

Московский государственный университет имени М.В.Ломоносова

На правах рукописи

Бакина Ирина Геннадьевна

**Морфологическое сравнение изображений
гибких объектов на основе циркулярных
моделей при биометрической идентификации
личности по форме ладони**

05.13.17 – Теоретические основы информатики

ДИССЕРТАЦИЯ

на соискание учёной степени

кандидата физико-математических наук

Научный руководитель

д. т. н., проф.

Местецкий Леонид Моисеевич

Москва – 2011

Содержание

Введение	5
Глава 1. Задача сравнения формы гибких объектов и известные методы её решения	13
1.1. Задача сравнения формы гибких объектов	13
1.2. Методы сравнения формы гибких объектов	14
1.3. Задача распознавания формы ладони	31
1.4. Методы распознавания формы ладони	34
1.4.1. Распознавание по признакам геометрии ладони	36
1.4.2. Распознавание на основе контурной информации	40
1.5. Постановка задач исследования	43
1.5.1. Сравнение формы гибких объектов	43
1.5.2. Сравнение формы ладоней	45
1.6. Выводы к первой главе	48
Глава 2. Циркулярная модель гибкого объекта и методы её идентификации и подгонки	50
2.1. Модель гибкого объекта	50
2.2. Метод построения силуэта модели	53
2.2.1. Алгоритм парных пересечений	54
2.2.2. Восстановление силуэта модели по силуэтам фиксированных компонент	57
2.3. Метод идентификации (разметки) модели	63
2.4. Сравнение модели и изображения объекта	67
2.5. Алгоритм решения прикладных задач	71
2.6. Выводы ко второй главе	73

Глава 3. Алгоритмы идентификации и подгонки моделей гибких объектов при распознавании формы ладони	75
3.1. Модель ладони	75
3.1.1. Описание ладони гибким объектом	75
3.1.2. Параметризация модели	77
3.1.3. Построение силуэта ладони	77
3.2. Алгоритм идентификации модели ладони	79
3.2.1. Определение ветвей — кандидатов в пальцы	79
3.2.2. Определение кончиков и оснований пальцев	80
3.2.3. Определение точек поворотов пальцев	82
3.3. Алгоритм сравнения формы ладоней	87
3.3.1. Параметризация трансформаций модели	88
3.3.2. Начальная укладка моделей	89
3.3.3. Подгонка моделей	91
3.4. Выводы к третьей главе	94
Глава 4. Система распознавания личности по форме ладони и голосовому паролю	96
4.1. Многомодальный подход к распознаванию	96
4.2. Описание и общая схема двумодального подхода	97
4.3. Фильтрация на основе голосового пароля	99
4.4. Описание установки и комплекса программ	102
4.4.1. Экспериментальная установка	102
4.4.2. Программная реализация	103
4.5. Вычислительные эксперименты	105
4.5.1. Распознавание личности по форме ладони	106
4.5.2. Разметка моделей ладоней	112

4.5.3.	Распознавание личности по голосовому паролю	114
4.5.4.	Двумодальный подход	115
4.6.	Выводы к четвёртой главе	116
Заключение	119
Литература	122

Введение

Морфологическое сравнение — сравнение изображений объектов, дающее количественную оценку сходства этих объектов по их форме. Необходимость морфологического сравнения объектов возникает во многих прикладных задачах: в системах компьютерного зрения, при распознавании образов, в компьютерной графике. Примерами таких задач являются: распознавание рукописных символов, жестов руки, позы человека, объектов на аэрофотоснимке местности и т.д. Морфологическое сравнение даёт оценку сходства, на основе которой может решаться задача классификации формы объектов.

Для проведения сравнения форм объектов необходима некоторая мера/метрика, позволяющая оценить сходство этих форм. Она может быть задана как некоторая функция, определённая на так называемых признаках формы (признаковое сравнение). Под признаками формы понимаются различные топологические и метрические характеристики объекта: количество граничных контуров, площадь, периметр, диаметр и т.д. Другой способ задания меры сходства формы основывается на сравнении объектов при их совмещении (беспризнаковое сравнение). Человек, классифицируя формы в повседневной жизни, чаще всего делает это путём её сравнения с некоторой известной ему «эталонной» формой: «здание в форме буквы Н», «грушевидная форма лица», «облако в форме верблюда» и т.д. При этом он неявно «накладывает» эти формы друг на друга и оценивает их сходство. В этом случае сходство форм — это совпадение или их близость при совмещении.

Однако существуют так называемые *гибкие объекты*, форма которых может значительно меняться. Примерами гибких объектов являются фигуры животных и человека. Объектом интереса может быть и часть фигуры человека, например, ладонь. При этом изменения формы гибкого объекта не могут

быть абсолютно произвольными, а подчиняются определённым естественным ограничениям. В фигуре животного и человека конечности изгибаются в суставах, но при этом отдельные части фигуры (голень, бедро, плечо и т.д.) сохраняют свою форму. Для сравнения формы гибких объектов признаковое описание оказывается неинформативным, поскольку значения признаков могут сильно варьироваться при изменении формы. Оценка сходства на основе простого совмещения объектов также неприемлема, т.к. части объекта могут занимать различное взаимное положение (рисунок 1а). Однако такое совмещение возможно при определённом изменении положения компонент объекта (рисунок 1б). Таким образом, *оценку сходства изображений гибких объектов* можно было бы выполнить в положении наилучшего совмещения с учётом допустимых трансформаций этих объектов.

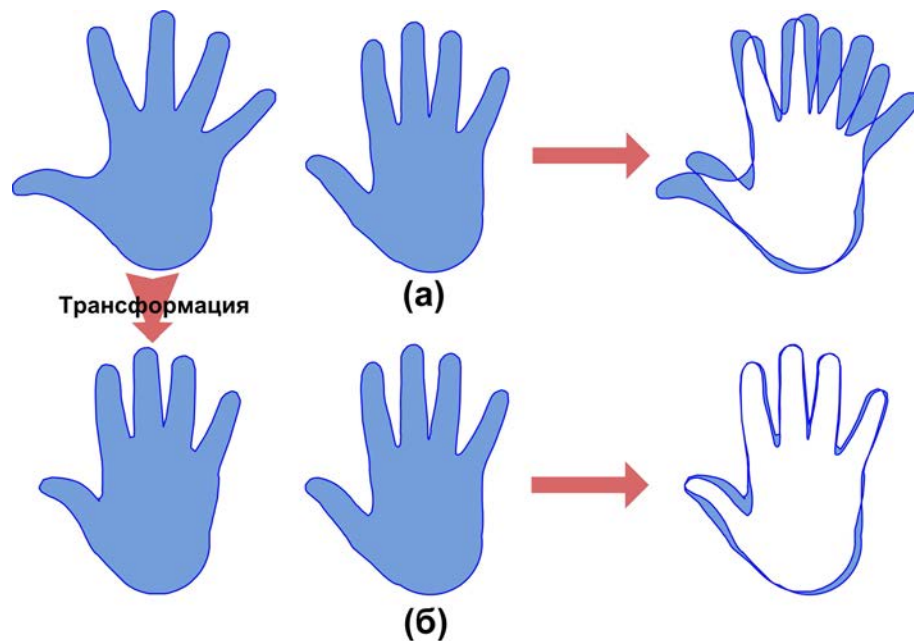


Рис. 1. Сравнение формы гибких объектов — ладоней человека: (а) простое совмещение силуэтов; (б) применение трансформаций к первой ладони и совмещение силуэтов после этого. В последней колонке тёмным цветом выделена симметрическая разность силуэтов ладоней.

Однако существующие модели и методы описания таких объектов пере-

менной формы развиты недостаточно. Традиционные способы представления формы объектов через задание границы их силуэтов не позволяют корректно описать допустимые трансформации объектов, поскольку в терминах локального изменения границ фигуры невозможно задать требования к сохранению формы её отдельных компонент.

Более адекватным инструментом для описания формы гибких объектов является циркулярное представление [1]. Суть циркулярного представления заключается в построении осевого графа (скелета) формы, с каждой точкой которого связан круг, вписанный в форму (рисунок 2.1). В подходе, описанном в [1], предлагается для каждой формы строить циркулярное представление на основе скелетизации и далее, используя это представление, преобразовывать форму в некоторое стандартное положение. Этот процесс называется нормализацией. После чего нормализованные фигуры сравниваются непосредственным наложением. Недостатки такого подхода:

1. высокие требования к качеству изображения объекта, которые не всегда легко выполнить на практике (необходимо, чтобы отсутствовали помехи и окклюзии);
2. низкая точность циркулярной аппроксимации объекта (в частности, по одному изображению весьма сложно сформировать описание возможных трансформаций).

Указанные недостатки существующих методов и практическая востребованность морфологического сравнения гибких объектов обуславливают **актуальность темы** данного исследования.

Целью диссертационной работы является разработка теоретического аппарата и практических методов моделирования, сравнения и классификации формы гибких объектов.

Предлагаемый подход к достижению цели:

1. Формирование эталонной модели гибкого объекта на основе циркулярного представления по нескольким изображениям без окклюзий, что должно обеспечить более высокую точность модели.
2. Сравнение наблюдаемой формы с эталонной моделью на основе подгонки этой модели под рассматриваемую форму. Этот путь позволяет классифицировать форму объектов в изображениях низкого качества, а также допускает наличие окклюзий.

Таким образом, **новизна подхода** состоит в использовании принципа «подгонки эталона под тестовый образец» вместо традиционной подгонки тестового образца под эталоны, представленные в нормализованном виде.

Реализация этого подхода ставит несколько **новых задач**:

1. Построение циркулярной модели гибкого объекта по нескольким эталонным изображениям.
2. Подгонка модели под наблюдаемое изображение (с окклюзиями) с учётом допустимых трансформаций.

Предлагаемый подход к решению задач состоит из следующих шагов:

1. Построение циркулярной модели гибкого объекта:
 - а. формирование набора эталонных изображений гибкого объекта без окклюзий;
 - б. описание группы допустимых трансформаций модели и её параметризация;

- в. идентификация параметров модели на основе эталонных изображений объекта.

2. Сравнение модели и тестовой формы:

- а. подгонка циркулярной модели под тестовую форму путём применения ряда допустимых деформаций к модели и поиска наилучшего совмещения силуэтов модели и теста;
- б. вычисление меры сходства силуэтов модели и тестовой формы в положении наилучшего совмещения.

На основе полученной меры сходства далее может быть построен метрический классификатор и решаться задача распознавания формы объектов.

Методы исследований. В работе использованы методы вычислительной геометрии, вычислительной математики, теории анализа изображений и компьютерной графики.

Для обоснования **достоверности** и **практической полезности** предлагаемого решения в диссертации рассматривается применение разработанных алгоритмов к решению практической задачи. В качестве прикладной задачи, на которой отрабатывались все шаги предлагаемого метода сравнения форм с окклюзиями, рассматривается задача биометрического распознавания личности по форме ладони. Как показано в диссертации, ладонь человека хорошо описывается предлагаемой моделью. Возможные трансформации ладони при её позиционировании в картинной плоскости видеокамеры описываются группой движений пальцев, что даёт возможность для построения достаточно простой и корректной математической модели. Для этой модели все элементы предлагаемого подхода доведены до рабочих алгоритмов и программной реализации.

Работа организована следующим образом. В главе 1 рассматривается задача сравнения формы гибких объектов, формулируется её общая постановка и приводится обзор существующих методов её решения. Также рассматривается применение этих методов к решению задачи сравнения формы ладоней. Указываются достоинства и недостатки методов. В конце главы формулируется общая постановка задач исследований. В главе 2 рассматривается общий подход к решению задачи морфологического сравнения изображений гибких объектов. Описывается параметрическая модель гибкого объекта и способ её построения, предлагается способ автоматической идентификации параметров модели и подгонки модели под заданную тестовую форму. Глава 3 иллюстрирует применение разработанного метода к решению конкретной прикладной задачи — задачи распознавания человека по форме его ладони. Глава 4 содержит описание экспериментальной установки и комплекса программ, реализующего предложенный подход в рамках системы двумодальной идентификации личности по форме ладони и голосовому паролю. Приводятся результаты экспериментов на реальных базах бинарных изображений ладоней и голосовых паролей. Далее следует заключение, содержащее основные выводы по данной работе.

На защиту выносятся следующие основные результаты и положения :

1. Модель гибкого объекта и метод идентификации её параметров (разметки) по серии бинарных изображений объекта в разных положениях.
2. Метод сравнения формы гибких объектов, основанный на подгонке размеченной эталонной модели под тестовое изображение.
3. Метод идентификации модели ладони по серии изображений.
4. Метод биометрической идентификации личности по форме ладони.

Апробация работы. Результаты работы докладывались на следующих конференциях:

- 7-ая международная научная конференция «Интеллектуализация обработки информации» ИОИ'08 (Алушта, 2008) [2];
- XVI международная научная конференция студентов, аспирантов и молодых учёных «ЛОМОНОСОВ-2009» (Москва, 2009) [3];
- всероссийская конференция «Математические методы распознавания образов» ММРО-14 (Суздаль, 2009) [4];
- 19-ая международная конференция по компьютерной графике и зрению «ГрафиКон'09» (Москва, 2009) [5];
- 8-ая международная научная конференция «Интеллектуализация обработки информации» ИОИ'10 (Пафос, 2010) [6];
- международная конференция по теории и приложениям компьютерного зрения "VISAPP—2011" (Португалия, 2011) [7];
- 2-ая научно-техническая конференция «Техническое зрение в системах управления» TVCS'2011 (Москва, 2011) [8];
- международная конференция по анализу и распознаванию изображений "ICIAР—2011" (Канада, 2011) [9];
- всероссийская конференция «Математические методы распознавания образов» ММРО-15 (Петрозаводск, 2011) [10].

Материалы диссертации использованы в работах по грантам РФФИ №№ 08-01-00670, 08-07-00305, 08-07-00270, 10-07-00609, 11-01-00783 и 11-07-00462.

Публикации. Материалы диссертации опубликованы в 11 печатных работах, из них 2 статьи в рецензируемых журналах [11, 12], 7 статей в сборниках трудов конференций [2, 4–7, 9, 10] и 2 тезиса докладов [3, 8]. В изданиях, рекомендованных ВАК РФ, опубликовано 2 статьи [9, 12].

Структура и объём диссертации. Диссертация общим объёмом 130 страниц состоит из введения, 4 глав, заключения и списка литературы из 73 наименований. В работе содержится 41 рисунок и 1 таблица.

Глава 1

Задача сравнения формы гибких объектов и известные методы её решения

1.1. Задача сравнения формы гибких объектов

Большую роль в восприятии и понимании мира человеком играет зрительная информация. Современное развитие компьютерной техники приводит к тому, что образуется большое количество изображений, которые нужно уметь быстро обрабатывать, и, что важно, находить среди всего этого многообразия изображения одинаковых или похожих объектов. Таким образом, возникает задача сравнения объектов на разных изображениях. При этом форма объекта, наряду с другими его характеристиками (например, текстурой и цветом), играет важную роль при принятии решения. Однако большинство объектов реального мира имеют форму, изменяющуюся в определённых пределах. Поэтому актуальной становится задача сравнения формы так называемых «гибких» объектов. Например, силуэтов животных и человека, форм микроорганизмов, технических объектов и т.д.

Однозначного определения понятия формы не существует [1]. Интуитивно человек понимает форму плоского объекта как некоторую область в двумерном пространстве, ограниченную непрерывной кривой. Двумерное пространство является моделью картинной плоскости, предстающей перед глазами человека. В толковом словаре русского языка [13] приводится следующее определение, наиболее подходящее для использования его в контексте анализа и обработки изображений: *форма* — это внешние очертания, наружный вид предмета. Близким по значению является слово «фигура». *Фигура* —

внешнее очертание, форма чего-либо. В данной работе слова «форма», «фигура», «силуэт» мы будем использовать как синонимы. При этом в качестве строгого математического определения фигуры будем рассматривать следующее, предложенное в [1]: *Фигурой* называется связная замкнутая область на плоскости, ограниченная конечным числом непересекающихся жордановых кривых.

Обычно задача распознавания формы гибких объектов на основе сравнения, где под *сравнением* понимается получение количественной оценки сходства/различия объектов, ставится следующим образом. Имеется либо бинарное изображение исследуемого объекта, либо его контур. При этом необходимо разработать метод сравнения таких представлений, устойчивый к различным трансформациям объекта. В простейшем случае, к аффинным преобразованиям и масштабированию. В более сложных случаях рассматриваются трансформации, связанные с изменением формы отдельных частей объекта и окклюзиями. Рассмотрим известные методы, применяющиеся для решения такого рода задач.

1.2. Методы сравнения формы гибких объектов

На сегодняшний день разными авторами предложено большое число подходов к описанию и сравнению формы объектов [14, 15]. На рисунке 1.1 представлена общая схема, взятая из работы [14], в которой существующие подходы к описанию формы сгруппированы в зависимости от типа используемых признаков. Соответственно, за каждым способом описания стоит свой метод сравнения формы. Эти методы широко используются сегодня для решения конкретных прикладных задач. Каждый из них ориентирован на распознавание формы определённого класса объектов и имеет свои достоинства

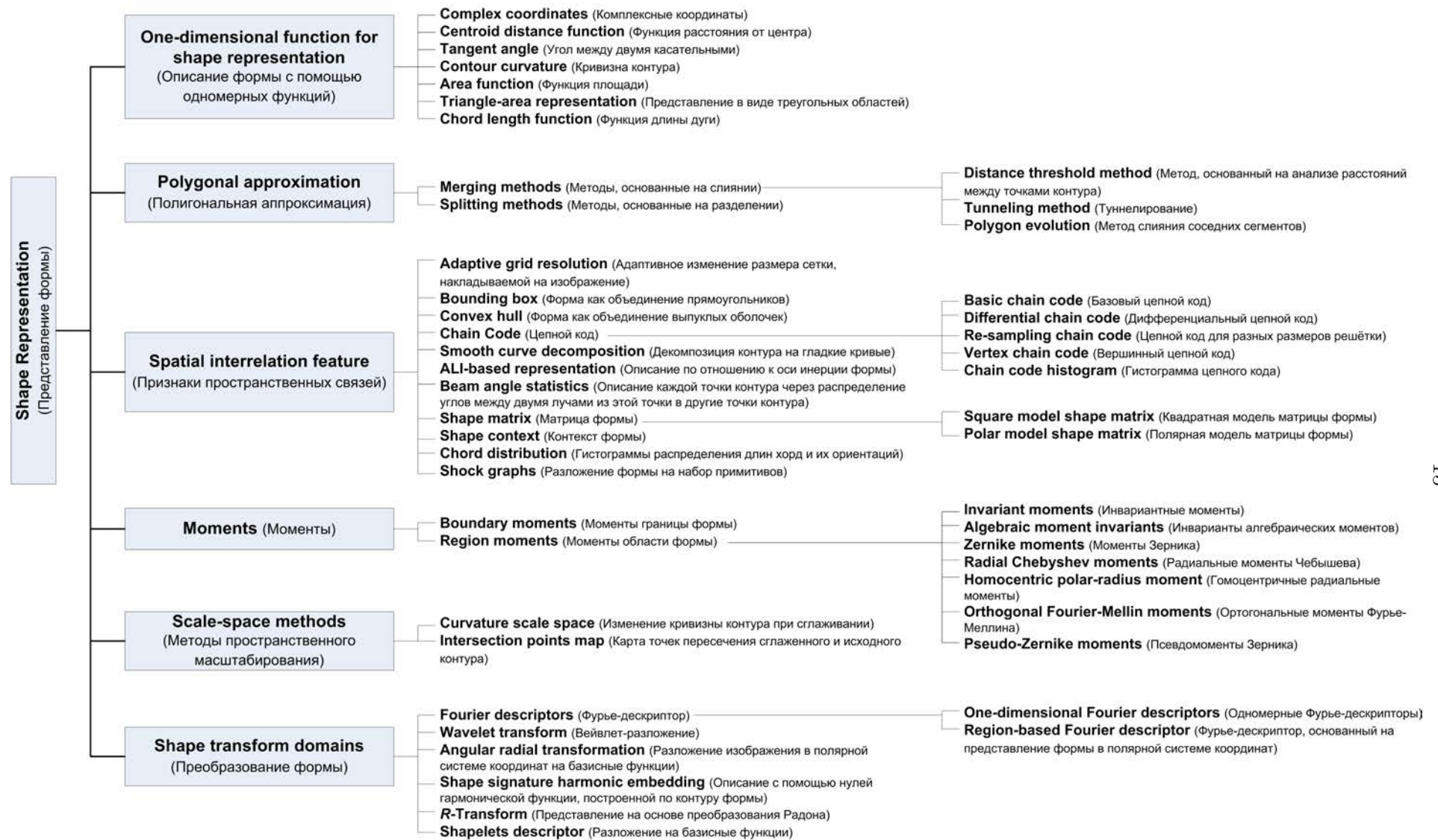


Рис. 1.1. Подходы к описанию формы объектов.

и недостатки. Рассмотрим применимость этих методов для решения задачи сравнения формы гибких объектов.

Дескрипторы группы **”One-dimensional function for shape representation”** (представление формы с помощью одномерных функций) в основном служат для описания формы самой по себе. Они часто используются на этапе начальной предобработки перед вычислением основных признаков формы, например, вейвлет- и Фурье-дескрипторов. Представителями данного класса дескрипторов являются: комплексные координаты; функция расстояния до центра тяжести формы; значения угла между двумя касательными к границе формы при её обходе; функция кривизны; функция площади; описание в виде совокупности площадей треугольников, образованных точками границы формы; функция длины хорды и т.д. Обычно эти описания инвариантны к параллельному переносу (т.к. используются относительные измерения) и масштабированию (за счёт нормализации). Чтобы компенсировать возможные различия в ориентации формы, при сравнении контуров проводится их подгонка. Кроме того, построение таких описаний вычислительно просто.

Однако такие дескрипторы чувствительны к шумам, и небольшие изменения границы приводят к большим ошибкам при подгонке и сравнении форм. Поэтому эти описания напрямую практически не используются. Они также неприменимы для сравнения формы гибких объектов, поскольку значения признаков могут значительно меняться при различных трансформациях гибкого объекта и окклюзиях.

”Polygonal Approximation” (полигональная аппроксимация) позволяет игнорировать небольшие изменения границы объекта и регистрировать только общую информацию. Существует два подхода к построению такого дескриптора — методом слияния и разделения.

В методе слияния рассматриваются две соседние точки контура, и далее

к ним последовательно добавляются точки контура до тех пор, пока отклонение текущих точек от прямой линии невелико. Затем полученный набор точек объединяется в одну сторону будущего многоугольника.

В методе разделения сначала строится одна линия, соединяющая две точки контура. Далее вычисляется расстояние от точек контура до этой линии. Если оно превышает некоторый порог, контур разбивается на две части точкой, в которой отклонение от прямой было наибольшим. Затем для каждого подконтура указанная процедура повторяется. Процесс завершается, когда для каждого подконтура отклонение от прямой линии, ему соответствующей, не превосходит заданный порог.

Полигональная аппроксимация уменьшает сложность описания границы формы, приводит к устранению шумов, а вершинами получаемого многоугольника являются исходные точки контура. Однако, как и одномерные функции, она чаще рассматривается как этап предобработки и подходит для описания объектов фиксированной формы. Полигональная аппроксимация не предназначена для описания и сравнения формы гибких объектов.

Дескрипторы группы **”Spacial interrelation feature”** (признаки пространственных связей) описывают форму или её контур на основе взаимного положения её пикселей или кривых. Обычно для этого используются такие геометрические признаки, как длина, кривизна, относительная ориентация, площадь, расстояние и т.д. Рассмотрим подход, называемый в литературе **”Shape Context”**, который хорошо зарекомендовало себя на практике как мощное средство для распознавания формы объектов.

Shape Context был предложен в работе [16] как метод сравнения контуров объектов, основанный на поиске соответствия между точками двух контуров. Каждая точка контура описывается гистограммой распределения расстояний от этой точки до остальных точек контура в полярной системе

координат. На рисунке 1.2а,б представлены два контура одной буквы, а на рисунках 1.2г,д,е — гистограммы, построенные для точек \circ , \triangleleft и \diamond соответственно. Рисунок 1.2в содержит разбиение полярной системы координат на ячейки, используемые при построении гистограммы (5 ячеек для $\log r$ и 12 — для угла θ). Результат сопоставления точек контуров изображён на рисунке 1.2ж.

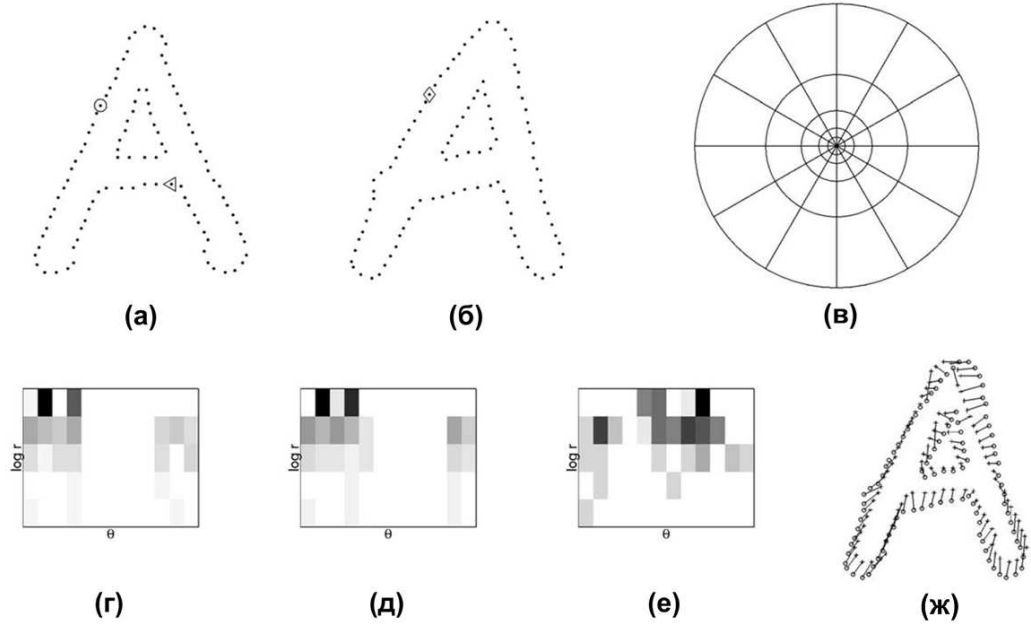


Рис. 1.2. (а),(б) Контура сравниваемых букв; (в) разбиение полярной системы координат на ячейки, используемые при построении гистограммы; (г),(д),(е) гистограммы, построенные для точек \circ , \triangleleft и \diamond ; (ж) найденное соответствие точек контуров.

В качестве меры сходства точки p_i одного контура и точки q_j другого рассматривается следующая величина: $C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^K \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$, где $h_i(k)$ — гистограмма распределения расстояний для точки p_i первого контура, $h_j(k)$ — для точки q_j второго, а $k = 1, \dots, K$ — ячейки, на которые разбивалась плоскость при построении гистограмм (рисунок 1.2в). Предлагаемый авторами метод позволяет найти соответствие между точками обоих контуров, при котором суммарное сходство точек контуров является наибольшим. Такой подход инвариантен к параллельному переносу контуров, небольшим

геометрическим искажениям и выбросам. Авторы также указывают на возможность адаптирования метода к случаям поворота и масштабирования контуров. Предлагаемый алгоритм также применим к распознаванию силуэтов трёхмерных объектов. В работе проведены эксперименты по распознаванию силуэтов объектов, рукописных цифр, товарных знаков и т.д., взятых из крупных баз, таких как *MNIST* [17], *COIL-20* [18] (силуэты трёхмерных объектов), *MPEG-7* [19] и др. Распознавание проводилось модифицированным методом k ближайших соседей. На *MNIST* и *MPEG-7* было показано качество распознавания, превосходящие существующие на тот момент подходы.

Shape Context является хорошим дескриптором формы объектов, если эта форма меняется незначительно. Эксперименты, представленные в работе [16], проводятся именно на таких объектах (изображения птиц, рыб и т.д.). Признаки Shape Context отражают взаимное расположение точек контура формы. Но поскольку для гибкого объекта допустимо изменение положения отдельных частей по отношению к другим его частям, эти признаки могут сильно варьироваться для одного и того же объекта. В результате, исходный гибкий объект и он же после трансформации оказываются совершенно непохожими. Поэтому данный подход для сравнения формы гибкого объекта практически непригоден.

Дерево формы. Авторы [20] предлагают метод сравнения контуров объектов, позволяющий оценивать сходство силуэтов как заранее сегментированных объектов, так и обнаруживать объекты заданной формы на изображениях реальных сцен. Основная идея подхода заключается в построении дерева, описывающего структуру и взаимное расположение точек контура. Пусть $A = (a_1, \dots, a_n)$ есть некоторый разомкнутый контур, составленный из n точек. Сначала определяется средняя точка контура a_i (например, с номером $\lceil n/2 \rceil$), и строится корень дерева, в который сохраняется информация о поло-

жении точки a_i по отношению к концевым точкам контура — a_1 и a_n . Далее исходный контур разбивается на два $A_1 = (a_1, \dots, a_i)$ и $A_2 = (a_i, \dots, a_n)$, и для каждого повторяется та же процедура. Полученные вершины становятся дочерними у корня и т.д. На рисунке 1.3 приведён пример такого дерева. По такому дереву контур восстанавливается единственным образом. Моделирование деформаций проводится путём изменения информации об относительном положении точек, хранящейся в вершинах дерева.

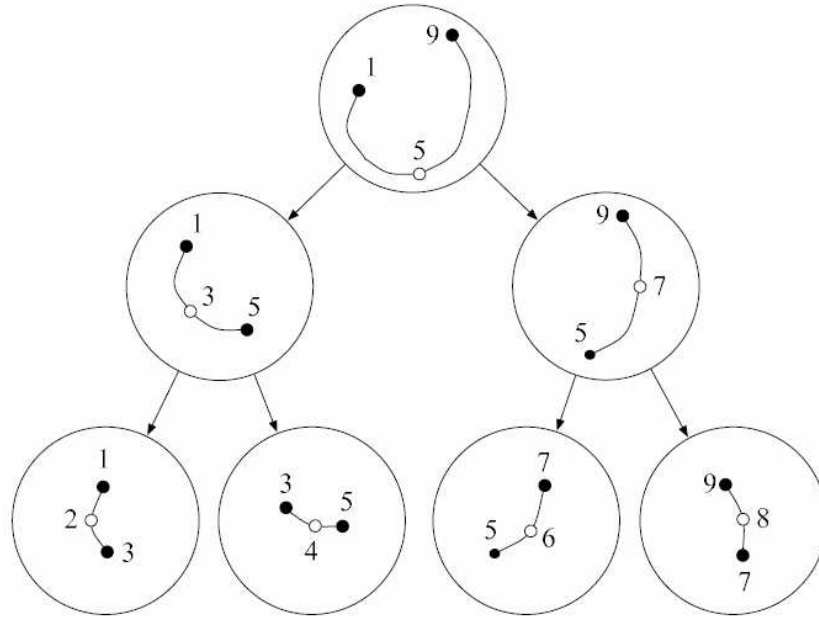


Рис. 1.3. Дерево контура, состоящего из 9 точек. Жирным цветом выделены концевые точки подконтура. Каждая вершина дерева хранит информацию о положении средней точки подконтура по отношению к его концевым вершинам.

Сравнение двух контуров $A = (a_1, \dots, a_n)$ и $B = (b_1, \dots, b_n)$ осуществляется следующим образом. Рассматривается средняя точка контура A и ищется наиболее похожая на неё точка контура B — точка b_j . Далее контур A разделяется средней точкой на два подконтура A_1 и A_2 , а контур B — на B_1 и B_2 , и аналогичным образом ищется соответствие точек контуров A_1 и B_1 , A_2 и B_2 . В качестве меры сходства точек a_i и b_j рассматривается квадрат рас-

стояния Прокруста¹ (Procrustes distance) между тройками точек (a_1, a_i, a_n) и (b_1, b_i, b_n) .

Задача поиска соответствия точек двух контуров решается методом динамического программирования. В случае, когда на одном из контуров присутствует/отсутствует некоторая часть в отличие от другого контура, мера сходства таких участков задаётся специальной функцией.

В случае замкнутых контуров строится несколько деревьев. Выбираются разные начальные точки контура, и строится отдельное дерево для каждого такого выбора.

Работа содержит результаты экспериментов на таких базах, как *Swedish Leaf Database* [22], *Brown Database* [23] и *MPEG-7*, а также сравнение с другими методами. Некоторые из них, такие как Фурье-дескрипторы и CSS, рассматриваются ниже.

Рассмотренный метод может быть адаптирован для сравнения формы гибких объектов без окклюзий. Для этого необходимо разделить контур на участки, которые остаются неизменными при выполнении трансформаций (считая, что эта информация известна). И далее строить дерево формы таким образом, чтобы первоначальные разрезы контура совпадали с концами этих участков. Тогда деревья исходного и трансформированного объекта будут содержать лишь небольшое число непохожих вершин, соответствующих

¹ Прокрустов анализ [21] — один из статистических методов, применяемый для анализа распределения множества форм. Считается, что возможные формы объекта принадлежат одному классу эквивалентности, если нивелировать все различия, связанные с аффинными преобразованиями и масштабированием. Соответственно, в рамках этого анализа предлагается алгоритм, позволяющий выполнить такую нормализацию формы. Пусть после выполнения нормализации два контура формы описываются следующими последовательностями точек: $((x_1, y_1), \dots, (x_n, y_n))$ и $((u_1, v_1), \dots, (u_m, v_m))$, и указано соответствие между точками одного и второго контура. Тогда расстояние Прокруста вычисляется по формуле (считая, что точке (x_1, y_1) одного контура соответствует точка (u_1, v_1) другого, а точке (x_n, y_n) — точка (u_m, v_m)):
$$d = \sqrt{(x_1 - u_1)^2 + (y_1 - v_1)^2 + \dots + (x_n - u_n)^2 + (y_m - v_m)^2}.$$

концам участков постоянной формы. Однако для случая окклюзий, связанных с перекрытием одной части объекта другой его частью, довольно сложно провести такую адаптацию (т.к. невозможно корректно сегментировать объект), и в этом случае метод неприменим.

Признаки пространственных связей часто используются для сжатия исходных изображений, а также облегчения процедуры сравнения формы объектов.

К группе дескрипторов **”Moments”** (моменты) относятся описания, которые основаны на вычислении моментов формы — моментов Хана, Чебышева, Ху, Зерника и т.д. *Моменты Ху* были предложены в работе [24] и представляют собой семь величин, вычисляемых для бинарного изображения объекта. Они инвариантны к аффинным преобразованиям и масштабированию. Обычный момент изображения вычисляется по следующей формуле: $m_{pq} = \sum x^p y^q f(x, y)$, где суммирование ведётся по всем точкам рассматриваемой области, а $f(x, y) = 1$, если точка (x, y) принадлежит объекту, и $f(x, y) = 0$, если соответствующая точка является точкой фона. Тогда m_{00} — площадь области, а $m_x = m_{10}/m_{00}$, $m_y = m_{01}/m_{00}$ — координаты центра тяжести. Инвариантность к параллельному переносу достигается за счёт вычисления центральных моментов: $\mu_{pq} = \sum (x - m_x)^p (y - m_y)^q f(x, y)$. Для сохранения инвариантности к поворотам и масштабированию рассматриваются нормализованные моменты: $\eta_{pq} = \mu_{pq} / \mu_{00}^\gamma$, где $\gamma = (p+q)/2 + 1$. Первые три момента Ху выглядят следующим образом: $\phi_1 = \eta_{20} + \eta_{02}$, $\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2$, $\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{12} - \eta_{03})^2$.

Таким образом, изображение описывается вектором из семи компонент: $\Phi = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7)$. Мерой сходства эталонного изображения Φ и тестового изображения $O = (o_1, o_2, o_3, o_4, o_5, o_6, o_7)$ является величина $\|O - \Phi\|$, где $\|\cdot\|$ обычно является евклидовым расстоянием. Распознавание формы

проводится методом одного ближайшего соседа.

Методы сравнения формы, основанные на вычислении моментов, обычно инвариантны к аффинным преобразованиям и масштабированию. Кроме того, они малочувствительны к небольшим изменениям формы объекта (например, искажениям его границы при наличии шумов). Однако для описания формы гибких объектов этот подход малопригоден по тем же причинам, что и признаки пространственных связей, рассмотренные выше.

В **”Scale-space methods”** (пространственное шкалирование) рассматриваемая кривая (граница формы) погружается в некоторое непрерывное семейство кривых $\Gamma_\sigma : \sigma \geq 0$. Основная идея этих методов заключается в том, что исходная кривая $\Gamma = \Gamma_0$ всё более и более упрощается, а незначительные элементы кривой исчезают при росте параметра σ . При разных значениях σ удаётся отделить общие элементы от индивидуальных особенностей кривой. Этот подход активно применяется на практике благодаря своей возможности сглаживать контура и устранять влияние шумов.

Представителем данного подхода является *CSS-описание* (Curvature Scale Space) границы контура, предложенное в работе [25]. Модифицированное *CSS*-представление было принято в качестве стандарта описания формы объекта в *MPEG-7* [26]. Фактически, *CSS*-описание хранит информацию о точках, в которых кривизна границы объекта меняет свой знак, т.е. переходит через 0.

Пусть $\Gamma = \{x(u), y(u) | u \in [0, 1]\}$ есть исходный контур объекта. Если сделать ядерное сглаживание для каждой компоненты, $x(u)$ и $y(u)$, с ядром Гаусса $G(u, \sigma)$, то получим более гладкую кривую. При увеличении параметра σ кривая Γ превращается в выпуклую кривую, и число точек, в которых кривизна границы переходит через 0, становится равным 0. В результате можно построить бинарное изображение (*CSS*-изображение), где по оси абсцисс

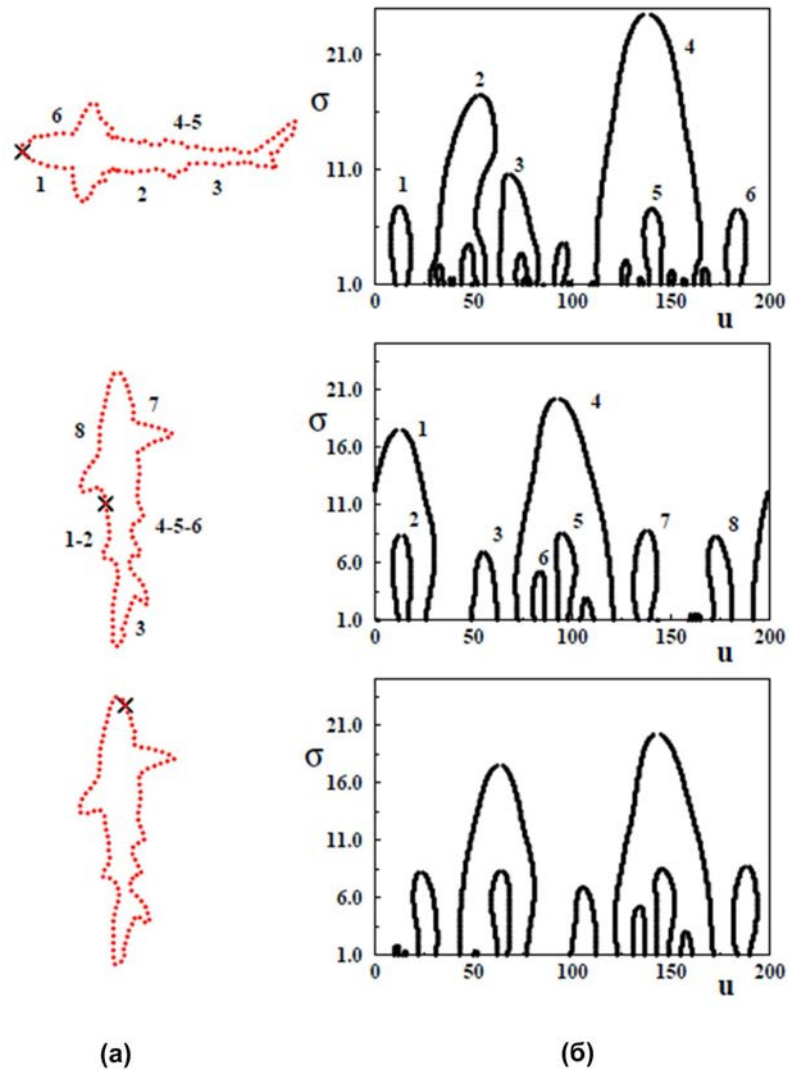


Рис. 1.4. (а) Граница рассматриваемого объекта при различной ориентации и выборе начальной точки (обозначена \times); (б) построенное для него CSS-изображение.

отложены значения u , а по оси ординат — σ . Точками отмечены значения u , в которых кривизна границы меняет свой знак. Пример CSS-изображения представлен на рисунке 1.4.

Пересечение любой горизонтальной линии с линиями CSS-изображения определяет положение точки перехода кривизны контура через 0. При построении такого представления исходные контура сначала дискретизируются (формируется 200 точек границы) и масштабируются (периметр контура

имеют фиксированную длину). В качестве дескриптора формы рассматривается положение точек локальных максимумов кривых CSS-изображения.

При сравнении двух контуров производится сопоставление точек максимумов CSS-описания. Мера сходства контуров есть сумма расстояний между сопоставленными точками плюс значение вертикальных координат для точек, которым не нашлось соответствующей точки в другом контуре.

Эксперименты проводились на базе изображений морских животных и листьев разных видов хризантем. При этом для каждого анализируемого объекта формировалось два списка — список похожих объектов из базы, составленный человеком, и список, полученный предложенным алгоритмом. Далее эти списки сравнивались. Оказалось, что люди по-разному выбирают похожие объекты. При этом ранжирование, полученное алгоритмом, согласуется с оценками небольшой группы людей. Однако более поздние исследования [27] показали, что CSSD-представление (модификация CSS для MPEG-7) в некоторых случаях значительно уступает существующим альтернативным описаниям, таким как Фурье-дескрипторы, Shape Context или моменту X_u .

Степень влияния трансформаций на изменение CSS-представления зависит от того, насколько меняется кривизна контура при их выполнении, сколько новых точек перехода кривизны границы через 0 при этом возникает. Однако рассматриваемый метод не подходит для сравнения гибких объектов с окклюзиями, поскольку они способны существенно изменить CSS-картину.

Группа методов **”Shape-transform domains”** (области преобразования формы) основана на анализе трансформаций обнаруженного объекта или всего изображения. Описанием формы являются коэффициенты–параметры трансформаций: Фурье-дескрипторы, вейвлет-разложение, R -преобразование (преобразование Радона) и т.д. Характерным представителем данной группы являются Фурье-дескрипторы [27, 28], которые используются уже на протя-

жени 40 лет. Они легко вычисляются, устойчивы к шумам и компактны. Для их построения достаточно получить некоторое одномерное описание границы контура (примеры рассмотрены выше), к которому применяется дискретное преобразование Фурье. Вектор нормированных коэффициентов Фурье-разложения называется Фурье-дескриптором формы.

Аналогично методам пространственного шкалирования данные методы позволяют сглаживать границу объекта, контролируя при этом точность аппроксимации, которая определяется числом рассматриваемых коэффициентов разложения. Однако они чувствительны к изменению формы всего объекта и, тем более, появлению окклюзий. Такие дескрипторы не могут быть напрямую использованы для описания формы гибких объектов.

Скелеты. Ещё одним подходом к описанию и распознаванию формы объектов является использование скелетного представления [29–32]. В [29] для исходного бинарного изображения строится скелет, который очищается от шумовых ветвей, после чего каждая цепочка из ветвей графа (на которые граф разбивается точками ветвления) взвешивается. Терминальные цепочки считаются менее важными по сравнению с внутренними, поскольку обычно они несут информацию о граничных особенностях объекта, которые несущественны при распознавании категорий объектов. Поэтому все терминальные цепочки графа получают наименьший вес, а чем более «внутренней» является цепочка графа, тем больший вес она имеет. С каждой вершиной графа связан круг максимального радиуса, вписанный в силуэт объекта, с центром в этой вершине.

Далее по скелету строится граф связности, в котором множество вершин есть терминальные вершины и точки ветвления скелетного графа, а каждой цепочке скелетного графа сопоставлено ребро в графе связности. На рисунке 1.5 приведён пример скелетного графа лошади и соответствующего ему

графа связности.

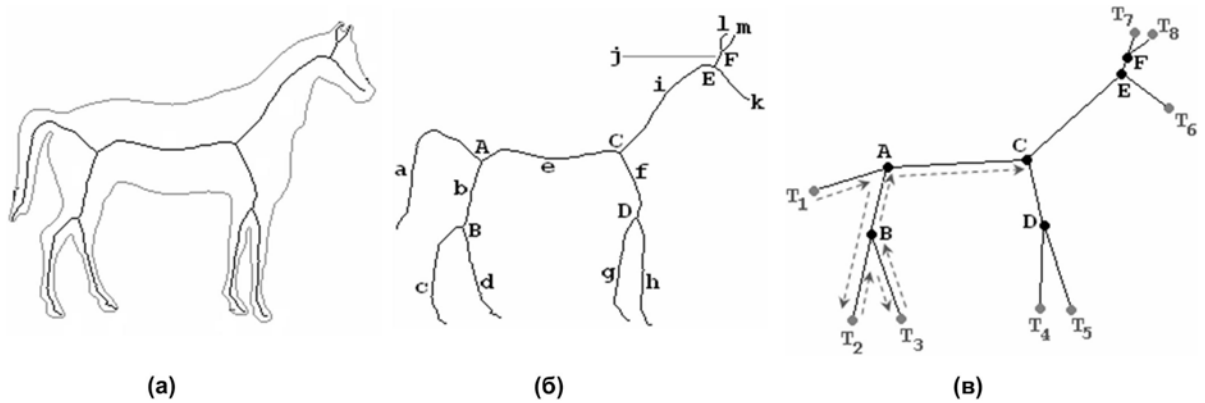


Рис. 1.5. (а) Силуэт лошади; (б) его скелетный граф и (в) граф связности.

Пусть есть два графа связности G_1 и G_2 , которые необходимо сравнить. Сначала проводится трассировка каждого графа против часовой стрелки и строится последовательность из радиусов кругов, соответствующих вершинам следа трассировки. При вычислении сходства двух элементов последовательностей учитывается как радиус круга, так и вес соответствующего ребра. Далее ищется сопоставление элементов последовательностей друг другу, при котором суммарное сходство всех вершин последовательностей является наибольшим. Поскольку в качестве начальной вершины трассировки графа может быть выбрана произвольная, последовательности сравниваются с учётом циклического сдвига.

Идея разметки и сопоставления графов также используется в работах [30, 31]. В работе [32] сравнение двух объектов происходит путём «обрезания» исходных скелетных графов до тех пор, пока они не станут изоморфными. При этом мера сходства определяется тем, сколько и каких ветвей пришлось удалить.

Методы, основанные на описании формы объекта с помощью скелета, лучше всего подходят для распознавания структуры объекта. Сам по себе скелет не позволяет учитывать допустимые преобразования объекта, что су-

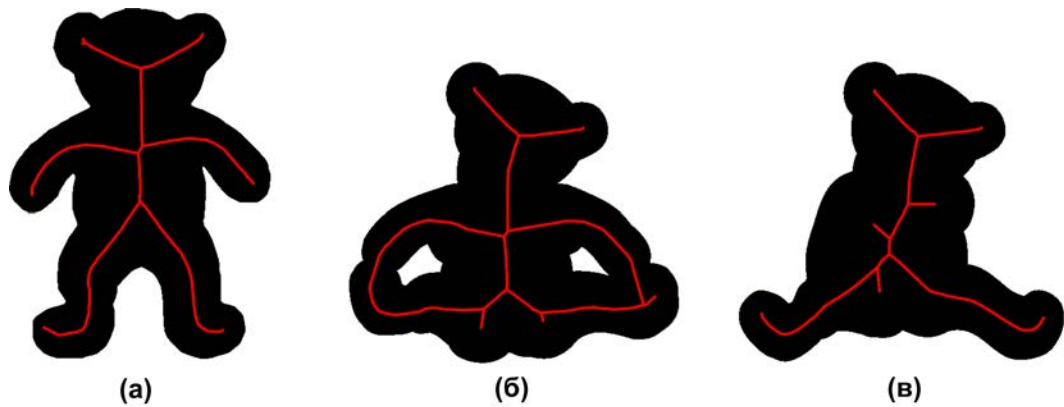


Рис. 1.6. (а) Скелет изображения исходного гибкого объекта; (б),(в) скелет того же гибкого объекта после трансформаций.

цественно при анализе гибких объектов. Более того, при трансформациях форма объекта может измениться настолько, что с точки зрения скелетных графов две формы (исходная и полученная) будут практически несравнимы (рисунок 1.6).

Описание формы через набор примитивов более всего подходит для описания формы гибких объектов. Такой подход присутствует в нескольких работах [1, 33, 34]. Рассмотрим их более подробно.

FORMS. Интересный подход к распознаванию формы деформируемых объектов представлен в работе [33]. В методе *FORMS* предлагается раскладывать исходную форму на набор составных компонент. Для описания компоненты используются два деформируемых примитива: червяк и круг (рисунок 1.7). Авторами предлагается метод разложения исходной формы на набор таких примитивов и способ их описания.

Сначала собирается эталонная база, которая содержит набор возможных компонент и информацию о том, к какому объекту данная компонента относится. Например, «передняя лапа собаки», «плавник рыбы», «голова пингвина» и т.д. Предполагается, что параметры примитивов подчиняются многомерному распределению Гаусса, поэтому в качестве меры сходства двух

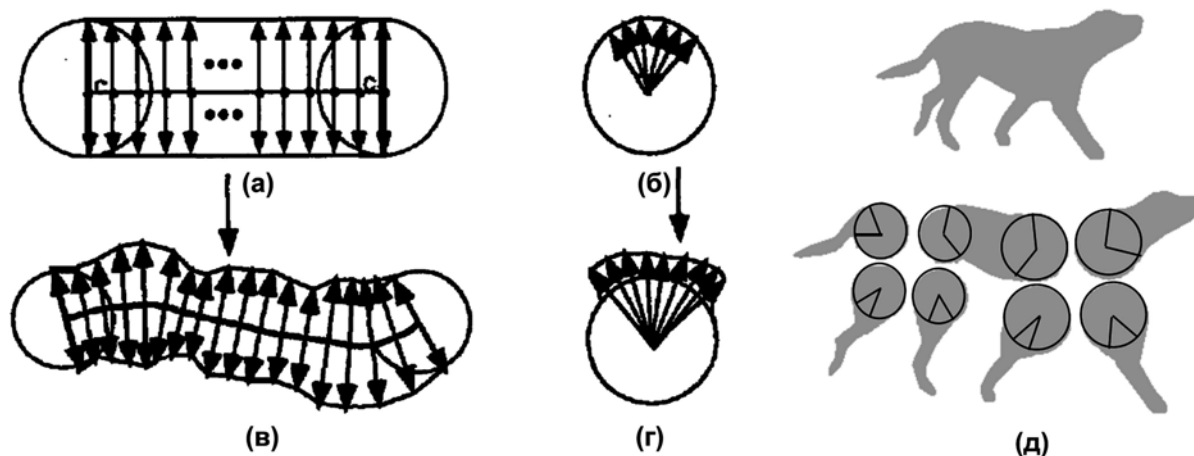


Рис. 1.7. Прimitives (а) червяк и (б) круг, а также (в),(г) их деформации. (д) Разложение силуэта собаки на компоненты.

компонент рассматривается их совместная вероятность.

Процесс распознавания тестовой формы состоит из двух шагов. На первом шаге тестовая форма раскладывается на составные компоненты, для каждой из которых определяется наиболее похожая из базы эталонов. В результате формируется список объектов, у которых хотя бы одна компонента попала в группу похожих. На втором шаге из полученного списка последовательно перебираются все объекты и сравниваются с тестовой формой. При этом проводится сопоставление скелетных графов форм и, соответственно, компонент эталона и теста, при котором совместная вероятность (равная произведению совместных вероятностей сопоставленных компонент) является наибольшей. Результатом выполнения второго шага является набор оценок сходства тестовой формы и каждого из объектов. Считается, что тестовая форма является силуэтом того же объекта, который оказался наиболее похожим с точки зрения полученных оценок. Эксперименты проводились на собственной базе изображений 35 категорий объектов (люди, ладони, животные, рыбы, насекомые, листья и т.д.). Полученные оценки сходства категорий объектов показали, что объекты отличимы, хотя некоторые категории довольно похожи друг

на друга.

Предложенный авторами метод ориентирован на распознавание силуэтов одушевлённых объектов (animated objects), в отличие от неодушевлённых (man-made), таких как дома, машины и т.д. Достоинством метода является возможность сравнения силуэтов деформируемых объектов. Например, человека или животных в разных позах.

Идея разложения исходного объекта на ряд компонент и их сопоставление при сравнении присутствует также в работах [1, 34]. В [34] предлагается аппроксимировать каждую компоненту жирной Б-сплайновой кривой. Для каждой такой кривой строится график её ширины, а в качестве меры сходства двух кривых рассматривается отношение пересечения областей под графиками к их объединению.

В [1] предлагается описывать объект циркулярным графом. В целом, циркулярный граф представляет собой осевой граф (иногда его называют скелетом формы), с каждой точкой которого связан круг, отражающий ширину объекта (рисунок 2.1). Циркулярный граф может быть построен для произвольного бинарного изображения объекта. Такое представление объединяет в себе скелетное и контурное, поскольку осевой граф является скелетом формы, а огибающая семейства кругов является силуэтом объекта. Кроме того, циркулярное представление позволяет моделировать различные трансформации гибкого объекта (например, сдвиги и повороты его частей). Сравнение двух гибких объектов производится путём сравнения их силуэтов, восстановленных по циркулярному графу.

Таким образом, описание формы через набор примитивов позволяет перейти от объектов фиксированной структуры к гибким объектам. Появляется возможность сравнения исходных и трансформированных объектов. Однако применение этих подходов к сравнению гибких объектов с окклюзиями за-

труднительно, поскольку они все подразумевают наличие сегментации формы на компоненты. А в случае окклюзий её получение не всегда возможно.

Идея циркулярного объекта в принципе способна привести к преодолению недостатков рассмотренных методов. Циркулярный граф является удобным инструментом для анализа формы [10, 35, 36], позволяя моделировать её различные трансформации и анализировать силуэт объекта после их выполнения. Поэтому по исходному изображению объекта без окклюзий можно построить циркулярное представление и определить множество допустимых трансформаций. А при сравнении искать такую трансформацию, в результате которой форма гибкого объекта наилучшим образом совпадёт с тестовой. На основе полученной меры сходства далее можно определить, является ли тестовая форма силуэтом рассматриваемого объекта. Данный подход может работать с окклюзиями, поскольку все трансформации выполняются над эталонным объектом (без окклюзий), а тестовая форма остаётся неподвижной.

1.3. Задача распознавания формы ладони

Системы распознавания личности, основанные на анализе ладони человека, относятся к числу наиболее старых систем автоматического распознавания личности. Устройства контроля доступа изготавливаются и внедряются уже с конца 1970-х годов. Они используются на ядерных заводах, в аэропортах и отелях. По сравнению с системами, использующими другие биометрические параметры человека, системы распознавания по форме ладони обладают рядом преимуществ. Они дружелюбны для пользователей (человеку привычно предъявлять свою ладонь, в отличие от радужной оболочки глаза или отпечатков пальцев); гигиеничны (в случае, когда не требуется позиционирование ладони на специальной платформе); менее зависимы от условий

окружающей среды и состояния человека, чем, например, системы распознавания по лицу (которые чувствительны к выражению лица и освещённости помещения); не требуют специальных сканеров (высокого разрешения, специальной конструкции и т.д.), для получения изображения ладони вполне подходят широко используемые недорогие web-камеры. У взрослого человека форма ладони меняется незначительно, за исключением случаев травм. К недостаткам можно отнести чувствительность таких систем к распознаванию в ситуациях, когда на ладони присутствуют украшения: кольца, большие браслеты; кроме того, такие системы чувствительны также к возможным изменениям формы ладони при перемене температуры или при физической нагрузке. Рука человека значительно меняется в период роста у детей.

Как отмечается в литературе, геометрия руки не обладает высокой отличительной способностью и, потому, может быть использована для проведения надёжного распознавания лишь в небольшой группе людей (несколько сот человек). Системы распознавания личности по форме ладони могут использоваться в больницах, университетах, в системах электронной торговли или при проведении банковских операций путем их внедрения в традиционные системы наряду с паролями и/или PIN-кодами.

Задача распознавания формы ладони обычно ставится следующим образом. Имеется набор эталонных изображений ладоней некоторой группы людей (пользователей). Для каждого человека может присутствовать одно или несколько изображений его ладони. Ладонь представлена на изображении либо связной областью чёрных пикселей на белом фоне, либо своей границей. Понятно, что эти представления взаимозаменяемы. Также имеется изображение ладони неизвестного человека. Необходимо определить, принадлежит ли эта ладонь одному из пользователей (и кому именно) или нет.

Для оценки качества распознавания того или иного метода обычно рас-

смаатриваются два показателя — это FAR (False Acceptance Rate) и FRR (False Rejection Rate). Первый показатель (коэффициент ложного пропуска) характеризует вероятность того, что система примет «чужого» человека за «своего». На практике измеряется (в процентах) как отношение числа «чужих», принятых за «своих», к общему числу людей в выборке. Второй показатель (коэффициент ложного отказа доступа) — вероятность отказа доступа человеку, имеющего допуск. Измеряется также в процентах, как отношение числа «своих», принятых за «чужих», к общему числу людей в выборке. Система тем лучше, чем меньше значение FRR при одинаковых значениях FAR .

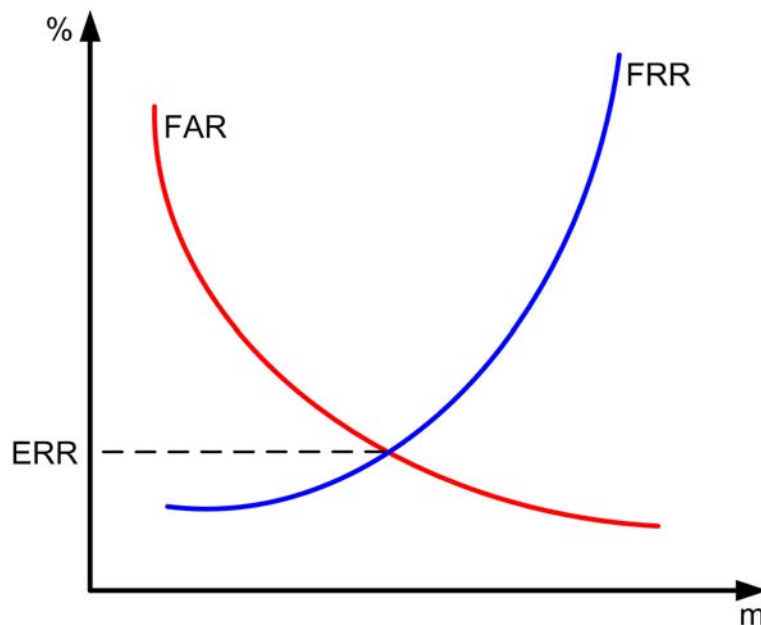


Рис. 1.8. Графики FRR и FAR , а также вычисление ошибки ERR , m — некоторый параметр системы.

Обычно системы распознавания зависят от некоторых параметров, варьируя которые можно управлять качеством распознавания. В этом случае строят графики FRR и FAR , где по оси абсцисс отложены возможные значения одного из параметров системы, а по оси ординат — значения FRR и FAR , получаемые при таком выборе параметра (рисунок 1.8). Например,

если распознавание проводится на основе применения порогового правила к получаемой оценке сходства объектов, то величина порога будет выступать в качестве такого параметра. Иногда используется сравнительная характеристика EER (Equal Error Rate), определяющая точку в которой графики FRR и FAR пересекаются.

Задачи распознавания личности обычно разделяют на две группы — задачи верификации и задачи идентификации. Под *верификацией* личности (сравнение «один к одному») подразумевается проверка, является ли человек действительно тем, кем себя представляет. Эта проверка проводится на основе сопоставления представленных человеком данных с эталонными данными представляемого человека. В системах верификации личности этот процесс обычно состоит из двух шагов. Сначала человек указывает, кем он хочет представиться системе. Это может быть выполнено различными средствами, например, путём ввода имени, пароля или использования смарт-карты. Далее система проводит сравнение один к одному и определяет подлинность сделанного человеком заявления. При *идентификации* (сравнении «один ко многим») производится сравнение предъявляемых данных с данными всех пользователей в базе. В результате чего система выдает решение о том, является ли данный человек «своим» и, в случае положительного ответа, кем именно. В отличие от случая верификации, человек не сообщает системе о том, кем он представляется. С этой точки зрения, задача идентификации является более сложной.

1.4. Методы распознавания формы ладони

Существующие подходы к распознаванию человека по форме ладони можно разделить на две группы:

- подходы, основанные на использовании признаков геометрии ладони;
- подходы, основанные на сравнении контуров ладоней.

На практике эти подходы используются как независимо, так и в комбинации друг с другом [37, 38].

Общая идея сравнения ладоней для методов из первой группы заключается в следующем. Сначала для ладони вычисляется ряд геометрических признаков и далее для распознавания применяется какой-либо из классификаторов. К числу признаков геометрии ладони относятся: длина пальцев, ширина пальцев на различных расстояниях от его кончика, кривизна пальцев, ширина ладони, периметр ладони и т.д. Рассматриваемый набор признаков и правила их вычисления варьируются между работами. На рисунке 1.9 представлен пример двух различных способов вычисления ширины пальцев. Так, на картинке слева ширина пальцев измеряется в достаточно большом количестве точек — 14-18 точек на каждый палец; тогда как на картинке справа таких точек всего три.

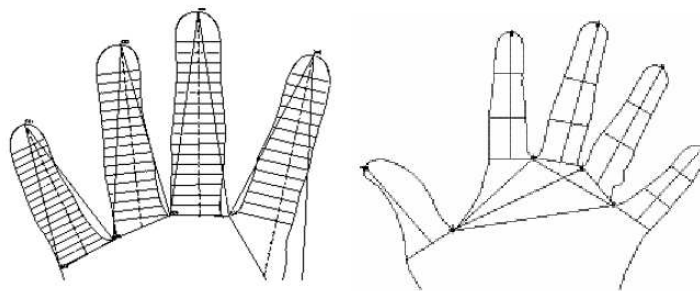


Рис. 1.9. Вычисление признаков геометрии ладони.

Методы из второй группы основаны либо на описании ладони набором признаков, получающихся в результате применения различных разложений к контуру ладони (Фурье или вейвлет-разложение, метод главных компонент и т.д.), либо на подгонке силуэтов ладоней и/или отдельных пальцев

с последующим вычислением меры их различия.

1.4.1. Распознавание по признакам геометрии ладони

В литературе представлено большое количество работ, посвященных распознаванию человека по признакам геометрии ладони. В работе [39] авторы формируют три вектора признаков отдельно для трех средних пальцев ладони. Вектор определяется 40 компонентами, каждый из которых характеризует ширину пальца на определенном расстоянии от кончика этого пальца. Обучение проводится методом опорных векторов. В процессе верификации принятие решения осуществляется на основе порогового правила. Экспериментальные данные содержат 200 изображений ладоней 20 человек (10 изображений для каждого человека). Возраст участников эксперимента составляет от 21 до 33 лет. Качество верификации оценивается по двум показателям — FRR и FAR . В результате серии экспериментов авторами получено значение $ERR = 3.4\%$.

Похожие признаки рассматриваются в работе [40] — длины пяти пальцев ладони, ширина среднего и безымянного пальцев, ширина ладони. Указывается устойчивость выбранных признаков для ладони одного человека, а также соответствие вычисленных значений признаков реальным измерениям. Количество правильно распознанных персон составило 95%.

В работе [41] распознавание проводится по средней ширине и длине четырех пальцев ладони. Большой палец исключен из рассмотрения, т.к. вычисляемые по нему характеристики оказываются неустойчивыми. На основе вычисляемых признаков формируется биокод человека. Биокод представляет собой совокупность двух показателей, первый из которых есть взвешенная линейная комбинация средней ширины пальцев, а второй — длины пальцев. В процессе верификации человек определяется как «свой», если расстояние

по обоим показателям не превышает стандартного отклонения этих показателей. Эксперименты проводились на базе из 80 человек в возрасте от 17 до 38 лет, содержащей не менее четырех ладоней для каждого участника. Полученное значение ERR составляет 4.5%.

Авторы [42] предлагают двухступенчатую модель распознавания. На первом шаге проводится сравнение признаков геометрии ладони и, в случае сходства ладоней по этим признакам, осуществляется дополнительное сопоставление областей кончиков пальцев. В работе анализируется 13 признаков геометрии ладони — длины всех пальцев, а также ширина для каждого пальца (за исключением большого), измеренная в двух точках. Для обучения используется модель Гауссовых смесей, решающее правило является пороговым. Эксперименты проводились на данных 20 человек, всего имелось 288 изображений (от 12 до 15 на человека). Наилучшие полученные показатели качества распознавания составили $FRR = 0\%$ и $FAR = 0.1222\%$ при использовании только признаков геометрии ладони и $FRR = 0.1111\%$, $FAR = 0.022\%$ в случае двухступенчатой модели.

Аналогичные признаки геометрии ладони рассматриваются в работе [43]. К ним относятся длина каждого из пальцев, ширина каждого из пальцев, измеренная в трех точках, и ширина ладони. Таким образом, всего вычисляется 21 признак. Авторами рассматривается как задача идентификации, так и задача верификации личности. Для обеих задач решающее правило является пороговым. В работе представлен сравнительный анализ качества распознавания в зависимости от используемой функции расстояния между векторами признаков. Наилучшее полученное значение ERR в случае верификации составило 2.999%, в случае идентификации — 5.83%. Оба значения были получены на одной и той же функции расстояния. Для проведения экспериментов использовались изображения ладоней 96 человек, по 5 изображений на каж-

дого.

Ширина пальцев в трех точках (в двух для большого пальца), длина каждого пальца и два измерения ширины ладони рассматриваются в качестве признаков в работе [44]. Сравнительный анализ качества верификации проводится для трех случаев: использование Евклидова расстояния, расстояния Хэмминга и модели Гауссовых смесей. Эксперименты показали явное преимущество модели Гауссовых смесей, лучший результат которой составил $ERR = 4.62\%$. В эксперименте участвовало 408 изображений ладоней.

В работе [45] ладонь описывается как вещественными (длина пальцев), так и функциональными признаками (функция ширины пальца в зависимости от параметра длины оси пальца). При распознавании используется модифицированный метод k ближайших соседей, позволяющий либо отнести ладонь к одному из существующих классов, либо определить, что она не принадлежит ни одному классу. Сходные признаки ладони используются в [11].

Эксперименты по идентификации и верификации личности на основе 25 признаков геометрии ладони были также проделаны в [46]. В работе используется специальная установка, представляющая собой горизонтальную платформу со штырьками-разделителями для позиционирования пальцев; с боковым зеркалом, позволяющим получать боковое изображение ладони; и ПЗС-камерой. Таким образом, наряду с характеристиками ладони, наблюдаемыми в горизонтальной плоскости, в вектор признаков включаются толщина ладони, среднего пальца и мизинца. Обучение проводилось на основе модели Гауссовых смесей, при распознавании использовалось пороговое правило. Эксперименты подтвердили преимущество использования выбранной модели перед простым вычислением расстояния Хэмминга с последующим использованием порогового правила. Результаты проведенных экспериментов показали качество верификации $ERR = 4.9\%$ для модели Гауссовых смесей против

$ERR = 8.3\%$ при использовании расстояния Хэмминга. Экспериментальные данные содержали изображения ладоней 20 человек, по меньшей мере 10 изображений на каждого участника.

Толщина пальцев ладони включается в вектор признаков также в работе [47]. Авторы приводят наглядную иллюстрацию (рисунок 1.10) вычисляемых 15 признаков геометрии ладони с помощью лиц Чернова [48]. Рассматривается пример, показывающий визуальную отличимость ладоней 20 человек по признакам геометрии.

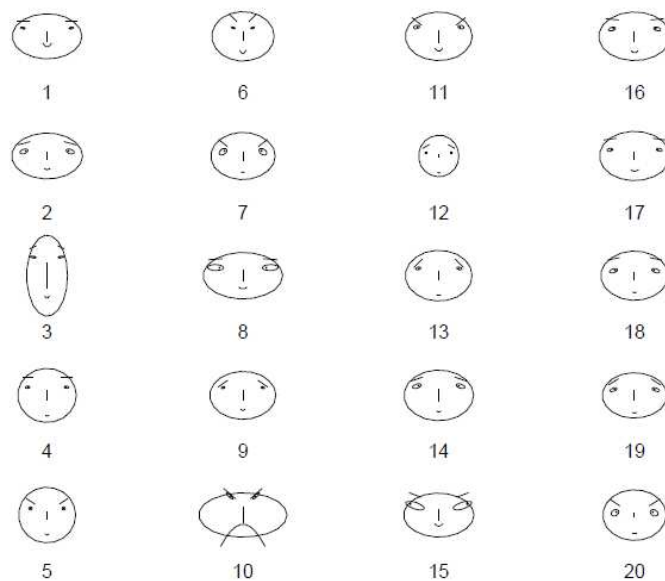


Рис. 1.10. Визуализация признаков геометрии ладони 20 человек с помощью лиц Чернова.

В [49] представлена зависимость качества распознавания от размера базы пользователей. Рассматриваются варианты с 10, 20, 30 и 40 персонами. При этом число правильно распознанных персон меняется с 92.5% до 91.67%.

Интересный подход представлен в работе [50]. Для исходного изображения ладони строится ее граничное описание, выделяются пальцы, строятся их оси. Каждый сегмент контура пальца проецируется на его ось, и вычисляются значения проекций. Далее строится гистограмма, описывающая плотность распределения полученных проекций в 100 точках. Для сглаживания плотно-

сти распределения используется оператор Гаусса. Таким образом, ладонь описывается 5×100 признаками. При сравнении ладоней проводится вычисление симметричного варианта расстояния Кульбака-Лейблера² между парами соответствующих пальцев эталонной и тестовой ладоней. В качестве меры сходства ладоней рассматривается среднее трех наименьших расстояний среди пяти полученных. Тестирование проводилось на базе ладоней Biosecure [52], содержащей 4700 изображений ладоней 750 человек, из них для 642 имелись изображения как для правой, так и для левой ладоней. Для 114 человек имелись снимки, выполненные с разрешением 150, 400 и 600 dpi. На разрешении 150 dpi были получены следующие результаты для 750 человек: при верификации доля правильно распознанных персон составила в среднем 95.6%, при идентификации — 87%.

Все методы, основанные на вычислении признаков геометрии ладони хорошо работают в отсутствии окклюзий. Однако, если окклюзии есть, не всегда возможно сегментировать ладонь (выделить пальцы и пясть), поэтому не хватает данных для построения признакового описания, либо получаемые значения признаков искажены.

1.4.2. Распознавание на основе контурной информации

При использовании признаков контура ладони удается получать результаты с более высоким качеством распознавания. Такие подходы представлены в работах [35, 37, 53–56].

В работе [54] авторами рассматриваются признаки разложения нормализованного представления ладони. Нормализация включает в себя ориентацию

² Если есть две дискретные случайные величины X, Y , принимающие значения в одном множестве $\chi \subset \mathbb{R}$, и их распределения задаются функциями вероятности p и q соответственно, то расстояние Кульбака-Лейблера [51] $D_{KL}(p, q)$ задаётся формулой: $D_{KL}(p, q) = \sum_{x \in \chi} p(x) \ln \frac{p(x)}{q(x)}$. Симметричный вариант получается при рассмотрении в качестве расстояния величины $D(p, q) = D_{KL}(p, q) + D_{KL}(q, p)$.

изображения ладони и каждого пальца по некоторому заданному направлению после регистрации. При сравнении ладоней производится совмещение их центров и основных осей (некоторых прямых, строящихся по изображению ладони). Авторами рассматривается случай присутствия колец-украшений, и предлагается метод их сегментации и удаления с получаемого изображения ладони. Рассматривается специальная процедура обработки запястья ладони, т.к. эта область меняет свои очертания в зависимости от освещения, присутствия рукавов одежды, часов и т.д. Отметим, что ранее [57] авторами исследовались четыре подхода — метод главных компонент, метод независимых компонент (метод Карунена-Лоева), метод осево-радиальных преобразований и распознавание по признакам, получаемым с помощью дискретного преобразования Фурье. Однако метод независимых компонент показал явное преимущество. В работе [54] исследуются расстояние Хаусдорфа и предлагаются две модификации метода независимых компонент. Наилучший результат был показан для одной модификации метода независимых компонент, составив 98 — 99% правильно распознанных персон в зависимости от размера вектора признаков. Рассматривались эксперименты как по верификации, так и по идентификации личности. Экспериментальные данные содержали 354 изображения ладоней 118 человек в возрасте от 20 до 35 лет.

Идея разложения контурного представления ладони представлена в [37]. Граница ладони рассматривается как сигнал, который раскладывается на составные части. В работе проведено сравнение трех способов генерации признакового описания — с помощью метода независимых компонент, вейвлет- и косинусного разложения. Задача классификации решалась двумя способами — нейронной сетью и методом ближайших соседей. Наилучшие оценки качества распознавания были получены при использовании метода ближайших соседей. Результаты идентификации в зависимости от выбранных признаков

разложения оказались сравнимыми и составили порядка 97 – 98%.

Другой подход представлен в [55]. Авторами предлагается выделять пары пальцев на сравниваемых изображениях ладони и сопоставлять их независимо друг от друга. Выравнивание пальцев проводится путем нахождения попарного соответствия точек границы пальцев. Среднее расстояние между точками границы пальцев рассматривается в качестве меры сходства ладоней. Полученные оценки FRR и FAR для задачи верификации составили 3.5% и 2% соответственно. На рисунке 1.11 представлен пример выравнивания контуров ладоней: слева представлены исходные контура, а справа — их совмещение. Экспериментальные данные содержали 353 изображения 53 человек. Результаты верификации показали $FRR = 3.5\%$ при $FAR = 2\%$.

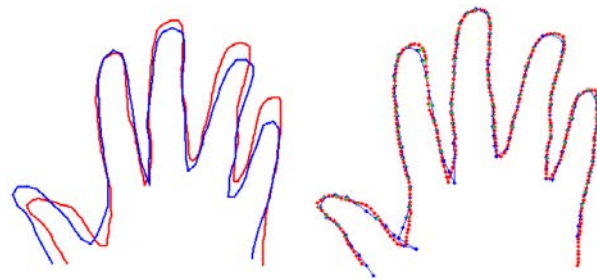


Рис. 1.11. Пример сопоставления контуров ладоней.

Идея попарного сравнения пальцев рассматривается также в [53]. В сравнении ладоней участвуют силуэты трех средних пальцев. Каждый из них приводится в некоторое стандартное положение, далее силуэты сравниваемых пальцев совмещаются, и их изображения вычитаются друг из друга. Полученная точность распознавания — 97%.

Предварительная нормализация ладоней применяется и в работе [35]. Для сравниваемых ладоней строятся циркулярные представления, позволяющее достаточно легко и быстро проводить различные трансформации объектов. При сравнении ладоней их циркулярные графы приводятся в положение

с фиксированными углами между пальцами и совмещаются. В качестве меры сходства ладоней используется нормированная симметрическая разность их силуэтов. Результаты проведенных экспериментов показали точность распознавания $ERR = 5\%$.

В работе [56] предлагается сегментировать пальцы на изображении ладони и описывать их жирными линиями. При сравнении ладоней проводится попарное сравнение их пальцев (как жирных линий), после чего вычисляется среднее значение близости по всем пальцам. Метрика, определяющая сходство жирных линий, есть отношение площади пересечения к площади объединения областей под графиками радиальных функций этих кривых. Распознавание проводилось на основе метода k ближайших соседей. Наилучший результат составил 96.92% правильно распознанных персон при $k = 5, 7$. Исходная база содержала 65 бинарных изображений ладоней 13 человек (5 изображений для каждого).

Исходя из анализа существующих методов распознавания формы ладони на основе контурной информации, можно сделать вывод, что использование подгонки ладоней даёт дополнительную возможность для сравнения формы гибких объектов. Однако, как и в случае признаков геометрии ладони, рассмотренные методы не работают при наличии окклюзий на изображении ладони.

1.5. Постановка задач исследования

1.5.1. Сравнение формы гибких объектов

Как видно из обзора литературы, существует достаточно большой спектр подходов и методов распознавания формы двумерных объектов. С точки зрения способа описания формы они значительно отличаются друг от друга.

Одни учитывают особенности границы объекта (например, её кривизну), другие — структуру формы (скелетный граф), статистические показатели (моменты) и т.д. Однако все эти методы объединяет следующее:

1. рассмотренные методы либо ориентированы на распознавание категорий объектов (когда важно выявить общие элементы у объектов одного класса, а не учитывать их индивидуальные особенности), либо конкретных объектов, форма которых меняется незначительно (негибкие объекты);
2. методы не предназначены для распознавания объектов с окклюзиями, когда некоторые части объекта перекрывают друг друга, но силуэт объекта известен полностью.

При этом, как отмечалось выше, необходимость распознавания объектов с окклюзиями на практике встречается довольно часто. Поэтому основная задача, рассматриваемая в рамках данной работы, — задача сравнения формы гибких объектов с окклюзиями.

Общая **постановка задачи** выглядит следующим образом. *Дано* множество бинарных изображений n гибких объектов: $I_1^r, I_2^r, \dots, I_N^r$. Для i -ого объекта имеется k_i его изображений без окклюзий: $\sum_{i=1}^n k_i = N$. Объект представлен чёрными пикселями на белом фоне, на каждом изображении присутствует только один объект, силуэты объекта на его изображениях различны. *Необходимо* разработать метод сравнения формы гибких объектов, допускающий присутствие окклюзий на тестовых изображениях.

Основная идея предлагаемого в работе *подхода* заключается в следующем:

1. Для каждого объекта строится подвижная модель. Она содержит информацию о структуре объекта, позволяет моделировать различные

движения объекта и анализировать его форму при их выполнении.

2. Для определения меры сходства эталонной модели и изображения объекта предлагается «шевелить» модель и искать такое её положение, при котором различие в силуэтах модели и тестового объекта является минимальным.

На основе полученной меры сходства в дальнейшем можно решать задачу классификации. Отметим, что аналогичным образом можно сравнивать модели друг с другом. Одна модель при этом считается эталонной (подвижной), а вторая — тестовой (неподвижной). Для тестовой модели строится силуэт, и далее он указанным выше способом сравнивается с эталонной моделью.

Задача сравнения формы гибких объектов в рамках предлагаемого подхода разбивается на **несколько подзадач**. Необходимо:

1. разработать модель гибкого объекта;
 - а. предложить метод построения силуэта модели с учётом допустимых деформаций;
 - б. разработать метод идентификации параметров (разметки) модели;
2. разработать метод сравнения модели и изображения объекта, основанный на подгонке их силуэтов.

Каждая из этих задач подробно рассматривается в главе 2.

1.5.2. Сравнение формы ладоней

Обзор существующих методов распознавания личности по форме и/или геометрии ладони показывает, что для функционирования этих методов необ-

ходимо, чтобы пальцы на исходном изображении ладони были хорошо отделены друг от друга, либо качество изображения допускало их корректную сегментацию. В противном случае признаки ладони и пальцев (например, длина и ширина пальцев, ширина ладони) вычисляются неверно и, соответственно, распознавание такой ладони получается некорректным. В некоторых системах проблема частично соприкасающихся пальцев решается путём использования специальных платформ со штырьками-ограничителями, маркеров или других средств фиксации положения ладони. Однако такие системы не очень удобны в использовании, особенно для людей старшего возраста. Предпочтительными и востребованными являются системы, допускающие максимально произвольное положение ладони на платформе [58].

В отсутствии средств фиксации положения ладони возникает проблема частично соприкасающихся пальцев. Она становится актуальной, когда имеется недостаточная освещённость или используются камеры с невысоким разрешением. В этой ситуации достоверно выделяется лишь силуэт ладони. Поэтому можно считать, что исходными данными являются бинарные изображения, на которых ладонь представлена чёрными пикселями на белом фоне. На рисунке 1.12 показаны примеры рассматриваемых бинарных изображений ладоней (на последних двух изображениях присутствуют частично соприкасающиеся пальцы).



Рис. 1.12. Примеры бинарных изображений ладоней.

Таким образом, рассматривается следующая **постановка задачи**. *Имеется* база эталонных изображений ладоней n человек: $I_1^r, I_2^r, \dots, I_N^r$. Для i -ого человека имеется k_i изображений его ладони: $\sum_{i=1}^n k_i = N$. На всех изображениях пальцы хорошо отделены друг от друга и могут быть легко сегментированы. Также имеется набор m тестовых изображений, возможно, с частично соприкасающимися пальцами: $I_1^t, I_2^t, \dots, I_m^t$. *Необходимо* разработать метод сравнения эталонных и тестовых изображений ладоней. *Решение* поставленной задачи основано на применение предлагаемого подхода. Для его реализации необходимо:

- описать модель ладони — задать группу допустимых деформаций и способ параметризации модели;
- ввести меру сходства силуэтов сравниваемых ладоней и способ её вычисления.

Каждый из этапов применительно к задаче сравнения формы ладоней подробно описывается ниже в главе 3. Преимущество подхода заключается в том, что:

- построение модели ладони производится в *offline* режиме и, соответственно, для получения более точной модели могут быть использованы более трудоёмкие процедуры;
- при сравнении происходит «шевеление» модели, а не тестовой ладони, тем самым снимается задача сегментации пальцев на тестовом изображении.

1.6. Выводы к первой главе

1. Задача сравнения формы гибких объектов возникает во многих прикладных задачах (биометрической идентификации, медицинской диагностике, биологических исследованиях и т.д.) и является актуальной на сегодняшний день.
2. Как показывает обзор литературы, известные методы не предназначены для распознавания формы гибких объектов, допускающих появление окклюзий при их трансформациях, хотя такие ситуации присутствуют на практике. Например, при биометрической верификации/идентификации личности по форме ладони, когда разрешается свободное позиционирование руки на платформе, возникают случаи частично соприкасающихся пальцев. А известные методы распознавания для них работают некорректно.
3. Существует метод описания формы в виде циркулярного графа, который потенциально имеет возможность решить задачу сравнения формы гибких объектов. Идея предлагаемого подхода заключается в построении модели гибкого объекта на основе циркулярного представления формы и её применении для моделирования допустимых трансформаций объекта и сравнения с другими формами. Для реализации предлагаемого подхода необходимо:
 - а. предложить и описать модель гибкого объекта;
 - б. предложить метод построения силуэта модели;
 - в. разработать метод идентификации параметров (разметки) модели;
 - г. разработать метод сравнения модели и изображения объекта, основанный на подгонке их силуэтов.

4. В качестве практического применения разработанного метода морфологического сравнения изображений гибких объектов рассматривается задача биометрической идентификации личности по форме ладони. Ладонь человека хорошо описывается предлагаемой моделью гибкого объекта. При этом структура и множество допустимых трансформаций достаточно просты.
5. Разработка нового подхода к сравнению гибких объектов применительно к задаче сравнения ладоней позволяет исследовать все элементы решения подобных задач, практически обосновать реализуемость и эффективность предлагаемых методов.

Глава 2

Циркулярная модель гибкого объекта и методы её идентификации и подгонки

Здесь и далее, говоря о «гибком объекте», мы будем понимать реальный объект, силуэт которого присутствует на изображении. А под словами «модель гибкого объекта» будем подразумевать предлагаемую в работе аппроксимацию реального объекта.

2.1. Модель гибкого объекта

В качестве модели гибкого объекта предлагается рассматривать его представление, которое вводится и подробно рассматривается в [1]. Приведём здесь основные определения, связанные с понятием гибкого объекта.

Пусть T — множество точек на евклидовой плоскости \mathbb{R}^2 , имеющее вид связного планарного графа. Граф имеет конечное множество вершин, а его ребра являются непрерывными линиями. При этом ребра не имеют точек пересечения, не совпадающих с их концами.

С каждой точкой $t \in T$ графа T связан некоторый круг c_t с центром в этой точке. Семейство кругов $C = \{c_t, t \in T\}$ называется *циркулярным графом* (или *циркуляром*). Граф T называется *осевым графом* (*скелетом*) циркулярного графа. Объединение $S = \bigcup_{t \in T} c_t$ всех кругов семейства C , как точечных множеств, называется *силуэтом* циркулярного графа. Силуэт циркулярного графа представляет собой замкнутое связное множество точек евклидовой плоскости $S \subset \mathbb{R}^2$, имеющее в качестве границы огибающую всего семейства кругов C .

Примеры циркулярных графов представлены на рисунке 2.1. Для удобства восприятия на нём присутствуют только круги циркулярного графа, определенные в вершинах осевого графа.

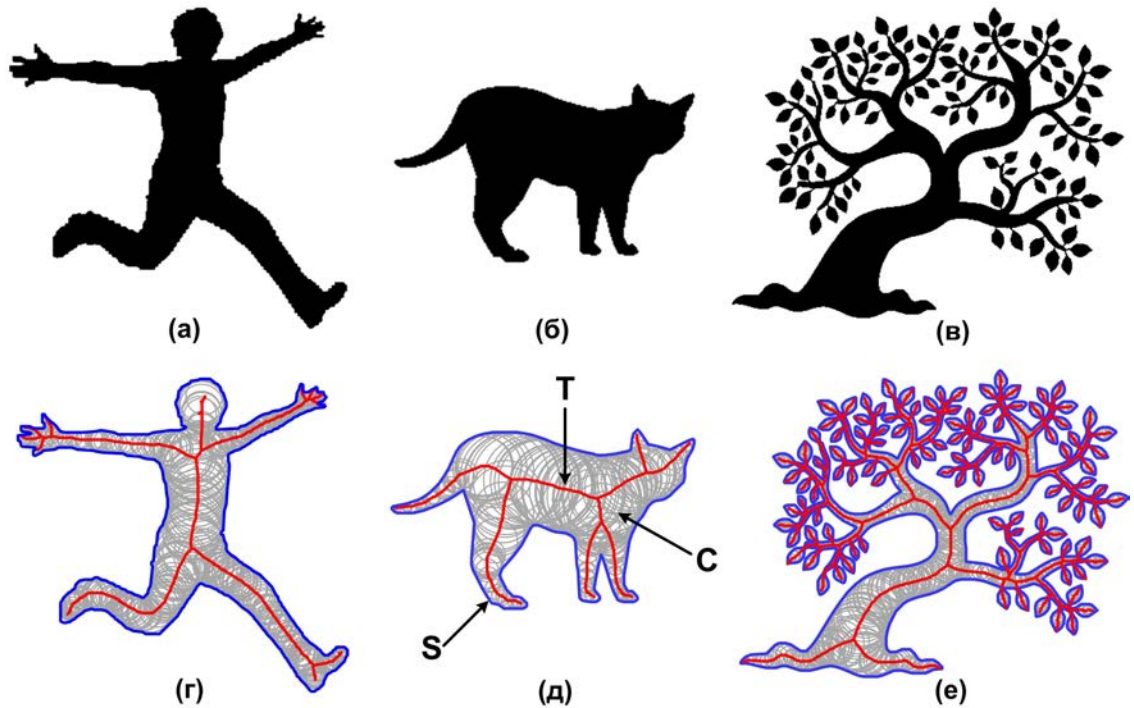


Рис. 2.1. Бинарные изображения объектов (а), (б), (в) и их циркулярные графы (г), (д) и (е) соответственно.

Множество допустимых преобразований циркулярного графа, обладающее следующими свойствами:

1. каждое преобразование сохраняет односвязность силуэта циркулярного графа;
2. в рамках множества преобразований для каждого преобразования существует обратное, т.е. множество преобразований образует группу;

называется *деформациями*³. Деформации могут включать в себя изменение

³ В отличие от работы [1], где класс деформаций существенно уже, поскольку там требуется также свойство сохранения планарности осевого графа

осевого графа с соответствующим перемещением семейства кругов (например, можно менять длины рёбер, поворачивать ветви осевого графа относительно его вершин), а также изменение размера кругов. В результате применения деформаций может меняться и силуэт циркулярного графа. Иногда вместо слова «деформация» мы будем использовать синонимы: «трансформация», «преобразование».

Два циркулярных графа называются *эквивалентными*, если для них существует деформация, в результате которой их силуэты совпадут при наложении. Обозначим через V группу деформаций циркулярного графа C .

*Циркулярной моделью гибкого объекта*⁴ называется циркулярный граф и его группа деформаций: $G = (C, V)$. Обозначим через $\mathcal{S}(\cdot)$ — оператор, ставящий в соответствие гибкому объекту G его силуэт, построенный как силуэт циркулярного графа этого объекта. Тогда $\mathcal{S}(G)$ — силуэт G . Циркулярный граф произвольного бинарного изображения может быть построен на основе непрерывного скелета [1]. В дальнейшем слово «циркулярный» мы будем опускать и говорить просто о модели гибкого объекта.

Циркулярный граф может быть построен для произвольного бинарного изображения, а вот множество допустимых деформаций определяется, исходя из конкретной решаемой задачи. Деформации могут применяться ко всему объекту сразу (например, сдвиг или поворот осевого графа) или к некоторой его части (например, поворот отдельных ветвей осевого графа). Во втором случае модель должна быть как-то размечена, т.е. указаны участки приложения деформаций.

Обозначим через G рассматриваемую модель объекта, а через \mathbf{p} — вектор её параметров. Также будем обозначать модель через $G(\mathbf{p})$, когда необходимо подчеркнуть выбранную параметризацию.

⁴ В отличие от работы [1], где этим определением вводится понятие гибкого объекта.

Таким образом, под моделью гибкого объекта будем подразумевать рассмотренное выше представление (циркулярный граф и множество допустимых деформаций), построенное для наблюдаемого силуэта. Для одного гибкого объекта может быть построено несколько моделей. Гибкий объект будем обозначать через \mathcal{F} .

2.2. Метод построения силуэта модели

По определению, силуэт гибкого объекта есть огибающая семейства кругов, его составляющих. Круг определён в каждой точке осевого графа. Следовательно, мы имеем бесконечное число кругов, для которых нужно построить общую огибающую.

Рассмотрим произвольное ребро O_1O_2 осевого графа (рисунок 2.2а). Согласно построению циркулярного графа [1] огибающей кругов, относящихся к этому ребру, будет замкнутая кривая $A_1A_2A_4A_3$, где A_1A_3 и A_2A_4 — отрезки внешних касательных к кругам с центрами в концевых вершинах ребра O_1O_2 , а A_1A_2 , A_3A_4 — дуги соответствующих окружностей. Фигуру, ограниченную кривой $A_1A_2A_4A_3$, будем называть *бициклом*. Считаем, что точки A_i в описании бицикла перечислены таким образом, что при движении от одной точки к другой бицикл остаётся слева.

Таким образом, для каждого ребра осевого графа может быть построен бицикл, и весь объект разбивается на конечное число пересекающихся бициклов (рисунок 2.2б). Тогда задача построения силуэта циркулярного графа превращается в задачу построения огибающей объединения конечного множества бициклов.

Формальная постановка задачи выглядит следующим образом. Дано множество бициклов $B = \{b_i\}_{i=1}^{n_b}$ на плоскости. Бицикл b_i образован двумя

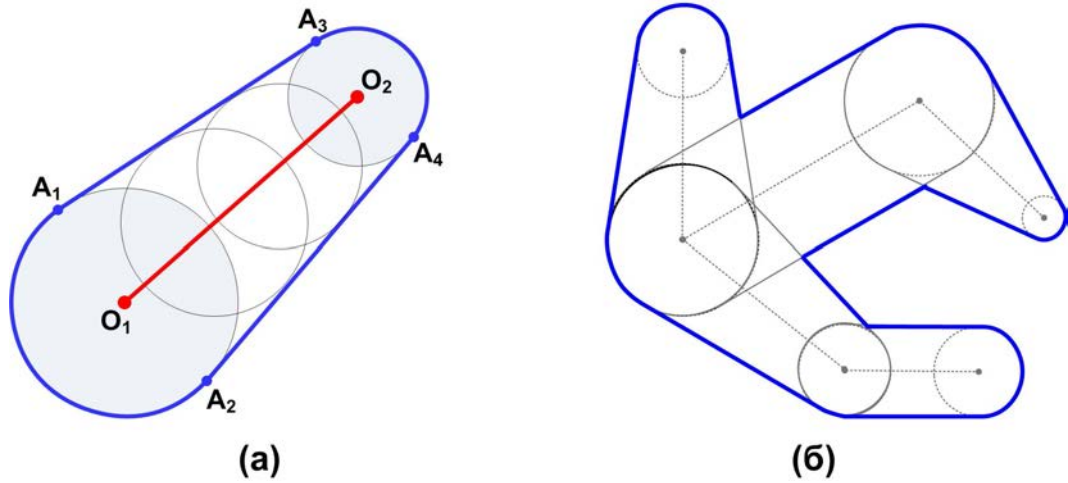


Рис. 2.2. (а) Бицикл $A_1A_3A_4A_2$ и (б) представление циркулярного графа в виде множества пересекающихся бициклов (жирной линией выделена их общая огибающая).

кругами с центрами в точках $O_{1i} = (x_{1i}, y_{1i})$ и $O_{2i} = (x_{2i}, y_{2i})$ и радиусами r_{1i} и r_{2i} соответственно. *Необходимо* построить общую огибающую заданного множества бициклов.

2.2.1. Алгоритм парных пересечений

Покажем, что данная задача может быть решена за время $O(n_b^2)$. **Наивный алгоритм** выглядит следующим образом:

1. перебираем все неупорядоченные пары бициклов (всего $n_b(n_b - 1)/2$ пар);
2. находим точки пересечения бициклов в паре (время $O(1)$ для пары);
3. прослеживаем внешнюю огибающую бициклов (время $O(n_b^2)$).

Рассмотрим шаги алгоритма на примере множества бициклов $B = \{b_i\}_{i=1}^5$, изображённых на рисунке 2.3. Сначала для каждой пары бициклов определяются их точки пересечения. Например, бициклы b_1 и b_2 пересекаются в двух точках, C_1 и C_{10} ; бициклы b_1 и b_5 — в точках C_5 и C_7 и т.д. Для каждой точки

C_i пересечения бициклов сохраняется информация о том, к каким бициклам она относится.

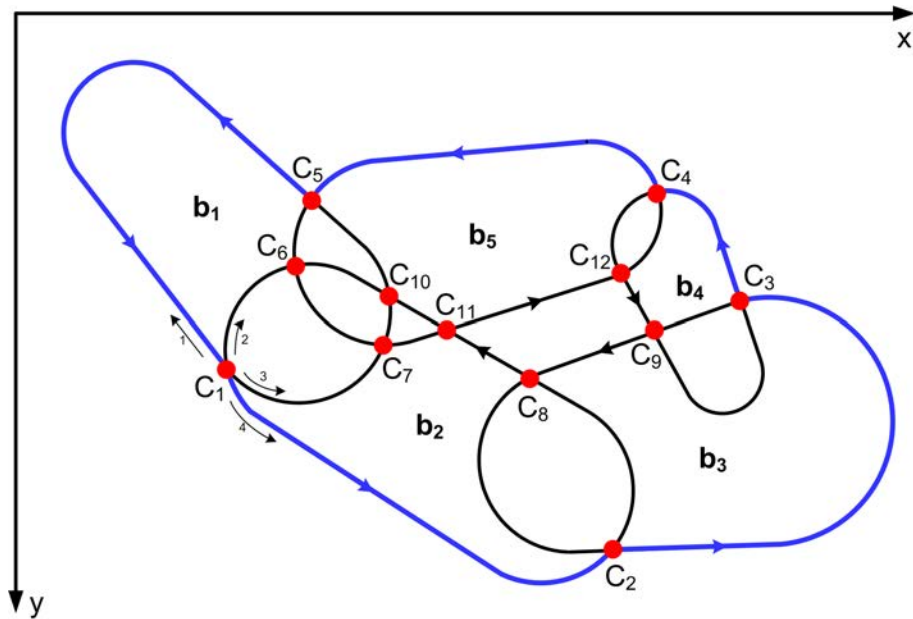


Рис. 2.3. Построение внешней огибающей объединения бициклов.

Построение внешней огибающей объединения бициклов начинается с точки пересечения C_i , которая заведомо в эту огибающую попадёт. Например, эта может быть самая левая, правая и т.д. из имеющихся точек. В данном примере начнём с самой левой точки (точки C_1) и будем строить огибающую против часовой стрелки. Из точки C_1 мы можем пойти по четырём направлениям, которые обозначены стрелками 1, 2, 3, 4 на рисунке 2.3. Поскольку огибающая строится против часовой стрелки, то выбираем направление, при котором угол между осью ординат и соответствующим направлением минимален. В данном случае это направление 4. Далее мы идём по выбранному направлению пока не дойдём до следующей точки пересечения — точки C_2 . Здесь уже достаточно «перескочить» с бицикла b_1 на бицикл b_2 и продолжить движение до следующей точки пересечения бициклов по направлению обхода бицикла b_2 . Процесс завершается тогда, когда мы вернёмся в исходную

точку C_1 . Поскольку в каждой точке пересечения мы просто перескакиваем с одного бицикла на другой, то прослеживание внешней огибающей линейно по числу точек пересечения. Два произвольных бицикла пересекаются не более, чем в четырёх точках. Всего имеется $n_b(n_b - 1)/2$ пар бициклов, поэтому сложность шага прослеживания внешней огибающей в худшем случае составляет $O(n_b^2)$. В среднем число точек пересечения есть $O(n_b)$. Однако общая сложность алгоритма остаётся $O(n_b^2)$.

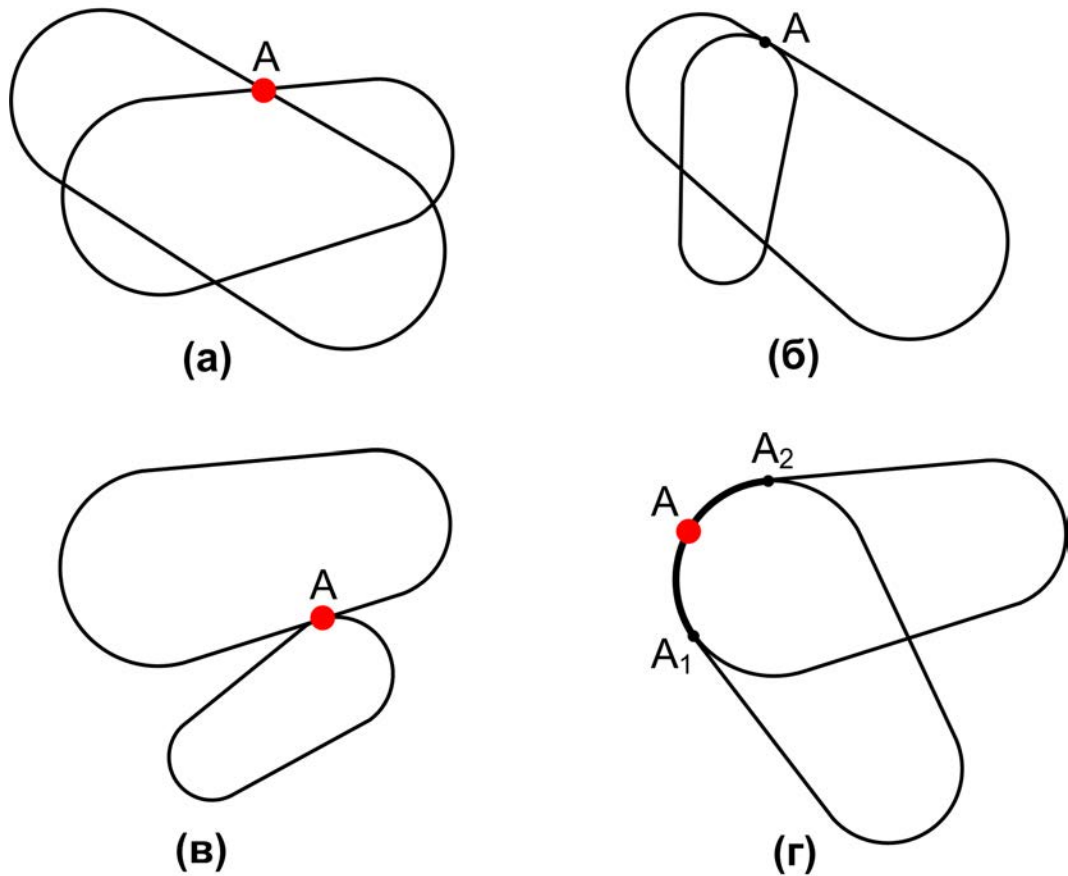


Рис. 2.4. Возможные варианты пересечения бициклов: (а) пересечение в точке A ; (б) внутренне касание в точке A ; (в) внешнее касание в точке A ; (г) два бицикла касаются целой дугой $\sim A_1A_2$, точкой касания считается точка A , совпадающая с серединой этой дуги.

Отметим, что касание бициклов является частным случаем их пересечения. На рисунке 2.4 представлены возможные варианты пересечения бициклов. В случае 2.4а мы имеем обычное пересечение в точке A . В случае 2.4б

имеется внутреннее касание, однако при прослеживании внешней огибающей в точке A не нужно переходить с одного бицикла на другой, поэтому эта точка исключается из рассмотрения. На рисунке 2.4в приведён пример внешнего касания в точке A . Более интересный случай представлен на рисунке 2.4г, когда два бицикла касаются целой дугой $\smile A_1A_2$. Такая ситуация возникает, если два бицикла имеют общий круг на концах. В этом случае в качестве точки касания бициклов предлагается рассматривать точку A , являющуюся серединой дуги $\smile A_1A_2$.

2.2.2. Восстановление силуэта модели по силуэтам фиксированных компонент

Для некоторого класса гибких объектов построение внешней огибающей бициклов может быть выполнено с меньшей вычислительной сложностью. Опишем этот случай. Будем рассматривать *гибкие объекты, которые могут быть представлены в виде совокупности небольшого числа компонент фиксированной структуры* (при выполнении трансформаций они не меняют свою форму). Эти компоненты образуются путём «разрезания» осевого графа в некоторых вершинах. При этом круг, определённый в вершине разреза, остаётся в обеих компонентах.

Примером такого гибкого объекта может выступать ладонь человека, фигура человека или животного. На рисунке 2.5 представлен гибкий объект, построенный для бинарного изображения игрушки «медвежонок», и его разложение на шесть компонент фиксированной структуры: четыре лапы, голова и туловище. В качестве допустимых трансформаций можно рассматривать повороты головы и лап, при которых не образуются «дырки» в силуэте. Так, силуэт медвежонок на рисунке 2.5а является односвязным. После трансформации, представленной на рисунке 1.6в свойство односвязности сохранилось,

а вот при трансформации 1.6г это свойство было потеряно (в силуэте появились две «дырки»).

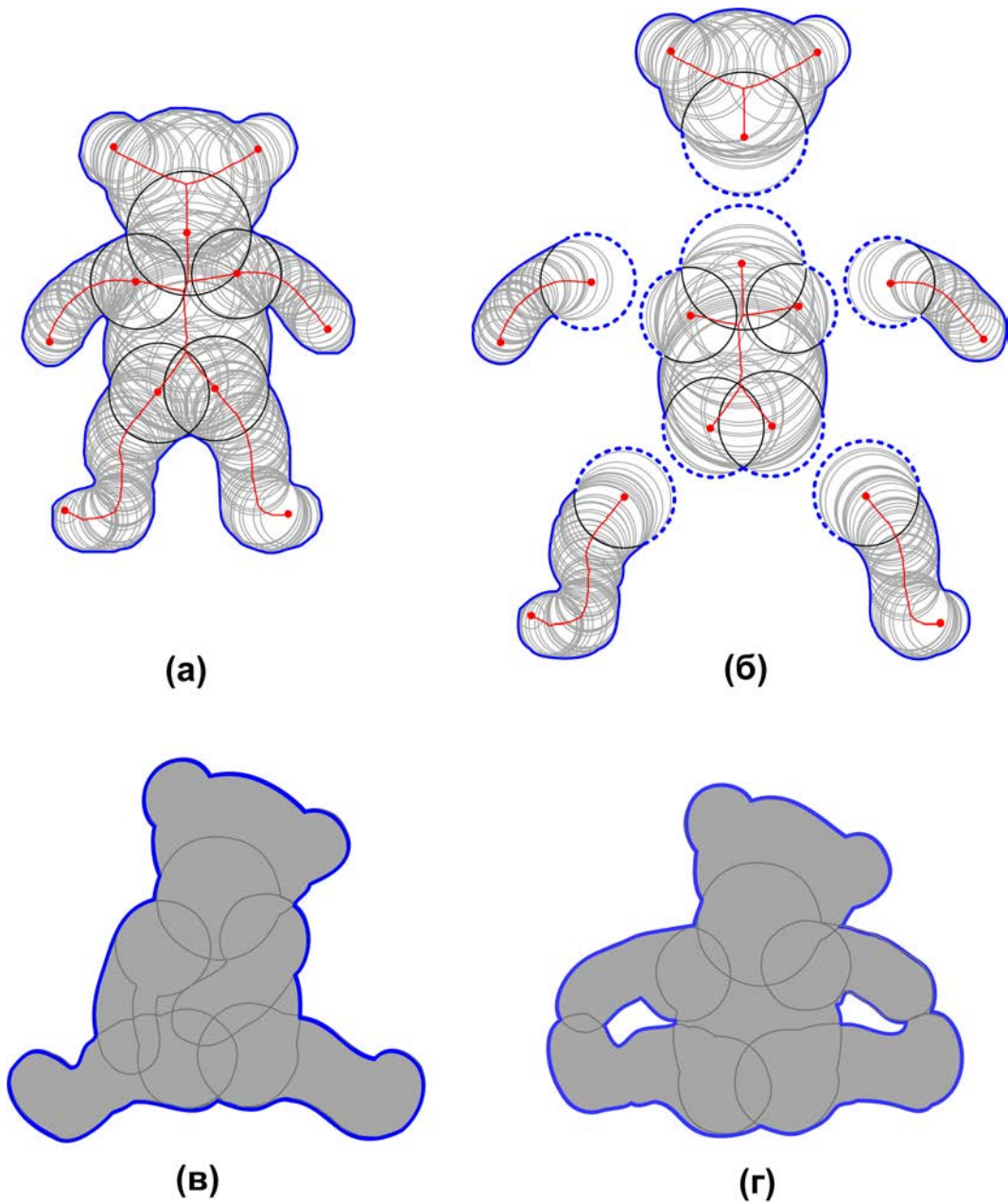


Рис. 2.5. Игрушка «медвежонок»: (а) циркулярное представление; (б) разложение на компоненты фиксированной структуры; (в),(г) силуэт объекта после выполнения трансформаций.

Если рассматривать процесс построения циркулярного графа по бинарному изображению объекта, то одним из шагов является построение много-

угольника минимального периметра. Фактически, этот многоугольник является силуэтом исходного (нетрансформированного) объекта, и нет необходимости восстанавливать его как огибающую множества бициклов.

Пусть гибкий объект разделяется на несколько компонент фиксированной структуры (рисунок 2.5б). Тогда граница такой компоненты будет состоять из частей исходной границы объекта (выделены сплошной линией) и дуг окружностей (выделены пунктирной линией). Аппроксимировав дуги ломаными линиями без самопересечений, мы получим границу компоненты в виде простого многоугольника. При этом исходный гибкий объект превращается в объединение конечного числа многоугольников. При выполнении трансформаций может меняться только взаимное расположение компонент, но форма каждой из них остаётся неизменной. Тогда задача построения силуэта трансформированного гибкого объекта сводится к задаче построения объединения конечного числа многоугольников (рисунок 2.5в), которая может быть решена известными методами [59].

Таким образом, *алгоритм построения силуэта гибкого объекта* выглядит следующим образом:

1. При построении модели гибкого объекта необходимо:
 - а. Разделить объект на компоненты фиксированной структуры (рисунок 2.6а); границей компоненты будет последовательность ломаных и дуг окружностей (рисунок 2.6б).
 - б. Аппроксимировать каждую дугу ломаной линией (рисунок 2.6в). В результате граница компоненты будет представлять собой простой многоугольник.
2. При трансформациях для построения силуэта гибкого объекта необходимо:

- а. Трансформировать гибкий объект — изменить взаимное расположение компонент (рисунок 2.6г).
- б. Построить огибающую компонент как огибающую объединения конечного числа простых многоугольников (рисунок 2.6д).

Аппроксимировать дуги окружностей ломаными можно разными способами. Например, можно строить вписанную ломаную, звенья которой имеют одинаковую длину, а расстояние от центра окружности до звеньев ломаной отличается от радиуса окружности не более, чем на ε . Рассмотрим пример окружности с центром в точке O и радиусом R (рисунок 2.6е). Пусть есть дуга $\smile AB$ этой окружности, которую нужно аппроксимировать ломаной линией. Пусть n_a — число её звеньев, а отрезок AE — одно из звеньев, $\angle AOB = \alpha$. Проведём прямую OC , проходящую через точку O и перпендикулярную AE : $C \in \smile AB, OC \perp AE$. Обозначим через D точку пересечения AE и OC : $D = AE \cap OC$. Поскольку длины всех звеньев ломаной равны, то $\angle AOE = \alpha/n_a$. Поскольку мы хотим построить такую аппроксимацию, что $OC - OD < \varepsilon$, то имеем следующее уравнение: $R - R \cos \frac{\alpha}{2n_a} < \varepsilon$. Откуда сразу же следует решение: $n_a > \frac{\alpha}{2 \arccos(1 - \frac{\varepsilon}{R})}$, при условии, что $\frac{\alpha}{2n_a} \in [0, \frac{\pi}{2}]$. Длина звена ломаной должна быть не более, чем $2\sqrt{\varepsilon(2R - \varepsilon)}$.

Замечание. В целом, аппроксимировать дуги ломаными линиями необязательно, поскольку алгоритм [60] поиска точек пересечения конечного числа отрезков может быть адаптирован для поиска точек пересечения конечного числа отрезков и дуг окружностей. Однако аппроксимация дуг ломаными даёт сведение задачи построения объединения компонент к задаче построения объединения конечного числа простых многоугольников, для которой существует известное решение.

Оценим *временную сложность* процесса построения силуэта трансфор-

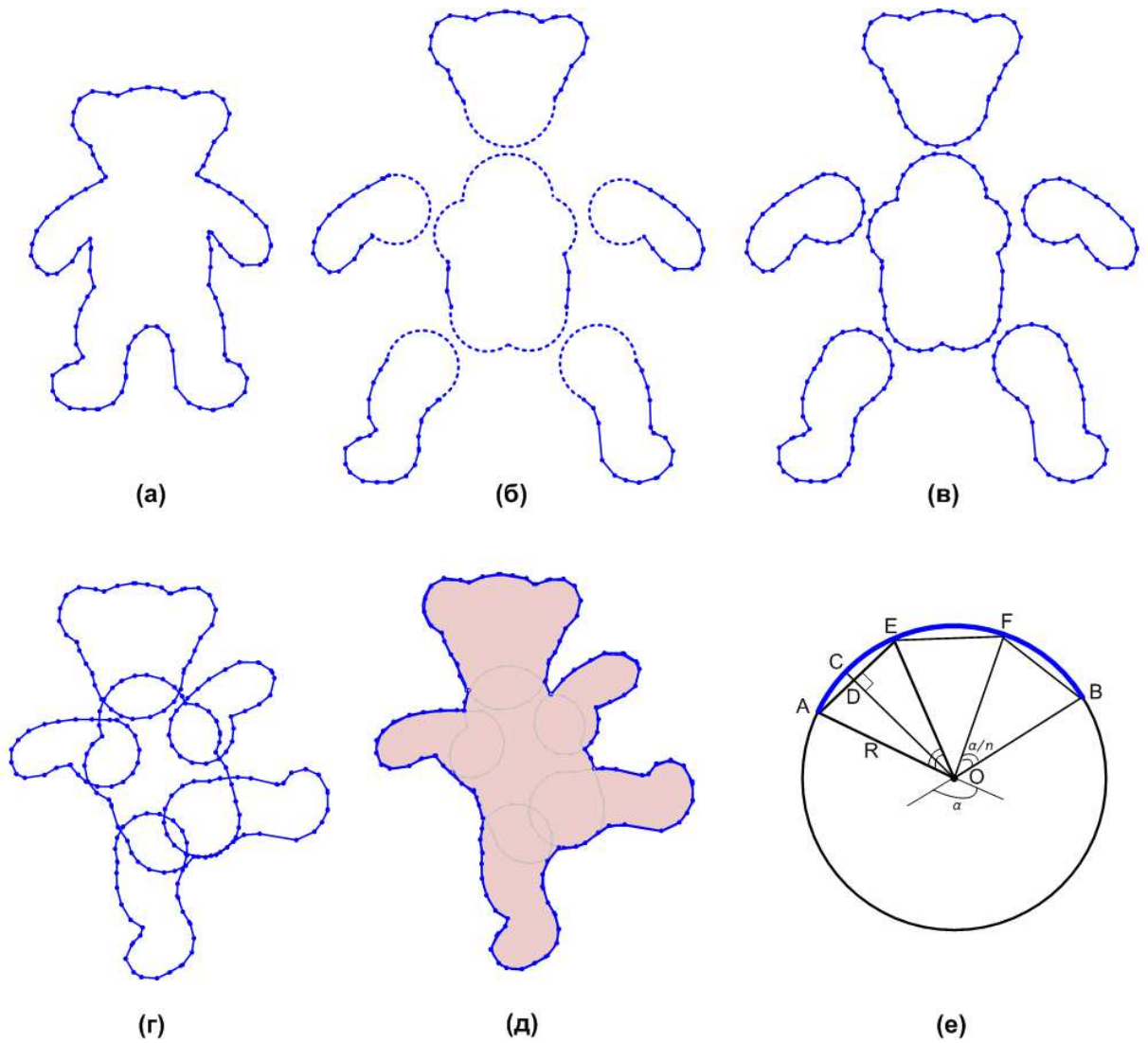


Рис. 2.6. (а) Силуэт исходного объекта, совпадающий с многоугольником минимального периметра (вершины многоугольника выделены жирными точками); (б) разложение гибкого объекта на компоненты фиксированной структуры: силуэт компоненты состоит из ломаных (выделены сплошной линией) и дуг окружностей (выделены пунктирной линией); (в) аппроксимация дуг ломаными, в результате чего границей каждой компоненты является простой многоугольник; (г) гибкий объект после трансформации; (д) построение силуэта гибкого объекта как объединения конечного числа многоугольников; (е) аппроксимации дуги AB ломаной линией $AEFB$.

мированного гибкого объекта по предложенному алгоритму, который фактически заключается в построении объединения конечного числа простых

многоугольников.

Пусть n_b — число бициклов, k — число компонент объекта, а n_0 — общее число точек пересечений границ компонент. Известно, что объединение i -ой пары многоугольников строится за время $O(n_p^i \log(n_1^i + n_2^i))$ в худшем случае и $O(n_p^i)$ в среднем, где n_1^i, n_2^i — число вершин в первом и втором многоугольнике соответственно, n_0^i — число их пересечений, а $n_p^i = n_0^i + n_1^i + n_2^i$ [59]. Так как число компонент объекта равно k , то необходимо рассмотреть $k - 1$ пар, поскольку сначала строим объединение первого и второго многоугольника, затем полученного объединения и третьего многоугольника и т.д. Так как $n_1^i + n_2^i \leq n$, $n_0^i \leq n_0$, где n — общее число сторон во всех многоугольниках, то грубая оценка вычислительной сложности будет в худшем случае $O((k - 1)(n + n_0) \log n)$ (или $O((k - 1)(n + n_0))$ в среднем). Так как k — константа, то оценка будет составлять $O((n + n_0) \log n)$ в худшем случае (или $O(n + n_0)$ в среднем).

Оценим n — общее число сторон в многоугольниках. Поскольку имеется n_b бициклов, то всего может быть $2n_b$ дуг окружностей. При разделении гибкого объекта на компоненты фиксированной структуры (рисунок 2.6б) новые дуги возникнуть не могут (в силу построения), поэтому общее число дуг, описывающих границу компонент, не превосходит $2n_b$. Каждая дуга аппроксимируется ломаной линией, пусть n_l — общее число звеньев этих ломаных. Далее, каждому ребру осевого графа соответствует один бицикл. С другой стороны, каждому ребру осевого графа соответствует не более двух сторон многоугольника минимального периметра (в силу построения), поэтому общее число прямолинейных отрезков, образующих границу компонент, не превосходит $2n_b$. Следовательно, после аппроксимации компонент мы будем иметь многоугольники, оценка общего числа сторон которых есть $n \leq 2n_b + n_l$.

Таким образом, силуэт гибкого объекта может быть построен за вре-

мя $O((n_b + n_l + n_0) \log(n_b + n_l))$ в худшем случае и $O(n_b + n_l)$ в среднем. Эксперименты показывают, что общее число n_l звеньев ломаных, аппроксимирующих дуги окружностей, на порядок меньше общего число прямых отрезков, формирующих границу компонент. То есть можно считать, что $n_l = O(n_b)$. Поэтому оценка сложности алгоритма будет составлять $O((n_b + n_0) \log n_b)$ в худшем случае и $O(n_b + n_0)$ в среднем (n_b — число бициклов, n_0 — общее число точек пересечений границ компонент).

2.3. Метод идентификации (разметки) модели

В случае, когда анализируемые объекты являются достаточно простыми или их небольшое количество, разметка их моделей может быть выполнена экспертом вручную. Однако в общем случае такой вариант может оказаться трудоёмким. Например, если рассматривать задачу распознавания личности по форме ладони и считать, что имеется хотя бы 100 пользователей, для каждого из которых есть по 10 изображений его ладони, то эксперту необходимо разметить 1000 моделей. Соответственно, при увеличении размера базы объём работы возрастает. Кроме того, не существует критерия точности разметки, выполняемой экспертом. Поэтому целесообразно иметь механизм, позволяющий производить такую разметку автоматически, т.е. идентифицировать параметры модели программно на основе имеющихся изображений гибкого объекта, а также дающий оценку качества получаемой разметки и способ её оптимизации.

Формальная постановка задачи выглядит следующим образом.

Дано n бинарных изображений I_1, \dots, I_n одного объекта в разных положениях. Считается, что фигуры на изображениях I_1, \dots, I_n имеют одинаковый масштаб. Для каждого такого изображения может быть построено циркулярное

разложение способом, описанным в разделе 2.1, и, соответственно, n незамеченных моделей G_1, \dots, G_n . *Задача* состоит в том, чтобы вычислить вектор параметров \mathbf{p} модели. Этот вектор считается одинаковым для всех моделей, т.к. они являются моделями одного и того же объекта.

Предлагаемое *решение* заключается в следующем. Пусть задана мера $\mu^* = \mu^*(G, S)$, которая позволяет определить сходство модели G и силуэта S некоторого объекта. Данная мера вводится ниже в разделе 2.4. Фактически, под «мерой сходства» понимается мера различия между объектами, т.е. чем она меньше, тем более объекты похожи друг на друга, и наоборот. Отметим, что значение меры μ^* зависит от параметризации модели G , поскольку при её вычислении модель подвергается различным разрешённым для неё деформациям. Если считать, что для исходных моделей G_i и G_j вектор \mathbf{p} определён, то степень сходства G_j и G_i может быть вычислена по формуле: $\mu^*(G_i(\mathbf{p}), S(G_j(\mathbf{p})))$. В общем случае $\mu^*(G_i(\mathbf{p}), S(G_j(\mathbf{p}))) \neq \mu^*(G_j(\mathbf{p}), S(G_i(\mathbf{p})))$.

Тогда в качестве интегральной меры сходства моделей одного объекта предлагается рассматривать среднее значение мер сходства моделей этого объекта, вычисляемое по следующей формуле:

$$\mu_{av}(\mathbf{p}) = \frac{1}{n(n-1)} \sum_{\substack{i,j=1,\dots,n \\ i \neq j}} \mu^*(G_i(\mathbf{p}), S(G_j(\mathbf{p}))) \quad (2.1)$$

Будем исходить из естественного предположения, что модели одного объекта должны быть максимально похожи друг на друга. Считая, что параметризация моделей неизвестна, можно поставить следующую оптимизационную задачу — необходимо найти такой вектор параметров \mathbf{p}^* , при котором инте-

гральная мера сходства моделей минимальна:

$$\mu_{av} : \mathbb{R}^{n_p} \rightarrow \mathbb{R}_+ \quad (2.2)$$

$$\mathbf{p}^* : \mu_{av}(\mathbf{p}^*) = \min_{\mathbf{p} \in P} \mu_{av}(\mathbf{p}) \quad (2.3)$$

где n_p обозначает размер вектора параметров \mathbf{p} .

Вектор \mathbf{p}^* , полученный в результате решения задачи (2.2) определяет необходимую разметку ладони. Символом $P \subseteq \mathbb{R}^{n_p}$ в формуле обозначено некоторое допустимое множество значений вектора параметров. Например, если мы хотим определить точки поворотов пальцев ладони по серии её снимков, на которых углы между пальцами меняются от изображения к изображению, в качестве \mathbf{p} можно рассматривать вектор, составленный из расстояний от кончика каждого пальца до его точки поворота. При этом из общих соображений понятно, что это расстояние должно быть больше длины пальца, но и не превосходить, например, двух длин пальца. В этом случае допустимое множество P будет являться пространственным параллелепипедом в пространстве параметров \mathbb{R}^5 .

Проиллюстрируем это на примере игрушки «медвежонок» (рисунок 2.5а). Пусть имеется 4 изображения медвежонка в разных положениях, представленных на рисунке 2.7а, для каждого из которых построена неразмеченная модель $G_i, i = 1, 2, 3, 4$, S_i — силуэт этой модели. Модель состоит из шести компонент фиксированной структуры (четырёх лап, головы и туловища). В качестве допустимых трансформаций будем рассматривать повороты лап и головы, считая, что туловище не меняет своего положения. Вектор параметров \mathbf{p} , задающий разметку модели, будет состоять из пяти элементов: $\mathbf{p} = (R_1, R_2, R_3, R_4, R_5)$, где R_i определяет положение точки поворота для соответствующей компоненты гибкого объекта. В качестве меры сходства μ^* будем рассматривать минимальную симметрическую разность, которую уда-

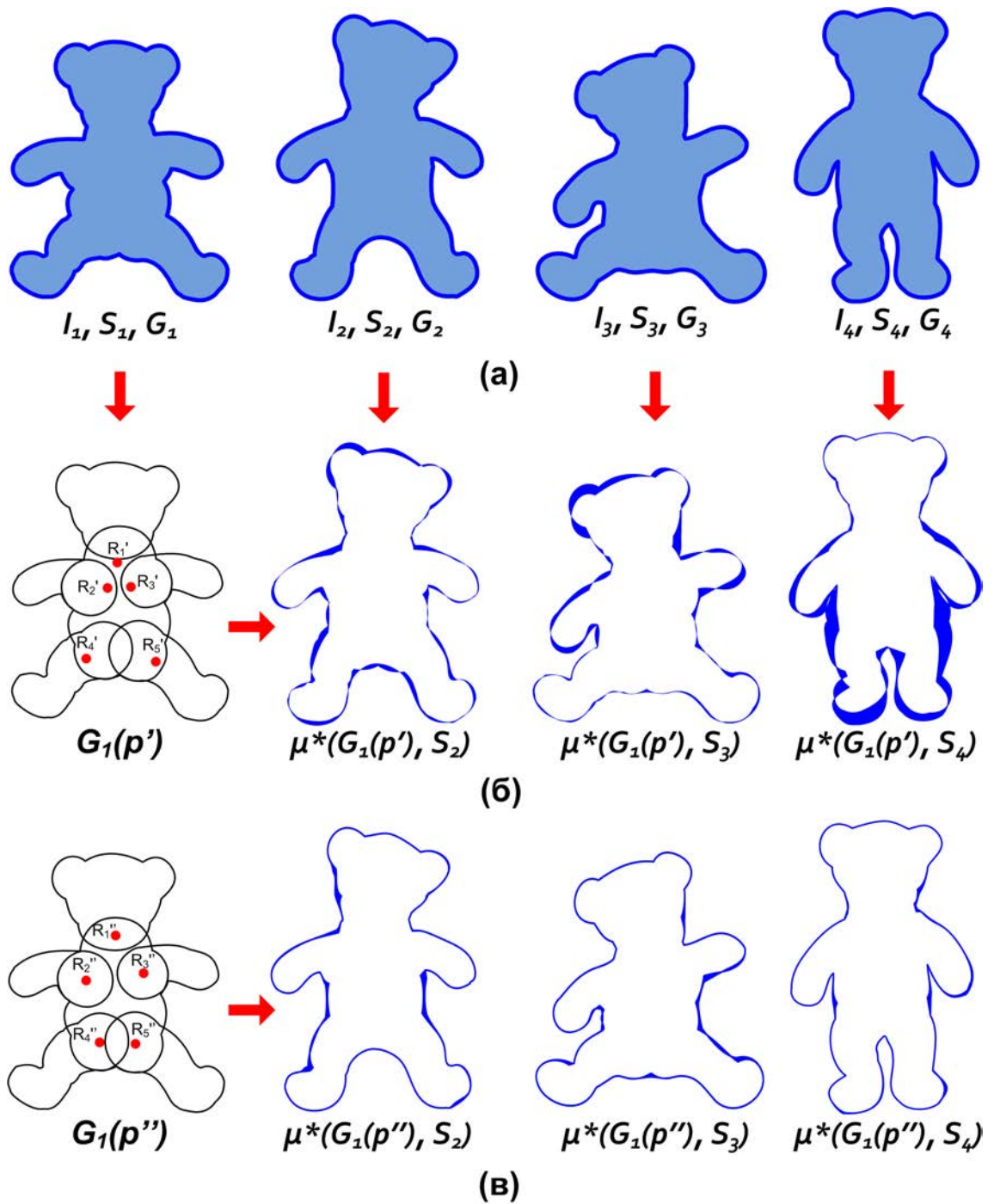


Рис. 2.7. (а) Исходные изображения гибкого объекта; (б),(в) совмещение силуэтов трансформированной модели гибкого объекта и исходных силуэтов при различном выборе точек R_i поворотов компонент (тёмным цветом выделена симметрическая разность силуэтов).

ётся получить при совмещении силуэтов объектов.

Рассмотрим модель G_1 гибкого объекта и некоторое произвольное положение точек поворотов $\mathbf{p}' = (R'_1, R'_2, R'_3, R'_4, R'_5)$, представленное на рисунке 2.7б. Последовательно сравним эту модель со всеми остальными моделями (G_2, G_3 и G_4), вычисляя меру их сходства $\mu^*(G_1(\mathbf{p}'), S(G_j(\mathbf{p}'))) = \mu^*(G_1(\mathbf{p}'), S_j), j = 2, 3, 4$. Аналогичное сравнение можно провести для всех остальных пар моделей и определить среднюю меру сходства $\mu_{av}(\mathbf{p}')$ моделей при заданном выборе точек поворотов по формуле (2.1). Уже на модели G_1 видно, что её силуэт после трансформации и силуэты остальных моделей отличаются друг от друга при совмещении, особенно это заметно при сравнении с моделью G_4 . Рассмотренное положение точек поворотов \mathbf{p}' вряд ли можно назвать оптимальным.

Далее рассматривается другое положение точек поворотов, например как на рисунке 2.7в, и снова вычисляется мера сходства $\mu_{av}(\mathbf{p}'')$, где $\mathbf{p}'' = (R''_1, R''_2, R''_3, R''_4, R''_5)$. Возможные положения P точек поворотов компонент ограничим областями кругов, определённых в вершинах разреза осевого графа, выполненного при построении компонент (рисунок 2.5а,б). Перебрав допустимые положения точек поворотов, мы можем найти наилучшее положение \mathbf{p}^* , для которого выполнено (2.2). В данном примере оно совпадает с \mathbf{p}'' .

Таким образом, разметка моделей гибких объектов сводится к решению оптимизационной задачи выбора значений параметров, при которых объекты одного класса являются максимально похожими друг на друга.

2.4. Сравнение модели и изображения объекта

Формальная постановка задачи. *Имеется* размеченная модель G (эталон) и бинарное изображение (тестового) объекта или, то же самое, его си-

луэт S . *Задача* состоит в том, чтобы оценить сходство G и S . На основе этой оценки в дальнейшем уже можно решать задачу классификации формы объектов.

Так как для тестового объекта нам известен лишь его силуэт, *предлагается* строить оценку сходства на основе сравнения силуэта модели G и силуэта S . Модель G является подвижной, т.е. нам известна группа её допустимых деформаций, мы можем применить их к модели и посмотреть её силуэт после их выполнения.

Пусть μ есть некоторая мера, позволяющая определить сходство силуэтов двух произвольных объектов. Например, в качестве меры μ может быть рассмотрена симметрическая разность силуэтов или Хаусдорфово расстояние между их границами. Мера μ подбирается под конкретную задачу, чтобы отразить существенные для неё различия в силуэтах объектов.

Тогда для силуэта эталонной модели G и тестового силуэта S можно вычислить меру сходства как $\mu(\mathcal{S}(G), S)$. Однако понятно, что простое совмещение силуэтов может не дать хорошей оценки сходства, если даже исходные объекты похожи. Например, на рисунке 2.8а представлена эталонная модель гибкого объекта G и тестовый силуэт S . Если в качестве μ рассмотреть симметрическую разность силуэтов, то указанные объекты будут сильно отличаться друг от друга (рисунок 2.8б). Это происходит потому, что на изображениях медвежонок принимает разные позы. При этом, если «пошевелить» модель медвежонок (применить деформации, связанные с движениями лап и головы), то можно получить силуэт, почти совпадающий с силуэтом S (рисунок 2.8в). Задача заключается в том, чтобы найти такую деформацию.

По определению гибкий объект есть пара $G = (C, V)$, где C — циркуляр, а V — множество допустимых деформаций. Будем считать, что семейство допустимых деформаций является параметрическим. Обозначим через \mathbf{a} набор

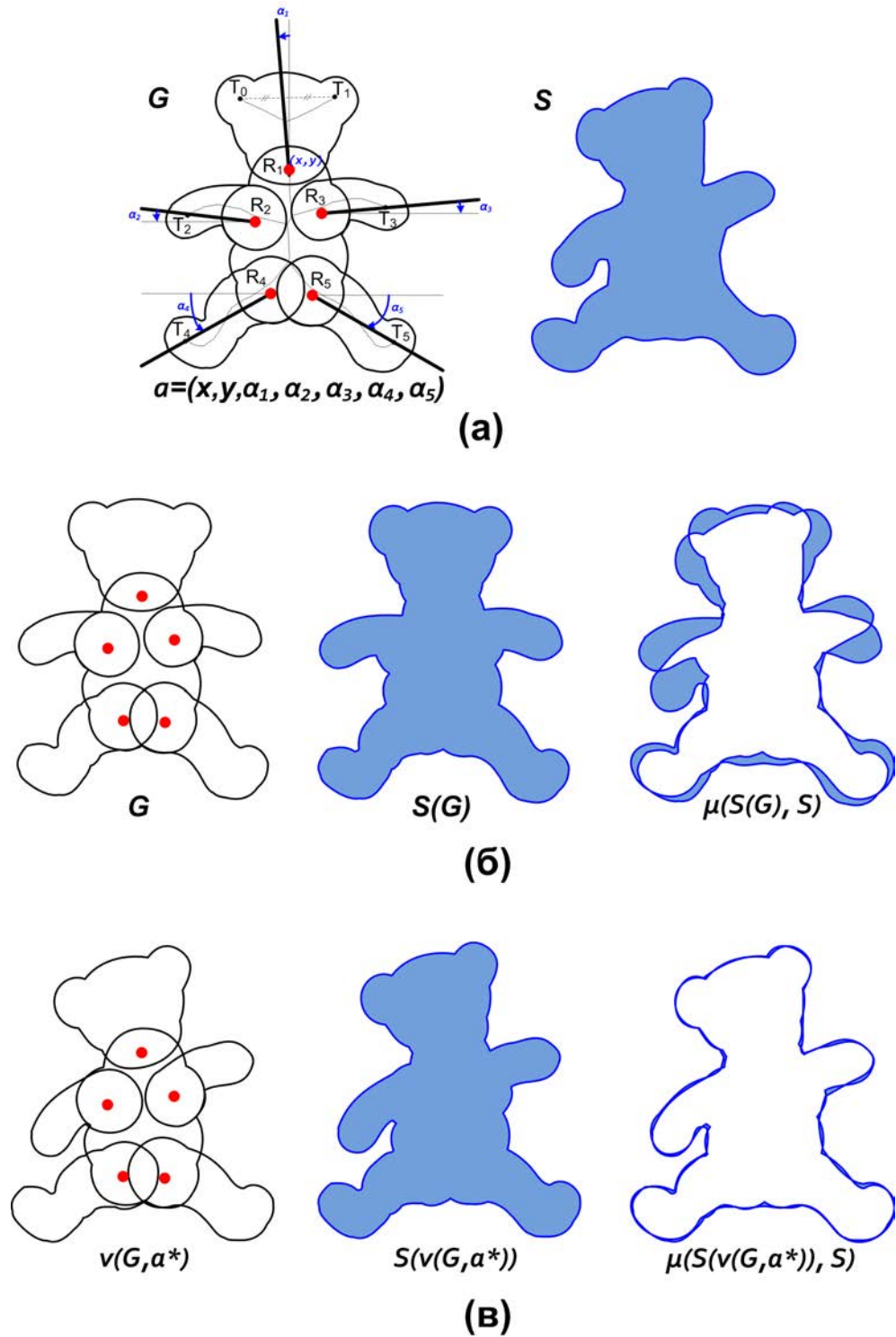


Рис. 2.8. (а) Эталонная модель гибкого объекта G и параметризация её трансформаций, тестовый силуэт S ; (б) сравнение силуэтов G и S ; (в) применение оптимальной трансформации $v(\mathbf{a}^*)$ к модели и результат совмещения силуэтов эталона и теста.

параметров деформаций, а через A — множество допустимых значений \mathbf{a} . Тогда $v(\mathbf{a}) \in V$ есть трансформация, однозначно определяемая вектором параметров \mathbf{a} . Далее, обозначим через $v(G, \mathbf{a})$ объект, получаемый в результате применения деформации $v(\mathbf{a})$ к гибкому объекту G , а через $\mathcal{S}(v(G, \mathbf{a}))$ — его силуэт.

Рассмотрим произвольную деформацию $v(\mathbf{a}) \in V$, $\mathbf{a} \in A$. Мера сходства силуэта деформированной модели G и силуэта S есть $\mu(\mathcal{S}(v(G, \mathbf{a})), S)$. Тогда поиск деформации, реализующей наилучшее совмещение силуэтов эталона G и теста S , можно сформулировать как следующую оптимизационную задачу:

$$\mathbf{a}^* : \mu(\mathcal{S}(v(G, \mathbf{a}^*)), S) = \min_{\mathbf{a} \in A} \mu(\mathcal{S}(v(G, \mathbf{a})), S) \quad (2.4)$$

Решение задачи (2.4) позволяет найти нужную деформацию и соответствующее ей значение меры сходства силуэтов.

Тогда предлагается определять меру сходства μ^* модели G и силуэта S как минимальное значение меры сходства силуэтов деформированной модели G и S :

$$\mu^*(G, S) = \min_{\mathbf{a} \in A} \mu(\mathcal{S}(v(G, \mathbf{a})), S) \quad (2.5)$$

Рассмотрим для модели гибкого объекта, изображённой на рисунке 2.8а, следующие допустимые трансформации: поворот лап и головы в соответствующих им точках R_i , а также параллельный перенос на некоторый вектор. Параметризуем эти трансформации. Пусть $T_i, i = 1, \dots, 5$ — терминальные вершины осевого графа. Проведём прямые $R_i T_i, i = 2, 3, 4, 5$ и обозначим через α_i угол между соответствующей прямой и горизонтальной осью. Также соединим отрезком точки T_0 и T_1 и проведём прямую, проходящую через середину этого отрезка и точку R_1 . Через α_1 обозначим угол между этой прямой и вертикальной осью. Пусть далее (x, y) — координаты точки R_1 поворота головы. Тогда в качестве вектора параметров рассмотрим следующий:

$\mathbf{a} = (x, y, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$. Ограничим допустимые трансформации следующим множеством $A = \{\mathbf{a} : (x, y) \in [-D; D]; \alpha_i \in [-\frac{\pi}{2}; \frac{\pi}{2}]\}$, где D — максимальное евклидово расстояние между точками силуэта гибкого объекта. Решение задачи (2.4) для рассматриваемого примера изображено на рисунке 2.8в.

Замечание. В предлагаемом подходе подразумевается, что гибкий объект \mathcal{F} представлен несколькими эталонными моделями G_1, \dots, G_n . Алгоритм, описанный выше, позволяет оценить сходство некоторого тестового силуэта S и каждой из моделей. Таким образом, для гибкого объекта может быть получено n значений меры сходства: $\mu^*(G_i, S)$, $i = 1, \dots, n$. При этом итоговая мера сходства $\mu_{\mathcal{F}}$ самого гибкого объекта \mathcal{F} и силуэта S может быть вычислена как некоторая функция от имеющихся n значений. Например, в качестве меры $\mu_{\mathcal{F}}(\mathcal{F}, S)$ может рассматриваться минимальное из значений $\mu^*(G_i, S)$, $i = 1, \dots, n$ или их среднее.

2.5. Алгоритм решения прикладных задач

Предложенный подход позволяет решать задачи, **общая постановка** которых может быть сведена к следующей.

Дано n бинарных изображений гибкого объекта \mathcal{F} в разных положениях: I_1, \dots, I_n , а также силуэт S . *Необходимо* определить меру сходства гибкого объекта \mathcal{F} и силуэта S .

Алгоритм решения для конкретной прикладной задачи представлен ниже. При этом обычным шрифтом выделены общие элементы подхода, а курсивом отмечены те, которые должны быть определены для конкретной решаемой задачи.

1. Построение моделей гибкого объекта по изображениям I_1, \dots, I_n . *Необходимо:*

- а. Построить циркулярный граф C_i по исходному изображению I_i . В результате будем иметь n неразмеченных моделей гибкого объекта $\mathcal{F}: G_1, \dots, G_n$.
- б. *Определить (вербально) допустимые преобразования модели. На основе этого определить способ параметризации модели, т.е. указать вектор параметров \mathbf{p} и множество его допустимых значений P .*
- в. Сформировать интегральную меру сходства $\mu_{av}(\mathbf{p})$ для моделей G_1, \dots, G_n по формуле (2.1) и поставить задачу минимизации (2.2) для поиска оптимального вектора параметров \mathbf{p}^* .
- г. *Решить оптимизационную задачу из предыдущего пункта, выбрав наиболее подходящий метод минимизации.* В результате будем иметь n размеченных моделей гибкого объекта.

2. Сравнение модели G_i и силуэта S . Необходимо:

- а. *Выбрать меру μ для определения сходства пары совмещённых силуэтов и метод её вычисления.*
- б. *Ввести способ параметризации деформаций модели, т.е. указать вектор параметров \mathbf{a} и множество его допустимых значений A . Также необходимо указать способ построения объекта, получающегося в результате применения деформации, задаваемой вектором параметров \mathbf{a} к модели G_i , т.е. $v(G_i, \mathbf{a})$.*
- в. Поставить задачу определения сходства модели гибкого объекта G_i и силуэта S как задачу поиска наилучшего совмещения силуэтов деформированной модели G_i и представленного силуэта S по формуле 2.4).

- г. Решить оптимизационную задачу из предыдущего пункта, выбрав наиболее подходящий метод минимизации.
- д. Численная оценка сходства модели и силуэта при этом будет иметь вид (2.5).

3. Определение сходства гибкого объекта \mathcal{F} и представленного силуэта S .
Необходимо:

- а. Сформировать оценку сходства гибкого объекта \mathcal{F} и силуэта S на основе имеющихся n оценок сходства этого силуэта и каждой из моделей G_1, \dots, G_n .

Оценка в пункте 1в строится на основе сравнения моделей друг с другом. Следовательно, для его выполнения шаги из пункта 2 также должны быть определены.

Таким образом, предлагаемый подход определяет общую методологию решения задачи. Ниже мы рассмотрим полномасштабную реализацию подхода — его применение в задаче сравнения формы ладони при биометрической идентификации личности — и на нём покажем работу всей схемы.

2.6. Выводы ко второй главе

1. Сформулирована задача сравнения формы гибких объектов и предложен метод её решения, основанный на построении подвижной модели гибкого объекта и подгонке этой модели под силуэт тестового объекта.
2. Разработана модель гибкого объекта, основанная на построении циркулярного представления гибкого объекта, являющегося удобным инструментом описания структуры объекта и множества его трансформаций.

Циркулярное представление может быть построено на основе регуляризации непрерывного скелета бинарного изображения объекта и его радиальной функции.

3. Предложен метод идентификации параметров (разметки) модели по серии изображений объекта в разных положениях. Идентификация параметров осуществляется на основе решения оптимизационной задачи, в которой целевой функцией является средняя мера сходства моделей одного гибкого объекта, а аргументами — параметры модели.
4. Предложен метод сравнения эталонной модели и силуэта тестового объекта, заключающийся в нахождении такой деформации модели, при которой силуэты эталона и теста совпадают наилучшим образом; задача поиска деформации также сформулирована в терминах оптимизационной задачи.
5. Разработан алгоритм решения прикладных задач на основе предложенного метода сравнения формы гибких объектов.

Глава 3

Алгоритмы идентификации и подгонки моделей гибких объектов при распознавании формы ладони

3.1. Модель ладони

В качестве прикладной задачи, на которой демонстрируется применимость и реализуемость предлагаемого подхода сравнения формы гибких объектов, рассматривается задача распознавания личности по форме ладони. Во-первых, ладонь человека хорошо описывается моделью гибкого объекта. А, во-вторых, её трансформации достаточно просты и, потому, шаги предлагаемого подхода легко иллюстрируются и интерпретируются. При этом можно отработать всю методику. Рассмотрим, каким образом может быть построена модель ладони, основываясь на предложенном в разделе 2.1 описании.

3.1.1. Описание ладони гибким объектом

Рассмотрим бинарное изображение произвольной ладони. Пример такого изображения представлен на рисунке 3.1а. Чтобы построить модель гибкого объекта ладони по исходному бинарному изображению, необходимо получить циркулярное представление ладони (рисунок 3.1б) и задать множество допустимых деформаций. Циркулярное представление может быть построено способом, предложенным в [1]. Поэтому перейдём к описанию деформаций.

Если проанализировать основные движения, которые может совершать ладонь при её позиционировании на горизонтальной поверхности, то можно выделить следующие:

- сдвиг ладони;
- поворот ладони;
- поворот пальцев ладони.

Определяемое множество допустимых деформаций должно позволять моделировать эти движения. В терминах циркулярного представления ладони эти движения описываются как:

- сдвиг циркулярного графа;
- поворот циркулярного графа;
- поворот ветвей циркулярного графа.

Первые две трансформации, являющиеся аффинными преобразованиями формы, могут быть достаточно просто применены к любому гибкому объекту. Для применения третьей трансформации — поворота ветвей — необходимо иметь дополнительную информацию о структуре анализируемого объекта. В частности, нужно уметь выделять пальцы и определять положение их точек поворотов.

Таким образом, в качестве *модели ладони* предлагается рассматривать гибкий объект, на котором выделены части ветвей циркулярного графа в качестве пальцев, и указаны их точки поворотов. Иллюстрирующий пример представлен на рисунке 3.1в: для каждого пальца ладони закрашена выделенная для него область на циркулярном графе, а через $O_i, i = 1, \dots, 5$ обозначены точки поворотов пальцев. Процесс поворота i -ого пальца при этом моделируется следующим образом: часть ветви осевого графа, относящаяся к пальцу, поворачивается относительно точки O_i как единое целое; естественно, что с переносом точек осевого графа происходит перенос кругов с центрами в ЭТИХ точках.

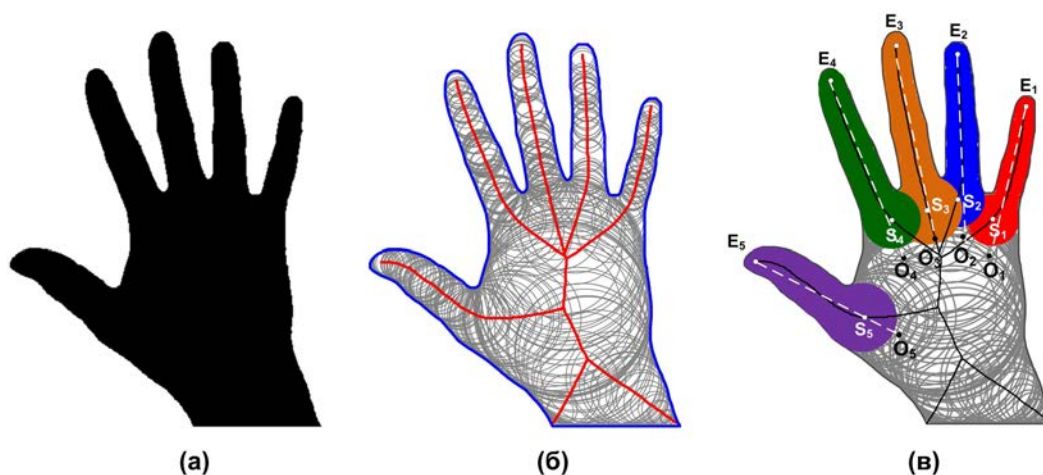


Рис. 3.1. (а) Бинарное изображение ладони, (б) его циркулярный граф и (в) модель ладони.

3.1.2. Параметризация модели

Модель ладони (рисунок 3.2а) может быть представлена как объединение шести компонент фиксированной формы — пяти пальцев и пясти (рисунок 3.2б). К пальцу относится часть ветви циркулярного графа, начинающаяся с некоторой вершины S_i степени 2 или 3 и заканчивающаяся концевой вершиной E_i этой ветви (рисунок 3.1в). То есть, чтобы выделить палец, достаточно указать точку S_i , а точка E_i определится автоматически. Пясть получается путём удаления ветвей $S_i E_i$ из циркулярного графа всей ладони. Допустимые трансформации модели — повороты пальцев относительно точек O_i на некоторый угол. Таким образом, модель ладони описывается вектором из десяти компонент: $\mathbf{p} = (\{S_i, O_i\}_{i=1}^5)$. Будем считать, что модель ладони *идентифицирована* (или *размечена*), если для неё указано численное значение вектора \mathbf{p} .

3.1.3. Построение силуэта ладони

Рассматривается следующая **постановка задачи**. Дана размеченная модель ладони $G(\mathbf{p})$, т.е. выделены пальцы и указаны их точки поворотов.

Модель ладони подвергается некоторой трансформации, связанной с поворотом компонент-пальцев на некоторый угол относительно точек O_i . *Необходимо* построить силуэт модели ладони после выполнения этой трансформации.

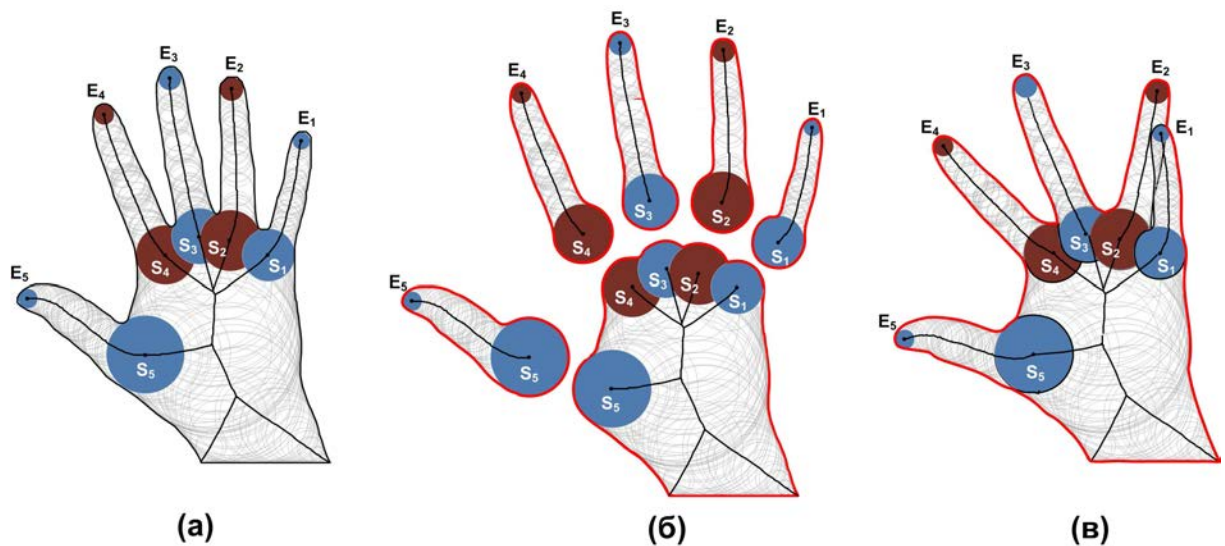


Рис. 3.2. (а) Модель ладони; (б) разложение на компоненты с неизменяемым силуэтом и (в) построение огибающей шести компонент.

В разделе 2.2.2 рассматривался алгоритм построения силуэта гибкого объекта для случая, когда объект представляется в виде объединения конечного числа компонент фиксированной формы, а трансформации таковы, что не влияют на связность силуэта объекта. Для предложенной модели ладони данные требования выполнены. Поскольку ладонь позиционируется на горизонтальной поверхности, её допустимые трансформации допускают только окклюзии, связанные со «слипанием» пальцев. В редких случаях при выполнении трансформаций могут возникать «дырки» в силуэте ладони. Это происходит, когда пальцы соприкасаются не всей своей границей, а лишь её частью. Например, на рисунке 3.2в возникает «дырка» между безымянным пальцем и мизинцем. Однако такие «дырки» обычно малы, и можно считать, что они попросту отсутствуют. Таким образом, связность силуэта модели ладони при трансформациях не меняется и метод построения силуэта модели

по силуэтам её компонент применим.

3.2. Алгоритм идентификации модели ладони

Рассматривается следующая **постановка задачи**. Дано n бинарных изображений ладони одного человека, по которым построено n неразмеченных моделей G_1, \dots, G_n . Необходимо разметить их, т.е. указать значение вектора параметров \mathbf{p} , одинакового для всех моделей.

Предлагаемый *алгоритм* решения этой задачи основан на методе идентификации параметров модели, описанном в разделе 2.3. Идея заключается в поиске такого набора параметров \mathbf{p} , при котором модели G_1, \dots, G_n являются наиболее похожими друг на друга. Основные шаги алгоритма:

1. определение ветвей — кандидатов в пальцы;
2. определение кончиков и оснований пальцев;
3. определение точек поворотов пальцев.

Первые два шага могут быть выполнены независимо для каждой из имеющихся n неразмеченных моделей. Для определения точек поворотов пальцев предлагаются два подхода — «простое назначение» этих точек и «оптимизационный подход». «Простое назначение» (или экспертный метод) может проводиться отдельно для каждой модели G_i ; в «оптимизационном подходе» точки поворотов ищутся по информации обо всех n моделях сразу.

3.2.1. Определение ветвей — кандидатов в пальцы

Будем рассматривать ветви осевого графа ладони G_i , начинающиеся с вершины степени 3 (максимально возможная степень вершины осевого графа) и заканчивающиеся вершиной степени 1, при этом все остальные верши-

ны этой ветви имеют степень 2. На рисунке 3.3а цифрами обозначены все вершины степени 1 и 3. Так, вершины 1, 2, 3, 4, 5, 6, 7 и 8 имеют степень 1, тогда как вершины 9, 10, 11, 12, 13 и 14 имеют степень 3. Таким образом, для данного осевого графа имеем 8 ветвей: 13–1, 13–2, 14–3, 14–4, 11–5, 10–6, 9–7 и 9–8 (для краткости указаны только начальная и конечная вершины).

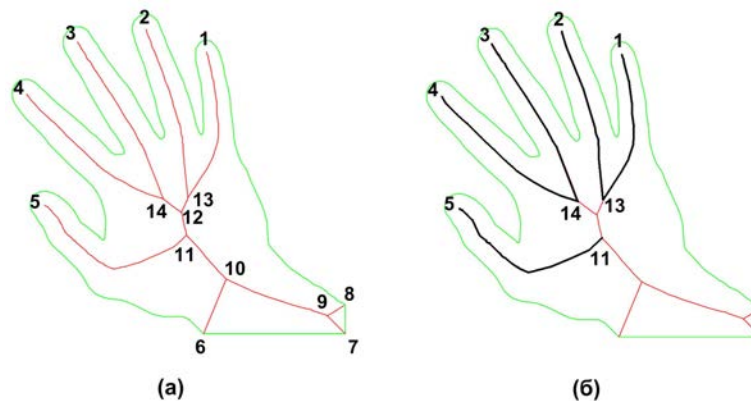


Рис. 3.3. Выделение пальцев по циркулярному графу ладони.

Пять самых длинных ветвей осевого графа определяют *ветви пальцев* ладони. Ветви пальцев выделены жирным цветом на рисунке 3.3б. Это ветви 13–1, 13–2, 14–3, 14–4, 11–5. Если количество ветвей меньше пяти, то пальцы выделить невозможно, и в этом случае модель не может быть размечена. Далее считаем, что пять самых длинных ветвей определены. Однако после выполнения этого шага ещё неизвестно, какая из ветвей относится к большому пальцу, какая — к указательному и т.д.

3.2.2. Определение кончиков и оснований пальцев

Рассмотрим одну из ветвей осевого графа ладони G_i , относящуюся к анализируемому пальцу, и начнём последовательно перебирать все соседние пары кругов, заданных в вершинах циркулярного графа, от точки ветвления до конечной вершины. Для каждой пары (предыдущий и текущий круги)

проверяется выполнение следующих условий:

$$\begin{cases} r \leq r_0 \\ \left[\begin{array}{l} r - r_p \leq 0 \\ \alpha \geq \alpha_0 \end{array} \right. \end{cases} \quad (3.1)$$

Здесь α — угол между двумя радиусами, проведёнными из центра круга в точки касания этого круга с силуэтом ладони; r — радиус текущего круга, а r_p радиус предыдущего круга в рассматриваемой паре; α_0 и r_0 — пороговые константы, определённые в результате проведения ряда экспериментов. Вершина осевого графа, соответствующая центру круга из пары, удовлетворяющей всем указанным условиям, объявляется основанием пальца.

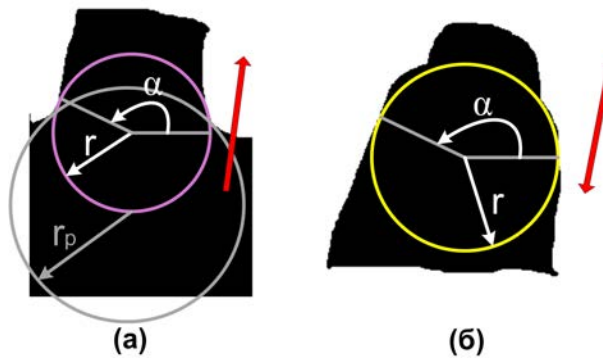


Рис. 3.4. Поиск (а) основания и (б) кончика пальца.

Далее определяется положение кончика пальца. Просматриваются последовательно все круги циркулярного графа из концевой вершины в точку ветвления до тех пор, пока не найдётся круг, для которого выполнены условия:

$$\begin{cases} \alpha \geq \alpha_0 \\ r \geq R(r_{root}) \end{cases} \quad (3.2)$$

В этом случае α снова угол между двумя радиусами, проведёнными из центра круга в точки касания этого круга с силуэтом ладони; r — радиус рассматриваемого круга, а r_{root} — радиус круга в основании пальца; α_0 — пороговая константа; $R(r_{root})$ — некоторая эмпирически подобранная функция, зависящая от радиуса круга в основании пальца (например, $R(r_{root}) = 0.65 r_{root}$). Иллюстрация поиска кончика и основания пальца приведена на рисунке 3.4. Такой подход к нахождению кончика пальца позволяет «обрезать» ногти на изображении и, тем самым, минимизировать их влияние при сравнении ладоней. Обозначим через P_1, P_2, P_3 и P_4 найденные кончики пальцев, а через R_1, R_2, R_3 и R_4 — основания пальцев. Далее через выделенные точку основания R_i и кончик пальца P_i проводится прямая R_iP_i , называемая *осью пальца*.

Теперь можно идентифицировать пальцы модели, т.е. определить, какая из ветвей относится к большому пальцу, какая — к среднему и т.д. Для этого сначала имеющиеся пальцы упорядочиваются в соответствии с обходом осевого графа по часовой стрелке. Затем определяется пара соседних пальцев, угол между осями которых максимален. Это есть большой палец и мизинец. Обозначим их номера через i и j . Чтобы отличить их, вычисляется две величины: d_1 и d_2 , где d_1 — евклидово расстояние между основаниями i -ого пальца и его соседа (не совпадающего с j), а d_2 , соответственно, — расстояние между основаниями j -ого пальца и его соседа (не совпадающего с i). Если $d_1 \geq d_2$, то палец с номером i считается большим, а с номером j — мизинцем; если $d_1 < d_2$, то наоборот.

3.2.3. Определение точек поворотов пальцев

Если посмотреть на строение ладони (рисунок 3.5), то можно отметить, что наиболее подвижным является основание проксимальной фаланги (за ис-

ключением большого пальца, для которого подвижным является также место сочленения проксимальной и дистальной фаланги). Для простоты изложения рассмотрим вариант определения точек поворотов четырёх пальцев ладони, исключив при этом большой палец.



Рис. 3.5. Строение кисти руки и запястья человека⁵.

Далее предлагаются два способа определения точек поворотов пальцев — экспертный метод (по одной модели ладони) и оптимизационный (на основе нескольких моделей ладони). Рассмотрим каждый из них в отдельности. В обоих подходах предполагается, что точка поворота пальца O_i находится на его оси R_iP_i . А в качестве точки S_i рассматривается вершина осевого графа, ближайшая (в смысле евклидовой метрики) к точке O_i . Указанные предположения позволяют перейти от вектора параметров модели \mathbf{p} к вектору параметров $\mathbf{l} = (l_1, l_2, l_3, l_4)$, где $l_i = |P_iO_i|$ есть расстояние от кончика i -ого пальца до точки поворота этого пальца вдоль его оси (рисунок 3.6б).

⁵ «Кисти рук и запястья». Плакат. ISBN 978-5-17-042700-0, ООО «Издательство АСТ», 2007 г.

Экспертный метод

В качестве точки поворота пальца назначается точка на оси пальца, отстоящая от его основания на $k\%$ его длины, где k — параметр метода (например, в [4] рассматривается случай $k = 30$). Таким образом:

$$l_i = \left(1 + \frac{k}{100}\right) \cdot len_i \quad (3.3)$$

где $len_i = |P_i R_i|$ — длина i -ого пальца ладони (рисунок 3.6б).

Эксперименты показали (раздел 4.5.2), что такое определение точки поворота с параметром $k \in [20; 40]$ позволяет получать адекватные результаты при моделировании поворотов пальцев. На рисунке 3.7 приведена иллюстрация разметки ладони.

Оптимизационный подход

При назначении точек поворотов пальцев в экспертном метода достаточно одного изображения ладони. В случае, когда имеется несколько моделей ладони в разных положениях, задача определения точек поворотов пальцев может рассматриваться в постановке, предложенной в разделе 2.3.

Пусть задана функция $\mu^* = \mu^*(G_1, G_2)$, определяющая степень сходства гибких объектов G_1 и G_2 . Эта функция вводится ниже в разделе 3.3.3. Значение этой функции неявно зависит от выбора точек поворотов пальцев ладони, т.к. при её вычислении проводятся трансформации гибкого объекта G_1 с целью получения наилучшего совмещения силуэтов G_1 и G_2 . В число таких трансформаций входят также повороты пальцев и, следовательно, результат трансформации зависит от выбора точек поворотов. Чтобы указать, что рассматривается гибкий объект G с некоторым выбранным положением точек поворотов пальцев \mathbf{l} , будем писать $G(\mathbf{l})$.

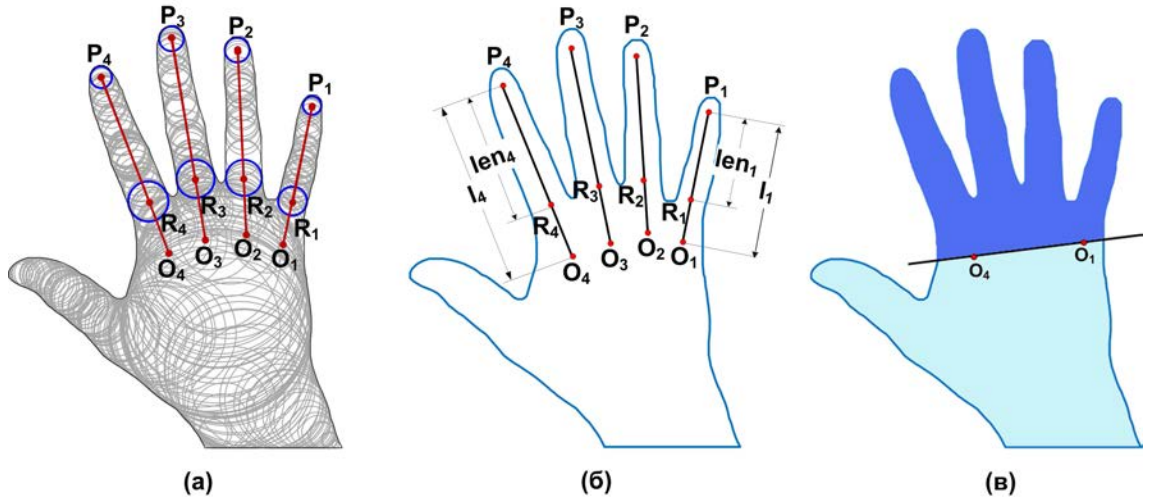


Рис. 3.6. (а) Поиск кончиков и оснований пальцев; (б) поиск точек поворотов и (в) выделение области ладони для сравнения.

Тогда интегральная мера сходства n ладоней согласно (2.1) будет выглядеть следующим образом:

$$\mu_{av}(\mathbf{l}) = \frac{1}{n(n-1)} \sum_{\substack{i,j=1,\dots,n \\ i \neq j}} \mu^*(G_i(\mathbf{l}), G_j(\mathbf{l})) \quad (3.4)$$

Значение $\mu_{av}(\mathbf{l})$ напрямую зависит от того, каким образом выбираются точки поворотов пальцев. Тогда оптимизационная задача (2.2) имеет вид:

$$\mu_{av} : \mathbb{R}^4 \rightarrow \mathbb{R}_+ \quad (3.5)$$

$$\mathbf{l}^* : \mu_{av}(\mathbf{l}^*) = \min_{\mathbf{l} \in L} \mu_{av}(\mathbf{l}) \quad (3.6)$$

где $L \subseteq \mathbb{R}^4$ задаёт допустимое множество для \mathbf{l} . Экспериментально показано, что оно может быть ограничено пространственным параллелепипедом: $L = \{l_i \in [1.2 len_i, 1.5 len_i] | i = 1, \dots, 4\}$.

Таким образом, имеем задачу минимизации функции четырёх переменных. Известно, что функция неотрицательна и конечна, при этом дифференцируемость и гладкость не гарантированы. Функция μ_{av} , также как и μ (которая вводится ниже в разделе 3.3.2), является многоэкстремальной, поскольку

при любом прикладывании сравниваемых ладоней можно так «пошевелить» пальцы, что будет достигнут её локальный минимум. При решении задачи поиска точек поворотов пальцев могут быть использованы методы как условной (с учётом того, что $\mathbf{l} \in L$), так и безусловной оптимизации.

В диссертации задача (3.5) была решена двумя способами — полным перебором каждого l_i с шагом 0.05 len_i в рамках указанного выше пространственного параллелепипеда L , а также методом Нелдера-Мида [61]. Начальный 4-симплекс для метода Нелдера-Мида содержал следующие 5 пространственных точек: $\mathbf{l}^1, \mathbf{l}^2, \mathbf{l}^3, \mathbf{l}^4, \mathbf{l}^5$, где:

$$\mathbf{l}^i = (k_1^i \cdot \text{len}_1, k_2^i \cdot \text{len}_2, k_3^i \cdot \text{len}_3, k_4^i \cdot \text{len}_4) \quad (3.7)$$

$$k_j^i = \begin{cases} 1.45, & \text{если } i = j; \\ 1.35, & \text{если } i \neq j. \end{cases} \quad (3.8)$$

Параметры метода: коэффициент отражения $\alpha = 1$, коэффициент сжатия $\beta = 0.5$ и коэффициент растяжения $\gamma = 2$.

На рисунке 3.7 представлен пример определения точек поворотов пальцев по $n = 4$ снимкам ладони в разных положениях. Отдельно представлены результаты для случая полного перебора допустимых значений \mathbf{l} , а также полученные точки поворотов пальцев при оптимизации методом Нелдера-Мида.

Эксперименты, представленные ниже в разделе 4.5.2, показали, что поиск точек поворотов пальцев на основе решения задачи минимизации функции $\mu_{av}(\mathbf{l})$ приводит к меньшим ошибкам FRR и FAR при распознавании личности. При этом использование метода Нелдера-Мида позволяет ускорить этот поиск без существенного влияния на качество распознавания.

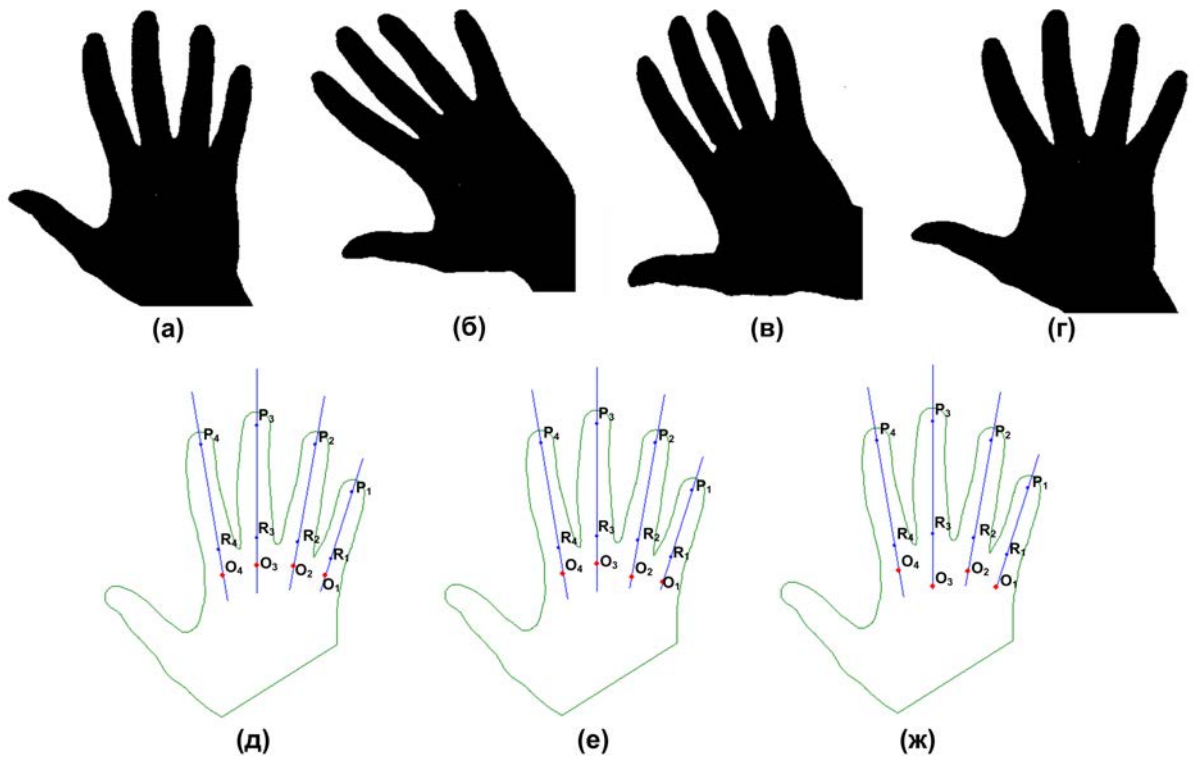


Рис. 3.7. (а),(б),(в),(г) — исходные изображения ладони, $n = 4$. Найденные точки поворотов пальцев (д) при простом назначении с $k = 25$, (е) при оптимизационном подходе с полным перебором на сетке и (ж) оптимизация методом Нелдера-Мида.

3.3. Алгоритм сравнения формы ладоней

Рассматривается следующая **постановка задачи**. Пусть *даны* размеченная модель ладони G_1 и модель (возможно, неразмеченная) G_2 . Назовем первую эталонной, а вторую — тестовой. Тестовая модель может быть неразмеченной в силу того, что на предъявляемом для распознавания изображении присутствуют касающиеся друг друга пальцы. *Необходимо* ввести меру сходства моделей G_1 и G_2 и предложить метод её вычисления. *Решение* этой задачи предлагается строить на основе подхода, предложенного в разделе 2.4. Для этого требуется определить меру μ сходства силуэтов моделей G_1 и G_2 , а также алгоритм поиска их наилучшего совмещения — вычисление меры μ^* .

Процесс сравнения моделей ладоней состоит из двух основных шагов:

1. начальная укладка моделей;
2. подгонка моделей.

Фактически, начальная укладка обеспечивает нам попадание в область притяжения локального минимума функции μ , а подгонка моделей — уточнение положения этого минимума.

3.3.1. Параметризация трансформаций модели

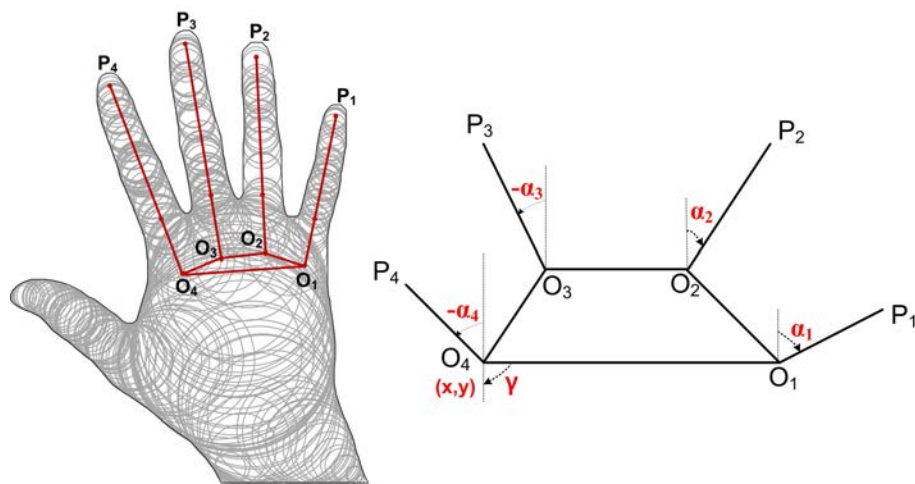


Рис. 3.8. Параметризация трансформаций ладони.

Определим параметры трансформации следующим образом. Рассмотрим углы между осями пальцев и вертикальной осью ординат. Предполагается, что ломаная линия $O_1O_2O_3O_4$, определяющая положение оснований проксимальных фаланг пальцев, имеет фиксированную структуру и не меняется при возможных поворотах пальцев в горизонтальной плоскости. То есть длины отрезков ломаной и углы между этими отрезками фиксированы. При поворотах пальцев возможно лишь изменение углов между отрезками O_iP_i и вертикальной прямой. Обозначим эти углы через α_i . Параметризация ладони проиллюстрирована на рисунке 3.8.

Будем обозначать через $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ набор углов между осями пальцев и вертикальной осью ординат, через (x, y) — координаты точки O_1 в глобальной системе координат, а через γ — угол между прямой O_1O_4 и осью ординат. Таким образом, предлагается определять трансформации ладони набором из семи параметров: $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, x, y, \gamma)$.

3.3.2. Начальная укладка моделей

Начальная укладка моделей включает в себя следующие шаги:

1. совмещение кончиков средних пальцев;
2. совмещение осей средних пальцев;
3. исключение области большого пальца;
4. исключение области запястья.

Как отмечалось выше, пункт 3 включён сюда только для простоты изложения.

Таким образом, при укладке ладоней предлагается совмещать оси R_3P_3 и кончики P_3 средних пальцев. Понятно, что на изображениях ладони одного и того же человека форма запястья может быть различной. Это обусловлено как разницей в позиционировании ладони, так и присутствием рукавов одежды, часов, браслетов и т.д. С целью исключения влияния разницы в форме запястья человека и подвижности кожи в области большого пальца при сравнении учитывается лишь область, лежащая выше прямой O_1O_4 эталонной ладони. Эта область выделена тёмным цветом на рисунке 3.6в.

Рассмотрим в качестве меры сходства двух нетрансформированных ладоней G_1 и G_2 нормированную симметрическую разность их силуэтов:

$$\mu(G_1, G_2) = \frac{\text{Area}(\mathcal{S}(G_1) \setminus \mathcal{S}(G_2)) + \text{Area}(\mathcal{S}(G_2) \setminus \mathcal{S}(G_1))}{\text{Area}(\mathcal{S}(G_1) \cup \mathcal{S}(G_2))} \quad (3.9)$$

Здесь $\mathcal{S}(G_1)$ и $\mathcal{S}(G_2)$ — силуэты гибких объектов G_1 и G_2 соответственно. Таким образом, при начальном сравнении ладони совмещаются, строятся их силуэты, выделяется область для сравнения, вычисляется площадь каждой из полученных фигур, а также их симметрическая разность. Силуэтом произвольного гибкого объекта, согласно построению (раздел 2.2.2), является многоугольная фигура. Поэтому задача вычисления меры μ сводится к задаче поиска нормированной симметрической разности двух многоугольных фигур. Для решения этой задачи в литературе предложено несколько подходов [59]. В работе реализован метод, при котором сначала определяется множество точек пересечения сторон двух многоугольных фигур (методом плоского заметания [60]), а затем строится их внешняя огибающая (объединение многоугольников). Тогда симметрическая разность вычисляется по формуле:

$$\mu(G_1, G_2) = 2(\text{Area}(\mathcal{S}(G_1) \cup \mathcal{S}(G_2))) - \text{Area}(\mathcal{S}(G_1)) - \text{Area}(\mathcal{S}(G_2)) \quad (3.10)$$

Однако понятно, что при простом совмещении ладоней хорошего совпадения их силуэтов получить не удастся. Укладка ладоней позволяет устранить разницу в смещении и повороте ладоней друг относительно друга, но не учитывает возможные шевеления пальцев. Поясняющий пример приведён на рисунке 1. В первом ряду представлены два изображения ладони одного человека. Если построить их циркулярные представления, выполнить начальную укладку и совместить, то получим картинку в третьей колонке. С точки зрения функции близости μ эти ладошки окажутся совсем непохожими друг на друга. Однако понятно, что все дело в том, что при повторном предъявлении человек просто пошевелил пальцами. Если смоделировать движение пальцев, то можно добиться хорошего совмещения силуэтов ладоней, как, например, на рисунке 3.9б. Таким образом, необходима процедура, которая бы позволяла моделировать «шевеление» пальцев ладони с целью получения

наилучшего совмещения силуэтов сравниваемых ладоней.

3.3.3. Подгонка моделей

Согласно (2.5) мера μ^* будет выглядеть следующим образом:

$$\mu^*(G_1, G_2) = \min_{\mathbf{a} \in A} \mu(\mathcal{S}(v(G_1, \mathbf{a})), G_2) = \mu(v(G_1, \mathbf{a}^*), G_2) \quad (3.11)$$

Таким образом, сходство гибких объектов G_1 и G_2 есть минимальное возможное сходство силуэтов трансформированного объекта G_1 и объекта G_2 .

Определим множество A . В силу способа укладки ладони однозначно определяются три компоненты вектора параметров — это координаты точки $O_1 = (x, y)$ и угол γ . Таким образом, вектор параметризации \mathbf{a} существенно зависит от четырёх величин $\mathbf{a} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. Вычислим значения этих углов по тестовой модели и обозначим их через $\alpha_1^t, \alpha_2^t, \alpha_3^t$ и α_4^t . Введём допустимое множество A как:

$$A = \{\alpha_i \in [\alpha_i^t - \Delta\alpha_i, \alpha_i^t + \Delta\alpha_i] | i = 1, \dots, 4\} \quad (3.12)$$

Здесь мы исходим из того, что при сравнении пары ладоней одного человека наилучшее совмещение их силуэтов произойдёт, если углы между соответствующими пальцами на эталонной и тестовой ладонях окажутся одинаковыми. Однако рассчитать точные значения углов α_i^t для тестовой ладони не всегда возможно (например, в случае частично соприкасающихся пальцев). Величина $\Delta\alpha_i$ отражает небольшие локальные «шевеления» пальцев и связана с погрешностью вычисления осей пальцев тестовой ладони. Проведённые эксперименты показали, что значения $\Delta\alpha_i = 5^\circ$ вполне достаточно для получения хорошего совмещения силуэтов сравниваемых ладоней.

Процесс вычисления μ^* является процессом подгонки одного гибкого объ-

екта под другой, т.е. подгонкой эталонной ладони под тестовую. Предлагаемый метод подгонки включает в себя два основных шага:

- построение начального приближения;
- поиск локального минимума.

Построение начального приближения

На данном шаге проводится начальное «шевеление» эталонной ладони, и определяется локальная область поиска минимума функции μ . Обозначим через $\mathbf{a}^0 = (\alpha_1^0, \alpha_2^0, \alpha_3^0, \alpha_4^0)$ начальное приближение для вектора параметров \mathbf{a} . Начальное приближение для вектора \mathbf{a} строится следующим образом:

$$\alpha_i^0 = \alpha_i^t, i = 1, \dots, 4 \quad (3.13)$$

В силу начальной укладки гибкого объекта $\alpha_3^t = 0$.

Поиск локального минимума

После того, как начальное совмещение сравниваемых ладоней выполнено, запускается процесс поиска локального минимума. Отметим, что углы отклонения осей пальцев от вертикальной оси $\alpha_1, \alpha_2, \alpha_3$ и α_4 можно рассматривать независимо.

Поиск локального минимума функции μ проводился в работе двумя способами — полным перебором с заданным шагом и методом Нелдера-Мида [61].

В процессе поиска локального минимума первым способом производится последовательный перебор значений α_i из отрезка $[\alpha_i^t - \Delta\alpha_i, \alpha_i^t + \Delta\alpha_i]$ с шагом в 2° . Набор параметров $\mathbf{a}^* = (\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_4^*)$, на котором получено минимальное значение функции μ , считается оптимальным. На рисунках 3.9 и 3.10

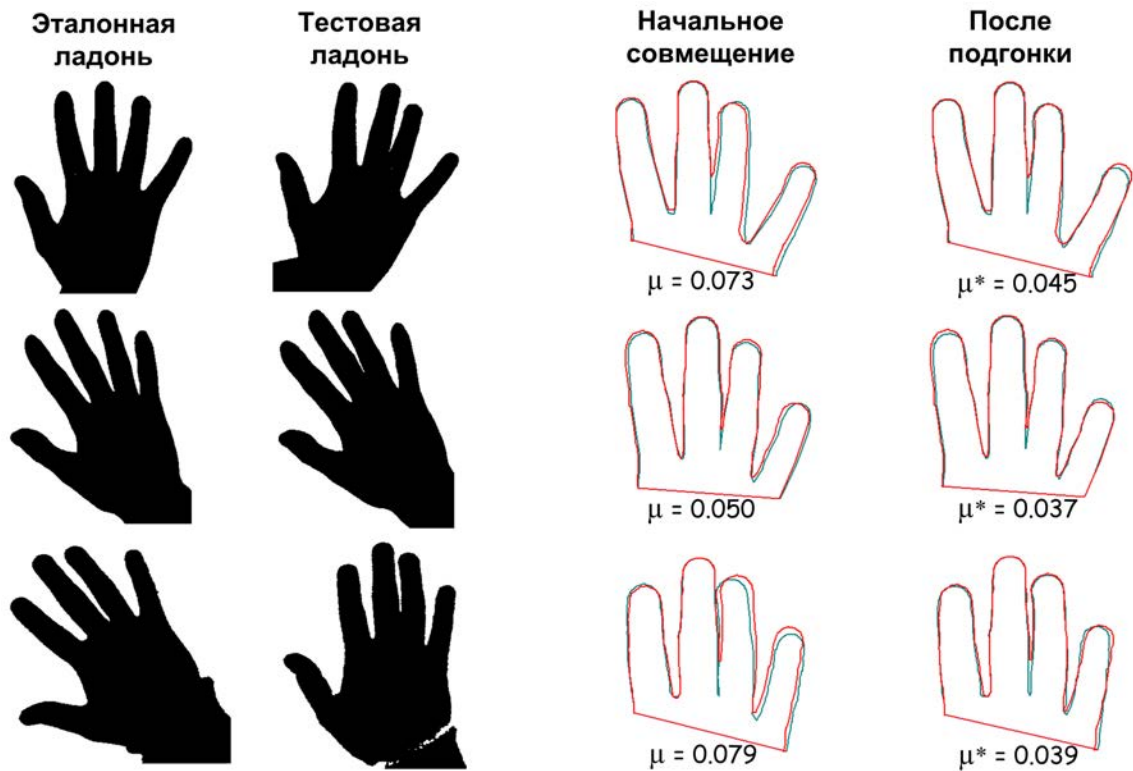


Рис. 3.9. Сравнение ладоней одного человека.

представлены примеры определения наилучшего совмещения ладоней одного человека и разных людей.

Метод Нелдера-Мида запускался с начальным 4-симплексом, содержащим 5 точек из A : $\mathbf{a}^1, \mathbf{a}^2, \mathbf{a}^3, \mathbf{a}^4, \mathbf{a}^5$, где:

$$\mathbf{a}^i = (\alpha_1^t + k_1^i \cdot \Delta\alpha_1, \alpha_2^t + k_2^i \cdot \Delta\alpha_2, \alpha_3^t + k_3^i \cdot \Delta\alpha_3, \alpha_4^t + k_4^i \cdot \Delta\alpha_4) \quad (3.14)$$

$$k_j^i = \begin{cases} 1, & \text{если } i = j; \\ -1, & \text{если } i \neq j. \end{cases} \quad (3.15)$$

Параметры метода: коэффициент отражения $\alpha = 1$, коэффициент сжатия $\beta = 0.5$ и коэффициент растяжения $\gamma = 2$.

Эксперименты показали, что метод Нелдера-Мида позволяет получить оптимальные значения μ , близкие к случаю полного перебора, однако процесс сходится за гораздо меньшее время. Поэтому в конечной программе исполь-

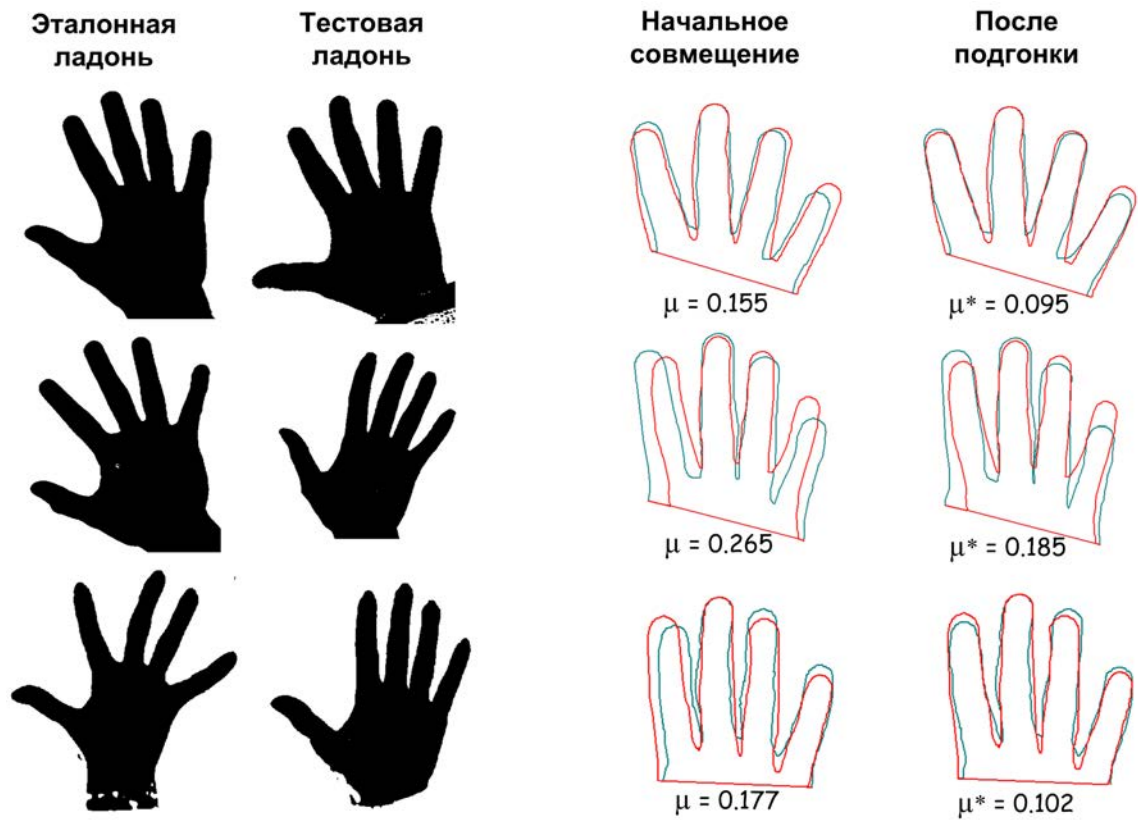


Рис. 3.10. Сравнение ладоней разных людей.

зован именно ЭТОТ подход.

3.4. Выводы к третьей главе

Для решения задачи распознавания формы ладони при наличии частично соприкасающихся пальцев в работе предложен метод, реализующий подход к сравнению формы гибких объектов, представленный в главе 2. А именно:

1. Предложена модель ладони человека. Модель строится по бинарному изображению ладони и представляется в форме гибкого объекта, к которому применим ряд деформаций, отражающих основные движения ладони и пальцев при позиционировании на горизонтальной поверхности.

сти. К таким движениям относятся сдвиг и поворот ладони, поворот пальцев. Предложен способ моделирования этих движений с помощью представленной модели ладони.

2. Предложен способ параметризации модели ладони, задающий положение пальцев и их точек поворотов.
3. Предложено два способа автоматической идентификации параметров модели — простое назначение точек и оптимизационный подход. В первом случае точка поворота пальца вычисляется как линейная функция (заданная заранее) от длины пальца; для её определения достаточно одного эталонного изображения ладони. В оптимизационном подходе разметка модели строится по серии снимков ладони в разных положениях.
4. Разработан метод сравнения моделей ладоней. Предложенный метод основан на подгонке эталонной (размеченной) модели под силуэт тестовой (возможно неразмеченной) модели. Соответственно, была введена мера сходства моделей ладоней и способ ее вычисления.

Глава 4

Система распознавания личности по форме ладони и голосовому паролю

Целью данной главы является практическая проверка рассмотренных ранее принципов и реализация предложенного подхода.

4.1. Многомодальный подход к распознаванию

Признаки формы ладони не являются уникальными. Существует люди с реально похожими ладонями, поэтому геометрия ладони не может быть использована для проведения надёжного распознавания в большой группе людей [62]. Системы распознавания личности по форме ладони могут быть использованы лишь в местах, требующих низкий и средний уровень безопасности. В силу этого, признаки ладони начинают использоваться как составные части многомодальных систем.

Способ математического выражения информации об объектах в задачах анализа данных принято называть *модальностью* представления объектов [63]. В терминах выбранной модальности каждый объект реального мира характеризуется некоторым значением в пространстве соответствующего обобщённого признака, например, в виде сигнала, изображения и т.д. Рассмотрим существующие многомодальные системы.

В работах [64–66] представлены системы распознавания личности по признакам геометрии ладони и её отпечатка. При использовании этих признаков были получены следующие оценки качества распознавания: $FAR = 0\%$, $FRR = 0.2\%$ в [64]; $FAR = 0.01\%$, $FRR = 4\%$ в [65]; $FAR = 0\%$, $FRR =$

1.41% в [66].

В [67] дополнительно к признакам геометрии ладони и её отпечатка используются отпечатки пальцев. Результаты по верификации $FRR = 0\%$ и $FAR = 0.01\%$.

Комбинация геометрии ладони, отпечатка ладони и текстуры пальцев представлена в [68]. Полученные оценки составляют $FAR = 0.31\%$, $FRR = 2.9\%$

Использование изображения лица и геометрии ладони рассматривается в [69]. В работе представлены результаты экспериментов на различных базах и способах комбинирования модальностей. Процент ошибок составил приблизительно 3%.

Таким образом, использование дополнительной модальности позволяет повысить качество распознавания, даже если модальности в отдельности могут не обладать высокой отличительной способностью. Далее предлагается один вариант реализации этой идеи — двумодальный метод, основанный на комбинировании признаков ладони и голоса человека, основанный на сведении задачи идентификации к задаче верификации личности.

4.2. Описание и общая схема двумодального подхода

В рамках предлагаемого двумодального подхода для каждого пользователя хранится не только набор эталонных размеченных моделей его ладони, а также набор эталонных записей парольного слова. Считается, что каждый пользователь системы обладает собственным паролем, причем пароли разных людей могут совпадать. В качестве пароля может выступать, например, фамилия человека. При распознавании человек предъявляет свою ладонь перед камерой и произносит парольное слово. Таким образом, в распознавании

участвуют бинарное изображение ладони и запись парольного слова с микрофона.

Рассматриваемые характеристики ладони и голоса человека могут быть вычислены и проанализированы по отдельности. Возможность независимого рассмотрения модальностей обусловила выбор типа их комбинирования — каскадная модель.

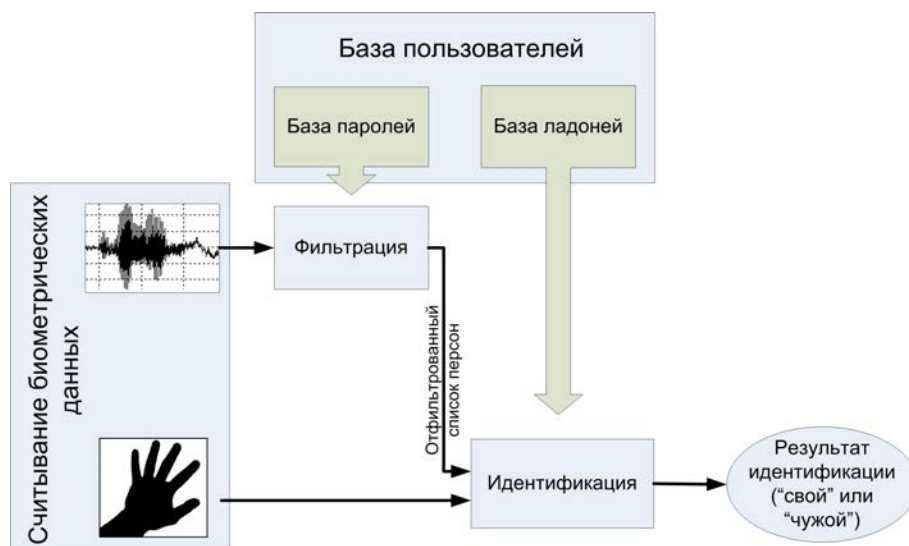


Рис. 4.1. Общая схема двумодальной системы распознавания личности.

На рисунке 4.1 представлена общая схема работы системы. Ее основные компоненты:

- модуль считывания биометрических данных;
- база персон — пользователей системы;
- модуль фильтрации;
- модуль идентификации.

Модуль считывания биометрических данных позволяет получать изображения ладони и осуществлять запись парольного слова для дальнейшего

распознавания. В состав модуля входит web-камера и микрофон. Следующий компонент системы — база персон, содержащая в себе всю биометрическую информацию, а также некоторые дополнительные сведения о пользователях системы. И, наконец, основные компоненты системы — модуль фильтрации и модуль идентификации, которые реализуют предлагаемый метод распознавания. Модули разделены, и каждый работает со своей биометрической характеристикой: фильтрация производится по голосовым признакам, идентификация — по форме ладони.

На этапе фильтрации выполняется исключение заведомо непохожих персон из дальнейшего рассмотрения. Здесь осуществляется последовательное сравнение произнесенного парольного слова со всеми эталонными записями парольных слов, хранящимися в базе. Далее составляется список «похожих» паролей и, следовательно, персон. Отфильтрованный список персон и предъявленное изображение ладони поступают в модуль идентификации. Здесь выносится окончательное решение о том, является ли человек пользователем системы («свой») или нет («чужой»). Такое разделение позволяет осуществлять распознавание человека со скоростью, приемлемой для работы системы в реальных условиях. Это достигается, во-первых, за счет использования достаточно быстрого метода фильтрации персон по парольному слову и, во-вторых, проведения идентификации уже в рамках небольшой группы персон.

4.3. Фильтрация на основе голосового пароля

Для фильтрации персон по парольному слову была использована модификация метода, предложенного в [70]. Дадим его краткое описание.

Рассмотрим дискретный звуковой сигнал $S = \{s(i)\}_{i=1}^n$, где n — число из-

мерений сигнала. Построение признакового описания такого сигнала состоит из следующих шагов:

1. Сигнал разбивается на перекрывающиеся фреймы. Размер фрейма составляет $t_f = 512$ измерений, величина перекрытия — $t_0 = 100$ измерений. Тогда общее число «полных» фреймов есть $n_f = \lfloor \frac{n-t_0}{t_f-t_0} \rfloor$. Будем обозначать через $s_p(j) = s((t_f - t_0)p + j)$ множество измерений сигнала в p -ом фрейме: $p = 1, \dots, n_f$, а $j = 0, \dots, t_f - 1$.
2. Каждый фрейм умножается на оконную функцию Хэмминга:

$$w(j) = 0.54 - 0.46 \cos\left(\frac{2j\pi}{t_f - 1}\right) \quad (4.1)$$

В результате имеем фрейм с измерениями $x_p(j) = w(j)s_p(j)$.

3. Для каждого полученного на предыдущем шаге фрейма рассчитывается набор $y_p(m), m = 0, \dots, t_f - 1$ кепстральных коэффициентов [71] по следующей формуле:

$$y_p(m) = \text{Re} \left[\frac{1}{t_f} \sum_{l=0}^{t_f-1} \exp\left(i \frac{2\pi l m}{t_f}\right) \ln \left[\left| \sum_{j=0}^{t_f-1} x_p(j) \exp\left(-i \frac{2\pi l j}{t_f}\right) \right| + 1 \right] \right] \quad (4.2)$$

где i — мнимая единица, Re — действительная часть, а \exp — экспонента.

4. Для каждого фрейма из полученного набора кепстральных коэффициентов отбираются первые 12 (за исключением 0-ого, т.к. он содержит лишь информацию об энергии сигнала) [72], т.е. $\mathbf{y}_p = (y_p(1), \dots, y_p(12))$.
5. Весь сигнал есть совокупность фреймов, каждый из которых описывается 12 параметрами.

Пусть теперь необходимо сравнить два звуковых сигнала S_1 и S_2 , описываемых наборами кепстральных коэффициентов $(\mathbf{y}_1^1, \dots, \mathbf{y}_{n_1}^1)$ и $(\mathbf{y}_1^2, \dots, \mathbf{y}_{n_2}^2)$

соответственно, где n_1 — число фреймов в первом сигнале, а n_2 — во втором. В качестве расстояния между двумя фреймами рассматривается обычное евклидово ($i = 1, \dots, n_1, j = 1, \dots, n_2$):

$$\rho(\mathbf{y}_i^1, \mathbf{y}_j^2) = \sqrt{\sum_{k=1}^{12} (y_i^1(k) - y_j^2(k))^2} \quad (4.3)$$

Идея сравнения двух сигналов заключается в нахождении такого соответствия между их фреймами, при котором суммарное различие между фреймами является наименьшим:

$$\tau^*(f) : \sum_{f=1}^{n_1} \rho(\mathbf{y}_f^1, \mathbf{y}_{\tau^*(f)}^2) = \min_{\tau \in T} \sum_{f=1}^{n_1} \rho(\mathbf{y}_f^1, \mathbf{y}_{\tau(f)}^2) \quad (4.4)$$

здесь $\tau(f)$ — перестановка индексов фреймов, а T — множество всех возможных перестановок, сохраняющих монотонность, т.е. если $f_1 \leq f_2$, то $\tau(f_1) \leq \tau(f_2)$.

Постановка задачи (4.4) сравнения пары паролей позволяет применить метод динамического программирования [70, 73] для её решения.

Для каждого пользователя в системе хранится несколько эталонных записей парольных слов. При распознавании нового (тестового) парольного слова определяется степень его принадлежности к каждой из персон в базе. Она вычисляется как минимальное расстояние между тестовым словом и всеми эталонными записями пароля конкретной персоны. Чем меньше это расстояние, тем более тестовое слово похоже на пароль персоны.

Таким образом, тестовое слово сравнивается со всеми эталонными, и далее строится список персон, упорядоченный по убыванию степени сходства. Из вершины этого списка отбираются k персон, среди которых будет проводиться идентификация по форме ладони (k — параметр метода).

4.4. Описание установки и комплекса программ

Для демонстрации работоспособности предложенного подхода к распознаванию личности по форме ладони и голосовому паролю была собрана установка и разработан комплекс программ. Этот подход нашел свое воплощение в экспериментальной системе "Time & Attendance", регистрирующей посещение студентами занятий в университете.

4.4.1. Экспериментальная установка

Система состоит из следующих основных компонент:

- *Web-камера со встроенным микрофоном Logitech QuickCam[®] Pro 4000.* Размер получаемых с камеры снимков ладони составляет 640×480 пикселей, разрешение — 1.3 мегапикселя, позволяет делать до 30 кадров в секунду. Запись парольных слов производится в моно режиме (8 бит) с частотой 22050 Гц.
- *Компьютер.* Процессор Intel(R) Core(TM)2 Duo CPU T8100, частота — 2.1 ГГц, память (RAM) — 2 Гб.
- *Площадка для позиционирования ладони.* В работе это был стол с однотонной поверхностью, на которой ладонь легко различима.

Камера устанавливается над площадкой и калибруется. Под калибровкой здесь понимается определение относительных масштабов получаемого изображения ладони. А именно, для круга фиксированного радиуса в сантиметрах (R_{cm}) замеряется его радиус в пикселях (R_{pix}) на получаемом с камеры изображении. Далее вычисляется величина $z = R_{cm}/R_{pix}$. В процессе распознавания все изображения ладоней, поступающие с камеры, масштаби-

руются с коэффициентом z . При изменении положения камеры параметр z пересчитывается.

При предъявлении ладони человек помещает её на площадке тыльной стороной вверх. Как только в области видимости камеры появляется ладонь, система просит произнести пароль. После чего записанный сигнал и изображение ладони подаются на вход программе двумодального распознавания. По результатам распознавания определяется, является ли человек пользователем системы или посторонним. При неудачной попытке распознавания человек может её повторить. Все результаты регистрируются в электронном виде в журнале посещаемости.

На рисунке 4.2 представлены фотографии системы: слева — процесс предъявления ладони, а справа — общий вид установки.



Рис. 4.2. Установка для двумодального распознавания личности по форме ладони и голосовому паролю.

4.4.2. Программная реализация

Система двумодального распознавания личности по форме ладони и голосовому паролю была написана на языке Object Pascal в среде разработки

Delphi 7.0. Программа имеет оконный интерфейс.

Система поддерживает следующие режимы работы:

- Работа с базой пользователей системы:
 - создание новой базы пользователей;
 - загрузка существующей базы пользователей;
 - добавление и удаление пользователей из базы;
 - регистрация биометрических данных пользователей (эталонных записей паролей и изображений ладоней);

- Настройка окружения:
 - настройка параметров камеры (контрастность, яркость, экспозиция и т.д.);
 - настройка параметров микрофона (громкость, формат записи);

- Режимы распознавания личности:
 - верификации;
 - идентификации.

Данные режимы работы устанавливаются из главного окна программы. После установки одного из режимов появляется дочернее окно программы, в котором содержатся необходимые инструменты для работы в выбранном режиме. На рисунке 4.3 представлен основной пользовательский интерфейс программы.

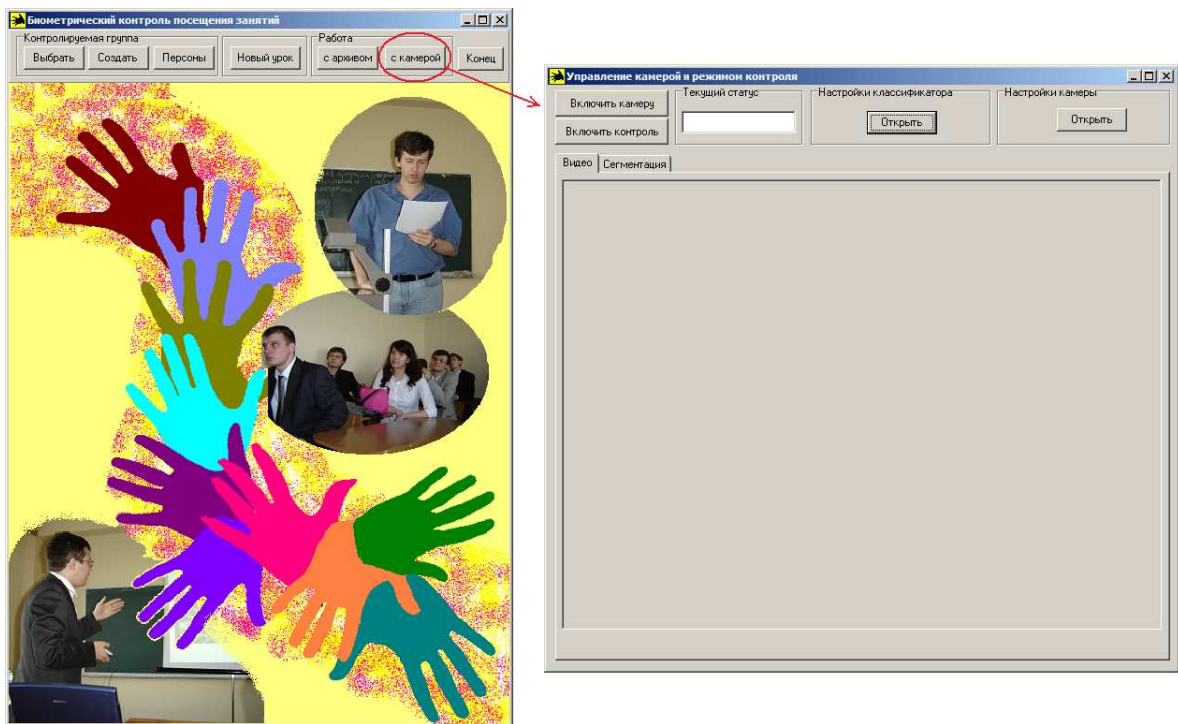


Рис. 4.3. Основной пользовательский интерфейс программы.

4.5. Вычислительные эксперименты

На основе полученной оценки сходства формы сравниваемых ладоней может решаться задача распознавания человека. При этом можно рассматривать оба случая — как задачу верификации, так и задачу идентификации личности [62]. Кроме того, наличие численной оценки позволяет настраивать и применять хорошо известные классификаторы, такие как метод k ближайших соседей, пороговые правила, нейронные сети и т.д, а также их комбинации.

Представим один вариант решения задачи распознавания личности по форме ладони, основанный на использовании порогового правила. Рассмотрим последовательно задачи верификации и идентификации личности.

Отметим, что для каждого человека в базе эталонных ладоней содержится несколько моделей ладони. Это происходит потому, что для каждого сде-

ланного бинарного изображения ладони строится отдельный гибкий объект, и затем проводится его параметрическая идентификация либо отдельно по этому объекту, либо по всей построенной совокупности гибких объектов (раздел 3.2.3). Пусть для i -ого пользователя имеется n моделей его ладони: G_1^i, \dots, G_n^i . Будем считать, что все ладони одного человека образуют один класс. Пусть G — некоторая произвольная ладонь. Будем определять степень сходства ладони G и ладоней i -ого человека как:

$$\mu_{cl} = \min_{j=1, \dots, n} \mu^*(G_j^i, G) \quad (4.5)$$

где μ^* было определено выше в разделе 3.3.3.

Таким образом, сходство некоторой ладони и класса ладоней человека есть минимальное из расстояний между этой ладонью и каждой из моделей в классе.

Эксперименты проводились на установке, которая подробно описывается в разделе 4.4.2. С ее помощью была собрана и сформирована база ладоней группы людей, а именно, 241 изображение ладоней 52 человек. Эти изображения были вручную разделены на две группы: 152 были отнесены к эталонам (2-3 для каждого человека) и 89 к контролю (1-3 для каждого человека). Среди тестовых ладоней 11% были с частично соприкасающимися пальцами.

4.5.1. Распознавание личности по форме ладони

Плотность распределения

Для имеющейся базы ладоней были построены гистограммы плотности распределения внутриклассовых и межклассовых расстояний. Каждая из тестовых ладоней сравнивалась с каждой из эталонных, и вычислялась мера μ^* . При этом, если сравнение проводилось для ладоней одного человека, то

расстояние считалось внутриклассовым, иначе — межклассовым.

На рисунках 4.4 и 4.5 представлены гистограммы плотности распределения внутриклассовых и межклассовых расстояний после начальной укладки ладоней (раздел 3.3.3) и в результате подгонки (раздел 3.3.3). По горизонтальной оси отложены значения μ^* , по вертикальной — значения плотности распределения. При построении модели ладони для определения точек поворотов пальцев использовался оптимизационный подход.

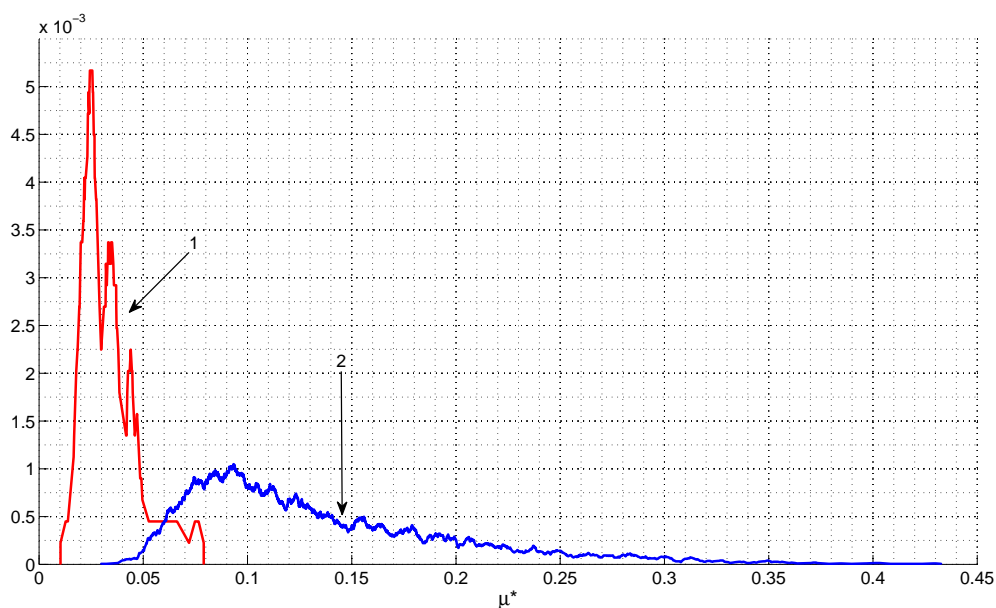


Рис. 4.4. Гистограмма плотности распределения 1 — внутриклассовых и 2 — межклассовых расстояний при начальной укладке.

Как видно по гистограммам, предложенный процесс подгонки позволяет увеличить различие между ладонями разных людей и уменьшить между ладонями одного человека. Ладони одного человека и разных людей достаточно хорошо отделимы.

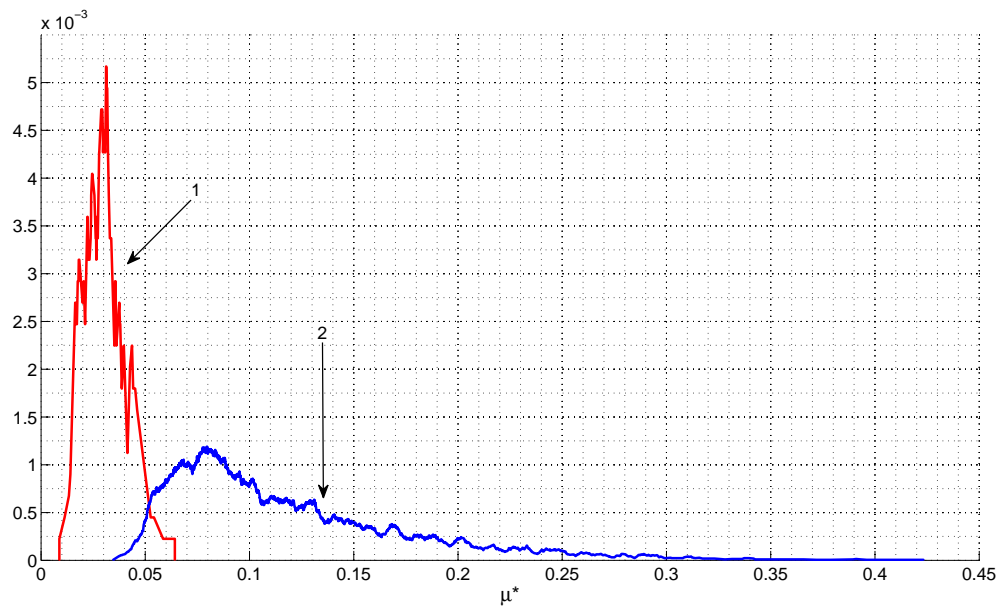


Рис. 4.5. Гистограмма плотности распределения 1 — внутриклассовых и 2 — межклассовых расстояний в результате подгонки.

Верификация

Верификация личности на основе формы ладони человека была реализована в системе следующим образом:

1. Для каждого пользователя системы заводится уникальный идентификационный номер. В качестве такого номера был выбран порядковый номер регистрации человека в базе.
2. При распознавании человек вводит этот номер и, далее, предъявляет свою ладонь.
3. Сходство μ_{cl} представленного изображения ладони и эталонных моделей ладони этого человека рассчитывается по формуле 4.5.
4. При принятии решения используется пороговое правило: если $\mu_{cl} \leq t$, то верификация считается успешной и человек рассматривается как

«свой», в противном случае он считается «чужим». Здесь m — величина порога.

Качество верификации определялось на основе двух критериев — FRR и FAR . При оценке FRR рассчитывалась доля тестовых ладоней, для которых расстояние μ_{cl} до кластера эталонных ладоней человека, к которому относилась тестовая ладонь, было больше величины порога m . При оценке FAR для каждой тестовой ладони из эталонов исключались все ладони того же человека. Далее вычислялось расстояние между данной тестовой ладонью и кластерами всех остальных пользователей. Если расстояние хотя бы до одного кластера было не больше величины порога m , то регистрировалась ошибка ложного узнавания. Ошибка FAR есть доля ладоней, для которых было зафиксировано ложное узнавание.

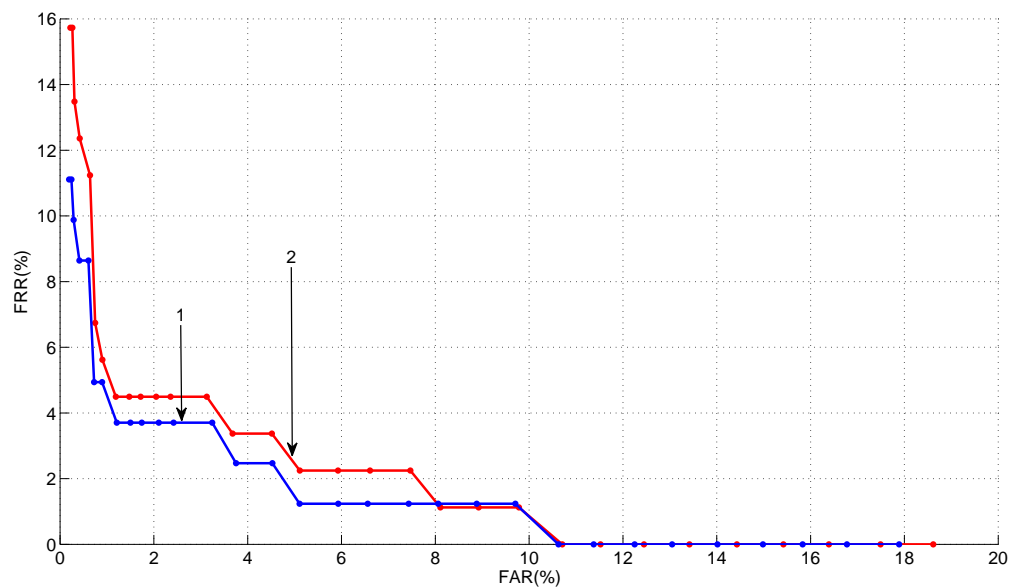


Рис. 4.6. Результаты верификации: кривая с меткой "1" — результат распознавания ладоней без частично соприкасающихся пальцев, кривая с меткой "2" — с частично соприкасающимися пальцами.

На рисунке 4.6 представлены полученные результаты по верификации личности.

Величины ERR для обоих случаев почти совпадают и составляют 3.5%. Как видно из приведённого графика, ладони с частично соприкасающимися пальцами распознаются практически также, как и ладони с хорошо разделёнными пальцами.

Идентификация

Задача идентификации личности по форме ладони также решалась путём применения порогового правила:

1. Предъявленное тестовое изображение ладони последовательно сравнивается со всеми классами эталонных изображений ладоней пользователей.
2. Определяется ближайший класс с точки зрения меры μ_{cl} .
3. Если расстояние до ближайшего класса не больше величины порога t , то ладонь считается принадлежащей человеку, чей класс является ближайшим. В противном случае считается, что на изображении представлена ладонь, не принадлежащая ни одному человеку из базы.

При оценке FRR и FAR использовались те же соображения, что и в случае верификации. При оценке FRR рассчитывалась доля тестовых ладоней, для которых идентификация прошла с ошибкой, т.е. либо ладонь была отнесена не к тому пользователю, либо не отнесена ни к одному человеку в базе. При оценке FAR для каждой тестовой ладони из эталонов исключались все ладони того же человека. Далее проводилась идентификация, и, если ладонь была отнесена хотя бы к какому-нибудь пользователю, фиксировалась ошиб-

ка ложного узнавания. Ошибка FAR есть доля ладоней, для которых было зафиксировано ложное узнавание.

На рисунке 4.7 представлены полученные результаты по идентификации личности.

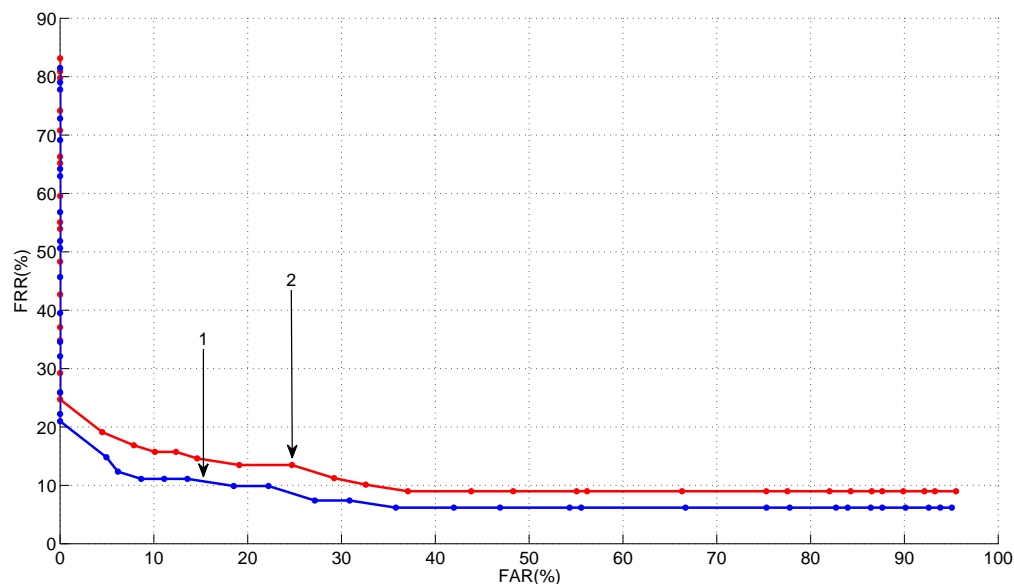


Рис. 4.7. Результаты идентификации: кривая с меткой "1" — результат распознавания ладоней без частично соприкасающихся пальцев, кривая с меткой "2" — с частично соприкасающимися пальцами.

В случае идентификации личности видно небольшое отличие качества распознавания на ладонях с окклюзиями и без. Полученное значение ERR для ладоней с хорошо разделёнными пальцами составляет 11% против 14.5% для ладоней с частично соприкасающимися пальцами.

Как показывают результаты проведённых экспериментов, форма ладони обладает невысокой отличительной способностью при использовании её для идентификации личности. Так, наилучшее полученное значение ERR составляет 11%. При $FRR = 0\%$ величина FAR достаточно высока и составляет 20%. При этом она не опускается ниже 6%. Таким образом, для проведе-

ния надежной идентификации выбранный способ сравнения форм ладоней не подходит. Однако он может быть использован как составной элемент много-модальных систем. Например, когда задача идентификации может быть сведена к задаче верификации личности. Один вариант использования формы ладони в составе двумодальной системы рассмотрен в этой главе (раздел 4.2), а результаты — в разделе 4.5.4.

4.5.2. Разметка моделей ладоней

Качество разметки модели ладони оценивалось на основе экспериментов по идентификации и верификации личности 4.5.1. При этом сравнивалось качество распознавания при экспертном назначении точек поворотов пальцев и в случае оптимизационного подхода.

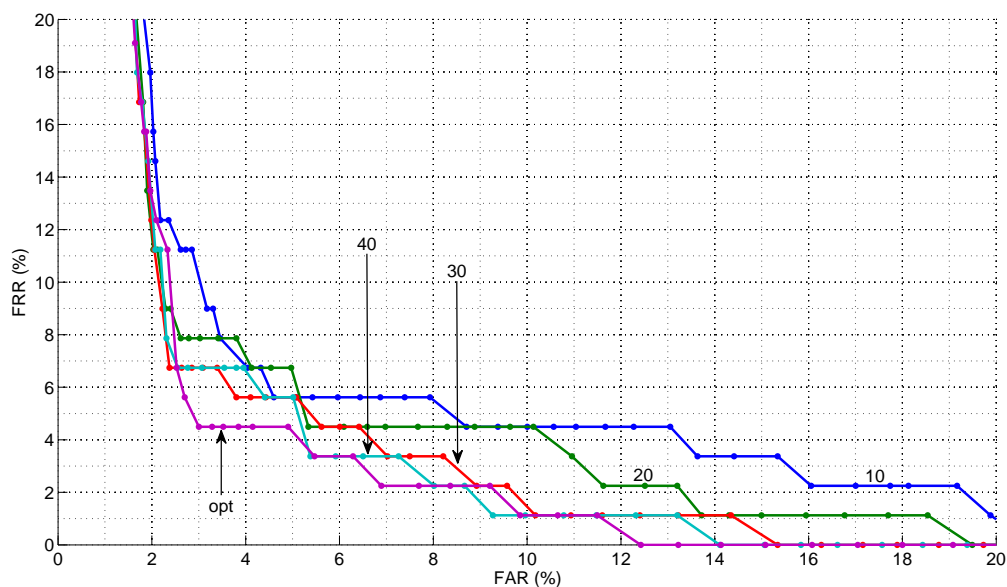


Рис. 4.8. Сравнение результатов верификации при экспертном назначении точек поворотов пальцев и в случае оптимизационного подхода. Четыре кривые — назначение точек с параметром k равным соответственно 10, 20, 30 и 40; пятая кривая — оптимизационный подход при полном переборе значений l с шагом в 5% (кривая с меткой "opt").

На рисунке 4.8 представлены результаты верификации при экспертном назначении точек поворотов пальцев и в случае оптимизационного подхода. Условные обозначения такие: четыре кривые — назначение точек с параметром k равным соответственно 10, 20, 30 и 40; пятая — оптимизационный подход при полном переборе значений \mathbf{l} с шагом в 5% (кривая с меткой "opt").

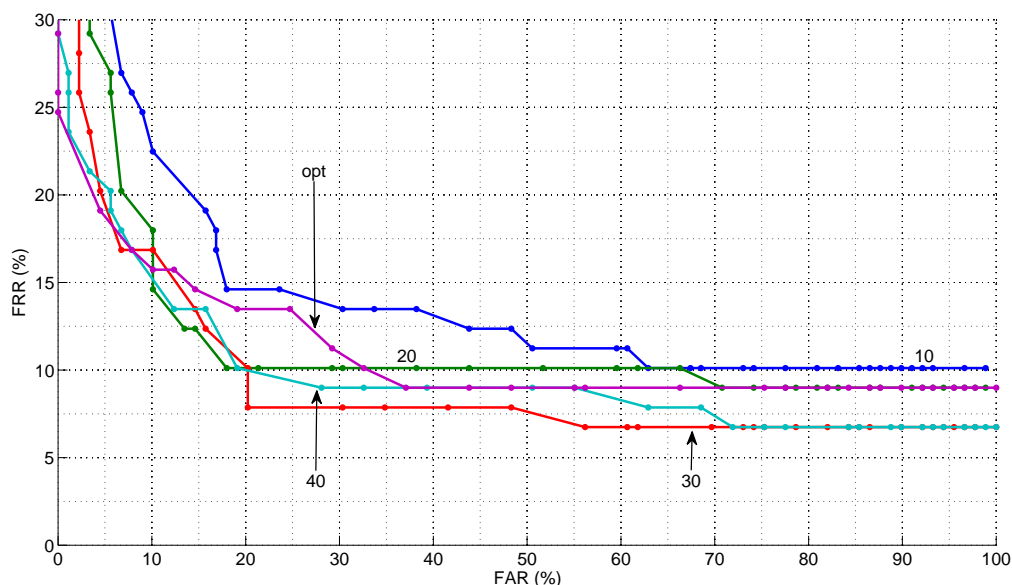


Рис. 4.9. Сравнение результатов идентификации при экспертном назначении точек поворотов пальцев и в случае оптимизационного подхода. Четыре кривые — назначение точек с параметром k равным соответственно 10, 20, 30 и 40; пятая кривая — оптимизационный подход при полном переборе значений \mathbf{l} с шагом в 5% (кривая с меткой "opt").

Как видно, в случае верификации определение точек поворотов пальцев по серии снимков ладони в разных положениях показывает явное преимущество по сравнению с экспертным назначением. Полученное значение ERR составляет для него 4.5%. Однако в случае идентификации (рисунок 4.9), когда тестовая ладонь сравнивается с ладонями всех пользователей, а не одного человека, как при верификации, рассматриваемые подходы ведут себя примерно одинаково. ERR составляет примерно 15% для оптимизационного

подхода и 13% в случае $k = 30$.

4.5.3. Распознавание личности по голосовому паролю

Для оценки качества фильтрации персон по парольному слову был организован следующий эксперимент. Была собрана база записей парольных слов группы людей — для каждого человека было сделано несколько записей его парольного слова. Далее последовательно перебирались все записи из базы, и каждая сравнивалась со всеми остальными записями. Если в число k ближайших рассматриваемой записи попадала хотя бы одна запись того же самого человека, фильтрация считалась успешной.

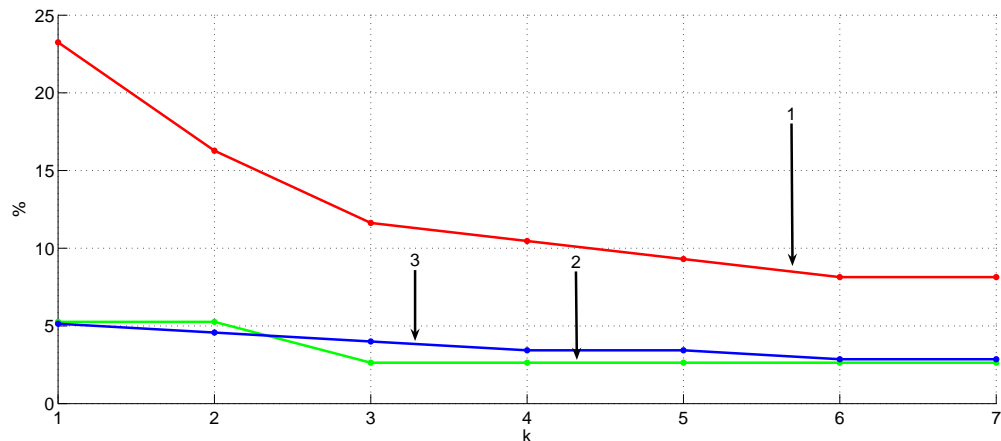


Рис. 4.10. Доля ошибок фильтрации на 1 — Базе 1, 2 — Базе 2 и 3 — Базе 3.

Эксперименты проводились на трех базах:

- База 1: 86 записей 18 человек (4-5 записей на каждого человека), сильно зашумленная
- База 2: 76 записей 20 человек (2-5 записей на каждого человека), зашумленная

- База 3: 175 записей 26 человек (6-7 записей на каждого человека), уровень шума минимален

На рисунке 4.10 представлены результаты по оценке качества фильтрации. По горизонтальной оси отложены рассматриваемые значения k , по вертикальной оси — доля записей (в %), для которых фильтрация выполнялась с ошибкой. Как видно по графику, на записях с невысоким уровнем шума ошибка фильтрации составляет порядка 3-5%.

4.5.4. Двумодальный подход

Для проведения эксперимента была составлена объединенная база, содержащая в себе ладони и записи парольных слов 20 человек. Из исходных данных было сформировано 38 эталонных пар вида «ладонь-пароль» и 117 тестовых. Оценка величины FAR осуществлялась так же, как и при идентификации по форме ладони: ошибка фиксировалась, если система признавала человека как «своего» при распознавании по базе, из которой этот человек был заведомо исключен.

Результаты эксперимента представлены в таблице 4.1. Рассматривались значения, при которых одномодальные классификаторы показывали наилучшие результаты. При фильтрации персон по голосу рассматривалось 3-5 ближайших, в качестве порога для распознавания по форме ладони брались значения в окрестности ERR .

Из полученных результатов можно сделать вывод о том, что рассмотренное комбинирование позволяет значительно улучшить качество распознавания с точки зрения FRR и FAR . Оба этих показателя снизились по сравнению с одномодальным методом. Например, для некоторых значений параметров k и m величина FAR вовсе оказалось равной 0%. Для $k = 3$ и $m = 0.060$ мы имеем $ERR = 1.7\%$.

k	m	<i>FRR</i> , %	<i>FAR</i> , %
3	0.060	1.7	1.7
4	0.060	1.7	6.7
5	0.060	1.7	8.4
3	0.065	3.4	0
4	0.065	3.4	5
5	0.065	3.4	6.7
3	0.070	4.2	0
4	0.070	4.2	0
5	0.070	4.2	0
3	0.075	6.7	0
4	0.075	6.7	0
5	0.075	6.7	0

Таблица 4.1. Результаты идентификации личности при двумодальном подходе при различном выборе параметров m и n .

4.6. Выводы к четвёртой главе

1. Проведённые эксперименты по верификации на реальной базе ладоней людей показали, что с точки зрения качества распознавания (критериев *FRR* и *FAR*) предложенный оптимизационный подход разметки моделей ладоней имеет преимущество. Это может быть объяснено тем, что в предложенном подходе точки поворотов назначаются на основе ладоней одного человека, без участия информации о ладонях других людей. Тогда как в случае задачи идентификации оптимизационный подход работает почти также, как и простое назначение точек поворотов пальцев.

2. Качество распознавания на выборке ладоней с хорошо разделёнными пальцами и на выборке, где было 11% ладоней с частично соприкасающимися пальцами, для задачи верификации оказалось одинаковым ($ERR = 3.5\%$). При идентификации хорошие ладони распознавались лучше по сравнению с ладонями, имеющими окклюзии ($ERR = 11\%$ против 14.5%).
3. Эксперименты по оценке качества распознавания предлагаемой системы распознавания личности по форме ладони показывают, что только формы ладони недостаточно для проведения надёжного распознавания в группе людей. Существуют люди, формы ладоней которых похожи, и, соответственно, для их корректной идентификации нужна дополнительная информация.
4. В качестве такой информации может использоваться «внешняя» информация (пароль, PIN-код), и в этом случае мы имеем задачу верификации личности.
5. Альтернативный вариант — использование других биометрических характеристик человека. Так, в работе предложен двумодальный подход, основанный на комбинировании признаков ладони и голоса человека. Распознавание состоит из двух шагов: на первом шаге происходит фильтрация заведомо непохожих персон по голосовому паролю, а на втором — проводится распознавание по форме ладони в оставшейся небольшой группе персон.
6. Проведённые эксперименты по оценке качества распознавания в случае двумодального подхода показали, что удаётся существенно снизить ошибки FRR и FAR по сравнению с одномодальным подходом. Полу-

ченное в экспериментах значение ERR составляет 1.7%.

7. Представлена экспериментальная система "Time & Attendance", реализующая предложенный в работе двумодальный подход распознавания личности по форме ладони и голосовому паролю.
8. Система позволяет проводить верификацию и идентификацию личности в режиме реального времени на группе персон в количестве порядка 50 человек.

Заключение

По итогам проведённой работы можно сделать следующие выводы:

1. При решении актуальных задач распознавания формы гибких объектов возникают ситуации, когда приходится сравнивать силуэты объектов, имеющих окклюзии. Например, такие ситуации возникают при распознавании поз и жестов человека. Однако существующие методы на такие случаи не ориентированы.
2. Случаи с окклюзиями удобно анализировать, когда известен исходный объект и семейство его допустимых деформаций. Тогда можно пробовать применять эти деформации и смотреть получающийся при этом силуэт объекта. Далее этот силуэт уже можно сравнивать с силуэтами других объектов и оценивать их сходство.
3. Предложенный в работе способ описания формы гибкого объекта в виде совокупности циркулярного графа и множества его допустимых преобразований предоставляет широкие возможности для анализа объекта и его силуэта. В частности, предложенная модель гибкого объекта задаёт его структуру, позволяет легко моделировать различные деформации объекта и восстанавливать его силуэт после их выполнения.
4. Силуэт модели есть фигура, ограниченная внешней огибающей бесконечного семейства кругов циркулярного графа. В работе показано сведение задачи построения такой огибающей к построению огибающей конечного связного множества бициклов. Приводится решение этой задачи для случая, когда исходный объект может быть разделён на небольшое число компонент фиксированной структуры. Разработанный метод построения силуэта такой модели работает за время $O((n_b + n_0) \log n_b)$

в худшем случае и $O(n_b + n_0)$ в среднем, где n_b — число бициклов, а n_0 — общее число точек пересечений границ компонент.

5. Для автоматической идентификации параметров (разметки) модели предложен метод, позволяющий определять оптимальный набор параметров по серии бинарных изображений объекта в разных положениях. Он основан на минимизации среднего различия между моделями одного и того же объекта.
6. Для сравнения модели и силуэта некоторого тестового объекта предложен новый метод, заключающийся в подгонке деформированной модели под представленный силуэт. Применении деформаций к эталонной модели, а не тесту обусловлено тем, что восстановить исходный объект по тестовому изображению не всегда возможно, как, например, в случае окклюзий. В процессе подгонки производится поиск такой деформации модели, при которой достигается наилучшее совмещение её силуэта с силуэтом теста.
7. В работе рассмотрено применение представленного метода сравнения формы гибких объектов к решению задачи распознавания человека по форме ладони. В частности, описана модель ладони человека и множество её допустимых деформаций, введена параметризация модели, представлен упрощённый способ построения силуэта ладони, предложена мера сходства модели и теста (симметрическая разность их совмещённых силуэтов) и метод её вычисления. Рассмотрены задачи верификации и идентификации личности по форме ладони и представлены результаты проведённых экспериментов на реальных данных по оценке критериев FRR и FAR . Эксперименты показали, что форма ладони подходит для проведения верификации личности в небольшой группе

персон, но является ненадёжным признаком для задачи идентификации. Однако качество распознавания может быть улучшено за счёт использования дополнительной биометрической характеристики человека, например, голоса.

8. В работе представлена разработанная экспериментальная система "Time & Attendance" двумодального распознавания личности по форме ладони и голосовому паролю, работающая в режиме реального времени. Она реализует каскадную модель комбинирования классификаторов, когда голос служит фильтром заведомо непохожих персон, а распознавание проводится по форме ладони среди небольшой группы (3 – 7 человек) оставшихся персон. Эксперименты, проведённые на реальных данных, показали существенное улучшение качества распознавания по сравнению с одномодальным подходом — распознаванием по форме ладони, а также приемлемое общее качество распознавания.

Предложенная в работе методология сравнения формы на основе гибкого объекта может быть использована при решении большого класса задач, возникающих в системах машинного зрения. Продемонстрировано удобство использования модели гибкого объекта при описании формы объекта и его допустимых деформаций, а также возможность сравнения объектов с окклюзиями. Применение рассмотренной методологии к решению конкретной прикладной задачи, проведённые эксперименты и собранный прототип системы двумодального распознавания личности показали реализуемость и работоспособность разработанных методов.

Литература

- [1] *Местецкий Л. М.* Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. — ФИЗМАТЛИТ, 2009.
- [2] *Бакина И. Г., Местецкий Л. М.* Многомодальная идентификация личности по форме ладони и голосу // *Таврический Вестник Информатики и Математики.* — 2008. — Т. 1. — С. 59–65.
- [3] *Бакина И. Г.* Генерация признаков при наличии артефактов для задачи распознавания личности по форме ладони // Сборник тезисов XVI Международной научной конференции студентов, аспирантов и молодых учёных «ЛОМОНОСОВ-2009». — М.: МАКС Пресс, 2009. — С. 14.
- [4] *Бакина И. Г., Местецкий Л. М.* Метод сравнения ладоней при наличии артефактов // Доклады 14-ой Всероссийской конференции «Математические методы распознавания образов» ММРО-14. — М.: МАКС Пресс, 2009. — С. 301–304.
- [5] *Бакина И. Г., Цискаридзе А. К., Местецкий Л. М.* Метод активного скелета в задаче распознавания формы изображений // Доклады 19-ой Международной конференция по компьютерной графике и зрению «ГрафиКон'2009». — Москва, МГУ: 2009. — С. 279–282.
- [6] *Бакина И. Г.* Идентификация личности по форме ладони и голосу // Доклады 8-ой международной конференции «Интеллектуализация обработки информации» ИОИ-2010. — М.: МАКС Пресс, 2010. — С. 438–441.
- [7] *Bakina I.* Palm shape comparison for person recognition // Proceedings of the International Conference on Computer Vision Theory and Applications, "VISAPP 2011". — Vilamoura, Algarve—Portugal: 2011. — Pp. 5–11.

- [8] *Бакина И. Г.* Определение точек поворотов пальцев при создании шаблона ладони // Тезисы докладов научно-технической конференции «Техническое зрение в системах управления» TVCS'2011. — Москва, ИКИ РАН: 2011. — Рр. 72–74.
- [9] *Mestetskiy L., Bakina I., Kurakin A.* Hand geometry analysis by continuous skeletons // Lecture Notes in Computer Science.
- [10] *Бакина И. Г.* Идентификация модели ладони по серии её снимков в разных положениях // Доклады 15-ой Всероссийской конференции «Математические методы распознавания образов» ММРО-15. — М.: МАКС Пресс, 2011. — С. 424–427.
- [11] *Бакина И. Г.* Генерация признаков для сравнения ладоней при наличии артефактов // *Научно-практический журнал «Прикладная информатика»*. — 2009. — № 4(22). — С. 68–75.
- [12] *Бакина И. Г.* Моделирование ладони гибким объектом при биометрической идентификации // *Вестник Тверского государственного университета*.
- [13] Словарь русского языка: В 4-х т. Институт русского языка АН СССР. — М.: «Русский язык», 1981.
- [14] *Yang M., Kpalma K., Ronsin J.* A survey of shape feature extraction techniques // *Pattern Recognition Techniques, Technology and Applications*. — IN-TECH, 2008. — Рр. 43–90.
- [15] *Loncaric S.* A survey of shape analysis techniques // *Pattern Recognition*. — 1998. — Vol. 31, no. 8. — Рр. 983–1001.

- [16] *Belongie S., Malik J., Puzicha J.* Shape matching and object recognition using shape contexts // *IEEE Trans. Pattern Anal. Mach. Intell.* — April 2002. — Vol. 24. — Pp. 509–522.
- [17] *LeCun Y., Bottou L., Bengio Y., Haffner P.* Gradient-based learning applied to document recognition. — 1998. — Vol. 86. — Pp. 2278–2324. <http://yann.lecun.com/exdb/mnist/>.
- [18] *Murase H., Nayar S. K.* Visual learning and recognition of 3-d objects from appearance // *Int. J. Comput. Vision.* — January 1995. — Vol. 14. — Pp. 5–24. <http://www1.cs.columbia.edu/CAVE/software/softlib/coil-20.php>.
- [19] *Jeannin S., Bober M.* Description of core experiments for mpeg-7 motion/shape // *Technical Report ISO/IEC JTC 1/SC 29/WG 11 MPEG99/N2690.* — March 1999. <http://knight.cis.temple.edu/~shape/MPEG7/dataset.html>.
- [20] *Felzenszwalb P. F., Schwartz J. D.* Hierarchical matching of deformable shapes // *CVPR.* — 2007.
- [21] *Dryden I., Mardia K.* *Statistical Shape Analysis.* — John Wiley and Sons, 1998.
- [22] *Soderkvist O.* — *Computer Vision Classification of Leaves from Swedish Trees.* — Master's thesis, Linköping University, SE-581 83 Linköping, Sweden, September 2001. — LiTH-ISY-EX-3132. http://www2.cvl.isy.liu.se/ImageDB/images/internal_images/blad/.
- [23] *Sebastian T. B., Klein P. N., Kimia B. B.* Recognition of shapes by editing

- their shock graphs // *IEEE Trans. Pattern Anal. Mach. Intell.* — 2004. — Vol. 26, no. 5. — Pp. 550–571.
- [24] *Hu M.-K.* Visual pattern recognition by moment invariants // *Information Theory, IRE Transactions on.* — 1962. — Vol. 8, no. 2. — Pp. 179–187.
- [25] *Mokhtarian F., Abbasi S., Kittler J.* Efficient and robust retrieval by shape content through curvature scale space. — 1996. — Pp. 35–42.
- [26] *Mokhtarian F., Bober M.* Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization. — Kluwer, 2003.
- [27] *Zhang D., Lu G.* A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval // *J. Visual Communication and Image Representation.* — 2003. — Vol. 14, no. 1. — Pp. 39–57.
- [28] *Zhang D., Lu G.* A comparative study of fourier descriptors for shape representation and retrieval // *Proc. of 5th Asian Conference on Computer Vision (ACCV.* — Springer, 2002. — Pp. 646–651.
- [29] *Zaboli H., Rahmati M., Mirzaei A.* Shape recognition by clustering and matching of skeletons // *JCP.* — 2008. — Vol. 3, no. 5. — Pp. 24–33.
- [30] *Siddiqi K., Shokoufandeh A., Dickinson S. J., Zucker S. W.* Shock graphs and shape matching // *International Journal of Computer Vision.* — 1999. — Vol. 35, no. 1. — Pp. 13–32.
- [31] *Быстров М. Ю.* Применение структурного подхода к распознаванию скелетов бинарных изображений // *Ученые записки Петрозаводского Государственного Университета.* — 2011. — Т. 2, № 115. — С. 76–80.

- [32] *Domakhina L., Okhlopkov A.* Shape comparison based on skeleton isomorphism // *VISAPP* (2). — 2009. — Pp. 237–242.
- [33] *Zhu S. C., Yuille A. L.* Forms: A flexible object recognition and modelling system // *International Journal on Computer Vision*. — 1996. — Vol. 20(3).
- [34] *Местецкий Л. М., Семёнов А.* Сравнение формы изображений на основе циркулярного разложения // Международная Конференция по Компьютерной Графике и Зрению «ГрафиКон'2004». — 2004.
- [35] *Mestetskiy L.* Shape comparison of flexible objects // *VISAPP*. — 2007. — Vol. 1. — Pp. 390–393.
- [36] *Цискаридзе А. К.* Восстановление пространственных циркулярных моделей по силуэтным изображениям: Ph.D. thesis. — 2010.
- [37] *Gonzalez S., Travieso C. M., Alonso J. B., Ferrer M. A.* Automatic biometric identification system by hand geometry // *Proceedings of the 37th Annual International Carnahan Conference on Security Technology*. — 2003. — Pp. 281–284.
- [38] *H.M.El-Bolok, M.I.El-Adawy, M.E.AbouEl-Wafa, R.I.Mubarak.* Person identification through hand geometry measurements // *Journal of Engineering Sciences, Assiut University*. — 2004. — Vol. 32(4). — Pp. 1661–1672.
- [39] *Morales A., Ferrer M. A., Diaz F., Alonso J. B., Travieso C. M.* Contact-free hand biometric system for real environments // *Proceedings of the 16th European Signal Processing Conference (EUSIPCO 2008)*.
- [40] *Jianxia W., Wanzhen Z., Xiaojun W., Min Q.* The feature parameter extraction in palm shape recognition system // *Journal of Communication and Computer*. — 2005. — Vol. 2. — Pp. 25–28.

- [41] *Boreki G., Zimmer A.* Hand geometry: A new approach for feature extraction // *Proceedings of the 4th IEEE Workshop on Automatic Identification Advanced Technologies.* — 2005. — Pp. 149–154.
- [42] *Wong A. L., Shi P.* Peg-free hand geometry recognition using hierarchical geometry and shape matching // *Proceedings of IAPR International Conference on Machine Vision Applications.* — 2002. — Pp. 281–284.
- [43] *Covavisaruch N., Prateepamornkul P., Ruchikachorn P., Taksaphan P.* Personal verification and identification using hand geometry // *ECTI Transactions on Computer and Information Technology.* — 2005. — Vol. 1. — Pp. 134–140.
- [44] *Varchol P., Levicky D.* Using of hand geometry in biometric security systems // *Radioengineering.* — 2007. — Vol. 16. — Pp. 82–87.
- [45] *Местецкий Л. М., Мехедов И.* Комбинированное правило ближайших соседей при классификации формы ладоней // Тезисы докладов Международной конференции Интеллектуализация обработки информации. — 2006. — С. 142–144.
- [46] *Sanchez-Reillo R.* Hand geometry pattern recognition through gaussian mixture modeling // *15th International Conference on Pattern Recognition (ICPR'00).* — 2000. — Vol. 2. — Pp. 937–940.
- [47] *Jain A. K., Ross A., Pankanti S.* A prototype hand geometry-based verification system // *Proceedings of 2nd International Conference on Audio- and Video-based Biometric Person Authentication.* — 1999. — Pp. 166–171.
- [48] *Chernoff H.* The use of faces to represent points in k-dimensional space

graphically // *Journal of the American Statistical Association*. — 1973. — Vol. 68(342). — Pp. 361–368.

- [49] *Roy V., Jawahar C. V.* Feature selection for hand-geometry based person authentication // *Proceedings of International Conference on Advanced Computing and Communication*. — 2005.
- [50] *Fouquier G., Likforman L., Darbon J., Sankur B.* The biosecure geometry-based system for hand modality // *ICASSP (1)*. — 2007. — Pp. 801–804.
- [51] *Duda R. O., Hart P. E., Stork D. G.* *Pattern Classification (2nd Edition)*. — 2 edition. — Wiley-Interscience, 2001.
- [52] Biosecure web page. <http://biosecure.it-sudparis.eu/AB/>.
- [53] *Su C.-L.* Index, middle, and ring finger extraction and identification by index, middle, and ring finger outlines // *VISAPP*. — 2008. — Vol. 1. — Pp. 518–520.
- [54] *Konukoglu E., Yoruk E., Darbon J., Sankur B.* Shape-based hand recognition // *Image Processing, IEEE*. — 2006. — Vol. 15. — Pp. 1803–1815.
- [55] *Jain A. K., Duta N.* Deformable matching of hand shapes for verification // *Proceedings of ICIP '99, Kobe, Japan*. — 1999. — Pp. 857–861.
- [56] *Mestetskiy L., Semenov A.* Palm shape comparison based on fat curves // *Pattern Recognition and Image Analysis. (Advances in Mathematical Theory and Applications)*. — 2005. — Vol. 15(4).
- [57] *Yoruk E., Dutagaci H., Sankur B.* Hand based biometry // *SPIE Electronic Imaging Conference: Image and Video Communications and Processing*. — 2005. — Vol. 5685. — Pp. 1106–1115.

- [58] *Duta N.* A survey of biometric technology based on hand shape // Pattern Recognition. — Vol. 42. — 2009. — Pp. 2797–2806.
- [59] *Скворцов А.* Построение объединения, пересечения и разности произвольных многоугольников в среднем за линейное время с помощью триангуляции // *Вычислительные методы и программирование.* — 2002. — Т. 3. — С. 116–123.
- [60] *Франко П., Майкл Ш.* // Вычислительная геометрия: Введение. — М.: Мир, 1989.
- [61] *Nash J. C.* Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation // International Series on Biometrics. — Adam Hilger, 1990.
- [62] *Ross A. A., Nandakumar K., Jain A. K.* Handbook of Multibiometrics // International Series on Biometrics. — Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. — Vol. 6.
- [63] *Татарчук А. И., Елисеев А. П., Моттль В. В.* Комбинирование классификаторов и потенциальных функций в многомодальном распознавании образов // Доклады 13-ой Всероссийской конференции «Математические методы распознавания образов, ММРО-13». — М.: МАКС Пресс, 2007. — С. 220–222.
- [64] *Ribaric S., Ribaric D., Pavesic N.* A multimodal biometric-based user-identification system for internet use // IEE Proceedings Vision, Image & Signal Processing. — Vol. 150(6). — 2003. — Pp. 409–416.
- [65] *Hanmandlu M., Kumar A., Madasu V. K., Yarlagadda P.* Fusion of hand based biometrics using particle swarm optimization // Proceedings

of the Fifth International Conference on Information Technology: New Generations. — Washington, DC, USA: IEEE Computer Society, 2008. — Pp. 783–788.

- [66] *Kumar A., Wong D. C. M., Shen H. C., Jain A. K.* Personal verification using palmprint and hand geometry biometric // Proceedings of 4th International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA). — Guildford, UK: 2003. — Pp. 668–678.
- [67] *Row R. K., Uludag U., Demirkus M., Parthasaradhi S., Jain A. K.* A Multi-spectral Whole-Hand Biometric Authentication System // Biometrics Symposium. — 2007. — Pp. 1–6.
- [68] *Sanches T., Antunes J., Correia P. L.* A single sensor hand biometric multimodal system // 15th European Signal Processing Conference, EUSIPCO'2007. — 2007.
- [69] *Ross A., Govindarajan R.* Feature level fusion using hand and face biometrics // Proceedings of SPIE Conference on Biometric Technology for Human Identification II. — Vol. 5779. — 2005. — Pp. 196–204.
- [70] *Theodoridis S., Koutroumbas K.* Pattern Recognition. Second Edition. — Elsevier, 2003.
- [71] *Ouzounov A.* Cepstral features and text-dependent speaker identification—a comparative study // *Cybernetics and Information Technologies*. — 2010. — T. 10.
- [72] *Karpov E.* Real-time speaker identification // *Master's Thesis*. — 2003.
- [73] *Rabiner L., Juang B.-H.* Fundamentals of Speech Recognition. — Prentice-Hall International Inc., 1993. — P. 507.