

Институт проблем передачи информации им. А.А. Харкевича Российской  
академии наук (ИППИ РАН)

На правах рукописи

Ершов Егор Иванович

**Быстрое преобразование Хафа как инструмент анализа  
двумерных и трехмерных изображений в задачах поиска  
прямых и линейной кластеризации**

Специальность 05.13.17 —  
«Теоретические основы информатики»

Диссертация на соискание учёной степени  
кандидата физико-математических наук

Научный руководитель:  
кандидат физико-математических наук  
Николаев Дмитрий Петрович

Москва — 2018

## Оглавление

	Стр.
<b>Введение</b> . . . . .	<b>4</b>
<b>Глава 1. Алгоритмы быстрого преобразования Хафа</b> . . . . .	<b>9</b>
1.1 Обзор существующих алгоритмов преобразования Хафа . . . . .	10
1.1.1 Ранние работы . . . . .	10
1.1.2 Способы параметризации прямых на плоскости . . . . .	16
1.1.3 Дискретизация и её влияние на точность . . . . .	20
1.1.4 Связь преобразования Радона и преобразования Хафа . . . . .	24
1.1.5 Алгоритмы быстрого вычисления ПХ и поиска прямых . . . . .	26
1.1.6 Преобразование Хафа в задачах анализа изображений . . . . .	28
1.2 Диадический паттерн и двумерное преобразование Хафа . . . . .	31
1.2.1 Свойства диадических паттернов . . . . .	32
1.2.2 Алгоритм быстрого преобразования Хафа для двумерного случая . . . . .	39
1.3 Диадическая плоскость и трехмерное быстрое преобразование Хафа для плоскостей . . . . .	41
1.3.1 Свойства диадических плоскостей . . . . .	41
1.3.2 Алгоритм трехмерного быстрого преобразования Хафа для плоскостей . . . . .	44
1.4 Диадическая прямая и трехмерное быстрое преобразование Хафа для прямых . . . . .	47
1.4.1 Свойства диадических прямых . . . . .	47
1.4.2 Алгоритм трехмерного быстрого преобразования Хафа для прямых . . . . .	49
<b>Глава 2. Поиск прямых и плоскостей с помощью БПХ в         двумерном и трехмерном случаях</b> . . . . .	<b>53</b>
2.1 Методы поиска прямых на изображении . . . . .	53
2.2 Связь между вычислением М-оценок и преобразованием Радона . . . . .	55
2.3 Метод поиска прямых и плоскостей с использованием свертки Хаф-образа . . . . .	59

2.4	Метод поиска прямых и плоскостей с использованием свертки исходного изображения . . . . .	62
2.5	Оценка точности метода поиска прямых . . . . .	66
2.6	Экспериментальное исследование предложенных алгоритмов . . .	67
2.6.1	Модель входных данных . . . . .	67
2.6.2	Сравнительное исследование точности определения параметров истинного сигнала с помощью ПХ и БПХ . . .	70
2.6.3	Сравнение методов решения задачи робастной линейной регрессии для двумерных и трехмерных изображений . . .	70
2.6.4	Результаты экспериментов . . . . .	72
<b>Глава 3. Быстрая линейная бинарная кластеризация . . . . .</b>		<b>75</b>
3.1	Метод Оцу для двумерных гистограмм . . . . .	75
3.2	Обобщения критерия Оцу . . . . .	80
3.3	Метод линейной бинарной кластеризации . . . . .	82
3.4	Оценка точности метода линейной бинарной кластеризации . . .	86
<b>Заключение . . . . .</b>		<b>89</b>
<b>Список сокращений и условных обозначений . . . . .</b>		<b>91</b>
<b>Список литературы . . . . .</b>		<b>92</b>
<b>Список рисунков . . . . .</b>		<b>107</b>
<b>Список таблиц . . . . .</b>		<b>109</b>
<b>Приложение А. Доказательство теоремы об оценке точности ортотропной ошибки аппроксимации диадическим паттерном . . . . .</b>		<b>110</b>

## Введение

В последние десятилетия неизменно растёт интерес к технологиям зрительных систем, в том числе компьютерного зрения, анализа и обработки изображений. Сначала главной причиной этого был прогресс вычислительной техники, затем – внедрение робототехнических систем в промышленности, а в последнее время – повсеместное распространение мобильных устройств, оснащенных видеокамерами. На всех этапах развития технологий зрительных систем пополнялся список необходимых для этого базовых инструментов, таких как морфологическая фильтрация, быстрое вычисление свёрток, выделение границ, цветовая адаптация и так далее. Среди зарубежных учёных, внесших существенный вклад в развитие анализа изображений, стоит отметить Р. Дериша, Б. Хорна, К. Шапиро, Г. Финлейсона, а среди отечественных – В. Л. Арлазарова, М. М. Бонгарда, Ю. В. Визильтера, С. Ю. Желтова, Ю. И. Журавлева, Д. С. Лебедева, Б. М. Миллера, В. А. Сойфера и П. А. Чочиа.

Один из таких алгоритмических инструментов – это дискретное преобразование Радона (ПР), именуемое также преобразованием Хафа (ПХ). Каждой прямой, проходящей через область изображения, ПХ ставит в соответствие сумму значений ближайших к этой прямой пикселей. ПХ используется для детекции прямолинейных объектов или их различных конфигураций на изображении, – например, для детекции дорожной разметки, поиска границ документа, цветовой сегментации, вычислительной томографии и прочих. Фундаментальные основы этой области заложены в работах Д. Балларда, Д. Донохо, О. Дуды, Дж. Иллингворта, Х. К. Йена, В. Г. Лабунца, Н. Г. Федотова и В. М. Чернова. О повышенном интересе к созданию схем быстрого вычисления ПХ свидетельствует обзор П. Мукхопадхья и А. Хассенейна 2015 года. Стремительно развиваются также алгоритмы вычисления ПХ для анализа трехмерных изображений (например, в медицине и робототехнике), к трудоёмкости которых также предъявляются высокие требования.

Одним из алгоритмов, вычисляющих преобразование Хафа на двумерном изображении, является быстрое преобразование Хафа (БПХ). В то время как вычислительная сложность стандартного ПХ (СПХ) для плотного равностороннего изображения равна  $\Theta(n^3)$ , где  $n$  – сторона изображения, у БПХ – всего  $\Theta(n^2 \log n)$  при несущественном искажении вычисляемого Хаф-образа (для ал-

проксимации части прямой на изображении используется дискретный паттерн специального вида, так называемый диадический паттерн (ДП)).

Ввиду вышесказанного ясно, что задачи обобщить БПХ на трехмерный случай, создать новые алгоритмы анализа изображений на его основе, а также изучить влияние структуры диадического паттерна на точность этих алгоритмов весьма актуальны.

**Цель** работы – создать алгоритмы быстрого преобразования Хафа для трехмерных изображений, исследовать их свойства, а также разработать на их основе методы анализа двумерных и трехмерных изображений и гистограмм.

Для достижения этой цели поставлены и решены следующие **задачи**:

1. Исследование свойств диадического паттерна как дискретной модели прямой для двумерных и трехмерных изображений.
2. Исследование возможных способов обобщения БПХ на случаи размерности три.
3. Разработка и исследование методов поиска прямых на двумерных, а также прямых и плоскостей на трехмерных изображениях путём вычисления М-оценок в задаче ортогональной линейной регрессии (ОЛР) с помощью БПХ.
4. Разработка и исследование методов быстрой линейной бинарной кластеризации на основе БПХ для двумерных и трехмерных изображений и гистограмм.

**Методология и методы исследования.** В диссертации используются методы цифровой обработки и анализа изображений, математической статистики, вычислительной оптимизации, интегральной геометрии, функционального анализа и численного моделирования.

**Научная новизна:** впервые получено теоретически обоснованное аналитическое выражение (а не рекуррентное, как ранее) для координат диадического паттерна; установлена зависимость оценки его ортогональной ошибки аппроксимации геометрической прямой от размера изображения, а также показано, что система ДП (набор паттернов, суммации по которым составляют БПХ) покрывает все пары пикселей изображения.

Впервые предложены алгоритмы быстрого преобразования Хафа для трёхмерных изображений, позволяющие производить быстрые вычисления сумм по всем дискретным диадическим прямым и плоскостям трёхмерного изображения; для первого асимптотическая сложность не может быть уменьшена.

Впервые предложен и исследован метод приближенного вычисления М-оценок с использованием БПХ в задаче ортогональной линейной регрессии для двумерных и трёхмерных изображений.

Впервые предложен и исследован метод линейной бинарной кластеризации с помощью БПХ для обобщенного метода глобальной бинаризации Оцу.

**Практическая значимость.** Предложенный метод детекции прямых используется в проекте компании ООО «Визиллект Сервис» по созданию беспилотного автобуса для детекции дорожной разметки, а предложенный метод линейной бинарной кластеризации используется в проекте по созданию визуального классификатора проезжающих транспортных средств для фильтрации ложных проездов путем кластеризации на гистограмме срабатываний детектора колёс, что подтверждено соответствующими актами.

### Основные положения, выносимые на защиту:

1. Доказано, что максимальный ортогональный разброс диадического паттерна ограничен снизу величиной  $\frac{\log_2 n}{2\sqrt{10}}$ , а сверху  $-\frac{\log_2 n}{6}$ ; для диадических плоскостей ограничен снизу  $-\frac{\log_2 n}{\sqrt{11}}$ , а сверху  $-\frac{\log_2 n}{3}$  и, наконец, для диадических прямых ограничен снизу величиной  $\frac{\sqrt{2}\log_2 n}{6}$ , а сверху  $-\frac{\log_2 n}{\sqrt{22}}$ , где  $n = 2^k$  – сторона изображения.
2. Доказано, что система диадических паттернов в БПХ обладает свойством *полноты*, то есть для любой пары пикселей на изображении найдется проходящий через них паттерн; аналогичное верно и для диадических прямых ТБПХ.
3. Предложены два алгоритма вычисления трехмерного быстрого преобразования Хафа (ТБПХ) для прямых и для плоскостей. Доказано, что ТБПХ для прямых обладает вычислительной сложностью  $\Theta(n^4)$ , ТБПХ для плоскостей –  $\Theta(n^3 \log n)$ .
4. Предложен метод поиска прямых для двумерных, а также прямых и плоскостей для трехмерных гистограмм путём вычисления М-оценки в задаче ОЛР с помощью алгоритмов БПХ. Для двумерных изображений вычислительная сложность предложенного метода составляет  $\Theta(N + n^2 \log n)$ , для трехмерных прямых –  $\Theta(N + n^4)$ , а для плоскостей –  $\Theta(N + n^3 \log n)$ , где  $N$  – число наблюдений. Получены оценки точности вычисления М-оценок.
5. Предложен метод бинарной кластеризации гиперплоскостью с помощью алгоритмов БПХ для любого критерия, выразимого через меры,

заданные на области определения гистограммы, с вычислительной сложностью  $\Theta(n^2 \log n)$  для двумерных гистограмм и  $\Theta(n^3 \log n)$  – для трехмерных. Приведен способ оценки его точности разделения на основе входной гистограммы и параметров разделения.

**Достоверность** полученных аналитических результатов диссертации обеспечена использованием строгого математического аппарата и методов функционального анализа, математической статистики и теории оптимизации. Помимо этого, достоверность подтверждается результатами тестирования, апробации и внедрения разработанных методов. Полученные результаты согласуются с результатами, полученными другими исследователями.

**Апробация работы.** Основные результаты работы докладывались на международных конференциях: 29th European Conference on Modelling and Simulation (ECMS 2015, Альбена, Болгария), ECMS 2016 (Регенсбург, Германия), ECMS 2017 (Будапешт, Венгрия), 8th International Conference on Machine Vision (ICMV 2015, Барселона, Испания), 4th Professors day in Huawei (2017, Москва, Россия), были дважды доложены на совместном научном семинаре Лабораторий №2 и №11 Института проблем передачи информации имени А.А. Харкевича РАН, а также на семинаре «Анализ и понимание изображений (математические, когнитивные и прикладные проблемы анализа изображений и сигналов)» ВЦ РАН.

**Личный вклад.** Все результаты диссертации, вынесенные на защиту, получены автором самостоятельно. Постановка задач и обсуждение результатов проводилось совместно с научным руководителем, численные сравнения методов поиска прямых осуществлялись под руководством диссертанта младшим научным сотрудником лаборатории 11 ИППИ РАН Е.Н. Асватовым. Работа по теоретическому обоснованию максимальной ортотропной ошибки аппроксимации диадическим паттерном производилась совместно с С.М. Карпенко.

**Публикации.** Основные результаты по теме диссертации опубликованы в 3 статьях в журналах из перечня ВАК, 1 из которых индексируется системой Web of Science. Помимо этого результаты доложены на 6 международных конференциях, из них 1 – российская, а остальные – зарубежные. Все доклады опубликованы в трудах конференций, 4 из которых индексируется системой Web of Science. Кроме того по теме диссертации опубликован 1 препринт на портале arXiv.

**Объем и структура работы.** Диссертация состоит из введения, трёх глав, заключения и одного приложения. Полный объём диссертации составляет 118 страниц, включая 24 рисунка и 1 таблицу. Список литературы содержит 154 наименования.

## Глава 1. Алгоритмы быстрого преобразования Хафа

Глава посвящена созданию и исследованию алгоритмов быстрого преобразования Хафа в постановке М. Брейди [1; 2] для двумерного и трехмерного изображений.

**Определение 1.** *Изображением ( $m$ -мерным изображением со сторонами  $\langle n \rangle = \langle n_1, n_2, \dots, n_m \rangle$  будем называть отображение  $I_{\langle n \rangle}^m : \mathbb{Z}_{\langle n \rangle}^m \rightarrow G$ , где  $\mathbb{Z}_{\langle n \rangle}^m \stackrel{\text{def}}{=} \prod_{i=1}^m \{0, 1, \dots, n_i - 1\} \subset \mathbb{R}^m$ ,  $m, n_i \in \mathbb{N}$ ,  $(G, +)$  – аддитивная абелева группа.*

Изображение  $I_{\langle n' \rangle}^m$  будем называть  $m$ -мерным равносторонним изображением со стороной  $n$  ( $n'_i = n, i = \overline{1, m}$ ) или изображением размера  $n^m$ . Такие изображения будем обозначать  $I_n^m$ , а их область определения –  $\text{dom } I_n^m = \mathbb{Z}_n^m$ . Аналогичную нотацию будем использовать и для прочих символов, значение которых зависит от размеров изображения. В диссертации, если это не оговорено отдельно, рассматриваются равносторонние изображения размерности 2 и 3 со стороной  $n = 2^k, k \in \mathbb{N} \cup \{0\}$ .

Пару  $\langle z, v \rangle$ , где  $z \in \mathbb{Z}_{\langle n \rangle}^m, v \in G$  будем называть пикселем, кортеж  $z$  – позицией, его элементы – координатами (пикселя), а  $v$  – значением пикселя. Позицию с нулевыми координатами будем называть началом координат и обозначать  $\mathbf{0}$ .

Для дальнейшего обсуждения требуется внести некоторую терминологическую ясность касательно терминов «прямая», «паттерн» и «отрезок», используемых при обсуждении этих алгоритмов. Изображение математически может быть представлено в виде дискретной решётки заданной размерности, на ограниченном множестве, в каждом узле (пикселе) которой содержится некоторое значение. Таким образом словоупотребление «поиск прямой на изображении» некорректно, поскольку прямая выходит за его пределы. Логично было бы заменить в этой фразе слово «прямая» на слово «отрезок», однако и в этом случае выражение будет содержательно неверным, поскольку Хаф-образ двумерного изображения в постановке М. Брейди двумерный, а множество всех отрезков, задаваемых концами, четырёхмерно. Следовательно, при разговоре о БПХ нужно либо дополнительно оговариваться о каких именно отрезках идёт речь, либо использовать другую терминологию.

В англоязычном научном сообществе и в некоторых коллективах русскоязычного эту проблему принято решать с помощью термина «паттерн». Паттерн – это множество пикселей на изображении, координаты которых связаны некоторым параметрическим уравнением или заданы перечислением. В такой манере назовём паттерн, аппроксимирующий сегмент прямой, проходящей через область изображения, *прямолинейным паттерном*, а паттерн, аппроксимирующий гиперплоскость – *плоскостным*. Недостаток такой терминологии заключается в заимствовании англоязычного термина, однако другого термина, который с достаточной точностью бы определял объект и был бы лаконичным (ввиду предмета исследования, частота использования этого термина велика) найти не удалось.

Таким образом фраза «поиск прямолинейного паттерна на изображении» проходит оба критерия: проверку на математическую корректность и применимость для описания объектов, которыми оперируют алгоритмы быстрого преобразования Хафа в постановке М. Брейди.

В связи с вышесказанным в работе будет использоваться слово *паттерн*. Результаты данной главы опубликованы в работах [3–7].

## 1.1 Обзор существующих алгоритмов преобразования Хафа

### 1.1.1 Ранние работы

В 1959 году выходит в свет публикация Пола Хафа с описанием метода поиска треков частиц в пузырьковой камере [8] по фотографии. Метод заключается в реализации двух идей. Первая состоит в разбиении изображения на прямоугольную сетку, в каждом элементе (фреймлете) которой производится поиск отрезка, близкого в некотором смысле к наибольшему количеству сигналов. Полученные отрезки, образующие кусочно-непрерывную линию, считаются аппроксимацией трека (не обязательно прямолинейного) на изображении. Вторая идея заключается в способе поиска отрезка внутри фреймлета: каждому ненулевому сигналу ставится в соответствие прямая, так, что для сигналов, лежащих на одной прямой в исходном фреймлете, соответствующие им прямые

пересекались в одной точке, называемой «узлом». Таким образом, чем больше прямых «голосует» за узел, тем больше соответствующих пикселей лежат на одной прямой, в результате задача сводится к поиску узла, набравшего большее число голосов. И, наконец, автор предлагает решение проблемы параметризации угловым коэффициентом (невозможность задания прямой  $x = const$ ), а именно больших значений наклона  $k$ , путем выполнения процедуры детекции прямых с наклоном  $|k| \leq 1$  для исходного и повернутого фреймлета.

Основным вкладом работы является вторая идея, поскольку именно она объединяет множество созданных к этому моменту методов и алгоритмов в названии которых фигурирует фамилия Пола Хафа. Спустя три года преобразование было запатентовано автором в США [9].

Интересным с исторической точки зрения является факт, что ни в публикации, ни в патенте П. Хаф не дал точного описания второй идеи, а лишь описал основные свойства, которыми должно обладать обсуждаемое им «плоскостное преобразование». Более того автор обошёл стороной вопросы дискретизации исходного и целевого пространств и способ дискретизации прямых (из которых в дальнейшем собираются более сложные паттерны). Во многом это, конечно, связано с низким общим технологическим уровнем того времени, отчего эти вопросы, по всей видимости, легли на плечи специалистов по работе с ЭВМ и остались без внимания автора. В конечном счете во многом благодаря культуре цитирования научного и патентного сообщества сейчас преобразование носит фамилию Пола Хафа, хотя тот, строго говоря, его не изобретал.

Следующей значимой вехой в истории преобразования Хафа, согласно [10], является известная книга Азриеля Розенфельда [11]. В ней автор со ссылкой на патент впервые описал способ применения второй из предложенных идей для детекции линий на изображении. А. Розенфельд ещё вскользь упомянул о другой интерпретации ПХ как способа обновления аккумулятора в пространстве параметров для каждой ненулевой точки входного сигнала (метод прямой проекции). Этим автор подчеркнул, что при работе с большим числом точек во входном изображении для поиска наилучшей прямой эффективнее, дискретизовав пространство параметров, обновлять аккумулятор для каждого ненулевого пикселя, а не перебирать все пары точек на изображении. Иначе говоря, сложность решения задачи поиска прямой с использованием ПХ определяется не объемом входных данных, а шагом дискретизации в исходном и конечном пространствах. Таким образом, для любого алгоритма детекции прямой, сложность

которого зависит от числа точек, есть порог по этому числу, начиная с которого его эффективность будет ниже, чем у ПХ. В то время рассматривались преимущественно бинарные изображения, но эта идея очевидным образом обобщается и на изображения в оттенках серого. С исторической точки зрения данная ремарка важна, поскольку это первая работа, где в описании преобразования Хафа появляется операция сложения, что существенно (с точностью до вопроса дискретизации исходного пространства и пространства параметров) сближает ПХ с дискретным преобразованием Радона. Впоследствии было опубликовано множество работ, исследующих вопрос сходства и различия преобразования Хафа и преобразования Радона, но об этом позже.

Третья значимая публикация – работа П. Харта и Р. Дуды [12]. В ней авторы предлагают альтернативное решение проблемы бесконечного аккумулятора (для прямой, близкой к вертикальной,  $k$  очень велико) за счет параметризации прямой нормалью, заданной в полярной системе координат. Описан метод вычисления сумм по всем прямым, продемонстрированы некоторые свойства двойственности исходного пространства и его синопаммы. В работе описан способ обобщения предложенного метода для поиска любой аналитической кривой. Существенный вклад данной работы еще и в подробном описании преимущества использования ПХ с точки зрения вычислительной эффективности. История изобретения и модификации ПХ до 1972 года описана в интервью П. Харта в статье 2009 года [10].

Работу П. Харта и Р. Дуды продолжил И. Дэвис [13]. По-существу авторы предложили два нововведения: 1. изменить способ дискретизации пространства Хафа путём округления вниз параметров нормали (расстояние и угол) рассматриваемой прямой в полярной системе координат; 2. переместить начало системы координат в центр изображения. По заявлению авторов, эти нововведения в сочетании с базовыми техниками предварительной фильтрации изображений позволили улучшить точность детекции параметров прямой, а также уменьшить объем памяти, занимаемой Хаф-образом. Аналогичные исследования велись и в работе [14]. Однако в последней автор заметил необходимость изменить классическую полярную параметризацию, чтобы избежать возникающей избыточности Хаф-образа.

Наряду с этой работой было предпринято множество попыток разработать новые алгоритмы (и улучшить существующие) вычисления ПХ в различных задачах анализа изображений, о чем свидетельствует высокая публикационная

активность научных коллективов по всему миру. Например, по запросу «hough transform» ресурс «google.scholar.com» выдаёт 121000 результатов. К настоящему моменту уже написано семь обзорных работ о способах его обобщения, ускорения и применения в практических задачах [15–21]. Стоит упомянуть также ряд отечественных работ, посвященных описанию свойств и способов применения ПХ [22–24]

В первую очередь стоит отметить, что идея обновлять аккумулятор для каждой входной точки (или пикселя) при подсчете преобразования Хафа использовалась для конструирования не только алгоритмов детекции прямолинейных паттернов, но и для детекции других форм, как заданных аналитически (окружности, парабола и пр.), так и перечислением. Первые идеи детекции паттернов произвольной формы на изображении были изложены ещё в патенте П. Хафа и детальнее описаны в вышеупомянутой статье П. Харта [12], а затем получили развитие в работе [25], где предложен способ ускорения процедуры детекции за счет параллелизации вычислений.

Однако предложенные алгоритмы обладают двумя недостатками, существенными на практике: неточностью аппроксимации детектируемого паттерна и слишком высокими требованиями к вычислительным ресурсам. Это привело к различным попыткам ускорить процедуру детекции за счет учета формы детектируемого объекта (что позволяло уменьшить размерность собираемого Хаф-образа), за счет рандомизированных и вероятностных техник и других алгоритмических приемов. Так, в 1972 году, была предложена принципиальная идея уменьшить размерность пространства Хафа для эллипсов путем учета информации о возможном расположении центра эллипса, согласно градиенту, вычисленному в данной точке [26], которая может быть распространена на любые аналитические кривые. Предложены решения для ускорения поиска кругов [27; 28]. Спустя 18 лет была опубликована обзорная работа о различных методах детекции кругов и эллипсов на изображении [29].

На задачу вычисления ОПХ можно посмотреть иначе, – как на вариацию задачи о вычислении семейства. Вы предлагаете поставить тире

В 1981 году Д. Баллард опубликовал обобщение для любых форм линий (паттернов) [30], вычисляемых на сером изображении, впервые введя термин обобщенное преобразование Хафа (ОПХ). Спустя год в работе [31] был предложен иерархический метод вычисления ОПХ для пирамиды масштабов изображений, что позволило сократить количество вычислений, чего, однако,

оказалось недостаточно для использования алгоритма в системах реального времени. Позднее этот подход развивался в работе [32]. И уже в 2003 Уйлрих предложил использовать иерархическое составление и вычисление  $R$ -таблиц [33], что, по заявлениям авторов, позволяет использовать данный алгоритм в режиме реального времени с целью формирования пространства признаков изображения и решения задач распознавания образов.

Альтернативной веткой развития ОПХ является рандомизация выбора паттернов для вычисления [34; 35]. Данный подход позволяет значительно уменьшить количество необходимых операций и объем используемой памяти, но появляется вероятность ошибиться в подборе данных при построении модели.

На задачу вычисления ОПХ можно посмотреть иначе, как на вариацию задачи о вычислении семейства [7; 36]: для наперед заданного паттерна (прямая, круг и прочее) найти такой набор подмножеств множества пикселей, подлежащих суммации, чтобы количество операций суммирования было минимально. Стоит отметить, что ключевой операцией может быть не только суммирование, но и любая другая двухоперандная (двуместная) операция (вычитание, максимум, минимум), так как данное обстоятельство существенно расширяет множество признаков изображения, вычисляемых средствами ОПХ. Таким образом, стоит задача создать генератор алгоритмов, параметризуемый размером изображения, формой целевого паттерна и типом операции. Важно заметить, что такой подход, в сущности, является обобщением подхода Уйлриха, где работа исследователя по созданию алгоритмических конструкций, ускоряющих полный перебор, подлежит автоматизации средствами создаваемого генератора алгоритмов. Задача о вычислении семейства является NP-полной [36]. В работе Пиппинжера 1980 года предложен обобщенный алгоритм на примере вычисления множества мономов [37]. Несмотря на NP-полноту данной задачи, вопрос о построении аппроксимации, решение которой имеет полиномиальную сложность, остаётся открытым и рассматривается в работе [38]. Подробности исследования и развития ОПХ изложены в обзорах [16; 18; 21].

В своей обзорной работе Мукхопадхья [21] указал 7 недостатков стандартного преобразования Хафа (СПХ) для прямых и для паттернов общего вида (автор так называет алгоритм, предложенный П. Хартом и Р. Дудой в упомянутой работе [12]):

1. Время вычисления и объем требуемой памяти растет экспоненциально с числом параметров паттерна, необходимых для его задания на изображении.
2. Дискретизация входного изображения и пространства Хафа ограничивают максимальную точность определения параметров модели. Для повышения точности необходимо уменьшать шаг дискретизации (когда это возможно), что, однако, увеличивает время выполнения и объем необходимой для вычисления памяти. Не предложено оптимальной стратегии для определения баланса между точностью вычисления ПХ на конкретном изображении и временем вычисления.
3. Равномерная дискретизация пространства параметров ведёт к неравномерному распределению точности детекции паттерна на изображении.
4. Процедура голосования в окрестности истинных параметров устроена таким образом, что максимум «размывается», что влияет на точность его локализации.
5. Для детекции паттерна другого типа необходимо повторно запускать процедуру вычисления преобразования.
6. Независимо от оценки шумности вклад одного пикселя в элементы пространства параметров одинаков (слепое голосование), что приводит к снижению точности детекции.
7. Преобразование Хафа не может автоматически детектировать концы отрезков.

Поскольку в данной диссертации рассматриваются алгоритмы БПХ в постановке П. Брейди (для прямолинейных и плоскостных паттернов), далее критика Мухопадхая будет детальнее обсуждаться в контексте работ, посвященных созданию методов быстрого вычисления сумм по всем прямолинейным паттернам (и плоскостным паттернам) на изображении (или по их аппроксимациям). Поэтому и пункты 1, 5 из данного списка не относятся к предмету исследования диссертации.

Резюмируя описанную критику, все пункты, применимые к ПХ для прямолинейных паттернов, можно разделить содержательно на две части: либо СПХ требует слишком много вычислительных ресурсов, либо точность для решения некоторой задачи недостаточна (будь то поиск прямолинейного паттерна или отрезка на изображении, либо вычисление всего Хаф-образа). Соответственно, исследования в массе были направлены на компенсацию этих недостатков.

### 1.1.2 Способы параметризации прямых на плоскости

Среди работ, посвященных выбору наилучшей параметризации для решения задачи вычисления Хаф-образа или детекции прямой, стоит выделить статью [39]. В ней авторы предлагают три критерия: «уникальность», «ограниченность», «равномерность» – которые позволяют охарактеризовать любую параметризацию. Авторы утверждают, что хорошая параметризация обладает всеми тремя свойствами. Наиболее важным и редко встречающимся условием авторы называют «равномерность». Из уже упомянутых работ, в которых предложен новый для того времени способ параметризации, всем критериям удовлетворяет только две параметризации: та, что предложена Р. Дудой и П. Хартом, и предложенная в статье [13], продолжающей работу. Однако, ни та, ни другая не линейны, то есть образом точки в Хаф-образе является не прямая.

Существует некоторое количество работ, целью которых было создать хорошую (согласно [39]) параметризацию с ограниченным Хаф-образом. Ясно, что свойством равномерности параметризация угловым коэффициентом не обладает. Для решения этой проблемы в работе [40] было предложено очевидное решение: определять точки прямой углом и сдвигом, используя уравнение вида  $y = tg(\alpha)x + c$ . Однако, как и обычная параметризация угловым коэффициентом, данная параметризация для некоторых входных данных порождает Хаф-образ неограниченного размера.

Несмотря на некоторое продвижение П. Харта и Р. Дуды к параметризации, обладающей хорошими свойствами, они утратили очень существенное свойство: двойственность точек и прямых. При использовании параметризации угловым коэффициентом точке в исходном пространстве соответствует прямая в двойственном и наоборот. Свойство ценное, поскольку на его основе возможно разрабатывать более сложные конструкции, такие как повторное преобразование Хафа для детекции точки схода на исходном изображении, и другие.

Одной из первых работ, в которых предложен новый способ параметризации, оказалась работа 1998 года Т. Тьютелаарса [41]. В работе ставится задача анализа аэрофотоснимков, а также рассматривается вопрос детекции точки схода на этих изображениях. Авторы справедливо заметили, что если съемка производится с камеры, оптическая ось которой направлена вертикально вниз, то точки схода параллельных линий на плоскости земли (соответствующей

щие параллельным линиям домов и дорог) будут расположены в бесконечности, а значит, существует такой пучок линий, ориентацию которого в обычной параметризации угловым коэффициентом невозможно определить. Для решения этой проблемы авторы предложили разбить исходное пространство координат изображения и пространство параметров прямых таким образом, чтобы любая точка схода находилось в области определения одного из изображений, а точка схода любого пучка линий была в пределах одного из полученных Хаф-образов. Вместе с предложенной параметризацией авторы предложили новый метод детекции прямых и точек схода – так называемое каскадное преобразование Хафа – однако в работе не содержится исследования эффективности предложенного метода и, судя по всему, дальнейшего развития данный подход не получил.

Следующей работой, где был предложен новый способ параметризации прямых, проходящих через область изображения, стала статья Р. Уолоса [42] под названием «Модифицированное преобразование Хафа», или короче «Маф-преобразование». В сущности, предлагается задавать прямую точками пересечения с границами изображения (параметризацию, которая связывает прямые с пиксельной решеткой изображения называют *натуральной*), координаты которых отсчитывать вдоль его периметра по часовой стрелке. А чтобы одной прямой соответствовал ровно один набор параметров, достаточно ограничить диапазон изменения координат величиной периметра и ввести между ними порядок. К сожалению, в открытом доступе исходного текста данной публикации нет, но, согласно [18], основным преимуществом предложенного метода стало то, что разрешение представленных линий равномерно и совпадает с множеством линий, которые могут быть изображены при помощи цифровой графики. Из контекста научных достижений и техник того времени становится понятно, что уже тогда уровень проработки техники цифровой графики был достаточно высок (к этому времени алгоритм Дж. Брезенхема был уже предложен и общеизвестен [43]). По всей видимости, автор имел в виду, что такой способ задания прямой позволяет естественным образом задавать дискретизацию пространства Хафа при заданной дискретизации исходного пространства, что является основным достоинством натуральной параметризации. Однако нельзя не заметить, что такое преобразование обладает существенным недостатком: образ точки на исходном изображении имеет хоть и аналитическую, но сложную форму.

В работе [44] предложена альтернативная параметризация, в некотором смысле, промежуточная между Маф-параметризацией и параметризацией, предложенной П. Хартом. Авторы предлагают параметризовать паттерн точкой на изображении и углом. Основная цель, которую пытались достичь таким нововведением авторы – уменьшение размера памяти для хранения Хаф-образа. Такой способ задавать прямые, как и Маф-параметризация, однако, порождает сложно структурированный Хаф-образ, что затрудняет его анализ. По всей видимости, поэтому данная ветвь исследования способов параметризации не получила дальнейшего развития.

Выбор параметризации посвящены также работы в области томографии [45], что естественно, поскольку преобразования Хафа и задача томографического восстановления тесно связаны между собой. Авторы этой работы предлагают другую модификацию Маф-преобразования, а именно определение двух точек на окружности, описанной вокруг изображения, с помощью их угловых координат. Заметим, что такой способ параметризации выглядит естественным в томографии, поскольку тесно согласован с позиционированием при томографировании. В обзорной работе [20] справедливо замечено, что при использовании такого преобразования Хаф-образ получается более гладким, чем при Маф-параметризации.

В работе [46] предлагается каждой точке на изображении ставить в соответствие окружность такую, что для точек, лежащих на одной прямой, соответствующие им окружности пересекались бы в одной точке. Несмотря на изящество предложенного способа параметризации и на то, что он позволяет хранить весь Хаф-образ изображения, используя гарантированно ограниченный объем памяти, его мало используют на практике. Причины низкой популярности предложенного метода, по всей видимости, две: 1. нетривиальность самой математической конструкции, 2. центральная область аккумулятора, согласно [20], содержит значительно бóльшие значения ввиду конструктивных особенностей такого Хаф-образа, что затрудняет его анализ.

Общим недостатком параметризаций, при использовании которых точке на исходном изображении ставится в соответствие не прямая, является нетривиальность анализа полученного Хаф-образа, – в частности, невозможность повторного взятия преобразования Хафа. Рассмотрим исследования по созданию линейной параметризации, где образом точки является прямая.

В работе [47] продемонстрирован интересный способ использования параллельной параметризации [48], обычно используемой для визуализации многомерных данных. Для двумерного изображения предлагается построить следующую конструкцию: на плоскости строятся две параллельные оси, первая соответствует абсциссе исходной точки, вторая – ординате. Далее для каждой ненулевой точки на исходном изображении выполняется операция растрового построения прямой в новом пространстве (инкрементируются все точки изображения, лежащие между соответствующими точками на новых осях). Таким образом, точки, лежащие на одной прямой в исходном пространстве, порождают пучок пересекающихся линий. Для включения несобственной прямой в вычисляемый Хаф-образ авторы дополняют свою СК еще одной (третьей) осью, направленной противоположно. Существенное преимущество такого способа – тривиальность его обобщения на случаи большей размерности. Тем не менее он остаётся слишком затратным вычислительно, а потому, несмотря на оригинальность идеи, редко используется на практике.

Как было сказано выше, одним из недостатков полярной параметризации нормали к прямой является сложная структура получаемого Хаф-образа: точке на исходном изображении соответствует синусоида в пространстве Хафа. Однако такая параметризация порождает менее наглядный Хаф-образ, более того – затрудняет выполнение повторного взятия преобразования Хафа, что выполняется при поиске точки схода. В связи с этим «голубой мечтой» специалистов этой области было найти такую параметризацию, которая сохраняла бы двойственность прямой и точки, а также не приводила бы к проблемам с хранением Хаф-образа (первоначальная проблема с изломом пространства Хафа в силу его неограниченности). Однако в работе [49] было теоретически доказано, что создание такой параметризации невозможно. Идея предложенного доказательства очень проста: достаточно рассмотреть множество прямых в области изображения, пересекающихся в одной точке. Ясно, что для задания прямой, проходящей через некоторую точку, достаточно одного числа – угловой координаты. Нетрудно понять, что угловая координата является одномерным континуумом, а значит, нет гарантии, что соответствующие пики в Хаф-образе будут находиться в пределах изображения.

Вопрос сравнения разных параметризаций между собой, перечисления их свойств и визуализация результатов их использования подробно изложены в работе [20]. В этой работе также приведены обоснования существования “рав-

номерной” дискретизации, что закрывает пункт 3 в вышеизложенном списке Мукхопадхая.

### 1.1.3 Дискретизация и её влияние на точность

Исследование методов дискретизации исходного пространства осмысленно, только если входными данными является облако точек, а не изображения или гистограмма с уже заданными параметрами дискретизации. В этом направлении исследования ведутся либо чтобы повысить точность детекции прямой (гиперплоскости), либо для ускорения.

Согласно [18], при работе с разреженными данными нужно использовать неравномерность распределения точек в пространстве: возможно как ускорить вычисление Хаф-образа (или поиска линии на изображении), так и уменьшить объем требуемой памяти за счет неравномерной дискретизации пространства. Пионерами исследований в данном направлении являются Дж. О’Рурк [50] и К. Слоун [51; 52]. По существу эти авторы предложили две новые структуры хранения данных: динамически дискретизирующееся пространство (DQS – dynamically quantized space) и динамически дискретизирующаяся пирамида (DQP – dynamically quantized pyramid). Первая структура – это бинарное дерево, каждый узел которого соответствует некоторому прямоугольному региону изображения. При сборе данных регионы пространства Хафа разбиваются пополам или сливаются воедино так, чтобы количество данных в каждом регионе было примерно одинаковым. Общее число узлов дерева – это параметр метода. В итоговом дереве большее количество регионов будет сконцентрировано в местах сгустков данных. Вторая структура основана на использовании многомерного дерева квадратов, где параметры связанности регионов зафиксированы, а число регионов мало. Недостатком второй структуры является именно это свойство, поскольку это приводит к ограничению точности по пространственному разрешению. В работе [52] авторы экспериментально исследовали обе предложенные структуры. По результатам были сформулированы следующие выводы: структуру DQS сложнее реализовать, но она позволяет получать более точный результат. Однако немногим позднее в работе [53] предложен способ модификации входного множества данных, позволяющий использовать DQP с большей

точностью, однако такая процедура довольно тяжеловесна (выполняется в три раза дольше, чем формирование DQP).

Другая полезная идея, предложенная независимо различными исследователями, – использовать пирамиды масштабов. Идея заключается в том, чтобы последовательно анализировать одно и то же изображения в разном разрешении. По структуре можно принимать решение о том для каких областей требуется более детальный анализ, а какие могут быть отброшены с целью экономии вычислительных ресурсов. Эта стратегия эффективна при анализе многомерных Хаф-образов (для ОПХ или СПХ изображений размерности три и более). На практике данный подход использовался в работе Дж. Адива [54], как эффективный способ поиска параметров движения в пространстве высокой размерности и Терезой Силберг [55] с целью распознавания трехмерных объектов на изображении.

Опубликована работа по ускорению вычисления ПХ, предложенных Х. Ли и его коллегами [56]. Авторы разработали быстрый алгоритм вычисления преобразования Хафа, пригодный для решения задач детекции линий, плоскостей, сегментации и оценки движения. Предложенный алгоритм авторы называют быстрым преобразованием Хафа, однако это принципиально другой подход, существенно отличающийся от того, который используется в данной диссертации (быстрого преобразования, предложенного П. Брейди [1]). Поэтому, во избежание путаницы, будем называть данный алгоритм «иерархическое быстрое преобразование Хафа» (ИБПХ). Этот алгоритм заключается в рекурсивном разделении пространства параметров на области, ограниченные гиперкубом, повышая разрешение и вычисляя преобразования Хафа только для областей с числом голосов, превышающим некоторый заданный порог. Голосует ли точка изображения за данный гиперкуб или нет, определяется пересечением гиперплоскости соответствующей данной точки с гиперкубом в пространстве параметров. Согласно авторам, ИБПХ позволяет значительно уменьшить количество вычислительных операций и объём используемой памяти. Более того, ИБПХ может быть реализован без использования операций перемножения, а только со сложением (сложение на вычислителях классической архитектуры выполняется быстрее). Одна из проблем данного подхода состоит в необходимости подбора вышеупомянутого порога. Согласно [18], на практике это может быть затруднительно: если его установить слишком низким, то процедура вычисления становится чрезмерно трудоёмкой, если слишком высоким – точность

алгоритма может быть недостаточна. В обзоре данный метод критикуют за сложность процедуры голосования, которая требует вычисления расстояния от гиперплоскости до различных кубов, структура которых в пространстве параметров динамически меняется. В конечном счете данная структура становится статичной и не позволяет оптимально адаптироваться к любой возможной конфигурации входных данных, в том числе поддерживать различную дискретизацию пространства по разным осям. Для решения последней проблемы авторы предложили использовать другую структуру данных [56].

Позднее в работах [57; 58] Дж. Илингворт и Дж. Китлер разработали итеративный подход, работающий по принципу «от простого к сложному», для детекции линий и кругов, определяемых в двумерном и трехмерном пространствах параметров соответственно. Идея, лежащая в основе данных работ, аналогична предыдущим: адаптировать точность подсчета Хаф-образа лишь там, где больше голосов входного сигнала. В работе [57] авторы утверждают, что полученный алгоритм значительно превосходит СПХ, однако экспериментальное подтверждение этому они приводят, используя синтетические изображения с малой долей равномерного шума на изображении, что плохо моделирует реальные изображения даже бинарные. Позднее авторы в своем знаменитом обзоре [18] и сами отмечают, что на практике использование этих методов не дает ожидаемого прироста в скорости вычислений и эффективности использования памяти.

Важно понимать, что в интерпретации многих ученых преобразование Хафа есть некоторый алгоритм детекции объекта, а не алгоритм расчета Хаф-образа. Из таких интересно отметить работы К. Брауна и Д. Шер [59; 60]. Авторы предложили уменьшать объем вычислительных ресурсов с использованием адресной памяти (хэш-таблица). Голоса точек постепенно хранятся в памяти, пока она не заполнится, после чего из памяти выбрасываются наименее популярные кандидаты (предложены также другие стратегии опустошения памяти). Понятно, что в этом случае сложно говорить о явной дискретизации пространства Хафа, однако неявным образом она происходит. Авторы провели подробное исследование эффективности такого подхода и пришли к выводу, что использование кэша не приводит к значительному улучшению, а для изображений, содержащих шум, такая стратегия часто приводит к ошибкам.

В статье [61] автор заметил, что если при задании дискретизации аккумулятора определять её некоторым шагом, не привязанным к решётке на

исходном изображении, то различные прямые в некоторых случаях будут голосовать за одну и ту же ячейку аккумулятора. Потому Томас Риссе предложил так называемую «hifi» дискретизацию. Основное её свойство – что для каждого рассматриваемого линейного паттерна, аппроксимирующего прямую, на изображении вычисляется отдельная сумма. Изначально эта идея была плохо принята из-за того, что количество разных дискретных паттернов на изображении равно  $\Theta(n^5)$ , где  $n$  – линейный размер изображения, а позднее в работе [62] было показано, что всевозможных прямых, построенных по правилу, реализованному в алгоритме Брезенхема, всего  $\Theta(n^4)$ . Позднее стали использовать координаты пикселей на краях изображения для задания прямых, в этом случае число прямых составляет  $\Theta(n^2)$ . Именно поэтому появилась серия работ по созданию способов параметризации прямых, при которых дискретная решётка исходного изображения определяет дискретизацию в пространстве параметров (Маф-параметризация и её модификации).

Рассмотрим работы в которых исследовалось влияние способа дискретизации на точность решения задачи детекции прямой или вычисления Хаф-образа.

Первыми учеными, которые поставили и предприняли попытки исследования вопросов о влиянии дискретизации на точность детекции параметров паттерна были С. Шапиро и А. Янино [15; 63]. Отдельно стоит отметить работу авторов, где они описывают геометрический способ оценки степени влияния шумности входного сигнала на точность определения параметров детектируемой прямой [64]. Фокус работ этих авторов смещен в большей степени на анализ точности определения параметров Хаф-образа для сложных паттернов (не прямолинейных). Годом позже авторы предложили идею использования ПХ для сжатия изображения [65]. Суть метода заключается в том, чтобы детектировать прямые и другие линии, которые можно аппроксимировать прямолинейными сегментами и производить их векторизацию, уменьшая объем хранимых данных. Позднее идею использования информации о наличии прямой в некоторой области изображения применили в системах сжатия видео “H.264”.

Стоит упомянуть другие работы, исследующие влияние шума на точность детекции параметров прямой с помощью преобразования Хафа [66; 67]. Аналогичные исследования выполнялись применительно и к не прямолинейным паттернам, об этом достаточно подробно написано в обзоре [57] в разделе 2 “Ранние разработки”.

При обзоре результатов, полученных в области изучения преобразования Хафа, его свойств и модификаций, невозможно не упомянуть Н. Кирьятти и А. Брукштейна. В одной из своих многочисленных работ они исследуют деградацию точности детекции параметров прямой в зависимости от шага дискретизации исходного пространства [68]. Авторы предлагают *цифровое преобразование Хафа* (ЦПХ), суть которого заключается в построении максимально точных аппроксимаций прямых. Годом ранее в работе [62] было показано, что число различных дискретных паттернов, аппроксимирующих все возможные прямые, проходящие через сторону изображения, составляет  $O(n^4)$ , где  $n$  – линейный размер квадратного изображения. Авторы показали, что с использованием ЦПХ на равномерной дискретной решетке точность определения параметров прямой выше, чем при использовании меньшего количества различных дискретных паттернов. Однако в дальнейшем от этого направления отказались, ввиду высокой асимптотической вычислительной сложности алгоритмов суммации по всем этим прямым.

Важно также упомянуть работу В. Ниблэка в этой области [69]. Автор рассматривает различные методы повышения точности определения параметров прямой по Хаф-образу. В работе проведена большая серия вычислительных экспериментов по изучению точности детекции прямой при использовании полярной параметризации аккумулятора [12]. В работе авторам удалось показать, что путём интерполяции Хаф-образа возможно достичь большей точности в определении пространства параметров детектируемого паттерна на исходном изображении.

#### 1.1.4 Связь преобразования Радона и преобразования Хафа

Преобразование Радона на евклидовой плоскости – это интегральное преобразование, относящее к функции  $f$  на плоскости её интегралы по всевозможным прямым (относительно евклидовой длины) [70]. Преобразование определено для функции  $f$ , заданной на конечном носителе, и впервые предложено Иоганном Радонам в 1917 году [71]. Позднее это преобразование было обобщено Р. Пенроузом [72] для описания кванто-механических событий в пространстве-времени. ПР широко используется при решении задач томогра-

фии [73; 74]. Дискретное преобразование Радона определяется аналогично ПР для дискретного плоского пиксельного изображения. Ввиду вышеописанных свойств ПХ ясно, что существует различное множество способов как дискретизации исходного пространства, так и пространства Радона; более того – нетривиален и способ построения дискретного паттерна, аппроксимирующего прямую. Все эти вопросы детально изучались при разработке алгоритмов ПХ, а вот аналогичных работ, посвященных ПР, в публичном доступе найти не удалось. Возможно, это связано с тем, что ПР обычно рассматривают на евклидовой плоскости, а не на дискретной решётке.

Именно к такому выводу приходит автор в работе [75]. В этой статье показано сначала, что преобразование Радона обладает всеми свойствами из списка Р. Дуды и П. Харта [12], после чего показывает, что в некотором смысле ПХ можно назвать частным случаем ПР. Все рассуждения автор проводит аккуратно, с использованием аппарата функционального анализа.

Спустя 23 года выходит другая работа, согласно которой ПР и ПХ – это одно и то же [76]. Несмотря на это, авторы предприняли попытку концептуального разделения ПР и ПХ. Верно замечено, что ПР рационально использовать при работе в *парадигме чтения* (рассматривая вопрос формирования значения в точке в пространстве Радона из значений элементов в исходном пространстве), а ПХ – в *парадигме записи* (рассматривается вопрос преобразование значения в точке исходного пространства в образ пространства Хафа). Однако авторы утверждают, что природа ПХ исключительно дискретная, что неверно, и было показано в начале данной главы.

Только спустя 30 лет развития методов преобразования Хафа в 1992 была опубликована работа, где впервые даётся его формальное определение [77]. В этой работе авторы обобщили ПХ на непрерывный случай, в виде комбинации преобразования Радона с предварительной свёрткой исходного изображения (которое формируется как сумма дельта-функций с координатами точек – узлов изображения). Этот формализм устанавливает конкретную связь между ПР и М-оценками, что впервые было описано в работе [78].

Стоит отметить, что аналогичный взгляд на ПХ развивается российской группой ученых из Пензы, в частности Федотовым Н.Г., авторы называют его «трейс-преобразованием» [79; 80].

Резюмируя существующие публикации на эту тему можно описать отличие следующим образом. Дискретное преобразование Радона и стандартное

преобразование Хафа вычисляют одну и ту же величину, но разными способами: первое вычисляет суммы по всем прямым, а второе обновляет Хаф-образ для каждого ненулевого пикселя. Таким образом, несмотря на идентичный результат работы, время работы этих двух алгоритмов отличается.

В данной работе, согласно названию М. Брейди и внутреннему устройству алгоритма, рассматривается быстрое дискретное преобразование Радона, однако, с учетом сложившейся терминологии этот алгоритм чаще называют быстрым преобразование Хафа. В связи с этим в данной диссертации тоже используется второе название.

### 1.1.5 Алгоритмы быстрого вычисления ПХ и поиска прямых

Содержательно способы ускорения можно разделить на две категории в зависимости от конечной цели: быстрое вычисление Хаф-образа или быстрый поиск прямой на изображении.

Для начала опишем основные подходы к решению задач второй категории. В работе [81] Н. Кирьяти и А. Брукштейн предложили так называемое *Вероятностное преобразование Хафа (ВПХ)*. Суть алгоритма заключается в выборе случайных параметров прямой и оценке её кандидатуры. Показано аналитически и экспериментально, что данный метод незначительно хуже СПХ по точности, даже в случае большой доли шума. Процедуру выбора параметров можно ускорить путём учета плотности и структуры [82] входных точек. Аналогичный метод предложен в работе [83]

В работах [34; 35] предложены алгоритмы подбора случайной пары точек на исходном изображении, – так называемое *Рандомизированное* преобразование Хафа (РПХ). Существуют некоторые теоретические работы по оценке эффективности этого и предыдущего методов [84], однако они предполагают опосредованное существование некоторой модели изображения. Подробнее об оценке рандомизированных и вероятностных подходов написано в работе [85].

В статье [86] предложен метод, который является комбинацией алгоритмов кластеризации точек исходного пространства и процедур голосования для наиболее выделяющихся кластеров. Важной особенностью метода является зависимость вычислительной сложности от числа точек  $k$ , а не от размера изоб-

ражения, и составляет  $O(k \log k)$ . Правда, в задачах, где входом является не облако точек, а массив значений, выполнение этого алгоритма в режиме реального времени невозможно уже для трехмерного изображения с линейным размером 64 пикселя. Иными словами, алгоритм удобен для использования только в случае разреженных данных.

Существует также множество алгоритмов, целью которых является ускорение вычисления Хаф-образа (первый подход), используя быстрое преобразование Фурье [87–90]. Согласно этим работам сложность вычисления таких алгоритмов составляет  $\Theta(n^2 \log n)$ .

Отдельно интересно отметить, что в работы по ускорению вычисления ПХ с помощью нейронных сетей [91–93].

Рассмотрим теперь способы ускорения вычисления ПХ в постановке М. Брейди [94]. В этом случае на вход алгоритму подаётся изображение (или производится дискретизация пространства, в результате чего возникает изображение), используется  $(s, t)$ -параметризация, а для аппроксимации прямой в области изображения – диадический паттерн (будет описано в следующем разделе). Развитие данного направления актуально для случая, когда входными данными либо является изображение (данные уже были округлены и не существует возможности это исправить), либо когда точек настолько много, что для алгоритмов, асимптотическая вычислительная сложность которых зависит именно от числа точек, расчет ПХ становится чрезмерно длительным. В этом смысле критика, содержащаяся во втором пункте вышеприведённого списка Мукхопадхая, оказывается несостоятельной: использование изображения вместо множества точек в качестве входных данных позволяет за счет уменьшения точности (в результате огрубления координат) значительно ускорить поиск наилучшей прямой.

Итак, несмотря на то, что СПХ стал очень популярным, его оказалось недостаточно для применения на практике: для квадратного изображения без нулей с линейным размером  $n$  вычислительная сложность алгоритма составляет  $O(n^3)$ . Время выполнения алгоритма было слишком велико для применения его в промышленных системах и системах реального времени (насколько известно автору, его не используют и сейчас), а также для применения в анализе многомерных данных.

Для преодоления этой проблемы, был предложен эффективный алгоритм вычисления аппроксимации преобразования Хафа – быстрое преобразование

Хафа (БПХ), в некоторых работах он носит название быстрого дискретного преобразования Радона. Можно сказать, что БПХ вычисляет обобщенное преобразование Хафа для диадических паттернов, которое для классического ПХ является аппроксимацией. Другой особенностью БПХ является то, что он позволяет заменять сложение любой другой ассоциативной операцией, например, вычислением максимума.

Интересно, что алгоритм изобретался несколько раз. Впервые его предложил М. Брейди [1] в 1992 году. Позднее он повторно был предложен австрийским ученым Готсом в диссертации 1993 года [95], затем в 1995 работа была опубликована на английском языке [96]. Третьим изобретателем стал французский профессор информатики Вюйемэн в 1994 [97] году.

Обзор алгоритмов точного и приближенного вычисления преобразования Хафа [21] показывает, что в настоящее время БПХ остается малоизвестным инструментом. Как следствие, его потенциал в системах компьютерного зрения реального времени остается не реализованным. Причинами этому служат как недостаток популяризации этого инструмента, так и малое число исследований на тему применимости алгоритма БПХ для решения различных практических задач.

Стоит отметить, что первая общедоступная программная реализация быстрого преобразования Хафа, приспособленная для использования в системах анализа изображений с открытым кодом, была опубликована лишь в 2015 году усилиями сотрудников лаборатории зрительных систем ИППИ РАН и в том числе членами научного коллектива, в открытой библиотеке OpenCV с целью популяризации данного инструмента. Позднее автор диссертации опубликовал коды алгоритма вычисления БПХ и его обобщений на трехмерный случай на языке MATLAB.

### 1.1.6 Преобразование Хафа в задачах анализа изображений

За десятилетия своего существования этот алгоритм зарекомендовал себя как ценный инструмент специалиста по анализу изображений и техническому зрению, что засвидетельствовано в обзорных работах [18; 21]. В настоящее время множество научных коллективов публикуют работы о том, как использовать

ПХ при решении различных исследовательских и инженерных задач. Очевидно, что в этом случае есть возможность заменить СПХ на БПХ, значительно уменьшив время вычисления соответствующего этапа. Однако в силу того, что у диадических паттернов есть некоторая специфическая ошибка аппроксимации прямой на изображении, точность решения задачи может измениться. Как показывают исследования, проведенные в процессе работы над диссертацией, искажения Хаф-образа на размерах изображения незначительны, что позволяет сделать качественный вывод о допустимости замены СПХ на БПХ. Таким образом, область применения БПХ не исчерпывается только работами, где непосредственно использовался этот алгоритм, но и содержит работы, эксплуатирующие СПХ, а также некоторые его модификации.

Основным способом применения, для которого используется ПХ – это детекция линий на изображении, часто прямых. Одной из формализаций данной задачи является задача робастной ортогональной линейной регрессии. Детальнее об этом вопросе приведено описание в следующей главе. Помимо этого существует множество работ, предлагающих использование ПХ для детекции линий на изображении.

ПХ применяется для анализа изображения документа [98], в частности для определения границ документа [99; 100]. Определив ориентацию документа на изображении, можно проективно исправить изображение так, будто оно снято вертикально, а границы изображения параллельны границам фотографической матрицы. Решение задачи виртуального поворота документа к этому виду позволяет существенно упростить дальнейший анализ документа и реализацию процедур распознавания его структуры и полей. В контексте анализа изображений документов интересно упомянуть работу в которой с использованием БПХ предлагается определять параметры наклона шрифта в тексте [101]. В работе [102] рассмотрены метод использования БПХ для сегментации текста. Сейчас уже возможно конструировать системы технического зрения промышленного качества, используя БПХ.

Другой обширной областью применения ПХ является техническое зрение автономных систем, функционирующих на антропогенных территориях (содержащих большое количество прямолинейных образов). Например, ПХ часто используется для детекции дорожного полотна и нанесенной на нём дорожной разметки [103; 104] и для детекции дорожной инфраструктуры города по снимкам с высоко летящего беспилотника [105].

ПХ часто используется при решении различных задач отслеживания объектов на изображении. Так, например, в одной из пионерских работ [106] предложено, используя покадровую разность изображений, отслеживать равномерно движущиеся объекты, которые на таких треках, очевидно, оставляют сильно выраженные прямые линии. Подробнее о методах отслеживания объектов с использованием преобразование Хафа написано в работе [107].

Ещё одной критически важной задачей является детекция точки схода на изображении. Точка схода – это точка на плоскости изображения в которой пересекаются проекции параллельных в трехмерной сцене прямых. Среди работ, рассматривающих данную задачу, можно упомянуть работы [108; 109].

Интересным способом применения БПХ стало его использование для реализации алгоритмов слепой калибровки радиальной дисторсии [110]. Авторы предлагают новый функционал Хаф-образа изображения, оптимизируя который подбирают параметры компенсации радиальной дисторсии. Основное и единственное предположение, в рамках которого метод функционирует – наличие прямолинейных объектов в сцене.

Ещё одним очень популярным и одним из старейших способов применения преобразования Хафа и преобразования Радона, является задача томографического восстановления [111]. Процедуру формирования проекций изображения анализируемого объекта с использованием рентгеновских лучей принято моделировать преобразованиями Радона и Хафа. В результате задача восстановления трехмерной структуры объекта на практике часто сводится к задаче обращения преобразования Хафа или преобразования Радона, – как говорят, к решению обратной задачи. Данной научной области посвящено большое количество работ, в том числе отечественных [74].

В последние несколько десятилетий стоит отметить значительное повышение интереса к анализу медицинских изображений, подробнее об этом написано в работе [112].

Вообще говоря, список разнообразных применений ПХ и БПХ для анализа изображений и в качестве одного из этапов функционирования системы технического зрения очень большой. Перечислим некоторые ещё не упомянутые, но интересные и оригинальные способы применения ПХ.

ПХ предлагали использовать для распознавания штрих-кодов [113; 114], анализа деформаций подземного профиля [115], определения степени сходства форм [116], детекции треков спутников по изображениям с длинной выдержкой

[117], для обнаружения радиолокационных целей [118; 119], для распознавания рукописного текста [120], для анализа структуры топологии микросхем по фотографиям [121], для подсчета количества осей транспортного средства по трекам детектора колес изображений, полученных с камеры, снятой сбоку [122], для трехмерной реконструкции плоских граней прозрачных минералов по набору изображений [123], для векторизации трехмерных изображения [124], для анализа SAR снимков [125] и так далее.

## 1.2 Диадический паттерн и двумерное преобразование Хафа

Как и любой другой объект, прямая при дискретизации меняет свои свойства, многие из которых зависят от способа дискретизации. В разделе исследуются свойства диадического паттерна и конструкций, построенных на его основе.

**Определение 2.** Паттерном  $p$  размера  $l$  на изображении  $I_{\langle n \rangle}^m$  будем называть множество позиций  $p \subset \mathbb{Z}_{\langle n \rangle}^m$ ,  $|p| = l$ .

Для паттерна  $p$  на изображении  $I_{\langle n \rangle}^m$  введем операцию циклического сдвига на  $s \in \mathbb{Z}_{\langle n \rangle}^m$ :  $p \nearrow s \stackrel{\text{def}}{=} \{ \langle z_i + s_i \pmod{n_i} \rangle_{i=1}^m \mid z \in p \}$ , при этом кортеж  $s$  будем называть величиной сдвига или сдвигом.

Множеством всевозможных сдвигов  $S_{\langle n \rangle}^m(d)$  по осям  $d$ , где  $d \in \{0, 1\}^m$  – индикатор индексов осей, будем называть множество  $S_{\langle n \rangle}^m(d) \stackrel{\text{def}}{=} \bigcup_{s \in \mathbb{Z}_{\langle n \rangle}^m} s \circ d$ , где символ « $\circ$ » обозначает произведение Адамара.

**Определение 3.** Будем говорить, что паттерн  $p$  сдвигами по осям  $d$  порождает на изображении  $I_{\langle n \rangle}^m$  пучок паттернов  $T_{\langle n \rangle}^m(p, d)$ , если  $\mathbf{0} \in p$  и множество паттернов  $T_{\langle n \rangle}^m(p, d) = \{p + s_i \mid s_i \in S_{\langle n \rangle}^m(d)\}$  является разбиением  $\text{dom } I_{\langle n \rangle}^m$ :  $p_i \neq p_j \in T_{\langle n \rangle}^m(p, d) \implies p_i \cap p_j = \emptyset$ ,  $\bigcup_{p_i \in T_{\langle n \rangle}^m(p, d)} p_i = \mathbb{Z}_{\langle n \rangle}^m$ .

### 1.2.1 Свойства диадических паттернов

**Определение 4.** Порождающими диадическими паттернами на изображениях  $I_{2^k}^2$  будем называть паттерны, задаваемые следующим рекуррентным выражением:

$$P_t^k = \begin{cases} \{\langle 0, 0 \rangle\}, & k = 0 \\ P_{\lfloor t/2 \rfloor}^{k-1} \cup \left( P_{\lfloor t/2 \rfloor}^{k-1} \nearrow \langle 2^{k-1}, \lceil t/2 \rceil \rangle \right), & k > 0 \end{cases}, \quad (1.1)$$

где  $t \in \{0, 1, \dots, 2^k - 1\}$  – параметр, который будем называть наклоном паттерна.

Опишем подробнее, что именно за правило приведено в этой формуле. На каждом шаге рекурсии изображение делится вертикально на две равные части, и для каждой из этих частей определяется наклон диадического подпаттерна и сдвиг (0 или 1) между ними. Эти значения задают координаты пикселей на границе разделения:  $(n/2 - 1, \lfloor t/2 \rfloor)$  и  $(n/2, \lceil t/2 \rceil)$ . Затем процедура проводится для каждого подпаттерна до тех пор, пока длина изображения не будет равна единице. Диадический паттерн обладает самоподобной структурой. Используя данное определение, можно показать связь между битовым представлением наклона  $t$  и структурой паттерна. Если в битовом представлении в  $k$ -м разряде стоит единица, значит на  $k$ -м шаге рекурсии каждая пара подпаттернов будет сдвинута на единицу относительно друг друга. Этот способ определения неудобен для анализа структуры паттерна, поэтому в работе [4] дано определение в более привычном виде, приведём его.

**Утверждение 1.** Порождающий диадический паттерн  $P_t^k$  является графиком следующей функции:

$$z_2 = D_t^k(z_1) \stackrel{\text{def}}{=} \sum_{b=0}^{k-1} \left( \left\lfloor \frac{t}{2^b} \right\rfloor \pmod{2} \right) \left[ \frac{2^b z_1}{2^k - 1} \right], \quad (1.2)$$

то есть  $\langle z_1, z_2 \rangle \in P_t^k \iff z_2 = D_t^k(z_1)$ .

*Доказательство.* Докажем, сначала, что утверждение для  $t = 2^r$ .

Очевидно, что  $t \pmod{2} = 1$  при  $t = 2^r$  только на  $r$ -м шаге рекурсии для отрезка длины  $(2^k - 1)/2^r$ .

Множество абсцисс  $z_1$  разделяющих линий для каждого такого отрезка выражается формулой  $(N_z + \frac{1}{2})\frac{2^k-1}{2^r}$ , где  $N_z \in [0, 2^r - 1]$ . Тогда определим множество абсцисс пикселей, располагающихся по одну сторону от разделяющих прямых, как  $\hat{X} = \lfloor (N_z + \frac{1}{2})\frac{2^k-1}{2^r} \rfloor$  и назовем такие пиксели *сдвиговыми*.

Покажем, что множество абсцисс сдвиговых пикселей, полученных согласно определению 2 равно  $\hat{X}$ . Чтобы пиксель с абсциссой  $x$  был сдвиговым, должно выполняться:

$$\begin{cases} \lfloor \frac{2^r x}{2^k-1} \rfloor = m, \\ \lfloor \frac{2^r(x+1)}{2^k-1} \rfloor = N_z + 1, \\ N_z \in [0, 2^r - 1], \end{cases}$$

где  $N_z$  – целое число. Преобразуем данную систему:

$$\begin{cases} \frac{2^r x}{2^k-1} \geq N_z - \frac{1}{2}, \\ \frac{2^r x}{2^k-1} < N_z + \frac{1}{2}, \\ \frac{2^r(x+1)}{2^k-1} \geq N_z + \frac{1}{2}, \\ \frac{2^r(x+1)}{2^k-1} < N_z + \frac{3}{2}, \\ N_z \in [0, 2^r - 1], \end{cases} \rightarrow \begin{cases} x < (N_z + \frac{1}{2})\frac{(2^k-1)}{2^r} \\ x > (N_z + \frac{1}{2})\frac{(2^k-1)}{2^r} - 1 \\ N_z \in [0, 2^r - 1]. \end{cases}$$

Множество решений системы можно записать в виде  $X = \lfloor (N_z + \frac{1}{2})\frac{2^k-1}{2^r} \rfloor$ . Так как  $X \equiv \hat{X}$  утверждение верно.

Теперь докажем утверждение для любого наклона  $t \in [0, 2^k - 1]$ .

Множество сдвиговых пикселей для наклона произвольного  $t$  есть объединение множеств сдвиговых пикселей для каждого базисного наклона  $X = X_{t_0} \cup X_{t_1} \cup X_{t_2} \cup \dots \cup X_{t_{p-1}}$ .

Кумулятивное построение по множеству  $X$  эквивалентно покоординатному суммированию соответствующих базисных диадических паттернов с соответствующими наклонами.

Что и требовалось доказать. □

Из утверждения 1 следует, что порождающие диадические паттерны порождают пучки вида  $T_{2^k}^2(P_t^k, \langle 0, 1 \rangle)$ .

**Определение 5.** Системой паттернов  $Y$  будем называть любое непустое объединение пучков паттернов на изображении  $I_{\langle n \rangle}^m: Y = \bigcup_{i=1}^N T_{\langle n \rangle}^m(p_i, d_i)$ , где  $p_i \subset \mathbb{Z}_{\langle n \rangle}^m$ ,  $d_i \in \{0, 1\}^m$ ,  $N \in \mathbb{N}$ .

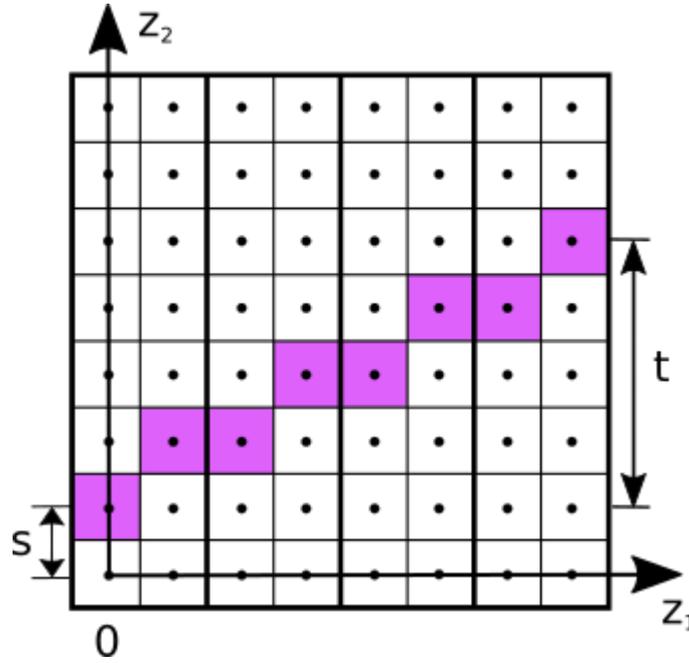


Рисунок 1.1 — Диадический паттерн  $\langle 1, 4 \rangle$  в st-параметризации системы  $Y_d^{2^3, 2}$ .

Множество всех систем паттернов на изображении  $I_{\langle n \rangle}^m$  будем обозначать  $\mathbb{Y}_{\langle n \rangle}^m$ .

Основной системой БПХ на двумерном изображении  $I_{2^k}^2$  будем называть систему  $Y_d^{2^k, 2} \stackrel{\text{def}}{=} \bigcup_{t=0}^{2^k-1} T_{2^k}^2(P_t^k, \langle 0, 1 \rangle)$ , элемент которой будем называть диадическим паттерном.

Суммой по паттерну будем называть сумму значений пикселей в позициях, принадлежащих паттерну:  $\sum_p I_{\langle n \rangle}^m = \sum_{z \in p} I_{\langle n \rangle}^m(z)$ .

**Определение 6.** Хаф-образом изображения  $I_{\langle n \rangle}^m$  для системы паттернов  $Y \in \mathbb{Y}_{\langle n \rangle}^m$  будем называть изображение  $J_{\langle n' \rangle}^{m'} = \sum_{Q(z)} I_{\langle n \rangle}^m$ , где  $Q : \mathbb{Z}_{\langle n' \rangle}^{m'} \rightarrow Y$  — сюръекция,  $z \in \mathbb{Z}_{\langle n' \rangle}^{m'}$ . Сюръекцию  $Q$  будем называть параметризацией системы  $Y$  для Хаф-образа  $J_{\langle n' \rangle}^{m'}$ .

St-параметризацией системы  $Y_d^{2^k, 2}$  для Хаф-образа  $J_{2^k}^2$  будем называть сюръекцию  $\langle s, t \rangle \mapsto P_t^k \nearrow \langle 0, s \rangle$ . Пример паттерна, задаваемого этой параметризацией, приведен на рис. 1.1.

В 1992 году Брейди опубликовал алгоритм БПХ, вычисляющий Хаф-образ с st-параметризацией для системы  $Y_d^{2^k, 2}$  с вычислительной сложностью  $\Theta(n^2 \log n)$ .

Группа симметрий множества всех позиций изображения  $I_{2^k}^m$  — это гипероктаэдрическая группа  $B_m$  (группа ортогональных преобразований в  $\mathbb{R}^m$

с целочисленными матрицами), также называемая группой симметрий гиперкуба.

**Определение 7.** Симметрическим расширением множества паттернов  $P$  на изображении  $I_n^m$  будем называть объединение всех множеств паттернов, симметричных  $P$ :  $Sym(P) \stackrel{def}{=} \bigcup_{\pi \in B_m, p \in P} \pi(p)$ .

Симметрическое расширение  $Sym(Y_d^{2^k, 2})$  будем называть системой БПХ на плоскости. Поскольку  $|B_2| = 8$ , то  $|Sym(Y_d^{2^k, 2})| \leq 8|Y_d^{2^k, 2}|$ .

**Утверждение 2.**  $|Sym(Y_d^{2^k, 2})| \leq 4|Y_d^{2^k, 2}|$ .

Утверждение 2 обосновывается симметричностью паттернов (смотри утверждение 3), задаваемых правилом (1.1). При отражении по обеим координатным осям любой порождающий паттерн из  $Y_d^{2^k, 2}$  переходит в отличающийся только сдвигом, а следовательно, порожденный им пучок переходит в себя. Подмножество симметрий, которые не переводят пучок в себя, обозначим  $B_2^s$ .

**Определение 8.** Будем называть систему  $k$ -полной, если для любого множества позиций из  $k$  элементов на изображении в системе найдется паттерн, включающий это множество.

Любая система паттернов 1-полна, поскольку любой пучок паттернов покрывает область определения изображения.

**Теорема 1.** Система БПХ на плоскости 2-полна.

*Доказательство.* Рассмотрим произвольный паттерн  $p_2 = \{z', z''\}$  размера 2 на изображении  $I_{2^k}^2$ . Координаты позиций паттерна обозначим следующим образом:  $z' = \langle z'_1, z'_2 \rangle$  и  $z'' = \langle z''_1, z''_2 \rangle$ , а разности их координат –  $\Delta z_1 = z''_1 - z'_1$ ,  $\Delta z_2 = z''_2 - z'_2$ . Поскольку  $p_2$  – множество, будем считать без потери общности, что  $\Delta z_1 \geq 0$ , причем  $\Delta z_1 = 0 \implies \Delta z_2 \geq 0$ . Тогда среди всех  $p_2$  можно выделить четыре подмножества, объединение которых включает любые  $p_2$ :

1.  $\Delta z_2 \leq 0, |\Delta z_2| \leq |\Delta z_1|$ ;
2.  $\Delta z_2 \geq 0, |\Delta z_2| \leq |\Delta z_1|$ ;
3.  $\Delta z_2 \leq 0, |\Delta z_2| \geq |\Delta z_1|$ ;
4.  $\Delta z_2 \geq 0, |\Delta z_2| \geq |\Delta z_1|$ .

Рассмотрим паттерн второго типа.

Будем называть первую координату позиции абсциссой, вторую – ординатой. Назовём множество пикселей паттерна  $p \in Y_d^{2^k, 2}$ , абсциссы которых лежат в диапазоне  $[z'_1, z''_1]$  подпаттерном. Будем считать, что  $t_p$  – наклон подпаттерна – равен разнице ординат пикселей паттерна, с абсциссами  $z'_1$  и  $z''_1$ .

Покажем, что если  $z''_1 - z'_1 \leq n - 1$ , то  $\exists p : z', z'' \in p$ . Если  $z''_1 - z'_1 = n - 1$ , то  $\exists p : z', z'' \in p$  по определению (1.1), такой паттерн называется диагональным. Особенностью диагонального паттерна является, что любой его подпаттерн является диагональным, аналогичное верно и для прямого паттерна ( $t = 0$ ).

Рассмотрим зависимость наклона подпаттерна от наклона паттерна  $t_p(t)$ : если  $\Delta t_p(t) = t_p(t + 1) - t_p(t)$  принимает только значения 0 или 1, то  $t_p$  будет пробегать все значения в диапазоне  $[0, z''_1 - z'_1]$  (в силу существования граничных паттернов – прямого и диагонального), а значит, для любой пары пикселей  $z', z''$  найдется паттерн, проходящий через них.

Докажем, что  $\Delta t_p(t)$  принимает значения только 0 или 1 методом математической индукции: для диадического паттерна размером  $n = 2^1$  утверждение верно, тогда пусть для  $n = 2^k$  оно также верно, докажем его для  $n = 2^{k+1}$ .

Рассмотрим два случая приращения наклона  $t$  на единицу: четный и нечетный. Пусть  $t$  – четный, в этом случае паттерн с наклоном  $t + 1$  отличается от предыдущего тем, что ординаты всех пикселей его правой половины увеличены на единицу. Следовательно, для любого его подпаттерна  $0 \leq \Delta t_p(t) \leq 1$ , что следует из определения (1.1).

Если  $t$  – нечетный, при увеличении наклона на единицу  $t$  становится четным, и, как следует из определения (1.1), сдвиг в центре паттерна исчезает, однако прибавляется по одной единице сдвига на каждую половину паттерна (в силу симметрии). Поскольку из предположения индукции для обеих половин длинны  $n = 2^k$  изменение наклона каждого подпаттерна удовлетворяет ограничению  $0 \leq \Delta t_p(t) \leq 1$ , и эти половины не сдвинуты друг относительно друга, то и в этом случае любой подпаттерн всего паттерна удовлетворяет данному ограничению.

Таким образом, для пары пикселей первого типа, величина  $t_p(t)$  пробегает все значения в диапазоне  $[0, z''_1 - z'_1]$ , а, следовательно поскольку  $z''_2 - z'_2 \leq z''_1 - z'_1$ , то  $\exists p : z', z'' \in p$ .

Аналогично для симметрического расширения паттернов основной системы БПХ и паттернов  $p_2$  соответствующего типа.  $\square$

Каждому порождающему диадическому паттерну ставится в соответствие корреспондентная прямая  $z_2 = az_1$ , где  $a = \frac{t}{2^k - 1}$ .

**Определение 9.** *Ортотропный разброс (отклонение от корреспондентной прямой) порождающего диадического паттерна  $p \in Y_d^{2^k, 2}$  с наклоном  $t$  выражается следующим образом:*

$$E_{t,k}^p(\hat{z}_1) = \hat{z}_2 - a\hat{z}_1, \quad (1.3)$$

где  $\langle \hat{z}_1, \hat{z}_2 \rangle \in p$

**Утверждение 3.** *Для любого  $y \in [0, 2^{k-1} - 1]$  и  $t \in [0, 2^k - 1]$  верно, что  $E_{t,n}^p(2^{k-1} - 1 - y) = -E_{t,n}^p(2^{k-1} + y)$ .*

*Доказательство.* Перепишем утверждение согласно формуле 9:

$$\frac{(2^{k-1} - 1 - y)2^r}{2^k - 1} - \left\lceil \frac{(2^{k-1} - 1 - y)2^r}{2^k - 1} \right\rceil = \left\lceil \frac{(2^{k-1} + y)2^r}{2^k - 1} \right\rceil - \frac{(2^{k-1} + y)2^r}{2^k - 1}.$$

Преобразуем данное выражение:

$$\left\lceil \frac{(2^{k-1} + y)2^r}{2^k - 1} \right\rceil + \left\lceil \frac{(2^{k-1} - 1 - y)2^r}{2^k - 1} \right\rceil = 2^r.$$

Покажем, что данное уравнение есть тождество:

$$\left\lceil \frac{2^{k-1}2^r}{2^k - 1} + \frac{2^r y}{2^k - 1} \right\rceil + \left\lceil \frac{2^{k-1}2^r}{2^k - 1} - \frac{(y + 1)2^r}{2^k - 1} \right\rceil = 2^r,$$

$$\left\lceil 2^{r-1} + \frac{2^{r-1}}{2^k - 1} + \frac{2^r y}{2^k - 1} \right\rceil + \left\lceil 2^{r-1} + \frac{2^{r-1}}{2^k - 1} - \frac{(y + 1)2^r}{2^k - 1} \right\rceil = 2^r.$$

В результате получим:

$$\left\lceil \frac{2^{r-1} + 2^r y}{2^k - 1} \right\rceil + \left\lceil \frac{-2^{r-1} - 2^r y}{2^k - 1} \right\rceil = 0,$$

что есть тождественно верное уравнение, что и требовалось доказать.  $\square$

**Определение 10.** *Максимальный ортотропный разброс (отклонение от корреспондентной прямой) порождающего диадического паттерна  $p \in Y_d^{2^k, 2}$  с наклоном  $t$  выражается следующим образом:*

$$E_{t,k} = \max_{z \in p} \left( \left| \hat{z}_2 - a\hat{z}_1 \right| \right), \quad (1.4)$$

где  $\langle \hat{z}_1, \hat{z}_2 \rangle \in p$

Брейди показал, что максимальный ортотропный разброс (10) не превышает величины  $\frac{k}{2}$ . Позднее С. Карпенко и Е. Ершов показали, что его точная верхняя грань для четных  $k$  составляет  $\frac{k}{6}$ , для нечетных  $k$  составляет  $\frac{k}{6} - \frac{1}{18}$ . Максимальный разброс достигается при наклонах  $t_1 = \frac{n-1}{3}$ ,  $t_2 = \frac{2(n-1)}{3}$ .

Предлагается новый способ количественной характеристики разброса диадического паттерна.

**Определение 11.** *Максимальный ортогональный разброс (отклонение от корреспондентной прямой) порождающего диадического паттерна  $p \in Y_d^{2^k, 2}$  с наклоном  $t$  – это наибольшее среди ортогональных расстояний от  $\hat{z} \in p$  до корреспондентной прямой:*

$$N_{t,k} = \max_{\hat{z} \in p} \left( \frac{|a\hat{z}_1 - \hat{z}_2|}{\sqrt{a^2 + 1}} \right). \quad (1.5)$$

Из определений 10 и 11 следует, что максимальные ортогональный и ортотропный разброс порождающего диадического паттерна связаны выражением

$$N_{t,k} = \frac{E_{t,k}}{\sqrt{1 + a^2}}. \quad (1.6)$$

Утверждение следует из теоремы косинусов, поскольку для фиксированного диадического паттерна максимальные ортогональная и ортотропная ошибки формируют треугольник, в котором первая – катет, а вторая – гипотенуза.

Согласно формуле (1.6) максимальный ортогональный разброс ограничен сверху ортотропным, то есть не превышает  $\frac{k}{6}$ . Максимальный ортогональный разброс ограничен снизу величиной  $\frac{k}{2\sqrt{10}}$ . Это нетрудно посчитать, подставив в формулу (1.6) наклон  $t = 1/3$  (наклон, на котором достигается наибольшая ортотропная ошибка) для наибольшего ортотропного разброса:

$$N_{t,k} = \frac{k}{6} \cdot \frac{1}{\sqrt{1 + \frac{1}{3^2}}} = \frac{k}{2\sqrt{10}}. \quad (1.7)$$

## 1.2.2 Алгоритм быстрого преобразования Хафа для двумерного случая

Перейдем непосредственно к описанию алгоритма БПХ. Входом алгоритма является двумерное изображение, а выходом – соразмерный двумерный Хаф-образ, в каждом пикселе которого содержится значение суммы по соответствующему координатам этого пикселя диадическому паттерну. Код функции *fht2* для вычисления БПХ представлен ниже на языке MATLAB.

Листинг 1.1 — Алгоритм вычисления быстрого преобразования Хафа, записанный на языке MATLAB.

```

function h = fht2(I)
    n = size(I, 2);
    if n < 2
        h = m;
5     return
    end
    h = mergeHT(fht2(I(:, 1 : n / 2, :)), fht2(I(:, (n / 2 + 1) :
        n, :)));
end

10 function h = mergeHT(h0, h1)
    [m, n0, o] = size(h0);
    n = 2 * n0;
    h = zeros(m, n, o);
    r = (n0 - 1) / (n - 1);
15 for i = 1 : n
        t = i - 1;
        t0 = round(t * r);
        s = t - t1;
        h(:, i, :) = h0(:, t0 + 1, :) + [h1(s + 1 : m, t0 + 1, :);
        h1(1 : s, t0 + 1, :)];
20 end
end

```

Входное изображение  $I_n^2$  рекурсивно разбивается на два равных изображения вертикальной прямой, пока ширина его не будет равна 2. Затем каждый столбец суммируется с парным поэлементно, после чего правый столбец циклически сдвигается вниз на один пиксель и выполняется второе поэлементное

суммирование. В результате вычисляются суммы по паттернам длины 2 с наклонами:  $t = 0$  и  $t = 1$ . Полученные суммы используются на следующем шаге рекурсии для вычисления сумм по паттернам высоты 4 в каждой четверке строк. Так, чтобы вычислить все суммы по паттернам длины 4 и наклоном  $t = 0$ , достаточно сложить суммы по паттернам длины 2 у соседних пар строк. Процедура продолжается, пока ширина результирующего изображения не будет составлять  $n$ .

В своей работе П. Брейди показал верность следующего утверждения.

**Утверждение 4.** Алгоритм БПХ вычисляет суммы по всем паттернам основной системы БПХ изображения  $I_{\langle n', n \rangle}$ , где  $n = 2^k$ , причем  $n' = 2^a \cdot n$ , где  $a \in \mathbb{N}$  – произвольно.

Оценим вычислительную сложность БПХ. Сумма по всем паттернам высоты 2 для пары строк требует  $2n$  операций, и таких пар –  $n/2$ , для следующего уровня рекурсии требуется  $4n$  операций, но четверок на изображении в два раза меньше, чем пар –  $n/4$ , продолжая, получим следующую сумму  $2n \cdot \frac{n}{2} + 4n \cdot \frac{n}{4} + \dots + n^2$ , в силу того, что таких слагаемых  $\log_2 n$  (по числу уровней рекурсии), то получаем оценку сложности  $n^2 \log n$ . Соответственно, для всех типов прямых сложность будет составлять  $4n^2 \log_2 n$ . Легко проиллюстрировать за счет чего достигается ускорения с использованием рисунка 1.2. На основе принципа динамического программирования эксплуатируя структуру диадического паттерна удаётся увеличить количество переиспользований промежуточных вычислений: одна сумма по подпаттерну используется большее количество раз.

Количество необходимой памяти для посчитанного Хаф-образа составляет  $\Theta(n^2)$ , в силу линейной зависимости диапазона изменений  $s$  и  $t$  от  $n$ . Следовательно, вычислительная сложность в пересчете на одну прямую (пиксель Хаф-образа) составляет  $\Theta(\log n)$ .

Вычисление БПХ может быть ускорено за счет параллелизации вычислений. Поскольку на каждом шаге алгоритма выполняется в точности  $n^2$  сложений, то с использованием  $n^2$  процессоров данные вычисления можно провести за константное время. А так как глубина рекурсии равна  $\log_2 n$ , то время вычисления может быть в этом случае оценено как  $\Theta(\log_2 n)$  при том же количестве используемой памяти.

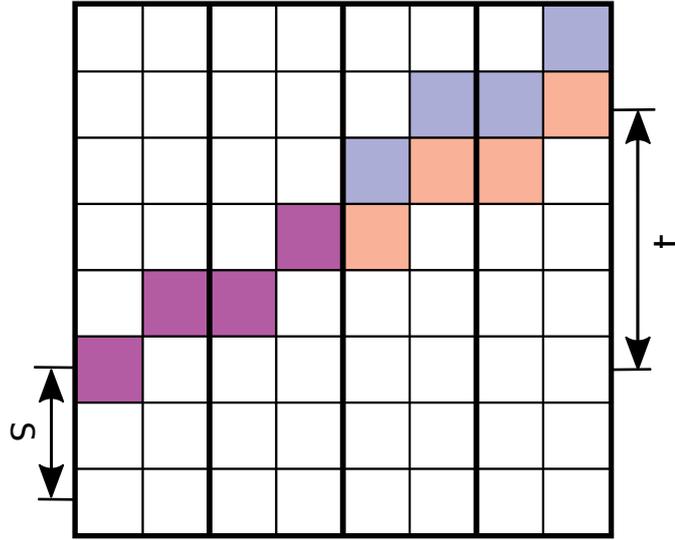


Рисунок 1.2 — Два диадических паттерна с наклонами, отличающимися на единицу.

### 1.3 Диадическая плоскость и трехмерное быстрое преобразование Хафа для плоскостей

В трехмерном случае термин «диадический паттерн» может трактоваться двояко: с одной стороны как аппроксимация прямой, с другой – плоскости. Во избежание путаницы, в первом случае будем использовать термин «диадическая прямая», а во втором – «диадическая плоскость».

#### 1.3.1 Свойства диадических плоскостей

**Определение 12.** Порождающими диадическими плоскостями  $DP_{(t_1, t_2)}^{(e)}$  на изображениях  $I_{\langle e \rangle}^3$ , где  $\langle e \rangle = \langle 2^k, 2^k, 2^{k+1} \rangle$ , будем называть паттерны, являющиеся графиками функций следующего вида:

$$z_3 = D_{t_1}^k(z_1) + D_{t_2}^k(z_2), \quad (1.8)$$

где  $t_1, t_2 \in \mathbb{Z}_{2^k}$  – наклоны плоскости.

Основной системой диадических плоскостей на трехмерном изображении  $I_{\langle e \rangle}^3$  будем называть систему  $Y_{dp}^{\langle e \rangle, 3} \stackrel{\text{def}}{=} \bigcup_{t_1=0}^{2^k-1} \bigcup_{t_2=0}^{2^k-1} T_{\langle e \rangle}^3(DP_{(t_1, t_2)}^{(e)}, \langle 0, 0, 1 \rangle)$ . St-пара-

метризацией системы  $Y_{dp}^{(e),3}$  для Хаф-образа  $J_{(e)}^3$  будем называть сюръекцию  $\langle s, t_1, t_2 \rangle \mapsto DP_{t_1, t_2}^{(e)} \nearrow \langle 0, 0, s \rangle$ . Диадической плоскостью будем называть элемент множества  $Sym(Y_{dp}^{(e),3})$ .

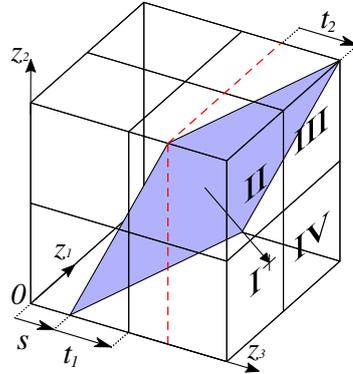


Рисунок 1.3 — Иллюстрация  $(s, t)$ -параметризации плоскости в трехмерном пространстве.

**Утверждение 5.**  $|Sym(Y_{dp}^{(e),3})| \leq 12|Y_{dp}^{(e),3}|$ .

Поскольку  $|B_3| = 48$ , то  $|Sym(Y_{dp}^{(e),3})| \leq 48|Y_{dp}^{(e),3}|$ . В  $B_3$  существует две конечные подгруппы отражений порядка 2, переводящие пучок в себя: первая — отражение оси  $z_3$ , вторая — отражение сразу по трём координатным плоскостям, следовательно  $|Sym(Y_{dp}^{(e),3})| \leq 12|Y_{dp}^{(e),3}|$ . Множество симметрий, которые не переводят пучок в себя, обозначим  $B_3^s$ .

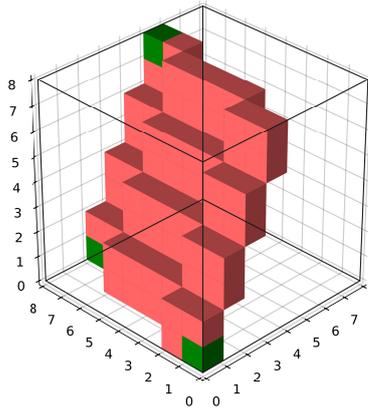


Рисунок 1.4 — Пример структуры диадической плоскости. Зеленым цветом отмечены пиксели, задающие плоскость.

Ортотропная ошибка аппроксимации диадической плоскости вычисляется как сумма ошибок её образующих, являющихся диадическими паттернами (что следует из (1.8)), поскольку для них эта ошибка измеряется в направлении общей оси  $x$ . Таким образом, для равностороннего изображения с линейным

размером  $n$  в худшем случае (четное  $m$ ) диадическая плоскость с параметрами  $s = 0, t_1 = \frac{n-1}{3}, t_2 = \frac{n-1}{3}$  имеет максимальную ортотропную ошибку, равную  $\frac{\log_2 n}{3}$ .

Для порождающей диадической плоскости вводится корреспондентная плоскость, которая описывается уравнением  $z_3 = a_1 z_1 + a_2 z_2$ , где  $a_1 = \frac{t_1}{n-1}$ ,  $a_2 = \frac{t_2}{n-1}$ .

**Определение 13.** *Максимальный ортотропный разброс (отклонение от корреспондентной плоскости) порождающей диадической плоскости  $p \in Y_{dp}^{(e),3}$  с наклонами  $\langle t_1, t_2 \rangle$  выражается следующим образом:*

$$E_{t_1, t_2, k}^{dp} = \max_{\hat{z} \in p} \left( \left| \hat{z}_3 - a_1 \hat{z}_1 - a_2 \hat{z}_2 \right| \right), \quad (1.9)$$

где  $\hat{z} = \langle \hat{z}_1, \hat{z}_2, \hat{z}_3 \rangle \in p$ .

Максимальный ортотропный разброс  $E_{t_1, t_2, k}^{dp}$  складывается из разбросов образующих – диадических паттернов и ограничен сверху  $\frac{k}{3}$ . Поскольку образующие основной диадической плоскости, диадические паттерны, лежат в плоскостях  $0z_1z_3, 0z_2z_3$ , их разбросы складываются.

**Определение 14.** *Максимальный ортогональный разброс (отклонение от корреспондентной плоскости)  $N_{t_1, t_2, k}^{dp}$  порождающей диадической плоскости  $p \in Y_{dp}^{(e),3}$  с наклонами  $\langle t_1, t_2 \rangle$  – это наибольшее среди ортогональных расстояний от точек  $z \in p$  до корреспондентной плоскости.*

**Утверждение 6.** *Максимальный ортогональный  $N_{t_1, t_2, k}^{dp}$  и ортотропный  $E_{t_1, t_2, k}^{dp}$  разбросы основных диадических плоскостей связаны следующим выражением:*

$$N_{t_1, t_2, k}^{dp} = \frac{E_{t_1, t_2, k}^{dp}}{\sqrt{1 + a_1^2 + a_2^2}}. \quad (1.10)$$

Утверждение следует из теоремы косинусов, поскольку для фиксированной диадической плоскости максимальные ортогональная и ортотропная ошибки формируют треугольник, в котором первая – катет, а вторая – гипотенуза.

Из утверждения непосредственно следует, что максимальный ортогональный разброс  $N_{t_1, t_2, k}^{dp}$  ограничен сверху величиной  $\frac{k}{3}$ . Аналогичным двумерному

случаю образом оценим эту величину снизу для наклонов  $\langle \frac{2^k-1}{3}, \frac{2^k-1}{3} \rangle$ :

$$N_{t_1, t_2, k}^{dp} = \frac{E_{t_1, t_2, k}^{dp}}{\sqrt{1 + \frac{1}{9} + \frac{1}{9}}} = \frac{k}{\sqrt{11}}. \quad (1.11)$$

### 1.3.2 Алгоритм трехмерного быстрого преобразования Хафа для плоскостей

Впервые алгоритм был предложен в работе [5] в 2016 году. В данной главе предлагается более подробное описание, в том числе и его свойств. Алгоритм вычисления БПХ для плоскостей для одного из 12 подтипов приведен на языке MATLAB ниже (смотри алгоритм 1.2). Входом алгоритма является трехмерное изображение, а выходом соответствующий Хаф-образ.

```

function h = fht3plane(m)
    sz = size(m);
    n = sz(1);

5   hs = zeros(sz);
    for i = 1:n
        hs(:, :, i, :) = fht2(m(:, :, i, :));
    end

10   h = zeros(sz);
    for i = 1:n
        h(:, i, :, :) = fht2(hs(:, i, :, :));
    end
end

```

Листинг 1.2 — Алгоритм трехмерного быстрого преобразования Хафа.

Предложен новый алгоритм трехмерного БПХ (ТБПХ) для плоскостей. Входом алгоритма является трехмерное изображение, а выходом – трехмерный Хаф-образ. Алгоритм описывается следующим образом:

1. Формируется трёхмерное изображение  $HS$  с координатными осями  $z'_1, z'_2, z'_3$  так, что для всех  $c \in [0, n - 1]$  в каждое двумерное изображение  $z'_2 = c$  записывается Хаф-образ слоя исходного изображения  $z_2 = c$ , причем  $z'_1$  – наклон паттерна, а  $z'_3$  – сдвиг.

2. Формируется трёхмерный Хаф-образ с координатными осями  $t_1, t_2, s$  так, что для всех  $c \in [0, n - 1]$  в каждое двумерное изображение  $t_1 = c$  записывается Хаф-образ слоя  $z'_1 = c$  изображения  $HS$ , причем  $t_2$  – наклон паттерна, а  $s$  – сдвиг.

Показано, что число сложений для выполнения каждого этапа составляет  $2n^3 \log_2 n$ , соответственно суммарное число сложений равно  $4n^3 \log_2 n$ .

**Утверждение 7.** Алгоритм 1.2 вычисляет Хаф-образ в  $st$ -параметризации системы  $Y_{dp}^{n,3}$  изображения  $I_{\langle e \rangle}^3$ , где  $\langle e \rangle = \langle 2^k, 2^k, 2^{k+1} \rangle$ .

*Доказательство.* Рассмотрим пиксель Хаф-образа с координатами  $(s, t_1, t_2)$ . Согласно описанному алгоритму, в этом пикселе содержится сумма по диадическому паттерну в изображении  $HS$ , лежащему в плоскости  $z'_1 = t_1$  с наклоном  $t_2$  и сдвигом  $s$ , то есть по паттерну  $z'_3 = D_{t_2}^k(z'_2) + s$ .

Каждый пиксель  $HS$  данного паттерна содержит сумму пикселей исходного изображения  $I$ , принадлежащих диадическому паттерну с наклоном  $t_1$  и сдвигом  $s + z_3$ , то есть по паттерну  $z_3 = D_{t_1}^k(z_1) + s + z'_3$ .

Совмещая уравнения паттернов для разных систем координат, получим уравнение диадической плоскости:

$$z_3 = D_{t_1}^k(z_1) + D_{t_2}^k(z_2) + s,$$

что и требовалось доказать. □

Из утверждения 5 следует, что трудоёмкость вычисления Хаф-образа  $Sym(Y_{dp}^{\langle e \rangle, 3})$  составляет  $48n^3 \log_2 n$  операции сложения, однако её можно уменьшить за счет повторного использования результатов первого этапа алгоритма до  $36n^3 \log_2 n$  операций.

Нетрудно оценить вычислительную сложность предложенного алгоритма: для  $2n$  плоскостей выполнялся расчет БПХ, сложность которого составляет  $2n^2 \log_2 n$ . Таким образом, для всех подтипов диадических плоскостей, результирующая вычислительная сложность алгоритма последовательного подсчета БПХ составляет  $48n^3 \log_2 n$ . На рисунке 1.5, показано как изменяется время вычисления сумм при переходе от ТБПХ для плоскостей к простому полному подсчету сумм для всех диадических плоскостей. Эксперимент проводился на процессоре Intel i7-4790, реализация алгоритмов выполнена на языке MATLAB в среде Octave 4.3.

Для случая параллелизации вычислений результирующая сложность равна  $\Theta(\log_2 n)$ , но с использованием уже  $2n^3$  последовательных вычислителей: алгоритм плоского БПХ вычисляется за время  $\Theta(\log_2 n)$  с использованием  $\Theta(n^2)$  последовательных вычислителей, а число независимых наборов таких вычислений составляет  $\Theta(n)$ .

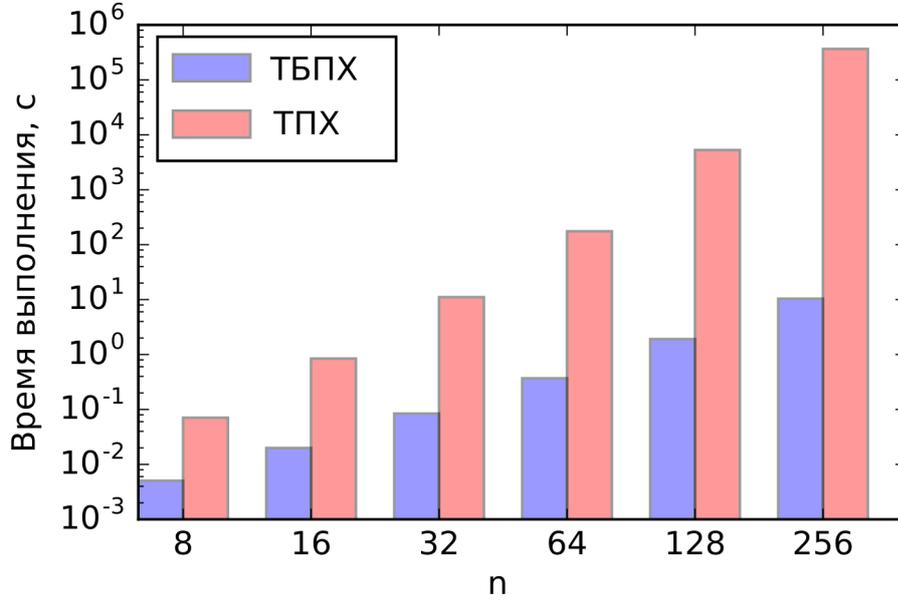


Рисунок 1.5 — Сравнения времени вычисления ТБПХ для плоскостей и алгоритма полного подсчета сумм по диадическим плоскостям для изображений разного размера. Показания приведены в логарифмической шкале.

Асимптотическая оценка величины Хаф-образа составляет  $\Theta(n^3)$ , что определяется размерностью линейной пропорциональностью каждого элемента вектора параметров  $(s, t_1, t_2)$  величине стороне изображения  $n$ .

Интересно заметить, что алгоритм достаточно прямолинейно обобщаются на случай любой размерности. Для входного  $m$ -мерного изображения послойные двумерные БПХ вычисляются поочередно, таким образом, сложность такого алгоритма будет составлять  $\Theta(n^m \log n)$ .

## 1.4 Диадическая прямая и трехмерное быстрое преобразование Хафа для прямых

В данном разделе предлагается новый алгоритм быстрого трехмерного преобразования Хафа для прямых, соответствующий преобразованию Йона для непрерывного случая [70]. Приведена асимптотическая оценка вычислительной сложности и показана, что она неуллучшаема.

### 1.4.1 Свойства диадических прямых

**Определение 15.** Порождающими диадическими прямыми  $DL_{t_1, t_2}^k$  на изображениях  $I_{2^k}^3$  будем называть паттерны, задаваемые системой уравнений:

$$\begin{cases} z_1 = D_{t_1}^k(z_3), \\ z_2 = D_{t_2}^k(z_3), \end{cases} \quad (1.12)$$

где  $z_3 \in \mathbb{Z}_n$ , а  $t_1, t_2 \in \mathbb{Z}_{2^k}$  – наклоны прямой.

Основной системой диадических прямых на трехмерном изображении  $I_n^3$  будем называть систему  $Y_{dl}^{n,3} \stackrel{\text{def}}{=} \bigcup_{t_1=0}^{n-1} \bigcup_{t_2=0}^{n-1} T_n^3(DL_{(t_1, t_2)}^n, \langle 1, 1, 0 \rangle)$ . St-параметризацией системы  $Y_{dl}^{n,3}$  для Хаф-образа  $J_{\langle n \rangle}^4$  будем называть сюръекцию  $\langle s_1, s_2, t_1, t_2 \rangle \mapsto DL_{t_1, t_2}^n \nearrow \langle s_1, s_1, 0 \rangle$  (смотри рисунок 1.6). Полной системой диадических прямых будем называть  $Sym(Y_{dl}^{n,3})$ . Диадической прямой будем называть элемент множества  $Sym(Y_{dl}^{(e),3})$  (смотри рисунок 1.6).

**Утверждение 8.**  $|Sym(Y_{dl}^{n,3})| \leq 12|Y_{dl}^{n,3}|$ .

Утверждение 8 доказывается аналогично 5, поскольку у порождающих диадических прямых такое же множество  $B_3^s$ .

**Теорема 2.** Полная система диадических прямых 2-полна.

Теорема 2 доказывается от противного: если утверждение ложно, то одна из проекций прямой на координатные плоскости  $0z_1z_3$  и  $0z_2z_3$  не будет

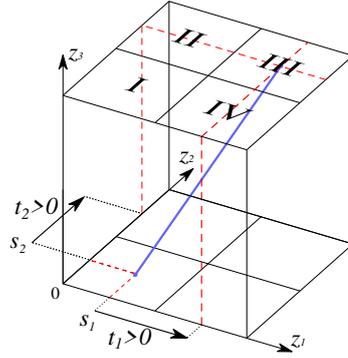


Рисунок 1.6 — Иллюстрация параметризации преимущественно-параллельной оси  $z$  прямой в кубе.

проходить через соответствующие проекции двух пикселей, что противоречит теореме 1.

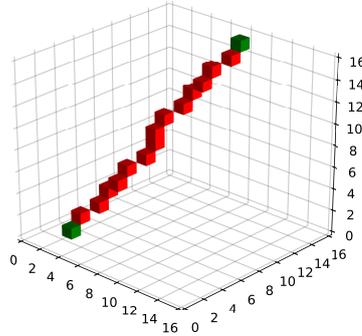


Рисунок 1.7 — Иллюстрация структуры диадической прямой в кубе. Зеленым отмечены воксели, задающие диадическую прямую.

Для порождающей диадической прямой вводится корреспондентная прямая, проходящая через точку  $(0,0,0)$  в направлении  $(t_1, t_2, n - 1)$ .

**Определение 16.** Максимальный ортотропный разброс (отклонение от корреспондентной прямой) порождающей диадической прямой  $p \in Y_{dl}^{(e),3}$  с наклонами  $\langle t_1, t_2 \rangle$  выражается следующим образом:

$$E_{t_1, t_2, k}^{dl} = \max_{\hat{z} \in p} \sqrt{(\hat{z}_1 - a_1 \hat{z}_3)^2 + (\hat{z}_2 - a_2 \hat{z}_3)^2}. \quad (1.13)$$

Максимальный ортотропный разброс  $E_{t_1, t_2, k}^{dl}$  ограничен сверху величиной  $\frac{k\sqrt{2}}{6}$ , это следует из его определения и теоремы Пифагора. Разбросы проекций прямой ортогональны, следовательно при их сложении суммарная ошибка складывается по теореме Пифагора:  $\sqrt{\left(\frac{k}{6}\right)^2 + \left(\frac{k}{6}\right)^2} = \frac{k\sqrt{2}}{6}$ .

**Определение 17.** Максимальный ортогональный разброс (отклонение от корреспондентной прямой)  $N_{t_1, t_2, k}^{dl}$  порождающей диадической прямой  $p \in Y_{dl}^{n,3}$  с наклонами  $\langle t_1, t_2 \rangle$  – это наибольшее среди ортогональных расстояний от точек  $z \in p$  до корреспондентной прямой.

**Утверждение 9.** Максимальный ортогональный  $N_{t_1, t_2, k}^{dl}$  и ортотропный  $E_{t_1, t_2, k}^{dl}$  разброс основных диадических прямых связаны следующим образом:

$$N_{t_1, t_2, k}^{dl} = \frac{E_{t_1, t_2, k}^{dl}}{\sqrt{1 + a_1^2 + a_2^2}}. \quad (1.14)$$

Утверждение следует из теоремы косинусов, поскольку для фиксированной диадической прямой максимальные ортогональная и ортотропная ошибки формируют треугольник, в котором первая – катет, а вторая – гипотенуза.

Из 1.14 следует, что максимальная ортогональная ошибка ограничена сверху максимальной ортотропной, а следовательно величиной  $\frac{k\sqrt{2}}{6}$ .

Аналогичным трехмерному случаю для плоскостей, оценим эту величину снизу, для наклонов  $\langle \frac{2^k-1}{3}, \frac{2^k-1}{3} \rangle$ :

$$N_{t_1, t_2, k}^{dl} = \frac{E_{t_1, t_2, k}^{dl}}{\sqrt{1 + \frac{1}{9} + \frac{1}{9}}} = \frac{k}{\sqrt{22}}. \quad (1.15)$$

#### 1.4.2 Алгоритм трехмерного быстрого преобразования Хафа для прямых

Рассмотрим алгоритм вычисления трехмерного быстрого преобразования Хафа для прямых. Код данного алгоритма приведен ниже на языке MATLAB (смотри алгоритм 1.3). Основной функцией алгоритма является *fht3line*, входом которой является четырехмерное многоканальное изображение  $m$ , а выходом – Хаф-образ для него.

```

function h = fht3line(m)
    n = size(m, 3);
    if n < 2
        h = m;
    return
end
5

```

```

n0 = round(n / 2);
h = mergeHT(fht3line(m(:, :, 1 : n0, :, :)), fht3line(m(:, :,
n0 + 1 : n, :, :)));
10 end

function h = mergeHT(h0, h1)
m = size(h0, 1);
n0 = size(h0, 3);
15 o = size(h0, 5);

n = 2 * n0;
h = zeros(m, m, n, n, o);
r = (n0 - 1) / (n - 1);
20

for i = 1 : n
for j = 1 : n
t = [i - 1, j - 1];
t0 = round(t .* r);
25 s = t - t0;

h(:, :, i, j, :) = h0(:, :, t0(1) + 1, t0(2) + 1, :) + ...
[[
h1(s(1) + 1 : m, s(2) + 1 : m, t0(1)
+ 1, t0(2) + 1, :), ...
30 h1(s(1) + 1 : m, 1 : s(2), t0(1) +
1, t0(2) + 1, :)
]];
[
h1(1 : s(1), s(2) + 1 : m, t0(1) +
1, t0(2) + 1, :), ...
h1(1 : s(1), 1 : s(2), t0(1) + 1, t0
(2) + 1, :)
35 ]];

end
end
end

```

Листинг 1.3 — Алгоритм быстрого преобразования Хафа для одного из 12 подтипов прямых.

В диссертации предложен новый алгоритм ТБПХ для прямых. Входом является трехмерное изображение, записанное в четырехмерное при  $z_4 = 1$ , а выходом – Хаф-образ для него. Сначала изображение разбивается на па-

ры соседних слоёв толщины 2 по  $z_3$ . Затем, для каждой пары выполняется поэлементное суммирование для всех конфигураций их взаимных сдвигов (паттернов высоты 2), что требует  $4n^2 \cdot n/2 = 2n^3$  сложений. На следующем шаге суммирование повторяется для паттернов высоты четыре, но уже с использованием посчитанных на предыдущем шаге сумм –  $4n^3$  сложений, на третьем шаге –  $16n^3$ , число шагов составляет  $\log_2 n$ . Таким образом, трудоёмкость алгоритма быстрого трехмерного преобразования Хафа составляет  $2n^3 + 4n^3 + 16n^3 + \dots + n^4 = 2n^3(n - 1)$  операций. То есть вычислительная сложность алгоритма равна  $\Theta(n^4)$  для симметрического расширения системы диадических прямых.

**Утверждение 10.** Алгоритм 1.3 вычисляет Хаф-образ изображения  $I_n^3$  в  $st$ -параметризации системы  $Y_{dl}^{n,3}$ , где  $n = 2^k$ ,  $k \in \mathbb{Z}$ .

*Доказательство.* Пусть на вход функции *mergeHT* алгоритма 1.3 поданы изображения  $B_{\langle n,n,n',n' \rangle}^4$  и  $T_{\langle n,n,n',n' \rangle}^4$  (массивы  $h0$  и  $h1$  в теле функции), где  $n' = 2^{k'}$ ,  $k' < k$ ,  $k' \in \mathbb{Z}$ . Тогда функция вернет изображение  $H_{\langle n,n,2n',2n' \rangle}^4$  (массив  $h$ ). Рассмотрим оператор на строках 27 – 35 алгоритма 1.3. В нем выходное изображение заполняется следующим образом:

$$H(s_1, s_2, t_1, t_2) = B\left(s_1, s_2, \left\lfloor \frac{t_1}{2} \right\rfloor, \left\lfloor \frac{t_2}{2} \right\rfloor\right) + \\ + T\left(s_1 + \left\lfloor \frac{t_1}{2} \right\rfloor \pmod{n}, s_2 + \left\lfloor \frac{t_2}{2} \right\rfloor \pmod{n}, \left\lfloor \frac{t_1}{2} \right\rfloor, \left\lfloor \frac{t_2}{2} \right\rfloor\right). \quad (1.16)$$

Отсюда следует, что, если изображения  $B$  и  $T$  являются Хаф-образами для прямых изображений  $J_{\langle n,n,n' \rangle}^3$  и  $K_{\langle n,n,n' \rangle}^3$  соответственно, то, согласно определению 4, утверждению 1 и определению 15,  $H$  будет Хаф-образом для прямых конкатенации изображений  $J$  и  $K$  по третьей координате.

Рассмотрим теперь основную функцию алгоритма 1.3 – *fht3line*. Пусть на вход ей подано изображение  $I_{\langle n,n,n,1 \rangle}^4$ :  $I_{\langle n,n,n,1 \rangle}^4(x, y, z, 1) = I_n^3(x, y, z)$ ,  $x, y, z \in \overline{0, n-1}$ , где  $I_n^3$  – изображение, Хаф-образ для прямых которого требуется построить. Функция *fht3line* рекурсивно разбивает входное изображение  $I_{\langle n,n,n,1 \rangle}^4$  попалам по третьей координате до размера 1 по ней. На обратном ходу рекурсии *fht3line* объединяет смежные фрагменты попарно, используя функцию *mergeHT*.

Поскольку изображения размерами  $\langle n, n, 1, 1 \rangle$ , получаемые функцией *fht3line* на последнем шаге рекурсивного деления, являются корректными

(тривиальными) Хаф-образами соответствующих фрагментов размера  $\langle n, n, 1 \rangle$  изображения  $I_n^3$ , то, по индукции, на шаге  $i$  обратного хода рекурсии будут строиться Хаф-образы фрагментов  $I_{\langle n, n, 2^i \rangle}^3$  исходного изображения  $I_n^3$ . На последней итерации  $2^i = n$  и выходное изображение  $H_n^4$  является требуемым Хаф-образом.  $\square$

На основе этого утверждения в диссертации показано, что сложность алгоритма, вычисляющего полную систему диадических прямых, составляет  $\Theta(n^4)$  и эта оценка асимптотически неуллучшаема.

Обращаясь к вопросу параллелизации вычисления данного преобразования, мы ясно видим, что располагая числом процессоров порядка  $\Theta(n^4)$  возможно вычислить ПХ используя наивную схему вычисления. Если такого числа процессоров нет в наличии (что так для большинства вычислителей уже при  $n = 16$ :  $16^4 = 65536$ ), можно, воспользовавшись алгоритмом 1.3, вычислить ПХ с использованием  $\Theta(n^3)$  процессоров (т.е. используя порядка 4096 последовательных вычислителей) за линейное время  $\Theta(n)$ .

Эта алгоритмическая схема аналогичным образом обобщается и на случаи большей размерности, увеличивается только число возможных сдвигов между парой паттернов в слое. Вычислительная сложность алгоритма в  $m$ -мерном случае для  $m > 2$  составляет  $\Theta(n^{2m-2})$ .

## Глава 2. Поиск прямых и плоскостей с помощью БПХ в двумерном и трехмерном случаях

### 2.1 Методы поиска прямых на изображении

В литературе выделяют один основной подход поиска прямых на изображении – преобразование Хафа [18; 21]. В обзорной части предыдущей главы (смотри раздел 1.1.5) приведено множество алгоритмов быстрого поиска прямых, среди которых в контексте данной работы выделяется быстрое преобразование Хафа [1]. В своей работе автор его называл быстрым дискретным преобразованием Радона, поскольку результатом работы алгоритма является дискретный Радон-образ – Хаф-образ, где выполняется поиск координат максимального значения. В диссертации дальше будет показано, что именно благодаря тому, что Хаф-образ вычисляется полностью, для приведенного метода можно получить верхние оценки точности.

В статистике задачу поиска прямой формулируют как задачу линейной регрессии. Линейная регрессия (ЛР) – это регрессионная модель зависимости одной (объясняемой, зависимой) переменной от другой или нескольких других переменных (факторов, регрессоров, признаков, независимых переменных) с линейной функцией зависимости [126]. Задача линейной регрессии заключается в поиске параметров этой зависимости.

Линейная регрессия стала первой строго изученной моделью анализа, активно используемой на практике [127]. Для классической линейной регрессии доказано, что МНК-оценки являются несмещенными, состоятельными и наиболее эффективными в классе всех линейных несмещённых оценок.

Список возможных применений ЛР можно разбить на две категории:

- для задач предсказания, прогнозирования или уменьшения ошибки;
- для оценки параметров влияния факторов на зависимую переменную.

Подробно о задаче линейной регрессии и её практической значимости написано в книге [127].

В работе [128] предложена модификация линейной регрессии в двумерном случае, так называемая регрессия Деминга. Автор избавился от предположения, что координаты факторов известны точно, таким образом устранив

различие между зависимыми и независимыми переменными. Обобщение этой модели на случай размерности три и более, а также для более сложных функций ошибок называют ортогональной линейной регрессией (ОЛР) [129]. Для её решения, аналогично ЛР, предложен метод наименьших полных квадратов [130].

Перечисленные задачи являются частным случаем М-оценок [131]. К списку М-оценок, обладающих свойством робастности, можно отнести М-оценки Уэлша, Тьюки, Джиммана-МакКлюра, Коши и другие.

Задача вычисления М-оценки по множеству наблюдений  $X = \{(x_{i1}, x_{i2}, \dots, x_{im})\}_{i=1}^N$  может быть записана следующим образом:

$$\hat{\theta} = \operatorname{argmin}_{\vec{\theta}} \sum_{i=1}^N \rho(\varepsilon_{\vec{\theta}}(\vec{x}_i)), \quad (2.1)$$

где  $N$  – число наблюдений,  $\varepsilon_{\vec{\theta}}(\vec{x}_i)$  – расстояние между точкой  $\vec{x}_i \in \mathbb{R}^m$  и гиперплоскостью с нормалью  $\vec{\theta}$  (или прямой, некоторым образом задаваемой  $\vec{\theta}$ ),  $\rho$  – неубывающая на  $[0, \infty)$  дифференцируемая функция с ограниченной производной, равная нулю в нуле.

Для приближенного решения задачи (2.1) обычно используют методы итеративного перевзвешивания (МИП), сложность которых составляет  $\Theta(Nt)$ , где  $t$  – число итераций метода. Из-за того, что целевой функционал не выпуклый МИП не даёт гарантии сходимости [132].

На стыке анализа изображений и статистики был предложен альтернативный метод вычисления М-оценки в задаче ОЛР для поиска прямых на изображении. Из работы [78] Н. Кирьяти и А. Брукштейн следует, что максимум Радон-образа функции вида

$$I_K(\vec{x}) = K(\|\vec{x}\|) * I(\vec{x}), \quad (2.2)$$

где

$$I(\vec{z}) = \sum_{\vec{x} \in X} \delta(\vec{z} - \vec{x}), \quad (2.3)$$

а  $K$  – радиально-симметричное ядро свёртки форма которого определяется типом М-оценки, является М-оценкой в задаче ОЛР. Авторы также показали, что для получения  $K$  достаточно вычислить обратное преобразование Абеля от конкретной функции  $\rho$ ; подробный вывод этого утверждения есть в статье [70]. Вычислительная сложность предложенного метода составляет  $\Theta(Nn_\theta n_\rho)$ , где  $n_\theta, n_\rho$  – размеры Хаф-образа. По сравнению с МИП он позволяет, управляя

дискретизацией пространства Хафа, выбрать подходящий для конкретной задачи баланс между временем вычисления и точностью определения  $\hat{\theta}$ , оценки которых известны заранее.

Предложенный метод стал популярен в анализе изображений и позднее исследовался [133], уточнялся, обобщался [134] и применялся для решения задачи поиска прямых (в том числе множества прямых) на изображении [135]. В работе [136] авторы провели серию вычислительных экспериментов для оценки точности работы данного метода, однако аналитических оценок точности в настоящее время не получено.

В диссертации предлагается развитие метода Н. Кирьяти и А. Брукштейна с использованием БПХ для двумерных и трехмерных изображений [137]. Предложенный метод на синтетических данных сравнивался с другими методами ЛР, устойчивыми к выбросам: методом главных компонент (МГК) [129], МГК с итеративным пересчетом весов [138], методом Тейла-Сена [139], методом Зигеля [140].

## 2.2 Связь между вычислением М-оценок и преобразованием Радона

ПР функции  $f$ , интеграл которой по любой гиперплоскости существует, может быть записано следующим образом:

$$Rf(\vec{\theta}) = \int_{\mathbb{R}^m} \delta(\vec{x}\vec{\theta}_a - \theta_0) f(\vec{x}) d\vec{x}, \quad (2.4)$$

где  $\theta_0$  – расстояние от начала координат,  $\vec{\theta}_a = (\theta_1, \dots, \theta_m)$  – координаты единичного вектора нормали к искомой гиперплоскости, а  $\delta(\vec{x})$  – дельта-функция Дирака. Определим *индикаторную функцию*

$$f^I(\vec{z}) = \sum_{\vec{x} \in X} \delta(\vec{z} - \vec{x}), \quad (2.5)$$

и *свертку индикаторной функции*

$$f_K^I(\vec{x}) = K(\|\vec{x}\|) * f^I(\vec{x}), \quad (2.6)$$

где  $K : \mathbb{R} \rightarrow \mathbb{R}$ .

Следуя работе [141], введем понятие Преобразование Радона с послойной свёрткой:

$$R_{\kappa}f^I(\vec{\theta}) = \int_{\mathbb{R}^n} \kappa(\vec{x}\vec{\theta}_a - \theta_0)f^I(\vec{x})d\vec{x}, \quad (2.7)$$

где  $\kappa(x) = C - \rho(x)$ , а  $C$  такая, что  $\forall \vec{\theta} \in \mathbb{R}^{m+1} : R_{\kappa}f^I(\vec{\theta}) \geq 0$ .

В двумерном случае метод Н. Кирьяти и А. Брукштейна заключается в прямой проекции точек  $X$  на аккумулятор размером  $n_{\theta}n_{\rho}$ , так, что каждой точке ставится в соответствие синусоидальный образ (в полярной параметризации прямой), профиль которого задаётся функцией М-оценки  $\rho$ , смотри рисунок 2.3. Образы всех точек складываются и формируют Хаф-образ изображения, являющийся дискретной аппроксимацией Радон-образа  $R_{\kappa}f^I$ . Искомой М-оценкой  $\hat{\theta}$  считаются координаты элемента Хаф-образа с максимальным значением, а точность её определения зависит от дискретизации Хаф-образа.

Заметим, что задача (2.1) может быть переписана следующим образом:

$$\hat{\theta} = \arg \max_{\theta} \left[ C \left( 1 - \sum_{i=0}^m \rho(\varepsilon_{\vec{\theta}}(\vec{x}_i)) \right) \right], \quad (2.8)$$

где  $C$  - любая константа.

**Утверждение 11.** Для индикаторной функции  $f^I$ , любой неубывающей ограниченной сверху функции  $\rho(x)$ , равной нулю при  $x = 0$ , и некоторой константы  $C$  существует такое одномерное ядро свертки вида  $\kappa(x) = C(1 - \rho(x))$ , что координаты максимума в пространстве  $R_{\kappa}f^I(\theta)$  являются М-оценкой для соответствующего множества наблюдений  $X$ .

*Доказательство.* Без ограничения общности рассмотрим доказательство для двумерного случая.

Докажем верность данного утверждение для случая, когда  $X$  состоит из одного наблюдения  $\vec{x}$ . Тогда  $f^I = \delta(\vec{x})$ , смотри рисунок 2.1, а в пространстве Радона  $Rf^I(\theta)$  ненулевые значения будут принимать лишь прямые, проходящие через точку  $\vec{x}$  (смотри рисунок 2.2).

Тогда, для функции (2.8) для любого  $\theta$  есть  $R_{\kappa}f^I(\theta)$  с ядром  $\kappa(x) = c(1 - \rho(x))$  (смотри рисунок 2.3), где  $c$  вычисляется по формуле

$$c = \max_{\theta} (\rho(\varepsilon_{\vec{\theta}}(\vec{x}_i)0)), \quad (2.9)$$

чтобы выполнялось условие  $\kappa(y) \geq 0, \forall y \in \mathbb{R}^+$ . Такое число всегда существует и конечно, поскольку размер гистограммы - конечен.

Покажем теперь верность утверждения для случая, когда  $X$  содержит более одного измерения. В этом случае для разных наблюдений константа  $c$ , вычисленная по формуле (2.9), может быть различной для М-оценок с неограниченным носителем. Выберем максимальную  $c$  из них, обозначим её символом  $C$ . Поскольку ПР является интегральным, то

$$R_\kappa(f + g)(\theta) = R_\kappa f(\theta) + R_\kappa g(\theta),$$

и, следовательно:

$$R_\kappa f^I(\theta) = \sum_{\vec{x} \in X} R\delta(\vec{x})(\theta)$$

Таким образом, для вычисления функционала (2.8) достаточно вычислить  $R_\kappa f^I(\theta)$  с ядром  $\kappa(x) = C(1 - \rho(x))$ , что и требовалось доказать.

Аналогичные рассуждения верны и для случаев большей размерности.  $\square$

Таким образом,

$$\hat{\theta} = \operatorname{argmax}_{\vec{\theta}} R_\kappa f^I(\vec{\theta}). \quad (2.10)$$

Поскольку  $K$  и  $\kappa$  связаны преобразованием Абеля, то  $Rf_K^I(\vec{\theta}) = R_\kappa f^I(\vec{\theta})$ . Для проверки этого достаточно рассмотреть Радон образы для фиксированного угла  $\theta_a$  (в случае радиально-симметричных ядер проекции для разных углов не различаются), которые совпадают и равны  $\kappa$ . Следовательно для множества точек  $X$  и функции расстояния  $\rho$ , убывающей быстрее, чем  $1/r$  (или финитной), можно, используя обратное преобразование Абеля, построить такое ядро свёртки  $K$

$$K(r) = A\rho = \frac{1}{\pi} \int_r^\infty \frac{\rho'(x)}{\sqrt{x^2 - r^2}} dx, \quad (2.11)$$

что М-оценка (2.1) в задаче ОЛР равна координатам максимума Радон-образа  $Rf_K^I$ :

$$\hat{\theta} = \operatorname{argmax}_{\vec{\theta}} Rf_K^I(\vec{\theta}), \quad (2.12)$$

где  $\vec{\theta} = (\theta_0, \theta_1, \dots, \theta_m)$ .

Заметим, что при  $\rho(e) = e^2$  задача (2.1) сводится к регрессии Деминга. Если  $\rho(e)$  является выпуклой функцией, то функционал в задаче (2.1) является выпуклым, так как является суммой выпуклых функций с одинаковой областью определения. Примером этому служит М-оценка Хьюбера [131]:

$$\rho(\varepsilon) = \begin{cases} \varepsilon^2/2, & |\varepsilon| \leq k \\ k(|\varepsilon| - k/2), & |\varepsilon| > k, \end{cases} \quad (2.13)$$

где  $k$  – действительный положительный параметр. В работе [132] показано существование сходящегося к глобальному оптимуму итеративного метода для задачи (2.1) с М-оценкой Хьюбера.

Однако, для М-оценок Уэлша

$$\rho(\varepsilon) = \frac{c^2}{2}(1 - \exp(-(\varepsilon/c)^2)), \quad (2.14)$$

Тьюки

$$\rho(\varepsilon) = \begin{cases} \frac{c^2}{2}(1 - (1 - (\varepsilon/c)^2)^3), & |\varepsilon| \leq c \\ \frac{c^2}{6}, & |\varepsilon| > c, \end{cases} \quad (2.15)$$

где  $c$  – положительная константа, Коши

$$\rho(\varepsilon) = \frac{c}{2} \log_2 \left( 1 + \frac{\varepsilon^2}{c^2} \right) \quad (2.16)$$

и Джимана-МакКлюра

$$\rho(\varepsilon) = \frac{\varepsilon^2}{2(1 + \varepsilon^2)} \quad (2.17)$$

функция в задаче 2.1 не выпуклая, а следовательно, при использовании МИП нет гарантии выбора правильного начального приближения, что может привести к выбору оптимума в неверном бассейне значений. Тем не менее возможность вычисления оценок Уэлша и Тьюки и других невыпуклых М-оценок представляет особый интерес для случаев данных, содержащих выбросы разного типа, так как они обладают высоким показателем асимптотической толерантности.

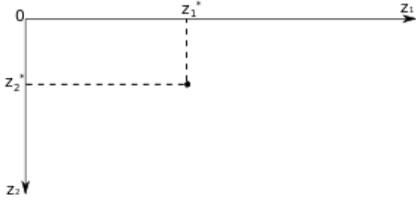


Рисунок 2.1 — Функция изображения  $I$  для одной точки с координатами  $(z_1^*, z_2^*)$  в двумерном пространстве

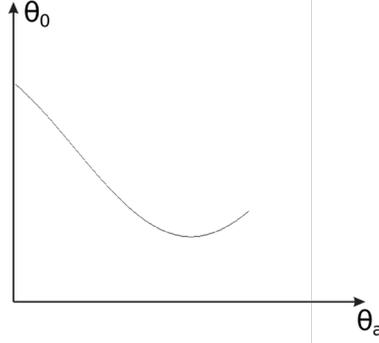


Рисунок 2.2 — Пространство радона функции  $I$

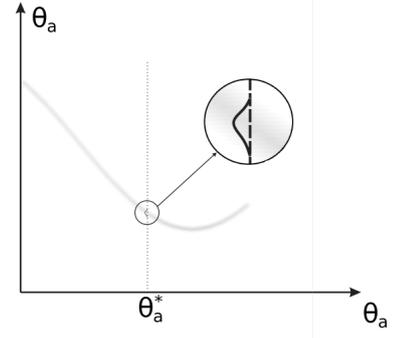


Рисунок 2.3 — Свертка пространства Радона вдоль координаты  $\theta_0$

### 2.3 Метод поиска прямых и плоскостей с использованием свертки Хаф-образа

В этом разделе предложен новый метод поиска прямых и плоскостей на двумерных и в трехмерных изображениях, путём вычисления М-оценок в задаче ОЛР. Метод является развитием метода Н. Кирьяти и А. Брукштейна, только в качестве входных данных используется гистограмма наблюдений, а вместо СПХ используются алгоритмы, предложенные в первой главе.

Общим первым этапом алгоритмов является формирование гистограммы по множеству входных наблюдений. Пусть  $a_j = \min_{i \in [1; N]}(x_{ij})$ ,  $b_j = \max_{i \in [1; N]}(x_{ij}) - \min_{i \in [1; N]}(x_{ij})$ . Пусть  $\hat{n} = \max_{j \in [1; m]}(n_j)$ , где  $n_j = b_j/h$ , притом шаг дискретизации  $h$  одинаковый по всем измерениям и выбирается таким, что  $\forall j \in [1, m] : n_j > 2$ .

**Определение 18.** Гистограмма  $H$  множества  $X$  — это равностороннее изображение  $I_n^m$  с  $n = 2^{\lceil \log_2 \hat{n} \rceil}$ , значение пикселя  $v$  — элемент абелевой группы  $(\mathbb{R}, +)$ , равный  $|\lambda_x|$ , где  $\lambda_x = \{x \in X | z_j - \frac{1}{2} < \frac{x_j - a_j}{b_j} n \leq z_j + \frac{1}{2}, \forall z \in \mathbb{Z}_n^m, \forall j \in [1, m]\}$ .

Для любого диадического паттерна с ненулевым сдвигом существует один пиксель, такой, что координата  $z_2$  его соседа отличается на  $n - 1$ . Для устранения этого эффекта «разрыва» в области определения изображения и решения на ней задачи ОЛР вводится расширенное изображение.

**Определение 19.** *Расширенным изображением  $I_{\langle n \rangle}^m$  по компоненте  $j$  на величину  $l$  назовём отображение  $\text{Aug}_j^l(I_{\langle n \rangle}^m) : \mathbb{Z}_{\langle n^a \rangle}^m \rightarrow \mathbb{R}$ , такое что  $\text{Aug}_j^l(I_{\langle n \rangle}^m)(z) = I_{\langle n \rangle}^m(z)$  для  $z \in \mathbb{Z}_{\langle n \rangle}^m$  и  $\text{Aug}_j^l(I_{\langle n \rangle}^m)(z) = 0$  для  $z \in \mathbb{Z}_{\langle n^a \rangle}^m \setminus \mathbb{Z}_{\langle n \rangle}^m$ , где  $n_j^a = n_j + l$ , а также  $\forall i \in \overline{1, m}, i \neq j : n_i^a = n_i$ .*

При переходе к расширенному изображению асимптотическая вычислительная сложность алгоритмов БПХ и ТБПХ не меняется, если  $l = cn$ , где  $c$  – неотрицательная константа, что непосредственно следует из устройства алгоритмов. В двумерном случае расширение изображения приводит к тому, что попарно суммируются столбцы длиной, не  $n$ , а  $l + n$ . Минимальное значение  $l$  для устранения «разрыва» всех паттернов основной системы БПХ составляет  $n$ , таким образом вычислительная сложность БПХ удваивается. В случае ТБПХ для плоскостей минимальное значение  $l$  аналогично БПХ равно  $n$ . Наконец в случае ТБПХ для прямых, для устранения эффекта «разрыва» расширение изображения производится в двух направлениях  $z_1, z_2$  на величину  $n$ , что увеличивает трудоёмкость в 4 раза, но не меняет асимптотическую вычислительную сложность.

Поскольку основная система диадических паттернов расширенного изображения содержит в себе основную систему исходного изображения, то и соответствующий алгоритм БПХ вычисляет суммы в том числе по всем паттернам из последней.

Паттерну (или плоскости, или прямой) из основной системы диадических паттернов (или плоскостей, или прямых), содержащих пиксели расположенные в области ненулевых значений, ставятся в соответствие корреспондентные прямые или плоскости с теми же свойствами разброса (любой диадический паттерн системы – сдвинутый основной, а от сдвига структура паттерна и, соответственно, его ошибки не меняется), что и для порождающих диадических паттернов (или плоскостей, или прямых). В двумерном случае корреспондентная прямая проходит через точки  $(0, s)$  и  $(n - 1, s + t)$ , в трехмерном случае корреспондентная плоскость –  $(0, 0, s)$ ,  $(n - 1, 0, s + t_1)$  и  $(n - 1, n - 1, s + t_1 + t_2)$ , а прямая –  $(s_1, s_2, 0)$  и  $(s_1 + t_1, s_2 + t_2, n - 1)$ .

Методы приближенного вычисления М-оценок в задаче ОЛР, предлагаемые в данной диссертации, делятся на два, в зависимости от порядка вычисления свёртки и Хаф-образа. На первом шаге выполняется формирование гистограммы  $H = I_{\hat{n}}^m$  множества наблюдений  $X$ . Далее описывается первый

метод, последующие шаги которого одинаковы и выполняются для всех элементов множества симметрий  $(B_2^s, B_3^s)$ .

1. Вычисление Хаф-образа  $J$  от расширенного изображения гистограммы<sup>1</sup>.
2. Свёртка Хаф-образа для каждого фиксированного набора угловых координат с ядром, соответствующим некоторой М-оценке, коэффициент растяжения линейного размера которого равен  $c(t)$  (для каждого из трёх алгоритмов коэффициент выписан ниже).
3. Умножение каждого элемента Хаф-образа с угловыми координатами  $t$  на  $1/c(t)$  для компенсации искажения значений, вызванных анизотропностью пиксельной решетки входной гистограммы.
4. Поиск координат  $\hat{\theta}$  максимума в Хаф-образе.

На заключительном шаге из полученных векторов  $\hat{\theta}$  выбирается тот, значение соответствующего Хаф-образа для которого максимально. Вид  $c(t)$  зависит от того, какой из трёх алгоритмов рассматривается. Пусть  $\alpha_t = \arctg(\frac{t}{n-1})$ , тогда  $\cos(\alpha_t) = (n-1)\sqrt{\frac{1}{t^2+(n-1)^2}}$ , тогда для двумерного случая коэффициент коррекции составляет  $c(t) = c_2(t) = \cos(\alpha_t)$ , для плоскостей в трехмерном случае  $c(t) = c_3^p(t) = \cos(\alpha_{t_1}) \cdot \cos(\alpha_{t_2})$ , а для прямых в трехмерном случае  $c(t) = c_3^l(t) = \sqrt{\frac{(n-1)^2}{(n-1)^2+t_1^2+t_2^2}}$ .

**Утверждение 12.** *Вычислительная сложность метода поиска прямых со свёрткой Хаф-образа составляет  $\Theta(N + n^2 \log n)$  для двумерного случая;  $\Theta(N + n^3 \log n)$  для плоскостей в трехмерном случае;  $\Theta(N + n^4 \log n)$  или  $\Theta(N + w^2 n^4)$  (без использования БПФ), где  $w$  – линейный размер носителя ядра свертки для прямых в трехмерном случае.*

Если ядро свёртки соответствует М-оценке Уэлша, сложность вычисления в последнем случае составляет  $\Theta(N + n^4)$  (с использованием алгоритма Дериса). Доказательство следует из следующих соображений: формирование гистограммы требует  $N$  сложений, БПХ вычисляется за  $\Theta(n^2 \log n)$ , для свёртки необходимо  $\Theta(n^2 \log n)$  (с использованием быстрого алгоритма свёртки через БПФ) и, наконец, для поиска максимального значения  $\Theta(n^2)$ . Аналогично утверждение доказывается для трехмерных случаев. Алгоритмы позволяют

<sup>1</sup>Для двумерного случая используется расширенное изображение  $\text{Aug}_2^n(H)$ , для плоскостей в трехмерном случае  $\text{Aug}_3^{2n}(H)$ , а для прямых –  $\text{Aug}_1^n(\text{Aug}_2^n(H))$ .

адаптивно подбирать параметры аддитивного шума, контролируя окно свёртки, без пересчёта Хаф-образа.

Предложенные алгоритмы предпочтительнее при работе с большим числом наблюдений. Например, в двумерном случае если  $N \gg n^2 \log_2 n$ , сложность предложенного алгоритма  $\Theta(N)$ . При этом предложенные алгоритмы обладают тем же преимуществом, что и исходный метод Н. Кирьяти и А. Брукштейна: возможностью балансировки между требуемой точностью и временем вычисления.

## 2.4 Метод поиска прямых и плоскостей с использованием свертки исходного изображения

В этом разделе описан второй метод, согласно которому процедура свёртки выполняется в исходном пространстве перед вычислением БПХ. В этом случае не требуется выполнять адаптацию размера ядра свёртки (шаг 2 метода с использованием свёртки Хаф-образа).

На первом шаге выполняется формирование гистограммы  $H = I_{\hat{n}}^m$  множества наблюдений  $X$ . Последующие шаги метода выполняются поочередно для всех элементов множества симметрий  $(B_2^s, B_3^s)$ .

1. Свёртка изображения с ядром, соответствующим некоторой М-оценке.
2. Вычисление Хаф-образа  $J$  от расширенного изображения гистограммы<sup>2</sup>.
3. Умножение каждого элемента Хаф-образа с угловыми координатами  $t$  на  $1/c(t)$  для компенсации искажения значений, вызванных анизотропностью пиксельной решетки входной гистограммы.
4. Поиск координат  $\hat{\theta}$  максимума в Хаф-образе.

**Утверждение 13.** *Вычислительная сложность метода поиска прямых с использованием свёртки исходного пространства составляет  $\Theta(N + n^2 \log n)$  для двумерного случая;  $\Theta(N + n^3 \log n)$  для плоскостей в трехмерном случае;  $\Theta(N + n^4)$  для прямых в трехмерном случае.*

<sup>2</sup>Для двумерного случая используется расширенное изображение  $\text{Aug}_2^n(H)$ , для плоскостей в трехмерном случае  $\text{Aug}_3^{2n}(H)$ , а для прямых –  $\text{Aug}_1^n(\text{Aug}_2^n(H))$ .

Сбор гистограммы наблюдений требует  $N$  операций сложения. На шагах 2-4 выполняется анализ собранной гистограммы, из которых у 2 шага самая высокая асимптотическая вычислительная сложность, равная сложности алгоритма быстрого вычисления соответствующего Хаф-образа, смотри главу 1.

Для реализации 1 шага требуется знать формы ядер свёрток в исходном пространстве. Для начала определим эти формы для БПХ и ТБПХ по прямому поскольку оба эти алгоритма вычисляют суммы по прямому, форма ядер для них будет одинаковая.

Известна проекция ядра, полученная в результате применения преобразования Радона (или Йона); и требуется найти образ этого ядра, иными словами, обернуть ПР для этой функции. Воспользуемся утверждениями, полученными в [70] и рассмотрим плоский случай. Для плоского случая ПР принято записывать в виде:

$$Rf(\theta) = \int_{-\infty}^{\infty} f(-x \sin(\theta_a) + \theta_0 \cos(\theta_a), x \cos(\theta_a) + \theta_0 \sin(\theta_a)) dx, \quad (2.18)$$

где  $f$  функция свертки. Так как, по предположению,  $f$  радиальная и  $\sqrt{x_1^2 + x_2^2} = r$ , то обозначим  $f(x_1, x_2) = F(r)$

Ясно, что проекция радиальной функции по всем направлениям одинакова (не зависит от угла), таким образом вместо двумерного ПР можно рассматривать одномерную функцию  $\hat{\mathbb{R}}f(\theta_0)$ , где  $p \in \mathbb{R}$ :

$$Rf(\theta_0) = \int_{-\infty}^{\infty} F\left(\sqrt{x^2 + \theta_0^2}\right) dx = \int_{|\theta_0|}^{\infty} \frac{F(r)}{\sqrt{r^2 - \theta_0^2}} r dr.$$

Данное интегральное преобразование является преобразованием Абеля и может быть обращено для функций  $F(r)$  спадающих быстрее, чем  $1/r$ :

$$F(r) = -\frac{1}{\pi} \int_r^{\infty} \frac{Rf'(\theta_0)}{\sqrt{\theta_0^2 - r^2}} d\theta_0 \quad (2.19)$$

Используя уравнение (2.19), аналогично работе [142], получим уравнения свертки Тьюки, Уэлша, Джимана-МакКлюра и Коши в исходном пространстве. Для М-оценки Тьюки (2.15).

$$Rf(x) = \begin{cases} \frac{c^2}{6} (1 - (x/k)^2)^3, & |x| \leq c \\ 0, & |x| > c. \end{cases} \quad (2.20)$$

Возьмем производную от данной функции:

$$Rf'(x) = \begin{cases} -x(1 - (\frac{x}{c})^2)^2, & |x| \leq c \\ 0, & |x| > c. \end{cases} \quad (2.21)$$

Таким образом ядро свертки для М-оценки Тьюки записывается следующим образом:

$$F(r) = \frac{1}{\pi} \int_r^c \frac{-x(1 - (\frac{x}{c})^2)^2}{\sqrt{x^2 - r^2}} dx = \frac{8}{15\pi c^4} (c^2 - r^2)^{\frac{5}{2}} \quad (2.22)$$

Производная для М-оценки Джимана-МакКлюра (2.17) равна:

$$Rf'(x) = -\frac{x}{(1 + x^2)^2}, \quad (2.23)$$

тогда ядро свертки записывается следующим образом:

$$F(r) = -\frac{1}{\pi} \int_r^{+\infty} -\frac{x}{(1 + x^2)^2} \frac{dx}{\sqrt{x^2 - r^2}} = \frac{1}{4(r^2 + 1)^{3/2}}. \quad (2.24)$$

Производная для М-оценки Уэлша (2.14) равна:

$$Rf'(x) = -x \exp\left(-\frac{x^2}{c^2}\right), \quad (2.25)$$

тогда ядро свертки записывается следующим образом:

$$F(r) = -\frac{1}{\pi} \int_r^{+\infty} -x \exp\left(-\frac{x^2}{c^2}\right) \frac{dx}{\sqrt{x^2 - r^2}} = \sqrt{\frac{4c^2}{\pi}} \exp\left(-\frac{r^2}{c^2}\right). \quad (2.26)$$

Производная для М-оценки Коши (2.16) равна:

$$Rf'(x) = -\frac{x}{1 + (x/c)^2}, \quad (2.27)$$

тогда ядро свертки, если максимальное расстояние между прямой и точкой известно и равно  $r_{max}$  (что является нормальным предположением при работе с изображениями), записывается следующим образом:

$$F(r) = -\frac{1}{\pi} \int_r^{r_{max}} -\frac{x}{1 + (x/c)^2} \frac{dx}{\sqrt{x^2 - r^2}} = \frac{c^2 \arctg \frac{\sqrt{r_{max}^2 - r^2}}{\sqrt{c^2 + r^2}}}{\pi \sqrt{c^2 + r^2}}. \quad (2.28)$$

Наконец рассмотрим случай М-оценки  $\rho(x) = x^2$ . Производная в этом случае равна:

$$Rf'(x) = -2x, \quad (2.29)$$

тогда ядро свертки, если максимальное расстояние между прямой и точкой известно и равно  $r_{max}$  (что является нормальным предположением при работе с изображениями), записывается следующим образом:

$$F(r) = -\frac{1}{\pi} \int_r^{r_{max}} -2x \frac{dx}{\sqrt{x^2 - r^2}} = \frac{1}{\pi} \sqrt{r_{max}^2 - r^2}. \quad (2.30)$$

Аналогичным образом могут быть получены выражения и для других М-оценок.

Для преобразования Радона в трехмерном случае требуется к полученным ядрам повторно применить обратное преобразование Абеля. Таким образом для М-оценки Уэлша получим, что производная для М-оценки Уэлша (2.14) равна:

$$Rf'(x) = -\frac{4x}{\pi c^2} \exp\left(-\frac{x^2}{c^2}\right), \quad (2.31)$$

тогда ядро свертки записывается следующим образом:

$$F(r) = \frac{2}{\pi} \exp\left(-\frac{r^2}{c^2}\right). \quad (2.32)$$

Получим аналогичный результат для М-оценки Джимана-МакКлюора, в этом случае производная равна:

$$Rf'(x) = -\frac{3x}{4(x^2 + 1)^{5/2}} \quad (2.33)$$

тогда ядро свертки записывается следующим образом:

$$F(r) = \frac{1}{2\pi(r^2 + 1)^2}. \quad (2.34)$$

Аналогичным образом могут быть получены выражения и для других М-оценок с радиальной функцией на конечном носителе в задаче ортогональной линейной регрессии на двумерных и трехмерных гистограммах можно.

В описанном в начале раздела методе используются дискретные аппроксимации этих свёрток.

## 2.5 Оценка точности метода поиска прямых

В разделе получены оценки точности аппроксимации корреспондентных прямых и плоскостей предложенным в 2.4 методом на основе результатов первой главы (о величинах максимальных ортогональных разбросов диадических паттернов).

Оптимизируемая в задаче (2.1) функция может быть записана в следующем виде:

$$f_{\rho}^X(\vec{\theta}) = \sum_{i=1}^N \rho(\varepsilon_{\vec{\theta}}(\vec{x}_i)) / N. \quad (2.35)$$

и определяется типом М-оценки. Ниже в аргументе функции  $\varepsilon_{\vec{\theta}}(\vec{x}_i)$  значок вектора опускается. Из-за дискретизации пространства признаков при формировании гистограммы, использования дискретной аппроксимации свертки и диадического паттерна (или плоскости, или прямой) функция  $\rho$  вычисляется с погрешностью  $\Delta\rho(\varepsilon_{\vec{\theta}}(x_i)) = \left| \rho(\varepsilon_{\vec{\theta}}(x_i)) - \hat{\rho}(\varepsilon_{\hat{\theta}}(\hat{x}_i)) \right|$ , где  $\hat{\rho}$  – дискретизированная функция  $\rho$ ,  $\hat{x}_i$  – позиция ближайшего к точке узла дискретной решётки, а  $\hat{\theta}$  – параметры найденной корреспондентной прямой. Оценим эту величину для двумерного случая сверху: поскольку  $\rho$  – неубывающая функция, то  $\rho(x + h\sqrt{2}) \geq \hat{\rho}(x)$ , тогда

$$\begin{aligned} \Delta\rho(\varepsilon_{\vec{\theta}}(x_i)) &\leq \left| \rho(\varepsilon_{\vec{\theta}}(x_i)) - \rho(\varepsilon_{\hat{\theta}}(\hat{x}_i) + h\sqrt{2}) \right| \leq \\ &\rho'_{max} \cdot \left| \varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i) - h\sqrt{2} \right| \leq \rho'_{max} h\sqrt{2} + \rho'_{max} \cdot \left| \varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i) \right|, \end{aligned}$$

где  $\rho'_{max}$  супремум производной функции  $\rho$ .

**Утверждение 14.** *Ошибка вычисления (2.35) для прямой с параметрами  $\vec{\theta}$  в двумерном случае ограничена следующим выражением:*

$$\Delta f_{\rho}^X(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left( \frac{\log_2 n}{6} + 2\sqrt{2} + 2 \right). \quad (2.36)$$

Верность утверждения следует из того, что ошибка вычисления ортогонального расстояния от прямой до точки  $\left| \varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i) \right|$ , в свою очередь складывается, из ортогонального разброса диадического паттерна, ограниченного величиной  $\frac{\log_2 n}{6}$ , погрешности дискретизации точки при сборе гистограммы

$h\sqrt{2}$  и максимально возможного изменения расстояния до точки  $x \in X$  при переходе от корреспондентной прямой с параметрами  $\hat{\theta}$  к прямой  $\vec{\theta}$  (в двумерном случае оно ограничено сверху величиной  $2h$ ).

**Утверждение 15.** *Ошибка вычисления функции (2.35) для плоскости с параметрами  $\vec{\theta}$  в трехмерном случае с помощью ТБПХ для плоскостей ограничена следующим выражением:*

$$\Delta f_{\rho}^X(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left( \frac{\log_2 n}{3} + 2\sqrt{3} + 2 \right). \quad (2.37)$$

Доказательство этого утверждения выполняется аналогично, только в свойства разброса диадического паттерна заменяются свойствами разброса диадической плоскости, а также погрешность округления на двумерном изображении – погрешностью округления на трехмерном.

**Утверждение 16.** *Ошибка вычисления функции (2.35) для прямой с параметрами  $\vec{\theta}$  в трехмерном случае с помощью ТБПХ для прямых ограничена следующим выражением:*

$$\Delta f_{\rho}^X(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left( \frac{\sqrt{2} \log_2 n}{6} + 2\sqrt{3} + 2\sqrt{2} \right). \quad (2.38)$$

Доказательство этого утверждения выполняется аналогично, только в свойства разброса диадического паттерна заменяются свойствами разброса диадической прямой, а погрешность округления на двумерном изображении – погрешностью округления на трехмерном.

## 2.6 Экспериментальное исследование предложенных алгоритмов

### 2.6.1 Модель входных данных

Опишем модель данных, используемую для генерации тестовых гистограмм. В данной работе рассматриваются двумерные и трехмерные гистограммы. Для простоты описания эксперимента здесь и далее будем считать, что все грани гистограммы одинакового размера  $n$ .

Для простоты изложения, когда размерность несущественна, будем использовать одно слово – гистограмма.

Модель данных задает непрерывное распределение на области определения гистограммы, согласно которому выполняется генерация тестовых точек. Список точек  $X$  генерируется в диапазоне  $[1, n]^m$ , где  $m \in \{2, 3\}$ . После этого координаты точек округляются до ближайшего целого  $\hat{X} = \text{round}(X)$ . Если в списке  $\hat{X}$  содержится  $k$  одинаковых точек, то  $k$  записывается в качестве значения в соответствующую точку гистограммы.

Подробнее о распределении тестовых точек. Точки делятся на три типа:

1.  $\hat{X}_{signal}^\theta$  – точки, нормально распределенные с нулевым математическим ожиданием и известным стандартным отклонением  $\sigma_a$  поперёк исходной восстанавливаемой гиперплоскости.
2.  $\hat{X}_{out}^\theta$  – точки выбросового шума, распределенные нормально относительно центра с диагональной матрицей ковариации, ненулевые элементы которой одинаковы и равны  $\sigma_o$ . Радиус кластера определяется стандартным отклонением и равен удвоенному значению этого параметра. Все точки выбросового кластера принадлежат гистограмме.
3.  $\hat{X}_{uni}^\theta$  – точки, распределенные равномерно по гистограмме.

В таблице 1 приведен список параметров модели данных.

Обозначение	Описание параметра
$n$	линейный размер изображения
$\theta$	параметры истинной гиперплоскости
$w_1$	количество точек, принадлежащих $\hat{X}_{signal}^\theta$
$\sigma_a$	стандартное отклонение аддитивного шума
$w_2$	число выбросовых точек $\hat{X}_{out}^\theta$
$x_o$	координаты центра выбросового кластера
$\sigma_o$	стандартное отклонение точек от центра кластера
$w_3$	количество точек равномерного шума $\hat{X}_{uni}^\theta$

Таблица 1 — Параметры модели данных.

Для всех экспериментов параметры истинной гиперплоскости  $\theta$  генерируются с помощью  $m$  удалённых друг от друга точек внутри гистограммы, через

которые можно провести только одну соответствующую размерности пространства гиперплоскость. Эти точки выбираются внутри гистограммы случайно, но с некоторыми ограничениями. Для двумерного случая выбор первой точки ограничен областью  $x \in [1, \frac{n}{8}]$  и  $y \in [1, \frac{n}{2}]$ , а второй -  $x \in [\frac{7n}{8} + 1, n]$  и  $y \in [1, \frac{n}{2}]$ . Для трехмерного случая первая точка находится внутри области, ограниченной  $x \in [1, \frac{n}{8}]$ ,  $y \in [1, n]$ ,  $z \in [1, \frac{n}{2}]$ , а вторая и третья - в области  $x \in [\frac{7n}{8} + 1, n]$ ,  $y \in [1, n]$ ,  $z \in [1, \frac{n}{2}]$ .

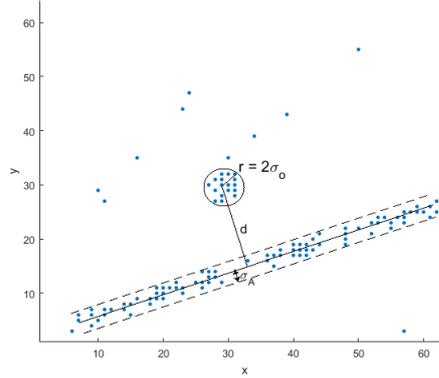


Рисунок 2.4 — Пример тестовой двумерной гистограммы

Кроме того, для приведённых экспериментов с некоторыми ограничениями случайно выбирается центр точек  $\hat{X}_{out}^\theta$  таким образом, чтобы все точки кластера с радиусом  $2\sigma_o$  не выходили за пределы гистограммы и не пересекались с  $\hat{X}_{signal}^\theta$ .

Метрика, по которой измерялось качество в двумерном случае, представляет из себя нормированный на линейный размер гистограммы максимум между длинами дуг  $l1$  и  $l2$ , образованных пересечением описанной вокруг гистограммы окружности и двух прямых: одна из прямых - целевая с истинными параметрами, а вторая - с оцененными параметрами  $\hat{\theta}$ .

$$d(\hat{\theta}) = \frac{\max(l1, l2)}{n} \quad (2.39)$$

Такая метрика показывает близость двух прямых на гистограмме и инвариантна к размеру изображения. Выбор этой метрики обусловлен тем, что она нечувствительна к повороту и её значения можно сравнивать для разных линейных размеров гистограмм  $n$ .

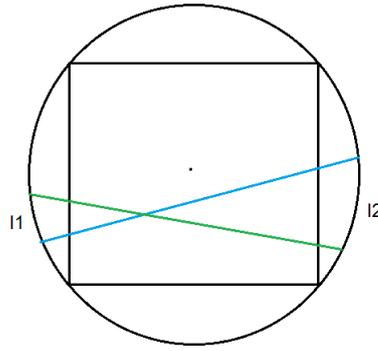


Рисунок 2.5 — Дуги  $l_1$  и  $l_2$  между пересечениями прямых и описанной вокруг гистограммы окружности

### 2.6.2 Сравнительное исследование точности определения параметров истинного сигнала с помощью ПХ и БПХ

Цель эксперимента заключается в изучении зависимости точности регрессии при переходе от ближайшей аппроксимации геометрической прямой в дискретном пространстве к аппроксимации диадическими паттернами. Для этого проводилось сравнение значений метрики (2.39) восстановленных параметров М-оценки Уэлша на гистограмме с использованием СПХ и БПХ.

Эксперимент проводился для следующих параметров:  $n = 512$ ,  $w_1 = 500$ ,  $w_2 = 0$ ,  $\sigma_a = 32$ ,  $\sigma_o = 0$ ,  $w_3 = 0$ .

Для эксперимента генерировалось 1000 тестовых гистограмм и для каждой из них вычислялась аппроксимация М-оценки Уэлша с помощью ПХ и БПХ. На тестовой гистограмме, размытой гауссовским ядром с параметром стандартного отклонения равным  $\sigma_a$ , восстанавливались параметры линейной модели методами ПХ и БПХ. Далее рассчитывались разницы оценок параметров, полученных методом ПХ и БПХ, с настоящими параметрами сигнала по введённой выше метрике (2.39).

### 2.6.3 Сравнение методов решения задачи робастной линейной регрессии для двумерных и трехмерных изображений

Данный эксперимент приводился для исследования качества оценок параметров сигнала с помощью БПХ в сравнении с альтернативными методами

линейной регрессии. Среди методов мы выбрали МНК (ols), Тейла-Сена (ts), Зигеля (siegel), М-оценки Хьюбера (huber), Тьюки (tuckey) и Уэлша (welsch), посчитанных с помощью метода итеративного перевзвешивания. С этой целью были получены значения ошибок каждого из рассматриваемых методов на данных с разной долей выбросового шума. Эксперимент проводился отдельно для двумерного и трехмерного случаев.

Были выбраны следующие параметры модели для данного эксперимента в двумерном случае:  $n = 512$ ,  $\sigma_a = 32$ ,  $\sigma_o = 16$ ,  $w_3 = 0.1$ ; подобранными случайно  $\theta$ ,  $x_o$ ; и изменяемыми вручную  $w_1$  и  $w_2$  так, что  $w_2$  пробегает значения от 0 до 0.7 с шагом 0.1. При этом общее количество точек из  $\hat{X}_{signal}^\theta$  и  $\hat{X}_{out}^\theta$  не менялось и равнялось 500.

Для каждого значения изменяемого параметра  $w_2$  генерировалось по 500 тестовых гистограмм и для каждой из них применялись метод восстановления регрессии с помощью быстрого преобразования Хафа и другие сравниваемые методы (МНК, Тейла-Сена, Зигеля, М-оценки Хьюбера, Таки, Уэлша). Для оценок параметров сигнала каждого из методов рассчитывается максимальное расстояние до истинных параметров гиперплоскости описанной выше метрикой  $d(\hat{\theta})$ .

Генерация тестовых гистограмм в трехмерном случае отличалась от двумерного значениями фиксированных параметров модели данных. Фиксированные параметры в трехмерном случае выглядят так:  $n = 64$ ,  $\sigma_a = 8$ ,  $\sigma_o = 4$ ,  $w_3 = 0.1$ . Кроме того, в трехмерном случае выброшены из рассмотрения методы восстановления регрессии Тейла-Сена и Зигеля по причине их долгой работы на 510 точках. Для трехмерного случая была также использована другая метрика ошибки регрессии:

$$d_3(\hat{\theta}) = \sqrt{\sum_{i=0}^2 (\theta_i - \hat{\theta}_i)^2}$$

## 2.6.4 Результаты экспериментов

### Эксперимент 1

На рисунках 2.6 и 2.7 изображены гистограммы значений ошибки по метрике 2.39 для идеальной аппроксимации геометрической прямой алгоритмом Брезенхема (ПХ) и диадической аппроксимации прямой (БПХ) соответственно. Гистограммы построены для диапазона значений ошибок от 0.00001 до 0.1 с шагом 0.001.

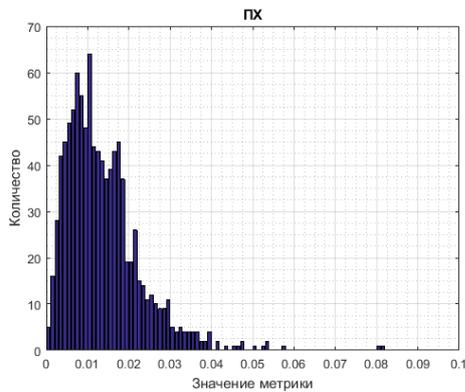


Рисунок 2.6 — Гистограмма ошибок аппроксимации истинной гиперплоскости наилучшими дискретными прямыми

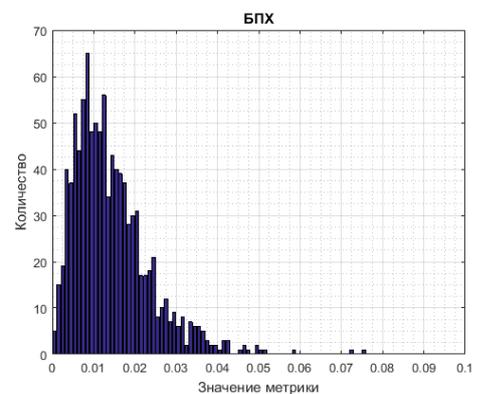


Рисунок 2.7 — Гистограмма ошибок аппроксимации истинной гиперплоскости диадическими паттернами

Из полученных гистограмм 2.6 и 2.7 видно, что вблизи нуля нерепрезентативное количество наблюдений.

Будем сравнивать полученные гистограммы в диапазоне значений ошибок от 0.00101 до 0.9. Такой выбор продиктован практической целесообразностью. Действительно, при значениях метрики до 0.00101 для размера изображения  $512 \times 512$  размер ошибки меньше пикселя, что несущественно.

Сравним две эмпирические функции распределения ошибок восстановления регрессии  $\Phi_{\hat{\theta}}(t) = \sum_{i=1}^m [d_i(\hat{\theta}) < t]$ , где  $d_i(\hat{\theta})$  – ошибка  $i$ -ого наблюдения по метрике 2.39,  $m$  – количество наблюдений. Для этого построим график (2.8)

отношения эмпирических функций распределения ошибок восстановления регрессии методами БПХ и ПХ  $\frac{\Phi_{FHT}}{\Phi_{HT}}$ . Минимум здесь достигается при  $d(\hat{\theta}) = 0.003$  и равен 0.7959.

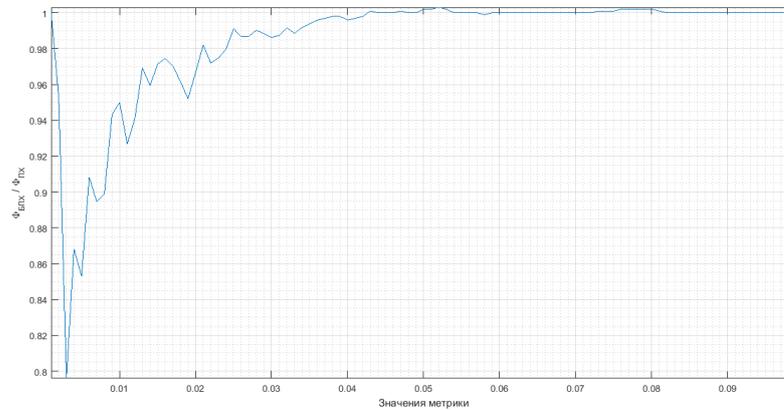


Рисунок 2.8 — График отношения  $\frac{\Phi_{FHT}}{\Phi_{HT}}$  эмпирических функций распределения ошибок

## Эксперимент 2

Сравнение методов ортогональной линейной регрессии для двумерных гистограмм показало, что БПХ превосходит все рассматриваемые методы по предложенной метрике для доли выбросового шума  $\frac{w_2}{w_1+w_2} \leq 0.6$  на модели данных, описанной в подразделе 2.6.1. На рисунке 2.9 проиллюстрированы результаты эксперимента для двумерного случая.

Аналогичное сравнение для трехмерных гистограмм показало, что БПХ превосходит все рассматриваемые в трехмерном случае методы по метрике  $L_2$  для  $\frac{w_2}{w_1+w_2} \leq 0.5$  на введенной модели данных. На рисунке 2.10 показаны результаты эксперимента по сравнению методов для трехмерного случая.

На обоих графиках можно заметить резкий рост ошибки после предельного значения доли выбросового шума. Это объясняется тем, что плотность выбросового кластера для такой конфигурации параметров модели данных становится достаточно большой, чтобы восстанавливаемая гиперплоскость всегда проходила через него.

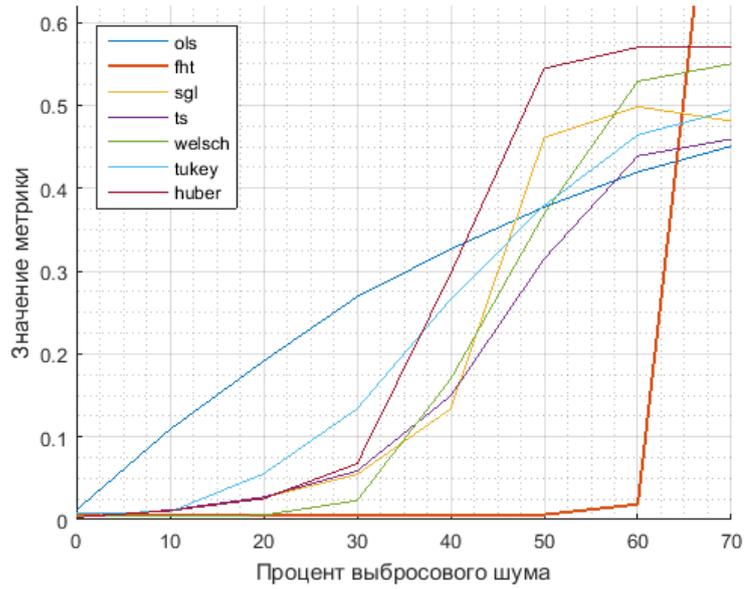


Рисунок 2.9 — Зависимость значения метрики оценок от веса  $\frac{w_2}{w_1+w_2} 100\%$  выбросового кластера

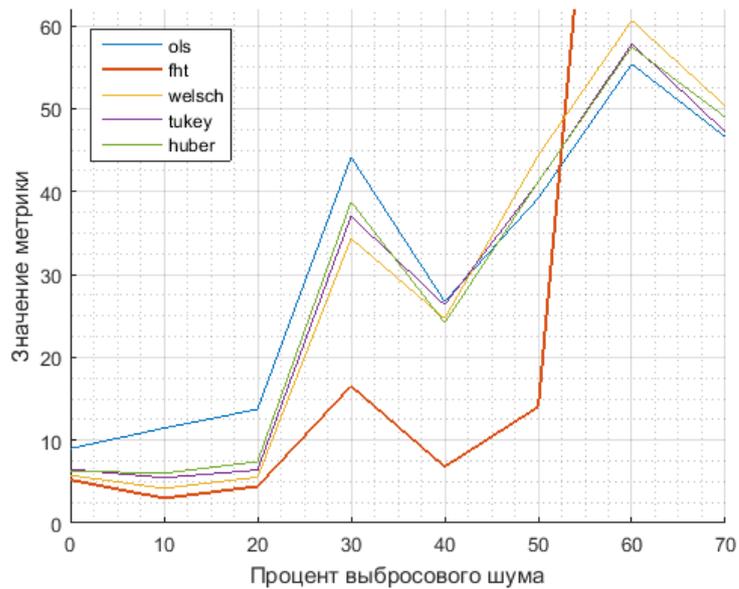


Рисунок 2.10 — Зависимость расстояния по метрике  $L_2$  параметров истинной и восстановленных гиперплоскостей от веса  $\frac{w_2}{w_1+w_2} 100\%$  выбросового кластера

### Глава 3. Быстрая линейная бинарная кластеризация

Глава посвящена созданию метода бинарной кластеризации на двумерных и трехмерных гистограммах в контексте развития обобщенных методов глобальной бинаризации Оцу.

Эта задача – один из этапов большинства алгоритмов, реализующего обобщенный метод глобальной бинаризации Оцу. Предложенный метод также может быть полезен при решении задач цветовой сегментации и константности путём анализа цветowych гистограмм.

В главе предложен метод линейного разделения на пару кластеров в двумерных или трехмерных гистограммах, причем критерий разделения – входной параметр алгоритма. Метод основан на комбинации быстрого преобразования Хафа и кумулятивного суммирования. Множество допустимых критериев ограничено требованием вычислимости по конечному набору статистик, задаваемых в виде функции на области определения гистограммы. Показано, что для двумерного случая вычислительная сложность предложенного метода составляет  $\Theta(n^2 \log n)$  операций, а для трехмерного –  $\Theta(n^3 \log n)$ , где  $n$  – линейный размер гистограммы.

#### 3.1 Метод Оцу для двумерных гистограмм

Метод бинаризации Оцу [143] был предложен в 1979 году японским ученым Н. Оцу. Среди его преимуществ можно выделить то, что он не требует настройки параметров и отличается низким временем вычисления. Это и делает его востребованным, подталкивая ученых исследовать возможности его обобщения для повышения качества. Для гистограммы входного скалярного изображения метод Оцу позволяет найти порог, разделяющий гистограмму на два класса, так, чтобы сумма дисперсий классов, умноженная на их веса, была максимальна. Такой критерий носит название «максимизация внутриклассовой дисперсии», он оптимален в классе распределений, являющихся смесью из пары нормальных распределений с равными весами и дисперсиями. По этой теме написана обзорная статья [144].

Позднее метод Оцу был обобщен на случай двумерных гистограмм, что позволило повысить качество бинаризации, но привело к значительному увеличению времени вычисления. Джиан Гонг с соавторами [145] в 1998 году предложили ускоренный обобщенный двумерный метод бинаризации Оцу. Приведем его вывод на основе метода максимального правдоподобия (ММП).

Пусть дано множество наблюдений  $X = \{(x_i^1, x_i^2, \dots, x_i^m)\}_{i=1}^N$ , пусть  $x_i = (x_i^1, x_i^2, \dots, x_i^m)^T$  – вектор-столбец из некоторого распределения  $f$ . Будем считать, что распределение точек  $X$  является смесью двух нормальных распределений с равными матрицами ковариации. В этом случае известно, что оптимальной границей их разделения является гиперплоскость. Интересно, что гиперплоскость оптимальна и для других параметрических семейств, в частности для экспоненциального, то есть дальнейшие рассуждения могут быть перенесены и на эти случаи.

В своей работе Дж. Гонг и соавторы использовали критерий оптимальный для смеси двух нормальных распределений  $f_1(x) \sim \mathcal{N}(\mu_1, \Sigma)$  и  $f_2(x) \sim \mathcal{N}(\mu_2, \Sigma)$  с равными весами (в каждом классе по  $N/2$  точек) и матрицами ковариации вида  $\Sigma = \lambda I$ , где  $I$  – единичная матрица, а  $\lambda \in \mathbb{R}^+ \cup \{0\}$ . Проведем аналогичные работе [146] выкладки для размерности  $m$ , чтобы вывести критерий. Для начала будем считать, что метка принадлежности точки к классу известна. Определим  $\alpha_i$  как двумерный вектор, кодирующий принадлежность к одному из двух классов: если первая компонента равна единице, а вторая нулю, то  $x_i$  из первого класса, иначе – из второго, и пусть  $A = \{\alpha_i\}_{i=1}^N$ . С учетом вышесказанного плотность вероятности наблюдений может быть записана следующим образом:

$$f(x_i | \alpha_i) = \sum_{j=1}^2 \alpha_{ji} f_j(x_i) = \prod_{j=1}^2 f_j^{\alpha_{ji}}(x_i). \quad (3.1)$$

Запишем правдоподобие для множества наблюдений  $X$ :

$$\mathcal{L}(\mu_1, \mu_2, \lambda | A) = \prod_{i=1}^N f(x_i | \alpha_i) = \prod_{i=1}^N \prod_{j=1}^2 f_j^{\alpha_{ji}}(x_i). \quad (3.2)$$

Поскольку плотность вероятности каждого из классов, по условиям, имеет вид

$$\begin{aligned} f_j(x) &= \frac{1}{\sqrt{2\pi}} \det(\Sigma)^{-1/2} \exp\left(-\frac{1}{2}(x - \mu_j)^T \Sigma^{-1}(x - \mu_j)\right) \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\lambda^{m/2}} \exp\left(-\frac{1}{2\lambda}(x - \mu_j)^T I(x - \mu_j)\right), \end{aligned}$$

подставим это выражение в формулу правдоподобия (3.2); получим

$$\begin{aligned}
\mathcal{L}(\mu_1, \mu_2, \lambda|A) &= \prod_{i=1}^N \prod_{j=1}^2 \frac{1}{(\sqrt{2\pi\lambda^{m/2}})^{\alpha_{ji}}} \exp\left(-\frac{\alpha_{ji}}{2\lambda}(x_i - \mu_j)^T I(x_i - \mu_j)\right) \\
&= \prod_{i=1}^N \frac{1}{(\sqrt{2\pi\lambda^{m/2}})^2} \prod_{j=1}^2 \exp\left(-\frac{\alpha_{ji}}{2\lambda}(x_i - \mu_j)^T I(x_i - \mu_j)\right) = \\
&= (2\pi)^{-\frac{Nm}{2}} \lambda^{-\frac{Nm}{2}} \prod_{i=1}^N \prod_{j=1}^2 \exp\left(-\frac{\alpha_{ji}}{2\lambda}(x_i - \mu_j)^T (x_i - \mu_j)\right) = \\
&= (2\pi)^{-\frac{Nm}{2}} \lambda^{-\frac{Nm}{2}} \exp\left(-\frac{1}{2\lambda} \sum_{i=1}^N \sum_{j=1}^2 \alpha_{ji}(x_i - \mu_j)^T (x_i - \mu_j)\right).
\end{aligned}$$

Введем обозначение для дисперсии каждого из классов  $D_j(A) = \sum_{i=1}^N \alpha_{ji}(x_i - \mu_j)^T (x_i - \mu_j)$ , тогда правдоподобие можно записать в виде

$$\mathcal{L}(\mu_1, \mu_2, \lambda|A) = (2\pi)^{-\frac{Nm}{2}} \lambda^{-\frac{Nm}{2}} \exp\left(-\frac{1}{2\lambda}(D_1(A) + D_2(A))\right).$$

Прологарифмируем данное выражение

$$\ln \mathcal{L}(\mu_1, \mu_2, \lambda|A) = -\frac{Nm}{2} \ln 2\pi - \frac{Nm}{2} \ln \lambda - \frac{1}{2\lambda}(D_1(A) + D_2(A)). \quad (3.3)$$

Для фиксированных  $X$  и  $A$  правдоподобие является функцией  $\mu_1, \mu_2$  и  $\lambda$  оптимальные значения которые вычисляются с использованием следующей системы уравнений:

$$\begin{cases} \frac{\partial \ln \mathcal{L}(\mu_1, \mu_2, \lambda|A)}{\partial \mu_j} = -\frac{1}{\lambda} \sum_{i=1}^N \alpha_{ji}(x_i - \mu_j) = 0 \\ \frac{\partial \ln \mathcal{L}(\mu_1, \mu_2, \lambda|A)}{\partial \lambda} = -\frac{mN}{\lambda} + \frac{1}{\lambda^2}(D_1(A) + D_2(A)) = 0. \end{cases}$$

Из данной системы следует, что оптимальные параметры равны

$$\begin{aligned}
\mu_j(A) &= \frac{2}{N} \sum_{i=1}^N \alpha_{ji} x_i \\
\lambda(A) &= \frac{D_1(A) + D_2(A)}{mN}.
\end{aligned}$$

Подставим полученные значения в уравнение (3.3):

$$\begin{aligned}
\ln \mathcal{L}(\hat{\mu}_1, \hat{\mu}_2, \hat{\lambda}|A) &= \\
&= -N \ln \sqrt{2\pi} - N \ln \lambda^{\frac{m}{2}} - \frac{1}{2\lambda} (D_1(A) + D_2(A)) = \\
&= -N \ln \sqrt{2\pi} - \frac{mN}{2} \ln \frac{D_1(A) + D_2(A)}{mN} - \frac{mN}{2D_1(A) + 2D_2(A)} (D_1(A) + D_2(A)) = \\
&= -N \ln \sqrt{2\pi} - \frac{mN}{2} \ln \frac{D_1(A) + D_2(A)}{mN} - \frac{mN}{2}.
\end{aligned} \tag{3.4}$$

Если метки неизвестны, будем выбирать такое множество  $\hat{A}$ , которое максимизирует правдоподобие

$$-N \ln \sqrt{2\pi} - \frac{mN}{2} \ln \frac{D_1(A) + D_2(A)}{mN} - \frac{mN}{2} \rightarrow \max_{A \in \mathbb{A}}, \tag{3.5}$$

где  $\mathbb{A}$  – множество всевозможных способов присвоения меток класса. Убирая константные члены, получим

$$F_1(A) = D_1(A) + D_2(A) \rightarrow \min_{A \in \mathbb{A}}. \tag{3.6}$$

Поскольку в данной задаче разделяющей поверхностью является гиперплоскость, то именно её параметры  $\theta$  определяют каким образом присваиваются метки: все точки по одну сторону от гиперплоскости принадлежат одному классу, тогда  $D_j(A(\theta)) = D_j(\theta)$ . Таким образом для решения задачи кластеризации множества точек  $X$  из вышеописанного распределения достаточно выполнить поиск параметров гиперплоскости, доставляющий максимум правдоподобию, записанному в уравнении (3.4), или же решить эквивалентную задачу:

$$F_1(\theta) = D_1(\theta) + D_2(\theta) \rightarrow \min_{\theta}. \tag{3.7}$$

Этот критерий именуют обобщением критерия Оцу и используют при решении задачи бинаризации по двумерной гистограмме входного серого изображения [147; 148]. Правда, в работах обычно говорят о минимизации следа внутречласовой матрицы ковариации, оценкой которой в данной задаче является сумма матриц разброса каждого из классов.

Для оптимизации этого критерия в работе [145] был предложен быстрый метод кластеризации гистограммы на основе предподсчета кумулятивных изображений, где в качестве разделяющей поверхности использовался прямоугольник, одна вершина которого 0, а вторая – искомый параметр  $(z_1^t, z_2^t)$ .

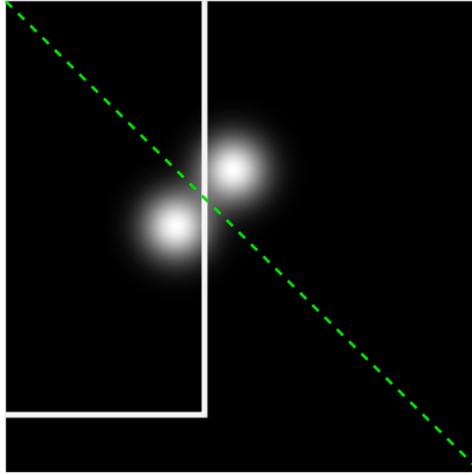


Рисунок 3.1 — Иллюстрация результата работы алгоритма бинарной кластеризации Дж. Гонг [145]. Белый прямоугольник — оптимальный ответ метода, зелёная штриховка — оптимальный ответ.

Авторы использовали прямоугольник вместо прямой потому, что только для него удалось создать быстрый алгоритм. Как при этом искажается критерий, авторы оставили за скобками (смотри рисунок 3.1).

Позднее в работе [149] предложили развитие метода [145]. Авторы, вычисляя тот же критерий на двумерной гистограмме, предлагают разбивать её на два прямоугольника: один с вершинами в 0 и в точке  $(z_1^t, z_2^t)$ , вершины второго находятся в точках  $(z_1^t, z_2^t)$  и  $(n-1, n-1)$ . Иными словами, авторы предлагают уменьшить область второго класса с потерей части собранной гистограммы в угоду скорости вычисления. Авторы отмечают, что точность вычисления критерия уменьшается по сравнению с предыдущей работой, однако дают только качественную оценку этому изменению, демонстрируя его на конкретных примерах.

В работе [150] был предложен способ разделения двумерной гистограммы прямой линией согласно критерию линейного дискриминанта Фишера. Основной сложностью такого подхода является необходимость полного перебора параметров разделяющей поверхности из-за невыпуклости целевого функционала. Позднее в работе [151] данный подход был развит, а для поиска оптимального значения функционала использовался эвристический метод оптимизации, названный «роевым интеллектом». Как следствие, для данного алгоритма отсутствует гарантия оптимального значения функционала.

В диссертации предложен метод быстрого линейного бинарного разделения двух кластеров в плоском случае и плоскостью в трехмерном, основанный на алгоритмах, предложенных в 1 главе.

### 3.2 Обобщения критерия Оцу

В этом разделе приведены другие обобщения критерия Оцу на случай размерности 2 и более. Рассмотрим случай смеси с разными весами и матрицей ковариации вида  $\Sigma = \lambda I$ . Максимальное правдоподобие записывается следующим образом:

$$\ln \mathcal{L}(X_H|A) \sim \sum_{j=1}^2 p_j(\theta) \ln p_j(\theta) - \ln \left( D_1(\theta) + D_2(\theta) \right), \quad (3.8)$$

то есть задача поиска оптимального критерия разделения записывается следующим образом

$$F_2(\theta) = \sum_{j=1}^2 p_j(\theta) \ln p_j(\theta) - \ln \left( D_1(\theta) + D_2(\theta) \right) \rightarrow \min_{\theta}. \quad (3.9)$$

Теперь рассмотрим случай смеси двух нормальных распределений  $f_1(x) \sim \mathcal{N}(\mu_1, \Sigma)$  и  $f_2(x) \sim \mathcal{N}(\mu_2, \Sigma)$  с равными весами, но без ограничений на структуру матрицы ковариации. В этом случае правдоподобие имеет вид

$$\begin{aligned} \mathcal{L}(X_H|A) &= \prod_{i=1}^N \prod_{j=1}^2 \frac{1}{(\sqrt{2\pi} \det(\Sigma))^{-1/2}} \alpha_{ji} \exp \left( -\frac{\alpha_{ji}}{2} (x_i - \mu_j)^T \Sigma^{-1} (x_i - \mu_j) \right) = \\ &= \prod_{i=1}^N \frac{1}{(\sqrt{2\pi} \det(\Sigma))^{-1/2}} \prod_{j=1}^2 \exp \left( -\frac{\alpha_{ji}}{2} (x_i - \mu_j)^T \Sigma^{-1} (x_i - \mu_j) \right), \end{aligned}$$

а его логарифм

$$\begin{aligned} \ln \mathcal{L}(X_H|A) &= -N \ln \sqrt{2\pi} - \frac{N}{2} \ln \det \Sigma - \\ &- \frac{1}{2} \left( \sum_{i=1}^N \alpha_{1i} (x_i - \mu_1)^T \Sigma^{-1} (x_i - \mu_1) + \sum_{i=1}^N \alpha_{2i} (x_i - \mu_2)^T \Sigma^{-1} (x_i - \mu_2) \right). \end{aligned} \quad (3.10)$$

Как и в предыдущем случае нетрудно видеть, что  $\mu_j = \frac{2}{N} \sum_{i=1}^N \theta_{ji} x_i$ , а оценка матрицы ковариации выполняется чуть более сложным способом. Преобразуем

уравнение (3.10)

$$\begin{aligned} \ln \mathcal{L}(X_H|A) &= -N \ln \sqrt{2\pi} - \frac{N}{2} \ln \det \Sigma - \\ &\quad - \frac{1}{2} \left( \sum_{i=1}^N \text{tr}(\alpha_{1i}(x_i - \mu_1)(x_i - \mu_1)^T \Sigma^{-1}) + \sum_{i=1}^N \text{tr}(\alpha_{2i}(x_i - \mu_2)(x_i - \mu_2)^T \Sigma^{-1}) \right) = \\ &= -N \ln \sqrt{2\pi} - \frac{N}{2} \ln \det \Sigma - \\ &\quad - \frac{1}{2} \left( \text{tr} \left( \sum_{i=1}^N \alpha_{1i}(x_i - \mu_1)(x_i - \mu_1)^T \Sigma^{-1} \right) + \text{tr} \left( \sum_{i=1}^N \alpha_{2i}(x_i - \mu_2)(x_i - \mu_2)^T \Sigma^{-1} \right) \right). \end{aligned}$$

Запишем это уравнение с учетом переобозначения  $S_j = \sum_{i=1}^N \alpha_{ji}(x_i - \mu_j)(x_i - \mu_j)^T$

$$\ln \mathcal{L}(X_H|A) = -N \ln \sqrt{2\pi} - \frac{N}{2} \ln \det \Sigma - \frac{1}{2} \left( \text{tr} (S_1 \Sigma^{-1}) + \text{tr} (S_2 \Sigma^{-1}) \right).$$

Дифференциал правдоподобия при нулевом приращении средних имеет вид

$$\begin{aligned} d \ln \mathcal{L}(X_H|A) &= -\frac{N}{2} \text{tr} \Sigma^{-1} \{d\Sigma\} - \\ &\quad - \frac{1}{2} \text{tr} \left[ -\Sigma^{-1} \{d\Sigma\} \Sigma^{-1} S_1 \right] - \frac{1}{2} \text{tr} \left[ -\Sigma^{-1} \{d\Sigma\} \Sigma^{-1} S_2 \right] = \\ &= \text{tr} \left[ \Sigma^{-1} \{d\Sigma\} \left( \frac{N}{2} I - \frac{1}{2} \Sigma^{-1} S_1 - \frac{1}{2} \Sigma^{-2} S_2 \right) \right] = 0. \end{aligned}$$

Откуда получаем оценку для матрицы ковариации

$$\Sigma = \frac{1}{N} (S_1 + S_2) \quad (3.11)$$

Подставив полученные  $\mu_j$  и  $\Sigma$  в уравнение (3.10) и упростив выражение, получим

$$\ln \mathcal{L}(X_H|A) = -\frac{mN}{2} \ln \pi - \frac{N}{2} \ln \det \left( \frac{S_1(\theta) + S_2(\theta)}{N} \right) - \frac{mN}{2}.$$

Принимая во внимание, что для данной модели распределения матрица внутри-классовой ковариации равна  $\Sigma_w(\theta) = \frac{1}{2} S_1(\theta) + \frac{1}{2} S_2(\theta)$ , мы получаем критерий разделение для модели смеси двух равновесных нормальных распределений с равными матрицами ковариации, позволяющий отыскать наиболее правдоподобное разделение множества точек  $X_H$ :

$$F_3(\theta) = \det(S_1(\theta) + S_2(\theta)) \rightarrow \min_{\theta}. \quad (3.12)$$

Согласно работе [146], если нормальные распределения смешаны с разными весами, в максимальное правдоподобие добавляется еще одно слагаемое:

$$\ln \hat{\mathcal{L}}(X_H|A) = -\frac{mN}{2} + N \sum_{j=1}^2 p_j(\theta) \ln p_j(\theta) - \frac{N}{2} \ln \det \left( \frac{S_1(\theta) + S_2(\theta)}{N} \right) - \frac{mN}{2}, \quad (3.13)$$

то есть задача поиска оптимального критерия разделения записывается в виде

$$F_4(\theta) = \sum_{j=1}^2 \hat{p}_j(\theta) \ln \hat{p}_j(\theta) - \frac{1}{2} \ln \det(\hat{S}_1(\theta) + \hat{S}_2(\theta)) \rightarrow \max_{\theta}. \quad (3.14)$$

### 3.3 Метод линейной бинарной кластеризации

Опишем метод для двумерного случая  $(H_{2^k}^2)$ , для основной системы БПХ. Каждому  $p \in Y_d^{2^k, 2}$  с параметрами  $\hat{\theta} = \langle s, t \rangle$  ставится в соответствие множество позиций расширенного изображения  $\text{Aug}_2^{2^k}(H_{2^k}^2)$ , такое, что:  $C_1(\hat{\theta}) = \{z^c : z_1^c = z_1, z_2^c \leq z_2, \forall \langle z_1, z_2 \rangle \in P\}$  множество всех остальных позиций обозначим  $C_2(\hat{\theta})$ .

**Утверждение 17.** *Вычислительная сложность подсчета сумм по множествам  $C_1(\hat{\theta})$  и  $C_2(\hat{\theta})$  изображения  $H_{2^k}^2$  составляет  $\Theta(1)$  при вычислительной сложности предподсчета  $\Theta(n^2 \log n)$ .*

Опишем кратко доказательство этого утверждения. Для изображения  $U = H_{2^k}^2$  формируется кумулятивное изображение  $U^c$  такое, что  $\forall \langle z_1, z_2 \rangle \in \mathbb{Z}_n^2 : U^c(z_1, z_2) = \sum_{z_2^* \in \overline{1, z_2}} U(z_1, z_2^*)$ . Трудоёмкость его создания составляет  $n^2$

операций. Таким образом, сумма по любому  $p \in Y_d^{2^k, 2}$  на расширенном изображении  $\text{Aug}_2^{2^k}(U^c)$  является суммой по  $C_1(\hat{\theta})$ , а сумма по  $C_2(\hat{\theta})$  равна разнице  $\Omega(U) = \sum_{\langle z_1, z_2 \rangle \in \mathbb{Z}_n^2} U(z_1, z_2)$  и суммы по  $C_1(\hat{\theta})$ . Вычислив БПХ от  $\text{Aug}_2^{2^k}(U^c)$  по-

лучим Хаф-образ, содержащий суммы по  $C_1(\hat{\theta})$  для всех  $p \in Y_d^{2^k, 2}$ , доступ к которым осуществляется за константное время. Итоговая вычислительная сложность предподсчета определяется самым трудным этапом – вычислением БПХ и составляет  $\Theta(n^2 \log n)$ .

**Утверждение 18.** *Вычислительная сложность подсчета критериев (3.7), (3.9) для гистограммы  $H_{2^k}^2$  и паттернов  $p \in Y_d^{2^k, 2}$  равна  $\Theta(1)$  при вычислительной сложности предподсчета –  $\Theta(n^2 \log n)$ .*

Доказательство основано на утверждении 17. Для вычисления критерия (3.9) в точке  $\hat{\theta}$  требуется вычислить вес  $\omega_j(\hat{\theta})$  и дисперсию  $D_j(\hat{\theta})$  каждого кластера. Покажем способ вычисления  $\omega_j(\hat{\theta})$ ; для этого достаточно все значения пикселей гистограммы  $H_{2^k}^2$  разделить на  $\Omega(H_{2^k}^2)$ , тогда  $\omega_j(\hat{\theta}) = \sum_{\langle z_1, z_2 \rangle \in C_j(\hat{\theta})} \text{Aug}_2^{2^k}(H_{2^k}^2)(z_1, z_2)$ , а эти суммы, согласно утверждению 17, вычисляются с необходимой асимптотической вычислительной сложностью.

Для вычисления выборочных средних значений кластеров  $\mu_j(\hat{\theta}) = (\mu_{j1}(\hat{\theta}), \mu_{j2}(\hat{\theta}))^T$  сформируем изображения  $M_i : \text{dom}(M_i) = \text{dom}(H_{2^k}^2)$ ,  $M_i(z_1, z_2) = z_i H_{2^k}^2(z_1, z_2)$ , тогда

$$\mu_{ji}(\hat{\theta}) = \frac{1}{\omega_j(\hat{\theta})} \sum_{\langle z_1, z_2 \rangle \in C_j(\hat{\theta})} \text{Aug}_2^{2^k}(M_i)(z_1, z_2), \quad (3.15)$$

$\forall i, j = \overline{1, 2}$ . Согласно утверждению 17, сумма в формуле (3.15) вычисляется с необходимой асимптотической вычислительной сложностью. Аналогично формуле (3.15) вычисляется вектор  $\nu_j(\hat{\theta}) = (\nu_{j1}(\hat{\theta}), \nu_{j2}(\hat{\theta}))^T$ ,  $\forall i, j = \overline{1, 2}$  для изображений  $L_i : \text{dom}(L_i) = \text{dom}(H_{2^k}^2)$ ,  $L_i(z_1, z_2) = z_i^2 H_{2^k}^2(z_1, z_2)$ . В результате легко видеть, что  $D_j(\hat{\theta}) = \nu_j(\hat{\theta}) - \mu_j(\hat{\theta}) \circ \mu_j(\hat{\theta})$ . Следовательно, поскольку асимптотическая вычислительная сложность подсчета составляющих равна  $\Theta(1)$  при сложности предподсчета  $\Theta(n^2 \log n)$ , то и для подсчета дисперсии это верно, а значит и для подсчета критерия (3.9). Так же можно вычислить элементы эмпирической матрицы ковариации  $S_j(\hat{\theta})$ , а после и критерий (3.14).

Аналогично вычисляется любой критерий разделения, выразимый через меры, заданные на области определения гистограммы.

Опишем алгоритмы для трехмерного случая ( $H_{(e)}^3$ ), для основной системы диадических плоскостей. Каждому  $p \in Y_{dp}^{(e), 3}$  с параметрами  $\hat{\theta} = \langle s, t_1, t_2 \rangle$  ставится в соответствие множество позиций расширенного изображения  $\text{Aug}_3^{2^k}(H_{(e)}^3)$ , разбитое на два:  $V_1(\hat{\theta}) = \{z^c : z_1^c = z_1, z_2^c = z_2, z_3^c \leq z_3, \forall \langle z_1, z_2, z_3 \rangle \in P\}$ ; множество всех остальных позиций обозначим  $V_2(\hat{\theta})$ .

**Утверждение 19.** *Вычислительная сложность подсчета сумм по множествам  $V_1(\hat{\theta})$  и  $V_2(\hat{\theta})$  изображения  $H_{(e)}^3$  составляет  $\Theta(1)$  при вычислительной сложности предподсчета  $\Theta(n^3 \log n)$ .*

Опишем кратко доказательство этого утверждения. Для изображения  $U = H_{\langle e \rangle}^3$  формируется кумулятивное изображение  $U^c$  такое, что  $\forall \langle z_1, z_2, z_3 \rangle \in \mathbb{Z}_n^3 : U^c(z_1, z_2, z_3) = \sum_{z_3^* \in \overline{1, z_3}} U(z_1, z_2, z_3^*)$ . Трудоёмкость его создания составляет  $n^3$  операций. Таким образом, сумма по любому  $p \in Y_{dp}^{\langle e \rangle, 3}$  на расширенном изображении  $\text{Aug}_3^{2^k}(U^c)$  является суммой по  $V_1(\hat{\theta})$ , а сумма по  $V_2(\hat{\theta})$  равна разнице  $\Omega(U) = \sum_{\langle z_1, z_2, z_3 \rangle \in \mathbb{Z}_n^2} U(z_1, z_2, z_3)$  и суммы по  $V_1(\hat{\theta})$ . Вычислив БПХ от  $\text{Aug}_3^{2^k}(U^c)$  получим Хаф-образ, содержащий суммы по  $V_1(\hat{\theta})$  для всех  $p \in Y_{dp}^{\langle e \rangle, 3}$ , доступ к которым осуществляется за константное время. Итоговая вычислительная сложность подсчета определяется самым трудным этапом – вычислением БПХ – и составляет  $\Theta(n^3 \log n)$ .

**Утверждение 20.** *Вычислительная сложность подсчета критериев (3.7), (3.9) для гистограммы  $H_{2^k}^2$  и паттернов  $p \in Y_d^{2^k, 2}$  равна  $\Theta(1)$  при вычислительной сложности подсчета  $\Theta(n^3 \log n)$ .*

Доказательство основано на утверждении 19. Для вычисления критерия (3.9) в точке  $\hat{\theta}$  требуется вычислить вес  $\omega_j(\hat{\theta})$  и дисперсию  $D_j(\hat{\theta})$  каждого кластера. Покажем способ вычисления  $\omega_j(\hat{\theta})$ ; для этого достаточно все значения пикселей гистограммы разделить на  $\Omega(H_{\langle e \rangle}^3)$ , тогда  $\omega_j(\hat{\theta}) = \sum_{\langle z_1, z_2, z_3 \rangle \in V_j(\hat{\theta})} \text{Aug}_3^n(H_{\langle e \rangle}^2)(z_1, z_2, z_3)$ , а эти суммы, согласно утверждению 19, вычисляются с необходимой вычислительной сложностью.

Для вычисления выборочных средних значений кластеров  $\mu_j(\hat{\theta}) = (\mu_{j1}(\hat{\theta}), \mu_{j2}(\hat{\theta}), \mu_{j3}(\hat{\theta}))^T$  сформируем изображения  $M_i : \text{dom}(M_i) = \text{dom}(H_{\langle e \rangle}^3)$ ,  $M_i(z_1, z_2, z_3) = z_i H_{\langle e \rangle}^3(z_1, z_2, z_3)$ . Тогда

$$\mu_{ji}(\hat{\theta}) = \frac{1}{\omega_j(\hat{\theta})} \sum_{\langle z_1, z_2, z_3 \rangle \in V_j(\hat{\theta})} \text{Aug}_3^{2^k}(M_i)(z_1, z_2, z_3), \quad (3.16)$$

$\forall i = \overline{1, 2}, j = \overline{1, 3}$ . Согласно утверждению 19, сумма в формуле (3.15) вычисляется с необходимой вычислительной сложностью. Аналогично формуле (3.16) вычисляется вектор  $\nu_j(\hat{\theta}) = (\nu_{j1}(\hat{\theta}), \nu_{j2}(\hat{\theta}), \nu_{j3}(\hat{\theta}))^T$ . В результате легко видеть, что  $D_j(\hat{\theta}) = \nu_j(\hat{\theta}) - \mu_j(\hat{\theta}) \circ \mu_j(\hat{\theta})$ . Следовательно, поскольку вычислительная сложность подсчета составляющих равна  $\Theta(1)$  при сложности подсчета  $\Theta(n^3 \log n)$ , то и для подсчета дисперсии это верно, а значит, и для подсчета критерия (3.9). Так же можно вычислить элементы эмпирической матрицы ковариации  $S_j(\hat{\theta})$ , а после и критерии (3.12), (3.14).

Аналогично вычисляется любой критерий разделения, выразимый через меры, заданные на области определения гистограммы.

Таким образом, общая схема метода линейной бинарной кластеризации для размерности два и три может быть описана следующей последовательностью шагов:

1. Формирование изображений, содержащих в себе меры, необходимые для вычисления входного критерия;
2. Кумулятивный подсчет для всех изображений с первого шага;
3. Вычисление БПХ (ТБПХ для плоскостей) для всех кумулятивных изображений;
4. Формирование критерия из полученных изображений;
5. Поиск координат  $\hat{\theta}$  оптимального значения критерия.

Как было сказано выше, вычислительная сложность такого метода определяется самым трудоёмким этапом – БПХ, и составляет  $\Theta(n^2 \log n)$  для двумерного случая и  $\Theta(n^3 \log n)$  – для трехмерного.

На рисунке 3.2 приведён пример работы предложенного алгоритма и алгоритма Дж. Гонг [145] для критерия 3.7. На рисунке 3.3 представлен пример работы предложенного метода в трехмерном случае.

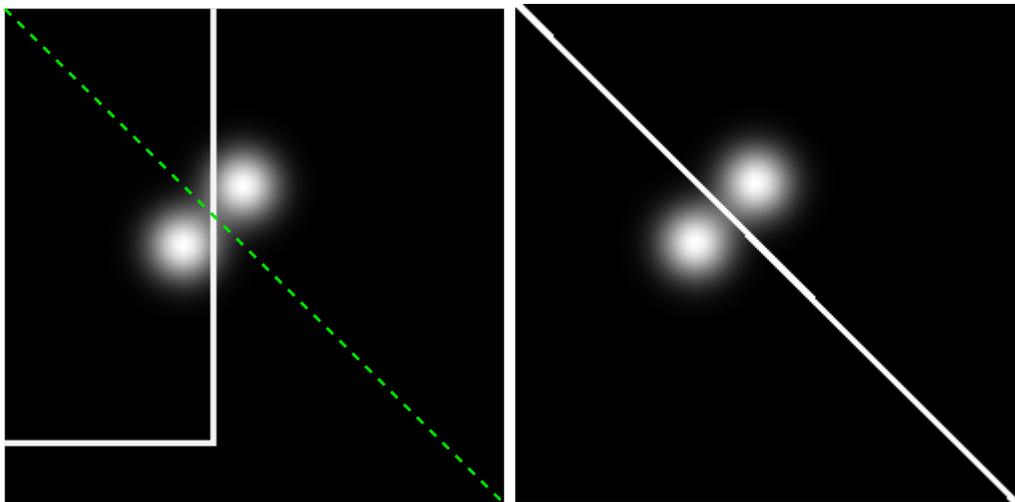


Рисунок 3.2 — Иллюстрация результата работы алгоритма бинарной кластеризации Дж. Гонг [145] и предложенного в диссертации алгоритма.

Описанный в разделе метод изложен в работах [152; 153].

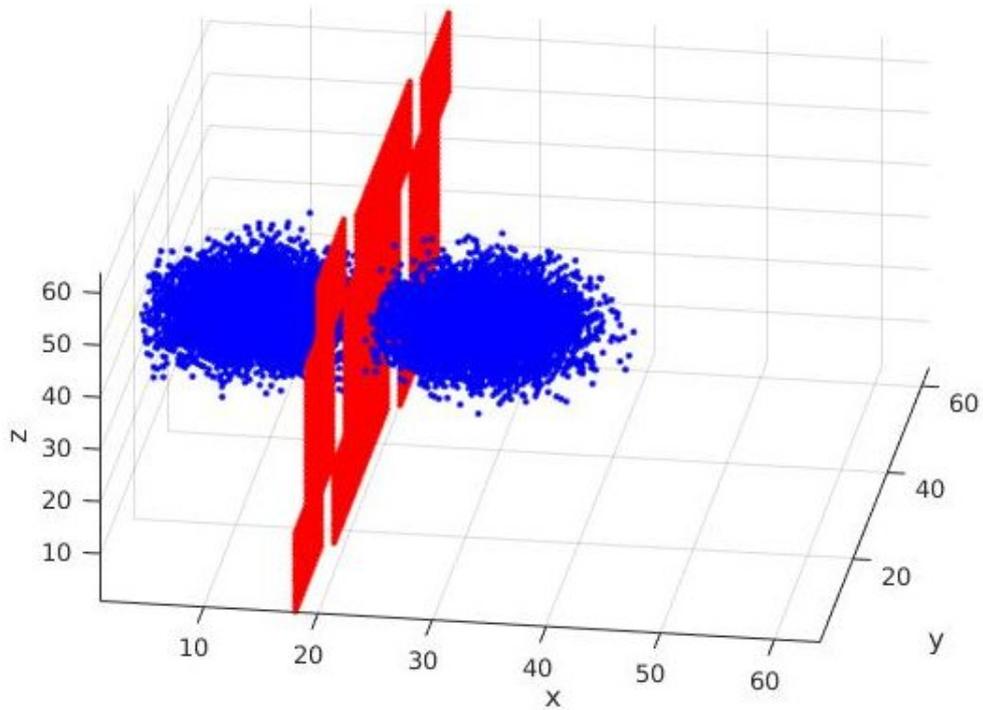


Рисунок 3.3 — Иллюстрация предложенного в диссертации алгоритма в трехмерном случае для критерия (3.7).

### 3.4 Оценка точности метода линейной бинарной кластеризации

В данном разделе предложен метод получения оценок точности вычисления критериев разделения в точке  $\theta$  по заданной гистограмме  $H$ . Рассмотрим метод на примере двумерного случая и критерия (3.9).

Погрешность подсчета критерия состоит из погрешностей подсчета его составляющих. Рассмотрим основные правила вычисления погрешности сложной величины, которые потребуются для анализа точности критерия (3.9).

Пусть для пары величин  $a_1, a_2 \in \mathbb{R}^+$  известны погрешности их вычисления  $\Delta a_1, \Delta a_2 \in \mathbb{R}^+$ , такие, что  $a_1 > \Delta a_1$  и  $a_2 > \Delta a_2$ . Тогда погрешность произведения равна:

$$\begin{aligned}
 \Delta(a_1 \cdot a_2) &= \\
 &= \max((a_1 + \Delta a_1)(a_2 + \Delta a_2) - a_1 a_2, a_1 a_2 - (a_1 - \Delta a_1)(a_2 - \Delta a_2)) = \\
 &= a_1 \Delta a_2 + a_2 \Delta a_1 + \Delta a_1 \Delta a_2.
 \end{aligned}
 \tag{3.17}$$

Аналогично получим погрешность вычисления частного:

$$\Delta\left(\frac{a_1}{a_2}\right) = \max\left(\frac{a_1}{a_2} - \frac{a_1 - \Delta a_1}{a_2 - \Delta a_2}, \frac{a_1 + \Delta a_1}{a_2 + \Delta a_2} - \frac{a_1}{a_2}\right) = \frac{a_1 \Delta a_2 + a_2 \Delta a_1}{a_2(a_2 - \Delta a_2)}. \quad (3.18)$$

Для логарифма от величины погрешность вычисления составляет:

$$\Delta \ln a_1 = \max\left(\ln a_1 - \ln(a_1 - \Delta a_1), \ln(a_1 + \Delta a_1) - \ln a_1\right) = \frac{a_1}{a_1 - \Delta a_1}. \quad (3.19)$$

Для вычисления погрешности всех четырех критериев, приведённых в этой работе, потребуется вычисление погрешности следующей операции:

$$\Delta(a_1 \ln a_1) = \max\left(a_1 \ln a_1 - (a_1 - \Delta a_1) \ln(a_1 - \Delta a_1), (a_1 + \Delta a_1) \ln(a_1 + \Delta a_1) - a_1 \ln a_1\right). \quad (3.20)$$

Максимум данной операции не раскрывается, аналогично предыдущим случаям.

Подход основан на оценке погрешностей мер, участвующих в вычислении критерия. Рассмотрим для начала погрешность вычисления веса в точке  $\hat{\theta} = \langle s, t \rangle$  по его Хаф-образу  $J_{\omega_j}(\hat{\theta})$ , полученному в процессе выполнения алгоритма, предложенного в предыдущей главе. Обозначим  $\Delta s = \max\left(1, \lceil \frac{\log_2 n}{6} \rceil\right)$ , где второй член соответствует величине ортотропной ошибки диадического паттерна, тогда  $\Delta \omega_j(\hat{\theta}) = J_{\omega_j}(s + \Delta s, t) - J_{\omega_j}(s - \Delta s, t)$ . Пример подсчета погрешности вычисления приведен на рисунке 3.4

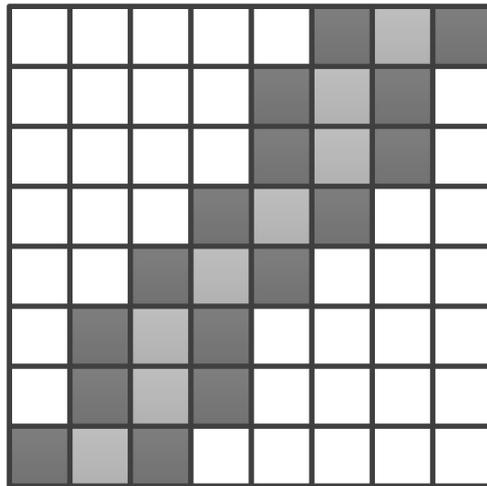


Рисунок 3.4 — Погрешность вычисления суммы по светло-серому паттерну вычисляется как модуль разности сумм по тёмно-серым паттернам.

Аналогично оценим погрешность вычисления сумм по изображениям  $M_i$  по их Хаф-образам:  $J_{M_i}(\hat{\theta}) = \sum_{\langle z_1, z_2 \rangle \in C_j(\hat{\theta})} \text{Aug}_2^{2^k}(M_i)(z_1, z_2)$ , получим  $\Delta J_{M_i}(\hat{\theta}) = J_{M_i}(s + \Delta s, t) - J_{M_i}(s - \Delta s, t)$ . Тогда погрешность вычисления оценки среднего получается на основании формулы (3.18) и величин  $\Delta \omega_j(\hat{\theta}), \Delta J_{M_i}(\hat{\theta})$ .

Для вычисления оценки точности дисперсии достаточно аналогично предыдущим двум примерам рассчитать погрешность для  $\nu_j(\hat{\theta})$ , а затем сложить её с погрешностью квадрата величины среднего, рассчитанной согласно формуле (3.17).

Повторяя те же рассуждения, только для второго кластера, получим оценки точности определения его веса, среднего и дисперсии. Наконец, все рассчитанные погрешности складываются сообразно формуле (3.9) с использованием приведенных в начале раздела правил. Аналогичным образом получают оценки точности остальных трех критериев.

Таким образом, предложена схема оценки точности критерия для конкретной гистограммы с помощью метода бинарной линейной кластеризации с использованием результатов промежуточных вычислений, которые в нём возникают. Для случаев большей размерности вывод оценки точности определения критерия выполняется аналогичным образом.

## Заключение

Основные результаты работы заключаются в следующем.

1. Предложены два обобщения диадического паттерна на трехмерный случай – диадическая плоскость и диадическая прямая.
2. Показано, что максимальный ортогональный разброс диадического паттерна для БПХ ограничен снизу величиной  $\frac{\log_2 n}{2\sqrt{10}}$ , а сверху –  $\frac{\log_2 n}{6}$ , ТБПХ для плоскостей ограничен снизу –  $\frac{\log_2 n}{\sqrt{11}}$ , а сверху –  $\frac{\log_2 n}{3}$  и, наконец, ТБПХ для прямых ограничен снизу величиной  $\frac{\log_2 n}{\sqrt{22}}$ , а сверху –  $\frac{\sqrt{2}\log_2 n}{6}$ , где  $n$  – сторона равностороннего изображения соответствующей размерности.
3. Доказано, что система диадических паттернов в БПХ 2-полна, то есть для любой пары пикселей на изображении найдется проходящий через них ДП.
4. Предложены два алгоритма вычисления трехмерного быстрого преобразования Хафа (ТБПХ) для прямых и плоскостей.
5. Показано, что ТБПХ для прямых обладает вычислительной сложностью  $\Theta(n^4)$ , а ТБПХ для плоскостей вычисляется за  $\Theta(n^3 \log n)$ , где  $n$  – линейный размер изображения.
6. Предложен метод поиска прямых на двумерных, а также прямых и плоскостей в трехмерных гистограммах путём вычисления М-оценки в задаче ОЛР с помощью алгоритмов БПХ и ТБПХ. Для двумерных изображений доказано, что вычислительная сложность предложенного метода составляет  $\Theta(N + n^2 \log n)$ , для трехмерных прямых –  $\Theta(N + n^4)$ , а для плоскостей –  $\Theta(N + n^3 \log n)$ , где  $N$  – число наблюдений.
7. Предложен метод линейной бинарной кластеризации с помощью алгоритмов БПХ для двумерных и трехмерных гистограмм и любых критериев, выразимых через меры, заданные на области определения гистограммы. Вычислительная сложность метода составляет  $\Theta(n^2 \log n)$  для двумерных гистограмм, и  $\Theta(n^3 \log n)$  – для трехмерных. Предложен метод оценки сверху ошибки вычисления величины критерия для заданной входной гистограммы и параметров разделения.

В заключение благодарю Д. П. Николаева за поддержку, обсуждение результатов и научное руководство, а также весь научный коллективу лаборатории №11 «Зрительные системы» ИППИ РАН. Кроме того благодарю С. М. Карпенко за плодотворные научные дискуссии и за пример математика, которому хочется соответствовать.

Отдельно благодарю авторов шаблона \*Russian-Phd-LaTeX-Dissertation-Template\* за помощь в оформлении диссертации.

В наибольшей степени благодарю своего деда Ершова Ю.А. за увлекательное путешествие в мир науки, которое бы без него не началось.

## Список сокращений и условных обозначений

$n$	размер стороны изображения в пикселях
$m$	размерность изображения или признакового пространства
$\langle n_1, n_2, \dots, n_m \rangle$	кортеж размеров сторон изображения в пикселях
$I$	изображение
$J$	Хаф-образ
$z_i$	координатная ось
$H$	гистограмма
$p$	паттерн
$T$	пучок паттернов
$Y$	система паттернов
$X$	множество наблюдений
$x$	одно наблюдение
$N$	число наблюдений
<b>ПХ</b>	преобразование Хафа
<b>ПР</b>	преобразование Радона
<b>ДП</b>	диадический паттерн
<b>СПХ</b>	стандартное преобразование Хафа
<b>ОПХ</b>	обобщенное преобразование Хафа
<b>ВПХ</b>	вероятностное преобразование Хафа
<b>БПХ</b>	быстрое преобразование Хафа
<b>ТБПХ</b>	трёхмерное быстрое преобразование Хафа
<b>ИБПХ</b>	иерархическое быстрое преобразование Хафа
<b>ЦПХ</b>	цифровое преобразование Хафа
<b>БПФ</b>	быстрое преобразование Фурье
<b>ЛР</b>	линейная регрессия
<b>МНК</b>	метод наименьших квадратов
<b>ОЛР</b>	ортогональная линейная регрессия
<b>МИП</b>	метод итеративного перевзвешивания
<b>МГК</b>	метод главных компонент
<b>ММП</b>	метод максимального правдоподобия

## Список литературы

1. *Brady, M. L.* Fast parallel discrete approximation algorithms for the Radon transform / M. L. Brady, W. Yong // Proceedings of the fourth annual ACM symposium on Parallel algorithms and architectures. — ACM. 1992. — С. 91—99.
2. *Brady, M.* A fast discrete approximation algorithm for the Radon transform / M. Brady // SIAM J. Computing. — 1998. — Т. 27, № 1. — С. 107—119.
3. *Ershov, E. I.* Generalization of the Fast Hough Transform for Three-Dimensional Images / E. I. Ershov, A. P. Terekhin, D. P. Nikolaev // Journal of Communications Technology and Electronics. — 2018. — Vol. 63, no. 6. — P. 626—636.
4. Fast Hough transform analysis: pattern deviation from line segment / E. Ershov [и др.] // Eighth International Conference on Machine Vision. — International Society for Optics, Photonics. 2015. — С. 987509I 1—5.
5. Fast 3D Hough Transform Computation. / E. Ershov [и др.] // 30 European Conference on Modelling and Simulation. — 2016. — С. 227—230.
6. О точной оценке неточностей аппроксимации прямых в алгоритме быстрого преобразования Хафа / Е. И. Ершов [и др.] // ИТиС 2015. 2015. — С. 858—868.
7. Generation algorithms of fast generalized Hough transform / E. Ershov [и др.] // 31st European Conference on Modelling and Simulation. — 2017. — С. 534—538.
8. *Hough, P. V.* Machine analysis of bubble chamber pictures / P. V. Hough // International conference on high energy accelerators and instrumentation. Т. 73. — 1959. — С. 2.
9. *Hough., P. V.* Method and means for recognizing complex patterns : Patent / P. V. Hough., A. Arbor. — US, 12.1962.
10. *Hart, P. E.* How the Hough transform was invented [DSP History] / P. E. Hart // IEEE Signal Processing Magazine. — 2009. — Т. 26, № 6.

11. *Rosenfeld, A.* Picture processing by computer / A. Rosenfeld // ACM Computing Surveys (CSUR). — 1969. — T. 1, № 3. — С. 147—176.
12. *Duda, R. O.* Use of the Hough transformation to detect lines and curves in pictures / R. O. Duda, P. E. Hart // Communications of the ACM. — 1972. — T. 15, № 1. — С. 11—15.
13. *Davies, E.* Image space transforms for detecting straight edges in industrial images / E. Davies // Pattern Recognition Letters. — 1986. — T. 4, № 3. — С. 185—192.
14. *Immerkær, J.* Some remarks on the straight line Hough transform / J. Immerkær // Pattern Recognition Letters. — 1998. — T. 19, № 12. — С. 1133—1135.
15. *Iannino, A.* A survey of the Hough transform and its extensions for curve detection / A. Iannino, S. D. Shapiro // Proc. IEEE Comput. Soc. Conf. Pattern Recognition and Image Processing. — 1978. — С. 32—38.
16. *Brown, L. G.* A survey of image registration techniques / L. G. Brown // ACM computing surveys (CSUR). — 1992. — T. 24, № 4. — С. 325—376.
17. *Leavers, V.* Which hough transform? / V. Leavers // CVGIP: Image understanding. — 1993. — T. 58, № 2. — С. 250—264.
18. *Illingworth, J.* A survey of the Hough transform / J. Illingworth, J. Kittler // Computer vision, graphics, and image processing. — 1988. — T. 44, № 1. — С. 87—116.
19. A comparison of Hough transform methods / J. Princen [и др.] // Image Processing and its Applications, 1989., Third International Conference on. — IET. 1989. — С. 73—77.
20. *Herout, A.* Review of hough transform for line detection / A. Herout, M. Dubská, J. Havel // Real-Time Detection of Lines and Grids. — Springer, 2013. — С. 3—16.
21. *Mukhopadhyay, P.* A survey of Hough Transform / P. Mukhopadhyay, B. B. Chaudhuri // Pattern Recognition. — 2015. — T. 48, № 3. — С. 993—1010.

22. *Фурсов, В. А.* Локализация контуров объектов на изображениях при вариациях масштаба с использованием преобразования Хафа / В. А. Фурсов, С. А. Бибииков, П. Ю. Якимов // Компьютерная оптика. — 2013. — Т. 37, № 4.
23. *Дегтярева, А.* Преобразование Хафа (Hough transform) / А. Дегтярева, В. Вежневек // компьютерная графика и мультимедиа.—2003.—Выпуск. — 2003. — № 1. — С. 2.
24. *Кудрина, М. А.* Использование преобразования Хафа для обнаружения прямых линий и окружностей на изображении / М. А. Кудрина // Известия Самарского научного центра Российской академии наук. — 2014. — Т. 16, № 4—2.
25. *Merlin, P. M.* A parallel mechanism for detecting curves in pictures / P. M. Merlin, D. J. Farber // IEEE Transactions on Computers. — 1975. — Т. 100, № 1. — С. 96—98.
26. *Kimme, C.* Finding circles by an array of accumulators / C. Kimme, D. Ballard, J. Sklansky // Communications of the ACM. — 1975. — Т. 18, № 2. — С. 120—122.
27. *Guil, N.* Lower order circle and ellipse Hough transform / N. Guil, E. L. Zapata // Pattern Recognition. — 1997. — Т. 30, № 10. — С. 1729—1744.
28. *Chen, T.-C.* An efficient randomized algorithm for detecting circles / T.-C. Chen, K.-L. Chung // Computer Vision and Image Understanding. — 2001. — Т. 83, № 2. — С. 172—191.
29. Comparative study of Hough transform methods for circle finding / H. Yuen [и др.] // Image and vision computing. — 1990. — Т. 8, № 1. — С. 71—77.
30. *Ballard, D. H.* Generalizing the Hough transform to detect arbitrary shapes / D. H. Ballard // Pattern recognition. — 1981. — Т. 13, № 2. — С. 111—122.
31. *Davis, L. S.* Hierarchical generalized Hough transforms and line-segment based generalized Hough transforms / L. S. Davis // Pattern Recognition. — 1982. — Т. 15, № 4. — С. 277—285.
32. *Jeng, S.-C.* Fast generalized Hough transform / S.-C. Jeng, W.-H. Tsai // Pattern Recognition Letters. — 1990. — Т. 11, № 11. — С. 725—733.

33. *Ulrich, M.* Real-time object recognition using a modified generalized Hough transform / M. Ulrich, C. Steger, A. Baumgartner // Pattern Recognition. — 2003. — Т. 36, № 11. — С. 2557—2570.
34. *Xu, L.* A new curve detection method: randomized Hough transform (RHT) / L. Xu, E. Oja, P. Kultanen // Pattern recognition letters. — 1990. — Т. 11, № 5. — С. 331—338.
35. *Xu, L.* Randomized Hough transform (RHT): basic mechanisms, algorithms, and computational complexities / L. Xu, E. Oja // CVGIP: Image understanding. — 1993. — Т. 57, № 2. — С. 131—154.
36. *Garey, M. R.* Computers and intractability: a guide to the theory of NP-completeness. 1979 / M. R. Garey, D. S. Johnson // San Francisco, LA: Freeman. — 1979. — Т. 58.
37. *Pippenger, N.* On the evaluation of powers and monomials / N. Pippenger // SIAM Journal on Computing. — 1980. — Т. 9, № 2. — С. 230—250.
38. Exact Fast Algorithm For Optimal Linear Separation Of 2D Distribution. / E. Ershov [и др.] // 29 European Conference on Modelling and Simulation. — 2015. — С. 469—474.
39. *Hu, Z.* The three conditions of a good line parameterization / Z. Hu, S. De Ma // Pattern Recognition Letters. — 1995. — Т. 16, № 4. — С. 385—388.
40. *Biland, H. P.* The recognition and volumetric description of three-dimensional polyhedral scenes by analysis of Hough-space structures : дис. . . . канд. / Biland Hans Peter. — ETH Zurich, 1987.
41. The cascaded Hough transform as an aid in aerial image interpretation / T. Tuytelaars [и др.] // Computer Vision, 1998. Sixth International Conference on. — IEEE. 1998. — С. 67—72.
42. *Wallace, R. S.* A modified Hough transform for lines. / R. S. Wallace // proc. CVPR. — 1985.
43. *Bresenham, J. E.* Algorithm for computer control of a digital plotter / J. E. Bresenham // IBM Systems journal. — 1965. — Т. 4, № 1. — С. 25—30.
44. *Forman, A. V.* A modified Hough transform for detecting lines in digital imagery / A. V. Forman // Applications of artificial intelligence III. Т. 635. — International Society for Optics, Photonics. 1986. — С. 151—161.

45. *Natterer, F.* The Mathematics of Computerized Tomography Wiley / F. Natterer // New York. — 1986.
46. *Sewisy, A. A.* Graphical techniques for detecting lines with the Hough transform / A. A. Sewisy // International journal of computer mathematics. — 2002. — T. 79, № 1. — C. 49–64.
47. *Dubská, M.* PClines—line detection using parallel coordinates / M. Dubská, A. Herout, J. Havel // Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. — IEEE. 2011. — C. 1489–1494.
48. *d’Ocagne, M.* Coordonnées parallèles & axiales: méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles / M. d’Ocagne. — Gauthier-Villars, 1885.
49. *Bhattacharya, P.* Point-to-line mappings as Hough transforms / P. Bhattacharya, A. Rosenfeld, I. Weiss // Pattern Recognition Letters. — 2002. — T. 23, № 14. — C. 1705–1710.
50. *O’Rourke, J.* Dynamically Quantized Spaces for Focusing the Hough Transform. / J. O’Rourke // IJCAI. T. 81. — 1981. — C. 24–28.
51. *Sloan Jr, K. R.* Dynamically Quantized Pyramids. / K. R. Sloan Jr // IJCAI. — 1981. — C. 734–736.
52. *O’Rourke, J.* Dynamic quantization: Two adaptive data structures for multidimensional spaces / J. O’Rourke, K. R. Sloan // IEEE transactions on pattern analysis and machine intelligence. — 1984. — № 3. — C. 266–280.
53. *Blanford, R. P.* Dynamically quantized pyramids for Hough vote collection / R. P. Blanford // Proc. IEEE Workshop Computer Architecture Pattern Anal. Machine Intell. — 1987. — C. 145–152.
54. *Adiv, G.* Recovering 2-D Motion Parameters in Scenes Containing Multiple Moving Objects. Tex. отч. / G. Adiv ; MASSACHUSETTS UNIV AMHERST DEPT OF COMPUTER ; INFORMATION SCIENCE. — 1983.
55. *Silberberg, T. M.* Object recognition using oriented model points / T. M. Silberberg, D. A. Harwood, L. S. Davis // Computer Vision, Graphics, and Image Processing. — 1986. — T. 35, № 1. — C. 47–71.

56. *Li, H.* Fast Hough transform: A hierarchical approach / H. Li, M. A. Lavin, R. J. Le Master // Computer Vision, Graphics, and Image Processing. — 1986. — T. 36, № 2/3. — C. 139—161.
57. *Illingworth, J.* The adaptive Hough transform / J. Illingworth, J. Kittler // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1987. — № 5. — C. 690—698.
58. *Illingworth, J.* Shape detection using the adaptive Hough transform / J. Illingworth, J. Kittler, J. Princen // Real-time object measurement and classification. — Springer, 1988. — C. 119—142.
59. *Brown, C. M.* Advanced Hough transform implementations : тех. отч. / C. M. Brown, M. B. Curtiss, D. B. Sher ; ROCHESTER UNIV NY DEPT OF COMPUTER SCIENCE. — 1983.
60. *Brown, C. M.* Hierarchical cache accumulators for sequential mode estimation : тех. отч. / C. M. Brown ; ROCHESTER UNIV NY DEPT OF COMPUTER SCIENCE. — 1983.
61. *Risse, T.* Hough transform for line recognition: complexity of evidence accumulation and cluster detection / T. Risse // Computer Vision, Graphics, and Image Processing. — 1989. — T. 46, № 3. — C. 327—345.
62. *Koplowitz, J.* The number of digital straight lines on an  $N^* N$  grid / J. Koplowitz, M. Lindenbaum, A. Bruckstein // IEEE Transactions on Information Theory. — 1990. — T. 36, № 1. — C. 192—197.
63. *Shapiro, S.* Aspects of transform method for curve detection / S. Shapiro // Joint Workshop on Pattern Recognition and Artificial Intelligence. — 1976. — C. 90—97.
64. *Shapiro, S. D.* Geometric constructions for predicting Hough transform performance / S. D. Shapiro, A. Iannino // IEEE transactions on pattern analysis and machine intelligence. — 1979. — № 3. — C. 310—317.
65. *Shapiro, S. D.* Use of the Hough transform for image data compression / S. D. Shapiro // Pattern Recognition. — 1980. — T. 12, № 5. — C. 333—337.
66. *Cohen, M.* On the detection of structures in noisy pictures. / M. Cohen, G. T. Toussaint // Pattern Recognition. — 1977. — T. 9, № 2. — C. 95—98.

67. *Van Veen, T.* Discretization errors in the Hough transform / T. Van Veen, F. C. Groen // Pattern recognition. — 1981. — Т. 14, № 1—6. — С. 137—145.
68. *Kiryati, N.* Digital or analog Hough transform? / N. Kiryati, M. Lindenbaum, A. M. Bruckstein // Pattern Recognition Letters. — 1991. — Т. 12, № 5. — С. 291—297.
69. *Niblack, W.* On improving the accuracy of the Hough transform / W. Niblack, D. Petkovic // Machine vision and applications. — 1990. — Т. 3, № 2. — С. 87—106.
70. *Гельфанд, И.* Избранные задачи интегральной геометрии / И. Гельфанд, С. Гиндикин, М. Граев. — Добросвет М., 2000.
71. *Radon, J.* Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten / J. Radon // Berichte über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften zu Leipzig. — 1917.
72. *Penrose, R.* Twistor algebra / R. Penrose // Journal of Mathematical physics. — 1967. — Т. 8, № 2. — С. 345—366.
73. *Gordon, R.* A tutorial on ART (algebraic reconstruction techniques) / R. Gordon // IEEE Transactions on Nuclear Science. — 1974. — Т. 21, № 3. — С. 78—93.
74. Effective regularized algebraic reconstruction technique for computed tomography / V. Prun [и др.] // Crystallography Reports. — 2013. — Т. 58, № 7. — С. 1063—1066.
75. *Deans, S. R.* Hough transform from the Radon transform / S. R. Deans // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 1981. — Т. 1, № 2. — С. 185—188.
76. *Ginkel, M. van.* A short introduction to the Radon and Hough transforms and how they relate to each other / M. van Ginkel, C. L. Hendriks, L. J. van Vliet // Quantitative Imaging Group, Imaging Science & Technology Department, TU Delft. — 2004.
77. *Princen, J.* A formal definition of the Hough transform: Properties and relationships / J. Princen, J. Illingworth, J. Kittler // Journal of Mathematical Imaging and Vision. — 1992. — Т. 1, № 2. — С. 153—168.

78. *Kiryati, N.* Antialiasing the Hough transform / N. Kiryati, A. M. Bruckstein // CVGIP: Graphical Models and Image Processing. — 1991. — Т. 53, № 3. — С. 213—222.
79. Предварительная обработка изображений на основе трейс-преобразований / Н. Федотов [и др.] // Труды Международного симпозиума «Надежность и качество». — 2011. — Т. 2.
80. *Fedotov, N.* Trace transform of three-dimensional objects: Recognition, analysis, and database search / N. Fedotov, S. Ryndina, A. Semov // Pattern recognition and image analysis. — 2014. — Т. 24, № 4. — С. 566—574.
81. *Kiryati, N.* A probabilistic Hough transform / N. Kiryati, Y. Eldar, A. M. Bruckstein // Pattern recognition. — 1991. — Т. 24, № 4. — С. 303—316.
82. *Galamhos, C.* Progressive probabilistic Hough transform for line detection / C. Galamhos, J. Matas, J. Kittler // Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. Т. 1. — IEEE. 1999. — С. 554—560.
83. *Ben-Tzvi, D.* A combinatorial Hough transform / D. Ben-Tzvi, M. B. Sandler // Pattern Recognition Letters. — 1990. — Т. 11, № 3. — С. 167—174.
84. *Kiryati, N.* Randomized or probabilistic Hough transform: unified performance evaluation / N. Kiryati, H. Kälviäinen, S. Alaoutinen // Pattern Recognition Letters. — 2000. — Т. 21, № 13/14. — С. 1157—1164.
85. Probabilistic and non-probabilistic Hough transforms: overview and comparisons / H. Kälviäinen [и др.] // Image and vision computing. — 1995. — Т. 13, № 4. — С. 239—252.
86. *Fernandes, L. A.* Real-time line detection through an improved Hough transform voting scheme / L. A. Fernandes, M. M. Oliveira // Pattern recognition. — 2008. — Т. 41, № 1. — С. 299—314.
87. *Лабунец, В. Г.* Алгебраическая теория сигналов и систем: Цифровая обработка сигналов / В. Г. Лабунец. — Изд-во Красноярского университета, 1984.
88. New fast algorithms of multidimensional Fourier and Radon discrete transforms / E. V. Labunets [и др.] // icassp. — IEEE. 1999. — С. 3193—3196.

89. *Yang, D.* Fast discrete Radon transform and 2-D discrete Fourier transform / D. Yang // *Electronics Letters*. — 1990. — Т. 26, № 8. — С. 550—551.
90. Fast slant stack: A notion of radon transform for data in a cartesian grid which is rapidly computible, algebraically exact, geometrically faithful and invertible / A. Averbuch [и др.] // *SIAM Scientific Computing*. — 2001.
91. Artificial neural network implementation of the Hough transform / E. McVey [и др.] // *Signals, Systems and Computers*, 1989. Twenty-Third Asilomar Conference on. Т. 1. — IEEE. 1989. — С. 128—132.
92. *Chan, C.* A neural network shape recognition system with Hough transform input feature space / C. Chan, M. Sandler // *Image Processing and its Applications*, 1992., International Conference on. — IET. 1992. — С. 197—200.
93. *Basak, J.* Hough transform network / J. Basak, S. K. Pal. — 1999.
94. *Brady, M. L.* Fast parallel discrete approximation algorithms for the Radon transform / M. L. Brady, W. Yong // In *Proceedings of the fourth annual ACM symposium on Parallel algorithms and architectures*. ACM. — 1992. — С. 91—99.
95. *Gotz, W. A.* Eine Schnelle Diskrete Radon Transformation basierend auf rekursiv definierten Digitalen Geraden / W. A. Gotz // *Dissertation*. — 1993. — University of Innsbruck.
96. *Gotz, W. A.* A fast digital Radon transform—An efficient means for evaluating the Hough transform. / W. A. Gotz, H. J. Druckmuller // *Pattern Recognition*. — 1995. — Т. 28.12. — С. 1985—1992.
97. *E, V. J.* Fast linear Hough transform / V. J. E // *Application Specific Array Processors*, 1994. *Proceedings. International Conference on*. — IEEE. — 1994. — С. 1—9.
98. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения / Т. С. Чернов [и др.] // *Вестник Российского фонда фундаментальных исследований*. — 2016. — № 4. — С. 55—71.
99. *Hinds, S. C.* A document skew detection method using run-length encoding and the Hough transform / S. C. Hinds, J. L. Fisher, D. P. D'Amato // [1990] *Proceedings. 10th International Conference on Pattern Recognition*. Т. 1. — IEEE. 1990. — С. 464—468.

100. Анализ особенностей использования стационарных и мобильных мало-размерных цифровых видео камер для распознавания документов / В. В. Арлазаров [и др.] // Информационные технологии и вычислительные системы. — 2014. — Т. 3. — С. 71—81.
101. *Bezmaternykh, P.* Textual blocks rectification method based on fast Hough transform analysis in identity documents recognition / P. Bezmaternykh, D. Nikolaev, V. Arlazarov // Tenth International Conference on Machine Vision (ICMV 2017). Т. 10696. — International Society for Optics, Photonics. 2018. — С. 1069606.
102. A Hough transform based technique for text segmentation / S. Saha [и др.] // arXiv preprint arXiv:1002.4048. — 2010.
103. *Cathey, F.* A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras / F. Cathey, D. Dailey // Intelligent Vehicles Symposium, 2005. Proceedings. IEEE. — IEEE. 2005. — С. 777—782.
104. *Krohina, D.* Usage of Otsu criterion in geometric cut space for horizon line search (in Russian) / D. Krohina, V., V. Postnikov // 37th Conference and School on Information Technologies and Systems (ITaS 2014). — 2014. — С. 303—307.
105. *Zheltoy, S. Y.* Robust computer image analysis for flight vehicles navigation and guidance / S. Y. Zheltoy, Y. V. Vizilter // 16th IFAC symposium on automatic control in aerospace. Т. 2. — 2004. — С. 164—167.
106. *Cowart, A. E.* The detection of unresolved targets using the Hough transform / A. E. Cowart, W. E. Snyder, W. H. Ruedger // Computer Vision, Graphics, and Image Processing. — 1983. — Т. 21, № 2. — С. 222—238.
107. *Mitiche, A.* Computation and analysis of image motion: A synopsis of current problems and methods / A. Mitiche, P. Bouthemy // International journal of computer vision. — 1996. — Т. 19, № 1. — С. 29—55.
108. *Lutton, E.* Contribution to the determination of vanishing points using Hough transform / E. Lutton, H. Maitre, J. Lopez-Krahe // IEEE transactions on pattern analysis and machine intelligence. — 1994. — Т. 16, № 4. — С. 430—438.

109. Hough transform: underestimated tool in the computer vision field / D. Nikolaev [и др.] // Proceedings of the 22th European Conference on Modelling and Simulation. — 2008. — С. 238—246.
110. *Kunina, I. A.* Blind compensation of radial distortion in a single image using fast Hough transform / I. A. Kunina, S. A. Gladilin, D. P. Nikolaev // Computer Optics. — 2016. — Т. 40, № 3. — С. 395—403.
111. *Toft, P.* The Radon transform. Theory and implementation / P. Toft. — 1996.
112. *Cho, Z.-H.* Foundations of medical imaging / Z.-H. Cho, J. P. Jones, M. Singh. — Wiley New York, 1993.
113. *Фурсов, В. А.* Локализация контуров объектов на изображениях при вариациях масштаба с использованием преобразования Хафа / В. А. Фурсов, С. А. Бибииков, П. Ю. Якимов // Компьютерная оптика. — 2013. — Т. 37, № 4.
114. *Muniz, R.* A robust software barcode reader using the Hough transform / R. Muniz, L. Junco, A. Otero // Information Intelligence and Systems, 1999. Proceedings. 1999 International Conference on. — IEEE. 1999. — С. 313—319.
115. *Долгий, А.* Гибридная модель интерпретации деформаций в балластной призме и основной площадке земляного полотна на основе целевого преобразования Хафа и нейронной сети Кохонена / А. Долгий, А. Хатламджиян // Известия Южного федерального университета. Технические науки. — 2007. — Т. 77, № 2.
116. Морфологическое сравнение по форме точечных паттернов и контурных изображений на основе преобразования Хафа и его модификаций / А. Рубис [и др.] // Вестник компьютерных и информационных технологий. — 2011. — № 7. — С. 9—16.
117. *Vandame, B.* Fast Hough transform for robust detection of satellite tracks / B. Vandame // Mining the Sky. — Springer, 2001. — С. 595—597.
118. *Семенов, А.* Обнаружение радиолокационных целей с помощью преобразования Хафа / А. Семенов // Наука и образование: научное издание МГТУ им. НЭ Баумана. — 2014. — № 12.

119. *Carlson, B.* Search radar detection and track with the Hough transform. III. Detection performance with binary integration / B. Carlson, E. Evans, S. Wilson // IEEE transactions on Aerospace and Electronic Systems. — 1994. — Т. 30, № 1. — С. 116—125.
120. *Мозговой, А.* Преобразование Хафа в задачах автоматического распознавания рукописного текста / А. Мозговой // Вестник Воронежского института высоких технологий. — 2012. — № 9. — С. 62—64.
121. *Дудкин, А.* Выделение контуров на полутоновых изображениях топологических слоев интегральных микросхем / А. Дудкин, Д. Вершок, А. Селиханович // Искусственный интеллект. — 2004. — № 3. — С. 453—458.
122. *Григорьев, А.* Определение количества осей транспортного средства по видеоряду проезда / А. Григорьев, Т. Ханипов, Д. Николаев // 54-й научной конференции МФТИ. Т. 10. — 2011. — С. 31.
123. *Гаганов, В.* Трехмерная реконструкция плоских граней прозрачных минералов по набору изображений с микроскопа / В. Гаганов, А. Игнатенко, М. Ломоносова // Труды конференции Графикон. — 2008. — С. 227—233.
124. Векторизация растровых изображений с использованием преобразования Хафа / М. Кудрина [и др.] // Труды Международного симпозиума «Надежность и качество». — 2013. — Т. 1.
125. *Skingley, J.* The Hough transform applied to SAR images for thin line detection / J. Skingley, A. Rye // Pattern Recognition Letters. — 1987. — Т. 6, № 1. — С. 61—67.
126. *Лагутин, М.* Наглядная математическая статистика / М. Лагутин // М.: Бином. Лаборатория знаний. — 2013.
127. *Yan, X.* Linear regression analysis: theory and computing / X. Yan, X. Su. — World Scientific, 2009.
128. *Deming, W. E.* Statistical adjustment of data. / W. E. Deming. — 1943.
129. *Markovsky, I.* Overview of total least-squares methods / I. Markovsky, S. Van Huffel // Signal processing. — 2007. — Т. 87, № 10. — С. 2283—2302.
130. *Van Huffel, S.* The total least squares problem: computational aspects and analysis. Т. 9 / S. Van Huffel, J. Vandewalle. — Siam, 1991.

131. *Huber, P. J.* Robust statistics / P. J. Huber. — Springer, 2011.
132. *Khurram Aftab, R. H.* Convergence of Iteratively Re-weighted Least Squares to Robust M-estimators / R. H. Khurram Aftab // 2015 IEEE Winter Conference on Applications of Computer Vision. — 2015. — С. 480—487.
133. *Ballester, P.* Hough transform for robust regression and automated detection / P. Ballester // Astronomy and Astrophysics. — 1994. — Т. 286. — С. 1011—1018.
134. *Kiryati, N.* Heteroscedastic Hough transform (HtHT): An efficient method for robust line fitting in the ‘errors in the variables’ problem / N. Kiryati, A. M. Bruckstein // Computer Vision and Image Understanding. — 2000. — Т. 78, № 1. — С. 69—83.
135. *Безматерных, П.* Решение задачи линейной регрессии с помощью быстрого преобразования Хафа / П. Безматерных, Т. Ханипов, Д. Николаев // Информационные технологии и системы (ИТиС’12). Петрозаводск. — 2012. — С. 354.
136. *Niblack, W.* On improving the accuracy of the Hough transform: theory, simulations, and experiments / W. Niblack, D. Petkovic // Computer Vision and Pattern Recognition, 1988. Proceedings CVPR’88., Computer Society Conference on. — IEEE. 1988. — С. 574—579.
137. *Асватов, Е. Н.* Робастная ортогональная линейная регрессия для маломерных гистограмм / Е. Н. Асватов, Е. И. Ершов, Д. П. Николаев // Сенсорные системы. — 2017. — Т. 31, № 4. — С. 331—342.
138. *Васильев, Ф.* Численные методы решения экстремальных задач / Ф. Васильев. — Наука М., 1988.
139. *Rousseeuw, P. J.* Robust regression and outlier detection. Т. 589 / P. J. Rousseeuw, A. M. Leroy. — John wiley & sons, 2005.
140. *Siegel, C. L.* Uber einige anwendungen diophantischer approximationen / C. L. Siegel // Abh. Pruess. Akad. Wiss., Phys.-Math. Kl. — 1929. — Т. 1. — С. 14—72.
141. *Казанцев, И. Г.* Численные и геометрические методы математического моделирования в многомерных задачах томографии и обработки изображений : дис. ... канд. / Казанцев Иван Гаврилович. — Диссертация на соискание ученой степени доктора физико-математических наук ..., 2014.

142. *Bracewell, R. N.* Strip integration in radio astronomy / R. N. Bracewell // Australian Journal of Physics. — 1956. — Т. 9, № 2. — С. 198—217.
143. *Otsu, N.* A threshold selection method from gray-level histograms / N. Otsu // IEEE transactions on systems, man, and cybernetics. — 1979. — Т. 9, № 1. — С. 62—66.
144. *Vala, M. H. J.* A review on Otsu image segmentation algorithm / M. H. J. Vala, A. Baxi // International Journal of Advanced Research in Computer Engineering & Technology (IJARCET). — 2013. — Т. 2, № 2. — pp—387.
145. *Gong, J.* Fast recursive algorithms for two-dimensional thresholding / J. Gong, L. Li, W. Chen // Pattern recognition. — 1998. — Т. 31, № 3. — С. 295—300.
146. *Kurita, T.* Maximum likelihood thresholding based on population mixture models / T. Kurita, N. Otsu, N. Abdelmalek // Pattern recognition. — 1992. — Т. 25, № 10. — С. 1231—1240.
147. *Jianzhuang, L.* Automatic thresholding of gray-level pictures using two-dimension Otsu method / L. Jianzhuang, L. Wenqing, T. Yupeng // Circuits and Systems, 1991. Conference Proceedings, China., 1991 International Conference on. — IEEE. 1991. — С. 325—327.
148. *Zhang, J.* Image segmentation based on 2D Otsu method with histogram analysis / J. Zhang, J. Hu // Computer Science and Software Engineering, 2008 International Conference on. Т. 6. — IEEE. 2008. — С. 105—108.
149. Modified two-dimensional Otsu image segmentation algorithm and fast realisation / Q. Chen [и др.] // IET Image Processing. — 2012. — Т. 6, № 4. — С. 426—433.
150. *Li, L.* Gray level image thresholding based on fisher linear projection of two-dimensional histogram / L. Li, R. Gong, W. Chen // Pattern Recognition. — 1997. — Т. 30, № 5. — С. 743—749.
151. Image segmentation using thresholding and swarm intelligence / Z. Ye [и др.] // Journal of Software. — 2012. — Т. 7, № 5. — С. 1074—1082.
152. *Ершов, Е. И.* Алгоритм бинарной линейной кластеризации маломерных гистограмм / Е. И. Ершов // Сенсорные системы. — 2017. — Т. 31, № 3. — С. 261—269.

153. *Ершов, Е. И.* Критерий и алгоритм кластеризации при обработке изображений / Е. И. Ершов, Н. Д.П.. // . — 57-я научная конференция МФТИ. 2017. — С. 45—47.
154. *Ershov, E.* Fast Hough Transform and approximation properties of dyadic patterns / E. Ershov, S. Karpenko // arXiv preprint arXiv:1712.05615. — 2017.

## Список рисунков

1.1	Диадический паттерн $\langle 1, 4 \rangle$ в $st$ -параметризации системы $Y_d^{2^3, 2}$ . . . . .	34
1.2	Два диадических паттерна с наклонами, отличающимися на единицу. . . . .	41
1.3	Иллюстрация $(s, t)$ -параметризации плоскости в трехмерном пространстве. . . . .	42
1.4	Пример структуры диадической плоскости. Зеленым цветом отмечены пиксели, задающие плоскость. . . . .	42
1.5	Сравнения времени вычисления ТБПХ для плоскостей и алгоритма полного подсчета сумм по диадическим плоскостям для изображений разного размера. Показания приведены в логарифмической шкале. . . . .	46
1.6	Иллюстрация параметризации преимущественно-параллельной оси $z$ прямой в кубе. . . . .	48
1.7	Иллюстрация структуры диадической прямой в кубе. Зеленым отмечены воксели, задающие диадическую прямую. . . . .	48
2.1	Функция изображения $I$ для одной точки с координатами $(z_1^*, z_2^*)$ в двумерном пространстве . . . . .	59
2.2	Пространство радона функции $I$ . . . . .	59
2.3	Свертка пространства Радона вдоль координаты $\theta_0$ . . . . .	59
2.4	Пример тестовой двумерной гистограммы . . . . .	69
2.5	Дуги $l_1$ и $l_2$ между пересечениями прямых и описанной вокруг гистограммы окружности . . . . .	70
2.6	Гистограмма ошибок аппроксимации истинной гиперплоскости наилучшими дискретными прямыми . . . . .	72
2.7	Гистограмма ошибок аппроксимации истинной гиперплоскости диадическими паттернами . . . . .	72
2.8	График отношения $\frac{\Phi_{FHT}}{\Phi_{HT}}$ эмпирических функций распределения ошибок . . . . .	73
2.9	Зависимость значения метрики оценок от веса $\frac{w_2}{w_1+w_2} 100\%$ выбросового кластера . . . . .	74

2.10	Зависимость расстояния по метрике $L_2$ параметров истинной и восстановленных гиперплоскостей от веса $\frac{w_2}{w_1+w_2}100\%$ выбросового кластера . . . . .	74
3.1	Иллюстрация результата работы алгоритма бинарной кластеризации Дж. Гонг [145]. Белый прямоугольник – оптимальный ответ метода, зелёная штриховка – оптимальный ответ.	79
3.2	Иллюстрация результата работы алгоритма бинарной кластеризации Дж. Гонг [145] и предложенного в диссертации алгоритма. . . . .	85
3.3	Иллюстрация предложенного в диссертации алгоритма в трехмерном случае для критерия (3.7). . . . .	86
3.4	Погрешность вычисления суммы по светло-серому паттерну вычисляется как модуль разности сумм по тёмно-серым паттернам. .	87
A.1	Зависимость максимальной ошибки аппроксимации диадическим паттерном от логарифма размера изображения. . . . .	110
A.2	Зависимость максимальной ошибки аппроксимации для каждого наклона для размеров изображений $2^6, 2^7, 2^8$ соответственно: по горизонтальной оси - наклон, по вертикальной - величина ошибки. .	111
A.3	Распределение ошибки аппроксимации всех диадических паттернов для изображения с линейным размером $2^{11}, 2^{12}, 2^{13}$ соответственно. .	111

## Список таблиц

1	Параметры модели данных. . . . .	68
---	----------------------------------	----

## Приложение А

### Доказательство теоремы об оценке точности ортотропной ошибки аппроксимации диадическим паттерном

Результаты, изложенные в этом приложении, получены совместно с С.М. Карпенко и приняты в журнал ППИ [6; 154]. Полученная теоретическая оценка максимального ортотропного разброса (максимально ошибки аппроксимации) диадического паттерна, используется, но не заявляется в качестве результата.

Для начала приведём серию качественных и количественных результатов вычислительных экспериментов по исследованию структуры и точности аппроксимации диадического паттерна.

Ошибка аппроксимации вычислялась по формуле 9. На графике А.1 показана зависимость величины максимальной ошибки аппроксимации от логарифма изображения по основанию 2. Видно, что данная ошибка мажорируется прямой  $\frac{k}{6}$ , что подтверждает результаты полученные в работах [2; 96]. На основании полученных экспериментальных данных удалось показать, что эта оценка является точной для изображений со стороной  $2^{2m}$  пикселей, для изображений со стороной  $2^{2m+1}$  точная оценка -  $\frac{k}{6} - \frac{2^k+1}{18(2^k-1)}$  [4], где  $w \in [0, \frac{k}{2}]$ .

Кроме того, производилось исследование величины максимальной ошибки для каждого наклона диадического паттерна. Данная ошибка имеет фрактальную природу, при том для чётных  $k$  ошибка достигает пика только в двух точках, в то время как для нечетных  $k$  - имеется  $2(k-1)$  пиков (смотри рисунок А.2.).

Получено распределение ошибки аппроксимации диадическими паттернами всех наклонов  $t$ . Экспериментально показано, что дисперсия ошибки аппроксимации диадическими паттернами не зависит от размера изображения.

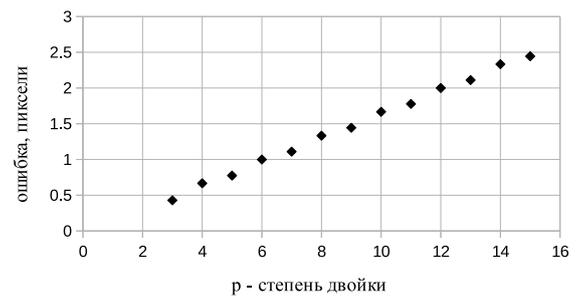


Рисунок А.1 — Зависимость максимальной ошибки аппроксимации диадическим паттерном от логарифма размера изображения.

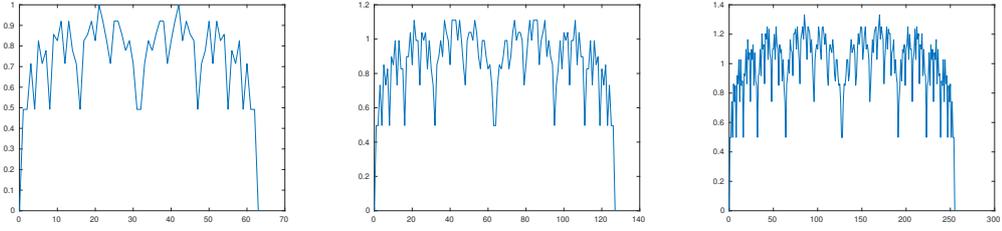


Рисунок А.2 — Зависимость максимальной ошибки аппроксимации для каждого наклона для размеров изображений  $2^6$ ,  $2^7$ ,  $2^8$  соответственно: по горизонтальной оси - наклон, по вертикальной - величина ошибки.

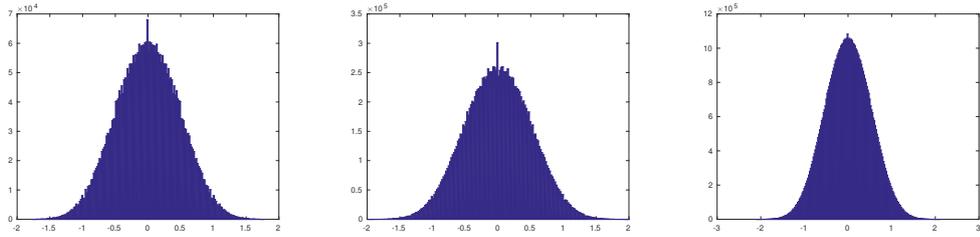


Рисунок А.3 — Распределение ошибки аппроксимации всех диадических паттернов для изображения с линейным размером  $2^{11}$ ,  $2^{12}$ ,  $2^{13}$  соответственно.

Теоретически покажем, что величина ортотропной ошибки аппроксимации никогда не превышает  $\frac{k}{6}$ .

Запишем выражение ошибки аппроксимации диадическим паттерном, согласно выражению 9:

$$E(x, t) = \sum_{r=0}^{k-1} t_r E_r(x), \quad (\text{A.1})$$

где  $E_r(x) = \left( \left[ \frac{2^r x}{2^k - 1} \right] - \frac{2^r x}{2^k - 1} \right)$ .

Заметим, что для фиксированного  $x$  модуль ошибки  $E(x, t)$  будет максимален, когда  $t_r \neq 0$  только для слагаемых одного знака. Так как согласно утверждению 2.3 ошибка аппроксимации симметрична, далее будем рассматривать только такие  $(r, x)$  для которых  $E_r(x) < 0$ . Иными словами, достаточно выбрать те  $r$ , при которых двоичное представление числа  $s_r = 2^r x$  длины  $k$  имеет ноль в старшем разряде. Двоичное представление  $s_{r+1}$  может быть получено из  $s_r$  циклическим сдвигом на один разряд.

Поэтому сформируем таблицу  $T(x)$  всевозможных циклических сдвигов двоичной записи  $x$ :

$$\begin{array}{cccc}
x_{k-1} & x_{k-2} & \dots & x_0 \\
x_{k-2} & x_{k-3} & \dots & x_{k-1} \\
x_{k-3} & x_{k-4} & \dots & x_{k-2} \\
\dots & \dots & \dots & \dots \\
x_0 & x_{k-1} & \dots & x_1
\end{array}$$

Определим функцию от таблицы  $S(T(x))$  как арифметическую сумму по строкам с нулевым старшим битом. Итак, поиск максимума модуля ошибки эквивалентен решению задачи:

$$S(T(x)) \rightarrow \max_x \quad (\text{A.2})$$

**Определение 20.** *Максимайзером будем называть любой  $\hat{x}$ , такой, что*

$$\hat{x} = \operatorname{argmax}_{x \in [0, 2^k - 1]} S(T(x)). \quad (\text{A.3})$$

Проанализируем структуру таблицы  $T(x)$ . Символами  $\infty$  и  $\text{---}\ddagger$  обозначим произвольную двоичную последовательность. Символами  $\overline{\infty}$  и  $\overline{\text{---}\ddagger}$  обозначим инверсию, соответственно, последовательностей  $\infty$  и  $\text{---}\ddagger$ , записанных в обратном порядке. Например, если  $\infty = 010000$ ,  $\overline{\infty} = 111101$ .

По построению, для любого числа  $x$  таблица  $T(x)$  будет иметь следующую структуру, где  $\infty = x_{k-2}, \dots, x_p$ :

$$\begin{array}{cccc}
x_{k-1} & & \infty & \\
& & x_{k-3} & \dots & x_{k-1} \\
& \text{⌈} & x_{k-4} & \dots & x_{k-2} \\
& & \dots & \dots & \dots \\
& & x_{k-1} & \dots & x_1,
\end{array}$$

заметим, что битовое представление  $x$  содержится как в первой строке, так и в первом столбце.

**Лемма 1.** *О замене.*

*Пусть символ  $T_\varepsilon(x)$  обозначает таблицу вида с  $x_{k-1} = \varepsilon$ :*

$$\begin{array}{cccc}
\varepsilon & & \infty & \\
& x_{k-3} & \dots & \varepsilon \\
\textcircled{\text{R}} & x_{k-4} & \dots & x_{k-2} \\
& \dots & \dots & \dots \\
& \varepsilon & \dots & x_1,
\end{array}$$

тогда  $S(T_0(x)) > S(T_1(x))$  эквивалентно  $\infty > \overline{\infty}$ .

*Доказательство.* Для доказательства леммы достаточно показать, что

$$S(T_0(x)) - S(T_1(x)) = \infty - \overline{\infty}.$$

Так как суммирование производится только по строкам с нулевым старшим битом, то разность первых строк равна  $\infty$ . Далее, для пары строк с номером  $r > 1$  в таблице  $T_0(x)$  значение в разряде  $k - r$  равно нулю, в то время как в правой таблице  $T_1(x)$  – единице, значения в остальных разрядах строк совпадают, т.е. вклад такой пары строк есть отрицательное число с единственной единицей в  $k - r$  разряде. Тогда сумма вкладов всех таких пар строк есть  $-\overline{\infty}$ . Что и требовалось доказать.  $\square$

**Лемма 2.** *О перестановке.*

Пусть символ  $1_k$  обозначает битовую последовательность единиц длины  $k$ , а символ  $T_{\varepsilon\delta}(x)$  – таблицу вида:

$$\begin{array}{cccc}
\varepsilon & \delta & & \infty \\
\delta & x_{k-3} & \dots & \varepsilon \\
& x_{k-4} & \dots & \delta \\
\textcircled{\text{R}} & \dots & \dots & \dots \\
& \varepsilon & \dots & x_1,
\end{array}$$

тогда  $S(T_{01}(x)) > S(T_{10}(x))$  эквивалентно  $\infty + \overline{\infty} < 1_k$ .

*Доказательство.* Аналогично доказательству леммы 1, запишем разность сумм по таблицам:

$$S(T_{01}(x)) - S(T_{10}(x)) = 01\infty - 0\infty1 - \overline{\infty},$$

здесь первое слагаемое есть вклад пары строк с  $r = 0$ , второе слагаемое – пары строк с  $r = 1$ , а третье слагаемое – это вклад оставшихся пар строк, полученный аналогичным лемме 1 способом. После преобразования получим:

$$S(T_{01}(x)) - S(T_{10}(x)) = 1_k - \infty - \overline{\infty},$$

что и требовалось доказать.  $\square$

Далее последовательно проверим всевозможные  $x$  на предмет принадлежности к множеству максимайзеров. Обозначим символом  $\overset{w}{\sim}$  последовательность чередующихся нулей и единиц длины  $w$ , а символом  $\overset{\varepsilon}{\sim}^{\delta}$  последовательность чередующихся нулей и единиц длины  $w$  со значениями  $\varepsilon$ ,  $\delta$  в старшем и младшем разрядах соответственно. Например,  $1 \overset{5}{\sim} 1 = 10101$ .

**Утверждение 1.** *Если в двоичной записи числа  $x$  встречаются три подряд идущих равных бита (далее битовая тройка), то такой  $x$  не может быть максимайзером.*

*Доказательство.* Без ограничения общности рассмотрим число в котором есть как минимум одна битовая тройка единиц  $111 \text{---} \mathbf{1}$ . Пусть  $\hat{x} = 11 \text{---} \mathbf{1}$ . Так как  $T(x)$  содержит все циклические сдвиги, то  $S(T(x)) = S(T(\hat{x}))$ , следовательно, достаточно показать, что

$$S(T(11 \text{---} \mathbf{1})) < S(T(01 \text{---} \mathbf{1})).$$

Обозначим  $\infty = 1 \text{---} \mathbf{1}$ , тогда  $\overline{\infty} = 0 \text{---} \mathbf{1}$ , видно, что  $\infty > \overline{\infty}$ . Следовательно, согласно лемме 1,  $S(T(11 \text{---} \mathbf{1})) < S(T(01 \text{---} \mathbf{1}))$  и  $111 \text{---} \mathbf{1}$  не может быть максимайзером.  $\square$

**Утверждение 2.** *Число вида  $x = 11 \overset{l}{\sim} 0 \overset{w}{\sim} 11 \text{---} \mathbf{1}$  для любых положительных  $l, w$  не может быть максимайзером.*

*Доказательство.* Пусть  $\hat{x}$  получен в результате циклического сдвига  $x$ :

$$\hat{x} = 11 \overset{w}{\sim} 0 \overset{l}{\sim} 11 \text{---} \mathbf{1} \overset{l}{\sim} 0.$$

Обозначим

$$\infty = 1 \overset{w}{\sim} 0 \overset{l}{\sim} 11 \text{---} \mathbf{1} \overset{l}{\sim} 0 = 1 \overset{w+2}{\sim} 1 \text{---} \mathbf{1} \overset{l}{\sim} 0,$$

тогда

$$\overline{\infty} = 1 \overset{l}{\sim} 1 \text{---} \mathbf{0} \overset{w}{\sim} 1 \overset{l}{\sim} 0,$$

Для сравнения  $\infty$  и  $\overline{\infty}$  необходимо рассмотреть три случая:  $w+2 < l$ ,  $w+2 = l$ ,  $w+2 > l$ . Рассмотрим, например, случай, когда  $w+2 < l$ . Будем побитово сравнивать  $\infty$  и  $\overline{\infty}$  начиная со старшего бита. Видно, что в первом несовпавшем разряде  $\infty$  содержит единицу, а  $\overline{\infty}$  - ноль, следовательно,  $\infty > \overline{\infty}$ . Аналогичным образом рассматриваются остальные два случая. Следовательно, согласно лемме 1:

$$S(T(11^0 \overset{l}{\sim} 011^0 \overset{w}{\sim} 011 \text{---} \text{---})) < S(T(01^0 \overset{l}{\sim} 001^0 \overset{w}{\sim} 011 \text{---} \text{---})),$$

то есть  $x$  не является максимайзером.  $\square$

**Утверждение 3.** Если в битовом представлении  $x$  содержится подпоследовательность вида 1100, то  $x$  не может быть максимайзером.

*Доказательство.* Пусть  $x = 100 \text{---} \text{---} 1$ , а  $\hat{x} = 010 \text{---} \text{---} 1$ .

Пусть  $\infty = 0 \text{---} \text{---} 1$ , тогда  $\overline{\infty} = 0 \text{---} \text{---} 1$ . Очевидно, что  $\infty + \overline{\infty} < 1_{k-2}$ , так как у обоих слагаемых старший разряд имеет нулевое значение.

Тогда, согласно лемме 2,  $S(T(x)) < S(T(\hat{x}))$  и, следовательно,  $x$  не может быть максимайзером.  $\square$

**Утверждение 4.** Число вида  $x = 11^0 \overset{w}{\sim} 011^0 \overset{n}{\sim} 100 \text{---} \text{---}$  не может быть максимайзером.

*Доказательство.* Пусть  $\hat{x}$  - циклический сдвиг числа  $x$ :

$$\hat{x} = 10^1 \overset{n-1}{\sim} 100 \text{---} \text{---} 11^0 \overset{w}{\sim} 01.$$

Докажем данное утверждение для  $\hat{x}$ , заметим, что  $x = 10 \infty$ , где

$$\infty = 1^0 \overset{n-2}{\sim} 100 \text{---} \text{---} 11^0 \overset{w}{\sim} 01,$$

тогда

$$\overline{\infty} = 0^1 \overset{w}{\sim} 100 \text{---} \text{---} 11^0 \overset{n-2}{\sim} 10.$$

Видно, что независимо от соотношения величин  $n-2$  и  $w$ ,  $\infty + \overline{\infty} < 1_{k-2}$ . Следовательно, согласно лемме 2,  $S(T(10^1 \overset{n-1}{\sim} 100 \text{---} \text{---} 11^0 \overset{w}{\sim} 01)) < S(T(00^1 \overset{n-1}{\sim} 100 \text{---} \text{---} 11^0 \overset{w}{\sim} 01))$  и  $\hat{x}$  - не максимайзер.  $\square$

**Утверждение 5.** Число вида  $x = 11^0 \overset{w+k}{\sim} 1001^w \overset{w}{\sim} 011$  не может быть максимайзером ( $w > 1$ ).

*Доказательство.* Пусть  $\hat{x}$  - циклический сдвиг числа  $x$ :

$$\hat{x} = 01^0 \overset{w-1}{\sim} 011 \overset{w+k}{\sim} 11^0 \overset{w+k}{\sim} 10,$$

докажем данное утверждение для  $\hat{x}$ , так как  $S(T(x)) = S(T(\hat{x}))$ . Пусть

$$\infty = 0 \overset{w-1}{\sim} 011 \overset{w+k}{\sim} 11^0 \overset{w+k}{\sim} 10,$$

тогда

$$\overline{\infty} = 1 \overset{w+k+1}{\sim} 100 \overset{w-1}{\sim} 001^w.$$

Поскольку для любых положительных  $w, n$  верно  $w + k + 1 > w - 1$ , ясно, что  $\infty + \overline{\infty} > 1_k$ . Следовательно, согласно лемме 2,  $x$  не может быть максимайзером.  $\square$

**Теорема 3.** Максимальная ошибка аппроксимации дивидендом паттерном достигается при  $x = \lfloor \frac{2^k}{3} \rfloor$  и  $x = \lfloor \frac{2^{k+1}}{3} \rfloor$  при наклонах  $t = \lfloor \frac{2^k}{3} \rfloor$  и  $t = \lfloor \frac{2^{k+1}}{3} \rfloor$ , и её модуль равен  $\frac{k}{6}$ , для чётных  $k$ .

*Доказательство.* Ясно, что  $x_1 = 0 \overset{k}{\sim} 1 = \lfloor \frac{2^k}{3} \rfloor$ ,  $x_2 = 0 \overset{k}{\sim} 1 = \lfloor \frac{2^{k+1}}{3} \rfloor$ . Любые другие  $x$  не могут быть максимайзерами, согласно утверждениям 1 – 5. Заметим, что  $x = 11 \overset{n}{\sim} 11 \overset{w}{\sim}$  является частным случаем утверждения 2, а  $x = 11 \overset{n}{\sim} 00 \overset{w}{\sim}$  являются частным случаем утверждения 4.

В силу симметрии (см утверждение 2.3),  $t_1 = \lfloor \frac{2^k}{3} \rfloor$ ,  $t_2 = \lfloor \frac{2^{k+1}}{3} \rfloor$  являются наклонами для данных максимайзеров, а для каждого из этих наклонов пиковая ошибка достигается при  $x_1 = \lfloor \frac{2^k}{3} \rfloor$ ,  $x_2 = \lfloor \frac{2^{k+1}}{3} \rfloor$ .

Рассчитаем теперь значение пиковой ошибки, к примеру для случая  $x_1 = \lfloor \frac{2^k}{3} \rfloor$ . Для этого достаточно вычислить сумму по таблице  $S(T(x_1))$  и поделить результат на  $2^k - 1$ , согласно формуле 9. Сумма по одной значимой строке в таблице есть сумма геометрической прогрессии из  $\frac{k}{2}$  элементов со знаменателем  $q = 4$ , с первым членом, равным 1. Нетрудно вычислить, что сумма по одной строке равна  $s = \frac{4^{k/2} - 1}{3} = \frac{2^k - 1}{3}$ , тогда для всех  $\frac{k}{2}$  значимых сумма по таблице равна  $\frac{k(2^k - 1)}{6}$ . В результате пиковая ошибка аппроксимации, согласно формуле A.1 равна  $\frac{k}{6}$ .

Аналогично, доказательство применимо для остальных пар  $(x, t)$ , где достигается пиковое значение ошибки.  $\square$

**Теорема 4.** *Модуль максимальной ошибки аппроксимации дивидендом паттерном равен  $\frac{k}{6} - \frac{1}{18}$ , для нечётных  $k \rightarrow \infty$  и достигается при  $x = 00^1 \overset{k-2}{\rightsquigarrow} 1$  или  $x = 11^0 \overset{k-2}{\rightsquigarrow} 0$ .*

*Доказательство.* Согласно утверждениям 1-5 для нечетных  $k$  максимайзером может быть только число, в битовом представлении которого содержится только одна пара соседствующих единичных или нулевых битов.

Для определения величины этой ошибки при  $k \rightarrow \infty$  аналогично Теореме 1 подсчитаем сумму по таблице. Пусть  $x_1 = 0 \overset{k-2}{\rightsquigarrow} 011$  и  $t_1 = 0 \overset{k-2}{\rightsquigarrow} 011$ . Видно, что такое число содержит  $w = \frac{k-1}{2}$  нулей и  $n = \frac{k-1}{2} + 1$  единиц.

Для начала вычислим сумму  $S_1$  по первой строке таблицы. Ясно, что  $0 \overset{k-2}{\rightsquigarrow} 011 = 0 \overset{k}{\rightsquigarrow} 0 + 1$ . Заметим, что сумма по  $0 \overset{k}{\rightsquigarrow} 0$  является суммой геометрической прогрессии с первым элементом равным  $a_1 = 2$  и знаменателем  $q = 4$ , и равна  $\frac{2}{3}(2^{k-1} - 1)$ . Тогда  $S^1 = \frac{2}{3}(2^{k-1} - 1) + 1$ .

Заметим, что разность соседних значимых строк  $\Delta_i$  в таблице формирует геометрическую последовательность,  $\Delta_1 = T_1(x_1) - T_3(x_1) = 10$ ,  $\Delta_2 = T_3(x_1) - T_5(x_1) = 1000$ ,  $\Delta_i = 2 \cdot 4^{i-1}$ , где  $i \in [1, \frac{k-1}{2} - 1]$ .

Тогда, значение суммы по каждой значимой строке  $k$  (индекс пробегает по строкам с нулевым старшим битом) можно представить как сумму по первой строке и добавки к каждой строке, равной соответствующей сумме  $S_k = \sum_{i=1}^{k-1} \Delta_i$ . Таким образом, сумма по таблице равна

$$S = \left[ \frac{k-1}{2} \cdot S^1 + \sum_{k=1}^{k-1} S_k \right].$$

Преобразовав данное выражение, получим

$$S = \left[ \frac{k-1}{6}(2^k - 1) + \frac{k-1}{3} + \frac{1}{9}(2^k - 1) - \frac{7}{9} - \frac{2}{3}(k-1) \right]. \quad (\text{A.4})$$

Тогда, в результате деления выражения A.4 на  $2^k - 1$  согласно формуле A.1, и устремления  $k \rightarrow \infty$  получим  $E(x_1, x_2) = \frac{k}{6} - \frac{1}{18}$ , что и требовалось доказать.

Аналогичным образом производится доказательство и для всех остальных пар  $(x, t)$ , где достигается пиковое значение ошибки.  $\square$

В работе проведены вычислительные и аналитические исследования структуры диадического паттерна, в частности ошибка его аппроксимации геометрической прямой. Показано, что пиковая ошибка для изображения линейным размером  $n$  не превосходит  $\frac{\log(n)}{6}$  для случая четных  $n$ , и  $\frac{\log(n)}{6} - \frac{1}{18}$  для нечетных.