Гаганов Виктор Александрович

Исследование и разработка программных средств распознавания образов для решения задачи трехмерного моделирования в микроскопии

Специальность 05.13.11 – математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата физико-математических наук

> Научный руководитель к. ф.-м. н. Баяковский Ю.М.

Содержание

Введе	ние		6
Глава	1. 3a ₂	цача построения трехмерной модели по изображе-	
ния	им с ми	кроскопа	13
1.1.	Модел	ирование микрообъектов	13
	1.1.1.	Особенности изображений с микроскопа	13
	1.1.2.	Методы на основе изображений	15
1.2.	Обзор	существующих методов	17
	1.2.1.	Постановка задачи	17
	1.2.2.	Методы семейства SFD	19
	1.2.3.	Методы семейства SFF	21
	1.2.4.	Мера резкости	22
	1.2.5.	Базовая схема методов SFF	24
	1.2.6.	Уточнение модели	27
	1.2.7.	Сравнение подходов SFF и SFD	30
1.3.	Недост	гатки существующих методов семейства SFF	31
1.4.	Мультифокус изображения		
1.5.	Заклю	учение	36
Глава	2. Vc'	тойчивый метод построения поверхности произ-	
вол	ьной ф	ормы	38
2.1.	Маркс	эвские случайные поля	38
	2.1.1.	Оценка параметров поля	40
	2.1.2.	Теорема Хаммерсли-Клифорда	41
2.2.	Предл	агаемый метод	42
	2.2.1.	Статистическая формулировка задачи	42

	2.2.2.	Функция правдоподобия	43
	2.2.3.	Структура поля	44
	2.2.4.	Итоговая функция энергии	47
	2.2.5.	Унарный потенциал	48
	2.2.6.	Парный потенциал	54
	2.2.7.	Оптимизация функционала	60
	2.2.8.	Усреднение меры резкости	62
	2.2.9.	Итоговый алгоритм	63
2.3	3. Экспеј	рименты	64
	2.3.1.	Тестовая базы для сравнения	64
	2.3.2.	Метрика для сравнения методов	66
	2.3.3.	Результаты сравнения	67
	2.3.4.	Визуальное сравнение результатов	70
2.4	4. Заклю	чение	72
Dese	- 2 0		
Глав	а 3. Оп	енка положения и ориентации плоского участка	74
Глава по	а 3. Оц оверхнос ^и 2.0.1	енка положения и ориентации плоского участка ги	74
Глава	а 3. Оп оверхнос ^и 3.0.1.	енка положения и ориентации плоского участка ги	74 74 75
Глава по 3.1	а 3. Оц эверхнос ^и 3.0.1. 1. Сущес	енка положения и ориентации плоского участка ги	74 74 75
Глава по 3.1	а 3. Он эверхнос 3.0.1. I. Сущес 3.1.1.	енка положения и ориентации плоского участка гипостановка задачипостановка задачи твующие методы оценки плоскости Метод наименьших квадратов	74 74 75 75
Глава по 3.1	а 3. Он эверхнос 3.0.1. 1. Сущес 3.1.1. 3.1.2.	енка положения и ориентации плоского участка гиПостановка задачи	74 74 75 75 75
Глава по 3.1	а 3. Он оверхнос 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77
Глава по 3.1	а 3. Он оверхнос 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78
Глава по 3.1	а 3. Он Эверхнос 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4. 3.1.5.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78 80
Глава по 3.1	а 3. Он DBEPXHOC 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4. 3.1.5. 3.1.6.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78 80 81
Глава по 3.1	а 3. Он Эверхнос 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4. 3.1.5. 3.1.6. 2. Метод	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78 80 81 82
Глава по 3.1 3.2	а 3. Он DBEPXHOC 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4. 3.1.5. 3.1.6. 2. Метод 3.2.1.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78 80 81 82 82
Глава по 3.1	а 3. Он DBEPXHOC 3.0.1. 1. Сущес 3.1.1. 3.1.2. 3.1.3. 3.1.4. 3.1.5. 3.1.6. 2. Метод 3.2.1. 3.2.2.	енка положения и ориентации плоского участка ги	74 74 75 75 76 77 78 80 81 82 82 82 83

	3.2.3.	Фильтрация ложных замеров	85
	3.2.4.	Идея предлагаемого метода	86
	3.2.5.	Машинное обучение	88
	3.2.6.	Признаки для классификации	89
	3.2.7.	Формирование обучающей выборки	91
	3.2.8.	Обучение классификатора	92
	3.2.9.	Оценка параметров плоскости	95
3.3.	Провер	ка надежности оценки плоскости	97
3.4.	Экспер	именты	100
	3.4.1.	Оценка качества алгоритма фильтрации ложных за-	
		меров	100
	3.4.2.	Точность оценки плоскости	102
	3.4.3.	Качество процедуры проверки надежности плоскости	103
3.5.	Заключ	ение	104
0.0.			101
Глава 4.	Про	граммная реализация предложенных методов	105
Глава 4. 4.1.	Про Алмазо	граммная реализация предложенных методов обрабатывающая промышленность	105 105
Глава 4 . 4.1.	Про Алмазо 4.1.1.	граммная реализация предложенных методов обрабатывающая промышленность Открытие окон	105 105 106
Глава 4. 4.1.	Про Алмазо 4.1.1. 4.1.2.	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107
Глава 4 . 4.1.	Про Алмазо 4.1.1. 4.1.2. 4.1.3.	граммная реализация предложенных методов обрабатывающая промышленность Открытие окон Построение трехмерной модели камня Моделирование дефектов внутри камня и уточнение	105 105 106 107
Глава 4. 4.1.	Про Алмазо 4.1.1. 4.1.2. 4.1.3.	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107
Глава 4 . 4.1.	Про Алмазо 4.1.1. 4.1.2. 4.1.3. 4.1.4.	граммная реализация предложенных методов обрабатывающая промышленность Открытие окон Построение трехмерной модели камня Моделирование дефектов внутри камня и уточнение модели Планирование огранки камня	105 105 106 107 108 111
Глава 4. 4.1. 4.2.	Про Алмазо 4.1.1. 4.1.2. 4.1.3. 4.1.4. Интерф	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107 108 111
Глава 4. 4.1. 4.2.	Про Алмазо 4.1.1. 4.1.2. 4.1.3. 4.1.4. Интерф Server	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107 108 111
Глава 4. 4.1. 4.2. 4.3.	Про Алмазо 4.1.1. 4.1.2. 4.1.3. 4.1.4. Интерф Server Детали	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107 108 111 111 113
Глава 4 . 4.1. 4.2. 4.3.	Про Алмазо 4.1.1. 4.1.2. 4.1.3. 4.1.4. Интерф Server Детали 4.3.1.	граммная реализация предложенных методов обрабатывающая промышленность	105 105 106 107 108 111 111 113

	4.3.2.	Архитектура подсистемы: уточнение модели поверх-	
		ности камня	116
	4.3.3.	Архитектура подсистемы: моделирование включений	119
	4.3.4.	Замеры производительности разработанных алгорит-	
		МОВ	120
4.4.	Заклю	чение	122
Заключ	чение .		124
Благод	арност	'N	125
Литера	атура .		126

Введение

Объект исследования и актуальность работы

С момента появления первого оптического микроскопа в конце XVI – начале XVII века это устройство плотно вошло в жизнь человека. Сейчас практически невозможно представить себе современное общество без этого изобретения. В данный момент оптические микроскопы находят применение в широком спектре приложений, начиная от рутинных медицинских и биологических исследований, заканчивая контролем качества на высокотехнологичных производствах.

Распространение цифровых видеокамер во второй половине XX века привело к появлению нового класса задач, связанных с оптическими микроскопами. Появление возможности захвата изображений в цифровом формате дало толчок развитию технологий автоматического улучшения и анализа изображений, полученных с помощью микроскопа [1]. Благодаря стремительному росту вычислительных мощностей процессоров круг задач по обработке и анализу изображений с микроскопа расширяется с каждым годом, ставя перед исследователями все новые и новые задачи.

В данной работе рассматривается одна из важнейших проблем в задачах автоматического анализа изображений, полученных с помощью оптического микроскопа, — задача построения трехмерной модели сцены, наблюдаемой с помощью микроскопа. Трехмерная модель сцены незаменима в таких приложениях, как, например, анализ качества на производстве печатных плат, где требуются измерения различных показателей формы поверхности наблюдаемого объекта. Стоит заметить, что реконструкция трехмерных моделей микрообъектов традиционными методами, такими как лазерное сканирование, является затруднительной из-за физических ограничений, налагаемых размерами объектов. Этот факт придает дополнительную актуальность методам построения трехмерных моделей, которые оперируют непосредственно изображениями, полученными с помощью микроскопа.

Задача построения трехмерных моделей объектов по изображениям с микроскопа рассматривается в работе в двух формулировках.

- Построение модели поверхности общего вида в данной формулировке предполагается, что форма участка поверхности объекта, наблюдаемого с помощью микроскопа, представима в виде карты глубины¹. Тогда по набору изображений с микроскопа требуется построить трехмерную модель участка поверхности объекта.
- Построение модели плоского участка поверхности объекта предполагается, что наблюдаемый под микроскопом участок объекта плоский, и требуется оценить положение и ориентацию данного участка поверхности объекта.

Цель диссертационной работы

Целью работы является исследование и разработка методов и алгоритмов построения трехмерных моделей объектов по изображениям с микроскопа, а также создание на основе разработанных методов программной системы для построения трехмерных моделей микрообъектов.

Основные задачи работы:

• Исследование существующих алгоритмов построения трехмерных моделей объектов по изображениям с микроскопа. Разработка метода

¹ Картой глубины называется представление трехмерной модели в виде матрицы, в каждой ячейке которой записано расстояние от некоторой фиксированной плоскости в трехмерном пространстве до поверхности объекта.

построения трехмерной модели участка поверхности объекта, который представим в виде карты глубины.

- Разработка специализированного метода для оценки положения и ориентации плоского участка поверхности объекта по набору изображений с микроскопа.
- Верификация разработанных алгоритмов путем оценки качества их работы на реальных данных и сравнение предложенных методов с существующими аналогами.
- Разработка программной системы построения трехмерных моделей микрообъектов с применением предложенных методов.

Научная новизна

В рамках диссертации разработан новый метод построения трехмерных моделей участка поверхности объекта, который представим в виде карты глубины. Отличительной чертой предложенного метода является его повышенная устойчивость к ложным данным, часто возникающим в областях поверхности объектов, которые не содержат текстуры, слабо освещенных областях поверхности и в местах присутствия бликов. В работе представлены результаты сравнительного анализа, подтверждающие, что предложенный метод позволяет добиться более устойчивых и точных результатов, чем существующие аналоги.

В рамках решения задачи оценки положения и ориентации плоского участка поверхности объекта предложен новый метод для определения того, можно ли считать замер положения участка сцены надежным. Также предложен метод, позволяющий определить, можно ли по набору трехмерных замеров, содержащему ошибки измерения, осуществить надежную оценку положения и ориентации плоскости. Кроме того, предложенные методы позволяют сделать оценку положения и ориентации плоского участка поверхности объекта более устойчивой к ошибкам в измерениях.

Практическая значимость

Предложенный метод построения трехмерной модели участка поверхности объекта по изображениям с микроскопа существенно расширяет класс объектов, для которых применим данный подход. Существующие методы в основном ориентированы на работу с объектами, отражающие свойства поверхности которых близки к Ламбертовской модели. Предложенный метод позволяет строить корректные модели объектов, состоящих из таких материалов, как металл и стекло, на поверхности которых возможно наличие бликов. Также предложенный метод способен корректно моделировать объекты, на поверхности которых присутствуют области без текстуры, в то время как существующие методы ориентированы на работу с объектами, на поверхности которых видимая текстура присутствует повсеместно.

Разработанный метод определения положения и ориентации плоского участка поверхности объекта позволяет производить оценку плоскости даже при наличии во входных данных шумов и большого количества ошибочных измерений. Разработанный метод определения надежности оценки плоскости позволяет для любых входных данных определить, удалось ли по ним корректно оценить положение и ориентацию плоского участка поверхности объекта.

На основе предложенных методов была создана программная система для моделирования поверхности алмазов и дефектов внутри алмазного сырья. Данная система была интегрирована в программный продукт Oxygen Microscope Server компании-заказчика Octonus Software Ltd. Данный программный продукт активно используется:

- в реальной работе более чем на 20 ограночных заводах таких центров алмазообрабатывающей промышленности, как г. Сурат (Индия) и г. Антверпен (Бельгия);
- в процессе подготовки и обучения специалистов по огранке алмазов в Геммологическом центре Московского государственного университета им. М.В. Ломоносова.

Апробация работы

Основные результаты диссертации докладывались и обсуждались на:

- 18-ой международной конференции по компьютерной графике и машинному зрению «Graphicon'2008», Россия, Москва, 2008;
- 15-ой международной конференции студентов аспирантов и молодых ученых «Ломоносов-2008», Россия, Москва, 2008;
- семинаре по компьютерной графике и мультимедиа под руководством Ю.М. Баяковского (ф-т ВМиК МГУ), Россия, Москва, 2009;
- секции вычислительной математики и кибернетики научной конференции «Ломоносовские чтения-2009», Россия, Москва, 2009;
- 19-ой международной конференции по компьютерной графике и машинному зрению «Graphicon'2009», Россия, Москва, 2009;
- семинаре по компьютерной графике под руководством А.В.Игнатенко (ф-т ВМиК МГУ), Россия, Москва, 2010;
- объединенном семинаре по робототехническим системам ИПМ им.
 М.В. Келдыша РАН, МГУ им. М.В. Ломоносова, МГТУ им. Н.Э.

Баумана, ИНОТиИ РГГУ и отделения "Программирование"ИПМ им. М.В.Келдыша РАН, Россия, Москва 2010.

Публикации

Материалы диссертации опубликованы в 5 печатных работах: 2 статьи в рецензируемых журналах, рекомендованных ВАК [2, 3], две статьи в сборниках трудов конференций [4, 5], а также тезисы доклада [6].

Структура и объем диссертации

Диссертация состоит из введения, 4 глав, заключения и библиографии. Общий объем диссертации составляет 133 страницы, включая 33 рисунка. Библиография включает 73 наименования.

В первой главе описывается математическая модель формирования изображения и дается понятие области резкости, описываются базовые принципы работы методов построения трехмерных моделей по изображениям с микроскопа и дается обзор существующих методов. Также в главе описываются основные недостатки существующих методов - неустойчивое поведение методов в областях объекта, где на поверхности отсутствует текстура, либо присутствуют крупные блики.

Во второй главе предлагается новый метод построения трехмерных моделей по изображениям с микроскопа семейства shape from focus, отличающийся от существующих аналогов повышенной устойчивостью к ложным данным, которые возникают в областях поверхности объекта без текстуры, слабо освещенных областях поверхности и в местах присутствия бликов. Также в главе приводятся результаты экспериментов на реальных данных, которые подтверждают, что предлагаемый метод позволяет получить более устойчивый результат, чем существующие аналоги. В третьей главе описывается предлагаемый метод определения положения и ориентации плоского участка поверхности объекта по набору изображений с микроскопа. Предлагается новая процедура фильтрации ложных замеров, основанная на методах машинного обучения. Также предлагается процедура проверки надежности оценки плоскости, которая позволяет обнаружить ситуации, в которых по входным данным невозможно осуществить корректную оценку положения и ориентации плоского участка поверхности объекта.

В четвертой главе описывается интерфейс и функциональность программной системы Oxygen Microscope Server, в которую были встроены предлагаемые в работе методы. Также в главе приводятся детали программной реализации предлагаемых методов и результаты замеров их производительности.

В заключении формулируются основные результаты работы.

Глава 1

Задача построения трехмерной модели по изображениям с микроскопа

1.1. Моделирование микрообъектов

Задача построения трехмерных моделей микрообъектов является актуальной для многих приложений. При работе и изучении таких объектов, как биологические образцы, природные минералы, печатные платы и другие мелкие объекты, зачастую используются оптический микроскоп. Поэтому крайне желательно иметь методы, который могут строить трехмерные модели таких объектов непосредственно по изображениям с микроскопа.

Использование для построения трехмерных моделей методов, требующих дополнительного оборудования, таких как например лазерное сканирование или томография, является нежелательным, а иногда и невозможным в связи с физическими ограничениями, связанными с малыми размерами объектов. По тем же причинам для моделирования макрообъектов получили широкое распространение методы, использующие изображения, такие как, например, стереореконструкция [7, 8].

1.1.1. Особенности изображений с микроскопа

Для понимания специфики построения трехмерных моделей объектов по изображениям с микроскопа необходимо понимание принципов формирования изображений в реальных оптических системах. На рис. 1.1(а) изображена простейшая модель оптической системы – оптическая система с тонкой линзой. За линзой находится поверхность сенсора, с помощью которого осуществляется считывание изображения. Все световые лучи, исходящие из точки сцены P, и попадающие на линзу, сходятся в точке Q. При этом расстояния от точек P и Q до линзы подчиняются т.н. уравнению тонкой линзы:

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \tag{1.1}$$

где f - фокусное расстояние линзы. В случае если для какой-то точки P', точка схода преломленных линзой лучей Q' не попадает на сенсор, энергия лучей, исходящих из точки P', рассеивается по небольшому участку поверхности сенсора (рис. 1.1,6).



Рис. 1.1. На рисунке (a) изображена оптическая система, и точка Р изображение которой попадает на сенсор с идеальным фокусом. На рисунке (б) изображение точки Р' на сенсоре будет расфокусированным

Размер участка поверхности сенсора, по которому рассеиваются лучи, исходящие из точки P', пропорционален расстоянию от точки Q' до поверхности сенсора. Поэтому при получении фотографии сцены, только некоторая часть сцены, оказывается на фотографии в резкости. Эта часть сцены называется областью резкости, и чем дальше точка сцены находится от области резкости, тем более размытой будет выглядеть изображение этой точки на фотографии [9]. Расстояние между самой близкой и самой дальней границами пространства, которые видны на изображении в резкости, называется глубиной резкости.

В то время как получение изображения макрообъекта, на котором он целиком виден в резкости, в большинстве случаев является возмож-

ным, для микрообъектов это является проблемой. Из-за особенности конструкции оптических микроскопов по мере усиления увеличения происходит уменьшение глубины резкости оптической системы. Глубину резкости оптической системы можно увеличить за счет закрытия диафрагмы, но добиться серьезного увеличения глубины резкости без значительной потери качества изображения не представляется возможным.

На рисунке 1.2 изображена пара типичных изображений трехмерного объекта, полученных с помощью оптического микроскопа. Как можно заметить на каждом изображении в резкости виден лишь незначительный срез наблюдаемой сцены, в то время как большая часть сцены является размытой.



a)

б)

Рис. 1.2. Пара изображений природного минерала, полученные с помощью черно-белой камеры, с разными положениями области резкости. Видно что на каждом из изображений в резкости видна лишь незначительная часть сцены.

1.1.2. Методы на основе изображений

Описанный эффект глубины резкости является серьезной преградой для использования традиционных методов построения трехмерных моде-

лей по изображениям при моделировании микрообъектов. Методы стереореконструкции [7, 8] требует два и более изображения сцены, полученных с разных точек зрения. Эти методы находят соответствия между точками на изображениях и с помощью процедуры *триангуляции* получают трехмерную модель сцены. Методы стереореконструкции способны получить точную модель сцены при условии, что на поверхности объектов сцены присутствует большое количество видимой текстуры. Именно поэтому применение стереореконструкции при моделировании микрообъектов затруднительно, из-за того, что большая часть сцены трехмерного объекта на изображениях видна не в резкости, и видимая текстура на изображениях размыта.

Методы на основе согласования цветов [10–12], которые строят модель исходя из предположения, что цвета проекций каждой точки поверхности объекта должны отличаться для разных изображений незначительным образом, тоже не применимы из-за эффекта глубины резкости. Методы фотометрического стерео [13, 14] также не позволяют обеспечить достаточную точность реконструкции нормалей поверхности объекта для размытых участков входного изображения.

Наиболее подходящими методами для построения трехмерных моделей микрообъектов являются методы трехмерной реконструкции по фокусировке [9, 15–21]. Точка трехмерной сцены выглядит на изображении тем более размыто, чем больше расстояние от этой точки до области резкости (см. ст. 14). Поэтому понятно, что по степени размытости точки на фото-изображении, можно судить о том, насколько велико расстояние от области резкости до данной точки в трехмерном пространстве. А значит и осуществить построение трехмерной модели поверхности объекта.

16

1.2. Обзор существующих методов

Методы реконструкции трехмерных моделей объектов по фокусировке делятся на два семейства: это методы семейства shape from focus (SFF) и методы семейства shape from defocus (SFD). Методы SFF [9, 15–17] для оценки положения в трехмерном пространстве некоторого участка сцены получают множество по-разному сфокусированных изображений сцены. После этого ищется изображение, на котором данный участок сцены виден в резкости и положение области резкости оптической системы для найденного изображения сцены принимается за оценку расстояния до данного участка сцены. Далее оценки расстояния для различных участков сцены комбинируются для получения трехмерной модели всей сцены.

Методы семейства SFD [18–21] моделируют степень резкости участка изображения как функцию расстояния до поверхности объекта. Затем, по двум или более по-разному сфокусированным изображениям сцены, формулируется обратная задача, решением которой является карта расстояний до поверхности сцены, которая полностью определяет трехмерную модель сцены.

1.2.1. Постановка задачи

Сформулируем задачу построения трехмерной модели поверхности микрообъекта по фокусировке более конкретно. Пусть у нас имеется монокулярный оптический микроскоп, на котором установлена цифровая камера. Предполагается, что у нас имеется возможность программно управлять фокусировкой микроскопа и его диафрагмой. Управление фокусировкой может достигаться как за счет взаимного перемещения линз в оптической системе, так и за счет перемещения всей оптической системы, или установленного под микроскопом образца, с помощью механической подвижки. Не ограничивая общности будем считать, что управление фокусировкой достигается за счет перемещения оптической системы относительно наблюдаемого объекта в направлении, параллельном оптической оси микроскопа. В случае если управление фокусировкой достигается за счет взаимного перемещения линз в оптической системе микроскопа, фотографиями с разными настройками фокуса будут иметь разное увеличение, что может быть скомпенсированно, как показано в [22].

Пусть с помощью микроскопа получен набор изображений сцены $\hat{I} =$ $\{I_1, \ldots, I_n\}$ с разным положением области резкости и, возможно, различным диаметром диафрагмы. Будем считать, что изображения в наборе \hat{I} являются полутоновыми. При этом будем считать, что для каждого из изображений известно положение области резкости в некоторой глобальной системе координат. Будем считать, что область резкости изображения I_i представляет собой область пространства между двумя плоскостями $L_{1,i}$ и $L_{2,i}$, расстояние между которыми равно глубине резкости оптической системы F_i . Плоскость, находящуюся посередине между $L_{1,i}$ и $L_{2,i}$ будем называть плоскостью резкости. Предположение о том, что область резкости ограниченна двумя плоскостями в пространстве является оправданным упрощением. Для реальных объективов состоящих из нескольких линз, область резкости может быть не плоской. Однако как показано в работе [22] кривизна области резкости может быть скомпенсирована за счет предварительной калибровки с помощью плоского калибровочного шаблона, расположенного перпендикулярно оптической оси микроскопа.

Тогда по такому набору изображений Î требуется построить трехмерную модель наблюдаемого участка объекта. В связи с тем, что съемка всех изображений ведется с одного направления, для представления трехмерной модели объекта удобно использовать карту глубины. То есть для представления трехмерной модели объекта можно использовать матрицу $D = \{d_{i,j}\}$, размеры которой равны размеру изображений набора \hat{I} .

1.2.2. Методы семейства SFD

Методы семейства shape from defocus [18–21, 23] моделируют степень резкости участка изображения как функцию расстояния до поверхности объекта. По двум или более по-разному сфокусированным изображениям сцены, формулируется обратная задача, решением которой является карта глубины, которая полностью определяет трехмерную модель сцены.

Для иллюстрации данной идеи рассмотрим один из первых методов shape from defocus, который был описан Пентландом в работе [23]. В своей работе Пентланд предложил использовать для построения трехмерной модели сцены пару изображений, то есть $\hat{I} = \{I_1, I_2\}$. В методе предложенном Пентландом оба изображения сцены сфокусированы одинаково. Однако эти изображения имеют разный диаметр диафрагмы. Первое изображение I_1 получается с максимально закрытой диафрагмой, то есть на этом изображении вся наблюдаемая сцена видна в резкости ¹. Второе изображение сцены I_2 получается с полностью открытой диафрагмой, то есть на этом изображении участки сцены, которые не находятся в пределах глубины резкости, расфокусированы.

Для моделирования эффекта размытия изображения Пентланд в своей работе предложил использовать фильтр Гаусса. То есть если рассмотреть некоторый небольшой участок сцены, то его расфокусированное изображение может быть представлено как:

$$I_{real}(x,y) = I_{sharp}(x,y) * G(x,y,\sigma)$$
(1.2)

¹ Стоит заметить, что в связи с тем, что для микрообъектов получение изображения, на котором вся сцена видна в резкости, невозможно из-за физических ограничений, данный метод применим только для макрообъектов

где $G(r, \sigma)$ представляет собой двумерный Гауссиан, параметр σ представляет собой степень размытия участка сцены, а * обозначает операцию свертки. В данной формуле I_{sharp} - идеально сфокусированная текстура поверхности участка сцены. В связи с тем, что в методе, предложенном Пентландом, первое изображение получается с полностью закрытой диафрагмой можно можно считать, что $I_{sharp} = I_1$.

Обозначим небольшой участок изображений I_1 и I_2 с центром в точек (x_0, y_0) как H_1 и H_2 . Тогда понятно, что $H_2(x, y) = H_1(x, y) * G(x, y, \sigma)$. Если рассмотреть Фурье-образы участков H_2 и H_1 имеем, что:

$$F_2(\lambda,\tau) = F_1(\lambda,\tau)G(\lambda,\frac{1}{\sqrt{2\pi\sigma}})$$
(1.3)

Благодаря тому, что свертка в пространственной области соответствует умножению в частотно-временной области можно перенести F_1 в левую часть равенства (1.3) и тем самым получить уравнение для поиска σ .

Тогда алгоритм построения карты глубины можно записать следующим образом:

- 1. Каждое из входных изображений I_1 , I_2 разбивается равномерно на небольшие блоки (размером, например, 10х10 пикселей)
- Для каждой пары соответствующих блоков рассчитывается преобразование Фурье
- 3. Далее для каждого блока с помощью уравнения (1.3) вычисляется степень размытия *σ*
- Далее для каждого блока степени размытия ставится в соответствие расстояние от центра области резкости изображения I₁. Это может быть сделано исходя из известных параметров модели линзы, как предлагается в работе [23]

Существует большое количество различных алгоритмов SFD. Большинство из них могут быть разбиты на две группы - локальные и глобальные. Примером локального алгоритма является описанный выше алгоритм Пентланда. Локальные алгоритмы [19, 21, 23, 24] разбивают входные изображения на набор небольших блоков и далее путем сравнения степени размытия пары соответствующих блоков строят оценку карты глубины для наблюдаемой сцены.

В отличии от локальных алгоритмов глобальные алгоритмы [18, 20] не осуществляют разбиение сцены на блоки. Глобальные алгоритмы SFD формулируют обратную задачу, записывая зависимость между полученными изображениями I_1 , I_2 и неизвестной картой глубины D

$$I_1 = f(D, s_1), I_2 = f(D, s_2)$$
(1.4)

где f - представляет собой функцию, которая моделирует процесс формирования изображения, а s₁ и s₂ представляют собой настройки фокусировки и диафрагмы для пары входных изображений. Исходя из формулы (1.4) глобальные методы SFD решают обратную задачу и тем самым получают целевую карту глубины D.

1.2.3. Методы семейства SFF

В отличии от методов семейства SFD, методы SFF [9, 15, 21, 22, 24] для оценки положения в трехмерном пространстве некоторого участка сцены получают множество по-разному сфокусированных изображений сцены. После этого ищется изображение, на котором данный участок сцены виден в резкости и положение плоскости резкости оптической системы для найденного изображения сцены принимается за оценку положения данного участка сцены в трехмерном пространстве. Для определения того, на каком



Рис. 1.3. Схематичная иллюстрация принципа работы алгоритмов семейства shape from focus. На графике справа горизонтальная ось соответствует номеру картинки, а вертикальная - значению меры резкости

изображении некоторый участок сцены виден наиболее четко используется т.н. *мера резкости* (см. рис 1.3). Далее оценки для различных участков сцены комбинируются для получения трехмерной модели всей сцены.

Таким образом в случае методов семейства SFF набор входных изображений $\hat{I} = \{I_1, \ldots, I_n\}$ состоит из большого количества входных изображений, потому как для получения корректной оценки геометрии сцены требуется, чтобы каждый участок наблюдаемой сцены был виден в резкости хотя бы на одном из входных изображений. Обычно набор изображений \hat{I} получается путем сканирования сцены с равномерным шагом (см. рис 1.3). То есть каждое изображение I_{k+1} отличается от изображения I_k положением плоскости резкости на некоторый шаг h, одинаковый для всех изображений.

1.2.4. Мера резкости

Методы семейства SFF опираются на поиск наиболее резкого изображений некоторого участка сцены среди набора фотографий $\hat{I} = \{I_1, \ldots, I_n\}$. Для того чтобы осуществить такой поиск алгоритмам SFF требуется численная мера, для измерения того, насколько резким выглядит участок изображения в окрестности пикселя (x_0, y_0) . Такая численная мера называется мерой резкости. В работах по методам SFF предлагается ряд различных мер резкости [25, 26].

Большинство предлагаемых мер резкости основаны на применении ко входным изображениям высокочастотных фильтров. Высокочастотные фильтры реагируют на наличие резких перепадов интенсивности на изображениях. Понятно, что если на изображении участок сцены находится за пределами области резкости оптической системы, то изображение этого участка будет выглядеть размытым, и резких деталей на изображении не будет. Следовательно отклик высокочастотного фильтра будет низким. В то же время если участок поверхности объекта виден на изображение в резкости изображение деталей поверхности объекта не будет размытым, и высокочастотный фильтр даст сильный отклик.

В работе Наяра [25] предлагается в качестве меры резкости использовать дискретный аналог т.н. модифицированного оператора Лапласа:

$$L(I) = \left| \frac{\partial^2 I}{\partial x^2} \right| + \left| \frac{\partial^2 I}{\partial y^2} \right|$$
(1.5)

Для расчета вторых производных по дискретной картинке используется свертка с ядрами $\begin{pmatrix} 1, -2, 1 \end{pmatrix}$ и $\begin{pmatrix} 1, -2, 1 \end{pmatrix}^T$. В работе Найяра экспериментальным путем было показано, что добавление модулей в формулу (1.5) улучшает качество работы меры резкости по сравнению с обычным оператором Лапласа.

В работе Кроткова [26] предлагается в качестве меры резкости использовать следующий оператор:

$$T(I) = \sqrt{\frac{\partial I^2}{\partial x} + \frac{\partial I^2}{\partial y}}$$
(1.6)

Данный оператор представляет собой модуль градиента в точек изображения и называется оператор Тененград. Для вычисления производных первого порядка в формуле (1.6) в работе [26] используется свертка с ядрами d_x и d_y , задаваемыми следующей формулой:

$$d_y = d_x^T, d_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$
(1.7)

На рис. 1.4 можно видеть пример результата расчет меры резкости Тененград для изображения природного минерала.

Основным плюсом оператора Тененград по отношению к оператору Лапласа является тот факт, что первые производные менее чувствительны к шуму на изображениях, чем вторые. Однако стоит заметить, что не существует универсальной меры резкости, которая была бы лучшей во всех ситуациях. Поэтому для каждого конкретного приложения, при реализации метода SFF, зачастую приходится подбирать меру резкости экспериментальным путем.

1.2.5. Базовая схема методов SFF

После того как введено понятие меры резкости можно описать базовую схему работы методов семейства SFF. Разработка методов трехмерной реконструкции семейства SFF началась еще в 80х годах прошлого века. С тех пор было разработано большое количество алгоритмов SFF, но большинство из них используют следующую схему, предложенную в одной из первых работы по методам семейства SFF [25]



Рис. 1.4. Изображение природного минерала, полученное с помощью оптического микроскопа, и мера резкости Тененград, рассчитанная по этому изображению. Можно заметить, что значения меры резкости больше в областях изображения, на которых присутствуют четкие детали.

- 1. Для каждого изображения из набора \hat{I} вычисляем меру резкости, тем самым получая карты значения меры резкости $\hat{S} = \{S_1, \ldots, S_n\}.$
- 2. Для каждого S_k проводим усреднение значений меры резкости $S_k \to \tilde{S}_k$ с целью подавления шумов
- 3. Для каждого пикселя (i, j) находим \tilde{S}_k , на котором достигается максимум величины меры резкости
- Положение плоскости резкости оптической системы для изображения *k* принимаем за оценку положения поверхности объекта в пикселе (*i*, *j*)

Второй шаг алгоритма является необходимым, для того, чтобы подавить колебания в значениях меры резкости, возникающие из-за шума камеры. Шумы во входных изображениях неизбежно оказывают сильное влияние на значения меры резкости, потому как большинство мер резкости основаны на высокочастотных фильтрах, которые чувствительны к шуму.

В работе Наяра [25] перед выполнением шага 3), который заключается в поиске максимального значения меры резкости в пикселе (x_0, y_0) , предлагается осуществить усреднение значений меры резкости по квадратному окну размера R с центром в пикселе (x_0, y_0) . Это эквивалентно обработке карт меры резкости S_k низкочастотным box-фильтром с ядром размера R.

$$\tilde{S}(x_0, y_0) = \sum_{i=1}^R \sum_{j=1}^R S(x_0 + i - R/2, y_0 + j - R/2)$$
(1.8)



a)

б)

Рис. 1.5. Пример усреднения меры резкости с помощью Гауссовского фильтра. На рисунке а) можно видеть изображение меры резкости, а на рисунке б) можно видеть результат обработки меры резкости фильтром

Другим популярным методом усреднения меры резкости является усреднение значений меры резкости с помощью фильтра Гаусса [16] (см. рис 1.5). Фильтр Гаусса превосходит box-фильтр потому что его отклик является поворотно-инвариантным. Также при использовании фильтра Гаусса веса значений меры резкости убывают по мере удаления от (x_0, y_0) что позволяет обеспечить более высокую точность по сравнению с box-фильтром. В работе [22] предлагается использовать для усреднения меры резкости расчет среднеквадратичного отклонения меры резкости по окну с центром в пикселе (x_0, y_0)

$$\tilde{S}(x_0, y_0) = \sum_{i=1}^R \sum_{j=1}^R (S(x_0 + i - R/2, y_0 + j - R/2) - S_{avg})^2$$
(1.9)

где S_{avg} - среднее по окрестности, в которой ведется расчет. Плюсом такой схемы является заниженный отклик меры резкости в областях с плавными перепадами яркости изображения, которые обычно соответствуют сильно размытой текстуре.

1.2.6. Уточнение модели

Большое количество работ по SFF посвящено уточнению модели, получаемой с помощью описанной базовой схемы. Главным недостатком описанной базовой схемы является необходимость производить усреднение меры резкости по окну в окрестности пикселя (x_0, y_0) .

В работе [15] авторы показали, что усреднение меры резкости по окну размером $R \ge R$ с помощью методов, описанных ранее, соответствует предположению о том, что наблюдаемая сцена в окрестности (x_0, y_0) является плоской и данная плоскость перпендикулярна оптической оси. В случае если участок поверхности имеет более сложную форму, или даже является плоскостью, которая не перпендикулярна оптической оси, усреднение меру резкости в окрестности пикселя (x_0, y_0) приводит к ухудшению точности оценки формы поверхности объекта.

Для решения данной проблемы авторы [15] предложили следующий алгоритм. На первом шаге происходит построение модели объекта с помощью базовой схемы SFF. Обозначим как S_v трехмерную матрицу, значения которой определяются по формуле $S_v(x, y, k) = S_k(x, y)$. То есть в ячейке (x, y, k) находится значение меры резкости изображения k в пикселе (x, y) без усреднения. Пусть оценка глубины, полученная с помощью базовой схемы в пикселе (x_0, y_0) равна d_0 . В этом случае авторы работы [15] предлагают осуществлять уточнение следующим образом.

В окрестности точки (x_0, y_0, d_0) размером $R \ge R \ge R \ge R$ методом полного перебора ищется плоскость суммарное значение меры резкости по которой максимально. При этом в суммарную меру резкости по плоскости вносят вклад только элементы $S_v(x, y, k)$, которые находятся в пределах данной окрестности. После того, как данная плоскость найдена, ищется пересечение этой плоскости и прямой проходящей через (x_0, y_0) параллельно оптической оси. Глубина полученного пересечения \tilde{d}_0 принимается за новую оценку глубины в пикселе (x_0, y_0) карты глубины D.

Применяя описанную процедуру для каждого пикселя карты глубины по отдельности можно осуществить уточнение всей модели поверхности объекта. Стоит заметить, что описанная процедура позволяет ослабить изначальное ограничение базовой схемы SFF до предположения, что поверхность объекта в окрестности точки (x_0, y_0) является плоской, но не обязательно перпендикулярной оптической оси.

Данная идея развивается в работах [17, 27, 28]. В работе [28] предлагается для моделирования участка поверхности в окрестности пикселя (x_0, y_0) использовать более сложную модель чем плоскость, а именно функцию, представляющую собой выход трехслойной нейронной сети. В работе [27] предлагается для каждого пикселя карты глубины искать оптимальную поверхность общего вида в окрестности (x_0, y_0, d_0) с помощью метода динамического программирования. Эти подходы позволяют еще больше повысить локальную точность получаемой модели по сравнению с [15].

В работе [29] рассматривается еще одна из проблем, которая связана с

использованием усреднения по небольшому окну с центром в (x_0, y_0) - это краевые эффекты. В случае если у наблюдаемой сцены в окрестности точки (x_0, y_0) проходит граница объекта, либо имеется сильный перепад глубины на поверхности объекта, то усреднение по окну также приводит к серьезной потери точности. Причем ни одна из обобщенных моделей поверхности [17, 27, 28] не в состоянии корректно обработать данную ситуацию.

Авторы [29] для решения этой проблемы предлагают использовать следующий метод. Пусть помимо изначального набора изображений с разной фокусировкой $\hat{I} = \{I_1, \ldots, I_n\}$ у нас имеется еще и изображение I_f на котором вся сцена видна в резкости (изображение с максимально закрытой диафрагмой). Тогда для усреднения значений меры резкости предлагается использовать следующий адаптивный метод

$$w(i,j) = G(I_f(x_0+i, y_0+j) - I_f(x_0, y_0), \sigma)$$
(1.10)

$$\tilde{S}(x_0, y_0) = \frac{\sum_{i,j=-R/2}^{R/2} w(i,j) S(x_0 + i, y_0 + j)}{\sum_{i,j=-R/2}^{R/2} w(i,j)}$$
(1.11)

где $G(x, \sigma)$ - одномерный Гауссиан с центром в нуле. Благодаря взвешиванию значений меры резкости с помощью формулы (1.11) достигается эффект адаптивного усреднения. Точки изображения, цвет которых существенно отличается от цвета пикселя (x_0, y_0) имеют низкий вес при усреднении. В связи с тем, что обычно в областях резкого перепада глубины и на границах объекта происходит резкая смена цвета данный метод позволяет добиться существенного улучшения точности работы метода SFF в областях резких перепадов глубины, даже при использовании довольно большого значения размера окна R. Данную идею можно также сочетать с использованием взвешивания функцией Гаусса в зависимости от расстояния до пикселя (x_0, y_0) . В работе [22] для решения проблемы краевых эффектов предлагается использовать уменьшение размер окна усреднения. При условии, что в окрестности точки (x_0, y_0) на изображении I_f находятся резкая граница авторы [22] предлагают использовать размер окна в два раза меньшей чем базовый.

1.2.7. Сравнение подходов SFF и SFD

Основным достоинством методов семейства SFD является тот факт, что для построения трехмерной модели сцены методам этого семейства требуется всего два изображения сцены. Данный факт позволяет осуществлять трехмерную реконструкцию быстро, и даже производить трехмерную реконструкцию для динамических сцен [24]. Методам семейства SFF для корректной работы требуется получить большой набор изображений сцены, что делает их не применимыми для динамических сцен. Однако для огромного количества приложений, таких как например дефектоскопия, где наблюдаемый объект статичен, данное ограничение не играет столь большой роли.

Главным недостатком методов семейства SFD является тот факт, что они в явном виде моделируют процесс формирования изображения, что может быть затруднительно для сложных оптических систем, таких как оптика микроскопа. А в случае если используемая методами SFD модель формирования изображения отличается от реальной методы SFD могут выдавать крайне нестабильные результаты. В тоже время методы семейства SFF могут быть без каких-либо изменений алгоритма и учета настроек и модели камеры использованы для любых оптических систем.

Именно эта универсальность делает методы семейства SFF столь привлекательными. Поэтому в данной работе рассматриваются именно эти методы и задача построения трехмерных моделей микрообъектов с помощью микроскопа решается именно с помощью SFF.

1.3. Недостатки существующих методов семейства SFF

Пожалуй, основной проблемой описанных выше методов SFF является неустойчивое поведение в областях сцены со слабой текстурой, или вовсе без текстуры. В случае если в пределах окна, на поверхности объекта, нет текстуры, изменения значений меры резкости, от изображения к изображению, будут возникать в основном из-за шума камеры и итоговая оценка глубины по окну будет крайне неустойчивой. Поэтому в областях без текстуры замеры, получаемые с помощью базовой схемы SFF зачастую оказываются *ложсными*. Ложные замеры в литературе по реконструкции трехмерных моделей по изображениям также называются *выбросами*.

Частичного решения проблемы можно добиться путем увеличения размера окна, по которому происходит усреднение меры резкости. В случае если размер нетекстурированной области меньше чем *R* из формулы (1.8) то при любом положении окна в окрестности данной нетекстурированной области в него попадет некоторое количество видимой текстуры. Однако стоит заметить, что для нетекстурированных областей большого размера использование данной идеи является нецелесообразным в связи с тем, что увеличение размера окна *R* ухудшает точность получаемой трехмерной модели.

Методы уточнения трехмерной модели [15, 17, 27, 28] также не в состоянии улучшить качество обработки областей без текстуры. Данные методы опираются на результат работы базовой схемы SFF как на первое приближение, и если это первое приближение было ошибочным, то ошибочным будет и результат уточнения получаемый с помощью этих методов. Метод с адаптивным взвешиванием, предложенный в [29] позволяет сделать размер окна достаточно большим без существенного ухудшения точности получаемой трехмерной модели. Однако данный метод все же не в состоянии корректно обрабатывать нетекстурированные области произвольного размера. Также данный метод требует для своей работы изображение сцены, на котором вся сцена находится в резкости. Такое изображение проблематично получить при работе с микрообъектами из-за физических ограничений.

Проблема, аналогичная нетекстурированным областям, возникает при работе с объектами, поверхность которых не является диффузной. Зачастую такие поверхности невозможно запечатлеть с помощью микроскопа без появления на изображениях бликов. В областях бликов на изображениях также полностью отсутствует текстура, и следовательно существующие алгоритмы SFF при работе с не диффузными объектами также будут давать большое количество ложных замеров. Аналогичная ситуация будет возникать при условии, что на изображении присутствуют недоэкспонированные области, яркость которых слишком мала. В таких областях результаты работы методов SFF также будут ошибочными.

Описанная проблема существенно ограничивает применимость методов семейства SFF для моделирования объектов реального мира. В работе [22] для решения данной проблемы предлагается процедура фильтрации ложных замеров. Авторы работы предлагают обработать изображение I_f , на котором вся сцена видна в резкости, с помощью меры резкости, и затем применить к этому изображению усреднение. Далее итоговое усредненное изображение \tilde{S}_f преобразуется в маску ненадежных замеров по следующей формуле

$$M = \begin{cases} \frac{1, \tilde{S}_f(i, j) > s_t}{0, \tilde{S}_f(i, j) \le s_t} \end{cases}$$
(1.12)

где s_t - порог на значение меры резкости, который подбирается вручную. После получения карты глубины D с помощью базовой схемы SFF авторы работы [22] предлагают отбросить все замеры в карте глубины, которые соответствуют нулям в M и провести интерполяцию для получения итоговой карты глубины D_M . Несмотря на то, что в оригинальном методе предлагается использовать для отброса ложных замеров изображение I_f данный метод можно применять даже в случае, если получение I_f невозможно из-за физических ограничений. В данном случае следует использовать в формуле (1.12) вместо \tilde{S}_f величину $max(\tilde{S}_1, \ldots, \tilde{S}_n)$, которая представляет собой матрицу, в каждой ячейке которой находится максимум значений меры резкости в данном пикселе по всем изображениям входного набора $\hat{I} = \{I_1, \ldots, I_n\}$.

Стоит заметить, что локальная процедура фильтрации с помощью правила (1.12) не позволяет обеспечить отброс всех ложных замеров. В процессе фильтрации будет отброшена часть верных замеров, и неизбежно будет пропущена часть ложных замеров.

В работе [30] авторами рассматривается задача трехмерного моделирования поверхности минералогических образцов, при работе с которыми также возникает проблема ложных замеров. Для решения этой проблемы авторами работы предлагается схема пост-обработки карты глубины. К итоговой карте глубины, получаемой с помощью базовой схемы SFF, сначала применяется медианный фильтр, для того чтобы удалить грубые ошибки измерения. После медианного фильтра авторы [30] предлагают применить фильтр Гаусса для получения итоговой карты глубины

К сожалению, при наличии на поверхности сцены крупных нетекстурированных областей, радиус фильтров для подавления ложных замеров должен быть довольно большим. А использование медианного и Гауссовского фильтра большого радиуса приводит к тому, что мелкие детали поверхности объекта теряются, что приводит к существенному ухудшению точности результатов метода.

Рисунок 1.6 иллюстрирует проблему ложных замеров при работе методов SFF с объектами реального мира. На рисунках а) и б) приведена пара входных изображений из набора $\hat{I} = \{I_1, \ldots, I_n\}$, состоящего из 100 изображений. При получении данного набора производилось сканирование небольшого участка поверхности природного минерала. В центральной части поверхности, видимой на изображениях с микроскопа, имеется углубление. Данное углубление освещено хуже, чем остальная поверхность объекта, поэтому при получении изображений области вне углубления были частично засвечены.



Рис. 1.6. На рисунках а) и б) можно видеть изображения природного минерала с двумя разными положениями области резкости. На рисунке в) изображена итоговая карта глубины, красными эллипсами отмечены ошибки в карте глубины

На рисунке в) приведены результаты работы метода SFF с применением схемы отброса ложных замеров (1.12) и применением медианного фильтра к итоговой карте глубины, для данного примера. Можно заметить, что полученная оценка карты глубины имеет резкие скачки, которые отмечены на рисунке в) с помощью красных эллипсов. Подобные ошибки являются иллюстрацией нестабильности существующих методов семейства SFF и являются серьезным ограничением при использовании методов SFF во многих реальных приложениях.

1.4. Мультифокус изображения

В связи с тем, что в работе будет рассматриваться большое количество примеров сцен, наблюдаемых с помощью оптического микроскопа для их описания требуется некоторое контактное представление. Необходимость компактного представления связана с эффектом глубины резкости. Для многих из рассматриваемых сцен, требуется получение порядка 100 изображений, для того чтобы каждый участок сцены был виден в резкости хотя бы на одном изображении. При таком количестве входных изображений вставка в текст изначального набора фото-изображений является невозможной. Именно поэтому для удобства требуется более сжатое представление.

В качестве такого компактного представления будут использоваться так называемые мультифокус-изображения. Алгоритмы синтеза мультифокус-изображений [31–35] получают на вход набор изображений, на каждом из которых в резкости видна лишь незначительная часть сцены, а на выходе выдают изображение, на котором все участки сцены видны в резкости. Некоторые методы построения мультифокус-изображений опираются на методы семейства SFF [31].

На рисунке 1.7 приведен пример мультифокус изображения для участка поверхности природного минерала. Мультифокус изображение было получено методом [31]. Видно, что на мультифокус-изображении видна в резкости сразу вся сцена целиком, в то время как на входных изображениях (рис. 1.7, а-д) в резкости виден лишь небольшой срез сцены.





Рис. 1.7. На рисунках (а-д) приведены изображения из входного набора, состоящего из 100 изображений, а на рисунке е) приведено мультифокус-изображение, рассчитанное по этому набору

д)

e)

1.5. Заключение

г)

В данной главе была описана актуальная для многих приложений задача построения трехмерной модели микрообъектов по изображениям с микроскопа. Обзор традиционных методов построения трехмерных моделей объектов по изображениям свидетельствует о том, что оптимальными методами для решения этой задачи являются методы семейств shape from focus (SFF) и shape from defocus (SFD). Также приведенный сравнительный анализ подходов SFF и SFD показывает что для задачи моделирования микрообъектов по изображениям больше пригоден метод SFF из-за того факта, что он не использует прямого моделирования процесса формирования изображения, что делает его более стабильным и универсальным.
В главе приведен обзор существующих методов SFF и отмечены достоинства и недостатки существующих алгоритмов. Основным недостатком существующих методов SFF является тот факт, что все они по сути отталкиваются от локальной схемы выбора наиболее резкого изображения по небольшому окну в окрестности пикселя (x_0, y_0) . По этим причинам существующие методы SFF дают крайне нестабильные результаты в областях без текстуры, пересвеченных либо недоэкспонированных областях.

Это приводит к существенным ограничениям в применимости существующих методов семейства SFF при работе с объектами, состоящими из таких материалов как металл, стекло и т.д. На поверхности таких объектов часто встречаются области с малым содержанием текстуры, пересвеченные либо недоэкспонированные области. Именно поэтому для успешного применения методов семейства SFF в реальных приложениях требуются новые, более устойчивые методы и алгоритмы.

Глава 2

Устойчивый метод построения поверхности произвольной формы

Для решения проблемы неустойчивости существующих методов *shape from focus* предлагается новый метод. Как было отмечено в предыдущей главе большинство существующих методов SFF локальны по своей природе. Ключевым отличием предлагаемого метода SFF от существующих аналогов является тот факт, что метод ищет решение задачи SFF с помощью глобальной минимизации функционала качества, называемого *функцией энергией*.

Используемая функция энергии состоит из двух частей - члены отвечающие за правдоподобие и члены отвечающие за регуляризацию решения. За счет использования в предлагаемом методе регуляризации удается добиться стабильного поведения алгоритма в областях с низким содержанием текстуры и областях, которые не содержат текстуры вовсе.

Для построения функции энергии, с помощью которой решается задача SFF используется теория Марковских случайных полей *(MRF)*.

2.1. Марковские случайные поля

Многие задачи машинного зрения могут быть успешно сформулированы в терминах теории Марковских случайных полей [36–40]. Теория Марковских случайных полей позволяет строить оценки различных пространственно-переменных величин по изображениям, при этом накладывая на эти величины определенные априорные ограничения. В качестве таких пространственно переменных величин могут выступать, например, значение диспаритета, в задаче стереореконструкции [37] и интенсивность изображения, в задаче подавления шума на изображениях [38].

Пусть у нас имеется некоторая неизвестная векторная величина $\hat{x} = \{x_1, \ldots, x_m\}$ значение которой требуется оценить. Будем называть эту величину состоянием и будем считать, что \hat{x} представляет собой некоторую случайную величину. Будем считать, что каждая из случайных величин x_i может принимать значения из заданного набора $\{1, \ldots, n\}$. Пусть помимо неизвестного состояния у нас для каждого x_i имеется т.н. множество сосседей N_i , которое представляет собой некоторое подмножество множества $\{1, \ldots, m\}$.

Будем обозначать $P(\hat{x} = \hat{X})$ как $P(\hat{x})$, а $P(x_i = X_i)$ как $P(x_i)$. Тогда в случае если для любого $i = \overline{1, m}$ верно, что

$$P(x_i|x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_m) = P(x_i|x_{N_i})$$
(2.1)

где x_{N_i} - подмножество элементов \hat{x} которое является соседями x_i , то случайная величина $\hat{x} = \{x_1, \ldots, x_m\}$ является Марковским случайным полем. Говоря другими словами \hat{x} представляет собой Марковское случайное поле в случае если плотность вероятности для элемента x_i зависит только от его непосредственных соседей, задаваемых множеством N_i .

Марковское случайное поле может выступать в качестве модели в задаче стереореконструкции, тогда неизвестное нам $\hat{x} = \{x_1, \ldots, x_m\}$ представляет собой вектор, в который собраны значения диспаритета для всех пикселей первого изображения стереопары [37]. При этом каждое x_i принимает значения от 0 до значения d_{max} , где d_{max} представляет собой максимальный допустимый диспаритет. В случае решения задачи подавления шума на полутоновом изображении в качестве $\hat{x} = \{x_1, \ldots, x_m\}$ выступает вектор значений реальной интенсивности пикселей без шума и каждое x_i принимает значения от 0 до 255.

2.1.1. Оценка параметров поля

Зачастую в большинстве приложений нет возможности наблюдать, или как-то непосредственно измерить неизвестное нам состояние Марковского поля $\hat{x} = \{x_1, \ldots, x_m\}$. Будем считать, что вместо этого у нас есть некоторое наблюдение $O = \{o_1, \ldots, o_m\}$. При этом будем считать, что нам известна функция правдоподобия

$$P(O|\hat{x}) = \prod_{i=1}^{m} P(o_i|x_i)$$
(2.2)

иметь наблюдение O при условии состояния поля \hat{x} . Функция $P(o_i|x_i)$ отвечает за моделирование шумов и зависимости наблюдаемых данных от неизвестного нам состояния системы. Например для задачи устранения шума на изображении в качестве наблюдения O выступает исходное зашумленное изображение I. А в качестве модели шума $P(o_i|x_i)$ можно выбрать например Гауссовское распределение

$$P(o_i|x_i) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(o_i - x_i)^2}{2\sigma^2}}$$
(2.3)

Принимая во внимание модель наблюдений O которая задается с помощью формулы (2.2) можно записать формулу Байеса

$$P(\hat{x}|O) = \frac{P(O|\hat{x})P(\hat{x})}{P(O)}$$
(2.4)

и тогда можно построить оценку максимума апостериорной вероятности для неизвестного нам \hat{x} максимизируя вероятность $P(\hat{x}|O)$. Если логорифмировать правую часть (2.4) и взять ее со знаком минус то описанную оценку \hat{x}_{est} можно получить путем

$$\hat{x}_{est} = \arg\min_{\hat{x}} (-\ln(P(O|\hat{x})) - \ln(P(\hat{x})))$$
(2.5)

Первое слагаемое выражения в правой части (2.4) можно разложить следующим образом исходя из (2.2)

$$\ln(P(O|\hat{x})) = \sum_{i=1}^{m} \ln(P(o_i|x_i))$$
(2.6)

Однако со вторым слагаемым формулы (2.5) в общем случае никакого упрощающего разложения сделать нельзя. И в общем случае оценка неизвестного состояния $\hat{x} = \{x_1, \ldots, x_m\}$ путем минимизации (2.5) представляет собой крайне сложную задачу. Однако в случае если допустить, что \hat{x} представляет собой Марковское случайное поле, минимизация (2.5) становится осуществимой.

2.1.2. Теорема Хаммерсли-Клифорда

Пожалуй самым главным результатом теории Марковских случайных полей для приложений компьютерного зрения является теорема Хаммерсли-Клифорда [41]. Пусть у нас имеется граф $G = \langle V, E \rangle$. Количество вершин данного графа равно m и каждой вершине графа поставлена в соответствие случайная величина x_i . Ребра графа E соответствуют множествам соседей N_i . То есть из вершины, которая соответствует x_i выходит $|N_i|$ ребер, которые идут в вершины, которые соответствуют соседям x_i .

В соответствии с теоремой Хаммерсли-Клифорда в случае если \hat{x} представляет собой Марковское случайное поле, то для плотности вероятности $P(\hat{x})$ верно

$$P(\hat{x}) = \frac{1}{Z} \exp(-\sum_{C} V_{C}(\hat{x}))$$
(2.7)

Суммирование в формуле (2.7) идет по всем кликам C в графе G, причем слагаемое $V_C(\hat{x})$ зависит только от тех переменных из \hat{x} , которые соответствуют вершинам графа G, которые принадлежат клике C. Величины $V_C(\hat{x})$ называются *потенциалами клик*.

Таким образом с учетом формул (2.7) и (2.6) можно переписать (2.5) в виде

$$\hat{x}_{est} = \arg\min(\sum_{i=1}^{m} -\ln(P(o_i|x_i)) + \sum_{C} V_C(\hat{x}))$$
(2.8)

Таким образом для случая Марковского случайного поля формулу (2.5) можно существенно упростить. И для определенных видов функции правдоподобия $P(o_i|x_i)$ и графа $G = \langle V, E \rangle$ существуют эффективные методы [36, 37, 39, 41–43] поиска минимума (2.8).

Отличительной чертой методов, построенных на основе Марковских случайных полей и минимизации (2.8) является то, что данные методы ищут решение задачи с помощью минимизации функционала, который включает в себя правдоподобие и члены, которые отвечают за регуляризацию решения. Использование регуляризации позволяет получать устойчивые результаты, даже в случае если во входных наблюдениях *O* имеются грубые ошибки. Именно этот факт делает теорию Марковских случайных полей столь привлекательной для создания устойчивого метода *shape from focus*.

2.2. Предлагаемый метод

2.2.1. Статистическая формулировка задачи

Для построения метода SFF с использованием теории Марковских случайных полей сформулируем задачу SFF в терминах математической статистики. Будем считать, что карта глубины D, которую нам необходимо построить по набору изображений, представляет собой Марковское случайное поле. Набор входных изображений $\hat{I} = \{I_1, \ldots, I_n\}$ в этом случае представляет собой наблюдаемые данные, тогда как D представляет собой неизвестное нам состояние случайного поля, оценку которого нам требуется построить. Будем считать что каждый элемент карты глубины может принимать значения из набора $\{1, \ldots, n\}$, то есть каждому пикселю карты глубины в результате решения должен быть приписан индекс одного из изображений из набора \hat{I} . Тогда по формуле Байеса имеем

$$P(D|\hat{I}) = \frac{P(\hat{I}|D)P(D)}{P(\hat{I})}$$
(2.9)

Использование входных изображений \hat{I} в качестве наблюдений напрямую при построении трехмерных моделей по фокусу является неудобным. Также как и все существующие методы SFF предлагается отказаться от использование входных изображений \hat{I} напрямую, и вместо этого использовать карты меры резкости $\hat{S} = \{S_1, \ldots, S_n\}$ (см. ст. 23). Тогда формулу (2.9) можно будет записать как

$$P(D|\hat{S}) = \frac{P(\hat{S}|D)P(D)}{P(\hat{S})}$$
(2.10)

Исходя из формулы (2.10) оценку максимальной апостериорной вероятности (MAP) для карты глубины *D* можно записать в виде

$$D_{map} = \arg\min_{D}(-\ln(P(\hat{S}|D)) - \ln(P(D)))$$
(2.11)

2.2.2. Функция правдоподобия

Для того, чтобы успешно оптимизировать функционал (2.11) необходимо в первую очередь осуществить упрощение первого слагаемого формулы, которое отвечает за правдоподобие решения. Для того, чтобы сделать это предлагается ввести предположение о том, что значения меры резкости в некотором пикселе (x_0, y_0) зависят статистически только от значения глубины d_0 в данном пикселе, но не зависят от значений глубины во всех остальных пикселях карты глубины D. Будем называть вектор

$$\hat{s}_{x_0,y_0} = \{S_1(x_0, y_0), \dots, S_n(x_0, y_0)\}$$
(2.12)

в котором расположены по порядку значения меры резкости в пикселе (x_0, y_0) профилем меры резкости в пикселе (x_0, y_0) . Тогда при условии сделанного предположения можно разложить $P(\hat{S}|D)$ как

$$P(\hat{S}|D) = \prod_{i,j} P(\hat{s}_{i,j}|D_{i,j})$$
(2.13)

где каждый из членов $P(\hat{s}_{i,j}|D_{i,j})$ представляет собой вероятность наблюдать заданный профиль меры резкости при условии определенного значения глубины D(i,j).

2.2.3. Структура поля

Для того, чтобы осуществить разложение второго члена в формуле (2.11) нам необходимо определить множества соседей для Марковского поля D, состояние которого нам необходимо оценить. На рис. 2.1 изображена предлагаемая структура графа $G_{sff} = \langle V, E \rangle$ который определяет множества соседей для элементов поля.

Граф представляет собой регулярную решетку, вершины которой соответствуют пикселям карты глубины. Каждый пиксель карты глубины соединен ребрами со своими непосредственными соседями по горизонтали и вертикали. То есть у угловых пикселей карты глубины имеется два соседа, у граничных пикселей карты глубины имеется по три соседа, а у



Рис. 2.1. На рисунке изображена структура графа $G_{sff} = \langle V, E \rangle$ который задает систему соседей Марковского случайного поля D для решения задачи SFF. Зеленые отметки на графе указывает на клики в графе размера один и два

пикселей карты глубины, которые не примыкают к границам имеется по четыре соседа. То есть для не граничного пикселя карты глубины $N_{i,j} = \{(i-1,j), (i+1,j), (i,j-1), (i,j+1)\}.$

При такой структуре соседей Марковского поля исходя из формулы (2.7) *P*(*D*) может быть записано следующим образом

$$P(D) = \frac{1}{Z} \exp\left(-\sum_{C_1} V_{C,1}(D) - \sum_{C_2} V_{C,2}(D)\right)$$
(2.14)

где $V_{C,1}(D)$ и $V_{C,2}(D)$ представляют собой потенциалы клик размера один и два соответственно. Потенциалы клик более высокого порядка в формуле (2.14) отсутствуют в связи с тем, что максимальный размер клики, задаваемой графом, который изображен на рисунке 2.1, равен двум.

Потенциал клик первого порядка $V_{C,1}(D) = V_{C,1}(D_{i,j})$ зависит от значения одного элемента карты глубины *D*. Потенциал первого порядка для элемента D_{x_0,y_0} заключает в себе априорную вероятность наблюдать некоторое конкретное значение глубины в пикселе (x_0, y_0) . Для некоторых приложений возможно наличие априорной информации о том, какой может быть глубины в пикселе (x_0, y_0) . Например в случае, если у нас уже есть какая-то оценка карты глубины полученная другим методом, информацию об этой оценке можно задать в виде потенциала клик первого порядка. Но в общем случае какая либо априорная информация о том, какие значения глубины в пикселе (x_0, y_0) более вероятны отсутствует. Поэтому в предлагаемом методе все потенциалы клик первого порядка равны нулю. ¹.

Потенциал клик второго порядка зависит сразу от пары соседних значений карты глубины D. То есть $V_{C,2}(D) = V_{C,2}(D_{i,j}, D_{k,p})$ где (i, j) и (k, p)являются соседями в графе G_{sff} . Данный потенциал заключает в себе априорную вероятность наблюдать конкретную конфигурацию пары соседних элементов карты глубины. В связи с тем, что форма многих объектов природного происхождения и объектов, сделанных человеком, зачастую не является абсолютно случайной, потенциал клик второго порядка можно использовать для учета априорной информации о форме поверхности объекта. Например для гладких объектов значения глубины в соседних пикселях не должны отличаться друг от друга слишком сильно. Поэтому в отличии от потенциала клик первого порядка в предлагаемом методе $V_{C,2}(D)$ не равен нулю и будет учитываться при минимизации (2.11). Таким образом

$$P(D) = \frac{1}{Z} \exp(-\sum_{C_2} V_{C,2}(D))$$
(2.15)

¹ Стоит заметить, что это не является ограничением предлагаемого метода. В случае если предлагаемый метод применяется к задаче, для которой присутствует априорная информация о вероятности наблюдать определенные значения карты глубины в некоторых пикселях эта информация может быть учтена без изменения метода

2.2.4. Итоговая функция энергии

С учетом допущений, сделанных в прошлых двух подразделах, запишем итоговую функцию для оптимизации (2.11) в более удобном виде. Для простоты обозначений запишем карту глубины D в виде одномерного вектора $\hat{d} = \{d_1, \ldots, d_m\}$. Также поставим в соответствие каждому элементу d_i профиль меры резкости \hat{s}_i соответствующего ему пикселя карты глубины.

Исходя из (2.13) можно записать $-\ln(P(\hat{S}|D))$ как сумму

$$-\ln(P(\hat{S}|D)) = -\sum_{i=1}^{m} \ln(P(\hat{s}_i|d_i)) = \sum_{i=1}^{m} V_i(d_i)$$
(2.16)

Величины $V_i(d_i)$ в формуле (2.16) принято называть унарными потенциалами поля [37]. Унарный потенциал поля заключает в себе информацию о правдоподобии ².

Исходя из (2.15) можно записать $-\ln(P(D))$ как сумму

$$-\ln(P(D)) = -\ln(\exp(-\sum_{C_2} V_{C,2})) = \sum_{i=1}^{m} \sum_{d_k \in N(d_i)} U_{i,k}(d_i, d_k)$$
(2.17)

где d_k - соседи d_i . Величины $U_{i,k}(d_i, d_k)$ называются *парными потенциалами* поля [37] и заключают в себе априорную информацию о форме поверхность объекта, трехмерную модель которого нам требуется построить. С учетом (2.16) и (2.17) можно переписать (2.11) как

$$E(\hat{d}) = \sum_{i=1}^{m} (V_i(d_i) + \lambda \sum_{d_k \in N(d_i)} U_{i,k}(d_i, d_k))$$
(2.18)

$$\hat{d}_{map} = \arg\min_{\hat{d}} \left(E(\hat{d}) \right) \tag{2.19}$$

² В случае если в приложении имеется априорная информация о вероятности наблюдать определенные значения карты глубины в некоторых пикселях то данную информацию можно также включить в унарный потенциал путем простого суммирования

Функция $E(\hat{d})$ называется функцией энергии. Величина λ в формуле в (2.18) представляет собой вес парных потенциалов и определяет силу регуляризации решения. Определив унарные и парные потенциалы, а затем оптимизировав функцию энергии в формуле (2.18) можно получить оценку карты глубины (2.19), которая является оценкой максимума апостериорной вероятности.

2.2.5. Унарный потенциал

Так как унарный потенциал отвечает за правдоподобие, он должен быть тем меньше, чем больше вероятность наблюдать имеющийся набор входных данных при условии некоторой заданной конфигурации поля. Исходя из упрощающего предположения, сделанного в разделе 2.2.2 унарный потенциал $V_i(d_i)$ должен иметь вид

$$V_i(d_i) = V(d_i, \hat{s}_i) \tag{2.20}$$

где $\hat{s}_i = \{s_1, \ldots, s_n\}$ - профиль меры резкости в пикселе, который соответствует d_i . То есть унарный потенциал элемента поля d_i не должен зависеть от значений меры резкости в каких-то других пикселях изображения, кроме пикселя, который соответствует d_i . Простейший пример унарного потенциала, который можно применить при решении задачи SFF выглядит следующим образом

$$V(d_i) = \begin{cases} 0, d_i = i_{max} \\ 1, d_i \neq i_{max} \end{cases}$$
(2.21)

где i_{max} это индекс максимального элемента в профиле меры резкости. При использовании унарного потенциала вида (2.21) функция энергии (2.18) будет штрафовать карту глубины D на величину 1, в случае если она проходит не через точку с максимальным значением меры резкости в пикселе, который соответствует d_i . В случае если карта глубины D проходит через максимум меры резкости в пикселе, который соответствует d_i то за данный пиксель карты глубины будет получать нулевой штраф.



Рис. 2.2. Примеры модельных профилей меры резкости для иллюстрации недостатков унарного потенциала из формулы (2.21). Профиль на рисунке а) имеет один ярко выраженный главный пик слева, а также один менее значительный пик справа. Профиль на рисунке б) имеет практически случайную структуру

Однако потенциал, задаваемый формулой (2.21) не является оптимальным. На рисунке 2.2 изображены два примера профилей меры резкости. Профиль на рисунке 2.2 а) имеет один главный пик, но помимо главного пика в профиле также присутствует менее значительный пик слева. При условии использования унарного потенциала (2.21) любая карта глубины, которая не проходит через главный пик будет оштрафована.

Однако понятно, что может оказаться, что главный пик в профиле меры резкости может соответствовать не изображению, в которой входит в резкости поверхность объекта. Главный пик в профиле меры резкости может соответствовать изображению, на котором в окрестности пикселя оказалась граница резкого блика³. В этом случае реальной поверхности

³ При изменении положения области резкости/перефокусировке границы бликов могут смешать-

объекта будет соответствовать второй, более слабый, пик в профиле меры резкости. Поэтому крайне желательно, чтобы унарный потенциал штрафовал прохождение через второй по величине пик в профиле на рисунке 2.2 а) слабее, чем области, которые удалены от обоих пиков.

Для примера, изображенного на рисунке 2.2 а) также является желательным, чтобы унарный потенциал $V(d_i, \hat{s}_i)$ давал менее значительный штрафы в непосредственной близости от обоих пиков на графике. Это связано с тем, что в профилях меры резкости для реальных объектов могут присутствовать колебания, возникающий в следствии шума камеры. Предположим, что в модельный профиль меры резкости 2.2 а) добавлен нормально распределенный шум с некоторой дисперсией σ . Понятно, что из-за шума положение главных пиков в профиле меры резкости может сместиться на одно-два изображения. Поэтому важно, чтобы унарный потенциал давал незначительный штрафы в непосредственной близости от существенных пиков в профиле резкости.

Модельный профиль меры резкости на рисунке 2.2 б) иллюстрирует еще одну проблему простейшего унарного потенциала (2.21). Профиль, изображенный на рисунке 2.2 б) имеет практически случайную структуру. В нем имеется множество пиков, которые приблизительно равнозначны, и предпочесть какой-то конкретный из пиков в этом профиле меры резкости невозможно. Ситуация, изображенная на рисунке 2.2 б) является типичной для областей без текстуры. В таких областях все вариации в мере резкости связаны исключительно с шумом камеры, и имеют полностью случайный характер. Крайне желательно, чтобы в ситуации, которая изображена на рисунке 2.2 б) унарный потенциал не отдавал явного предпочтения никакому конкретному значению глубины *d*. В этом случае значение глубины в данном пикселе будет полностью определяться парным потенциалом, ко-

ся на изображении. Данный эффект проиллюстрирован на рис. 2.3

торый отвечает за регуляризацию.

С учетом сформулированных выше пожеланий к поведению унарного потенциала предлагается следующий вид унарного потенциала

$$V(d_i) = \min\left(\frac{(\max \hat{s}_i - \hat{s}_i(d_i))^2}{\sigma_s^2}, T_d\right)$$
(2.22)

Видно, что данный вид унарного потенциала равен нулю в точке максимальной резкости, также как и потенциал (2.21). Поэтому данная форма унарного потенциала способствует тому, чтобы итоговая поверхность проходила через резкие области. Но в отличии от (2.21) штраф, накладываемый унарным потенциалом (2.22) зависит от величины меры резкости, которая соответствует значению глубины d_i . Именно поэтому данный вид унарного потенциала не будет накладывать столь большие штрафы за прохождение поверхности рядом с максимумом в профиле меры резкости и через пики в мере резкости, которые не являются главными.

Величина σ_s^2 представляет собой дисперсию шума в профиле меры резкости, который вызван шумом камеры. Величина дисперсии шума может быть измерена для любой конкретной цифровой камеры априорно, перед построением трехмерной модели. Деление на σ_s^2 позволяет сделать унарный потенциал менее чувствительным к перепадам меры резкости, которые связаны исключительно с шумом камеры. Поэтому в областях с низким содержанием текстуры, или полным ее отсутствием, данный вид унарного потенциала не будет отдавать предпочтения какому-то конкретному значению глубины.

Величина T_d ограничивает значение штрафа, накладываемого унарным потенциалом сверху и является входным параметром алгоритма. Такое ограничение сверху необходимо для того, чтобы штрафы за прохождение через область которая выглядит абсолютно размытой, не были слишком большими. Необходимость такого ограничения сверху иллюстрируется примером, приведенным на рисунке 2.3.



Рис. 2.3. На рисунке проиллюстрирован пример смещения блика при перефокусировке микроскопа. На рисунках а) и б) приведены два изображения природного минерала с разными положениями области резкости. На рисунках в) и г) приведены увеличенные версии участков изображений а) и б), которые отмечены красными квадратами.

На данном рисунке приведен пример двух изображений поверхности природного минерала с разными положениями области резкости, а также увеличенные версии изображений участка поверхности, который содержит блик. Участок поверхности вокруг блика, который изображен на рисунке 2.3 в) практически не содержит текстуры. Видно, что на изображении 2.3 г) граница блика существенно смещается относительно своего положения на рисунке 2.3 в). При этом граница блика на рисунке 2.3 г) (отмечено зеленой линией) лишь слегка размыта. В областях, куда смещается граница блика, на поверхности объекта отсутствует текстура. Поэтому в данных областях профили меры резкости будут содержать единственный сильный пик, которые вызван исключительно смещение блика из-за перефокусировки, а не наличием на поверхности объекта реальной текстуры. Поэтому благодаря ограничению штрафа за прохождение через размытые области сверху с помощью константы T_d можно снизить негативное влияние из-за подобных грубых ошибок измерения.



Рис. 2.4. Поведение унарного потенциала (2.22) на реальном примере. На рисунке а) изображено мультифокус-изображение участка поверхности природного минерала. На рисунке б) приведено изображение значений унарного потенциала по линии, которая отмечена на рисунке а)

На рисунке 2.4 приведена иллюстрация поведения предлагаемого в работе унарного потенциала на реальном примере. На рисунке 2.4 а) приведен пример мультифокус-изображения поверхности природного минерала. Также в нижней части рисунка 2.4 а) отмечена горизонтальная линия, для которой иллюстрируются значения унарного потенциала. На рисунке 2.4 б) приведено изображение значений унарного потенциала по данной линии. Ось x на рисунке 2.4 б) соответствует оси x на изображении а). Ось *у* на рисунке 2.4 б) соответствует глубине *d*. Белый цвет на рисунке 2.4 б) значению унарного потенциала, равному *T_d*, а черный цвет соответствует нулевому значению унарного потенциала.

Можно заметить, что большая часть поверхности под линией имеет достаточную текстуру и в значениях унарного потенциала наблюдаются четкие минимумы. В тоже время в области, где присутствует блик (правая часть линии) предлагаемая форма унарного потенциала не отдает предпочтения никакому конкретному значению глубины. Поэтому в области блика форма поверхности будет определяться в основном за счет парного потенциала.

2.2.6. Парный потенциал

Парный потенциал в формуле (2.18) заключает в себе априорную вероятность наблюдать некоторую карту глубины *D*. Этот потенциал отвечает за регуляризацию решения. Парный потенциал имеет решающее значение в ситуациях, когда во входных данных содержатся грубые ошибки измерений. Именно благодаря регуляризации в областях без текстуры, областях бликов и их окрестностях можно добиться стабильного поведения алгоритма оценки формы поверхности объекта.

Стоит заметить, что если приравнять парный потенциал в формуле (2.18) нулю, то решением задачи минимизации (2.18) будет поверхность, которая в каждой точке проходит через минимум унарного потенциала. Таким образом без использования возможностей парного потенциала метод основанный на (2.18) превращается в обычный локальный метод SFF, который в каждой точке ищет наиболее резкое изображение поверхности объекта.

Понятно, что поверхность объектов реального мира, наблюдаемых с

помощью микроскопа, обычно не является абсолютно произвольной. Зачастую поверхность объектов является гладкой, либо же кусочно-гладкой. Задав в форме парного потенциала ограничение кусочной гладкости можно получить метод, который будет существенно превосходить локальные методы SFF за счет использования априорных знаний о наиболее вероятной форме поверхности объекта.

В литературе по компьютерному зрению встречается целый ряд различных парных потенциалов. Примером популярного парного потенциала является т.н. обобщенная модель Потса *(GPM)*

$$U_{i,k}(d_i, d_k) = \begin{cases} 0, d_i = d_k \\ 1, d_i \neq d_j \end{cases}$$
(2.23)

Видно, что данный вид парного потенциала накладывает штраф на любую пару соседних пикселей, значения глубины в которых отличается друг от друга. Таким образом данный парный потенциал способствует тому, чтобы карта глубины имела как можно меньше вариаций в значениях соседних пикселей.

Обобщенная модель Потса активно применяется при решении различных задач машинного зрения, таких как, например, многоклассовая семантическая сегментация [44]. Однако, для построения моделей гладких поверхностей данная форма потенциала не совсем подходит. Как отмечено в работе [39] обобщенная модель Потса соответствует априорному предположению о том, что наблюдаемая сцена представляется в виде набора областей, каждой из которых соответствует некоторое константное значение глубины *d*. В то время как для задачи семантической сегментации такое предположение является оправданным, в задаче построения модели объекта по фокусу более реалистичной моделью является предположение о кусочной гладкости реконструируемой поверхности. В работе [37] описываются парные потенциалы, которые больше соответствуют предположению о кусочной гладкости поверхности.

$$U_{i,k}(d_i, d_k) = \min(|d_i - d_k|, T_s)$$
(2.24)

$$U_{i,k}(d_i, d_k) = \min((d_i - d_k)^2, T_s)$$
(2.25)

Отличием этих парных потенциалов от обобщенной модели Потса состоит в том, что штраф, накладываемый ими тем больше, чем больше разность между значениями глубины в соседних пикселях изображения. Именно поэтому данные виды парных потенциалов более предпочтительны при решении задачи трехмерной реконструкции объектов [37].

Дополнительно стоит отметить, что парные потенциалы в формулах (2.24) и (2.25) ограниченны сверху величиной T_s . Данное ограничение необходимо для того, чтобы в областях резких перепадов глубины на объекте, или в области границы объекта, парный потенциал не накладывал на перепад глубины слишком большой штраф. Парные потенциалы, которые ограничивают штраф сверху таким образом называются *сохраняющими скачки* [40] ⁴. Если отказаться от использования ограничения сверху значения парных потенциалов (2.24) и (2.25) то области сильных перепадов глубины на реальном объекте будут сглаживаться, что будет приводить к ошибкам [37].

Парный потенциал предлагаемый автором для решения задачи SFF основывается на (2.25). Причина, по которой данный вид парного потенциала предпочтительней чем (2.24) проиллюстрирована на следующем модельном примере.

⁴ Стоит заметить, что обобщенная модель Потса также является парным потенциалом, сохраняющим скачки



Рис. 2.5. Иллюстрация проблем в поведении парного потенциала, задаваемого формулой (2.24) при реконструкции трехмерных моделей гладких поверхностей. Черно-белый фон представляет собой значения унарного потенциала. Красными линиями отмечены различные решения, которые имеют штраф, накладываемый парным потенциалом, такой же как и зеленое решение, которое является оптимальным

Пусть имеется одномерная задача оценки формы поверхность объекта и ее решение осуществляется путем минимизации функционала

$$E(\hat{x}) = \sum_{i=1}^{N} V_i(x_i) + \sum_{i=1}^{N-1} U_i(x_i, x_{i+1})$$
(2.26)

Пусть элементы вектора $\hat{x} = \{x_1, \ldots, x_m\}$ принимают значения из множества $\{1, \ldots, n\}$. Пусть унарные потенциалы $V_i(x_i)$ принимают значения в соответствии с изображением на рисунке 2.5. На модельном примере из рисунка 2.5 унарные потенциалы в левой и правой частях картинки имеют ярко выраженный минимум, в то время как унарные потенциалы по центру картинки не отдают предпочтения ни какому конкретному значению x.

Понятно, что желаемым результатом оценки является кривая, которая в левой и правой частях картинки проходит через области минимума унарных потенциалов, а в центральной области гладко связывает левую и правую части. В частности минимум обоих частей формулы (2.26), при использовании парного потенциала (2.24), достигается, если связать левую и правую части с помощью прямой линии. Данное решение изображено на рисунке 2.5 зеленой линией. Однако стоит заметить, что существует множество решений, которые обладают точно таким-же штрафом парного потенциала, как и оптимальное зеленое решение. Эти решения изображены на рисунке красными линиями. Поэтому в случае если во входных данных, через которые рассчитывается унарный потенциал, будет присутствовать шум, то выбор между этими различными вариантами решения будет зависеть исключительно от того как будет распределен шум на конкретном примере. Именно эта нестабильность потенциала (2.24) делает более предпочтительным для использования в задаче SFF потенциала (2.25).

Исходя из приведенных выше доводов парный потенциал, предлагаемый в работе, имеет следующий вид

$$U(d_p, d_k) = \begin{cases} \frac{\nu \min\left((d_p - d_k)^2, T_s\right), \min(w_p, w_k) = 1}{\nu (d_p - d_k)^2, \min(w_p, w_k) = 0} \end{cases}$$
(2.27)

Коэффициент ν представляет собой квадрат отношения между размером пикселя изображений, и шагом по глубине между парой соседних изображений в наборе $\hat{I} = \{I_1, \ldots, I_n\}$. Использование нормировочного коэффициента ν позволяет сделать разность $(d_p - d_k)^2$ пропорциональной квадрату Евклидова расстояния в трехмерном пространстве.

Величина w_p представляет собой индикатор наличия сильной текстуры в окрестности пикселя, который соответствует d_p . В случае если сильная текстуры присутствует как в окрестности пикселя d_p , так и в окрестности его соседа d_k , парный потенциал между этими пикселями рассчитывается в соответствии с верхней частью формулы (2.27). Видно что эта часть формулы ограничена сверху величиной T_s , таким образом предлагаемый парный потенциал разрешает скачки для пикселей, в окрестности которых присутствует сильная текстура. Для пикселей в окрестности которых нет сильной текстуры величина парного потенциала не ограничена сверху, поэтому в таких пикселях резкие перепады глубины блокируются парным потенциалом (2.27).



Рис. 2.6. На рисунке приведен пример расчета карты наличия сильной текстуры на реальном примере. На рисунке а) приведено мультифокус-изображение поверхности природного минерала, а на рисунке б) приведена карта наличия сильной текстуры, рассчитанная описанным методом

Индикатор наличия сильной текстуры в окрестности элемента поля d_p рассчитывается следующим образом. Вычислим матрицу W, элемент (i, j)которой равен максимальному значение меры резкости в пикселе (i, j) среди всех изображений набора \hat{I} . После этого рассчитаем маску W_t , каждый элемент которой равен 1 если соответствующий пиксель W превосходит некоторый порог μ , и равен 0 в противном случае. После этого полученную таким образом маску предлагается отфильтровать с помощью математической морфологии операцией "открытие"с круглым структурным элементом радиуса R. После этого полученную отфильтрованную маску предлагается записать в виде одномерного вектора w_p , каждый элемент которого является индикатором наличия существенной текстуры в окрестности соответствующего элемента поля. За счет такой формы w_p предлагаемый парный потенциал допускает сильные перепады в карте глубины только в областях, где присутствует значительное количество текстуры. Величины μ и R являются входными параметрами алгоритма. На рисунке 2.6 приведен пример расчета карты наличия сильной текстуры для реального примера.

2.2.7. Оптимизация функционала

В прошлых разделах были определены унарный и парный потенциалы $V_i(d_i)$ и $U_{i,k}(d_i, d_k)$. Таким образом функция энергии из (2.18) теперь полностью определена и теперь необходимо определить, как функцию энергии данного типа оптимизировать. Оптимизация функций энергии вида (2.18) является темой активных исследований. Как показано, например, в работе [37] поиск точного минимума функции энергии вида (2.18) в общем случае представляет собой *NP-hard* задачу.

Однако, несмотря на то что поиск точного минимума данной энергии не представляется реальным, с учетом мощностей современных компьютеров, существуют и активно развиваются методы приближенного поиска минимума (2.18). К таким алгоритмам относятся метод Iterated Conditional Modes (ICM) [45], методы семейства Belief Propagation (BP) [46], метод Tree-Reweighted Message Passing (TRW) [47] а также метода на основе разрезов графов α - expansion и $\alpha\beta$ - swap [37]. Каждый из упомянутых методов обладает своими достоинствами и недостатками.

В работе [45] приводится сравнительный анализ сходимости и качества результатов этих методов на реальных данных из пяти различных задач машинного зрения. Метод ICM представляет в основном исторический интерес, в связи с тем, что по сути является простейшим итерационным методом локальной минимизации. Для того, чтобы данный метод сошелся корректно ему требуется очень точное локальное приближение, которое в большинстве случаем затруднительно получить. В соответствии со сравнительным анализом в [45] метод ВР по сравнению с методами на основе разрезов графов и методом TRW сходится крайне медленно и зачатую минимум, к которому он сходится, обладает существенно большей энергией, чем результаты работы методов α - expansion, αβ - swap и TRW. Также авторы работы отмечают, что на их тестовых данных алгоритм αβ - swap практически никогда не достигает минимумов лучше, чем α - expansion и зачастую сходится медленнее.

Из оставшейся пары методов α - ехраnsion и TRW для решения задачи SFF предлагается использовать первый метод. В соответствии со сравнением в работе [45] метод TRW сходится существенно дольше, чем метод α -ехраnsion. Однако минимумы к которым он сходится зачастую имеют меньшую энергию, чем результаты α - ехраnsion. Однако авторы работы [45] отметили следующий факт. Несмотря на то, что минимумы к которым сходится TRW обладают меньшей энергией чем минимумы α - ехраnsion качественно результаты работы этих алгоритмов практически неотличимы. Для большинства тестовых задач авторы [45] имели эталонный результат, полученный путем ручной или автоматической разметки ⁵. В соответствии с замерами, приведенными в [45] результаты алгоритмов α - ехраnsion и TRW зачастую имеют существенно меньшую энергию, чем эталон. Поэтому ошибки в результатах этих методов связаны не с качеством минимизации функционала, а скорее с несовершенством модели задаваемая энергией (2.18).

Поэтому из соображений компромисса между вычислительной сложностью и качеством результата в данной работе используется алгоритм α

61

⁵ Для задачи стереореконструкции, например, проводилось сравнение с результатами лазерного сканирования

- expansion. Несмотря на то, что из-за формы парного потенциала, предложенного в (2.27), функция энергии (2.18) является нерегулярной [37] алгоритм α -expansion по прежнему применим для минимизации, как было показано в [48].

2.2.8. Усреднение меры резкости

Как можно заметить, предлагаемый метод, в отличии от существующих аналогов [9, 15, 21, 22, 24], не использует усреднение меры резкости. В то время как существующие методы сначала осуществляют усреднение меры резкости по окну, унарные потенциалы в предлагаемом методе рассчитываются напрямую по изначальным картам меры резкости $\hat{S} = \{S_1, \ldots, S_n\}.$

Данная особенность связана с тем, что в формуле энергии (2.18) фигурирует сумма всех унарных потенциалов $V_i(d_i)$. Поэтому предлагаемый метод осуществляет неявное усреднение значений меры резкости по всей поверхности итоговой карты глубины D в процессе решения оптимизационной задачи (2.19).

Однако усреднение меры резкости в предлагаемом методе можно использовать при условии что есть необходимость ускорить работу алгоритма. Пусть нам требуется построить трехмерную модель участка поверхности микрообъекта, наблюдаемый с помощью микроскопа. Путь величина глубины резкости оптической системы составляет 10 микрон, а размер пикселя на изображении составляет один микрон ⁶.

Понятно, что из-за того факта, что глубина резкости составляет 10 микрон точность, предлагаемого метода не может быть больше, чем половина глубины резкости, то есть абсолютная точность реконструкции огра-

⁶ Предполагается, что модель проекции ортографическая

ничена величиной в 5 микрон. Понятно что при этом вместо того, чтобы использовать для расчета унарных потенциалов карты меры резкости изначального разрешения можно уменьшить карты меры резкости в 5 раз, и при этом потери в абсолютной точности получаемого результата не произойдет. Однако при этом удастся сократить количество неизвестных в (2.18) в 25 раз, что позволит существенно быстрее минимизировать функцию энергии.

2.2.9. Итоговый алгоритм

С учетом выкладок и замечаний приведенных в данном разделе итоговый алгоритм SFF на основе Марковских случайных полей выглядит следующим образом. Пусть у нас имеется набор входных изображений $\hat{I} = \{I_1, \ldots, I_n\}$, полученный в результате сканирования поверхности наблюдаемого объекта с регулярным шагом

- 1. Для каждого изображения набора вычисляем меру резкости $I_k \to S_k$
- 2. При наличии возможности осуществляем уменьшение карт меры резкости как описано в 2.2.8
- По картам меры резкости S_k вычисляем для каждого пикселя изображения значение унарного потенциала V_i(d_i) с помощью формулы (2.22), тем самым получая трехмерный массив значений унарного потенциала V_{i,j,k}
- 4. По картам меры резкости S_k вычисляем карту наличия существенной текстуры в соответствии с написанным в 2.2.6
- 5. С помощью алгоритма α-expansion осуществляем оптимизацию (2.18) с парным потенциалом в виде, задаваемом (2.27)

2.3. Эксперименты

2.3.1. Тестовая базы для сравнения

В отличии от задачи стереореконструкции [7] для тестирования методов shape from focus не существует стандартной тестовой базы изображений. Поэтому для оценки качества работы предлагаемого метода и сравнения метода с существующими аналогами автором была собрана тестовая база из 27 наборов изображений с микроскопа. Для создания тестовой базы использовался микроскоп *Leica Z6 APOA*, с установленной на него камерой разрешения 1280х1024. Все изображения для экспериментов были получены в градациях серого. Сканирование объектов производилось с максимальным увеличением, глубина резкости для которого на данном микроскопе составляет 15 микрон. Величина пикселя изображения, для максимального увеличения, в данной системе составляет порядка 2 микрон. Шаг между соседними изображениями наборов по фокусу был равен половине глубины резкости. В каждом из тестовых наборов содержалось от 50 до 200 изображений.

В качестве объектов для тестирования использовались полупрозрачные природные минералы, в частности алмазы. Такой выбор объектов обусловлен тем, что поверхность полупрозрачных минералов часто не содержит текстуры. Также, часто при их съемке возникают блики, поэтому, моделирование таких объектов представляет сложность для существующих методов трехмерной реконструкции семейства SFF. На максимальном увеличении минерал обычно целиком не помешается в поле зрения, поэтому для каждого образца сканировались несколько участков его поверхности с разных углов зрения.

На рисунке 2.7 приведена иллюстрация одного из тестовых наборов из собранной тестовой базе. Как можно заметить по мультифокус-изоб-



Рис. 2.7. На рисунке а) приведено изображение алмаза, установленного под микроскопом. Красным кругом на рисунке отмечена каверна на поверхности алмаза. На рисунке б) приведено мультифокус-изображение этой каверны, полученное по набору изображений с микроскопа с максимальным увеличением

ражению участка поверхности минерала, приведенному на 2.7 б), данный пример содержит большое количество областей без текстуры, и поэтому представляет серьезный интерес для тестирования алгоритмов SFF.

Для оценки качества работы предложенного метода и сравнения его с существующими аналогами крайне желательно иметь объективный способ сравнения результатов работы методов. Поэтому в среде программирования MATLAB автором был реализован инструмент для ручной разметки карты глубины по набору изображений. Данный инструмент позволяет просматривать изображения набора и расставлять на этих изображениях контрольные точки. Контрольные точки необходимо выставлять на изображениях в тех местах, где на них видны резкие детали.

По набору размеченных контрольных точек разработанный инструмент позволяет построить эталонную карту глубины набора. В областях, где контрольные точки не были установлены (в частности это области без текстуры), значение глубины для эталонной карты глубины вычисляется с помощью интерполяции . В областях без текстуры, которые примыкают к границам изображения, провести интерполяцию корректно невозможно. Поэтому в такие области отмечались вручную и результаты работы алгоритмов в этих областях не учитывались.



Рис. 2.8. На рисунке а) приведено мультифокус-изображение участка поверхности алмаза. На рисунке б) приведена эталонная карта глубины, полученная с помощью ручной разметки

С помощью разработанного инструмента были размечены эталонные карты глубины для всех 27 тестовых наборов. На рисунке 2.8 можно видеть пример вручную размеченной эталонной карты глубины.

2.3.2. Метрика для сравнения методов

Для сравнения результатов работы различных методов с эталонной картой глубины предлагается использовать следующую метрику

$$Q(\hat{d}, \tilde{d}) = 100 \frac{\sum_{i=1}^{m} \omega(\hat{d}_i - \tilde{d}_i, T_e)}{m} \omega(x, t) = \begin{cases} \frac{1, |x| > t}{0, |x| \le t} \end{cases}$$
(2.28)

где \hat{d} и \tilde{d} это результат работы метода и эталонная карта глубины, за-

писанные в виде одномерного вектора. Данная метрика была предложена в работе [7] для сравнения и оценки качества работы различных алгоритмов стереореконструкции и представляет собой процент пикселей, ошибка оценки глубины в которых больше некоторого порога. Со времени появления статьи [7] эта метрика стала практически стандартом для сравнения различных алгоритмов стреореконструкции.

В данной работе при оценке использовался порог T_e , который соответствует величине двух глубин резкости микроскопа. Выбор столь высокого порога обусловлен тем, что основной целью экспериментов было оценить устойчивость работы методов. Поэтому порог был выбран таким образом, чтобы он обнаруживал грубые ошибки, а незначительные отклонения поверхности от эталона игнорировал.

2.3.3. Результаты сравнения

Для сравнения на встроенном языке сценариев MATLAB был реализован предложенный алгоритм, алгоритм Наяра [9] и доработанный алгоритм Наира и Стюарта [22]. Алгоритм Наира и Стюарта частично опирается на фотографию сцены с большой глубиной резкости, на которой сцена целиком видна в резкости. Например, на базе фотографии сцены подбирается размер окна для усреднения меры резкости. При сильном увеличении микроскопа получить качественную фотографию сцены с большой глубиной резкости не представляется возможным. Поэтому при реализации использовались только те особенности алгоритма, которые не опираются на фотографию сцены с большой глубиной резкости. Также алгоритм Наира и Стюарта был дополнен процедурой постобработки карты глубины, описанной в [30].

Каждый из этих трех алгоритмов имеет несколько входных парамет-

67

ров. Для предложенного алгоритма это параметры $\{\lambda, T_d, T_s, \mu, R\}$, для алгоритма Наяра это размер окна для усреднения меры резкости, а для доработанного алгоритма Наира и Стюарта это размер окна, порог для отброса ложных замеров и радиусы фильтров для постобработки. Все свободные параметры для алгоритмов в экспериментах были подобраны так, чтобы алгоритмы давали в среднем наилучший результат с точки зрения метрики (2.28). Фактически результаты экспериментов показывают лучшее, на что способны данные алгоритмы, на собранной тестовой базе.

Название метода	Ошибка Q
Найяра	15.6%
Наира и Стюарта	6.4%
Предложенный метод	3.6%

Таблица 2.1. В данной таблице приводятся результаты сравнения предлагаемого метода с существующими аналогами

С помощью каждого из алгоритмов были реконструированы поверхности объектов для всех тестовых наборов из базы. Для каждого тестового набора результаты работы алгоритмов сравнивались с эталоном по метрике (2.28). Далее для каждого из алгоритмов значение метрики усреднялось по всем наборам, для получения средней ошибки алгоритмов. В таблице 2.1 приведены полученные таким образом средние ошибки для трех реализованных алгоритмов. Видно, что предлагаемый алгоритм в среднем дает более стабильные результаты, чем существующие аналоги и практически в два раза превосходит метод Наира и Стюарта по количеству грубых ошибок в результирующей карте глубины.









г)



Ж

к)









м)



л)

2.3.4. Визуальное сравнение результатов

Для визуальной оценки качества работы предложенного алгоритма на рисунке 2.9 приведены результаты работы всех алгоритмов на трех примерах из тестовой базы.

Первый вывод, который можно сделать из результатов, приведенных на рисунке 2.9 это то, базовый алгоритм SFF, предложенный Наяром, дает крайне нестабильные результаты при работе на реальных данных. Видно что в результатах работы алгоритма Наяра присутствуют как крупные области с грубыми ошибками, которые соответствуют областям без текстуры, так и небольшие области с сильными скачками даже в тех областях, где визуально текстура присутствует. Эти наблюдения полностью согласованы с результатами объективного сравнения, приведенными в таблице 2.1.

Результаты работы модифицированного алгоритма Наира и Стюарта выглядят существенно лучше. Видно, что в результатах данного метода практически полностью отсутствуют мелкие, локализованные скачки, которых так много в результатах работы метода Наяра. Однако также видно, что алгоритм Наира и Стюарта дает грубые ошибки в больших областях без текстуры и в местах присутствия крупных бликов.

Результаты предложенного метода визуально незначительно отличаются от эталонных карт глубины. Поверхности, получаемые предложенным методом выглядят гладкими и согласованными. В результатах работы предложенного алгоритма практически полностью отсутствуют сильные скачки, характерные для результатов методов Наяра и Наира/Стюарта, что подтверждает высокую устойчивость предложенного метода.

Для оценки характера ошибок, допускаемых предложенном методом, на рисунке 2.10 приведены примеры разности с эталонной картой глубины для результатов предложенного метода и метода Наира и Стюарта. При-



Рис. 2.10. На рисунке а) приведен пример разности результатов работы предложенного метода с эталонной картой глубины. На рисунке б) приведен пример разности для метода Наира и Стюарта

мер, для которого приведена разность, можно видеть на рисунке 2.9 а). Видно, что предложенный метод допустил на данном существенно меньше ошибок, чем метод Наира и Стюарта. Тем не менее некоторые ошибки в результатах предложенного метода все же присутствуют. Наиболее характерной является ошибка в центральной части карты глубины, которая связана с тем, что предложенной метод слишком сильно регуляризовал поверхность и из-за этого сгладил нижнюю часть углубления по центру карты глубины. Реальную форму данного углубления можно найти на рисунке 2.9 г) а результат работы предложенного метода на рисунке 2.9 н).

На рисунке 2.11 изображены результаты работы алгоритма для еще трех примеров из собранной базы. На данных примерах практически нет крупных нетекстурированных областей, поэтому на этих данных результаты метода Наира и Стюарта не сильно уступают результатам предложенного алгоритма.

71

2.4. Заключение

В данной главе был описан новый метод построения трехмерных моделей семейства SFF, отличающийся повышенной устойчивостью. Ключевым отличием предложенного метода от существующих аналогов является тот факт, что он ищет решение задачи SFF путем глобальной минимизации функции энергии. Предложенный вид функции энергии основан на теории Марковских случайных полей (MRF) и включает в себя как члены, отвечающие за правдоподобие решения, так и члены, отвечающие за регуляризацию.

Использование в предложенном методе регуляризации позволяет методу строить корректные модели объектов даже в случае, если на поверхности объекта присутствуют крупные не текстурированные области и блики. Это подтверждают приведенные в главе эксперименты, который также демонстрируют то, что предложенный метод обладает большей устойчивостью, чем существующие аналоги. Высокая устойчивость предлагаемого метода позволяет существенно расширить класс объектов, для которых применима трехмерная реконструкция по изображениям, полученным с помощью оптического микроскопа.


Рис. 2.11. Примеры результатов работы алгоритмов для визуального сравнения. Первый ряд - мультифокус-изображения, ряды далее это эталонные карты глубины и результаты методов Наяра, Наира/Стюарта и предложенного метода соответственно.

Глава З

Оценка положения и ориентации плоского участка поверхности

Важным частным случаем задачи построения моделей объектов по изображениям с микроскопа является задача оценки положения и ориентации плоского участка поверхности объекта. При работе с объектами антропогенного происхождения в таких приложениях, как например дефектоскопия на производстве, приходится часто сталкиваться с поверхностями, которые являются плоскими. Понятно, что алгоритм предложенный в прошлой главе, непригоден для решения задачи оценки положения и ориентации плоского участка поверхности объекта.

За счет разработки специализированного алгоритма для моделирования плоского участка поверхности объекта можно добиться выполнения сразу двух целей. Известно, что оценку положения плоскости в трехмерном пространстве можно построить, имея в распоряжении лишь три точки. В случае если количество замеров точек плоскости больше трех, оценку плоскости можно сделать более точной и менее чувствительной к шуму. Также, в связи с тем, что для построения оценки плоскости достаточно всего трех точек, специализированный алгоритм можно сделать более устойчивым к ложным замерам, которые так часто возникают при работе с реальными объектами.

3.0.1. Постановка задачи

Сформулируем задачу оценки положения и ориентации плоского участка поверхности объекта по изображениям, полученным с помощью оптического микроскопа. Пусть у нас имеется набор изображений $\hat{I} = \{I_1, \ldots, I_n\}$. Также как и для случая построения модели поверхности общего вида с помощью методов SFF (см. 1.2.3) будем считать, что изображения сделаны с различными положениями области резкости и регулярным шагом между областями резкости соседних изображений.

Также будем считать, что для каждого изображения у имеется преобразование $(X, Y, Z) = B_i(x, y)$, которое переводит точку на изображении i в точку трехмерного пространства, которая получена как пересечение плоскости резкости изображений i и прямой, пущенной в направлении проекции из точки (x, y)¹. Функции B_i позволяют преобразовать замер по глубине в пикселе (x_0, y_0) в точку трехмерного пространства (X_0, Y_0, Z_0) таким образом карта глубины, получаемая с помощью методов семейства SFF может быть трансформирована в множество точек в трехмерном пространстве $P = \{x_i, y_i, z_i\}, i = \overline{1, m}$.

Также будем считать, что наблюдаемый под микроскопом участок объекта представляет собой плоскость. Тогда по входным данным $\hat{I} = \{I_1, \ldots, I_n\}$ требуется построить оценку параметров плоского участка наблюдаемого объекта.

3.1. Существующие методы оценки плоскости

3.1.1. Метод наименьших квадратов

Задача оценки параметров плоскости по набору трехмерных точек $P = \{x_i, y_i, z_i\}, i = \overline{1, m}$, который содержит шум, является известной задачей оценивания параметров модели и существует множество различных

¹ Данная функция задается однозначным образом при условии, что выполнено требование из 1.2.3 о том, что для каждого изображения известно положение плоскостей $L_{1,i}$ и $L_{2,i}$ которые ограничивают области резкости изображения *i*

методов ее решения. Классическим методом решения данной задачи является метод наименьших квадратов, который ищет решение задачи оценки плоскости путем минимизации суммы квадратов расстояний до точек входного набора

$$\hat{\theta} = \arg\min(\sum_{i=1}^{m} d(P_i, \theta)^2)$$
(3.1)

где θ это параметрическое представление плоскости в трехмерном пространстве, а $d(., \theta)$ представляет собой евклидово расстояние от точки в трехмерном пространстве до плоскости θ .

Оценка плоскости, задаваемая (3.1), представляет собой оценку максимального правдоподобия при условии, что шум во входных данных P имеет нормальное распределение. Однако в случае если во входных данных присутствуют элементы, которые не сгенерированы нормальным распределением, оценка (3.1) становится крайне нестабильной. В частности такое происходит если входные данные *загрязнены* ложными замерами. Наличие во входных данных даже одного единственного ложного замера может крайне негативно сказаться на качестве оценки, получаемой с помощью (3.1).

3.1.2. Устойчивые методы оценки плоскости

Для решения данной проблемы был разработан целый ряд подходов, которые могут справиться с загрязнением во входных данных. Методы оценки параметров модели, который являются устойчивыми по отношению к ложным замерам, называются *устойчивыми или робастными методами*. К методам данного класса относятся такие методы как M-estimators [49–51] и методы на основе медианной ошибки (LMedS) [52–54], методы на основе голосования [55–57](преобразование Хафа) и методы на основе случайных выборок [58-62].

3.1.3. M-estimators и медиана квадратов разностей

Ключевым недостатком метода наименьших квадратов, который делает его столь чувствительным к наличию во входных данных ложных замеров, является вид целевой функции для оптимизации в (3.1). В случае если какая-то точка из входного набора P окажется далеко от плоскости θ штраф $d(P_i, \theta)^2$ будет крайне большим в связи с тем, что $d(P_i, \theta)^2$ ничем не ограничена сверху.

Для решения этой проблемы авторы методов семейства M-estimators [49–51] предлагают изменить исходную целевую функцию следующим образом

$$\hat{\theta} = \arg\min(\sum_{i=1}^{m} \rho(d(P_i, \theta)^2))$$
(3.2)

Единственным отличием (3.2) от (3.1) является введение функции $\rho(.)$, которая называется *функцией насыщения*. Задав определенным образом функцию насыщения можно сделать оценку вида (3.2) мало чувствительной к наличию во входных данных загрязнения. Примером функции насыщения является функция следующего вида

$$\rho(x) = \begin{cases} \frac{x, x < T}{T, x \ge T} \end{cases}$$
(3.3)

Видно, что при использовании функции насыщения данного вида насколько далеко точки бы не отстояли от плоскости штраф, накладываемый ими, не может превосходить величины T. Поэтому если предположить, что шум во входных данных имеет гауссовское распределение с дисперсией σ^2 и установить $T = 3\sigma$ оценка, получаемая с помощью (3.2), станет менее чувствительна к ложным замерам. В литературе описано большое количество различных функций насыщения [51]. Однако все известные виды функций насыщения делают прямое решение (3.2) невозможным, и поэтому для поиска решения в методах M-estimators применяется локальная нелинейная оптимизация. В качестве первого приближения при оптимизации обычно используется решение, полученное методом наименьших квадратов. Поэтому зачастую при наличии во входных данных существенной доли ложных замеров (более 30-40%) методы M-estimators не справляются с поиском минимума (3.2) из-за низкого качества начального приближения.

Методы LMedS [52–54] заменяют суммирование в формуле (3.1) на медиану

$$\hat{\theta} = \arg\min(median(d(P_i, \theta)^2, \dots, d(P_m, \theta)^2))$$
(3.4)

При использовании медианы вместо суммирования наличие больших невязок $d(P_i, \theta)^2$ не сильно сказывается на итоговом значении медианы. Несмотря на то, что функция (3.4) является нелинейной, существуют эффективные методы оптимизации (3.4) для моделей θ малых размерностей, таких как, например, плоскость. Методы LMedS могут осуществить успешную оценку параметров модели при условии, что процент ложных замеров во входных данных не превосходит 50% [51].

3.1.4. Методы на основе случайных выборок

Первым методом оценки параметров модели, основанным на случайных выборках был метод RANSAC, предложенный Фишером и Болесом [58]. Алгоритм RANSAC выглядит следующим образом:

1. Из множества входных данных P случайным образом выбираются k различных элементов, где k - минимальное количество элементов, по

которым можно оценить модель θ . В случае плоскости k=3

- 2. По выбранным k элементам строится модель θ_i
- 3. Оценивается степень согласованности модели θ_i со всеми входными данными P. В алгоритме RANSAC для расчета степени согласованности модели со входными данными предлагается использовать количество элементов из P, расстояние от которых до модели не превосходит некоторого заданного порога T, т.е. $d(P_i, \theta) < T$

Описанный выше алгоритм повторяется N раз и из всех гипотез θ_i выбирается та, которая наиболее согласована с входными данными P. Основная идея данного алгоритма заключается в следующем. В случае если среди выбранных случайным образом k элементов из P не присутствует ни одного ложного замера, то оценка θ_i получаемая по таким данным будет верной. И степень согласованности модели, построенной по k надежным замерам, будет превосходить согласованности моделей, которые построены по ложным замерам.

Понятно, что перебрать все возможные варианты по k элементов из P не представляется возможным из соображений эффективности. Поэтому выбор k элементов из P осуществляется случайным образом. Вероятность того, что за N итераций будет выбран хотя бы один набор из k элементов, в котором нет ложных замеров, равна P(N) и возрастает вместе с N. И как показано в [58] можно сделать число итераций N таким, чтобы P(N) было сколь угодно близко к единице ².

Существует множество модификаций базовой схемы [59–62], которые отличаются от нее функцией согласованности гипотез и тем, каким образом выбираются k элементов из P. Однако для всех схем на основе случайных

 $^{^2}$ Это верно при условии, что во входных данных есть не мене
еkнадежных замеров

выборок характерны следующие свойства.

Схемы, основанные на случайных выборках способны корректно работать даже при условии, что загрязнение во входных данных превышает 50%. Однако при достаточно большом проценте ложных замеров (порядка 90%) количество итераций, требуемое алгоритмам данного семейства, становится крайне большим. Также, в случае большого процента ложных замеров, эти замеры могут образовать ложное решение, для которого степень согласованности со входными данными будет выше, чем у реальной модели. Однако стоит заметить, что это является проблемой для любого метода устойчивой оценки параметров модели и решить данную проблему можно лишь путем фильтрации и отброса части ложных замеров перед проведением оценки.

Также минусом методов данного семейства является то, что алгоритмы семейства RANSAC рандомизированны и результаты работы алгоритма от запуска к запуску могут отличаться друг от друга, что неприемлемо для некоторых приложений.

3.1.5. Преобразование Хафа

Методы оценки параметров плоскости на основе преобразования Хафа [55–57] работают по следующему принципу. Пространство возможных значений параметров плоскости θ дискретизуется и создается таблица, количество элементов в которой равно количеству возможных значений параметров модели после дискретизации. Данная таблица называется *аккумулятором*.

На первом шаге алгоритма значения аккумулятора обнуляются. Далее для всех возможных значений параметров плоскости θ для каждой точки из входных данных проверяется выполнение условия $d(P_i, \theta) < T$. В случае если данное условие выполнено, то ячейка аккумулятора, которая соответствует данному значению параметров θ , увеличивается на единицу. После того, как данная процедура проделана для всех возможных значений параметров модели, в аккумуляторе ищется ячейка, которой соответствует наибольшее количество голосов. Значения параметров модели для данной ячейки принимается за результат работы алгоритма.

Основным недостатком схем на основе преобразования Хафа является тот факт, что сложность этих схем растет экспоненциально по мере увеличения количества параметров модели. Поэтому применение таких методов для моделей с большим количеством параметров является невозможным. Для плоскости существуют параметризации из всего лишь трех параметров, поэтому применение методов на основе преобразования Хафа для оценки плоскости является возможным.

Также как и методы на основе случайных выборок, методы на основе преобразования Хафа способны успешно получать оценку даже в случае если число ложных замеров превышает 50%. Однако методы на основе преобразования Хафа не являются рандомизированными, и результат работы алгоритма от запуска к запуску остается неизменным.

3.1.6. Выводы

Из обзора видно, что наиболее устойчивыми методами являются методы на основе случайных выборок и методы на основе преобразования Хафа. Эти методы способны корректно работать при условии что процент ложных замеров во входных данных превышает 50%.

Однако даже эти методы не в состоянии корректно решить задачу построения плоскости при условии, что процент ложных замеров очень велик в связи с тем, что в этом случае ложные замеры могут образовать ложное решение, которое более согласованно со входными данными, чем реальная модель.

3.2. Метод оценки положения и ориентации плоскости

3.2.1. Схема предлагаемого метода

Схема метода, предлагаемого для решения задачи оценки положения и ориентации плоского участка поверхности объекта по изображениям с микроскопа, выглядит следующим образом:

- 1. Получение набора трехмерных замеров. С помощью базовой схемы методов семейства SFF (см. 1.2.3) получить набор трехмерных замеров положения точек плоскости $P = \{x_i, y_i, z_i\}, i = \overline{1, m}$
- 2. **Фильтрация замеров**. В связи с ограничениями на количество ложных замеров, с которыми могут справить устойчивые алгоритмы оценки плоскости, очень важно осуществить отброс как можно большего количества ложных замеров перед процедурой оценки параметров плоскости.
- 3. Оценка положения и ориентации плоскости. По отфильтрованному набору точек в трехмерном пространстве $P_f = \{x_i, y_i, z_i\}, i = \overline{1, n}$ осуществить оценку параметров плоскости $\hat{\theta}$

Новизна предлагаемого метода заключается в новой схеме фильтрации ложных замеров (пункт 2), а также во введении дополнительного шага проверки надежности итоговой оценки плоскости (пункт 4). Для того, чтобы сделать метод оценки положения и ориентации плоского участка поверхности объекта более устойчивым к ложным замерам крайне важно создать процедуру фильтрации, которая работает оптимальным образом. Важно отбросить максимальное число ложных замеров, сохранив при этом как можно больше корректных замеров положения плоскости. Дополнительный шаг проверки надежности полученной оценки плоскости позволяет придать методу еще большую устойчивость. Это достигается за счет того, что метод способен обнаружить ситуации, в которых он не в состоянии сработать корректно.

В последующих подразделах данной главы будут подробно разобраны все четыре пункта предлагаемого алгоритма оценки положения и ориентации плоского участка поверхности объекта по набору изображений, полученных с помощью оптического микроскопа.

3.2.2. Получения множества замеров положения плоскости

В первую очередь предлагаемый метод должен осуществить построение набора трехмерных замеров положения наблюдаемого плоского участка поверхности объекта по набору входных изображений $\hat{I} = \{I_1, \ldots, I_n\}$. Для того чтобы получить набор трехмерных замеров предлагается использовать следующий метод, который основана на базовой схеме SFF (см. 1.2.3).

В соответствии с базовой схемой методов семейства SFF для каждого изображения из набора \hat{I} вычисляется карта меры резкости S_k . В качестве меры резкости предлагается использовать модифицированный оператор Лапласа (1.5).

Для усреднения меры резкости, в отличии от базовой схемы, предлагается использовать следующий метод. Карты меры резкости разбиваются с помощью регулярной сетки на набор ячеек одинакового размера. Далее по каждой ячейке происходит усреднение значений меры резкости с равными весами, то есть

$$\tilde{S}(i,j) = \sum_{k=Ri}^{R(i+1)} \sum_{p=Rj}^{R(j+1)} S(k,p)$$
(3.5)

где R это размер ячейки, по которой производится усреднение. Важным отличием данной схемы усреднения меры резкости от базового варианта является тот факт, что усредненная карта меры резкости \hat{S}_i имеет в Rраз меньше разрешение, чем входные изображения \hat{I}_i . Для входных изображений разрешения один мегапиксель базовая схема усреднения меры резкости выдаст миллион замеров положения точек плоскости. Понятно, что применение методов оценки плоскости к столь большому количеству замеров может быть очень накладным с вычислительной точки зрения. Именно поэтому предлагается использовать схему (3.5), которая дает в R^2 меньше замеров. Также стоит заметить, что расчет (3.5) можно осуществить за один проход по карте меры резкости S_i .

После того, как получены усредненные и уменьшенные карты меры резкости \tilde{S}_i для каждой ячейки (i, j) выбирается изображение k с максимальным значением меры резкости и далее происходит преобразование в трехмерный замер с помощью формулы

$$(X, Y, Z) = B_k(R(i+0.5), R(j+0.5))$$
(3.6)

где B_k это функция которая преобразует точку на изображении в соответствующую трехмерную точку находящуюся на плоскости резкости изображения k (см. 3.0.1).

3.2.3. Фильтрация ложных замеров

На рис 3.1 изображен пример замеров положения точек плоской грани объекта, полученных с помощью описанной схемы. В данных можно заметить присутствие плоской структуры, но также в них присутствует некоторое количество ложных замеров.



Рис. 3.1. Примеры, иллюстрирующие результаты работы первого этапа предложенного алгоритма. На рисунках а) и б) изображен набор точек в трехмерном пространстве, полученный с помощью схемы 3.2.2. На рисунке в) изображен тот же набор точек, в котором синим выделены точки, принадлежащие плоскости

В то время как на приведенном примере количество ложных замеров сравнительно невелико, зачастую количество ошибочных замеров может достигать 90-95%. В этом случае построение плоскости по таким данным становится невозможным даже с использованием устойчивых методов математической статистики. Поэтому крайне важным является построение эффективной схемы фильтрации ложных замеров, для того чтобы минимизировать степень загрязнения входных данных.

В соответствии с анализом недостатков существующих методов SFF, приведенном в 1.3, основными методами борьбы с ложными замерами является отброс ложных замеров по порогу на максимальное значение меры резкости [22], а также применение к карте глубины локальных фильтров, таких как медианный [30].

Методы локальной фильтрации полученной карты глубины [30] для случая большого количества ложных замеров во входных данных не только не способны улучшить ситуацию, а даже могут увеличить количество ошибок во входных данных. В случае большого количества ложных замеров во входных данных корректный замер положения точки плоскости может находиться в окрестности большого количества неверных замеров. Тогда при применении медианного фильтра произойдет замена корректного замера на ложный в связи с тем, что в окрестности корректного замера находится большое количество ложных замеров.

Как было показано в 1.3 отброс ложных замеров по порогу на максимальное значение меры резкости [22] зачастую не в состоянии отбросить все ошибки измерения. Данный метод фильтрации применим в случае, если количество ложных замеров во входных данных невелико. Однако целью разработки нового метода является попытка сделать его устойчивым и работоспособным даже в случае крайне большого количества ложных замеров во входных данных.

3.2.4. Идея предлагаемого метода

Предлагаемая процедура фильтрации развивает и усовершенствует метод отсева по порогу на максимум меры резкости. В отличии от метода отброса по порогу на максимальное значение резкости предлагается использовать для фильтрации ложных замеров форму профиля меры резкости. Понятно, что для случая надежного замера, профиль меры резкости должен иметь один ярко выраженный пик, который соответствует изображению, на котором поверхность видна в резкости. Во всех остальных областях профиль меры резкости должен иметь значения, существенно меньшие максимума.



Рис. 3.2. На рисунке изображены примеры профилей меры резкости. На рисунке а) изображен надежный профиль. На рисунке (б) находится пример ненадежного профиля с большим количеством локальных максимумов, найти по которому наиболее резкое положение надежно не представляется возможным. На в) изображен еще один пример надежного профиля

На рисунке 3.2 проиллюстрировано то, почему за счет использования формы профиля меры резкости можно добиться улучшения качества фильтрации ложных замеров. На данном рисунке приведены примеры трех различных профилей меры резкости, измеренных по изображениям реальных объектов. Первый профиль, который изображен на рисунке 3.2 a) является надежным. Как видно из рисунка он содержит ярко выраженный пик, и низкие значения меры резкости в областях, удаленных от пика. Максимальное значение меры резкости, достигаемое в данном профиле составляет порядка 220.

Профиль, изображенный на рисунке 3.2 б) является примером ненадежного профиля меры резкости. Видно, что форма данного профиля обладает случайной природой. Имеется большое количество локальных максимумов сравнимой величины и выбрать какой-либо из них надежно не представляется возможным. При этом максимальное значение резкости в данном профиле меры резкости составляет порядка 35. В тоже время профиль, изображенный на рисунке 3.2 в) имеет достаточно ярко выраженный основной пик и скорее всего является надежным. Однако, стоит заметить, что максимальное значение меры резкости для данного профиля по своей величине меньше 30. Тогда понятно, что в случае использования фильтрации по порогу на максимальное значение меры резкости для любого значения порога *T* классификация одного из профилей б) и в) будет ошибочной. Однако если удастся создать эффективную процедуру, которая способна учитывать форму профиля меры резкости замеры б) и в) могут быть классифицированы корректно.

3.2.5. Машинное обучение

Таким образом требуется построить правило, которое принимает на вход профиль меры резкости \hat{s} , а на выходе выдает является ли данный профиль меры резкости надежным или нет. В связи с тем, что сформулировать такое правило в явном виде затруднительно, предлагается для этой цели использовать методы машинного обучения с учителем [63–67]. Методы машинного обучения являются крайне популярным инструментом для решения задачи классификации, в которой на входе имеется некоторый вектор наблюдений, и по данному вектору требуется отнести наблюдаемый объект к одному из двух или более классов. Ярким примером успешного применения методов машинного обучения для решения задачи такого вида является задача обнаружения лиц на изображении [68, 69].

Пусть у нас имеется объект, который описывается некоторым численным вектором x. Пусть известно, что данный объект принадлежит к одному из классов $\{1, \ldots, k\}$. Тогда методы классификации с учителем позволяют автоматически построить правило H(x), которое получает на вход вектор, описывающий объект, а на выходе выдает номер класса, к которому данный объект принадлежит. Для автоматического построения правила H(x) методы машинного обучения с учителем используют т.н. обучающую выборку, которая представляет собой набор пар (x_i, c_i) где x_i это значение вектора, который описывает объект, а c_i это верное значение класса, к которому данный объект принадлежит.

3.2.6. Признаки для классификации

Известно, что зачастую вместо того, чтобы использовать в процессе машинного обучения непосредственно вектор, который задает объект x, можно добиться существенно более высокого качества классификации за счет вычисления по x так называемого *вектора признаков* f(x) [69]. В этом случае итоговое правило для классификации будет принимать на вход значение вектора признаков, вместо значения x. В случае задачи классификации профилей меры резкости \hat{s} на надежные и ненадежные требуется вычислить по профилю меры резкости набор признаков, который описывает форму профиля меры резкости.

В качестве признаков для классификации в данной работе предлагается использовать ряд характеристик одномерного сигнала \hat{s} , которые приведены в таблице 3.1. В формулах таблицы вектор $\hat{i}^{max} = \{i_1^{max}, \ldots, i_k^{max}\}$ представляет собой вектор индексов всех локальных максимумов в профиле меры резкости \hat{s} , упорядоченных по величине от большего к меньшему.

Как можно заметить каждый из предлагаемых признаков характеризует форму и поведения одномерного сигнала \hat{s} . Например признак 1, который представляет собой количество локальных максимумов в профиле меры резкости характеризует случайность профиля. Понятно, что чем больше в профиле меры резкости \hat{s} локальных максимумов, тем более вероятно, что данный профиль является ненадежным.

Признак 4 представляет собой расстояние между двумя самыми круп-

89

Номер	Формула	Номер	Формула
1	$len(\hat{i}^{max})$	6	$ i_1^{max} - i_2^{max} /N$
2	$len(\hat{i}^{max})/N$	7	$\hat{s}(i_1^{max})/max(\hat{s})$
3	$\hat{s}(i_1^{max})/\hat{s}(i_2^{max})$	8	$ argmax(\hat{s}) - i_1^{max} $
4	$ i_1^{max} - i_2^{max} $	9	$ argmax(\hat{s}) - i_1^{max} /N$
5	$max(\hat{s})$		

Таблица 3.1. Список признаков, использованных для классификации профилей меры резкости на надежные и ненадежные

ными локальными максимумами в профиле меры резкости \hat{s} . Понятно, что если это расстояние велико, то профиль меры резкости может быть ненадежным т.к. он содержит два сильных локальных максимума на большом расстоянии друг от друга и реальное положение поверхности может соответствовать как первому локальному максимуму, так и второму.

Признак 3 представляет собой отношение значения самого сильного локального максимума в профиле меры резкости к значению второго по величине локального максимума. Понятно, что чем ближе данное отношение к 1, тем более вероятно, что замер, который соответствует данному профилю является ненадежным.

Теперь обозначим функцию, которая вычисляет признаки из таблицы 3.1 по профилю меры резкости как $f = f(\hat{s})$. Тогда итоговый вектор признаков для профиля меры резкости \hat{s} предлагается вычислять следующим образом

$$F(\hat{s}) = \{ f(\hat{s} * G(\sigma_1)), \dots, \hat{s} * G(\sigma_v)) \}$$
(3.7)

где * обозначает свертку, а $G(\sigma_v)$ - гауссовское ядро свертки. В работе использовались четыре значения параметра сигма, а именно 1, 2, 4 и 8. Вычисление признаков $f(\hat{s})$ для различных уровней сглаживания σ дает возможность алгоритму машинного обучения выбрать тот уровень сглаживания, который обеспечивает наибольшую информативность каждого из признаков в таблице 3.1.

3.2.7. Формирование обучающей выборки

Для того, чтобы осуществить построение решающего правила $H(F(\hat{s}))$ по набору признаков, описанных в предыдущей главе, требуется собрать обучающую выборку из различных профилей меры резкости, для каждого из которых известно, соответствует ли он надежному замеру, или нет. Для сбора обучающей выборки использовалось порядка 20 наборов изображений. Каждый из наборов изображений включал в себя от 40 до 100 изображений с различными положениями области резкости. В качестве тестовых объектов использовались частично обработанные природные минералы, на поверхности которых присутствуют плоские грани. На рисунке 3.3 можно видеть изображение одного из тестовых объектов, полученное при слабом увеличении микроскопа.

По набору изображений плоской грани объекта осуществлялось получение набора трехмерных замеров для всех наборов изображений. Далее по полученному множеству точечных замеров осуществлялось построение плоскости с помощью метода RANSAC. В качестве порога *T* для обнаружения ложных замеров использовалась величина, равная половине глубины резкости микроскопа, а количество итераций метода было установлено равным 10⁵.

Далее для каждого набора изображений вручную осуществлялась проверка того, является ли оценка плоскости, полученная с помощью RANSAC, верной. Для всех верных оценок плоскости в обучающую выборку добав-



Рис. 3.3. На рисунке приведено изображение тестового объекта, который использовался для составления обучающей выборки. Видно, что на поверхности объекта присутствует большое количество плоских граней

лялись профили меры резкости, которые соответствуют трехмерным замерам, в соответствии с расстоянием до полученной плоскости. Все замеры, расстояние от которых до плоскости больше, чем половина глубины резкости, были помечены как ложные, а все замеры, расстояние до которых меньше, чем половина глубины резкости, были помечены как надежные.

3.2.8. Обучение классификатора

По составленной таким образом обучающей выборке был построен классификатор с помощью алгоритма Gentle AdaBoost [63]. В качестве слабых классификаторов для алгоритма Gentle AdaBoost в данной работе предлагается использовать порог на один признак из вектора $F(\hat{s})$. При этом итоговое решающее правило, получаемой с помощью алгоритма Gentle AdaBoost имеет следующий вид

$$C(F,t) = \begin{cases} \frac{1, H(F) > t}{0, H(F) \le t} \ H(F) = \sum_{i=1}^{P} \alpha_i \omega_i(F) \end{cases}$$
(3.8)

где *P* - количество итераций алгоритма Gentle AdaBoost, *t* представляет собой порог классификации, а ω_i представляет собой пороговую функцию вида

$$\omega_i(F) = \begin{cases} \frac{a, F(k_i) > t_i}{b, F(k_i) \le t_i} \end{cases}$$
(3.9)

где k_i представляет собой индекс признака, к которому применяется порог, а t_i представляет собой пороговое значение для признака k_i . В случае, если функция C(F,t) возвращает единицу, считается что замер с вектором признаков F является надежным. В противном случае он считается ненадежным.

В обучающей выборке, собранной в соответствии с описанием из раздела 3.2.7, процентная доля ненадежных замеров была существенно больше, чем доля надежных замеров ³. Данная проблема известна в задачах машинного обучения, как проблема *несбалансированных классов*. Известно, что качество работы многих алгоритмов машинного обучения с учителем, включая Gentle AdaBoost, существенно падает при условии того, что классы в обучающей выборке не сбалансированы. Для решения проблемы несбалансированных классов предлагается использовать метод random undersampling [70].

При использовании метода random undersampling происходит прореживание обучающей выборки, и из нее выбрасываются замеры, который принадлежат доминантному классу. Для обучающей выборки, описанной в 3.2.7, это является приемлемым, благодаря тому, что разметка базы про-

³ Количество ложных замеров в собранной обучающей выборке приблизительно в 10 раз больше, чем количество надежных замеров

исходит полуавтоматически и выборка, собранная по 20 наборам изображений, состоит более чем из 5 · 10⁵ примеров. Благодаря столь большому размеру обучающей выборки можно отбросить часть замеров для того, чтобы решить проблему несбалансированных классов, при этом не потеряв в качестве классификации.



Рис. 3.4. На рисунке приведен пример работы предлагаемого алгоритма фильтрации ложных замеров. На рисунке а) можно видеть изначальный набор трехмерных замеров. На рисунке б) изображены замеры после применения процедуры фильтрации

На рисунке 3.4 изображен пример работы предлагаемого алгоритма фильтрации ложных замеров на реальном примере. На рисунке 3.4 а) можно видеть изначальное множество трехмерных замеров без фильтрации. Можно заметить достаточно большое количество ложных замеров среди этого множества. На рисунке 3.4 б) приведено это множество замеров после применения процедуры фильтрации. Красным цветом отмечены ошибки измерений, которые не удалось отбросить с помощью предложенной процедуры, а синим цветом отмечены верные замеры, оставшиеся после фильтрации.

3.2.9. Оценка параметров плоскости

После фильтрации в соответствии с описанной схемой по результирующему множеству замеров можно осуществлять оценку плоскости. В связи с тем, что предложенная локальная процедура фильтрации не в состоянии откинуть абсолютно все ложные замеры, для оценки плоскости предлагается использовать устойчивый метод оценки. Предлагаемый метод оценки плоскости основан на преобразовании Хафа.

Для параметрического представления плоскости используется параметры (ϕ, τ, ρ), где первые два параметра отвечают за нормаль плоскости, а параметр ρ представляет собой расстояние от начала координат до плоскости ⁴. Используемое представление нормали выглядит следующим образом. Для получения нормали, задаваемой параметрами (ϕ, τ) необходимо повернуть вектор $\mathbf{N_0} = (1, 0, 0)$ вокруг начала координат на угол ϕ относительно оси Z, тем самым получив вектор $\mathbf{N_1}$. После этого вектор $\mathbf{N_1}$ поворачивается на угол τ вокруг начала координат относительно оси, которая перпендикулярна Z и вектору $\mathbf{N_1}$. В итоге получается финальный вектор нормали плоскости $\mathbf{N_f}$.

Предлагаемый метод основан на последовательном применении нескольких преобразований Хафа. На каждой итерации осуществляется дискретизация параметров (ϕ, τ, ρ) с регулярным шагом. При этом на каждой новой итерации шаги дискретизации параметров становятся меньше. Также на каждом шаге осуществляется уменьшение окрестности, в котором осуществляется перебор. Результат работы каждой итерации используется как начальное приближение для последующей итерации.

Алгоритм выглядит следующим образом:

1. В качестве начального приближения зададим $\phi_0 = \pi$, $\tau_0 = 0$ и ρ_0 рав-

⁴ Расстояние откладывается в направлении нормали плоскости и является знаковым параметром

ное среднему расстоянию от начала координат до множества точек входного набора

- 2. В качестве начальной величины окрестности поиска выбираются π для параметра ϕ , $\pi/2$ для параметра τ , и максимум модуля разницы между ρ_0 и расстоянием от начала координат до точек входного набора для параметра ρ
- 3. Применяется поиск плоскости с помощью преобразования Хафа с окрестностью поиска, задаваемой начальным приближением и величиной окрестности. Шаги по углам и расстоянию при этом равны s_a и s_d. Обозначим результат работы преобразования Хафа на текущей итерации как (\$\phi_h\$, \$\tau_h\$, \$\rho_h\$)
- После этого величины окрестности поиска и шага уменьшаются в α раз и процедура поиска решения повторяется с начальным приближением, равным (φ_h, τ_h, ρ_h)
- 5. После того, как шаги по углу и расстоянию достигли пороговых значений t_a и t_d, поиск прекращается и результата работы предыдущей итерации считается итоговой оценкой плоскости

Использование адаптивной схемы с последовательным уменьшением шага дискретизации позволяет сократить количество значений, которые требуется перебрать, при этом сохраняя возможность осуществить построение итоговой оценки с достаточно мелким шагом дискретизации. В данной работе в качестве коэффициента уменьшения размера окрестности поиска использовалось $\alpha = 4$. Параметры t_a и t_d являются специфическими для конкретного приложения и отражают уровень точности, который необходимо получить ⁵.

⁵ Для приложения в которое непосредственно внедрялся предлагаемый метод использовался по-

3.3. Проверка надежности оценки плоскости

В связи с тем, что предлагаемая локальная процедура фильтрации трехмерных замеров не в состоянии отбросить все ложные замеры полностью, в задаче оценки положения и ориентации плоского участка поверхности объекта по набору изображений с микроскопа возникает еще одна важная подзадача.

Допустим, что наблюдаемый на изображениях плоский участок поверхности объекта практически полностью засвечен, или например на его поверхности полностью отсутствует видимая текстура. В этом случае после построения набора трехмерных замеров получится облако точек, которое практически полностью состоит из случайных ошибок измерения. После применения локальной процедуры фильтрации существенная часть ошибок будет отброшена, однако некоторая часть ложных замеров будет пропущена. По данному отфильтрованному множеству, состоящему практически полностью из ложных замеров, будет произведена оценка плоскости и результат оценки будет ложным в связи с тем, что все, или почти все, замеры, использованные при построении, были ложными.

Таким образом для обеспечения надежности результатов работы метода для каждого входного набора данных требуется определить возможно ли построить по таким входным данным плоскость корректно. Для решения данной проблемы предлагается метод, основанный на теории статистической проверки гипотез.

Пусть у нас есть набор точек $P = \{x_i, y_i, z_i\}, i = \overline{1, m}$ в трехмерном пространстве. Будем считать, что все точки из этого набора представляют собой случайные ошибки измерения т.е. что эти точки являются выбросами. Традиционно для моделирования поведения выбросов используется $\overline{p_{\text{ог}} t_a = 0.01^{\circ}}$ и порог $t_d = 2\mu m$

равномерное распределение [59], таким образом $p_i \sim U(\Omega)$, где Ω это область пространства, в которой расположены точки P. В качестве Ω можно взять например сферу, ограничивающую все точки из P.



Рис. 3.5. Схематическая иллюстрация областей Ω и Ω₀

Теперь зафиксируем в пространстве плоскость θ , которая проходит через область Ω и обозначим через Ω_0 множество точек Ω , которые лежат от зафиксированной плоскости на расстоянии не большем некоторого заданного порога

$$\Omega_0 = \{ p | p \in \Omega, d(p, \theta) < t \}$$
(3.10)

где t - это величина порога. Тогда понятно, что каждая конкретная точка из P принадлежит области Ω_0 с вероятностью $P(p_i \in \Omega_0) = |\Omega_0|/|\Omega|$ и таким образом количество ложных замеров, попавших в окрестность плоскости, подчиняется биномиальному закону $c \sim B(|\Omega_0|/|\Omega|, m)$.

Тогда если использовать в качестве основной гипотезы H_0 тот факт,

что точки из P распределены равномерно, а в качестве критической статистики использовать c, то можно построить тест, проверяющий надежность плоскости, следующим образом. Найдем количество точек c_0 такое, что вероятность того, что c_0 точек или более принадлежат окрестности зафиксированной плоскости при условии H_0 равно β и положим β достаточно маленьким, например равным 0.01. Тогда если количество точек, которые попали в окрестность плоскости, превосходит c_0 то можно утверждать, что с большой вероятностью гипотеза H_0 неверна и наблюдаемые точки имеют неравномерное распределение.

Пусть у нас имеется набор трехмерных точек P после применения фильтрации и оценка положения плоскости θ , построенная по этим точкам. Тогда предложенная идея реализуется в виде следующего алгоритма оценки надежности плоскости:

- 1. Рассчитаем по набору точек P выпуклую оболочку и будем считать, что она представляет собой область Ω
- 2. Для расчета Ω_0 положим в формуле (3.10) порог t равным половине глубины резкости
- Вычислим количество точек из набора P, которые находятся от плоскости θ на расстоянии меньшем, чем половина глубины резкости. Обозначим это количество точек как c_p
- 4. Вычислим со в соответствии с выкладками, описанными выше
- Тогда в случае если c_p > c₀ будем считать, что оценка плоскости θ является надежной, и ненадежной в противном случае

3.4. Эксперименты

Для оценки качества работы и верификации предложенного метода он был реализован на языке программирования MATLAB. Для верификации предложенного метода было произведено три эксперимента. В первом эксперименте оценивалось качество фильтрации ложных замеров. Во втором и третьем экспериментах оценивались точности оценки плоскости, которую позволяет получить предложенный алгоритм, а также качество работы предложенной процедуры проверки надежности плоскости.

Для экспериментов использовался микроскоп *Leica Z6 APOA*, с установленной на него камерой разрешения 1280х1024. Все изображения, полученные для экспериментов, были полутоновыми. Сканирование объектов производилось с максимальным увеличением, глубина резкости для которого на данном микроскопе составляет 15 микрон. В качестве тестовых объектов использовались частично обработанные природные минералы, на поверхности которых присутствуют плоские грани ⁶.

3.4.1. Оценка качества алгоритма фильтрации ложных замеров

Для оценки качества предложенной процедуры фильтрации на обучающей выборке, описанной в разделе 3.2.7, была проведена оценка качества работы двух классификаторов. Первый классификатор использовал все признаки, описанные в разделе 3.2.6, а второй классификатор использовал один признак - максимальное значение резкости в профиле меры резкости.

Для оценки качества обоих классификаторов на обучающей выборке для них была построена ROC-кривая [71] методом скользящего контроля. При этом обучающая выборка разбивалась случайным образом на две ча-

⁶ Пример изображения такого тестового объекта был приведен на рисунке 3.3

сти - на первой проводилось обучение классификаторов, а на второй оценка ошибок классификации. Данная процедура была произведена 10 раз и результаты для каждого классификатора были усреднены.



Рис. 3.6. ROC-кривый для предложенного метода (синяя сплошная линия), и метода фильтрации на основе максимального значения резкости в профиле (красная пунктирная линия)

На рисунке 3.6 приведены изображения ROC-кривых для обоих классификаторов. По оси X на рисунке отложен процент ложных срабатываний классификаторов, а по оси Y отложен процент верных срабатываний. В данном случае ложные срабатывания это ложные замеры, который были классифицированные как надежные, а верные срабатывания это надежные замеры, которые были верно классифицированы.

Как можно заметить при одинаковом уровне ложных срабатываний предлагаемый метод позволяет обеспечить больший процент верных срабатываний, чем метод, основанный на максимальном значении меры резкости.

3.4.2. Точность оценки плоскости

Во втором эксперименте оценивалась реальная точность построения плоскости предложенным методом. Для оценки точности сканировалась одна и та же грань минерала с разных направлений и с разными параметрами освещения. Если грани были достаточно большими также для чистоты эксперимента сканировались различные участки грани. По каждому из полученных наборов оценивалось положение и ориентация плоскости грани.

N	α_{avg} (градусы)	α_{max} (градусы)	V_0	V_{f}
8	0.12	0.24	85%	12%
8	0.16	0.22	76%	3%
7	0.20	0.31	87%	10%
8	0.15	0.22	89%	2%

Таблица 3.2. Иллюстрирует статистику точности работы методов построения плоскости и фильтрации замеров. Каждая строка таблицы соответствует одной грани, которая сканировалась с разных направлений и при разных условиях освещенности. Расшифровка обозначений: N - количество наборов на грань, α_{avg} - средняя ошибка по углу, α_{max} - максимальная ошибка по углу, V_0 - процент ложных замеров до фильтрации, V_f - процент ложных замеров после фильтрации.

Далее результаты построения по всем наборам усреднялись и считался разброс оценок плоскости относительно этого среднего. В частности для каждой грани были вычислены средний разброс относительно среднего значения и максимальное отклонение от среднего значения.

Результаты данного эксперимента для нескольких граней приведены в таблице 3.2. В таблице приведен разброс оценок по углу нормали, так как это наиболее важный показатель точности построения плоскости. Также в таблице приведена статистика среднего количества выбросов в замерах до и после фильтрации. Как можно заметить, при использовании оптической системы с глубиной резкости 15 микрон, предлагаемый метод позволяет осуществлять оценку плоскости со средней точностью порядка 0.16 градуса по углу нормали плоскости.

3.4.3. Качество процедуры проверки надежности плоскости

В третьем эксперименте оценивалось качество работы предложенного теста для определения надежности плоскости. Было снято несколько наборов фотографий, на которых грани минерала достаточно грязные и детали хотя бы на части их поверхности различимы под микроскопом. Также было снято некоторое количество наборов фотографий, на которых грань практически не видна. К каждому из полученных наборов фотографий был применен описанный алгоритм построения плоскости и для каждого набора вручную было установлено, удалось ли построить плоскость корректно. После ручной разметки к каждому из наборов был применен алгоритм автоматического определения надежности оценки плоскости.

Ручная классификация	Общее количество	Ошибки
Оценка корректна	36	3
Оценка некорректна	29	0

Таблица 3.3. Во второй колонке приведено общее количество надежно/ненадежно построенных граней, участвовавших в эксперименте. В третей приведено количество ошибок автоматического определения надежности

Статистика, демонстрирующая согласованность результатов ручной разметки и результатов работы теста на надежность приведена в таблице 3.3. Как можно заметить, предложенный тест успешно обнаруживает ситуации, когда грани на изображениях видны слишком плохо, но при этом иногда он отбрасывает грани, по которым можно построить плоскость надежно.

3.5. Заключение

В данной главе был описан новый метод оценки положения и ориентации плоского участка поверхности объекта по набору изображений с микроскопа. Был предложен новый метод отсева ложных замеров, который позволяет существенно повысить качество фильтрации по сравнению с существующими аналогами.

Использование предлагаемого метода фильтрации ложных замеров позволяет осуществлять оценку положения и ориентации плоскости даже в случаях, когда процент ложных замеров во входных данных велик. Это позволяет существенно расширить класс объектов, к которым применим метод оценки положения и ориентации плоского участка поверхности по изображениям с микроскопа.

Предложенная схема проверки надежности плоскости позволяет обнаруживать ситуации, когда корректная оценка плоскости по входным данным невозможна. Данная проверка крайне важна для практических приложений, когда получение абсолютно неверной оценки плоскости существенно хуже, чем явное сообщение о том, что построение оценки плоскости по входным данным является невозможным.

Глава 4

Программная реализация предложенных методов

Разработанные в рамках диссертационной работы методы построения трехмерной модели участка поверхности, представимого в виде карты глубины, и оценки положения плоской грани на поверхности объекта, были реализованы в рамках программной подсистемы InclEngine на языке C++. Реализованная подсистема была внедрена в программный продукт Oxygen Microscope Server компании-заказчика Octonus Software Ltd и используется в реальных практических приложениях в процессе обработки алмазного сырья. Ключевыми функциями разработанной подсистемы являются:

- 1. Уточнение трехмерной модели поверхности алмазного сырья
 - а. Моделирование мелких каверн на поверхности алмазов
 - б. Уточнение положения плоских граней на поверхности частично обработанных алмазов для осуществления точного расчет преломлений внутри камня
- 2. Полуавтоматическое моделирование неоднородностей и дефектов внутри алмазного сырья

4.1. Алмазообрабатывающая промышленность

После прибытия на алмазообрабатывающую фабрику алмазное сырье проходит несколько стадий обработки. В процессе обработки строится трехмерная модель поверхности алмаза, моделируются неоднородности внутри камня планируется оптимальный распил и только после этого осуществляется огранка камня. В данном разделе будут кратко описаны все стадии процесса обработки алмазного сырья.

4.1.1. Открытие окон

Алмазное сырье, в том виде в котором оно присутствует в природе и добывается, зачастую имеет непрозрачную поверхность. В связи с тем, что в процессе формирования алмаза внутри него могут оказаться инородные элементы, для определения оптимального способа огранки алмаза требуется иметь возможность заглянуть внутрь камня и оценить положение, размер и форму неоднородностей.



Рис. 4.1. Изображение алмаза, на поверхности которого открыты окна, с двух различных ракурсов. Можно заметить, что через окна на поверхности камня видны неоднородности внутри алмаза

Для того, чтобы сделать это, после прихода на фабрику алмазное сырье подвергается первичной обработке. В процессе первичной обработки некоторые участки поверхности камня шлифуются. При этом, зачастую, шлифовка осуществляется таким образом, чтобы минимально уменьшить массу камня. По результатам шлифовки на поверхности камня образуется набор плоских граней, которые называются окнами. Данные грани прозрачны и через них имеется возможность анализировать внутренность алмазного сырья. Данная стадия обработки называется *открытием окон*. На рисунке 4.1 можно видеть пример камня после процедуры открытия окон на его поверхности.

4.1.2. Построение трехмерной модели камня

После открытия окон камень приклеивается на специальную подставку, называемую холдер (holder) (см. рис 4.2). Жесткая фиксация на холдере упрощает дальнейшую обработку камня. Благодаря своей конструкции холдер может быть установлен в устройство для построения трехмерной модели камня и под микроскоп единственным образом. Поэтому, после построения трехмерной модели камня, он может быть перенесен под микрсокоп без потери регистрации трехмерной модели относительно фотографии.



Рис. 4.2. Изображение алмаза с открытыми окнами, жестко зафиксированного на подставке - холдере.

После того, как камень приклеен на холдер, он устанавливается в специальное устройство, которое осуществляет построение трехмерной модели камня. В подавляющем большинстве случаев для построения трехмерной модели камня используется метод Visual Hull [72]. При использовании этого метода камень вращается на подставке и делаются его снимки с различных направлений. После этого по каждому снимку вычисляется форма силуэта камня и формы силуэтов пересекаются в трехмерном пространстве для получения трехмерной модели поверхности камня.

Основным преимуществом данного подхода является его скорость один камень может быть обработан менее чем за 10 секунд. Однако недостатком данного подходя является то, что метод Visual Hull не позволяет корректно моделировать каверны и невыпуклости на поверхности алмаза. Для камней сложной формы также используются сканеры, оснащенные лазером, где модель камня может быть уточнена с помощью метода лазерного сканирования. Однако в этом случае на время сканирования камень необходимо покрывать матовой краской. При этом узкие и глубокие каверны на поверхности будут заполнены краской, и даже с помощью лазерного сканирования не удастся получить корректную модель камня. На рисунке 4.3 можно видеть пример установки, которая используется для построения трехмерной модели алмаза.

4.1.3. Моделирование дефектов внутри камня и уточнение модели

После того, как построена трехмерная модель, камень переносится под микроскоп для дальнейшего анализа и уточнения трехмерной модели. На рисунке 4.4 можно видеть изображение микроскопа, который используется для анализа камня. Благодаря использованию холдера после переноса

108


Рис. 4.3. Изображение установки с помощью которой осуществляется сканирование алмаза и построение его трехмерной модели.

под микроскоп модель камня, построенная на сканере остается зарегистрированной относительно фотографий, получаемых с помощью микроскопа. Холдер устанавливается под микроскопом на специальной подвижке, которая позволяет перемещать и вращать камень. Тем самым достигается возможность изучать камень с произвольного ракурса. За счет того, что на подвижку установлена специальная считывающая аппаратура, при вращении и перемещении камня под микроскопом модель камня остается зарегистрированной относительно фотоизображений.

Основными задачами, которые решаются на данном этапе являются:

- 1. Полуавтоматическое моделирование неоднородностей и дефектов внутри алмаза. Такие неоднородности называются *включениями*. Оператор анализирует внутренность камня через окна на его поверхности, ищет включения и осуществляет их разметку.
- 2. Уточнение трехмерной модели поверхности алмаза в областях, про-



Рис. 4.4. Изображение микроскопа, с помощью которого осуществляется анализ неоднородностей внутри алмазного сырья и уточнение трехмерной модели поверхности камня.

блемных для построения с помощью сканера.

Оптические микроскопы на максимальном увеличении могут иметь очень маленькую глубину резкости, вплоть до 5µm. Поэтому за счет использования методов, разработанных в рамках диссертационной работы можно решить следующие задачи:

- 1. Высокоточное построение невыпуклых участков поверхности алмазов с помощью предложенного в работе метода построения трехмерной модели участка поверхности, представимого в виде карты глубины.
- 2. Повышение точности трехмерных моделей включений в алмазе за счет использования предложенного метода оценки положения и ориентации плоского участка поверхности объекта. В связи с тем, что при моделировании включений в алмазе необходимо рассчитывать преломления внутри камня, точность построения трехмерных моде-

лей включений напрямую зависти от точности оценки положения и ориентации окон на поверхности камня.

4.1.4. Планирование огранки камня

После того, как произведено уточнение трехмерной модели поверхности алмаза, и построены трехмерные модели включений внутри камня, задача оптимальной огранки камня может быть сформулирована как задача нелинейной минимизации. В результате решения этой задачи получается план распила и огранки камня, в соответствии с которым он гранится.

4.2. Интерфейс и функциональность системы Oxygen Microscope Server

Программный продукт Oxygen Microscope Server предназначен для работы пользователя с алмазом, установленным под микроскоп. Программный продукт Oxygen Microscope Server имеет развитый пользовательский интерфейс (см. рис 4.5) и предоставляет оператору широкий спектр возможности для работы с алмазным сырьем.

У программного продукта Oxygen Microscope Server можно выделить три основных режима работы:

- 1. **Режим настройки микроскопа.** Режим настройки микроскопа позволяет осуществлять калибровку параметров микроскопа.
 - а. Калибровка параметров объектива и камеры с помощью автоматического анализа специального двумерного калибровочного шаблона, который устанавливается под микроскоп на время калибровки

- б. Выравнивание системы координат, связанной с подвижкой микроскопа, с системой координат сканера, за счет использования специального калибровочного объекта в форме куба с метками
- 2. Режим моделирования включений и уточнения модели камня. Основной режим работы продукта Oxygen Microscope Server в котором доступны следующие возможности:
 - а. Панель моделирования включений, с помощью которой можно создавать различные типы включений в камне, а также осуществлять разметку этих включений на изображениях. Также присутствует возможность просматривать репроекцию смоделированных включений на изображения для анализа точности полученных моделей включений (см. рис 4.5, б).
 - б. Панель работы с поверхностью камня, которая позволяет осуществлять сканирование каверн на поверхности камня, а также уточнение положения и ориентации окон на поверхности камня.
 - в. Возможность просмотра трехмерной модели камня со всеми включениями, которые были построены внутри него (см. рис 4.5, а), а также с учетом всех смоделированных на поверхности камня каверн.
- 3. Режим поиска оптимального плана распила и огранки камня. В данном режиме работы пользователю доступны различные алгоритмы поиска оптимального распила камня, возможность оценки стоимости каждого из предлагаемых программой решений, а также возможность просмотра и анализа предлагаемых вариантов огранки камня в 3D.



Рис. 4.5. На рисунке приведены снимки экрана программы Oxygen Microscope Server. На рисунке а) приведено изображения камня и включений размеченных в нем. На рисунке б) приведено изображение камня, выровненное относительно трехмерной модели, построенной на сканере

4.3. Детали реализации подсистемы InclEngine

Разработанные в рамках диссертации методы были реализованы как часть подсистемы InclEngine. Помимо предложенных методов построения участка поверхности, представимого в виде карты глубины, и оценки положения и ориентации плоского участка поверхности объекта, разработанная система включает в себя функциональность, предназначенную для моделирования включений различных типов внутри алмазного сырья. Суммарный объем кода подсистемы составляет более 200000 строк на языке C++.

В рамках подсистемы InclEngine был программно реализован предлагаемый в диссертационной работе метод построения трехмерной модели участка поверхности, представимого в виде карты глубины. Данный метод используется в продукте Oxygen Microscope Server для уточнения модели поверхности алмаза в областях, где присутствуют каверны, которые невозможно построить с помощью сканера. Также в рамках подсистемы InclEngine был программно реализован предлагаемый метод оценки положения и ориентации плоской грани на поверхности объекта. Данный метод используется в продукте Oxygen Microscope Server для высокоточного оценивания положения окон на поверхности алмазного сырья.

4.3.1. Взаимодействие реализованной подсистемы с Oxygen Microscope Server

Подсистема InclEngine реализована в виде динамически подключаемую библиотеки (DLL). Доступ подсистемы к ресурсам Oxygen Microscope Server, таким как камера, микроскоп, модель камня и список включений в камне, осуществляется за счет интерфейсов автоматизации, предоставляемых Oxygen Microscope Server.

На рисунке 4.6 можно видеть UML-диаграмму компонент, которая иллюстрирует то, каким образом разработанная подсистема интегрирована в программный продукт Oxygen Microscope Server. Пунктирные линии на диаграмме иллюстрируют зависимости компонент между собой.

- Компонент **OxygenAutomationSDK** включает в себя набор интерфейсов автоматизации, предоставляемых программным продуктом Oxygen Microscope Server. С помощью предоставляемого набора интерфейсов автоматизации внешние модули могут взаимодействовать с ресурсами Oxygen Microscope Server, в частности управлять камерой, фокусом и увеличением микроскопа, взаимодействовать с внутренними структурами такими как трехмерная модель камня и т.д.
- Компонент ModelingPanel включает в себя реализацию графического интерфейса, с помощью которой пользователь программного продукта Oxygen Microscope Server может использовать программные возможности реализованной подсистемы InclEngine. В компоненте ModelingPanel присутствуют реализация панели для работы со включениями. Также в данной компоненте присутствуют реализации



Рис. 4.6. UML-диаграмма компонент, иллюстрирующая то, каким образом реализованная подсистема InclEngine интегрирована в продукт Oxygen Microscope Server и то, каким образом данная подсистема взаимодействует с ресурсами Oxygen Microscope Server

интерфейса для осуществления построения кавери на поверхности камня и для уточнения положения окон на поверхности камня.

- Компонент InclVisualization предоставляет функциональность для визуализации и ввода данных, которые необходимо задать пользователю при полуавтоматическом моделировании включений. Также данный компонент отвечает за визуализацию репроекции трехмерных моделей включений, которая требуется для анализа точности моделирования включений.
- Компонент InclEngine, предоставляет возможности для полуавтоматического моделирования включений внутри камня, а также для осуществления уточнения трехмерной модели поверхности алмаза с по-

мощью методов, предложенных в рамках диссертационной работы. Подсистема InclEngine интенсивно использует интерфейсы автоматизации, представленные в компоненте OxygenAutomationSDK, для решения возложенных на данную подсистему задач. Как можно заметить из диаграммы 4.6 реализованный компонент полностью независим от графического интерфейса в ModelingPanel и InclVisualization, что дает возможность модифицировать и дорабатывать графический интерфейс без внесения изменений в код подсистемы InclEngine.

4.3.2. Архитектура подсистемы: уточнение модели поверхности камня

На рисунке 4.7 приведена диаграмма классов, которая иллюстрирует организацию внешних интерфейсов компоненты InclEngine, которые позволяют осуществлять уточнение трехмерной модели камня с помощью методов, предложенных в диссертационной работе.

Интерфейс IInclEngineManager представляет собой главный класс подсистемы InclEngine, с помощью которого можно получить доступ ко всем другим интерфейсам, предоставляемым подсистемой InclEngine. В частности с помощью методов CreateCavityScanner и CreateWinodowCorrector можно получить доступ к классам, которые позволяют осуществлять уточнение трехмерной модели.

Метод оценки положения и ориентации окон на поверхности камня и метод построения моделей каверн на поверхности камня представляют собой две различных реализации интерфейса IModelCorrector. При вызове метода PerformScan выполняется сканирование участка поверхности камня и осуществляется уточнение модели камня методом, который определяется тем, какая реализация интерфейса используется. Классы-реализации



Рис. 4.7. На рисунке изображена UML-диаграмма классов, которая иллюстрирует устройство внешнего интерфейса, который предоставляет подсистема InclEngine для осуществления уточнения трехмерной модели камня. Зеленая черта представляет собой границу между подсистемой InclEngine и интерфейсами автоматизации из компоненты ОхуgenAutomationSDK.

CCavityScanner и CWindowCorrector осуществляют уточнение трехмерной модели камня путем активного взаимодействия с интерфейсами автоматизации Oxygen Microscope Server, как показано на рисунке 4.7.

На рисунке 4.8 приведена UML-диаграмма классов, которая иллюстрирует внутреннюю организацию метода построения трехмерной модели участка поверхности, представимого в виде карты глубины. Архитектурное решение, использованное при реализации метода оценки положения и ориентации плоского участка поверхности объекта по изображениям с микроскопа выглядит аналогичным образом.

Класс CCavityScanner, который представляет собой главный класс,



Рис. 4.8. На рисунке изображена UML-диаграмма классов, которая иллюстрирует внутреннее устройство реализации метода построения трехмерной модели участка поверхности, представимого в виде карты глубины. Зеленая черта представляет собой границу между подсистемой InclEngine и интерфейсами автоматизации из компоненты OxygenAutomationSDK.

осуществляющий построение трехмерной модели участка поверхности объекта, является связующим звеном между тремя классами. Класс CScanningManage осуществляет работу с интерфейсом микроскопа IHardwareAutomation и отвечает за определение пределов сканирования по фокусу, шага сканирования и за организацию процесса сканирования. С помощью метода MakeNextShot осуществляется получения нового кадра в процессе сканирования.

Класс CModelIntersector отвечает за преобразование карты глубины поверхности в трехмерный многогранник, а также за коррекцию трехмерной модели камня в соответствии с результатами сканирования участков поверхности. Работа с моделью камня осуществляется путем взаимодействия с интерфейсом IInclusionAutomation. Класс CMRFShapeFromFocus заключает в себе реализацию предложенного в диссертационной работе метода. Класс CMRFShapeFromFocus использует класс CLaplacian, с помощью которого осуществляется расчет меры резкости по методу модифицированный оператор Лапласа (см. формулу 1.5)¹.

Можно заметить, что разбиение реализации на три класса позволяет сделать класс CMRFShapeFromFocus, в котором содержится внутренняя логика предлагаемого метода, полностью независимым от интерфейсов автоматизации Oxygen Microscope Server. Благодаря этому реализацию предлагаемого метода можно легко вычленить из системы и использовать в других программных продуктах, с другими типами камер и другими моделями микроскопов.

4.3.3. Архитектура подсистемы: моделирование включений

На рисунке 4.9 приведена UML-диаграмма классов, которая иллюстрирует архитектуру системы моделирования включений, предоставляемую реализованной подсистемой InclEngine. Класс IInclusionsManager позволяет создавать и открывать на редактирование включения различных типов. Для каждого типа включений, который требуется моделировать, подсистема InclEngine предоставляет отдельный интерфейс.

Все типы включений, поддерживаемый в подсистеме InclEngine, являются наследниками от интерфейса IInclusion, который предоставляет базовую для всех включений функциональность, такую как получение и запись трехмерной модели включения, именование включения и т.д. Конкретные интерфейсы включений, такие как например IPointIncl, I3DInclusion и IFlatCrack предоставляют специализированные функции для моделирова-

¹ Класс CLaplacian также используется в реализации метода оценки положения и ориентации плоского участка поверхности объекта



Рис. 4.9. На рисунке приведена UML-диаграмма классов, которая иллюстрирует организацию иерархии включений, предоставляемую реализованной системой InclEngine. Каждый интерфейс включения позволяет осуществлять моделирование конкретного типа включения, осуществлять обработку пользовательских данных и построение трехмерной модели включения в камне. Для удобства на диаграмме не приводится большая часть методов, предоставляемых интерфейсами включений.

ния конкретного типа включений.

4.3.4. Замеры производительности разработанных алгоритмов

Для программных реализаций алгоритмов, разработанных в рамках диссертационной работы, были произведены замеры производительности. Замеры проводились на компьютере с процессором Intel Core2 Quad 2.5Ghz (4 ядра) с объемом оперативной памяти 3.25RAM. Стоит добавить что при реализации алгоритмов, разработанных в рамках данной работы, использовалась библиотека Intel Integrated Performance Primitives (IPP) [73].

Данная библиотека предоставляет большое количество функций для

фильтрации изображений и обработки массивов данных. При этом реализации функций в библиотеке IPP используют процессорно-зависимую оптимизацию (такую как использование ОКМД инструкций SSE, SSE2 и т.д.). Также большинство функций библиотеки IPP распараллеленны и позволяют при выполнении задействовать все ядра процессора Core2 Quad.

При произведении замеров не учитывалось время, затрачиваемое на получение кадров с помощью камеры, а также время, требуемое на изменение фокусировки микроскопа. Это связано с тем, что данное время существенно зависит от конкретного типа микроскопа и камеры, которые используются при сканировании.

Метод	Среднее время
Карта глубины	8.4 сек
Плоскость	2.3 сек

Таблица 4.1. Результаты замеров производительности реализованных методов. Во второй строчке приводятся результаты замеров производительности метода построения модели участка поверхности, представимого в виде карты глубины. В третей строке таблицы приводятся результаты замеров производительности для метода оценки положения и ориентации плоского участка поверхности объекта.

Замеры производительность осуществлялись на наборах из 50 изображений, разрешение которых равнялось 1280х1024. При осуществлении замеров производительности предложенного метода построения трехмерной модели участка поверхности, представимого в виде карты глубины, использовалось разрешение карты глубины равное 125х100.²

В таблице 4.1 приводятся результаты описанных замеров производительности. Для каждого из методов для оценки производительности ис-

² Это было достигнуто уменьшением разрешения карт меры резкости в 10 раз с помощью линейного метода изменения разрешения изображений.

пользовалось 3 набора входных данных. Для каждого набора входных данных метод запускался 10 раз, после чего все 30 временных замеров усреднялись.

4.4. Заключение

В данной главе были описаны детали программной реализации методов, предложенных в диссертационной работе. Было приведено описание технического процесса обработки алмазного сырья, в рамках которого используются реализации предложенных методов. Также были приведены детали архитектуры реализованной в рамках работы подсистемы InclEngine и замеры производительности реализованных методов.

Программный продукт Oxygen Microscope Server, в который была интегрирована реализованная в рамках работы подсистема InclEngine, успешно внедрен в реальное производство. Продукт Oxygen Microscope Server:

- Используется на алмазообрабатывающем производстве более чем двадцати ограночных заводов в таких центрах алмазообрабатывающей промышленности, как г. Сурат (Индия) и г. Антверпен (Бельгия)
- Используется в процессе подготовки и обучения специалистов по огранке алмазов в Геммологическом Центре Московского Государственного Университета им. М.В. Ломоносова

Стоит отметить что объекты, с которыми работает алмазообрабатывающая промышленность, являются сложными для методов построения трехмерных моделей по изображениям с микроскопа. Алмазное сырье может быть частично прозрачным, а также на поверхности алмазов часто может содержаться множество бликов. Поэтому успешное внедрение предложенных методов в реальное производство еще раз подтверждает их высокую устойчивость.

Заключение

Основные результаты работы состоят в следующем:

- На основе проведенного исследования были разработаны и программно реализованы новые устойчивые алгоритмы моделирования участков, представимых в виде карты глубины, и плоских участков поверхности объекта, позволяющие существенно расширить класс объектов, к которым применимы методы построения трехмерных моделей по изображениям с микроскопа.
- 2. Произведен анализ разработанных методов и алгоритмов на реальных данных, который показывает, что предложенные методы позволяют строить трехмерные модели объектов точнее, чем существующие аналоги.
- 3. На базе разработанных алгоритмов реализован программный модуль для построения формы поверхности микрообъектов, позволяющий ускорить и автоматизировать производственный процесс в реальных практических приложениях.

Благодарности

Автор выражает благодарность научному руководителю Ю.М. Баяковскому и научным консультантам А.В. Игнатенко, В.А. Галактионову и А.Г. Волобою за содействие и помощь в работе, а также всему коллективу лаборатории компьютерной графики и мультимедиа Московского государственного университета им М.В. Ломоносова за плодотворные совместные обсуждения, способствовавшие реализации идей диссертации.

Литература

- Юрковец Д. И. Программное обеспечение для количественного морфологического анализа структур по полутоновым изображениям // Кандидатская диссертация. ВМиК МГУ им. М.В. Ломоносова, Москва. 1998.
- Гаганов В. А., Игнатенко А. В. Устойчивое построение трехмерных моделей объектов по набору изображений с микроскопа // Информационные технологии в проектировании и производстве. 2009. Т. 4. С. 56–64.
- Гаганов В. А., Игнатенко А. В. Реконструкция плоских объектов по изображениям с микроскопа // Программные продукты и системы. 2009. Т. 3. С. 78–82.
- Гаганов В. А., Игнатенко А. В. Трехмерная реконструкция плоских граней прозрачных минералов по набору изображений с микроскопа // Труды конференции Графикон. 2008. С. 227–233.
- Gaganov V., Ignatenko A. Robust Shape from Focus via Markov Random Fields // Proceedings of Graphicon Conference. 2009. Pp. 74–80.
- 6. Гаганов В. А., Игнатенко А. В. Трехмерная реконструкция плоских граней прозрачных объектов по набору изображений с микроскопа // Сборник трудов XV Международной конференции студентов аспирантов и молодых ученых «Ломоносов-2008». 2008. С. 58.
- Scharstein D., Szeliski R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms // Proceedings of ICCV. Vol. 1-3. 2002. Pp. 7–42.

- Cox I., Hingorani S., Rao S. Maximum Likelihood Stereo Algorithm // Computer Vision and Image Understanding. 1996. Vol. 63.
- Nayar S., Nakagawa Y. Shape from focus // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1994. Vol. 16. Pp. 824–831.
- Seitz S., Dyer C. Photorealistic Scene Reconstruction by Voxel Coloring // Proceedings of CVPR. 1997. Pp. 1067–1073.
- Kutulakos K. N., Seitz S. M. A Theory of Shape by Space Carving // International Journal of Computer Vision. 2000. Vol. 38, no. 3. Pp. 199–218.
- Eisert P., Steinbach E., Girod B. Multi-hypothesis, Volumetric Reconstruction of 3-D Objects from Multiple Calibrated Camera Views // Proceedings of ICASSP. 1999. Pp. 3509–3512.
- Hertzmann A., Seitz S. Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs. // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2005. Vol. 27, no. 8. Pp. 1254–1264.
- Woodham R. Reflectance Map Techniques for Analyzing Surface Defects in Metal Castings, Technical Report AI-TR-457: Tech. rep.: 1978.
- Subbarao M., Choi T. Accurate recovery of three-dimensional shape from image focus // IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995. Vol. 17, no. 3. Pp. 266–274.
- Darrell T., Wohn K. Pyramid Based Depth from Focus // Proc. Computer Vision and Pattern Recognition. 1988. Pp. 504–509.
- Yun J., Choi T. Accurate 3-d shape recovery using curved window focus measure // Proc. IEEE International Conference on Image Processing. Vol. 3. 1999. Pp. 910–914.

- Jin H., Favaro P. A variational approach to shape from defocus // Proc. IEEE European Conference on Computer Vision. 2002. Pp. 18–30.
- Subbarao M., Surya G. Depth from defocus by changing camera aperture: A spatial domain approach // Proc. Computer Vision and Pattern Recognition. 1993. Pp. 61–67.
- Rajagopalan A. N., Chaudhuri S. Optimal Recovery of Depth from Defocused Images Using an MRF Model // In Proc. International Conference on Computer Vision. 1998. Pp. 1047–1052.
- Xiong Y., Shafer S. A. Depth from Focusing and Defocusing // In Proc. of the DARPA Image Understanding Workshop. 1993. Pp. 68–73.
- Nair H., Stewart C. Robust focus ranging // Proc. Computer Vision and Pattern Recognition. 1992. Pp. 309–314.
- Pentland A. P. A new sense for depth of field // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1987. Vol. 9. Pp. 523–531.
- Nayar S. K., Noguchi M., Watanabe M., Nakagawa Y. Focus Range Sensors // Proc. of Intl. Conf. on Computer Vision. 1995. Pp. 995–1001.
- Nayar S. K., Nakagawa Y. Shape from focus: An efficient Approach for Rough Surfaces // Proc. of IEEE Conference on Robotics and Automation. Vol. 16. 1990. Pp. 218–225.
- Krotkov E. Focusing // International Journal on Computer Vision. 1987.
 Vol. 1. Pp. 223–237.
- Ahmad M., Choi T.-S. A heuristic approach for finding best focused shape // IEEE Transactions on Circuits and Systems for Video Technology. 2005. Vol. 15, no. 4. Pp. 566–574.

- Asif M., sun Choi T. Shape from Focus Using Multilayer Feedforward Neural Networks // IEEE Transactions on Iamge Processing. 2000. Vol. 10. Pp. 1670–1675.
- Aydin T., Akgul Y. A New Adaptive Focus Measure for Shape From Focus // Proc. British Machine Vision Conference. 2008.
- 30. Niederoest M., Niederoest J., Scucky J. Automatic 3D Reconstruction and Visualization of Microscopic Objects from a Monoscopic Multifocus Image Sequence // Proc. International Workshop on Visualization and Animation of Reality based 3D Models. 2002.
- Fedorov D., Sumengen B., Manjunath B. S. Multi-focus imaging using local focus estimation and mosaicking // Proc. IEEE International Conference on Image Processing. 2006. Pp. 227–231.
- 32. Li H., Manjunath B. S., Mitra S. K. Multisensor image fusion using the wavelet transform // Proc. IEEE International Conference on Image Processing. 1994. P. 51–55.
- Eltoukhy H. A., Kavusi S. A computationally efficient algorithm for multi-focus image reconstruction // In Proc. SPIE. Electronic imaging. 2003. Pp. 332–341.
- 34. Forster B., Ville D. V. D., Berent J. et al. Extended Depth of Focus for Multi-Channel Microscopy Images: A Complex Wavelet Approach // Proceedings of the 2004 IEEE International Symposium on Biomedical Imaging. 2004. Pp. 660–663.
- Seales W., Dutta S. Everywhere-in-focus image fusion using controllable cameras // In Proc. of SPIE. 1996. Pp. 227–234.

- 36. Boykov Y., Jolly M. Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images. // In Proc. International Conference On Computer Vision. 2001. Pp. 105–112.
- Boykov Y., Veksler O., Zabih R. Fast Approximate Energy Minimization via Graph Cuts // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001. Vol. 23, no. 11. Pp. 1222–1239.
- 38. Geman S., Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1984. Vol. 6. Pp. 721–741.
- Boykov Y., Veksler O., Zabih R. Markov Random Fields with Efficient Approximations. // In Proc. Computer Vision and Pattern Recognition. 1998. Pp. 648–655.
- 40. Kolmogorov V., Zabih R. What energy functions can be minimized via graph cuts // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004. Vol. 26. Pp. 65–81.
- Perez P. Markov random fields and images // CWI Quarterly. 1998.
 Pp. 413–437.
- Ishikawa H. Exact optimization for markov random fields with convex priors // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003. Vol. 25. Pp. 1333–1336.
- 43. Boykov Y., Kolmogorov V. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001. Vol. 26. Pp. 359–374.

- 44. Shotton J., Winn J., Rother C., Criminisi A. TextonBoost: Joint Appearance, Shape and Context Modeling for Multi-Class Object Recognition and Segmentation // In Proc. of ECCV. 2006. Pp. 1–15.
- 45. Szeliski R., Zabih R., Scharstein D. et al. A comparative study of energy minimization methods for Markov random fields // In Proc. of ECCV. 2006. Pp. 16–29.
- 46. Yedidia J., Freeman W., Weiss Y. Generalized belief propagation // In Proc. of NIPS. 2000. P. 689–695.
- 47. Kolmogorov V. Convergent Tree-reweighted Message Passing for Energy Minimization // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006. Vol. 28. Pp. 1568–1583.
- Rother C., Kumar S., Kolmogorov V., Blake A. Digital tapestry // In Proc. of Computer Vision and Pattern Recognition. 2005. Pp. 589–596.
- Hampel F. R., Rousseeuw P. J., Ronchetti E., Stahel W. A. Robust Statistics: The Approach Based on Influence Functions // Wiley Series in Probability and Statistics. 1986.
- Haralick R. M., Joo H., Lee C.-N. et al. Pose estimation from corresponding data // IEEE Trans. Systems Man Cybernetics. 1989. Vol. 19. Pp. 1426–1446.
- Stewart C. Robust parameter estimation in computer vision // SIAM Review. 1999. Vol. 1, no. 3. Pp. 513-537.
- P.J.Rousseeuw. Least median of squares regression // Journal of the American Statistical Association. 1984. Vol. 79. Pp. 871–880.

- Bernholt T. Computing the least median of squares estimator in time O(nd) // Proceedings of ICCSA 2005, LNCS 3480. 2005. Pp. 697–706.
- Olson C. An Approximation Algorithm for Least Median of Squares Regression // Information Processing Letters. 1997. Vol. 63. Pp. 237–241.
- Duda R. O., Hart P. E. Use of the Hough Transformation to Detect Lines and Curves in Pictures // Communications of the ACM. 1979. Vol. 15. P. 11–15.
- J.Illingworth, J.Kittler. A survey of the Hough transform // In Proc. of CVGIP. Vol. 44. 1988. Pp. 87–116.
- 57. Fernandes L. A., Oliveira M. M. Real-time line detection through an improved Hough transform voting scheme // Pattern Recognition. 2004.
 Vol. 41, no. 9. Pp. 299–314.
- 58. M.A.Fischler, R.C.Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography // In Communications of the ACM. 1981. Vol. 24, no. 6. Pp. 381–395.
- P.H.S.Torr, A.Zisserman. MLESAC: A New Robust Estimator with Application to Estimating Image Geometry // In Proc. of CVIU. Vol. 78. 2000. Pp. 138–156.
- Myatt D., Torr P., Nasuto S., Craddock. R. NAPSAC: High Noise, High Dimensional Robust Estimation // In Proc. BMVC. 2002. Pp. 458–467.
- Chum O., Matas J. Randomized RANSAC with T(d,d) test. // In Proc. BMVC. 2002. Pp. 448–457.
- Chum O., Matas J., Kittler J. Locally Optimized RANSAC. // In Proc.
 25th Pattern Recognition Symposium (DAGM). 2003. Pp. 236–243.

- Friedman J., Hastie T., Tibshirani R. Additive Logistic Regression: a Statistical View of Boosting // Annals of Statistics. 1998. Vol. 28. P. 2000.
- 64. Schapire R., Freund Y., Bartlett P., Lee. W. S. Boosting the margin: A new explanation for the effectiveness of voting methods // In Machine Learning: Proceedings of the Fourteenth International Conference. 1997.
- 65. Breiman L. Bagging Predictors // Machine Learning. 1996. Pp. 123–140.
- Burges C. J. C. A tutorial on support vector machines for pattern recognition // Data Mining and Knowledge Discovery. 1998. Vol. 2. Pp. 121–167.
- Ninomiya H., Sasaki A. A Study on Generalization Ability of 3-Layer Recurrent Neural Networks // Proc. of IJCNN. 2002. Pp. 1063 – 1068.
- Rowley H. A., Member S., Baluja S., Kanade T. Neural Network-Based Face Detection // IEEE Transactions On Pattern Analysis and Machine intelligence. 1998. Vol. 20. Pp. 23–38.
- Viola P., Jones M. Robust real-time face detection // International Journal of Computer Vision. 2004. Vol. 57. Pp. 137–154.
- 70. Japkowicz N., Stephen S. The class imbalance problem: a systematic study // In Proc. of IDAJ. 2002. Pp. 429–450.
- 71. Вежневец В. П. Оценка качества работы классификаторов. http:// cgm.computergraphics.ru/content/view/106.
- 72. Laurentini A. The visual hull concept for silhouette-based image understanding // IEEE Trans. Pattern Analysis and Machine Intelligence. 1994.
 P. 150–162.
- 73. Intel Integrated Performance Primitives. http://software.intel.com/ en-us/articles/intel-ipp.