

华章程序员书库

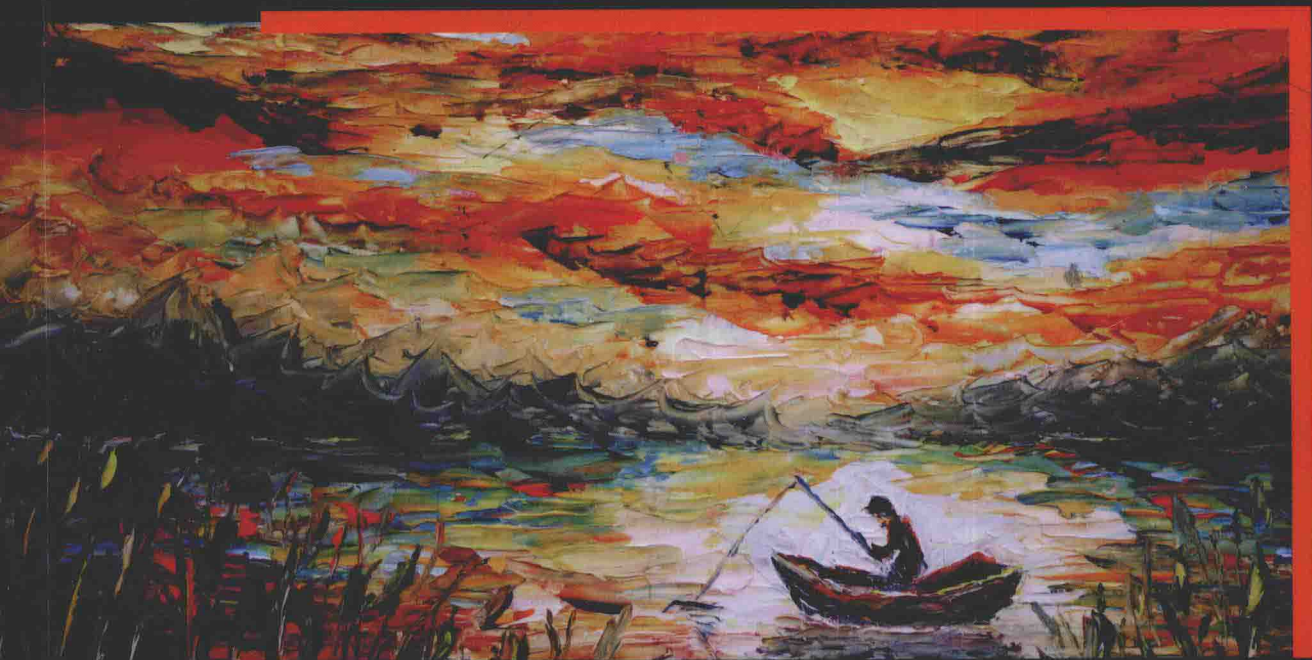
[PACKT]
PUBLISHING

HZ Books
华章IT

OpenCV

实例精解

OpenCV By Example



[美] 普拉蒂克·乔希 (Prateek Joshi)

[西班牙] 大卫·米兰·埃斯克里瓦 (David Millán Escrivá) 著

[巴西] 维尼修斯·戈多伊 (Vinicius Godoy)

呆萌院长 李风明 李翰阳 译



机械工业出版社
China Machine Press

华章程序员书库



OpenCV By Example

OpenCV

实例精解



[美] 普拉蒂克·乔希 (Prateek Joshi)

[西班牙] 大卫·米兰·埃斯克里瓦 (David Millán Escrivá) 著

[巴西] 维尼修斯·戈多伊 (Vinícius Godoy)

呆萌院长 李风明 李翰阳 译



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

OpenCV 实例精解 / (美) 普拉蒂克·乔希 (Prateek Joshi) 等著; 呆萌院长, 李风明, 李翰阳译. —北京: 机械工业出版社, 2016.8

(华章程序员书库)

书名原文: OpenCV By Example

ISBN 978-7-111-54741-9

I. O… II. ①普… ②呆… ③李… ④李… III. 图像处理软件—程序设计 IV. TP391.413

中国版本图书馆 CIP 数据核字 (2016) 第 200466 号

本书版权登记号: 图字: 01-2016-1584

OpenCV By Example (ISBN: 978-1-78528-094-8).

Copyright © 2016 Packt Publishing. First published in the English language under the title “OpenCV By Example”.

All rights reserved.

Chinese simplified language edition published by China Machine Press.

Copyright © 2016 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。

OpenCV 实例精解

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: 陈佳媛

责任校对: 董纪丽

印 刷: 北京诚信伟业印刷有限公司

版 次: 2016 年 9 月第 1 版第 1 次印刷

开 本: 186mm×240mm 1/16

印 张: 14

书 号: ISBN 978-7-111-54741-9

定 价: 59.00 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) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@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

The Translator's Words 译者序

通俗地讲，计算机视觉就是给计算机安装上眼睛（照相机）和大脑（算法），让其能够感知周围的环境。它是对生物视觉的一种模拟，通常的做法是通过对采集的图像或者视频进行处理来获得相应场景的三维信息。计算机视觉不仅应用在计算机科学和工程、信号处理、物理学、应用数学和统计学中，也广泛应用在神经生理学和认知科学等领域，发展前景可见一斑。

工欲善其事，必先利其器。作为如今开发计算机视觉应用最流行的库之一，OpenCV 不但能够实时运行许多不同的计算机视觉算法，而且几乎可以兼容所有的平台。

本书本着学以致用用的精神，每章都包含现实世界的例子和示例代码，以帮助读者更好地了解它们在现实生活中的应用。桃李不言，下自成蹊，对本书最真实的评价，来自于广大的读者朋友。

本书翻译的过程并不短暂，作为译者，我们尽可能地忠于原著。对于本书中大量的专业术语尽量遵循已有的标准，并参阅了大量文献，以便于读者理解。

全书由呆萌院长、李风明和李翰阳共同完成翻译。由于水平有限，书中出现的错误和不妥之处，恳请读者批评指正。

译者

2016 年 6 月

前 言 *Preface*

OpenCV 是开发计算机视觉应用最流行的库之一。它使我们能够实时运行许多不同的计算机视觉算法。它已经存在了很多年，并成为这个领域的标准库。OpenCV 的主要优点之一是它的高度优化和几乎可以在所有平台上兼容。

本书首先介绍了计算机视觉中的各个领域和在 C++ 中相关的 OpenCV 功能。每章都包含真实世界的例子和示例代码帮助你轻松地掌握主题，并了解它们在现实生活中的应用。总之，本书是一部实用指南，会教你如何在 C++ 中使用 OpenCV，并建立各种应用程序。

本书的主要内容

第 1 章涵盖各种操作系统的安装步骤，介绍了人类视觉系统，以及计算机视觉中的各种主要内容。

第 2 章讨论如何在 OpenCV 中读 / 写图像和视频，并且介绍如何使用 CMake 建立一个项目。

第 3 章介绍如何通过创建一个图形用户界面和鼠标事件检测器来实现交互式应用程序。

第 4 章探讨直方图和滤波器，也演示了如何卡通化图像。

第 5 章描述了各种图像的预处理技术，如去除噪声、阈值化，以及轮廓分析。

第 6 章处理对象识别和机器学习，并学习如何使用支持向量机建立一个对象分类系统。

第 7 章讨论了人脸检测和 Haar 级联，并解释如何使用这些方法来检测人脸的各个

部分。

第 8 章探索背景差分、视频监控和形态学图像操作，并描述了它们如何彼此关联。

第 9 章介绍如何使用不同的技术跟踪对象，如基于颜色和基于特征。

第 10 章介绍光学字符识别、文本分割和 Tesseract OCR 引擎。

第 11 章深入研究 Tesseract OCR 引擎，介绍如何将它应用于文本检测、提取和识别。

你需要准备什么

本书的例子会用到以下技术：

- OpenCV 3.0 或更新的版本
- CMake 3.3.x 或更新的版本
- Tesseract
- Leptonica (Tesseract 依赖包)
- QT (可选)
- OpenGL (可选)

相关章节提供了详细的安装说明。

本书的读者对象

本书面向 OpenCV 初学者，以及希望在 C++ 中使用 OpenCV 进行计算机视觉应用开发的开发人员。懂得 C++ 的基础知识将有助于理解本书。本书对于想要开始学习计算机视觉，并了解基本概念的人来说同样适用。他们应该知道基本的数学概念，如向量、矩阵、矩阵乘法，等等，这样才能最大限度地利用本书。在阅读本书的过程中，你将从头学习如何使用 OpenCV 创建各种计算机视觉应用。

下载示例代码

可登录 <http://www.hzbook.com>，下载本书示例代码。

目 录 *Contents*

译者序

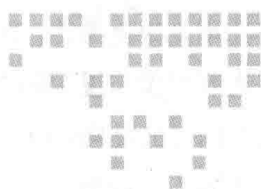
前言

第 1 章 OpenCV 的探险之旅	1
1.1 理解人类视觉系统	1
1.2 人类是怎么理解图像内容的	3
1.3 OpenCV 可以做什么	4
1.4 安装 OpenCV	11
1.5 总结	14
第 2 章 OpenCV 基础知识介绍	15
2.1 CMake 基本配置文件	15
2.2 创建库	16
2.3 管理依赖关系	17
2.4 脚本复杂化	19
2.5 图像和矩阵	21
2.6 读写图像	23
2.7 读取视频和摄像头	27
2.8 其他基本对象类型	30
2.9 矩阵的基本运算	33

2.10	基本数据持久性和存储	36
2.11	总结	38
第3章	图形用户界面和基本滤波	39
3.1	介绍 OpenCV 的用户界面	39
3.2	使用 OpenCV 实现基本图形用户界面	40
3.3	QT 的图形用户界面	45
3.4	在界面上添加滑动条和鼠标事件	47
3.5	在用户界面上添加按钮	51
3.6	支持 OpenGL	55
3.7	总结	60
第4章	深入研究直方图和滤波器	61
4.1	生成 CMake 脚本文件	62
4.2	创建图形用户界面	63
4.3	绘制直方图	65
4.4	图像色彩均衡化	69
4.5	LOMO 效果	71
4.6	卡通效果	76
4.7	总结	80
第5章	自动光学检测、目标分割和检测	81
5.1	隔离场景中的目标	82
5.2	创建 AOI 应用程序	84
5.3	输入图像的预处理	86
5.4	分割输入图像	92
5.5	总结	101
第6章	学习目标分类	102
6.1	介绍机器学习的概念	103

6.2 计算机视觉和机器学习的工作流程	106
6.3 自动检测目标分类的示例	108
6.4 特征提取	110
6.5 总结	120
第 7 章 识别人脸部分并覆盖面具	121
7.1 理解 Haar 级联	121
7.2 积分图	123
7.3 在实时视频中覆盖上面具	124
7.4 戴上太阳镜	127
7.5 跟踪鼻子、嘴和耳朵	130
7.6 总结	131
第 8 章 视频监控、背景建模和形态学操作	132
8.1 理解背景差分	132
8.2 简单背景差分法	133
8.3 帧差值法	137
8.4 混合高斯方法	141
8.5 形态学图像操作	144
8.6 图像细化	145
8.7 图像加粗	146
8.8 其他形态学运算	147
8.9 总结	152
第 9 章 学习对象跟踪	153
9.1 跟踪特定颜色的对象	153
9.2 建立交互式对象跟踪器	156
9.3 使用 Harris 角点检测器检测点	161
9.4 Shi-Tomasi 角点检测器	163
9.5 基于特征的跟踪	166

9.6 总结	175
第 10 章 文本识别中的分割算法	176
10.1 OCR 简介	176
10.2 预处理步骤	178
10.3 在你的操作系统上安装 Tesseract OCR	186
10.4 使用 Tesseract OCR 库	190
10.5 总结	195
第 11 章 使用 Tesseract 识别文本	196
11.1 文本识别 API 工作原理	196
11.2 使用文本识别 API	200
11.3 总结	212



OpenCV 的探险之旅

计算机视觉应用是很有趣也很有用的，但是它的基础算法是计算密集型的。伴随着云计算的到来，我们拥有越来越多处理这种算法能力。在实际情况下，OpenCV 库可以更有效地运行计算机视觉算法。它已经存在很多年了，并且已经成为这个领域的一个标准库了。OpenCV 的一个主要优势是它已经被高度优化，并且几乎支持在所有平台上使用。这本书即将介绍 OpenCV 的方方面面包括：我们使用的算法，为什么使用 OpenCV，以及怎么整合 OpenCV 到各个领域。

本章下面将要介绍如何在多操作系统环境下安装 OpenCV。OpenCV 提供的可以立即使用的功能有哪些，以及可以使用内置函数做到的事情。

在本章结束时候，你可以回答出以下几个问题：

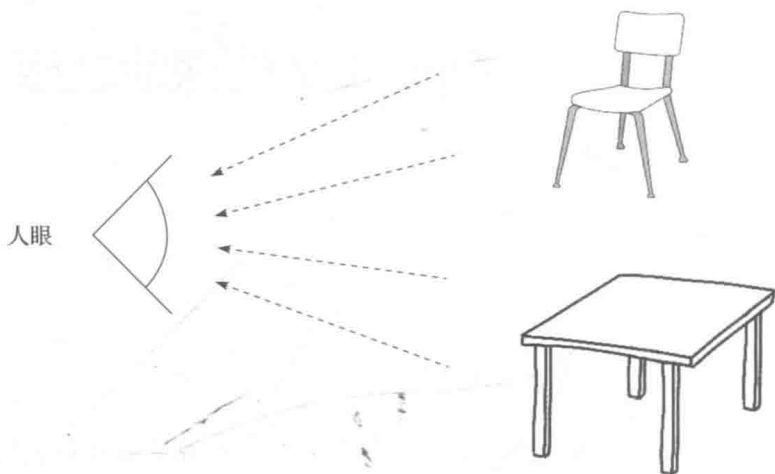
- 人们怎么处理视觉数据，以及怎么理解图像内容？
- OpenCV 可以做些什么，在 OpenCV 提供的大量模块中哪些可以用来完成这些事情？
- 如何在 Windows、Linux 以及 Mac OS X 上安装 OpenCV？

1.1 理解人类视觉系统

在接触 OpenCV 功能之前，我们首先需要了解这些功能为什么要创建。了解人类视

觉系统的工作原理是很重要的，因为只有这样你才能够开发出正确的算法。计算机视觉算法的目的是理解图像和视频的内容。人类似乎可以毫不费力地做到这一点！那么，如何使机器也具有同样的准确性呢？

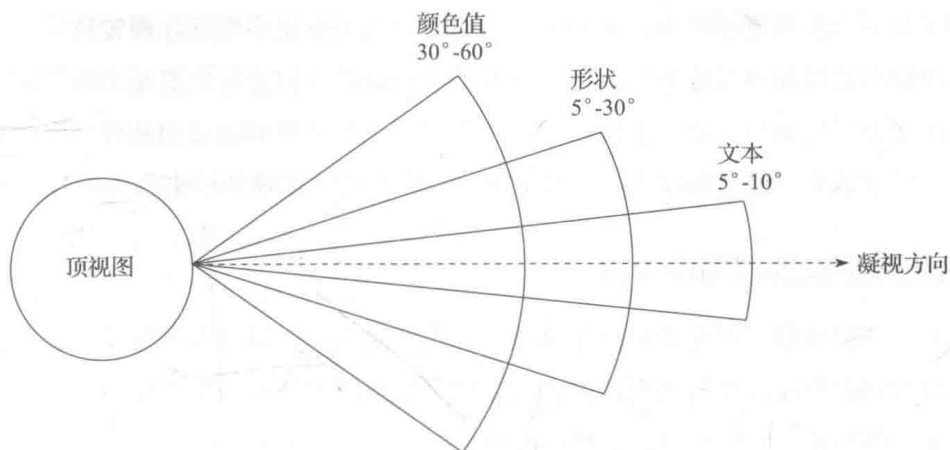
接下来，看看下面的图：



人类的眼睛会同时捕获色彩、形状、亮度等信息。在上图中，人眼捕获了这两个主体的所有信息，并以某种方式将它们存储起来。一旦明白了人类的视觉系统是如何工作的，我们就可以利用这个来实现预期结果。举个例子，下面有几件事情需要知道：

- 人类视觉系统对低频内容敏感程度高于高频内容。低频内容是指像素值不迅速改变的平面区域，高频内容是指像素值波动很大的角落和边缘地区。你可能已经注意到，如果在平坦的表面有斑点，就可以很容易地被发现，但是在质地不平的表面就很难被发现。
- 人眼对亮度的变化敏感程度高于颜色的变化。
- 人类视觉系统对运动的事物很敏感。如果有东西在视野中移动，即使人们没有直视它，也能很快地意识到。
- 人们往往会用心记住视野内突出的点。下面来想象一下，有一个白色的桌子，它的四条腿是黑色的，并且表面的某个角落上有一个红点。当看着这张桌子时，你会立即记住表面和腿有对立的颜色，并且其中一个角落有一个红点。人类大脑是很聪明的！它会立即做这些，以便下次遇到的时候能够对其进行快速识别。

为了解人类视野，接下来看看人类看不同事物的角度：



人类视觉系统实际上能够做更多的事情，但这样足以开始下面的内容了。你可以通过在互联网上阅读人类视觉系统模型来深入探索它。

1.2 人类是怎么理解图像内容的

环顾四周，你会看到很多对象。每天可能会遇到各种各样的对象，但你会毫不费劲地一眼识别出它们。事实上，当看到一张椅子，你不会等待几分钟才意识到那是张椅子。对，你会立马意识到那是张椅子。实际上从另一方面说，计算机很难做好这件事。研究者们进行了多年研究才找出为什么计算机不擅长做这种人类相当擅长的事情。

为了得到这个问题的答案，首先要理解人们是怎么做到的。视觉数据处理发生在腹侧视觉通路。这个腹侧视觉通路涉及与对象识别相关联的人类视觉系统回路。这是人类大脑中一块区域的基本层次结构，它会有助于对象识别。人们可以毫无费劲地认知不同事物，并且还可以将相似的对象归类成组。之所以可以做到这个，是因为人类开发对相同类别对象的不变性排序。当人们观察某个对象时，他们的大脑提取了一些特征点，例如方向、尺寸、观点，以及不要紧的光照等因素。

比正常大一倍尺寸并且倾斜 45 度角放着的椅子仍然是一张椅子。因为处理方式的原因，我们可以很轻松地识别出它。机器反而不能轻松处理好这样的情况。人们趋向于通过形状和一些重要的特征记忆一个对象。不管这个对象是如何摆放的，人们仍然可以认出它。在人类视觉系统中，大脑创建了可以帮助我们的稳健有关位置、缩放和角度方面的不变性层次结构。

如果你对人类视觉系统有深入研究，就会发现人们有很多细胞在视觉皮层。这些细胞可以识别出曲线和直线等形状。当深入腹侧通路时，我们会看到更复杂的细胞。这些细胞被训练去反应更为复杂的事物，例如树和门等。人类腹面通路中的神经元会在感觉域上显示尺寸增长。这也与加上它们首选的刺激的复杂性增加的事实相关联。

为什么机器很难理解图像内容

现在，我们理解了视觉数据是怎么进入人类视觉系统，以及人类视觉系统怎么处理它。目前的问题是还没理解透彻人类大脑如何识别和组织这些视觉数据。人们仅从图像中提取出一些特征，并且要求计算机通过机器学习算法学习人类。仍有很多变化例如形状、尺寸、观点、角度、光照、遮挡等。例如，在机器眼里，同样的椅子从侧面看起来不一样。不管它如何呈现，人们可以很容易地识别出它是一张椅子。但是应该如何跟计算机解释这个呢？

一种处理方法是将一个对象不同的变化存储起来，包括大小、角度、光照等。但是这样处理过于麻烦又太耗时。而且事实上，它不能将能遇到的每一种变化数据收集起来。为了识别出这些对象，计算机会消耗大量内存和时间去构建模型。即使能满足所有这些，当存在特殊遮挡的，计算机仍不能够识别出它，因为计算机会认为它是一个新事物。所以，在构建一个计算机视觉库时，我们需要构建基本功能块，那样就可以在各种各样的情况下结合成复杂的算法。OpenCV 提供了很多功能，并且这些功能得到了很好的优化。所以，一旦理解 OpenCV 提供的立即使用方法，我们就可以高效地使用它创建有趣的应用。OpenCV 的方法将在下一章具体介绍。

1.3 OpenCV 可以做什么

使用 OpenCV，你可以做相当多能够想象出的计算机视觉任务。现实生活中的问题需要使用很多函数块来完成预期结果。所以，还需要理解哪些模块和函数能达到预期的效果。下面开始介绍 OpenCV 提供的可以立即使用的方法。

1.3.1 内置数据结构和输入 / 输出

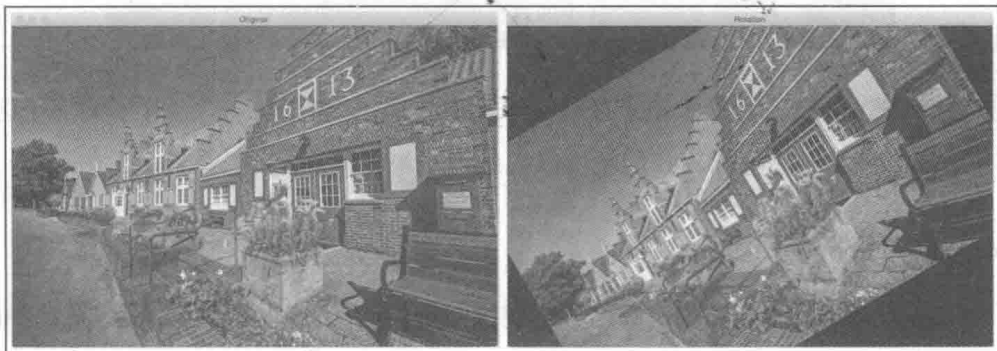
OpenCV 中最利好的消息是它提供了大量内置基元去处理涉及图像处理和计算机视觉的操作。如果从零开始写一些东西，你需要定义一些对象包括图像、点、矩形等。这

些几乎是任何计算机视觉算法的基础。OpenCV 提供了这些可以立即使用的基本框架，并且在核心模块中包含了它们。另一个优势是这些基本框架已经在运行速度和内存使用上进行了优化，所以不需要担心实现细节。

`imgcodecs` 模块处理图像文件的读写。当处理写入图像和创建图像文件时，你可以通过简单的命令将图像保存为 JPG 或者 PNG 格式的文件。当使用摄像机的时候，需要处理大量的视频文件。`videoio` 模块可以处理视频文件所有读写相关的操作。你可以很容易地从摄像头中获取视频，或者读取不同种格式的视频文件。甚至可以通过设置每秒帧播放速度、帧的大小等属性将一大堆的帧保存为视频文件。

1.3.2 图像处理方法

当编写计算机视觉算法时，会有一堆能反复使用的基本图像处理操作。`imgproc` 模块展示了大部分函数。你可以处理例如图像滤波，形态学操作，几何变换，色彩变换，绘制图像，结构分析，直方图，形状分析，运动分析，特征检测等事情。接下来思考下面的图：



右图是左图的一个旋转的版本，可以通过 OpenCV 中的一行代码做到这种转换。OpenCV 有个叫作 `ximgproc` 的模块，它包含了高级图像处理算法，例如基于结构森林的边缘检测，域变换滤波，自适应流形滤波等。

1.3.3 构建 GUI

OpenCV 提供了一个叫作 `highgui` 的模块，它是用来处理高级用户交互操作的。在处理下一步之前，讨论处理的问题和想要检查图像的样子。这个模块包括了创建用于展示图像或者视频的窗口等一系列函数。它还包括等待功能，那是等到用户触发键盘上按

键才能进行下一步。还有一个函数可以检测鼠标移动，它对开发交互应用很有帮助。使用这个功能就可以在输入窗口中绘画出长方形，处理被选择的区域。

考虑如下图：



从图中可以看出在图像上绘制了一个长方形，并且提供了一个底片影响那个区域。一旦有了这个长方形的坐标，我们就可以仅处理这块区域。

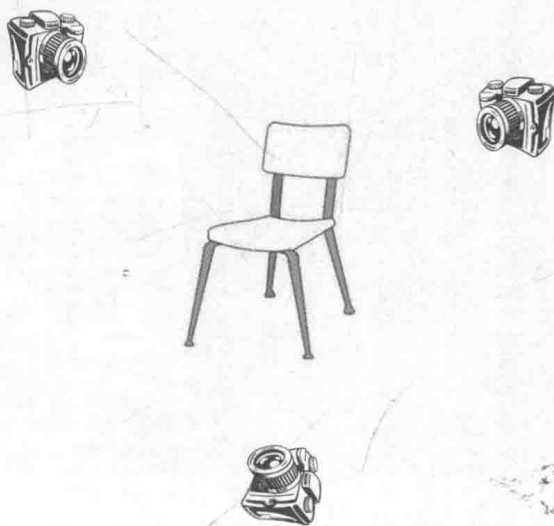
1.3.4 视频分析

视频分析包括了如下任务，例如分析视频中的连续帧之间的运动，跟踪视频中的不同对象，创建视频监控模型等。OpenCV 提供了 video 模型，它可以处理上面种种问题。还有个 videostab 模型用于视频去抖动。视频去抖动是摄像机中的一个重要组成部分。当用手举起相机捕捉视频，你很难保持手保持绝对静止。事实上，当观看这个视频时，会发现它看上去很差、摇摇晃晃的。所有现代设备在最后将视频展现给用户之前均使用了视频去抖动技术处理视频。

1.3.5 三维重建

三维重建是计算机视觉中的一个重要课题。通过使用相关算法，就可以将一系列的二维图像重建出三维场景。OpenCV 提供了可以发现二维图像中大量事物的相关性来

计算它们三维位置的算法。calib3d 模块可以处理所有这些。这个模块同样处理摄像机标定，它是一个相机的必要估计参数。这些参数是基本的内在参数，这些内在参数主要是将相机拍摄到的捕捉幕转化为图像。需要知道这些参数以便设计算法，否则会得到意想不到的结果。接下来看下面的图：



上图从不同角度捕捉相同的对象。接下来的任务就是通过 2D 图像重建原始对象。

1.3.6 特征提取

正如之前讨论的，人类视觉系统趋向于从一个给予场景中提取特征点，这样可以方便以后检索。为了模仿这点，人们开始设计很多特征提取器，目的是为了从已知图像上提取这些特征点。一些流行的算法包括 SIFT（尺度不变特征变换）、SURF（加速鲁棒特征）和 FAST（加速分段测试特征）等。features2d 模块提供了检测和提取这些特征的函数。xfeatures2d 模块提供了一些更多的特征提取器，其中一些还在实验中。如果想挑战，你可以试试这些试验中的特征提取器。其中一个叫作 bioinspired 的模块提供了计算机视觉仿生模型方面的算法。

1.3.7 目标检测

目标检测是指在给定图像中检测目标的位置。这一过程不关心目标的类型。如果设计一个椅子检测器，它只会告诉你在给定图像中椅子的位置，而不会告知是张红色高背