



普通高等教育“十二五”规划教材

电子信息类精品教材

图像处理基础教程

Fundamental Course of Image Processing

• 章毓晋 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

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

普通高等教育“十二五”规划教材
电子信息类精品教材

图像处理基础教程

章毓晋 编著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书主要介绍对图像进行加工以改善其质量,并从中获取更多信息的一些基础概念、基本原理、典型方法、实用技术和应用成果。

本书主要内容分为10章。第1章是对图像技术基础和对全书的概括介绍,第2章介绍图像采集的模型以及坐标变换,第3章介绍图像空域增强技术,第4章介绍一些典型和常用的图像变换,第5章介绍图像频域增强技术,第6章介绍多类图像恢复技术,第7章介绍基本的图像编码技术,第8章介绍各种图像分割技术,第9章介绍目标表达和描述技术,第10章介绍特性分析和特征测量。附录A对图像国际标准进行了概括介绍。书中提供了大量例题和练习题,并对部分练习题提供了解答。书末除给出了相关的参考文献外还列出了主题名词索引(包括英文)。

本书可作为信号与信息处理、通信与信息系统、电子与通信工程、模式识别与智能系统、计算机视觉等学科的本科和研究生专业基础课教材,也可供信息与通信工程、电子科学与技术、计算机科学与技术、测控技术与仪器、机器人自动化、生物医学工程、光学、电子医疗设备研制、遥感、测绘和军事侦察等领域的师生和科技工作者参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

图像处理基础教程 / 章毓晋编著. —北京: 电子工业出版社, 2012.12

电子信息类精品教材

ISBN 978-7-121-18688-2

I. ①图… II. ①章… III. ①图像处理—高等学校—教材 IV. ①TP391.41

中国版本图书馆CIP数据核字(2012)第239919号

责任编辑: 韩同平 特约编辑: 李佩乾

印 刷: 三河市鑫金马印装有限公司

装 订: 三河市鑫金马印装有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路173信箱 邮编: 100036

开 本: 787×1092 1/16 印张: 18.25 字数: 470千字

印 次: 2012年12月第1次印刷

印 数: 3000册 定价: 39.80元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zltz@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

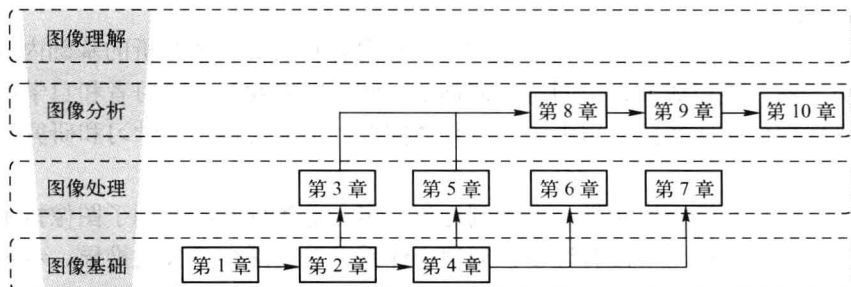
服务热线:(010)88258888。

前 言

本书是一部介绍图像处理和分析的基本原理、典型方法和实用技术的专门教材,希望为高等工科院校相关专业的本科或研究生开设第一门图像课程服务。

图像技术领域涉及范围很广,整个学科可统称为图像工程。图像工程所研究和应用的图像技术可划分到由低到高的三个层次中:图像处理、图像分析和图像理解。作为教材,尤其是为开设第一门图像课程服务,还需要在开始介绍一些图像基础内容。

本书共 10 章。它们所处的层次和相互的联系及学习的顺序如下图所示。



先将各章内容概括如下。

第 1 章是对图像处理的概括介绍以及从图像技术到图像工程学科的发展概述;对图像采集设备和图像打印中的半调和抖动输出技术也进行了介绍;还概括了本书的范围、主要内容及各章要点。

第 2 章介绍一些有关图像采集的基础知识,包括几何和亮度成像模型;讨论了数字图像的空间和幅度分辨率以及它们对视觉质量等的影响;介绍了像素的邻域,像素间的邻接、连接和连通,以及距离等概念;还介绍了坐标变换和形态变换的方式和方法。

第 3 章介绍图像空域增强技术,包括:直接利用像素灰度映射进行灰度变换的方法,利用图像间算术运算和逻辑运算的方法,利用图像直方图的统计信息对直方图进行变换修正的方法,以及一些典型的线性滤波技术和非线性滤波技术。

第 4 章介绍基本的图像变换,包括傅里叶变换、沃尔什变换和哈达玛变换,以及近年在图像编码中得到广泛使用的离散余弦变换;最后讨论了比较有特色的哈尔变换和在纹理图像分析中有优异性能的盖伯变换。

第 5 章介绍图像频域增强的技术,首先分析了频域增强的原理,并列出增强所涉及的各个步骤;然后依次介绍了低通滤波器,高通滤波器,带通和带阻滤波器,以及陷波滤波器和同态滤波器。

第 6 章介绍图像恢复技术,包括对几何失真的校正技术,将不同滤波器结合使用的方法,无约束恢复技术——逆滤波和有约束恢复技术——维纳滤波;最后对近年得到较多重视和应用的图像恢复新技术——图像修补给予了较全面的介绍。

第 7 章介绍图像编码技术,讨论了 3 种数据冗余形式,并给出了几种评判图像编解码质量的方法;介绍了典型的哈夫曼编码、算术编码,位平面分解等编码方法;最后对国际图像

压缩标准中应用最广泛的预测编码和变换编码技术进行了详细讨论。

第 8 章介绍图像分割技术,先结合图像分割的定义给出了一种有特色和通用的分割技术分类方法,将所有分割技术分成四大类(并行边界类、串行边界类、并行区域类和串行区域类),然后介绍了各类中一些近年得到较广泛应用的典型算法。

第 9 章介绍对分割后所提取出目标的表达和描述方法,结合对目标分割所采用的策略,介绍了两种表达方法:基于目标边界像素和目标区域像素的方法;也介绍了两种描述方法:基于边界的和基于区域的描述方法。

第 10 章介绍各种图像特征提取和测量技术,涉及到典型的颜色特征、纹理特征、形状特征和运动特征;最后介绍了影响特征测量精度的一些主要因素以及测量误差分析方面的一些内容。

另外,附录 A 介绍了各种图像国际标准的情况和所采用的技术,有些也涉及到图像理解的层次。

本书从选材上主要覆盖了狭义图像处理的主要内容和狭义图像分析的基础内容。本书基本针对信息类专业,能自成体系,同时也兼顾了一些不同专业背景学习者和自学读者学习的需要。读者既能据此解决实际应用中的一些具体问题,也能为进一步学习和研究高层技术打下基础。

本书从编写上比较注重实用性。考虑到近年许多专业领域都涉及到了图像技术,但并非专门研究图像技术,所以本书没有过多强调理论体系,尽量减少了公式推导,着重介绍了常用的方法。本书有较多的例题,通过直观的解释来帮助读者理解抽象的概念。书末给出了术语索引(文中标为黑体),对每个术语均在索引中给出了对应的英文,方便读者对本书的查阅,也方便读者联网搜索相关资料。

本书从结构上看,共有 10 章正文,1 个附录以及部分练习题解答、参考文献和索引。在这 14 个一级标题下共有 71 个二级标题(节),再之下还有 158 个三级标题(小节)。全书共有文字(也包括图片、绘图、表格、公式等)约 50 万字,共有编了号的图 292 个、表格 31 个、公式 572 个。为便于教学和理解,本书共给出各类例题 103 个,练习题 120 个,并对其中的 20 个练习题提供了参考答案。另外书末列出了直接相关的约 100 篇参考文献的目录和用于索引的 600 多个术语(中英文)。

本书从教学内容上可分为基本内容和选学内容(在对应节号上加了星号)。如果只使用基本内容,应可作为一门 32 学时基本课程的教材;如果包括选学内容,应可作为一门 48 学时扩充课程的教材。两种情况下对各章建议的学时数如下表所示:

学时	第 1 章	第 2 章	第 3 章	第 4 章	第 5 章	第 6 章	第 7 章	第 8 章	第 9 章	第 10 章	附录 A	复习
32	2	2	3	3	3	2	4	3	4	4		2
48	2	5	4	4	4	5	5	5	4	5	2	3

本书从先修课程知识要求看一般考虑 3 个方面:①数学:包括线性代数和矩阵理论,以及有关统计学、概率论和随机建模的基础知识。②计算机科学:包括对计算机软件技术的掌握,对计算机结构体系的理解,以及对计算机编程方法的应用。③电子学:一方面电子设备的特性原理;另一方面电路设计等内容。另外,学习本书最好在学习完有关信号处理的课程后进行。

本书的电子教案可通过华信教育资源网 <http://www.hxedu.com.cn> 或本书作者主页 <http://oa.ee.tsinghua.edu.cn/~zhangyujin/> 下载。

感谢出版社编辑的精心组稿，认真审阅和细心修改。

最后，作者感谢妻子何芸、女儿章荷铭以及父母家人在各方面的理解和支持。

章毓晋

2012 年暑假于清华大学

通信：北京清华大学电子工程系（100084）

办公：清华大学，罗姆楼，6 层 305 室

电话：(010) 62798540

传真：(010) 62770317

邮箱：zhang-yj@tsinghua.edu.cn

主页：oa.ee.tsinghua.edu.cn/~zhangyujin/

目 录

第 1 章 绪论	(1)	2.3.1 光度学	(30)
1.1 图像基础	(1)	2.3.2 图像亮度成像模型	(33)
1.1.1 图像表示和显示	(1)	2.4 图像分辨率与质量	(34)
1.1.2 空间分辨率和幅度分辨率	(4)	2.4.1 空间分辨率的影响	(34)
1.2 图像工程	(5)	2.4.2 灰度分辨率的影响	(35)
1.2.1 图像技术和图像工程	(5)	2.4.3 空间和灰度分辨率的影响	(35)
1.2.2 图像工程的 3 个层次	(6)	2.5 像素间联系	(36)
1.2.3 图像工程相关学科和领域	(7)	2.5.1 像素邻域	(36)
1.2.4 图像工程的技术应用	(8)	2.5.2 像素间距离	(37)
1.2.5 图像工程文献统计分类	(8)	2.6 图像坐标变换	(39)
1.3 图像采集设备	(10)	2.6.1 基本坐标变换	(39)
1.3.1 采集设备性能指标	(10)	2.6.2 坐标变换扩展	(41)
1.3.2 固态采集器件	(11)	*2.7 图像仿射变换	(43)
1.4 图像半调和抖动输出	(12)	2.7.1 一般仿射变换	(43)
1.4.1 半调输出	(12)	2.7.2 特殊仿射变换	(47)
1.4.2 抖动输出	(14)	2.7.3 变换间的联系	(48)
1.5 图像存储	(15)	练习题	(49)
1.5.1 图像存储器	(15)	第 3 章 图像空域增强	(51)
1.5.2 图像文件格式	(16)	3.1 增强技术分类	(51)
1.6 关于本书	(17)	3.2 灰度映射	(52)
1.6.1 整体框架和各章概述	(18)	3.2.1 灰度映射原理	(52)
1.6.2 编写特点	(19)	3.2.2 灰度映射示例	(53)
1.6.3 先修基础	(20)	3.3 图像间运算	(55)
练习题	(20)	3.3.1 算术运算	(55)
第 2 章 图像采集	(22)	3.3.2 逻辑运算	(57)
2.1 成像和变换	(22)	3.4 直方图修正	(59)
2.2 几何成像模型	(22)	3.4.1 直方图和累积直方图	(59)
2.2.1 投影成像几何	(23)	3.4.2 直方图均衡化	(60)
2.2.2 基本几何成像模型	(24)	3.4.3 直方图规定化	(62)
2.2.3 一般几何成像模型	(26)	*3.5 空域滤波	(65)
2.2.4 通用几何成像模型	(28)	3.5.1 原理和分类	(65)
2.3 亮度成像模型	(30)	3.5.2 线性平滑滤波器	(67)

3.5.3 线性锐化滤波器	(68)	5.4.3 带通和带阻滤波器的联系	(104)
3.5.4 非线性平滑滤波器	(69)	5.5 陷波滤波器	(105)
3.5.5 非线性锐化滤波器	(71)	*5.6 同态滤波器	(107)
练习题	(72)	*5.7 空域技术与频域技术	(109)
第4章 图像变换	(74)	练习题	(111)
4.1 变换特性	(74)	第6章 图像恢复	(113)
4.1.1 变换核	(74)	6.1 图像退化	(113)
4.1.2 可分离性	(74)	6.1.1 图像退化示例	(113)
4.1.3 对称性	(75)	6.1.2 基本退化模型	(114)
4.1.4 正交性	(75)	6.1.3 退化和恢复	(115)
4.2 傅里叶变换	(76)	6.2 几何失真校正	(115)
4.2.1 2-D 傅里叶变换	(76)	6.2.1 空间变换	(116)
4.2.2 傅里叶变换定理	(78)	6.2.2 灰度插值	(117)
4.3 沃尔什变换和哈达玛变换	(81)	6.3 噪声滤波器	(118)
4.3.1 沃尔什变换	(81)	6.3.1 噪声及来源	(118)
4.3.2 哈达玛变换	(83)	6.3.2 噪声概率密度函数	(120)
4.3.3 两种变换的联系	(84)	6.3.3 均值滤波器	(122)
4.4 离散余弦变换	(86)	6.3.4 排序统计滤波器	(124)
4.4.1 变换定义	(86)	6.4 组合滤波器	(125)
4.4.2 变换计算	(87)	6.4.1 混合滤波器	(125)
4.5 哈尔变换	(88)	6.4.2 选择性滤波器	(126)
*4.6 盖伯变换	(90)	*6.5 逆滤波	(128)
4.6.1 窗函数	(90)	6.5.1 逆滤波原理	(128)
4.6.2 短时傅里叶变换	(91)	6.5.2 消除匀速直线运动模糊	(129)
4.6.3 连续盖伯变换	(91)	*6.6 维纳滤波	(130)
4.6.4 离散盖伯表达	(92)	*6.7 图像修补	(131)
练习题	(94)	6.7.1 图像修补原理	(132)
第5章 图像频域增强	(95)	6.7.2 全变分模型修复	(133)
5.1 频域增强原理	(95)	6.7.3 混合模型修复	(134)
5.2 低通滤波器	(96)	6.7.4 基于样本的补全	(135)
5.2.1 理想低通滤波器	(96)	6.7.5 结合稀疏表达的补全	(137)
5.2.2 实用低通滤波器	(98)	练习题	(140)
5.3 高通滤波器	(100)	第7章 图像编码	(142)
5.3.1 基本高通滤波器	(100)	7.1 图像压缩原理	(142)
5.3.2 特殊高通滤波器	(101)	7.2 数据冗余	(143)
5.4 带通和带阻滤波器	(103)	7.2.1 像素相关冗余	(143)
5.4.1 带通滤波器	(103)	7.2.2 编码冗余	(144)
5.4.2 带阻滤波器	(104)	7.2.3 心理视觉冗余	(145)

7.2.4 编解码器模型	(145)	8.4.3 局部阈值的选取	(186)
7.3 编码定理	(146)	8.4.4 动态阈值的选取	(189)
7.3.1 信息和信源描述	(146)	8.5 串行区域技术	(190)
7.3.2 无失真编码定理	(147)	8.5.1 区域生长	(190)
7.4 图像保真度	(148)	8.5.2 分裂合并	(191)
7.4.1 客观保真度准则	(148)	*8.6 SUSAN 检测算子	(192)
7.4.2 主观保真度准则	(149)	8.6.1 USAN 原理	(192)
7.5 变长编码	(150)	8.6.2 SUSAN 算子边缘检测	(193)
7.5.1 哈夫曼编码	(150)	*8.7 主动轮廓模型	(196)
7.5.2 算术编码	(153)	8.7.1 主动轮廓	(196)
7.5.3 香农-法诺编码	(155)	8.7.2 能量函数	(196)
*7.6 位平面编码	(156)	*8.8 过渡区分割法	(199)
7.6.1 位面分解	(156)	8.8.1 有效平均梯度	(199)
7.6.2 位面编码	(158)	8.8.2 过渡区边界	(200)
7.7 预测编码	(160)	8.8.3 利用过渡区取阈值	(201)
7.7.1 无损预测编码	(160)	练习题	(202)
7.7.2 有损预测编码	(162)	第9章 目标表达和描述	(203)
7.8 变换编码	(165)	9.1 表达与描述	(203)
7.8.1 基于 DCT 的变换编码	(166)	9.2 目标标记	(203)
7.8.2 子图像尺寸选择	(166)	9.2.1 像素标记	(203)
7.8.3 变换选择	(167)	9.2.2 游程连通性分析	(204)
7.8.4 比特分配	(169)	9.3 基于边界的表达	(205)
练习题	(171)	9.3.1 边界表达分类	(205)
第8章 图像分割	(173)	9.3.2 链码	(206)
8.1 定义和技术分类	(173)	9.3.3 边界段和凸包	(207)
8.1.1 图像分割定义	(173)	9.3.4 边界标记	(208)
8.1.2 图像分割技术分类	(174)	9.3.5 多边形	(209)
8.2 并行边界技术	(174)	9.3.6 地标点	(210)
8.2.1 边缘及检测原理	(174)	9.4 基于区域的表达	(211)
8.2.2 一阶导数算子	(175)	9.4.1 区域表达分类	(211)
8.2.3 二阶导数算子	(177)	9.4.2 空间占有数组	(211)
8.2.4 边界闭合	(179)	9.4.3 四叉树	(211)
8.3 串行边界技术	(180)	9.4.4 金字塔	(212)
8.3.1 图搜索	(180)	9.4.5 围绕区域	(213)
8.3.2 动态规划	(181)	9.4.6 骨架	(214)
8.4 并行区域技术	(183)	9.5 基于边界的描述	(216)
8.4.1 原理和分类	(183)	9.5.1 简单边界描述符	(216)
8.4.2 全局阈值的选取	(184)	9.5.2 形状数	(217)

9.5.3 边界矩	(218)	10.4.2 形状复杂性	(240)
9.6 基于区域的描述	(218)	10.5 运动特征测量	(242)
9.6.1 简单区域描述符	(218)	10.5.1 基于图像求差的运动检测 ...	(242)
9.6.2 拓扑描述符	(220)	10.5.2 基于模型的运动检测	(245)
9.6.3 区域不变矩	(220)	*10.6 特征测量的准确度	(247)
练习题	(222)	10.6.1 准确度和精确度	(247)
第 10 章 特征提取和测量	(224)	10.6.2 影响测量准确度的因素	(248)
10.1 特性分析	(224)	10.6.3 直线长度测量	(249)
10.2 颜色特征测量	(224)	练习题	(250)
10.2.1 面向硬设备的颜色模型	(225)	附录 A 图像国际标准	(251)
10.2.2 面向视觉感知的颜色模型 ...	(226)	A.1 国际标准	(251)
10.2.3 直方图特征	(229)	A.2 二值图像压缩国际标准	(252)
10.2.4 颜色布局特征	(230)	A.3 静止图像压缩国际标准	(253)
10.3 纹理特征测量	(231)	A.4 运动图像压缩国际标准	(256)
10.3.1 统计法	(231)	A.5 多媒体国际标准	(262)
10.3.2 结构法	(234)	部分练习题解答	(265)
10.3.3 频谱法	(235)	参考文献	(268)
10.4 形状特征测量	(238)	索引	(272)
10.4.1 形状紧凑性	(238)		

第1章 绪 论

本书是一本介绍基本图像技术的教材。

本章将对相关基础知识给予概括介绍,包括对图像、像素、矩阵表达、图像显示等基本概念的介绍(见 1.1 节),对整体图像技术框架和相关学科的介绍(见 1.2 节),对常见图像获取设备的介绍(见 1.3 节),以及对图像打印输出方式和技术的介绍(见 1.4 节)。最后,还对全书的整体内容和各章的概况等给予概述(见 1.5 节)。

1.1 图 像 基 础

人们一般将图像看作对场景或景物的一种可视表现形式。例如,字典里对**图像**的一个定义是“对物体的表达、表象、模仿,一个生动的视觉描述,为了表达其他事物而引入的事物”[Bow 2002]。严格一点说,图像是用各种观测系统以不同形式和手段观测客观世界而获得的,可以直接或间接作用于人眼并进而产生视知觉的实体[章 1996a]。人的视觉系统就是一个观测系统,通过它得到的图像就是客观景物在人心目中形成的影像。

图像含有丰富的信息。我们生活在一个信息时代,科学研究和统计表明,人类从外界获得的信息约有 75%来自视觉系统,也就是从图像中获得的。这里图像的概念是比较广义的,包括照片、绘图、动画、视频,甚至文档等。中国有句古话,“百闻不如一见”。人们常说,“一图值千字”。这些都说明图像中所含的信息内容非常丰富,是我们最主要的信息源。

1.1.1 图像表示和显示

本书主要讨论对自然场景成像得到的图像,不考虑人或计算机绘制出来的人工图形等。这样的图像也包含许多类别。例如用数码相机拍摄的照片(人,风景等),用数码摄像机拍摄的视频(家庭联欢、足球比赛等),用监控系统记录下的各种序列(交通管理、导弹飞行等),以及用太空望远镜摄取的各种电磁辐射图像,用雷达依靠反射波形成的回波图像,医学上常使用的 X 射线图像、B 超图像、CT 图像、磁共振图像(MRI)等。不仅有灰度的、彩色的图像,还可以有纹理的、深度的图像等。下面先来介绍如何表示和显示图像。

1. 图像和像素

客观世界在空间上是三维(3-D)的,但一般从客观景物得到的图像是二维(2-D)的。一幅图像可以用一个 2-D 数组 $f(x, y)$ 来表示,这里 x 和 y 表示 2-D 空间 XY 中一个坐标点的位置,而 f 则代表图像在点 (x, y) 的某种性质 F 的数值。实用中, $f(x, y)$ 有时代表的是一幅图像,但有时也可以指在位置 (x, y) 的点的性质值。这一般根据上下文就可以区别。

性质 F 的含义常取决于图像的类别。例如, **灰度图像**中 f 表示灰度值,它常对应客观景物被观察到的亮度。文本图像常为**二值图像**, f 的取值只有两个,分别对应文字和空白。图像在点 (x, y) 也可同时具有多种性质,此时可用矢量 f 来表示。例如一幅**彩色图像**在每一个图

像点同时具有红、绿、蓝 3 个值，可记为 $[f_r(x, y), f_g(x, y), f_b(x, y)]$ 。需要指出，人们总是根据图像内不同位置所具有的不同性质来利用图像的。

例 1.1.1 一般的图像表达函数

图像可表示一种辐射能量的空间分布，这种分布可以是 5 个变量的函数， $T(x, y, z, t, \lambda)$ ，其中 x, y, z 是空间变量， t 代表时间变量， λ 是波长(对应频谱变量)。例如，红色的物体反射波长在 $0.57 \sim 0.70 \mu\text{m}$ 的光并吸收几乎所有其他波长的能量，绿色的物体反射波长在 $0.48 \sim 0.57 \mu\text{m}$ 的光，蓝色的物体反射波长在 $0.40 \sim 0.48 \mu\text{m}$ 的光，紫外(色)的物体反射波长在 $0.25 \sim 0.40 \mu\text{m}$ 的光，红外(色)的物体反射波长在 $0.9 \sim 1.5 \mu\text{m}$ 的光。它们合起来所覆盖的波长为 $0.25 \mu\text{m} < \lambda < 1.5 \mu\text{m}$ 。由于实际图像在时空上都是有限的，所以 $T(x, y, z, t, \lambda)$ 是一个 5-D 有限函数。□

早年获取的图像多是连续(模拟)的，即 f, x, y 的值可以是任意实数。进入 21 世纪后，所获取的图像基本为离散(数字)的，可直接用计算机进行加工。曾有人用 $I(r, c)$ 来表示数字图像，其中 I, c, r 的值都是整数。这里 I 代表离散化后的 f ， (r, c) 代表离散化后的 (x, y) ，其中 r 代表图像的行(row)， c 代表图像的列(column)。本书讨论的都是数字图像，在不至引起混淆的情况下均使用图像或 $f(x, y)$ 代表数字图像，如不做特别说明，这里 f, x, y 都在整数集合中取值的。

早期英文书籍里一般用 picture 来指图片，随着数字技术的发展，现都用 image 代表离散化了的“图象”，因为“计算机存储人像或场景的数字图象(computers store numerical images of a picture or scene)”[Zhang 1996]。这样看来，应该使用“数字图象”而不是“数字图像”。事实上“图象”比图像的含义更广，覆盖面更宽。图像中每个基本单元叫做图像元素，简称为像素。对 2-D 图像，英文里常用简称 pixel 代表像素(也有用 pel 的)。如果采集一系列的 2-D 图像或利用一些特殊设备还可得到 3-D 图像。对 3-D 图像，英文里常用 voxel 代表其基本单元，简称体素。

2. 图像的矩阵和矢量表示

一幅 $M \times N$ (其中 M 和 N 分别为图像的总行数和总列数) 的 2-D 图像既可以用一个 2-D 数组 $f(x, y)$ 来表示，也可用一个 2-D 矩阵 F 来表示：

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{MN} \end{bmatrix} \quad (1.1.1)$$

上述矩阵表达也可以化成矢量表达，例如上式可写成

$$F = [f_1 \quad f_2 \quad \cdots \quad f_N] \quad (1.1.2)$$

其中

$$f_i = [f_{1i} \quad f_{2i} \quad \cdots \quad f_{Mi}]^T \quad i = 1, 2, \dots, N \quad (1.1.3)$$

需要注意数组运算和矩阵运算是不同的。以两幅 2×2 的图像 $f(x, y)$ 和 $g(x, y)$ 为例，它们的数组积为

注：□为例题结束符，全书同。

$$f(x,y)g(x,y) = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} f_{11}g_{11} & f_{12}g_{12} \\ f_{21}g_{21} & f_{22}g_{22} \end{bmatrix} \quad (1.1.4)$$

而它们的矩阵积为

$$FG = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} f_{11}g_{11} + f_{12}g_{21} & f_{12}g_{12} + f_{12}g_{22} \\ f_{21}g_{11} + f_{22}g_{21} & f_{21}g_{12} + f_{22}g_{22} \end{bmatrix} \quad (1.1.5)$$

3. 图像的显示方式

对 2-D 图像的显示可以采取多种形式,其基本思路是将 2-D 图像看作在 2-D 空间中的一种幅度分布。根据图像的不同,采取的显示方式也可不同。

例 1.1.2 灰度图像显示示例

图 1.1.1 (a) 和 (b) 所示为两幅典型的灰度图像 (Lena 和 Cameraman)。图 (a) 的坐标系统常在屏幕显示中采用 (屏幕扫描是从左向右、从上向下进行的), 它的原点 (origin) O 在图像的左上角, 纵轴标记图像的列, 横轴标记图像的行。图 (b) 的坐标系统常在图像计算中采用, 它的原点在图像的左下角, 横轴为 X 轴, 纵轴为 Y 轴 (与常用的笛卡儿坐标系相同)。图 1.1.1 (c) 给出对图 (a) 的一个 3-D 透视显示, 其中将各个像素的灰度显示在对应的垂直高度上。

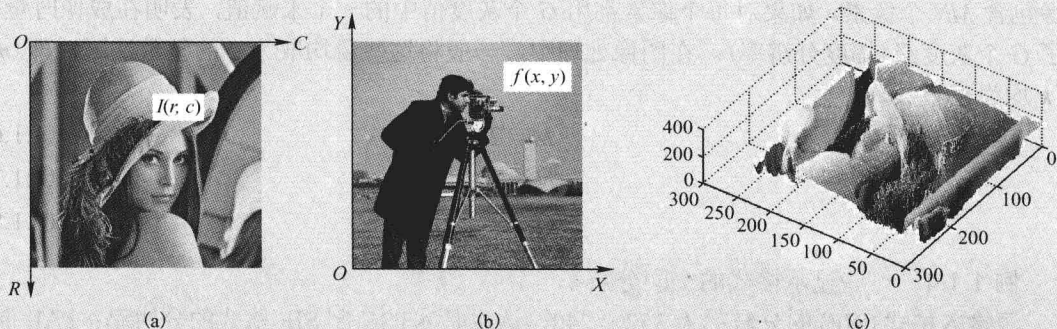
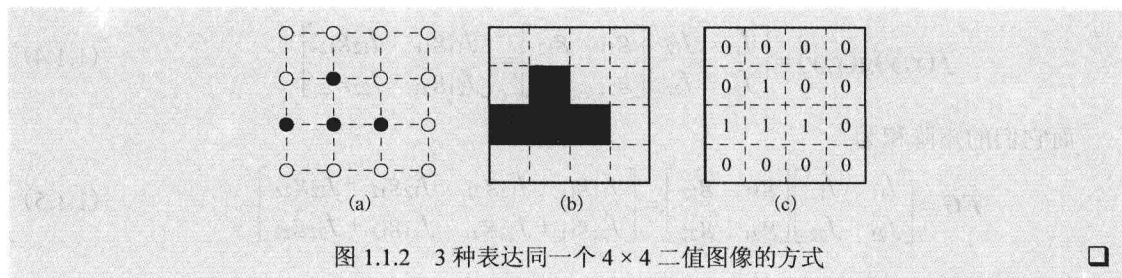


图 1.1.1 灰度图像显示示例

顺便指出, 常见的彩色图像可分解为 3 幅灰度图像 (参见 10.2 节), 若把对同一场景用不同波长的传感器获得的 3 幅灰度图像同时输入特定的显示装置, 就可获得一幅彩色图像。

例 1.1.3 二值图像表示示例

二值图像可看作特殊 (仅有两个灰度) 的灰度图像 (见 7.6 节)。图 1.1.2 给出了对同一个二值图像矩阵的 3 种不同的可视表达方式。在图像表达的数学模型中, 一个像素区域常用其中心来表示, 这样得到的表达形式就是平面上的离散点集, 对应图 (a)。如果将像素区域用填充方块来表示, 就得到图 (b)。若把幅度值标在图像中相应的位置, 就得到如图 (c) 所示的类似方阵的结果。用图 (b) 的形式还可表示有多个灰度的图像, 此时需要用不同深浅的色调表示不同的灰度。用图 (c) 的形式也可表示有多个灰度的图像, 此时将不同灰度用不同的数值表示。



1.1.2 空间分辨率和幅度分辨率

一幅图像必须要在空间和灰度上都离散化后才能被计算机处理。空间坐标的离散化称为空间采样(简称采样),它确定了图像的空间分辨率;而灰度值的离散化称为灰度量化(简称量化),它确定了图像的幅度分辨率。它们对图像视觉质量的影响将在 2.4 节讨论。这里仅考虑空间分辨率和幅度分辨率与图像数据量的关系。

设 X 、 Y 和 F 均为实整数集。采样过程可看作将图像平面划分成规则网格,每个网格中心点的位置由一对笛卡儿坐标 (x, y) 所决定,其中 x 是 X 中的整数, y 是 Y 中的整数。令 $f(\cdot)$ 是给点对 (x, y) 赋予灰度值 (f 是 F 中的整数) 的函数,那么 $f(x, y)$ 就是一幅图像,而这个赋值过程就是量化过程。

如果一幅图像的尺寸(空间分辨率)为 $M \times N$,表明在成像时采集了 MN 个样本,或者说图像包含 MN 个像素。如果对每个像素都用 G 个灰度值中的一个来赋值,表明在成像时量化成了 G 个灰度级(幅度分辨率)。在图像处理中,一般将这些量均取为 2 的整数次幂,即 (m 、 n 、 k 均为正整数)

$$M = 2^m \quad (1.1.6)$$

$$N = 2^n \quad (1.1.7)$$

$$G = 2^k \quad (1.1.8)$$

例 1.1.4 一些显示格式的空间分辨率

源输入格式(SIF)的分辨率为 352×240 ,这也是 NTSC 制 SIF 格式的分辨率;PAL 制 SIF 格式的分辨率为 352×288 ,这也是 CIF 格式的分辨率;QCIF 格式的分辨率为 176×144 ;VGA 的分辨率为 640×480 ;CCIR/ITU-R 601 的分辨率为 720×480 (NTSC) 或 720×576 (PAL);而 HDTV 的分辨率可达 1440×1152 甚至 1920×1152 。

存储一幅图像所需的数据量由图像的空间分辨率和幅度分辨率决定。根据式 (1.1.6) ~ 式 (1.1.8),存储一幅图像所需的位数 b (单位是比特, bit) 为

$$b = M \times N \times k \quad (1.1.9)$$

如果 $N = M$ (以下一般都设 $N = M$), 则

$$b = N^2 k \quad (1.1.10)$$

例 1.1.5 分辨率与存储和处理

存储一幅图像所需的比特数常很大。假设有一幅 512×512 , 256 个灰度级的图像,它需要用 2 097 152 比特来存储。一个字节是 8 个比特,为表示 256 个灰度级需用一个字节(即用一个字节表示一个像素的灰度),这样上面的图像需要 262 144B 来存储。如果彩色图像的分辨率为 1024×1024 ,则需要 3.15MB 来存储,这相当于存储一本 750 页的书。视频是

由连续的图像帧组成的(PAL制为每秒25帧)。假设彩色视频的每帧图像为 512×512 ,则一秒钟的数据量为 $512 \times 512 \times 8 \times 3 \times 25$ 比特或19.66MB。

为实时处理每帧图像为 1024×1024 的彩色视频,需要每秒处理 $1024 \times 1024 \times 8 \times 3 \times 25$ 比特的数据,对应的处理速度要达到约每秒78.64MB。如果设对一个像素的处理需要10个浮点运算(floating-point operations, FLOPS),则对一秒钟视频的处理就需要近8亿个浮点运算。并行运算策略通过利用多个处理器同时工作来加快处理速度。最乐观的估计认为并行运算的时间可减少为串行运算的 $1/J$,其中 J 为并行处理器的个数。按照这种估计,如果使用100万个并行处理器来处理一秒钟的视频,每个处理器还要具有每秒78万多次运算的能力。□

1.2 图像工程

对图像的采集和加工技术近年来得到极大的重视和长足的进展,出现了许多有关的新理论、新方法、新算法、新手段和新设备。图像已在科学研究、工业生产、医疗卫生、教育、娱乐、管理和通信等方面得到了广泛的应用,对推动社会发展、改善人们生活水平都起到了重要的作用。

为了对各种图像技术进行综合研究、集成应用,有必要建立一个整体框架——**图像工程(IE)**。图像工程是一门系统地研究各种图像理论、技术和应用的新的交叉学科。它的研究方法 与数学、物理学、生理学、心理学、电子学、计算机科学等学科相互借鉴,它的研究范围与模式识别、计算机视觉、计算机图形学等专业互相交叉,它的研究进展与人工智能、神经网络、遗传算法、模糊逻辑等理论和技术密切相关,它的发展应用与生物医学、遥感、通信、文档处理等许多领域紧密结合。

1.2.1 图像技术和图像工程

尽管图像技术的历史可追溯到1946年世界上第一台电子计算机的诞生(借助打印设备对图片编码获得数字图像甚至可追溯到20世纪20年代[Gonzalez 2008]),但在20世纪50年代计算机主要还是用于数值计算,满足不了处理大数据量图像的要求。20世纪60年代,第三代计算机的研制成功,以及快速傅里叶变换算法的发现和应用使得对图像处理的某些计算得以实际实现,人们从而逐步开始利用计算机对图像进行加工利用。20世纪70年代,图像技术有了长足的进展,图像分析借助图像测量和统计模式识别独立出来,而且第一本重要的图像技术书籍[Rosenfeld 1976]也得以出版,对图像的计算方法不再与传感器捆绑在一起并更为通用。20世纪80年代,各种硬件和系统的研制成功使得人们不仅能处理2-D图像,而且能获取3-D图像并进行处理分析,同时基于模型的图像技术也得到了广泛的重视,图像理解在几何建模和人工智能等的辅助下渐露头角。20世纪90年代,网络的发展进一步提升了人们对各种图像应用的兴趣,当时得到广为宣传的多媒体中图像其实占据了最主要的地位。广义上来说,文本、图形、视频等都属于图像,也都需要借助图像技术才能充分获得利用。进入21世纪的第一个十年,立体图像的广泛采集、认知理论的迅速发展等使图像技术影响到了人类生活和社会发展的各个方面。展望21世纪20年代,图像技术将得到更迅猛的发展,其应用将渗透到各个领域,并在改变人们的生活方式及社会结构等方面都将起到更重要的作用。

图像技术在广义上是各种与图像有关的技术的总称。本书主要讨论计算机图像技术,包

括利用计算机和其他电子设备进行的一系列计算工作，例如图像的采集、获取、编码、存储和传输，图像的合成和产生，图像水印的嵌入和提取，图像的显示和输出，图像的变换、增强、恢复、修复和投影重建，图像的分割，目标的检测、跟踪、表达和描述，目标特征的提取和测量，序列图像的校正，3-D 景物的重建复原，图像数据库的建立、索引和抽取，图像的分类、表示和识别，图像模型的建立和匹配，图像和场景的解释和理解，以及基于它们的判断决策和行为规划等(相关名词定义可见[章 2009b])。另外，图像技术还可包括为完成上述功能而进行的硬件设计及制作等方面的技术。随着人们研究的深入和应用的广泛，已有的图像技术在不断更新和扩展，许多新的图像技术也在不断诞生。对各种典型的图像技术的原理和方法的介绍是本书的主要内容。

由于图像技术近年来得到极大的重视和长足的进展，出现了许多新理论、新方法、新算法、新手段、新设备、新应用。对各种图像技术进行综合集成的研究和应用应当在一个整体框架下进行，这个框架就是**图像工程**[章 1996b]。众所周知，工程是指将自然科学的原理应用到工业部门而形成的各学科的总称。图像工程学科则是一个将数学、光学等基础科学的原理结合在图像应用中积累的经验而发展起来的，将各种图像技术集中结合起来的，对整个图像领域进行研究应用的新学科。事实上，图像技术多年来的发展和积累为图像工程学科的建立打下了坚实的基础，而各类图像应用也对图像工程学科的建立提出了迫切的需要。本书将各种图像技术集合在图像工程的框架下来介绍。

1.2.2 图像工程的 3 个层次

图像工程的研究内容非常丰富，覆盖面也很广，可以分为 3 个层次(见图 1.2.1)：图像处理、图像分析和图像理解。这 3 个层次在操作对象和语义层次上都各有特点，而各层次在数据量和抽象性方面均有不同。实际上，图像工程是既有联系又有区别的图像处理、图像分析及图像理解 3 者的有机结合，另外还包括对它们的工程应用。

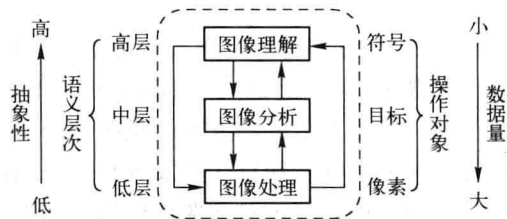


图 1.2.1 图像工程 3 层次示意图

图像处理 (IP) 可看作一大类图像技术，着重强调在图像之间进行的变换。虽然人们常用图像处理泛指各种图像技术，但比较狭义的图像处理技术的主要目标是要对图像进行各种加工以改善图像的视觉效果并为其后的目标自动识别打基础，或对图像进行压缩编码以减少图像存储所需的空间或图像传输所需的时间(从而也降低了对传输通路的要求)。

图像分析 (IA) 主要是对图像中感兴趣的目标进行检测和测量，以获得它们的客观信息从而建立对图像和目标的描述。如果说图像处理是一个从图像到图像的过程，则图像分析是一个从图像到数据的过程。这里数据可以是对目标特征测量的结果，或是基于测量的符号表示。它们描述了图像中目标的特点和性质。

图像理解 (IU) 的重点是在图像分析的基础上，进一步把握图像中各目标的性质和它们之

