

新一代

视频压缩编码标准

H.264/AVC

毕厚杰 主编

图书在版编目 (CIP) 数据

新一代视频压缩编码标准——H.264/AVC/毕厚杰主编. —北京: 人民邮电出版社, 2005.5

ISBN 7-115-13064-7

I. 新... II. 毕... III. 视频信号—数据压缩 IV. TN919.81

中国版本图书馆 CIP 数据核字 (2005) 第 014784 号

内 容 提 要

H.264 视频压缩标准被称为新一代视频压缩编码标准。本书在介绍数字视频和视频编码的基本原理基础上, 论述了 H.264 的特点、编码器原理、解码器原理以及编解码器的实现方案。为了便于读者更好地理解 H.264 的编解码技术, 书中介绍了 H.264 码流的句法和语义, 并对 H.264 视频编码传输的服务质量进行了专门论述。

本书取材新颖、内容全面, 重点论述了 H.264, 还介绍了已有的若干视频编码国际标准 (特别是 MPEG-4)。本书可作为高等院校通信、广播电视专业本科生教材或教学参考书, 也可供该领域的硕士生、博士生深入研究用, 还可作为视频技术、视频服务领域广大技术人员的必备参考书。

新一代视频压缩编码标准——H.264/AVC

- ◆ 主 编 毕厚杰
责任编辑 陈万寿
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
读者热线 010-67129258
北京鸿佳印刷厂印刷
新华书店总店北京发行所经销
- ◆ 开本: 787×1092 1/16
印张: 18.5
字数: 453 千字
印数: 1—3 600 册
- 2005 年 5 月第 1 版
2005 年 5 月北京第 1 次印刷

ISBN 7-115-13064-7/TN · 2415

定价: 36.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

前 言

数字视频技术在通信和广播领域获得了日益广泛的应用，特别是 20 世纪 90 年代以来，随着 Internet 和移动通信的迅猛发展，视频信息和多媒体信息在 Internet 和移动网络中的处理和传输技术成为了当前我国信息化中的热点。

众所周知，视频信息具有一系列优点，如直观性、确切性、高效性、广泛性等等。但是视频信息的信息量太大，要使视频得到有效的应用，必须首先解决视频压缩编码问题，其次解决压缩后视频质量保证的问题。这两者是相互矛盾的，是矛盾的两个方面。我们的任务是既要有较大的压缩比，又要保证一定的视频质量。

为此，人们付出了巨大的辛勤的劳动，现在已结出丰硕的成果。从 1984 年 CCITT 公布第一个视频编码国际标准以来，至今已有 20 年了。ITU-T 等国际标准化组织陆续颁布了接近十个视频编码国际标准，大大推动了视频通信和数字电视广播的发展，这也是有目共睹的事实。但是严格地讲，视频通信和数字电视广播这两大领域至今的发展仍不能令人满意，总起来讲，应用的范围不广，主要是视频压缩与质量之间的矛盾不能很好解决。例如，可视电话一直被认为是一种理想的通信设备，可近 30 年来至今未能普及，就是因为性能价格比不高。

2003 年 3 月，ITU-T/ISO 正式公布了 H.264 视频压缩标准，由于其具有比以往标准更出色的性能，被人们称为新一代视频编码标准。具体讲，与 H.263 或 MPEG-4 相比，在同样质量下，其数码率能降低一半左右；或者说在同样码率下，其信噪比明显提高。这样一来，H.264 标准在国际上受到了广泛地重视和欢迎。在这样的背景下，我们编写了这一本书，对其进行详细介绍。

本书的特点是取材新颖、内容全面。它不仅重点论述了 H.264，而且首先介绍了数字视频和视频编码的基础知识，介绍了已有的若干视频编码国际标准（特别是 MPEG-4），以便为进一步学习 H.264 打下良好的基础。

全书共 9 章，在 H.264 部分（第 6~9 章）详尽地论述了 H.264 的特点、编码器原理、解码器原理、编解码器的实现方案。为了更好地理解 H.264 编解码器的原理及其实现方案，第 7 章详细介绍了 H.264 码流的句法和语义。最后对 H.264 视频编码传输的 QoS 进行了专门地论述。

本书可作为通信、广播电视专业高校本科生教材，并可供该领域的硕士生、博士生深入研究用，也可供广大的视频技术、视频服务领域的技术人员参考用。

本书由毕厚杰教授主编，撰写了第 1~4 章及第 6 章前 3 节。左雯撰写了第 5 章。马国强、徐苏珊撰写了第 7 章。焦良葆、王健撰写了第 9 章。其余部分由方晖、焦良葆、王健、马国强、左雯、李涛、徐苏珊、鹿宝生等人共同编写完成。全书由毕厚杰负责审稿，左雯负责校对。

本书编写过程中，得到了南京大学章德教授、陈启美教授的多方支持，特此深表谢意。本书的编写也引用了不少参考文献，特向这些作者表示诚挚的谢意。

由于时间仓促及水平有限，书中难免出现不当之处，恳请广大读者批评指正，以便再版时进一步修正。

编者

目 录

第 1 章 绪论	1
1.1 信息化与视频通信	1
1.1.1 什么是信息	1
1.1.2 什么是信息化	1
1.1.3 我国的信息化和视频通信	2
1.2 视频信息和信号的特点	2
1.2.1 直观性	2
1.2.2 确定性	2
1.2.3 高效性	2
1.2.4 广泛性	3
1.2.5 视频信号的高带宽性	3
1.3 视频压缩编码的要求和可能性	3
1.3.1 视频压缩编码的目标	3
1.3.2 视频压缩的可能性	3
1.4 视频压缩编码技术综述	4
1.4.1 基本结构	4
1.4.2 基于波形的编码	5
1.4.3 基于内容的编码	5
1.4.4 三维(立体)视频编码	6
参考文献	6
第 2 章 数字视频	7
2.1 数字电视的基本概念	7
2.1.1 数字电视的优越性	7
2.1.2 数字电视的 PCM 原理	8
2.2 数字电视信号	11
2.2.1 电视信号的时间和空间取样	11
2.2.2 彩色空间	12
2.2.3 彩色电视取样格式	13
2.2.4 数字电视信号的编码参数	13
2.3 视频信号的预处理	14
2.3.1 色彩插值(Color Interpolation)	14
2.3.2 色彩校正(Color Correction)	15

2.3.3 伽马校正 (Gamma Correction)	16
2.3.4 图像增强 (Image Enhancement)	17
2.3.5 白平衡 (White Balance)	21
2.4 视频质量	22
2.4.1 主观质量的评定	22
2.4.2 客观质量的测量	23
参考文献	24
第3章 视频压缩编码的基本原理	25
3.1 预测编码	25
3.1.1 预测编码的基本概念	25
3.1.2 帧内预测编码	26
3.1.3 帧间预测编码	30
3.1.4 运动估计	33
3.2 变换编码	46
3.2.1 变换编码的基本概念	46
3.2.2 K-L 变换	46
3.2.3 离散余弦变换 DCT	47
3.2.4 锯齿形扫描和游程编码	49
3.3 变换编码与预测编码的比较	49
3.4 熵编码	49
3.4.1 变长编码	50
3.4.2 算术编码	50
参考文献	52
第4章 视频编码标准简介	53
4.1 视频编码发展简史	53
4.2 H.261 标准	53
4.2.1 图像格式	53
4.2.2 H.261 视频编解码器	55
4.3 H.263 标准	58
4.3.1 H.263 标准图像格式	58
4.3.2 H.263 视频信源编码算法	59
4.4 MPEG-1 标准	59
4.4.1 功能	59
4.4.2 图像类型和编码结构	60
4.5 MPEG-2	61
4.5.1 MPEG-2 编码复用系统	61
4.5.2 MPEG-2 的档次和级别	61
4.5.3 MPEG-2 视频编解码器	62

4.6 JPEG 标准	64
参考文献	65
第 5 章 MPEG-4 压缩编码标准	66
5.1 MPEG-4 标准概述	66
5.1.1 MPEG-4 标准的特性	66
5.1.2 AVO 及数据结构	66
5.2 MPEG-4 标准构成	67
5.2.1 系统	68
5.2.2 音频	69
5.2.3 视频	70
5.2.4 网格动画	73
5.2.5 其余	73
5.3 MPEG-4 编码技术	74
5.3.1 形状编码	74
5.3.2 可扩展性编码	76
5.3.3 sprite 编码	79
5.3.4 视频系统合成	80
5.4 MPEG-4 的档次和级	81
参考文献	83
第 6 章 H.264/AVC 编码器原理	84
6.1 H.264/AVC 的应用	84
6.2 H.264/AVC 编解码器	85
6.2.1 H.264 编解码器的特点	85
6.2.2 H.264 编码器	86
6.2.3 H.264 解码器	86
6.3 H.264/AVC 的结构	87
6.3.1 名词解释	87
6.3.2 H.264 的档次和级	87
6.3.3 编码数据格式	87
6.3.4 参考图像	89
6.3.5 片和片组	89
6.4 帧内预测	92
6.4.1 4×4 亮度预测模式	92
6.4.2 16×16 亮度预测模式	94
6.4.3 8×8 色度块预测模式	95
6.4.4 帧内预测模式信号化	96
6.5 帧间预测	96
6.5.1 树状结构运动补偿	97

6.5.2	运动矢量	97
6.5.3	MV 预测	99
6.5.4	B 片预测	100
6.5.5	加权预测	103
6.6	H.264 的 SP/SI 帧技术	103
6.6.1	SP/SI 帧的应用	104
6.6.2	SP/SI 帧的基本原理	106
6.6.3	实验结果和性能分析	110
6.7	整数变换与量化	111
6.7.1	整数变换	112
6.7.2	量化	115
6.7.3	DCT 直流系数的变换量化	117
6.8	CAVLC(基于上下文自适应的可变长编码)	118
6.8.1	熵编码的基本原理	118
6.8.2	CAVLC 的基本原理	119
6.8.3	CAVLC 的上下文模型	119
6.8.4	CAVLC 的编码过程	119
6.8.5	CAVLC 的解码过程	121
6.8.6	CAVLC 编解码过程实例	123
6.8.7	CAVLC 与 UVLC 比较	124
6.9	CABAC(基于上下文的自适应二进制算术熵编码)	126
6.9.1	自适应算术编码	126
6.9.2	上下文模型	129
6.9.3	对输入流预编码	131
6.9.4	初始化	131
6.9.5	结论	132
6.10	码率控制	132
6.10.1	基于 Lagrangian 优化算法的 H.264 编码控制模型	132
6.10.2	实验结果和性能分析	135
6.11	去方块滤波	138
6.11.1	去方块滤波基本概念	139
6.11.2	边界分析	139
6.11.3	滤波过程	142
6.12	其余特征	144
6.12.1	参考图像的管理	144
6.12.2	重排序	145
6.12.3	隔行视频	145
6.12.4	数据分割片	146
6.12.5	H.264 传输	146
	参考文献	147

第 7 章 H.264 的句法和语义	149
7.1 句法	149
7.1.1 句法元素的分层结构	149
7.1.2 句法的表示方法	152
7.2 句法表	154
7.3 语义	168
7.3.1 NAL 层语义	168
7.3.2 序列参数集的语义	170
7.3.3 图像参数集的语义	173
7.3.4 片头的语义	174
7.3.5 参考图像序列重排序的语义	178
7.3.6 加权预测的语义	180
7.3.7 参考图像序列标记 (Marking) 操作的语义	180
7.3.8 片数据的语义	181
7.3.9 宏块层的语义	182
7.3.10 宏块预测的语义	186
7.3.11 子宏块预测的语义	187
7.3.12 用 CAVLC 方式编码的残差数据的语义	188
7.3.13 用 CABAC 方式编码的残差数据的语义	188
7.4 总结	188
参考文献	188
第 8 章 H.264/AVC 解码器的原理和实现	190
8.1 解码器原理	190
8.2 NAL 单元	191
8.2.1 NAL 单元结构	191
8.2.2 NAL 单元解码过程	192
8.3 图像序列号 (Picture Order Count) 的计算	192
8.3.1 图像序列号 (POC)	192
8.3.2 POC 类型为 0 的 POC 计算	194
8.3.3 POC 类型为 1 的 POC 计算	194
8.3.4 POC 类型为 2 的 POC 计算	196
8.4 宏块片组映射图的产生	197
8.5 片数据分割的解码	199
8.6 参考图像列表的初始化	200
8.6.1 图像序号的计算	200
8.6.2 参考图像列表的初始化	201
8.6.3 参考帧列表的重排序	204
8.7 解码的参考图像的标记过程	206

8.7.1	frame_num 不连续的解码过程	207
8.7.2	参考图像滑窗标记过程	208
8.7.3	参考图像的自适应内存控制标记过程	208
8.8	帧内预测	210
8.8.1	4×4 亮度块预测方式的提取	210
8.8.2	4×4 亮度块的帧内预测编码方式	211
8.8.3	16×16 亮度块的帧内预测方式	214
8.8.4	8×8 色度块的帧内预测方式	216
8.9	帧间预测解码处理	218
8.9.1	MV 分量及参考索引获取	219
8.9.2	帧间预测像素解码处理	221
8.10	变换系数解码	227
8.10.1	变换系数逆扫描过程	227
8.10.2	DCT 变换系数中直流系数的逆变换量化	228
8.10.3	残差变换系数的反量化	230
8.10.4	残差变换系数的逆 DCT 变换	230
8.10.5	去方块滤波前的图像恢复与重建	230
8.11	SP 片中的 P 宏块和 SI 片中的 SI 宏块的解码过程	231
8.11.1	主 SP 片中 P 宏块的解码过程	231
8.11.2	辅 SP/SI 片的解码过程	232
	参考文献	233

第 9 章 H.264 视频编码传输的 QoS 235

9.1	互联网视频传输 QoS	235
9.1.1	错误恢复在视频通信中的重要性和实现途径	235
9.1.2	基于块的混合视频编码框架	236
9.1.3	视频通信中提高 QoS 的抗误码和错误恢复技术	237
9.2	无线网视频传输 QoS	243
9.2.1	无线视频通信系统	243
9.2.2	无线信道编码和错误控制	244
9.2.3	无线视频通信的应用	248
9.2.4	H.264 无线通信中传输结构	249
9.2.5	无线视频传输的健壮性研究	250
9.3	H.264 视频编、解码的错误恢复	251
9.3.1	H.264 的视频编码层的错误恢复措施	251
9.3.2	H.264 的网络提取层的错误恢复	253
	参考文献	256

术语及缩略语 259

附录 1 CAVLC 相关码表 263

附录 2 CABAC 相关码表 267

附录 3 H.264 的档次和级 277

第1章 绪 论

1.1 信息化与视频通信

本书在讨论视频编码之前,先简要介绍一些与信息化有关的问题,即研讨当前信息社会背景,然后讨论信息化与视频通信的关系。

1.1.1 什么是信息

众所周知,人类社会的三大支柱是物质、能量和信息。具体而言,农业现代化的支柱是物质,工业现代化的支柱是能量,而信息化的支柱是信息。

广义地讲,信息就是客观世界的描述和分析,它无所不在,无时不在,具有广泛性和通用性,这是信息的一个性质。信息没有重量,没有长度,具有抽象性,但它确实存在,这是信息的另一个特性。

信息的第三个特性是无限性。比如,关于物质的信息,物质具有无限的不可分性,物质由分子组成,分子由原子组成,下面还有中子、质子、电子、中微子等等;关于通信网络的信息,为了增加通信容量,最初一对电话线只能通一对电话,后来利用 N-ISDN 技术,可在一对电话线上同时通两对电话,以后又发展 B-ISDN、ATM、IP、MPLS,直到今天,通信网络技术仍在不断地发展,应该说,它具有无限性。

总之,信息具有通用性、抽象性、无限性三个特征,其中无限性最重要。

从以上描述可知,信息是很有用的,它是客观世界中最本质的客观规律描述和分析,是人类社会可用的重要资源。信息资源如果能被充分开发和利用,人类社会的物质和精神文明水准将大大提高。

地球上的物质资源是有限的。石油、煤等在一定期限内总会开采完,但信息资源是无限的,对客观规律性的探讨是无尽的。而信息是要通过不断研究、不断分析,通过科学研究、反复实践才能掌握的。

1.1.2 什么是信息化

信息化是这样-一个过程:“在现代信息技术广为普及的基础上,通过开发和利用信息资源,各种社会和经济活动的功能和效益得以显著提高,人类社会的物质和精神文明达到了一个新的水平。”

可见,信息化的目的是提高人类的社会和经济效益,而实现信息化的关键在于开发和利用各种信息资源。

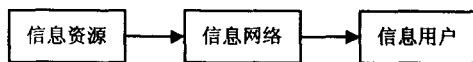


图 1.1 信息系统

信息化的实现依赖于如图 1.1 所示的信息系统来构建。

信息化的实现首先要开发和利用各种信息资源,其次要有一个信息网络(如宽带通信网),

通过该网络将大量的信息传送到信息用户。

1.1.3 我国的信息化和视频通信

我国已成为规模上世界第一的通信大国。固定电话的普及率已达到 32%，移动电话普及率已达 39%。宽带网络正在不断建设中，宽带用户普及率达 17%。8 纵 8 横的光纤构成了我国骨干传输网。以 MSTP（多业务传输平台）技术为主的城域网正在大力建设，以 ADSL 为主的宽带接入网也在迅速发展。应该指出，我国信息化的瓶颈在于信息资源的开发和利用。

经常有这样的疑问：有了宽带网络，传送什么内容呢？实际上，世界上的信息资源是十分丰富的，问题在于人们没有努力开发或者充分利用它。

信息资源中，视频信息的开发、利用更具有重要的理论意义和应用价值。

20 世纪 90 年代初以来，我国的会议电视技术获得了巨大发展，短短几年，从最初的一个中央到各省的会议电视骨干网，逐步发展为铁道、电力、石油、公安各系统、各省以至各地县约数千个的会议电视网。

20 世纪 90 年代 Internet 的迅猛发展，导致 IP 技术应用普及到各个方面，网上教育、远程医疗、电子商务、电子政务、电子游戏、网上证券等如雨后春笋。一句话，IP 视频通信已十分引人注目。

通信业务已从电话、传真、电报发展为可视电话、会议电视、视频点播等，即由音频为主发展到以视频为主的多媒体通信。

为了更深入地理解视频通信，有必要对它的特点进行进一步探讨。

1.2 视频信息和信号的特点

1.2.1 直观性

利用人的视觉获取的信息称为视频信息，它具有直观性的特点。话音信息是利用人的听觉获取的。两者相比，视频信息给人印象深刻、具体，话音信息则相对较浅。从交流信息的客观效果讲，视频信息的效果更好。

1.2.2 确定性

视频信息直观具体，不易与其他内容相混淆，能保证信息的准确性。而话音则会由于地方口音的不同产生歧义，导致不必要的损失。

1.2.3 高效性

由于利用视觉，人们可以并行地观察一幅图像的各个像素，因而获取视频信息的效率比音频信息高得多。例如，人们通过一幅电机构建的图，能很快搞清楚定子、转子及其相关位置等信息，从而很快弄清电机的结构及其原理；如果人们没有这样的图，只是一味地讲课，听者从这样的音频信息中反复理解电机结构，也可能仍搞不清楚，其接受的效率要低得多。

1.2.4 广泛性

据统计,人们每天通过视觉获取的信息约占外界信息总量的 70%左右,即人们每天获得的信息大部分是视觉信息。

1.2.5 视频信号的高带宽性

视频信息的表示形式是视频信号,通常为视频的电信号,这种信号通过网络传送至终端用户,并在屏幕上显示。

视频信号所包含的信息量大,其内容可以是活动的,也可以是静止的;可以是彩色的,也可以是黑白的;有时变化多、细节多,有时变化十分平坦。一般而言,视频信号信息量大,传输网络所需要的带宽相对较宽。例如,一路可视电话或会议电视信号,由于其活动内容较少,所需带宽较窄,但要达到良好质量,不压缩约需若干 Mbit/s,压缩后需要 384kbit/s;又如,一路高清晰度电视信号(HDTV),由于其信息量相当巨大,不压缩需 1Gbit/s,利用 MPEG-2 压缩后,尚需 20Mbit/s。可见,视频信息虽然具有直观性、确定性、高效性等优越性能,但要传送包含视频信息的信号却需要较高的网络带宽。这就是为获得视频信息所需付出的代价。

1.3 视频压缩编码的要求和可能性

1.3.1 视频压缩编码的目标

如上所述,视频信号由于信息量大,传输网络带宽要求高,就像一辆庞大的货车只有在宽阔的马路上才能行驶一样。于是出现一个问题:能否将视频信号在传送前先进行压缩编码,即进行视频源压缩编码,然后在网络上进行传送,以便节省传送的带宽和存储空间。这里有两个要求:

(1) 必须压缩在一定的带宽内,即视频编码器应具有足够压缩比的功能;

(2) 视频信号压缩之后,经恢复应保持一定的视频质量。这个视频质量有两个标准:一个为主观质量,由人从视觉上进行评定;一个为客观质量,通常用信噪比(S/N)表示。

如果不考虑质量,一味地压缩,虽然压缩比很高,但压缩后信息严重失真,显然达不到要求;反之,如只关心质量,压缩比太小,也不符合要求。

当然,在以上两个要求下,视频编码器的实现方式应力求简单、易实现、成本低、可靠性高,这也是其基本的要求。

1.3.2 视频压缩的可能性

1.3.2.1 预测编码

前面讨论了视频信息的优越性,视频信号压缩的必要性,也提出了视频压缩的目标(要求),但实现这些目标的可能性如何?

众所周知,一幅图像由许多个称为像素的点组成,如图 1.2 中的“○”表示一个像素。大量的统计表明,同一幅图像中像素之间具有较强的相关性,两个像素之间的距离越短,则其相关性越强,通俗地讲,即两个像素的值越接近。换言之,两个相邻像素的值发生突变的概

率极小，“相等、相似或缓变”的概率则极大。

于是，人们可利用这种像素间的相关性进行压缩编码。例如当前像素 X （设为立即传送的像素）可用前一个像素 a 或 b 、 c ，或三者的线性加权来预测。这些 a 、 b 、 c 被称为参考像素。在实际传送时，把实际像素 X （当前值）和参考像素（预测值）相减，例如只传送 $X-a$ ，到了接收端再把 $(X-a)+a$ 恢复为 X 。由于 a 是已传送的（在接收端被存储），于是得到当前值 X 。由于 X 与 a 相似， $(X-a)$ 值很小，视频信号被压缩。这种压缩方式称为帧内预测编码。

不仅如此，还可利用如图 1.3 所示的帧间相关性进行压缩编码。由于邻近帧之间的相关性一般比帧内像素间的相关性更强，可得到的压缩比也更大。

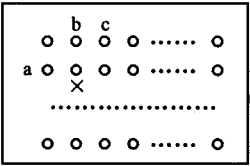


图 1.2 像素间相关性解释

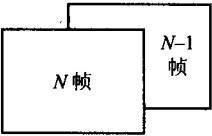


图 1.3 帧间相关性解释

由此可见，利用像素之间（帧内）的相关性和帧间的相关性，找到相应的参考像素或参考帧作为预测值，即可以实现视频压缩编码。

1.3.2.2 变换编码

大量统计表明，视频信号中包含着在能量上占主要部分的直流和低频成分（即图像的平坦部分），还有少量的高频成分，即图像的细节。因此，可以用另一种方法进行视频编码，将图像经过某种数学变换后，得到变换域中的图像（如图 1.4 所示），其中 u 、 v 分别是空间频率坐标。在图 1.4 中，用“o”表示的低频和直流成分占图像能量中的大部分，而高频成分（用“x”表示，要用码传送）则是少量的，其余均为零值，用“O”表示，于是可用较少的码表示直流低频以及高频，而“O”则不必用码传递，结果完成了压缩编码。

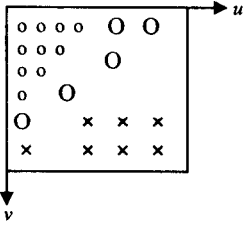


图 1.4 变换域图像

1.4 视频压缩编码技术综述

1.4.1 基本结构

视频编码系统的基本结构如图 1.5 所示。

由图 1.5 可见，视频编码方法与所采用的信源模型有关。如果采用“一幅图像由许多像素构成”的信源模型，这种信源模型的参数就是每个像素的亮度和色度的幅度值。对这些参数进行压缩编码的技术称为基于波形的编码。如果采用一个分量有几个物体构成的信源模型，这种信源模型的参数就是各个物体的形状、纹理和运动。对这些参数进行压缩编码的技术被称为基于内容的编码。

由此可见，根据采用的信源模型，视频编码可以分为两大类：基于波形的编码和基于

内容的编码。它们利用不同的压缩编码方法，得到相应的量化前的参数；再对这些参数进行量化，用二进制码表示其量化值；最后，进行无损熵编码进一步压缩码率。解码则为编码的逆过程。

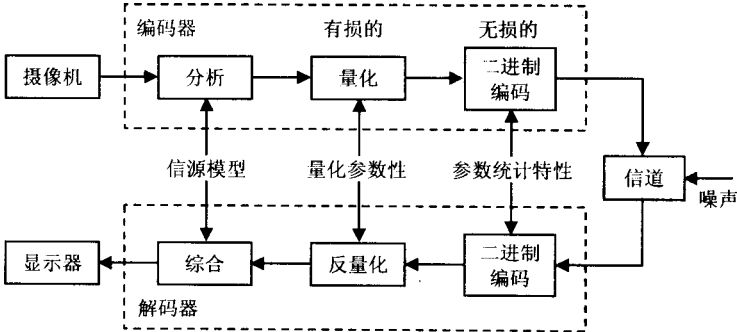


图 1.5 视频编码系统

1.4.2 基于波形的编码

如上所述，利用像素间的空间相关性和帧间的时间相关性，采用预测编码和变换编码技术可大大减少视频信号的相关性，从而显著降低视频序列的码率，实现压缩编码的目标。

基于波形的编码采用了把预测编码和变换编码组合起来的基于块的混合编码方法。

为了减少编码的复杂性，使视频编码操作易于执行，采用混合编码方法时，首先把一幅图像分成固定大小的块，例如块 8×8（即每块 8 行，每行 8 个像素）、块 16×16（每块 16 行，每行 16 个像素）等等，然后对块进行压缩编码处理。

自 1989 年 ITU-T 发布第一个数字视频编码标准——H.261 以来，已陆续发布了 H.263 等视频编码标准及 H.320、H.323 等多媒体终端标准。ISO 下属的运动图像专家组（MPEG）定义了 MPEG-1、MPEG-2、MPEG-4 等娱乐和数字电视压缩编码的国际标准。

2003 年 3 月份，ITU-T 颁布了 H.264 视频编码标准。它不仅使视频压缩比较以往标准有明显提高，而且具有良好的网络亲和性，特别是对像 IP 互联网、无线移动网等易误码、易阻塞、QoS 不易保证的网络，其视频传输性能有明显的改善。本书的主要内容即 H.264 视频编码。

所有这些视频编码都采用了基于块的混合编码法，都属于基于波形的编码。

1.4.3 基于内容的编码

如上所述，基于块的编码易于操作，但由于人为地把一幅图像划分成许多固定大小的块，当包含边界的块属于不同物体时，它们分别具有不同的运动方式，便不能用同一个运动矢量表示该边界块的运动状态。如果强制划分成固定大小的块，这种边界块必然会产生高的预测误差和失真，严重影响了压缩编码信号的质量。

于是产生了基于内容的编码技术。这时，先把视频帧分成对应于不同物体的区域，然后分别对其进行编码。具体说来，即对不同物体的形状、运动和纹理进行编码。在最简单情况下，利用二维轮廓描述物体的形状；利用运动矢量描述其运动状态；而纹理则用颜色的波形进行描述。

当视频序列中的物体种类已知时，可采用基于知识或基于模型的编码。例如，对人的脸部，已开发了一些预定义的线框对脸的特征进行编码，这种编码效率很高，只需少数比特就能描述其特征。

对于人脸的表情（如生气、高兴等）可能的行为可用语义编码，由于物体可能的行为数目非常小，可获得非常高的编码效率。

MPEG-4 采用的编码方法就既基于块的混合编码，又有基于内容的编码方法。

1.4.4 三维（立体）视频编码

立体视频编码是视频编码的发展方向之一，它在平面信息外增加了深度信息，数据量非常庞大。

立体视频编码也有两种类型的方法：第一种是基于波形的，它组合了运动补偿预测和位差补偿预测。所谓位差估计即在两幅不同图像中寻找对应的点，对预测残差图像、位差和运动矢量进行编码；第二种是基于物体的，它直接对成像物体的三维结构和运动进行编码。当物体结构简单时，可获得非常高的压缩比，其编码结构如图 1.6 所示。

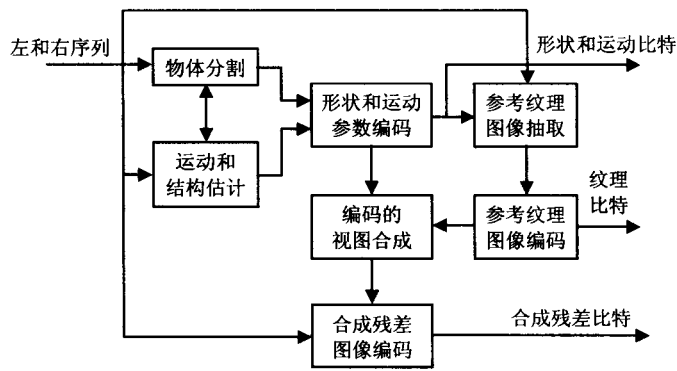


图 1.6 三维物体序列编码系统

参 考 文 献

- 1 Yao Wang, Jorn Ostermann, Ya-Qin Zhang. Video Processing and Communication. Pearson Education, 2002
- 2 毕厚杰. 多媒体信息的传输与处理. 北京：人民邮电出版社，1999
- 3 毕厚杰. 信息和信息化. 中国工程科学，2003（5），92

第2章 数字视频

视频压缩编码技术就是对数字视频信号进行压缩和解压缩的技术。因此讨论视频压缩编码前,必须先了解数字视频信号的概念和构成。这是本章的主要内容。

什么是数字视频信号?它是自然的电视景象的数字表示。具体说来,在本章中将依次讨论:(1)数字电视的概念;(2)彩色空间;(3)数字电视景象标准格式;(4)A/D和D/A变换;(5)取样和亚取样;(6)量化;(7)数字视频的质量。

2.1 数字电视的基本概念

2.1.1 数字电视的优越性

现在,模拟彩色电视已经相当普及,在一定程度上满足了人们的生活需要。但是,模拟电视的缺陷日益暴露,为此,数字电视就应运而生。与模拟电视相比,数字电视具有许多突出的优点。

(1) 失真小、噪声低、质量高

模拟电视信号在放大、处理、传输、存储过程中,难免会引入失真和噪声,而且多种噪声和失真叠加到电视信号上后,不易去除,并会随着处理次数和传输距离的增加不断累积,导致图像质量及信噪比的下降。

相反,数字电视信号没有上述的噪声累积效应。如图 2.1 所示,只要噪声电平不超过有用信号脉冲幅度的一半,就可用脉冲再生技术对其进行整形,并恢复成“0”或“1”两种电平,不会引入噪声。这样说来,是否数字电视信号就没有任何失真和噪声呢?从下面的叙述可知,它会引入“量化噪声”,这是因为信号的数字化过程必须要经过取样、量化、编码三个基本步骤。“量化”是不可缺少的,量化噪声就不可避免,但可控制在相当低的电平以下。

可见,数字电视的第一个突出优点是噪声低、失真小、视频质量好。

(2) 易处理、易校正

模拟电视信号要进行压缩编码处理、加密处理、校正处理都不是一件容易的事情。

数字电视信号利用 VLSI 芯片进行压缩编码处理、加密处理、白平衡调整、 γ 校正、彩色校正和轮廓校正,相对来说容易得多。随着大规模集成电路设计水平和工艺的不断发展,现在利用专用芯片和通用 DSP 来实现以上处理的成本不断下降,这是视频数字压缩编码能取得不断发展的一个重要原因。

(3) 容量大、节目多

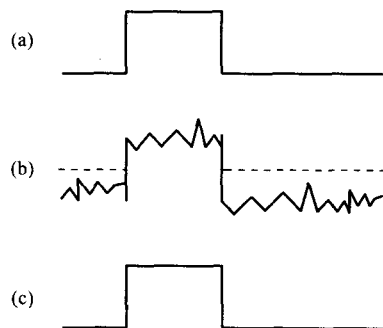


图 2.1 数字电视信号抗噪声能力