

УДК 004.93Python

ББК 32.972.1

С60

Содем, Ян Эрик.

С60 Программирование компьютерного зрения на языке Python / Я. Э. Содем ; пер. с англ. А. А. Слинкина. — 2-е изд., эл. — 1 файл pdf : 313 с. — Москва : ДМК Пресс, 2023. — Систем. требования: Adobe Reader XI либо Adobe Digital Editions 4.5 ; экран 10". — Текст : электронный.

ISBN 978-5-89818-310-3

Если вы хотите разобраться в основах теории и алгоритмов компьютерного зрения, то эта книга — как раз то, что вам нужно. Вы узнаете о методах распознавания объектов, трехмерной реконструкции, обработке стереоизображений, дополненной реальности и других приложениях компьютерного зрения. Изложение сопровождается понятными примерами на языке Python. При этом объяснения даются в общих чертах, без погружения в сухую теорию.

Издание идеально подходит для студентов, исследователей и энтузиастов-любителей с базовыми знаниями математики и навыками программирования.

УДК 004.93Python

ББК 32.972.1

Электронное издание на основе печатного издания: Программирование компьютерного зрения на языке Python / Я. Э. Содем ; пер. с англ. А. А. Слинкина. — Москва : ДМК Пресс, 2016. — 312 с. — ISBN 978-5-97060-200-3. — Текст : непосредственный.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

Материал, изложенный в данной книге, многократно проверен. Но поскольку вероятность технических ошибок все равно существует, издательство не может гарантировать абсолютную точность и правильность приводимых сведений. В связи с этим издательство не несет ответственности за возможные ошибки, связанные с использованием книги.

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 978-5-89818-310-3

© Jan Erik Solem, 2012.

© Оформление, перевод на русский язык,
ДМК Пресс, 2016



ОГЛАВЛЕНИЕ

Предисловие	11
Требования к читателю и структура книги	12
Необходимые знания и навыки	12
Чему вы научитесь	12
Структура книги	13
Введение в компьютерное зрение	14
Python и NumPy	14
Обозначения и графические выделения	15
О примерах кода	16
Как с нами связаться	17
Благодарности	17
Об авторе	18
Глава 1. Основы обработки изображений	19
1.1. PIL – библиотека Python Imaging Library	19
Преобразование изображения в другой формат	20
Создание миниатюр	21
Копирование и вставка областей	21
Изменение размера и поворот	22
1.2. Библиотека Matplotlib	22
Рисование точек и прямых линий	22
Изолинии и гистограммы изображений	25
Интерактивное аннотирование	26
1.3. Пакет NumPy	27
Представление изображения в виде массива	27
Преобразование уровня яркости	29
Изменение размера изображения	31
Выравнивание гистограммы	31
Усреднение изображений	33
Метод главных компонент для изображений	34
Использование модуля pickle	37
1.4. Пакет SciPy	39
Размытие изображений	39
Производные изображения	40
Морфология – подсчет объектов	43
Полезные модули в пакете SciPy	46

1.5. Более сложный пример: очистка изображения от шумов	47
Упражнения.....	50
Соглашения в примерах кода	51

Глава 2. Локальные дескрипторы изображений 53

2.1. Детектор углов Харриса	53
Нахождение соответственных точек в изображениях	57
2.2. SIFT – масштабно-инвариантное преобразование признаков . 62	
Особые точки.....	62
Дескриптор.....	63
Обнаружение особых точек.....	63
Сопоставление дескрипторов.....	67
2.3. Сопоставление изображений с геометками	70
Загрузка изображений с геометками из Panoramio	71
Сопоставление с помощью локальных дескрипторов.....	74
Визуализация связанных изображений.....	76
Упражнения.....	78

Глава 3. Преобразования изображений 80

3.1. Гомографии	80
Алгоритм прямого линейного преобразования	82
Аффинные преобразования.....	84
3.2. Деформирование изображений.....	85
Изображение внутри изображения	86
Кусочно-аффинное деформирование	91
Регистрация изображений.....	95
3.3. Создание панорам.....	101
RANSAC	101
Устойчивое вычисление гомографии	102
Сшивка изображений.....	106
Упражнения.....	109

Глава 4. Модели камер и дополненная реальность... 110

4.1. Модель камеры с точечной диафрагмой	110
Матрица камеры	111
Проецирование точек трехмерного пространства.....	113
Вычисление центра камеры	116
4.2. Калибровка камеры	116
Простой метод калибровки	117
4.3. Оценивание положения по плоскостям и маркерам	119
4.4. Дополненная реальность	123
PyGame и PyOpenGL.....	124

От матрицы камеры к формату OpenGL.....	125
Помещение виртуальных на изображение	127
Собираем все вместе	129
Загрузка моделей	132
Упражнения.....	134
Глава 5. Многовидовая геометрия	135
5.1. Эпиполярная геометрия	135
Демонстрационный набор данных	138
Построение трехмерных графиков в Matplotlib.....	140
Вычисление F – восьмиточечный алгоритм	141
Эпиполус и эпиполярные прямые	142
5.2. Вычисления, относящиеся к камерам и трехмерной структуре	145
Триангуляция.....	145
Вычисление матрицы камеры по точкам в пространстве.....	148
Вычисление матрицы камеры по фундаментальной матрице.....	150
5.3. Многовидовая реконструкция.....	153
Устойчивое вычисление фундаментальной матрицы.....	154
Пример трехмерной реконструкции.....	156
Обобщения и случай более двух видов.....	159
5.4. Стереοизображения	161
Вычисление карт диспаратности.....	163
Упражнения.....	167
Глава 6. Кластеризация изображений	170
6.1. Кластеризация методом K средних.....	170
Пакет кластеризации в SciPy.....	171
Кластеризация изображений	172
Визуализация проекций изображений на главные компоненты.....	174
Кластеризация пикселей	175
6.2. Иерархическая кластеризация.....	178
Кластеризация изображений	182
6.3. Спектральная кластеризация.....	186
Упражнения.....	191
Глава 7. Поиск изображений	193
7.1. Поиск изображений по содержанию	193
Векторная модель – инструмент анализа текста	193
7.2. Визуальные слова.....	195
Создание словаря.....	195
7.3. Индексирование изображений	198
Подготовка базы данных.....	198

Добавление изображений.....	200
7.4. Поиск изображений в базе данных.....	202
Использование индекса для получения кандидатов	203
Запрос по изображению	205
Эталонное тестирование и построение графика	206
7.5. Ранжирование результатов с применением геометрических соображений.....	209
7.6. Создание демонстраций и веб-приложений	212
Создание веб-приложений с помощью CherryPy.....	212
Демонстрация поиска изображений	212
Упражнения.....	215
Глава 8. Классификация изображений по содержанию	217
8.1. Метод k ближайших соседей	217
Простой двумерный пример	218
Плотные SIFT-дескрипторы в качестве признаков изображения	222
Классификация изображений – распознавание жестов	223
8.2. Байесовский классификатор	227
Использование метода главных компонент для понижения размерности.....	231
8.3. Метод опорных векторов	232
Использование библиотеки LibSVM	233
И снова о распознавании жестов	235
8.4. Оптическое распознавание символов.....	237
Обучение классификатора	238
Отбор признаков.....	238
Выделение клеток и распознавание символов	240
Выпрямление изображений	243
Упражнения.....	245
Глава 9. Сегментация изображений	247
9.1. Разрезание графов.....	247
Графы изображений	249
Сегментация с привлечением пользователя	254
9.2. Сегментация с применением кластеризации.....	258
9.3. Вариационные методы	264
Упражнения.....	265
Глава 10. OpenCV.....	268
10.1. Интерфейс между OpenCV и Python	268
10.2. Основы OpenCV	269

Чтение и запись изображений	269
Цветовые пространства	270
Отображение изображений и результатов обработки	270
10.3. Обработка видео	273
Ввод видео	273
Чтение видео в массивы NumPy	275
10.4. Трассировка	276
Оптический поток	276
Алгоритм Лукаса-Канаде	279
Использование трассировщика	283
Применение генераторов	284
10.5. Другие примеры	285
Ретуширование	285
Сегментация по морфологическим водоразделам	286
Обнаружение фигур с помощью преобразования Хафа	288
Упражнения	288

Приложение А. Установка пакетов 291

A.1. NumPy и SciPy	291
Windows	291
Mac OS X	291
Linux	292
A.2. Matplotlib	292
A.3. PIL	292
A.4. LibSVM	293
A.5. OpenCV	293
Windows и Unix	293
Mac OS X	294
Linux	294
A.6. VLFeat	294
A.7. PyGame	295
A.8. PyOpenGL	295
A.9. Pydot	295
A.10. Python-graph	296
A.11. Simplejson	296
A.12. PySQLite	297
A.13. CherryPy	297

Приложение Б. Наборы изображений 298

Б.1. Flickr	298
Б.2. Panoramio	299
Б.3. Оксфордская группа Visual Geometry	300

Б.4. Эталонные изображения для распознавания Кентуккийского университета	301
Б.5. Другие наборы	301
Пражский генератор данных и эталонный набор для сегментации текстур	301
Набор данных Grab Cut научно-исследовательского центра Microsoft в Кембридже	301
Caltech 101.....	302
База данных статических положений руки.....	302
Наборы стереоизображений Мидлбери-колледжа.....	302
Приложение В. Благодарности авторам изображений	303
В.1. Изображения с сайта Flickr	303
В.2. Прочие изображения	304
В.3. Иллюстрации	304
Литература.....	305
Предметный указатель	308