

视频压缩漫谈

谢青林 彭卫红 张金标

(广东省气象影视中心, 广州 510080)

提 要

视频压缩是数字视频处理过程中不可或缺的技术, 视频压缩的过程是消除存在于视频信号里的冗余成分, 减少图像或图像组的内容信息, 视频编码器可以分为帧内编码器和帧间编码器两种, 常见的编码器有 M-JPEG, MPEG-1, MPEG-2, MPEG-4 等, 各种编码器各有其特点和应用领域。

关键词: 视频压缩 编码器 帧内编码 帧间编码

近年来, 非线性编辑系统特别是基于微机的非编系统被大量引入到气象电视节目制作中, 对提高制作效率和节目质量起了很大的作用。非编系统中不可避免地涉及到视频压缩的问题, 我们根据收集到的资料和使用体会整理了这篇文章, 供大家参考。

1 视频压缩的原因

未压缩的电视视频数据量是巨大的, 例如, PAL 制式的视频图像以 720×576 的分辨率, 24bits/像素, 每秒 25 帧的质量传输时, 其数据率达 29.7M 字节/秒或 237M 位/秒。若再考虑音频信号的传输, 则数据量更大, 多数情况下传输数字图像的带宽不能达到传输未压缩图像的要求。例如, 目前 IDE 硬盘的数据传输速度大约是 66.7M ~ 133M 位/秒, 远小于 237M 位/秒。此外, 若以 29.7M 字节/秒的数据率计算, 1 分钟的未压缩视频图像占用 1782M 字节的存储空间, 尽管近年来存储技术有了长足的发展, 但相对于如此巨大的数据, 大多数台式计算机来是不可接受的。因此, 视频压缩是视频处理过程中不可或缺的技术。

2 视频压缩的原理

实际上视频信号的压缩并不是一个新问题, 而是随着彩色电视的引入应用。在 PAL

制式中, 彩色图像采用色彩 YUV 色彩空间表示, YUV 色彩空间是将亮度信号 Y 和红差信号 R-Y 及蓝差信号 B-Y 进行编码, 然后用同一信道发送, 这种方式解决了黑白电视和彩色电视兼容性, 便于大面积着色的实现。所谓大面积着色就是利用人眼对亮度信号敏感而对颜色信号不敏感的原理, 将亮度信号全额发送而对颜色信号配额发送。我们平常见到的 4:2:2 就是一种色彩编码方式, 通常所用的色彩编码方式有 4:4:4、4:2:2、4:1:1、4:2:0, 除第一种外, 其他几种都使用了大面积着色原理, 对图像都有压缩能力; 比如 4:2:0 就是指在水平和垂直方向上每两个亮度信号共用一个 U 样本和一个 V 样本, 平均每个像素用 1.5 个字节表示, 原本一个像素需要 3 个字节, 现在减少了一半, 压缩率为 2:1。

随着数字视频技术的发展, 通过模拟信号转换为数字信号, 获得更稳定连续的图像质量及更高更高的压缩串。压缩基本上是这样—一个过程: 通过消除存在于视频信号里的冗余成分, 来减少图像或图像组的内容信息。实际的节目内容中有两种信号分量: 一种是异常而无法预见的, 另一种是可以预见的。异常分量叫作熵, 它是信号中的真实信息。

余下的部分叫作冗余,它不是必需的。压缩系统的工作方式是编码器将熵从冗余中分离出来,只有熵被录制或传输,而解码器则计算传输信号中的冗余。理想的编码器会提取出所有的熵,但实际上没有理想的编码器。

编码器可以简单地分为帧内编码器和帧间编码器。帧内编码对图像中的空间频率进行分析,使用诸如弱波和 DCT(离散余弦变换)的变换对图像进行压缩。帧间编码则依赖于找到连续画面的相似之处。如果解码器中有了一个画面,那么下一个画面可以通过仅仅发送画面差异来创建。当物体移动时,画面差异会增加,但由于移动物体在画面之间一般不大改变其外形,所以画面差异的大小可以通过运动补偿来抵销,通过将前面画面中的部分内容移动到新位置上的方法来创建当前画面中的近似值,这个移动处理过程由通过传送到解码器中的矢量来控制,矢量传送比发送画面差异数据所需的数据要小得多,从而使图像数据得到压缩。

3 常见的视频编码器简介

典型的帧内编码格式是 MOTION-JPEG 压缩技术,它是基于静态视频压缩发展起来的技术,它的主要特点是不考虑视频流中不同帧之间的变化,只单独对某一帧进行压缩。可以获取清晰度很高的视频图像,而且可以灵活设置每路视频清晰度和压缩帧数。但由于没有考虑到帧间变化,造成大量冗余信息被重复存储,因此单帧视频的占用空间较大,在保持相同清晰度的前提下,整体数据压缩率明显偏低,数据存储耗费大量的硬盘空间。但在非线性编辑系统中可很容易地实现精确到帧的定位,因而在非线性编辑系统中被广泛运用。

常见的 MPEG 压缩是帧间编码,MPEG(即 Moving Picture Experts Group 运动图像专家小组)是个国际标准。目前至少有 MPEG-1, MPEG-2, MPEG-4 三个标准与视

频压缩有关。

MPEG 采用非平衡编码,这意味着压缩一幅图像比解压缩慢的多,MPEG 压缩将数字视频按时序分组,然后每个图像组(GOP: group of pictures)选定一个基准图像利用运动估计减少图像间的时间冗余,最后将基准图像和运动估计误差进行离散余弦变换(DCT: discrete cosin transform)、系数量化和熵编码(VLC&RLC: 'variable length coding' and 'run length coding')以消除空间冗余

MPEG 图像编码包含 3 个成分:I 帧,P 帧和 B 帧。MPEG 编码过程中,一些图像压缩成 I 帧,一些压缩成 P 帧,另一些压缩成 B 帧。I 帧压缩去掉图像的空间冗余度,P 帧和 B 帧去掉时间冗余度。I 帧压缩采用基准帧模式,使用离散余弦变换(DCT)进行帧内压缩,即仅仅考虑了帧内的图像。I 帧压缩不能除去帧间冗余。I 帧压缩在不产生可觉察的模糊现象同时可以得到 6:1 的压缩比。P 帧采用提供帧间编码,利用相邻帧的一般统计信息预测当前帧与前面最近的 I 帧或 P 帧的差别,P 帧可达到比 I 帧更高的压缩比。B 帧为双向帧间编码,基于当前帧与前一帧和后一帧图像之间的差别进行压缩,B 帧压缩可以达到 200:1 的压缩比,其文件尺寸一般为 I 帧压缩尺寸的 15%,不到 P 帧压缩尺寸的一半。B 帧和 P 帧要求计算机有更强的功能。

MPEG-1 标准适用于数码率在 1.5Mbit/s 左右的应用环境,也就是为 VCD 所制定的。MPEG-1 的基本算法对于压缩水平方向 352 个像素、竖直方向 288 个像素的空间分辨率及每秒 24~30 帧幅的运动图像有很好的效果。在 MPEG-1 的标准中的帧是逐行扫描的图像,如果待处理信号是隔行扫描的图像,则编码前必须将其转换成逐行扫描的格式。

MPEG-2 是由 MPEG 开发的第 2 个标

准,是“活动图像及有关声音信息的通用编码”(Generic Coding of Moving Pictures Audio Information)标准。MPEG-2 用于宽带传输的图像,图像质量达到电视广播甚至 HDTV 的标准。和 MPEG-1 相比,MPEG-2 支持更广的分辨率和比特率范围,成为数字图像盘(DVD)和数字广播电视的压缩方式。MPEG-2 标准制定了一个涵盖大部分应用的型/级体系。每个型都是 MPEG 语法的一个完整的子集合,型中选定不同的参数形成不同的级。MPEG-2 标准共分 5 型:简单型(SP),只有基准帧 I 和预测帧 P;主型(MP),比 SP 增加了双向推测帧 B;信杂比分层型(SNRP);空间可分层型(SSP)和高型(HP)。

MPEG-2 标准共分 4 级:低级(LL),输入信号的像素为 ITU-R 601 格式的四分之一;主级(ML),输入信号的像素为 ITU-R 601 格式;高级-1440(H14L:)为 4:3 模式电视高清晰度格式;高级(HL)为 16:9 模式电视的高清晰度格式。

MPEG-4 是近年来比较热门的 MPEG 压缩标准,是一种适用在低带宽下进行信息交换的音视频处理技术,可以动态的侦测图像各个区域变化,通过基于对象的而调整压缩方法的技术获得比 MPEG1 更大的压缩比,更低的压缩码流。MPEG-4 引入 AV 对象(Audio/Visaul Objects),使得更多的交互操作成为可能。MPEG-4 目前广泛应用于视频点播,网络流媒体传输,监控系统,视频会议等场合。在电视节目制作中,许多非编网络系统将其作为低质量载体,供无卡工作站粗编。

表 1 是 MPEG1 MPEG2 MPEG4 的一些技术参数。

当然,图像压缩标准不仅这些,H.261, Cinepak 和 Indeo 都是常见的视频压缩技术。

表 1 MPEG1 MPEG2 MPEG4 技术系数

	MPEG1	MPEG2	MPEG4
开始成为标准的年份	1992	1995	1999
最大视频分辨率	352×288	1920×1152	720×576
缺省视频分辨率(PAL)	352×288	720×576	720×576
缺省视频分辨率(NTSC)	352×288	640×480	640×480
最大音频范围	48 kHz	96 kHz	96 kHz
声道通道的最大数目	2	8	8
最大数据率	3 Mbit/秒	80 Mbit/秒	至 10 Mbit/秒
	1380 kbit/s	6500 kbit/s	880 kbit/s
一般使用的数据率	(352 × 288)	(720 × 576)	s (720 × 576)
帧频(PAL)	25	25	25
帧频(NTSC)	30	30	30
视频质量	满意	很好	好、很好

H.261 与 Motion-JPEG 非常相似,但它支持通过 P 帧的帧间压缩,不支持 B 帧压缩。其压缩率很高,但高压压缩率是以牺牲部分图像质量为代价的。当图像质量和运动很重要时,H.261 将不再是好的选择。

Indeo 允许双向预测(B 帧)和缩放。它支持 B 帧编码,但速度很慢。Indeo 的 B 帧压缩会造成帧丢失。缩放时还会造成突变边缘有可觉察的象素化现象及帧的丢失。不用 B 帧压缩和缩放功能时,320x240 分辨率,每秒 15 帧的图像可以压缩到每秒 200Kbytes。

Cinepak 是由 Radius 公司发展的一种压缩技术,压缩速度很慢。一般它提供每秒 15 帧的 CD 而不象 MPEG 为每秒 30 帧。

视频压缩是一项复杂的技术,涉及的内容远非一篇文章可以写完,本文试图对其进行一次全景式的描绘,希望它能让大家对此有所了解,并在节目制作中有所帮助。

参考文献

- 1 刘惠芬编著. 数字媒体传播基础. 北京:清华大学出版社,2000 年.
- 2 李昱,周进,编著. 数字影视后期制作高级教程. 北京:中国青年出版社 2001 年.