

一种基于单环结构的扩展基本层 FGS 视频编码方法

江涛 张兆扬 马然 石旭利

(上海大学通信学院电子信息工程系, 上海 200072)

摘要 可分级编码是解决 Internet 流视频应用中网络带宽不断波动的一种有效方法,所以 MPEG-4 标准中采用了 FGS(fine granularity scalability)编码方法来获得精细颗粒可分级能力,但其代价是编码效率的下降。为解决此问题,现提出在增强层中采用运动补偿的 MC 加 FGS(motion compensation 加 FGS)结构,用于去除 FGS 方案中增强层在时域上的冗余,以提高 FGS 方案编码效率的双环和单环两种方法。在比较了两种结构各自的优缺点后,选定了一种复杂度小、实现简单、效率高的单环结构,并提出了对单环结构的缺陷进行改善的方法。实验结果表明,该方法的编码性能优于 MPEG-4 FGS 方法。

关键词 流视频 精细颗粒可分级 编码效率 运动补偿 MPEG-4

中图分类号: TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2004)07-0869-04

An Expanded Base-layer FGS Video Coding Scheme Based on One-loop Structure

JIANG Tao, ZHANG Zhao-yang, MA Ran, SHI Xu-li

(Department of Electronic & Information Engineering, School of Communication, Shanghai University, Shanghai 200072)

Abstract Transmitting video streaming over Internet is a challenging problem because the bandwidth of network varies all the time. Scalable coding seems to be an effective solution to resolve the bandwidth varying on the streaming video application over Internet. Though traditional scalable video coding schemes get scalability, the scalability is coarse and need much more memory expense. To obtain fine granularity scalability, the MPEG-4 video coding standard adopts FGS coding scheme. It permits bitstreams after being coded be truncated and transmitted arbitrarily with the available bandwidth. However, the fine granularity scalability of the FGS is obtained at sacrifice of coding efficiency. The reason that causes the FGS coding scheme's coding efficiency is, there exists much temporal redundancy on the enhancement layer. In this paper, two MC FGS structure based on using motion compensation on the enhancement layer are developed to get higher coding efficiency. Both of them exploit the temporal redundancy of the enhancement layer. After comparing both structures, one is recommended by this paper. The experimental results prove that the performance of scheme presented by this paper is better than that of MPEG-4 FGS.

Keywords video streaming, FGS, coding efficiency, motion compensation, MPEG-4

1 引言

传统视频编码的目的就是通过对量化步长的控制来获得一个给定的比特率,但对于 Internet 网络流媒体的应用,视频编码的目的并不是为了获得一个给定的比特率,而是为了获得一给定的比特码率

范围,这是由 Internet 网络本身的特性决定的。由于用户接入网络形式的异构性,致使用户能够获得的可用带宽各不相同,且由于网络带宽本身自有的波动性,使得用户的可用带宽是不断变化的,因此视频编码后得出的码流必须具有可伸缩性,以适应网络带宽的变化;另一方面,由于现有网络本身不提供任何 QoS(quality of service)保障,而只是“尽其所能”

(best-effort),因此丢包是不可避免的,而可伸缩性编码就是为解决带宽波动及丢包问题而提出来的。传统的可伸缩性编码只是简单地分成一个基本层和一个或多个增强层,当可用带宽大于增强层阈值码率时,可以获得最好的增强层的图像质量,当可用带宽大于基本层阈值码率而小于增强层阈值码率时,则仅能获得基本层的图像质量。由此可见,传统的基于图像质量的编码方法 SNR(signal noise ratio)可分级虽然是分层的,但由于每个层的比特率在编码时是预先已经定好的,因此它只能提供固定分级或称为粗糙的可分级能力,而无法精细地匹配网络的即时带宽,且在增强层码率发生跳变的临界点上,图像质量过渡不够平滑^[1]。为获得精细可分级的编码码流,MPEG-4 标准中采用了 FGS 编码方案^[2,3]。MPEG-4 FGS 是一种基于图像质量可分级的视频编码,编码时,它是把原始视频序列压缩成两个码流——基本层码流和增强层码流。其中基本层是采用传统的基于运动补偿的 DCT 来编码,以产生一个固定的小于网络最小可用带宽的低码率码流,并提供给用户可接受的图像质量;增强层则采用位平面技术来编码源图像与基本层重构图像之间的差值^[4],由于位平面技术提供了嵌入的可分级能力,它不仅能够对增强层码流进行任意码率的截取、传输,而且用户接收到的码流越多,图像质量就越好,因此它可提供非常精细的图像质量可分级能力。在鲁棒性方面,由于基本层码率小于网络最小可用带宽,因此基本层码流发生丢包的概率大大降低,同时对基本层码流进行较强的差错复原保护技术可使得

基本层图像质量得到很好的保障,而对于增强层码流可根据应用的具体要求采用不同等级的差错复原机制^[5],但这种 FGS 编码的可伸缩性和鲁棒性是以图像质量下降为代价的。文献^[6]中指出,对某些具有高时域相关性的序列,FGS 编码与非伸缩的传统编码相比,在重建的基本层图像质量上要损失 2~3dB。这主要是因为 FGS 的增强层编码没有使用运动补偿而未能去除增强层的时域相关性的缘故。

为提高 MPEG-4 FGS 的编码效率,故本文提出对增强层也使用运动补偿的“双环 MC 加 FGS”和“单环 MC 加 FGS”两种结构,以通过去除增强层在时域上的相关性来提高 FGS 的编码性能。其中单环结构的增强层可称为扩展基本层,该结构具有开销小、实现更为简单的特点。

2 MC 加 FGS 的结构

图 1(a)是 MPEG-4 建议的 FGS 结构框架(以下简称原始 FGS),图中灰度部分代表实际传输被接收到的用于图像质量增强的比特码流。从图中可以看出,原始 FGS 只对基本层进行了运动补偿,并去除了基本层在时域的相关性,而对于增强层则没有进行任何去除时域相关性的处理,这势必降低了压缩效率。为去除增强层在时域上的相关性,本文认为可用如图 1(b)、图 1(c)所示的两种结构来实现,即用图 1(b)的双环运动补偿结构或图 1(c)的单环运动补偿结构来实现。图 1 中 I、P、B 分别代表 I 帧、P 帧和 B 帧。

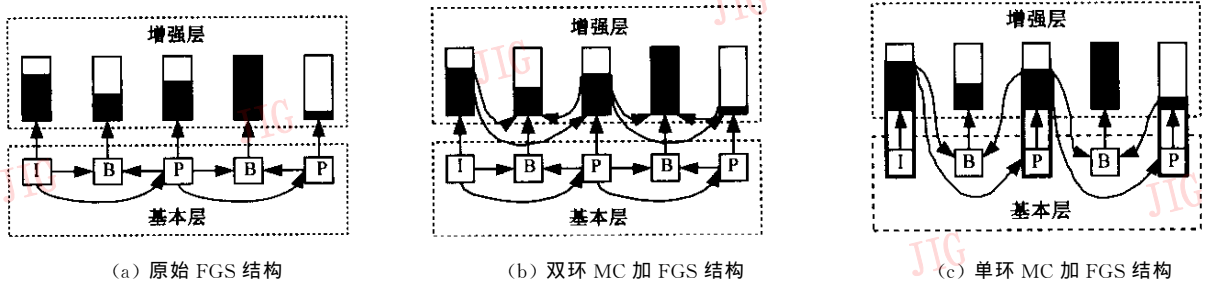


图 1 几种不同的 MC 加 FGS 结构

2.1 双环结构及其编码方案

图 1(b)所示的双环运动补偿结构有两个相互独立的环路——基本层环路和增强层环路。由于计算复杂度集中在运动估计上,因此为降低编码的计算复杂度,增强层环路中所用的运动估计算法和预测模式可与基本层中的所用相同。双环结构优点是:(1)基本层

和增强层是相互独立的,对于基本层来说,完全与原始 FGS 基本层一样,由于无需做任何改变,而仅仅改变增强层的实现方式,因此扩展容易,且具有良好的向前兼容性;(2)增强层传输接收发生的差错不会因累积而影响后续基本层的预测参考。但双环结构有如下的缺点:(1)增强层所用的运动参考只使用了接收到的增强

层码流而不包括接收到的基本层的码流,因此参考不够精确,由于只是部分地去除了增强层时域上的相关性,因此编码效率提高不多。(2)计算复杂度高和编码器结构复杂。图 2 是双环结构方案的编码器框图。由该图可见,与原始 FGS 方案相比,增强层多出了一个运动补偿模块及一个帧存储器模块。图中的 $n(m)$ 代表用于做运动参考的比特平面数。

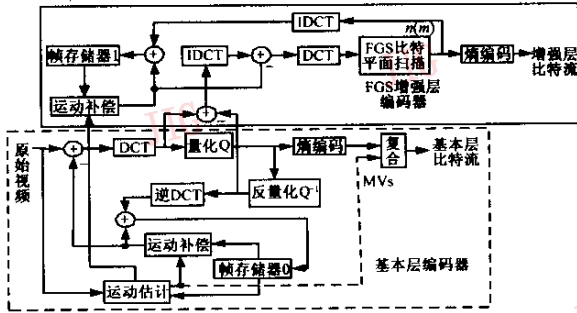


图 2 双环 MC 加 FGS 编码器实现框图

2.2 单环结构及其编码方案

图 1(c)的单环 MC 加 FGS 结构,由于其思想是使用更精确的运动参考来得到更精确的预测,从而可得到更精确的运动补偿,以提高编码效率。图 3 给出单环结构方案的编码器框图,由于其中增强层与基本层共用大多数模块,因此也将增强层称为扩展基本层。基本层和扩展基本层都可被用来对后面基本层的帧进行运动预测和参考。单环结构优点是:(1)由于使用扩展基本层作为运动参考,可提高运动参考的精确度,因此编码效率比双环结构提高得多;(2)复杂度较低、编码器结构简单和实现容易。单环结构也有以下缺点:(1)当容许比特率小于扩展基本层比特率的时候,由于预测漂移(drift)会累积,因而会影响到后面的预测精度;(2)增强层发生的丢包会影响到基本层帧的图像质量。

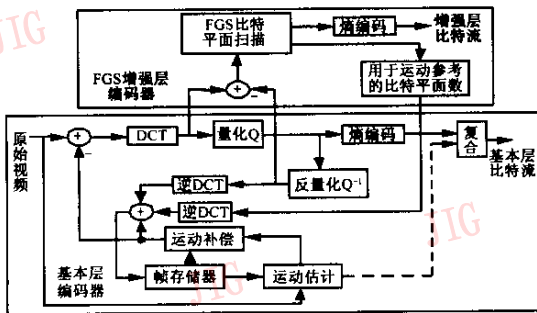


图 3 单环 MC 加 FGS 结构的编码器框图

由于单环结构的实现比双环结构简单、复杂度低,且具有较高的压缩效率,因此单环结构比双环结构

更有实用价值。虽然单环结构也有一定缺陷,但采取以下措施可抑制单环结构的缺陷:(1)减小低比特率时的预测漂移,会对影响到预测精度的 I 帧和 P 帧,其用于运动参考的比特平面数不宜过多(如 2~3 个比特平面)。当容许比特率较高时,用于运动参考的比特平面数可恢复正常;(2)增强层发生丢包或差错时,为减低其对基本层图像质量的影响,可在单环结构中添加一帧存储器用来保存从基本层中重构出的图像,这不仅可以获得一个等同于原始 FGS 结构中的保守图像质量,并可减小其对后面的预测参考所带来的差错累积的影响。这样即使在容许比特率很低时,至少也可以得到原始 FGS 的图像质量。

3 实验结果及分析

为验证 MC 加 FGS 的编码效果,本文对单环 MC 加 FGS 的编码性能进行了仿真实验,并与 MPEG-4 的 FGS 编码方案进行了实验比较。实验的测试序列为 Foreman、Coastguard、Akiyo (CIF 格式)序列,采用 TM5 速率控制方式,帧率为 10 f/s。GOP(group of vop)结构为 $N_I=1, N_P=4, N_B=12$ 。设定基本层比特率 $R_{BL}=128$ kbits/s,基本层采用 MPEG-4 视频压缩编码方案。增强层的码率可以根据网络可用带宽进行任意码率的截取,本实验中分别截取增强层码率为 64 kbits/s、172 kbits/s、...直到 872 kbits/s。表 1 所列为原始 FGS 结构与单环 MC 加 FGS 结构中 3 个序列的第 5 帧随码流比特率变化的亮度分量 Y 的峰值信噪比,需指出的是,此峰值信噪比的值是图像质量最好时的峰值信噪比 PSNR(非重构出的基本层图像质量),即当增强层所有码流全部被接收,并被正确解码而得出的图像的峰值信噪比 PSNR。由表中的实验结果可知,本文方案重构图像的峰值信噪比 PSNR 比 MPEG-4 FGS 方案的 PSNR 有所提高,其中,对 Forman 可提高 1.23~5.83 dB,对 Coastguard 可提高 0.66~3.05 dB,对 Akiyo 可提高 4.43~9.64 dB。

图 4 为这 3 个视频序列的峰值信噪比 PSNR 随容许比特率的变化曲线。对于低速运动视频序列(Akiyo),由于其在时域上的相关性很强,因此当增强层去除了时域相关性后的图像质量,本方案比 MPEG-4 FGS 方案的图像质量有明显增强;而对于高速运动序列(Coastguard),由于其在时域上的相关性相对较弱,因此当去除了增强层的时域相关性

表1 不同码率情况下 Foreman、Coastguard、Akiyo 序列亮度分量 Y 的峰值信噪比 PSNR

单位: dB

比特率 (kbits/s)	Foreman		Coastguard		Akiyo	
	原始 FGS	MC 加 FGS	原始 FGS	MC 加 FGS	原始 FGS	MC 加 FGS
192	37.404 7	42.017 5	32.335 7	33.265 7	43.236 3	52.044 6
300	40.313 9	43.874 7	33.244 5	34.573 9	46.876 8	56.521 1
450	41.362 2	47.195 7	35.222 2	36.945 5	47.949 2	56.905 8
600	43.038 9	47.123 6	36.418 9	37.082 9	49.704 1	56.905 8
1000	46.391	47.616 2	38.837 2	41.886 3	52.540 5	56.905 8

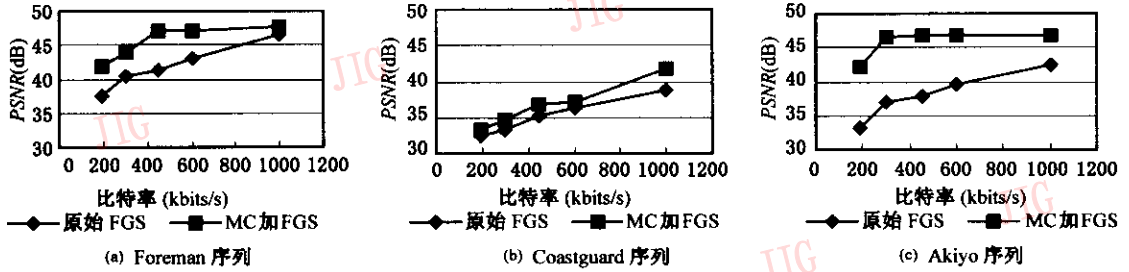


图4 原始 FGS 结构与 MC 加 FGS 结构的峰值信噪比 PSNR 比较

后,图像质量增强不是非常明显;而对于中速运动序列(Foreman),由于其在时域上的相关性介于低运动序列与高运动序列之间,因此当去除增强层时域相关性后,其图像质量增强的幅度也介于两者之间。

4 总结

本文提出了两种基于运动补偿的 FGS 改进结构,并分析比较了两种结构各自的优缺点,其中基于单环的 MC 加 FGS 结构具有实现简便、编码效率更高的特点,实验表明,对其缺陷进行改进后,其编码性能明显高于 MPEG-4 FGS 方案,其对于例如 over Internet 视频会议那样的需要进行实时视频编码的应用场合十分适用。

参考文献

- Gardos T. Efficient receiver-driven layered video multicast using H.263 + SNR scalability[A]. In: Proceedings of International Conference Image Processing ICIP'98 [C], Chicago, Illinois USA, 1998; 32~35.
- ISO/IEC. Coding of audio-visual object, part-2 visual amendment 2: Streaming video profile[S]. 14496-2/FDAM2, Pisa, Italy, 2001; 207~212.
- Li Wei-ping. Overview of fine granularity scalability in MPEG-4 video standard[J]. IEEE Transactions Circuit System Video Technology, 2001, 11(3):301~317.
- Ling F, Li W, Sun H. Bitplane coding of DCT coefficients for image and video compression[A]. In: Proceedings of SPIE Visual Communications and Image Processing[C], San Jose, CA, USA, 1999;25~27.
- Van M, Radha H. Unequal packet loss resilience for fine-

granular-scalability video[J]. IEEE Transactions on multimedia, 2001, 3(4): 381~394.

- Radha H, Van M, Chen Y. The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP [J]. IEEE Transactions on Multimedia, 2001, 3(1):53~67.



江涛 1979 年生,现为上海大学通信与信息系统专业博士研究生,硕士期间曾参与 MPEG-4 视频解码器 DSP 实现课题的研究。现主要从事网络流视频的研究,包括视频编码、传输控制及网络带宽估计算法等的研究。

E-mail:taojiangsh@163.com



张兆扬 1938 年生,教授、博士生导师,1962 年毕业于交大无线电工程系,长期从事数字视频处理与图像编码的研究,先后于 1995 年与 1997 年到日本大阪大学与美国加州大学讲学。完成科研项目 29 项,发表论文 85 篇、出版著作 8 部。其中获省部级以上奖 9 项。现主要研究领域为视频图像编码、网络多媒体通信等。

马然 1974 年生,讲师,在上海大学获硕士学位后留校从事科研与教学工作,现为上海大学通信与信息系统专业博士研究生。现主要从事流媒体网络传输中差错码的检测、控制、恢复等研究。

石旭利 1970 年生,讲师,2003 年获上海大学博士学位,并留校任教。现主要从事 MPEG-4 视频的帧内编码、传输及 H.264 等方面的研究。