

数字图像处理

原理与应用

Tinku Acharya Ajoy K. Ray 著

田浩 葛秀慧 王顶 等译



IMAGE PROCESSING

Principles and Applications

清华大学出版社



Tinku Acharya, Ajoy K. Ray
Image Processing: Principles and Applications
EISBN: 0-471-71998-6

Copyright © 2007 by Wiley Publishing, Inc.

All Rights Reserved. This translation published under license.

Simplified Chinese translation edition is published and distributed exclusively by Tsinghua University Press under the authorization by McGraw-Hill Education (Asia) Co., within the territory of the People's Republic of China only, excluding Hong Kong, Macao SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书中文简体字翻译版由美国 John Wiley & Sons, Inc. 公司授权清华大学出版社在中华人民共和国境内(不包括中国香港、澳门特别行政区和中国台湾)独家出版发行。未经许可之出口, 视为违反著作权法, 将受法律之制裁。未经出版者预先书面许可, 不得以任何方式复制或抄袭本书的任何部分。

北京市版权局著作权合同登记号 图字 01-2007-1846 号

本书封面贴有 John Wiley & Sons 公司防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

数字图像处理: 原理与应用/ () 阿查里雅 (Acharya, T.) 等著; 田浩等译. —北京: 清华大学出版社, 2007.7

(世界著名计算机教材精选)

书名原文: Image Processing: Principles and Applications

ISBN 978-7-302-15222-4

I. 数… II. ①阿… ②田… III. 数字图像处理—教材 IV. TN911.73

中国版本图书馆 CIP 数据核字 (2007) 第 073431 号

责任编辑: 龙啟铭

责任校对: 张 剑

责任印制: 何 芊

出版发行: 清华大学出版社

<http://www.tup.com.cn>

c-service@tup.tsinghua.edu.cn

地 址: 北京清华大学学研大厦 A 座

邮 编: 100084

社 总 机: 010-62770175

邮购热线: 010-62786544

投稿咨询: 010-62772015

客户服务: 010-62776969

印 刷 者: 北京市清华园胶印厂

装 订 者: 三河市溧源装订厂

经 销: 全国新华书店

开 本: 185×260 印张: 20.25 字数: 487 千字

版 次: 2007 年 7 月第 1 版 印次: 2007 年 7 月第 1 次印刷

印 数: 1~3000

定 价: 39.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题, 请与清华大学出版社出版部联系调换。联系电话:

010-62770177 转 3103 产品编号: 023630-01

译者序

我们正在一个视觉迷人的世界中，图像处理越来越受到人们的重视，它广泛应用于自动视觉检测、遥感场景解译、防御观测、基于内容的图像检索、运动对象跟踪和生物医学成像等各种应用中，数字图像处理的重要性日益显著。

本书涵盖了数字处理的基本理论并且涉及许多先进应用，它的主要特色在于：

- 图像处理技术，包括小波变换和发展
- 图像增强和复原，包括噪声建模和滤波
- 分割方案及对象分类与识别
- 纹理与形状分析技术
- 在图像处理中模糊集理论方案及神经网络技术
- 基于内容的图像检索与图像挖掘
- 生物图像分析和解译，包括人脸识别和签名验证等
- 遥感图像及应用
- 动态场景分析和运动目标检测与追踪的基本原理和应用
- 图像压缩基础，包括 JPEG 标准和新的 JPEG2000 标准

本书由 Tinku Acharya 博士和 Ajoy K. Ray 博士执笔，Acharya 博士是 Avisere 公司的首席技术执行官，在加入 Avisere 之前，Acharya 博士在英特尔公司工作，他是 88 个美国专利的发明者和 15 项欧洲专利的拥有者。他的研究方向是图像处理、计算机视觉、生物测定、多媒体数据采集等。Ajoy K. Ray 博士是 Avisere 公司资深研究科学家。主要研究自动化和计算机视觉。

本书的特点在于系统性和实用性。本书除了给出数字图像处理的基本原理及相关的研究方向之外，还给出了最新的标准和技术，并通过详细的例子加以阐述说明。使许多原本抽象的理论具体化、实例化。

我们在本书的翻译过程中力求忠实于原著，在专业术语第一次出现的地方给出了英文注释，以方便读者理解和参照。通过翻译我们进一步认为这是一本非常值得深入研究的好书。

参加本书翻译的有：田浩、葛秀慧、王顶、刘秋红、刘朝晖、焦仁普、朱书敏、盖俊飞和张桂香等。由于本书涉及各个应用领域中的专业知识，如医学方面，内容全面，覆盖范围广，加之译者水平有限，书中的错误和不妥之处在所难免，敬请广大读者批评指正。

特别感谢清华大学出版社的龙啟铭编辑，也感谢为这本书辛勤工作的许多工作人员所付出的努力，使这本书能够与广大读者见面。

译者

前言

在不同应用领域中对图像处理的需求呈现出不断增长的趋势，例如多媒体计算、安全图像数据通信、生物医学成像、生物统计学、遥感、纹理理解、模式识别、基于内容的图像检索和压缩，等等。因此，人们迫切需要一本以全新的视角和内容介绍图像处理的书籍。我们尽可能地对这些最新的发展进行了介绍，还保留了经典的图像处理部分。

第 1 章介绍了图像处理技术的基本原理，并且还给出了整本书的组织结构。第 2 章涉及数字成像的基本原理和表示。第 3 章专门介绍了颜色和彩色影像。除了介绍色彩感知和色彩空间变换的基本原理之外，还介绍了颜色插值法（或去马赛克）的概念，目前，它是任何彩色成像设备必然集成的一部分。第 4 章介绍了多种图像变换技术。由于小波的突出特性，最近小波变换变得非常受欢迎。第 5 章将专门介绍小波变换。

理解在各种类型图像中都存在噪声的本质是很重要的。第 6 章详细介绍了图像增强和复原问题，包括噪声建模和滤波。在图像处理和模式识别中图像分割是一项很重要的任务。第 7 章将详尽阐述多种分割方法。一旦完成图像的适当分割之后，下一项重要任务就是图像中对象的分类与识别。第 8 章介绍了多种模式分类和对象识别技术。在图像理解中，纹理和形状将扮演重要角色，第 9 章详细介绍了一些纹理和形状分析技术。

与分类处理图像分析形成鲜明对比，模糊集合理论方法为许多图像处理任务提供了一流的方法学。第 10 章涉及了一些模糊集的理论方法。第 11 章介绍了基于内容的图像检索和图像挖掘。在生物医学工程中，生物医学的图像（像 X 光、超声波扫描术和 CT 扫描图像）为医疗诊断提供了大量信息。第 12 章专门讨论生物医学的图像分析和解译，描述了一些生物统计算法，特别是人脸识别、签名验证等。第 13 章介绍了用于遥感图像及其应用的技术。第 14 章描述了动态场景分析、运动对象检测与跟踪的基本原理和应用。在图像存储和传输中，图像压缩将起到很重要的作用。第 15 章专门介绍图像压缩的基本原理。第 16 章介绍了图像压缩的 JPEG 标准。第 17 章和第 18 章介绍了新的 JPEG2000 标准。

本书的读者对象是全世界的大学生和研究生、教师、科学家、工程师、在 R&D 和研究实验室的专家，本书可以作为一本很好的参考用书。

我们衷心地感谢 Chittabrata Mazumdar 先生，他使我们团结在一起，并合作完成这个项目。在我们努力完成这个项目过程中，还要感谢他一如既往的支持和鼓励。

感谢我们的编辑 Val Moliere 和她的助手 Wiley 对本项目的帮助。感谢所有在 Avisere 和 Indian Institute of Technologic, Kharagpur 的同事，特别是 Roger Undhagen 先生、Andrew. Griffiths 博士、G.S.Sanyal 教授、N.B. Chakrabarti 教授和 Arun Majumdar 教授，感谢他们一如既往的支持和鼓励。特别要感谢 Odala Nagaraju、Shyama P. Choudhury、Brojeswar

Bhowmick、Ananda Datta、Pawan Baheti、Milind Mushrif、Vinu Thomas、Arindam Samanta、Abhik Das、Abha Jain、Arnab Chakraborti、Sangram Ganguly、Tamalika Chaira、Anindya Moitra、Kaushik Mallick 和其他人以不同的方式直接或间接地对我们原稿的准备工作提供支持。感谢本书的匿名评论者和他们建设性的意见。

最后，还要感激我们家庭成员对这个项目不遗余力的支持。特别是 Baishali Acharya 夫人和 Supriya Ray 夫人，她们是我们强有力的后盾。我们还想对我们的孩子 Arita、Arani、Aniruddha 和 Ananya 表示衷心的感谢，因为他们总是很关心本项工作，并使我们为之骄傲。

Tinku Acharya
Ajoy K. Ray

目 录

第1章 概述	1	2.6 图像文件格式	22
1.1 图像处理基本原理	1	2.7 重点要点	23
1.2 图像处理应用	2	2.8 本章小结	24
1.2.1 自动视觉检测系统	2	参考文献	24
1.2.2 遥感场景解译	3	第3章 颜色和彩色影像	26
1.2.3 生物医学成像技术	3	3.1 概述	26
1.2.4 防御观测	4	3.2 色彩知觉	26
1.2.5 基于内容的图像检索	4	3.3 色彩空间量化和最小可觉差 (JND)	27
1.2.6 运动对象跟踪	4	3.4 色彩空间与变换	28
1.2.7 图像与视频压缩	5	3.4.1 CMYK 空间	28
1.3 人类视觉知觉	5	3.4.2 NTSC 或 YIQ 色彩空间	28
1.3.1 人眼	6	3.4.3 YC_bC_r 色彩空间	29
1.3.2 视觉的神经机制	6	3.4.4 视觉均匀色彩空间	29
1.4 图像处理系统的组成	7	3.4.5 CIELAB 色彩空间	31
1.5 本书的结构	9	3.5 颜色插值或去马赛克	32
1.6 与其他书的不同之处	10	3.5.1 非自适应颜色插值算法	32
1.7 本章小结	10	3.5.2 自适应算法	34
参考文献	11	3.5.3 新的基于自适应的色彩插值算法	38
第2章 成像与表示	12	3.5.4 实验结果	41
2.1 概述	12	3.6 本章小结	41
2.2 成像	12	参考文献	42
2.2.1 照明	12	第4章 图像变换	43
2.2.2 反射模型	13	4.1 概述	43
2.2.3 点扩散函数	14	4.2 傅立叶变换	43
2.3 采样与量化	15	4.2.1 一维傅立叶变换	43
2.3.1 图像采样	16	4.2.2 二维傅立叶变换	44
2.3.2 图像量化	17	4.2.3 离散傅立叶变换	45
2.4 二值图像	18	4.2.4 变换核	45
2.4.1 几何特性	19	4.2.5 矩阵形式表示	46
2.4.2 二值对象的链码表示	20	4.2.6 性质	46
2.5 三维成像	21	4.2.7 快速傅立叶变换	48
2.5.1 立体图像	21		
2.5.2 深度图像采集	22		

VI 数字图像处理：原理与应用

4.3 离散余弦变换	50	6.4.6 中值滤波	84
4.4 Walsh Hadamard 变换	51	6.5 图像增强的频域方案	84
4.5 Karhunen-Loeve 变换或主分量分析	52	6.6 噪声建模	87
4.5.1 协方差矩阵	53	6.7 图像复原	88
4.5.2 特征向量和特征值	54	6.7.1 含冲激噪声图像的复原	89
4.5.3 主分量分析	54	6.7.2 模糊图像的复原	90
4.5.4 奇异值分解	54	6.7.3 反向滤波	90
4.6 本章小结	56	6.7.4 Wiener 滤波器	91
参考文献	56	6.8 图像处理的其他方法	93
第 5 章 离散小波变换	57	6.8.1 重谱图像复原	93
5.1 概述	57	6.8.2 层析成像重建	94
5.2 小波变换	58	6.9 本章小结	94
5.2.1 离散小波变换	59	参考文献	94
5.2.2 Gabor 滤波	60	第 7 章 图像分割	96
5.2.3 多分辨率分析的概念	61	7.1 预处理	96
5.2.4 滤波和金字塔算法的实现	62	7.2 边缘、线与点检测	97
5.3 推广到二维信号	64	7.3 边缘检测	99
5.4 DWT 的提升实现	65	7.3.1 基于 Robert 算子的边缘 检测器	99
5.4.1 有限冲激响应滤波与 Z 变换	66	7.3.2 基于 Sobel 算子的边缘 检测器	99
5.4.2 Laurent 多项式的 Euclidean 算法	67	7.3.3 基于 Prewitt 算子的边缘 检测器	100
5.4.3 完全重建和滤波多相位表示	68	7.3.4 Kirsch 算子	101
5.4.4 提升	69	7.3.5 Canny 的边缘检测器	101
5.4.5 用于提升计算的数据关系图	74	7.3.6 基于二阶导数的算子	103
5.5 基于提升的 DWT 的优点	75	7.3.7 基于边缘分割的局限性	105
5.6 本章小结	75	7.4 图像阈值化技术	106
参考文献	76	7.4.1 两级阈值化	106
第 6 章 图像增强与复原	77	7.4.2 多级阈值化	107
6.1 概述	77	7.4.3 基于熵的阈值化	108
6.2 图像增强与复原的区别	78	7.4.4 遇到的问题与可能的解决方案	108
6.3 空间图像增强技术	78	7.5 区域生长	109
6.3.1 空间低通滤波和高通滤波	78	7.5.1 区域邻接图	110
6.3.2 均值和空间低通滤波	79	7.5.2 区域合并和分裂	110
6.3.3 模糊掩盖锐化与匀边	79	7.5.3 基于分割的聚类	111
6.3.4 方向平滑	80	7.6 用于分割的瀑布算法	112
6.4 基于直方图的对比增强	80	7.7 连通分量标记	112
6.4.1 图像的直方图	81	7.8 文档图像分割	113
6.4.2 直方图均衡化	81	7.9 本章小结	114
6.4.3 局部直方图均衡化	83	参考文献	114
6.4.4 直方图规定化	83		
6.4.5 双曲线直方图	83		

第 8 章 图像模式识别	116	9.5.2 多边形形状描述符	141
8.1 概述	116	9.5.3 形状描述中的控制点	142
8.2 决策理论的模式分类	116	9.5.4 在形状确定中曲率的角色	142
8.3 贝叶斯决策理论	117	9.5.5 用于形状分析的多边形逼近	142
8.3.1 参数估计	118	9.6 主动轮廓建模	143
8.3.2 最短距离分类	118	9.7 形状复原与归一化	145
8.4 非参数分类	120	9.7.1 形状散布矩阵	146
8.5 线性判别分析	121	9.7.2 平移和旋转坐标轴	146
8.6 无监督分类策略——聚类	121	9.7.3 改变基本尺度	147
8.6.1 单链路聚类	122	9.8 基于轮廓的形状描述符	148
8.6.2 完全链路聚类	122	9.9 基于区域的形状描述符	149
8.6.3 平均链路聚类	122	9.9.1 Zernike 矩	150
8.7 K-均值聚类算法	123	9.9.2 径向 chebyshev 矩	150
8.8 句法模式分类	123	9.10 知觉的格式塔理论	150
8.8.1 基元选择策略	124	9.11 本章小结	150
8.8.2 高维模式文法	125	参考文献	151
8.9 句法推断	125		
8.10 符号投影方法	125	第 10 章 图像处理中的模糊集理论	153
8.11 人工神经网络	126	10.1 模糊集理论概述	153
8.11.1 神经网络的进化	127	10.2 为什么图像会模糊	153
8.11.2 多层感知器	127	10.3 模糊集理论介绍	154
8.11.3 Kohonen 的自组织特征 映射	129	10.4 预备知识和背景	154
8.11.4 对向传播神经网络	130	10.4.1 模糊化	155
8.11.5 网络的全局特征	131	10.4.2 基本术语与运算	155
8.12 本章小结	131	10.5 将图像作为模糊集	156
参考文献	132	10.6 对比度增强的模糊方法	158
第 9 章 纹理与形状分析	133	10.6.1 使用模糊化增强对比度	158
9.1 概述	133	10.6.2 去噪的模糊空间过滤器	159
9.1.1 纹理基元	133	10.6.3 平滑算法	159
9.1.2 纹理分类	134	10.7 使用模糊方法的图像分割	160
9.2 灰度共生矩阵	135	10.8 像素分类的模糊方法	162
9.2.1 基元的空间关系	136	10.9 模糊的 C-均值算法	162
9.2.2 广义共生	136	10.10 神经网络模糊逻辑的融合	163
9.3 纹理光谱	137	10.11 本章小结	165
9.4 使用分形的纹理分类	138	参考文献	165
9.4.1 分形线和形状	138		
9.4.2 纹理分类中的分形	138	第 11 章 图像挖掘与基于内容的图像检索	167
9.4.3 使用地毯覆盖方法计算分形 维数	139	11.1 概述	167
9.5 形状分析	140	11.2 图像挖掘	168
9.5.1 定位点	141	11.3 检索和挖掘图像的特征	170
		11.3.1 颜色特征	170
		11.3.2 纹理特征	172
		11.3.3 形状特征	173

VIII 数字图像处理: 原理与应用

11.3.4 拓扑结构.....	175	12.11.1 乳腺超声波.....	201
11.3.5 多维索引.....	176	12.11.2 乳腺 X 光照片图像分析步骤.....	201
11.3.6 简单 CBIR 系统的结果.....	177	12.11.3 乳腺 X 光照片增强.....	202
11.4 图像检索系统中的模糊相似性 测度.....	179	12.11.4 可疑区域检测.....	202
11.5 视频挖掘.....	180	12.11.5 肿瘤分割.....	203
11.5.1 MPEG7: 多媒体内容描述 接口.....	181	12.11.6 特征选择与提取.....	205
11.5.2 基于内容的视频检索系统.....	182	12.11.7 乳腺 X 照片图像的小波分析.....	205
11.6 本章小结.....	183	12.12 本章小结.....	206
参考文献.....	183	参考文献.....	206
第 12 章 生物医学与生物医学图像处理.....	185	第 13 章 遥感多光谱场景分析.....	209
12.1 概述.....	185	13.1 概述.....	209
12.2 生物医学模式识别.....	185	13.2 卫星传感器与影像.....	209
12.2.1 特征选择.....	186	13.2.1 LANDSAT 卫星图像.....	210
12.2.2 正面面部特征提取.....	186	13.2.2 印度遥感卫星影像.....	210
12.2.3 侧面面部特征提取.....	187	13.2.3 中分辨率成像光谱仪 (MODIS).....	211
12.2.4 人脸识别.....	188	13.2.4 合成孔径雷达 (SAR).....	211
12.3 使用特征脸的人脸识别.....	188	13.3 多光谱图像的特征.....	212
12.4 签名鉴别.....	190	13.3.1 数字卫星影像的数据格式.....	212
12.5 签名模式预处理.....	190	13.3.2 失真与校正.....	212
12.6 生物医学图像分析.....	193	13.4 不同地球对象的反射光谱.....	213
12.6.1 显微图像分析.....	193	13.4.1 水区域.....	213
12.6.2 宏观图像分析.....	193	13.4.2 植被区域.....	213
12.7 生物医学成像模式.....	193	13.4.3 土壤.....	214
12.7.1 磁共振成像 (MRI).....	194	13.4.4 人造/人工目标.....	214
12.7.2 计算机轴断层摄影术.....	194	13.5 场景分类策略.....	215
12.7.3 核与超声波成像.....	195	13.5.1 使用误差反向传播的基于 神经网络的分类器.....	215
12.8 X 光成像.....	195	13.5.2 对向传播网络.....	215
12.8.1 用于肺疾病辨识的 X 光图像.....	195	13.5.3 试验与试验结果.....	216
12.8.2 胸部 X 光增强.....	196	13.5.4 分类准确性.....	216
12.8.3 用于肺结节检测的 CT-扫描.....	196	13.6 基于知识方法的光谱分类.....	217
12.8.4 用于心脏病识别的 X 光图像.....	197	13.6.1 自然/人造目标的光谱信息.....	217
12.8.5 用于先天性心脏病识别的 X 光图像.....	197	13.6.2 训练点选择与特征提取.....	217
12.8.6 用梯度算子增强胸部 X 光 照片.....	197	13.6.3 系统实现.....	218
12.8.7 骨疾病识别.....	198	13.6.4 规则生成.....	218
12.8.8 肋骨识别.....	198	13.6.5 基于规则的开发.....	219
12.9 牙齿 X 光图像分析.....	199	13.7 空间推理.....	220
12.10 牙龋齿分类.....	199	13.7.1 证据积累.....	220
12.11 乳腺 X 光照片图像分析.....	201	13.7.2 空间规则生成.....	221
		13.8 遥感的其他应用.....	222
		13.9 本章小结.....	223
		参考文献.....	223

第 14 章 动态场景分析: 运动对象的检测与跟踪	225	16.3.1 色彩空间转换	260
14.1 概述	225	16.3.2 源图像数据阵列	260
14.2 问题定义	225	16.3.3 基线压缩算法	262
14.3 自适应背景建模	225	16.3.4 DCT 系数编码	262
14.3.1 基本背景建模策略	226	16.4 本章小结	268
14.3.2 背景建模鲁棒性方法	226	参考文献	268
14.4 连通域标记	228	第 17 章 图像压缩的 JPEG2000 标准	269
14.5 阴影检测	229	17.1 概述	269
14.6 对象跟踪原理	229	17.2 为什么使用 JPEG2000	269
14.7 跟踪系统模型	230	17.3 部分 JPEG 标准	272
14.8 离散 Kalman 滤波	230	17.4 JPEG2000 概述: 第一部分编码系统	272
14.9 扩展 Kalman 滤波	232	17.5 图像处理	273
14.10 基于对象跟踪的粒子滤波	234	17.5.1 拼接	273
14.10.1 粒子属性	236	17.5.2 DC 电平偏移	273
14.10.2 粒子滤波器算法	236	17.5.3 多尺度变换	273
14.10.3 对象跟踪结果	237	17.6 压缩	274
14.11 聚缩算法	237	17.6.1 离散小波变换	275
14.12 本章小结	238	17.6.2 量化	277
参考文献	239	17.6.3 感兴趣区域编码	278
第 15 章 图像压缩概述	241	17.6.4 码率控制	280
15.1 概述	241	17.6.5 熵编码	280
15.2 信息论概念	242	17.7 第二层编码与比特流格式	281
15.2.1 离散无记忆模型与熵	242	17.8 本章小结	281
15.2.2 无噪声信源编码定理	243	参考文献	282
15.2.3 唯一判读	244	第 18 章 JPEG2000 标准中的编码算法	284
15.3 压缩算法分类	245	18.1 概述	284
15.4 信源编码算法	246	18.2 编码的数据分割	284
15.5 霍夫曼编码	247	18.3 在 JPEG2000 中的第一层编码	285
15.6 算术编码	249	18.3.1 分数位平面编码	285
15.6.1 编码算法	249	18.3.2 BPC 编码器举例	296
15.6.2 解码算法	251	18.3.3 二进制算术编码——MQ 编码器	300
15.6.3 QM 编码器	251	18.4 JPEG2000 的第二层编码	302
15.7 本章小结	254	18.4.1 比特流格式	303
参考文献	255	18.4.2 包头信息编码	305
第 16 章 JPEG 静态图像压缩标准	256	18.5 本章小结	306
16.1 概述	256	参考文献	307
16.2 JPEG 无损编码算法	257		
16.3 基线 JPEG 压缩	259		

第 1 章

概 述

1.1 图像处理基本原理

我们正处于一个视觉迷人的世界之中，主要体现在这个世界具有各种各样的形式和形状、颜色和纹理、运动和宁静。人类知觉有能力获取、集成和解译我们周围所有的、丰富的视觉信息。在我们的知觉世界里，为了解译嵌入在静止图像、图表、视频或运动图像中的视觉信息，并将这种解译能力赋予计算机是非常富有挑战性的。所以理解存储技术、处理、传输、识别和对这样的视觉场景的最终解译都是非常重要的。本书将尽可能地提供视觉信息分析技术在不同领域中的应用。

设计图像分析系统的第一步是使用光传感器或热量波长传感器来采集数字图像。由这些传感器记录的二维图像来映射三维视觉世界。对采集到的二维信号进行采样和量化来生成数字图像。

有时我们接收到由某种退化机构退化后的噪声图像。最常见的图像退化的根源是数码相机在获取视觉信息时使用的光学透镜系统。如果相机聚焦错误，那么我们将得到模糊图像。这里的模糊机构是指聚焦错误的相机。人们很可能处于有雾天气的室外。任何在有雾的冬天早晨、在室外拍摄的场景都必然是模糊图像。在这个例子中，退化是由于大气中的雾和薄雾，这种类型的退化称为大气退化。在其他的例子中，引起退化的可能是对象与相机间的相对运动。因此，如果在照相期间，被摄对象是静止的，而照相机突然被移动，拍摄的结果一定是模糊和噪点很多的图像。在上述的例子中，我们需要有适当的技术来改善图像，以便最终图像具有更好的视觉质量，从变型和噪点中解脱出来。在这个领域的早期，图像增强、滤波和复原就已经是一些很重要的图像处理应用了[1]~[4]。

分割是将图像细分成许多统一、均匀区域的过程。每个均匀区域都是整个场景的要素部分或对象。换句话说，图像的分割是定义相连且不相互重叠区域的集合，以便在图像段的每个像素获取表明区域的一个独特的、属于哪个区域的标记。在自动化图像分析中，分割是最重要的元素之一，主要是因为在这一步将从图像中提取感兴趣的对象或其他实体，用于随后的处理，如描述和识别。例如，航空图像包含海洋和陆地，最初，分割的问题是将图像分为两部分——陆地段和水体（或海洋段）。其后，需要适当地分割图像中陆地部分对象，并进行相应的分类。

在提取每段之后，接下来的任务是提取重要的特征值集，如纹理、颜色和形状。这些

是重要的可测量的实体，给测量图像段提供了各种各样的特征。纹理的一些特征是粗糙度、平滑性、规律性等，而通常形状描述都包括长度、广度、长宽比、面积、位置、周长、压缩性等。在场景中的各个分割区域都可以通过一系列这样的特征来描述。

最后是基于这些提取特征的集合，每个分割对象都被归类为重要类集合中的一类。在海洋数字图像中，这些类也许是大船或小船，甚至是海军舰艇和水体等大类。在计算机视觉中，场景分割和对象分类问题是两个集成研究的领域。目前已经知道，专家系统、语义网络和神经网络系统对执行这样高水平视觉任务是相当高效的。

图像处理的另一个方面涉及压缩和视觉信息编码。随着各种各样图像应用需求的增长，数字影像的存储需求也在爆发式地增长。在目前，图像数据和它们存储压缩共同证明，传输所需通信带宽是关键的、发展活跃的领域。更有趣的是，在图像的标准表示中，图像数据一般包含大量的冗余信息。图像压缩技术有助于减少在原始图像中数据的冗余，减少存储所需要的容量及通信带宽。

1.2 图像处理应用

大多数图像处理应用是在人类不同的光谱活动中——从遥感场景解译到生物医学图像解译。在本节里，我们将只提供一些这类应用的概述。

1.2.1 自动视觉检测系统

在制造业和相关产业中，自动视觉检测系统是提高生产力和改进产品质量的根本要素[5]。在这里我们对视觉检测系统做一些简要的介绍。

- **白炽灯灯丝的自动检测 (automatic inspection of incandescent lamp filaments)**: 在自动视觉检测中很有趣的应用包括电灯泡制造过程的检测。由于灯丝的错误几何结构，例如在灯中灯丝定位不一致等，在短时间后，电灯泡的灯丝通常会熔化。人工检测不能有效的检测这样的象差。
在自动的基于视觉的检测系统中，将产生灯丝的一个二进制图像切片，从而生成灯丝廓影。这个廓影是用来分析和识别在电灯泡里灯丝几何结构定位是否不一致。这样的系统由通用电器公司 (General Electric Corporation) 设计和安装。
- **故障元件识别 (faulty component identification)**: 自动视觉检测也可以用于在电子或机电系统中识别故障元件。故障元件通常会产生更多的热能。在装配中，从热能的分布产生红外 (infra-red, IR) 图像。通过分析这些 IR 图像，我们就在装配中识别出故障元件。
- **自动表面检测系统 (automatic surface inspection systems)**: 在许多金属工业中，检测表面的缺陷是很重要的。例如，钢厂的热轧板机或冷轧板机必须查出在被轧金属表面上的任何象差。这可以使用图像处理技术 (如边缘检测、纹理识别、分形分析等) 来完成。

1.2.2 遥感场景解译

有关自然资源的信息，例如农业、水文学、矿物、森林、地质资源等，可以通过遥感图像分析提取。为了进行遥感场景分析，地球的表面图像由在遥感卫星中的传感器或者安置在航空器中的多光谱扫描器采集，然后传送到地面工作站做进一步处理[6,7]。图 1.1 中给出了两幅遥感图像。图 1.1(a)是印度的恒河三角洲。浅蓝段代表恒河三角洲区域的沉积物，深蓝色段代表水体，深红区域是毗邻海岛的美洲红树林沼泽。图 1.1(b) 是喜马拉雅山脉的冰川流。白色区域代表低速的停滞冰川。



(a) 恒河三角洲



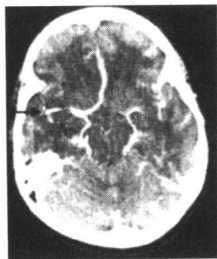
(b) 喜马拉雅山脉的冰川流

图 1.1 遥感图像的例子

在卫星图像中对区域和对象的解译技术可以用于市政规划、资源募集、防洪、农业生产监控等。

1.2.3 生物医学成像技术

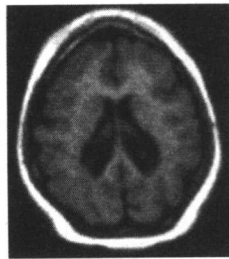
各种各样成像设备，如 X 光、计算机辅助断层摄影 (computer aided tomographic, CT)、超声波等广泛地用于医疗诊断中[8~10]。生物医学图像采集的例子可以使用不同的成像模式，例如 CT 扫描、X 光和 MRI，如图 1.2 所示。



(a) 大脑的 CT 扫描图像



(b) 手腕的 X 光图像



(c) 大脑的 MRI 图像

图 1.2 图像实例

- 考虑局部对象，即不同的器官。
- 采取提取对象（即在图像中的肿瘤）的措施。
- 用于诊断的对象解译。

下面介绍一些生物医学成像的应用。

- 肺病识别：在胸部X光片中，黑色部分代表空气，而固体组织看上去更亮一些。骨头比软组织具有更高的辐射不透明度。在正常胸部X光片中，可清晰地看见肋骨、胸部脊椎、胸腔和腹腔的横隔膜。通过分析相应部分，这些在胸腔X光片中的区域可以用于病变的诊断。
- 心脏病识别：定量测量（例如心脏大小和形状）是对心脏病分类的重要诊断特征。图像分析技术可以被用于X光图像用来改善心脏病的诊断。
- 数字乳腺X光片：为了诊断乳腺肿瘤，数字乳腺X光片对查出特征（例如微钙化）是非常有用的。图像处理技术，例如对比度增强、分割、特征提取、形状分析等，可以用于乳腺X光片的分析。肿瘤形状的规律性被用来确定肿瘤为良性还是恶性。

1.2.4 防御观测

防御观测应用是图像处理技术中一个重要的研究领域。为了满足对土地和海洋的连续观测的需求，我们可以使用空中观测技术。

假设我们要在海洋表面的上空图像中找出海军舰艇类型和形态。在这时的主要任务是将图像中的水体部分分割成不同的对象。在提取段之后，使用参数（像面积、位置、周长、压缩度、形状、长度、广度和长宽比等），对每个段中的对象进行分类。

这些对象的范围从小船到巨型军舰。使用以上的特征可以识别和定位这些对象。为了描述船的所有可能的形态，我们需要在8个可能的方向上识别这些对象，这8个方向即为北部、南部、东部、西部、东北部、西北部、东南部和西南部。从这些对象的空间分布就可以解译整个海洋场景，这对海洋观测也是极为重要的。

1.2.5 基于内容的图像检索

在图像处理中，从大量图像存档中查询图像是一种重要的应用。大量多媒体采集机构和数字图书馆的出现导致了对搜索工具开发的需求，使用这些搜索工具，人们可以从大量多媒体采集机构和数字图书馆中检索信息。目前可用的一些优秀的搜索引擎是以机读的形式检索文本，而尚没有用于检索亮度和彩色图像的快速搜索工具。搜索和索引图像的传统方法速度慢而且昂贵。因而目前急需开发基于内容的图像检索算法。

从大型图像数据库中检索图像信息，数字图像的特征（例如形状、纹理、颜色、对象的拓扑结构等）可以作为关键索引值。基于这种图像内容的检索是很广泛的，它也被称为基于内容的图像检索[11,12]。

1.2.6 运动对象跟踪

为了测量运动参数和获得运动对象的视觉记录，跟踪运动对象是图像处理中的一个重要应用领域[13,14]。总的来说，用于对象跟踪的有两种不同方法：

- 基于识别的跟踪。
- 基于运动的跟踪。

为跟踪快速对象（即，军用飞机、导弹等）而开发的系统是基于运动预测技术的，例如 Kalman 滤波、扩展的 Kalman 滤波、粒子滤波等。在基于对象跟踪系统的自动图像处理中，如果没有人工干预，将自动地获取进入传感器视野的目标对象。在基于识别的跟踪中，是在连续图像帧之中识别对象模式，并且使用位置信息实现跟踪。

1.2.7 图像与视频压缩

图像和视频压缩是图像处理中一个活跃的应用领域[12,15]。图像和视频的压缩技术的发展继续多媒体通信和应用中充当重要角色。虽然在近 20 年间，存储费用显著降低，但图像和视频数据存储的需求也成指数增长。以 $70\mu\text{m}$ 扫描的 $36\text{cm}\times 44\text{cm}$ 的数字 X 光片需要大约 45MB 的存储空间。同样，以每秒 60 传输帧传输的分辨率为 1280×720 的高清晰电视的存储空间需要每秒超过 1250MB。不经过任何压缩而在目前的信道中实时传输这些视频图像实在是很困难的。有趣的是，不论是静止图像还是视频图像，在它们的标准表示中都有大量视觉上的冗余信息。这些冗余存在于正常图像平滑均匀区域的相邻像素之间，相邻像素间的值差异很小，人类的视觉很难觉察到其值的变化。同样，在慢速移动的视频序列中的连续帧中，也有冗余嵌入在这些连续帧中。图像和视频压缩技术主要是减少数据表示中这样的视觉冗余，使用少数的重要的数据位来表示图像帧，这样就减少了对存储空间和有效通信带宽的需求。

1.3 人类视觉知觉

通过眼睛和附属在视网膜的敏感细胞，光带产生的电磁辐射从我们的视觉环境进入视觉系统。活动从视网膜开始，来自于相邻接收器的信号的比较，编码信息通过视觉神经传输到在我们耳后的皮层。在[16]中我们可以知道人类视觉知觉是很优秀的。在参考文献[17,18]中已经对我们视觉系统的空间特征建立了非线性模型。

虽然眼睛可以查看静态图像，但本质上它们是运行物体探测器。眼睛有识别静态对象的能力，并且它能建立静态场景中各种对象和区域间的空间关系。它们的基本功能取决于刺激比较相邻的细胞，它将导致对运动的解译。当观察一个静态场景时，眼睛进行的小反复运动称为眼跳，移动边缘通过感受器。然而，我们的视觉感知、识别和解译是发生在我们的大脑里。在场景中的对象和不同区域在我们的大脑中重新组织，从边缘或边界处来封装在场景中的对象和不同区域。在这些边缘或边界处嵌入着对象的大量信息。识别的过程是我们神经系统组织学习的结果。线的方位和移动方向也可以用于对象识别过程。

1.3.1 人眼

人眼的结构如图 1.3 所示。视觉信号从眼睛的视网膜传入到大脑需要通过大约 150 万个视觉神经的神经元。视网膜包含大量的光感受器，这些光感受器紧密地以较规则或较不规则六角阵列排列。视网膜阵列包含 3 种类型颜色传感器称为视锥细胞，视网膜中央部分称做中央凹。视锥细胞以以下方式分布：它们密集地分布在视网膜附近的中央部分并且在中央凹附近的周边部分密度减少。在这里有 3 种不同类型的视锥细胞，即红色、绿色和蓝色视锥细胞，它们用于颜色视觉。这 3 类明显不同的视锥细胞包含不同的光敏色素。3 种色素都有最大吸收能力，分为大约为 430nm（紫罗兰色）、530nm（青绿色）和 560nm（黄绿色）。

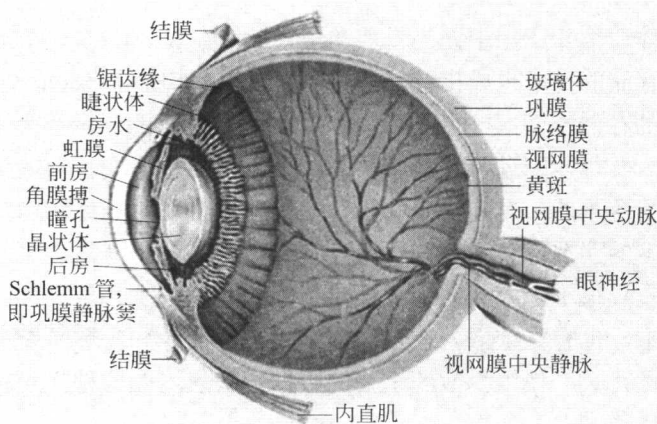


图 1.3 人眼的结构

另一个类型的小传感器填充在视锥细胞之间。这些传感器叫做视杆细胞，它们用于灰色视觉。这些感受器在数量上比视锥细胞要多。

视杆细胞对较暗的亮度非常敏感，它负责实现我们在昏暗的光线（暗视觉）中看见物体的能力。另一方面，视锥细胞或明视觉系统负责在强光亮级识别物体，这时可使用的光子很多，并且可以降低敏感度为代价达到最大分辨率。

1.3.2 视觉的神经机制

视觉系统中的视觉神经进入眼珠并与位于眼睛后面的视杆细胞与视锥细胞连接。

神经元包含神经树（输入），在末端有长的树枝状神经轴突（输出）。神经元通信通过突触。信号传输同横跨表面的化学药品扩散联系在一起，对于化学药品，接收神经元要么是兴奋性突触，要么是抑制性突触，然后在表面扩散。视觉神经在视网膜一侧的神经节细胞开始作为轴突。另一方面，视杆细胞和视锥细胞通过双极细胞连接到神经节细胞，并且还有水平神经细胞建立侧向联系。

来自于视网膜内相邻感受器的信号被水平细胞分组，并在中心和边缘形成反向反应的接收区域，以便区域均匀的照明没有引起净刺激。在不均匀的照明情况下，在中心和边缘

产生的刺激是有差异的。一些接收区域使用色差,例如红色-绿色或黄色-蓝色,因此对色差刺激应用也适用于亮度。在外侧膝状体对接收区域反应进一步分组,并且将视觉外皮层用作定向边缘检测和眼睛偏好检测。这是低级处理,我们还不是很清楚在这之前的高级解译所使用的机制。然而,它说明了在知觉中差异所扮演的重要角色,它的本质在于对比现象。如果视网膜被照明,在亮度和色彩上都很均匀的情况下,发生的神经活动将很少。

正常人的视网膜有 600 万~700 万个视锥,1100 万~1300 万个视杆。视杆、视锥的光信号是通过视觉神经的纤维进行传输的。视觉神经经过视交叉,所有来自于两个右侧视网膜的信号将被传送给右半脑,所有来自于左侧视网膜的信号,都将被传送到左半脑。在两个半脑中分别形成半个图像。这样能确保即使眼睛出现问题,还能保留视觉系统能力。视神经末端在外侧膝状体中,它中途通过大脑,并且信号将从这里扩散到视觉外皮层。视觉外皮层仍然有视网膜的拓扑结构,并且这仅仅是知觉的第一阶段,在这个阶段信息变得可用。两个大脑半球视觉区域在胼胝体中连接,胼胝体使两半的视觉区域联接成一个整体。

1.4 图像处理系统的组成

图像处理系统由多个组成部分。图像处理系统的第一个主要组成部分是采集三维对象图像的照相机。

数码相机

在大多照相机中使用的传感器是电耦合器件(charge coupled device, CCD)或 CMOS 传感器。CCD 照相机包含大量非常小的光侦测器,这种小的光侦测器叫做感光单元。在进行适当的模拟到数字转换之后,累积在图像中各单元上的电荷将被传输和记录。

在 CMOS 传感器里,将使用一定数量的晶体管来放大在每个像素点位置的信号。每个像素位置的合成信号都能被单独读出来。因为使用几个晶体管使光敏感度变得更低。这是由于一些光子在与色彩传感器邻接的晶体管上入射(用于信号放大)。与 CCD 传感器相比,当前达到最新技术发展水平的 CMOS 传感器具有更多噪声。但它们耗电量低,相对便宜。

在明亮阳光下,摄像机镜头的光圈不需要很大,因为此时我们不需要太多的光,当在多云的天气里,我们需要更多光来生成图像时,光圈就应该被放大。这与我们眼睛的作用是相同的。快门速度给出了光线通过光圈所用的时间。打开快门和关闭快门的时间取决于对光需求。数码相机的焦距是透镜焦点和传感器阵列表面之间的距离。在选择需要从照相机获得必需的放大时,焦距是非常重要的信息。

在隔行扫描摄像机中,每个图像帧都分成两个场,每个场包含偶数(前场)或奇数(后场)水平视频线。这两个场由视频显示器设备组合。在隔行扫描照相机中前场和后场的组合方式如图 1.4 所示。另一方面,在改进的扫描照相机中,整个帧输出将作为一个唯一帧。当对移动的场景成像时,例如在机器人视觉中,它使用闪光灯脉冲来照亮场景中的对象并采集它。在这样的成像应用中,改进的扫描照相机是更可取的。

隔行扫描照相机不会用于这样的应用,因为如果使用隔行扫描,照明时间比帧时间短,