

Московский государственный университет имени М. В. Ломоносова



Факультет Вычислительной Математики и Кибернетики

Кафедра Математических Методов Прогнозирования

## **ДИПЛОМНАЯ РАБОТА СТУДЕНТА 517 ГРУППЫ**

### **«Методы анализа формы ладони на основе морфологического спектра»**

Выполнил:

студент 5 курса 517 группы

*Зак Евгений Михайлович*

Научный руководитель:

д.т.н., профессор

*Местецкий Леонид Моисеевич*

# Содержание

<b>1 Введение</b>	<b>3</b>
1.1 Определения и обозначения . . . . .	4
1.2 Известные результаты исследований анализа формы объектов . . . . .	14
1.3 Постановка задачи классификации ладоней с помощью морфологического спектра . . . . .	17
<b>2 Известные и новые подходы к решению задачи анализа формы</b>	<b>19</b>
<b>3 Вычислительные эксперименты</b>	<b>28</b>
3.1 Исходные данные и условия эксперимента . . . . .	30
3.2 Результаты эксперимента . . . . .	32
<b>4 Заключение</b>	<b>41</b>
<b>Список литературы</b>	<b>43</b>

## Аннотация

В работе исследуется вопрос применимости спектрального подхода для решения задачи классификации как признакового описания формы на примере анализа изображения ладони.

Рассматриваются морфологический спектр Марагоса (Pattern spectrum), медиальный спектр (Medial width spectrum) и спектр на основе карты расстояний (Distance map). Сравниваются их вычислительная сложность, свойства и полезность их использования как признаков объектов для классификации.

В тексте работы введены основные определения, проведен краткий обзор литературы по теме, рассмотрены известные и предложены новые алгоритмы построения спектра, рассмотрены основные подходы к сравнению спектров, предложены метрики и проведено их сравнение на реальных данных, приведены результаты применения спектрального подхода к описанию формы ладоней человека для построения признаков и классификации данных. Проведены эксперименты на данных, состоящих из выборки 36 людей и 160 ладоней.

Эксперименты по анализу ладоней производятся по предложенному в работе алгоритму в несколько этапов: бинаризация изображения ладони, обработка шумов изображения, выделение кисти руки путем удаления предплечья, построение спектра и непосредственная классификация. В разделе работы, посвященном экспериментам, показана эффективность предлагаемых подходов, описаны полученные результаты.

В качестве алгоритма классификации рассмотрен алгоритм  $k$  ближайших соседей (kNN). Для оценки точности и выбора алгоритма используется скользящий контроль. Проведены эксперименты с выборкой на верификацию и идентификацию данных.

# 1 Введение

Трудно не согласиться с тем фактом, что появление ЭВМ и развитие информационных технологий открыло много новых возможностей перед человеком. Одной из самых интересных задач является задача биометрической идентификации.

Известны современные подходы к решению задач биометрической идентификации, обеспечивающие высокую точность распознавания, например, идентификация по радужной оболочке глаза или по рисунку папиллярных линий пальцев. Однако зачастую такие технологии являются либо дорогостоящими, либо связаны с рядом ограничений и не доступны для простого пользователя. Вместе с тем, востребованы также более простые методы идентификации, не требующие специального оборудования. Разработке одного из таких методов посвящена настоящая работа.

Предлагается рассмотреть задачу классификации людей по форме их ладоней [11]. В настоящее время известно значительное количество подходов к анализу формы объекта, в том числе использование математической морфологии Серра [3]. В данном исследовании рассматриваются подходы, основанные на построении спектра, в первую очередь морфологических спектров, предложенных П. Марагосом в рамках математической морфологии [1], а также медиальный спектр и спектр на основе карты расстояний (Distance map).

Актуальность предлагаемого подхода обусловлена прежде всего тем, что рассматриваемые спектры являются мощными дескрипторами формы, отражающие свойства толщины объекта. Важен тот факт, что для подсчета морфологического спектра или его аналога достаточно иметь фотоснимок, сделанный стандартной цифровой камерой, которая сейчас доступна любому обывателю: современные персональные компьютеры имеют встроенные веб-камеры, и при необходимости она может быть приобретена даже с доставкой на дом. Предлагаемый подход не предполагает физического контакта ладони человека с устройством съемки, что является важным фактором с психологической точки зрения для значительного количества людей. Таким образом, описанные требования позволяют реализовывать классификацию для большого количества объектов и классов.

Хотя задача выглядит достаточно узкой, её рассмотрение позволяет оценить, насколько спектральный подход в целом применим для создания признакового описа-

ния формы какого-либо класса объектов и насколько это описание применимо для классификации и анализа данных.

В ходе проведения данного исследования получен результат, подтверждающий обоснованность использования спектрального подхода для решения задачи классификации. Подробное описание полученных результатов приводится в главе «Вычислительные эксперименты».

В следующих разделах работы рассматриваются основные определения, описываются предлагаемые подходы к построению спектров, формализуется постановка задачи и приводится обзор релевантной литературы. В главе «Известные и новые подходы к решению задачи анализа формы» вводится общая схема алгоритма, предложенного автором, состоящая из обработки изображения, выделения кисти руки от предплечья с помощью алгоритма, построения спектра и классификации; далее подробно раскрывается каждый из этапов. После описания алгоритма рассматриваются эксперименты над выборкой, состоящей из 160 изображений ладоней 36 различных людей.

## 1.1 Определения и обозначения

В настоящее время существует различные подходы к анализу форм объектов. В данной работе рассматриваются методы, основанные на описании формы областей с помощью методов морфологии Серра [3] а также с помощью непрерывной морфологии в соответствии с книгой [2].

**Определение 1.1.** *Морфологическое сравнение — сравнение изображений объектов, дающее количественную оценку сходства этих объектов по их форме. [3]*

Форма является одним из тех фундаментальных понятий, смысл которых понимают все, но формализовать его затруднительно. Вместе с тем, определение формы является важным для понимания других определений. Поэтому введем понятие формы в соответствии со словарем русского языка:

**Определение 1.2.** *Форма — это внешнее очертание, наружный вид предмета [15].*

В этом же словаре приводится близкое по смыслу значение - фигура. Понятие фигуры может быть введено строго.

**Определение 1.3.** Область на евклидовой плоскости — непустое связное открытое множество точек.

**Определение 1.4.** Нормальная область — ограниченная замкнутая область, граница которой представляет собой объединение конечного числа замкнутых контуров, каждый из которых в свою очередь состоит из конечного числа участков аналитических кривых.

**Определение 1.5.** Фигурой будет считаться любая нормальная область.

Будем использовать также следующие обозначения. Пусть  $\mathbf{R}^2$  — евклидова плоскость. Тогда фигуры будем обозначать так же, как принято обозначать множества — то есть заглавными латинскими буквами, например  $A$  или  $X$ .

Введенные выше определения в общем случае являются непрерывными. Вместе с тем в задачах, связанных с анализом форм объектов, чаще всего используются объекты, изображенные на растровом изображении. Таким образом, вводится следующее определение:

**Определение 1.6.** Растровой решеткой называется квадратная решетка  $\mathbf{Z}^2$ , состоящая из точек евклидовой плоскости с целочисленными координатами.

В данной работе рассматриваются именно растровые изображения, то есть изображения, представленные в виде сетки, каждым узлом которой является пиксель с сопоставленным ему значением цвета. В данном случае, мы рассматриваем некоторое подмножество растровых изображений, а именно бинарные изображения, то есть растровое изображение, каждому пикселю которого поставлено в соответствие одно из двух значений «1» или «0», что соответствует определению метки объекта и фона. Для наглядности и простоты изложения будем называть точки с меткой «1» черными, а с меткой «0» — белыми.

Отметим, что каждое растровое изображение вне зависимости от цветности может быть переведено в бинарное. О методах подобного отражения рассказывается, в частности, в [4].

Поскольку растровая решетка является подмножеством евклидовой плоскости, то введенные понятия области и фигуры корректны и в этом случае, при соответствующем определении связности, которое приведено, в том числе, в [2].

Введем понятие морфологической операции. Бинарное изображение — это множество точек, принимающих одно из двух значений. В общем случае, морфологическая операция является некоторым преобразованием этого множества. В качестве входных данных морфологическая операция принимает фигуру  $X$  и параметры операции, выходными данными является бинарное изображение.

**Определение 1.7.** *Перенос фигуры  $X$  на вектор  $z$  называется фигура*

$$X_z = \{y | x \in X, y = x + z\}$$

**Определение 1.8.** *Эрозией (от лат. *erosio* — разведение) фигуры  $X$  с примитивом (структурным элементом)  $B$  называется отображение вида*

$$X \ominus B = \{z | B_z \subseteq X\}$$

**Определение 1.9.** *Дилатацией (от лат. *dilato* — расширять) фигуры  $X$  с примитивом (структурным элементом)  $B$  называется отображение вида*

$$X \oplus B = \{\cup B_z | z \in X\}$$

В качестве параметра для морфологической операции чаще всего выступает другое изображение, называемое примитивом или структурным элементом. Структурный элемент обычно меньше исходного изображения и является некоторой простой и симметричной формой, например кругом или квадратом. В таких случаях структурный элемент задается радиусом и своим названием, например  $Box(r)$ ,  $Disk(r)$  и т.д. Пример структурного элемента в виде круга с радиусом  $r = 2$  изображен на Рис. 1.

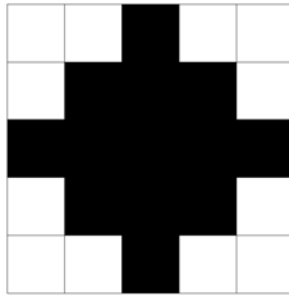


Рис. 1: Пример структурного элемента в виде круга с радиусом  $r=2$

Если рассматривать введенные выше операции с точки зрения работы с растровой решеткой, то можно сказать, что дилатация является морфологической операцией,

которая служит, в основном, для того, чтобы закрыть мелкие дыры и узкие щели у объекта. Работает она по следующему принципу: в каждую точку фигуры  $A$  помещается центр структурного элемента  $B$ . Далее, проходя по всем точкам объекта и помещая центр структурного элемента, производится логическое сложение структурного элемента в данном положении с бинарным изображением. Принцип работы операции дилатации изображен на Рис. 2. Очевидно, что помимо того, что применение операции позволяет избавиться от мелких дыр и щелей, результатом ее применения станет увеличение площади объекта.

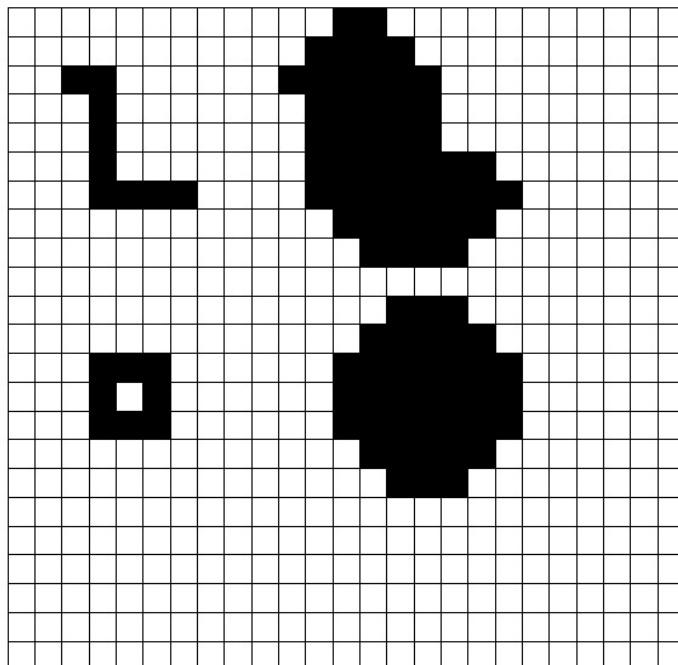


Рис. 2: Пример применения операции дилатации со структурным элементом круг с  $r=2$

Результатом действия эрозии является логическое пересечение множества точек фигуры  $A$  и множества центров переносов  $B_z, \forall z$  структурного элемента  $B$ , которые полностью лежат в множестве объектов. Таким образом, результатом данной операции является удаление тонких линий и мелких деформаций объекта. Примеры работы эрозии со структурным элементом кругом, изображенным на Рис. 1, показаны на Рис. 3. Стоит отметить, что одним их эффектов от применения эрозии является уменьшение площади фигуры.

Для того, чтобы сохранить полезные свойства этих операций и при этом оставить неизменными основные свойства формы фигуры, были предложены операции



замыкания (закрытия) и размыкания (открытия). Применение этих операций заключается в последовательном применении эрозии и дилатации (в случае с операцией открытия) и последовательном применении дилатации и эрозии (в случае операции закрытия).

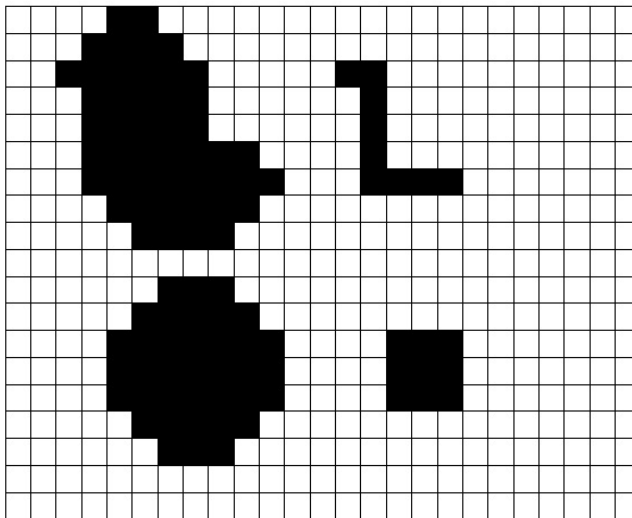


Рис. 3: Пример применения операции эрозии со структурным элементом круг с  $r=2$

**Определение 1.10.** *Открытием фигуры  $X$  с примитивом (структурным элементом)  $B$  называется отображение вида:*

$$X \circ B = (X \ominus B) \oplus B = \{x | x \in B_z, B_z \subseteq X\}$$

**Определение 1.11.** *Закрытием фигуры  $X$  с примитивом (структурным элементом)  $B$  называется отображение вида*

$$X \bullet B = (X \oplus B) \ominus B = \{x | B_x \subseteq \{\cup B_z | z \in X\}\}$$

Рассмотрим более подробно операцию открытия. Применение операции эрозии позволяет избавиться от линий, которые тоньше структурного элемента, при этом уменьшая общую площадь фигуры. Вместе с тем применение дилатации с тем же структурным элементом позволяет почти полностью восстановить площадь объекта. Таким образом, меняя размер структурного элемента, мы удаляем от объекта части той или иной толщины, не меняя площади основной части объекта (Рис. 4).

Последним и самым важным определением, которое необходимо ввести, является морфологический спектр. Впервые это понятие было введено П. Марагосом [1].

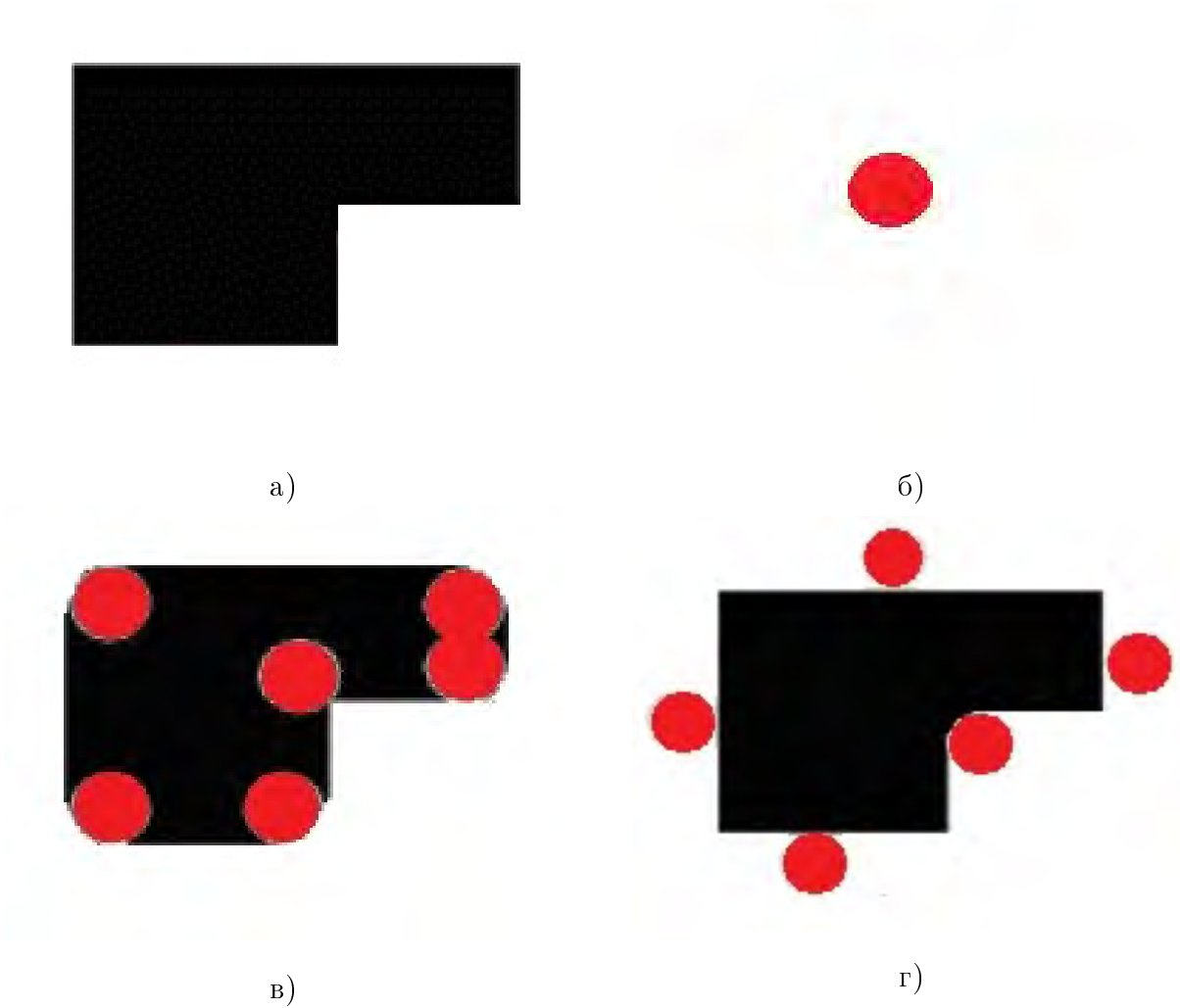


Рис. 4: Пример применения операций открытия (в) и закрытия (г) к фигуре (а) со структурным элементом круг (б).

Марагосом, по аналогии с преобразованием Фурье, вводится представление фигуры изображения в виде форморазмерного спектра, который вычисляется с помощью морфологических операций.

Итак, введем в соответствии с [1] несколько интуитивно понятных, но при этом необходимых определений.

Пусть на евклидовой плоскости задано выпуклое множество  $B$  с центром масс в начале координат. Будем считать, что мощность  $B$  считается единичной, тогда для  $r \geq 0$  определим множество

$$rB = \{rb | b \in B\}$$

**Определение 1.12.** Говорят, что множество (фигура)  $rB$  (или  $B(r)$ ) имеет относительно  $B$  размер  $r$ .

Очевидно, что форма  $rB$  повторяет форму  $B$  и фактически является ее увеличением в  $r$  раз.

Определим морфологический спектр (patter spectrum). Итак, в соответствии с [1]:

**Определение 1.13.** Морфологическим спектром  $PS_X$  фигуры  $X$  относительно примитива (структурного элемента)  $B$  называется функция

$$PS_X(r, B) = -\frac{dS(X \circ rB)}{dr},$$

где  $r \geq 0$ ,  $S(X) = \|X\|$  - площадь  $X$

Поскольку использовать непрерывные функции над объектами растрового изображения неудобно, введем определение морфологического спектра для дискретного случая:

**Определение 1.14.** (Морфологический спектр для дискретного случая) Морфологическим спектром  $PS_X$  фигуры  $X$  относительно примитива (структурного элемента)  $B$  называется функция

$$PS_X(r_i, B) = -\frac{S(X \circ B(r_i)) - S(X \circ B(r_{i+1}))}{r_i - r_{i+1}},$$

где  $(r_i - r_{i+1})$  - шаг растровой решетки,  $r_i \geq 0$ .

Итак, морфологический спектр - это одно из описаний формы фигуры. Принцип построения спектра заключается в последовательном применении операции открытия с увеличивающимся радиусом структурного элемента (в нашем случае радиуса круга). На каждом шаге мы удаляем все части объекта, имеющие толщину меньше двух радиусов структурного элемента, при этом оставляя площадь остальной части объекта без изменений. Таким образом, фиксируя на каждом шаге количество пикселей (площадь) объекта, которое изменилось после применения морфологического открытия, мы строим морфологический спектр. Другими словами, морфологический спектр — это распределение площади объекта по толщинам, поэтому часто морфологический спектр называют картой толщин. Спектр принято изображать в виде

гистограммы, удалившейся при применении морфологического открытия площади объекта от радиуса структурного элемента. Чаще всего значением столбца является не абсолютная разница площади (накопительная гистограмма спектра), а именно разность в площадях между результатом от применения открытия со структурным элементом радиуса  $r-1$  и радиуса  $r$ . Пример построения морфологического спектра в соответствии с определением 1.14 изображен на Рис. 5

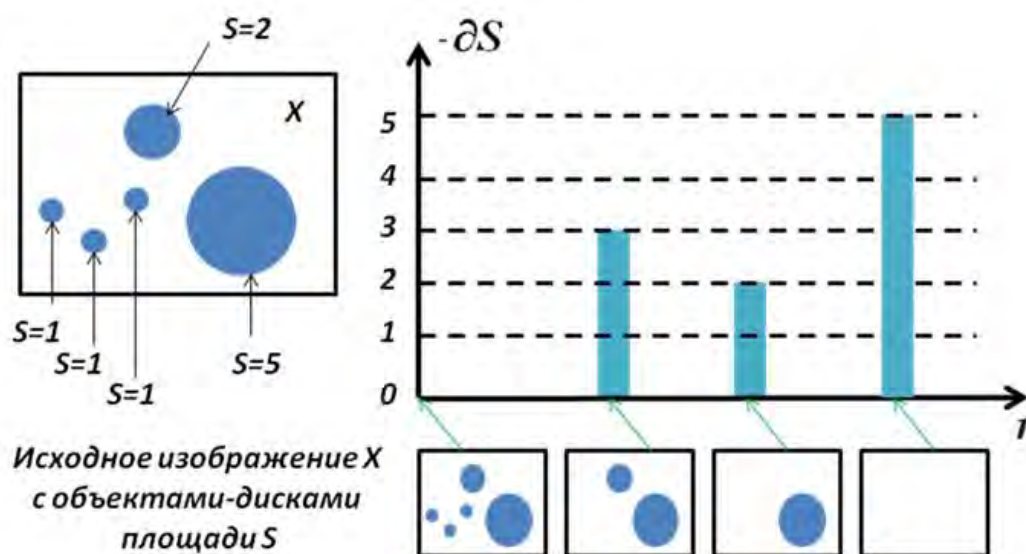


Рис. 5: Пример построения морфологического спектра и его визуализация в виде гистограммы [9].

Еще одним подходом к классификации формы объектов являются срединные оси, или скелеты, которые вводятся с помощью непрерывной морфологии [2]. Скелетное представление основано на представлении фигуры через точки вписанных в неё окружностей и имеет вид плоского геометрического графа. Используя этот граф можно строить различные топологические и метрические признаки формы объекта.

**Определение 1.15.** *Скелетом фигуры называется множество точек-центров всех вписанных кругов фигуры.*

Можно дополнить модель скелета, поставив в соответствие каждой точке скелета радиус вписанной окружности, центром которой она является. Полученная конструкция называется радиальной функцией скелета. Таким образом, появляются новые признаки, с помощью которых мы можем описывать ширину объекта относительно

срединных осей. Совокупность срединных осей и радиальной функции называется медиальным представлением фигуры.

**Определение 1.16.** *Радиальная функция определена в точках скелета и ставит в соответствие точке скелета радиус вписанного круга с центром в этой точке.*

Очевидно, что медиальное представление в целом аналогично морфологическому спектру, который строится с использованием структурного элемента в виде диска.

В работе [10], которая на момент написания диплома находилась в печати, предлагается расширить понятие радиальной функции на все точки фигуры. Для этого вводится понятие медиальной ширины фигуры в точке.

**Определение 1.17.** *Спицей называется отрезок прямой, соединяющий точку скелета с ближайшей точкой границы фигуры.*

Легко доказать [10], что для каждой точки скелета найдется хотя бы одна спица, которая проходит через эту точку.

**Определение 1.18.** *Медиальная ширина  $\varphi(x)$  фигуры  $X$  в точке  $x$  равна длине минимальной спицы, инцидентной этой точке.*

Из определения следует, что в точках скелета медиальная ширина совпадает с радиальной функцией, и при этом медиальная ширина определена во всех точках фигуры. Обозначим множество точек фигуры, медиальная ширина которых меньше заданного целого числа  $w$  через  $G_w = \{x \in X : \varphi(x) < w\}$ .

**Определение 1.19.** *Функция медиальной ширины  $\mathcal{F}(w)$  фигуры  $X$  есть площадь множества точек фигуры, имеющих медиальную ширину  $< w$*

$$\mathcal{F}(w) = |G_w|$$

По аналогии с морфологическим спектром Марাগоса, определим медиальный спектр фигуры [10].

**Определение 1.20.** *Функция  $MWS_X(w) = \frac{d\mathcal{F}(w)}{dw}$  называется медиальным спектром фигуры  $X$ .*

Медиальный спектр, в отличие от морфологического спектра, при построении не требует выполнения сложных вычислительных операций и, соответственно, вычисляется на несколько порядков быстрее. Один из алгоритмов для быстрого вычисления медиального спектра предлагается в статье [10]. Вместе с тем, смысл, который отражают спектры, практически одинаковый: медиальная функция, в соответствии с Опр. 1.19, отражает площадь фигуры, которая удалится при применении морфологического открытия с примитивом в форме диска с радиусом  $r = w$ . Пример медиального спектра изображен на Рис. 6.

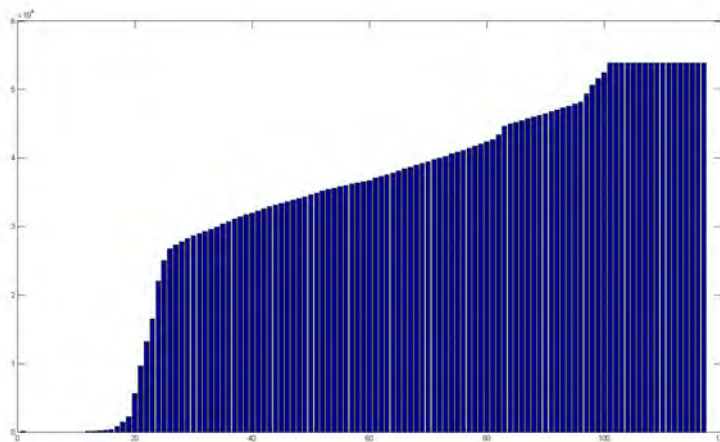


Рис. 6: Пример построенного медиального спектра.

Еще одним способом описать фигуру с помощью дескрипторов, основанных на ширине фигуры, является вычисление карты расстояний (Distance transform (map), далее по тексту - DT) [13]. Карта расстояний (Distance map) — это такое представление изображения фигуры, в котором для каждой точки объекта известно расстояние до ближайшей точки фона. Другими словами, каждой точке объекта ставится в соответствие минимальное расстояние до точек фона. Пример построенной карты расстояний изображен на Рис. 7.

**Определение 1.21.** *Функция построения карты расстояний  $DT$  является отображением фигуры  $X$ , при котором каждой точке фигуры  $x$  ставится в соответствие минимальное расстояние до точек фона, то есть:*

$$DT(x) = \min_p(d(x,p)|x \in X, p \notin X)$$

Карта расстояний является еще одним описанием объекта с точки зрения толщины, причем вычисление ДТ является операцией, выполняемой за линейное время  $O(n)$ , где  $n$  - площадь фигуры. В современных прикладных системах, например Matlab, реализованы эффективные методы построения карты расстояний. Некоторые эффективные алгоритмы приводятся в статьях [13, 14]. ДТ легко перевести в спектральный вид, отложив по горизонтальной оси область значений  $DT(x)$ , а по вертикальной оси - количество точек, имеющих соответствующую глубину. Пример такого спектра изображен на Рис. 8.

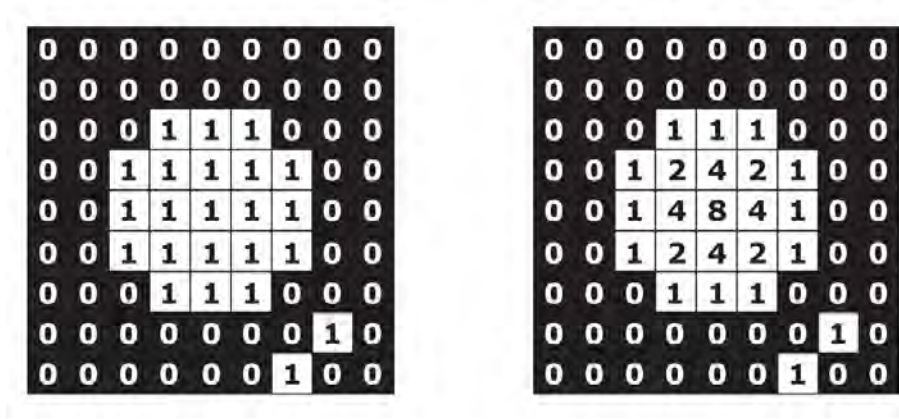


Рис. 7: Пример построения карты расстояний.

Параметром функции  $DT(x)$  является выбор метрики, с помощью которой будет вычисляться расстояние. В статье [14] приводятся основные метрики, например евклидова, Манхэттена ( $L_1$ ) или шахматной доски ( $L_\infty$ ).

В работе с растровой графикой удобнее всего использовать Евклидову метрику, то есть:

$$d(q, p) = \sqrt{(q_x - p_x)^2 + (q_y - p_y)^2}.$$

## 1.2 Известные результаты исследований анализа формы объектов

Отметим те исследования морфологических спектров, которые были проведены на момент написания работы автором. Во-первых, это введение самого понятия спектра Петросом Марагосом [1]. Им было введено понятие морфологического спектра

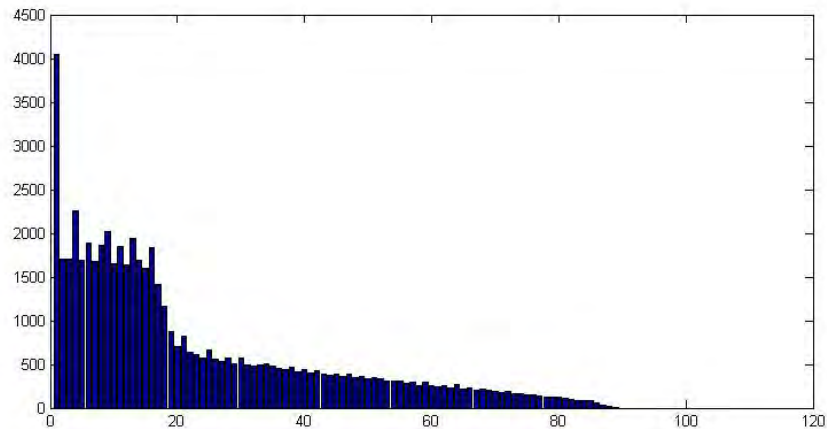


Рис. 8: Пример спектра, построенного на основе ДТ.

и приведены основные свойства такого представления формы объекта. Во-вторых, исследования морфологического спектра, в частности его устойчивости, были приведены в работах Визильтера Ю.В. и Сидякина С.В. в частности [5] и [9]. В данных работах была доказана устойчивость спектра к изменению положения и формы фигуры, а также предложены подходы к сравнению, в частности EMD метрика сравнения гистограмм. В работах справедливо отмечается высокая вычислительная сложность операций по построению спектра, что сильно сужает круг задач, к которым указанный подход применим. В основном это задачи, где не требуется проведение анализа в режиме «реального времени».

Достаточно обширным, по мнению автора, исследованием морфологического спектра является диссертационная работа Сидякина С.В. [9]. В ней рассмотрены известные и новые подходы к вычислению спектра, способы сравнения спектров и применение морфологического спектра к анализу видеопоследовательностей.

Также применение морфологического спектра к задаче анализа формы фигур описано группой ученых из Национального института Астрофизики Мехико, Мексика [7]. В данной работе авторы рассматривают применение морфологического спектра к задаче классификации человека по форме ладони, однако с рядом ограничений: изображения рассматриваются в высоком разрешении, изображения «снимаются» с помощью современного сканера, положение рук и освещенность фиксировано, обучающая выборка для каждого класса состоит из не менее 20 объектов. В работе [7]



получены, по мнению автора, хорошие результаты - при проведении скользящего контроля более 90 % объектов были верно отнесены к соответствующему классу.

В настоящей работе будет рассмотрена аналогичная задача, однако без указанных выше ограничений, с гораздо более реалистичными условиями: изображения снимаются с помощью веб-камеры, освещенность не равномерная, присутствуют шумовые эффекты при снятии объекта, положение фигуры не фиксировано. Более подробно об этом будет сказано во главе 3.

Применение медиального спектра является новым подходом к анализу формы объектов, предложенный Л.М. Местецким. При сохранении всех полезных свойств морфологического спектра Марагоса, медиальный спектр обладает существенно более низкой вычислительной сложностью. На момент написания дипломной работы статья [10] находилась в стадии печати, однако ее упоминание крайне важно в рамках данного исследования. В статье приводится алгоритм построения медиального спектра на основе скелетного представления объектов. Скелетное представление, в свою очередь, является исследованной областью, имеются эффективные алгоритмы построения скелета фигуры, которые можно изучить в книге [2].

При проведении обзора релевантной литературы не было обнаружено публикаций на тему применения карты расстояний (DT) к анализу формы объектов, имея виду классификацию. В основном, DT используется для приближенной скелетизации объектов, анализа изменения положения объекта на последовательности кадров, поиске различных путей внутри объекта и т.д.

Таким образом, в настоящее время исследованы свойства морфологического спектра, доступны публикации на тему применения подхода на основе морфологического спектра к прикладным задачам. Исследование применимости медиального спектра находится на начальной стадии, однако приведены позитивные результаты применения такого подхода к задачам, аналогичным исследованным для морфологического спектра. Применимость спектрального подхода, основанного на построении карты расстояний (DT), к решению задачи анализа формы объектов и их классификации по мнению автора является неисследованным.

### 1.3 Постановка задачи классификации ладоней с помощью морфологического спектра

Целью данной работы является обоснование применимости к задаче классификации (в частности, к задачам верификации и идентификации) и сравнение морфологического, медиального и основанного на карте расстояний спектров. Стоит отметить, что на данный момент эффективность медиального спектра и спектра на основе вычисления расстояний глубоко не исследована. В ходе вычислительных экспериментов (глава 3) данный вопрос будет рассмотрен на задаче классификации ладоней различных людей, постановка которой приводится в работе [11].

Задача состоит в следующем: имеется выборка фотографий ладоней различных людей (Рис. 9), необходимо построить алгоритм классификации (верификации или идентификации), который при поступлении новой фотографии мог бы определить к какому человеку относится изображенная ладонь или отклонить ее, сообщив, что данная ладонь относится к человеку, отсутствующему в базе.



Рис. 9: Пример изображения ладони.

Формально задача идентификации выглядит следующим образом: имеется множество объектов  $X$  цифровых изображений ладоней, множество классов  $Y = \{1, \dots, M\}$  людей. Необходимо, имея обучающую выборку  $\{x_1, \dots, x_l\} \subset X$ ,  $\{y(x_1), \dots, y(x_l)\} \subset Y$ , построить алгоритм  $a : X \rightarrow Y$ .

Предлагаемый в данной работе метод основывается на предположении, что с точки зрения распределения толщин по кисти человека ладони разных людей будут

существенно отличаться. Подходы, описанные в работе, заключаются в обработке изображения и построении его спектров, которые выступают признаковыми описаниями объекта. С помощью полученных признаков обучается модель классификации.

## 2 Известные и новые подходы к решению задачи анализа формы

Прежде чем описать новые подходы, отметим некоторые особенности морфологических спектров. Во-первых, исследуемые в настоящей работе спектры устойчивы и инварианты к изменению положения объекта на изображении (Рис. 10). То есть спектр описывает, в первую очередь, распределение толщин по объекту, можно сказать «вместимость» в него структурных элементов разного радиуса, и его регулярность в этом смысле крайне полезна при решении задачи анализа форм. Спектр также отображает и другую структурную информацию: например, сумму значений морфологического спектра Марагоса, которая является площадью объекта, а значение первого элемента спектра на основе карты расстояний есть длина границы объекта. Такая информация также может быть использована при формировании признаков для задачи классификации.

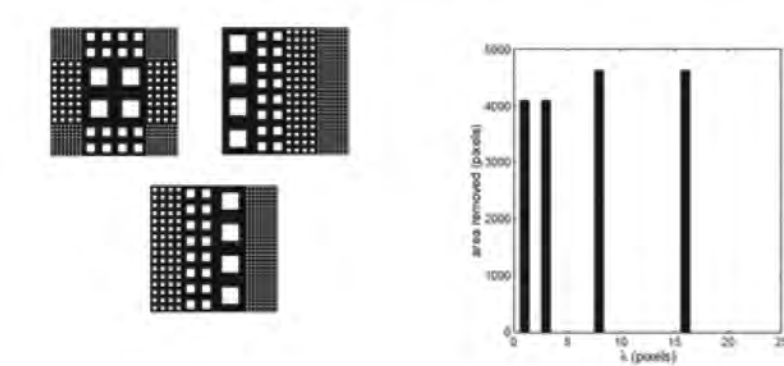


Рис. 10: Различные бинарные изображения, имеющие одинаковый спектр.

Рассмотрим основные полезные свойства морфологического спектра Марагоса, которые также могут быть отнесены и к другим спектрам, рассматриваемым в данной работе.

Наличие больших значений в начале спектра означает, что объект обладает большим количеством мелких неровностей, наличие больших значений в конце спектра означает наличие крупных частей объекта, вмещающих в себя структурный элемент  $rB$  с большим радиусом  $r$ .

В зависимости от выбора формы структурного элемента спектр может показывать «*B*-форменность» объекта, то есть максимальная степень содержания структурного элемента в объекте. На практике чаще всего удобно использовать структурный элемент в виде диска (круга), но существуют задачи, когда структурным элементом может выступать квадрат, овал и т.д.

Описанные выше особенности позволяют говорить об актуальности применения морфологического спектра в целом классе задач, связанных с анализом формы объектов и их классификацией или кластеризацией. Применение спектра позволяет ослабить ограничения на положение ладони в кадре, так как сам спектр инвариантен к положению объекта. В связи с тем, что спектр строится на основе бинарного изображения, качество классификации не будет зависеть от изменения освещенности объекта и т.д. Таким образом, единственными ограничениями в применимости спектра являются различные изменения формы объекта в связи с некорректной съемкой (неправильный ракурс, сжатые пальцы и т.д.) или в связи с физиологическими изменениями (в нашем случае примером может являться утолщение кисти и пальцев в связи с травмой).

Основная идея применимости спектров заключается в построении дескрипторов для каждого из объектов, вычислении расстояния между объектами с помощью метрик и построении алгоритма классификации.

Для того, чтобы настроить алгоритм классификации, необходимо определить меру сравнения спектров, то есть определить метрику. О том, какие конкретно метрики были рассмотрены и оказались оптимальными при решении задачи, поставленной в данной работе, будет описано в главе 3. Так как спектры визуализируются гистограммами, рассмотрим основные подходы к их сравнению.

Первый и самый простой способ сравнить морфологические спектры — это сумма поэлементной разности, то есть *L1*-метрика

$$d_{simple} = \sum_i |H_2(i) - H_1(i)|$$

Метрика проста и легко интерпретируется: чем меньше значение *d*, тем больше сходство.

Следующим способом сравнения гистограмм, похожим на предыдущий, является метрика на основе пересечения, то есть

$$d_{intersection} = \sum_i \min(H_1(i), H_2(i))$$

Для пересечения высокая оценка указывает на хорошее совпадение, а низкая - на плохое. Если обе гистограммы нормированы на 1, то идеальное совпадение - это 1, а совершенное несовпадение - 0.

Ещё один незаменимый инструмент для работы с гистограммами был предложен Карлом Пирсоном [Pearson], который предложил метрику хи-квадрат:

$$d_{\chi^2} = \sum_i \frac{(H_1(i) - H_2(i))^2}{(H_1(i) + H_2(i))}$$

Для хи-квадрата более низкая оценка означает лучшее совпадение, чем более высокая. Идеальное совпадение равно 0, а совершенное несовпадение не ограничено (зависит от размера гистограммы).

Одним из хороших способов сравнения гистограмм, но при этом не сильно распространённым, является расстояние Бхаттачария:

$$d_{Bhattacharyya} = \sqrt{1 - \frac{\sum_i \frac{\sqrt{H_1(i) \cdot H_2(i)}}{\sqrt{\sum_i H_1(i) \cdot \sum_i H_2(i)}}}{1}}$$

Низкие показатели свидетельствуют о хорошем совпадении, а высокие о плохом. Идеальное совпадение - это 0, а совершенное несовпадение - 1.

В работах [9] и [5] описывается высокая эффективность «транспортных» метрик, например EMD-метрики (Earth mover's distance). Основная идея данной метрики заключается в оценке стоимости преобразования гистограммы  $H_1$  в  $H_2$ . В общем случае метрика задается следующим образом: пусть есть некоторая базовая метрика  $d_E$ , а также для каждого  $i$ -ого элемента  $j$ -ой гистограммы задан его «вес»  $h_{ij}$ , при этом вариантов распределения весов несколько (например  $m$  штук). Тогда:

$$d_{EMD} = \min_{h_{ij}} \sum_{i=1, \dots, n} \sum_{j=1, \dots, m} h_{ij} \cdot d_E(H_1(i), H_2(j))$$

В нашем случае, наиболее важным является не конкретный элемент  $i$  гистограммы (то есть отвечающий примитиву радиуса  $i$ ), а те элементы, в которых происходит изменение гистограммы. Поэтому логичным выглядит установить все веса равными 1.

Тогда если мы работаем с накопительными гистограммами (Рис. 15), то евклидова метрика является частным случаем EMD-метрики, то есть:

$$d_{EMD} = \min_{h_{ij}} \sum_{i=1, \dots, n} \sum_{j=1, \dots, m} h_{ij} \cdot d_E(H_1(i), H_2(j)) = \sum_{i=1, \dots, n} \sqrt{(H_1(i) - H_2(i))^2}$$

Сравнение метрик на модельном примере приведено на Рис. 11. В таблице для каждой из гистограмм в правом столбце указано значение метрики при сравнении с исходной гистограммой(моделью).





Model:					
	Simple	Пересечение	$\chi^2$	Бхаттачария	EMD
	<b>0.0</b>	<b>1.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<b>1.0</b>	<b>0.5</b>	<b>0.67</b>	<b>0.55</b>	<b>0.5</b>
	<b>2.0</b>	<b>0.0</b>	<b>2.0</b>	<b>1.0</b>	<b>1.0</b>

Рис. 11: Сопоставление метрик сравнения гистограмм по выдаваемым ответам на модельные данные.

Соответственно, в зависимости от вида решаемой задачи, исследователю необходимо будет сделать выбор одного или нескольких методов сравнения гистограмм.

Рассмотрев обоснование применимости спектрального подхода и основные способы сравнения гистограмм объектов, приступим к рассмотрению задачи, поставленной в разделе 1.3, разделив её на несколько подзадач:

- бинаризация изображения;
- улучшение бинаризованного изображения;
- сегментация ладони, удаление запястья и шумов;

- построение морфологического(медиального, на основе карты расстояний) спектра;
- настройка алгоритма классификации.

Рассмотрим алгоритм решения каждой из подзадач. Для решения задачи бинаризации изображения использовались стандартные методы пороговой бинаризации. Подробно останавливаться на этом в рамках данной работы не будем и отметим, что достаточное множество методов описано в книге Местецкого Л.М. [2], а также в фундаментальном труде Гонсалеса Р. и Вудса Р. [4]. Кроме того, большинство методов эффективно реализовано в различных прикладных средах, например, в пакетах Matlab.

Следующим этапом решения нашей задачи является улучшение изображения после бинаризации. Из определения морфологического спектра следует, что он чувствителен к изменению формы ладони, в том числе и к шумовым эффектам. Наличие шумов можно определить по спектру, например, в случае морфологического спектра это возможно определить по наличию в левой части гистограммы столбцов, значение которых не близко к нулю (то есть в области спектра, отвечающей малым значениям радиуса структурного элемента) (Рис. 12).

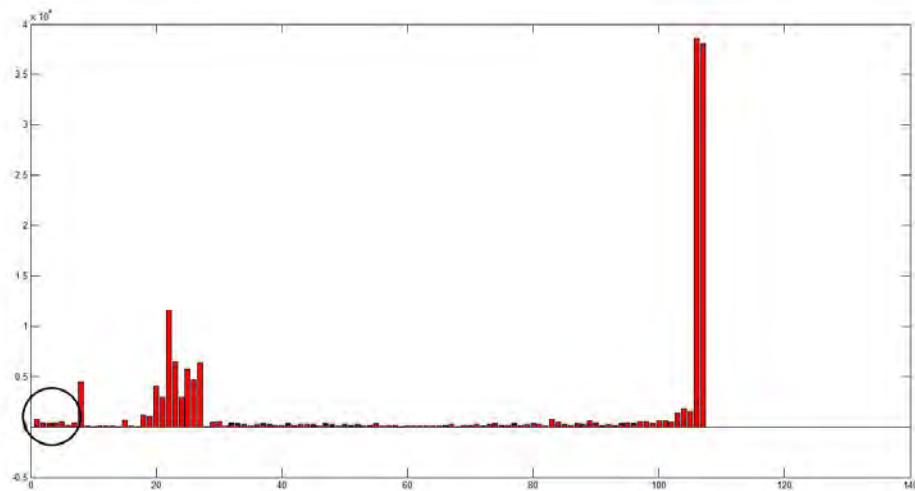


Рис. 12: Пример влияния шумовых элементов изображения на морфологический спектр.

На практике такие эффекты легко убираются различными фильтрами. В различной литературе, посвященной обработке изображений, предлагается применение



медианного фильтра и морфологического открытия со структурным элементом малого радиуса, что позволяет добиться минимизации шумовых эффектов. Однако на выборке, рассматриваемой в данной работе, этих мер не всегда достаточно (пример сильно зашумленного изображения на Рис. 13 а). Применение фильтров позволяет избавиться от мелких неровностей, чего явно не достаточно на Рис. 13 а), поэтому в данной работе предлагается избавляться от шумовых эффектов также с помощью поиска связных компонент. Очевидно, что наибольшая по площади компонента - это интересующая нас область, объект, а все остальные компоненты можно отбросить.

Следующим этапом решения задачи является сегментация ладони и удаление запястья. Необходимость удаления ненужных частей руки обусловлена влиянием лишних элементов на площадь объекта, и, как следствие, на значения спектра. В отличие от выборки, рассматриваемой в статье [7], которая получена с помощью съемки ладони на сканер, наша выборка была получена с помощью веб-камеры, и поэтому практически на всех изображениях, кроме ладоней, присутствует запястье человека. На Рис. 13 а) видно, что запястье может иметь значительные размеры, а значит, оказывать существенное влияние на вид морфологического спектра. Такая ситуация недопустима в рамках нашей задачи, поэтому ставится задача построения алгоритма, удаляющего запястье.

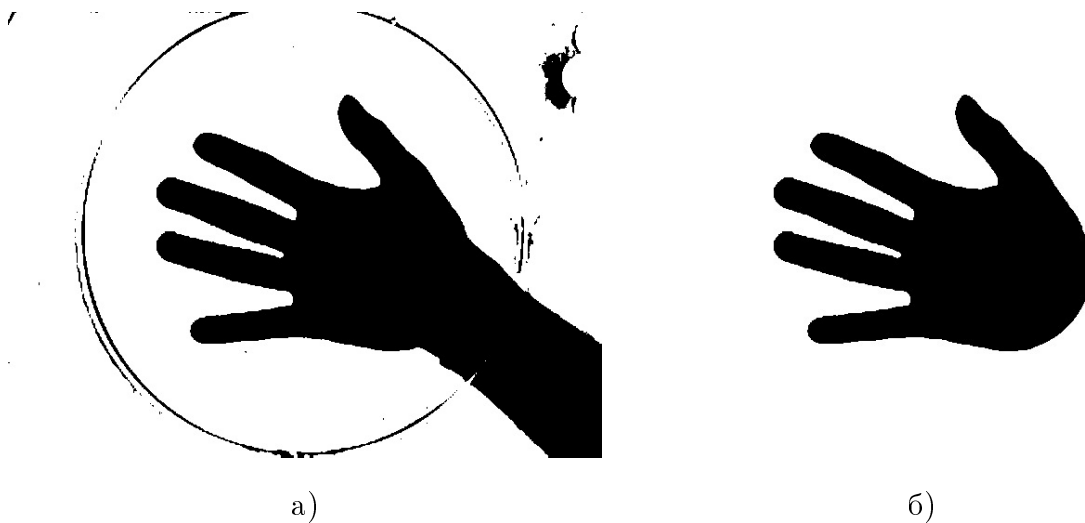


Рис. 13: Пример зашумленного изображения до и после обработки.

Для решения указанной подзадачи предлагается отдельный алгоритм, разработанный автором, который работает достаточно эффективно на представленной выборке. Смысл алгоритма заключается в следующем: необходимо сегментировать

пять (часть ладони между пальцами и запястьем), удалить ее из объекта и произвести заливку запястья пикселями фона, после чего вернуть пять в изображение. Выделение пясти, в нашем случае, произведем на основе нахождения центра самой крупной вписанной окружности. Найти центр этой окружности можно, например, проводя итерационное уточнение ладони, пока от объекта не останется одна точка, которая и будет искомым центром. После нахождения самой большой вписанной окружности, которую мы назовем контрольной областью, необходимо её удалить из объекта. После удаления контрольной области объект распадается минимум на 3 связанные компоненты (Рис. 14), одной из которых является запястье, которое нам необходимо удалить (или, в терминах растровой графики, залить значениями фона). После заливки запястья контрольная область возвращается.

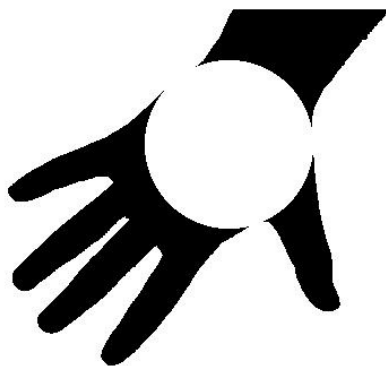


Рис. 14: Изображение ладони после удаления контрольной области.

Итогом проведения подготовительных этапов к построению спектра для Рис. 13 а является изображение Рис. 13 б. Можно наблюдать существенные улучшения в сегментации объекта на изображении.

Теперь опишем один из самых важных этапов – построение по изображению ладони спектра. Алгоритм построения морфологического спектра Марагоса полностью соответствует определению 1.14 для морфологического спектра для дискретного слу-

чая. То есть, последовательно увеличивая радиус структурного элемента, мы проводим операцию открытия и рассчитываем изменение площади.

В качестве структурного элемента оптимальным является круг, так как форма ладони достаточно гладкая и имеет много округлостей.

Как отмечалось выше, морфологический спектр может быть представлен в 2 видах: разностном и накопительном (пример обоих видов морфологического спектра приведен на Рис. 15). В зависимости от условий задачи удобно представлять спектр в том или другом виде. В данной работе везде и далее мы будем использовать накопительный вид спектра. Такой подход удобен для вычисления расстояния между объектами с помощью EMD-метрики.

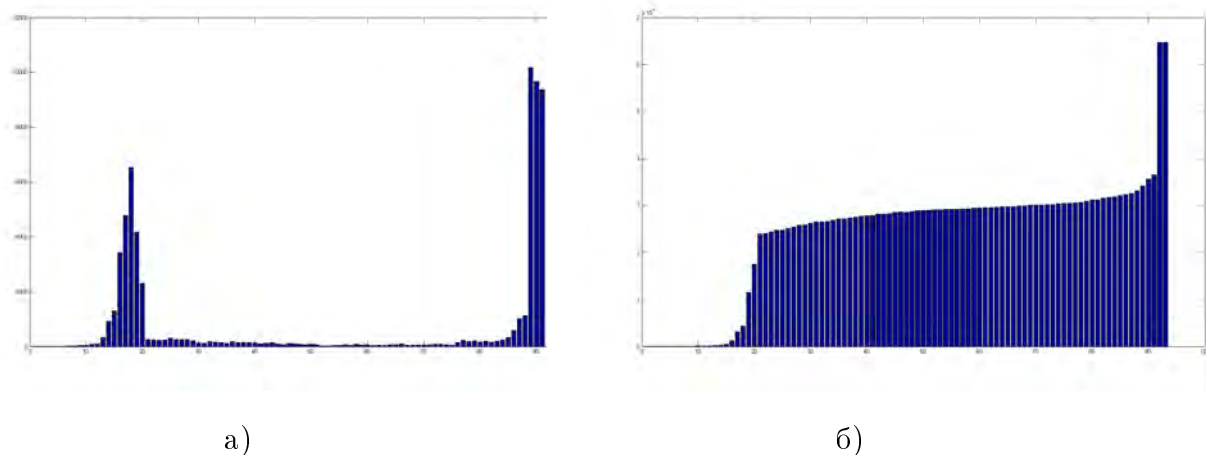


Рис. 15: Морфологический спектр в 2 представлениях: а) разностный вид; б) накопительный вид.

Алгоритм построения медиального спектра заключается в нахождении для каждой точки растрового изображения значения ее медиальной функции. Эффективная реализация алгоритма представлена в статье [10]. Основная идея предлагаемого алгоритма заключается в аппроксимировании исходной фигуры многоугольниками и циркулярами, для которых известны быстрые алгоритмы построения диаграммы Вороного, с помощью которой строится скелет и радиальная функции. Имея скелет и соответствующую ему радиальную функцию построение медиального спектра происходит понятным способом в соответствии с введенными определениям.

Для вычисления спектра на основе карты расстояний необходимо с помощью функции расстояния определить для каждой точки ее минимальное расстояние до точек фона. Алгоритмы нахождения значений данной функции могут быть разнооб-

разными и зависеть, как уже говорилось выше, от выбранной метрики и связности (2-, 4-, 6- или 8-связность). Наиболее простой алгоритм заключается в итерационном «прохождении в глубину» фигуры с маркировкой на каждом шаге слоя соответствующей глубины.

Алгоритмы построения карты расстояний также описаны в статьях [13, 14]. После построения карты расстояний она преобразовывается в спектр с помощью простого подсчета количества точек карты, значение которых равны соответствующему делению горизонтальной оси спектра.

### 3 Вычислительные эксперименты

Логичным является вопрос о том, почему при всех достоинствах спектрального представления со времени первой публикации на данную тему [1] в 1989 году морфологический спектр не нашел широкого применения. Это, в первую очередь, связано с высокой вычислительной сложностью морфологических операций. Фактически вычисление спектра Марагоса состоит из многократного морфологического открытия с примитивами разного радиуса, а операция открытия практически не может быть распараллелина (то есть является последовательной операцией). При этом время выполнения открытия зависит от размера изображения. В результате проведенных автором экспериментов (таблица 1) можно сделать вывод, что в среднем время выполнения операции открытия растет на 10 % при росте размера изображения на 100 px по каждому из измерений. При том факте, что в среднем для построения спектра на изображении 400x400 px необходимо сделать 80 операций открытия, рост по времени выполнения может оказаться существенным.

Размер изображения (px)	Усредненное время выполнения операции открытия с радиусом примитива 5 px (сек)
200x200	0.03044
300x300	0.03154
400x400	0.03772
500x500	0.04118

Таблица 1: Данные экспериментального сравнения времени выполнения операции открытия с круглым примитивом размером 5 px.

В рассматриваемой работе вычислительные эксперименты проводятся на выборке из изображений размером 640x480 px. Среднее время построения спектра для такого изображения на компьютере с процессором Intel Core i5 2.67 GHz составляет 2.43 секунды, что является существенными временными затратами. С этой точки зрения говорить об использовании морфологического спектра при решении задач анализа формы объекта в режиме реального времени говорить не приходится.

Вместе с тем, с помощью пакета среды Matlab Parallel Toolbox можно провести расчеты операции открытия на графическом процессоре (GPU Computing). Операция открытия с растровыми изображениями и примитивом фактически происходит с помощью умножения матриц, и с этой точки зрения графический процессор подходит для выполнения данной операции. Результаты эксперимента по применению алгоритма построения морфологического спектра с помощью графической карты Nvidia GTX 650 на изображениями разных размеров приведены на таблице 2.

Размер изображения (px)	Время выполнения операции на CPU (сек)	Время выполнения операции на GPU (сек)
200x200	2.22314	1.61922
300x300	2.24416	1.62874
400x400	2.2908	1.65776
500x500	2.38306	1.67442

Таблица 2: Данные экспериментального сравнения времени построения морфологического спектра на основном и графическом процессорах.

В работе [9] рассматриваются алгоритмы, с помощью которых можно сократить время расчета операции открытия. В частности, автором исследования [9] выделяется алгоритм, основанный на использовании циркулярного представления объекта с помощью непрерывной морфологии Л.М. Местецкого, описанной в книге [2]. Однако рассмотрение данного подхода останется за пределами данной работы.

С другой стороны, медиальный спектр вычисляется за гораздо меньшее время. В статье [10] приводятся результаты временных замеров для медиального спектра и получившиеся результаты демонстрируют уменьшение вычислительной сложности в 100 раз.

Для сравнения вычислительной сложности были проведены эксперименты по вычислению спектров на 100 случайно выбранных изображениях (таблица 3). Стоит отметить, что морфологический спектр существенно отстает от других видов спектров по времени расчета. Разница между медиальным спектром и спектром на основе карты расстояний в 4 раза связана, по мнению автора, с относительной сложностью

операции построения карты глубины, которая сравнима по времени с алгоритмом обхода в глубину графа. Вместе с тем, думается что для объектов более сложной формы, чем ладонь человека, расчет медиальной функции может происходить дольше.

Морфологический спектр (сек)	Медиальный спектр (сек)	Спектр на основе карты расстояний (сек)
2.458	0.025	0.093

Таблица 3: Усредненное по 100 вычислениям на CPU время построения спектров на изображениях выборки(640x480 px).

### 3.1 Исходные данные и условия эксперимента

Для проверки применимости спектра для решения задач анализа формы объектов и классификации была собрана выборка из 160 фотографий ладоней 36 различных людей. Выборка была получена в рамках проведения Фестиваля науки — 2013 с помощью веб-камеры Logitech HD Pro Webcam C920. Выборка состоит из изображений с разрешением 640x480 px. В выборке присутствуют изображения ладоней людей преимущественно в возрасте от 16 до 22 лет с равным гендерным соотношением. Изображения снимались в цвете и автоматически переводились в бинарное изображение. Пример некоторых из объектов выборки приведен на Рис. 16. Как видно из примеров, положение ладони на изображении и другие ограничения не фиксированы.

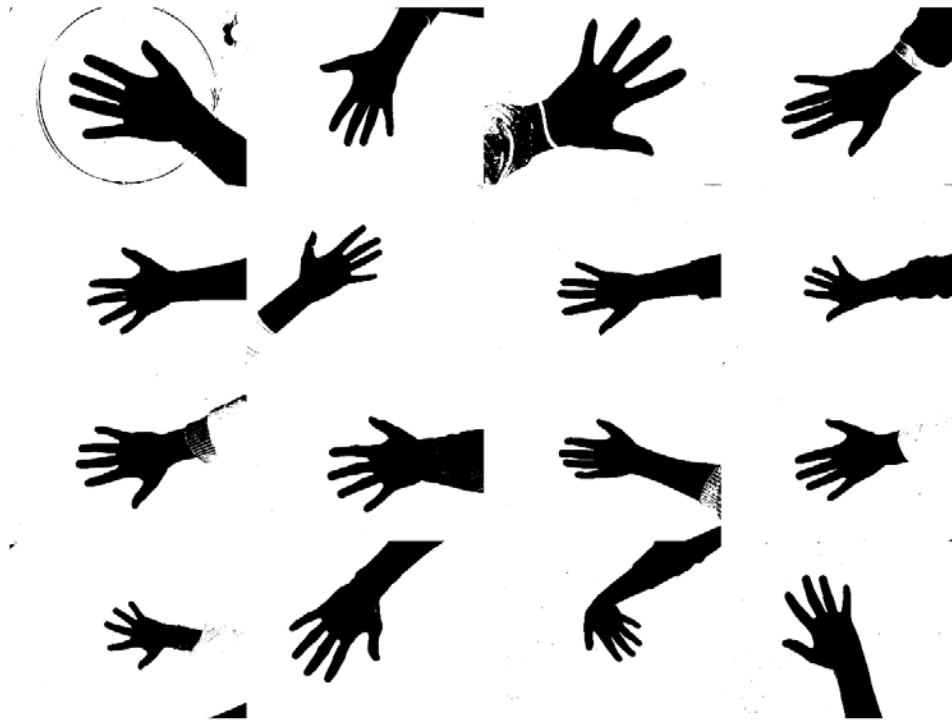


Рис. 16: Элементы выборки.

Напомним, что в соответствии с предложенным в разделе 2 разделением задачи на подзадачи, первые 4 этапа состоят из преобразования изображения в морфологический спектр. Предложенный алгоритм удаления шумов и запястья был применен на выборке и показал свою эффективность: 156 из 160 изображений (97.5 %) были автоматически подготовлены к вычислению на них спектра. 4 изображения из 160 были обработаны в ручную. Примеры успешной и неуспешной обработки изображения и вычисления спектра приведены на Рис. 17, 18 соответственно. Неверная обработка изображения на Рис. 18 заключается в некачественном удалении шумов и, как следствие, неверном удалении кисти руки (на приведенном примере алгоритм вообще не смог ее удалить).



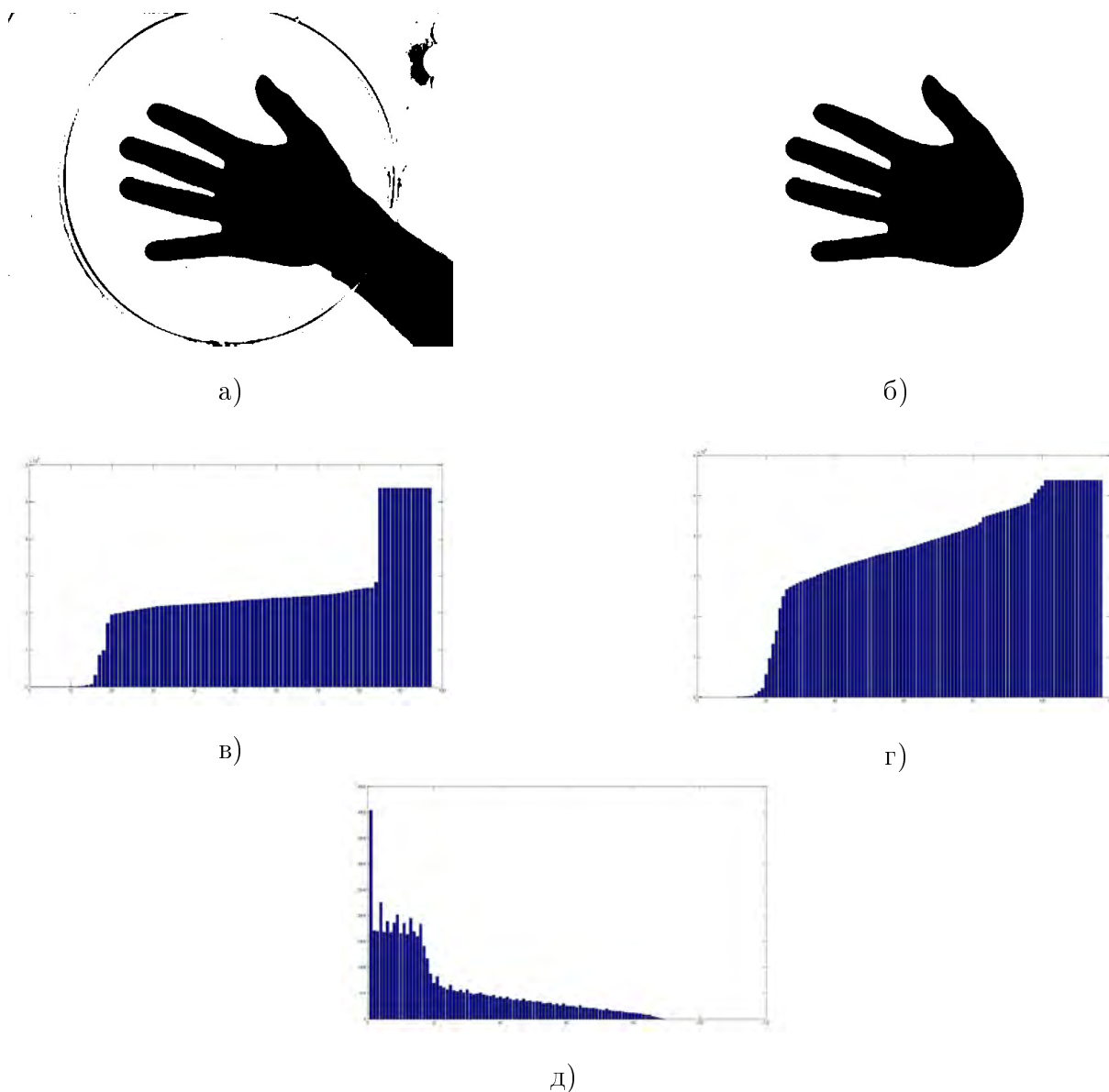


Рис. 17: Пример элемента выборки, который верно предобработан алгоритмом: а)исходный вид; б)обработанный вид; в) морфологический спектр; г) медиальный спектр; д) спектр на основе карты расстояний.

### 3.2 Результаты эксперимента

В соответствии с статьей [7] метрический алгоритм в целом успешно решает задачу классификации. Поэтому необходимо провести сравнение метрик, приведенных в разделе 2, для выбора оптимальной функции расстояния между объектами. Используя простой метрический алгоритм классификации, который определяет класс

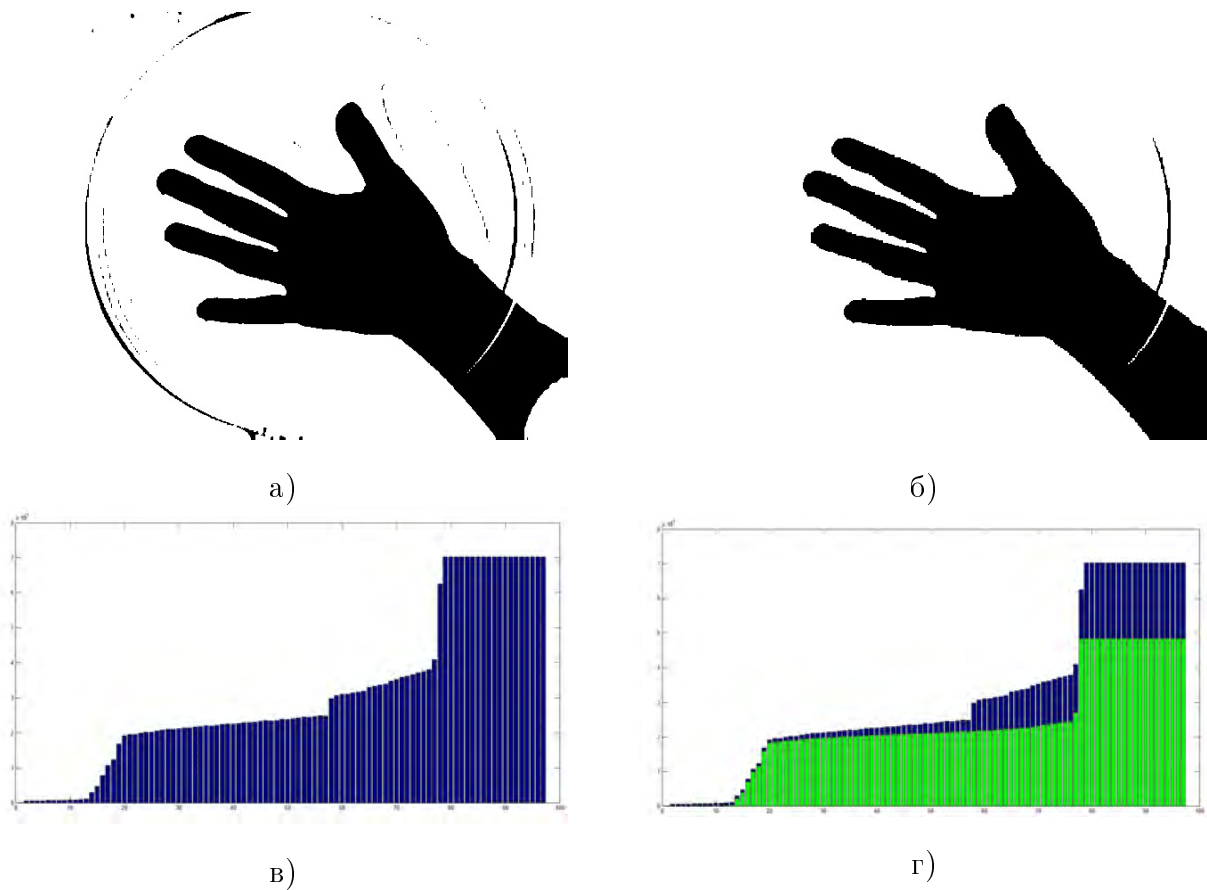


Рис. 18: Пример элемента выборки, который неверно обрабатывается алгоритмом: а) исходный вид; б) обработанный вид; в) морфологический спектр; г) сравнение морфологических спектров объекта, обработанного автоматически и некорректно (синим цветом), и после ручного удаления шумов, то есть корректная обработка (зеленым цветом).

объекта по ближайшему объекту обучающей выборки, мы сможем определить оптимальную метрику.

Для оценки качества воспользуемся скользящим контролем по схеме *leave-one-out*, которая заключается в следующем: фиксируется некоторое множество разбиений исходной выборки на две подвыборки: обучающую и контрольную, которая в нашем случае состоит из одного объекта. Для каждого разбиения выполняется настройка алгоритма по обучающей подвыборке, затем оценивается его средняя ошибка на объекте контрольной подвыборки. Оценкой скользящего контроля называется средняя по всем разбиениям величина ошибки на контрольных подвыборках. В качестве объектов будем использовать морфологические спектры, результаты сравнения для других спектров аналогичны.

Отметим сразу, что в рассмотрении не участвовала метрика на основе пересечения, так как очевидно, что результаты ее применения не будут адекватны. Результаты сравнения остальных методов расчета расстояний между объектами приведены в таблице 4.

Название метрики	Качество
Simple	0.8313
$\chi^2$	0.8663
Бхаттачария	0.7375
EMD	0.9012

Таблица 4: Сравнение метрик сравнения морфологических спектров на выборке по качеству классификации.

Таким образом, экспериментально выявлено, что для решения поставленной в данной работе задачи имеет смысл выбрать EMD-метрику, так как она доставляет наилучшее качество классификации.

### Эксперименты с верификацией

Задача верификации состоит в том, чтобы при предъявлении обученному алгоритму объектов он определял, принадлежит ли он классам объектов из обучения или нет. Часто задачу верификации называют неформально задачей определения «свой-чужой». Основой решения задачи верификации является подбор порогового значения, при достижении которого объект отклоняется, то есть классифицируется как «чужой». Чаще всего верификация отвечает на вопрос: является ли предъявляемый объект представителем класса, за который он себя пытается выдать.

Для оценки применимости спектра к решению задачи верификации ладоней мы случайным образом разделим классы из имеющейся выборки на обучение и контроль. После этого из новой выборки обучения исключим случайным образом выбранные объекты разных классов и также добавим их в контроль. Таким образом, получившаяся контрольная выборка будет содержать как объекты, которые должны быть определены как «свои», так и «чужие» объекты, то есть не принадлежащие обучающей выборке.

В нашем случае поделим выборку на 120 объектов обучения и 40 объектов контроля, из которых половина будет «чужих». Используя EMD-метрику и перебирая пороговое значение от 1 000 до 50 000 с шагом 250, построим ROC-кривую (ROC — *англ.* receiver operating characteristic). Одним из представлений ROC-кривой является график зависимости коэффициента ложного приема (FAR) от коэффициента ложного отклонения (FRR). Более подробно об оценках качества верификации, в частности о построении ROC-кривой можно увидеть в [16].

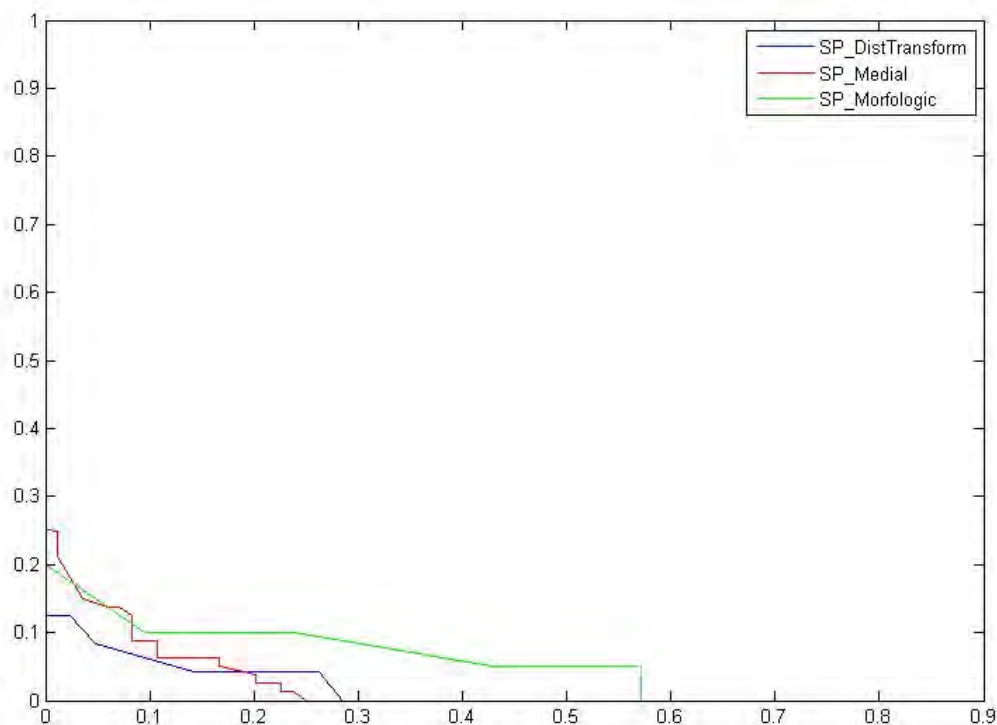


Рис. 19: ROC-кривые, полученные при верификации выборки на основе: морфологического спектра (зеленым цветом); медиального спектра (красным цветом); спектра на основе карты расстояний (синим цветом).

На Рис. 19 изображены ROC-кривые, полученные при верификации с помощью морфологического, медиального и на основе карты расстояний спектров соответственно. Из него следует, что на задаче верификации лучше всего работает модель классификации с спектрами на основе карты расстояний, у которой равный уровень

ошибок (коэффициент ERR [16]) равен  $\approx 8\%$ . Результаты верификации на основе морфологических спектров оказались самыми неэффективными.

### Эксперименты с идентификацией

Задача идентификации, в отличие от задачи верификации, заключается в определении номера класса поступившего на вход алгоритму объекта. В нашем случае количество классов равно 36, в среднем по 4 объекта на класс. Так как основной целью исследования (раздел 1.3) является обоснование эффективности использования спектров к задачам классификации, а также сравнение 3 видов спектров, в рамках данной работы будет рассмотрено использование спектрального подхода вместе с алгоритмом классификации по  $k$ -ближайшим соседям (алгоритм  $kNN$ , далее будем использовать сокращенное наименование). За рамками данной работы оставим рассмотрение других алгоритмов классификации, в работе которых использование спектра может быть обосновано, так как, во-первых, спектральный подход в связке с алгоритмом  $kNN$  позволяет достичь высокого качества работы, что будет показано ниже, а во-вторых, в статье [7] показано, что отсутствует существенная разница между алгоритмами классификации при применении морфологических спектров к анализу классификации форм объектов на изображении.

При использовании алгоритма  $kNN$  основной нашей задачей является настройка алгоритма с помощью варьирования его параметров, а именно количества «соседей» и функции расстояния [8]. Также в силу специфики спектров еще одним параметром будет являться количество элементов спектра, которые будут использоваться в классификации. Обосновать использование такого параметра, к примеру для морфологического спектра, можно с помощью Рис. 20, на котором изображено сопоставление объектов из одного и разных классов с помощью морфологического спектра. Отметим, что на Рис. 20 четко видно, что основная разница наблюдается в правом конце морфологического спектра (то есть в области значений, отвечающих общей площади объекта) и в первой половине спектра, то есть элементам спектра, отвечающим за толщину пальцев и перепады толщины в основных элементах ладони (пять, фаланги и т.д.). Аналогичная ситуация присутствует при сопоставлении медиального спектра и спектра на основе карты расстояний.

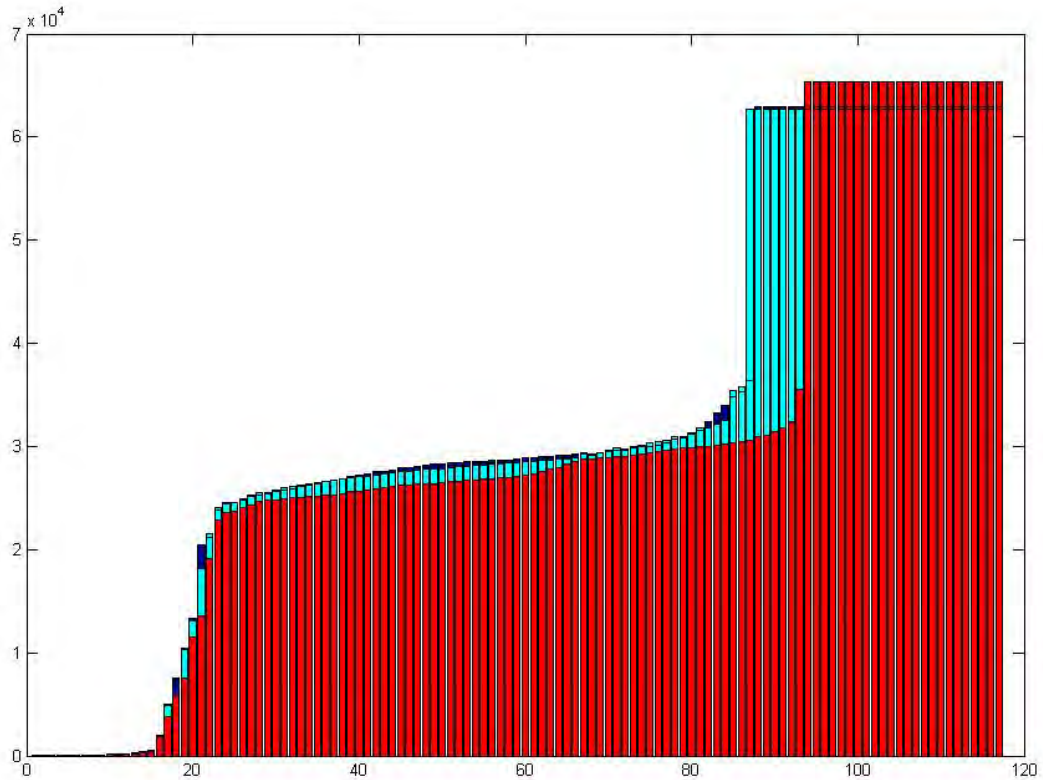
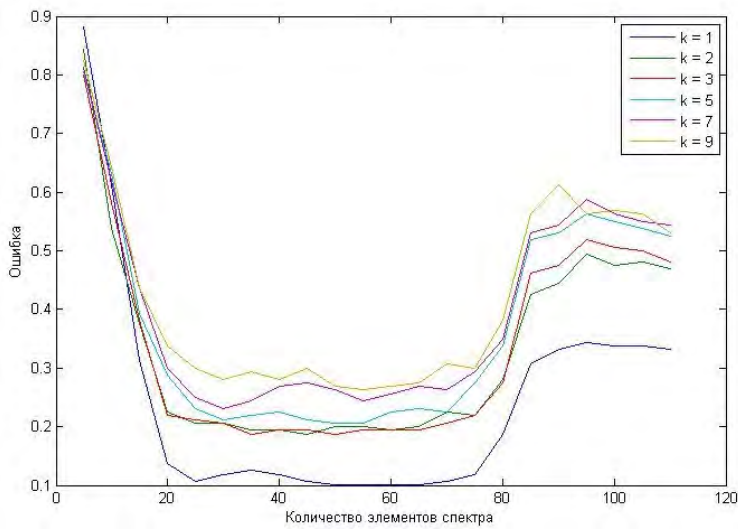


Рис. 20: Сопоставление морфологических спектров объектов разных классов: синим и голубым цветами обозначены объекты одного класса, красным - другого.

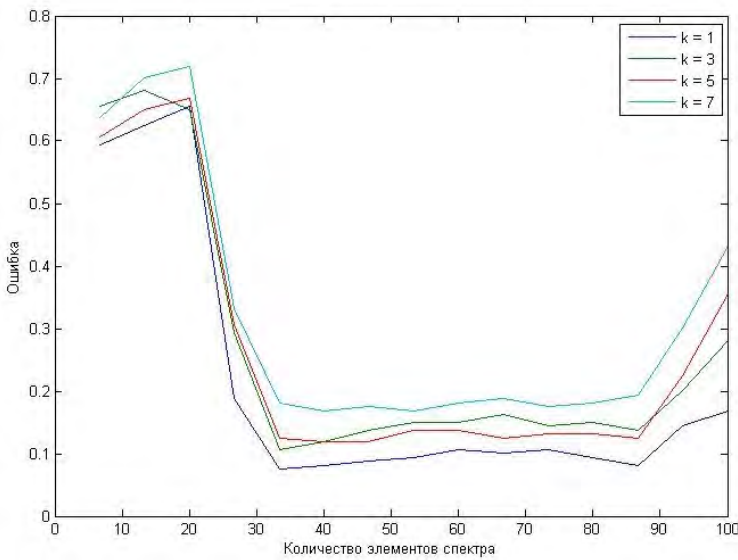
Таким образом, в предлагаемой модели варьируются 3 параметра. Определение оптимальной метрики, то есть функции расстояния, уже описано в начале раздела 3.2.

Для определения количества элементов спектра, входящих в оптимальный набор, и количества  $k$  ближайших соседей была проведена серия экспериментов, которая заключалась в переборе значений параметров и оценке качества классификации с помощью скользящего контроля. Значение качества ( $Q$ ) вычислялось как соотношение числа верно классифицированных объектов ( $N_{true}$ ) от общего числа объектов ( $N$ ):  $Q = \frac{N_{true}}{N}$ . Значение  $1 - Q$  отражает соотношение количества объектов, классификация которых прошла неверно, к общему числу объектов.

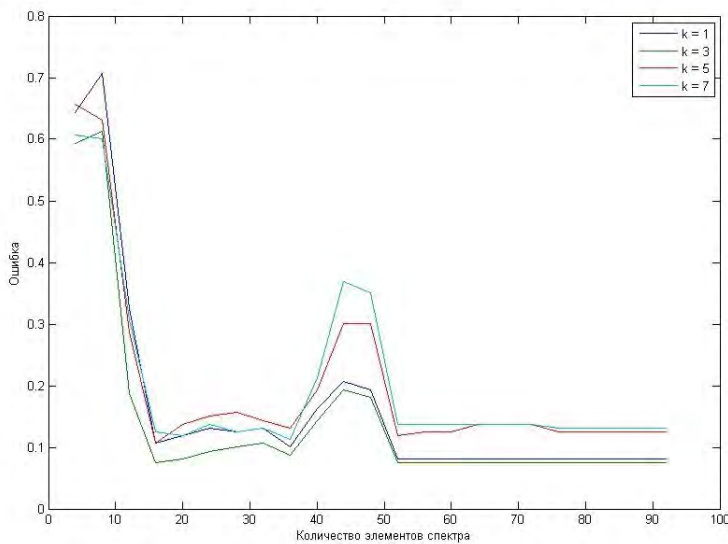
Результаты экспериментов для морфологического, медиального спектров и спектра на основе карты расстояний приведены на Рис. 21.



морфологический спектр



медиальный спектр



спектр на основе карты расстояний

Рис. 21: Значение ошибки классификации алгоритмом kNN в зависимости от выбора параметров: количество «соседей», количество первых элементов морфологического спектра, участвующих в рассмотрении алгоритма.

Из Рис. 21 можно определить оптимальные значения параметров. Во-первых, для всех 3 видов спектра эксперименты показали, что на имеющихся данных наиболее эффективно работает алгоритм, определяющий класс по 1 ближайшему объекту обучающей выборки. Это, в частности, объясняется компактностью объектов внутри класса.

В отличие от количества объектов-«соседей», длина оптимального набора элементов спектра для спектров оказалась разной. Для морфологического спектра, как видно из Рис. 21 а, оптимальным является набор из первых 60 элементов. Также, отдельные эксперименты показали, что включение последнего элемента, отвечающего за общую площадь объекта, позволяет улучшить качество классификации.

Для медиального спектра оптимальный набор состоит из первых 85 элементов и, аналогично морфологическому спектру, последнего элемента спектра. Для спектра на основе карты расстояний оптимальными являются все наборы, состоящие из не менее 54 первых элементов. Такая разница по сравнению со спектрами Марагоса и медиальным заключается в том, что хотя все спектры описывают свойства ширины объекта, интерпретируют они эти свойства по-разному. В частности, морфологический и медиальный спектры описывают в первую очередь «форменность» фигуры относительно структурного элемента (в нашем случае диска), а спектр на основе карты расстояний описывает связь (удаленность) внутренних точек фигуры с границей.

С учетом подобранных параметров, были проведены эксперименты по вычислению качества идентификации с помощью каждого вида спектра. На Рис. 22 представлен график зависимости ошибки классификации с ростом объема выборки для моделей классификации на основе каждого из 3 спектров. График построен на основании проведения 100 экспериментов, в ходе каждого из которых исходная выборка случайным образом делилась на 40 блоков. После чего, в каждом из экспериментов рассматривались результаты классификации для моделей обученных на 1, 2, 3 и т.д. блоках.

Итоговое значение качества классификации для моделей на основе морфологического спектра Марагоса, медиального спектра и спектра на основе карты расстояний составило **93,7 %**, **94,5 %**, **95,2 %** соответственно.



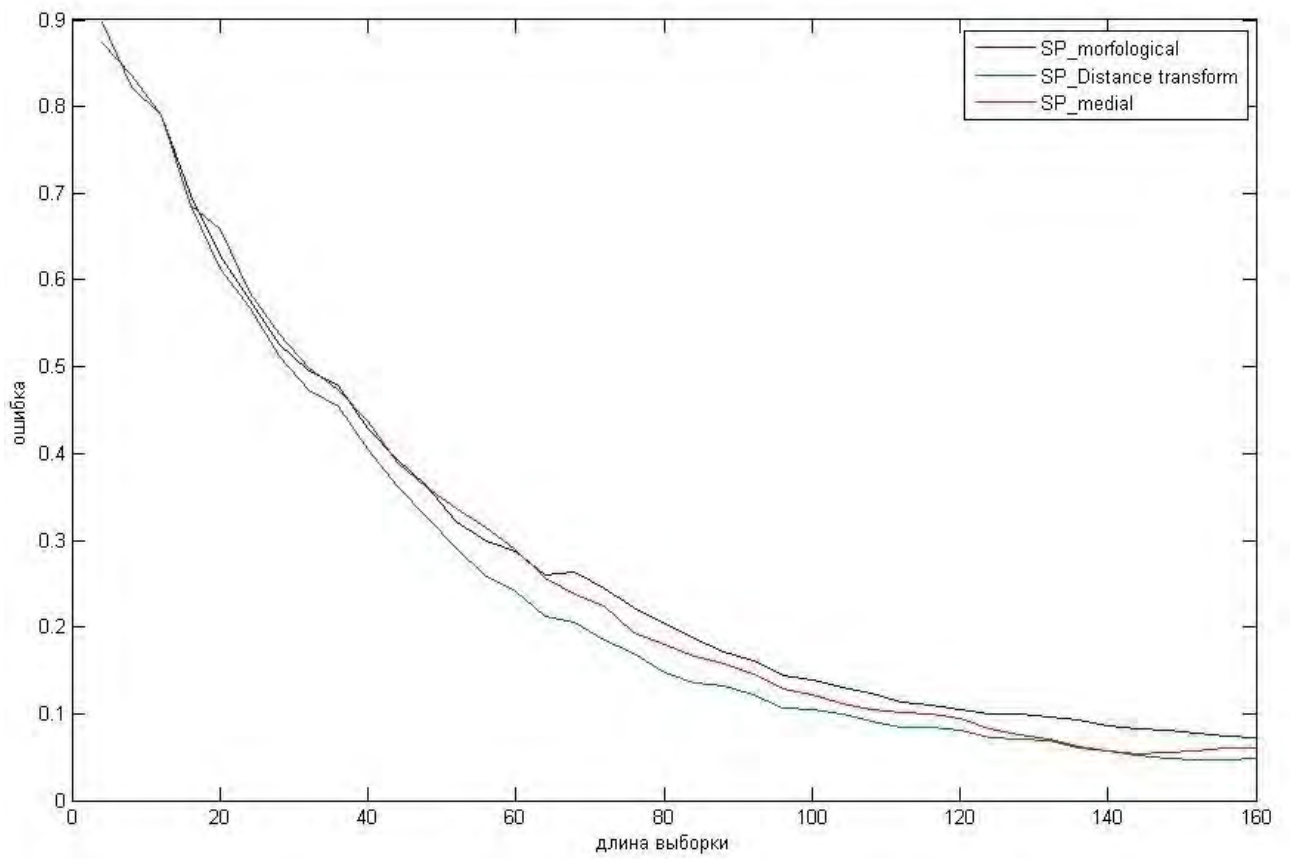


Рис. 22: Сравнение качества классификации в зависимости от длины выборки с использованием морфологического спектра (синий цвет), медиального спектра (красный цвет) и спектра на основе карты расстояний (зеленый цвет) .

## 4 Заключение

В данной работе рассмотрен один из подходов к решению задачи классификации объектов по их форме. Метод, основанный на построении спектров, то есть дескрипторов ширины фигуры, не подразумевает существенных ограничений к объектам выборки и вычислительным ресурсам, при этом он обладает набором полезных свойств, например регулярность, инвариантность относительно положения объекта на изображении и т.д. Подход может свободно применяться для решения задач, аналогичных исследуемой в данной работе, например в задачах классификации животных по их силуэтам, или верификация бракованных деталей на производстве по их изображению.

В работе рассмотрены 3 подхода к построению спектра: морфологический, медиальный и основанный на вычислении расстояний (Distance map). Первый из них достаточно исследован, но при этом отличается высокой сложностью вычисления (так как основная операция – операция вычисления морфологического открытия - имеет сложность порядка  $O(n^2)$ , где  $n$  - количество пикселей на изображениях выборки). Медиальный спектр, в свою очередь, имеет существенно более низкую вычислительную сложность. Использование медиального спектра позволяет решать задачи анализа формы в режиме реального времени, так как время построения спектра составляет  $\approx \frac{1}{40}$  секунды. Для сравнения, стандартная частота кадров в видеопоследовательности равна  $\frac{1}{24}$ .

Спектр на основе карты расстояний также вычисляется за время, существенно меньшее, чем морфологический спектр, но уступает медиальному примерно в 4 раза. При этом он обладает рядом полезных свойств: построение карты расстояний является легко реализуемой операцией, эффективная реализация которой присутствует во всех современных прикладных системах, например в Matlab. Также, экспериментально показано, что спектр на основе карты расстояний показывает наилучший результат верификации и идентификации среди 3 рассматриваемых спектров на имеющейся выборке.

В дипломной работе экспериментально показана эффективность применения описания объектов с помощью спектров, в том числе для решения задачи классификации. Введен неисследованный, по мнению автора, вид спектра фигуры, который

показал более высокие результаты, чем морфологический спектр Марагоса. В этой связи можно предположить, что такой подход применим к решению задач классификации не только человека на основе изображения ладоней, но и объектов другого рода, имеющих выраженные формы, в том числе для анализа последовательностей изображений или видео, где задача верификации до сих пор весьма актуальна. Описанные в данной работе подходы и результаты могут быть использованы в исследовании выявления дескрипторов формы объектов и фигур.

## Список литературы

- [1] Petros Maragos. Pattern Spectrum and Multiscale shape representation // IEEE transaction on pattern analysis and machine intelligence, VOL II, № 7 — 1989.
- [2] Местецкий Л.М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры // Москва, Физматлит — 2009.
- [3] Serra J. Image Analysis and Mathematical Morphology // Academic Press, London — 1982.
- [4] Гонзалес Р., Вудс Р. Цифровая обработка изображений. // М., Техносфера — 2006.
- [5] Визильтер Ю.В., Сидякин С.В. Робастное сравнение морфологических спектров двумерных фигур и изображений, семинар «Морфологический анализ данных» под руководством проф. Ю.П. Пытьева., — 2012.
- [6] Ю.В. Визельтер. Обобщенная проективная морфология // Журнал Компьютерная Оптика, том 32 № 4 — 2008
- [7] J.M. Ramirez-Cortes, P. Gomez-Gil, G. Sanchez-Perez, D. Baez-Lopez, A Feature Extraction Method Based on the Pattern Spectrum for Hand Shape Biometry // Proceedings of the World Congress on Engineering and Computer Science, WCECS 2008, October 22 - 24, San Francisco, USA — 2008
- [8] Журавлев Ю. И., Рязанов В. В., Сенько О. В. «Распознавание». Математические методы. Программная система. Практические применения. — М.: Фазис, 2006
- [9] Сидякин С.В. Разработка алгоритмов построения морфологических спектров для анализа цифровых изображений и видеопоследовательностей // Диссертация на соискание ученой степени кандидата технических наук, Москва, 2013.
- [10] Mestetskiy L.M., Medial width of polygonal and circular figures approach via line segment Voronoi diagram // [в печати].
- [11] Бакина И.Г. Морфологическое сравнение изображений гибких объектов на основе циркулярных моделей при биометрической идентификации личности по

- форме ладони // Диссертация на соискание ученой степени кандидата физико-математических наук, Москва, 2011.
- [12] Домахина Л.Г. Скелетная сегментация и циркулярная морфология многоугольников// Диссертация на соискание ученой степени кандидата физико-математических наук, Москва, 2014.
- [13] A. Meijster, J. B. T. M. Roerdink, W. H. Hesselink, A General Algorithm for Computing Distance Transforms in Linear Time // Mathematical Morphology and its Applications to Image and Signal Processing Computational Imaging and Vision Volume 18 — 2000
- [14] R. Fabbri, L. Da F. Costa, J. C. Torelli, O. M. Bruno, 2D Euclidean Distance Transform Algorithms: A Comparative Survey // ACM Computing Surveys, Vol. 40, No. 1, Article 2 — 2008.
- [15] Словарь русского языка: в 4-х т. Институт русского языка АН СССР. // М.: «Русский язык» — 1981.
- [16] CHARACTERISTICS OF BIOMETRIC SYSTEMS. Cernet // <http://www.ccert.edu.cn/education/cissp/hism/039-041.html>