

图像信息处理丛书



图像压缩与投影重建

孙即祥 等 编著

 科学出版社
www.sciencep.com



图像信息处理丛书

图像压缩与投影重建

孙即祥 等 编著



科学出版社

北京

内 容 简 介

《中国科学院规划教材——图像信息处理丛书》共分三册,丛书之一为《图像处理》,丛书之二为《图像分析》,丛书之三即本书。

本书涉及图像信息处理的重要内容:图像数据压缩、图像的投影重建及视频模型与检索,其中深入系统地阐述和论证了图像数据压缩理论、方法与技术。

本书适于电子科学与工程类、控制理论与工程类、计算机理论与技术类、仪器科学与技术类及其他有关专业和研究方向的研究生、本科高年级学生作为教材或学习参考书,也可供相关专业的科研人员工作中参考。

图书在版编目(CIP)数据

图像压缩与投影重建/孙即祥等编著. —北京:科学出版社,2005
ISBN 7-03-014670-0

I. 图… II. 孙… III. 图像处理-数据压缩-高等学校-教材
N. TN919.81

中国版本图书馆 CIP 数据核字(2005)第 000107 号

责任编辑:匡 敏 潘继敏/责任校对:张 琪
责任印制:钱玉芬/封面设计:耕者设计工作室 蔡 璐

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2005 年 7 月第 一 版 开本:B5 (720×1000)

2005 年 7 月第一次印刷 印张:15 1/2

印数:1—3 000 字数:297 000

定价:28.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

序 言

人类所获取的外界知识中约有 80% 以上的信息来自于视觉,很多情况下图像所承载的信息比任何其他形式的信息都更真切、更丰富,获取也更便捷。图像信息处理是一门多学科交叉的学科,它涉及数学、物理学、控制理论、信息处理、生物生理学、心理学、计算机科学等多个领域的知识。20 世纪 50 年代以来,国内外各领域的学者、专家对其高度重视,并投入了极大的研究热情,有关的机构也投入了大量的经费。经过广大研究人员几十年的努力,图像信息处理现已发展成为一个活跃的学科体系,并形成了多个研究方向,它的理论成果已广泛地应用于经济建设、国防建设等诸多社会领域中,产生了巨大的经济和社会效益。因此,关于图像信息的理论与应用研究极为重要。

孙即祥教授根据多年的教学经验和科研成果,撰写了这套既面向高校教学使用又适于科研参考的优秀系列著作。为了深入系统地介绍这门学科的有关知识,且便于读者选用,孙教授将其所涉及的内容合理安排,分三册(《图像处理》、《图像分析》、《图像压缩与投影重建》)出版。这套书中关于基础性理论的阐述充实严谨,有助于学生打下坚实的理论基础,同时书中介绍的一般性的和较新的图像信息处理技术又为广大读者提供了广泛的解决实际问题的思路和方法。相信本套书必将为图像信息处理的教学和相关科研工作的发展起到一定的促进和推动作用。

郭桂林 院士

前 言

视觉活动是人类最重要的基本活动之一,在人们的日常生活、社会活动、工作学习和科研生产中,无时无刻不在进行着视觉活动。视觉信息是人类获取外部知识、了解世界的主要途径和重要形式。据统计,人类大约有 80% 以上的信息来自于视觉;在许多情况下,没有任何其他形式比图像所传递的信息更丰富、更真切。由于人类视觉具有完善的感知能力,许多科研、工作的中间或最后的结果数据都要以可视的图像形式表示出来,以利于人们对其理解、分析和应用,还由于人们不断地对太空外层世界及粒子微观内部世界的探索,以及我们周围的科研、工作的需要,可以断言,人们以图像形式表达信息,利用视觉获取知识的比例将会进一步增加。图像信息处理是研究人或机器对图像信号的产生与采集,信息的形成、提取、分析、综合、表达和利用的理论与方法的科学,是一门综合性、交叉性学科。它在理论上涉及数学、物理学、生理心理学、认知心理学、信号处理、控制科学、计算机科学与技术等众多学科知识,它还与许多工程应用密切联系着,其内涵可以归结为信息处理、分析、决策与表达。该学科的理论任务是研究和发现信息的各种形式及各个阶段的内在转换规律,而其应用目标是研发出能帮助或代替人类视觉的智能机器系统。自 20 世纪 60 年代以来,它始终受到学术界、应用界的密切关注,并投入了极大的研究热情。随着相关学科的发展和有关技术的成熟,图像信息处理已经发展成为一门较完善的学科(群),其主要标志就是现在已有了许多相关的科技刊物,许多国内外的优秀论著,许多重大的应用成果,及许多高校设置了相关的专业。

由于教学与科研的需要,笔者曾在 1990 年完稿,并于 1993 年出版了《数字图像处理》一书。除国防科技大学研究生、本科生教学使用本书外,其他一些高校及研究单位也采用了此书,反映良好;国防科技大学至今仍使用该书,甚至近几年某些关于图像处理的计算机程序书中也直接转录了本书的许多内容。由于学科及科研、教学的发展,笔者在原《数字图像处理》一书的基础上,依据教学和科研的需要及学科发展趋势,结合该课程及有关课程(如计算机视觉、现代模式识别、机器学习)多年教学实践经验与相关科研成果,并参考大量他人的学术著作和科技文献,撰写、扩充而成本书系。笔者将它分为四部分(《图像处理》,《图像分析》,《图像压缩与投影重建》,《机器视觉》)分册出版,以利于不同课程及不同层面的读者使用。本书系是一套关于图像信息处理的理论和方法的著作,较全面系统地论述了图像信息处理的主流技术。本书系是笔者已经出版的《现代模式识别》及《模式识别中的特征提取及计算机视觉不变量》的姊妹篇,这些书涉及了信息科学与技术中的许多重要核心内容。

在本书系撰写过程中,笔者遵循以下四个原则:①在结构安排上,尽量使学科知识表达体系与学科体系相一致;②在阐述方式上遵循人的认知规律;③在选材上满足使读者掌握本学科经典和现代的重要学科知识,使之学习后能达到解决实际问题 and 能汇入学科发展潮流之目的;④使本书系受众者多,使用期长。笔者的目标就是使本书系可读性好、学术性强、实用价值大。在内容深度与表述形式上,定位于教材与专著之间,兼顾理科与工科使用。

较详细地了解本书系的特点,对于读好本书系,学习这门课程,掌握这门知识是有益的。本书系具有如下特点:

1) 内容广泛新颖。图像信息处理是一门非常活跃的重要学科,其发展非常迅速,所涉及的各种理论、方法也十分丰富,新理论、新方法、新技术和新应用不断涌现。本书系较全面系统地阐述了图像信息处理中几乎所有主流领域的知识内容,除了包括传统的、经典的重要内容之外,还收入了经实践证明有重要理论意义和应用价值的新理论、新方法 and 新技术。本书系涵盖面大,内容多,以益于不同专业的教师和学生有更多的选择空间,也满足了现在课堂讲授信息量大的教学要求。本书系的某些内容超出了有关课程当前教学大纲,内容在质和量上都具有“超前性”,不仅能扩大读者的知识面,提高读者深入学习的兴趣,而且还能适应教学与时俱进、不断更新发展的需要,同时,也为科技工作者提供较广泛的解决工作中实际问题的思路和方法。

2) 选材考究、精细。如前所述,这门学科(群)的各种理论、方法、技术纷繁众多,而且还在不断地出现和发展。笔者根据多年科研与教学工作实践的体会和认识,对图像信息处理的各理论、方法、技术及应用的成果进行梳理、沉淀、提炼、归纳和总结,以是否具有当前或潜在的理论意义或应用价值为标准,在众多的知识中选取那些或具基础理论性,或具思维训练性,或具有实用价值,或具思想启发性,或具方向前瞻性的有代表性的重要内容。

3) 结构清晰、合理。学科知识的组织结构是否得当将直接影响读者的学习效果,合理、清晰的学科知识表述体系有益于读者对各种理论、方法的理解和记忆。为了处理好本书系的内容与其他学科知识的关系,处理好图像信号与信息处理各个阶段知识点的布局,本书系在内容组织上呈现层次化、模块化。全书系及各章都尽力做到由浅入深,先易后难,先具体后抽象安排内容。笔者对结构进行了精心安排,前部分是共性的知识,然后主要以平面目标识别为目的展开论述,再后是以图像压缩及图像重构为中心展开讨论,最后一部分是以 3D 目标识别为中心的计算机视觉方面的内容。当所运用的理论和数学工具读者普遍熟知时,笔者以处理目的为纲建立知识表述结构,这样便于明了知识在学科体系中的位置,对于普遍感到不太熟知的理论和数学工具,笔者以技术为中心展开论述。《图像处理》的第一至四章分别从学科知识总体、所需数学知识、人类视觉特性以及图像的数学描述四个方面较详

尽地给出了必备的知识,然后是以目的为纲的知识表述,其次序是:图像数字化,图像正交变换,图像增强,图像恢复,图像分割,图像描述和分析,图像数据压缩,图像投影重建。当具有了这些概念、理论和方法后,再给出形态学理论和方法,小波分析的理论与方法等内容,这两部分放在书系的中后部是恰当的,以技术为中心来安排这两部分内容,首先论述基本理论,然后再介绍在各方面的应用。关于机器视觉方面的内容,它涉及射影几何,微分几何,成像的几何模型,成像的光学模型,摄像机的定标,目标的定位,运动分析,目标 3D 形状恢复,几何特征点及模型的匹配,不变量,目标描述,建模,人的视觉知识等。

本书系主要是为研究生教学使用撰写的,但也兼顾到了其他层次的读者。处理好“合”与“分”的对立与统一。正如模糊数学中的模糊集没有明确界线,论域中元素只能以某一隶属度属于某一集合一样,本书系的第一册《图像处理》大部分及其他册各章的前部分适于高年级本科生。第一册中介绍了较深入的理论及方法,第二册《图像分析》、第三册《图像压缩与投影重建》主要是面向硕士研究生。同时,这三册书中的许多内容对博士研究生也是应该了解掌握的。另外,本书系也适用于不同层面的从事相关科研的技术人员参考。

4) 注重基础。打好基础是教育经验的总结,也是科技高速发展的需要,本书系始终非常注重强化基本概念、基础思想、基本理论、基本方法和基本技能。本书系基于“新”与“恒”的对立与统一的观点和方法进行取材和论述,两者统一在知识的实在价值上。实际上,任何“新”都是相对的,事物本质的性质和规律是最有魅力、最基本和最有用的知识,这才是最有意义的、最长久的。只有掌握了基本而有用的知识,读者才能在日后的学习和科研中有“后劲”,才能持续前进,才能应对科技发展日新月异的挑战,才能以“不变”应“万变”,才能在浩瀚的知识海洋中畅游,始终站在科技潮流的最前沿。

5) 注重讲明道理,突出学术思想。为了便于读者真正掌握学科知识和实实在在地提高解决问题的能力,本书系在阐述知识时,注重不仅讲其然,还要讲其所以然。本书系不是一些定理、结论、算法步骤的简单罗列,而是注重理论的严密性与表述浅显性的统一。对一些通常认为较“简单”的内容,尽量挖掘其理论依据,使之有理论深度,而对于涉及较深奥理论的部分,在严格、严密的前提下,尽量用平实直白的语言进行论述,以避免因不必要的符号猜解和复杂推导,使读者忽略或淡化更为重要的基本概念、学术思想和技术思路。本书系各部分在深浅上基本达到一致。

6) 详略得当。由于本书系涉及的知识面广,所以必须在论述上有详有略,详略得当。对于基本的、重要的、决定性的、关键性的理论和方法要论述清楚、说深说透;而对于相关的、类似的或不“稳定”的方法,或知识层次或结构限制的内容,则进行适当的论述或点到为止,以给出联系、给出启示、给出方向。这样既实现了知识在面上的宽广,又达到知识点处的理论深度,读者通过知识的“内插”和“外推”,就可以

形成巨大的“三维”知识体。在此注意既不因面上的宽广而泛泛空谈,也不为内容的深度使读者感到晦涩难懂。内容撰写上兼顾论述深透和点到为止的“平衡”。

以上努力的目的在于让读者对图像信息处理知识能深刻理解、融会贯通、牢固记忆、全面掌握,花费较少的代价而学到较多有价值和有意义的知识;在学到知识的同时,提高自学能力、活用能力和创新能力,有益于读者综合素质的提高。

为了加深对正文的理解,在本书系各章节中给出了一些例题和图片。

与“图像信息处理”直接相关或者可作为先导课程的有:概率论与数理统计、矩阵论、各种最优化理论与方法、模糊数学、形式语言、人工智能原理、计算机软件设计、射影几何、微分几何、代数学、图论、张量分析、运动学、估计理论,等等。对于某些读者来讲,不可能也不必花费大量的精力和时间去学完上述每门课程,由于本书系在结构、选材及论述上的特点,读者只要具备一些必要的理论基础和相关基本知识便可以顺利地学完每一章的主要内容。对于希望深入学习、进行学术理论研究或从事技术应用的读者,应先行修完上述课程,并在学习该课程时根据本书系提供的参考文献参阅其他相关的科技资料。

图像信息处理是一个多阶段、多途径、多目标的信息处理过程。本册书深入、系统、广泛地阐述和论证了图像信息处理的重要内容——图像数据压缩、图像投影重建与视频图像检索,本册书是本书系的第一册《图像处理》、第二册《图像分析》的姊妹篇。第一册主要论述了关于图像信息处理的概貌,与图像处理有关的数学知识、视觉知识,图像的数学描述,图像的数字化,图像变换,图像增强,图像恢复。第二册主要论述了图像分割,图像描述与分析,数学形态学及其应用,小波变换及其应用等。

正如第一册书的前言所述,本套书得以出版是众多人员努力、支持的结果。本书由孙即祥主编,第一章的 1.14.2,1.14.3,1.14.5,1.14.6,1.14.7 由季虎撰写,1.15 由姚伟、林成龙编写,第一章的其余部分和第二章由孙即祥撰写,第三章由唐波撰写,最后由孙即祥统稿。

本书的内容除了包含我们的某些科研成果之外,还取材于国内外诸多专家学者的研究成果、论文著作,对他们以及我们的研究生表示由衷的感谢。感谢原校长郭桂蓉院士、院长庄钊文教授以及其他院、系领导对本书的撰写给予的极大支持。

由于图像信息处理是一门不断发展的学科,新的理论、方法和技术、新的应用成果不断出现,再加上我们的学识水平及时间有限,可能没有完全达到笔者所希望的目标,也不可避免地存在各种错误和疏漏,敬请读者给予批评指正。

孙即祥

2005 年 5 月

于国防科技大学

目 录

前言

第一章 图像数据压缩	1
1.1 概述	1
1.2 信息论基本知识	4
1.2.1 信息量 熵	4
1.2.2 条件熵	6
1.2.3 互信息量	7
1.2.4 编码	9
1.2.5 信源	10
1.2.6 率失真函数及限失真编码	13
1.3 轮廓编码压缩	14
1.3.1 T 算法	14
1.3.2 IP 算法	15
1.3.3 编码	15
1.4 行程编码压缩	17
1.4.1 一维行程编码	17
1.4.2 二维行程编码	17
1.5 预测误差编码	18
1.5.1 基本原理	18
1.5.2 预测器	20
1.5.3 量化器	24
1.6 正交变换压缩编码	27
1.6.1 基本原理	27
1.6.2 正交变换	28
1.6.3 量化	30
1.6.4 自适应编码压缩	38
1.7 混合编码压缩与运动补偿预测	39
1.7.1 帧内混合编码压缩	39
1.7.2 帧间混合编码压缩	39
1.7.3 预测模式的选择	39

1.7.4 运动补偿预测方法	40
1.8 子带编码技术	43
1.9 位平面编码	47
1.10 编码技术	48
1.10.1 霍夫曼码	48
1.10.2 香农-费诺码	50
1.10.3 B 码	51
1.10.4 算术编码	52
1.10.5 LZW 码	55
1.10.6 方块编码	57
1.11 标量量化与矢量量化	58
1.11.1 标量最佳量化	58
1.11.2 矢量量化原理	59
1.11.3 矢量量化某些算法	61
1.12 基于人工神经网络的压缩技术	62
1.12.1 人工神经网络的基础知识	63
1.12.2 利用多层前向网络进行数据压缩	72
1.12.3 人工神经网络的非线性预测差值编码的实现	72
1.12.4 运用 Hopfield 网络实现图像正交变换	74
1.12.5 运用神经网络实现主分量分析(KLT)	74
1.12.6 运用神经网络实现矢量量化	76
1.13 分形方法	78
1.13.1 分形的概念	78
1.13.2 分形维数——分维	80
1.13.3 有关的数学知识——度量空间、不动点、压缩映射	83
1.13.4 拼贴定理	85
1.13.5 迭代函数系统(IFS)	86
1.13.6 自动分形图像压缩方法	93
1.14 基于小波变换的图像数据压缩	97
1.14.1 图像分解与合成在图像压缩中的应用	97
1.14.2 提升小波变换	102
1.14.3 整数提升小波变换	117
1.14.4 嵌入零树编码	120
1.14.5 SPIHT 算法	124
1.14.6 基于提升小波变换和人眼视觉特性的图像压缩编码	125

1.14.7 基于整数小波的图像压缩编码	129
1.15 图像压缩的若干国际标准简介	131
1.15.1 概述	131
1.15.2 H.261/H.263 建议	135
1.15.3 JPEG 标准	140
1.15.4 MPEG-1 标准	141
1.15.5 MPEG-2 标准	151
1.15.6 MPEG-4 标准	158
1.15.7 JPEG2000 标准	160
习题	168
参考文献	170
第二章 图像的投影重建	171
2.1 引言	171
2.2 图像投影重建的基本原理	174
2.3 离散图像的傅里叶变换重建法	178
2.4 卷积逆投影法重建图像	180
2.4.1 卷积逆投影法原理	180
2.4.2 重建滤波器的设计	181
2.5 扇束投影的滤波逆投影方法	184
2.6 代数重建法	187
2.7 三维物体投影重建原理	189
2.8 投影重建技术应用简介	190
习题	192
参考文献	194
第三章 图像、视频模型与检索	195
3.1 概述	195
3.1.1 传统的检索方法	195
3.1.2 基于内容的检索	197
3.1.3 CBR 系统简介	198
3.2 图像与视频的数据模型	203
3.2.1 图像数据模型	204
3.2.2 视频数据模型	205
3.3 基于颜色特征的内容检索	213
3.3.1 颜色特征的描述	213
3.3.2 基于颜色特征的检索算法与实现	215

3.4 基于形状特征的内容检索	220
3.4.1 MPEG-7 中形状特征的描述	220
3.4.2 形状检索的实现	221
3.5 基于纹理特征的内容检索	224
3.5.1 纹理特征的描述	224
3.5.2 同类纹理描述符检索算法	226
3.6 基于运动特征的内容检索	229
3.6.1 运动特征的描述	229
3.6.2 基于运动块匹配算法的运动特征检索	231
小结	235
习题	235
参考文献	235

第一章 图像数据压缩

1.1 概 述

一幅实用的数字图像的数据量是非常巨大的,这给图像的传输和存储带来相当大的困难。鉴于此,人们自然会考虑下面的问题,在满足一定图像质量条件下,是否能减少图像的比特数,最少能少到多少,用什么方法减少等。

事实上,减少图像数据量是可能的,因为一般原始图像中通常存在很大的各类冗余度。例如,一般地,相邻的两个像素具有相近的灰度,它们在灰度上存在相关性,这表明图像存在空间性冗余;由于先验知识,人们知道图像中的一部分信息就可推知另一部分信息,例如,某些物体是关于某轴线或点对称的,某物体的一个物面灰度是相同的,这表明这些图像存在结构性冗余;还有,由于人的视觉特性,当图像中某些信息去掉后,对人的视觉和关于图像内容的理解影响不大,比如,人的视觉具有掩盖效应对空缺直线填充的能力,那些因视觉的有限性和联想性可以丢弃的不重要的信息称为心理视觉冗余;另外,用户通常只对原始图像的一部分信息感兴趣,不感兴趣的另一部分可以降低分辨率从而减少关注度冗余;当系统各部分分辨率不一致时,保持某一部分的高分辨率是没有意义的,系统存在匹配性冗余,如当信道的分辨率低于输入图像的分辨率时,降低输入图像的分辨率后输出图像质量变化不大;对于序列图像,相邻帧中同一对应点处的灰度也是相近的,存在相关性,这表明图像序列灰度存在时间性冗余。研究表明,通常还存在着量化性冗余,这种冗余表现为,改变量化判决灰度和重建灰度,用更少的量化等级数可达到同样的视觉效果;还可能存在着所谓的编码性冗余,如果图像的灰度在编码时所用的符号数多于表示每个灰度级实际所需的最少符号数,用这种编码方式得到的图像具有编码冗余。

在保证一定的图像质量和满足任务要求条件下,减少原始图像数据量的处理过程称为图像压缩。

研究表明,原始图像的灰度分布越有规律,图像内容的结构性越强,各像素间的相关性越大,它可能被压缩的数据量就越多。

图像数据实际被压缩多少,图像质量如何,或者说,压缩技术的优劣,取决于所采用的方法与给定图像的结构、分布、特征等匹配的程度。为了使压缩方法更为有效,理论上讲,应首先了解原始图像性质,然后选择方法。

有些原始图像具有简单的图形结构,如工件图像,灰度级较少的气象云图,大幅人头像等;另一些图像它们的图形结构不明显,或很复杂呈现出随机性,但存在统计上的规律性;另有一类图像以上两种特点兼有之;还有一类图像在灰度分布结构上存在相对整体与相对局部的相似性,如蕨类植物图像。以后各节将针对上述四类图像介绍压缩技术。在实际应用中,一个系统通常要对各种类型的图像数据进行压缩处理,因此要求系统的各环节所采用方法和技术通用性要好。

图像数据压缩过程或系统无论采用什么具体结构和技术,从原理上讲有三个基本环节:变换,量化,编码,如图 1.1.1 所示,这三个环节都有数据压缩的任务。首先对原始图像进行变换,然后对变换结果进行量化和编码。这里的变换是广义的,变换的作用是将原始图像在一个新的域中用另一种形式表示,改变图像数据的特性,使在变换域中图像能用较少的像点或比特表示或较少的比特数量化和编码。对变换器的要求是:有效地压缩数据,同时信息保真度要高,具体地讲应是高度去相关的;还要求是广义可逆的,可重建原图像,重建均方误差要小;以及方法简单,经济实用。对量化器的要求是:在一定的允许客观误差或主观察觉图像损伤条件下,总的量化等级数要尽可能的少,同时也应便于实现。编码器的编码方式应和信号或符号的分布特性相适应以求得较大的压缩比且抗扰性强。

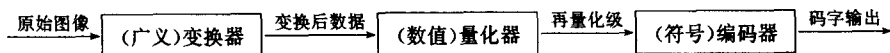


图 1.1.1 图像数据压缩的一般过程

除了用于分析和理解外,图像数据压缩的目的在很多的情况下是为了存储和传输(如图像通信),为了能利用和还原原始图像,通常还要对已压缩的图像数据进行解压缩处理,在原理上通常解压缩是图像数据压缩的逆过程。

设一幅图像在压缩前和压缩后所占的比特数分别为 B 和 B_d ,则压缩比 C 定义为

$$C = \frac{B}{B_d} \quad (1.1.1)$$

每个像素所占的比特数称为比特率,单位是 bit/s,它也是刻画压缩技术或系统的重要性能指标。

图 1.1.2 示出不同压缩比下的图像分辨率,可以看出,压缩比越大图像的分辨率越低。

下面我们将较重要的基本的变换方法、量化方法、编码方法列写在图 1.1.3 中。需要指出的是,图 1.1.3 中有的方法实际上是几种技术的综合,这样只能根据它的关键技术思想进行归类。例如,表中的子带变换、子块编码与位平面编码都是将图像分解,子带变换是将图像在频域中分解,子块编码是在空间域中分解,位平面编码是在灰度域中分解,这里只能根据它们的关键思想分别归类于变换、编码和变换环节。

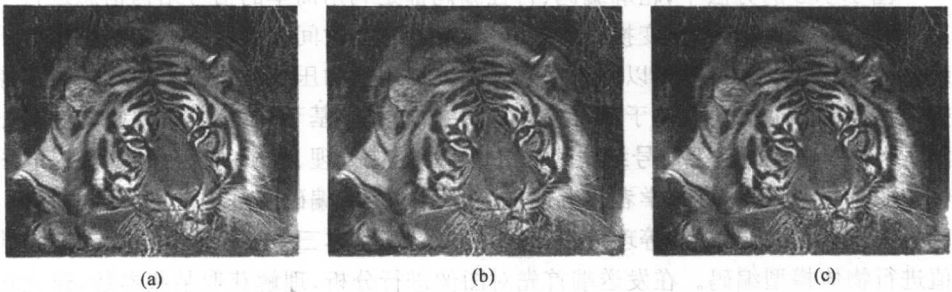


图 1.1.2 图像数据压缩的一个实例

(a) 457×385 的原始图像(175945 字节) (b) 压缩到 29269 字节的图像(压缩比为 6)

(c) 压缩到 10309 字节的图像(压缩比为 17)

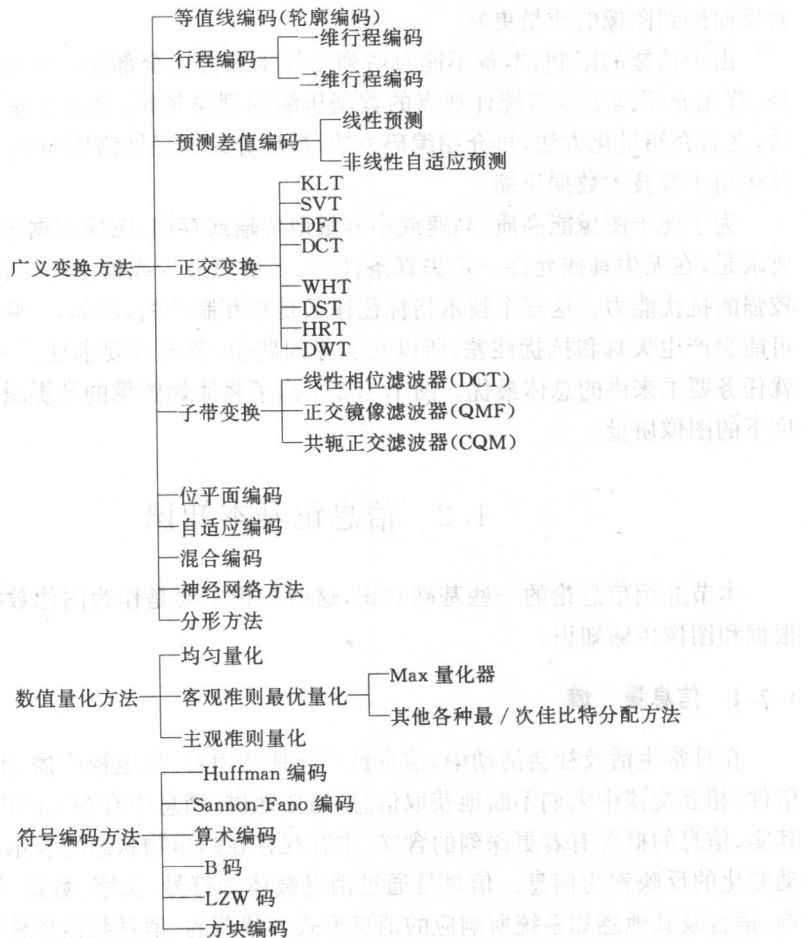


图 1.1.3 图像压缩的一些主要技术

图 1.1.3 的方法中,轮廓编码、行程编码都是利用简单的信号结构信息进行编码;预测差值编码和正交变换编码是分别利用空间、时间和频域中的相关性进行压缩编码的,从学术上讲,可以说是基于概率统计理论的压缩方法。当前及今后图像数据压缩的研究方向是基于信号或信息的结构模型、基于知识、基于视觉特性的方法。分形方法就是基于信号结构的方法。利用图像处理、分析的理论和方法提取图像中的边缘或区域,有的学者相继提出了基于轮廓的编码和基于区域分割的编码,利用模式识别、机器视觉等理论和方法,研究基于物体三维表面形状与运动的估计值进行物体模型编码。在发送端首先对图像进行分析,理解获取某些参数,建立它的模型,然后对它们进行编码,在接收端再基于这些参数和模型重新合成图像。

图像数据压缩的另一个新途径是,接收端在收到高压缩比图像之后采用“图像修复技术”提高图像“美感”,虽然,可能不真实,但只要修复得自然,合乎常理,观看者反而感到图像的质量更好。

由于篇幅的限制,本章不能对所列出技术和方法全部深入介绍。本章的结构是,首先介绍基于概率统计观点的数据压缩的理论依据,然后介绍有关的变换方法,之后介绍量化方法,再介绍编码方法,最后介绍神经网络压缩技术、分形技术以及运用小波技术数据压缩。

为了便于图像能高质、高速或小容量地传输或存储,图像数据压缩的基本技术要求是,在无失真或允许一定失真条件下,尽量用最少或较少的比特数,并且具有较强的抗扰能力。这三个技术指标往往又是相互制约的,例如,一味追求比特数少可能会产生失真和抗扰性差,所以在实际问题中,要统筹要求这三个指标,以达到就任务要求来讲的总体最优。图 1.1.2 示出了将原始图像的数据量压缩到不同程度下的图像质量。

1.2 信息论基本知识

本节介绍信息论的一些基础知识,这些知识主要是作为图像数据压缩的理论根据和图像传输知识。

1.2.1 信息量 熵

在日常生活及社会活动中,信息被广泛地使用着。从电视广播、报纸书刊、来往信件、相互交谈中人们不断地获取信息,通俗地讲,消息中有意义的内容就是信息。其实,信息的概念有着更深刻的含义,主客观存在的事物状态的表示及事物发生运动变化的反映称为信息。信息是通过消息载体如符号、文字、数据、波形、图像、音响、语言及其他感知系统所响应的消息形式等载荷的,消息载体俗称信号。对自然界和社会,人们存在太多的不知和疑惑,一个消息若能提供我们许多原来不知道的

新内容,我们就认为这个消息很有意义,信息量大,收信活动很有价值;反之,消息的内容若是众所周知的事实,我们就认为这个收信活动没有价值,从而消息的信息量为零;若某事件不易发生,一旦发生的消息传来,使人意外,我们也认为这个消息信息量大。将研究对象进行数学抽象,撇开信息的物理载体、信号形式、具体的内容、对人的情感影响及人们的反映,认为收方所获得的信息在数量上等于收方所消除的不确定度。不确定度在概率论中是用随机事件发生的概率来描述的,因此某事件所蕴含的信息量应考虑用该事件出现的概率的倒数的某个函数表示。根据信息量取值的一些约定,如基本信息量是非负的,完全预知或确定的事物蕴含信息量为零,另外,根据生物感知系统对信号响应的模型,以及我们通过不同的途径计算某一消息的信息量时发现,输入按指数变化时输出按线性变化,所以应采用按事件出现概率的倒数的对数度量信息量。如果某事件 x 出现的概率为 $p(x)$,则此事件提供的信息量为

$$I(x) = \log_2 \frac{1}{p(x)} = -\log_2 p(x) \quad (1.2.1)$$

若对数以 2 为底,则信息量的单位为比特(binary unit, bit),若以 e 为底,则为奈特(nature unit, nat),若以 10 为底,则为哈特(hart,以纪念 Hartley)。下面有关论述总是以 2 为底取对数,选用这种定义也符合实际,因如此定义的信息量等于记录信息所用的最小位数。

能够提供信息的事物称为信源,在数学上信源是一概率场。若信源 X 可能产生的信息是 $x_1, x_2, \dots, x_n$, 这些信息出现的概率分别是 $p_1, p_2, \dots, p_n$, 则该信源可表示为

$$X = \begin{Bmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{Bmatrix} \quad (1.2.2)$$

而信源随机出现的符号集 $\{x_i\}$ 用 X 表示,即

$$X = \{x_1, x_2, \dots, x_n\}$$

由单一事件的信息量的定义我们可以得出信源 X 的平均信息量

$$H(X) = \sum_{i=1}^n p_i I(x_i) = - \sum_{i=1}^n p_i \log_2 p_i \triangleq H_n(p_1, \dots, p_n) \triangleq H_n(p) \quad (1.2.3)$$

$H(X)$ 称为信源 X 的熵(entropy)。这是因为信息量的定义内涵与统计物理中的热熵的概念相似,两者的计算式结构也相同。这里规定 $0 \log_2 0 = 0$, 该规定是根据洛必达法则按 $\lim_{p \rightarrow 0} p \log_2 p = 0$ 而建立的。信息熵的单位是: n 制单位/符号,因此根据对数底的不同分别是:比特/符号,奈特/符号,哈特/符号。

由上述关于信息量及平均信息量的定义我们可以得出下面的一些结论:

- 1) 单个事件比较,小概率事件信息量大。
- 2) $H_n(p) \geq 0$, 即 $H_n(p)$ 是非负的。当且仅当某个 i 有 $p_i = 1$, 而其余 $p_i =$