

◆ 视 ◆ 频 ◆ 通 ◆ 信 ◆ 系 ◆ 列 ◆ 丛 ◆ 书 ◆

静止图像编码的 基本方法与国际标准

JINGZHI TUXIANG BIANMA DE JIBEN FANGFA YU GUOJI BIAOZHUN

胡 栋 编著



1-65



北京邮电大学出版社
www.buptpress.com

静止图像编码的基本方法 与国际标准

胡 栋 编著

北京邮电大学出版社
·北京·

内 容 简 介

本书主要介绍目前广泛用于连续色调静止图像编码的三个主要标准,即 JPEG、JPEG-LS 和 JPEG2000。为了便于读者对各个标准的理解,本书将内容分为两大部分:第一部分是关于静止图像编码的基础知识,包括基本理论和标准中涉及到的一些基本概念和主要编码方法;第二部分则详细介绍各个标准的基本算法和码流语法等内容。

本书既可适用于普通高等院校信息类专业高年级本科生和研究生,也可适用于从事图像相关方向的应用开发的工程技术人员。

图书在版编目(CIP)数据

静止图像编码的基本方法与国际标准/胡栋编著. —北京:北京邮电大学出版社,2003
ISBN 7-5635-0715-9

I. 静… II. 胡… III. 图像处理—数据压缩—图像编码—国际标准 IV. TN919.81-65

中国版本图书馆 CIP 数据核字(2003)第 098334 号

书 名: 静止图像编码的基本方法与国际标准
作 者: 胡 栋
责任编辑: 刘 洋
出 版 者: 北京邮电大学出版社(北京市海淀区西土城路 10 号)邮编:100876
电话:(010)62282185 62283578(传真)
经 销: 各地新华书店
印 刷: 北京通州皇家印刷厂
开 本: 787 mm × 1 092 mm 1/16
印 张: 16.5
字 数: 389 千字
印 数: 1—3 000 册
版 次: 2003 年 12 月第 1 版 2003 年 12 月第 1 次印刷

ISBN 7-5635-0715-9/TN·314

定价:29.00 元

如有印装质量问题请与北京邮电大学出版社发行部联系

总 序

进行信息交流的最主要的手段之一就是通信,和人类生产、生活相伴随而产生的通信技术发展到今天,已经从面对面的语言、手势的信息交流,发展到相隔万里的声音、文字及图像信息的交流。

在丰富多彩的图像信息中,以充分表现活动彩色场景的视频图像最为引人注目。因此,视频信息的交流在当今的信息社会中备受欢迎,在现代通信中占据了重要的一席之地。日趋成熟的视频通信技术在很多方面已得到广泛的应用,并给人们的生活、学习、工作带来极大的方便。如用视频通信技术来实现实时现场信息的交流,在英特网上进行视频广播,用卫星视频传输系统建立的远程作战指挥系统,等等。

正是由于图像、视频信息的加入,使得传统的通信系统从单一媒体的传输发展为多媒体传输;同样是由于图像、视频信息的加入,使得传统的通信技术难以胜任多种媒体传输的需要。由此可知,在视频通信领域,既有良好的发展机遇,又存在相当严峻的挑战。本套丛书正是从视频通信这个角度来向读者介绍这一领域中的主要应用技术、实用系统、国际标准、最新发展以及简单的基本理论。当然,本丛书并不仅仅局限于此,还包含一些和视频通信密切相关的内容,如 IP 宽带通信网、静止图像通信等。

必须说明的是,本套丛书并未包括视频通信领域的所有的重要内容,仅涉及这一领域最主要且比较热门的一些专题。例如,如何保证图像、多媒体信息在通信中的安全之类的一些迫在眉睫的技术与应用问题,当前国内外主要的图像通信应用系统,有关视频、静止图像的国际标准、压缩处理技术近年来突飞猛进的发展,代表将来视频通信发展趋势的 IP 宽带通信系统,等等。

本套丛书以从事计算机技术、通信技术以及电视技术的工程技术人员、高等院校的有关专业高年级学生或研究生为主要阅读对象。

本套丛书的编著,从内容上主要体现以下四方面特点:

其一,密切联系实际。本丛书将重点放在经典的和新型的实用系统和实用技术上,对于所介绍的各个应用系统给出它的全貌以及主要的技术内容,对基础原理部分只是简要地提及。

其二,紧跟世界视频通信及相关技术发展的新潮流,在丛书中尽量体现当前视频通信技术方面的新技术、新应用、新系统、新标准和新观点。

其三,本套丛书在内容的编排上,既注重全面性和系统性,力求每一本书都能给读者以清晰简要的全貌;又注重关键技术的细节,对重要的部分、核心的概念给出了不少的实例和图解,力图使读者对它们有较为深刻的理解。

最后,本丛书没有明显的层次结构,不存在先读哪一本、后读哪一本的要求,每一本书都自成体系。之所以以丛书的形式出版,是鉴于这套书内容的内在联系,它们较为全面地向读者介绍了视频通信领域的基本概况。

本套丛书共6本,包括《多媒体信息安全技术与应用》、《图像通信应用系统》、《视频图像编码技术及国际标准》、《静止图像编码的基本方法与国际标准》、《二值图像编码的基本方法与国际标准》和《IP宽带通信网络系统》,将陆续出版。

尽管本丛书的作者们一直从事这方面的科研和教学工作,并在这套丛书的编著中付出了辛勤的劳动,但由于通信和视频技术的发展日新月异,作者个人的视野和水平毕竟有限,再加之编写时间仓促,书中难免存在疏漏和不足之处,真诚地欢迎广大读者予以批评指正。

“视频通信系列丛书”编委会

2003年5月

前 言

生命在不断运动,生活充满着变化,多少美丽的瞬间在我们的眼前稍纵即逝?过去我们无法挽留时间老人匆匆的脚步,常常只能在无数记忆中遐想、惋惜,但今天我们已经可以轻松让时间定格,从容品味生活中那精彩的分分秒秒……

对了,也许你已经知道,这一切都得益于图像压缩编码技术神奇的力量。

图像压缩编码在经过了半个多世纪的发展后,已经从研究室走进了我们的生活。当我们坐在桌旁,用 Web 浏览器浏览 Web 网站各种精彩的图片时;当我们举家外出郊游,用数码相机拍下迷人的风景时;当我们进入医疗影像室,用数字 CT 做成像检查时;当“神舟五号”飞船从遥远的太空给我们发来神秘的宇宙图像时……,我们都在享受图像编码技术带来的种种便利。

推动图像编码技术走向大规模应用的基础是图像压缩编码的国际标准。从 20 世纪 80 年代以来,国际标准化组织 ISO 和国际电联 ITU 的图像专家工作组针对各种不同类型的图像和应用,制定了一系列国际标准,其中包括:视频压缩编码的 H.261/263/264, MPEG-1/2/4;二值图像压缩编码的 JBIG, JBIG2;静止图像压缩编码的 JPEG, JPEG-LS, JPEG2000 等。在今天各种数字图像设备中,我们都能看到它们的身影。把有关的国际标准及相关技术介绍给对图像编码感兴趣和从事相关行业的读者是笔者的初衷。

本书主要介绍了目前三个主要的连续色调静止图像(即所谓自然图像)编码标准,即 JPEG、JPEG-LS 和 JPEG2000,重点是各个标准的基本算法和码流语法等内容。其中, JPEG 标准是第一个连续色调静止图像压缩编码的国际标准。由于 IJG 小组在免费软件方面所做的努力, JPEG 标准的核心——JPEG 基本系统得到了广泛的普及应用,是目前主流的编码标准。 JPEG-LS 主要用于一些保真度要求高、但实现相对简单的场合。 JPEG2000 是最新的静止图像压缩编码标准,其基本系统在制定时秉承了 JPEG 基本系统的 Royalty Free 思想,但编码效率更高,而且提供了渐进显示、分等级显示和感兴趣区域等功能,可以被认为是 JPEG 的更新版本,一些图像显示软件上已经增加了支持 JPEG2000 压缩格式的插件。

本书编写过程中得到了许多专家、同行的帮助,朱秀昌教授在编写过程中给予了宝贵的意见,刘泓女士完成了许多文字录入和编辑工作,在此谨表示衷心谢意。此外,张颖希、王滨、张希光、邢有涛、黄昌龙等同学帮助绘制了许多插图在此也一并表示谢意。

本书的编写对笔者是一个巨大的挑战,同时也是一个学习和提高的机会。在编写过程中,笔者参阅了大量相关的国际标准和文献资料,力图使介绍的内容更加准确,对读者更具参考价值。限于编者水平,同时由于编写时间较为仓促,书中一定存在不妥之处。在此,笔者非常恳切地希望得到各位读者的批评指正,并通过相互交流使本书内容得以逐步完善。

胡 栋

2003 年 10 月于南京

目 录

第1章 绪 论

1.1 图像、图像压缩编码与静止图像压缩编码	1
1.2 静止图像压缩编码及其国际标准的发展简史	3
1.3 内容安排	4

第2章 静止图像压缩编码基础

2.1 编码模型	5
2.1.1 理想信源的无失真编码	5
2.1.2 图像信源的建模与编码	6
2.2 Huffman 编码	7
2.3 算术编码	8
2.4 游长编码	9
2.5 预测编码	9
2.6 变换编码	10
2.7 比特面与比特面编码	11
2.8 具有多种表示能力的编码	11
2.8.1 渐进显示能力	12
2.8.2 分辨率分等级能力	14
2.9 小波变换编码与零树编码	15
2.9.1 图像的小波变换与多分辨率、渐进显示	16
2.9.2 小波变换图像编码	18
2.9.3 零树编码	18

第3章 JPEG 压缩编码标准

3.1 概 述	23
3.1.1 JPEG 建议中的基本要素	23
3.1.2 有损与无损压缩	25
3.1.3 基于 DCT 的编码	25
3.1.4 无损编码	27

3.1.5	工作模式	27
3.1.6	熵编码选择项	29
3.1.7	采样精度	29
3.1.8	多分量控制	29
3.1.9	压缩数据结构	31
3.1.10	图像、帧与扫描	32
3.2	JPEG 标准中的数学定义	32
3.2.1	输入原图像	32
3.2.2	原图像数据编码顺序	33
3.2.3	DCT 压缩	34
3.2.4	点变换	36
3.2.5	无损和分等级工作模式中的算术处理	36
3.3	压缩数据格式	36
3.3.1	压缩数据格式规范概述	37
3.3.2	通用的顺序处理和渐进处理的语法	40
3.3.3	分等级编码语法	49
3.3.4	压缩图像数据的简约格式	51
3.3.5	表格数据定义的简约格式	51
3.3.6	小 结	51
3.4	Huffman 码表定义	53
3.4.1	Huffman 码表定义的标记字段	54
3.4.2	将 Huffman 码表定义转换为码表和码字长度	54
3.4.3	字节中的比特顺序	54
3.4.4	产生用于定义 Huffman 码表的列表的过程的示例	56
3.4.5	Huffman 码表的示例	57
3.5	编码器和解码器的控制过程	59
3.5.1	编码器的控制	59
3.5.2	解码器的控制	61
3.6	基于 DCT 的顺序工作模式	62
3.6.1	编码过程	63
3.6.2	解码过程	68
3.7	基于 DCT 的渐进工作模式	72
3.7.1	编码过程	72
3.7.2	解码过程	79
3.8	无损工作模式	79
3.8.1	编码过程	80

3.8.2 解码过程	81
3.9 分等级编码工作模式	82
3.9.1 分等级编码	83
3.9.2 分等级解码	85

第4章 JPEG-LS 压缩编码标准

4.1 概 述	88
4.1.1 编码概要	88
4.1.2 原图像定义	89
4.1.3 编码处理过程	90
4.1.4 解码处理过程	91
4.1.5 多分量图像的编码	91
4.1.6 数据流内容	91
4.1.7 交换格式	91
4.1.8 其 他	92
4.2 单分量编码过程	92
4.2.1 编码参数和压缩图像数据	93
4.2.2 初始化过程和一些约定	93
4.2.3 上下文的决定	95
4.2.4 预 测	97
4.2.5 预测误差的编码	99
4.2.6 变量的更新	100
4.2.7 游长模式的编码	101
4.2.8 编码过程的流程	105
4.3 多分量图像的编码过程	107
4.3.1 概 述	107
4.3.2 行交织模式	107
4.3.3 采样交织模式	108
4.3.4 最小编码单元 MCU	109
4.4 压缩数据的格式	109
4.4.1 概 述	110
4.4.2 JPEG-LS 的通用编码语法	110
4.4.3 压缩图像数据的简约格式	116
4.4.4 用于表格说明数据的简约格式	116
4.5 编码控制过程	116
4.5.1 编码一幅图像的控制过程	116

4.5.2 编码一帧图像的控制过程	117
4.5.3 编码一个扫描的控制过程	117
4.5.4 编码一个重启区间的控制过程	118
4.5.5 对一个最小编码单元 MCU 的编码控制过程	118
4.6 解码控制过程	119
4.6.1 处理流程	119
4.7 编解码处理过程的描述	121
4.7.1 上下文建模	121
4.7.2 正常模式下的编码	122
4.7.3 游长模式的编码	124
4.8 编解码实例和指南	125

第5章 JPEG2000 压缩编码标准

5.1 概 述	139
5.2 JPEG2000 基本系统的组成及工作原理	140
5.3 码流语法	144
5.3.1 码头和标记字段	144
5.3.2 标记字段信息	146
5.3.3 码流结构	148
5.3.4 分界标记	150
5.3.5 固定信息的标记字段	151
5.3.6 功能标记字段	153
5.3.7 指针标记字段	160
5.3.8 比特流内部的标记字段	164
5.3.9 信息标记字段	165
5.4 数据排列	166
5.4.1 将图像分为不同的分量	166
5.4.2 将图像分为若干叠块和叠块分量	167
5.4.3 分量与参考栅格映射的示例	168
5.4.4 将叠块分量分为若干不同的分辨率和子带	171
5.4.5 将分辨率分为围地	171
5.4.6 将子带分为编码子块	172
5.4.7 层次	172
5.4.8 数据包	173
5.4.9 数据包头的信息编码	173
5.4.10 叠块数据与叠块分部	179

5.4.11 渐进次序	179
5.5 算术熵编码	182
5.5.1 二进制编码	182
5.5.2 算术编码器的描述	183
5.5.3 算术解码过程	190
5.6 小波变换系数比特的建模	195
5.6.1 编码子块的内部扫描模式	195
5.6.2 变换系数比特与系数的重要性	196
5.6.3 比特面上的解码通过	196
5.6.4 初始化与终止	200
5.6.5 抗误码分割符号	202
5.6.6 选择性算术解码的旁路	202
5.6.7 垂直方向因果性上下文的形成	203
5.6.8 编码子块的编码流程图	204
5.7 量 化	205
5.7.1 标量系数的反量化	206
5.7.2 标量系数的量化	208
5.8 叠块分量的离散小波变换	209
5.8.1 概 述	209
5.8.2 离散小波反变换	209
5.8.3 离散小波正变换	215
5.9 直流电平的位移和分量变换	220
5.9.1 叠块分量的直流电平位移	220
5.9.2 可逆分量变换 RCT	221
5.9.3 不可逆分量变换 ICT	222
5.9.4 色差分量的亚抽样和图像的参考栅格	222
5.10 带有感兴趣区域 ROI 的图像编码	223
5.10.1 Maxshift 方法的描述	223
5.10.2 ROI 的模板生成	224
5.10.3 关于 ROI 编码的注释	225
5.10.4 对 Maxshift 方法解释的一个示例	226
附 录	227
参考文献	249

第1章 绪 论

1.1 图像、图像压缩编码与静止图像压缩编码

什么是图像？其实我们天天都在和图像打交道——我们每天接收的外部信息中，大部分内容是通过视觉系统接收的：我们用肉眼来观察、了解外部世界，所见的就是景物的图像。虽然实际的景物是客观存在的实体，但只有当光辐射能量照在物体并经过它的反射或透射，或者发光物体本身可以发出光能量，才能在人的视觉器官中重现出的物体的视觉信息——图像。如果没有了光源的照射，例如设想你呆在一间漆黑的小屋，“伸手不见五指”，就不可能看到任何物体。照片、电影、电视、图画等都属于图像的范畴。

据分析，在人类的感观所接收的外部信息中，声音和图像占据了其中的主要部分。其中，通过耳朵接收的听觉信息约占总信息量的 20% 左右，而通过人眼的视觉接收的信息则高达 60% 以上。由于图像信息具有直观性强、信息量大的特点，并且特别适应人们在信息交流中所需要的直接、亲近等生理和心理的要求特点，图像信息在通信中的地位非常突出。人们用“百闻不如一见”来表示图像信息交流的重要性和迫切性。然而，由于图像信息是高维信息（二维或二维以上），图像信号的内容复杂，它的数据量非常大，这成为图像信息交流的主要障碍。例如，一幅 256×256 大小，8 比特/像素的黑白图像需要 8 秒钟才能在一个话音信道上传送；而数字化后的活动图像信号数据率可能高达 200 Mbit/s，需占用近千条普通话路。因此，在通信领域中，图像通信的发展比起我们熟悉的语音通信——电话的发展——在时间上要滞后得多。

为了实现图像的存贮和传输，图像的压缩编码成为技术的关键。图像的压缩编码最直接的目的就是尽量降低一幅图像的数据量，同时：(1) 或者保持图像的内容不变；(2) 或者使内容的差别在一定的控制范围内；(3) 或者保证一定观察效果的主观质量。第一种情况被称为无失真或无损编码，后两种情况被称为有限失真或有损编码。这几种编码具有不同的应用场合，例如对于一些医学诊断和航天探测图像可能需要无损压缩编码，而对于供一般观赏用的图像多数都采用有损压缩。一般地，有损压缩的效率比无损压缩的效率高许多，在压缩比达到数十倍甚至更高时也可能看不出有明显的质量损伤。

由于我们所处的是一个丰富多彩的世界，图像的内容是复杂多变的。为了能有效地对图像信息进行压缩，必须首先根据图像的特征以及应用的需要进行分类，然后进行行之有效的处理。一般而言，对图像的分类可以从以下几个方面考虑：

(1) 根据图像色调的不同，把图像分为单色（黑白、灰度）图像和彩色图像。在本书中，我们只考虑（空间）二维数字图像，此时灰度图像从数学上说就是一个二维数组，也称

之为单分量;彩色图像则包含了多个彩色分量,在国际标准中也称为多分量。每个分量都是一个二维数组,例如,显示器通常使用的是 RGB 三基色彩色图像,而在彩色印刷时通常使用 CMYK 四个分量,对于资源卫星的探测图像,可能使用了多个探测频段,每一个频段均对应一个图像分量。对多分量图像进行编码实际上就是分别对每个分量进行编码,但为了获得最佳的压缩效率,通常在编码之前使用合适的变换去除分量之间的相关性。例如对于 RGB 彩色图像通常变换为 YUV 彩色图像。

(2) 根据每个分量像素的幅度层次不同,把图像分为连续色调图像和二值图像。当图像的分量只有两种幅度等级时,则为二值图像;当为两种以上幅度时,则称为连续色调图像。一般情况下,连续色调数字图像每个分量的电平幅度有 $2^P = 256$ 种, $P = 8$ 称为比特深度;但对于一些特殊的应用,例如医学中的人体软组织图像,可能要求 10 比特甚至 12 比特。由于连续色调图像与二值图像在信号的复杂性和统计特性等方面差别很大,因此需要分别使用不同的编码方法。此外,使用比特面编码技术,可以把连续色调图像转换为二值图像,然后使用二值图像编码方法。

(3) 根据图像内容,即景物随时间变化性质的不同,将图像分为静态图像和活动图像。以灰度图像为例,静态图像只需要一幅图像就足以代表所考虑的时间段,因为图像的内容保持不变(设摄像机引入的背景噪声可以忽略);而活动的灰度图像由一系列按时间顺序排列的给定时间段内的二维图像序列,即序列图像。与此相对应,编码可分为静止图像编码和活动图像编码。静止图像编码就是以每一幅图像为单位独立地进行编码,而活动图像编码则要求充分考虑序列图像的相邻图像之间的相关性,例如使用帧间预测等技术,提高序列图像的压缩效率。当然,静止图像编码方法可以用于序列图像的压缩,因为序列图像本身是由一幅幅图像构成的,我们可以把它们看作在各时刻的“凝固”图像;但采用这种方法的压缩效率将会大大低于活动图像压缩方法,因为它只利用了每幅图像自身内部的相关特性而没有充分利用图像之间的相关特性。

以上我们已经明确了与静止图像压缩编码有关的概念。随着图像编码应用的普及,人们对于静止图像编码除了要求高数据压缩率以外,还要求编码能满足以下要求:

首先,编码能提供足够高的图像清晰度。在大部分应用中,静止图像压缩的最终目标都用于图像的静态显示,此时人眼对图像的细节观察得更仔细,因此希望图像的尺寸(分辨率)足够大,逼真度(准确性)足够高。例如:高质量的照片扫描图像尺寸可能高达 2048×2048 以上;在一些特殊的场合要求图像能被无失真地压缩编码。

其次,编码器能提供灵活的数据组织和表示功能,以满足特殊的显示要求。当图像尺寸较大而传输带宽较窄时,采用普通的顺序扫描的编码传输方式将使观察者等待很长的时间才能看到整幅图像。符合人的视觉心理的方式则是先尽快提供一个原图像的一个粗略的图像,然后随着时间的增加不断传送新的码流,它们在解码后能和已经解码的图像叠加,使重建的图像质量不断改善。具有这种能力的编码被称为逐渐显示的图像编码,或渐进编码,现在也称之为嵌入式编码。这种编码在压缩码流的组织上进行了分层,其核心数据提供基本图像质量,被最先传输,外层数据提供质量的增量,在后面依次传输。此外,目前一种新的要求是对图像中某些感兴趣的区域进行特殊的编码处理,包括:编码中对这部分区域要求较高的清晰度,而其他部分(背景)区域则用较低的精度进行编码;或者在传输

时首先突出这部分区域(前景)的显示,然后根据要求逐步增加背景的清晰度。这种区域称为感兴趣区域 ROI。例如新的 JPEG2000 标准就提供了对这一功能的支持。

第三,编码码流能适应抗误码传输的要求。由于静态图像的显示时间较长,因传输误码引起重建图像的差错对视觉观察的效果影响更大,因此要求编码中采用一定措施减少误码的影响。目前常用的做法是采用再同步和误码隔离等措施。例如:JPEG 中使用重启标记 RST 对码流进行重新置位, JPEG2000 用编码子块 Code-block 限制误码的作用范围。

在具体实现时,还要考虑一些实际情况,例如实现的复杂度以及成本等因素。此外硬件实现与软件实现也会有很多差异之处。

1.2 静止图像压缩编码及其国际标准的发展简史

图像编码的发展至今已走过了半个多世纪的历程。在 20 世纪 40 年代末脉冲编码调制技术 PCM 出现后不久,人们就开始了对电视信号的数字化研究。经典的图像编码方法基于 Shannon 信息论,其中最基本的 Huffman 编码(熵编码)、预测编码和变换编码理论就产生、发展于 20 世纪的五六十年代,并且影响至今,在目前已知的图像压缩编码的国际标准中,仍然被普遍使用。之后,人们在探索一些新的高效的编码方法方面不断取得进展。例如,算术编码已经被许多编码标准所采用,它的一个最大的优点就是可以在编码过程中调整信号的概率模型,是一类高效、普适的无失真编码;其他一些编码方法,不考虑其出现的时间顺序,包括矢量量化、方块截断编码、比特面编码、亚抽样与内插、子带编码和小波变换编码等。此外,由于人们对视觉特性的认识更加深入,出现了许多结合人的视觉系统特性 HVS 和多种编码算法的综合算法,编码效率被不断改善。20 世纪 90 年代以来, Kunt 等人提出了所谓第二代编码的概念。他们认为传统的编码方法是基于信号波形的方法,衡量编码算法的效果主要以重建信号与原始信号的波形一致性程度为评价准则;而新一代编码方法则是基于对象模型的描述方法,因此有可能获得比经典的算法高得多的压缩效率。有代表性的方法包括分形编码、模型基编码、轮廓编码等。在一定的实验条件下,这些编码方法可以取得非常出色的编码性能,但如果考虑实用化,它们还受到许多限制,其中最主要的一点就是处理的复杂度,而且在对一般的自然图像进行处理时,所获得的编码增益也没有理论上预期的好。与此相反,一类仍属“经典”的编码方法——小波变换编码——自 90 年代以来日益受到人们的重视,特别是 J.M.Shapiro 提出嵌入式零树小波变换编码 EZW 算法,向人们展示了它对之前的各种算法优异的压缩性能,更提供了天然的多尺度、多分辨率的图像描述方法。在此之后, A.Said 等人提出了改进的所谓分等级树的集分割 SPIHT 算法,在运算复杂度显著降低的同时,获得了与 EZW 算法相当或更好的压缩率。由此确立了小波变换在新的图像编码标准中的重要地位。

自 20 世纪 80 年代以来,无论是从技术的发展还是从社会的需求来看,图像编码技术已经逐步进入了较大范围的应用阶段。但是由于没有统一的压缩算法和码流格式,在进行图像信息的交流中遇到了很多困难。为了解决这一问题,国际标准化组织 ISO 和国际电报电话咨询委员会 CCITT(现国际电联的电信委员会 ITU-T)的图像专家组于 1986 年前后开始进行标准的制定,其主要目的包括:(1)提供高效的压缩编码算法;(2)提供统一的

压缩数据流格式。经过对多个方案进行的大量严格的实验测试,从算法压缩性能到实现的复杂度等综合因素的考虑比较之后,最终形成了两个著名的里程碑式的国际标准,这就是人们熟知的用于连续色调静止图像压缩编码的 JPEG 标准和码率为 $p \times 64 \text{ kbit/s}$ 的数字视频压缩编码标准 H.261 建议。这两个标准中都选择了经典的编码方法,但通过标准测试图像确定了最佳的编码参数。两者均使用 8×8 的 DCT 变换、均匀量化和 Huffman 编码等处理,但 H.261 中使用了带有运动补偿的帧间预测编码,以提高视频信号的压缩效率;而 JPEG 建议中推荐使用基于人的视觉阈值的量化矩阵,以获得更好的主观图像质量。这两个标准的制定颁布,极大地推动了图像通信的大规模普及应用,而这种大规模的应用又对图像编码技术提出了新的、更高的要求,例如:更高的压缩性能、使用更加灵活、表现能力更加丰富等。此外,随着 Internet 和无线通信的发展普及,还要考虑更加优良的抗误码性能等。由此促进了标准化工作的进一步发展。迄今为止,国际标准化组织和国际电联已经制定了适用于不同类型图像的压缩编码标准。例如对于简单的二值图像压缩标准,由早期用于三类传真机 G3 的 T.4 建议和四类传真机 G4 的 T.6 建议,发展到新的 JBIG、JBG2 标准等;对于数字视频的压缩标准有 H.261、H.263、H.264、MPEG-1/2/3/4 等,对于连续色调静止图像压缩标准有 JPEG、JPEG-LS、JPEG2000 等。这些国际标准和建议的制定满足了不同类型数字图像传输的应用要求。

1.3 内容安排

本书的主要内容是关于连续色调静止图像的压缩编码的三个重要的国际标准。目前在 Internet 上的图片浏览和下载、电子图书出版业、医学图像处理以及数字照相机等诸多领域,这三个编码标准已经被普遍采用。在一些有特殊要求的场合,静止图像编码标准也可用于数字视频的压缩,典型的应用是视频编辑器。虽然这样做牺牲了一部分压缩效率,但优点是可以准确定位到每一个视频帧。

本书的内容是这样安排的:第 2 章将首先介绍有关静止图像压缩编码的基础知识,主要对标准中所使用的编码方法进行简要的介绍,包括预测编码、DCT 变换编码、Huffman 编码、算术编码、游长编码、比特面编码、多分辨率编码、小波变换编码、嵌入式零树小波编码 EZW 以及 SPIHT 等。这些内容的介绍将有助于对后面压缩标准的理解。第 3 章将介绍 JPEG 标准,主要是 JPEG 基本系统以及采用 Huffman 编码的编码模式。有关采用算术编码的编码模式没有包括在这里,因为在 JPEG 的实际应用中,基本都使用 Huffman 编码。第 4 章将介绍 JPEG-LS 标准。这是一种无损及受控失真下的压缩编码标准,特点是信号模型简单,易于实现。目前在一些特殊的环境中采用这一标准。例如,在医学图像的成像与传输的国际标准 DICOM3.0 中,就要求支持 JPEG-LS。第 5 章将介绍最新的 JPEG2000 标准,主要包括了 JPEG2000 的基本系统,即 JPEG2000 的第一部分。该标准采用了新的变换——离散小波变换,能够实现的压缩性能和功能比其前身——JPEG 标准——均有明显的改进和增强。

第2章 静止图像压缩编码基础

本章将介绍静止图像编码的一些基础理论和概念,其中涉及的编码方法主要针对目前的几个静止图像压缩编码的国际标准。笔者在此不作严格的理论分析和讨论,仅仅将算法的基本概念和方法作比较浅显的介绍,目的是有助于对国际标准相关内容的理解。如果感兴趣的话,建议读者直接查阅更加深入的编码理论专著和专题论文。另外,以下所介绍的内容均为数字图像的编码与处理,笔者假定读者已经掌握了有关图像数字化的预备知识。

2.1 编码模型

2.1.1 理想信源的无失真编码

根据信息论可知,当信源的特性,即信源符号的概率分布已知时,就确定了编码可能达到的码率下界。对于离散信源,可以将其分为两类:一种是简单的离散无记忆信源;另一种就是离散有记忆信源。

对于离散无记忆信源,将信源抽象地描述为信源 X 从 K 个符号集中 $\{a_1, a_2, \dots, a_k\}$ 发出符号序列,在时刻 N 发出符号 u_N 与前面发出的符号 $u_1, u_2, \dots, u_{N-1}$ 无关。对无记忆信源编码时,无失真编码的每个符号的平均码长,即码率的下界是该信源的一阶熵 H_1 。在编码时采用不等长码字进行编码,也称变字长编码 VLC,原则是给概率大的符号用短码,概率小的符号用长码,使最终的平均码长最短^{*}。为了使不等长码字满足惟一可译和实时译码的要求,不等长编码采用前缀码,也称为非续长码,它的构造特点是,任何一个码字都不为符号集中任何其他符号码字的前缀。设 $\bar{L}$ 为不等长编码的平均码长,则有以下不等式成立:

$$H_1 \leq \bar{L} < H_1 + 1 \quad (2.1)$$

其中,一阶熵 H_1 为:

$$H_1 = - \sum_i P(a_i) \log P(a_i) \quad (2.2)$$

$P(a_i)$ 为信源发出符号 a_i 的概率。由此可见,平均码长最多不会超过信源熵 1 比特。在编码系统中,无失真(无损)编码也称为熵编码。

对于离散有记忆信源,情况就复杂得多,信源 X 由 K 个符号集 $\{a_1, a_2, \dots, a_k\}$ 发出符号序列,在时刻 N 发出符号 u_N 与前面发出的符号 $u_{-m}, \dots, u_1, \dots, u_{N-1}$ 的情况有关。目前

^{*} 本书中只涉及二进制码字,码字单位为比特。