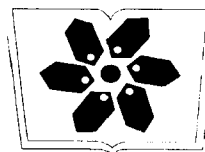


余松煜 张文军 孙 军 编著

现代 图像信息压缩技术

科学出版社





中国科学院科学出版基金资助出版

现代图像信息压缩技术

余松煜 张文军 孙 军 编著

科学出版社

1998

00/3/22

内 容 简 介

数字图像压缩编码是在经典信息论中的信源编码理论上发展起来的“古老而又新兴”的学科。随着数字化信息社会的到来,如何有效地压缩数字(活动)图像的庞大数据量,以方便图像信息的传输和存贮,已成为整个电子信息技术领域必须研究解决的首要问题之一。

本书根据作者近10年来从事数字图像压缩编码的教学和实际开发经验,系统地介绍了图像压缩编码的基本原理和方法,并从实用性角度重点介绍了图像压缩编码的国际标准,以及具备实际应用意义的图像压缩编码硬件系统。对近年来在该领域中国内外所研究的主要算法也作了一定的介绍。

本书可作为信息工程、通信工程、广播电视、计算机应用、信号处理等专业高年级本科生或研究生的教学用书,也可供上述领域的研究开发人员参考。

图书在版编目(CIP)数据

现代图像信息压缩技术/余松煜,张文军,孙军编著.-北京:科学出版社,1998

ISBN 7-03-006603-0

I. 现… II. ①余… ②张… ③孙… III. 图像编码 IV. TN919.8

中国版本图书馆 CIP 数据核字(98)第 06276 号

科学出版社 出版

北京北京东黄城根北街16号
邮政编码:100717

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1998年11月第 一 版	开本: 787×1092 1/16
1998年11月第一次印刷	印张: 11 1/2
印数: 1—2 000	字数: 261 000

定价: 20.00 元

(如有印装质量问题,我社负责调换〈新欣〉)

19981116

前 言

随着现代信息社会对通信业务要求的不断增长,图像通信与通信网容量的矛盾日益突出.特别是具有庞大数据量的数字图像通信,更难以传输与存贮,极大地制约了图像通信的发展,已成为图像通信发展中的“瓶颈”问题.图像压缩编码的目的就是要以尽量少的比特数表征图像,同时保持复原图像的质量,使它符合特定应用场合的要求.

数字图像通信有数字通信的一系列优点,如:可以中继传输和多次复制,而不会造成噪声和非线性失真的累积;便于进行加密;便于用大规模集成电路芯片实现,制作方便、成本低、可靠性高;便于和计算机联网等.然而数字图像通信却迟迟未得到推广和应用,其主要原因就是数字图像的数据量非常巨大.实际上,图像编码研究最早始于20世纪40年代末,但图像编码技术真正走向实用化、产业化并得到飞速发展,则是最近10年的事,其标志就是国际上集图像编码40年研究成果制订的一系列图像编码国际标准.同时,超大规模集成电路工艺的发展,使得研制出集成度更高、性能更好的图像编码专用芯片成为可能,并迅速转化为各种新产品,从而迎来了数字图像通信的黄金时代.现在的情况已不是简单地满足于某些用户的图像压缩的要求,而是正在以图像编码为主要核心技术大规模地积极开拓新的产品和应用领域,并已开始影响到人们的日常生活方式.其中最具代表性的应用就是数字电视,如数字会议电视系统、VCD、DVD、数字卫星电视广播机顶盒,以及美国的数字高清晰度电视和欧洲的DVB等等.目前,图像压缩编码正在寻求新的技术突破以使视听信息融入计算机和电话,真正促成实现广播、通信、计算机的相互“汇聚”.可以预见,图像压缩编码将是正在建设的数字信息化社会所依赖的主要技术基础之一.

本书是作者在上海交通大学图像所近10年来从事数字图像压缩编码的研究生教学 and 实际研究开发工作的基础上编写的.书中较全面系统地介绍了图像压缩编码的基本原理和方法,并且结合作者主持和参与研制的数字会议电视系统、国家重大科技产业工程“高清晰度电视功能样机”视频编码器和信源解码器等项目的实际开发经验,从实用性角度重点介绍了图像压缩编码的国际标准以及具备现实应用意义的图像压缩编码硬件系统.

本书大部分章节是在余松煜为其研究生开设的“图像编码”课程讲稿基础上修补增订而成的.其中,第七章由张文军编写,第五章由孙军编写,第四章一部分由傅立言编写.谭经薇协助完成了本书部分内容的打印工作,在此谨向他们表示深切的谢意.

全书初稿承蒙北京大学无线电电子学系徐孟侠教授审阅,提出了许多宝贵的修改意见.作者借此机会对徐教授多年来就学术研究和学术动态所提供的帮助表示衷心的感谢.

本书的出版得到了科学出版社出版基金的资助,谨致谢忱.

由于作者学识有限以及篇幅的限制,难免有不够严谨和错漏之处,敬请读者指正.

余松煜 张文军 孙 军

1998年1月于上海

常用术语缩略表

ADSL(Asymmetric Digital Subscriber Line)	不对称数字用户线
ASIC(Application Specific Integrated Circuit)	专用芯片
ATM(Asynchronous Transfer Mode)	异步转移模式
ATSC(Advanced Television System Committee)	先进电视系统委员会
BMA(Block Matching Algorithm)	块匹配法
BTC(Block Truncation Coding)	方块截尾编码
CATV(Cable Television)	有线电视
CCIR(Consultative Committee of International Radio)	国际无线电咨询委员会
CD-ROM(Compact Disc-Read Only Memory)	只读光盘
CIF(Common Intermediate Format)	公共中间格式
CSPS(Constrained System Parameter Stream)	限定的参数流
DAB(Digital Audio Broadcasting)	数字音频广播
DCT(Discrete Cosine Transform)	离散余弦变换
DFT(Discrete Fourier Transform)	离散傅里叶变换
DPCM(Differential Pulse Code Modulation)	差值脉冲编码调制
DTS(Decoding Time Stamp)	解码时间标志
DVB(Digital Video Broadcasting)	数字视频广播
DVD(Digital Versatile Disc)	高密度数字通用光盘
FIR(Finite Impulse Response)	有限冲击响应
GOB(Group Of Blocks)	宏块组
GOP(Group Of Pictures)	图像组
HDSL(High-speed Digital Subscriber Line)	高速数字用户线
HDTV(High Definition Television)	高清晰度电视
HVS(Human Visual System)	人类视觉系统
IFS(Iterative Function System)	迭代函数系统
IIR(Infinite Impulse Response)	无限冲击响应
IRD(Integrated Receiver/Decoder)	综合接收机/解码器
ISDN(Integrated Services Digital Network)	综合业务数字网
JPEG (Joint Photographic Experts Group)	联合图片专家组
LDTV (Low Definition Television)	低清晰度电视
MB (Macro Block)	宏块
MC(Motion Compensation)	运动补偿
MDL(Minimal Description Length)	最小描述长度准则
MDU (Minimum Data Unit)	最小数据单元

ME(Motion Estimation) 运动估值
MHC(Modified Huffman Code) 修正霍夫曼码
MPEG(Motion Pictures Expert Group) 活动图片专家组
MTF(Modulation Transfer Function) 调制转移数
NVOD (Near Video On Demand) 准点播电视
PCM (Pulse Code Modulation) 脉冲编码调制
PES(Packetised Elementary Streams) 打包基本流
PTS(Presentation Time Stamp) 显示时间标志
QAM(Quadrature Amplitude Modulation) 正交幅度调制
QMF(Quadrature Mirror Filters) 正交镜像滤波器组
RLC(Run Length Coding) 行程码
SCR(System Clock Reference) 系统时钟基准
SDTV(Standard Definition Television) 标准清晰度电视
SQ(Scale Quantization) 标量量化
STC(System Time Clock) 系统时钟
VCD(Video-CD) 数字激光视盘
VLC(Variable Length Coding) 变长码
VLSI(Very Large Scale Integration) 超大规模集成电路
VOD(Video On Demand) 点播电视
VQ(Vector Quantization) 矢量量化
VSB(Vestigial Side-Band) 残留边带
WHT(Walsh-Hadamard Transform) 沃尔什-哈德玛变换

目 录

前言

常用术语缩略表

第一章 概论	(1)
1.1 图像压缩编码的必要性和可能性	(1)
1.2 图像压缩编码一般框图与分类	(2)
1.3 图像压缩编码技术的发展历史与现状	(6)
1.4 图像压缩编码技术的应用	(8)
1.5 参考文献介绍	(11)
参考文献	(12)
第二章 图像压缩编码基本原理和方法	(13)
2.1 图像压缩编码的理论基础	(13)
2.2 预测编码	(21)
2.3 变换编码	(28)
2.4 量化	(40)
2.5 变长码技术	(46)
参考文献	(52)
第三章 运动估值和运动补偿	(53)
3.1 象素递归法	(53)
3.2 块匹配法和一致性块匹配法	(55)
3.3 多分辨率运动估值	(60)
3.4 运动补偿帧内插	(63)
参考文献	(65)
第四章 图像编码的其它算法	(66)
4.1 模型编码	(66)
4.2 金字塔编码	(71)
4.3 子带编码	(73)
4.4 小波变换编码	(82)
4.5 矢量量化编码	(89)
4.6 人工神经网络在图像编码中的应用	(90)
4.7 数学形态学方法用于二值图像编码	(93)
4.8 分形图像编码	(95)
参考文献	(104)
第五章 图像压缩编码的国际标准	(105)
5.1 概述	(105)

5.2	会议电视和 H. 261	(108)
5.3	MPEG-1 标准	(117)
5.4	MPEG-2 标准	(122)
5.5	甚低码率图像编码国际建议 H. 263	(132)
5.6	JPEG 标准	(134)
	参考文献	(140)
第六章	图像压缩编码实时硬件系统	(141)
6.1	图像编解码 ASIC 介绍	(141)
6.2	会议电视系统视频编解码器控制策略举例	(152)
6.3	本章小结	(158)
	参考文献	(159)
第七章	高清晰度电视图像压缩编码	(160)
7.1	高清晰度电视简介	(160)
7.2	常用的高清晰度电视图像压缩编码方法	(163)
7.3	实际应用范例	(169)
	参考文献	(176)

第一章 概 论

1.1 图像压缩编码的必要性和可能性

图像压缩编码的目的是以尽量少的比特数表征图像,同时保持复原图像的质量,使它符合预定应用场合的要求.压缩数据量,提高有效性是图像压缩编码的首要目的.通常把图像压缩编码简称为图像编码.图像编码是一种信源编码,其信源是各种类型的图像信息.图像信息的压缩有模拟和数字两种方法,但是由于数字压缩方法具有模拟压缩方法无可比拟的优越性,目前绝大多数系统都采用数字压缩方法,模拟压缩方法将逐渐被淘汰.本书只讨论数字图像压缩技术,它的最大优点是可以实现很高的压缩比.目前已达到能用一个话路传输彩色活动图像的水平.用模拟压缩技术实现这样高的压缩是不可想象的.

数字图像通信有数字通信的一系列优点,如:可以中继传输和多次复制,不会造成噪声和非线性失真的累积;便于进行加密;便于用 VLSI 芯片实现,制作方便、成本低、可靠性高;便于和计算机联网等.既然数字图像通信有这么多好处,为什么没有更早地得到推广和应用呢?这主要是因为数字图像的数据量非常巨大,若不经压缩,数字图像传输所需的高传输速率和数字图像存贮所需的巨大容量将成为推广应用数字图像通信的最大障碍.这就是为什么要进行图像压缩编码的根本原因.

现以一路电视信号为例,看看将它数字化后的数码率.按 CCIR601 标准,数字化后的分辨率为 720×576 ,每秒 25 帧, $Y:U:V$ 为 $4:2:2$.若以 8bit 表示 Y 信号,则每像素占 16bit,数码率为 165.9Mbps.以 64kbps 作为一个数字话路,若不加压缩,为传输一路电视要占用 2 592 个有效数字话路.这在实际应用中是难以接受的.若用一个容量为 1GB 的硬盘或 CD-ROM 来存贮这样的数据,则只能存贮不到 1 分钟的图像.并且,所需的高数据吞吐率是一般的硬盘和 CD-ROM 难以达到的.若不加压缩,HDTV 信号的数码率可接近 1Gbps,更加惊人.再以指纹库为例,若以 $(512 \times 512) \times 8\text{bit}$ 的灰度图像来存贮一个手指的指纹,一个 40 万人的指纹库,每人十指,则共需 1 000GB 的存贮量.由这些例子就可以看出图像压缩编码的必要性.图像编码技术的进展已使这些制约因素不再成为瓶颈,从而推动了各类图像通信系统的推广和应用.图像编码是各类图像信息传输、存贮产品的一项核心技术.

图像数据可以进行压缩有几方面的原因.首先,原始图像数据是高度相关的,存在很大的冗余度.数据冗余造成比特数浪费,消除这些冗余可以节约码字,也就是达到了数据压缩的目的.大多数图像内相邻像素之间有较强的相关性,这称为空间冗余度.序列图像前后帧之间有较强的相关性,这称为时间冗余度.多光谱遥感图像各谱间有相关性,这称为频率域冗余度.其次,若用相同码长表示不同出现概率的符号也会造成比特数的浪费,这种浪费称为符号冗余度.如果采用可变长编码技术,对出现概率高的符号用短码字表示,对出现概率低的符号用长码字,就可消除符号冗余度,从而节约码字.

允许图像编码有一定的失真也是图像可以压缩的一个重要原因.在许多应用场合,并

不要求经压缩及复原以后的图像和原图完全相同,而允许有少量失真.只要这些失真并不被人眼所察觉,在许多情况下是完全可以接受的.这就给压缩比的提高提供了十分有利的条件.图像质量允许的损失愈多,可以实现的压缩比就愈大.这种有失真的编码称为限失真编码.在多数应用中,人眼往往是图像信息的最终接收者(信宿).如果能充分利用人眼的视觉特性,就可以在保证所要求的图像主观质量的前提下实现较高的压缩比.这就是利用了视觉冗余度.其实人类的视觉系统(HVS)是有缺陷的,对某些失真不敏感,难以觉察.一个图像编码方法如果能充分利用这些特性,就可取得较好的效果,即在复原图像主观质量较好的前提下得到较高的压缩比.我们举几个例来说明:①人眼对颜色的空间分辨率低于对亮度信号的分辨率.在CCIR601标准中, $Y:U:V$ 可选用 $4:2:2$ 就是利用了这个特性,将色差信号的空间分辨率减半,仍可得到非常高的图像质量.如果将它和三个分量都是8bit的方案相比,得到 $24:16$ 即 $3:2$ 的压缩比,而图像主观质量几乎没有降低.②人眼对画面静止部分的空间分辨率高于活动部分.所以,可以对静止部分减少时间分辨率,而对快运动部分减少空间分辨率,达到数据压缩的目的.③人眼不能觉察亮度的细小变化.即存在视觉阈值,而且此阈值随着图像内容的变化而变化.在平坦区,阈值低,对失真较敏感.在边缘和纹理区,对失真不敏感,这就是视觉掩盖效应.这种特性被广泛用来提高压缩比.④人眼对屏幕中心区的失真敏感,对屏幕四周的失真不敏感.因此对四周的粗糙量化也可以节约码字.

此外,还可以利用先验知识实现图像编码.在某些特定的应用场合,编码对象的某些特性可预先知道.例如,在可视电话中,编码对象为人的头肩部像.这时,可以利用对编码对象的先验知识为编码对象建立模型.通过提取模型参数,对参数进行编码而不对图像直接进行编码,可以达到非常高的压缩比.这是模型编码的基本思想.在这里,利用了知识冗余度.

图像编码技术就是要把种种压缩的可能性变为现实.关于图像编码的现实性,我们将在第1.4节中结合应用实例展开讨论,从中可以看到各种图像编码产品已经得到广泛应用,并将以更高的速度发展.

1.2 图像压缩编码一般框图与分类

1.2.1 图像压缩编码一般框图

图像编码的过程可以概括成图1.1所示的三个步骤,原始图像经映射变换后的数据再经量化器和熵编码器成为码流输出.现对编码框图作下列说明:

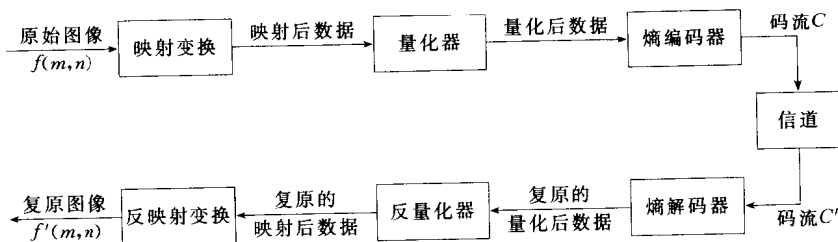


图 1.1 图像编码的一般框图

1. 映射变换

其目的是通过映射改变图像数据的特性,使之更有利于压缩编码。

例 以左邻像素灰度值作为当前像素灰度值的预测值,以预测差作为映射后数据,并对其进行量化编码。

这样做一般优于直接对原图进行量化编码。因为图像数据在相邻像素间有较大相关性,不管原图明暗度如何,相邻像素灰度差总是分布在零附近。进一步的统计表明,差值信号的分布接近拉普拉斯分布,其标准差比原始图像的标准差要小得多,因而对它进行量化编码所需的比特数就较少。实际上,求差值的过程减少了相邻像素间的相关性,从而减少了冗余度,因此可以实现压缩。

如果采用更好的预测方法,如由空间相邻四像素点预测当前像素值(帧内 DPCM),或用前一帧的像素值预测当前像素值(帧间 DPCM),可望得到更好的预测值,从而进一步提高压缩比。在实际应用中,映射变换的方法种类繁多,还可能更复杂。如:在变换编码中,先将图像分成若干个 $n \times n$ 大小的子块,然后进行映射变换。这种情况下的映射变换是对各子块进行某种正交变换,而量化和编码是对变换后所得系数进行的。事实上,映射变换是图像编码的一个核心部分,它决定了量化和编码的对象类型。所以可以据此对编码方法进行分类。在上面的例 1 中,若在映射变换部分中进行了帧内预测,则称其为帧内 DPCM 编码方法,若进行的映射是帧间预测则称其为帧间 DPCM 编码。二者都属于波形编码,因为是直接在空间域中对图像灰度值或其简单运算后的结果进行量化编码的。而变换编码则是对经映射后得到的变换域中的数据进行量化编码的。在模型编码中,量化编码的对象是模型参数。这里的映射变换实际上是一个复杂的建模和模型参数估值的过程。它把原始图像映射成数据量极小的模型参数,从而实现了极高的压缩比。从这些例子可以体会到映射变换在图像编码中的核心作用。一个好的映射变换通过与合适的量化器相配合,应能充分消除图像信源的各种冗余度。

2. 量化器

在限失真编码中要对映射后的数据进行量化。若量化是对映射后的数据逐个地进行的,则称标量量化。若量化是成组地进行的,则称矢量量化。量化总会造成某些信息丢失,形成失真,即量化失真或量化噪声。为使失真小,应量化得精细,但压缩比就高不了。这是一对矛盾。应选用恰当的量化级数和量化曲线形状来缓解这对矛盾。量化器的引入是图像编码产生失真的根源。在要求复原图像与原图完全一致的无失真编码器中必须不用量化器。但这样一来,压缩比难以提高。在多数应用中,存在少量失真并不可怕,只要把失真的程度和性质控制在允许的范围内。也就是把复原图像的主观质量控制在允许的程度内,就可以在满足应用要求的前提下提高压缩比。值得注意的是,对于同样的量化失真,由不同的映射变换与反变换会引起不同性质的复原图像的失真,人眼对某些性质的失真敏感而对另一些性质的失真不敏感。

3. 熵编码器

这一步是用来消除符号编码冗余度的。它一般不产生失真。理想的情况是使编出的码流的平均码长等于量化后数据的信息熵。常用的编码方法有许多种,例如分组码:行程码(RLC)和变长码(VLC);不分组码:算术码(Arithmetic Coding)。

行程码传输的是数据行程长度而不是数据本身。例如要传 108 个 0,用行程码时改为

传行程码标记和行程长度值“108”。

在变长码中最常用的是霍夫曼(Huffman)码。基本原则是对出现概率大的数值用短码来编,对出现概率小的数值用长码来编,从而使平均码长减小,降低码率。在实际编码器中常把 RLC 和 Huffman 码结合起来,称为霍夫曼行程码。在 H. 261, JPEG, MPEG 等国际标准中就用到它。而算术码则用于 JPEG, H. 263, JBIG 等国际标准中。

上述三个步骤之间是相互联系相互制约的。对有些编码方法,如预测编码或变换编码,映射变换后数据量并没有减少,甚至因动态范围的加大而使数据量略有增加。但它为后两步作了准备,使它们能有效发挥作用。而在模型编码中,经映射变换后得到的模型参数,其数据量已大大小于原始图像,即第一步已实现了很大的压缩。后面的量化编码则是作进一步的压缩。其情况和波形编码、变换编码有很大的不同。我们举一个例子来说明各部分之间的相互影响:如果第二步使用了不均匀量化器,它可使每层量化值的出现概率接近相等,那么后面再用变长码,这第三步的压缩效果就会大打折扣。或者反过来说,若后面用了变长码,前面的量化器可以不妨采用线性量化器,总的效果仍然很好,量化器实现起来也要方便得多。

此外,如果码流在信道中传输时产生误码,就会造成复原图像质量的下降。有时一个比特的误码会造成一大片复原图像的错误,这种现象称误码扩散现象。在上述例子中,由于编码端对与左邻像素的差值进行量化编码,则在解码端,若因误码造成了某个像素复原值有误,则从该像素起,以后复原的右面所有像素都有误,按信息论的观点,压缩得越多、冗余度越小的码流越难抵抗误码的影响。在实用的图像编码算法中必须对误码扩散提出有效制约对策。

值得注意的是,误码可能导致码流结构的变化,从而不能正确解码。一个好的熵编码器应该有自同步能力,能够在受到误码影响后经过一定的码以后仍能自动回到正确解码路径上来,而不会死机或一直错下去。

1.2.2 图像压缩编码的分类

图像编码应用广泛,有许多编码方法。对它们进行分类有助于对问题的理解和解决。基于不同的角度,有许多种分类方法。如:按待压缩图像的性质分类,按复原图像是否与原图一致分类,按所用方法原理分类等等。下面分别对它们进行说明:

1. 按复原图像是否与原图一致分类

可将编码方法分为两大类:无失真编码和限失真编码。无失真编码又称信息保持编码(lossless 编码)或可逆编码。限失真编码又称非信息保持编码(Lossy 编码)或不可逆编码。编码会造成失真,不过这些失真可以被控制在一定的限度内,不致影响使用效果。在无失真编码中不可使用量化器,因为量化总会带来不可恢复的失真。无失真编码的压缩比低,可达到的最高压缩比取决于信源熵。在限失真压缩中允许的失真愈大,可达到的压缩比愈高,即与量化的粗细有很大关系。在限失真编码中把失真控制在视觉阈值以下或控制在可容忍的程度是有效地提高压缩比的关键。

2. 按待压缩图像的不同属性分类

有不同的分类方法。如:以图像的光学特征分,有单色、彩色、多光谱图像的压缩编码;以灰度等级分,有二值图、多值图与灰度图像编码;若以动静来分,有静止图像与序列图像

编码. 在静止图像编码中, 只能进行帧内编码, 而且人眼对静止图像的失真较运动图像更为敏感, 压缩比就没有运动图像编码来得高. 但运动图像编码对实时性的要求很高, 这是它特有的难点, 对不同要求(无失真、小失真、较大失真), 不同性质的图像用适当的编码方法和编码参数(如量化级大小)进行压缩编码是达到预期目标的关键.

3. 按所用方法的原理分类

可将图像编码方法分为基于图像统计特性, 基于人眼视觉特性, 基于图像特征提取等方法. 在实际编码中, 常常要同时用到图像的统计特性和人眼视觉特性才能进行有效的编码, 难于把它们截然分开. 而有些编码方法的类别还可进一步细分, 如基于统计特性的编码方法还可分为帧内预测、帧间预测、变换编码等. 许多新的国际图像编码标准使用了混合编码方法, 即同时用到了帧内、帧间预测和 DCT 编码.

为便于理解和叙述, 可以把图像编码方法分为四大类: 波形编码、变换编码、模型编码和拟合编码. 它们都是以映射变换的性质为依据进行分类的. 前三种在上一小节已说明过了. 属于拟合编码的方法有: 方块截尾编码(BTC), 三角形逼近法和迭代函数系统法(IFS)等.

1.2.3 从 MDL 的观点看图像编码方法的分类¹⁾

最短描述长度准则(MDL)是 Rissanen 于 1978 年提出的, 它可用于寻找数据的内在相关性, 其中包括图像数据的相关性. 若用 X 表示原始图像数据, 用 M 表示模型, 此模型反映了 X 的内在规律性. 可以把图像编码过程看成用模型 M 对原始图像数据 X 的描述. 若把原始数据的直接描述(即不用模型)记作 $D(X)$, 其长度记作 $|D(X)|$. 用了模型 M 后, 对 X 的描述可记作 $D(X, M)$, 其长度为 $|D(X, M)|$. 若 $|D(X, M)| < |D(X)|$, 则我们说用模型 M 对图像 X 进行了压缩编码. 我们有:

$$D(X, M) = D(M) + D(X|M) \quad (1.2-1)$$

$$|D(M)| + |D(X|M)| = |D(X, M)| < |D(X)| \quad (1.2-2)$$

式(1.2-1)的含义是对 X 的描述(编码)可转化为对模型的描述 $D(M)$ 及利用模型对原始数据的描述 $D(X|M)$. 在使用参数模型的情况下, $D(M)$ 就是对模型参数 θ 的编码. $D(X|M)$ 就是对原始数据 X 与模型 M 的误差 $X-M$ 的编码. 我们可以用 MDL 观点将图 1.1 改画成图 1.2 的形式. 把映射变换看成建模表达. 不论实际的映射变换是什么, 是预测也好, 变换也好还是真的对景物建立模型, 在这里均把它看成用一个广义模型来表示图像. 图中画出采用参数模型的情况. 即映射后数据由模型参数 θ 和模型对原图的描述误差 $X-M$ 两部分构成. 对于模型编码, 图 1.2 的表达方式是最直观的. 此图同样也适于其它编码方法. 例如对帧内预测编码, 模型参数就是预测系数, 由于它已事先设定, 不必对它进行编码传输, 故 $|D(M)| = 0$, 而只需对描述误差 $X-M$ 进行量化编码. 在帧间预测编码时, 模型参数就是运动矢量, 若运动估计得较准确, 则 $D(X|M)$ 可以较小, 但总的码流长度中包括了运动矢量码的长度 $|D(M)|$.

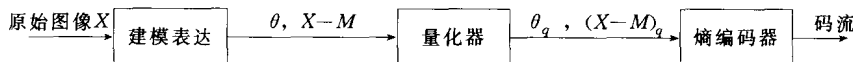


图 1.2 从 MDL 观点看图像编码的三个步骤

1) 初学者不必阅读本小节.

若有两个模型(对应了两种编码方法) M_1 和 M_2 ,有

$$|D(M_1)| + |D(X|M_1)| < |D(M_2)| + |D(X|M_2)|$$

则从 MDL 观点看 M_1 比 M_2 好,即第一种编码方法有更高的压缩比. 从 MDL 观点看最优的模型应是使描述长度为最短的,即

$$M^* = \arg \min_{M_i \in m} \{|D(M_i)| + |D(X|M_i)|\} \quad (1.2-3)$$

式中 M^* 为最优模型, $m = \{M_1, M_2, \dots\}$ 是模型论域. 对此有两点需要强调说明:

(1) 对模型本身的描述长度 $|D(M)|$ 必须包括进去,而不应只看 $|D(X|M)|$ 来定优劣. 举个极端的例子,若取 $M=X$,则这时 $|D(X|M)|=0$,但这并不好,因为 $|D(X,M)| = |D(X)| + |D(X|M)| = |D(X)| + 0 = |D(X)|$,反而没有起到压缩作用.

(2) 最优模型实际上是在模型的复杂性和模型对图像数据的描述能力之间取得最优平衡的模型. 一般来说,模型越复杂,对数据的描述能力越强,即 $|D(M)|$ 大而 $|D(X|M)|$ 小. 关键就在于找到一个既简单又有对图像有很强描述能力的模型.

由于迄今为止,人们对人眼视觉特性的认识还很肤浅,无法精确地定量描述,且主观感觉不仅与误差本身有关,还与误差所在区域灰度及其强弱变化等等有关,加之式(1.2-3)通常是一个非凸问题,其最优解是个不可计算问题,所以不可能用 MDL 准则直接导出一个与主观质量相吻合的最优模型. 但这并不影响用它来指导对各种编码方法的分类,表 1.1 总结了用 MDL 准则对几种典型的图像编码算法的分类情况,颇有启发意义.

表 1.1 用 MDL 准则对图像编码算法分类

编码方法	类别	模型类型	模型参数
帧内预测	预测	图像本身(空间邻域)	预测系数
帧间预测	预测	前帧图像	运动矢量
变换编码	变换	基图像	变换系数
方块截尾编码	拟合	分段常数	分段定义及各段灰度值
三角形逼近编码	拟合	三角形面片	顶点坐标
分形编码	拟合	图像本身	IFS 代码
模型编码	模型	景物几何模型	几何量和亮度、色度

在预测编码中,模型是原图像本身,而在变换编码中,模型是一基图像. 基图像在 DCT、小波变换等中是固定的,在 K-L 等变换中是可变的. 这两类编码方法都把图像作为一个二维点集. 而拟合编码类方法把图像作为三维空间中的一张曲面,通过对该曲面进行拟合实现图像编码. 模型编码则把图像及图像序列作为三维景物及其运动变化在二维图像平面上的投影,通过对三维景物或其二维图像平面上的投影的建模表达,达到图像编码的目的.

1.3 图像压缩编码技术的发展历史与现状

1.3.1 图像编码技术发展历史

1948 年提出电视信号数字化后,就同时开始了对图像压缩编码的研究工作,至今已

有 40 多年的历史. 开始时, 即 50 和 60 年代, 限于客观条件, 仅对帧内预测法和亚取样内插复原法进行研究, 对视觉特性也做了一些极为有限但可贵的工作. 1966 年 J. B. O' Neal 对比分析了 DPCM 和 PCM 并提出了用于电视的实验数据. 1969 年进行了线性预测编码的实际试验. 1969 年举行首届图像编码会议 (Picture Coding Symposium). 70 年代开始进行了帧间预测编码的研究. 80 年代初开始对作运动补偿 (MC) 所用的运动估值 (ME) 进行研究. 变换编码是 1968 年 H. C. Andrews 等人提出的, 采用的是二维离散傅里叶变换. 此后相继出现了用其它变换方法的变换编码, 其中包括二维 DCT. 对模型编码的研究始于 80 年代初.

于 1988 年形成草案、1990 年通过的 ITU-T H. 261 建议, 是图像编码技术走向实用化的重要一步. 它是图像编码 40 年研究成果的结晶. 90 年代初相继提出的 MPEG-1, MPEG-2, H. 263 等都是在 H. 261 基础上的发展和改进的. 这些国际建议普遍采用的混合编码技术是当今最实用的高效编码方法, 得到了广泛的推广应用, 业已成为当今图像编码方法的主流.

1.3.2 图像编码技术的现状

进入 90 年代以后, ITU-T 和 ISO 制定了一系列图像编码国际建议, 如:

- (1) 1990 年为会议电视和可视电话制定的 H. 261 建议.
- (2) 1991 年为静止图像编码制定的 JPEG 建议 (ISO 10918) (1991 年草案; 1994 年正式通过).
- (3) 1991 年为二值图像编码制定的 JBIG 标准 (ISO CD 11544).
- (4) 1991 年为电视图像数字存储 (数码率高达 1.5Mbps) 而制定的 MPEG-1 (ISO CD 11172) (1991 年草案, 1993 年正式通过).
- (5) 1993 年为活动图像及其伴音压缩而制定的通用编码国际建议 MPEG-2 (其中视频部分, 即 ITU-T H. 262, ISO/IEC 13818-2, 1993 年草案, 1994 年正式通过).
- (6) 美国“大联盟” (Grand Alliance, 简写 GA) 公布数字 HDTV 系统的说明书草稿 (1993 年草稿, 1994 年正式文本). 美国“先进电视系统委员会” (ATSC) 拟定“数字电视标准” (1996 年).
- (7) 1995 年 ITU-T 为甚低码率视频编码而制定的 H. 263 建议.

这些建议的制定极大地推动了图像编码技术的实用化和产业化. 会议电视等各类使用图像压缩技术的产品纷纷推出, 数字激光视盘 (VCD) 等产品以百万台的数量级走向市场, 进入家庭, 从而迎来了数字图像通信的黄金时代.

另一方面, 图像编码技术产业化进程的加快也推动了图像编码技术以更快的速度发展. 目前的研究工作分为两个方向:

1. 更好地实现现有的图像编码国际建议

(1) 研制出集成度更高、性能更好的图像编码专用芯片 (ASIC), 使编码系统成本更低, 可靠性更高. 几个国际建议的单片 (或两片) 系统级解码 ASIC 均已推出, 其中包括对运算速度要求很高的 MPEG-2 解码芯片. 不久还将推出要求更高的 HDTV 的单片视频解码芯片. 符合 MPEG-2 标准的单片或多片视频编码 ASIC 也已推出.

(2) 解决好现有的图像编码系统开发中的技术问题. 例如: 提高图像质量, 提高抗误

码能力,在 ATM 网等变速信道上的应用等.如果拿现在生产的符合 MPEG-1 的 VCD 图像质量和几年前 MPEG-1 刚制定时的 VCD 图像质量相比,就可以看到虽然用的是同一个国际建议和同样的数码率(1.5Mbps),但图像质量大大提高了.这就是近几年来对 MPEG-1 编码器具体实现算法作深入研究的成果.国际建议的开放性结构为这种深入的改进提供了前提,它允许人们在不影响兼容性的前提下发挥自己的创造性,对建议中的开放部分进行改进.这些开放性部分包括运动估值和运动补偿方法,自适应量化系数和缓存器控制策略等.在国际建议规定的约束下,对这些部分的算法作更合理的细化有助于提高编解码器的性能.如:复原图像质量、抗误码能力等.在本书下面的有关章节中我们还会对这些内容做详细的讨论.

2. 对图像编码理论和其它图像编码方法的研究

正如文献[8]所指出的那样,基于线性方法(预测法和分块 DCT 相结合)的 H. 261/MPEG 核心算法既不能有效地反映图像的固有统计特性,也不能有效地反映人们的视觉特性,这两方面问题还有待理论上的突破.

目前已提出和正在进行研究的图像编码方法择要列举如下:

(1) 多分辨率编码.最早提出的是金字塔编码,后来是子带编码(Subband Coding),最近是用小波变换进行图像编码.

(2) 基于表面描述的编码方法(三角形逼近法).

(3) 模型编码.它可分为物体模型未知的物体基编码和物体模型已知的语义基编码.

(4) 利用人工神经网的压缩编码.

(5) 利用分形几何的图像编码(IFS 编码).

(6) 利用数学形态学的图像编码等等.

本书第四章将对这些方法的原理作简要介绍,并对其中某些方法作进一步的介绍.已提出的许多图像编码方法中有的仍脱离不了线性方法的范畴.有不少方法,原理新、想法合理,也有不少成功的报道,但它们的成熟程度如何目前还很难说.举个例子,对 MPEG-4 的许多候选方案进行的测试评比虽未见公开报道,但显然并没有某个新方法是目前最流行的混合编码方法相比有显著的优越性.我们期待着图像编码技术的新突破.德国 Musmann 教授 1995 年秋在来华讲学中,对图像压缩技术的现状作了分析,他在列举了不少实例后指出,现已成熟的压缩算法所达到的有效压缩比(不把亚抽样的数据压缩能力计算在内)约为 26 倍.如果这个数字还能再提高三至四倍,则可以把电视信号经亚抽样及压缩挤入电话信道,其意义将十分巨大.然而这三至四倍的压缩比的提高(当然是在复原图像质量满足要求的前提下)难以用现有技术的框架来实现,需要新的技术突破.

1.4 图像压缩编码技术的应用

图像编码一系列国际建议的提出标志着图像编码技术已经成熟,开始由学术研究走向产业化,前景十分诱人.早在 1991 年就有人预言,图像编码技术的突破具有十分巨大的意义.其意义之大已大到可以促使现有信息产业的结构发生巨变的程度,它使通信、广播、计算机产业的界限变得模糊了.近年来的实践也证实了这个预言.目前,国外已出现了有线电视公司和通信、计算机公司之间的相互合作(甚至合并),如:美国最大的有线电视

公司(Telecommunication)和软件公司 Microsoft 合作开发交互式(interactive)电视等.文献[3]在解释制定新的国际建议 MPEG-4 的目标和功能要求时,一开始就指出原来的广播电视,计算机,电话三者分立的局面正在改观,而 MPEG-4 面临的将是视听信息进入计算机和电话这两个“世界”的新环境.现在的情况已不是简单地满足某些用户的图像压缩的要求,而是正在以图像编码为核心技术之一大规模地积极开拓新的产品和应用领域,并已开始影响到人们的生活方式.

1.4.1 数字电视

数字电视具有图像质量高,可以更合理地充分利用各种类型的频谱资源,可以实现多种业务的动态组合和统计复用,易于加密,具有可扩展性、可分组性和互操作性,可以灵活组成交互式电视系统等优点.这些优点是模拟电视所无法比拟的.可以断言,下一代电视必将是数字电视.但是,数字电视受到数据量庞大的约束,所以近年来图像编码研究主要集中在视频压缩上,以改进数字电视的性能,从而促使各等级(level)数字电视的成功实现,扩大其应用领域.这从近年来制订的图像编码国际建议中视频编码和静止图像编码国际建议的数量比例中即可可见一斑.以分辨率为基础,数字电视可分为四个等级:可视电话、会议电视、数字标准清晰度电视和高清晰度电视.表 1.2 总结了它们的有关情况(摘自文献[18]),其中 VCD 已经批量投产.在国外,H. 263 和常规电视分辨率的 MPEG-2 产品也开始推出,而 HDTV 接收机的批量生产则是眼前就可实现的事.现对各类数字电视产品择要分别加以介绍.

(1) 数字激光视盘(VCD).它是各类数字电视产品中投入市场最多的产品之一,在国内已经普及,目前正以百万台的批量进入家庭.它以 MPEG-1 为基础,分辨率 ≥ 280 线(水平) $\times 260$ 线(垂直).有关产品包括节目制作设备,盘片生产和 VCD 放像机.目前可以用单片实现 MPEG-1 系统级码流的解码.随着对 MPEG-1 研究的深入,复原图像质量提高很快,已略优于 VHS 录像机而成本却低于它.

(2) 数字低清晰度电视(LDTV).这是一种普及型数字电视,采用的仍是 MPEG-1 技术.与数字标准清晰度电视相比,它的成本低得多,因而更适于我国国情.它可用于有线电视网中,以改善电视图像质量.其信道包括同轴电缆、光缆、卫星等.它还有便于加扰的优点,适于做付费电视.由于它的分辨率正好与我国现有电视机相匹配,因而可不必更新现有电视机而达到图像质量的提高,在我国有良好的推广前景.

(3) 会议电视和可视电话.会议电视系统的图像编码采用 ITU-T H. 261 建议,数码率在 64kbps 到 2Mbps 之间.目前我国已建立会议电视骨干网(到省会城市),现正向下延伸到地、县,以及铁道、交通、石油等部门,其需求量愈来愈大.目前会议电视较多使用固定速率的数字通信网,例如用 E1(一次群),2B+D 等信道.随着计算机网和 ATM 技术的推广,联网会议电视系统在今后将得到较大发展.目前利用 HDSL,ADSL 技术可在双绞线上传输会议电视码流,从而大大扩大了它的应用范围.H. 261 除了用于会议电视,可视电话外,还可用于各种监控系统中,减轻对传输信道的压力.采用新制订的 H. 263 建议可使数码率低于 64kbps,从而可以用普通电话网(如:28.8kbps)或移动网传输活动图像.到那时可视电话的普及率将极大提高.