

*Саманте и Дженис,
Дэвиду и Джонатану*

Содержание

Предисловие к английскому изданию	11
Благодарности	15
Сайт книги в сети Интернет	16
Об авторах	17
Рафаэл С. Гонсалес.....	17
Ричард Е. Вудс.....	18
Предисловие научного редактора перевода	19
Глава 1. Введение	22
1.1. Что такое цифровая обработка изображений?.....	22
1.2. Истоки цифровой обработки изображений.....	24
1.3. Примеры областей применения цифровой обработки изображений.....	29
1.3.1. Формирование изображений с помощью гамма-лучей.....	30
1.3.2. Рентгеновские изображения.....	32
1.3.3. Изображения в ультрафиолетовом диапазоне.....	34
1.3.4. Изображения в видимом и инфракрасном диапазонах.....	36
1.3.5. Изображения в микроволновом диапазоне.....	43
1.3.6. Изображения в диапазоне радиоволн.....	44
1.3.7. Примеры, иллюстрирующие другие способы формирования изображений.....	46
1.4. Основные стадии цифровой обработки изображений.....	51
1.5. Компоненты системы обработки изображений.....	54
Заключение.....	58
Ссылки и литература для дальнейшего изучения.....	58
Литература, добавленная при переводе.....	61
Глава 2. Основы цифрового представления изображений	65
Введение.....	65
2.1. Элементы зрительного восприятия.....	65
2.1.1. Строение человеческого глаза.....	66
2.1.2. Формирование изображения в глазу.....	69
2.1.3. Яркостная адаптация и контрастная чувствительность.....	70
2.2. Свет и электромагнитный спектр.....	74
2.3. Считывание и регистрация изображения.....	78
2.3.1. Регистрация изображения с помощью одиночного сенсора.....	79
2.3.2. Регистрация изображения с помощью линейки сенсоров.....	80
2.3.3. Регистрация изображения с помощью матрицы сенсоров.....	82
2.3.4. Простая модель формирования изображения.....	83
2.4. Дискретизация и квантование изображения.....	85
2.4.1. Основные понятия, используемые при дискретизации и квантовании.....	85
2.4.2. Представление цифрового изображения.....	87
2.4.3. Пространственное и яркостное разрешения.....	92
2.4.4. Интерполяция цифрового изображения.....	99

2.5. Некоторые фундаментальные отношения между пикселями.....	102
2.5.1. Соседи отдельного элемента.....	102
2.5.2. Смежность, связность, области и границы.....	102
2.5.3. Меры расстояния.....	105
2.6. Введение в математический аппарат, применяемый в цифровой обработке изображений.....	107
2.6.1. Поэлементные и матричные операции.....	107
2.6.2. Линейные и нелинейные преобразования.....	108
2.6.3. Арифметические операции.....	109
2.6.4. Теоретико-множественные и логические операции.....	115
2.6.5. Пространственные операции.....	121
2.6.6. Векторные и матричные операции.....	129
2.6.7. Преобразования изображений.....	130
2.6.8. Вероятностные методы.....	133
Заключение.....	135
Ссылки и литература для дальнейшего изучения.....	136
Задачи.....	138
Глава 3. Яркостные преобразования и пространственная фильтрация.....	143
Введение.....	143
3.1. Предпосылки.....	143
3.1.1. Основы яркостных преобразований и пространственной фильтрации.....	143
3.1.2. О примерах, приводимых в данной главе.....	146
3.2. Некоторые основные градационные преобразования.....	147
3.2.1. Преобразование изображения в негатив.....	147
3.2.2. Логарифмическое преобразование.....	148
3.2.3. Степенные преобразования (гамма-коррекция).....	150
3.2.4. Кусочно-линейные функции преобразований.....	155
3.3. Видоизменение гистограммы.....	160
3.3.1. Эквиализация гистограммы.....	162
3.3.2. Приведение гистограммы (задание гистограммы).....	170
3.3.3. Локальная гистограммная обработка.....	179
3.3.4. Использование гистограммных статистик для улучшения изображения.....	181
3.4. Основы пространственной фильтрации.....	186
3.4.1. Механизмы пространственной фильтрации.....	186
3.4.2. Пространственная корреляция и свертка.....	188
3.4.3. Векторное представление линейной фильтрации.....	193
3.4.4. Формирование масок пространственных фильтров.....	193
3.5. Сглаживающие пространственные фильтры.....	194
3.5.1. Линейные сглаживающие фильтры.....	195
3.5.2. Фильтры, основанные на порядковых статистиках (нелинейные фильтры).....	199
3.6. Пространственные фильтры повышения резкости.....	200
3.6.1. Основы.....	201
3.6.2. Повышение резкости изображений с использованием вторых производных: лапласиан.....	203
3.6.3. Нерезкое маскирование и фильтрация с подъемом высоких частот.....	207
3.6.4. Использование производных первого порядка для (нелинейного) повышения резкости изображений: градиент.....	209
3.7. Комбинирование методов пространственного улучшения.....	212

3.8. Применение нечетких методов для яркостных преобразований и пространственной фильтрации.....	217
3.8.1. Введение.....	217
3.8.2. Начала теории нечетких множеств.....	218
3.8.3. Использование нечетких множеств.....	222
3.8.4. Использование нечетких множеств для яркостных преобразований.....	232
3.8.5. Использование нечетких множеств для пространственной фильтрации.....	235
Заключение.....	238
Ссылки и литература для дальнейшего изучения.....	238
Задачи.....	239
Глава 4. Фильтрация в частотной области.....	246
Введение.....	246
4.1. Основы.....	247
4.1.1. Краткая история ряда и преобразования Фурье.....	247
4.1.2. О примерах, приводимых в данной главе.....	248
4.2. Предварительные понятия.....	249
4.2.1. Комплексные числа.....	249
4.2.2. Ряды Фурье.....	250
4.2.3. Импульсы и их свойство отсеивания.....	250
4.2.4. Преобразование Фурье функции одной непрерывной переменной.....	252
4.2.5. Свертка.....	256
4.3. Дискретизация и преобразование Фурье дискретных функций.....	257
4.3.1. Дискретизация.....	258
4.3.2. Преобразование Фурье дискретизованных функций.....	258
4.3.3. Теорема отсчетов.....	261
4.3.4. Наложение спектров.....	264
4.3.5. Реконструкция (восстановление) функции из отсчетов.....	267
4.4. Дискретное преобразование Фурье (ДПФ) одной переменной.....	268
4.4.1. Получение ДПФ из непрерывного преобразования дискретизованных функций.....	268
4.4.2. Взаимосвязь между шагом дискретизации и частотными интервалами.....	271
4.5. Расширение на функции двух переменных.....	272
4.5.1. Двумерный импульс и его свойство отсеивания.....	272
4.5.2. Пара двумерных непрерывных преобразований Фурье.....	273
4.5.3. Двумерная дискретизация и двумерная теорема отсчетов.....	274
4.5.4. Наложение спектров при преобразовании изображений.....	276
4.5.5. Двумерное дискретное преобразование Фурье и его обращение.....	284
4.6. Некоторые свойства двумерного дискретного преобразования Фурье.....	284
4.6.1. Взаимосвязи пространственных и частотных интервалов.....	284
4.6.2. Сдвиг и поворот.....	285
4.6.3. Периодичность.....	285
4.6.4. Свойства симметрии.....	287
4.6.5. Фурье-спектр и фаза.....	293
4.6.6. Двумерная теорема о свертке.....	298
4.6.7. Краткое изложение свойств двумерного дискретного преобразования Фурье.....	304
4.7. Основы фильтрации в частотной области.....	304
4.7.1. Дополнительные характеристики частотной области.....	304

4.7.2. Основы частотной фильтрации	306
4.7.3. Последовательность шагов частотной фильтрации	313
4.7.4. Соответствие между пространственными и частотными фильтрами	314
4.8. Частотные фильтры сглаживания изображения	320
4.8.1. Идеальные фильтры низких частот	320
4.8.2. Фильтры низких частот Баттерворта	324
4.8.3. Гауссовы фильтры низких частот	327
4.8.4. Дополнительные примеры низкочастотной фильтрации	330
4.9. Повышения резкости изображений частотными фильтрами	332
4.9.1. Идеальные фильтры высоких частот	334
4.9.2. Фильтры высоких частот Баттерворта	336
4.9.3. Гауссовы фильтры высоких частот	336
4.9.4. Лапласиан в частотной области	338
4.9.5. Нерезкое маскирование, высокочастотная фильтрация с подъемом частотной характеристики, фильтрация с усилением высоких частот	339
4.9.6. Гомоморфная фильтрация	342
4.10. Избирательная фильтрация	346
4.10.1. Режекторные и полосовые пропускающие фильтры	346
4.10.2. Узкополосные фильтры	347
4.11. Вопросы реализации	351
4.11.1. Разделимость двумерного ДПФ	351
4.11.2. Вычисление обратного ДПФ при помощи алгоритма прямого ДПФ	352
4.11.3. Быстрое преобразование Фурье	353
4.11.4. Некоторые замечания по поводу построения фильтров	356
Заключение	357
Ссылки и литература для дальнейшего изучения	357
Задачи	359
Глава 5. Восстановление и реконструкция изображений	366
Введение	366
5.1. Модель процесса искажения/восстановления изображения	367
5.2. Модели шума	368
5.2.1. Пространственные и частотные свойства шума	369
5.2.2. Функции плотности распределения вероятностей для некоторых важных типов шума	369
5.2.3. Периодический шум	374
5.2.4. Построение оценок для параметров шума	375
5.3. Подавление шумов — пространственная фильтрация	378
5.3.1. Усредняющие фильтры	378
5.3.2. Фильтры, основанные на порядковых статистиках	382
5.3.3. Адаптивные фильтры	386
5.4. Подавление периодического шума — частотная фильтрация	393
5.4.1. Режекторные фильтры	393
5.4.2. Полосовые фильтры	395
5.4.3. Узкополосные фильтры	396
5.4.4. Оптимальная узкополосная фильтрация	398
5.5. Линейные трансляционно-инвариантные искажения	403
5.6. Оценка искажающей функции	406
5.6.1. Оценка на основе визуального анализа изображения	407
5.6.2. Оценка на основе эксперимента	407
5.6.3. Оценка на основе моделирования	408
5.7. Инверсная фильтрация	411

5.8. Фильтрация методом минимизации среднего квадрата отклонения (винеровская фильтрация).....	414
5.9. Фильтрация методом минимизации сглаживающего функционала со связью.....	418
5.10. Среднегеометрический фильтр.....	424
5.11. Реконструкция изображения по проекциям.....	424
5.11.1. Введение.....	425
5.11.2. Принципы компьютерной томографии (КТ).....	428
5.11.3. Проекции и преобразование Радона.....	432
5.11.4. Теорема о центральном сечении.....	438
5.11.5. Реконструкция по проекциям в параллельных пучках методом фильтрации и обратного проецирования.....	439
5.11.6. Реконструкция на основе фильтрованных обратных проекций с верным пучком.....	445
Заключение.....	452
Ссылки и литература для дальнейшего изучения.....	453
Задачи.....	454
Глава 6. Обработка цветных изображений.....	460
Введение.....	460
6.1. Основы теории цвета.....	461
6.2. Цветовые модели.....	468
6.2.1. Цветовая модель RGB.....	469
6.2.2. Цветовые модели CMY и CMYK.....	474
6.2.3. Цветовая модель HSI.....	475
6.3. Обработка изображений в псевдоцветах.....	483
6.3.1. Квантование по яркости.....	484
6.3.2. Преобразование яркости в цвет.....	489
6.4. Основы обработки цветных изображений.....	495
6.5. Цветовые преобразования.....	496
6.5.1. Постановка задачи.....	496
6.5.2. Цветовое дополнение.....	501
6.5.3. Вырезание цветового диапазона.....	502
6.5.4. Яркостная и цветовая коррекция.....	504
6.5.5. Обработка гистограмм.....	509
6.6. Сглаживание и повышение резкости.....	511
6.6.1. Сглаживание цветных изображений.....	511
6.6.2. Повышение резкости цветных изображений.....	514
6.7. Сегментация изображения, основанная на цвете.....	515
6.7.1. Сегментация в цветовом пространстве HSI.....	515
6.7.2. Сегментация в цветовом пространстве RGB.....	517
6.7.3. Обнаружение контуров на цветных изображениях.....	520
6.8. Шум на цветных изображениях.....	524
6.9. Сжатие цветных изображений.....	527
Заключение.....	529
Ссылки и литература для дальнейшего изучения.....	529
Задачи.....	530
Глава 7. Вейвлеты и кратномасштабная обработка.....	536
Введение.....	536
7.1. Предпосылки.....	537

7.1.1. Пирамиды изображений	538
7.1.2. Субполосное кодирование	543
7.1.3. Преобразование Хаара	550
7.2. Кратномасштабное разложение	554
7.2.1. Разложения в ряды	554
7.2.2. Масштабирующие функции	557
7.2.3. Вейвлет-функции	562
7.3. Одномерные вейвлет-преобразования	566
7.3.1. Разложение в вейвлет-ряды	566
7.3.2. Дискретное вейвлет-преобразование	569
7.3.3. Интегральное вейвлет-преобразование	571
7.4. Быстрое вейвлет-преобразование	574
7.5. Двумерные вейвлет-преобразования	583
7.6. Вейвлет-пакеты	593
Заключение	604
Ссылки и литература для дальнейшего изучения	604
Задачи	606
Глава 8. Сжатие изображений	611
Введение	611
8.1. Основы	612
8.1.1. Кодовая избыточность	614
8.1.2. Пространственная и временная избыточность	616
8.1.3. Лишняя информация	617
8.1.4. Измерение содержащейся в изображении информации	618
8.1.5. Критерии верности воспроизведения	621
8.1.6. Модели сжатия изображений	624
8.1.7. Форматы изображений, контейнеры и стандарты сжатия	626
8.2. Некоторые основные методы сжатия	630
8.2.1. Кодирование Хаффмана	630
8.2.2. Кодирование Голомба	633
8.2.3. Арифметическое кодирование	637
8.2.4. LZW-кодирование	640
8.2.5. Кодирование длин серий	644
8.2.6. Кодирование на базе шаблонов	651
8.2.7. Кодирование битовых плоскостей	655
8.2.8. Блочное трансформационное кодирование	659
8.2.9. Кодирование с предсказанием	679
8.2.10. Вейвлет-кодирование	700
8.3. Нанесение цифровых водяных знаков на изображение	713
Заключение	722
Ссылки и литература для дальнейшего изучения	722
Задачи	724
Глава 9. Морфологическая обработка изображений	728
Введение	728
9.1. Начальные сведения	728
9.2. Эрозия и дилатация	732
9.2.1. Эрозия	732
9.2.2. Дилатация	734
9.2.3. Двойственность	737

9.3. Размыкание и замыкание.....	737
9.4. Преобразование «попадание/пропуск».....	742
9.5. Некоторые основные морфологические алгоритмы.....	745
9.5.1. Выделение границ.....	745
9.5.2. Заполнение дырок.....	746
9.5.3. Выделение связанных компонент.....	748
9.5.4. Выпуклая оболочка.....	751
9.5.5. Утончение.....	753
9.5.6. Утолщение.....	754
9.5.7. Построение остова.....	755
9.5.8. Усечение.....	758
9.5.9. Морфологическая реконструкция.....	761
9.5.10. Сводная таблица морфологических операций.....	768
9.6. Морфология полутоновых изображений.....	771
9.6.1. Эрозия и дилатация.....	772
9.6.2. Размыкание и замыкание.....	774
9.6.3. Некоторые основные алгоритмы полутоновой морфологии.....	777
9.6.4. Полутоновая морфологическая реконструкция.....	784
Заключение.....	787
Ссылки и литература для дальнейшего изучения.....	788
Задачи.....	788
Глава 10. Сегментация изображений.....	797
Введение.....	797
10.1. Основы.....	798
10.2. Обнаружение точек, линий и перепадов.....	800
10.2.1. Основы.....	801
10.2.2. Обнаружение изолированных точек.....	805
10.2.3. Обнаружение линий.....	806
10.2.4. Модели перепадов.....	811
10.2.5. Простые методы обнаружения контурных перепадов.....	817
10.2.6. Более совершенные методы обнаружения контуров.....	826
10.2.7. Связывание контуров и нахождение границ.....	840
10.3. Пороговая обработка.....	854
10.3.1. Обоснование.....	854
10.3.2. Обработка с глобальным порогом.....	858
10.3.3. Метод Оцу оптимального глобального порогового преобразования.....	860
10.3.4. Применение сглаживания изображения для улучшения обработки с глобальным порогом.....	865
10.3.5. Использование контуров для улучшения обработки с глобальным порогом.....	867
10.3.6. Обработка с несколькими порогами.....	870
10.3.7. Обработка с переменным порогом.....	874
10.3.8. Пороги, основанные на нескольких переменных.....	880
10.4. Сегментация на отдельные области.....	881
10.4.1. Выращивание областей.....	882
10.4.2. Разделение и слияние областей.....	886
10.5. Сегментация по морфологическим водоразделам.....	889
10.5.1. Исходные предпосылки.....	889
10.5.2. Построение перегородок.....	892
10.5.3. Алгоритм сегментации по водоразделам.....	894

10.5.4. Использование маркеров.....	896
10.6. Использование движения при сегментации.....	899
10.6.1. Пространственные методы.....	899
10.6.2. Частотные методы.....	903
Заключение.....	907
Ссылки и литература для дальнейшего изучения.....	907
Задачи.....	909
Глава 11. Представление и описание.....	919
Введение.....	919
11.1. Представление.....	919
11.1.1. Прослеживание границы.....	920
11.1.2. Цепные коды.....	922
11.1.3. Аппроксимация ломаной линией минимальной длины.....	925
11.1.4. Другие методы аппроксимации ломаной линией.....	932
11.1.5. Сигнатуры.....	934
11.1.6. Сегменты границы.....	936
11.1.7. Остовы областей.....	938
11.2. Дескрипторы границ.....	941
11.2.1. Некоторые простые дескрипторы.....	941
11.2.2. Нумерация фигур.....	942
11.2.3. Фурье-дескрипторы.....	944
11.2.4. Статистические характеристики.....	947
11.3. Дескрипторы областей.....	948
11.3.1. Некоторые простые дескрипторы.....	949
11.3.2. Топологические дескрипторы.....	950
11.3.3. Текстурные дескрипторы.....	954
11.3.4. Инварианты моментов двумерных функций.....	966
11.4. Использование главных компонент для описания.....	969
11.5. Реляционные дескрипторы.....	980
Заклучение.....	983
Ссылки и литература для дальнейшего изучения.....	984
Задачи.....	985
Глава 12. Распознавание объектов.....	990
Введение.....	990
12.1. Образы и классы образов.....	990
12.2. Распознавание на основе методов теории принятия решений.....	995
12.2.1. Сопоставление.....	996
12.2.2. Статистически оптимальные классификаторы.....	1003
12.2.3. Нейронные сети.....	1013
12.3. Структурные методы распознавания.....	1036
12.3.1. Сопоставление номеров фигур.....	1036
12.3.2. Сопоставление строк символов.....	1037
Заклучение.....	1039
Ссылки и литература для дальнейшего изучения.....	1040
Задачи.....	1041
Приложения.....	1045
Кодовые таблицы для сжатия изображений.....	1045
Литература.....	1050
Предметный указатель.....	1081

Предисловие к английскому изданию

Если что-то удастся читать без усилий,
значит больших усилий стоило это написать.

Эрике Хардиел Понсела

Настоящее издание является результатом значительной переработки книги «Цифровая обработка изображений». При подготовке этого издания, как и при подготовке предыдущих четырех редакций (Гонсалес и Уинтц, 1977 г. и 1978 г.; Гонсалес и Вудс, 1992 г. и 2002 г.), мы ориентировались прежде всего на студентов и преподавателей. Как и прежде, основные цели этой книги — служить введением в основные понятия и методы цифровой обработки изображений, а также создать основу для последующего изучения и проведения самостоятельных исследований в этой области. С учетом этих целей мы вновь сконцентрировались на материале, который считаем фундаментальным и применимым не только для решения узкоспециальных задач. Уровень математической сложности книги остался в рамках программы колледжа высшей ступени или первого курса университета и предполагает начальную подготовку в области математического анализа, линейной алгебры и линейных систем, теории вероятности и математической статистики, а также компьютерного программирования. Дополняющий книгу Интернет-сайт содержит учебные материалы, помогающие читателям при необходимости освежить в памяти эти темы.

Одна из важнейших причин, почему настоящая книга уже более 30 лет является мировым лидером в своей области, — это высокая степень внимания к изменению образовательных потребностей наших читателей. Нынешнее издание базируется на самом обширном из когда-либо нами проводившихся исследований читательского мнения. К исследованию были привлечены профессорско-преподавательский состав, студенты и независимые читатели из 134 организаций в 32 странах. Исследование показало, что необходимо:

- расширить введение в начале книги касательно математических инструментов, используемых в обработке изображений;
- развернуто объяснить технику обработки гистограмм;
- излагать сложные алгоритмы в форме пошагового краткого описания;
- подробнее объяснить пространственную корреляцию и свертку;
- добавить введение в теорию нечетких множеств и применение ее в цифровой обработке изображений;
- пересмотреть материал, относящийся к частотной области, начиная с базовых принципов и демонстрации того, как дискретное преобразование Фурье следует из дискретизации данных;
- охватить тематику, связанную с компьютерной томографией;
- разъяснить основные понятия в главе о вейвлет-анализе;
- переработать главу о сжатии данных, чтобы шире охватить методы сжатия видеoinформации, обновленные стандарты и технологию цифровых водяных знаков;
- расширить главу о морфологии с целью охвата морфологического восстановления изображений и переработать раздел о полутоновой морфологии;

- расширить тематику сегментации изображений за счет более совершенных методов обнаружения контуров, таких как алгоритм Кэнни, и более широкой трактовки пороговой обработки изображений;
- обновить содержание главы о представлении и описании изображений;
- упростить материал, относящийся к структурным методам распознавания образов.

Переработка и дополнение книги, предпринятые в этом издании, отражают наше стремление соблюсти разумный баланс между строгостью и ясностью изложения и конъюнктурными соображениями, оставаясь при этом в рамках приемлемого объема книги. Наиболее важные изменения, сделанные в этом издании книги, перечислены ниже.

Глава 1: Внесены изменения в несколько иллюстраций, а текст частично переработан в соответствии с изменениями в последующих главах.

Глава 2: Приблизительно половина главы переработана, чтобы включить новые изображения и более понятные объяснения. В число крупных изменений входят новый раздел, посвященный интерполяции изображений, и обширный новый раздел, обобщающий главные математические инструменты, которые используются в книге. Однако вместо сухого перечисления математических концепций мы использовали эту возможность, чтобы во второй главе ввести в оборот ряд прикладных задач обработки изображений, которые прежде были разбросаны по всей книге. Например, в эту главу мы перенесли усреднение и вычитание изображений в качестве иллюстрации арифметических операций. Здесь мы следуем тенденции, начатой во втором издании книги: как можно раньше вводить в обсуждение приложения, не только для иллюстрации, но и для мотивации студентов. Закончив изучение главы 2 в ее обновленном виде, читатель получит базовое понимание того, как осуществляется манипулирование цифровыми изображениями и их обработка. Это та твердая основа, на которой строится остальная часть книги.

Глава 3: Основная переработка этой главы связана с подробным обсуждением пространственной корреляции и свертки и их применением для фильтрации изображений с помощью пространственных масок. Исследование рынка также выявило потребность в численных примерах, иллюстрирующих технику эквализации и задания гистограммы, поэтому мы добавили несколько примеров для демонстрации действия этих инструментов обработки. При исследовании часто высказывалось пожелание рассмотреть нечеткие множества и их приложения к обработке изображений. Мы включили в эту главу новый раздел, посвященный основам теории нечетких множеств и ее применениям для ярких преобразований и пространственной фильтрации — двум главным областям использования этой теории в обработке изображений.

Глава 4: За последние четыре года больше всего замечаний и предложений мы слышали в связи с изменениями, сделанными в главе 4 при переходе от первого ко второму изданию. Целью этих изменений было упростить изложение преобразования Фурье и частотной области. Очевидно, мы зашли слишком далеко, и многие пользователи книги жаловались на излишнюю поверхностность материала. В настоящем издании мы исправили этот дефект. Теперь изложение материала начинается с преобразования Фурье одной непрерывной переменной и далее переходит к выводу дискретного фурье-преобразования,

исходя из базовых понятий дискретизации и свертки. Побочным результатом такой формы подачи материала стал наглядный вывод теоремы отсчетов и ее следствий. После этого изложение для одномерной постановки распространяется на двумерный случай, где мы приводим ряд примеров, иллюстрирующих влияние дискретизации на цифровые изображения, в том числе эффекты, вызванные наложением спектров и муаром. Затем мы иллюстрируем двумерное дискретное преобразование Фурье, выводя и суммируя ряд его важных свойств. Эти понятия далее используются в качестве основы для фильтрации в частотной области. В заключение обсуждаются вопросы реализации, такие как разложение преобразований и вывод алгоритма быстрого преобразования Фурье. К концу этой главы читатель перейдет от дискретизации одномерных функций через понятный вывод основ дискретного Фурье-преобразования к некоторым важнейшим его применениям для цифровой обработки изображений.

Глава 5: Основным изменением в этой главе стало добавление раздела, касающегося реконструкции изображения по проекциям, с упором на компьютерную томографию (КТ). Изложение КТ начинается с наглядного примера, демонстрирующего основополагающие принципы реконструкции изображения по проекциям, и описания различных применяемых на практике методов формирования изображения. Затем мы выводим преобразование Радона и теорему о центральном сечении, используя их как основу для формулировки концепции фильтрации и обратного проецирования. Обсуждается реконструкция изображений с помощью как параллельных, так и веерных пучков; обсуждение иллюстрируется примерами. Включение этого уже давно ожидавшегося материала представляет собой важное добавление к книге.

Глава 6: Изменения в этой главе ограничивались только уточнениями и незначительной коррекцией обозначений без добавления каких-либо новых понятий.

Глава 7: Многочисленные полученные замечания относились к тому, что переход от предшествующих глав к вейвлетам оказался трудным для начинающих. Чтобы сделать материал более понятным, мы заново переписали несколько подготовительных разделов.

Глава 8: Эта глава полностью переработана с учётом достигнутого современного уровня. Основные изменения касаются новых методов кодирования, более широкого обсуждения видеоданных, полностью переработанного раздела о стандартах сжатия и введения в технологию цифровых водяных знаков. Новая организация материала упрощает его освоение начинающими студентами.

Глава 9: Основные изменения в этой главе состоят в добавлении нового раздела о морфологической реконструкции и полной переработке раздела, посвященного полутоновой морфологии. Благодаря рассмотрению морфологической реконструкции как для двоичных, так и для полутоновых изображений, появилась возможность изложить более сложные и полезные морфологические алгоритмы, чем прежде.

Глава 10: Эта глава также подверглась значительной переработке. Сохранив прежнюю организацию, мы добавили новый материал, лучше подчеркивающий основные принципы, а также обсуждение более совершенных методов сегментации. Модели перепадов теперь рассматриваются и иллюстрируются более подробно, как и свойства градиента. Детекторы контуров Марра-Хилдрета

и Кэнни добавлены в качестве иллюстрации более совершенных методов обнаружения контуров. Также переработан раздел о пороговой обработке, в который добавлен метод Оцу нахождения оптимального порога, популярность которого значительно возросла за последние несколько лет. Мы включили этот подход вместо выбора оптимального порога на основе байесовского правила классификации не только потому, что его легче понять и реализовать, но и поскольку он значительно чаще используется на практике. Байесовский подход перенесен в главу 12, где решающее правило Байеса обсуждается более подробно. Мы также добавили обсуждение того, как использовать информацию о контурах для улучшения пороговой обработки, и несколько новых примеров обработки с адаптивным порогом. В разделы, посвященные морфологическим водоразделам и использованию движения для сегментации, внесены только незначительные уточнения.

Глава 11: Главные изменения в этой главе связаны с добавлением алгоритма прослеживания границы, более подробным изложением алгоритма аппроксимации дискретной границы с помощью ломаной минимальной длины и новым разделом об использовании матриц сочетаемости яркостей для описания текстуры. Новыми являются многие примеры в разделах 11.2 и 11.3, а также все примеры в разделе 11.4.

Глава 12: Изменения в этой главе включают новый раздел о корреляционном сопоставлении и новый пример, демонстрирующий использование байесовского классификатора для распознавания интересных областей в мультиспектральных изображениях. Раздел о структурной классификации теперь ограничен обсуждением только сопоставления строк символов.

Все вышеуказанные изменения выразились в появлении в книге более 400 новых изображений, свыше 200 новых рисунков и таблиц, и более 80 новых задач и упражнений для самостоятельной работы. Сложные процедуры обработки по мере необходимости сведены в сжатую форму пошаговых алгоритмов. Также обновилась ссылка на литературу в конце каждой главы.

Интернет-сайт книги, основанный при выпуске второго издания, успешно действует, привлекая более 20 тысяч посетителей ежемесячно. Этот сайт был переделан и пополнен новой информацией в соответствии с выпуском настоящего издания книги. Более подробно возможности сайта и его наполнение представлены в разделе «Сайт книги в сети Интернет», сразу после раздела «Благодарности».

Это издание книги «Цифровая обработка изображений» отражает то, как изменились с 2002 г. образовательные потребности наших читателей. Как обычно для подобных проектов, прогресс продолжается и после окончания работы над рукописью. Одна из причин, по которым эта книга была так хорошо принята, начиная с ее появления в 1977 г., — это ее постоянный акцент на фундаментальных концепциях. Такой подход является одной из попыток указать устойчивые ориентиры в стремительно развивающейся совокупности знаний. Мы старались соблюсти тот же принцип при подготовке этого издания книги.

Р. С. Г.
Р. Е. В.

Благодарности

Мы в неоплатном долгу перед множеством лиц, как из академических кругов, так и из промышленности и правительства, которые внесли большой вклад в это издание книги. Их помощь была столь важной и разносторонней, что нам трудно выделить их как-то иначе, чем просто перечислить в алфавитном порядке. В частности, мы глубоко признательны за те многочисленные советы по улучшению формы изложения и содержания книги, которые предложили наши коллеги Монги А. Абиди, Стивен Л. Эддинс, Йонг Мин Ким, Брайан Морс, Эндрю Олдройд, Али М. Реза, Эдгардо Фелипе Риверон, Хосе Руиз Шульклопер и Кэмерон Х. Г. Райт.

Многие лица и организации оказали нам ценную помощь во время написания настоящего издания. Мы перечисляем их опять-таки по алфавиту. Мы особо признательны Кертни Эспозито и Наоми Фернандес из компании The MathWorks, предоставившим нам пакет программ MATLAB и оказавшим помощь в его использовании, что позволило построить и уточнить многие примеры и экспериментальные результаты, включенные в настоящее издание. Значительная доля использованных в нем новых изображений (а в некоторых случаях — их история и объяснение) поступили к нам благодаря усилиям лиц, чей вклад воистину неоценим, среди которых, в частности: Серж Бехер, Меллисса Д. Байнд, Джеймс Бланкеншип, Уве Боос, Эрнесто Брибиеска, Майкл Е. Кейси, Майкл У. Дэвидсон, Сьюзен Л. Форсбург, Томас Р. Гест, Лалит Гупта, Дэниел А. Хаммер, Жонг Хе, Роджер Хиди, Хуан А. Херрера, Джон М. Хьюдек, Майкл Гурвиц, Крис Й. Йохансен, Ронда Найтон, Эшли Мохамед, А. Моррис, Кертис С. Обер, Джозеф Е. Пассенте, Дэвид Р. Пикенс, Майкл Робинсон, Барретт А. Шефер, Майкл Шаффер, Пит Сайтс, Салли Стоу, Крейг Уотсон, Дэвид К. Уи и Роберт А. Вест. Мы также признательны другим лицам и организациям, упоминающимся в подписях ко многим фотографиям, за любезное разрешение использовать в книге предоставленные ими материалы.

Сотрудники издательства Прентис-Холл Винс О'Брайен, Роуз Кернан, Скотт Дисанно, Майкл Мак-Дональд, Джо Руддик, Хитер Скотт и Элис Дворкин заслуживают особой благодарности. Их творческие способности, помощь и терпение при выпуске книги были действительно неоценимыми.

Р. С. Г.
Р. Е. В.

Сайт книги в сети Интернет

www.prenhall.com/gonzalezwoods

или зеркальный сайт

www.imageprocessingplace.com

Хотя книга «Цифровая обработка изображений» содержит в себе весь необходимый материал, дополняющий ее сайт в Интернет предлагает помощь в ряде важных вопросов.

Для студентов и самостоятельных читателей на сайте имеются :

- краткие обзоры по таким областям, как теория вероятностей, математическая статистика, линейная алгебра;
- подробные решения к некоторым из предлагаемых в книге упражнений;
- учебные проекты для самостоятельной компьютерной реализации;
- раздел «Учебные пособия» с десятками обучающих материалов по большинству тем, рассматриваемых в книге;
- база данных, содержащая все изображения из книги.

Для преподавателей сайт предлагает:

- «Пособие для преподавателя» с полными решениями всех упражнений из книги, а также примерами учебных планов и методическими указаниями к проведению лабораторных работ. Это руководство бесплатно предоставляется преподавателям, которые принимают данную книгу в качестве основы для учебного курса;
- материалы для учебных презентаций в формате PowerPoint;
- изъятые из книги части прежних ее изданий в формате PDF, которые преподаватель может при желании получить по сети и использовать в учебном процессе;
- многочисленные ссылки на другие образовательные ресурсы.

Для практиков на сайте имеются дополнительные специализированные разделы, например:

- ссылки на коммерческие сайты Интернет по тематике;
- отдельные новые литературные ссылки;
- ссылки на коммерческие базы данных изображений.

Данный Интернет-сайт является идеальным инструментом для актуализации книги в промежутках между изданиями, позволяя добавлять новые темы, цифровые изображения и прочий материал, появившийся после выхода книги из печати. Хотя выпуск книги осуществлялся с большим вниманием и заботой, сайт также позволяет удобно размещать информацию об обнаруженных ошибках. Ссылки на материалы сайта отмечаются на полях следующим значком:



Об авторах

Рафаэл С. Гонсалес

Р. С. Гонсалес получил степень бакалавра в университете шт. Майами в 1965 г., а затем степени магистра и доктора философии по электротехнике в университете шт. Флорида (г. Гейнсвилл) в 1967 и 1970 гг. соответственно. В 1970 г. он поступил на работу на факультет электронной и компьютерной техники университета шт. Теннесси (УТК) в г. Ноксвилл, где последовательно получил звания адъюнкт-профессора (1973 г.), профессора (1978 г.) и профессора с особыми заслугами (1984 г.). С 1994 по 1997 гг. он был президентом факультета, а в настоящее время является почетным профессором УТК в отставке.

Гонсалес был основателем двух лабораторий УТК: Лаборатории анализа изображений и образов и Лаборатории робототехники и машинного зрения. Он также основал в 1982 г. компанию Perceptics Corporation и вплоть до 1992 г. был ее президентом. В течение трех последних лет этого периода он являлся сотрудником Westinghouse Corporation, которая приобрела компанию Perceptics в 1989 г.

Под руководством Гонсалеса компания Perceptics достигла больших успехов в обработке изображений, машинном зрении и технологии запоминающих устройств на основе лазерных дисков. За первое десятилетие своей работы компания Perceptics внедрила ряд новаторских изделий, в том числе: первую в мире коммерческую систему машинного зрения для автоматического чтения номерных знаков транспортных средств; ряд крупномасштабных систем обработки и хранения изображений, которые применяются на шести различных заводах ВМФ США при контроле реактивных двигателей ракет «Трайидент 2» для подводных лодок; семейство плат обработки изображений для модернизации компьютеров «Макинтош», занимающее передовые позиции на рынке; а также линейку запоминающих устройств на базе лазерных дисков терабайтной емкости.

Он часто привлекается промышленными предприятиями и правительством в качестве консультанта в области распознавания образов, обработки изображений и обучающихся машин. Его академические заслуги в этих областях отмечены в 1977 г. премией технического факультета УТК за высокие достижения; в 1978 г. премией ректора УТК для ученых-исследователей; в 1980 г. — премией им. М. Е. Брукса как выдающийся профессор и почетным званием профессора Magnavox Engineering. В 1981 г. он становится профессором IBM в университете шт. Теннесси, а в 1984 г. получил там звание профессора с особыми заслугами. Он отмечен званием выдающегося выпускника университета шт. Майами (1985 г.), премией научного общества «фи-каппа-фи» (1986 г.) и премией им. Натана В. Догерти университета шт. Теннесси за высокие достижения в области техники (1992 г.).

Среди наград за промышленные достижения — премия выдающемуся инженеру IEEE 1987 г. за коммерческие разработки в Теннесси; Национальная премия им. Альберта Роуза 1988 г. за успехи в создании коммерческих систем обработки

изображений; премия Б. Отто Уили за достижения в переносе технологий; звание «предприниматель года» агентства Куперс и Либранд в 1989 г.; премия выдающемуся инженеру IEEE 1992 г.; и Национальная премия 1993 г. за развитие технологии Ассоциации по автоматизированной обработке изображений.

Гонсалес является автором и соавтором свыше 100 технических статей, двух монографий и четырех учебников в области распознавания образов, обработки изображений и робототехники. Его книгами пользуются в более чем 1000 университетах и исследовательских институтах по всему миру. Он включен в престижные справочники «Кто есть кто в Америке», «Кто есть кто в технике», «Кто есть кто в мире», а также еще в 10 национальных и международных биографических перечней. Он совладелец двух патентов США и член редколлегий научных журналов «Труды по теории систем» (IEEE Transactions on Systems), «Человек и кибернетика» (Man and Cybernetics), Международного журнала по вычислительной технике и информатике (International Journal of Computer and Information Sciences). Он является членом IEEE и состоит в многочисленных профессиональных и почетных обществах (в том числе «тау-бета-пи», «фи-каппа-фи», «эта-каппа-ню» и «сигма-кси»).

Ричард Е. Вудс

Ричард Е. Вудс получил степени бакалавра, магистра и доктора философии по электротехнике в университете шт. Теннесси (г. Ноксвилл). Его профессиональный опыт охватывает широкий диапазон от предпринимательской деятельности до исполнения более традиционных обязанностей в академической, консультативной, правительственной и промышленной сферах. Недавно он основал работающую в области высоких технологий компанию MedData Interactive, которая специализируется на разработке портативных компьютерных систем медицинского назначения. Он также является соучредителем и вице-президентом компании Perceptics Corporation, отвечая в ней за разработку многих видов продукции для количественного анализа изображений и автономного принятия решений.

До работы в компаниях Perceptics и MedData д-р Вудс был доцентом по электротехнике и вычислительной технике в университете Теннесси, а еще раньше — инженером по компьютеризации в компании Union Carbide Corporation. В качестве консультанта он привлекался к разработке ряда процессоров специального назначения для многочисленных космических и военных ведомств, включая агентство NASA, Центр управления баллистическими ракетами и Национальную лабораторию Оук Ридж (Oak Ridge).

Д-р Вудс опубликовал множество статей по цифровой обработке сигналов и состоит в ряде профессиональных обществ, включая «тау-бета-пи», «фи-каппа-фи» и IEEE. В 1986 г. он был признан выдающимся выпускником университета шт. Теннесси.

Предисловие научного редактора перевода

Тот факт, что не менее 90 % информации человек получает при помощи зрения, давно превратился в банальное утверждение, однако не потерял своей значимости. Более того, объемы зрительной информации, которую приходится получать, воспринимать и анализировать, постоянно возрастают. Все бóльшая роль при этом отводится техническим системам обработки и анализа видеoinформации — как с целью преобразования и видоизменения поступающих данных для последующего предоставления человеку, так и с целью автоматического их анализа и извлечения информации, в том числе для выбора последующей реакции.

Число областей, в которых исходные данные поступают в виде изображений, огромно. Здесь и различные системы наблюдения, технического зрения, мониторинга, видеотелефонии, которые регистрируют и передают огромные объемы видеоданных, и различные автономные системы (роботы), принимающие решение на основе анализа видеoinформации, и персонализированное телевизионное вещание, и полиграфия со значительно возросшим объемом иллюстраций в печатной продукции, и медицина, и Интернет, и многие другие направления. При этом наряду со значительным повышением уровня развития техники, самую существенную роль играют методы обработки видеoinформации. Они обеспечивают улучшение изображений для их наилучшего восприятия, сжатие видеоданных для хранения и передачи по каналам связи, а также анализ, распознавание и интерпретацию зрительных образов для принятия решений и управления поведением автономных технических систем.

Широкое распространение как профессиональных, так и бытовых средств фото- и видеорегистрации, а также достаточно мощных вычислительных машин, привело к появлению разнообразных программных средств обработки изображений. Они предоставляют богатые возможности преобразования и улучшения данных, но, как правило (кроме простейших случаев), не дают понимания того, какие именно операции над изображениями производятся. В такой ситуации у пользователя нет возможности ни объяснить выполняемые действия, ни быть уверенным, что при повторном их выполнении над другими данными будет получен соответствующий результат, ни составить правильный план последовательности необходимых действий для достижения какой-то поставленной цели. Решение этих задач возможно только на основе научного подхода.

Монография известных американских специалистов в области цифровой обработки изображений Р. Гонсалеса и Р. Вудса является введением в данную область знаний и создает основы для дальнейшего изучения этой многогранной и быстро развивающейся науки. В книге излагаются основы теории и методов преобразования и анализа видеoinформации, и она является одним из наиболее известных и распространенных за рубежом учебников в данной области науки. Многие из алгоритмов, которые в ней приведены, реализованы в различных широко используемых пакетах компьютерного редактирования изображений. В США данная книга вышла уже третьим изданием (первое появилось в 1992 г.).

Перед этим было два издания книги с тем же названием, написанной Гонсалесом и Винцем (в 1977 и в 1987 гг.).

Круг рассматриваемых в книге вопросов достаточно широк. Изучаются все основные направления обработки и анализа изображений, включая основы теории восприятия и регистрации видеоинформации, основные методы фильтрации и вейвлет-преобразования, методы улучшения, восстановления и сжатия черно-белых и цветных изображений, восстановления изображения по проекциям; также обсуждаются вопросы сегментации, распознавания образов, описания и представления деталей, морфологического анализа изображения. При этом необходимо отметить, что некоторые из вопросов, например, вопросы вейвлет-преобразования изображений или сопоставление разнообразных подходов к кодированию изображений и видеопоследовательностей до настоящего момента были опубликованы на русском языке лишь в виде отдельных разрозненных статей. Без сомнения, проведенные в книге рассмотрения не являются исчерпывающими в каждом из отдельных вопросов, однако они дают достаточно хорошее представление о сущности проблем и о наиболее распространенных способах их решения.

Все разделы книги содержат достаточно подробные обсуждения, приводится большое количество примеров и обработанных изображений, иллюстрирующих рассматриваемые методы. Оригинал книги сопровождается Интернет-сайтом (англоязычным), расположенным по адресу <http://www.imageprocessingplace.com> или <http://www.prenhall.com/gonzalezwoods> (адрес зеркального сайта). На нем представлены вспомогательные и дополнительные материалы, решения предложенных авторами задач, методические материалы по составлению учебных планов, а также база данных, содержащая многие из приведенных в книге иллюстраций.

Материал книги излагается на вполне доступном уровне. Предполагается, что читатель имеет математическую подготовку в объеме первых курсов технического вуза и знаком с основами математического анализа, линейной алгебры, теории вероятностей и математической статистики, а также владеет минимальными навыками программирования.

Данная монография создавалась авторами как введение в основные понятия и методы цифровой обработки изображений, дающее основы для последующего углубленного изучения проблематики обработки изображений, а также проведения самостоятельных исследований в этой области. В качестве стиля изложения авторы выбрали постепенный переход от простого к сложному и снабдили материал большим количеством задач, приведенных в конце каждой главы. Эти качества, вместе с полнотой охвата и достаточной подробностью рассмотрения изучаемых вопросов, позволяют рекомендовать эту книгу в качестве основы для построения учебного курса по теории и методам цифровой обработки изображений для студентов соответствующих специальностей. Данная книга будет полезна широкому кругу читателей: профессионалам в области анализа и переработки видеоинформации, студентам и аспирантам, специализирующимся в области информатики, а также всем интересующимся вопросами компьютерной обработки и анализа неподвижных изображений и видеопоследовательностей.

Следуя стилю авторов, мы сочли уместным привести список наиболее важных книг, изданных на русском языке, связанных с проблематикой обработки изображений. Данный список помещен в конце главы 1 вслед за аналогичным списком монографий, рекомендованных авторами. К сожалению, в нашей стране в последние годы заметно ощущается недостаток учебной литературы, отражающей как фундаментальные основы, так и изменения, происходящие в области цифровой обработки видеoinформации. Мы надеемся, что книга Р. Гонсалеса и Р. Вудса сможет восполнить возникший в отечественной учебной литературе дефицит книг из области теории и методов обработки изображений.

При переводе книги пришлось столкнуться с рядом трудностей, прежде всего терминологического характера; встречались также и не отмеченные авторами опечатки и неточности в формулировках. В русскоязычном издании устранены многие из них: как те, которые были отмечены авторами, так и те, которые встретились уже при переводе. Переводчиками была проделана большая работа по проверке значительной части формулировок, алгоритмов и выкладок. Насколько было возможно, переводчики старались сохранить стиль авторов; однако книга написана не строго формально, в результате чего в некоторых местах для достижения большей точности и ясности изложения приходилось несколько отклоняться от текста оригинала. В тех случаях, когда по мнению переводчиков полезным являлось дать некоторые разъяснения, были написаны дополнительные примечания. Нумерация всех примечаний сделана сплошная по каждой из глав: т. е. по умолчанию ссылка указывает на примечание внутри той же самой главы.

Предыдущее издание настоящей книги (в том числе и его перевод на русский язык) вызвало большой интерес читателей. Хочется надеяться, что перевод третьего издания также будет востребован.

Перевод книги выполнен канд. техн. наук Л. И. Рубановым (вводная часть, гл. 1, 2, 9—12) и канд. техн. наук П. А. Чочиа (гл. 3—8). Данный перевод в значительной степени основан на переводе второго издания той же книги. В нем, кроме двух указанных переводчиков, принимал участие канд. физ.-мат. наук Д. В. Сушко (гл. 4—7), который кроме того оказал существенную помощь при переводе гл. 5 настоящего издания.

П. А. Чочиа

ГЛАВА I

ВВЕДЕНИЕ

Лучше один раз увидеть, чем сто раз услышать.

Пословица

Интерес к методам цифровой обработки изображений происходит из двух основных областей ее применения, которыми являются повышение качества изображений для улучшения его визуального восприятия человеком и обработка изображений для их хранения, передачи и представления в автономных системах машинного зрения. Эта глава преследует несколько целей: (1) определить границы области, называемой «обработка изображений»; (2) нарисовать историческую перспективу развития этой области; (3) дать представление о современном состоянии предмета, рассмотрев несколько важнейших областей, где применяется обработка изображений; (4) кратко обсудить основные подходы, используемые в цифровой обработке изображений; (5) дать общее представление о компонентах типичной многоцелевой системы обработки изображений и (6) указать книги и периодические издания, в которых публикуются работы по тематике обработки изображений.

1.1. Что такое цифровая обработка изображений?

Изображение можно определить как двумерную функцию $f(x, y)$, где x и y — координаты в *пространстве* (конкретно на плоскости) и значение f которой в любой точке, задаваемой парой координат (x, y) , называется *интенсивностью* или *уровнем серого*¹ изображения в этой точке. Если величины x , y и f принимают конечное число дискретных значений, то говорят о *цифровом изображении*. *Цифровой обработкой изображений* называется обработка цифровых изображений с помощью цифровых вычислительных машин (компьютеров). Заметим, что цифровое изображение состоит из конечного числа элементов, каждый из которых расположен в конкретном месте и принимает определенное значение. Эти элементы называются *элементами изображения* или *пикселями*². Чаще всего для элементов цифрового изображения используется термин «пиксель». Более строгие формальные определения будут даны в главе 2.

¹ В отечественной литературе для этого понятия широко распространен термин *яркость*, который мы и будем использовать во всех случаях, когда это не приводит к недоразумениям. — *Прим. перев.*

² Слово *pixel* образовано от английского словосочетания *picture element* — «элемент изображения». В англоязычной литературе также широко употребляется сокращенное написание *pel*. — *Прим. перев.*



Зрение является наиболее совершенным из наших органов чувств, поэтому неудивительно, что зрительные образы играют важнейшую роль в человеческом восприятии. Однако, в отличие от людей, способных воспринимать электромагнитное излучение лишь в видимом диапазоне, машинная обработка изображений охватывает практически весь электромагнитный спектр от гамма-излучения до радиоволн. Обработываемые изображения могут порождаться такими источниками, которые для человека непривычно связывать с наблюдаемыми изображениями. Таковы, например, ультразвуковые изображения, изображения, получаемые в электронной микроскопии или генерируемые компьютером. Таким образом, цифровая обработка изображений охватывает широкие и разнообразные области применения.

Не существует общепринятой точки зрения, где заканчивается обработка изображений и начинаются другие смежные области, например анализ изображений и машинное зрение. Иногда разграничение делается здесь по тому принципу, что обработка изображений определяется как дисциплина, в которой на входе и на выходе процесса присутствуют изображения. Мы убеждены, что такое вычленение является ограниченным и отчасти искусственным. Например, в рамках этого определения даже такая тривиальная задача, как определение средней интенсивности по полю изображения (при решении которой ищется единственное число), не может рассматриваться как операция обработки изображения. С другой стороны, существуют такие области, как машинное зрение, где конечной целью является компьютерная имитация человеческого зрения, включая обучение, способность к умозаключениям и действиям на основе наблюдаемой информации. Эта область сама по себе образует лишь одно из направлений искусственного интеллекта, целью которого является имитация интеллектуальной деятельности человека. Искусственный интеллект находится еще на начальной стадии своего развития, причем следует отметить, что прогресс идет значительно медленнее, чем первоначально ожидалось. Область, связанная с анализом изображений (иначе называемая «понимание» или «интерпретация» изображений), занимает промежуточное положение между обработкой изображений и машинным зрением.

Во всем диапазоне от обработки изображений до машинного зрения нет четких границ, тем не менее можно различать в нем компьютеризованные процессы низкого, среднего и высокого уровня. Процессы низкого уровня касаются только примитивных операций типа предобработки с целью уменьшения шума, повышения контраста или улучшения резкости изображений. Для низкоуровневых процессов характерен тот факт, что на входе и на выходе присутствуют изображения. Обработка изображений на среднем уровне охватывает такие задачи, как сегментация (разделение изображения на области или выделение на нем объектов), описание объектов и сжатие их в удобную для компьютерной обработки форму, а также классификация (распознавание) отдельных объектов. Для процессов среднего уровня характерно наличие изображений только на входе, на выход же поступают признаки и атрибуты, извлекаемые из этих изображений (например границы областей, линии контуров, отличительные признаки конкретных объектов). Наконец, высокоуровневая обработка включает в себя «осмысление» набора распознанных объектов, как это делается в анализе

изображений, и в пределе осуществление познавательных функций, которые принято связывать со зрением.

Исходя из вышеприведенного замечания мы видим, что естественным этапом перехода от обработки изображений к их анализу выступает распознавание отдельных областей или объектов на изображении. Таким образом, то, что в этой книге называется «цифровой обработкой изображений», включает процессы с изображениями на входе и на выходе, а также процессы извлечения признаков из изображений, вплоть до (или включая) распознавания индивидуальных объектов. В качестве иллюстрации, разъясняющей указанные понятия, рассмотрим область автоматизированного анализа печатного или рукописного текста. В сферу цифровой обработки изображений, которая рассматривается в данной книге, входят процессы получения изображения области, содержащей этот текст, предварительной обработки полученного изображения, выделения (сегментации) отдельных символов текста, описания символов в подходящей для компьютерной обработки форме и, наконец, распознавания этих символов. Что касается осмысления содержимого страницы, то оно может быть отнесено уже к сфере анализа изображений или даже машинного зрения, в зависимости от уровня сложности, который подразумевается за словом «осмысление». Как мы вскоре увидим, определенная таким образом цифровая обработка изображений успешно применяется в широком круге областей, важных с социально-экономической точки зрения. Концепции, развиваемые в последующих главах, служат основой для методов, которые используются в этих прикладных областях.

1.2. Истоки цифровой обработки изображений

Одно из первых применений цифровых изображений было опробовано в газетном деле для передачи иллюстраций по трансокеанскому подводному кабелю между Лондоном и Нью-Йорком. В начале 1920-х годов была внедрена система «Бартлейн» для передачи изображений по кабелю, что позволило уменьшить время доставки иллюстраций через Атлантику с обычной недельной задержки до менее чем трех часов. С помощью специального печатающего оборудования осуществлялось кодирование исходного изображения для передачи по кабелю и последующее восстановление этого изображения на приемной стороне. На рис. 1.1 показано изображение, переданное таким образом и распечатанное затем на телеграфном буквопечатающем аппарате со специальным шрифтом, имитирующим различные уровни почернения.

Ряд начальных проблем, связанных с улучшением визуального качества этих первых цифровых изображений, относился к выбору процедуры печати и распределению уровней интенсивности. Способ печати, использованный для получения рис. 1.1, в конце 1921 г. был отвергнут в пользу технологии фотографической репродукции с использованием перфоленты, формируемой телеграфным аппаратом на приемной стороне линии. На рис. 1.2 показано изображение, полученное таким способом. Улучшения по сравнению с рис. 1.1 очевидны как в отношении качества передачи полутонов, так и в пространственном разрешении.



Рис. 1.1. Цифровое изображение, полученное в 1921 г. с кодовой ленты на телеграфном аппарате с особым шрифтом [McFarlane³]

Ранние системы «Бартлейн» были способны кодировать изображения с помощью пяти градаций яркости. В 1929 г. эти возможности были увеличены до 15 градаций. На рис. 1.3 приведено типичное изображение, которое могло быть получено с помощью оборудования, использующего 15 градаций яркости. За этот период были внедрены системы, в которых фотопластинка экспонировалась с помощью световых лучей, модулируемых закодированной на перфоленте информацией, что позволило значительно улучшить процесс репродукции.

Хотя в вышеприведенных примерах фигурируют цифровые изображения, их нельзя рассматривать как результат цифровой обработки изображений в контексте нашего определения, коль скоро при их получении не использовались компьютеры. Таким образом, история цифровой обработки изображений тесно связана с развитием цифровой вычислительной техники. В самом деле, для цифровых изображений требуется такая большая память и вычислительная мощность, что прогресс в области цифровой обработки изображений в значительной степени определяется развитием компьютеров и вспомогательных технологий для хранения, отображения и передачи данных.

Идея компьютера восходит к абаку, изобретенному на полуострове Малая Азия более 5000 лет назад. Ближе к нашим дням, в последние два столетия делались продвижения, заложившие основу для появления компьютеров. Однако принципы того, что мы называем *современным* цифровым компьютером, были

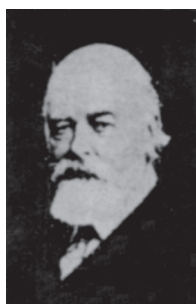


Рис. 1.2. Цифровое изображение, полученное в 1922 г. с использованием перфоленты после прохождения сигнала через Атлантику дважды [McFarlane]

³ Список литературы в конце книги расположен в алфавитном порядке по фамилии первого автора.



Рис. 1.3. Неретушированная фотография генералов Першинга и Фоша, переданная в 1929 г. по кабелю из Лондона в Нью-Йорк с помощью 15-градационного оборудования [McFarlane]

заложены всего лишь в 1940-х годах, когда Джон фон Нейман ввел в рассмотрение два ключевых понятия: (1) равноправное хранение в памяти и данных, и программ и (2) условный переход в программе. Эти две идеи заложены в фундаменте центрального процессора, который является сердцем современных компьютеров. Вслед за фон Нейманом был целый ряд ключевых продвижений, которые привели к появлению компьютеров, достаточно мощных для использования при цифровой обработке изображений. Кратко перечислим эти достижения: (1) изобретение транзистора в компании Bell Laboratories в 1948 г.; (2) изобретение в 1950-х и 1960-х гг. языков программирования высокого уровня КОБОЛ (COBOL, Common Business-Oriented Language — общий язык программирования, ориентированный на бизнес) и ФОРТРАН (FORTRAN, Formula Translator — «транслятор формул»); (3) изобретение интегральной микросхемы компанией Texas Instruments в 1958 г.; (4) разработка операционных систем в начале 1960-х гг.; (5) выход на рынок персонального компьютера IBM в 1981 г. и (7) последовательная миниатюризация электронных компонентов, начиная с появления больших интегральных схем (БИС) в конце 1970-х гг., затем сверхбольших интегральных схем (СБИС) в 1980-х гг., вплоть до сегодняшних ультрабольших интегральных схем (УБИС). Одновременно с перечисленными продвижениями шло развитие в области запоминающих устройств и систем отображения, наличие которых необходимо для цифровой обработки изображений.

Первые компьютеры с мощностью, достаточной для выполнения осмысленных задач цифровой обработки изображений, появились в начале 1960-х гг. Рождение того, что мы сегодня называем цифровой обработкой изображений, прослеживается с момента возникновения таких машин и появления программ изучения космоса. Параллельный прогресс в этих двух областях привел в действие мощный потенциал идей цифровой обработки изображений. Работы по использованию вычислительной техники для улучшения визуального качества изображений, получаемых с помощью беспилотных космических аппаратов, были развернуты в Лаборатории реактивного движения в Пасадене, шт. Калифорния, в 1964 г., когда переданные космическим аппаратом «Рейн-

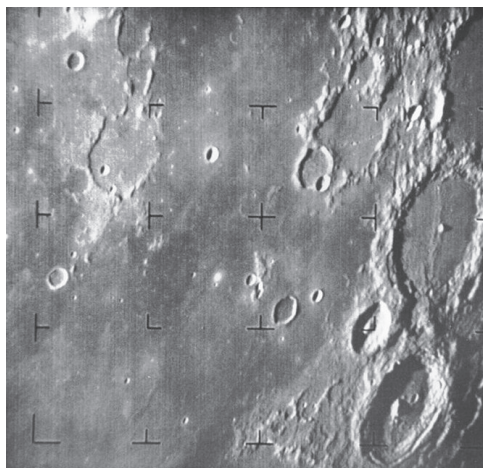


Рис. 1.4. Изображение лунной поверхности, переданное космическим аппаратом «Рейнджер-7» 31 июля 1964 г. (Снимок предоставлен агентством NASA)

жер-7» изображения лунной поверхности были подвергнуты компьютерной обработке для исправления различных искажений, обусловленных конструкцией бортовой телевизионной камеры. На рис. 1.4 приведено изображение Луны, полученное «Рейнджером-7» 31 июля 1964 г. в 9 час. 9 мин. восточно-американского времени (EDT), приблизительно за 17 мин. до его удара о лунную поверхность (на фотографии видны метки так называемой *ризо-маркировки*, применяемой для коррекции геометрических искажений; эта техника обсуждается подробнее в главе 2). Это было первое изображение Луны, полученное американским космическим аппаратом⁴. Опыт, накопленный при обработке первых космических изображений, послужил основой для разработки усовершенствованных методов восстановления и улучшения изображений. Эти методы позже применялись при обработке изображений, полученных в ходе полетов к Луне космических аппаратов «Сервейер», полетов аппаратов «Маринер» по пролетной траектории вблизи Марса, пилотируемых полетов космических кораблей «Аполлон» на Луну и т. д.

Параллельно с космическими исследованиями в конце 1960-х — начале 1970-х гг. методы цифровой обработки изображений начали применяться в медицине, дистанционном исследовании земных ресурсов, астрономии. В начале 1970-х гг. была изобретена рентгеновская вычислительная томография, кратко называемая также компьютерной томографией (КТ), что стало важнейшим событием в области применения обработки изображений для медицинской диагностики. При компьютерной томографии набор детекторов излучения и рентгеновский источник располагаются на кольце, внутрь которого помещается исследуемый объект (т. е. пациент), и кольцо вращается вокруг объекта. Проходящее через объект рентгеновское излучение улавливается детекторами,

⁴ Напомним, что первые космические снимки лунной поверхности (в том числе обратной стороны Луны) были сделаны советской автоматической межпланетной станцией «Луна 3» в 1959 г. — *Прим. перев.*

находящимися на противоположной стороне кольца; этот процесс повторяется по мере вращения источника. Томография состоит из алгоритмов, которые на основе использования данных от детекторов строят изображения «сечения» объекта в плоскости кольца. При движении объекта вдоль оси кольца создается набор таких сечений, которые в совокупности образуют трехмерное представление внутреннего строения объекта. Томографию независимо друг от друга предложили сэр Годфри Н. Хаунсфилд и проф. Ален М. Кормак, которые в 1979 г. были удостоены за это изобретение Нобелевской премии по медицине. Интересно отметить, что рентгеновские лучи были открыты в 1895 г. Вильгельмом Конрадом Рентгеном, получившим за это Нобелевскую премию по физике в 1901 г. Эти два открытия, которые разделяет почти 100 лет, привели к одному из наиболее важных сегодняшних применений цифровой обработки изображений.

С 1960-х гг. до настоящего времени область применения обработки изображений значительно расширилась. Помимо медицинских и космических приложений, методы цифровой обработки изображений сегодня используются в широком круге областей. Компьютеризированные процедуры применяются для облегчения восприятия рентгеновских и иных изображений в промышленности, медицине и биологии путем повышения контраста или цветового кодирования различных уровней интенсивности (представления изображений в псевдоцветах). Аналогичные методы применяются в географии для изучения картины загрязнений окружающей среды по данным аэрофотосъемки и космическим снимкам. Методы улучшения и восстановления изображений применяются при обработке некачественных изображений утраченных объектов или трудновоспроизводимых экспериментальных результатов. Например, в археологии с помощью методов цифровой обработки изображений удалось по имеющимся нечетким фотографиям успешно восстановить первоначальный вид раритетов, которые со времени съемки были утрачены или повреждены. В физике и смежных областях компьютерная обработка является обычным способом улучшения качества изображений, получаемых в ходе экспериментов, как, например, в электронной микроскопии или физике высокотемпературной плазмы. Аналогичные примеры успешного применения технологий обработки изображений можно найти в астрономии, биологии, медицинской радиологии, промышленности, в оборонной и правоохранительной сфере.

Приведенные примеры относились к случаям, когда результаты обработки предназначены для восприятия человеком. Другая крупная область применения методов обработки изображений, упоминавшаяся в начале этой главы, — это решение задач, связанных с машинным восприятием изображений. В этом случае интерес вызывают процедуры, извлекающие из изображения некоторую информацию и представляющие ее в форме, подходящей для компьютерной обработки. Часто эта информация весьма мало похожа на визуальные признаки, используемые людьми при интерпретации содержимого изображения. Примерами информации такого рода, часто применяемой при машинном восприятии изображений, могут быть статистические моменты, коэффициенты преобразования Фурье, значения многомерного расстояния и т. д. Типичными задачами машинного восприятия, в которых интенсивно используются методы обработки изображений, являются автоматическое распознавание символов, системы

машинного зрения для автоматизации сборки и контроля продукции, задачи опознавания «свой—чужой» для военных объектов, автоматическая обработка отпечатков пальцев, проверка анализов крови и результатов рентгеновских исследований, компьютерная обработка аэрофотоснимков и спутниковых изображений с целью прогнозирования погоды и экологического мониторинга. Продолжающееся уменьшение соотношения цена/производительность современных компьютеров, рост пропускной способности сетей телекоммуникаций и развитие Интернета создали беспрецедентные возможности для дальнейшего расширения сферы цифровой обработки изображений. Некоторые из прикладных областей рассматриваются в следующем разделе.

1.3. Примеры областей применения цифровой обработки изображений

Сегодня в технике нет почти ни одной области, которую в той или иной мере не затрагивала бы цифровая обработка изображений. Наше обсуждение коснется лишь сравнительно небольшого числа таких прикладных областей, однако, несмотря на вынужденную ограниченность материала, у читателя не должно остаться сомнений в широте и важности применения обработки изображений. В этом разделе демонстрируются многочисленные прикладные области, в каждой из которых повседневно используются методы цифровой обработки изображений, описываемые в последующих главах. Многие из приведенных здесь изображений впоследствии упоминаются во встречающихся в книге примерах. Все эти изображения являются цифровыми.

Области применения цифровой обработки изображений столь разнообразны, что попытка охватить их во всей широте требует какой-то системы изложения. Один из простейших способов — классификация прикладных областей в соответствии с видами источников, формирующих соответствующие изображения (например оптические, рентгеновские и т. д.). Главным источником энергии для формирования применяемых сегодня изображений является электромагнитное излучение. Среди других важных энергетических источников, которые могут создавать изображения, упомянем акустические и ультразвуковые (механические) колебания, а также электронные пучки, применяемые в электронной микроскопии. Кроме того, имеется целый класс синтетических (искусственных) изображений, которые синтезируются компьютерными программами и используются для моделирования и визуализации. В этом разделе мы кратко обсудим, как регистрируются изображения этих многочисленных категорий и каковы области их применения. Методы преобразования изображений в цифровую форму будут рассмотрены в следующей главе.

Наиболее привычны изображения, создаваемые электромагнитным излучением, особенно в видимом спектре или в рентгеновском диапазоне. Электромагнитные волны можно трактовать как распространяющиеся синусоидальные колебания определенной частоты, а можно — как поток частиц, движущихся со скоростью света. Каждая такая частица обладает определенной энергией, но нулевой массой и называется *квантом излучения (фотоном)*. Если располо-

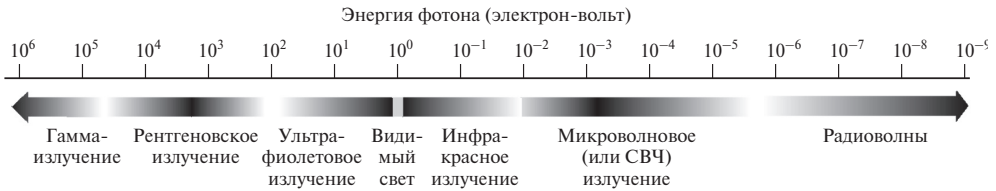


Рис. 1.5. Спектр электромагнитного излучения в порядке убывания энергии фотона

жить диапазоны излучения в порядке убывания энергии фотона, то получим изображенный на рис. 1.5 спектр, простирающийся от гамма-лучей (обладающих максимальной энергией) до радиоволн. Плавное изменение окраски интервалов диапазона на этом рисунке призвано подчеркнуть тот факт, что спектр электромагнитного излучения не разграничен строго, а имеет смысл говорить скорее о плавном переходе одного участка диапазона в другой.

1.3.1. Формирование изображений с помощью гамма-лучей

Изображения, полученные с помощью гамма-излучения, используются главным образом в медицинской радиологии и астрономических наблюдениях. В медицинской радиологии применяется подход, при котором пациенту вводится радиоактивный изотоп, распад которого сопровождается гамма-излучением. Это излучение регистрируется детекторами гамма-излучения, сигналы которых и используются для формирования изображения. На рис. 1.6(а) приведен полный снимок скелета, полученный с помощью гамма-лучей описанным образом. Изображения такого вида используются для обнаружения участков различных патологий костей, в частности при инфекционных или онкологических заболеваниях. Рис. 1.6(б) демонстрирует другой важный вид медицинских радиологических изображений, получаемых методом позитронной эмиссионной томографии (ПЭТ). Используется тот же принцип, что и при рентгеновской томографии, кратко описанный в разделе 1.2, однако вместо использования внешнего источника рентгеновского излучения пациент принимает радиоактивный изотоп, распад которого сопровождается позитронным излучением. При встрече позитрона с электроном они аннигилируют с выделением двух гамма-квантов. Это гамма-излучение регистрируется, и формируется томографическое изображение в соответствии с основными принципами томографии. Приведенное на рис. 1.6(б) изображение представляет собой один кадр из последовательности изображений, которые в совокупности дают трехмерное представление тела пациента. На этом кадре хорошо видны небольшие белые скопления — опухоли в мозге и в легком пациента.

Около 15 тыс. лет назад в созвездии Лебедя произошел взрыв сверхновой, что привело к образованию расширяющегося облака газа сверхвысокой температуры, которое получило название Петли Лебедя. Столкновение этого облака с окружающими газовыми облаками порождает излучение широкого спектра, которое в видимом диапазоне дает эффектную цветовую картину. На рис. 1.6(в) приведено изображение Петли Лебедя в диапазоне гамма-излучения. В отличие от примеров, показанных на рис. 1.6(а, б), это

<i>a</i>	<i>б</i>
<i>в</i>	<i>г</i>

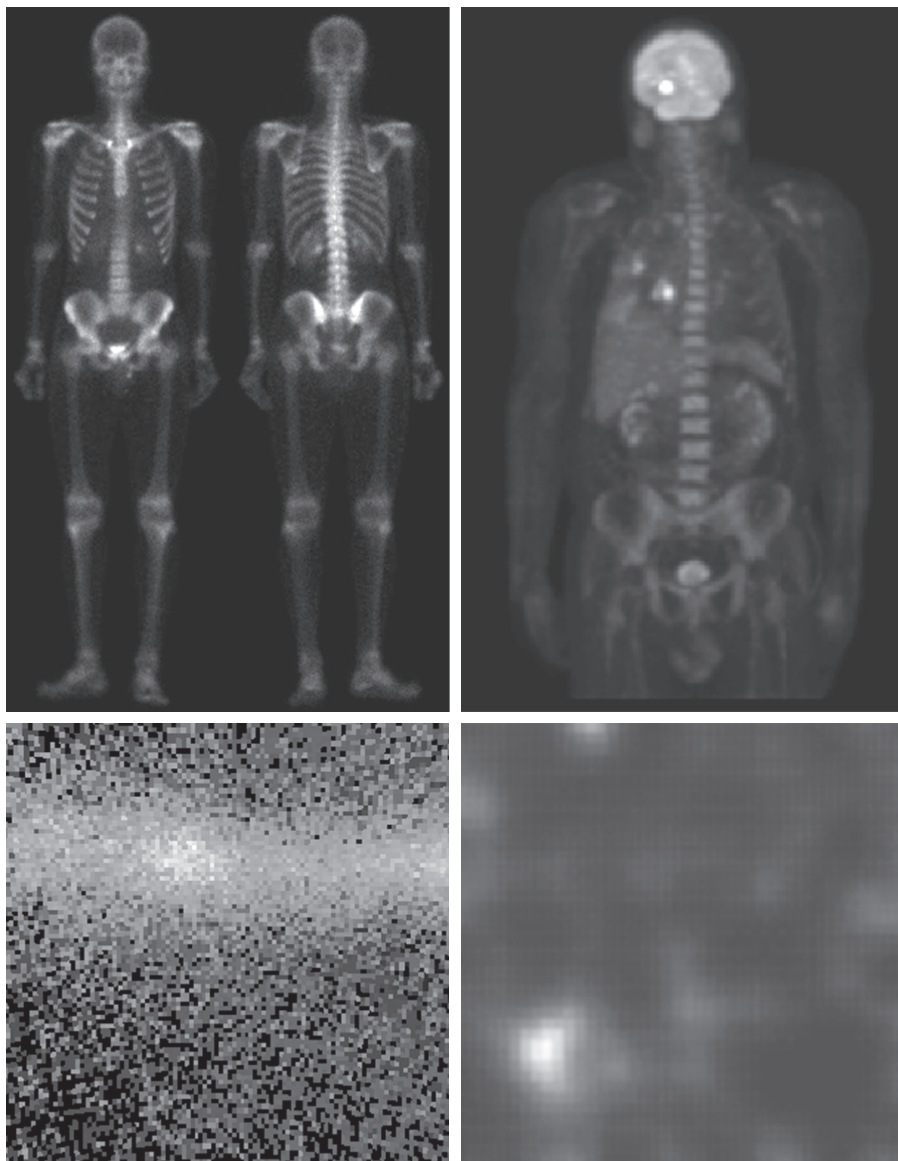


Рис. 1.6. Примеры изображений, полученных с помощью гамма-лучей. (а) Снимок скелета. (б) ПЭТ-изображение. (в) Петля Лебеда. (г) Гамма-излучение из клапана реактора (яркое пятно). (Изображения предоставили: (а) компания G. E. Medical Systems; (б) д-р Майкл Е. Кейси, компания STI PET Systems; (в) профессора Жонг Хи и Дэвид К. Уи, университет шт. Мичиган)

изображение было получено с использованием естественного излучения изображаемого объекта. Наконец, на рис. 1.6(г) демонстрируется изображение гамма-излучения из клапана ядерного реактора. В левой нижней части изображения видна область сильной радиации.

1.3.2. Рентгеновские изображения

Рентгеновские лучи — один из самых старых источников электромагнитного излучения, используемых для получения изображений. Хорошо известно применение рентгеновских лучей для медицинской диагностики, однако они также широко используются в промышленности и других областях, в частности астрономии. Рентгеновское излучение для формирования изображений в медицине и промышленности генерируется с помощью рентгеновской трубки — вакуумного прибора с катодом и анодом. Катод находится в нагретом состоянии, вследствие чего испускает свободные электроны, которые с высокой скоростью летят к положительно заряженному аноду. При соударении электронов с ядрами атомов материала анода энергия выделяется в форме рентгеновского излучения. Энергия рентгеновских лучей (часто называемая «жесткостью»), определяющая их проникающую способность, регулируется изменением приложенного к аноду напряжения, а интенсивность излучения (количество рентгеновских лучей) регулируется изменением тока, проходящего через нить накала катода. На рис. 1.7(а) показан хорошо знакомый рентгеновский снимок грудной клетки, получаемый при помещении пациента между рентгеновской трубкой и чувствительной к рентгеновскому излучению пленкой. При прохождении рентгеновских лучей через тело пациента их интенсивность изменяется в зависимости от степени поглощения (рассеяния), и окончательный уровень энергии фиксируется на рентгеновской пленке, экспонируя ее почти так же, как лучи света формируют изображение на фотопленке. В цифровой рентгенографии применяются два способа получения цифровых изображений: (1) дискретизация (оцифровка) обычных рентгеновских пленок или (2) непосредственная регистрация прошедших через тело пациента рентгеновских лучей устройством, преобразующим рентгеновское излучение в световое (например с помощью фосфоресцирующего экрана). Полученный световой сигнал затем считывается с помощью цифровой системы, работающей в оптическом диапазоне. Вопросы дискретизации изображений подробнее рассматриваются в главах 2 и 4.

Другое важное применение рентгеновских изображений — ангиография, которая является одним из видов контрастной рентгенографии. Эта процедура используется для получения изображений кровеносных сосудов; такие изображения называются *ангиограммами*. В артерию или вену в паховой области вводится катетер (тонкая гибкая трубка), который продвигается вдоль сосуда, пока не достигнет обследуемой зоны. Затем через катетер впрыскивается контрастное вещество, хорошо поглощающее рентгеновские лучи. Благодаря этому усиливается контраст рентгеновского изображения кровеносных сосудов, что позволяет врачу-радиологу видеть аномалии кровоснабжения или места закупорки сосудов. На рис. 1.7(б) демонстрируется ангиограмма аорты, на которой виден катетер, введенный через крупный кровеносный сосуд внизу слева. Обратим внимание на высокий контраст изображения крупного сосуда в направлении потока контрастного вещества к почкам, которые тоже видны на изображении. Как описано в главе 2, в ангиографии широко используется цифровая обработка изображений, в частности вычитание изображений с целью дальнейшего повышения контраста исследуемых кровеносных сосудов.

<i>a</i>	<i>з</i>
<i>б</i>	
<i>в</i>	<i>д</i>

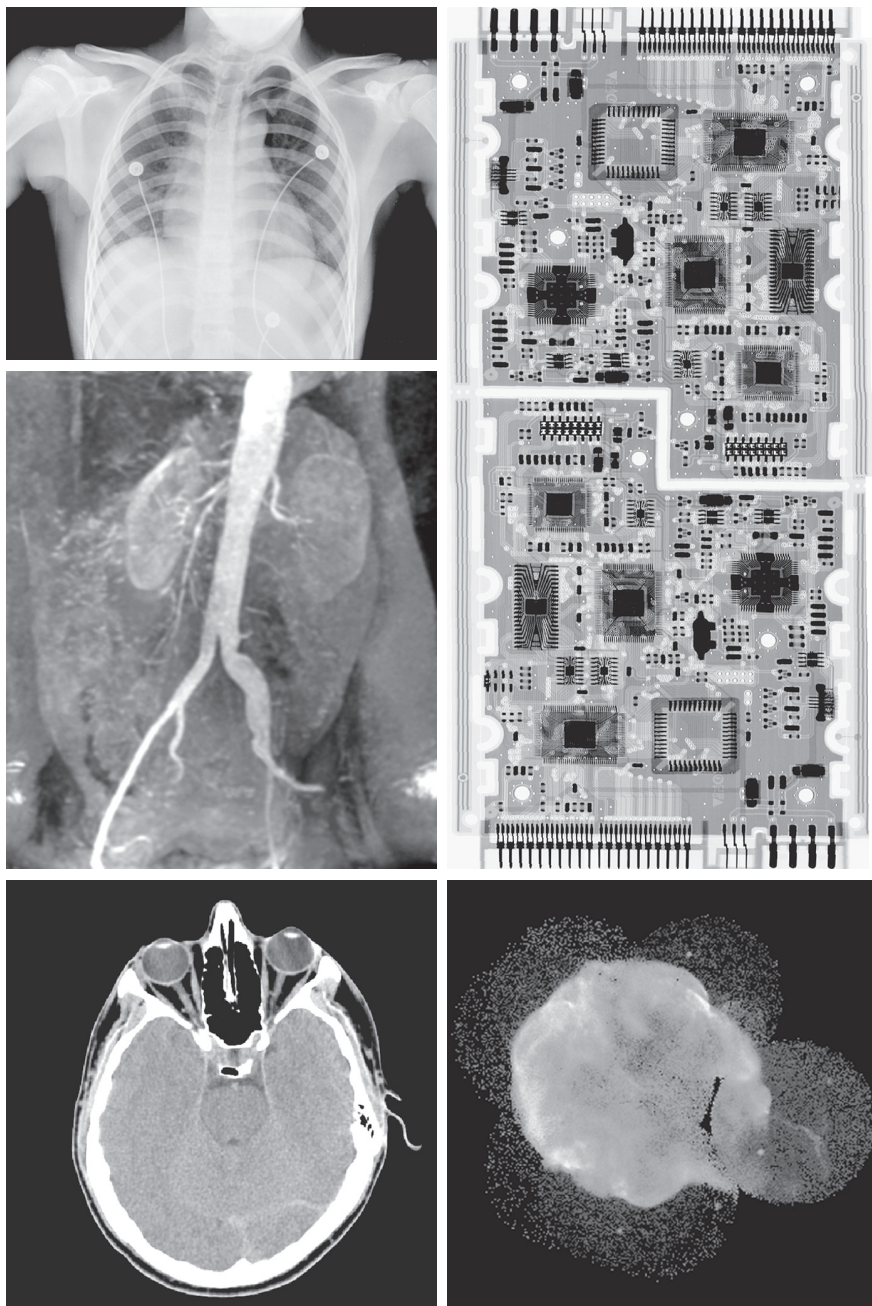


Рис. 1.7. Примеры рентгеновских изображений. (а) Рентгенограмма грудной клетки. (б) Ангиограмма аорты. (в) Компьютерная томограмма головы. (з) Печатные платы. (д) Петля Лебедя. (Изображения предоставили: (а, в) д-р Дэвид Р. Пикенс, департамент радиологии медицинского центра университета Вандербилта; (б) д-р Томас Р. Гест, отделение анатомии медицинской школы университета шт. Мичиган; (з) Джозеф Е. Пассенте, компания Lixi, Inc.; (д) агентство NASA)

Еще одно важное применение рентгеновских лучей для формирования изображений в медицине — это компьютерная томография. Благодаря высокому разрешению и возможности трехмерного представления, компьютерная томография с момента своего первого появления в начале 1970-х гг. произвела революцию в медицине. Как отмечалось в разделе 1.2, каждое КТ-изображение передает поперечный срез тела пациента. При продольном перемещении пациента формируется множество таких срезов, которые в совокупности образуют трехмерное представление внутреннего строения тела с продольным разрешением, пропорциональным количеству срезов. На рис. 1.7(б) показано изображение типичного среза, получаемого при компьютерной томографии головы.

Аналогичная технология используется и в промышленном производстве, хотя там обычно применяется рентгеновское излучение с большей энергией. На рис. 1.7(в) приведено рентгеновское изображение печатной платы радиоэлектронного прибора. Подобные изображения, будучи лишь одним из сотен возможных промышленных применений рентгеновских изображений, используются для контроля печатных плат на наличие дефектов, таких как отсутствие деталей или разрывы контактных дорожек. Промышленная компьютерная томография применима, когда детали проницаемы для рентгеновских лучей, что очевидно в случае пластмассовых узлов, но возможна даже и при контроле крупных изделий вроде твердотопливных реактивных двигателей. На рис. 1.7(д) показан пример применения рентгеновских изображений в астрономии. Здесь изображена та же Петля Лебедя, что и на рис. 1.6(б), но на этот раз в диапазоне рентгеновского излучения.

1.3.3. Изображения в ультрафиолетовом диапазоне

Ультрафиолетовый «свет» находит разнообразные применения, в частности в литографии, производственном контроле, микроскопии, лазерной технике, биологических и астрономических наблюдениях. Мы проиллюстрируем использование изображений ультрафиолетового диапазона на примерах из области микроскопии и астрономии.

Ультрафиолетовое освещение используется во флуоресцентной микроскопии — одном из наиболее быстро развивающихся направлений микроскопии. Явление флуоресценции было открыто в середине XIX в., когда впервые было замечено, что минерал флюорит (плавиковый шпат) излучает свет при направлении на него ультрафиолетового излучения. Сами по себе ультрафиолетовые лучи невидимы, но при столкновении фотона ультрафиолетового излучения с электроном атома флуоресцентного материала электрон переходит на более высокий энергетический уровень. Последующее возвращение возбужденного электрона на нижний уровень сопровождается излучением фотона с меньшей энергией, что соответствует видимому (ближе к красному) диапазону спектра. Принцип работы флуоресцентного микроскопа заключается в облучении подготовленного препарата ярким активизирующим освещением и последующем выделении значительно более слабого флуоресцентного свечения. Таким образом, глаз наблюдателя или другой детектор будет воспринимать только вторичное излучение. Свечение флуоресцирующих участков должно наблюдаться на темном фоне, чтобы обеспечивался достаточный для их обнаружения кон-

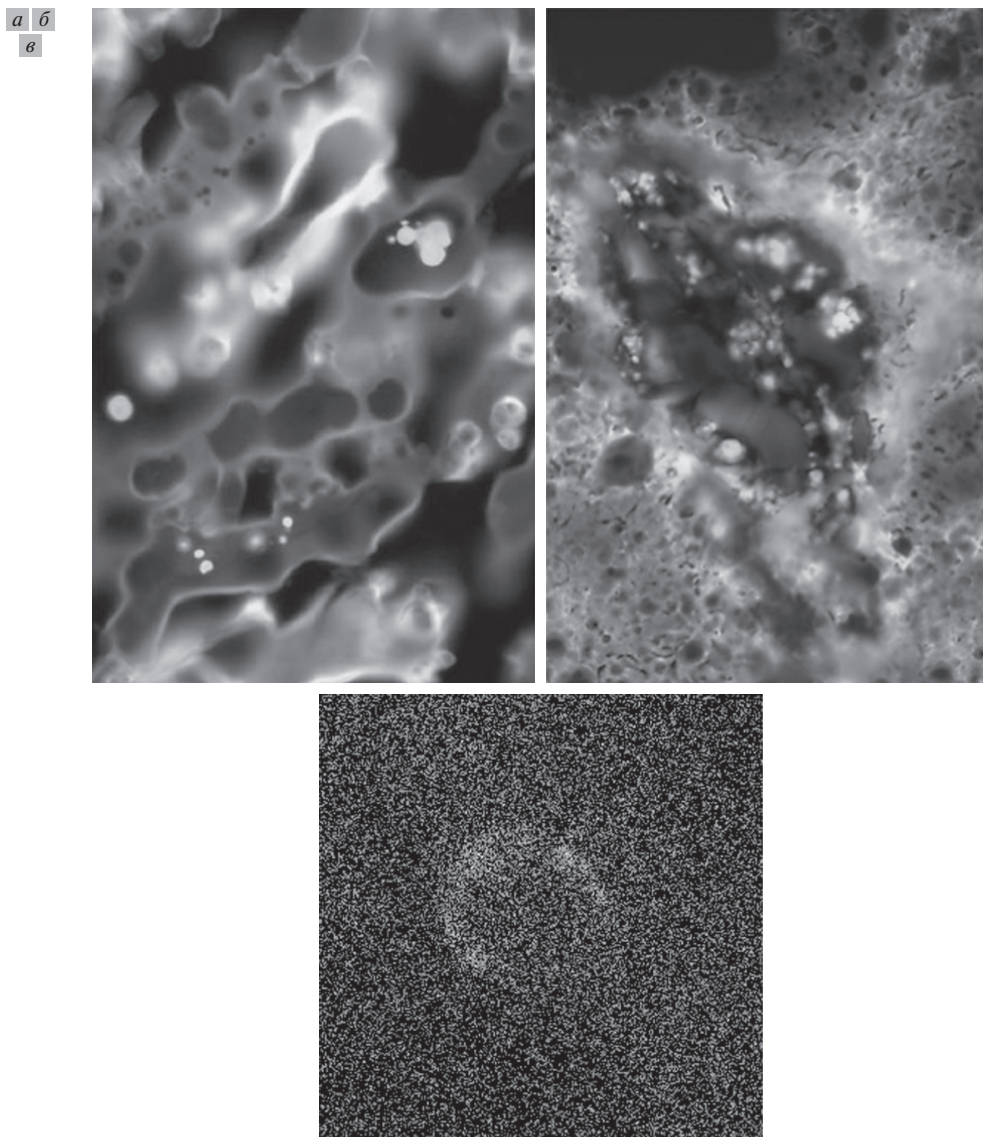


Рис. 1.8. Примеры изображений в ультрафиолетовом диапазоне. (а) Нормальное зерно. (б) Зерно, зараженное головней. (в) Петля Лебеда. (Изображения предоставили: (а, б) д-р Майкл У. Дэвидсон, университет шт. Флорида; (в) агентство NASA)

траст. Чем темнее фон, изготовленный из нефлуоресцирующего материала, тем выше эффективность прибора.

Флуоресцентная микроскопия — прекрасный метод исследования материалов, обладающих флуоресцирующими свойствами, либо в естественной форме (первичная флуоресценция), либо в результате обработки флуоресцирующими химикатами (вторичная флуоресценция). Рис. 1.8(а, б) демонстрируют типичные возможности флуоресцентной микроскопии. На рис. 1.8(а) по-

казан полученный с помощью флуоресцентного микроскопа снимок здорового зерна, а на рис. 1.8(б) — снимок зерна, зараженного головней — заболеванием зерновых и бобовых культур, трав, а также луковичных растений, вызываемым более чем 700 видами паразитических грибов. Особенно опасна головня злаковых культур, поскольку зерновые — один из важнейших источников пищи людей. В качестве иллюстрации другой области применения на рис. 1.8(в) показано изображение Петли Лебеда в высокочастотной полосе ультрафиолетового диапазона.

1.3.4. Изображения в видимом и инфракрасном диапазонах

Учитывая, что видимый диапазон электромагнитного спектра для нас наиболее привычен, неудивительно, что область использования изображений этого диапазона оказывается намного шире, чем всех остальных вместе взятых. Инфракрасные изображения часто используются совместно с видимыми, поэтому для иллюстрации мы объединили оба эти диапазона в одном разделе. В нижеприведенном обсуждении в качестве примеров областей использования будут обсуждаться: световая микроскопия, астрономия, дистанционное зондирование, промышленность и правоохранительная деятельность.

На рис. 1.9 показано несколько примеров изображений, полученных с помощью оптического микроскопа. Диапазон примеров простирается от фармацевтики и микроскопических методов производственного контроля до определения характеристик материалов. Даже в пределах одной микроскопии множество возможных прикладных областей слишком обширно, чтобы его детально описать. Несложно вообразить варианты обработки, которые могут потребоваться в применении к таким изображениям, от улучшения их визуального качества до проведения различных измерений.

Еще одна важная область обработки видимых изображений — дистанционное зондирование земной поверхности, охватывающее обычно несколько зон в видимом и инфракрасном диапазонах спектра. В табл. 1.1 перечислены такие тематические зоны, в которых осуществляет зондирование спутник LANDSAT,

Таблица 1.1. Тематические зоны американского спутника LANDSAT

№	Наименование	Длины волн (мкм)	Характеристики и назначение
1	Видимый синий цвет	0,45—0,52	Максимальное проникание воды
2	Видимый зеленый цвет	0,52—0,60	Измерение плотности растительного покрова
3	Видимый красный цвет	0,63—0,69	Различение формы растительности
4	Ближнее ИК-излучение	0,76—0,90	Съемка береговой линии и распределения биомассы
5	Средний ИК-диапазон	1,55—1,75	Содержание влаги в почве и растительности
6	Тепловое ИК-излучение	10,4—12,5	Влажность почвы и температурная карта
7	Средний ИК-диапазон	2,08—2,35	Поиск полезных ископаемых

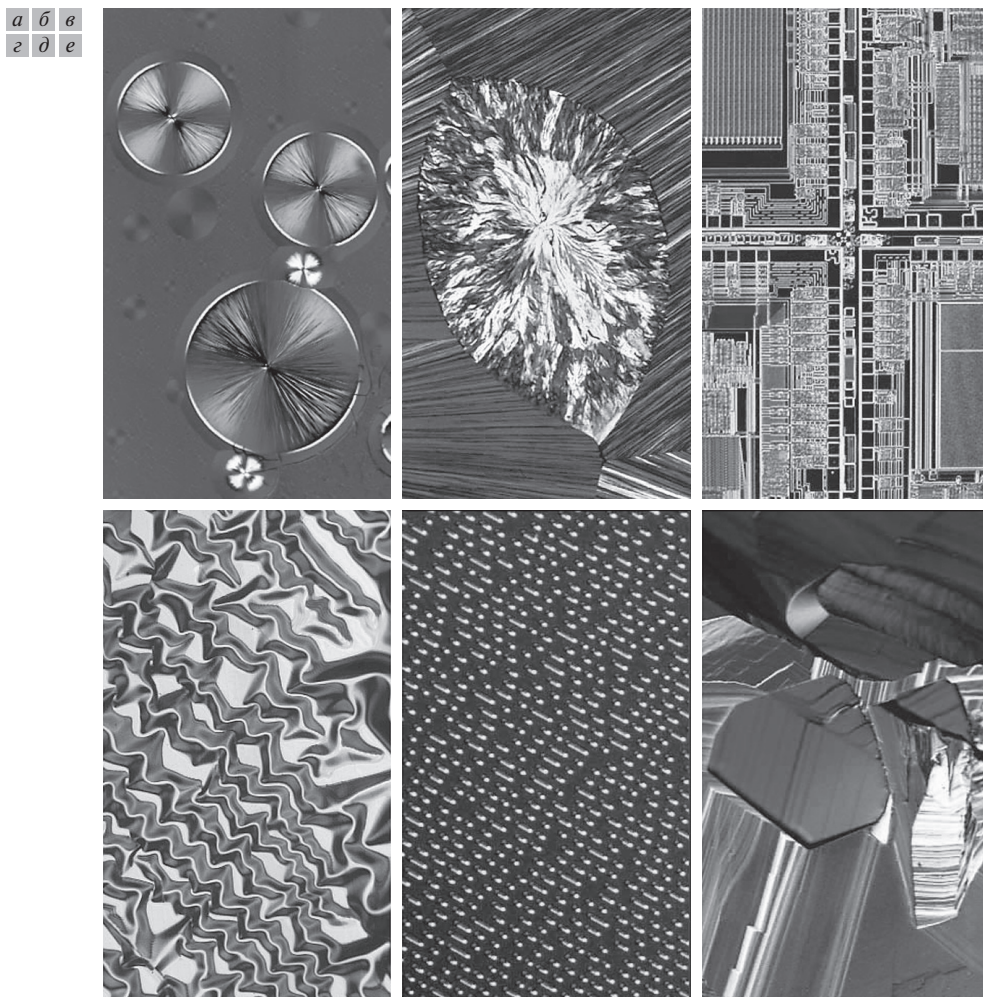


Рис. 1.9. Примеры изображений в оптической микроскопии. (а) Таксол (противораковый препарат), увеличение 250х. (б) Холестерин, увеличение 40х. (в) Микропроцессор, увеличение 60х. (г) Тонкая пленка окиси никеля, увеличение 600х. (д) Поверхность музыкального компакт-диска, увеличение 1750х. (е) Органический сверхпроводящий материал, увеличение 450х. (Изображения предоставил д-р Майкл У. Дэвидсон, университет шт. Флорида)

запущенный агентством NASA. Главная задача LANDSAT состоит в получении и передаче изображений Земли из космоса с целью глобального экологического мониторинга. Интервалы спектра представлены длинами волн в микронах (мкм); напомним, что $1 \text{ мкм} = 10^{-6} \text{ м}$ (длины волн, отвечающих различным диапазонам электромагнитного спектра, подробно обсуждаются в главе 2). Обратим внимание на указанные в таблице характеристики и назначение каждой из тематических зон.

Чтобы получить начальное представление о возможностях таких *многозональных* изображений, взглянем на рис. 1.10, где приведено по одному изображе-

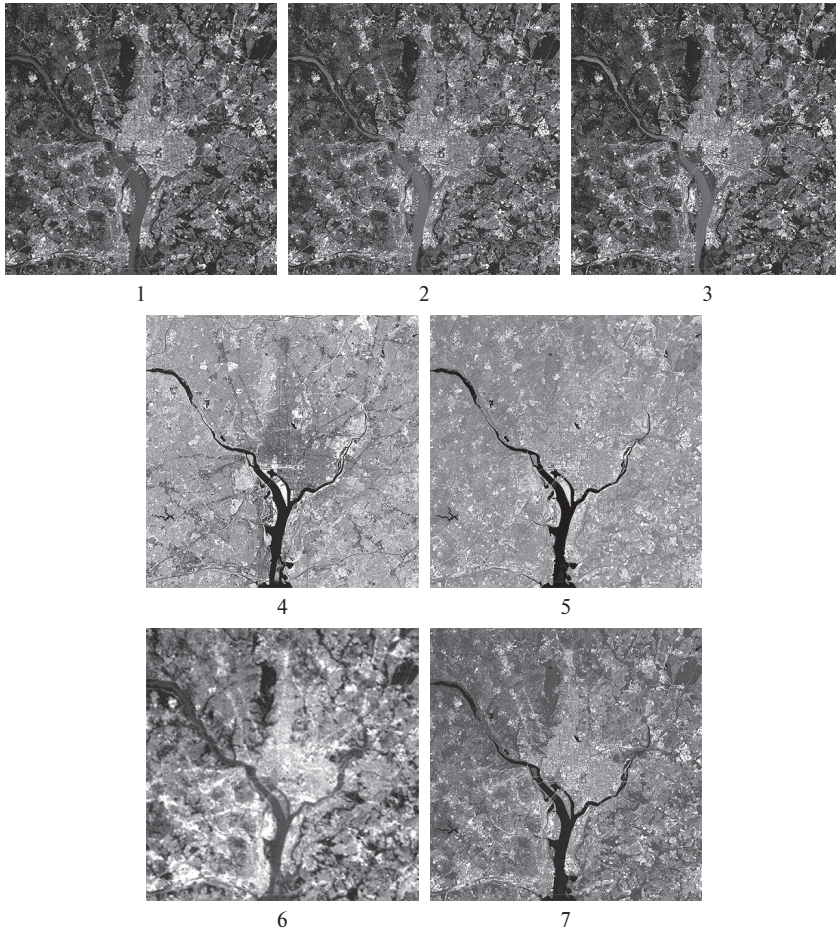


Рис. 1.10. Полученные со спутника LANDSAT изображения г. Вашингтон, округ Колумбия. Номера снимков соответствуют номерам тематических зон в табл. 1.1. (Изображения предоставлены агентством NASA)

нию для каждой из зон таблицы. Изображен район г. Вашингтон, округ Колумбия; видны здания, дороги, участки растительности и протекающая через город крупная река Потомак. Изображения населенных пунктов часто (и давно) используются для оценки роста численности населения, динамики загрязнений и прочих факторов, вредно влияющих на экологию. Примечательна разница между видимым и инфракрасным изображениями на этих снимках. Обратим, например, внимание, насколько хорошо выделяется река на фоне берегов в изображениях, соответствующих 4-й и 5-й зонам.

Наблюдение за погодой и составление прогнозов также является важным применением многозональных спутниковых изображений. Например, на рис. 1.11 приводится изображение урагана Катрина — одного из самых разрушительных ураганов в западном полушарии за последнее время. Оно было получено спутником Национальной океанографической и атмосферной службы США (NOAA) с помощью датчиков, работающих в видимом и инфракрасном

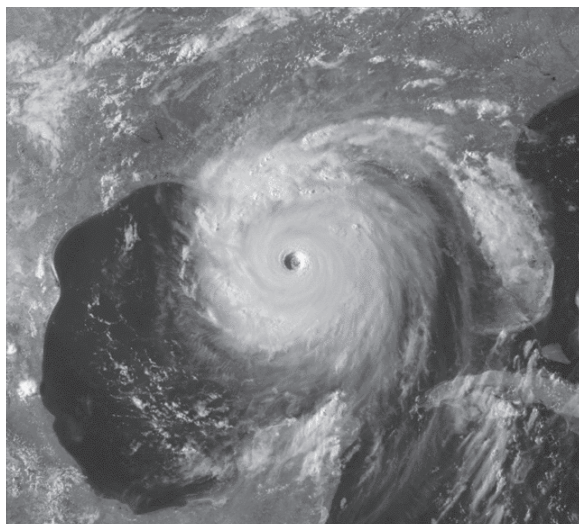


Рис. 1.11. Спутниковое изображение урагана Катрина, полученное 29 августа 2005 г. (Изображение предоставлено службой NOAA)

диапазонах. На этом снимке хорошо виден так называемый «глаз» урагана (соответствует центру циклона).

Рис. 1.12 и 1.13 демонстрируют применение инфракрасных изображений. Эти снимки были получены инфракрасной системой регистрации изображений, установленной на спутнике DMSP, запущенном по оборонно-метеорологической программе службы NOAA, и представляют собой фрагменты обширного набора данных «Ночные огни Земли» — глобального реестра населенных пунктов. Устройство регистрации инфракрасных изображений у этого спутника работает в диапазоне длин волн 10,0—13,4 мкм и обладает уникальной способностью фиксировать находящиеся на земной поверхности слабые источники ближнего инфракрасного излучения, в том числе города, поселки, деревни, газовые факелы и пожары. Даже не будучи специалистом в обработке изображений, легко представить себе компьютерную программу, которая использовала бы такие изображения для оценки относительной доли электроэнергии, потребляемой в различных районах Земли.

Важной областью применения изображений, регистрируемых в видимом диапазоне, является автоматический контроль выпускаемой продукции. На рис. 1.14 даны несколько примеров такого применения. Рис. 1.14(а) демонстрирует плату контроллера дисководов CD-ROM. Для подобных изделий типичной задачей обработки изображений может быть контроль наличия всех компонентов (на данном примере черный квадрат в правой верхней части изображения демонстрирует отсутствие микросхемы). На рис. 1.14(б) показана упаковка таблеток. Здесь задача состоит в компьютерном визуальном контроле отсутствия пустых мест в упаковке. Рис. 1.14(в) иллюстрирует пример обработки изображений для выявления недостаточно заполненных бутылок на производственной линии. На рис. 1.14(г) показана прозрачная пластмассовая деталь с недопустимым содержанием пузырьков воздуха. Обнаружение подобных аномалий составляет важную область промышленного контроля



Рис. 1.12. Инфракрасные спутниковые изображения американского континента. Для наглядности рядом приведена небольшая карта. (Изображение предоставлено службой NOAA)

различных материалов, например тканей и дерева. На рис. 1.14(д) изображен пакет кукурузных хлопьев, проходящий контроль по цвету и наличию брака в виде отдельных подгоревших хлопьев. Наконец, на рис. 1.14(е) показано изображение внутриглазного имплантата (вживляемого в глаз искусственного хрусталика) при специальном структурированном освещении. Эта техноло-



Рис. 1.13. Инфракрасные спутниковые изображения населенных регионов на других континентах. Для наглядности рядом приведена небольшая карта. (Изображение предоставлено службой NOAA)

гия применяется для простоты визуального обнаружения плоских деформаций имплантата. Следы в положениях «1 час» и «5 часов» (по аналогии с часовым циферблатом) — повреждения, оставленные пинцетом, а большинство других мелких пятнышек на изображении созданы пылинками и остатками материала. Цель данного вида контроля состоит в автоматическом обнаружении бракованных или поврежденных имплантатов перед упаковкой готовой продукции.

В качестве заключительной иллюстрации обработки изображений видимого спектра рассмотрим рис. 1.15. На рис. 1.15(а) изображен отпечаток большого пальца. Изображения отпечатков пальцев в массовом порядке подвергаются компьютерной обработке как с целью их улучшения, так и для поиска признаков, помогающих автоматически выбирать из базы данных похожие отпечатки. На рис. 1.15(б) приведено изображение бумажной купюры. Цифровая обработка таких изображений находит применение при автоматическом подсчете наличности и в правоохранительной деятельности для чтения номеров купюр с целью

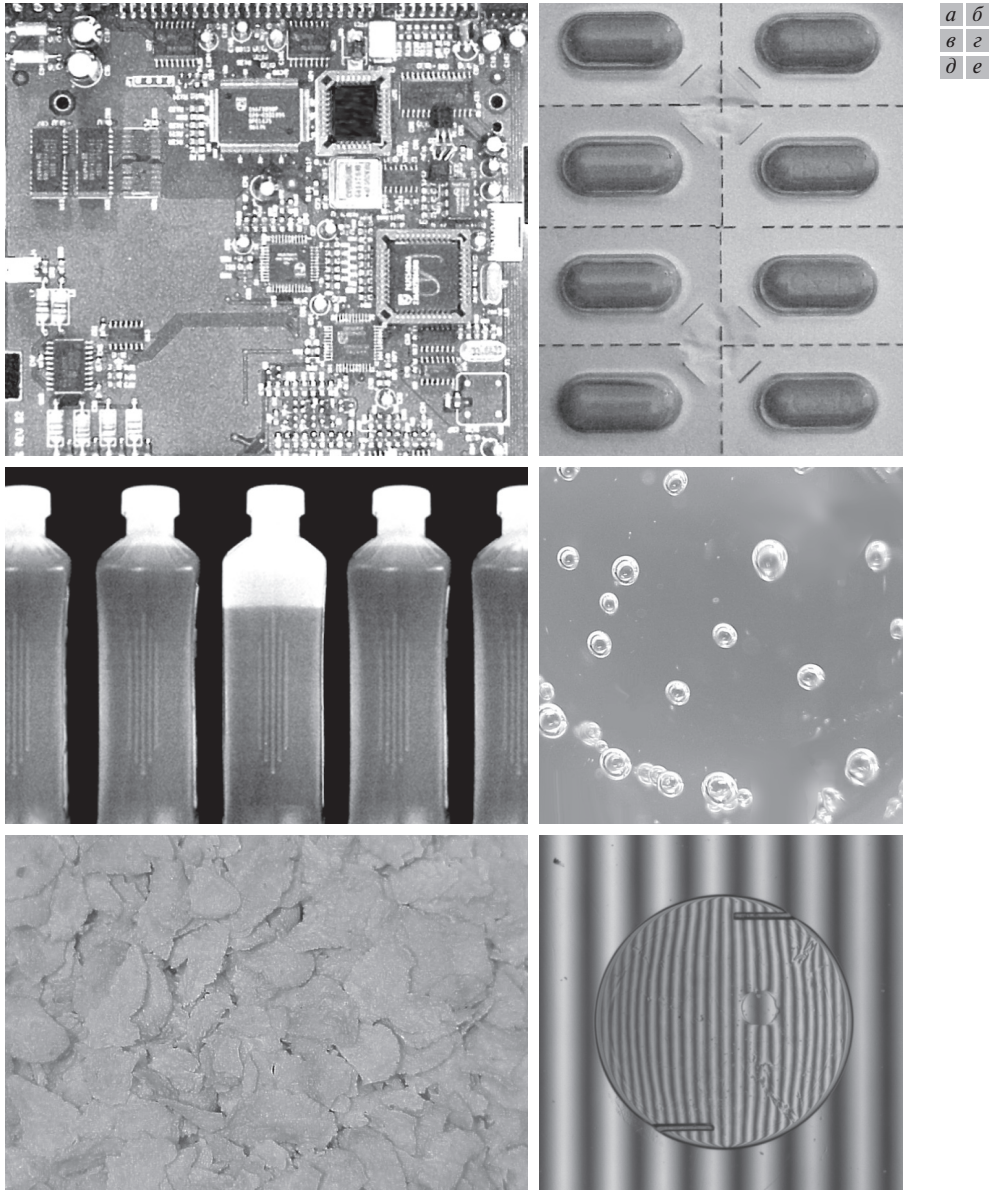


Рис. 1.14. Некоторые примеры промышленной продукции, часто контролируемой с помощью цифровой обработки оптических изображений. (а) Печатная плата контроллера. (б) Упаковка таблеток. (в) Бутылки. (г) Пузырьки воздуха в изделии из прозрачной пластмассы. (д) Кукурузные хлопья. (е) Изображение искусственного хрусталика. (Изображение (е) предоставил Пит Сайтс, Perceptics Corporation)

их прослеживания и идентификации. Два изображения транспортных средств, показанные на рис. 1.15(в, г), являются примерами автоматического чтения регистрационных номеров. Светлые прямоугольники указывают области, в кото-



Рис. 1.15. Некоторые дополнительные примеры обработки изображений видимого спектра. (а) Отпечаток пальца. (б) Бумажная купюра. (в, г) Автоматическое чтение номерных знаков. (Изображения предоставили: (а) Национальный институт стандартов и технологии США; (в, г) д-р Хуан Херрера, Perceptics Corporation)

рых система обработки изображений опознала номер транспортного средства, а в черных прямоугольниках выводятся результаты автоматического распознавания этой системой содержимого номера. Чтение автомобильных номеров и другие применения автоматического распознавания символов широко используются для контроля дорожного движения и надзора правоохранительных органов.

1.3.5. Изображения в микроволновом диапазоне

Изображения микроволнового диапазона применяются главным образом в радиолокации. Уникальным качеством радиолокации является возможность получения изображения любого района независимо от условий освещения и погоды. Микроволновое излучение некоторых диапазонов способно проникать

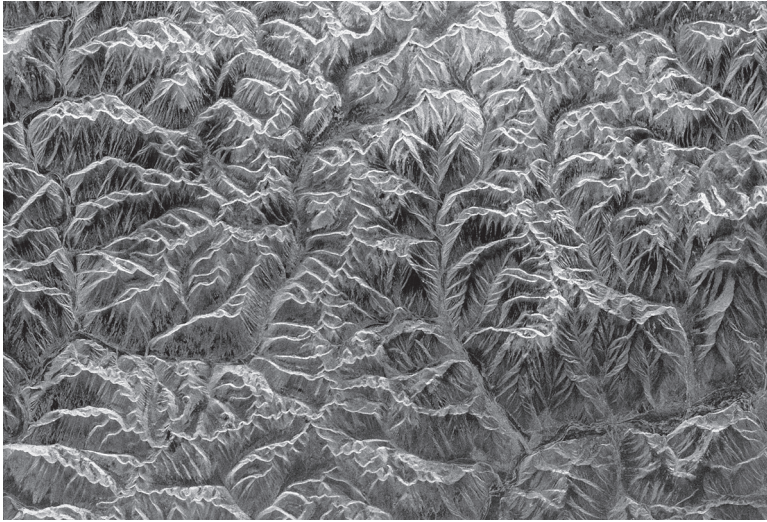


Рис. 1.16. Космическое радиолокационное изображение горного массива на юго-востоке Тибета. (Изображение предоставлено агентством NASA)

даже сквозь облака, растительность, лед и сухой песок. Во многих случаях радиолокация остается единственным способом исследования труднодоступных районов Земли. Применяемый для получения изображения радиолокатор работает аналогично фотоаппарату со вспышкой, в том смысле, что он использует собственный источник освещения (микроволновые импульсы), которое направляется на снимаемый участок поверхности. Роль объектива фотоаппарата в радиолокаторе играет антенна, сигнал от которой проходит через компьютерную систему, осуществляющую регистрацию и обработку изображения. Радиолокационное изображение отображает распределение интенсивностей отраженной энергии микроволнового диапазона, которую уловила антенна локатора.

На рис. 1.16 показано полученное из космоса радиолокационное изображение труднодоступного горного массива в юго-восточном Тибете, приблизительно в 90 км к востоку от г. Лхаса. В правом верхнем углу видна широкая долина реки Лхаса, населенная тибетскими фермерами, разводящими яков; в этой долине расположена деревня Менба. Высота гор в этом районе достигает 5800 м над уровнем моря, а дно долины располагается на высоте около 4300 м. Обратим внимание на четкость изображения и качество воспроизведения деталей, независимо от облаков и других атмосферных неоднородностей, которые обычно мешают получить изображение сходного качества в оптическом диапазоне.

1.3.6. Изображения в диапазоне радиоволн

Как и в случае изображений, получаемых на противоположной стороне электромагнитного спектра (гамма-лучи), основными областями применения изображений в диапазоне радиоволн выступают медицина и астрономия. В меди-

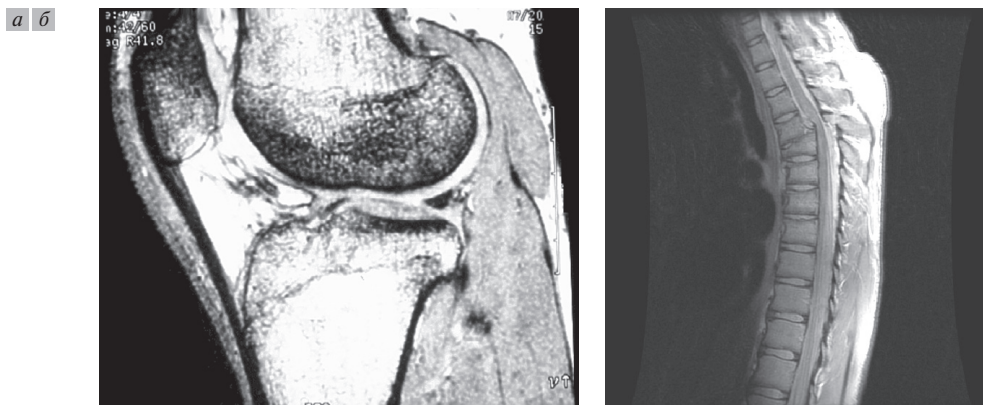


Рис. 1.17. ЯМР-изображения человеческого (а) колена и (б) позвоночника. (Изображения предоставили: (а) д-р Томас Р. Гест, отделение анатомии медицинской школы университета шт. Мичиган; (б) д-р Дэвид Р. Пикенс, департамент радиологии медицинского центра университета Вандербилта)

цине радиоволны используются для получения изображений методом ядерного магнитного резонанса (ЯМР). По этой технологии пациента помещают в сильное магнитное поле, и через его тело пропускают радиоволны в форме коротких импульсов. В ответ на каждый такой импульс ткани тела пациента реагируют, излучая свои радиоволновые сигналы. Место возникновения и сила этих сигналов регистрируются компьютерной системой обработки, генерирующей двумерное изображение среза тела пациента. С помощью ЯМР можно получить срез вдоль любой плоскости. На рис. 1.17 показаны ЯМР-изображения человеческого коленного сустава и позвоночника.

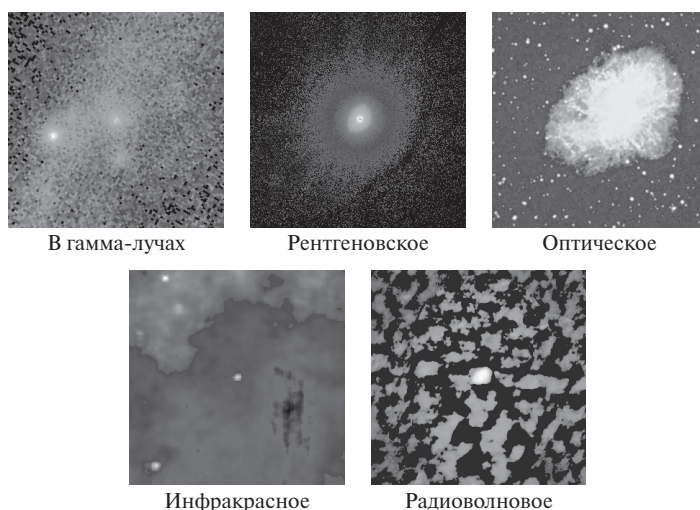


Рис. 1.18. Изображения пульсара Крабовидной туманности (находится в центре снимков) в различных диапазонах электромагнитного спектра. (Изображения предоставлены агентством NASA)

Крайний справа снимок на рис. 1.18 демонстрирует изображение пульсара Крабовидной туманности в диапазоне радиоволн. Интересно сравнить его с приведенными на этом рисунке изображениями того же района, полученными в других обсуждавшихся ранее диапазонах электромагнитного спектра. Заметим, что каждое изображение дает свой, совершенно отличающийся вид этого пульсара.

1.3.7. Примеры, иллюстрирующие другие способы формирования изображений

Хотя чаще всего используются изображения, полученные в электромагнитном спектре, существует ряд других важных способов формирования изображений. В частности, в этом разделе мы обсудим акустические изображения, электронную микроскопию и искусственные изображения, синтезированные с помощью компьютеров.

Построение изображений с помощью звуковых волн находит применение в геологических изысканиях, промышленности и медицине. В геологии используются звуковые колебания с частотами у нижней границы звукового спектра (до сотен герц), а в других областях для получения изображений применяются ультразвуковые колебания с частотами порядка мегагерц (миллионов герц). Наиболее важные коммерческие применения обработки изображений в геологии касаются поисков нефти и других полезных ископаемых. При формировании изображения, несущего информацию о земных недрах, один из основных методов состоит в использовании тяжелого грузовика и большой плоской стальной платформы. Грузовик давит на землю через платформу и одновременно является источником вибраций в спектре частот до 100 Гц. Мощность и скорость распространения отраженных звуковых волн определяются геологическим составом грунта под поверхностью. В результате компьютерного анализа этих звуковых колебаний строится цифровое изображение.

В морской геологии для получения изображений в звуковом диапазоне обычно используют источник энергии в виде пары пневмопушек, буксируемых позади судна. Отраженные звуковые волны детектируются гидрофонами, помещенными внутрь кабелей, которые либо также буксируются судами, либо укладываются на океанское дно или вертикально подвешиваются к буям. Пневмопушки поочередно выстреливают, создавая импульс давления порядка 150 атмосфер. Картина отраженных звуковых волн в совокупности с постоянным движением судна (что дает продольную составляющую) используется для формирования трехмерной карты состава земной коры под дном океана.

На рис. 1.19 изображено поперечное сечение известной трехмерной модели, на которой проверяются характеристики алгоритмов построения изображений по данным сейсморазведки. Стрелка указывает на углеводородный пласт (место залегания нефти и/или газа). Указываемое место выглядит ярче окружающих пластов, поскольку изменения плотности в этом месте оказываются выше. Анализ изображений при сейсморазведке состоит в поиске подобных «ярких пятен», соответствующих вероятным нефтяным и газовым месторождениям. Вышележащие пласты также выделяются своей яркостью, однако в этих случаях изменения яркости в поперечном направлении не столь сильны. Многие

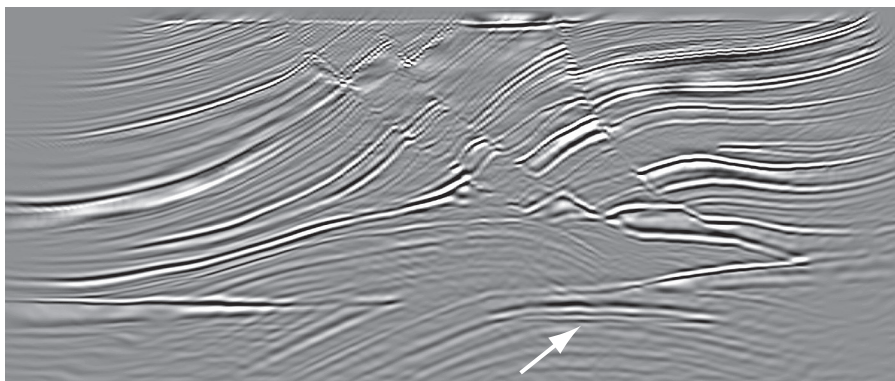


Рис. 1.19. Поперечное сечение модели данных сейсморазведки. Стрелкой указан пласт углеводородов (место залегания нефти и/или газа). (Изображение предоставил д-р Кертис С. Обер, Sandia National Laboratories)

алгоритмы реконструкции данных сейсморазведки испытывают трудности при обнаружении отмеченной области именно из-за ошибок, возникающих в вышележащих областях.

Хотя ультразвуковые изображения широко используются в промышленности, наиболее известно применение этой технологии в медицине, особенно в акушерстве, где изображения еще не рожденных детей изучаются на предмет отсутствия аномалий их развития. Дополнительным результатом такого исследования является определение пола будущего ребенка. Ультразвуковые изображения формируются следующим образом:

1. Ультразвуковая система (состоящая из компьютера, ультразвукового зонда с излучателем и приемником и дисплея) передает в тело ультразвуковые импульсы высокой частоты (от 1 до 5 МГц).
2. Звуковые волны проходят сквозь тело пациента, и на границах между тканями (например между жидкостью и мягкой тканью, мягкой тканью и костью скелета) происходит частичное отражение. Часть звуковых волн отражается обратно в сторону зонда, часть волн затухает, а остальные распространяются дальше, пока не достигнут следующей границы раздела и снова частично отразятся и т. д.
3. Отраженные волны улавливаются приемником зонда и передаются в компьютер.
4. Исходя из времени прихода каждого эхо-сигнала и известной скорости звука в тканях (1500 м/с), компьютер вычисляет расстояние от зонда до соответствующей границы ткани или внутреннего органа.
5. Вычисленные расстояния и интенсивности принятых отраженных сигналов выводятся на дисплее в виде двумерного изображения.

В типичных ультразвуковых исследованиях каждую секунду генерируются и принимаются миллионы звуковых импульсов и эхо-сигналов. Зонд можно двигать вдоль поверхности тела и наклонять, получая изображения в различных проекциях. На рис. 1.20 приводятся несколько примеров таких изображений.

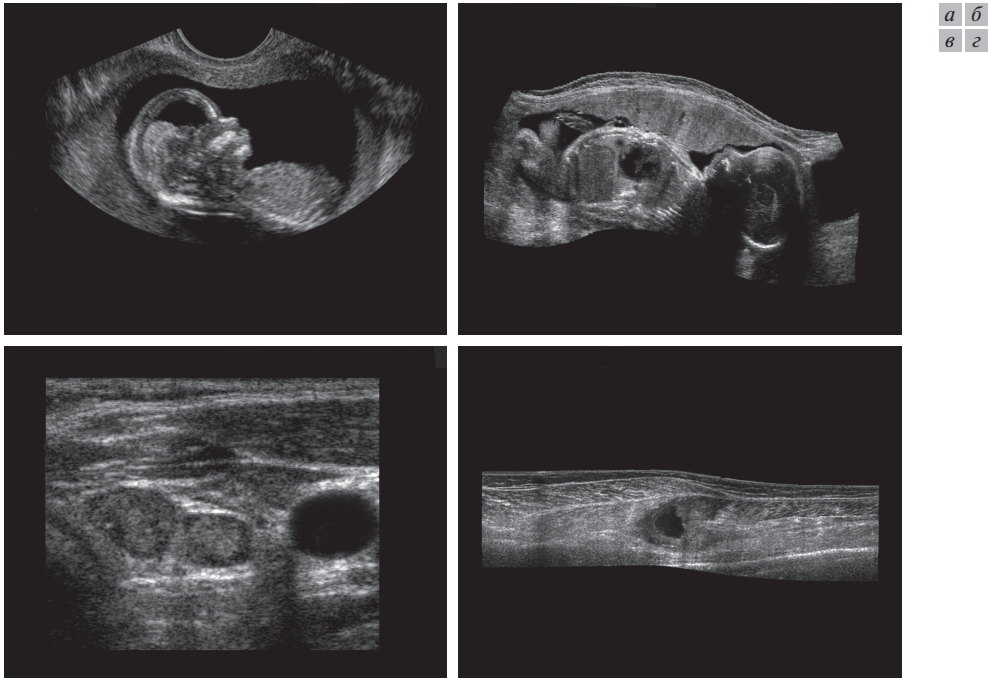


Рис. 1.20. Примеры ультразвуковых изображений. (а) Ребенок. (б) Ребенок в другой проекции. (в) Щитовидная железа. (з) Мышечные слои с заметным повреждением. (Изображения предоставила группа по ультразвуку компании Siemens Medical Systems, Inc.)

Мы продолжим обсуждение других способов получения изображений на примерах из электронной микроскопии. Электронный микроскоп действует аналогично оптическому с той разницей, что вместо световых лучей для получения изображения исследуемого объекта применяется сфокусированный пучок электронов. Работа электронного микроскопа складывается из следующих основных шагов. Источник испускает поток электронов, которые, благодаря приложенному положительному напряжению, движутся с ускорением в направлении исследуемого образца. С помощью металлических щелевых диафрагм и магнитных линз этот поток ограничивается и фокусируется, образуя тонкий пучок, сфокусированный на образце. Внутри облучаемого образца происходит взаимодействие, оказывающее влияние на прохождение пучка электронов, эффект от которого обнаруживается и преобразуется в изображение аналогично тому, как свет отражается или поглощается объектами наблюдаемой сцены. Перечисленные основные шаги имеют место во всех электронных микроскопах независимо от их типа.

Принцип работы *просвечивающего электронного микроскопа* (ПЭМ) во многом аналогичен проектору слайдов. Проектор направляет (передает) пучок света на слайд; при проходе этого пучка сквозь слайд свет модулируется содержимым слайда. После этого пучок света проецируется на экран, формируя увеличенное изображение слайда. ПЭМ работает точно так же, за исключением того, что сквозь образец, играющий роль слайда, направляется пучок электронов. Часть

пучка, прошедшая сквозь исследуемый образец, проецируется на экран из фосфоресцирующего материала. Взаимодействие электронов с этим материалом приводит к появлению света и, следовательно, видимого изображения.

Сканирующий (растровый) электронный микроскоп (СЭМ или РЭМ) осуществляет действительное сканирование образца электронным пучком и запись отраженного результата взаимодействия электронного пучка с каждой точкой поверхности образца. Полное изображение формируется путем растрового сканирования образца электронным пучком аналогично телевизионной развертке. Отраженные электроны при попадании на фосфоресцирующий экран создают на нем видимое изображение. СЭМ лучше подходят для объемных образцов, тогда как для ПЭМ необходим очень тонкий образец.

Электронные микроскопы способны дать очень большое увеличение. Если в оптической микроскопии кратность увеличения ограничена приблизительно $1000\times$, то в электронной микроскопии достигается увеличение $10000\times$ и более. На рис. 1.21 показаны полученные с помощью СЭМ два изображения образцов с дефектами, возникшими вследствие температурных перегрузок.

Завершим обсуждение различных способов формирования изображений, кратко рассмотрев изображения, полученные не от какого-то физического объекта или явления, а сгенерированные компьютером. Примечательным примером изображений, синтезированных с помощью компьютеров, являются *фракталы* [Lu, 1997]. По существу, фракталы представляют собой не что иное, как повторяющееся воспроизведение некоторого исходного образа по определенным математическим правилам. Например, *мозаика* из квадратных элементов является одним из простейших способов генерации фрактальных изображений. Квадрат можно разделить на четыре квадратных подобласти, каждую из которых, в свою очередь, можно разбить снова на четыре еще более мелких квадрата и так далее. В зависимости от сложности правил заполнения каждой

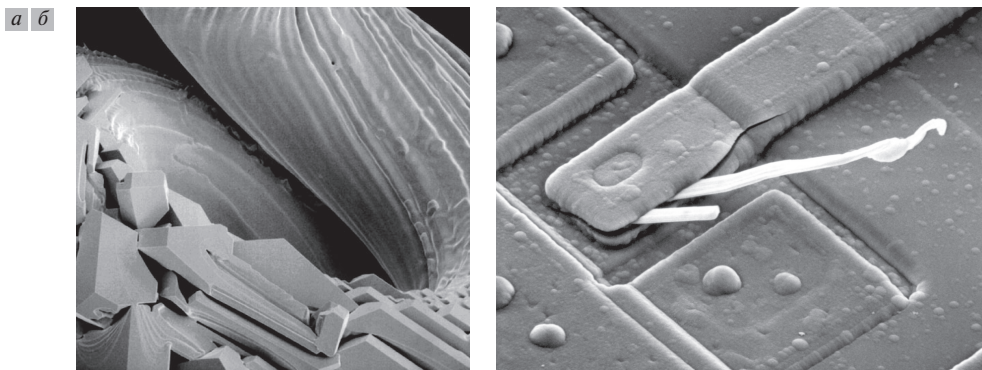


Рис. 1.21. (а) Изображение поврежденной при перегреве вольфрамовой нити накаливания, полученное на СЭМ при увеличении $250\times$ (обратите внимание на осколки в левом нижнем углу). (б) СЭМ-изображение вышедшей из строя интегральной микросхемы при увеличении $2500\times$. Белые нити — это полосы окалины, возникшие при тепловом разрушении. (Изображения предоставили: (а) Майкл Шаффер, факультет геологии университета шт. Орегон, г. Юджин; (б) д-р Джон М. Хьюдек, университет Мак-Мастер, г. Гамильтон, шт. Онтарио, Канада)

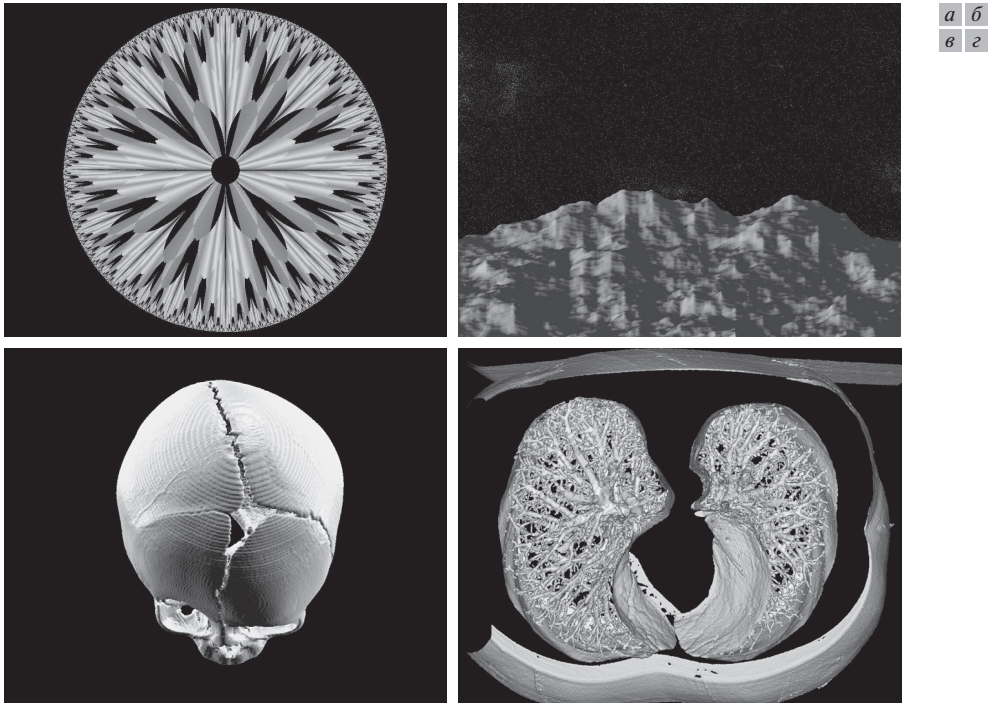


Рис. 1.22. (а, б) Фрактальные изображения. (в, з) Изображения, построенные по трехмерным компьютерным моделям показанных объектов. (Изображения предоставили: (а, б) Мелисса Д. Байнд, Swarthmore College; (в, з) агентство NASA)

квадратной подобласти, таким способом могут быть получены некоторые красивые мозаичные изображения. Разумеется, геометрия может быть произвольной. Например, фрактальное изображение может вырастать из одной центральной точки — такой пример изображен на рис. 1.22(а). Рис. 1.22(б) демонстрирует другое фрактальное изображение («лунный ландшафт»), являющееся интересной аналогией некоторых космических снимков, приведенных в качестве иллюстрации в предыдущих разделах.

Фрактальные изображения, при всей их художественности, несут в себе элементы математического описания процесса «выращивания» изображения из его более мелких элементов в соответствии с некоторыми правилами. Иногда они находят применение в качестве случайных текстур. Более систематический подход к генерации изображений с помощью компьютеров состоит в построении трехмерных моделей объектов. Именно эта важная область, находящаяся на стыке обработки изображений и компьютерной графики, является основой для построения множества систем трехмерной визуализации (например авиационных тренажеров). Рис. 1.22(в, з) демонстрируют примеры изображений, синтезированных с помощью компьютерных программ. Поскольку объекты моделировались как трехмерные, с помощью плоской проекции трехмерного объема можно построить изображения этих объектов в любом ракурсе. Изображения подобного вида используются в медицинском образовании, а также как

основа для решения других прикладных задач, например в криминалистике, судебной медицине или для создания спецэффектов.

1.4. Основные стадии цифровой обработки изображений

Было бы полезно разделить изложенный в последующих главах материал на две большие категории, упомянутые в разделе 1.1: методы, в которых на входе и на выходе имеются изображения, и методы, где на вход поступают изображения, а на выходе возникают признаки и атрибуты, выделенные на основании этих изображений. Такая организация материала книги сведена в схему, изображенную на рис. 1.23. Эта схема не подразумевает, что к изображению применяется каждый из описанных процессов, напротив, целью было донести принципы всех методов обработки, которые могут применяться к изображениям в различных целях и, возможно, с различными получаемыми результатами. Проводимое в данном разделе обсуждение можно рассматривать как краткий обзор материала, представленного в остальной части книги.

Регистрация изображения — первый из процессов, показанных на рис. 1.23. Обсуждение, проведенное в разделе 1.3, дает некоторую информацию относительно возможных источников цифровых изображений, однако значительно более подробно эта тема рассматривается в главе 2, где также вводится ряд базовых понятий, относящихся к цифровым изображениям и используемых далее



Рис. 1.23. Основные стадии цифровой обработки изображений. Внутри блоков указаны главы, в которых рассматривается соответствующий материал

на протяжении всей книги. Заметим, что регистрация изображения может оказаться предельно простой, как в случае, когда исходное изображение уже представлено в цифровой форме. В общем случае стадия регистрации изображения включает некоторую предобработку, например масштабирование.

Улучшение изображения — это процесс манипулирования изображением, в результате которого оно становится более подходящим для конкретного применения, чем оригинал. Здесь важно слово «конкретного», поскольку оно с самого начала устанавливает, что методы улучшения изображений являются проблемно-ориентированными. Так, например, метод, который весьма полезен для улучшения рентгеновских изображений, может оказаться не лучшим подходом для улучшения спутниковых изображений, снятых в инфракрасном диапазоне электромагнитного спектра.

Общей «теории» улучшения изображений не существует. Если изображение обрабатывается с целью визуальной интерпретации, то оценку, насколько хорошо работает конкретный метод, дает в конечном счете наблюдатель. Методы улучшения настолько разнообразны и используют так много различных подходов к обработке изображения, что трудно собрать осмысленную совокупность подходящих для улучшения методов в одной главе, не проводя отдельное обширное исследование. Для новичков в области обработки изображений все связанное с улучшением изображений представляется визуально привлекательным, интересным и относительно простым для понимания. По указанным выше причинам мы используем улучшение изображений в качестве примеров, когда вводим в рассмотрение новые концепции в части главы 2, а также в главах 3 и 4. Материал двух последних глав охватывает многие методы, традиционно применяемые для улучшения изображений. Таким образом, привлекая примеры из области улучшения изображений для представления новых методов обработки изображений уже в этих начальных главах, мы не только избавляемся от необходимости включения в книгу особой главы об улучшении изображений, но, что еще важнее, эффективно вовлекаем начинающих в детали методов обработки изображений на ранней стадии изучения. Однако, как будет видно по мере освоения остальной части книги, введенный в этих главах материал применим для намного более широкого класса задач, чем только лишь улучшение изображений.

Восстановление изображений — это область, также связанная с повышением визуального качества изображения, однако, в отличие от собственно улучшения, критерии которого субъективны, восстановление изображения является объективным в том смысле, что методы восстановления изображений опираются на математические или вероятностные модели искажений изображения. Напротив, улучшение изображений основано на субъективных предпочтениях человеческого восприятия, которые связаны с тем, *что* именно считается «хорошим» результатом улучшения.

Обработка цветных изображений приобрела особую важность в связи со значительным расширением использования цветных изображений в Интернете. В главе 6 излагается ряд фундаментальных понятий, относящихся к цветовым моделям и основным видам цифровых преобразований цветов. Цвет также используется в последующих главах как основа для выделения из изображения некоторых интересующих признаков.

Вейвлеты образуют фундамент для представления изображений в нескольких масштабах одновременно. В частности, этот аппарат используется в книге применительно к сжатию данных изображения, а также для построения пирамидального представления, при котором изображение поэтапно разбивается на все более мелкие фрагменты.

Сжатие, как следует из самого названия, относится к методам уменьшения объема памяти, необходимого для хранения изображения, или сужения полосы пропускания канала, требуемой для его передачи. Хотя техника запоминающих устройств за последнее десятилетие была значительно усовершенствована, этого нельзя сказать в отношении пропускной способности линий связи. Это особенно справедливо по отношению к информации в Интернете, где изобразительная составляющая является существенным элементом содержимого. Со сжатием изображений знакомы (возможно, не отдавая себе в этом отчета) большинство пользователей компьютеров, встречающих в именах графических файлов определенные расширения; например, *jpg* используется в стандарте сжатия изображений, разработанном Объединенной группой экспертов по фотографии (Joint Photographic Experts Group — JPEG).

Морфологическая обработка связана с инструментами для извлечения таких компонентов изображения, которые могут быть полезны для представления и описания формы. Приведенный в этой главе материал дает основы перехода от процессов, имеющих на выходе *изображение*, к процессам, имеющим на выходе *атрибуты изображения*, как это указывалось в разделе 1.1.

Сегментация разделяет изображение на составные части или объекты. В целом автоматическая сегментация принадлежит к числу самых трудных задач цифровой обработки изображений. Излишне подробная сегментация направляет процесс на длинный путь решения задачи, требуя идентификации объектов по отдельности. С другой стороны, недостаточно подробная или же неверная сегментация почти неизбежно приведет к возникновению ошибок на финальной стадии обработки. В общем, чем точнее сегментация, тем больше шансов на успех при распознавании.

Представление и описание почти всегда следуют непосредственно за этапом сегментации, на выходе которого обычно имеются лишь необработанные данные о пикселях, которые либо образуют границу области (т. е. дается множество пикселей, отделяющих одну область изображения от другой), либо представляют все точки самих областей. В обоих случаях необходимо преобразовать данные в форму, пригодную для компьютерной обработки. Первое решение, которое следует принять, — должны ли эти данные представляться в форме только границ областей или областей целиком. Представление в виде границ подходит для тех случаев, когда в центре внимания находятся внешние характеристики формы областей, например углы и изгибы. Представление в виде областей является предпочтительным, если акцент делается на внутренних свойствах объектов, например текстуре или форме скелета. В некоторых приложениях эти представления дополняют друг друга. Выбор способа представления — лишь часть принятия решения по преобразованию «сырых» пиксельных данных в подходящую для дальнейшей компьютерной обработки форму. Должен быть еще указан метод описания данных, при котором бы выдвигались на передний план интересующие признаки. *Построение описания*, иначе называемое *выбором*

признаков, связано с выделением атрибутов, которые бы выражали интересующую информацию в количественном виде или бы могли служить основой для различения классов объектов.

Распознавание представляет собой процесс, который присваивает некоторому объекту идентификатор (например «транспортное средство») на основании его описателей. Как подробно разъяснялось в разделе 1.1, мы считаем, что сфера цифровой обработки изображений заканчивается разработкой методов распознавания отдельных объектов.

До настоящего момента ничего не говорилось о необходимости априорных знаний или, в терминах рис. 1.23, о взаимосвязи между *базой знаний* и модулями обработки. На самом деле знание о проблемной области (т.е. база знаний) некоторым образом закодировано внутри самой системы обработки изображений. Это знание может быть очень простым, как, например, детальное указание участков изображения, где должна находиться интересующая информация, что позволит сузить область ее поиска. База знаний может быть и очень сложной, как, например, взаимосвязанный список всех наиболее вероятных дефектов в задаче контроля материалов либо база данных спутниковых изображений некоторого района с высоким разрешением в прикладных задачах мониторинга. Помимо того, что база знаний руководит работой каждого модуля обработки, она также управляет взаимодействием между модулями. Эта особенность показана на рис. 1.23 с помощью двунаправленных стрелок между обрабатываемыми модулями и базой знаний.

Хотя мы не обсуждаем в этом месте задачу визуализации изображений, важно иметь в виду, что на выходе любой из показанных на рис. 1.23 стадий может выполняться отображение результатов обработки. Отметим также, что не во всех прикладных задачах обработки изображений требуется вся сложность взаимодействия, подразумеваемого рис. 1.23. На самом деле во многих случаях даже не все эти модули необходимы. Например, улучшение изображений для визуальной интерпретации человеком редко нуждается в использовании каких-либо других стадий из числа показанных на рис. 1.23. В общем случае, однако, чем выше сложность задачи обработки изображений, тем большее число процессов требуется привлекать для решения этой задачи.

1.5. Компоненты системы обработки изображений

Еще в середине 1980-х гг. большинство поступающих на рынок образцов систем обработки изображений были относительно крупными периферийными устройствами, которые соединялись со столь же громоздкими компьютерами. Позднее — в конце 1980-х — начале 1990-х гг. — на рынке произошел переход к аппаратуре обработки изображений в виде одиночной платы, конструктивно совместимой с какой-либо из ставших стандартом магистралей и пригодной для установки в индустриальные и персональные компьютеры. Наряду со снижением стоимости такой аппаратуры рынок также стал катализатором для возникновения значительного числа новых компаний, ориентированных на разработку программного обеспечения специально для обработки изображений.



Рис. 1.24. Компоненты универсальной системы обработки изображений

Хотя для массовой обработки изображений большого размера (например спутниковых) еще производятся крупные системы цифровой обработки, общая тенденция направлена в сторону миниатюризации и оснащения обычных малых компьютеров специализированным оборудованием для решения задач обработки изображений. На рис. 1.24 изображены основные компоненты, из которых состоит типичная *универсальная* система цифровой обработки изображений, а ниже обсуждаются функции каждого из ее компонентов начиная с системы регистрации.

Что касается *регистрации*, то для получения цифровых изображений в общем случае необходимы два элемента. Первый из них — это *чувствительный элемент (сенсор)*, т. е. физическое устройство, обладающее чувствительностью к тому виду излучаемой объектом энергии, который мы хотим отобразить. Второй элемент, *аналого-цифровой преобразователь (АЦП)*, представляет собой устройство, преобразующее аналоговый выходной сигнал чувствительного элемента в цифровую форму⁵. Например, в цифровой видеокамере элементы светочувствительной матрицы вырабатывают электрический сигнал, пропорцио-

⁵ Здесь необходимо отметить, что для преобразования непрерывного сигнала в цифровую форму необходимы два, вообще говоря, независимых процесса: *дискретизация*, т. е. пространственное разложение непрерывного сигнала на некоторое число отсчетов, и *квантование* — перевод непрерывного значения сигнала каждого из отсчетов в конечный диапазон значений. — Прим. ред. перев.

нальный силе света. Цифровой преобразователь трансформирует эти сигналы в цифровые данные. Более подробно эта тема рассматривается в главе 2.

Специализированные устройства для обработки изображений обычно включают вышеупомянутый цифровой преобразователь, а также оборудование, с помощью которого выполняются другие элементарные операции, как например, арифметико-логическое устройство (АЛУ), которое позволяет выполнять арифметические и логические операции параллельно для всего изображения. Один из вариантов использования АЛУ — локальное усреднение изображений одновременно с оцифровкой — может быть полезен для снижения уровня шума. Оборудование такого типа иногда называют *подсистемой предобработки* (или *препроцессором*); ее отличительной характеристикой является высокая скорость работы. Иначе говоря, этот блок выполняет функции обработки данных, требующие высокой производительности (например оцифровка и усреднение видеоизображений со скоростью 25 кадров в секунду), с чем не справляется типичный управляющий компьютер системы.

Под *компьютером* в системе обработки изображений подразумевается универсальная ЭВМ в диапазоне от обычного персонального компьютера (ПК) до суперкомпьютера. В специализированных приложениях для достижения требуемой производительности иногда используются компьютеры специальной конструкции, однако мы рассматриваем здесь именно универсальную систему обработки изображений. В таких системах практически любой хорошо оснащенный ПК пригоден для решения задач обработки изображений, не требующих работы в реальном масштабе времени.

Программное обеспечение для обработки изображений состоит из специализированных модулей, выполняющих конкретные операции. В развитых пакетах программ имеются также средства, позволяющие пользователю самостоятельно разрабатывать программы, которые как минимум запускают в работу специализированные модули системы. Более сложные программные пакеты позволяют сочетать вызов этих модулей с обычными операторами какого-либо из универсальных языков программирования.

Наличие *массовой памяти* большого объема обязательно для практических задач обработки изображений. Для хранения изображения размером 1024×1024 пикселя, в котором яркость каждого пикселя представляется 8-битовой величиной, необходим один мегабайт памяти, если не используются средства сжатия изображений. При работе с тысячами или даже миллионами изображений наличие достаточной внешней памяти в системе обработки изображений может оказаться проблематичным. Цифровые запоминающие устройства для задач обработки изображений делятся на три основные категории: (1) временная память для краткосрочного использования в ходе обработки, (2) внешняя память, обладающая относительно коротким временем обращения, и (3) архивная память, для которой характерны редкие обращения. Емкость запоминающих устройств измеряется в байтах (8 бит), килобайтах (тысяча байтов), мегабайтах (миллион байтов), гигабайтах (миллиард байтов) и терабайтах (триллион байтов)⁶.

⁶ Зачастую при выражении емкости памяти в байтах коэффициент увеличения при переходе в каждый следующий класс принимается равным $2^{10} = 1024$, а не $10^3 = 1000$, что, впрочем, здесь несущественно. — *Прим. перев.*

Одним из вариантов реализации временной памяти может быть оперативная память компьютера. Другой вариант состоит в использовании специальных плат, называемых *буферами кадров*, которые хранят одно или более изображений, обеспечивая высокую скорость чтения/записи, обычно соответствующую частоте кадров видеосигнала (например 25 кадр/с). Этот способ позволяет практически мгновенно выполнять *увеличение* изображения либо сдвигать его в вертикальном (*прокрутка*) или горизонтальном (*панорамирование*) направлениях. Буферы кадров обычно расположены в показанном на рисунке блоке «Специализированные устройства обработки изображений». Внешняя память, как правило, представлена накопителями на магнитных или оптических дисках и характеризуется частыми обращениями к хранящейся информации; т. е. важнейшим для нее является фактор быстродействия. Напротив, обращение к архивной памяти за информацией происходит редко, но требуется очень большая емкость памяти. В запоминающих устройствах, используемых в качестве архивной памяти, в качестве носителей информации обычно применяются магнитные ленты и оптические диски внутри многодисковых хранилищ с автоматической установкой и сменой дисков (*jukebox*).

Графические дисплеи, используемые в настоящее время, в основном оснащаются электронно-лучевыми трубками по типу телевизионных, предпочтительно с плоским экраном. Сигнал на монитор подается с платы отображения (видеоадаптера), входящей в состав компьютера. В редких случаях к системе отображения предъявляются такие требования, которым не отвечают встроенные видеоадаптеры современных компьютеров. Иногда необходим стереоскопический режим отображения; это может достигаться с помощью закрепляемой на голове гарнитуры с двумя малогабаритными дисплеями, встроенными в оправу, похожую на защитные очки, в которые и смотрит пользователь.

К числу *устройств получения твердых копий* относятся лазерные и струйные принтеры, устройства термопечати, пленочные фотокамеры и цифровые устройства, например оптические диски⁷. Максимальное разрешение достигается при выводе на пленку, однако для письменных и печатных материалов более естественным носителем является бумага. Для показа в ходе презентаций изображения выводятся на прозрачную пленку или цифровой носитель (если используется подключаемый к компьютеру проектор). Последний вариант постепенно становится общепринятым стандартом презентации изображений.

Соединение с *телекоммуникационной сетью* уже стало почти подразумеваемой функцией в любой сегодняшней компьютерной системе. Учитывая большие объемы данных, связанные с задачами обработки изображений, важнейшим фактором для передачи изображений является пропускная способность сети. В локальных сетях и на выделенных каналах телекоммуникации трудностей обычно не возникает, однако обмен информацией с удаленными пунктами через Интернет далеко не всегда оказывается столь же эффективным. К счастью, в результате развития оптоволоконных сетей и других технологий широкополосной связи положение в этой сфере быстро исправляется.

⁷ Оптические диски, вообще говоря, относятся к устройствам хранения информации. — *Прим. перев.*

Заключение

Главной целью материала этой главы было показать в исторической перспективе истоки возникновения цифровой обработки изображений и, что более важно, нынешние и будущие возможности применения этой технологии. Хотя эти вопросы невозможно исчерпывающе рассмотреть в пределах одной главы, у читателя должно остаться ясное ощущение широты области практических применений цифровой обработки изображений. Поскольку последующие главы книги посвящены теории и приложениям обработки изображений, мы привели большое число примеров, чтобы более четко представлять перспективы и пользу от применения таких методов. Закончив изучение последней главы книги, читатель должен прийти к такому уровню понимания предмета, который является фундаментом для большинства работ, ныне проводимых в этой области.

Ссылки и литература для дальнейшего изучения

В конце последующих глав даются относящиеся к конкретным темам каждой из этих глав ссылки на общий список литературы, приведенный в конце книги. Однако в этой главе использована другая форма, чтобы собрать в одном месте всю совокупность научных журналов, публикующих материалы по обработке изображений и смежным вопросам. Приводится также список книг, по которым читатель легко может представить себе историческую и современную перспективу деятельности в этой области. Таким образом, приводимый в этой главе список литературных источников следует рассматривать как доступный путеводитель по изданной литературе в области обработки изображений.

К числу наиболее важных реферируемых журналов, публикующих статьи по обработке изображений и смежным вопросам, относятся: *IEEE Transactions on Image Processing*; *IEEE Transactions on Pattern Analysis and Machine Intelligence*; *Computer Vision, Graphics, and Image Processing* (до 1991 г.); *Computer Vision and Image Understanding*; *IEEE Transactions on Systems, Man and Cybernetics*; *Artificial Intelligence*; *Pattern Recognition*; *Pattern Recognition Letters*; *Journal of the Optical Society of America* (до 1984 г.); *Journal of the Optical Society of America — A: Optics, Image Science and Vision*; *Optical Engineering*; *Applied Optics — Information Processing*; *IEEE Transactions on Medical Imaging*; *Journal of Electronic Imaging*; *IEEE Transactions on Information Theory*; *IEEE Transactions on Communications*; *IEEE Transactions on Acoustics, Speech and Signal Processing*; *Proceedings of the IEEE*; а также выпуски журнала *IEEE Transactions on Computers* до 1980 г. Также представляют интерес публикации Международного общества по оптической технике (SPIE).

Следующие книги, перечисленные в обратном хронологическом порядке (и с уклоном в сторону более свежих публикаций), содержат материал, дополняющий наш взгляд на цифровую обработку изображений. Эти книги дают общую картину обсуждаемой области за последние 30 с лишним лет и были выбраны так, чтобы представить разнообразие существующих трактовок. В списке присутствуют и учебники, излагающие фундаментальный материал, и руковод-

ства по применению определенных методов, и, наконец, научные монографии, представляющие достигнутый уровень исследований в этой области.

Prince J. L., Links J. M. [2006]. *Medical Imaging, Signals, and Systems*, Prentice Hall, Upper Saddle River, NJ.

Bezdek J. C. et al. [2005]. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*, Springer, New York.

Davies E. R. [2005]. *Machine Vision: Theory, Algorithms, Practicalities*, Morgan Kaufmann, San Francisco, CA.

Rangayyan R. M. [2005]. *Biomedical Image Analysis*, CRC Press, Boca Raton, FL.

Umbaugh S. E. [2005]. *Computer Imaging: Digital Image Analysis and Processing*, CRC Press, Boca Raton, FL.

Gonzalez R. C., Woods R. E., Eddins S. L. [2004]. *Digital Image Processing Using MATLAB*, Prentice Hall, Upper Saddle River, NJ. [Имеется перевод: Гонсалес Р., Вудс Р., Эддинс С. *Цифровая обработка изображений в среде MATLAB*. — М.: Техносфера, 2006.

Snyder W. E., Qi Hairong [2004]. *Machine Vision*, Cambridge University Press, New York.

Klette R., Rosenfeld A. [2004]. *Digital Geometry — Geometric Methods for Digital Picture Analysis*, Morgan Kaufmann, San Francisco, CA.

Won C. S., Gray R. M. [2004]. *Stochastic Image Processing*, Kluwer Academic/Plenum Publishers, New York.

Soille P. [2003]. *Morphological Image Analysis: Principles and Applications*, 2nd ed., Springer-Verlag, New York.

Dougherty E. R., Lotufo R. A. [2003]. *Hands-on Morphological Image Processing*, SPIE — The International Society for Optical Engineering, Bellingham, WA.

Gonzalez R. C., Woods R. E. [2002]. *Digital Image Processing*, 2nd ed., Prentice Hall, Upper Saddle River, NJ.

Forsyth D. F., Ponce J. [2002]. *Computer Vision — A Modern Approach*, Prentice Hall, Upper Saddle River, NJ.

Duda R. O., Hart P. E., Stork D. G. [2001]. *Pattern Classification*, 2nd ed., John Wiley & Sons, New York.

Pratt W. K. [2001]. *Digital Image Processing*, 3rd ed., John Wiley & Sons, NY. [Имеется перевод 1-го издания: Прэйт У. *Цифровая обработка изображений*. — М.: Мир, 1982. Кн. 1 и 2.]

Ritter G. X., Wilson J. N. [2001]. *Handbook of Computer Vision Algorithms in Image Algebra*, CRC Press, Boca Raton, FL.

Shapiro L. G., Stockman G. C. [2001]. *Computer Vision*, Prentice Hall, Upper Saddle River, NJ.

Dougherty E. R. (ed.) [2000]. *Random Processes for Image and Signal Processing*, IEEE Press, New York.

Etienne E. K., Nachttegael M. (eds.) [2000]. *Fuzzy Techniques in Image Processing*, Springer-Verlag, New York.

Goutsias J., Vincent L., Bloomberg D. S. (eds.) [2000]. *Mathematical Morphology and Its Applications to Image and Signal Processing*, Kluwer Academic Publishers, Boston, MA.

Mallot A. H. [2000]. *Computational Vision*, The MIT Press, Cambridge, MA.

- Marchand-Maillet S., Sharaiha Y. M. [2000]. *Binary Digital Image Processing: A Discrete Approach*, Academic Press, New York.
- Mitra S. K., Sicuranza G. L. (eds.) [2000]. *Nonlinear Image Processing*, Academic Press, New York.
- Edelman S. [1999]. *Representation and Recognition in Vision*, The MIT Press, Cambridge, MA.
- Lillesand T. M., Kiefer R. W. [1999]. *Remote Sensing and Image Interpretation*, John Wiley & Sons, New York.
- Mather P. M. [1999]. *Computer Processing of Remotely Sensed Images: An Introduction*, John Wiley & Sons, New York.
- Petrou M., Bosdogianni P. [1999]. *Image Processing: The Fundamentals*, John Wiley & Sons, UK.
- Russ, J.C. [1999]. *The Image Processing Handbook*, 3rd ed., CRC Press, Boca Raton, FL.
- Smirnov A. [1999]. *Processing of Multidimensional Signals*, Springer-Verlag, New York.
- Sonka M., Hlavac V., Boyle R. [1999]. *Image Processing, Analysis, and Computer Vision*, PWS Publishing, New York.
- Haskell B. G., Netravali A. N. [1997]. *Digital Pictures: Representation, Compression, and Standards*, Perseus Publishing, New York.
- Jahne B. [1997]. *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*, Springer-Verlag, New York.
- Castleman K. R. [1996]. *Digital Image Processing*, 2nd ed., Prentice Hall, Upper Saddle River, NJ.
- Geladi P., Grahn H. [1996]. *Multivariate Image Analysis*, John Wiley & Sons, NY.
- Bracewell R. N. [1995]. *Two-Dimensional Imaging*, Prentice Hall, Upper Saddle River, NJ.
- Sid-Ahmed M. A. [1995]. *Image Processing: Theory, Algorithms, and Architectures*, McGraw-Hill, NY.
- Jain R., Rangachar K., Schunk B. [1995]. *Computer Vision*, McGraw-Hill, New York.
- Mitiche A. [1994]. *Computational Analysis of Visual Motion*, Perseus Publishing, New York.
- Baxes G. A. [1994]. *Digital Image Processing: Principles and Applications*, John Wiley & Sons, New York.
- Gonzalez R. C., Woods R. E. [1992]. *Digital Image Processing*, Addison-Wesley, Reading, MA.
- Haralick R. M., Shapiro L. G. [1992]. *Computer and Robot Vision*, vols. 1 & 2, Addison-Wesley, Reading, MA.
- Pratt W. K. [1991] *Digital Image Processing*, 2nd ed., Wiley-Interscience, New York. [Имеется перевод 1-го издания: Прэтт У. *Цифровая обработка изображений*. — М.: Мир, 1982. Кн. 1 и 2.]
- Lim J. S. [1990]. *Two-Dimensional Signal and Image Processing*, Prentice Hall, Upper Saddle River, NJ.
- Jain A. K. [1989]. *Fundamentals of Digital Image Processing*, Prentice Hall, Upper Saddle River, NJ.

- Schalkoff R. J. [1989]. *Digital Image Processing and Computer Vision*, John Wiley & Sons, New York.
- Giardina C. R., Dougherty E. R. [1988]. *Morphological Methods in Image and Signal Processing*, Prentice Hall, Upper Saddle River, NJ.
- Levine M. D. [1985]. *Vision in Man and Machine*, McGraw-Hill, New York.
- Serra J. [1982]. *Image Analysis and Mathematical Morphology*, Academic Press, New York.
- Ballard D. H., Brown C. M. [1982]. *Computer Vision*, Prentice Hall, Upper Saddle River, NJ.
- Fu K. S. [1982]. *Syntactic Pattern Recognition and Applications*, Prentice Hall, Upper Saddle River, NJ. [Имеется перевод более ранней книги: Фу К. *Структурные методы в распознавании образов*. — М.: Мир, 1977.]
- Nevatia R. [1982]. *Machine Perception*, Prentice Hall, Upper Saddle River, NJ.
- Pavlidis T. [1982]. *Algorithms for Graphics and Image Processing*, Computer Science Press, Rockville, MD.
- Rosenfeld R., Kak A. C. [1982]. *Digital Picture Processing*, 2nd ed., vols. 1 & 2, Academic Press, New York.
- Hall E. L. [1979]. *Computer Image Processing and Recognition*, Academic Press, New York.
- Gonzalez R. C., Thomason M. G. [1978]. *Syntactic Pattern Recognition: An Introduction*, Addison-Wesley, Reading, MA.
- Andrews H. C., Hunt B. R. [1977]. *Digital Image Restoration*, Prentice Hall, Upper Saddle River, NJ.
- Pavlidis T. [1977]. *Structural Pattern Recognition*, Springer-Verlag, New York, 1977.
- Tou J. T., Gonzalez R. C. [1974]. *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, 1974.
- Andrews H. C. [1970]. *Computer Techniques in Image Processing*, Academic Press, New York.

Литература, добавленная при переводе

Ниже приводится список литературы по вопросам цифровой обработки изображений, который, по нашему мнению, будет способствовать расширению кругозора читателя данной книги. Из российских специализированных периодических изданий, посвященных рассматриваемой проблематике, следует отметить журнал «Компьютерная оптика», и журнал *Pattern Recognition and Image Analysis*; к сожалению, последний публикуется издательством Pleiades Publishing исключительно на английском языке. К числу отечественных периодических изданий, в которых достаточно регулярно публикуются научные работы по данной тематике, относятся журналы «Автометрия», «Исследование Земли из Космоса» и «Кибернетика». Большинство современных результатов публикуются в трудах различных научных конференций, симпозиумов, ведомственных изданий, или же в виде сборников статей, издаваемых институтами Российской Академии наук или иными научными и учебными учреждениями. Из их числа без сомнения следует отметить такие сборники как «Иконика» и «Цифровая оптика», издававшиеся Институтом проблем передачи информации РАН, а также сбор-

ники «Вопросы кибернетики», издаваемые научным советом по комплексной проблеме «Кибернетика» РАН.

Ранее (до начала 1990 гг.) издавался перевод ежемесячного журнала «Труды института инженеров по электротехнике и радиоэлектронике (ТИИЭР)», в котором с достаточной регулярностью появлялись как отдельные статьи, так и тематические выпуски, посвященные обработке изображений. Без сомнения, особого внимания заслуживают следующие тематические выпуски ТИИЭР.

Сокращение избыточности. — ТИИЭР, 1967, т. 55, № 3.

Цифровая обработка изображений. — ТИИЭР, 1972, т. 60, № 7.

Распознавание образов и обработка изображений. — ТИИЭР, 1979, т. 67, № 5.

Цифровое кодирование. — ТИИЭР, 1980, т. 68, № 7.

Обработка изображений. — ТИИЭР, 1981, т. 69, № 5.

Системы видеосвязи. — ТИИЭР, 1985, т. 73, № 4 и № 5.

Реконструктивная и вычислительная томография. — ТИИЭР, 1983, т. 71, № 3.

Из числа изданных на русском языке книг, представляющих интерес при изучении вопросов цифровой обработки изображений (как оригинальных, так и переводных изданий), можно рекомендовать следующие.

Ахмед Н., Рао К. Р. *Ортогональные преобразования при обработке цифровых сигналов.* — М.: Связь, 1980.

Александров В. В., Горский Н. Д. *Представление и обработки изображений: Рекурсивный подход.* — Л.: Наука, 1985.

Бейтс Р., Мак-Доннелл М. *Восстановление и реконструкция изображений.* — М.: Мир, 1989.

Бендат Дж., Пирсол А. *Прикладной анализ случайных данных.* — М.: Мир, 1989.

Бонгард М. М. *Проблема узнавания.* — М.: Наука, 1967.

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.

Василенко Г. И. *Теория восстановления сигналов: О редукации к идеальному прибору в физике и технике.* — М.: Сов.радио, 1979.

Василенко Г. И., Тараторин А. М. *Восстановление изображений.* — М.: Радио и связь, 1986.

Ватолин Д. С. *Алгоритмы сжатия изображений. Методическое пособие.* М.: МГУ, 1999.

Введение в цифровую фильтрацию / Ред. Богнер Р., Константи́нидис А. — М.: Мир, 1976.

Виттих В. А., Сергеев В. В., Сойфер В. А. *Обработка изображений в автоматизированных системах научных исследований.* — М.: Наука, 1982.

Возенкрафт Дж., Джейкобс И. *Теоретические основы техники связи.* — М.: Мир, 1968.

Голд Б., Рэйдер Ч. *Цифровая обработка сигналов.* — М.: Сов. радио, 1973.

Гонсалес Р., Вудс Р. *Цифровая обработка изображений,* М.: Техносфера, 2005

Горелик А. Л., Скрипкин В. А. *Методы распознавания: Учебное пособие.* — М.: Высшая школа, 1989.

Гренандер У. *Лекции по теории образов*. — М.: Мир. т. 1: *Синтез образов*. — 1979. т. 2: *Анализ образов*. — 1981. т. 3: *Регулярные структуры*. — 1983.

Даджион Д., Мерсеро Р. *Цифровая обработка многомерных сигналов*. — М.: Мир, 1988.

Добеши И. *Десять лекций по вейвлетам*. — Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001.

Дуда Р., Харт П. *Распознавание образов и анализ сцен*. — М.: Мир, 1976.

Жуков А. И. *Метод Фурье в вычислительной математике*. — М.: Наука, 1992.

Завалишин Н. В., Мучник И. Б. *Модели зрительного восприятия и алгоритмы анализа изображений*. — М.: Наука, 1974.

Ковалевский В. А. *Методы оптимальных решений в распознавании изображений*. — М.: Наука, 1976.

Кодирование и обработка изображений / Ред. Лебедев Д. С., Зяблов В. В. — М.: Наука, 1988.

Коды с обнаружением и исправлением ошибок. — М.: ИЛ, 1956.

Лебедев Д. С., Цуккерман И. И. *Телевидение и теория информации*. М. — Л.: Энергия, 1965.

Марр Д. *Зрение. Информационный подход к изучению представления и обработки зрительных образов*. — М.: Радио и связь, 1987.

Методы компьютерной обработки изображений / Ред. Сойфер В. А. — М.: Физматлит, 2001.

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.

Минский М., Пейперт С. *Перцептроны*. — М.: Мир, 1971.

Нильсон Н. *Искусственный интеллект. Методы поиска решений*. — М.: Мир, 1973.

Нильсон Н. *Обучающие машины*. — М.: Мир, 1967.

Обработка изображений и цифровая фильтрация. / Ред. Хуанг Т. — М.: Мир, 1979.

Обработка изображений при помощи цифровых вычислительных машин / Ред. Эндрюс Г., Инло Л. — М.: Мир, 1973. (См. также ТИИЭР, 1972, т. 60, № 7).

Оппенгейм А. В., Шафер Р. В. *Цифровая обработка сигналов*. — М.: Связь, 1979.

Осовский С. *Нейронные сети для обработки информации*. — М.: Финансы и статистика, 2002.

Павлидис Т. *Алгоритмы машинной графики и обработки изображений*. — М.: Радио и связь, 1986.

Применение цифровой обработки сигналов / Ред. Оппенгейм Э. — М.: Мир, 1980.

Прэтт У. *Цифровая обработка изображений*. — М.: Мир, 1982. Кн. 1 и 2.

Психология машинного зрения / Ред. Уинстон П. — М.: Мир, 1978.

Рабинер Л., Гоулд Б. *Теория и применение цифровой обработки сигналов*. — М.: Мир, 1978.

Распознавание образов при помощи цифровых вычислительных машин / Ред. Хармон Л. — М.: Мир, 1974. (См. также ТИИЭР, 1972, т. 60, № 10).

Реконструкция изображений / Ред. Старк Г. — М.: Мир, 1992.

- Розенблатт Ф. *Принципы нейродинамики. Перцептрон и теория механизмов мозга*. — М.: Мир, 1966.
- Розенфельд А. *Распознавание и обработка изображений с помощью вычислительных машин*. — М.: Мир, 1972.
- Столниц Э., ДеРоуз Т., Салезин Д. *Вейвлеты в компьютерной графике. Теория и приложения*. — Ижевск: НИЦ «Регулярная и хаотическая динамика», 2002.
- Теория информации и ее приложения*. М.: Физматгиз, 1959.
- Тихонов А. Н., Арсенин В. Я. *Методы решения некорректных задач*. — М.: Наука, 1979.
- Ту Дж., Гонсалес Р. *Принципы распознавания образов*. — М.: Мир, 1978.
- Ульман Ш. *Принципы восприятия подвижных объектов*. — М.: Радио и связь, 1983.
- Фу К. *Последовательные методы в распознавании образов и обучении машин*. — М.: Наука, 1979.
- Фу К. *Структурные методы в распознавании образов*. — М.: Мир, 1977.
- Фукунага К. *Введение в статистическую теорию распознавания образов*. — М.: Наука, 1979.
- Хермен Г. *Восстановление изображений по проекциям. Основы реконструктивной томографии*. — М.: Мир, 1983.
- Хорн Б. К. П. *Зрение роботов*. — М.: Мир, 1989.
- Хэмминг Р. В. *Теория кодирования и теория информации*. — М.: Радио и связь, 1983.
- Цифровая обработка изображений: Специальный выпуск*. — Зарубежная радиоэлектроника. 1985, № 10.
- Цифровая обработка изображений в информационных системах: Учебное пособие* / Грузман И. С., Киричук В. С. и др. — Новосибирск, НГТУ, 2002.
- 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.
- Цифровое кодирование телевизионных изображений* / Цуккерман И. И., Кац Б. М., Лебедев Д. С. и др. — М.: Радио и связь, 1981.
- Чисар И., Кёрнер Я. *Теория информации: теоремы кодирования для дискретных систем без памяти*. — М.: Мир, 1985.
- Чуи К. *Введение в вэйвлеты*. — М.: Наука, 2001.
- Шашлов Б. А. *Цвет и цветовоспроизведение*. — М.: МГАП «Мир книги», 1995.
- Эндрюс Г. *Применение вычислительных машин для обработки изображений*. — М.: Энергия, 1977.
- Яне Б. *Цифровая обработка изображений*. — М.: Техносфера, 2007.
- Яншин В. В. *Анализ и обработка изображений: принципы и алгоритмы*. — М.: Машиностроение, 1994.
- Ярославский Л. П. *Введение в цифровую обработку изображений*. — М.: Сов. радио, 1979.
- Ярославский Л. П. *Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику*. — М.: Радио и связь, 1987.

ГЛАВА 2

ОСНОВЫ ЦИФРОВОГО ПРЕДСТАВЛЕНИЯ ИЗОБРАЖЕНИЙ

Желающий добиться успеха должен задавать правильные предварительные вопросы.

Аристотель

Введение

Цель этой главы — познакомить читателя с набором основных понятий в области цифровой обработки изображений, которые используются на протяжении всей книги. В разделе 2.1 описывается действие зрительной системы человека, включая механизмы формирования изображения в глазу, и возможности зрения в плане яркостной адаптации и контрастной чувствительности. В разделе 2.2 обсуждается свет и другие составляющие электромагнитного спектра, а также их характеристики в отношении формирования изображений. Раздел 2.3 посвящен обсуждению различных чувствительных элементов и способов регистрации цифровых изображений. В разделе 2.4 вводятся понятия равномерной пространственной дискретизации изображений и квантования по яркости. Кроме того, в этом разделе обсуждаются вопросы представления цифровых изображений и эффекты при изменении частоты дискретизации и количества уровней квантования. Вводятся понятия пространственного и яркостного разрешения и рассматриваются принципы интерполяции изображений. В разделе 2.5 рассматриваются некоторые основные взаимоотношения между пикселями. Наконец, раздел 2.6 служит введением в основной математический аппарат, используемый на протяжении всей книги. Другая цель этого раздела — помочь читателю почувствовать, как эти инструменты используются при решении разнообразных элементарных задач обработки изображений. Область рассмотрения и применения указанных инструментов по мере необходимости расширяется в оставшейся части книги.

2.1. Элементы зрительного восприятия

Хотя цифровая обработка изображений как прикладная дисциплина строится на основе математических и вероятностных формулировок, человеческая интуиция и анализ играют центральную роль при выборе того или иного метода

среди других, и этот выбор часто совершается на основе субъективного визуального оценивания. Поэтому первым шагом нашего путешествия по этой книге будет приобретение элементарных знаний о зрительном восприятии человека. Понимая сложность и широту этой темы, мы сможем охватить лишь самые начальные аспекты зрения человека. В частности, мы опишем физические механизмы и параметры, связанные с процессом формирования изображения и его восприятия человеком. Нам также интересно рассмотреть физические ограничения человеческого зрения с точки зрения факторов, используемых при работе с цифровыми изображениями. Поэтому такие вопросы, как соотношение человеческого и электронного зрения по разрешающей способности и возможности адаптироваться к изменениям освещенности, не только представляют познавательный интерес, но и важны в практическом плане.

2.1.1. Строение человеческого глаза

На рис. 2.1 в упрощенном виде показано горизонтальное сечение человеческого глаза. Глаз имеет почти сферическую форму со средним диаметром около 20 мм. Глаз окружен тремя оболочками: *роговица* со *склерой* образуют внешнюю оболочку, под которой последовательно расположены сосудистая оболочка (*хороидея*) и нейроглиальная оболочка (*сетчатка*). Роговица — это плотная прозрачная ткань, закрывающая переднюю поверхность глаза. Продолжением ее является склера — непрозрачная оболочка, закрывающая остальную часть оптической сферы глаза.

Хороидея расположена непосредственно под склерой. В этой оболочке расположена сеть кровеносных сосудов, обеспечивающих питание глаза. Даже незначительное повреждение хороиdei, часто кажущееся неопасным, может привести к серьезному нарушению зрения из-за воспаления, препятствующего нормальному кровотоку. Поверхностный слой хороиdei сильно пигментирован, что снижает интенсивность попадающего через склеру внешнего света, мешающего восприятию из-за его отражения и рассеяния внутри оптической сферы. Передняя область (зубчатая линия) хороиdei непосредственно вплетена в *цилиарное тело* и *радужную оболочку* (или *радужку*). Отверстие в центре радужной оболочки (*зрачок*) может сужаться или расширяться, регулируя тем самым количество попадающего через роговицу света. Диаметр зрачка может изменяться в пределах от 2 до 8 мм. От цвета пигмента на передней поверхности радужки зависит цвет глаза человека, а пигмент на задней поверхности радужки имеет черный цвет, что также снижает внутреннее рассеяние света.

Хрусталик (*хрусталиковая линза*), состоящий из наружной капсулы и внутривнутрихрусталикового вещества, закреплен внутри глаза с помощью передней и задней порций волокон ресничного пояса хрусталика, которые проходят между отростками цилиарного тела и вплетаются в зубчатую линию хороиdei. Капсула и внутривнутрихрусталиковое вещество состоят из коллагеновых волокон и содержат от 60 % до 70 % воды, около 6 % жиров и больше белков, чем любые другие ткани глаза. Внутривнутрихрусталиковое вещество имеет слабо-желтую пигментацию, которая с возрастом усиливается. Ускоренное помутнение вещества хрусталика, связанное с нарушением его питания, приводит к заболеванию, называемому *катарактой*, при котором ухудшается цветовое восприятие и остро-

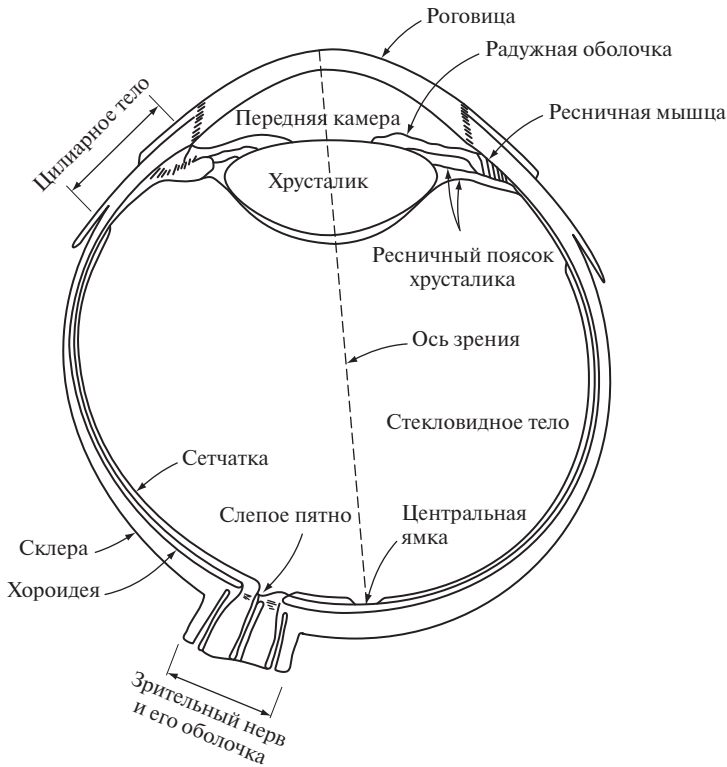


Рис. 2.1. Упрощенная схема глаза человека в разрезе

та зрения. Хрусталик поглощает около 8% света в видимом диапазоне спектра и практически не пропускает более коротковолновое излучение. Свет инфракрасного и ультрафиолетового диапазонов существенно поглощается белком хрусталика и при высокой интенсивности может привести к необратимому нарушению зрения.

Самая внутренняя оболочка глаза — сетчатка — выстилает изнутри задний отдел глаза. При правильной оптической фокусировке глаза свет от наружного объекта проецируется в виде изображения на сетчатку. Зрительное восприятие образов становится возможным благодаря распределению дискретных светочувствительных клеток (*рецепторов*) по внутренней поверхности сетчатки. Существуют рецепторы двух видов — *колбочки* и *палочки*. В глазу насчитывается от 6 до 7 миллионов колбочек, которые обладают высокой чувствительностью к спектральным составляющим света и располагаются преимущественно в центральной области сетчатки, называемой *желтым пятном*. В центре желтого пятна имеется так называемая *центральная ямка* — область наибольшей остроты зрения. Человек различает мелкие детали изображения в основном благодаря колбочкам, поскольку каждая из них соединена с отдельным нервным окончанием. Наружные мышцы глаза обеспечивают вращение глазного яблока так, чтобы изображение интересующего объекта попадало в область желтого пятна. Колбочки обеспечивают *фотопическое* зрение, или зрение в ярком свете.

Количество палочек в глазу намного больше: по поверхности сетчатки их распределено от 75 до 150 миллионов. Бóльшая, чем у колбочек, область распределения и тот факт, что к одному нервному окончанию присоединено сразу несколько палочек (в среднем около 10), уменьшают возможности различения деталей с помощью этих рецепторов. Палочки позволяют сформировать общую картину всего поля зрения. Они наиболее чувствительны при низких уровнях освещенности и не участвуют в обеспечении функций цветного зрения. Например, предметы, имеющие яркую окраску при дневном свете, при сумеречном освещении выглядят как лишённые цветов образы, поскольку возбуждаются только палочки. Это явление известно как *скотопическое* (или сумеречное) зрение.

Рис. 2.2 иллюстрирует зависимость плотности распределения палочек и колбочек по сетчатке в зависимости от величины угла между зрительной осью и линией, проведенной из центра хрусталика до сетчатки. Изображено горизонтальное сечение правого глаза в месте выхода зрительного нерва. Отсутствие рецепторов в этой области приводит к появлению так называемого *слепого пятна* (см. рис. 2.1). В остальной области сетчатки распределение рецепторов центрально симметрично относительно центра желтого пятна. Из рис. 2.2 видно, что максимальная плотность колбочек наблюдается в центре сетчатки (в центральной ямке), а плотность палочек возрастает от этой точки приблизительно до угла в 20° , после чего плавно снижается вплоть до периферии сетчатки.

Центральная ямка представляет собой углубление круглой формы в сетчатке с диаметром около 1,5 мм. В контексте дальнейшего обсуждения более естественно говорить о прямоугольных массивах чувствительных элементов. В несколько вольной интерпретации можно рассматривать центральную ямку как квадратный массив чувствительных элементов на площади $1,5 \times 1,5$ мм. Плотность колбочек в этой области сетчатки приблизительно равна 150 тыс. на 1 мм^2 , следовательно, общее количество колбочек в области наибольшей остроты зре-

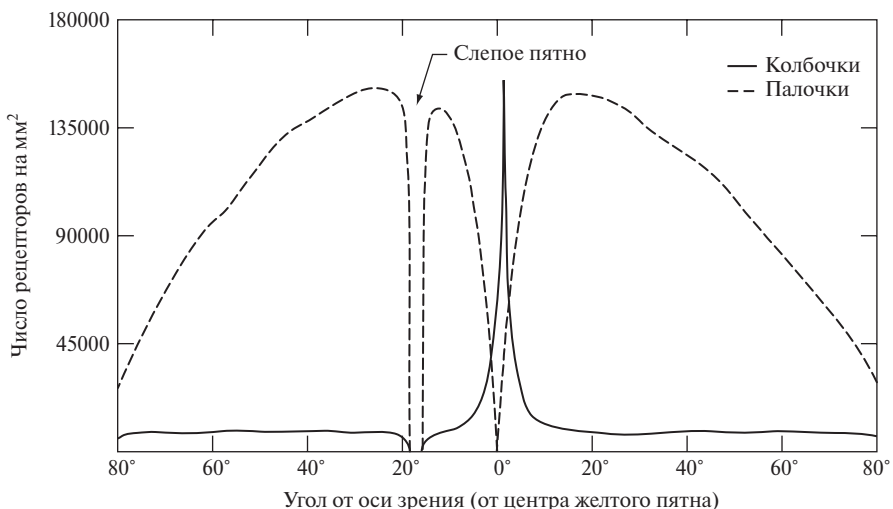


Рис. 2.2. Распределение палочек и колбочек по сетчатке

ния составляет около 337 тыс. элементов. Если рассуждать только в терминах разрешающей способности, то широко применяемые в современной технике светочувствительные матрицы среднего разрешения на основе приборов с зарядовой связью (ПЗС) содержат такое же количество чувствительных элементов при площади кристалла не более 5×5 мм. Однако такое чисто количественное сравнение было бы несколько поверхностным, поскольку не учитывает способность человека объединять зрение с интеллектом и опытом. В дальнейшем будем просто иметь в виду, что глаз человека по своей разрешающей способности вполне сопоставим с современными электронными устройствами получения изображений.

2.1.2. Формирование изображения в глазу

В обычном фотоаппарате объектив имеет постоянное фокусное расстояние, и фокусировка на различную дальность достигается изменением расстояния от объектива до плоскости изображения, в которой расположена фотопленка (или светочувствительная матрица в случае цифровой камеры). Напротив, в человеческом глазу расстояние между линзой (хрусталиком) и местом формирования изображения (сетчаткой) постоянное, а необходимое для получения четкой картины фокусное расстояние достигается за счет вариации формы хрусталика. Это осуществляется с помощью волокон ресничного пояса, придающих хрусталику уплощенную или округленную форму для фокусировки зрения, соответственно, на удаленном или близкорасположенном предмете. Расстояние от центра хрусталика до сетчатки вдоль оси зрения составляет приблизительно 17 мм, а диапазон изменения фокусного расстояния — от 14 мм до 17 мм (последнее значение соответствует расслабленному состоянию ресничной мышцы, когда глаз сфокусирован для рассматривания предметов на удалении более 3 м).

Рис. 2.3 иллюстрирует, как вычислить размеры изображения некоторого объекта на сетчатке. Пусть, например, наблюдатель смотрит на дерево высотой 15 м с расстояния 100 м. Обозначая h высоту изображения дерева на сетчатке (в мм), получаем пропорцию $15/100 = h/17$, откуда $h = 2,55$ мм. Как указывалось в разделе 2.1.1, проецирующееся на сетчатку изображение воспринимается главным образом областью желтого пятна. Расположенные в ней рецепторы возбуждаются в соответствии с интенсивностью падающего света, что приводит к преобразованию энергии светового излучения в электрические нервные импульсы, которые в конечном счете декодируются в мозге человека.

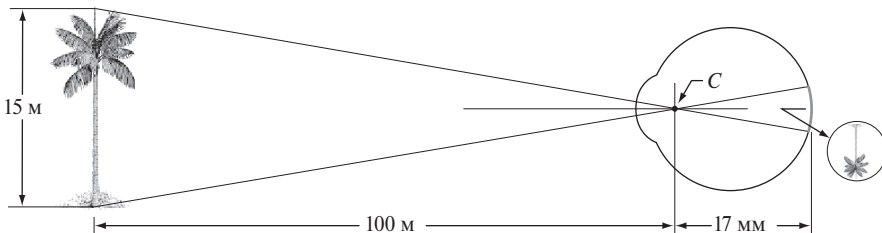


Рис. 2.3. Схематическое изображение глаза, наблюдающего дерево (точка C — оптический центр хрусталика)

2.1.3. Яркостная адаптация и контрастная чувствительность

Поскольку цифровые изображения воспроизводятся как дискретное множество элементов с различной яркостью, при представлении результатов обработки изображений необходимо учитывать способность глаза различать отличающиеся уровни яркости. Зрительная система человека способна адаптироваться к огромному, порядка 10^{10} , диапазону яркостей — от порога чувствительности скотопического зрения до предела ослепляющего блеска. Эксперименты со всей очевидностью показывают, что *субъективная яркость* (т. е. яркость как она воспринимается зрительной системой человека) является логарифмической функцией от физической яркости света, попадающего в глаз. На рис. 2.4 изображен график этой зависимости субъективной яркости от истинной яркости. Длинная сплошная кривая представляет диапазон яркостей, в котором способна адаптироваться зрительная система. При использовании одного фотопического зрения этот диапазон составляет около 10^6 . Постепенный переход от скотопического зрения к фотопическому происходит в диапазоне приблизительно от 0,003 до 0,3 кд/м² (соответствует диапазону от -3 до -0 на графике), что показано в виде двух ветвей кривой адаптации в этом диапазоне яркостей.

Для правильной интерпретации столь впечатляющего динамического диапазона, изображенного на рис. 2.4, важно понимать, что зрительная система не способна работать во всем этом диапазоне *одновременно*. Вместо этого она охватывает такой большой диапазон за счет изменения общей чувствительности. Это явление известно как *яркостная адаптация*. Общий диапазон одновременно различаемых уровней яркости относительно мал по сравнению со всем диапазоном адаптации. Для любого данного набора внешних условий текущий уровень чувствительности зрительной системы, называемый *уровнем яркостной адаптации*, соответствует некоторой яркости, например точке B_a на рис. 2.4.

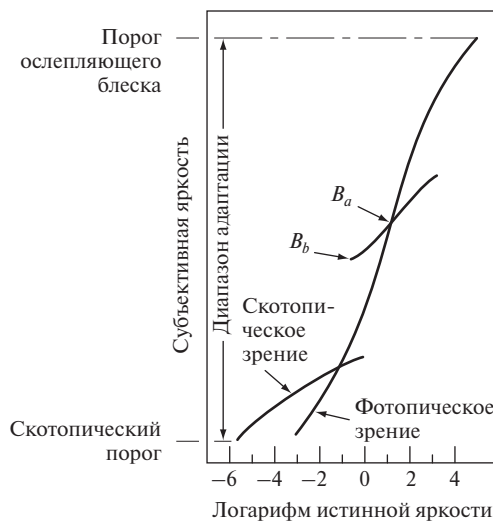


Рис. 2.4. Диапазон субъективно воспринимаемой яркости и конкретный уровень адаптации

Короткая кривая, пересекающая основной график, представляет диапазон субъективной яркости, которую способен воспринимать глаз при адаптации к указанному уровню. Этот диапазон достаточно ограничен: все уровни яркости ниже B_b субъективно воспринимаются зрением как черное и, значит, неразличимы. Верхняя часть этой кривой реально не ограничена, но теряет смысл при большой длине, поскольку при повышении яркости просто повышается уровень адаптации B_a .

Способность зрения различать *изменения яркости* при данном уровне адаптации также представляет значительный интерес. Классический эксперимент для определения способности зрительной системы человека различать разные уровни яркости состоит в том, что испытуемый смотрит на плоский равномерно освещенный экран достаточно больших размеров, такой, что он занимает все поле зрения. Как правило, это рассеиватель из матового стекла, освещаемый со стороны, противоположной наблюдателю, световым источником, яркость I которого можно регулировать. На это равномерное поле накладывается добавочная яркость ΔI в форме кратковременной вспышки в области круглой формы, расположенной в центре равномерно освещенного экрана, как изображено на рис. 2.5.

Если приращение ΔI недостаточно велико (неразличимо), испытуемый говорит «нет», указывая тем самым на отсутствие видимых изменений. По мере увеличения ΔI в какой-то момент он начнет говорить «да», подтверждая тем самым восприятие изменений яркости. Наконец, при достаточно большом значении ΔI испытуемый станет говорить «да» на каждую вспышку. Величина $\Delta I_c / I$, где ΔI_c — величина приращения яркости, различимая в 50% случаев на фоне яркости I , называется *отношением Вебера*. Малое значение $\Delta I_c / I$ означает, что различаются очень малые относительные изменения яркости, т.е. имеет место «высокая» контрастная чувствительность. Наоборот, большое значение $\Delta I_c / I$ означает, что требуется большое относительное изменение яркости, чтобы его заметить; это говорит о «низкой» контрастной чувствительности.

График зависимости величины $\log(\Delta I_c / I)$ от $\log I$ имеет общую форму, изображенную на рис. 2.6. Эта кривая показывает, что низкая контрастная чувствительность (т.е. большое отношение Вебера) наблюдается при малых уровнях яркости, и контрастная чувствительность заметно возрастает (т.е. отношение Вебера уменьшается) при увеличении фоновой яркости. Наличие двух ветвей кривой отражает тот факт, что при малых уровнях яркости зрение осуществляется благодаря действию палочек, тогда как при больших уровнях яркости (которым соответствует высокая контрастная чувствительность) зрительные функции выполняют колбочки сетчатки.

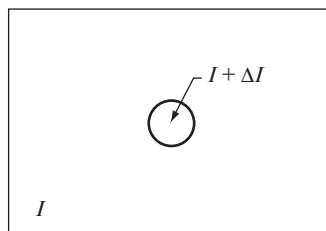


Рис. 2.5. Постановка простого эксперимента для определения характеристик контрастной чувствительности

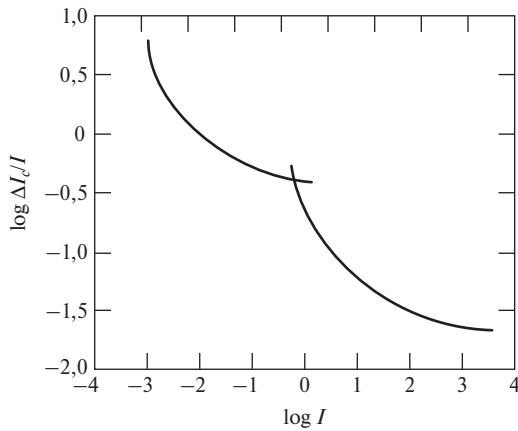


Рис. 2.6. Типичная зависимость отношения Вебера как функции яркости

Если поддерживать фоновую яркость постоянной, а яркость добавочного источника варьировать не вспышками, а ступенчатым изменением яркости от неотличимого до заметного всегда, то типичный наблюдатель способен различить всего 10—20 различающихся ступеней яркости. Грубо говоря, этот результат относится к числу различных уровней яркости, которые человек способен различить в произвольной точке монохромного изображения. Это не означает, что изображение может быть представлено таким небольшим числом градаций яркости, так как по мере движения взгляда по изображению меняется

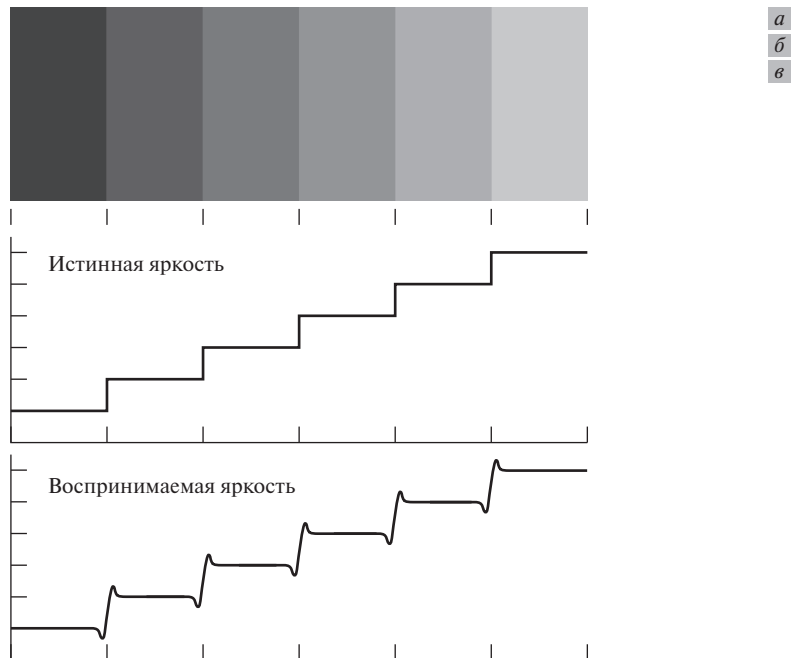


Рис. 2.7. Пример, показывающий, что воспринимаемая яркость не является просто функцией от истинной яркости

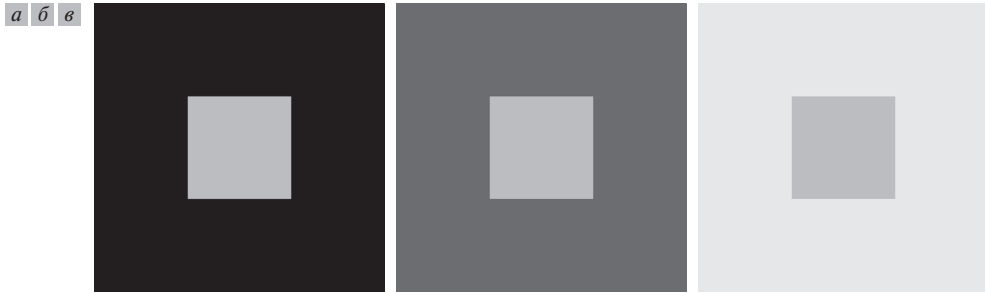


Рис. 2.8. Примеры одновременного контраста. Яркость всех центральных квадратов одинакова, но они кажутся все темнее, чем светлее становится фон

среднее значение яркости фона, что позволяет обнаруживать *различные* множества относительных изменений яркости для каждого нового уровня адаптации. Конечным следствием является способность глаза различать яркости в намного более широком *общем* диапазоне. В действительности, как мы покажем в разделе 2.4.3, глаз способен обнаруживать нежелательные ложные контуры в монохромных изображениях, общий диапазон яркостей которых представляется значительно большим количеством, чем 20 уровней.

Известны два явления, ясно доказывающие, что воспринимаемая яркость не является простой функцией истинной яркости. Первое основывается на том факте, что вблизи границ соседних областей с отличающимися, но постоянными яркостями зрение человека склонно «подчеркивать» яркостные перепады, как бы добавляя несуществующие выбросы яркости, что убедительно демонстрирует пример на рис. 2.7(а). Хотя яркость каждой из полос постоянна, мы, кроме действительно ступенчатого изменения яркости, видим характерные выбросы вблизи краев полос (рис. 2.7(в)). Эти полосы с кажущимися изменениями яркости на краях называются *полосами Маха* в честь Эрнста Маха, впервые описавшего этот феномен в 1865 г.

Второе явление, называемое *одновременным контрастом*, связано с тем фактом, что воспринимаемая яркость некоторой области не определяется просто ее яркостью, как показывает рис. 2.8. Здесь все центральные квадраты имеют в точности одинаковую яркость, однако зрительно воспринимаются тем более темными, чем светлее фон. Еще более знакомым примером является лист бумаги, который кажется белым, когда он лежит на столе, но может показаться совершенно черным, если им закрывать глаза, глядя на яркое небо.

Другими примерами феноменов человеческого зрительного восприятия являются оптические иллюзии, в которых глаз восполняет несуществующую информацию или ошибочно воспринимает геометрические свойства объектов. Некоторые примеры оптических иллюзий изображены на рис. 2.9. На рис. 2.9(а) ясно видны очертания квадрата, вопреки тому факту, что на изображении отсутствуют линии, определяющие такую фигуру. Аналогичный эффект, на этот раз в виде круга, виден на рис. 2.9(б); заметим, как всего нескольких линий достаточно для получения иллюзии полного круга. Два горизонтальных отрезка на рис. 2.9(в) имеют одинаковую длину, но один кажется короче другого. Наконец, все проведенные под углом 45° линии на рис. 2.9(г) параллельны

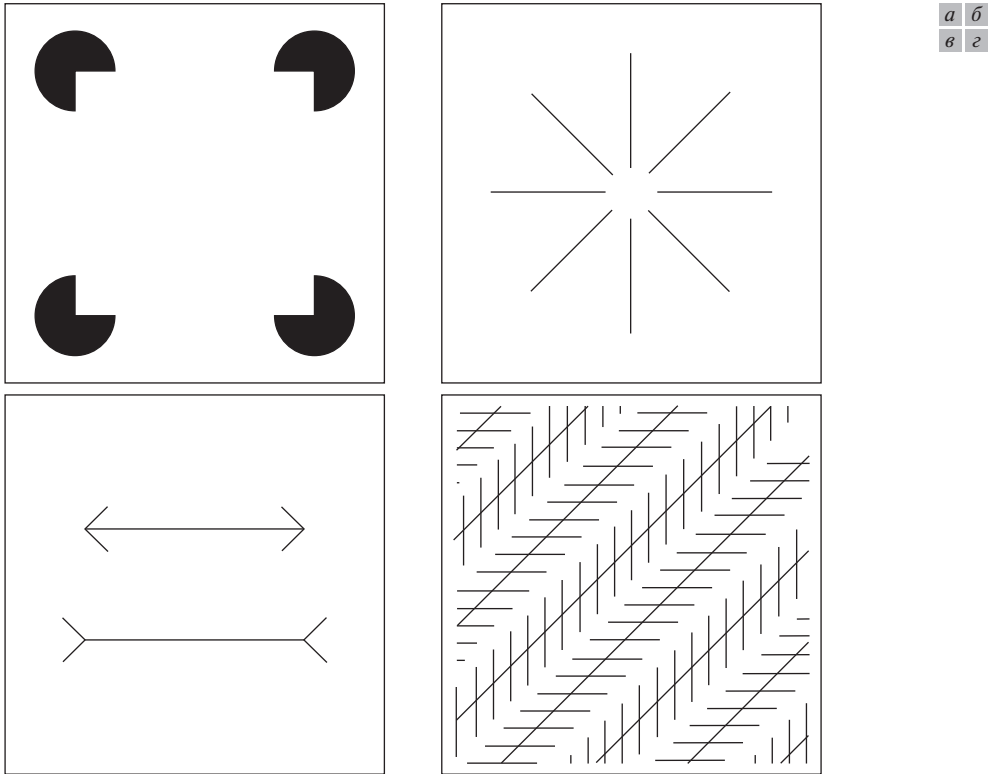


Рис. 2.9. Некоторые хорошо известные оптические иллюзии

и расположены на одинаковых расстояниях друг от друга. Однако штриховка создает иллюзию, что эти линии далеки от параллельности. Оптические иллюзии относятся к числу не вполне понятных характеристик зрительной системы человека.

2.2. Свет и электромагнитный спектр

Электромагнитный спектр был кратко представлен в разделе 1.3; теперь мы рассмотрим эту тему более подробно. В 1666 г. сэр Исаак Ньютон открыл, что при прохождении луча солнечного света сквозь стеклянную призму возникает световой пучок, который имеет не белый цвет, а состоит из непрерывного цветового спектра, цвет которого меняется от фиолетового на одном конце до красного на другом. Как видно из рис. 2.10, диапазон цветов, которые мы воспринимаем как видимый свет, составляет очень малую часть спектра электромагнитного излучения. На одном конце этого спектра находятся радиоволны, длина которых в миллиарды раз превышает длины волн видимого света, а на другом конце — гамма-лучи, длина волны которых в миллионы раз меньше длины световых волн. Компоненты электромагнитного спектра можно выражать в терминах длины волны, частоты колебаний или энергии. Длина волны (λ) и частота (ν) связаны соотношением

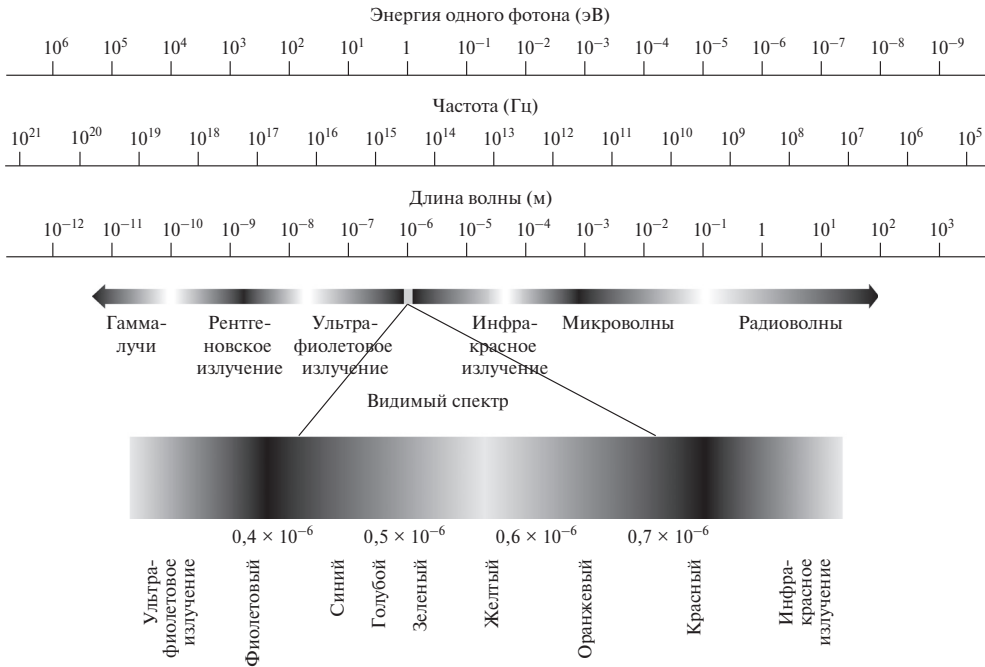


Рис. 2.10. Спектр электромагнитных колебаний. Видимый спектр показан в растянутом виде, но следует подчеркнуть, что он занимает весьма узкий участок всего электромагнитного спектра

$$\lambda = \frac{c}{\nu}, \quad (2.2-1)$$

где c — скорость света ($2,998 \cdot 10^8$ м/с). Энергия составляющей электромагнитного спектра определяется выражением

$$E = h\nu, \quad (2.2-2)$$

где h — постоянная Планка. Длина волны измеряется в метрах, но столь же часто в качестве единиц измерения употребляются *микрон* ($1 \text{ мкм} = 10^{-6}$ м) и *нанометр* ($1 \text{ нм} = 10^{-9}$ м)¹. Частота измеряется в герцах (Гц); 1 Гц соответствует колебанию с частотой один период в секунду. Общеупотребительной единицей измерения энергии фотонов является электрон-вольт (эВ).

Электромагнитные волны можно трактовать как распространяющиеся синусоидальные колебания с длиной волны λ (рис. 2.11), а можно — как поток частиц с нулевой массой, движущихся со скоростью света. Каждая такая частица не имеет массы, но обладает определенной энергией и называется *квантом излучения (фотоном)*. Из соотношения (2.2-2) видно, что энергия пропорциональна частоте, поэтому электромагнитные колебания более высокой частоты (т. е. с более короткой длиной волны) обладают большей энергией фотона. Таким образом, радиоволны характеризуются малой энергией фотона, у микроволн энергия больше, у инфракрасного излучения еще больше, далее энергия фотона

¹ Весьма распространенной единицей измерения длины световой волны является также ангстрем ($1 \text{ \AA} = 10^{-10}$ м). — *Прим. перев.*

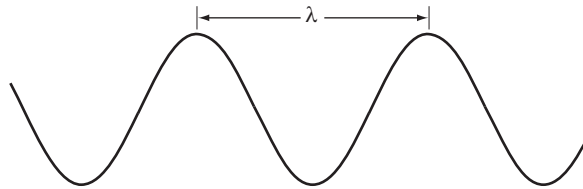


Рис. 2.11. Графическое представление длины волны колебаний

последовательно возрастает для диапазонов видимого спектра, ультрафиолетового излучения, рентгеновских лучей и, наконец, гамма-лучей, обладающих самой большой энергией. Именно по этой причине гамма-излучение так опасно для живых организмов.

Свет является особым видом электромагнитного излучения, которое воспринимается человеческим глазом. Видимый (цветовой) спектр приведен на рис. 2.10 в растянутом виде только для сведения; более подробно цвет рассматривается в главе 6. Этот видимый диапазон электромагнитного спектра охватывает длины волн приблизительно от 0,43 мкм (фиолетовый цвет) до 0,79 мкм (красный цвет). Для удобства цветовой спектр делят на семь широких полос (цветов): фиолетовый, синий, голубой, зеленый, желтый, оранжевый и красный; но это разграничение не резкое, а скорее, один цвет плавно переходит в другой, как показано на рис. 2.10, подобно любой другой составляющей электромагнитного спектра.

Различаемые зрением человека цвета предметов определяются характером света, *отраженного* от этих предметов. Тело, которое отражает свет приблизительно одинаково во всем видимом диапазоне волн, представляется наблюдателю белым, тогда как тело, отражающее свет в каком-то ограниченном диапазоне длин волн, воспринимается с некоторым цветовым оттенком. Например, зеленый предмет в основном отражает свет с длинами волн 500—570 нм, поглощая большинство энергии в других интервалах длин волн.

Свет, лишенный цветовой окраски, называется *монохроматическим* или *ахроматическим*. Единственным параметром такого освещения является его интенсивность, или *яркость*. Для описания монохроматической яркости также используется термин *уровень серого*, поскольку яркость изменяется от черного до белого, с промежуточными серыми оттенками. В дальнейшем термины «яркость» и «уровень серого» будут использоваться в одинаковом значении. Диапазон измеренных значений монохроматической яркости от черного до белого принято называть *шкалой полутонов*, и монохроматические изображения часто называют *полутоновыми изображениями*.

Хроматический (цветной) свет охватывает, как уже отмечалось, спектр электромагнитного излучения в диапазоне длин волн приблизительно от 0,43 мкм до 0,79 мкм. Помимо частоты, хроматические источники света характеризуются тремя основными величинами: энергетическим потоком, световым потоком и (субъективной) яркостью. *Энергетический поток* — это общее количество энергии, излучаемой источником света, обычно измеряемое в ваттах (Вт). *Световой поток*, измеряемый в люменах (лм), характеризует количество энергии, которое наблюдатель *воспринимает* от светового источника. Например, световой источник, работающий в дальнем инфракрасном диапазоне, может давать

значительный энергетический поток, но наблюдатель его практически не ощущает, так что световой поток такого источника почти нулевой. Наконец, как уже обсуждалось в разделе 2.1, *яркость* описывает субъективное восприятие света и практически не поддается измерению. Она олицетворяет понятие интенсивности в ахроматическом случае и является одним из ключевых факторов при описании цветового ощущения.

Продолжая обсуждение рис. 2.10, заметим, что коротковолновая сторона спектра электромагнитного излучения представлена гамма- и рентгеновскими лучами. Мы уже обсуждали в разделе 1.3.1 важность использования изображений в гамма-лучах для медицины, астрономии и ядерной энергетики. Жесткое (с более высокой энергией) рентгеновское излучение используется для получения изображений в промышленности. Для получения рентгеновских изображений грудной клетки и в стоматологии используется мягкое рентгеновское излучение (с меньшей энергией). Мягкие рентгеновские лучи плавно переходят в дальний ультрафиолетовый диапазон, длинноволновый участок которого, в свою очередь, — в видимый спектр. Двигаясь дальше в сторону увеличения длин волн, мы встретим инфракрасный диапазон, в котором излучается тепло, что делает его полезным для получения изображений на основе тепловой картины объекта. Участок инфракрасного диапазона, соседствующий с видимым спектром, называется *ближним инфракрасным* диапазоном, а противоположный участок — *дальним инфракрасным* диапазоном. Последний плавно переходит в микроволновый диапазон, хорошо известный благодаря кухонным микроволновым печам, но также используемый во многих других целях, в том числе для связи и радиолокации. Наконец, в диапазоне радиоволн осуществляется теле- и радиовещание, а в области высоких энергий этого диапазона проводятся астрономические наблюдения радиосигналов, испускаемых некоторыми звездными телами. Примеры изображений для большинства перечисленных диапазонов излучения были приведены в разделе 1.3.

В принципе, если сконструировать чувствительный элемент, способный обнаруживать излучаемую энергию в некотором диапазоне электромагнитного спектра, то можно получить изображение интересующих объектов в этом диапазоне. Однако важно заметить, что длина электромагнитных волн, используемых для «наблюдения» некоторого объекта, должна быть меньше его размера. Например, размер молекулы воды равен порядка 10^{-10} м, поэтому для исследования этих молекул необходимо применять источник излучения в диапазонах ультрафиолетового или мягкого рентгеновского излучения. Подобные ограничения, наряду с физическими свойствами материала, из которого изготовлен чувствительный элемент, определяют физические пределы возможностей сенсоров, применяемых для регистрации изображений, в частности оптических, инфракрасных или других.

Хотя подавляющее большинство получаемых цифровых изображений основано на энергии излучения электромагнитных волн, это не единственный способ генерации изображений. Например, как говорилось в разделе 1.3.7, отраженные от объектов звуковые волны могут использоваться для построения ультразвуковых изображений. Другие важные источники цифровых изображений — электронные пучки, применяемые в электронной микроскопии, и компьютерный синтез, используемый для визуализации и в компьютерной графике.

2.3. Считывание и регистрация изображения

Большинство интересующих нас изображений есть двумерное отображение наблюдаемой сцены (как правило, двух- или трехмерной), возникающее как результат регистрации лучистой энергии, исходящей из наблюдаемой сцены, с помощью некоторого устройства — сенсора (или совокупности сенсоров одновременно). Мы предполагаем, что регистрируемый сенсором сигнал возникает в результате взаимодействия источника «освещения» с элементами изображаемой «сцены» в условиях эффектов отражения и поглощения энергии этого источника. Мы берем слова *освещение* и *сцена* в кавычки, чтобы подчеркнуть тот факт, что они носят значительно более общий характер, чем в привычной ситуации, когда источник видимого света освещает обычную трехмерную бытовую сцену. Например, освещение не только может порождаться источником друго-

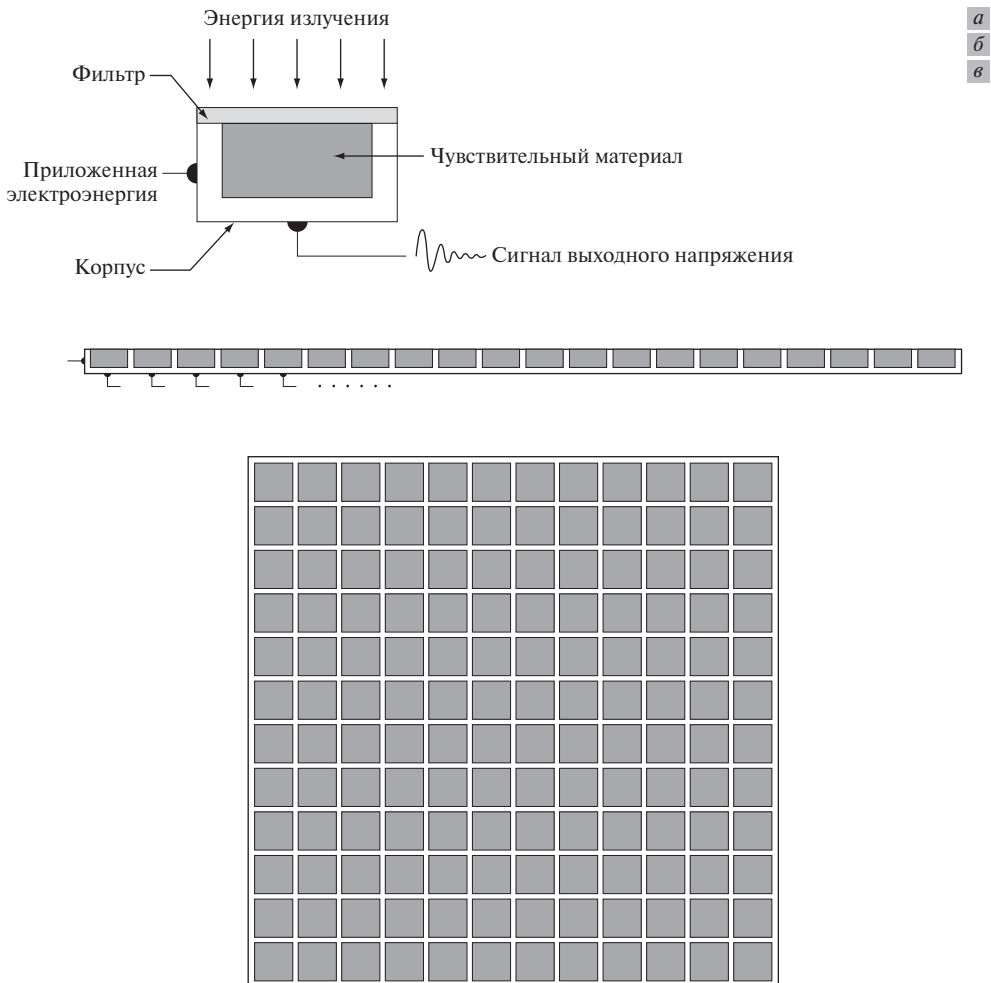


Рис. 2.12. (а) Одиночный чувствительный элемент. (б) Линейка чувствительных элементов. (в) Матрица чувствительных элементов

го диапазона электромагнитного излучения, например радиолокационным, инфракрасным или рентгеновским, но и происходить из менее традиционных источников, например ультразвукового или даже виртуального, синтезированного компьютерной программой. В роли элементов сцены могут выступать знакомые предметы, но вполне могут быть и молекулы, структуры подземных пластов или мозг человека. В зависимости от природы источника и особенностей сцены, энергия освещения отражается от объектов сцены или проходит сквозь них. Примером первого вида может быть свет, отраженный от поверхности предметов. Второй вид взаимодействия имеет место, например, при пропускании рентгеновских лучей через тело пациента для получения диагностического рентгеновского снимка на пленке. В некоторых прикладных задачах отраженная или проходящая энергия направляется на фотопреобразователь (например экран, покрытый флуоресцирующим материалом), который преобразует эту энергию в видимый свет. Такой подход обычен для электронной микроскопии и регистрации изображений в гамма-лучах.

На рис. 2.12 изображены три основные схемы размещения чувствительных элементов (сенсоров), которые используются для преобразования энергии «освещения» в цифровое изображение. Сама идея преобразования очень проста: падающая энергия преобразуется в напряжение благодаря сочетанию материала, обладающего чувствительностью к интересующему виду излучения, и приложенной к нему электрической энергии. В ответ на энергию внешнего излучения такой чувствительный элемент выдает сигнал выходного напряжения, который затем преобразуется в цифровую форму. В этом разделе мы рассмотрим основные способы получения и регистрации изображений, а вопросы дискретизации и квантования изображений обсуждаются в разделе 2.4.

2.3.1. Регистрация изображения с помощью одиночного сенсора

На рис. 2.12(a) показаны компоненты одиночного сенсора (чувствительного элемента). Вероятно, наиболее известным сенсором такого типа является фотодиод, изготовленный из полупроводникового материала (кремния), напряжение выходного сигнала которого пропорционально освещенности. Установка фильтра перед чувствительным элементом обеспечивает избирательность сенсора. Например, если установить перед сенсором зеленый пропускающий фильтр, то выходной сигнал будет выше для зеленого участка видимого спектра, чем для всех остальных.

Для получения двумерного изображения с помощью одиночного сенсора необходимо обеспечить его перемещение в двух взаимно перпендикулярных направлениях (по осям x и y) относительно регистрируемой области. На рис. 2.13 изображена конструкция, применяемая в прецизионных сканерах, где пленочный негатив закрепляется на барабане, вращение которого обеспечивает перемещение по одной оси. Одиночный сенсор закреплен на ходовом винте, вращение которого приводит к линейной подаче в перпендикулярном направлении. Поскольку механическим перемещением можно управлять с большой точностью, такой способ позволяет регистрировать изображения с высоким разрешением при небольших затратах (но медленно). Другой вид

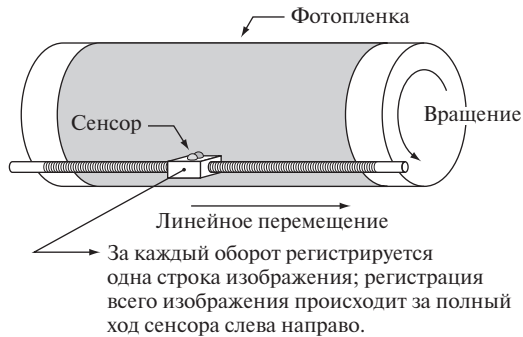


Рис. 2.13. Перемещение одиночного сенсора при регистрации двумерного изображения

механической конструкции аналогичного назначения — это планшет, по которому чувствительный элемент линейно передвигается в двух направлениях. Такие устройства с последовательным механическим сканированием всего поля изображения (как барабанные, так и планшетные) иногда называют *микроденситометрами*.

При другом способе получения изображения с помощью одиночного сенсора используется лазерный источник света, конструктивно совмещенный с сенсором. Механическая развертка осуществляется с помощью движения зеркал, направляющих луч источника на сканируемый плоский объект и возвращающих отраженный луч на чувствительный элемент. Та же конструкция может применяться для регистрации изображений с помощью линеек и матриц чувствительных элементов, что обсуждается в следующих двух параграфах.

2.3.2. Регистрация изображения с помощью линейки сенсоров

Более часто, чем одиночный сенсор, для считывания изображений используется одномерный массив сенсоров, обычно располагаемых вдоль прямой, как это показано на рис. 2.12(б). Такая линейка обеспечивает одновременную регистрацию элементов изображения в одном направлении (условно говоря, по строке), а перемещение всей линейки в перпендикулярном направлении позволяет получить все строки изображения (рис. 2.14(а)). Подобная конструкция применяется в большинстве планшетных сканеров. Удаётся изготавливать линейки, состоящие из 4000 и более расположенных в ряд чувствительных элементов. Расположение сенсоров в ряд широко используется при аэрофотосъемке, когда система регистрации устанавливается на самолете, летящем с постоянной скоростью и на неизменной высоте над интересующим районом. Одномерные линейки сенсоров, чувствительных к излучениям в различных участках электромагнитного спектра, располагаются перпендикулярно направлению полета. В каждый момент времени линейка сенсоров регистрирует одну строку изображения, а движение всей системы в перпендикулярном направлении позволяет заполнить второе измерение двумерного изображения. Для проекции сканируемой области на линейку сенсоров применяются линзы или другие фокусирующие устройства.

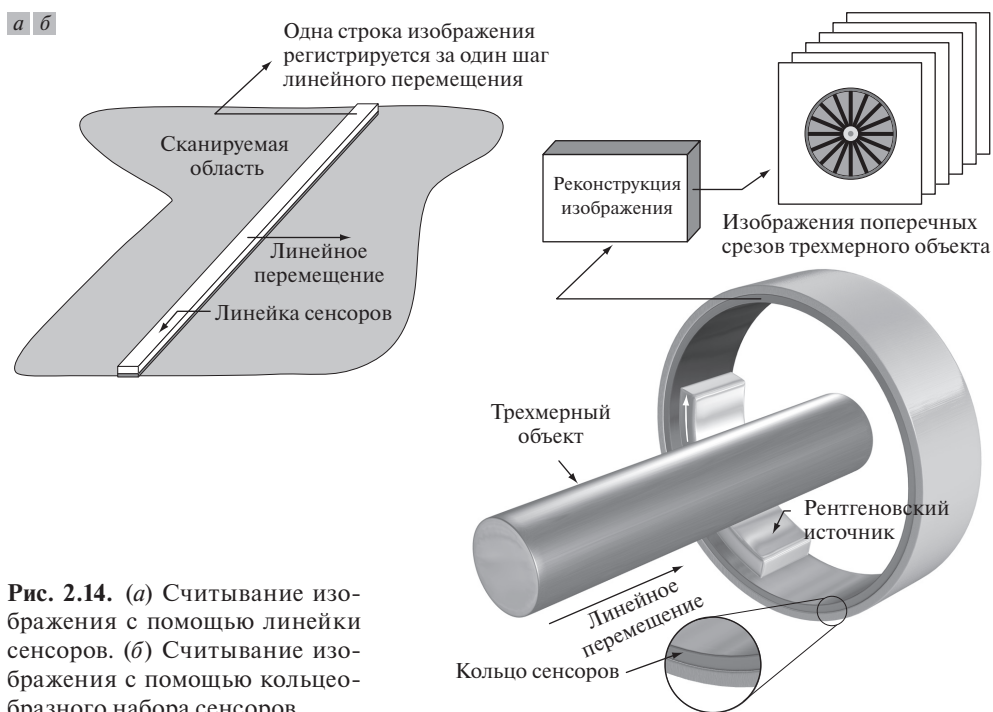


Рис. 2.14. (а) Считывание изображения с помощью линейки сенсоров. (б) Считывание изображения с помощью кольцеобразного набора сенсоров

Кольцеобразные наборы сенсоров применяются в медицине и промышленности для получения изображений поперечного сечения («срезов») трехмерных объектов, как показано на рис. 2.14(б). Вращающийся рентгеновский источник освещает объект, а расположенные на противоположной стороне кольца детекторы рентгеновского излучения улавливают энергию рентгеновских лучей, прошедших сквозь объект. Таков основной принцип получения изображений в компьютерной томографии (КТ), ранее упоминавшейся в разделах 1.2 и 1.3.2. Важно отметить, что выходные сигналы сенсоров подлежат обработке с помощью алгоритмов реконструкции, задача которых состоит в преобразовании регистрируемых данных в осмысленные изображения поперечных срезов (см. раздел 5.11). Другими словами, изображение среза не может быть получено простой регистрацией принимаемых сигналов одновременно с движением источника, а необходима значительная по объему вычислений обработка этих первичных данных. Трехмерное представление исследуемого объекта, состоящее из серии последовательных срезов, полученных с некоторым шагом, генерируется путем перемещения объекта в направлении, перпендикулярном к плоскости кольца. Существуют и другие способы регистрации изображений с использованием принципа КТ, но на базе иных физических процессов, в частности получение изображений методом ядерного магнитного резонанса (ЯМР) и позитронной эмиссионной томографии (ПЭТ). В них используются источники освещения и чувствительные элементы других типов, отличается и вид получаемых изображений, но принципиально эти способы регистрации изображений в значительной степени основаны на базовой схеме, показанной на рис. 2.14(б).

2.3.3. Регистрация изображения с помощью матрицы сенсоров

На рис. 2.12(в) изображено расположение отдельных сенсоров в форме двумерного массива (матрицы). Многочисленные электромагнитные и некоторые ультразвуковые устройства ввода данных современных систем обработки изображений используют именно матрицу сенсоров. Такая же конструкция находится внутри подавляющего числа цифровых камер, в которых типичным чувствительным элементом является матрица на основе приборов с зарядовой связью (ПЗС), которые выпускаются в виде монолитной конструкции, объединяющей 4000×4000 элементов (и более) с широким диапазоном чувствительных свойств. ПЗС-матрицы широко используются в цифровых фото- и видеокамерах, а также других светочувствительных приборах. Отклик каждого сенсора пропорционален интегралу световой энергии, попадающей на его поверхность за время экспозиции; это свойство используется в астрономии и других приложениях, где требуется получать изображения с низким уровнем шума. Уменьшение шума достигается за счет того, что чувствительным элементам дают возможность интегрировать принимаемый световой сигнал в течение минут или даже часов. Коль скоро изображенная на рис. 2.12(в) матрица сенсоров двумерна, ее главное достоинство состоит в том, что можно сфокусировать сразу все изображение, если сфокусировать на поверхности матрицы отвечающий ему пространственный поток лучистой энергии. Легко видеть, что в таком случае отпадает необходимость в механическом перемещении сенсоров, как это было в рассмотренных выше случаях одиночного сенсора или линейки таких сенсоров.

В некоторых случаях регистрируется изображение самого источника, например при получении снимков Солнца.

Рис. 2.15 иллюстрирует главный способ использования матриц сенсоров. Здесь показано, что энергия, излучаемая источником освещения, отражается от объекта сцены (но, как отмечалось в начале этого раздела, энергия также может и проникать сквозь объекты сцены). Первая функция, выполняемая системой формирования изображения (рис. 2.15(в)), состоит в том, чтобы собрать поступающую энергию и сфокусировать ее на плоскости изображения. Если для освещения используется источник видимого света, то на входе системы формирования изображения имеется оптическая линза, которая проецирует наблюдаемую сцену на плоскость изображения (рис. 2.15(е)). Совмещенная с этой плоскостью чувствительная матрица генерирует набор выходных сигналов, каждый из которых пропорционален интегралу световой энергии, принятой соответствующим сенсором. С помощью цифровой и аналоговой электроники эти выходные сигналы поочередно преобразуются в комплексный видеосигнал. Тот факт, что регистрация двумерного сигнала осуществляется дискретно расположенными в пространстве сенсорами, обеспечивает пространственную дискретизацию сигнала; квантование его осуществляется в следующем блоке системы формирования изображения. На выходе ее получается цифровое изображение, схематически показанное на рис. 2.15(д). Преобразование изображения в цифровую форму является темой раздела 2.4.

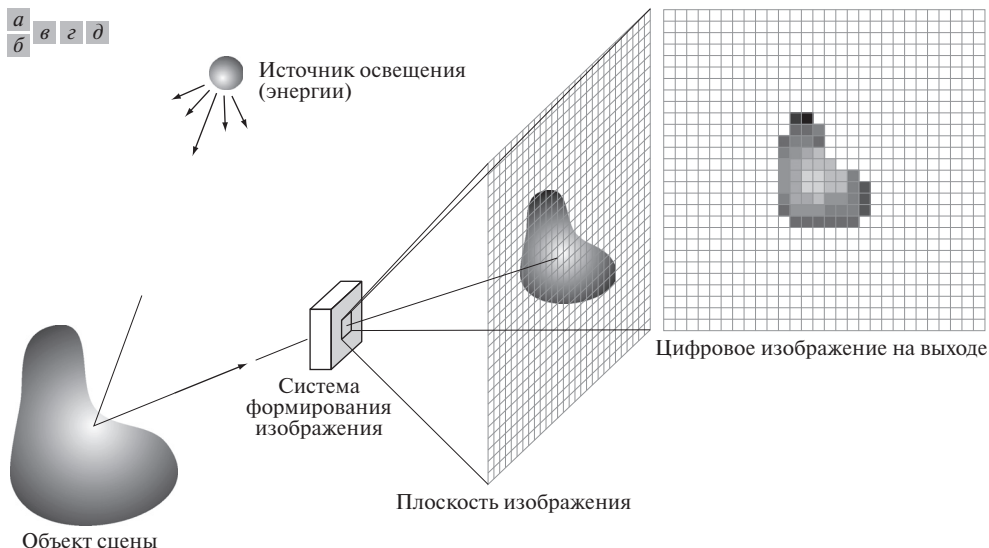


Рис. 2.15. Процесс регистрации цифрового изображения (пример). (а) Источник энергии («освещения»). (б) Элемент сцены. (в) Система формирования изображения. (г) Проекция сцены на плоскость изображения. (д) Оцифрованное изображение

2.3.4. Простая модель формирования изображения

Как уже говорилось в разделе 1.1, мы рассматриваем изображение как двумерную функцию вида $f(x, y)$. Значение функции f в точке с пространственными координатами (x, y) является положительной скалярной величиной, физический смысл которой определяется источником изображения. Если изображение генерируется в результате физического процесса, его значения пропорциональны энергии излучения некоторого физического источника, например энергии электромагнитных колебаний, вследствие чего функция $f(x, y)$ должна быть ненулевой и конечной, т. е.

$$0 < f(x, y) < \infty. \quad (2.3-1)$$

Яркость элементов изображения может принимать отрицательные значения в ходе обработки или в результате интерпретации. Например, в радиолокационных изображениях объекты, движущиеся в направлении локатора, часто интерпретируются как имеющие отрицательную скорость, а объекты, удаляющиеся от локатора, — как имеющие положительную скорость. Таким образом, изображение поля скоростей может кодироваться элементами как с положительными, так и отрицательными значениями. При хранении и визуализации изображений диапазон яркостей обычно преобразуется так, чтобы минимальному отрицательному значению соответствовала нулевая яркость (см. раздел 2.6.3 относительно преобразования масштаба яркостей).

Функцию $f(x, y)$ можно характеризовать двумя компонентами: (1) величиной светового потока, который падает на наблюдаемую сцену от источника,

и (2) относительной долей светового потока, отраженного от объектов этой сцены. Мы будем называть эти компоненты *освещенностью* и *коэффициентом отражения*, обозначая их соответственно $i(x, y)$ и $r(x, y)$. Произведение этих функций дает функцию изображения:

$$f(x, y) = i(x, y) r(x, y), \quad (2.3-2)$$

где

$$0 < i(x, y) < \infty \quad (2.3-3)$$

и

$$0 < r(x, y) < 1. \quad (2.3-4)$$

Соотношение (2.3-4) указывает, что коэффициент отражения может меняться в пределах от 0 (полное поглощение) до 1 (полное отражение). Природа функции $i(x, y)$ зависит от источника освещения, тогда как функция $r(x, y)$ определяется свойствами объектов изображаемой сцены. Примечательно, что приведенные выражения в равной мере применимы также и к изображениям, сформированным в проходящем освещении (сквозь наблюдаемый объект), как, например, при рентгеноскопии грудной клетки. В подобном случае в качестве функции $r(x, y)$ мы имеем дело с *коэффициентом пропускания* вместо коэффициента отражения, но пределы ее изменения будут те же, что и в (2.3-4), и функция изображения формируется по той же модели — как произведение (2.3-2).

Пример 2.1. Некоторые типичные значения освещенности и коэффициента отражения.

- Значения, указанные в соотношениях (2.3-3) и (2.3-4), представляют собой теоретические границы. Ниже приводятся средние числовые значения, иллюстрирующие типичный интервал изменения функции $i(x, y)$ для видимого света. В ясный день солнце создает на земной поверхности освещенность 90000 лм/м² и выше, а в пасмурную погоду эта величина падает до 10000 лм/м². Безоблачной ночью в полнолуние освещенность земной поверхности составляет около 0,1 лм/м². В типичных служебных помещениях поддерживается уровень освещенности порядка 1000 лм/м². Типичные значения коэффициента отражения (т.е. функции $r(x, y)$) составляют: 0,01 для черного бархата; 0,65 для нержавеющей стали; 0,80 для поверхности стены, окрашенной в ровный белый цвет; 0,90 для посеребренной металлической поверхности и 0,93 для снега. ■

Обозначим величину яркости (уровня серого) монохроматического изображения в произвольной точке (x_0, y_0)

$$\ell = f(x_0, y_0). \quad (2.3-5)$$

Из соотношений (2.3-2) — (2.3-4) видно, что ℓ лежит в некотором интервале

$$L_{\min} \leq \ell \leq L_{\max}. \quad (2.3-6)$$

Теоретически к границам этого интервала предъявляются только требования, чтобы L_{\min} было положительно, а L_{\max} — конечно. На практике $L_{\min} = i_{\min} r_{\min}$, а $L_{\max} = i_{\max} r_{\max}$. С учетом вышеприведенных типичных значений освещенности в служебных помещениях и коэффициента отражения можно ожидать типич-

ных пределов $L_{\min} \approx 10$ и $L_{\max} \approx 1000$ для изображений, наблюдаемых в помещениях в отсутствие дополнительного освещения.

Интервал $[L_{\min}, L_{\max}]$ называется *диапазоном яркостей*. На практике его обычно сдвигают по числовой оси, получая интервал $[0, L - 1]$, края которого принимаются за уровень черного ($\ell = 0$) и уровень белого ($\ell = L - 1$). Все промежуточные значения в этом интервале соответствуют некоторым оттенкам серого при изменении от черного до белого.

2.4. Дискретизация и квантование изображения

Из сказанного в предыдущем разделе ясно, что, несмотря на многочисленные возможные способы регистрации изображений, задача всегда одна и та же: сформировать цифровое изображение на основе данных, воспринимаемых чувствительными элементами. От большинства сенсоров поступает аналоговый выходной сигнал в форме непрерывно меняющегося напряжения, форма и амплитуда которого связаны с регистрируемым физическим явлением. Чтобы получить цифровое изображение, необходимо преобразовать непрерывно поступающий сигнал в цифровую форму. Эта операция включает в себя два процесса — *дискретизацию и квантование*.

В этом разделе дискретизация обсуждается на интуитивном уровне. Более глубокое рассмотрение проводится в главе 4.

2.4.1. Основные понятия, используемые при дискретизации и квантовании

Главный принцип, лежащий в основе дискретизации и квантования изображений, проиллюстрирован на рис. 2.16. Здесь (рис. 2.16(a)) приведено исходное изображение $f(x, y)$, которое мы хотим преобразовать в цифровую форму. Изображение непрерывно по координатам x и y , а также по амплитуде. Чтобы преобразовать эту функцию в цифровую форму, необходимо представить ее отсчетами по обеим координатам и по амплитуде. Представление координат в виде конечного множества отсчетов называется *дискретизацией*, а представление амплитуды значениями из конечного множества — *квантованием*.

Изображенная на рис. 2.16(b) одномерная функция представляет собой график изменения значений яркости непрерывного изображения вдоль отрезка AB на рис. 2.16(a). Случайные отклонения на графике вызваны наличием шумов в изображении. Чтобы дискретизовать эту функцию, разобьем отрезок AB на равные интервалы, как показано засечками на рис. 2.16(b) внизу. Значения в выбранных точках отсчета представлены небольшими квадратиками на графике функции. Построенный набор значений в точках дискретизации описывает функцию в виде совокупности ее дискретных отсчетов, однако сами эти значения пока еще охватывают весь непрерывный диапазон яркостей (по вертикали). Чтобы построить цифровую функцию, диапазон яркостей также необходимо преобразовать в дискретные величины (*проквантовать*). Справа

на рис. 2.16(а) изображена шкала яркостей, разбитая на восемь дискретных интервалов от черного до белого. Засечки указывают конкретное значение яркости, присвоенное каждому интервалу. Квантование непрерывных значений яркости в точках дискретизации осуществляется простым сопоставлением каждому отсчету одного из восьми отмеченных значений — того, к которому ближе найденное значение яркости. В результате совместных операций дискретизации и квантования возникает отвечающий одной строке изображения дискретный набор цифровых отсчетов, показанный на рис. 2.16(з). Выполняя такую процедуру построчно, с верхней по нижнюю строки, получаем двумерное цифровое изображение. Рис. 2.16 наводит на мысль, что, помимо числа применяемых дискретных уровней, получаемая точность квантования сильно зависит от шумовой составляющей дискретизованного сигнала.

Выполнение дискретизации описанным выше способом предполагает, что нам доступно непрерывное по обеим координатам и по яркости изображение. На практике, однако, способ оцифровки определяется конструкцией сенсорного устройства, применяемого для регистрации изображения. Если изображение формируется одиночным сенсором в сочетании с механическим его перемещением (рис. 2.13), выходной сигнал сенсора квантуется, как описано выше, а дискретизация определяется выбором шагов механического перемещения сенсора в процессе сбора данных. Механическое перемещение может выполняться с очень

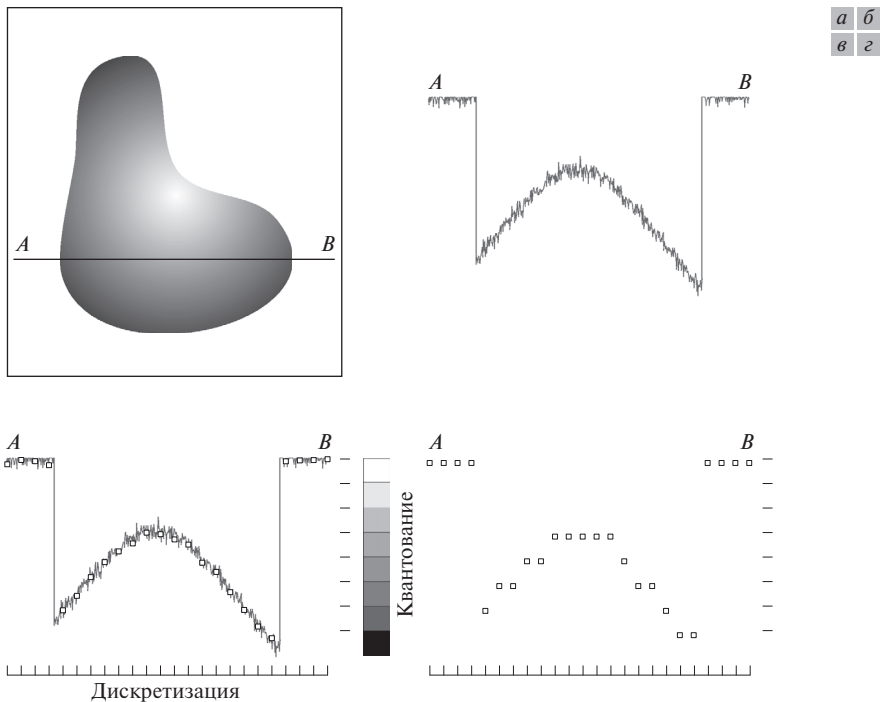


Рис. 2.16. Формирование цифрового изображения. (а) Непрерывное изображение. (б) Профиль вдоль линии сканирования между точками A и B на непрерывном изображении, который используется для иллюстрации понятий дискретизации и квантования. (в) Дискретизация и квантование. (г) Цифровое представление строки изображения

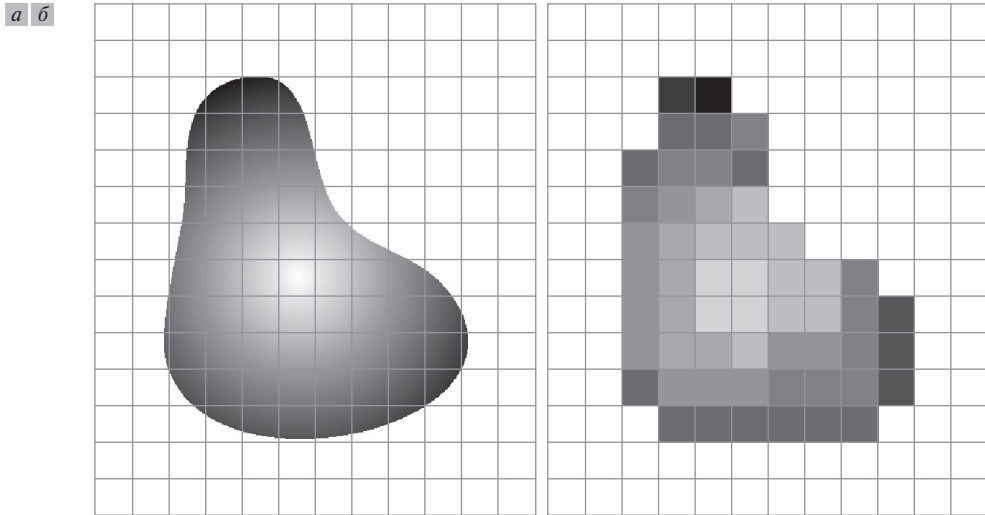


Рис. 2.17. (а) Проекция непрерывного изображения на матрицу чувствительных элементов. (б) Результат дискретизации и квантования изображения

высокой точностью, так что в принципе почти нет пределов для уменьшения шага дискретизации, однако на практике предельная точность дискретизации определяется другими факторами, например качеством оптической системы.

Если для формирования изображения используется линейка сенсоров, то число сенсоров в ней определяет предел дискретизации изображения по одному направлению. Механическим перемещением в другом направлении можно управлять и с более высокой точностью, но нет особого смысла пытаться повышать частоту дискретизации в одном направлении, коль скоро в другом направлении она жестко ограничена. Выходные сигналы всех элементов линейки подвергаются однотипному квантованию, в результате чего строится цифровое изображение.

В случае регистрации изображения с помощью матрицы сенсоров механическое перемещение отсутствует и пределы дискретизации изображения по обоим направлениям определяются числом сенсоров в матрице. Квантование их выходных сигналов осуществляется так же, как и раньше. Рис. 2.17 иллюстрирует этот случай. На рис. 2.17(а) показано непрерывное изображение, спроецированное на плоскость сенсорной матрицы, а рис. 2.17(б) демонстрирует то же изображение после дискретизации и квантования. Ясно, что качество получаемого представления в большой степени зависит от числа отсчетов в разбиении и числа уровней квантования. Однако, как мы увидим в разделе 2.4.3, при выборе значений этих количественных параметров важно учитывать содержательное наполнение изображения.

2.4.2. Представление цифрового изображения

Пусть $f(s, t)$ представляет непрерывную функцию изображения от двух непрерывных переменных s и t . Преобразуем эту функцию в *цифровое изображение* с помощью операций дискретизации и квантования, как описано в предыду-

шем разделе. Предположим, что непрерывное изображение преобразовано в двумерный массив $f(x, y)$ — матрицу из M строк и N столбцов, где (x, y) — дискретные координаты. Для ясности обозначений и большего удобства мы будем использовать для этих координат целочисленные значения: $x = 0, 1, 2, \dots, M - 1$ и $y = 0, 1, 2, \dots, N - 1$, принимая за начало координат левый верхний угол изображения, где $(x, y) = (0, 0)$. Следующим значением координат вдоль первой строки изображения будет точка $(x, y) = (0, 1)$. Важно иметь в виду, что обозначение $(0, 1)$ используется лишь для указания на второй отсчет в первой строке и *не означает*, что это фактические значения физических координат точек дискретизации. В общем случае значение изображения в произвольной координатной точке (x, y) обозначается $f(x, y)$, где x и y — целые числа. Область действительной координатной плоскости, охватываемая координатами изображения, называется *пространственной областью*, а x и y — *пространственными переменными* или *пространственными координатами*.

Как видно из рис. 2.18, есть три основных способа представления $f(x, y)$. На рис. 2.18(a) представлен график функции, у которого две оси определяют поло-

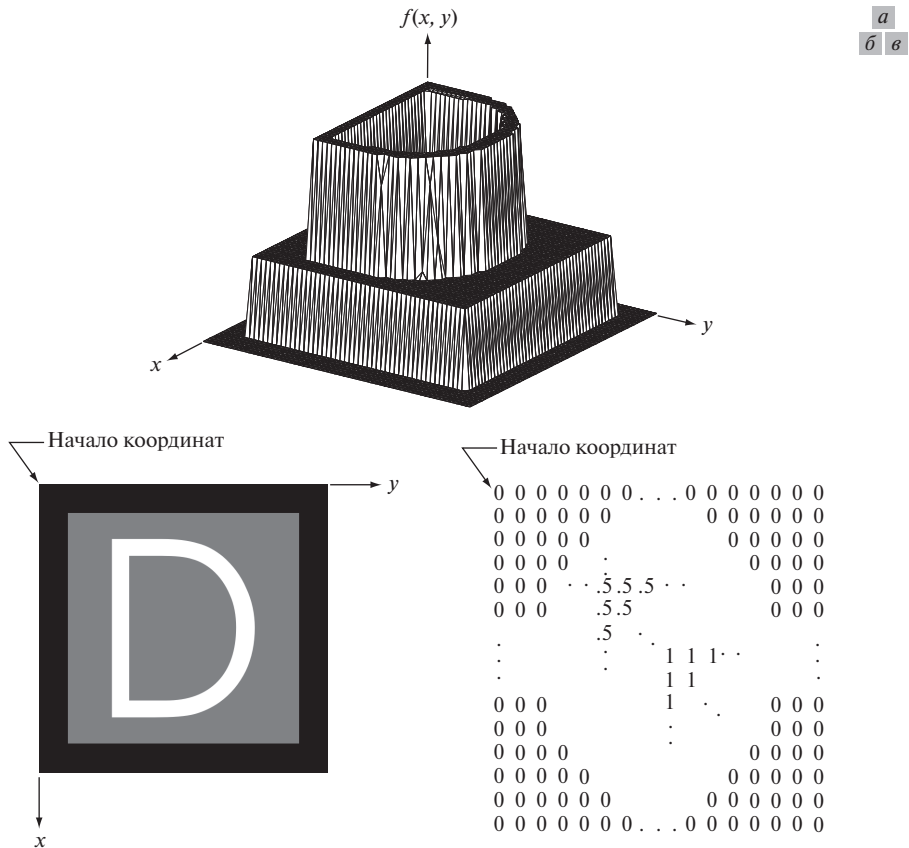


Рис. 2.18. (а) График изображения как объемной поверхности. (б) Отображение в форме массива видимых яркостей. (в) Изображение в виде числовой матрицы (значения 0, 0,5 и 1 представляют соответственно черное, серое и белое)

жение в пространственной области, а на третьей оси откладываются значения яркости f как функции двух пространственных координат x и y . Хотя в этом примере по графику можно представить себе структуру изображения, сложные изображения обычно содержат большое число деталей, которые трудно интерпретировать на основании подобных графиков. Такое представление полезно для работы с множествами полутонов, элементы которых представляются тройками вида (x, y, z) , где x и y — пространственные координаты, а z — значение яркости f в точке с координатами (x, y) . Это представление будет использоваться в разделе 2.6.4.

Представление на рис. 2.18(б) намного привычнее. Оно отображает $f(x, y)$ в таком виде, как изображение выглядит на мониторе или фотографии. Здесь яркость каждой точки пропорциональна значению f в этой точке. На данном рисунке есть только три значения яркости, взятые через равные промежутки. Если нормализовать яркость на интервал $[0, 1]$, то каждая точка изображения имеет одно из значений — 0, 0,5 или 1. Монитор или принтер просто преобразуют эти три значения в черное, серое или белое соответственно, что и показано на рис. 2.18(б). Третье представление изображает численные значения $f(x, y)$ в виде матрицы. В этом примере f имеет размеры 600×600 элементов, т. е. 360 000 чисел. Ясно, что печать всей матрицы была бы громоздкой и малоинформативной. Однако при разработке алгоритмов такое представление вполне полезно, если в числовой форме печатаются и анализируются только фрагменты изображения, как показано на рис. 2.18(в).

Из сказанного можно заключить, что наиболее полезны представления, изображенные на рис. 2.18(б) и (в). Графические дисплеи позволяют мгновенно увидеть результаты. Числовые массивы применяются для обработки изображений и разработки алгоритмов. Представление числовой матрицы размерами $M \times N$ можно записать в форме соотношения

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}. \quad (2.4-1)$$

Обе части этого равенства — эквивалентные способы количественного представления цифрового изображения. В правой части стоит матрица действительных чисел. Каждый элемент этой матрицы называется *элементом изображения* или *пикселем*. Далее повсюду будут употребляться термины *изображение* и *пиксель* для указания на цифровое изображение и его элементы.

В ряде случаев для обозначения цифрового изображения и его элементов бывает полезно использовать более традиционную матричную запись:

$$\mathbf{A} = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}. \quad (2.4-2)$$

Ясно, что $a_{ij} = f(x = i, y = j) = f(i, j)$, поэтому матрицы (2.4-1) и (2.4-2) идентичны. Можно даже представлять изображение как вектор \mathbf{v} , например, формируя вектор-столбец размерами $MN \times 1$ так, что первые M элементов \mathbf{v} совпадают

с первым столбцом матрицы A , следующие M — со вторым столбцом и т. д. Альтернативно для формирования такого вектора мы можем использовать строки матрицы A . Любое представление будет правильным, коль скоро обработка изображения ведется соответственно.

Возвращаясь ненадолго к рис. 2.18, отметим, что начало координат расположено в левом верхнем углу цифрового изображения, ось x направлена вниз, а ось y — вправо. Такое общепринятое представление основано на том факте, что многие графические дисплеи (например телевизионные мониторы) осуществляют развертку изображения, начиная с левого верхнего угла и двигаясь затем вправо вдоль каждой строки. Еще важнее то, что первым элементом матрицы считается элемент в левом верхнем углу, так что выбор в качестве начала координат этой точки является математически осмысленным. Следует иметь в виду, что такое представление является привычной правосторонней² декартовой системой координат, в которой просто оси направлены вниз и вправо, а не вправо и вверх.

Иногда может быть полезно выражать операции дискретизации и квантования в более формальных математических терминах. Пусть Z и R обозначают соответственно множества целых и действительных чисел. Процесс дискретизации можно рассматривать как разбиение плоскости xy на множество ячеек прямоугольной сетки, координаты центра каждой ячейки которой суть элементы декартова произведения Z^2 , т. е. множества всех пар (z_i, z_j) , где z_i и z_j — элементы множества Z . Следовательно, $f(x, y)$ — цифровое изображение, если (x, y) суть целочисленные пары из Z^2 , и функция f приписывает каждой паре координат (x, y) конкретное значение яркости, т. е. действительное число из множества R . Такое сопоставление, осуществляемое функцией f , отвечает описанному выше процессу дискретизации. Если (как обычно в этой и последующих главах) значения яркости также целочисленные и вместо множества R используется Z , тогда цифровое изображение становится двумерной цифровой функцией, у которой как обе координаты, так и значения — целые числа.

Для вычислений или разработки алгоритмов часто бывает полезно отобразить шкалу L значений яркости на интервал $[0, 1]$, так что они перестают быть целочисленными. Однако при хранении и визуализации изображений в большинстве случаев полученные значения масштабируются обратно на интервал целых значений $[0, L - 1]$.

Для выполнения процесса оцифровки изображения необходимо принять решения относительно значений M и N , а также числа дискретных уровней (градаций) яркости L . Для M и N не устанавливаются специальных ограничений помимо того, что они должны быть положительными целочисленными значениями. Однако значение L по соображениям удобства построения оборудования для хранения и квантования обычно выбирают равным целочисленной степени двойки:

$$L = 2^k. \quad (2.4-3)$$

² Напомним, что правосторонней называется такая система координат (x, y, z) , для которой при вытянутом указательном пальце правой руки вдоль оси x и согнутом под прямым углом среднем пальце вдоль оси y большой палец направлен вдоль оси z . Как видно из рис. 2.18(a), система координат изображения именно такая.

Мы предполагаем, что дискретные уровни яркости расположены с постоянным шагом (т.е. используется равномерное квантование) и принимают целые значения в интервале $[0, L - 1]$. Иногда интервал значений яркости неформально называют динамическим диапазоном изображения. Этот термин используется по-разному в различных областях. Здесь мы определяем *динамический диапазон* системы регистрации изображений как отношение максимально возможного значения измеренной яркости к минимальному уровню яркости, обнаруживаемому этой системой. Как правило, верхний предел определяется *насыщением*, а нижний предел — *шумом* (см. рис. 2.19). По существу, динамический диапазон устанавливает минимальный и максимальный уровни яркости, которые система способна представлять и, следовательно, могут присутствовать в изображении. С этим понятием тесно связан *контраст* изображения, который мы определяем как разность между максимальным и минимальным уровнями яркости в данном изображении. Если заметная доля пикселей обладает большим динамическим диапазоном, можно ожидать, что такое изображение имеет высокий контраст. Наоборот, изображение с малым динамическим диапазоном обычно выглядит тусклым, размытым и серым. Более подробно эта тема обсуждается в главе 3.

Общее количество бит b , необходимое для хранения цифрового изображения, определяется по формуле

$$b = M \times N \times k. \quad (2.4-4)$$

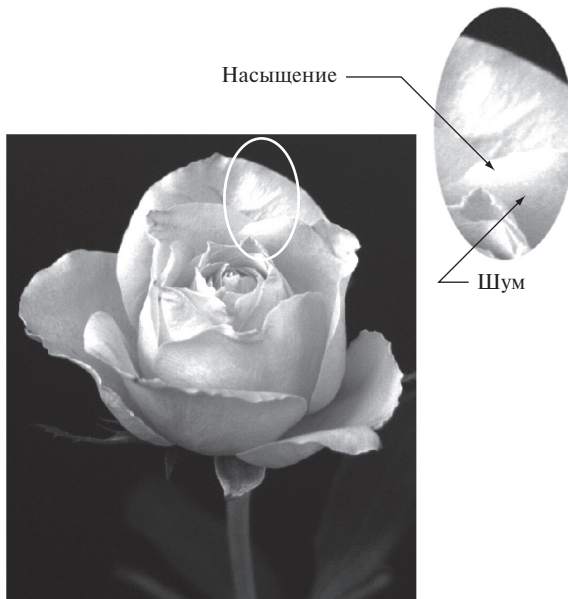


Рис. 2.19. Изображение, в котором присутствуют насыщение и шум. Насыщение определяется тем максимальным уровнем, выше которого все уровни яркости неразличимы (обрезаются). Обратите внимание, что вся область насыщения имеет высокий *постоянный* уровень яркости. Шум в данном случае проявляется в виде зернистой текстуры. Шум, особенно в темных областях изображения (например на стебле розы), маскирует минимально различимый уровень истинной яркости

Таблица 2.1. Число бит для хранения изображения при различных значениях N и k . L — число градаций яркости

N/k	1 ($L=2$)	2 ($L=4$)	3 ($L=8$)	4 ($L=16$)
32	1.024	2.048	3.072	4.096
64	4.096	8.192	12.288	16.384
128	16.384	32.768	49.152	65.536
256	65.536	131.072	196.608	262.144
512	262.144	524.288	786.432	1.048.576
1024	1.048.576	2.097.152	3.145.728	4.194.304
2048	4.194.304	8.388.608	12.582.912	16.777.216
4096	16.777.216	33.554.432	50.331.648	67.108.864
8192	67.108.864	134.217.728	201.326.592	268.435.456
N/k	5 ($L=32$)	6 ($L=64$)	7 ($L=128$)	8 ($L=256$)
32	5.120	6.144	7.168	8.192
64	20.480	24.576	28.672	32.768
128	81.920	98.304	114.688	131.072
256	327.680	393.216	458.752	524.288
512	1.310.720	1.572.864	1.835.008	2.097.152
1024	5.242.880	6.291.456	7.340.032	8.388.608
2048	20.971.520	25.165.824	29.369.128	33.554.432
4096	83.886.080	100.663.296	117.440.512	134.217.728
8192	335.544.320	402.653.184	469.762.048	536.870.912

В случае квадратного изображения $M = N$ и это равенство приобретает вид

$$b = N^2 k. \quad (2.4-5)$$

В табл. 2.1 приводится число бит, необходимых для хранения квадратных изображений при различных значениях N и k . Количество градаций яркости, соответствующее каждому значению k , указано в скобках. Если пиксели изображения могут принимать 2^k значений яркости, то такое изображение часто называют « k -битовым»; например, изображение с возможными 256 градациями яркости называют восьмибитовым. Из таблицы видно, что для хранения 8-битовых изображений размерами 1024×1024 и более элементов требуется существенный объем памяти.

2.4.3. Пространственное и яркостное разрешения

Интуитивно понятно, что пространственное разрешение — это размер мельчайших различимых деталей на изображении. Количественно *пространственное разрешение* можно определять многими способами; к наиболее употребительным относятся *число пар линий на единицу длины* и *число точек (пикселей) на единицу длины*. Предположим, что построен чертеж, состоящий из чередующихся

черных и белых вертикальных линий, каждая шириной W единиц длины (W может быть меньше 1). Таким образом, ширина *пары линий* составляет $2W$, и на единице длины размещается $1/2W$ таких пар. Например, если ширина линии равна 0,1 мм, то на единицу длины (миллиметр) приходится пять пар линий. Широко используемое определение разрешения состоит именно в указании максимального числа *различимых* пар линий на единицу длины; например 100 пар линий на миллиметр. Число точек на единицу длины — мера разрешения изображения, которая повсеместно применяется в полиграфии и издательском деле. В США эту характеристику обычно выражают в *точках на дюйм* (dpi). Представление о качестве дают следующие цифры: газеты печатаются с разрешением 75 dpi, журналы — 133 dpi, глянцевые брошюры — 175 dpi, а эта страница книги (имеется в виду англоязычный оригинал — *прим. ред. перев.*) напечатана с разрешением 2400 dpi.

В предыдущем абзаце главное то, что содержательная мера пространственного разрешения должна определяться по отношению к пространственным единицам измерения. Размеры изображения сами по себе не дают полной информации. Сказать, например, что изображение имеет разрешение 1024×1024 пикселя — не показательно, если при этом не указаны пространственные размеры, занимаемые этим изображением. И размеры изображения сами по себе полезны только для сравнения возможностей получения изображений. Например, можно ожидать, что цифровой фотоаппарат с 20-мегапиксельной светочувствительной ПЗС-матрицей обеспечит более высокую детальность изображения, чем 8-мегапиксельный фотоаппарат, в предположении, что оба аппарата оснащены сравнимыми объективами и сравниваются снимки, сделанные с одинакового расстояния.

Яркостным (или *полутоновым*) *разрешением* аналогично называется мельчайшее различимое изменение яркости. Если при дискретизации изображений имеется относительно большая свобода действий при выборе числа отсчетов (т. е. частоты дискретизации), то при выборе числа градаций яркости приходится в значительной степени учитывать особенности аппаратуры; по этим причинам число градаций обычно выбирается равным степени 2. Наиболее частым решением является выбор 8-битового представления (256 градаций яркости), но в некоторых приложениях используется 16 бит, если необходимо иметь более точное представление полутонов. Квантование уровней яркости с точностью 32 бит используется редко. Иногда можно встретить системы, в которых квантование уровней яркости изображения проводится с точностью 10 или 12 бит, но это скорее исключение, чем правило³. В отличие от пространственного разрешения, которое имеет смысл только если дается в отношении к единице длины, *яркостным разрешением* принято называть число бит, используемых при квантовании величины яркости. Например, об изображении, у которого яркость квантуется на 256 градаций, обычно говорят как об имеющем яркостное разрешение 8 бит. Поскольку фактически различимые изменения яркости зависят

³ Отметим, что *точность квантования* яркости, определяемую используемым АЦП, во многих случаях путают с *точностью представления*, которая в значительной степени зависит от способа хранения. Весьма часто сигнал квантуется на 2^6 , 2^{10} , 2^{12} или 2^{14} градаций, но сохраняется при этом с точностью 8 или 16 бит соответственно. — *Прим. ред. перев.*

не только от уровней шума и насыщения, но и от возможностей человеческого восприятия (см. раздел 2.1), то заявление, что изображение имеет яркостное разрешение 8 бит, означает лишь то, что формирующая его система способна квантовать величину яркости через равные ступени, каждая величиной в $1/256$ максимальной амплитуды яркости.

Следующие два примера иллюстрируют влияние по отдельности пространственного и яркостного разрешений на степень детализации изображения. Позже в этом разделе мы обсудим, как эти два параметра совместно определяют осязаемое качество цифрового изображения.

Пример 2.2. Демонстрация эффектов при уменьшении пространственного разрешения цифрового изображения.

■ Рис. 2.20 демонстрирует эффекты, возникающие в результате снижения пространственного разрешения цифрового изображения. Изображения рис. 2.20(а)—(г) приведены с разрешением 1250, 300, 150 и 72 dpi соответственно. Естественно, по числу элементов изображения с более низким разрешением меньше оригинала. Например, изображение при 72 dpi представляет собой матрицу размерами 213×162 , тогда как размеры оригинала — 3692×2812 пикселей. Для большей наглядности сравнения все изображения с меньшими размерами были увеличены до размеров оригинала (метод увеличения рассматривается в разделе 2.4.4). В некотором смысле это эквивалентно рассмотрению уменьшенных изображений с более близкого расстояния, чтобы объективно судить о сравнительной степени их детальности.

Между рис. 2.20(а) и (б) имеются небольшие видимые отличия, наиболее заметные в виде легких искажений формы большой черной стрелки, но по большей части изображение на рис. 2.20(б) вполне приемлемо. Фактически 300 dpi — это типичное минимальное пространственное разрешение в книгоиздании, так что здесь не следовало ожидать больших видимых отличий. На рис. 2.20(в) начинает проявляться ухудшение визуального качества изображения (например вид круглых краев хронометра и малой стрелки в правой части, которая указывает на 60). Рис. 2.20(г) демонстрирует заметное ухудшение большинства деталей изображения. Как говорится в разделе 4.5.4, при печати с таким низким разрешением в полиграфии применяется ряд хитростей (таких как локальное изменение размеров печатной точки), которые позволяют получить значительно лучшие результаты, чем показаны на рис. 2.20(г). Кроме того, как мы увидим в разделе 2.4.4, приведенные на рис. 2.20 результаты можно улучшить за счет выбора метода интерполяции. ■

Пример 2.3. Типичные эффекты при изменении числа градаций яркости в цифровом изображении.

■ В этом примере мы сохраняем число отсчетов дискретизации постоянным, но постепенно уменьшаем число уровней квантования с 256 до 2, двигаясь по степеням 2. На рис. 2.21(а) показан полученный методом проекционной КТ рентгеновский снимок размерами 452×374 пикселей с 256 градациями яркости ($k = 8$). Подобные изображения получаются при закреплении рентгеновского источника в некотором положении, что дает двумерное изображение в нужном

<i>a</i>	<i>б</i>
<i>в</i>	<i>г</i>



Рис. 2.20. Типичные эффекты при уменьшении пространственного разрешения. Приведены изображения с разрешением (а) 1250 dpi, (б) 300 dpi, (в) 150 dpi и (г) 72 dpi. Тонкие рамки добавлены для удобства, они не являются частью изображений

направлении. Проекционные КТ-изображения используются для настройки параметров компьютерного томографа, например угла наклона, количества срезов и диапазона.

Изображения на рис. 2.21(б—з) были получены путем уменьшения числа бит представления с $k = 7$ до $k = 1$ при сохранении постоянного пространственного разрешения 452×374 пикселей. Изображения с 256, 128 и 64 градациями яркости

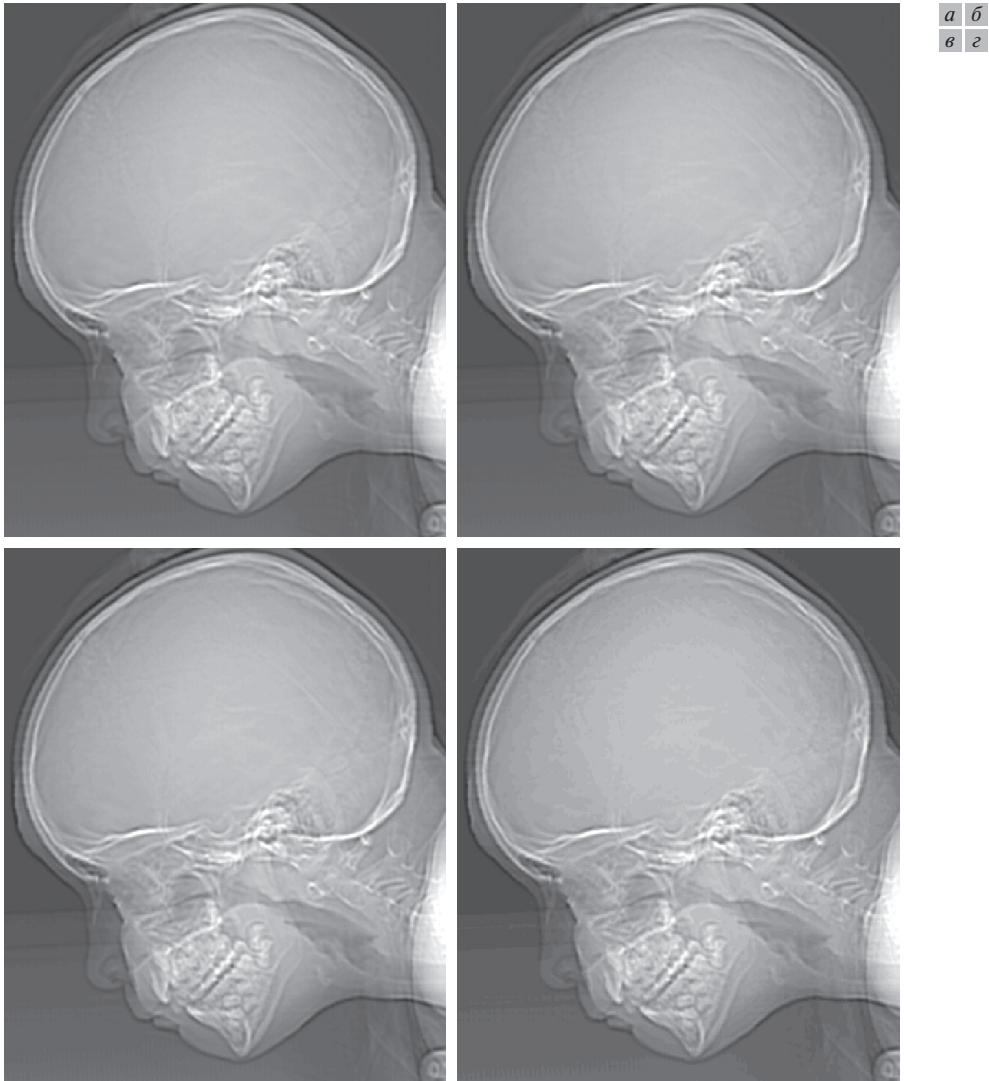


Рис. 2.21. (а) Изображение размерами 452×374 пикселей с 256 градациями яркости. (б)—(г) Изображения того же пространственного разрешения, представленные с 128, 64 и 32 градациями яркости

оказываются визуально неотличимыми и в равной мере применимыми. Однако на 32-градационном изображении (рис. 2.21(г)) появляются почти незаметные мелкие рубчатые структуры в области постоянной или мало меняющейся яркости, в частности внутри черепной коробки. Этот эффект, вызванный использованием недостаточного числа градаций яркости в областях плавных переходов полутонов, называется появлением *ложных контуров*, поскольку эти линии напоминают контурные линии горизонталей на топографической карте. Ложные контуры становятся отчетливо видны на изображениях, представленных с равномерным разбиением всего яркостного диапазона на 16 градаций и менее, как видно из рис. 2.21(д—э).

d	e
$ж$	$з$

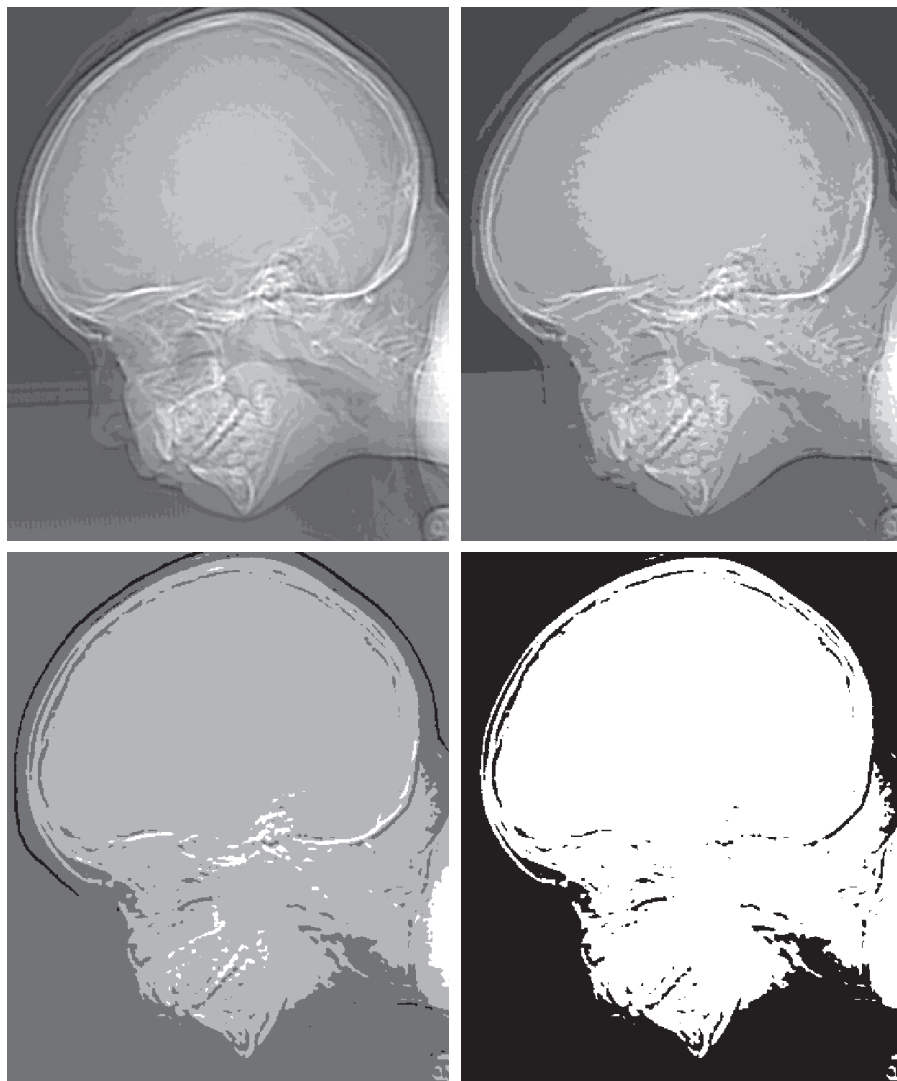


Рис. 2.21. (Продолжение.) (д)—(з) Изображения, представленные с 16, 8, 4 и 2 градациями яркости. (Исходное изображение предоставил д-р Дэвид Р. Пикенс, департамент радиологии медицинского центра университета Вандербилта)

В качестве очень грубого эмпирического правила можно считать, что минимальные пространственное и яркостное разрешение, при котором цифровое изображение, напечатанное с размерами порядка 5×5 см, будет относительно свободным от дефектов типа ложных контуров и ступенчатости, составляет около 256×256 пикселей с 64 градациями яркости. ■

Результаты, показанные на примерах 2.2 и 2.3, иллюстрируют влияние независимого изменения значений N и k на качество изображения. Но это дает лишь частичный ответ на вопрос о влиянии N и k на свойства цифровых изображений,



Рис. 2.22. (а) Изображение с малым числом деталей. (б) Изображение со средним уровнем детализации. (в) Изображение с относительно большим количеством деталей. (Изображение (б) предоставлено Массачусетским технологическим институтом)

поскольку мы пока не рассматривали возможные взаимосвязи между этими двумя параметрами. В ранних исследованиях [Huang, 1965] делалась попытка измерить экспериментально, как влияет на качество изображения одновременное изменение значений N и k . Эксперимент состоял из набора субъективных тестов, в которых использовались изображения, подобные приведенным на рис. 2.22. Женское лицо относится к классу изображений с относительно малым содержанием деталей, снимок толпы, напротив, содержит большое число деталей, а изображение кинооператора занимает промежуточное положение по степени наличия деталей.

Генерировался набор изображений указанных трех классов для различных значений N и k , после чего наблюдателю предлагалось упорядочить их в соответствии

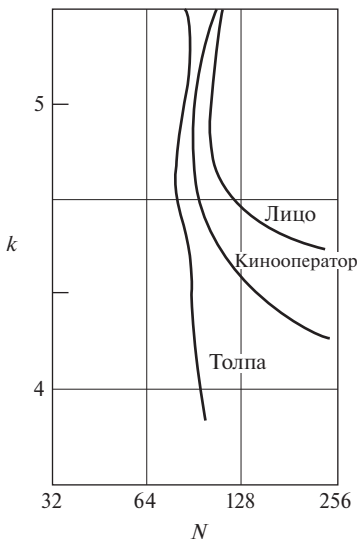


Рис. 2.23. Образцы кривых равного предпочтения для трех типов изображений, показанных на рис. 2.22

с субъективно ощущаемым качеством. Результаты эксперимента представлялись в форме так называемых *кривых равного предпочтения* на плоскости Nk (рис. 2.23 демонстрирует типичные образцы кривых равного предпочтения для изображений, приведенных на рис. 2.22). Каждая точка на плоскости Nk соответствует изображению со значениями параметров N и k , равными координатам этой точки. Точки, лежащие на кривой равного предпочтения, соответствуют изображениям одинакового субъективно воспринимаемого качества. В описываемых экспериментах было обнаружено, что кривые равного предпочтения имеют тенденцию к смещению вправо и вверх, однако форма их для каждого из трех рассматриваемых классов изображений отличается и в целом подобна изображенной на рис. 2.23. В целом результаты не являются неожиданными, поскольку сдвиг кривой вправо и вверх означает просто увеличение значений N и k , что неминуемо ведет к повышению качества изображения.

В контексте нашего обсуждения наиболее важным представляется тот факт, что кривые равного предпочтения приобретают все более вертикальную форму по мере увеличения степени детализации изображения. Такой результат наводит на мысль, что для изображений с большим количеством деталей может оказаться достаточным лишь небольшое число градаций яркости. Например, приведенная на рис. 2.23 кривая равного предпочтения для изображения толпы почти вертикальна. Это показывает, что для фиксированного значения N воспринимаемое качество изображения такого типа почти не зависит от числа использованных градаций яркости (в показанном на рис. 2.23 диапазоне). Интересно также отметить, что для других двух типов изображений визуальное качество остается одинаковым в некоторых интервалах, где пространственное разрешение увеличивается при одновременном уменьшении числа градаций яркости. Наиболее вероятная причина такого результата состоит в том, что при уменьшении k повышается видимый контраст изображения, а этот эффект часто субъективно воспринимается человеком как улучшение качества изображения.

2.4.4. Интерполяция цифрового изображения

Интерполяция является основным инструментом, широко используемым при решении таких задач, как увеличение и уменьшение изображений, их поворот и коррекция геометрических искажений. Главная цель этого раздела — ввести в рассмотрение интерполяцию и ее использование для изменения (уменьшения или увеличения) размеров цифрового изображения, что по существу означает *повторную дискретизацию (передискретизацию)* исходного изображения. Применение интерполяции в таких приложениях, как поворот и коррекция геометрических искажений, рассматривается в разделе 2.6.5. Мы также вернемся к этой теме в главе 4, где повторная дискретизация цифрового изображения обсуждается более подробно.

В своей основе *интерполяция* — это процесс, при котором имеющиеся данные используются для оценки значений в неизвестных точках. Начнем обсуждение этой темы с простого примера. Предположим, что имеется изображение размерами 500×500 пикселей, которое необходимо увеличить в 1,5 раза до размеров 750×750 пикселей. Простейший способ визуально наблюдать увеличение — это вообразить сетку размерами 750×750 элементов с таким же шагом, что и у пикселей исходного изображения, и затем сжимать ее вплоть до точного совпадения краев с краями исходного изображения. Очевидно, что шаг сжатой сетки 750×750 будет меньше одного пикселя исходного изображения. Чтобы присвоить значение яркости любому элементу наложенного изображения, найдем ближайший к нему пиксель исходного изображения и припишем его яркость данному элементу сетки 750×750 . Присвоив так значения яркости всем элементам наложенной сетки, после растяжения ее до первоначальных размеров получим увеличенное цифровое изображение.

Такой способ называется *интерполяцией по ближайшему соседу*, поскольку в нем каждому пикселю нового изображения присваивается значение яркости ближайшего соседнего пикселя в исходном изображении (вопросы соседства пикселей рассматриваются в разделе 2.5). Этот подход отличается простотой,

но, как мы увидим позже в данном разделе, он часто приводит к появлению на изображении нежелательных артефактов вроде значительного искажения прямолинейных контуров и поэтому редко применяется на практике. Более подходящим методом является *билинейная интерполяция*, в которой для присвоения яркостей элементам нового изображения используются четыре ближайших соседа данной точки. Пусть (x, y) — координаты точки нового изображения, которой нужно присвоить значение яркости (эту точку можно представить себе как точку ранее упоминавшейся наложенной сетки). Обозначим этот уровень яркости $v(x, y)$. В случае билинейной интерполяции он задается соотношением

$$v(x, y) = ax + by + cxy + d, \quad (2.4-6)$$

где коэффициенты a, b, c, d находятся из системы четырех линейных уравнений с четырьмя неизвестными, выписанной для четырех ближайших соседей точки (x, y) . Как мы скоро увидим, билинейная интерполяция дает значительно лучшие результаты, чем интерполяция по ближайшему соседу, ценой умеренного усложнения вычислений.

Заметим, что, вопреки названию, билинейная интерполяция *не является* линейной ввиду наличия члена xy , однако она линейна отдельно по каждой переменной при фиксированном значении другой переменной.

Следующей по уровню сложности является *бикубическая интерполяция*, в которой участвуют шестнадцать ближайших соседей рассматриваемой точки. Значение яркости, которое будет приписано точке (x, y) , определяется соотношением

$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j, \quad (2.4-7)$$

где шестнадцать коэффициентов a_{ij} находятся из системы 16 уравнений с 16 неизвестными, выписанной для шестнадцати ближайших соседей точки (x, y) . Заметим, что соотношение (2.4-7) сводится к виду (2.4-6), если в обеих суммах стоят пределы суммирования от 0 до 1. В общем случае при бикубической интерполяции мелкие детали изображения сохраняются лучше, чем при билинейной, почему первая стандартно используется в коммерческих программах обработки изображений, таких как Adobe Photoshop и Corel Photopaint.

Пример 2.4. Сравнение методов интерполяции для целей увеличения и уменьшения изображений.

- На рис. 2.24(а) снова повторено изображение из рис. 2.20(э), которое было получено уменьшением разрешения 1250 dpi исходного изображения на рис. 2.20(а) до 72 dpi (в результате чего размеры изображения уменьшились с 3692×2812 до 213×162 пикселей) и последующим увеличением при печати до первоначальных размеров. Чтобы получить изображение рис. 2.20(э), в обоих случаях, при уменьшении и увеличении, использовалась интерполяция по ближайшему соседу. Как уже говорилось, показанный на рис. 2.24(а) результат довольно плохого качества. На рис. 2.24(б) и (в) приведены результаты аналогичной процедуры, но с использованием соответственно билинейной и бикубической

<i>a</i>	<i>б</i>	<i>в</i>
<i>г</i>	<i>д</i>	<i>е</i>



Рис. 2.24. (а) Изображение, уменьшенное до разрешения 72 dpi и затем увеличенное до исходных размеров 3692×2812 пикселей с помощью интерполяции по ближайшему соседу. Это изображение повторяет рис. 2.20(г). (б) Изображение, уменьшенное и увеличенное с помощью билинейной интерполяции. (в) То же, что и (б), но с использованием бикубической интерполяции. (г)—(е) Аналогичная последовательность для случая, когда разрешение уменьшается до 150 dpi вместо 72 dpi (рис. 2.24(г) совпадает с рис. 2.20(в)). Сравните рис. 2.24(д) и (е), особенно последний, с исходным изображением на рис. 2.20(а)

интерполяции при уменьшении и увеличении. Результат, полученный с помощью билинейной интерполяции, значительно лучше по сравнению с интерполяцией по ближайшему соседу. Резкость изображения, полученного методом бикубической интерполяции, несколько выше, чем при билинейной интерполяции. На рис. 2.24(г) снова повторено изображение из рис. 2.20(в), полученное из исходного путем снижения разрешения до 150 dpi методом интерполяции по ближайшему соседу (при уменьшении и при увеличении). Обсуждая выше этот рисунок, мы говорили, что при уменьшении разрешения до 150 dpi начинает ухудшаться визуальное качество изображения. Рис. 2.24(д) и (е) демонстрируют результаты использования соответственно билинейной и бикубической ин-

терполяции. Несмотря на снижение исходного разрешения 1250 dpi до 150 dpi, эти два последние изображения вполне сопоставимы с оригиналом, что говорит о мощи двух обсуждаемых методов интерполяции. Как и прежде, результаты бикубической интерполяции отличаются большей резкостью. ■

Для интерполяции может использоваться большее число соседей, и существуют более сложные методы, например с использованием сплайнов и вейвлетов, что в некоторых случаях может дать лучшие результаты, чем описанные методы интерполяции. Хотя сохранение мелких деталей оказывается исключительно важным соображением при генерации изображений в трехмерной компьютерной графике [Watt, 1993; Shirley, 2002] и при обработке медицинских изображений [Lehmann et al., 1999], однако ввиду повышенной вычислительной сложности применение такого подхода в универсальных системах обработки изображений редко бывает оправданным, и обычно для операций увеличения и уменьшения ограничиваются билинейной или бикубической интерполяцией.

2.5. Некоторые фундаментальные отношения между пикселями

В этом разделе мы рассмотрим некоторые важные взаимосвязи между элементами цифрового изображения. Как указывалось выше, мы будем обозначать изображение в виде функции $f(x, y)$. Ссылаясь в пределах этого раздела на конкретные пиксели, мы будем пользоваться строчными буквами, например p и q .

2.5.1. Соседи отдельного элемента

У элемента изображения p с координатами (x, y) имеются четыре соседа по *вертикали* и *горизонтали*, координаты которых даются выражениями

$$(x+1, y), (x-1, y), (x, y+1), (x, y-1).$$

Это множество пикселей называется *четверкой соседей* p и обозначается $N_4(p)$. Каждый его элемент находится на единичном расстоянии от (x, y) ; если же точка (x, y) лежит на краю изображения, то некоторые из соседей оказываются за пределами изображения. Этот вопрос обсуждается в главе 3.

Четыре соседа p по *диагонали* имеют координаты

$$(x+1, y+1), (x+1, y-1), (x-1, y+1), (x-1, y-1)$$

и обозначаются $N_D(p)$. Вместе с четверкой соседей эти точки образуют так называемую *восьмерку соседей*, обозначаемую $N_8(p)$. Как и выше, некоторые точки множеств $N_D(p)$ и $N_8(p)$ могут оказаться за пределами изображения, если точка (x, y) лежит на его краю.

2.5.2. Смежность, связность, области и границы

Пусть V — множество значений яркости, используемое при определении понятия смежности. В бинарном изображении $V = \{1\}$, если смежными счита-

ются соседние пиксели с единичным значением яркости. Для полутоновых изображений идея та же, но множество V обычно состоит из большего числа элементов. Например, при определении понятия смежности для пикселей с диапазоном возможных значений яркости от 0 до 255 множество V может быть любым подмножеством этих 256 значений. Мы будем рассматривать три вида смежности:

1) *4-смежность*. Два пикселя p и q со значениями из множества V являются 4-смежными, если q входит в множество $N_4(p)$;

2) *8-смежность*. Два пикселя p и q со значениями из множества V являются 8-смежными, если q входит в множество $N_8(p)$;

3) *m -смежность (mixed, смешанная)*. Два пикселя p и q со значениями из множества V являются m -смежными, если:

а) элемент q входит в множество $N_4(p)$ или

б) элемент q входит в множество $N_D(p)$ и множество $N_4(p) \cap N_4(q)$ не содержит элементов изображения со значением яркости из множества V .

Мы используем символы \cap и \cup для обозначения соответственно пересечения и объединения множеств. Напомним, что пересечением множеств A и B называется множество, элементы которого входят одновременно в A и B , а объединение этих двух множеств состоит из элементов, принадлежащих или множеству A , или B , или одновременно обоим множествам. Множества и операции над ними подробнее рассматриваются в разделе 2.6.4.

Смешанная смежность представляет собой модификацию 8-смежности с целью исключения неоднозначности, часто возникающей при использовании 8-смежности в чистом виде. Рассмотрим, например, изображенную на рис. 2.25(а) конфигурацию пикселей при $V = \{1\}$. Три элемента в верхней части рис. 2.25(б) демонстрируют неоднозначную 8-смежность, как указано пунктирными линиями. Эта неоднозначность устраняется при использовании m -смежности, что иллюстрирует рис. 2.25(в).

Дискретным путем (или *кривой*) от пикселя p с координатами (x, y) до пикселя q с координатами (s, t) называется неповторяющаяся последовательность пикселей с координатами

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n),$$

где $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$ и пиксели (x_i, y_i) и (x_{i-1}, y_{i-1}) являются смежными при $1 \leq i \leq n$. В этом случае n называется *длиной* пути. Если $(x_0, y_0) = (x_n, y_n)$, то путь называется *замкнутым*. Можно определить 4-, 8- или m -*путь* в соответствии с заданным типом смежности. Например, на рис. 2.25(б) изображено два 8-пути между правым верхним и правым нижним элементами, а на рис. 2.25(в) показан m -путь.

Пусть S — некоторое подмножество элементов изображения. Два его элемента p и q называются *связными в S* , если между ними существует путь, целиком состоящий из элементов подмножества S . Для любого пикселя p из S множество всех пикселей, связанных с ним в S , называется *связной компонентой* (или *компонентой связности*) S . Если множество S содержит только одну компоненту связности, оно называется *связным множеством*.

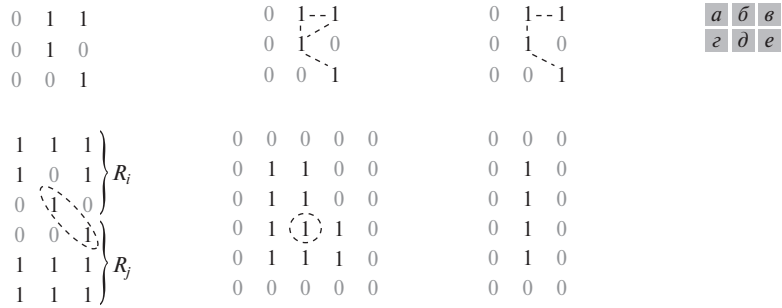


Рис. 2.25. (а) Конфигурация пикселей. (б) Элементы, являющиеся 8-смежными между собой (показано пунктиром; видна неоднозначность). (в) Отношения m -смежности. (г) Две области единичных пикселей, являющиеся смежными, если используется 8-смежность. (д) Элемент в кружке является частью границы области единичных пикселей, если только между областью и фоном рассматривается 8-смежность. (е) Внутренняя граница области с единичными значениями не образует замкнутого пути, а ее внешняя граница — образует

Пусть R — подмножество элементов изображения. Будем называть его *областью*, если R — связное множество. Две области R_x и R_y называются *смежными*, если их объединение является связным множеством. Области, не являющиеся смежными, называются *раздельными (несмежными)*. При рассмотрении областей подразумевают либо 4-, либо 8-смежность. Чтобы такие определения были содержательными, необходимо указать используемый тип смежности. Например, две области единичных пикселей на рис. 2.25(г) являются смежными только в смысле 8-смежности (в соответствии с данным выше определением, не существует 4-пути между этими областями, и потому их объединение не есть связное множество).

Предположим, что изображение содержит K несмежных областей R_k , $k = 1, 2, \dots, K$, ни одна из которых не примыкает к краю изображения⁴. Обозначим R_u объединение этих K областей, а $(R_u)^c$ — его дополнение (напомним, что *дополнением* множества S называется множество точек, не принадлежащих S). Назовем совокупность точек R_u *интересующей областью* изображения, а совокупность точек $(R_u)^c$ — его *фоновой областью (фоном)*.

Границей области R (также называемой *замкнутым контуром* или *краем*) назовем множество точек этой области, которые являются смежными к точкам дополнения R . Иначе говоря, граница области — это множество пикселей этой области, у которых один или более соседей принадлежат фону. Опять-таки здесь необходимо еще указать вид связности, используемый для определения смежности. Например, точка, обведенная кружком на рис. 2.25(д), не является частью границы области единичных пикселей, если при рассмотрении смежности интересующей области и фона используется 4-связность. Во избежание

⁴ Такое предположение делается, чтобы избежать рассмотрения особых случаев. Потери общности при этом не происходит, так как если одна или более областей примыкают к краям изображения, мы всегда можем дополнить изображение рамкой шириной в один пиксель с фоновым значением яркости.

подобных ситуаций, как правило, смежность точек области и фона определяют в терминах 8-связности.

Предыдущее определение иногда называют определением *внутренней границы* области, в отличие от *внешней границы*, определяемой аналогичным образом для фона. Такое различие важно при разработке алгоритмов, в которых прослеживаются границы. Обычно такие алгоритмы строятся для обхода внешней границы, чтобы получить в результате замкнутый путь. Например, внутренняя граница области пикселей с единичным значением на рис. 2.25(e) совпадает с самой областью. Такая граница не удовлетворяет данному выше определению замкнутого пути. С другой стороны, внешняя граница этой области образует замкнутый путь вокруг нее.

Если подмножество R есть все изображение (которое, напомним, является прямоугольной матрицей элементов), то его граница по определению состоит из элементов первого столбца и первой строки, а также элементов последнего столбца и последней строки. Такое доопределение необходимо, поскольку у этих элементов нет соседей за пределами изображения. Обычно, говоря об области, имеют в виду подмножество всего изображения, а элементы на границе области, которые совпадают с краем изображения, безусловно включаются в состав границы этой области.

При обсуждении областей и границ часто возникает понятие *контура*. Между контуром (в общем случае незамкнутым) и границей существует принципиальная разница. Граница конечной области всегда образует замкнутый путь и поэтому является «глобальным» понятием. Контур же, как подробно обсуждается в главе 10, состоит из пикселей, на которых значения производной яркости превышают заранее заданный порог. Поэтому по самой своей идее контур является «локальным» понятием, основанным на мере непрерывности уровня яркости в некоторой точке. Контурные точки могут соединяться, образуя сегменты контуров, и эти сегменты иногда соединяются подобно границам, но такой случай имеет место не всегда. Единственным исключением, когда контуры и границы соответствуют друг другу, являются бинарные изображения. В зависимости от используемых вида связности и оператора выделения контуров (эти операторы рассматриваются в главе 10), выделение контуров в бинарной области дает результаты, совпадающие с границей этой области. Ограничимся здесь этим изложением на интуитивном уровне и, пока мы не достигли главы 10, будем понимать контуры как разрывы яркости, а границы — как замкнутые пути.

2.5.3. Меры расстояния

Пусть элементы изображения p , q и z имеют координаты (x, y) , (s, t) и (v, w) соответственно. Функция D называется *функцией расстояния* или *метрикой*, если:

- а) $D(p, q) \geq 0$, причем $D(p, q) = 0$ тогда и только тогда, когда $p = q$,
- б) $D(p, q) = D(q, p)$,
- в) $D(p, z) \leq D(p, q) + D(q, z)$.

Евклидово расстояние (метрика L_2) между элементами p и q определяется следующим образом:

$$D_e(p, q) = \left[(x-s)^2 + (y-t)^2 \right]^{\frac{1}{2}}. \quad (2.5-1)$$

При такой метрике пиксели, находящиеся на расстоянии не более r от заданной точки (x, y) , образуют круг радиуса r с центром в этой точке.

Расстояние D_4 (метрика L_1) между элементами p и q определяется следующим образом:

$$D_4(p, q) = |x - s| + |y - t|. \tag{2.5-2}$$

В этом случае пиксели, находящиеся на расстоянии D_4 , меньшем или равном r , от заданной точки (x, y) , образуют повернутый на 45° квадрат с центром в этой точке. Например, пиксели с расстоянием $D_4 \leq 2$ от центральной точки образуют следующие замкнутые линии равных расстояний:

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

Пиксели с расстоянием $D_4 = 1$ являются четверкой соседей для элемента (x, y) .

Расстояние D_8 (метрика L_∞) между элементами p и q определяется следующим образом:

$$D_8(p, q) = \max(|x - s|, |y - t|). \tag{2.5-3}$$

В этом случае пиксели, находящиеся на расстоянии D_8 , меньшем или равном r , от заданной точки (x, y) , образуют квадрат с центром в этой точке. Например, пиксели с расстоянием $D_8 \leq 2$ от центральной точки образуют следующие замкнутые линии равных расстояний:

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

Пиксели с расстоянием $D_8 = 1$ являются восьмеркой соседей для элемента (x, y) .

Заметим, что расстояния D_4 и D_8 между двумя элементами p и q не зависят от каких-либо путей, которые могли существовать между этими пикселями, поскольку в определении этих расстояний участвуют только координаты элементов. Однако, если мы выбираем в качестве меры m -смежность, то расстояние D_m между двумя элементами изображения определяется как длина кратчайшего m -пути между этими элементами. В этом случае расстояние между пикселями будет зависеть от значений всех пикселей вдоль этого пути, равно как и от значений их соседей. Например, рассмотрим следующую конфигурацию пикселей, и пусть элементы p, p_2 и p_4 имеют значение 1, а элементы p_1 и p_3 могут принимать значения 0 или 1:

$$\begin{array}{cc} p_3 & p_4 \\ p_1 & p_2 \\ p & \end{array}$$

Предположим, что рассматривается смежность пикселей со значением 1, т. е. $V = \{1\}$. Если оба элемента p_1 и p_3 имеют значения 0, то длина кратчайшего m -пути (т. е. расстояние D_m) между p и p_4 равна 2. Если значение p_1 равно 1, то элементы p и p_2 больше не являются m -смежными (см. определение отношения m -смежности) и длина кратчайшего m -пути становится равной 3 (этот путь проходит через точки p, p_1, p_2, p_4). Аналогичные рассуждения имеют место в том случае, если значение p_3 равно 1 (а значение p_1 равно 0). В этом случае длина кратчайшего m -пути также равна 3. Наконец, если оба пикселя p_1 и p_3 имеют единичные значения, то длина кратчайшего m -пути между p и p_4 станет равной 4. В таком случае путь проходит через последовательность точек p, p_1, p_2, p_3, p_4 .

2.6. Введение в математический аппарат, применяемый в цифровой обработке изображений

Этот раздел преследует две главные цели: (1) познакомить читателя с различными математическими инструментами, которые используются на протяжении всей книги и (2) помочь читателю ощутить, как именно этот аппарат используется, применяя его в разнообразных простых задачах обработки изображений, часть которых будет неоднократно появляться в дальнейших обсуждениях. По мере необходимости область применения этих инструментов будет расширяться в последующих главах.

Перед тем, как двигаться дальше, читателю может быть полезно загрузить и изучить обзорный материал из раздела «Обучающие материалы» на сайте книги в Интернете. Этот обзор содержит вводный материал о матрицах и векторах, линейных системах, теории множеств и теории вероятности.

2.6.1. Поэлементные и матричные операции

Поэлементные операции, в которых участвуют одно или более изображений, всегда выполняются *попиксельно* над соответствующими элементами изображений. Ранее в этой главе упоминалось, что изображения можно также эквивалентно рассматривать как матрицы. И в самом деле, во многих случаях операции над изображениями выполняются по правилам матричной алгебры (см. раздел 2.6.6). Именно по этой причине необходимо четко разграничить поэлементные и матричные операции. Рассмотрим, например, следующие два изображения размерами 2×2 :

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ и } \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Поэлементное произведение этих двух изображений вычисляется так:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}.$$

Напротив, *матричное произведение* изображений определяется выражением

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}.$$

В последующем на протяжении книги мы всюду подразумеваем, что операции выполняются попиксельно, если не оговорено иное. Например, говоря о возведении цифрового изображения в степень, мы имеем в виду, что значение каждого пикселя по отдельности возводится в эту степень; говоря об операции деления одного изображения на другое, мы на самом деле подразумеваем, что деление выполняется для соответственных пикселей двух изображений и т. д.

2.6.2. Линейные и нелинейные преобразования

Одна из важнейших характеристик любого метода обработки изображений — это является ли он линейным или нелинейным. Рассмотрим оператор общего вида H , который строит выходное изображение $g(x, y)$ для данного входного изображения $f(x, y)$:

$$H[f(x, y)] = g(x, y). \quad (2.6-1)$$

Говорят, что оператор H *линейный*, если

$$H[a_i f_i(x, y) + a_j f_j(x, y)] = a_i H[f_i(x, y)] + a_j H[f_j(x, y)] = a_i g_i(x, y) + a_j g_j(x, y), \quad (2.6-2)$$

где $f_i(x, y)$ и $f_j(x, y)$ — любые изображения одинаковых размеров, а a_i и a_j — произвольные константы. Соотношение (2.6-2) показывает, что результат применения линейного оператора к сумме двух входных изображений совпадает с суммой результатов применения такого оператора к этим изображениям по отдельности. А также результат применения линейного оператора к изображению, умноженному на константу, идентичен умножению на эту константу результата применения оператора к исходному изображению. Первое свойство называется *аддитивностью*, а второе — *однородностью*.

Рассмотрим простой пример, где в качестве H используется оператор суммы Σ , функция которого состоит просто в суммировании его входов. Для проверки линейности этого оператора начнем с левой части (2.6-2) и попытаемся доказать, что она равна правой части:

$$\begin{aligned} \sum [a_i f_i(x, y) + a_j f_j(x, y)] &= \sum a_i f_i(x, y) + \sum a_j f_j(x, y) = \\ &= a_i \sum f_i(x, y) + a_j \sum f_j(x, y) = a_i g_i(x, y) + a_j g_j(x, y), \end{aligned}$$

где первый шаг преобразования вытекает из дистрибутивности сложения. Таким образом, левая часть (2.6-2) равна правой, и можно заключить, что оператор суммы линейный.

Здесь всюду используется поэлементное суммирование, а не сумма всех пикселей изображения. Применение оператора суммы к одному изображению дает само это изображение.

Напротив, рассмотрим оператор \max , функция которого состоит в нахождении максимального значения пикселя во входном изображении. Для наших целей простейший способ доказать, что этот оператор нелинейный, — найти контрпример, для которого нарушается условие (2.6-2). Рассмотрим следующие два изображения:

$$f_1 = \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \text{ и } f_2 = \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix},$$

и предположим, что константы $a_1 = 1$ и $a_2 = -1$. Проверку линейности опять начнем с левой части (2.6-2):

$$\max \left\{ (1) \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1) \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix} \right\} = \max \left\{ \begin{bmatrix} -6 & -3 \\ -2 & -4 \end{bmatrix} \right\} = -2.$$

Действуя теперь с правой частью, получим

$$(1) \max \left\{ \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \right\} + (-1) \max \left\{ \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix} \right\} = 3 + (-1)7 = -4.$$

В данном случае левая и правая части (2.6-2) не равны друг другу, и тем самым доказано, что в общем случае оператор \max является нелинейным.

Как мы увидим в трех следующих главах, особенно в главах 4 и 5, линейные операторы исключительно важны для обработки изображений, поскольку они опираются на значительную совокупность хорошо изученных теоретических и практических результатов. Нелинейные операторы исследованы значительно хуже, поэтому область их применения более ограничена. Однако мы познакомимся в последующих главах с несколькими нелинейными операциями обработки изображений, результаты которых значительно превосходят те, которые достигаются с помощью линейных аналогов.

2.6.3. Арифметические операции

Арифметические операции над изображениями являются поэлементными операциями, т. е., как уже говорилось в разделе 2.6.1, они применяются к паре соответствующих пикселей двух изображений. Эти четыре арифметические операции обозначаются следующим образом:

$$\begin{aligned} s(x, y) &= f(x, y) + g(x, y), \\ d(x, y) &= f(x, y) - g(x, y), \\ p(x, y) &= f(x, y) \times g(x, y), \\ v(x, y) &= f(x, y) \div g(x, y). \end{aligned} \tag{2.6-3}$$

Понятно, что эти операции применяются к соответственным парам элементов изображений f и g для $x = 0, 1, 2, \dots, M - 1$ и $y = 0, 1, 2, \dots, N - 1$, где, как обычно, M и N — число строк и столбцов изображений соответственно. Ясно, что s, d, p и v тоже являются изображениями с размерами $M \times N$. Заметим, что в так опреде-

ленной арифметике участвуют изображения одинаковых размеров. Следующие примеры показывают важную роль, которую играют арифметические операции в цифровой обработке изображений.

Пример 2.5. Сложение (усреднение) зашумленных изображений для уменьшения шума.

■ Рассмотрим зашумленное изображение $g(x, y)$, формируемое прибавлением шума $\eta(x, y)$ к исходному изображению $f(x, y)$, то есть

$$g(x, y) = f(x, y) + \eta(x, y), \quad (2.6-4)$$

где предполагается, что значения шума в каждой точке (x, y) являются некоррелированными⁵ и имеют нулевое среднее значение. Целью нижеследующей процедуры является уменьшение уровня шума путем суммирования серии зашумленных изображений $\{g_i(x, y)\}$. Этот метод часто применяется для улучшения изображений.

Если шум удовлетворяет только что сформулированным условиям, то можно показать следующее (задача 2.20). Пусть изображение $\bar{g}(x, y)$ получено усреднением K изображений $g_i(x, y)$, отличающихся лишь шумом $\eta_i(x, y)$,

$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^K g_i(x, y), \quad (2.6-5)$$

откуда следует, что

$$E\{\bar{g}(x, y)\} = f(x, y) \quad (2.6-6)$$

и

$$\sigma_{\bar{g}(x, y)}^2 = \frac{1}{K} \sigma_{\eta(x, y)}^2, \quad (2.6-7)$$

где $E\{\bar{g}(x, y)\}$ есть математическое ожидание \bar{g} , а $\sigma_{\bar{g}(x, y)}^2$ и $\sigma_{\eta(x, y)}^2$ — дисперсии \bar{g} и η , все в точке (x, y) . Стандартное отклонение в каждой точке усредненного изображения будет

$$\sigma_{\bar{g}(x, y)} = \frac{1}{\sqrt{K}} \sigma_{\eta(x, y)}. \quad (2.6-8)$$

Как следует из уравнений (2.6-7) и (2.6-8), при увеличении K величина отклонения (уровень шума) значения элемента в каждой точке (x, y) уменьшается. Поскольку $E\{\bar{g}(x, y)\} = f(x, y)$, это означает, что $\bar{g}(x, y)$ приближается к $f(x, y)$ с увеличением числа суммируемых зашумленных изображений. На практике изображения $g_i(x, y)$ должны являться *совмещенными*, чтобы уменьшить влияние расфокусировки или других искажений на выходном изображении.

Усреднение изображений находит важное применение в астрономии, где обычным является получение изображений при очень малом уровне освещенности и, как следствие, высоком уровне шума датчика, что зачастую делает отдельные снимки почти бесполезными для анализа. На рис. 2.26(a) представле-

⁵ Вспомним, что дисперсия случайной переменной z со средним m определяется как $E\{(z - m)^2\}$, где $E\{\cdot\}$ есть математическое ожидание аргумента. Ковариация двух случайных переменных z_i и z_j определяется как $E\{(z_i - m_i)(z_j - m_j)\}$. Если переменные *некоррелированы*, их ковариация равна 0.