

ViCowel :一种支持视频组播应用的可伸缩性小波视频编码方法^{*}

张 军 柳 伟 王 晖 李国辉

(国防科技大学人文与管理学院 湖南 长沙 410073)

摘 要 :WebLearning+ 是作者开发的基于 IP 组播的协同学习环境,可伸缩性视频编解码及其自适应组播传输是系统研制中的核心技术。针对 WebLearning+ 中自适应视频组播对编解码器的要求,提出了一种基于行扫描的可伸缩性小波视频编码方法 ViCowel。实验结果表明该方法对于视频组播应用具有较好的适应性。

关键词 :视频组播;可伸缩视频编码;ViCowel

中图分类号 :TP391 文献标识码 :A

ViCowel : A Scalable Wavelet Based Video Coding Approach for Video Multicast Applications

ZHANG Jun , LIU Wei , WANG Hui , LI Guo-hui

(College of Humanities and Management , National Univ. of Defense Technology , Changsha 410073 ,China)

Abstract :WebLearning-plus is a multicast-based collaborative learning environment we developed , in which scalable video coding and adaptive multicast transmission are the key issues . According to the demand of adaptive video multicast in WebLearning-plus for the coder , we present a new scalable wavelet based video coding approach named ViCowel , which is a kind of set partition based on line scanning . Experiments show that it has better flexibility in video multicast applications .

Key words :video multicast ;scalable video coding ;ViCowel

WebLearning+ 是作者开发的基于 IP 组播的协同学习环境^[1],该系统采用了 MPEG-2 视频组播技术实现教师视频到多个学生终端的实时递送。

针对学生终端处理能力(CPU、内存与显示等)与传输链路带宽的异构性问题,系统采用了一种基于视频网关的自适应层次组播体系^[2],将 MPEG-2 视频序列中的 I 帧作为基层,P 帧与 B 帧为增强层,采用类似于接收者驱动的层次组播方法 RLM^[3]实现自适应的 MPEG-2 视频组播。但基于 MPEG-2 的视频层次编码具有先天的码率非伸缩性,很难满足网络教育异构环境下接收者视频接收质量精细控制的要求。

1 层次化视频组播体系

自适应视频组播技术是近年来多媒体网络领域研究的热点问题。为解决网络与接收者终端的异构性问题,视频组播自适应机制可分为发送者驱动、接收者驱动和基于网关或代理的层次化驱动三种基本模式^[4]。针对大规模(接收节点一般在 100 个以上)的视频组播应用来说,基于网关的层次化视频组播方法在扩展性、公平性以及快速收敛性方面具有较好的性能。WebLearning+ 就是采用该方法实现自适应 MPEG-2 视频组播体系,其原理如图 1 所示。

该体系采用层次编码、接收方驱动的层次组播以及自适应视频过滤的视频网关相结合的方法解决异构性问题。即将视频流中的 I 帧作为基层,P 帧、B 帧作为增强层,并采用层次映射自适应算法^[5]解决多视频源组播会话的共享带宽分配和速率自适应问题;同时在广域网边缘设置应用视频网关,为广域网上加入同一实时课堂的教师或学生端视频发送源(S_1, S_2, S_3)建立相应的视频过滤代理(F_1, F_2, F_3),多个视频过滤代理共享输出的会话带宽,加入课堂的学生(R_1, R_2, R_3, R_4)可以根据自己的兴趣选择观

^{*} 收稿日期 2004-04-05
作者简介 张军(1975—),女,讲师,博士生。

看某个教师或学生的视频流。但是,基于 MPEG-2 的层次组播接收者必须接收基层,且要么全部接收增强层,要么全部不接收,码率控制颗粒度较大,很难得到与接收者接收能力和到达网络链路带宽相匹配的视频接收质量。解决这一问题的有效方法就是采用可伸缩视频编码^[6,7],基于小波的视频编码是一种能够保证视频质量,又具有很高压缩率的编码方法。同时由于实用的小波编码属于嵌入式编码,允许任意长度的码流截断,为自适应视频组播提供了保障。为此,我们提出了一个基于小波的可伸缩性视频编码方法 ViCowel。

2 视频编解码 ViCowel

2.1 小波提升

小波变换的基本思想是利用信号间存在的相关性,建立一种信号的稀疏表示。经典的小波变换通常是通过傅立叶分析对其进行时频分析。在信号的空间域,对信号实现分裂、预测、更新,并进行频率分解,称为信号的小波提升^[8]。

(1) 分裂

对于整数集 Z 和实数集 R ,考虑信号 $X = X_k(k \in Z), X_k \in R$ 。首先对信号进行分裂,得两子集:

$$X = X_1 Y X_2, \quad X_1 \perp X_2 = \phi \quad (1)$$

可知,两子集是高度相关的。这样可以通过子集 X_2 来预测 X_1 。设这样的预测算子为 p ,而预测结果与原 X_1 的差异即 X_1 的细节部分或高频部分记为 d 。

(2) 预测

由上述可知, $d = X_1 - p(X_2)$ 。给定细节部分 d 及信号的一部分 X_2 ,可恢复信号的另一部分 $X_1 = p(X_2) + d$ 。由于预测算子的选择也许不能使预测得来的 X_1 和原信号具有相同的整体性质,因而可考虑对细节 d 进行修正,设更新算子为 U 。

(3) 更新

$$S = X_1 + U(d) \quad (2)$$

则更新后

$$X_1 = S - U(d) \quad (3)$$

至此完成了信号的一次提升,相当于小波一层分解。

对低频部分 X_1 重复实现新的提升,进行公式(1)(2)和(3)的计算,即完成了信号的多分辨率分解。重构时,将低频分量和各高频分量进行合并,数据的重构公式为

$$X_1 = S - U(d), \quad d = X_1 - p(X_2) \quad (4)$$

$$X = Merge(X_1, d) \quad (5)$$

用提升方法来实现小波变换的一个最大优点是:将小波滤波器过程分解成极其简单的基本步骤,且分解的每一步都是可逆的。提升方法的重构过程是分解过程的逆过程。

小波提升分解的优点也就是把小波基系数用另一些 2 幂次的倍数来代替,所以小波变换可以通过系数的移位与加法实现,这样变换速度与效率都有较大的提高,减少了能量损失。

对于(9,3)小波,可以通过提升算法将它简化为

$$P = (1, 1)/2 \quad (6)$$

$$U = (-3, 19, 1, 9, -3)/64 \quad (7)$$

2.2 空间域上基于行扫描的集合分割编码

对小波变换后的系数进行处理时,结构化编码是一种性能比较优良的编码方法。它可以充分利用

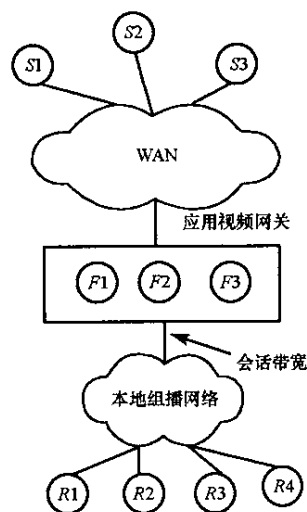


图1 WebLearning+ 的层次化视频组播原理图

Fig.1 Layered video multicast scenario in WebLearning-plus

小波系数的能量分布特性,不仅能够获得较好的压缩效果,还具备良好的可伸缩性。我们描述的编码算法其基本原理是一种集合划分算法。该算法在 SPECK^[9]的基础上加以改进,采用了与 SPECK 相同的基于块的划分策略。与 SPECK 的区别在于引入了基本块的概念。对于非正方形图像,系数矩阵为 $M \times N$ ($M > N$), 设其 n 次小波迭代后的 LL_n 系数矩阵为 $M_n \times N_n$ 。按照基于块的集合划分算法,对 LL_n 进行分块。划分到子块的大小为基本块时,停止划分。设基本块尺寸为 $m' \times n'$, 对系数进行行扫描,根据行内系数的重要性,对其输出比特进行霍夫曼编码。对于非矩形系数矩阵,分块算法无法到达单个系数,就像 SPECK 算法,这时只能对块内的系数进行输出。而且,相对于整个块,块内的系数行在重要性分布上更规则,出现全 0 和全 1 的概率很大。除了全 0 和全 1,其他的分布也很集中。

对于 320×240 YUV12 格式的一帧图像,经过 5 次迭代,其 LL_5 为 10×8 , 经过一次分块,得到的子块为 5×4 , 将这个大小的块定义为基本块。对块内的 4 行系数,在每个位平面其重要性输出有 32 种组合。表 1 是随机抽取的几帧图像的系数分布。

表 1 系数分布表
Tab. 1 Significant map sample

00000	572	600	640	637	655
10000	175	204	199	189	175
01000	140	133	146	152	125
00100	115	101	118	119	128
00010	113	88	113	115	110
00001	116	109	148	152	136

另外 26 种组合出现的概率很低,一般在 0~16 之间,使用霍夫曼编码,在不降低 PSNR 的前提下,至少可以减少 4000 bit 的数据量。

在我们提出的算法中,对搜索策略作了一些改进,算法利用了小波系数的能量分布特点,用预测来代替部分搜索。使用集合划分编码算法,对整幅图像作编码能取得较高的压缩率。在这类算法中,对重要集合和非重要集合的判别需要的运算量很大,通常要对整块的系数矩阵进行搜索、比较。尤其是非重要集合,需要对剩下的所有系数进行判断。对于多次迭代的小波,其能量分布是比较规律的, LL_1 通常占图像能量分布的 90% 以上。在对非重要集合进行判别时,取最小的系数集进行判断。其中,设待编码的图像系数集合为 X , 将 X 分割为两个集合: $S = ROOT^{[10]}$, $I = X - S$ 。经测试发现,对越高的位平面,搜索集合越紧凑,误差也越小,对于效率的提高也越显著,因为这时 S 集合很小,而 I 集合比较大。随着位平面的降低,搜索集合逐渐放大,误差也随之增大,这些误差主要集中在 LL_2 以上,只占图像很少的信息,主要是一些细节信息,误差对图像恢复的影响并不明显。从总体来讲,我们的搜索策略可以提高 15% 的效率,当然图像恢复质量也有所下降,但在连续视频的条件下,视觉质量的降低不明显。

2.3 基于时域的帧间差分预测

本文所研究的算法是针对远程教学系统的应用背景,需要考虑到远程教学视频的特征。在系统中,摄像头安置在固定的位置,那么它所拍摄到的空间多数情况下是固定的,也就是说背景基本是一致的或者背景的变化比较缓慢。即使有运动像块存在,每个像块的幅值各不相同,但像块内各样值的幅度是相近的或相同的,幅值跃变部分对应于像块的轮廓,只占整幅图像的很小一部分。相邻帧间的背景位置的像素几乎完全一样,这意味着前后像素之差或前后帧间相应位置像素之差为零或差值小的概率大,差值大的概率小。但是差值编码的缺点就是只用到了与前一帧的相关性,所以差值相对就比较大,相关性比较弱。而预测编码利用像素的相关性,可进一步减小差值。

从前面的分析可以看出,如果差值编码中小幅度出现的机会增加,由于其对应的码长较短,总数码率会进一步减小。如果能预测出下一个样值,那么差值就会是零,当然这种情况是没有意义的。但可以肯定,如果不仅利用前后样值的相关性,同时也利用其它行和相邻帧的像素的相关性,用更接近当前样值的预测值与当前样值相减,小幅度差值就会增加,总数码率就会减小,这就是预测编码的思想。

在本文提出的这个算法中,用到了预测编码的思想,采用二维预测,即预测值等于 $1/2$ 前一帧像素加 $1/4$ 上一行相应像素,再加上 $1/4$ 上一行相应的前一帧像素。这样不仅利用了前一像素的相关性,也利用了上一行相应像素的相关性,这样做要比差值编码有更大的码率压缩。如果再用上前一帧的像素,会进一步降低码率。但为了得到前一帧的像素,必须使用帧存储器,这样会占用缓存。

把原始帧(第一帧)进行小波变换,将第二帧与预测帧的值进行差分,将差分再进行小波变换,如果令 t_0 表示需要预测的像素点值, t_1 表示前一帧的相应位置的像素点, t_2 表示上一行相应像素值, t_3 表示上一行相应的前一帧像素值,则差分 d 表示为:

$$d = t_0 - \left(\frac{1}{2} t_1 + \frac{1}{4} t_2 + \frac{1}{4} t_3 \right) \quad (8)$$

$$d = t_0 - t_1 \quad (\text{如果 } t_0 \text{ 是第一行的像素点}) \quad (9)$$

第一行的像素点只与前一帧的相应的像素点做差值。经过分析,发现 d 的值都比较小,对 d 进行小波变换之后,会发现 0 的概率很大,那么编码将会很容易。这样,就只需将 1 帧进行小波变换,编码方法用 2.2 节提出的算法。后面的帧以此方法继续,这样,只需用前一帧对后面一帧进行预测。在解码端,只需将第一帧解码,差分解码,然后用同样的公式计算出后续帧的像素值。为了减少误差,可以在几帧之后传送一个完整的 I 帧作为参考帧。这种方法简单,不需要额外的缓存,而且编码效率高,速度快,对于远程教学系统,可以满足视频质量的要求,并且实时性好。编解码算法流程图见图 2。

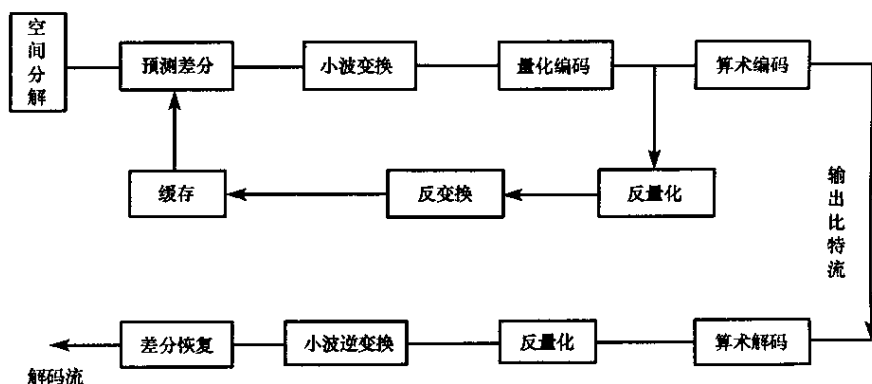


图 2 算法流程图

Fig.2 Algorithm flow chart

3 测试结果

本文在 WebLearning+ 系统的研究中,从变换和编码两个方面入手,综合信噪比、压缩率、算法复杂性三个指标,采用了多种能够提高性能的手段,包括利用(9.3)小波的提升,使得算法复杂度降低;使用 Intel 的单指令多数据技术,从算法程序实现上进行优化。提出的空间域上基于行扫描的集合分割编码方法和时间域上帧间差分预测算法,在牺牲信噪比的较小代价下,能够实现较高的压缩比,大大降低了算法复杂性。通过与 Open-Divx 的 MPEG-4 编码器作对比实验,得到如下对比结果,见表 2。本算法与 3D-SPIHT^[11] 视频编码方法进行对比实验,得出了 PSNR 对比结果见图 3,从图中可以发现 ViCowel 在同一码率下有着较高的信噪比。3D-SPIHT 虽然是很有价值的编码思想,但是在实际应用中,因其计算复杂度较高,且需要缓存比较多的视频帧,难以实现视频的实时处理。如果缓存的视频帧比较少,那么时间域的小波变换后能量集中不明显,编码效率相对就会较低。ViCowel 的算法复杂度比 3D-SPIHT 低,并且不需要缓存过多的帧,能满足实时处理视频的要求。

表2 对比结果

Tab.2 Comparison result

编码方法(相同 PSNR)	YUV320 压缩后大小	每秒压缩帧数	压缩倍数
Open-Divx	4.2KB	5~7 帧	0.2bpp(40 倍)
ViCowel	2.5KB	≥15 帧	0.1bpp(80 倍)

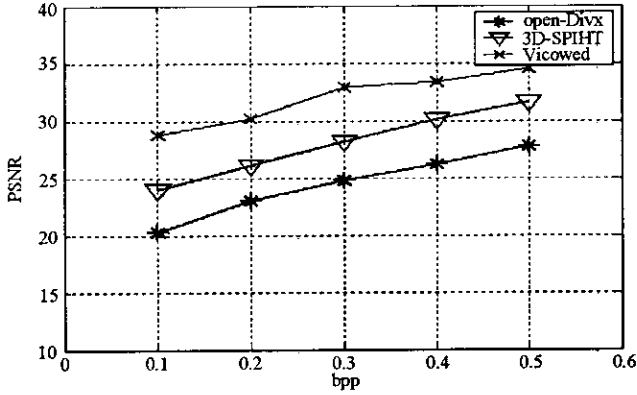


图3 编码方法 PSNR 对比图(320×240)

Fig.3 Compared chart in PSNR among diverse coding method(320×240)

4 结束语

针对自适应视频组播应用提出了一种精细比特控制、低复杂性的实时视频编码方法 ViCowel 及其算法。空间域上的编码算法是在 SPECK 的基础上,提出了一种基于行扫描的集合分割编码算法。时间域上提出了一种差分预测视频编码方法,与以往差分编码方法不同的是用到了预测编码的相关理论,不仅考虑了帧间像素点的相关性,还考虑了同一帧相邻像素点之间的相关性。在传输过程中,提出了一种对比特流进行伸缩性控制的方法,在分层的比特流中加识别码,对于带宽的变化具有很好的自适应性。实验结果表明,本文的算法具有较好的压缩率和信噪比,计算复杂度低,能满足远程教学系统中视频的实时压缩与传输要求。

参考文献:

- [1] 王晖,姜志宏,张军,等. WebLearning+: 一个基于 IP 组播的协同学习环境[J]. 计算机工程, 2002, (3).
- [2] 王晖,陈文伟,等. 多视频源组播会话中编码层次映射自适应模型与算法研究[J]. 小型微型计算机系统, 2003, 9.
- [3] McCanne S, Jacobson V, Vetterli M. Receiver-driven Layered Multicast[C]. In Proceedings of ACM SIGCOMM '96, Stanford, CA, 1996: 27.
- [4] Li X, Ammar M, Paul S. Video Multicast over the Internet[J]. IEEE Network Magazine, 1999, 13(2): 46-60.
- [5] Amir E, McCanne S, Katz R. Receiver-driven Bandwidth Adaptation for Light-weight Session[C]. In Proceedings of ACM Multimedia '97, 1997.
- [6] Li S, Wu F, Zhang Y Q. Study of a New Approach to Improve FGS Video Coding Efficiency[S]. ISO/IEC JTC1/SC29/WG11, MPEG98/M5583, 1999.
- [7] Li W. Overview of the Fine Granularity Scalability in MPEG-4 Video Standard[J]. IEEE Trans. Circuits and Sys. for Video Tech., 2001, 11(3): 301-17.
- [8] Daubechies I, Sweldens W. Factoring Wavelet Transforms into Lifting Steps[J]. Fourier Anal. Appl., 1998, 4(3).
- [9] Said A, Pearlman W A. A New, Fast and Efficient Image Code Based on Set Partitioning in Hierarchical Trees[J]. IEEE Trans. Circuits System, Video Technol., 1996, 6: 243-250.
- [10] 张军. 基于小波的伸缩性视频编码研究[D]. 国防科技大学硕士论文, 2002.
- [11] Choi S J, Woods J W. Motion-compensated 3-D Subband Coding of Video[J]. IEEE Transactions on Image Processing, 1999, 8: 155-167.

