

Python计算机视觉编程 (图灵程序设计丛书)

作者: [美]Jan Erik Solem

版权信息

书名: Python计算机视觉编程

作者: (美)索利姆 (Solem J. E.)

译者: 朱文涛, 袁勇

ISBN: 978-7-115-35232-3

本书由北京图灵文化发展有限公司发行数字版。版权所有, 侵权必究。

您购买的图灵电子书仅供您个人使用, 未经授权, 不得以任何方式复制和传播本书内容。

我们愿意相信读者具有这样的良知和觉悟, 与我们共同保护知识产权。

如果购买者有侵权行为, 我们可能对该用户实施包括但不限于关闭该帐号等维权措施, 并可能追究法律责任。

图灵社区会员 ptpress (libowen@ptpress.com.cn) 专享 尊重版权

[版权声明](#)

[O'Reilly Media, Inc. 介绍](#)

[业界评论](#)

[推荐序](#)

[前言](#)

[先决条件和概述](#)

[读者须知](#)

[本书内容](#)

[各章概览](#)

[计算机视觉简介](#)

[Python和NumPy](#)

[排版约定](#)

[使用代码示例](#)

[Safari® Books Online](#)

[联系我们](#)

[致谢](#)

[第 1 章 基本的图像操作和处理](#)

[1.1 PIL: Python图像处理类库](#)

[1.1.1 转换图像格式](#)

[1.1.2 创建缩略图](#)

[1.1.3 复制和粘贴图像区域](#)

[1.1.4 调整尺寸和旋转](#)

[1.2 Matplotlib](#)

[1.2.1 绘制图像、点和线](#)

[1.2.2 图像轮廓和直方图](#)

[1.2.3 交互式标注](#)

[1.3 NumPy](#)

[1.3.1 图像数组表示](#)

[1.3.2 灰度变换](#)

[1.3.3 图像缩放](#)

[1.3.4 直方图均衡化](#)

[1.3.5 图像平均](#)

[1.3.6 图像的主成分分析 \(PCA\)](#)

[1.3.7 使用pickle模块](#)

[1.4 SciPy](#)

[1.4.1 图像模糊](#)

[1.4.2 图像导数](#)

[1.4.3 形态学: 对象计数](#)

[1.4.4 一些有用的SciPy模块](#)

[1.5 高级示例: 图像去噪](#)

[练习](#)

[代码示例约定](#)

[第 2 章 局部图像描述子](#)

[2.1 Harris角点检测器](#)

[在图像间寻找对应点](#)

[2.2 SIFT \(尺度不变特征变换\)](#)

[2.2.1 兴趣点](#)

[2.2.2 描述子](#)

[2.2.3 检测兴趣点](#)

[2.2.4 匹配描述子](#)

[2.3 匹配地理标记图像](#)

[2.3.1 从Panoramio下载地理标记图像](#)

[2.3.2 使用局部描述子匹配](#)

[2.3.3 可视化连接的图像](#)

[练习](#)

[第3章 图像到图像的映射](#)

[3.1 单应性变换](#)

[3.1.1 直接线性变换算法](#)

[3.1.2 仿射变换](#)

[3.2 图像扭曲](#)

[3.2.1 图像中的图像](#)

[3.2.2 分段仿射扭曲](#)

[3.2.3 图像配准](#)

[3.3 创建全景图](#)

[3.3.1 RANSAC](#)

[3.3.2 稳健的单应性矩阵估计](#)

[3.3.3 拼接图像](#)

[练习](#)

[第4章 照相机模型与增强现实](#)

[4.1 针孔照相机模型](#)

[4.1.1 照相机矩阵](#)

[4.1.2 三维点的投影](#)

[4.1.3 照相机矩阵的分解](#)

[4.1.4 计算照相机中心](#)

[4.2 照相机标定](#)

[4.2.1 一个简单的标定方法](#)

[4.3 以平面和标记物进行姿态估计](#)

[4.4 增强现实](#)

[4.4.1 PyGame和PyOpenGL](#)

[4.4.2 从照相机矩阵到OpenGL格式](#)

[4.4.3 在图像中放置虚拟物体](#)

[4.4.4 综合集成](#)

[4.4.5 载入模型](#)

[练习](#)

[第5章 多视图几何](#)

[5.1 外极几何](#)

[5.1.1 一个简单的数据集](#)

[5.1.2 用Matplotlib绘制三维数据](#)

[5.1.3 计算F：八点法](#)

[5.1.4 外极点和外极线](#)

[5.2 照相机和三维结构的计算](#)

[5.2.1 三角剖分](#)

[5.2.2 由三维点计算照相机矩阵](#)

[5.2.3 由基础矩阵计算照相机矩阵](#)

[5.3 多视图重建](#)

[5.3.1 稳健估计基础矩阵](#)

[5.3.2 三维重建示例](#)

[5.3.3 多视图的扩展示例](#)

[5.4 立体图像](#)

[计算视差图](#)

[练习](#)

[第 6 章 图像聚类](#)

[6.1 K-means聚类](#)

[6.1.1 SciPy聚类包](#)

[6.1.2 图像聚类](#)

[6.1.3 在主成分上可视化图像](#)

[6.1.4 像素聚类](#)

[6.2 层次聚类](#)

[图像聚类](#)

[6.3 谱聚类](#)

[练习](#)

[第 7 章 图像搜索](#)

[7.1 基于内容的图像检索](#)

[从文本挖掘中获取灵感——矢量空间模型](#)

[7.2 视觉单词](#)

[创建词汇](#)

[7.3 图像索引](#)

[7.3.1 建立数据库](#)

[7.3.2 添加图像](#)

[7.4 在数据库中搜索图像](#)

[7.4.1 利用索引获取候选图像](#)

[7.4.2 用一幅图像进行查询](#)

[7.4.3 确定对比基准并绘制结果](#)

[7.5 使用几何特性对结果排序](#)

[7.6 建立演示程序及Web应用](#)

[7.6.1 用CherryPy创建Web应用](#)

[7.6.2 图像搜索演示程序](#)

[练习](#)

[第 8 章 图像内容分类](#)

[8.1 K邻近分类法 \(KNN\)](#)

[8.1.1 一个简单的二维示例](#)

[8.1.2 用稠密SIFT作为图像特征](#)

[8.1.3 图像分类：手势识别](#)

[8.2 贝叶斯分类器](#)

[用PCA降维](#)

[8.3 支持向量机](#)

[8.3.1 使用LibSVM](#)

[8.3.2 再论手势识别](#)

[8.4 光学字符识别](#)

[8.4.1 训练分类器](#)

[8.4.2 选取特征](#)

[8.4.3 多类支持向量机](#)

[8.4.4 提取单元格并识别字符](#)

[8.4.5 图像校正](#)

[练习](#)

[第 9 章 图像分割](#)

[9.1 图割 \(Graph Cut\)](#)

[9.1.1 从图像创建图](#)

[9.1.2 用户交互式分割](#)

[9.2 利用聚类进行分割](#)

[9.3 变分法](#)

[练习](#)

[第 10 章 OpenCV](#)

[10.1 OpenCV的Python接口](#)

[10.2 OpenCV基础知识](#)

[10.2.1 读取和写入图像](#)

[10.2.2 颜色空间](#)

[10.2.3 显示图像及结果](#)

[10.3 处理视频](#)

[10.3.1 视频输入](#)

[10.3.2 将视频读取到NumPy数组中](#)

[10.4 跟踪](#)

[10.4.1 光流](#)

[10.4.2 Lucas-Kanade算法](#)

[10.5 更多示例](#)

[10.5.1 图像修复](#)

[10.5.2 利用分水岭变换进行分割](#)

[10.5.3 利用霍夫变换检测直线](#)

[练习](#)

[附录 A 安装软件包](#)

[A.1 NumPy和SciPy](#)

[A.1.1 Windows](#)

[A.1.2 Mac OS X](#)

[A.1.3 Linux](#)

[A.2 Matplotlib](#)

[A.3 PIL](#)

[A.4 LibSVM](#)

[A.5 OpenCV](#)

[A.5.1 Windows 和 Unix](#)

[A.5.2 Mac OS X](#)

[A.5.3 Linux](#)

[A.6 VLEat](#)

[A.7 PyGame](#)

[A.8 PyOpenGL](#)

[A.9 Pydot](#)

[A.10 Python-graph](#)

[A.11 Simplejson](#)

[A.12 PySQLite](#)

[A.13 CherryPy](#)

[附录 B 图像集](#)

[B.1 Flickr](#)

[B.2 Panoramio](#)

[B.3 牛津大学视觉几何组](#)

[B.4 肯塔基大学识别基准图像](#)

[B.5 其他](#)

[B.5.1 Prague Texture Segmentation Datagenerator与基准](#)

[B.5.2 微软剑桥研究院Grab Cut数据集](#)

[B.5.3 Caltech 101](#)

[B.5.4 静态手势数据库](#)

[B.5.5 Middlebury Stereo数据集](#)

[附录 C 图片来源](#)

[C.1 来自Flickr的图像](#)

[C.2 其他图像](#)

[C.3 插图](#)

[参考文献](#)

版权声明

© 2012 by O'Reilly Media, Inc.

Simplified Chinese Edition, jointly published by O'Reilly Media, Inc. and Posts & Telecom Press, 2014. Authorized translation of the English edition, 2014 O'Reilly Media, Inc., the owner of all rights to publish and sell the same.

All rights reserved including the rights of reproduction in whole or in part in any form.

英文原版由 O'Reilly Media, Inc. 出版 2012。

简体中文版由人民邮电出版社出版，2014。英文原版的翻译得到 O'Reilly Media, Inc. 的授权。此简体中文版的出版和销售得到出版权和销售权的所有者——O'Reilly Media, Inc. 的许可。

版权所有，未得书面许可，本书的任何部分和全部不得以任何形式复制。

O'Reilly Media, Inc. 介绍

O'Reilly Media 通过图书、杂志、在线服务、调查研究和会议等方式传播创新知识。自 1978 年开始，O'Reilly 一直都是前沿发展的见证者和推动者。超级极客们正在开创着未来，而我们关注真正重要的技术趋势——通过放大那些“细微的信号”来刺激社会对新科技的应用。作为技术社区中活跃的参与者，O'Reilly 的发展充满了对创新的倡导、创造和发扬光大。

O'Reilly 为软件开发人员带来革命性的“动物书”；创建第一个商业网站（GNN）；组织了影响深远的开放源代码峰会，以至于开源软件运动以此命名；创立了 Make 杂志，从而成为 DIY 革命的主要先锋；公司一如既往地通过多种形式缔结信息与人的纽带。O'Reilly 的会议和峰会集聚了众多超级极客和高瞻远瞩的商业领袖，共同描绘出开创新产业的革命性思想。作为技术人士获取信息的选择，O'Reilly 现在还将先锋专家的知识传递给普通的计算机用户。无论是通过书籍出版、在线服务还是面授课程，每一项 O'Reilly 的产品都反映了公司不可动摇的理念——信息是激发创新的力量。

业界评论

“O'Reilly Radar 博客有口皆碑。”

——Wired

“O'Reilly 凭借一系列（真希望当初我也想到了）非凡想法建立了数百万美元的业务。”

——Business 2.0

“O'Reilly Conference 是聚集关键思想领袖的绝对典范。”

——CRN

“一本 O'Reilly 的书就代表一个有用、有前途、需要学习的主题。”

——Irish Times

“Tim 是位特立独行的商人，他不光放眼于最长远、最广阔的视野，并且切实地按照 Yogi Berra 的建议去做了：‘如果你在路上遇到岔路口，走小路（岔路）。’回顾过去 Tim 似乎每一次都选择了小路，而且有几次都是一闪即逝的机会，尽管大路也不错。”

——Linux Journal

推荐序

计算机视觉是一门实践性很强的课程，虽然已经有了不少教科书，但大多数都只重视其中的理论和算法，很少有实践指导书。因而对于学习者而言，如果希望在实践中学习，往往需要编写大量的程序。在这方面，本书的出版可以算是一个有益的补充，相信本书将成为计算机视觉学习者的一个重要参考。

作为一本面向计算机视觉编程的书，本书涉及了这一学科中相对成熟并且被以往实践验证有效的部分典型算法，因而具有很好的实用性。例如第 2 章描述子部分选择了 Harris 角点检测器和 SIFT 描述子及其实现加以介绍；第 3 章则以全景图的创建为例，给出了 RANSAC 的实现；第 9 章图像分割中讨论了 Graph Cut 的实现等。这些方法大多数具有很好的通用性，因而为读者提供了一种实现范例。

本书的另一个特点是对介绍的单一方法，通过综合运用提升学习者灵活应用这些方法的能力。例如第 4 章给出的增强现实的例子，以及第 8 章给出的图像校正的例子。这些例子能够帮助进一步提升学习者对前述方法的感性认识。

与早期计算机视觉领域多数程序都是由 C/C++ 写就的情形不同。随着计算机硬件速度越来越快，研究者在考虑选择实现算法语言的时候会更多地考虑编写代码的效率和易用性，而不是像早年那样把算法的执行效率放在首位。这直接导致近年来越来越多的研究者选择 Python 来实现算法。与此同时，Python 的开放性使不同领域的研究者能够有机会在 Python 中加入他们需要的特性，甚至可以纳入 Python 的标准库，这也大大吸引了众多研究者对 Python 的参与。

本书的第三个特点是提供了与 OpenCV 接口的介绍。这为利用 OpenCV 中的资源提供了方便的途径。

今天在计算机视觉领域，越来越多的研究者使用 Python 开展研究。本书中文版的出版一方面能够鼓励更多的研究者采用这一语言，另一方面则为 Python 的学习者提供了一种尝试不同领域算法的机会。

2014 年 4 月 21 日

欢迎访问：电子书学习和下载网站 (<https://www.shgis.cn>)

文档名称：《Python计算机视觉编程》[美]Jan Erik Solem.epub

请登录 <https://shgis.cn/post/1848.html> 下载完整文档。

手机端请扫码查看：

