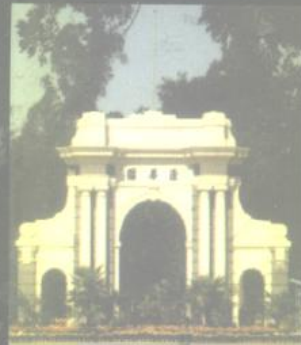




清华大学电子与信息技术系列教材

图象工程 上册

图象处理和分析

章毓晋 编著



清华大学出版社

<http://www.tup.tsinghua.edu.cn>



清华大学电子与信息技术系列教材

图象工程（上册）——

图象处理和分析

章毓晋 编著

清华大学出版社

(京)新登字158号

内 容 提 要

本书主要介绍了图象处理和分析的基本原理、典型方法和实用技术。考虑到图象技术的飞速发展和广泛应用,本书在讲解基本理论的同时还介绍了许多近年来国际上有关的最新研究成果和应用实例。

本书主要包括三大部分。第一部分(包含第1, 2, 3章)是图象基础,论述了图象工程的定义,图象技术整体概况和分类以及有关视觉和图象模型,数字图象采集、表达和象素关系,图象的各种基本变换技术等。第二部分(包含第4, 5, 6章)论述了图象处理的各重要分支,如图象增强,图象恢复,由投影重建图象和图象压缩编码等基础理论、技术和方法。第三部分(包含第7, 8章和附录A)介绍了图象分析的基本原理和技术,如图象分割、目标表达和描述、特征测量、形态学方法等。书中还提供了大量例题与习题。

本书可作为信息和信号处理、通信与电子系统、模式识别、生物医学工程等学科大学本科和研究生专业基础课教材,也可供上述学科及信息工程、电子工程、计算机科学与技术、机器人自动化、遥感和军事侦察等领域的科技工作者和高等院校的师生参考。

0025/10

图书在版编目(CIP)数据

图象工程 上册:图象处理和分析/章毓晋编著. —北京:清华大学出版社, 1999. 2
ISBN 7-302-03343-9

I. 图… II. 章… III. 图象处理—计算机应用 IV. TP391.4

中国版本图书馆CIP数据核字(1999)第03712号

出版者: 清华大学出版社(北京清华大学校内, 邮编100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 中国科学院印刷厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 **印张:** 18.75 **字数:** 465千字

版 次: 1999年3月第1版 1999年3月第1次印刷

书 号: ISBN 7-302-03343-9/TN·100

印 数: 0001~4000

定 价: 19.00 元

0000100

前 言

图象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。科学研究和统计表明，人类从外界获得的信息约有75%是从图象中获得的。随着计算机的发展，数字图象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，并已在科学研究、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。

为了对各种图象技术进行综合研究、集成应用，我们有必要建立一个整体框架——图象工程。图象工程是一门系统地研究各种图象理论、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。

图象工程是一门综合学科。考虑到其内容多、跨度大、覆盖面广以及教学的特点，本书分为相对独立又互相联系的上下两册：《图象工程（上册）—— 图象处理和分析》和《图象工程（下册）—— 图象理解与计算机视觉》。

本书在编写的方针上，力求具有理论性、实用性、系统性、实时性。考虑到图象技术的飞速发展，本书在内容上，既认真选取了有代表性的经典内容，又注意选取了 90 年代国际上的最新研究成果。本书既参考了许多有关文献，也结合了作者的一些研究工作和成果以及教案。本书主要介绍图象工程的基本内容，但又有一定的深度和广度，希望读者能通过本书的学习独立地和全面地了解该领域的基本理论、技术、应用和发展。本书在内容叙述上，力求理论概念严谨，论证简明扼要。考虑到图象技术的应用特点，书中列举了大量实例使原理和概念具体化，引导读者把这些概念和原理应用于实际。

考虑到专业课教学的特点，本书从方便读者学习的角度出发做了两项工作：

第一，在主要内容的编排上充分注意了理论的系统性和内容的结构化。全书将不同的图象技术分支均纳入图象工程大框架进行介绍，而且正文都统一按照技术分章节，互相关系清晰明确。除将第一章作为全书总纲外，之后每章（包括附录）又将其第一节作为全章总纲，并以结构图表示所属内容及分类。希望这些能使读者在学习时既有一个总体概念又可时刻把握当前学习内容在整个结构中的位置。

第二，增加了两方面的内容。其一是为方便读者进一步查阅（尤其是联网查阅）英文原文资料增加了英文目录，这也相当于提供了英文关键词。其二是本书不仅列举了书中直接引用的具体文献，还增加了一些重要专著、期刊和国际会议的名称。这不仅对阅读本书有益，为进一步深入学习提供了线索，而且对将来从事科研工作也会有帮助。

本册书为上册, 主要介绍了图象处理和分析的基本概念、基本原理、典型方法、实用技术以及国际上有关研究的新成果。读者能借此建立一个比较完整的图象处理和分析理论体系并了解和掌握常用的图象处理和分析技术。这样一方面可帮助读者进一步学习和研究图象工程高层技术, 另一方面使读者能据此解决实际图象应用中的具体问题。

本册书主要包括三大部分。第一部分(包含第 1, 2, 3 章)是图象和图象工程基础, 论述了图象技术整体概况和分类以及有关视觉和图象模型, 数字图象采集、表达和象素关系, 对图象的各种基本变换技术等。第二部分(包含第 4, 5, 6 章)论述了图象处理的各重要分支, 如图象增强, 图象恢复, 由投影重建图象和图象压缩编码等基础理论, 技术和方法。第三部分(包含第 7, 8 章和附录 A)介绍了图象分析的基本原理和技术, 如图象分割、目标表达描述、特征测量、形态学方法等。

本册书共有 8 章正文和 2 个附录, 在 10 个一级标题下有 62 个二级标题(节), 再之下还有 121 个三级标题(小节), 从内容上大体覆盖了图象处理的各个重要分支和图象分析的基本技术。全册书共有文字(也包括图象、绘图、表格、公式等)约 46 万字。本册书共有编了号的图 244 个、表格 34 个、公式 645 个。为便于教学, 本册书共给出各类例题 97 个。为便于检查教学效果, 各章后均有练习题, 本册书共有练习题 143 个。另外书末列出了提供参考和直接引用的 100 多篇文献的目录。

本册书的编写始于 1996 年, 1997 年开始作为课程教材陆续在多个学校不同年级试用。这次在此基础上, 主要结合教学实践和反馈意见进行了修改, 增加了许多例题、229 幅图片, 并在每章后都加入了练习题。徐寅、李勍、黄英、杨劲波、罗沅、吴玮、葛菁华等多名学生先后参与了可与本册书配合使用的计算机辅助教学软件 IP&A-Lab 和多媒体教学课件 IP&A-CAI(将由高教出版社出版)的编程制作。本册书采用的图片除作者制作的外主要是由徐寅、罗沅、吴玮等制作的。另外本册书还采用了刘青棣, 刘忠伟, 罗惠韬, 薛景浩和姚玉荣等学生在学位论文工作中收集和实验得到的图片。

本册书在编写和使用中得到吴国威教授的关心和帮助, 他还结合教学使用中的反馈意见对本册书提出了很好的建议, 作者愿借此机会表示衷心的感谢。

全书在编写过程中得到王仁康编辑认真的审阅和精心的修改, 这既使全书增色不少也使作者受益匪浅, 作者自行排版也得到了她的热心支持。

最后要指出, 作者妻子何芸的直接帮助和牺牲精神, 家人在各方面的理解和支持, 是作者最终能把本书写到这一页的保证!

章毓晋

1998 年国庆节于清华园

通信: 北京清华大学电子工程系, 100084
电话: (010) 62781430
传真: (010) 62770317
电子邮件: zhangyj@image.ee.tsinghua.edu.cn
主页: www.ee.tsinghua.edu.cn/~zhangyujin

目 录

第1章 绪论	1	2.7.3 等价关系和传递闭包	33
1.1 从图象到图象工程	1	2.7.4 距离量度	35
1.1.1 图象和数字图象	1	2.8 算术和逻辑运算	36
1.1.2 图象技术和图象工程	2	2.9 坐标变换	38
1.1.3 相关学科和领域	4	练习题	41
1.2 图象处理和分析	5	第3章 图象变换	43
1.2.1 图象处理和分析系统	5	3.1 概述和分类	43
1.2.2 图象采集模块	6	3.2 傅里叶变换和性质	43
1.2.3 图象显示模块	8	3.2.1 1-D 傅里叶变换	43
1.2.4 图象存储模块	8	3.2.2 2-D 傅里叶变换	44
1.2.5 图象通信模块	9	3.2.3 2-D 傅里叶变换的性质	45
1.2.6 图象处理和分析模块	9	3.2.4 由采样重建图象	50
1.3 本书安排和概要	10	3.3 快速傅里叶变换	55
练习题	12	3.3.1 快速傅里叶变换算法原理	55
第2章 图象和视觉基础	13	3.3.2 运算次数	56
2.1 概述和分类	13	3.3.3 反变换	57
2.2 人眼与亮度视觉	14	3.3.4 算法实现	57
2.2.1 人眼成象	14	3.4 其它可分离图象变换	58
2.2.2 亮度适应和区分	15	3.4.1 可分离变换	58
2.3 颜色视觉	17	3.4.2 沃尔什变换	59
2.3.1 颜色基础	17	3.4.3 哈达码变换	61
2.3.2 颜色模型	19	3.4.4 离散余弦变换	64
2.4 光度学和成象模型	22	3.4.5 哈尔变换	65
2.5 成象变换	24	3.4.6 斜变换	67
2.6 采样和量化	27	3.5 霍特林变换	68
2.6.1 均匀采样和量化	27	练习题	70
2.6.2 非均匀采样和量化	31	第4章 图象增强	72
2.7 象素间联系	31	4.1 概述和分类	72
2.7.1 象素的邻域	31	4.2 空域变换增强	73
2.7.2 连通性	32		

4.2.1 直接灰度变换	73	5.7.3 卷积逆投影重建	126
4.2.2 直方图处理	76	5.7.4 级数展开	129
4.2.3 图象间运算	81	5.7.5 综合重建方法	130
4.3 空域滤波增强	82	练习题	132
4.3.1 原理和分类	82		
4.3.2 平滑滤波器	84	第6章 图象编码	134
4.3.3 锐化滤波器	85	6.1 概述和分类	134
4.4 频域增强	87	6.2 基本概念和理论	135
4.4.1 原理和分类	87	6.2.1 数据冗余	135
4.4.2 低通滤波	87	6.2.2 图象保真度和质量	138
4.4.3 高通滤波	90	6.2.3 图象编码模型	139
4.4.4 带通和带阻滤波	92	6.2.4 信息论简介	141
4.4.5 同态滤波	93	6.2.5 基本编码定理	144
4.5 局部增强	95	6.3 简单编码方法	149
4.6 彩色增强	96	6.3.1 变长编码	149
4.6.1 伪彩色增强	96	6.3.2 位平面编码	153
4.6.2 真彩色增强	98	6.4 预测编码	158
练习题	99	6.4.1 无损预测编码	158
		6.4.2 有损预测编码	159
第5章 图象恢复和重建	101	6.5 变换编码	166
5.1 概述和分类	101	6.5.1 变换编码系统	166
5.2 退化模型和对角化	102	6.5.2 变换选择	166
5.2.1 退化模型	102	6.5.3 子图象尺寸选择	168
5.2.2 循环矩阵对角化	105	6.5.4 比特分配	169
5.2.3 无约束和有约束恢复	107	6.6 国际标准简介	172
5.3 无约束恢复	107	6.6.1 二值图象压缩标准	172
5.3.1 逆滤波	107	6.6.2 静止图象压缩标准	173
5.3.2 消除匀速直线运动模糊	109	6.6.3 序列图象压缩标准	174
5.4 有约束恢复	111	练习题	176
5.4.1 维纳滤波器	111		
5.4.2 有约束最小平方恢复	113	第7章 图象分割	179
5.5 交互式恢复	116	7.1 概述和分类	179
5.6 几何失真校正	118	7.2 并行边界技术	181
5.6.1 空间变换	119	7.2.1 边缘检测	181
5.6.2 灰度插值	120	7.2.2 微分算子	182
5.7 投影重建	121	7.2.3 边界闭合	186
5.7.1 原理和分类	121	7.2.4 哈夫变换	187
5.7.2 傅里叶反变换重建	124	7.3 串行边界技术	192

7.3.1 图搜索	192	8.5.5 不变矩	240
7.3.2 动态规划	193	8.6 关系描述	241
7.4 并行区域技术	195	8.6.1 字符串描述符	241
7.4.1 原理和分类	195	8.6.2 树结构描述符	243
7.4.2 依赖象素的阈值选取	197	8.7 特征测量误差	244
7.4.3 依赖区域的阈值选取	199	8.7.1 影响测量精确度的因素	244
7.4.4 依赖坐标的阈值选取	204	8.7.2 采样密度的影响	245
7.4.5 连通区域标记	205	8.7.3 分割算法的影响	248
7.5 串行区域技术	207	8.7.4 特征量计算公式的影响	249
7.5.1 区域生长	207	8.7.5 综合影响	251
7.5.2 分裂合并	208	练习题	252
7.6 分割评价简介	209	附录 A 图象代数——数学形态学	254
7.6.1 评价方法及分类	209	A.1 概述和分类	254
7.6.2 最终测量精度	211	A.2 基本集合定义	254
7.6.3 实用分割评价框架	213	A.3 二值形态学基本运算	256
练习题	215	A.3.1 膨胀和腐蚀	256
第 8 章 目标表达和描述	218	A.3.2 开启和闭合	261
8.1 概述和分类	218	A.3.3 基本运算性质	264
8.2 边界表达	219	A.4 二值形态学实用算法	265
8.2.1 链码	219	A.5 二值形态学小结	269
8.2.2 边界段	220	A.6 灰度形态学基本运算	269
8.2.3 多边形	221	A.6.1 膨胀和腐蚀	270
8.2.4 标记	223	A.6.2 开启和闭合	272
8.3 区域表达	223	A.6.3 基本运算性质	275
8.3.1 空间占有数组	223	A.7 灰度形态学实用算法	275
8.3.2 四叉树	224	练习题	277
8.3.3 骨架	224	附录 B 参考文献	279
8.4 边界描述	226	B.1 概述和分类	279
8.4.1 简单描述符	226	B.2 主要英文专著	280
8.4.2 形状数	228	B.3 重要英文期刊	281
8.4.3 矩	229	B.4 一些定期国际会议	281
8.4.4 傅里叶描述符	229	B.5 本书引用中文文献	282
8.5 区域描述	231	B.6 本书引用英文文献	282
8.5.1 简单描述符	231	英文目录	287
8.5.2 拓扑描述符	232		
8.5.3 形状描述符	233		
8.5.4 纹理描述符	236		

第 1 章 绪 论

1.1 从图象到图象工程

1.1.1 图象和数字图象

图象对我们并不陌生。它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 1996b]。人的视觉系统就是一个观测系统，通过它得到的图象就是客观景物在人心目中形成的影象。

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，也就是从图象中获得的。这里图象是比较广义的，例如照片、绘图、视象等等。图象带有大量的信息，百闻不如一见，一图值千字都说明了这个事实。

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象是二维 (2-D) 的。一幅图象可以用一个 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 则代表图象在点 (x, y) 的某种性质 F 的数值。例如常用的图象一般是灰度图，这时 f 表示灰度值，它常对应客观景物被观察到的亮度。需要指出，我们一般是根据图象内不同位置的不同性质来利用图象的。

常见图象是连续的，即 f, x, y 的值可以是任意实数。为了能用计算机对图象进行加工，需要把连续的图象在坐标空间 XY 和性质空间 F 都离散化。这种离散化了的图象是数字图象，可以用 $I(r, c)$ 来表示。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 代表图象的行 (row)， c 代表图象的列 (column)。这里 I, c, r 的值都是整数。本书以后主要讨论数字图象，在不至引起混淆的情况下我们用 $f(x, y)$ 代表数字图象，如不特别说明， f, x, y 都在整数集合中取值。

早期英文书籍里一般用 picture 代表图象，随着数字技术的发展，现都用 image 代表离散化了的数字图象[Zhang 1996b]。图象中每个基本单元叫做图象元素，简称像素 (picture element)。对 2-D 图象，英文里常用 pixel 代表像素。如果采集一系列的 2-D 图象或利用一些特殊设备还可得到 3-D 图象。对 3-D 图象，英文里常用 voxel 代表其基本单元，简称体素 (volume element)。

例 1.1.1 数字图象示例

图 1.1.1 给出 2 幅典型的数字图象。图(a)所用的坐标系统常在屏幕显示中采用，它的原点 O (origin) 在图象的左上角，纵轴标记图象的行，横轴标记图象的列。 $I(r, c)$ 既可代表这幅图象，也可表示在 (r, c) 行列交点处的图象值。图(b)所用的坐标系统常在图象计算

中采用，它的原点在图象的左下角，横轴为 X 轴，纵轴为 Y 轴。 $f(x, y)$ 既可代表这幅图象，也可表示在 (x, y) 坐标处象素的值。

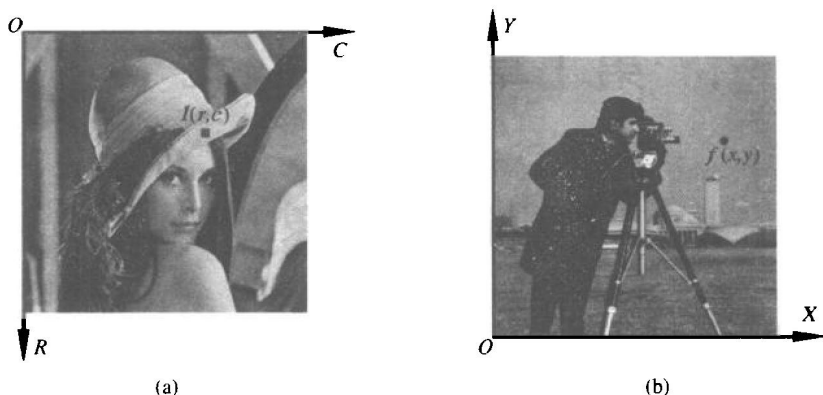


图 1.1.1 数字图象示例

□

1.1.2 图象技术和图象工程

图象技术在广义上是各种与图象有关的技术的总称。目前人们主要研究的是数字图象，主要应用的是计算机图象技术。这包括利用计算机和其它电子设备进行和完成的一系列工作，例如图象的采集、获取、编码、存储和传输，图象的合成和产生，图象的显示和输出，图象的变换、增强、恢复（复原）和重建，图象的分割，目标的检测、表达和描述，特征的提取和测量，序列图象的校正，3-D 景物的重建复原，图象数据库的建立、索引和抽取，图象的分类、表示和识别，图象模型的建立和匹配，图象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等。另外，图象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。

尽管计算机图象技术的历史可追溯到 1946 年世界上第 1 台电子计算机的诞生（借助打印设备进行数字图象处理甚至可追溯到 20 年代[Gonzalez 1992]），但在 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象的要求。在 60 年代，第 3 代计算机的研制成功，以及快速傅里叶变换算法的发现和运用使得对图象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象进行加工利用。在 70 年代，图象技术有了长足的进展，而且第 1 本重要的图象处理专著[Rosenfeld 1976]也得以出版。在 80 年代，各种硬件的发展使得人们不仅能处理 2-D 图象而且开始处理 3-D 图象。许多能获取 3-D 图象的设备和处理分析 3-D 图象的系统研制成功，图象技术得到了广泛的应用。进入 90 年代，图象技术已逐步涉及人类生活和社会发展的各个方面。以近年得到广为宣传和应用的多媒体为例，图象在其中占据了主要地位。广义上来说，文本、图形、视频等都需要借助图象技术才能充分利用。展望即将到来的 21 世纪，图象技术必将得到进一步的发展和应用，从而改变人们的生活方式以及社会结构。

由于图象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。图象界一致认为亟需对它们进行综合研究和集成应用。我们以

为这个工作需要在一个整体框架下进行, 这个框架就是图象工程[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。图象工程学科则是将数学、光学等基础科学的原理, 结合在图象应用中积累的技术经验而发展起来的。“图象工程”的概念在 1982 年首先提出, 当时主要包括有关图象的理论技术, 对图象数据的分析管理以及各种应用[Fu 1982], 但其后一段时间并未得到广泛响应。现我们重新使用这个概念, 并将图象工程看作一个对整个图象领域进行研究应用的新学科。事实上, 图象技术多年来的发展和积累为图象工程学科的建立打下了坚实的基础, 而各类图象应用也对图象工程学科的建立提出了迫切的需要。

图象工程的内容非常丰富。根据抽象程度和研究方法等的不同可分为三个层次(见图 1.1.2): 图象处理、图象分析和图象理解。换句话说, 图象工程是既有联系又有区别的图象处理、图象分析及图象理解三者的有机结合, 另外还包括对它们的工程应用。

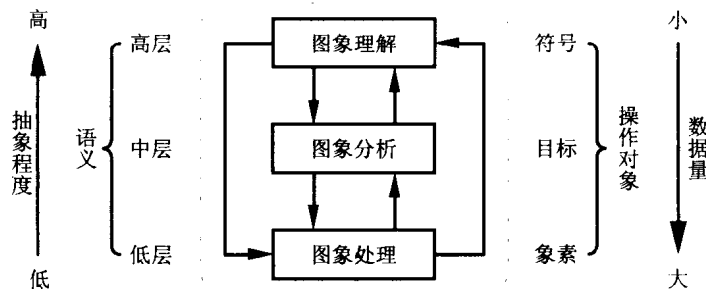


图 1.1.2 图象工程三层次示意图

图象处理着重强调在图象之间进行的变换。虽然人们常用图象处理泛指各种图象技术, 但比较狭义的图象处理主要指对图象进行各种加工以改善图象的视觉效果并为自动识别打基础, 或对图象进行压缩编码以减少对其所需存储空间或传输时间、传输通路的要求。

图象分析则主要是对图象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象的描述。如果说图象处理是一个从图象到图象的过程, 则图象分析是一个从图象到数据的过程。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。它们描述了图象中目标的特点和性质。

图象理解的重点是在图象分析的基础上, 进一步研究图象中各目标的性质和它们之间的相互联系, 并得出对图象内容含义的理解以及对原来客观场景的解释, 从而指导和规划行动。如果说图象分析主要是以观察者为中心研究客观世界(主要研究可观察到的事物), 那么图象理解在一定程度上是以客观世界为中心, 借助知识、经验等来把握整个客观世界(包括没有直接观察到的事物)。

由上所述, 图象处理、图象分析和图象理解是处在三个抽象程度和数据量各有特点的不同层次上, 可参见图 1.1.2。图象处理是比较低层的操作, 它主要在图象像素级上进行处理, 处理的数据量非常大。图象分析则进入了中层, 分割和特征提取把原来以像素描述的图象转变成比较简洁的非图形式的描述。图象理解主要是高层操作, 基本上是对从描述抽象出来的符号进行运算, 其处理过程和方法与人类的思维推理可以有许多类似之处。另外由图 1.1.2 可见, 随着抽象程度的提高数据量是逐渐减少的。具体说来, 原始图象数

99!0075

据经过一系列的处理过程逐步转化为更有组织和用途的信息。在这个过程中，语义不断引入，操作对象发生变化，数据量得到了压缩。另一方面，高层操作对低层操作有指导作用，能提高低层操作的效能。

1.1.3 相关学科和领域

图象工程是一门系统地研究各种图象理论、技术和应用的新的交叉学科。从它的研究方法来看，它与数学、物理学、生理学、心理学、电子学、计算机科学等许多学科可以相互借鉴，从它的研究范围来看，它与模式识别、计算机视觉、计算机图形学等多个专业又互相交叉，图 1.1.3 给出图象工程与相关学科和领域的联系和区别[章 1996d]。另外，图象工程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都有密切的联系，它的发展应用与医学、遥感、通信、文档处理和工业自动化等许多领域也是不可分割的。

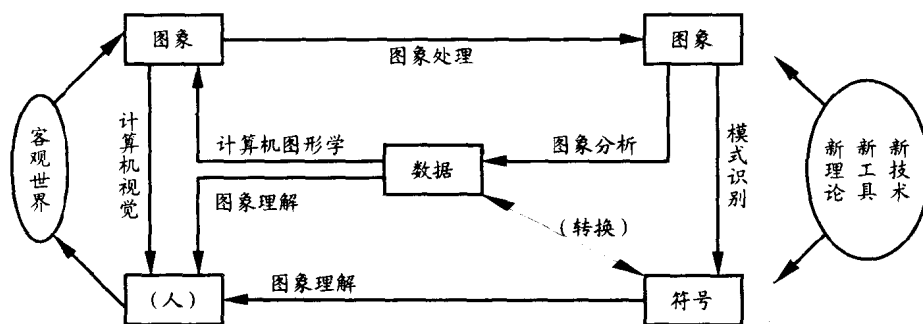


图 1.1.3 图象工程与相关学科和领域的联系和区别

从图 1.1.3 可以看到图象工程三个层次各自不同的输入输出内容以及它们与计算机图形学、模式识别、计算机视觉等的关系。图形学原本指用图形（graph）、图表（chart）、绘图（drawing）等形式表达数据信息的科学，而计算机图形学研究的就是如何利用计算机技术来产生这些形式。如果将它和图象分析对比，两者的处理对象和输出结果正好对调。计算机图形学试图从非图象形式的数据描述来生成（逼真的）图象。另一方面，（图象）模式识别与图象分析则比较相似，只是前者试图把图象分解成可用符号较抽象地描述的类别。它们有相同的输入，而不同的输出结果可以比较方便地进行转换。至于计算机视觉主要强调用计算机实现人的视觉功能，这中间实际上用到图象工程三个层次的许多技术，但目前的研究内容主要与图象理解相结合。

由此看来，以上学科互相联系，覆盖面有所重合。事实上这些名词也常混合使用，它们在概念上或实用中并没有绝然的界限。在许多场合和情况下，它们只是专业和背景不同的人习惯使用的不同术语。它们虽各有侧重但常常是互为补充的。另外以上各学科都得到了包括人工智能、神经网络、遗传算法、模糊逻辑等新理论、新工具、新技术的支持（见图 1.1.3），所以它们又都在近年得到了长足进展。总的来说，图象工程既能较好地许多

相近学科兼蓄并容，也进一步强调了图象技术的应用，所以我们选用图象工程来概括整个图象领域的研究应用。

例 1.1.2 中国图象工程文献分类

图象工程是一门综合学科，它的研究内容非常广泛，覆盖面也很大。从 1996 年开始，《中国图象图形学报》上每年都刊登对国内发表的前一年有关图象工程文献统计分类的综述文章，已形成系列。表 1.1.1 给出了对 1996 年国内 15 种期刊上发表的 1205 篇文献中选出的与图象工程有关的 212 篇文献的分类情况[章 1997c]。这里共分了 5 大类 18 小类，从中既可对图象工程的内容分类及在我国发展的情况有个概括了解，也可看出当前研究应用的热点和焦点。

表 1.1.1 中国图象工程文献分类

大类及名称	文献数量	小类及名称	文献数量
A：图象处理	52	A1：图象采集和获取（包括成象方法、摄像机校正等）	6
		A2：图象重建（从投影重建图象）	9
		A3：图象变换，滤波，增强，恢复或复原等	15
		A4：图象压缩编码（包括算法研究、国际标准实现等）	22
B：图象分析	72	B1：边缘检测，图象分割	27
		B2：目标表达，描述，测量（包括二值图处理、数学形态学等）	16
		B3：目标形状，纹理，空间，运动等的分析	11
		B4：（2-D）目标识别，分类和提取	18
C：图象理解	30	C1：（序列、立体）图象（特征点）匹配	13
		C2：3-D 建模，客观场景恢复	15
		C3：图象解释，推理	2
D：技术应用	55	D1：硬件，（硬件）系统	8
		D2：视频，通信	6
		D3：文档（包括文字，数字，符号等）	5
		D4：生物医学	17
		D5：遥感	11
		D6：其它	8
E：综述评论	3	E1：综述（介绍图象处理、分析、理解或它们的综合进展）	3

□

1.2 图象处理和分析

1.2.1 图象处理和分析系统

图象工程包含的内容很多，跨度也很大，覆盖了许多方面。本书主要介绍图象处理

和分析的一些基本技术。通过综合这些技术可构建各种图象处理和分析系统，并具体应用于实际问题（如[Perry 1990], [Takashi 1992], [Ten Kate 1990], [Zhang 1991a]）。

一个基本的图象（处理和分析）系统构成可由图 1.2.1 表示。图中各模块都有特定的功能，分别是采集、显示、存储、通信、处理和分析。图中还列出了完成各功能所需的一些设备。图象采集可采用电荷耦合器件（charge coupled devices, CCD）照相机、带有视象管（vidicon）的视频摄像机和扫描仪（scanners）等。图象显示可用电视显示器（TV monitors）、随机读取阴极射线管（cathode ray tubes, CRT）和各种打印机（printers）等。图象存储可采用磁带（magnetic tape）、磁盘（magnetic disks）、光盘（optical disks）和磁光盘（magneto-optical disks）等。图象通信可借助综合业务网（ISDN）、计算机局域网（LAN），甚至普通电话网（PSTN）等。最后，图象处理和分析主要是运算，所使用的设备主要是计算机，当然必要时还可借助专用硬件。

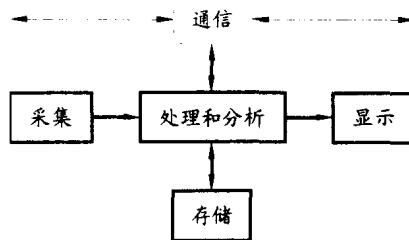


图 1.2.1 图象（处理和分析）系统的构成示意图

以下对图 1.2.1 中的各个模块分别给予简单介绍。

1.2.2 图象采集模块

为采集数字图象，需要两种装置（器件）。一种是对某个电磁能量谱波段（如 X 射线、紫外线、可见光、红外线等）敏感的物理器件，它能产生与所接受到的电磁能量成正比的（模拟）电信号。另一种称为数字化器，它能将上述（模拟）电信号转化为数字（离散）的形式。所有采集数字图象的设备都需要这两种装置。

以常见的 X 光透视成像仪为例，由 X 光源发出的射线穿越物体到达另一端的对 X 光敏感的媒体。这个媒体能获得物体材料对 X 光不同吸收的图象。它可以是胶片，一个带有将 X 光转化为光子的电视摄影机，或其它能输出数字图象的离散检测器。

用于可见光和红外线成像的设备主要有显微密度计（micro-densitometers）、析象管（image dissector）、视象管和对光子敏感的固态阵（solid-state arrays）等。使用显微密度计时，需要数字化的图象应是透明底片或照片的形式。视象管和对光子敏感的固态阵除可接受以上形式的图象外，还可以将有足够光强入射检测器的自然图象数字化。

在使用显微密度计时，需要将透明底片或照片放在一个平板上或卷在一个圆鼓上。当光线聚焦在图象上时，平移平板或转动圆鼓就可以完成扫描。如果是透明底片，光穿过透明底片；如果是照片，光从表面反射。在这两种情况下，光束都聚焦在光子检测器上，各个检测器根据光强度记录下对应当前位置的图象灰度值。如果灰度值和位置坐标都取整

数, 就得到一幅数字图象。尽管显微密度计速度比较慢, 但由于机械平移过程的连续本质, 它的空间精确度很高。

视象管摄影机的工作原理基于光导性质。聚焦在视象管表面的图象能形成与光学图象的辉度分布相对应的光导模式。用另一个独立的电子束扫描光导管的另一面, 由于中和作用, 这个电子束在一个接受器上产生与输入光亮度模式对应的压差信号。如果量化这个信号并记录下对应的扫描束位置, 就可得到一幅数字图象。

固态阵是由称为感光基元 (photosites) 的离散硅成象元素构成的。这样的感光基元能产生与所接受的输入光强成正比的输出电压。固态阵可按几何组织形式分为两种: 线扫描器和平面扫描器。线扫描传感器包括一行感光基元, 它靠场景和检测器之间的相对运动来获得 2-D 图象。平面扫描传感器由排成方阵的感光基元组成, 可直接得到 2-D 图象。固态平面传感器阵的一个显著特点是它具有非常快的快门速度 (可达 10^{-4} s), 所以能将许多运动定格下来。

固态阵中主要元件是电荷耦合器件 CCD。图 1.2.2 给出一个线扫描 CCD 传感器的示意图。这个传感器由一行感光基元, 两个定时地将感光基元中的内容传给传输寄存器的传输门, 以及一个定时地将传输寄存器中的内容传给放大器的输出门构成。放大器输出的电压信号与感光基元行的内容成比例。

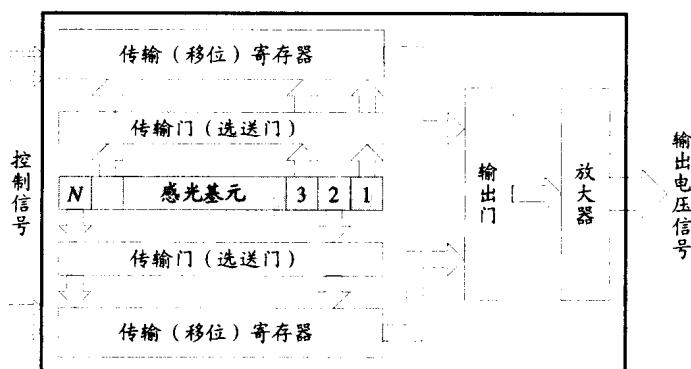


图 1.2.2 线扫描传感器

电荷耦合平面阵的工作原理与线阵相似, 但这里感光基元排列成一个矩阵形式 (如图 1.2.3 所示)。感光基元列由传输门和传输寄存器隔开。先将奇数列感光基元的内容顺序送进垂直传输寄存器, 然后再送进水平传输寄存器。把水平传输寄存器的内容送进放大器就得到 1 帧隔行的视频信号。对偶数列感光基元重复以上过程就可得到另 1 帧隔行的视频信号。将 2 帧合起来就得到隔行扫描电视的 1 场 (f)。NTSC 制的扫描速度是 30 f/s, PAL 制的扫描速度是 25 f/s。

现在常用的线扫描 CCD 一般有 512 到 4096 个像素或更多, 而 4096×4096 个像素的平面扫描 CCD 也已在使用。利用图象处理和分析的手段, 还可以通过图象的拼接用较小分辨率的 CCD 获得较大视场的图象[章 1997b]。电视摄像机一般用 CCD 阵组成, 要得到数字图象需把摄像机的视频输出送到一个数字化器中。这常通过在计算机中插专门的硬件卡来实现。

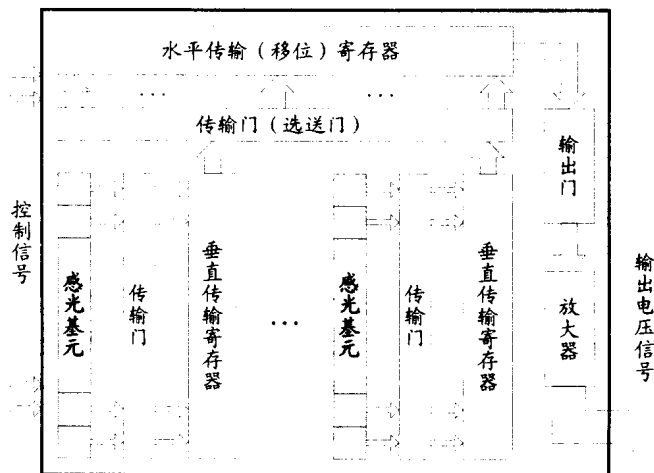


图 1.2.3 平面扫描传感器

1.2.3 图象显示模块

对图象处理来说，处理的结果主要用于显示给人看。对图象分析来说，分析的结果也可以借助计算机图形学技术转换为图象形式直观地展示。所以图象显示对图象处理和分析系统来说是非常重要的。

常用图象处理和分析系统的主要显示设备是电视显示器。输入显示器的图象也可以通过硬拷贝转换到幻灯片、照片或透明胶片上。除了电视显示器，可以随机存取的阴极射线管（CRT）和各种打印设备也可用于图象输出和显示。

在 CRT 中，电子枪束的水平垂直位置可由计算机控制。在每个偏转位置，电子枪束的强度是用电压来调制的。每个点的电压都与该点所对应的灰度值成正比。这样灰度图就转化为光亮度变化的模式，这个模式被记录在阴极射线管的屏幕上。

打印设备一般用于输出较低分辨率的图象。在纸上打印灰度图象的一种简便方法是利用标准行打印机的重复打印能力。输出图象上任一点的灰度值可由该点打印的字符数量和密度来控制。近年来使用的各种热敏、喷墨和激光打印机等具有更高的能力，已可打印较高分辨率的图象。

1.2.4 图象存储模块

图象包含有大量的信息，因而存储图象也需要大量的空间。在图象处理和分析系统中，大容量和快速的图象存储器是必不可少的。在计算机中，图象数据最小的量度单位是比特（bit）。存储器的存储量常用字节（1 byte = 8 bit）、千字节（k byte）、兆（ 10^6 ）字节（Mbyte）、吉（ 10^9 ）字节（Gbyte）、太（ 10^{12} ）字节（Tbyte）等表示。存储 1 幅 1024×1024 的 8bit 图象就需要 1Mbyte 的存储器。用于图象处理和分析的数字存储器可分为 3 类：

- (1) 处理和分析过程中使用的快速存储器；
- (2) 用于比较快地重新调用的在线或联机存储器；

(3) 不经常使用的数据库(档案库)存储器。

计算机内存就是一种提供快速存储功能的存储器。目前一般微型计算机的内存常为几十 M byte。另一种提供快速存储功能的存储器是特制的硬件卡,也叫帧缓存。它常可存储多幅图象并可以视频速度(每秒 25 或 30 幅图象)读取。它也可以允许对图象进行放大缩小,以及垂直翻转和水平翻转。目前常用的帧缓存容量在几十 Mbyte 到上百 Mbyte。

磁盘是比较通用的在线存储器,常用的 Winchester 磁盘已可存储多个 G byte 的数据。近年还常用磁光(magneto-optical, MO)存储器,它可在 5¼英寸的光片上存储上 G byte 的数据。在线存储器的一个特点是需要经常读取数据,所以一般不采用磁带一类的顺序介质。对更大的存储要求,还可以使用光盘塔,一个光盘塔可放几十个到几百个光盘,利用机械装置插入或从光盘驱动器中抽取光盘。

数据库存储器的特点是要求非常大的容量,但对数据的读取不太频繁。这里常用磁带和光盘。长 13 英尺的磁带可存储达 G byte 的数据。但磁带的储藏寿命较短,在控制很好的环境中也只有 7 年。一般常用的一次写多次读(write-once-read-many, WORM)光盘可在 12 英寸的光盘上存储 6 G byte 数据,在 14 英寸的光盘上存储 10 G byte 数据。另外 WORM 光盘在一般环境下可储藏 30 年以上。在主要是读取的应用中,也可将 WORM 光盘放在光盘塔中。一个存储量达到 T byte 级的 WORM 光盘塔可存储上百万幅 1024×1024 的 8 bit 图象。

1.2.5 图象通信模块

近年来随着信息高速公路的建设,各种网络的发展非常迅速。因而,图象的通信传输也得到了极大的关注。另一方面,图象传输可使不同的系统共享图象数据资源,也极大地推动了图象在各个领域的广泛应用。图象通信可分成近程的和远程的,它们的发展情况不太一致。

近程图象通信主要指在不同设备间交换图象数据,现已有许多用于局域通信的软件和硬件以及各种标准协议。

远程图象通信主要指在图象系统间传输图象。长距离图象通信遇到的首要问题是图象数据量大而传输通道通常比较窄。例如目前常用的电话线的速率为 9 600 bit/s,如果以这样的速率传输 1 幅 512×512 的 8 bit 图象就需要 300 s。利用中继站的无线传输速率比较高,但费用增加得更高。解决这个问题需要对图象数据进行压缩,这要靠下面介绍的图象处理和分析模块完成。

1.2.6 图象处理和分析模块

对图象的处理和分析(参见 1.1.2 小节)一般可用算法的形式描述,而大多数的算法可用软件实现,只有在为了提高速度或克服通用计算机限制的情况下才用特制的硬件。进入 90 年代,人们设计了各种与工业标准总线兼容的可以插入微机或工作站的图象卡。这不仅减少了成本,也促进了图象处理和分析专用软件的发展。这些图象卡包括用于图象数字化和临时存储的图象采集卡,用于以视频速度进行算术和逻辑运算的算术逻辑单元,以