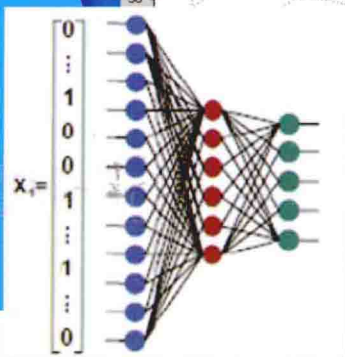
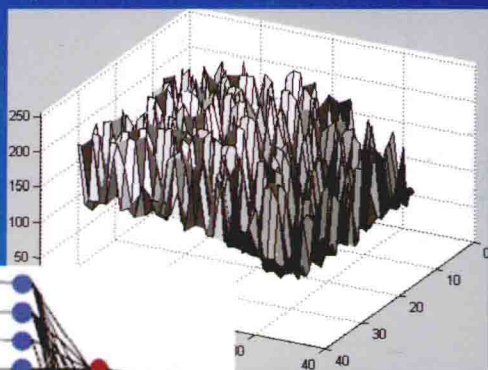


数字图像处理 与机器视觉

— Visual C++与Matlab实现
(第2版)



■ 张铮 徐超 任淑霞 韩海玲 编著

■ 为想要了解数字图像处理与机器视觉领域的读者提供了一条扎实的进阶之路，从基本的图像处理算法，到投影校正、顶帽变换等高级应用，再到特征提取、分类器设计，以及光学字符识别、人脸识别等综合案例，一步步地引导读者从阅读中获得知识，于实践中提升技能。

■ Visual C++与Matlab两种语言开发无缝连接，体现出科学研究和工程实践在图像处理与机器视觉领域的完美结合。

- ✓ 各种标准图像处理算法 (Matlab/Visual C++)
- ✓ 汽车牌照的投影失真校正 (Matlab/Visual C++)
- ✓ 顶帽变换 (Matlab/Visual C++)
- ✓ 主成份分析 (Matlab)
- ✓ LBP特征提取 (Matlab)
- ✓ 基于人工神经网络的数字识别系统 (Visual C++)
- ✓ 基于支持向量机的人脸识别系统 (Matlab)



源程序

人民邮电出版社
POSTS & TELECOM PRESS

数字图像处理 与机器视觉

— Visual C++与Matlab实现
(第2版)

张铮 徐超 任淑霞 韩海玲 编著

人民邮电出版社

北京

图书在版编目 (C I P) 数据

数字图像处理与机器视觉 : Visual C++与Matlab实现 / 张铮等编著. -- 2版. -- 北京 : 人民邮电出版社, 2014.5

ISBN 978-7-115-34668-1

I. ①数… II. ①张… III. ①C语言—程序设计②数字图像处理③计算机视觉 IV. ①TP312②TN911.73③TP302.7

中国版本图书馆CIP数据核字(2014)第038575号

内 容 提 要

本书将理论知识、科学研究和工程实践有机结合起来,内容涉及数字图像处理和识别技术的方方面面,包括图像的点运算、几何变换、空域和频域滤波、小波变换、图像复原、彩色图像处理、形态学处理、图像分割、图像压缩以及图像特征提取等;同时对机器视觉进行了前瞻性的探究,重点介绍了3种目前在工程技术领域非常流行的分类技术——人工神经网络(ANN)、支持向量机(SVM)和AdaBoost,并在配套给出的识别案例中直击光学字符识别(OCR)、人脸识别和性别分类等热点问题。

全书结构紧凑,内容深入浅出,讲解图文并茂,适合于计算机、通信和自动化等相关专业的本科生、研究生,以及工作在图像处理和识别领域一线广大工程技术人员阅读参考。

-
- ◆ 编 著 张 铮 徐 超 任淑霞 韩海玲
责任编辑 张 涛
责任印制 彭志环 杨林杰
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京中新伟业印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 37.25
字数: 1 058 千字 2014年5月第2版
印数: 12 301-16 300册 2014年5月北京第1次印刷
-

定价: 79.00元(附光盘)

读者服务热线: (010)81055410 印装质量热线: (010)81055316
反盗版热线: (010)81055315

前 言

图像处理与机器视觉是当今计算机科学中的一个热门研究方向，应用广泛，发展前景乐观。近年来，伴随着人工智能、模式识别学科以及人机智能接口技术的飞速发展，机器视觉的研究正在不断升温——从日常生活中与人类息息相关的光学字符识别（OCR）和汽车自动驾驶，到医学应用中的病灶检测与分析，再到未来人机智能交互领域中的人脸识别、情感计算等，图像处理作为机器视觉研究中必不可少的图像预处理环节是读者需要掌握的首要技术。

和图像相关的东西往往容易引起计算机初学者的兴趣，笔者在读本科的时候就觉得能让计算机理解所“看”到的东西是一件非常神秘和令人兴奋的事情；但同时它的理论性较强，门槛较高，在各个高校中，这门课程大多也是作为计算机专业研究生的选修课程。要理解该领域的知识，读者需要具有一定的数学基础，除此之外还涉及信号处理、统计分析、模式识别和机器学习等专业领域知识，因此令很多人望而却步。

其实“难以理解”的关键在于缺乏必要的先序知识，造成了读者在相关知识上难以跨越的鸿沟。在撰写本书过程中，对于可能造成读者理解困难的地方，均尽可能地给出了必要的基本知识，深入浅出，尽量定性地去描述，对于那些并不一目了然的结论均给出了思路 and 解释，必要的还提供了证明，对于某些非常专业已经超过本书讨论范围的相关知识在最后给出了参考文献，供有兴趣的读者进一步学习和研究。

本书的宗旨是在向读者介绍知识的同时，培养读者的思维方法，使读者知其然还要知其所以然，并在解决实际问题中能有自己的想法。

1. 内容安排

本书在第一版的基础上，根据技术的发展和应用的需求进行改版，主要内容介绍如下。

第0~2章介绍了数字图像处理的基础知识和编程基础，使读者第一步能够建立起对于数字图像本质的正确认识，了解和掌握必要的术语和预备知识，并且熟悉本书自始至终需要使用的两大工具 MATLAB 和 Visual C++，给出了功能强大的 C++ 数字图像处理基类 CImg 及其派生类 CImgProcess 的框架。

第3~4章分别介绍了图像的灰度变换和几何变换。通过灰度变换可以有效改善图像的外观，并在一定程度上实现图像的灰度归一化；几何变换则主要应用在图像的几何归一化和图像校准当中。总体而言，这些内容大多作为图像的前期预处理工作的一部分，是图像处理中相对固定和程式化的内容。

第5~6章分别从空间域和频率域两个角度去考量图像增强的各个主要方面。图像增强作为数字图像处理中相对简单却最具艺术性的领域之一，可理解为根据特定的需要突出一幅图像中的某些信息，同时，削弱或去除某些不需要的信息的处理方法。其主要目的是使处理后的图像对某种特定的应用来说，比原始图像更适用。

第7章小波变换继第6章之后继续在频率域中研究图像。傅里叶变换一直是频率域图像处理的基石，它能用正余弦函数之和表示任何分析函数，而小波变换则基于一些有限宽度的基小波，这些小波不仅在频率上是变化的，而且具有有限的持续时间。比如对于一个乐谱，小波变换不仅能提供要演奏的音符，而且说明了何时演奏等细节信息，但是傅里叶变换只提供了音符，局部信息在变换中丢失。

第 8 章图像复原与图像增强相似，其目的也是改善图像质量。但是图像复原是试图利用退化过程的先验知识使已被退化的图像恢复本来面目，而图像增强是用某种试探的方式改善图像质量，以适应人眼的视觉与心理。引起图像退化的因素包括由光学系统、运动等造成的图像模糊，以及源自电路和光学因素的噪声等。图像复原是基于图像退化的数学模型，复原的方法也建立在比较严格的数学推导上。

第 9 章是本书中相对独立的一章，以介绍色彩模型之间的相互转换以及彩色图像处理方面的基本概念和基本方法为主。随着基于互联网的图像处理应用在不断增长，彩色图像处理已经成为一个重要领域。

第 10 章图像压缩旨在减少表示数字图像时需要的数据量。主要介绍了图像压缩的基本原理、DCT 变换、预测编码、霍夫曼编解码，以及算术编码和游程编码等知识，并给出了有关 JPEG 和 JPEG2000 压缩标准的内容。

第 11~13 章（形态学图像处理、图像分割、特征提取）是从单纯图像处理向图像识别（机器视觉）过渡，这一阶段的特点是输入是图像，输出则是在识别意义上读者感兴趣的图像元素。形态学处理是提取图像元素的有力技术，它在表现和描述形状方面非常有用；分割过程则将一幅图像划分为组成部分或目标对象；研究特征提取则是要将前面提取出来的图像元素或目标对象表示为适合计算机后续处理的数值形式，最终形成能够直接供分类器使用的特征。

第 14 章在前面知识的基础之上，引出了机器视觉的前导性内容，给出了解决识别问题的一般思路。

最后 3 章（人工神经网络、支持向量机和 AdaBoost）介绍了 3 种十分强大的分类技术，并在手写数字字符识别、人脸识别和性别分类这样的经典案例中进行讲解。

2. 读者对象

- 大学二年级以上（具备必要的数学基础）的相关专业的本科生、研究生。
- 工作在图像处理和识别领域一线的工程技术人员。
- 对于数字图像处理和机器视觉感兴趣的并且具备必要预备知识的所有读者。

3. 在阅读本书之前，读者最好具有如下的预备知识

- 读者应该熟悉 C++ 编程语言和面向对象的基本思想。书中的相当一部分示例是以 C++ 语言描述的。
- 读者应具备一定的数学基础，主要的高等数学知识、少量的线性代数基本概念加上对于概率理论主要思想的理解（识别部分）。

4. 附书光盘和读者反馈

本书所有 Visual C++ 实例的源代码和大部分 Matlab 实例的源代码均可在随书附赠的光盘中找到。为了尽可能满足所有读者的需要，同时提供了 C++ 代码的 Visual C++ 6.0 版本和 Visual Studio 2010/2012 版本。虽然本书中的所有例子都已经在 Windows XP、Windows 2003 和 Windows 7 下测试通过，但由于许多算法比较复杂且笔者水平局限，也有存在缺陷的可能，即便正确也很可能存在更加优化的算法或更加合理的程序结构，如发现任何上述问题，欢迎读者通过以下方式联系本书作者，以便做出改进。

邮箱：book95_editor@qq.com

腾讯微博：@Book95_Aaron

5. 致谢

首先要感谢我的授业恩师——南开大学的白刚教授和天津大学的赵政教授，是他们引导我进入了图像处理与机器视觉的研究领域。同时，他们在我写作过程中的指点和教诲确保了本书的权威性和严谨性。

感谢我的好友王焯阳提供并调试了许多实例代码；感谢徐超、陈明、张阳、李广鹏、马惠来、马宏、卞长迪、郑琦、王命达、杜强、陈香凝、苑春苗、宁升、李鸿鹏、李明剑等参与了本书部分章节的编写和修改；感谢罗小科先生为本书制作了很多插图；感谢我的兄长张钊为本书提供了部分照片；还要感谢朱毅、丁宗尧、刘群忠、贾万宏等为本书的编写提出了很多的宝贵意见和建议。

最后感谢我的妻子马宏、儿子张垚淼及我的父母家人，没有你们的鼓励和支持就不会有我的这部作品。

编 著

目 录

第 0 章 初识数字图像处理与机器视觉.....1	1.4 读取和写入图像文件..... 32
0.1 数字图像.....1	1.5 图像的显示..... 34
0.1.1 什么是数字图像.....1	第 2 章 Visual C++图像处理编程基础..... 37
0.1.2 数字图像的显示.....1	2.1 位图文件及其 C++操作..... 37
0.1.3 数字图像的分类.....2	2.1.1 设备无关位图..... 37
0.1.4 数字图像的实质.....3	2.1.2 BMP 图像文件数据结构..... 37
0.1.5 数字图像的表示.....4	2.2 认识 CImg 类..... 40
0.1.6 图像的空间和灰度级分 辨率.....4	2.2.1 主要成员函数列表..... 40
0.2 数字图像处理与机器视觉.....5	2.2.2 公有成员..... 41
0.2.1 从图像处理到图像识别.....5	2.3 CImg 类基础操作..... 41
0.2.2 什么是机器视觉.....6	2.3.1 加载和写入图像..... 41
0.2.3 数字图像处理和识别的应用 实例.....7	2.3.2 获得图像基本信息..... 44
0.3 数字图像处理的预备知识.....8	2.3.3 检验有效性..... 45
0.3.1 邻接性、连通性、区域和 边界.....8	2.3.4 按像素操作..... 45
0.3.2 距离度量的几种方法.....9	2.3.5 改变图像大小..... 47
0.3.3 基本的图像操作.....10	2.3.6 重载的运算符..... 47
第 1 章 MATLAB 数字图像处理编程 基础.....11	2.3.7 在屏幕上绘制位图图像..... 48
1.1 MATLAB R2011a 简介.....11	2.3.8 新建图像..... 48
1.1.1 MATLAB 软件环境.....11	2.3.9 图像类型的判断与转化..... 50
1.1.2 文件操作.....12	2.4 DIPDemo 工程..... 51
1.1.3 在线帮助的使用.....13	2.4.1 DIPDemo 主界面..... 51
1.1.4 变量的使用.....15	2.4.2 图像操作和处理类—— CImg 和 CImgProcess..... 52
1.1.5 矩阵的使用.....17	2.4.3 文档类——CDIPDemoDoc..... 53
1.1.6 细胞数组 (Cell Array) 和 结构体 (Structure).....19	2.4.4 视图类——CDIPDemoView..... 53
1.1.7 关系运算与逻辑运算.....20	2.5 CImg 应用示例..... 54
1.1.8 常用图像处理数学函数.....21	2.5.1 打开图像..... 54
1.1.9 MATLAB 程序流程控制.....22	2.5.2 清空图像..... 55
1.1.10 M 文件编写.....25	2.5.3 像素初始化方法..... 56
1.1.11 MATLAB 函数编写.....26	2.5.4 保存图像..... 57
1.2 MATLAB 图像类型及其存储方式.....28	第 3 章 图像的点运算..... 58
1.3 MATLAB 的图像转换.....30	3.1 灰度直方图..... 58
	3.1.1 理论基础..... 58
	3.1.2 MATLAB 实现..... 59
	3.1.3 Visual C++实现..... 62
	3.2 灰度的线性变换..... 63

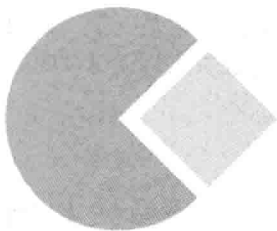
3.2.1	理论基础	63	4.6.1	以原点为中心的图像 旋转	103
3.2.2	MATLAB 程序的实现	64	4.6.2	以任意点为中心的图像 旋转	104
3.2.3	Visual C++实现	66	4.6.3	图像旋转的实现	105
3.3	灰度对数变换	67	4.7	插值算法	106
3.3.1	理论基础	67	4.7.1	最近邻插值	106
3.3.2	MATLAB 实现	68	4.7.2	双线性插值	107
3.3.3	Visual C++实现	69	4.7.3	高阶插值	109
3.4	伽玛变换	70	4.8	图像配准简介	111
3.4.1	理论基础	70	4.8.1	图像配准	112
3.4.2	MATLAB 编程实现	70	4.8.2	人脸图像配准的 MATLAB 实现	112
3.4.3	Visual C++实现	72	4.9	Visual C++高级应用实例—— 汽车牌照的投影失真校正	115
3.5	灰度阈值变换	73	4.9.1	系统分析与设计	116
3.5.1	理论基础	73	4.9.2	系统实现	117
3.5.2	MATLAB 编程实现	74	4.9.3	功能测试	122
3.5.3	Visual C++实现	75	第 5 章	空间域图像增强	126
3.6	分段线性变换	76	5.1	图像增强基础	126
3.6.1	理论基础	76	5.2	空间域滤波	127
3.6.2	MATLAB 编程实现	77	5.3	图像平滑	133
3.6.3	Visual C++编程实现	81	5.3.1	平均模板及其实现	133
3.7	直方图均衡化	82	5.3.2	高斯平滑及其实现	134
3.7.1	理论基础	82	5.3.3	通用平滑滤波的 Visual C++ 实现	138
3.7.2	MATLAB 编程实现	83	5.3.4	自适应平滑滤波	139
3.7.3	Visual C++实现	85	5.4	中值滤波	140
3.8	直方图规定化(匹配)	86	5.4.1	性能比较	140
3.8.1	理论基础	86	5.4.2	一种改进的中值滤波策略	144
3.8.2	MATLAB 编程实现	87	5.4.3	中值滤波的工作原理	145
3.8.3	Visual C++实现	89	5.5	图像锐化	145
第 4 章	图像的几何变换	92	5.5.1	理论基础	145
4.1	解决几何变换的一般思路	92	5.5.2	基于一阶导数的图像 增强——梯度算子	145
4.2	图像平移	94	5.5.3	基于二阶微分的图像 增强——拉普拉斯算子	149
4.2.1	图像平移的变换公式	94	5.5.4	基于一阶与二阶导数的 锐化算子的比较	151
4.2.2	图像平移的实现	94	5.5.5	高提升滤波及其实现	152
4.3	图像镜像	96	5.5.6	高斯-拉普拉斯变换(Laplacian of a Gaussian, LoG)	156
4.3.1	图像镜像的变换公式	96			
4.3.2	图像镜像的实现	97			
4.4	图像转置	99			
4.4.1	图像转置的变换公式	99			
4.4.2	图像转置的实现	99			
4.5	图像缩放	101			
4.5.1	图像缩放的变换公式	101			
4.5.2	图像缩放的实现	101			
4.6	图像旋转	103			

第 6 章 频率域图像增强	159	内在联系	204
6.1 频率域滤波——与空间域滤波殊途同归	159	附录	205
6.2 傅里叶变换基础知识	159	第 7 章 小波变换	207
6.2.1 傅里叶级数	159	7.1 多分辨率分析	207
6.2.2 傅里叶变换	161	7.1.1 多分辨率框架	207
6.2.3 幅度谱、相位谱和功率谱	163	7.1.2 分解与重构的实现	213
6.2.4 傅里叶变换的实质——基的转换	165	7.1.3 图像处理中分解与重构的实现	214
6.3 快速傅里叶变换及实现	166	7.2 Gabor 多分辨率分析	220
6.3.1 FFT 变换的必要性	167	7.3 常见小波分析	223
6.3.2 常见的 FFT 算法	167	7.3.1 Haar 小波	223
6.3.3 按时间抽取的基-2 FFT 算法	168	7.3.2 Daubechies 小波	225
6.3.4 离散反傅里叶变换的快速算法	171	7.4 高维小波	227
6.3.5 N 维快速傅里叶变换	171	第 8 章 图像复原	230
6.3.6 MATLAB 实现	171	8.1 图像复原的理论模型	230
6.3.7 Visual C++ 实现	175	8.1.1 图像复原的基本概念	230
6.4 频域滤波基础	183	8.1.2 图像复原的一般模型	232
6.4.1 频域滤波与空域滤波的关系	183	8.2 噪声模型	232
6.4.2 频域滤波的基本步骤	184	8.2.1 噪声种类	233
6.4.3 频域滤波的 MATLAB 实现	184	8.2.2 MATLAB 实现	237
6.4.4 频域滤波的 Visual C++ 实现	185	8.2.3 Visual C++ 实现	239
6.5 频率域低通滤波器	187	8.3 空间滤波	244
6.5.1 理想低通滤波器及其实现	187	8.3.1 空域滤波原理	244
6.5.2 高斯低通滤波器及其实现	191	8.3.2 MATLAB 实现	245
6.6 频率域高通滤波器	195	8.3.3 Visual C++ 实现	247
6.6.1 高斯高通滤波器及其实现	195	8.4 逆滤波复原	250
6.6.2 频域拉普拉斯滤波器及其实现	198	8.4.1 逆滤波原理	250
6.7 MATLAB 综合案例——利用频域滤波消除周期噪声	201	8.4.2 MATLAB 实现	251
6.7.1 频域带阻滤波器	201	8.4.3 Visual C++ 实现	253
6.7.2 带阻滤波器消除周期噪声	202	8.5 维纳滤波复原	256
6.8 频域滤波器与空域滤波器之间的		8.5.1 维纳滤波原理	256
		8.5.2 MATLAB 实现	257
		8.5.3 Visual C++ 实现	260
		8.6 有约束最小二乘复原	262
		8.7 Lucky-Richardson 复原	265
		8.8 盲去卷积图像复原	266
		8.9 MATLAB 图像复原综合案例——去除照片的运动模糊	268
		第 9 章 彩色图像处理	270
		9.1 彩色基础	270
		9.2 彩色模型	272

9.2.1	RGB 模型	272	11.2.4	闭运算及其实现	356
9.2.2	CMY、CMYK 模型	274	11.3	二值图像中的形态学应用	357
9.2.3	HSI 模型	276	11.3.1	击中与未击中变换及其实现	357
9.2.4	HSV 模型	282	11.3.2	边界提取与跟踪及其实现	359
9.2.5	YUV 模型	287	11.3.3	区域填充及其 Visual C++ 实现	363
9.2.6	YIQ 模型	292	11.3.4	连通分量提取及其实现	365
9.2.7	Lab 模型简介	296	11.3.5	细化算法及其 Visual C++ 实现	370
9.3	全彩色图像处理基础	296	11.3.6	像素化算法及其 Visual C++ 实现	374
9.3.1	彩色补偿及其 MATLAB 实现	296	11.3.7	凸壳及其 Visual C++ 实现	379
9.3.2	彩色平衡及其 MATLAB 实现	298	11.3.8	bwmorph() 函数	382
第 10 章	图像压缩	300	11.4	灰度图像中的基本形态学运算	383
10.1	图像压缩理论	300	11.4.1	灰度膨胀及其实现	383
10.1.1	图像冗余	300	11.4.2	灰度腐蚀及其实现	386
10.1.2	香农定理	303	11.4.3	灰度开、闭运算及其实现	389
10.1.3	保真度评价	304	11.4.4	顶帽变换 (top-hat) 及其实现	392
10.2	DCT 变换与量化	304	小结		394
10.2.1	DCT 变换原理	304	第 12 章	图像分割	395
10.2.2	量化	306	12.1	图像分割概述	395
10.2.3	DCT 变换和量化的 Visual C++ 实现	307	12.2	边缘检测	396
10.3	预测编码	312	12.2.1	边缘检测概述	396
10.4	霍夫曼编码	313	12.2.2	常用的边缘检测算子	397
10.4.1	霍夫曼编码原理	313	12.2.3	MATLAB 实现	400
10.4.2	霍夫曼编码的 Visual C++ 实现	316	12.2.4	Visual C++ 实现	402
10.5	算术编码	324	12.3	霍夫变换	409
10.5.1	算术编码原理	324	12.3.1	直线检测	409
10.5.2	算术编码的 Visual C++ 实现	327	12.3.2	曲线检测	411
10.6	游程编码	330	12.3.3	任意形状的检测	411
10.7	JPEG 和 JPEG2000 压缩标准	331	12.3.4	Hough 变换直线检测的 MATLAB 实现	412
10.8	Visual C++ 综合案例——类似 JPEG 的图像压缩	332	12.3.5	Hough 变换直线检测的 Visual C++ 实现	415
第 11 章	形态学图像处理	341	12.4	阈值分割	418
11.1	预备知识	341	12.4.1	阈值分割方法	419
11.2	二值图像中的基本形态学运算	342	12.4.2	MATLAB 实现	422
11.2.1	腐蚀及其实现	343			
11.2.2	膨胀及其实现	350			
11.2.3	开运算及其实现	353			

12.4.3	Visual C++实现	423	14.3.1	最小距离分类器及其 MATLAB实现	476
12.5	区域分割	425	14.3.2	基于相关的模板匹配	477
12.5.1	区域生长及其实现	425	14.3.3	相关匹配的计算效率	482
12.5.2	区域分裂与合并及其 MATLAB实现	429	第 15 章	神经网络	484
12.6	小结	433	15.1	神经网络简介	484
第 13 章	特征提取	434	15.1.1	仿生学动机	484
13.1	图像特征概述	434	15.1.2	神经网络的应用 实例	486
13.2	基本统计特征	436	15.2	神经网络的理论基础	487
13.2.1	简单的区域描绘子及其 MATLAB实现	436	15.2.1	训练线性单元的梯度 下降算法	487
13.2.2	直方图及其统计特征	437	15.2.2	多层神经网络	492
13.2.3	灰度共现矩阵及其 Visual C++实现	439	15.2.3	Sigmoid 单元	492
13.3	特征降维	442	15.2.4	反向传播(Back Propagation, BP)算法	493
13.3.1	维度灾难	442	15.2.5	训练中的问题	496
13.3.2	特征选择简介	443	15.3	基于 ANN 的数字字符识别系统 DigitRec——分析与设计	498
13.3.3	主成分分析	444	15.3.1	任务描述	498
13.3.4	快速 PCA 及其实现	450	15.3.2	数据集简介	498
13.4	综合案例——基于 PCA 的人脸 特征抽取	451	15.3.3	设计要点	498
13.4.1	数据集简介	452	15.4	基于 ANN 的数字字符识别 系统——DigitRec 的实现	500
13.4.2	生成样本矩阵	452	15.4.1	构建神经元结构—— SNeuron	500
13.4.3	主成分分析	453	15.4.2	构建神经网络网络层—— SNeuronLayer	501
13.4.4	主成分脸可视化分析	454	15.4.3	神经网络信息头—— NeuralNet_Header	502
13.4.5	基于主分量的人脸重建	456	15.4.4	神经网络类—— CNeuralNet	502
13.5	局部二进制模式	457	15.4.5	神经网络的训练数据类—— CNeuralData	513
13.5.1	基本 LBP	457	15.4.6	误差跟踪类—— CValueTrack	518
13.5.2	圆形邻域的 LBP _{PR} 算子	458	15.4.7	训练对话框类—— CTrainDlg	520
13.5.3	统一化 LBP 算子—— Uniform LBP 及其 MATLAB 实现	459	15.4.8	测试对话框类—— CTestDlg	523
13.5.4	MB-LBP 及其 MATLAB 实现	462	15.5	基于 ANN 的数字字符识别系统 ——DigitRec 的测试	526
13.5.5	图像分区及其 MATLAB 实现	467			
第 14 章	图像识别初步	470			
14.1	模式识别概述	470			
14.2	模式识别方法分类	474			
14.3	最小距离分类器和模板匹配	476			

15.5.1	训练	526	16.3.3	应用实例	557
15.5.2	测试	526	16.4	综合案例——基于 PCA 和 SVM 的人脸识别系统	557
15.6	改进的 DigitRec	527	16.4.1	人脸识别简介	558
15.6.1	数字字符图像的预处理类——COCRImageProcess	527	16.4.2	前期处理	558
15.6.2	输入图像的预处理——实现	528	16.4.3	数据规格化	558
15.6.3	输入图像的预处理——测试	539	16.4.4	核函数的选择	561
15.7	神经网络参数对训练和识别的影响	540	16.4.5	参数选择	562
15.7.1	隐藏层单元数目的影响	540	16.4.6	构建多类 SVM 分类器	564
15.7.2	学习率的影响	541	16.4.7	实验结果	566
15.7.3	训练时代数目的影响	542	16.5	SVM 在线资源	571
第 16 章	支持向量机	544	16.5.1	MATLAB 的 SVM 工具箱	572
16.1	支持向量机的分类思想	544	16.5.2	LibSVM 的简介	572
16.2	支持向量机的理论基础	545	第 17 章	AdaBoost	573
16.2.1	线性可分情况下的 SVM	545	17.1	AdaBoost 分类思想	573
16.2.2	非线性可分情况下的 C-SVM	548	17.2	AdaBoost 理论基础	575
16.2.3	需要核函数映射情况下的 SVM	550	17.3	构建 AdaBoost 的 MATLAB 工具箱	577
16.2.4	推广到多类问题	553	17.4	MATLAB 综合案例——基于 AdaBoost 的面部图像男女性别分类	580
16.3	SVM 的 MATLAB 实现	554	17.4.1	关于数据集	580
16.3.1	训练——svmtrain	555	17.4.2	数据的预处理	581
16.3.2	分类——svmclassify	556	17.4.3	算法流程实现	581
			参考文献		583



第0章 初识数字图像处理与机器视觉

图像是指能在人的视觉系统中产生视觉印象的客观对象，包括自然景物、拍摄到的图片、用数学方法描述的图形等。图像的要素有几何要素（刻画对象的轮廓、形状等）和非几何要素（刻画对象的颜色、材质等）。

本章中，主要讲解数字图像和数字图像处理的实质内容和一般步骤，以及一些后面会经常使用到的基本概念。

0.1 数字图像

自然界中的图像都是模拟量，在计算机普遍应用之前，电视、电影、照相机等图像记录与传输设备都是使用模拟信号对图像进行处理。但是，计算机只能处理数字量，而不能直接处理模拟图像。所以要在使用计算机处理图像之前进行图像数字化。

0.1.1 什么是数字图像

简单地说，数字图像就是能够在计算机上显示和处理的图像，可根据其特性分为两大类——位图和矢量图。位图通常使用数字阵列来表示，常见格式有 BMP、JPG、GIF 等；矢量图由矢量数据库表示，接触最多的就是 PNG 图形。



提示

本书只涉及数字图像中位图图像的处理与识别，如无特别说明，后文提到的“图像”和“数字图像”都仅仅是指位图图像。一般而言，使用数字摄像机或数字照相机得到的图像都是位图图像。

将一幅图像视为一个二维函数 $f(x, y)$ ，其中 x 和 y 是空间坐标，而在 $x - y$ 平面中的任意一对空间坐标 (x, y) 上的幅值 f 称为该点图像的灰度、亮度或强度。此时，如果 f 、 x 、 y 均为非负有限离散，则称该图像为数字图像（位图）。

一个大小为 $M \times N$ 数字图像是由 M 行 N 列的有限元素组成的，每个元素都有特定的位置和幅值，代表了其所在行列位置上的图像物理信息，如灰度和色彩等。这些元素称为图像元素或像素。

0.1.2 数字图像的显示

不论是 CRT 显示器还是 LCD 显示器，都是由许多点构成

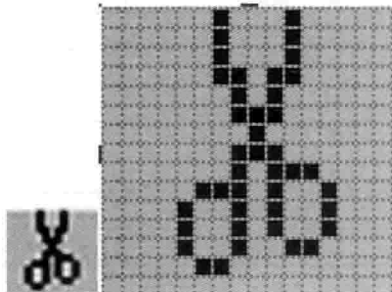


图 0.1 位图图像示例

的,显示图像时这些点对应着图像的像素,称显示器为位映像设备。所谓位映像,就是一个二维的像素矩阵,而位图也就是采用位映像方法显示和存储的图像。当一幅数字图像被放大后就可以明显地看出图像是由很多方格形状的像素构成的,如图0.1所示。

0.1.3 数字图像的分类

根据每个像素所代表信息的不同,可将图像分为二值图像、灰度图像、RGB图像以及索引图像等。

1. 二值图像

每个像素只有黑、白两种颜色的图像称为二值图像。在二值图像中,像素只有0和1两种取值,一般用0来表示黑色,用1表示白色。

2. 灰度图像

在二值图像中进一步加入许多介于黑色与白色之间的颜色深度,就构成了灰度图像。这类图像通常显示为从最暗黑色到最亮的白色的灰度,每种灰度(颜色深度)称为一个灰度级,通常用 L 表示。在灰度图像中,像素可以取 $0\sim L-1$ 之间的整数值,根据保存灰度数值所使用的数据类型不同,可能有256种取值或者说 2^k 种取值,当 $k=1$ 时即退化为二值图像。

3. RGB图像

众所周知,自然界中几乎所有颜色都可以由红(Red, R)、绿(Green, G)、蓝(Blue, B)3种颜色组合而成,通常称它们为RGB三原色。计算机显示彩色图像时采用最多的就是RGB模型,对于每个像素,通过控制R、G、B三原色的合成比例决定该像素的最终显示颜色。

对于三原色RGB中的每一种颜色,可以像灰度图那样使用 L 个等级来表示含有这种颜色成分的多少。例如对于含有256个等级的红色,0表示不含红色成分,255表示含有100%的红色成分。同样,绿色和蓝色也可以划分为256个等级。这样每种原色可以用8位二进制数据表示,于是3原色总共需要24位二进制数,这样能够表示出的颜色种类数目为 $256\times 256\times 256=2^{24}$,大约有1600万种,已经远远超过普通人所能分辨出的颜色数目。

RGB颜色代码可以使用十六进制数减少书写长度,按照两位一组的方式依次书写R、G、B三种颜色的级别。例如:0xFF0000代表纯红色,0x00FF00代表纯绿色,而0x00FFFF是青色(这是绿色和蓝色的加和)。当RGB三种颜色的浓度一致时,所表示的颜色就退化为灰度,比如0x808080就是50%的灰色,0x000000为黑色,而0xFFFFFFFF为白色。常见颜色的RGB组合值如表0.1所示。

表0.1 常见颜色的RGB组合值

颜色	R	G	B
红(0xFF0000)	255	0	0
蓝(0x00FF00)	0	255	0
绿(0x0000FF)	0	0	255
黄(0xFFFF00)	255	255	0
紫(0xFF00FF)	255	0	255
青(0x00FFFF)	0	255	255
白(0xFFFFFFFF)	255	255	255
黑(0x000000)	0	0	0
灰(0x808080)	128	128	128

未经压缩的原始BMP文件就是使用RGB标准给出的3个数值来存储图像数据的,称为RGB图像。在RGB图像中每个像素都是用24位二进制数表示,故也称为24位真彩色图像。

4. 索引图像

如果对每个像素都直接使用 24 位二进制数表示, 图像文件的体积将变得十分庞大。来看一个例子, 对一个长、宽各为 200 像素, 颜色数为 16 的彩色图像, 每个像素都用 RGB 三个分量表示。这样每个像素由 3 个字节表示, 整个图像就是 $200 \times 200 \times 3 = 120\text{kB}$ 。这种完全未经压缩的表示方式, 浪费了大量的存储空间, 下面简单介绍另一种更节省空间的存储方式: 索引图像。

同样还是对 200×200 像素的 16 色图像, 由于这张图片中最多只有 16 种颜色, 那么可以用一张颜色表 (16×3 的二维数组) 保存这 16 种颜色对应的 RGB 值, 在表示图像的矩阵中使用那 16 种颜色在颜色表中的索引 (偏移量) 作为数据写入相应的行列位置。例如, 颜色表中第 3 个元素为 $0xAA1111$, 那么在图像中所有颜色为 $0xAA1111$ 的像素均可以由 $3-1=2$ 表示 (颜色表索引下标从 0 开始)。这样一来, 每一个像素所需要使用的二进制数就仅仅为 4 位 (0.5 字节), 从而整个图像只需要 $200 \times 200 \times 0.5 = 20\text{kB}$ 就可以存储, 而不会影响显示质量。

上文所指的颜色表就是常说的调色板 (Palette), 另一种说法叫作颜色查找表 (Look Up Table, LUT)。Windows 位图中应用到了调色板技术。其实不仅是 Windows 位图, 许多其他的图像文件格式比如 PCX、TIF、GIF 都应用了这种技术。

在实际应用中, 调色板中通常只有少于 256 种的颜色。在使用许多图像编辑工具生成或者编辑 GIF 文件的时候, 常常会提示用户选择文件包含的颜色数目。当选择较低的颜色数目时, 将会有效地降低图像文件的体积, 但也会一定程度上降低图像的质量。

使用调色板技术可以减小图像文件体积的条件是图像的像素数目相对较多, 而颜色种类相对较少。如果一个图像中用到了全部的 24 位真彩色, 对其使用颜色查找表技术是完全没有意义的, 单纯从颜色角度对其进行压缩是不可能的。

0.1.4 数字图像的实质

实际上, 0.1.1 小节中对于数字图像 $f(x, y)$ 的定义仅适用于最为一般的情况, 即静态的灰度图像。更严格地说, 数字图像可以是 2 个变量 (对于静止图像, Static Image) 或 3 个变量 (对于动态画面, Video Sequence) 的离散函数。在静态图像的情况下是 $f(x, y)$, 而如果是动态画面, 则还需要时间参数 t , 即 $f(x, y, t)$ 。函数值可能是一个数值 (对于灰度图像), 也可能是一个向量 (对于彩色图像)。

提示

静态的灰度图像是本书研究的主要对象, 对于函数值为向量的情况会在第 9 章彩色图像处理中阐述。

图像处理是一个涉及诸多研究领域的交叉学科, 下面就从不同的角度来审视数字图像。

(1) 从线性代数和矩阵论的角度, 数字图像就是一个由图像信息组成的二维矩阵, 矩阵的每个元素代表对应位置上的图像亮度和/或色彩信息。当然, 这个二维矩阵在数据表示和存储上可能不是二维的, 这是因为每个单位位置的图像信息可能需要不只一个数值来表示, 这样可能需要一个三维矩阵来对其进行表示 (参见 1.2 节关于 Matlab 中 RGB 图像表示的介绍)。

(2) 由于随机变化和噪声的原因, 图像在本质上是统计性的。因而有时将图像函数作为随机过程的实现来观察其存在的优越性。这时有关图像信息量和冗余的问题可以用概率分布和相关函数来描述和考虑。例如, 如果知道概率分布, 可以用熵 (Entropy) $H^{\textcircled{1}}$ 来度量图像的信息量, 这是信息论中一个重要的思想。

(3) 从线性系统的角度考虑, 图像及其处理也可以表示为用狄拉克冲激公式表达的点展开函数

^① 熵 (Entropy): 熵是信息论中用于度量信息量的一个概念。一个系统越是有序, 信息熵就越低; 反之, 一个系统越是混乱, 信息熵就越高。所以, 信息熵也可以说是系统有序化程度的一个度量。

的叠加, 在使用这种方式对图像进行表示时, 可以采用成熟的线性系统理论研究。在大多数时候, 都考虑使用线性系统近似的方式对图像进行近似处理以简化算法。虽然实际的图像并不是线性的, 但是图像坐标和图像函数的取值都是有限的和非连续的。

0.1.5 数字图像的表达

为了表述像素之间的相对和绝对位置, 通常还需要对像素的位置进行坐标约定。本书中所使用的坐标约定如图 0.2 所示。但在 MATLAB 中坐标的约定会有变化, 具体请参见 1.1.5 小节。

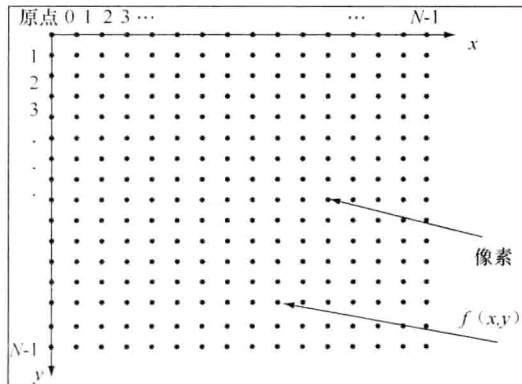


图 0.2 数字图像的坐标约定

在这之后, 一幅物理图像就被转化成了数字矩阵, 从而成为计算机能够处理的对象了。数字图像 f 的矩阵表示如下所示。

$$f(y,x) = \begin{bmatrix} f(0,0) & \cdots & f(0,N-1) \\ \vdots & \ddots & \vdots \\ f(M-1,0) & \cdots & f(M-1,N-1) \end{bmatrix} \quad (0-1)$$

有时也可以使用传统矩阵表示法来表示数字图像和像素, 如下式所示。

$$A = \begin{bmatrix} a_{0,0} & \cdots & a_{0,N-1} \\ \vdots & \ddots & \vdots \\ a_{M-1,0} & \cdots & a_{M-1,N-1} \end{bmatrix} \quad (0-2)$$

其中行列 (M 行 N 列) 必须为正整数, 而离散灰度级数目 L 一般为 2 的 k 次幂, k 为整数 (因为使用二进制整数值表示灰度值), 图像的动态范围为 $[0, L-1]$, 那么图像存储所需的比特数为 $b = M \times N \times k$ 。注意到在矩阵 $f(y,x)$ 中, 一般习惯于先行下标, 后列下标的表示方法, 因此这里先是纵坐标 y (对应行), 然后才是横坐标 x (对应列)。

而有些图像矩阵中, 很多像素的值都是相同的。例如在一个纯黑背景上使用不同灰度勾勒的图像, 大多数像素的值都会是 0。这种矩阵称为稀疏矩阵 (Sparse Matrix), 可以通过简单描述非零元素的值和位置来代替, 大量地写入 0 元素。这时存储图像需要的比特数可能会大大减少。

0.1.6 图像的空间和灰度级分辨率

1. 图像的空间分辨率 (Spatial Resolution)

图像的空间分辨率是指图像中每单位长度所包含的像素或点的数目, 常以像素/英寸 (pixels per inch, ppi) 为单位来表示。如 72ppi 表示图像中每英寸包含 72 个像素或点。分辨率越高, 图像将越

清晰，图像文件所需的磁盘空间也越大，编辑和处理所需的时间也越长。

像素越小，单位长度所包含的像素数据就越多，分辨率也就越高，但同样物理大小范围内所对应图像的尺寸也会越大，存储图像所需要的字节数也越多。因而，在图像的放大缩小算法中，放大就是对图像的过采样，缩小是对图像的欠采样，这些会在 4.5 节图像缩放中进一步介绍。

一般在没有必要对涉及像素的物理分辨率进行实际度量时，通常会称一幅大小为 $M \times N$ 的数字图像的空间分辨率为 $M \times N$ 像素。

图 0.3 给出了同一幅图像在不同的空间分辨率下呈现出的不同效果。当高分辨率下的图像以低分辨率表示时，在同等的显示或者打印输出条件下，图像的尺寸变小，细节变得不明显；而当将低分辨率下的图像放大时，则会导致图像的细节仍然模糊，只是尺寸变大。这是因为缩小的图像已经丢失了大量的信息，在放大图像时只能通过复制行列的插值的方法来确定新增像素的取值。

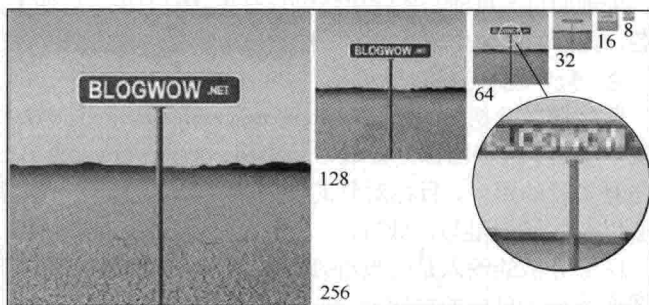


图 0.3 图像的空间分辨率——一幅分辨率为 1024×1024 的图像逐次减少至 32×32 的分辨率

2. 图像的灰度级/辐射计量分辨率 (Radiometric Resolution)

在数字图像处理中，灰度级分辨率又叫色阶，是指图像中可分辨的灰度级数目，即前文提到的灰度级数目 L ，它与存储灰度级别所使用的数据类型有关。由于灰度级度量的是投射到传感器上光辐射值的强度，所以灰度级分辨率也叫辐射计量分辨率。

随着图像的灰度级分辨率逐渐降低，图像中包含的颜色数目变少，从而在颜色的角度造成图像信息受损，同样使图像细节表达受到了一定的影响，如图 0.4 所示。

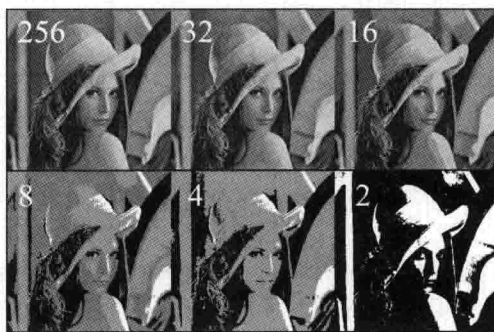


图 0.4 图像的灰度级分辨率——分别具有 256、32、16、8、4 和 2 个灰度级的一幅图像

0.2 数字图像处理与机器视觉

0.2.1 从图像处理到图像识别

图像处理、图像分析和图像识别是认知科学与计算机科学中的一个令人兴奋的活跃分支。从 1970 年这个领域经历了人们对其兴趣的爆炸性增长以来，到 20 世纪末逐渐步入成熟。其中遥感、技术诊断、智能车自主导航、医学平面和立体成像以及自动监视领域是发展最快的一些方向。这种进展最集中地体现在市场上多种应用这类技术的产品纷纷涌现。事实上，从数字图像处理到数字图像分析，再发展到最前沿的图像识别技术，其核心都是对数字图像中所含有的信息的提取及与其