

一样的软件，不一样的学习方法

精通

Matlab

数字图像处理与识别

张铮 倪红霞 苑春苗 杨立红 编著

- 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、形态学处理、图像分割，以及图像特征提取的相关内容。
- 重点介绍了目前在工程技术领域非常流行的方尖技术——人工神经网络（ANN）和支持向量机（SVM），并讲解了人脸识别技术。



CD-ROM 源程序



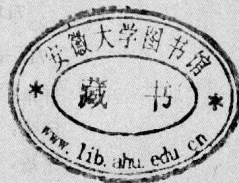
人民邮电出版社
POSTS & TELECOM PRESS

精通

Matlab

数字图像处理与识别

■ 张铮 倪红霞 苑春苗 杨立红 编著



人民邮电出版社
北京

图书在版编目 (C I P) 数据

精通Matlab数字图像处理与识别 / 张铮等编著. --
北京: 人民邮电出版社, 2013. 4
ISBN 978-7-115-30463-6

I. ①精… II. ①张… III. ①Matlab软件—应用—数字图象处理 IV. ①TN911.73

中国版本图书馆CIP数据核字(2012)第307738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、形态学处理、图像分割以及图像特征提取的相关内容; 同时, 对于机器视觉进行了前瞻性的探究, 重点介绍了两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)和支持向量机(SVM), 并讲解了人脸识别的热点技术。

全书结构紧凑, 内容深入浅出, 讲解图文并茂, 适合计算机、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的广大工程技术人员参考使用。

精通 Matlab 数字图像处理与识别

-
- ◆ 编 著 张 铮 倪红霞 苑春苗 杨立红
责任编辑 张 涛
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 25.75
字数: 711 千字
印数: 1-3 500 册
- 2013 年 4 月第 1 版
2013 年 4 月河北第 1 次印刷

ISBN 978-7-115-30463-6

定价: 65.00 元 (附光盘)

读者服务热线: (010)67132692 印装质量热线: (010)67129223
反盗版热线: (010)67171154

前 言

图像处理与识别是当今计算机科学中一个热门研究方向，其应用广泛，发展前景乐观。Matlab 是 MathWorks 公司开发的一款工程数学计算软件。由于其强大的科学运算、灵活的程序设计流程、高质量的图形可视化与界面设计，以及便捷的与其他程序和语言接口的功能，Matlab 已成为当今国际上科学界最具影响力、最有活力的软件。本书正是紧密结合大量的 Matlab 代码和案例展开对数字图像处理和识别技术的介绍。

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。但要掌握好数字图像处理与识别技术却并非易事，它的理论性较强、门槛较高，在各个高校中，这门课程多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。我们在撰写本书过程中，对于可能造成读者理解困难的地方，均给出了必要的先序知识，尽量定性地描述；对于那些并不一目了然的结论均给出了思路和解释；对于某些非常专业、已经超过本书讨论范围的相关知识也都在全书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，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。

—— 内容安排

全书共分 15 章，主要内容介绍如下。

第 1~2 章介绍了数字图像处理的基础知识和 Matlab 数字图像编程基础，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，并且熟悉本书自始至终需要使用的工具 Matlab，重点介绍其数字图像处理工具箱。

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。傅立叶变换一直是频率域图像处理的基石，它能用正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。例如，对于一张乐谱，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，但傅立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 10~12 章（形态学处理、边缘检测与图像分割、特征提取）是从单纯图像处理向图像识别（机器视觉）的过渡，这一阶段的特点是，输入是图像，输出则是在识别意义上我们感兴趣的图像元素。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 13 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。

第 14~15 章（人工神经网络、支持向量机）介绍了两种十分强大的分类技术，并介绍了人脸识别技术及经典案例应用。

—— 读者对象

- 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。

—— 在阅读本书之前，读者最好具有如下的预备知识

• 读者应具备一定的数学基础，如高等数学知识、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）。

—— 在线支持和读者反馈

本书所有 Matlab 实例的源代码均可在 bbs.book95.com 网站的“金羽图书与答疑板块”板块与本书同名的主题帖子附件中提供下载。虽然本书中的所有例子都已经在 Windows XP、Windows 2003 和 Windows 7 等操作系统下的 Matlab2006 到 MatlabR2011a 的各个版本中测试通过，但由于交稿时间要求和笔者水平的局限，也有存在 Bug 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，请您不吝告知本书的作者（aaron@book95.com），以便我们做出改进。

—— 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内容的严谨。

本书的第 7 章、第 8 章由王艳艳和赵国宇编写，在此向他们表示感谢；同时感谢我的好友王艳平提供并调试了许多实例代码；感谢我的同事——闵卫东、陈香凝、任淑霞、姚清爽、王佳欣、孙连坤、孙学梅、张振和王作为等为本书所做出的工作；感谢我昔日的师弟和师妹——王杉、闫丽霞、刘旭、赵国宇和李宏鹏等参与了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；还要感谢徐超、王欣和郭朋博士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、儿子张垚森以及我的家人，没有你们的鼓励和支持就不会有我的这部作品。

编者

目 录

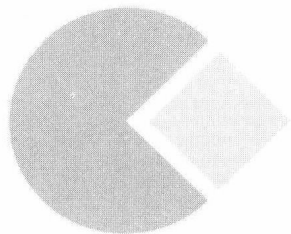
第 1 章 初识数字图像处理与识别	1	2.4 读取和写入图像文件	37
1.1 数字图像	1	2.5 图像的显示	39
1.1.1 什么是数字图像	1	第 3 章 图像的点运算	42
1.1.2 数字图像的显示	1	3.1 灰度直方图	42
1.1.3 数字图像的分类	2	3.1.1 理论基础	42
1.1.4 数字图像的实质	3	3.1.2 Matlab 实现	43
1.1.5 数字图像表示	4	3.2 灰度的线性变换	46
1.1.6 图像的空间和灰度级 分辨率	5	3.2.1 理论基础	46
1.2 数字图像处理与识别	6	3.2.2 Matlab 实现	46
1.2.1 从图像处理到图像识别	6	3.3 灰度对数变换	49
1.2.2 数字图像处理与识别的 应用实例	7	3.3.1 理论基础	49
1.2.3 数字图像处理与识别的 基本步骤	9	3.3.2 Matlab 实现	50
1.3 数字图像处理的预备知识	10	3.4 伽玛变换	51
1.3.1 邻接性、连通性、区域 和边界	10	3.4.1 理论基础	51
1.3.2 距离度量的几种方法	11	3.4.2 Matlab 实现	51
1.3.3 基本的图像操作	12	3.5 灰度阈值变换	53
第 2 章 Matlab 数字图像处理基础	13	3.5.1 理论基础	53
2.1 Matlab R2011a 简介	13	3.5.2 Matlab 实现	54
2.1.1 Matlab 软件环境	13	3.6 分段线性变换	55
2.1.2 文件操作	14	3.6.1 理论基础	55
2.1.3 在线帮助的使用	15	3.6.2 Matlab 实现	56
2.1.4 变量的使用	18	3.7 直方图均衡化	60
2.1.5 矩阵的使用	20	3.7.1 理论基础	60
2.1.6 细胞数组和结构体	23	3.7.2 Matlab 实现	61
2.1.7 关系运算与逻辑运算	24	3.8 直方图规定化	63
2.1.8 常用图像处理数学函数	25	3.8.1 理论基础	63
2.1.9 Matlab 程序流程控制	26	3.8.2 Matlab 实现	64
2.1.10 M 文件编写	29	第 4 章 图像的几何变换	66
2.1.11 Matlab 函数编写	30	4.1 解决几何变换的一般思路	66
2.2 Matlab 图像类型及其存储方式	32	4.2 图像平移	67
2.3 Matlab 的图像转换	35	4.2.1 图像平移的变换公式	68
		4.2.2 图像平移的 Matlab 实现	68
		4.3 图像镜像	70
		4.3.1 图像镜像的变换公式	70
		4.3.2 图像镜像的 Matlab 实现	71

4.4	图像转置	72	5.5.2	基于一阶导数的图像增强——梯度算子	98
4.4.1	图像转置的变换公式	72	5.5.3	基于二阶微分的图像增强——拉普拉斯算子	101
4.4.2	图像转置的 Matlab 实现	72	5.5.4	基于一阶与二阶导数的锐化算子的比较	104
4.5	图像缩放	73	5.5.5	高提升滤波及其实现	105
4.5.1	图像缩放的变换公式	73	5.5.6	高斯-拉普拉斯变换 (Laplacian of a Gaussian, LoG)	106
4.5.2	图像缩放的 Matlab 实现	74			
4.6	图像旋转	75	第 6 章	频率域图像增强	109
4.6.1	以原点为中心的图像旋转	75	6.1	频率域滤波——与空间域滤波殊途同归	109
4.6.2	以任意点为中心的图像旋转	75	6.2	傅立叶变换基础知识	109
4.6.3	图像旋转的 Matlab 实现	77	6.2.1	傅立叶级数	110
4.7	插值算法	77	6.2.2	傅立叶变换	111
4.7.1	最近邻插值	78	6.2.3	幅度谱、相位谱和功率谱	114
4.7.2	双线性插值	78	6.2.4	傅立叶变换的实质——基的转换	115
4.7.3	高阶插值	79	6.3	快速傅立叶变换及实现	116
4.8	Matlab 综合案例——人脸图像配准	81	6.3.1	FFT 变换的必要性	117
4.8.1	什么是图像配准	81	6.3.2	常见的 FFT 算法	117
4.8.2	人脸图像配准的 Matlab 实现	81	6.3.3	按时间抽取的基-2 FFT 算法	118
			6.3.4	离散反傅立叶变换的快速算法	121
第 5 章	空间域图像增强	85	6.3.5	N 维快速傅立叶变换	121
5.1	图像增强基础	85	6.3.6	Matlab 实现	122
5.1.1	为什么要进行图像增强	85	6.4	频域滤波基础	126
5.1.2	图像增强的分类	85	6.4.1	频域滤波与空域滤波的关系	126
5.2	空间域滤波	86	6.4.2	频域滤波的基本步骤	126
5.2.1	空间域滤波和邻域处理	86	6.4.3	频域滤波的 Matlab 实现	127
5.2.2	边界处理	87	6.5	频域低通滤波器	128
5.2.3	相关和卷积	88	6.5.1	理想低通滤波器及其实现	128
5.2.4	滤波操作的 Matlab 实现	88	6.5.2	高斯低通滤波器及其实现	131
5.3	图像平滑	90	6.6	频率域高通滤波器	135
5.3.1	平均模板及其实现	90	6.6.1	高斯高通滤波器及其实现	135
5.3.2	高斯平滑及其实现	92			
5.3.3	自适应平滑滤波	95			
5.4	中值滤波	95			
5.4.1	性能比较	95			
5.4.2	一种改进的中值滤波策略	98			
5.4.3	中值滤波的工作原理	98			
5.5	图像锐化	98			
5.5.1	理论基础	98			

6.6.2	频域拉普拉斯滤波器及其实现	137	9.3.1	彩色补偿及其 Matlab 实现	215
6.7	Matlab 综合案例 —— 利用频域滤波消除周期噪声	139	9.3.2	彩色平衡及其 Matlab 实现	217
6.7.1	频域带阻滤波器	139	第 10 章	形态学图像处理	220
6.7.2	带阻滤波消除周期噪声	141	10.1	预备知识	220
6.8	频域滤波器与空域滤波器之间的内在联系	143	10.2	二值图像中的基本形态学运算	222
第 7 章	小波变换	146	10.2.1	腐蚀及其实现	222
7.1	多分辨率分析	146	10.2.2	膨胀及其实现	227
7.1.1	多分辨率框架	146	10.2.3	开运算及其实现	229
7.1.2	分解与重构的实现	153	10.2.4	闭运算及其实现	231
7.1.3	图像处理中分解与重构的实现	155	10.3	二值图像中的形态学应用	232
7.2	Gabor 多分辨率分析	160	10.3.1	击中与击中不中变换及其实现	232
7.3	常见小波分析	163	10.3.2	边界提取与跟踪及其实现	234
7.3.1	Haar 小波	164	10.3.3	区域填充	235
7.3.2	Daubechies 小波	166	10.3.4	连通分量提取及其实现	237
7.4	高维小波	168	10.3.5	细化算法	240
第 8 章	图像复原	171	10.3.6	像素化算法	242
8.1	图像复原的一般理论	171	10.3.7	凸壳	243
8.1.1	图像复原的基本概念	171	10.3.8	bwmorph 函数	243
8.1.2	图像复原的一般模型	172	10.4	灰度图像中的基本形态学运算	244
8.2	实用图像复原技术	190	10.4.1	灰度膨胀及其实现	244
8.2.1	图像复原的数值计算方法	190	10.4.2	灰度腐蚀及其实现	247
8.2.2	非线性复原	193	10.4.3	灰度开、闭运算及其实现	248
第 9 章	彩色图像处理	197	10.4.4	顶帽变换 (top-hat) 及其实现	250
9.1	彩色基础	197	10.5	小结	252
9.1.1	什么是彩色	198	第 11 章	图像分割	253
9.1.2	我们眼中的彩色	198	11.1	图像分割概述	253
9.1.3	三原色	198	11.2	边缘检测	254
9.1.4	计算机中的颜色表示	199	11.2.1	边缘检测概述	254
9.2	彩色模型	200	11.2.2	常用的边缘检测算子	255
9.2.1	RGB 模型	200	11.2.3	Matlab 实现	258
9.2.2	CMY、CMYK 模型	202	11.3	霍夫变换	261
9.2.3	HSI 模型	203	11.3.1	直线检测	261
9.2.4	HSV 模型	207	11.3.2	曲线检测	264
9.2.5	YUV 模型	210	11.3.3	任意形状的检测	264
9.2.6	YIQ 模型	213	11.3.4	Hough 变换直线检测的 Matlab 实现	265
9.2.7	Lab 模型简介	214			
9.3	全彩色图像处理基础	215			

11.4	阈值分割	268	12.5.3	统一化 LBP 算子—— Uniform LBP 及其 Matlab 实现	316
11.4.1	阈值分割方法	268	12.5.4	MB-LBP 及其 Matlab 实现	319
11.4.2	Matlab 实现	272	12.5.5	图像分区及其 Matlab 实现	324
11.5	区域分割	273	第 13 章	图像识别初步	328
11.5.1	区域生长及其实现	273	13.1	模式识别概述	328
11.5.2	区域分裂与合并及 其 Matlab 实现	275	13.1.1	模式与模式识别	328
11.6	基于形态学分水岭算法的 图像分割	280	13.1.2	图像识别	329
11.6.1	形态学分水岭算法	280	13.1.3	关键概念	329
11.6.2	Matlab 实现	283	13.1.4	识别问题的一般描述	330
11.7	Matlab 综合案例—— 分水岭算法	284	13.1.5	过度拟合	331
11.8	小结	289	13.1.6	模式识别系统结构	332
第 12 章	特征提取	290	13.1.7	训练/学习方法分类	332
12.1	图像特征概述	290	13.2	模式识别方法分类	332
12.1.1	什么是图像特征	290	13.2.1	统计模式识别	333
12.1.2	图像特征的分类	290	13.2.2	句法模式识别	333
12.1.3	特征向量及其几何 解释	291	13.2.3	小结	334
12.1.4	特征提取的一般原则	291	13.3	最小距离分类器和模板匹配	334
12.1.5	特征的评价标准	291	13.3.1	最小距离分类器及其 Matlab 实现	334
12.2	基本统计特征	292	13.3.2	基于相关的模板匹配	336
12.2.1	简单的区域描绘子及 其 Matlab 实现	292	13.3.3	相关匹配的计算效率	339
12.2.2	直方图及其统计特征	293	第 14 章	人工神经网络	341
12.2.3	灰度共现矩阵	295	14.1	人工神经网络简介	341
12.3	特征降维	299	14.1.1	仿生学动机	341
12.3.1	维度灾难	299	14.1.2	人工神经网络的 应用实例	343
12.3.2	特征选择简介	299	14.2	人工神经网络的理论基础	344
12.3.3	主成分分析	300	14.2.1	训练线性单元的梯度 下降算法	344
12.3.4	快速 PCA 及其实现	307	14.2.2	多层人工神经网络	350
12.4	综合案例——基于 PCA 的 人脸特征抽取	309	14.2.3	sigmoid 单元	351
12.4.1	数据集简介	309	14.2.4	反向传播 (BP, back propagation) 算法	352
12.4.2	生成样本矩阵	309	14.2.5	训练中的问题	356
12.4.3	主成分分析	310	14.3	神经网络算法的可视化实现	357
12.4.4	主成分脸可视化分析	311	14.3.1	NNTool 的主要功能 及应用	357
12.4.5	基于主分量的人脸重建	313	14.3.2	神经网络的仿真测试	361
12.5	局部二进制模式	315			
12.5.1	基本 LBP	315			
12.5.2	圆形邻域的 LBP _{PR} 算子	315			

14.4	Matlab 神经网络工具箱	365	15.3	SVM 的 Matlab 实现	380
14.4.1	网络的创建	365	15.3.1	训练——svmtrain	380
14.4.2	网络初始化	365	15.3.2	分类——svmclassify	382
14.4.3	网络训练	366	15.3.3	应用实例	382
14.4.4	网络仿真测试	366	15.4	综合案例——基于 PCA 和 SVM 的人脸识别系统	383
14.4.5	网络性能分析	367	15.4.1	人脸识别简介	383
第 15 章	支持向量机	368	15.4.2	前期处理	383
15.1	支持向量机的分类思想	368	15.4.3	数据规格化	384
15.1.1	分类模型的选择	368	15.4.4	核函数的选择	387
15.1.2	模型参数的选择	369	15.4.5	参数选择	388
15.2	支持向量机的理论基础	369	15.4.6	构建多类 SVM 分类器	390
15.2.1	线性可分情况下的 SVM	369	15.4.7	实验结果	392
15.2.2	非线性可分情况下的 C-SVM	373	15.5	SVM 在线资源	399
15.2.3	需要核函数映射情况 下的 SVM	375	15.5.1	Matlab 的 SVM 工具箱	399
15.2.4	推广到多类问题	378	15.5.2	LIBSVM 的简介	399
			参考文献		401



第1章 初识数字图像处理与识别

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，我们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。

1.1 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。所以我们要在使用计算机处理图像之前进行图像数字化。

1.1.1 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，我们接触最多的就是 PNG 图形。



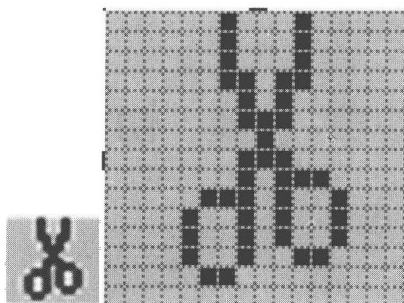
本书只涉及数字图像中位图图像的处理与识别，如无特别说明，后文提到的“图像”和“数字图像”都仅仅是指位图图像。一般而言，使用数字摄像机或数字照相机得到的图像都是位图图像。

我们可以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**。

1.1.2 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 1.1 所示。



▲图 1.1 位图图像示例

1.1.3 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 表示不含红色成份，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，于是三原色总共需要 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 1 600 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，例如 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 1.1 所示。

表 1.1 常见颜色的 RGB 组合值

颜 色	R	G	B
红 (0xFF0000)	255	0	0
蓝 (0x00FF00)	0	255	0

续表

颜色	R	G	B
绿 (0x0000FF)	0	0	255
黄 (0xFFFF00)	255	255	0
紫 (0xFF00FF)	255	0	255
青 (0x00FFFF)	0	255	255
白 (0xFFFFFF)	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 3 个分量表示。这样每个像素由 3 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)**，另一种说法叫做**颜色查找表 (LUT: Look Up Table)**。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。

1.1.4 数字图像的实质

实际上，1.1.1 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，数字图像可以是两个变量 (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，下面就让我们从不同的角度来审视数字图像。

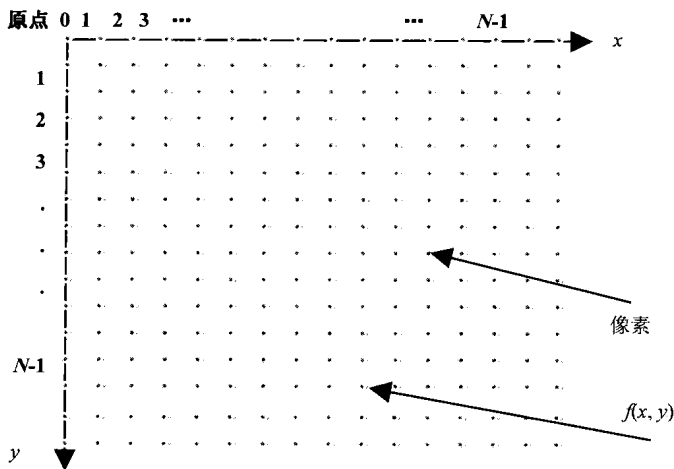
- 从线性代数和矩阵论的角度，数字图像是由图像信息组成的二维矩阵，矩阵的每个元素代表对应位置上的图像亮度和/或色彩信息。当然，这个二维矩阵在数据表示和存储上可能不是二维的，这是因为每个单位位置的图像信息可能需要不只一个数值来表示，这样可能需要一个三维矩阵来对其进行表示（参见 2.2 节关于 Matlab 中 RGB 图像表示的介绍）。

- 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存在其优越性。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^{D} 来度量图像的信息量，这是信息论中一个重要的思想。

- 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，我们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，图像坐标和图像函数的取值都是有限的和非连续的。

1.1.5 数字图像的表达

为了表述像素之间的相对和绝对位置，我们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 1.2 所示。但在 Matlab 中坐标的约定会有变化，具体请参见 2.1.5 小节。



▲图 1.2 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 的矩阵的表示如下。

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

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。

有时也可以使用传统矩阵表示法来表示数字图像和像素。

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

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大大减少。

1.1.6 图像的空间和灰度级分辨率

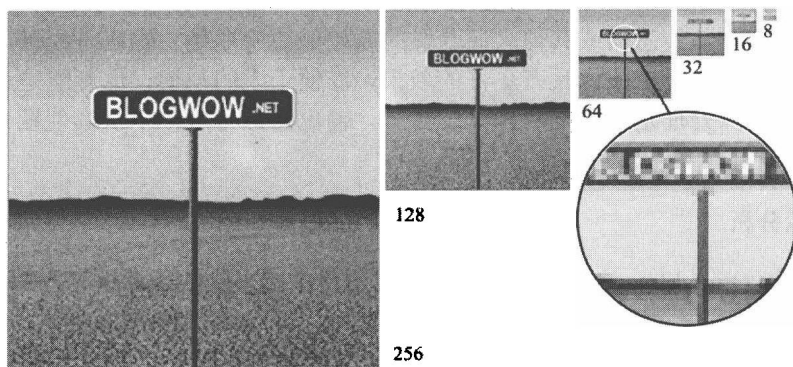
1. 图像的空间分辨率

图像的空间分辨率 (Spatial Resolution) 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 1.3 所示为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方法来确定新增像素的取值。



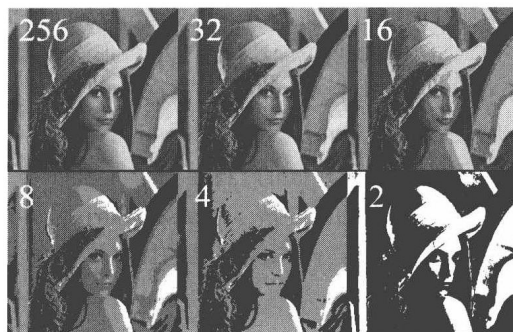
▲图 1.3 图像的空间分辨率 (一幅分辨率为 256×256 的图像逐次减少至 8×8 的分辨率)

2. 图像的灰度级/辐射计量分辨率

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 (Radiometric Resolution)**。

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影响，如图 1.4 所示。



▲图 1.4 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)

12 数字图像处理与识别

1.2.1 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相关的各种辅助过程。

1. 数字图像处理

数字图像处理 (Digital Image Processing) 就是指使用电子计算机对量化的数字图像进行处理，具体地说就是通过对图像进行各种加工来改善图像的外观，是对图像的修改和增强。

图像处理的输入是从传感器或其他来源获取的原始的数字图像，输出是经过处理后的输出图像。处理的目的是使输出图像具有更好的效果，以便于人的观察；也可能是为图像分析和识别做准备，此时的图像处理是作为一种**预处理**步骤，输出图像将进一步供其他图像分析、识别算法使用。

2. 数字图像分析

数字图像分析 (Digital Image Analyzing) 是指对图像中感兴趣的目标进行检测和测量，以获得客观的信息。数字图像分析通常是指将一幅图像转化为另一种非图像的抽象形式，如图像中某物体与测量者的距离以及目标对象的计数或其尺寸等。这一概念的外延包括边缘检测和图像分割、特征提取以及几何测量与计数等。

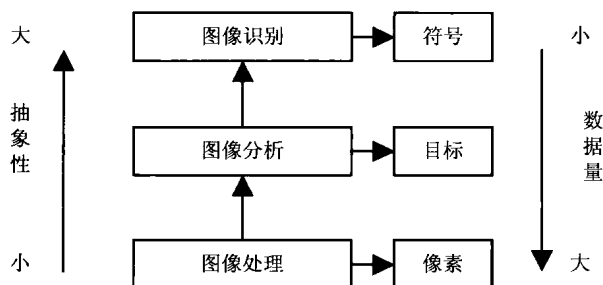
图像分析的输入是经过处理的数字图像，其输出通常不再是数字图像，而是一系列与目标相关的图像特征 (目标的描述)，如目标的长度、颜色、曲率和个数等。

3. 数字图像识别

数字图像识别 (Digital Image Recognition) 主要研究图像中各目标的性质和相互关系, 识别出目标对象的类别, 从而理解图像的含义。这往往囊括了使用数字图像处理技术的很多应用项目, 如光学字符识别 (OCR)、产品质量检验、人脸识别、自动驾驶、医学图像和地貌图像的自动判读理解等。

图像识别是图像分析的延伸, 它根据从图像分析中得到的相关描述 (特征) 对目标进行归类, 输出我们感兴趣的目标类别标号信息 (符号)。

总而言之, 从图像处理到图像分析再到图像识别这个过程, 是一个将所含信息抽象化, 尝试降低信息熵, 提炼有效数据的过程, 如图 1.5 所示。



▲图 1.5 数字图像处理、分析和识别的关系

从信息论的角度上说, 图像应当是物体所含信息的一个概括, 而数字图像处理侧重于将这些概括的信息进行变换, 例如升高或降低熵值。数字图像分析则是将这些信息抽取出来以供其他过程调用。当然, 在不太严格时, 数字图像处理也可以兼指图像处理和分析。

您或许也听过另一个概念, **计算机图形学 (Computer Graphics)**。此概念与数字图像分析大致相反, 它是一个对由概念或数学表述的物体图像进行处理和显示的过程。

1.2.2 数字图像处理与识别的应用实例

如今, 数字图像处理与机器视觉的应用越来越广泛, 已经渗透到国家安全、航空航天、工业控制、医疗保健等各个领域乃至我们的日常生活和娱乐当中, 在国民经济中发挥着举足轻重的作用。它的一些典型的应用如表 1.2 所示。

表 1.2 图像处理与识别的典型应用

相关领域	典型应用
安全监控	指纹验证, 基于人脸识别的门禁系统
工业控制	产品无损检测, 商品自动分类
医疗保健	X 光照片增强, CT, 核磁共振, 病灶自动检测
生活娱乐	基于表情识别的笑脸自动检测, 汽车自动驾驶, 手写字符识别

下面结合 2 个典型的应用来说明。

1. 图像处理的典型案例——X 光照片的增强

图 1.6 所示为两幅图片, (a) 是一幅直接拍摄未经处理的 X 光照片, 对比度较低, 图像细节难以辨识; (b) 中呈现了 (a) 经过简单的增强处理后的效果, 图像较为清晰, 可以有效地指导诊断