

视频细粒度可分级编码研究进展^{*})

王相海

(辽宁师范大学计算机与信息技术学院 大连 116029)

(计算机软件新技术国家重点实验室(南京大学) 南京 210093)

摘要 Internet 的迅速发展使越来越多的应用使用流媒体技术。作为流媒体的核心技术之一,视频的可分级编码技术已经成为一个重要的研究领域。本文首先对 MPEG-4 修订版中 FGS 的编码机制、可扩展特性和所存在的问题进行了讨论,然后对细粒度可分级视频编码的研究进展进行了分析,最后对视频细粒度可分级编码的未来发展趋势进行了展望。

关键词 视频编码,细粒度可分级,运动补偿,宏块

Research Progress on Fine Granularity Scalability Video Coding

WANG Xiang-Hai

(College of Computer and Information Technology, Liaoning Normal University, Dalian 116029)

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract With the development of Internet, more and more applications start to use the media streaming. The video scalable coding technique, one of core technology in media streaming, have been thought as the important study field. This paper first provides an overview of the fine granularity scalability (FGS) video coding technique in the Amendment of the MPEG-4, and discusses the several extensible features of FGS coding and its defects. In addition, the research progress on fine granularity scalability video coding for Internet is analyzed. Finally, the state of arts and possible new directions of FGS are stated.

Keywords Video coding, Fine granularity scalability, Motion compensation, Macro block

1 引言

视频点播(VoD)、视频会议、远程教育和数字图书馆等网络多媒体应用的迅速发展,导致了流媒体(streaming media)研究的迅速发展。在流媒体技术的研究过程中,流媒体的编解码技术作为最重要的核心技术之一,近年来得到了蓬勃发展^[1],出现了用于 64kbps 视频传输的 H. 261、面向 1.5Mbps 数字视频音频传输和存储的 MPEG-1、面向高品质数字视频音频传输和存储的 MPEG-2、适应于低码率视频编码的 H. 263 以及基于内容视频编码的 MPEG-4^[2]。而到目前为止,最有代表性的流媒体编解码技术为视频的可分级(Scalable)编解码技术。一个视频编码比特流具有可分级性(Scalability)是指利用该比特流的一部分(或子集)可以产生对该视频的一个有用表示的能力,即解码器能够根据分辨率的情况对码流的一部分进行解码。视频的可分级性通常有 3 种情况,即空间分辨率可分级、时间分辨率可分级和质量可分级(或称 SNR 可分级)^[3,4]。空间分辨率可分级是指以不同空间分辨率解码图像的能力;时间分辨率可分级是指视频帧率的更新可以被调整;而质量可分级则是指对一个压缩的码流,用户可根据带宽或系统能力的具体情况在一定范围内按任意比率从码流的开始端来截取相应的一段码流进行解码。

视频的可分级编码思想在 MPEG-2 中已经有所体现,但 MPEG-2 和 MPEG-4 早期的可分级编码机制采用的是基于

DCT 的分层编码模式,即视频流被编码为一个基层和一个增强层,在解码过程中首先解码基层码流,然后根据解码分辨率的情况,通过解码增强层来提高解码精度。上述分层可分级编码技术的一个共同特性是要想使得解码精度得以提高,增强层的码流必须完全解码,否则不会增强解码的质量,这样一般不能很好地实现解码器端更细粒度的解码,并且对帧内和预测帧差采用传统的分块 DCT 编码所产生的“块效应”往往会使可分级编码不能取得很好的视觉效果,特别对低码率情况。

MPEG 组织注意到了网络传输对视频编码的新的要求,开始征集精细的可分级视频编码方案,该方案要求将视频编码成一个可以单独解码的基层码流和一个可以在任何地点截断的增强层码流,即实现解码的“连续可分级”,其中基层码流适应最低的网络带宽,而增强层码流用来适应网络带宽变化的动态范围。世界各地的学者提出了许多解决方案,最终 Weiping Li 的 FGS 方案^[5]被选中。MPEG-4 在其修订版本中提出了该编码方案,它突破了传统视频编码方法的局限性,以其编码的高效性和可分级的灵活性而受到关注。近年来,人们对该领域进行了积极的研究,不断完善和改进了所提出的 FGS 编码方案,出现了许多有效的 FGS 编码策略。本文首先介绍了在 MPEG-4 的修订版中所给出的 FGS 的编解码机制,分析了它的可扩展特性和存在的不足,然后对视频细粒度可分级编码技术的研究进展进行了分析和讨论,并提出了

^{*})国家自然科学基金项目(60372071)、辽宁省自然科学基金项目(20032105)、辽宁省高等学校优秀人才支持计划和大连市科技基金项目资助。王相海 博士、教授,主要研究领域为多媒体信息处理、多媒体数字水印和计算机图形学。

视频细粒度可分级编码的下一步发展方向。

2 MPEG-4 中 FGS 视频编码情况分析

2.1 FGS 视频编码的基本思想

FGS 视频编码的基本思想是将一个视频序列编码为一个基层码流和一个增强层码流,基层采用非可分级编码方式进行编码,其码率作为解码端可分级解码的下界;增强层则采用了基于 DCT 系数的位平面编码机制,对原始图像和重构图像的帧差图像进行编码。MPEG-4 修订版中 FGS 编码器和解码器框架参见图 1 和图 2。

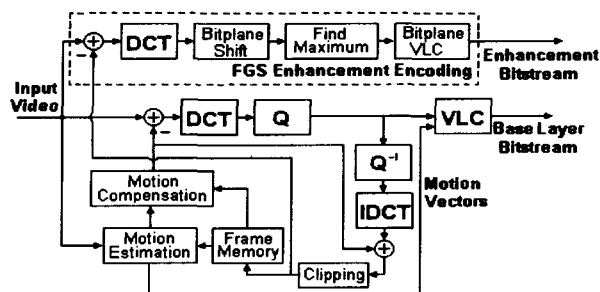


图 1 FGS 编码器结构

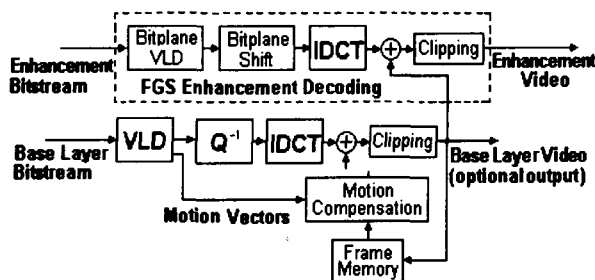


图 2 FGS 解码器结构

尽管上述 FGS 编码机制也是将视频序列分为两层,但与传统的可分级编码不同,FGS 增强层可以根据对每一帧的解码精度要求,任意截取其码流,从而实现连续可分级的解码机制。

在上述 FGS 的编、解码器框架中,其基层采用的是 MPEG-4 ASP 编解码方案,而增强层则是对原始 DCT 系数与基层粗量化系数之差用 Bit-Planed 的编码方式。由于采用了位平面编码技术,使得每一个系数的较高位,即该系数的较重要部分优先得到编码。这样,在码流截断时可以保证当前所获得的解码信息是重要信息,从而提供了精细可分级的特性。

2.2 FGS 中所采用的 Bit-Plane 编码思想

FGS 的基层编码与 MPEG-4 中非可分级编码是相同的,都是由运动估计、运动补偿、DCT 变换、量化和变长编码组成,而其增强层编码时对参差图像则采用了基于 8×8 分块的 Bit-Plane 编码。

Bit-Plane 编码方式是将每一个量化的 DCT 系数看成一个由若干个 bit 组成的二进制数^[5,6],对每一个 8×8 的 DCT 块,64 个绝对值系数被沿着“Z”字形的顺序排成一个序列。所个块的 Bit-Plane,是指这 64 个系数的具有相同重要位置的比特位所形成的数组。基于 Bit-Plane 模式的编码是首先从最重要的 Bit-Plane(MSB)开始,形成(RUN, EOP)符号,其中“RUN”表示一个“1”之前连续“0”的个数,而“EOP”取

“0”或“1”则分别表示该 Bit-Plane 中自该“1”后没有“1”和存在“1”;然后再对这些(RUN, EOP)符号连同其符号位进行变长编码。下面的例子给出了形成(RUN, EOP)的过程。

假设一个块的 64 个系数的绝对值经“Z”字形排序后为: 8, 0, 5, 0, 0, 4, 2, 2, 0, 1, 0, 0, 5, 6, 0, ..., 0, 0, 由于系数中最大数为 8,其二进制数为(1000),故 4bit-plane 被考虑,具体形式如下:

1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ..., 0, 0 (MSB)
0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 0, ..., 0, 0 (MSB-1)
0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 0, ..., 0, 0 (MSB-2)
0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, ..., 0, 0 (MSB-3)

对每一 Bit-Plane 所形成的(RUN, EOP)如下,全为“0”的 bit-plane 可用 ALL-ZERO 表示:

(0, 1) (MSB)
(2, 0), (2, 0), (6, 0), (0, 1) (MSB-1)
(6, 0), (0, 0), (5, 1) (MSB-2)
(2, 0), (6, 0), (2, 1) (MSB-3)

2.3 FGS 的可扩展特性

前面所讨论的 FGS 只是考虑了要编码系数的值,而实际上上述的 FGS 编码机制中蕴含了许多很好的性质。利用这些性质可以使编码更加灵活,同时可以进一步改进编码效果。

(1) 频率加权特性

所谓频率加权特性是指对不同的 DCT 系数加以不同的权重,从而使不同的 DCT 系数上升不同的位平面^[7,8]。因为不同的 DCT 系数对重构画面的可视质量的影响是不同的,比如一般低频 DCT 系数的精度要较高频 DCT 系数更加重要。这样,若通过加权的方式使重构图像的质量相对重要的频率系数尽可能早地放到增强比特流的前端,势必会在可分级截取比特流时增加了它们被选中的概率,如图 3 所示。

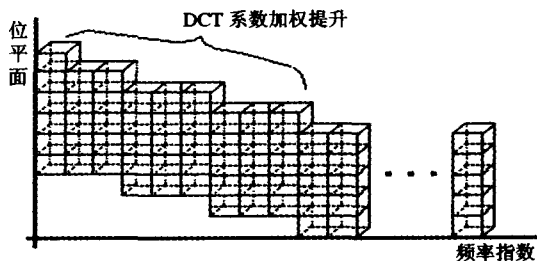


图 3 DCT 系数加权位平面上升示意图

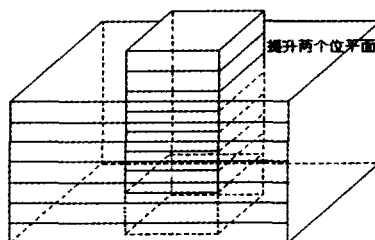


图 4 选择增强特性示意图

(2) 选择增强特性

MPEG-4 编码与 MPEG-1 和 MPEG-2 的最大区别是前者提供了基于对象的编码机制。与 JPEG-2000 中 ROI 的 Maxshift 编码方法^[9]相类似,可以通过对感兴趣区域宏块中系数的位平面进行“移位”提升,即对这些系数进行“增强”处

理(参见图4),从而保证了在对图像的DCT系数进行位平面编码时,感兴趣区域的系数优先得到传输。

(3) 码流的错误恢复特性

当FGS码流在信道特别是无线信道上传输时,同样可能会受到信道噪声的干扰。为了很快地孤立,FGS在其增强层的码流中设置了再同步标志 fgs_bp_start_code 和 fgs_resync_marker,图5给出了带有再同步标志的FGS码流的形式。实际上,fgs_bp_start_code除了可以用于再同步标志而使码流尽快从错误中恢复过来外,它还可以帮助服务器或解码器在不用完全解码所有的VLC码字的情况下可以尽快地识别出一个位平面的开始位置。标志位的具体使用参见文[10]。

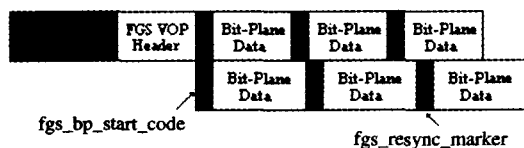


图5 FGS增强层码流图同步标志位示意图

(4) 时间可分级特性

在前面FGS编码结构中,由于对时间增强帧的时间预测是在基层进行的,这样每个时间增强帧的质量不会影响到其它帧的质量,从而可以在上述FGS编码结构中结合时间可分级编码机制。这种混合的可分级编码方案被称作FGST^[11]。采用该可分级编码方案,不仅保持了FGS的精细可分级特性,而且支持帧率的变化。图6、7给出了FGST示意图。

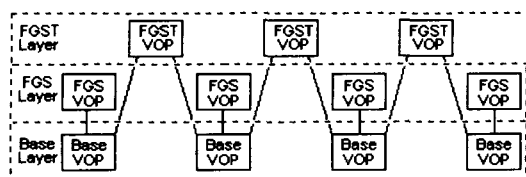


图6 FGST与FGS相分离的编码结构

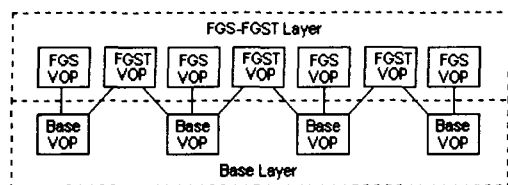


图7 FGST与FGS相合并的编码结构

3 细粒度可分级视频编码研究进展

前面介绍的FGS编码结构的一个主要特点是预测帧中的基层和增强层都是以重构的参考帧中的基层作为预测信息进行预测的,这样FGS编码方案对于偶尔的数据丢失或增强层中的数据错误通常具有很好的错误恢复能力。此外,由于利用基层信息来预测增强层,从而使得在码流传输过程中增强层信息的丢失或被侵蚀不会像传统的SNR可分级编码方案那样引起误差漂移现象。然而,同样是由于预测总是在于最低质量的基层上进行,这样对增强层的预测一般不是很准确,形成的DCT参差较大,编码效率不是很高,有时甚至比传统的SNR可分级方案的编码质量还要糟糕^[12]。近年来,

人们在FGS编码结构的基础上对细粒度的视频可分级编码技术进行了较深入的研究,提出了许多有效的改进算法。以下我们对目前最为典型的几种改进算法进行分析和讨论。

3.1 渐进的精细可分级编码(PFGS)

为了改进FGS的编码效率,微软亚洲研究院的李世鹏和吴枫等人提出了一种渐进的精细可分级编码方案(PFGS, Progressive Fine Granular Scalable)^[13,14],该方法在保持了FGS所具有的细粒度可分级特性、对信道的自适应特性和对错误的可恢复能力的同时,提高了较FGS方案1dB左右的编码效率。

(1) 基本的PFGS框架

PFGS编码结构有两个关键点:一是使用了两种不同质量的参考图像,即基本层图像使用前一帧重建的基本层图像为参考,进行运动估计和补偿,而对于增强层图像则是尽可能地采用前一帧某个增强层图像为参考进行运动估计和补偿;二是保持了一个从基本层到最高质量层的预测路径,这使得PFGS编码方案能够从包丢失或数据错误中得以恢复,高层信息的丢失或误差也可以通过从低层开始跨越几帧渐渐地得以恢复,而不会带来错误的漂移。

图8给出了PFGS的基本编码结构。其中,不同帧间的实线箭头表示当前帧利用前一帧中的基层或增强层进行预测的情况,比如第3帧是利用第2帧的基层和奇数增强层来进行预测的;图中的虚线追踪了实际被传输的视频层的情况,比如在第2帧处带宽发生了减少,服务器便降低了所传输的高层比特数,而到了第3帧,随着带宽的逐步回升,所传输的增强层的比特数也逐渐增多。

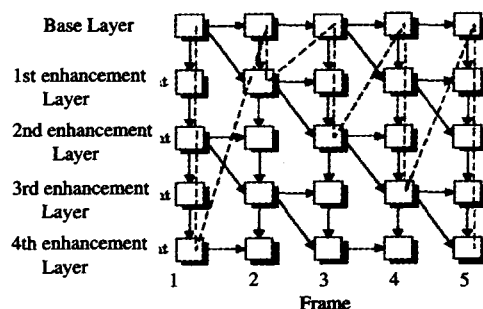


图8 PFGS的框架

(2) 对上述基本PFGS框架的改进

尽管图8所给出的PFGS编码框架具有码流自适应信道的变化和错误恢复的优点,但它却需要较多的额外的帧缓存器来保存多个重构的增强层,以便对后面帧进行预测,从而增加了PFGS编码器和解码器的内存消耗和计算复杂度。此外,采用多个不同的参考图像时,当从一个低质量的参考图像过渡到一个高质量的参考图像,将会产生一个DCT系数的波动,从而会对编码效率产生影响。

幸运的是,并不是所有的增强层图像都适合做参考图像。事实上,由于相邻帧间仅仅中间增强层图像的DCT系数才具有很强的相关性,这样只有通过使用中间增强层图像作为参考图像,一般才能使编码效率得到改进,这样只需要增加一个缓冲区来保存高质量的参考图像;此外,为了克服从低质量的参考图像到高质量参考图像所带来的数据波动问题,可以采用利用高质量的参考对所有增强层进行编码,而利用低质量的参考替代高质量的参考来重构解码图像,从而解决了误差传播问题^[14]。具体改进的PFGS编码框架参见图9。其中,

实线实箭头表示相邻两帧和帧内的相邻层的预测关系;实线虚箭头表示用低质量的基层来重构一个增强层,作为下一帧的高质量的参考;虚线虚箭头表示在得不到高质量的参考时,用低质量的参考来代替。

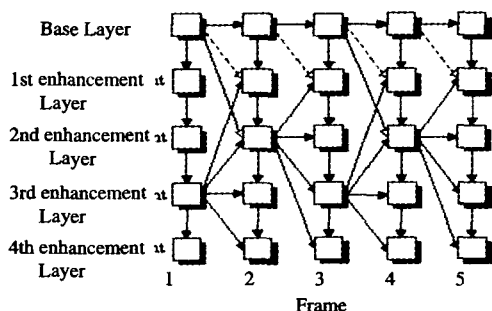


图9 改进的PFGS框架

(3)讨论

与FGS编码方案类似,PFGS也将视频帧编码成了相对低质量的基层和可以逐渐提高质量的、多个任意的、具有可分级特性的增强位平面层。与FGS不同的是,PFGS编码方案在增强层编码中有选择地使用了基层参考帧的同时,更多地使用了高质量的参考帧,进行运动补偿和重建,从而提高了编码效率。另一方面,PFGS还采用了利用高质量的参考对所有增强层进行编码,而利用低质量的参考替代高质量的参考来重构解码图像的机制,有效地克服了采用不同类型参考图像转换所带来的在低码率下的误差传播。然而,上述PFGS编码机制也存在着一些问题:一是由于采用低码率的基层参考帧进行重建,得到的高质量参考帧也存在着一定的质量损失,这在一定程度上影响了PFGS的编码效率;另一方面,图12的PFGS编码框架在增强层每一帧编码的运动补偿和重建过程中,预测参考和重构参考的选择是基于整帧图像的。也就是说,在增强层编码的预测和重建过程中,或者全部使用低质量的参考帧,或者全部使用高质量参考帧。这样,PFGS编码方法很难在有效控制误差传递和累积的同时获得较高的编码效率。

3.2 基于宏块的渐进细粒度可分级编码

所谓基于宏块的渐进细粒度可分级编码^[16,17]是指在上述PFGS的运动预测和重建中,在宏块一级选择所使用的参考信息,即每个增强层宏块在运动补偿和重建中使用的参考宏块可灵活地在高质量参考帧和低质量的参考帧间选择,从而给PFGS方案带来更大的灵活性,并且能够更好地在消除误差传递和提高编码效率间寻求平衡。

(1)增强层宏块的3种编码方式

在文[16,17]中,根据宏块运动补偿和重构时所选用的参考图像的不同,提出了3个增强层的帧间编码方式(参见图10,其中实心箭头实线表示运动补偿时所用的参考图像,虚箭头的实线表示重构当前高质量参考时所用图像)。在LPLR方式中,增强层宏块在运动补偿中使用前一个低质量的参考图像进行预测,并使用相同的低质量参考图像进行重构。这种编码方式不会产生误差传播和累积。但如果所有的增强层宏块都使用这种方式进行编码,则PFGS编码方式也就退化为FGS方式了,其编码效率是最低的;在HPHR编码方式中,增强层宏块在运动补偿时使用前一个高质量的参考图像,并使用相同高质量的参考图像进行重构。如果所有的增强层宏块都采用该方式进行编码,那么PFGS在高码率下

将得到最高的编码效率。然而,该方式当网络带宽下降或出现传输错误而无法得到高质量的参考时,解码器将不得不采用低质量的图像来作为参考图像,这样将会带来解码误差以及误差的传播。针对上述两种编码方式所存在的问题,HPLR将PFGS中所采用的在增强层编码中使用高质量的参考图像进行运动补偿,而是用低质量的参考图像进行重建的编码机制扩展到宏块一级进行控制,从而有效地消除了误差的积累和传播。但由于HPLR编码方式中采用了低质量的参考图像进行重构,因此这种方式也会影响PFGS的编码效率。

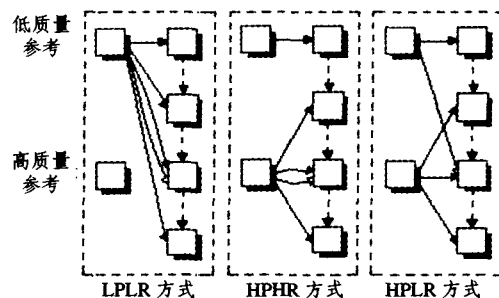


图10 增强宏块的编码方式

综上,上述3种编码方式虽然直观上是将在FGS和PFGS的编码机制简单地扩展到了宏块一级,但这种扩展为在编码过程中灵活地选择最优的编码模式,使得在提高PFGS编码效率的同时有效地控制误差的传播和积累提供了可能。

(2)增强层宏块编码方式的选择

首先利用公式 $\min(\|X_b - \tilde{X}_b\|, \|X_e - \tilde{X}_e\|)$ 来确定某个宏块是否采用LPLR编码方式,其中 X_b 和 X_e 分别表示用低质量参考宏块和高质量参考宏块进行运动补偿所得到的预测误差的DCT系数, $\tilde{X}_b$ 表示基本层的重建DCT系数。当 $\|X_b - \tilde{X}_b\|$ 小于 $\|X_e - \tilde{X}_e\|$ 时,增强层宏块采用LPLR方式,否则进入到下一步的判断。

第二步,利用准则 $\|p_e(n) - p_b(n)\| > k' \times \|x_0 - p_e(n)\|$ 来区分是采用HPHL方式还是HPHR方式,其中 $p_e(n)$ 和 $p_b(n)$ 分别表示高质量和低质量的预测宏块, x_0 是当前原始图像的宏块信息, k' 是可接受的损失因子。具体满足上述公式时,则采用HPLR编码方式,否则采用HPHR方式。

所给出的宏块编码方式选择算法实现起来非常简单,并且非常容易编码。算法的具体理解可参见文[17]。

(3)讨论

上述基于宏块的PFGS编码方案与MPEG-4的FGS和原有的PFGS相比,在低码率下基本保持原有的编码性能,而在中、高码率时,其编码效率要较PFGS高出1.1dB,较MPEG-4的FGS高出1.5dB。

3.3 基于宏块的时域可分级编码(PFGST)

在上述基于宏块的PFGS方案的基础上,文[18]给出了一种同时支持时域和精细SNR可分级的PFGST编码方案。

(1)PFGST的基本思想

与前面所介绍的FGST方案相类似,PFGST方案中基层仅包含以基本层帧率进行编码的I帧和P帧,所有的时域可分级帧采用B帧形成增强层码流。其中I帧和P帧的编码与基于宏块的PFGS相同,B帧的方法与FGST类似。经过双向预测后,其帧差使用位平面技术形成时域可分级的增强层码流。与FGST不同的是,在B帧的编码过程中,每个宏块

都可以选择 SNR 可分级的低质量或高质量的参考宏块来进行运动补偿,因而可以获得更好的编码性能。

(2) 时域增强层宏块编码的两种模式

根据使用的参考宏块的不同, PFGST 方案提出了两种时域增强层宏块的编码模式: LP 模式和 HP 模式(参见图 11)。在 LP 模式中, 增强层宏块预测和重建中使用的都是低质量的参考宏块, 而在 HP 模式中增强层宏块预测和重建中使用的则都是高质量的参考宏块。两种编码模式的选择是通过宏块 DCT 域预测残差的绝对值的均值来确定的。

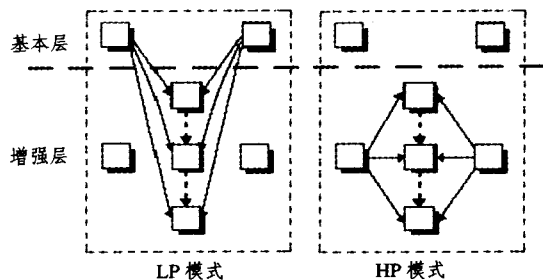


图 11 PFGST 方案中增强宏块的两种编码模式

(3) 讨论

与 MPEG-4 中的 FGST 相比, PFGST 编码方法能提高编码效率 2.8dB, 并且同样具有 FGST 所提供的时域可分级和 SNR 可分级的特性。

3.4 精细的空域可分级编码 (FGSS)

前面所介绍的 PFGS 编码方案中, 总是假设基本层能够得到充分的解码。这样, 对于带宽小于基本层码流的码率的用户而言, PFGS 就不再具有网络带宽的自适应能力了。为此, 文[19]中提出了一种精细的空域可分级编码方案 (FGSS)。

(1) FGSS 的基本思想

FGSS 将 PFGS 和传统的空域可分级编码相结合, 与 PFGS 相同, 所有基层图像使用重建的基层图像来做运动预测和补偿, 而增强层图像则使用前一帧重建的某个增强层图像来做运动预测和补偿, 即 FGSS 使用两个参考图像, 做两次运动预测和两次运动补偿。与 PFGS 相比, FGSS 的基本层码流的码率更低。而传统的空域可分级编码不同, 低分辨率的视频和高分辨率的视频质量可以随码率的增加连续地加以改进。

(2) FGSS 的体系结构及算法

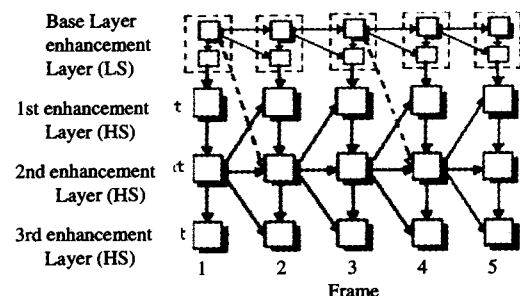


图 12 FGSS 的体系结构

图 12 描述了 FGSS 编码的体系结构。在 FGSS 编码方案中, 输入的视频首先进行下采样, 然后采用任何非可分级编码技术将视频压缩到给定的码率下 (参见图 12 中虚线矩形框内的小矩形), 对增强层则采用位平面编码技术使得要编码的

系数从高位到低位按位平面依次编码, 这使得增强层码流可以在任何地点阶段, 从而具有质量精细变价的特性。

(3) 讨论

FGSS 与传统的空域可分级编码最主要的差别在于增强层的编码。传统的空域可分级增强层一般被看作 B 帧编码, 生成的码流不具有任何可分级性。如果只解码增强层的一部分, 则会产生严重的误差传播。而 FGSS 与 PFGS 一样, 增强层编码使用了位平面技术, 这保证了重要信息尽可能地排在码流的前端, 增强层码流可以解码到任何位置而停止。

在 FGSS 中一个必须解决的问题是如何在增强层编码时选择高分辨率的参考图像以及如何避免误差的传播。与前面图 12 的 PFGS 结构相类似, 方案中采用了第 2 或第 3 个增强层的解码图像作为高分辨率下编码的参考图像, 同时在重建高质量参考图像时交替使用前一帧的高质量参考图像和基本层参考图像, 很好地解决了这些问题。具体原因参见文[14]。

3.5 其它的改进算法

除了前面介绍的典型 FGS 视频编码技术以外, 利用增强层信息进行预测编码的一些改进算法也被提出。比如在文[20]中提出了利用前面增强层的运动补偿重构来预测当前增强层的方案, 但所提出的方法并没有利用当前基层的正差信息。文[21]对该方法进行了改进, 它利用过增强层的重构信息和基层信息, 通过一个估计理论框架对当前帧进行优化预测, 通过优选选择预测帧, 所获得的可分级编码性能较单独使用基层信息进行预测和单独使用增强层信息进行预测提高了几个 dB, 这一方法可以很容易地扩充到 FGS 结构中。该方法的主要缺点是由于在增强层编码中需要用到多个运动补偿环, 因而具有较高的复杂性。G. G. Ding 等人在文[16]的基于宏块的编码方案的基础上, 将漏预测 (Leaky Prediction) 技术^[23]引入其中, 提出了一种既与宏块的 FGS-LP 编码方案, 该方案首先得增强层采用了漏预测帧间模式, 然后采用一种基于时间预测的决策机制选择优化编码模式和预测漏的数量, 该方案进一步提高了文[16]中基于宏块的可分级编码效率。

4 展望

尽管过去几年已经出现了许多有效的视频精细粒度可分级编码方案, 但在许多地方还需要进一步完善。综合前面的分析和讨论, 我们认为未来除了进一步完善已有的细粒度可分级编码技术外, 这种编码技术将会在以下方面得到进一步的发展:

(1) 基于 HVS 的视频细粒度可分级编码技术

纯粹从像素值出发的 MSE 准则与人眼的视觉判断通常有一定的误差。近年来人们利用 HVS 的视觉掩盖效应提出了一些有效的图像编码技术。但在视频编码方面, 所做的工作还不是很多。目前见到一些将 HVS 与小波变换相结合的视频编码方案^[25], 而对于与 MPEG-4 相适应的、基于 HVS 的精细粒度可分级编码技术还很少见到。目前已有的精细粒度可分级视频编码技术总体而言效率还不是很高, 而结合 HVS 的编码技术及有可能在保持视频可分级特性的基础上, 使编码的视觉效果有较大幅度的提高。

(2) 面向对象的视频细粒度可分级编码技术

在 JPEG2000 中提出了基于 ROI 的图像可分级编码机制^[9], 而基于对象的视频可分级编码方面的研究还很少。这一方面是由于视频对象的自动化分割没有标准, 如何建立对象的模型是有待研究的问题; 另一方面, 适应于细粒度可分级编码的视频对象的运动估计也是一个难解的问题。但基于对

象的视频细粒度可分级编码作为与 MPEG-4 相适应的编码技术,具有重要的理论研究价值和实际应用背景。

(3) 空间和质量混合细粒度可分级视频编码技术

所谓空间和质量混合细粒度可分级是指在空间细粒度可分级确定前提下的数率细粒度可分级,或在数率细粒度可分级确定前提下的空间细粒度可分级。Internet 的发展将对视频的可分级编码技术提出更高的要求,信这种混合的细粒度可分级技术一定会得到很好的发展。

参 考 文 献

- 1 钟玉琢,向哲,沈洪. 流媒体和视频服务器. 北京:清华大学出版社,2003
- 2 ISO/IEC JTC1/SC29/WG11. Overview of the MPEG-4 standard. N3536, Beijing, 2000
- 3 Sikora T. MPEG digital video-coding standards. IEEE Signal Processing Magazine, 1997, 14(5): 82~110
- 4 王相海. 基于小波的图像和视频可分级编码研究: [南京大学博士后研究报告]. 2001
- 5 Li W. Overview of fine granularity scalability in MPEG-4 video standard. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(3): 301~317
- 6 ISO/IEC 14496-2/ PDAM4. Coding of Audio-Visual Objects, Part-2 Visual, Amendment 4: Streaming Video Profile, 2000
- 7 Jiang H, Thayer G M. Using frequency weighting in FGS bit-plane coding for natural video. ISO/IEC JTC1/SC29/WG11, MPEG99/M5489, 1999
- 8 Li W. Frequency weighting for FGS. ISO/IEC JTC1/ SC29/WG11, MPEG99/M5589, 1999
- 9 王相海,张福炎. 静态图像编码研究进展. 计算机研究与发展, 2001, 38(11): 1315~1326
- 10 Yan R, Wu F, Li S, et al. Error resilience methods in the FGS enhancement bitstream. ISO/IEC JTC1/SC29/ WG11, MPEG00/M6207, July 2000
- 11 van der Schaar M, Radha H, Chen Y. An all FGS solution for hybrid temporal-SNR scalability. ISO/IEC JTC1/ SC29/ WG11, MPEG99/M5552, Dec. 1999
- 12 Macnicol J, Frater M, Arnold J. Results on fine granularity scalability. Melbourne, Australia: ISO/IEC JTC1/SC29/ WG11, MPEG99/m5122, Oct. 1999
- 13 Li S, Wu F, Zhang Y Q. Study of a new approach to improve FGS video coding efficiency. ISO/IEC JTC1/ SC29/ WG11, MPEG99/M5583, Dec. 1999
- 14 Wu F, Li S, Zhang Y Q. A framework for efficient progressive fine granularity scalable video coding. IEEE Trans on Circuits and Systems for Video Technology, 2001, 11(3): 332~344
- 15 Li S, Wu F, Zhang Y Q. Experimental results with progressive fine granularity scalable (PFGS) coding. ISO/IEC JTC1/SC29/ WG11, MPEG99/ M5742, 2000
- 16 Sun X, Wu F, Li S, et al. Macrobloc-based progressive fine granularity scalable video coding. IEEE International Conference on Multimedia and Expo (ICME), Tokyo, August, 2001
- 17 孙晓艳, 高文, 吴枫, 等. 基于宏块的渐进. 精细可伸缩的视频编码. 软件学报, 2002, 13(11): 2134~2141
- 18 孙晓艳, 高文, 吴枫, 等. 基于宏块的具有时域和 SNR 精细可伸缩的视频编码. 计算机学报, 2003, 26(3): 345~352
- 19 Wang Q, Wu F, Li S, et al. Fine-granularity spatially scalable video coding. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Salt Lake City, 2001, 3: 1801~1804
- 20 Horn U, Stuhlmüller K W, Link M, et al. Robust internet video transmission based on scalable coding and unequal error protection. Signal Processing Image Commun, 1999, 15: 77~94
- 21 Rose K, Regunathan S L. Toward optimality in scalable predictive coding. IEEE Trans Image Processing, 2001, 10(7): 965~976
- 22 Wilson D, Ghanbari M. Transmission of SNR scalable two layer MPEG-2 coded video through ATM networks. In: Proc. 7th Int Workshop Packet Video, Mar 1996, 185~189
- 23 Ghanbari M, Seferidis V. Efficient H. 261-based two-layer video codes for ATM networks. IEEE Trans Circuits Syst Video Technol, 1995, 5: 171~175
- 24 Ding G G, Guo B L. Macrobloc-based fine granularity scalable video coding with leaky prediction. <http://www.compscip-reprints.com/Computer-vision/0306001>
- 25 Wang Z, Lu L, Bovik A C. Rate scalable video coding using a foveation-based human visual system model. In: IEEE International Conference on Acoustics, Speech, & Signal Processing, 2001, 3: 1785~1789
- (上接第 59 页)
- 19 Afrati F N, Gergatsoulis M, Kavalieros T. Answering queries using materialized views with disjunction. In: Proc. of the 7th Int Conf on Database Theory (ICDT'99), Vol 1540 of Lecture Notes in Computer Science, Springer, 1999, 435~452
- 20 Cohen S, Nutt W, Serebrenik A. Rewriting aggregate queries using views. In: Proc. of the 18th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'99), 1999, 155~166
- 21 Grumbach S, Rafanelli M, Tininini L. Querying aggregate data. In: Proc. of the 18th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'99), 1999, 174~184
- 22 Srivastava D, Dar S, Jagadish H V, et al. Answering queries with aggregation using views. In: Proc of the 26th Int Conf on Very Large Data Bases (VLDB'96), 1996, 318~329
- 23 Duschka O M, Genesereth M R. Answering recursive queries using views. In: Proc. of the 16th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'97), 1997, 109~116
- 24 Beerl C, Levy A Y, Rousset M-C. Rewriting queries using views in description logics. In: Proc. of the 16th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'97), 1997, 99~108
- 25 Calvanese D, De Giacomo G, Lenzerini M, et al. Rewriting of regular expressions and regular path queries. In: Proc of the 18th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'99), 1999, 194~204
- 26 Calvanese D, De Giacomo G, Lenzerini M, et al. Query processing using views for regular path queries with inverse. In: Proc. of the 19th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS 2000), 2000, 58~66
- 27 Calvanese D, De Giacomo G, Lenzerini M, et al. View-based query processing and constraint satisfaction. In: Proc. of the 15th IEEE Symp on Logic in Computer Science (LICS 2000), 2000, 361~371
- 28 De Giacomo G. Intensional query answering by partial evaluation. J of Intelligent Information Systems, 1996, 7(3): 205~233
- 29 Lenzerini M. Data integration: a theoretical perspective. In: Proc. of the 21st ACM SIGACT SIGMOD SIGART Symposium on Principles of Database Systems (PODS 2002), Madison, WI, USA, 2002, 233~246
- 30 Hammer J, Brenning M, Garcia-Molina H, et al. Template-based wrappers in the TSIMMIS system. In: Proc. of ACM SIGMOD Conference, Tucson, USA, 1997, 532~535
- 31 Sahuguet A, Azavant F. Building Light-Weight Wrappers for Legacy web Data-Sources Using W4F. In: Proc. of the Very Large Data Bases (VLDB), 1999, 738~741
- 32 Meng X F, Lu H J, Wang H Y, et al. SGWRAP: A Schema-Guided Wrapper Generator. In: Proc. of the 18th International Conference on Data Engineering (ICDE'02), San Jose, USA, 2002, 331~332
- 33 Arasu A, Garcia-Molina H. Extracting Structured Data from Web Pages. SIGMOD 2003, June 9~12
- 34 Abiteboul S, Quass D, McHugh J, et al. The lorel query language for semistructured Data. International Journal on digital libraries, 1997, 1(1): 68~88
- 35 Buneman P. Semistructured data. In: Proc. of the 16th ACM SIGACT SIGMOD SIGART Symp on Principles of Database Systems (PODS'97), 1997, 117~121
- 36 Arocena G O, Mendelzon A O. WebOQL: Restructuring documents, Databases and Webs. In: Proc. of ICDE. Orlando, FL, 1998, 24~33
- 37 Fernandez M, et al. A query language for a web-site management system. SIGMOD Record, 1997, 26(3): 4~11
- 38 Hansen M, Madnick S. Data integration using Web Services. In: Proc. of the VLDB, 2002
- 39 Bayardo J, Bohrer M. InfoSleuth: Agent-Based semantic integration of information in open and dynamic environment. In: Proc. of the ACM SIGMOD, 1997
- 40 Cabrera F, Copeland G. Web services coordination. BEA System & IBM Corporation & Microsoft Corporation, 2002
- 41 Knoblock C A, Minton S. The ariadne approach to web-based information integration. Int'l Journal on Intelligent Cooperative Information Systems (IJCIS), 2001, 10(1-2): 145~169
- 42 Tsur S. Are web services the next revolution in E-Commerce. In: Apers P. ed. Proc. of the 27th Int'l Conf on Very Large Data Bases. Roma: Morgan Kaufmann Publishers, 2001, 614~617