

УДК 519.7
ББК 32.81
Ш23

Серия основана в 2006 г.

Шапиро Л.

Ш23 Компьютерное зрение / Л. Шапиро, Дж. Стокман ; пер. с англ. — 5-е изд., электрон. — М. : Лаборатория знаний, 2024. — 763 с. — (Лучший зарубежный учебник). — Систем. требования: Adobe Reader XI ; экран 10". — Загл. с титул. экрана. — Текст : электронный.

ISBN 978-5-93208-725-1

В данной книге теоретические аспекты обработки зрительных данных рассматриваются с привлечением большого количества примеров из практических задач. Наряду с классическими темами, в книге рассматриваются базы данных изображений и системы виртуальной и дополненной реальности. Приведены примеры приложений в промышленности, медицине, землепользовании, мультимедиа и компьютерной графике.

Студентам старших курсов и аспирантам, интересующимся современным состоянием дел в машинном зрении.

**УДК 519.7
ББК 32.81**

Деривативное издание на основе печатного аналога: Компьютерное зрение / Л. Шапиро, Дж. Стокман ; пер. с англ. — М. : БИНОМ. Лаборатория знаний, 2006. — 752 с. : ил., [8] с. цв. вкл. — (Лучший зарубежный учебник). — ISBN 5-94774-384-1.

В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации

ISBN 978-5-93208-725-1

Authorized Translation from the English language edition, entitled COMPUTER VISION; by LINDA SHAPIRO; and by GEORGE STOCKMAN; published by Pearson Education, Inc, publishing as Prentice Hall. Copyright © 2001 by Prentice Hall, Inc. All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc. Electronic RUSSIAN language edition published by BKL PUBLISHERS. Copyright © 2013.

Авторизованный перевод издания на английском языке, озаглавленного COMPUTER VISION, авторы LINDA SHAPIRO и GEORGE STOCKMAN, опубликованного Pearson Education, Inc, осуществляющим издательскую деятельность под торговой маркой Prentice Hall © 2001 by Prentice Hall, Inc. Все права защищены. Воспроизведение или распространение какой-либо части/частей данной книги в какой-либо форме, какими-либо способами, электронными или механическими, включая фотокопирование, запись и любые поисковые системы хранения информации, без разрешения Pearson Education, Inc запрещены. Электронная русскоязычная версия издана BKL Publishers, Copyright © 2013.

© Перевод на русский язык, Лаборатория знаний, 2015

Оглавление

Предисловие редактора перевода	5
Предисловие к русскому изданию	7
Предисловие	8
Замечания относительно языка программирования	10
Варианты работы с книгой	10
Глава 1. Введение	13
1.1. Могут ли машины видеть?	14
1.2. Прикладные задачи	15
1.2.1. Предварительное рассмотрение структуры цифрового изображения	15
1.2.2. Поиск в базе данных изображений	16
1.2.3. Контроль отверстий в поперечных балках	17
1.2.4. Контроль медицинских изображений головы человека	20
1.2.5. Обработка сканированных страниц текста	21
1.2.6. Оценка снежного покрова по спутниковым изображениям	22
1.2.7. Анализ сцен, содержащих детали промышленного производства	23
1.3. Операции обработки изображений	24
1.3.1. Модификация пикселей в малых окрестностях	25
1.3.2. Глобальное улучшение качества изображения	26
1.3.3. Комбинация нескольких изображений	26
1.3.4. Вычисление характерных признаков изображения	27
1.3.5. Формирование неграфических описаний	28
1.4. Успехи, сложности и неудачи компьютерного зрения	29
1.5. Применение компьютеров и программного обеспечения	30
1.6. Смежные области	31
1.7. Структура оставшейся части книги	32
1.8. Литература	32
1.9. Дополнительные упражнения	34
Глава 2. Формирование и представление изображений	37
2.1. Восприятие света	38
2.2. Устройства для формирования изображений	39
2.2.1. Камеры на основе ПЗС	39
2.2.2. Формирование изображения	41

2.2.3. Видеокамеры	43
2.2.4. Человеческий глаз	44
2.3. Проблемы формирования цифровых изображений	45
2.3.1. Геометрические искажения	45
2.3.2. Дисперсия	46
2.3.3. Блюминг (избыточная яркость)	46
2.3.4. Неоднородности ПЗС-матрицы	46
2.3.5. Отсечение и циклический возврат	47
2.3.6. Хроматическая дисторсия	47
2.3.7. Эффекты дискретизации	47
2.4. Функции интенсивности и цифровые изображения	47
2.4.1. Типы изображений	47
2.4.2. Дискретизация изображений и пространственные измерения	50
2.5. Форматы цифровых изображений*	55
2.5.1. Заголовок файла изображения	56
2.5.2. Данные изображения	56
2.5.3. Сжатие данных	57
2.5.4. Часто используемые форматы	57
2.5.5. Групповое кодирование бинарных изображений	58
2.5.6. Формат PGM: Portable Gray Map	58
2.5.7. Формат файлов изображений GIF	60
2.5.8. Формат файлов изображений TIFF	60
2.5.9. Формат JPEG для хранения фотографий	60
2.5.10. Язык PostScript	62
2.5.11. Файловый формат MPEG для хранения видеоданных	62
2.5.12. Сравнение форматов файлов изображений	63
2.6. Особенности и проблемы получения изображений естественных сцен	64
2.7. Оценка пространственных свойств объектов по двумерным изображениям	65
2.8. Пять систем координат	67
2.8.1. Пиксельная система координат I	68
2.8.2. Система координат объекта O	68
2.8.3. Система координат камеры C	68
2.8.4. Действительная система координат изображения F	69
2.8.5. Мировая система координат W	69
2.9. Другие типы датчиков*	69
2.9.1. Микроденситометр*	70
2.9.2. Цветные и мультиспектральные изображения*	70
2.9.3. Рентгеновские изображения*	72
2.9.4. Получение изображений методом ядерного магнитного резонанса (ЯМР)*	73
2.9.5. Дальномеры и дальнометрические изображения*	74
2.10. Литература	76
Глава 3. Анализ бинарных изображений	78
3.1. Пикселы и окрестности пикселов	79
3.2. Применение масок к изображениям	80
3.3. Подсчет объектов на изображении	82

3.4. Маркировка связных компонент	84
3.5. Морфология бинарных изображений	93
3.5.1. Структурирующие элементы	93
3.5.2. Основные операции	95
3.5.3. Некоторые приложения бинарной морфологии	98
3.5.4. Условное наращивание.	103
3.6. Свойства областей.	103
3.7. Графы смежности областей бинарного изображения.	114
3.8. Пороговая бинаризация полутоновых изображений.	116
3.8.1. Выбор порога бинаризации по гистограмме	117
3.8.2. Автоматический выбор порога бинаризации: метод Оцу*	119
3.9. Литература	124
Глава 4. Основные понятия распознавания образов.	126
4.1. Задачи распознавания образов	126
4.2. Общая модель классификации	128
4.2.1. Классы.	128
4.2.2. Датчик/преобразователь	128
4.2.3. Экстрактор характерных признаков	129
4.2.4. Классификатор.	130
4.2.5. Построение системы классификации	130
4.2.6. Оценка ошибок системы	130
4.2.7. Ложные срабатывания и ложные пропуски	132
4.3. Взаимосвязь точности системы и полноты выборки.	133
4.4. Признаки, используемые для описания объектов	133
4.5. Представление объектов в виде векторов признаков	136
4.6. Реализация классификатора	136
4.6.1. Классификация по ближайшему среднему значению	137
4.6.2. Классификация по расстоянию до ближайших соседей.	139
4.7. Структурные методы распознавания	140
4.8. Матрица неточностей	143
4.9. Деревья решений.	144
4.10. Байесовский подход к принятию решений	152
4.10.1. Параметрические модели распределений.	154
4.11. Принятие решений при обработке многомерных данных.	156
4.12. Самообучающиеся машины	158
4.13. Искусственные нейронные сети*	159
4.13.1. Персептрон	160
4.13.2. Многослойная сеть прямого распространения	165
4.14. Литература	167
Глава 5. Фильтрация и улучшение изображений	169
5.1. Что необходимо сделать с изображением?	170
5.1.1. Повышение качества изображений.	170
5.1.2. Обнаружение низкоуровневых признаков	170
5.2. Изменение тонового распределения	172
5.2.1. Выравнивание гистограммы	173

5.3. Удаление малых областей изображения.....	176
5.3.1. Удаление шума «соль и перец».....	177
5.3.2. Удаление малых компонент.....	178
5.4. Сглаживание изображения.....	179
5.5. Медианная фильтрация.....	181
5.5.1. Вычисление выходного изображения по данным входного изображения.....	184
5.6. Обнаружение краев с помощью дифференциальных масок.....	186
5.6.1. Дифференцирование одномерных сигналов.....	186
5.6.2. Дифференциальные операторы для двумерных изображений.....	190
5.7. Гауссовская фильтрация и ЛОГ-фильтрация для обнаружения краев.....	196
5.7.1. Обнаружение краев с помощью ЛОГ-фильтра.....	200
5.7.2. Обнаружение краев в зрительной системе человека.....	201
5.7.3. Теория Марра—Хилдрет.....	204
5.8. Детектор краев Кэнни.....	207
5.9. Использование масок в качестве согласованных фильтров*.....	207
5.9.1. Векторное пространство всех сигналов, заданных в виде n дискретных значений.....	208
5.9.2. Использование ортогонального базиса.....	210
5.9.3. Неравенство Коши—Шварца.....	212
5.9.4. Векторное пространство изображений размерами $m \times n$	213
5.9.5. Базис Робертса для окрестностей 2×2	213
5.9.6. Базис Фрея—Чена для окрестностей 3×3	214
5.10. Свертка и кросс-корреляция*.....	219
5.10.1. Определение операций посредством масок.....	219
5.10.2. Операция свертки.....	221
5.10.3. Возможности параллельной реализации.....	225
5.11. Анализ пространственных частот с использованием гармонических функций*.....	225
5.11.1. Базис Фурье.....	227
5.11.2. Двумерные функции интенсивности.....	230
5.11.3. Дискретное преобразование Фурье.....	234
5.11.4. Полоснопропускающая фильтрация.....	237
5.11.5. Обсуждение преобразования Фурье.....	238
5.11.6. Теорема о свертке*.....	238
5.12. Итоги и обсуждение.....	240
5.13. Литература.....	241
Глава 6. Цвет и освещенность.....	244
6.1. Физические свойства цвета.....	245
6.1.1. Восприятие освещенных объектов.....	246
6.1.2. Дополнительные факторы.....	247
6.1.3. Чувствительность рецепторов.....	248
6.2. Цветовая система RGB.....	249
6.3. Другие цветовые системы.....	252
6.3.1. Субтрактивная цветовая система CMY.....	252
6.3.2. Цветовая система HSI: Тон-Насыщенность-Интенсивность (Hue-Saturation-Intensity).....	252

6.3.3. Телевизионные цветовые системы YIQ и YUV	256
6.3.4. Классификация с использованием цветовой информации.	257
6.4. Цветовые гистограммы.	259
6.5. Сегментация цветных изображений	261
6.6. Освещенность	264
6.6.1. Излучение одиночного источника света.	264
6.6.2. Диффузное отражение.	266
6.6.3. Зеркальное отражение	267
6.6.4. Потемнение при увеличении расстояния	269
6.6.5. Дополнительные факторы.	270
6.6.6. Модель освещения Фонга*.	271
6.6.7. Учет освещенности зрительной системой человека	271
6.7. Смежные темы*	272
6.7.1. Приложения	272
6.7.2. Человеческое цветовое восприятие.	272
6.7.3. Многоспектральные изображения	273
6.7.4. Классификационные изображения	273
6.8. Литература	274
 Глава 7. Текстура	276
7.1. Текстура, текстелы и текстурные статистики	277
7.2. Тексельное описание текстур.	278
7.3. Количественные характеристики текстур	280
7.3.1. Плотность и направление краев	280
7.3.2. Локальное двоичное разбиение.	282
7.3.3. Вычисление характерных признаков на основе матрицы вхождений	282
7.3.4. Энергетические текстурные характеристики Лавса.	285
7.3.5. Автокорреляция и спектр мощности	287
7.4. Текстуриная сегментация.	289
7.5. Литература	291
 Глава 8. Поиск изображений на основе содержания	293
8.1. Примеры баз данных изображений.	294
8.2. Запросы к базам данных изображений	295
8.3. Запросы по образцу	297
8.4. Меры расстояния между изображениями	298
8.4.1. Характеристики цветового сходства.	299
8.4.2. Характеристики текстурного сходства.	303
8.4.3. Характеристики сходства формы.	305
8.4.4. Характеристики сходства, учитывающие присутствующие на изображении объекты и их взаимное расположение	313
8.5. Организация базы данных.	318
8.5.1. Стандартные индексы	318
8.5.2. Пространственная индексация	322

8.5.3. Индексация для систем поиска изображений на основе содержания, в которых применяется несколько различных мер расстояния	322
8.6. Литература	324
Глава 9. Движение на двумерных изображениях	327
9.1. Явления и прикладные задачи, связанные с движением на изображениях	328
9.2. Вычитание изображений	329
9.3. Вычисление векторов перемещения	331
9.3.1. Игра Decathlete	333
9.3.2. Использование соответствующих точек	334
9.3.3. Алгоритм MPEG для сжатия потока видеоданных	339
9.3.4. Вычисление потока изображения*	341
9.3.5. Уравнение потока изображения*	342
9.3.6. Распространение ограничений при вычислении потока изображения*	344
9.4. Вычисление траекторий движущихся точек	345
9.4.1. Слежение с учетом априорных знаний из предметной области	351
9.5. Обнаружение значительных изменений условий видеосъемки	353
9.5.1. Сегментация видеопоследовательностей	354
9.5.2. Игнорирование некоторых эффектов камеры	357
9.5.3. Хранение фрагментов видеопоследовательностей	359
9.6. Литература	359
Глава 10. Сегментация изображений	362
10.1. Обнаружение областей	363
10.1.1. Методы кластеризации	364
10.1.2. Нарращивание областей	375
10.2. Способы представления областей	377
10.2.1. Оверлейные представления	377
10.2.2. Маркированные изображения	378
10.2.3. Кодирование границ	379
10.2.4. Квадрантные деревья	380
10.2.5. Таблицы свойств	381
10.3. Обнаружение контуров	382
10.3.1. Прослеживание границ существующих областей	382
10.3.2. Детектор и компоновщик краев Кэнни	386
10.3.3. Группировка согласующихся соседних контурных фрагментов в кривые	387
10.3.4. Преобразование Хафа для обнаружения прямых и дуг окружностей	393
10.4. Подбор моделей сегментов	403
10.5. Обнаружение высокоуровневых структур	410
10.5.1. Ленты	410
10.5.2. Обнаружение углов	413

10.6. Сегментация на основе согласованного движения	414
10.6.1. Границы движущихся объектов	414
10.6.2. Накопление траекторий движения	415
10.7. Литература	418
Глава 11. Сопоставление в двумерном пространстве	421
11.1. Совмещение двумерных данных	422
11.2. Представление точек	424
11.3. Аффинные геометрические преобразования	425
11.4. Наилучшее двумерное аффинное преобразование*	438
11.5. Распознавание двумерных объектов с использованием аффинных преобразований	439
11.6. Распознавание двумерных объектов с использованием реляционных моделей	451
11.7. Нелинейные методы деформации изображений	469
11.8. Итоги	473
11.9. Литература	474
Глава 12. Восприятие трехмерных сцен по двумерным изображениям	476
12.1. Внутренние изображения	477
12.2. Маркировка контурных изображений объектов с плоскими гранями	483
12.3. Трехмерные признаки на двумерных изображениях	491
12.4. Другие способы определения пространственных свойств объектов	497
12.4.1. Определение формы объектов по одному признаку	497
12.4.2. Точки схода	502
12.4.3. Определение глубины с помощью фокусировки	503
12.4.4. Признаки, связанные с движением	504
12.4.5. Контурные и виртуальные прямые	504
12.4.6. Совмещенность	505
12.5. Модель формирования изображений на основе перспективной проекции	506
12.6. Определение глубины с помощью стереоскопической системы	509
12.6.1. Обнаружение соответствующих признаков	513
12.7. Формула тонкой линзы*	518
12.8. Итоги	522
12.9. Литература	522
Глава 13. Восприятие трехмерных сцен. Оценка пространственного положения и ориентации объектов	526
13.1. Устройство стереоскопической системы компьютерного зрения общего назначения	527
13.2. Аффинные преобразования в трехмерном пространстве	530
13.2.1. Системы координат	530
13.2.2. Перенос	531
13.2.3. Масштабирование	532
13.2.4. Поворот	532
13.2.5. Поворот вокруг произвольной оси	535
13.2.6. Совмещение фигур посредством аффинных преобразований	537

13.3. Модель камеры	539
13.3.1. Матрица перспективной проекции	540
13.3.2. Прямоугольная и слабая перспективная проекции.	544
13.3.3. Вычисление трехмерных координат с использованием нескольких камер	547
13.4. Наилучшая аффинная калибровочная матрица	549
13.4.1. Калибровочный стенд	550
13.4.2. Задача для решения методом наименьших квадратов.	550
13.4.3. Обсуждение аффинного метода	555
13.5. Использование структурной подсветки	557
13.6. Простая процедура оценки положения наблюдаемого объекта	560
13.7. Улучшенный метод калибровки камеры*.	567
13.7.1. Внутренние параметры камеры	567
13.7.2. Внешние параметры камеры	568
13.7.3. Пример калибровки	573
13.8. Оценка положения объектов *	578
13.8.1. Оценка положения на основе соответствия двумерных и трех- мерных координат точек	578
13.8.2. Линейная оптимизация при наличии ограничений.	580
13.8.3. Вычисление преобразования $Tr = \{\mathbf{R}, \mathbf{T}\}$	581
13.8.4. Верификация и оптимизация при вычислении положения объектов.	583
13.9. Реконструкция трехмерных объектов.	584
13.9.1. Сбор дальнометрических данных	585
13.9.2. Совмещение видов	588
13.9.3. Реконструкция поверхности	589
13.9.4. Алгоритм объемного удаления	590
13.10. Вычисление формы по данным освещенности.	594
13.10.1. Стереофотометрический метод	598
13.10.2. Интегральный учет пространственных ограничений	599
13.11. Геометрическая структура объектов по данным о движении	600
13.12. Литература	603

Глава 14. Трехмерные модели. Распознавание объектов на изображениях на основе моделей. 607

14.1. Обзор распространенных разновидностей моделей.	608
14.1.1. Трехмерные каркасные модели.	608
14.1.2. Модели типа «поверхность-ребро-вершина»	609
14.1.3. Модели на основе обобщенных цилиндров	612
14.1.4. Модели на основе октантных деревьев.	615
14.1.5. Модели на основе суперквадрик.	615
14.2. Модели на основе классов видимости как альтернатива истинным трехмерным моделям	618
14.3. Физические и деформационные модели	619
14.3.1. Модели на основе активных контуров (snakes-модели)	620
14.3.2. Трехмерные оболочечные модели.	624
14.3.3. Моделирование движения человеческого сердца	626

14.4. Основные методы распознавания трехмерных объектов	627
14.4.1. Распознавание по трехмерным моделям с использованием процедуры совмещения	629
14.4.2. Распознавание на основе сопоставления исходных данных с реляционными моделями	639
14.4.3. Распознавание на основе сопоставления с функциональными моделями	651
14.4.4. Распознавание по внешнему виду объектов	654
14.5. Литература	663
Глава 15. Системы виртуальной реальности	667
15.1. Основные признаки систем виртуальной реальности	668
15.2. Приложения систем виртуальной реальности	669
15.3. Системы дополненной реальности	671
15.4. Дистанционное управление	674
15.5. Устройства для систем виртуальной реальности	677
15.6. Краткий обзор датчиков для систем виртуальной реальности	682
15.7. Генерация простых трехмерных моделей	685
15.8. Сочетание реальных и искусственных изображений	687
15.9. Психофизиологические аспекты человеко-машинного интерфейса ..	691
15.10. Литература	692
Глава 16. Примеры прикладных задач	694
16.1. Veggie Vision: система для распознавания овощей и фруктов	694
16.1.1. Прикладная область и требования к системе	695
16.1.2. Устройство системы	696
16.1.3. Процедура идентификации	698
16.1.4. Более подробное описание процесса обработки	698
16.1.5. Производительность	703
16.2. Идентификация личности человека по радужной оболочке глаза ...	704
16.2.1. Требования к системам идентификации личности	705
16.2.2. Устройство системы	707
16.2.3. Производительность системы	712
16.3. Литература	713
Аннотированный список литературы, добавленный при переводе	714
1. Распознавание образов и компьютерное зрение	714
2. Представление и обработка изображений	714
3. Компьютерная графика и геометрическое моделирование	715
4. Смежные вопросы	716
Предметный указатель	717