

文章编号:1672-3961(2011)02-0058-04

基于迭代译码算法的分级分布式视频编码

卿粼波,吕瑞,欧先锋,郑敏,何小海

(四川大学电子信息学院图像信息研究所,四川 成都 610064)

摘要:分布式视频编码(distributed video coding, DVC)是一种新的视频编码算法,与传统视频编码系统相比,具有低编码复杂度和高鲁棒性的优点。但它的压缩率比较低,对画面组(group of picture, GOP)的长度依赖性比较大。将H.264解码算法引入像素域视频编码系统(pixel domain Wyner-Ziv codec, PDWZ)中,提出了一种混合视频编码框架。系统采用了并行编码和迭代解码结构,同时引入了预测编码和一种新的边信息生成算法。仿真结果表明,提出的结构比最优的PDWZ有更高的压缩效率。同时它使视频能顺序编解码,并提供了多层的可分级性。

关键词:画面组(GOP);迭代译码;预测编码;可分级视频编码;WZ视频编码技术

中图分类号:TP919.81 **文献标志码:**A

An iterative decoding based scalable distributed video coding

QING Lin-bo, LÜ Rui, OU Xian-feng, ZHENG Min, HE Xiao-hai

(Institute of Image Information, College of Electronics and Information Engineering,
Sichuan University, Chengdu 610064, China)

Abstract: As a new video coding paradigm, distributed video coding (DVC) has lower encoding complexity and higher robustness to transmission error than conventional video coding. However, its compression ratio was comparatively low and highly dependent on group-of-pictures (GOP) size. A combined video coding scheme was proposed by introducing H.264 coding into typical pixel domain Wyner-Ziv (PDWZ) coding. The coder has a parallel encoding and iterative decoding structure, in which a predictive coding scheme and a novel side-information (SI) generation algorithm are also adopted. Simulation showed that the proposed architecture has higher compression ratio than an optimized PDWZ coder. Moreover, it enables a video to be processed in sequential order and provides multilayer in scalability.

Key words: group of picture(GOP); iterative decoding; predictive coding; scalable video coding; Wyner-Ziv codec

0 引言

DVC也称之为WZ(Wyner-Ziv)视频编码技术,主要基于2个理论基础:Slepian-Wolf^[1]理论和Wyner-Ziv^[2]理论。它在解码端用边信息(side information, SI)估计原始信息的统计信息。和传统视频编码技术用高复杂度的编码器来压缩数据相比不同的是WZ视频编码算法把相关性估计的工作从编码端移到了解码端,极大地降低了编码端的复杂度和能量消耗。WZ解码器还消除了预测回路的

问题,具有更强的传输误码鲁棒性。

WZ视频编码系统中,边信息质量对整个系统性能影响很大。Ascenso提出的边信息内插生成算法^[3]是目前最典型的边信息生成算法之一,然而这种内插结构要求非顺序解码,同时限制了GOP的值,因为大的GOP将导致系统性能下降和比较大的解码延时。Yaacoub根据视频信息自适应设置GOP值来解决GOP问题^[4]。但是,GOP值还受关键帧相关性的限制。

除了编码复杂度低以外,分布式视频编码的另一个特点就是无线信道误码的高鲁棒性,这使得

收稿日期:2010-12-01

基金项目:四川大学青年科学基金资助项目(校青2008005);欧盟FP7-PEOPLE-IRSES资助项目(247083)

作者简介:卿粼波(1982-),男,讲师,博士,主要研究方向是通信与信息系统,模式识别。E-mail:qing_lb@scu.edu.cn

它非常适用于无线多媒体终端设备中。因此针对异构网络, 学者们开始将 DVC 应用于分级视频编码 (scalable video coding, SVC) 领域, Q. Xu 等提出了类似于 MPEG4-FGS 的分层 WZ 结构^[5], 将传统的视频编码器 H. 26L 作为基本层编码, WZ 作为增强层, 获得了较好的分级性能。Tagliasacchi 提出了时间和空间可分级系统^[6], 在采用传统 H. 264 对基本层进行编码的同时, 在增强层 (enhanced layer, EL) 采用了 DVC 编码器来增强系统对误码的鲁棒性, 同时系统还具有可分级特性。然而, 这些系统主要关注的是系统的可分级性, 或对误码的鲁棒性, 并未就提高编码效率进行讨论。

本文以提高整个系统编码效率为目标, 将传统视频编码和 DVC 编码结合, 提出了一种基于并行编码和迭代解码结构的空间可分级视频编码系统。在基本层对小部分视频数据采用 H. 264 编码, 在增强层采用 DVC 编码方式。同时, 本文还提出了新的边信息生成算法, 结合 H. 264 编码器的信息, 使系统压缩率对 GOP 的依赖性大大降低。

1 提出的视频编码系统

本文提出的视频编码框架主要基于 Ascenso 提出的像素域 WZ 视频编码器 PDWZ^[3], 系统框图如图 1 所示。系统总体结构上分为两层, 基本层 (base layer, BL) 视频序列和 EL 视频序列。在编码前, 首先将原始视频帧按相邻像素分为 4 层, 分层示意图如图 2 所示。编码端, BL 层采用传统的 H. 264 编码器进行编码; EL 层的编码采用基于预测的并行 PDWZ 编码器进行编码; 3 个 EL 子层在并行 PDWZ 编码器中进行独立量化和信道编码, 产生 3 个独立的比特流。解码端, BL 层数据经过 H. 264 解码, 生成低质量低分辨率的视频序列; 在 EL 层, 解码器结合 BL 层的数据完成基于迭代结构的 WZ 解码生成一个高分辨率高质量的视频序列。当带宽或者终端处理能力受限时, BL 优先传输以保证基本的视频通信。而当 EL 层的比特流部分解码时, 分别能提供不同的空间分辨率。

1.1 EL 层基于预测编码的并行 PDWZ 编码器

如图 2 所示, 系统包括 3 个 EL 层像素帧 EL1, EL2, EL3, 其尺寸与 BL 层相同。在对各个 EL 层进行编码时, 以 BL 层编码过程中产生的解码图像作为预测帧, 对“BL-EL” (即残差帧) 进行 PDWZ 编码。预测的引入可以降低信源的熵, 从而使得压缩效率大幅提升。预测编码中一个关键点是编码端和

解码端的参考信息必须一致, 以避免误差漂移。因此, 本文在编解码端都用 BL 编解码器中重建的 BL 帧来生成 EL 层的参考帧信息。由于 EL 和 BL 都来源于同一帧, 临近像素的相关性保证了预测质量。另外该框架使预测性能也不受 GOP 的影响。

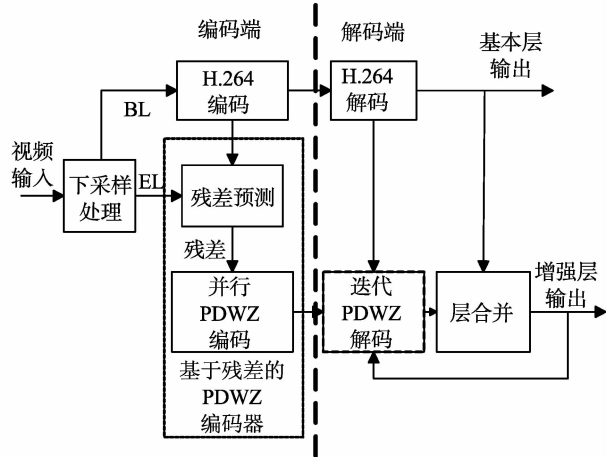


图 1 文中提出的视频编码框架
Fig. 1 The proposed video coding scheme

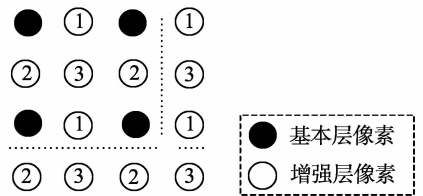


图 2 BL 层和 EL 层示意图
Fig. 2 The dividing of the BL and the EL

1.2 迭代解码结构

针对 EL 层的数据进行解码, 本文设计了一种基于迭代结构的并行解码器, 其框图如图 3 所示。

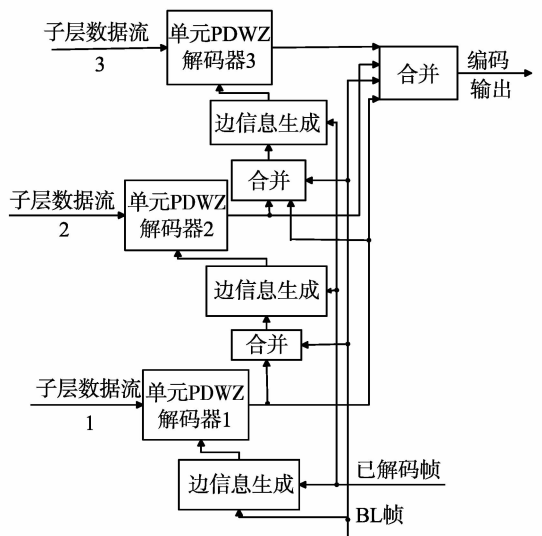


图 3 迭代 PDWZ 解码结构
Fig. 3 The iterative PDWZ decoder structure

首先, 在 EL1 的解码过程中, 只有 BL 和前面已解码帧可用于边信息的生成。然后, 结合已解码的

EL1 和 BL 层联合生成边信息用于 EL2 的解码。EL3 的解码以此类推。显然随着解码过程的继续,边信息越来越精确,后面子帧解码所需校验位也越来越少。

1.3 搜索和补偿边信息生成算法

边信息是影响 WZ 编码系统 RD 性能的重要因素,其质量越高解码所需要的校验信息越少,因此本文提出了适合本系统的高性能边信息生成算法。传统的 WZ 系统中,解码端没有当前需要解码图像帧的信息,必须基于前后已解码的关键帧,这使得 GOP 越大,边信息重建质量越差。而在本文系统中,随着各层数据的解码,可以用来为边信息重建提供更多的信息。因此针对本文系统特点,提出了一个搜索补偿算法来利用这些信息。

Step 1 用已重建信息 R 作为参考信息生成当前帧 I 的一个粗边信息 C 。首先,已解码的 BL 和子 EL 联合上采样获得 I 的近似帧 A ,并将 A 分成多个子块。对每一子块,用亚像素搜索算法找到其在 R 中的近似块。搜索过程中,只将已重建像素点用于相似度匹配计算。因此,越多子 EL 帧可用,搜索结果越精确。然后,将 R 中的匹配块复制到 C 中。与传统的内/外插补算法相比,本系统中用于匹配的像素块是直接从前帧提取到的,运动矢量更精确。另外,因为该过程中只需已解码的前一帧,不需限制 GOP 来保证高相关性。同时,该系统可以顺序编解码,没有额外的延时。

Step 2 通过误差补偿过程更新粗边信息 C 。首先,计算 A 中已解码像素和 C 中相同位置像素的误差 E_A 。对于 A 中未解码位置像素的误差 E_U 用邻近 E_A 的像素插值得到。 E_A 和 E_U 联合得到一个误差帧 E 。 E 和 C 相加得到一个更新后的边信息。最后,将与待解码子帧相同位置处的边信息提取出来作为后面解码的边信息。误差估计的引入基于下面的原因:虽然运动搜索框架允许大 GOP 值,但如果将前一帧的块直接用做边信息,虚拟信道就包含 2 种误差。一个是原始前一帧和已重建前一帧的误差;另一个是原始前一帧和原始当前帧的误差。因为边信息的失真会导致重建结果的失真,当前帧解码过程中未纠正的虚拟信道噪声就成为下一帧的第一种误差。再加上第二种误差,随着误差的累积,重建失真很快会达到量化的水平,进而极大地降低视频通信质量,这样 GOP 值就会受限。上述纠正更新过程有效纠正了先前重建帧和当前帧已解码部分的误差。在第一个子 EL 帧的解码过程中,误差补偿算法保证了生成的边信息失真和已重建 BL 在同一

个失真水平上。因为 BL 的 H.264 编码器采用了一个出色的量化器来保证一个稳定的解码质量,所以 PDWZ 编码器能保证一个独立于 GOP 的、稳定的性能。另外,在迭代解码过程中,随着更多的子 EL 帧解码成功,可以得到更多的像素值,下一个子帧的插补越准确。

2 仿真结果

为了测试本系统的性能,对 CIF(352 × 288)格式、15fps 的 Foreman, Silent 和 Coastguard 序列的亮度信号进行算法仿真,采用比特率失真进行系统性能评估。比特流用每秒传送比特数进行度量,失真用峰值信噪比(peak signal to noise ratio, PSNR)进行度量。BL 用 H.264 参考软件 JM14.0 进行编解码。

Foreman 序列前 32 帧 GOP = 8 的 BL 和 EL 性能如图 4 所示。其中子序列的失真是将相应的序列图像内插成 CIF 尺寸后进行计算的。从图中可以看出,提出的系统提供了视频编码中的多层可分级性。另外,注意到 BL 和 EL1 都有“性能上限”效应,这是因为插补不能恢复原始视频中的高频部分,因此仅仅提高传输比特流不能持续提高重建性能。

为了验证本系统在总体压缩性能方面的效果,本文将提出的系统与传统的非可分级 PDWZ 编码器系统进行了比较,结果如图 5、6、7 所示。其中非可分级 PDWZ 编码系统采用的是已有的 PDWZ 编码器。在对本文系统进行测试时采用了 BL 层和所有 EL 层的数据。由图可知,对于 3 个序列而言,提出的算法都比非可分级 PDWZ 系统更优。

因为 Forman 序列包含剧烈并且细碎的运动,PDWZ 编码器通过运动补偿可以有效地挖掘帧与帧之间的相关性,因此这两个系统性能之间的差距主要来自提出算法对预测误差进行的补偿。对于 Silent 序列,背景区域包含很多高频成分,导致提出系统在低比特率条件下性能的降低。但是,由于序列高速度、大尺度和不规则运动的特点,PDWZ 中的运动插补算法不能利用帧与帧之间的相关性,因而在中高比特率条件下,本文算法性能更好。Coastguard 序列包含多种高速运动物体,并且互相交叉。插值算法较难获得一个有效的边信息。但是,本文提出的算法利用较低层得到的信息和误差补偿算法有效解决了该问题,并且获得了较好的性能。另外,在大 GOP 情况下,随着关键帧间相关性的降低,非可分级 PDWZ 系统性能降低得很明显,而本文系统性能受 GOP 的影响则较小。

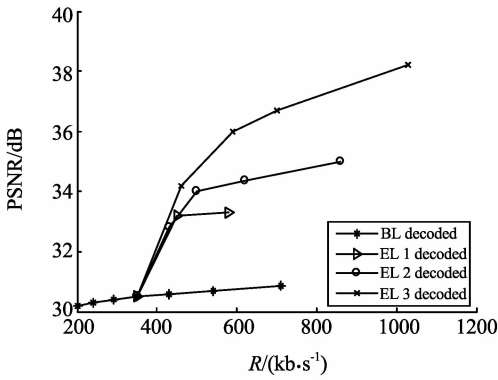


图4 Forman的可分级RD性能

Fig. 4 Scalable RD performance for foreman

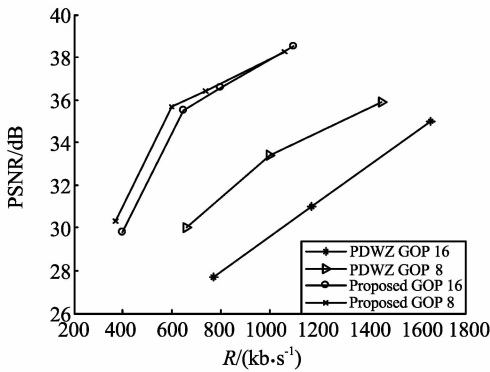


图5 Forman的RD性能比较(1~100帧)

Fig. 5 RD performance comparison for foreman sequence (frame 0 ~ 100)

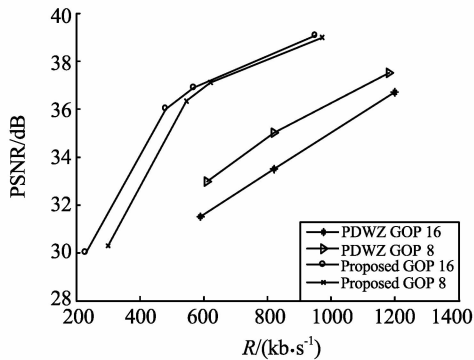


图6 Silent的RD性能比较(100~200帧)

Fig. 6 RD performance comparison for silent sequence (frame 100 ~ 200)

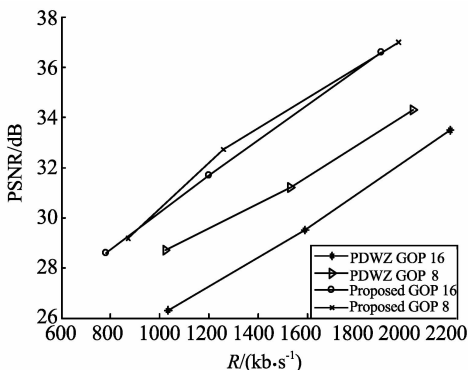


图7 Coastguard的RD性能比较(1~100帧)

Fig. 7 RD performance comparison for coastguard sequence (frame 0 ~ 100)

3 结论

本文提出了一个基于并行编码迭代译码的空间可分级PDWZ系统,来解决大GOP的问题,并具有多层可分级性。仿真结果验证了本文提出系统的有效性。另外,虽然本文在编码器端引入了传统的H.264编码算法,但是由于其尺寸远小于原始视频,对整个系统的计算复杂度增加有限。凭借DVC自身编码复杂度低的特点,该系统与传统视频编码相比,编码复杂度大大降低。另外,本系统EL层没有帧间预测环路,某些帧EL层的错误不会传递到后面的EL解码过程中,系统的鲁棒性得到了很大的提高。

参考文献:

- [1] SLEPIAN D, WOLF J. Noiseless coding of correlated information sources[J]. IEEE Trans Inform Theory, 1973, 19(4): 471-480.
- [2] WYNER A, ZIV J. The rate-distortion function for source coding with side information at the decoder[J]. IEEE Trans Inform Theory, 1976, 22(1): 1-10.
- [3] ASCENSO J, BRITES C, PEREIRA F. Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding [C/OL]//5th EURASIP Conf. Speech and Image Processing, Multimedia Communications and Services, Smolenice, Slovak Republic, June 29-July 2, 2005; The European Association for Speech, Signal and Image Processing (EURASIP). [2010-10-23]. http://amalia.img.lx.it.pt/~cicb/papers/EURASIP'05_ImprovingFI.pdf
- [4] YAACOUB C, FARAH J, PESQUET-POPESCU B. New adaptive algorithms for GOP size control with return channel suppression in Wyner-Ziv video coding[J]. International Journal of Digital Multimedia Broadcasting, 2009(2009):1-11.
- [5] XU Q, XIONG Z. Layered Wyner-Ziv video coding[J]. IEEE Trans Image Process, 2006, 15(12): 3791-3803.
- [6] TAGLIASACCHI M, MAJUMDAR A, RAMCHANDRAN K. A distributed-source-coding based robust spatio-temporal scalable video codec [C/OL]//Proceedings of Picture Coding Symposium (PCS'04), San Francisco, CA, USA, Dec. 2004; Picture Coding Symposium. [2010-10-23]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.4953&rep=rep1&type=pdf>.

(编辑:孙培芹)