# 基于宏块的渐进、精细可伸缩的视频编码*

孙晓艳[1], 高 文[1], 吴 枫[2], 李世鹏[2], 张亚勤[2]

[1](哈尔滨工业大学 计算机科学与工程系,黑龙江 哈尔滨 150001);
[2](微软亚洲研究院,北京 100080)
E-mail: xysun_cs@hotmail.com

摘要: 渐进精细可伸缩(PFGS)的视频编码是面向 Internet 的视频流化应用中的一项重要技术,同 MPEG-4 标准中的精细可伸缩(FGS)的视频编码相比,PFGS 视频编码具有更高的编码效率.然而,由于现有的 PFGS 方法以帧为单位来选择编码时的参考图像,因此该方法很难同时在编码效率和误差控制方面都取得好的效果.提出了一种灵活、高效的基于宏块的 PFGS 编码方法,并提出 3 种帧间(INTER)编码方式用于增强层宏块编码,其中的一个编码方式把原有的 PFGS 视频编码中的误差控制技术扩展到了宏块层.同时也提出了一个编码方式选择算法以确定每个增强层宏块的编码方式,由于该算法仅需用到时域预测信息,因此实现起来非常简单,并且性能稳定.实验结果表明,基于宏块的 PFGS 视频编码既能有效地消除低码率下的误差传递和累积,也能提高在其他码率下的编码效率.

关 键 词: 视频编码;精细可伸缩的视频编码;宏块的编码方式;运动补偿;位平面编码
中图法分类号: TP391 文献标识码: A

在 Internet 发展之初,由于传输带宽的限制,其应用主要是面向文本和图像.但是近 10 年来随着相关技术的迅猛发展,网络的访问速度越来越快,面向视频流和音频流的应用日益增多[1,2].然而,在互联网上传输视频流需要解决的一个基本问题是网络带宽的波动,即相同的终端不同的时间或者不同的终端相同时间访问互联网时所得到的传输带宽存在很大的差异.而传统的视频编码技术通常希望在给定的某个码率下获得最好的视频质量,因此其产生的码流很难适应网络带宽的波动,不能充分利用信道的容量.

随着视频流在互联网上应用的日益广泛,视频编码技术也在发生巨大的变化,其编码的目标从单纯地追求在某个码率下的编码效率发展为在一个较宽的码率范围内提高传输效率、优化视觉质量.例如,MPEG-4 标准中的 FGS 视频编码技术就具有这样的功能[3,4].在 FGS 编码方案中,基本层采用传统的编码技术,而基本层的量化误差采用位平面编码技术形成增强层.由于位平面编码技术提供了一个嵌入式码流和精细可伸缩的能力,从而使得 FGS 增强层的码流可以在任何码率下解码,因此 FGS 方法可以很容易地适应网络带宽的波动,并且提供了更高的传输效率.但是,由于 FGS 编码技术总是用最低质量的基本层作为运动估计的参考图像,所以其编码效率要远远低于传统的单层视频编码方法,甚至低于传统的 PSNR 多层编码方法[5].与传统的单层视频编码技术相比,在相同的码率下,FGS 的编码效率会降低 2.0dB~3.0dB[3].

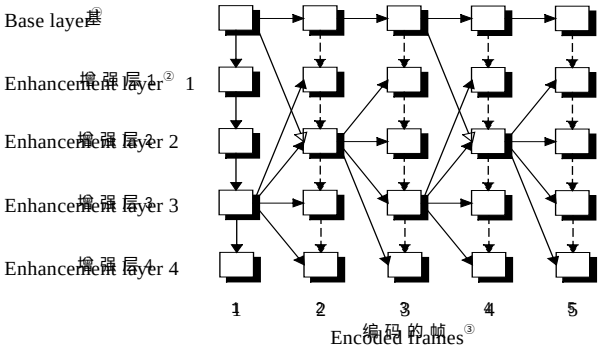
PFGS 编码技术是在 FGS 编码技术基础上的一个改进,它能提供更高的编码效率[6].PFGS 的编码结构如图 1 所示,其中带实箭头的实线表示运动补偿所用的参考图像,带空心箭头的实线表示从前一个基本层重构当前的高质量的参考图像,带实箭头的虚线表示同一帧的不同层之间的预测关系.由于在增强层的编码中使用了

* 收稿日期: 2001-02-15; 修改日期: 2001-05-15
基金项目: 国家自然科学基金资助项目(69789301);国家 863 高科技发展计划资助项目(863-306-ZT03-01-2)

作者简介: 孙晓艳(1974－),女,北京人,博士生,主要研究领域为视频编码,图像压缩;高文(1956－),男,辽宁大连人,博士,教授,博士生导师,主要研究领域为多媒体技术,人工智能,多模式接口,计算机视觉,虚拟现实;吴枫(1969－),男,湖北天门人,博士,主要研究领域为多媒体数据压缩;李世鹏(1967－),男,山东威海人,博士,主要研究领域为多媒体信息处理;张亚勤(1966－),男,山西太原人,博士,主要研究领域为多媒体信息处理和通信.

由部分增强层构成的高质量的参考图像,与 FGS 相比,在中、高码率下,PFGS 的编码效率可以提高 1.0dB.类似于 FGS 编码技术,PFGS 也生成两个码流:基本层码流和增强层码流.一般情况下,基本层视频流的码率总是被控制到与网络的最小传输带宽相适应,因此可以认为基本层在解码端总是可以得到的.然而,由于高质量的参考图像包含了一些低的增强层,因而需要更多的带宽来传输这些增强层.当网络带宽波动到不足以传输这部分增强层码流时,解码端将有可能部分或全部丢掉高质量的参考图像.这种编码器和解码器之间的高质量参考图像的差异必然会导致误差传递和累积,并且会严重地影响解码图像的视觉质量.

从图 1 可以看出,为了有效地控制这种误差传递和累积,现有的 PFGS 在编码过程中交替地使用前一个低质量的参考图像和前一个高质量的参考图像来重构当前高质量的参考图像.当使用前一个低质量的参考图像来重构当前的高质量的参考图像时,解码器和编码器总可以得到相同的时域预测,从而有效地抑制了误差传递.然而在用这种方法消除误差传递和累积的同时,由于高质量的参考图像不能获得它本应得到的最好的重构质量,PFGS 的编码效率也会受到影响.此外,由于现有的 PFGS 编码技术中预测参考和重构参考的选择是基于整帧图像的,因此要想在高编码效率和低误差累积之间取得一个较好的平衡也是比较困难的.




①基本层,②增强层,③编码的帧.

Fig.1 The architecture of PFGS scheme

图 1 PFGS 编码结构

本文将现有的 PFGS 编码方法的基本技术扩展到增强层中的宏块层,提出了 3 种编码方式以选择每个宏块编码时使用的参考图像.这种基于宏块的增强层编码技术能提供更大的灵活性,使得 PFGS 编码方法既能有效地消除在低码率下的误差传递和累积,又能进一步提高其编码效率.另外,本文也提出了一个编码方式选择算法,以优化地确定每个增强层宏块的编码方式.由于该算法仅需用到时域预测信息,因而实现起来非常简单,并且性能稳定.

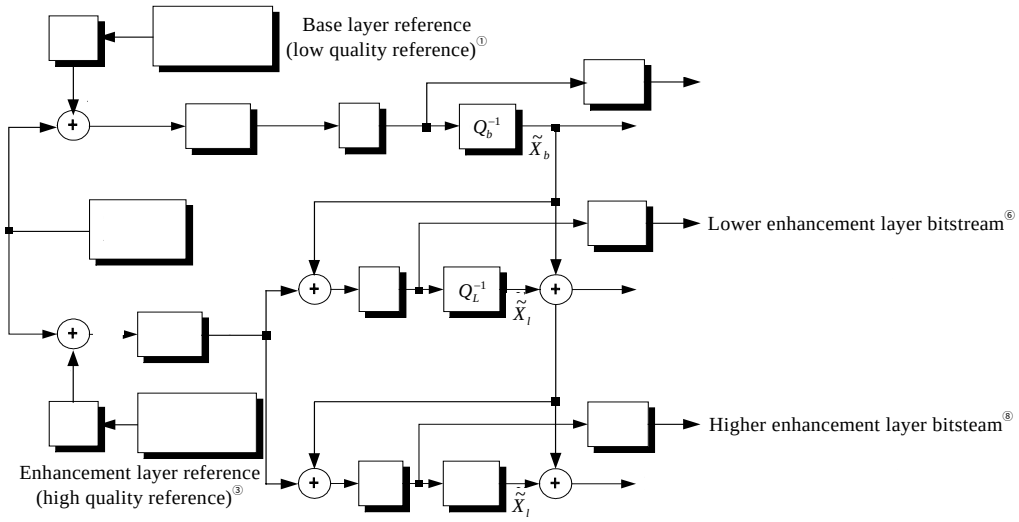
## 1 增强层宏块的编码方式

图 2 是一个 PFGS 编码器的示意图,为简单起见,在该编码器的框图中略去了运动估计模块,图中的两个运动补偿模块使用相同的运动向量.由于 PFGS 编码器有两个不同质量的参考图像,因而生成了两组预测 DCT 系数:使用前一个低质量的参考图像得到的预测 DCT 系数 $X_b$ 和使用前一个高质量参考图像得到的预测 DCT 系数 $X_e$,其中 $X_b$ 使用传统的变长编码技术生成基本层码流, $X_e$ 和重构的 $\tilde{X}_b$ 之间的差值经过位平面编码技术生成增强层码流.在所有的增强层中,只有较低的几个增强层用于重构高质量的参考图像.

在现有的 PFGS 编码方案中(如图 1 所示),增强层是在整帧图像上来选择运动补偿和重构时使用的参考图像的,也就是说,在每一帧图像的增强层编码过程中,或者全部宏块都使用低质量的参考图像,或者全部宏块都使用高质量的参考图像.事实上,由于在 PFGS 中有两个不同质量的参考图像,如果每个宏块都能选择一个最合适的参考图像来进行运动补偿和重构,PFGS 方法将能得到更大的灵活性和更好的综合性能,这也就是让增强层的每一个宏块在运动补偿和重构过程中都可以灵活地选择使用高质量的参考或低质量的参考.显然,这种基于宏块的思想能够更为灵活地在消除误差传递和提高编码效率之间寻求平衡.

根据宏块运动补偿和重构时所选用的参考图像的不同,本文提出了 3 个增强层宏块的帧间编码方式,即 LPLR 方式、HPHR 方式和 HPLR 方式.这 3 种编码方式如图 3 所示,灰色的方块表示这些层的重构图像将用作下一帧图像编码时的参考,带实箭头的实线表示运动补偿时所用的参考图像,带虚箭头的实线表示重构当前高质量的参考时所用的参考图像,带实箭头的虚线表示各层之间在 DCT 域的预测关系.下面我们将分别讨论这 3

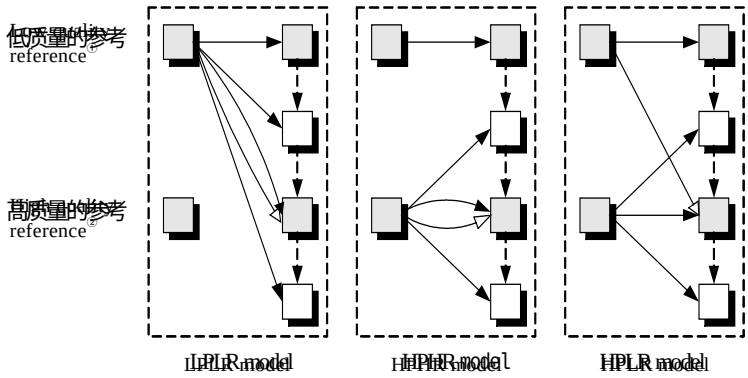
种编码方式的特点.




①基本层参考(低质量参考),②源图像,③增强层参考(高质量参考),④基本层码流,⑤重构的 $X_b$,
⑥低的增强层码流,⑦重构的低层的 $X_e$,⑧高的增强层码流,⑨重构的高层的 $X_e$.

Fig.2 The diagram of a PFGS encoder

图 2 PFGS 编码器示意图




①低质量的参考,②高质量的参考.

Fig.3 The coding modes for enhancement macroblock

图 3 强层宏块的编码方式

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.由于在可伸缩的视频编码方法中总是假设基本层码流能够被正确地传输到用户端,因而编码器和解码器总可以得到相同的低质量参考图像,因此这种编码方式不会产生误差传递和累积.如果所有的增强层宏块都使用这种方式进行编码,那么 PFGS 编码方法也就和 FGS 编码方法一样了.由于在运动补偿和重构中都使用了低质量的参考图像,LPLR 方式的编码效率是最低的.

在 HPHR 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,并使用相同的高质量参考图像进行重构.由于使用高质量参考图像使得运动补偿更为有效,因此 HPHR 编码方式的编码效率最高.如果所有的增强层宏块都采用这种方式进行编码,那么 PFGS 在高码率下将得到最高的编码效率.但是,这种编码方式存在一个严重的缺点:当网络带宽下降或者出现传输错误造成解码器无法得到前一个高质量的参考时,解码器将不得不用低质量的参考图像来代替出现错误的高质量参考图像来进行重构.在这种情况下,解码器重构的高质量参考图像 $\hat{r}_e(n)$ 为

$$\hat{r}_e(n) = p_b(n) + f^{-1}(\tilde{X}_b(n) + \tilde{X}_l(n)), \tag{1}$$

其中 $p_b(n)$ 表示低质量的预测图像, $\tilde{X}_b(n)$ 表示在基本层编码的 DCT 系数, $\tilde{X}_l(n)$ 表示在低的增强层中编码的 DCT 系数,公式 $f^{-1}(*)$ 表示逆 DCT 变换.然而,在编码器重构相同的高质量的参考图像时却采用了前一个高质量的参考图像,即 $r_e(n)$ 为

$$r_e(n) = p_e(n) + f^{-1}(\tilde{X}_b(n) + \tilde{X}_l(n)), \tag{2}$$

其中 $p_e(n)$ 表示高质量的预测图像.可见解码器在重构当前的高质量参考时产生了误差,这个误差可以用式(3)来表示.

$$e(n) = p_e(n) - p_b(n). \tag{3}$$

值得注意的是,这个差值 $e(n)$ 不仅影响当前帧解码图像的质量,如果在后面帧的增强层编码中还用这个高质量的参考图像作为预测,这个误差还将继续传递下去,影响到后面帧图像的解码质量,直到遇到下一个 $I$ 帧时才会结束.通常情况下,当网络带宽发生波动和出现错误时会有连续多帧图像的低增强层都发生错误,如果这些帧都采用 HPHR 方式编码,那么每一帧解码中都会产生这样的误差并传递下去,这样,后面帧的解码图像的质量会由于误差传递和累积迅速降低.

HPLR 是一个新颖的增强层宏块编码方式,用来消除这种误差的传递和累积.在这种编码方式下,增强层宏块在运动补偿过程中使用了高质量的参考图像,而在重构过程中使用的是低质量的参考图像.这样做的目的是为了确保编码器和解码器总能得到相同的时域预测(因为编码器和解码器总能得到相同的低质量的参考图像),从而有效地消除了误差 $e(n)$ 的积累和传递.HPLR 的显著特点就是在宏块的运动补偿和重构过程中使用了不同质量的参考图像.事实上,这种编码方式是把原来的 PFGS 编码技术中的误差控制方法扩展到了增强层宏块的编码之中.

由于在 HPLR 编码方式中使用了低质量的参考图像进行重构,使得当前帧的高质量的参考得不到它本应得到的最好的重构质量,因此这种编码方式也会影响到 PFGS 的编码效率.实际上,这种编码方式是把在 HPHR 编码方式中存在于编码器和解码器的高质量参考图像间的误差(定义见公式(3))转移为编码过程中重构的高质量参考图像的质量损失.在下一帧图像的运动补偿中,使用这个重构高质量参考图像得到的时域预测也会有同样的质量损失,从而使运动补偿后的预测残差增大,这样,在下一帧图像编码中将不得不使用更多的比特来编码这些预测残差.可见,HPLR 编码方式在有效地消除误差积累的同时也会对 PFGS 的编码效率产生负面影响.

在基于宏块的 PFGS 编码技术中不得不解决的另一个问题是,如何在编码过程中优化地选择每个增强层宏块的编码方式.这种编码控制的目的是在有效地限制 PFGS 编码方法在低码率下误差传递的同时,尽可能地提高它在高码率下的编码效率.

## 2 增强层宏块的编码方式选择算法

为了优化地确定每个增强层宏块的编码方式,我们提出了一个增强层宏块编码方式选择算法.在传统的编码方法中,宏块的编码方式可分为两类:帧内(INTRA)编码和帧间(INTER)编码.在 PFGS 的基本层编码中,一个宏块是采用帧间编码还是采用帧内编码是通过运动估计模块来决定的.而在 PFGS 的增强层编码中,如果一个宏块在基本层编码时使用帧内编码方式,那么在增强层中相同位置的那个宏块在编码时也采用帧内模式,无须任何时域预测;否则,在增强层中相同位置的那个宏块则采用帧间编码方式.根据运动补偿和重构参考图像的不同,上一节中定义了增强层宏块编码的 3 种方式,下面我们将详细讨论如何优化地选择每个增强层宏块的编码方式.

从 3 个帧间编码方式的定义可知,采用 LPLR 编码方式的增强层宏块在帧间运动补偿时使用的是低质量的参考图像,而采用 HPHR 或 HPLR 编码方式的增强层宏块在帧间运动补偿时则使用高质量的参考图像.因此,在编码方式选择算法中第 1 步是首先决定某个宏块是否采用 LPLR 编码方式.公式(4)提供了一个简单的准则来区分 LPLR 和另外两个帧间编码方式

$$\min\left(\left\|X_b - \tilde{X}_b\right\|, \left\|X_e - \tilde{X}_b\right\|\right). \tag{4}$$

换句话说,当一个增强层宏块使用低质量参考图像进行预测时,在增强层中编码的 DCT 残差为 $X_b - \tilde{X}_b$,而当一个增强层宏块使用高质量参考图像进行预测时,在增强层中编码的 DCT 残差为 $X_e - \tilde{X}_b$.我们知道,DCT 残差越小,编码时所需字节数就越少.假设使用低质量参考图像预测时得到 DCT 残差的绝对值的均值为 $\left\|X_b - \tilde{X}_b\right\|$,使用高质量参考图像预测时得到的 DCT 残差的绝对值的均值为 $\left\|X_e - \tilde{X}_b\right\|$.显然,当 $\left\|X_b - \tilde{X}_b\right\|$ 小于

$\left\| X_e - \tilde{X}_b \right\|$时,该增强层宏块应采用 LPLR 编码方式;否则,这个宏块应该在运动补偿过程中采用高质量的参考图像.选择算法将在 HPHR 和 HPLR 之间进一步确定该宏块的编码方式.

在 HPHR 和 HPLR 编码方式中,增强层宏块在运动补偿中都使用了高质量的参考图像,所不同的是它们在重构过程中分别使用了高质量的参考和低质量的参考.由于使用 HPHR 可以获得较高的编码效率,因此在增强层宏块的编码中应尽可能地多采用这种编码方式.然而,过多地使用这种编码方式也会在低码率的情况下造成严重的误差传递和累积.因此,编码方式选择算法必须估计在最坏情况下可能的误差传递和累积的影响,从而对其进行控制.如果当某个增强层宏块使用 HPHR 编码方式可能造成的误差大于一个给定的阈值时,这个宏块就采用 HPLR 编码方式,否则采用 HPHR 编码方式.

在 PFGS 编码方案中造成误差传递的原因是,当增强层宏块采用 HPHR 方式进行编码时,用于重构高质量参考的那些低的增强层码流没有被正确地传到解码端,从而造成编码器和解码器重构的高质量参考图像不一致,导致了当前解码图像出现质量损失,并且由于运动补偿的作用,编码器和解码器之间这种重构的高质量参考图像的不匹配还会造成误差的传递和累积,进而影响同一个图像组(group of picture,简称 GOP)中后续帧的解码图像质量.可见,为了精确地估计可能出现的误差传递和累积的影响,编码器需要知道在重构过程中使用低质量的参考代替高质量的参考时所造成的误差以及这种误差的传递所产生的累积效应,为此编码器将不得不额外增加一个运动补偿模块,这将增加 PFGS 编码方法的复杂度.在实际应用中,精确地计算出可能出现的误差及其累积效应是没有必要的.本文提出的编码方式选择算法中采用了一个简单的误差估计方法以便有效地控制 PFGS 在低码率下可能出现的误差传递和累积.

根据公式(3),编码方式选择算法利用前一个低质量参考图像和前一个高质量参考图像之间的差值来估计可能出现的误差传递.显然,这两个时域预测之间的差值越大,编码时用低质量参考图像来代替丢失的高质量参考图像所导致的质量损失也就越多.编码方式选择算法须把这个差值控制到一定的范围,因此定义了如下准则:

$$\left\| p_e(n) - p_b(n) \right\| > k \times \left\| x_o - r_e(n) \right\|, \tag{5}$$

其中 $x_o$ 是当前的原始图像,$k$ 是可接受的损失因子.这里,$k$ 是一个可调节的参数,它控制着 PFGS 编码方法在低码率和高码率下的综合性能.$\left\| x_o - r_e(n) \right\|$是重构高质量参考图像的均方误差(MSE).当两个时域预测的差值大于公式(5)中右边的值时,这个宏块就使用 HPLR 编码方式,因为如果在这种情况下宏块依然采用 HPHR 编码方式将有可能造成严重的误差传递.然而值得注意的是,宏块编码方式应该在每个宏块编码前就已经决定了,因此在确定编码方式时,公式(5)中的 $r_e(n)$ 往往无法得到.所以我们使用公式(6)来近似公式(5),这样 HPHR 和 HPLR 两种编码方式的选择准则为

$$\left\| p_e(n) - p_b(n) \right\| > k' \times \left\| x_o - p_e(n) \right\|, \tag{6}$$

其中 $p_e(n)$ 代替了 $r_e(n)$ .由于这两个时域预测图像在当前帧图像开始编码时就已经得到了,因此我们提出的编码方式选择算法无须引入任何附加的计算量.

显然,本文提出的编码方式选择算法实现起来非常简单,并且宏块编码方式的信息非常容易编码.每个增强层宏块的编码方式包含在增强层码流的宏块头中,由于在基本层码流中已经提供了宏块的 INTRA 编码方式的信息,在增强层码流中只需描述每个宏块的 INTER 编码方式,采用一个简单的 VLC 表即可.

## 3 试验结果和分析

下面我们将通过实验来验证基于宏块的 PFGS 编码技术的性能.在这个实验中共采用了 4 种不同的编码器:基于宏块的 PFGS 编码器、原有的基于帧的 PFGS 编码器、MPFG-4 标准中的 FGS 编码器以及 MPEG-4 的传统的单层编码器.

所有的实验条件都是根据 MPFGE-4 核心实验[7]设计的.在实验中使用了 4 个 MPEG 标准的视频测试序列:图像格式为 CIF 的 Foreman 和 Coastguard 以及图像格式为 QCIF 的 Foreman 和 Carphone.在编码每个视频序列时,只有第 1 帧采用帧内编码($I$ 帧),其余所有的帧都采用预测方式进行帧间编码($P$ 帧).运动估计是在前后两帧

源图像上进行的,估计得到的运动矢量用于该帧的基本层和增强层运动补偿.编码的帧率为 10Hz.

对于 CIF 格式的视频序列,运动估计过程中最大的运动矢量为±31.5 个像素,基本层的码率是 128kb/s,通过 TM5 码率控制方法来实现.增强层在编码过程中没有任何码率限制,生成的码流可以任意截断,在本文的实验中增强层的码流被分别截成 128kb/s,256kb/s,…,直到 1024k/s,间隔为 128kb/s.在编码 CIF 格式的视频序列时,PFGS 编码方法把每帧编码生成的比特数首次大于 20 000 时的那个增强层重构为高质量的参考图像,可接受的损失因子为 1.6.

对于 QCIF 格式的视频序列,运动估计过程中最大的运动矢量为±15.5 个像素,基本层的码率是 32kb/s,通过 TM5 码率控制方法来实现,增强层的码流被分别截成 16kb/s,32kb/s,…,直到 160kb/s,间隔为 16kb/s.在编码 QCIF 格式的视频序列时,PFGS 编码方法把每帧编码生成的比特数首次大于 5 000 时的那个增强层重构为高质量的参考图像,可接受的损失因子为 2.3.

实验结果如图 4~图 6 所示.图 4 中显示了 4 种编码方法分别对 Foreman(CIF/QCIF)序列进行编解码时在各个码率下得到的平均 PSNR.图 5 显示了 Carphone(QCIF)和 Coastguard(CIF)序列用基于宏块的 PFGS,FGS 和传统的单层编码方法在各个码率下的平均 PSNR.可以看出,与传统的单层的编码方法相比,FGS 编码方法的编码效率降低了 2.0dB~3.0dB.同时实验结果也显示,尽管使用了高质量的参考图像,通过采用本文提出的 3 个 INTER 编码方式和对应的编码方式选择算法,基于宏块的 PFGS 编码方法仍然在低码率下取得了与 FGS 编码方法基本相同的编码效率.而在中或高码率时,基于宏块的 PFGS 编码方法的编码效率比 FGS 方法和基于帧的 PFGS 方法分别高出 1.5dB 和 1.1dB.特别是,在中码率或高码率时基于宏块的 PFGS 方法得到的 PSNR 曲线基本上位于 FGS 和传统的单层视频编码方法之间,缩小了非伸缩性的单层视频编码方法和可伸缩视频编码方法之间编码效率上的差距.此外,在 CIF 和 QCIF 格式的视频序列编码过程中,可损失因子的值是固定的,图 4 的结果也验证了本文提出的编码方式选择算法对于不同格式、不同特点的视频序列都能够提供良好的控制.

图 6 显示了在相同的帧率下,基于宏块的 PFGS 和 FGS 两种编码方法得到的视频图像的 PSNR 值.在左图中显示的是 Carphone(QCIF 格式)编码码流被截成 48kb/s 和 176kb/s 时基于宏块的 PFGS 和 FGS 解码所得到的各帧的图像质量;右图为 Foreman(CIF 格式)编码码流被截成 256kb/s 和 768kb/s 时基于宏块的 PFGS 和 FGS 解码所得到的各帧的图像质量.在低码率时,基于宏块的 PFGS 和 FGS 两种编码方法得到的解码图像的 PSNR 值都很接近,看不到明显地由于误差传递和累积所造成的质量损失,可见在基于宏块的 PFGS 编码方法中低码率下的误差传递和积累得到了良好的控制.在中、高码率时,基于宏块的 PFGS 编码方法的 PSNR 值明显高于 FGS 编码方法.

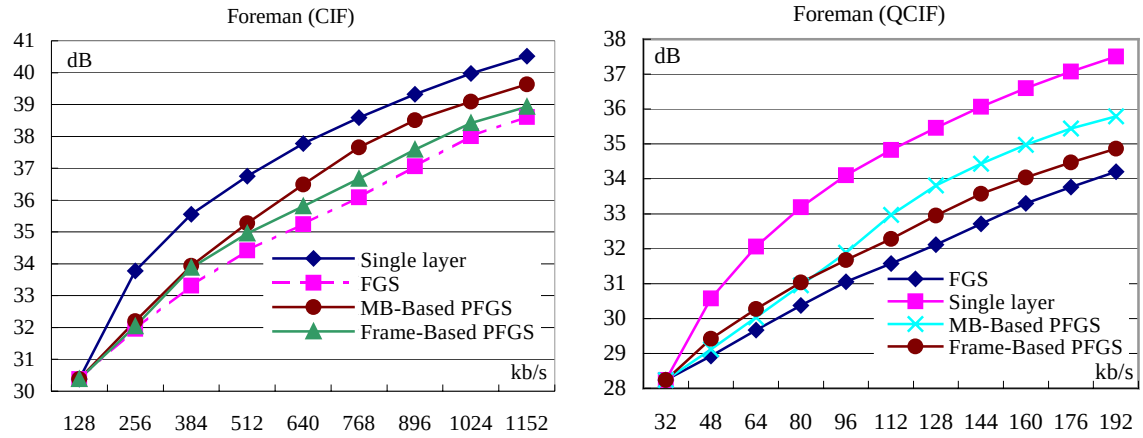



Fig.4 Comparison of coding efficiency among MB-based PFGS, frame-based PFGS, FGS and non-scalable coding

图 4 基于宏块的 PFGS、基于帧的 PFGS、FGS 和非可伸缩的编码方法的编码性能比较

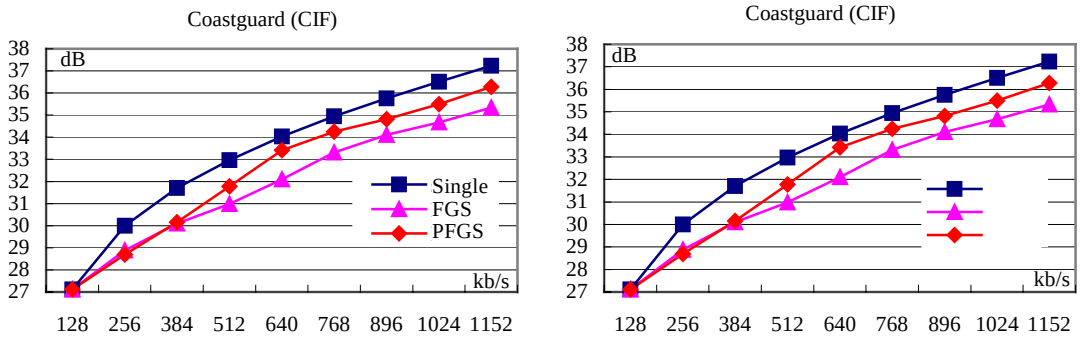



Fig.5 Comparison of coding efficiency among MB-based PFGS, FGS and non-scalable coding
图 5 基于宏块的 PFGS、FGS 和非可伸缩的视频编码方法的编码性能比较

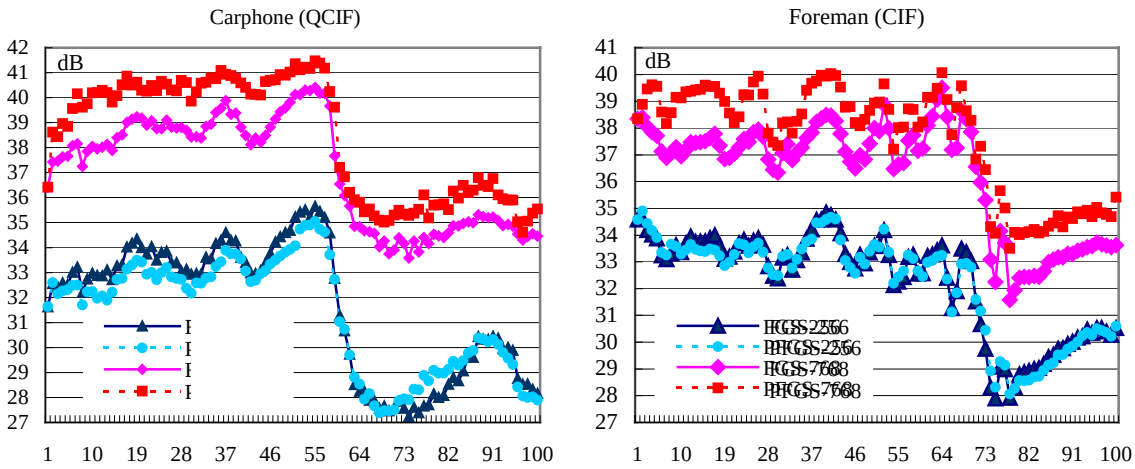



Fig.6 Comparison of PSNR between MB-based PFGS and FGS at the same bit rate
图 6 相同帧率下基于宏块的 PFGS 和 FGS 单帧 PSNR 结果比较

## 4 结 论

在本文中,我们提出了一个灵活、有效的基于宏块的 PFGS 视频编码方法.在该编码方法中,根据运动补偿和重构时选用的参考图像的不同,定义了 3 种增强层宏块的 INTER 编码方式.为了在提高编码效率和消除误差传递之间获得良好的平衡,我们也提出了一个增强层宏块选择算法,以确定每个增强层宏块的编码方式.增强层宏块编码方式和对应的选择算法给 PFGS 方案带来了很大的灵活性,使其能够在确保编码效率的同时有效地控制误差传递和积累.与 MPEG-4 的 FGS 编码技术和原有的基于帧的 PFGS 编码技术相比,本文的方法能够在低码率下基本保持原有的编码性能.而在中、高码率时,基于宏块的 PFGS 编码方法的编码效率比原有的 PFGS 编码方法高出 1.1dB,比 FGS 编码方法高出 1.5dB.基于宏块的 PFGS 编码方案进一步缩短了可伸缩视频编码方法和单层视频编码方案之间编码效率上的差距.

**References**:

[1] Luthra, A. Need for simple streaming video profile. ISO/IEC JTC1/SC29/WG11, MPEG doc M5800, Noordwijkerhout, 2000.

[2] Lu, Jian. Signal processing for Internet video streaming: a review. In: Vasudev, B., Hsing, R.T., Tescher, A.G., *et al*., eds. Image and Video Communication and Processing 2000. SPIE Vol. 3974, 2000. 246~258. http://spie.org/scripts/abstract.pl?bibcode=2000SPIE.3974..246L&db_key.

[3] Li, Wei-ping. Fine granularity scalability in MPEG-4 for streaming video. In: Circuits and Systems, Proceedings of the ISCAS 2000 IEEE International Symposium, Vol 1. 299~302. http://ieeexplore.ieee.org/Xplore/RecentCon.jsp?Punumber=6910.

[4] Li, Wei-ping. Streaming video profile in MPEG-4. IEEE Transactions on Circuit and Systems for Video Technology, 2001,11(3): 301~317.

[5] Arnold, J., Frater, M., Wang, Y. Efficient drift-free signal-to-noise ratio scalability. IEEE Transactions on Circuit and Systems for Video Technology, 2000,10(1):70~82.

[6] Wu, Feng, Li, Shi-peng, Zhang, Ya-qin. A framework for efficient progressive fine granularity scalable video coding. IEEE Transactions on Circuit and Systems for Video Technology, 2001,11(1):332~344.

[7] ISO/IEC JTC1/SC29/WG11. FGS core experiments. MPEG doc M3316. Noordwijkerhout, 2000.

# Macroblock-Based Progressive Fine Granularity Scalable Video Coding*

SUN Xiao-yan[1], GAO Wen[1], WU Feng[2], LI Shi-peng[2], ZHANG Ya-qin[2]

[1](*Department of Computer Science and Technology, Harbin Institute of Technology, Harbin* 150001, *China*);

[2](*Microsoft Research Asia, Beijing* 100080, *China*)

E-mail: xysun_cs@hotmail.com

**Abstract:** The progressive fine granularity scalable (PFGS) coding is a promising technique for streaming video applications due to its much higher coding efficiency than the FGS (Fine Granularity Scalable) coding in MPEG-4 standard. However, since the original PFGS scheme only chooses its reconstructed references as frame-based, it is very difficult to achieve a good trade-off between high coding efficiency and low drifting errors. In this paper, a flexible and effective scheme is presented to control the PFGS coding at macroblock level. Three INTER modes with distinct characteristics are first proposed for the enhancement macroblock coding. One coding mode among them provides a novel method to effectively reduce the drifting errors that may occur at low bit rates. Moreover, a decision-making mechanism based on temporal predictions is developed to choose the optimal coding mode for each enhancement macroblock, which offers a consistent improved performance for sequences with different contents and formats. Furthermore, the proposed control scheme can be easily implemented without any additional computations. The experimental results show that the macroblock-based enhancement layer coding technique using the three INTER modes and the proposed decision-making mechanism can effectively reduce the drifting errors at low bit rates, while providing high coding efficiency at moderate or high bit rates.

**Key words:** video coding; granularity scalable video coding; macroblock coding mode; motion compensation; bit plane coding

* Received February 15, 2001; accepted May 15, 2001

Supported by the National Natural Science Foundation of China under Grant No.69789301; the National High Technology Development 863 Program of China under Grant No.863-306-ZT03-01-2