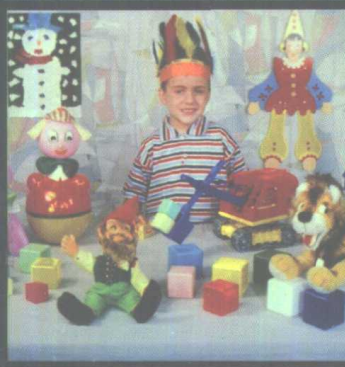
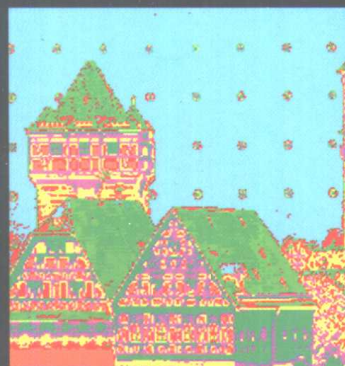
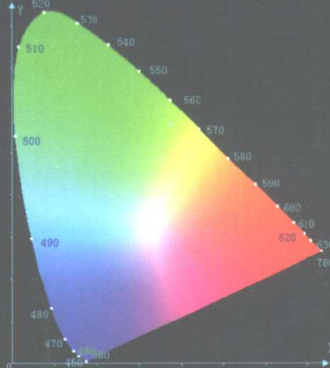
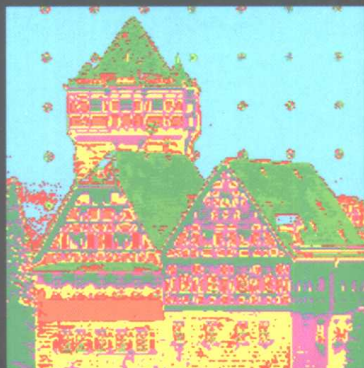




清华大学电子与信息技术系列教材

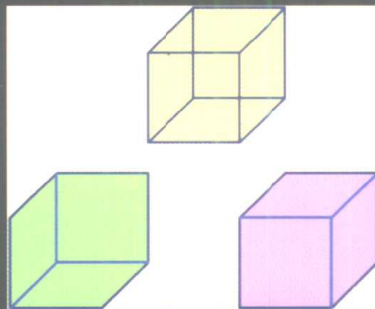
图象工程 附册

教学参考及习题解答

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(附册)——

教学参考及习题解答

章毓晋 编著

清华大学出版社

(京)新登字 158 号

内 容 提 要

本书是《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》的附册。

本书主要包括三个部分。第一部分是教学参考材料,这部分内容的选取包括多个方面,有作者通过教学实践认为可帮助教学的补充材料,有根据其他教师的反馈意见而进行的扩展,有考虑到学生专业和基础不同而补充的内容,还有作者近期相关工作的总结和该领域的最新研究成果。第二部分是前两册书中所有习题的解答步骤和参考答案。第三部分是学生的一些典型问题及简短的解答。这样,该书既可提供教学和学习参考,方便教师扩展内容选择使用,又可帮助读者进一步学习和了解自我掌握的情况。

本书可用作在信息工程、电子工程、计算机科学与技术、机器人、自动化、医学工程、遥感、测绘和军事侦察等应用图象技术的领域使用前两册书教学的补充材料,学习前两册书的参考材料,自学前两册书的辅导材料。

书 名: 图象工程(附册)——教学参考及习题解答

作 者: 章毓晋 编著

出版者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

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

印刷者: 北京密云胶印厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 印张: 15 字数: 340 千字

版 次: 2002 年 1 月第 1 版 2002 年 1 月第 1 次印刷

书 号: ISBN 7-302-04935-1/TN·123

印 数: 0001~5000

定 价: 19.00 元

前 言

本书为《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》的附册。

《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》两册书出版后,已有相当多的读者。上册 1999 年 3 月出版,现已印刷 6 次,共达 2 万多册;下册 2000 年 8 月出版,现已印刷 2 次,共达近 1 万册。

前两册书均是按教材编写和出版的。据了解,除有不少大学选作教材外,还有许多人用于自学和科研参考。通过作者的教学实践和不少同行的意见反馈,在分析了读者的来信来电(包括几百个 E-mail)以及学生的提问和要求后,作者感到有必要进一步提供一些教学参考和补充材料,以方便教学工作和读者自学及科研参考使用。

本书的内容主要包括 3 个部分。第 1 部分是教学参考材料,这部分内容的选取包括多个方面,有作者通过教学实践认为可帮助教学的材料,有根据其他教师的反馈意见而进行的扩展,有考虑到学生专业和基础不同而补充的内容,还有作者近期相关工作的总结和该领域的最新研究成果。第 2 部分是前两册书中所有习题的解题步骤和参考答案。第 3 部分是学生的一些常见问题及简短的解答。这样,该书既可提供教学和学习参考,方便教师扩展内容选择使用,又可帮助读者进一步学习和了解自我掌握的情况。

本书在编排次序上与前两册书对应,分为两大篇,第 1 篇对应上册,第 2 篇对应下册。对每一篇,各章均用原书的标题,节都为两节,第 1 节是教学参考,是增加的教学参考材料;第 2 节为习题解答,完全对应前两册书。学生的常见问题及解答则放在第二篇后面。

教学参考的内容根据知识点结合,互相之间相对独立,基本上与前两册书按章对应,但不具体对应节。本书的教学参考材料可作教学补充使用,就如同上、下册的附录 A 一样。事实上,上、下册的附录 A 也是按章内的结构编写的。

习题解答均是根据前两册书中的内容对各章后的习题进行逐一解答。上、下册的习题中都有有一些是由以前学习该课程的学生编出来的,或是在历届学生作业和课堂讨论的基础上提炼出来的[章 2000d]。原习题也列在本书中,以方便读者独立使用本书。

上、下册在编写时的考虑有相当区别,上册原定对象为首次学习图象理论和技术的读者,内容介绍比较具体,习题多以复习书中内容为主;下册则主要面向有一定图象知识基础的读者,内容介绍相对概括和抽象一些,有些习题是以扩充书中内容为目的的,有些则是为了提供课外设计型的题目[章 1999c]。在解答时,对上册的习题均提供了比较详细的步骤,而对下册的习题有些则仅给出思路或方向。

本书第 1 篇有对应上册书 8 章正文和 1 个附录的内容,第 2 篇有对应下册书 9 章正文和 1 个附录的内容。这样全书共有 21 个一级标题、38 个二级标题、65 个三级标题。全书共有文字(包括图象、绘图、表格、公式等)约 35 万字。本书共有编了号的图 169 幅、表格 27 个、公式 209 个,还有 4 个例题,对上、下册的 285 个习题提供了较详细的解答,有些

题甚至还提供了两个解答。

顺便对前两册书的教学学时提出一个参考,见表 0.1。上册用 48 学时应可完成 8 章正文,下册 32 学时应可完成 9 章正文。下册主要面向研究生,所以可进一步精讲多练。

表 0.1 上、下册教学学时参考表

上 册		下 册	
章	学时	章	学时
1	2	1	1
2	7	2	2
3	6	3	4
4	7	4	5
5	7	5	4
6	7	6	5
7	6	7	4
8	6	8	5
		9	2
小计	48	小计	32

另外,与上册书配套的“图象处理和分析”计算机辅助多媒体教学课件已由高等教育出版社和高等教育电子音像出版社出版[章 2001d],其中既有例题,又有习题[章 1999d],如把它与本书结合使用,将对学习很有帮助[Zhang 1999]。

值得指出的是,书中汇集了几年来不少听课学生的贡献,许多例题和一些图片还直接由学生帮助制作,在选材上也从学生的反馈中受到许多启发。另外,本书还采用了俞天力等学生在学位论文工作中收集和实验得到的图片。感谢学生戴声扬、葛菁华、李娟、刘惟锦、谭华春、朱小青试做了部分习题和制作了部分图片,以及谭华春回复 E-mail 的助教工作。感谢读者(既有同行又有学生)的意见反馈和讨论,是他们既推动了本册书的编写,又为作者提供了不少思路和帮助。同时感谢王仁康编审在构思和规划上的热心指导和陈国新编辑的认真审阅。

最后要指出,作者妻子何芸的直接帮助和牺牲精神,父母和孩子在各方面的理解和支持,为作者的写作提供了保证!

章毓晋

2001 年 5 月 1 日于清华园

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@ee.tsinghua.edu.cn

个人主页: www.ee.tsinghua.edu.cn/~zhangyujin

研究室网: image.ee.tsinghua.edu.cn

目 录

第 1 篇(上册)

第 1 章 绪论.....	1	4.1.1 直接灰度变换后的 直方图示例.....	32
1.1 教学参考	1	4.1.2 单映射规则和组映射 规则的映射示例.....	32
1.1.1 图象工程文献的作者 情况统计	1	4.1.3 图象间运算的应用.....	33
1.1.2 近 6 年图象工程文献 选取和分类概况	2	4.1.4 中值滤波器.....	34
1.1.3 半调图象	2	4.1.5 锐化滤波器效果 示例.....	35
1.1.4 图象文件格式	3	4.1.6 色度分量的增强.....	36
1.2 习题及解答	5	4.2 习题及解答.....	37
第 2 章 图象和视觉基础.....	8	第 5 章 图象恢复和重建	44
2.1 教学参考	8	5.1 教学参考.....	44
2.1.1 红绿蓝三色与 色度图	8	5.1.1 灰度插值的实现.....	44
2.1.2 从 RGB 到 HSI 的转换	9	5.1.2 高阶灰度插值.....	44
2.1.3 辐射度量学和光度 学名词.....	10	5.2 习题及解答.....	45
2.1.4 亮度和照度.....	11	第 6 章 图象编码	52
2.2 习题及解答.....	11	6.1 教学参考.....	52
第 3 章 图象变换	20	6.1.1 率失真定理.....	52
3.1 教学参考.....	20	6.1.2 解码的惟一性和 即时性.....	53
3.1.1 Gabor 变换.....	20	6.1.3 香农-法诺编码	54
3.1.2 小波变换.....	20	6.1.4 图象数字水印.....	54
3.1.3 小波包变换.....	22	6.2 习题及解答.....	59
3.2 习题及解答.....	23	第 7 章 图象分割	71
第 4 章 图象增强	32	7.1 教学参考.....	71
4.1 教学参考.....	32	7.1.1 坎尼算子.....	71
		7.1.2 分析直方图凹凸性 确定阈值.....	72
		7.1.3 基于小波变换的多 分辨率阈值选取	73

7.2 习题及解答	74	2.1.2 随机点立体图	111
第8章 目标表达和描述	83	2.2 习题及解答	112
8.1 教学参考	83	第3章 立体成象	115
8.1.1 积分光密度	83	3.1 教学参考	115
8.1.2 一种建立四叉树 的方法	83	3.1.1 深度图象和亮度图象 的同时采集	115
8.1.3 二叉树	84	3.1.2 摄象机外部参数 的校准	115
8.1.4 骨架的比较	85	3.1.3 立体镜旋转采集 图象	119
8.1.5 小波轮廓描述符	85	3.1.4 显微镜 3-D 分层 成象	120
8.1.6 八边形区域的 形状参数	86	3.2 习题及解答	123
8.2 习题及解答	88	第4章 扩展分析技术	133
附录 A 图象代数——数学形 态学	98	4.1 教学参考	133
A.1 教学参考	98	4.1.1 HSI 空间中的彩色 图象阈值分割	133
A.1.1 用位移实现膨胀 和腐蚀	98	4.1.2 基于小波变换的 有监督纹理分割	134
A.1.2 距离变换	99	4.1.3 立体图投影	136
A.1.3 开启和闭合的 几何解释	100	4.1.4 灰度共生矩阵和 纹理描述符	137
A.2 习题及解答	101	4.1.5 灰度-梯度共生矩阵和 纹理描述符	139
第2篇(下册)		4.2 习题及解答	141
第1章 绪论	105	第5章 立体视觉	150
1.1 教学参考	105	5.1 教学参考	150
1.1.1 一个有关图象分析和 计算机视觉的综述系列	105	5.1.1 深度图象与灰度 图象的比较	150
1.1.2 计算机视觉技术在体育 训练中的应用示例	105	5.1.2 基于边缘线段的 三目立体匹配	151
1.2 习题及解答	107	5.2 习题及解答	152
第2章 视感觉和视知觉	110	第6章 三维形状信息恢复	156
2.1 教学参考	110	6.1 教学参考	156
2.1.1 对一些视错觉 假说的讨论	110		

6.1.1 摄象机的角运动 与照度变化	156	8.2 习题及解答	175
6.1.2 视频序列中的全局运动 和局部运动	156	第9章 视觉理论和信息系统	183
6.1.3 基于高阶梯度 的光流计算	158	9.1 教学参考	183
6.1.4 平面上圆形纹理的形变 与面元朝向的关系	159	9.1.1 从针图恢复深度	183
6.2 习题及解答	159	9.1.2 虚拟现实和增强 现实	184
第7章 知识和表达	164	9.1.3 基于图象的绘制	185
7.1 教学参考	164	9.2 习题及解答	187
7.1.1 基于知识的建模	164	附录 A 相关技术	190
7.1.2 属性超图	165	A.1 教学参考	190
7.2 习题及解答	166	A.1.1 贝叶斯规则和 判决误差	190
第8章 匹配和理解	170	A.1.2 聚类分割	193
8.1 教学参考	170	A.1.3 用神经网络辨识 目标的步骤	194
8.1.1 运动特征匹配	170	A.2 习题及解答	195
8.1.2 结构匹配量度	173	上册常见问题解答	202
8.1.3 检索结果的计量 准则	174	参考文献	226

第 1 篇 上 册

第 1 章 绪 论

1.1 教学参考

1.1.1 图象工程文献的作者情况统计

为了总结和推动图象工程这一新学科的发展,需要了解和掌握该学科有关科技人员的情况。为此,根据文献计量学的原理[章 2001c],对 1996—2000 年《中国图象图形学报》上连续刊登的对图象工程文献统计分类的综述文章([章 1996a]、[章 1996b]、[章 1997a]、[章 1998]、[章 1999b]、[章 2000b])中索引文献的作者数量、作者发文量、作者合著量、篇均作者数、活跃作者群、作者增复量等进行了统计分析。根据统计,在 1995—1999 年的 5 年期间共有 1333 篇有关论文在 15 个中国图象工程重要期刊上发表,涉及的作者共 2026 人(指所有排名作者)。下面根据作者的发文量对作者的数量和水平进行讨论。

作者发文量在一定程度上反映了他们对图象工程学科的贡献[林 1996]。表 1.1 给出对不同发文量作者数的统计情况。

表 1.1 作者发文量统计

发文量/篇		1	2	3	4	5	6	7	8
作者	数量/人	1447	302	102	58	32	17	18	11
	比例	71.42%	14.91%	5.03%	2.86%	1.58%	0.839%	0.888%	0.543%
发文量/篇		9	10	11	12	13	14	15	16
作者	数量/人	8	3	1	5	3	2	2	4
	比例	0.395%	0.148%	0.049%	0.247%	0.148%	0.099%	0.099%	0.197%
发文量/篇		17	18	19	20	21	22	合计	人均
作者	数量/人	4	2	2	1	1	1	2026	1.779
	比例	0.197%	0.099%	0.099%	0.049%	0.049%	0.049%	100%	

科技人员的数量和水平是衡量一门学科发展程度的重要测度指标之一。从表 1.1 可

知,在这5年里,有2000多名作者取得了科研成果,这是一个不小的数量。但在2000多名作者中,发文量仅1篇的作者占了71.42%(根据洛特卡定律应为60%[丁1993]),这既表明图象工程仍然属于一个新兴和前进中的学科,又从一个侧面反映了我国当前图象工程学科领域的研究人员变动比较频繁的特点。对表1.1的进一步分析表明,一方面平均每个作者5年内仅发表了1.78篇论文,这是一个比较小的数值;另一方面发文量在5篇及5篇以上(即平均每年1篇以上)的作者人数有117人,约占作者总数的5.8%,表明目前图象工程学科在中国已有一个基本的有一定水平的论文作者队伍。

1.1.2 近6年图象工程文献选取和分类概况

表1.2给出了对1995—2000年期间发表在图象工程重要期刊上的图象工程文献统计分类的概况[章2001b]。表1.2中,论文总数表示文献综述系列(参见1.1.1小节)所选15种期刊上发表的论文总数,选取总数为其中有关图象工程文献的数量,分类栏中括号内为该类文献数量在总选取文献数量中所占的比例。

表1.2 近6年图象工程文献的选取和分类

年度	论文总数	选取总数	选取率	图象处理	图象分析	图象理解	技术应用	综述评论
1995	997	147	14.74%	35(23.8%)	52(35.4%)	14(9.52%)	46(31.3%)	
1996	1205	212	17.59%	52(24.5%)	72(34.0%)	30(14.2%)	55(25.9%)	3(1.42%)
1997	1438	280	19.47%	104(37.1%)	76(27.1%)	36(12.9%)	60(21.4%)	4(1.43%)
1998	1477	306	20.72%	108(35.3%)	96(31.4%)	28(9.15%)	71(23.2%)	3(0.98%)
1999	2048	388	18.95%	132(34.0%)	137(35.3%)	42(10.8%)	73(18.8%)	4(1.03%)
2000	2117	464	21.92%	165(35.6%)	122(26.3%)	68(14.7%)	103(22.2%)	6(1.29%)
小计	9282	1797		596(33.2%)	555(30.9%)	218(12.1%)	408(22.7%)	20(1.11%)
平均	1547	300	19.36%	99.3	92.5	36.3	68	3.3

从表1.2的统计数据可以看出:

(1) 这6年中无论是论文总数还是选取总数都是逐年增长的。论文总数的增加表明这些刊物自身近年来在不断发展,而选取总数的增长则表明对图象工程的研究和应用在我国越来越广泛。

(2) 文献选取率反映了图象工程在各刊覆盖专业范围中的重要性。这6年来,文献选取率基本上是单增的,图象工程文献在这么多学会和专业的刊物中不断增加,并已占到一定的比例,这充分说明图象技术已得到了广泛的重视和应用。

(3) 有关图象处理的文献数量和图象分析的文献数量均超过文献选取总数的30%,它们是图象工程目前发展的主体。

1.1.3 半调图象

图象的原始灰度常有几十到几百级甚至上千级,但有些图象输出设备的灰度只有两

级(或者打印,输出黑,或者不打印,输出白),如激光打印机。为了在这些设备上输出图象并保持其原有灰度级,常采用的一种技术称为半调输出(halftoning)[Gonzalez 1992]。半调输出的原理是利用人眼的集成特性,在每个象素位置打印一个其尺寸反比于该象素灰度的圆点,即在亮的图象区域打印的点小,而在暗的图象区域打印的点大。这里要求点足够小,观察距离足够远,观看时分不出各个点。一般报纸上图片的分辨率约在每英寸 100 点(dot per inch, DPI),而书或杂志上图片的分辨率约在每英寸 300 点。

为得到半调模式的输出可有几种近似方法。早期利用打印机的方法是将字符重叠打印,即在亮的图象区域重叠的次数少一些,而在暗的图象区域重叠的次数多一些。另一种方法则将图象区域先细分,取邻近的结合起来,这样在每个输出区域内包含若干个单元,通过把一些单元输出黑,而把其他单元输出白来得到不同灰度的效果。例如利用一个 2×2 区域按照图 1.1 输出 5 种灰度,利用一个 3×3 区域按照图 1.2 输出 10 种灰度。这里如果一个区域在某个灰度为黑,则在所有大于这个灰度的输出中仍为黑。按这种方法,要输出 256 种灰度需要使用一个 16×16 的象素区域。需要注意的是,这个方法是以减少图象的空间分辨率来增加图象的幅度分辨率的,所以有可能导致图象采样过粗而影响图象的显示质量。

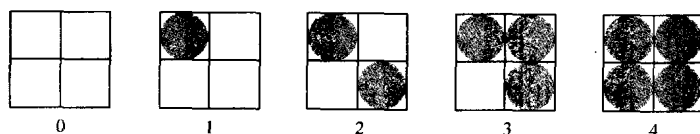


图 1.1 用 2×2 区域输出 5 种灰度

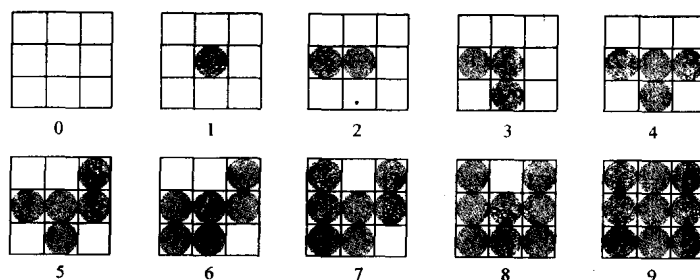


图 1.2 用 3×3 区域输出 10 种灰度

与上述方法密切相关的一种技术是抖动(dithering)技术,它是通过变动图象的幅度值来改善量化过粗图象显示质量的一种方法。抖动的实现一般是给原始图象 $f(x, y)$ 加一个随机的小噪声 $d(x, y)$,即将两者加起来进行显示,由于 $d(x, y)$ 的值与 $f(x, y)$ 没有任何有规律的联系,所以可帮助消除量化不足而导致的图象中出现虚假轮廓的现象。

下面是实现抖动的一种具体方法。设 b 为图象显示的比特数,则 $d(x, y)$ 的值可以从以下 5 个数中以均匀概率取得: $-2^{(5-b)}$, $-2^{(5-b)}$, 0 , $2^{(5-b)}$, $2^{(6-b)}$ 。将 $f(x, y)$ 加一个随机的小噪声 $d(x, y)$ 的 b 个最高有效比特作为象素的值。

1.1.4 图象文件格式

图象文件指包含图象数据的文件。文件内除图象数据本身以外,一般还有对图象的

描述信息,以便读取、显示图象。

表示图象常使用两种不同的方式,一种是矢量形式,另一种是光栅(也称位图或像素图)形式。图象文件可以采用任一种形式,也可以结合使用两种形式。

在矢量形式中,图象是用一系列线段或线段的组合体来表示的。线段的灰度(色度)可以是均匀的或变化的,在线段的组合体中各部分也可用不同灰度填充。矢量文件一般像程序文件,里面有一系列命令和数据,执行这些命令就可根据数据画出图案。矢量文件主要用于图形数据文件。

图象数据文件主要使用光栅形式。该形式与人对图象的理解一致(一幅图象是许多图象点的集合),比较适合色彩、阴影或形状变化复杂的图象。它的主要缺点是缺少对象素间相互关系的直接表示,且限定了图象的分辨率(见上册 2.6 节)。有两个问题是由后者带来的,一个是将图象放大到一定程度就会出现方块效应,另一个是如果将图象缩小再恢复到原尺寸则图象会变得模糊。

图象数据文件的格式已有很多种,不同的系统平台和软件常使用不同的图象文件格式。例如,Macintosh 机上普遍使用 MacPaint 格式(固定大小,宽 576 像素,高 720 像素),PC Paintbrush 支持 PCX 格式(包括单色、16 色、256 色),Digital Research(现 Novell)支持 GEM IMG 格式,Sun Microsystems 支持 Sun 光栅格式等[凯 1994]。

下面简单介绍 4 种应用比较广泛的格式[董 1994]、[凯 1994]。

1. BMP(bitmap)格式

BMP 格式是 Windows 环境中的一种标准(但很多 Macintosh 应用程序不支持它),它的全称是 Microsoft 设备对立位图(device independent bitmap)。BMP 图象文件也称位图文件,包括 3 个部分:位图文件头(也称表头)、位图信息(常称调色板)、位图阵列(即图象数据)。一个位图文件只能存放一幅图象。

位图文件头长度固定为 54 个字节,它给出图象文件的类型、大小和位图阵列的起始位置等信息。位图信息给出图象的长、宽、每个像素的位数(可以是 1,4,8,24,分别对应单色、16 色、256 色和真彩色的情况)、压缩方法、目标设备的水平和垂直分辨率等信息。位图阵列给出原始图象里每个像素的值(每 3 个字节表示一个像素,分别是蓝、绿、红的值),它的存储格式可以有压缩(仅用于 16 色和 256 色图象)和非压缩 2 种。

2. GIF(graphics interchange format)格式

GIF 格式是一种公用的图象文件格式标准,它是 8 位文件格式(一个像素一个字节),所以最多只能存储 256 色图象。GIF 文件中的图象数据均为压缩过的。

GIF 文件结构较复杂,一般包括 7 个数据单元:文件头,通用调色板,图象数据区,以及 4 个补充区。其中表头和图象数据区是不可缺少的单元。

一个 GIF 文件中可以存放多幅图象(这个特点对实现网页上的动画很有利),所以文件头中会包含适用于所有图象的全局数据和仅属于其后那幅图象的局部数据,当文件中只有一幅图象时,全局数据和局部数据一致。多幅图象存放时,每幅图象集中成一个图象数据块,每块的第一个字节是标识符,指示数据块的类型(可以是图象块、扩展块或文件结束符)。

3. TIFF(tagged image format file)格式

TIFF 格式是一种独立于操作系统和文件系统的格式(在 Windows 环境和 Macintosh 机上都可使用),很便于在软件之间进行图象数据交换。TIFF 图象文件包括文件头(表头)、文件目录(标识信息区)和文件目录项(图象数据区)。文件头只有一个,且在文件前端。它给出数据存放顺序、文件目录的字节偏移信息。文件目录给出文件目录项的个数信息,并有一组标识信息,给出图象数据区的地址。文件目录项是存放信息的基本单位,也称域。域主要分为基本域、信息描述域、传真域、文献存储和检索域 5 类。

TIFF 格式的描述能力很强,可制定私人用的标识信息。TIFF 格式支持任意大小的图象,文件可分为 4 类:二值图象、灰度图象、调色板彩色图象和全彩色图象。一个 TIFF 文件中可以存放多幅图象,也可存放多份调色板数据。

4. JPEG(joint picture expert group)格式

JPEG 是对静止灰度或彩色图象的一种压缩标准(参见上册 6.6.2 小节),在使用有损压缩方式时其可节省的空间是相当大的。

JPEG 标准只是定义了一个规范的编码数据流,并没有规定图象数据文件的格式。Cube Microsystems 公司定义了一种 JPEG 文件交换格式(JPEG file interchange format, JFIF)。JFIF 图象是一种或使用灰度表示,或使用 Y, C_b, C_r 分量彩色表示的 JPEG 图象。它包含一个与 JPEG 兼容的头。一个 JFIF 文件通常包含单个图象,图象可以是灰度的,其中的数据为单个分量,也可以是彩色的,其中的数据是 Y, C_b, C_r 分量。 Y, C_b, C_r 分量与常见的 R, G, B 三原色的关系如下:

$$Y = 0.299R + 0.587G + 0.114B \quad (1.1)$$

$$C_b = 0.1687R - 0.3313G + 0.5B \quad (1.2)$$

$$C_r = 0.5R - 0.4187G - 0.0813B \quad (1.3)$$

TIFF6.0 也支持用 JPEG 压缩的图象, TIFF 文件可以包含直接 DCT 的图象,也可以包含无损 JPEG 图象,还可以是用 JPEG 编码的条或块的系列(这样允许只恢复图象的局部而不用读取全部内容)。

1.2 习题及解答

1.1 连续图象 $f(x, y)$ 与数字图象 $I(r, c)$ 中各量的含义各是什么? 它们有什么联系和区别? 它们的取值各在什么范围?

解答:

连续图象 $f(x, y)$ 对应一个 2-D 数组, 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。这里 f, x, y 的值可以是任意实数。

数字图象 $I(r, c)$ 也对应一个 2-D 数组, 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, c 代表图象的列。这里 I, c, r 的值都是整数。

1.2 图象处理、图象分析和图象理解各有什么特点? 它们之间有哪些联系和区别?

解答:

图象处理、图象分析和图象理解处在 3 个抽象程度和数据量各有特点的不同层次上。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。图象处理着重强调在图象之间进行的变换。图象分析则进入了中层,分割和特征提取把原来以象素描述的图象转变成比较简洁的非图形式的描述。图象分析主要是对图象中感兴趣的目标进行提取和测量,以获得它们的客观信息,从而建立对图象的描述。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。所以图象理解主要是高层操作,基本上是对从描述中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。

1.3 图象工程主要与哪些学科相关? 它们之间如何互相影响?

解答:

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。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。

1.4 近年来有哪些新的数学工具在图象工程中得到应用?

解答:

如人工智能、神经网络、遗传算法、模糊逻辑、小波变换等。

1.5 简要叙述图象采集、图象显示、图象存储以及图象通信与图象处理和分析的联系。

解答:

图象采集、图象显示、图象存储以及图象通信都是围绕图象处理和分析而工作的,反过来,图象处理和分析又对图象采集、图象显示、图象存储以及图象通信提供支持。

1.6 近年来在图象采集、图象显示、图象存储各方面最突出的进展是什么? 对图象处理和分析产生了哪些影响?

解答:

在图象采集方面,除 CCD 速度加快、分辨率提高外,近年来固体摄像器件发展迅速,电荷注射器件(charge-injection device, CID)和 CMOS 摄像机等都已投入应用。

在图象显示方面,各种新型打印机,如热敏、热升华打印机得到应用。

在图象存储方面,微型计算机的内存多已超过 100Mbyte, Winchester 磁盘的存储量已多用几十个 Gbyte 来衡量。闪速存储器(flash)等新型存储器得到广泛使用,读写光驱已成为许多便携式计算机的标准配制。

上述进展加快了图象处理和分析的速度,并扩展了图象处理和分析的应用范围。

1.7 本书主要可分成哪些层次? 互相之间有什么关系?

解答：

全书主要可分为3个层次,第一层是基础知识,第二层是图象处理技术,第三层是图象分析技术。第一层的基础知识是支持第二层图象处理技术和第三层图象分析技术的。

1.8 附录 B 中的哪些文献你听说过或查阅过?

解答：

大部分文献在清华大学图书馆都可查到。

第2章 图象和视觉基础

2.1 教学参考

2.1.1 红绿蓝三色与色度图

人们早就认识到:红、绿、蓝三色相加会成为白色,所以对红、绿、蓝三色的感受是人眼的特性,或者说人眼中有3种相应的感受器。但在以三色为基础的视觉机制的反应曲线中(见图2.1),三曲线均有较宽的分布,互相有重叠,即某一种光可兴奋2~3种感受器,所以,即使刺激是单纯的,人眼的反应也并不单纯。

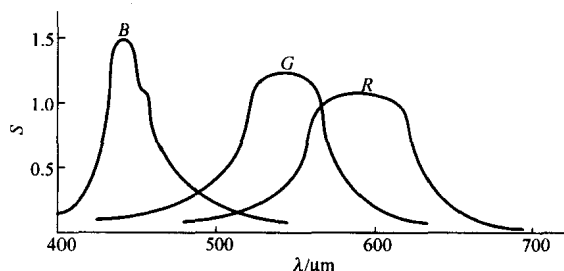


图2.1 三种感受器的反应曲线

当把红、绿、蓝三色光射到屏幕上混合时,通过改变三者各自的强度比例可得到白色以及其他各种色调和饱和度的彩色,用公式可表示为

$$C \equiv rR + gG + bB \quad (2.1)$$

其中 C 代表某一特定色, $\equiv$ 表示匹配, R, G, B 表示三原色, r, g, b 代表比例系数,且有 $r+g+b=1$ 。

在匹配某些颜色时,把 R, G, B 相加并不能得出相等的配对,此时可将三原色之一加到被匹配颜色一方(即写成负值),以达到相等的颜色配对,例如

$$C \equiv rR + gG - bB \quad (2.2)$$

$$bB + C \equiv rR + gG \quad (2.3)$$

国际照度委员会 CIE 设定了3种设想的颜色 X, Y, Z 以产生光谱的各种颜色,并消除负值(见图2.2中三刺激值曲线)。换句话说,任一波长的颜色总可用 X, Y, Z 三者叠加而得到。

进一步,把 X, Y, Z 用色系数 x, y, z 归一化就得到色度图(见上册图2.3.1)。注意,在色度图中,过 C 点直线端点的两彩色为互补色,例如对一种紫红色段的非光谱颜色,可用直线另一端光谱颜色的补色来表示,写成 $510C$ 。

对各种颜色复原系统,需要选择合适的 R, G, B ,例如 PAL 和 NTSC 两种制式的色度三角形,见图2.3。这里考虑的因素主要有:

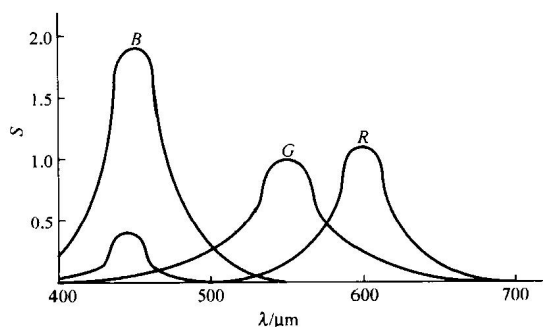


图 2.2 三刺激值曲线

- (1) 技术上产生高饱和度颜色较困难；
- (2) 三角形应较大，以包括较大面积，即包含较多的颜色；
- (3) 饱和的蓝绿色不常用，所以 RG 的连线靠近光谱轨迹。

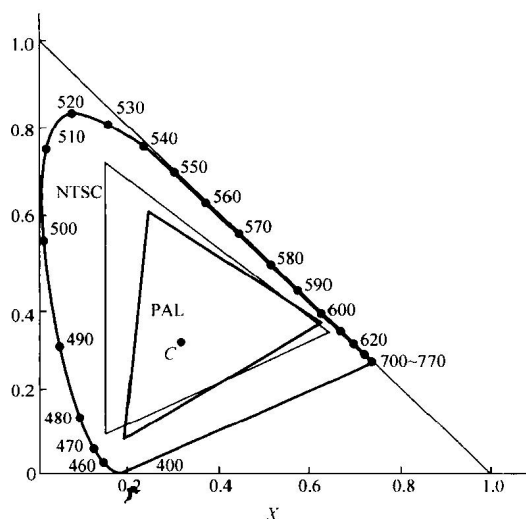


图 2.3 PAL 和 NTSC 两种制式的色度三角形

以上讨论中，匹配指颜色在视觉感受上一致，称“同貌异质(metameric)”的颜色配对。所以在色度图中，颜色的色度坐标只表达了一个颜色的外貌，而不能表达它的光谱能量分布情况。

2.1.2 从 RGB 到 HSI 的转换

参见图 2.4，图中三角形的中心为 $O(1/3, 1/3, 1/3)$ ，设考虑的颜色对应 OP ，延长 OP 到 RG 连线，交于 Q 点，再从 P 点向 OY 作垂线，交 OY 于 T 点。

令从原点到 O 的矢量为 $\mathbf{o}$ ，从原点到 P 的矢量为 $\mathbf{p}$ ，从原点到 R 的矢量为 $\mathbf{r}$ ，则有

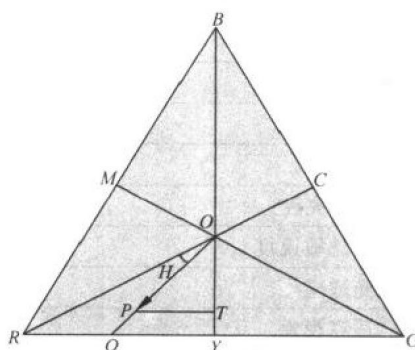


图 2.4 推导转换公式的三角形