

Федеральное государственное бюджетное учреждение науки
Институт проблем передачи информации им. А.А. Харкевича
Российской академии наук (ИППИ РАН)

На правах рукописи

Чочиа Павел Антонович

**ТЕОРИЯ И МЕТОДЫ ОБРАБОТКИ ВИДЕОИНФОРМАЦИИ
НА ОСНОВЕ ДВУХМАСШТАБНОЙ МОДЕЛИ ИЗОБРАЖЕНИЯ**

Специальность 05.13.18 —

математическое моделирование, численные методы и комплексы программ

ДИССЕРТАЦИЯ

на соискание ученой степени

доктора технических наук

Москва – 2015

Оглавление

ВВЕДЕНИЕ	8
ГЛАВА 1. МОДЕЛИ ИЗОБРАЖЕНИЙ.....	19
Введение	19
1.1. Модель формирования изображения	21
1.1.1. Непрерывное и дискретное представление изображения.....	21
Представление изображения в непрерывной модели.....	21
Представление изображения в дискретной модели.....	23
1.1.2. Области изображения и области анализа.....	24
1.2. Двухмасштабная многокомпонентная модель изображения	26
1.2.1. Подход к построению модели	26
Два масштаба — масштаб окрестности и масштаб фрагмента.....	28
Многокомпонентное представление изображения.....	28
1.2.2. Модель окрестности (масштаб элементов).....	29
Статистические связи элементов окрестности	32
1.2.3. Модель фрагмента (масштаб объектов).....	35
Связь параметров модели с характеристиками изображения	38
1.2.4. Модель цветного изображения.....	40
1.3. Вероятностная модель контурного изображения.....	40
1.3.1. Общая постановка задачи	41
Особенности дискретного представления	42
1.3.2. Построение контурного изображения.....	44
Параметры модели и их соотношения	44
1.3.3. Эксперименты по синтезу контурных изображений	47
1.4. Оценка сложности изображения	49
Сложность как характеристика изображения.....	49
О сложности сигналов и функций.....	50
Использование оценки сложности в задачах восстановления	51
1.4.1. Применение двумерных вариаций для оценки сложности.....	52
Двумерные вариации и их дискретные аналоги	52
Двумерная вариация Кронрода	54

Показатель размеров объектов изображения.....	57
Двумерная вариация как оценка сложности изображения.....	57
О сравнении и интерпретации оценок	59
1.4.2. Экспериментальные исследования.....	59
Влияние шума на оценку сложности изображения	61
Зависимости двумерных вариаций от размеров и поворота деталей	62
Выводы касательно оценки сложности изображения	63
1.5. Выводы и результаты	64
ГЛАВА 2. ДЕКОМПОЗИЦИЯ ИЗОБРАЖЕНИЯ.....	66
Введение	66
2.1. Задача сглаживания изображения при сохранении контурных перепадов	67
2.1.1. Искажения контурных перепадов при использовании арифметического среднего и медианы по фрагменту	68
2.1.2. Сглаживание на основе анализа распределения	70
Сигма-фильтр.....	71
2.2. Разработка метода декомпозиции изображения.....	72
О размере центральной области.....	74
2.2.1. Алгоритм декомпозиции «D»	75
2.2.2. Алгоритм декомпозиции «D*» без повышения резкости.....	79
2.2.3. О компонентах, формируемых при декомпозиции.....	81
2.3. Декомпозиция цветного изображения	85
2.4. Изменение сложности изображения при сглаживании и декомпозиции	88
2.4.1. Эксперименты с модельными изображениями.....	89
2.4.2. Эксперименты с реальными изображениями.....	91
2.5. Выводы и результаты	94
ГЛАВА 3. РАЗРАБОТКА МЕТОДОВ ФИЛЬТРАЦИИ, КОРРЕКЦИИ И УЛУЧШЕНИЯ	
ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДВУХМАСШТАБНОЙ МОДЕЛИ	96
Введение	96
3.1. Фильтрация помех	97
3.1.1. Фильтрация импульсных помех	97
Постановка задачи	97
Наиболее распространенные алгоритмы фильтрации.....	98
Вероятности ошибок предсказания.....	100

Алгоритм ранговой пороговой фильтрации	103
Экспериментальные данные	104
3.1.2. Фильтрация периодических помех	106
Фильтрация в частотной области	106
Построение фильтра-маски	108
Выполнение операции фильтрации	110
Фильтрация в пространственной области	111
3.2. Автоматическая яркостная (градационная) коррекция	113
3.2.1. Стандартные подходы	113
Задача градационной коррекции	113
Формализация и используемые методы	114
3.2.2. Алгоритм автоматической градационной коррекции	116
Функция локальных контрастов	116
Гипотеза о константности функции локальных контрастов	117
Алгоритм выравнивания функции локальных контрастов	119
Эквализация клиппированной гистограммы	121
Эксперименты по автоматической градационной коррекции	122
3.3. Улучшение изображений повышением локальных контрастов	125
3.3.1. Обзор используемых методов	126
3.3.2. Классификация методов повышения локальных контрастов	127
3.3.3. Сравнение методов	129
3.3.4. Применение декомпозиции для улучшения изображений	130
3.4. Улучшение цветных изображений	131
3.4.1. Градационная коррекция	133
3.4.2. Методы глобальной цветовой коррекции	134
3.4.3. Повышение локальных контрастов цветных изображений	135
3.5. Выводы и результаты	139
ГЛАВА 4. РАЗРАБОТКА МЕТОДОВ АНАЛИЗА ИЗОБРАЖЕНИЙ НА ОСНОВЕ	
ДВУХМАСШТАБНОЙ МОДЕЛИ	141
Введение	141
4.1. Обнаружение объектов заданной площади	141
4.1.1. Обнаружение объектов с площадью больше заданной	141
4.1.2. Обнаружение объектов с площадью меньше заданной	142

4.1.3. Обнаружение объектов с площадью в интервале $[Q_1, Q_2]$	142
4.2. Обнаружение различных объектов на изображениях.....	144
4.3. Применение декомпозиции для нахождения границ объектов.....	146
4.4. Быстрое корреляционное совмещение изображений.....	149
4.5. Обнаружение дефектов на снимках электронных микросхем.....	152
4.5.1. Постановка задачи.....	152
4.5.2. Подход к построению алгоритма обнаружения дефектов.....	153
4.5.3. Предварительная обработка изображений.....	156
4.5.4. Формирование двумерного массива несовпадений.....	157
4.5.5. Формирование списка областей несовпадений.....	158
4.5.6. Удаление ложных пятен	160
4.6. Выводы и результаты	163
ГЛАВА 5. ПРИМЕНЕНИЕ МОДЕЛИ И МЕТОДОВ ОБРАБОТКИ ДЛЯ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ.....	165
Введение	165
5.1. Сегментация изображений на основе прослеживания контуров	167
5.1.1. Особенности задачи сегментации снимков микросхем.....	167
5.1.2. Получение выровненных изображений.....	169
5.1.3. Алгоритм прослеживания контуров.....	171
5.1.4. Формирование объединенной карты контуров.....	175
5.1.5. Обработка карты контуров и формирование результата.....	177
5.2. Цвето-текстурная сегментация изображений.....	180
5.2.1. Особенности задачи сегментации изображений.....	182
5.2.2. Построение пространства признаков	184
5.2.3. Метрика в пространстве признаков.....	187
5.2.4. Структура алгоритма сегментации.....	188
5.2.5. Первичная сегментация — пирамидальный алгоритм	189
5.2.6. Вторичная сегментация — упрощение графа кластеров.....	193
5.2.7. Эксперименты по выбору параметров преобразований.....	195
5.3. Выводы и результаты	197
ГЛАВА 6. ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ МЕТОДОВ ДЛЯ ОБРАБОТКИ И АНАЛИЗА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ.....	201
Введение	201

6.1. Предварительная обработка видеок кадров.....	203
6.1.1. Особенности получаемых видеоданных	203
6.1.2. Устранение систематических искажений	203
6.1.3. Устранение периодических помех	205
6.1.4. Выравнивание и усиление локальных контрастов сигнала.....	206
6.2. Компенсация дрейфа и получение усредненного изображения	207
6.3. Формирование карты контурных линий капилляров	210
6.3.1. Формирование опорного множества областей	211
6.3.2. Нахождение множества контурных точек	212
6.3.3. Формирование первичной карты контурных линий.....	212
6.3.4. Синтаксическая фильтрация и корректировка линий	213
6.3.5. Упрощение карты линий и карты областей	215
6.4. Выбор мажоритарного капилляра.....	216
6.5. Морфологический анализ и корректировка контурных линий.....	217
6.5.1. Векторное представление контурных линий.	217
6.5.2. Анализ морфологической характеристики контурной линии.....	218
6.5.3. Корректировка расположения контурных линий	219
6.6. Построение распрямляющего отображения	220
6.7. Определение границ отделов капилляра.....	222
6.8. Преобразование кадров видеопоследовательности.....	223
6.9. Определение параметров кровотока	224
6.10. Выводы и результаты.....	226
ГЛАВА 7. МОДИФИКАЦИЯ МОДЕЛИ И МЕТОДОВ ОБРАБОТКИ ДЛЯ ТРЕХМЕРНЫХ ИЗОБРАЖЕНИЙ.....	228
Введение	228
7.1. Особенности трехмерных изображений	229
7.1.1. Области анализа и соседство элементов	230
7.2. Модификация двухмасштабной многокомпонентной модели.....	231
7.2.1. Модель окрестности (масштаб малого размера).....	232
7.2.2. Модель фрагмента (масштаб большого размера)	233
7.3. Модификация методов частотной фильтрации при переходе в 3D.....	234
7.4. Модификация методов пространственной обработки и анализа при переходе в 3D.....	235
7.4.1. Методы, использующие оценку среднего по фрагменту	236

7.4.2. Операторы контурных перепадов.....	236
7.4.3. Фильтрация импульсных помех	238
7.4.4. Декомпозиция изображения	240
7.4.5. Обнаружение объектов заданного объема	241
7.5. Модификация некоторых вычислительных алгоритмов.....	242
7.5.1. Вычисление суммы по прямоугольному параллелепипеду.....	242
7.5.2. Вычисление порядковых статистик по прямоугольному параллелепипеду.....	244
7.6. Параллельный алгоритм вычисления порядковых статистик.....	245
7.7. Выводы и результаты	250
ЗАКЛЮЧЕНИЕ И ОСНОВНЫЕ РЕЗУЛЬТАТЫ.....	252
ЛИТЕРАТУРА.....	260
ПРИЛОЖЕНИЯ.....	284
Приложение 1. Обработка снимков планет	284
Приложение 2. Обнаружение дефектов на снимках поверхностей микросхем	292
Приложение 3. Номинация программного комплекса на Гран-при СеВит	295
Приложение 4. Документы о внедрении	296

ВВЕДЕНИЕ

Важнейшим средством получения исходных сведений о наблюдаемой сцене является видеoinформация (визуальная информация), у живых организмов поступающая через органы зрения, а у технических устройств — через те или иные сенсоры, выполняющие, по сути, аналогичные функции. Объемы поступающей визуальной информации и сложность решаемых задач анализа постоянно возрастают. Основная роль при этом отводится различным системам технического зрения, обеспечивающим регистрацию, переработку и анализ видеoinформации. Если раньше преобразование поступающих видеоданных осуществлялось в основном для последующего их визуального предъявления человеку, то в последнее время акценты сместились в сторону автоматического принятия решений. В связи с этим главной целью этапа обработки и анализа данных становится извлечение необходимой информации и предоставление ее системе более высокого уровня для выбора соответствующей реакции.

Число областей, в которых видеoinформация выступает как основной вид исходных и/или передаваемых данных, огромно. К ним относятся системы видеонаблюдения, мониторинга, видеотелефонии, регистрирующие и передающие огромные объемы данных, разнообразные автономные системы, принимающие решение на основе анализа визуальной информации, персонализированное телевизионное вещание, полиграфия с возросшим объемом иллюстраций, медицина, Интернет, и многие другие направления. При этом наряду со значительным повышением уровня развития техники самую существенную роль продолжают играть цифровые методы обработки видеoinформации. Они обеспечивают улучшение изображений для зрительного восприятия, сжатие видеоданных для хранения и передачи по каналам связи, а также анализ, распознавание и интерпретацию зрительных образов для принятия решений при управлении поведением автономных технических систем [188].

Понятие *видеoinформация* является обобщающим для неподвижных двумерных, трехмерных, а также движущихся изображений (видеопоследовательностей). Поскольку средством обработки и анализа данных, как правило, является компьютер, предполагающий представление данных в цифровой форме, то, если не оговорено иначе, под изображением будем всегда понимать *дискретизованное изображение*, представимое в виде двумерного массива чисел. Такие изображения могут быть получены как различными оптическими системами с оконечной цифровой регистрацией амплитуды сигнала интересующих диапазонов электромагнитных волн, так и сформированы иным образом. К последним относятся, например, изображения, получаемые радиолокационными сканерами, компьютерными томографами, компьютерным синтезом и др. Общим для всех видов является то, что любое изображение есть отображение (образ, проекция) какой-то реальной или гипотетической пространственной сцены

(совокупности исходных объектов) на пространство изображения (двумерное или трехмерное), а также то, что значение каждого из элементов изображения есть интегральная характеристика выбранного физического параметра на некотором пространственном участке сцены, который соответствует данному элементу, согласно выбранному способу отображения.

Научное направление цифровой обработки и анализа видеоинформации развивается весьма высокими темпами. Ежегодно издаются десятки книг и сотни статей по вопросам обработки и анализа изображений как российских так и зарубежных исследователей, из которых реально возможно упомянуть лишь небольшую часть. Из монографий отметим [3, 5, 19, 21, 23, 32, 40, 47, 48, 52, 53, 57, 58, 86, 90, 92, 93, 98, 100, 108, 113, 122, 132, 134, 139, 147, 151, 154, 157, 158, 205, 206, 207, 208, 209, 219, 263, 266, 270, 274, 284, 289, 300, 326, 330, 333, 343, 353, 357, 364]. За рубежом регулярно издаются печатные журналы по этой тематике, из которых наиболее известны «Computer Vision, Graphics and Image Processing», «Computer Vision and Image Understanding», «Pattern Recognition», «International Journal of Computer Vision», «Computer Vision and Pattern Recognition», «Image and Vision Computing», «Journal of Real-Time Image Processing», «Journal of Mathematical Imaging and Vision», «Machine Vision and Applications», «Pattern Analysis and Applications», «IEEE Transactions on Image Processing», «IEEE Transactions on Pattern Analysis and Machine Intelligence», «IEEE Transactions on Medical Imaging». Многие статьи печатаются также в других журналах, издаваемых Институтом инженеров по электротехнике и электронике (IEEE), а кроме того в различных специализированных сборниках Института, издаваемых не регулярно, например [102, 119, 120, 149].

В нашей стране из периодических изданий, в той или иной мере специализированных в области обработки изображений, можно упомянуть лишь журналы «Компьютерная оптика» («Computer optics» в переводном варианте) и «Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications», последний из которых издается, к сожалению, только на английском языке. На русском языке отдельные статьи публикуются в журналах «Автометрия» («Optoelectronics, Instrumentation and Data Processing»), «Исследование Земли из космоса» («Izvestiya, Atmospheric and Oceanic Physics»), «Радиотехника и электроника» («Journal of Communications Technology and Electronics»). Ранее издавались различные сборники научных трудов, как например [1, 2, 4, 118, 153, 156], в том числе специализированные сборники «Иконика», в частности [60, 61, 68], выходившие под разными названиями. Поскольку обработка видеоинформации является прикладным направлением науки, многие результаты, относящиеся к данной области, публикуются в тематических научных изданиях по соответствующим отраслям, среди которых необходимо отметить журналы «Дефектоскопия» («Russian Journal of Nondestructive Testing») и «Медицинская техника» («Biomedical Engineering»).

Кроме того в последние годы появились многочисленные зарубежные и отечественные интернет-издания и журналы, публикующие в электронном виде статьи, книги, а также другие материалы по вопросам обработки, анализа и интерпретации видеоинформации.

В развитие вопросов анализа и обработки видеоинформации значительный вклад внесли отечественные и зарубежные ученые: Г.И. Василенко, С.Б. Гуревич, В.П. Дворкович, Ю.И. Журавлев, Ю.Б. Зубарев, В.С. Киричук, Д.С. Лебедев, В.В. Моттль, И.Б. Мучник, В.П. Пяткин, Ю.П. Пытьев, С.С. Садыков, В.В. Сергеев, В.А. Соيفер, А.А. Спектор, И.И. Цуккерман, В.В. Яншин, Л.П. Ярославский, Н. Ахмед, Р. Вудс, А. Гагалович, Р. Гонсалес, Б. Гоулд, А. Джайн, Р. Дуда, Д. Марр, А. Нетравали, А. Оппенгейм, Т. Павлидис, У. Прэтт, А. Рабинер, К. Рао, А. Розенфельд, М. Сондхи, О. Фожра, К. Фу, А. Хабиби, Р. Харалик, Я. Харт, Б. Хорн, Т. Хуанг, Г. Эндрюс, Б. Яне, А. Bovik, U. Grenander, E. Hall, R. Jain, J.-S. Lee, G. Nagy, L. Shapiro, J. Togiwaki и многие другие.

Актуальность проблемы. В технических устройствах анализ поступающей видеоинформации выполняют системы технического зрения, области применения которых чрезвычайно широки. Ключевую роль при этом играют методы обработки, анализа и извлечения требуемой информации из изображений.

Несмотря на множество исследований, единого подхода к описанию свойств изображения как многомерного сигнала не выработано. Это мешает как обоснованию и сравнению, так и разработке новых и эффективных методов анализа и обработки видеоинформации. Решение данного вопроса является важным и необходимым.

Разработка и применение методов преобразования и анализа любого сигнала строится на основе априорной информации о его свойствах. Формулировка свойств служит моделью сигнала и составляет один из фундаментальных аспектов теории и методов его обработки. Математически модель сигнала и его предполагаемых искажений (помех) можно представить в виде функций, описывающих их существенные характеристики и зависимости. В области обработки и анализа видеоинформации таковой является модель изображения. Как и в случае с любой моделью сигнала, построение модели цифрового изображения необходимо для достижения двух целей: более точного и полного описания межэлементных связей и свойств изображения как исходного информационного объекта, а также предоставления удобного фундамента для разработки эффективных методов его анализа и преобразования.

Моделей дискретных изображений предложено много. Наиболее известные из них можно объединить в следующие классы: модели, основанные на особенностях зрительного восприятия, в том числе цветового [42, 58, 90, 127, 247, 256, 265], стационарные и нестационарные статистические модели [47, 83, 114, 258, 278], авторегрессионные и другие

модели линейного предсказания [32, 49, 125], различные марковские модели [85, 97, 207, 300, 370], модели двухкомпонентного источника [371], разрывные модели, описывающие совокупности протяженных областей, в том числе фасеточные и мозаичные модели [123, 145, 212, 272, 293, 323].

Общий недостаток предложенных моделей состоит в том, что они не универсальны с позиции величины области анализа (масштаба). В каждой из них выбирается некоторая одна совокупность соотношений, описывающая свойства изображения либо только для малых областей, в пределах нескольких элементов (модели линейного предсказания, стохастические), либо только для протяженных областей (разрывные, двухкомпонентные модели). Основная проблема заключается в том, что статистические характеристики областей малых и больших размеров (масштабов) существенно различаются. Причина этого исходит из важнейших свойств, присущих изображениям — наличию протяженных областей с малыми изменениями яркости, разделенных контурными границами. Модели малого масштаба не могут описать свойства протяженных областей, и наоборот, модели большого масштаба не в состоянии корректно учитывать особенности контурных участков изображения. Корректно решить проблему универсальности, вводя в какую-то модель масштабный параметр при использовании одних и тех же соотношений, не удастся. Данный недостаток существенно ограничивает возможности применения моделей как при изучении свойств изображений, так и при разработке методов их анализа и обработки.

Таким образом актуальной научной проблемой является создание теории, позволяющей, адекватно описывать свойства изображений как на малых, так и на больших расстояниях, которая бы позволила повысить точность представления данных и стала фундаментом для разработки новых эффективных методов и алгоритмов обработки и анализа видеоинформации.

Цель и задачи исследования: построение модели цифрового изображения, способной описывать его свойства в пределах областей анализа различной протяженности (масштаба), и разработка на основе такой модели эффективных методов и алгоритмов обработки и анализа видеоинформации. При этом *объектом исследований* является цифровая видеоинформация, а *предметом исследований* — модели, методы и алгоритмы ее обработки и анализа.

Для достижения поставленной цели решаются следующие задачи:

— разработка математической модели, описывающей свойства изображения в пределах области анализа требуемого масштаба (размера) путем аппроксимации значений его элементов участками аналитически задаваемых поверхностей, и предоставляющую возможность создавать на основе модели эффективные методы и алгоритмы обработки и анализа видеоинформации;

- разработка контурной модели изображения путем введения формальных топологических свойств границ объектов, требуемых для моделирования источника контурных изображений;
- базируясь на исследуемой модели изображения и нелинейных методах локального оценивания данных разработка метода и алгоритма декомпозиции (разделения) изображения на компоненты с различным информационным содержанием;
- разработка способа оценивания сложности изображения двумерными вариациями;
- разработка новых подходов к построению методов обработки и анализа видеoinформации на основе двухмасштабной модели изображения и методе декомпозиции, а также построение соответствующих алгоритмов преобразования;
- исследование возможностей применения разрабатываемых моделей, методов и алгоритмов к данным более высокой сложности: цветным (многозональным), движущимся (видеопоследовательностям) и трехмерным (объемным) изображениям;
- решение при помощи созданной теоретической и алгоритмической базы важных научно-технических задач обработки и анализа данных.

Методы исследования базируются на использовании теории обработки непрерывных и дискретных сигналов и изображений, математической статистики и теории статистических решений, теории информации, цифрового моделирования, статистического анализа, в том числе с использованием двумерных вариаций, экспертного оценивания.

Научная новизна работы.

Получены следующие научные результаты, являющиеся новыми на период их разработки и опубликования.

1. Разработана двухмасштабная многокомпонентная математическая модель изображения, описывающая его свойства в пределах локальных областей анализа малого и большого масштабов, отличающаяся представлением сигнала изображения комбинацией участков аналитически задаваемых поверхностей и случайных текстурной, детальной и шумовой компонент.
2. Построена вероятностная модель контурного изображения, отличающаяся аксиоматикой задания свойств границ объектов и топологических элементов. Разработан алгоритм источника дискретных изображений, позволяющий моделировать контурные изображения с задаваемыми вероятностными характеристиками.
3. Предложен способ оценивания сложности изображения при помощи двумерных вариаций; введена новая характеристика, названная *показатель размеров объектов*.
4. Разработаны метод и алгоритм декомпозиции изображения на компоненты с различным информационным содержанием согласно используемой модели изображения: кусочно-гладкую

компоненту, несущую информацию о протяженных объектах и резких границах между ними, и текстурно-детальную компоненту, содержащую малоразмерные детали, текстуру и шум.

5. На основе двухмасштабной модели и алгоритма декомпозиции изображения разработаны новые подходы и модифицирован ряд известных методов и алгоритмов фильтрации, коррекции, улучшения и анализа изображений. В том числе методы: фильтрации импульсных помех, фильтрации периодических помех в частотной и пространственной областях, автоматической градиентной коррекции, повышения локальных контрастов и улучшения изображений, обнаружения объектов заданной площади, обнаружения различий набора объектов на изображении, выделения границ объектов; предложены способы применения разработанных методов для преобразования цветных и многоканальных изображений.

6. Используя разработанные методы и алгоритмы решен ряд важных научно-технических задач: обнаружения дефектов на снимках электронных микросхем; цвето-текстурной сегментации изображений на основе анализа расстояний в пространстве признаков; сегментации прослеживанием контуров сложного вида; анализа видеоданных, формируемых капилляроскопом.

Основные положения и результаты, выносимые на защиту:

1. Предложенная двухмасштабная многокомпонентная модель цифрового изображения позволяет более точно описывать свойства цифровых изображений и предоставляет основу для разработки новых эффективных методов и алгоритмов обработки и анализа видеoinформации.
2. Вероятностная модель двумерного контурного изображения описывает основные топологические свойства контуров объектов на изображении и позволяет построить источник случайных дискретных контурных изображений с задаваемыми свойствами и характеристиками.
3. Разработанные метод и алгоритм декомпозиции изображения обеспечивают разделение изображения на компоненты с различным информационным содержанием, что позволяет применять специфические для них алгоритмы дальнейшего анализа или обработки.
4. Разработанные методы и алгоритмы обработки и анализа изображений, основанные на двухмасштабной многокомпонентной модели, обеспечивают более высокую эффективность методов фильтрации, коррекции, улучшения изображений и обнаружения объектов, подтверждая адекватность предложенной модели и метода декомпозиции.
5. Разработанные методы и алгоритмы успешно применены для решения следующих важных научно-технических задач, что подтверждает эффективность предложенных методов:
 - задачи обнаружения дефектов на снимках электронных микросхем;
 - задачи сегментации снимков микросхем на основе прослеживания границ сложного вида;
 - задачи сегментации изображений общего вида на основе анализа расстояний в пространстве яркостно-цвето-текстурных признаков;

— задачи обработки видеоданных на примере автоматического анализа видеопоследовательностей, формируемых капилляроскопом.

Научная значимость.

Разработаны теоретические положения, относящиеся к области теории и методов цифровой обработки видеоинформации. Полученные результаты могут использоваться для разработки новых методов анализа, обработки и сжатия неподвижных, движущихся или объемных изображений; при проектировании вновь создаваемых систем анализа и переработки видеоданных, в частности промышленных роботов, автоматических систем анализа в дефектоскопии и медицинском приборостроении, а также в других приложениях. Теоретические результаты могут быть использованы в дальнейших научных исследованиях, а также в учебном процессе при изучении основ обработки видеоинформации.

Реализация и практическая значимость результатов работы.

Реализация результатов диссертационной работы осуществлялась в виде комплексов программ на базе разработанных автором специализированных систем обработки изображений, созданных в разное время для ЭВМ различных типов. Целью создания таких систем являлась как разработка готового комплекса программ для обработки и анализа видеоинформации, так и предоставление программисту-исследователю операционной среды, специализированной для проектирования новых алгоритмов и программ обработки изображений [77, 164, 180, 191, 196, 202, 231]. Разработанный автором специализированный программный комплекс обработки изображений «IRBIS» [164, 180, 191, 196], современная модификация которого и в настоящее время используется в Институте проблем передачи информации РАН, выставлялся на Международной выставке компьютерных и информационных технологий CeBit (г. Ганновер, Германия) в 1992 г. и был признан Национальным номинантом на Гран-при «Golden Softies» в области инноваций и разработки программного обеспечения в Европе (см. Приложение 3).

Методы и алгоритмы, изложенные в диссертационной работе, могут быть непосредственно использованы для восстановления и улучшения искаженных изображений, для анализа видеоинформации, поступающей в виде отдельных изображений или видеопоследовательностей. Они применялись и применяются в ИППИ РАН и ряде других организаций при создании систем переработки, хранения, и отображения видеоинформации и обеспечили их высокую эффективность. Для реализации алгоритмов использовалось программирование на языках “С”, “С++” и на языках ассемблера. В Приложениях 1 и 2 приведена иллюстрация некоторых практических применений созданных методов.

Разработанные методы и программные комплексы применялись для обработки снимков поверхности планет, переданных отечественными автоматическими межпланетными станциями

«Марс-4, -5», «Венера-9, -10», «Венера-13, -14», «Фобос-2», «Венера-15, -16» (совместно с НИИ космического приборостроения и ОКБ МЭИ), для обработки данных, получаемых системами наблюдения поверхности Земли (АО «Российские космические системы»), при разработке систем автоматического контроля и диагностики автоматизированных систем дефектоскопии (НИИ Интроскопии МНПО «Спектр»), для восстановления архивных фотоснимков (НИЦ технической документации СССР), для анализа видеоданных, формируемых капилляроскопом (ЗАО Центр «Анализ веществ»). Результаты проведенных исследований использовались при выполнении работ по темам, выполнявшимся по правительственным программам, распоряжениям Президиума Академии наук, договорам с зарубежными компаниями и отечественными организациями, в частности с компаниями L.H. Conceil Optronic (France) — разработка алгоритмов и программ согласования и коррекции изображений в реальном времени для видеокамеры кругового обзора, Samsung Advanced Institute of Technology (South Korea) — разработка алгоритмов цвето-текстурной сегментации изображений общего вида, MicroSpec Technologies Ltd. (Carl Zeiss Group) — разработка комплекса алгоритмов и программ анализа снимков поверхности микросхем, получаемых электронными и оптическими микроскопами.

Публикации.

Основные научные результаты, изложенные в диссертации, опубликованы в следующих 84 печатных работах и переводах [15, 17, 51, 55, 64, 77, 78, 95, 99, 101, 105, 109, 117, 129, 143, 150, 155, 159-163, 165-171, 173-179, 181-186, 189, 190, 192-195, 197-201, 203, 216-218, 226-230, 232-245, 252, 253, 255, 311, 312, 314, 337, 338], включая один патент СССР [55] и два зарубежных патента [311, 312]. В публикациях, совместных с соавторами, соискателю принадлежат основные результаты, относящиеся к тематике диссертационной работы. Кроме того, часть результатов применения изложенных теории и методов анализа и обработки видеоинформации опубликована в статьях: [16, 38, 128, 261, 354]. Всего по теме диссертации, применению разработанных методов, а также по смежным вопросам автором опубликовано более 100 печатных работ.

Личный вклад автора. Научные положения и результаты, составляющие основное содержание диссертации, получены автором лично.

Апробация работы. Основные положения диссертационной работы на разных стадиях докладывались и обсуждались на многих отечественных и международных конференциях и симпозиумах, в том числе: Всесоюзной конференции «Оптическое изображение и регистрирующие среды» (Ленинград, 1982), Всесоюзных конференциях «Обработка изображений и дистанционные исследования» (Новосибирск, 1985 и 1987), Всесоюзном симпозиуме «Зрение организмов и роботов» (Вильнюс, 1985), Всесоюзной конференции

«Автоматизированные системы обработки изображений» (Ленинград, 1989), Всесоюзных конференциях «Проблемы создания систем обработки, анализа и понимания изображений» (Ташкент, 1989, 1990 и 1991), Международном семинаре «Digital Image Processing in medicine, remote sensing and visualization of information» (Рига, 1992), International Workshop «Image Processing and Computer Optics» (Самара, 1994), Всероссийских конференциях «Математические методы распознавания образов» (Ленинградская область, 2007, Казань, 2013, Светлогорск, 2015), Международной научно-практической конференции «Современные информационные технологии и ИТ-образование» (Москва, 2014), International Conference on Computer Graphics and Vision (Москва, 2009), Lunar and Planetary Science Conferences №№ XVII, XIX, and XXI (Houston, USA, 1986, 1988, 1990), International Workshop on Digital Image Processing and Computer Graphics (Vienna, Austria, 1997).

Кроме того, результаты работы докладывались на семинарах Института проблем передачи информации РАН, Института систем обработки изображений РАН (Самара), Института проблем информатики РАН, Института космических исследований РАН, Вычислительного центра СО РАН (Новосибирск), ВНИИ полиграфии, НИИ Интроскопии МНПО «СПЕКТР», НИЦ Технической документации СССР, National Institute of Research in Computer Science and Control (INRIA) (Rocquencourt, France), Fraunhofer Institute for Computer Graphics Research IGD (Darmstadt, Germany), Institute of Information Processing Austrian Academy of Sciences (Vienna, Austria), Samsung Advanced Institute of Technology (Suwon, South Korea), Компании L.H. Conceil Optronique (Plaisir, France), Компании Huawei Technologies (Москва), а также на многих других конференциях, симпозиумах и семинарах.

Краткое содержание глав диссертации

Первая глава посвящена построению и исследованию необходимых математических моделей изображения. Прежде всего рассматривается модель формирования изображения. Определяются особенности непрерывного и дискретного представлений изображения, а также его основные характеристики. Дается определение локальной области анализа, типичной для пространственных методов. Формулируется двухмасштабная многокомпонентная модель изображения, позволяющая описывать взаимосвязи элементов как на малых расстояниях, в пределах нескольких шагов дискретизации, так и на больших, соразмерных объектам изображения. Верность предложенной модели подтверждена проведенными исследованиями. Построена вероятностная модель двумерного контурного изображения, позволяющая создавать топологически различные контурные изображения. Реализованный на ее основе алгоритм дает возможность формировать двумерные массивы данных, отвечающие основным свойствам обычного изображения. Исследована возможность оценки сложности изображения при помощи

двумерных вариаций. Показано, что предложенная оценка хорошо отражает морфологическую структуру изображения.

Во второй главе исследуется проблема декомпозиции изображения, которая заключается в разделении его на компоненты согласно принятой двухмасштабной многокомпонентной модели. Декомпозиция рассматривается как задача сглаживания изображения при условии сохранения контурных перепадов; ее решение ищется в области локальных методов анализа. Предлагаются метод и алгоритм декомпозиции для одноканальных изображений, а также его модификации для многоканальных (цветных) изображений. Исследуется изменение сложности изображения в результате различных вариантов сглаживания. Показывается, что алгоритм декомпозиции дает наилучшие результаты как с позиции восстановления кусочно-гладкой составляющей, так и с позиции оценки сложности изображения.

Третья глава посвящена разработке и исследованию методов и алгоритмов фильтрации, коррекции и улучшения изображений на основе двухмасштабной многокомпонентной модели. Рассмотрены методы удаления импульсных помех, частотные и пространственные методы фильтрации периодических помех. Исследованы методы автоматической градационной коррекции; предложена функция локальных контрастов и разработан алгоритм улучшения изображения, основанный на ее выравнивании. Базируясь на многокомпонентной модели, проведены классификация и сравнение методов повышения локальных контрастов как монохромных, так и цветных изображений; показано, что наилучшие результаты достигаются при использовании предварительной декомпозиции изображения.

В четвертой главе исследуется применение двухмасштабной многокомпонентной модели изображения для построения алгоритмов анализа видеoinформации. Рассматриваются задачи выделения объектов по величине их площади, нахождение различий в наборах объектов на парах изображений. Показывается эффективность применения декомпозиции для нахождения границ объектов. Рассматривается быстрый алгоритм корреляционного совмещения квазирегулярных изображений. Исследуется научно-техническая задача обнаружения дефектов на снимках электронных микросхем и предлагается способ ее решения.

В пятой главе исследуются проблемы сегментации изображений. Исследуется способ сегментации прослеживанием контуров объектов при сложном виде границ, разработанный применительно к задаче анализа снимков электронных микросхем. Исследуется научно-техническая задача цвето-текстурной сегментации изображений. Предлагается и исследуется метод иерархической пирамидальной сегментации изображений, основанный на оценке расстояния в объединенном пространстве яркостных, цветовых и текстурных признаков.

В шестой главе исследуются вопросы применения методов обработки изображений для анализа видеоданных на примере автоматической обработки и извлечения информации из видеопоследовательностей, формируемых компьютерным капилляроскопом. Рассматриваются вопросы устранения систематических и случайных искажений, пространственной синхронизации кадров видеопоследовательности, компенсации дрейфа, обнаружения контурных линий капилляров. Предложены алгоритмы нахождения мажоритарного капилляра, анализа морфологических характеристик контурных линий, выполнения гомеоморфного распрямляющего отображения и формирования выровненной видеопоследовательности. Корреляционный анализ последней позволяет строить карту мгновенных скоростей и определять динамические параметры кровотока.

Седьмая глава посвящена вопросам перехода от двумерных к трехмерным изображениям. Рассматриваются особенности модификации двухмасштабной многокомпонентной модели изображения и видоизменения ряда алгоритмов применительно к трехмерным изображениям. Среди них: алгоритмы частотной и пространственной фильтраций, декомпозиции, обнаружения контуров и объектов на изображении, и другие. Рассмотрены модификации алгоритмов быстрого вычисления значений арифметического среднего, дисперсии и порядковых статистик по трехмерному фрагменту в форме прямоугольного параллелепипеда. Разработан параллельный алгоритм вычисления гистограммы и порядковых статистик по скользящему фрагменту для двумерных и трехмерных изображений.

ГЛАВА 1. МОДЕЛИ ИЗОБРАЖЕНИЙ

Введение

При проведении теоретических исследований и практических разработок необходимо, чтобы изображение, как многомерный сигнал, было представлено в виде формализованного объекта, допускающего математическое описание основных своих свойств, признаков и характеристик. Формулировка такого описания составляет *математическую модель изображения*, которая позволяет абстрагироваться от физической природы получаемых изображений, а также создавать методы и алгоритмы анализа и обработки видеoinформации.

Для описания сигналов одного и того же вида могут применяться различные модели. Критериями выбора служат: назначение модели, точность представления реальных сигналов и описания их свойств, сложность и возможности использования модели и др. Имеются три классических подхода к построению математических моделей сигналов: детерминированный, стохастический и комбинированный. В первом случае сигнал, как правило, задается математической функцией вида $y(1, \dots, k) = \Phi(1, \dots, k; x_1, \dots, x_m; c_1, \dots, c_n)$, где $y(1, \dots, k)$ — значение сигнала в k -мерном пространстве, x_1, \dots, x_m — независимые аргументы, а c_1, \dots, c_n — параметры. Второй подход рассматривает сигнал как порождение случайного процесса, реализации которого приобретают конкретные значения с некоторой вероятностью. Модель такого сигнала представляет собой описание статистических характеристик случайного процесса путем задания законов распределения вероятностей, корреляционных функций и т.д. Третий подход предполагает комбинацию детерминированной и случайной составляющих.

Для описания изображений, как правило, используют модели, относящиеся ко второму или третьему типу. Модель сигнала также требует задания области определения возможных значений аргументов и функции, что в нашем случае удобно сделать на основании физических характеристик и особенностей процесса формирования изображения.

Математические модели, используемые для обработки изображений, подразделяются на два больших класса: класс непрерывных и класс дискретных моделей [336]. Непрерывные модели удобны для аналитических исследований. Они облегчают теоретические формулировки и проведение аналитических оценок, однако в действительности описывают некий идеальный объект, оперировать с которым у нас возможности нет. Несмотря на это, непрерывные модели достаточно плодотворно используются для разработки ряда методов обработки и анализа изображений, в частности, при решении обратных задач.

При использовании ЭВМ имеется возможность сохранять и обрабатывать сигналы только в дискретном виде, поэтому реальным цифровым изображениям много больше соответствуют дискретные модели. Такие модели дают простор для построения различных алгоритмов обработки изображений, но приносят трудности при теоретическом анализе.

Сказанное выше позволяет сделать вывод, что ни непрерывные, ни дискретные модели не дают возможности для решения всех проблем [245]. Учитывая особенности каждого из подходов, нами был сделан выбор в пользу дискретных моделей. Почти все исследования, приводимые ниже, основаны на дискретном представлении сигнала и дискретных моделях, однако в некоторых случаях для удобства и наглядности будут применяться также и непрерывные представления сигнала.

Анализируемое изображение является отображением произвольной сцены, которая может содержать достаточно большое число объектов. Разумно предположить, что между близко расположенными участками изображения, в большинстве случаев соответствующим одному и тому же объекту сцены, должны наблюдаться значительно более тесные корреляционные связи, чем между дальними, которые с высокой степенью вероятности соответствуют разным объектам. Поэтому во многих процедурах анализа полезно придавать больший вес свойствам участков первого типа (близким) и ослаблять влияние участков второго типа (дальних). Отсюда следует вывод о перспективности использования *локальной области анализа* — т.е. некоторой пространственно ограниченной области изображения. В зависимости от задач, размеры области анализа могут различаться в десятки и более раз. Ясно, что размер области описания модели должен быть адекватен размеру области анализа алгоритмов, разрабатываемых на ее основе.

Рассмотрение многочисленных известных моделей изображения приводит к выводу, что универсального размера или масштабного параметра, при котором можно было бы построить модель, единым и приемлемым образом описывающую необходимые свойства изображения для областей анализа как малых, так и больших размеров (масштабов), не существует. Исходя из этого был выбран подход, при котором модель изображения объединяет две различные части — модели для малого и для большого масштаба области анализа. Интегральная модель изображения, объединяющая модель окрестности (области малого размера, масштаба близлежащих элементов) и модель фрагмента (области большого размера, масштаба отдельных объектов), названа *двухмасштабной моделью изображения*.

В многочисленных работах, как например [53,90,113,138,147,270,289,347], отмечается, что основная часть информации, которую несет изображение, содержится не в уровнях яркости отдельных объектов, а в контурах или совокупности границ его областей. В связи с этим

несомненный интерес представляет построение модели, способной описывать множество контурных линий на изображении — *модели контурного изображения*.

В контурной модели изображение считается сформированным из графических примитивов — отрезков прямых линий. Особыми при этом являются точки пересечений и изломов линий — так называемые *узловые точки*. Заданием вероятностей простого набора возможных конфигураций узловых точек удастся получать топологически совершенно различные контурные изображения. Закрашиванием образуемых связных областей случайной яркостью (цветом, текстурой) можно достичь формирования двумерного массива, отвечающего всем основным статистическим свойствам обычного изображения.

В процессе развития теории обработки изображений возникает необходимость в оценке сложности изображения. Основной проблемой здесь является отыскание таких характеристик многомерных функций, по которым можно было бы отличить менее сложные функции от более сложных. Сложность изображения интерпретируется как характеристика, отражающая число, размеры и заметность деталей на изображении. В ряде работ для оценки сложности метрических пространств и многомерных аналитических функций предлагается использовать многомерные вариации [33]. Рассматриваются различные известные варианты двумерных вариаций в применении к дискретным изображениям и возможности их использования для оценки сложности. Представляется плодотворным применение для этого двумерных вариации Кронрода [75]. Рассмотрены различные известные варианты двумерных вариаций в применении к дискретным сигналам и исследованы возможности их использования для оценки сложности изображений. Проанализированы изменения значений вариаций при различных трансформациях изображения, предложена модификация оценок, названная «показатель размеров объектов».

1.1. Модель формирования изображения

1.1.1. Непрерывное и дискретное представление изображения

Представление изображения в непрерывной модели

Как уже отмечалось, изображение является отображением (проекцией) некоторой реальной или виртуальной пространственной сцены (совокупности ее объектов) на зрительную плоскость. Изображение формируется падающим электромагнитным излучением (или соответствующей моделирующей системой) в виде двумерного сигнала и несет информацию об интенсивности излучения, исходящего от объектов сцены. Двумерное монохроматическое изображение в непрерывном представлении есть действительная функция двух переменных

$F(u,v)$; $(u,v) \in D$, где носитель D — ограниченная на плоскости UV область. Согласно принципу формирования изображения, естественными ограничениями $F(u,v)$ должны быть неотрицательность и ограниченность. Очевидно, что набор всевозможных подобных функций оказывается слишком большим. В частности, он включает ряд функций, к которым многие математические операции не применимы. Чтобы устранить подобные проблемы, будем полагать, что $F(u,v)$ является «хорошей», т.е. удовлетворяет всем необходимым свойствам функций, которые потребуются для проведения тех или иных математических операций.

Считая, что функция $F(u,v)$ отражает плотность энергии электромагнитного излучения, т.е. квадрата амплитуды, можно отметить ее некоторые очевидные свойства [127]. Во-первых, $F(u,v)$ является неотрицательной, ограниченной и действительной функцией:

$$0 \leq F(u,v) \leq C < \infty. \quad (1.1.1)$$

Вторым свойством является ее интегрируемость на $D \in UV$. Энергия, как физическая величина, есть функция не одной точки, а некоторой области пространства; с другой стороны, изменение ее плотности, определяемое через производную, является функцией точки. Значит функция $F(u,v)$, которая является ограниченной и интегрируемой на замкнутом ограниченном множестве D , также является непрерывной на D всюду, кроме быть может множества меры 0.

Третье свойство вытекает из используемой физической модели. Находясь в рамках геометрической оптики, мы должны в каждой из малых областей пространства рассматривать только плоскую электромагнитную волну. Это означает, что по крайней мере на расстояниях порядка длины волны λ или разрешения прибора, амплитуда волны является постоянной или слабо меняющейся функцией [81]. Следовательно, на малых участках можно рассматривать амплитуду измеряемого излучения и изображение $F(u,v)$ как кусочно-постоянную или кусочно-гладкую функцию на D .

Таким образом показано, что изображение $F(u,v)$ является действительной функцией двух переменных $(u,v) \in D$ и удовлетворяет следующим условиям:

1. $F(u,v)$ является ограниченной: $0 \leq F(u,v) \leq C$, где C — константа;
2. $F(u,v)$ является непрерывной на D всюду, кроме, быть может, множества меры 0;
3. $F(u,v)$ является кусочно-гладкой функцией на D .

Указанных условий достаточно, чтобы формулировать многие необходимые свойства $F(u,v)$. В частности, на их основе можно показать, что $F(u,v)$ является функцией с ограниченной вариацией [245]. Эти формулируемые свойства могут быть использованы при решении многих задач обработки изображений.

В реальности изображение создается в результате некоторого физического процесса и значение $F(u,v)$ пропорционально энергии излучения, например, энергии электромагнитных

колебаний. Если изображение создается не самосветящейся сценой, то функция $F(u,v)$ задается произведением двух составляющих: освещенности $l(u,v)$ и коэффициента отражения $r(u,v)$:

$$F(u,v) \sim l(u,v)r(u,v), \quad (1.1.2)$$

где $l_{\min} \leq l(u,v) \leq l_{\max}$ и $r_{\min} \leq r(u,v) \leq r_{\max}$. Природа функции $l(u,v)$ зависит от источника освещения, тогда как функция $r(u,v)$ определяется свойствами объектов изображаемой сцены.

Представление изображения в дискретной модели

Формирование цифрового изображения осуществляется *дискретизацией* (представлением непрерывных значений конечным множеством) пространственных координат u и v , а также дискретизацией значений самой функции F , которая называется *квантованием*. В результате совместных операций пространственной дискретизации и квантования возникает двумерный набор цифровых отсчетов $F(m,n)$. Способ оцифровки, а также шаги дискретизации и квантования определяются конструкцией устройства регистрации изображения. Для определенности будем считать, что шаги дискретизации Δu и Δv постоянны и одинаковы по координатам: $\Delta u = \Delta v = \Delta$, а диапазоны изменения значений координат, обозначаемых для дискретного сигнала через m и n соответственно, составляют $0 \leq m \leq (M-1)\Delta$ и $0 \leq n \leq (N-1)\Delta$. Такой способ называется дискретизацией на квадратной решетке.

Будем полагать [86], что квантование функции F в диапазоне $[F_{\min}, F_{\max}]$ осуществляется в соответствии с законом Вебера-Фехнера:

$$(F_{k+1} - F_k)/F_k = \gamma = \text{const}, \quad \text{или} \quad F_{k+1} = F_k(1 + \gamma), \quad (1.1.3)$$

что означает формирование уровней квантования по следующему закону:

$$F_k = F_{\min}(1 + \gamma)^k.$$

Общее число уровней квантования диапазона $[F_{\min}, F_{\max}]$, также называемого *динамическим диапазоном*, при этом составит

$$K = [\log(F_{\max}/F_{\min})/\log(1 + \gamma)], \quad (1.1.4)$$

где операция $[\cdot]$ означает взятие целой части. Получаемое значение x после квантования сигнала F будет равно:

$$x(F) = [\log(F/F_{\min})/\log(1 + \gamma)]. \quad (1.1.5)$$

При этом говорят о *логарифмическом преобразовании* сигнала и *равномерном квантовании*. С учетом квантования, убирающего масштабный множитель, можно переписать (1.1.5) как

$$x_{mn} = [\log((lr)_{mn}/F_{\min})]. \quad (1.1.6)$$

Подразумевая под lr_{mn} логарифм $(lr)_{mn}/F_{\min}$, опуская знак операции взятия целой части и добавляя обычно присутствующий случайный шум ξ , получим:

$$x_{mn} = lr_{mn} + \xi_{mn}. \quad (1.1.7)$$

Именно такой сигнал считается исходным в большинстве работ, связанных с рассмотрением изображений; величину x_{mn} при этом называют просто *яркостью*. Мы также в дальнейшем будем следовать такому представлению сигнала и терминологии.

Отметим, что аддитивный в (1.1.7) шум соответствует добавлению мультипликативного шума на аналоговом уровне (1.1.2) формирования поля изображения.

В соответствии со сказанным выше, изображение в дискретном виде представляется матрицей $\mathbf{X} = [x_{mn}]$ размерами $M \times N$. Значение каждого элемента x_{mn} ($0 \leq x_{mn} \leq K - 1$) есть квантованное на K уровней значение яркости. Считается, что элементы заданы на квадратном растре в точках пересечения растровых линий (m, n) , а расстояния между ними (ρ) определены в обычном геометрическом смысле, причем единицей является шаг дискретизации:

$$\rho(x_{mn}, x_{m'n'}) = ((m - m')^2 + (n - n')^2)^{1/2}. \quad (1.1.8)$$

В случае, если \mathbf{X} — цветное (Z -зональное) изображение, значением его элемента будет являться цветовой вектор $\mathbf{x}_{mn} = \{x_{mn}^1, \dots, x_{mn}^Z\}$. В случае $Z = 1$ говорят, что изображение является одноцветным (монохромным, черно-белым). Если специально не оговорено, приводимые ниже рассуждения и выводы относятся к одноцветным изображениям. Большинство из них без особого труда могут распространяться и на цветные (многозональные) изображения. Случаи, где многоцветность изображения существенна, рассматриваются особо.

1.1.2. Области изображения и области анализа

Изображение, как средство двумерного отображения информации о некоторой сцене или предмете, можно рассматривать состоящим из отдельных связанных областей (объектов), которые соответствуют различаемым деталям сцены или предмета, и плотно покрывают изображение. Пространственной *областью* (или *объектом*) изображения называют максимальное по размеру отдельно воспринимаемое связанное множество его элементов, имеющих близкие, возможно плавно меняющиеся значения яркости (цвета). Области могут соприкасаться произвольным образом, в том числе одна область может быть полностью окружена другой. Яркости соседних областей должны заметно различаться. Не соприкасающиеся области могут иметь произвольные, в том числе и совпадающие яркости. Границы между соседними областями, как объектами различающейся яркости, называют *контурами*.

Одним из важнейших понятий в обработке изображений является область анализа. *Область анализа* V_d — подмножество исходных данных, используемое в процедуре анализа или обработки. Методы, в которых при анализе исходных данных для каждого элемента (или малых фрагментов) изображения используется своя область, как правило, центрированная в соответствующей точке, называют *локальными методами*.

Рассмотрим связное множество элементов $x_{ij} \in V_d(m, n)$, которые отстоят от центрального элемента x_{mn} на расстояние не далее *радиуса анализа* d и вместе составляют фигуру некоторой задаваемой формы. При $d \lesssim 1 \div 3$ множество $V_d(m, n)$, окружающее центральный элемент x_{mn} , будем называть *окрестностью* и обозначать V_{mn} , а при $d \gg 1$ — *фрагментом* и обозначать W_{mn} . Отметим, что в зависимости от выполняемых операций, сам центральный элемент x_{mn} может как принадлежать, так и не принадлежать $V_d(m, n)$. Соответственно, операции вида

$$y_{mn} = f\{x_{ij} \mid x_{ij} \in V_d(m, n)\}, \quad (1.1.9)$$

в которых результат в каждой точке (m, n) зависит лишь от значений элементов x_{ij} , входящих в $V_d(m, n)$, называют *локальными операциями*. Будем также говорить, что $V_d(x)$ — область порядка k , если k — максимальный из ее размеров по координатным осям.

Окрестность V_{mn} является подмножеством фрагмента W_{mn} и, как правило, расположена в его центре. Таким образом, каждый элемент изображения имеет окружающие его фрагмент и окрестность, а также попадает в окрестности и фрагменты его близлежащих элементов. В принципе, формы V_{mn} и W_{mn} могут быть произвольными, но для удобства вычислений, как правило, выбирают прямоугольные, причем чаще всего квадратные формы. Для определенности укажем приблизительные соотношения и типичные диапазоны размеров изображения, фрагмента и окрестности: $M, N \approx 100 \div 10000$; $d_w \approx 5 \div 200$; $d_v \approx 1 \div 3$ шагов дискретизации.

В ряде случаев интерес представляет не весь набор точек, попадающих в область анализа, а лишь некоторое его подмножество, включающее элемент x_{mn} . Такое подмножество будем называть *областью принадлежности* элемента x_{mn} . Способ выбора области принадлежности зависит от задачи и алгоритма, но наиболее типичным примером области принадлежности является пересечение области анализа и того объекта изображения, к которому относится элемент x_{mn} . Однако в большинстве случаев предварительного разбиения изображения на области не имеется, и тогда выбор множества элементов, относящихся к области принадлежности, может осуществляться, например, статистическими методами, базируясь на значении x_{mn} и распределении значений элементов в области анализа.

Гистограммой $H_{mn}(V_d) = \{h_{mn}(k)\}$ области анализа V_d (окрестности, фрагмента изображения) будем называть выборочное распределение вероятностей значений его элементов:

$$h_{mn}(k) = P\{x = k \mid x \in V_{mn}\}; \quad \sum_{k=0}^{K-1} h_{mn}(k) = 1. \quad (1.1.10)$$

Будем говорить, что дискретная случайная величина ξ , принимающая K значений на множестве $[0, K - 1]$, распределена нормально $N(\mu, \sigma^2)$ со средним μ и дисперсией σ^2 , если ее функция распределения во всех точках k отличается от функции распределения непрерывной нормальной случайной величины η с теми же значениями μ и σ^2 не более, чем на пренебрежимо малую ε .

1.2. Двухмасштабная многокомпонентная модель изображения

1.2.1. Подход к построению модели

Разработка методов обработки и анализа видеoinформации базируется на использовании математических моделей сигнала. Как было сказано выше, изображение отображает некоторую сцену, состоящую из множества объектов, и представляет собой набор областей разнообразной формы, яркостно-цветовых характеристик и текстуры, соответствующих отдельным объектам или их частям. Модель необходима для краткой формулировки сведений об основных свойствах изображений — наличии протяженных областей и контурных перепадов между ними. Чтобы быть полезной, с одной стороны она должна описывать статистические связи как ближних так и дальних элементов, находящихся на расстояниях, обусловленных особенностями задач, с другой — обеспечивать возможности построения достаточно быстрых алгоритмов обработки и анализа изображений.

Тематике построения математических моделей, описывающих свойства изображений, посвящено большое количество публикаций. Из них в первую очередь следует отметить обзоры [49,123] и симпозиум по моделированию изображений, труды которого изданы в [336]. Среди значительного количества известных моделей дискретного изображения распространены: авторегрессионные (каузальные), некаузальные, статистические, марковские, основанные на визуальном восприятии и др. [49, 123, 127, 300].

В [123] рассматриваются два основных класса моделей. Первый — статистические модели, описывающие совокупности нескольких элементов изображения. Из них наиболее часто используются модели состояний элементов и их групп, модели случайных полей, в том числе марковские вероятностные модели [85, 97, 300, 370], одномерные и двумерные модели линейного предсказания (авторегрессии). Данные модели определяют вероятностные связи близлежащих элементов, что удобно при описании «тонкой» структуры (текстуры) внутри однородных областей изображения. Большинство таких моделей достаточно хорошо описывает вероятностные связи элементов, находящихся внутри одной области изображения на расстоянии нескольких шагов дискретизации. Проблемы возникают при необходимости совместного описания групп элементов, составленных из представителей статистически различных множеств — внутренних и контурных элементов, или элементов из нескольких соседних областей. Это подтверждается в [84], где показана непригодность гауссовых случайных полей для моделирования большинства ансамблей реальных изображений. Второй класс — пространственные модели, описывающие разбиение изображения на составные части или области, т. е. учитывающие в основном его макроструктуру; они удобны при

морфологическом описании изображения, позволяют рассматривать его в виде совокупности отдельных протяженных областей [323] и контурных перепадов между ними.

Некоторое промежуточное положение занимают модели, рассматривающие изображение как сумму двух независимых компонент — кусочно-гладкой пространственной компоненты, определяющей яркостные изменения протяженных деталей, и детальной компоненты, задающей текстуру, шум, мелкие детали [272,371]. Рассматриваемая нами двухмасштабная многокомпонентная модель относится именно к такому классу.

Статистические связи элементов изображения, находящихся на больших расстояниях, существенно отличаются от аналогичных связей близлежащих элементов. Истоки таких отличий — в различии самих физических причин корреляционных связей таких элементов, поскольку они зависят главным образом от того, попадают ли элементы в одну и ту же область изображения, или в разные. В случае малых расстояний, когда элементы с высокой степенью вероятности попадают в одну область, определяющими являются мелкие детали, текстура и контурные перепады. В случае больших расстояний определяющими становятся различия яркостей протяженных областей изображения и переходы полутонов внутри каждой из них.

Исследования [323], в которых измерялись размеры связных областей на изображении, показали, что наиболее существенную долю изображения занимают области площадью в несколько десятков и более элементов изображения. Очевидно, что корреляция элементов внутри одной области может быть описана через межэлементные статистические связи, тогда как элементы разных областей, даже расположенные близко, слабо коррелированы. Модели, описывающие статистические связи соседних элементов, не в состоянии описать участки изображения размерами в несколько десятков элементов, поскольку не учитывают яркостные различия отдельных областей изображения. С другой стороны, модели протяженных областей неудовлетворительно описывают корреляционные связи близкорасположенных элементов.

Можно сделать вывод, что ни модель малого, ни большого масштабов по отдельности не дают достаточно полного описания свойств реальных изображений, и необходимо сочетание качеств обеих моделей. Таким образом требуется построение модели, которая бы адекватно описывала свойства изображения как на малых, так и на сравнительно больших расстояниях.

Поскольку статистические взаимосвязи элементов существенно меняются в зависимости от расстояния между ними, то очевидно, что свойства модели также должны зависеть от размеров описываемого участка изображения. Корректно решить данную проблему используя одни и те же соотношения и лишь вводя в модель масштабный параметр, изменяющий ее свойства в зависимости от расстояния, не удастся. Поэтому выбран двухмасштабный подход, объединяющий модели для областей анализа разных размеров — модель окрестности и модель

фрагмента. Такой подход хорошо согласуется с имеющимися данными о строении рецептивных полей зрительных участков коры головного мозга [308], где также присутствует анализ изображений разных масштабов.

Два масштаба — масштаб окрестности и масштаб фрагмента

Поскольку единую универсальную модель изображения создать не удастся, предлагается использовать сочетание моделей двух различных масштабов: модель окрестности, учитывающую свойства изображения на масштабе малого размера — группы из нескольких соседних элементов, и модель фрагмента, связывающую элементы изображения на масштабе большого размера и позволяющую объединять группы из элементов нескольких разных объектов. Вместе они составят единую *двухмасштабную многокомпонентную модель изображения*.

Одновременно хотелось бы, чтобы математические представления моделей окрестности и фрагмента оказались достаточно простыми, что в дальнейшем дало бы возможность разрабатывать эффективные в вычислительном отношении алгоритмы обработки изображений, способные работать в условиях сложной смеси статистически различных сигналов. В качестве отправной точки для построения такой двухмасштабной модели представляется удобным воспользоваться моделью Харалика и Ватсона [272], описывающей изображение как совокупность отдельных наклонных участков-фасетов небольших размеров.

На масштабе малого размера (*масштабе элементов окрестности*) рассматривается сравнительно небольшое связанное множество элементов, расположенных на расстоянии нескольких шагов дискретизации. Используется понятие окрестности V_{mn} элемента x_{mn} , как группы из R элементов $x_{mn}^r \in V_{mn}$, $r = 1, \dots, R$, ближайших к x_{mn} и попадающих в то же множество (контурное или фоновое), что и x_{mn} .

На масштабе большого размера (*масштабе объектов фрагмента*) рассматривается пространственно ограниченное связанное множество элементов, которое может одновременно покрывать несколько объектов. Тем самым состав элементов во фрагменте является смесью групп элементов нескольких объектов с различающимися статистическими характеристиками.

Многокомпонентное представление изображения

Типичное изображение характеризуется наличием протяженных областей (объектов) разных размеров и формы, имеющих приблизительно постоянную или плавно меняющуюся яркость, и разделенных контурными перепадами. Каждая такая протяженная область, как правило, содержит мелкие детали, текстуру, и искажена некоторым шумом. Иногда такое изображение рассматривается как некоторая двумерная поверхность, являющаяся функцией яркости в плоскости координат [113]. Протяженным объектам соответствуют гладкие и близкие

к горизонтальным области, а контурам — крутые участки, на которых угол наклона поверхности и величина яркостного перепада превышают некоторые пороговые значения.

Будем представлять изображение $x_{mn} \in \mathbf{X}$ в виде суммы статистически независимых компонент, каждая из которых характеризуется уменьшающимся масштабным параметром (размерами содержащихся в ней деталей):

$$x_{mn} = s_{mn} + u_{mn} + v_{mn} + \dots \quad (1.2.1)$$

Первый член суммы s_{mn} — кусочно-гладкая компонента, определяющая средний уровень яркости в области принадлежности элемента x_{mn} . Остальные члены несут информацию о текстуре, мелких деталях, шуме и т.д. Без существенных ограничений их можно объединить в одну текстурно-детальную компоненту: $t_{mn} = u_{mn} + v_{mn} + \dots$ и записать [170,371]:

$$x_{mn} = s_{mn} + t_{mn}. \quad (1.2.2)$$

При регистрации и оцифровке изображение подвергается воздействию помех, вызываемых зернистостью фотопленки, шумами регистратора, аналого-цифрового преобразователя и др. Поэтому в выражение (1.2.2) следует добавить шумовую компоненту ξ :

$$x_{mn} = s_{mn} + t_{mn} + \xi_{mn}. \quad (1.2.3)$$

Все компоненты полагаются независимыми и аддитивными, а t_{mn} и ξ_{mn} — нормально распределенными и несмещенными.

1.2.2. Модель окрестности (масштаб элементов)

Для элемента x_{mn} в точке (m,n) рассмотрим окружающую его окрестность V_{mn} . Перенумеруем элементы окрестности V_{mn} , обозначив индексом r положение элемента внутри нее: $x_{mn}^r \in V_{mn}$; $r = 1, \dots, R$. Расстояние между центральным элементом x_{mn} и x_{mn}^r будет обозначаться ρ^r . Методом наименьших квадратов проведем плоскость, наиболее близкую значениям элементов из V_{mn} . Данная плоскость будет составлять с горизонтальной двугранный угол величиной θ_{mn} . Тем самым каждой точке изображения можно приписать вектор \mathbf{g}_{mn} с амплитудой $g_{mn} = \text{tg } \theta_{mn}$ и углом поворота v_{mn} . Через g_{mn}^r обозначим величину проекции \mathbf{g}_{mn} на вектор из x_{mn} в x_{mn}^r . Проведенная плоскость в точке r отличается от значения x_{mn}^r на случайную величину ξ_{mn}^r . Такое представление позволяет связать значения элементов окрестности $x_{mn}^r \in V_{mn}$ при помощи формулы:

$$x_{mn}^r = \mu_{mn} + \rho^r g_{mn}^r + \xi_{mn}^r, \quad (1.2.4)$$

где μ_{mn} — значение проведенной плоскости в центральной точке окрестности (m,n) .

Проведение методом наименьших квадратов плоскости, ближайшей к элементам окрестности, несложно [272]. Пусть α и β — тангенсы углов наклона такой плоскости в

вертикальном (i) и горизонтальном (j) направлениях, а i и j — расстояния по координатным осям от центра окрестности ($-l \leq i, j \leq l$). Тогда уравнение плоскости можно записать как $z_{ij} = \mu + \alpha i + \beta j$, а формула (1.2.4) будет выглядеть:

$$x_{ij} = \mu + \alpha i + \beta j + \gamma_{ij}.$$

Величина среднеквадратического отклонения будет равна

$$\lambda(\alpha, \beta, \mu) = \sum_i \sum_j (\alpha i + \beta j + \mu - x_{ij})^2.$$

Минимум $\lambda(\alpha, \beta, \mu)$ достигается в точке, в которой все частные производные равны нулю:

$$\partial \lambda / \partial \alpha = 2 \sum \sum i (\alpha i + \beta j + \mu - x_{ij}) = 0;$$

$$\partial \lambda / \partial \beta = 2 \sum \sum j (\alpha i + \beta j + \mu - x_{ij}) = 0;$$

$$\partial \lambda / \partial \mu = 2 \sum \sum (\alpha i + \beta j + \mu - x_{ij}) = 0.$$

Для квадратной окрестности $\sum_{z=-l}^l z = 0$ и $\sum_{z=-l}^l z^2 = l(2l^2 + 3l + 1)/3$; тогда

$$\alpha = \frac{3 \sum_{i=-l}^l i \sum_{j=-l}^l x_{ij}}{l(2l^2 + 3l + 1)}; \quad \beta = \frac{3 \sum_{i=-l}^l j \sum_{j=-l}^l x_{ij}}{l(2l^2 + 3l + 1)}; \quad \mu = \frac{\sum_{i=-l}^l \sum_{j=-l}^l x_{ij}}{(2l+1)^2}. \quad (1.2.5)$$

Для окрестности 3×3 элементов $l = 1$ и получим

$$\alpha = \frac{1}{6} \left(\sum_{j=-1}^1 x_{1,j} - \sum_{j=-1}^1 x_{-1,j} \right); \quad \beta = \frac{1}{6} \left(\sum_{i=-1}^1 x_{i,1} - \sum_{i=-1}^1 x_{i,-1} \right) \quad \mu = \frac{1}{9} \sum_i \sum_j x_{ij}. \quad (1.2.6)$$

Зная α и β , параметры вектора \mathbf{g} (модуль g и угол поворота ν) находятся как:

$$g = \sqrt{\alpha^2 + \beta^2}; \quad \nu = \arctg(\beta/\alpha). \quad (1.2.7)$$

Естественно полагать, что для точек внутри протяженных областей значения g будут малыми, а для точек на границах областей — большими. Выберем некоторый порог g_0 и разделим все элементы изображения на два множества: контурные точки, для которых $g_{mn} \geq g_0$, и остальные (внутренние точки областей, фоновые), для которых $g_{mn} < g_0$. Введем понятие контурной маски $\mathbf{E} = [e_{mn}]$:

$$e_{mn} = \begin{cases} 1, & \text{если элемент } x_{mn} \text{ контурный,} \\ 0, & \text{если элемент } x_{mn} \text{ фоновый.} \end{cases}$$

Пусть для элемента изображения задана некоторая исходная форма окрестности (области анализа) A_V . Введем следующий способ модификации формы окрестности. В области анализа A_V элемента x_{mn} оставим лишь те элементы, которые относятся к тому же множеству (контурному или фоновому), что и x_{mn} ; т.е. для которых выполняется: $e_{mn}^r = e_{mn}$; $r = 1, \dots, R$. Также отбросим элементы, отделенные от x_{mn} проходящим между ними контуром, если таковой имеется. Оставшиеся элементы будут составлять связное множество, не выходящее за границы A_V . Его будем считать окрестностью V_{mn} элемента x_{mn} .

Рассмотрим модуль разностей значений соседних элементов из окрестности V_{mn} :

$$P(d) = P\{|x_{m,n} - x_{m,n-1}|\} \quad (1.2.8)$$

и построим распределение вероятностей $P(d)$ для реального изображения. Получаемая зависимость для всех точек изображения показана кривой 1 на Рис. 1.1.1; ее особенностью является наличие пика в области малых d . Построим также распределения $P(d)$ отдельно для фоновых $P_{\Phi}(d)$ и контурных $P_{\text{К}}(d)$ участков изображения. Такие зависимости показаны кривыми 2 и 3. Два последних графика нетрудно приблизить кривыми нормального распределения (4 и 5), которые в используемом полулогарифмическом масштабе выглядят как параболы вида $ax^2 + c$ ($a < 0$) в линейной системе координат. Таким образом множество значений d можно интерпретировать как смесь двух нормально распределенных случайных величин, а множество точек изображения, описываемых кривой 1, — как объединение двух множеств точек: фоновых и контурных, в каждом из которых распределения $P(d)$ близки нормальному.

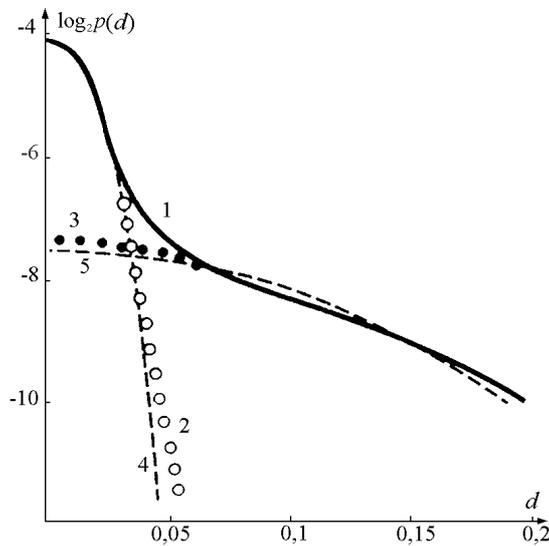


Рис. 1.1.1. Распределение вероятностей модуля разности $P(d)$ (1.2.8): 1) для всего изображения; 2) для внутренних точек областей; 3) для контурных точек; 4) и 5) кривые нормальных распределений с разными дисперсиями.

Проводилась экспериментальная проверка изложенной двухкомпонентной гипотезы на реальных изображениях, представленных на Рис. 1.1.2. Исследования показали, что для каждого из изображений вид распределений $P_{\text{К}}(d)$ и $P_{\Phi}(d)$ близок к нормальному, но дисперсии (σ^2) их резко отличаются: $\sigma_{\text{К}}/\sigma_{\Phi} \approx 3 \div 8$; значения параметров асимметрии и эксцесса находятся в диапазонах $0,05 \div 0,2$ и $0,1 \div 0,3$ соответственно. Причиной несколько завышенных значений последних параметров по-видимому является недостаточно точное разделение множества элементов изображения на контурные и фоновые, что само по себе является нетривиальным.

Таким образом, фоновую и контурную составляющие можно рассматривать как компоненты, распределенные на непересекающихся множествах (фоновом и контурном), при объединении дающие полное изображение. Обозначим значение проекции вектора \mathbf{g}_{mn} на направление из центра окрестности в точку r для окрестностей контурных и фоновых

элементов через φ_{mn}^r и ψ_{mn}^r соответственно. Используя контурную маску e_{mn} и представляя g_{mn}^r и ξ_{mn}^r в виде сумм $g_{mn}^r = e_{mn}^r \varphi_{mn}^r + (1 - e_{mn}^r) \psi_{mn}^r$ и $\xi_{mn}^r = e_{mn}^r \chi_{mn}^r + (1 - e_{mn}^r) \eta_{mn}^r$, получим окончательную формулу модели окрестности [169,170], описывающую статистические связи ее элементов:

$$x_{mn}^r = \mu_{mn} + e_{mn}^r (\varphi_{mn}^r \rho^r + \chi_{mn}^r) + (1 - e_{mn}^r) (\psi_{mn}^r \rho^r + \eta_{mn}^r). \quad (1.2.9)$$

Здесь χ_{mn}^r — стохастическое возбуждение в точке r для окрестности контурных элементов, а η_{mn}^r — для фоновой окрестности. Отметим, что то или иное значение контурной маски e_{mn} по сути определяет принадлежность центрального элемента окрестности x_{mn} контурному или фоновому множеству. Случайные величины φ_{mn} , ψ_{mn} , χ_{mn} , и η_{mn} в (1.2.9) считаются некоррелированными и несмещенными; шумовые составляющие χ_{mn} , и η_{mn} кроме того будем считать нормально распределенными: $N(0, \sigma_\chi^2)$ и $N(0, \sigma_\eta^2)$.



Рис. 1.1.2. Тестовые изображения и их нумерация ($0 \div 15$)

Статистические связи элементов окрестности

Определим дисперсии и ковариации разностей соседних элементов внутри одной и той же компоненты — фоновой или контурной. Рассмотрим окрестность 3×3 элементов и возьмем три ее элемента: центральный x_0 и два из восьми оставшихся: x_i и x_j . Согласно модели (1.2.9) их значения связаны следующими соотношениями (для простоты опустим лишние индексы):

$$x_0 = \mu + \xi_0; \quad x_i = \mu + a_i \rho_i + \xi_i.$$

Обозначим через α_{ij} угол между векторами $\overrightarrow{x_0, x_i}$ и $\overrightarrow{x_0, x_j}$. Пусть по элементам окрестности методом наименьших квадратов проведена плоскость и задан вектор \mathbf{g} , и пусть β — угол между векторами \mathbf{g} и $\overrightarrow{x_0, x_i}$. Можно показать, что тогда

$$a_i = g \cos \beta; \quad a_j = g \cos (\alpha_{ij} + \beta). \quad (1.2.10)$$

Разность элементов окрестности будет равна

$$x_0 - x_i = \xi_0 - \xi_i - a_i \rho_i$$

со средним $M\{x_0 - x_i\} = 0$ и дисперсией $M\{(x_0 - x_i)^2\} = 2M\{\xi^2\} + \rho_i^2 M\{a_i^2\}$.

Учитывая (1.2.10), а также что $M\{\cos^2 \beta\} = 1/2$, получим:

$$M\{(x_0 - x_i)^2\} = 2M\{\xi^2\} + (\rho_i^2/2)M\{g^2\}. \quad (1.2.11)$$

Ковариация разностей $x_0 - x_i$ и $x_0 - x_j$ составит величину

$$\text{cov}\{(x_0 - x_i), (x_0 - x_j)\} = M\{\xi^2\} + \rho_i \rho_j M\{a_i a_j\}. \quad (1.2.12)$$

На основании (1.2.10) можно записать:

$$M\{a_i a_j\} = M\{g^2 \cos \beta \cos(\alpha + \beta)\} = (1/2) \cos \alpha M\{g^2\}. \quad (1.2.13)$$

Подставляя (1.2.13) в (1.2.12), получим

$$\text{cov}\{(x_0 - x_i), (x_0 - x_j)\} = M\{\xi^2\} + (\rho_i \rho_j / 2) \cos \alpha_{ij} M\{g^2\}. \quad (1.2.14)$$

Для удобства сократим запись ковариации $\text{cov}\{(x_0 - x_i), (x_0 - x_j)\}$, обозначив разности стрелками, совпадающими с направлениями векторов. Так, ковариация (1.2.14) векторов направлений 0 и $\pi/2$ обозначится как $\text{cov}(\uparrow \rightarrow)$. В основании векторов находится элемент x_0 , а в вершинах — x_i и x_j . Для окрестности 3×3 элементов квантом поворота векторов является угол $\pi/4$, так что положение векторов ясно из рисунка. Вообще, ниже вектором обозначается разность пары элементов: элемент в основании минус элемент в вершине. Для наиболее важных частных случаев, подставляя в (1.2.14) конкретные значения α_{ij} , а также ρ_i и ρ_j , получим следующие соотношения:

$$1. \alpha_{ij} = \pi/2: \quad (\uparrow \rightarrow \downarrow \leftarrow \nearrow \nwarrow \searrow \swarrow \nwarrow \searrow) \\ \text{cov}(\uparrow \rightarrow) = \dots = \text{cov}(\nwarrow \searrow) = M\{\xi^2\}. \quad (1.2.15)$$

$$2. \alpha_{ij} = \pi; \quad \rho_i = \rho_j = 1: \quad (\leftrightarrow \updownarrow) \\ \text{cov}(\leftrightarrow) = \text{cov}(\updownarrow) = M\{\xi^2\} - M\{g^2\}/2. \quad (1.2.16)$$

$$3. \alpha_{ij} = \pi; \quad \rho_i = \rho_j = \sqrt{2}: \quad (\nearrow \nwarrow) \\ \text{cov}(\nearrow) = \text{cov}(\nwarrow) = M\{\xi^2\} - M\{g^2\}. \quad (1.2.17)$$

$$4. \alpha_{ij} = \pi/4; \quad \rho_i = 1, \quad \rho_j = \sqrt{2}: \quad (\swarrow \nwarrow \nearrow \nwarrow \searrow \swarrow \nwarrow \searrow) \\ \text{cov}(\swarrow) = \dots = \text{cov}(\nwarrow) = M\{\xi^2\} + M\{g^2\}/2. \quad (1.2.18)$$

Значения величин $M\{\xi^2\}$ и $M\{g^2\}$ нетрудно получить, зная дисперсии разностей пар элементов, подсчитанные, например, по направлениям (\uparrow) и (\nearrow). Согласно (1.2.11):

$$\begin{aligned} M\{(\uparrow)^2\} &= 2M\{\xi^2\} + M\{g^2\}/2, \\ M\{(\nearrow)^2\} &= 2M\{\xi^2\} + M\{g^2\} \end{aligned} \quad (1.2.19)$$

Отсюда:

$$\begin{aligned} M\{\xi^2\} &= M\{(\uparrow)^2\} - M\{(\nearrow)^2\}/2, \\ M\{g^2\} &= 2(M\{(\nearrow)^2\} - M\{(\uparrow)^2\}). \end{aligned} \quad (1.2.20)$$

Некоторые из полученных соотношений представляют отдельный интерес. Так, измерив ковариацию $\text{cov}(\uparrow \rightarrow)$ по фоновой компоненте (1.2.15), можно оценить величину $M\{\xi^2\}$, т.е. дисперсию шума на изображении. Среднеквадратическое отклонение (СКО) шума $\sigma = (M\{\xi^2\})^{1/2}$ будет многократно использоваться в дальнейшем. Измерение дисперсии $M\{g^2\}$ по контурной компоненте (1.2.20) позволяет оценить степень дискретизации изображения: на изображении высокого пространственного разрешения значение $M\{g^2\}$ должно быть сравнительно мало, а на изображении низкого разрешения — велико.

В качестве проверки согласованности модели окрестности с реальными изображениями, представляет интерес измерение на типичных изображениях дисперсий разностей пар соседних элементов $M\{(\uparrow)^2\}$ и $M\{(\nearrow)^2\}$, определение параметров $M\{\xi^2\}$ и $M\{g^2\}$ и последующее сравнение значений ковариаций, подсчитанных по формулам (1.2.15)–(1.2.18), со значениями, измеренными экспериментально. Такие данные, измеренные отдельно по фоновым и контурным компонентам ряда изображений (Рис. 1.1.2), приведены в Таблице 1.1.1.

Из полученных данных видно, что наблюдается значительное различие статистических свойств элементов на контурных и фоновых участках. Среднеквадратические отклонения случайных величин g_K и g_F (тангенсов углов наклона плоскости, проводимой по элементам окрестностей на контурных и фоновых участках реальных изображений) различаются в $6 \div 25$ раз. Сравнение показывает, что наблюдается достаточно хорошее совпадение теоретических и экспериментальных значений ковариаций. Определенные различия, по-видимому, вызваны следующими факторами: а) ошибками в разделении изображения на множество контурных и фоновых элементов; б) погрешностями в измерении значений дисперсий разностей пар элементов по направлениям (\uparrow) и (\nearrow); в) помехами при оцифровке изображения, приводящими к возможным систематическим отклонениям средней яркости строки от истинного значения.

Таблица 1.1.1. Среднеквадратические отклонения (СКО) разностей по направлениям (\uparrow) и (\nearrow), подсчитанные по ним значения СКО для ξ и g , а также теоретические (K_T) и экспериментальные (K_ε) значения ковариаций разностей пар соседних элементов (отдельно по контурным и фоновым элементам). Нумерация изображений соответствует Рис. 1.1.2.

Контурная компонента:

№ изображения:	0	2	4	6	12	13	14
СКО (\uparrow)	14,39	5,08	6,61	5,54	7,60	10,70	11,15
СКО (\nearrow)	18,18	7,11	8,65	7,21	10,04	13,40	14,25
СКО ξ_k	9,35	1,03	2,57	2,16	2,61	4,97	6,27
СКО g_k	15,70	7,02	7,91	6,02	9,3	11,41	14,10
K_T \updownarrow	-39,6	-23,6	-24,8	-16,6	-35,6	-40,3	-39,4
K_ε \updownarrow	-33,7	-19,6	-24,3	-13,3	-20,5	-31,7	-51,9
K_T \nearrow	-163,0	-48,4	-55,9	-37,8	-78,7	-105,4	-118,1
K_ε \nearrow	-119,5	-33,4	-44,5	-21,3	-45,2	-78,0	-90,4
K_T \curvearrowright	83,8	1,07	6,28	4,70	7,36	24,7	39,4
K_ε \curvearrowright	65,1	1,06	5,38	4,77	7,31	21,6	31,1
K_T \curvearrowleft	83,8	1,07	6,28	4,7	7,36	24,7	39,4
K_ε \curvearrowleft	120,9	1,96	10,39	6,11	16,3	40,8	39,1
K_T \nearrow	207,1	25,8	37,4	26,0	50,4	89,8	119,1
K_ε \nearrow	175,2	25,5	37,9	24,2	50,6	102,9	117,2

Фоновая компонента:

№ изображения:	0	2	4	6	12	13	14
СКО (\uparrow)	1,45	1,02	2,63	1,77	2,50	3,29	2,23
СКО (\nearrow)	1,52	1,18	2,69	1,87	2,64	3,53	2,41
СКО ξ_Φ	0,98	0,59	1,82	1,17	1,66	2,14	1,44
СКО g_Φ	0,65	0,56	0,80	0,86	1,18	1,84	1,30
K_T \updownarrow	0,74	-0,02	2,98	1,01	2,07	3,05	1,23
K_ε \updownarrow	0,78	-0,18	2,59	1,04	2,03	3,14	1,08
K_T \nearrow	0,53	-0,37	2,66	0,64	1,37	1,31	0,39
K_ε \nearrow	0,77	-0,32	1,74	0,74	1,66	3,42	1,04
K_T \curvearrowright	0,95	0,34	3,30	1,38	2,77	4,59	2,07
K_ε \curvearrowright	0,82	0,21	3,45	1,18	2,74	4,58	1,85
K_T \curvearrowleft	0,95	0,34	3,30	1,38	2,77	4,59	2,07
K_ε \curvearrowleft	1,00	0,35	2,82	1,37	2,73	4,90	2,29
K_T \nearrow	1,16	0,69	3,62	1,75	3,47	6,23	2,91
K_ε \nearrow	1,28	0,67	3,63	1,50	3,47	6,24	3,33

1.2.3. Модель фрагмента (масштаб объектов)

Рассмотренная выше модель окрестности статистически связывает значения элементов изображения, находящихся на расстояниях не более нескольких шагов дискретизации. Однако во многих задачах обработки изображений такие размеры анализируемого участка оказываются недостаточными и приходится использовать фрагменты гораздо больших размеров — до нескольких десятков элементов. В такой ситуации модель окрестности неприменима и требуется разработка соответствующей модели фрагмента.

Традиционный подход к описанию изображения, используемый в задачах сегментации, распознавания, анализа сцен и ряде других, основывается на представлении изображения как множества непересекающихся связных областей, плотно покрывающих двумерное пространство, на котором оно задано. Поэтому на макроуровне будем рассматривать изображение как объединение достаточно гладких связных областей, разделенных контурными границами. Каждая из таких областей обладает какими-то свойствами, отличающими ее от соседних. Согласно исследованию [323], значительную долю площади изображения занимают протяженные области, содержащие от 10^2 до 10^4 и более элементов каждая. Сопоставляя размеры фрагмента ($d_w \approx 5 \div 200$) и областей на реальном изображении, очевидно, что произвольно взятый фрагмент может как попадать целиком внутрь некоторой большой области, так и покрывать две или более независимых областей, включая их части, или даже захватывая небольшие области целиком. Возможный вариант расположения областей во фрагменте показан на Рис. 1.1.3,а. Фрагмент находится на границе двух различных протяженных областей (U^2 и U^3) и целиком захватывает третью область (U^1).

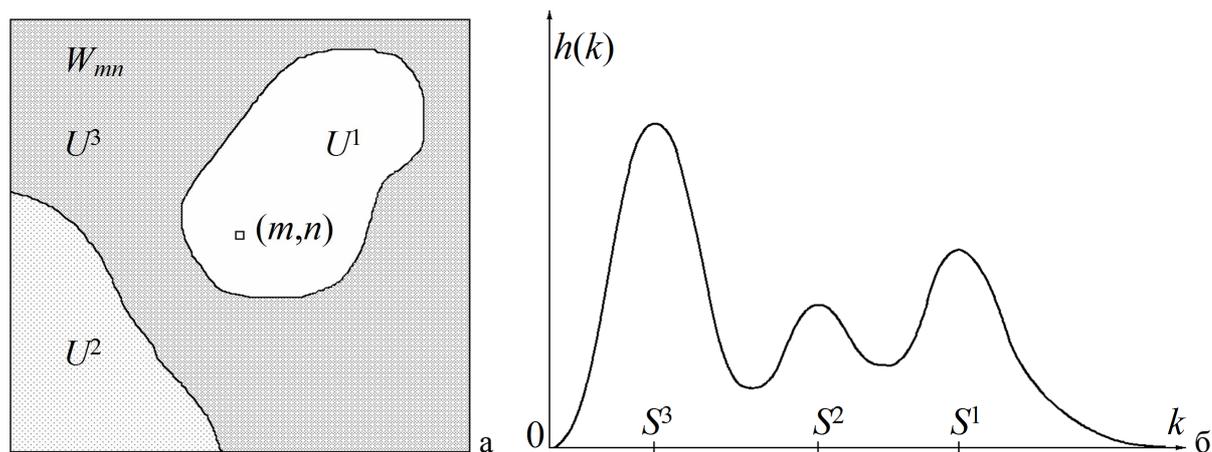


Рис. 1.1.3. Фрагмент, содержащий части трех областей (а), и его гистограмма (б).

Очевидно, элементы, относящиеся к одной и той же области, будут статистически зависимы не только на малых расстояниях (внутри окрестности), но и на достаточно больших — по крайней мере в пределах размеров фрагмента. В то же время средние яркости различных областей и, соответственно, значения входящих в них элементов являются статистически независимыми, пусть даже расстояния между ними невелики.

Рассмотрим фрагмент W_{mn} размерами $(2L+1) \times (2L+1)$, захватывающий R различных областей изображения: U^1, \dots, U^R . Обозначим через x'_{ij} элемент области U^r фрагмента, имеющий координаты (i,j) внутри фрагмента ($-L \leq i, j \leq L$). Таким образом для анализа доступны лишь те элементы и части областей, которые оказываются внутри границ фрагмента.

Основываясь на представлении изображения в виде суммы трех компонент (1.2.3) — гладкой s_{ij} , текстурной t_{ij} и шумовой ξ_{ij} , предполагается, что гладкие составляющие тех частей R областей, которые попадают во фрагмент, могут быть представлены полиномом степени не выше ω . Тогда составляющая s_{ij} внутри фрагмента W_{mn} может быть описана формулой

$$s_{ij}^r = \sum_{U=1}^R \delta_{U^r} \sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^r i^{p-q} j^q ; \quad (1.2.21)$$

здесь (i, j) — точка фрагмента W_{mn} ; δ — индикаторная функция множества: $\delta_{U^r} = 1$, если точка $(i, j) \in U^r$ и $\delta_{U^r} = 0$ в остальных случаях. Вводя в (1.2.21) текстурную t_{mn} и шумовую ξ_{mn} составляющие, получим выражение для значения элемента изображения:

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} \left(\sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^r i^{p-q} j^q + t_{ij}^r + \xi_{ij} \right) \quad (1.2.22)$$

Это основная формула модели фрагмента, которая описывает внутренние элементы областей. Для элементов, попадающих на контурные перепады, такой подход не применим и для них следует использовать модель окрестности (1.2.9).

На большинстве реальных изображений протяженные области имеют приблизительно постоянные средние локальные яркости, заметно не изменяющиеся в пределах типичного окна анализа, что подтверждается проведенными экспериментами [170]. По этой причине, а также имея в виду требование вычислительной эффективности алгоритмов обработки, выберем кусочно-постоянную модель с минимальной степенью полинома: $\omega = 0$. Тогда $s_{ij}^r(W_{mn}) = s_{mn}^r$ и (1.2.22) преобразуется к виду

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} (s_{mn}^r + t_{ij}^r + \xi_{ij}). \quad (1.2.23)$$

Мы получили формулу кусочно-постоянной модели фрагмента для представления участков областей изображения, попадающих во фрагмент W_{mn} . Отметим, что при смещении фрагмента изменятся и доли областей, которые покрываются фрагментом частично, а значит, вообще говоря $s_{mn}^r \neq s_{m+\Delta m, n+\Delta n}^r$.

Модель (1.2.23) подводит нас к использованию гистограммы фрагмента как основного его статистического описания. Случайные величины t и ξ можно считать независимыми, несмещенными и нормально распределенными: $N(0, \sigma_t^2)$ и $N(0, \sigma_\xi^2)$. Тогда их сумма $\tau = t + \xi$ также распределена нормально: $N(0, \sigma_\tau^2)$, причем $\sigma_\tau^2 = \sigma_t^2 + \sigma_\xi^2$, и (1.2.23) упростится до

$$x_{ij}^r = \sum_{U=1}^R \delta_{U^r} (s_{mn}^r + \tau_{ij}^r). \quad (1.2.24)$$

Эта модель проста, достаточно близка большинству реальных изображений и приемлема для построения на ее основе различных алгоритмов фильтрации, в том числе ранговых. Если

фрагмент попадает внутрь одной протяженной области изображения, то его гистограмма будет одномодовой со средним s_{mn}^1 и дисперсией σ_r^2 . Если же во фрагмент попадают участки двух или более (R) областей, то число мод (локальных максимумов) гистограммы может достигать R . Положения мод определяются средними значениями яркостей s^1, s^2, \dots, s^R участков областей U^1, U^2, \dots, U^R , попадающих во фрагмент, а мощности мод — долей площади, занимаемой каждой областью во фрагменте (Рис. 1.1.3,б). Заполнение промежутков между модами зависит от дисперсий мод σ_r^2 (вообще говоря, различных), расстояний между ними, определяющих наложения «хвостов» распределений, и долей контурных элементов в общем распределении.

Связь параметров модели с характеристиками изображения

При оценке изображения, кроме визуального качества — неформального, но достаточно часто используемого субъективного показателя, используются такие критерии как четкость, контрастность, уровень шума (помех). Данные характеристики непосредственно связаны с моделью окрестности (1.2.9) и ее параметрами: $\sigma_\phi, \sigma_\psi, \sigma_\chi, \sigma_\eta$.

Четкость изображения определяет максимально возможное число деталей, которые можно различить на изображении, т.е. фактически крутизну контурных перепадов, которая задается параметром σ_ϕ : чем более четким является изображение, тем выше значение σ_ϕ .

Контрастностью изображения называется отношение яркостей самого светлого и самого темного его участков. В начале главы было сказано, что значения элементов изображения суть логарифмы яркостей, поэтому в нашем случае контрастностью является величина $(s_{\max} - s_{\min})$, определяющая динамический диапазон значений яркостей областей. Основываясь на известных свойствах зрительной системы человека, в частности частотно-контрастной характеристике и размерах фовеальной области [247], а также модели зрительной системы [265], можно сделать вывод, что визуальная контрастность изображения оценивается локально — на участках размерами порядка $1/20 \div 1/10$ линейных размеров изображения. Таким образом можно предложить следующую оценку визуальной контрастности изображения:

$$C = \sum_{W \subset X} (s_{\max}^W - s_{\min}^W) / N_W. \quad (1.2.25)$$

Здесь $W \subset X$ означает, что суммирование ведется по всем возможным фрагментам W выбранных размеров, $(s_{\max}^W - s_{\min}^W)$ — максимальная разность значений яркостей областей по каждому из фрагментов, N_W — общее число расположений таких фрагментов на изображении.

Про связь уровня шума со значением σ_η уже было сказано выше. Уровень шума на изображении в основном определяет величину параметра σ_η и значительно меньше влияет на σ_ψ . Параметры σ_ϕ и σ_χ , которые по абсолютной величине значительно больше, чем σ_ψ и σ_η ,

сравнительно мало зависят от величины шума. Это довольно естественно, поскольку заметность шума на контурных участках существенно ниже, чем на гладких фоновых участках.

Параметры модели зависят также от частоты отсчетов при дискретизации изображения. Уменьшение шага дискретизации приводит к увеличению σ_χ и σ_η из-за возрастания шумов зернистости и ухудшения соотношения сигнал/шум аппаратуры ввода (меньше энергии при измерении яркости). Параметры σ_ϕ и σ_ψ сами по себе не должны при этом меняться, но члены $\phi^r \rho^r$ и $\psi^r \rho^r$ в уравнении модели (1.2.9) уменьшаются за счет уменьшения ρ^r . Следует также отметить, что частота дискретизации и четкость изображения являются взаимосвязанными (обратной зависимостью) величинами.

Сравним распределения вероятностей разностей соседних элементов $d_{mn} = |x_{m,n} - x_{m,n-1}|$ для одного и того же изображения, но оцифрованного с различными шагами дискретизации (25, 50, 100 и 200 мк.). Зависимости показаны на Рис. 1.1.4,а.

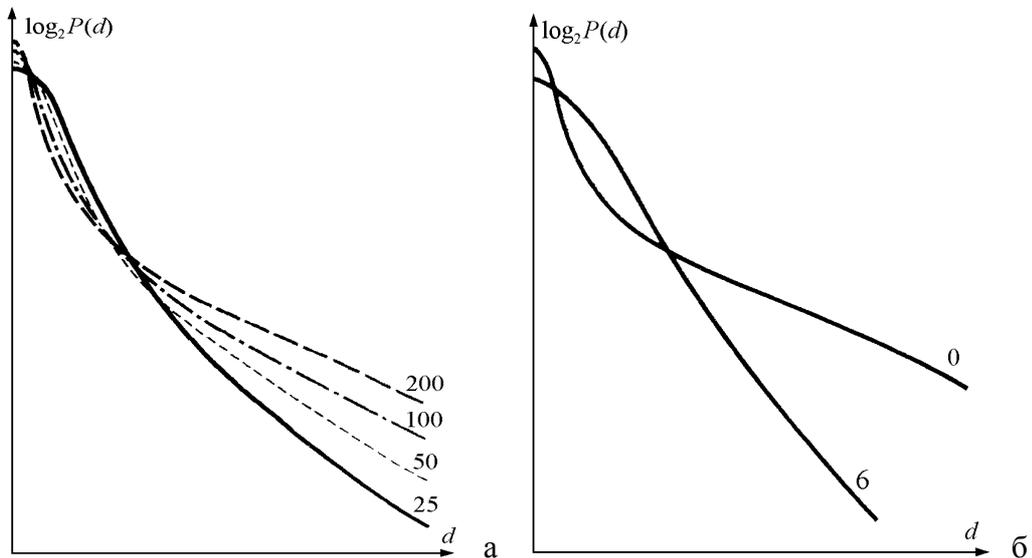


Рис. 1.1.4. Распределение вероятностей модуля разности соседних элементов изображения d :
 а) для одного изображения, оцифрованного с различной дискретизацией (25, 50, 100, 200 мк);
 б) для различных изображений (0 и 6 на Рис. 1.1.2)

Значения малых d соответствуют фоновым, а больших d — контурным участкам изображения. Видно, что с увеличением шага дискретизации значения $P(d)$ при больших d (контур) возрастают, что вызвано увеличением межэлементного расстояния ρ^r , а при малых d (фон) уменьшается за счет сглаживания и уменьшения шума. Распределения вероятностей модуля разностей соседних элементов для различных изображений (0 и 6 на Рис. 1.1.2) показаны на Рис. 1.1.4,б. Как видно, изображение 0 характеризуется меньшим значением σ_ψ и большим значением σ_ϕ , чем изображение 6. Это означает, что изображение 0 имеет четкость выше и вероятно оно допускает более высокое пространственное разрешение при дискретизации; тогда как изображение 6 оцифровано с «переразрешением». Измерить значения параметров σ_ϕ , σ_ψ , σ_χ и σ_η на реальном изображении, возможно по формулам (1.2.15) или (1.2.20).

1.2.4. Модель цветного изображения

Цветное или многоканальное изображение \mathbf{X} представляется двумерным массивом векторов $\mathbf{x}_{mn} = \{x_{mn}^1, \dots, x_{mn}^Z\}$ размерами $M \times N$. Z означает число *каналов* изображения (в некоторых случаях называемых также или *компонентами* или *зонами*). В случае цветного RGB изображения $Z = 3$ и, как правило, x^1 соответствует красному, x^2 – зеленому, а x^3 – синему каналам. Часто все изображения с $Z \geq 2$ обобщают названием *цветные*. Количество каналов обычно ясно из контекста, поэтому путаницы не возникает.

Предполагается, что каждый из каналов цветного изображения отвечает всем свойствам одноканального (полутонного, черно-белого) изображения. При этом как статистические, так и информационные свойства каналов могут быть различными. Соответственно, различными могут являться и результаты, получаемые при анализе каждого из каналов по отдельности (например, обнаруживаемые контурные участки и области).

По аналогии с (1.2.3) цветное изображение \mathbf{X} может быть представлено в виде

$$\mathbf{x}_{mn} = \mathbf{s}_{mn} + \mathbf{t}_{mn} + \boldsymbol{\xi}_{mn}, \quad (1.2.26)$$

где \mathbf{x}_{mn} , \mathbf{s}_{mn} , \mathbf{t}_{mn} и $\boldsymbol{\xi}_{mn}$ суть Z -мерные векторы. В векторном виде можно записать и многие другие выражения модели изображения, так, аналогично (1.2.24) можно записать выражение для модели фрагмента цветного изображения:

$$\mathbf{x}_{ij}^r = \sum_{U=1}^R \delta_{U^r} (\mathbf{s}_{mn}^r + \boldsymbol{\tau}_{ij}^r). \quad (1.2.27)$$

Здесь так же, как и выше, \mathbf{x}_{ij}^r — векторное значение элемента (i, j) фрагмента с центром в (m, n) , \mathbf{s}_{mn}^r — среднее значение части области r во фрагменте, которой принадлежит элемент \mathbf{x}_{ij}^r , а $\boldsymbol{\tau}_{ij}^r$ — вектор, определяющий значение текстуры и шума в точке (i, j) .

Во многих случаях преобразования цветного изображения могут осуществляться по каждому из каналов по отдельности. В тех же случаях, когда возможен или необходим вариант совместного преобразования или анализа, это будет рассматриваться особо.

Изложенные в разделах 1.1–1.2 результаты опубликованы в работах [169, 170, 171, 179, 235, 236].

1.3. Вероятностная модель контурного изображения

Как уже говорилось выше, изображение можно представить в виде множества связанных областей, плотно покрывающее некоторое ограниченное двумерное пространство. Каждая из областей обладает какими-то свойствами, отличающими ее от соседних; контурами при этом называют границы между областями [49, 113, 123]. Можно сказать, что именно контуры

(границы областей) определяют форму объектов [53,90,147]. В данном разделе мы не будем касаться свойств элементов внутри областей, а остановимся только на рассмотрении картины контуров. Для упрощения терминологии, под словом *изображение* здесь будем понимать *контурное изображение*, представляющее собой двумерный бинарный сигнал $x(u,v)$, принимающий значения 1 в контурных и 0 в остальных точках. Иногда такое изображение называют также *рисунком* [85,138].

Возможность описания произвольного контурного препарата при помощи отрезков прямых линий, формируемых при обработке реального изображения, обсуждается в работах [40,53], где предлагается приближать кривую линию последовательностью линейных сегментов. Для этого на кривой следует отметить точки высокой кривизны, которые затем соединить прямыми линиями. Приемлемость такого подхода подтверждается и работой [214], в которой приводится пример построения изображения согласно указанному правилу. Данные работы подтверждают приемлемость используемого ниже подхода для построения модели контурного изображения.

В работах [32,85,125,298] рассмотрены модели, учитывающие контуры лишь двух направлений — горизонтальные и вертикальные. Хотя и утверждается, что любой контур можно сколь угодно точно приблизить отрезками таких линий, тем не менее свойства изображения, описываемого вертикальными и горизонтальными линиями, кардинально отличаются от свойств изображения, описываемого линиями произвольного направления.

1.3.1. Общая постановка задачи

Задача состоит в разработке вероятностной модели контурного изображения и способа его построения [166]. Под контурным понимается изображение со следующими свойствами:

- контуры состоят из отрезков прямых линий произвольных направлений, которые могут иметь изломы и ветвления, но не имеют обрывов (каждая линия начинается и заканчивается либо соединением с какими-то другими линиями, либо на границах изображения);
- изображение является однородным и изотропным, т.е. статистические свойства не меняются в зависимости от выбранной точки или направления.

Также потребуем, чтобы для задания характеристик изображения использовалось минимальное число параметров.

Варианты конфигураций соединений контурных линий показаны на Рис. 1.2.1. Хотя некоторые из приведенных конфигураций и похожи — являются зеркальными по отношению друг к другу или же результатом поворота (например, 2-5 или 6 и 7), тем не менее для последующей разработки алгоритма построения контурного изображения их необходимо

разделить. Точку, в которой происходит излом (2-5) или стык (6-9) нескольких линий будем называть узловой точкой (узлом). Назовем узлы, из которых исходят две, три или четыре линии, узлами 'V', 'Y' и 'X' соответственно.

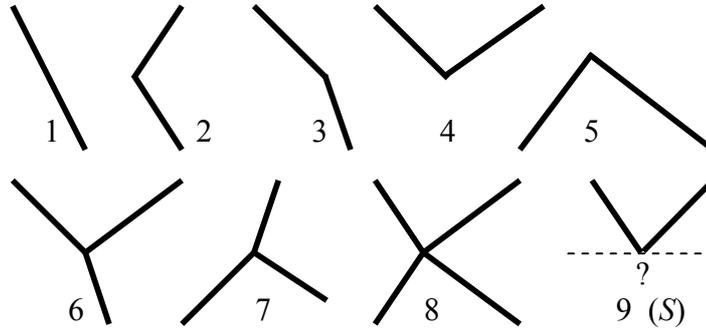


Рис. 1.2.1. Конфигурация узлов при изломах и ветвлениях линий: 1) линия без излома; 2-5) изломы; 2, 3) узлы '<'; 4) узел 'V'; 5) узел 'Λ'; 6) узел 'Y'; 7) узел 'λ'; 8) узел 'X'; 9) узел 'S' (промежуточный), когда имеется стык линий сверху, а нижние линии еще не определены.

В принципе допустимо существование и более сложных узлов, из которых исходят пять или большее число линий, однако можно считать, что вероятности таких узлов ничтожно малы и что они представимы в виде объединения нескольких близко расположенных узлов типа 'Y' или 'X', поэтому ограничимся представленными конфигурациями. Аналогичный набор возможных конфигураций узлов, несколько расширенный за счет рассмотрения частных случаев, например узлов типа 'T', 'K' и некоторых других, приводится в работах [49,136,138,307] при решении задач анализа сцены по контурным рисункам.

Особенности дискретного представления

Требования к модели были сформулированы выше в достаточно общем виде. Необходимо дискретное описание, без которого невозможна алгоритмизация источника контурных изображений с заданными свойствами. Будем полагать, что дискретизация осуществляется на квадратной решетке размерами $M \times N$, в узлах которой находятся элементы изображения X : $x_{mn} \in \{0,1\}$, $0 \leq m < M$, $0 \leq n < N$. (В действительности же способ дискретизации не существен и выбор квадратной решетки обусловлен лишь соображениями удобства и традиции; так, можно было бы рассматривать дискретизацию, например, на гексагональной решетке).

Введем понятие ε -окрестности (ε_{mn}) растровой точки (m,n) как области из непрерывного пространственного множества (u,v) , которая окружает точку (m,n) . Будем считать, что размеры ε -окрестностей ограничены $x(u,v) \in \varepsilon_{mn}$, $(m-u)^2 + (n-v)^2 \leq \varepsilon^2$, что множество ε -окрестностей точек дискретного раstra является полным на (u,v) , и что пересечения ε -окрестностей растровых точек есть множество меры ноль.

Припишем события, происходящие в ε_{mn} , самой точке (m,n) . Таким образом, если через ε_{mn} проходит линия или в ε_{mn} находится узел, то будем говорить, что линия проходит через точку

(m,n) , или что узел находится в точке (m,n) . Для квадратного раstra с шагом h естественной ε -окрестностью точки (m,n) становится квадрат, который определяется как совокупность точек $x(u,v)$, удовлетворяющих условию: $m - h/2 \leq u < m + h/2$, $n - h/2 \leq v < n + h/2$. Наклонная линия в этом случае отобразится ступенчато в виде серии горизонтальных или вертикальных сегментов.

Рассмотрим дискретное представление прямой линии, проходящей под углом α к вертикали (Рис. 1.2.2). Предполагая формирование изображения сверху вниз, считаем, что угол α ограничен диапазоном $-\pi/2 \leq \alpha \leq \pi/2$. Рассмотрим строку изображения m , пересекаемую линией. Линия заметает на этой строке горизонтальный отрезок в несколько элементов, который назовем *сегментом*. Первый элемент сегмента, имеющий координаты (m,n) , будем называть корнем сегмента. Пусть l — длина сегмента. Предполагаем, что линия формируется из сегментов по принципу 4-смежности [40] так, как представлено на Рис. 1.2.2. Математическое ожидание длины сегмента линии, проходящей под углом α , равно

$$M(l) = \operatorname{tg} \alpha + 1, \quad (1.3.1)$$

а вероятности возможных длин сегментов такой линии составят:

$$P\{l = [\operatorname{tg} \alpha] + 1\} = 1 - (\operatorname{tg} \alpha - [\operatorname{tg} \alpha]), \quad \text{а} \quad P\{l = [\operatorname{tg} \alpha] + 2\} = \operatorname{tg} \alpha - [\operatorname{tg} \alpha],$$

где $[\cdot]$ означает операцию взятия целой части.

Точку (m,n) , являющуюся корнем сегмента, можно рассматривать как источник линии на строке m ; причем здесь перестает быть существенной вся предыстория данной линии на строках от 0 до $m-1$ (существенно лишь значение угла ее наклона α). На следующей строке корень сегмента данной линии будет находиться в точке $(m+1, n_1)$, причем $M(n_1 - n) = \operatorname{tg} \alpha$. Отвлекаясь от стохастичности длин сегментов, можно сделать вывод, что состояние элементов (0 или 1) любой строки изображения полностью задано координатами корней всех проходящих через нее линий n_1, \dots, n_k и значениями углов наклона линий $\alpha_1, \dots, \alpha_k$. Заметим, что общее число элементов, заметаемых линией из точки (m,n) в точку (m',n') , равно $N_L = |m' - m| + |n' - n|$.

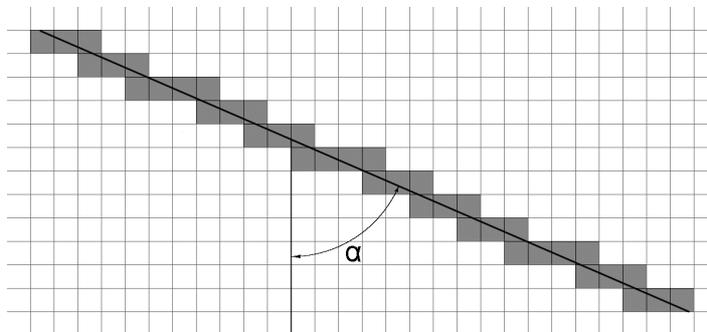


Рис. 1.2.2. Представление прямой линии, имеющей наклон α , на квадратном растре.

Для формирования линий на границах изображения нам потребуется знать распределение вероятностей углов α . Пусть уже имеется некоторая реализация изображения, удовлетворяющая заданным выше свойствам. В силу изотропности, распределение вероятностей длин отрезков

линий до узла или излома одинаково для всех направлений α . Проведем через изображение под произвольным углом прямую, называемую секущей. Вероятность пересечения секущей и линии изображения пропорциональна $l \cos \alpha$, где l — длина линии, а α — угол между линией и перпендикуляром к секущей. Очевидно, что плотность распределения вероятностей углов $p(\alpha)$ для секущей любого направления составит

$$p(\alpha) = \cos \alpha / \int_{-\pi/2}^{\pi/2} \cos \alpha \, d\alpha = 1/2 \cos \alpha, \quad -\pi/2 < \alpha \leq \pi/2. \quad (1.3.2)$$

Воспользовавшись (1.3.1) и (1.3.2), найдем среднюю длину сегмента L для множества дискретных линий под всевозможными углами α к вертикали и при условии 4-смежности:

$$L = \int_{-\pi/2}^{\pi/2} (1 + \operatorname{tg} \alpha) \cos \alpha \, d\alpha / \int_{-\pi/2}^{\pi/2} \cos \alpha \, d\alpha = 2. \quad (1.3.3)$$

1.3.2. Построение контурного изображения

Задача состоит в создании алгоритма построения изображения. Для определенности будем считать, что формирование осуществляется в процессе строчной последовательной развертки слева направо и сверху вниз (каузальная модель [49]), причем строки растра располагаются горизонтально. Из этого следует, что одиночные контурные линии могут порождаться лишь на верхней и боковых границах изображения. При этом направление линии может быть произвольным, но на верхней границе оно ограничено третьим и четвертым квадрантами, на левой границе — четвертым, а на правой — третьим квадрантами. Вся остальная картина изображения должна формироваться случайным образом в соответствии с задаваемыми вероятностями событий, состоящих в ветвлениях и изломах линий, показанных на Рис. 1.2.1. При таком построении узел 'S' (Рис. 1.2.1, вариант 9), возникающий при стыке двух идущих сверху линий, является основным; далее он может быть продолжен одной линией, двумя линиями, или не продолжаться вовсе. Нетрудно видеть, что остальные возможные узлы соответствуют повороту получающейся конфигурации на тот или иной угол. Поскольку базовыми принципами построения являются однородность и изотропность изображения, очевидно, что вероятности появления конфигураций узлов, различающихся только поворотом, должны быть одинаковыми для всех точек изображения.

Параметры модели и их соотношения

Зададимся вероятностью p появления корня сегмента в произвольной точке изображения, в том числе в точке границы изображения. Учитывая (1.3.3) сразу же отметим формальное ограничение: $p < 0,5$; реально же предполагается, что p еще значительно меньше. Как говорилось выше, возможны лишь три принципиальных вида конфигураций узлов: '<', 'Y' и 'X'. Соответственно, вероятности их появления должны быть взаимосвязаны. Будем считать,

что задаются условные вероятности появления узлов 'Y' и 'X' при стыке двух линий, т.е. $p_Y = P(Y|S)$ и $p_X = P(X|S)$; вероятности появления остальных возможных узлов будут производными от них. При формировании изображения потребуются вероятности ветвления линии $P(\lambda)$, излома линии $P(<)$, а также возникновения узла «на пустом месте» $P(\wedge)$.

Возьмем любую строку изображения, которую пересекают несколько линий. Рассмотрим, каковы окажутся средние расстояния между их сегментами в строке. Вероятность того, что корень сегмента некоторой пересекающей строку линии (номера k), расположенный в точке n_k , отстоит от точки корня предыдущей линии $k-1$ на расстояние x , равна $p(1-p)^{x-1}$. Можно показать, что при малых p среднее расстояние между корнями соседних линий Δ будет

$$\Delta = M(n_k - n_{k-1}) \approx 1/p. \quad (1.3.4)$$

Средний размер свободного промежутка $(\Delta - L)$ между сегментами соседних линий на одной строке при этом составит: $\Delta - L = M(n_k - n_{k-1}) - L \approx 1/p - 2$.

Найдем вероятность стыка пары линий p_S (узел S на Рис. 1.2.1), т.е. вероятность того, что сегмент некоторой линии пересечется с сегментом другой линии, также проводимой сверху вниз. Чтобы данное событие произошло, корень сегмента второй проводимой линии должен попасть в отрезок средней длины $2L - 1$ внутри промежутка между соседними корнями (который в среднем равен Δ). Таким образом, вероятность p_S того, что сегмент линии будет пересечен какой-то другой линией, равна

$$p_S = (2L - 1)/\Delta = 3/\Delta \approx 3p. \quad (1.3.5)$$

Событие S , как уже было сказано выше, является объединением трех возможных событий: узла Y , узла X и узла V . Сумма их условных вероятностей p_Y, p_X, p_V равна единице:

$$p_Y + p_X + p_V = P(Y|S) + P(X|S) + P(V|S) = 1; \quad p_V = 1 - p_Y - p_X. \quad (1.3.6)$$

Таким образом, вероятности того, что произвольно взятый сегмент окажется точкой стыка и ветвления того или иного вида, будут равны $p_S p_Y, p_S p_X$ и $p_S p_V$ соответственно.

Значение $P(\lambda)$, т.е. вероятность ветвления линии на две, найдем следующим образом. Поскольку узел λ является симметричным по отношению к узлу Y , то, очевидно, их вероятности должны быть взаимосвязаны. Возьмем две соседние строки изображения (m и $m+1$). Пусть строку m пересекает n_1 линий, причем с вероятностью $p_S p_Y$ две линии на строке n_1 сливаются в одну линию на строке n_2 (узел Y), а с вероятностью p_λ одна линия на строке n_1 разветвляется на две на строке n_2 (узел λ). Тогда на строке $m+1$ общее число линий будет равно $n_2 = n_1(1 - p_S p_Y - P(\lambda)) + n_1 p_S p_Y / 2 + 2n_1 P(\lambda)$. Приравнявая в силу однородности $n_1 = n_2$, и считая, что ветвление возможно в любой точке сегмента, вероятность его возникновения составит

$$P(\lambda) = p_S p_Y / 2L = 3/4 p p_Y. \quad (1.3.7)$$

Перейдем к узлам ‘ V ’, ‘ \wedge ’ и ‘ $<$ ’. Очевидно, что если узлы ‘ V ’ или ‘ \wedge ’ повернуть на 90° , то получится излом линии, т.е. узел ‘ $<$ ’. Для определения соотношения вероятностей рассмотрим точку, из которой исходят две линии под углами α_1 и α_2 . Очевидно, если α_1 и α_2 находятся в диапазоне $[-\pi/2, \pi/2]$, то это будет узел ‘ \wedge ’; если α_1 и α_2 в диапазоне $[\pi/2, -\pi/2]$, то это узел ‘ V ’. В остальных случаях, которых, как очевидно, в 2 раза больше, чем узлов ‘ V ’ или ‘ \wedge ’, это будет излом (узел ‘ $<$ ’). Согласно однородности и изотропности изображения, для средних количеств узлов этих конфигураций на любом участке изображения будут справедливы соотношения:

$$n(V) = n(\wedge), \quad n(<) = 2n(V). \quad (1.3.8)$$

Строку длиной в N элементов в среднем пересекает Np линий. Т.е. на строке в среднем NpL точек занято сегментами проходящих линий и $N(1-pL)$ точек свободно. Число узлов вида ‘ S ’ будет равно $n(S) = Npp_S = 3Np^2$, а число узлов вида ‘ V ’ будет равно $n(V) = n(S)p_V = 3Np^2p_V$. Поскольку $n(V) = n(\wedge)$, то в $N(1-pL)$ оставшихся свободных точках строки должно быть $3Np^2p_V$ узлов вида ‘ \wedge ’. Отсюда находится вероятность возникновения узла в свободной точке:

$$P(\wedge) = 3Np^2p_V / (N(1-pL)) = 3p^2p_V / (1-2p). \quad (1.3.9)$$

Вероятность излома $P(<)$ найдем аналогичным образом. В строке из N элементов в среднем NpL точек занято сегментами проходящих линий, из которых в свою очередь $3Np^2p_VL$ точек занято узлами вида ‘ V ’. Согласно (1.3.8), в оставшихся $(NpL - 3Np^2p_VL)$ точках должно произойти $2 \times 3Np^2p_V$ изломов. Отсюда

$$P(<) = 6Np^2p_V / (NpL - 3Np^2p_VL) = 3pp_V / (1 - 3pp_V). \quad (1.3.10)$$

Тем самым требуемые при генерации вероятности ветвления линии $P(\lambda)$, излома линии $P(<)$ и возникновения узла $P(\wedge)$ задаются через базовые вероятности p , p_V и p_X , что полностью определяет свойства контурного изображения и алгоритм его формирования.

Некоторой особенностью является задание линий на границе. В силу формирования изображения слева направо и сверху вниз, необходимо задавать конфигурацию линий лишь на верхней и боковых границах. При этом на левой боковой границе должны генерироваться лишь линии в диапазоне углов $0 < \alpha < \pi/2$, поскольку линии в диапазоне $\pi/2 < \alpha < \pi$ будут на этой границе заканчиваться при построении. Аналогично на правой боковой границе должны генерироваться линии в диапазоне углов $\pi/2 < \alpha < \pi$. Таким образом, вероятность генерации корня линии в точке составляет величину p для верхней и лишь $p/2$ для каждой из боковых границ. Распределение вероятностей углов наклона α линий должно быть равномерным на $[-\pi/2, \pi/2]$ для верхней границы, $[0, \pi/2]$ для левой и $[\pi/2, \pi]$ для правой границы.

Анализируя полученные соотношения для вероятностей p_S , p_λ , $p_<$ и $P(\wedge)$, отметим, что все они естественным образом зависят от вероятности появления корня p . Данный параметр может также рассматриваться в качестве масштабного коэффициента или, что то же самое, параметра

точности дискретизации. Интересно подметить, что согласно (1.3.3) L не зависит от p , т.е. средняя длина сегмента, независимо от дискретизации, составляет два элемента.

1.3.3. Эксперименты по синтезу контурных изображений

В соответствии с изложенной моделью была реализована программа, синтезирующая случайные контурные изображения согласно задаваемым значениям параметров. Некоторые примеры сформированных случайных изображений с различными значениями параметров p , p_Y и p_X представлены на Рис. 1.2.3. Так, изображение (а) содержит только ветвления вида Y и λ , а изображение (б) — только ветвления вида X . На изображении (в) ветвления отсутствуют, а изображение (г) включает все варианты возможных ветвлений.

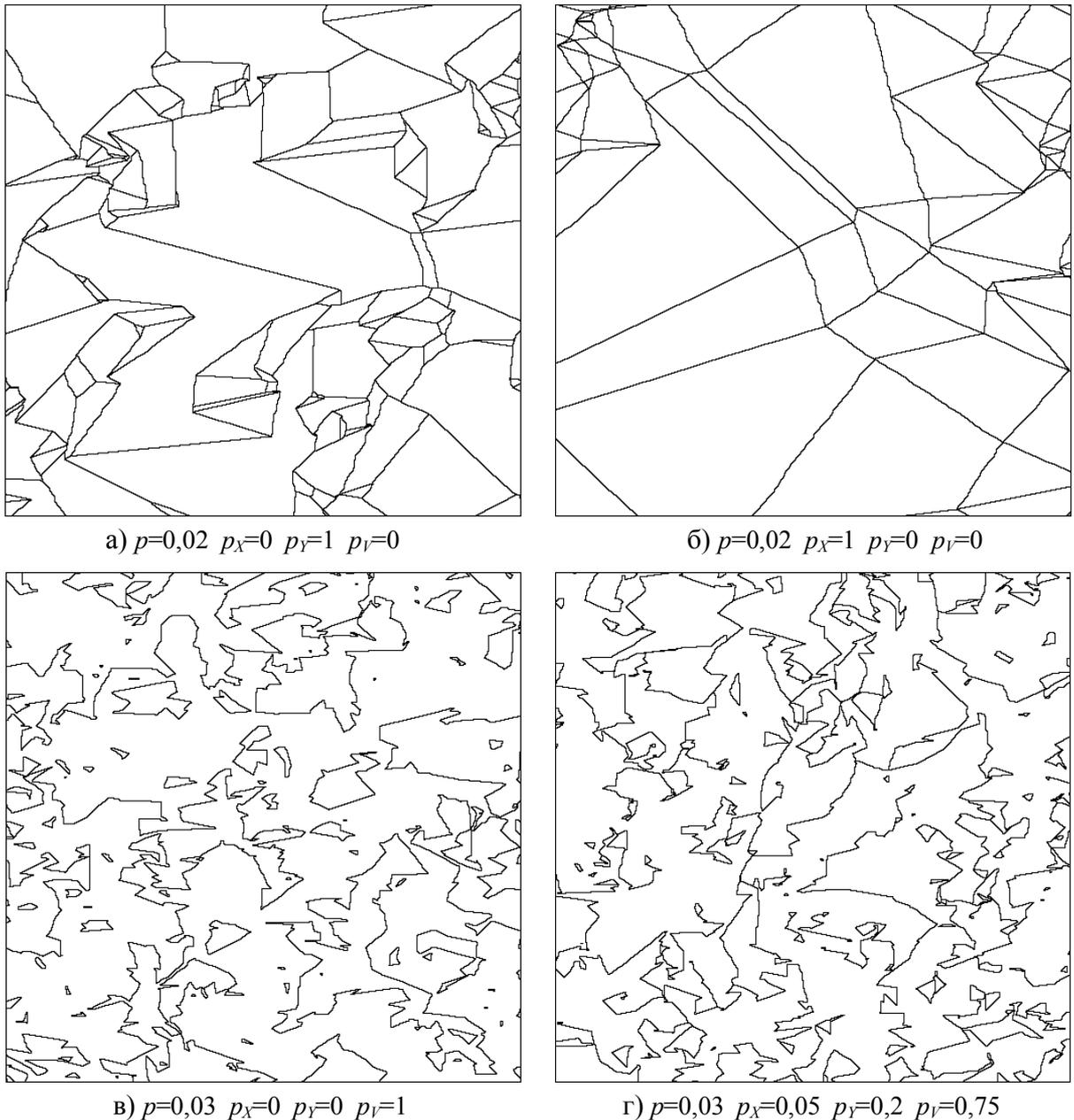


Рис. 1.2.3. Моделирование контурного изображения при различных параметрах.

Проведение экспериментов по генерации случайных контурных изображений показывает, что данная вероятностная модель описывает довольно широкий спектр контурных изображений, некоторые из которых по своей структуре напоминают контурные препараты, получаемые обработкой реальных изображений. Остается открытым вопрос, насколько изображения, генерируемые согласно столь простым принципам и параметрам, оказываются близки контурным структурам реальных изображений. Для ответа на него требуется выполнение морфологического исследования, выходящего за рамки настоящей работы.

Контурное изображение может быть легко преобразовано в обычное многоградационное изображение путем прослеживания связанных областей, ограниченных проведенными контурами, и заполнением их теми или иными значениями яркости. Именно так получено изображение, представленное на Рис. 1.2.4, в котором каждая из областей заполнена случайным постоянным значением яркости. Дополнительно возможно добавление случайного шума, приближающее изображение к модели (1.2.24), или даже заполнение областей участками случайных полиномиальных поверхностей согласно модели (1.2.22).

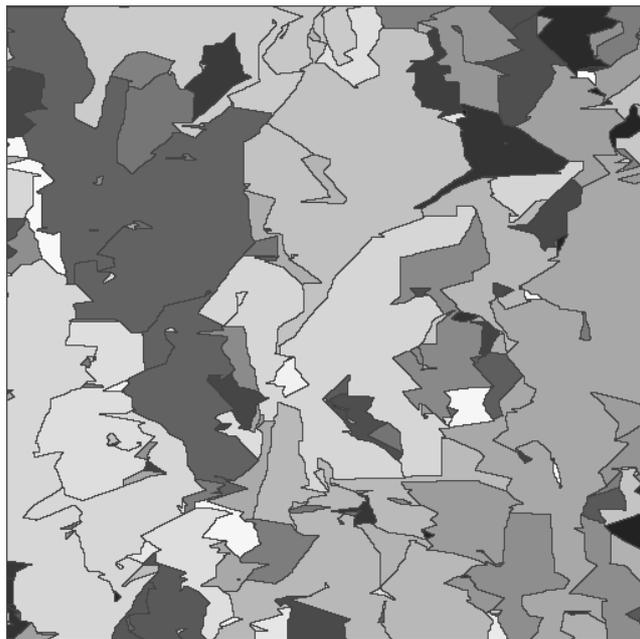


Рис. 1.2.4. Моделирование контурного изображения с заполнением областей случайной яркостью. $p=0,03$ $p_x=0,05$ $p_y=0,6$ $p_v=0,35$.

Возможно дальнейшее многоуровневое развитие контурной модели, заключающееся в следующем. Случайное изображение, сформированное с некоторыми заданными параметрами, содержит целый ряд связанных областей. Рассматривая данные области как самостоятельные, внутри них по тем же правилам можно сформировать контурную картину с другими свойствами (значениями параметров). Этот процесс может быть повторен многократно.

Изложенные в разделе 1.3 результаты опубликованы в работе [166].

1.4. Оценка сложности изображения

Сложность как характеристика изображения

В теоретических исследованиях и при практическом применении методов обработки видеoinформации требуются оценки достигаемого результата. Для задач, которые удается сформулировать как задачи восстановления или фильтрации сигнала, часто используются различные варианты отклонений (MSE, PSNR, SSIM и др.), которые удобны для теоретических разработок или в тех случаях, когда исходный сигнал известен. В реальных же ситуациях неискаженный сигнал недоступен. При таких условиях хотелось бы найти численную оценку результата, не использующую оригинальный сигнал, которая была бы адекватна изображению как двумерной функции.

Как правило, для решения данной задачи предлагаются методы, использующие один из следующих двух подходов. Первый заключается в попытке решения задачи оценивания «качества» изображения [281,364,365]. При этом отмечаются три уровня анализа, отличающиеся степенью доступности к неискаженному эталонному изображению для сравнения с ним результата. На первом уровне эталонное изображение доступно полностью, на втором лишь частично, на третьем полностью недоступно. Последний уровень назван «слепым» оцениванием. Внимание исследователей почти всегда фокусируется на первом, иногда на втором из указанных уровней, однако они оба предполагают использование оригинального сигнала. Другой подход чаще всего связан с независимым оцениванием сложности изображения. Обычно он используется при существенных ограничениях, например, в задаче сжатия изображений [373] или в задаче автоматического опознавания целей [329].

Применительно к изображению, которое является двумерным сигналом, также хотелось бы выбрать оценку, характеризующую его пространственную изменчивость. Плодотворным здесь представляется подход, основанный на рассмотрении изображения как функции с ограниченной вариацией [94] и оценивании ее изменчивости при помощи механизма двумерных и многомерных вариаций [33,59,75,211]. При этом постановка вопроса оценивания изменчивости близка вопросу оценивания сложности изображения [95].

Сложность изображения будет интерпретироваться как характеристика, отражающая число, размеры и заметность (контраст) деталей на изображении. Разумно также полагать, что при увеличении числа деталей и их контраста сложность должна возрастать, а при увеличении размеров деталей — уменьшаться.

О сложности сигналов и функций

Изучая вопрос сложности изображений, первоначально рассмотрим, как трактуется понятие сложности для сигналов, функций и потока сообщений.

Одномерный сигнал, являющийся действительной функцией времени, характеризуется энергией, длительностью и шириной спектра [22]. В теории могут использоваться различные определения длительности Δt и ширины спектра сигнала Δf , но формально эти две величины не могут быть конечными одновременно. Для реальных сигналов, имеющих начало и конец во времени, шириной спектра как правило считают полосу частот, в которой сосредоточена почти вся энергия сигнала (скажем, 99%). В теоретических исследованиях в качестве Δf часто принимают положение первого нуля спектра. Как известно из теории сигналов [146], независимо от выбранного определения длительности и ширины спектра сигнала, минимальное значение произведения $\Delta f \Delta t$, которое принято называть *базой*, для простых сигналов примерно равно единице. Соответственно, сигналы, для которых $\Delta f \Delta t \gg 1$, называют сложными [22]. Но к изображениям данное определение применить затруднительно. Во-первых, изображение является существенно двумерной функцией, а во-вторых (отвлекаясь от технологии формирования в поэлементной развертке), точки изображения, вообще говоря, соответствуют одному и тому же моменту или интервалу времени.

Сложность аналитической функции многих переменных, как отмечается в [34], характеризуется числом ее переменных n , а для s раз дифференцируемых функций — отношением n/s . Этот подход не применим к изображению как функции яркости пространственных координат, поскольку в данном случае имеем всего $n = 2$.

Теоретико-информационный подход к оценке сложности потока сообщений, представимого в дискретном виде, основан на вычислении энтропии $H = -\sum_b p(z) \log(p(z))$, где $p(z)$ — вероятность элемента изображения со значением z , а $b \in [b_{\min}, b_{\max}]$ — диапазон значений z . Такой вариант по сути предполагает независимость значений в точках изображения, в то время как изображение любой сцены является сигналом с сильными корреляционными связями. Поэтому подходы, используемые при оценивании сложности потока независимых сообщений, например энтропия, оказываются несостоятельными. Использование в качестве значения z некоторой функции, вычисляемой по близлежащим элементам, также не дает существенного улучшения и ограничивает универсальность оценки.

Для оценивания сложности дискретного сигнала иногда используется коэффициент сокращения объема данных (количества бит), необходимых для его восстановления. Такой подход, идеологически близкий понятию колмогоровской сложности [69], в реальности связан с

необходимостью выбора того или иного алгоритма сжатия и по сути является вариантом энтропийного оценивания.

Каждый из указанных способов оценки сложности по тем или иным причинам мало или даже совсем не применим к изображению как двумерной функции пространственных координат, отражающей некоторую физическую характеристику наблюдаемой сцены. В частности, существенный общий недостаток перечисленных вариантов оценок заключается в том, что они не являются метрическими характеристиками и не зависят от амплитуды сигнала.

Использование оценки сложности в задачах восстановления

Численные оценки сигнала актуальны при решении различных задач, в частности, классической оптимизационной задачи восстановления искаженного сигнала методом регуляризации [133], которая формулируется следующим образом. Пусть исходный сигнал z искажен оператором A (линейным или нелинейным) и шумом ξ , в результате чего наблюдается сигнал $u = A(z) + \xi$. Полагая оператор A с определенной точностью известным, строится восстанавливающий оператор \hat{A} (алгоритм) $\hat{A}(u) \Rightarrow \{z_\alpha\}$, отображающий u на множество допустимых результатов $z_\alpha \in Z$. Вычисляется функционал

$$\varphi_\alpha(u, \lambda) = \|Az_\alpha - u\| + \lambda\Omega(z_\alpha), \quad (1.4.1)$$

где $\Omega(z_\alpha)$ — неотрицательный стабилизирующий функционал, а λ — регуляризирующий параметр. Оптимальный на Z результат z^* находится путем минимизации функционала $\varphi_\alpha(u, \lambda)$ при выбранном значении λ :

$$\varphi_{z^*}(u, \lambda) = \inf_{z^a \in Z} \varphi_\alpha(u, \lambda). \quad (1.4.2)$$

Восстанавливающий оператор \hat{A} строится исходя из предположений о свойствах искажающего оператора A и зависит от конкретных условий. В каких-то случаях, как в задаче линейного восстановления [133], он может быть обратным оператором $\hat{A} = A^{-1}$, в других, как, например, в задаче фильтрации шума на основе полной вариации [344] — единичным оператором $\hat{A} = E$, или каким-то другим.

Для нас важным в данном алгоритме является то, что в отличие от восстанавливающего оператора \hat{A} , стабилизирующий функционал $\Omega(z)$, вообще говоря, не зависит от искажающего преобразования и выбирается как некоторая характеристика сигнала z , которая может задаваться исходя из модели сигнала. Для одномерных функций обычно предлагается использовать оценки типа нормы или вариации [133]. Наряду с этим можно интерпретировать функционал $\Omega(z)$ как некоторую оценку сложности сигнала.

1.4.1. Применение двумерных вариаций для оценки сложности

Двумерные вариации и их дискретные аналоги

Распространенной метрической характеристикой изменчивости и сложности одномерной функции на отрезке $[a, b]$ является вариация. Считая функцию $z(x)$ отражением некоей реальной физической величины (например, энергии или плотности) можно полагать, что $z(x)$ на наблюдаемом отрезке ограничена и имеет конечное число точек разрыва первого рода. Тогда $z(x)$ является функцией с ограниченной полной вариацией, которая по определению есть величина

$$V_a^b(z(x)) \triangleq \sup_P \sum_{k=0}^{K-1} \|z(x_{k+1}) - z(x_k)\|, \quad (1.4.3)$$

т.е. точная верхняя грань по всем возможным разбиениям P отрезка $[a, b]$. В контексте кусочно-постоянной модели изображения существенно отметить, что значение вариации не изменяется при «вставке» в функцию $z(x)$ участков вида $z(x) = \text{const} \{x \in [c, d] \subset [a, b]\}$, если при этом не добавляются разрывы первого рода.

Полная вариация $V_a^b(z(x))$ может служить удовлетворительной оценкой сложности одномерной функции. Предложено много ее обобщений на случай функции многих переменных (вариации Витали, Арцела, Фреше, Тонелли и др.). Все они сводятся к тому, что определяется некоторый один функционал, ограниченность которого гарантирует наличие у функции ряда свойств, аналогичных свойствам одномерной функции с конечным изменением. Однако перечень таких свойств оказывается ощутимо беднее набора свойств вариации для одномерных функций, хотя при этом и удается формулировать отдельные выводы и теоремы [33]. К тому же возникает определенная неоднозначность многомерных вариаций, связанная во многом с тем, что они существенно зависят от выбора системы координат.

Многомерные вариации (как и одномерные) формулируются для непрерывных функций через точные верхние грани по множеству допустимых разбиений носителя секущими гиперплоскостями на элементарные параллелепипеды. В случае дискретной функции, минимальное разбиение ограничено снизу пространственной точностью ее представления. Поэтому формулы вариации дискретных функций являются лишь соответствующими аналогами и приближениями формул вариации непрерывных функций, что в некоторых случаях может вызвать определенные неточности. Также необходимо отметить, что далеко не все операции могут быть применимы к дискретной функции без потери информации: например, операции поворота на произвольный угол, или операции сжатия как в пространстве носителя, так и в области значений.

Рассмотрим, как будут выглядеть наиболее известные определения вариаций [33,59,91] в применении к двумерной дискретной функции $\mathbf{F} = \{f(i,j)\}$ ($i=1,\dots,I; j=1,\dots,J$) на прямоугольном носителе D с соответствующим разбиением.

Арцела вариация [33,91] по сути является суммой модулей градиентов по координатным осям и представляется формулой

$$A(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} (|f(i+1,j) - f(i,j)| + |f(i,j+1) - f(i,j)|). \quad (1.4.4)$$

Данный способ оценки изменчивости двумерной функции используется в большинстве теоретических и практических разработок.

Витали вариация [33,91] формально записывается как

$$V(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |f(i+1,j+1) + f(i,j) - f(i,j+1) - f(i+1,j)|. \quad (1.4.5)$$

В качестве существенного недостатка данного дискретного представления вариации следует отметить, что на участках функции $f(i,j)$, где наблюдаются перепады ее значений в направлении только одной из координатных осей (i или j), отклик оператора, соответствующего формуле (1.4.5), будет равен нулю.

Пьерпонта вариация [91,332] предполагает равномерное разбиение каждой из сторон прямоугольника D на равные между собой отрезки, что идеологически соответствует равномерной дискретизации носителя. В каждом образуемом элементарном параллелепипеде $\delta(i,j)$ измеряется величина колебания $\omega(\mathbf{F}, \delta(i,j))$, а вариация вычисляется как

$$P(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} \omega(\mathbf{F}, \delta(i,j)). \quad (1.4.6)$$

Согласно [332], колебание ω есть $\omega(\mathbf{F}, \delta(i,j)) = \max\{f(x,y) \in \delta(i,j)\} - \min\{f(x,y) \in \delta(i,j)\}$, что в дискретном случае можно определить как $\omega(\mathbf{F}, \delta(i,j)) = \max\{|f(i+1,j) - f(i,j)|, |f(i,j+1) - f(i,j)|\}$.

Тонелли вариация [33,91] определяется как сумма одномерных вариаций (1.4.3) по всем значениям i и j каждой из координат при постоянном значении второй координаты. Пусть $\psi_1(x_i)$ — одномерная вариация функции $f(i,j)$ по j при $i = \text{const}$, а $\psi_2(j)$ — одномерная вариация функции $f(i,j)$ по i при $j = \text{const}$. Тогда

$$T(\mathbf{F}, D) = \sum_{i=1}^I \psi_1(i) + \sum_{j=1}^J \psi_2(j). \quad (1.4.7)$$

Нетрудно убедиться, что значения, вычисляемые по формулам (1.4.7) и (1.4.4) будут равны, т.е. вариации Арцела и Тонелли для дискретных функций совпадают.

Фреше вариация [91] в формулировке для двумерной дискретной функции совпадает с *Харди вариацией* [33,91] и вычисляется следующим образом:

$$\begin{aligned}
\Delta_1(f;(i,j)) &= f(i+1,j) - f(i,j), \\
\Delta_2(f;(i,j)) &= \Delta_2(\Delta_1(f;(i,j))) = (f(i+1,j+1) - f(i,j+1)) - (f(i+1,j) - f(i,j)), \\
F(\mathbf{F}, D) &= H(\mathbf{F}, D) = \sum_{i=1}^{I-1} \sum_{j=1}^{J-1} |\Delta_2(\Delta_1(f;(i,j)))|. \tag{1.4.8}
\end{aligned}$$

Как легко заметить, формулы (1.4.8) и (1.4.5) совпадают, т.е. вариации Витали, Фреше и Харди для дискретных функций между собой равны, в то время как для непрерывных функций они различаются.

Рассмотренные вариации так или иначе имеют в своей основе значение модуля градиента функции в точке, поэтому их величины и поведение близки между собой.

Кронрода вариация [75] в ряду вариаций многомерных функций стоит особняком по той причине, что не использует градиентные характеристики функции, а также дает в результате значения не одного, а нескольких функционалов. Ее рассмотрению посвящен следующий раздел.

Двумерная вариация Кронрода

Обобщения отдельных выводов и теорем, которые удавалось формулировать используя определения вышеперечисленных вариаций (для непрерывных функций), привели к выводу, что функция многих переменных должна характеризоваться не одним, а несколькими функционалами, которые в определенном смысле независимы. Данный тезис был обоснован А.С. Кронродом при изучении функций двух переменных [75]. Основываясь на понятии множеств уровня он предложил для функции двух переменных использовать два функционала, определяемых следующим образом:

$$w_1(f) = \int_{-\infty}^{\infty} v_0(e_t) dt, \quad w_2(f) = \int_{-\infty}^{\infty} v_1(e_t) dt. \tag{1.4.9}$$

Здесь множество e_t — это t -уровень функции $f(x,y)$, т.е. множество тех точек (x,y) , в которых $f(x,y) = t$; $v_0(e_t)$ — число компонент множества e_t ; $v_1(e_t)$ — длина e_t (по Хаусдорфу). Тем самым значение w_1 характеризует число и выраженность локальных экстремумов функции.

А.Г. Витушкиным было найдено удачное определение вариации множества, причем предложенный подход, также основанный на рассмотрении множеств уровня, был расширен для функций многих переменных [33]. Для плоского множества e_t , которое является t -уровнем функции $f(x,y)$, вариация задается не одним, а тремя значениями: v_0 , v_1 и v_2 , которые определяются следующим образом.

Пусть E — замкнутое ограниченное множество на плоскости; тогда вариация $v_0(E)$ есть число компонент E , вариация $v_2(E)$ есть суммарная площадь компонент E , а значение вариации $v_1(E)$ определяется следующим выражением [33,59]:

$$v_1(E) = c \int_0^{2\pi} v_0(E, L_\alpha) d\alpha, \quad \text{где} \quad v_0(E, L_\alpha) = \int_{L_\alpha} v_0(E \cap (L_\alpha)_z^\perp) dz.$$

Здесь L_α — прямая $x \cos \alpha + y \sin \alpha = 0$, а $(L_\alpha)_z^\perp$ — прямая, проходящая через точку $z \in L_\alpha$ перпендикулярно к L_α . Константа c выбирается таким образом, чтобы для отрезка I единичной длины на оси Ox выполнялось соотношение: $v_1(I) = 1$. Для случая конечной $v_0(e_t)$, вариацию $v_1(e_t)$ можно интерпретировать как длину границ компонент e_t . Вариации v_0 , v_1 и v_2 названы соответственно *нулевой*, *линейной* и *плоской* согласно тому, как изменяются их значения при гомеоморфизме: v_0 не меняется, v_1 меняется линейно с изменением коэффициента растяжения K , а v_2 меняется пропорционально K^2 , т.е. как площадь плоской фигуры.

Важно отметить, что вариация v_0 множества уровня e_t , входящая в определение вариации $w_1(f)$, является не метрической, а в определенном смысле топологической характеристикой функции, основанной на понятии связности и не изменяющейся при гомеоморфизме. Касательно второй вариации w_2 следует отметить, что для непрерывно дифференцируемой на D функции $f(x, y)$ значение $w_2(f)$ будет равно [94]:

$$w_2(f) = \iint_D |\text{grad}(f(x, y))| dx dy,$$

т.е. $w_2(f)$ совпадает с интегралом модуля градиента функции по области носителя. В нашем случае $f(x, y)$ есть значение сигнала (яркости) изображения, которое можно считать неотрицательным и ограниченным, а значит пределы интегрирования в (1.4.9) можно ограничить диапазоном $[0, T]$, где T — максимальное допустимое значение.

В отличие от v_0 и v_1 , характеристика $v_2(e_t)$ не представляет интереса для оценки сложности функции. Рассматривая $v_2(e_t)$ как сумму площадей всех компонент сечения уровня t функции яркости $f(x, y)$, разделенных границами, очевидно, что $v_2(e_t) = S(D)$, т.е. равно площади всего множества D , и тем самым дополнительной информации не несет.

Можно предположить, что полезным окажется использование подмножества $\tilde{v}_2(e_t)$ ($\tilde{v}_2(e_t) \subset v_2(e_t)$), состоящего, скажем, только из тех точек (x, y) , для которых выполняется $f(x, y) \geq t$. По аналогии с (1.4.9) определим третий функционал как интеграл плоской вариации

$$w_3(f) = \int_0^T \tilde{v}_2(e_t) dt,$$

значение которого будет равно объему трехмерной фигуры, ограниченной функцией $f(x, y)$ и плоскостью $z(x, y) = 0$. Поскольку $f(x, y)$ задана на ограниченном двумерном множестве $D \ni (x, y)$, справедливо выражение $\tilde{v}_2(e_t)/S(D) = 1 - F_f(t)$, где $S(D)$ — площадь D , а $F_f(t)$ — функция распределения значений $f(x, y)$. Отсюда

$$\tilde{v}_2(e_t) = S(D)(1 - \int_0^t h(z) dz) = S(D) \int_t^T h(z) dz,$$

где $h(z) = p\{f(x,y) = z\}$ есть плотность вероятности значений функции $f(x,y)$. Легко видеть, что при этом соотношение $w_3(f)/S(D)$ будет равно среднему значению $f(x,y)$ на D , или, в интерпретации изображения, его средней яркости. Это означает, что функционал $w_3(f)$ также не содержит информации о структуре $f(x,y)$. Тем самым, для оценки сложности функции $f(x,y)$ остаются только две вышеприведенные вариации $w_1(f)$ и $w_2(f)$.

В случае дискретной функции $f(i,j)$, двумерные вариации (1.4.9) выражаются следующими формулами:

$$w_1(f) = \sum_{x=0}^{T-1} v_0(e_t) / T \quad \text{и} \quad w_2(f) = \sum_{x=0}^{T-1} v_1(e_t) / T, \quad (1.4.10)$$

где T — общее число возможных значений функции $f(i,j)$ (для изображения — число градаций яркости). При вычислении вариаций, в качестве множества e_t выступает бинарная матрица $\mathbf{B} = \{b(i,j)\}$, строящаяся следующим образом:

$$b(i,j) = \begin{cases} 0, & \text{если } f(i,j) < t, \\ 1, & \text{если } f(i,j) \geq t. \end{cases}$$

Дискретизация изображения осуществляется, как правило, на квадратной решетке. При этом значение $v_0(e_t)$ определяется как сумма числа связных компонент, составленных из соседствующих (в смысле 4-соседства [40]) элементов изображения со значением $b_{i,j} = 0$ и числа аналогичных связных компонент из элементов с $b(i,j) = 1$. Значение $v_1(e_t)$ есть суммарная длина границ компонент $v_0(e_t)$. В случае дискретной функции $f(i,j)$ значение $v_1(e_t)$ зависит от способа вычисления длины границы. На квадратной решетке длину границ чаще всего измеряют в метрике L_1 , определяя $v_1(e_t)$ как суммарное число сегментов решетки, разделяющих элементы с несовпадающими значениями.

Формальное вычисление значений вариаций согласно (1.4.10) на ограниченном носителе D предполагает, что сама область D также учитывается в качестве отдельной компоненты. Это приводит к следующему противоречию. Вариации w_1 и w_2 для функции постоянного значения $f(i,j) = \text{const}$, для которой они, очевидно, должны быть равны нулю, оказываются ненулевыми: $w_1(f) = 1$, а $w_2(f) = P(D)$ — периметру области D . Для устранения данного противоречия следует модифицировать формулы (1.4.10) следующим образом:

$$w_1(f) = \left(\sum_{x=0}^{T-1} v_0(e_t) / T \right) - 1 \quad \text{и} \quad w_2(f) = \left(\sum_{x=0}^{T-1} v_1(e_t) / T \right) - P(D). \quad (1.4.11)$$

В дальнейшем значения w_1 и w_2 для дискретной функции $f(i,j)$ на ограниченном носителе будут пониматься именно в смысле (1.4.11). Будучи нормированным на величину диапазона T , значение w_1 является характеристикой числа и амплитуды объектов изображения, образующих в сечениях отдельные компоненты. Поэтому удобно назвать значение w_1 *показателем числа*

объектов изображения. Можно также показать, что значение второй вариации $w_2(f)$ в (1.4.11) будет совпадать со значением дискретной вариации Тонелли (1.4.7).

Показатель размеров объектов изображения

Получаемые значения вариаций Кронрода позволяют оценить некоторые параметры изображения, которые отражают как амплитудные, так и морфологические его характеристики. Значение первой вариации w_1 отражает число и контраст деталей на изображении, а значение второй вариации, w_2 , — сумму периметров деталей. Наряду с этим, важным выглядит соотношение вариаций $q_w(f) = w_2(f)/w_1(f)$, которое отражает средний периметр деталей на изображении (здесь полагаем $w_1(f) > 0$, иначе $q_w(f) = 0$).

В дискретном изображении наименьшим возможным объектом является один элемент (пиксель). Это значит, что минимально возможная компонента, из числа входящих в множество уровня e_t , имеет линейный размер равный одному шагу дискретизации. Периметр такой компоненты из одного элемента будет равен 4 — значению, которое есть аналог числа π в метрике L_1 и является минимально возможным для соотношения периметр/площадь объекта. Предполагая, что в дискретном случае длина границ компонент, составляющих $v_1(e_t)$, измеряется в метрике L_1 , очевидно, что и для $q_w(f)$ также существует минимально возможное значение, которое тоже равно 4, причем минимум достигается в случае, когда все компоненты множеств уровня e_t для всех t являются одноэлементными. В связи со сказанным удобно ввести нормированную характеристику

$$d(f) = q_w(f)/4, \quad (1.4.12)$$

которую по аналогии с показателем числа компонент w_1 в (9) естественно назвать *показатель размеров объектов* изображения; для нее будет выполняться соотношение $d(f) \geq 1$.

Двумерная вариация как оценка сложности изображения

Для изображений с преобладанием мелких деталей, которые на множествах уровня отображаются компонентами небольших размеров, значения d будут малыми, а для изображений с крупными деталями — большими. Отметим, что к деталям изображения относятся также и шумы, являющиеся локальными выбросами сигнала, и проявляющиеся на множествах уровня как самостоятельные компоненты с минимально возможным периметром. В связи с этим значение d должно убывать при увеличении числа и амплитуды шумовых выбросов. Ниже этот факт подтверждается экспериментально.

Особый интерес представляет поведение значений $w_1(f)$ и $d(f)$ при возможных преобразованиях функции $f(x,y)$. Для некоторых важных частных случаев можно сформулировать следующие утверждения.

1. Значение w_1 не зависит от выбора системы координат, а d зависит от выбора или поворота системы координат относительно носителя D лишь в степени точности дискретизации $f(x,y)$.
2. При линейных амплитудных преобразованиях вида $Cf(x,y)$, где C — константа, значение w_1 увеличивается пропорционально C , а значение d не изменяется.
3. При линейном пространственном растяжении носителя D в K раз, значение w_1 не изменяется, а значение d увеличивается также в K раз.
4. Расширим D добавлением области U , на которой $f(x,y)$ имеет постоянное значение; это означает, что $f(x,y) = \text{const} \quad \{(x,y) \in U \subset D' = D \cup U\}$. Условием выполнения операции расширения является то, что она не должна добавлять новых участков разрыва первого рода. Такое расширение не приводит к изменению значений w_1 и d .
5. При увеличении количества деталей на изображении (но при условии сохранения необходимых статистических соотношений и распределений) значение w_1 пропорционально увеличивается, а значение d не изменяется.
6. Пусть область D прямоугольна, и по выбранной оси, скажем оси Y , ограничена отрезком $[a,b]$. Расширим ее областью U , заданной на полуотрезке $(b, 2b - a]$, на котором $f(x,y)$ имеет зеркальное продолжение, т.е. $f(x, b + y) = f(x, b - y)$. При таком расширении значение w_1 увеличивается пропорционально изменению площади, тогда как значение d не меняется.

Таким образом, при указанных изменениях двумерной функции $f(x,y)$, как минимум одно из значений w_1 и d оказывается инвариантом, второе же значение является предсказуемой функцией преобразования. Применительно к изображениям утверждение 2 соответствует линейному изменению контраста; утверждение 3 — линейной геометрической трансформации; утверждение 4 — добавлению/удалению участков изображения с ровным фоном без деталей; утверждение 5 близко предыдущему и соответствует концентрации/разреженности деталей на изображении; утверждение 6 — вариант зеркальной пролонгации изображения за рамки области D , часто используемый алгоритмами локального анализа. Кроме того, повторение свойств части изображения на всем изображении по существу означает однородность (стационарность), т.е. достаточность определения значений w_1 и d на доверительном участке и интерпретации их для изображения в целом.

Отметим важную особенность введенного показателя размеров объектов $d(f)$, имеющую следствие в утверждениях 2 и 5. Характеристика $d(f)$ обладает тем свойством, что отражает лишь средние размеры объектов на изображении и не зависит от количества и контраста самих объектов — эту часть информации несет вариация $w_1(f)$.

Из сказанного ясно, что первая вариация $w_1(f)$ вместе с характеристикой $g(f)$ отражают пространственную и яркостную изменчивость изображения и в совокупности могут служить

показателями *сложности изображения* в соответствии с требованиями, сформулированными выше. Различное поведение указанных характеристик подтверждает ту изначальную гипотезу А.С. Кронрода, что для описания изменчивости двумерных функций какой-то одной вариации недостаточно и следует использовать два независимых функционала.

О сравнении и интерпретации оценок

Поскольку характеристик сложности оказывается две, причем они вообще говоря независимы, то возникает естественный вопрос, как с их помощью сравнивать различные изображения между собой. Достаточно очевидно, что если характеристика $d(f)$ постоянна, то с увеличением $w_1(f)$ сложность изображения будет возрастать; также сложность будет возрастать и с уменьшением $d(f)$ при постоянном $w_1(f)$. Тогда, если для пары изображений f_1 и f_2 выполняются соотношения $w_1(f_1) < w_1(f_2)$ и $d(f_1) > d(f_2)$, то естественно полагать, что изображение f_2 является более сложным, чем f_1 . Остается вопрос, как сравнивать другие случаи.

Рассмотрим изображение с малыми значениями w_1 и d ; такие соотношения означают, что на изображении имеются детали малых размеров и их либо немного, либо они мало контрастные. Наоборот, большие значения w_1 и d свидетельствуют, что изображение содержит детали большого размера, и вероятно они высококонтрастные. В подобных случаях обычно принято говорить, что оценка сложности зависит от конкретной задачи.

В некоторых случаях требуется из нескольких функционалов тем или иным образом формировать обобщенное значение. Одним из простейших вариантов обобщения является следующая линейная комбинация:

$$w(z) = w_1(z) + Cw_2(z) = w_1(z)(1 + Cd(z)), \quad (1.4.13)$$

которая может применяться в (1.4.1) в качестве стабилизирующего функционала $\Omega(z)$.

1.4.2. Экспериментальные исследования

Представляет интерес, насколько измеряемые значения сложности $w_1(f)$ и $d(f)$ могут меняться от изображения к изображению, как они зависят от уровня шумов и как изменяются при сглаживании. При проведении подобных экспериментов требуется сравнение сигнала, получаемого при преобразовании, с неким эталонным и неискаженным. Сформировать такое эталонное изображение можно основываясь на двухмасштабной многокомпонентной модели изображения (Раздел 1.2). Согласно данной модели (1.2.24), изображение $f(i,j)$ представляется суммой двух компонент: кусочно-гладкой $S(i,j)$ и текстурно-шумовой $\tau(i,j)$. Последняя в свою очередь несет информацию о текстуре, мелких деталях $t(i,j)$ и шуме $\xi(i,j)$:

$$f(i,j) = S(i,j) + \tau(i,j) = S(i,j) + t(i,j) + \xi(i,j).$$

Для локальных алгоритмов обработки результат преобразования в каждой точке (i,j) зависит лишь от элементов, попадающих в ограниченную и сравнительно небольшую область анализа $R(i,j)$. В разделе 1.2.3 показано, что в таком случае для большинства реальных изображений можно считать, что $S(u,v) = S(i,j)$ при условии $(u,v) \in R(i,j)$. Таким образом в пределах области $R(i,j)$ значение $S(u,v)$ полагается постоянным. Это позволяет формировать эталонное изображение как кусочно-постоянный двумерный сигнал, который впоследствии будет подвергаться искажениям и процедурам сглаживания.

Для экспериментов был выбран набор из реальных и синтезированных изображений различной детальности и тематического содержания. Часть изображений представлена на Рис. 1.4.1: (а) и (б) — портреты с мелкими и крупными деталями; (в), (г), (д) — городской ландшафт, аэрофотоснимок и натюрморт соответственно; (е) — компьютерная томограмма; (ж) — фрагмент изображения текста; (з) — синтезированный тестовый сигнал с кусочно-постоянными деталями различных размеров и яркости на ровном фоне.



Рис. 1.4.1. Тестовые изображения.

В Таблицах 1.4.1 и 1.4.2 представлены значения двумерных вариаций, измеренных для тестовых изображений на Рис. 1.4.1. Двумерные вариации Арцела (A), Тонелли (T), Пьерпонта (P), Витали (V), Фреше (F), Харди (H), Кронрода ($K-w_1$ и $K-w_2$) вычислялись согласно формулам (1.4.4)-(1.4.8) и (1.4.11). Приведены также соответствующие показатели размеров объектов (d) согласно (1.4.12) и значения энтропии (E).

Как видно, значения одних и тех же вариаций для реальных изображений могут различаться в несколько раз, тогда как значения энтропии отличаются очень мало. Это

подтверждает сделанный ранее вывод, что энтропия не является подходящей оценкой сложности изображения. Сравнение значений вариаций в Таблицах 1.4.1 и 1.4.2 демонстрирует достаточно точное совпадение вариаций Арцела / Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) для всех изображений.

Для синтезированных изображений, (Рис. 1.4.1,з) и аналогичных, значения вариаций и энтропии оказываются существенно отличающимися от соответствующих значений для реальных изображений. Объясняется это тем, что подобные изображения, являясь кусочно-постоянными, не содержат шума и мелких деталей. Добавление же даже небольшого шума, который обозначен составляющей ξ в (1.2.3), резко меняет значения оценок и делает их мало отличимыми от тех, которые были получены для реальных изображений. Это демонстрируется в последней строке Таблицы 1.4.1, соответствующей изображению Рис. 1.4.1,з, искаженному аддитивным гауссовым шумом с $\sigma = 2$ градации яркости. Более подробно влияние шума рассматривается ниже.

Таблица 1.4.1. Значения двумерных вариаций, показателя размеров объектов (d) и энтропии (E) для изображений на Рис. 1.4.1.

Изображение	A/T	P	$V/F/H$	$K-w_1$	$K-w_2$	d	E
а	5242,72	3715,70	2859,43	469,85	5260,25	2,80	7,59
б	2583,56	1825,14	1038,43	129,23	2589,58	5,01	7,60
в	7928,75	6072,46	3497,91	618,79	7947,70	3,21	7,59
г	4804,46	3399,38	2323,97	380,03	4804,38	3,16	7,16
д	3777,20	2844,99	1497,63	178,63	3777,15	5,29	7,29
е	5107,64	3642,48	2499,84	419,62	5118,65	3,05	6,76
ж	6462,14	4989,61	2165,84	231,24	6462,04	6,99	5,61
з	811,67	778,54	132,52	33,13	811,67	6,12	0,58
з+шум	1944,96	1539,15	945,04	233,96	1944,93	2,08	3,63

Влияние шума на оценку сложности изображения

Для определения влияние шума на оценку сложности изображения было выбрано синтезированное изображение с кусочно-постоянными объектами (Рис. 1.4.1,з) разных размеров и контраста. На него наносился нормально распределенный аддитивный шум $N(0, \sigma^2)$ с величиной среднеквадратического отклонения $\sigma = 1; 2; 3$ и 5 градаций яркости. Результаты измерения различных двумерных вариаций, показателя размеров объектов (d) и энтропии (E) численно приведены в Таблице 1.4.2, а графически представлены на Рис. 1.4.2.

Важно отметить, что наличие даже слабого шума с СКО в 1 градацию существенно меняет значения всех вариаций. Разницы особенно заметны на первой вариации Кронрода ($K-w_1$), возрастающей в 10 раз, и показателе размеров объектов (d), убывающем в 7 раз.

Таблица 1.4.2. Значения двумерных вариаций, показателя размеров объектов (d), энтропии (E), а также их зависимости от уровня добавленного шума для изображения на Рис. 1.4.1,з.

Шум, град.	A/T	P	$V/F/H$	$K-w_1$	$K-w_2$	d	E
$\sigma = 0$	1160,50	1021,17	528,25	11,34	1160,48	25,58	3,11
$\sigma = 1$	1641,15	1327,58	909,12	110,27	1644,71	3,73	5,06
$\sigma = 2$	2198,37	1696,72	1303,17	210,29	2204,56	2,62	5,94
$\sigma = 3$	2758,17	2072,65	1697,21	309,89	2766,76	2,23	6,43
$\sigma = 5$	3867,95	2816,74	2487,14	511,75	3892,58	1,90	7,01

Из приведенных данных и графиков можно сделать следующие выводы: а) с увеличением уровня шума (для выбранного диапазона шумов) значения всех вариаций растут практически линейно; б) показатель размеров объектов (d) сначала резко падает, затем убывает медленно, стремясь к предельному значению 1; в) энтропия (E) асимптотически возрастает до предельного значения 8, равного числу битов в двоичной записи значений яркости; г) в подтверждение сказанному выше, значения вариаций Арцела / Тонелли (A/T) и второй вариации Кронрода ($K-w_2$) с большой степенью точности совпадают.

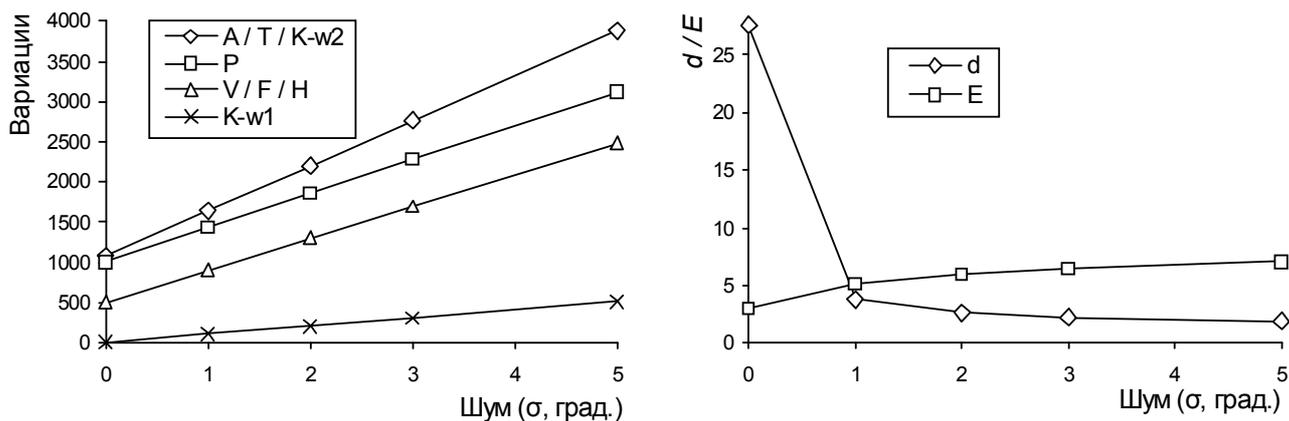


Рис. 1.4.2. Зависимость значений двумерных *вариаций*, показателя размеров объектов d и энтропии E от величины нанесенного гауссова шума (σ , град.) для изображения на Рис. 1.4.1,з.

Зависимости двумерных вариаций от размеров и поворота деталей

На Рис. 1.4.3 показаны зависимости значений двумерных вариаций от размеров деталей и их поворота. Для этого было сформировано несколько изображений, аналогичных показанному на Рис. 1.4.1,з. Изображения имели постоянный фон и содержали одинаковое число прямоугольных объектов случайной яркости, имеющих размеры от 1x1 до 5x5 элементов. Значения вариаций и показателей размеров объектов измерялись как для изображений с нормальным положением объектов, когда их стороны параллельны строкам и столбцам, так и для изображений с повернутыми на 45° объектами без интерполяции значений элементов, но при соответствующем приближенном проведении прямых контурных линий.

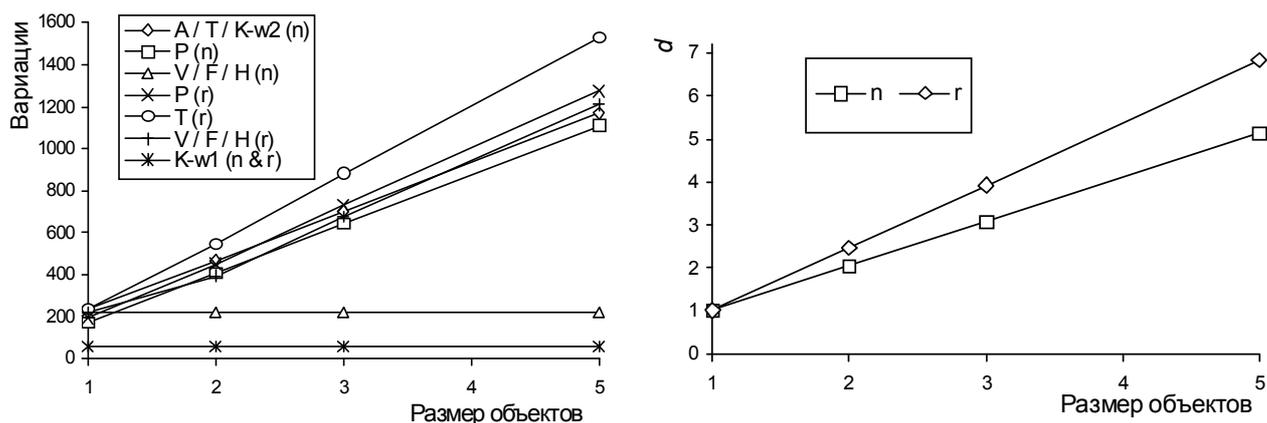


Рис. 1.4.3. Значения двумерных *вариаций* и показателя размеров объектов d при нормальном (n) и повернутом (r) положениях объектов в зависимости от размеров объектов.

На левом графике Рис. 1.4.3 хорошо видно, что график $K-w_1(n&r)$ горизонтален и совпадает для нормальных и повернутых объектов. Это подтверждает, что показатель числа объектов w_1 инвариантен как к изменению размеров объектов, так и к их повороту. Постоянство графика значений $V/F/H(n)$, который соответствует объектам в нормальной ориентации, подтверждает сделанный выше вывод, что локальные операторы, используемые в дискретных аналогах вариаций Витали / Фреше / Харди, дают нулевые отклики на тех участках функции, где она имеет перепад значений в направлении лишь одной из координатных осей. Если же контурные линии не вертикальны или не горизонтальны, то такой особенности не наблюдается (график $V/F/H(r)$). Графики всех остальных вариаций расположены достаточно плотным пучком, что говорит о том, что соответствующие значения вариаций различаются слабо.

Как видно из правого графика на Рис. 1.4.3, показатель размеров объектов d изменяется пропорционально размерам объектов, тем самым иллюстрируя утверждение 3, что значение этой характеристики линейно зависит от коэффициента растяжения носителя. Тот факт, что график, соответствующий повернутым объектам, имеет несколько большую крутизну, противореча тем самым утверждению 1, объясняется тем, что длина контура повернутого дискретного объекта, измеряемая в метрике L_1 , оказывается больше, чем длина контура объекта с вертикальным и горизонтальным расположением краев.

Выводы касательно оценки сложности изображения

Полученные результаты позволяют сделать вывод, что при условии справедливости двухмасштабной многокомпонентной модели, выбранная совокупность значений w_1 и d достаточно хорошо отражает морфологическую структуру изображения, в частности, такие важные параметры, как число, размеры и заметность деталей изображения, и поэтому может являться оценкой сложности изображения. Определяемая таким способом характеристика

может быть использована в качестве оценки сложности изображения, например, как стабилизирующий функционал при решении задач восстановления сигнала.

Изложенные в разделе 1.4 результаты опубликованы в работах [95, 168, 182, 203, 244, 245, 314].

1.5. Выводы и результаты

Необходимость различных подходов к описанию свойств изображения для малого и большого масштабов вызвана разницей самих физических причин, лежащих в основе корреляционных связей элементов на ближних и дальних расстояниях. На малых расстояниях определяющими являются шумы, мелкие детали, текстура и контурные перепады, а на больших расстояниях — отличия яркостей протяженных областей изображения, отвечающих разным объектам наблюдаемой сцены, и переходы полутонов внутри каждой из них, вызываемые, чаще всего, изменением формы и освещения объекта, а также собственными изменениями яркости и/или цвета его поверхности.

Разработана двухмасштабная многокомпонентная математическая модель цифрового многоградационного изображения. Данная модель рассматривает изображение в пределах локальных областей анализа как сумму разрывной кусочно-гладкой компоненты, аппроксимируемой участками аналитических поверхностей, и случайных текстурно-детальной и шумовой составляющих. Модель описывает свойства изображения для областей анализа как малого размера (масштаба), в пределах нескольких шагов дискретизации, так и для областей большого масштаба, соразмерного объектам изображения. Двухмасштабная модель является локальной. Даны определения локальной операции, локальной области анализа и области принадлежности. Модель применима как к одноканальным (черно-белым многоградационным изображениям), так и многоканальным (цветным изображениям).

Проведены исследования модели на реальных изображениях, подтверждающие ее соответствие реальной видеоинформации. Проанализированы статистические связи элементов изображения, показано различие свойств элементов, расположенных на внутренних участках протяженных областей (фоновых) и на участках, разграничивающих области (контурных).

Двухмасштабная многокомпонентная модель изображения может являться фундаментом для исследования и построения новых и эффективных алгоритмов обработки и анализа видеоинформации. Показана связь параметров модели с характеристиками изображения. Предложен способ определения случайного независимого шума на изображении, а также подходы к оцениванию визуальной контрастности, четкости и степени дискретизации изображения.

Построена вероятностная модель двумерного контурного изображения. Модель базируется на простой аксиоматике задания свойств границ объектов и имеет всего три параметра: вероятность появления контурной линии и вероятности возникновения двух видов узловых точек из трех принципиально возможных. Разработан алгоритм моделирования источника, порождающего дискретные контурные изображения с задаваемыми вероятностными характеристиками. На основе реализованного источника проиллюстрировано формирование контурных изображений с различными сочетаниями указанных вероятностей. Показана возможность получения многоградационных изображений со случайным рисунком областей, которые в свою очередь могут разбиваться контурами с другими статистическими свойствами, заполняться участками аналитических поверхностей, случайным шумом и т.д.

Формируемые контурные и многоградационные изображения могут быть использованы в различных задачах, связанных с обработкой, анализом и сжатием видеоинформации в качестве модельного сигнала с заданными свойствами.

Дана интерпретация сложности изображения как характеристики, отражающей число, размеры и заметность деталей на изображении. Исследована возможность оценивания сложности изображения на основе использования аппарата двумерных вариаций. Показана редукция многомерных вариаций непрерывных функций к двумерным дискретным функциям. Проведено сравнение различных известных вариаций; показано преимущество двумерной вариации Кронрода. На основе двумерных вариаций введена новая характеристика, соответствующая среднему периметру деталей на изображении, и предложена ее модификация, названная *показатель размеров объектов*.

Проанализированы изменения значений двумерных вариаций при различных трансформациях изображения. Теоретические выводы подтверждены экспериментально. Исследовано влияние шума и геометрических трансформаций изображения на оценку его сложности. Исследования подтверждают, что комбинация показателя числа объектов w_1 и показателя размеров объектов d хорошо отражает морфологическую структуру изображения и может являться оценкой сложности изображения. Предложен вариант редукции двух значений вариаций к единственному обобщенному функционалу, который может использоваться в классических оптимизационных схемах, например, как стабилизирующий функционал при решении задач восстановления сигнала.

ГЛАВА 2. ДЕКОМПОЗИЦИЯ ИЗОБРАЖЕНИЯ

Введение

Согласно рассмотренной в первой главе модели, изображение представляет собой сложный объект, сочетающий несколько составляющих различного информационного содержания. В исходном сигнале они представлены в виде смеси, что существенно затрудняет анализ поступающей видеoinформации. Разделение изображения на соответствующие информационные компоненты, сложность каждой из которых меньше, чем сложность исходного сигнала, является перспективным подходом, упрощающим последующую обработку и анализ.

В настоящей главе рассматривается задача разделения изображения на две компоненты согласно модели (1.2.2) — кусочно-гладкую компоненту s_{mn} (будем называть ее *сглаженной*), несущую информацию о протяженных областях изображения и границах между ними, и *текстурно-детальную* (иногда называемой просто *текстурной*) компоненту t_{mn} , содержащую текстуру, мелкие детали и шум. Такую операцию разделения будем называть *декомпозицией*.

Выделение только интересующей компоненты позволяет встать на некоторую новую ступень на пути к решению многих традиционных задач обработки и анализа изображений. Так например, выделяя текстурную компоненту и удаляя информацию о средней яркости, можно более эффективно решать задачу текстурной сегментации. Варьируя параметрами, управляющими разделением (т.е. относя объекты к той или иной компоненте), можно решать некоторые задачи обнаружения объектов. Спектр таких задач, в которых может использоваться предварительная декомпозиция, достаточно широк, и рассмотрение всех возможных вариантов выходит за рамки настоящего исследования.

Поскольку компонента s_{mn} на внутренних участках областей является гладкой, операцию по ее выделению можно формулировать как сглаживание каждой из областей изображения по отдельности и называть *сглаживанием изображения при сохранении контуров*. Нахождение компоненты, отвечающей таким условиям, можно интерпретировать как решение задачи достижения минимальной сложности изображения при максимальном сохранении необходимой информации о содержащихся в нем объектах.

Для отыскания алгоритма декомпозиции изображения будем основываться на двухмасштабной многокомпонентной модели изображения, рассмотренной в разделе 1.2. Модель является локальной, т.е. статистическая оценка сглаженного значения для элемента x_{mn} полностью задается значениями элементов, входящих в окружающую его область $V_d(x_{mn})$. Следуя одному из сформулированных выше подходов — сглаживанию каждой из областей

изображения по отдельности — для его реализации требуется найти эффективный способ вычленения из общей области анализа того подмножества элементов, которое является *областью принадлежности* элемента x_{mn} . Предварительное разбиение изображения на области невозможно, поэтому выбор множества элементов, относящихся к области принадлежности, будем осуществлять статистическими методами, базируясь на значении распределения значений элементов в области анализа $H_{mn}(V_d)$, которое часто называют *гистограммой по локальному фрагменту* или просто *локальной гистограммой*.

Как будет показано в дальнейшем, в процессе отыскания области принадлежности по $H_{mn}(V_d)$ происходит выбор некоторого ограниченного диапазона значений $[x_l, x_r]$, по которому в дальнейшем и производится оценивание. Такой подход нетрудно интерпретировать как введение локальности в области значений. Считая совокупность пространственной области определения изображения и области значений элементов как единое пространство, методы, в которых при анализе используется множество элементов, локальное по координатам двух различных типов, принято называть *билатеральными методами*.

2.1. Задача сглаживания изображения при сохранении контурных перепадов

Сглаживание — одна из наиболее распространенных операций, применяемых при обработке и анализе видеoinформации для устранения помех. К сожалению наряду с устранением помех, большинство методов сглаживания одновременно приводят к размыванию контурных перепадов. Отысканию метода, который при достаточно хорошем сглаживании протяженных деталей обеспечивал бы сохранение контурных (яркостных) перепадов между ними, посвящены многочисленные исследования. В ряде работ сглаживание формулируется как устранение шума на изображении. Среди них такие, как медианный фильтр [358], адаптивный линейный фильтр [296], метод масок [320], близкий ему метод сглаживания по выборочным соседям [355], сглаживание, обратное градиенту [363], сигма-фильтр [297]. Сравнению данных методов посвящена обзорная статья [309]. Однако во всех случаях предполагается, что сглаживание происходит по сравнительно небольшой окрестности обрабатываемого элемента, со стороной размерами не более $5 \div 7$ элементов, в то время как часто требуется сглаживание по значительно большему участку, размерами в несколько десятков и более элементов. В такой ситуации большинство упомянутых методов оказываются по тем или иным причинам неприменимыми.

Различие задач сглаживания по участкам малых и больших размеров проясняется, если воспользоваться двухмасштабной многокомпонентной моделью изображения, изложенной в первой главе. Оно заключается в том, что в первом случае сглаживанием определяется значение

составляющей μ модели окрестности (1.2.9), а во втором — компоненты s модели фрагмента (1.2.23). Нас будет интересовать вторая задача — нахождение метода сглаживания изображения по достаточно большому фрагменту, обеспечивающего при этом сохранение контуров. В связи с условием независимой обработки каждого элемента изображения и требованием построения эффективного алгоритма, ограничимся классом методов, основанных на анализе распределения вероятностей элементов по области анализа $V_d(m,n)$ — гистограмме $H_{mn}^{V_d} = [h_{mn}(k)]$:

$$h_{mn}(k) = P\{x = k \mid x \in V_d(m,n)\}; \quad 0 \leq k < K; \quad \sum_{k=0}^{K-1} h_{mn}(k) = 1. \quad (2.1.1)$$

Выбор гистограммы по фрагменту, как основного средства анализа связан с тем, что для обычной последовательной обработки элементов изображения известен простой и достаточно эффективный алгоритм пересчета гистограммы и порядковых статистик по скользящему фрагменту при переходе от элемента n к элементу $n+1$.

2.1.1. Искажения контурных перепадов при использовании арифметического среднего и медианы по фрагменту

Из сглаживающих операторов, основанных на анализе гистограммы по фрагменту W_{mn} размерами $L \times L$; $L = 2l + 1$ (для удобства будем считать фрагмент квадратным), простейшими и наиболее распространенными являются два: арифметическое среднее значений элементов

$$A(W_{mn}) = \frac{1}{L^2} \sum_{i=m-l}^{m+l} \sum_{j=n-l}^{n+l} x_{ij} = \sum_{k=0}^{K-1} k h_{mn}(k) \quad (2.1.2)$$

и порядковые статистики вида $R(q)$; $0 \leq q \leq 1$, определяемые как

$$R(q) = y, \quad \text{если} \quad \sum_{k=0}^{y-1} h_{mn}(k) < q \leq \sum_{k=0}^y h_{mn}(k), \quad (2.1.3)$$

из которых чаще всего используется медиана, т. е. значение

$$\text{med}(W_{mn}) = R(0,5). \quad (2.1.4)$$

Если фрагмент W покрывает элементы лишь некоторой одной области изображения, то, согласно модели (1.2.24), его гистограмма будет иметь нормальное распределение $N(s^1, \sigma_1^2)$. Тогда и среднее $A(W)$ и медиана $\text{med}(W)$ также будут иметь нормальное распределение со средним s^1 и дисперсиями $\sigma_A^2 = \sigma_1^2/L^2$ и $\sigma_{\text{med}}^2 = \pi \sigma_1^2/2L^2$. При этом первая оценка оказывается более точной: $\sigma_{\text{med}}^2/\sigma_A^2 = \sqrt{\pi/2} \approx 1,25$. Однако устойчивость оценок среднего и медианы к изменению формы распределения, которое происходит при попадании фрагмента на границу нескольких областей изображения, совершенно различна [70].

Если во фрагмент попадают несколько областей (U^1, \dots, U^j), то математические ожидания $M\{A(W)\}$ и $M\{\text{med}(W)\}$ будут равны некоторым промежуточным значениям внутри интервала

(s_{\min}, s_{\max}). В этом случае множество элементов фрагмента можно рассматривать как совокупность элементов нескольких (J) нормальных распределений: $N(s^1, \sigma_1^2), \dots, N(s^J, \sigma_J^2)$. Пусть p_j — вероятность того, что произвольно взятый элемент фрагмента принадлежит области U^j ; ($\sum p_j = 1$). Тогда математическое ожидание арифметического среднего по фрагменту будет равно

$$M\{A(W)\} = M\left\{\sum_W x_{ij} / L^2\right\} = \sum_{j=1}^J p_j s^j. \quad (2.1.5)$$

Математическое ожидание медианы $M\{\text{med}(W)\}$ будет равно значению z_0 , являющемуся корнем следующего уравнения:

$$\sum_{j=1}^J p_j F_j(z) = 0,5, \quad (2.1.6)$$

где $F_j(z)$ — функция распределения значений элементов из области U^j .

Сравним результаты для случая, когда фрагмент покрывает две области: U^1 и U^2 . Построим зависимости $M\{A(W)\}$ и $M\{\text{med}(W)\}$ от p_2 при различных $\Delta s = s_2 - s_1$. Такие зависимости для $\sigma_1 = \sigma_2$ показаны на Рис. 2.1.1, а.

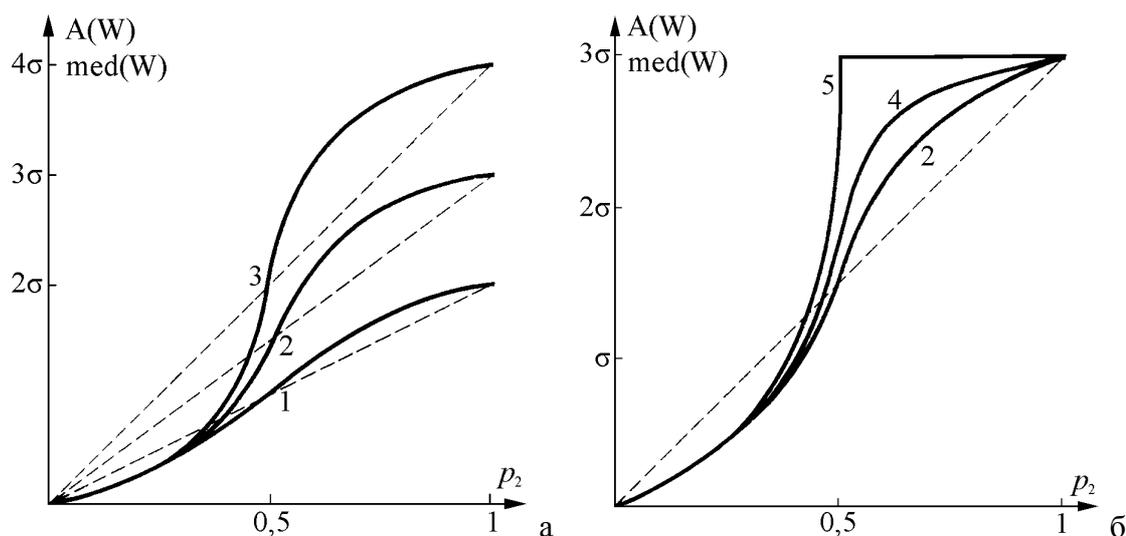


Рис. 2.1.1. Зависимости среднего (пунктир) и медианы (непрерывная линия) для двухмодового распределения в зависимости от соотношения дисперсий мод:

а) $\sigma_1 = \sigma_2$: 1 — $\Delta s = 2\sigma$, 2 — $\Delta s = 3\sigma$, 3 — $\Delta s = 4\sigma$; б) $\Delta s = 3\sigma$: 2 — $\sigma_1 = \sigma_2$, 4 — $\sigma_2 = 0,5\sigma_1$, 5 — $\sigma_2 = 0$.

При $\Delta s = 2\sigma$ значения среднего и медианы различаются слабо. С увеличением Δs производная $M\{\text{med}(W)\}$ почти не меняется на краях при $p_2 < 0,3$ и $p_2 > 0,7$, но появляется резкий перепад вблизи $p_2 \approx 0,5$, в то время как математическое ожидание среднего $M\{A(W)\}$ всегда изменяется линейно. При различающихся σ_1 и σ_2 крутизна графика медианы для малых p_2 будет определяться дисперсией σ_1 а для больших p_2 — дисперсией σ_2 ; график для среднего по фрагменту не изменится (см. Рис. 2.1.1, б).

Физический смысл полученных зависимостей в следующем. Пусть на изображении имеются две области, разграниченные контурным перепадом, значения элементов которых

описываются нормальными распределениями: $N(s^1, \sigma_1^2)$ и $N(s^2, \sigma_2^2)$. Тогда представленные графики будут показывать изменения значений среднего и медианы по скользящему фрагменту вдоль линии, перпендикулярной контурному перепаду.

Сделанные выводы подтверждаются экспериментально — сглаживанием тестового изображения локальным средним и медианным фильтром (Рис. 2.1.2, б, в).

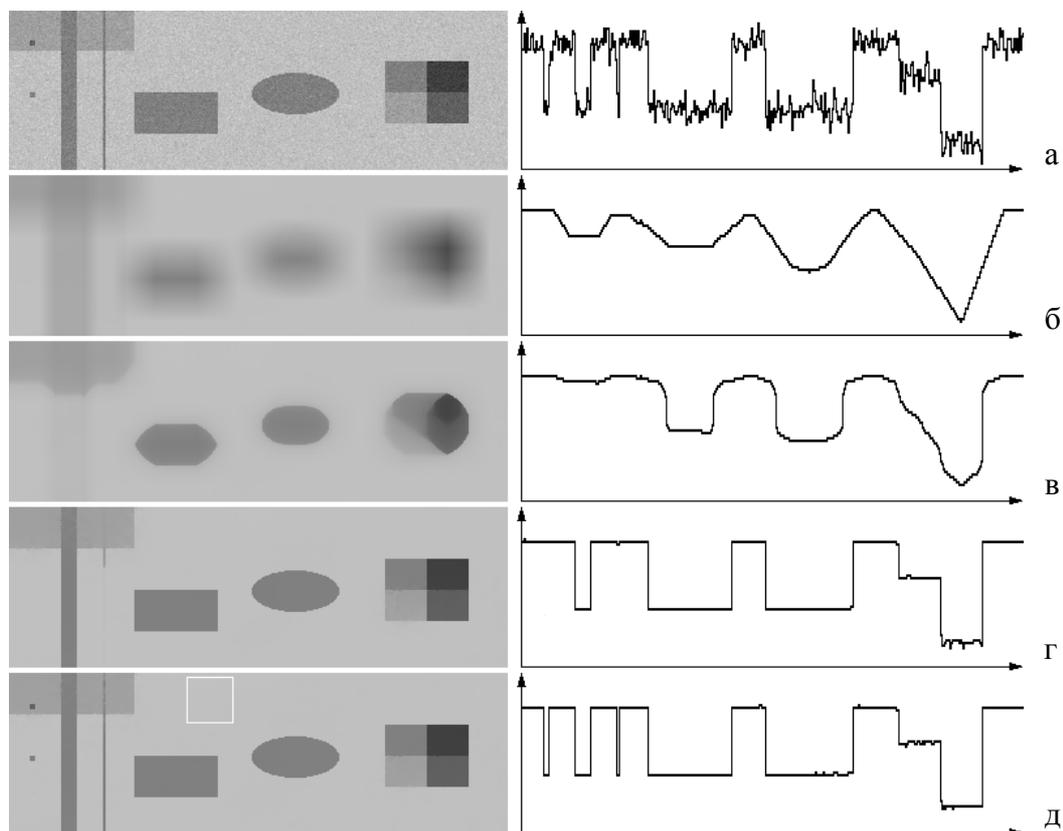


Рис. 2.1.2. Сглаживание тестового изображения (128×384 эл.); размеры фрагмента сглаживания (33×33 эл.) показаны белым контуром на (д); а) модельное изображение с шумом и график одной из его строк, б) среднее (2.1.2) и график той же строки; в) медиана (2.1.4); г) сглаживание алгоритмом декомпозиции «D» (раздел 2.2.1) с $n^W = 80$; д) то же с $n^W = 10$.

Хорошо видно, как изменились контуры и формы объектов. Среднее «размыло» все контуры на изображении, тогда как медиана сохранила протяженные линейные контуры, но «стерла» углы деталей. Эти искажения вызваны тем, что в углах число элементов объекта (его части, попадающей во фрагмент анализа) меньше числа фоновых элементов. Это подтверждает так называемые «структурные свойства» медианных фильтров [213]. Графики строк видеосигнала (а–в) показывают совпадение теоретических и экспериментальных результатов.

2.1.2. Сглаживание на основе анализа распределения

Как было сказано выше, будем рассматривать сглаживание изображения x_{mn} согласно модели (1.2.2) как задачу выделения кусочно-гладкой компоненты s_{mn} . При поиске решения

ограничимся классом методов, основанных на анализе распределения вероятностей значений элементов (гистограммы) по скользящему фрагменту W_{mn} . Типичный фрагмент, содержащий несколько областей U^j ($j = 1, \dots, J$) изображения, а также график соответствующего распределения $h_{mn}(k)$ показаны на Рис. 1.1.3. Согласно модели фрагмента (1.2.24), задачу сглаживания можно переформулировать как нахождение средней яркости s^1 части области U^1 , попадающей во фрагмент W_{mn} , содержащей центральный элемент x_{mn} . В распределении $h_{mn}(k)$ имеется J локальных максимумов (мод), число и расположение которых определяется набором средних значений s^j , а мощность — вероятностью принадлежности точки фрагмента той или иной моде, т.е. долей площади, занимаемой участком области во фрагменте W_{mn} . Элемент x_{mn} , если не является контурным, принадлежит центральной области U^1 фрагмента W_{mn} которой может соответствовать любая из мод гистограммы $H_{mn}^{V_d}$. Таким образом, задача определения значения компоненты s_{mn}^1 , или, что то же самое, средней яркости области U^1 , сводится к нахождению в гистограмме фрагмента положения моды, соответствующей U^1 .

Сделанная формулировка типична для задач распознавания и классификации, в частности использующих статистические модели [6,53,56], поэтому для решения можно было бы воспользоваться известными методами. Но поскольку операцию анализа необходимо проводить для каждого элемента изображения, выбирать необходимо среди простейших и эффективных в вычислительном отношении методов. Подходящим байесовским классификатором, основанным на сходстве между объектами, является метод ближайших соседей, а именно — метод парзеновского окна [327], в котором используется следующая функцию оценки:

$$f(u) = \frac{1}{Kh} \sum_{k=0}^{K-1} F\left(\frac{x_j - u}{h}\right), \quad (2.1.7)$$

где $F(x)$ — ядро анализа, h — параметр ширины ядра. На этой идее основаны многие алгоритмы классификации, в частности, метод ближайших соседей. При этом известно, что вид функции ядра F практически не влияет на точность классификации, а определяющим является параметр его ширины h . Указанный подход можно применить для оценивания среднего по множеству элементов области анализа следующим образом:

$$\bar{x}_{mn} = \frac{\sum_{k=0}^{K-1} x_k F(x_k - x_0)}{\sum_{k=0}^{K-1} F(x_k - x_0)}, \quad (2.1.8)$$

где $F(x)$ — весовая функция, x_0 — опорное значение для точки (m, n) .

Сигма-фильтр

В случае, когда в качестве x_0 используется значение центрального элемента области анализа x_{mn} , а весовая функция $F_i(x) = 1$ для $-\delta \leq x \leq \delta$, и $F_i(x) = 0$ для остальных x , получим известный сигма-фильтр Ли [297], являющийся частным случаем парзеновского окна [327].

Суть его сводится к тому, что из элементов окрестности V'_{mn} выбираются те значения x_{mn}^r ($r = 1, \dots, R$), которые попадают в интервал $(x_{mn} - \delta, x_{mn} + \delta)$, где x_{mn} — центральный (обрабатываемый) элемент, а δ — параметр, определяющий ширину интервала. По значениям x_{mn}^r подсчитывается арифметическое среднее:

$$\bar{x}_{mn} = A(V'_{mn}, x_{mn}, \delta) = \frac{1}{R} \sum_{r=1}^R x_{mn}^r, \quad x_{mn} - \delta \leq x_{mn}^r \leq x_{mn} + \delta, \quad (2.1.9)$$

которое и записывается как результат. Если же в интервал $x_{mn} \pm \delta$ попадает точек меньше, чем некоторое заданное r_0 , то \bar{x}_{mn} определяется как среднее по четырем ближайшим соседям:

$$\bar{x}_{mn} = (x_{m-1,n} + x_{m,n-1} + x_{m,n+1} + x_{m+1,n}) / 4; \quad (2.1.10)$$

Сигма-фильтр Ли предлагается использовать с $\delta = 2\sigma$, где σ^2 — дисперсия шума [297].

В обзоре [309] сигма-фильтр был признан одним из лучших фильтров, дающих удовлетворительное сглаживание при сохранении контуров (для небольших окрестностей усреднения, размерами в $3 \div 7$ элементов). Слабое звено сигма-фильтра в том, что в качестве опорного выбирается значение центрального элемента окрестности x_{mn} . Это является простой, но недостаточно точной оценкой для зашумленных изображений. С заметной долей вероятности разность $(\mu_{mn} - x_{mn})$ может быть большой, и в результате искомым пик распределения окажется вне диапазона $-\delta \leq x \leq \delta$. Из-за этого полученная оценка может оказаться неверной и изображение будет искажено. Влияние этого эффекта усиливается вблизи контуров.

2.2. Разработка метода декомпозиции изображения

Выбор моды, ближайшей к x_{mn} , далеко не всегда оказывается правильным. Это связано с тем, что на реальных изображениях дисперсия σ_t в (1.2.24) может быть сравнима с разницей средних яркостей двух различных областей во фрагменте $|s^1 - s^j|$ и достаточно вероятно, что ближайшей к значению окажется мода, соответствующая другой области. Анализ реальных изображений показывает, что значения СКО текстурной σ_t и шумовой σ_ξ составляющих в (1.2.23) могут различаться в 2–3 раза в ту или иную сторону в зависимости от попадания точки (m,n) внутрь однородной области или на ее границы. Более хорошей оценкой для s_{mn}^1 является значение μ_{mn} в (1.2.9), которое можно получить анализом элементов окрестности V_{mn} (согласно определению в разделе 1.2.2, в окрестность V_{mn} попадают лишь те элементы, которые относятся к той же области, что и x_{mn}). Это приводит к выводу, что в качестве предварительной оценки для s_{mn}^1 следует использовать значение μ_{mn} .

Улучшить оценку опорного значения x_0 на границах объектов, где распределение x_j имеет большую дисперсию и далеко от симметричного, можно заменой среднего по R элементам внутри окна $x_{mn} \pm \delta$ на значение медианы по тем же элементам:

$$\bar{x}_{mn} = \text{med}(V', x_{mn}, \delta) = \text{med}\{x_{mn}^1, \dots, x_{mn}^R\}; \quad (2.2.1)$$

в случае малого числа точек вместо (2.2.1) берется медиана по окрестности из пяти элементов:

$$\bar{x}_{mn} = \text{med}(x_{m-1,n}, x_{m,n-1}, x_{m,n}, x_{m,n+1}, x_{m+1,n}). \quad (2.2.2)$$

Для нахождения s_{mn}^1 при уже известном \bar{x}_{mn} формируется аналогичное парзеновское окно по фрагменту. Имея заданное Δ , выбирают те элементы фрагмента W_{mn} (или, что то же самое, точки гистограммы по фрагменту H_{mn}^W), значения которых попадают в интервал $(\bar{x}_{mn} - \Delta, \bar{x}_{mn} + \Delta)$. По выбранным элементам подсчитывается $A(W, \bar{x}, \Delta)$ или $\text{med}(W, \bar{x}, \Delta)$.

Сравним точности оценок $A(\bar{x}, \Delta)$ и $\text{med}(\bar{x}, \Delta)$ в предположении, что значения $x_{ij} \in W_{mn}$ имеют нормальное распределение $N(s^1, \sigma^2)$, т.е. предполагается, что фрагмент W_{mn} либо попадает на одну область изображения, либо соседние моды отстоят достаточно далеко и их влияния можно не учитывать. Для простоты выберем $\sigma = 1$ и обозначим:

$$\int_a^b = \int_a^b e^{-\frac{v^2}{2}} dv / \sqrt{2\pi}.$$

Математическое ожидание арифметического среднего элементов, попадающих в интервал $(\bar{x}_{mn} - \Delta, \bar{x}_{mn} + \Delta)$, равно:

$$M\{A\} = \int_{x-\Delta}^{x+\Delta} v e^{-\frac{v^2}{2}} dv / \sqrt{2\pi} \int_{x-\Delta}^{x+\Delta} = (e^{-\frac{(x-\Delta)^2}{2}} - e^{-\frac{(x+\Delta)^2}{2}}) / \sqrt{2\pi} \int_{x-\Delta}^{x+\Delta}. \quad (2.2.3)$$

Математическое ожидание медианы $M\{\text{med}\}$ по тому же интервалу совпадает с корнем уравнения:

$$\int_{x-\Delta}^{M\{\text{med}\}} = \frac{1}{2} \int_{x-\Delta}^{x+\Delta}; \quad \text{отсюда:} \quad \int_{x-\Delta}^{M\{\text{med}\}} = \frac{1}{2} \left(\int_0^{x+\Delta} - \int_0^{x-\Delta} \right). \quad (2.2.4)$$

Зависимости $M\{A, x\}$ и $M\{\text{med}, x\}$ для $\Delta = \sigma, 1,4\sigma, 2\sigma$ и 3σ представлены графиками на Рис. 2.2.1.

Определим понятие скорости сходимости оценки как отношение $\alpha_A = M\{A\}/x$ или $\alpha_{\text{med}} = M\{\text{med}\}/x$. Как видно, для всех Δ оценки $M\{A\}$ и $M\{\text{med}\}$ весьма близки, но $M\{\text{med}\}$ является несколько более точной: $\alpha_{\text{med}} > \alpha_A$. Как и следовало ожидать, при увеличении Δ наблюдается увеличение α от $1,4 \div 1,5$ при $\Delta = \sigma$, до $3 \div 5$ при $\Delta = 2\sigma$ и $15 \div 30$ при $\Delta = 3\sigma$. Следует отметить, что если в реальных распределениях оказываются достаточно близкие моды, «хвосты» распределений которых существенно перекрываются, то вероятность попадания в интервал $(\bar{x}_{mn} - \Delta, \bar{x}_{mn} + \Delta)$ заметной доли второй моды тем выше, чем больше Δ . В связи с этим не рекомендуется выбирать Δ больше, чем 2σ .

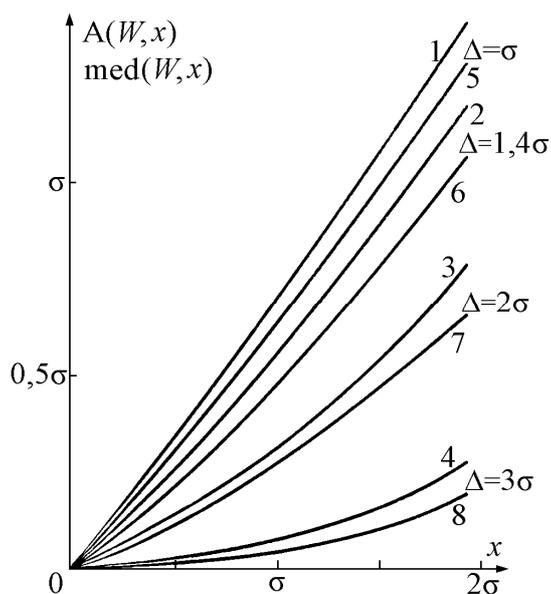


Рис. 2.2.1. Зависимости математических ожиданий среднего $A(W, x)$ и медианы $\text{med}(W, x)$ от x (кривые 1–4 и 5–8 соответственно), получаемые фильтром (2.1.9) при $\delta = \sigma, 1,4\sigma, 2\sigma$ и 3σ .

Для повышения точности определения s^1 можно использовать несколько итераций. Тогда принимается $\bar{x}_1 = \bar{x}$, а на очередном шаге $k+1$ при уже найденном \bar{x}_k определяется \bar{x}_{k+1} . Это позволяет повысить первоначальную скорость сходимости приблизительно до α^k , где k — число итераций. Так, при $\Delta = 2\sigma$ и $k = 3$ получим $\alpha \approx 100$. Можно также предложить другой итерационный подход. Заметим, что при последовательных итерациях у нас не меняется знак разности $(\bar{x}_k - s^1)$, т. е. если $\bar{x}_k > s^1$, то и $\bar{x}_{k+1} > s^1$, и наоборот, если $\bar{x}_k < s^1$, то и $g_{mn}^r < s^1$, а также, что зависимости $M\{A\}$ и $M\{\text{med}\}$ от x при $x \approx 1 \div 2\sigma$ почти линейны. Тогда, выбрав некоторое среднее значение $\alpha(\Delta)$, можно находить значение \bar{x}_{k+2} по следующей формуле:

$$\bar{x}_{k+2} = \bar{x}_{k+1} - \alpha(\Delta)(\bar{x}_k - \bar{x}_{k+1}) \quad (2.2.5)$$

При итерационном подходе на первом шаге следует использовать меньшее значение Δ (порядка $1 \div 1,2\sigma$), а на втором шаге увеличить Δ до 2σ ; тогда в принципе достаточно двух итераций.

О размере центральной области

При рассмотрении вопроса определения среднего уровня яркости s^1 центральной области U^1 нигде не учитывались ее размеры, т. е. число элементов n_1 , в ней содержащихся. Тем не менее этот вопрос существен, поскольку при достаточно малых n , возможно, область U^1 не следует выделять как самостоятельную и нужно отнести к текстуре, а в качестве сглаженного значения в данных точках необходимо принять некоторое другое s^2 , являющееся средним по области U^2 , окружающей U^1 . Обозначим такой порог разделения по числу элементов n_p ; выбор его, вообще говоря, является неформальной задачей и во многом зависит от конкретного сюжета, конечной цели обработки и степени дискретизации изображения.

Величину порога n_p нетрудно ввести как параметр в алгоритм сглаживания следующим образом. Пусть имеются гистограмма H_{mn}^W , заданное n_p и первоначальная оценка среднего уровня яркости \bar{x}_{mn} . Определим по n_p значения порядковых статистик $R_1 = R(n_p/L^2)$ и $R_2 = R(1 - n_p/L^2)$. Опорное значение \bar{x}_1 для первой итерации выбирается алгоритмом:

$$\bar{x}_1 = \begin{cases} \bar{x}, & \text{если } R_1 \leq \bar{x} \leq R_2; \\ R_1, & \text{если } \bar{x} < R_1; \\ R_2, & \text{если } \bar{x} > R_2. \end{cases} \quad (2.2.6)$$

Заметим, что при использовании вместо H_{mn}^W и \bar{x} гистограммы по окрестности H_{mn}^V и значения центрального элемента x_{mn} , данный алгоритм (2.2.6) становится фильтром импульсных помех.

2.2.1. Алгоритм декомпозиции «D»

На основе изложенного можно сформулировать метод сглаживания изображения, сохраняющий контурные перепады, в виде следующего алгоритма декомпозиции [192], который при заданных размерах $l \times l$ окрестности V_{mn} и $L \times L$ ($l < L$) фрагмента W_{mn} , центрированных в обрабатываемой точке (m, n) , ширине яркостного интервала анализа Δ^V и Δ^W , а также параметрах размеров $n^V < l^2/2$ и $n^W < L^2/2$ соответственно, позволяет находить значение S_{mn} в (1.2.23). Алгоритм включает следующие операции, выполняемые для каждой точки (m, n) изображения:

1. Подсчитываются гистограммы по окрестности H_{mn}^V и фрагменту H_{mn}^W с центром в (m, n) .
2. По гистограмме окрестности H_{mn}^V и заданному значению n^V находятся ранговые границы $R_1^V = R^V(n^V/l^2)$ и $R_2^V = R^V(1 - n^V/l^2)$, где $R(x)$ определяется как решение уравнения $\sum_{i=0}^{R(x)} h_{mn}^V(i) = x$. Сравнением значения x_{mn} с R_1^V и R_2^V определяется промежуточное усеченное значение \tilde{x}^V :

$$\tilde{x}^V = \begin{cases} x_{mn}, & \text{если } R_1^V \leq x_{mn} \leq R_2^V; \\ R_1^V, & \text{если } x_{mn} < R_1^V; \\ R_2^V, & \text{если } x_{mn} > R_2^V. \end{cases}$$

3. С помощью фильтра (2.1.9), являющегося частным случаем парзеновского окна [327], находится урезанное среднее \bar{x}_{mn} . Для этого из элементов окрестности V_{mn} выбираются те n значений $x_{mn}^r \in V_{mn}$ ($r = 1, \dots, n$), которые попадают в интервал $(\tilde{x}^V - \Delta^V, \tilde{x}^V + \Delta^V)$, где Δ^V — полуширина интервала. По значениям x_{mn}^r в данном интервале подсчитывается среднее:

$$\bar{x}_{mn} = A(V_{mn}, x_{mn}, n^V, \Delta^V) = \frac{1}{n} \sum_{r=1}^n x_{mn}^r, \quad \tilde{x}^V - \Delta^V \leq x_{mn}^r \leq \tilde{x}^V + \Delta^V. \quad (2.2.7)$$

4. Аналогично п. 2, по гистограмме фрагмента H_{mn}^W и заданному n^W находятся ранговые границы $R_1^W = R^W(n^W/L^2)$ и $R_2^W = R^W(1 - n^W/L^2)$. Сравнением значения \bar{x}_{mn} с R_1^W и R_2^W определяется усеченное значение \tilde{x}^W :

$$\tilde{x}^W = \begin{cases} \bar{x}_{mn}, & \text{если } R_1^W \leq \bar{x}_{mn} \leq R_2^W; \\ R_1^W, & \text{если } \bar{x}_{mn} < R_1^W; \\ R_2^W, & \text{если } \bar{x}_{mn} > R_2^W. \end{cases}$$

5. Сглаженное значение s_{mn} находится по гистограмме фрагмента H_{mn}^W как среднее или медиана значений H_{mn}^W , попадающих в интервал $(\tilde{x}^W - \Delta^W, \tilde{x}^W + \Delta^W)$:

$$s_{mn} = A(W_{mn}, \bar{x}_{mn}, n^W, \Delta^W) \quad \text{или} \quad s_{mn} = \text{med}(W_{mn}, \bar{x}_{mn}, n^W, \Delta^W). \quad (2.2.8)$$

Полученное значение считается искомой сглаженной компонентой S_{mn} в (1.2.23). Возможно дополнительное уточнение значения s_{mn} несколькими шагами итераций, используя в качестве \tilde{x}^W значение s_{mn} , получаемое на предыдущем шаге. Выбор между значением урезанного среднего или медианы в (2.2.8) можно сделать исходя из того, что вычисление среднего проще, но медиана дает более точную оценку вблизи контурных перепадов. Касаясь вычислительной эффективности приведенного алгоритма можно отметить, что для его реализации достаточно в среднем $C_1(l + L) + C_2$ операций на один элемент изображения, где C_1 и C_2 — константы, а l и L — число строк в окрестности V и фрагменте W .

Результаты обработки тестового изображения описанным алгоритмом при различных n_p показаны на Рис. 2.1.2, г, д. В первом случае (г) параметр n_p выбран относительно большим и на изображении удалились как шум, так и мелкие детали. Во втором случае (д) при малом n_p изображение сглажено, но все детали сохранены, несмотря на то что мелкие детали имели размеры 4×4 элемента, а фрагмент сглаживания — 33×33 элемента. Точность восстановления оригинала (без шума) после первой итерации, подсчитанная по среднеквадратическому критерию, составила 0,3%. На графиках видна хорошая степень сглаживания шума при полном сохранении контурных перепадов. После второй итерации изображение–результат почти полностью совпадает с исходным изображением. Уровень шума, измеренный согласно (1.2.15), на исходном изображении (а) составлял 8,15, а на сглаженном (д) — 0,32 градации яркости.

В Таблице 2.2.1 приведены значения стандартного отклонения (СКО) тестового кусочно-постоянного изображения, искаженного гауссовым шумом величиной $\sigma = 10$ градаций яркости (Рис. 2.1.2, а), от незашумленного оригинала в зависимости от размера фрагмента сглаживания. Приведены значения, полученные следующими алгоритмами сглаживания: «Среднее» — локальное арифметическое среднее (2.1.2), «Медиана» — локальная медиана (2.1.4), « σ -фильтр» — сигма-фильтр (2.1.9), «Декомп.» — алгоритм декомпозиции «D» (2.2.8), «Декомп.(3)» — 3

итерации алгоритма декомпозиции «D». Те же самые значения изображены в виде зависимостей на графиках Рис. 2.2.2.

Таблица 2.2.1. Значения стандартных отклонений результатов сглаживания (в единицах градаций яркости) разными методами зашумленного изображения от неискаженного оригинала.

Размер фрагмента	Среднее	Медиана	σ -фильтр	Декомп.	Декомп. (3)
3×3	6,06	4,66	4,59	–	–
5×5	7,20	5,26	4,05	2,83	1,82
9×9	9,20	6,14	3,53	2,05	1,39
15×15	11,49	7,79	3,44	1,80	1,25
25×25	14,81	12,64	3,54	1,78	1,23

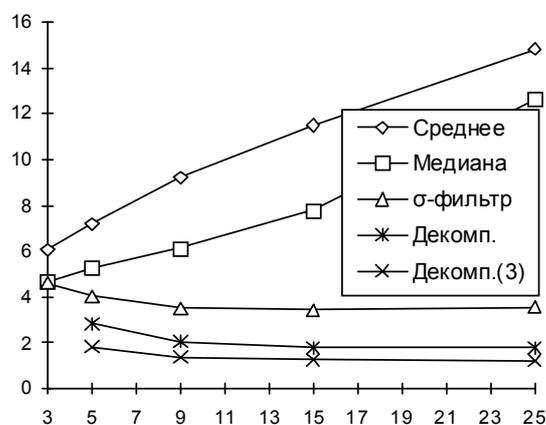


Рис. 2.2.2. Зависимости стандартных отклонений результатов сглаживания изображения с шумом разными методами от незашумленного оригинала.

Приведенные в Таблице 2.2.1 и на Рис. 2.2.2 данные показывают, что алгоритм декомпозиции «D» восстанавливает зашумленный двумерный кусочно-постоянный сигнал лучше остальных методов, а ближайший из остальных сигма-фильтр дает результаты в 2÷2,5 раза хуже. Увеличение размеров фрагмента сглаживания более 9×9 элементов почти не меняет точности восстановления декомпозиции, тогда как сглаживание локальным средним или медианой только ухудшает результаты восстановления сигнала. Тем самым можно сделать вывод, что в отношении изображений, удовлетворяющих кусочно-постоянной модели, операция сглаживания при помощи алгоритма декомпозиции «D» соответствует лучшему восстановлению исходного изображения из зашумленного сигнала.

Примеры обработки реального изображения (классического снимка «Lenna») различными алгоритмами сглаживания с одинаковыми размерами сглаживающего фрагмента (15×15 элементов) — локальным средним, локальной медианой и алгоритмом декомпозиции «D»

показаны на Рис. 2.2.3. Уровень шума (1.2.15) с 6,13 градации яркости на исходном изображении (а) уменьшился до 0,98 на сглаженном (б), обработанном алгоритмом декомпозиции «D».

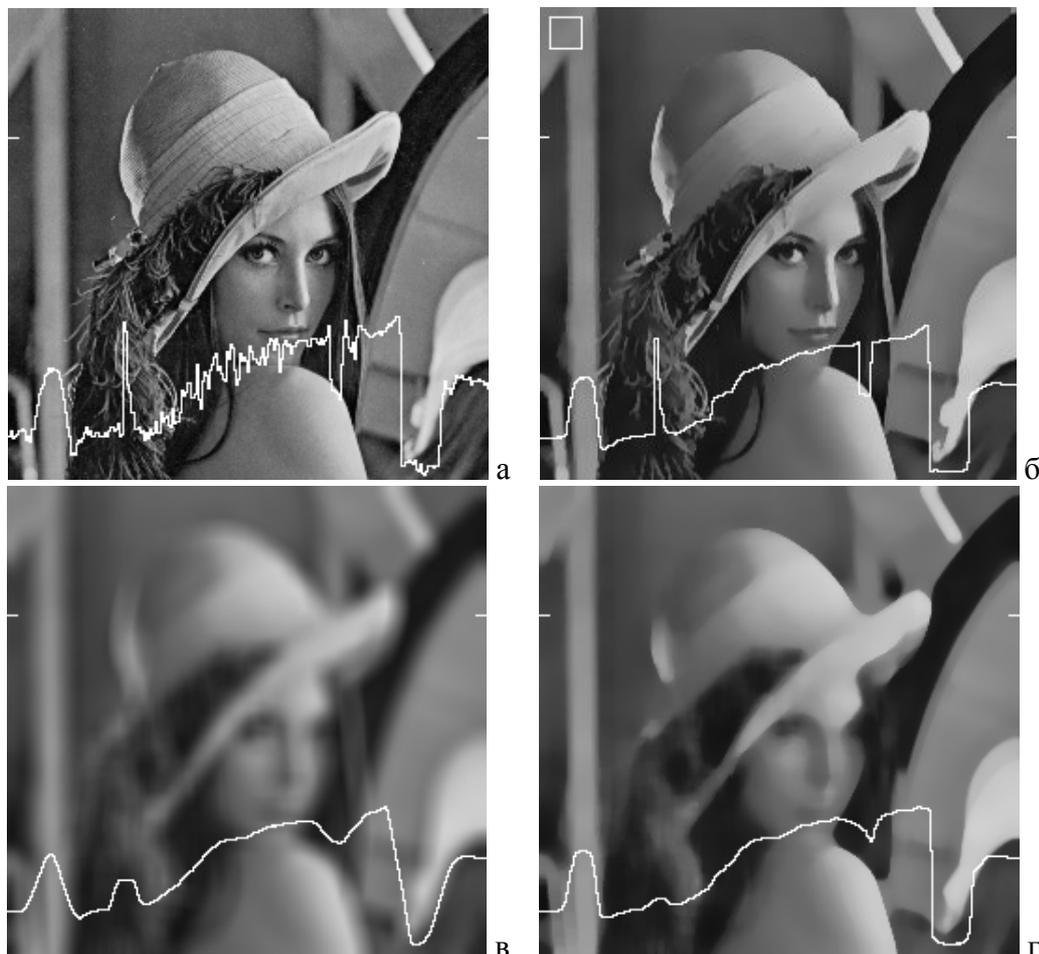


Рис. 2.2.3. Декомпозиция изображения (показаны профили отмеченной строки): а) исходное изображение (256×256 эл.); б) сглаженная алгоритмом «D» компонента s_{mn} (фрагмент 15×15 эл. показан вверху); в) локальное среднее и г) локальная медиана по фрагменту того же размера.

Сглаживание реальных изображений показывает, что наряду с удалением текстуры, шума и мелких деталей полностью сохраняются контурные перепады между протяженными областями, которые становятся даже несколько более резкими. Данный эффект вызван тем, что алгоритм декомпозиции «D» как бы «приписывает» элементы протяженного контурного перепада одной или другой граничной области изображения в зависимости от близости яркостных значений. Рассмотрим следующий простой пример. Возьмем две области U^1 и U^2 , разделенные протяженным контурным перепадом (Рис. 2.2.4,а); s^1 и s^2 — их средние яркости, причем $s^1 > s^2$, а элементы контурного перепада имеют промежуточные яркости из интервала (s^2, s^1) . При обработке такого сигнала алгоритмом «D» элементы со значениями $x(n) > (s^1 + s^2)/2$ после итерационного приближения будут стремиться к s^1 тем точнее, чем ближе к нему была исходная величина $x(n)$. Соответственно, элементы $x(n) < (s^1 + s^2)/2$ будут стремиться к s^2 . График сигнала после сглаживания $s(n)$ показан пунктиром (б). Кроме повышения резкости

контурных перепадов, данный эффект в некоторых случаях может приводить к появлению ложных контуров на участках с плавным изменением яркости.

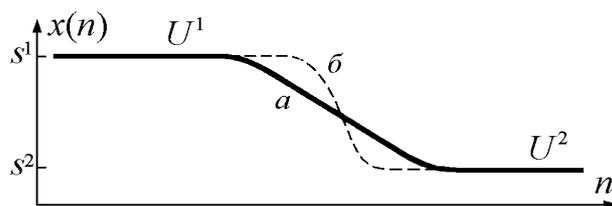


Рис. 2.2.4. К объяснению причины увеличения резкости при сглаживании изображения:
а) исходный сигнал; б) сигнал после обработки алгоритмом «D».

Основной причиной возникновения данного эффекта является то, что алгоритм сглаживания построен на основе кусочно-постоянной модификации модели фрагмента (1.2.24), которая, естественно, не в состоянии описывать наклонные участки с плавным яркостным перепадом. Рецепт здесь может быть существенное уменьшение размеров фрагмента W_{mn} при попадании его на наклонные участки. Задать какую-нибудь функцию изменения величины фрагмента в зависимости от наклона несложно, но это требует дополнительного анализа. Тем не менее, можно предложить модификацию алгоритма сглаживания, которая не повышает резкости контурных перепадов. Изменения сводятся к тому, что нужно диагностировать контурный перепад, оценивая величину g_{mn} (1.2.7), и в случае контура, вместо сглаженного значения по фрагменту s_{mn} брать в качестве результата значение \bar{x}_{mn} (2.2.1), являющееся сглаженным по окрестности V_{mn} . Модифицированный соответствующим образом алгоритм декомпозиции (без повышения резкости) рассмотрен в следующем разделе.

2.2.2. Алгоритм декомпозиции «D*» без повышения резкости

Модификация алгоритма декомпозиции «D» сводится к следующему.

1. Алгоритмом декомпозиции «D» (2.2.8) находится сглаженное значение S_{mn} .
2. Определяется амплитуда вектора g_{mn} (1.2.7) по окрестности V_{mn} .
3. Значение g_{mn} сравнивается с порогом δ . Если $g_{mn} < \delta$, то элемент x_{mn} считается внутренним элементом области (фоновым), если же $g_{mn} \geq \delta$, то контурным; это соответствует оператору:

$$\hat{S}_{mn} = S_{mn}, \text{ если } g_{mn} < \delta, \text{ и } \hat{S}_{mn} = \bar{x}_{mn}, \text{ если } g_{mn} \geq \delta. \quad (2.2.9)$$

Полученное значение \hat{S}_{mn} считается новым значением сглаженного изображения в точке (m, n) .

На Рис. 2.2.5 приведены результаты обработки реального изображения алгоритмами декомпозиции «D» (Рис. 2.2.5,б), а также алгоритмами сглаживания средним (2.1.2) и медианой (2.1.4) по локальному фрагменту (Рис. 2.2.5,в,г). На Рис. 2.2.5,д,е,ж,з показаны усиленные разности исходного и сглаженного изображения $t_{mn} = x_{mn} - s_{mn}$, где s_{mn} получена локальным

средним (2.1.2), локальной медианой (2.1.4), алгоритмом декомпозиции «D» (2.2.8) и алгоритмом декомпозиции «D*» (2.2.9).

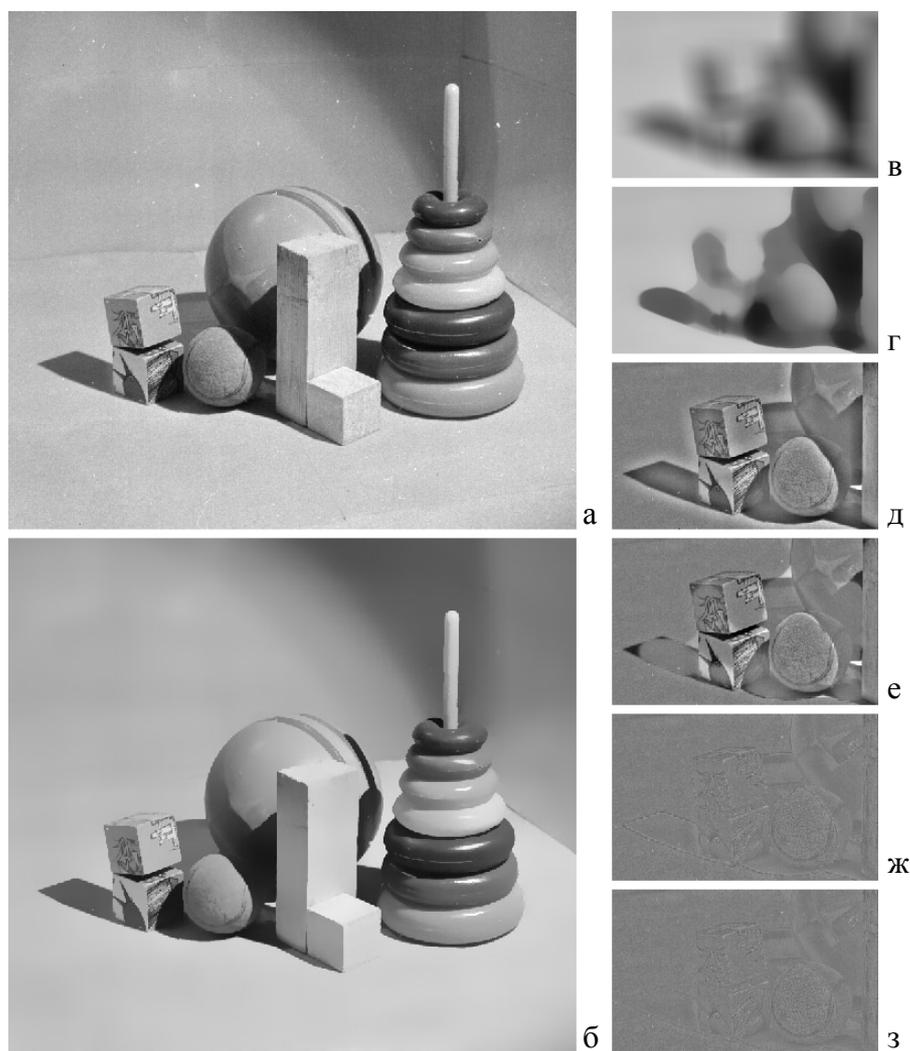


Рис. 2.2.5. Пример сглаживания изображения: а) исходное; б) сглаживание алгоритмом «D»; в) участок изображения, сглаженный локальным средним (2.1.2); г) то же, медианой (2.1.4); д-з) усиленные разности исходного и сглаженного изображений: д) сглаживание локальным средним; е) медианой; ж) алгоритмом «D»; з) алгоритмом «D*» (сравнить границы тени).

Полученные разностные сигналы (Рис. 2.2.5, в-з) содержат всю ту информацию, которая отнесена к текстурной компоненте t_{mn} . Как видно, в (д) и (е), кроме собственно текстуры и мелких деталей, присутствуют также все контуры протяженных объектов, что говорит о плохом качестве сглаживания изображения. Уровень шума, измеренный согласно (1.2.15), на исходном изображении (а) составлял 4,62, а на сглаженном (б) — 0,39 градации яркости.

Яркость точек, возникающих на текстурной компоненте t_{mn} вблизи границ объектов, постепенно уменьшается от (д) до (ж), и на (з) уже практически полностью отсутствует. Полученные результаты подтверждают тот вывод, что алгоритмом декомпозиции «D*» удастся правильно разделить исходное изображение на гладкую фоновую и текстурно-детальную

составляющие, решив таким образом задачу сглаживания изображения по большому фрагменту при полном сохранении контурных перепадов.

Упрощенный вариант построения алгоритма сглаживания изображения на основе метода парзеновского окна (2.1.8) с использованием весовой функции $w(x)$ гауссовой формы и одного уровня оценивания s_{mn} без предварительной оценки по окрестности и без введения ранговых ограничений стал впоследствии известен под названием *билатеральная фильтрация* [356].

На основе полученных результатов, сформулированных в виде алгоритмов декомпозиции «D» и «D*», может быть построен целый ряд различных методов и алгоритмов обработки и анализа изображений. Некоторые из них рассмотрены в Главах 3 и 4; кроме того алгоритм декомпозиции использован при построении алгоритмов сегментации изображения (Глава 5). Вопросы модификации алгоритма декомпозиции для применения к трехмерным изображениям изложены в Главе 7.

2.2.3. О компонентах, формируемых при декомпозиции

Представляет интерес сравнение двух способов разделения изображения на компоненты: рассмотренного алгоритма декомпозиции изображения и классического варианта частотной фильтрации [26,113], также позволяющего разделить изображение на две составляющие — низкочастотную и высокочастотную.

Для выполнения частотной фильтрации изображения первоначально вычисляется его спектр при помощи двумерного дискретного преобразования Фурье (ДПФ) по формуле

$$F_{uv} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{mn} e^{-i2\pi(um/M+vn/N)}, \quad (2.2.10)$$

где x_{mn} — исходное изображение размерами $M \times N$, а F_{uv} — его Фурье-спектр. Нужная нам фильтрация соответствует умножению спектра на маску:

$$F'_{uv} = F_{uv} \times Z_{uv}. \quad (2.2.11)$$

Для выделения низких частот маска Z_{uv} должна иметь вид:

$$Z_{uv} = \begin{cases} 1, & \text{если } u \leq M/L \text{ и } v \leq N/L; \\ 0, & \text{если } u > M/L \text{ и } v > N/L. \end{cases} \quad (2.2.12)$$

Здесь L — размер фрагмента, используемый при выполнении декомпозиции по алгоритму (2.2.8). Такой выбор порога обеспечивает согласованность степени сглаживания одним и другим способом, что позволяет сравнивать получаемые результаты.

Обратным дискретным преобразованием Фурье по формуле

$$y_{low}(m,n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F'_{uv} e^{i2\pi(um/M+vn/N)} \quad (2.2.13)$$

получим низкочастотную составляющую исходного изображения x_{mn} . Высокочастотную составляющую $y_{high}(m,n)$ можно получить либо инвертируя маску Z_{uv} в (2.2.12), либо взяв разность x_{mn} и $y_{low}(m,n)$:

$$y_{high}(m,n) = x_{mn} - y_{low}(m,n). \quad (2.2.14)$$

Результаты, близкие к частотной фильтрации, можно достичь также при помощи операций в пространственной области. Сглаживание на основе локального среднего (2.1.2) соответствует преобразованиям (2.2.10)–(2.2.13), когда в качестве Z_{uv} в (2.2.12) выбирается функция вида $\sin(x)/x$:

$$Z_{uv} = \sin(uL/M)\sin(vL/N)/[uvL^2/(MN)]. \quad (2.2.15)$$

Здесь предполагается, что используется фрагмент тех же размеров, т.е. $L \times L$.

Примеры частотного и пространственного разделения изображения на компоненты представлены на Рис. 2.2.6. Исходное изображение (аэрофотоснимок) Рис. 2.2.6,а обрабатывалось тремя разными алгоритмами, в результате чего получены следующие варианты: пространственно сглаженное локальным арифметическим средним по формуле (2.1.2) изображение (Рис. 2.2.6,б), низкочастотная компонента (Рис. 2.2.6,г), полученная фильтрацией по формулам (2.2.10)–(2.2.13), а также кусочно-гладкая компонента (Рис. 2.2.6,е), полученная алгоритмом декомпозиции «D» (2.2.8). Разностные (высокочастотные) компоненты после частотной фильтрации (2.2.10)–(2.2.13) и декомпозиции (2.2.8) представлены на Рис. 2.2.6,в и Рис. 2.2.6,д соответственно. Сравнивая получаемые компоненты можно сделать вывод, что алгоритмом декомпозиции «D» обеспечивается наилучшее разделение исходного сигнала на компоненту с протяженными объектами и компоненту с мелкими деталями.

В отличие от частотной фильтрации, при декомпозиции «D» не происходит разделения сигнала на так называемые высокочастотные и низкочастотные компоненты. И если текстурно-детальную компоненту t_{mn} в модели (1.2.2) после декомпозиции еще можно называть высокочастотной (в том смысле, что низких частот в ней действительно не содержится), то сглаженную компоненту S_{mn} называть низкочастотной никак нельзя, поскольку высокочастотные спектральные составляющие в ней частично остаются. Последнее утверждение очевидно хотя бы из того, что в кусочно-гладкой компоненте S_{mn} сохраняются резкие контурные перепады, которые дают вклад во все пространственные частоты.

И сглаженная, и текстурно-детальная компоненты могут быть использованы для последующего анализа или обработки изображения — как совместно, так и по отдельности. Отдельные варианты их использования рассмотрены в следующих главах.

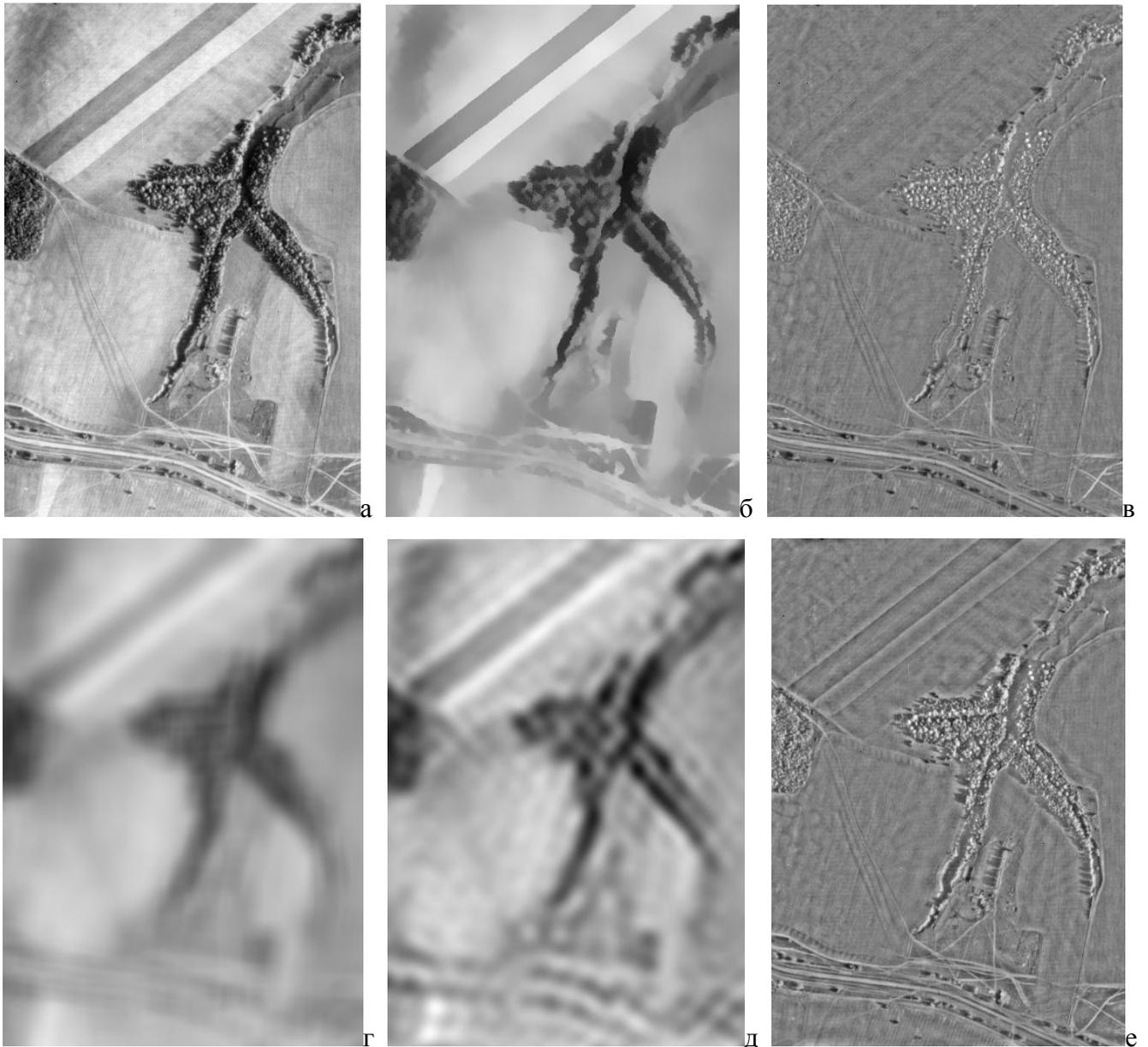


Рис. 2.2.6. Разделение на компоненты. а) Исходное изображение. Декомпозиция: б) кусочно-гладкая в) текстурно-детальная компоненты. г) Пространственное разделение — локальное среднее (2.1.2). Частотное разделение: д) низкочастотная, е) высокочастотная компоненты.

Еще один пример выделения и простейшего использования текстурно-детальной компоненты для анализа показан на Рис. 2.2.7. На исходное изображение с текстурой березовой коры, взятой из альбома Бродатца [221], была аддитивно наложена искусственная картина из четырех пятен разной формы и яркости (Рис. 2.2.7,а). Из смоделированного таким образом изображения x_{mn} алгоритмом декомпозиции «D» были выделены сглаженная компонента S_{mn} (Рис. 2.2.7,б) и текстурно-детальная компонента t_{mn} (1.2.2) как разность исходного изображения x_{mn} и сглаженной компоненты S_{mn} (Рис. 2.2.7,в). На таком изображении детали текстуры уже легко обнаруживаются простым пороговым оператором (Рис. 2.2.7,г), после чего, например, их

нетрудно пересчитать, определить процент занимаемой ими площади, или провести какой-то другой интересующий анализ.

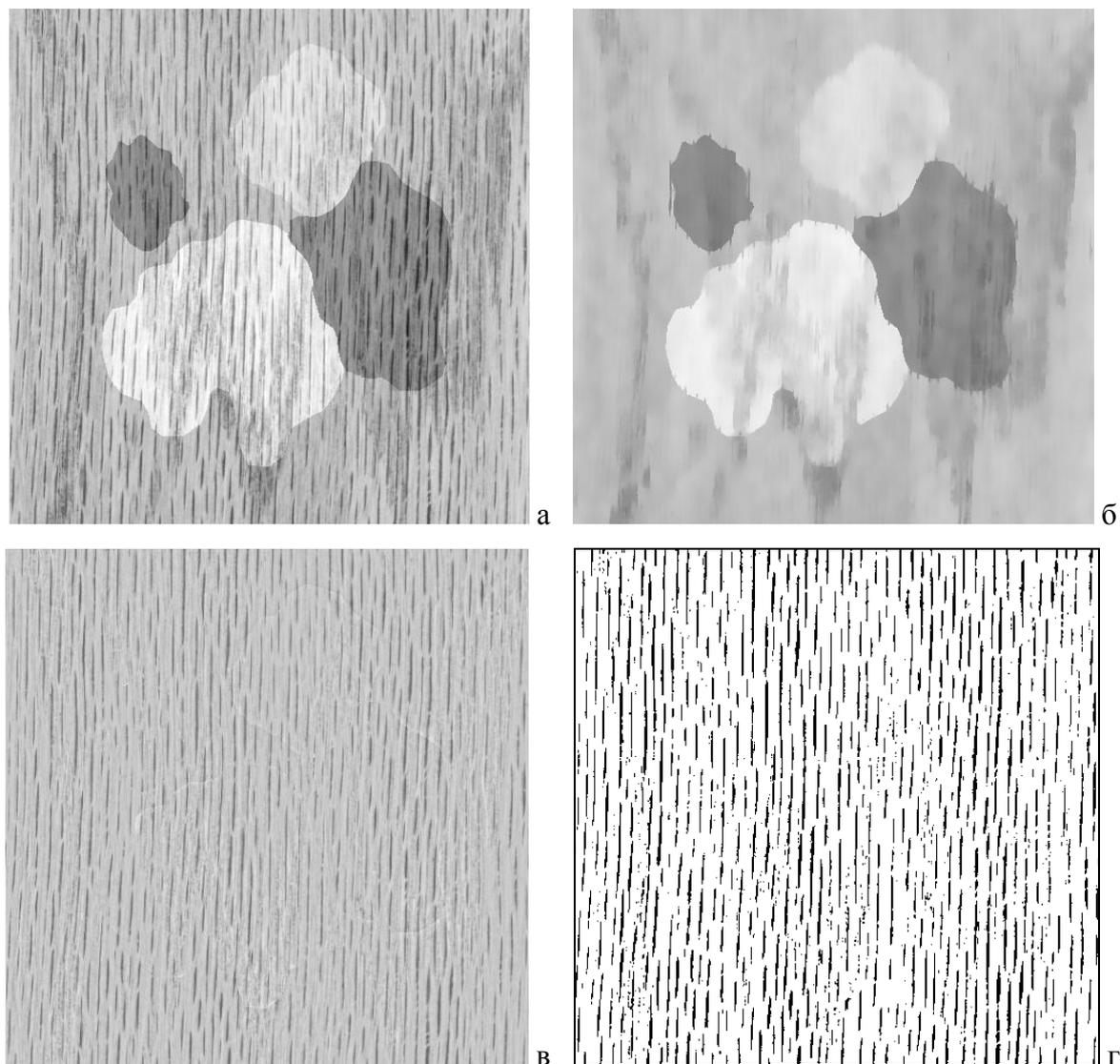


Рис. 2.2.7. Выделение текстурной компоненты: а) исходное изображение, полученное суммированием реальной текстуры с искусственным кусочно-постоянным фоном; б) сглаженная компонента S_{mn} ; в) текстурно-детальная компонента t_{mn} ; г) пороговое обнаружение текстурных деталей по изображению (в).

Рассматривая алгоритм декомпозиции, основанный на анализе распределения вероятностей значений элементов по фрагменту, следует сделать следующее замечание. Если в локальной области анализа содержатся две или более не граничащие между собой области, имеющие близкие средние яркости (например, $S^1 \approx S^2$ для областей U^1 и U^2 на Рис. 1.1.3,а), то на анализируемом распределении (Рис. 1.1.3,б) они окажутся неразличимы. Вследствие этого значение сглаженной компонентой s_{mn} , как результат алгоритма декомпозиции «D» (2.2.8), будет являться оценкой среднего для смеси значений элементов областей U^1 и U^2 , попадающих во фрагмент анализа. На уровне рассмотренного алгоритма такая ситуация неразрешима.

Преодолеть ее возможно следующим образом. На получаемой в результате алгоритма декомпозиции (при малом размере фрагмента W) сглаженной компоненте s_{mn} находятся линии границ между областями (см. Раздел 4.3), что позволяет индексировать области и принадлежащие им элементы. Это дает принципиальную возможность для алгоритма декомпозиции считать локальную гистограмму H_{mn}^W не по всему фрагменту, а лишь по тому набору входящих во фрагмент элементов, которые имеют тот же индекс, что и центральный элемент x_{mn} . Однако такой подход существенно усложняет алгоритм, и к тому же для многих реальных случаев разница получаемых результатов не является принципиальной.

2.3. Декомпозиция цветного изображения

Согласно модели цветного изображения (1.2.27), рассмотренной в разделе 1.2.4, значение элемента изображения следует рассматривать как вектор \mathbf{x}_{mn} в Z -мерном пространстве цветовых координат, например трехмерном (**RGB**):

$$\mathbf{x}_{mn} = \mathbf{S}_{mn} + \mathbf{t}_{mn}, \quad (2.3.1)$$

где \mathbf{S}_{mn} — значение сглаженной компоненты в точке (m,n) , а \mathbf{t}_{mn} — разностный вектор, несущий яркостно-цветовую информацию о текстуре, мелких деталях и шуме. Рассмотрим особенности определения значения вектора \mathbf{S}_{mn} на примере трехмерного цветового пространства (**RGB**).

Множество возможных значений элементов цветного изображения составляет цветовой куб размерами $K \times K \times K$, каждая точка которого с координатами (r,g,b) соответствует точкам на изображении со значениями цветовых составляющих $\{r,g,b\}$. Цветовой куб пространства (**RGB**) изображен на Рис. 2.3.1. Диагональное его сечение плоскостью составляет так называемый *треугольник Максвелла*, также показанный на рисунке. Каждая точка данного треугольника имеет уникальные цветность и насыщенность по отношению к остальным точкам. Главная диагональ, обозначенная вектором \mathbf{L} , представляет множество оттенков серого цвета от черного в начале координат до белого [40].

Гистограмма по фрагменту W_{mn} для одноканального изображения задается формулой (1.1.10). В случае цветного изображения она представляется следующим образом:

$$h_{mn}(r,g,b) = P(x = \{r,g,b\} | x \in W_{mn}); \quad \sum_{r=0}^{K-1} \sum_{g=0}^{K-1} \sum_{b=0}^{K-1} h_{mn}(r,g,b) = 1. \quad (2.3.2)$$

Гистограмма $H(W_{mn})$ также занимает цветовой куб, каждая точка которого (r,g,b) содержит вероятность того, что элемент $x_{ij} \in W_{mn}$ фрагмента имеет значение данной точки. Для фрагмента, изображенного на Рис. 1.1.3,а и захватывающего части трех областей, гистограмма будет выглядеть как показано на Рис. 2.3.1. Модам одномерной гистограммы черно-белого изображения (пикам распределения $h(k)$ на Рис. 1.1.3,б) здесь соответствуют сгущения

ненулевых точек U^1, U^2, U^3 , отвечающие тем частям областей, которые попадают во фрагмент, а центры мод в позициях S^1, S^2 и S^3 представляются векторами центров сгущений: S^1, S^2, S^3 .

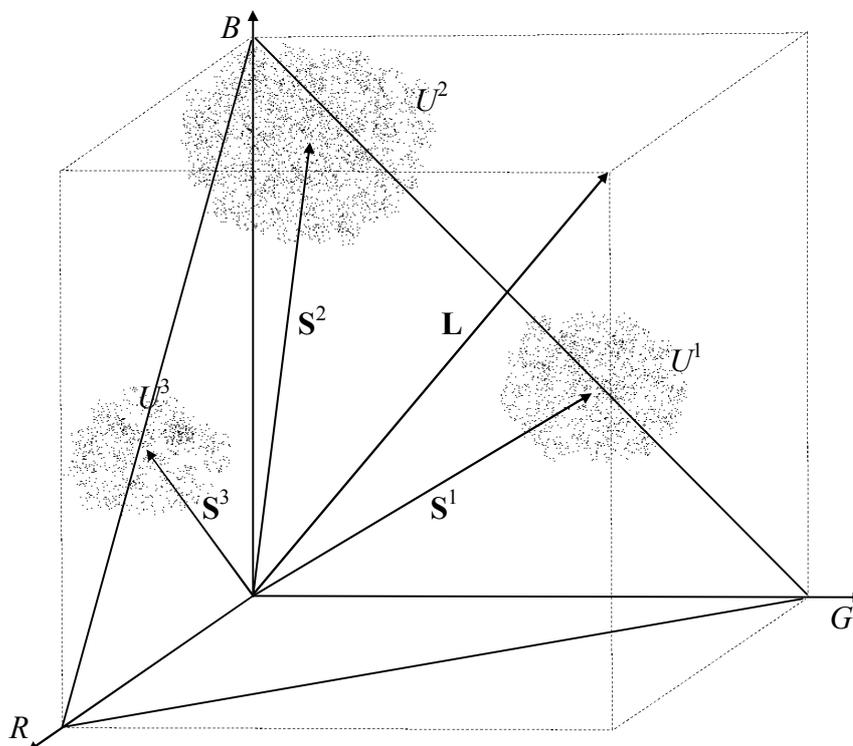


Рис. 2.3.1. Трехмерное цветовое пространство и гистограмма $H(R, G, B)$ по фрагменту W_{mn} .

Декомпозиция цветного изображения, т.е. определение значения сглаженной компоненты S_{mn} , подобно нахождению значения гладкой компоненты S_{mn} одноканального изображения, которое рассмотрено в разделе 2.1.2. Оно заключается в выборе в гистограмме $H(W_{mn})$ сгущения, соответствующего центральной области U^1 фрагмента, и определении положения центра сгущения, обозначенного вектором S^1 , который и считается сглаженным значением S_{mn} цветного изображения в точке (m, n) .

Алгоритм декомпозиции Z -канального изображения X строится по той же схеме, что и алгоритм декомпозиции «D» (2.2.8) одноканального изображения. Отличие в том, что в качестве яркостных интервалов анализа Δ^V и Δ^W должны использоваться *гиперсферы* соответствующих радиусов с центром в точке x_{mn} или \bar{x}_{mn} , а также в том, что в (2.2.8) следует использовать среднее

$$s_{mn} = A(W_{mn}, \bar{x}_{mn}, n^V, \Delta^W). \quad (2.3.3)$$

Медиану использовать невозможно, поскольку вариационный ряд можно построить только для скалярных значений; для векторных же значений x_{mn} потребуется выбрать какую-то функцию отображения вектора в скаляр. Ранговые параметры n^V и n^W можно интерпретировать как число

точек, отстоящих от центральной оценки далее, чем некоторое расстояние R^V или R^W , но это приводит к обременительным вычислениям.

Для ряда оценок по множеству окружающих элементов удастся предложить многомерные аналоги. В предположении изотропности Z -мерного пространства, вместо модуля разности значений следует использовать норму разности векторов. Общее выражение для вычисления среднего по множеству окружающих элементов изображения (2.1.8) будет следующим:

$$\bar{\mathbf{x}}_{mn} = \frac{\sum_{z=1}^Z \mathbf{x}_z w(\|\mathbf{x}_z - \mathbf{x}_0\|)}{\sum_{z=1}^Z w(\|\mathbf{x}_z - \mathbf{x}_0\|)}, \quad (2.3.4)$$

а сигма-фильтр (2.1.9) в многомерном виде будет

$$\bar{\mathbf{x}}_{mn} = A(V'_{mn}, \mathbf{x}_{mn}, \delta) = \frac{1}{R} \sum_{r=1}^R \mathbf{x}'_{mn}, \quad \|\mathbf{x}'_{mn} - \mathbf{x}_{mn}\| \leq \delta. \quad (2.3.5)$$

Норму разности векторов можно интерпретировать как расстояние между точками в цветовом пространстве. Такой подход предоставляет возможности построения метрики в цветовом пространстве. Использование расстояния в цветовом пространстве будет обсуждаться далее в Главе 5 при рассмотрении задачи сегментации изображений.

Практическая реализация алгоритма декомпозиции цветного изображения, требующего в каждой точке анализ Z -мерной гистограммы, является задачей, требующей весьма больших вычислительных ресурсов. Более простым может быть отдельная поканальная обработка цветного изображения, когда вектор \mathbf{S}_{mn} составляется из результатов одноканальных декомпозиций по каждому из R, G, B каналов. Такой подход является значительно более простым в вычислительном отношении; по сути он означает анализ проекций цветного куба гистограммы $H(W_{mn})$ на каждую из R, G, B осей. Очевидный недостаток такого подхода — возможное наложение в проекции двух или более сгущений цветовой гистограммы, что может привести к невозможности их разделения. Однако такие события редки и проведенные эксперименты показывают, что упрощенный подход дает приемлемые результаты.

Цветное изображение на Рис. 2.3.2,а было обработано алгоритмом декомпозиции «D» (2.2.8) по каждому из R, G, B каналов по отдельности. Размер фрагмента сглаживания W_{mn} был 25×25 пикселей (показан соответствующим квадратом в правом верхнем углу изображения). Результат сглаживания, представленный на Рис. 2.3.2,б, является вполне удовлетворительным. Уровень шума (СКО), измеренный согласно (1.2.15) и усредненный по всем каналам, составлял 4,15 градации яркости на исходном изображении (а) и уменьшился до 0,92 на сглаженном (б).

Результаты исследований, изложенных в разделах 2.1–2.3, опубликованы в работах [171,179,192,193,235,236].



Рис. 2.3.2. Декомпозиция цветного изображения: а) исходное изображение (в правом верхнем углу показаны размеры фрагмента сглаживания); б) сглаживание алгоритмом «D» (2.1.18).

2.4. Изменение сложности изображения при сглаживании и декомпозиции

Во введении к настоящей главе говорилось, что нахождение сглаженной компоненты при декомпозиции можно интерпретировать как решение задачи достижения минимальной сложности изображения при максимальном сохранении информации о содержащихся в нем объектах. Вопрос сложности изображений рассматривался в разделе 1.4, где было показано, что подходящей мерой сложности могут быть оценки, основанные на двумерных вариациях. Также в разделе 1.4.2 было показано, как изменяется сложность изображения при добавлении шума. Представляет интерес поведение оценок сложности (показателя числа и показателя размеров объектов) при сглаживании и декомпозиции изображения. Исследования были проведены как на зашумленном эталонном изображении, для которого известен неискаженный оригинал, так и на реальных изображениях.

2.4.1. Эксперименты с модельными изображениями

Для исследования изменений оценки сложности изображения при сглаживании было сформировано кусочно-постоянное изображение с множеством объектов разных размеров и контраста относительно фона (Рис. 1.4.1,з). Изображение искажалось аддитивным гауссовым шумом с $\sigma = 5$ градаций яркости, после чего для удаления шума и восстановления изображения применялись сглаживающие алгоритмы локального среднего (2.1.2), локальной медианы (2.1.4), а также алгоритм получения сглаженной компоненты $S(i,j)$ при декомпозиции (2.2.8).

Результаты преобразований в виде значений показателя числа объектов (w_1), второй вариации Кронрода (w_2) и показателя размеров объектов d приведены в Таблице 2.4.1. Строка 1 соответствует эталонному изображению; 2 — искаженному шумом. Остальные строки отражают результаты фильтрации зашумленного изображения различными алгоритмами сглаживания при размере фрагмента анализа 15×15 элементов: 3 — локальное среднее (2.1.2); 4 — локальная медиана (2.1.4); 5 — гауссова фильтрация; 6 — сглаженная компонента $S(i,j)$ в (2.1.8) после одной итерации декомпозиции; 7 — сглаженная компонента $S(i,j)$ после трех итераций декомпозиции. Остальные значения вариаций и показателя размеров объектов, получаемые при использовании фрагментов сглаживания других размеров, представлены графически на Рис. 1.4.4-1.4.6. Измерения, проведенные по другим синтезированным изображениям (в частности, оригинала, использовавшегося на Рис. 2.1.2), а также изображениям, искаженным шумами с другими распределениями, показали те же зависимости.

Таблица 2.4.1. Значения СКО, показателя числа объектов (w_1), второй вариации Кронрода (w_2), показателя размеров объектов (d) и энтропии (E).

	Вид изображения	СКО	w_1	w_2	d	E
1	Исходное	0,00	33,13	811,87	6,13	0,58
2	С шумом, $\sigma=5$ град.	5,00	533,71	3644,04	1,71	4,93
3	Локальное среднее	19,30	8,83	463,95	13,13	4,08
4	Локальная медиана	20,26	6,80	253,46	10,51	2,09
5	Гауссова фильтрация	18,91	4,40	340,26	14,71	4,10
6	Декомпозиция (1 итер.)	0,64	36,55	872,95	5,97	1,68
7	Декомпозиция (3 итер.)	0,63	32,89	822,22	6,25	1,62

На Рис. 2.4.1–2.4.2 показаны изменения значений показателя числа объектов (w_1), второй вариаций Кронрода (w_2), показателя размеров объектов (d), а также среднеквадратического отклонения (СКО) результата фильтрации от неискаженного оригинала в зависимости от размеров фрагмента сглаживания (размер варьировался от 3×3 до 15×15 элементов) для результатов фильтрации алгоритмами локального среднего (Mean), локальной медианы (Median), гауссового сглаживания (Gauss) и декомпозиции (Decomp).

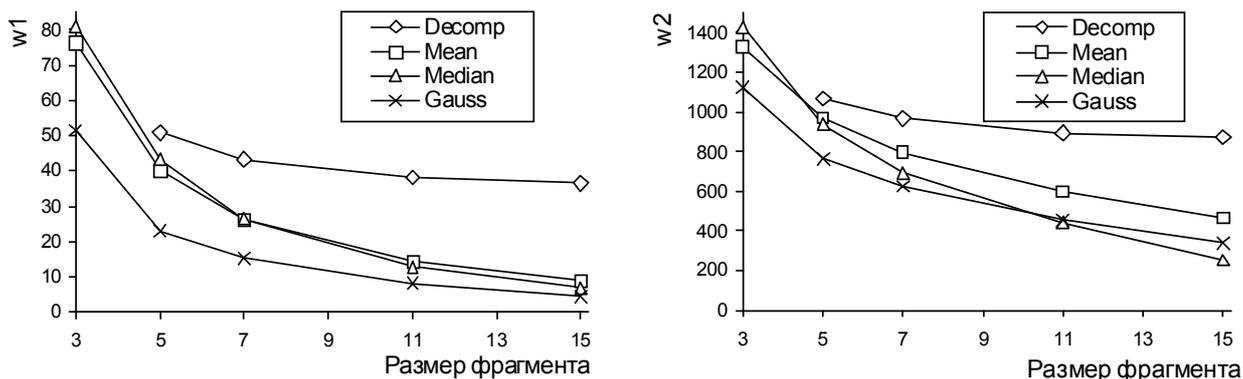


Рис. 2.4.1. Зависимости показателя числа объектов (w_1) и второй вариаций Кронрода (w_2) от размера фрагмента сглаживания.

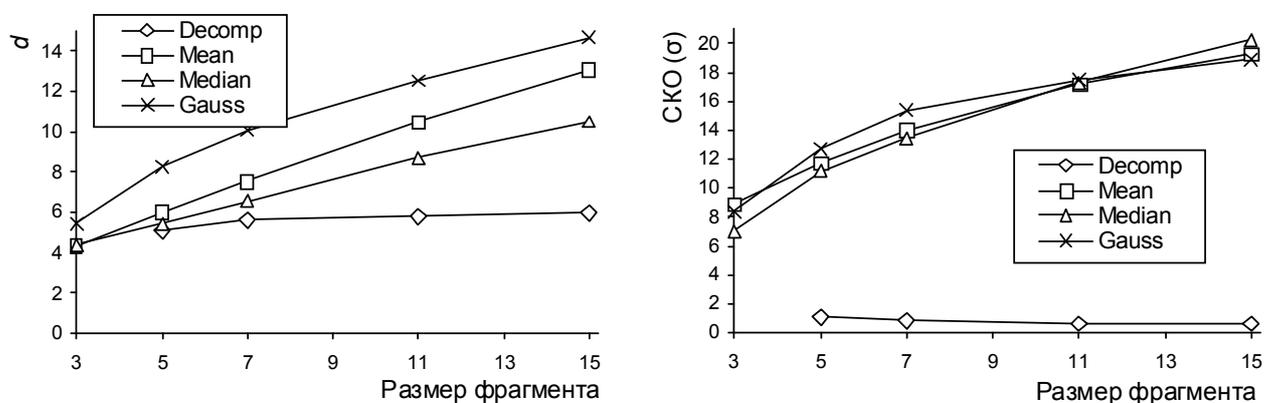


Рис. 2.4.2. Изменение показателя размеров объектов (d) и СКО (σ) результата фильтрации от неискаженного оригинала в зависимости от размера фрагмента сглаживания.

Анализируя полученные результаты можно сделать следующие выводы.

1. Алгоритм декомпозиции обеспечивает лучшее восстановление зашумленного сигнала по сравнению с другими методами.
2. С увеличением размеров фрагмента сглаживания, значения показателя числа объектов (w_1), второй вариации (w_2), а также показателя размеров объектов (d) для сглаженной компоненты декомпозиции стремятся к постоянным значениям. Для остальных вариантов сглаживания значения w_1 и w_2 стремятся к нулю, а d — возрастают почти линейно (Рис. 2.4.1–2.4.2).
3. Точность восстановления изображения, искаженного шумом, при использовании алгоритма декомпозиции возрастает с увеличением размера фрагмента сглаживания и числа итераций, а остальные алгоритмы при этом дают все более отличающиеся от оригинала результаты (Таблица 2.4.1 и Рис. 2.4.2, правый).
4. Значения показателя числа объектов (w_1), второй вариаций (w_2) и показателя размеров объектов (d) в результате декомпозиции возвращаются к значениям, близким первоначальному неискаженному изображению (Таблица 2.4.1). То, что данные значения выходят на плато

(Рис. 2.4.1) одновременно с асимптотическим приближением к нулю величины отклонения результата сглаживания от оригинала (Рис. 2.4.2, правый) говорит о стабилизации преобразования вблизи исходного неискаженного изображения.

5. Стремление для обычных вариантов сглаживания значений w_1 и w_2 к нулю, а значений d к возрастанию, означает уменьшение сложности изображения, что говорит о все большем размывании деталей и разрушении сигнала изображения. Выход тех же значений на плато в случае алгоритма декомпозиции означает, что основная структура деталей изображения сохраняется.

6. Для сглаженной компоненты декомпозиции значения w_1 , w_2 и d выходят на плато при размерах фрагмента около 11×11 элементов, т.е. при дальнейшем увеличении фрагмента сглаживаемое изображение практически не изменяется. Точка выхода значений на плато может служить критерием выбора размера фрагмента сглаживания.

7. Малое значение энтропии (2,09) в строке 4 Таблицы 2.4.1 объясняется особенностью медианной фильтрации сигнала, имеющего ограниченное число значений.

Эксперименты по сглаживанию изображения (подавлению шумов) показали, что при постепенном увеличении числа итераций и размеров фрагмента (степени сглаживания), значения w_1 и w_2 сначала спадают достаточно резко, затем выходят на некоторое плато и почти перестают уменьшаться. Значения параметров сглаживания, при которых происходит выход на плато, можно интерпретировать как точку достижения искомого результата. Получаемый результат достаточно устойчив к последующему сглаживанию и соответствующее ему сочетание значений показателя числа объектов (w_1) и показателя размеров объектов (d) можно считать оценкой, характеризующей сложность изображения.

2.4.2. Эксперименты с реальными изображениями

Представляет интерес, как будут вести себя оценки сложности при сглаживании реальных изображений алгоритмом декомпозиции «D» (2.2.8). В экспериментах использовались изображения, представленные на Рис. 1.4.1, имеющие размеры 256×256 пикселей. Размер фрагмента сглаживания составлял 15×15 элементов. Сглаженные компоненты, полученные указанным алгоритмом, представлены на Рис. 2.4.3. Значения показателя числа объектов (w_1), второй вариации Кронрода (w_2), показателя размеров объектов (d) и энтропии (E) как для исходных изображений, так и для сглаженных компонент (S) после операции декомпозиции (2.2.8) приведены в Таблице 2.4.2. Изображения, обозначенные буквами (а – ж), соответствуют тем же изображениям на Рис. 1.4.1.

Анализируя значения вариаций, приведенные в Таблице 2.4.2 для исходного изображения и его сглаженной компоненты, в целом отметим следующее.

1. Значения показателей числа объектов (w_1) снижаются существенно — от 3,5 до 9 раз.
2. Значения показателей размеров объектов (d) возрастают для всех изображений, причем диапазон изменений от 1,5 до 5 раз.
3. Значения вторых вариаций Кронрода (w_2) уменьшаются для всех изображений, и диапазон изменений сравнительно мал — от 1,2 до 1,9 раз.
4. Значения энтропии (E) меняются слабо и мало информативны.



Рис. 2.4.3. Сглаженные компоненты (S) изображений на Рис. 1.4.1 после декомпозиции (2.2.8).

Таблица 2.4.2. Значения показателей числа объектов (w_1), показателя размеров объектов (d), второй вариации Кронрода (w_2) и энтропии (E) для исходных изображений на Рис. 1.4.1 и сглаженной компоненты (Рис. 2.4.3) после декомпозиции (2.2.8) изображения.

Изображение (Рис. 2.4.3)	Исходное изображение				Сглаженная компонента (S)			
	w_1	w_2	d	E	w_1	w_2	d	E
а	469,85	5260,25	2,80	7,59	148,66	3272,62	5,50	7,44
б	129,23	2589,58	5,01	7,60	15,12	1476,61	24,42	7,48
в	618,79	7947,70	3,21	7,59	276,97	5516,53	4,98	7,48
г	360,03	4804,38	3,16	7,16	110,68	2597,77	5,87	7,03
д	256,84	3260,76	3,17	7,34	70,27	2838,65	10,10	7,17
е	419,62	5118,65	3,05	6,76	51,33	2703,48	13,17	5,93
ж	231,24	6462,04	6,39	5,61	65,40	5258,05	20,10	4,73

Сравнивая данные в Таблице 2.4.2, в первую очередь те, которые относятся к сглаженной компоненте, можно обнаружить две различных группы изображений. Первую составляют изображения 'а', 'в' и 'г', а вторую — 'б', 'д', 'е' и 'ж'. В первой группе сглаженные компоненты характеризуются сравнительно большими значениями показателей числа объектов (w_1) и малыми значениями показателей размеров объектов (d). Во второй наоборот, значения w_1

оказываются сравнительно малыми, а d — большими. Рассматривая, как меняются значения w_1 и d при сглаживании, видно, что для первой группы коэффициенты уменьшения w_1 находятся в диапазоне от 2,2 до 3,2, а коэффициенты увеличения d — в диапазоне от 1,5 до 2,0. Для второй группы данные диапазоны совершенно другие — от 3,5 до 8,6 и от 3,2 до 4,9.

Воспользуемся сделанными в разделе 1.4.1 объяснениями, касающимися сравнения и интерпретации оценок сложности. Малые значения w_1 и большие d у изображений ‘б’, ‘е’ и ‘ж’ второй группы указывают, что они характеризуются сравнительно малым числом деталей с большими размерами. Значения w_1 и d для первой группы показывают, что данные изображения наоборот, содержат много мелких объектов. Про изображение ‘в’, у которого наблюдается сравнительно небольшое изменение значения d и самое большое значение w_1 , можно отдельно сделать вывод, что оно содержит значительное число мелких объектов. В заключение сравнения значений w_1 и d можно сказать, что у изображений первой группы сглаженные компоненты являются более сложными, чем у изображений второй группы.

На Рис. 2.4.4–2.4.6 приведены зависимости показателей числа объектов (w_1) и размеров объектов (d) от размера фрагмента анализа для трех сглаженных изображений, приведенных на Рис. 1.4.1,а,б,в. Сглаживания производились следующими локальными алгоритмами: средним (Mean), медианой (Median), гауссовым (Gauss) и декомпозицией (Decomp).

Нетрудно видеть, что несмотря на различия абсолютных значений, в целом на графиках прослеживаются те же закономерности, которые отмечались для модельных изображений и были иллюстрированы на Рис. 2.4.1–2.4.2. А именно: сглаживание алгоритмами локального среднего, медианы и гауссовым фильтром с увеличением размеров фрагмента дает результаты, для которых значения показателей числа объектов (w_1) стремятся к нулю, а значения показателей размеров объектов (d) уверенно возрастают. Такое поведение значений w_1 и d говорит о разрушении деталей изображения.

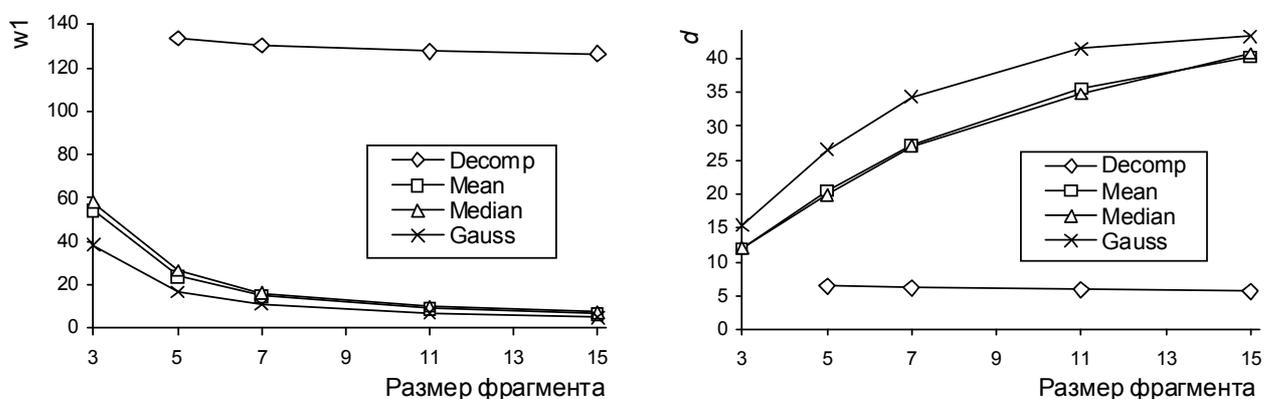


Рис. 2.4.4. Зависимости показателей числа объектов (w_1) и размеров объектов (d) от размера фрагмента для изображения на Рис. 2.4.3,а при различных способах сглаживания.

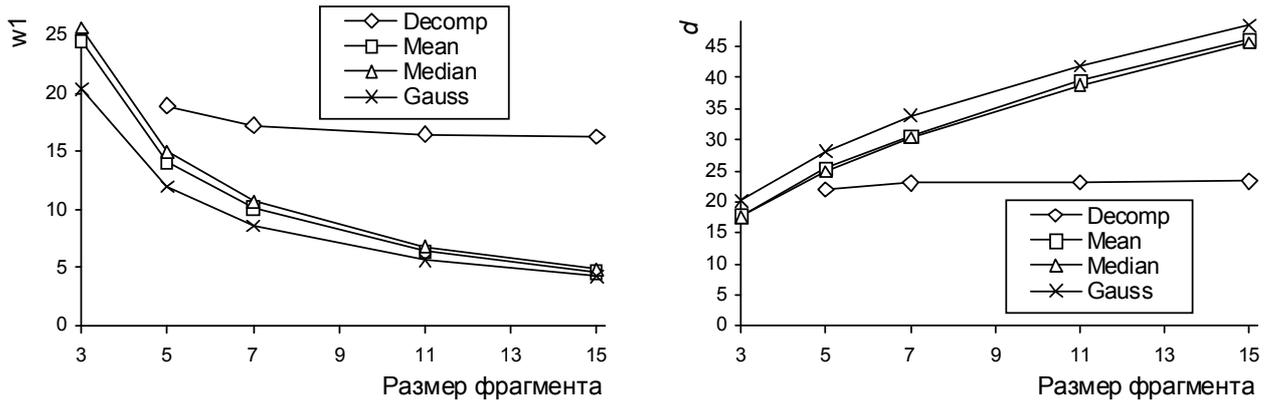


Рис. 2.4.5. Зависимости показателей числа объектов (w_1) и размеров объектов (d) от размера фрагмента для изображения на Рис. 2.4.3,б при различных способах сглаживания.

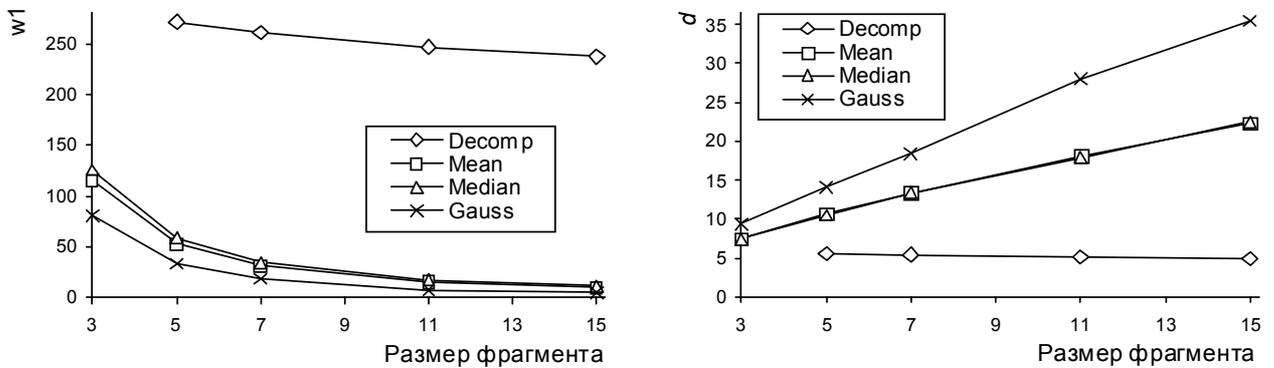


Рис. 2.4.6. Зависимости показателей числа объектов (w_1) и размеров объектов (d) от размера фрагмента для изображения на Рис. 2.4.3,в при различных способах сглаживания.

Наряду с этим значения w_1 и d для сглаженных компонент, получаемых алгоритмом декомпозиции, показывают стабилизацию результатов, что говорит о сохранении деталей изображения. Получаемые при этом значения показателя числа объектов (w_1) и показателя размеров объектов (d) являются оценкой сложности сглаженной компоненты изображения.

Изложенные в разделе 2.4 результаты опубликованы в [95, 168, 182, 203, 244, 245, 314].

2.5. Выводы и результаты

Исследована задача сглаживания изображения при условии сохранения контурных перепадов. Показано, что в рамках разработанной двухмасштабной многокомпонентной модели задача соответствует разделению изображения на различные информационные компоненты. Выбран класс оценок, основанных на использовании статистических методов анализа распределения значений элементов в локальной пространственной области анализа. Показано, что локальные арифметическое среднее и медиана не приводят к удовлетворительному

решению. Из ранее известных методов сглаживания наиболее подходящим является сигма-фильтр, также имеющий недостатки.

На основе двухмасштабной многокомпонентной модели разработаны метод и алгоритм декомпозиции изображения на две различные информационные составляющие — кусочно-гладкую компоненту, несущую информацию о яркостях протяженных областей и границах между ними, и текстурно-детальную компоненту, содержащую мелкие детали, текстуру и шум. Как и модель, алгоритм декомпозиции также является двухмасштабным: первоначально осуществляется анализ по области малых размеров (окрестности), а затем по области больших размеров (фрагменту). По сути такой подход позволяет учесть при анализе не только распределения значений признаков (яркости), но и пространственную корреляцию.

Проведены исследования точности восстановления модельного кусочно-постоянного двумерного сигнала, искаженного шумом, и осуществлено сравнение разработанного алгоритма декомпозиции с различными алгоритмами сглаживания. Исследовалось изменение сложности изображения и точности восстановления сигнала в зависимости от алгоритма и размеров области анализа. Показано, что с увеличением размеров области анализа (окна сглаживания) ошибка восстановления сигнала алгоритмом декомпозиции стремится к нулю, а сложность изображения снижается лишь до некоторого уровня, что говорит о сохранении деталей на изображении. При использовании локального среднего или локальной медианы ошибка только возрастает, а к нулю стремится сложность изображения, что свидетельствует о разрушении деталей изображения. Показано, что алгоритм декомпозиции обеспечивает в $2 \div 2,5$ раза более точное восстановление искаженного шумом изображения по сравнению с другими алгоритмами сглаживания, в том числе и с сигма-фильтром.

Показаны возможности применения алгоритма декомпозиции для сглаживания цветных и многоканальных изображений. Предложено несколько вариантов его реализации: через введение метрики в цветовом пространстве или через проецирование и покомпонентный анализ; последний вариант менее точен, но более эффективен в вычислительном отношении.

На основе двухмасштабной многокомпонентной модели и метода декомпозиции, позволяющего разделять изображение на компоненты с различным информационным содержанием, предлагается новый подход к анализу и обработке изображений. Он заключается в том, что для анализа следует выбирать лишь ту компоненту изображения, в которой должна содержаться искомая информация. При обработке возможно применение преобразований, специфических для каждой из полученных компонент.

ГЛАВА 3. РАЗРАБОТКА МЕТОДОВ ФИЛЬТРАЦИИ, КОРРЕКЦИИ И УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДВУХМАСШТАБНОЙ МОДЕЛИ

Введение

В настоящей главе рассматриваются методы и алгоритмы устранения помех, коррекции и улучшения изображений, разработанные на основе изложенной в Главе 1 двухмасштабной многокомпонентной модели. Целью процедур фильтрации, коррекции и улучшения изображения является устранение искажений и приведение видеоинформации к виду, наилучшему для последующего анализа или восприятия. Подобные процедуры обычно применяются на этапе предварительной обработки данных, и их использование предполагает знания некоторых априорных сведений. Фильтрация, как правило, применяется в случае, если известны необходимые свойства сигнала и помехи, коррекция — когда известны характеристики искажающего преобразования. Задача улучшения во многом является проблемно ориентированной; ее решение в определенной степени связано с вопросами оценивания визуального качества изображения и зависит от конечной цели обработки.

Тем не менее, опираясь на двухмасштабную модель и разработанный на ее основе метод декомпозиции, удастся использовать ряд свойств изображений, позволяющих разрабатывать методы фильтрации коррекции и улучшения изображений на основе лишь апостериорной информации, т.е. той, которую можно извлечь из уже полученного искаженного изображения. К числу таковых можно отнести алгоритмы фильтрации помех с автоматической оценкой свойств помехи, методы автоматической градационной коррекции видеоинформации, а также алгоритмы повышения локальных контрастов (как адаптивные, так и не адаптивные).

На процедуре улучшения изображения (*image enhancement*), применяемой в основном на этапе предварительной обработки, необходимо остановиться особо. Во многих работах предлагаются различные методы и алгоритмы улучшения изображений; обсуждение наиболее известных можно найти в публикациях [1,40,113,274,315,319,343,346]. Однако оценивать их затруднительно из-за отсутствия четких критериев сравнения получающихся результатов — основной оценкой является визуальное восприятие получаемых видеоданных. В последнее время появился ряд работ, в которых предлагаются функционалы качества изображения [281,294,319,348,349,365], с помощью которых делаются попытки отыскать эмпирическую формулу для оценки изображения, но единый общепринятый подход тем не менее отсутствует.

Резюмируя предлагаемое множество оценок отметим, что визуальное качество изображения (монохромного, отвлекаясь от самостоятельной задачи цветокоррекции) в целом определяется глобальным контрастом (то есть тем, что изображение занимает весь диапазон возможных яркостей), локальным контрастом (заметностью деталей относительно окружающего фона) а также резкостью контурных перепадов.

Таким образом задачу улучшения изображений можно разделить на две самостоятельные части: *градационную коррекцию*, глобально воздействующую на изображение целиком, и *усиление локальных контрастов*, обеспечивающее повышение заметности деталей [113]. Почти всегда для изменения общего контраста используются глобальные методы анализа и преобразования [40,257,260,276,277], а для усиления локального контраста, соответственно, локальные методы [31,66,150,155,179,236,343,346]. Ниже будет показано, что при использовании двухмасштабной многокомпонентной модели изображения возможно их сочетание.

На основе модели, рассматривающей изображение как смесь независимых компонент различного информационного содержания, удастся записать каноническую формулу, обобщающую большинство применяемых методов повышения локальных контрастов. Применение рассмотренного в Главе 2 метода декомпозиции, обеспечивающего разделение изображения на указанные компоненты, позволяет строить алгоритмы улучшения изображений, в которых каждая из компонент может подвергаться независимому преобразованию.

3.1. Фильтрация помех

3.1.1. Фильтрация импульсных помех

Постановка задачи

Фильтрация импульсных помех — одна из классических задач устранения искажений, возникающих на изображениях. Импульсными помехами называют случайные и сравнительно редкие искажения отдельных элементов изображения, когда значения помехи значительно отличаются от истинных значений сигнала и не коррелированы с ними. Подобного рода помехи были типичны для цифровых систем регистрации и передачи сигнала первых поколений [29,87,157] и в настоящее время возникают не часто, тем не менее вопрос их устранения важен с позиции методологии разработки алгоритмов обнаружения некоррелированных шумовых выбросов со статистикой, резко отличающейся от статистики сигнала.

На изображении импульсные помехи выглядят как хаотически расположенные точки, хорошо различимые визуально и, как правило, случайные по яркости. Различие статистических свойств помех и изображений позволяет разработать методы фильтрации, использующие

межэлементные статистические связи последних. Алгоритмы фильтрации импульсных помех, большинство из которых являются эвристическими, рассматриваются во многих работах, например [40,78,87,157,372]. Улучшить их результаты возможно, если учесть неоднородности изображения и особенности визуального обнаружения помехи.

Модель искажения изображения импульсными помехами проста. Значение каждого из элементов x_{mn} изображения независимо от значений остальных элементов с вероятностью p заменяется на случайное значение ξ_{mn} . Обозначим через $\mathbf{X}' = [x'_{mn}]$ исходное неискаженное изображение, через $\mathbf{X} = [x_{mn}]$ — искаженное импульсной помехой, а через $\mathbf{Y} = [y_{mn}]$ — результат фильтрации. Процесс искажения представится в виде:

$$x_{mn} = \begin{cases} x'_{mn} & \text{с вероятностью } (1-p) \text{ для неискаженного элемента;} \\ \xi_{mn} & \text{с вероятностью } p \text{ для элемента, искаженного помехой.} \end{cases} \quad (3.1.1)$$

Без существенного ограничения общности можно полагать, что значения импульсных помех ξ_{mn} распределены равномерно в диапазоне яркостей $[0, K-1]$. Задача фильтрации импульсных помех состоит в обнаружении помех и последующей коррекции искаженных отсчетов яркости.

Наиболее распространенные алгоритмы фильтрации

Известные методы фильтрации основываются на предсказании значения элемента \tilde{x}_{mn} по окружающей его окрестности V_{mn} . При этом используются локальные корреляционные связи ближайших элементов изображения [96] и учитывается, что шум пространственно декоррелирован. При обнаружении наблюдаемое значение элемента x_{mn} сравнивается с предсказываемым \tilde{x}_{mn} , и если они отличаются более, чем на величину некоторого заданного порога δ , считается, что x_{mn} — помеха, после чего осуществляется ее исправление на значение \tilde{x}_{mn} . Общая формула вычисления предсказываемого значения \tilde{x}_{mn} для большинства методов:

$$\tilde{x}_{mn} = f\{x_{ij} \mid x_{ij} \in V_{mn}\}.$$

Формула алгоритма фильтрации при этом выглядит следующим образом:

$$y_{mn} = \begin{cases} x_{mn}, & \text{если } |x_{mn} - \tilde{x}_{mn}| < \delta; \\ \tilde{x}_{mn}, & \text{если } |x_{mn} - \tilde{x}_{mn}| \geq \delta. \end{cases} \quad (3.1.2)$$

Параметр δ есть порог обнаружения; в некоторых алгоритмах этап обнаружения отсутствует и исправляются все элементы изображения, что соответствует $\delta = 0$.

Обычно используется один из следующих способов предсказания: линейная оценка; марковская модель изображения, и методы, основанные на порядковых статистиках.

1. *Линейная оценка.* Предсказание \tilde{x}_{mn} определяется в результате свертки изображения $\mathbf{X} = [x_{mn}]$ с матрицей $\mathbf{V} = [v(i,j)]$, имеющей порядок $2l + 1$:

$$\tilde{x}_{mn} = \sum_{i=-l}^l \sum_{j=-l}^l x(m+i, n+j) v(i, j). \quad (3.1.3)$$

Основное различие методов, описанных в литературе, заключается в выборе матрицы V . Наиболее распространенные варианты следующие:

$$\mathbf{V}_1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}; \quad \mathbf{V}_2 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}; \quad \mathbf{V}_3 = \frac{1}{8} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}; \quad \mathbf{V}_4 = \frac{1}{20} \begin{bmatrix} 2 & 3 & 2 \\ 3 & 0 & 3 \\ 2 & 3 & 2 \end{bmatrix}.$$

Как правило, фильтрация выполняется итеративно с постепенным уменьшением параметра δ в формуле (3.1.2). Описан метод фильтрации с адаптивным изменением матрицы V : суммируются лишь те элементы изображения \mathbf{X} , которые (на предыдущей итерации) были отмечены как неискаженные [343]. Предложен каузальный алгоритм вычисления \tilde{x}_{mn} по уже обработанным элементам изображения x_{mn} в последовательной развертке [87]. Матрица V при этом приобретает вид:

$$\mathbf{V}_5 = \left(\sum C_i \right)^{-1} \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \text{ рекомендуются значения: } c_1 = c_3 = 2, c_2 = c_4 = 3.$$

2. Марковская модель. Предполагается, что изображение есть реализация марковского случайного поля и x'_{mn} в (3.1.1) зависит лишь от значений элементов локальной окрестности $x_{mn}^k \in V_{mn}$. Это приводит к следующему алгоритму предсказания [12,370]:

$$\tilde{x}_{mn} = \left(\sum_{k=1}^K v_{mn}^k \right)^{-1} \sum_{k=1}^K v_{mn}^k x_{mn}^k, \quad (3.1.4)$$

где $v_{mn}^k = (\alpha + \beta |x_{mn} - x_{mn}^k|)^{-1}$, а параметры α и β характеризуют статистические связи между элементами марковского поля.

3. Порядковые статистики. Методы основаны на ранжировании элементов окрестности $x_{mn}^k \in V_{mn}$: $x_{mn}^{(1)} \leq x_{mn}^{(2)} \leq \dots \leq x_{mn}^{(R)}$. Как правило, в качестве предсказания выбирают медиану:

$$\tilde{x}_{mn} = \text{med}(V_{mn}) = x_{mn}^{((R+1)/2)}. \quad (3.1.5)$$

Наиболее распространен метод исправления всех точек изображения [78,292,358], т.е. $\delta = 0$ в соотношении (3.1.2); нами использовался метод медианной фильтрации с пороговым обнаружением помех [78,172] при $\delta > 0$. Известен также подход, использующий линейные комбинации медиан [292] по вложенным окрестностям $V_{mn}^{(1)}, \dots, V_{mn}^{(K)}$:

$$\tilde{x}_{mn} = \sum_{k=1}^K a_k \text{med}(V_{mn}^{(k)}) / \sum_{k=1}^K a_k. \quad (3.1.6)$$

В [372] предложен метод, в котором явная оценка $\tilde{x}(m, n)$ отсутствует, а обнаружение помехи производится в зависимости от расположения x_{mn} в вариационном ряду элементов окрестности:

$$x_{mn} \text{ — помеха, если } x_{mn} \leq x_{mn}^{(r)} \text{ или } x(m,n) \geq x_{mn}^{(R-r+1)}; \quad (3.1.7)$$

исправление осуществляется заменой на значение медианы $\text{med}(V_{mn})$.

Вероятности ошибок предсказания

Воспользуемся моделью окрестности (1.2.9) и рассмотрим множество фоновых элементов, временно положив $\psi_{mn} = 0$. Поскольку для фоновых элементов $e_{mn} = 0$, формула модели, связывающая R элементов локальной окрестности $x_{mn}^r \in V_{mn}$, упрощается до

$$x_{mn}^r = \mu(V_{mn}) + \eta_{mn}^r, \quad (3.1.8)$$

где $\mu(V_{mn})$ — величина, постоянная для элементов окрестности V_{mn} , η_{mn}^k — независимая нормально распределенная случайная величина с нулевым средним и дисперсией σ^2 : $N(0, \sigma^2)$.

Предположим, что известна бинарная маска импульсных помех $\Lambda(p) = [\lambda_{mn}]$: $\lambda_{mn} = 1$ для искаженных точек, и $\lambda_{mn} = 0$ для неискаженных, причем $P\{\lambda_{mn} = 1\} = p$, а $P\{\lambda_{mn} = 0\} = 1 - p$. Тогда (3.1.1) можно записать в виде:

$$x_{mn} = (1 - \lambda_{mn})x'_{mn} + \lambda_{mn}\xi_{mn}. \quad (3.1.9)$$

Аналогично можно переписать и модель (3.1.8):

$$x_{mn}^r = (1 - \lambda_{mn}^r)(\mu(V_{mn}) + \eta_{mn}^r) + \lambda_{mn}^r \xi_{mn}^r. \quad (3.1.10)$$

Воспользуемся моделью (3.1.8) для сравнения точности предсказаний значения элемента x_{mn} по медиане $\hat{x}_{mn} = \text{med}(V_{mn})$ и среднему значению $\bar{x}_{mn} = \sum_R x_{mn}^r / R$ элементов окрестности V_{mn} . Зная соседние элементы, мы можем определить лишь $\mu(V_{mn})$, поэтому рассмотрим ошибки предсказаний: $\hat{\zeta}_{mn} = |\hat{x}_{mn} - \mu(V_{mn})|$, $\bar{\zeta}_{mn} = |\bar{x}_{mn} - \mu(V_{mn})|$. В [70] показано, что медиана выборки K значений из нормальной совокупности асимптотически нормальна: $N(\mu, \pi\sigma^2/2K)$. Известно также, что, для нормально распределенного шума арифметическое среднее по выборке является более точной оценкой для $\mu(V_{mn})$, чем медиана: $\sigma_{\bar{\zeta}}/\sigma_{\hat{\zeta}} = \sqrt{\pi/2} \approx 1,25$. Однако устойчивость среднего значения к появлению резко выделяющихся наблюдений, какими являются импульсные помехи и контурные элементы изображения, совершенно различна.

Рассмотрим изображение с импульсными помехами. Предполагая распределение η_{mn}^r нормальным, выберем такое ε , что

$$\sum_{-\infty}^{-\varepsilon} P(\eta_{mn}^r) + \sum_{-\varepsilon}^{\infty} P(\eta_{mn}^r) \ll \sum_{-\varepsilon}^{\varepsilon} P(\eta_{mn}^r). \quad (3.1.11)$$

При $|\xi_{mn}^r - \mu(V_{mn})| \leq \varepsilon < \delta$, где δ — порог визуального обнаружения мелких деталей, помеха на изображении будет неразличима; в этом случае будем считать, что искажения не произошло.

Найдем вероятности ошибок оценок среднего $\bar{x}(m, n)$ и медианы $\hat{x}(m, n)$ по окрестности, т.е. значений $P\{|\bar{x}(m, n) - \mu(V_{mn})| > \varepsilon\}$ и $P\{|\hat{x}(m, n) - \mu(V_{mn})| > \varepsilon\}$, как функций от вероятности искажения p . Можно показать, что при условии $2R\varepsilon \leq 1$:

$$\bar{P}_{\text{ош}}(p) \geq \sum_{r=1}^R p^r (1-p)^{R-r} C_R^r 2^{1-r} (1 - 2R\varepsilon/r); \quad (3.1.12)$$

$$\hat{P}_{\text{ош}}(p) = \sum_{r=(R+1)/2}^R p^r (1-p)^{R-r} C_R^r 2^{1-r} \sum_{s=(R+1)/2}^r C_r^s. \quad (3.1.13)$$

На Рис. 3.1.1 представлены зависимости $\bar{P}_{\text{ош}}(p)$ и $\hat{P}_{\text{ош}}(p)$. Здесь выбраны реальные значения: $R = 9$ и $\varepsilon = 1/2R$. Как видно на Рис. 3.1.1, при $0 < p < 0,3$ оценка медианы $\hat{x}(m, n)$ обеспечивает значительно меньшую вероятность ошибки по сравнению с $\bar{x}(m, n)$; при $p > 0,5$ обе оценки оказываются недостаточно точными.

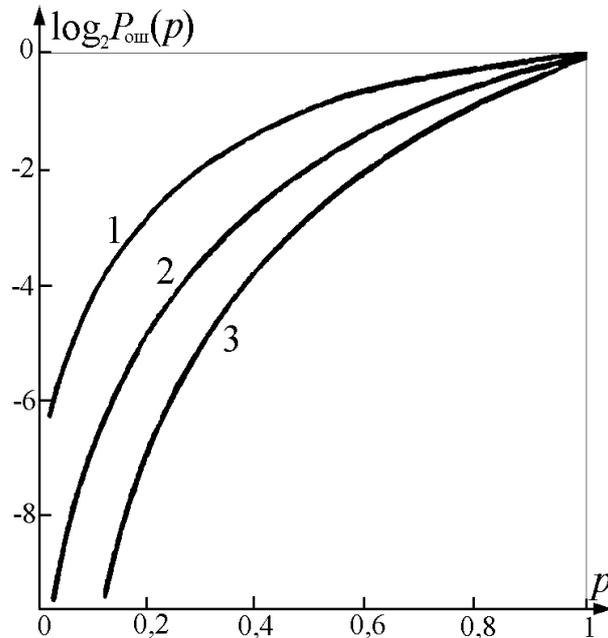


Рис. 3.1.1. Вероятности ошибок предсказания по среднему $\bar{P}_{\text{ош}}$ и медиане $\hat{P}_{\text{ош}}$ от вероятности помехи p : 1) среднее по 9 элементам; 2) медиана по 5 элементам; 3 — медиана по 9 элементам.

Проводились экспериментальные проверки точности предсказания значений элементов исходного изображения $[x'_{mn}]$ по изображению с импульсными помехами $[x_{mn}]$. Для этого на реальные изображения наносилась импульсная помеха заданной вероятности искажения p , находились оценки \bar{x}_{mn} и \hat{x}_{mn} , затем определялись распределения вероятностей отклонения $\bar{P}(z) = P\{|\bar{x}_{mn} - x'_{mn}| = z\}$ и $\hat{P}(z) = P\{|\hat{x}_{mn} - x'_{mn}| = z\}$. Полученные распределения вероятностей отклонения предсказаний по окрестности 3×3 при различных вероятностях искажения элемента импульсной помехой (p), усредненные по 16 изображениям (Рис. 1.1.2), представлены на Рис. 3.1.2. Среднеквадратические значения отклонений оценок \bar{x}_{mn} и \hat{x}_{mn} приведены ниже в

Табл. 3.1.1 (методы «а» и «д»). Эксперименты подтверждают теоретические выводы, что оценка по медиане обеспечивает существенно лучшую точность, чем оценка по среднему, и это проявляется, как при больших, так и при малых вероятностях помех. Заметим, что и при полном отсутствии помех ($p = 0$) ошибка предсказания велика.

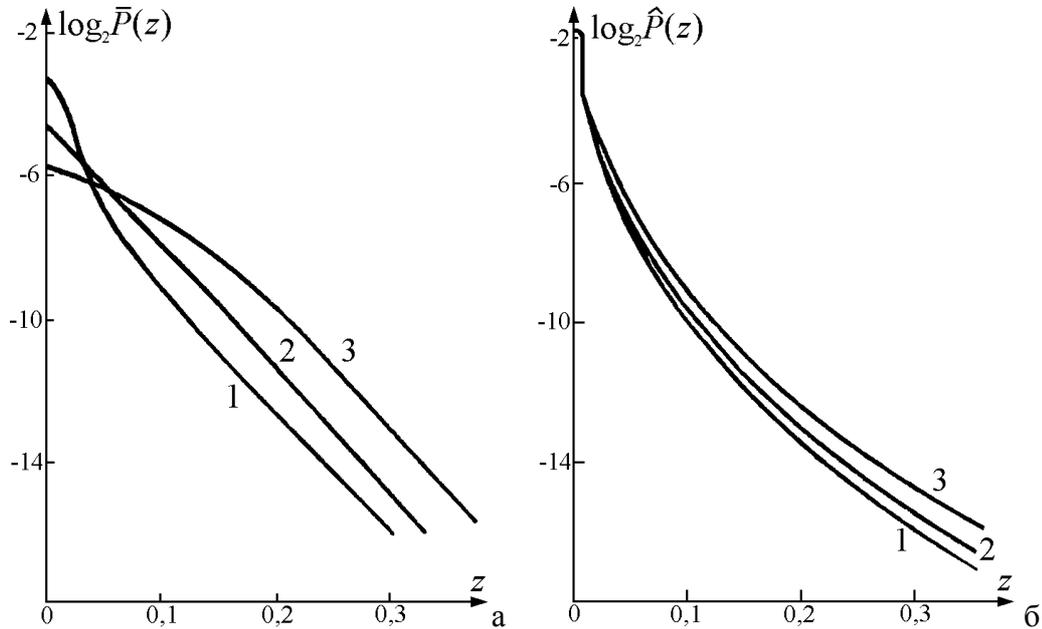


Рис. 3.1.2. Вероятности отклонения предсказания от истинного значения элемента: а) оценка по среднему; б) оценка по медиане. Вероятность помехи (p): 1) $p = 0$; 2) $p = 0,1$; 3) $p = 0,2$:

Анализ и экспериментальные проверки различных методов фильтрации показывают, что основная проблема сводится к нахождению эффективного способа обнаружения помех с возможно более низкой вероятностью ошибки. Места ошибок второго рода (точки ложного обнаружения помех) концентрируются вблизи контуров. Это говорит о том, что однородная модель (3.1.8), не учитывающая контурные перепады, оказывается неудовлетворительной. В разделе 1.2.2 было показано, что изображение следует рассматривать как смесь независимых случайных компонент, распределенных на фоновом и контурном множествах. С учетом разделения на такие множества, модель (3.1.8) следует переписать в виде

$$x_{mn}^r = \mu(V_{mn}) + e_{mn}^r \alpha_{mn}^r + (1 - e_{mn}^r) \eta_{mn}^r, \quad (3.1.14)$$

где $\mathbf{E} = \{e_{mn}\}$ — бинарная контурная маска изображения: $e_{mn} = 1$, если (m, n) — контурная точка, и $e_{mn} = 0$, если (m, n) — точка фонового участка; α_{mn} — нормально распределенная случайная величина $N(0, \sigma_a^2)$, описывающая статистику значений элементов на контурных участках. Для изображения с импульсными помехами, аналогично выражению (3.1.10), можно записать взаимосвязь соседних элементов:

$$x_{mn}^r = (1 - \lambda_{mn}^r) [\mu(V_{mn}) + e_{mn}^r \alpha_{m,n}^r + (1 - e_{mn}^r) \eta_{mn}^r] + \lambda_{mn}^r \xi_{mn}^r, \quad (3.1.15)$$

где $\Lambda(p) = [\lambda_{mn}]$ — бинарная маска импульсных помех.

Алгоритм ранговой пороговой фильтрации

До сих пор мы рассматривали вероятности ошибок предсказания в предположении, что порог визуального обнаружения δ в (2.1.7) по всему изображению одинаков. Из свойств зрения известно, что контрастная чувствительность глаза вблизи контурных перепадов снижается; также согласно модели окрестности (1.2.9) существенно отличаются дисперсии распределений значений элементов для фоновых и контурных точек. Очевидно, что и ε для них также должны различаться. Предположим, что значение ε в (3.1.11) для точки (m,n) определяется дисперсией элементов окрестности σ_{mn}^2 как $\varepsilon_{mn} = u\sigma_{mn}$. Тогда для неискаженного изображения справедливо:

$$P\{|x_{mn}^r - \mu_{mn}| \geq u\sigma_{mn}\} = \theta = \text{const}. \quad (3.1.16)$$

Используя μ_{mn} и σ_{mn} , выберем следующий критерий обнаружения импульсных помех:

$$\text{если } |x_{mn} - \mu_{mn}| \geq (u\sigma_{mn} + v), \text{ то } x_{mn} \text{ — помеха.} \quad (3.1.17)$$

Здесь v — порог визуального обнаружения отдельной точки на ровном фоне, аналогичный по сути порогу δ в соотношении (3.1.2). В соответствии с (1.2.4) σ_{mn} определяется согласно соотношению $\sigma_{mn}^2 = \sigma_{pg}^2 + \sigma_{\xi}^2$.

Будем считать, что неискаженные элементы внутри окрестности $x_{mn}^r \in V_{mn}$ распределены нормально: $N(\mu_{mn}, \sigma_{mn}^2)$. Наличие в V_{mn} импульсных помех, значительно отличающихся от μ_{mn} , приводит к искажению среднего и дисперсии общей совокупности элементов окрестности. Для оценки μ и σ нормальной совокупности (неискаженных элементов окрестности), зашумленной элементами другого распределения со значительно большей дисперсией, воспользуемся порядковыми статистиками. Пусть имеется ранжированная выборка R элементов из нормальной совокупности: $x^{(1)} \leq x^{(2)} \leq \dots \leq x^{(R)}$. Известно [54], что можно оценить среднее и дисперсию как

$$\mu_r = \sum_r^{R-r+1} x_r / (R - 2r); \quad \sigma_r = |x^{(R-r+1)} - x^{(r)}| / M_r.$$

Здесь μ_r — урезанное среднее, а M_r — математическое ожидание разности элементов того же ранга (r) из совокупности в R элементов с распределением $N(0,1)$. Данный способ определения среднего и дисперсии устойчив к появлению резких выбросов и эффективен при вычислении.

На основании изложенного предлагается следующий ранговый пороговый алгоритм фильтрации импульсных помех [179,201]. Для каждой точки выбираются элементы окрестности $x_{mn}^r \in V_{mn}$ и осуществляется их ранжирование: $x_{m,n}^{(1)} \leq \dots \leq x_{m,n}^{(R)}$. Зная параметр r , вычисляются $\mu_r(m,n)$ и $d_r(m,n) = (x_{mn}^{(R-r+1)} - x_{mn}^{(r)}) / M_r$. Алгоритм фильтрации будет выглядеть:

$$y_{mn} = \begin{cases} x_{mn}, & \text{если } |x_{mn} - \mu_r(m,n)| < c_1 d_r(m,n) + c_2, \\ \mu_r(m,n), & \text{если } |x_{mn} - \mu_r(m,n)| \geq c_1 d_r(m,n) + c_2. \end{cases} \quad (3.1.18)$$

В данном алгоритме имеются три параметра: r , c_1 и c_2 ; при их выборе можно руководствоваться следующими соображениями. Пусть в окрестность V_{mn} попадает s таких помех, что все они либо меньше, либо больше μ_{mn} . Тогда, очевидно, необходимо выбрать $r = s + 1$. Таким образом, r определяется исходя из вероятности искажения p . Зная r и задавшись определенным θ в выражении (3.1.16), можно определить c_1 исходя из $c_1 M\{d_r(m,n)\} = u\sigma$; в реальных задачах фильтрации рекомендуется выбирать $c_1 \approx 1 \div 1,5$. Чтобы выбрать c_2 , построим распределение вероятностей величины $\zeta_r(m,n) = |x_{mn} - \mu_r(m,n)| - c_1 d_r(m,n)$.

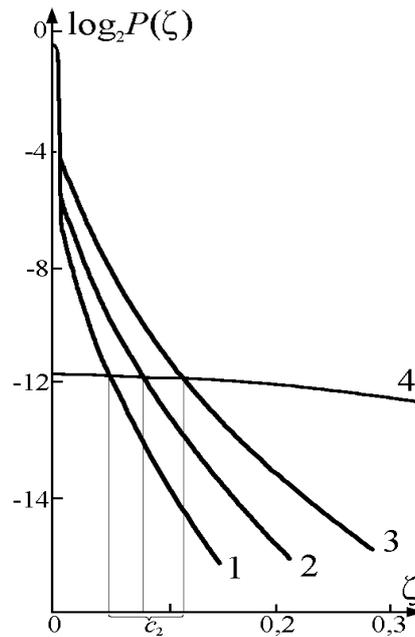


Рис. 3.1.3. Вероятности отклонения значения элемента от границы $u\varepsilon$ -окрестности. 1) $r = 1$; 2) $r = 2$; 3) $r = 3$; 4) распределение вероятности модуля разности элемента и помехи.

Кривые 1–3 на Рис. 3.1.3 показывают значения $\zeta_r(m,n)$ для неискаженных изображений, вычисленные по окрестности 3×3 , а кривая 4 показывает распределение вероятностей величины $|\xi_{mn} - x_{mn}|$, где ξ_{mn} — значение помехи. Точки пересечения кривой 4 с кривыми 1–3 определяют оптимальные значения параметра c_2 при заданных r и c_1 . При больших p , когда вероятность $P\{s_{mn} \geq 2\}$ достаточно высока, рекомендуется итеративный процесс фильтрации: первоначально с большими значениями r , c_1 и c_2 , затем постепенное их уменьшение до $r = 1$ и оптимальных c_1 и c_2 . Заметим, что при $r = n/2$, где n — число элементов в окрестности, алгоритм (3.1.18) будет пороговой медианной фильтрацией.

Экспериментальные данные

Для проверки и сравнения методов фильтрации импульсных помех использовались 16 изображений достаточно высокого качества с различными сюжетами (Рис. 1.1.2), имеющих размеры 256×256 элементов. При реализации алгоритма (3.1.18) окрестность составляла 3×3 элементов. В Табл. 3.1.1 приведены данные, усредненные для указанного набора изображений.

В соответствии с формулой (3.1.1) на изображения наносились импульсные помехи с вероятностью искажения p . Помехи устранялись различными методами; среднеквадратическое отклонение (СКО, σ) от оригинального изображения $\mathbf{X}' = [x'(m,n)]$ измерялось для каждого (i) полученного изображения $\mathbf{X} = [x_{mn}]$:

$$\sigma_i = \left(\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (x_{mn} - x'_{mn})^2 \right)^{1/2}.$$

Усредненное значение СКО подсчитывалось по формуле: $\sigma = (\sum \sigma_i) / 16$. В Таблице 3.1.1 приведены результаты фильтрации алгоритмами (3.1.3)–(3.1.7), (3.1.18) при вероятностях искажения $p = 0,0 \div 0,2$. При $p = 0,0$ фильтрация производилась с теми же значениями параметров, что и при $p = 0,05$. Визуально качество фильтрации можно оценить на Рис. 3.1.4.

Таблица 3.1.1. Усредненные значения СКО фильтрованного изображения от неискаженного оригинала (в градациях сигнала).

	Использованный алгоритм фильтрации	Уровень импульсного шума			
		$p=0,00$	$p=0,05$	$p=0,10$	$p=0,20$
а	линейный алгоритм (3.1.3), матрица \mathbf{V}_1 , $\delta = 0$	3,9	4,8	6,0	7,9
б	линейный алгоритм (3.1.3) с \mathbf{V}_4 , $\delta > 0$	3,6	4,3	5,1	7,1
в	линейный каузальный алгоритм с \mathbf{V}_5	3,3	4,1	4,6	5,6
г	алгоритм (3.1.4) (марковская модель)	–	–	–	6,5
д	алгоритм медианной фильтрации (3.1.5) при $\delta = 0$	3,5	3,7	3,9	4,5
е	пороговая медианная фильтрация (3.1.5), $\delta > 0$	1,9	2,7	3,1	4,0
ж	алгоритм (3.1.7)	2,9	3,1	3,5	4,2
з	алгоритм (3.1.18)	1,2	1,7	2,3	3,2



Рис. 3.1.4. Фильтрация изображения с импульсными помехами: а) исходное изображение; б) изображение с нанесенной импульсной помехой ($p = 0,20$); в) обработка алгоритмом (3.1.18) с 3 итерациями, СКО от оригинала $\sigma = 2,6$ градации.

Эксперименты показали, что разработанный метод ранговой пороговой фильтрации (3.1.18) обеспечивает в $1,5 \div 2$ раза более высокую точность восстановления, чем остальные. Следует отметить, что точные значения среднеквадратического отклонения зависят от конкретного изображения и при прочих равных условиях могут различаться на $20 \div 30\%$;

причинами разницы служат различие резкости, уровня шумов, пространственное разрешение. Данный факт необходимо учитывать при сопоставлении результатов различных исследований.

В развитие изложенного метода можно предложить новый подход к фильтрации помех, основанный на использовании метода декомпозиции (Раздел 2.2). Он заключается в том, что изображение исходно разлагается на кусочно-гладкую и текстурно-детальную компоненты. Первая остается без изменений, а фильтрации подвергается лишь вторая, поскольку шумы будут содержаться именно в ней. После фильтрации компоненты складываются.

Изложенные в разделе 3.1.1 результаты опубликованы в работах [78,179,201].

3.1.2. Фильтрация периодических помех

Периодические помехи — сравнительно частый вид искажений сигнала. На изображениях они могут появляться во время формирования, передачи или дискретизации сигнала вследствие интерференции различных электромеханических или электрических процессов [40]. Наиболее распространенный способ удаления подобных помех — частотная фильтрация. Периодичность помех приводит к появлению всплесков в спектре Фурье, которые могут быть обнаружены и использованы для построения фильтра. Отфильтрованный спектр затем подвергается обратному преобразованию Фурье для получения изображения без помех. Для проведения анализа с целью обнаружения пиков спектра, амплитудный сигнал спектра следует интерпретировать как изображение. Рассмотрим этот вопрос на примере реставрации архивных фотографических снимков с тиснением фотобумаги [190,227].

Фильтрация в частотной области

Фотографические отпечатки с тиснением (текстурным рельефом) фотобумаги не выглядят дефектными, но из-за рельефа поверхности при оцифровке возникает непостоянство углов отражения на каждом из элементов тиснения. В результате изображение искажается множеством мелких пятен, форма и структура расположения которых зависят от конкретной матрицы тиснения. Пример того, как выглядит подобное изображение, представлен на Рис. 3.1.5.

В такой ситуации становятся полезными методы частотной фильтрации. Известно, что в предположении аддитивности, некоррелированности и стационарности сигнала и помехи, наилучшей линейной фильтрацией, с точки зрения минимума среднеквадратического отклонения, является винеровская фильтрация [14,35,368]. Для аддитивной модели, в которой получаемый сигнал есть сумма исходного сигнала и шума

$$g(t) = s(t) + n(t), \quad (3.1.19)$$

где $s(t)$ — полезный сигнал, а $n(t)$ — шум, оптимальный линейный фильтр должен иметь следующую частотную характеристику:

$$C(f) = \frac{P_s(f)}{P_s(f) + P_n(f)}. \quad (3.1.20)$$

Здесь $P_s(f)$ — энергетический спектр сигнала, а $P_n(f)$ — энергетический спектр шума. Этот подход может быть с успехом применен и к двумерным дискретным сигналам — изображениям [40,208], с тем лишь отличием, что частота f становится значением координат в двумерном спектре: $f \Rightarrow (x,y)$, где $x \in [0, X-1]$, $y \in [0, Y-1]$; X, Y — размеры изображения.



Рис. 3.1.5. Исходное изображение с помехами, вызванными тиснением фотобумаги.

При построении фильтра (3.1.20) основная трудность состоит в оценке значений $P_s(x,y)$ и $P_n(x,y)$. К сожалению ни одна, ни другая функция нам неизвестны. Все, чем мы можем воспользоваться — это измерить спектр получаемого сигнала $P_g(x,y) = P_s(x,y) + P_n(x,y)$, т.е. их смеси, и применить какие-то предположения о характеристиках $P_s(x,y)$ и $P_n(x,y)$, исходя из априорных сведений о сигнале и помехе.

Исходный сигнал (изображение) $g(x,y)$ является действительным, следовательно, его Фурье-спектр $F(x,y)$ будет комплексно сопряженным, а энергетический спектр — центрально симметричным. Компоненты спектра, соответствующие низким пространственным частотам изображения, будут располагаться по углам матрицы значений двумерного спектра, а соответствующие высоким пространственным частотам — в ее центре. Такое расположение не вполне удобно для анализа. Как правило, его меняют, диагонально переставляя квадранты матрицы: $\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$; при этом низкие частоты оказываются в центре, а высокие — по углам спектра. В результате нулевая частота будет находиться в точке с координатами $(X/2, Y/2)$. Такая

перестановка эквивалентна предварительному умножению изображения на $(-1)^{x+y}$ [40]. Далее мы будем придерживаться именно такой интерпретации спектра.

На Рис. 3.1.6,а показан энергетический спектр $P_g(x,y)$ изображения, представленного на Рис. 3.1.5,. Изображению без помех соответствует центральное светлое пятно в области низких частот спектра. Спектральные составляющие сравнительно быстро убывают по амплитуде к краям практически до нуля и сливаются с уровнем черного. Тектурная помеха отображена в энергетическом спектре группой ярких пятен (пиков), составляющих вокруг центра периодическую структуру. Для удобства будем называть их *сателлитами*. Геометрическое положение этих пятен определяется пространственной периодичностью и формой элементов тиснения на исходном фотоснимке.

Амплитуда пиков, соответствующих текстурной помехе, во много раз больше амплитуды сигнала в тех же точках, а значит, фазовая составляющая этих компонент спектра практически полностью определяется помехой. Следовательно, даже отыскав и применив оптимальный фильтр $C(f)$ в (3.1.20), восстановить правильную фазу в этих точках не удастся, поскольку фаза меняться не будет, а как известно [107], в формировании структуры изображения фаза спектральных составляющих играет значительно более важную роль, чем амплитуда. Поэтому ненулевые частотные составляющие с не восстановленной фазой по-прежнему будут являться помехой, хотя и со значительно меньшей амплитудой.

Построение фильтра-маски

Энергетический спектр $P(x,y) = P_s(x,y) + P_n(x,y)$ является действительным сигналом, поэтому он может рассматриваться как изображение с соответствующими свойствами. Используем его в качестве исходных данных для обнаружения участков пятен-сателлитов и построения фильтра, применяя к нему методы преобразования и анализа изображений. Будем полагать, что компоненты спектра текстурной помехи, дающие основной вклад, почти полностью сосредоточены в мелких пятнах-сателлитах, причем $P_n(x,y) \gg P_s(x,y)$, если точка (x,y) принадлежит одному из таких пятен, и $P_n(x,y) \ll P_s(x,y)$ в противном случае. С учетом сделанного выше замечания по поводу фазы, не будет большой ошибкой вместо винеровского использовать идеальный (режекторный) фильтр [40], полностью обнуляющий компоненты спектра в заданных точках.

Поскольку значения элементов в пятнах-сателлитах значительно отличаются от значений окружающего фона ($P_n(x,y) \gg P_s(x,y)$), то для обнаружения пятен-сателлитов воспользуемся методами обнаружения объектов заданной площади, разработанными на базе двухмасштабной многокомпонентной модели изображения и алгоритма декомпозиции «D» (2.2.8), и изложенными в разделе 4.1. В данном случае подойдет алгоритм обнаружения объектов с

площадью меньше заданной (4.1.2). Фрагмент анализа при этом рекомендуется выбирать таких размеров, чтобы его площадь не менее чем в 6–8 раз превышала площадь попадающих во фрагмент пятен-спутников. Выберем в (4.1.2) константу отсечения $\delta \approx 0,2$ диапазона яркостей и $C = 0$. В результате сформируется предварительная маска u_{mn} со следующими свойствами:

$$u(x, y) = \begin{cases} 1, & \text{если точка } (x, y) \text{ принадлежит пятну;} \\ 0, & \text{в противном случае.} \end{cases} \quad (3.1.21)$$

Центральное пятно в области низких частот энергетического спектра $P_g(x, y)$ изображения вообще говоря может быть неоднородным. Из-за этого в формируемую методом (3.1.21) маску $u(x, y)$ могут попасть и некоторые участки спектра полезного сигнала, расположенные вблизи центрального пятна, которые необходимо удалить.

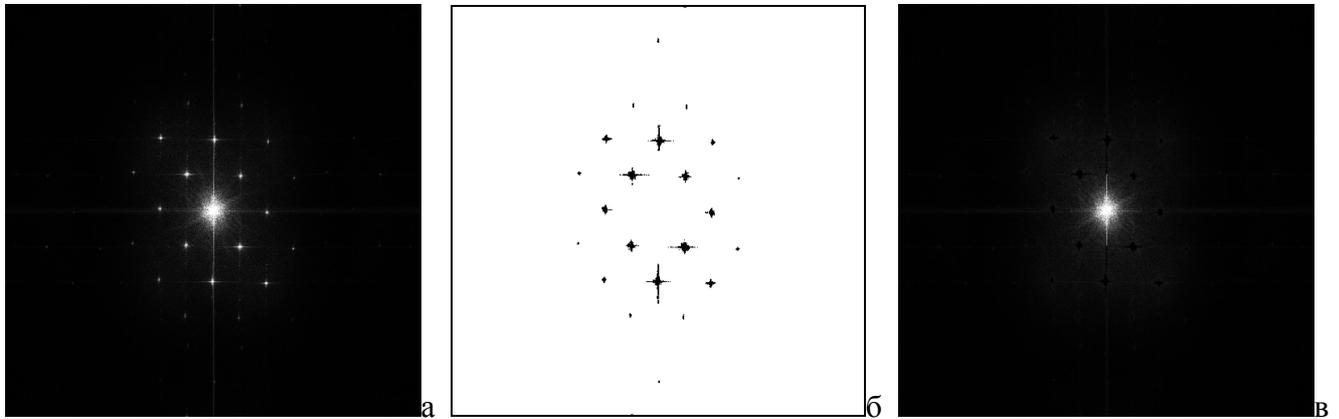


Рис. 3.1.6. а) Спектр изображения с помехой; б) маска фильтра; в) спектр после фильтрации.

Чтобы осуществить это, найдем расстояние от центра спектра $(X/2, Y/2)$ до ближайшего пятна-спутника, соответствующего периодической помехе. Как правило, пятна с наиболее низкой частотой являются наиболее мощными, и соответственно, наибольшими по размерам из образовавшихся пятен в $u(x, y)$. Найти максимальное по размерам сечения пятно можно последовательным применением операции эрозии [40], заключающейся в том, что на каждом шаге происходит стирание граничных элементов пятна. Пятно, стираемое последним, и будет являться наибольшим по размеру. Расстояние от его центра (последней оставшейся точки) до центра спектра $(X/2, Y/2)$ обозначим через R . Стирание точек в центральной части спектра выполняется преобразованием

$$v(x, y) = \begin{cases} u(x, y), & \text{если } ((x - X/2)^2 + (y - Y/2)^2)^{1/2} < \rho R, \\ 0, & \text{для остальных точек.} \end{cases}$$

Коэффициент ρ рекомендуется выбирать в диапазоне $0,8 \div 0,85$.

В завершение $v(m, n)$ подвергается операциям расширения и сглаживания. Расширение осуществляется алгоритмом дилатации [40], состоящим в том, что нулевым элементам маски, являющимся соседями ненулевых элементов, присваивается значение 1. Для сглаживания

используется обычный фильтр, состоящий в вычислении среднего (2.1.2) по небольшой окрестности (например, 3×3 элементов): $C(x,y) = \sum_{ij} v(x+i,y+j)/9$. Сформированный фильтр $C(x,y)$ (белый цвет соответствует 1 фильтра, а черный — 0) показан на Рис. 3.1.6,б.

Выполнение операции фильтрации

Собственно фильтрация осуществляется поэлементным произведением спектра исходного изображения с периодической помехой $F(x,y)$ и полученного фильтра-маски $C(x,y)$:

$$f(x,y) = C(x,y)F(x,y). \quad (3.1.22)$$

Результат операции (3.1.22) показан на (Рис. 3.1.6,в). Обратным Фурье-преобразованием спектра $f(x,y)$ получаем фильтрованное изображение (Рис. 3.1.7).



Рис. 3.1.7. Изображение Рис. 3.1.5 после удаления периодических помех.

Полученные результаты демонстрируют вполне приемлемое качество фильтрованных изображений. Тем не менее, заметно, что абсолютно полного удаления помехи на всех участках изображения достичь не удастся. Основные причины здесь две: во-первых, нестационарность (пространственная неоднородность) помехи на изображении, а во-вторых, возможные нарушения предполагаемой аддитивной модели сигнала и шума (3.1.19), возникающие вследствие неопределенных яркостных преобразований, предваряющих процесс оцифровки изображения. Также из-за эффекта Гиббса остаются следы помех на краях изображения.

Использование для удаления периодических помех подхода, основанного на частотной фильтрации, позволяет корректировать искажения почти не затрагивая полезную информацию и не внося дополнительных артефактов. Методы преобразования, использующие частотную фильтрацию, могут быть достаточно эффективно реализованы при помощи алгоритма быстрого преобразования Фурье (БПФ) [8].

Фильтрация в пространственной области

Важным частным случаем периодических помех являются помехи вдоль линий развертки изображения (одномерные), иногда возникающие в радиотехнических системах передачи информации. Обычно периоды подобных помех сравнительно большие, из-за чего всплески в спектре, соответствующие помехе, располагаются в области низких частот, где расположена основная информационная и энергетическая часть сигнала изображения. По спектру можно определить величину периода помехи p , однако применение для устранения таких помех частотной фильтрации с идеальным фильтром [208], вариант которой был рассмотрен выше, приводит к заметным искажениям исходного сигнала [200] и поэтому нежелательно.

Рассмотрим вопрос фильтрации аддитивной периодической помехи, считая известным ее период p . При этом будем основываться на кусочно-постоянную модель изображения (1.2.24). Т.к. рассматриваемая помеха одномерна, то и изображение в последовательной развертке также будем считать одномерным кусочно-постоянным сигналом: $x(u) = S(u) + \tau(u)$, где $S(u)$ — кусочно-гладкая фоновая составляющая изображения, $\tau(u) = t(u) + \xi(u)$ — сумма текстурной и шумовой составляющих — независимая случайная величина с нулевым средним: $M(\tau) = 0$. На сигнал $x(u)$ накладывается аддитивная периодическая помеха $w(v) = w(v + np)$ ($0 \leq v < p$) с периодом p . Регистрируемый сигнал $y(u)$ имеет вид:

$$y(u) = x(u) + w(u \bmod(p)). \quad (3.1.23)$$

Для подавления неравномерностей по полю кадра воспользуется операцией (3.3.1):

$$\hat{y}(u) = y(u) - \frac{1}{p} \sum_{v=u-p/2}^{u+p/2} y(v) = y(u) - \bar{y}_p(u). \quad (3.1.24)$$

Черта сверху $\bar{y}_p(u)$ означает усреднение на отрезке $[u - p/2, u + p/2]$. Тогда

$$\hat{y}(u) = S(u) - \bar{S}_p(u) + \tau(u) - \bar{\tau}_p(u) + w(u) - \bar{w}_p(u).$$

Форму помехи можно найти усреднением $\hat{y}(u)$ за N периодов $z(v) = (1/N) \sum_{n=0}^{N-1} \hat{y}(v+np)$, ($0 \leq v < p$).

Раскрывая $\hat{y}(u)$ получим:

$$z(v) = \frac{1}{N} \sum_{n=0}^{N-1} [S(v+np) - \bar{S}_p(v+np) + \tau(v+np) - \bar{\tau}_p(v+np) + w(v+np) - \bar{w}_p(v+np)]. \quad (3.1.25)$$

Рассмотрим члены под знаком суммы:

$$\bar{w}_p(v + np) = (1/p) \sum_{x=u-p/2}^{u+p/2} w(x) = C, \text{ где } C \text{ — среднее значение помехи } w(v) \text{ за период;}$$

$$(1/N) \sum_{n=0}^{N-1} w(v + np) = w(v);$$

поскольку $\tau(u)$ — случайная величина с нулевым средним, то

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \tau(v + np) = 0 \quad \text{и} \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \bar{\tau}(v + np) = 0.$$

Тогда, при $N \rightarrow \infty$, (3.1.25) представится в виде:

$$z(v) = w(v) + \lim_{N \rightarrow \infty} L(v) - C, \quad (0 \leq v < p), \quad \text{где } L(v) = \frac{1}{N} \sum_{n=0}^{N-1} (S(v+np) - \bar{S}_p(v+np)). \quad (3.1.26)$$

Для реальных изображений $L(v)$ — случайная величина с нулевым средним. Считая $\lim L(v) = 0$ при $N \rightarrow \infty$, получим: $w(v) = z(v) + C$, ($0 \leq v < p$), где C — среднее значение периодической помехи $w(v)$ за период p . После того, как найдена $w(v)$, фильтрация сигнала $y(u)$ согласно (3.1.23) сводится к операции вычитания:

$$x(u) = y(u) - z(u \bmod(p)) - C. \quad (3.1.27)$$

Формула фильтрации (3.1.27) получена для одномерного сигнала $y(u)$, т.е. для изображения в последовательной развертке. При этом неявно предполагается непрерывность фазы периодической помехи при переходе от конца одной строки к началу следующей. Если длина строки изображения кратна периоду помехи, или если фаза помехи для всех строк одинакова, то проблем не возникает — во втором случае при анализе достаточно просто отбрасывать элементы в конце каждой строки, не захватывающие полный период.

Сложности возникают, когда при переходе от строки к строке фаза помехи изменяется и необходима корректировка фазы на каждой строке отдельно. Для этого требуется привязка фазы, которую можно осуществить методами корреляционного совмещения сигнала $z(v)$, получаемого согласно формуле (3.1.25) по каждой из строк. Поскольку в таком случае N сравнительно мало, рекомендуется в (3.1.25) вместо среднего вычислять значение медианы: $z_M(v) = \text{med}_n(\hat{y}(v + np))$, после чего осуществлять *циклическое* корреляционное совмещение массивов $z_M(v)$ по строкам. Именно таким способом корректировалась фаза при устранении периодической помехи на изображении Рис. 3.1.8,а. Результат фильтрации показан на Рис. 3.1.8,б. Точность восстановления (СКО от оригинала) составила $\sigma = 0,19\%$.

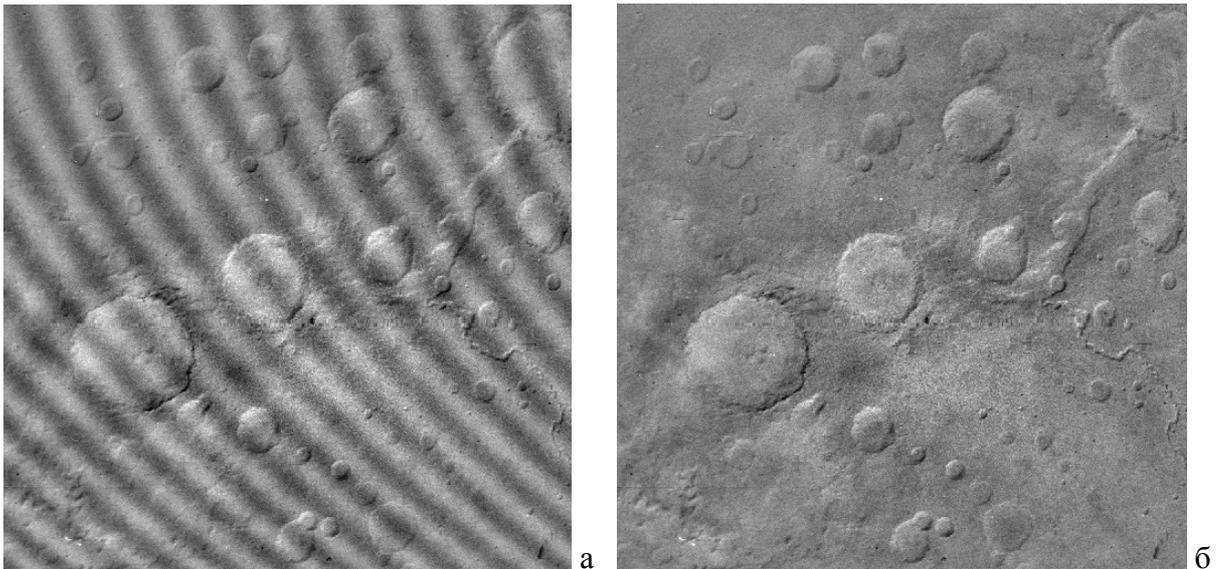


Рис. 3.1.8. а) Снимок поверхности Марса, искаженный периодической помехой.
б) Результат фильтрации.

Изложенные в разделе 3.1.2 результаты опубликованы в работах [190,200,227].

3.2. Автоматическая яркостная (градационная) коррекция

3.2.1. Стандартные подходы

Задача градационной коррекции

Процессы регистрации или переноса изображения с одного носителя на другой, как, например, фотопроцесс, часто сопровождаются заметными амплитудными (фотометрическими) искажениями, которые могут возникать на всех этапах, вплоть до преобразования изображения в цифровую форму. Причинами возникновения таких искажений могут являться неточности экспозиции, чрезмерный контраст деталей объекта съемки, нелинейности характеристик регистрирующего устройства и аналого-цифрового преобразователя и т.д. Для исправления подобных искажений используется градационная коррекция; кроме того, она позволяет согласовать характеристики изображения и процесса его последующей обработки и/или воспроизведения.

Задача автоматической градационной коррекции видеоинформации, поступающей в виде неподвижных или движущихся изображений, является одной из наименее исследованных, несмотря на то, что ей уделяется внимание почти в каждой монографии, посвященной обработке изображений. В классической постановке данная задача часто формулируется как приведение распределения вероятностей значений элементов изображения, т.е. его гистограммы, к желаемому виду [40,113,343]. Однако традиционный подход, основанный лишь на анализе гистограммы получаемого изображения [260,264,277,315], не в состоянии предложить удовлетворительного решения проблемы, что связано с существенным влиянием самого сюжета сцены на вид гистограммы, а значит и на форму градационного преобразования сигнала. Основной недостаток такого подхода в том, что форма распределения значений получаемого сигнала одновременно зависит как от характеристик регистрирующего устройства, так и от самой наблюдаемой сцены. Поэтому попытаемся подойти к решению данного вопроса с другой стороны — отыскать какие-то измеримые характеристики реальных изображений, которые могли бы лечь в основу построения *функции градационной коррекции* (иногда называемой *функцией гамма-коррекции*).

Ниже осуществляется поиск некоторых общих статистических закономерностей, характерных для изображений высокого визуального качества, которые могли бы помочь оценить градационные искажения и построить подходящую корректирующую функцию для улучшения изображения. Настоящий подход базируется на оценивании локальных контрастов изображения, с помощью которых формулируется метод градационной коррекции, названный

метод выравнивания локальных контрастов. Данный метод может быть использован в условиях, когда априорные сведения о виде градационного искажения, которому было подвергнуто полученное изображение, отсутствуют, и единственным источником информации может служить лишь получаемое (предположительно искаженное) изображение.

Наряду с этим рассматривается алгоритм, названный *эквализацией клиптированной гистограммы*, эффективность и сравнительно невысокая сложность которого позволяют реализовывать его в системах реального времени (например, в устройствах телевизионного ввода, видеокамерах с автоматической коррекцией сигнала и т.п.).

Формализация и используемые методы

Говоря об амплитудных искажениях предполагается, что исходное изображение $\tilde{\mathbf{X}} = [\tilde{x}_{mn}]$ подвергается поэлементному монотонному пространственно-инвариантному преобразованию $F(z)$, в результате чего регистрируется изображение $\mathbf{X} = [x_{mn}]$, являющееся функцией

$$x_{mn} = F(\tilde{x}_{mn}). \quad (3.2.1)$$

Чтобы вернуться к исходному сигналу, необходимо выполнить обратное градационное преобразование. Для этого требуется найти функцию $f(z) = F^{-1}(z)$, определив которую можно восстановить изображение

$$y_{mn} = f(x_{mn}).$$

Строгое выполнение равенства $y_{mn} = \tilde{x}_{mn}$, да и само существование $f(z)$ возможны только лишь при условии непрерывного представления значений x_{mn} . В случае квантованных значений яркости, когда $f(z)$ задана на дискретном множестве $[0, K-1]$, восстановление может быть осуществлено лишь с определенной степенью точности. Так, если на участке $[k_1, k_2]$ градационной характеристики требуется поднять контраст в R раз, то результирующие значения окажутся отстоящими друг от друга в среднем на R и $(k_1 - k_2)(R - 1)$ градаций окажутся неиспользованными, что может привести к появлению на изображении так называемых «ложных контуров». Подобных искажений зачастую удается избежать, если добавить к значению сигнала изображения случайный шум. В [30] показано, что оптимальным является равномерный шум, дисперсия которого составляет $1/3$ величины интервала квантования.

Выбор обратной функции $f(z)$ является наиболее сложным вопросом градационной коррекции, т.к. прямая функция определяется суммарными искажениями, возникающими на всех последовательных этапах формирования, регистрации и преобразования каждого отдельного изображения, которые часто остаются неизвестными. В таких случаях приходится находить характеристику преобразования на основе анализа полученного (искаженного) изображения, пользуясь лишь свойствами изображений и предположениями о возможном виде

преобразования $F(z)$. Рассмотрим виды преобразований, обычно используемых для градационной коррекции.

Наиболее распространенным и простейшим вариантом градационной коррекции является линейное преобразование вида

$$y_{mn} = ax_{mn} + b \quad \text{или} \quad y_{mn} = (K - 1)(x_{mn} - k_1)(k_2 - k_1), \quad (3.2.2)$$

где $k_1 = -b/a$ и $k_2 = (K - b)/a$. Способ задания преобразования через параметры k_1 и k_2 часто удобнее, поскольку они измеряются в тех же единицах, что и x_{mn} . Если $x_{mn} < k_1$ или $x_{mn} > k_2$, то результату y_{mn} присваиваются значения 0 или $(K - 1)$ соответственно. Применение подобных преобразований оправдано в тех случаях, когда нелинейности искажающей функции $F(z)$ малы по сравнению с требуемой точностью воспроизведения изображения.

Значения k_1 и k_2 обычно находят по гистограмме $h(k)$ изображения, которая определяет распределение вероятностей значений элементов:

$$h(k) = P(x_{mn} = k); \quad \sum_{k=0}^{K-1} h(k) = 1. \quad (3.2.3)$$

Чаще всего используется метод квантилей, когда k_1 и k_2 вычисляются через порядковые статистики (2.1.3) при задаваемом значении параметра ε : $k_1 = R(\varepsilon)$; $k_2 = R(1 - \varepsilon)$. Иногда основываются на предположениях инвариантности статистических параметров гистограммы, например, что она должна иметь заданные среднее значение M_0 и дисперсию σ_0^2 . Тогда, определив соответствующие параметры M и σ^2 исходного изображения \mathbf{X} , легко найти параметры преобразования как

$$k_1 = \frac{K}{2} \left(\frac{M}{M_0} - \frac{\sigma}{\sigma_0} \right) \quad \text{и} \quad k_2 = \frac{K}{2} \left(\frac{M}{M_0} + \frac{\sigma}{\sigma_0} \right).$$

Из нелинейных градационных преобразований используются методы модификации (приведения) гистограммы [276], а именно: результат обработки исходного изображения \mathbf{X} есть изображение \mathbf{Y} с заданной формой гистограммы $h_y(k)$. Пусть $f(k)$ — неубывающая функция (т. е. если $k_1 < k_2$, то $f(k_1) \leq f(k_2)$), определяющая поэлементное градационное преобразование изображения \mathbf{X} в \mathbf{Y} : $y_{mn} = f(x_{mn})$, а $h_x(k)$ и $h_y(k)$ — гистограммы изображений \mathbf{X} и \mathbf{Y} . В этом случае с точностью, определяемой числом уровней квантования, справедливо:

$$\sum_{k=0}^z h_x(k) = \sum_{k=0}^{f(z)} h_y(k).$$

Выберем желаемую функцию $G(k)$ распределения вероятностей значений элементов результирующего изображения \mathbf{Y} . Пусть $G^{-1}(z)$ — обратная к ней функция, т. е. $G^{-1}[G(k)] = k$. Тогда формула градационного преобразования $f(k)$, приводящая изображение \mathbf{X} с функцией распределения $H_x(k)$ в изображение \mathbf{Y} с функцией распределения $G(k)$, будет иметь вид

$$f(k) = G^{-1}[H_x(k)]. \quad (3.2.4)$$

Выбор функции распределения $G(k)$ может осуществляться на основании требований, предъявляемых к результатам обработки.

Формула (3.2.4) является канонической для построения алгоритмов модификации гистограммы. Для важного частного случая эквализации гистограммы [276], когда $h_y = \text{const}$ а, $G(k) = 1$, получим:

$$f_0(k) = H_x(k) = \sum_{z=0}^k h_x(z). \quad (3.2.5)$$

Иногда решается задача цифрового препарирования изображений [9] — выделения одних информационных элементов изображения при ослаблении некоторых других. В таких случаях используют иные критерии выбора $h_y(k)$, среди которых следует упомянуть гиперболизацию и степенную интенсификацию гистограммы [9,257,367].

3.2.2. Алгоритм автоматической градационной коррекции

Функция локальных контрастов

Если исходную гистограмму $h_x(k)$ измерять не по всему изображению, а лишь по отдельным его участкам, то это приведет к соответствующим изменениям как ее самой, так и обратной функции $f(z)$. Выбор таких сюжетно важных участков индивидуален для каждого изображения. В некоторых случаях предлагается для этой цели использовать так называемые информационные участки изображения — вблизи контуров [88]; утверждается, что такие гистограммы должны быть близки к распределению нормальной случайной величины.

Рассмотренные выше методы коррекции основаны на априорных предположениях о распределении вероятностей значений элементов и не учитывают межэлементные статистические связи на изображении, хотя достаточно очевидно, что амплитудные искажения должны приводить также и к их изменениям. Поставим задачу определения функции $f(z)$ по искаженному изображению, пользуясь моделью, рассмотренной в разделе 1.1.

Учтем амплитудное преобразование (3.2.1) в модели окрестности (1.2.4). Постоянная составляющая μ_{mn} изменится на $F(\mu_{mn})$; текстурная $\rho^r g^r$, в предположении достаточной гладкости $F(z)$, на $F'(\mu_{mn})\rho^r g^r + o(F'(\mu_{mn}))$, где $F'(\mu_{mn})$ — производная функции $F(\mu_{mn})$. Полагаем, что составляющая ξ^r , определяемая зернистостью фотопленки, шумами устройства ввода и т.д., добавляется после амплитудного преобразования, поэтому не изменяется. Считая $o(F'(\mu_{mn}))$ пренебрежимо малым, получим:

$$x_{mn}^r = F(\mu_{mn}) + F'(\mu_{mn})\rho^r g_{mn}^r + \xi_{mn}^r. \quad (3.2.6)$$

Аналогично, значения параметров μ_0, α_0, β_0 в (1.2.5) изменятся на:

$$\mu = F(\mu_0); \quad \alpha = F'(\mu_0)\alpha_0; \quad \beta = F'(\mu_0)\beta_0; \quad (3.2.7)$$

а также

$$g = F'(\mu_0)g_0. \quad (3.2.8)$$

Таким образом определено, как амплитудные искажения влияют на изменение значений параметров модели окрестности. Усредним g_{mn} по изображению следующим образом. Для каждого значения k выделим множество v_k состоящее из $N(v_k)$ элементов изображения таких, что для их окрестностей $\mu_{mn} = k$, и подсчитаем величину

$$C(k) = \frac{1}{N(v_k)} \sum_{x_{mn} \in v_k} g_{mn}. \quad (3.2.9)$$

Будем называть зависимость $C(k)$ *функцией локальных контрастов* изображения. Для изображения с исходной функцией локальных контрастов $C_0(k)$, искаженного амплитудным преобразованием $F(z)$, получим:

$$C(F(k)) = F'(k)C_0(k) + o(F'(k)C_0(k)). \quad (3.2.10)$$

Последнее слагаемое опустим, считая его достаточно малым. Полагая, что для $F(z)$ существует обратная функция $f(z)$, запишем формулу восстановления функции локальных контрастов:

$$C_0(f(k)) \approx f'(k)C(k). \quad (3.2.11)$$

Зависимость $C(k)$ определяется на основании статистического анализа полученного изображения X . Если мы сумеем задаться какими-либо априорными сведениями о виде функции локальных контрастов $C_0(k)$ неискаженного изображения, то с помощью (3.2.11) легко найти требуемую функцию градиентной коррекции $f(k)$.

Гипотеза о константности функции локальных контрастов

Согласно модели (1.2.9), статистические свойства случайной величины g_{mn} будут существенно различаться в зависимости от принадлежности точки (m,n) множеству контурных или фоновых элементов. Соответственно, вид функции локальных контрастов $C(k)$ также зависит от выбора того или иного множества элементов изображения. Если для определения $C(k)$ использовать фоновые элемент, то возникают два затруднения: а) статистические свойства текстур различных областей изображения могут заметно различаться, а следовательно, g_{mn} и $C(k)$ будут в значительной степени определяться ими, а не реальными амплитудными искажениями; б) поскольку дисперсия текстурной компоненты сравнительно невелика, то вероятна ситуация, когда моды распределения величины μ_{mn} по множеству фоновых элементов не перекрываются, из-за чего появятся незаполненные участки диапазона яркостей, что в конечном итоге приведет к неопределенности функции локального контраста в данных точках диапазона.

Более предпочтительным выглядит использование множества контурных элементов изображения. Значения g_{mn} на контурных участках определяются яркостной разницей $|S^i - S^j|$ на

контурном перепаде между соседними областями U^i и U^j и пространственной шириной перепада. Количество элементов изображения, попадающих на контурные участки, зависит от двух факторов: резкости исходного изображения и степени его дискретизации, определяющей пространственную ширину контурных перепадов. Отметим, что величина шага дискретизации одинаково (линейно) влияет на g_{mn} во всех точках изображения.

В силу влияния разнообразных случайных факторов и при условии достаточно большой репрезентативности множества контурных элементов естественно сделать предположения: а) на контурном перепаде между соседними областями U^i и U^j вероятность значения g_{mn} распределена равномерно на $[S^i, S^j]$; б) крутизна контурных перепадов между соседними областями не зависит от их абсолютных значений S^i или S^j и может считаться одинаковой на всем диапазоне значений $[S_{\min}, S_{\max}]$; в) множество отрезков $[S^i, S^j]$, определяемых различием яркостей соседних областей и, соответственно, задающих амплитуду контурного перепада, покрывают почти все множество яркостей изображения $[B_{\min}, B_{\max}]$.

Из сделанных предположений вытекает, что для неискаженного изображения значения функции $C_0(k)$, определяемые по множеству контурных элементов, должны быть одинаковыми во всем диапазоне значений яркостей $[k_{\min} \leq k \leq k_{\max}]$. Таким образом можно принять гипотезу о константности функции локальных контрастов на изображении $C_0(k) \approx C_0$.

Представляет безусловный интерес проверка гипотезы о константности функции локальных контрастов. Для этого необходимо измерить $C(k)$ по формуле (3.2.9) на каких-либо неискаженных изображениях, которые можно было бы принять за эталонные. К сожалению, выбрать изображения, про которые можно было бы с уверенностью сказать, что они не подвергались неконтролируемым градационным искажениям (либо форма таких искажений была достаточно точно известна — в этом случае их легко исправить), оказалось затруднительным. Поэтому для анализа был выбран набор изображений высокого визуального качества, предполагавшихся искаженными незначительно.

Эксперименты показали, что функция локальных контрастов $C(k)$, измеренная на множестве контурных точек, действительно близка к константной. И наоборот, при измерении по всему изображению, наблюдается антикорреляция данной функции и гистограммы изображения: впадины $C(k)$ совпадают с пиками гистограммы $H(k)$. В качестве примеров на Рис. 3.2.1,а,б показаны графики зависимостей функций локальных контрастов $C(k)$ для изображений 6 и 13 на Рис. 1.1.2. Графики (1) изображают функции $C_1(k)$, подсчитанные по всем элементам изображения, на графиках (2) — функции $C_2(k)$, подсчитанные по множествам контурных элементов. Гистограммы $h(k)$ изображений представлены графиками (3). Видно, что зависимости $C_2(k)$, измеренные по контурным участкам изображения, близки к константным.

Наряду с этим графики $C_1(k)$, соответствующие измерениям по всем элементам изображения, считать равномерными нельзя. Также наблюдается заметная корреляция локальных максимумов $h(k)$ и минимумов $C_1(k)$. Это говорит о снижении локальных контрастов в области пиков гистограммы, вызванном малой дисперсией текстурной компоненты на фоновых участках изображения.

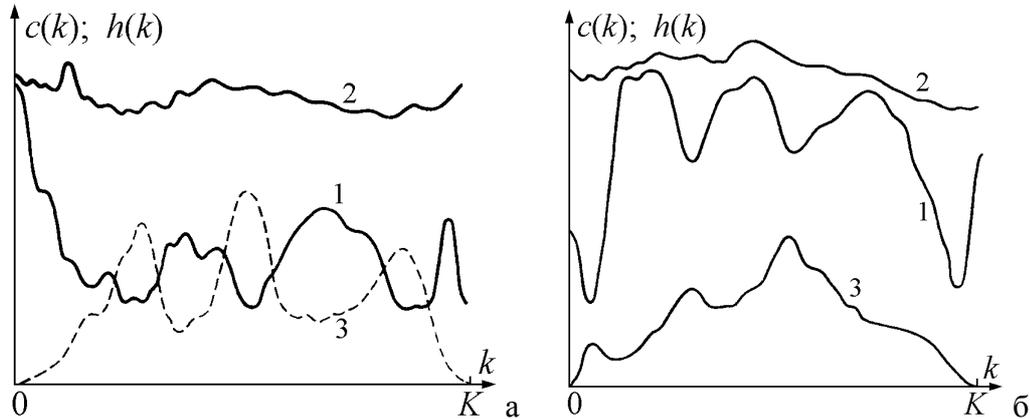


Рис. 3.2.1. Графики функций локальных контрастов $C(k)$ двух реальных изображений — №6 (а) и №13 (б) на Рис. 1.1.2: 1 — $C_1(k)$ по всем элементам изображения; 2 — $C_2(k)$ по контурным элементам; 3 — гистограмма $h(k)$ по всему изображению.

Таким образом проведенные исследования подтверждают гипотезу о константности функция локальных контрастов $C(k)$, измеренной на множестве контурных точек изображения.

Алгоритм выравнивания функции локальных контрастов

На основании гипотезы о константности функции локальных контрастов, из (3.2.11) следует: $f'(k) \approx C_0/C(k)$. Переходя к дискретным значениям, просуммировав от 0 до k , получим:

$$f(k) \approx C_0 \sum_{z=0}^k \frac{1}{C(z)}.$$

Вводя естественную нормировку $f(K-1) = (K-1)$, найдем:

$$C_0 \approx (K-1) \left/ \sum_{k=0}^{K-1} \frac{1}{C(k)} \right.$$

Отсюда получим окончательное приближенное выражение для функции градационной коррекции $f(k)$, выравнивающей функцию локальных контрастов на изображении [173,229,243]:

$$f(k) \approx (K-1) \sum_{z=0}^k \frac{1}{C(z)} \left/ \sum_{k=0}^{K-1} \frac{1}{C(k)} \right. \quad (3.2.12)$$

Проводились эксперименты по обработке ряда изображений алгоритмом (3.2.12), выравнивающим функцию локальных контрастов для контурных участках. Визуальная экспертиза показала, что обработанные таким образом изображения являются для большинства наблюдателей предпочтительными как по сравнению с исходными, так и по сравнению с обработанными другими алгоритмами градационной коррекции.

Определенный недостаток разработанного алгоритма — необходимость предварительного выделения контурных участков на изображении. Действительно, поскольку градационные искажения заметно влияют на результаты выделения контуров, решение данной задачи не может быть вполне корректным. Преодолеть указанную сложность удастся построением серии итераций, объединяющей в каждом цикле операции обнаружения контуров и выравнивания локальных контрастов.

При выполнении итераций, изменению подвергается только первый шаг цикла, на котором происходит обнаружения контуров. Осуществляется это путем модификации значения порога δ , используемого в алгоритме обнаружения контуров для разделения множества элементов изображения на контурные и фоновые. Второй шаг — алгоритм выравнивания локальных контрастов — остается без изменений. На первой итерации следует выбирать уменьшенное значение порога δ и повышать его при каждой следующей итерации до оптимального для используемого алгоритма обнаружения контуров.

Если выделить контурную компоненту на изображении достаточно надежно не удастся а итерационный подход по каким-то причинам не применим, то можно предложить небольшую модификацию алгоритма, позволяющую использовались все точки изображения без их предварительного разделения. Заметим, что величина $N(v_k)$ в (3.2.9) эквивалентна числу точек изображения яркости k , т.е. значению гистограммы изображения $h(k)$. Как уже говорилось выше, функция локальных контрастов $C(k)$ в этом случае будет иметь заметные спады при значениях аргумента, соответствующих яркостям протяженных объектов (пикам гистограммы), и потребуются компенсация влияния множества внутренних точек областей.

Возможны различные подходы к тому, как подавить это влияние. Одним из наиболее простых и распространенных является введение компенсационного члена в формулу (3.2.9), для чего в знаменателе к $h(k)$ прибавляется небольшая по сравнению с общим числом точек изображения $M \times N$ величина $o(M \times N)$:

$$C_H(k) = \frac{1}{h(k) + o(M \times N)} \sum_{x_{mn} \in v_k} g_{mn} .$$

Такое добавление $o(M \times N)$ уменьшает $C_H(k)$ при малых $h(k)$, компенсируя тем самым подавление функции локальных контрастов $C_H(k)$ при больших $h(k)$. Полученное значение $C_H(k)$ должно быть подставлено в (3.2.12), как и ранее.

Экспериментальные оценки алгоритма выравнивания локальных контрастов (3.2.12) показали его эффективность для улучшения изображения и предпочтительность по сравнению с другими алгоритмами автоматической градационной коррекции. Можно рекомендовать данный способ коррекции для обработки большинства изображений низкого качества. Дальнейшим

шагом по пути улучшения визуального качества изображений может являться применение методов повышения локальных контрастов, которое рассматривается в разделе 3.3.

Эквализация клипированной гистограммы

Рассматривая задачу градационной коррекции в контексте улучшения визуального качества изображений следует заметить, что наблюдателю чаще требуется скомпенсировать недостатки освещения реальной сцены, чем возможные нелинейности регистрирующих устройств. Более того, существуют многие задачи, в которых наблюдатель при визуальном анализе хотел бы получить изображение с максимально возможным контрастом, причем зачастую он готов мириться с искажениями абсолютной разницы яркостей объектов, если сами объекты и их детали при этом оказываются хорошо различимыми.

В ряде приложений не требуется решения задачи компенсации искажений яркости. Достаточным является максимальное использование имеющегося диапазона яркостей, однако требуется использовать максимально простой в реализации алгоритм. Подобные задачи ставятся при разработке устройств реального времени, когда алгоритм (3.2.12) оказывается недопустимо сложным. В других случаях влияют специфические условия съемки, например при наличии яркого объекта на темном фоне возникает значительный перепад яркости, который хотелось бы сократить, но не удается сделать при использовании алгоритма (3.2.12).

Сформулируем задачу градационной коррекции в следующем виде: усилить контраст в представительных участках диапазона яркостей за счет сжатия незанятых его участков, при условии, что контраст результата не может превышать некоторого порогового значения.

Данную задачу можно решить различными путями. В простейшем виде это можно сделать следующим образом. Распределение значений элементов изображения описывается его гистограммой: $P\{x = k\} = h(k)$ (1.2.1). В соответствии с поставленной задачей, весь диапазон значений яркости $[0, K-1]$ может быть разбит на набор интервалов $[k_i, k_{i+1}]$, где $h(k)$ считается или имеющим некоторое представительное значение (т.е. больше некоторого порога значимости T_0), или равно нулю (в случае меньше порога). Примем, что $h(k)$ равно нулю в каждом четном интервале $[k_{2i}, k_{2i+1}]$, и отлично от нуля в каждом нечетном $[k_{2i+1}, k_{2i+2}]$. Тогда для решения задачи достаточно удалить все четные интервалы, затем слить вместе все нечетные, и наконец линейно растянуть полученный набор на весь диапазон значений $[0, K]$.

Можно предложить другой простой и обладающий достаточной устойчивостью алгоритм [161, 243]. Существует хорошо известный метод эквализации гистограммы (3.2.5) [276], обладающий способностью удалять неиспользуемые участки диапазона яркостей и растягивать оставшиеся зоны. Его недостаток состоит в значительной неравномерности преобразования контраста — коэффициент усиления контраста значений яркости k пропорционален частоте

появления точки данной яркости, т.е. значению гистограммы $h(k)$. Однако данный алгоритм удается применить к нашей задаче, модифицировав следующим образом.

Пусть заданы два порога: T_0 в качестве порога значимости, и T_1 в качестве порога ограничения (клиппирования): $T_0 \leq T_1$. Если $P\{k\} \geq T_0$, будем считать, что яркость k является значимой и данная точка добавляется к ближайшему нечетному интервалу (или будет началом нового такого интервала). Наоборот, если $P\{k\} < T_0$, точка добавляется к четному интервалу нулей. Дополнительно, если $P\{k\} \geq T_1$, то значение $h(k)$ ограничивается до уровня T_1 . Тем самым мы можем записать следующую формулу преобразования (клиппирования) гистограммы изображения:

$$h_c(k) = \begin{cases} T_0, & \text{если } h(k) < T_0; \\ h(k), & \text{если } T_0 \leq h(k) \leq T_1; \\ T_1, & \text{если } h(k) > T_1. \end{cases} \quad (3.2.13)$$

Теперь для получения функции градационного преобразования $F(k)$ достаточно применить к $h_c(k)$ обычный алгоритм эквализации гистограммы (3.2.5). Окончательное уравнение для вычисления преобразования *эквализации клиппированной гистограммы* $F(k)$ будет следующим:

$$F(k) = (K-1) \sum_{i=0}^k h_c(i) / \sum_{k=0}^{K-1} h_c(k). \quad (3.2.14)$$

Предложенный алгоритм обеспечивает близкий к максимально возможному контраст изображения при отсутствии ограничения сигнала. Преимуществом его является достаточно простой алгоритм анализа — требуется лишь подсчет гистограммы. Данный алгоритм может быть с эффективностью использован в системах реального времени.

Эксперименты по автоматической градационной коррекции

Для тестирования алгоритма выравнивания локальных контрастов (3.2.12) были выбраны различные по сюжету изображения, разделенные на следующие три группы: а) изображения низкого визуального качества; б) эталонные изображения высокого качества, искусственно искаженные градационным преобразованием неизвестной формы; в) те же эталонные изображения в неискаженном виде. После обработки алгоритмом (3.2.12) для каждого изображения определялись две оценки: оценка абсолютного визуального качества и оценка относительного качества по сравнению с исходным изображением (лучше или хуже оригинала).

Общая оценка применимости алгоритма (3.2.12) показала, что он является вполне удовлетворительным для большинства изображений. Изображения низкого качества (а) были улучшены как по абсолютной так и по относительной оценкам. Искаженные изображения (б) также были улучшены по сравнению с искаженным экземпляром, но в некоторых случаях качество обработанного изображения оставалось хуже первоначального неискаженного ориги-

нала. Эталонные изображения (в) после преобразования разделились на три приблизительно равные группы: группу практически не изменившихся изображений, группу измененных, но имеющих близкое визуальное качество, и группу изображений ухудшенного качества.

Объяснение факту ухудшения некоторых эталонных изображений можно дать следующее. Изображения высокого качества, как правило, занимают полный диапазон яркостей и обладают высоким контрастом между протяженными объектами. Алгоритм (3.2.12) повышает контраст деталей в областях, соответствующих участку гамма-характеристики пониженного контраста (такowymi, как правило, являются затененные или яркие области). Для усиления контраста указанных деталей необходимо растянуть соответствующую часть диапазона, а значит снизить контраст между объектами в остальных частях диапазона яркостей, что часто и воспринимается как ухудшение визуального качества изображения.

Пример коррекции показан на Рис. 3.2.2. Исходное изображение (а) было преобразовано алгоритмом (3.2.12). Результат представлен на рисунке (б). На рисунке (в) показаны графики зависимостей: 1) $C(k)$ — функция локальных контрастов исходного изображения (а), полученная по формуле (3.2.9), 2) $f(k)$ — функция градационной коррекции (3.2.12), 3) $h(k)$ — гистограмма исходного изображения (а); 4) распределение значения $\mu(k)$ в (3.2.7) для множества контурных элементов, по которым происходит анализ при определении функции $C(k)$. Масштаб графика по оси ординат соответствует кривой 2; оставшиеся кривые (1, 3 и 4) нормированы по максимальным значениям — их масштабные параметры не существенны.

Как можно видеть, форма графика $C(k)$ совершенно не похожа как на форму гистограммы $h(k)$, так и на форму величины $\mu(k)$ — распределения значений точек, попадающих на контуры. Визуальная экспертная оценка показала, что изображение улучшилось.

Алгоритм эквализации клиппированной гистограммы (3.2.14) тестировался как на неподвижных изображениях, так и в реальных видеосъемках. Визуальное качество выходных изображений было оценено как вполне удовлетворительное. Пример обработки изображения алгоритмом (3.2.14) представлен на Рис. 3.2.3. Исходное изображение (а), представляющее собой светлый объект на практически ровном темном фоне, было обработано стандартным алгоритмом эквализации гистограммы (3.2.5) и алгоритмом эквализации клиппированной гистограммы (3.2.14). Как видно на рисунке (б), стандартный алгоритм эквализации осветляет фон и практически не приводит к улучшению изображения. Предлагаемый вариант (3.2.14), показанный на рисунке (в), является предпочтительным.

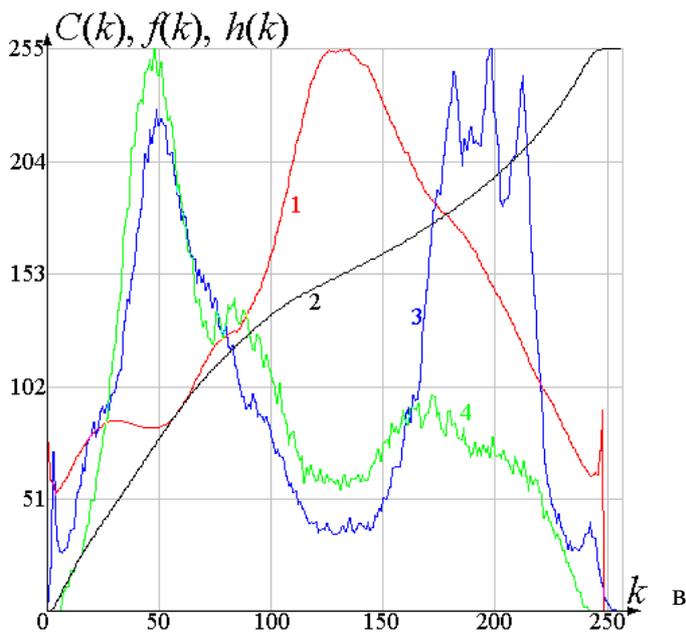


Рис. 3.2.2. Выравнивание локальных контрастов: а) исходное изображение; б) результат коррекции; в) графики: 1) функция локальных контрастов $C(k)$; 2) функция градационной коррекции $f(k)$; 3) гистограмма исходного изображения $h(k)$; 4) распределение $\mu(k)$ по контурным точкам.

Учитывая эффективность и простоту реализации алгоритма эквализации клипированной гистограммы (3.2.14), он может быть рекомендован для применения в приложениях, связанных с обработкой видеоинформации в реальном (телевизионном) масштабе времени и при обработке данных в системах технического зрения. Он был реализован в экспериментальной цифровой видеокамере кругового обзора, которая разрабатывалась компанией L.H. Conseil Optronic

(Франция) с участием автора. Для подсчета гистограммы $h(k)$ и выполнения табличного преобразования использовалась интегральная микросхема L64250 (LSI Logic Corporation). Статистические данные передавались процессору Motorola 68020 для выполнения клиппирования и эквализации по формулам (3.2.13) и (3.2.14). Результирующая функция градационного преобразования $F(k)$ загружалась обратно в L64250 для обработки видеосигнала в реальном времени. Задержка по времени между получением анализируемого сигнала и загрузкой таблицы градационного преобразования составляла один видео кадр [243].



Рис. 3.2.3. а) Исходное изображение; б) стандартный алгоритм эквализации гистограммы (3.2.5); в) эквализация клиппированной гистограммы (3.2.14).

Другим возможным применением алгоритма эквализации клиппированной гистограммы (3.2.14) может быть сокращение диапазона значений сигнала — преобразование входного сигнала, имеющего изначально более 8 значащих разрядов яркости, в стандартный диапазон из 2^8 значений. Примером тому может являться так называемая HDR-технология (High Dynamic Range Imaging), в которой для формирования сцены с широким диапазоном яркостей регистрируются нескольких изображений, каждое из которых покрывает лишь часть диапазона. Впоследствии из них синтезируется единое изображение, охватывающее полный диапазон яркостей. Возможности применения рассмотренных алгоритмов градационной коррекции к обработке цветных изображений рассматриваются ниже в разделе 3.4.

Изложенные в разделе 3.2 результаты опубликованы в работах [17,161,173,179,189,197, 229, 241,243].

3.3. Улучшение изображений повышением локальных контрастов

В начале главы говорилось, что одним из способов улучшения изображений является повышение заметности деталей, достигаемое при помощи усиления их локальных контрастов по отношению к окружающему фону. Из этого с очевидностью вытекает подход к решению: разделить сигнал изображения на две компоненты (фоновую и детальную), осуществить

контрастирование детальной и добавить ее обратно к фоновой компоненте. Такой подход весьма хорошо согласуется с рассмотренными в Главе 1 моделью и в Главе 2 алгоритмом декомпозиции изображения. Для начала рассмотрим другие известные алгоритмы улучшения изображений путем повышения локальных контрастов и сделаем некоторые обобщения.

3.3.1. Обзор используемых методов

Работ, описывающих различные методы и алгоритмы улучшения изображений путем усиления локальных контрастов, опубликовано достаточно большое количество. Наиболее распространенными из них являются алгоритмы, использующие статистические параметры некоторого пространственно ограниченного участка (локального фрагмента) изображения, окружающего обрабатываемый элемент [31,66,113,150,155,179,236,343,346]. Такие преобразования основываются на статистических данных, подсчитанных не по всему изображению, а лишь по локальному фрагменту вокруг обрабатываемого элемента. Из них наиболее известны скользящая эквализация гистограммы [76,277] и усиление локальных контрастов с использованием значений среднего [155] или медианы [346] по скользящему фрагменту. Можно сказать, что данные методы основаны на анализе гистограммы H_{mn} (1.2.1) по фрагменту W_{mn} вокруг обрабатываемого элемента x_{mn} , как основной статистической характеристики. При этом, как правило, используются либо локальное среднее и дисперсия, либо локальная медиана и интерквартильное расстояние [346] по анализируемому фрагменту.

Наиболее часто используются методы, по существу аналогичные известной операции нерезкого маскирования [40], которую в общем виде можно представить как

$$y_{mn} = a(x_{mn} - S_{mn}) + bS_{mn} + c, \quad (3.3.1)$$

где S_{mn} является локальным средним или локальной медианой, a есть коэффициент усиления локальных контрастов, b и c — параметры преобразования (обычно константы).

В широко цитируемой статье [362] Уоллис предложил оператор, обобщающий операции нерезкого маскирования и статистического дифференцирования [122] путем приведения изображения к выбранным значениям среднего S_D и стандартного отклонения σ_D :

$$y_{mn} = (\alpha\sigma_D/(\alpha\sigma_{mn} + \sigma_D))(x_{mn} - S_{mn}) + \beta S_{mn} + (1 - \beta)S_D. \quad (3.3.2)$$

Здесь S_{mn} и σ_{mn} — локальные среднее и среднеквадратическое отклонение (СКО) по фрагменту исходного изображения, S_D и σ_D — локальные среднее и СКО после обработки, α — коэффициент усиления малоразмерных объектов, β — коэффициент усиления фона. Похожий алгоритм фильтрации был предложен в [346], где вместо локального среднего и дисперсии предлагалось использовать медиану и интерквартильное расстояние.

Анализ и классификация алгоритмов улучшения изображений показали, что большинство используемых методов повышения локальных контрастов могут быть обобщены следующей канонической формулой [173,243]:

$$y_{mn} = f(x_{mn} - S_{mn}, v_{mn}) + b_{mn}S_{mn} + c, \quad (3.3.3)$$

Членами данной формулы являются: S_{mn} — оценка средней яркости области, находящейся в центре локальной области анализа; v_{mn} — оценка изменчивости изображения по области анализа (ширина функции распределения значений, дисперсия); $f(u, v)$ — функция контрастирования — монотонная по аргументу u и зависящая от параметра v .

Формула (3.3.3) описывает достаточно широкий класс преобразований, задаваемых как способом оценивания локального среднего S_{mn} , так и выбором функции контрастирования $f(u, v)$. В качестве S_{mn} могут выбираться не только локальное среднее или медиана, но и другие оценки. Сравнивая формулу (3.3.3) с формулой модели фрагмента (1.2.24) очевидно, что значение S_{mn} по своей сути соответствует s_{mn}^r — среднему значению по области, в которую попадает центральный элемент x_{mn} . Такой подход к улучшению изображений можно обобщить следующим образом: сначала исходный сигнал разделяется на сглаженную и разностную компоненты (т.е. осуществляется декомпозиция изображения), после чего выполняется совместное преобразование над полученными компонентами. Так, вместо S_{mn} и $(x_{mn} - S_{mn})$ мы можем подставить в (3.3.3) сглаженную S_{mn} компоненту (2.2.8), получаемую в результате декомпозиции изображения алгоритмом «D», изложенным в разделе 2.2.1, и соответствующую разностную текстурно-детальную компоненту $t_{mn} = x_{mn} - S_{mn}$.

3.3.2. Классификация методов повышения локальных контрастов

Существует несколько разновидностей данных методов; предлагается способ их классификации по двум признакам: линейные/нелинейные и детерминированные/адаптивные методы [174]. Первый признак указывает на линейность усиления контрастов относительно значения фона S_{mn} в (3.3.3), определяемого отдельно для каждого элемента изображения, второй — на возможность вариации функции f по полю изображения в зависимости от значений параметров $t_{mn} = (x_{mn} - S_{mn})$ и v_{mn} , подсчитываемых по анализируемому фрагменту.

Линейные детерминированные (ЛД-методы).

ЛД-методы являются классическими методами повышения локальных контрастов, обобщающими алгоритм нерезкого маскирования [40,155,173,179]. Описываются формулой

$$y_{mn} = at_{mn} + bS_{mn} + c, \quad (3.3.4)$$

где параметры a , b и c постоянны по всему изображению. Нормирующая добавка c в большинстве случаев выбирается как $c = (1 - b)/2$.

В зависимости от выбора значений параметров a и b можно существенно менять свойства преобразования. Так, если S_{mn} является локальным средним, $a = 0$ и $b = 1$, то получим фильтр нижних частот (сглаживание изображения) (2.1.2); если $a = 1$ и $b = 0$, то фильтр верхних частот; при $a > 1$ получим формулу усиления локальных контрастов; при $a = b$ будем иметь обычную формулу линейной градационной коррекции (3.2.2).

Все различие ЛД-методов сводится к возможным способам определения гладкой фоновой компоненты S_{mn} . Из класса линейных методов в основном используется арифметическое среднее (2.1.2) по скользящему фрагменту [155], из нелинейных — медиана (2.1.4) по тому же фрагменту [346], являющаяся частным случаем более общего класса оценок, основанных на порядковых статистиках [54,65,78]. Интересно отметить, что в последнем случае в основе линейного метода усиления локальных контрастов (3.3.4) оказываются нелинейные методы сглаживания изображения.

Линейные адаптивные (ЛА-методы).

Формула обработки изображения ЛА-методами аналогична (3.3.4), но параметр a , определяющий коэффициент усиления локальных контрастов, вычисляется для каждого элемента изображения независимо, на основании оценки параметра v_{mn} :

$$y_{mn} = a_{mn}t_{mn} + bS_{mn} + c, \quad (3.3.5)$$

Наиболее распространены два способа (а и б) определения a_{mn} [346]:

а) если S_{mn} — среднее, то a_{mn} определяется дисперсией яркостей элементов по фрагменту σ_{mn}^2 :

$$a_{mn} = \sigma_{рез} / (\sigma_{mn} + \delta); \quad (3.3.6)$$

здесь σ_{mn}^2 — дисперсия значений элементов (гистограммы H_{mn}) по фрагменту W_{mn} ; $\sigma_{рез}^2$ — дисперсия результата; δ — параметр ограничения максимальной степени усиления контраста;

б) если S_{mn} — медиана, то a_{mn} определяется интерквартильным расстоянием распределения значений элементов (гистограммы H_{mn}) по фрагменту W_{mn} : $\Delta R = R(0,75) - R(0,25)$, где $R(0,25)$ и $R(0,75)$ — значения первого и третьего квартилей распределения (гистограммы) (2.1.3). Тогда

$$a_{mn} = \Delta R_{рез} / (\Delta R_{исх} + \delta), \quad (3.3.7)$$

где $\Delta R_{исх}$ — интерквартильное расстояние распределения по фрагменту исходного изображения, а $\Delta R_{рез}$ — результата.

Способы задания a_{mn} возможны самые различные; на основании экспериментов удачным представляется следующий вариант (в):

в) S_{mn} — медиана, a_{mn} определяется по следующей формуле:

$$a_{mn} = 1 + A / (1 + \sigma_{mn}/Q)^2, \quad (3.3.8)$$

где A — прирост коэффициента усиления контраста, σ_{mn}^2 — дисперсия яркостей элементов по фрагменту, Q — параметр скорости убывания a_{mn} при возрастании σ_{mn} : $a = 1 + A$ при $\sigma = 0$ и убывает до $a = 1$ при увеличении σ .

Нелинейные адаптивные (НА-методы).

Алгоритмы модификации скользящей гистограммы [9,10] не укладываются в приведенную классификацию, поскольку в них не происходит разделения сигнала изображения на две компоненты (3.3.3), но по сути они являются именно нелинейными адаптивными. Основой их служит рассмотренная в разделе 3.2.1 каноническая формула для построения алгоритмов модификации гистограммы (3.2.4). По аналогии с (3.2.4) можно записать следующую формулу модификации скользящей гистограммы:

$$y_{mn} = G^{-1}[R_{mn}(x_{mn})], \quad (3.3.9)$$

здесь $R_{mn}(z)$ — ранг значения z в распределении $h_{mn}(k)$, которая является гистограммой по фрагменту W_{mn} , окружающему x_{mn} ; а $G^{-1}(z)$ — функция, обратная к желаемой функции распределения вероятностей значений элементов по фрагменту. Как и в случае глобальной модификации гистограммы, выбор $G(k)$ может осуществляться на основании требований, предъявляемых к результату обработки.

Нелинейные детерминированные (НД-методы).

Методы данного класса в литературе не встречались и упоминаются здесь для полноты, но они могут быть реально построены, например, по следующей формуле:

$$y_{mn} = f(t_{mn}) + bS_{mn} + c; \quad (3.3.10)$$

функция $f(t_{mn})$ может быть задана в виде произвольной нелинейной функции.

3.3.3. Сравнение методов

Выше были сформулированы лишь известные по литературе методы, а также достаточно очевидные их модификации. Необходимо сказать, что методы приведены в формулировках, отличающихся от оригинальных, но адекватных рассматриваемой модели изображения, что, на наш взгляд, позволяет лучше отразить их сходство.

Сравним рассмотренные методы. Общей чертой ЛД- и ЛА-методов наряду с линейностью усиления компоненты t_{mn} является линейная модификация среднего уровня яркости фона S_{mn} при помощи параметров b и c ; последняя особенность характерна и для НД-методов в приведенной формулировке. Наряду с этим существенное различие детерминированных и адаптивных методов заключается в выборе коэффициента усиления a_{mn} . Если в (3.3.4) он постоянен по изображению, то в (3.3.5) является величиной переменной, причем a_{mn} будет

принимать близкие к максимальным значения внутри протяженных и гладких областей изображения со слабо выраженной текстурой, и минимальные — вблизи контурных перепадов между областями. Совершенно аналогичное неравномерное усиление компоненты t_{mn} , но не столь просто описываемое, происходит и при использовании НА-методов, однако информация о средних яркостях протяженных объектов изображения при этом теряется. Важно отметить, что большинство методов предполагает разделение изображения на две компоненты — детальную компоненту t_{mn} , которая подлежит усилению, и гладкую компоненту S_{mn} , подвергаемую незначительным изменениям. Такой подход весьма близок представлению изображения в модели (1.2.24).

Наибольший интерес, по нашему мнению, представляют ЛД-методы, причем выше уже отмечалось, что основные различия между ними сводятся к способу определения гладкой компоненты S_{mn} . В работе [192] и Главе 2 было показано, что подсчет среднего значения по фрагменту $A(W_{mn})$ или медианы $\text{med}(W_{mn})$ дает неудовлетворительные оценки для яркости фоновой компоненты s_{mn} . По этой причине использование значений среднего или медианы в (3.3.4) приводит к появлению заметных артефактов: подчеркиваются яркостные перепады на границах протяженных областей происходит чрезмерное контрастирование деталей небольших размеров, иногда приводящее к их ограничению из-за лимитированности динамического диапазона яркостей (Рис. 3.3.1,в,г).

3.3.4. Применение декомпозиции для улучшения изображений

Очевидно, искажений удастся избежать, если правильно определить значение гладкой фоновой компоненты согласно модели (1.2.24). Такой метод сглаживания изображения, обеспечивающий сохранение контурных перепадов, предложен в [192]. Усиление локальных контрастов будет достигнуто подстановкой в (3.3.4) значения сглаженной s_{mn} и текстурной $t_{mn} = x_{mn} - s_{mn}$ компонент. В Главе 2 было предложено два варианта алгоритма сглаживания. Первый из них (алгоритм «D», (2.2.8)) позволяет сглаживать изображение при повышении резкости контуров. Однако при использовании в алгоритме усилении локальных контрастов (3.3.12), этот эффект отрицательно сказывается на результатах, вызывая в некоторых случаях заметное ложное оконтуривание, приводящее к разрушению визуально монотонного контурного перепада (хорошо видны на границах ложки на Рис. 3.3.1,д).

Причина оконтуривания заключается в том, что в текстурную компоненту t_{mn} попадает часть контурной информации — разность между первоначальным контурным перепадом и его более резким аналогом. Использование второго варианта сглаживания (алгоритм «D*», (2.2.9)) позволяет достичь успешных результатов. В этом случае сглаженная компонента обеспечивает

гораздо лучшее приближение протяженных контурных перепадов, возникает более точная декомпозиция на гладкую и текстурную компоненты. Результаты обработки представлены на Рис. 3.3.1,б. Визуальный экспертный анализ показал, что таким путем достигается лучшее качество результата по сравнению с остальными рассмотренными методами, причем заметных артефактов не возникает.

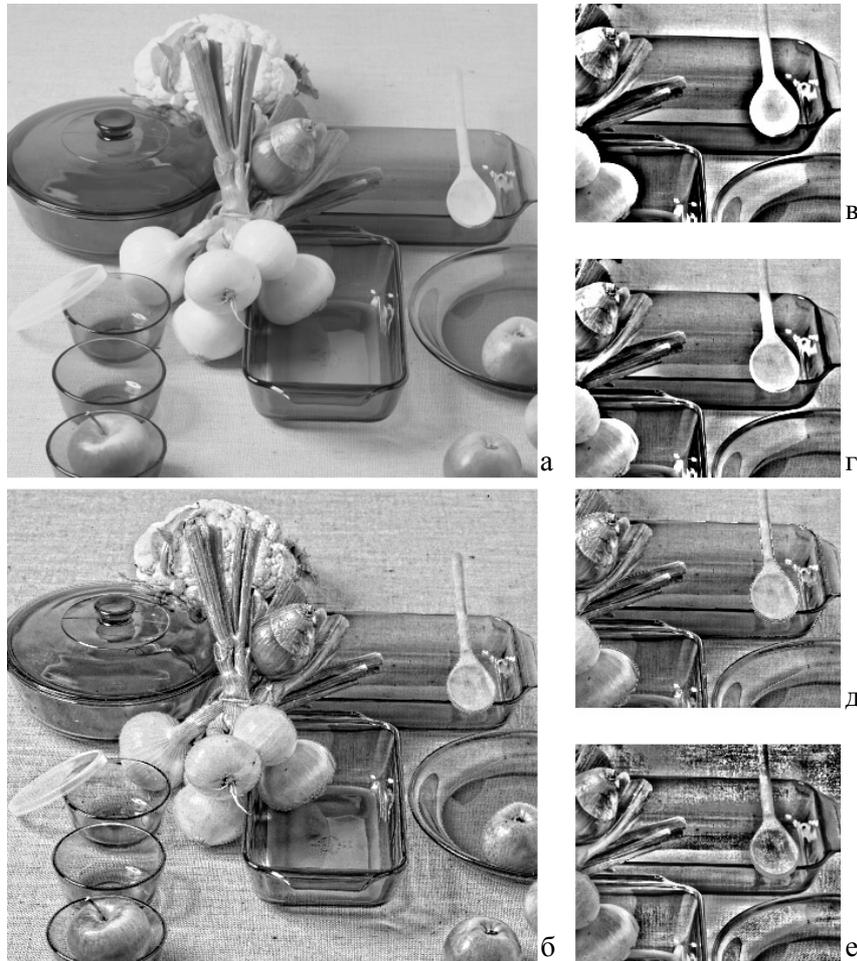


Рис. 3.3.1. Повышение локальных контрастов: а) исходное изображение; б) обработка по формуле (3.3.4), S_{mn} — сглаженная компонента по алгоритму «D*» (2.2.9); в) то же, но S_{mn} — локальное среднее по фрагменту (2.1.2); г) S_{mn} — медиана (2.1.4); д) S_{mn} — сглаживание по алгоритму «D» (2.2.8), видно ложное оконтуривание; е) скользящая эквализация гистограммы (3.3.9).

Изложенные в разделе 3.3 результаты опубликованы в работах [150, 155, 171, 173, 174, 179, 189, 199, 235, 236, 239, 243].

3.4. Улучшение цветных изображений

Алгоритмы улучшения, изложенные выше в применении к черно-белым (монохромным) изображениям, в ряде случаев могут быть применены и к цветным (многоканальным) изображениям. Поскольку настоящая глава посвящена вопросам коррекции и улучшения

изображений, то необходимо подчеркнуть, что градационные преобразования цветного изображения во многих случаях приводят к изменениям цветового тона изображения, что неизбежно сказывается на оценке визуального качества изображения.

Отметим также, что вопрос оценивания качества изображения, а также вопросы, связанные с колориметрией, выбором равноконтрастной цветовой метрики и некоторые другие являются самостоятельными обширными темами, изучаемыми на протяжении многих лет, и выходят за рамки настоящего исследования. Обсуждая возможные применения рассмотренных выше алгоритмов для преобразования цветных изображений, мы будем касаться этих тем в минимально необходимой степени.

Фундаментальным вопросом при работе с цветным изображением является выбор цветового пространства (часто называемого цветовой моделью) [40,74,113,204,333]. Число цветовых пространств, используемых для описания сигнала, полученного в видимом спектре, достаточно велико — в литературе их рассматривается не менее двух десятков. Большинство цветовых моделей ориентированы либо на устройства ввода или воспроизведения, либо на прикладные задачи, например, задачи анализа [40]; есть также модели, используемые в основном для целей колориметрии [74,204].

Наиболее часто используемая модель **RGB** соответствует зрительному восприятию и многим устройствам регистрации видеоинформации, а ее базисные цвета соответствуют реальным цветам. Модель **RGB** и дуальная ей модель **CMY** несут яркостно-цветовую информацию в каждом из трех своих каналов. Во многих других моделях производится разделение на яркостную (как правило обозначаемую символом **L** или **V**, иногда **W**) и цветовые составляющие; среди наиболее используемых следует упомянуть модели **Lab** и **Luv**, несущие информацию о цвете в виде цветоразностей (**ab** или **uv**), а также модели **HSI**, **BHS**, **HSV**, **S₀W**, которые по сути являются тем или иным представлением пространства **RGB** в полярных координатах (**I**, **B**, **V**, **W** — яркость, **H**, **Θ** — цветовой тон, **S** — насыщенность). Используется также модель **XYZ**, как базовая и удобная для трансформации из одного цветового пространства в другое, и система координат, основанная на методе главных компонент (преобразовании Карунена-Лоэва). Последняя чаще всего применяется для сокращения числа первоначальных цветовых составляющих, особенно в случае многоканальной исходной видеоинформации ($J > 3$). Описание вышеупомянутых цветовых пространств, а также ряда других можно найти в монографиях [48,113,204,333]. От одного информационного представления точки в каком-то пространстве нетрудно перейти к представлению в другом пространстве при помощи соответствующих преобразований значений по каналам [204,333].

Из обсуждения ясно, что можно предложить по меньшей мере три различных подхода, позволяющих применить к цветным изображениям изложенные выше алгоритмы улучшения. *Первый* заключается в предварительной трансформации изображения в одно из цветовых пространств, в котором яркостная информационная составляющая отделена от цветовых (**Lab**, **Luv**, **HSI**, **BHS**, **HSV** или **S₀W**), выполнению необходимого преобразования над полученным яркостным каналом, и последующей обратной трансформацией в первоначальное цветовое пространство. Согласно *второму* подходу осуществляется независимая поканальная обработка изображения, что предполагает повторение J раз одного и того же алгоритма в применении к каждому из J цветовых каналов. *Третий* подход предполагает модификацию самого алгоритма в применении его к J -мерному цветовому пространству.

Сравнивая указанные подходы следует отметить, что первый подход, когда преобразованию подвергается только яркостная составляющая, обеспечивает сохранение цветовых характеристик объектов, тогда как второй, в силу независимой поканальной трансформации, вообще говоря не гарантирует сохранения цветового тона объектов изображения. В зависимости от приложения, это может являться как его недостатком, так и преимуществом. Применение в данном подходе алгоритмов выравнивания локальных контрастов и эквализации клиппированной гистограммы обеспечивает, по сути, растяжение диапазона значений по каждому из каналов до максимально возможного.

3.4.1. Градационная коррекция

Алгоритмы выравнивания функции локальных контрастов и эквализации клиппированной гистограммы могут быть использованы для улучшения цветного изображения согласно как первому подходу (преобразованию яркостной составляющей), так и второму подходу (независимой поканальной трансформации). Проведенные эксперименты показали эффективность предложенных алгоритмов для улучшения визуального качества цветных изображений [243].

Пример обработки цветного изображения алгоритмом выравнивания локальных контрастов (3.2.12) представлен на Рис. 3.4.1 (б) и (в). Изображение (б) получено трансформацией из формата **RGB** представления в **BHS**, преобразованием яркостной составляющей и обратной трансформацией в **RGB** формат. Результат выглядит более ярким по сравнению с исходным, на нем стали лучше видны детали, но одновременно стал более заметен неестественный красно-коричневый тон изображения. Изображение (в) получено независимым преобразованием **R**, **G** и **B** каналов. Детали на изображениях (б) и (в) видны одинаково хорошо, но цветовой баланс изображения (в) выглядит значительно более естественно.

Справедливости ради следует отметить, что независимое поканальное преобразование хотя и обеспечивает бóльшую заполненность яркостно-цветового пространства, тем не менее не всегда обеспечивает получение результата с визуально лучшим цветовым балансом, поэтому предполагает последующее визуальное оценивание результата. Эксперименты по применению алгоритма эквализации клиппированной гистограммы (3.2.14) для обработки цветных изображений показали, что в ряде случаев он также может быть использован для улучшения визуального качества в соответствии как с первым, так и со вторым подходами.



Рис. 3.4.1. Выравнивание функции локальных контрастов цветного изображения алгоритмом (3.2.12): а) исходное изображение; б) выравнивание по каналу яркости в **BHS** представлении; в) независимое выравнивание по каждому из цветовых каналов **RGB**.

3.4.2. Методы глобальной цветовой коррекции

К методам глобальной цветовой коррекции относятся те, в которых происходит усиление цветовых различий без учета локальных характеристик изображения. Подобные методы часто используются при анализе аэрокосмических и медицинских снимков.

В [251] описан метод, направленный на усиление цветовой насыщенности без изменения интенсивности и цветности. По существу он является лапласианом в цветовом пространстве. Данный метод изменяет как локальные цветовые различия, так и цветовые оттенки протяженных областей изображения. К его достоинствам следует отнести простоту.

В [215] предложен метод, названный «цветовой эквализацией», и который сводится к следующему. Осуществляется переход из декартовой системы цветовых координат (R, G, B) в цилиндрическую (Y, φ, I) , где Y —яркость, φ — цветность, а I — насыщенность. Диапазон цветности (φ) разбивается на N секторов (в [215] выбрано $N=96$); подсчитываются гистограммы по яркости и каждому сектору цветности, затем находятся и выполняются преобразования, выравнивающие эти гистограммы. В результате цветовая эквализация приводит к получению

изображения с равномерными распределениями как яркости элементов, так и их насыщенности, но при неизменной цветности. Тем самым данный метод преобразования изменяет средние яркости и цветовые оттенки протяженных областей изображения; недостатком его является и то, что он непосредственно не применим для многозональных изображений с большим, чем три, числом цветовых компонент.

Для преобразования многозональных изображений иногда используется метод главных компонент [273], называемый также преобразованием Карунена–Лоэва. Первоначально для J компонент многозонального изображения подсчитывается ковариационная матрица и находятся ее собственные векторы, которые совпадают с осями эллипсоида рассеяния значений исходных многозональных измерений. Выбирается новая система координат, оси которой параллельны направлениям собственных векторов, и осуществляется линейное преобразование (поворот) в новую систему координат [71]. В результате такого преобразования вновь полученные значения оказываются некоррелированными и дисперсии распределений яркостей отдельных зон убывают в соответствии с убыванием собственных значений ковариационной матрицы. Метод главных компонент удобен в случае, если требуется перейти из цветового пространства размерности J в пространство пониженной размерности L ; для этого достаточно выбрать первые L компонент нового разложения. К недостаткам данного метода следует отнести несоответствие новых координатных векторов каким-либо «чистым» цветам цветового пространства, а также весьма существенную неустойчивость координат собственных векторов к относительному изменению дисперсий первоначальных цветовых компонент; заметим также, что реально метод приводит лишь к незначительному усилению цветовых и яркостных различий деталей на изображении.

Искажения средних цветовых оттенков протяженных областей изображения мешают правильному последующему анализу и интерпретации видеoinформации. Естественно ставить задачу усиления локальных цветовых различий мелких деталей по отношению к окружающему фону при сохранении общих средних яркостей и цветовых оттенков протяженных участков изображения.

3.4.3. Повышение локальных контрастов цветных изображений

Метод улучшения изображений повышением локальных контрастов, рассмотренный в разделе 3.3 для монохромных изображений, может быть применен для обработки цветных и многозональных изображений согласно каждому из трех указанных выше подходов. Для определенности возьмем за основу более общий алгоритм (3.3.5) и будем рассматривать его видоизменение при переходе к многомерному цветовому пространству. Согласно модели

фрагмента цветного изображения (1.2.27), рассмотренной в разделе 1.2.4, значение элемента следует рассматривать как вектор \mathbf{x}_{mn} в J -мерном пространстве цветовых координат, например трехмерном (**RGB**):

$$\mathbf{x}_{mn} = \mathbf{S}_{mn} + \mathbf{t}_{mn}, \quad (3.4.1)$$

где \mathbf{S}_{mn} — оценка среднего значения в точке (m,n) , а \mathbf{t}_{mn} — разностный вектор, несущий яркостно-цветовую информацию о текстуре, мелких деталях и шуме. Преобразование (3.3.5), где для простоты положим $b = 1$ и $\mathbf{c} = 0$, представится в векторном виде следующим образом:

$$\mathbf{y}_{mn} = a_{mn}\mathbf{t}_{mn} + \mathbf{S}_{mn}. \quad (3.4.2)$$

В многомерном случае важным является способ нахождения вектора \mathbf{S}_{mn} . В разделе 2.3 было показано, как будет выглядеть в цветовом пространстве гистограмма по фрагменту. Она будет составлять цветовой куб размерами $K \times K \times K$, каждая точка которого $h(r,g,b)$ есть вероятность точки \mathbf{x}_{mn} фрагмента изображения иметь значение $x = \{r,g,b\}$. На Рис. 3.4.2 показано, как будет выглядеть цветовая гистограмма фрагмента, представленного на Рис. 1.1.3,а и захватывающего части трех областей.

Модам одномерной гистограммы черно-белого изображения (пикам распределения $h(k)$ на Рис. 1.1.3,б) здесь соответствуют сгущения ненулевых точек U^1, U^2, U^3 , отвечающие тем частям областей, которые попадают во фрагмент, а центры мод в позициях S^1, S^2 и S^3 представляются векторами центров сгущений: $\mathbf{S}^1, \mathbf{S}^2, \mathbf{S}^3$. Согласно формуле повышения локальных контрастов (3.4.2), результирующий вектор \mathbf{y}_{mn} находится как сумма вектора среднего по сгущению \mathbf{S}^1 , которому соответствует исходный вектор \mathbf{x}_{mn} , и произведения вектора разности $(\mathbf{x}_{mn} - \mathbf{S}^1)$ на скалярный коэффициент усиления a_{mn} (см. Рис. 3.4.2).

Согласно первому подходу, каналы (rgb) вектора \mathbf{x}_{mn} должны быть преобразованы в один яркостной канал $x_{mn} = \|\mathbf{x}_{mn}\|$ и два цветowych, которые остаются неизменными. Многомерная гистограмма $h(r,g,b)$, отвечающая яркостной составляющей, будет преобразована в одномерное распределение $h(k)$, что соответствует проецированию всего объема куба гистограммы $H(R,G,B)$ на центральный вектор \mathbf{L} с направлением $(1, 1, 1)$ (Рис. 3.4.2). Очевидна проблема, возникающая при таком сжатии пространства: проекции разных сгущений могут отображаться в накладывающиеся моды, которые невозможно будет различить. По полученной одномерной гистограмме $h(k)$ алгоритмом декомпозиции «D*» (2.2.9), приведенном в разделе 2.2, находится значение S_{mn} , после чего формула повышения локальных контрастов по (3.3.5) может быть выражена следующим образом:

$$\mathbf{y}_{mn} = (a_{mn}(1 - S_{mn}/x_{mn}) + 1)\mathbf{x}_{mn}. \quad (3.4.3)$$

Такое преобразование соответствует коллинеарному изменению каждого из векторов значений \mathbf{x}_{mn} исходного изображения, т.е. вектор \mathbf{y}_{mn} будет проходить через ту же точку цветового

треугольника Максвелла, что и вектор \mathbf{x}_{mn} . Особенность данного подхода — цветовые характеристики точек изображения не изменяются, а сферы сгущений, показанные на Рис. 3.4.2, вытягиваются в овоиды вдоль направлений векторов \mathbf{S}^j каждого из сгущений U^j .

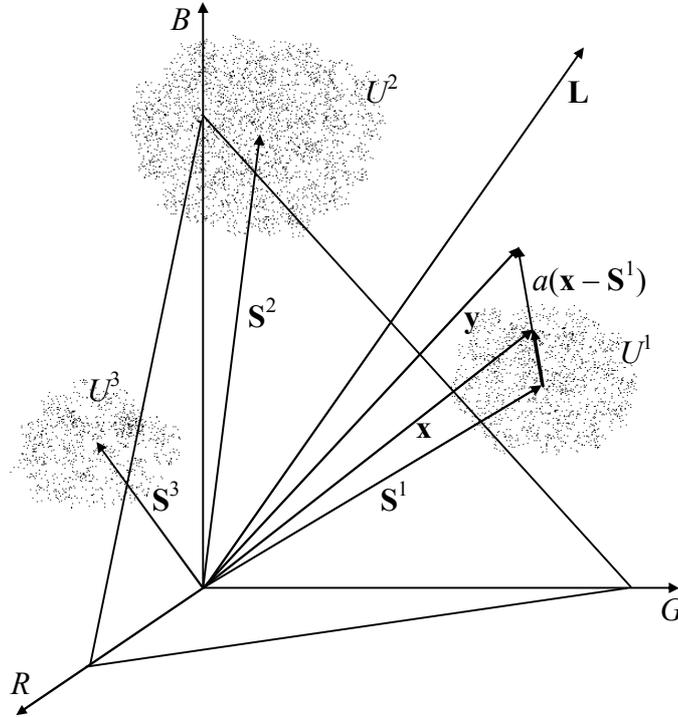


Рис. 3.4.2. Трехмерная цветовая гистограмма $H(R, G, B)$ и формирование вектора результата \mathbf{y} .

Второй подход — независимое J -кратное поканальное преобразование изображения. Для этого осуществляется проецирование многомерной гистограммы $h(k_1, \dots, k_J)$ на каждую из J координатных осей, после чего применяется алгоритм декомпозиции «D*» (2.2.9) для анализа каждой из проекций. Из полученных частных значений S^j формируется вектор \mathbf{S}_{mn} , который и подставляется в (3.4.2). С точностью до возможного различия значений a_{mn} для разных каналов, данная процедура эквивалентна применению формулы (3.3.5) для каждого из каналов изображения по отдельности.

Третий подход заключается в модификации алгоритма к J -мерному цветовому пространству. Осуществить это можно применяя для определения значения \mathbf{S}_{mn} способа цветового сглаживания, описанного в разделе 2.3, основанного на использовании расстояния в цветовом пространстве. Графическая интерпретация преобразования повышения локальных контрастов цветного изображения показана на Рис. 3.4.2 векторами \mathbf{x} , \mathbf{S}^1 и \mathbf{y} .

Описанный способ улучшения цветных изображений приводит к усилению яркостных и цветовых контрастов за счет изменения соотношений локальных яркостей и локальных цветовых тонов мелких деталей с окружающим фоном. При этом сохраняются общие средние яркости и средние цветовые тона протяженных областей изображения. Как уже говорилось,

задача улучшения изображения является проблемно ориентированной, поэтому к оценкам результатов следует относиться с осторожностью. Тем не менее эксперименты показывают, что изображение, преобразованное описанным способом, во многих случаях позиционируется выше оригинала.

Пример улучшения цветного изображения представлен на Рис. 3.4.3. На рисунке (а) показан фрагмент исходного изображения, на рисунке (б) — сглаженная компонента S_{mn} , полученная поканальной декомпозицией при помощи алгоритма «D*» (2.2.9), а на рисунке (в) — результат повышения локальных контрастов по формуле (3.4.2).

Применение данного метода улучшения цветных изображений, подобно методу усиления локальных контрастов черно-белого изображения (3.3.4)–(3.3.5), значительно повышает возможности последующего визуального анализа или интерпретации полученной видеоинформации. В результате такой обработки проявляются локальные яркостные и цветовые различия, практически не заметные на исходном изображении. Как показали эксперименты, усиление локальных контрастов по формулам (3.3.4), (3.3.5) и (3.4.2) может быть успешно применено при анализе аэрокосмической видеоинформации с целью изучения природных ресурсов или исследования планет, в полиграфической промышленности — для улучшения черно-белых или цветных оригиналов, в дефектоскопии и рентгенодиагностике — для выявления малоконтрастных объектов, а также во многих других случаях.

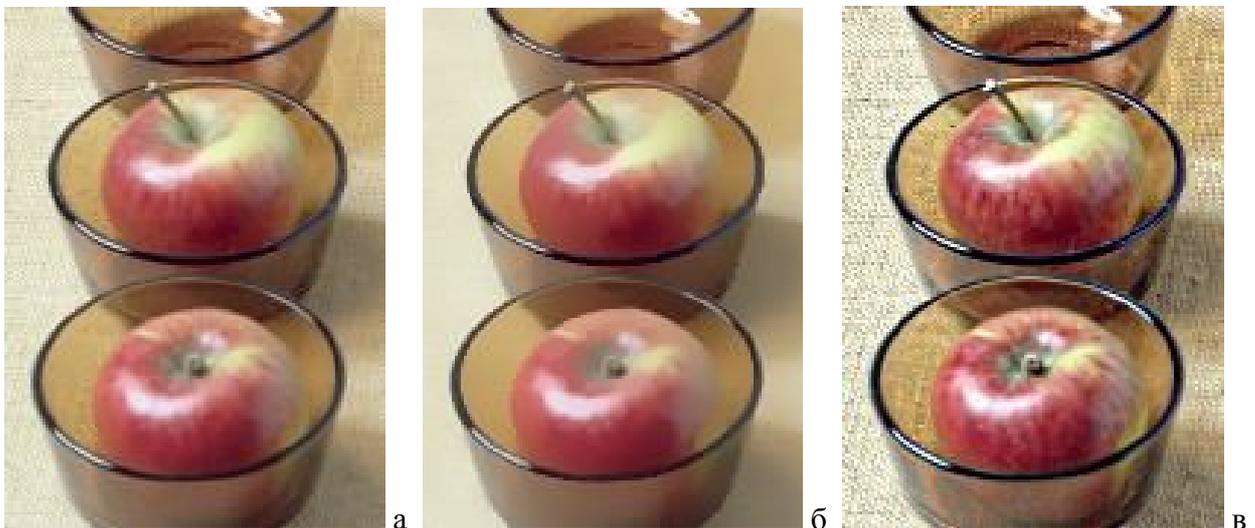


Рис. 3.4.3. Декомпозиция и улучшение цветного изображения: а) исходное изображение; б) сглаженная компонента после декомпозиции; в) улучшение, основанное на декомпозиции.

Изложенные в разделе 3.4 результаты опубликованы в работах [173,179,199,235,236,243].

3.5. Выводы и результаты

Проведены исследования методов фильтрации, коррекции и улучшения изображений, разработанных на основе двухмасштабной многокомпонентной модели. Предложены новые алгоритмы фильтрации, а также способы автоматического определения необходимых параметров для существующих методов. Показано, что постулатов локальной модели, предполагающей использование лишь ограниченной области анализа изображения, в ряде случаев достаточно для построения не только локальных, но и глобальных методов его обработки.

Разработан ранговый алгоритм фильтрации импульсных помех с динамическим определением пороговых параметров для каждого элемента изображения, вычисляемых на основе анализа элементов окрестности. Алгоритм показал в 1,5÷2 раза более высокую точность восстановления изображения по сравнению с другими известными методами. Предложен вариант алгоритма фильтрации, основанный на использовании декомпозиции изображения. Он сводится к разложению изображения на кусочно-гладкую и текстурно-детальную компоненты, фильтрацию только второй и обратное сложение компонент.

Для фильтрации периодических помех в частотной области предложен алгоритм автоматического формирования маски фильтра, базирующийся на многокомпонентной модели. При этом энергетический спектр интерпретируется как изображение, а всплески отыскиваются применением разработанных алгоритмов обнаружения объектов по их площади.

Для устранения одномерных аддитивных периодических помех, возникающих вдоль линий развертки изображения в радиотехнических системах, предложен алгоритм анализа изображения и устранения помех, основанный на корреляционном совмещении, усреднении и пространственной фильтрации сигнала.

Исследована задача автоматической яркостной (градационной) коррекции изображения. Для оценивания градационных искажений предложена функция локальных контрастов, позволяющая оценивать яркостные искажения. Выдвинута гипотеза о ее константности, подтвержденная экспериментально. Разработан алгоритм автоматической градационной коррекции изображений, обеспечивающий выравнивание функции локальных контрастов. Для случая большой разности яркостей между объектами (т.н. «разрыва шкалы яркостей») предложен алгоритм эквализации клиппированной гистограммы. Проведены исследования алгоритмов градационной коррекции и проиллюстрирована эффективность предложенных методов.

Исследована задача улучшения изображения усилением локальных контрастов. На основе многокомпонентной модели изображения сформулирована обобщенная формула для большинства используемых методов. Проведена классификация методов повышения локальных контрастов и выполнено их сравнение. Согласно обобщенной формуле сигнал изображения

разделяется на сглаженную и разностную составляющие. Предложена модификация алгоритмов повышения локальных контрастов путем использования компонент, формируемых при помощи модифицированного алгоритма декомпозиции изображения.

Показаны возможности и предложены различные варианты применения методов градационной коррекции и повышения локальных контрастов для улучшения цветных и многоканальных изображений.

Полученные результаты иллюстрируют важный вывод Главы 2, заключающийся в том, что эффективным подходом к обработке и анализу изображения является предварительная его декомпозиция на различные информационные компоненты и последующее применение специфических преобразований к каждой из полученных компонент.

ГЛАВА 4. РАЗРАБОТКА МЕТОДОВ АНАЛИЗА ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДВУХМАСШТАБНОЙ МОДЕЛИ

Введение

Рассматривается применение двухмасштабной многокомпонентной модели изображения и созданного на ее основе алгоритма декомпозиции для анализа видеоинформации и обнаружения объектов на изображении. Сглаженная компонента S_{mn} (1.2.24), формируемая алгоритмом декомпозиции «D» (2.2.7)–(2.2.8), используется для построения алгоритмов обнаружения объектов по их площади и для исследования метода нахождения различий в наборе объектов на паре изображений. На основе использования сглаженной компоненты предлагается новый подход к решению одной из классических задач предварительной обработки — обнаружению контуров на изображении. В последней части главы рассматривается использование модели для решения важной научно-технической задачи — обнаружения дефектов на поверхности электронных микросхем по снимкам, получаемых микроскопом высокого разрешения.

4.1. Обнаружение объектов заданной площади

Можно сформулировать три варианта задачи обнаружения объектов на изображении по их площади: обнаружение объектов площади больше заданной, меньше заданной и обнаружение объектов, имеющих площадь в заданном интервале. Другими словами, задается некоторое пороговое значение Q (или пара значений: $Q_1 < Q_2$) и требуется найти все области изображения, число элементов N^j внутри которых либо больше, либо меньше Q , либо находится в интервале $Q_1 < N^j < Q_2$. Следует отметить, что понятие площади здесь используется несколько в необычном смысле — как «локальная» площадь, т.е. площадь той части объекта, которая попадает внутрь области анализа.

4.1.1. Обнаружение объектов с площадью больше заданной

В данной постановке предполагается, что изображение представляет собой достаточно ровный фон (одна большая область U^0), на котором имеется определенный набор сравнительно небольших областей U^1, \dots, U^j . Требуется найти области с $N^j > Q$. Воспользуемся алгоритмом декомпозиции «D» (2.2.8). Предположим, что области U^1, \dots, U^j отстоят друг от друга достаточно далеко и можно выбрать некоторый размер фрагмента L ($L^2/2 > Q$) такой, что в любой фрагмент

W_{mn} попадает не более одной области с $N^j > Q$, либо несколько меньших, но при условии $\Sigma N^j < Q$ ($U^j \in W$). Тогда в п. 5 алгоритма декомпозиции можно выбрать $n^W = Q$, а R_1^W и R_2^W как $R(Q/L^2)$ и $R(1 - Q/L^2)$ соответственно. Обработкой исходного изображения по алгоритму (2.2.8) с указанными значениями R_1^W и R_2^W получим [192]:

$$y_{mn} = S_{mn}, \quad (4.1.1)$$

т.е. сглаженную компоненту исходного изображения, на которой остались лишь области с $N^j > Q$. Эти области уже легко обнаруживаются детектором со значением порога $S(U^0) \pm \delta$, где $S(U^0)$ — яркость фона, а $\delta < \min_j |S(U^j) - S(U^0)|$; ($S(U^j)$ — яркости соответствующих областей):

$$z_{mn} = \begin{cases} y_{mn}, & \text{если } |y_{mn} - S(U^0)| \geq \delta; \\ S(U^0), & \text{если } |y_{mn} - S(U^0)| < \delta. \end{cases}$$

Данная задача имеет важное приложение в неразрушающем контроле (дефектоскопии) при обнаружении дефектов в исследуемых объектах. На Рис. 4.1.1,а представлена рентгенограмма детали с внутренними кавернами. Согласно правилам контроля, дефектами считаются лишь те каверны, размеры которых превышают некоторый порог. Обработка изображения описанным алгоритмом с указанием параметра n^W равным площади максимального из допустимых дефектов позволила отфильтровать все мешающие мелкие пятна (изображения мелких каверн). Результат такой обработки (4.1.1) показан на Рис. 4.1.1,б; на котором, однако, осталась большая область, вызванная особенностями формы детали.

4.1.2. Обнаружение объектов с площадью меньше заданной

Требуется найти объекты с площадью $N^j < Q$. Сглаженная компонента S_{mn} , в (4.1.1) содержит лишь области с $N^j > Q$, а значит области с $N^j < Q$ будут содержаться в текстурной компоненте $t_{mn} = x_{mn} - S_{mn}$. Для ликвидации возможных шумовых выбросов полезно осуществить фильтрацию x_{mn} , которая в нашем случае фактически происходит на шаге 3 алгоритма декомпозиции (2.2.7): $\bar{x}_{mn} \approx x_{mn} - \xi_{mn}$. Собственно выделение объектов с $N^j < Q$ достигается пороговым их обнаружением в тех точках, где $|t_{mn}| \geq \delta$ (δ — порог обнаружения):

$$z_{mn} = \begin{cases} \bar{x}_{mn}, & \text{если } |\bar{x}_{mn} - S_{mn}| \geq \delta; \\ C, & \text{если } |\bar{x}_{mn} - S_{mn}| < \delta. \end{cases} \quad (4.1.2)$$

4.1.3. Обнаружение объектов с площадью в интервале $[Q_1, Q_2]$

Нахождение объектов с площадью в интервале $[Q_1, Q_2]$. Возможны два варианта решения.

В первом случае сначала выберем $n^W = Q_1$. Тогда сглаженная компонента S_{mn} в (4.1.1) будет содержать объекты с $N^j > Q_1$ и не будет содержать объекты с $N^j < Q_1$. Осуществим повторную ее обработку алгоритмом (4.1.2), но с $n^W = Q_2$ ($Q_2 > Q_1$). Очевидно, y_{mn} будет

содержать лишь искомые объекты, попадающие в диапазон $Q_1 < N^j < Q_2$. Недостаток данного решения — алгоритм получается двухпроходовым.

Второй вариант. Обратим внимание, что при анализе гистограмм по окрестности и фрагменту используются два разных порога (n^V и n^W). Выберем размеры окрестности l и фрагмента L больше обычного — такими, чтобы $l^2 > 2Q_1$ и $L^2 > 2Q_2$. Тогда, задав R_1^V и R_2^V как $R_1^V = R^V(Q_1/l^2)$ и $R_2^V = R^V(1 - Q_1/l^2)$, после операции (2.2.7) будем иметь \bar{x}_{mn} , которое уже не содержит области с $N^j < Q_1$. Далее в п. 5 алгоритма декомпозиции, при анализе H_{mn}^W , зададим R_1^W и R_2^W как $R_1^W = R^W(Q_2/L^2)$ и $R_2^W = R^W(1 - Q_2/L^2)$. Получив значение S_{mn} в (2.2.8), воспользуемся решающим правилом, аналогичным (4.1.2):

$$z_{mn} = \begin{cases} \bar{x}_{mn}, & \text{если } |\bar{x}_{mn} - S_{mn}| \geq \delta; \\ C, & \text{если } |\bar{x}_{mn} - S_{mn}| < \delta. \end{cases} \quad (4.1.3)$$

Как видно, (4.1.3) почти полностью совпадает с (4.1.2). Действительно, они отличаются лишь тем, что в (4.1.3) выбирается значение \bar{x}_{mn} , получаемое после п. 3 алгоритма декомпозиции. Результат показан на Рис. 4.1.1, в. Устранены как верхнее пятно, так и большая область.

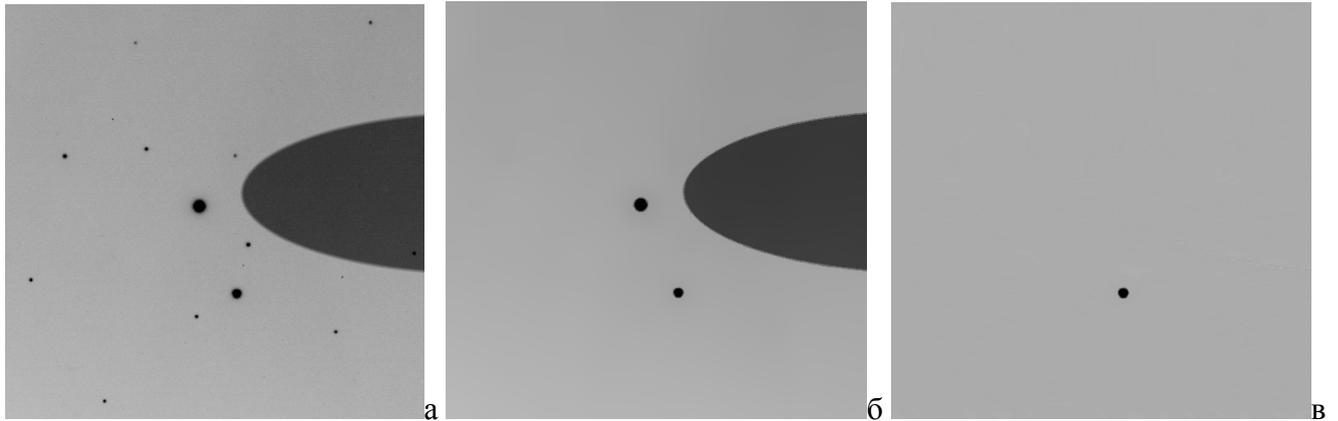


Рис. 4.1.1. Обнаружение каверн детали: а) исходное изображение; б) обнаружение объектов с площадью больше Q ; в) обнаружение объектов с площадью в интервале от Q_1 до Q_2 .

Другой пример обнаружения областей на изображении данным алгоритмом показан на Рис. 4.1.2. Здесь аэрофотоснимок (Рис. 4.1.2,а) обработан по формуле (4.1.3) с различными парами значений Q_1 и Q_2 . В первом случае диапазон выбран в области малых значений и обнаружены объекты малой локальной площади (Рис. 4.1.2,б); во втором случае параметры увеличены и в результате обнаружены объекты большой площади (Рис. 4.1.2,в). Поскольку, как говорилось выше, используется понятие «локальной» площади — т.е. площади части объекта, попадающей внутрь области анализа, на Рис. 3,б обнаружены протяженные тонкие линии, общее число элементов в которых велико, но во фрагмент одновременно их попадает немного (приблизительно $d \times L$, где d — толщина линии).

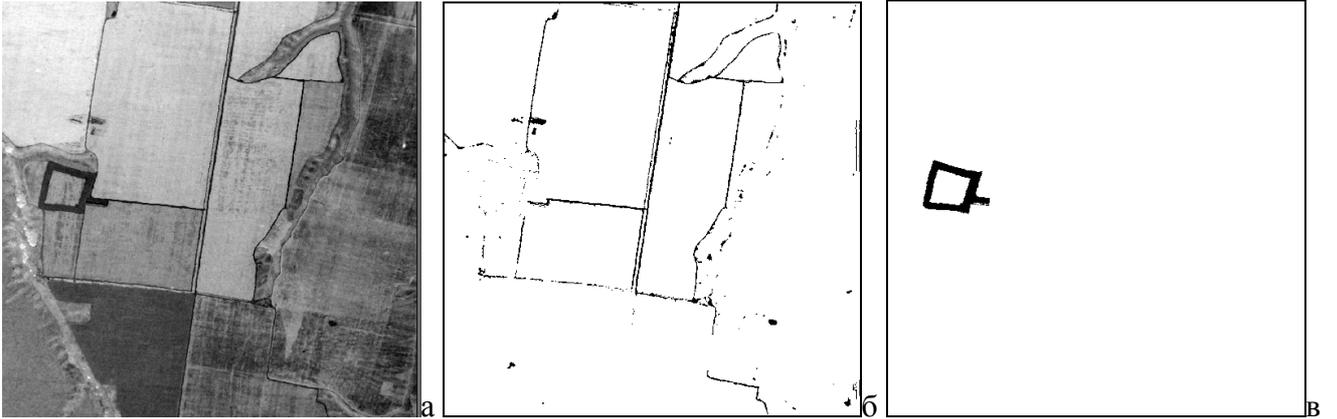


Рис. 4.1.2. Обнаружение объектов: а) исходный аэрофотоснимок; б) объекты малой локальной площади по формуле (4.1.2); в) объекты большой локальной площади по формуле (4.1.3).

4.2. Обнаружение различных объектов на изображениях

Цель состоит в обнаружении различий двух изображений \mathbf{X}^1 и \mathbf{X}^2 , отображающих сцену, на которой произошли некоторые изменения. Сформулируем задачу следующим образом. Пусть имеется некоторая исходная сцена \mathbf{V}_0 , состоящая из объектов v_1, \dots, v_M . Рассматриваемые изображения \mathbf{X}^1 и \mathbf{X}^2 отображают в действительности сцены \mathbf{V}^1 и \mathbf{V}^2 , отличающиеся от \mathbf{V}_0 добавлением в одном случае объектов v_{11}, \dots, v_{1K} , а в другом — v_{21}, \dots, v_{2L} . Отметим, что отсутствие объекта на одном изображении можно трактовать как добавление его на другом. Подобным же образом можно трактовать и изменение формы объектов.

На тех участках изображения, где не произошло появления новых объектов, \mathbf{X}^1 и \mathbf{X}^2 являются двумя реализациями одной и той же сцены \mathbf{V}_0 ; обозначим их через \mathbf{X}_0^1 и \mathbf{X}_0^2 . Тем самым \mathbf{X}^1 и \mathbf{X}^2 отличаются от \mathbf{X}_0^1 и \mathbf{X}_0^2 добавлением областей u_{11}, \dots, u_{1K} и u_{21}, \dots, u_{2L} , соответствующих v_{11}, \dots, v_{1K} и v_{21}, \dots, v_{2L} .

Введем некоторые ограничения. Будем считать, что на изображениях \mathbf{X}^1 и \mathbf{X}^2 требуется обнаружить области, площади которых лежат в пределах $N_1 \leq N(u_{ij}) \leq N_2$. Представим изображение \mathbf{X}^1 как сумму изображения \mathbf{X}_0^1 , областей u_{11}, \dots, u_{1p} , площади которых находятся в пределах $N_1 \leq N(u_{ij}) \leq N_2$, и областей w_{11}, \dots, w_{1p} , имеющих площади вне указанных пределов: $N(w_{ij}) < N_1$ или $N(w_{ij}) > N_2$ ($P + p = K$):

$$\mathbf{X}^1 = \mathbf{X}_0^1 \oplus (w_{11}, \dots, w_{1p}) \oplus (u_{11}, \dots, u_{1p}), \quad (4.2.1)$$

где операция $a \oplus b$ означает, что в тех точках, где определены элементы как множества a , так и множества b , берутся элементы из b . Аналогично для изображения \mathbf{X}^2 ($Q + q = L$):

$$\mathbf{X}^2 = \mathbf{X}_0^2 \oplus (w_{21}, \dots, w_{2q}) \oplus (u_{21}, \dots, u_{2q}). \quad (4.2.2)$$

Можно считать, что имеются промежуточные изображения \mathbf{X}_w^1 и \mathbf{X}_w^2 :

$$\mathbf{X}_w^1 = \mathbf{X}_0^1 \oplus (w_{11}, \dots, w_{1p}) \quad \text{и} \quad \mathbf{X}_w^2 = \mathbf{X}_0^2 \oplus (w_{21}, \dots, w_{2q}),$$

отличающиеся наличием областей w_{11}, \dots, w_{1p} и w_{21}, \dots, w_{2q} , но не различимые в нашей постановке задачи. Отсюда:

$$\mathbf{X}^1 = \mathbf{X}_w^1 \oplus (u_{11}, \dots, u_{1p}) \text{ и } \mathbf{X}^2 = \mathbf{X}_w^2 \oplus (u_{21}, \dots, u_{2q}).$$

Выделить на изображении области, имеющие площадь в диапазоне $[N_1, N_2]$, можно при помощи алгоритма (4.1.3), рассмотренного в разделе 4.1. Преобразовав таким способом \mathbf{X}^1 и \mathbf{X}^2 , получим сглаженные изображения $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$. На них будут выделены как искомые области u_{11}, \dots, u_{1p} , и u_{21}, \dots, u_{2q} , так и области $u_{11}^0, \dots, u_{1r}^0$ и $u_{21}^0, \dots, u_{2r}^0$ компонент \mathbf{X}_w^1 и \mathbf{X}_w^2 , отображающие повторяющиеся объекты исходной сцены с площадью в том же диапазоне; в данном случае они являются мешающими. Из (4.1.3) следует, что $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$ должны представлять собой ровный фон яркости C , на котором будут содержаться сглаженные области \bar{u}_{ij}^0 и \bar{u}_{ij} :

$$\bar{\mathbf{X}}^1 = C \oplus (\bar{u}_{11}^0, \dots, \bar{u}_{1r}^0, \bar{u}_{11}, \dots, \bar{u}_{1p}) \text{ и } \bar{\mathbf{X}}^2 = C \oplus (\bar{u}_{21}^0, \dots, \bar{u}_{2r}^0, \bar{u}_{21}, \dots, \bar{u}_{2q}). \quad (4.2.3)$$

Согласно модели (1.2.23):

$$x_{mn}^1 = S_{mn}^1 + t_{mn}^1 + \xi_{mn}^1 \text{ и } x_{mn}^2 = S_{mn}^2 + t_{mn}^2 + \xi_{mn}^2.$$

Для \bar{x}_{mn} в (4.1.3) можно показать, что

$$\bar{x}_{mn} = S(u_{mn}^i) + o(D), \quad (4.2.4)$$

где $S(u_{mn}^i)$ — средняя яркость той части области u^i во фрагменте анализа, в которую попадает элемент x_{mn} , а D — общий диапазон яркостей.

Пусть изображения \mathbf{X}^1 и \mathbf{X}^2 согласованы по градиационным характеристикам, т.е. для областей u_{1i}^0 и u_{2i}^0 , отображающих один и тот же реальный объект v_i на разных изображениях, выполняется:

$$S(u_{1i}^0) = S(u_{2i}^0) + o(D).$$

Тогда разность изображений $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$, $z_{mn} = \bar{x}_{mn}^1 - \bar{x}_{mn}^2$, согласно (4.1.3) будет выглядеть:

$$z_{mn} = \begin{cases} o(D), & \text{если объекты совпадают,} \\ (\bar{x}_{mn}^1 - \bar{x}_{mn}^2), & \text{если объекты не совпадают.} \end{cases} \quad (4.2.5)$$

Таким образом, $z_{mn} = o(D)$ в точках, где \mathbf{X}^1 и \mathbf{X}^2 отображают одни и те же объекты исходной сцены, и, согласно (4.2.4), $z_{mn} \approx S(u_{mn}^1) - S(u_{mn}^2)$, если x_{mn}^1 и x_{mn}^2 соответствуют разным объектам.

Для наиболее частого случая, когда новый объект возникает лишь на одном из изображений:

$$z_{mn} = \begin{cases} o(D), & \text{если объекты совпадают,} \\ (\bar{x}_{mn}^1 - C), & \text{если объект возник на изображении 1,} \\ (C - \bar{x}_{mn}^2), & \text{если объект возник на изображении 2.} \end{cases} \quad (4.2.6)$$

Для удобства обнаружения объектов положим $C = 0$. Тогда $z_{mn} \leq o(D)$ в точках, где объекты интересующих размеров отсутствуют либо совпадают; $z_{mn} = \bar{x}_{mn}^1$, если объект возник на

изображении 1; и $z_{mn} = -\bar{x}_{mn}^2$, если объект возник на изображении 2. Выбрав значение порога $\delta \approx \sigma_t + \sigma_\xi$, где σ_t и σ_ξ — стандартные отклонения текстурной t_{mn} и шумовой ξ_{mn} компонент в (1.2.23), очевидным становится решающее правило об отсутствии различий или наличии объекта на одном или другом изображении в зависимости от факта попадания z_{mn} в интервал $[-\delta, \delta]$, либо знака z_{mn} при выходе за рамки данного интервала.

Тем самым мы получили способ отыскания различий на изображениях. Однако он может давать некоторые погрешности, вызванные следующим. Пусть u_{1i}^0 и u_{2i}^0 — две области изображений \mathbf{X}^1 и \mathbf{X}^2 , отображающие один и тот же объект. В силу неточностей дискретизации и совмещения изображений возможны несовпадения границ данных областей. Из-за этого на разностном изображении \mathbf{Z} могут возникнуть отдельные точки или контурные линии на границах областей. Избавиться от них нетрудно, обнаружив предварительно контурные участки на $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$, и воспользовавшись следующим алгоритмом:

1. Определяется значение разности $d_{mn} = |\bar{x}_{mn}^1 - \bar{x}_{mn}^2|$.
2. Если $d_{mn} < \delta$, то x_{mn}^1 и x_{mn}^2 — элементы одной и той же области и в этом случае $z_{mn} = C$.
3. Если $d_{mn} \geq \delta$, то: а) если (m,n) — контурная точка как на \mathbf{X}^1 , так и на \mathbf{X}^2 , то $z_{mn} = C$; б) если (m,n) — контурная точка на \mathbf{X}^1 и фоновая на \mathbf{X}^2 , то $z_{mn} = \bar{x}_{mn}^1 - C$, если же наоборот, то $z_{mn} = C - \bar{x}_{mn}^2$; в) если (m,n) — фоновая точка и на \mathbf{X}^1 и на \mathbf{X}^2 , то z_{mn} равно тому значению \bar{x}_{mn}^1 или \bar{x}_{mn}^2 , которое не попадает в диапазон $C \pm \delta$; если же оба выходят за этот диапазон, то значит соответствующие области на изображениях (объекты сцены) перекрываются. В этом случае можно в качестве значения выбрать либо \bar{x}_{mn}^1 , либо \bar{x}_{mn}^2 в зависимости от априорного приоритета изображений, либо взять их разность. В некоторых случаях может потребоваться вместо сглаженных \bar{x}_{mn}^1 или \bar{x}_{mn}^2 в соответствующих точках брать значения элементов исходных изображений x_{mn}^1 или x_{mn}^2 .

Изложенные в разделах 4.1–4.2 результаты опубликованы в работах [167,173,176,179,236].

4.3. Применение декомпозиции для нахождения границ объектов

Традиционной задачей предварительного анализа изображений является нахождение границ объектов, которая обычно решается через обнаружение контурных перепадов. Вопросы выделения контуров на изображении многократно рассматривались в литературе (см., например, [40,113,206,270,308]). Классический подход состоит в формировании градиентного отклика тем или иным контурным оператором, его квантовании на два уровня по выбранному

порогу бинаризации δ , и последующим упрощении путем удаления ложных линий и слияния отдельных разорванных контурных сегментов в более крупные и связные. Альтернативными подходами являются алгоритмы так называемого «зеро-кроссинга» [90,246] и прослеживания контуров [350], в том числе метод Canny [223].

Существенным недостатком большинства подходов является высокая чувствительность к имеющимся на изображении шуму и текстуре, отсюда повышенная вероятность ошибок как первого, так и второго рода, вызывающих появление ложных точек, идентифицируемых как контурные, а также разрывам протяженных контурных линий.

Опираясь на многокомпонентную модель изображения предлагается новый подход к решению задачи обнаружения контуров. Как уже говорилось, изображение считается состоящим из набора отдельных областей U^1, \dots, U^N , соответствующих объектам сцены, границы между которыми и являются контурами. Информацию о данных границах несет составляющая S_{mn} в (1.2.23), которая считается гладкой внутри отдельных областей изображения и разрывной на контурах. Текстурная t_{mn} и шумовая ξ_{mn} компоненты в данном случае являются мешающими, поскольку в них подобной информации не содержится, а значит они могут лишь вызывать появление дополнительных искажений при обнаружении контуров.

Из сказанного с очевидностью вытекает, что из изображения следует предварительно выделить сглаженную компоненту S_{mn} и затем уже на ней обнаруживать границы областей при помощи того или иного контурного оператора. Способ выделения компоненты S_{mn} был изложен выше, в разделе 2.2. Алгоритм декомпозиции (2.2.8) отделяет текстурную и шумовую составляющие, приводит к значительному сглаживанию изображения и даже к некоторому повышению резкости контурных перепадов [192]. На таком изображении проведение границ между областями достигается практически любым контурным оператором. Меняя в алгоритме декомпозиции (2.2.8) значение параметра n^W , можно задавать минимальный размер областей, которые выделятся на изображении как отдельные и будут оконтурены. Заметим, что такой подход близок рассмотренному выше обнаружению объектов.

В качестве примера на Рис. 4.3.1 демонстрируются результаты обнаружения контуров на изображении, представленном на Рис. 4.1.2,а. Для формирования контурного отклика использовался простейший дифференциальный оператор Робертса [121]:

$$y_{mn} = \begin{cases} 1, & \text{если } z_{mn} = ((x_{m,n} - x_{m+1,n+1})^2 + (x_{m,n+1} - x_{m+1,n})^2)^{1/2} \geq \delta, \\ 0, & \text{если } z_{mn} < \delta, \end{cases} \quad (4.3.1)$$

с одним и тем же значением порога бинаризации δ . На Рис. 4.3.1,а показано выделение контуров по исходному, а на Рис. 4.3.1,б — по сглаженному алгоритмом декомпозиции (2.2.8) изображению. На (а) видна разрывность некоторых протяженных контурных границ, наличие

многочисленных ложных контурных линий и отдельных точек, вызванных шумом и текстурой. Изображение (б) отличается большей слитностью контурных линий, меньшей их толщиной и значительно меньшим количеством помех.



Рис. 4.3.1. Обнаружение контуров: а) по исходному изображению (Рис. 4.1.2,а);
б) по сглаженной компоненте S_{mn} .

Еще один пример сглаживания и обнаружения контуров показан на Рис. 4.3.2. На Рис. 4.3.2,а представлено обнаружение контуров по исходному изображению (Рис. 2.2.3,а), а на Рис. 4.3.2,б — по сглаженной компоненте S_{mn} (Рис. 2.2.3,б). Контурные выделялись тем же оператором Робертса (4.3.1) с одинаковыми значениями порога бинаризации δ .



Рис. 4.3.2. Обнаружение контуров: а) по исходному изображению (Рис. 2.2.3,а);
б) по сглаженной компоненте S_{mn} (Рис. 2.2.3,б).

4.4. Быстрое корреляционное совмещение изображений

Задача совмещения пары изображений, различающихся пространственным сдвигом без деформаций [110,318], является достаточно частой. Она возникает, например, при контроле изделий в том или ином технологическом процессе, когда в качестве одного из этапов требуется совмещение анализируемого изображения с эталоном. В ряде подобных случаев изображения имеют выраженные структурные особенности, например: значительная часть контуров объектов имеет определенную направленность, или же объекты расположены периодически вдоль каких-то направления. Совокупность подобных особенностей будем называть *квазирегулярной структурой объектов* анализируемых изображений, или просто *квазирегулярностью*. Использование квазирегулярности позволяет разработать быстрый алгоритм корреляционного совмещения изображений [165,234].

Пример пары таких изображений — эталонного и изображения с дефектом — показан ниже. В данном случае контуры объектов имеют горизонтальную и вертикальную направленность, но задача может быть обобщена на случай любых, в том числе и не ортогональных направлений.

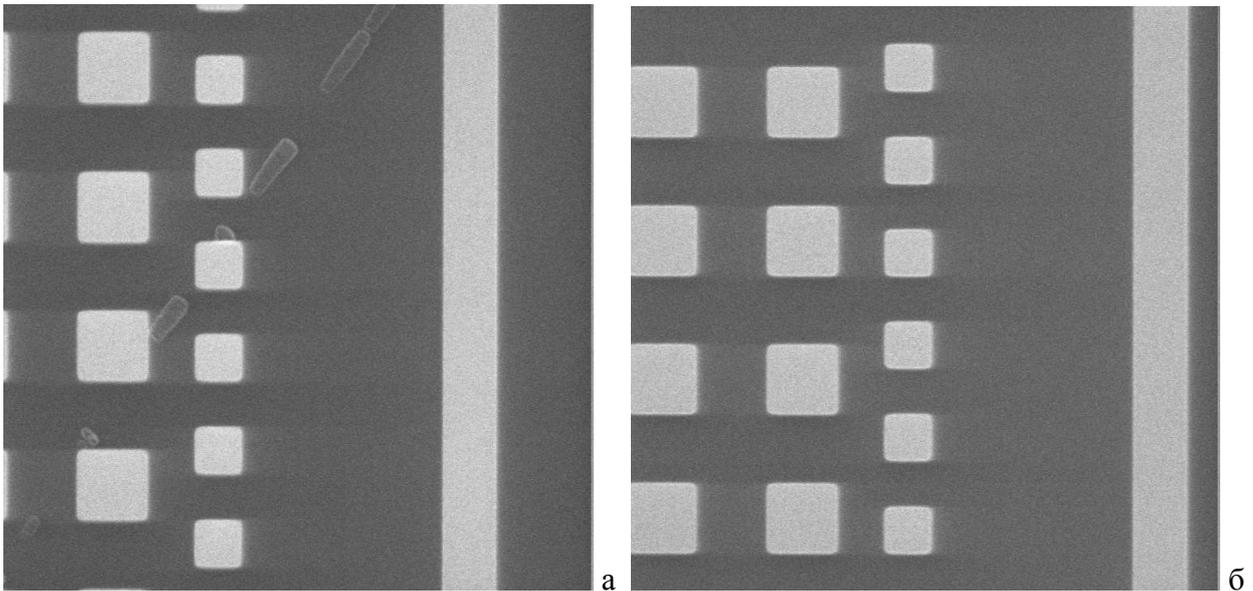


Рис. 4.4.1. Совмещаемые квазирегулярные изображения: а) базовое; б) сдвигаемое.

Задача совмещения двумерных сигналов (изображений) известна давно. Без использования распознавания и семантической интерпретации объектов, классическим средством является корреляционное совмещение изображений [110,210,262]. Оно сводится к поиску максимума двумерной функции ковариации изображений **S** и **T**:

$$\text{cov}_{T,S}(v,h) = \frac{1}{W_{v,h}} \sum_{m,n} (b_T(m,n) - \bar{b}_T)(b_S(m-v,n-h) - \bar{b}_S). \quad (4.4.1)$$

Здесь $b_S(m,n)$ и $b_T(m,n)$ — значения элементов в точке (m,n) ; \bar{b}_S и \bar{b}_T — средние значения изображений **S** и **T**; (v,h) — величина относительного сдвига изображений по вертикали и горизонтали; $W_{v,h}$ — число точек в пересечении изображений. Легко видеть, что для подсчета ковариационной функции по всему множеству возможных параметров (v,h) требуется порядка M^2N^2 количества операций, где $M \times N$ — размеры изображения.

Основная идея сокращения объема вычислений заключается в том, что в случае квазирегулярных изображений можно вычислять корреляцию не для всех возможных значений сдвига (v,h) , а только для некоторого множества точек. Для того, чтобы их выбрать, осуществим некоторые преобразования. Вычислим массивы средних значений по строкам и столбцам для совмещаемых изображений **S** и **T**, считая для простоты, что контуры объектов на изображениях имеют горизонтальную и вертикальную направленность:

$$V_S(m) = \frac{1}{N} \sum_{n=1}^N b_S(m,n) \quad \text{и} \quad V_T(m) = \frac{1}{N} \sum_{n=1}^N b_T(m,n) \quad \text{— массивы средних по строкам,}$$

$$H_S(n) = \frac{1}{M} \sum_{m=1}^M b_S(m,n) \quad \text{и} \quad H_T(n) = \frac{1}{M} \sum_{m=1}^M b_T(m,n) \quad \text{— массивы средних по столбцам.}$$

Чтобы подавить влияние несущественных яркостных различий, но сохранить контурные перепады, каждый из полученных массивов $V_S(m)$, $V_T(m)$, $H_S(n)$, $H_T(n)$ подвергается операции линейной фильтрации (нерезкого маскирования) по следующей формуле, аналогичной (3.3.1):

$$\tilde{f}(x) = f(x) - \frac{1}{2n+1} \sum_{y=x-n}^{x+n} f(y).$$

В результате этих операций получены проекции контурных перепадов, причем среднее значение каждой из полученных функций равно нулю: $M\{\tilde{f}(x)\} = 0$. По вертикальным $\tilde{V}_S(m)$, $\tilde{V}_T(m)$ и горизонтальным $\tilde{H}_S(n)$, $\tilde{H}_T(n)$ парам массивов подсчитываются одномерные ковариационные функции для соответствующих направлений:

$$\text{cov}_V(v) = \frac{1}{W} \sum_{m \in W} \tilde{V}_S(m) \tilde{V}_T(m-v) \quad \text{— ковариационная функция по вертикали,}$$

$$\text{cov}_H(h) = \frac{1}{W} \sum_{n \in W} \tilde{H}_S(n) \tilde{H}_T(n-h) \quad \text{— ковариационная функция по горизонтали.}$$

Локальные максимумы $\text{cov}_V(v)$ и $\text{cov}_H(h)$ соответствуют совпадению проекций контурных перепадов изображений **S** и **T**. Для приведенных выше изображений ковариационные функции по вертикальному $\text{cov}_V(v)$ и горизонтальному $\text{cov}_H(h)$ направлениям показаны на Рис. 4.4.2:

Точка глобального максимума двумерной ковариации (4.4.1) одновременно будет одной из точек локальных максимумов проекций контурных перепадов $\text{cov}_V(v)$ и $\text{cov}_H(h)$, что позволяет искать ее только среди множества точек $\{v_i, h_j\}$, составленного из координат локальных максимумов проекций v_i и h_j . Таким образом, алгоритм поиска очевиден. Находятся наборы

локальных максимумов v_i и h_j — ковариаций по вертикали $\text{cov}_V(v)$ и горизонтали $\text{cov}_H(h)$. Множество точек, среди которых ищутся координаты совмещения изображений, определяется как набор пар координат (v_i, h_j) , $\{i = 1, \dots, N_i; j = 1, \dots, N_j\}$, где v_i и h_j суть координаты соответствующих положительных локальных максимумов по вертикали и горизонтали. Также разумно ограничить относительный сдвиг изображений; нами был выбран диапазон $\pm 1/4$ размеров изображения. Для вышеприведенных изображений набор возможных координат совмещения изображений определяется как множество комбинаций из следующих наборов значений v_i и h_j : $v_i = \{-97, -40, 18, 78, 130\}$, $h_j = \{-76, 35, 95\}$.

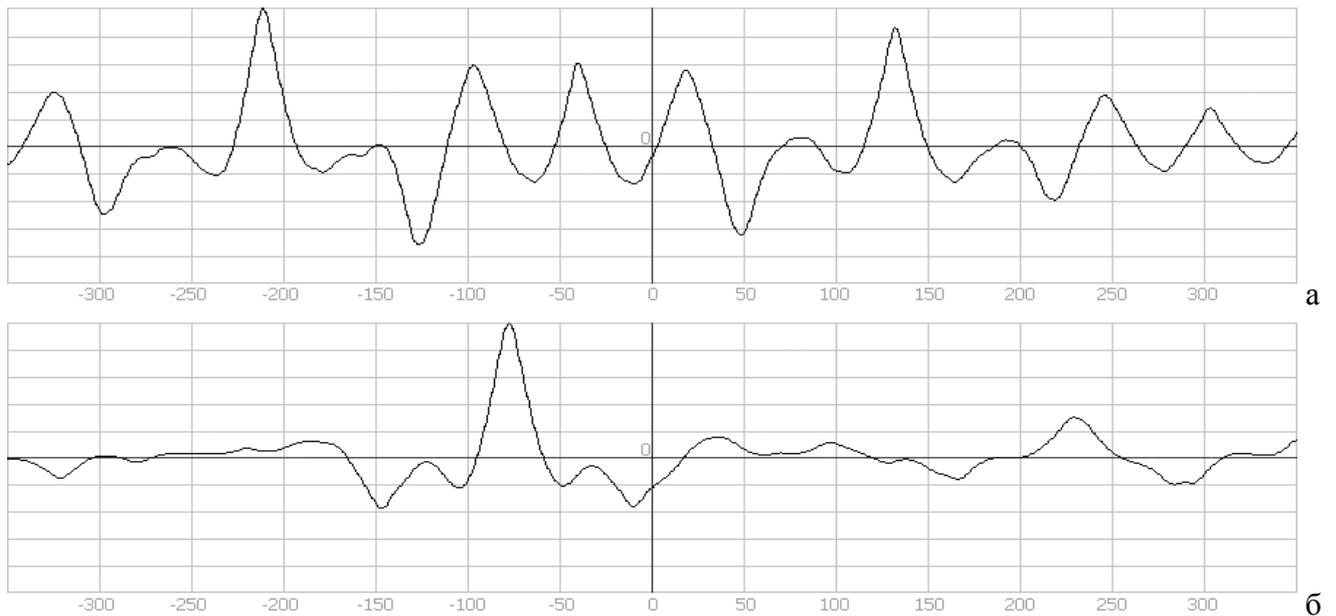


Рис. 4.4.2. Графики функций ковариации: а) по вертикали $\text{cov}_V(v)$; б) по горизонтали $\text{cov}_H(h)$.

В каждом из узлов построенной сетки осуществляется подсчет двумерных ковариаций по формуле (4.4.1), после чего выбирается узел с максимальным значением $\text{cov}_{T,S}(v, h)$. Необходимо отметить, что совершенно не обязательно таковым должен быть узел с координатами двух глобальных максимумов. Так, в данном примере оптимальным оказался узел с координатами $(130, -76)$. Отметим, что точка глобального максимума $\text{cov}_V(v)$, имеющая значение -210 , дает второй приемлемый вариант для совмещения $(-210, -76)$, но она выходит за границы $1/4$ размеров изображения. Затем в окрестности выбранного узла методом малых смещений уточняется точка локального максимума ковариации. Координаты найденной точки считаются параметрами сдвига, требуемого для совмещения изображений.

Предложенный алгоритм быстрого корреляционного совмещения изображений требует $MN(2 + k_m k_n + v)$ операций. Здесь MN — размеры изображения, k_m и k_n — число обнаруженных локальных максимумов проекций при условии, что значение локального максимума не менее $1/2$ значения глобального максимума, v — число итераций при поиске максимума вблизи точки

оптимума. Очевидно, это число зависит от изображений (в основном от k_m и k_n); обычно $k \approx 1 \div 3$, $v \approx 4 \div 7$. Использование для корреляционного совмещения изображений алгоритма быстрого преобразования Фурье требует не менее $MN(3C(\log M + \log N) + 2)$ операций, что приблизительно в $3C$ раз больше. Для изображений, приведенных в примере, коэффициент сокращения объема вычислений по сравнению с полным перебором составил около 8000.

Согласно многокомпонентной модели изображения (1.2.3), информация о контурах объектов содержится только в кусочно-гладкой компоненте s_{mn} . Поэтому можно рекомендовать выполнение декомпозиции (2.2.8) изображений и совмещение только s_{mn} – компонент.

Описанный способ корреляционного совмещения [165,234] может быть применен не только к квазирегулярным изображениям, таких как на Рис. 4.4.1. Так, он с успехом был использован для совмещения изображений капилляров при разработке методов анализа видеопоследовательностей [162,163], рассматриваемых в Главе 6.

Изложенные в разделе 4.4 результаты опубликованы в работах [165,234].

4.5. Обнаружение дефектов на снимках электронных микросхем

4.5.1. Постановка задачи

Исследуется научно-техническая задача обнаружения дефектов электронных микросхем по снимкам их поверхности, формируемым сканирующим электронным или оптическим микроскопами высокого разрешения. Для этого производится микросъемка двух аналогичных участков — один с вероятным дефектом, второй без дефекта. Изображения имеют одно и то же увеличение и отображают одинаковые структуры. Между объектами на снимках, как правило, имеется сдвиг, но нет поворота. Первое изображение (дефектное) с высокой вероятностью может содержать один или несколько дефектов неопределенной формы и размеров, второе (эталонное) дефектов не содержит. Задача состоит в нахождении местоположения и формы дефектов на первом изображении.

Виды дефектов

Реальные дефекты на электронных микросхемах возникают либо вследствие появления посторонних частиц (по сути, «грязи») на каком-то из этапов их изготовления, либо из-за нарушения формы деталей схемы. На изображениях они выглядят следующим образом:

- дефекты в виде одиночных или множественных пятен разнообразной формы, размеров и яркости на ровном фоне или перекрывающие несколько первоначальных объектов;
- разрывы протяженных объектов (т.е. проводников и пр.);
- нарушения формы контуров объектов.

Диапазон возможных размеров дефектов — от нескольких элементов до изображения целиком.

Особенности изображений микросхем

Формы и размеры объектов на поверхности электронных микросхем могут быть совершенно различными, однако разнообразие одновременно возможных их характеристик (имея в виду яркость или отражающую способность) весьма невелико; они сохраняются на протяжении всего объекта и повторяются в других частях изображения. Можно сказать, что изображения микросхем характеризуются постоянством яркостей каждой из областей. Такое представление вполне соответствует рассмотренной выше кусочно-постоянной модели изображения (1.2.24). Более того, поскольку текстурная составляющая на самих деталях практически отсутствует, то модель (1.2.24) может быть упрощена до следующей:

$$x_{ij}^v = \sum_{u=1}^V \delta_{uv} (S_{mn}^v + \xi_{ij}^v). \quad (4.5.1)$$

Согласно модели, фон может восприниматься как один (пусть и достаточно большой) объект.

Шум ξ считается нормально распределенным с нулевым средним и дисперсией $\sigma_\xi^2: N(0, \sigma_\xi^2)$.

Полагаем также, что разность яркостей p и q соседних объектов удовлетворяет неравенству:

$$|S^p - S^q| > 2\sigma_\xi. \quad (4.5.2)$$

Из особенностей реальных изображений и дефектов важно отметить следующие:

- шум на изображениях весьма заметный: $\sigma_\xi \approx 2,5\%$ и более диапазона яркостей;
- градационные характеристики и контрасты объектов на парах изображений (дефектном и эталонном) могут существенно различаться;
- формы объектов на эталонном и дефектном изображениях повторяются неточно; возможны отличия положений границ объектов на совмещенных изображениях в 2-3 элемента;
- возможна неравномерность (неконстантность) яркости протяженных объектов;
- дефекты часто имеют размытые очертания, что затрудняет точное определение их границ;
- контраст пятна дефекта может быть весьма мал и близок уровню шума.

Примеры пар изображений (дефектное–эталонное) представлены на Рис. 4.5.1 и 4.5.2.

4.5.2. Подход к построению алгоритма обнаружения дефектов

Поскольку изображения поверхности электронных микросхем удовлетворяют достаточно простой кусочно-постоянной модели (4.5.1), то для обнаружения различия пары изображений хотелось бы применить подход, описанный в разделе 4.2. К сожалению ряд особенностей как дефектов (низкий контраст и большие возможные размеры дефектов, превышающие допустимые величины), так и изображений в целом (неточности повторения форм объектов) не позволяют им воспользоваться.

Указанные причины, а также требования к вычислительной скорости алгоритма анализа (число операций, требуемое процедурой декомпозиции, несколько превышает допустимый предел), привели к решению использовать алгоритм, основанный на прямом поиске различий эталонного и дефектного изображений. При таком подходе алгоритм обнаружения дефектов по паре изображений может быть разделен на два последовательных и независимых этапа. Первый — совмещение и предобработка пары изображений, и второй — нахождение области дефекта на уже совмещенных дефектном и эталонном изображениях.

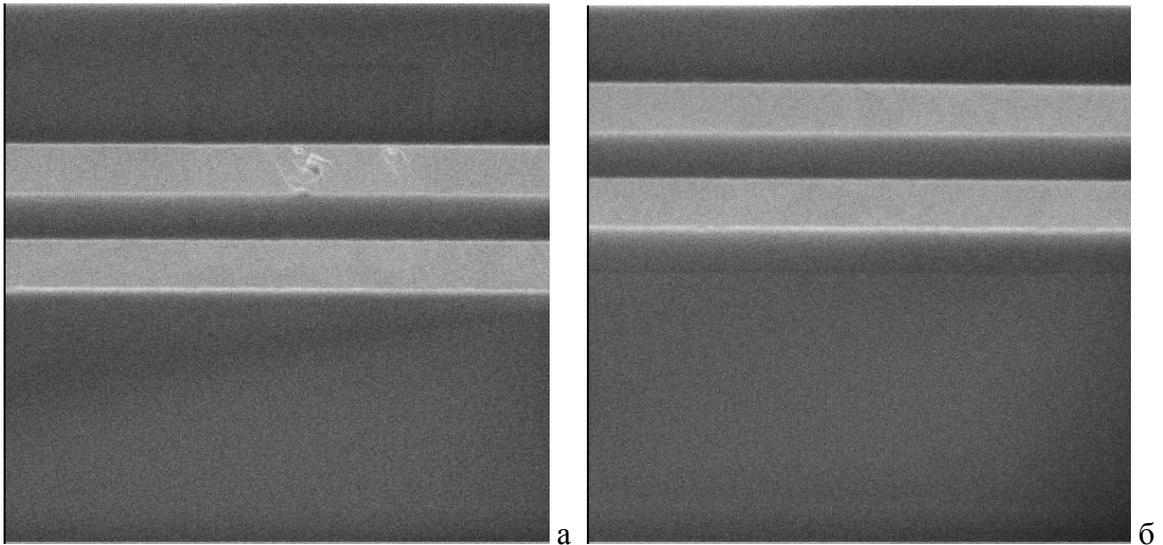


Рис. 4.5.1. Фрагменты изображений микросхем: а) с дефектом; б) без дефекта.

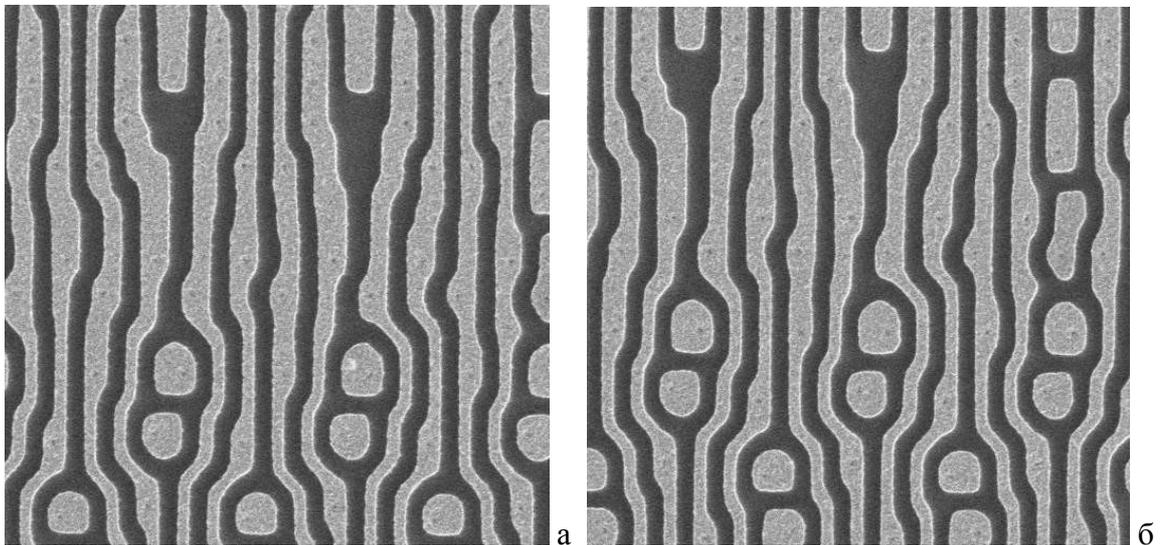


Рис. 4.5.2. Фрагменты изображений микросхем: а) с дефектом; б) без дефекта.

Совмещение изображений

На первом этапе сопоставления изображений необходимо осуществить их пространственное совмещение. Наиболее подходящим средством для этого является корреляционное совмещение [210,262]. При всем разнообразии возможных форм контуров объектов удалось

подметить две морфологические особенности, такие, что для каждого из анализируемых изображений выполняется хотя бы одна из них: а) значительная часть контуров объектов имеет горизонтальную и/или вертикальную направленность; б) объекты расположены на изображении периодически, и периодичность также имеет вертикальную и горизонтальную составляющие. Алгоритм быстрого корреляционного совмещения изображений с подобными особенностями изложен в разделе 4.4, где совокупность таких характеристик названа квазирегулярной структурой объектов (квазирегулярностью).

Быстрый алгоритм совмещения изображений, использующий свойство квазирегулярности изображений [165,234], сводится к формированию перпендикулярных проекций изображений, нахождению двумерного набора точек возможного совмещения изображений на основе одномерного корреляционного совмещения проекций, выбора оптимальной точки и окончательному уточнению позиции совмещения (раздел 4.4).

Возможности анализа совмещенных изображений

Предполагаем, что дефектное X_D и эталонное X_E изображения уже совмещены. Однако надежно обнаружить дефект по разности совмещенных изображений простым пороговым сравнением не удастся — на разностном сигнале возникает довольно много ложных пятен. Причины их появления кроются в указанных выше особенностях изображений, а именно: в отличии форм объектов дефектного X_D и эталонного X_E изображений, в неравномерностях яркостей объектов; в высоком уровне шума, в наличии мелких пятен, по своим характеристикам похожих на искомые дефекты. Примеры получаемых таким путем разностных изображений показаны на Рис. 4.5.3. Отыскать области дефектов на подобных изображениях без морфологического анализа удастся далеко не всегда.

Обнаружение области дефекта становится возможным при совместном анализе пары изображений. Для этого предлагается алгоритм, включающий следующие стадии:

1. Анализ и предобработка исходных изображений. На данной стадии осуществляется взаимное приведение градационных характеристик пары изображений, оценивается уровень шума и выполняется градационная коррекция.
2. Формирование двумерного массива несовпадений. Определяются точки различий пары изображений, формируется двумерная маска несовпадений и осуществляется первичная фильтрация мелких пятен.
3. Формирование и обработка списка областей несовпадений. Осуществляется переход от исходного представления данных, как двумерного массива пикселей, к списку описаний каждой отдельной области. Сформированный список областей подвергается последовательным этапам анализа, фильтрации, и упрощения.

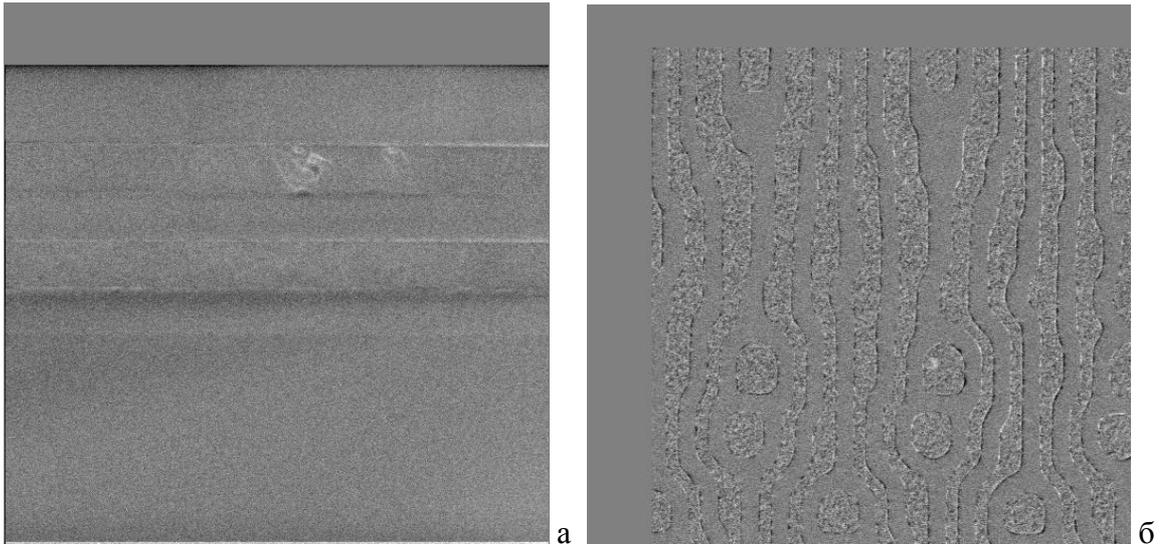


Рис. 4.5.3. Разности пар совмещенных изображений: а) на Рис. 4.5.1; б) на Рис. 4.5.2.

4.5.3. Предварительная обработка изображений

Градационная коррекция

Для согласования амплитудных характеристик дефектного и эталонного изображений используется следующий алгоритм коррекции. Пусть M_E и σ_E^2 — среднее и дисперсия эталонного, а M_D и σ_D^2 — дефектного изображения. Для согласования изображений выполняется поэлементное преобразование:

$$x_{mn} = \frac{\sigma_E}{\sigma_D} (\tilde{x}_{mn} - M_D) + M_E. \quad (4.5.3)$$

Здесь \tilde{x}_{mn} — элемент исходного, а x_{mn} — преобразованного дефектного изображения.

Определение параметров шума

Согласно (4.5.1) шум считается независимым и нормально распределенным: $N(0, \sigma_\xi^2)$. Для его определения следовало бы воспользоваться выражениями (1.2.15) или (1.2.20), но для этого необходимо предварительно отделить фоновые участки изображения от контурных. По этой причине значение СКО шума σ_ξ на дефектном и эталонном изображениях определяется через статистику распределения разности соседних элементов:

$$\Delta x_{mn} = x_{mn} - x_{m,n+1}. \quad (4.5.4)$$

Поскольку шум имеет нормальное распределение, распределение разности элементов (4.5.4), попадающих в одну область, будет также нормальным: $N(0, 2\sigma_\xi^2)$. К сожалению границы отдельных областей изображения нам не известны, а значит при подсчете статистики распределения Δx заметное влияние будут оказывать элементы, попадающие на границы областей (контурные

участки); это влияние особенно сильно на «хвостах» распределения. Решить задачу удастся следующим способом, основанным на рассмотренном в разделе (1.2.2) подходе.

Будем считать, что Δx есть смесь двух нормальных распределений Δx^1 и Δx^2 со значительно различающимися дисперсиями, причем вероятности элементов $P\{\Delta x^2\} \ll P\{\Delta x^1\}$, что справедливо для соотношения числа контурных и фоновых элементов. Выберем некоторые два значения F_1 и F_2 вблизи средних значений: $0,3 < F_1 < F_2 < 0,7$. Подсчитаем функцию распределения Δx в (4.5.4): $F(z) = P\{\Delta x < z\}$ и найдем u_1 и u_2 такие, что $F(u_1) = F_1$, а $F(u_2) = F_2$. Зная функцию нормированного нормального распределения $\Phi(z)$, найдем v_1 и v_2 такие, что $\Phi(v_1) = F_1$, а $\Phi(v_2) = F_2$. Тогда при вышеуказанном условии соотношения вероятностей Δx^1 и Δx^2 можно оценить дисперсию распределения Δx^1 как: $\sigma_{\Delta x} \approx (u_2 - u_1)/(v_2 - v_1)$. При реализации были выбраны $v_1 = -0,6$ и $v_2 = 0,6$; отсюда $F_1 \approx 0,274$ и $F_2 \approx 0,726$, следовательно точки u_1 и u_2 находятся по функции распределения $F(\Delta x)$ как удовлетворяющие равенствам: $F(u_1) = 0,274$ и $F(u_2) = 0,726$. С учетом того, что дисперсия распределения Δx равна $2\sigma_{\xi}^2$, получим:

$$\sigma_{\xi} \approx (u_2 - u_1)/(\sqrt{2}(v_2 - v_1)). \quad (4.5.5)$$

Формирование разностного сигнала выравнивания

Как уже отмечалось, на реальных изображениях наблюдается заметная неравномерность (неконстантность) фоновой составляющей внутри объектов, которая может привести к появлению ложных пятен. Для устранения таких артефактов при формировании массива несовпадений используется сигнал выравнивания, вычисляемый как разность дефектного и эталонного изображений $d_{mn} = x_D(m,n) - x_E(m,n)$, которая затем подвергается фильтрации низких частот путем двумерного сглаживания по фрагменту большого размера [155]:

$$\bar{d}_{mn} = \frac{1}{(2h+1)^2} \sum_{i=m-h}^{m+h} \sum_{j=n-h}^{n+h} (x_D(i,j) - x_E(i,j)). \quad (4.5.6)$$

Здесь $h \times h$ — размеры фрагмента усреднения ($h \approx 1/6$ размеров изображения).

4.5.4. Формирование двумерного массива несовпадений

Формирование двумерного массива несовпадений является фундаментальным этапом обнаружения различий пары изображений. Выше уже были приведены факторы, мешающие использовать для этого значение разности изображений; для уменьшения их влияния разработан следующий алгоритм.

Формирование первичного массива несовпадений

Основная идея алгоритма формирования массива несовпадений в следующем. Осуществляется поэлементное сравнение дефектного \mathbf{X}_D и эталонного \mathbf{X}_E изображений. В каждой точке (m,n) вычисляются интервалы допустимых значений яркости: $[l_D, r_D]$ для дефектного и $[l_E, r_E]$ для

эталонного изображений. Считается, что в точке (m, n) наблюдается совпадение, если значения элементов $x_D(m, n)$ и $x_E(m, n)$ каждого из изображений \mathbf{X}_D и \mathbf{X}_E оказываются внутри допустимого интервала, вычисленного по второму изображению. Для этого в каждой точке (m, n) по окрестности 3×3 , окружающей обрабатываемый элемент, находятся оценки: $\min_D(m, n)$ и $\max_D(m, n)$ — минимальное и максимальное значения дефектного изображения; $\min_E(m, n)$ и $\max_E(m, n)$ — то же для эталонного изображения. По значению \bar{d}_{mn} сглаженного разностного сигнала выравнивания (4.5.6) определяются инкременты d^+ и d^- :

$$d^+ = \begin{cases} \bar{d}_{mn}, & \text{если } \bar{d}_{mn} > 0, \\ 0, & \text{если } \bar{d}_{mn} \leq 0; \end{cases} \quad \text{и} \quad d^- = \begin{cases} 0, & \text{если } \bar{d}_{mn} \geq 0, \\ \bar{d}_{mn}, & \text{если } \bar{d}_{mn} < 0. \end{cases}$$

Границы интервалов l_D , r_D , l_E и r_E вычисляются по следующим формулам, в которых коэффициенты K_1 и K_2 определяются исходя из величины измеренного шума σ на изображениях (4.5.5):

$$\begin{aligned} l_D &= \min_D(m, n) - K_1[\max_D(m, n) - \min_D(m, n)] - K_2 - d^+; \\ r_D &= \max_D(m, n) + K_1[\max_D(m, n) - \min_D(m, n)] + K_2 - d^-; \\ l_E &= \min_E(m, n) - K_1[\max_E(m, n) - \min_E(m, n)] - K_2 + d^+; \\ r_E &= \max_E(m, n) + K_1[\max_E(m, n) - \min_E(m, n)] + K_2 + d^-. \end{aligned}$$

Считается, что в точке (m, n) имеется совпадение, если одновременно выполняются условия:

$$\begin{cases} l_D(m, n) \leq x_E(m, n) \leq r_D(m, n); \\ l_E(m, n) \leq x_D(m, n) \leq r_E(m, n). \end{cases} \quad (4.5.7)$$

Иначе точка (m, n) относится к области возможного дефекта. По окончании процедуры формируется первичная маска несовпадений, элементы которой равны нулю в точках совпадений и $(x_D - x_E)$ в точках несовпадений.

Удаление шумовых выбросов

Высокий уровень шума на исходных изображениях приводит к появлению в первичном массиве несовпадений большого количества мелких всплесков, состоящих из одного или нескольких элементов. Для их удаления используется следующий простой алгоритм.

Для каждого ненулевого элемента по маске несовпадений проверяются 8 его ближайших соседей. Если число окружающих ненулевых элементов, имеющих тот же знак, что и центральный элемент, меньше или равно некоторого заданного порога, то текущий элемент удаляется из маски несовпадений и обнуляется.

4.5.5. Формирование списка областей несовпадений

Во многих случаях пятна реальных дефектов не являются самыми яркими, или самыми большими по площади областями несовпадений. Для их обнаружения требуется анализ участ-

ков несовпадений, определение и удаление ложных областей, и окончательный выбор одной или нескольких областей, идентифицируемых как дефектные области.

На данной стадии осуществляется переход от анализа исходных данных как двумерного изображения, к описанию и анализу данных как набора отдельных пятен, являющихся кандидатами на искомые области дефекта. При этом, конечно же, остаются доступными исходные двумерные данные (изображения) но их использование становится вспомогательным, а основным материалом анализа становится одномерный набор дескрипторов, характеризующий области несовпадений.

Операции производятся со следующими объектами:

- а) **пятно** — дескриптор, описывающий отдельную область на изображении и содержащий всю требуемую информацию об области несовпадений;
- б) **список пятен** — структура, объединяющая все пятна на изображении и обеспечивающая их упорядочение, нумерацию и ассоциативные связи;
- в) **список ассоциаций** — промежуточная таблица, объединяющая нескольких объектов в один.

Формирование первичного списка пятен

Формирование списка пятен несовпадений включает следующие операции: анализ двумерной маски несовпадений, построение первичного списка пятен и списка ассоциаций для объединения соприкасающихся пятен; объединение ассоциированных пятен и корректировка списка пятен; подсчет требуемых характеристик пятен. удаление пустых мест в списке пятен и сортировка списка по качеству.

При формировании списка пятен для каждого пятна формируются и сохраняются следующие характеристики:

N — число точек в пятне (площадь пятна);

m_t, m_b, n_l, n_r — координаты прямоугольника, описывающего пятно;

ζ^+ — сумма положительных разностей значений элементов $(x_D - x_E) > 0$;

ζ^- — сумма отрицательных разностей значений элементов $(x_D - x_E) < 0$;

Q — «качество» пятна, определяемое функционалом $Q = ((\zeta^+ - \zeta^-) \log(k+32)) / (N+16)$. Здесь $k = N$ для малых ($N < 64$), и $k = 64$ для больших N .

Удаление периферийных и мелких пятен

Первоначально удаляются пятна, отстоящие от краев изображения на расстояние менее 1/20 размеров изображения, а также пятна, отстоящие от краев изображения на расстояние менее 1/10 размеров изображения и не попадающие в число первых 16 по качеству. Затем удаляются пятна низкого качества, расположенные достаточно далеко от пятен высокого качества. Находится пятно с максимальным уровнем качества и определяется порог, выше

которого пятно считается приемлемым, а ниже — слабым. Удаляются те слабые пятна, которые не находятся вблизи пятен приемлемого качества. Оставшиеся пятна малого размера либо ассоциируются с соседними, либо удаляются. Текущий результат показан на Рис. 4.5.4.

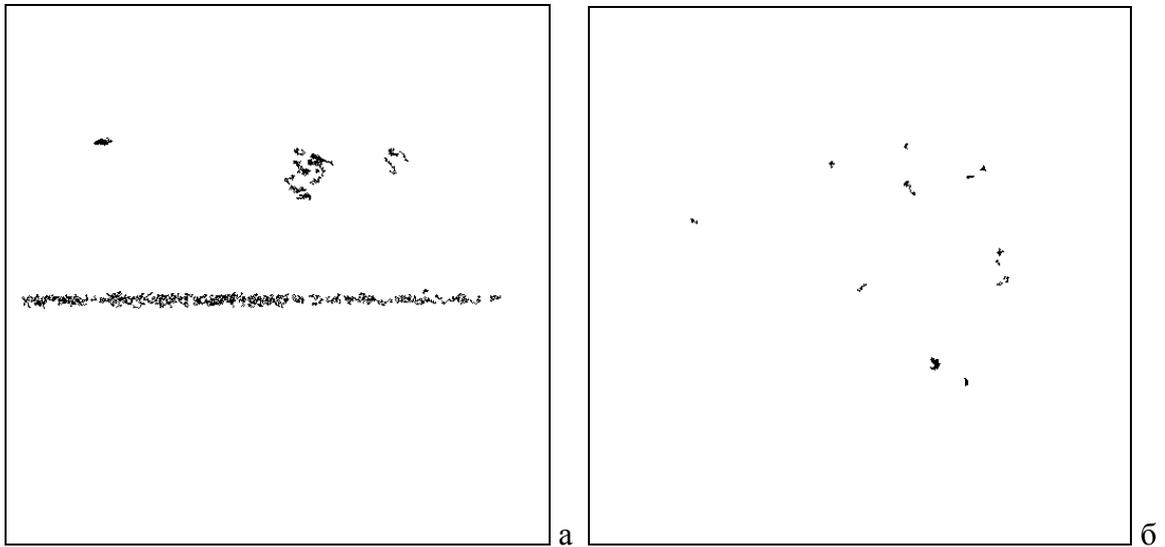


Рис. 4.5.4. Вид маски после удаления периферийных и мелких пятен

Над полученной маской пятен выполняется морфологическая операция «замыкание», состоящая в последовательном выполнении операций дилатации и эрозии [40]. Эта пара операций достаточно часто используется в задачах обнаружения объектов и формирования масок. После чего производится еще один этап обнаружения и удаления пятен малого размера и формируется обновленный список пятен.

4.5.6. Удаление ложных пятен

После выполнения вышеуказанных процедур на сформированной маске еще остается заметное число ложных пятен. Их удаление происходит на финальном этапе. Для этого подсчитываются дополнительные характеристики оставшихся пятен и формируются некоторые новые признаки. Сочетание этих признаков позволяет принять решение о том, какие пятна действительно соответствуют дефектами, а какие являются ложными и должны быть удалены.

Формируются следующие дополнительные характеристики пятен:

L — длина границы пятна, измеряемая как суммарное число сторон пограничных элементов пятна, контактирующих с фоном;

Σ_D^+ , Σ_D^- — суммы положительных $(x_d - x_b) > 0$ и отрицательных $(x_d - x_b) < 0$ разностей элементов на всех граничных элементах пятна (x_d — элемент пятна, x_b — соседний элемент фона) по дефектному изображению ;

Σ_E^+ , Σ_E^- — суммы положительных $(x_d - x_b) > 0$ и отрицательных $(x_d - x_b) < 0$ разностей элементов на всех граничных элементах пятна по эталонному изображению;

L_M — число контурных элементов с совпадающим знаком разности $(x_d - x_b)$;

R — удаленность пятна от центра анализируемого участка изображения ($0 \leq R \leq 1$);

D — разность яркостей, определяемая как $D = (\zeta^+ - \zeta^-)(1 - R^2)/N$;

I_{true}, I_{false} — два набора битовых индикаторов, в которых пятно помечается как «правильное» или «ложное» по тем или иным признакам.

C — функционал «контраста» пятна $C = \left(\left(K \left\| \sum_D^+ - \sum_D^- \right\| - \left\| \sum_E^+ - \sum_E^- \right\| \right) / L \right) + ((\zeta^+ - \zeta^-) / N)$; здесь K — коэффициент, выравнивающий соотношение между средней контурной разницей и средней разницей элементов по всему пятну дефектного и эталонного изображений (значение оценивается по тестовым изображениям и равно $K = 6$).

F — фактор пятна (функционал характеристик, $0 \leq F \leq 1$), определяющий его «заметность»; изначально устанавливается в 1 и изменяется при анализе.

При подсчете характеристик осуществляется нормализация значений качества Q и контраста C по их максимальным значениям: $Q = Q_0 / Q_{max}$, $C = C_0 / C_{max}$. Таким образом, как качество, так и контраст принимают значения в диапазоне от 0 до 1.

Стирание пятен, ограниченных совпадающими контурами

Проверяется соотношение $L_M \geq 0,5L$ для каждого из оставшихся пятен. Если оно выполняется, то считается, что данный контур очерчивает повторяющиеся по форме детали, находящиеся на анализируемом и эталонном изображениях, но различающиеся по контрасту с фоном. Поскольку разница в контрастах деталей не считается дефектом, такие объекты удаляются из маски дефектов.

Подавление пятен, удаленных от центра изображения

Подсчитывается расстояние r между центром пересекающейся части изображений и центром каждого пятна (под этим понимается центр прямоугольника, описывающего пятно). Затем формируется фактор R удаленности пятна от центра изображения, который нормируется на размер половины диагонали изображения: $R = 2r/d$. Формируется значение разности яркостей D . Подавление пятен, удаленных от центра, состоит в уменьшении значения фактора пятна F по формуле $F = (1 - R)F_0$.

Выбор главного пятна и подавление удаленных от него пятен

Главное пятно определяется как то, на котором достигается максимум функционала

$$\varphi = QCF \log N, \quad (4.5.8)$$

являющегося произведением следующих характеристик пятна: качества, контраста, фактора, и логарифма размеров. Далее также будут использованы значения размера N_{main} и функционала Φ_{main} главного пятна.

Как и выше, подавление пятен сводится к модификации (уменьшению) значений факторов F пятен, удаленных от главного пятна. Для этого первоначально находится базовый допустимый радиус R_0 как функция линейных размеров главного пятна и выбранного минимального значения R_{min} :

$$R_0 = 2(m_b - m_t + n_r - n_l) + R_{min};$$

здесь $2(m_b - m_t + n_r - n_l)$ — периметр прямоугольника, описывающего главное пятно. Если оказывается, что $R_0 > 2R_{min}$, то значение ограничивается $R_0 = 2R_{min}$.

Осуществляется проверка всех оставшихся пятен. Первоначально задается допустимый радиус R . Для обычных пятен он выбирается как $R = R_0$. Для мощных пятен, удовлетворяющих условиям: $N > 0,3N_{main}$ и $\varphi > 0,5\varphi_{main}$ и не имеющих маркированных битовых полей в I_{false} , выбирается большее значение: $R = 3R_0$.

Затем подсчитывается расстояние r между центром анализируемого пятна и центром главного пятна. Фактор пятна F подвергается преобразованию:

$$F = \begin{cases} F, & \text{если } r < R, \\ F(2R - r)/R, & \text{если } R < r < 2R, \\ 0, & \text{если } r > 2R. \end{cases}$$

Удаление оставшихся лишних пятен и формирование результата

Значения контраста C , качества Q , размера N , а также разности яркостей D для каждого из пятен проверяются на условия: $C \geq k_c$, $Q \geq k_q$, $N \geq k_N N_{max}$ и $D \geq k_d D_{max}$. В зависимости от выполнения этих условий устанавливаются соответствующие индикаторы I_{true} или I_{false} . Все пятна, имеющие ненулевые значения индикатора I_{false} , а также те, у которых низкий фактор пятна ($F < 0,1$), удаляются.

Формируется двухградационное изображение, на котором оставшиеся пятна составляют области дефектов. Полученное изображение считается окончательным результатом — маской пятен дефектов (Рис. 4.5.5).

При тестировании на серии из 464 снимков алгоритм показал следующие результаты: правильное обнаружение (результат удовлетворяет требуемым критериям) — 92,03%, частично удовлетворительное обнаружение (результат не соответствует требуемому процентному критерию обнаруживаемой площади дефекта) — 6,25%, неправильное обнаружение — 1,72%. Алгоритм обнаружения дефектов на снимках электронных микросхем был разработан по заказу

компании MicroSpec Technologies Ltd. (Carl Zeiss Group) для использования в процедуре диагностики при промышленном производстве интегральных микросхем.

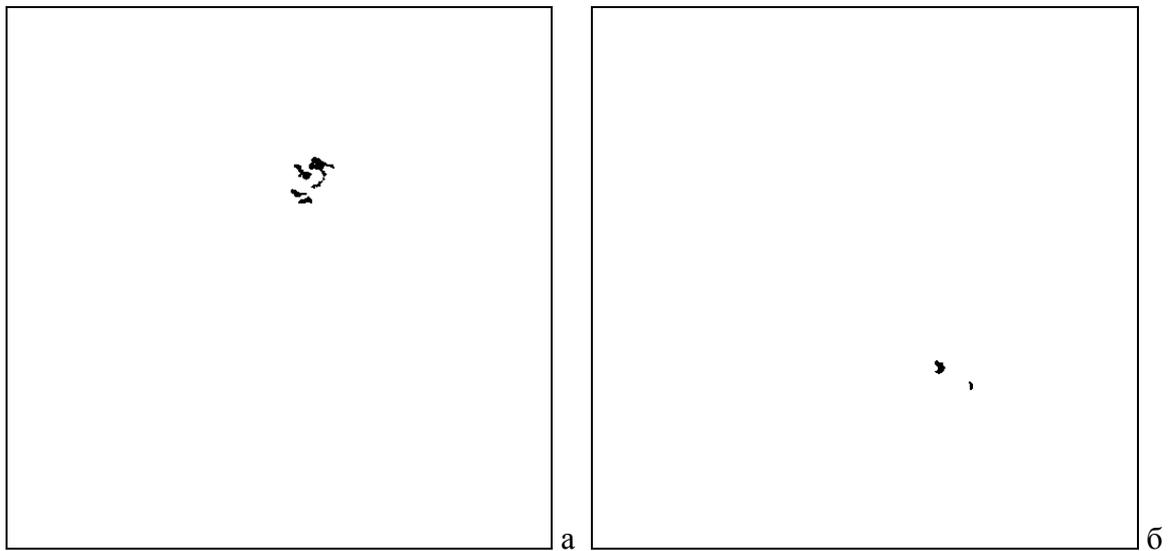


Рис. 4.5.5. Сформированные маски пятен дефектов по изображениям: а) Рис. 4.5.1; б) Рис. 4.5.2.

Изложенные в разделе 4.5 результаты опубликованы в работе [176].

4.6. Выводы и результаты

Показано, что двухмасштабная многокомпонентная модель и процедура декомпозиции изображения являются подходящим фундаментом для построения ряда алгоритмов анализа видеoinформации, в числе которых: обнаружение объектов, дефектов, контуров, различий на изображениях и другие.

На основе многокомпонентной модели и методе декомпозиции предложен новый подход и разработаны алгоритмы решения задачи обнаружения объектов на изображении по их площади, в которой рассмотрены три возможных варианта: обнаружения объектов с площадью меньше задаваемого порога, больше порога и с площадью в задаваемом интервале. Введено понятие «локальной площади», позволяющее обнаруживать тонкие протяженные объекты.

Рассмотрены теоретические аспекты обнаружения различий изображений с малыми изменениями. Предложен метод нахождения различий на изображениях, основанный на предварительном обнаружении объектов.

Предложен новый подход к процедуре нахождения границ объектов как участков с перепадами яркостей, состоящий в том, что контурные перепады следует обнаруживать не непосредственно по исходному изображению, а по предварительно сглаженной компоненте, формируемой в результате операции декомпозиции. Это позволяет с одной стороны повысить

слитность контурных линий и обеспечить меньшую их толщину, а с другой — значительно снизить количество ошибок первого и второго рода, вызываемых шумом и текстурой.

Для изображений с квазирегулярной структурой объектов разработан алгоритм быстрого их корреляционного совмещения. Количество операций, требуемое предлагаемым алгоритмом, равно $MN(2 + k_m k_n + v)$, где $M \times N$ — число элементов изображения, k_m и k_n — число локальных максимумов проекций, v — число итераций при поиске максимума вблизи точки оптимума (как правило $k \approx 1 \div 3$, $v \approx 4 \div 7$). Для реальных изображений размерами 512×512 элементов коэффициент сокращения по сравнению с полным перебором составляет величину порядка 10^4 .

Исследована и решена важная научно-техническая задача обнаружения широкого класса дефектов на поверхности электронных микросхем по снимкам, формируемым оптическим или сканирующим электронным микроскопом высокого разрешения. Разработаны алгоритмы обнаружения дефектов в широком диапазоне их проявлений, которыми могут быть посторонние пятна разнообразной формы и размеров, разрывы протяженных объектов (проводников), нарушения формы контуров объектов. В разработанной для обнаружения дефектов последовательности операций использованы алгоритмы совмещения и обработки изображений, построенные на основе предложенной двухмасштабной многокомпонентной модели изображения. Проверка алгоритма на серии из 464 тестовых снимков показала его полное соответствие предъявляемым требованиям в 92% и частичное — в 6,25% случаев.

Полученные результаты иллюстрируют важный вывод Главы 2, заключающийся в том, что эффективным подходом к анализу и обработке изображения является предварительная его декомпозиция на различные информационные компоненты и последующее использование той или другой полученной компоненты.

ГЛАВА 5. ПРИМЕНЕНИЕ МОДЕЛИ И МЕТОДОВ ОБРАБОТКИ ДЛЯ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ

Введение

Цель сегментации изображения заключается в разбиении исходного изображения на множество непересекающихся связных областей, ассоциируемых с объектами наблюдаемой сцены или их частями в соответствии с некоторыми выбранными критериями. Результатом является получение карты областей (сегментов) изображения, т.е. отображение множества точек изображения в конечный и сравнительно небольшой набор значений $1, \dots, Q$, где Q есть финальное число сегментов. Такое разбиение может служить как предварительным шагом для последующего морфологического или семантического анализа и интерпретации наблюдаемой сцены, так и самостоятельным результатом.

Задача сегментации изображений имеет длительную историю. Подходов к ее решению предложено много, и исследованию данного вопроса посвящено большое количество работ. Тем не менее, сегментация по-прежнему остается одной из проблем анализа поступающей видеоинформации. Разработано много различных методов и алгоритмов, но применимость каждого из них сильно зависит от решаемой задачи. В большинстве случаев алгоритмы сегментации строятся либо на использовании базовых свойств изображений, таких как однородность на связных областях и разрывность на их границах [40,113,343], либо на преобразовании сигнала в пространство признаков и классификации в этом пространстве [53].

Среди подходов к построению алгоритмов сегментации можно выделить четыре класса. Первые два из них в основном ориентированы на однокомпонентные — монохромные или черно-белые изображения, тогда как третий и четвертый классы допускают анализ многокомпонентных данных.

1. Разбиение изображения проведением границ. Заключается в обнаружении границ контурными операторами, в их прослеживании, связывании, и составлении из них замкнутых границ областей. Большинство алгоритмов обнаружения контурных границ основано на фильтрации, преобразующей изображение в двумерный массив, значения которого соответствуют вероятности нахождения контура в соответствующей точке. Наиболее известные контурные операторы основаны на вычислении первой производной (операторы Робертса [121], Собеля [351] и др. [25,40,113,335,343]) или второй производной (оператор Лапласа [40,113,343]). Результаты их применения выглядят неплохо с точки зрения визуальной оценки, однако не удовлетворяют ряду формальных критериев, в частности требованиям непрерывности и

минимальной толщины контурных линий. Альтернативой алгоритмам фильтрации является прослеживание контуров [25,121]. Хорошо известен так называемый метод “zero crossing”, основанный на обнаружении точек перехода через ноль второй производной [246] — он создает контурные линии минимальной толщины, но при этом появляется много ложных линий, слабо соответствующих реальным контурам. Неплохие результаты достигаются алгоритмом Кэнни [223], однако получаемый им контурный препарат тоже не идеален. Известны также и другие методы и алгоритмы [40,113,335,343]. Сегментация на основе проведения границ как правило используется для специфических классов анализируемых изображений.

2. Морфологический подход. Из подобных методов наиболее известна сегментация по водоразделам [280,352], идея которой состоит в интерпретации гладких областей как локальных бассейнов, а контуров между ними как водоразделов. Близкие результаты достигаются так называемой сегментацией «спуском» (tobogganing) [353]. В большинстве случаев достигаемым результатом является излишне подробное разбиение, что требует дополнительного анализа. Оба подхода эффективны для однокомпонентных изображений или признаков, но слабо применимы к многокомпонентным цветным и текстурным изображениям.

3. Разбиение изображения на однородные по значению области. Здесь в первую очередь необходимо упомянуть пороговые методы, использующие как глобальные, так и адаптивные подходы [40,288]. Подобные методы просты в реализации, но подобрать удовлетворительные уровни разделения не всегда возможно. Известны методы выращивания областей [259,271], заключающиеся в группировке элементов и мелких областей изображения в более крупные, начиная из так называемых «центров кристаллизации». Проблемы данного метода — выбор подходящих мер близости элементов и критериев остановки процесса выращивания областей [270]. Альтернативой выращиванию служат методы, основанные на первичном разбиении изображения на множество малых областей и последующем их слиянии или делении [347].

4. Классификация в пространстве признаков. Состоит в выборе отображения набора входных данных в некоторое пространство признаков и последующем решении задачи кластеризации — разбиении выбранного пространства на классы, базируясь на плотности распределения в нем [53,286]. Данные методы удобны, если количество классов объектов в изображении заранее известно, или когда имеется возможность предварительного обучения классификатора. Если отображения различных областей исходного изображения в пространстве признаков перекрываются, классическая процедура сегментации может оказаться неверной. Преодолеть проблему предлагается расширением пространства признаков, например, добавлением значений координат изображения [53].

Как и при распознавании, которому задача сегментация изображений весьма близка, кроме собственно выбора алгоритма, существенными вопросами являются выбор пространства признаков, а также использование специфической информации об анализируемом сигнале. В нашем случае последний из указанных вопросов сводится к использованию пространственной корреляции близкорасположенных элементов изображения. Помочь в этом может применение рассмотренной в разделе 1.2 двухмасштабной многокомпонентной модели изображения. Ниже будут рассмотрены алгоритмы решения двух задач сегментации: сегментации прослеживанием контуров объектов при сложном виде границ на примере сегментации снимков поверхности микросхем, формируемых электронным микроскопом, а также сегментации изображений общего вида на основе цвето-текстурных различий объектов.

Обсуждая вопросы сегментации приходится согласиться с выводом, сделанным в обзоре [369], что не существует и не может существовать какой-то общий и универсальный метод сегментации, который одинаково хорошо работал бы применительно ко всем задачам. Связано это с тем, что проблема сегментации изображения сложна по той причине, что является некорректной. Согласно определению, задача в математике считается корректно поставленной, если выполняются три условия: 1) ее решение существует в некотором заданном метрическом пространстве решений, 2) решение единственно, и 3) решение непрерывно зависит от данных на этом пространстве (т.е. устойчиво по отношению к допустимым возмущениям исходных данных). Обращаясь к проблеме сегментации изображений можно ожидать выполнение первого критерия, т.е. существование решения. Однако в задаче сегментации, которая могла бы быть сформулирована ограниченным образом, оставшиеся два критерия обычно не будут выполняться. Поэтому при возникновении новой задачи обработки (сегментации) изображений, важно тщательно оценивать различные доступные методы и выбирать тот, который решает задачу наилучшим образом.

5.1. Сегментация изображений на основе прослеживания контуров

5.1.1. Особенности задачи сегментации снимков микросхем

При разработке электронных микросхем требуется контроль качества изготавливаемых пластин, для чего производится анализ их поверхности с помощью снимков, формируемых сканирующим электронным микроскопом (СЭМ). Задачей является разработка алгоритма сегментации, т.е. разбиения изображений участков микросхем, получаемых СЭМ, на связанные области, при котором осуществляется присвоение каждому элементу изображения индекса, соответствующего тому или иному объекту анализируемой поверхности.

Получаемое изображение является монохромным. Оно содержит множество областей (паттернов), образов тех или иных деталей микросхемы, разделенных контурными границами. Внутри паттернов значения меняются слабо, но характеризуются наличием шума достаточно высокого уровня. Пример исходного снимка, получаемого электронным микроскопом, показан на Рис. 5.1.1. Изображения подобного типа хорошо соответствуют кусочно-постоянному варианту двухмасштабной многокомпонентной модели, изложенной в Главе 1.

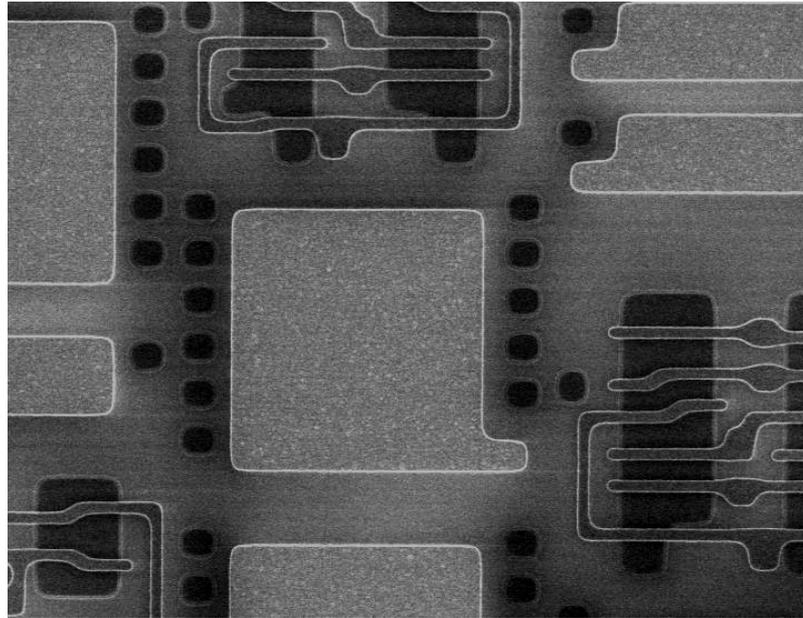


Рис. 5.1.1. Исходное изображение поверхности микросхемы.

Для сегментации был выбран подход, основанный на обнаружении и прослеживании замкнутых контурных линий на границах паттернов, построении бинарной контурной карты сегментов и последующей индексации элементов внутри полученных областей. Основным здесь является формирование корректной карты контуров. Особенность рассматриваемой задачи состоит в том, что на исходных изображениях наблюдаются два различных типа границ между паттернами: границы в виде разделяющих контрастных светлых линий (назовем их границами первого типа), и границы в виде яркостного перепада между паттернами, как это и бывает между областями обычного изображения (границы второго типа). Во многих случаях на изображении присутствуют границы обоих типов одновременно.

Два типа границ объектов требуют различных подходов к построению алгоритмов их обнаружения и прослеживания. Чтобы не разрабатывать два самостоятельных алгоритма, было решено применить операцию декомпозиции исходного изображения (раздел 2.2), применив которую удастся преобразовать границы первого типа и границы второго типа к одному виду; это позволяет использовать один и тот же алгоритм прослеживания контуров. Получаемые

промежуточные карты контуров затем подвергаются операциям объединения и корректировки, в результате чего формируется финальная карта контуров [195,240].

5.1.2. Получение выровненных изображений

Задача данного этапа — формирование сигнала (изображения), отражающего вероятность нахождения контура в каждой точке исходного снимка. На таком изображении внутренние области паттернов должны иметь близкие к нулю значения, а границы между паттернами — близкие к максимальному. Будем называть его *выровненным по нулю изображением* или просто *выровненным изображением*. Поскольку, как сказано выше, возможны два типа границ, то потребуется сформировать два изображения разных типов.

Рассмотрим границы первого типа — разделяющие контрастные светлые линии. Их можно представить как тонкие объекты, что позволяет для их выделения воспользоваться подходом, рассмотренным в разделе 4.1 для обнаружения объектов малой локальной площади, и основанном на алгоритме декомпозиции «D». Согласно модели (1.2.24) и алгоритму (4.1.2), сигнал с границами первого типа можно получить как разность

$$t_{mn} = x_{mn} - S_{mn}, \quad (5.1.1)$$

где x_{mn} — значение элемента исходного изображения, S_{mn} — сглаженное значение, получаемое алгоритмом декомпозиции «D» (2.2.7)–(2.2.8), а t_{mn} — компонента с контрастными контурными линиями. Отметим, что значение n^W в (2.2.8) при этом следует выбирать исходя из соотношения $n^W > dL$, где L — размер используемого фрагмента, а d — максимально возможная ширина контрастной линии между паттернами. Получаемое выровненное изображение с откликами контуров первого типа показано на Рис. 5.1.2.

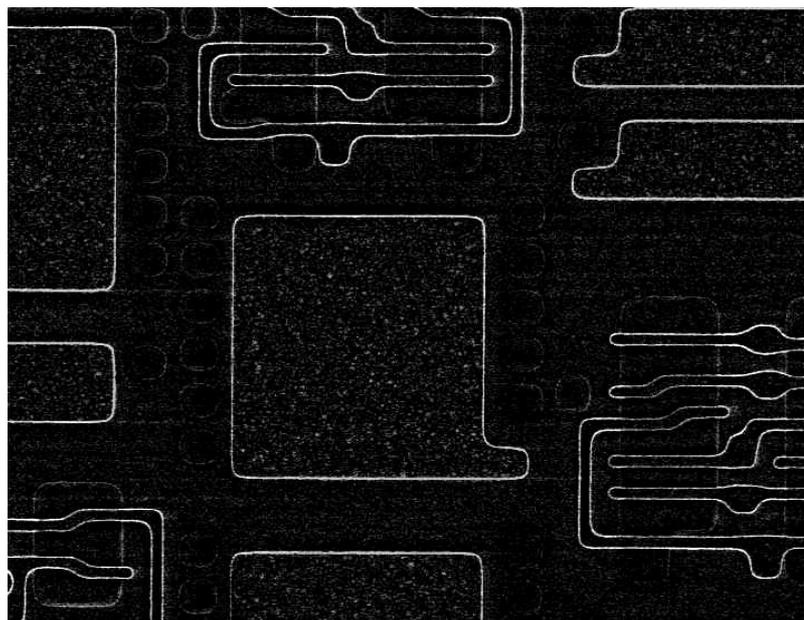


Рис. 5.1.2. Выровненное изображение с контурными откликами первого типа.

Изображение с границами второго типа, т.е. границами между протяженными объектами, можно получить согласно подходу к обнаружению контуров, рассмотренному в разделе 4.3. Для этого нужно использовать ту же самую сглаженную компоненту S_{mn} , к которой следует применить один из контурных операторов. Нами был использован вариант градиентного фильтра, вычисляемый как величина размаха значений S_{mn} в окрестности v_{mn} точки (m,n) , состоящей из 9 (3×3) элементов:

$$t_{mn} = \max \{x_{ij} \in v_{mn}\} - \min \{x_{ij} \in v_{mn}\}. \quad (5.1.2)$$

Выровненное изображение с контурными откликами второго типа показано на Рис. 5.1.3.

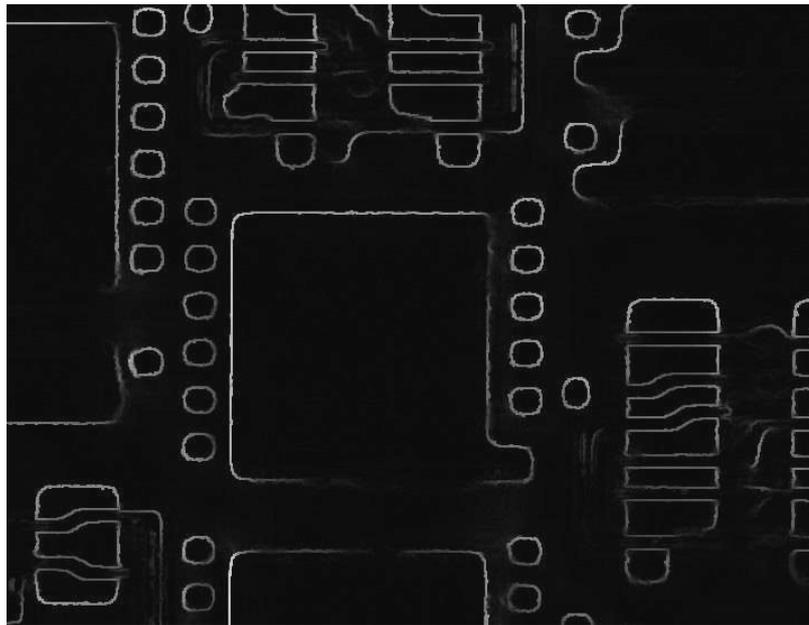


Рис. 5.1.3. Выровненное изображение с контурными откликами второго типа.

Представляемое очевидным возможное сложение полученных выровненных изображений с контурными откликами первого и второго типа оказывается не слишком удачным вариантом. Причина в том, что существует ряд границ, находящихся между объектами различающейся яркости, которые также имеют контрастные разделяющие линии. Такие границы относятся к каждому из двух типов, а значит проявляются на обоих выровненных изображениях. В силу разницы способов обнаружения, получаемые множества контурных точек хотя и располагаются вблизи друг от друга, однако не совпадают. Это приводит к увеличению ширины контурных участков и, как следствие, к последующим неточностям проведения контурной линии.

Таким образом границы первого и второго типов приведены к единому виду, что позволяет в дальнейшем использовать единый алгоритм прослеживания контуров. Обе карты контуров затем объединяются в одно контурное изображение.

5.1.3. Алгоритм прослеживания контуров

Прослеживание контуров по выровненному изображению и состоит из следующих этапов:

1. Формирование поля направлений контуров.
2. Определение контурного порога и отыскание начальных точек прослеживания контурных линий изображении.
3. Проведение контурных линий.

Формирование поля направлений контуров

При прослеживании границ требуется информация о линии направления контуров, для чего предварительно строится *поле направлений контуров*, содержащее информацию о направлении контурной линии в каждой точке изображения. Известно, что максимальная изменчивость значений элементов наблюдается в направлении перпендикулярном контуру, а минимальная — вдоль контура, поэтому направление контурного перепада определяется как то, на котором достигается максимум вариации значений пикселей. Для анализа были приняты достаточными 4 направления: 0, 45, 90 и 135 градусов.

Выберем некоторое направление анализа p . Пусть x_- , x и x_+ — значения, соответствующие трем точкам вдоль направления p , подсчитанные как средние по окрестностям размерами 5×1 пикселей, вытянутым по направлению, перпендикулярному по отношению к p . При этом окрестность, соответствующая значению x , проходит через анализируемую точку (m, n) , а окрестности x_- и x_+ отстоят от нее на расстояние $\pm \delta$. Найдем функционалы по направлениям p , являющиеся, по сути, аналогами второй производной:

$$f_p(m, n) = x - (x_- + x_+)/2. \quad (5.1.3)$$

Вычисляются значения функционала $f_p(m, n)$ по четырем направлениям: $f_0(m, n)$, $f_{45}(m, n)$, $f_{90}(m, n)$ и $f_{135}(m, n)$ по выровненному изображению. Значение, на котором достигается максимум $f_p(m, n)$, указывает направление контурного перепада в точке, сама же контурная линия ориентирована в перпендикулярном направлении. Отметим, что разность $(\max f_p(m, n) - \min f_p(m, n))$ отражает степень «остроты» контура (эксцесс). Полученные значения p_{mn} , $p = \{0, 45, 90, 135\}$ составляют поле направлений, используемое для прослеживания контурных линий.

Нахождение контурного порога

Прослеживание контура может начинаться с любой принадлежащей ему точки, поэтому нет необходимости обнаруживать все или большинство точек контурной линии, а достаточно хотя бы одной *начальной точки* для прослеживания всего контура. Множество начальных точек находится сравнением значений элементов выровненного изображения с *контурным порогом*. Поскольку уровень шума на изображениях сильно варьируется — среднеквадратическое

отклонение шума может быть от 6-10 градаций на изображении с контрастными деталями до 25-35 градаций на изображении без деталей, то подобрать единый порог невозможно, и его выбор осуществляется автоматически на основе анализа получаемого выровненного изображения, используя следующие его свойства.

На Рис. 5.1.4 в логарифмическом масштабе показаны гистограммы двух выровненных изображений; первое из них (а) содержит детали, а второе (б) не содержит. Хорошо видно, что график (а) имеет характерную точку перегиба в районе значения яркости 30 — точка 't'. На втором же графике подобной точки не имеется. Согласно модели изображения, рассмотренной в первой главе (см. раздел 1.2.2 и Рис. 1.1.1.), такой перегиб возникает в случае, когда наблюдается смесь двух случайных процессов со значительно различающимися распределениями (дисперсиями). В нашем случае это смесь значений элементов на ровных участках фона (с малой дисперсией) и на контурных линиях (с большой дисперсией).

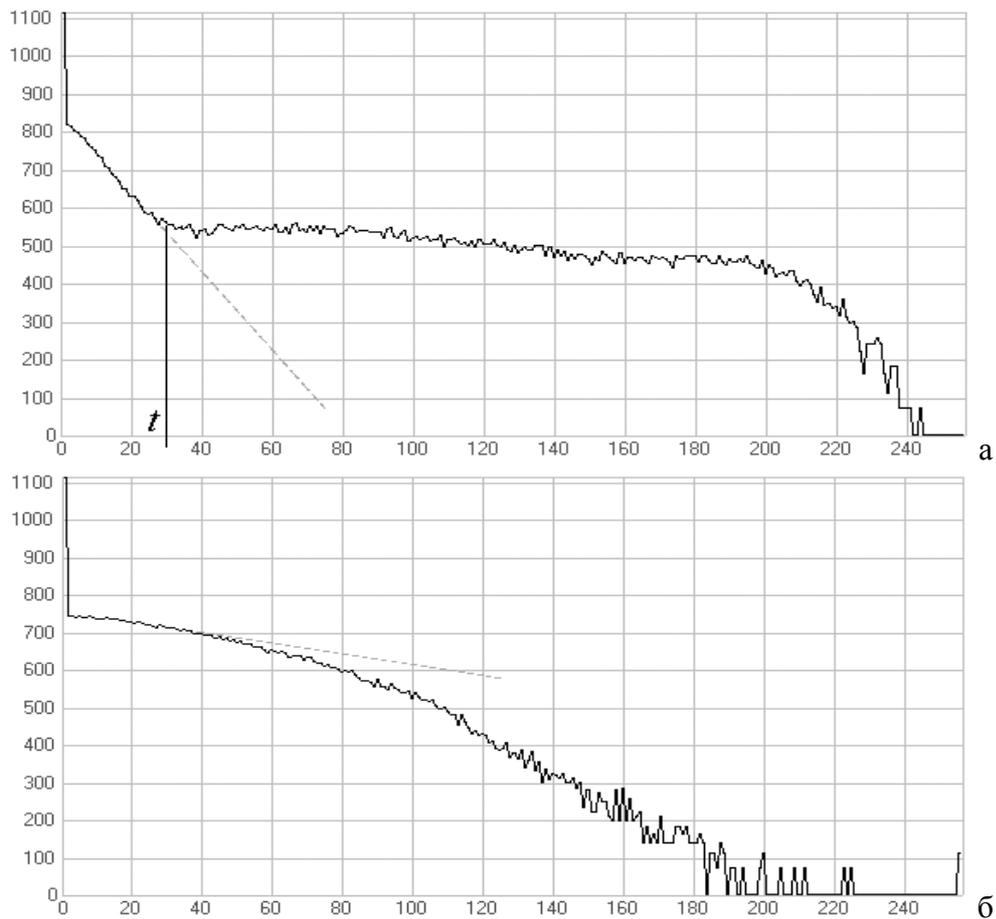


Рис. 5.1.4. Гистограммы изображений с деталями (а), и без (б) в логарифмическом масштабе.

Точка перегиба графика отыскивается как точка, где график существенно отклоняется в сторону больших значений от прямой линии (показана серым пунктиром), аппроксимирующей точку графика в области малых яркостей. Найденная величина t используется как значение контурного порога. Результатом порогового обнаружения является двоичное изображение, в

котором ненулевые точки, где $x_{mn} > t$, суть кандидаты на начальные точки прослеживания контурных линий.

Выбор направления линии контура

Направление контура выбирается по выровненному изображению вдоль гребня линии. Для этого окрестность точки прослеживания с радиусом 5 элементов делится на 8 секторов по 45 градусов с биссектрисами под углами $45n$ ($n = 0, \dots, 7$) (Рис. 5.1.5); т.е. каждый сектор заметает область с углами $45n \pm 25,5$ градусов. Вычисляются средние значения элементов внутри трех секторов: сектора, направление которого совпадает с текущим направлением контура, а также двух соседних. Для повышения устойчивости прямых линий, среднее значение в секторе, совпадающим по углу с текущим направлением контура, умножается на коэффициент 1,2. Направление, на котором достигается максимум, выбирается как новое направление контура. В начальной точке прослеживания линии в качестве направления контура выбирается то, которое записано в поле направлений. На основании этого определяются предварительные координаты следующей точки контура (точки прослеживания).

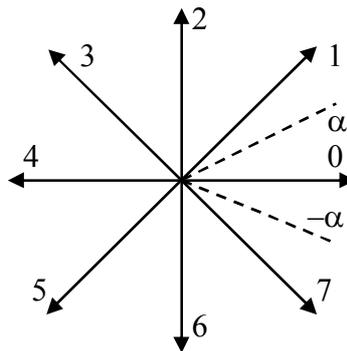


Рис. 5.1.5. Направления контура и сектор анализа $[-\alpha, \alpha]$.

Ширина линии контурного отклика и координаты следующей точки контура

На выровненных изображениях, содержащих контурную информацию реальных снимков, ширина линии контурного отклика варьируется в пределах от 1 до 10 пикселей и может вдвое превышать выбранные размеры секторов анализа. Поэтому найденное на предыдущем шаге направление контура еще не принимается как окончательное, а производится уточнение координат точки прослеживания путем отыскания точки середины линии.

Рассмотрим профиль выровненного по нулю изображения вдоль перпендикуляра к направлению линии контурного отклика (по секущей). Для контуров второго типа форма такого сигнала близка к колоколу с максимумом в точке середины гребня, а для контуров первого типа может меняться от колоколообразного в случае узкой разделяющей линии, до плато с плавно спадающими краями в случае широкой разделяющей линии. Обозначим набор точек, располо-

женных вдоль такой секущей, как $x(n)$, $-k \leq n \leq k$; $k \geq 2w$, где $w = 2u + 1$ — предсказываемая ширина линии. Возьмем функцию $h_u(n)$ с прямоугольным всплеском шириной $2u + 1$:

$$h_u(n) = 1, \text{ если } -u \leq n \leq u;$$

$$h_u(n) = -1, \text{ если } n < -u \text{ или } n > u.$$

Рассмотрим функционал:

$$f(u, v) = \frac{1}{2u + 1} \sum_{n=-k}^k x(v+n)h_u(v+n). \quad (5.1.4)$$

Максимум этого функционала (по всем допустимым u и v) достигается в точке (u', v') , где v' — положение центра, а u' — полуширина гребня. Перевод точки оптимума v' в координаты изображения дает координаты (i_{n+1}, j_{n+1}) следующей точки контура, а значение $w = u'$ сохраняется и используется в дальнейшем.

Проведение контурной линии к следующей точке прослеживания

Проводится сегмент контурной линии из предыдущей точки контура (i_n, j_n) в следующую точку прослеживания (i_{n+1}, j_{n+1}) . Корректная контурная линия должна иметь толщину в один пиксель и проходить через границы вертикального/горизонтального соседства (4-соседства) пикселей, как это показано на Рис. 5.1.6,а.

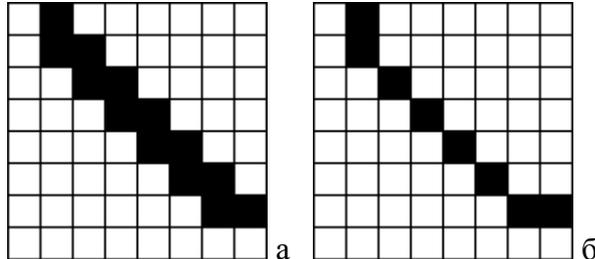


Рис. 5.1.6. Корректный (а) и некорректный (б) варианты проведения контурной линии.

Такой вариант проведения контурной линии гарантирует, что соседние области, разделенные контуром, не будут иметь диагональных контактов пикселей. Процесс проведения контура прекращается, если происходит касание другого контура или границы изображения.

Как отмечалось выше, набор начальных точек — кандидатов прослеживания контурных линий — включает значительно больше точек, чем необходимо. Проводимый контур может пройти вблизи таких точек, а не через них. Поэтому в процессе проведения контура необходимо обеспечить обнуление некоторого «коридора» вблизи проводимого контура и удалить лишние начальные точки. Ширина такого коридора d выбирается в зависимости от локальной ширины линии контурного отклика: $d = 2w + 4$, где w — ширина линии, определяемая согласно максимуму функционала (5.1.4).

Если контурная линия не коснулась границ изображения, другой линии или самой себя, алгоритм заменяет координаты терминальной точки (i_n, j_n) новыми (i_{n+1}, j_{n+1}) и продолжает процедуру прослеживания линии. Если же касание произошло, процесс прослеживания линии прекращается и осуществляется переход к следующей начальной точке-кандидату и прослеживанию новой линии. Процедура продолжается до тех пор, пока остается хотя бы одна начальная точка. В результате окончания полного цикла формируется соответствующая карта контуров первого или второго типа.

Параметры линии контурного отклика

При последующем анализе корректности контурных линий требуются некоторые измеряемые параметры. Наиболее удобно их измерять на этапе прослеживания контурных линий. Такими параметрами являются оценки, близкие статистическим параметрам асимметрии и эксцесса профиля линии контурного отклика, которые удобно получать во время вычисления функционала (5.1.4).

Пусть $x(n)$ — элементы исходного изображения, лежащие на прямой, перпендикулярной контуру в некоторой точке. Пусть также $2u + 1$ — ширина линии контурного отклика, а v — размеры левого и правого «хвостов». Вычислим следующие суммы:

$$S_{\text{left}} = x(-v - u) + \dots + x(-u - 1);$$

$$S_{\text{center}} = x(-u) + \dots + x(u);$$

$$S_{\text{right}} = x(u + 1) + \dots + x(u + v).$$

Асимметрией профиля A назовем оценку

$$A = |S_{\text{right}} - S_{\text{left}}|/v, \quad (5.1.5)$$

а эксцессом профиля E — оценку

$$C = S_{\text{center}}/(2u + 1) - (S_{\text{right}} + S_{\text{left}})/2v. \quad (5.1.6)$$

Средние значения асимметрии и эксцесса профиля линии будут равны:

$$A_{\text{aver}} = (A_1 + \dots + A_N)/N, \quad (5.1.7)$$

$$E_{\text{aver}} = (E_1 + \dots + E_N)/N. \quad (5.1.8)$$

Здесь N — число точек в линии. Средние значения асимметрии и эксцесса профиля линии используются по окончании процесса прослеживания линий для оценивания корректности проблемных линий.

5.1.4. Формирование объединенной карты контуров

Объединение карт контуров первого и второго типов имеет некоторые особенности. Выше уже говорилось о существовании границ, находящихся между объектами различающейся яркости, которые также имеют контрастные разделяющие линии. Такие границы относятся к

каждому из типов и появляются на картах контуров обоих типов. Однако контурные линии, соответствующие одной и той же границе, на картах разных типов не совпадают: на карте границ первого типа контурные линии проводятся по центру светлых линий, а на карте границ второго типа — фактически по краю такой линии.

Поэтому при объединении промежуточных карт контуров, полученных прослеживанием границ первого типа X_1 и второго типа X_2 производится удаление паразитных линий на карте границ второго типа вблизи контурных линий, имеющих на карте границ первого типа. Для этого используется уже найденное выше значение ширины «коридора» обнуления $d = 2w + 4$ и выполняется следующая операция:

$$x_2(m, n) = \begin{cases} 0, & \text{если точка } (m, n) \text{ находится в } d\text{-окрестности одной из линий } x_1, \\ x_2(m, n), & \text{в противном случае.} \end{cases}$$

По окончании процедуры обнуления производится слияние карт X_1 и X_2 и формируется объединенное контурное изображение X :

$$x_{mn} = \begin{cases} 1, & \text{если } x_1(m, n) \vee x_2(m, n) = 1, \\ 0, & \text{в противном случае.} \end{cases} \quad (5.1.9)$$

Условия корректности карты контуров предусматривают, что каждая сегментированная область должна быть окружена непрерывной и замкнутой контурной линией. Из-за наличия шумов и особенностей самого алгоритма полученная объединенная карта контуров может не удовлетворять полностью этим требованиям (так, использование процедуры удаления паразитных линий на карте X_2 может привести к разрыву линий на объединенной карте). Некоторые контурные линии оказываются незамкнутыми или «оборванными», возможно также появление ложных «петель». Такие линии должны быть либо продолжены и соединены с той или иной контурной линией, либо удалены. Коррекция осуществляется продлением оборванных линий и последующей обработкой карты контуров.

Продление и замыкание разорванных линий

Данная процедура осуществляется вышеописанным алгоритмом прослеживания контуров, начиная с точек разрыва и до замыкания на тот или иной контур. Для анализа используется синтезированное изображение T , получаемое взвешенным суммированием выровненных изображений с границами первого (T_1) и второго (T_2) типов, а также карты K уже прослеженных контуров:

$$t_{mn} = at_1(m, n) + bt_2(m, n) + ck(m, n). \quad (5.1.10)$$

Здесь t_1 и t_2 — элементы выровненных изображений с контурами первого и второго типов, k — элемент объединенной карты контуров, a , b , c — весовые коэффициенты, а t_{mn} — элемент формируемого изображения.

5.1.5. Обработка карты контуров и формирование результата

В силу сравнительно высокой зашумленности исходного изображения, на карте контуров возможно появление ложных линий. Иногда происходит замыкание линии на саму себя с образованием на конце небольшой петли или сплошной «капли». В некоторых точках (как правило, на участках соединения линий) ширина контурной линии может оказаться больше, чем один элемент. Такие некорректные участки и линии должны быть соответствующим образом обработаны или удалены. Обработка карты контуров включает утончение контурных линий, нумерацию сегментов и удаление ложных линий и сегментов.

Утончение контурных линий

Для каждого элемента контурной линии анализируется окрестность из 3×3 элементов. Если рассматриваемый элемент граничит с элементами ограничиваемой области и толщина линии больше, чем 1 пиксель, то такой элемент присоединяется к элементам области. Алгоритм состоит из следующих шагов:

1. Выбирается очередной контурный элемент. Из значений 8 окружающих элементов формируется циклически замкнутый массив. Он рассматривается как состоящий из некоторого числа подразделов, каждый из которых в свою очередь включает пару множеств: последовательность элементов, принадлежащих контурной линии, и последовательность элементов области. Заметим, что число таких подразделов не может превышать 4.
2. Подсчитывается число таких подразделов и общее число контурных элементов.
3. Если контурных элементов не обнаружено, то значит, центральный элемент является изолированной контурной точкой, которая должна быть удалена.
4. Если имеется только один подраздел и а) число контурных элементов больше чем 2, или б) число контурных элементов равно 2 и имеется контурный элемент непосредственно над или под обрабатываемым, то такой контурный элемент удаляется.
5. В остальных случаях центральный элемент остается без изменений.

Алгоритм применяется итеративно. Он позволяет утончить линии и привести «капли» на концах некоторых линий в статус «оборванных» линий. После этого становится возможным использование общих операций по обнаружению оборванных линий и последующему их удалению или продлению до касания других контурных линий.

Нумерация сегментов

Нумерация сегментов осуществляется прогрессивным сканированием полученного изображения карты контуров (слева направо и сверху вниз) с присвоением каждому элементу области некоторого номера. Производится проверка и нумерация образовавшихся сегментов (внутренних областей, ограниченных контурами) и удаляются оставшиеся незамкнутыми «оборванные» контурные линии.

Алгоритм нумерации анализирует для каждого внутреннего элемента области (не контурного) 4 предыдущих, уже обработанных соседних элемента: 3 сверху и один слева. При этом возможны следующие ситуации.

1. Среди предыдущих элементов имеются элементы области с одним и тем же номером сегмента. В этом случае текущий обрабатываемый элемент получает тот же номер.
2. Среди предыдущих элементов имеются элементы области с разными номерами сегментов. Текущему элементу присваивается минимальный из этих номеров, и в таблице списка сегментов появляются соответствующие ассоциативные связи, указывающие на совпадение одного сегментного номера с другим, т.е. означающие, что фрагменты сегментов с этими номерами составляют один большой сегмент.
3. Все предыдущие элементы являются контурными или граничными. В этом случае текущему элементу присваивается новый сегментный номер.

После такой нумерации различные части сегмента сложной формы (фрагменты) могут получить разные номера, а в сложных случаях может возникнуть цепочка ассоциаций. Полный список всех ассоциаций представляет собой *таблицу контактов и ассоциаций фрагментов*. Число входов таблицы совпадает с полным числом обнаруженных фрагментов. Значение в точке таблицы, соответствующей каждому номеру фрагмента, есть номер сегмента, с которым ассоциирован данный фрагмент; большие номера ассоциируются с меньшими. После упрощения и установления всех ассоциаций, данная таблица превращается в таблицу перенумераций, в каждой ячейке которой содержится новый номер, который должен быть приписан элементам данного фрагмента. При последующем сканировании номера элементов поправляются в соответствии с данной таблицей, и сегментация изображения становится корректной. Одновременно подсчитывается число пикселей в каждом сегменте, и определяются координаты прямоугольника, описанного вокруг сегмента, что используется для сокращения последующих вычислений.

Удаление ложных контурных линий и сегментов

Ложные контурные линии могут быть двух типов. Первый тип — «оборванные» линии — характеризуются не замкнутыми на другие контурные линии концами. Второй тип —

корректные с точки зрения замыкания линии, но лежащие целиком внутри одного сегмента. Для пояснения данного случая рассмотрим два замкнутых контура, соединенных третьей линией — «перемычкой»; топологически это выглядит как структура вида \bigcirc — \bigcirc . Поскольку она возникает внутри некоторой большой области, то элементы с обеих сторон перемычки имеют один и тот же номер. Это является признаком ложной линии.

Для удаления таких ложных линий алгоритм проверяет сегментные номера областей по обе стороны каждой контурной линии. Если они одинаковые, то линия считается ложной и удаляется. Это осуществляется простой заменой значений элементов контурной линии на значения окружающих элементов сегмента.

Одной из особенностей рассмотренного выше алгоритма прослеживания контуров является то, что он прекращает работу только в том случае, когда произошло касание другого контура, границы изображения, или возникло замыкание контура на самого себя. Как результат, из-за влияния шума в некоторых случаях на изображениях могут появиться ложные сегменты, очерченные формально корректными контурными границами. Для обнаружения и удаления подобных ложных сегментов используется следующий способ анализа.

Рассматриваются сегменты, целиком расположенные внутри некоторого большего, окружающего их сегмента. Осуществляется обход вдоль границы каждого сегмента, и по формулам (5.1.7) и (5.1.8) вычисляются средние значения асимметрии A_{aver} и выпуклости E_{aver} , которые затем сравниваются с экспериментально найденными порогами T_A и T_C . Если $E_{aver} < T_C$ или $A_{aver} < T_A$, то сегмент считается ложным и удаляется вместе со своей границей. В этом случае значения внутренних элементов сегмента и его границы изменяются на значение номера сегмента, их окружающего.

Формирование результата сегментации

Результатом сегментации является символическое представление исходного изображения, т.е. массив того же размера, в котором каждому элементу области приписан тот или иной номер сегмента. Способы визуализации могут быть различными; мы в данном случае выбрали вариант, при котором внутренним элементам областей присвоены значения, равные средней яркости по сегменту, а контуры проведены тонкими белыми линиями между ними. На Рис. 5.1.7 показаны результаты сегментации изображения, представленного на Рис. 5.1.1.

Тестирование алгоритма на серии из 758 снимков показало, что удовлетворяющие требуемым критериям точности разделения объектов результаты достигаются в 98,3% случаев. Алгоритм сегментации изображений на основе прослеживания контуров на снимках электронных микросхем был разработан по заказу компании MicroSpec Technologies Ltd. (Carl

Zeiss Group) для использования в процедуре диагностики при промышленном производстве интегральных микросхем.

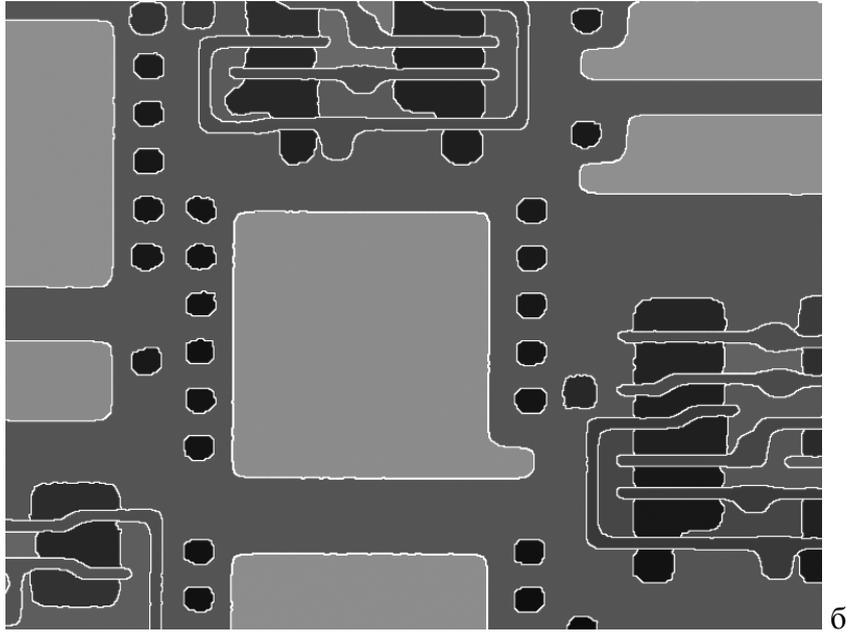


Рис. 5.1.7. Результат сегментации изображения на Рис. 5.1.1.

Изложенные в разделе 5.1 результаты опубликованы в работах [195,238,240].

5.2. Цвето-текстурная сегментация изображений

Цвето-текстурная сегментация означает разбиение изображения на множество связных областей, отличающихся от соседствующих областей по своим яркостным, цветовым и/или текстурным характеристикам в соответствии с некоторыми выбранными критериями. Во введении к данной главе уже говорилось, что различных вариантов и подходов к решению задачи сегментации изображений имеется много. Как и для задач распознавания или дискриминации, важнейшим вопросом также является выбор (построение) пространства признаков, в которое будет отображена информация о наиболее существенных различиях объектов наблюдаемой сцены. Каждой точке изображения в таком пространстве соответствует некоторый вектор. Существенное отличие задач сегментации изображений от многих задач распознавания в том, что видеоинформация представляет собой сигнал с сильными пространственными корреляционными связями, и этот факт необходимо использовать при разработке алгоритмов анализа.

Возможны два пути улучшения уже имеющихся результатов сегментации: первый состоит в нахождении более соответствующих задаче признаков, а второй — в разработке алгоритма, более глубоко использующего свойства изображений. В настоящем исследовании преследуются оба подхода. В части выбора признаков исследуется комбинация яркостных, цветовых и

текстурных пространств признаков с выбором метрики в обобщенном пространстве. Для построения алгоритма сегментации используется иерархический подход, основанный на анализе расстояний в пространстве признаков, преобразовании изображения в граф первичных кластеров (для этого используется пирамидальный алгоритм) и последующем упрощении карты первичных кластеров. В части учета свойств изображений упор сделан на использовании двухмасштабной многокомпонентной модели изображения, рассмотренной в первой главе, и методов декомпозиции изображения, позволяющих выделять текстурную компоненту (раздел 2.2), которая согласно модели представляется разностью исходного изображения и получаемой в результате декомпозиции сглаженной яркостно-цветовой компоненты.

Признаки *яркость* и *цвет* являются измеримыми и удобными для классификации [324,268], в то время как понятие *текстура* является достаточно расплывчатым и нечетко формулируемым [310]. Обобщенно говоря, под текстурой понимают смесь большого числа в той или иной степени упорядоченных и похожих элементов или образов (примитивов), когда ни один из этих конкретных образов не имеет принципиального значения, и когда интерес представляет только общий вид и эффект, производимый достаточно большими областями. При описании текстур обычно используются следующие характеристики: регулярность, направленность, крупность, повторяемость (эффективный период или размер), форма, цвет и контраст примитивов, и некоторые другие [268,325]. Одни из упомянутых характеристик могут быть четко сформулированы, тогда как другие — нет.

Проблема текстурного анализа исследуется много лет, и основная трудность заключается в отыскании подходящих характеристик описания текстуры; основными являются три подхода: структурный, спектральный и статистический. Структурный основан на предварительном выделении примитивов и последующей оценке их размещения, что эффективно при анализе регулярных текстур. Спектральные методы используют свойства Фурье-спектра и подходят для отыскания периодичности в изображении. Более универсальным является статистический подход. В качестве статистических характеристик чаще всего используют матрицу смежности, автокорреляцию, разности уровней серого, длины серий по направлениям, локальную частоту флуктуаций яркости [220,268,269,325,361]. Иногда применяется иерархический подход, позволяющий оценивать характеристики на разных масштабах. После 1980-х годов возник интерес к многоканальному и многомасштабному [339] текстурному анализу, а позже — к использованию вейвлетов [225,331] и Габоровских фильтров [254,287,331,366], в свою очередь тесно связанных с пирамидальным представлением изображения [331].

5.2.1. Особенности задачи сегментации изображений

Сегментацию изображения можно переформулировать как задачу разделения двумерного сигнала на связные области (гладкие в смысле анализируемых характеристик), отличающиеся от соседствующих областей по своим яркостным, цветовым и/или текстурным характеристикам. Можно отметить следующие принципиальные отличия задачи сегментации изображений общего вида от задачи классификации или распознавания.

1. Изображения являются двумерным сигналом и отличаются высокой пространственной корреляцией характеристик соседствующих элементов.
2. Области на изображениях могут иметь достаточно большие размеры и на их протяжении значения характеристик областей могут заметно изменяться.
3. Возможное число классов объектов на изображениях, как правило, заранее не известно.
4. При сегментации изображения не ставится задача объединения в один класс близких по характеристикам, но не соседствующих между собой областей.

Из-за указанных особенностей элементы одного и того же объекта изображения в пространстве признаков могут отображаться в достаточно протяженную область сложной формы. В свою очередь области, соответствующие различным объектам, могут оказаться близкими или даже пересекаться. При поэлементном отображении также теряется информация о пространственном соседстве точек изображения. Все это свидетельствует, что сегментация изображения путем решения классической задачи распознавания имеет ограниченное применение.

Удовлетворить указанным особенностям должен подход, использующий критерии близости как выбранных признаков, так и пространственного расположения элементов изображения. В качестве такового был принят иерархический подход, основанный на пирамидальном алгоритме [185], позволяющем сравнивать признаки пространственно соседствующих элементов и объединять те, которые в этом смысле оказываются близки. Согласно данному алгоритму, изображение представляется в виде графа, на нижнем уровне которого расположены элементы изображения, а ветви соответствуют отдельным объектам. Ключевым является выбор оценки близости соседних вершин графа. В применении к рассматриваемой задаче это означает выбор пространства признаков и способа оценки близости точек в этом пространстве. Сопоставлять вершины графа, соответствующие соседним элементам и/или сегментам изображения, удобно при помощи расстояния в пространстве признаков, которое и будет являться критерием при принятии решения об их слиянии или же проведении между ними границы.

Начальным шагом является выбор того пространства, в котором будет решаться задача сегментации: в пространстве признаков либо в пространстве координат изображения.

В первом случае производится отображение элементов исходного изображения в пространство признаков (вообще говоря многомерное), в котором уже осуществляется анализ получаемого распределения вероятностей векторов признаков, отыскание областей с повышенной плотностью вероятностей, в том числе обладающих определенными свойствами, кластеризация на известное или неизвестное заранее число признаков, и присваивание точкам изображения идентификатора соответствующего кластера. Иногда при таком подходе, опираясь на пространственные свойства изображений, неявно предполагается, что если однородные области изображения отображаются в компактные образования в пространстве признаков, то и наоборот, компактные области в пространстве признаков отображаются в некоторое множество связанных областей изображения. Однако если не предпринимать специальных мер, то информация о пространственной корреляции элементов изображения теряется. Данный подход используют в случаях, когда множество возможных объектов и условий формирования изображения хорошо известны, и когда на многие параметры (в том числе и на допустимое число классов) удастся наложить сравнительно жесткие условия. Однако для получения достаточно четких, так сказать «хороших», границ объектов в данном подходе все равно требуется привлечение пространственной информации.

Во втором случае осуществляется поиск решения в пространстве координат исходного двумерного изображения. Такой подход обычно допускает проведение границ между возможными областями, анализ получаемой первоначальной карты областей, и последующее уточнение их границ (например, путем слияния или деления областей). При этом удастся избежать многих ограничений, но оборотной стороной такого подхода является то, что число возможных классов объектов становится равным или близким общему числу обнаруживаемых объектов. По этой причине в некоторых задачах, связанных с классификацией, может потребоваться дополнительный анализ в пространстве выбранных признаков.

В настоящем исследовании выбор сделан в пользу второго подхода — разбиения изображения на однородные области путем анализа сходства в пространстве признаков. Поскольку задача ставится как цвето-текстурная сегментация, очевидно, что каждому элементу изображения можно приписать не менее четырех признаков (3 яркостно-цветовых и 1 текстурный); тем самым пространство признаков должно быть многокомпонентным. В качестве критерия близости элементов выбрано расстояние между их образами в пространстве признаков. Отметим, что хотя при таком подходе и используется пространство признаков, но в отличие от классического решения задачи кластеризации анализируется не картина распределения, а лишь оценивается близость признаков точек. Тем самым одним из важнейших вопросов становится выбор метрики в этом пространстве.

5.2.2. Построение пространства признаков

Требуется отыскать такое пространство признаков, которое бы адекватно отображало необходимые признаки и в котором было бы возможно выбрать удобную и не слишком сложную меру расстояния. Для этого было принято решение строить полное пространство признаков как объединение нескольких более простых пространств (подпространств): яркостного, цветового и текстурного.

Яркостное пространство является наиболее простым. Оно одномерно, и на всем диапазоне яркостей можно предполагать постоянство контрастной чувствительности [40,113]. В такой ситуации удачной метрикой является модуль разности яркостей: $|B_x - B_y|$.

С выбором *цветового* пространства ситуация существенно сложнее. Предлагается много различных цветовых представлений, таких как XYZ , $U^*V^*W^*$, SOW^* , $L^*a^*b^*$, $L^*u^*v^*$ и др. [113,333]. Тестирование метрик в ряде указанных пространств показало, что при их использовании достигаются близкие результаты сегментации. Было решено использовать пространство **BHS**, в котором хорошо разделяются яркостная и цветовые составляющие.

В качестве *текстурного* пространства в настоящей работе использован набор многомасштабных оценок изменчивости яркости по множеству направлений, которые вычисляются для каждой точки изображения на основе значений локальной вариации и локальной осцилляции (изменения знака) градиента вдоль выбранных направлений. Такой подход существенно развивает метод разностей яркостей, обсуждаемый в обзорах [224,361], и близок идее оценки функций при помощи вариаций.

Пространство **BHS** отличается от обычного цветового пространства **RGB** координатным преобразованием. По сути, координаты B , H и S являются полярными координатами пространства **RGB**. Пусть r, g, b — координаты точки в исходном пространстве **RGB** и пусть $u = \min(r, g, b)$. Тогда

$$\begin{aligned}
 B &= \sqrt{(r^2 + g^2 + b^2)}/3, \\
 H &= \begin{cases} (2\pi/3)(b-u)/(g+b-2u) + \pi/3, & \text{если } u = r, \\ (2\pi/3)(r-u)/(b+r-2u) + \pi, & \text{если } u = g, \\ (2\pi/3)(g-u)/(r+g-2u) + (5\pi/3), & \text{если } u = b. \end{cases} \\
 S &= 1 - u / (r + g + b).
 \end{aligned} \tag{5.2.1}$$

Подпространства **B** и **HS** составляют яркостное и цветовое пространства признаков.

Текстурные характеристики оцениваются по множеству значений яркости B текстурной компоненты τ из (1.2.24), получаемой в результате операции декомпозиции (2.2.8). Проводится анализ профиля на локальном участке (число подъемов-спусков). С одной стороны это подобие частотного анализа, с другой — аналог первой вариации Кронрода [75], предложенной для

оценки двумерных сигналов и характеризующей число «выступов». Применимость такой характеристики к реальным изображениям была проанализирована в разделе 1.4.

Значения текстурных признаков вычисляются следующим образом. Рассмотрим отрезок длиной $2L$, центрированный в анализируемой точке и проходящий под углом $\alpha_k = k\pi/K$, где $(k = 0, \dots, K-1)$. Расположим на нем равномерно $2L+1$ точек (отсчетов), и каждой точке i отрезка $(-L \leq i \leq L)$ присвоим значение x_i , равное значению ближайшего элемента изображения. Пусть $d_i = x_{i+1} - x_i$ есть разность между соседними точками на отрезке (другими словами, градиент в точке i вдоль направления отрезка). Выберем также спадающую к краям весовую функцию $q_i = \cos(i\pi/2L)/\sum \cos(i\pi/2L)$. Тогда взвешенные суммы положительных v^+ и отрицательных v^- перепадов вдоль отрезка будут равны

$$v^+ = \sum_{i=-L}^{L-1} q_i d_i \text{ по всем } d_i > 0, \quad \text{и} \quad v^- = -\sum_{i=-L}^{L-1} q_i d_i \text{ по всем } d_i < 0. \quad (5.2.2)$$

Анализируемый отрезок может попасть на контурную границу объектов или же участок с заметным яркостным трендом, что внесет существенные изменения в значения одной из сумм (5.2.2). В предположении, что на коротком отрезке длиной $2L+1$ имеется не более одного яркостного перепада, который необходимо подавить, убрать его влияние на значения текстурных откликов можно выбором наименьшего из значений v^+ и v^- . Таким образом, значение *локальной вариации* v_k по направлению k определяется как

$$v_k = \min(v^+, v^-). \quad (5.2.3)$$

Значение *локальной осцилляции градиента* g_k определяется как число перемен знака градиента d_i на отрезке $(-L \leq i \leq L)$ таких, что $|d_i| > \delta$, где δ — порог чувствительности. Для уменьшения влияния шума можно рекомендовать в качестве x_i брать значение среднего вдоль небольших отрезков, перпендикулярных основному отрезку анализа текстуры и центрированных в каждой из i $(-L \leq i \leq L)$ точек отрезка.

Произведение локальных вариации v_k и осцилляции g_k градиента $\tilde{t}_k = v_k g_k$ дает предварительное значение текстурного признака по направлению k . Для выравнивания признаков высоко- и малоcontrastных текстур используется, как это предлагается в [254], функция арктангенса:

$$t_k = \arctg(\alpha \tilde{t}_k). \quad (5.2.4)$$

Полученное значение t_k считается *текстурным признаком* по направлению k в соответствующей точке изображения. Из-за локальной изменчивости текстурных элементов значения t_k оказываются зашумленными; для повышения однородности можно рекомендовать

использование низкочастотной фильтрации какого-нибудь вида, например, вычисления медианы по небольшому локальному двумерному фрагменту размерами $h \times h$ точек.

Текстура является масштабно-зависимой характеристикой. Для различения текстур с разными масштабами необходимо выделение признаков, соответствующих различным пространственным частотам. Можно, например, пропорционально изменить размеры отрезка, взяв отсчеты с другим шагом и сохранив общее число точек на отрезке равным $(2L+1)$. Так, увеличив шаг в n раз, получим соответствующий сдвиг диапазона текстурной чувствительности в сторону более низких частот. При этом следует подавить влияние высоких частот при помощи низкочастотной фильтрации (например, путем сглаживания). Другим способом изменить диапазон частот можно при помощи предварительного масштабирования (сжатия) самого изображения в те же n раз при сохранении размеров отрезков анализа неизменными. С точки зрения количества вычислений такой подход представляется предпочтительным.

Таким образом, текстурные признаки вычисляются по формулам (5.2.2)–(5.2.4) для каждой точки изображения по всем направлениям $0, \dots, K-1$ и масштабам $1, \frac{1}{2}, \frac{1}{4}, \dots, 2^{-(N-1)}$, где K и N — общее число направлений и масштабных уровней. Тем самым, мы получаем набор масштабно-зависимых значений локальной вариации v_k^n и осцилляции g_k^n локального градиента. Текстурные признаки при этом будут равны

$$t_k^n = \arctg(\alpha v_k^n g_k^n). \quad (5.2.5)$$

Можно также оценить общий уровень текстуры в точке следующим образом:

$$T = \sum_{n=1}^N w^n \sum_{k=0}^{K-1} t_k^n; \quad (5.2.6)$$

здесь w^n есть весовой коэффициент масштабного уровня n .

Суммарное число текстурных признаков для каждой точки изображения равно $M \times N$. В реальности достаточными представляются $K = 4$ и $N = 4$, т.е. всего 16 формируемых и анализируемых текстурных признаков. Объединенный набор текстурных, яркостных и цветовых признаков для точки изображения составляет полный вектор признаков, а всё возможное многообразие таких векторов — пространство яркостно-цвето-текстурных признаков. Такую комбинацию яркостного, цветового и текстурного пространств можно назвать *объединенным пространством признаков*. Весьма существенным при таком подходе становится определение близости точек в пространстве признаков, т.е. расстояния между ними. Для удобства мы будем называть это введением метрики в пространстве признаков.

5.2.3. Метрика в пространстве признаков

Поскольку задачей сегментации изображения является выделение однородных областей, то важнейшую роль при ее решении играет вопрос «похожести» элементов и/или областей изображения, т.е. близости их признаков. Характеристики как элементов, так и областей могут быть отображены в одно и то же пространство признаков, поэтому вопрос измерения их близости может быть переформулирован как введение расстояния $D(x,y)$ в пространстве признаков. Для удобства будем называть его *метрикой*, хотя нам необходимо выполнение лишь первых двух аксиом из трех: $D(x,y) = 0$ если $x = y$, а также $D(x,y) = D(y,x)$. Т.е. термин «метрика» используется здесь несколько в редуцированном смысле — только как некоторый способ вычисления расстояния между парой точек в пространстве признаков.

Важнейшей частью иерархического алгоритма сегментации является принятие решения о слиянии соседних кластеров (вершин графа), что осуществляется сравнением близости их образов в пространстве признаков. Уже говорилось, что данная задача может быть сформулирована как определение расстояния (выбор метрики) в пространстве признаков. Универсальной метрики в объединенном пространстве признаков к сожалению не существует, таковой нет даже для цветового подпространства. Предлагаемая в настоящей работе метрика разработана специально для задачи сегментации.

Поскольку между яркостными, цветовыми и текстурными признаками имеется существенное различие, а также поскольку допустима редукция пространства признаков, было решено использовать отдельные метрики в каждом из подпространств, а затем объединить их в одну общую. Отметим также, что выбор параметров нижеприведенной формулы вычисления расстояния (метрики) производился в применении к одному и тому же пирамидальному алгоритму сегментации [185,228].

Задача различения цветов и построения метрики в цветовом пространстве известна давно и обычно решается поиском равноконтрастной цветовой метрики [113,283,306,333]. Как уже говорилось, тестирование метрик в различных пространствах показало, что при их использовании в задаче сегментации достигаются близкие и во многих случаях приемлемые результаты. Нами было выбрано пространство **BHS**, поскольку в нем удастся хорошо разделить яркостную и цветовую составляющие, а также построить модифицированную метрику, которая лучше соответствует задаче сегментации и позволяет достичь более точных результатов в темных областях изображения. Метрики в каждом из подпространств задаются следующим образом.

Яркостная метрика $D_B(x,y)$ определяется как модуль разности яркостей B_x и B_y в двух точках x и y одномерного подпространства **B**:

$$D_B(x,y) = |B_x - B_y|. \quad (5.2.7)$$

Цветовая метрика задается следующей формулой:

$$D_C(x,y) = \left[w_H \varphi_B(x,y) \varphi_S(x,y) (\min(|H_x - H_y|, 2\pi - |H_x - H_y|))^2 + w_S \varphi_B(x,y) (S_x - S_y)^2 \right]^{1/2}. \quad (5.2.8)$$

Здесь H_i и S_i ($i = x; y$) суть значения компонент цветности и насыщенности соответственно, w_H и w_S — весовые коэффициенты; $\varphi_B(x,y) = F((B_x + B_y)/B_0)$, $\varphi_S(x,y) = F((S_x + S_y)/S_0)$, где B_0 и S_0 — значения уровней нечувствительности, а функция

$$F(z) = (2/\pi) \arctg(z) \quad (5.2.9)$$

служит ингибитором: $0 \leq F(z) < 1$. Тем самым обеспечивается подавление избыточного влияния цветности в области малых значений яркости и насыщенности, а также подавление насыщенности в области малых яркостей.

Текстурная метрика определяется формулой:

$$D_T(x,y) = \sum_{n=1}^N w^n \sum_{k=1}^K |t_k^n(x) - t_k^n(y)|, \quad (5.2.10)$$

где $t_k^n(x)$ и $t_k^n(y)$ — значения текстурных откликов уровня n и направления t в точках x и y , а w^n — весовой коэффициент уровня n .

Объединенная яркостно-цвето-текстурная метрика задается как взвешенная сумма расстояний в яркостном (5.2.7), цветовом (5.2.8) и текстурном (5.2.10) подпространствах:

$$D(x,y) = w_B D_B(x,y) + w_C D_C(x,y) + w_T D_T(x,y). \quad (5.2.11)$$

Здесь x и y — две точки объединенного подпространства, являющиеся отображениями точек или участков изображения, а w_B , w_C и w_T , — весовые коэффициенты отдельных метрик. Если входное изображение не полностью задействует пространство признаков, можно использовать часть подпространств: например, яркостное и текстурное в случае исходного монохромного изображения или яркостное и цветовое, если изображение заведомо не содержит текстуры.

5.2.4. Структура алгоритма сегментации

Используемый алгоритм содержит два уровня (стадии). На первом изображении при помощи пирамидального алгоритма разбивается на множество первичных мелких кластеров. На втором происходит сравнение их близости и слияние. Каких-либо принципиальных ограничений на пространство признаков при этом не накладывается. Это может быть одномерное пространство яркостей (или какого-то аналогичного параметра), обычное трехмерное цветовое RGB пространство, многомерное пространство условных цветов, пространство, расширенное дополнительными (например, текстурными) признаками, и пр. Такой подход позволяет

использовать единую алгоритмическую основу для применения к изображениям различного типа. Принципиальная структура алгоритма следующая.

Первоначально строится объединенное N -мерное пространство признаков, содержащее преобразованные цветовые и выделенные текстурные характеристики, и осуществляется переход к так называемому *векторному изображению*, которое является двумерным массивом N -мерных векторов в этом пространстве. Полученное векторное изображение обрабатывается пирамидальным алгоритмом, преобразующим его в первичный *граф кластеров*, в котором каждый элемент изображения оказывается соотнесенным тому или иному кластеру — соответствующему узлу графа. Тем самым, каждый узел графа отображает соответствующий кластер, а ребра графа — возможные связи между соседствующими кластерами. Необходимые характеристики кластеров и их связей с соседними кластерами объединяются в список кластеров, используемый позже, на второй стадии. Полученный на данной стадии граф кластеров представляет лишь предварительный результат сегментации, поскольку размеры первичных кластеров оказываются значительно меньше размеров объектов на изображении, а число их, соответственно, значительно больше числа интересующих объектов. В последнее время подобные первичные кластеры называют «суперпикселями».

На второй стадии анализируются и попарно сравниваются соседние кластеры, имеющие общие границы. Осуществляется слияние тех кластеров, расстояние между которыми в пространстве признаков достаточно мало, и происходит упрощение построенного графа. Поскольку слияние кластеров приводит к изменению их средних параметров, то этот процесс выполняется по итеративной схеме. Критерием окончания процедуры слияния служит отсутствие пар кластеров с расстоянием, меньше некоторого заданного порога. Результатом данной операции и является сегментированное изображение.

5.2.5. Первичная сегментация — пирамидальный алгоритм

Стадия первичной сегментации служит для преобразования исходного изображения в набор кластеров (подсегментов) небольшого размера. На данном этапе каждый объект изображения может еще отображаться не одним кластером, а набором из нескольких кластеров. Каждый сформированный кластер при этом должен характеризоваться как собственными параметрами, так и параметрами связи с соседними кластерами. Границы между объектами должны проходить по границам между кластерами.

Такую задачу удобно решать при помощи пирамидально-рекурсивного алгоритма [279,342]. Пирамидальный алгоритм обработки изображения требует задания способа вычисления уменьшенного изображения уровня $n+1$ на основе имеющегося изображения уровня n и

применения этой процедуры рекурсивно до предельного уменьшения размера изображения. Часто для этого выбирают элементарный домен некоторой формы, обеспечивающей плотное покрытие всей площади изображения, и задают способ нахождения значения элемента следующего уровня по значениям элементов домена предыдущего уровня. Рекурсивное применение данной процедуры позволяет построить дерево (граф), в котором каждый элемент изображения некоторого уровня (кроме самого нижнего) является узлом, связанным с элементами домена предыдущего уровня, а число нисходящих связей определяется размерами выбранного домена. В случае, если в качестве домена выбирается квадрат размерами 2×2 элемента, мы получаем так называемое *квадродерево* (Рис. 5.2.1). Оно и будет использоваться в дальнейшем.

Алгоритм пирамидальной сегментации осуществляет прямой (вверх) и обратный (вниз) проходы по квадродереву.

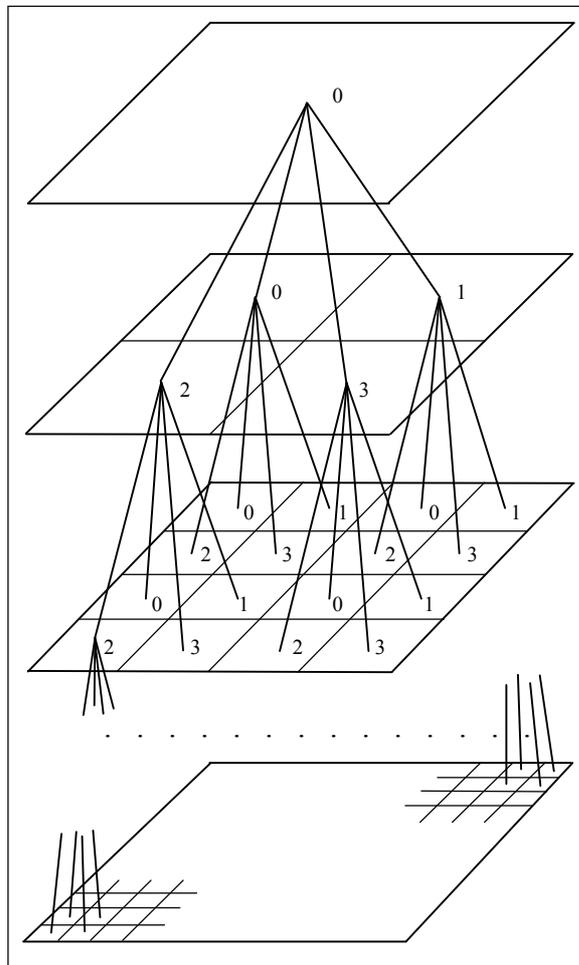


Рис. 5.2.1. Пирамидальная структура представления данных (квадродерево).

При *прямом проходе* вверх по квадродереву происходит рекурсивный анализ всех уровней пирамиды, начиная с самого нижнего (исходного изображения) и заканчивая верхним уровнем, состоящим из одного узла; одновременно с этим строится само квадродерево. На каждом шаге

на основе анализа четырех нижних узлов уровня n , составляющих один домен, создается узел уровня $n+1$ (см. Рис. 5.2.1). В новом узле запоминается информация об узлах, входящих в домен предыдущего уровня, соединенных с данным узлом, средних значениях по каждому из параметров, а также о прохождении контуров между узлами. Таким образом, каждый узел служит вершиной некоторого квадродерева, охватывающего расположенные под ним элементы изображения, и содержит информацию о поддеревьях предыдущего уровня. Для проведения контуров используется рассмотренный в разделе 4.3 подход, основанный на предварительной декомпозиции изображения на сглаженную и текстурную компоненты с последующим обнаружением контуров по сглаженной компоненте.

Основной задачей пирамидального этапа сегментации является объединение соседних элементов, имеющих близкие признаки и не разделенных контуром. Эта процедура требует прослеживания контурных линий на всех уровнях пирамиды. Считается, что два вертикально или горизонтально соседствующих элемента разделены контуром в том случае, если расстояние между их отображениями в пространстве признаков в смысле выбранной метрики (5.2.11) превышает некоторый заданный порог. Для второго и последующих уровней процедура обнаружения контуров также учитывает наличие контура на предыдущем уровне.

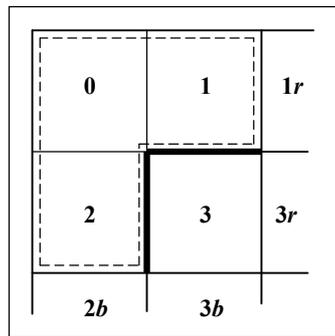


Рис. 5.2.2. Домен из четырех элементов и его соседние элементы справа и снизу.

Домен размерами 2×2 элементов показан на Рис. 5.2.2. Анализ и объединение его элементов состоит из следующих шагов.

1. Анализируется наличие контурных перепадов между парами соседних (в смысле 4-соседства [40]) элементов внутри домена: **0–1**, **2–3**, **0–2**, и **1–3**. На Рис. 5.2.2 показан контур с границами между элементами **1** и **3** и между элементами **2** и **3**. Тем самым, элементы **0**, **1**, и **2** на приведенном рисунке соответствуют некоторому одному кластеру, а элемент **3** — другому кластеру. Элементы (узлы графа), не разделенные контуром, считаются принадлежащими одному и тому же кластеру.

2. Производится выбор мажоритарного набора элементов, принадлежащих домену нижнего уровня (n); (в данном случае в такой набор входят элементы **0**, **1**, и **2**). Они связываются с узлом

следующего ($n+1$) уровня и их усредненные значения записываются в этот узел. Оставшиеся элементы (номер **3** на рисунке) считаются «подвешенными»; информация о них на следующий уровень не передается и сохраняется только в описании текущего узла на уровне n .

3. Строится карта контуров домена с помощью анализа множества элементов, состоящего из набора внутренних элементов домена, дополненного четверкой соседних с ней элементов, следующих в последовательной развертке, т.е. ближайших справа и снизу. Согласно примеру на Рис. 5.2.2, будет проверяться наличие контурных перепадов между элементами **1–1r** и **2–2b**. Карта контуров каждого домена записывается в узел следующего уровня, расположенный над данным доменом.

4. Производится усреднение яркостных, цветовых и текстурных параметров элементов, составляющих мажоритарный набор элементов домена, которые также записываются в узел следующего уровня.

По окончании прямого прохода по всему квадродереву, каждый узел любого уровня будет либо принадлежать поддереву, начинающемуся на одном из верхних уровней, либо сам являться корнем некоторого поддерева.

При *обратном проходе* вниз по квадродереву происходит синтез первичных кластеров. На каждом из уровней вершинам новых поддеревьев (т.е. тем, которые оказались «подвешены» на шаге 2 прямого прохода) присваиваются новые номера кластеров; данные номера распространяются вниз на все элементы, входящие в данное поддерево. Эта процедура рекурсивно выполняется для всего дерева начиная с верхнего уровня, т.е. вершины дерева, до нижнего — самого изображения; элементы изображения, получившие при этом одинаковые номера, образуют первичные кластеры (первичные сегменты) обрабатываемого изображения.

Таким образом процедура синтеза на каждом из узлов уровня n дерева сводится к следующему. Просматриваются все элементы очередного домена, т.е. узлы уровня $n-1$. Если рассматриваемый узел относится к числу связанных с узлом уровня n , ему присваивается номер узла n . Если же узел оказывается «подвешенным», он получает следующий свободный номер. Таким образом номер кластера, присвоенный некоторому текущему узлу, присваивается всем узлам следующего, более низкого, уровня, связанным с текущим узлом при прямом проходе.

Результатом данной процедуры является первичная карта кластеров сегментируемого изображения. Согласно структуре алгоритма сегментации, разбиение образовавшихся кластеров на более мелкие части в дальнейшем невозможно. Следовательно, результаты первой стадии должны удовлетворять следующим качественному и количественному требованиям. Качественное требование состоит в том, что первичный набор кластеров должен быть достаточно детализированным, чтобы обеспечить разделение всех областей изображения,

которые должны оказаться отдельными на финальной карте сегментов. Количественное требование заключается в том, что число первичных кластеров не должно быть излишне большим, чтобы предотвратить перегруженность процедуры анализа на второй стадии. Отметим также, что одним из преимуществ использования предлагаемого пирамидального подхода к сегментации является тот факт, что он во многих случаях позволяет решить проблему разделения областей, имеющих границу с разрывами.

То, как выглядят изображения после операции пирамидальной (первичной) сегментации (полученный набор «суперпикселей»), можно видеть на Рис. 5.2.6,б и 5.2.7,б.

5.2.6. Вторичная сегментация — упрощение графа кластеров

Если целью первой стадии сегментации является максимальная детализация изображения, то задача второй стадии — максимальное сокращение числа сегментов при условии минимальной потери (т.е. ошибочного слияния) объектов. Для дальнейшего анализа формируется связный граф соседства кластеров, полученных на нижнем уровне по окончании первичной сегментации. Каждому кластеру соответствует узел графа. Вычисляется информация, описывающая свойства каждого кластера. В нее входят число элементов, объединенных в кластере, и средние значения признаков — яркости, цвета и текстуры. Ребра графа соединяют пары соседствующих кластеров, имеющих общую пространственную границу.

Вторичная сегментация заключается в упрощении графа кластеров, построенного во время первичной сегментации. Это осуществляется сравнением признаков соседних узлов и слиянием тех, признаки которых близки. Если расстояние между парой соседних узлов меньше некоторого задаваемого порога, то эти узлы сливаются, а их признаки усредняются. Признаки новых кластеров, получаемых в результате слияния, формируются как усреднение признаков объединяемых кластеров или элементов. Тем самым они могут быть отображены в то же пространство признаков, а значит, допустимо использовать ту же самую метрику (5.2.11) и для измерения межкластерного расстояния. Однако такой подход никак не использует доступную *пространственную* информацию о кластере и его соседях, такую как число элементов в кластере, длина границы и т.д. По аналогии с уже рассмотренной выше метрикой, для сравнения близости кластеров в пространстве признаков вводится модифицированная *мера расстояния* между ними.

Предложенная метрика (5.2.11) не охватывает всех возможностей сравнения близости кластеров — на данном этапе доступна дополнительная информация. Можно модифицировать данную метрику путем добавления в пространство признаков новых характеристик — размеров сравниваемых кластеров (число элементов в кластере) $P(x)$. Модифицированное расстояние

между кластерами u и v в расширенном пространстве признаков в результате задается формулой:

$$\begin{aligned} \tilde{D}(u, v) = & \left[w_C + w_T (1 - t(u, v) p(u, v)) \right] s(u, v) (w_B D_B(u, v) + w_C D_C(u, v)) + \\ & + \left[(1 - s(u, v)) (w_C + w_T) + w_T s(u, v) t(u, v) p(u, v) \right] D_T(u, v). \end{aligned} \quad (5.2.12)$$

Здесь $D_B(u, v)$, $D_C(u, v)$ и $D_T(u, v)$ — яркостная, цветовая и текстурная метрики в соответствующих подпространствах признаков, задаваемые формулами (5.2.7)–(5.2.10), с той лишь разницей, что вместо признаков отдельных элементов или узлов подставляются соответствующие усредненные значения признаков кластеров. Функции p , t и s зависят от размеров, текстуры и цветовой насыщенности кластеров u и v , и задаются выражениями:

$$\begin{aligned} p(u, v) &= F(\min(P(u), P(v))/P_0), \\ t(u, v) &= (T(u) + T(v))/2T_{\max}, \\ s(u, v) &= 0.5 + 0.5F(\max(S(u), S(v))/S_0). \end{aligned}$$

Здесь T_{\max} — константа, определяющая максимальное значение уровня текстуры, функция подавления $F(z)$ задается тем же выражением (5.2.9), а S_0 и P_0 — уровни нечувствительности, имеющие размерности цветовой насыщенности и размеров кластера.

Формула (5.2.12) заменяет в расширенном пространстве признаков формулу (5.2.11). Множители перед $(D_B(u, v) + D_C(u, v))$ и $D_T(u, v)$ суть модифицированные веса яркостно-цветовой и текстурной составляющих соответственно. Сумма этих коэффициентов постоянна и равна $(w_C + w_T)$. Введение функций p , t и s в формулу (5.2.12) преследует цель сбалансировать метрику вблизи вырожденных ситуаций, когда один или несколько значений признаков, таких как уровень текстуры, цветовая насыщенность или размеры кластера становятся малы. Так, например, если размеры кластера u или v становятся меньше значения P_0 , вес цветовой части возрастает, а вес текстурной части одновременно уменьшается. Если цветовая насыщенность обоих кластеров u и v уменьшается, или же увеличивается уровень текстуры, то соответственно уменьшается модифицированный вес цветовой части и увеличивается вес текстурной.

Процедура вторичной сегментации построена по следующей итеративной схеме. Набор узлов графа (кластеров) анализируется в порядке убывания размеров кластеров. Для каждого очередного узла u , который в данном случае выступает в роли *базового узла* (базового кластера), просматривается набор соединенных с ним узлов v_i (соседних кластеров). Если согласно (5.2.12) расстояние между базовым узлом и очередным соседним узлом v_i не превышает заданного порога θ (т.е. $\tilde{D}(u, v_i) < \theta$), то узел v_i сливается с базовым узлом u . Это означает, что все связи узла v_i переориентируются на узел u , а также корректируются все связи и характеристики объединенного узла (кластера). Затем операция сравнения повторяется со

следующим соседним узлом v_{i+1} . По окончании списка соседних узлов осуществляется переход к следующему базовому узлу. Процедура циклически обрабатывает все узлы графа и заканчивается, когда обработаны все узлы.

Подобрать такое значение порога θ , чтобы получить приемлемые результаты за один проход, не удастся. Поэтому процедура вторичной сегментации выполняется в несколько итераций, каждая из которых в свою очередь включает этапы слияния малых кластеров, слияния больших кластеров, темных кластеров и текстурных кластеров. Значение порога θ задается уровнем итерации и размерами кластеров:

$$\theta = \alpha(1/(P_{\min} + P_0) + \beta). \quad (5.2.13)$$

Здесь P_{\min} — размер наименьшего кластера в паре, P_0 — порог чувствительности размеров кластера, β ограничивает уменьшение порога для «больших» кластеров ($P_{\min} \gg P_0$), а коэффициент α задает общую регулировку порога. Слияние кластеров выполняется итеративно с постепенным уменьшением P_0 и, соответственно, увеличением θ .

Процедура *слияния темных кластеров* обрабатывает только кластеры с низким значением яркости B . Вводятся также некоторые дополнительные изменения: текстурный весовой коэффициент w_T в (5.2.12) принимается равным нулю, и вчетверо уменьшаются весовые коэффициенты w_H и w_S в (5.2.8).

Процедура *слияния текстурных кластеров* применяется только к множеству кластеров с высоким значением уровня текстуры $T(u)$ согласно (5.2.6). При этом значение коэффициента w_T в (5.2.12) увеличивается вчетверо, а значения w_B и w_C уменьшаются вдвое.

Окончательные результаты сегментации представлены в конце главы на иллюстрациях Рис. 5.2.3-5.2.9.

5.2.7. Эксперименты по выбору параметров преобразований

Сегментация изображений общего вида является трудно формализуемой задачей с точки зрения выбора критериев оптимизации, необходимых для автоматического поиска параметров, предполагающего количественную оценку результатов преобразования. Для этого при сегментации изображений потребовалась бы априорная информация о числе и форме сегментов, которая может быть точной лишь при использовании синтезированных или же искусственно скомпилированных изображений-мозаик известных форм. Оценка качества сегментации произвольных изображений неизбежно требует морфологического анализа и интерпретации наблюдаемой сцены — задача, еще весьма далекая от автоматического решения в общем виде. Поэтому вопрос о качестве сегментации таких изображений решается при помощи экспертного визуального оценивания.

Проверка алгоритма и эксперименты по выбору параметров рассмотренных метрик и преобразований проводились как на искусственных, так и на реальных изображениях без существенных ограничений. Для этого использовалось более 250 изображений различных классов: а) синтезированные изображения трехмерных сцен с геометрическими фигурами, освещенными несколькими цветными источниками, полученные компьютерным моделированием; б) текстурные мозаики с границами заданной формы, составленные из классических текстур альбома Бродатца [221]; в) фотоснимки сцен ограниченной сложности с цветовыми и текстурными объектами известной простой формы; г) реальные изображения произвольного характера и сложности, включающие пейзажи, портреты, текстуры, изображения животных и пр. из различных доступных баз данных и библиотек. Изображения классов (а), (б), (в) и частично (г) — около 120 изображений — составляли тестовую группу настройки параметров.

По классам тестовых изображений ограниченной сложности (а), (б) и (в) качество сегментации оценивалось сравнением количества обнаруженных сегментов и точности проведения их границ с известным разбиением сцены, а по классу (г) — экспертным сравнением. Критерием являлась минимизация расхождения разбиений по всему ансамблю изображений группы настройки параметров. Параметры подбирались так, чтобы получить наилучшие результаты в среднем; выбранные значения фиксировались и применялись далее для произвольных изображений. Приемлемость выбранных параметров проверялась на оставшихся изображениях из класса (г).

Выбор коэффициентов первоначально осуществлялся для яркостно-цветового и для текстурного подпространств признаков по отдельности на основе изображений из классов (а) и (б) соответственно, после чего добавлялись изображения классов (в) и (г), при сегментации которых уже необходимо задействовать полное пространство признаков. В результате экспериментов коэффициенты цветовой метрики (5.2.8) составили значения $w_H = 6,7$ и $w_S = 1,0$; текстурные коэффициенты в (5.2.6) и (5.2.10) — значения $w^n = 1 - 0,12n$, где n — номер масштабного уровня. Коэффициенты объединенной метрики в (5.2.11) были выбраны $w_B = 3,3$, $w_C = 1,0$, $w_T = 0,65$. Значения α , и P_0 в (5.2.13) зависят от номера итерации: $\alpha = 1 - 0,7i/I$ и $P_0 = S^{(0,2+0,8i/I)}$, где i — номер итерации, I — общее число итераций, а S — общее число элементов в изображении.

Примеры сегментации синтезированных тестовых цветного и текстурного изображений показаны на Рис. 5.2.3, сегментации тестовых цветных изображений сцен ограниченной и повышенной сложности — на Рис. 5.2.4 и 5.2.5, сегментации реальных цветных изображений — на Рис. 5.2.6 и 5.2.7, сегментации реальных цвето-текстурных изображений — на Рис. 5.2.8 и 5.2.9. Внутренние части каждой из сегментированных областей заполнены средним значением

по сегменту (за исключением Рис. 5.2.3,в и Рис. 5.2.9,в где границы нанесены поверх исходных изображений); границы между сегментами проведены белыми линиями.

Важно заметить, что поскольку алгоритм ориентирован на анализ изображений в обобщенном текстурно-цветовом пространстве, то его возможности сокращаются, если исходное изображение отображается лишь в его подпространство, что происходит, например, при сегментации черно-белого изображения. Тем не менее, даже в таком случае алгоритм демонстрирует приемлемое качество (Рис. 5.2.3,в).

Исследование задачи и разработка методов цвето-текстурной сегментации изображений проводились в рамках договора с Samsung Advanced Institute of Technologies (SAIT, Samsung co.).

Изложенные в разделе 5.2 результаты опубликованы в работах [185,194,228,237,311,312].

5.3. Выводы и результаты

Исследовано применение модели и разработанных методов обработки видеоинформации для решения проблемы сегментации изображений. Сегментация означает разбиение изображения на множество связанных областей в соответствии с выбранными критериями. Рассмотрены два различных подхода, обусловленные принципиальными различиями исходных данных и критериев разделения областей.

Первой исследована научно-техническая задача сегментации снимков поверхности электронных микросхем, формируемых сканирующим электронным микроскопом. Особенность задачи в том, что контурные границы объектов не однозначны — допускается наличие двух разных типов границ между объектами и совпадение значений яркостей соседних сегментов.

Разработан алгоритм, основанный на декомпозиции и обнаружении объектов, позволяющий привести разные типы границ к одному виду. Процедура обработки включает этапы: формирования препаратов с контурными откликами допустимых типов границ, прослеживания откликов и получение слоев контурных линий соответствующих типов, формирования единой карты контурных линий, морфологической обработки карты линий, утончения линий, удаления ложных линий и продления контурных линий. Результатом является символическое описание, т.е. разбиение изображения на связанные области, в которых каждому из элементов сегмента приписан один и тот же индекс. Тестирование алгоритма показало, что результаты сегментации удовлетворяют требуемым критериям точности разделения объектов в 98,3% случаев.

Второй исследована задача цвето-текстурной сегментации изображений общего вида, в которой изображение должно разбиваться на однородные области, отличающиеся от соседствующих по своим яркостным, цветовым и/или текстурным характеристикам.

В качестве критерия однородности или различия элементов и областей изображения выбрано расстояние между их образами в пространстве признаков. Предложен способ построения объединенного пространства яркостно-цвето-текстурных признаков. В качестве текстурных признаков предложена оценка локальных текстурных характеристик, основанная на многомасштабной ориентированной локальной вариации и локальной осцилляции градиента, вычисляемого по текстурно-детальной компоненте, формируемой при декомпозиции. Предложена метрика в объединенном пространстве признаков, допускающая как редукцию, так и расширение пространства.

На основе анализа расстояний в пространстве признаков разработан иерархический пирамидальный алгоритм сегментации изображений, использующий как признаки для отдельных точек, так и суммарные признаки для формируемых пространственных кластеров.

Предложенный иерархический подход и разработанный алгоритм показали эффективность и приемлемые результаты на подавляющем большинстве доступных изображений, среди которых: тестовые изображения сцен ограниченной сложности, изображения из различных баз данных и библиотек, а также целый ряд изображений реальных сцен, портретов, пейзажей и пр. без существенных ограничений.

Разработанный алгоритм допускает модификации путем применения различных цветовых, текстурных или обобщенных признаков и метрик. Он может быть применен не только к обычным цветным *RGB* изображениям, но также и к любым другим одно- или многоканальным массивам данных, удовлетворяющим свойствам изображений. Необходимо только выбрать адекватный набор признаков и метрику, а также подобрать соответствующие параметры.

Эффективность разработанных методов сегментации еще раз подтверждает вывод Главы 2, что при анализе изображения следует выбирать лишь ту компоненту изображения, в которой должна содержаться искомая информация, а также применять преобразования, специфические для каждой из компонент.

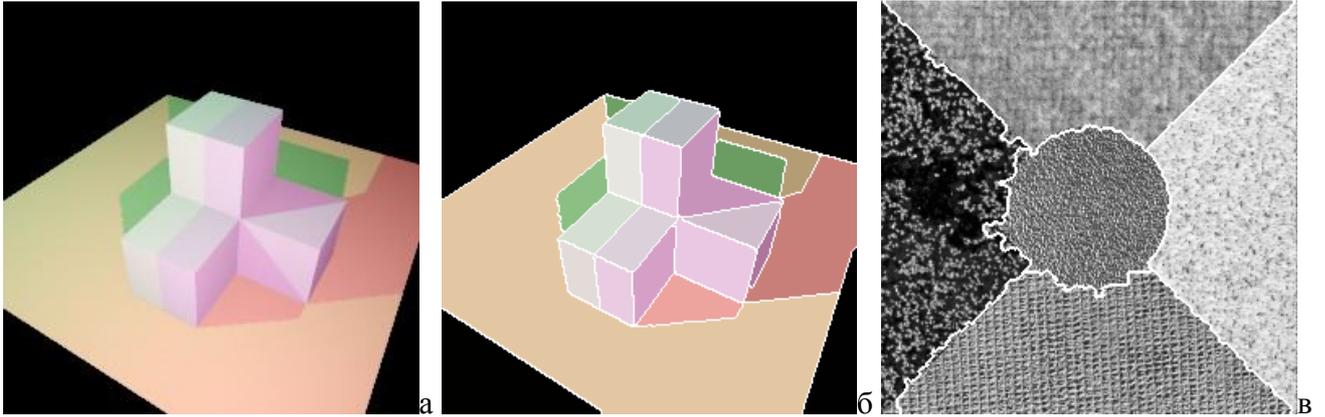


Рис. 5.2.3. Тестовые синтезированные изображения: а) исходное цветное изображение, полученное компьютерным моделированием (сцена освещена двумя источниками разного цвета); б) результат сегментации (23 сегмента); в) коллаж из альбома текстур Бродатца — сегментация на 5 сегментов.

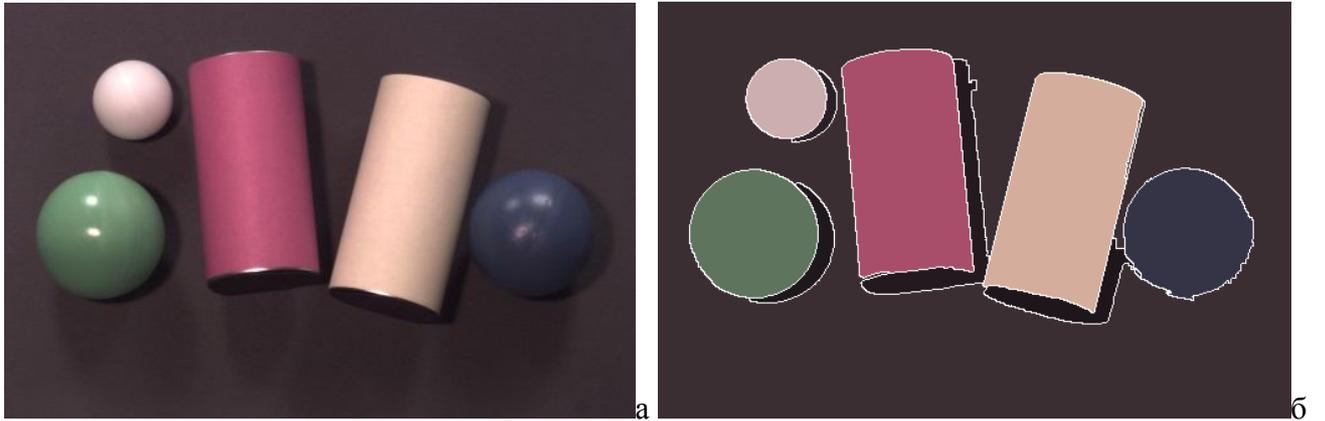


Рис. 5.2.4. Тестовое изображение цветных объектов ограниченной сложности: а) исходное изображение; б) результат сегментации (11 сегментов).



Рис. 5.2.5. Тестовое изображение цветных объектов повышенной сложности: а) исходное изображение; б) результат сегментации (75 сегментов).

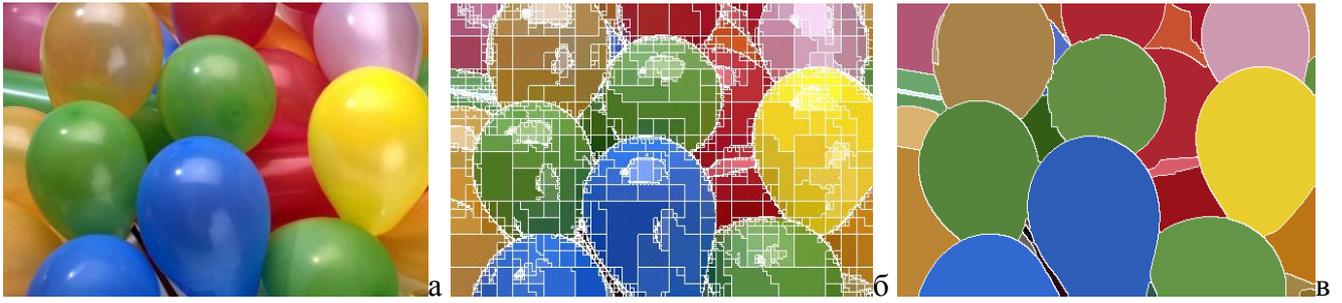


Рис. 5.2.6. Сегментация реального цветного изображения из альбома Adobe: а) исходное изображение; б) результат первичной сегментации; в) окончательный результат (35 сегментов).



Рис. 5.2.7. Сегментация реального цветного изображения: а) исходное изображение «Тукан»; б) результат первичной сегментации; в) окончательный результат (22 сегмента).

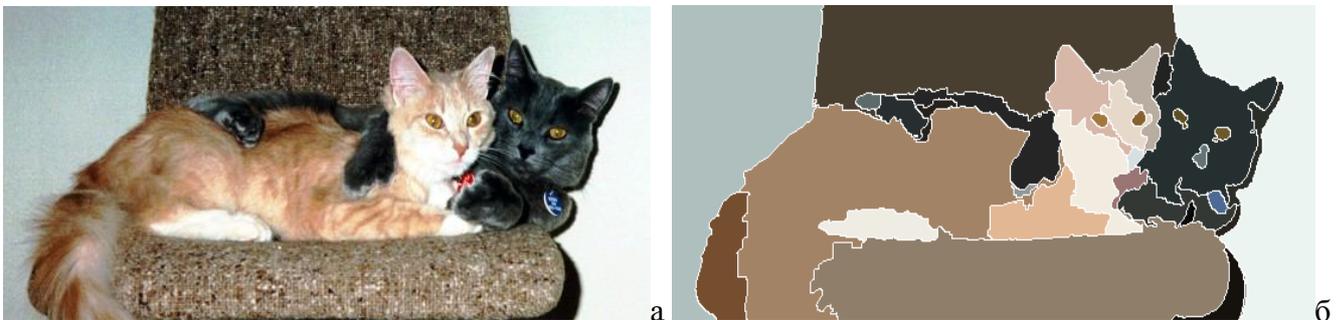


Рис. 5.2.8. Сегментация реального цвето-текстурного изображения: а) исходное изображение «Кошки»; б) результат сегментации (33 сегмента).



Рис. 5.2.9. Сегментация цвето-текстурных изображений: а) исходное изображение; б) результат его сегментации (37 сегментов); в) сегментация изображения «Зёрна» (12 сегментов).

ГЛАВА 6. ПРИМЕНЕНИЕ РАЗРАБОТАННЫХ МЕТОДОВ ДЛЯ ОБРАБОТКИ И АНАЛИЗА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ

Введение

Рассматривается применение модели и методов обработки изображений для анализа видеоданных на примере обработки видеопоследовательностей, получаемых компьютерными капилляроскопами. Капилляроскопия — способ неинвазивного (без повреждения кожных покровов и проникновения внутрь организма) исследования микроциркуляторного кровяного русла с целью диагностики функционирования периферического отдела сердечно-сосудистой системы человека в кожных и слизистых покровах. В последнее время она является одним из активно развивающихся направлений в медицине, использующих методы цифровой обработки изображений. Скорость, объем кровотока, форма капилляров и некоторые другие параметры позволяют судить о состоянии здоровья пациента и его предрасположенности к заболеваниям [36]. Согласно обзору [290], большинство методов неинвазивного измерения скорости протекания крови основаны на использовании томографии (МРТ, ПЭТ), а также доплеровского эффекта в области ультразвука или видимом спектре [80]. Микровидеосъемка капилляров в основном используется при визуальном [36] и интерактивном [7] анализе. В статье [302] говорится об автоматическом измерении скорости кровотока на основе прослеживания промежутков между группами эритроцитов, однако такой подход применим только к почти прямолинейным участкам капилляров, на которых удастся выделить прямоугольную область для анализа. В настоящей главе рассматривается решение задачи автоматического измерения динамики кровотока на протяжении всей видимой части капилляра, включая его артериальный, венозный и переходный отделы.

В капилляроскопии используются специальные приборы — компьютерные капилляроскопы [7,46,62,63], позволяющие осуществлять микровидеосъемку объекта в отраженном свете. Большинство капилляроскопов разрабатывается для наблюдения ногтевого ложа пальцев рук, как наиболее удобной области из возможных. Это обусловлено тем, что в этих областях капиллярные петли выстраиваются параллельно поверхности кожи, и в последнем слое их можно достаточно четко рассмотреть по всей длине.

При обследовании пучок видимого света падает на палец пациента, отражается от капилляров и других тканей, регистрируется устройством ввода и поступает в компьютер. Капилляры выглядят как вытянутые петли с темными участками, изображающими группы эритроцитов, разделенных более светлыми промежутками, в которых находятся прозрачные

лейкоциты и плазма крови; стенки капилляров также прозрачны (Рис. 6.0.1). Движение крови и, соответственно, смещение положения эритроцитов, приводит к изменению картины объектов в области капилляра. Типичные параметры формируемого капилляроскопами видеосигнала следующие: размер кадра 640×480 пикселей, частота 25-100 кадров/сек, поле зрения около $0,6 \times 0,4$ мм, разрешение около 1 мкм.



Рис. 6.0.1. Отдельный кадр исходной видеопоследовательности после устранения искажений.

Основное назначение капилляроскопа — измерение скорости капиллярного кровотока. Вычисление скорости по видеоданным требует определения межкадрового смещения объектов, которое можно найти корреляционным методом [301]. Однако проблема в том, что корреляционный анализ требует предварительного задания области анализа; использовать в качестве таковой капилляр с его сложной и не описываемой аналитически формой, затруднительно. Решением может быть преобразование области капилляра в область простой формы, на которой затем и будет выполняться корреляционный анализ.

В начале настоящей главе рассматриваются вопросы предварительной обработки видеоданных, пространственной синхронизации кадров и формирования усредненного изображения, которое используется для получения и фильтрации карты контурных линий капилляров, нахождения мажоритарного капилляра, составления морфологических характеристик контурных линий. В качестве геометрического преобразования предлагается гомеоморфное распрямляющее отображение области капилляра в область прямоугольной формы, используемое для преобразования всех кадров и формирования новой видеопоследовательности, по которой уже проводится корреляционный анализ и определяются характеристики капиллярного кровотока. Излагается алгоритм вычисления мгновенной скорости и объема капиллярного кровотока как функции времени и местоположения в капилляре. В экспериментах использовались видеоданные, получаемые следующими капилляроскопами: [62,63].

6.1. Предварительная обработка видеок кадров

6.1.1. Особенности получаемых видеоданных

Видеосигнал, получаемый с капилляроскопов, особенно разработанных на основе использования КМОП–матриц, содержит ряд различных искажений и помех, поэтому требует довольно существенной предварительной обработки. Искажения можно разделить на два класса: систематические, повторяющиеся в видеопоследовательности от кадра к кадру, и случайные, индивидуальные для каждого кадра.

К систематическим искажениям относятся рассогласования измерений, вызванные различиями характеристик каждого из сенсоров КМОП–матрицы [304,305], неравномерность освещения по полю кадра, артефакты (пыль, пятна) на матрице, малый контраст полезного сигнала, наличие периодических помех. К случайным искажениям — индивидуальные шумы сенсоров и аналого-цифровых преобразователей, а также дрейф наблюдаемых объектов (капилляров) в поле изображения, вызываемый дрожанием пальца обследуемого [186], достигающий сотни и более микрон.

Как выглядит исходный кадр типичной видеопоследовательности, полученной капилляроскопом с КМОП–матрицей, можно видеть на Рис. 6.1.1,а. Видно, что интересующие детали (капилляры) достаточно слабо различимы на фоне имеющихся помех, среди которых доминируют регулярные вертикальные линии. Исследования показывают, что значения элементов данного исходного изображения имеют среднеквадратическое отклонение (СКО) 12,55 градаций яркости (из полного диапазона в 256 градаций). При этом собственно полезный сигнал оказывается заметно слабее уровня помех — его СКО равно всего 7,51 градаций, в то время как СКО систематических помех равно 9,94 градаций, а СКО случайных шумов датчиков равно 1,77 градации (что составляет 23% от величины полезного сигнала). Подсчет СКО осуществлялся методом, рассмотренным в разделе 1.2.2, по формулам (1.2.15)–(1.2.20). Таким образом, коррекция сигнала, выравнивание общего фона, а также повышение контрастов информационных деталей на кадрах видеопоследовательности являются абсолютно необходимыми этапами обработки видеоданных.

6.1.2. Устранение систематических искажений

Использованная в экспериментальном приборе КМОП–матрица [304] представляет собой двумерный массив независимых сенсоров, каждый из которых имеет свои индивидуальные, отличающиеся от остальных сенсоров параметры. Для коррекции сигнала, получаемого каждым из сенсоров, рекомендуется [305] использовать линейное преобразование вида

$$y = 256(x - X_B)/(X_W - X_B), \quad (6.1.1)$$

где x и y — исходное и скорректированное значения сигнала, X_W и X_B — опорные значения белого (W) и черного (B) эталонных стимулов, индивидуальные для каждого отдельного сенсора матрицы (дающего значение одного пикселя изображения), полученные при вводе калибровочных стимулов яркости W и B . Такие стимулы должны представлять собой черную и белую поверхности с круговой диаграммой направленности отраженного света. Оцифровку необходимо осуществлять при одной и той же яркости ламп подсветки, причем оптическая система должна быть расфокусирована относительно положения этих поверхностей [305].

Эксперименты однако показывают, что рекомендуемый алгоритм (6.1.1) не позволяет устранить помехи полностью. После контрастирования изображения на нем остаются заметными следы вертикальных полос. Было сделано предположение, что это связано с определенной нелинейностью характеристик сенсоров, а также с тем, что использование черного и белого стимулов может приводить к попаданию сигнала на края характеристической кривой сенсоров и выходить за границы ее линейного участка.

Для компенсации нелинейности и уменьшения заметности помех было принято решение вместо линейного преобразования по двум точкам использовать, во-первых, кусочно–линейное преобразование по нескольким точкам, а во-вторых, калибровочные значения W и B выбирать не на самых краях яркостного диапазона, а вблизи максимального и минимального значений яркости исходного сигнала. Эксперименты показали, что трех калибровочных точек кусочно–линейного преобразования вполне достаточно для получения приемлемых результатов. Преобразование (6.1.1) в этом случае примет следующий вид:

$$\begin{aligned} y &= (G - B)(x - X_B)/(X_G - X_B) + B, & \text{если } x < X_G, \\ y &= (W - G)(x - X_G)/(X_W - X_G) + G, & \text{если } x \geq X_G. \end{aligned} \quad (6.1.2)$$

Здесь W , B , G — калибровочные значения белого, черного и серого стимулов, т.е. те, которые подавались на вход системы и, соответственно, должны быть получены на выходе после ввода, оцифровки и преобразования; X_W , X_B , X_G , — опорные значения каждого отдельного сенсора (пикселя) матрицы, получаемые по калибровочным стимулам.

Возникает вопрос, откуда взять значения W , B и G для каждого сенсора, поскольку у нас нет иных измерений, кроме значений x , возникающих на выходе сенсоров матрицы. Получить их можно путем ввода светлого, темного и серого эталонных стимулов соответственно, при сохранении одной и той же схемы освещения, и последующим усреднением значений x по всему полю изображения. Дополнительным плюсом преобразования (6.1.2) является то, что при таком подходе автоматически устраняются неравномерности освещения стимула по полю кадра

(вызванные особенностями конкретной системы подсветки), а также значительное подавление артефактов на изображении, вызываемых наличием «грязи» в оптической части устройства.

После того, как найдены значения W , B и G для каждого сенсора, коррекция (6.1.2) выполняется для каждого из кадров всех видеопоследовательностей, получаемых тестируемым устройством при одном и том же освещении.

6.1.3. Устранение периодических помех

На кадрах ряда экспериментальных видеопоследовательностей, даже после коррекции систематических искажений, обнаружилась горизонтальная периодическая помеха с периодом в P элементов ($P \ll N$, где N — длина строки кадра), хотя и слабая, но значительно мешающая правильной корреляционной привязке. Для устранения помехи использовался рассмотренный в разделе 3.1.2 алгоритм пространственной фильтрации периодических помех. В упрощенном виде он состоит в следующем. На каждой строке m подсчитывается набор из $P+1$ значений U_0, \dots, U_P , где U_p есть среднее значение элементов $x(m, n_k)$, для которых $n_k = p + Pk$ ($k = 0, 1, \dots, N/P$). Пусть также V есть среднее значение полученных U_0, \dots, U_P . Тогда фильтрация сводится к следующей простой операции:

$$x_{mn} = x_{mn} - U_{n(\text{mod } P)} + V. \quad (6.1.3)$$

Результат коррекции систематических искажений и удаления периодических помех показан на Рис. 6.1.1,б. Легко заметить, что наряду с устранением полос, являющихся следствием разброса характеристик сенсоров, устраняются также и пятна. Также становятся заметными периферийные капиллярные петли, почти не различимые на исходном изображении.

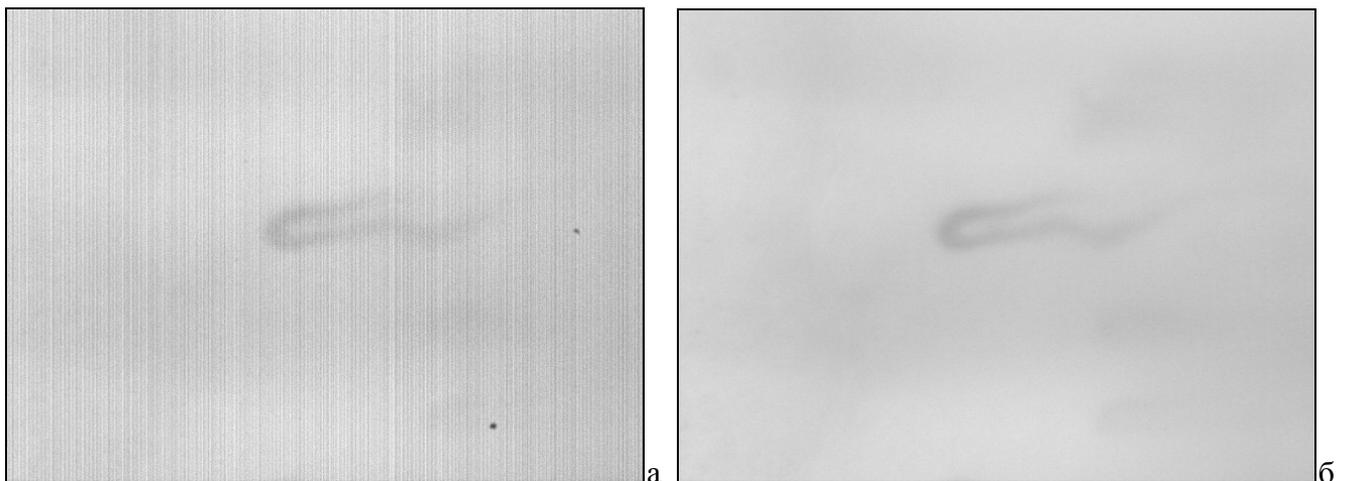


Рис. 6.1.1. а) Исходный кадр видеопоследовательности; б) результат устранения систематических искажений и периодических помех.

6.1.4. Выравнивание и усиление локальных контрастов сигнала

Устранение дрейфа наблюдаемых объектов в поле изображения требует совмещения кадров видеопоследовательности. Само совмещение осуществляется корреляционным методом. Как отмечалось выше, контраст объектов (капилляров) даже после устранения искажений и периодических помех остается весьма низким. Для повышения точности совмещения сигналов следует усилить те признаки, по которым должно осуществляться совмещение, и ослабить мешающие признаки. В нашем случае это означает необходимость повысить амплитуду (локальный контраст) капилляров относительно фона, а также убрать неравномерности фона из-за непостоянности освещения по полю кадра, если таковые остаются, например, если корректирующее преобразование (6.1.2) не применялось.

Выровнять фон изображения с одновременным усилением контрастов деталей можно при использовании одной из операций повышения локальных контрастов, которые были рассмотрены в разделе 3.3. Наиболее удобен в нашем случае линейный детерминированный алгоритм (3.3.4), иначе называемый операцией нерезкого маскирования [40,155,173]:

$$y_{mn} = A(x_{mn} - \bar{x}_{mn}) + Bx_{mn} + C \quad (6.1.4)$$

с соответствующим выбором параметров A , B и C . Здесь x_{mn} и y_{mn} — значения исходного и обработанного элементов изображения, а

$$\bar{x}_{mn} = \sum_{h=-H}^H \sum_{w=-W}^W x(m+h, n+w) / S \quad (6.1.5)$$

есть аналогичное (2.1.2) среднее значение по фрагменту размерами $S = (2H+1) \times (2W+1)$ с центром в точке (m, n) . Для выравнивания изображения следует выбрать $B = 0$, а параметры размеров окрестности W и H — около 1/15 размеров изображения.

Коэффициенты A и C найдем анализируя усредненную видеопоследовательность с выровненными при помощи (6.1.4) кадрами, причем для сокращения времени можно брать кадры с достаточно большим прореживанием — например, лишь каждый пятнадцатый–двадцатый. Пусть $h(x)$, $(x = 0, \dots, 255)$ — нормализованная гистограмма изображения (вероятность точки со значением x), т.е. $\sum h(x) = 1$. Выберем достаточно малое δ и найдем значения квантилей порядков δ , 0,5 и $(1 - \delta)$, обозначив их L , M и R соответственно:

$$\sum_{x=0}^{x<L} h(x) \leq \delta < \sum_{x=0}^{x<L+1} h(x), \quad \sum_{x=0}^{x<M} h(x) \leq 0.5 < \sum_{x=0}^{x<M+1} h(x) \quad \text{и} \quad \sum_{x=0}^{x<R-1} h(x) < 1 - \delta \leq \sum_{x=0}^{x<R} h(x).$$

В нашем случае выбиралось значение $\delta = 0,001$. Параметры A и C определяются как

$$A = 256\alpha / (R - L); \quad C = 256(M - L) / (R - L). \quad (6.1.6)$$

Параметр α служит для некоторого уменьшения размаха сигнала и должен быть меньше единицы; нами было выбрано $\alpha = 0,8$. Затем все кадры видеопоследовательности обрабатываются по формуле (6.1.4) с выбранными значениями параметров A , B и C .

6.2. Компенсация дрейфа и получение усредненного изображения

Как уже говорилось выше, получаемые видеопоследовательности характеризуются заметным дрейфом наблюдаемых объектов в поле изображения, причина которого — естественное смещение и дрожание пальца обследуемого [46,186], в том числе и из-за пульсации крови в результате сердцебиения. Так, отклонение наблюдаемого объекта всего на несколько сотых долей миллиметра приводит к смещению изображения (всей наблюдаемой сцены целиком) на несколько десятков пикселей, а в реальности смещения достигают четверти и более размеров кадра. С достаточной степенью точности можно считать, что дрейф приводит только к сдвигу кадров, без их поворота или изменения масштаба. Компенсация такого дрейфа, т.е. нахождение параметров совмещения отдельных кадров и соответствующий сдвиг изображений, является необходимым этапом обработки данных.

На Рис. 6.2.1,а показана трасса дрейфа кадров видеопоследовательности, а на Рис. 6.2.1,б — графики зависимости смещения от номера кадра. Черный график соответствует смещению по вертикали, серый — по горизонтали.

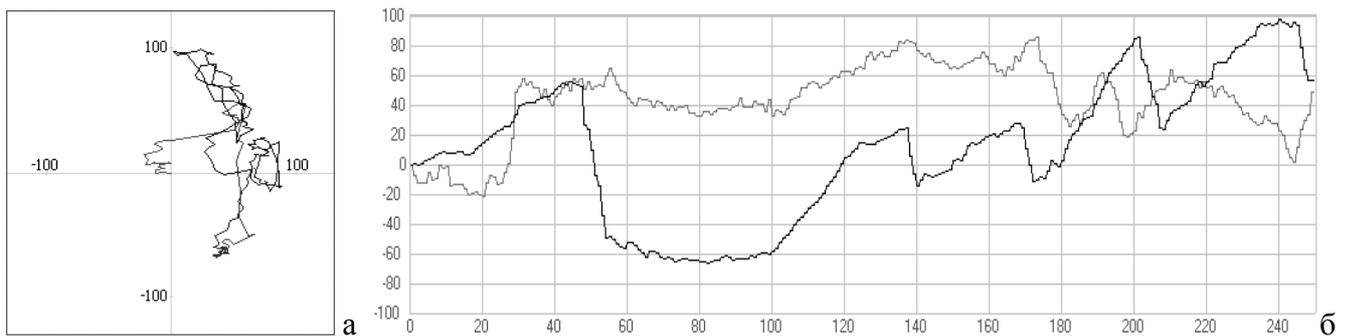


Рис. 6.2.1. Трасса дрейфа кадров видеопоследовательности (а); графики значений смещения кадров по вертикали (черный) и по горизонтали (серый) (б).

Задача автоматического совмещения кадров видеопоследовательности осложняется несколькими факторами. Очевидно, что основой для совмещения кадров должны служить изображения капилляров, однако их контуры, как правило, достаточно размыты. Кровоток, присутствующий в капиллярах, выглядит как перемещение объекта внутри сцены, что в свою очередь мешает правильному совмещению. Дополнительной спецификой задачи является сравнительно высокий уровень шума, остающийся после устранения систематических помех и улучшения изображений.

С достаточной степенью точности можно считать, что дрейф кадров приводит только к их сдвигу без поворота или изменения масштаба. Процедура компенсации дрейфа состоит из следующих шагов, выполняемых для каждого кадра.

1. Перед совмещением выполнение выравнивания фона алгоритмом (6.1.4) с параметрами $A = 1,3$, $B = 0$ и $C = 128$. Локальное среднее \bar{x} находится по окрестности размерами $M/16 \times N/16$, где M и N — размеры изображения.

2. Двухэтапное сглаживание изображения перед совмещением. Поскольку уровень шума на видеокдрах весьма высок, перед корреляционным анализом осуществляется сглаживание изображения алгоритмом псевдомедианной фильтрации [173] (для устранения возможных импульсных помех) и фильтром арифметического среднего (6.1.5) по окрестности 3×3 пикселей.

3. Совмещение изображений и определение параметров сдвига. Классическим средством является корреляционный анализ [210,262], который сводится к отысканию максимума двумерной функции ковариации базового S и сдвигаемого T изображений:

$$\text{cov}_{T,S}(v, h) = \frac{1}{W_{v,h}} \sum_{m,n} (x_S(m, n) - C)(x_T(m - v, n - h) - C). \quad (6.2.1)$$

Здесь $x_S(m, n)$ и $x_T(m, n)$ — значения элементов изображений S и T в точке (m, n) ; C — средние значения яркостей выровненных изображений S и T после шага 1; (v, h) — величина сдвига изображения T по вертикали и горизонтали; $W_{v,h}$ — число точек в области пересечения изображений при сдвиге на v и h .

Для подсчета ковариационной функции по всему множеству возможных параметров (v, h) требуется порядка $Z = M \times N \times V \times H$ операций, где $M \times N$ — размеры изображения, а $V \times H$ — размеры области анализа, внутри которой ищется максимум функционала (6.2.1). Смещения, даже относительно предыдущего кадра, могут быть достаточно заметными, поэтому Z оказывается весьма большим. Сокращения объема вычислений можно добиться использованием быстрого корреляционного совмещения изображений (раздел 4.4), основанного на предварительном одномерном совмещении проекций изображений по вертикали и горизонтали, и окончательным двумерным корреляционным совмещением в малой окрестности найденной точки [165].

4. Сдвиг кадра исходной видеопоследовательности в соответствии с параметрами смещения, найденными на шаге 3. Очевидно, что при сдвиге часть изображения выходит за края кадра и пропадает, а противоположная часть кадра оказывается неопределенной. Для устранения данной проблемы неопределенные части кадра заполняются значениями, которые берутся из усредненного изображения, формирующегося на шаге 5 по предыдущим кадрам. Обработанный кадр записывается обратно в видеопоследовательность на прежнее место.

5. Формирование текущего базового изображения из фильтрованных на шаге 1 кадров путем сдвига, суммирования и деления на текущее число кадров. Такое изображение, получаемое суммированием по предыдущим обработанным кадрам, используется в качестве

базового изображения S в (6.2.1) при совмещении с очередным кадром. Начальным базовым изображением выбирается первый кадр видеопоследовательности.

В результате совмещения всех кадров видеопоследовательности, суммированием значений соответствующих элементов всех получаемых на шаге 4 сдвинутых кадров и делением на общее число кадров, формируется *усредненное изображение*.

Компенсация дрейфа позволяет получить видеопоследовательность, в которой полностью устранено видимое дрожание объектов. К обработанным таким образом и синхронизованным видеоданным могут быть применены методы как внутрикадровой, так и межкадровой фильтрации, позволяющие дополнительно снизить уровень помех.

На Рис. 6.2.2,а показано усредненное по кадрам изображение, получаемое в результате обработки видеопоследовательности из 250 кадров. Как видно, на нем капилляры приобрели вид непрерывных полос и значительно уменьшен шум, имеющийся на всех кадрах видеопоследовательности. Считая шум случайным и независимым, степень его подавления составляет приблизительно \sqrt{K} , где K — общее число кадров. На усредненном изображении стали видны многие детали, незаметные на исходных зашумленных кадрах.

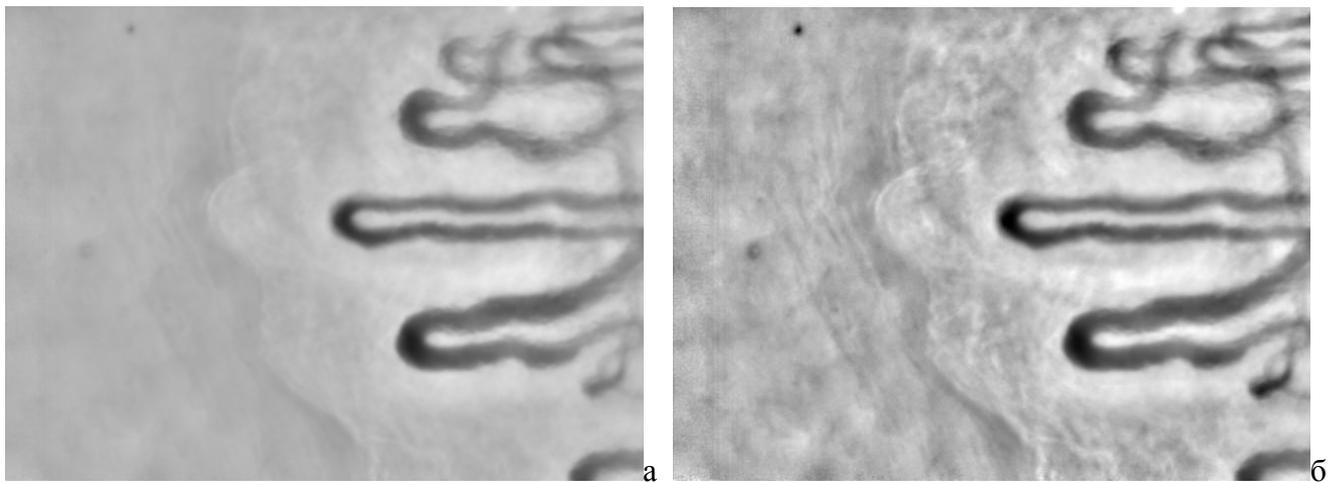


Рис. 6.2.2. Усредненное изображение (а) и адаптивное повышение локальных контрастов (б).

Для возможного визуального анализа полученного усредненного изображения врачом-диагностом заметность деталей можно дополнительно повысить применением усиления локальных контрастов [173], например, линейным адаптивным преобразованием (3.3.5) в следующей модификации:

$$y_{mn} = f(R - L)(x_{mn} - M_{mn}) + B M_{mn} + C, \quad (6.2.2)$$

в которой используются квантили распределения значений элементов по локальному фрагменту с центром в обрабатываемой точке (m, n) . L и R — левый и правый квантили (например, порядков 0,3 и 0,7), а M_{mn} — медиана по фрагменту, т.е. квантиль порядка 0,5. Функция $f(v)$

адаптивно задает изменение коэффициента усиления в зависимости от величины контрастов в локальном фрагменте исходного изображения, определяемых разностью квантилей ($R - L$); функция $f(v)$ должна иметь максимум при $v = 0$ и убывать с возрастанием v . В качестве таковой была выбрана следующая зависимость:

$$f(v) = 1 + A/(1 + v/V)^2.$$

Здесь A — коэффициент усиления, аналогичный такому же в формуле (6.1.4), а V — параметр, управляющий скоростью спада $f(v)$. Результат применения такого алгоритма показан на Рис. 6.2.2,б. Использовались значения: $A = 4$; $V = 20$; $B = 0,5$; $C = 0,4$. Размеры локального фрагмента составляли $1/8$ размеров изображения.

Усредненное изображение того вида, который показан на Рис. 6.2.2,а, является базовой информацией, на основе которой происходит весь последующий процесс обнаружения и анализа формы капилляров.

6.3. Формирование карты контурных линий капилляров

Вследствие способа формирования образа капилляра из отдельных видимых пятен (групп эритроцитов), эластичности стенок капилляров, приводящей к небольшим вариациям их формы, погрешностей совмещения кадров, а также некоторых других причин, области капилляров на усредненном изображении имеют непостоянную яркость и нечеткие границы. Прямое обнаружение областей и границ капилляров на таком изображении при помощи пороговых или контурных методов [40], включая методы прослеживания границ, в частности алгоритма Кэнни [223], не обеспечивает корректного результата [89].

В связи с этим нахождение границ капилляров строится как процедура отыскания множества контурных линий без ветвлений, расположенных вблизи так называемого опорного набора областей, обнаруживаемого пороговым методом. Весь набор контурных линий будем называть картой контурных линий капилляров. Процедура ее получения включает следующие этапы: формирование опорного набора областей приближенной формы, нахождение множества контурных точек, формирование первичной карты контурных линий, синтаксическая фильтрация с корректировкой линий, упрощение карты линий и карты областей.

Далее будут использоваться операции фильтрации и анализа, в которых потребуются масштабные привязки к размерам объектов на изображении. В качестве такого параметра выбран средний видимый диаметр капилляра D в единицах пикселей изображения.

6.3.1. Формирование опорного множества областей

Сформируем двумерный массив бинарных элементов, в котором единица означает, что с вероятностью, превышающей некоторый порог, элемент относится к области капилляра. Это достигается выравниванием фона усредненного изображения, пороговым обнаружением областей и последующей морфологической коррекцией.

Для выравнивания фона усредненное изображение обрабатывается фильтром высоких частот при помощи преобразования (6.1.4) с параметрами $A = 1$, $B = 0$ и $C = 0,5$. Параметры окрестности в (6.1.5) выбираются следующими: $W = H \approx 1,5D$. Поскольку яркость капилляров ниже яркости фона, обнаружение областей капилляров состоит в пороговом разделении элементов фильтрованного изображения следующим адаптивным алгоритмом, близким классическому методу бинаризации текстовых документов [321]:

$$y_{mn} = \begin{cases} 1 & \text{если } x_{mn} < C - (\sigma + \delta), \\ 0 & \text{иначе.} \end{cases} \quad (6.3.1)$$

Здесь σ — значение СКО (корня из дисперсии) яркости по фрагменту с центром в точке (m, n) , а δ — постоянная небольшая константа. Полученные единичные элементы отмечают области капилляров, а нулевые — область фона.

Морфологическая коррекция предназначена для заполнения мелких разрывов между областями, сглаживания неровностей краев и удаления мелких областей. Для этого выполняются две последовательные операции *замыкания* и *размыкания*, каждая из которых в свою очередь состоит из пары операций *дилатации* и *эрозии* [40]. Замыкание строится как дилатация, результат которой подвергается эрозии, а размыкание — как эрозия, результат которой подвергается дилатации. Дилатация и эрозия основаны на анализе окрестности $V_8(m, n)$ каждого элемента x_{mn} изображения, состоящей из 8 окружающих его элементов, т.е. $\{x_i \mid i = 1, \dots, 8\} = \{x(m \pm 1, n \pm 1)\}$. Проверяется наличие хотя бы одного $x_i \in V_8(m, n)$, ($i = 1, \dots, 8$), удовлетворяющего следующим условиям.

$$\text{Операция дилатации: } y_{mn} = \begin{cases} 1 & \text{если существует } x_i \neq 0, \\ 0 & \text{иначе.} \end{cases} \quad (6.3.2)$$

$$\text{Операция эрозии: } y_{mn} = \begin{cases} 0 & \text{если существует } x_i = 0, \\ 1 & \text{иначе.} \end{cases} \quad (6.3.3)$$

Получаемое изображение все еще содержит большое количество областей малых размеров, а также ложных областей на границах изображения. Был выбран пороговый параметр N_R , ограничивающий максимально допустимое число опорных областей (в нашем случае было выбрано $N_R = 15$). Анализируются размеры, вытянутость и расположение областей, а сами области ранжируются в порядке убывания площади. Удаляются вытянутые области вблизи границ изображения, а также мелкие, ранг которых превышает значение N_R .

В результате преобразований формируется карта *опорного множества областей*, вид которой показан на Рис. 6.3.1,а. Это множество в дальнейшем будет использоваться в качестве вспомогательного при анализе контурных линий и выборе мажоритарного капилляра.

6.3.2. Нахождение множества контурных точек

Для обнаружения контурных точек используется подход, близкий первой части алгоритма Кэнни [223]. Сначала усредненное изображение сглаживается гауссовым фильтром $y_{mn} = \sum_{u,v} x(m-u, n-v) \exp\{-((m-u)^2 + (n-v)^2)/2\sigma^2\}/2\pi\sigma$, снижающим шум и повышающим устойчивость нахождения максимума градиента в центре контурного перепада; рекомендуется значение $\sigma = 1,5$. Затем происходит обнаружение множества контурных точек. В каждой точке (m,n) сглаженного изображения находятся направление α_{mn} и модуль g_{mn} градиента яркости; для этого анализируются 4 возможные направления градиента от 0 до 135° с шагом 45° и выбирается то, на котором достигается максимальное значение модуля градиента яркости. Далее сравниваются значения модулей градиента g_{mn} в обрабатываемой точке и в двух ближайших точках по линии направления α_{mn} . Точка (m,n) считается контурной, если g_{mn} больше, чем значения модулей градиента в обеих соседних точках по выбранной линии.

Множество соседствующих контурных точек можно рассматривать как единую область, а весь набор контурных точек — как множество областей. Для устранения возможных небольших разрывов между областями, соответствующими одному контуру, следует выполнить морфологическую операцию замыкания, рассмотренную в п. 6.3.1. Однако поскольку в дальнейшем следует процедура скелетизация областей, состоящая из последовательных операций утончения, то достаточно выполнить только первую часть замыкания — операцию дилатации (6.3.2).

6.3.3. Формирование первичной карты контурных линий

Преобразование области в линию означает получение *остова области* [40,124], который формируется путем *утончения (скелетизации)* области. Алгоритм утончения является двухпроходным. Он оперирует с бинарным изображением, элементы которого могут иметь значения либо 0 либо 1, однако в процессе его работы вводится третье значение, которое возникает на первом проходе и уничтожается на втором.

На первом проходе для каждого контурного пикселя x_{mn} анализируется восьмерка его соседей $V_8(m,n)$, которая интерпретируется как кольцо. Формируются всевозможные тройки, состоящие из соседних элементов кольца (ясно, что таковых всего 8). Составляются пары из не соседствующих троек, расположенных на противоположных сторонах кольца. Если имеется

хотя бы одна пара, в которой одна тройка состоит из всех контурных элементов, а вторая — из всех фоновых элементов, то тогда центральный пиксель x_{mn} помечается как *промежуточный* (т.е. ни фоновый, ни контурный). На втором проходе последовательно просматриваются все пиксели изображения и те, которые отмечены как промежуточные, стираются и становятся фоновыми. Алгоритм утончения итеративный и оканчивается тогда, когда на первом проходе ни один пиксель не переведен в ранг промежуточного. В результате формируются линии с 4-смежным соседством пикселей [40], допускающим только вертикальные и горизонтальные соседства пикселей.

Интерес представляют контурные линии, проходящие на небольшом расстоянии от сформированных в п. 6.3.1 опорных областей. Поэтому линии, отстоящие от какой-либо из опорных областей далее, чем заданный размер капилляра D , удаляются. В итоге формируется *первичная карта контурных линий*. Результат операций дилатации набора контурных точек, утончения и удаления далеко расположенных линий показан на Рис. 6.3.1,б.

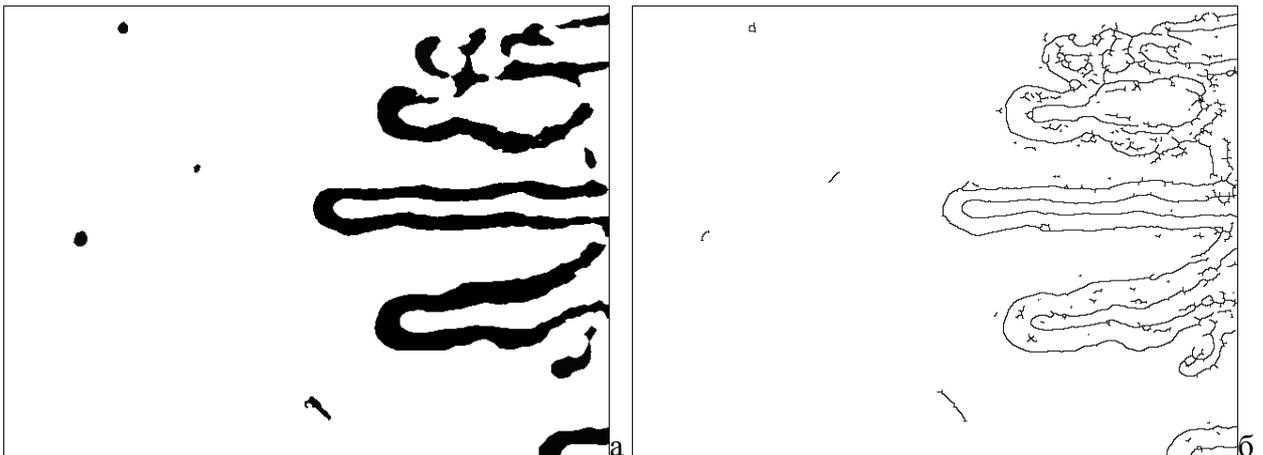


Рис. 6.3.1. Карта опорного множества областей (а) и первичная карта контурных линий (б).

6.3.4. Синтаксическая фильтрация и корректировка линий

Первичная карта контурных линий имеет целый ряд недостатков: количество линий все еще слишком велико, причем многие из них малой длины; кроме того имеется много ветвлений, образующих короткие ветви и петли различной сложности. Необходима корректировка и упрощение этого набора линий.

Удаление линий малой длины. Аналогично максимально допустимому числу опорных областей N_R на изображении, вводится ограничение на допустимое число контурных линий $N_L = 2N_R$. Для выполнения операции удаления, контурные линии интерпретируются как области. Той же операцией, что и выше в случае областей, осуществляется анализ размеров

линий и производится их упорядочение по убыванию размеров. Линии, получившие номер более, чем N_L , или имеющие размер менее, чем $1,5D$, удаляются.

Удаление коротких ветвей. Просмотром изображения составляется список точек ветвления линий и длин ветвей. Ветви, длина которых меньше предела D , стираются.

Корректировка линий. Карта линий содержит ряд некорректных участков. Анализируется топология линий и осуществляется корректировка. Дефекты, часть из которых показаны на Рис. 6.3.2,а-г, можно классифицировать так: концевая петля (а), внутриконтурная петля вида **О** (б), ветвление вида **У** с короткой ветвью, межконтурное соединение вида **Н**, внутриконтурное соединение вида **θ** (в), дефект сложного вида (г). Цель — приведение линий в незамкнутый вид без ветвлений для представления каждой линии последовательным списком точек, каждая из которых имеет не более двух соседей.

Алгоритм корректировки итерационный. Первоначально находятся координаты всех особых точек — точек ветвления и терминальных точек T концов линий, после чего осуществляется циклическая проверка принадлежности точек ветвления каждому из видов поочередно. Переход к следующему виду дефектов происходит, если отсутствуют особые точки предыдущего вида. Полный цикл итераций завершается, когда не оказалось ни одной особой точки последнего вида.

Корректировка начинается с выбора вида и проверки соответствия ему каждой точки ветвления; при совпадении лишняя ветвь стирается, а координаты граничных точек такой ветви удаляются из списка особых точек. Рассмотрим очередную особую точку, состоящую из 3-х ветвей. Вычисляется длина каждой ветви, исходящей из центральной точки, до ее окончания (терминальной точки T) или до следующей точки ветвления. Обозначим ветви через B_i ($i = 0,1,2$), их длины через L_i , координаты обрабатываемой особой точки через C , а ближайшей особой точки на ветви i — через C_i . Если $L_i > 8D$, то такая ветвь считается длинной без особой точки. Проверяются и обрабатываются следующие случаи.

1. *Концевая петля* (Рис. 6.3.2,а). Если $C_i = C$ при $L_i < 5D$ ($i = 0,1,2$), то ветвь считается концевой петлей и удаляется (Рис. 6.3.2,д).
2. *Внутриконтурная петля вида О* (Рис. 6.3.2,б). Если $C_i = C_j$ ($i = 0,1,2; j \neq i$) и $C_i \neq C$, то ветви B_i и B_j образуют петлю вида **О**. Удаляется наиболее длинная из двух ветвей (Рис. 6.3.2,е).
3. *Вилка* (ветвление вида **У**). Если одна или несколько из ветвей B_i заканчиваются терминальной точкой T , то удаляется наиболее короткая из них — на которой минимизируется значение L_i .
4. *Межконтурное соединение вида Н*. Рассматривается случай, когда только одна из ветвей (i) имеет точку ветвления, а любая из двух других либо не имеет особой точки, либо имеет терминальную точку вида T . В таком случае проверяются ветви ближайшей точки ветвления C_i .

Если для них также выполняется условие, что только одна из ветвей имеет точку ветвления (ясно, что это может быть только предыдущая точка C), а любая из двух оставшихся либо не имеет особой точки, либо имеет терминальную точку вида T , то такая пара точек $[C, C_i]$ является соединением вида \mathbf{H} и линия между C и C_i удаляется.

5. *Внутриконтурное соединение вида $\mathbf{\theta}$* (Рис. 6.3.2,в). Рассматривается точка ветвления C , у которой каждая из ветвей также оканчивается точкой ветвления. Проверяется, что среди этих трех точек ветвления имеется такая C_i , что 2 оставшиеся ее ветви также заканчиваются точками ветвления, причем они попарно совпадают с оставшимися двумя точками ветвления исходной точки C . Такая конструкция считается соединением вида $\mathbf{\theta}$ и отрезок между C и C_i удаляется. Тем самым соединение вида $\mathbf{\theta}$ переводится во внутриконтурную петлю вида \mathbf{O} , устранимую на следующем итерационном шаге (Рис. 6.3.2,ж).

6. *Удаление оставшихся особых точек.* Сложные структуры, как, например, на Рис. 6.3.2,г, обычно удается упростить при помощи рассмотренных выше вариантов корректировки, но иногда этого не происходит. Тогда производится стирание оставшихся точек ветвлений; линии в этих точках разрываются, после чего короткие отрезки удаляются (Рис. 6.3.2,з). В результате все линии приводятся в незамкнутый вид без ветвлений. Образовавшиеся короткие линии будут удалены на следующем этапе.

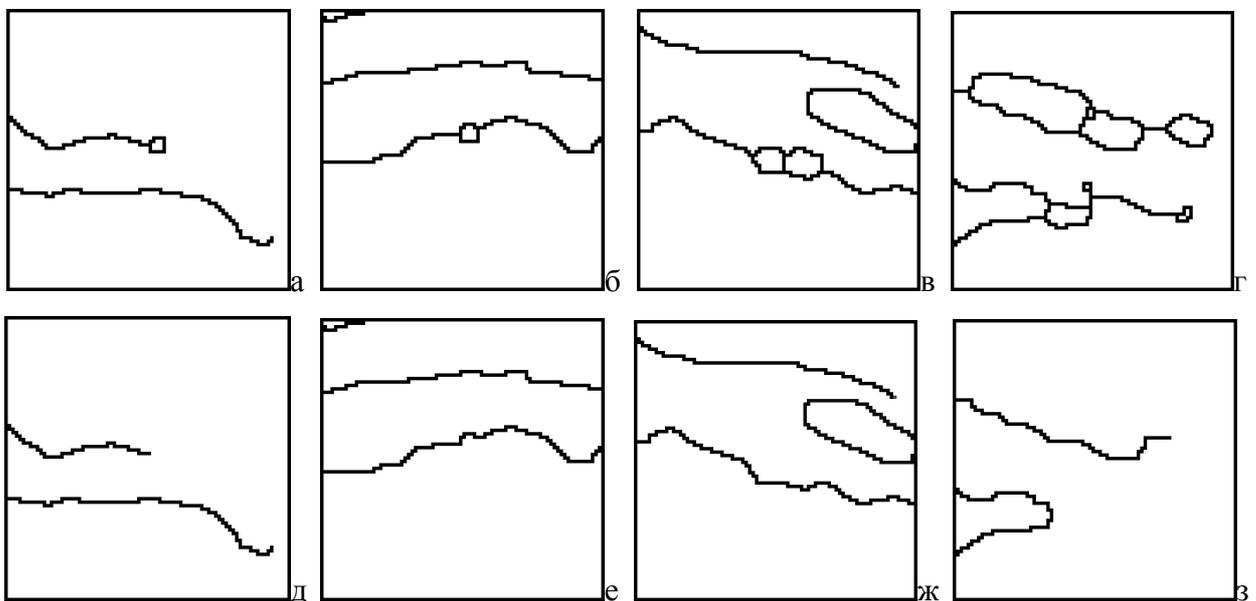


Рис. 6.3.2. Дефекты линий: концевая петля (а), внутриконтурная петля (б), соединение вида $\mathbf{\theta}$ (в), дефект сложного вида (г). Результаты корректировки: (д)-(з).

6.3.5. Упрощение карты линий и карты областей

Цель данного этапа — дальнейшее упрощение карты линий и карты областей путем удаления коротких линий, а также линий и небольших областей, расположенных далеко от

областей наибольшей площади. Для этого на основе сопоставления текущей карты линий и карты опорных областей, полученной в разделе 6.3.1, подсчитываются расстояния между областями и контурными линиями. На основе этого строится матрица расстояний, каждый элемент которой равен минимальному расстоянию между парой “область–линия”.

Сформированная матрица анализируется. Линии, отстоящие от ближайшей области, а также области, отстоящие от ближайшей линии на расстояние далее, чем D , помечаются как удаляемые. Также помечаются области, площадь которых не превышает D^2 , и линии, размеры которых не превышают значения $4D$. Помеченные области и линии удаляются. На Рис. 6.4.1,а представлен результат корректировки и упрощения карты линий.

6.4. Выбор мажоритарного капилляра

Под мажоритарным понимается капилляр, являющийся наиболее выраженным на изображении и подходящим для последующего анализа параметров кровотока. При выборе мажоритарного капилляра, характеризующегося совокупностью соответствующих ему опорных областей и контурных линий, принимались во внимание критерии:

- опорная область должна быть наибольшей (или одной из наибольших) по размеру;
- опорная область должна иметь вытянутую форму;
- оптимальное число контурных линии, соседствующих выбираемой области, — две;
- сумма длин соседствующих контурных линий должна быть близка периметру области;
- приоритет расположения капилляра — ближе к центру изображения.

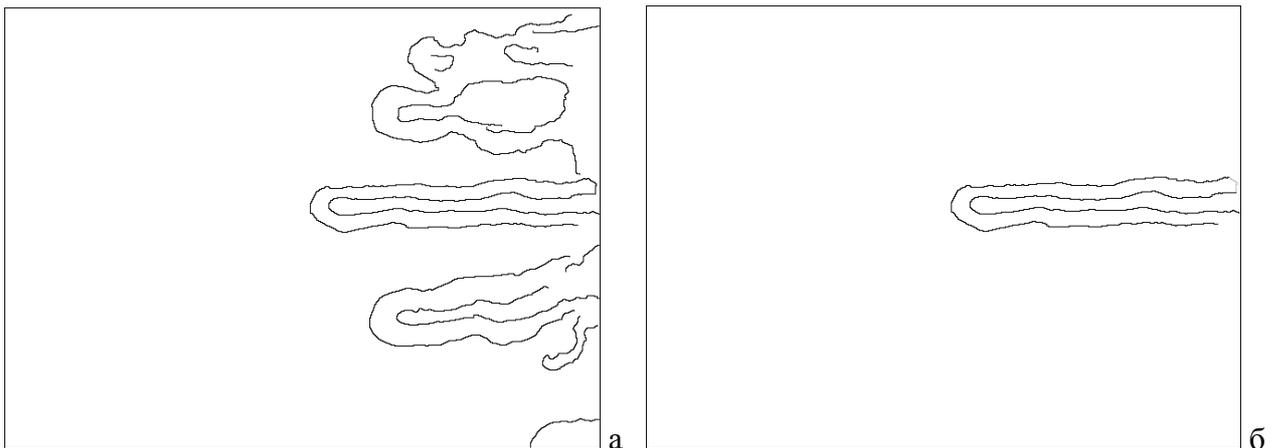


Рис. 6.4.1. Результирующая карта линий (а); контурная линия мажоритарного капилляра (б).

Для учета указанных критериев был подобран так называемый «показатель мажоритарности» в виде функционала $F(r)$, вычисляемого для каждой из опорных областей r , $r = 1, \dots, R$:

$$F(r) = S(r)E(r)N(r)D(r)L(r). \quad (6.4.1)$$

Параметры S , E , N , D и L определяются следующим образом:

$S(r)$ — площадь опорной области (число содержащихся в ней пикселей).

$E(r)$ — параметр вытянутости: $E(r) = \ln(1 + P(r) / \sqrt{S(r)})$, где $P(r)$ — периметр области r .

$N(r)$ — параметр числа соседних с областью линий n : $N(r) = 1/(\ln(i + 2) + 1)$, где $i = n$, если $n \geq 2$, иначе $i = 4 - n$; максимум находится в точке $n = 2$.

$D(r)$ — параметр соотношения периметра области $P(r)$ и суммы длин соседних контуров C_{len} : если $P(r) < C_{len}$, то $D(r) = P(r)/C_{len}$; иначе $D(r) = C_{len}/P(r)$.

$L(r)$ — показатель расположения области: равен 1, если центр области r расположен от края далее, чем $1/3$ размера изображения, и спадает до 0 к краю изображения.

Мажоритарной считается область, на которой достигается максимум функционала $F(r)$. Все остальные области, а также контурные линии, не соседствующие с мажоритарной областью, удаляются. На Рис. 6.4.1,б показана оставшаяся контурная линия, соответствующая выбранному мажоритарному капилляру, а все остальные линии удалены.

Формируемые линии границ не всегда оказываются корректными. Как видно на Рис. 6.4.1,б, мажоритарным выбран капилляр, окруженный всего одной контурной линией (в данном случае контурная линия показана двумя черными линиями и соединяющим их серым отрезком). Подобный дефект требует устранения. Морфологический анализ и корректировка дефектов контурных линий рассматриваются в следующем разделе.

6.5. Морфологический анализ и корректировка контурных линий

6.5.1. Векторное представление контурных линий.

Для анализа формы линий перейдем от растрового к векторному их описанию — в виде последовательностей соединенных точек. Представим линию L как упорядоченный набор точек $p(z) \in L$, ($z = 1, \dots, Z$), где Z — общее число точек (длина) линии L . Каждая точка $p(z)$ соответствует пикселю $x(m(z), n(z))$ изображения с координатами $m(z)$ и $n(z)$, т.е. между ними допустимо ставить тождество. В соответствии со способом формирования контурных линий (раздел 6.3.3), соседние точки линии $p(z)$ и $p(z+1)$ находятся в 4-смежном пиксельном соседстве.

Выберем подмножество L' множества точек линии L из элементов $q(z) \in L'$, взятых с шагом d : $q(z) = p(z \cdot d)$, ($z = 1, \dots, Z'$), где $Z' = Z/d$. Проведем векторы $\mathbf{V}(i)$, ($i = 1, \dots, Z'-1$) между парами точек $q(i)$ и $q(i+1)$: $\mathbf{V}(i) = \overline{q(i)q(i+1)}$. Обозначим через $\gamma(i)$, ($i = 1, \dots, Z'-2$) угол между векторами $\mathbf{V}(i)$ и $\mathbf{V}(i+1)$, а через $\beta(i)$ — величину накопленного угла: $\beta(i) = \sum_{j=1}^i \gamma(j)$. Приводя среднее значение $\beta(i)$ к нулю, получим последовательность:

$$\alpha(i) = \beta(i) - \sum_{z=1}^{Z-2} \beta(z)/(Z-2). \quad (6.5.1)$$

Найденную последовательность $\alpha(i)$, ($i = 1, \dots, Z-2$) назовем *морфологической характеристикой линии*; ее анализ позволяет в дальнейшем делать выводы о форме линии.

6.5.2. Анализ морфологической характеристики контурной линии

Морфологическая характеристика линии $\alpha(i)$ является основной информацией для анализа формы контурной линии. Рассмотрим этот вопрос на примере морфологической характеристики ранее полученной контурной линии мажоритарного капилляра (Рис. 6.4.1,б). График данной функции показан на Рис. 6.5.1; по горизонтали отложен номер i точки подмножества L' , а по вертикали — значение $\alpha(i)$ в градусах.

Как можно видеть, график содержит четыре почти горизонтальных участка (0-55, 80-140, 155-210 и 225-285), разделенных тремя ступеньками. Примечательно, что средние значения на соседних горизонтальных участках различаются приблизительно на 180° . Интерпретировать эту особенность несложно. Горизонтальные участки графика соответствуют длинным прямым участкам контурной линии (ограничивающим артериальный и венозный отделы капилляра), которые между собой почти параллельны, а ступеньки между ними — участкам поворота линии (переходный отдел капилляра). Знак градиента на ступеньке говорит о направлении изгиба линии на данном участке.

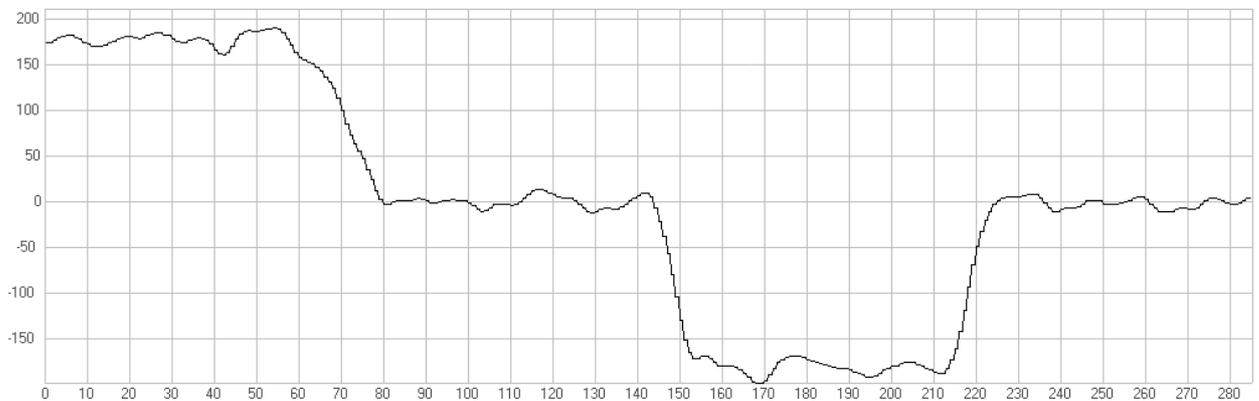


Рис. 6.5.1. Морфологическая характеристика контурной линии на Рис. 6.4.1,б.

Как нетрудно понять, морфологическая характеристика корректной контурной линии должна содержать два горизонтальных участка, соответствующих артериальному и венозному отделам капилляра, и один переходный участок, соответствующий переходному отделу. На основе этого можно сделать заключение, что кривая на Рис. 6.5.1 составлена из характеристик двух контурных линий, которые соединены участком поворота при значениях $i \approx 145$. Если удалить из контурной линии участок, соответствующий данной области морфологической характеристики, получим две корректные контурные линии, окружающие капилляр.

Рассматривая возможные случаи, можно сформулировать следующий алгоритм корректировки морфологической характеристики контурной линии.

1. В последовательности значений $\alpha(i)$ отыскиваются переходные области TR как участки высокой крутизны, на которых $\alpha(i)$ меняется почти на 180° : $TR(i)$, ($i = 1, \dots, K$). Здесь K — общее число переходных областей; определяются также координаты их начала и конца.
2. Проверяется найденное значение K . Если $K = 1$, то значит рассматриваемая контурная линия соответствует модели внешнего или внутреннего контура капилляра.
3. Случай $K = 2$ означает, что линия имеет дополнительный поворот. Он может возникнуть, если при первичном формировании контуров происходит захват края видимой области капилляра. Такой поворот обычно располагается на одном из концов линии. В такой ситуации выбирается переходная область $TR(i)$, располагающаяся ближе к одному из концов контурной линии, и соответствующий участок линии удаляется.
4. Случай $K = 3$ означает объединение двух контурных линий, вариант которого показан на Рис. 6.4.1,б. Подобное объединение возможно, если на одном из краев области капилляра детектируется ложный контурный участок, приводящий к соединению внутренней и внешней контурных линий капилляра. Такой участок, соответствующий средней переходной области $TR(2)$, удаляется (отмечен серым цветом на контуре мажоритарного капилляра на Рис. 6.4.1,б). В результате картина линий приводится в корректное состояние. В такой ситуации требуется пересчет морфологических характеристик образованных линий, т.е. возврат к п. 6.5.1.
5. Если же $K = 0$, то анализируемая контурная линия не соответствует принятой модели, т.е. контур капилляра не был правильно обнаружен. Вероятно, в такой ситуации требуется повторное формирование исходной видеопоследовательности.

6.5.3. Корректировка расположения контурных линий

Как говорилось выше, корректная контурная линия состоит из двух почти параллельных участков и участка поворота между ними — переходной области. Можно сказать, что с точностью до поворота контурная линия повторяет форму буквы U (т.е. имеет U-форму). Карта контурных линий мажоритарного капилляра должна включать две контурные линии — внутреннюю и внешнюю. Будем считать каноническим такое расположение, при котором первой является внутренняя, а второй — внешняя линия, причем последовательность точек векторного представления каждой из контурных линий упорядочена таким образом, что начало ее находится в вершине левой ветви U-формы.

Для приведения линий в каноническое расположение используем уже полученные на предыдущих этапах результаты и вычисленные морфологические характеристики. Корректная

контурная линия имеет лишь одну переходную область. Обозначим через TR_{beg} и TR_{end} координаты начала и конца переходной области линии. Введем параметр крутизны перехода:

$$G = (\alpha(TR_{end}) - \alpha(TR_{beg})) / (TR_{end} - TR_{beg}). \quad (6.5.2)$$

Обратим внимание, что абсолютная величина G будет обратно пропорциональна радиусу закругления переходной области. Если значение $G > 0$, то контурная линия ориентирована правильно и может быть оставлена без изменения. Если $G < 0$, то линия является обратно ориентированной и последовательность точек должна быть зеркально перевернута. Также соответствующим образом должны быть пересчитаны значения $\alpha(i)$, TR_{beg} и TR_{end} .

Поскольку внешняя контурная линия капилляра является огибающей по отношению к внутренней линии, то радиус закругления переходной области внешней линии должен быть больше, чем радиус закругления переходной области внутренней линии. На основании этого достаточно сравнить абсолютные значения G_1 и G_2 двух контурных линий. Если $|G_1| > |G_2|$, то первая линия — внутренняя и расположение является каноническим. Если же $|G_1| < |G_2|$, то необходимо переставить местами линии вместе со своими параметрами. Полученные в результате таких преобразований графики морфологических характеристик внутренней и внешней контурных линий представлены на Рис. 6.5.2.

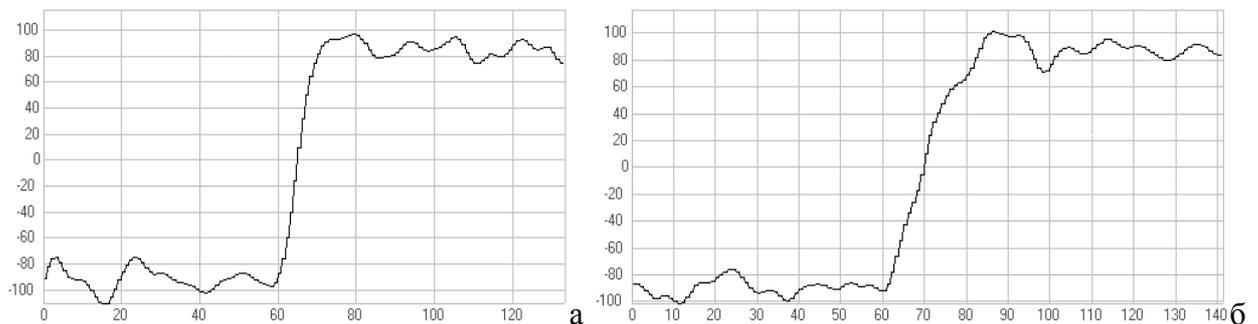


Рис. 6.5.2. Морфологические характеристики внутренней (а) и внешней (б) контурных линий.

6.6. Построение распрямляющего отображения

Для определения скорости по видеоданным обычно применяется корреляционный подход [301]. Мгновенная скорость объекта определяется как $v = \Delta s / \Delta t$, где Δs — величина межкадрового смещения объекта, а Δt — период кадра, обычно постоянный во время наблюдения. Для корреляционного анализа необходимо использовать единую аналитически задаваемую область анализа, однако выбрать такую для фигуры сложной конфигурации, каковой является капилляр, затруднительно. Решением может быть нахождение способа геометрического преобразования области капилляра в область простой формы, на которой затем и будет выполняться корреляционный анализ с использованием прямоугольной окрестности.

Задача ставится как нахождение гомеоморфного отображения связной области, ограниченной двумя параметрически задаваемыми кривыми линиями, концы которых соединены прямыми, в область прямоугольной формы. Предполагаем также, что граничные линии удовлетворяют некоторым условиям, в частности таким как отсутствие пересечений и самопересечений (*условие-1*). Еще одно условие, формулируемое ниже, возникает естественным образом при построении преобразования.

Будем считать, что исходная область ограничена парой непрерывных параметрически заданных кривых **A** и **B**: $a(l)$, $(0 \leq l \leq L_A)$ и $b(l)$, $(0 \leq l \leq L_B)$, а также отрезками $[a(0), b(0)]$ и $[a(L_A), b(L_B)]$ (Рис. 6.6.1,а). Задача — преобразовать данную область в прямоугольник размерами $(M+1) \times (N+1)$ (Рис. 6.6.1,б).

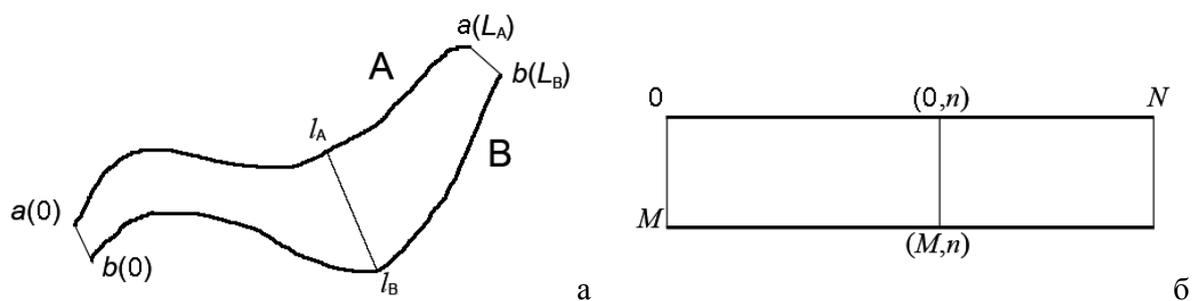


Рис. 6.6.1. Иллюстрация распрямляющего преобразования.

Преобразование осуществим следующим образом. Отобразим кривые **A** и **B** линейно в верхнюю и нижнюю строки результирующего прямоугольника. Выберем столбец n массива результата: $[(0,n), \dots, (M,n)]$. В начальную его точку $(0,n)$ отобразим точку первой кривой $a(l_A)$, $l_A = (n/N)L_A$ с координатами (x_A, y_A) на исходном изображении, а в конечную (M,n) — точку $b(l_B)$, $l_B = (n/N)L_B$ второй кривой с координатами (x_B, y_B) . Промежуточная точка столбца результирующего массива с координатами (m,n) будет отображением точки исходной области с координатами (x,y) , где $x = (M - m)x_A + mx_B$, а $y = (M - m)y_A + my_B$. Таким образом отрезок $[l_A, l_B]$ исходной области пропорционально отображается в столбец $[(0,n), \dots, (M,n)]$ результата. Теперь очевидна необходимость *условия-2*, а именно: никакой из отрезков $[l_A(n), l_B(n)]$ не должен иметь дополнительных пересечений с кривыми **A** и **B**. Легко видеть, что данное условие зависит не только от формы каждой из кривых, но и от их взаимного расположения.

Полученное в результате отображение будем называть *распрямляющим*.

Отрезки $[l_A(n), l_B(n)]$ по сути являются сечениями капилляра. Важнейшим следствием, непосредственно вытекающим из *условия-2*, является то, что никакие сечения не пересекаются между собой. Проводимые подобным образом сечения вообще говоря не будут пересекать по нормали некоторую гипотетическую срединную линию капилляра (о ней речь ниже). Однако рассматривая капилляр как трубу с жидкостью, легко видеть, что объем жидкости,

протекающий за единицу времени через сечение трубы, не зависит от наклона сечения. Поэтому для нашей задачи указанный факт не внесет погрешности.

6.7. Определение границ отделов капилляра

Длины отделов внутренней и внешней контурных линий могут заметно отличаться, причем длины переходных отделов могут различаться в два и более раз. Это может привести к неблагоприятному расположению сечений и нарушению *условия-2*. Уменьшить вероятность такого события можно выполняя распрямляющее преобразование по отдельности для каждого из трех отделов капилляра (артериального, венозного и переходного), и затем соединяя результирующие прямоугольные массивы в один. Для разграничения указанных отделов необходимы дополнительные построения.

Срединная линия капилляра находится как множество точек, равноудаленных от каждой из двух контурных линий капилляра. Подобная линия формально может выходить за пределы реальной области капилляра. Чтобы избежать этого, анализируется множество пар точек контурных линий, ближайших для каждой точки срединной линии. Те точки срединной линии, для которых одной из ближайших является точка начала или конца любой из контурных линий, считаются выходящими за границы капилляра и удаляются. Дополнительно производится небольшое укорочение концов срединной линии (Рис. 6.7.1,а).

Найденная срединная линия $c(l)$, $(0 \leq l \leq L_C)$ служит основой для определения границ отделов капилляра. Так же, как и при анализе контурной линии в п. 6.5.3, находятся точки начала и конца переходного отдела срединной линии. Таким образом мы имеем четыре ее характерные точки: точки начала и конца линии c_0 и c_{end} , а также точки начала и конца переходного отдела c_{Tbeg} и c_{Tend} . Найдем на контурных линиях капилляра **A** и **B** точки, являющиеся ближайшими к характерным точкам срединной линии: a_0 , a_{Tbeg} , a_{Tend} , a_{end} и b_0 , b_{Tbeg} , b_{Tend} , b_{end} . Отрезки $[a_{Tbeg}, b_{Tbeg}]$ и $[a_{Tend}, b_{Tend}]$ будем считать границами начала и конца переходного отдела капилляра, а отрезки $[a_0, b_0]$ и $[a_{end}, b_{end}]$ — соответственно началом и концом капилляра.

Зная границы всех трех отделов капилляра, можно выполнить распрямляющее отображение для каждого из них по отдельности, объединить получаемые результаты и сформировать полное распрямляющее отображение для всего капилляра. Срединная линия и найденные границы отделов показаны на Рис. 6.7.1,а, а общая картина образуемых сечений капилляра, которые на результирующем прямоугольном массиве будут отображаться столбцами с равным шагом, — на Рис. 6.7.1,б. Отметим, что в данный момент мы еще не можем различить артериальный и венозный отделы капилляра.

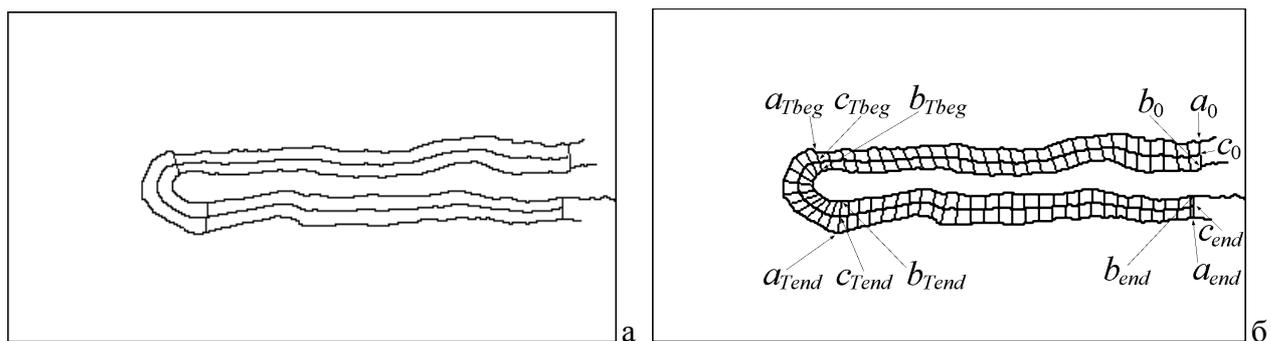


Рис. 6.7.1. Срединная линия и границы отделов капилляра (а); сечения, иллюстрирующие распрямляющее преобразование (б).

6.8. Преобразование кадров видеопоследовательности

Размеры M и N результирующего массива в принципе можно выбрать любыми, но естественно взять их равными размерам самого капилляра. Поэтому M выбиралось как среднее значение видимого диаметра капилляра, а N — как длина срединной линии. Строится матрица отображения размерами $(M+1) \times (N+1)$, в каждый элемент которой заносятся координаты соответствующей точки на исходном (синхронизованном) кадре. Согласно построенной матрице, по каждому из исходных кадров формируется кадр новой видеопоследовательности, размер которого равен размеру матрицы отображения. Поскольку координаты отображаемой точки на исходном изображении оказываются дробными, значение яркости вычисляется методом билинейной интерполяции по четырем ближайшим соседям.

В результате преобразований формируется новая видеопоследовательность, содержащая только информативную часть исходных видеоданных и имеющая равное с ней число кадров. Тем самым исходные данные приводятся к виду, удобному для анализа динамики кровотока: определению скорости движения и объема протекающей крови, а также выполнению иных необходимых измерений. На Рис. 6.8.1,а показан кадр с изображением распрямленного капилляра, полученный из исходного кадра, представленного на Рис. 6.0.1. Обратим внимание, что ориентация деталей в результирующем кадре соответствует определенному выше каноническому расположению контурных линий.

Видимые на полученном изображении темные участки суть группы эритроцитов; они разделены более светлыми участками, в которых находятся прозрачные лейкоциты и плазма крови. В такой ситуации определять скорость протекания крови фактически возможно только по фронтам видимых групп эритроцитов [302]. Однако контраст деталей во многих случаях низкий и недостаточный для надежного сопоставления фрагментов, поэтому требуется дополнительная обработка кадров. Она сводится к выравниванию яркости, пороговому разделению сигнала и низкочастотной фильтрации.

Выравнивание яркости и пороговое разделение можно объединить в одну операцию:

$$y_{mn} = \begin{cases} 0, & \text{если } x_{mn} \leq S_{mn}, \\ 255, & \text{если } x_{mn} > S_{mn}. \end{cases} \quad (6.8.1)$$

Здесь S_{mn} — значение среднего по локальному фрагменту (6.1.5), размеры которого выбирались исходя из размеров капилляра как $H = M/8$, $W = 2M$. Результатом операции (6.8.1) является бинаризованное изображение. Поскольку межкадровая изменчивость форм групп эритроцитов высока, для ослабления влияния возникающей погрешности осуществлялась низкочастотная фильтрация при помощи операции двумерного гауссова сглаживания полученного изображения $Y = \{y_{mn}\}$ с параметрами $\sigma \approx 2$ по вертикали и $\sigma \approx 5$ по горизонтали. Результат показан на Рис. 6.8.1,б.



Рис. 6.8.1. Вид капилляра из кадра на Рис. 6.0.1 после распрямляющего преобразования (а); тот же кадр после обработки (б).

6.9. Определение параметров кровотока

Основным параметром, отыскание которого и представляет главную цель исследования, является скорость кровотока. Для измерения скорости вдоль капилляра воспользуемся функцией ковариации:

$$R_k(n, t) = \frac{1}{(M+1)(2Z+1)} \sum_{m=0}^M \sum_{z=-Z}^Z b_k(m, n+z) b_{k-1}(m, n+z+t),$$

где $b_k(m, n)$ и $b_{k-1}(m, n)$ — значения элементов текущего и предыдущего кадров в точке (m, n) , $(M+1) \times (2Z+1)$ — размер окрестности анализа, $R_k(n, t)$ — значение функции ковариации в точке n для смещения t . Анализируемый диапазон значений t также был выбран равным размеру окна: $-Z \leq t \leq Z$. Заметим, что несмотря на то, что нас интересует скорость только в одном направлении — вдоль капилляра, значение ковариации считается по двумерной окрестности. Величина межкадрового смещения для точки n определяется как такое значение t_n , при котором достигается максимум функции ковариации:

$$d_k(n) = t_n \mid R_k(n, t_n) = \max_{-Z \leq t \leq Z} (R_k(n, t)).$$

Межкадровое смещение $d_k(n)$ находится для каждой точки n вдоль капилляра и для каждого кадра k . Тем самым формируется двумерный массив размерами $K \times N$, где K — число кадров видеопоследовательности, а N — длина капилляра. Делением значения $d_k(n)$ на время между кадрами получаем значение мгновенной скорости $V(k, n)$ как функции номера кадра и

положения вдоль капилляра. Из-за высокой изменчивости форм групп эритроцитов от кадра к кадру, о которой уже говорилось, получаемые значения смещений (скорости) оказываются сильно зашумлены. Из физической интерпретации задачи ясно, что значения скорости не могут иметь скачков и изменяются достаточно плавно как по времени, так и вдоль капилляра. Это позволяет для подавления шума применить фильтрацию низких частот при помощи операции двумерного сглаживания (2.1.2). Сформированная и сглаженная карта мгновенных скоростей представлена на Рис. 6.9.1.

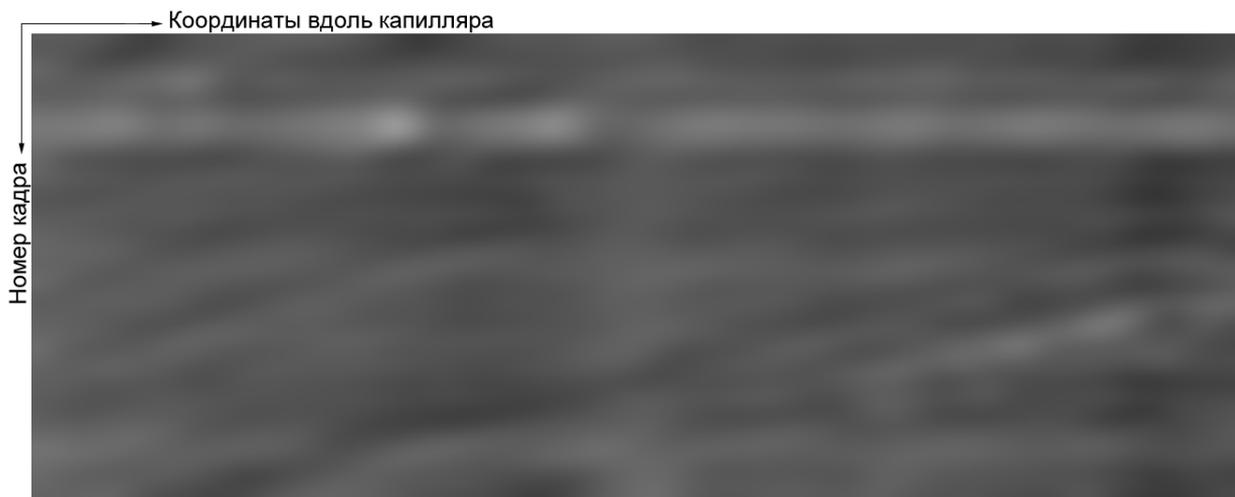


Рис. 6.9.1. Карта мгновенных скоростей.

Полученный двумерный массив мгновенных скоростей может быть использован для анализа динамики кровотока. Так, среднее значение по строкам (усреднение вдоль капилляра) дает зависимость средней скорости в капилляре от времени. На Рис. 6.9.2 она представлена как зависимость межкадрового смещения в элементах изображения от номера кадра. Знак значения скорости указывает на направление кровотока. Как видно, в районе 50-го кадра скорость движения крови уменьшается почти до нуля. Средние отрицательные значения, как в данном случае, позволяют идентифицировать левую ветвь U-формы капилляра (верхний горизонтальный участок на Рис. 6.7.1,а) как его венозный, а правую (нижний горизонтальный участок) — как артериальный отделы.

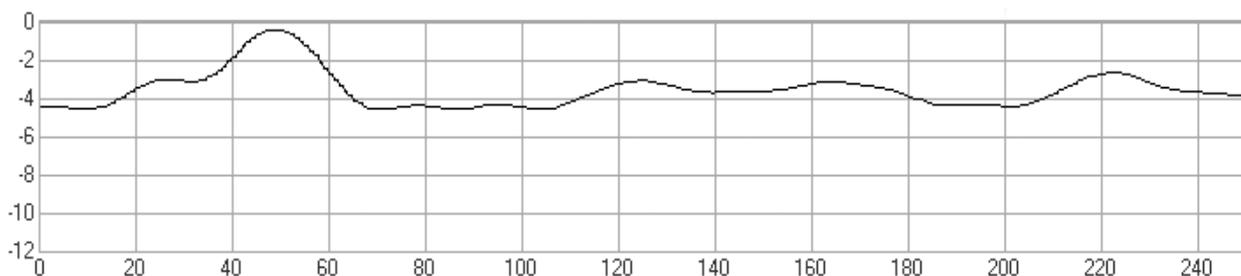


Рис. 6.9.2. Зависимость средней скорости в капилляре от времени (от номера кадра).

Усреднение карты мгновенных скоростей по столбцам (т.е. по кадрам) дает среднюю скорость в разных точках капилляра. Зная масштаб съемки, на основе имеющихся контурных линий обнаруженного капилляра легко вычислить размер, диаметр и площадь сечения капилляра. Умножением значения площади сечения на скорость потока находится мгновенный объем кровотока в соответствующей точке капилляра.

Исследование задачи и разработка методов обработки и автоматического анализа видеопоследовательностей, формируемых компьютерным капилляроскопом, проводились в сотрудничестве с ООО «Новые энергетические технологии» и ЗАО Центр «Анализ веществ». Экспериментальный анализ доступных видеопоследовательностей (51 шт.) показал, что вероятность правильного обнаружения мажоритарного капилляра, проведения контурных линий и определения скорости кровотока составляет около 87%.

Изложенные в настоящей главе результаты опубликованы в работах [162,163,177,178,181,226,230,232,233].

6.10. Выводы и результаты

Показаны возможности использования разработанных в главах 2–4 методов для решения задач обработки и анализа видеоданных.

Решена важная научно-техническая задача автоматического анализа и извлечения необходимой информации из видеопоследовательностей, формируемых компьютерным капилляроскопом, позволяющая в результате определять ряд статических и динамических параметров капиллярного кровотока. Разработан комплекс алгоритмов и программ для решения следующих подзадач: а) предварительной обработки видеок кадров; б) компенсации дрейфа видеопоследовательности и формирования усредненного по кадрам изображения; в) нахождения мажоритарного капилляра; г) проведения морфологического анализа контурных линий; д) выполнения гомеоморфного распрямляющего преобразования; е) определения динамических параметров капиллярного кровотока.

Процедура предварительной обработки видеок кадров включает устранение систематических искажений, фильтрацию периодических помех, выравнивание фона кадра и усиление локальных контрастов видеосигнала. Разработан алгоритм компенсации дрейфа кадров, осуществляющий их пространственную синхронизацию на основе разработанного быстрого алгоритма совмещения. Суммированием синхронизованных кадров формируется усредненное изображение, на котором существенно снижен уровень помех, что позволяет значительно уменьшить количество ошибок при дальнейшем автоматическом анализе. При необходимости

визуального анализа капилляров врачом-диагностом, возможно улучшение усредненного изображения одним из методов повышения локальных контрастов.

Анализ усредненного изображения включает построение карты контурных линий, ее синтаксическую фильтрацию, выбор мажоритарного капилляра, формирование векторного представления контурных линий и составление их морфологических характеристик для последующей интерпретации. В результате определяется область мажоритарного капилляра, ограниченная параметрически заданными линиями. Получаемые морфологические характеристики позволяют делать выводы о форме капилляра, а также определять его статические параметры — диаметр и границы отделов.

Предложен способ выполнения гомеоморфного распрямляющего преобразования области изображения, часть границ которой составляют параметрически заданные и аналитически не описываемые линии. Такое преобразование обеспечивает эластичную деформацию и гладко отображает область капилляра в область прямоугольной формы. Используя полученную матрицу отображения из исходных видеокадров формируется новая видеопоследовательность, содержащая только интересующую выровненную область мажоритарного капилляра, по которой и производится корреляционный анализ.

Разработан алгоритм формирования карты мгновенных скоростей, которая служит основой для определения динамических параметров капиллярного кровотока, в том числе скорости и объема протекающей крови как функций времени и местоположения в капилляре.

Экспериментальный анализ показал, что вероятность правильного обнаружения мажоритарного капилляра, проведения контурных линий и определения скорости кровотока разработанным методом составляет около 87%.

ГЛАВА 7. МОДИФИКАЦИЯ МОДЕЛИ И МЕТОДОВ ОБРАБОТКИ ДЛЯ ТРЕХМЕРНЫХ ИЗОБРАЖЕНИЙ

Введение

Настоящая глава посвящена переходу от двумерных (2D) к трехмерным (3D) изображениям. Рассматриваются вопросы модификации двухмасштабной многокомпонентной модели изображения, изменения области анализа, базовых методов обработки видеoinформации, в том числе алгоритмов фильтрации, декомпозиции изображений и обнаружения объектов. Предложены модификации быстрых алгоритмов вычисления локального среднего и порядковых статистик по скользящему окну для 3D-изображений. Разработан параллельный алгоритм вычисления гистограммы и порядковых статистик по скользящему фрагменту для двумерных и трехмерных изображений.

Под термином «трехмерное изображение» зачастую понимают совершенно различные виды данных [72,248,303,322,357], основные из которых следующие.

1. Данные, задаваемые функцией трех координат и являющиеся отображением (проекцией) некоторого объемного участка реального или виртуального трехмерного пространства, включая все содержащиеся в нем объекты.
2. Стереоскопическое изображение — пара соответствующим образом получаемых двумерных изображений, которые за счет диспаратности дают представление о пространственном расположении объектов, видимых с точки наблюдения.
3. Изображение, являющееся некоторой двумерной проекцией объектов исходной трехмерной сцены (например, аксонометрической); оно дает наблюдателю представление о форме и расположении объектов, однако само по себе остается двумерным.
4. Двумерное изображение, каждая точка которого соответствует некоторым координатам в трехмерном пространстве, например, дальности или рельефу.
5. Особым способом формируемые двумерные изображения, создающие образы наблюдаемых объектов, например голограммы.
6. Видеопоследовательности, содержащие набор двумерных кадров наблюдаемых объектов. Такие данные могут представляться в виде трехмерного массива, аналогично первому типу, однако при этом одна из координат является не пространственной, а координатой времени.

В дальнейшем под *трехмерным* или *3D-изображением* (непрерывным или дискретным) будут пониматься изображения исключительно первого типа. По существу, 3D-изображение представляет собой расширение обычного 2D-изображения добавлением еще одного простран-

ственного измерения. Как и в случае двумерного изображения, предполагается, что изначальное трехмерное изображение можно рассматривать как непрерывную функцию вида $F(x,y,z)$, причем значение функции F в точке с пространственными координатами (x,y,z) является скалярной величиной, имеющей свойства плотности некоторой выбранной физической характеристики, смысл которой определяется источником формирования изображения.

Цифровое *трехмерное изображение* (или *3D-изображение*) как функция $f(m,n,k)$ в пространстве с дискретными координатами (m,n,k) , соответствующими координатам (x,y,z) непрерывного изображения, представляет собой трехмерный массив данных размерами $M \times N \times K$. Элементы 3D-изображения соответствуют точкам отображаемого участка непрерывного трехмерного пространства, а значение каждого элемента отражает величину выбранной физической характеристики в окрестности соответствующей точки пространства, усредненную согласно заданной апертурной функции. Иногда такие изображения также называют *объемными изображениями*. По аналогии с двумерным изображением, дискретный элемент которого называют *пиксель*, дискретный элемент 3D-изображения называют *воксель* (англ. *voxel* — volumetric pixel) — элемент объемного изображения.

Способы формирования трехмерного изображения могут быть совершенно различными; наиболее известным является томографическое сканирование — рентгеновская компьютерная томография (КТ) и магнитно-резонансная томография (МРТ). Возможно получение трехмерных изображений в результате сейсморазведки при проведении геологических исследований, в гистологическом анализе, в микроскопии при использовании объектива с переменным фокусным расстоянием, при компьютерном моделировании трехмерных объектов и сцен, или каким-то иным образом.

Обработка и анализ трехмерных изображений играют в настоящее время существенную роль во многих областях исследований, особенно в медицине и геологии. В данной главе будут рассмотрены вопросы расширения двухмасштабной многокомпонентной модели изображения, изложенной в разделах 1.1-1.2, и применения ее к трехмерным изображениям. Также будут рассмотрены вопросы модификации операций частотной и пространственной фильтраций при переходе в 3D, вопросы сглаживания и декомпозиции изображения [179,192], фильтрации помех, обнаружения контуров и объектов, а кроме того вычислительные аспекты реализации некоторых алгоритмов для трехмерных изображений.

7.1. Особенности трехмерных изображений

В дискретном виде 3D-изображение представляется трехмерным массивом $\mathbf{X} = [x_{mnk}]$ размерами $M \times N \times K$. Будем считать, что как и в 2D-изображении, значение каждого элемента x_{mnk}

есть квантованное на $(x_{\max}+1)$ градаций значение логарифма яркости (энергии) $0 \leq x_{mnk} \leq x_{\max}$, называемое просто *яркостью*. Трехмерное изображение, как средство отображения информации о некоторой наблюдаемой сцене или предмете, можно рассматривать состоящим из плотно упакованных связанных трехмерных областей (объектов), соответствующих деталям сцены или предмета. *Областью* или *объектом* 3D-изображения будем называть максимальное по размеру связанное множество его элементов, имеющих близкие, возможно плавно меняющиеся значения яркости. Области могут соприкасаться произвольным образом, в том числе одна область может быть полностью окружена другой. На границах соседних областей значения яркости должны заметно различаться. Не соприкасающиеся области могут иметь произвольные, в том числе и совпадающие яркости. *Контурами* при этом будем называть пространственные границы между соседними областями, как объектами различающейся яркости.

Отметим некоторые существенные свойства трехмерных изображений:

- в трехмерном изображении объект (3D-объект) — пространственная фигура, а само трехмерное изображение — пространственно ограниченная совокупность трехмерных объектов, плотно заполняющих пространство изображения;
- контуры в 3D-изображении суть пространственные границы между объектами;
- сечение 3D-изображения плоскостью любого направления, а также проекция трехмерного изображения на плоскость любого направления дают двумерный сигнал со всеми свойствами обычного 2D-изображения.

7.1.1. Области анализа и соседство элементов

Область анализа трехмерного изображения, как подмножество исходных данных, используемое в оценке параметров, является естественным расширением области анализа для двумерного изображения. По аналогии с определением, данным в разделе 1.1.2 для двумерного изображения, ее следует рассматривать как связанное множество элементов $x_{ijl} \in V_d(x_{mnk})$, отстоящих от центрального элемента x_{mnk} на расстояние не далее, чем d и вместе составляющих фигуру некоторой задаваемой формы. При $d \lesssim 1 \div 3$ множество $V_d(x_{mnk})$, окружающее центральный элемент (воксель) x_{mnk} , называется *окрестностью* и обозначается V_{mnk} , а при $d \gg 1$ — *фрагментом* и обозначается W_{mnk} . Также как и ранее, в зависимости от выполняемых операций сам центральный элемент x_{mnk} может как принадлежать, так и не принадлежать $V_d(x)$. Соответственно, локальные операции, в которых результат в каждой точке (m, n, k) зависит лишь от значений элементов x_{ijl} , входящих в $V_d(x_{mnk})$, описываются формулой

$$y_{mnk} = f\{x_{ijl} \mid x_{ijl} \in V_d(x_{mnk})\}. \quad (7.1.1)$$

При переходе из 2D в 3D варианты окрестностей и соседства элементов претерпевают следующие изменения. Окрестность из 2×2 элементов (4 пикселя) становится окрестностью из $2 \times 2 \times 2$ элементов (8 вокселей, Рис. 7.1.1,а), в которой каждый воксель соседствует с каждым. В двумерной окрестности из 3×3 элементов (9 пикселей), как известно, можно рассматривать два варианта соседства элементов: 4-соседство (только по сторонам пикселей) и 8-соседство (по сторонам и вершинам пикселей) [40]. Аналогом первого из них в 3D будет окрестность с 6-соседством вокселей (Рис. 7.1.1,б). Аналогом второго будет окрестность с 26-соседством вокселей (Рис. 7.1.1,г). Возможен также промежуточный вариант с 18-соседством вокселей (Рис. 7.1.1,в). Таким образом, в случае окрестности на Рис. 7.1.1,б соседними с центральным вокселем (который на рисунках не виден) считаются только те воксели, которые имеют с ним общие грани; в случае окрестности на Рис. 7.1.1,в — общие грани и ребра; а в случае окрестности на Рис. 7.1.1,г — общие грани, ребра и вершины. Выбор варианта окрестности обычно определяется контекстом задачи и используемым алгоритмом.

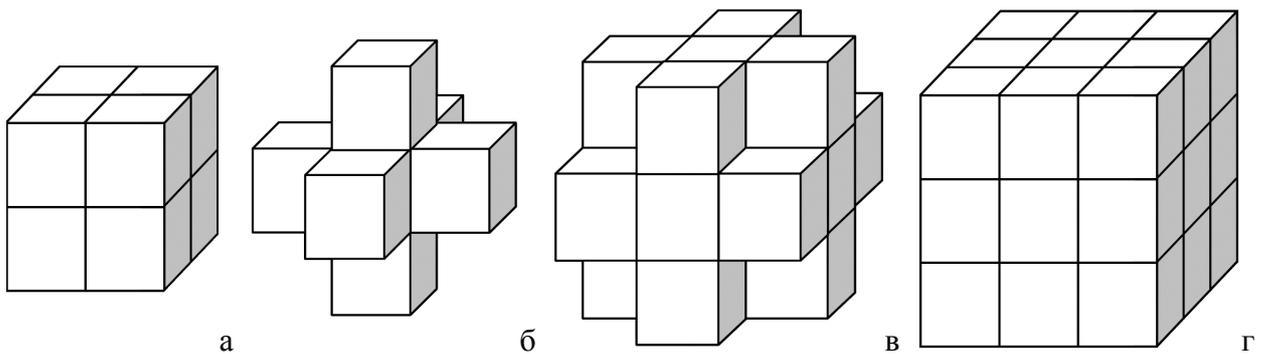


Рис. 7.1.1. Варианты окрестностей и соседства элементов (вокселей) в 3D: а) $2 \times 2 \times 2$ вокселей; б) 6-соседство вокселей; в) 18-соседство вокселей; г) 26-соседство вокселей.

По аналогии с 2D используется понятие *области принадлежности* как некоторого интересующего подмножества области анализа, включающего центральный элемент. Аналогично двумерному случаю, областью принадлежности, как правило, является пересечение области анализа и объекта изображения, к которому относится центральный элемент.

7.2. Модификация двухмасштабной многокомпонентной модели

По аналогии с моделью обычного 2D-изображения (раздел 1.2), 3D-изображение тоже можно описать двухмасштабной многокомпонентной моделью, представляющей статистические взаимосвязи элементов как на малых расстояниях в несколько шагов дискретизации, так и на больших расстояниях, соразмерных объектам изображения. Значения элементов 3D-изображения $\mathbf{X} = [x_{mnk}]$ также представляются суммой статистически независимых компонент:

$$x_{mnk} = S_{mnk} + t_{mnk} + \xi_{mnk}. \quad (7.2.1)$$

Первый член суммы — кусочно-гладкая компонента S_{mnk} , определяющая средние уровни яркости протяженных областей изображения; t_{mnk} — текстурно-детальная компонента, несущая информацию о текстуре и мелких деталях; ξ_{mnk} — шумовая компонента, определяемая шумами регистратора, аналого-цифрового преобразователя и др. Все компоненты предполагаются независимыми и аддитивными, а t_{mnk} и ξ_{mnk} — нормально распределенными и несмещенными.

Рассматриваемая модель 3D-изображения также является *локальной*, описывающей для каждого произвольно взятого элемента изображения x_{mnk} его взаимосвязи лишь с некоторым пространственно ограниченным множеством элементов $V_d(x_{mnk})$ или просто $V_d(x)$, которые его окружают и находятся от центрального элемента (выбранной метрики) на расстоянии, не превышающем d .

7.2.1. Модель окрестности (масштаб малого размера)

На масштабе малого размера (*масштабе элементов окрестности*), соответствующем масштабу окрестности двумерного изображения (раздел 1.2.2), рассматривается сравнительно небольшое связанное множество вокселей, расположенных на расстоянии нескольких шагов дискретизации. Как и в двумерной модели, воксели трехмерного изображения разделяются на два непересекающихся множества: попадающие на контурные участки (контурные элементы) и не попадающие (внутренние элементы), составляющие вместе полное изображение. Вводится понятие окрестности V_{mnk} вокселя x_{mnk} как группы из R вокселей $x_{mnk}^r \in V_{mnk}$, $r = 1, \dots, R$, ближайших к x_{mnk} , и попадающих в то же множество (контурное или внутренние), что и центральный воксель x_{mnk} .

Методом наименьших квадратов проводится гиперплоскость, наиболее близкая значениям вокселей из V_{mnk} , составляющая с гиперплоскостью, ориентированной вдоль осей координат **MNK**, некоторый угол, величина и направление которого в точке (m, n, k) характеризуется вектором \mathbf{g}_{mnk} . В точке r окрестности проведенная гиперплоскость отличается от значения x_{mnk}^r на случайную величину γ_{mnk}^r . Такое представление позволяет связать значения вокселей окрестности $x_{mnk}^r \in V_{mnk}$ формулой:

$$x_{mnk}^r = \mu_{mnk} + \rho^r \mathbf{g}_{mnk}^r + \gamma_{mnk}^r, \quad (7.2.2)$$

где μ_{mnk} — значение проведенной гиперплоскости в центральной точке окрестности (m, n, k) , ρ^r — расстояние между центральным вокселем x_{mnk} и x_{mnk}^r , \mathbf{g}_{mnk}^r — величина проекции \mathbf{g}_{mnk} на вектор из x_{mnk} в x_{mnk}^r , а γ_{mnk}^r — случайная величина.

Вводится понятие контурной маски $\mathbf{E} = [e_{mnk}]$, совпадающей по размерам с изображением: $e_{mnk} = 1$ для контурных и $e_{mnk} = 0$ для внутренних вокселей. Обозначая для контурных и

внутренних вокселей g_{mnk}^r через φ_{mnk}^r и ψ_{mnk}^r , а γ_{mnk}^r через ζ_{mnk}^r и η_{mnk}^r соответственно, представим g_{mnk}^r и γ_{mnk}^r в виде сумм $g_{mnk}^r = e_{mnk}^r \varphi_{mnk}^r + (1 - e_{mnk}^r) \psi_{mnk}^r$ и $\gamma_{mnk}^r = e_{mnk}^r \zeta_{mnk}^r + (1 - e_{mnk}^r) \eta_{mnk}^r$. В результате получим формулу модели трехмерной окрестности, описывающую статистические взаимосвязи входящих в нее вокселей:

$$x_{mnk}^r = \mu_{mnk} + e_{mnk}^r (\varphi_{mnk}^r \rho^r + \zeta_{mnk}^r) + (1 - e_{mnk}^r) (\psi_{mnk}^r \rho^r + \eta_{mnk}^r). \quad (7.2.3)$$

Здесь ζ_{mnk}^r — стохастическое возбуждение в точке r окрестности для контурных, а η_{mnk}^r — для внутренних вокселей. Случайные величины φ_{mnk} , ψ_{mnk} , ζ_{mnk} , и η_{mnk} считаются некоррелированными и несмещенными, а шумовые составляющие ζ_{mnk} , и η_{mnk} — нормально распределенными. Эксперименты, проводившиеся на двумерных изображениях [170,179], показывают, что значения дисперсий компонент φ и ψ различаются в 10-100 раз. Аналогичные цифры следует ожидать и для трехмерных изображений.

Как и для модели окрестности двумерного изображения, рассмотренной в разделе 1.2.2, вектор \mathbf{g}_{mnk} в дискретном пространстве отображает значение градиента непрерывной функции $F(x,y,z)$ в соответствующей точке (x,y,z) . Считая функцию $F(x,y,z)$ дважды дифференцируемой, ее полный дифференциал dF запишется в виде

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy + \frac{\partial F}{\partial z} dz,$$

где $\partial F/\partial j$ — частная производная по направлению j ($j = x; y; z$), а dj — дифференциалы соответствующих независимых переменных. Полным градиентом функции $F(x,y,z)$ будет

$$\nabla F(x,y,z) = (\partial F/\partial x)\mathbf{x} + (\partial F/\partial y)\mathbf{y} + (\partial F/\partial z)\mathbf{z}, \quad (7.2.4)$$

где \mathbf{x} , \mathbf{y} и \mathbf{z} — единичные векторы по направлениям x , y и z . Модуль градиента составит величину

$$\|\nabla F\| = \{(\partial F/\partial x)^2 + (\partial F/\partial y)^2 + (\partial F/\partial z)^2\}^{1/2}, \quad (7.2.5)$$

а угловое направление градиента будет задаваться двумя значениями:

$$\theta = \arctg[(\partial F/\partial y)/(\partial F/\partial x)], \quad \varphi = \arctg\{(\partial F/\partial z)/[(\partial F/\partial y)^2 + (\partial F/\partial x)^2]^{1/2}\}. \quad (7.2.6)$$

Лапласиан ($\Delta = \nabla \cdot \nabla$) определяется как сумма частных вторых производных функции $F(x,y,z)$ по направлениям:

$$\Delta F = (\partial^2 F/\partial x^2) + (\partial^2 F/\partial y^2) + (\partial^2 F/\partial z^2). \quad (7.2.7)$$

7.2.2. Модель фрагмента (масштаб большого размера)

На масштабе большого размера (*масштабе объектов фрагмента*), аналогично модели фрагмента (раздел 1.2.3), предполагается, что гладкие составляющие S^v тех частей областей u^v ($v = 1, \dots, V$), которые попадают во фрагмент W_{mnk} , могут быть представлены полиномом степени

не выше, чем ω . Тогда внутри фрагмента W_{mnk} составляющая S_{ijl} может быть описана следующей формулой:

$$S_{ijl}^v(W_{mnk}) = \sum_{v=1}^V \delta_{u^v} \sum_{p=0}^{\omega} \sum_{q=0}^{\omega-p} \sum_{r=0}^{\omega-p-q} a_{pqr}^v i^p j^q l^r; \quad (7.2.8)$$

здесь (i, j, l) — точка области u^v во фрагменте W_{mnk} ; $\delta_{u^v} = 1$, если точка $(i, j, l) \in u^v$, и $\delta_{u^v} = 0$ в остальных случаях. Добавляя в (7.2.8) текстурную t_{ijl} и шумовую ξ_{ijl} составляющие, получим выражение для значений вокселей внутри фрагмента:

$$x_{ijl}^v = \sum_{v=1}^V \delta_{u^v} \left(\sum_{p=0}^{\omega} \sum_{q=0}^{\omega-p} \sum_{r=0}^{\omega-p-q} a_{pqr}^v i^p j^q l^r + t_{ijl}^v + \xi_{ijl} \right). \quad (7.2.9)$$

Это общая формула модели, описывающей значения вокселей областей внутри фрагмента. Значения составляющих t_{ijl}^v и ξ_{ijl} считаются нормально распределенными, но значения дисперсий t^v вообще говоря различаются от области к области.

В силу свойств объектов отображаемого трехмерного пространства, многие трехмерные изображения в пределах типичного фрагмента анализа имеют приблизительно постоянные средние локальные яркости, не меняющиеся заметно. По этой причине во многих случаях допустимо выбрать минимальную степень полинома: $\omega = 0$. Тогда $S_{ijl}^v(W_{mnk}) = S_{mnk}^v$ и (7.2.9) преобразуется к следующему виду:

$$x_{ijl}^v = \sum_{u=1}^V \delta_{u^v} (S_{mnk}^v + t_{ijl}^v + \xi_{ijl}). \quad (7.2.10)$$

Это формула кусочно-постоянной модели фрагмента для представления участков областей изображения, попадающих во фрагмент W_{mnk} .

На масштабе большого размера обычно предполагается, что фрагмент W_{mnk} имеет форму прямоугольного параллелепипеда. Такой выбор обусловлен особенностями алгоритмической реализации методов обработки трехмерных изображений.

7.3. Модификация методов частотной фильтрации при переходе в 3D

Многие методы фильтрации при переходе от одномерных к двумерным сигналам модифицируются сравнительно просто. При переходе к трехмерным сигналам их изменение во многих случаях тоже несложно. Покажем это на примере наиболее распространенных алгоритмов, основанных на частотной и пространственной фильтрациях.

Одним из эффективных подходов в обработке сигналов является применение методов фильтрации, использующих ортогональные преобразования: Фурье, Уолша-Адамара, Хаара, косинусное, и пр. [8]. Наиболее распространен класс преобразований, обеспечивающих

разложение сигналов по гармоническим функциям, из которых важнейшим является преобразование Фурье; на его примере и покажем модификацию преобразования.

При переходе от 2D- к 3D-изображениям трехмерными становятся как пространственная, так и частотная области представления данных, тем самым вместо двумерного получаем трехмерное Фурье-преобразование; все основные свойства преобразования Фурье при этом сохраняются.

Пусть $f(x,y,z)$ — непрерывная функция трех переменных x , y и z . Пара трехмерных непрерывных преобразований Фурье (прямое и обратное) задается следующими выражениями:

$$F(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) e^{-i2\pi(xu+yv+zw)} dx dy dz \quad (7.3.1)$$

$$\text{и} \quad f(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v, w) e^{i2\pi(xu+yv+zw)} du dv dw, \quad (7.3.2)$$

где u , v и w — непрерывные частотные, а x , y и z — непрерывные пространственные переменные.

Трехмерное дискретное преобразование Фурье (ДПФ) может быть записано в виде:

$$F(u, v, w) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} f(m, n, k) e^{-i2\pi(um/M+vn/N+wk/K)}, \quad (7.3.3)$$

где $f(m,n,k)$ — трехмерный цифровой массив размерами $M \times N \times K$. Обратное трехмерное дискретное преобразование Фурье будет иметь вид:

$$f(m, n, k) = \frac{1}{MNK} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \sum_{w=0}^{K-1} F(u, v, w) e^{i2\pi(um/M+vn/N+wk/K)} \quad (7.3.4)$$

для $0 \leq m < M$, $0 \leq n < N$ и $0 \leq k < K$, а также $0 \leq u < M$, $0 \leq v < N$ и $0 \leq w < K$.

Как и в двумерном случае, частотная фильтрация трехмерного сигнала осуществляется выполнением прямого преобразования Фурье (7.3.3), модификацией полученного спектра $F(u,v,w)$, например умножением его на амплитуду спектра (трехмерного) того или иного фильтра, и последующим обратным преобразованием (7.3.4). По аналогии с преобразованием Фурье, другие ортогональные преобразования при переходе к трехмерному сигналу также несложно видоизменяются добавлением размерности.

7.4. Модификация методов пространственной обработки и анализа при переходе в 3D

Пространственная фильтрация обычно осуществляется локальными операторами согласно формуле (7.1.1) с определенными ограничениями на размеры локальной области анализа $V_d(x)$.

Это обусловлено тем, что в случае функции $f(\cdot)$ в (7.1.1) общего вида, когда не существует рекуррентных соотношений для определения последующих значений на основе ранее полученных, количество необходимых вычислений возрастает как минимум пропорционально объему $V_d(x)$, т.е. числу элементов, в него попадающих. Окрестности, которые в двумерном случае содержат порядка $(2d)^2$ пикселей, в 3D будут содержать уже порядка $(2d)^3$ вокселей. Тем не менее для некоторых случаев, например, когда в основе фильтрации лежит вычисление арифметического среднего или порядковых статистик по $V_d(x)$, в трехмерном случае, как и в двумерном, удастся построить быстрые алгоритмы.

7.4.1. Методы, использующие оценку среднего по фрагменту

Ряд алгоритмов пространственной обработки изображений, связанных со сглаживанием, выделением низко- или высокочастотных составляющих, улучшением изображения, из числа рассматривавшихся в разделе 3.3.1, обобщаются на трехмерные изображения следующей формулой, аналогичной формуле (3.3.3):

$$y_{mn} = f(x_{mn} - S_{mn}, D_{mn}) + bS_{mn} + c. \quad (7.4.1)$$

Здесь x_{mn} — значение центрального (обрабатываемого) элемента, S_{mn} — оценка сглаженного значения по фрагменту в точке (m, n) (например, значение арифметического среднего или медианы по фрагменту), D_{mn} — оценка ширины функции распределения значений по фрагменту, $f(u, v)$ — зависимость усиления контраста, b и c — параметры преобразования.

Очевидно, преобразования вида (7.4.1) легко переносятся на 3D-изображения добавлением размерности, но с учетом того, что S_{mnk} и D_{mnk} должны определяться уже не по двумерному, а по трехмерному фрагменту (прямоугольному параллелепипеду) с центром в точке (m, n, k) :

$$y_{mnk} = f(x_{mnk} - S_{mnk}, D_{mnk}) + bS_{mnk} + c. \quad (7.4.2)$$

Как и в двумерном случае, при $f(u, v) = 0$ и $c = 0$ получаем сглаживание (выделение низкочастотной составляющей), при $f(u, v) = u$, $b = 0$ и $c = 0,5$ — выделение высокочастотной составляющей, при $f(u, v) > u$, $b \lesssim 1$ и $c \approx (1-b)/2$ — улучшение изображения.

Модификации алгоритмов быстрого вычисления арифметического среднего или медианы S_{mnk} , а также дисперсии D_{mnk} по прямоугольному параллелепипеду рассмотрены в конце главы в разделе 7.5.

7.4.2. Операторы контурных перепадов

В основе большинства алгоритмов обнаружения контуров на изображении так или иначе лежат операторы контурных перепадов, в которых, как правило, используются оценки значений либо первой, либо второй производной [40, 113]. В случае 2D-изображения контуры объектов

суть линии, разделяющие плоские области, и в идеальном случае (когда контурные линии не имеют разрывов) получаем обычную плоскую карту изображаемой сцены.

В 3D-пространстве идеальное контурное изображение представляет собой множество разделяющих объекты поверхностей произвольной формы, и здесь возможно возникновение самых различных конфигураций контуров; например, становятся допустимыми узлы. Сечение 3D-контурного изображения плоскостью представляет собой двумерную карту контуров. Однако такая карта вообще говоря не обязана совпадать с картой, получаемой проведением контуров по 2D-изображению, являющемуся сечением 3D-изображения той же плоскостью. Следует также отметить, что, как показано в разделе 4.3, обнаружение контуров не по самому изображению, а по выделенной путем декомпозиции сглаженной компоненте S_{mnk} (о которой речь ниже) позволяет достичь более высокой слитности контурных линий, меньшей их толщины, а также существенного снижения количества ошибок при обнаружении контуров.

Рассмотрим модификации базовых операторов контурных перепадов при переходе от 2D-к 3D-изображениям на основе первой производной — модуля градиента (7.2.5), второй производной — Лапласиана (7.2.7), а также некоторые другие, которые могут быть построены в рамках рассмотренных выше трехмерных окрестностей.

На основе первой производной: оператор Робертса

Оператор Робертса для двумерного изображения описывается следующей формулой [121]:

$$y(m,n) = \{|x(m,n) - x(m+1,n+1)| + |x(m+1,n) - x(m,n+1)|\}/2,$$

т.е. как сумма модулей разностей значений диагональных элементов по квадрату 2×2 пикселей.

Для 3D-изображения оператор Робертса будет представлять собой сумму модулей разностей значений элементов в диагоналях куба из $2 \times 2 \times 2$ вокселей (Рис. 7.1.1,а):

$$y(m,n,k) = \{|x(m,n,k) - x(m+1,n+1,k+1)| + |x(m+1,n,k) - x(m,n+1,k+1)| + \\ + |x(m,n+1,k) - x(m+1,n,k+1)| + |x(m,n,k+1) - x(m+1,n+1,k)|\}/4. \quad (7.4.3)$$

На основе первой производной: оператор Собела

Оператор Собела для двумерного изображения задается выражением [40]:

$$y(m,n) = \{|x(m-1,n-1) + 2x(m-1,n) + x(m-1,n+1) - \\ - x(m+1,n-1) - 2x(m+1,n) - x(m+1,n+1)| + \\ + |x(m-1,n-1) + 2x(m,n-1) + x(m+1,n-1) - \\ - x(m+1,n-1) - 2x(m+1,n) - x(m+1,n+1)|\}/8. \quad (7.4.4)$$

Оператор Собела (и аналогичные ему) для 3D-изображения строятся в окрестности $3 \times 3 \times 3$ вокселей (Рис. 7.1.1,г). Для построения оператора Собела необходимо определить частные отклики как модули разностей значений элементов для каждого из направлений m , n и k , что соответствует выражению под одним из знаков модуля в (7.4.4). Для направления m такую

зависимость $y_m(m,n,k)$ с точностью до выбора коэффициентов a , b и c можно выразить следующей формулой:

$$\begin{aligned}
 y_m(m,n,k) = & |ax(m-1,n-1,k-1) + bx(m-1,n-1,k) + ax(m-1,n-1,k+1) + \\
 & + bx(m-1,n,k-1) + cx(m-1,n,k) + bx(m-1,n,k+1) + \\
 & + ax(m-1,n+1,k-1) + bx(m-1,n+1,k) + ax(m-1,n+1,k+1) - \\
 & - ax(m+1,n-1,k-1) - bx(m+1,n-1,k) - ax(m+1,n-1,k+1) - \\
 & - bx(m+1,n,k-1) - cx(m+1,n,k) - bx(m+1,n,k+1) - \\
 & - ax(m+1,n+1,k-1) - bx(m+1,n+1,k) - ax(m+1,n+1,k+1)| / (4a + 4b + c).
 \end{aligned} \tag{7.4.5}$$

Аналогичным образом записываются модули разностей $y_n(m,n,k)$ и $y_k(m,n,k)$ для составляющих n и k . Трехмерный оператор, аналогичный оператору Собела, окончательно выразим через сумму откликов по трем направлениям:

$$y(m,n,k) = y_m(m,n,k) + y_n(m,n,k) + y_k(m,n,k). \tag{7.4.6}$$

Коэффициенты a , b и c рекомендуется выбирать вблизи значений 1, 2 и 3 соответственно.

На основе второй производной: оператор Лапласа

Оператор Лапласа (лапласиан) для двумерного изображения [40] задается формулой

$$\begin{aligned}
 y(m,n) = & x(m,n) - \{x(m-1,n-1) + x(m-1,n) + x(m-1,n+1) + x(m,n-1) + \\
 & + x(m,n+1) + x(m+1,n-1) + x(m+1,n) + x(m+1,n+1)\} / 8.
 \end{aligned} \tag{7.4.7}$$

Модификация оператора Лапласа для 3D-изображения может быть выражена через набор элементов окрестности следующим образом:

$$y_{mnk} = x_{mnk} - (1/Q) \sum_{x_{ijl} \in V_{mnk}} x_{ijl}. \tag{7.4.8}$$

Здесь V_{mnk} — окрестность точки x_{mnk} , не включающая саму центральную точку x_{mnk} , Q_V — число точек в такой окрестности, $\{x_{ijl}\}$ — набор элементов окрестности V_{mnk} . В качестве V_{mnk} может выступать одна из окрестностей, показанных на Рис. 7.1.16–г: для окрестности (б) $Q_V = 6$, для окрестности (в) $Q_V = 18$, для окрестности (г) $Q_V = 26$.

Другие контурные операторы

Как и для двумерного случая, возможно построение и других детекторов, позволяющих обнаруживать контурные перепады [317], например, разности максимума и минимума значений элементов окрестности V_{mnk} : $y(m,n,k) = \max\{x_{ijl} \mid x_{ijl} \in V_{mnk}\} - \min\{x_{ijl} \mid x_{ijl} \in V_{mnk}\}$.

7.4.3. Фильтрация импульсных помех

Импульсными помехами называют искажения отдельных случайных элементов изображения, когда значения помехи не коррелированы с истинными значениями элементов сигнала. Вопросы фильтрации импульсных помех на двумерных изображениях подробно были рассмотрены в разделе 3.1.1. Для трехмерного изображения можно использовать расширение

двумерной модели искажения и подходов к фильтрации, которое в основном сводится к добавлению третьей пространственной размерности.

Модель искажения изображения импульсными помехами проста. Значение каждого из элементов x_{mnk} с вероятностью p заменяется на случайное значение ξ_{mnk} независимо от значений x_{mnk} или остальных элементов. Обозначим через $\mathbf{X}^o = [x_{mnk}^o]$ исходное неискаженное изображение, через $\mathbf{X} = [x_{mnk}]$ — искаженное импульсной помехой, а через $\mathbf{Y} = [y_{mnk}]$ — результат фильтрации. Как и в случае 2D-изображения, процесс искажения представляется в виде:

$$x_{mnk} = \begin{cases} x_{mnk}^o & \text{с вероятностью } (1-p); \\ \xi_{mnk} & \text{с вероятностью } p. \end{cases}$$

Как правило полагается, что значения импульсных помех ξ_{mnk} распределены равномерно в диапазоне яркостей $[0, x_{\max}]$. Задача фильтрации импульсных помех сводится к обнаружению помех и коррекции искаженных отсчетов яркости. В наиболее распространенных алгоритмах фильтрации как обнаружение, так и исправление значения элемента основываются на предсказании значения элемента \tilde{x}_{mnk} , получаемого на основе анализа окружающей его окрестности V_{mnk} . При этом используются локальные корреляционные связи близлежащих элементов изображения и предполагается, что шум пространственно декоррелирован. При обнаружении сравниваются наблюдаемое x_{mnk} и предсказываемое \tilde{x}_{mnk} значения. Если они отличаются более, чем на величину некоторого порога обнаружения δ , считается, что x_{mnk} — помеха и осуществляется ее исправление на значение \tilde{x}_{mnk} .

Саму процедуру фильтрации 3D-изображения можно обобщенно сформулировать следующим образом. На этапе детектирования помехи сравниваются наблюдаемое x_{mnk} и предсказываемое \tilde{x}_{mnk} значения. Если они отличаются более, чем на величину некоторого заданного порога обнаружения δ , считается, что x_{mnk} — помеха и осуществляется ее исправление на значение \tilde{x}_{mnk} . Общая формула вычисления \tilde{x}_{mnk} для большинства методов при этом практически совпадает с формулой (7.1.1):

$$\tilde{x}_{mnk} = f\{x_{ijl} \mid x_{ijl} \in V_{mnk}\}. \quad (7.4.9)$$

Значение функции $f\{\cdot\}$ в (7.4.9) будем находить следующим образом. Выберем из набора элементов окрестности $x_{ijl} \in V_{mnk}$ усеченное множество $\{x'_{ijl}\}$ путем выстраивания значений $\{x_{ijl}\}$ в вариационный ряд и отбрасывания n точек с наименьшими и n точек наибольшими значениями ($n \approx N_V/8$), где N_V — число элементов в окрестности V_{mnk} . Множество точек $\{x'_{ijl}\}$ будем рассматривать в качестве области принадлежности центрального элемента x_{mnk} . Проведем через выбранные точки ближайшую гиперплоскость и возьмем ее значение в точке (m, n, k) . Полученная величина и будет искомым предсказываемым значением \tilde{x}_{mnk} .

Обобщенная формула фильтрации при этом выглядит следующим образом:

$$y_{mnk} = \begin{cases} x_{mnk}, & \text{если } |x_{mnk} - \tilde{x}_{mnk}| < \delta; \\ \tilde{x}_{mnk}, & \text{если } |x_{mnk} - \tilde{x}_{mnk}| \geq \delta. \end{cases} \quad (7.4.10)$$

При больших p , когда вероятность искажения пары соседних вокселей достаточно высока, рекомендуется итеративный процесс фильтрации: первоначально с большим значением порога δ , затем с постепенным его уменьшением. Для первой итерации можно рекомендовать применение окрестности на Рис. 7.1.1,б для последующих — окрестностей на Рис. 7.1.1,в и г. При использовании последней окрестности достигается наибольшее сглаживание.

7.4.4. Декомпозиция изображения

Согласно формулировке, сделанной во введении Главы 2, в терминах модели (7.2.1) декомпозиция означает разделение изображения на сглаженную S и текстурно-детальную ($t + \xi$) компоненты. Вопросы декомпозиции двумерного изображения были подробно изложены в Главе 2 и опубликованы работах [176,179,192,236]. Суть изложенного в них алгоритма заключается в том, что для каждой точки изображения (m,n) на основе последовательного анализа окрестности V_{mn} и фрагмента W_{mn} , окружающих элемент x_{mn} , выбирается часть точек W_{mn} в качестве *области принадлежности*, и значение S_{mn} оценивается по выбранному множеству. Согласно предложенному алгоритму, анализируются лишь распределения значений элементов по окрестности V_{mn} и фрагменту W_{mn} .

Такой подход, использующий распределение значений элементов по области анализа, позволяет легко перейти от 2D- к 3D-изображению — достаточно вместо двумерных окрестности V_{mn} и фрагмента W_{mn} подставить в формулы трехмерные окрестность V_{mnk} и фрагмент W_{mnk} . Тогда, по аналогии с алгоритмом декомпозиции «D» двумерного изображения, рассмотренным в разделе 2.2.1, алгоритм декомпозиции для трехмерного изображения можно сформулировать следующим образом.

При заданных размерах $l \times l \times l$ окрестности V_{mnk} и $L \times L \times L$ фрагмента W_{mnk} ($l < L$), центрированных в точке (m,n,k) , ширине яркостных интервалов анализа Δ^V и Δ^W , а также ранговых параметрах $n^V < l^3/2$ и $n^W < L^3/2$, значение сглаженной компоненты S_{mnk} в (7.2.1) находится при помощи следующих операций, выполняемых для каждой точки (m,n,k) изображения:

1. Подсчитываются нормированные гистограммы (вероятности распределения) по трехмерным окрестности $H_{mnk}^V = \{h_{mnk}^V(i)\}$ и фрагменту $H_{mnk}^W = \{h_{mnk}^W(i)\}$ с центром в точке (m,n,k) .
2. По гистограмме окрестности H_{mnk}^V и значению n^V находятся ранговые границы $R_1^V = R^V(n^V/l^3)$ и $R_2^V = R^V(1 - n^V/l^3)$; здесь $R(x)$ определяется как решение уравнения $\sum_{i=0}^{R(x)} h_{mnk}^V(i) = x$, где $h_{mnk}^V(i)$

— гистограмма значений элементов в окрестности V_{mnk} . Сравнением значения центрального элемента x_{mnk} с R_1^V и R_2^V определяется промежуточное усеченное значение \tilde{x}^V :

$$\tilde{x}^V = \begin{cases} x_{mnk}, & \text{если } R_1^V \leq x_{mnk} \leq R_2^V; \\ R_1^V, & \text{если } x_{mnk} < R_1^V; \\ R_2^V, & \text{если } x_{mnk} > R_2^V. \end{cases}$$

3. С помощью сигма-фильтра [297], являющегося частным случаем парзеновского окна [327], находится урезанное среднее \bar{x}_{mnk} . Для этого из элементов окрестности V_{mnk} выбираются z значений $x_{mnk}^r \in V_{mnk}$ ($r=1, \dots, z$), попадающих в интервал $(\tilde{x}^V - \Delta^V, \tilde{x}^V + \Delta^V)$, где Δ^V — полуширина интервала. По значениям x_{mnk}^r из данного интервала подсчитывается среднее:

$$\bar{x}_{mnk} = A(V_{mnk}, x_{mnk}, n^V, \Delta^V) = \frac{1}{z} \sum_{r=1}^z x_{mnk}^r, \quad \tilde{x}^V - \Delta^V \leq x_{mnk}^r \leq \tilde{x}^V + \Delta^V. \quad (7.4.11)$$

4. Аналогично п. 2, по гистограмме фрагмента H_{mnk}^W и заданному n^W находятся ранговые границы $R_1^W = R^W(n^W/L^3)$ и $R_2^W = R^W(1 - n^W/L^3)$ и определяется усеченное значение \tilde{x}^W :

$$\tilde{x}^W = \begin{cases} \bar{x}_{mnk}, & \text{если } R_1^W \leq \bar{x}_{mnk} \leq R_2^W; \\ R_1^W, & \text{если } \bar{x}_{mnk} < R_1^W; \\ R_2^W, & \text{если } \bar{x}_{mnk} > R_2^W. \end{cases}$$

5. Сглаженное значение S_{mnk} находится по гистограмме фрагмента H_{mnk}^W как арифметическое среднее или медиана значений H_{mnk}^W , попадающих в интервал $(\tilde{x}^W - \Delta^W, \tilde{x}^W + \Delta^W)$:

$$S_{mnk} = A(W_{mnk}, \bar{x}_{mnk}, n^W, \Delta^W). \quad (7.4.12)$$

Полученное значение S_{mnk} считается искомой сглаженной компонентой.

7.4.5. Обнаружение объектов заданного объема

В разделе 4.1 было показано, как алгоритм декомпозиции может быть использован для обнаружения объектов на изображении. Аналогично 2D-изображениям, для которых решается задача обнаружения объектов по их площади, в трехмерной модификации задача формулируется как обнаружения объектов по их объему. По аналогии с двумерной, трехмерная задача также допускает формулировку в трех вариантах: обнаружение объектов с объемом (т.е. числом элементов) N^j меньше заданного T_1 , больше заданного T_2 , и обнаружение объектов, имеющих объем в интервале $T_1 < N^j < T_2$.

1. Обнаружение объектов с $N^j > T$

Предполагается, что изображение состоит из достаточно ровного фона (большая область U^0), внутри которой имеется набор небольших областей U^1, \dots, U^j , отстоящих друг от друга достаточно далеко, и можно выбрать некоторый размер фрагмента L ($L^3/2 > T$) такой, что в любой фрагмент W_{mnk} попадает не более одной области с $N^j > T$, либо несколько меньших, но

при условии $\Sigma N^j < T$ ($U^j \subset W$). В п. 5 алгоритма декомпозиции (7.4.12) выберем $n^W = T$, а $R_1^W = R(T/L^3)$ и $R_2^W = R(1-T/L^3)$. Обработкой с выбранными значениями R_1^W и R_2^W получим:

$$y_{mnk} = S_{mnk}, \quad (7.4.13)$$

т.е. сглаженную компоненту исходного изображения, на которой остались только области с $N^j > T$, обнаруживаемые детектором со значением порога $S(U^0) \pm \delta$, где $S(U^0)$ — средняя яркость фона, а $\delta < \min_j \{|S(U^j) - S(U^0)|\}$; ($S(U^j)$ — яркости соответствующих областей).

2. Обнаружение объектов с $N^j < T$

Сглаженная компонента S_{mnk} , в (7.4.12) содержит лишь области с $N^j > T$, а области с $N^j < T$ содержатся в компоненте \bar{x}_{mnk} , получаемой по формуле (7.4.11). Объекты с $N^j < T$ обнаруживаются в тех точках (m, n, k) , где $|\bar{x}_{mnk} - S_{mnk}| \geq \delta$ (δ — порог обнаружения).

3. Обнаружение объектов с $T_1 < N^j < T_2$. Как и в 2D, возможны два варианта решения.

В первом случае сначала выберем $n^W = T_1$. Тогда сглаженная компонента S_{mnk} в (7.4.12) будет содержать объекты с $N^j > T_1$. Осуществим повторную ее обработку алгоритмом (7.4.12) с $n^W = T_2$ ($T_2 > T_1$). Очевидно, что во вновь полученной сглаженной компоненте S'_{mnk} будут содержаться лишь объекты с $N^j > T_2$. Взяв разность $y_{mnk} = |S_{mnk} - S'_{mnk}|$ получим сигнал, содержащий объекты в диапазоне $T_1 < N^j < T_2$. Недостаток данного решения — алгоритм получается двухпроходным.

Второй вариант. Обратим внимание, что при анализе гистограмм по окрестности и фрагменту используются два разных порога (n^V и n^W). Выберем размеры окрестности l и фрагмента L больше обычного — такими, чтобы $l^3 > 2T_1$ и $L^3 > 2T_2$. Задав R_1^V и R_2^V как $R_1^V = R^V(T_1/l^2)$ и $R_2^V = R^V(1 - T_1/l^2)$, после операции (7.4.11) будем иметь \bar{x}_{mnk} , которое уже не содержит области с $N^j < T_1$. Далее в п. 5 алгоритма декомпозиции, при анализе H_{mn}^W , зададим R_1^W и R_2^W как $R_1^W = R^W(T_2/L^3)$ и $R_2^W = R^W(1 - T_2/L^3)$. Получив значение S_{mnk} в (7.4.12), возьмем разность $y_{mnk} = |\bar{x}_{mnk} - S_{mnk}|$, на которой объекты выделяются пороговым детектором. Отметим, что понятие объема в трехмерном варианте алгоритма, также как и как понятие площади в двумерном случае, используются в несколько необычном смысле — как «локальный» объем, т.е. объем той части объекта, которая попадает внутрь фрагмента W_{mnk} .

7.5. Модификация некоторых вычислительные алгоритмов

7.5.1. Вычисление суммы по прямоугольному параллелепипеду

Введем следующие обозначения. Сумму значений элементов по прямоугольному фрагменту двумерного изображения обозначим через

$$S_{(ij)(mn)} = \sum_{u=i}^{m-1} \sum_{v=j}^{n-1} x_{uv} .$$

Таким образом $S_{(ij)(mn)}$ — сумма значений элементов x_{uv} , попадающих в прямоугольный фрагмент, диагональные точки которого имеют координаты (i, j) и $(m-1, n-1)$. Обратим внимание, что фрагмент не включает точку с координатами (m, n) и соответствующие ей строку и столбец. Аналогично для трехмерного изображения, $S_{(ijl)(mnk)}$ — сумма значений элементов x_{ijl} в прямоугольном параллелепипеде с угловыми координатами (i, j, l) и $(m-1, n-1, k-1)$:

$$S_{(ijl)(mnk)} = \sum_{u=i}^{m-1} \sum_{v=j}^{n-1} \sum_{w=l}^{k-1} x_{uvw} .$$

Для двумерного изображения классический способ вычисления суммы $S_{(mn)(m+H, n+L)}$ по скользящему прямоугольному фрагменту размерами $H \times L$ элементов при переходе от элемента (m, n) к элементу $(m, n+1)$ сводится к формуле

$$S_{(m, n+1)(m+H, n+L+1)} = S_{(m, n)(m+H, n+L)} - S_{(m, n)(m+H, n+1)} + S_{(m, n+L)(m+H, n+L+1)}, \quad (7.5.1)$$

где последние два члена — суммы элементов по левому (удаляемому) и правому (добавляемому) столбцам фрагмента. Такой алгоритм требует в среднем 4 арифметических операции на один элемент изображения независимо от размеров фрагмента — 2 операции в выражении (7.5.1) и две операции на пересчет каждой из сумм по столбцу $S_{(m, n)(m+H, n+1)}$ при переходе от строки m к строке $m+1$. Дополнительно требуется N ячеек памяти для хранения массива сумм по столбцам.

При переходе к трехмерному изображению, формула (7.5.1) будет модифицирована для скользящего прямоугольного параллелепипеда размерами $H \times L \times J$ следующим образом:

$$S_{(m, n, k+1)(m+H, n+L, k+J+1)} = S_{(m, n, k)(m+H, n+L, k+J)} - S_{(m, n, k)(m+H, n+L, k+1)} + S_{(m, n, k+J)(m+H, n+L, k+J+1)}, \quad (7.5.2)$$

где последние два члена — суммы элементов по левой (удаляемой) и правой (добавляемой) граням параллелепипеда. Данный алгоритм требует 6 арифметических операции независимо от размеров фрагмента — 2 операции в выражении (7.5.2), две операции на пересчет каждой из сумм по граням и две на пересчет сумм по столбцам. Кроме того, требуется K ячеек памяти для хранения массива сумм по граням и $N \times K$ ячеек памяти для хранения сумм по столбцам.

Для двумерного изображения известен также другой алгоритм вычисления суммы по прямоугольнику произвольного размера. Пусть для каждой точки (m, n) подсчитаны суммы $S_{mn} = S_{(0,0)(m,n)}$ по прямоугольнику с диагональными элементами x_{00} и $x_{m-1, n-1}$. Тогда сумма $S_{(ij)(mn)}$ значений элементов внутри прямоугольника с координатами углов (i, j) и $(m-1, n-1)$ вычисляется следующим образом:

$$S_{(ij)(mn)} = S_{mn} - S_{m, j} - S_{i, n} + S_{ij}, \quad (7.5.3)$$

что в среднем для каждого элемента изображения требует 2 операции для вычисления суммы S_{mn} и 3 операции для вычислений по формуле (7.5.3). Однако для хранения сумм S_{mn} требуется

уже $M \times N$ ячеек памяти, т.е. столько же, каков размер изображения. Преимуществом является то, что имея массив S_{mn} , за те же 3 операции можно вычислять суммы $S_{(ij)(mn)}$ для любых значений угловых координат, а не только по скользящему фрагменту.

Алгоритм (7.5.3) также может быть распространен на трехмерное изображение. Пусть для каждой точки (m, n, k) подсчитаны суммы $S_{mnk} = S_{(000)(mnk)}$ по прямоугольному параллелепипеду с диагональными элементами x_{000} и $x_{m-1, n-1, k-1}$. Нетрудно показать, что в таком случае сумма $S_{(ijl)(mnk)}$ значений элементов внутри прямоугольного параллелепипеда с угловыми координатами (i, j, l) и $(m-1, n-1, k-1)$ вычисляется при помощи следующей операции:

$$S_{(ijl)(mnk)} = S_{mnk} - S_{mjk} - S_{mnl} - S_{ink} + S_{mjl} + S_{ijk} + S_{inl} - S_{ijl}. \quad (7.5.4)$$

Таким образом с учетом того, что для вычисления каждого из значений S_{mnk} требуется 3 арифметических операции, плюс для вычисления α сумм $S_{(ijl)(mnk)}$ по (7.5.4) требуется 7α операций для каждого элемента трехмерного изображения. Объем дополнительной требуемой памяти составит $M \times N \times K$ ячеек с разрядностью, достаточной для хранения значений сумм S_{mnk} . Аналогично двумерному варианту (7.5.3), имея массив S_{mnk} , вычисление суммы $S_{(ijl)(mnk)}$ для любых значений угловых координат требует дополнительно 7 операций.

Аналогичным образом можно находить дисперсии по фрагменту $D_{(ijl)(mnk)}$, вычисляя значения сумм квадратов $S_{(mnk)}(x^2)$ для каждой из точек (m, n, k) и для прямоугольного параллелепипеда $S_{(ijl)(mnk)}(x^2)$, после чего находим значение дисперсии D пользуясь формулой

$$D_{(ijl)(mnk)} = \{S_{(ijl)(mnk)}(x^2) - (S_{(ijl)(mnk)})^2\} / N_{(ijl)(mnk)}, \quad (7.5.5)$$

где $S_{(ijl)(mnk)}(x^2)$ — сумма квадратов значений элементов, попадающих в параллелепипед, а $N_{(ijl)(mnk)} = (m-i) \times (n-j) \times (k-l)$ — общее число точек в параллелепипеде.

7.5.2. Вычисление порядковых статистик по прямоугольному параллелепипеду

Основой для вычисления порядковых статистик по фрагменту W как двумерного, так и трехмерного изображения служит гистограмма распределения значений яркости по этому фрагменту $h^W(x)$, а также ее интегральная характеристика $F^W(x)$:

$$F^W(x) = \sum_{i=0}^x h^W(i); \quad F^W(x_{\max}) = N^W, \quad (7.5.6)$$

где x_{\max} — максимально возможное значение яркости, а N^W — число точек во фрагменте W . Порядковые статистики вида $R^W(n)$, где $0 \leq n \leq N^W$, представляют собой зависимость:

$$R^W(n) = z, \quad \text{если } F^W(z-1) < n \leq F^W(z). \quad (7.5.7)$$

Алгоритм скользящего вычисления гистограммы по фрагменту строится аналогично формулам (7.5.1) и (7.5.2), т.е. при смещении фрагмента к следующей точке производится удаление точек на одной грани фрагмента и добавление точек на противоположной грани [183].

В следующем разделе будет рассмотрен алгоритм, позволяющий вычислять гистограмму и порядковые статистики по скользящему фрагменту за конечное число тактов [55,183], однако он применим лишь для специализированных многопроцессорных вычислителей.

В случае однопроцессорного вычислителя и двумерного изображения, алгоритм скользящего вычисления гистограммы по фрагменту при переходе от точки (m,n) к соседней точке $(m,n+1)$ требует в среднем $2H$ числа операций (H — число строк во фрагменте). В трехмерном же случае при переходе от точки (m,n,k) к точке $(m,n,k+1)$ потребуется уже $2H \times L$ операций, где $H \times L$ — число точек в грани параллелепипеда, перпендикулярной направлению смещения K .

Как видно, в трехмерном случае число требуемых операций растет пропорционально произведению $H \times L$. Однако, если $(H \times L) > (x_{\max} + 1)$, то вместо операций со значениями отдельных точек становится выгодным предварительно сформировать гистограммы $h_{(ijk)(m,n,k+1)}^F(x)$ для соответствующих граней параллелепипеда, а затем осуществлять операции вычитания и прибавления таких гистограмм:

$$h_{(i,j,k+1)(m,n,k+Z+1)}^W(x) = h_{(ijk)(m,n,k+Z)}^W(x) - h_{(ijk)(m,n,k+1)}^F(x) + h_{(i,j,k+Z)(m,n,k+Z+1)}^F(x), \quad (7.5.8)$$

где Z — размер фрагмента в направлении смещения K . Выполнение действий по формуле (7.5.8) требует в среднем $2(x_{\max} + 1)$ арифметических операций на одну точку изображения независимо от размера параллелепипеда. Для пересчета гистограмм по граням параллелепипеда $h_{(ijk)(m,n,k+1)}^F(x)$ требуется дополнительно в среднем $2L$ операций на точку и $(x_{\max} + 1) \times K$ ячеек памяти для хранения K гистограмм.

Число операций $(x_{\max} + 1)$, требуемое для прибавления/вычитания каждой из гистограмм по граням параллелепипеда $h_{(ijk)(m,n,k+1)}^F(x)$, можно значительно уменьшить, если воспользоваться одним из свойств изображений — наличием протяженных областей с медленным изменением яркости. На таких участках, которых на реальных изображениях обычно большинство, размах значений элементов во фрагменте W сравнительно невелик — в несколько раз меньше полного диапазона, равного $(x_{\max} + 1)$ градаций. Добавив к каждой из гистограмм $h_{(ijk)(m,n,k+1)}^F(x)$ по 2 ячейки для запоминания минимального и максимального значений распределения, и, соответственно, обрабатывая лишь указываемый диапазон градаций, удастся дополнительно в несколько раз сократить общее число операций.

7.6. Параллельный алгоритм вычисления порядковых статистик

Ниже излагается достаточно простой параллельный алгоритм, который при мультипроцессорной реализации обеспечивает вычисление гистограммы и порядковых статистик

цифрового сигнала за конечное и сравнительно небольшое число операций, не зависящее от размеров фрагмента [55,183]. В дальнейшем будем для простоты рассматривать двумерный цифровой сигнал, квантованный на K уровней, однако описываемый алгоритм с очевидностью может быть расширен и для аналогичного трехмерного сигнала.

Пусть каждый элемент цифрового сигнала принимает значение, равное одному из K уровней квантования в диапазоне $0 \leq k < K$. Скользящая гистограмма $G(W_{mn};k)$ есть дискретная функция от k , указывающая число элементов исходного сигнала уровня k , попадающих в прямоугольный фрагмент W_{mn} , размеры которого H строк по L элементов в строке. Положение фрагмента определяется координатами текущего обрабатываемого элемента; для простоты будем считать, что оно задается координатами левого верхнего угла фрагмента (m,n) . Под интегральной гистограммой $F(W;k)$ понимается сумма

$$F(W;k) = \sum_{i=0}^k G(W;i), \quad (7.6.1)$$

а порядковые статистики $R(W;q)$, где $0 \leq q \leq HL$, представляют собой зависимость вида

$$R(W;q) = k, \quad \text{если } F(W;k-1) < q \leq F(W;k). \quad (7.6.2)$$

Формирование гистограммы по фрагменту W_{mn} требует, вообще говоря, $H \times L$ операций на элемент сигнала. Значительно более быстродействующими являются способы скользящего формирования гистограммы, в которых при переходе от элемента (m,n) к элементу $(m,n+1)$ используются результаты, полученные на предыдущем шаге.

Наиболее известен алгоритм [275], основанный на том, что при скользящем переходе от фрагмента W_{mn} к фрагменту $W_{m,n+1}$ из W_{mn} исключается столбец элементов S_{mn} и добавляется столбец элементов $S_{m,n+L}$, а центральное множество элементов W_0 сохраняется. Алгоритм состоит в следующем. При переходе от W_{mn} к $W_{m,n+1}$ из имеющейся гистограммы $G(W_{mn};k)$ вычитаются единицы в позициях $a_{mn}, \dots, a_{m+n-1,n}$, соответствующих значениям элементов левого столбца S_{mn} , и прибавляются единицы в позициях $a_{m,n+L}, \dots, a_{m+n-1,n+L}$, соответствующих значениям элементов правого столбца $S_{m,n+L}$. Порядковые статистики определяются по гистограмме путем последовательного прибавления (вычитания) соседних значений гистограммы и сравнением сформированных сумм со значениями искомым порядковым статистик. Данный алгоритм при переходе от точки (m,n) к точке $(m,n+1)$ требует выполнения $2H(1+t)+ct$ операций, где t — число искомым порядковым статистик.

В [249] рассматривается алгоритм, согласно которому одновременно формируются Q гистограмм по числу разрядов значения сигнала ($Q = \log_2 K$). Число операций, требуемых на формирование t порядковых статистик, составляет величину $Q(2H+t)$. В [79] для подсчета медианы по фрагменту осуществляется формирование гистограмм одновременно по нескольким (в среднем $\log_2 H$) полосам. Требуемое число операций приблизительно равно

$2H+(1+z)\log_2 H$, где z — среднее значение разности медиан для фрагментов W_{mn} и $W_{m,n+1}$. Известны и другие алгоритмы вычисления гистограмм и порядковых статистик, но во всех из них число требуемых операций не лучше, чем $2H(1+C)$, где $C > 0$ и может зависеть от H .

Алгоритмы вычисления медианы, реализуемые на основе сортирующих сетей [360], позволяют осуществлять обработку данных в темпе поступления. Недостаток их состоит в существенных аппаратных затратах: требуется $N(N-1)/2$ сортирующих элементов, где N — число элементов данных во фрагменте. Так, для фрагмента размером 10×10 потребовалось бы построить вычислительную структуру с 50000 сортирующих элементов.

Отличие предлагаемого ниже алгоритма состоит в том, что во всех операциях участвуют не сами значения элементов сигнала, а их *гистограммы* по фрагментам и столбцам и даже гистограммы отдельных элементов. Это позволяет проводить вычисления независимо для всех K уровней квантования сигнала и, как следствие, осуществить их распараллеливание для K процессорных элементов. Основу алгоритма составляют следующие соотношения, в которых действия осуществляются над значениями гистограмм по фрагментам W_{mn} и столбцам S_{mn} :

$$G(W_{mn};k) = G(W_0;k) + G(S_{mn};k); \quad (7.6.3)$$

$$G(W_{m,n+1};k) = G(W_{mn};k) - G(S_{mn};k) + G(S_{m,n+L};k), \quad (7.6.4)$$

где W_0 — часть фрагмента, которая сохраняется при переходе от точки (m,n) к точке $(m,n+1)$. Аналогично, для столбца элементов S_{mn} при переходе от строки t к строке $t+1$ исключается элемент a_{mn} , добавляется элемент $a_{m+H,n}$ и сохраняется центральная часть столбца S_0 :

$$G(S_{mn};k) = G(S_0;k) + G(a_{mn};k); \quad (7.6.5)$$

$$G(S_{m+1,n}) = G(S_{mn};k) - G(a_{mn};k) + G(a_{m+H,n};k). \quad (7.6.6)$$

Отметим, что гистограмма $G(a_{mn};k)$ одного элемента a_{mn} содержит единицу в позиции $k = a_{mn}$ и нули в остальных точках.

Таким образом, для пересчета гистограммы при переходе от точки (m,n) к точке $(m,n+1)$ необходимо выполнить действия, описываемые формулами (7.6.4) и (7.6.6) для каждого из уровней квантования цифрового сигнала k , т. е. всего $4K$ операций.

Выражения (7.6.3)–(7.6.6) останутся также справедливыми и в том случае, если вместо гистограмм $G(W;k)$ и $G(S;k)$ подставить интегральные гистограммы $F(W;k)$ и $F(S;k)$ (7.6.1):

$$F(S_{m+1,n};k) = F(S_{mn};k) - F(a_{mn};k) + F(a_{m+H,n};k); \quad (7.6.7)$$

$$F(W_{m,n+1};k) = F(W_{mn};k) - F(S_{mn};k) + F(S_{m,n+L};k). \quad (7.6.8)$$

В этом случае гистограмма $G(W;k)$ может быть получена как разность соседних значений интегральной гистограммы:

$$G(W;k) = F(W;k) - F(W;k-1). \quad (7.6.9)$$

Значения порядковых статистик вида $R(W;q)$ (7.6.2) при заданном q находятся следующим путем: а) устанавливается $u(k) = 1$, если $F(W;k) \geq q$, и $u(k) = 0$ в остальных случаях; б) определяется значение k , для которого выполняется условие несовпадения соседних значений: $u(k) \neq u(k-1)$.

Возможность распараллеливания алгоритма состоит в том, что гистограмма $F(W;k)$ может рассматриваться как K -мерный вектор и вычисления по формулам (7.6.7)–(7.6.8) могут производиться одновременно для всех K уровней квантования сигнала. Тем самым количество требуемых операций уменьшается до 4, что по порядку величины в H раз меньше, чем у последовательного алгоритма [275]. При реальных размерах фрагмента $H = 20 \div 50$ и более такое повышение быстродействия весьма ощутимо.

Параллельный алгоритм может быть реализован на мультипроцессоре, структурная схема которого показана на Рис. 7.5.1. Мультипроцессор содержит блок P из K процессорных элементов, каждый из которых осуществляет вычисления для отдельного уровня сигнала, блок регистров R и блок памяти гистограмм по столбцам B , имеющий матрицу памяти из KN элементов, где N — длина строки обрабатываемого сигнала; управление работой и синхронизацией элементарных процессоров осуществляется управляющим процессором CP , обеспечивающим также обработку и обмен данными, находящимися во внешней (оперативной) памяти M . Управление процессорными элементами, а также обмен данными между ними и управляющим процессором происходит через шины C (управления) и D (данных).

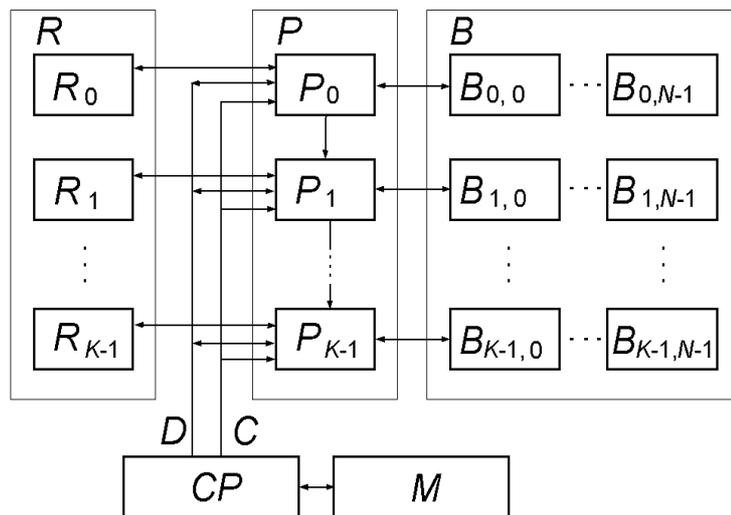


Рис. 7.5.1. Блок-схема мультипроцессора

Перейдем к рассмотрению алгоритма скользящего вычисления гистограммы. Рассмотрим момент времени, когда интегральная гистограмма $F(W_{mn};k)$ по фрагменту W_{mn} уже вычислена. Каждый элемент памяти B_{kn} содержит значение интегральной гистограммы $F(S_{mn};k)$ по столбцу S_{mn} , соответствующее уровню квантования k . Тем самым вся гистограмма по столбцу S_{mn}

содержится в столбце элементов памяти $B_{0n}, \dots, B_{K-1,n}$. Каждый из регистров R_k содержит сумму значений элементов $B_{kn}, \dots, B_{k,n+L-1}$. Таким образом, все регистры вместе составляют вектор значений гистограммы $F(W_{mn}; k)$ по фрагменту W_{mn} .

Для пересчета гистограммы по формуле (7.6.8) необходимо вычесть из вектора R_0, \dots, R_{K-1} вектор $B_{0n}, \dots, B_{K-1,n}$ и прибавить вектор $B_{0,n+L}, \dots, B_{K-1,n+L}$. Эти действия выполняются одновременно всеми процессорами блока P , что составляет два такта.

Пересчет гистограммы по столбцам $F(S_{mn}; k)$ при переходе от столбца S_{mn} к столбцу $S_{m+1,n}$ осуществляется следующим образом. В регистры блока P первоначально заносится вектор $B_{0n}, \dots, B_{K-1,n}$. На шину данных D подается значение элемента сигнала a_{mn} , которое берется из внешней памяти M . Каждый из элементарных процессоров P_k сравнивает значение на шине с имеющимся номером процессора k и вычитает единицу из B_{kn} , если $k \geq a_{mn}$. Затем на шину D подается значение сигнала $a_{m+H,n}$ и процессор P_k прибавляет к ранее полученному значению единицу, если $k \geq a_{m+H,n}$. Полученный результат записывается в тот же столбец $B_{0n}, \dots, B_{K-1,n}$. Эта процедура требует четырех тактов; таким образом, полный цикл пересчета гистограммы в расчете на один элемент требует шести тактов управляющего процессора CP независимо от размеров фрагмента.

В дальнейших вычислениях необходим обмен данными между соседними процессорами. Будем предполагать, что каждый из процессоров P_k может передать некоторое значение z_k соседнему процессору P_{k+1} . Тогда в силу (7.6.9) формирование гистограммы $G(W; k)$ будет происходить за две операции: а) в качестве z_k передается значение из регистра R_k ; б) вычисляется разность значений R_{k+1} и z_k .

Вычисление значения порядковых статистик (7.6.2) осуществляется следующим образом. На шину D подается требуемое значение q , которое сравнивается каждым из процессоров P_k со значением в регистре P_k , и формируется параметр x_k , равный единице, если $R_k \geq q$, и нулю в противном случае; значение x_k передается соседнему процессору P_{k+1} в качестве z_k . Далее каждый процессор P_k сравнивает значения x_k и z_{k-1} и, если $x_k \neq z_{k-1}$, на шину выдается значение номера процессора k , которое считывается управляющим процессором CP . Это условие выполняется только для одного из процессоров P_k , а значит, интерференция исключена. Для определения значения порядковой статистики по уже вычисленной интегральной гистограмме достаточно четырех тактов.

Таким образом, общее число тактов, требуемых для пересчета интегральной гистограммы и определения значений t порядковых статистик, не зависит от размеров фрагмента и равно $6+4t$. Полученные значения являются параметрами преобразования сигнала a_{mn} , находящегося в памяти M , которое выполняется управляющим процессором CP . Рассмотренный параллельный

алгоритм может быть использован для разработки мультипроцессора, обеспечивающего обработку сигнала в темпе поступления данных.

Разработка такого мультипроцессора является вполне разрешимой задачей, поскольку реальное число уровней квантования сигналов, для которых требуется подобная обработка (звуковых, видеосигналов и пр.), сравнительно невелико и, как правило, не превышает величины $K=256$. Сами же процессорные элементы должны выполнять лишь простейшие арифметико-логические операции.

Нами рассмотрена схема алгоритма для двумерного изображения, но без труда она переносится и на трехмерные изображения. В этом случае вместо столбца элементов S_{mn} в (7.6.3)–(7.6.8) следует рассматривать торцевую грань параллелепипеда в направлении смещения, и кроме того добавляется еще один уровень формирования гистограмм по граням из гистограмм по столбцам. В трехмерном случае для пересчета интегральной гистограммы и определения значений t порядковых статистик потребуется $8+4t$ тактов.

Изложенные в настоящей главе результаты опубликованы в работах [55, 175, 183, 184, 198, 242].

7.7. Выводы и результаты

Исследованы особенности перехода от двумерных (2D-) изображений к трехмерным (объемным, 3D-) изображениям. Определены области анализа, виды возможных трехмерных окрестностей и соседства элементов для таких изображений. Рассмотрены возможности модификации модели и алгоритмов обработки и анализа изображения. Предложены модификации быстрых алгоритмов вычисления локальных статистических параметров.

Осуществлено расширение двухмасштабной многокомпонентной модели изображения в применении к трехмерной видеоинформации.

На примере преобразования Фурье иллюстрирована общая схема модификация алгоритмов частотной фильтрации при переходе в 3D.

Модификация алгоритмов пространственной обработки и анализа показана на примере ряда важных методов и алгоритмов, среди которых: класс методов фильтрации, использующих оценку сглаженного значения сигнала по локальной области анализа, различные операторы контурных перепадов, алгоритм фильтрации импульсных помех, алгоритм декомпозиции изображения, алгоритм обнаружения объектов заданного объема.

Показаны модификации алгоритмов быстрого вычисления сумм и порядковых статистик значений элементов внутри прямоугольного параллелепипеда для трехмерных изображений. На одну точку вычисление сумм по n параллелепипедам произвольного размера требует не более

$3+7n$ операций; вычисление порядковых статистик — $2H \times L$ операций, где $H \times L$ — размер грани параллелепипеда перпендикулярной направлению смещения, но не более $2(x_{\max}+1)$, где $(x_{\max}+1)$ — число градаций яркости в изображении.

Разработан параллельный алгоритм вычисления гистограммы и порядковых статистик по скользящему фрагменту для двумерных и трехмерных изображений. Предложена схема его реализации на основе группы процессоров с простыми арифметико-логическими операциями. Требуемое число процессоров равно числу градаций яркости сигнала. Предложенная схема обеспечивает высокую скорость вычисления, не зависящую от размеров фрагмента. При переходе от текущего к соседнему элементу для пересчета гистограммы и определения значений t порядковых статистик требуется $6+4t$ тактов для двумерного и $8+4t$ тактов для трехмерного изображения.

ЗАКЛЮЧЕНИЕ И ОСНОВНЫЕ РЕЗУЛЬТАТЫ

В диссертационной работе решена научная проблема построения двухмасштабной модели изображения и разработки на ее основе новых эффективных методов и алгоритмов обработки и анализа видеоинформации, имеющая важное теоретическое и прикладное значение. Созданные алгоритмы реализованы и исследованы на базе разработанного автором специализированного программного комплекса обработки изображений. Получены следующие основные результаты:

Разработана двухмасштабная многокомпонентная математическая модель цифрового изображения, согласно которой изображение в пределах локальных областей анализа представляется как сумма разрывной кусочно-гладкой компоненты, аппроксимируемой участками аналитических поверхностей, и случайных текстурно-детальной и шумовой составляющих. Модель описывает свойства изображения для областей анализа как малого размера (масштаба), в пределах нескольких шагов дискретизации, так и большого масштаба, соразмерного объектам изображения. Исследования модели показали ее соответствие реальной видеоинформации. Проведено расширение модели на цветные и многоканальные изображения. Сделан вывод, что двухмасштабная многокомпонентная модель изображения может являться фундаментом для исследования и построения новых и эффективных алгоритмов обработки и анализа видеоинформации.

Построена вероятностная модель двумерного контурного изображения, базирующаяся на простой аксиоматике задания свойств границ объектов, и имеющая три параметра: вероятность появления контурной линии и вероятности узловых точек двух видов из трех возможных. Разработан алгоритм источника, позволяющего формировать дискретные контурные изображения с задаваемыми вероятностными характеристиками. Проиллюстрировано порождение контурных изображений при различных сочетаниях задаваемых вероятностей. Показана возможность получения многоградационных изображений со случайным рисунком областей.

Дана интерпретация сложности изображения как характеристики, отражающей число, размеры и заметность деталей на изображении. Исследована возможность оценивания сложности изображения на основе использования аппарата двумерных вариаций. Показана редукция многомерных вариаций непрерывных функций к двумерным дискретным функциям. Введена новая характеристика, соответствующая среднему периметру деталей изображения, и предложена ее модификация, названная *показатель размеров объектов*. Проанализированы изменения значений двумерных вариаций при различных трансформациях изображения. Теоретические выводы подтверждены экспериментально. Исследовано влияние шума и геометрических трансформаций на оценку сложности изображения. Показано, что комбинация

показателя числа объектов и показателя размеров объектов хорошо отражает морфологическую структуру изображения и может являться оценкой сложности изображения.

Исследована задача сглаживания изображения при условии сохранения контурных перепадов. Показано, что в рамках принятой двухмасштабной многокомпонентной модели она соответствует задаче разделения изображения на компоненты с различным информационным содержанием.

На основе двухмасштабной многокомпонентной модели разработаны метод и алгоритм декомпозиции изображения на две различные информационные составляющие — кусочно-гладкую компоненту, несущую информацию о яркостях протяженных областей и границах между ними, и текстурно-детальную компоненту, содержащую мелкие детали, текстуру и шум. Как и модель изображения, алгоритм декомпозиции двухмасштабный: первоначально осуществляется анализ по малой области (окрестности), затем по большой (фрагменту). По сути такой подход позволяет учесть при анализе не только распределения значений признаков (яркости), но и пространственную корреляцию.

Проведено сравнение предложенного алгоритма декомпозиции с другими алгоритмами сглаживания. Исследовано изменение сложности и точности восстановления изображения, искаженного шумом, в зависимости от алгоритма и величины области анализа. Показано, что с увеличением области анализа ошибка восстановления изображения алгоритмом декомпозиции стремится к нулю, при этом оценка сложности приближается к константному ненулевому значению, что означает сохранение деталей. При использовании локального среднего или локальной медианы ошибка только возрастает, а к нулю стремится сложность изображения, что свидетельствует о разрушении деталей изображения. Показано, что алгоритм декомпозиции обеспечивает в $2 \div 2,5$ раза более точное восстановление искаженного шумом изображения по сравнению с другими алгоритмами сглаживания.

Предложены модификации алгоритма декомпозиции для сглаживания цветных и многоканальных изображений. Рассмотрено несколько вариантов реализации: через введение метрики в цветовом пространстве или через проецирование и покомпонентный анализ; последний вариант менее точен, но более эффективен в вычислительном отношении.

На основе модели и метода декомпозиции предложен новый подход к анализу и обработке изображений, заключающийся в том, что для анализа следует выбирать лишь ту компоненту изображения, в которой должна содержаться искомая информация. При обработке возможно применение преобразований, специфических для каждой из полученных компонент.

Проведены исследования методов фильтрации, коррекции и улучшения изображений, разработанных на основе двухмасштабной многокомпонентной модели. Предложены новые алгоритмы фильтрации, а также способы автоматического определения необходимых параметров для существующих методов. Показано, что постулатов локальной модели изображения в ряде случаев достаточно для построения не только локальных, но и глобальных методов обработки видеoinформации.

Предложен ранговый алгоритм фильтрации импульсных помех с динамическим нахождением пороговых параметров для каждого элемента изображения, вычисляемых на основе рангового анализа элементов окрестности. Алгоритм показал в 1,5÷2 раза более высокую точность восстановления изображения по сравнению с другими известными методами.

Для фильтрации периодических помех в частотной области предложен алгоритм автоматического формирования маски фильтра. При этом энергетический спектр интерпретируется как изображение, а всплески находятся применением разработанных алгоритмов обнаружения объектов по их площади.

Для устранения периодических помех вдоль линий развертки изображения разработан алгоритм пространственного анализа изображения и устранения помех, основанный на корреляционном совмещении, усреднении и пространственной фильтрации сигнала.

Исследована задача автоматической градационной коррекции изображения. Предложена *функция локальных контрастов*, позволяющая оценивать яркостные искажения, и выдвинута гипотеза о ее константности, которая подтверждена экспериментально. Разработан алгоритм градационной коррекции изображений, выравнивающий функцию локальных контрастов. Для случая большой разницы яркостей между объектом и фоном (разрыва шкалы яркостей) предложен алгоритм эквализации клиппированной гистограммы, допускающий реализацию в телевизионном масштабе времени. Проведены исследования методов автоматической градационной коррекции, подтверждающие эффективность предложенных алгоритмов.

Исследованы методы улучшения изображений путем усиления локальных контрастов. На основе многокомпонентной модели изображения сформулирована общая каноническая формула для большинства используемых методов. Проведена классификация алгоритмов повышения локальных контрастов и выполнено их сравнение. Разработан алгоритм повышения локальных контрастов, использующий декомпозицию изображения.

Показаны возможности и предложены различные варианты применения методов градационной коррекции и повышения локальных контрастов для преобразования цветных и многоканальных изображений.

Показано, что двухмасштабная многокомпонентная модель и процедура декомпозиции изображения являются подходящим фундаментом для построения алгоритмов анализа видеоинформации, в числе которых: обнаружение объектов, дефектов, контуров, нахождения различий набора объектов и другие.

На основе многокомпонентной модели и метода декомпозиции предложен новый подход и разработаны алгоритмы решения задачи обнаружения объектов на изображении по их площади, в которой рассмотрены три возможных варианта: обнаружения объектов с площадью меньше задаваемого порога, больше порога и с площадью в задаваемом интервале. Введено понятие «локальной площади», позволяющее обнаруживать тонкие протяженные объекты.

Рассмотрены теоретические аспекты обнаружения различий изображений с малыми изменениями. Предложен метод нахождения различий на изображениях, основанный на предварительном обнаружении объектов.

Предложен новый подход к выделению границ объектов как участков с перепадами яркостей, состоящий в том, что перепады следует обнаруживать не непосредственно по исходному изображению, а по предварительно сглаженной компоненте, формируемой в результате операции декомпозиции. Это повышает слитность контурных линий и значительно снижает количество ошибок первого и второго рода.

Разработан алгоритм быстрого корреляционного совмещения квазирегулярных изображений, требующий $MN(2 + k_m k_n + \nu)$ операций, где MN — размеры изображения, k_m и k_n — число локальных максимумов проекций, ν — число итераций при поиске максимума вблизи точки оптимума (как правило $k \approx 1 \div 3$, $\nu \approx 4 \div 7$). Для реальных изображений размерами 512×512 элементов коэффициент сокращения числа операций по сравнению с полным перебором составляет величину порядка 10^4 .

Исследована и решена важная научно-техническая задача обнаружения широкого класса дефектов микросхем по снимкам, формируемым оптическим или сканирующим электронным микроскопами. Разработаны соответствующие алгоритмы и комплексы программ для обнаружения дефектов, в которых использованы алгоритмы совмещения и обработки изображений, построенные на основе предложенной двухмасштабной многокомпонентной модели изображения. Тестирование показало полное соответствие алгоритма предъявляемым требованиям к обнаружению дефектов в 92% и частичное — в 6,25% случаев.

Исследовано применение модели и разработанных методов обработки видеоинформации для решения проблемы сегментации изображений — разбиения изображения на множество

связных областей в соответствии с выбранными критериями. Исследованы два подхода к ее решению.

Первой рассмотрена научно-техническая задача сегментации снимков микросхем, получаемых электронным микроскопом. Особенность задачи в том, что допускается наличие двух различных типов контурных границ между объектами и возможно совпадение значений яркостей соседних сегментов. Разработан алгоритм, основанный на алгоритме декомпозиции и обнаружения объектов, позволяющий в результате привести два типа границ к одному виду. Он включает стадии формирования выровненных изображений с контурными откликами для двух допустимых типов границ, прослеживания линий, построения единой карты контуров и ее морфологической обработки. Результатом является разбиение изображения на связные области, в которых каждому элементу приписан один и тот же индекс. Тестирование алгоритма показало, что удовлетворяющие требуемым критериям точности разделения объектов результаты достигаются в 98,3% случаев.

Второй задачей рассмотрена цвето-текстурная сегментация изображений общего вида, в которой требуется разбиение изображения на однородные области, отличающиеся от соседствующих по своим яркостным, цветовым и/или текстурным характеристикам. В качестве критерия однородности или различия элементов и областей изображения выбрано расстояние между их образами в пространстве признаков. Предложен способ построения пространства яркостно-цвето-текстурных признаков и выбора метрики в нем. В качестве текстурных характеристик выбраны ориентированные многомасштабные локальные вариации и локальные осцилляции градиента. Разработан иерархический алгоритм сегментации, использующий построенную метрику.

Предложенный иерархический подход, основанный на анализе расстояний в пространстве признаков, показал эффективность сегментации изображений. Он допускает использование различных цветовых, текстурных или обобщенных метрик. Разработанный алгоритм может быть применен и к другим одно- или многоканальным массивам данных, удовлетворяющим свойствам изображений. Возможна его модификация путем использования различных других признаков при условии выбора соответствующей метрики.

Показаны возможности использования разработанных методов для решения задач обработки и анализа видеоданных. Решена важная научно-техническая задача автоматического анализа и извлечения необходимой информации из видеопоследовательностей, формируемых компьютерным капилляроскопом, позволяющая определять статические и динамические параметры капиллярного кровотока. Разработан комплекс алгоритмов и программ для

обработки, анализа и извлечения необходимой информации из получаемой видеопоследовательности.

Процедура предварительной обработки видеок кадров включает устранение искажений и помех, выравнивание фона кадра и усиления локальных контрастов видеосигнала. Разработан алгоритм компенсации дрейфа кадров, осуществляющий их пространственную синхронизацию на основе разработанного быстрого алгоритма совмещения. Усредненное изображение формируется суммированием синхронизированных кадров.

Анализ получаемого усредненного изображения включает алгоритмы построения карты контурных линий, ее синтаксической фильтрации, выбора мажоритарного капилляра, формирования векторного представления контурных линий, составления их морфологических характеристик для последующей интерпретации. В результате находится область мажоритарного капилляра, ограниченная параметрически заданными линиями. Получаемые морфологические характеристики позволяют делать выводы о форме капилляра и определять его статические параметры.

Предложен способ выполнения гомеоморфного распрямляющего преобразования области изображения, часть границ которой составляют параметрически заданные и аналитически не описываемые линии. Такое преобразование обеспечивает эластичную деформацию ограничиваемого участка и гладко отображает область капилляра в область прямоугольной формы. Используя полученную матрицу отображения из исходных видеок кадров формируется новая видеопоследовательность, содержащая только интересующую область выбранного капилляра, по которой и производится корреляционный анализ.

Разработан алгоритм формирования карты мгновенных скоростей, являющейся основой для определения динамических параметров капиллярного кровотока, в том числе скорости и объема протекающей крови как функций времени и местоположения в капилляре. Вероятность правильного обнаружения мажоритарного капилляра, проведения контурных линий и определения скорости кровотока составляет около 87%.

Исследованы особенности перехода от двумерных (2D-) изображений к трехмерным (объемным, 3D-) изображениям. Определены области анализа, виды возможных трехмерных окрестностей и соседства элементов для таких изображений. Показаны модификации модели изображения и различных алгоритмов обработки и анализа изображений.

Осуществлено расширение двухмасштабной многокомпонентной модели изображения применительно к трехмерной видеоинформации. На примере преобразования Фурье показана общая схема модификации алгоритмов частотной фильтрации при переходе в 3D. Модифика-

ции алгоритмов пространственной обработки и анализа показаны на примерах важных методов и алгоритмов, среди которых: класс методов фильтрации, использующих оценку сглаженного значения сигнала по локальной области анализа, различные контурные операторы, алгоритм фильтрации импульсных помех, алгоритм декомпозиции изображения, алгоритм обнаружения объектов заданного объема.

Показаны модификации алгоритмов быстрого вычисления сумм и порядковых статистик значений элементов внутри прямоугольного параллелепипеда для трехмерных изображений. На одну точку вычисление сумм по n параллелепипедам произвольного размера требует не более $3+7n$ операций; вычисление порядковых статистик — $2H \times L$ операций, где $H \times L$ — размер грани параллелепипеда перпендикулярной смещению, но не более $2(x_{\max}+1)$, где $(x_{\max}+1)$ — число градаций яркости в изображении.

Разработан параллельный алгоритм вычисления гистограммы и порядковых статистик по скользящему фрагменту для двумерных и трехмерных изображений. Предложена схема его реализации на основе группы процессоров с простыми арифметико-логическими операциями. Требуемое число процессоров равно числу градаций яркости сигнала. Предложенная схема обеспечивает высокую скорость вычисления, не зависящую от размеров фрагмента. При переходе от текущего к соседнему элементу для пересчета гистограммы и определения значений t порядковых статистик требуется $6+4t$ тактов для двумерного и $8+4t$ тактов для трехмерного изображения.

Полученные результаты подтверждают адекватность двухмасштабной многокомпонентной модели реальным изображениям и эффективность предложенного на ее основе подхода, состоящего в декомпозиции изображения на самостоятельные информационные составляющие с последующим специфическим для каждой из компонент анализом или обработкой. Потенциал модели и подхода продемонстрирован построением новых и модификацией известных методов обработки и анализа видеоинформации.

Разработанные методы и алгоритмы преобразования видеоинформации были реализованы и исследованы на базе созданного автором специализированного проблемно-ориентированного программного комплекса обработки изображений «IRBIS». Все приведенные в диссертации результаты экспериментальных исследований были в разное время получены на его основе. Данный специализированный комплекс включает программы, реализующие как изложенные в диссертации, так и многие другие важные и распространенные алгоритмы преобразования видеоданных. Он предоставляет пользователю специализированную операционную среду для разработки новых алгоритмов и программ обработки и анализа видеоинформации.

Разработанный автором специализированный комплекс программ обработки изображений «IRBIS» используется в Институте проблем передачи информации РАН в разных модификациях с 1988 г. по настоящее время. В 1992 г. он выставлялся на Международной выставке компьютерных и информационных технологий CeBit (г. Ганновер, Германия) и был признан Национальным номинантом на Гран-при «Golden Softies» в области инноваций и разработки программного обеспечения в Европе (Приложение 3).

Ряд рассмотренных в работе методов и алгоритмов применялся при обработке видеоинформации, переданной с орбитальных и автоматических межпланетных станций «Салют-6», «Марс-4, -5», «Венера-9, -10», «Венера-13, -14», «Венера-15, -16», «Фобос-2» и других. Многие из результатов проведенных исследований были успешно использованы при выполнении работ по темам, выполнявшимся по распоряжениям Президиума Академии наук, договорам с зарубежными компаниями и отечественными организациями. Некоторые иллюстрации результатов приведены в Приложениях 1 и 2.

ЛИТЕРАТУРА

1. Автоматизация анализа и распознавания изображений. Методы и средства. Рига: Зинатне, 1979.
2. Автоматизация анализа и распознавания изображений. Выпуск 2. Рига: Зинатне, 1980.
3. Автоматизированная обработка изображений природных комплексов Сибири. / Алексеев А.С., Дементьев В.Н., Пяткин В.П. и др. Новосибирск: Наука, Сибирское отделение, 1988, 223 с.
4. Адаптивные методы обработки изображений. Сборник научных трудов / Ред. Сифоров В.И., Ярославский Л.П. М.: Наука, 1988.
5. Александров В.В., Горский Н.Д. Представление и обработки изображений: Рекурсивный подход. Л.: Наука, 1985.
6. Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений. М.: Высшая школа, 1983.
7. Аракчеев А.Г., Гурфинкель Ю.И., Певгов В.Г. Компьютерный капилляроскоп для неинвазивных исследований параметров циркулирующей крови // Московский хирургический журнал, 2010, № 5, С. 27–30.
8. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. М.: Связь, 1980.
9. Беликова Т.П. Некоторые методы цифрового препарирования изображения // Цифровая обработка сигналов и ее применения. М.: Наука, 1981, С. 87–99.
10. Беликова Т.П., Гендлер В.Е., Ярославский Л.П. Геологическое дешифрирование в диалоговом режиме в автоматизированных системах цифровой обработки изображений // Исследование Земли из космоса, 1981, № 3, С. 102–112.
11. Белокуров А.А., Сечко В.В. Стохастические модели в задачах анализа и обработки изображений // Зарубежная радиоэлектроника. 1989, вып. 5, С. 3–18.
12. Безрук А.А., Лебедев Д.С. Исправление изображений, искаженных импульсными помехами // Обработка изображений и дистанционные исследования. Тез. докл. Новосибирск, 1981, С. 150–151.
13. Бендат Дж., Пирсол А. Прикладной анализ случайных данных. М.: Мир, 1989.
14. Боде Г., Шеннон К. Упрощенное изложение линейной минимально-квадратичной теории сглаживания и предсказания // Теория информации и ее приложения. – М.: Физматгиз, 1959, С. 113–137.

15. Бокштейн И.М., Кронрод М.А., Чочиа П.А. Методика составления радиолокационной карты Венеры по данным автоматических межпланетных станций «Венера-15» и «Венера-16» // Иконика. Цифровая обработка видеoinформации. – М.: Наука, 1989, С. 35–60.
16. Бокштейн И.М., Кронрод М.А., Чочиа П.А. Синтез радиолокационной карты Венеры // Современные вопросы информатики, вычислительной техники и автоматизации: Тез. докл. Всесоюзной конф., посвященной Дню советской науки. М., 1985, С. 43.
17. Борилин Б.Л., Чочиа П.А. Реставрация фотодокументов с помощью ЭВМ // Советские архивы, 1980, вып. 3, С. 45–48.
18. Борн М., Вольф Э. Основы оптики. М.: Наука, 1970.
19. Быков Р.Е., Гуревич С.Б. Анализ и обработка цветных и объемных изображений. М.: Радио и связь, 1984.
20. Быстрые алгоритмы в цифровой обработке изображений: Преобразования и медианные фильтры. / Хуанг Т.С., Эклунд Дж.-О., Нуссбаумер Г.Дж. и др. Ред. Хуанг Т.С. М.: Радио и связь, 1984.
21. Бутаков Е.А., Островский В.И., Фадеев И.Л. Обработки изображений на ЭВМ. М.: Радио и связь, 1987.
22. Варакин Л.Е. Теория сложных сигналов. М.: Сов. радио, 1970.
23. Василенко Г.И. Теория восстановления сигналов: о редукации к идеальному прибору в физике и технике. М.: Сов.радио, 1979.
24. Василенко Г.И., Тараторин А.М. Восстановление изображений. М: Радио и связь, 1986.
25. Введение в контурный анализ и его приложения к обработке изображений и сигналов. / Фурман Я.А., Кревецкий А.В., Передреев А.К. и др. М.: Физматлит, 2003.
26. Введение в цифровую фильтрацию. / Ред. Богнер Р., Константи́нидис А. М.: Мир, 1976.
27. Вендровский К.В., Вейцман А.И. Резкость фотографического изображения // Иконика. – М.: Наука, 1968, С. 97–130.
28. Вернер М. Основы кодирования. Учебник для ВУЗов. М.: Техносфера, 2004. – 288 с.
29. Виленчик Л.С. Искажения в цифровом телевидении // Труды НИИР, 1977, № 2, С. 36–40.
30. Виллингсли Ф. Влияние шума аппаратуры цифровой обработки изображений // Обработка изображений и цифровая фильтрация. – М.: Мир, 1979. С. 271–307.
31. Виткус Р.Ю., Ярославский Л.П. Адаптивные линейные фильтры для обработки изображений // Адаптивные методы обработки изображений. – М., Наука, 1988, С. 6–35.
32. Виттих В.А., Сергеев В.В., Соифер В.А. Обработка изображений в автоматизированных системах научных исследований. М.: Наука, 1982.
33. Витушкин А.Г. О многомерных вариациях. М.: Гостехиздат, 1955.

34. Витушкин А.Г. Оценка сложности задачи табулирования. М.: Физ.-мат. лит., 1959.
35. Возенкрафт Дж., Джейкобс И. Теоретические основы техники связи. М.: Мир, 1968.
36. Возможности использования неинвазивной компьютерной капилляроскопии в космической медицине и в клинической практике / Гурфинкель Ю.И., Каце Н.В., Макеева О.В., Михайлов В.М. // Методы нелинейного анализа в кардиологии и онкологии. Вып. 2. – М.: Университет, 2010, С. 111–121.
37. Гаек Я., Шидак З. Теория ранговых критериев. М.: Наука, 1979.
38. Геоморфологическая характеристика северного полушария Венеры / Барсуков В.Л., Базилевский А.Т., Богомолов А.Ф., Чочиа П.А. и др. // Геотектоника, 1986, № 4, С. 4-25.
39. Голд Б., Рэйдер Ч. Цифровая обработка сигналов. М.: Сов. радио, 1973.
40. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2012.
41. Горелик А.Л., Скрипкин В.А. Методы распознавания: Учебное пособие. М.: Высшая школа, 1989.
42. Гранат Д.Дж. Роль моделей зрения человека в обработке изображений // ТИИЭР, 1981, т. 69, вып. 5, С. 65-77.
43. Гренандер У. Лекции по теории образов. М.: Мир. т. 1: Синтез образов. 1979, т. 2: Анализ образов. 1981, т. 3: Регулярные структуры. 1983.
44. Грэхэм Д.Н. Передача изображений посредством кодирования двумерных контуров // ТИИЭР, 1967, т. 55, вып. 3, С. 102–113.
45. Гурфинкель Ю.И., Каце Н.В., Макеева О.В., Михайлов В.М. Возможности использования неинвазивной компьютерной капилляроскопии в космической медицине и в клинической практике. Методы нелинейного анализа в кардиологии и онкологии. Вып. 2. М.: Университет, 2010, С. 111–121.
46. Гурфинкель Ю.И., Кузнецов М.И., Певгов В.Г. Устройство и способ неинвазивного исследования характеристик капилляров и капиллярного кровотока. Патент РФ № 2008123909/14.
47. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988.
48. Дворкович А.В., Дворкович В.П. Цифровые видеоинформационные системы (теория и практика). М.: Техносфера, 2012.
49. Джайн А.К. Успехи в области математических моделей для обработки изображений // ТИИЭР, 1981, Т. 69, № 5, С. 9–39.
50. Джайн Р., Аггарвал Дж.К. Машинный анализ сцен, состоящих из объектов криволинейных очертаний. ТИИЭР, 1979, Т. 67, № 5, С. 121–130.

51. Динамический радиографический контроль с использованием отсеивающего раstra и цифровой обработки изображений / Веселовский Л.Н., Гусев Е.А., Петушков А.А., Соснин Ф.Р., Рузинова Е.В., Каплун Я.М., Чочиа П.А., Санпитер И.А. // Дефектоскопия, 1984, вып. 6, С. 29–32.
52. Домненко В.М., Бурсов М.В. Моделирование формирования оптического изображения. Учебное пособие. СПб.: СПбГУ ИТМО, 2005.
53. Дуда Р., Харт Я. Распознавание образов и анализ сцен. М.: Мир, 1976.
54. Дэйвид Г. Порядковые статистики. М.: Наука, 1979.
55. Елманов С.А., Чочиа П.А., Устройство для вычисления порядковых статистик. Патент СССР № 1704148 А1, кл. G 06 F 15/36, 1989. (АС СССР № 1196897, кл. G 06 F 15/36, 1984; АС СССР № 1354210, кл. G 06 F 15/36, 1986).
56. Журавлев Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики, вып. 33. – М.: Наука, 1978, С. 5–68.
57. Журавлев Ю.И., Гуревич И.Б.. Распознавание образов и анализ изображений // Искусственный интеллект: в 3-х книгах. Книга 2. Модели и методы: Справочник. М.: Радио и связь, 1990.
58. Завалишин Н.В., Мучник И.Б. Модели зрительного восприятия и алгоритмы анализа изображений. М.: Наука, 1974.
59. Иванов Л.Д. Вариации множеств и функций. М.: Наука, 1975.
60. Иконика. Теория и методы обработки изображений. М.: Наука, 1983.
61. Иконика. Цифровая обработка видеоинформации. М.: Наука, 1989.
62. Капилляроскоп «С12». URL:
<http://casmos.ru/index.php/ru/production/capillarscope> .
63. Капилляроскоп общеклинического применения «Капилляроскан-1». URL:
<http://www.aetechnologies.ru/production/2-details.htm> .
64. Карнаухов В.Н., Милюкова О.П., Чочиа П.А. Спектральные свойства искаженных изображений и системы распознавания. // Математические методы распознавания образов (ММРО-13). Докл. 13-й Всероссийской конф., М.: МАКС Пресс, 2007, С. 328–331.
65. Ким В. Ранговые алгоритмы сглаживания изображений // Анализ сложных информационных систем. – М.: ИППИ АН СССР, 1984, ч. 2, С. 11–13.
66. Ким В., Ярославский Л.П. Ранговые алгоритмы обработки изображений // Адаптивные методы обработки изображений. – М.: Наука, 1988, С. 35–73.
67. Ковалевский В.А. Методы оптимальных решений в распознавании изображений. М.: Наука, 1976.

68. Кодирование и обработка изображений. М.: Наука, 1988.
69. Колмогоров А.Н. Три подхода к определению понятия «количество информации». // Проблемы передачи информации, 1965, Т. 1, вып. 1, С. 3–11.
70. Крамер Г. Математические методы статистики. М.: Мир, 1975.
71. Красиков В.Л., Хатунцева М.В., Штарьков Ю.М. Метод главных компонент в задачах обработки многозональных данных // Аэрокосмические исследования Земли. Обработка видеоинформации на ЭВМ. М.: Наука, 1978, С. 143–151.
72. Красильников Н.Н. Цифровая обработка 2D- и 3D-изображений. С.-Пб.: ВНУ-Петербург, 2011.
73. Крашенинников В.Р. Основы теории обработки изображений: Учебное пособие. // Ульяновск: УлГТУ, 2003. 150 с.
74. Кривошеев М.И., Кустарев А.К. Цветовые измерения. М.: Энергоатомиздат, 1990.
75. Кронрод А.С. О функциях двух переменных. УМН, 1950, т. 5, № 1, С. 24–134
76. Кронрод М.А. Несколько задач обработки изображений // Вопросы кибернетики; вып. 38. Иконика. Цифровая обработка и фильтрация изображений. – М.: Наука, 1978, С. 49–59.
77. Кронрод М.А., Чочиа П.А. Математическое обеспечение диалоговой системы обработки изображений // Иконика. Теория и методы обработки изображений. – М.: Наука, 1983, С. 87-99.
78. Кронрод М.А., Чочиа П.А. Фильтрация помех на изображении с использованием медианы распределения // Иконика. Теория и методы обработки изображений. – М.: Наука, 1983, С. 100–108.
79. Кучеренко К.И., Очин Е.Ф. Мультимикроспроцессорная система медианной фильтрации изображений с поразрядным поиском медианы // Автометрия, 1986, № 1, С. 53–56.
80. Лазерная доплеровская флоуметрия микроциркуляции крови. / Ред.: Крупаткин А.И., Сидоров В.В. М.: Медицина, 2005.
81. Ландау Л.Д., Лифшиц Е.М. Теория поля. М.: Наука, 1967.
82. Лапий В.Ю., Калюжный А.Я., Красный Л.Г. Устройства ранговой обработки информации. Киев: Техніка, 1986.
83. Лебедев Д.С. Статистическая модель изображения // Иконика. Пространственная фильтрация изображений. Фотографические системы. М.: Наука, 1970, С. 53–65.
84. Лебедев Д.С. Теория и методы первичной обработки видеоинформации. Диссертация на соискание ученой степени доктора физико-математических наук в форме научного доклада. М. 1993

85. Лебедев Д.С., Безрук А.А., Новиков В.М. Марковская вероятностная модель изображения и рисунка // Препринт ИППИ АН СССР. М.: ВИНТИ, 1983. 40 с.
86. Лебедев Д.С., Цуккерман И.И. Телевидение и теория информации. М.-Л.: Энергия, 1965.
87. Лебедев Д.С., Ярославский Л.П. Нелинейная фильтрация импульсных помех на изображении. // Иконика. Пространственная фильтрация изображений. Фотографические системы. – М.: Наука, 1970, С. 26–34.
88. Литван Р.И., Аверьянов Ю.И., Быковская Ф.С. Оптимальное градационное преобразование изображений // Техника кино и телевидения, 1979, № 2, С. 38–41.
89. Лучко Д.В., Иванушкина Н.Г. Количественный анализ в коронарной ангиографии: алгоритм автоматизированного определения диаметра сосудов. Электроника и связь, 2009, № 2–3, Киев, С. 161–165.
90. Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. М.: Радио и связь, 1987.
91. Математическая энциклопедия. М.: Советская энциклопедия, 1977, Т. 1-5.
92. Методы компьютерной обработки изображений / Ред. Сойфер В.А. М.: Физматлит, 2001.
93. Методы передачи изображений. Сокращение избыточности / Прэтт У.К., Сакрисон Д.Д., Мусман Х.Г.Д. и др. / Ред. Прэтт У.К. М.: Радио и связь, 1983.
94. Милюкова О.П. Изображение как функция с ограниченной полной вариацией. Иконика. Цифровая обработка видеоинформации. М.: Наука, 1989, С. 19–25.
95. Милюкова О.П., Чочиа П.А. Об оценке сложности изображений с помощью двумерных вариаций. // Информационные процессы, 2012, Т. 12, № 4, С. 362–371.
96. Миркин Л.И. Внутрикадровые статистические характеристики телевизионных изображений. // Вопросы радиоэлектроники, сер. «Техника телевидения», 1975, вып. 3, С. 68–77.
97. Моттль В.В., Мучник И.Б. Скрытые марковские модели в структурном анализе сигналов. М.: Физматлит, 1999, 352 с.
98. Новейшие методы обработки изображений / Потапов А.А., Гуляев Ю.В., Никитов С.А. и др. М.: Физматлит, 2008.
99. Обработка данных полосовой съемки поверхности Венеры, переданных станциями «Венера-15» и «Венера-16» / Богомоллов А.Ф., Скрыпник Г.И., Бокштейн И.М., Кронрод М.А., Чочиа П.А., Бергман М.Ю., Кудрин Л.В., Башнин А.В. // Космические исследования, 1985, Т. 23, вып. 2, С. 179–190.
100. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий / Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В. и др. М.: Физматкнига, 2010. 672 с.

101. Обработка и топографическая интерпретация телевизионных панорам, полученных с посадочных аппаратов АМС «Венера-9» и «Венера-10» / Непоклонов Б.В., Лейкин Г.А., Селиванов А.С., Алексашин Е.П., Бокштейн И.М., Кронрод М.А., Чочиа П.А., Ярославский Л.П. // Первые панорамы поверхности Венеры. – М.: Наука, 1979. С.80–106.
102. Обработка изображений. Тематический выпуск. // ТИИЭР, 1981, т. 69, № 5.
103. Обработка изображений и цифровая фильтрация. / Ред. Хуанг Т. М.: Мир, 1979.
104. Обработка изображений при помощи цифровых вычислительных машин / Ред. Эндрюс Г., Инло Л. М.: Мир, 1973.
105. Обработка телевизионных панорам поверхности Венеры, переданных спускаемыми аппаратами станций «Венера-13» и «Венера-14» / Бокштейн И.М., Кронрод М.А., Чочиа П.А., Гектин Ю.М. // Космические исследования, 1983, Т. 21, вып. 2, С.190–199.
106. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. М.: Связь, 1979.
107. Оппенхайм А.В., Лим Дж.С. Важность фазы при обработке сигналов. // ТИИЭР, т. 69, № 5, 1981, С. 39–54.
108. Павлидис Т. Алгоритмы машинной графики и обработки изображений. М.: Радио и связь, 1986.
109. Первые цветные панорамы поверхности Венеры, переданные АМС «Венера-13,-14» / Селиванов А.С., Аваткова Н.А., Бокштейн И.М., Гектин Ю.М., Герасимов М.А., Давыдова И.Е., Кронрод М.А., Нараева М.К., Носов Б.И., Панфилов А.С., Свешникова О.М., Титов А.С., Файнберг И.С., Чемоданов В.П., Чочиа П.А. // Космические исследования, 1983, Т. 21, вып. 2, С. 183–189.
110. Потапов А.С, Малышев И.А., Луцив В.Р. Совмещение аэрокосмических изображений с субпиксельной точностью методом локальной корреляции // Оптический журнал, 2004, т. 71, № 5, С. 31-36.
111. Предобработка и распознавание двумерных изображений / Грибков И.В., Захаров А.В., Кольцов П.П. и др. // Вопросы кибернетики. Распознавание. Нейросети. Виртуальная реальность. – М.: НИИСИ РАН, 1997.
112. Применение цифровой обработки сигналов / Ред. Оппенгейм Э. М.: Мир, 1980.
113. Прэтт У. Цифровая обработка изображений. М.: Мир, 1982, Т. 1, 2.
114. Прэтт У.К., Фожра О.Д., Гагалович А. Применение моделей стохастических текстур для обработки изображений // ТИИЭР, 1981, т. 69, вып. 5, С. 54–64.
115. Психология машинного зрения. / Ред. Уинстон П. М.: Мир, 1978.
116. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978.

117. Радиографический контроль с обработкой изображений методом линейной фильтрации / Гусев Е.А., Петушков А.А., Соснин Ф.Р., Чочиа П.А. // Дефектоскопия, 1984, № 3, С. 46-48.
118. Развитие и использование аэрокосмических методов изучения природных явлений и ресурсов. Новосибирск, 1979.
119. Распознавание образов и обработка изображений. ТИИЭР, 1979, т. 67, № 5.
120. Распознавание образов при помощи цифровых вычислительных машин / Ред. Хармон Л. М.: Мир, 1974.
121. Робертс Л. Автоматическое восприятие трехмерных объектов. // Интегральные роботы. — М.: Мир, 1973, С. 162–208.
122. Розенфельд А. Распознавание и обработка изображений с помощью вычислительных машин. М.: Мир, 1969.
123. Розенфельд А., Дэйвис Л.С. Сегментация и модели изображений // ТИИЭР, 1979, Т. 67, № 5, С. 71–81.
124. Садыков С.С., Кан В.Н., Самандров И.Р. Методы выделения структурных признаков изображений. Ташкент: Фан, 1990.
125. Сергеев В.В., Сойфер В.А. Имитационная модель изображения и метод сжатия данных // Автоматика и вычислительная техника, 1978, № 3, С. 76–78
126. Смирнов А.Я. Математические описания изображений // Сер. Обработка оптических изображений, вып. 1. Л.: ГОИ, 1986. 74 с.
127. Стокхэм Т.Дж., мл. Обработка изображений в контексте моделей зрения // ТИИЭР, 1972, т. 60, № 7, С. 93–108
128. Структура неоднородностей температуры по наблюдениям атмосферной рефракции с орбитальной станции «Салют-6» / Гурвич А.С., Загоруйко С.В., Кан В., Попов Л.И., Рюмин В.В., Савченко С.А., Чочиа П.А. // Доклады АН СССР, 1981, Т. 259, вып.6, С. 1330–1333.
129. Съёмка облачного слоя Венеры с орбитального аппарата станции «Венера-9» / Селиванов А.С., Гектин Ю.М., Кержанович В.В., Нараева М.К., Панфилов А.С., Чемоданов В.П., Чочиа П.А. // Космические исследования, 1978, Т. 16, вып. 6, С.877–885.
130. Теребиж В.Ю. Введение в статистическую теорию обратных задач. М.: Физматлит, 2005.
131. Террайен Ч.У., Куатъери Т.Ф., Даджон Д.Е. Алгоритмы анализа изображений, основанные на статистических моделях // ТИИЭР, 1986, Т. 74, вып. 4. С. 4–25.
132. Техническое зрение роботов / Ред. Пью А. М.: Машиностроение, 1987.
133. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М., Наука, 1979.
134. Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978.
135. Тьюки Дж. Анализ результатов наблюдений. М.: Мир, 1981.

136. Уинстон П.Г. Построение структурных описаний по примерам. Психология машинного зрения. М.: Мир, 1978, С. 185–248.
137. Уинтц П.А. Кодирование изображений посредством преобразований // ТИИЭР, 1972, т. 60, вып. 7, С. 69–83.
138. Уолц Д. Интерпретация контурных рисунков, изображающих сцены с тенями. Психология машинного зрения. М.: Мир, 1978, С. 30–111.
139. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. М.: Вильямс, 2004.
140. Фу К. Последовательные методы в распознавании образов и обучении машин. М.: Наука, 1979.
141. Фу К. Структурные методы в распознавании образов. М.: Мир, 1977.
142. Фукунага К. Введение в статистическую теорию распознавания образов. М.: Наука, 1979.
143. Функциональное преобразование изображения при радиационной дефектоскопии / Гусев Е.А., Леонов Б.И., Петушков А.А., Соснин Ф.Р., Санпитер И.А., Чочиа П.А. // Дефектоскопия, 1984, вып. 10, С. 91–93.
144. Фурман Я.А., Кревецкий А.В., Передреев А.К., и др. Введение в контурный анализ и его приложения к обработке изображений и сигналов. М.: Физматлит, 2003.
145. Харалик Р.М. Статистический и структурный подходы к описанию текстур // ТИИЭР, 1979, т. 67, вып. 5, С. 98–120.
146. Харкевич А.А. Спектры и анализ. М.: Гостехиздат, 1957
147. Хорн Б.К.П. Зрение роботов. М.: Мир, 1989.
148. Хьюбер П. Робастность в статистике. М.: Мир, 1984.
149. Цифровая обработка изображений. ТИИЭР, 1972, т. 60, № 7.
150. Цифровая обработка изображений / Беликова Т.П., Кронрод М.А., Чочиа П.А., Ярославский Л.П. // Поверхность Марса. – М.: Наука, 1980, С. 45–62.
151. Цифровая обработка изображений в информационных системах: Учебное пособие / Грузман И.С., Киричук В.С., Косых В.П. и др. Новосибирск, НГТУ, 2002.
152. Цифровая обработка изображений: Специальный выпуск. // Зарубежная радиоэлектроника, 1985, № 10.
153. Цифровая обработка сигналов и ее применения. М.: Наука, 1981.
154. Цифровая обработка телевизионных и компьютерных изображений / Дворкович А.В., Дворкович В.П., Зубарев Ю.Б. и др. М.: Международный центр научной и технической информации, 1997.

155. Цифровая обработка фотоснимков поверхности Марса, переданных АМС «Марс-4» и «Марс-5» / Беликова Т.П., Кронрод М.А., Чочиа П.А., Ярославский Л.П. // Космические исследования, 1975, Т. 13, № 6, С. 898–906.
156. Цифровая оптика в медицинской интроскопии. М.: Наука, 1992.
157. Цифровое кодирование телевизионных изображений. / Цуккерман И.И., Кац Б.М., Лебедев Д.С. и др. М.: Радио и связь, 1981.
158. Цифровое телевидение / Кривошеев М.И., Виленчик Л.С., Красносельский И.Н. и др. Ред. Кривошеев М.И. М.: Связь, 1980.
159. Цифровой синтез изображения при радиографии объектов сложной конфигурации / Гусев Е.А., Кронрод М.А., Лебедев Д.С., Петушков А.А., Соснин Ф.Р., Чочиа П.А. // Дефектоскопия, 1982, вып. 6, С.86–91.
160. Цифровой синтез рентгенодиагностических изображений / Рабкин И.Х., Гусев Е.А., Петушков А.А., Соснин Ф.Р., Чочиа П.А. и др. // Медицинская техника, 1984, вып.1, С. 15-19.
161. Чочиа П.А. Автоматическая градационная коррекция видеоинформации // Компьютерная оптика, 1995, вып. 14–15, ч. 1, С. 37–45.
162. Чочиа П.А. Анализ видеоданных, формируемых капилляроскопом, и измерение динамики кровотока. // Информационные процессы, 2014, Т. 14, № 1, С. 79–86.
163. Чочиа П.А. Анализ видеопоследовательностей, формируемых капилляроскопом. // Математические методы распознавания образов (ММРО-16): 16-я Всероссийская конф., г. Казань, 2013 г. Тез. докл. – М.: Торус пресс, 2013, С. 56.
164. Чочиа П.А. Базовая система обработки изображений для персональной ЭВМ // Автоматизированные системы обработки изображений. Тез. докл. III Всесоюзной конф. – Л., 1989, С. 30–31.
165. Чочиа П.А. Быстрое корреляционное совмещение квазирегулярных изображений. // Информационные процессы, 2009, Т. 9, № 3, С. 117–120.
166. Чочиа П.А. Вероятностная модель контурного изображения // Иконика. Цифровая обработка видеоинформации. – М.: Наука, 1989, С. 25-34.
167. Чочиа П.А. Выделение областей на изображении // Зрение организмов и роботов: Тез. докл. Всесоюзного симпозиума. Т. 2. – Вильнюс, 1985, С. 94–96.
168. Чочиа П.А. Двумерные вариации как средство оценивания сложности изображений. // Машинное обучение и анализ данных, 2015, Т. 1, № 11. С. 1660–1676.

169. Чочиа П.А. Двухкомпонентная статистическая модель фрагмента изображения // Обработка изображений и дистанционные исследования: Тез. докл. Всесоюзной конф., ч. 1 – Новосибирск, 1984, С. 60–61.
170. Чочиа П.А. Двухмасштабная модель изображения // Кодирование и обработка изображений. – М.: Наука, 1988, С. 69-87.
171. Чочиа П.А. Декомпозиция изображения и возможности ее применения // Обработка изображений и дистанционные исследования: Тез. докл. Региональной конф. – Новосибирск, 1987, С. 191.
172. Чочиа П.А. Исследование диалоговых методов цифровой коррекции искаженных изображений и разработка средств их реализации в автоматизированных системах. Канд. дисс. М.: ИППИ АН СССР, 1982.
173. Чочиа П.А. Методы преобразования изображений, использующие двухмасштабную модель // Кодирование и обработка изображений. – М.: Наука, 1988, С. 98-112.
174. Чочиа П.А. Методы улучшения аэрокосмических изображений, использующие гистограмму по фрагменту // Исследование Земли из космоса, 1985, № 6, С. 66–78.
175. Чочиа П.А. Модификация модели и алгоритмов обработки при переходе от двумерных к трехмерным изображениям // IX Международная научно-практическая конф. Современные информационные технологии и ИТ-образование. Сборник избранных трудов. – М.: МГУ, 2014, С. 820-833.
176. Чочиа П.А. Некоторые алгоритмы обнаружения объектов на основе двухмасштабной модели изображения // Информационные процессы, 2014, Т. 14, № 2, С. 117–136.
177. Чочиа П.А. Обнаружение капилляров на изображениях, формируемых капилляроскопом. // Информационные процессы, 2013, Т. 13, № 2, С. 63–75.
178. Чочиа П.А. Обработка видеоданных, получаемых компьютерным капилляроскопом. // IX Международная научно-практическая конф. Современные информационные технологии и ИТ-образование. Сборник избранных трудов. – М.: МГУ, 2014, С. 834-846.
179. Чочиа П.А. Обработка и анализ изображений на основе двухмасштабной модели: Препринт ИППИ АН СССР. М.: ВИНТИ, 1986, 69 с.
180. Чочиа П.А. Операционная среда системы обработки изображений // Проблемы создания систем обработки, анализа и распознавания изображений. Тезисы докладов конференции. – Ташкент, 1990, С. 17–18.
181. Чочиа П.А. Определение параметров капиллярного кровотока на основе анализа видеоданных. // Медицинская техника, 2015, № 1, С. 14-17.

182. Чочиа П.А. Оценивание сложности изображений двумерными вариациями. // Математические методы распознавания образов ММРО-17. Тезисы докладов 17-й Всероссийской конференции с международным участием. г. Светлогорск, Калининградская обл. М.: Торус пресс, 2015. С. 140.
183. Чочиа П.А. Параллельный алгоритм вычисления скользящей гистограммы // Автометрия, 1990, № 2, С. 40–44.
184. Чочиа П.А. Переход от 2D- к 3D-изображениям: модификация двухмасштабной модели и алгоритмов обработки. // Информационные процессы, 2014, Т. 14, № 3, С. 242–255.
185. Чочиа П.А. Пирамидальный алгоритм сегментации изображений. // Информационные процессы, 2010, Т. 10, № 1, С. 23–35.
186. Чочиа П.А. Предварительная обработка видеопоследовательностей, формируемых капилляроскопом. // Информационные процессы, 2011, Т. 11, № 1, С. 76–85.
187. Чочиа П.А. Предисловие научного редактора перевода // Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005, С.12–14.
188. Чочиа П.А. Предисловие научного редактора перевода. // Гонсалес Р., Вудс Р. Цифровая обработка изображений. 3-е изд. – М.: Техносфера, 2012, С. 19–21.
189. Чочиа П.А. Применение методов цифровой обработки изображений для реставрации архивных документов // Иконика. Теория и методы обработки изображений. – М.: Наука, 1983, С. 115–125.
190. Чочиа П.А. Применение частотной фильтрации изображения для удаления помех, вызванных тиснением фотобумаги // Информационные процессы, 2010, Т. 10, № 4, С. 334-338.
191. Чочиа П.А. Принципы и структурная организация системы обработки изображений // Обработка изображений и дистанционные исследования: Тез. докл. международной конф. – Новосибирск, 1990.
192. Чочиа П.А. Сглаживание изображения при сохранении контуров // Кодирование и обработка изображений. – М.: Наука, 1988, С. 87–98.
193. Чочиа П.А. Сглаживание цветных изображений при сохранении контуров на основе анализа расстояний в цветовом пространстве. // Математические методы распознавания образов (ММРО-13). Докл. 13-й Всероссийской конф. (Дополнение). Ленинградская обл., г. Зеленогорск. – М.: МАКС Пресс, 2007, С.696–698.
194. Чочиа П.А. Сегментация изображений на основе анализа расстояний в пространстве признаков. // Автометрия, 2014, Т. 50, № 6, С. 97–110.

195. Чочиа П.А. Сегментация изображений на основе прослеживания контуров применительно к анализу снимков электронных микросхем. // Информационные процессы, 2009, Т. 9, № 2, С. 93–104.
196. Чочиа П.А. Система обработки изображений для персональной ЭВМ // Проблемы создания систем обработки, анализа и распознавания изображений. Тезисы докладов II Республиканского семинара. – Ташкент, 1989, С. 17–18.
197. Чочиа П.А. Согласование характеристик изображений при составлении фотомозаики // Обработка изображений и дистанционные исследования. Тез. докл. Всесоюзной конф. ч.1. – Новосибирск, 1985, С. 62.
198. Чочиа П.А. Трехмерные и двумерные изображения: модели, алгоритмы и области анализа. // International Journal of Open Information Technologies, 2014, vol. 2, no. 11, pp. 1–9.
199. Чочиа П.А. Улучшение многозональных цветных изображений путем усиления локальных контрастов // Исследование Земли из космоса, 1988, вып. 6, С. 95–99.
200. Чочиа П.А. Фильтрация изображения, искаженного периодической помехой // Всесоюзная конф. «Оптическое изображение и регистрирующие среды». – Л., 1982, С. 151.
201. Чочиа П.А. Цифровая фильтрация импульсных помех на телевизионных изображениях // Техника средств связи: сер. Техника телевидения, 1984, вып. 1, С. 26–36.
202. Чочиа П.А. Язык управления заданием системы обработки изображений // Проблемы создания систем обработки, анализа и понимания изображений. Тез. докл. конференции. – Ташкент, 1991, С. 26.
203. Чочиа П.А., Милюкова О.П. Сравнение двумерных вариаций в контексте оценки сложности дискретных изображений. // Информационные процессы, 2015, Т. 15, № 2, С. 169–182.
204. Шашлов Б.А. Цвет и цветовоспроизведение. М.: Мир книги, 1995.
205. Эндрюс Г. Применение вычислительных машин для обработки изображений. М.: Энергия, 1977.
206. Яне Б. Цифровая обработка изображений. – М.: Техносфера, 2007.
207. Яншин В.В. Анализ и обработка изображений: принципы и алгоритмы. М.: Машиностроение, 1994.
208. Ярославский Л.П. Введение в цифровую обработку изображений. М.: Сов. радио, 1979.
209. Ярославский Л.П. Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. – М.: Радио и связь, 1987.
210. Ackermann F. Digital image correlation: Performance and potential application in photogrammetry // Photogrammetric Record, 1984, No. 11(64), pp. 429-439.

211. Adams C.R., Clarkson J.A. On definitions of bounded variation for functions of two variables, *Trans, of Am. Math. Soc.* 36 (1933), pp. 824.
212. Ahuja N., Rosenfeld A. Mosaic Models for Textures // *IEEE trans. PAMI-3*, 1981, vol. 3, no. 1, pp. 1–20.
213. Astola J., Heinonen P., Neuvo Y. On Roots Structures of Median and Median-Type Filters // *IEEE Trans.*, vol. ASSP–35, 1987, no. 8, pp. 1199–1201.
214. Attneave F. Some informational aspects of visual perception // *Psychol. Rev.*, 1954, vol. 61, pp. 183–193.
215. Bockstein I.M. Color equalization — a perspective method of color image processing // *Acta Politechnica Scandinavica // Applied Physics Ser.*, 1985, № 149, pp. 132–135.
216. Bockstein I.M., Chochia P.A., Kronrod M.A. Interactive Processing of Venus Images // *Lunar and Planetary Science XVII Conf. – Houston: Lunar and Planetary Institute*, 1986, pp. 60–61.
217. Bockshtein I., Chochia P., Kronrod M. Methods of Venus Northern Area Radiolocation Map Synthesis Using Strip Images of «Venera–15» and «Venera–16» Space Stations // *Lunar and Planetary Science XIX Conf. – Houston: Lunar and Planetary Institute*, 1988, pp.108–109.
218. Bockstein I., Chochia P., Kronrod M. Methods of Venus Radiolocation Map Synthesys Using Strip Map of «Venera–15» and «Venera–16» Space Stations // *Earth, Moon, and Planets*, 1988, vol. 43, pp. 233–259.
219. Bovik A.C. *Handbook of image and video processing*. Academic Press, Orlando, FL, 2005.
220. Bovik A.C., Clark M., Geisler W.S. Multichannel Texture Analysis Using Localized Spatial Filters // *IEEE Trans. PAMI*, 1990, vol. 12, no. 1, pp. 55–73.
221. Brodatz P. *Textures: A Photographic Album for Artists and Designers*. New York: Dover Publications, 1966.
222. Bumbaca F., Smith K.C. Design and Implementation of a Colour Vision Model for Computer Vision Applications // *CVGIP*, 1987, vol. 39, no.2, pp. 226–245.
223. Canny J.A Computational Approach for Edge Detection // *IEEE Trans. Pattern Analysis Machine Intelligence*, 1986, vol. 8, no. 6. pp. 679–698.
224. Carton E.J., Weszka J.S., Rosenfeld A. Some Basic Texture Analysis Techniques, TR-288. Computer vision laboratory, Computer Science Center. Univ. of Maryland, 1974.
225. Chang T., Kuo C.J. Texture Analysis and Classification with Three-Structured Wavelet Transform // *IEEE Trans. Image Processing*. 1993, vol. 2, no. 4. pp. 429–441.
226. Chochia P.A. Analysis of Video Data Formed by the Capillarscope and Blood Flow Dynamics Measurements. // *Journal of Communications Technology and Electronics*, 2014, vol. 59, no. 12, pp. 1524–1529.

227. Chochia P.A. Application of Image Frequency Filtering to Elimination of the Noise Caused by the Embossing of the Photographic Paper // *Journal of Communications Technology and Electronics*, 2011, vol. 56, no. 12, pp. 1518–1521.
228. Chochia P.A. A Pyramidal Image Segmentation Algorithm // *Journal of Communications Technology and Electronics*, 2010, vol. 55, no. 12, pp. 1550–1560.
229. Chochia P.A. Automatic gray scale correction of video data. // 5'th International Workshop on Image Processing and Computer Optics (DIP-94). // *Proc. SPIE*, 1995, vol. 2363, pp. 82-88.
230. Chochia P.A. Automatic Processing and Analysis of Video Data Formed by a Capillaroscope. // *Pattern Recognition and Image Analysis*, 2015, vol. 25, no. 4. pp. 649-662. (в печати)
231. Chochia P.A. Control task language of the image processing system for personal computer // *Proceedings Third International Seminar on Digital Image Processing*. – Riga, 1992, pp. 61–64.
232. Chochia P.A. Detection of Capillaries in the Images Formed by a Capillaroscope // *Journal of Communications Technology and Electronics*, 2013, vol. 58, no. 12, pp. 1314–1323.
233. Chochia P.A. Determination of Parameters of Capillary Blood Flow Using Video Data Analysis. // *Biomedical Engineering*, 2015, vol. 49, no.1, pp. 19-23.
234. Chochia P.A. Fast Correlative Matching of Quasi-Regular Images // *Journal of Communications Technology and Electronics*, 2010, vol. 55, no. 12, pp. 1482–1484.
235. Chochia P.A. Image Decomposition and Enhancement Using Rank Filtering // *Computer and Holographic Optics and Image Processing. Proceedings SPIE*, 1998, vol. 3348, pp. 261–266.
236. Chochia P.A. Image Enhancement Using Sliding Histograms // *Computer Vision Graphics Image Processing*, 1988, vol. 44, no. 2, pp. 211-229.
237. Chochia P.A. Image Segmentation Based on the Analysis of Distances in an Attribute Space. // *Optoelectronics, Instrumentation and Data Processing*, 2014, vol. 50, no. 6, pp. 1–12.
238. Chochia P.A. Image Segmentation via Contour Tracking in Application to the Analysis of the Photographs of Electronic Microcircuits // *Journal of Communications Technology and Electronics*, 2010, vol. 55, no. 12, pp. 1466–1473.
239. Chochia P.A. Methods of enhancing aerospace images using fragment histograms // *Soviet journal of remote sensing*. 1989, vol. 5, no. 6, pp. 1103-1123.
240. Chochia P.A. Segmentation of chip microimages by contour tracking // 19th International Conference on Computer Graphics and Vision (GraphiCon 2009). Conference proceedings. – Moscow, 2009, pp. 309–310.
241. Chochia P.A. Some methods for automatic gamma–correction of video data // 5th International Workshop «Image Processing and Computer Optics». – Samara, 1994, pp. 6–7.

242. Chochia P.A. Transition from 2D- to 3D-images: modification of two-scale image model and image processing algorithms. // *Journal of Communications Technology and Electronics*, 2015, vol. 60, no. 6, pp. 678–687.
243. Chochia P.A. Two Tasks in Image Enhancement Technology // *Optical Memory and Neural Networks*, 1998, vol. 7, no. 1, pp. 37–50.
244. Chochia P.A., Milukova O.P. Comparison of Two-Dimensional Variations in the Context of the Digital Image Complexity Assessment. // *Journal of Communications Technology and Electronics*, 2015, vol. 60, no. 12, pp. 1432–1440.
245. Chochia P.A., Milukova O.P. Two-Dimensional Variation and Image Decomposition. // 6th International Workshop «Image Processing and Computer Optics (DIP-97)» // *Proceedings SPIE*, 1998, vol. 3346, pp. 329-339.
246. Clark J.J. Authenticating Edges Produced by Zero-Crossing Algorithms // *IEEE Trans. Pattern Analysis Machine Intelligence*, 1989, vol. 12, no. 8, pp. 830–831.
247. Cornsweet T.N. *Visual Perception*. N.Y.: Academic Press, 1970.
248. Cyganek B. *An Introduction to 3D Computer Vision Techniques and Algorithms*. NY.: Wiley, 2009.
249. Danielsson P.-E. Getting the median faster // *Computer Graphics Image Processing*, 1981, vol. 17, no. 1, pp. 71–78.
250. Davis L.S., Rosenfeld A. Noise cleaning by iterative local averaging // *IEEE Trans. System Man Cybernetics*, 1978, vol. SMC-8, pp. 705-710.
251. Digital processing of endoscopic color images / Ohyama N., Suzuki K., Honda T. et.al. // *Optics Communications*, 1985, vol. 55, no. 4, pp. 242-247.
252. Digital synthesis of diagnostic X-ray images / Rabkin I.Kh., Gusev E.A., Petushkov A.A., Sosnin F.R., Chochia P.A. et al. // *Biomedical Engineering*, 1984, vol. 18, no. 1, pp. 7-9.
253. Digital synthesis of images in the radiography of objects with complex configuration / Gusev E.A., Kronrod M.A., Lebedev D.S., Petushkov A.A., Sosnin F.R., Chochia P.A. // *Soviet journal of nondestructive testing*, 1982, vol. 18, no. 7, pp. 571–575.
254. Dunn D., Higgins W.E. Optimal Gabor Filters For Texture Segmentation // *IEEE Trans. Image Processing*. 1995, vol. 4, no. 7, pp. 947–964.
255. Dynamic radiographic monitoring with a grid and digital image processing / Veselovskii L.N., Gusev E.A., Petushkov A.A., Sosnin F.R., Ruzinova E.V., Kaplun Y.M., Chochia P.A., Sanpiter I.A. // *Soviet journal of nondestructive testing*, 1984, vol. 20, no. 4, pp. 246-249.
256. Faugeras O.D. Digital Color Image Processing Within the Framework of a Human Visual Model // *IEEE Trans. ASSP-27*, 1979, vol. 27, no. 4, pp. 380–393.

257. Frei W. Image Enhancement by Histogram Hyperbolisation // Computer Graphics Image Processing, 1977, vol. 6. no. 3. pp. 286–294.
258. Frieden B.R. Statistical Models for the Image Restoration Problems // CGIP, 1980, vol. 12, no. 1, pp. 40–59.
259. Fu K.S., Mui J.K. A Survey of Image Segmentation // Pattern Recognition, 1981, vol. 13, no. 1. pp. 3–16.
260. Fu-Nian Ku The Principles and Methods of Histogram Modification Adapted for Visual Perception // Computer Vision, Graphics, and Image Processing, 1984, vol.26, no.1, pp. 107-117.
261. Geology and morphology of the northern hemisphere of Venus. / Barsukov V.L., Bazilevskij A.T., Bogomolov A.F., Chochia P.A. et al. // Geotektonics. 1986. vol. 20, no. 4, pp. 256-270.
262. Gruen A.W. Adaptive least squares correlation: a powerful image matching technique // S. Afr. J. of Photogrammetry, Remote Sensing and Cartography, 1985, No. 14(3), pp. 175-187.
263. Grenander U., Miller M.I. Pattern Theory: From Representation to Inference. Oxford: University Press, 2007. 596 p.
264. Hall E.L. Almost Uniform Distribution for Computer Image enhancement // IEEE Trans. Computers, 1974, vol. C-23, no. 2, pp. 207-208.
265. Hall C.F., Hall E.L. A Nonlinear Model for the Spatial Characteristics of the Human Visual System // IEEE Trans. SMA–7, 1977. № 3. pp. 161–170.
266. Hall E.L. Computer Image Processing and Recognition. Academic Press, New York: 1979.
267. Haralick R.M. Edge and Region Analysis for Digital Image Data // CGIP, 1980, vol. 12, no. 1, pp. 60–73.
268. Haralick R.M. Image Texture Survey // Fundamentals in Computer Vision. – Cambridge: CUP, 1983, pp. 145–172.
269. Haralick R.M. Statistical and Structural Approaches to Textures // Proc. IEEE. 1979, vol. 67, no. 5. pp. 786–804.
270. Haralick R.M., Shapiro L.G. Computer and Robot Vision, vols. 1, 2. Reading, MA: Addison-Wesley, 1992.
271. Haralick R.M., Shapiro L.G. Image Segmentation Techniques // Computer Vision, Graphics, Image Processing, 1985, vol. 29, no. 2. pp. 100–132.
272. Haralick R.M., Watson L. A Facet Model for Image Data // Computer Graphics and Image Processing, 1981, vol. 15. no. 2. pp. 113–129.
273. Hotelling H. Analysis of a Complex of Statistical Variables into Principal Components // J. Educ. Physics, 1933, vol. 24, pp. 417–441, 498–520.
274. Huang T.S. Image Enhancement: A Review // Opto-Electronic, 1969, vol. 1, pp. 49-59.

275. Huang T.S., Yang G.J., Tang G.Y. A fast two-dimensional median filtering algorithm // IEEE Trans. on Acoustics, Speech, Signal Processing, 1979, vol. ASSP-27, no. 1, pp.13–18.
276. Hummel R.A. Histogram Modification Techniques // Computer Graphics Image Processing, 1975. vol. 4, no. 3. pp. 209—224.
277. Hummel R.A. Image Enhancement by Histogram Transformation // Computer Graphics Image Processing, 1977. vol. 6, no. 3. pp. 184–195.
278. Hunt B.R. Nonstationary Statistical Image Models (and Their Application to Image Data Compression) // CGIP, 1980, vol. 12, no. 2, pp. 173–186.
279. Hunter G.M., Steiglitz K. Operation of Images Using Quad Trees // IEEE Trans. PAMI-1, 1979, no. 2, pp. 145–153.
280. Image Analysis and Mathematical Morphology. Vol. 2: Theoretical Advances / Ed. J. Serra. NY.: Academic Press, 1988.
281. Image quality assessment: From error visibility to structural similarity / Wang Z., Bovik A.C., Sheikh H.R., Simoncelli E.P. // IEEE Transactions on Image Processing, 2004, vol. 13, no. 4, pp. 600–612.
282. Jagadish H.J., O'Gorman L. An Object Model for Image Recognition // Computer, 1989, vol. 22, no. 12. pp. 33–41.
283. Jain A.K. Color Distance and Geodesics in Color 3 Space // JOSA. 1972, vol. 62, no. 11, pp. 1287–1291.
284. Jain A.K. Fundamentals of Digital Image Processing. Prentice Hall, N.J., 1989.
285. Jain A.K. Image Coding via a nearest neighbors image model // IEEE Trans. COM-23, 1975, no. 3, pp. 318–331.
286. Jain A.K., Dubes R.C. Algorithms for Clustering Data. Prentice Hall, 1988.
287. Jain A.K., Farrokhnia F. Unsupervised Texture Segmentation Using Gabor Filters // Pattern Recognition. 1991, vol. 24, no. 12. pp. 1167–1186.
288. Jain R., Kasturi R., Schunk B. Machine Vision. New York: McGraw-Hill, 1995.
289. Jain R., Rangachar K., Schunk B. Computer Vision. New York: McGraw-Hill, 1995.
290. Jayanthi A.K., Sujatha N., Ramasubba Reddy M. Measuring Blood flow: techniques and applications – A review. // International journal of research and review in applied sciences, 2011, vol. 6, pp. 203-216.
291. Jobson D.J., Zia-ur Rahman, Woodell G.A. A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Scenes // IEEE Trans. Image Processing, 1997, vol. 6, no. 7, pp. 965–976.

292. Justusson B.I. Median filtering: statistical properties. // Topics in Applied Physics, vol. 43, «Two-Dimensional Signal Processing. Transforms and Median Filters» / ed. Huang T.S. – Berlin: Springer-Verlag, 1981.
293. Kashyap R.L., Chellapa R., Ahuja N. Decision Rules of Choice for Neighbors in Random Field Models of Images // CGIP, 1981, vol. 15, no. 4, pp. 301–318.
294. Keelan B.W. Handbook of Image Quality / New York: Marcel Dekker, 2002. 516 p.
295. Kim V., Yaroslavskii L.P. Rank algorithms for picture processing // Computer Graphics Image Processing, 1986, vol. 35, no. 3. pp. 234-258.
296. Lee J.-S. Digital Image Enhancement and Noise Filtering by Use of Local Statistics // IEEE Trans. PAMI-2, 1980, no. 2, pp. 165–168.
297. Lee J.-S. Digital Image Smoothing and the Sigma Filter // Computer Vision, Graphics, Image Processing, 1983, vol. 24, no. 2. pp. 255–269.
298. Lebedev D.S. Probabilistic characterization of images in filtration and restoration problems. // Signal processing: Theories and applications. – Amsterdam: EURASIP, 1980, p. 55–64.
299. Lebedev D.S, Ip H.H.-S, Potter D.J. Impulse noise cleaning by iterative threshold median filtering // Pattern Recognition Letters, 1983, vol. 2, pp. 89–93.
300. Li S.Z. Markov Random Field Modeling in Image Analysis. Springer, 2009.
301. Limb J.O., Murphy H.A. Measuring the Speed of Moving Objects from Television Signals. IEEE Trans. Communication COM-23, 1975, no. 4, pp. 474–478.
302. Lo L.-Ch., Chiang J.Y., Cai Yu-S. Three-dimensional Vision-based Nail-fold Morphological and Hemodynamic Analysis // 11th IEEE Int. Conf. on Bioinformatics and Bioengineering BIBE-2011. – Taichung, Taiwan, pp. 44-51.
303. Lohmann G., Volumetric Image Analysis. Teubner, Chichester: Wiley, 1998.
304. LUPA-300 CMOS Image Sensor. URL:
http://www.fastvideo.ru/info/sensor/cypress/lupa_300.pdf .
305. LUPA-300 Frequently Asked Questions – AN6019. URL:
<http://www.cypress.com/file/123371/download> .
306. MacAdam D.L. Projective Transformations of the ICI Color Specifications // JOSA. 1935, vol. 27, no. 9, pp. 294–299.
307. Mahabala H. Preprocessor for programs which recognize scenes. // MIT Artificial Intelligence Lab., Mem. 177, 1969.
308. Marr D., Hildreth E. Theory of Edge Detection // Proc. Royal Soc. London. Ser. B, 1980, vol. 207, pp. 187–217.

309. Mastin G.A. Adaptive Filters for Digital Image Noise Smoothing: An Evaluation // *Computer Vision, Graphics, Image Processing*, 1985, vol. 31. no. 1. pp. 103–121.
310. Matas J., Kittler J., Spatial and Feature Space Clustering: Applications in Image Analysis // 6th Int. Conf. on Computer Analysis and Patterns. – Prague, Czech Republic, 1995.
311. Method and apparatus for measuring color–texture distance, and method and apparatus for sectioning image into plurality of regions using measured color–texture distance / Kim S.-K., Lee S.-D., Kim C.-Y., Chochia P., Sushko D., Miller B. // Patent 01308622.8–2217. European Patent Office, 2001.
312. Method and apparatus for measuring color-texture distance, and method and apparatus for sectioning image into plurality of regions using measured color-texture distance / Kim S.-K., Lee S.-D., Kim C.-Y., Chochia P., Sushko D., Miller B. // Patent US20020090133 A1 United States Patent Application, July 11, 2002.
313. Milyukova O.P. On Justification of Image Model. *Proc. SPIE*, 1997, vol. 3348, pp. 283-289.
314. Milyukova O.P., Chochia P.A., On Estimation of the Image Complexity by Two-Dimensional Variations // *Journal of Communications Technology and Electronics*, 2013, vol. 58, no. 6, pp. 628–635.
315. Mitra S.K., Tian-Hu Yu Transform Amplitude Sharpening: A New Method of Image Enhancement // *Computer Vision, Graphics, Image Processing*, 1987, vol. 40, no. 2, pp. 205-218.
316. Modeling and Segmentation of Speckled Images Using Complex Data / Derin H., Kelly P., Vezina G., Labitt S.G. // *IEEE Trans. Geoscience and Remote Sensing*. 1990, vol. 28, no. 1, pp. 76–87.
317. Monga O., Deriche R., Rocchisani J. 3D edge detection using recursive filtering: application to scanner images // *CVGIP: Image Understanding*, 1991, vol. 53, no. 1, pp. 76–87.
318. Mottl V.V., Kopylov A.V. Algorithms for matching images with raster distortions. // *Pattern Recognition and Image Analysis. Advances in Mathematical Theory and Applications*, 1996, Vol. 6, № 4. P. 697–703.
319. Munteanu C., Rosa A. Gray-scale image enhancement as an automatic process driven by evolution // *IEEE Trans. on Systems, Man and Cybernetics. Part B: Cybernetics*. 2004, vol. 34. no. 2, pp. 1292–1298.
320. Nagao M., Matsuyama T. Edge Preserving Smoothing // *Computer Graphics Image Processing*, 1979, vol. 9, no. 4, pp. 394–407.
321. Niblack W. *An Introduction to Digital Image Processing*. Englewood Cliffs: Prentice Hall, 1986.
322. Nikolaidis N., Pitas I. *3-D Image Processing Algorithms*. NY.: Wiley, 2000.
323. Nishikawa S., Massa R.J., Mott-Smith J.C. Area Properties of Television Pictures // *IEEE Trans. IT–11*, 1965, no. 3, pp. 348–352.

324. Ohta Y., Kanade T., Sakai T. Color Information for Region Segmentation // *Computer Graphics, Image Processing*. 1980, vol. 13, no. 3, pp. 224–241.
325. Pal N.K., Pal S.K. A Review on Image Segmentation Techniques // *Pattern Recognition*. 1993, vol. 26, no. 9. pp. 1277–1293.
326. Pavlidis T. *Algorithms for Graphics and Image Processing*. – Computer Science Press, Rockville, 1982.
327. Parzen E. On the estimation of a probability density function and mode // *Annals of Mathematical Statistics*, 1962, vol. 33, pp. 1065–1076.
328. Perona P., Malik J. Scale-Space and Edge Detection Using Anisotropic Diffusion // *IEEE Trans. PAMI*, 1990, vol. 12, no. 7, pp. 629-639.
329. Peters R.A., Strickland R.N. Image Complexity Metrics for Automatic Target Recognizers // *Automatic Target Recognizer System and Technology Conference*, 1990. – Silver Spring, MD, pp. 30-31.
330. Petrou M., Bosdogianni P. *Image Processing: The Fundamentals*. John Wiley & Sons, UK, 1999.
331. Pichler O., Teuner A., Hosticka B.J. A Comparison Of Texture Feature Extraction Using Adaptive Gabor Filtering Pyramidal And Tree Structured Wavelet Transforms // *Pattern Recognition*, 1996, vol. 29, no. 5, pp. 733–742.
332. Pierpont J. *Lectures on the theory of functions of real variables*. vol.1. N.Y.: Dover Publications, 1959.
333. Plataniotis K.N., Venetsanopoulos A.N. *Color image processing and applications*. Berlin–Heidelberg: Springer, 2000.
334. Pomalaza-Raez C.A. McGillem C.D. An adaptive, nonlinear edge-preserving filter // *IEEE Trans. Acoustic, Speech, Signal processing*, 1984, vol. ASSP-32, no. 3, pp. 571–576.
335. Prewitt J.M.S. *Object Enhancement and Extraction*. // *Picture Processing and Psychopictorics*. – New York: Academic Press, 1970, pp. 75–150.
336. *Proceedings of symposium on image modeling*. Rosemont, Illinois, 1979 // *Computer Graphics Image Processing*, 1980, vol. 12 (Special issue on image models.), no. 1-4, 426 p.
337. *Processing of Mars Surface Images Received from Phobos–2 Space Station* / Bockstein I.M., Chochia P.A., Kronrod M.A., Gektin Yu.M. // *Lunar and Planetary Science XXI Conference*. Part 1. – Houston.: Lunar and Planetary Institute, 1990, pp. 101–102.
338. Radiographic monitoring with image-processing by linear filtration / Gusev E.A., Petushkov A.A., Sosnin F.R., Chochia P.A. // *Soviet journal of nondestructive testing*, 1984, vol. 20, no. 3, pp. 183–185.

339. Roan S.J., Aggarwal J.K. Multiple Resolution Imagery and Texture Analysis // *Pattern Recognition*, 1987, vol. 20, no. 1. pp. 17–31.
340. Roberts J.M. Attentive Visual Tracking and Trajectory Estimation for Dynamic Scene Segmentation. Ph.D thesis., University of Southampton, 1994.
341. Roberts L.G. Machine Perception of Three-Dimensional Solids // *Optical and Electro-Optical Information Processing*. – MIT Press, Cambridge, Mass, 1965.
342. Rosenfeld A. Quadrees and Pyramids for Pattern Recognition and Image Analysis // *Proc. of the 5th Int. Conference on Pattern Recognition*. – Miami Beach, 1980. pp. 802–811.
343. Rosenfeld A., Kak A.C., *Digital Picture Processing*, vol. 1-2. N.-Y.: Academic press, 1982.
344. Rudin L.I., Osher S., Fatemi E. Nonlinear total variation based noise removal algorithms. // *Physica D: Nonlinear Phenomena*, 1992, vol. 60, pp.259-268.
345. Sahoo P.K., Soltani S., Wong A.K.C., Chan Y.C. A Survey of Thresholding Techniques // *Computer Vision, Graphics, Image Processing*, 1988, vol. 4, pp. 233–260.
346. Scollar I., Wang B., Huang T.S. Image Enhancement Using the Median and the Interquartile Distance // *Computer Vision, Graphics, Image Processing*, 1984, vol. 25, no. 2, pp. 236–251.
347. Shapiro L.G., Stockman G.C. *Computer Vision*. Upper Saddle River, N.J.: Prentice Hall, 2001.
348. Sheikh H.R., Bovik A.C., De Veciana G. An Information Fidelity Criterion for Image Quality Assessment Using Natural Scene Statistics // *IEEE Trans. Image Proc.*, 2005, pp. 2117–2128.
349. Sheikh H.R., Sabir M.F., Bovik F.C. A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms // *IEEE Trans. Image Proc.*, 2006, vol. 15, no. 11. pp. 3440-3451.
350. Smith P., Drummond T., Cipolla R. Edge tracking for motion segmentation and depth ordering // *Proc. 10th British Machine Vision Conf.*, V. 2. – Nottingham, 1999, pp. 369–378.
351. Sobel I.E. Camera Models and Machine Perception. Ph.D. dissertation. – Stanford University, Palo Alto, Calif., 1970.
352. Special issue on Mathematical Morphology and Nonlinear Image Processing // *Pattern Recognition*, 2000, vol. 33, no. 6. pp. 875–1117.
353. Szeliski R. *Computer Vision: Algorithms and Applications*. New York: Springer, 2010.
354. The Geology and Geomorphology of the Venus Surface as Revealed by the Radar Images Obtained by «Venera 15 and 16» / Barsukov V.L., Basilevsky A.T., Bogomolov A.F., Chochia P.A. et al. // *Journal of Geophysical Research*, 1986, vol. 91, no. B4, pp. 378–398.
355. Tomita F., Tsuji S. Extraction of Multiple Regions by Smoothing in Selected Neighborhoods // *IEEE Trans. SMC-7*, 1977, no. 7, pp. 107–109.
356. Tomasi C., Manduchi R. Bilateral filtering for gray and color images. *Proc. IEEE 6th Int. Conf. on Computer Vision*. – Bombay, India, Jan. 4-7, IEEE, 1998, pp. 839–846.

357. Toriwaki J., Yoshida H. *Fundamentals of Three-Dimensional Digital Image Processing*. NY.: Springer, 2009.
358. Tukey J.W. *Exploratory Data Analysis* // Reading, Mass. Addison-Wesley, 1971.
359. Tyan S.G. Median Filtering: Deterministic Properties // *Topics in Applied Physics*, vol. 43, «Two-Dimensional Signal Processing. Transforms and Median Filters» / ed. Huang T.S. – Berlin, Springer-Verlag, 1981. pp. 197–217.
360. Vainio O, Neuvo Y., Butner S.E. A Signal Processor for Median-Based Algorithms // *IEEE Trans.* vol. ASSP–37, 1989, no. 9. pp. 1406–1414.
361. Van Gool L., Dewaele P., Oosterlinck A. Texture Analysis Anno 1983 // *Computer Vision, Graphics, Image Processing*, 1985, vol. 29, pp. 336–357.
362. Wallis R. An Approach to the Space-Variant Restoration and Enhancement of Images // *Image Science Mathematics; Proc. Sympos. Current Mathematical Problems in Image Science*. Monterey, Calif., 1976. – Western Periodicals, North Hollywood, 1977, pp. 10-12.
363. Wang D., Vagnucci A., Li C. Image Enhancement by Gradient Inverse Weighted Smoothing Scheme // *Computer Graphics Image Processing*, 1981, vol. 15, no. 2, pp. 167–181.
364. Wang Z., Bovik A.C. *Modern image quality assessment*. N.Y.: Morgan and Claypool, 2006.
365. Wang Z., Simoncelli E.P., Bovik A.C. Multi-scale structural similarity for image quality assessment // *IEEE ACSSC*, San Francisco, USA, 2003, pp. 1398-1402.
366. Weldon T.P., Higgins W.E., Dunn D.F. Efficient Gabor Filter Design For Texture Segmentation // *Pattern Recognition*, 1996, vol. 29, no. 12, pp. 2005–2015.
367. Werner F. Image enhancement by histogram hyperbolization. // *Computer Graphics Image Processing*, 1977, vol. 6, no. 3, pp. 286–294.
368. Wiener N. *Extrapolation, interpolation and smoothing of stationary time series*. N.-Y., John Wiley, 1949.
369. Wirjadi O. *Survey of 3d image segmentation methods*. Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM). – Kaiserslautern (Germany), 2007.
370. Woods J.W. Two-dimensional discrete Markovian fields // *IEEE Trans. Information Theory*, 1972, vol. 18, pp. 232–240.
371. Yan J.K., Sakrison D.J. Encoding of Images Based on a Two-Component Source Model // *IEEE Trans. COM–25*, 1977, no. 11, pp. 1315–1322.
372. Yaroslavskij L.P., Pribilova R.A. Vergleich von Algorithmen für die Filtering von Punktrauchen in Bildern. // *Bild und Ton*, 1985, no. 6, pp. 177–180.
373. Yu H., Winkler S. Image complexity and spatial information. // *Proc. 5th International Workshop on Quality of Multimedia Experience (QoMEX)*, 2013. – Klagenfurt, Austria, pp. 12-17.

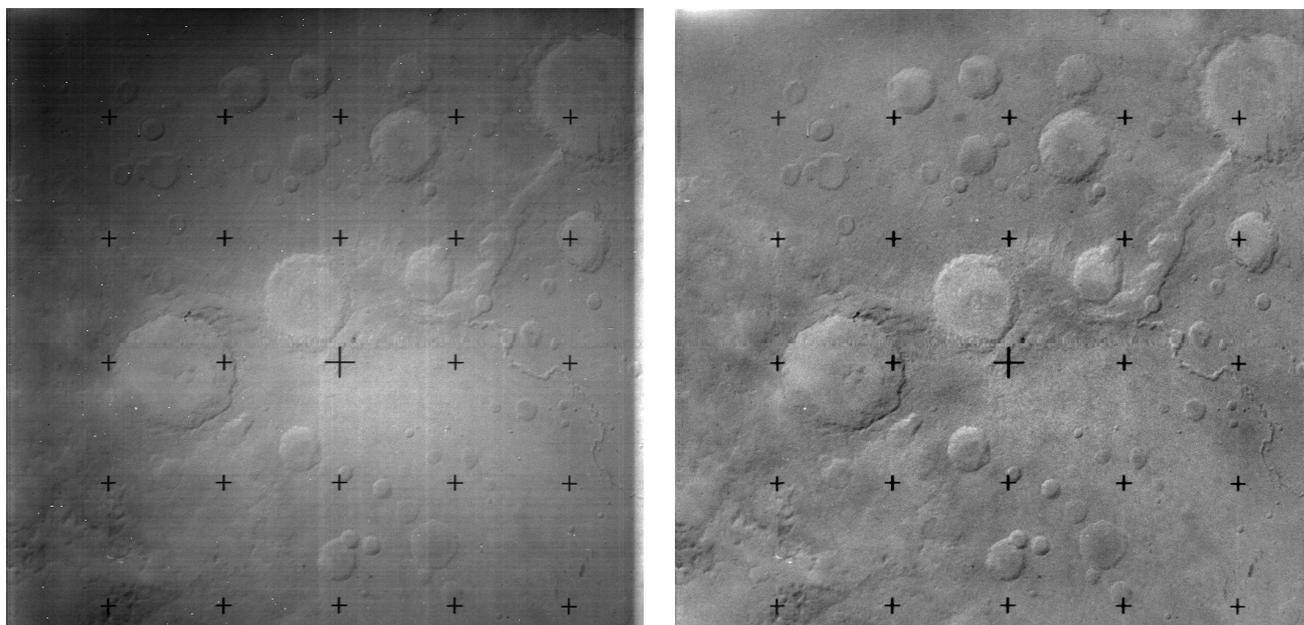
374. Zamperoni P. Model-Based Segmentation of Grey-Tone Images // *Image and Vision Comput.* 1984, vol. 2, no. 3, pp. 123–133.
375. Zhang T.Y., Suen C.Y. A Fast Parallel Algorithm for Thinning Digital Patterns. *Comm. ACM*, 1984, vol. 27, no. 3, pp. 236–239.

ПРИЛОЖЕНИЯ

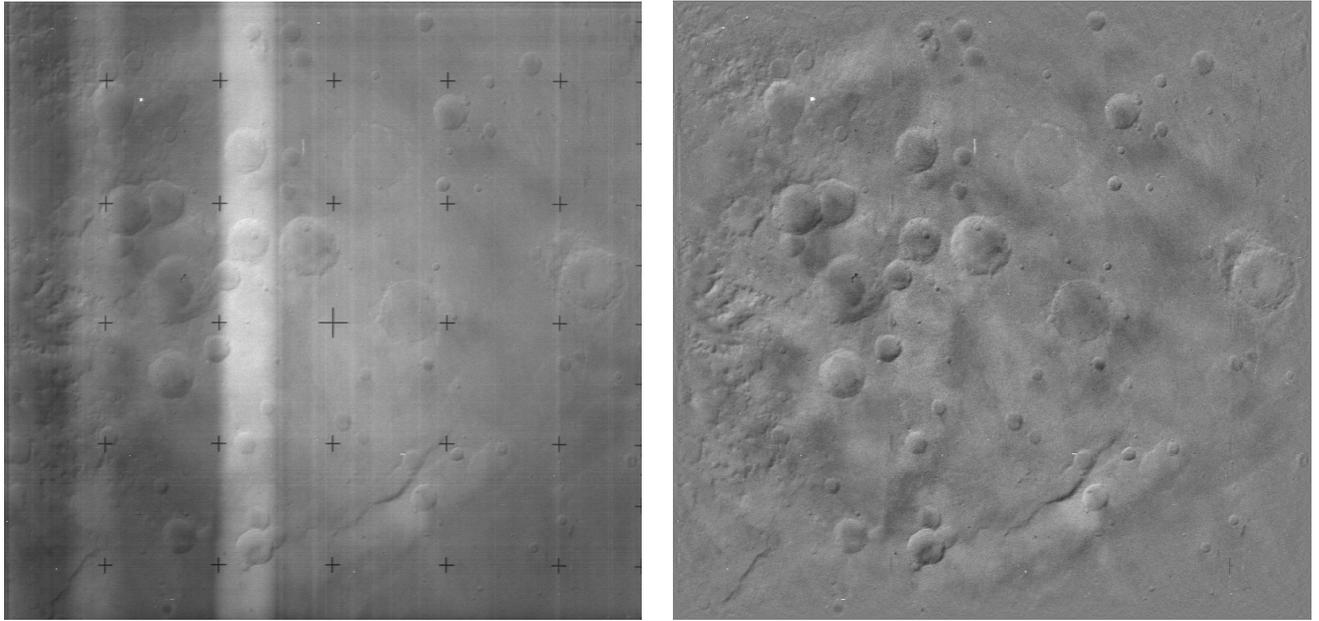
Приложение 1. Обработка снимков планет

Иллюстрируется применение методов и алгоритмов обработки видеоинформации, рассмотренных в основной части (в основном, изложенных в Главе 3), применительно к обработке снимков планет, полученных автоматическими межпланетными станциями. Преобразования осуществлялись на специализированном комплексе обработки изображений Института проблем передачи информации РАН (АН СССР).

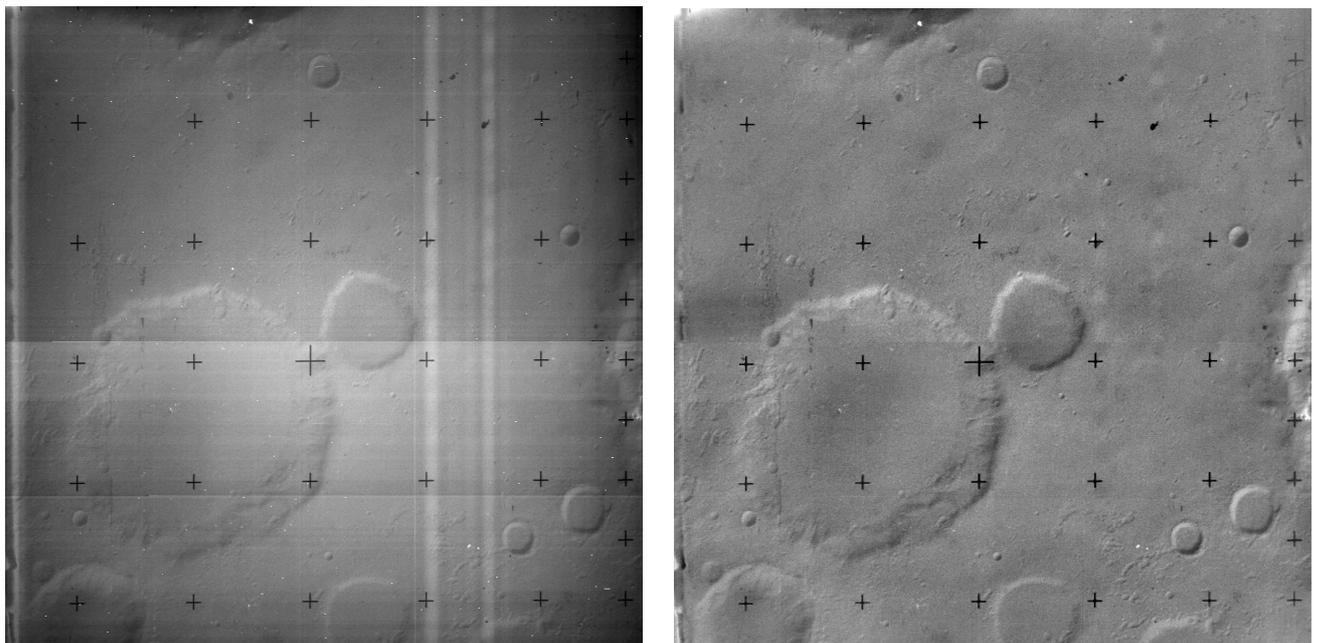
1.1. Обработка фототелевизионных снимков поверхности Марса, переданных АМС «Марс-4» и «Марс-5»



Исходный мелкомасштабный (1 : 5000000) снимок, полученный АМС «Марс-5», (слева) и результат обработки (справа).



Исходный мелкомасштабный (1 : 5000000) снимок, полученный АМС «Марс-5», (слева) и результат обработки (справа).



Исходный крупномасштабный (1 : 500000) снимок, полученный АМС «Марс-5», (слева) и результат обработки (справа).

Алгоритмы обработки снимков поверхности Марса, полученных АМС «Марс-4, -5», опубликованы в работах [150, 155].

1.2. Обработка первых панорам поверхности Венеры, полученных посадочными аппаратами АМС «Венера-9» и «Венера-10»



Результат обработки панорамы поверхности Венеры, переданной АМС «Венера-9».

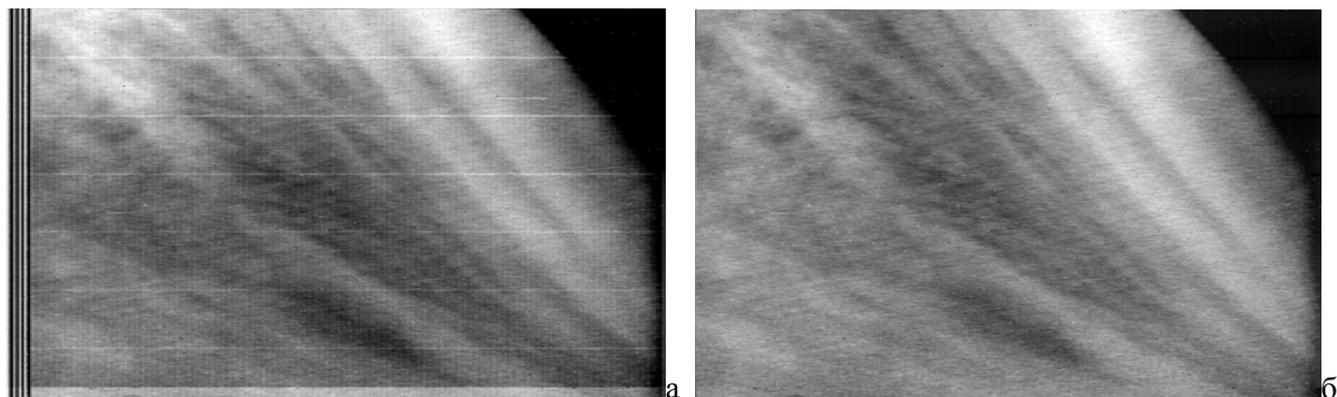
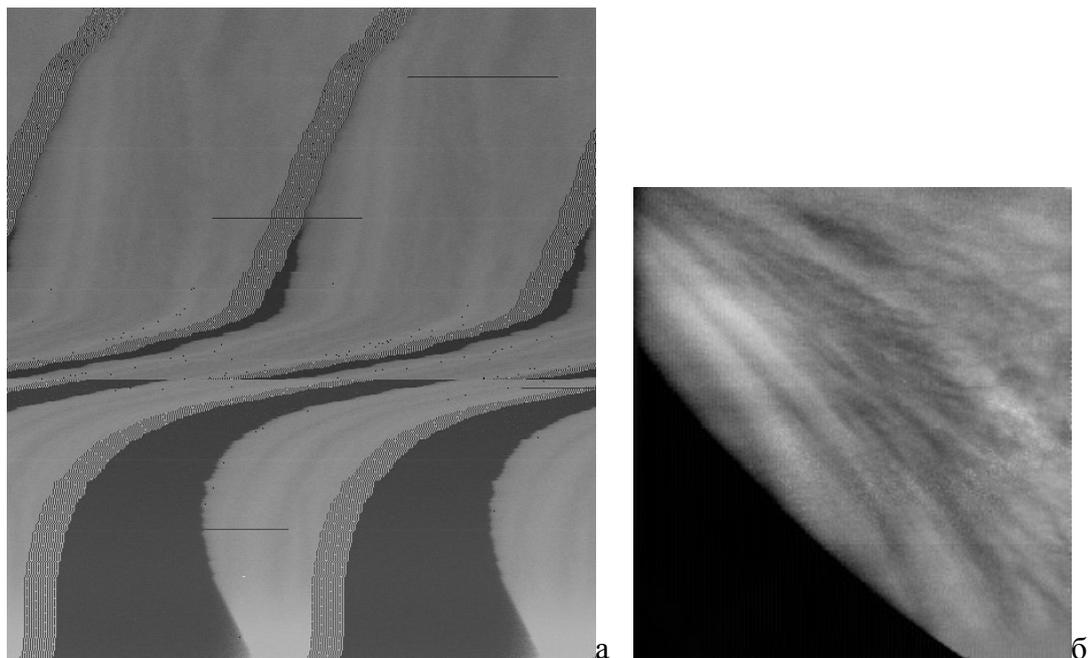


Результат обработки панорамы поверхности Венеры, переданной АМС «Венера-10».

Алгоритмы обработки телевизионных панорам поверхности Венеры, полученных АМС «Венера-9, -10», опубликованы в работе [101].

1.3. Обработка снимков облачного слоя Венеры, переданных орбитальным аппаратом станции «Венера-9»

Вид исходных панорамных снимков облачного покрова Венеры, переданных орбитальным аппаратом АМС «Венера-9» (слева), и результат обработки изображений (справа):



Алгоритм обработки панорам облачного слоя Венеры, полученных АМС «Венера-9», опубликован в работе [129].

1.4. Обработка панорам поверхности Венеры, полученных посадочными аппаратами АМС «Венера-13» и «Венера-14»



Результат обработки панорамы поверхности Венеры, переданной камерой I АМС «Венера-13».



Результат обработки панорамы поверхности Венеры, переданной камерой II АМС «Венера-13».



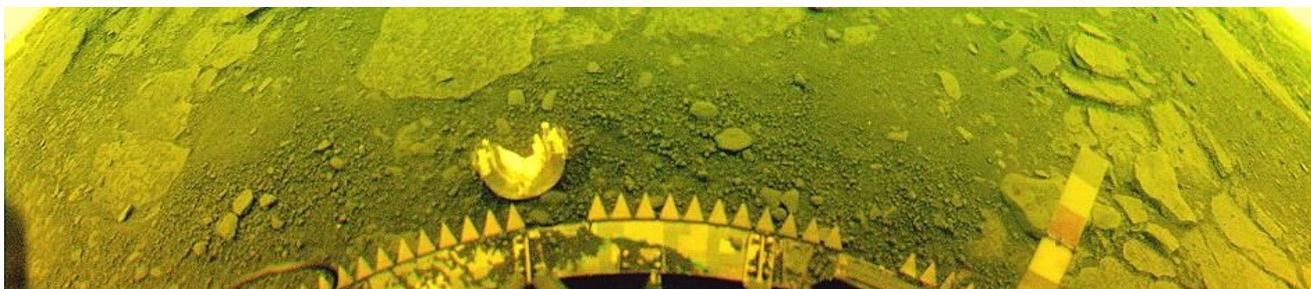
Результат обработки панорамы поверхности Венеры, переданной камерой I АМС «Венера-14».



Результат обработки панорамы поверхности Венеры, переданной камерой II АМС «Венера-14».

Алгоритмы обработки телевизионных панорам поверхности Венеры, полученных АМС «Венера-13, -14» опубликованы в работах [105, 216].

1.5. Синтез цветных снимков поверхности Венеры, полученных посадочными аппаратами АМС «Венера-13» и «Венера-14»



Синтез цветной панорамы поверхности Венеры, переданной камерой II АМС «Венера-13».



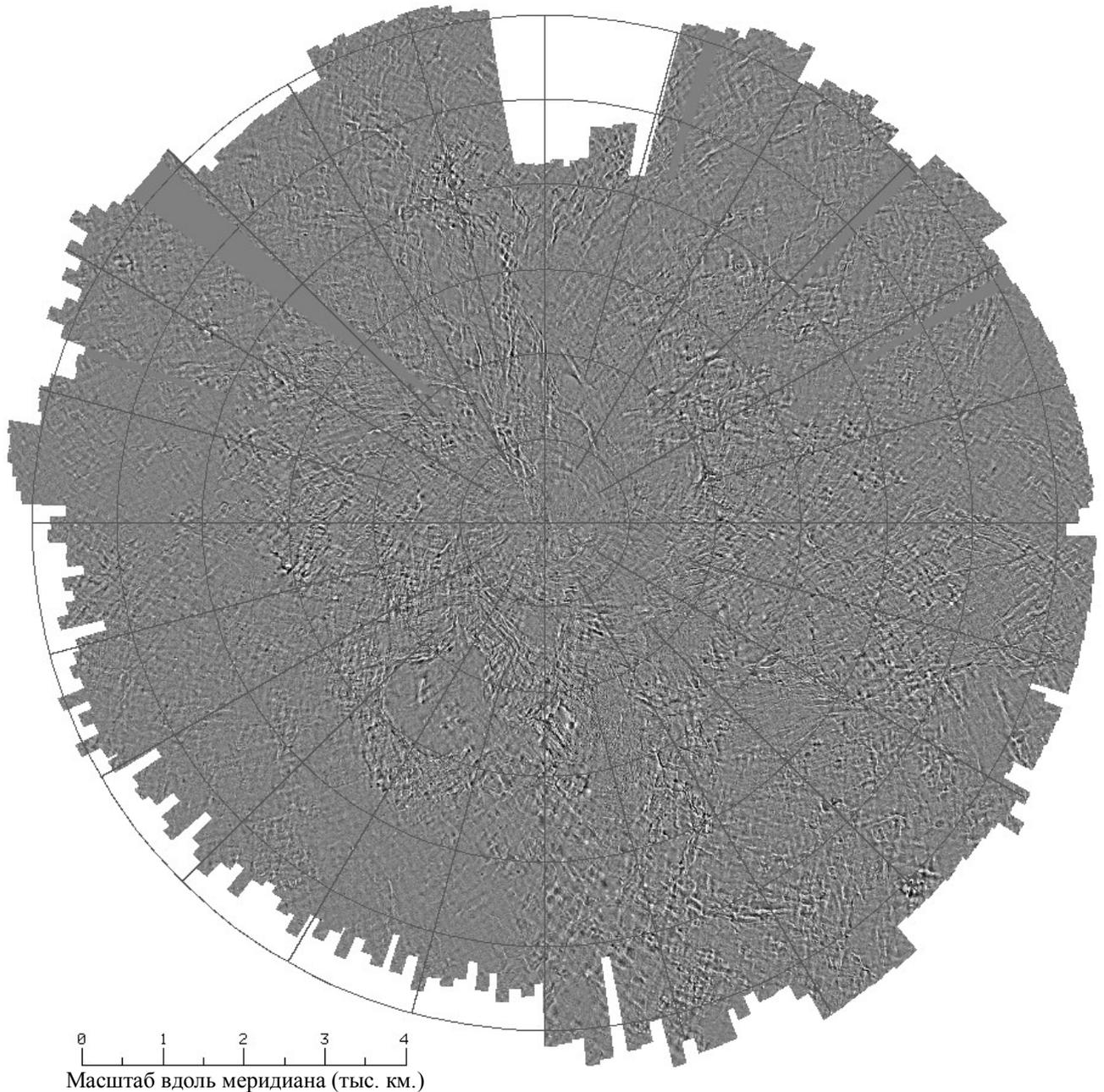
Синтез цветной панорамы поверхности Венеры, переданной камерой II АМС «Венера-14».



Синтез цветных панорам поверхности Венеры, переданных камерами I АМС «Венера-13» (слева) и «Венера-14» (справа).

Алгоритмы обработки и синтеза цветных панорам поверхности Венеры, полученных АМС «Венера-13, -14» опубликованы в работах [105, 109].

1.6. Синтез радиолокационной карты северного полушария Венеры по данным АМС «Венера-15» и «Венера-16»

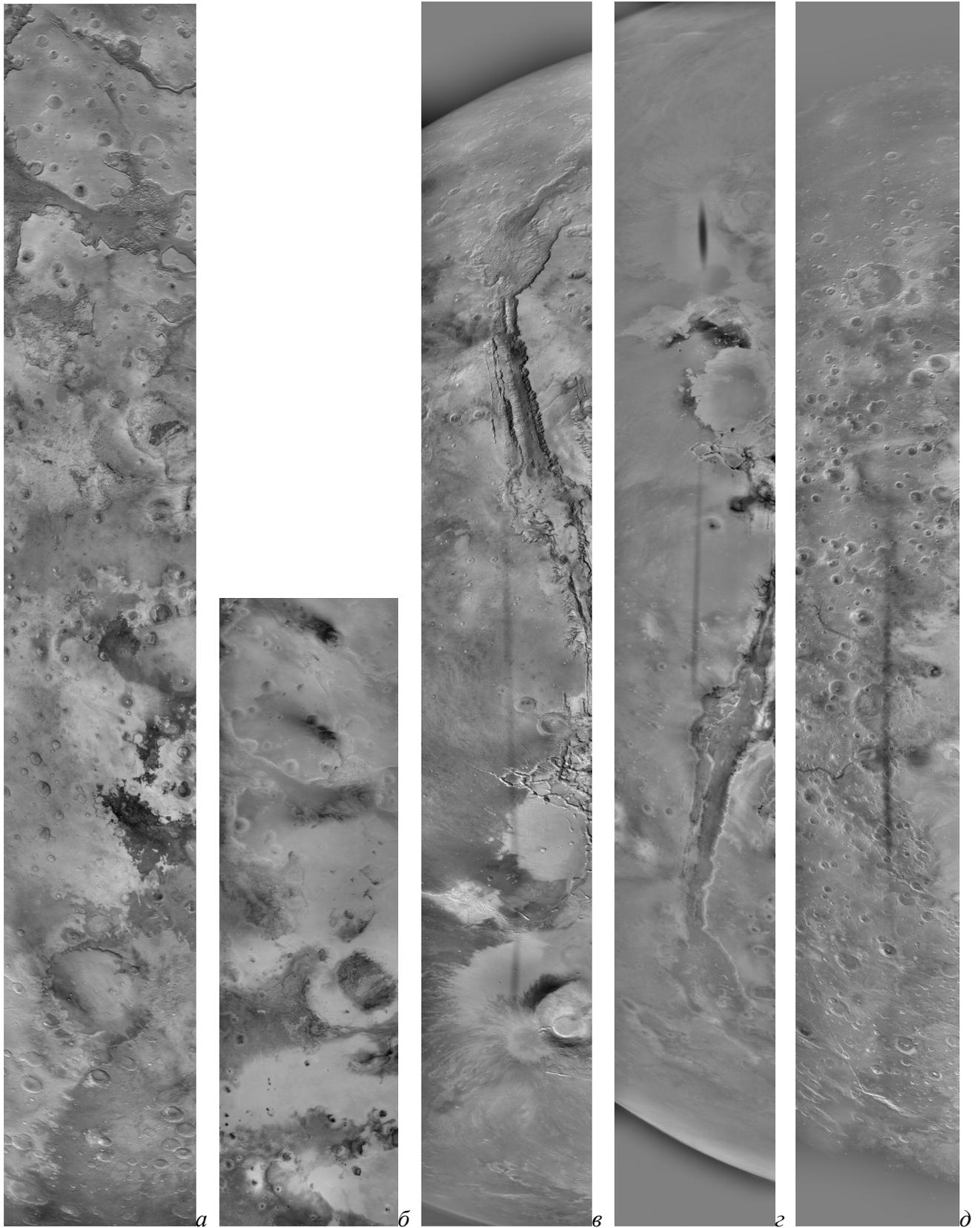


Карта северного полушария Венеры по данным, полученным радиолокаторами бокового обзора с синтезированной апертурой станций «Венера-15» и «Венера-16», разработанными ОКБ МЭИ. Обработка данных и синтез карты на системе обработки изображений ИППИ АН СССР.

Особенности полосовой радиолокационной съемки, алгоритмы обработки исходных данных, синтеза карты северного полушария Венеры и ее геологической интерпретации опубликованы в работах [15, 16, 38, 99, 217, 218, 261, 354].

1.7. Обработка панорамных снимков поверхности Марса, полученных АМС

«Фобос-2»



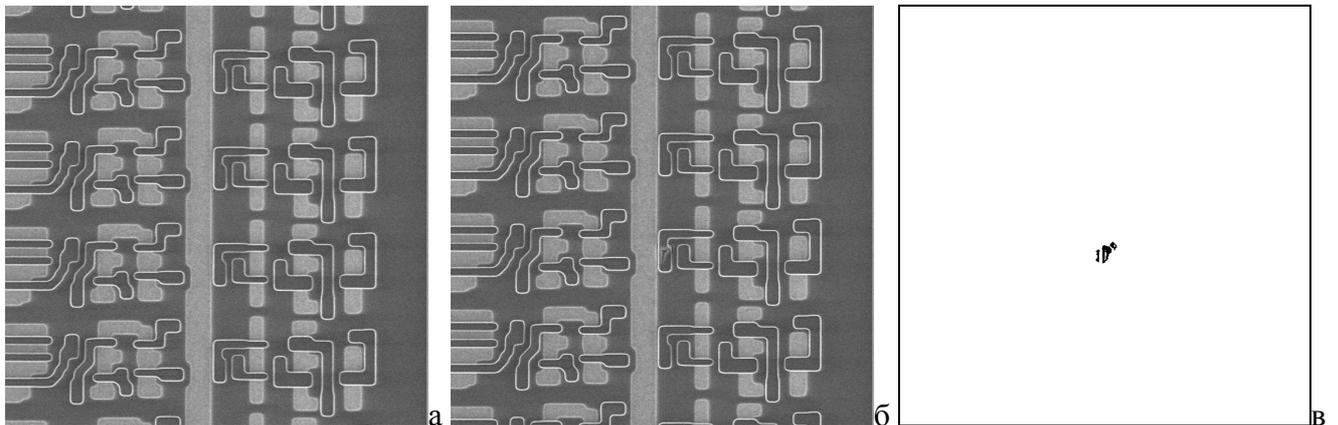
Результаты обработки панорамных снимков поверхности Марса в диапазоне 0,5-1,0 мкм. (панорамы *a* и *b*) и в диапазоне 8-13 мкм. (панорамы *v*, *z* и *d*). Алгоритмы опубликованы в [337].

Приложение 2. Обнаружение дефектов на снимках поверхностей микросхем

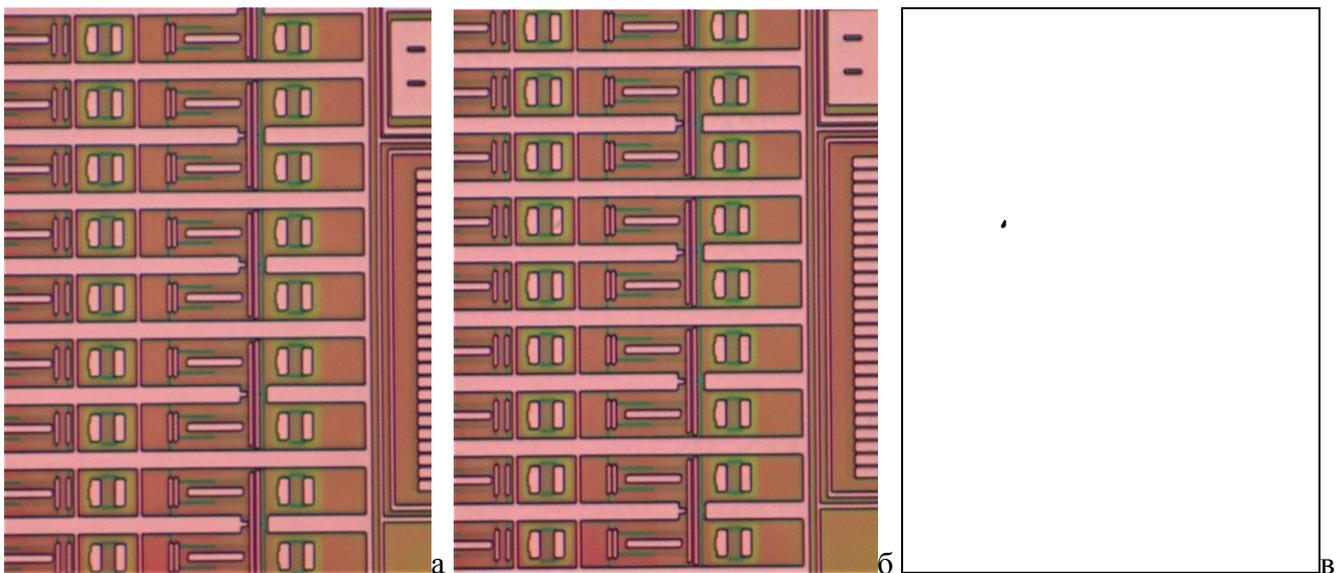
При разработке электронных микросхем требуется контроль качества формируемых пластин. Одним из способов контроля является анализ их поверхности при использовании снимков, формируемых оптическим или сканирующим электронным микроскопом. Часть задач анализа были подробно рассмотрены в основном тексте в разделах 4.5 и 5.1. Опубликованы в работах [176, 195, 238, 240].

2.1. Обнаружение дефектов на снимках, получаемых оптическим микроскопом

Постановка задачи обнаружения и алгоритм анализа для черно-белого изображения рассмотрены в разделе 4.5.



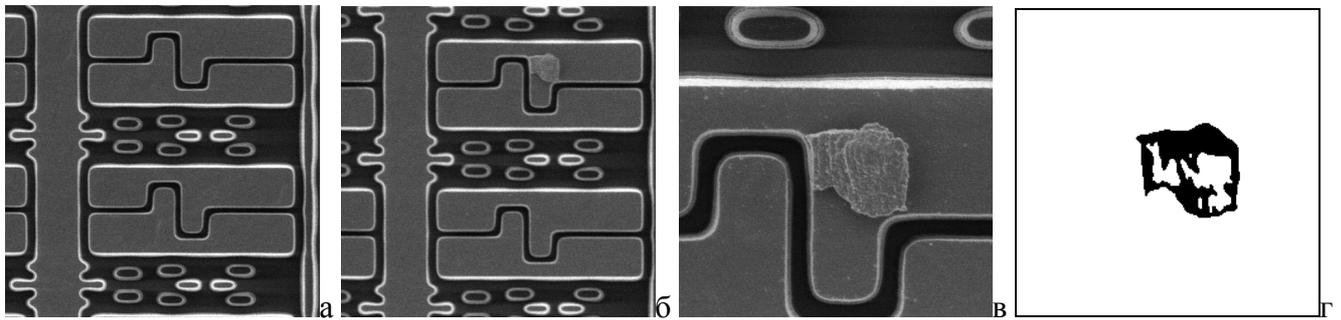
а) Эталонный снимок; б) снимок с дефектом; в) маска обнаруженного дефекта.



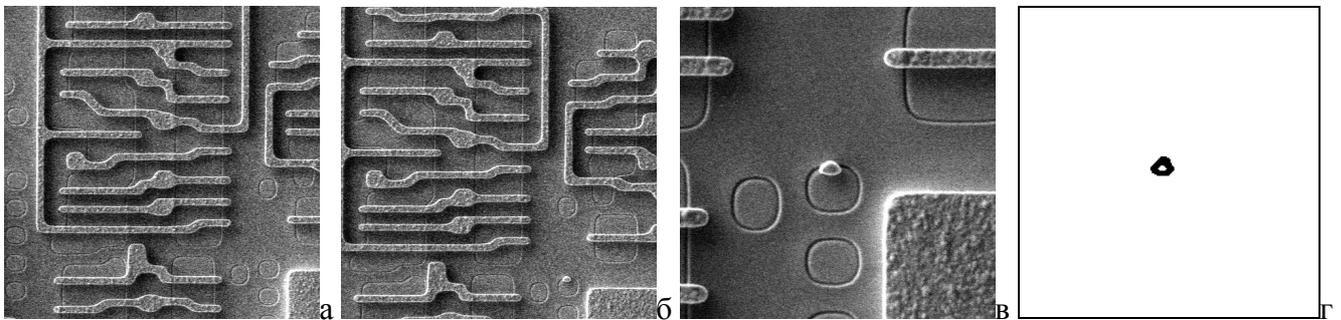
а) Эталонный снимок; б) снимок с дефектом; в) маска обнаруженного дефекта.

2.2. Обнаружение дефектов на снимках, получаемых сканирующим электронным микроскопом

Исходные данные состоят из трех снимков: два с малым увеличением (один эталонный без дефекта, другой с дефектом), а третий с большим увеличением, содержащий только участок с дефектом. Дефектами могут являться как посторонние образования, так и некорректности форм проводников; они могут находиться в любом месте изображения, иметь произвольные размеры, форму и яркость. Задача состоит в обнаружении дефекта путем анализа изображений малого увеличения и определении его формы в виде маски по изображению с дефектом большого увеличения.



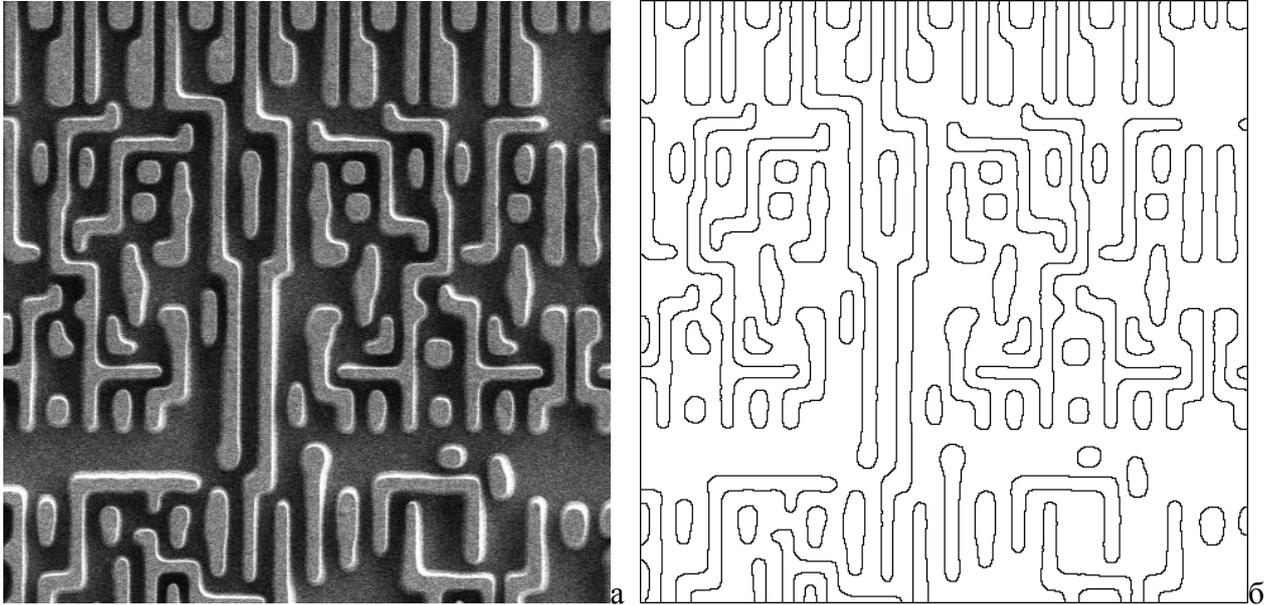
а) Эталонный снимок малого увеличения; б) снимок с дефектом малого увеличения; в) участок с дефектом большого увеличения; г) маска обнаруженного дефекта.



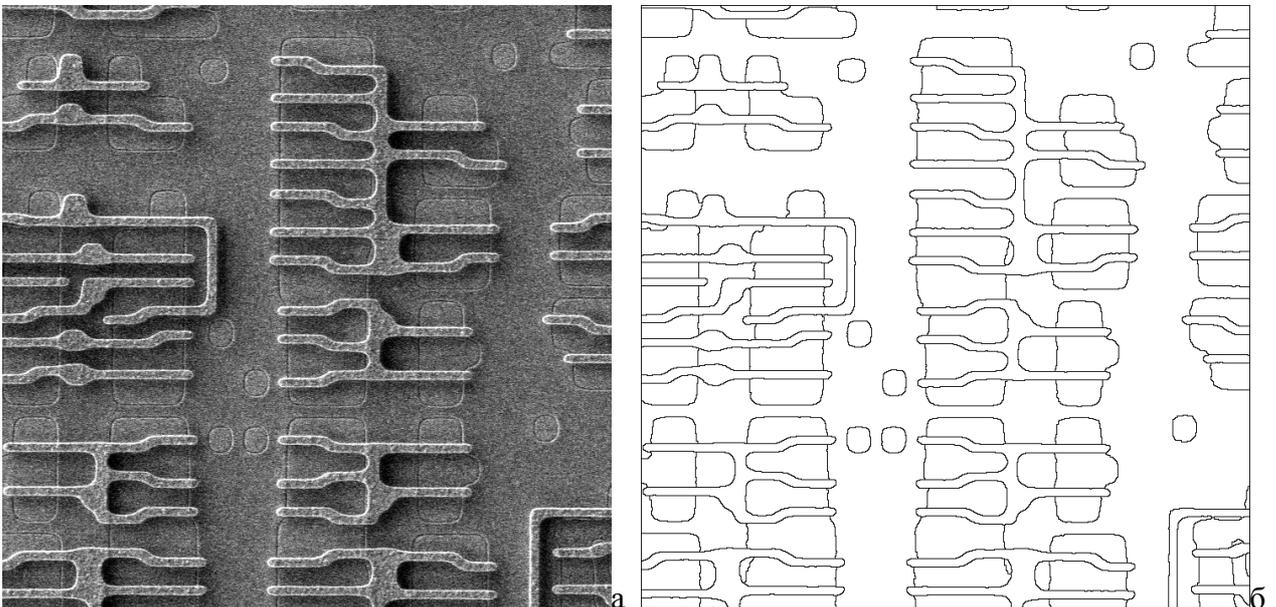
а) Эталонный снимок малого увеличения; б) снимок с дефектом малого увеличения; в) участок с дефектом большого увеличения; г) маска обнаруженного дефекта.

2.3. Сегментация областей на снимках, получаемых сканирующим электронным микроскопом

Постановка задачи и алгоритм сегментации изображений рассмотрены в разделе 5.1.



а) Исходный снимок; б) результат сегментации.



а) Исходный снимок; б) результат сегментации.

Разработанные методы анализа, обнаружения дефектов и сегментации использовались компанией MicroSpec Technologies Ltd. (Carl Zeiss Group) в процедуре диагностики изделий при промышленном производстве интегральных микросхем.

Приложение 3. Номинация программного комплекса на Гран-при CeBit

Разработанный автором программный комплекс — специализированная система обработки изображений «IRBIS» — на Международной выставке компьютерных и информационных технологий CeBit (г. Ганновер, Германия) в 1992 г. был признан Национальным номинантом на Гран-при «Golden Softies» в области инноваций и разработки программного обеспечения в Европе. Опубликован в работах [164, 180, 191, 196].



Software for Europe »Golden Softies«

IRBIS
IPPI AS Russia

has been recognized in the
following Status

National Nominee

for Software innovation
and development in Europe.

CeBIT 1992

Hannover, 16 March 1992


1&1 EDW Marketing GmbH
Ralph Dommermuth


IDG Communications Verlag AG
Eckhard Utpadel

Приложение 4. Документы о внедрении