} с условиями:

$$\left\{ \begin{array}{l} p_{01} + p_{02} + p_{11} + p_{21} + p_{30} = 1, \\ 6p_{02} - p_{11} - 6p_{30} \leq 0, \\ p_{01} + 2p_{02} - p_{21} - 3p_{30} = 0, \\ p_{ij} \geq 0 \end{array} \right. \quad \text{для всех } (i,j).$$

Отметим, что отрицательные оценки сверху в неравенствах заменены на нулевые, то есть ограничения ослаблены — впрочем, это несущественно при рассмотрении неограниченно больших n_{ij} . Решение этой задачи⁴ приводит к

⁴Так как это типовая задача линейного программирования, мы не будем детально расписывать процесс решения. Например, к данному ответу можно прийти вручную, выразив p_{21} и p_{30} через p_{01} , p_{02} и p_{11} из ограничений-равенств и перейдя к эквивалентной задаче с тремя переменными:

$$\left\{ \begin{array}{l} p_{02} \rightarrow \max, \\ 2p_{01} + 3p_{02} + p_{11} \geq 1, \\ 4p_{01} + 5p_{02} + 3p_{11} \leq 3, \\ 6p_{01} + 3p_{02} + 4p_{11} \geq 3, \\ p_{ij} \geq 0, \end{array} \right. \quad (i,j) \in \{(0,1), (0,2), (1,1)\}.$$

Окончательный ответ можно получить графическим методом при переборе троек возможных активных ограничений.

следующим значениям:

$$p_{01} = \frac{1}{3}, \quad p_{02} = \frac{1}{3}, \quad p_{11} = \frac{0}{3}, \quad p_{21} = \frac{0}{3}, \quad p_{30} = \frac{1}{3}.$$

Сформулировав аналогичную задачу для поиска оптимальной доли от числа рёбер — это можно представить как минимизацию $p_{11} + 2p_{21} + 3p_{30}$ с заменой первого из равенств на $p_{10} = 1$ — получим то же значение максимума, которое достигается в той же точке. ■

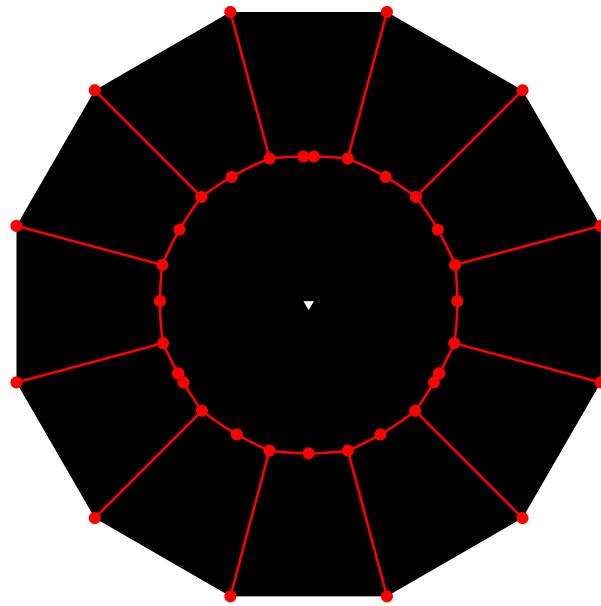


Рисунок 2.21 — Скелет общего положения с наибольшим удельным числом минимальных точек

Тип фигуры, на которой достигается это предельное значение, показан на рис. 2.21. Очевидно, что число «лопастей» в ней можно сделать неограниченно большим, а одна лопасть состоит из терминальной вершины, точки минимума и тройника с входящими рёбрами, как и было предсказано решением задачи линейного программирования.

Таким образом, число пересечений усечённых бициклов в общем случае ограничено третью от числа рёбер скелета. Несмотря на это, нет никакого более строгого с точки зрения порядка роста ограничения на число смежных бициклов у отдельного взятого бицикла. Эту ситуацию иллюстрирует рис. 2.22а, на котором у одного круга обнаруживаются сразу шесть пересечений. Хотя структура выглядит достаточно нетривиально, изображённые круги действительно могут быть концевыми кругами усечённых бициклов: концевые секторы у них не больше π , а концы дуг секторов, которые должны быть точками границы, другими кругами не покрываются. Принципиальная схема такой фигуры представляет собой фрагмент типа «веер», она приведена на рис. 2.22б и показывает,

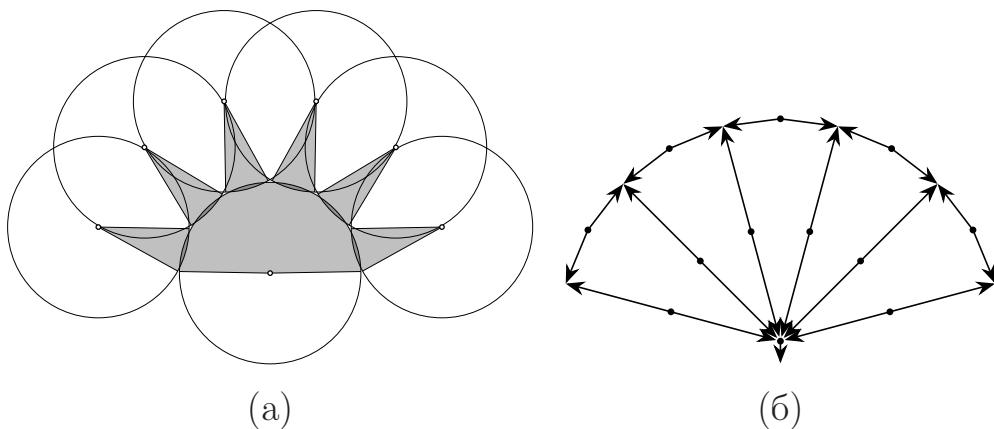


Рисунок 2.22 – Фрагмент фигуры типа «веер»: его концевые круги и общая схема скелета

что такой фрагмент в скелете возможен, так как нет вершин степени больше двух с не менее чем двумя исходящими рёбрами. Более того, число секций веера можно сделать неограниченно большим, настолько, что большинство рёбер всей фигуры сконцентрируется в веере, и число смежных пересечений у выбранного бицикла будет $O(n)$, если n — число рёбер.

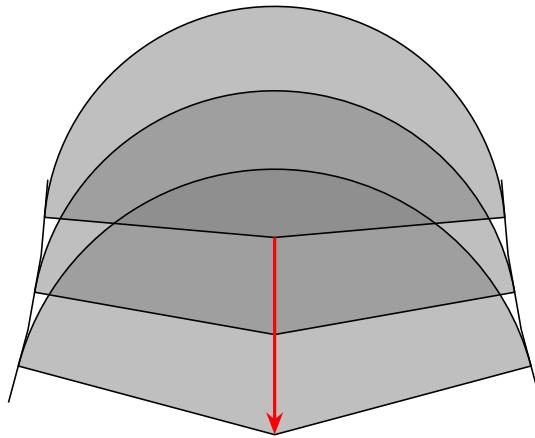


Рисунок 2.23 – Фрагмент фигуры типа «ручка» и его последовательные концевые круги

Таким образом, сложность шага обработки вершины при перестроении структуры пересечений может составить $O(n)$ (скелет — планарный граф, так что в качестве n можно рассматривать как число рёбер, так и число вершин). Это пока не даёт ответ на вопрос о сложности всей процедуры, так как необходимо оценить число таких шагов. Ситуацию проясняет рис. 2.23: при движении по расширяющемуся участку возможна передача смежных бициклов от более узкого бицикла более широкому. Но если граница состоит из множества мелких сегментов, длина которых значительно меньше радиуса круга, положение и радиус последующего круга будут незначительно отличаться от положения

предыдущего. В результате окажется возможной передача всех имевшихся пересечений по цепочке бициклов. Если прикрепить к рассмотренному ранее вееру расширяющуюся ручку, содержащую $O(n)$ вершин, то не исключено, что $O(n)$ бициклов придётся проверять $O(n)$ раз, что приводит к в худшем случае квадратичной сложности процедуры поддержания структуры пересечений. Хотя можно заметить, что, так как площадь линзы равна $r(\varphi - \sin \varphi)$, где r — радиус круга, а φ — его покрытая дуга, то для круга, имеющего k пересечений, и $k \rightarrow \infty$, дуга отдельной линзы в среднем не превышает $\frac{\pi}{k}$. Заметим также, что $\sin(x)$ при разложении в ряд Тейлора в окрестности нуля принимает вид $x - \frac{x^3}{6} + o(x^3)$, поэтому для площади всех линз S имеем формулу:

$$S \approx kr \left(\frac{\pi}{k} - \sin \frac{\pi}{k} \right) = kr \left(\frac{\pi}{k} - \frac{\pi}{k} + \frac{\pi^3}{6k^3} + o\left(\frac{1}{k^3}\right) \right) = \frac{\pi^3 r}{6k^2} + o\left(\frac{1}{k^2}\right).$$

Таким образом, сумма площадей линз имеет второй порядок малости по числу, обратному к числу линз. При практической реализации алгоритма вычисления площади покрытия это может оказаться значимым фактором, поскольку можно установить на k разумное ограничение, при превышении которого площадь линз становится сопоставимой с погрешностями машинных вычислений, и скорректировать процедуру поиска пересечений таким образом, чтобы бицикл имел не более k_{max} смежных: сохранять только k_{max} смежных бициклов с наибольшими линзами, а остальные отбрасывать. Тогда обработка одной вершины займёт $O(k_{max})$ операций, и вся процедура учёта пересечений потребует $O(nk_{max})$ операций. Также можно принять во внимание, что в результате аппроксимации дискретной сцены непрерывной получаются многоугольные фигуры с целочисленными координатами границ. Поэтому, так как фигура типа «веер» может возникнуть только при очень высокой точности приближения границ относительно размеров кругов, исходным изображением для этой фигуры может быть только изображение очень большого размера, который к тому же должен неограниченно увеличиваться с ростом числа линз, что делает ситуацию малореалистичной.

Также стоит отметить, что вся процедура поиска пересечений выполняется единственный раз за всё время работы алгоритма и ход её работы не зависит от числа значений радиуса, для которых мы должны вычислить значения функции морфологической ширины. Поэтому при большом количестве значений радиуса эффект наличия квадратичного по числу рёбер слагаемого в оценке числа операций будет нивелирован. При малой мощности множества

значений радиуса необходимость многократно проводить локальную обработку с проверками линз на сохранность в промежуточных значениях аргумента — радиусах кругов — может оказаться ощутимой. В этом случае можно задействовать альтернативную процедуру построения структуры пересечений, которая не проводит промежуточных проверок, а сразу строит для заданного радиуса весь необходимый граф смежности усечённых бициклов, руководствуясь геометрическими свойствами положения центров их меньших концевых кругов. Интуитивно ясно, что бицикл не может быть смежным к другому, если у того есть более близкий бицикл, лежащий примерно в том же направлении — неважно даже, как именно выглядит скелет в окрестности этого бицикла. Формальному изложению этой идеи посвящён следующий подраздел.

2.4.4 Лакуны дискового покрытия

Для начала введём понятие, которое поможет нам произвести группировку секторов и ограничить список кандидатов на пересечение секторами, принадлежащими одной и той же группе.

Определение 2.8. *Лакуной дискового r -покрытия называется замыкание связной компоненты в объединении спиц длины менее r .*

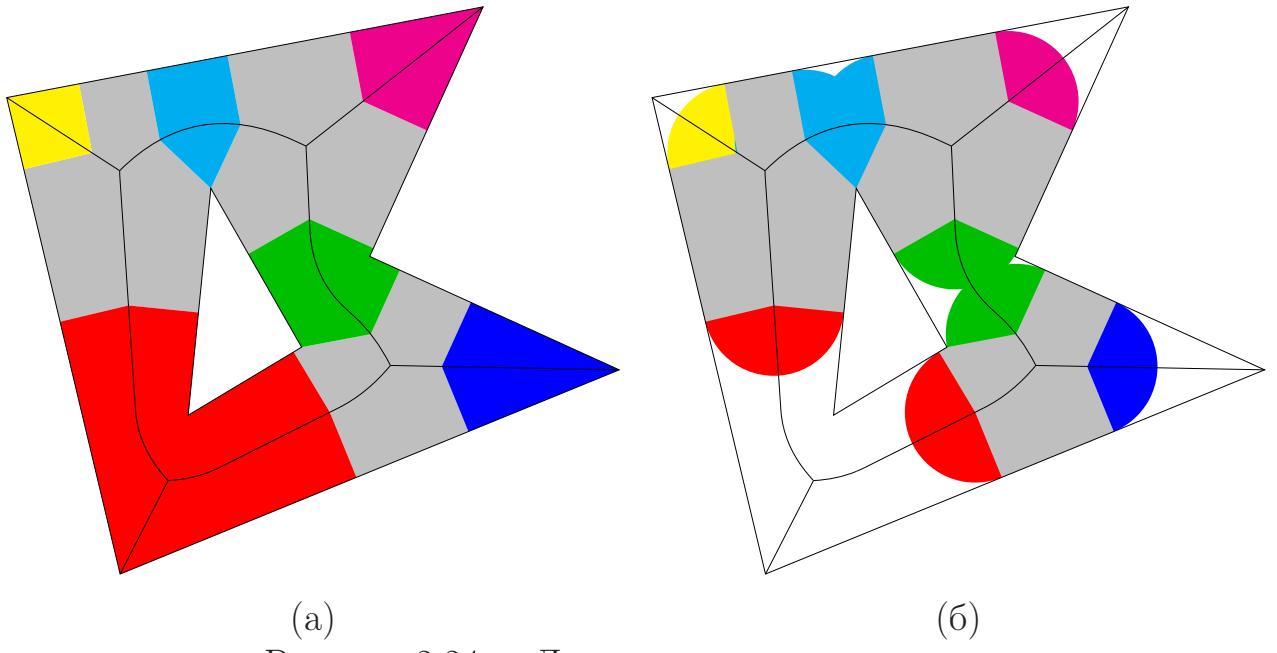


Рисунок 2.24 — Лакуны дискового покрытия

Фактически лакуны с точностью до своих границ представляют собой дополнение объединения собственных областей полных и усечённых бициклов до

всей фигуры. Сами лакуны также могут пересекаться только по своим границам.

Определение 2.9. *Осью лакуны называется множество точек скелета, расположенных в лакуне.*

Ось лакуны в r -покрытии можно получить как замыкание компоненты связности скелетного подграфа, содержащего точки скелета с радиальной функцией меньше r .

Определение 2.10. *Бицикл считается принадлежащим лакуне, если часть его ребра с радиальной функцией меньше r непуста и является частью оси лакуны.*

В силу связности осей лакун и монотонности рёбер бициклов, каждый узкий и усечённый бицикл принадлежит ровно одной из лакун, а каждый широкий бицикл не принадлежит ни одной из лакун.

Определение 2.11. *Терминальным сектором лакуны будем называть внешний сектор меньшего концевого круга усечённого бицикла, принадлежащего лакуне, а терминальной вершиной — центр этого круга.*

Если радиус большего концевого круга бицикла в точности равен r , а радиус меньшего концевого круга изначально был меньше r , то в результате коррекции, описанной в разделе 2.3, больший и меньший круги совпадут, а ребро бицикла будет состоять из единственной точки. При наличии нескольких подобных бициклов с общим большим концевым кругом центр круга может оказаться терминальной вершиной для осей сразу нескольких лакун. На рис. 2.24а серым закрашены собственные области бициклов, входящих в покрытие, а шесть образовавшихся лакун закрашены разными цветами. На рис. 2.24б теми же цветами закрашены терминальные сектора лакун.

С лакуной могут пересекаться и терминальные сектора других лакун, как показано на рис. 2.25, где пересечения секторов разных лакун закрашены смешанными цветами. Покажем, что области лакуны, покрытые секторами других лакун, покрываются секторами самой лакуны. Разобъём оси бициклов, которые становятся в r -покрытии усечёнными, на две части: с радиальной функцией не меньше r и не больше r , и будем считать, что каждая часть оси порождает свой бицикл. Рассмотрим группу бициклов, не принадлежащих лакуне, и скелетный

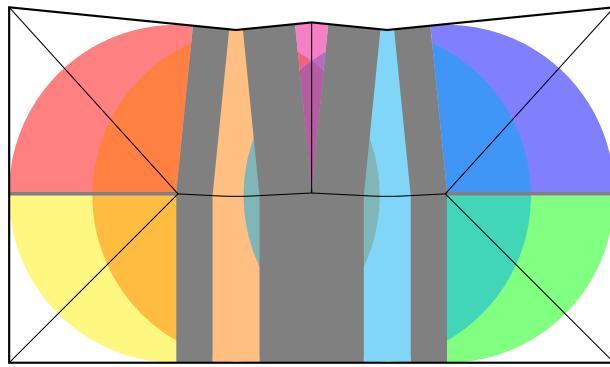


Рисунок 2.25 — Попадание «посторонних» терминальных секторов в лакуну

подграф, образованный их осями. Точки этих бициклов, лежащие в лакуне, не принадлежат ни одной из собственных областей этих бициклов, а значит, по рассуждениям из доказательства леммы 2.1 принадлежат внешним секторам концевых терминальных вершин подграфа. Но эти сектора и являются терминальными секторами лакуны, что доказывает исходное утверждение.

Так как пересечения лакун как между собой, так и с собственными областями являются множествами меры 0, при вычислении дискового покрытия можно рассматривать лакуны по отдельности. При этом для каждой лакуны достаточно определить площадь её области, покрытой связанными с ней секторами, так как прочие сектора эту область не расширяют, и добавить её к сумме площадей собственных областей полных и усечённых бициклов.

Процедура поиска смежных бициклов будет опираться на понятие триангуляции Делоне множества точек.

Определение 2.12. Триангуляцией $T(S)$ конечного множества точек S называется плоский граф, имеющий наибольшее возможное число рёбер (триангуляция $T(S)$ получена путём соединения точек из S непересекающимися прямолинейными отрезками так, что любая грань, лежащая внутри выпуклой оболочки $\text{Conv}(S)$, является треугольником).

Определение 2.13. Триангуляция $T(S)$ удовлетворяет условию Делоне, если внутри окружности, описанной вокруг любого построенного треугольника, нет ни одной точки из S . Такая триангуляция называется триангуляцией Делоне $DT(S)$.

Пример триангуляции Делоне для множества из 10 точек, сгенерированного случайным образом, приведён на рис. 2.26. Триангуляцию Делоне n точек

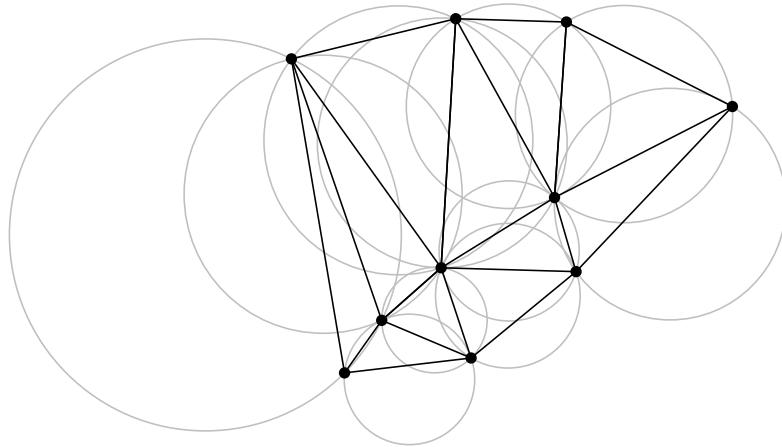


Рисунок 2.26 – Триангуляция Делоне

можно построить за время $O(n \log(n))$ с помощью алгоритма, следующего алгоритмической парадигме «разделяй и властвуй» [63].

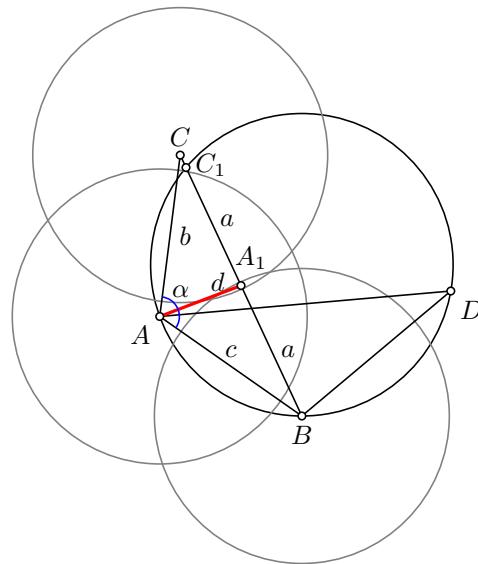


Рисунок 2.27 – К доказательству теоремы 2.3

Теорема 2.3. *Если бицикли r -покрытия являются смежными, центры их меньших концевых кругов связаны ребром в триангуляции Делоне множества терминальных точек их лакуны.*

Доказательство. Теорему будем доказывать от противного. Пусть бицикли с центрами меньших концевых кругов в точках C и B – смежные, но ребро CB отсутствует в триангуляции Делоне множества терминальных точек их лакуны. Тогда это ребро не принадлежит границе выпуклой оболочки множества терминальных точек, пересекает один из треугольников с вершиной в точке B , а также описанную окружность этого треугольника. Обозначим другие точки треугольника A и D , а точку пересечения отрезка CB с окружностью

назовём C_1 . В описанной окружности одна из дуг между точками C_1 и B не больше π . Не ограничивая общности, можно считать, что это дуга, содержащая точку A . Тогда

$$\angle BAC_1 \geq \frac{\pi}{2},$$

и

$$\angle BAC = \angle BAC_1 + \angle C_1 AC \geq \frac{\pi}{2}.$$

Покажем, что линза в пересечении концевых кругов с центрами в точках B и C захватывается кругом с центром в точке A . Проведём в треугольнике BAC медиану AA_1 . Точка A_1 лежит внутри указанной линзы. Пусть длина отрезков AB , BC , CA и AA_1 равны c , $2a$, b и d соответственно, а угол BAC имеет величину α .

По формуле длины медианы имеем:

$$d^2 = \frac{1}{2} \sqrt{b^2 + c^2 + 2bc \cos \alpha}.$$

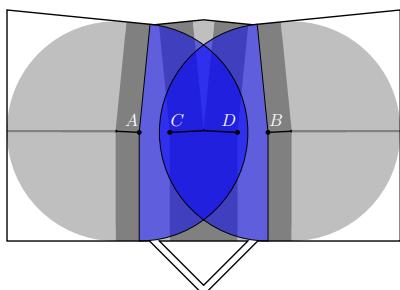
Согласно теореме косинусов,

$$b^2 + c^2 - 2bc \cos \alpha = (2a)^2,$$

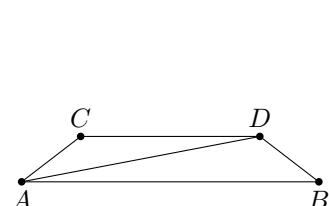
поэтому

$$d^2 = \frac{1}{2} \sqrt{(2a)^2 + 2bc \cos \alpha}.$$

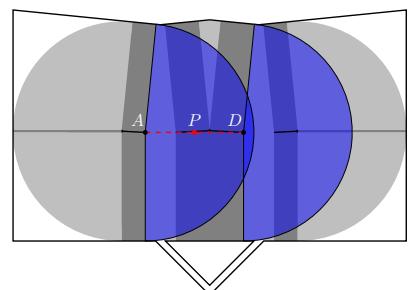
Но $\cos \alpha \leq 0$, следовательно, $d \leq a < r$ в силу пересечения концевых кругов в точках B и C . В итоге круги с центрами в точках A , B и C имеют непустое тройное пересечение, но тогда, в силу результата леммы 2.7, бициклы, соответствующие кругам B и C , не могут быть смежными. Полученное противоречие доказывает утверждение теоремы. ■



(а)



(б)



(в)

Рисунок 2.28 — Несмежные бициклы, соединённые ребром в триангуляции

Теорема 2.3 даёт возможность отфильтровать пары усечённых бициклов, заведомо не являющихся смежными. Однако присутствие соответствующего ребра в триангуляции Делоне — лишь необходимый, но не достаточный критерий. К примеру, на рис. 2.28а бициклы, соответствующие точкам A и B , принадлежат одной лакуне, но не являются смежными, однако триангуляция Делоне терминальных точек лакуны, схематично изображённая на рис. 2.28б, всё равно будет содержать ребро AB . Несмотря на это, существует простой критерий, позволяющий исключить из триангуляции лишние пары бициклов: пересечения концевых кругов смежных бициклов являются максимальными по включению. Так, пересечение кругов с центрами в точках A и C поглощает аналогичное пересечение для точек A и B . Чтобы убрать избыточные пересечения, воспользуемся следующим алгоритмом.

Рассмотрим множество усечённых бициклов. Опишем внешний сектор каждого концевого круга его начальным углом θ_0^i и угловым размером α_0^i , $0 < \alpha_0^i \leq \pi$, тогда все точки сектора будут расположены относительно центра круга под полярным углом φ , $\theta_0^i \leq \varphi \leq \theta_0^i + \alpha_0^i$. Предварительно исключим из списка рёбер в триангуляции пары бициклов, внешние сектора концевых кругов которых «не смотрят друга на друга», как это происходит для бициклов точек A и D на рис. 2.28в. Для этого воспользуемся правилом: точка P , середина отрезка, соединяющего центры кругов (пусть номера их бициклов — i_1 и i_2), должна принадлежать обоим секторам, то есть быть расположена относительно и одного, и другого центра под таким углом φ_i , что $\theta_0^i \leq \varphi_i \leq \theta_0^i + \alpha_0^i$, $i = i_1, i_2$.

Далее, поочерёдно рассматривая бициклы, вычислим соответствующие величины для всех линз, образованных выбранным сектором и секторами бициклов, которые остались связанными с текущим бициклом в триангуляции: (θ_j^i, α_j^i) , $i = 1, \dots, n$, $j = 1, \dots, m_i$. Будем считать, что пары отсортированы по первому элементу, так что $\theta_j^i \leq \theta_k^i$ при $j < k$ и что $\theta_0^i \leq \theta_j^i \leq \theta_0^i + \alpha_0^i$ для всех j (это всегда можно сделать, так как углы θ_j^i определены с точностью до периода 2π). Тогда, если

$$\theta_k^i < \theta_j^i + \alpha_j^i, \quad k > j,$$

то k -й бицикл не является смежным, так как его линза имеет пересечение с какой-то другой линзой, но не включает её полностью, следовательно, пересечение не является максимальным по включению, и ребро можно исключить из графа смежности. В результате этой процедуры в графе останутся только

ребра, задающие смежные бициклы. Сложность процедуры определяется общей сложностью сортировки списков углов для групп ребер, имеющих общую вершину, но она в худшем случае равна $O(n \log n)$, где n — число усечённых бициклов, то есть не превосходит затраты на построение самой триангуляции.

Сформировать лакуны и определить их терминальные точки можно с помощью поиска в ширину [64]. Для создания очередной лакуны необходимо найти ещё не просмотренное ребро, меньший концевой круг которого имеет радиус меньше радиуса покрытия r . Далее из центра этого круга будем двигаться во всех возможных направлениях, помечая ребра как просмотренные, до тех пор, пока не «уткнёмся» в вершину с радиусом не меньше r . Для этого предельного ребра необходимо будет произвести операцию усечения, новый центр меньшего круга и будет одной из терминальных точек лакуны. Процедура продолжается до тех пор, пока не останется ни одного ребра, из которого можно запустить процесс поиска: все ребра либо уже просмотрены, либо не входят ни в одну из лакун, так как их бициклы являются полными.

Отметим, что процедуру формирования лакун в таком виде необходимо будет запускать заново при изменении радиуса дискового покрытия. Поэтому, если предполагается вычисление площади покрытия для некоторого набора значений z_1, \dots, z_k , упорядоченного по возрастанию, можно организовать процедуру последовательного расширения лакун. Как это сделать, показано в алгоритме 1.

Такая процедура позволяет распределить ребра по лакунам за один проход по списку ребер в течение всей работы процедуры вычисления морфологической ширины. Множество терминальных точек каждой лакуны можно определить в конце каждой итерации по z путём сечения её терминальных ребер. В результате одно и то же ребро всё равно может быть рассмотрено несколько раз, если необходима его многократная коррекция.

Также возможна эвристика, позволяющая сократить сложность описанной выше процедуры и число потенциально значимых терминальных точек. Из предыдущих рассуждений известно, что более удалённые усечённые бициклы не могут пересекаться, если не пересекаются менее удалённые бициклы, расположенные на той же ветви скелета. Это значит, что иметь существенные пересечения могут только те терминальные сектора, которые пришли со стороны точек, имевших пересечения при предыдущем значении радиуса покрытия. Исходя из этого, будем определять лакуну набором её значимых ребер — рё-

Алгоритм 1 Алгоритм расширения лакун

Дано: Множество бициклов \mathcal{B} , $r(\beta)$ — радиус меньшего концевого круга, $R(\beta)$ — радиус большего концевого круга, $\ell(\beta)$ — ассоциированная лакуна для бицикла β , $\{z_j\}$ — множество аргументов функции ширины, упорядоченных по возрастанию.

Найти: Множество лакун \mathcal{L} , $\mathcal{T}(\ell)$ — множество усечённых бициклов лакуны ℓ .

- 1: Упорядочить бициклы в \mathcal{B} по $r(\beta)$
 - 2: **Для всех** $\beta \in \mathcal{B}$ **выполнять**
 - 3: $\ell(\beta) = \emptyset$
 - 4: **Конец цикла**
 - 5: $\beta = \mathcal{B}[1]$
 - 6: **Для всех** $z \in \{z_j\}$ **выполнять**
 - 7: **Для всех** $\ell \in \mathcal{L}$ **выполнять**
 - 8: **Для всех** $t \in \mathcal{T}(\ell)$ **выполнять**
 - 9: **Если** $z > R(t)$ **тогда**
 - 10: $\mathcal{T}(\ell) = \mathcal{T}(\ell) \setminus t$
 - 11: **Конец условия**
 - 12: **Конец цикла**
 - 13: **Конец цикла**
 - 14: **До тех пока** $r(\beta) < z$ **выполнять**
 - 15: $m = \{\ell(a) : a \in \mathcal{A}(\beta)\}$, где $\mathcal{A}(\beta)$ — множество бициклов, смежных с β по центру меньшего круга
 - 16: Создать новую лакуну p
 - 17: $\ell(\beta) = p$
 - 18: **Для всех** $a \in \mathcal{B}$ таких, что $\ell(a) \in m$ **выполнять**
 - 19: $\ell(a) = p$
 - 20: **Конец цикла**
 - 21: $\mathcal{T}(p) = \bigcup_{m \in m} \mathcal{T}(m)$
 - 22: **Если** $z \leq R(\beta)$ **тогда**
 - 23: $\mathcal{T}(p) = \mathcal{T}(p) \cup \{\beta\}$
 - 24: **Конец условия**
 - 25: $\mathcal{L} = (\mathcal{L} \setminus m) \cup \{p\}$
 - 26: Взять следующее β , если оно есть, иначе — закончить цикл
 - 27: **Конец цикла**
 - 28: **Конец цикла**
-

бер, задающих усечённые бициклы, имеющие существенные пересечения между собой. При необходимости ассоциировать очередное ребро с одной из лакун воспользуемся правилом: если ни одно из рёбер, смежных с заданным ребром по центру меньшего круга, не является значимым для какой-либо лакуны и не является тупиковым, создаём новую лакуну и добавляем ребро в список её значимых рёбер. Если же смежных значимых рёбер нет, но есть смежные тупиковые, считаем само ребро тупиковым. Наконец, если есть смежные значимые рёбра, объединяем лакуны и добавляем текущее ребро в список значимых рёбер общей лакуны. После того, как на данной итерации будут просмотрены все рёбра, статус которых должен измениться, необходимо произвести постобработку лакун: убрать из списка рёбра, которые заведомо не могут быть значимыми в силу неравенства $R < z_i$, а после построения триангуляции Делоне и анализа пересечений усечённых бициклов признать тупиковыми те рёбра, которые не задействованы в пересечениях. В итоге предложенная эвристика позволит исключить из рассмотрения заведомо невозможные пересечения.

Как результат, алгоритм построения графа смежных бициклов из триангуляций Делоне терминальных точек лакун даёт нам второе решение задачи 2.3 о поиске смежных бициклов.

2.5 Структура алгоритма и анализ вычислительной сложности

Результаты предыдущих разделов главы показывают, что для вычисления r -площади можно воспользоваться представлением дискового r -покрытия в виде объединения бициклов. Для этого нужно сделать следующие шаги.

1. Получить представление формы объектов на бинарном изображении в виде многоугольных фигур. Сложность данного этапа составляет $O(m)$ операций, где m — число пикселей изображения (будем считать этот шаг предварительным и не учитывать в общем анализе).
2. Построить медиальное представление многоугольной фигуры в виде скелета и радиальной функции. Найти все элементарные подграфы скелета — рёбра. Алгоритм описан в [43] и имеет сложность $O(n \log n)$, где n — число сторон многоугольной фигуры. При этом число рёбер скелета составляет $O(n)$.
3. Найти рёбра, на которых расположены точки минимума радиальной функции, и разбить их на монотонные части (раздел 1.5). Построить

множество монотонных бициклов, покрывающих многоугольную фигуру. Этот этап потребует $O(n)$ операций.

4. Для заданного значения r найти множество усечённых бициклов и вычислить положение их меньших концевых кругов (раздел 2.3). Этот этап потребует $O(n)$ операций.
5. Для полных и усечённых бициклов вычислить площади собственных областей и найти их сумму (раздел 2.2). Этот этап также потребует $O(n)$ операций.
6. Для усечённых бициклов определить внешние сектора и найти их суммарную площадь (раздел 2.3). Этот этап снова потребует $O(n)$ операций.
7. Найти все линзы в пересечении внешних секторов и вычислить их суммарную площадь (подраздел 2.4.1). Это можно сделать двумя способами: либо с помощью инкрементного прослеживания скелетного графа (подраздел 2.4.2), что потребует в худшем случае $O(n^2)$ операций на протяжении всей процедуры вычисления морфологической ширины, либо с помощью образования лакун и построения триангуляции Делоне их терминальных точек (подраздел 2.4.4). Рассмотрим второй вариант подробнее. Формирование лакун с помощью поиска в ширину потребует $O(n)$ операций для каждого значения r . Последовательное наращивание лакун будет иметь сложность $O(n \log n)$ на протяжении всей работы программы при использовании структуры «система непересекающихся множеств» с эвристикой объединения по рангу [64]. Действительно, при привязке очередного ребра к лакуне рассматриваются $O(1)$ рёбер, смежных к нему, для каждого из которых за $O(\log n)$ операций проверяется его родительская лакуна. Возможное объединение лакун потребует $O(1)$ операций, а каждое ребро при проходе по списку просматривается один раз, что приводит нас к оценке $O(n(\log n + 1)) = O(n \log n)$. Отметим, что та же самая сложность достигается при наивной реализации лакун в виде списков рёбер, но также с эвристикой объединения по рангу. В этом случае родительская лакуна определяется за $O(1)$ операций, а каждое ребро может поменять свою лакуну не более $O(\log n)$ раз, так как меньшая лакуналивается в большую. Таким образом, общие затраты на формирование лакун также составят $O(n(1 + \log n)) = O(n \log n)$ операций.

- 7.1. Однако построение триангуляции Делоне, требующее $O(n \log n)$ операций, понадобится для каждого значения аргумента.
8. Найти r -площадь как сумму площадей собственных областей и концевых секторов бициклов за вычетом суммарной площади линз. Из подраздела 2.4.1 известно, что число линз составляет $O(n)$, поэтому данный шаг потребует $O(n)$ операций.

Эти шаги обеспечивают решение задачи 2.1 — основной задачи главы. Проанализируем общую сложность алгоритма. Шаги 4–8 необходимо выполнить для каждого аргумента (за исключением тех случаев, когда используются инкрементные процедуры), число которых обозначим за k . Поэтому полные затраты составят (верхний индекс при O обозначает номера соответствующих шагов):

- $O^{(2)}(n \log n) + O^{(3)}(n) + O^{(4,5,6,8)}(kn) + O^{(7)}(n^2) = O((k+n)n)$ в случае инкрементного прослеживания скелетного графа;
- $O^{(2)}(n \log n) + O^{(3)}(n) + O^{(4,5,6,8)}(kn) + O^{(7)}(n \log n) + O^{(7.1)}(kn \log n) = O(kn \log n)$ в случае инкрементной обработки лакун.

Эти оценки позволяют заключить, что выбор конкретной процедуры зависит от величины k — числа различных радиусов, для которых мы считаем площадь покрытия. Если оно невелико, тогда квадратичная по n часть первой процедуры является неэффективной, и лучше воспользоваться реализацией через лакуны. Однако, когда k становится достаточно большим, эти затраты окупаются возможностью не строить на каждом шаге триангуляцию Делоне, и инкрементное прослеживание скелета оказывается предпочтительнее. Отметим также, что приведены оценки в худшем случае, который на практике может быть чрезвычайно редким, поэтому первая процедура может сработать достаточно быстро даже при небольших k .

2.6 Эксперименты с гибкими объектами

Для демонстрации устойчивости предложенного дескриптора к гибким деформациям была вычислена морфологическая ширина изображений из набора «Балерина», содержащего восемь рисунков, представляющих одну и ту же фигуру в разных позах (рис. 2.29). Эксперимент показал, что дескриптор изменяется достаточно слабо в случае отсутствия самоокклюзий. Небольшие вариации в результатах вызваны именно наложением элементов изображения

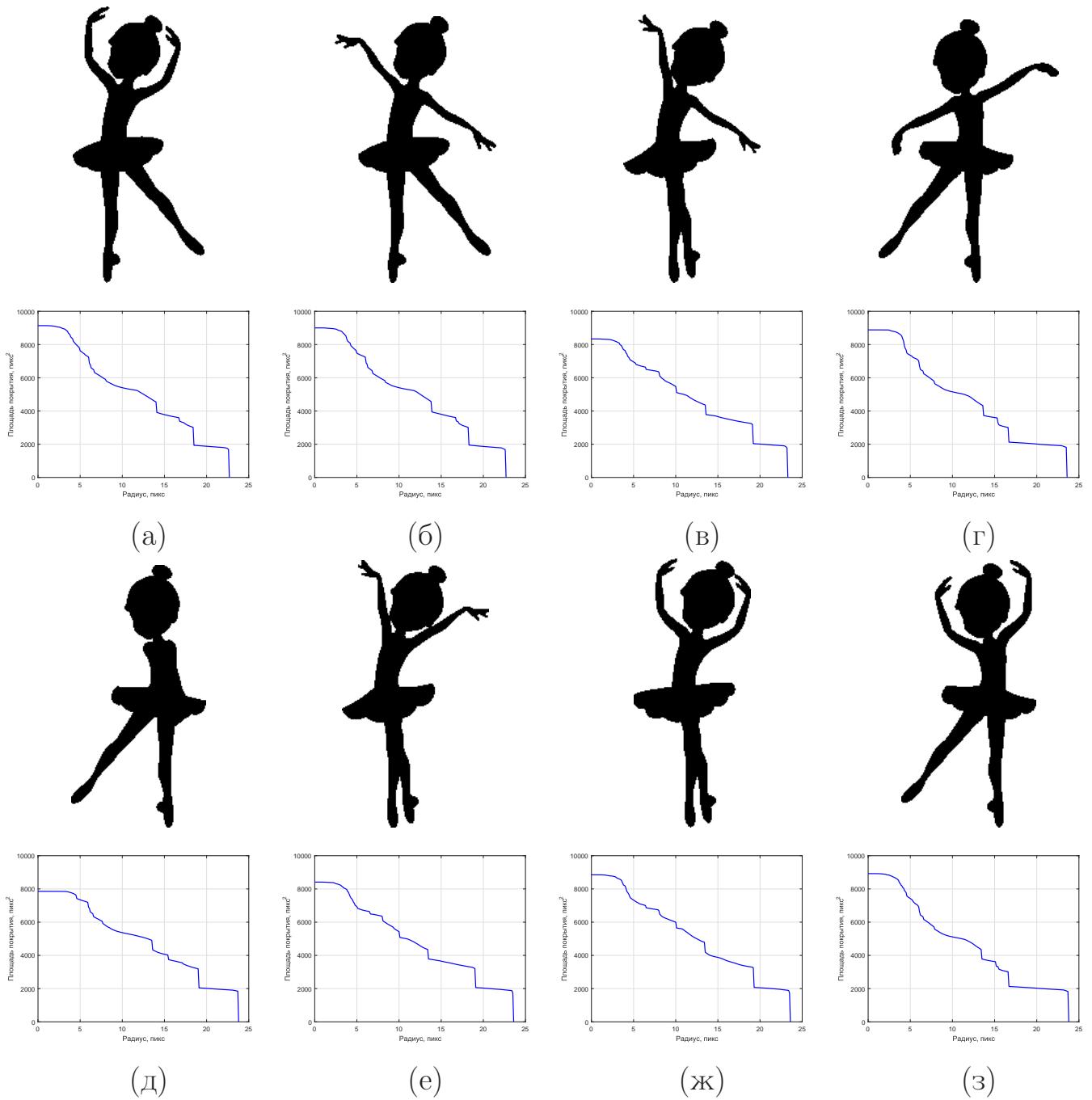


Рисунок 2.29 — Набор изображений «Балерина» (ряды 1 и 3) и их морфологическая ширина (ряды 2 и 4)

друг на друга (скрещивание ног или рук), а также изменением размеров балетной пачки (рис. 2.31).

Кроме того, с целью проверки гипотезы на реальных изображениях были вычислены диаграммы морфологической ширины для набора изображений «Лошадь», который представляет собой 395 кадров из анимированной последовательности (рис. 2.30). Для визуализации из этого набора с равномерным шагом были выбраны 8 изображений, также была вычислена средняя диаграм-

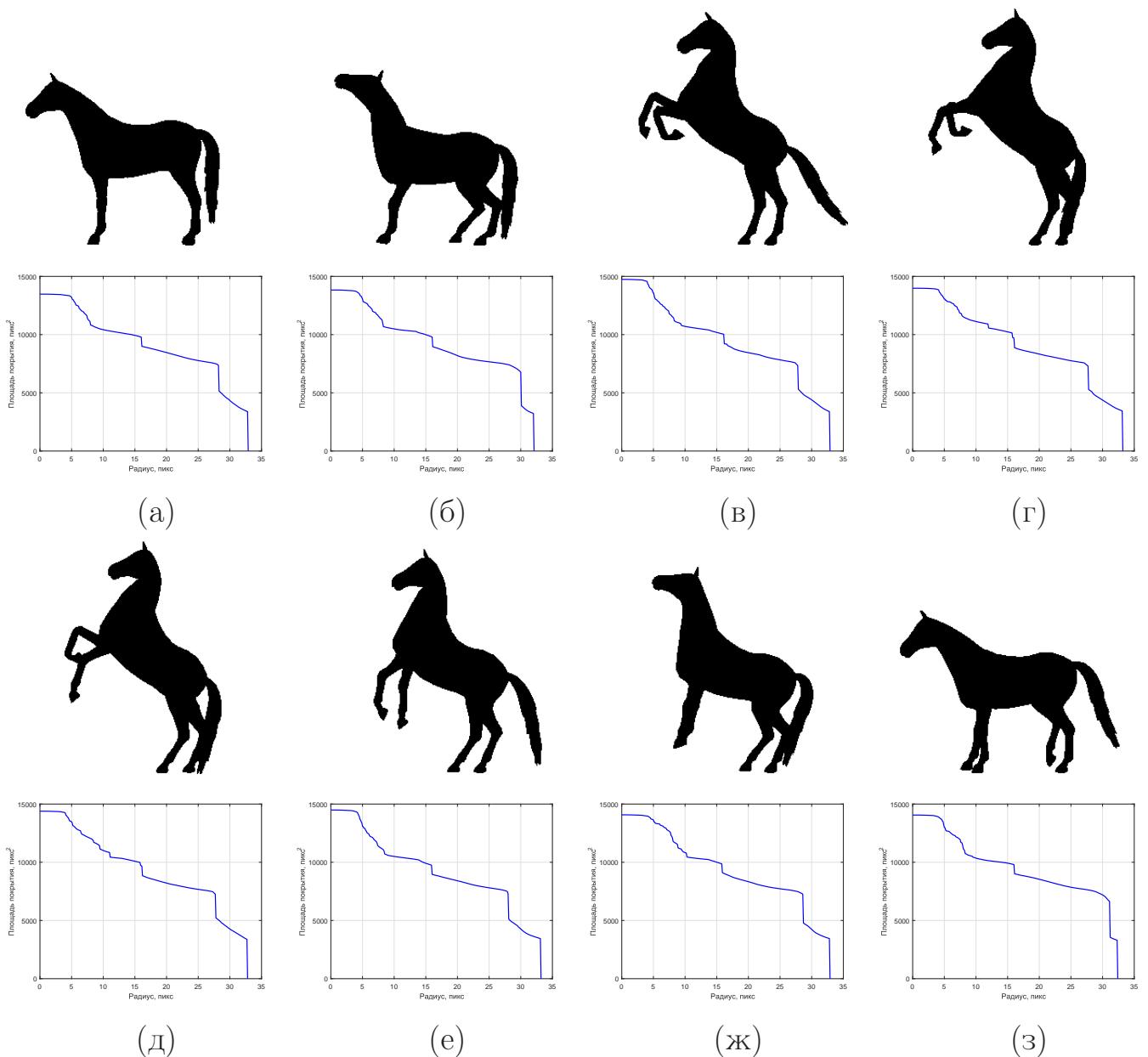


Рисунок 2.30 — Набор изображений «Лошадь» (ряды 1 и 3) и их морфологическая ширина (ряды 2 и 4)

ма, медианная диаграмма и диаграмма минимальных и максимальных значений по всем 395 кадрам, которые показаны на рис. 2.32б. Эксперимент снова показывает, что, несмотря на разнообразие поз, все результаты лежат в достаточно узком диапазоне, который позволяет выделить для объекта заданного класса узнаваемый признаковый профиль.

В завершение был проведён эксперимент по сравнению изображений из разных наборов и проверке гипотезы разделимости. Так как изображения были получены из разных источников и находятся в разных масштабах, возникла проблема нормировки функций ширины. Она была решена следующим обра-

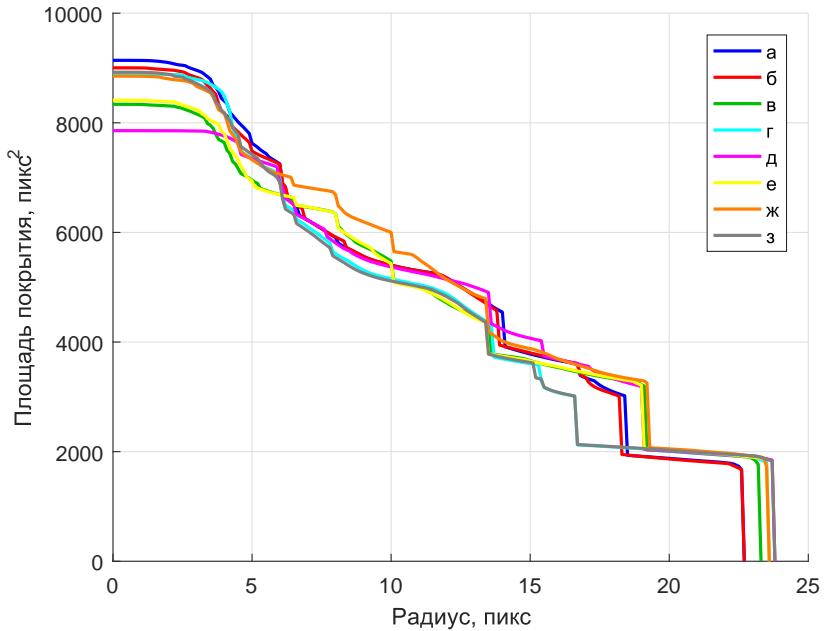


Рисунок 2.31 — Сводный график для набора изображений «Балерина»

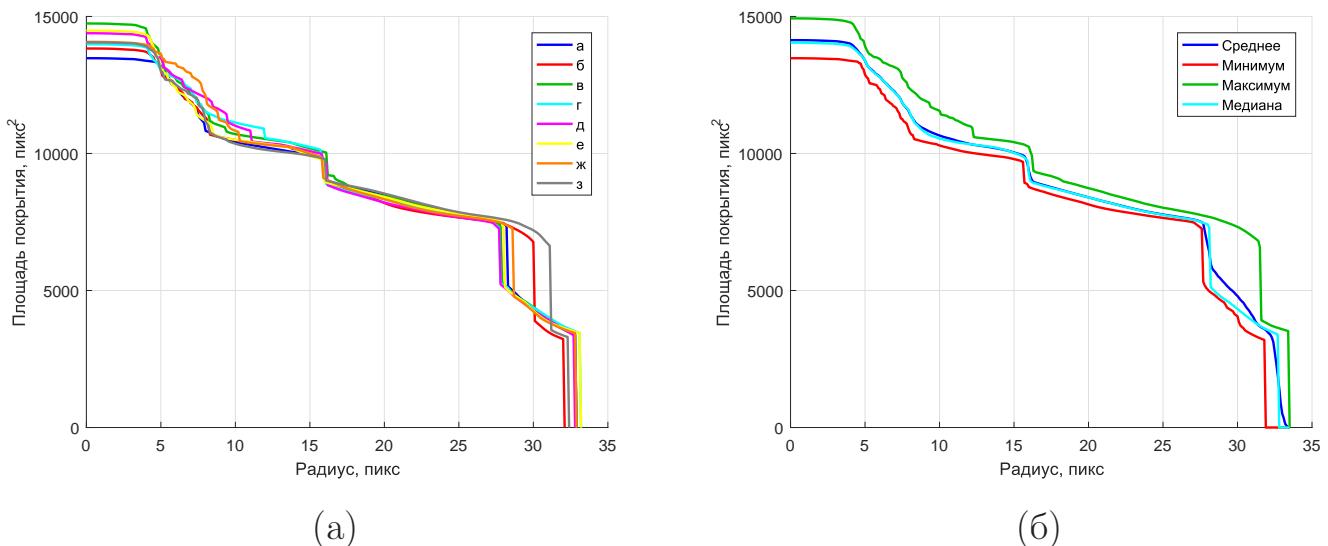


Рисунок 2.32 — Сводные графики для набора изображений «Лошадь»

зом: после нормировки площадь под графиком должна стать равной единице. Заметим, что если функция ширины фигуры X равна $\mathfrak{F}(r)$, то функция ширины фигуры X^* , полученной растяжением фигуры X в k раз, равна

$$\mathfrak{F}^*(r) = k^2 \mathfrak{F}\left(\frac{r}{k}\right).$$

Отметим также, что площадь под графиком при таком масштабировании увеличивается в k^3 раз (можно считать, что фигура, образованная графиком и осями, растягивается в k раз по оси x и в k^2 раз по оси y). Поэтому масштаб-

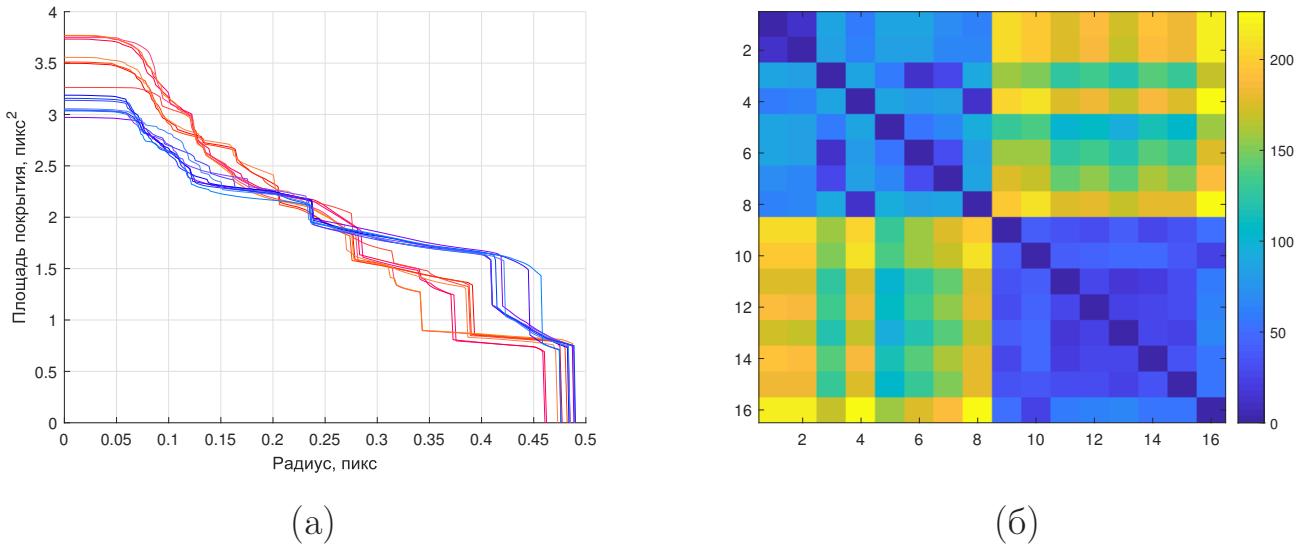


Рисунок 2.33 — Сравнение диаграмм ширины для изображений из наборов «Балерина» и «Лошадь»

ный коэффициент для требуемой нормировки вычисляется как

$$k = \left(\int_0^\infty \mathfrak{F}(r) dr \right)^{\frac{1}{3}}.$$

Нормализованные диаграммы для изображений из наборов «Балерина» и «Лошадь» показаны на рис. 2.33а оттенками красного и оттенками синего соответственно. Далее между диаграммами были посчитаны попарные расстояния в L_1 -метрике, они визуализированы на рис. 2.33б. В результате эксперимента объекты своего класса оказывались с точки зрения ширины ближе объектов чужого за единственным исключением — для изображения балерины с рис. 2.29д изображение лошади с рис. 2.30д оказалось ближе, чем изображение балерины с рис. 2.29з, что можно объяснить наличием сильных окклузий (руки на рис. 2.29д не видны). Впрочем, по рис. 2.33а нетрудно заметить, что нормализованные диаграммы классов разделимы, например, при значениях $r = 0$ и $r = 0,4$.

Итак, гипотеза разделимости в целом подтверждается, однако эксперимент также показывает, что задачи нормализации диаграмм ширины и разработки более мощных дескрипторов ширины являются актуальными.

2.7 Выводы к главе 2

- Предложен способ описания переменной ширины объекта в виде функции, отражающей зависимость площади вписанных в фигуру дисков от радиуса диска.

2. Для расчёта данного дескриптора введена модель дискового покрытия мноугольной фигуры, опирающаяся на медиальное представление формы — скелет и радиальную функцию. Проанализирована структура дискового покрытия, показано, что оно состоит из областей, соответствующих отдельным бициклам скелета, а оставшаяся часть покрывается секторами концевых кругов усечённых бициклов, которые могут пересекаться между собой. Введено понятие опорного пути, на основе которого определяются смежные бициклы, показано, что только пересечения смежных бициклов являются значимыми при расчёте площади покрытия.
3. Получены оценки числа максимального числа пар смежных бициклов по отношению как к числу рёбер, так и числу вершин скелета. В явном виде представлено строение фигур, на которых достигаются асимптотически оптимальные оценки.
4. Доказано, что имеют значение только попарные пересечения усечённых бициклов, даны оценки максимальной предельной доли смежных пар. В силу этого рассматривается понятие графа смежности бициклов и предложены две стратегии поддержания этого графа в актуальном состоянии: инкрементное прослеживание — локальное перестроение графа смежности с ростом радиуса диска, и обработка лакун — построение для каждой лакуны триангуляции Делоне множества её терминальных точек и поиск связной компоненты графа смежности как подграфа этой триангуляции.
5. Даны оценки вычислительной сложности обеих стратегий, которые свидетельствуют о том, что первая стратегия является асимптотически более эффективной по числу аргументов, для которых определяются значения функции, а вторая — по числу рёбер скелета.
6. Проведены эксперименты с изображениями гибких объектов, показывающие, что предложенный дескриптор ширины остаётся достаточно устойчивым к гибким деформациям в случае отсутствия у объекта самопересечений.

Глава 3. Производные дескрипторы ширины

В предыдущей главе было введено понятие морфологической ширины и описан алгоритм её вычисления для произвольной многоугольной фигуры. Разработанный алгоритм опирается на математическую модель области заданной ширины — дискового покрытия. Можно считать, что основным достижением предыдущей главы является именно всесторонний анализ этой модели. Сама же функция морфологической ширины является лишь одним из инструментов описания дискового покрытия в виде некоторого дескриптора — вектора чисел. В практических приложениях могут возникнуть задачи, требующие других дескрипторов дискового покрытия, либо вычисления морфологической ширины в нестандартных условиях. Можно выделить следующие типы таких задач.

В модели дискового покрытия все точки являются равноправными и вносят одинаковый вклад в расчёт площади. В некоторых практических задачах для фигуры может быть задана явно или выделена в процессе анализа формы область интереса, и описание ширины должно быть получено для части фигуры, лежащей внутри такой области.

Модель дискового покрытия строится для многоугольных фигур, которые служат представлением формы объектов на бинарных изображениях. При этом в ряде задач анализа формы в качестве исходных данных выступают полутонаовые изображения, для описания которых признаки ширины также могут оказаться очень полезными.

Морфологическая ширина не учитывает пространственную информацию, в частности, инвариантна к сдвигу, повороту и отражению как всей фигуры, так и отдельных её частей. В некоторых случаях характер задачи может потребовать учёта пространственной информации и чувствительности к подобным преобразованиям, так что морфологическая ширина сама по себе оказывается непригодной.

Наконец, многие методы распознавания и сравнения плоских форм используют характеристики формы — кривизну, расстояние, соответствие локальному шаблону — вычисленные для точек интереса, так как на базе таких оценок можно создавать весьма гибкие дескрипторы формы и применять к ним развитые техники машинного обучения. Одной из таких характеристик может выступать ширина, однако модель дискового покрытия оперирует на уровне

бициклов, а не точек, и не предоставляет информацию о ширине в отдельных точках напрямую.

Во всех этих случаях задача либо не сводится к вычислению площади покрытия, либо это вычисление не позволит добиться желаемого результата. Тем не менее вычисление морфологической ширины можно считать частным случаем более детальной постановки задачи — к примеру, можно считать, что это оценка ширины по области интереса, включающей фигуру целиком или вычисление ширины для объекта на полутонаром изображении, все пиксели которого принимают одно из двух крайних значений яркости. Таким образом, дескрипторы ширины для решения рассмотренных задач могут быть получены путём обобщения модели дискового покрытия и развития идеи морфологической ширины. Задача разработки таких дескрипторов будет исследована в данной главе.

3.1 Частная морфологическая ширина

В ряде задач анализа формы оценка ширины объекта совмещается с его сегментацией. К примеру, известны методы идентификации человека по форме кисти руки, основанные на признаках ширины [47; 62]. В рамках подобных подходов вся кисть рассматривается как единое целое, и вклад, который вносят в признаконое описание отдельные части объекта — пальцы, запястье и ладонь — усредняется, что приводит к потере локальной информации. Но более развитый способ идентификации [65] связан с сегментацией изображения и выделением на нём областей, соответствующих каждому пальцу. Поэтому можно ожидать от метода, вычисляющего тонкие признаки ширины — отдельно по сегментам изображения, более высокого качества работы. Такой дескриптор является более мощным, а вычисление признаков для полного объекта можно рассматривать как его частный случай.

Другим примером значимости частных признаков ширины может служить задача оценки морфогенеза плоских червей — планарий [66]. После поперечной перерезки планария способна восстановить свою форму и заново отрастить голову и хвостовую часть. При наличии качественных изображений планарии в период регенерации оценку степени регенерации можно рассматривать как задачу распознавания формы объекта на изображении. При этом особое внимание следует уделить областям, в которых должны располагаться выращенные заново части — голова или хвост, так что эти области можно рассматривать как области интереса. Как будет показано далее, эта задача может

быть успешно решена при помощи локальных признаков ширины — признаков, вычисленных для сегмента изображения, попадающего в область интереса.

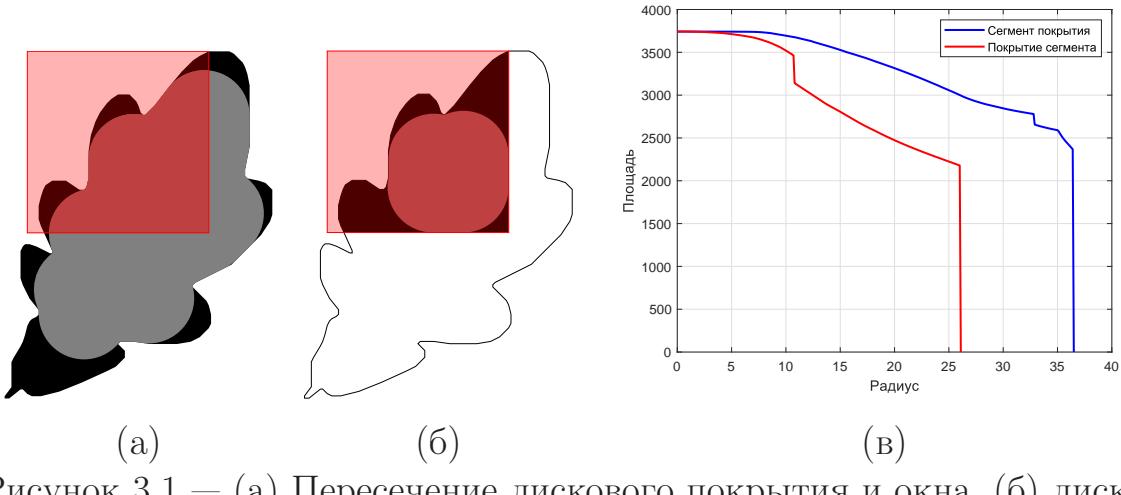


Рисунок 3.1 — (а) Пересечение дискового покрытия и окна, (б) дисковое покрытие пересечения фигуры и окна, (в) графики площадей сегментированных покрытий

В рассмотренных выше практических приложениях постановка задачи требует извлечения признаков для некоторой части изображения. Применив такую постановку к понятию морфологической ширины, мы столкнёмся с некоторыми трудностями. Пусть X — исходная многоугольная фигура, P — область интереса, которая представляет собой произвольный простой многоугольник (дальнейшие рассуждения легко обобщаются на случай многоугольника с дырами, а границы, представленные в другом виде, с достаточной точностью аппроксимируются многоугольниками). Необходимо получить распределение ширины той части фигуры X , которая попадает в окно P . Сделать это, однократно определив новую многоугольную фигуру (или набор многоугольных фигур) $X_P = X \cap P$, и рассчитав для её r -покрытия $X_P(r)$ площадь — не самая удачная идея, поскольку

$$\begin{aligned} X_P(r) &= \bigcup_{x \in X_P} \{B_x(r) \mid B_x(r) \subseteq X_P\} \subseteq \\ &\subseteq X(r) \cap P = \left(\bigcup_{x \in X} \{B_x(r) \mid B_x(r) \subseteq X\} \right) \cap P, \end{aligned}$$

где $B_x(r) = \{z \mid \|z - x\| \leq r\}$ — круг радиуса r с центром в точке x . Иначе говоря, r -покрытие фигуры X_P может оказаться меньше части r -покрытия фигуры X , попадающей в область P . При этом очевидно, что решение, считать ли

какую-либо область фигуры достаточно широкой, должно определяться лишь формой самой фигуры и не зависеть от конкретного вида области интереса. Поэтому вычисление дискового покрытия фигуры, предварительно пересечённой с окном, вообще говоря, занижает реальную оценку ширины сегмента фигуры. Особенно это характерно для областей около углов, возникающих при пересечении фигуры X с окном: в них ширина начинает считаться небольшой (области перестают попадать в покрытие), хотя в исходной фигуре X размещённый в этой области круг может оказаться достаточно большим. Ситуацию систематической недооценки ширины сегмента фигуры иллюстрирует рис. 3.1. Таким образом, возникает задача вычисления площади фигуры $X(r) \cap P$, то есть именно сегмента покрытия, а не покрытия сегмента.

Задача 3.1. *Дана многоугольная фигура X и многоугольная область интереса P . Для заданного набора радиусов $\{r_i\}_{i=1}^n$ вычислить площади фигур $X(r_i) \cap P$, то есть частей r -покрытия X , лежащих внутри P .*

Определение 3.1. *Функцию зависимости площади $X(r) \cap P$ от r будем обозначать $\mathfrak{F}_P(r)$ и называть частной морфологической шириной.*

В предыдущей главе дисковое покрытие $X(r)$ было разбито на области следующих типов:

1. Собственные области бициклов — полных и усечённых.
2. Внешние сектора меньших концевых кругов усечённых бициклов.
3. Линзы в пересечении внешних секторов.

Площадь дискового покрытия складывалась из суммы площадей областей первых двух типов за вычетом суммы площадей областей третьего типа. Для вычисления частной морфологической ширины можно оставить без изменений основной принцип вычисления, но рассматривать не просто площадь каждой области, а площадь области, пересечённой с окном. Заметим, что собственные области бициклов представляют собой простые многоугольники, внешние сектора — области, ограниченные двумя отрезками прямых и выпуклой круговой дугой, а линзы в их пересечении — области, ограниченные двумя выпуклыми круговыми дугами. Таким образом, границу любой интересующей нас области можно описать в виде последовательности участков двух типов — прямолинейных отрезков и круговых дуг. Более формально границу криволинейной области G , состоящей из подобных участков, можно описать в виде последовательности пар $\{g_i\}$, $i = 1, \dots, n$, где $g_i = (A_i(x_i, y_i), \theta_i)$. Первое значение в

паре задаёт начальную точку очередного участка, второе — угловую величину дуги, соединяющей текущую точку с последующей. При этом прямолинейный отрезок задаётся в общем виде как круговая дуга с угловой величиной, равной 0^1 . Если $\theta_i > 0$, будем считать центр окружности, содержащей дугу, расположенным слева от прямой, проходящей через точки (x_i, y_i) и (x_{i+1}, y_{i+1}) , если $\theta_i < 0$ — размещаем центр справа. Координаты центра (x_i^*, y_i^*) можно вычислить по этим значениям непосредственно:

$$\begin{aligned} x_i^* &= \frac{x_i + x_{i+1}}{2} - \frac{d}{2 \operatorname{tg} \frac{\theta_i}{2}} \frac{y_{i+1} - y_i}{d}, \\ y_i^* &= \frac{y_i + y_{i+1}}{2} + \frac{d}{2 \operatorname{tg} \frac{\theta_i}{2}} \frac{x_{i+1} - x_i}{d}, \end{aligned}$$

где $d = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$. Также будем считать, что $(x_{n+1}, y_{n+1}) = (x_1, y_1)$.

Попробуем описать в таком виде все необходимые области. Для этого там потребуется операция проецирования точки C на сайт-сегмент, концами которого являются точки A и B . Можно заметить, что $P_{AB}(C)$ — проекция точки на прямую, порождённую сайтом-сегментом, лежит на этом сайте, так что её можно считать и проекцией точки на сайт-сегмент:

$$P_{AB}(C) = A + t(B - A), \quad t = \frac{\langle \overrightarrow{AC}, \overrightarrow{AB} \rangle}{|\overrightarrow{AC}| \cdot |\overrightarrow{AB}|}.$$

Пусть A и B — концы оси линейного бицикла, а отрезки CD и EF — его сайты-сегменты. Тогда обход границы собственной области, опуская θ_i , так как все они равны нулю, можно произвести в следующем порядке (рис. 3.2а):

$$G = \{A, P_{CD}(A), P_{CD}(B), B, P_{EF}(B), P_{EF}(A)\},$$

или просто

$$G = \{A, P_{CD}(B), B, P_{EF}(B)\},$$

если A — терминальная точка скелета и изначально принадлежит образующим сайтам (рис. 3.2б):

Обход границы собственной области параболического бицикла с сайтом-точкой C и сайтом-сегментом DE принимает вид (рис. 3.2в):

$$G = \{A, P_{DE}(A), P_{DE}(B), B, C\}.$$

¹Можно считать прямолинейный отрезок частью дуги окружности с бесконечным радиусом.

Наконец, при обходе границы собственной области гиперболического бицикла с сайтами-точками C и D получаем (рис. 3.2г):

$$G = \{A, C, B, D\}.$$

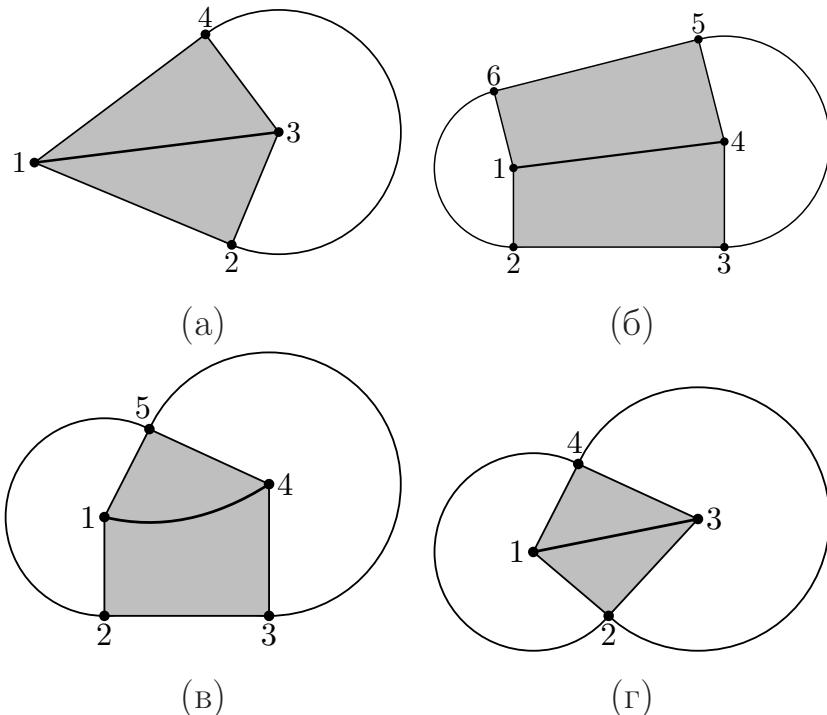


Рисунок 3.2 — Обход границ собственных областей (а)-(б) линейного, (в) параболического и (г) гиперболического бициклов

Как правило, в математических рассуждениях рассматриваются положительно ориентированные контуры — считается, что при обходе контура интересующая нас область остаётся слева, то есть контур обходится в направлении против часовой стрелки. Так как вершины ребра и образующие сайты изначально неупорядочены, может сложиться ситуация, когда точки в списке G записаны в порядке обхода по часовой стрелке. Выявить такую ситуацию можно следующим образом: известно, что площадь многоугольника G с n вершинами равна

$$S(G) = \frac{1}{2} \left| \sum_{i=1}^n (x_{i+1} - x_i)(y_i + y_{i+1}) \right|.$$

Фактически предыдущая формула является частным случаем формулы Грина в применении к многоугольнику. Положительное значение выражения под знаком модуля свидетельствует о том, что обход производится по часовой стрелке. В этом случае нужно инвертировать порядок элементов в списке G .

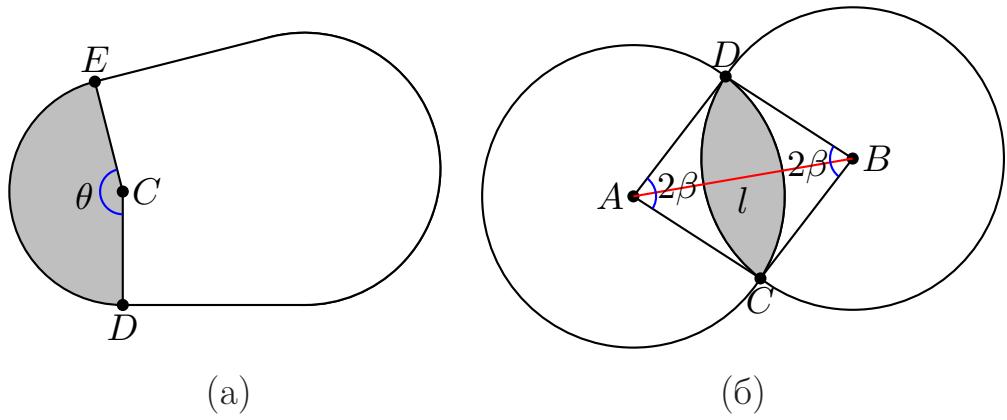


Рисунок 3.3 — Обход границы сектора концевого круга (а) и линзы в пересечении концевых кругов (б)

Теперь рассмотрим внешний сектор концевого круга усечённого бицикла. Центром этого круга является точка C — крайняя точка, полученная в результате процедуры сечения бицикла при построении r -покрытия. О том, как определить её точные координаты, можно узнать в разделе 2.3. Пусть также точки D и E — концы круговой дуги, которые можно найти как проекции точки C на пару образующих сайтов бицикла (в случае сайта-точки под проекцией понимается сама эта точка), причём $\{C,D,E\}$ задают порядок обхода против часовой стрелки (рис. 3.3а). Тогда список G сформируем в виде:

$$G = \{(C, 0), (D, \theta), (E, 0)\},$$

где θ — угловая величина кругового сектора, вычисляемая по формуле

$$\theta = 2 \arcsin \frac{|\overline{DE}|}{2r},$$

а r — радиус покрытия.

Отметим также, что возможна ситуация, когда точки C , D и E лежат на одной прямой, и направление обхода определить нельзя. В этом случае направление обхода можно установить по тройке $\{C', D, E\}$, где C' — другой конец оси бицикла.

Наконец, рассмотрим линзу, образованную пересечением кругов радиуса r с центрами в точках $A(x_1, y_1)$ и $B(x_2, y_2)$, находящихся друг от друга на расстоянии d , $0 < d < 2r$. Точка B расположена относительно точки A в направлении, заданном углом $\alpha = \text{atan}2(y_2 - y_1, x_2 - x_1)$ ². Угловая величина дуги линзы

²Функция $\text{atan}2(y, x) : \mathbb{R}^2 \rightarrow (-\pi, \pi]$ определяет угол в радианах между положительной x -полуосью и лучом из начала координат до точки $(x, y) \neq (0, 0)$. Обозначение взято из названия этой функции во многих языках программирования.

равна $2\beta = 2 \arccos\left(\frac{d}{2r}\right)$, следовательно, точки пересечения $C(x_3, y_3)$ и $D(x_4, y_4)$ концевых кругов имеют координаты:

$$(x_3, y_3) = (x_1 + r \cos(\alpha - \beta), y_1 + r \sin(\alpha - \beta)),$$

$$(x_4, y_4) = (x_1 + r \cos(\alpha + \beta), y_1 + r \sin(\alpha + \beta)).$$

Тогда линзу, показанную на рис. 3.3б, можно описать в виде $G = \{(C, 2\beta), (D, 2\beta)\}$.

Теперь все необходимые области описаны в стандартном формате, и мы можем обратиться к задаче вычисления площади их пересечения с областью интереса. Для нахождения области пересечения отсекаемой криволинейной области G по отсекающему многоугольнику P можно воспользоваться модификацией алгоритма Уайлера – Атертона [67], общий вид которого представлен как алгоритм 2.

Таким же образом можно найти и объединение областей, в этом случае обход нужно начинать с выходных точек. Собственные области, круговые сектора и линзы представляют собой выпуклые области, поэтому в случае выпуклого окна P область пересечения будет только одна.

Единственное существенное отличие от оригинального алгоритма Уайлера – Атертона заключается в том, что прямолинейный отрезок и дуга окружности могут иметь две точки пересечения. Пусть концы отрезка заданы точками (x_1, y_1) и (x_2, y_2) , а дуга является частью окружности радиуса r с центром в точке (x_3, y_3) , и полярный угол точек дуги изменяется в пределах от α до β , $\alpha < \beta$, $\beta - \alpha < 2\pi$. Тогда точки пересечения прямой, на которой лежит отрезок, и окружности можно получить из системы уравнений:

$$\begin{cases} x = x_1 + t(x_2 - x_1), \\ y = y_1 + t(y_2 - y_1), \\ (x - x_3)^2 + (y - y_3)^2 = r^2. \end{cases}$$

Подставляя первые два уравнения в третье, получаем:

$$(x_1 + t(x_2 - x_1) - x_3)^2 + (y_1 + t(y_2 - y_1) - y_3)^2 = r^2,$$

$$\begin{aligned} t^2(x_2 - x_1)^2 + 2t(x_2 - x_1)(x_1 - x_3) + (x_1 - x_3)^2 + \\ + t^2(y_2 - y_1)^2 + 2t(y_2 - y_1)(y_1 - y_3) + (y_1 - y_3)^2 = r^2, \end{aligned}$$

Алгоритм 2 Алгоритм Уайлера – Атертона

Дано: Криволинейная область G , многоугольная фигура P

Найти: Множество криволинейных фигур $\{G_i\}$ в пересечении G и P

- 1: Составить два списка из участков границы G и P
 - 2: Пометить вершины в каждом из списков в соответствии с тем, находятся ли они внутри другой фигуры или нет
 - 3: Добавить в оба списка точки пересечения границ областей; установить между совпадающими точками в разных списках двусторонние связи
 - 4: **Если** ни одного пересечения не найдено **тогда**
 - 5: **Если** G внутри P **тогда**
 - 6: вернуть G
 - 7: **иначе если** P внутри G
 - 8: вернуть P
 - 9: **иначе если** G и P не пересекаются
 - 10: вернуть пустое множество
 - 11: **Конец условия**
 - 12: **иначе**
 - 13: Составить список точек пересечения, в которых граница фигуры G при обходе входит в многоугольник P
 - 14: Следуя из каждой такой точки по часовой стрелке вдоль границ обеих областей G и P , найти множество областей пересечения
 - 15: **Конец условия**
-

$$t^2 \left((x_2 - x_1)^2 + (y_2 - y_1)^2 \right) + 2t \left((x_2 - x_1)(x_1 - x_3) + (y_2 - y_1)(y_1 - y_3) \right) + (x_1 - x_3)^2 + (y_1 - y_3)^2 - r^2 = 0.$$

В зависимости от знака дискриминанта у уравнения может быть два корня, один, или не быть корней вовсе. В первом случае прямая пересекает окружность, во втором является касательной к ней, в третьем — проходит в стороне от окружности. Далее все имеющиеся корни t_i нужно проверить на принадлежность заданных ими точек отрезку и дуге. Для принадлежности отрезку нужно, чтобы $0 \leq t_i \leq 1$, для принадлежности дуге — чтобы $\alpha \leq \gamma \leq \beta$, где $\gamma = \text{atan2}(y_1 + t_i(y_2 - y_1), x_1 + t_i(x_2 - x_1)) + 2\pi k$, а k выбрано так, чтобы $\alpha \leq \gamma$ и $\alpha > \gamma - 2\pi k$. Так же, как и на этапе проверки, снаружи или внутри фигуры находится точка, требуется особая обработка пересечений, расположенных на

конце сегмента (прямолинейного или кругового). Принцип работы алгоритма Уайлера – Атертона для фигур рассматриваемого типа показан на рис. 3.4.

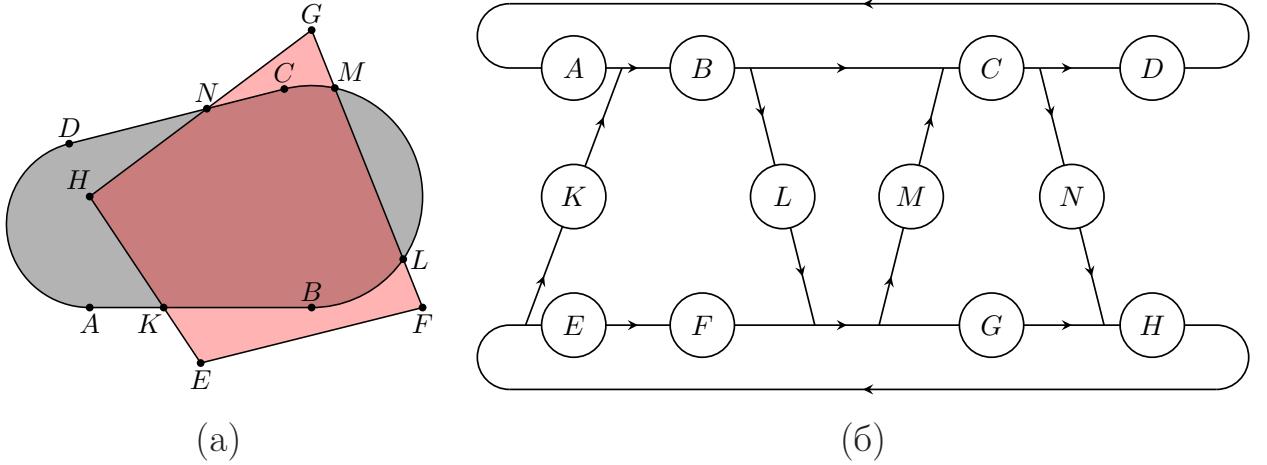


Рисунок 3.4 – Структуры данных для алгоритма Уайлера –Атертона с криволинейными фигурами. (а) Пересекаемые фигуры, (б) списки их вершин (верхний и нижний ряд) и список вершин пересечения (средний ряд) с двусторонними связями. Обход области пересечения можно начинать из любой точки среднего ряда в направлении, обозначенном стрелками.

Очевидно, что все участки границ фигур, которые получаются в пересечении, являются частями участков границ исходных фигур, поэтому новые фигуры можно описать в том же виде, что и исходные: с помощью пар вида (A_i, θ_i) . Хотя алгоритм пересечения применим для любых фигур с прямолинейными и круговыми участками границы, можно заметить, что при анализе дискового покрытия подобные фигуры имеют особенность: круговые участки границы являются выпуклыми, т.е. при обходе дуги против часовой стрелки область, принадлежащая фигуре, остаётся слева. Эта особенность пригодится нам при доказательстве следующей теоремы.

Теорема 3.1. Пусть G – замкнутая область без самопересечений, граница которой состоит из прямолинейных отрезков и выпуклых дуг окружностей и задана в порядке обхода против часовой стрелки парами вида $(A_i(x_i, y_i), \theta_i)$, $i = 1, \dots, n$. Тогда площадь области можно найти по формуле:

$$S(G) = \frac{1}{2} \sum_{i=1}^n ((x_i - x_{i+1})(y_{i+1} + y_i) + R_i^2(\theta_i - \sin(\theta_i))), \quad (3.1)$$

где

$$R_i = \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{2 \sin \frac{\theta_i}{2}}.$$

В случае $\theta_i = 0$ доопределим слагаемое с участием θ_i по непрерывности и будем считать его равным нулю.

Доказательство. Воспользуемся теоремой Грина. Для положительно ориентированной кусочно-гладкой замкнутой кривой на плоскости C , ограничивающей область D , и функций $P(x, y)$ и $Q(x, y)$, определенных в области D и имеющих непрерывные частные производные, справедлива формула

$$\oint_C P dx + Q dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy.$$

Будем считать, что $P = -y$, $Q = 0$, следовательно $\frac{\partial P}{\partial y} = -1$, $\frac{\partial Q}{\partial x} = 0$.

Получаем:

$$\oint_C -y dx = \iint_D 1 dx dy.$$

В формуле Грина предполагается, что C — положительно ориентированный контур, т.е. он обходится против часовой стрелки. Пусть этот контур разбит на участки $\{C_i\}_{i=1}^n$, соответствующие отрезкам и дугам. Тогда

$$\iint_D 1 dx dy = \oint_C -y dx = \sum_{i=1}^n \int_{C_i} -y dx. \quad (3.2)$$

Рассмотрим по отдельности каждое из слагаемых в сумме справа. Не ограничивая общности, можно считать, что фигура G полностью расположена не ниже оси OX , так как формула 3.1 инвариантна к параллельному переносу. Если слагаемое соответствует прямолинейному отрезку с концами (x_i, y_i) и

(x_{i+1}, y_{i+1}) , оно принимает вид:

$$\begin{aligned}
 \int_{C_i} -y \, dx &= - \int_{x_i}^{x_{i+1}} \left(y_i + \frac{x - x_i}{x_{i+1} - x_i} (y_{i+1} - y_i) \right) dx = \\
 &= \int_{x_i}^{x_{i+1}} \left(x \frac{y_{i+1} - y_i}{x_{i+1} - x_i} + y_i - x_i \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right) dx = \\
 &= \frac{x^2}{2} \cdot \frac{y_{i+1} - y_i}{x_{i+1} - x_i} + x \left(y_i - x_i \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right) \Big|_{x_{i+1}}^{x_i} = \\
 &= \frac{x_i^2 - x_{i+1}^2}{2} \cdot \frac{y_{i+1} - y_i}{x_{i+1} - x_i} + (x_i - x_{i+1}) \left(y_i - x_i \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right) = \\
 &= -\frac{(x_{i+1} + x_i)}{2} (y_{i+1} - y_i) + (x_i - x_{i+1}) y_i + x_i (y_{i+1} - y_i) = \\
 &= \frac{1}{2} (-x_{i+1} y_{i+1} + x_{i+1} y_i - x_i y_{i+1} + x_i y_i + 2x_i y_i - 2x_{i+1} y_i + 2x_i y_{i+1} - 2x_i y_i) = \\
 &= \frac{1}{2} (x_i - x_{i+1})(y_{i+1} + y_i).
 \end{aligned} \tag{3.3}$$

Этот результат легко получить и из геометрических соображений: с точностью до знака он равен площади под графиком линейной функции $y = y(x)$, то есть площади прямоугольной трапеции с основанием $|x_i - x_{i+1}|$ и боковыми сторонами y_i и y_{i+1} . Отметим, что этот результат равен i -му компоненту суммы 3.2, так как совпадает с его левой частью, а правая равна нулю.

Если C_i является дугой окружности, $\int_{C_i} y \, dx$ также с точностью до знака равен площади под графиком функции $y = y(x)$. Рассмотрим ситуацию, когда дуга является выпуклой, а x начинает убывать при движении по участку контура C_i . Возможны три варианта:

- x убывает на всём участке;
- x сначала убывает, а потом возрастает, и $x_{i+1} \leq x_i$;
- x сначала убывает, а потом возрастает, и $x_{i+1} > x_i$.

Все эти варианты показаны на рис. 3.5.

Покажем, что при любом из них интегрирование по участку исключает соответствующий компонент из суммы в утверждении теоремы.

В первом случае (рис. 3.5а) интеграл равен площади криволинейной трапеции $ABCDE$, которая складывается из площади трапеции $ACDE$ и кругового сегмента ABC . Так как $x_{i+1} < x_i$, площадь трапеции равна $\frac{1}{2}(x_i - x_{i+1})(y_{i+1} - y_i)$,

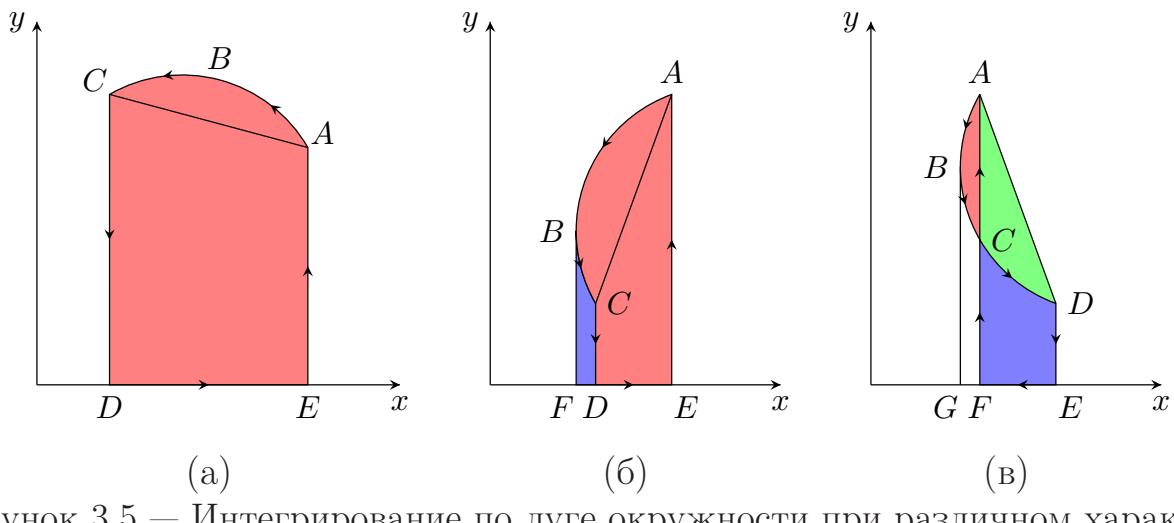


Рисунок 3.5 – Интегрирование по дуге окружности при различном характере изменения координаты x : (а) всё время убывает, (б) возрастает в конце участка, но не становится больше начального значения, (в) возрастает в конце участка, становясь больше начального значения

а радиус дуги сегмента R_i получается из рассмотрения равнобедренного треугольника ACP , где P – центр окружности:

$$\begin{aligned} \sin \frac{\angle APC}{2} &= \frac{AC}{2AP}, \\ \sin \frac{\theta_i}{2} &= \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{2R_i}, \\ R_i &= \frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{2 \sin \frac{\theta_i}{2}}. \end{aligned}$$

Площадь сегмента составляет $\frac{R_i^2}{2}(\theta_i - \sin \theta_i)$, так что случай соответствует утверждению теоремы.

Во втором случае (рис. 3.5б) интеграл равен разности площадей криволинейных трапеций $ABFE$ и $BFDC$, то есть площади фигуры с криволинейной границей $ABCDE$. Но она снова складывается из из площади трапеции $ACDE$ и кругового сегмента ABC .

Наконец, в третьем случае (рис. 3.5в) интеграл равен разности площадей криволинейных трапеций $ABGF$ и $BGEDC$. Так как трапеция $BGEDC$ складывается из двух трапеций – $BGFC$ и $CFED$, а $BGFC$ вложена в $ABGF$, получаем, что интеграл равен разности площадей сегмента ABC (красная область) и криволинейной трапеции $CFED$ (синяя область). Рассмотрим также криволинейный треугольник ACD (зелёная область), круговой сегмент $ABCD$ (объединение красной и зелёной областей) и трапецию $AFED$ (объединение

синей и зелёной). Можно заметить, что

$$S(ABC) + S(AFED) = S(ABCD) + S(CFED),$$

то есть

$$S(ABC) - S(CFED) = S(ABCD) - S(AFED).$$

Но здесь $-S(AFED) = \frac{1}{2}(x_i - x_{i+1})(y_{i+1} - y_i)$, так как $x_{i+1} > x_i$, поэтому случай также исключает пару слагаемых с индексами i в исходной формуле. Случаи, в которых x возрастает в начале движения по сегменту, а также интегрирование по вогнутой дуге рассматриваются аналогично.

В итоге мы получили взаимное соответствие между слагаемыми в формуле Грина и слагаемыми в формулировке теоремы, что подтверждает её справедливость. ■

Общую схему вычисления морфологической ширины с параметром r , описанную в главе 2.1, можно без изменений применить к частной ширине:

1. Найти все полные и усечённые бициклы.
2. Найти все пары смежных бициклов — либо с помощью инкрементного прослеживания, либо путём поиска лакун и построения триангуляции Делоне для множества крайних точек каждой лакуны.
3. Вычислить морфологическую ширину по сегменту (окну) как сумму площадей собственных областей полных и усечённых бициклов, а также внешних секторов усечённых бициклов, за вычетом суммы площадей линз в пересечении смежных бициклов. Для каждой области, площадь которой вычисляется, необходимо предварительно определить её пересечение с окном.

Процедура вычисления пересечения областей является достаточно сложной и вычислительно затратной, поэтому логично минимизировать случаи её вызова. Например, можно заметить, что если полная собственная область бицикла полностью попадает в окно, то и любая усечённая область того же бицикла также полностью попадает в окно. В целом у рассматриваемых областей можно выделить следующие свойства:

- для собственной области выделенного бицикла в r -покрытии $D_i(r)$ имеем $D_i(r_2) \subseteq D_i(r_1)$, если $r_1 \leq r_2$;
- для линзы в пересечении внешних секторов концевых кругов бициклов $L_{ij}(r)$ имеем $L_{ij}(r_2) \subseteq L_{ij}(r_1)$, если $r_1 \leq r_2$;

- для внешнего сектора меньшего концевого круга $S_i(r)$, вообще говоря, $S_i(r_2) \not\subseteq S_i(r_1)$ и $S_i(r_1) \not\subseteq S_i(r_2)$, если $r_1 \leq r_2$;
- также для выбранного бицикла имеем $D_i(r_2) \cup S_i(r_2) \subseteq D_i(r_1) \cup S_i(r_1)$, если $r_1 \leq r_2$. В случае, если r_i меньше радиуса меньшего концевого круга r , условимся считать, что $S(r_i) = S(r)$.

Это означает, что при вычислении морфологической ширины с возрастающим параметром r собственная область или линза либо изначально полностью попадает в окно, либо изначально полностью не попадает, либо сначала попадает частично, а затем, при достижении некоторого r , начинает попадать полностью. Поэтому предлагается поддерживать три списка областей: с полным пересечением, частичным и без пересечения, и по ходу вычисления при необходимости перебрасывать области из второго списка в первый. Вычислять пересечения с окном необходимо только для элементов второго списка, а вычислять площади — для элементов первых двух. Для секторов концевых кругов не выполняется свойство вложенности при возрастающем r , однако решение о включении сектора в один из списков можно принять по области вида $D_i(r) \cup S_i(r)$: если она попадает в окно полностью, сектор заносится в первый список, в если частично — во второй, также с возможностью затем перейти в первый.

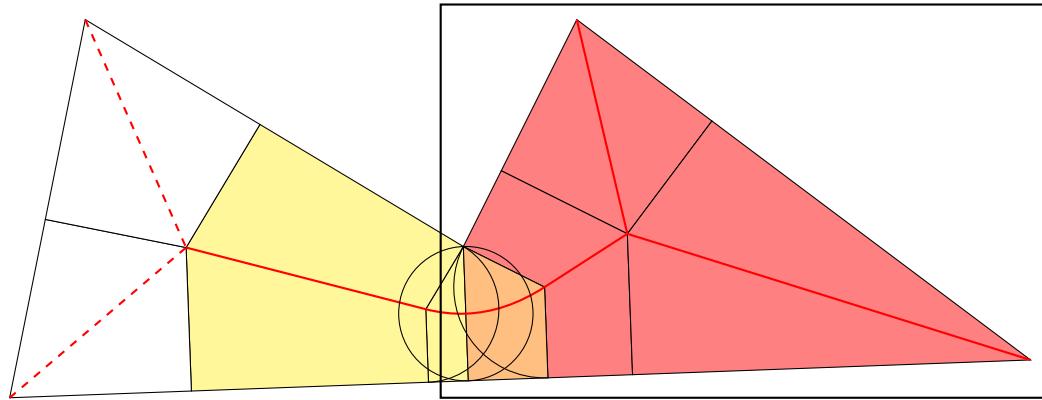


Рисунок 3.6 — Типы областей при вычислении частичного покрытия.

Собственные области для которых $(D_i(0) \cup S_i(0)) \setminus P = \emptyset$ отмечены красным, для которых $(D_i(0) \cup S_i(0)) \setminus P \subseteq S_i(0)$ — оранжевым, для которых $\emptyset \neq (D_i(0) \cup S_i(0)) \setminus P \not\subseteq S_i(0)$ — жёлтым, и для которых $(B_i(0) \cup S_i(0)) \cap P = \emptyset$ — белым. Рёбра редуцированного скелета показаны сплошной красной линией, отсечённые — пунктирной.

Так как поиск смежных бициклов также является достаточно сложным этапом, возникает вопрос: нельзя ли изначально редуцировать сам скелет,

чтобы оставить в нём только те элементы, которые могут внести вклад в вычисление площади? Следующая теорема даёт на него положительный ответ.

Теорема 3.2. *Алгоритм вычисления частной морфологической ширины останется корректным, если оставить в скелете только те рёбра, для которых области $D_i(0) \cup S_i(0)$ имеют непустое пересечение с окном.*

Доказательство. В результате редуции скелета мы не потеряем ни одну пересекающуюся с окном собственную область, сектор концевого круга или линзу. С другой стороны, и после процедуры упрощения скелета сектора $S_i(r)$ и $S_j(r)$, относящиеся к смежным бициклам и образующие линзу, будут принадлежать одной и той же лакуне r -покрытия, так как все промежуточные собственные области также будут задевать окно и, следовательно, их рёбра останутся в скелете.

Пусть в результате работы процедуры обрезки скелета \mathcal{S} по окну был получен скелет \mathcal{S}' . Очевидно, что каждая лакуна L'_i скелета \mathcal{S}' является подмножеством некоторой лакуны L_j скелета \mathcal{S} , а для множеств их крайних точек выполнено $P(L'_i) \subseteq P(L_j)$. Если точки $p_1, p_2 \in P(L'_i)$ связаны ребром в $DT(P(L_j))$, то они должны быть связаны ребром и в $DT(P(L'_i))$. Таким образом, проблема потери линз не возникает, но не исключена проблема появления лишних линз из-за рёбер в $DT(P(L'_i))$, которые отсутствуют в $DT(P(L_j))$. Согласно теореме 2.3, точки p_1 и p_2 , не связанные ребром в $DT(P(L_j))$, относятся к несмежным бициклам и могут образовывать линзу только в том случае, если эта линза вложена в другую линзу — образованную точками p'_1 и p'_2 , лежащими на опорном пути от p_1 к p_2 . Но это означает, что точки p'_1 и p'_2 также должны принадлежать множеству $P(L'_i)$, и линза, образованная точками p_1 и p_2 будет отфильтрована при постобработке триангуляции. Теорема доказана. ■

Итак, процедура упрощения скелета перед вычислением морфологической ширины фигуры по сегменту не приводит к потере имеющихся слагаемых или возникновению лишних, поэтому основной алгоритм вычисления остаётся корректным. Модификация алгоритма по правилам, описанным выше, даст нам решение задачи 3.1.

В завершение раздела рассмотрим аспект вычислительной сложности алгоритма. Согласно результатам, полученным в главе 2.1, исходный алгоритм вычисления площади дискового покрытия имеет следующую сложность по числу сегментов границы n и числу значений радиуса k :

- $O((k + n)n)$ в случае инкрементного прослеживания скелетного графа;
- $O(kn \log n)$ в случае инкрементной обработки лакун.

При этом непосредственное вычисление площадей областей, составляющих покрытие, в обоих случаях занимает $O(kn)$ операций. Можно заметить, что для криволинейной фигуры G с границей из t участков и области интереса P , представляющей собой многоугольник с s вершинами, алгоритм Уайлера – Атертона [2](#) и последующее вычисление области пересечения отработает за время $O(ts)$, так как между границами фигуры может быть обнаружено не более $2ts$ точек пересечения. Но фигура G представляет собой собственную область, круговой сектор или линзу, так что её граница состоит из не более чем шести участков. Поэтому сложность расчёта площадей можно считать зависящей только от сложности окна P и равной $O(kns)$, что даст нам оценки:

- $O(n^2 + kns)$ в случае инкрементного прослеживания скелетного графа;
- $O(kn(\log n + s))$ в случае инкрементной обработки лакун.

3.2 Морфологическая ширина полутонового изображения

Бинарное изображение является очень наглядным и эффективным способом представления формы фиксируемых визуально объектов. Тем не менее в задачах распознавания изображений бинарные изображения редко возникают естественным образом. Как правило, они являются результатом отдельной процедуры получения бинарного изображения из полутонового или полноцветного. Эта процедура слабо поддаётся универсализации, так как во многом обусловлена особенностями задачи. Можно выделить два случая, когда получение бинарного изображения из исходного нецелесообразно:

1. Процедура бинаризации не реализована или не даёт удовлетворительные результаты из-за сложности исходных данных.
2. Процедура бинаризации неуместна, так как на изображении нет объектов как таковых. Такая ситуация возникает при анализе текстур, рисунков, имеющих значимые перепады интенсивности, — в частности, такие данные распространены в анализе медицинских изображений. Бинаризация не просто не даст нужных очертаний, но уничтожит значительную часть полезной информации, связанную с перепадом яркости.

Между тем в подобных задачах также может возникнуть необходимость вычисления для содержимого изображения форморазмерных признаков, в том

числе признаков ширины. В бинарном случае этому сильно способствовал аппарат математической морфологии, описанный в разделе 1.1. Аналоги базовых морфологических операций — эрозии, дилатации, открытия и закрытия — были сформулированы для случая полутонового изображения и полутонового структурирующего элемента, основы полутоновой математической морфологии описаны в работе [68]. На основе базовых операций формулируется и последовательность морфологических фильтров, задающих полутоновые гранулометрии, эффективный алгоритм для расчёта которых был предложен в работе [69]. Полутоновые гранулометрии нашли применение в широком ряде задач анализа изображений, возникающих в медицине, материаловедении и фитоморфологии. К примеру, этот метод использовался в задачах анализа текстур [70], задаче контроля размера гранул корма для рыб [71] и компьютерной морфометрии синаптических бутонов [72].

Держа в уме тесную связь дискового покрытия и морфологической операции открытия, задающей гранулометрию, было бы полезным обобщить понятие морфологической ширины на случай полутонового изображения и предложить модель для её эффективного вычисления.

Кратко изложим основы полутоновой математической морфологии. В её рамках растровое изображение $I \in \mathbb{R}^{W \times H}$ можно считать некоторым измерением в точках целочисленной сетки вещественнозначного скалярного поля $f : C \rightarrow \mathbb{R}$, заданного на прямоугольном участке плоскости $C = \{(x, y) \mid 0 \leq x \leq W, 0 \leq y \leq H\}$. Будем называть такое скалярное поле f *векторным полутоновым изображением*. Также в виде скалярного поля представляется структурирующий элемент $g : D \rightarrow \mathbb{R}, D \subseteq \mathbb{E}^2$.

Определение 3.2. Эрозия векторного полутонового изображения f структурирующим элементом g задаётся как

$$(f \oplus g)(z) = \sup_{t \in C, z-t \in D} [f(t) + g(z-t)].$$

Определение 3.3. Дилатация векторного полутонового изображения f структурирующим элементом g задаётся как

$$(f \ominus g)(z) = \inf_{t \in C, z-t \in D} [f(t) - g(z-t)].$$

Также рассматриваются операции открытия со структурирующим элементом, заданным в виде множества точек евклидовой плоскости: $B \subset \mathbb{E}^2$:

$$\begin{aligned}(f \oplus B)(z) &= \sup_{z+t \in C, t \in B} f(z+t), \\ (f \ominus B)(z) &= \inf_{z+t \in C, t \in B} f(z+t).\end{aligned}\tag{3.4}$$

Можно заметить, что эти операции эквивалентны операциям с полями, если заменить элемент B полем $g(x,y) = 0$, $g : \hat{B} \rightarrow \mathbb{R}$, $\hat{B} = \{(x,y) : (-x, -y) \in B\}$.

Бинарные эрозию и дилатацию можно представить в виде, аналогичном уравнениям 3.4, с первым аргументом в виде скалярного поля. Для этого необходимо воспользоваться аппаратом характеристических функций. Пусть $\chi_A(z)$ — функция, принимающая на всей числовой плоскости значение 1, если точка z принадлежит множеству A , и значение 0 в противном случае. Тогда характеристические функции множеств $A \ominus B$ и $A \oplus B$ можно представить в следующем виде:

$$\begin{aligned}\chi_{A \ominus B}(z) &= \min\{\chi_A(z') \mid z' \in B_z\}, \\ \chi_{A \oplus B}(z) &= \max\{\chi_A(z') \mid z' \in B_z\}.\end{aligned}$$

Операция открытия в полутоновом случае снова представляет собой последовательное применение к изображению операций эрозии и дилатации. В случае замкнутого множества B его можно записать в виде

$$(f \circ B)(z) = ((f \ominus B) \oplus B)(z) = \max\{\min\{f(z'') \mid z'' \in B_{z'}\} \mid z' \in B_z\}.$$

Таким образом, операцию открытия кругом радиуса r (образования r -покрытия) можно рассматривать как последовательное применение к каждой точке функций минимума и максимума по круговому окну с центром в данной точке. Отметим также, что вычисление площади фигуры X можно представить как интегрирование её характеристической функции по некоторой объемлющей X области. Это позволяет ввести понятие морфологической ширины для полутоновых изображений, включая бинарные изображения как их частный случай.

Определение 3.4. Морфологической шириной векторного полутонового изображения $f : C \rightarrow \mathbb{R}$ называется функция

$$\mathfrak{F}(r) = \iint_C (f \ominus B(r)) \oplus B(r)(x,y) dx dy.$$

Иначе говоря, морфологическая ширина представляет собой совокупную интенсивность векторного изображения $f \ominus B(r)$.

Для интегрируемости потребуем, чтобы поле f было ограниченным ($f \leq f(x,y) \leq \bar{f}$), дифференцируемым почти всюду и принимало ненулевые значения в ограниченной области. Полученное в результате открытия кругом радиуса r скалярное поле $f_r(z)$ может принимать бесконечно много различных значений в различных своих точках. Заметим также, что в силу свойств функций максимума и минимума $f_r(z) \geq f_R(z)$ во всех точках, если $r \leq R$. Возникает вопрос, нельзя ли свести анализ открытия скалярного поля f к анализу открытий таких полей, которые могут служить характеристическими функциями некоторых фигур — то есть, фактически, к анализу открытий самих фигур. Ответ на этот вопрос даёт следующая теорема.

Теорема 3.3. *Результат операции открытия скалярного поля f с кругом радиуса r равен h в точке z тогда и только тогда, когда z входит в покрытие множества $T(h)$ и не входит в покрытие множества $T(h')$ при любых $h' > h$. Множество $T(h)$ определяется как $\{z : f(z) \geq h\}$.*

Доказательство. Пусть $(f \ominus B(r))(z) = h$. Это означает, что $\max_{B_z(r)}\{f \ominus B(r)\} = h$. Пусть $p \in B_z(r)$ и $(f \ominus B(r))(p) = h$. Тогда $\forall q \in B_p(r) : f(q) \geq h$, но это значит, что все точки круга $B_r(p)$, включая точку z , входят в r -покрытие $T(h)$. С другой стороны, $\forall h' > h, \forall p \in B_r(z)$ имеем $(f \ominus B(r))(p) < h'$, то есть нет ни одного покрывающего z круга радиуса r , входящего в r -покрытие $T(h')$, следовательно, сама точка z в него не входит.

Доказательство в обратную сторону не вызывает сложностей: z входит в r -покрытие для $T(h)$, но не входит для $T(h')$ при $h' > h$ — следовательно, $\exists p \in B_z(r) : \forall q \in B_p(r)$ имеем $f(q) \geq h$, то есть $(f \ominus B(r))(p) = \min\{f(z) | z \in B_p(r)\} \geq h$, но для любого $h' > h$ имеем $\forall p \in B_z(r) : (f \ominus B(r))(p) < h'$. Таким образом, максимум $f \ominus B(r)$ в круге $B_z(r)$ равен h , и $(f \ominus B(r))(z) = h$. ■

Будем считать растровое полутоновое изображение $G \in [0, \dots, 255]^{W \times H}$ дискретизацией скалярного поля f , заданного в области $C = \{(x, y) \mid 1 \leq x \leq W, 1 \leq y \leq H\}$. Проведём преобразование:

$$f(i, j) = \frac{255 - g_{ij}}{255}, \quad i = 1, \dots, W, \quad j = 1, \dots, H.$$

Инвертирование произведено для того, чтобы чёрным пикселям на бинарном изображении, которые считаются принадлежащими объекту, соответствовало нужное значение характеристической функции — 1. В целом поле f можно

рассматривать как поле вероятностей присутствия объекта в точках, так как все его значения принадлежат сегменту $[0, 1]$.

Возникает задача интерполяции поля f в точках с нецелочисленными координатами. При этом согласно теореме 3.3 нам необходимо рассмотреть дисковые покрытия множеств $T(h) = \{z : f(z) \geq h\}$ для всех h , поэтому желательно, чтобы поле принимало конечное число значений. С другой стороны, мы умеем вычислять дисковые покрытия для многоугольных фигур, поэтому логично задать множества $T(h)$ в виде многоугольных сцен. Скалярное поле, удовлетворяющее этим условиям, назовём *полутоновой многоугольной сценой*.

Покажем, что к желаемому результату приводит рассмотрение множеств точек $N_k = \{(i,j) \mid f_{ij} \geq \frac{k}{255}\}, k = 0, \dots, 255, i, j \in \mathbb{Z}$. Такое множество для k будем называть k -м срезом функции $f(x,y)$. Предлагается вычислять приближения многоугольными сценами для каждого N_k с помощью уже описанной в разделе 1.3 процедуры, так как срезы являются бинарными изображениями. Очевидно, что N_k как множество не расширяется с ростом k , а также $T(h) \subseteq T(h')$ при $h' \geq h$. Поэтому от процедуры оконтуривания, преобразующей дискретное множество точек N_k в непрерывную сцену X_k в виде набора многоугольных фигур, требуется, чтобы выполнялось $X_k \subseteq X_t$ для $k \geq t$. Тогда полутоновую многоугольную сцену \mathbf{X} можно восстановить по набору непрерывных срезов $T(\frac{k}{255}) = X_k, k = 0, \dots, 255$.

$$\mathbf{X}(z) = \begin{cases} 1, & z \in X_{255}, \\ \frac{k}{255}, & z \in (X_k \setminus X_{k+1}), \quad 0 \leq k < 255. \end{cases} \quad (3.5)$$

Следующая теорема доказывает, что переход от дискретной сцены к эквивалентной ей непрерывной, описанный в разделе 1.3, позволяет корректно восстановить полутоновую сцену.

Теорема 3.4. *Пусть I_1 и I_2 — бинарные изображения, а N_1 и N_2 — множества их чёрных пикселей, для которых выполнено $N_1 \subseteq N_2$. Если для этих изображений получены аппроксимации в виде наборов многоугольных фигур минимального периметра, то, рассматривая их как множества точек X_1 и X_2 , имеем $X_1 \subseteq X_2$.*

Доказательство. По построению граничного коридора любой лежащий в нём замкнутый контур обладает следующим свойством: все точки, лежащие внутри него, принадлежат оконтуриваемой компоненте, а все точки, лежащие

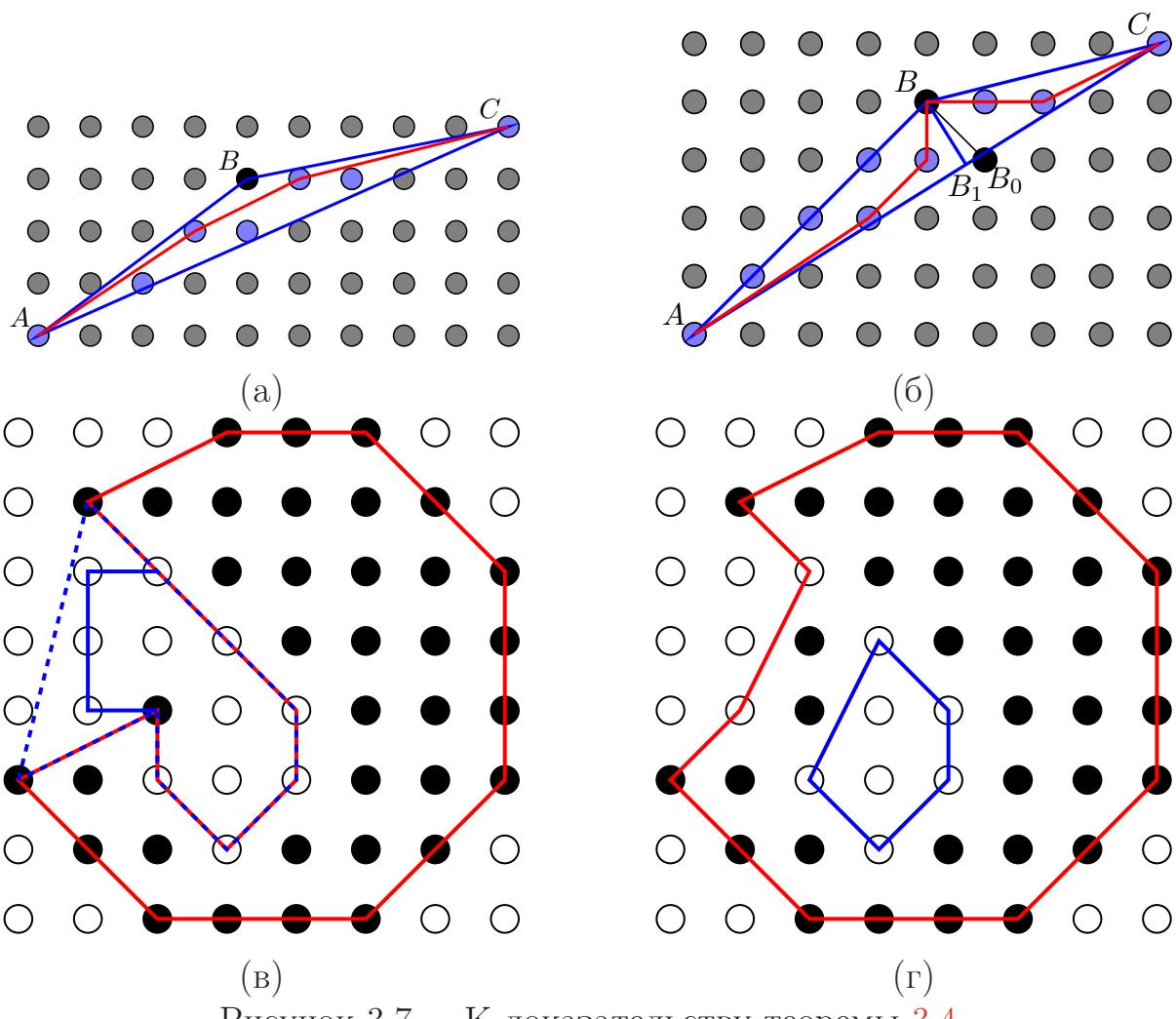


Рисунок 3.7 — К доказательству теоремы 3.4.

вне — не принадлежат. Не ограничивая общности, будем считать, что контур является внешним, а точки компоненты — чёрными. Покажем, что разделяющий многоугольник имеет минимальный периметр тогда и только тогда, когда выполняются также следующие свойства: все его вершины расположены в точках сетки (точках с целочисленными координатами), вогнутые точки многоугольника являются белыми, а выпуклые — чёрными.

Пусть последнее свойство не выполняется, и среди вершин многоугольника P есть выпуклая белая точка B . Рассмотрим две соседних с ней вершины многоугольника, обозначим их A и C так, чтобы вершины в треугольнике ABC обходились по часовой стрелке. Построим выпуклую оболочку точек, расположенных внутри треугольника ABC , включая границу, за исключением точки B . Заменим в многоугольнике P ломаную ABC на часть выпуклой оболочки, которая обходится от A к B по часовой стрелке (красная линия на рис. 3.7а). Заметим, что новый многоугольник удовлетворяет требованиям к расцветке внешних и внутренних точек (белая точка становится внешней;

также некоторые из внутренних точек могут стать граничными, для которых цвет не регламентирован), но по свойствам выпуклой оболочки имеет меньший периметр, чем исходный. Таким образом, при наличии выпуклой белой точки можно провести коррекцию и перейти к разделяющему многоугольнику меньшего периметра. Аналогичные рассуждения справедливы для вогнутых чёрных точек и для точек, не совпадающих с узлами сетки. Последовательность коррекций должна остановиться, так как разделяющих многоугольников с целочисленными координатами конечное число. Итак, многоугольник с требуемыми свойствами существует — покажем, что он единственен.

Пусть таких многоугольников два, P_1 и P_2 . Рассмотрим многоугольники $P_3 = P_1 \cup P_2$ и $P_4 = P_1 \cap P_2$, для которых также должны выполняться требования к раскраске внутренних и внешних точек. Так как P_1 и P_2 различны, существует многоугольник Q с вершинами из P_1 и P_2 , принадлежащий $P_3 \setminus P_4$, причём одна часть границы Q (будем называть её внешней) берётся из границы P_3 , а другая (внутренняя) — из границы P_4 . Так как любой многоугольник имеет не менее трёх выпуклых вершин, либо на внешней, либо на внутренней границе Q найдётся промежуточная выпуклая точка. В первом случае она должна иметь чёрный цвет, так как является выпуклой для P_3 , во втором — белый, так как является вогнутой для P_4 . Но тогда существует либо чёрная точка вне P_4 , либо белая — внутри P_3 , что приводит нас к противоречию. Таким образом, выполнение требований к раскраске выпуклых и вогнутых вершин является критерием минимальности периметра многоугольника.

Для того чтобы доказать утверждение теоремы, достаточно показать, что $X_1 \subseteq X_2$, если N_2 получено из односвязного множества N_1 добавлением единственной точки B , соседней с $B_0 \in X_1$. Для начала рассмотрим случай, когда такое перекрашивание не изменяет топологию, то есть не образует дыру. Очевидно, что если $B \in X_1$, то $X_2 = X_1$. В противном случае отрезок B_0B пересекает одну из сторон многоугольника X_1 , которую мы обозначим AC . Пусть B_1 — проекция B на AC . Построим выпуклые оболочки точек в треугольниках ABB_1 и BCB_1 и заменим отрезок AC в P ломаной из частей границ оболочек между A и B и между B и C (красная линия на рис. 3.7б). Требования к расцветке точек вне границы выполняются, так как новых таких точек не появилось, а некоторые из них, включая точку B , стали граничными. Заметим, что чёрная точка B , как и положено, оказалась выпуклой, а прочие новые граничные точки оказались вогнутыми. Отметим также, что внутренний угол

в точках A и C увеличился или остался прежним, то есть эти точки могли из выпуклых стать вогнутыми. Так как коррекция вогнутой точки порождает лишь новые вогнутые точки, нам понадобятся только коррекции такого типа. Но каждая такая коррекция расширяет многоугольник, поэтому $X_1 \subseteq X_2$.

Пусть теперь перекрашивание точки B образует дыру. Залив эту дыру, получим множество N_3 , к которому можно перейти от N_1 путём последовательных перекрашиваний, не изменяющих топологию (доказательство этого факта оставляем за рамками работы). По доказанному ранее $X_1 \subseteq X_3$ (красные контуры на рис. 3.7в-г). Рассмотрим выпуклую оболочку H_1 множества N_1 . Она содержит X_1 , а одна из связных компонент $H_1 \setminus X_1$, взятых вместе с границей, содержит все точки дыры (синий пунктирный контур на рис. 3.7в). Обозначим эту компоненту G_1 , её граница принадлежит границе X_1 за исключением отрезка, взятого из границы H_1 , соединяющего чёрные точки. При оконтуривании множества белых точек в H_1 можно отталкиваться от G_1 : выполняются все необходимые требования, кроме, возможно, вогнутости этих чёрных точек. В результате возможной коррекции выпуклых точек мы получим многоугольник $G'_1 \subseteq G_1$ (скорректированная часть границы — сплошная синяя линия на рис. 3.7в). Но по доказанному ранее при оконтуривании только точек дыры получится многоугольник $G_2 \subseteq G'_1$ (синий контур на рис. 3.7г). Так как $(X_1 \setminus \partial X_1) \cap (G_1 \setminus \partial G_1) = \emptyset$, а значит, $X_1 = \overline{X_1 \setminus G_1}$, $X_2 = \overline{X_3 \setminus G_2}$, $X_1 \subseteq X_3$, $G_2 \subseteq G_1$, получаем

$$X_1 \subseteq X_2,$$

что и требовалось доказать. ■

Замечание. Возможна ситуация, когда $N_1 \subset N_2$ (строгое вложение), но $X_1 = X_2$.

Пример установленной аппроксимации приведён на рис. 3.8, и его можно считать удачным как визуально — результат не искажает облик изображения — так и с точки зрения пригодности для дальнейшей обработки.

Пусть теперь у нас получены многоугольные сцены X_t , $t = 0, \dots, 255$, каждая из которых определяет множества, где значения \mathbf{X} не меньше $\frac{t}{255}$. Для каждого из этих множеств построим дисковое r -покрытие $X_t(r)$, тогда открытие полуточновой сцены \mathbf{X} представляется через систему вложенных покрытий:

$$(\mathbf{X} \circ B(r))(z) = \begin{cases} \frac{k}{255}, & z \in (X_k(r) \setminus X_{k+1}(r)), \quad 0 \leq k < 255. \\ 1, & z \in X_{255}(r), \end{cases}$$



Рисунок 3.8 — (а) Полутоновое изображение и (б) его аппроксимация полутоновой многоугольной сценой

Следовательно, морфологическую ширину данного векторного изображения можно рассчитать по формуле:

$$\mathfrak{F}(r) = \sum_{i=0}^{254} \frac{i}{255} (\mathfrak{F}_{X_i}(r) - \mathfrak{F}_{X_{i+1}}(r)) + \mathfrak{F}_{X_{255}}(r) = \frac{1}{255} \sum_{i=1}^{255} \mathfrak{F}_{X_i}(r), \quad (3.6)$$

нижний индекс у символа \mathfrak{F} указывает, для какой фигуры вычисляется площадь покрытия. Фактически, вышеприведённые рассуждения являются адаптацией к непрерывному случаю аппарата пороговой декомпозиции дискретной математической морфологии на операции с бинарными изображениями [73].

Это означает, что процедура вычисления морфологической ширины полуточновой сцены фактически сводится к усреднению функций морфологической ширины для всех её срезов (нулевой срез представляет собой всю область изображения и не учитывается). Таким образом, исходную задачу необходимо всего лишь решить многократно, что допускает простую программную реализацию, но оборачивается значительным увеличением вычислительных затрат. Один из путей ускорения процедуры расчёта ширины в этом случае — вычислять её приближенно, используя разложение изображения на меньшее, чем 256, число уровней. Способ определения разложения с наименьшей потерей информации на базе метода динамического программирования описан в работе [74].

Можно также отследить ситуации, когда $X_i = X_{i+1}$, или какие-то из многоугольников присутствуют и в X_i , и в X_{i+1} , чтобы не вычислять одно и то же покрытие несколько раз. Достаточно распространённой, однако, является такая ситуация, в которой X_i изменяется постепенно, и многоугольники имеют незначительные отличия. Хотя они по-прежнему могут отличаться лишь

локально, и их скелеты могут во многом совпадать, что свидетельствует о дублирующихся вычислениях, выявление таких случаев и определение общих ветвей вычислений является весьма нетривиальной задачей, лежащей за рамками данного исследования.

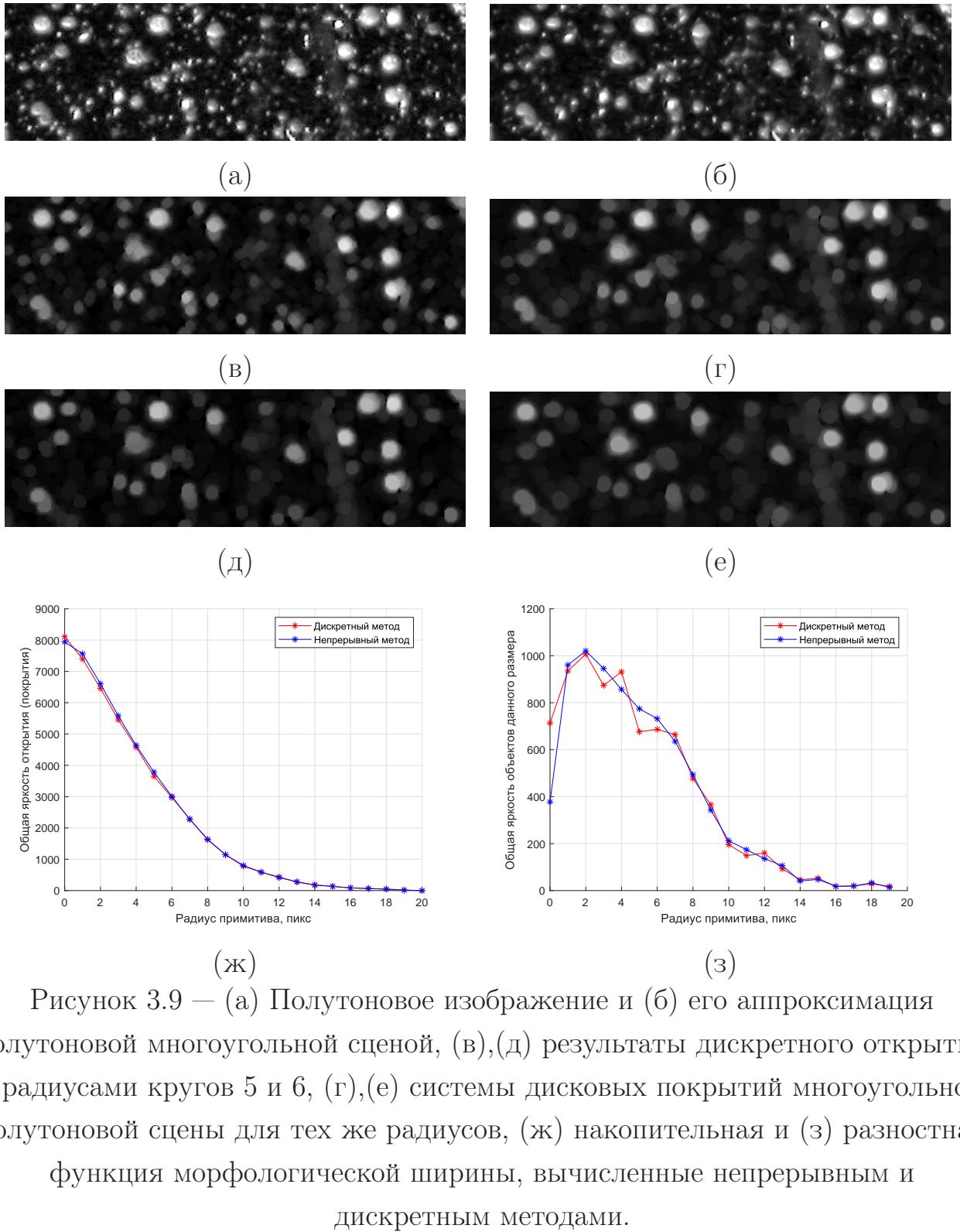


Рисунок 3.9 — (а) Полутоновое изображение и (б) его аппроксимация полутоновой многоугольной сценой, (в),(д) результаты дискретного открытия с радиусами кругов 5 и 6, (г),(е) системы дисковых покрытий многоугольной полутоновой сцены для тех же радиусов, (ж) накопительная и (з) разностная функция морфологической ширины, вычисленные непрерывным и дискретным методами.

Пример вычисления морфологической ширины полутонового изображения снежных хлопьев приведён на рис. 3.9. Небольшие различия в результатах между непрерывным и дискретным методом на базе морфологических операций из системы MATLAB вызваны особенностями аппроксимации — как формы фигур при построении контуров в непрерывном случае, так и растеризацией примитивов в дискретном. Отметим, что непрерывная модель позволяет посчитать как саму морфологическую ширину, так и статистики, производные от неё — например, среднюю интенсивность точек покрытия или её дисперсию. Таким образом, в большинстве задач анализа изображений, использующих морфологический спектр, можно заменить процедуру его вычисления на процедуру вычисления морфологической ширины.

3.3 Карты морфологических признаков

Введённая ранее морфологическая ширина фигуры является интегральным описанием некоторой её морфологической характеристики, которую мы условились называть шириной. Фактически она представляет собой функцию распределения некоторой величины, определённой в каждой точке фигуры, а при вычислении данной функции с заданным шагом мы вычисляем накопительную гистограмму. В практических приложениях нам может пригодиться не только производная статистика данной величины, но и сам результат её вычисления в конкретных точках фигуры. Можно найти этому несколько причин. Во-первых, нам могут понадобиться такие процедуры обработки информации о ширине, которые очень сложно применить непосредственно к дисковому покрытию как геометрическому объекту. К примеру, очень проблематично вычислить по дисковому покрытию аналитически моменты Цернике [75] — функции вида $Z_n^m(\rho, \varphi) = R_n^m(\rho) \cos(m\varphi)$ и $Z_n^{-m}(\rho, \varphi) = R_n^m(\rho) \sin(m\varphi)$, где $R_n^m(\rho)$ — многочлены. Возможный подход здесь — переход к работе с дисковым покрытием в растровом представлении. Во-вторых, большинство существующих методов компьютерного зрения и анализа изображений предназначены для работы с растровой информацией и могут использовать карты признаков, полученные в результате некоторого преобразования изображений, для их последующей обработки — обнаружения точек интереса, выделения характерных фрагментов, вычисления признаков высокого уровня и т.д. К примеру, точечные оценки ширины — ключевой этап работы при вычислении признакового описания в ряде методов распознавания формы объектов на изображениях [25; 27; 29; 76].

3.3.1 Расчёт карты расстояний

Пожалуй, самым известным примером морфологического преобразования бинарного изображения является *distance transform*, который мы будем называть *картой расстояний*. Формально картой расстояний бинарного изображения B является функция, которая каждой точке p переднего плана ($B(p) = 1$) присваивает расстояние до некоторого множества особых пикселей E :

$$DT(p) = \min_q \{d(p,q) \mid q \in E\}.$$

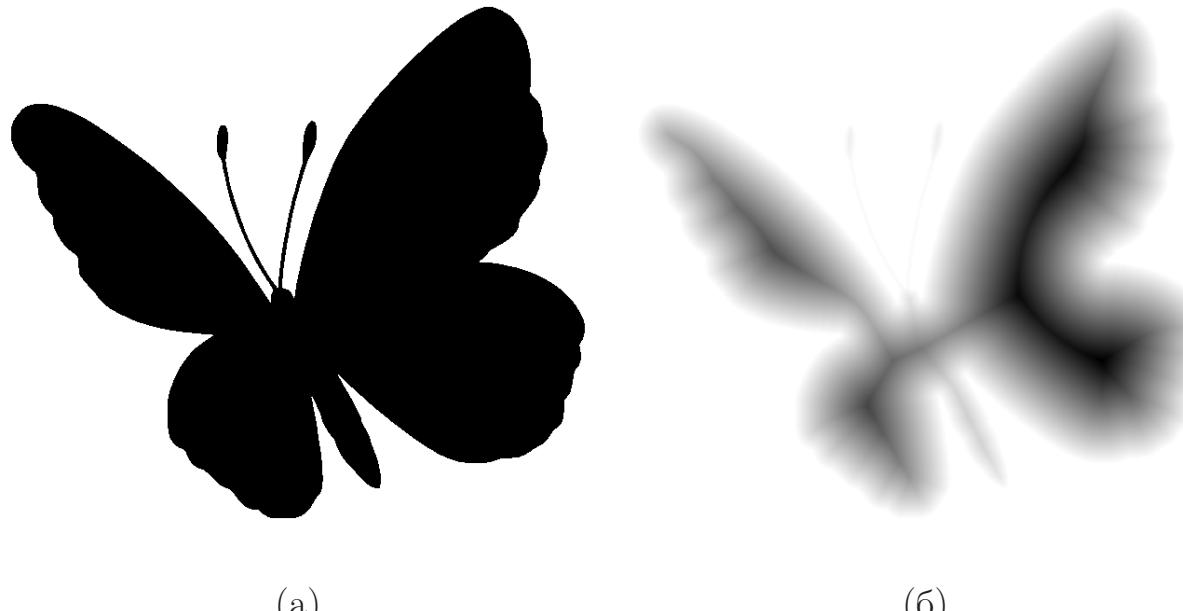


Рисунок 3.10 — Бинарное изображение (а) и его карта расстояний (б). Более тёмные точки соответствуют большим значениям функции.

Разнообразие вариаций карт расстояний обеспечивается выбором множества точек E и функции расстояния d . В самом распространённом случае под множеством E поднимается множество пикселей фона (при этом достаточно рассмотреть только пиксели границы), а расстояние считается евклидовым. Пример карты расстояний показан на рис. 3.10.

Задача 3.2. *Дано бинарное изображение B . Вычислить преобразование $DT(p) = \min_q \{\|p - q\|_2 \mid q \in \{s \mid B(s) = 0\}\}, p \in \{s \mid B(s) = 1\}\}$.*

Известен ряд способов быстрого вычисления карты расстояний, принципиально различающихся по своим подходам. Можно выделить следующие категории:

- алгоритмы, основанные на принципе распространения [77; 78]. Такие алгоритмы можно считать реализацией метафоры «пожара в прерии»: стартуя с граничных точек, огневой фронт последовательно продвигается внутрь фигуры, обновляя при этом текущее расстояние до границы, до момента столкновения с фронтом, пришедшим с другой стороны. Примечательно, что столкновение происходит в точках скелета, а расстояние до границы в этот момент определяет значение радиальной функции. Принцип распространения находит своё применение в задачах математической физики, так как может быть использован для решения уравнения эйконала [79];
- алгоритмы растрового сканирования, которые несколько раз производят преобразования изображения по строкам или по столбцам а итоговый результат получают путём агрегации результатов таких преобразований [80];
- алгоритмы, основанные на поиске пересечения семейства парабол. В случае евклидовой метрики расстояние от точки из определённого столбца до ближайшей точки фона в выбранной строке представляет собой квадратичную функцию. Таким образом, поиск общего минимума по всем строкам приводит к анализу нижней огибающей семейства парабол [81; 82];
- алгоритмы, использующие принципы математической морфологии. Известно, что результат дистанционного преобразования можно получить, применив к изображению операцию эрозии с достаточно большим полтоновым дисковым структурирующим элементом [83], которую также можно рассматривать как операцию свёртки двух функций с выбором минимального значения [84];
- алгоритмы, основанные на построении диаграммы Вороного. В данном случае сайтом считается отдельная точка границы, а заполнение карты расстояний производится в результате сканирования соответствующих ячеек Вороного и вычисления расстояния до этого сайта [85].

Как показывает последний из упомянутых подходов, результат вычисления карты расстояний тесно связан с диаграммой Вороного точек множества E . Действительно, если известно, каким ячейкам диаграммы принадлежат точки изображения, дальнейшая процедура вычисления становится тривиальной: нужно лишь определить расстояние до соответствующего сайта.

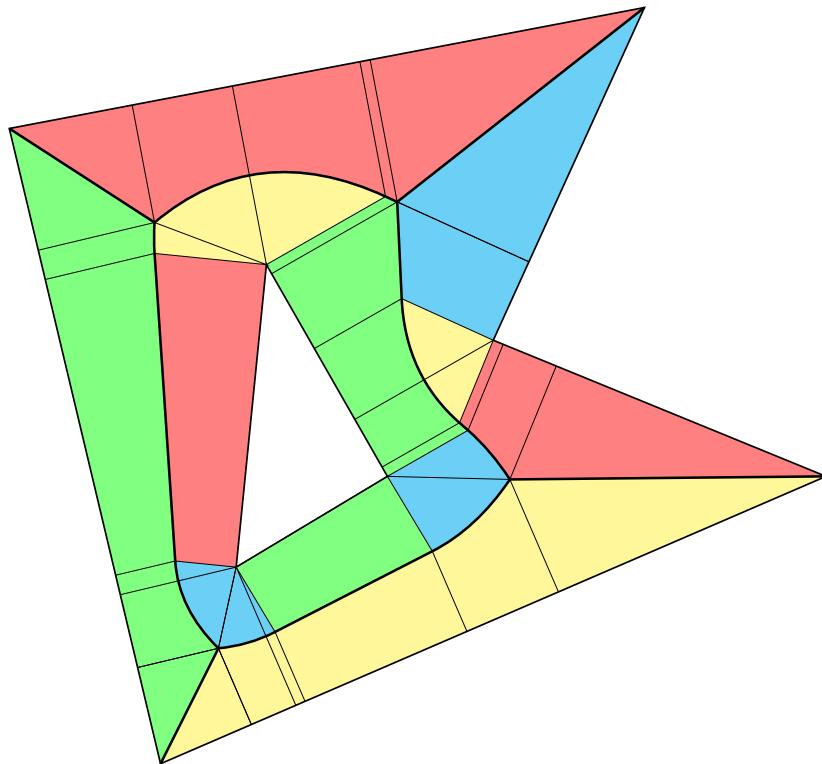


Рисунок 3.11 — Ячейки Вороного, складывающиеся из собственных подобластей

Но если ставится задача определения расстояния до ближайшей точки границы, то можно воспользоваться её многоугольной аппроксимацией, и процедура скелетизации фактически определит все ближайшие сайты. Рассмотрим ребро e скелета \mathcal{S} . У этого ребра есть пара образующих сайтов, для которых ребро является бисектором. Собственная область ребра образована множеством спиц, для которых ближайшие точки границы расположены на соответствующих сайтах. Значит, ребро делит собственную область на две подобласти, в каждой из которых ближайший сайт один и тот же — эти подобласти являются частями ячеек Вороного (рис. 3.11). Остаётся только проследить все точки дискретной сетки, попадающие в подобласть, и для каждой из них найти расстояние до соответствующего сайта. В случае сайта-точки вычисления очевидны, в случае сайта-сегмента расстояние от произвольной точки подобласти C до сегмента AB можно найти как расстояние от C до её проекции на сегмент — точки $P_{AB}(C)$.

Опишем вид необходимых подобластей.

1. Для линейного бицикла с ребром AB и образующими сайтами CD и EF это два четырёхугольника: $\{A, P_{CD}(A), P_{CD}(B), B\}$ и $\{A, P_{EF}(A), P_{EF}(B), B\}$.

2. Для гиперболического бицикла с образующими сайтами C и D это два треугольника: $\{A,C,B\}$ и $\{A,B,D\}$.
3. Для параболического бицикла с сайтом точкой C и сайтом-сегментом DE это криволинейный треугольник $\{A,C,B\}$ и криволинейный четырёхугольник $\{A, P_{DE}(A), P_{DE}(B), B\}$, так как сторона AB в этих фигурах представляет собой отрезок параболы.

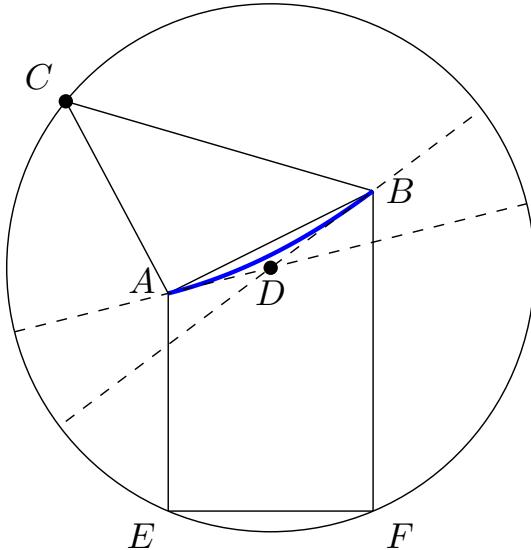


Рисунок 3.12 — Контрольная точка D параболического бицикла с концами ребра A и B , сайтом-точкой C и сайтом-сегментом, содержащим отрезок EF

Процедура прослеживания точек в фигуре с криволинейной границей представляет собой некоторую сложность, которую мы устраним следующим образом. Определим контрольную точку параболического бицикла — точку пересечения касательных к параболическому ребру в его концах. Известно, что касательная к параболе в точке содержит биссектрису угла, образованного фокусом параболы, самой точкой и её проекцией на директрису. Поэтому по построению контрольная точка равноудалена от сайта-точки и двух проекций концов ребра на сайт-сегмент, т.е. является центром описанной около треугольника из этих точек окружности. В обозначениях из рис. 3.12 система уравнений для поиска контрольной точки принимает вид:

$$\begin{cases} (x_D - x_C)^2 + (y_D - y_C)^2 = (x_D - x_E)^2 + (y_D - y_E)^2, \\ (x_D - x_C)^2 + (y_D - y_C)^2 = (x_D - x_F)^2 + (y_D - y_F)^2, \end{cases}$$

$$\begin{cases} -2x_D x_C + x_C^2 - 2y_D y_C + y_C^2 = -2x_D x_E + x_E^2 - 2y_D y_E + y_E^2, \\ -2x_D x_C + x_C^2 - 2y_D y_C + y_C^2 = -2x_D x_F + x_F^2 - 2y_D y_F + y_F^2, \end{cases}$$

$$\begin{bmatrix} x_D \\ y_D \end{bmatrix} = \begin{bmatrix} 2(x_E - x_C) & 2(y_E - y_C) \\ 2(x_F - x_C) & 2(y_F - y_C) \end{bmatrix}^{-1} \begin{bmatrix} x_E^2 + y_E^2 - x_C^2 - y_C^2 \\ x_F^2 + y_F^2 - x_C^2 - y_C^2 \end{bmatrix}.$$

Таким образом, внутри треугольника ABC ближайшим сайтом является сайт-точка, внутри четырёхугольника $ACFE$ — сайт-сегмент, а внутри треугольника ABD ближайший сайт под вопросом, поэтому нужно определить расстояния до обоих сайтов и взять минимальное. Отметим также, что бицикли могут быть немонотонными, и процедуру их разбиения проводить необязательно.

В итоге алгоритм обхода точек с целыми коэффициентами, лежащих внутри выпуклого многоугольника P (все подобласти в нашем случае выпуклые) можно описать в виде алгоритма 3. Принцип его работы основан на том свойстве, что непустое пересечение многоугольника прямой вида $y = c$ представляет собой единственный отрезок. Ось OY в алгоритме считается направленной вверх. Применение алгоритма для каждой подобласти каждого ребра скелета даёт нам решение задачи 3.2.

Алгоритм был реализован на языке C++, работа проводилась на компьютере с процессором Intel® Core i5™ и 6 ГБ оперативной памяти. Для оценки вычислительной сложности разработанного алгоритма можно обратиться к факту из [86]: каждую строку изображения с w столбцами может пересекать не более w ячеек Вороного сайтов-точек с целыми координатами, соответствующих пикселям. Но это значит, что при организации сканирования ячеек Вороного по h строкам мы произведём не более $wh = m$ операций. Это свойство сохраняется и при замене сайтов-точек на обобщённые сайты — точки или сегменты. Таким образом, общая сложность алгоритма построения карты расстояний на основе непрерывного скелета, включая само построение скелета, составляет $O(n \log n) + O(m)$ операций, где n — число сайтов диаграммы Вороного, а m — число пикселей бинарного изображения.

Эту оценку подтверждает таблица 1: эффективность метода относительно остальных напрямую связана с выигрышем от перехода от дискретного представления изображения к непрерывному. Если границы объектов на изображении достаточно простые и гладкие, метод значительно превосходит конкурентов в эффективности, однако если на изображении много мелких объектов или границы объектов слишком извилистые, переход к непрерывному представлению себя не окупает, и основное время работы алгоритма

Алгоритм 3 Прослеживание выпуклого многоугольника

Дано: Выпуклый многоугольник $P = \{(x_i, y_i)\}, i = 1, \dots, N$, в котором точки перечислены в порядке обхода против часовой стрелки

Найти: $DT(p)$ для точек внутри P

- 1: Определить b и t — номера нижней и верхней точек
 - 2: Построить левую и правую границы многоугольника:
 - 3: $L = \{x_i^l, y_i^l\} = \{(x_j, y_j)\}, j = b, b-1, \dots, 1, N, \dots, t+1, t$ и
 - 4: $R = \{x_i^r, y_i^r\} = \{(x_j, y_j)\}, j = b, b+1, \dots, N, 1, \dots, t-1, t$
 - 5: $i = 1, j = 1$
 - 6: **Для** $y = \lceil y_t \rceil, \dots, \lfloor y_b \rfloor$ **выполнять**
 - 7: **До тех пока** $y_i^l \leq y$ **выполнять**
 - 8: $i = i + 1$
 - 9: **Конец цикла**
 - 10: **До тех пока** $y_j^r \leq y$ **выполнять**
 - 11: $j = j + 1$
 - 12: **Конец цикла**
 - 13: **Если** $i \leq |L|$ и $j \leq |R|$ **тогда**
 - 14: $x_l = x_{i-1}^l + \frac{y - y_{i-1}^l}{y_i^l - y_{i-1}^l} (x_i^l - x_{i-1}^l)$
 - 15: $x_r = x_{j-1}^r + \frac{y - y_{j-1}^r}{y_j^r - y_{j-1}^r} (x_j^r - x_{j-1}^r)$
 - 16: **иначе**
 - 17: $x_l = x_{i-1}^l, x_r = x_{j-1}^r$
 - 18: **Конец условия**
 - 19: **Для** $x = \lceil x_l \rceil, \dots, \lfloor x_r \rfloor$ **выполнять**
 - 20: Вычислить расстояние до нужного сайта (пары сайтов) в точке (x, y)
 - 21: **Конец цикла**
 - 22: **Конец цикла**
-

Изображение					
Размер (пикс.)	640×640	442×1312	2156×2352	191×274	488×275
Пиксели объекта	168763	321297	1894986	2430	35775
Рёбра скелета	763	1035	4399	415	1303
Лотуфо-Зампиролли [83]	13,85	24,76	308,10	0,54	2,29
Майстер [82]	7,59	9,18	121,66	0,72	2,41
Маурер [85]	10,16	12,26	145,21	0,93	2,74
Кьюзенэр [78]	8,54	18,87	129,65	0,44	3,12
Фельценцвальб [84]	13,46	18,43	216,74	1,29	3,67
Ломов	5,33	8,27	35,77	1,41	4,79
Скелетизация	2,40	3,69	11,53	1,14	3,42

Таблица 1 — Временные затраты на расчёт карты расстояний в миллисекундах. Затраты на собственный метод измерены с учётом скелетизации.

уходит на построение скелета. Несмотря на это, для достаточно широкого класса изображений — имеющих хорошее разрешение и содержащих фигуры с отчётливой формой — метод оказывается крайне эффективным и притом весьма простым в реализации. Код для сравнения был взят из проекта <http://distance.sourceforge.net/>, служащего сопровождением к статье [87] с подробным обзором задачи и методов её решения.

3.3.2 Расчёт карты толщин

Вернёмся к понятию *гранулометрической функции*, рассмотренной в подразделе 1.1. Напомним, что гранулометрическая функция $G_\Psi(p)$, порождённая гранулометрией $\Psi = \{\Psi_r\}_{r \geq 0}$ определяется для точек фигуры X следующим образом:

$$G_\Psi(p) = \max\{r \geq 0 \mid p \in \Psi_r(X)\}.$$

В терминах дискового покрытия операция $\Psi_r(X)$ превращается в построение дискового покрытия фигуры X , поэтому гранулометрическая функция представляется как

$$G(p) = \max\{r \geq 0 \mid p \in X(r)\}. \quad (3.7)$$

Соответственно, дисковое r -покрытие можно восстановить по значению функции $G(p)$:

$$X(r) = \{p \mid G(p) \geq r\}.$$

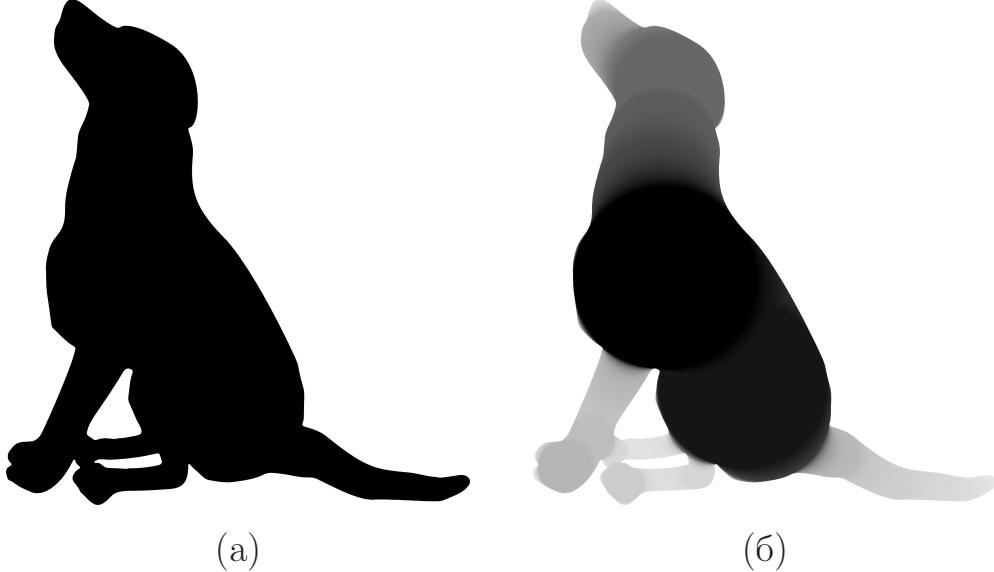


Рисунок 3.13 — (а) Бинарное изображение и (б) его карта толщин. Более тёмные точки соответствуют большим значениям функции.

Гранулометрические функции успешно применялись для решения задач оценки механических свойств углеродистой стали [88], сегментации изображений раздробленных горных пород [89] и поиска движущихся человеческих фигур на видеопоследовательностях [76]. Также весьма оригинальный пример использования этого морфологического преобразования можно найти в работе [90], в которой значения гранулометрической функции использовались для оценки сложности мозаичных форм, что позволило совместить методы двух математических морфологий — Пытьева [91] и Серра [14]. В практических приложениях, как правило, в силу дискретной природы входных данных значение гранулометрической функции определяется для точек, расположенных в узлах целочисленной сетки. Отметим, что в русскоязычной литературе за гранулометрической функцией, порожденной семейством открытий со структурирующим элементом, закрепилось название *карты толщин*, которое мы и будем использовать. Пример визуализации карты толщин представлен на рис. 3.13.

Задача 3.3. *Дана многоугольная фигура X . Вычислить значения функции $G(p) = \max\{r \geq 0 \mid p \in X(r)\}$ в точках множества $X \cap \mathbb{Z}^2$.*

Для эффективного вычисления карты толщин бинарного изображения использовался двупроходный рекурсивный алгоритм, основанный на принципе распространения [92]. Также, согласно принципам кодирования изображений с помощью скелета [19], карта толщин в дискретном случае может быть восстановлена из скелета растеризацией ассоциированных с точками скелета кругов. Этот подход использовался в работе [93], причём дискретный скелет был получен из непрерывного в результате растеризации.

Остановимся на связи карты толщин и скелета подробнее. Из уравнения 3.7 следует, что карта толщин в точке p принимает значение, равное радиусу максимального покрывающего p круга, лежащего целиком внутри фигуры. Это значение естественно интерпретировать как ширину фигуры в точке. Очевидно, можно рассматривать только максимальные по включению круги, поэтому карту толщин можно определить через скелет \mathcal{S} и радиальную функцию $r(q)$:

$$G(p) = \max_{q \in \mathcal{S}} \{r(q) \mid x \in B_q(r(q))\}. \quad (3.8)$$

Таким образом, вычисление ширины фигуры в точке можно свести к поиску конкретного максимального покрывающего круга. Представим скелет \mathcal{S} в виде произвольного объединения подграфов:

$$\mathcal{S} = \bigcup_{i=1}^k \mathcal{S}_i.$$

В этом случае можно задать функцию ширины в точке относительно подграфа:

$$G_i(p) = \begin{cases} \max\{r(q) \mid p \in B_q(r(q))\}, & \text{если } \exists q \in \mathcal{S}_i : p \in B_q(r(q)); \\ -1, & \text{иначе.} \end{cases} \quad (3.9)$$

и рассматривать функцию $G(p)$ в виде

$$G(p) = \max_{i=1,\dots,k} G_i(p). \quad (3.10)$$

Разбиение скелета на подграфы может упростить процедуру вычисления значения функции ширины в точке благодаря следующей теореме.

Теорема 3.5. Пусть $L = \mathcal{S}_i$ — простая кривая, принадлежащая скелету фигуры X , радиальная функция при движении по кривой L изменяется монотонно, и точка $p = (x_0, y_0)$ покрывается одним из вписанных кругов с

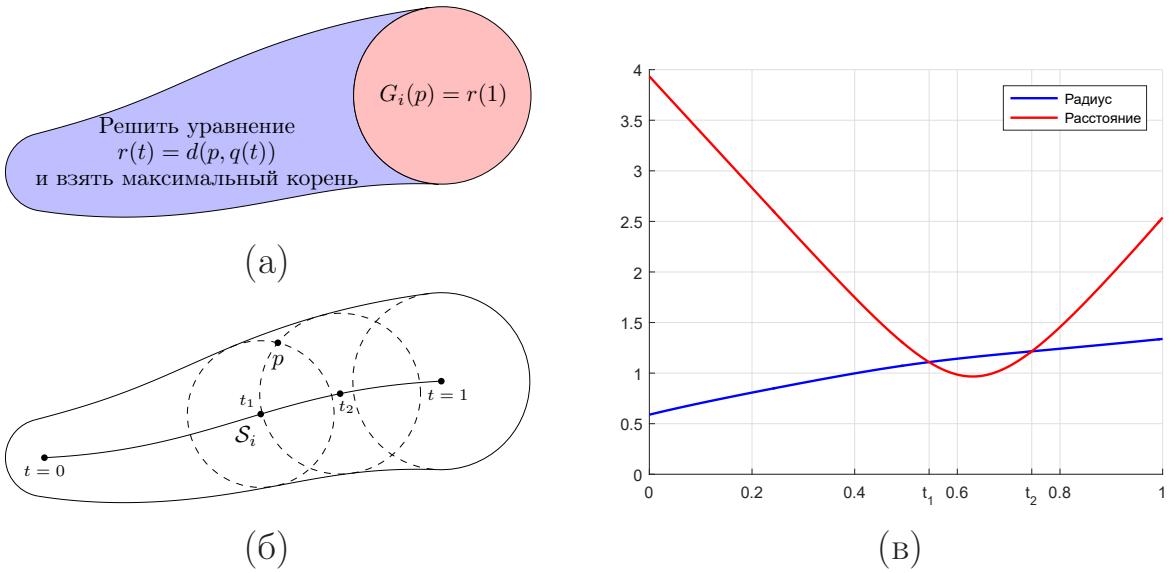


Рисунок 3.14 — Поиск максимального круга покрывающего радиуса в силуэте монотонного подграфа. (а) Схема подхода, (б) предельные круги для точки p , (в) графики функций $r(t)$ и $d(p, q(t))$. В этом случае $G_i(p) = r(t_2)$.

центром на L . Тогда максимальный покрывающий p круг с центром на L — это либо круг с наибольшим радиусом на всей кривой, либо такой круг, что p лежит на его границе.

Доказательство. Параметризуем кривую L в направлении роста радиальной функции: $q = q(t) = (x(t), y(t))$, $r = r(t)$, $t \in [0,1]$. Рассмотрим следующую функцию:

$$g(t) = r^2(t) - d^2(p, q(t)), \quad (3.11)$$

где d — евклидово расстояние между точками. Функция $g(t)$ неотрицательна, если точка p лежит внутри вписанного круга с центром в точке $q(t)$, равна нулю, если точка p лежит на границе круга, и отрицательна, если точка p не покрыта кругом. Также эта функция непрерывна в силу непрерывности скелета и радиальной функции. Пусть точка p не принадлежит максимальному кругу кривой. В силу условий задачи и непрерывности $g(t)$ множество $P = \{t : g(t) \geq 0\}$ является компактом, и максимальное число из этого множества $t' < 1$. Очевидно, что $q(t')$ — центр максимального покрывающего круга. Но $g(t') = 0$ также в силу непрерывности функции $g(t)$ и непостоянства её знака в окрестности точки t' . ■

Таким образом, для вычисления значения $G_i(p)$ достаточно двух шагов:

1. Если $g(1) \geq 0$, принять $G_i(p) = r(1)$.

2. Иначе решить уравнение $g(t) = 0$ и принять $G_i(p) = r(t_{\max})$, где t_{\max} — максимальный из корней.

Отметим, что для применения теоремы 3.5 скелет должен быть разбит на монотонные подграфы. Подходящим примером таких подграфов могут выступать оси монотонных бициклов. В этом случае необходимо прояснить вопрос решения уравнения $g(t) = 0$ для бициклов всех трёх типов.

Будем обозначать (x_1, y_1) , (x_2, y_2) — координаты концов оси бицикла, r_1, r_2 ($r_1 \leq r_2$) — радиусы концевых кругов, (x_3, y_3) — координаты образующего сайта-точки в случае гиперболического (можно взять любой из двух сайтов) или параболического бицикла, (x_0, y_0) — координаты точки p , в которой вычисляем значение карты толщин.

1. Радиальная функция линейного бицикла изменяется линейно при равномерном движении по оси, поэтому бицикл можно параметризовать следующим образом:

$$\begin{aligned} x(t) &= x_1 + t(x_2 - x_1); \\ y(t) &= y_1 + t(y_2 - y_1); \\ r(t) &= r_1 + t(r_2 - r_1). \end{aligned} \tag{3.12}$$

В результате параметризации уравнение $g(t) = 0$ становится квадратным:

$$\begin{aligned} (x_1 + t(x_2 - x_1) - x_0)^2 + (y_1 + t(y_2 - y_1) - y_0)^2 &= (r_1 + t(r_2 - r_1))^2, \\ t^2 \left((x_2 - x_1)^2 + (y_2 - y_1)^2 - (r_2 - r_1)^2 \right) + \\ + 2t \left((x_2 - x_1)(x_1 - x_0) + (y_2 - y_1)(y_1 - y_0) - r_1(r_2 - r_1) \right) + \\ \left((x_1 - x_0)^2 + (y_1 - y_0)^2 - r_1^2 \right). \end{aligned} \tag{3.13}$$

2. Отрезок параболы является частным случаем кривой Безье второго порядка. Две её крайних контрольных точки совпадают с концами отрезка, а промежуточная (x_4, y_4) является точкой пересечения касательных к параболе в концах отрезка (рис. 3.12). Параметризация имеет следующий вид:

$$\begin{aligned} x(t) &= (1-t)^2 x_1 + 2t(1-t)x_4 + t^2 x_2; \\ y(t) &= (1-t)^2 y_1 + 2t(1-t)y_4 + t^2 y_2; \\ r(t) &= \sqrt{(x(t) - x_3)^2 + (y(t) - y_3)^2}. \end{aligned} \tag{3.14}$$

Уравнение $g(t) = 0$ принимает вид

$$\begin{aligned}
 & (x(t) - x_0)^2 + (y(t) - y_0)^2 = (x(t) - x_3)^2 + (y(t) - y_3)^2 \Rightarrow \\
 & -2x(t)x_0 + x_0^2 - 2y(t)y_0 + y_0^2 = -2x(t)x_3 + x_3^2 - 2y(t)y_3 + y_3^2, \\
 & 2((1-t)^2x_1 + 2t(1-t)x_4 + t^2x_2)(x_3 - x_0) + \\
 & + 2((1-t)^2y_1 + 2t(1-t)y_4 + t^2y_2)(y_3 - y_0) + \\
 & + x_0^2 - x_3^2 + y_0^2 - y_3^2, \\
 & t^2(2(x_1 - 2x_4 + x_2)(x_3 - x_0) + 2(y_1 - 2y_4 + y_2)(y_3 - y_0)) + \\
 & + t(2(-2x_1 + 2x_4)(x_3 - x_0) + 2(-2y_1 + 2y_4)(y_3 - y_0)) + \\
 & + (2x_1(x_3 - x_0) + 2y_1(y_3 - y_0) + x_0^2 - x_3^2 + y_0^2 - y_3^2). \quad (3.15)
 \end{aligned}$$

В итоге снова имеем квадратное уравнение.

3. Наконец, гиперболический бицикл можно параметризовать так:

$$\begin{aligned}
 x(t) &= x_1 + t(x_2 - x_1); \\
 y(t) &= y_1 + t(y_2 - y_1); \\
 r(t) &= \sqrt{(x(t) - x_3)^2 + (y(t) - y_3)^2}.
 \end{aligned} \quad (3.16)$$

Интересующее нас уравнение снова можно представить в виде

$$-2x(t)x_0 + x_0^2 - 2y(t)y_0 + y_0^2 = -2x(t)x_3 + x_3^2 - 2y(t)y_3 + y_3^2.$$

Соответствующие преобразования:

$$\begin{aligned}
 & 2(x_1 + t(x_2 - x_1))(x_3 - x_0) + 2(y_1 + t(y_2 - y_1))(y_3 - y_0) + x_0^2 - x_3^2 + y_0^2 - y_3^2, \\
 & t(2(x_2 - x_1)(x_3 - x_0) + 2(y_2 - y_1)(y_3 - y_0)) + \\
 & + (2x_1(x_3 - x_0) + 2y_1(y_3 - y_0) + x_0^2 - x_3^2 + y_0^2 - y_3^2) \quad (3.17)
 \end{aligned}$$

на этот раз дают нам линейное уравнение.

Таким образом, для вычисления ширины в точке относительно бицикла любого вида достаточно решить квадратное или линейное уравнение. Итоговый алгоритм расчёта карты толщин сводится к поиску максимального значения ширины относительно бицикла по всем бициклам согласно формуле 3.10. При этом

из соображений эффективности выгоднее для каждого бицикла определять множество принадлежащих ему точек сетки, а не определять для каждой точки множество покрывающих её бициклов. Обход точек бицикла, в свою очередь, можно свести к растеризации кругов и выпуклых многоугольников, рассмотренных в разделе 3.3.1. В итоге получаем решение задачи 3.3.

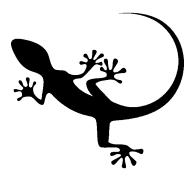
Изображение				
Размер (пикс.)	556×685	626×562	326×512	2000×1053
Пиксели объекта	189926	73182	57729	297438
Пиксели скелета	1876	2370	5075	19596
Рёбра скелета	693	1385	2896	11619
Дискретный [50]	3512	469	205	702
Дискретно-непрерывный [93]	45,45	11,01	11,68	64,19
Непрерывный	53,77	26,87	21,29	111,89
Скелетизация	2,82	5,91	9,13	46,91

Таблица 2 – Временные затраты на расчёт карты толщин в миллисекундах. Затраты на непрерывный и дискретно-непрерывный методы измерены с учётом скелетизации.

Реализованный на языке C++ алгоритм расчёта карты толщин непосредственно по непрерывному скелету сравнивался с алгоритмом из работы [93], использующим растеризованную версию непрерывного скелета, и методом на основе прямого вычисления операции открытия по алгоритму из [50]. Результаты, представленные в таблице 2 свидетельствуют о том, что метод работает систематически быстрее дискретного аналога, однако уступает комбинированному подходу, особенно в случае изображений с тонкими элементами. Хотя сложность непрерывного (число рёбер) скелетного представления, как правило, ощутимо меньше сложности дискретного (число пикселей), необходимость действовать в вещественнонозначной, а не целочисленной, арифметике, а также решать для каждой точки уравнение, а не присваивать наперёд заданное значение при растеризации круга, ведёт к общему снижению эффективности. Тем не менее очевидным достоинством предложенного алгоритма можно считать его

более высокую точность: к примеру, он способен выдавать для ширины фигуры в точке нецелые значения.

Карту толщин можно использовать для расчёта сложных дескрипторов формы, которые проблематично вычислить на основе исключительно непрерывного представления. Кратко опишем идею таких дескрипторов. Пусть помимо карты толщин $G(p) = F_1(p)$ имеются карты других признаков, вычисленных для каждой точки p из множества точек интереса $S \subseteq X$: $F_i(p)$, $i = 2, \dots, k$. Тогда можно вычислить дескрипторы (гистограммы) совместных значений признаков:

$$D(v_1, \dots, v_k) = |\{p \in S : F_i(p) = v_i, i = 1, \dots, k\}|.$$

Если же имеются парные характеристики $H_i(p, q)$, $i = 1, \dots, m$, можно вычислить гистограммы для каждой точки по отдельности:

$$D_p(v_1, \dots, v_m) = |\{q \in S : H_i(p, q) = v_i, i = 1, \dots, m\}|.$$

Некоторые из парных характеристик вычисляются для пар точек независимо, другие можно получить из индивидуальных: например, $H_i(p, q) = G(p) - G(q)$. В частности, в работе [42] исследуются скелетно-геодезические гистограммы толщин, направлений и расстояний, толщинная информация для которых извлекается из соответствующей карты. В докладах [39] и [35] был освещён усовершенствованный дескриптор, в котором информация о толщинах, расстояниях и углах вычисляется отдельно по сегментам фигуры, а затем набор полученных дескрипторов кодируется по принципу «мешка» [94], определяя таким образом содержание изображения в виде наличия в нём некоторых шаблонов гибкой формы.

3.4 Морфологические моменты многоугольных фигур

3.4.1 Задача пространственного описания ширины

Рассмотренная в предыдущей главе площадь дискового покрытия (морфологическая ширина) реализует на базе непрерывного представления формы идеи морфологического спектра, который является одной из распространённых техник морфологического анализа изображений. Она описывает вклад в формирование изображения примитивов различного размера и в силу выбора диска в качестве примитива и строго сегментированных входных может рассматриваться как интегральное описание ширины объектов на изображении.

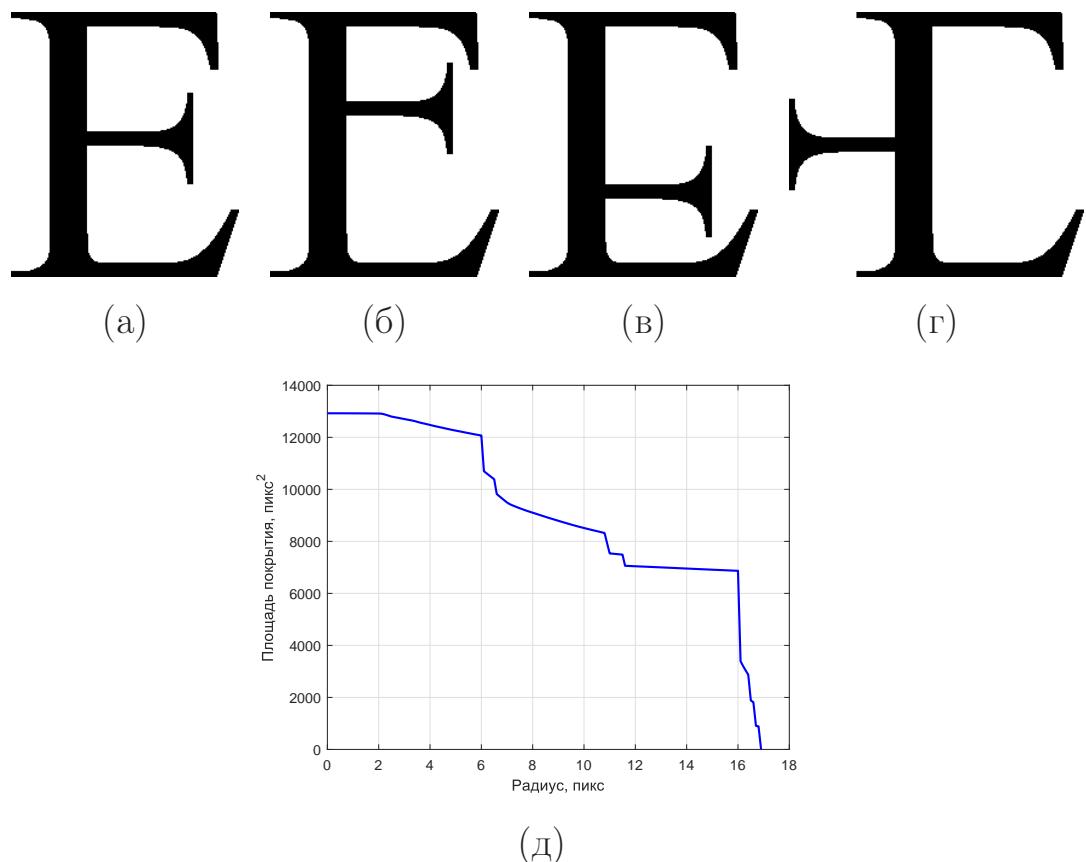


Рисунок 3.15 — Бинарные изображения с совпадающими функциями морфологической ширины. (а) Исходное изображение символа, (б)-(г) его искажения, (д) график функции ширины

Недостатком морфологической ширины можно считать её неспособность зафиксировать пространственную информацию, содержащуюся в изображении. В частности, если бинарное изображение содержит несколько объектов (связных компонент), то спектр никак не зависит от их взаимного расположения, если только компоненты остаются разделимыми. Более изощрённый пример преобразований, не влияющих на распределение ширины по фигуре, показан на рис. 3.15: здесь искажения символа имеют совершенно такие же функции морфологической ширины, что и сам исходный символ. Также в некоторых приложениях нечувствительность к повороту и гибким деформациям является нежелательным свойством процедуры анализа изображения. Таким образом, возникает необходимость разработки более гибкого и мощного инструмента анализа ширины объектов.

Чтобы исправить эти недостатки, были предложены несколько модификаций морфологического спектра. Обобщённый морфологический спектр [95] вычисляет не площадь (сумму интенсивностей в полутоновом случае) разности двух последовательных морфологических операций, а значение степенной

функции от координат этой разности. Метод анализа пространственного распределения размеров [96] производит сравнение геометрических ковариограмм в случае бинарных изображений или, в случае полутоновых изображений, анализ функции автокорреляции между исходным изображением и его гранулометрической трансформацией. Метод многомерного анализа связности [97] симулирует изменения смежности объектов на изображении при изменении масштаба. При вычислении спектра по размеру и плотности [98] используется менее строгий вариант операции открытия, в котором достаточно, чтобы точками объекта был покрыт не весь примитив, а лишь некоторая его доля, определяемая значением плотности. Во всех этих модификациях традиционный морфологический спектр предстаёт частным случаем более сложной и развитой техники обработки изображений.

Стоит заметить, что обобщённый морфологический спектр тесно связан с понятием геометрических моментов, широко используемых в обработке и анализе изображений. Геометрические моменты используются во многих задачах анализа изображений в диапазоне от инвариантного распознавания образов [99] и сжатия изображений [100] до определения позы [101]. В применении к изображениям они описывают распределение яркости изображения по отношению к его осям. Моменты способны учитывать как глобальную, так и локальную геометрическую информацию об изображении. Альтернативой классическим моментам могут служить моменты Цернике [75], в основе которых лежит ортогональная система полиномов, вследствие чего моменты являются функционально независимыми. Несмотря на то, что моменты могут быть использованы для восстановления изображений [102], признаки, полученные с помощью моментов, не являются достаточно информативными, так как момент заданной степени представляет собой лишь скалярное число.

Цель данного подраздела — совместить достоинства морфологического спектра и геометрических моментов и получить дескриптор, который отразит как пространственные, так и структурные особенности формы. Идея заключается в том, чтобы вычислять моменты не только для самих изображений, но и для результатов применения к ним морфологических операций. Далее будут представлены теоретические основания морфологических моментов, определены преобразования, обеспечивающие их инвариантность к сдвигу, масштабированию и повороту. Для класса многоугольных фигур будет разработан точный аналитический алгоритм вычисления морфологических моментов на

базе модели дискового покрытия. Наконец, в заключительной главе мы оценим временные затраты на вычисление моментов, а также покажем их продуктивность в задаче распознавания шрифтов.

3.4.2 Концепция морфологических моментов

Простейшие геометрические свойства изображения, такие как площадь, положение и ориентация, могут быть легко вычислены с помощью набора линейных функционалов изображения, известных под названием геометрических моментов. Будем считать, что изображение представляется в виде скалярного поля $f : \Omega \in \mathbb{R}^2 \rightarrow \mathbb{R}$, определённого на компактном множестве Ω (плоскости изображения) и определяющего интенсивность изображения в пространственной позиции $(x,y) \in \Omega$. Потребуем также ограниченности поля f : $\underline{f} \leq f(x,y) \leq \bar{f}$.

Определение 3.5. Величина

$$m_{pq} = \iint_{\Omega} x^p y^q f(x,y) dx dy.$$

называется моментом порядков (p,q) изображения f .

Так как изначально аналоговая функция изображения $f(x,y)$ определена на бесконечном множестве, в том случае, если она не допускает явного аналитического представления, она может быть оцифрована и сохранена в памяти компьютера в дискретном варианте $f(x_i, y_j)$ с массивом пикселей размера $W \times H$. Тогда двойное интегрирование в предыдущей формуле, как правило, заменяют на суммирование, в котором (x_i, y_i) — центральная точка пикселя (i,j) :

$$\tilde{m}_{pq} = \Delta^2 \sum_{i=1}^W \sum_{j=1}^H x_i^p y_j^q f(x_i, y_j), \quad (3.18)$$

где $\Delta = x_i - x_{i-1} = y_j - y_{j-i}$ — интервал семплирования. В частности, при $\Delta = 1$ именно эта формула определяет геометрический момент растрового изображения. Для быстрого расчёта \tilde{m}_{pq} был предложен ряд как программных [103; 104], так и аппаратных [105] реализаций. Очевидно, однако, что так как \tilde{m}_{pq} по существу задаёт результат приближённого вычисления m_{pq} методом средних прямоугольников, это не слишком точная оценка m_{pq} , особенно когда порядок момента (p,q) возрастает.

Далее нас будет особо интересовать случай бинарных входных данных в виде некоторой двумерной функции, которая принимает значения 0 или 1 в каждой точке. Примером такой функции является бинарное изображение. Его непрерывным аналогом является сцена, описывающая набор фигур, представляющих собой замкнутые области на плоскости, ограниченные конечным числом непересекающихся замкнутых жордановых кривых. Эта сцена задаётся функцией, которая принимает значение 1, если точка принадлежит некоторой фигуре, в противном случае функция принимает значение 0. Момент фигуры X в непрерывном случае задаётся следующим уравнением:

$$m_{pq}(X) = \iint_{(x,y) \in X} x^p y^q dx dy.$$

Напомним, что площадь фигуры X определяется по формуле:

$$S(X) = \iint_{(x,y) \in X} 1 dx dy.$$

В таком представлении морфологическая ширина фигуры X принимает вид (напомним, что $X(r)$ обозначает дисковое r -покрытие фигуры X):

$$\mathfrak{F}(r) = \iint_{(x,y) \in X(r)} 1 dx dy,$$

и её значение при аргументе, равном r , можно рассматривать как момент порядка $(0,0)$ фигуры $X(r)$. С другой стороны, моменты фигуры X ненулевых порядков можно считать моментами 0-покрытия фигуры X . Возникает естественная идея совместить наборы аргументов и построить функцию, которая зависела бы и от порядков (p,q) , и от радиуса покрытия r .

Определение 3.6. *Морфологическим моментом порядков (p,q) фигуры X считается функция, описывающая зависимость момента, вычисленного по открытию фигуры кругом радиуса r (по дисковому r -покрытию), от размера примитива:*

$$\mathfrak{M}_{pq}(r) = m_{pq}(X(r)). \quad (3.19)$$

В таком случае «классические» моменты можно определить как $\mathfrak{M}_{pq}(0)$, а морфологический спектр следующим образом:

$$PS(r) = -\partial \mathfrak{M}_{00}(X(r)) / \partial r, \quad r \geq 0. \quad (3.20)$$

Так как дисковое r -покрытие можно получить с помощью операции открытия с диском радиуса r , понятие морфологического момента можно адаптировать для дискретного случая — с дискретной фигурой $X \subset \mathbb{Z}^2$ и семейством дискретных структурирующих элементов $B(r) \subset \mathbb{Z}^2$, $r \in \mathbb{N}_0$:

$$\widetilde{\mathfrak{M}}_{pq}(r) = \sum_{(x,y) \in X \circ B(r)} x^p y^q.$$

Однако в силу невысокой вычислительной эффективности морфологических операций более полезной представляется формулировка с помощью карты толщин $G(x,y)$ (напомним, что это функция, которая каждой точке фигуры ставит в соответствие радиус максимального вписанного в фигуру круга, покрывающего точку):

$$\widetilde{\mathfrak{M}}_{pq}(r) = \sum_{(x,y) \in X : G(x,y) \geq r} x^p y^q. \quad (3.21)$$

Отметим, что выражение

$$\mathfrak{G}_{pq}(r) = \widetilde{\mathfrak{M}}_{pq}(r) - \widetilde{\mathfrak{M}}_{pq}(r+1) \quad (3.22)$$

в дискретном случае даст нам обобщённый морфологический спектр порядков (p,q) с дисковым структурирующим элементом, предложенный в работе [95]. Тем не менее задачи исследования свойств инвариантности морфологических моментов и разработки эффективного алгоритма их вычисления не были исследованы в должной мере.

3.4.3 Инварианты морфологических моментов

Представляет значительный интерес вопрос о нормировке морфологических моментов с целью обеспечить их инвариантность к параллельному переносу, масштабированию и повороту. Как показано в работе [106], для классических моментов инвариантность к параллельному переносу достигается следующим образом:

$$\mu_{pq} = \iint_{(x,y) \in X} (x - \bar{x})(y - \bar{y}) dx dy, \quad \bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}, \quad (3.23)$$

а к масштабированию — в результате преобразования

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\left(\frac{p+q}{2}\right)}}. \quad (3.24)$$

Заметим, что центральный классический момент порядков $(1,0)$ и $(0,1)$ всегда равен нулю, а инвариант масштаба нулевого порядка равен 1, так как в силу требования инвариантности положение и площадь объекта не должны иметь значения. Морфологические моменты добавляют зависимость функционала от радиуса r , что вносит свои особенности. В задачах, где требуется инвариантность к сдвигу, не имеет смысла использовать информацию об абсолютном положении центра масс. Однако информация об *изменении* центра масс с ростом радиуса покрытия r может оказаться достаточно полезной. Поэтому мы будем производить нормировку относительно центра масс не для каждого r в отдельности, а единственный раз — для всей фигуры. В этом случае центр масс (\bar{x}, \bar{y}) представляется как

$$\bar{x} = \frac{\mathfrak{M}_{10}(0)}{\mathfrak{M}_{00}(0)}, \quad \bar{y} = \frac{\mathfrak{M}_{01}(0)}{\mathfrak{M}_{00}(0)}, \quad (3.25)$$

и производятся сдвиги на это значение аналогично случаю инвариантов X_y в уравнении 3.23:

$$\mu_{pq}(r) = \iint_{(x,y) \in X(r)} (x - \bar{x})^p (y - \bar{y})^q dx dy. \quad (3.26)$$

Раскроем в этой формуле скобки (здесь $\binom{n}{k}$ — биномиальный коэффициент, равный $\frac{n!}{k!(n-k)!}$):

$$\begin{aligned} \mu_{pq}(r) &= \iint_{(x,y) \in X(r)} \sum_{i=0}^p \binom{p}{i} (-1)^{p-i} x^i \bar{x}^{p-i} \sum_{j=0}^q \binom{q}{j} (-1)^{q-j} y^j \bar{y}^{q-j} dx dy = \\ &= \sum_{i=0}^p \sum_{j=0}^q (-1)^{p+q-i-j} \iint_{(x,y) \in X} \binom{p}{i} \binom{q}{j} \bar{x}^{p-i} \bar{y}^{q-j} x^i y^j dx dy. \end{aligned} \quad (3.27)$$

С учётом того, что $\iint_{(x,y) \in X(r)} x^i y^j dx dy = \mathfrak{M}_{ij}(r)$, а \bar{x} и \bar{y} — константы, при начальных значениях p и q получаем следующие варианты:

$$\begin{aligned}
\mu_{00}(r) &= \mathfrak{M}_{00}(r), \\
\mu_{10}(r) &= \mathfrak{M}_{10}(r) - \bar{x}\mathfrak{M}_{00}(r), \\
\mu_{01}(r) &= \mathfrak{M}_{01}(r) - \bar{y}\mathfrak{M}_{00}(r), \\
\mu_{20}(r) &= \mathfrak{M}_{20}(r) - 2\bar{x}\mathfrak{M}_{10}(r) + \bar{x}^2\mathfrak{M}_{00}(r), \\
\mu_{11}(r) &= \mathfrak{M}_{11}(r) - \bar{x}\mathfrak{M}_{01}(r) - \bar{y}\mathfrak{M}_{10}(r) + \bar{x}\bar{y}\mathfrak{M}_{00}(r), \\
\mu_{02}(r) &= \mathfrak{M}_{02}(r) - 2\bar{y}\mathfrak{M}_{01}(r) + \bar{y}^2\mathfrak{M}_{00}(r), \\
\mu_{30}(r) &= \mathfrak{M}_{30}(r) - 3\bar{x}\mathfrak{M}_{20}(r) + 3\bar{x}^2\mathfrak{M}_{10}(r) - \bar{x}^3\mathfrak{M}_{00}(r), \\
\mu_{21}(r) &= \mathfrak{M}_{21}(r) - 2\bar{x}\mathfrak{M}_{11}(r) + \bar{x}^2\mathfrak{M}_{01}(r) - \bar{y}\mathfrak{M}_{20}(r) + 2\bar{x}\bar{y}\mathfrak{M}_{10}(r) - \bar{x}^2\bar{y}\mathfrak{M}_{00}(r), \\
\mu_{12}(r) &= \mathfrak{M}_{12}(r) - 2\bar{y}\mathfrak{M}_{11}(r) + \bar{y}^2\mathfrak{M}_{10}(r) - \bar{x}\mathfrak{M}_{02}(r) + 2\bar{x}\bar{y}\mathfrak{M}_{01}(r) - \bar{x}\bar{y}^2\mathfrak{M}_{00}(r), \\
\mu_{03}(r) &= \mathfrak{M}_{03}(r) - 3\bar{y}\mathfrak{M}_{02}(r) + 3\bar{y}^2\mathfrak{M}_{01}(r) - \bar{y}^3\mathfrak{M}_{00}(r).
\end{aligned} \tag{3.28}$$

Инварианты η_{pq} , устойчивые как к параллельному переносу, так и к изменению масштаба, могут быть получены из центральных моментов нормировкой площади фигуры к 1. Это достигается сжатием фигуры в $\lambda = \sqrt{\mu_{00}(0)}$ раз по каждой оси, в результате чего мы получаем моменты фигуры X' , которая не повёрнута, но масштабирована таким образом, что площадь её равна 1, а центр масс расположен в начале координат. Заметим также, что дисковое r -покрытие фигуры X' образуется сжатием в λ раз λr -покрытия фигуры X . В итоге имеем:

$$\begin{aligned}
\eta_{pq}(r') &= \iint_{(x',y') \in X'(r')} (x' - \bar{x}')^p (y' - \bar{y}')^q dx' dy' = \\
&= \left\{ x' = \frac{x}{\sqrt{\mu_{00}(0)}}, y' = \frac{y}{\sqrt{\mu_{00}(0)}}, r' = \frac{r}{\sqrt{\mu_{00}(0)}} \right\} = \\
&= \iint_{(x,y) \in X(r' \sqrt{\mu_{00}(0)})} \left(\frac{x - \bar{x}}{\sqrt{\mu_{00}(0)}} \right)^p \left(\frac{y - \bar{y}}{\sqrt{\mu_{00}(0)}} \right)^q \frac{1}{\sqrt{\mu_{00}(0)}} dx \frac{1}{\sqrt{\mu_{00}(0)}} dy = \\
&= \frac{\mu_{pq}(r' \sqrt{\mu_{00}(0)})}{(\mu_{00}(0))^{1+\frac{p+q}{2}}} \cdot
\end{aligned} \tag{3.29}$$

Инвариантность к повороту достигается в точности тем же способом, что и в случае инвариантов Ху, с той лишь разницей, что она должна быть проведена

для каждого r в отдельности. Напомним вид исходных инвариантов:

$$\begin{aligned}
 I_1 &= \eta_{20} + \eta_{02}, \\
 I_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\
 I_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2, \\
 I_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\
 I_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
 &\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\
 I_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}), \\
 I_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\
 &\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].
 \end{aligned} \tag{3.30}$$

В качестве доказательства можно заметить, что поворот фигуры X на угол α означает переход к новым координатам:

$$x' = x \cos \alpha - y \sin \alpha, \quad y' = x \sin \alpha + y \cos \alpha,$$

но в новых координатах формулы не меняются. Для начала отметим, что

$$\bar{x}' = \bar{x} \cos \alpha - \bar{y} \sin \alpha, \quad \bar{y}' = \bar{x} \sin \alpha + \bar{y} \cos \alpha.$$

Рассмотрим первый инвариант $I_1(r) = \eta_{20}(r) + \eta_{02}(r)$, вычисленный для фигуры X :

$$I_1(r) = \frac{1}{(\mu_{00}(0))^{1+\frac{p+q}{2}}} \iint_{(x,y) \in X(r\sqrt{\mu_{00}(0)})} ((x - \bar{x})^2 + (y - \bar{y})^2) dx dy,$$

и аналогичное выражение для фигуры X' , полученной в результате поворота:

$$I'_1(r) = \frac{1}{(\mu_{00}(0))^{1+\frac{p+q}{2}}} \iint_{(x',y') \in X'(r\sqrt{\mu_{00}(0)})} f(x',y') dx' dy'.$$

Здесь

$$\begin{aligned}
 f(x',y') &= (x' - \bar{x}')^2 + (y' - \bar{y}')^2 = \\
 &= ((x - \bar{x}) \cos \alpha - (y - \bar{y}) \sin \alpha)^2 + ((x - \bar{x}) \sin \alpha + (y - \bar{y}) \cos \alpha)^2 = \\
 &= (x - \bar{x})^2 \cos^2 \alpha - 2(x - \bar{x})(y - \bar{y}) \sin \alpha \cos \alpha + (y - \bar{y})^2 \sin^2 \alpha + \\
 &\quad + (x - \bar{x})^2 \sin^2 \alpha + 2(x - \bar{x})(y - \bar{y}) \sin \alpha \cos \alpha + (y - \bar{y})^2 \cos^2 \alpha = \\
 &= (x - \bar{x})^2 + (y - \bar{y})^2,
 \end{aligned} \tag{3.31}$$

а $dx' dy' = (\cos^2 \alpha + \sin^2 \alpha) dx dy = dx dy$. В итоге мы получили под интегралом ту же функцию, что и для неповёрнутой фигуры, следовательно $I'_1(r) = I_1(r)$. Аналогичную проверку можно провести и для остальных инвариантов поворота.

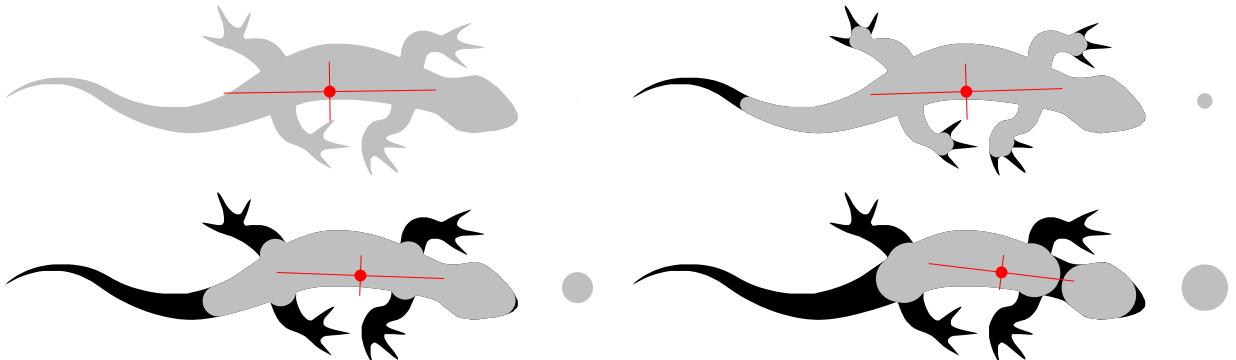


Рисунок 3.16 — Центры масс и оси инерции дисковых покрытий фигуры «ящерица» (справа изображены образцы покрывающих дисков)

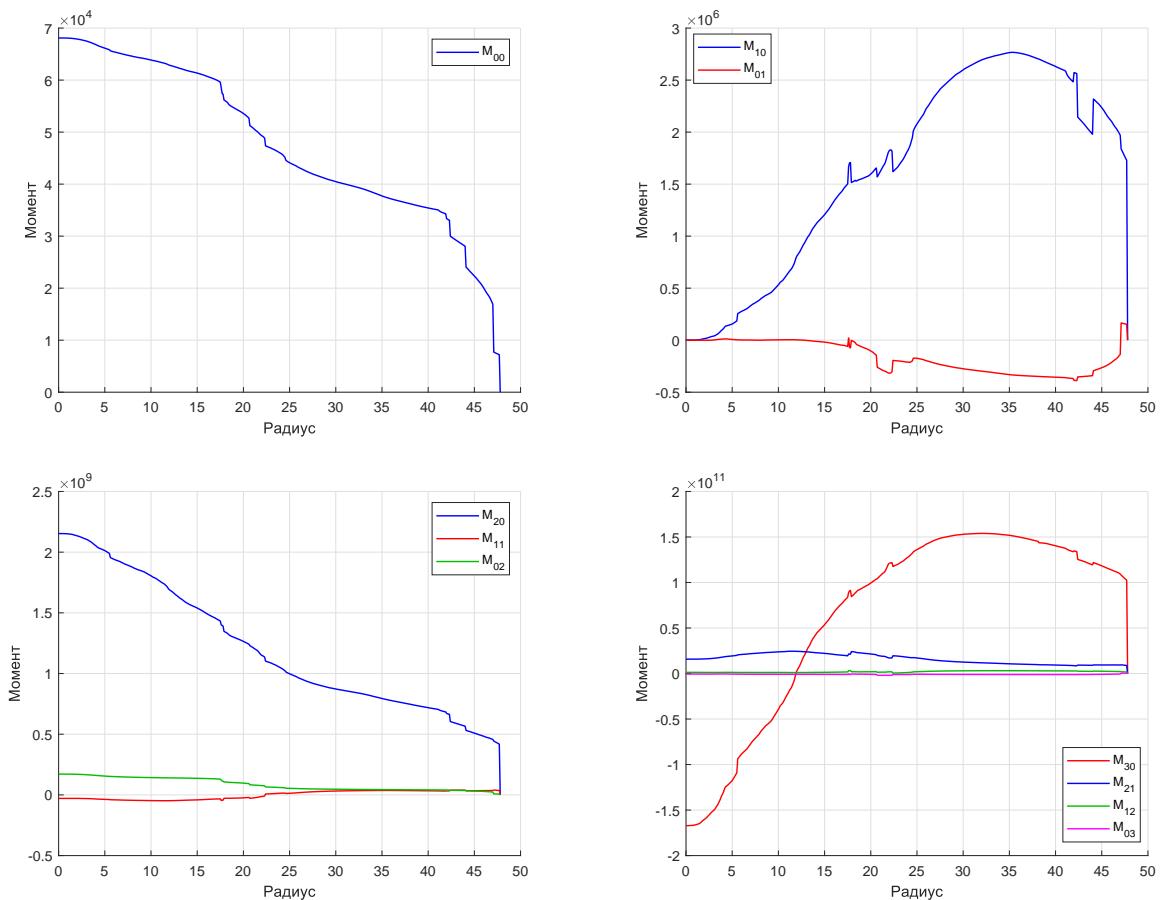


Рисунок 3.17 — Графики центральных морфологических моментов при $p + q \leq 3$ для фигуры «ящерица»

Иллюстрации введённых понятий показаны на рис. 3.16 и 3.17.

Так как все инварианты получаются из начальных моментов $\mathfrak{M}_{pq}(r)$, возникает задача их вычисления при заданных значениях r .

Задача 3.4. *Дана многоугольная фигура X , порядки p, q и набор значений радиуса $\{r_i\}_{i=1}^n$, $r_i \geq 0$ и $r_i < r_j$ при $i < j$. Вычислить $\mathfrak{M}_{pq}(r_i)$, $i = 1, \dots, n$.*

Наивный подход предполагает получение результата морфологической операции $X \circ B(r)$ явно в виде бинарного изображения. Хотя алгоритмы вычисления базовых морфологических фильтров хорошо разработаны, сама необходимость использовать дискретную модель не позволяет добиться алгоритма, который мог бы быть использован в системах реального времени. Кроме того, можно снова обратиться к модели на основе карты толщин, рассмотренной в предыдущем разделе, и вычислить моменты по формуле 3.21. Хотя сама карта вычисляется на основе непрерывного скелета, при таком подходе вновь рассматривается дискретное множество точек, расположенных в узлах целочисленной сетки. Однако возможен и полностью непрерывный подход к вычислению морфологических моментов, о котором мы расскажем в следующем подразделе.

3.4.4 Непрерывный алгоритм вычисления моментов

В первой главе работы показано, что дисковое r -покрытие совпадает с открытием фигуры диском радиуса r . Это значит, что для вычисления морфологического момента порядков (p,q) в точке r необходимо взять интеграл от функции $f(x,y) = x^p y^q$ по дисковому r -покрытию. Напомним, что представленный выше алгоритм подразумевает разбиение дискового покрытия многоугольной фигуры на области следующих типов:

1. Собственные области бициклов — полных и усечённых.
2. Внешние сектора меньших концевых кругов усечённых бициклов.
3. Линзы в пересечении внешних секторов.

Площадь покрытия складывается из площадей областей первых двух типов за вычетом площадей областей третьего типа. Благодаря свойству аддитивности определённого интеграла тот же подход может быть использован для вычисления не площадей, а моментов: достаточно лишь заменить для каждой области операцию вычисления площади на операцию вычисления момента.

Как показано ранее, собственная область любого типа представляет собой многоугольник, каждая вершина которого — это либо конец оси бицикла,

либо его проекция на образующий сайт (в случае сайта точки проекция совпадает с самой точкой). Все операции для вычисления точных координат вершин соответствующего многоугольника рассмотрены в разделе 3.1.

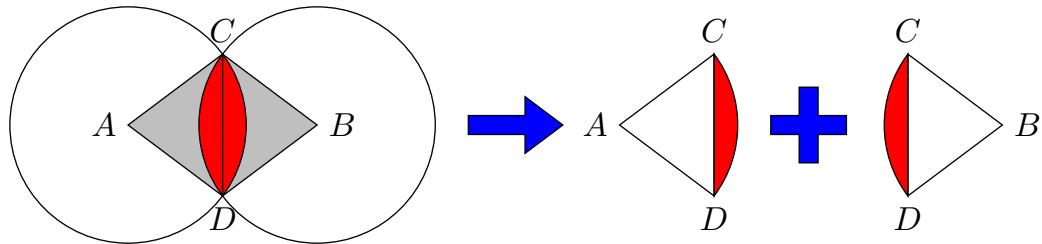


Рисунок 3.18 — Представление линзы через треугольники и круговые сектора

Области второго типа представляют собой круговые сектора. Заметим, что нас интересует не просто угловой диапазон сектора, а конкретные начальное и конечное значения угла. Их можно определить по координатам проекции центра меньшего концевого круга на образующие сайты. Наконец, область третьего типа можно разбить на две части, каждая из которых представляет собой дополнение треугольника до кругового сектора (рис. 3.18), так что момент каждой части получается как разность моментов сектора и треугольника.

Таким образом, проблема может быть сведена к вычислению морфологических моментов областей двух типов: многоугольников и круговых секторов.

3.4.5 Морфологические моменты примитивных областей

Точные формулы для расчёта произвольных моментов многоугольников выведены в работах [107] и [108]. Так, общая формула ненормированных моментов многоугольника с вершинами $\{(x_i, y_i)\}_{i=1}^n$ имеет вид:

$$m_{pq} = \frac{1}{(p+q+2)(p+q+1)\binom{p+q}{p}} \times \\ \times \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1}) \sum_{k=0}^p \sum_{l=0}^q \binom{k+l}{l} \binom{p+q-k-l}{q-l} x_i^k x_{i-1}^{p-k} y_i^l y_{i-1}^{q-l}. \quad (3.32)$$

Для вычисления начальных моментов при малых p и q удобно предварительно произвести преобразование формулы:

$$\begin{aligned}
 m_{00} &= \frac{1}{2} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1}), \\
 m_{10} &= \frac{1}{6m_{00}} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1})(x_{i-1} + x_i), \\
 m_{01} &= \frac{1}{6m_{00}} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1})(y_{i-1} + y_i), \\
 m_{20} &= \frac{1}{12m_{00}} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1})(x_{i-1}^2 + x_{i-1}x_i + x_i^2), \\
 m_{11} &= \frac{1}{24m_{00}} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1})(2x_{i-1}y_{i-1} + x_{i-1}y_i + x_iy_{i-1} + 2x_iy_i), \\
 m_{02} &= \frac{1}{12m_{00}} \sum_{i=1}^n (x_{i-1}y_i - x_iy_{i-1})(y_{i-1}^2 + y_{i-1}y_i + y_i^2).
 \end{aligned} \tag{3.33}$$

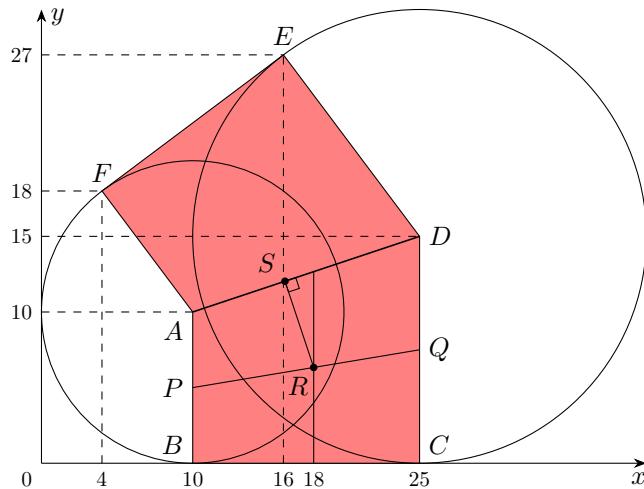


Рисунок 3.19 — Вычисление момента собственной области линейного бицикла

Проверим справедливость этих формул на конкретном примере. Рассмотрим шестиугольник с координатами $A(10; 10)$, $B(10; 0)$, $C(25; 0)$, $D(25; 15)$, $E(16; 27)$ и $F(4; 18)$, который может служить собственной областью для линейного бицикла с концами ребра в точках A и B и радиусами концевых кругов 10 и 15 (рис. 3.19). В этом случае применение формулы для площади много-

угольника даёт нам:

$$\begin{aligned} 2m_{00} = & (4 \cdot 10 - 10 \cdot 18) + (10 \cdot 0 - 10 \cdot 10) + (10 \cdot 0 - 25 \cdot 0) + \\ & + (25 \cdot 15 - 25 \cdot 0) + (25 \cdot 27 - 16 \cdot 15) + (16 \cdot 18 - 4 \cdot 27) = \\ = & -140 - 100 + 0 + 375 + 435 + 180 = 750. \end{aligned}$$

Таким образом, площадь многоугольника $ABCDEF$ равна 375, что в силу симметрии совпадает с результатом её вычисления как удвоенной площади трапеции $ABCD$ с основаниями AB и CD и высотой BC :

$$m_{00} = 2S_{ABCD} = (AB + CD) \cdot BC = (10 + 15) \cdot 15 = 375.$$

Теперь вычислим с помощью формул 3.33 центр масс многоугольника:

$$\begin{aligned} m_{10} = & \frac{1}{6 \cdot 375} (-140 \cdot (4 + 10) - 100 \cdot (10 + 10) + 0 \cdot (10 + 25)) + \\ & + \frac{1}{6 \cdot 375} (375 \cdot (25 + 25) + 435 \cdot (25 + 16) + 180 \cdot (16 + 4)) = \\ = & \frac{1}{2250} (-1960 - 2000 + 0 + 18750 + 17835 + 3600) = \frac{161}{10}, \\ m_{01} = & \frac{1}{6 \cdot 375} (-140 \cdot (18 + 10) - 100 \cdot (10 + 0) + 0 \cdot (0 + 0)) + \\ & + \frac{1}{6 \cdot 375} (375 \cdot (0 + 15) + 435 \cdot (15 + 27) + 180 \cdot (27 + 18)) = \\ = & \frac{1}{2250} (-3920 - 1000 + 0 + 5625 + 18270 + 8100) = \frac{361}{30}. \end{aligned}$$

Известно, что в прямоугольной трапеции расстояние от центра масс до меньшего из оснований равно $h \frac{2a+b}{3(a+b)}$, где a и b — длины большего и меньшего основания соответственно, а h — высота трапеции. В случае трапеции $ABCD$ имеем $a = BC = 15$, $b = AB = 10$, $h = BC = 15$, поэтому абсцисса центра масс, точки R , равна $10 + 15 \cdot \frac{2 \cdot 15 + 10}{3 \cdot (10 + 15)} = 18$. Кроме того, точка R лежит на отрезке, соединяющем центры оснований: точку P — середину отрезка AB с координатами $(10; 5)$ и точку Q — середину отрезка CD с координатами $(25; 7,5)$. Таким образом,

$$y_R = 5 + \frac{(18 - 10)}{15} \cdot \frac{5}{2} = \frac{19}{3}.$$

Остаётся заметить, что трапеции $ABCD$ и $ADEF$ симметричны относительно прямой AD , следовательно, отношение симметричности сохраняется и

для центра масс трапеций. Поэтому центр масс шестиугольника S равен середине отрезка, соединяющего центры масс трапеций, то есть проекции каждого из них на отрезок AD :

$$\begin{aligned} S &= A + t \cdot (D - A), \\ t &= \frac{(x_D - x_A)(x_R - x_A) + (y_B - y_A)(y_R - y_A)}{(x_D - x_A)^2 + (y_D - y_A)^2} = \\ &= \frac{(25 - 10)(18 - 10) + (15 - 10)(\frac{19}{3} - 10)}{(25 - 10)^2 + (15 - 10)^2} = \frac{61}{150}, \\ x_S &= 10 + \frac{61}{150}(25 - 10) = \frac{161}{10}, \\ y_S &= 10 + \frac{61}{150}(15 - 10) = \frac{361}{30}. \end{aligned}$$

Таким образом, в данном случае формулы 3.33 подтверждаются более простыми геометрическими рассуждениями. Отметим, что сложность алгоритма вычисления моментов по формуле 3.33 составит $O(npq)$, если вычислить величины $x_i^k, y_i^k, i = 1, \dots, n, k = 0, \dots, \max(p, q)$ предварительно.

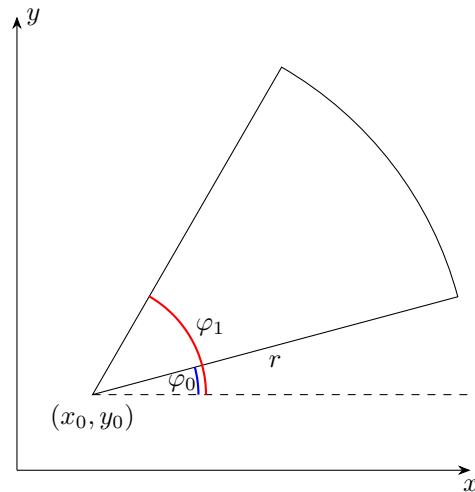


Рисунок 3.20 — Круговой сектор

Интеграл по круговому сектору S функции $f(x,y) = x^p y^q$ легче рассматривать в полярных координатах. В этом случае интеграл можно переписать как

$$m_{pq}(S) = \int_0^{R} \int_{\varphi_0}^{\varphi_1} (x_0 + r \cos \varphi)^p (y_0 + r \sin \varphi)^q r dr d\varphi.$$

Раскрыв скобки, мы можем преобразовать интегрируемую функцию к виду:

$$f(r, \varphi) = \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} x_0^{p-i} y_0^{q-j} r^{i+j+1} \cos^i \varphi \sin^j \varphi.$$

Поэтому

$$\begin{aligned}
 m_{pq}(S) &= \int_0^R \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} x_0^{p-i} y_0^{q-j} r^{i+j+1} dr \int_{\varphi_0}^{\varphi_1} \cos^i \varphi \sin^j \varphi d\varphi = \\
 &= \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} x_0^{p-i} y_0^{q-j} \frac{R^{i+j+2}}{i+j+2} \int_{\varphi_0}^{\varphi_1} \cos^i \varphi \sin^j \varphi d\varphi = \\
 &= \sum_{i=0}^p \sum_{j=0}^q \binom{p}{i} \binom{q}{j} x_0^{p-i} y_0^{q-j} \frac{R^{i+j+2}}{i+j+2} I_{ij}(\varphi_0, \varphi_1),
 \end{aligned} \tag{3.34}$$

где $I_{ij}(\varphi_0, \varphi_1) = \int_{\varphi_0}^{\varphi_1} \cos^i \varphi \sin^j \varphi d\varphi$.

Воспользуемся тригонометрическими формулами понижения степени в общем виде:

$$\begin{aligned}
 \cos^i \varphi &= \begin{cases} \frac{\binom{i}{\frac{i}{2}}}{2^i} + \frac{1}{2^{i-1}} \sum_{k=0}^{\frac{i}{2}-1} \binom{i}{k} \cos((i-2k)\varphi) & \text{при чётном } i, \\ \frac{1}{2^{i-1}} \sum_{k=0}^{\frac{i-1}{2}} \binom{i}{k} \cos((i-2k)\varphi) & \text{при нечётном } i; \end{cases} \\
 \sin^j \varphi &= \begin{cases} \frac{\binom{j}{\frac{j}{2}}}{2^j} + \frac{1}{2^{j-1}} \sum_{t=0}^{\frac{j}{2}-1} (-1)^{\frac{j}{2}-t} \binom{j}{t} \cos((j-2t)\varphi) & \text{при чётном } j, \\ \frac{1}{2^{j-1}} \sum_{t=0}^{\frac{j-1}{2}} (-1)^{\frac{j-1}{2}-t} \binom{j}{t} \sin((j-2t)\varphi) & \text{при нечётном } j, \end{cases}
 \end{aligned}$$

что можно представить как

$$\begin{aligned}
 \cos^i \varphi &= \text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^i} + \frac{1}{2^{i-1}} \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \binom{i}{k} \cos((i-2k)\varphi), \\
 \sin^j \varphi &= \text{mod}(j+1, 2) \frac{\binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^j} + \\
 &\quad + \frac{1}{2^{j-1}} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} (-1)^{\lfloor \frac{j}{2} \rfloor-t} \binom{j}{t} \cos \left(\text{mod}(j, 2) \frac{\pi}{2} + (-1)^j (j-2t)\varphi \right),
 \end{aligned}$$

где $\text{mod}(a, b) = a - b \lfloor \frac{a}{b} \rfloor$.

Таким образом, так как $\cos(\alpha) \cdot \cos(\beta) = \frac{1}{2} (\cos(\alpha + \beta) + \cos(\alpha - \beta))$,

$$\cos^i \varphi \sin^j \varphi =$$

$$\begin{aligned}
&= \left[\text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^i} + \frac{1}{2^{i-1}} \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \binom{i}{k} \cos((i-2k)\varphi) \right] \times \\
&\times \left[\text{mod}(j+1, 2) \frac{\binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^j} + \frac{1}{2^{j-1}} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} (-1)^{\lfloor \frac{j}{2} \rfloor - t} \binom{j}{t} \cos \left(\text{mod}(j, 2) \frac{\pi}{2} + (-1)^j (j-2t)\varphi \right) \right] = \\
&= \left(\text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^i} \right) \left(\text{mod}(j+1, 2) \frac{\binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^j} \right) + \\
&+ \left(\text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^i} \right) \frac{1}{2^{j-1}} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} (-1)^{\lfloor \frac{j}{2} \rfloor - t} \binom{j}{t} \cos \left(\text{mod}(j, 2) \frac{\pi}{2} + (-1)^j (j-2t)\varphi \right) + \\
&+ \left(\text{mod}(j+1, 2) \frac{\binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^j} \right) \frac{1}{2^{i-1}} \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \binom{i}{k} \cos((i-2k)\varphi) + \\
&+ \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \cos \left(\text{mod}(j, 2) \frac{\pi}{2} + ((-1)^j (j-2t) + i-2k)\varphi \right) + \\
&+ \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \cos \left(\text{mod}(j, 2) \frac{\pi}{2} + ((-1)^j (j-2t) - i+2k)\varphi \right).
\end{aligned}$$

Интегрируя это выражение, получаем

$$\begin{aligned}
 I_{ij}(\varphi_0, \varphi_1) &= \int_{\varphi_0}^{\varphi_1} \cos^i \varphi \sin^j \varphi d\varphi = \\
 &= \text{mod}(i+1, 2) \cdot \text{mod}(j+1, 2) \cdot \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor} \binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^{i+j-1}} (\varphi_1 - \varphi_0) + \\
 &+ \left(\text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^{i+j-1}} \right) \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} (-1)^{\lfloor \frac{j}{2} \rfloor - t} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + (-1)^j(j-2t)\varphi_1)}{(-1)^j(j-2t)} + \\
 &- \left(\text{mod}(i+1, 2) \frac{\binom{i}{\lfloor \frac{i}{2} \rfloor}}{2^{i+j-1}} \right) \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} (-1)^{\lfloor \frac{j}{2} \rfloor - t} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + (-1)^j(j-2t)\varphi_0)}{(-1)^j(j-2t)} + \\
 &+ \left(\text{mod}(j+1, 2) \frac{\binom{j}{\lfloor \frac{j}{2} \rfloor}}{2^{i+j-1}} \right) \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \binom{i}{k} \frac{\sin((i-2k)\varphi_1) - \sin((i-2k)\varphi_0)}{i-2k} + \\
 &+ \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + ((-1)^j(j-2t) + i-2k)\varphi_1)}{(-1)^j(j-2t) + i-2k} + \\
 &- \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + ((-1)^j(j-2t) + i-2k)\varphi_0)}{(-1)^j(j-2t) + i-2k} + \\
 &+ \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + ((-1)^j(j-2t) - i+2k)\varphi_1)}{(-1)^j(j-2t) - i+2k} + \\
 &- \sum_{k=0}^{\lfloor \frac{i-1}{2} \rfloor} \sum_{t=0}^{\lfloor \frac{j-1}{2} \rfloor} \frac{(-1)^{\lfloor \frac{j}{2} \rfloor - t}}{2^{i+j-1}} \binom{i}{k} \binom{j}{t} \frac{\sin(\text{mod}(j, 2)\frac{\pi}{2} + ((-1)^j(j-2t) - i+2k)\varphi_0)}{(-1)^j(j-2t) - i+2k},
 \end{aligned}$$

что позволяет вычислить значение момента сектора по формуле 3.34. Оценим вычислительную сложность алгоритма, использующего данную формулу. Вычисление I_{ij} требует $\Theta(i)\Theta(j)$ операций, так как формула содержит вложенные циклы длины $\lfloor \frac{i+1}{2} \rfloor$ и $\lfloor \frac{j+1}{2} \rfloor$. Вычисление момента сектора порядков (p, q) требует вычисления значений I_{ij} при $0 \leq i \leq p$ и $0 \leq j \leq q$. Следовательно, общая сложность составляет $\sum_{i=0}^p \sum_{j=0}^q \Theta(i)\Theta(j) = \sum_{i=0}^p \Theta(i) \sum_{j=0}^q \Theta(j) = \Theta(i^2 j^2)$.

В результате момент дискового покрытия многоугольной фигуры может быть вычислен как сумма моментов собственных областей бициклов — полных и усечённых — плюс сумма моментов внешних секторов меньших концевых кругов усечённых бициклов за вычетом суммы моментов линз в пересечении

смежных секторов. Все необходимые моменты вычисляются аналитически, что обеспечивает нам точное решение задачи 3.4

3.5 Выводы к главе 3

1. Представлен ряд задач морфологического анализа изображений, в которых возникает потребность получить информацию о ширине объектов, однако в силу особенностей задачи вычисление морфологической ширины не приводит к желаемому результату. Показано, что в этих случаях соответствие требованиям задачи достигается модификацией модели дискового покрытия или способа его измерения.
2. Предложен дескриптор, описывающий распределение ширины по сегменту фигуры и названный частной морфологической шириной, предназначенный для использования в задачах с выделенным окном интереса. Алгоритм вычисления дескриптора основан на расчёте площади пересечения каждой составной части дискового покрытия с окном интереса. Предложено обобщение алгоритма Уайлера — Атертона на случай криволинейных областей, границы которых включают дуги окружностей, позволяющее определить и описать фигуры в пересечении. Показано, что для снижения вычислительных операций можно заблаговременно убрать из скелета незначимые рёбра, области которых лежат целиком вне окна интереса.
3. Предложена модель системы вложенных дисковых покрытий, предназначенная для описания ширины полутонового изображения. Представлен способ перехода от дискретной полутоновой сцены к эквивалентной ей непрерывной, в результате которого каждый срез (результат отсечения по порогу) функции, задающей непрерывное изображение, представляет собой многоугольную сцену. Предложен дескриптор ширины полутонового изображения в виде функции, полученной усреднением функций морфологической ширины отдельных срезов.
4. Предложены способы задания локальных характеристик ширины в виде функций, определённых в каждой точке фигуры. Описан принцип построения мультимодальных дескрипторов формы, в которых ширину можно комбинировать с другими признаками, вычисленными в точке. Разработан алгоритм поиска максимального вписанного в фигу-

ру круга, покрывающего точку, для вычисления локальной ширины. Показано, что непрерывный скелет напрямую позволяет определить ближайшую точку границы фигуры, что обеспечивает вычисление дистанционного преобразования со скоростью, превосходящей существующие аналоги для широкого ряда фигур с достаточно плавной границей.

5. Разработан дескриптор геометрических моментов, позволяющий описать не просто площадь достаточно широкой части фигуры, а пространственное распределение этой части. Дескриптор совмещает принципы, лежащие в основе методов анализа плоских форм с помощью *морфологического спектра* и *геометрических моментов*. Показано, что инвариантность к сдвигу, масштабированию и повороту достигается в результате простейших преобразований ненормализованных моментов. Разработан алгоритм вычисления моментов, который строится на базе алгоритма вычисления морфологической ширины, отличие возникает лишь на этапе измерения каждой примитивной области в покрытии.
6. Все предложенные дескрипторы вычисляются аналитически по модели непрерывного скелета многоугольной фигуры, для алгоритмов их вычисления получена программная реализация.

Глава 4. Практические приложения алгоритмов вычисления признаков ширины

4.1 Распознавание компьютерных шрифтов

4.1.1 Описание задачи

Современная типографика использует тысячи различных шрифтовых гарнитур для дизайна и верстки газет, журналов, книг, объявлений, рекламных проспектов и другой печатной и медийной продукции. Единой классификации шрифтов нет, поэтому во многих случаях выбор шрифта для печати документов выполняется на основе образцов. При таком выборе часто возникает задача идентификации шрифта по образцу текста. Само происхождение образца может быть различным. Это может быть страничка текста на бумаге, изображение на экране, фотография вывески или объявления. Например, дизайнер хочет напечатать рекламу, используя шрифт, который он увидел на конфетной коробке в магазине. Задача состоит в том, чтобы определить название шрифта по имеющемуся изображению. В быстрой правильной идентификации шрифта заинтересованы дизайнеры, типографы, издатели для правильного подбора адекватного шрифта при производстве различной печатной продукции. В решении этой задачи также заинтересованы правообладатели шрифтов, поскольку они хотят обеспечить контроль лицензионного использования своих шрифтов.

Идентификация шрифта в общем виде ставится как задача поиска следующим образом. Имеется эталонный набор большого числа шрифтов, включающий несколько тысяч различных гарнитур. Такой набор называется файлом. Имеется также образец печатного текста, называемый запросом. Необходимо определить наиболее близкий к запросу шрифт в файле. Методов, предполагающих взаимодействие с обширной (более 1000 шрифтов) базой шрифтов не так много, некоторые из них основаны на формировании кодовой книги из векторов локальных признаков, в качестве которых может использоваться HOG [109] или SIFT [110]. Самым успешным оказался метод на основе глубинного обучения [111], внедрённый в редактор Adobe Photoshop в виде инструмента Match Font.

В рамках данного эксперимента мы рассматриваем подход к решению этой задачи путем построения меры сходства шрифта-запроса с эталонными

шрифтами и выбора «ближайшего соседа». Целью эксперимента является исследование меры сходства, основанной на морфологической ширине шрифта, рассмотренной в главе 2 и морфологических моментах, описанных в главе 3. Также значительный интерес представляет быстродействие метода, так как задача требует обработки внушительной базы изображений.

4.1.2 Сравнение шрифтов на основе диаграмм ширины

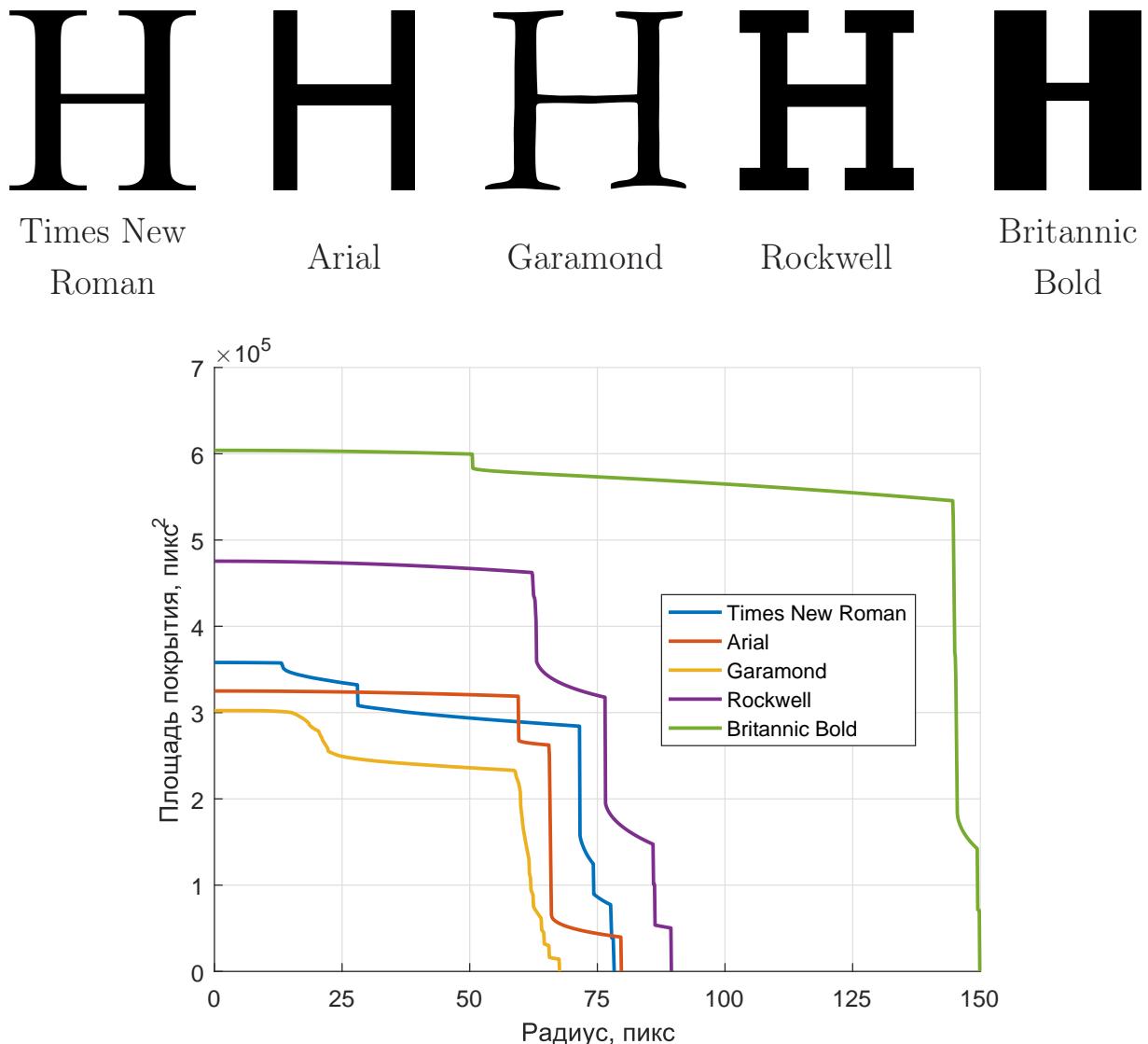


Рисунок 4.1 — Диаграммы морфологической ширины символа разных шрифтов

Проведём предварительную оценку возможностей использования морфологической ширины для решения указанных задач. Как показывает рис. 4.1, ширинные характеристики различных шрифтов при фиксированной высоте символов очень сильно отличаются друг от друга. Однако символы одного и

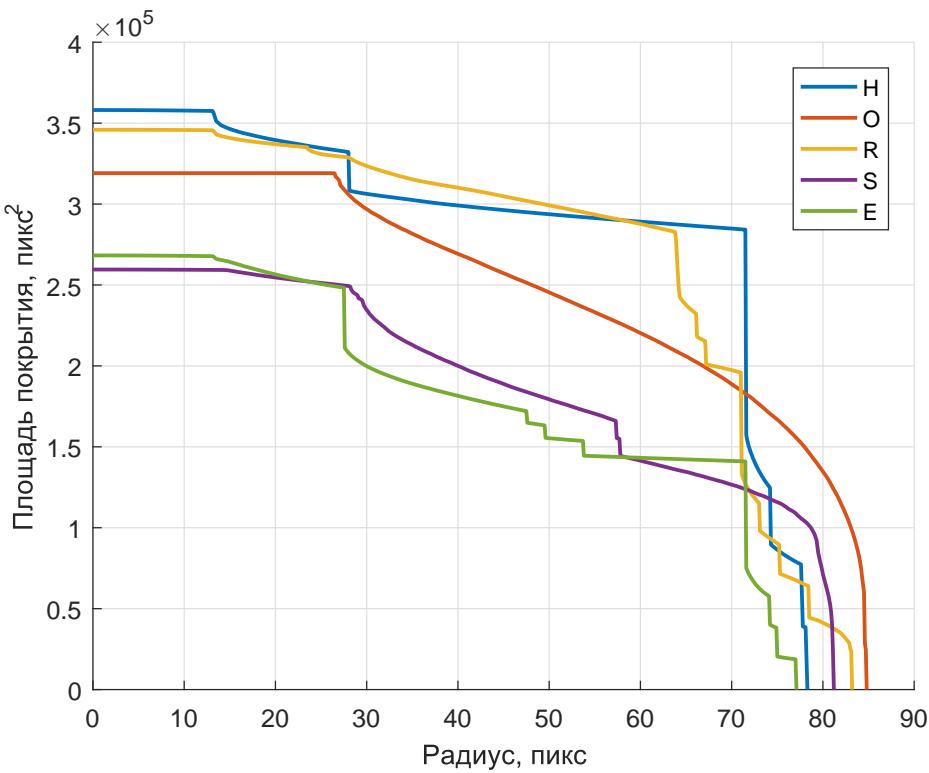


Рисунок 4.2 — Диаграммы морфологической ширины символов одного шрифта

того же шрифта также имеют индивидуальные хорошо различимые портреты (рис. 4.2). Это не позволяет составить совокупную характеристику ширины шрифта, не зависящую от конкретного символа. Вместо этого мы будем рассматривать в качестве признакового описания шрифта набор дескрипторов ширины его отдельных символов. В качестве таких символов можно взять буквы латинского алфавита, что обеспечит универсальность метода, так как латиницу поддерживает подавляющее большинство шрифтов.

Далее описывается логика работы системы идентификации шрифта. Она принимает на вход размеченное изображение $Q \in \{0, \dots, k\}^{W \times H}$, содержащее k символов. При $Q(x,y) = 0$ пиксель в позиции (x,y) считается принадлежащим к фону, при $Q(x,y) = i$, $1 \leq i \leq k$ — принадлежащим i -му символу. Можно также считать, что такая разметка определяет набор из k бинарных изображений $\{Q_i\}$, $1 \leq i \leq k$:

$$Q_i(x,y) = \begin{cases} 1, & \text{если } Q(x,y) = i, \\ 0, & \text{иначе.} \end{cases}$$

Совокупность объектов (символов) на входном изображении будем называть *надписью*.

Пусть также для каждого изображения Q_i проведено распознавание — известно, какая именно литера содержится на изображении. Будем считать, что эти литеры принадлежат пронумерованному ограниченному алфавиту и *код символа* c_i определяет номер в этом алфавите литеры на изображении Q_i . Распознавание и сегментацию символов на изображении обеспечивают современные системы оптического распознавания символов [112].

Для бинарного изображения существует процедура извлечения g признаковых описаний различных типов:

$$F_t(Q_i) = \mathbf{q}_{it}, \quad i = 1, \dots, k; t = 1, \dots, g,$$

признаковое описание каждого типа представляет собой конечномерный вектор: $\mathbf{q}_{it} \in \mathbb{S}$, $\mathbb{S} = \bigcup_{j=1}^{\infty} \mathbb{R}^j$. Между дескрипторами одного и того же типа можно вычислить расстояние с помощью функции $d_t : \mathbb{S} \times \mathbb{S} \rightarrow \mathbb{R}$.

База эталонных дескрипторов формируется следующим образом. Пусть в базе содержится m шрифтов, а алфавит имеет мощность l . Для каждого шрифтового начертания каждая литера растеризуется в высоком разрешении, образуя бинарное изображение R_{ij} , $i = 1, \dots, l$, $j = 1, \dots, m$. Для этого изображения вычисляются признаковые описания всех типов:

$$\mathbf{r}_{ijt} = F_t(R_{ij}), \quad t = 1, \dots, g,$$

Определим теперь расстояние от надписи до конкретного шрифта с номером j . Оно получается как взвешенная сумма расстояний по всем символам и всем типам дескрипторов:

$$\begin{aligned} \Delta_j &= \frac{1}{kg} \sum_{i=1}^k \sum_{t=1}^g w_{it} \Delta_{ijt}, \\ \Delta_{ijt} &= d_t(\mathbf{q}_{it}, \mathbf{r}_{c_{ijt}}), \quad i = 1, \dots, k; t = 1, \dots, g, \end{aligned} \tag{4.1}$$

здесь веса w_{it} , $i = 1, \dots, l$; $t = 1, \dots, g$ определяют важность дескриптора t -го типа i -й литеры в алфавите для распознавания шрифта.

Результатом распознавания изображения Q является номер шрифта с минимальным расстоянием:

$$A(Q) = \arg \min_j \Delta_j.$$

Кроме того, в результате распознавания имеющиеся в базе шрифты ранжируются по мере сходства с представленным образцом, т.е. в порядке убывания

расстояния, что позволяет получить как количественную (расстояния), так и качественную (номер в упорядоченном списке) оценку сходства.

Возникает задача настройки весов w_{it} с целью получить как можно более высокое качество распознавания. Пусть надпись на изображении Q набрана y -м шрифтом. Тогда алгоритм правильно идентифицирует шрифт, если

$$\Delta_y < \min_{j \neq y} \Delta_j.$$

Идентификация происходит тем увереннее, чем больше величина отступа:

$$M = \min_{j \neq y} \Delta_j - \Delta_y.$$

Это приводит нас к идее сформулировать задачу настройки весов как задачу квадратичной оптимизации по принципу максимизации отступа, в стиле, аналогичном многоклассовому методу опорных векторов [113]. Будем считать здесь, что обучающая выборка содержит n надписей с, вообще говоря, разным набором и количеством символов. Обозначим для s -й надписи k_s — число символов, c_{is} — код i -го символа, y_s — номер правильного шрифта, а Δ_{ijt}^s — расстояние по дескриптору t -го типа до i -го символа j -го шрифта, вычисляемое по второй формуле в 4.1. Требуемая задача выглядит так:

$$\begin{cases} \frac{1}{2} \sum_{i=1}^l \sum_{t=1}^g w_{it}^2 + C \sum_{s=1}^n \xi_s \rightarrow \min; \\ \frac{1}{k_s g} \sum_{i=1}^{k_s} \sum_{t=1}^g w_{c_{ist}} (\Delta_{ijt}^s - \Delta_{iy_{st}}^s) \geq 1 - \xi_s, \quad s = 1, \dots, n; \quad 1 \leq j \leq m; \quad j \neq y_s; \\ \xi_s \geq 0, \quad s = 1, \dots, n; \\ w_{it} \geq 0, \quad i = 1, \dots, l; \quad t = 1, \dots, g. \end{cases} \quad (4.2)$$

Выражение в левой части неравенства во второй строке системы задаёт отступ между правильным шрифтом и шрифтом-конкурентом, если он недостаточно велик или отрицателен, шрифт идентифицируется недостаточно уверенно или вовсе неправильно, и становятся активными ограничения на переменные штрафа ξ_s . Величина C определяет компромисс между шириной разделяющей полосы и тем, насколько нежелательно присутствие объектов в полосе и на стороне объектов чужого класса.

Построение дескрипторов будем осуществлять на основе функций морфологических моментов $\mathfrak{M}_{pq}(r)$ и морфологической ширины $\mathfrak{F}(r) = \mathfrak{M}_{00}(r)$ как её частного случая. Если бинарному изображению Q соответствует многоугольная сцена X , то из её морфологического момента порядков (p,q) получается

дескриптор номер $t = \frac{(p+q)(p+q+1)}{2} + q + 1^1$ следующим образом:

$$(F_t(Q))_i = \lambda^{2+p+q} \mathfrak{M}_{pq} \left(\frac{(i-1)d}{\lambda} \right), \quad i = 1, \dots, N, \quad (4.3)$$

где d — стандартный шаг радиуса, λ — масштабный коэффициент, необходимый для нормализации диаграмм ширины, полученных для изображений символов разного размера, а N подобрано таким образом, что

$$\mathfrak{F} \left(\frac{(N-1)d}{\lambda} \right) > 0, \quad \mathfrak{F} \left(\frac{Nd}{\lambda} \right) = 0.$$

Так как морфологическая ширина — невозрастающая функция, при заданных d и λ значение N определяется однозначно.

Наконец, сами дескрипторы будем сравнивать с помощью L_1 -метрики. Так как мы имеем дело с дескрипторами, вообще говоря, разной длины, дополним недостающую часть «короткого» дескриптора нулями, поскольку площадь дискового покрытия при соответствующих значениях аргумента равна нулю:

$$d_t(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{\min(l(\mathbf{p}), l(\mathbf{q}))} |\mathbf{p}_i - \mathbf{q}_i| + \sum_{i=\min(l(\mathbf{p}), l(\mathbf{q}))+1}^{l(\mathbf{p})} \mathbf{p}_i + \sum_{i=\min(l(\mathbf{p}), l(\mathbf{q}))+1}^{l(\mathbf{q})} \mathbf{q}_i,$$

здесь $l(\mathbf{p})$ обозначает длину вектора \mathbf{p} .

Итак, мы описали принцип работы системы распознавания и можем перейти к демонстрации её работы.

4.1.3 Эксперименты

Для создания базы эталонных дескрипторов были использованы 2543 шрифтовых начертания коллекции цифровых шрифтов компании Паратайл [114]. Отметим, что принадлежность начертаний к одному и тому же семейству с точки зрения нашего метода не играет никакой роли, и эти начертания рассматриваются как полностью различные. Размер шрифта был выбран таким образом, чтобы высота заглавной буквы Н составила 1000 пикселей.

Для создания базы тестовых образцов символы этих же шрифтов были растеризованы в более низком разрешении так, чтобы высота буквы Н составила 100, 70, 50 и 40 пикселей. Каждый образец содержал по одному изображению

¹Такое отображение определяет взаимно однозначное соответствие между парами (p,q) , $p,q \geq 0$ и положительными t .

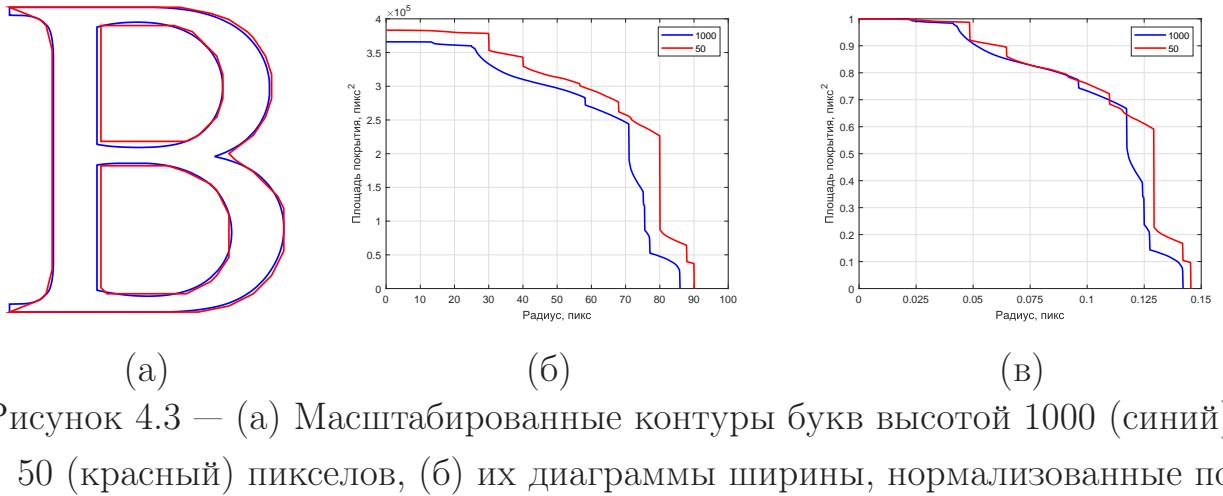


Рисунок 4.3 — (а) Масштабированные контуры букв высотой 1000 (синий) и 50 (красный) пикселов, (б) их диаграммы ширины, нормализованные по высоте и (в) по площади

всех 52 букв латинского алфавита. При этом возникла задача нормализации диаграмм ширины тестовых образцов. Отметим, что вычисление дескрипторов ширины по формуле 4.3 по сути представляет собой вычисление с шагом d значений функции ширины для фигуры, полученной из исходной растяжением в λ раз. Таким образом, нормализация сводится к подбору значений λ . Пусть h — высота заглавной буквы Н в надписи-образце, а S_1, \dots, S_k — площади символов в надписи. Были рассмотрены три стратегии нормализации:

- нормализация по высоте: $\lambda = \frac{1000}{h}$ (совмещается высота заглавных букв Н);
- индивидуальная нормализация по площади: $\lambda_i = \sqrt{10^5 S_i}$ (каждый символ приводится к фиксированной площади);
- совокупная нормализация по площади: $\lambda = \sqrt{\frac{10^5 k}{\sum_{i=1}^k S_i}}$, где k -число символов в надписи (площадь всей надписи приводится к фиксированному значению).

Во всех случаях шаг радиуса при вычислении диаграмм был равен 0,1 пикселя.

Результат первых двух методов нормализации изображён на рис. 4.3. Несмотря на то что результат второго типа нормализации выглядит явно более убедительным, при нём стираются различия в пропорциях между символами надписи.

Для определения лучшего типа нормализации было измерена точность классификации композиции, учитывающей все 52 символа. Можно считать, что в этом случае мы находимся в идеальных условиях и имеем отпечатанный образец шрифта. Веса классификаторов по отдельным буквам были получены на решения задачи 4.2 для обучения были случайным образом взяты 20% шриф-

тов из набора, параметр C был выбран равным 10^{-5} в силу того, что значения расстояний весьма велики. Результаты распознавания при применении нормализации каждого типа приведены в таблице 3.

Размер	По площади	По площади индивидуально	По площади в совокупности
100	96,23%	95,95%	96,69%
70	94,64%	93,71%	94,87%
50	92,20%	91,55%	92,58%
40	74,10%	71,49%	76,87%

Таблица 3 — Точность распознавания при различных типах нормализации

Таким образом, наиболее эффективной оказалась нормализация по общей площади надписи, с которой мы и будем работать далее. Также при анализе ошибок метода выяснилось, что значительная их доля связана с неразличением прямого и курсивного начертания одного и того же шрифта. Эту ситуацию иллюстрируют графики на рис. 4.4, из которых видно, что функции морфологической ширины символа в различных начертаниях могут быть очень похожи, а функции морфологических моментов — ощутимо отличаться. Поэтому с целью извлечения более тонких признаков ширины для тех же символов были вычислены центральные моменты порядков (1,0) и (0,1).

Для проведения эксперимента в более реальных условиях были взяты 1000 самых популярных английских слов (средняя длина слова составила 5,33 символа, в выборке были и очень короткие слова, например, I, be), из которых 40% были переведены в верхний регистр. Классификация производилась на основе метода ближайшего соседа, для определения шрифта по слову использовались только буквы, входящие в слово.

Размер	Площадь	Момент (1,0)	Момент (0,1)	Всё вместе
100	91,45%	94,24%	93,38%	96,20%
70	81,42%	88,02%	85,56%	91,62%
50	69,75%	80,60%	76,43%	86,92%
40	48,83%	67,32%	60,07%	77,59%

Таблица 4 — Точность распознавания для моментов начальных порядков и их комбинации

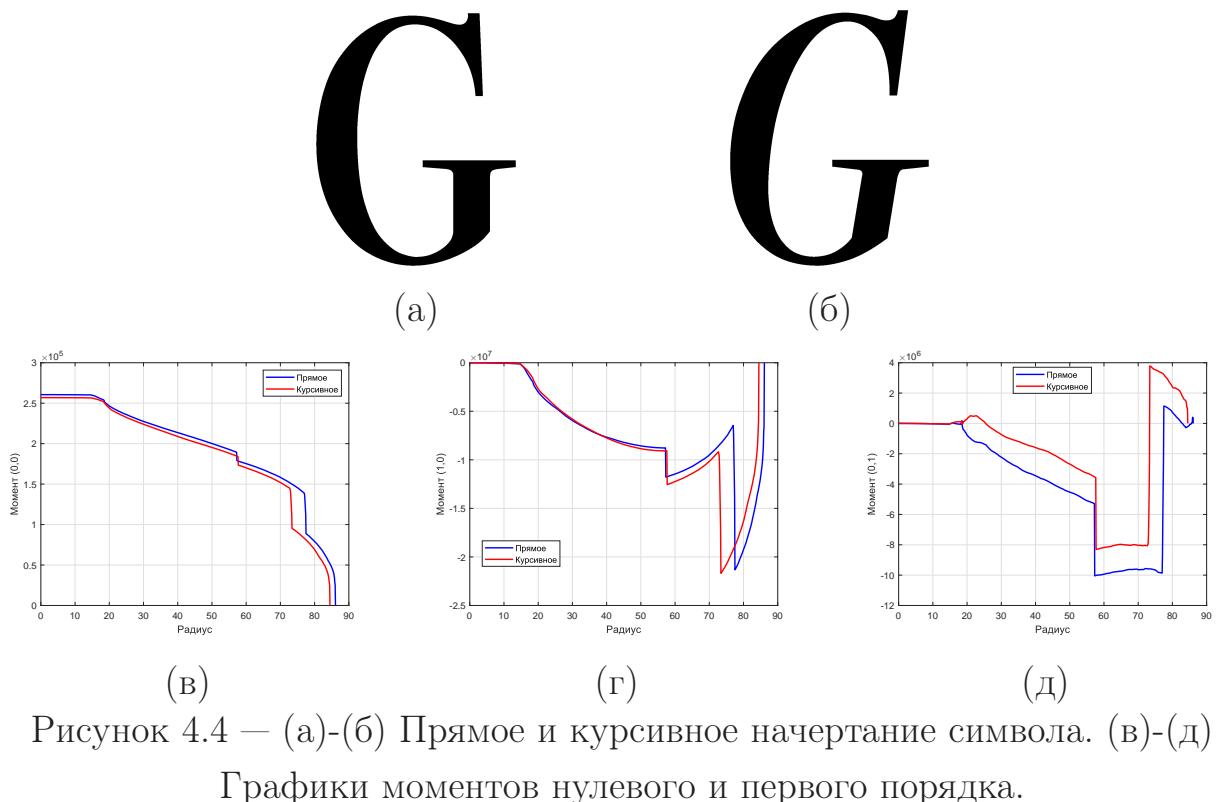


Рисунок 4.4 – (а)-(б) Прямое и курсивное начертание символа. (в)-(д)
Графики моментов нулевого и первого порядка.

Для проведения экспериментов использовалась реализация алгоритмов вычисления морфологической ширины и морфологических моментов на языке C++, обработка результатов проводилась в системе MATLAB. Создание базы эталонных диаграмм на компьютере с процессором Intel® Core i5™ и 6 ГБ оперативной памяти заняло 6,5 часов, причем почти все это время было затрачено на растеризацию букв большого размера, а построение скелетов и расчет диаграмм ширины заняли в сумме менее 15 минут. В ходе эксперимента были вычислены расстояния по 52 буквам между всеми парами из 2543 начертаний, что заняло 42 минуты. Это означает, что время запроса — проверки одного начертания по базе эталонов — для надписей, содержащих все 52 символа, составило около одной секунды.

Чтобы сравнить временные затраты с дискретным (реализация на основе алгоритма Урбаха-Уилкинсона [50]) и дискретно-непрерывным методом (на основе алгоритма [93]), было замерено время на обработку символов 20 случайно выбранных шрифтов при разных размерах. Шаг был выбран равным 1, так как дискретный и дискретно-непрерывный методы допускают только целые значения радиусов. Отметим, что в нашем алгоритме вычисление моментов различных порядков представляет собой отдельные процедуры, ход вычислений в которых не дублируется (однако все они требуют предварительной скелетизации), в то время как в алгоритмах-аналогах достаточно лишь по-разному

проводить постобработку результатов. Поэтому рост затрат непрерывного алгоритма при вычислении сразу нескольких моментов для одной фигуры будет значительно выше.

Тем не менее результаты, приведённые в таблице 5, демонстрируют убедительное превосходство в скорости предложенного метода, особенно на изображениях большого размера. Отметим, что в экспериментах использовался метод поиска смежных бициклов на основе инкрементного прослеживания, хотя выбор конкретного метода не играет существенной роли: для того, чтобы различия проявились, нужны специально сгенерированные «искусственные» фигуры со множеством линз, по сравнению с которыми форма печатных символов очень проста.

Метод \ Размер	1000	400	200
Дискретный [50]	5291	679	217
Дискретно-непрерывный [93]	71,48	15,91	8,84
Непрерывный (площадь)	5,75	2,32	1,12
Моменты 1-го порядка	6,89	2,78	1,34
Моменты 2-го порядка	8,28	3,32	1,59
Скелетизация	5,54	2,28	1,10

Таблица 5 — Среднее время на обработку одного символа, мс. Затраты методов, за исключением дискретного, указаны с учётом скелетизации.

Таким образом, проведенный эксперимент подтвердил работоспособность предложенного метода и показал его эффективность на практической задаче сравнения большого количества изображений с весьма высоким качеством распознавания.

4.2 Компьютерная морфометрия планарий

4.2.1 Описание задачи

Планарии — это плоские черви, по краям тела которых находятся реснички, обеспечивающие их движение по поверхности. В основном обитают в пресноводных водоёмах. Известны тысячи видов планарии, размерами взрослые особи могут быть от нескольких миллиметров до 35 см. Ресничные плоские черви планарии обладают выдающимися способностями к регенерации — отрастанию утраченных частей тела после поперечной перерезки. Так, даже при

поперечной перерезке в области головы выживают обе части тела планарии: головная часть доращивает хвостовую, а хвостовая — головной конец. Первый этап регенерации проходит за неделю: у планарии образуется новый «головной мозг» и она становится способна к поиску пищи. Окончательное восстановление занимает около трёх недель. Благодаря этому планарии широко используются в качестве подопытного материала в биологических экспериментах. В частности, их используют при изучении биологических основ процесса регенерации, а также процессов пролиферации и дифференцировки клеток, поскольку гены, отвечающие за эти явления у планарий гомологичны подобным генам человека (таких генов сейчас найдено более 30).

В зависимости от внешних воздействий и условий, например, качества воды, регенерация может замедляться или ускоряться, поэтому группой биологов было выдвинуто предложение использовать планарии как биоиндикатор экологического состояния среды. Идея состоит в том, чтобы регистрировать изменение скорости регенерации планарий в зависимости от условий среды, в которую они помещены. Для реализации этой идеи разрабатываются методы оценки скорости регенерации по изображениям планарий, полученным в течение всего периода регенерации с помощью фотокамеры [115]. Традиционно количественная оценка степени регенерации осуществляется исследователем на основе визуального анализа. При этом результат сильно зависит от того, как определяется граница между телом и вновь прирастающей областью (блестемой), а точно выделить эту границу сложно даже специалисту. В связи с этим актуальной задачей является разработка критериев и методов оценки скорости регенерации с помощью инструментов обработки и анализа цифровых изображений в автоматическом режиме. Основной задачей эксперимента была разработка метода количественной оценки скорости регенерации. Цель такой разработки — повышение точности оценок и сокращение трудозатрат по сравнению с традиционной техникой «ручной» обработки на основе визуального анализа.

4.2.2 Метод решения

В основе экспериментов лежит наблюдение, свидетельствующее о том, что в ходе регенерации форма головы планарии становится всё более и более заострённой. По этой причине подходящей оценкой степени регенерации планарии может служить морфологическая ширина её головы, что требует грамотного

выделения области интереса. Стоит также учитывать, что абсолютные размеры планарии на фотографии зависят как от ракурса съёмки, так и от изгиба самой планарии и особенностей процесса её регенерации, поскольку он протекает за счет внутренних пластических и энергетических запасов червя. Наконец, планарии фотографируются в процессе произвольного движения при помощи микроскопа, что ведёт к не самому высокому качеству входных изображений, так что оценка ширины индивидуальной планарии имеет достаточно высокий разброс, и более уместно оценивать степень регенерации по группе в совокупности. Таким образом, можно выделить следующие этапы решения задачи:

1. Сегментация — получение бинарного изображения, в котором чёрные точки соответствуют планарии, а белые — фону.
2. Выделение области интереса — области, в которой расположена голова.
3. Вычисление дескриптора ширины — площади достаточно узкой части планарии в пределах области интереса.
4. Нормализация дескриптора с учётом абсолютных размеров головы.
5. Вычисление морфологической функции для группы в целом.
6. Сопоставление результатов в разные дни и по группам.

Остановимся на каждом пункте подробнее. Можно заметить, что в наборе экспериментальных условия съёмки более или менее устойчивы, и яркость фона у границ кадра оказывается стабильно ниже, чем яркость в середине. Это свойство используем в процессе бинаризации, который состоит из следующих шагов:

1. Сглаживание полутонового изображения (рис. 4.5а) фильтром Гаусса со стандартным отклонением $\sigma = 3$ (рис. 4.5б).
2. Пороговая бинаризация изображения таким образом, что 10% пикселей становятся чёрными (рис. 4.5в).
3. Определение основы связной компоненты планарии — максимальной связной компоненты, которая не касается границ кадра (красная область на рис. 4.5в).
4. Определение предельного значения порога t_{max} — повышение порога бинаризации до тех пор, пока связная компонента планарии не будет касаться границ кадра (рис. 4.5г).
5. Бинаризация изображения с порогом $t_{max} - 4$ и выделение связной компоненты, содержащей основу (рис. 4.5д).

Процесс бинаризации проиллюстрирован рис. 4.5.

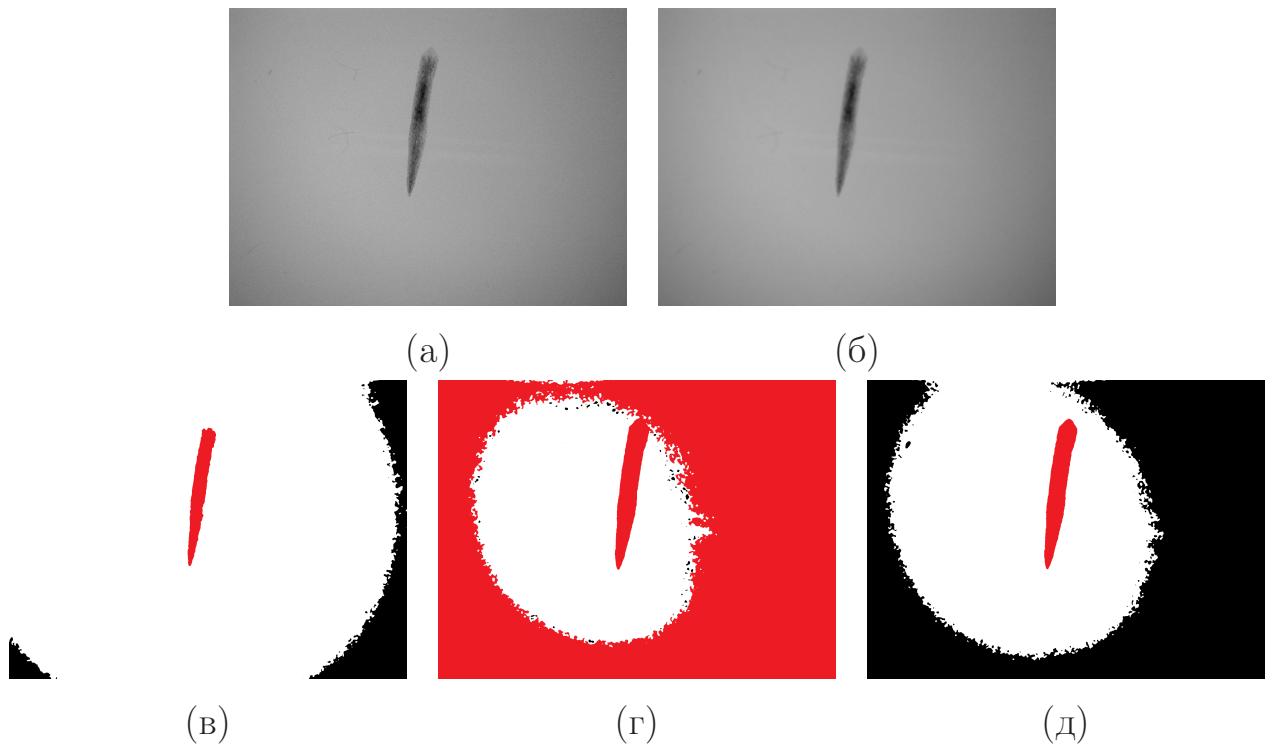


Рисунок 4.5 — Последовательность шагов при бинаризации

Для определения области интереса построим по бинаризованному изображению скелет. Применим к нему операцию стрижки (отсечения рёбер скелета, вносящих слабый вклад в образование формы фигуры), описанную в [43], с максимально возможным отклонением до исходной формы в 20 пикселей. Это приведёт к тому, что на большинстве изображений скелет будет редуцирован до единственной ветви, а на оставшихся мы будем рассматривать самую длинную ветвь. Было замечено, что на всех изображениях в выборке голова расположена выше хвоста. Обозначим r_{top} значение радиальной функции в верхнем конце скелета. Далее, двигаясь по скелету сверху вниз, пройдём расстояние, равное $0,2r_{top}$ и обозначим точку остановки p_{cut} , а значение радиальной функции в ней — r_{cut} . В этом месте мы проведём разрез — построим прямоугольник со сторонами $4r_{cut}$ на $3r_{cut}$, нижняя сторона длины $4r_{cut}$ которого делится точкой p_{cut} пополам и проходит перпендикулярно скелету в этой точке. Такая процедура обеспечивает успешное выделение области головы планарии на большинстве изображений, исключение составляют лишь изображения с откровенно неудачной бинаризацией в этой области из-за серьёзных помех, но их и без того затруднительно использовать для оценки ширины.

Теперь, имея область интереса D , определим функцию частной морфологической ширины $\mathfrak{F}_D(r)$ по алгоритмам из параграфа 3.1. Напомним, что нас интересует площадь достаточно узкой, а не достаточно широкой части, кото-

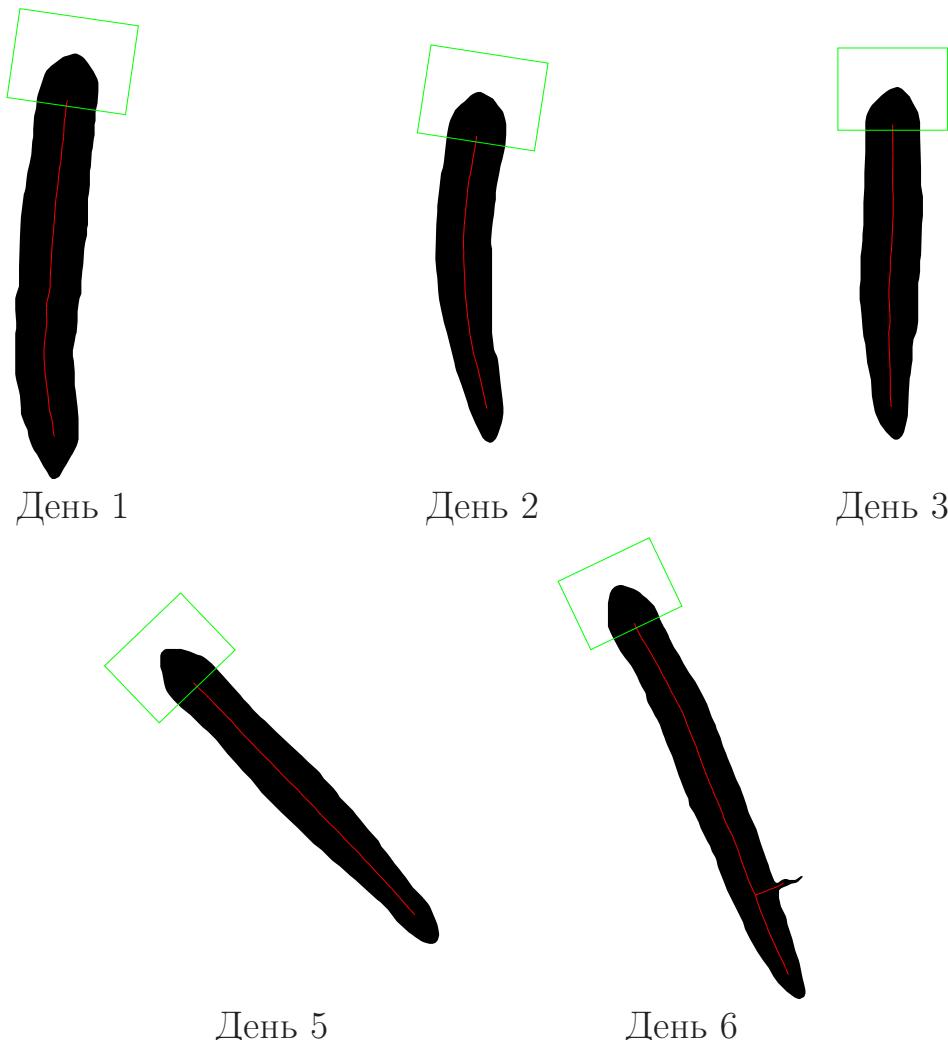


Рисунок 4.6 – Выделение областей интереса

ную можно определить как дополнение дискового покрытия до всей фигуры. Действительно, гранулометрическая функция (ширина в точке) $G(p)$ определяется для точек фигуры X как радиус максимального вписанного в фигуру круга, покрывающего точку. Дисковое r -покрытие фигуры восстанавливается из значений этой функции:

$$X(r) = \{p \mid G(p) \geq r\}.$$

Соответственно, точки, в которых локальная ширина меньше r , образуют область $X \setminus X(r)$, поэтому площадь узкой части определяется по формуле

$$S_0(r) = \mathfrak{F}_D(0) - \mathfrak{F}_D(r).$$

Чтобы сделать результат инвариантным к масштабу изображения, нормируем его таким образом, чтобы аргумент r_{cut} соответствовал в новой функции единице:

$$S(r) = \frac{1}{r_{cut}^2} S_0(r \cdot r_{cut}).$$

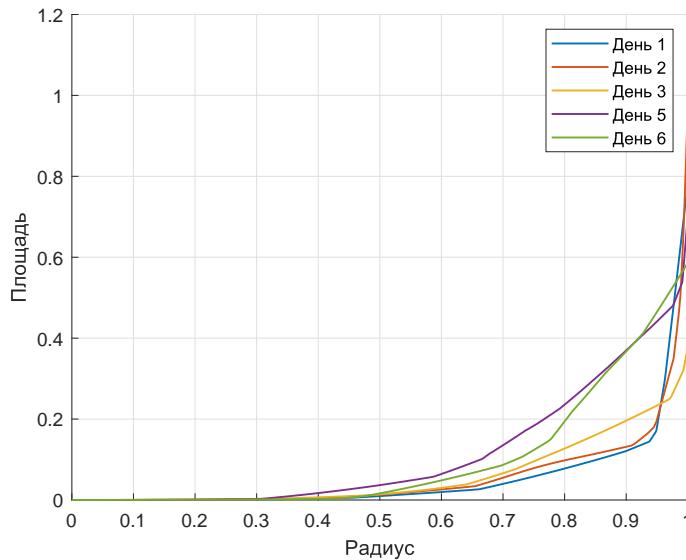


Рисунок 4.7 — Диаграммы ширины одной и той же особи в разные дни

Результат выделения области интереса для фотографии одной и той же особи в разные дни показан на рис. 4.6), а соответствующие диаграммы ширины — на рис. 4.7. Отметим, что поведение функции при значениях аргумента больше 0,9 является слишком неустойчивым, чтобы принимать его в расчёте.

Хотя результат хорошо согласуется с интуитивным восприятием «засторённости», возникает необходимость агрегации результатов по группам, которую мы произведём простым усреднением:

$$S_{dg}(r) = \frac{1}{n_{dg}} \sum_{i=1}^{n_{dg}} S_{dg}^i(r),$$

где d — индекс дня, g — индекс группы, n_{dg} — число фотографий, соответствующих этим дню и группе, и $\{S_{dg}^i(r)\}$ — морфологические функции, вычисленные по этим фотографиям.

4.2.3 Эксперименты

В ходе экспериментов исследовалась скорость регенерации опытной группы, помещенной в раствор пептида в концентрации 10^{-12} , и контрольной группы, помещенной в воду, с первого по шестой день после разреза, за исключением четвёртого.

Входными данными служили полутоновые изображения размера 1388×1040 , всего в наборе была 171 фотография, 8 из которых пришлось исключить

из выборки, так как они содержали дефекты, не позволяющие провести автоматическую сегментацию качественно. Число образцов каждой группы по дням представлено в таблице 6.

Группа	Контрольная	Опытная
День 1	10	8
День 2	13	19
День 3	24	18
День 5	16	19
День 6	18	18

Таблица 6 — Число отобранных фотографий в группах по дням

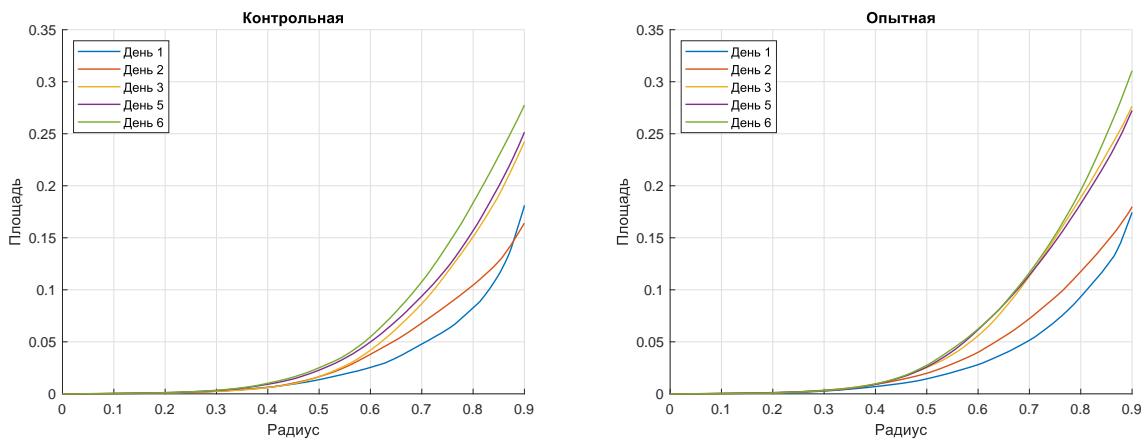


Рисунок 4.8 — Групповые диаграммы ширины одной и той же группы в разные дни

Морфологические функции для каждой из групп в отдельности показаны на рис. 4.8. Мы видим, что значения функции при начальных значениях аргумента стабильно возрастают с течением времени, хотя после третьего дня рост сильно замедляется. Так как сам факт регенерации планарий не подлежит сомнению, можно сделать вывод, что предложенный критерий вполне успешно оценивает степень регенерации. Это позволяет с достаточным вниманием отнестись к результатам на рис. 4.9, где показаны результаты для двух групп в каждый из дней. Здесь опытная группа демонстрирует стабильно более высокие показатели, поэтому эксперимент служит аргументом в пользу гипотезы, что пептид стимулирует регенерацию планарий. При этом наиболее показательным является третий день и значения аргумента в диапазоне [0,7; 0,9]. Отметим также, что основные вычислительные затраты метода связаны с бинаризацией, в

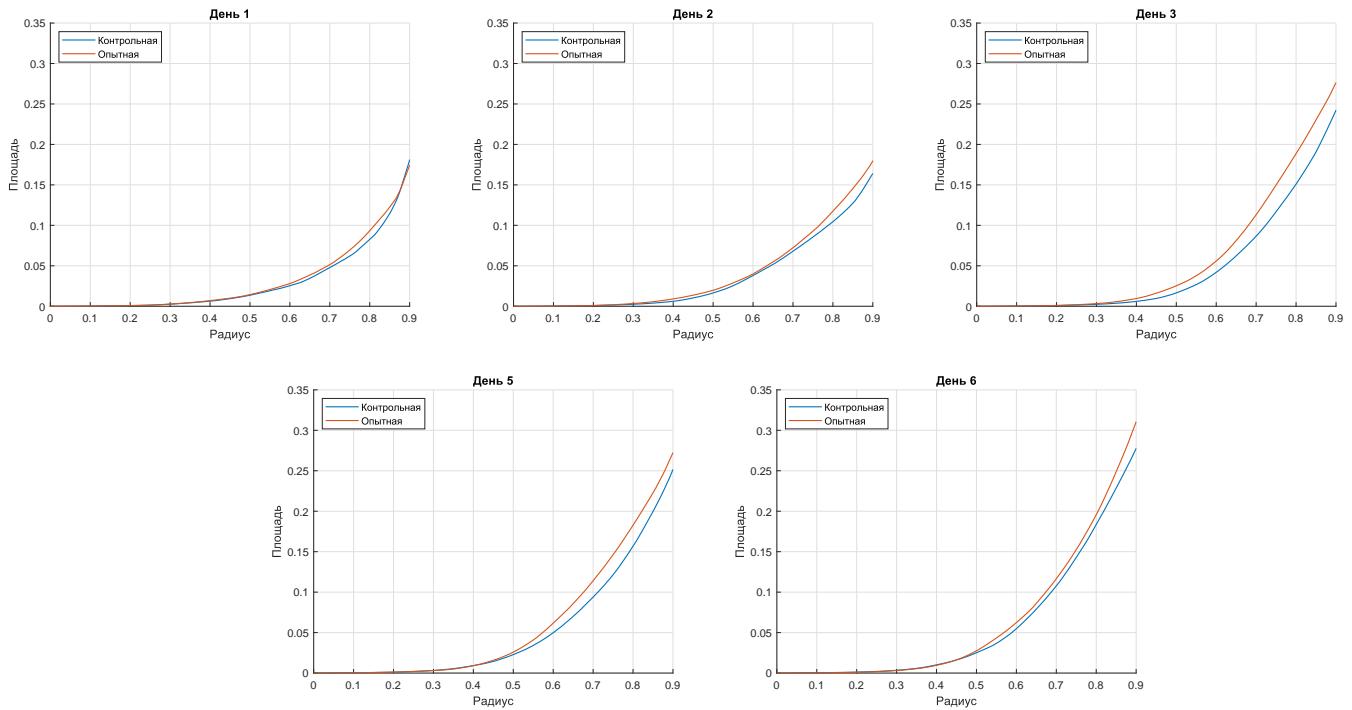


Рисунок 4.9 — Групповые диаграммы ширины для обеих групп в один и тот же день

особенности, выделением связных компонент. Реализованное на языке C++ вычисление морфологической функции занимает, включая скелетизацию, меньше трёх миллисекунд в среднем на изображение.

Таким образом, была разработана полностью автоматизированная система компьютерной морфометрии планарий, которая продемонстрировала возможность решения задач, связанных с оценкой ширины сегментированной фигуры. Предложенный морфологический критерий оказался эффективной численной оценкой качества окружающей среды, в которой происходит регенерация планарий.

4.3 Выводы к главе 4

1. Проведены вычислительные эксперименты на реальных и синтезированных данных, показывающие полезность разработанных методов оценки ширины для решения практических задач.
2. Разработана система распознавания шрифта по изображению текста на основе морфологических признаков, оценивающих ширину отдельных штрихов символов шрифта. Показано, что морфологические моменты действительно несут дополнительную информацию о форме объекта —

- например, о том, какая сторона символа (левая или правая) в целом шире, не учитываяющуюся морфологической шириной.
3. Скорость вычислений дескрипторов ширины оказалась достаточной для обработки обширной базы изображений символов, содержащей более ста тысяч образцов в высоком разрешении, за разумное время. Сравнение диаграмм ширины L_1 -метрикой даёт подходящую оценку различия объектов по ширине и обеспечивает достаточно быстрый и качественный поиск наиболее похожего шрифта в базе.
 4. Проведённый эксперимент по оценке заострённости головной части плоских червей-планарий после их поперечной перерезки свидетельствует в пользу гипотезы, утверждающей, что пептид стимулирует регенерацию планарий. Показано, что величину спада морфологической ширины по сравнению с начальным значением можно рассматривать как площадь достаточно узкой части фигуры и использовать как оценку заострённости определённого сегмента фигуры. При этом вычисление частной морфологической ширины по области интереса позволяет производить оценку ширины именно для нужного сегмента, а не для всей фигуры.
 5. Рассмотрены различные способы нормализации диаграмм ширины с целью их последующего сравнения — по площади объекта, габариту, максимальному вписанному кругу и т.д. Показано, что выбор способа нормализации не универсален и определяется особенностями задачи. Кроме того, стоит отметить, что разработанные методы достаточно требовательны к качеству входных данных — разрешению изображений и точности их сегментации.

Заключение

В ходе работы над исследованием был разработан комплекс методов, позволяющих описать ширину объектов сложной формы с целью использования в качестве интегрального морфологического дескриптора в задачах распознавания изображений. Предложен подход к описанию понятия «ширина фигуры» на основе покрытия фигуры дисками определённого размера. В качестве базового дескриптора формы используется функция, отражающая зависимость площади дискового покрытия от размера дисков. Дисковое покрытие является ключевой математической моделью в данной работе, полностью изучена его структура, а главное — впервые разработаны универсальные алгоритмы обработки пересечений примитивных областей покрытия, прежде казавшиеся «камнем преткновения» при разработке точных аналитических алгоритмов расчёта дескрипторов ширины, оперирующих непрерывными моделями формы.

Также был исследован ряд производных от дискового покрытия моделей и производных дескрипторов, которые могут быть использованы в зависимости от постановки задачи:

- для случая, когда входные данные являются полутоновыми, а не бинарными изображениями, предназначена система дисковых покрытий;
- если в фигуре выделена область интереса, для которой нужно получить отдельный дескриптор ширины, можно вычислить частную морфологическую ширину;
- если необходимо учесть пространственное распределение ширины, могут оказаться полезными морфологические моменты и их инварианты;
- если требуется выделить зависимости признаков ширины и других признаков формы, можно рассчитать морфологические карты из признаков ширины отдельных точек.

Также вышеприведённые методы можно комбинировать между собой, обеспечивая подходы к решению задач, требующих описания ширины объектов на изображении в самых различных формулировках. Алгоритмы расчёта всех предложенных дескрипторов имеют теоретическое обоснование, их точность и эффективность продемонстрирована в ходе вычислительных экспериментов с использованием оригинальной программной реализации.

Полезность предложенных способов описания ширины доказывают примеры решения практических задач, в которых разработанные методы оказываются способными выявить необходимые особенности формы объектов.

Список литературы

1. *Hayward, W. G.* Effects of Outline Shape in Object Recognition / W. G. Hayward // Journal of Experimental Psychology: Human Perception and Performance. — 1997. — т. 24. — с. 227—440.
2. *Lloyd-Jones, T. J.* Outline shape is a mediator of object recognition that is particularly important for living things / T. J. Lloyd-Jones, L. Luckhurst // Memory & Cognition. — 2002. — т. 30, № 4. — с. 489—498.
3. *Ожегов, С. И.* Толковый словарь русского языка, изд. 4-е, дополненное / С. И. Ожегов, Н. Ю. Шведова. — Российская академия наук, институт русского языка им. В. В. Виноградова, 2006.
4. *Lin, Y.* Contour Shape Description Based on an Arch Height Function / Y. Lin, J. Dou, H. Wang // Pattern Recognition. — New York, NY, USA, 1992. — т. 25, № 1. — с. 17—23.
5. Shape Matching and Classification Using Height Functions / J. Wang [и др.] // Pattern Recognition Letters. — New York, NY, USA, 2012. — т. 33, № 2. — с. 134—143.
6. *Zhang, D.* Review of shape representation and description techniques / D. Zhang, G. Lu // Pattern Recognition. — 2004. — т. 37, № 1. — с. 1—19.
7. Shape Feature Extraction Using Fourier Descriptors with Brightness in Content-Based Medical Image Retrieval / G. Zhang [и др.] // Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing. — Washington, DC, USA : IEEE Computer Society, 2008. — с. 71—74.
8. Elliptic Fourier Analysis of Cell and Nuclear Shapes / G. Diaz [и др.] // Computers and Biomedical Research. — San Diego, CA, USA, 1989. — т. 22, № 5. — с. 405—414.
9. Shape Distributions / R. Osada [и др.] // ACM Transactions on Graphics. — New York, NY, USA, 2002. — т. 21, № 4. — с. 807—832.
10. *Kindratenko, V.* On using functions to describe the shape / V. Kindratenko // Journal of Mathematical Imaging and Vision. — 2003. — т. 18, № 3. — с. 225—245.

11. *Belongie, S. J.* Shape Matching and Object Recognition Using Shape Contexts / S. J. Belongie, J. Malik, J. Puzicha // IEEE Trans. Pattern Anal. Mach. Intell. — 2002. — т. 24, № 4. — с. 509—522.
12. *Bai, S.* Aggregating contour fragments for shape classification / S. Bai, X. Wang, X. Bai // 2014 IEEE International Conference on Image Processing, ICIP 2014, Paris, France, October 27-30, 2014. — 2014. — с. 5252—5256.
13. *Ling, H.* Shape Classification Using the Inner-Distance / H. Ling, D. W. Jacobs // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2007. — т. 29, № 2. — с. 286—299.
14. *Serra, J.* Image Analysis and Mathematical Morphology / J. Serra. — Orlando, FL, USA : Academic Press, Inc., 1983.
15. *Matheron, G.* Random Sets and Integral Geometry / G. Matheron. — Wiley, 1975. — (Probability and mathematical statistics).
16. *Breen, E. J.* Attribute Openings, Thinnings, and Granulometries / E. J. Breen, R. Jones // Computer Vision and Image Understanding. — 1996. — т. 64, № 3. — с. 377—389.
21. *Maragos, P.* Pattern Spectrum and Multiscale Shape Representation / P. Maragos // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1989. — т. 11, № 7. — с. 701—716.
22. *Визильтер, Ю.* Морфологические спектры / Ю. Визильтер, С. Сидякин // Техническое зрение в системах управления — 2012. Труды научно-технической конференции // под ред. Р.Р. Назирова. — М.: ИКИ РАН, 2012. — с. 234—241.
23. *Vizilter, Y.* Comparison of shapes of two-dimensional figures with the use of morphological spectra and EMD metrics / Y. Vizilter, S. Sidyakin // Pattern Recognition and Image Analysis. — 2015. — т. 25. — с. 365—372.
24. *Визильтер, Ю. В.* Параметрические и морфологические спектры / Ю. В. Визильтер, С. В. Сидякин // Компьютерная оптика. — 2015. — т. 39, вып. 1. — с. 109—118.
25. *Shekar, B. H.* Shape Representation and Classification through Pattern Spectrum and Local Binary Pattern – A Decision Level Fusion Approach / B. H. Shekar, B. Pilar // Proceedings of 5th International Conference on Signal and Image Processing, ICSIP 2014. — 2014. — с. 218—224.

26. Grid structured morphological pattern spectrum for off-line signature verification / B. H. Shekar [и др.] // International Conference on Biometrics, ICB 2015, Phuket, Thailand, 19-22 May, 2015. — c. 430—435.
27. *Pilar, B.* Circular Map Pattern Spectrum – An Accurate Descriptor for Shape Representation and Classification / B. Pilar, B. H. Shekar // Proceedings of Data Analysis and Learning 2018. — Department of Studies on Computer Science, University of Mysore, 2019. — c. 207—219.
17. *Blum, H.* A Transformation for Extracting New Descriptors of Shape / H. Blum // Models for the Perception of Speech and Visual Form / под ред. W. Wathen-Dunn. — Cambridge : MIT Press, 1967. — c. 362—380.
18. *Lantuejoul, C.* La squelettisation et son application aux mesures topologiques de mosaiques polycristallines. PhD thesis / Lantuejoul C. — 1979.
19. *Maragos, P.* Morphological skeleton representation and coding of binary images / P. Maragos, R. W. Schafer // IEEE Trans. Acoustics, Speech, and Signal Processing. — 1986. — т. 34. — с. 1228—1244.
20. *Siddiqi, K.* Medial Representations: Mathematics, Algorithms and Applications / K. Siddiqi, S. Pizer. — 1st Edition. — Springer Publishing Company, Incorporated, 2008.
28. Shape Recognition by Combining Contour and Skeleton into a Mid-Level Representation / W. Shen [и др.] // Pattern Recognition - 6th Chinese Conference, CCPR 2014, Changsha, China, November 17-19, 2014. Proceedings, Part I. — 2014. — c. 391—400.
29. Shape recognition by bag of skeleton-associated contour parts / W. Shen [и др.] // Pattern Recognition Letters. — 2016. — т. 83. — с. 321—329.
30. *Xie, J.* Shape matching and modeling using skeletal context / J. Xie, P.-A. Heng, M. Shah // Pattern Recognition. — 2008. — т. 41, № 5. — с. 1756—1767.
31. *Зак, Е. М.* Сравнение формы изображений на основе спектра ширины объектов / Е. М. Зак, Н. А. Ломов, Л. М. Местецкий // Труды 24-й Международной конференции по компьютерной графике и зрению Графикон'2014. — Академия архитектуры и искусств ЮФУ, Ростов-на-Дону, 2014. — с. 150—154.

32. Ломов, Н. А. Идентификация цифровых шрифтов на основе морфологической ширины / Н. А. Ломов, Л. М. Местецкий // GraphiCon2016 Труды Международной научной конференции. — ННГАСУ, 2016. — с. 89—95.
33. Ломов, Н. А. Распознавание цифровых шрифтов по изображениям на основе дискового покрытия / Н. А. Ломов, Л. М. Местецкий // Интеллектуализация обработки информации: Тезисы докладов 11-й Международной конференции. — ТОРУС ПРЕСС, Москва, 2016. — с. 108—109.
34. Ломов, Н. А. Классификация двумерных фигур с использованием скелетно-геодезических гистограмм толщин-расстояний / Н. А. Ломов, С. В. Сидякин, Ю. В. Визильтер // Интеллектуализация обработки информации: Тезисы докладов 11-й Международной конференции. — ТОРУС ПРЕСС, Москва, 2016. — с. 106—107.
35. Ломов, Н. А. Классификация двумерных фигур с использованием контурных, скелетных и толщинных описаний / Н. А. Ломов, С. В. Сидякин, Ю. В. Визильтер // Тезисы 8-й научно-технической конференции «Техническое зрение в системах управления». — ИКИ РАН Москва, 2017. — с. 71—72.
36. Lomov, N. Morphological Moments of Binary Images / N. Lomov, S. Sidyakin // The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. — 2nd International ISPRS Workshop on PSBB Moscow, Russia, 2017. — с. 19—25. — (WG II/10, II/5).
37. Lomov, N. Pattern Width Description through Disk Cover — Application to Digital Font Recognition / N. Lomov, L. Mestetskiy // Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017). — Setubal, Portugal, 2017. — с. 484—492.
38. Липкина, А. Л. Распознавание шрифтов по изображениям на основе анализа структуры букв / А. Л. Липкина, Н. А. Ломов // ГРАФИКОН'2017 Труды 27-й Международной научной конференции. — Пермский государственный национальный исследовательский университет, Пермь, 2017. — с. 359—366.

39. Ломов, Н. А. Классификация бинарных фигур с использованием мешка контекстов ребер скелета / Н. А. Ломов, С. В. Сидякин // Математические методы распознавания образов: Тезисы докладов 18-й Всероссийской конференции с международным участием. — ТОРУС ПРЕСС, Москва, 2017. — с. 118—119.
40. Местецкий, Л. М. Распознавание формы гибких объектов изображений на основе спектров ширины / Л. М. Местецкий, Е. М. Зак, Н. А. Ломов // Приволжский научный журнал. — Нижний Новгород, 2014. — № 4. — с. 59—67.
41. Ломов, Н. А. Площадь дискового покрытия — дескриптор формы изображения / Н. А. Ломов, Л. М. Местецкий // Компьютерная оптика. — 2016. — т. 40, № 4. — с. 516—525.
42. Ломов, Н. А. Классификация двумерных фигур с использованием скелетно-геодезических гистограмм толщин-расстояний / Н. А. Ломов, С. В. Сидякин, Ю. В. Визильтер // Компьютерная оптика. — 2017. — т. 41, № 2. — с. 227—236.
43. Местецкий, Л. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры / Л. Местецкий. — М.: ФИЗМАТЛИТ, 2009.
44. Serra, J. Lecture Notes On Morphological Operators / J. Serra. — First French-Nordic Summer Course in Mathematics. Uppsala University. Sweden, 2001.
45. Kresch, R. Skeleton-based morphological coding of binary images / R. Kresch, D. Malah // IEEE Transactions on Image Processing. — 1998. — т. 7, № 10. — с. 1387—1399.
46. Real-time implementation of morphological filters with polygonal structuring elements / J. Bartovský [и др.] // Journal of Real-Time Image Processing. — 2015. — т. 10, № 1. — с. 175—187.
47. Shape-based hand recognition approach using the morphological pattern spectrum / J. M. Ramirez-Cortés [и др.] // Journal of Electronic Imaging. — 2009. — т. 18, № 1. — с. 013012/1—013012/6.

48. *Omata, M.* Lip Recognition Using Morphological Pattern Spectrum / M. Omata, T. Hamamoto, S. Hangai // Bigun J., Smeraldi F. (eds) Audio-and Video-Based Biometric Person Authentication. AVBPA 2001. Lecture Notes in Computer Science. т. 2091. — Springer, Berlin, Heidelberg, 2001. — c. 108—114.
49. *Barnich, O.* Robust Analysis of Silhouettes by Morphological Size Distributions / O. Barnich, S. Jodogne, M. Van Droogenbroeck // Proceedings of the 8th International Conference on Advanced Concepts For Intelligent Vision Systems. — Berlin, Heidelberg : Springer-Verlag, 2006. — c. 734—745. — (ACIVS'06).
50. *Urbach, E. R.* Efficient 2-D Grayscale Morphological Transformations With Arbitrary Flat Structuring Elements / E. R. Urbach, M. H. Wilkinson // Transactions on Image Processing. — Piscataway, NJ, USA, 2008. — т. 17, № 1. — c. 1—8.
51. *Moreaud, M.* Fast algorithm for dilation and erosion using arbitrary flat structuring element: Improvement of Urbach and Wilkinson's algorithm to GPU computing / M. Moreaud, F. Itthirad // 2014 International Conference on Multimedia Computing and Systems (ICMCS). — 2014. — c. 289—294.
52. *Torres-Huitzil, C.* FPGA-based fast computation of gray-level morphological granulometries / C. Torres-Huitzil // Journal of Real Time Image Processing. — 2016. — т. 11, № 3. — c. 547—557.
53. *Bentley, J. L.* Algorithms for Reporting and Counting Geometric Intersections / J. L. Bentley, T. A. Ottmann // IEEE Transactions on Computers. — 1979. — т. C—28, № 9. — c. 643—647.
54. Implementing continuous-scale morphology via curve evolution / G. Sapiro [и др.] // Pattern Recognition. — 1993. — т. 26, № 9. — c. 1363—1372.
55. *Ghosh, P. K.* Mathematical morphological operations of boundary-represented geometric objects / P. K. Ghosh, R. M. Haralick // Journal of Mathematical Imaging and Vision. — 1996. — т. 6, № 2/3. — c. 199—222.
56. *Agarwal, P. K.* Polygon decomposition for efficient construction of Minkowski sums / P. K. Agarwal, E. Flato, D. Halperin // Computational Geometry. — 2002. — т. 21, № 1. — c. 39—61. — Sixteenth European Workshop on Computational Geometry - EUROCG-2000.

57. Vector-Based Morphological Operations on Polygons Using Straight Skeletons for Digital Pathology / D. F. G. Obando [и др.] // Discrete Geometry for Computer Imagery - 21st IAPR International Conference, DGCI 2019, Marne-la-Vallée, France, March 26-28, 2019, Proceedings. — 2019. — c. 249—261.
58. *Huber, S.* A fast straight-skeleton algorithm based on motorcycle graphs / S. Huber, M. Held // International Journal of Computational Geometry and Applications. — 2013. — т. 22. — с. 471—498.
59. *Vizilter, Y.* Comparison of shapes of two-dimensional figures with the use of morphological spectra and EMD metrics / Y. Vizilter, S. Sidyakin // Pattern Recognition and Image Analysis. — 2015. — т. 25(3). — с. 365—372.
60. *Сидякин, С.* Морфологические дескрипторы формы бинарных изображений на основе эллиптических структурирующих элементов / С. Сидякин, Ю. Визильтер // Компьютерная оптика. — 2014. — т. 38, № 3. — с. 511—520.
61. *Макарова, Е.* Непрерывные алгоритмы морфологического анализа и сравнения листьев растений. Дипломная работа / Макарова Е.Ю. — МГУ, ВМК, 2012.
62. *Местецкий, Л. М.* Медиальная ширина фигуры — дескриптор формы изображений / Л. М. Местецкий // Машинное обучение и анализ данных. — 2014. — т. 1, № 9. — с. 1291—1318.
63. *Lee, D.* Two Algorithms for Constructing a Delaunay Triangulation / D. Lee, B. Schachter // International Journal of Parallel Programming. — 1980. — т. 9. — с. 219—242.
64. Алгоритмы: построение и анализ, 3-е издание (пер. с англ.) / Т. Кормен [и др.]. — Москва: Вильямс, 2012. — (Классические учебники: Computer science).
65. *Bakina, I.* Hand Shape Recognition from Natural Hand Position / I. Bakina, L. Mestetskiy // 2011 International Conference on Hand-Based Biometrics, ICHB 2011 - Proceedings. — 2011. — с. 1—6.
66. *Reddien, P. W.* Fundamentals of planarian regeneration / P. W. Reddien, A. S. Alvarado // Annual review of cell and developmental biology. — 2004. — т. 20. — с. 725—757.

67. *Weiler, K.* Hidden Surface Removal Using Polygon Area Sorting / K. Weiler, P. Atherton // SIGGRAPH Comput. Graph. — New York, NY, USA, 1977. — т. 11, № 2. — с. 214—222.
68. *Sternberg, S. R.* Grayscale morphology / S. R. Sternberg // Computer Vision, Graphics, and Image Processing. — 1986. — т. 35, № 3. — с. 333—355.
69. *Vincent, L.* Fast Grayscale Granulometry Algorithms / L. Vincent // Mathematical Morphology and Its Applications to Image Processing. Computational Imaging and Vision. т. 2. — Springer, Dordrecht, 1994. — с. 265—272.
70. *Asano, A.* Texture analysis using morphological pattern spectrum and optimization of structuring elements / A. Asano // Proceedings of 10th International Conference on Image Analysis and Processing. — 1999. — с. 209—214.
71. *Ljungqvist, M. G.* Image Analysis of Pellet Size for a Control System in Industrial Feed Production / M. G. Ljungqvist [и др.] // PLOS ONE. — 2011. — т. 6, № 10. — с. 1—10.
72. *Prodanov, D.* Automatic morphometry of synaptic boutons of cultured cells using granulometric analysis of digital images / D. Prodanov, J. Heeroma, E. Marani // Journal of Neuroscience Methods. — 2006. — т. 151, № 2. — с. 168—177.
73. *Shih, F.* Threshold Decomposition of Gray-Scale Morphology into Binary Morphology / F. Shih, O. Mitchell // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1989. — т. 11. — с. 31—42.
74. *Визильтер, Ю. В.* Проективные морфологии на базе операторов фильтрации и сегментации изображений, вычислимых методом динамического программирования / Ю. В. Визильтер, С. Ю. Желтов, Н. А. Ларетина // Вестник компьютерных и информационных технологий. — 2009. — т. 6. — с. 14—24.
75. *Khotanzad, A.* Invariant Image Recognition by Zernike Moments / A. Khotanzad, Y. H. Hong // IEEE Trans. Pattern Anal. Mach. Intell. — 1990. — т. 12, № 5. — с. 489—497.

76. Сидякин, С. В. Распознавание движущихся человеческих фигур на видеопоследовательности с использованием дисковой карты толщин / С. В. Сидякин, А. И. Егоров, И. К. Малин // Вестник МАИ. — 2014. — т. 21, № 5. — с. 157—166.
77. Eggers, H. Two Fast Euclidean Distance Transformations in Z2 Based on Sufficient Propagation / H. Eggers // Computer Vision and Image Understanding. — New York, NY, USA, 1998. — т. 69, № 1. — с. 106—116.
78. Cuisenaire, O. Fast Euclidean Distance Transformation by Propagation Using Multiple Neighborhoods / O. Cuisenaire, B. Macq // Computer Vision and Image Understanding. — New York, NY, USA, 1999. — т. 76, № 2. — с. 163—172.
79. Verbeek, P. W. Shading from Shape, the Eikonal Equation Solved by Grey-weighted Distance Transform / P. W. Verbeek, B. J. Verwer // Pattern Recognition Letters. — New York, NY, USA, 1990. — т. 11, № 10. — с. 681—690.
80. Shih, F. Y. The Efficient Algorithms for Achieving Euclidean Distance Transformation / F. Y. Shih, Y.-T. Wu // Transactions on Image Processing. — Piscataway, NJ, USA, 2004. — т. 13, № 8. — с. 1078—1091.
81. Saito, T. New algorithms for euclidean distance transformation of an n-dimensional digitized picture with applications / T. Saito, J.-I. Toriwaki // Pattern Recognition. — 1994. — т. 27. — с. 1551—1565.
82. Meijster, A. A General Algorithm For Computing Distance Transforms In Linear Time / A. Meijster, J. Roerdink, W. Hesselsink // Mathematical Morphology and Its Applications to Image and Signal Processing. — 2002. — с. 331—340.
83. Lotufo, R. A. Fast Multidimensional Parallel Euclidean Distance Transform Based on Mathematical Morphology / R. A. Lotufo, F. A. Zampirolli // Proceedings of the XIV Brazilian Symposium on Computer Graphics and Image Processing. — Washington, DC, USA : IEEE Computer Society, 2001. — с. 100—105. — (SIBGRAPI '01).
84. Felzenszwalb, P. F. Distance Transforms of Sampled Functions / P. F. Felzenszwalb, D. P. Huttenlocher // Theory of Computing. — 2004. — т. 8. — с. 415—428.

85. Maurer, C. R. A Linear Time Algorithm for Computing the Euclidean Distance Transform in Arbitrary Dimensions / C. R. Maurer, V. Raghavan, R. Qi // Information Processing in Medical Imaging, 17th International Conference, IPMI 2001, Davis, CA, USA, June 18-22, 2001, Proceedings. — 2001. — c. 358—364.
86. Linear time Euclidean distance transform algorithms / H. Breu [и др.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1995. — т. 17, № 5. — c. 529—533.
87. 2D Euclidean distance transform algorithms: A comparative survey. / R. Fabbri [и др.] // ACM Computing Surveys. — 2008. — т. 40, № 1. — 2:1—2:44.
88. Method for grain size determination in carbon steels based on the ultimate opening / C. A. Paredes-Orta [и др.] // Measurement. — 2019. — т. 133. — c. 193—207.
89. Outal, S. Controlling the ultimate openings residues for a robust delineation of fragmented rocks / S. Outal, S. Beucher // ECS10 — The 10th European Congress of Stereology and Image Analysis. — Milan, Italy, 2009.
90. Брянский, С. Условные морфологические меры сложности и информативности изображений / С. Брянский, Ю. Визильтер // Компьютерная оптика. — 2018. — т. 42, № 3. — с. 501—509.
91. Пытьев, Ю. Методы морфологического анализа изображений / Ю. Пытьев, А. Чуличков. — М.: Физматлит, 2010.
92. Haralick, R. M. Recursive opening transform / R. M. Haralick, S. Chen, T. Kanungo // Proceedings 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. — 1992. — с. 560—565.
93. Сидякин, С. Разработка алгоритмов построения морфологических спектров для анализа цифровых изображений и видеопоследовательностей. Диссертация на соискание учёной степени кандидата технических наук / Сидякин Сергей. — Вычислительный центр им. А.А. Дородницына Российской академии наук, Москва, Россия, 2013.
94. Locality-constrained Linear Coding for image classification. / J. Wang [и др.] // CVPR. — IEEE Computer Society, 2010. — с. 3360—3367.

95. *Wilkinson, M. H. F.* Generalized Pattern Spectra Sensitive to Spatial Information / M. H. F. Wilkinson // 16th International Conference on Pattern Recognition, ICPR 2002, Quebec, Canada, August 11-15, 2002. — 2002. — c. 21—24.
96. *Ayala, G.* Spatial Size Distributions: Applications to Shape and Texture Analysis / G. Ayala, J. Domingo // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2001. — т. 23, № 12. — c. 1430—1442.
97. *Braga-Neto, U.* Object-Based Image Analysis Using Multiscale Connectivity / U. Braga-Neto, J. Goutsias // IEEE Trans. Pattern Anal. Mach. Intell. — 2005. — т. 27, № 6. — c. 892—907.
98. *Zingman, I.* Size-density spectra and their application to image classification / I. Zingman, R. Meir, R. El-Yaniv // Pattern Recognition. — 2007. — т. 40, № 12. — c. 3336—3348.
99. *Flusser, J.* Moments and Moment Invariants in Pattern Recognition / J. Flusser, B. Zitova, T. Suk. — Wiley Publishing, 2009.
100. Shape reconstruction from moments: theory, algorithms, and applications / P. Milanfar [и др.] // Advanced Signal Processing Algorithms, Architectures, and Implementations X. т. 4116 / под ред. F. T. Luk. — International Society for Optics, Photonics. SPIE, 2000. — c. 406—416.
101. *Tahri, O.* Complex Objects Pose Estimation based on Image Moment Invariants / O. Tahri, F. Chaumette // Proceedings of the 2005 IEEE International Conference on Robotics and Automation. — 2005. — c. 436—441.
102. *Xin, Y.* Image Reconstruction with Polar Zernike Moments / Y. Xin, M. Pawlak, S. X. Liao // Pattern Recognition and Image Analysis, Third International Conference on Advances in Pattern Recognition, ICAPR 2005, Bath, UK, August 22-25, 2005, Proceedings, Part II. — 2005. — c. 394—403.
103. *Dai, M.* An efficient algorithm for computation of shape moments from run-length codes or chain codes / M. Dai, P. Baylou, M. Najim // Pattern Recognition. — 1992. — т. 25. — c. 1119—1128.
104. *Flusser, J.* Fast Calculation of Geometric Moments of Binary Images / J. Flusser // Pattern Recognition and Medical Computer Vision / под ред. M. Gengler. — Illmitz : ÖCG, 1998. — c. 265—274.

105. *Kotoulas, L. G.* Efficient hardware architectures for computation of image moments / L. G. Kotoulas, I. Andreadis // Real-Time Imaging. — 2004. — т. 10. — с. 371—378.
106. *Hu, M.-K.* Visual pattern recognition by moment invariants / M.-K. Hu // IRE Transactions on Information Theory. — 1962. — т. 8, № 2. — с. 179—187.
107. *Singer, M. H.* A general approach to moment calculation for polygons and line segments / M. H. Singer // Pattern Recognition. — 1993. — т. 26. — с. 1019—1028.
108. *Steger, C.* On the calculation of arbitrary moments of polygons / C. Steger // Technical Report FGBV-96-05. — Technische Universitat Munchen, 1996.
109. *Bui, T.* Font finder: Visual recognition of typeface in printed documents / T. Bui, J. P. Collomosse // 2015 IEEE International Conference on Image Processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015. — IEEE, 2015. — с. 3926—3930.
110. Large-Scale Visual Font Recognition / G. Chen [и др.] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014, Columbus, OH, USA, June 23-28, 2014. — IEEE Computer Society, 2014. — с. 3598—3605.
111. DeepFont: Identify Your Font from An Image / Z. Wang [и др.] // Proceedings of the 23rd Annual ACM Conference on Multimedia Conference, MM '15, Brisbane, Australia, October 26 - 30, 2015 / под ред. X. Zhou [и др.]. — ACM. — с. 451—459.
112. *Patel, C.* Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study / C. Patel, A. Patel, D. Patel // International Journal of Computer Applications. — 2012. — т. 55. — с. 50—56.
113. *Crammer, K.* On the Algorithmic Implementation of Multiclass Kernel-based Vector Machines / K. Crammer, Y. Singer // Journal of Machine Learning Research. — 2001. — т. 2. — с. 265—292.
114. 2008-2014. ПАРАТАЙП. Цифровые шрифты. Каталог / Э. Якупов [и др.]. — Москва, 2015.
115. Морфогенез у планарий *Dugesia tigrina* / И. Шейман [и др.] // Онтогенез. — 2004. — т. 35(4). — с. 285—290.