

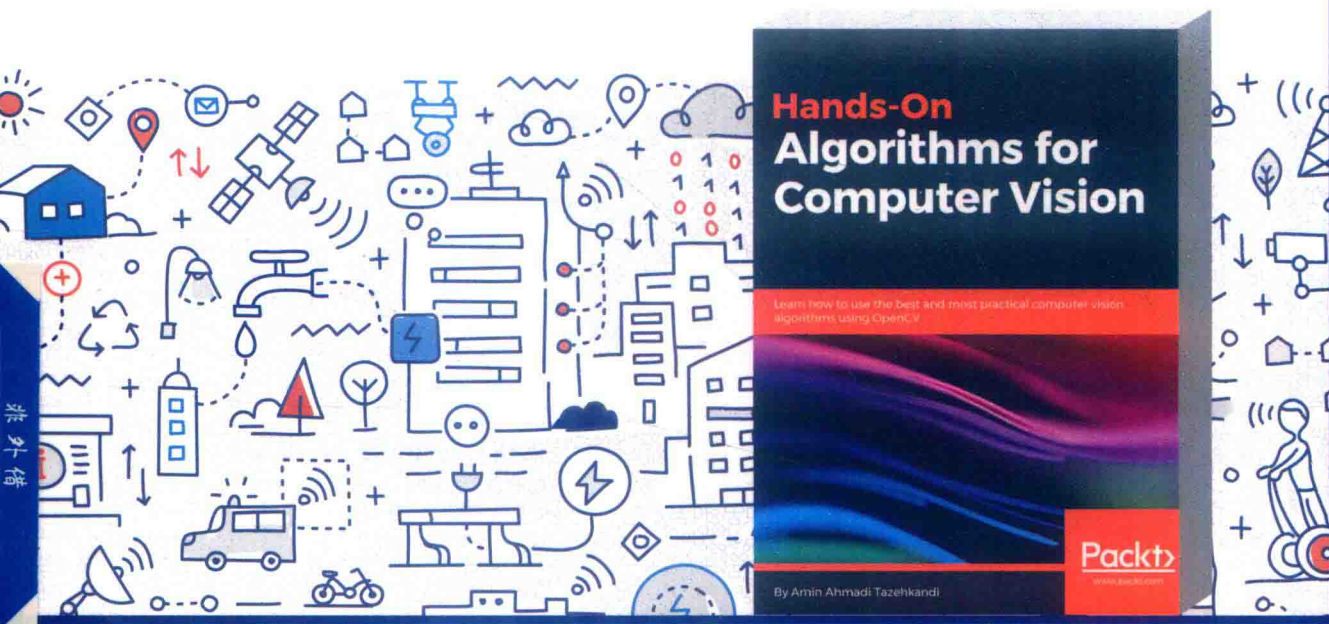
Hands-On Algorithms for Computer Vision

计算机视觉算法

基于 OpenCV 的计算机应用开发

[伊朗] 阿明·艾哈迈迪·塔兹赫孔迪
(Amin Ahmadi Tazehkandi) 著

陈靖 杨欣 译



Hands-On Algorithms for Computer Vision

计算机视觉算法

基于 OpenCV 的计算机应用开发

[伊朗] 阿明·艾哈迈迪·塔兹赫孔迪 著
(Amin Ahmadi Tazehkandi)

陈靖 杨欣 译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

计算机视觉算法：基于 OpenCV 的计算机应用开发 / (伊朗) 阿明·艾哈迈迪·塔兹赫孔迪著；陈靖，杨欣译．—北京：机械工业出版社，2019.3

(智能科学与技术丛书)

书名原文：Hands-On Algorithms for Computer Vision

ISBN 978-7-111-62315-1

I. 计… II. ①阿… ②陈… ③杨… III. 计算机视觉—算法 IV. TP302.7

中国版本图书馆 CIP 数据核字 (2019) 第 053575 号

本书版权登记号：图字 01-2018-6841

Amin Ahmadi Tazehkandi : *Hands-On Algorithms for Computer Vision* (ISBN: 978-1-78913-094-2) .

Copyright © 2018 Packt Publishing. First published in the English language under the title “Hands-On Algorithms for Computer Vision” .

All rights reserved.

Chinese simplified language edition published by China Machine Press.

Copyright © 2019 by China Machine Press.

本书中文简体字版由 Packt Publishing 授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

本书首先介绍计算机视觉开发所需的工具及其安装和配置，接着探索 OpenCV 框架及其强大的库和函数集。从最简单的图像修改、滤镜和变换开始，读者将逐步掌握各种算法知识，直到能执行更复杂的任务，例如使用深度学习算法进行实时对象检测等。本书适合所有对计算机视觉感兴趣并希望在实际工作中使用计算机视觉算法的读者。

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：冯秀泳

责任校对：殷虹

印刷：北京诚信伟业印刷有限公司

版次：2019 年 4 月第 1 版第 1 次印刷

开本：185mm×260mm 1/16

印张：13.5

书号：ISBN 978-7-111-62315-1

定价：69.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991 88379833

投稿热线：(010) 88379604

购书热线：(010) 68326294

读者信箱：hzjsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

美国东部时间 2012 年 6 月 25 日，当《纽约时报》作家 Jim Wilson 写下《How Many Computers to Identify a Cat? 16,000》这篇报道时，应该只有很少人能预料到：这个现象及其背后的技术，将如此迅速而深刻地为—个时代的变革拉开序幕。当时的我，在准备赴美读博时也并未把机器学习作为可能的研究方向。两年后，博士肄业加入谷歌的我发现，无论愿意与否，计算机专业的热点已转到了人工智能和机器学习（尤其是深度学习），想要将自己置身于这浪潮之外会越来越难：不管是什么方向的技术峰会、专业论坛，人工智能和机器学习都是人们关心的方向，也是嘉宾和讲师必然会涉及的话题。

在机器学习的诸多领域中，计算机视觉是目前发展最成熟的方向之一，也是许多人接触机器学习算法的起点。深度学习之所以火起来，不正是始于吴恩达博士在谷歌大脑用深度学习算法让机器学会识别猫吗（在薛定谔的猫之后，猫又一次为科学发展做出了贡献）？同样，让机器更聪明地看与听，仍将是未来几年人工智能的主要课题之一。而随着国家 AI+ 战略的推进，未来不只是计算机软硬件以及互联网行业的从业者会站在这波浪潮之巅，过去相对传统的行业也将被卷入其中。以在线教育为例：作业帮、一起作业网等能够拍照搜题；叮咚课堂、aiKID 和葡萄英语能够进行简单的语音识别，用 AI 老师的方式跟孩子互动；而在 51Talk，我们用人脸识别、表情识别来进行课堂质量监督并根据孩子的表现调整教学内容。在线教育只是冰山一角，在当下已经有成百上千的应用场景正在被 AI 深度影响甚至颠覆。

与其他计算机领域—样，理解计算机视觉算法需要坚实的数学基础。此外，读者最好还具备图像处理、信号与系统、计算机图形学等学科的背景及开发经验。开发者还要掌握各种基础算法和数据结构，以应对各种开发需求。当然，在移动设备成为主要个人计算平台的今天，针对不同硬件条件下图像和视频的采集、编解码、传输，以及不同处理器和片上系统（System on Chip, SoC）的优化与测试更是算法落地必不可少的技能。没有足够的理论知识和工程经验的初学者面对如此浩瀚繁多且异常艰深的内容将无从下手。

然而，本书和许多计算机视觉书籍的最大区别就在于隐藏了很多复杂的细节，绕开了相关算法的数学证明和实现，利用计算机视觉领域大名鼎鼎的 OpenCV（Open Source Computer Vision library），让读者自己动手，实现一个个具体且实用的例子。这样的好处显而易见，读者在完成这些实例后能熟悉计算机视觉算法开发的整体流程，逐步树立自信。相关算法的细节可以通过研究 OpenCV 源代码，结合阅读相关书籍、论文来逐步

补充学习。在译者看来，这是掌握一门技术的最佳路径。

本书特别适合计算机背景的学生和开发者学习阅读。全书在简单介绍了计算机视觉的基本概念和 OpenCV 的使用方法后，按照计算机视觉项目开发的难易程度，从基础的图像和视频的读写操作，到简单的图像矩阵运算，再到基础的图像处理、视频分析，最后到对象检测与机器学习分类算法，抽丝剥茧层层递进。本书还配备了完整的示例代码，有助于读者自主学习并验证各种算法的效果。同时，得益于 OpenCV 提供的工具、函数和类，这些示例代码都很轻量，非常易于理解。推荐读者在阅读本书时参考 OpenCV 的官方文档和其他图像处理、机器学习算法类的书籍，这样将会事半功倍。

译者在本书的翻译过程中，得到了北京大生在线科技有限公司（51Talk）的同事和华章公司编辑的大力协助，在此由衷感谢。

译者

我们正生活在一个激动人心的时代。每天你都能发现全新的软件或电子设备，它们为你完成特定工作、提供娱乐、联系亲朋好友、以图片或视频的形式记录你的生活等。不管是你口袋里巴掌大的智能手机、手腕上小巧的智能手表，还是你联网的智能汽车，大部分软件和设备的出现都归功于处理器技术的巨大发展。更便宜、更快速的处理器能运行优秀的软件库和软件框架，使得一些过去被认为只能由人类感知器官（例如眼睛和耳朵）来完成的任务，可以通过数学算法和计算机技术高效地实现。

计算机视觉是计算机科学的一个分支领域，近年来，随着硬件和软件技术的发展，这个领域有了革命性的进步。反过来，计算机视觉也同样深刻地影响了硬件和软件的发展。如今，从数码相机拍照这样简单的任务到理解自动驾驶汽车的环境从而躲避车祸这样复杂的任务，计算机视觉被应用到各个方面。它可以在视频录像中把人脸换成兔子，也可以在显微镜影像中检测癌变细胞和组织从而挽救人的生命。说到计算机视觉应用的可能性，可以说是没有限制，尤其考虑到软件库和软件框架的巨大进步，公司和开发者可以把他们无穷无尽的想法变成现实。

在过去的几年里，OpenCV（全称是 Open Source Computer Vision library，开源计算机视觉库）已经发展为一套完整的计算机视觉开发工具。它涵盖了几乎所有你能想到的解决计算机视觉问题所需的功能：从最基本的图像操作，如缩放和滤镜，到可用于快速准确地检测物体的机器学习算法训练模型。它也提供了开发计算机视觉应用需要的几乎所有模块，并支持一些最流行的编程语言，如 C++ 和 Python。你甚至能找到 .NET 框架的接口。OpenCV 能在所有主流操作系统上运行，包括移动系统和桌面平台。

本书的目标是用实际的例子和项目教会读者如何使用 OpenCV 来进行计算机视觉应用的开发。本书每一章都会讲解一些对应主题的计算机视觉算法，按照章节顺序，读者能学到大量可实际应用到项目中的计算机视觉算法。虽然本书大部分算法都是独立的，但强烈推荐读者从头开始阅读，并尝试按章节顺序建立计算机视觉知识体系。请读者亲手实践书中的每个例子，这些例子都很有趣，能让读者在实践过程中树立自信。

本书是数月辛勤工作的成果，而且如果没有 Tiksha Sarang 耐心而出色的编辑、Adhithya Haridas 精准深刻的技术评审和批注以及 Sandeep Mishra 提供的这次难得的出版机会，本书是不可能完成的。感谢技术评审 Zhuo Qingliang 先生的大力帮助和其他所有协助我完成并出版这本书的 Packt 出版社的员工以及计算机视觉开源社区的成员。

本书的目标读者

任何有扎实 C++ 编程语言基础和第三方软件库使用经验的开发人员都能轻松阅读本书和上手书中的例子。另外，熟悉 Python 编程语言的开发者也能使用本书学习使用 OpenCV 软件库，但需要自己把 C++ 的例子转为 Python，因为本书中出现的算法主要是 C++ 版的。

本书内容

第 1 章介绍计算机视觉科学的基础——它是什么，它用在哪，图像的定义和基本属性，例如像素、深度和通道等。该章比较短，是针对机器视觉领域的入门读者进行介绍的章节。

第 2 章通过概述 OpenCV 开发最重要的组成模块来介绍 OpenCV 软件库及其核心部件。读者也能了解如何获取和使用 OpenCV。该章会简略介绍用 CMake 创建 OpenCV 工程，让读者学会 Mat 类及其变体、图像视频的读写操作，以及获取摄像头输入（包括其他类型的输入方式）。

第 3 章涵盖了用于创建和修改矩阵的基础算法。读者将学习如何进行矩阵运算，比如向量乘积、标量乘积和矩阵求逆。除了求均值、求和和傅里叶变换这些数学运算，这一章将引入很多基于矩阵元素的运算。

第 4 章尽可能多地覆盖了本书涉及的各类图像处理算法。读者将学习如何在图像中绘制图形和文字，还将学习如何绘制直线、箭头、矩形等。该章还将展示各种广泛应用的图像滤镜操作算法，例如图像的平滑、膨胀、腐蚀和形态操作。学完该章，读者将熟悉功能强大的重映射算法和计算机视觉领域使用的各种色彩映射表。

第 5 章引入直方图的概念并介绍单通道和多通道图像计算直方图的方法。读者将学习灰度和彩色图像的直方图数据可视化，或者说，由像素的色调值来计算直方图。读者还将学习反向投影这种直方图的逆运算。直方图的比较和均衡也是该章的主题。

第 6 章解释了如何处理视频，尤其是实时目标检测和追踪，这些操作运用了一些流行的计算机视觉追踪算法。在简略介绍了视频处理的基础知识之后，读者将结合实例学习目标追踪应用场景下的均值偏移（Mean Shift）、CAM 偏移（CAM Shift）算法及卡尔曼滤波（Kalman filtering）。学完该章，读者将掌握背景和前景提取算法及其在实践中的使用方法。

第 7 章首先是目标检测的简介，接着引入各类用于形状分析的算法。该章涵盖的主题还包括关键点检测、描述符提取和描述符匹配，它们被用于基于特征（而不是基于简单的色调或亮度值）的目标检测中。

第 8 章涵盖了 OpenCV 里的机器学习 (ML)、神经网络 (DNN) 模块和一些重要的算法、类和函数。从 SVM 算法开始, 读者将学习如何用多种类似的训练流程来训练模型, 然后用模型来对输入进行分类。读者将学习如何用 HOG 描述符和 SVM 对图像分类。该章也涵盖了 OpenCV 中构建神经网络的方法, 并介绍了级联分类。8.5 节将教授读者如何使用第三方软件库生成的模型 (例如 TensorFlow 模型) 来实时检测多个物体。

如何从书中获益更多

尽管所有章节里需要的工具和软件在每章的开头都有描述, 作为一个简单快速的参考, 这里列表如下:

- 一台普通计算机, 要求安装了较新版本的 Windows、macOS 或 Linux (比如 Ubuntu) 操作系统
- Microsoft Visual Studio (Windows 系统)
- Xcode (macOS 系统)
- CMake
- OpenCV

首先, 读者可以通过网上搜索或是询问当地的计算机商店来了解目前主流计算机的配置, 不过读者手头用的机器一般足够学习使用了。

另外, 使用哪种集成开发环境 (IDE) 或是编译系统 (在本例中为 CMake) 与书中提供的例子关系不大。例如, 只要熟悉 OpenCV 软件库在其中的设置方法, 读者完全可以使用任何代码编辑器或编译系统来学习书中的例子。

下载示例代码及彩色图像

本书的示例代码及所有截图和样图, 可以从 <http://www.packtpub.com> 通过个人账号下载, 也可以访问华章图书官网 <http://www.hzbook.com>, 通过注册并登录个人账号下载。

本书的代码包在 <https://github.com/PacktPublishing/Hands-On-Algorithms-for-Computer-Vision> 也能找到。这个 GitHub 仓库里的代码会保持更新。

约定

 警告或要点使用这样的图标。

 提示和技巧使用这样的图标。

关于作者

Hands-On Algorithms for Computer Vision

阿明·艾哈迈迪·塔兹赫孔迪 (Amin Ahmadi Tazehkandi) 是一位伊朗作家、软件开发人员和计算机视觉专家。他在伊朗完成了计算机软件工程学习，并为世界各地的众多软件和工业公司工作。

我要感谢我的妻子 Senem，她是我的激情、爱和力量的象征。我还要感谢我的家人，他们是一群由一位发明家父亲和慈爱的母亲所生的杰出工程师，感谢他们无条件的爱和支持。

Zhuo Qingliang (KDr2 online) 目前在金融科技创业公司 paodingai 工作，该公司致力于通过使用人工智能技术改善金融业。他在 Linux、C、C++、Java、Python 和 Perl 开发方面拥有超过 10 年的经验。他对编程、咨询工作感兴趣并为开源社区提供帮助。

他拥有一个个人网站 KDr2，在那里你可以找到更多关于他的信息。

目 录

Hands-On Algorithms for Computer Vision

译者序

前言

关于作者

关于审稿人

第 1 章 计算机视觉概述 1

- 1.1 技术要求 1
- 1.2 理解计算机视觉 1
- 1.3 理解计算机图像 3
 - 1.3.1 色彩空间 5
 - 1.3.2 输入、处理和输出 7
- 1.4 计算机视觉框架和软件库 8
- 1.5 总结 9
- 1.6 习题 9

第 2 章 OpenCV 入门 11

- 2.1 技术要求 12
- 2.2 OpenCV 介绍 12
- 2.3 OpenCV 的下载、编译和安装 14
- 2.4 在 C++ 或 Python 项目中使用 OpenCV 17
- 2.5 理解 Mat 类 19
 - 2.5.1 创建一个 Mat 对象 20
 - 2.5.2 删除一个 Mat 对象 23
 - 2.5.3 访问像素 24
- 2.6 图像读写 27
- 2.7 视频文件读写 29
 - 2.7.1 使用摄像头 31
 - 2.7.2 使用 RTSP 和网络流媒体 32
- 2.8 Mat 类家族 32
- 2.9 总结 32
- 2.10 习题 33
- 2.11 补充阅读 33

第 3 章 数组和矩阵操作 34

- 3.1 技术要求 34
- 3.2 Mat 类中的操作 35
 - 3.2.1 克隆矩阵 35
 - 3.2.2 计算叉积 35
 - 3.2.3 提取对角线 36
 - 3.2.4 计算点积 36
 - 3.2.5 学习单位矩阵 37
 - 3.2.6 矩阵求逆 37
 - 3.2.7 元素级矩阵乘法 37
 - 3.2.8 全一和全零矩阵 37
 - 3.2.9 矩阵转置 38
 - 3.2.10 重塑 Mat 对象 39
- 3.3 元素级矩阵操作 39
 - 3.3.1 基本操作 39
 - 3.3.2 按位逻辑操作 43
 - 3.3.3 比较操作 47
 - 3.3.4 数学操作 48
- 3.4 矩阵和数组级操作 49
 - 3.4.1 为外推法生成边界 50
 - 3.4.2 翻转（镜像）和旋转图像 51
 - 3.4.3 使用图像通道 52
 - 3.4.4 数学函数 54
 - 3.4.5 搜索和定位功能 58
- 3.5 总结 60
- 3.6 习题 61

第 4 章 绘图、滤镜和变换 62

- 4.1 技术要求 62
- 4.2 在图像上绘图 63
 - 4.2.1 在图像上打印文字 63
 - 4.2.2 在图像上绘制形状 66
- 4.3 图像滤镜 72

4.3.1 模糊/平滑滤镜	72	6.6 如何提取背景和前景	129
4.3.2 形态滤镜	76	6.7 总结	132
4.3.3 基于导数的滤镜	79	6.8 习题	132
4.3.4 任意滤镜	80		
4.4 图像变换	81	第7章 对象检测——特征和描述符	133
4.4.1 阈值算法	81	7.1 技术要求	133
4.4.2 色彩空间和类型转换	83	7.2 用于对象检测的模板匹配	134
4.5 几何变换	84	7.3 检测角点和边缘	137
4.6 使用色彩表	86	7.3.1 学习 Harris 角点检测算法	137
4.7 总结	88	7.3.2 边缘检测算法	143
4.8 习题	88	7.4 轮廓计算和分析	147
第5章 反向投影和直方图	89	7.5 特征检测、描述和匹配	152
5.1 技术要求	89	7.6 总结	158
5.2 理解直方图	90	7.7 习题	158
5.3 直方图反向投影	95		
5.4 直方图比较	103	第8章 机器学习与计算机视觉	160
5.5 直方图均衡	105	8.1 技术要求	160
5.6 总结	106	8.2 支持向量机	161
5.7 习题	107	8.3 用人工神经网络训练模型	169
5.8 补充阅读	107	8.4 级联分类算法	171
		8.4.1 使用级联分类器进行对象检测	171
第6章 视频分析——运动检测和追踪	108	8.4.2 训练级联分类器	174
6.1 技术要求	108	8.5 使用深度学习模型	180
6.2 视频处理	109	8.6 总结	184
6.3 理解均值偏移算法	112	8.7 习题	184
6.4 使用连续自适应均值偏移算法	119		
6.5 使用卡尔曼滤波器进行运动追踪和降噪	122	习题答案	185

计算机视觉概述

毫无疑问，计算机科学，尤其是实现算法的方式，近年来得到了长足的发展。这是由于个人计算机，甚至你口袋里的智能手机都比之前的产品快了很多，也便宜了很多。同许多计算机重要的领域一样，计算机视觉也因此受到了深刻的影响。计算机视觉算法的实现和使用方法近年来变化很大。从本概述开始，本书会逐步揭开计算机视觉算法的面纱，展现实现这些算法的最新最前沿的方法。

作为概述，本章首先介绍绝大部分计算机视觉算法中会用到的基础概念。即使读者对计算机视觉及其基础（比如图像、像素、通道的概念等）已经比较熟悉了，我们仍然建议通读本章，以确保正确理解这些计算机视觉基础概念，或者是加深印象。

在本章，我们会从计算机视觉领域的简单介绍开始，然后通过例子介绍一些计算机视觉应用最重要的行业领域。之后，我们就从图像开始深入计算机视觉的一些基本概念。我们会带领读者学习计算机视觉中图像及其构成单元的本质。在此过程中，学习内容会覆盖像素、深度和通道，这些概念对于正确理解和使用计算机视觉算法至关重要。

到本章结束，读者会学到如下内容：

- 计算机视觉是什么？哪里会用到它？
- 计算机视觉中图像的概念是什么？
- 像素、深度与通道的概念及它们之间的关系是什么？

1.1 技术要求

本章是概述性的，我们主要专注理论的介绍，因此本章没有技术要求。

1.2 理解计算机视觉

给计算机视觉下定义不是件简单的事情，而且计算机视觉专家在这件事情上也难以达成共识。因此，定义计算机视觉不在本书的范围和兴趣之内，我们只会给出一个简单实用且恰当的定义。历史上，计算机视觉是图像处理的同义词，特指一类以一张图像作

为输入，在一系列操作之后，生成另一张图像或者一组输出值（或测量值）的方法和技术。而如今你会发现，计算机视觉工程师在谈论这个概念时，他们多半指的是能够模拟人类视觉的算法，比如看到（检测）图像中的一个物体或者人。

那么，本书要使用哪一个概念呢？答案是两者兼有。简而言之，计算机视觉是处理数字可视数据（或任何可视化数据）的算法、方法和技术。这里可视数据不一定是从传统摄像头获取的图像，它们可能是地图的图示或者海拔、热强度图，或者任何可视化的数据，而不管它们在现实世界中具体是什么含义。

根据这个定义，如下问题（以及更多其他问题）都可以用计算机视觉来解决：

- 如何柔化或者锐化一幅图像？
- 如何减小一幅图像的大小？
- 如何增加或者降低一幅图像的亮度？
- 如何检测一幅图像最亮的区域？
- 如何检测并追踪视频（或者连续图像）中的人脸？
- 如何识别监控视频中的人脸？
- 如何检测视频中的运动？

i 在现代计算机视觉领域，图像处理通常是计算机视觉算法的一部分，用来进行图像滤波、变换等。不过，很多人会互换使用“计算机视觉”和“图像处理”这两个词。

现在，计算机视觉是计算机科学和软件行业最火的主题之一。因为计算机视觉被应用在很多领域，在应用软件、数字设备或者工业机器上，它用计算机算法实现了原来依赖人眼才能完成的工作。在很多行业都有实际的例子，比如汽车、电影、生物医学仪器、安防、照片编辑、共享工具及视频游戏。我们会讨论其中的一部分例子，剩余的留给读者自行研究。

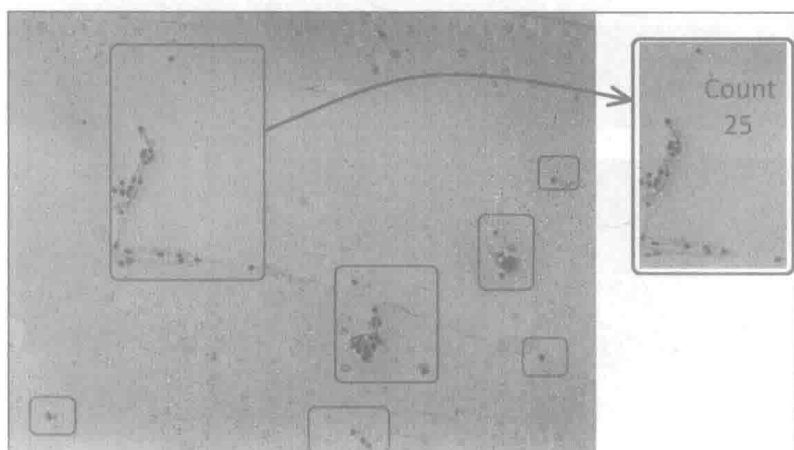
在汽车行业，计算机视觉被用来提高安全性和便捷性。计算机视觉使得汽车能够检测交通信号、提醒驾驶员不要超速甚至检测道路上的车道和障碍物来对可能的危险进行预警。我们无法穷尽计算机视觉颠覆汽车行业的所有可能，更不要说无人驾驶了。主要的科技公司都对此投入了巨量的资源，甚至把它们取得的成果共享给开源社区。在本书的最后一章，我们会使用这些成果展示如何实时检测多种类型的多个目标。

下图描述了汽车行业关心的一些物体、交通标志和其他感兴趣的区域，这些都是挂载在车上的摄像头常拍到的图像：



另一个面临技术革命的例子是生物医学行业。不只是器官和人体成像方法有了很大的提高，这些图像的理解和可视化也因为计算机视觉算法的进步而显著增强了。用极高精度的显微镜获取器官的图像，然后利用计算机可以检测器官是否癌变。机器人行业也涌现出许多有前景的方向，例如用机器人做手术。

下图中，计算机视觉算法可以在数字显微镜扫描出来的图像中统计生物组织的不同区域分别有多少感兴趣的特定生物目标（此例为细胞）：



除了汽车和生物医学之外，计算机视觉也被广泛用在移动和桌面应用程序中来解决各种各样的问题。浏览手机上的应用程序商店看一下与计算机视觉相关的应用，你会发现，只有你想不到的，没有未来计算机视觉做不到的。

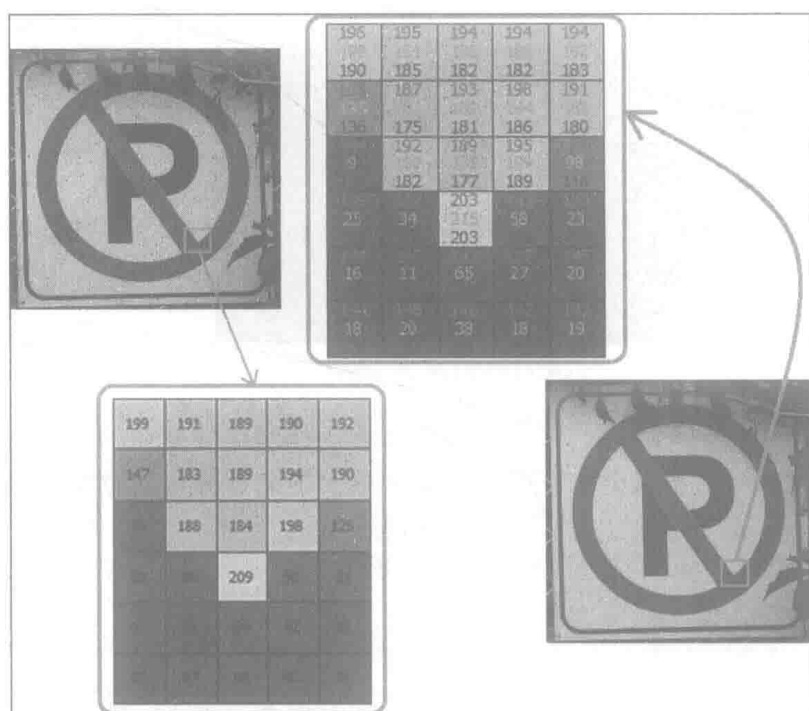
1.3 理解计算机图像

现在，我们从图像开始学习计算机视觉的基本概念。那么，图像究竟是什么？在计

计算机视觉中，一幅图像就是一个矩阵，换句话说，就是一个二维向量。用这种方式看待图像不但可以简化图像自身的表示，而且可以简化图像的组成元素：

- 图像的宽度对应矩阵的列数。
- 图像的高度就是矩阵的行数。
- 矩阵的每个元素代表一个像素，像素是图像最基本的组成部分，图像就是像素的集合。
- 每个像素（或矩阵的每个元素）可以包含一到多个数值来对应其视觉表示（颜色、亮度等）。我们后面讲到计算机视觉中色彩空间的概念时再涉及。但应注意，像素的每个数值表示一个通道。例如，灰度图像的像素通常用无符号 8 位整型来表示一个 0 ~ 255 的数值，因此，灰度图像是单通道的。在这种表示法中，0 代表黑色，255 代表白色，其他数值代表对应的灰度值。再如标准的 RGB 图像表示法，它用 3 个无符号 8 位整型数来表示，每个整型数介于 0 ~ 255 之间。RGB 图像的每个像素，用这三个通道对应其红、蓝、绿分量的强度值，这三个分量结合在一起可以表示任意可能的颜色。这样的图像被称为三通道图像。

下图是同一幅图像的灰度图版本和彩色图（RGB）版本的同一区域放大后的结果。读者应该注意到灰度图像（左侧）中越高的数值对应了越高的亮度，反之亦然。类似地，在彩色图像（右侧）中，读者能看到红色通道的数值很高，对应这个区域颜色是偏红的，当然白色区域的红色通道数值也很高。



除了上述提到的内容，图像还有如下属性：

- 每个像素（矩阵的元素）可以是整数或者浮点数，可以是8位、16位等。表示每个像素的数值的类型和通道数量一起构成了图像的深度。例如，假设一个四通道的图像使用16位整数表示每个通道，那么其深度就是16乘以4位，即64位（或者说4个字节）。
- 图像分辨率反映了图像中的像素数量。例如，一个宽度1920、高度1080的图像（也称为full-HD图像）的分辨率是1920乘以1080，包含超过200万个像素（约2兆像素）。

正是因为这种表示方式，图像可以轻松地转化为数学元素，进而在不同类型的算法中使用。回到最简单的图像表示法（灰度图像），通过几个简单的例子我们就可以看到，大多数图像编辑软件（以及计算机视觉算法）都用这种表示法，配合简单的算法及矩阵操作可以轻松修改图像。下图中，在输入图像（中间）的每个像素上加一个常数（此例中是80），就可以提高结果图像的亮度（右图），也可以从每个像素上减去一个常数来让图像变得更暗（左图）。



i 现在，我们先聚焦在计算机视觉的基本概念上，不深究图像编辑的实现细节。我们在后续章节再讨论其细节以及其他图像处理的技术和算法。

本节介绍的图像属性（宽度、高度、分辨率、深度和通道）在计算机视觉中被广泛使用。例如，在一些情况下，如果一个图像处理算法特别复杂耗时，我们会先缩小图像，这样处理起来就会快很多。等处理完毕，处理的结果可以被映射回原图大小并展示给用户。对于深度和通道也可以用类似的处理方法，如果算法只需要图像特定的通道，你可以提取出相应的通道，或者干脆把图像转成灰度图像。注意，一旦目标检测完成，应该在原来的彩色图像上展示检测结果。对这些图像属性的正确理解有助于你使用计算机视觉算法来解决各种各样的问题。言归正传，下面我们开始讨论色彩空间。

1.3.1 色彩空间

尽管有不同的定义方法，但通常色彩空间（有时称为色彩模型）是一种用来解释、