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»

Автореферат
диссертации на соискание учёной степени
кандидата физико-математических наук

Москва — 2019

Работа выполнена в Институте проблем передачи информации им. А.А. Харкевича Российской академии наук (ИППИ РАН).

Научный руководитель: кандидат физико-математических наук **Николаев Дмитрий Петрович**.

Официальные оппоненты: **Визильтер Юрий Валетинович,** доктор физико-математических наук, профессор РАН, Федеральное государственное унитарное предприятие «Государственный научно-исследовательский институт авиационных систем», старший научный сотрудник.

Фараджев Игорь Александрович, кандидат физико-математических наук, Федеральный исследовательский центр «Информатика и управление» Российской академии наук Институт системного анализа, главный научный сотрудник.

Ведущая организация: Институт систем обработки изображений РАН – филиал федерального государственного учреждения «Федеральный научно-исследовательский центр «Кристаллография и Фотоника» российской академии наук».

Защита состоится 15 апреля 2019 г. в 11 часов на заседании диссертационного совета Д 002.077.05 при Институте проблем передачи информации им. А.А. Харкевича Российской академии наук (ИППИ РАН) по адресу: 127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1.

С диссертацией можно ознакомиться на сайте http://iitp.ru и в библиотеке Института проблем передачи информации им. А.А. Харкевича РАН.

Отзывы на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 127051, г. Москва, Большой Каретный переулок, д. 19, стр. 1, ученому секретарю диссертационного совета Д 002.077.05.

Автореферат разослан «___» ___________ 2019 г.
Телефон для справок: +7 (985) 979-86-96.

Ученый секретарь
диссертационного совета
Д 002.077.05,
доктор физ.-мат. наук Цитович Иван Иванович

Общая характеристика работы

Актуальность темы. 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 индустрии, 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, М. Брейди, В. Готса, 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. Мукхопадхая и А. Хассенейна 2015 года. Стремительно развиваются также алгоритмы вычисления ПХ для анализа трёхмерных 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 аппроксимации части прямой на изображении используется дискретный паттерн специального вида, так называемый диадический паттерн (ДП)).

Ввиду вышесказанного ясно, что задачи обобщить БПХ на трёхмерный 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 БПХ для трёхмерных изображений, исследовать их свойства, а также разработать на их основе методы анализа двумерных и трёхмерных изображений и гистограмм.

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

1. Исследование свойств диадического паттерна как дискретной модели прямой для двумерных и трёхмерных изображений.
2. Исследование способов обобщения БПХ для трёхмерных изображений.
3. Разработка и исследование методов поиска прямых на двумерных, а также прямых и плоскостей на трёхмерных 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 трёхмерных 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 (а не рекуррентное, как ранее) выражение 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 БПХ 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 обобщённого 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 путём 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 величиной $\frac{\log_2 n}{\sqrt{11}}$, а сверху – $\frac{\log_2 n}{3}$ и, наконец, для диадических прямых ограничен снизу величиной $\frac{\log_2 n}{\sqrt{22}}$, а сверху – $\frac{\sqrt{2}\log_2 n}{6}$, 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 найдётся проходящий через них паттерн; аналогичное верно и для диадических прямых ТБПХ.
3. Предложены два алгоритма вычисления трёхмерного 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 трёхмерных 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 трёхмерного случая для 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 трёхмерных. Приведён 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. Достоверность подтверждается результатами тестирования, проверки 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 Huaweii (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.

Содержание работы

Во **введении** обоснована актуальность диссертационной работы, сформулирована цель и научная новизна исследований, показана практическая значимость полученных результатов, представлены выносимые на защиту научные положения.

В **первой главе** вводятся необходимые определения, предлагаются два обобщения алгоритма Брейди для трёхмерного пространства, устанавливаются их свойства. Вводится модель многомерного изображения, даются определения диадического паттерна (ДП) и системы паттернов для двумерного изображения, а также предлагаются два способа их обобщения для изображений произвольной размерности. В трёхмерном пространстве для каждого из обобщений предлагается алгоритм быстрого вычисления сумм по системе соответствующих паттернов, оценивается их вычислительная сложность. Вводится понятие разброса паттерна, строятся нижние и верхние оценки на максимальный для системы паттернов разброс в пространствах размерности 2 и 3.

Определение 1. *Изображением (m-мерным изображением со сторонами $\langle n\rangle = \langle n_1, n_2, ..., n_m\rangle$ будем называть отображение $I^m_{\langle n\rangle} : \mathbb{Z}^m_{\langle n\rangle} \to G$, где $\mathbb{Z}^m_{\langle n\rangle} \overset{def}{=} \prod\limits_{i=1}^{m} \{0, 1, ..., n_i - 1\} \subset \mathbb{R}^m$, $m, n_i \in \mathbb{N}$, $(G, +)$ – аддитивная абелева группа.*

Изображение $I^m_{\langle n'\rangle}$ 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^m_n, а их область определения – $\operatorname{dom} I^m_n = \mathbb{Z}^m_n$. 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 $z \in \mathbb{Z}^m_{\langle n\rangle}, v \in G$ будем называть пикселем, кортеж z – позицией, его элементы – координатами (пикселя), а v – значением пикселя. Позицию с нулевыми координатами будем называть началом координат и обозначать $\mathbf{0}$.

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

Для паттерна p на изображении $I^m_{\langle n\rangle}$ введём операцию циклического сдвига на $s \in \mathbb{Z}^m_{\langle n\rangle}$: $p \nearrow s \overset{\text{def}}{=} \{\langle z_i + s_i \pmod{n_i}\rangle^m_{i=1} | z \in p\}$, при этом кортеж s будем называть величиной сдвига или сдвигом.

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

Определение 3. *Будем говорить, что паттерн p сдвигами по осям d порождает на изображении $I^m_{\langle n\rangle}$ пучок паттернов $T^m_{\langle n\rangle}(p, d)$, если $\mathbf{0} \in p$ и множество паттернов $T^m_{\langle n\rangle}(p, d) = \{p + s_i | s_i \in S^m_{\langle n\rangle}(d)\}$ является разбиением $\operatorname{dom} I^m_{\langle n\rangle}$: $p_i \neq p_j \in T^m_{\langle n\rangle}(p, d) \implies p_i \cap p_j = \varnothing$, $\bigcup\limits_{p_i \in T^m_{\langle n\rangle}(p,d)} p_i = \mathbb{Z}^m_{\langle n\rangle}$.*

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

$$P^k_t = \begin{cases} \{\langle 0, 0\rangle\}, & k = 0 \\ P^{k-1}_{\lfloor t/2\rfloor} \cup \left(P^{k-1}_{\lfloor t/2\rfloor} \nearrow \langle 2^{k-1}, \lceil t/2\rceil\rangle\right), & k > 0 \end{cases}, \tag{1}$$

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

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

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

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

Утверждение 1 доказывается по индукции. Из утверждения 1 следует, что порождающие диадические паттерны порождают пучки вида $T^2_{2^k}(P_t^k, \langle 0, 1 \rangle)$.

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

Множество всех систем паттернов на изображении $I^m_{\langle n \rangle}$ будем обозначать $\mathbb{Y}^m_{\langle n \rangle}$.

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

Суммой по паттерну будем называть сумму значений пикселей в позициях, принадлежащих паттерну: $\sum\limits_p I^m_{\langle n \rangle} = \sum\limits_{z \in p} I^m_{\langle n \rangle}(z)$.

Определение 6. *Хаф-образом изображения $I^m_{\langle n \rangle}$ для системы паттернов $Y \in \mathbb{Y}^m_{\langle n \rangle}$ будем называть изображение $J^{m'}_{\langle n' \rangle} = \sum\limits_{Q(z)} I^m_{\langle n \rangle}$, где $Q : \mathbb{Z}^{m'}_{\langle n' \rangle} \to Y$ – сюръекция, $z \in \mathbb{Z}^{m'}_{\langle n' \rangle}$. Сюръекция Q будем называть параметризацией системы Y для Хаф-образа $J^{m'}_{\langle n' \rangle}$.*

St-параметризацией системы $Y_d^{2^k,2}$ для Хаф-образа $J^2_{2^k}$ будем называть сюръекцию $\langle s, t \rangle \longmapsto P_t^k \nearrow \langle 0, s \rangle$. Пример паттерна, задаваемого этой параметризацией, приведён на рис. 1.

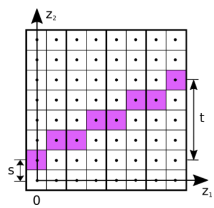

Рис. 1 — Диадический паттерн $\langle 1, 4 \rangle$ в st-параметризации системы $Y_d^{2^3,2}$.

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 $I^m_{2^k}$ – 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: $Sym(P) \stackrel{def}{=} \bigcup\limits_{\pi \in B_m, p \in P} \pi(p)$.*

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, задаваемых правилом (1). При отражении по обеим координатным осям любой порождающий паттерн из $Y_d^{2^k,2}$ переходит в отличающийся только сдвигом, а следовательно, порождённый 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 найдётся 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 теоремы выполняется в два этапа. Сначала по индукции доказывается, что $\forall z', z'' \in p \in Y_d^{2^k,2} : z_1' < z_1''$ при увеличении наклона t на единицу разница $\Delta z = z_2'' - z_2'$ не убывает. Затем, на основе этого утверждения, показывается как для любой пары позиций z' и z'' построить паттерн, через них проходящий.

Определение 9. *Порождающими диадическими плоскостями $DP_{(t_1,t_2)}^{\langle e \rangle}$ 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:*

$$z_3 = D_{t_1}^k(z_1) + D_{t_2}^k(z_2), \tag{3}$$

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

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

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

Утверждение 3. $|Sym(Y_{dp}^{\langle e\rangle,3})| \le 12|Y_{dp}^{\langle e\rangle,3}|$.

Поскольку $|B_3| = 48$, то $|Sym(Y_{dp}^{\langle e\rangle,3})| \le 48|Y_{dp}^{\langle e\rangle,3}|$. В B_3 существует две конечные подгруппы отражений порядка 2, первая – отражение сразу по трём координатным осям, отображающая каждый пучок в себя, а вторая – отражение по координатной оси, переставляющая пары пучков системы местами. Обозначим B_3^s множество нетривиальных симметрий.

Предложен новый алгоритм трёхмерного БПХ (ТБПХ) для плоскостей. Входом алгоритма является трёхмерное изображение, а выходом – трёхмерный Хаф-образ. Алгоритм состоит из двух этапов: на первом этапе формируется промежуточное трёхмерное изображение HS путём вычисления БПХ от каждого двумерного изображения значения z_3 всех пикселей которого в исходном трёхмерном изображении равны. В результате в значениях пикселей HS содержатся суммы по диадическому паттерну, из которых на втором шаге формируются суммы по диадическим плоскостям. 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 4. *Алгоритм ТБПХ для плоскостей вычисляет Хаф-образ в st-параметризации системы $Y_{dp}^{n,3}$ изображения $I_{\langle e\rangle}^3$, где $\langle e\rangle = \langle 2^k, 2^k, 2^{k+1}\rangle$.*

Утверждение 4 доказывается по индукции. Из утверждения 3 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 счёт повторного использования результатов первого этапа алгоритма до $36n^3 \log_2 n$ операций.

Определение 10. *Порождающими диадическими прямыми DL_{t_1,t_2}^k на изображениях $I_{2^k}^3$ будем называть паттерны, задаваемые системой уравнений:*

$$\begin{cases} z_1 = D_{t_1}^k(z_3), \\ z_2 = D_{t_2}^k(z_3), \end{cases} \tag{4}$$

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

Основной системой диадических прямых на трёхмерном изображении I_n^3 будем называть систему $Y_{dl}^{n,3} \stackrel{\text{def}}{=} \bigcup\limits_{t_1=0}^{n-1} \bigcup\limits_{t_2=0}^{n-1} T_n^3(DL_{(t_1,t_2)}^n, \langle 1,1,0\rangle)$.

St-параметризацией системы $Y_{dl}^{n,3}$ для Хаф-образа $J_{\langle n\rangle}^4$ будем называть сюръекцию $\langle s_1, s_2, t_1, t_2\rangle \longmapsto DL_{t_1,t_2}^n \nearrow \langle s_1, s_1, 0\rangle$. Полной системой диадических прямых будем называть $Sym(Y_{dl}^{n,3})$. Диадической прямой будем называть элемент множества $Sym(Y_{dl}^{\langle e\rangle,3})$.

Утверждение 5. $|Sym(Y_{dl}^{n,3})| \leq 12|Y_{dl}^{n,3}|$.

В работе показано, что у диадических прямых такое же множество нетривиальных симметрий B_3^s, что и у диадических плоскостей, откуда следует верность утверждения 5.

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

В диссертации предложен новый алгоритм ТБПХ для прямых. Входом является трёхмерное изображение, записанное в четырёхмерное при $z_4 = 0$, а выходом – Хаф-образ для него. Сначала изображение разбивается на пары соседних слоёв 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 трёхмерного преобразования Хафа составляет $2n^3 + 4n^3 + 16n^3 + ... + n^4 = 2n^3(n-1)$ операций. То есть асимптотическая вычислительная сложность алгоритма равна $\Theta(n^4)$ для симметрического расширения системы диадических прямых.

Утверждение 6. *Алгоритм ТБПХ для прямых вычисляет Хаф-образ изображения I_n^3 в st-параметризации системы $Y_{dl}^{n,3}$, где $n = 2^k$, $k \in \mathbb{Z}$.*

Утверждение 6 доказывается по индукции. На основе этого утверждения в диссертации показано, что асимптотическая сложность алгоритма, вычисляющего полную систему диадических прямых, составляет $\Theta(n^4)$, и эта оценка асимптотически неулучшаема.

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

Определение 11. *Ортотропный разброс (отклонение от корреспондентной прямой) порождающего диадического паттерна $p \in Y_d^{2^k,2}$ с наклоном*

t называется величина E, равная:

$$E_{t,k} \stackrel{def}{=} \max_{z \in p} \left(\left| \hat{z}_2 - a\hat{z}_1 \right| \right), \tag{5}$$

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

Брейди показал, что максимальный ортотропный разброс (11) не превышает величины $\frac{k}{2}$. Позднее С. Карпенко и автор диссертации показали, что его точная верхняя грань для чётных k составляет $\frac{k}{6}$, для нечётных k составляет $\frac{k}{6} - \frac{1}{18}$.

В диссертации предложен новый способ количественной характеристики разброса диадического паттерна.

Определение 12. *Максимальный ортогональный разброс $N_{t,k}$ 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, то есть*

$$N_{t,k} \stackrel{def}{=} \max_{\hat{z} \in p} \left(\frac{|a\hat{z}_1 - \hat{z}_2|}{\sqrt{a^2+1}} \right). \tag{6}$$

Из определений 11 и 12 следует, что максимальные ортогональный и ортотропный разбросы порождающего диадического паттерна связаны выражением

$$N_{t,k} = \frac{E_{t,k}}{\sqrt{1+a^2}}. \tag{7}$$

Согласно формуле (7) максимальный ортогональный разброс ограничен сверху ортотропным, то есть не превышает $\frac{k}{6}$. Показано, что максимальный ортогональный разброс ограничен снизу величиной $\frac{k}{2\sqrt{10}}$.

В диссертации характеристики разброса обобщаются на трёхмерный случай. 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. *Максимальным ортотропным разбросом (отклонение от корреспондентной плоскости) порождающей диадической плоскости $p \in Y_{dp}^{\langle e \rangle,3}$ с наклонами $\langle t_1, t_2 \rangle$ называется величина $E^{dp}_{t_1,t_2,k}$ равная:*

$$E^{dp}_{t_1,t_2,k} \stackrel{def}{=} \max_{\hat{z} \in p} \left(\left| \hat{z}_3 - a_1\hat{z}_1 - a_2\hat{z}_2 \right| \right), \tag{8}$$

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

Максимальный ортотропный разброс $E^{dp}_{t_1,t_2,k}$ складывается из разбросов образующих диадических паттернов – и ограничен сверху $\frac{k}{3}$.

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

Утверждение 7. *Максимальный ортогональный $N^{dp}_{t_1,t_2,k}$ и ортотропный $E^{dp}_{t_1,t_2,k}$ разбросы основных диадических плоскостей связаны следующим выражением:*

$$N^{dp}_{t_1,t_2,k} = \frac{E^{dp}_{t_1,t_2,k}}{\sqrt{1 + a_1^2 + a_2^2}}. \tag{9}$$

В диссертации показано, что $N^{dp}_{t_1,t_2,k}$ ограничен сверху величиной $\frac{k}{3}$, а снизу – $\frac{k}{\sqrt{11}}$.

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

Определение 15. *Максимальным ортотропным разбросом порождающей диадической прямой $p \in Y^{\langle e\rangle,3}_{dl}$ с наклонами $\langle t_1, t_2\rangle$ называется величина $E^{dl}_{t_1,t_2,k}$ равная:*

$$E^{dl}_{t_1,t_2,k} \stackrel{def}{=} \max_{\hat{z}\in p} \sqrt{\left(\hat{z}_1 - a_1\hat{z}_3\right)^2 + \left(\hat{z}_2 - a_2\hat{z}_3\right)^2}. \tag{10}$$

В работе показано, что $E^{dl}_{t_1,t_2,k}$ ограничен сверху величиной $\frac{k\sqrt{2}}{6}$.

Определение 16. *Максимальный ортогональный разброс (отклонение от корреспондентной прямой) $N^{dl}_{t_1,t_2,k}$ порождающей диадической прямой $p \in Y^{n,3}_{dl}$ с наклонами $\langle t_1, t_2\rangle$ – это наибольшее среди ортогональных расстояний от точек $z \in p$ до корреспондентной прямой.*

Утверждение 8. *Максимальный ортогональный $N^{dl}_{t_1,t_2,k}$ и ортотропный $E^{dl}_{t_1,t_2,k}$ разброс основных диадических прямых связаны следующим образом:*

$$N^{dl}_{t_1,t_2,k} = \frac{E^{dl}_{t_1,t_2,k}}{\sqrt{1 + a_1^2 + a_2^2}}. \tag{11}$$

В диссертации показано, что $N^{dl}_{t_1,t_2,k}$ ограничен сверху величиной $\frac{k\sqrt{2}}{6}$, а снизу – $\frac{k}{\sqrt{22}}$.

Результаты, изложенные в главе, опубликованы в работах [1, 5, 6, 8, 9, 10].

Вторая глава посвящена развитию метода поиска прямых на изображении, предложенного Н. Кирьяти и А. Брукштейном. Этот метод позволяет найти прямые путём решения задачи ортогональной линейной регрессии (ОЛР), вычисляя М-оценку с помощью стандартного преобразования Хафа. В диссертации предложены новые алгоритмы для решения задачи ОЛР на двумерных и трёхмерных изображениях с помощью алгоритмов быстрого преобразования Хафа, предложенных в предыдущей главе.

Раздел 2.1 содержит обзор существующих методов детекции прямых, рассмотрены как общеизвестные алгоритмы, так и антология алгоритмов, использующих ПХ. Подробно рассмотрена серия работ Н. Кирьяти и А. Брукштейна по развитию метода детекции прямой путём вычисления М-оценки в задаче ОЛР.

Описана М-оценка в задаче ОЛР по множеству наблюдений $X = \{(x_{i1}, x_{i2}, ..., x_{im})\}_{i=1}^{N}$:

$$\hat{\theta} = \underset{\vec{\theta}}{\operatorname{argmin}} \sum_{i=1}^{N} \rho(\varepsilon_{\vec{\theta}}(\vec{x_i})), \tag{12}$$

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 $\vec{\theta}$), ρ – неубывающая на $[0, \infty)$ дифференцируемая функция с ограниченной производной, равная нулю в нуле.

Для приближенного решения задачи (12) обычно используют методы итеративного перевзвешивания (МИП), асимптотическая вычислительная сложность которых составляет $\Theta(Nt)$, где t – число итераций метода. В своей работе Н. Кирьяти и А. Брукштейн предложили альтернативный метод для решения этой задачи в двумерном случае с помощью стандартного преобразования Хафа, показав, что его асимптотическая сложность составляет $\Theta(Nn_\theta n_\rho)$, где n_θ, n_ρ – размеры Хаф-образа. По сравнению с МИП этот метод позволяет, управляя дискретизацией пространства Хафа, выбирать подходящий для конкретной задачи баланс между временем вычисления и требуемой точностью определения $\hat{\theta}$.

В разделе 2.2 показано, что М-оценка в задаче ортогональной линейной регрессии является координатами максимума Радон-образа некоторой функции от X. ПР функции f, интеграл которой по любой гиперплоскости существует, может быть записано следующим образом:

$$Rf(\vec{\theta}) = \int\limits_{\mathbb{R}^m} \delta(\vec{x}\vec{\theta_a} - \theta_0) f(\vec{x}) d\vec{x}, \tag{13}$$

где θ_0 – расстояние от начала координат, $\vec{\theta_a} = (\theta_1, ..., \theta_m)$ – координаты единичного вектора нормали к искомой гиперплоскости, а $\delta(\vec{x})$ – дельта-функция Дирака. Определим *индикаторную функцию*

$$f^I(\vec{z}) = \sum_{\vec{x} \in X} \delta(\vec{z} - \vec{x}) \tag{14}$$

и *свёртку индикаторной функции*

$$f^I_K(\vec{x}) = K(||\vec{x}||) * f^I(\vec{x}), \tag{15}$$

где $K : \mathbb{R} \to \mathbb{R}$.

Из результатов Н. Кирьяти и А. Брукштейна непосредственно следует, что для множества точек X и функции расстояния ρ, убывающей быстрее, чем $1/r$ (или финитной), можно построить такое ядро свёртки K

$$K(r) = A\rho = \frac{1}{\pi} \int\limits_r^\infty \frac{\rho'(x)}{\sqrt{x^2 - r^2}} dx, \tag{16}$$

что М-оценка (12) равна координатам максимума Радон-образа Rf^I_K:

$$\hat{\theta} = \operatorname*{argmax}_{\vec{\theta}} Rf^I_K(\vec{\theta}), \tag{17}$$

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

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 нормальной параметризации прямой), профиль которого задаётся функцией М-оценки ρ. Образы всех точек складываются и формируют Хаф-образ изображения, являющийся дискретной аппроксимацией Радон-образа Rf^I_K. Искомой М-оценкой $\hat{\theta}$ считаются координаты элемента Хаф-образа с максимальным значением, а точность её определения зависит от дискретизации Хаф-образа.

В диссертации вводится понятие ПР с послойной свёрткой

$$R_\kappa f^I(\vec{\theta}) = \int\limits_{\mathbb{R}^n} \kappa(\vec{x}\vec{\theta_a} - \theta_0) f^I(\vec{x}) d\vec{x}, \tag{18}$$

для которого показано, что $Rf^I_K(\vec{\theta}) = R_\kappa f^I(\vec{\theta})$ и, таким образом,

$$\hat{\theta} = \operatorname*{argmax}_{\vec{\theta}} R_\kappa f^I(\vec{\theta}), \tag{19}$$

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

В разделе 2.3 описан новый метод поиска прямых и плоскостей на двумерных и в трёхмерных изображениях путём вычисления М-оценок в задаче ОЛР с помощью алгоритмов, предложенных в главе 1. Предложенный 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.

Первым этапом метода является формирование гистограммы по множеству входных наблюдений. Пусть $a_j = \min\limits_{i\in[1;N]}(x_{ij})$, $b_j = \max\limits_{i\in[1;N]}(x_{ij}) - \min\limits_{i\in[1;N]}(x_{ij})$. Пусть $\hat{n} = \max\limits_{j\in[1,m]}(n_j)$, где $n_j = b_j/h$, притом шаг дискретизации h одинаковый по всем измерениям и такой, что $\forall j \in [1, m] : n_j > 2$.

Определение 17. *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 $\big(\mathbb{R}, +\big)$, 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 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 18. *Расширенным изображением $I^m_{\langle n \rangle}$ по компоненте j на величину l назовём отображение $Aug_j^l(I^m_{\langle n \rangle}) : \mathbb{Z}^m_{\langle n^a \rangle} \to R$, такое что $Aug_j^l(I^m_{\langle n \rangle})(z) = I^m_{\langle n \rangle}(z)$ для $z \in \mathbb{Z}^m_{\langle n \rangle}$ и $Aug_j^l(I^m_{\langle n \rangle})(z) = 0$ для $z \in \mathbb{Z}^m_{\langle n^a \rangle} \setminus \mathbb{Z}^m_{\langle n \rangle}$, где $n_j^a = n_j + l$, а также $\forall i \in \overline{1, m}, i \neq j : n_i^a = n_i$.*

В работе показано, что при переходе к расширенному изображению асимптотическая вычислительная сложность алгоритмов БПХ и ТБПХ не меняется, если $l = cn$, где c – неотрицательная константа. Паттерну (или плоскости, или прямой) из основной системы диадических паттернов (или плоскостей, или прямых), содержащих пиксели, расположенные в области ненулевых значений, ставятся в соответствие корреспондентная прямая или плоскость с теми же свойствами разброса, что и для порождающих диадических паттернов (или плоскостей, или прямых).

В диссертации предложено два метода приближенного вычисления М-оценок, в задаче ОЛР делятся в зависимости от порядка вычисления свёртки и Хаф-образа. На первом шаге выполняется формирование гистограммы $H = I^m_{\hat{n}}$ множества наблюдений X. Далее описан метод со свёрткой Хаф-образа. Последующие шаги для группы одинаковы и выполняются для всех элементов множества симметрий (B_2^s, B_3^s):

1. Вычисление Хаф-образа J от расширенного изображения гистограммы[1].
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 решётки 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 $\alpha_t = \operatorname{arctg}(\frac{t}{n-1})$, 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 трёхмерном случае $c(t) = c_3^p(t) = \cos(\alpha_{t_1}) \cdot \cos(\alpha_{t_1})$, а для прямых в трёхмерном случае $c(t) = c_3^l(t) = \sqrt{\frac{(n-1)^2}{(n-1)^2+t_1^2+t_2^2}}$.

Утверждение 9. *Вычислительная сложность метода поиска прямой со свёрткой Хаф-образа составляет $\Theta(N + n^2 \log n)$ для двумерного случая; $\Theta(N + n^3 \log n)$ для плоскостей в трёхмерном случае; $\Theta(N + n^4 \log n)$ или $\Theta(N + w^2 n^4)$ (без использования БПФ), где w – линейный размер носителя ядра свёртки для прямых в трёхмерном случае.*

Доказательство строится на следующих соображениях. Формирование гистограммы требует N сложений, БПХ вычисляется за $\Theta(n^2 \log_2 n)$, 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 доказывается для трёхмерных случаев. Метод позволяет адаптивно подбирать параметры аддитивного шума без пересчёта Хаф-образа.

В разделе 2.4 описан второй метод, согласно которому процедура свёртки выполняется в пространстве изображения перед вычислением БПХ. В этом случае не требуется выполнять адаптацию размера ядра свёртки.

Утверждение 10. *Вычислительная сложность метода поиска прямой со свёрткой изображения составляет $\Theta(N + n^2 \log n)$ для двумерного случая; $\Theta(N + n^3 \log n)$ для плоскостей в трёхмерном случае; $\Theta(N + n^4)$ для прямых в трёхмерном случае.*

На основе (16) получены формы ядер свёртки с прямым пространством для М-оценок Тьюки, Коши, Джимана-МакКлюра и Уэлша.

[1] Для двумерного случая используется расширенное изображение $\mathrm{Aug}_2^n(H)$, для плоскостей в трёхмерном случае $\mathrm{Aug}_3^n(H)$, а для прямых – $\mathrm{Aug}_1^n(\mathrm{Aug}_2^n(H))$.

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

В разделе 2.5 получены оценки точности аппроксимации корреспондентных прямых и плоскостей предложенным в 2.5 методом на основе результатов первой главы (о величинах максимальных ортогональных разбросов диадических паттернов).

Оптимизируемая в задаче (12) функция может быть записана в следующем виде:

$$f_\rho^X(\vec{\theta}) = \sum_{i=1}^{N} \rho(\varepsilon_{\vec{\theta}}(\vec{x}_i))/N. \tag{20}$$

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

$$\Delta\rho(\varepsilon_{\vec{\theta}}(x_i)) \leq \Big|\rho(\varepsilon_{\vec{\theta}}(x_i)) - \rho(\varepsilon_{\hat{\theta}}(\hat{x}_i) + h\sqrt{2})\Big| \leq$$

$$\rho'_{max} \cdot \Big|\varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i) - h\sqrt{2}\Big| \leq \rho'_{max} h\sqrt{2} + \rho'_{max} \cdot \Big|\varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i)\Big|,$$

где ρ'_{max} супремум производной функции ρ. Ошибка вычисления ортогонального расстояния от прямой до точки $\Big|\varepsilon_{\vec{\theta}}(x_i) - \varepsilon_{\hat{\theta}}(\hat{x}_i)\Big|$, в свою очередь, складывается из ортогонального разброса диадического паттерна, ограниченного величиной $\frac{\log_2 n}{6}$, из 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 11. *Ошибка вычисления* (20) *для прямой с параметрами* $\vec{\theta}$ *в двумерном случае ограничена следующим выражением:*

$$\Delta f_\rho^X(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left(\frac{\log_2 n}{6} + 2\sqrt{2} + 2\right). \tag{21}$$

Утверждение 12. *Ошибка вычисления функции* (20) *для плоскости с параметрами* $\vec{\theta}$ *в трёхмерном случае с помощью ТБПХ для плоскостей ограничена следующим выражением:*

$$\Delta f_{\rho}^{X}(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left(\frac{\log_2 n}{3} + 2\sqrt{3} + 2 \right). \tag{22}$$

Утверждение 13. *Ошибка вычисления функции* (20) *для прямой с параметрами* $\vec{\theta}$ *в трёхмерном случае с помощью ТБПХ для прямых ограничена следующим выражением:*

$$\Delta f_{\rho}^{X}(\vec{\theta}) \leq h \cdot \rho'_{max} \cdot \left(\frac{\sqrt{2}\log_2 n}{6} + 2\sqrt{3} + 2\sqrt{2} \right). \tag{23}$$

В разделе 2.6 приведены результаты сравнения на модельных данных (с выбросовым и аддитивным шумом) предложенных методов вычисления М-оценок и других методов решения задачи линейной регрессии, таких как МНК, медианные алгоритмы Зигеля и Тейла-Сены, а также алгоритм итеративного перевзвешивания весов с функциями Тьюки, Уэлша и Хьюбера.

Расстояние между искомой прямой $\vec{\theta}$ и найденной алгоритмом $\hat{\theta}$ в двумерном случае измеряется как нормированный на n максимум между длинами дуг l_1, l_2 образованными в результате пересечения описанной вокруг гистограммы H окружности этих прямых. В трёхмерном случае расстояние мерилось как евклидово расстояние между точками $\vec{\theta}$ и $\hat{\theta}$.

Сравнение методов ортогональной линейной регрессии показало, что метод вычисления М-оценки Уэлша с помощью БПХ превосходит все рассматриваемые методы по предложенным метрикам в двумерном случае с долей выбросового шума до 0,7, в трёхмерном – до 0,5.

Результаты, изложенные в главе, опубликованы в работах [2, 7].

Третья глава диссертации посвящена созданию метода решения задачи бинарной кластеризации на двумерных и трёхмерных гистограммах в контексте развития обобщённых методов глобальной бинаризации Оцу.

Существует множество обобщений метода Оцу, где, в отличие от классического, анализируется не одномерная гистограмма, а двумерная H, что позволяет повысить качество бинаризации, но существенно замедляет алгоритм. В своей статье Цзянь Гон с соавторами предложили быстрый метод разделения области определения гистограммы с трудоёмкостью $\Theta(n^2)$, где n – сторона гистограммы. Как и в оригинальном методе, авторы выполняют поиск параметров разделяющей поверхности, оптимизирующих обобщенный критерий Оцу. Известно, что в этом случае оптимальной разделяющей поверхностью является прямая, поскольку критерий Оцу выводится из предположения, что сигнал порождён смесью двух нормальных распределений с равными матрицами ковариации. Однако авторы в качестве разделяющей поверхности используют прямоугольник одна вершина которого 0, а вторая – искомый параметр (z_1^t, z_2^t): пиксели, ограниченные этим прямоугольником, – это пиксели первого класса, остальные – второго. Причина выбора заключается в том, что только для таких разделяющих поверхностей существовал достаточно быстрый алгоритм подсчёта критерия Оцу, а степень искажения критерия при такой подмене выносилась за скобки в угоду уменьшению асимптотической вычислительной сложности.

В диссертации предложен метод линейного разделения гистограмм согласно критерию Оцу с помощью алгоритмов БПХ. Здесь разделяющей поверхностью является диадический паттерн или диадическая плоскость, максимальный ортотропный и ортогональный разброс которых относительно размера стороны стремится к нулю при $n \to \infty$.

В разделах 3.1 и 3.2 рассматривается классический критерий Оцу и приведены его обобщения на случай многомерных гистограмм, а также их вывод из принципа максимального правдоподобия.

Пусть множество точек X – наблюдения из 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 и матрицами ковариации такими, что $\Sigma = \lambda I$, где I – единичная матрица, а $\lambda \in \mathbb{R}^+$; тогда критерий разделения выглядит следующим образом:

$$F_1(\vec{\theta}) = \frac{m}{2} \ln \Big(\sum_{j=1}^{2} \omega_j(\vec{\theta}) D_j(\vec{\theta}) \Big). \tag{24}$$

Здесь $\omega_j(\vec{\theta})$ – вес j-го кластера, $D_j(\vec{\theta})$ – выборочная дисперсия j-го кластера.

В разделе 3.3 описан метод линейной бинарной кластеризации, согласно критерию (24) и любым другим критериям, выразимым через меры, задаваемые на области определения гистограммы.

Опишем метод на примере двумерного случая $(H^2_{2^k})$ 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 $\mathrm{Aug}_2^{2^k}(H^2_{2^k})$. Обозначим $C_1(\hat{\theta}) = \{z^c : z_1^c = z_1, z_2^c \leq z_2, \forall \langle z_1, z_2 \rangle \in P\}$, а множество всех остальных позиций – $C_2(\hat{\theta})$.

Утверждение 14. *Вычислительная сложность подсчёта сумм по множествам $C_1(\hat{\theta})$ и $C_2(\hat{\theta})$ изображения $H^2_{2^k}$ составляет $\Theta(1)$ при асимптотической вычислительной сложности предподсчёта $\Theta(n^2 \log n)$.*

Опишем кратко доказательство этого утверждения. Для изображения $U = H^2_{2^k}$ формируется кумулятивное изображение U^c, такое, что $\forall \langle z_1, z_2 \rangle \in \mathbb{Z}_n^2 : U^c(z_1, z_2) = \sum\limits_{z_2^* \in \overline{1, z_2}} U(z_1, z_2^*)$, для чего требуется n^2 операций. Таким образом, сумма по любому $p \in Y_d^{2^k,2}$ на расширенном изображении $\mathrm{Aug}_2^{2^k}(U^c)$ является суммой по $C_1(\hat{\theta})$, а сумма по $C_2(\hat{\theta})$ равна разнице $\Omega(U) = \sum\limits_{\langle z_1, z_2\rangle \in \mathbb{Z}_n^2} U(z_1, z_2)$ и суммы по $C_1(\hat{\theta})$. Вычислив БПХ от $\mathrm{Aug}_2^{2^k}(U^c)$: получим 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, как и к каждому значению $C_2(\hat{\theta})$. Итоговая асимптотическая вычислительная сложность предподсчёта определяется самым трудным этапом – вычислением БПХ – и составляет $\Theta(n^2 \log n)$.

Утверждение 15. *Вычислительная сложность подсчёта критерия (24) для гистограммы $H_{2^k}^2$ и паттернов $p \in Y_d^{2^k,2}$ равна $\Theta(1)$ при асимптотической вычислительной сложности предподсчёта $\Theta(n^2 \log n)$.*

Доказательство основано на утверждении 14. Для вычисления критерия (24) 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 $\omega_j(\hat{\theta}) = \sum\limits_{\langle z_1, z_2\rangle \in C_j(\hat{\theta})} \mathrm{Aug}_2^{2^k}(H_{2^k}^2)(z_1, z_2)$, а эти суммы, согласно утверждению 14, 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 $M_i : dom(M_i) = dom(H_{2^k}^2), M_i(z_1, z_2) = z_i H_{2^k}^2(z_1, z_2)$, тогда

$$\mu_{ji}(\hat{\theta}) = \frac{1}{\omega_j(\hat{\theta})} \sum_{\langle z_1, z_2\rangle \in C_j(\hat{\theta})} \mathrm{Aug}_2^{2^k}(M_i)(z_1, z_2), \tag{25}$$

$\forall i,j = \overline{1,2}$. Согласно утверждению 14, сумма в формуле (25) вычисляется с необходимой асимптотической вычислительной сложностью. Аналогично (25) вычисляется вектор $\nu_j(\hat{\theta}) = (\nu_{j1}(\hat{\theta}), \nu_{j2}(\hat{\theta}))^T$, $\forall i,j = \overline{1,2}$ для изображений $L_i : dom(L_i) = dom(H_{2^k}^2), L_i(z_1, z_2) = z_i^2 H_{2^k}^2(z_1, z_2)$. 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 подсчёта составляющих равна $\Theta(1)$ при сложности предподсчёта $\Theta(n^2 \log n)$, то и для подсчёта дисперсии это верно, а значит и для подсчёта критерия (24).

Аналогично вычисляется любой критерий разделения, выразимый через меры, заданные на области определения изображения, в том числе в трёхмерном случае (тогда асимптотическая вычислительная сложность предподсчёта составляет $\Theta(n^3 \log n)$). В работе приведено описание метода линейной бинарной кластеризации с использованием БПХ (или ТБПХ).

В разделе 3.4 рассмотрен вопрос влияния структуры диадического паттерна (и диадической плоскости) на точность вычисления критерия и предложен метод оценки точности критерия для заданной гистограммы и точки $\hat{\theta}$. Погрешность вычисления критерия определяется через погрешности вычисления составляющих его мер, заданных на области определения гистограммы. Оценка погрешности вычисления меры в точке $\hat{\theta}$ мажорируется модулем разности сумм по соседним параллельным паттернам (эти суммы уже подсчитаны в процессе работы алгоритма разделения и содержатся в соответствующем Хаф-образе).

Результаты, изложенные в главе, опубликованы в работах [3, 4].

В **заключении** приведены основные результаты работы:

1. Предложены два обобщения диадического паттерна на трёхмерный 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 величиной $\frac{\log_2 n}{\sqrt{11}}$, 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 найдётся проходящий через них ДП.
4. Предложены два алгоритма вычисления трёхмерного 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 трёхмерных 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 трёхмерных изображений для 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 трёхмерных 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 трёхмерных. Предложен метод оценки сверху ошибки вычисления величины критерия для заданной входной гистограммы и параметров разделения.

Публикации автора по теме диссертации

В изданиях из списка ВАК РФ

1. *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.
2. *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.
3. *Ершов, Е. И.* Алгоритм бинарной линейной кластеризации маломерных гистограмм / Е. И. Ершов // Сенсорные системы. — 2017. — Т. 31, № 3. — С. 261—269.

В сборниках трудов конференций

4. 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.
5. 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.
6. Generation algorithms of fast generalized Hough transform / E. Ershov [и др.] // 31st European Conference on Modelling and Simulation. — 2017. — С. 534—538.
7. *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.
8. О точной оценке неточностей аппроксимации прямых в алгоритме быстрого преобразования Хафа / Е. И. Ершов [и др.] //. — ИТиС 2015. 2015. — С. 858—868.
9. Fast 3D Hough Transform Computation. / E. Ershov [и др.] // 30 European Conference on Modelling and Simulation. — 2016. — С. 227—230.

В прочих изданиях

10. *Ershov, E.* Fast Hough Transform and approximation properties of dyadic patterns / E. Ershov, S. Karpenko // arXiv preprint arXiv:1712.05615. — 2017.

Ершов Егор Иванович

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

Автореф. дис. на соискание учёной степени канд. физ.-мат. наук

Подписано в печать ___.___._____. Заказ № ____________
Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.
Типография ______________________________________