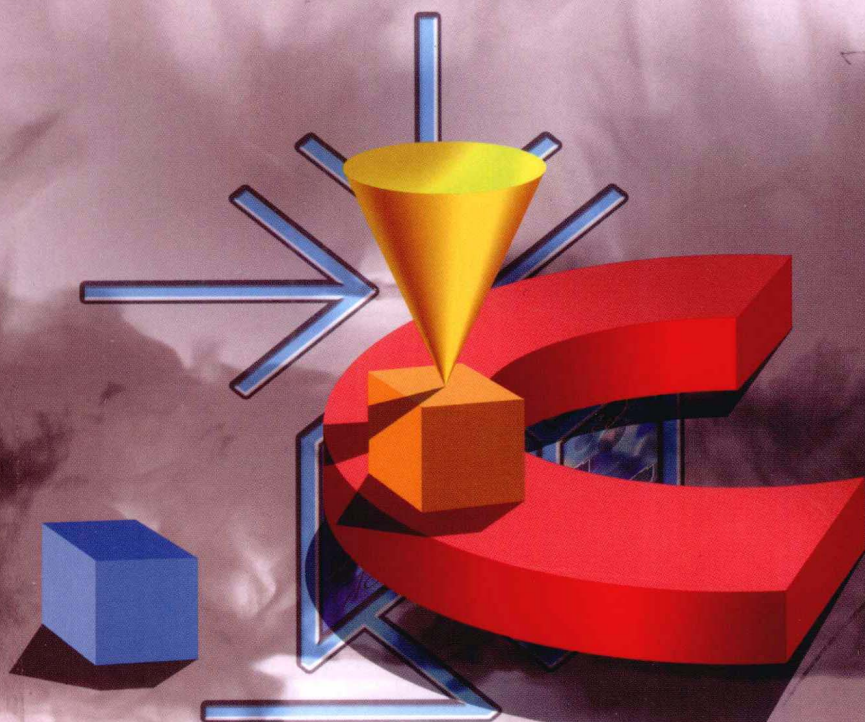


21世纪重点大学规划教材

孙兴华 郭丽 编著

数字图像处理

——编程框架、理论分析、实例
应用和源码实现



21 世纪重点大学规划教材

数字图像处理

——编程框架、理论分析、实例应用和源码实现

孙兴华 郭丽 编著



机械工业出版社

出版说明

“211 工程”是“重点大学和重点学科建设项目”的简称，是国家“九五”期间唯一的教育重点项目。

进入“211 工程”的 100 所学校拥有全国 32% 的在校本科生、69% 的硕士、84% 的博士生，以及 87% 的有博士学位的教师；覆盖了全国 96% 的国家重点实验室和 85% 的国家重点学科。相对而言，这批学校中的教授、教师有着深厚的专业知识和丰富的教学经验，其中不少教师对我国高等院校的教材建设做过很多重要的工作。为了有效地利用“211 工程”这一丰富资源，实现以重点建设推动整体发展的战略构想，机械工业出版社推出了“21 世纪重点大学规划教材”。

本套教材以重点大学、重点学科的精品教材建设为主要任务，组织知名教授、教师进行编写，教材适用于高等院校计算机及其相关专业，选题涉及公共基础课、硬件、软件、网络技术等，内容紧密贴合高等院校相关学科的课程设置和培养目标，注重教材的科学性、实用性、通用性，在同类教材中具有一定的先进性和权威性。

为了体现建设“立体化”精品教材的宗旨，本套教材为主干课程配置了电子教案、学习指导、习题解答、课程设计、毕业设计指导等内容。

机械工业出版社

前 言

数字图像处理是一门综合性的学科,涉及心理学、计算机视觉、机器学习、概率统计等不同的方向。如果沉浸在其中,就会发现,其不但能给从业者带来研究的乐趣,而且会产生巨大的商业应用价值。可以从多种角度来看待数字图像处理(或者说写“数字图像处理”的书),包括数字信号处理、计算机编程处理、普通用户的多媒体应用,以及计算机视觉等角度。从不同的角度编写“数字图像处理”的书,其内容也将是千差万别的。基于作者的从业经历,本书是从计算机编程处理的角度来编写的。

从作者 1997 年开始研究生阶段的学习,直到本书草稿的完成,对数字图像处理的研究和应用已经有十多年了。早在 3 年前,作者就收到过写数字图像处理书的邀请,但那时觉得自己的知识积累还不够,无论是在研究成果还是在教学经验上。因此,尽管一直怀着写一本数字图像处理书的冲动(因为这毕竟是自己长期的研究方向),却总是处于规划阶段。直到 2009 年,当作者从国外访问留学回来,手上的科研压力相对较小的时候,机械工业出版社的编辑与作者进行了联系,询问是否可以写一本数字图像处理方面的书,这重新激发了作者的热情,从而才有了现在的这本书。

开始写这本书的时候,觉得自己的知识积累应该可以满足写书所需了,但是随着编写工作的深入,作者不得不重新学习一些必要的知识。所以,写这本书一方面是为了给读者提供一些关于数字图像处理学习方面的帮助,另一方面也让作者自己在数字图像处理研究和应用上得到了进一步的提高。以前有一种说法,就是要真正学好一门课,就去教这门课。作者觉得这种说法是正确的。现在还觉得,要真正掌握一门学科,如果能尝试着写一本这方面的书,就会取得更好的成绩。这里也与本书的读者分享一下,就是学习数字图像处理时,要始终保持一颗学习的心,一是因为这门学科是无止境的,二是因为她会给你带来无穷的乐趣。

本书的第一大特色,就是提供了大量的全面的源码实例,小到简单的几何变换和边缘算子,大到复杂的兴趣点提取和基于内容的图像检索,所提供的源码几乎涵盖了本书的所有知识点。根据作者目前了解的情况,本书附带的源码实例之多在同类书中是少有的。为什么要提供如此多的源码呢?其原因在于,从作者多年的从业经验而言,很多书对数字图像处理的原理都有介绍,都能看得懂,但是却很难实现,或者是需要很长的时间去实现,这对各个层次的数字图像处理学习者来说都是极其令人沮丧的事情,所以源码实现就显得异常重要了。可以说,本书所提供的源码,凝聚了作者长期的科研经验,目的就是希望能够给众多的数字图像处理学习者提供尽可能多的帮助。当然,很多源码也参考了一些公开资料,在此向这些资料的作者表示感谢和致敬,并由衷地希望本书所提供的一切有朝一日也能成为他人参阅的资料。尽管所有的源码都是经过作者精心测试通过的,但可能还会有部分源码不尽如人意,请读者见谅,今后作者会予以改进。另外,本书所使用的数字图像处理编程框架,是作者在数字图像处理领域中开发并一直使用的。这个完整的基于内容的图像检索系统是作者攻读博士学位时的一个研究成果。

本书的结构安排,也与现有的同类书不太相同。本书以数字图像处理编程框架为主线,

从介绍编程框架开始（除本书的绪论外），以编程框架的接口扩展结束，中间串联着从易到难和由简至繁的数字图像处理功能，简单的有几何变换和图像分割等功能，复杂的有图像压缩编码和图像底层特征的提取，更复杂的有高级数字图像处理部分（包括光流分析、兴趣点提取和基于内容的图像检索）。作者相信，上述结构安排有利于读者的学习，本科生可以从最简单的几何变换开始学起，研究生可以直接参考图像特征提取和高级数字图像处理部分，工程项目开发人员可以在参考本书理论分析的基础上，尽可能地利用本书所提供的源码。所以，本书的适用范围是广泛的，这种以数字图像处理编程框架为轴线并按照从易到难和由简至繁的方式进行章节安排是本书的第二大特色。

本书是孙兴华和郭丽共同完成的，我们之间的分工并不是按照章节来进行的，而是将一切的合作贯穿于整本书的内容构思、文字撰写、资料查阅和源码实现中，所以对整本书的贡献我们应该是同等的。作者的研究生周靓、杨莉娜、付宗运和朱飞翔参与了本书最后的文字校对和细节检查，在此对他们的辛勤劳动表示感谢。

在编写本书的过程中，作者参考了很多的文献和资料，包括最新的学术著作、学术论文，以及互联网上的资料等，在此向这些文献的作者表示由衷的感谢！

本书的编写，大部分都是在工作之余完成的，而这部分时间本应该属于我的家人，所以怀着深深的歉意向我的妻子和孩子说声谢谢！

如果读者在看完这本书之后，感觉有所帮助，作者会感到无比的高兴，因为这是对作者最大的安慰。当然，由于作者的水平有限，所以非常欢迎和恳请读者能够就本书的所有方面给予反馈意见，以帮助作者改进书中的不足。

编 者

目 录

出版说明

前言

第 1 章 数字图像处理绪论	1
1.1 数字图像处理目的	2
1.1.1 面向终端用户	2
1.1.2 面向专业编程	4
1.2 数字图像处理内容	6
1.2.1 底层图像处理技术	6
1.2.2 中层图像处理技术	7
1.2.3 高层图像处理技术	8
1.3 数字图像处理应用	9
1.3.1 图像增强与恢复	9
1.3.2 基于图像的生物特征识别	10
1.3.3 基于内容的图像检索	11
1.3.4 图像序列分析	11
1.3.5 计算机视觉	11
1.4 本章小结	13
1.5 本章习题	13
参考文献	14
第 2 章 基于 Visual C++ 的编程框架	15
2.1 编程框架	15
2.1.1 C/C++ 语言在数字图像处理中的优势	15
2.1.2 C/C++ 语言在数字图像处理中的限制	16
2.2 Visual C++ 编程环境介绍	17
2.2.1 Microsoft Visual Studio 2008 简介	18
2.2.2 基于编程框架的数字图像处理工程建立示例	19
2.3 数字图像处理编程框架	26
2.3.1 数据结构	26
2.3.2 方案设计	27
2.3.3 框架实现	30
2.3.4 实验分析	33
2.4 编程框架中的具体类实现	37
2.4.1 CArray 类	39
2.4.2 CMatrix 类	41

2.4.3	CTArrayEx 类	43
2.4.4	CTMatrixEx 类	44
2.4.5	CImage 类	45
2.5	本章小结	48
2.6	本章习题	48
	参考文献	49
第 3 章	几何变换	51
3.1	简单几何变换	52
3.1.1	平移	53
3.1.2	旋转	56
3.1.3	镜像	58
3.1.4	转置	60
3.1.5	缩放	61
3.1.6	切变	63
3.2	复杂几何变换	66
3.2.1	仿射变换	67
3.2.2	投影变换	70
3.2.3	非刚体变换	72
3.3	本章小结	73
3.4	本章习题	73
	参考文献	74
第 4 章	图像分割	76
4.1	边缘提取	77
4.1.1	Robert 算子	80
4.1.2	Sobel 算子	81
4.1.3	Laplacian 算子	85
4.1.4	Prewitt 算子	86
4.1.5	Kirsch 算子	88
4.1.6	LoG 算子	89
4.1.7	Canny 算子	91
4.1.8	算子对比	100
4.2	区域分割	102
4.2.1	区域生长	102
4.2.2	分水岭算法	105
4.2.3	K-means 聚类算法	115
4.3	本章小结	121
4.4	本章习题	122
	参考文献	123
第 5 章	图像平滑与锐化	124

5.1	图像平滑	124
5.1.1	低通滤波法	125
5.1.2	高斯滤波法	135
5.1.3	均值滤波法	138
5.1.4	中值滤波法	140
5.2	图像锐化	144
5.2.1	高通滤波法	145
5.2.2	差值滤波法	148
5.3	本章小结	150
5.4	本章习题	150
	参考文献	151
第6章	形态学处理	153
6.1	基本运算	154
6.1.1	腐蚀	154
6.1.2	膨胀	156
6.1.3	开操作	158
6.1.4	闭操作	159
6.1.5	基本运算性质	160
6.2	复杂运算	160
6.2.1	细化算法	161
6.2.2	距离变换	175
6.2.3	击中与击不中变换	180
6.3	本章小结	186
6.4	本章习题	186
	参考文献	187
第7章	图像压缩与编码	189
7.1	基本理论概述	189
7.1.1	质量评价	190
7.1.2	无损压缩与有损压缩	192
7.1.3	预测编码与变换编码	192
7.1.4	离散余弦变换与小波变换	193
7.1.5	典型的图像压缩流程	194
7.2	无损压缩	197
7.2.1	熵编码	197
7.2.2	游程编码	202
7.2.3	差分脉冲编码调制	204
7.2.4	LZW 字典算法	206
7.3	有损压缩	209
7.3.1	离散余弦变换	210

7.3.2	色度抽样	213
7.3.3	向量量化	215
7.3.4	分形变换	216
7.4	JPEG 压缩标准	218
7.4.1	颜色空间转换	220
7.4.2	下采样	220
7.4.3	块分裂	220
7.4.4	离散余弦变换	220
7.4.5	量化	221
7.4.6	熵编码	222
7.5	JPEG 2000 压缩标准	223
7.5.1	瓦片拼贴与分量变换	224
7.5.2	小波变换与标量量化	224
7.5.3	块编码	225
7.5.4	分组与分层	225
7.6	本章小结	226
7.7	本章习题	226
	参考文献	227
第 8 章	图像特征提取	230
8.1	图像色彩与灰度	230
8.1.1	彩色信息处理	231
8.1.2	直方图处理	238
8.2	形状特征提取	245
8.2.1	Hu 不变矩	245
8.2.2	Zernike 不变矩	247
8.2.3	Hough 变换	250
8.2.4	Radon 变换	255
8.2.5	傅里叶描述子	260
8.3	纹理特征提取	263
8.3.1	纹理分析方法	263
8.3.2	用于纹理分析的频域变换	266
8.3.3	对应于人眼视觉感受的纹理特征	271
8.4	本章小结	277
8.5	习题	277
	参考文献	278
第 9 章	高级数字图像处理	281
9.1	光流分析	282
9.1.1	Horn & Schunck 算法	285
9.1.2	Lucas & Kanade 算法	290

9.1.3	多尺度块匹配算法	292
9.2	兴趣点检测与特征描述	299
9.2.1	角点检测	301
9.2.2	斑点检测	306
9.2.3	背脊检测	310
9.2.4	特征描述子	312
9.3	基于内容的图像检索	321
9.3.1	背景介绍	321
9.3.2	传统的检索技术	323
9.3.3	基于内容的检索技术	325
9.3.4	未来研究方向	327
9.3.5	图像检索实验系统	329
9.4	本章小结	334
9.5	本章习题	334
	参考文献	336
第 10 章	编程框架接口扩展	340
10.1	与 OpenCV 的接口	341
10.1.1	OpenCV 在 Visual C++ 2008 下的环境设置	342
10.1.2	具体接口实现	346
10.1.3	基于接口的编程实例	349
10.2	与 MATLAB 的接口	353
10.2.1	MATLAB 与 Visual C++ 2008 下的环境设置	354
10.2.2	具体接口实现	358
10.2.3	基于接口的编程实例	360
10.3	本章小结	363
10.4	本章习题	364
	参考文献	365

第1章 数字图像处理绪论

数字图像处理，就是对以数字信号形式出现的图像数据进行操作，以满足特定的应用或人眼视觉感受等要求的技术^[1]。所以说，数字图像处理属于数字信号处理在图像数据上的应用和特例。在过去，诸如图像和视频等多媒体数据都是以模拟的形式存储的，但是随着数字设备的发展和应用，现在大多数图像都是以数字的形式进行存储，在很多情况下图像处理就是指数字图像处理。在模拟形式时代，图像处理都是在光学设备中进行的，所采取的处理方法基于特定的光学理论，其本身具有良好的并行特性。在很多应用领域（如全息摄影等领域）中，这些光学方法也具有重要的作用。与模拟形式相比，数字形式下的图像处理在功能更加强大的计算机硬件条件的支撑下，更加的普及和容易重现。特别的是，针对数字图像处理的快速应用，推出了专用的硬件设备，包括基于流水线的计算机体系结构和基于 DSP 的图像处理专用芯片等。与这些专用硬件设备同步发展的是，很多研究和商用目的的数字图像处理任务都可以在个人电脑上以软件的形式来执行。随着摩尔定律的运用，很多过去看似算法复杂度很高的数字图像处理问题现在都能在很短的时间内得到解决，而且并行计算和网格计算技术的普及，使很多数字图像处理的问题都可以在并行计算的框架下得到解决，这就为数字图像处理技术的实时应用（如智能监控等）奠定了硬件基础。

图像处理的初次应用出现在 20 世纪 20 年代，得到普遍的重视则是在 20 世纪 60 年代中期，到了 60 年代末，图像处理已经发展为一门新兴的学科，其中电子计算机的大力发展起到了推波助澜的作用^[2]。早期的数字图像处理技术，出现在喷气推进实验室、麻省理工学院、贝尔实验室、马里兰大学和其他一些地方，而目前数字图像处理技术的研究和应用已经在全球范围内得到了推广。可以说，数字图像处理的应用几乎是无所不在的，例如图像修补、信息隐藏、图像压缩、目标识别等。借助于图像修补技术，可以将过去的破损纸质老照片恢复至原有的状态，也可以使得照片色彩更加符合人眼的视觉感受。基于图像的信息隐藏，就是将重要的信息（如密码）嵌入至图像本身的内容中，这样做的目的是使信息尽可能的隐秘，信息隐藏在版权保护和情报传递中有着极其重要的作用。图像压缩，现在已经极其普遍了，人们目前在互联网上所获取的图片几乎都是压缩格式的，这样可以在不损失太多视觉信息的条件下做到存储尺寸的极大缩减。目标识别，特别是生物特征识别，当前的发展非常迅速，其可以做到在很短的时间内从已有的数据库中搜寻到具有特定特征（如指纹和人脸特征等）的人。

信号处理，就是对信号进行有效的表示、变换和运算，并根据领域先验知识分析所包含的内容，其在计算机应用、电子学和药物分析等学科中都有着极其广泛的应用^[3]。所谓信号，就是在人与人、人与机器或者机器与机器之间进行联系的信息媒介，从狭义上讲可以是一个波形或者一个脉冲，从广义上讲包括所有物体的外在信息表现形式。图像处理属于信号处理的一个分支，但是其处理思路与普通的信号处理可能是大相径庭的，因为图像数据最本质的评价标准来自于人的视觉感知判断。通过多年的发展，信号处理技术已非常成熟，其应用也是遍及各行各业，包括石油勘探、机器人、数据通信、水下探测、地震预报等。信号处

理可以分为一维信号的处理、二维信号的处理和多维信号的处理，其中图像属于二维信号的范畴。除了二维信号的处理技术可以应用于图像处理以外，一维信号的处理技术同样也可以应用于图像数据，例如在语音（属于一维信号）识别领域中成熟应用的技术 HMM（隐马尔科夫模型），就已经成功地被应用到图像处理中。利用一维信号的处理技术进行图像处理有两种思路：一种思路是将图像分成多个子块，每个子块抽取一定的特征，将这些特征连接起来就构成了一维信号；另外一种思路，就是直接将诸如 HMM 的一维信号处理技术扩展至 2 维的情况，如 2 维 HMM 技术。当然，相对于一维信号的处理技术而言，图像处理本身也具有其独特的属性，如区域连通和旋转不变性等，这些属性是一维信号所不具备的。所以，虽然图像处理属于信号处理的范畴，但是其研究和应用还是具有很强的独特性。

1.1 数字图像处理目的

对数字图像处理的理解可以分为两种，一种是借助于专业的图像工具进行处理（面向终端用户）；另一种是使用某种计算机语言进行编程（面向专业编程）。专业图像工具很多，如 Windows 自带的画图程序，Adobe 公司的 Photoshop 软件，等等。这些工具能够完成非常多的数字图像处理功能，特别是一些特技处理。至于计算机编程，应该说，现有的语言，如 C++、Java 等，几乎都支持数字图像处理。上述两种理解，并不存在孰优孰劣的问题，仅是它们基于的侧重点不同而已，并且二者在很多应用中相辅相成。专业图像工具，侧重于现有的成熟的通用技术实现，并提供非常友好的交互界面；计算机编程，侧重于图像内在的表示和用户自定义的需求，适用于探索性和研究性的问题。所以，在非计算机领域，数字图像处理往往是指应用特定的工具来改变图像的显示；而在计算机领域，数字图像处理更侧重于计算机编程。当然，在实际工作中也常常使用专业的图像工具进行一些辅助性的工作，如查看一幅图像的哪个分量更符合用户的视觉感知等。

1.1.1 面向终端用户

Windows 自带的画图工具（Paint），是一个简单但极其实用的数字图像处理工具^[4]。借助 Paint 工具，用户可以绘制精美的图画。这些图画可以是黑白图像，也可以是彩色图像，并且可以保存为 BMP 图像格式。BMP 图像格式是 Windows 自带的图像格式，在 Windows 的平台下使用非常广泛，其与另一种图像格式 JPEG 的主要不同是 BMP 格式一般不带压缩。用户使用 Paint 工具绘制的图画，可以用做桌面背景，或者粘贴到其他的文档中。用户甚至可以使用 Paint 工具来浏览和编辑扫描的照片，并可以处理多种格式的图像，包括 JPEG、GIF 和 BMP 格式。Paint 工具最大的特点就是简单，并且不失一般通用的数字图像处理功能，但是其不会生成诸如边缘提取等特殊效果。使用 Paint 工具，用户可以完成如下任务：绘制直线；使用指定颜色进行区域填充；将指定图像设置为 Windows 桌面主题背景；显示栅格以进行像素信息的精细调整；处理从数码相机或扫描仪得来的图像；将图像作为电子邮件附件进行发送。

本书的作者，作为数字图像处理的专业人员，使用最多的图像处理工具就是 Paint 工具，当然作者更多的时间是花费在使用编程语言来实现自己所期望的图像效果，并且乐此不疲。这也是数字图像处理在专业编程方面给人所带来的极大乐趣和个人成就感。图 1-1 给出

了 Paint 工具的示意图, 其显示的是图像处理界的标准图像 Lena^[5]。关于 Lena 图像, 其来源一般不为人所知, 相应的出处可以在互联网上查找到, 这也是一个比较有趣的现象, 在学术界曾引起不小的争议。美国南加州大学图像处理研究所的 William K. Pratt 博士要找一幅图像来做图像压缩的测试, 就将 Lena 图像作为他研究使用的样例图像。自此以后, Lena 就渐渐成为了图像处理界的标准图像。可以说, 在图像处理界 Lena 图像就是学术界的标准。



图 1-1 Windows 自带的 Paint 工具

除了 Windows 自带的 Paint 工具, 另一款软件也得到了广泛的使用, 其用户既有非常专业的多媒体从业者也有业余的多媒体使用者, 这就是 Adobe Photoshop。Adobe Photoshop 是一款专业图像处理软件, 它由 Adobe Systems 公司出品^[6]。大多数用户所使用的 Photoshop 都是基于 Microsoft Windows 操作系统的, 但是该软件也有 Mac OS 和 UNIX 两种操作系统下的版本, 这点与该公司的其他软件是一样的。Photoshop 所具有的功能是极其强大的, 除了基本的图像编辑功能外, 它还提供了非常广泛的编辑和绘图工具, 如图层和滤镜功能等。图层功能可以方便用户进行效果测试, 滤镜功能可以使用户创造出各种奇妙的效果, 这远远胜过 Windows 自带的 Paint 工具。所以, 很多人认为 Photoshop 是最好的图像处理软件, 不过要完全掌握 Photoshop 中的奥妙则需要投入很多的学习时间。图 1-2 给出了 Adobe Photoshop 的一个应用窗口, 其中显示了两幅用于 MPEG-4 测试的标准视频中的图像截图, 左上的图像称为 Miss America, 右下的图像称为 Flowers。MPEG-4 是压缩编码标准 MPEG 中的一种, MPEG 的全称为 Motion Picture Experts Group, 其所处理的对象是音视频数据^[7], 而 MPEG-4 的目的则是为音视频数据的通信、存储和管理提供一个灵活和开放的框架和编码工具。除了 MPEG-4 以外, MPEG 标准还涉及 MPEG-1、MPEG-2、MPEG-7 和 MPEG-21。

另一款值得推荐的数字图像处理专业软件是 GIMP^[8]。GIMP 的全称为 GNU Image Manipulation Program, 中文意思为 GNU 图像处理程序。顾名思义, GIMP 是遵循 GNU GPL 协议的, 属于自由软件。很多情况下, 人们所见到的 GIMP 程序都是运行在 Linux 操作系统中的, 但是实际上, GIMP 也可以在 Microsoft Windows 和 Mac OS X 等操作系统中运行, 属于真正的跨平台软件。GIMP 的功能是极其强大的, 包括基本的图像处理功能和高级的照片润色和图像合成功能, 所以 GIMP 也被认为是 Photoshop 的代替品, 特别是在 Linux

操作系统中。但是在 Windows 平台上，GIMP 却是很少被使用的，这主要是与人们的长期习惯和 Windows 的窗口管理有关。

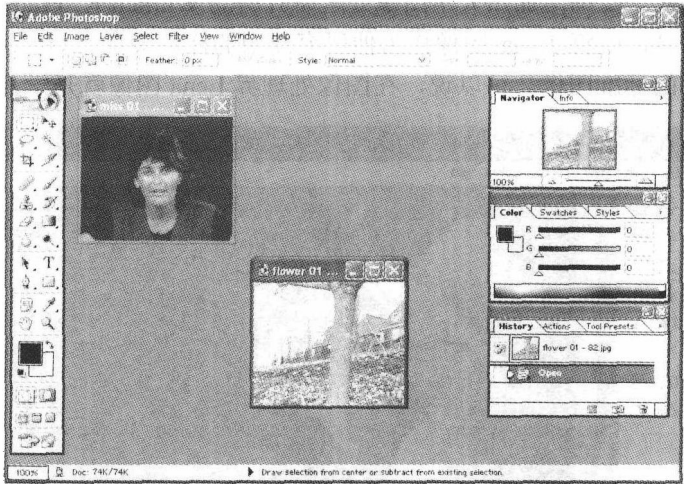


图 1-2 Adobe Photoshop 软件

1.1.2 面向专业编程

说到编程，肯定要涉及某种编程语言。目前编程语言很多，在选择上并不存在倾向性很强的指导性原则。有的时候往往基于程序员的个人喜好和编程习惯，其原因就是运用任意一种编程语言几乎都能实现所要的目标功能。当然，每种编程语言在形成的时候都会有自己特定的设计用途，使用与目标任务特性相匹配的编程语言会有事半功倍的效果。设想使用汇编语言去开发一个信息管理系统，肯定不会有人说一定不会成功，但也肯定不会有人实际去尝试。目前，对于数字图像处理来说，一般使用下述 3 种语言进行编程实现，包括 MATLAB^[9]、Java^[10]和 C++^[11]。

MATLAB 是一种商业数学软件，包括数值计算环境和编程语言，是由美国 MathWorks 公司出品的，其在数学类科技应用软件中属于国际控制界的标准数值计算软件^[12]。MATLAB 的名称，是从 MATrix LABoratory 而来的，从中可以看出 MATLAB 的基础与矩阵（Matrix）的计算是相关的，实际上也是这样的，例如图像数据就被自然地看做是一种二维矩阵。显然，对于 MATLAB 而言，图像数据仅仅是一种矩阵而已，更多的是数学意义上的概念，而很少涉及其中的语义特性。基于 MATLAB 的编程相对 C++语言来说要简单得多，其解决问题的方式类似于积木构建的样式，对于工程技术人员来说是非常容易理解和学习的。MATLAB 最为强大的部分是其所提供的工具箱。在 MATLAB 中，有两个工具箱，包括数字图像处理和数字信号处理，提供了非常强大的处理功能。往往一个数字图像处理算法，如果用 C++编写，可能需要上千行代码，而在 MATLAB 中只需要一个函数就可以实现，这都要归功于 MATLAB 所提供的强大的工具箱。所以，MATLAB 特别适合数字图像处理相关的算法设计，但是其执行效率很低，在实际应用中往往需要将 MATLAB 代码转换为 C++代码。鉴于 MATLAB 工具箱的强大，在本书的最后一章将专门介绍与 MATLAB 的接口实现。

Java 是一款真正意义上的面向对象的程序设计语言，其通过虚拟机的方式实现了跨平台

应用,而且随着互联网的迅速发展,它一度成为了首选的网络编程语言^[13]。Java 是由 SUN 公司(现已被 ORACLE 公司收购)出品的,作者主要为 James Gosling 等人,出品时间为 20 世纪 90 年代初。Java 语言在风格上与 C++语言是非常类似的,其将 C++语言中的面向对象技术发扬光大,并舍弃了一些 C++语言的特性,如指针、运算符重载、多重继承等。指针具有非常好的优点,但是容易引起错误,所以 Java 使用引用的概念来代替指针;对于多重继承,Java 则使用接口的概念来代替。使用 Java 语言编程,很多情况下是不需要关心内存是否释放这一问题的,因为其增加了垃圾回收器的功能,目的就是回收不再引用的对象所占据的内存空间。根据 SUN 公司对 Java 编程语言的解释,Java 语言是一种简单的、健壮的、安全的编程语言,其具有如下特性:面向对象、分布性、解释性、跨平台性、可移植性、高性能、多线程性等。Java 语言所在的平台称为 Java 平台,微软公司有个与之对应的平台——.NET 平台(其语言为 C#语言)。在解决数字图像处理问题方面,Java 和 C#应该说都没有太大的特色和优势,因为它们既没有 MATLAB 功能强大的工具箱,也没有 C++语言编程灵活和运行效率高的优点。当然,由于 Java 和 C#语言自身设计的特点,用户在使用这两种语言时能够很快上手,并且在编程时不需要过于严谨的编程风格,无须担心因为使用指针而导致的内存没有释放或者内存泄露等问题,这也是有人使用 Java 和 C#语言执行数字图像处理任务的原因,但这很难成为主流。

C++语言的使用是极其广泛的,曾一度成为计算机专业学生必用的编程语言,属于通用型的程序设计语言^[14]。C++的历史可以追溯至 20 世纪 80 年代,其由贝尔实验室发明,是作为 C 语言的增强型出现的。在不断发展的过程中,C++语言也添加了很多的特性,包括虚函数、运算符重载、多重继承、异常处理和命名空间等,并逐渐地被纳入国际标准。到 1998 年,ISO 颁布了 C++程序设计语言的国际标准 ISO/IEC 14882-1998。但是,随着更多新编程语言的推出,C++语言的使用比例总体上呈下降趋势。在 1998 年,以 C++为开发工具的人员的比例为 76%左右,而到了 2004 年,这一比例降为 46%,当然这与基于互联网的应用越来越多有关。有很多人认为,在众多的新编程语言中,Python 语言是其最为喜爱的一种,的确,Python 语言具有非常好的特性,而这些特性往往是 C++语言当初设计和后期发展中所不曾具有的。当然,在 C/C++阵营内部,也有很多不和谐的声音,很多 C 程序员就认为 C++有很多缺点,如编译器过于复杂,因此变得不太可靠,也就不适合构建至关重要的应用;也有人认为面向对象的技术不如面向过程的技术简单易学,在构建庞大的应用程序时,容易出现崩溃的现象,这点与 Java 语言类似。在数字图像处理领域,C++作为首选编程语言,得到了很多人的支持,原因就在于它能够非常高效地处理大内存数据,所编译程序的执行效率也是最高的,并且能够提供非常灵活的编程方式。当然,这种快的优点是基于 C++语言特性(指针)之上的,在编程中往往会表现为一种双刃剑的作用——快但不安全。本书会在后面的章节中详细地讨论这个问题,并试图去解决它。

Microsoft Visual C++,简称 VC++,由微软公司出品,是一个集成开发环境,提供了极其强大的编程功能,可以编译 C 和 C++等多种编程语言^[15]。在 VC++中进行编程,可以很轻松地使用 Windows 的 API 接口功能、DirectX API 功能和.NET 框架所提供的功能。VC++实际上是属于 Visual Studio 集成开发环境的一部分,目前的最新版本为 VC++ 2010。对于 VC++,微软公司提供了 4 种类型可供选择,分别为简易版(Express)、标准版(Standard)、专业版(Professional)和团队系统版(Team System),其中简易版(Express)为免费版本,

可以从 MSDN 的网站上直接下载，并且可以用于商业应用。由于 Microsoft Windows 操作系统的广泛使用（特别是在中国的广大院校中），所以如果在 Windows 平台上进行 C++ 开发，最理想的工具应该就是 VC++。事实上，本书就是基于 VC++ 来进行功能实现和源码组织的。相对 MATLAB，VC++ 似乎缺乏针对数字图像处理的专业开发包，这样对于执行大型的数字图像处理任务和进行专业的算法设计与应用将是一个很沉重的负担。幸运的是，在 C++ 语言的框架下，也存在多个数字图像处理相关的开发包或库，其中 OpenCV^[16] 就得到了非常广泛的应用。OpenCV 是由英特尔公司发起并参与开发的，是一种跨平台的、免费的、开源的计算机视觉开发库，其全称为 Open Source Computer Vision Library。OpenCV 所遵循的开源协议为 BSD，既可以用于研究目的，也可以用于商用目的。基于 OpenCV 的强大功能和易用性，本书在最后一章介绍了与 OpenCV 的接口，并以一些源码示例来说明。在 VC++ 平台下进行数字图像处理的开发时，同样能做到与 MATLAB 相似的很少源码编写与强大功能实现的完美结合。作者建议在从事计算机视觉的相关研究和开发时，要么使用 C++ 与 OpenCV 的结合，要么使用 MATLAB，因为它们能使开发者专注于自己的开发工作，并且能有效地进行算法验证和系统构建。

1.2 数字图像处理内容

数字图像处理允许对图像处理问题使用相对复杂的算法，这样对简单任务既提供了更好的性能，也可以实现通过模拟途径不可能完成的方法。对于数字图像处理所涉及的内容，不同的人有不同的视角和理解，正如本书上 1.1 节所阐述的面向终端用户和面向专业编程之间的区别。从涉及的特定任务出发，数字图像处理内容包括分类、特征提取、模式识别、投影和多尺度信号分析等；从所使用的技术手段出发，数字图像处理内容包括主成分分析、独立成分分析、自组织映射、隐马尔科夫模型和人工神经网络等。数字图像处理的最终目的是更好地辅助用户对图像进行分析，所以用户对数字图像处理结果的评价往往从人眼视觉感知出发，如基于内容图像检索的返回结果是否符合人眼的相似性标准。这样，就可以对数字图像处理内容从语义层次上进行分类，分别为底层图像处理技术、中层图像处理技术和高层图像处理技术。

1.2.1 底层图像处理技术

底层图像处理技术主要涉及底层图像特征，包括颜色、形状和纹理等。颜色是一种重要的视觉信息属性，相对于其他特征，颜色特征非常稳定，对于旋转、平移、尺度变化，甚至各种变形都不敏感，表现出相当强的鲁棒性，而且颜色特征计算简单。对彩色图像处理的研究，必须在特定的颜色空间中进行。实际应用中常用的颜色空间很多，有 RGB、YIQ 和 HSV 等。根据使用目的，常用的颜色空间可以分为两类，一类是面向硬设备的颜色空间，另一类是面向彩色处理的颜色空间，所谓硬设备，是指彩色显示器或打印机之类的设施。面向硬设备的最常用颜色空间是 RGB 颜色空间，而面向彩色处理的最常用颜色空间是 HSV 颜色空间。

纹理是与物体表面材质有关的图像特征，纹理分析一直是计算机视觉的一个重要研究方向，其方法主要分为结构方法和统计方法两类。结构方法假定图像由较小的纹理基元排列而

成，它采用句法分析方法，只适用于规则的结构纹理。统计方法又可进一步分为传统的统计方法、基于模型的方法以及基于频谱分析的方法。传统的统计分析方法始于 20 世纪 70 年代早期 Haralick 等人提出的基于二阶灰度统计特征的共生矩阵方法。20 世纪 80 年代开始，基于随机场模型的技术被用于纹理分析，多尺度的思想促进了基于频谱分析方法的发展，特别是基于 Gabor 滤波器的技术在纹理分析中被广泛采用。Gabor 滤波器可以通过调整获得图像不同方向和尺度的纹理信息。

形状的描述也是困难的问题，常用的方法有傅里叶描述子、矩不变量、各种简单的形状因子（如面积、圆度、偏心度、主轴方向）等。除了这些全局特征以外，有时也用一些局部特征（如直线段、圆弧、角点、高曲率点等），以解决遮挡问题。除这些常规方法外，人们还提出了许多基于不变性和变换的方法。由于基于简单特征的方法无法对形状作细致的匹配，也不能解决广泛存在的变形问题。因此实际系统中常常只用它们来作初步的过滤，最终的匹配结果由一些更复杂的方法给出，如变性模板、弹性匹配等。在很多情况下，人们感兴趣的并不是整幅图像，而是图像中的某些区域或目标，于是人们提出了基于区域或目标的图像表示方法，在这类方法中通常需要完成图像分割，然后再利用感兴趣部分的信息进行分析。

1.2.2 中层图像处理技术

顾名思义，中层图像处理技术主要涉及处理具有中层语义的图像元素。应该说，国内外对于中层语义的图像元素的理解也不尽相同，很多时候都是根据具体的应用和任务来定义。这种基于应用的图像理解往往能在特定的领域取得预期的效果，但是具有挑战性的却是通用性的图像理解。这里，作者倾向于将中层语义的图像元素理解为目标，主要涉及目标检测与识别。对于单帧图像而言，根据不同的应用领域，存在多种目标类型，例如人脸、指纹、手势、汽车、障碍物等，其中就数人脸识别^[17]得到了最为广泛的研究和应用。

人脸识别的概念可以分为广义的和狭义的。广义的人脸识别，包括涉及人脸识别系统的一系列技术，如数据采集、人脸定位、图像预处理、人脸验证和身份查询等；而狭义的人脸识别，是指根据人脸特征来确认当前人的身份。与其他生物特征识别技术一样，人脸识别技术属于非常热门的研究课题，既可以用于军事目的（如罪犯追踪），也可以用于一般的民用领域（如智能监控）。与其他生物特征识别技术不同的是，人脸识别具有非常好的特性，是非接触式的识别。例如，人脸识别不像指纹识别，指纹识别必须要求被识别人将手指接触至数据采集设备，这点会令很多人感到不舒服；而人脸识别，通过远处的摄像头就可以获取被识别人的人脸数据，这不但不会令被识别人不舒服，甚至不会引起被识别人的注意。当然，整容和浓妆对于人脸识别而言，也是一个很大的挑战。

虽然人脸识别具有很多的优点，而且在一些实际的应用中得到了发展，但是也面临着很多的挑战，其甚至被认为是生物特征识别领域中最困难的研究课题之一。应该说，人脸识别所面临的挑战，是由人脸自身的特点所决定的，有的人脸甚至使用人的肉眼都无法分辨，那么让计算机来识别就显得勉为其难了。人脸识别所面临的问题，与其他一些模式识别课题面临的问题一样，即类间差别小、类内差别大的问题。所谓类间差别小，是指不同人的脸型差别不大，不但结构相似，而且人脸器官的外形也是很相似的。而类内差别大，是指人脸是非刚体的，不但外形变化大，而且脸部表情极其丰富，并且随着观察角度的变化，人脸的成