

可伸缩视频编码的码率控制技术研究

刘 光¹, 房 胜^{1,2}, 尉希娟¹

(1. 山东科技大学 信息科学与工程学院, 山东 青岛 266510;

2. 清华大学 深圳研究生院, 广东 深圳 518055)

摘 要:可伸缩视频编码能生成空域、时域、质量可伸缩的视频码流,如何对其进行码率控制来生成最优的可伸缩性码流,对适应不同需求的终端十分重要。结合可伸缩视频编码码率控制的研究现状,描述了最新的研究进展,综述了码流控制的原理、经典算法,分析总结了可伸缩视频编码的关键技术及在可伸缩视频编码中的挑战。通过对增强层中各种码率控制算法的优缺点的分析,总结出增强层的时域可扩展和空域可扩展中码率控制的最优解决方案。最后对可伸缩编码中码率控制技术的发展趋势进行了展望。

关键词:可伸缩视频编码;码率控制;分层B帧

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2010)10-0026-04

Research in Rate Control Technique for Scalable Video Coding

LIU Guang¹, FANG Sheng^{1,2}, WEI Xi-juan¹

(1. College of Information Science and Engineering, Shandong University of

Science and Technology, Qingdao 266510, China;

2. Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China)

Abstract: Scalable video coding can generate spatial, temporal and quality scalable video streaming. Rate control plays an important role in how to generate the optimal scalable bit - streams to adapt to different terminals. Combined current situations for scalable video coding rate control, describe the latest research progress, review the principles and classical algorithms of rate control, and analyze and summarize the key technology for scalable video coding and the challenge of rate control in scalable video coding. Through the analysis of advantages and disadvantages for rate control algorithms in the enhancement layer, summarize the optimal solution for temporal and spatial scalable rate control in the enhancement layer. Finally, prospect the future development of rate control technology in scalable video coding.

Key words: scalable video coding; rate control; hierarchical B

0 引 言

目前,基于 H. 264 的可伸缩扩展的视频编码系统^[1]被 ISO MPEG 和 ITU VCEG 组成的联合视频编码小组(JVT)采纳为可伸缩视频编码(SVC, Scalable Video Coding)的标准,并提供时域、空域和质量可伸缩。其中,时域可伸缩是通过等级B帧编码结构实现的,而空域可伸缩则采用了分层的方法,质量可伸缩性采用了位平面编码的精细可伸缩(FGS, Fine Granularity Scalability)视频编码技术。

对视频通信而言,由于通信信道带宽有限,需要控

制编码的码流,来保证码流的顺利传输和信道带宽的充分利用。尽管 SVC 的质量可伸缩具有 FGS 的特性,并在理想的条件下,能替代传统码率控制中核心的算法获得特定码率的功能。但是,FGS 这种在解码端由用户“被动”地确定码率的方法与码率控制在编码端“主动”控制码率的策略相比,码率控制仍然具有 FGS 不可比拟的优势;另外,码率控制可以在率失真最优意义下选择合适的编码参数并通过宏块级的精确调整避免类似 FGS 仅仅简单的对某一质量层的所有 NAL 单元数据等比例分割带来的图像质量波动^[2]。总体来说,FGS 的率失真性能不及码率控制。因此,在 SVC 中实现码率控制十分必要。文中对基于 H. 264 的可伸缩视频编码中码率控制技术的研究进展进行了综述。

1 码率控制的问题描述

码率控制是视频编码的重要部分。一般来讲,它

收稿日期:2010-03-02;修回日期:2010-06-14

基金项目:国家自然科学基金重大研究计划项目(90718011)

作者简介:刘 光(1984-),男,山东聊城人,硕士研究生,研究方向为视频编码与传输;房 胜,副教授,研究方向为视频编解码技术、信息家电等。

由三个相互联系的部分组成:码率预分配、缓冲区管理和量化参数调节,其基本框架如图 1 所示。码率预分配通常在图像组(GOP)层、帧层和宏块组层进行。缓冲区管理一般都是建立在一个假想的解码器模型上,起到平滑码率波动的作用。量化参数调节可以在帧层或宏块组层进行。

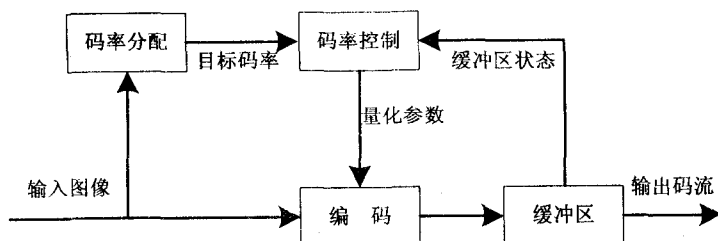


图 1 码率控制原理图

在视频压缩标准的发展过程中,码率控制算法也不断改进,各种经典的码率控制算法被提出。如:

MPEG-1 中的校验模型 SM3^[3]提供了一种简单的码率控制方法,首先根据前一帧图像编码产生的比特数确定下一帧的目标比特数,然后根据缓冲器的“充盈程度”来确定量化参数。

MPEG-2 的 TM5^[4]将码率控制分为三个步骤:目标比特分配、基于缓冲区饱和度的码率控制和自适应量化。由于 TM5 模型中不同编码类型帧有各自的虚拟编码缓冲区并且目标比特分配单位是图像,没有对帧中每一个宏块进行比特分配;这就存在场景切换适应性、缓冲区一致性和一帧图像中宏块比特分配不均匀等问题。

VM8^[5]根据二次 R-Q 模型,并用此模型为不同类型的帧进行目标码率控制,并引入滑动窗口来调整模型的参数,实现多尺度、不同复杂度的码率控制。

TMN8^[6]根据量化系数的经验熵模型,推导出宏块的目标码率预测公式,再结合经典的率失真模型,由拉格朗日最优化,求出 MSE 准则下的宏块层的最佳量化步长。和 VM8 相比,TMN8 可更精确地控制目标码率,保持缓冲区的稳定性。

与传统视频编码技术相比,作为最新编码标准的 SVC 在编码时采用了分级机制和层间预测机制,使得传统的码率控制技术不能很好地应用在 SVC 中。主要表现在以下几个方面:

(1)传统的码率控制算法可以使用在 SVC 的基本层,但只针对 IPPP 和 IBBP 结构。SVC 采用的 Hierarchical B 预测结构,从而产生了如何为 Hierarchical B 结构进行码率控制的新问题。

(2)SVC 采用了层间预测机制,不能简单地将传统视频编码中的码率控制方案应用于各增强层。

(3)与传统的单层编码相比,SVC 的整个编码效率的损失在基本层和增强层的分布是不均匀的。需要建立新的率失真模型。

2 SVC 中基本层的码率控制分析

Leontaris^[7]等人将码率控制扩展到 SVC 的基本层中,并在 JSVM 上实现了码率控制方案,取得不错的效果。但是,该方案欠缺对分层时域结构的特殊性的考虑。该方案的核心算法使用 H. 264 中的码率控制技术。H. 264 编码算法采用了 RDO(Rate Distortion Optimization)进行模式选择,使得它的码率控制困难,原因在于采用了 R-D 优化模式判别,

宏块的编码模式与量化参数的选取有关。这就造成以下问题:为了进行宏块的模式判别,必须先给出一个量化级。然而为了控制码率,量化级的计算一般要根据编码模式判别之后的帧内预测或者帧间运动估计残差来决定,造成了所谓的“蛋鸡悖论”,为了解决该问题,Ma^[8]等提出了 JVT-F086 提案,整个码率控制的流程与 TM5 算法类似,主要通过一个假想的参考解码器的充盈程度来确定量化参数。由于它不是一个单通道算法,所以时间复杂度过高。Li^[9]等人提出的 G012 提案,通过前一帧相同位置基本单元的 MAD(Mean Absolute Difference)来预测当前帧相同位置基本单元的 MAD 值来避免模式选择和码率控制之间存在的蛋鸡悖论,并利用 VM8 中的二次 R-Q 模型和滑动窗口来计算量化参数和调整模型的参数。H017^[10]提案在 G012 提案的基础上,将帧的目标比特数限制在一个上下界之间,来符合假设的参考解码器的要求。

3 增强层的码率控制

目前,涉及到增强层的码率控制工作还比较少,Xu^[11]等人提出了一个 SVC 的码率控制方案,在 JSVM 的早期版本 SVM3.0 实现,由于方案只是工作在时域子带上,所以 QP 的确定并没有影响率失真下的运动搜索和模式选择。Liu 等把他们之前提出的针对 H. 264/AVC 的码率控制算法^[12]扩展到了 SVC,在对当前帧编码时,通过从当前空间层的时域前帧和当前时刻的空间基本层两者之间选择其中最优的一种^[13]。根据 SVC 提供的分层 B 帧的预测和空间/CGS 可伸缩编码结构,文中将增强层的码率控制划分为时域可扩展的码率控制和空域可扩展的码率控制分别进行阐述。

3.1 时域可扩展的码率控制

可伸缩编码采用分层 B 帧(Hierarchical B)预测进

行时间可伸缩,如图 2 所示。基于分层 B 帧的时域可扩展中,帧的编码顺序和解码顺序是完全去耦合的。任何图像帧都可以作为后续帧的运动估计/运动补偿过程的参考帧。这就为码率控制带来困难,因为在进行目标比特分配时必须将时间分级考虑在内。

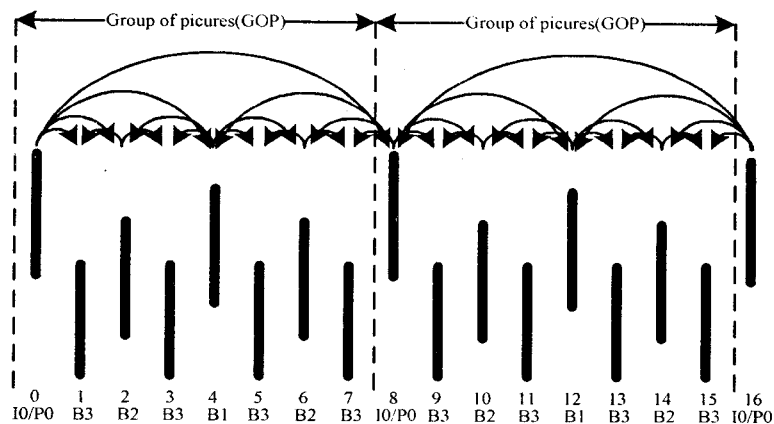


图 2 四级分层 B 帧(Hierarchical B)的结构

文献[13]提出了一种分级 B 帧的码率控制方案,该方案通过在一个 GOP 级别内对不同时间级别的时域增强层设定不同的权重因子,即对底层的时域增强帧设定较大的权重因子,高层的时域增强帧设定较小的权重因子,以保证底层被用作参考帧的帧分配更多的码率。然后由每帧分配的码率通过线性的 R-Q 模型和缓冲区的充盈度计算出帧级或者宏块级别的量化参数 QP。从而实现时域增强层的码率控制,但其权重因子是固定的经验值,并不是对所有的视频序列都是最优的。

文献[14]提出了一种分层和叠层(cascade)量化的概念,其每层的权重因子的设定使用到他们先前研究的方法[15],通过时域等级中的相关帧和运动补偿预测的滤波器来确定权重因子。第 K 层的权重是由前一层的权重计算得来的,然后由得到的第 K 层的权重来计算当前帧的权重因子(scaling-factors)。在此基础上,由权重因子和每帧的复杂度来得到各自的码率配额。其次,针对不同帧的统计特性分别使用不同的率失真模型,I/P 帧指定二次率失真模型,B 帧指定线性模型来确定各自的量化参数 QP。达到时域的码率控制。

3.2 空域可扩展的码率控制

SVC 提供了一系列层间预测机制,尽可能地应用各空间层的相关性提高编码效率。SVC 的多种层间预测技术为增强层的码率控制提供了丰富的空间。图 3 所示为一种分层 B 帧结构的层间预测编码的例子。如何更合理地利用层间的预测机制来提高码率控制的效率是当前的研究热点和难点。文献[16]提出了一种

时域和 CGS(Coarse-Grain-SNR)码率控制框架,通过改进的 TMN8 模型来对 I、P 和参考帧 B 帧进行码率控制。用二次细化量化参数的方法来预测分层 B 帧中的量化参数,在分层的 B 帧中,先利用层间标准差预测值来预测第一次量化参数,然后在率失真优化的模式下选择最优的运动估计和模式选择,最后用最优下的标准差来进行二次细化量化参数。由于需要对 B 帧的量化参数进行二次细化量化参数,所以计算量比较复杂,不适于实时的码率控制。

文献[13]把他们之前提出的针对 H.264/AVC 的码率控制算法扩展到了 SVC。通过对残差纹理信息的平均绝对值(MAD)的准确估计,来提高码率控制,在空间层,他们提出了一种动态的选择策略,针对当前帧,从当前空间层的时域前帧和当前时刻的空间基本层两者之间选择一种更加准确的 MAD 估计值。其中 i 表示帧数, j 表示层数, Y_1, Y_2, Z_1, Z_2 是通过线性回归方法更新的四个参数。

$$MAD_{pred,1}[j][i] = Y_1[j] \cdot MAD_{actual}[j][i-1] + Y_2[j] \quad (1)$$

$$MAD_{pred,2}[j][i] = Z_1[j] \cdot MAD_{actual}[j-1][i] + Z_2[j] \quad (2)$$

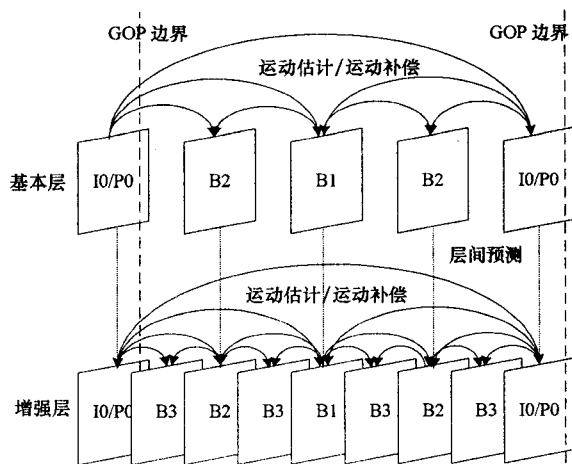


图 3 分层 B 帧结构的层间预测编码

由等式(1)可知,MAD 值是通过当前空间层中的前一帧来预测的,等式(2)中的 MAD 则是由当前时间层中的基本层来预测的。可以清楚的看到,当视频图像运动平滑时,等式(1)可以获得更好的预测信息,而当视频图像运动剧烈时,等式(2)可以获得更好的预测信息。其次,通过比较算法获得最优的 MAD 值,经过线性 R-Q 模型和缓冲区的充盈度计算出帧级或者宏块级别的量化参数 QP。比起文献[16]提到的方案,该

算法充分考虑了空间增强层中的层间预测、层间帧内预测机制,并且与二次R-Q模型相比,在PSNR和码率相差很小的情况下,算法复杂度明显降低。

4 结束语

码率控制的根本宗旨是确定适当的量化参数使得编码产生的码流适应给定的码率并使解码图像的失真最小。尽管可伸缩编码方案通过一次编码产生的高码率提供可分级,但只有在码率控制下才能产生各精细程度的码流。如何在可伸缩编码中设计更优的码率控制方案成为当前的研究重点。通过对比分析当前最新的可伸缩编码中的码率控制技术,认为未来码率控制技术将在以下方面得到进一步发展。

(1) 基本层和增强层的联合码率控制。

目前,码率控制算法在SVC上扩展的非常简单,基本上仅仅是对各个编码层进行重复性的延拓。文献[13]提到了基本层和增强层的联合码率控制思想,文献[17]在码率控制算法中引入预处理过程。但因这些算法只是对各个编码层进行重复的延拓,执行效率较低。如何在基本层与增强层间划分比特数,如何使用已编码的基本层来预测增强层的量化信息,使其既能达到可分级的要求又能获得更令人满意的视频效果是待解决的问题。

(2) 低比特率、低延时的码率控制算法。

当今,3G网络和无线通信大规模应用。低比特率传输和低延时业务成为可伸缩编码码率控制的主要挑战。基于感兴趣区域可伸缩编码的提出,使得如何在低比特率下对感兴趣区域的可伸缩编码^[18]进行码率控制以提高视频主观质量成为一个新的研究方向。如何简化现有R-Q模型的高复杂度,使其能适应低比特率下的实时通信需求,具有很高的研究价值。

(3) 增强层中各层间的比特分配。

可伸缩编码提供了丰富的分层机制,也为其码率控制带来困难。权重因子^[13,15]的提出为增强层的码率控制方案提供了思路。但文献中提到的权重因子大都是启发式的选择,并非对所有的视频序列最优。如何建立更优的权重因子模型,并合理地各增强层分配比特,还有待改进。

参考文献:

[1] Schwarz H, Marpe D, Thomas W. Overview of the scalable video coding extension of the H.264/AVC standard[J]. IEEE Transactions on CSVT, 2007, 17(9): 1103 - 1120.
[2] Li Weiping. Overview of Fine Granularity Scalability in MPEG

- 4 Video Standard[J]. IEEE Transactions on CSVT, 2001, 11(3): 301 - 307.
[3] ISO-IEC/JTC1/SC29/WG11, MPEG-1 Simulation Model 3(SM3)[S]. 1992.
[4] ISO-IEC/JTC1/SC29/WG11, MPEG-2 Test Model 5 Draft[S]. 1993.
[5] Clzhang T, Zhang Y Q. A New Rate Control Scheme Using Quadratic Rate Distortion Model[J]. IEEE Transactions on CSVT, 1997, 7(1): 287 - 311.
[6] Ribas C J, Lei S. Rate Control in DCT Video Coding for Low-delay Communications[J]. IEEE Transactions on CSVT, 1999, 9(1): 172 - 185.
[7] Leontaris A. Rate Control for the Joint Scalable Video Model (JSVM) [M]. Joint Video Team, JVT-W043, San Jose, California, USA: [s. n.], 2007.
[8] Ma Si wei, Gao Wen, Lu Yan. Proposed draft description of rate control [C]//Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, Doc. JVT-F086-L, 6th Meeting. Awaji, Island, JP: [s. n.], 2002.
[9] Li Z G, Pan Feng, Lim K P, et al. Adaptive Basic Unit Layer Rate Control for JVT [C]//In Proceedings of JVT-G012, 7th Meeting. Thailand: [s. n.], 2003.
[10] Ma Siwei, Li Z G, Wu Feng. Proposed Draft of Adaptive Rate Control [C]//Proceedings of JVT-H017, 8th Meeting. Geneva: [s. n.], 2003.
[11] Xu Long, Ma Siwei, Zhao Debin, et al. Rate control for scalable video model [C]//Proceedings of the SPIE. Beijing, China: [s. n.], 2005: 525 - 534.
[12] Liu Y, Li Z G. A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard [J]. IEEE Transactions on CSVT, 2007, 17(1): 68 - 78.
[13] Liu Y, Li Z G. Rate Control of H.264/AVC Scalable Extension [J]. IEEE Transactions on CSVT, 2008, 18(1): 116 - 121.
[14] 徐龙, 高文, 季向阳, 等. 一种面向SVC的码率控制算法[J]. 计算机学报, 2008, 31(7): 1175 - 1184.
[15] Xu Long, Gao Wen, Ji Xiangyang, et al. Rate control for hierarchical B-picture coding with scaling-factors [C]//Proceedings of the ISCAS' 2007. New Orleans, USA: [s. n.], 2007: 49 - 52.
[16] Xu Long, Gao Wen, Ji Xiangyang, et al. Rate control for spatial scalable coding in SVC [C]//Proceedings of the PCS2007. Lisboa, Portugal: [s. n.], 2007.
[17] 谈永敏. 可伸缩视频编码的率失真分析及其码率控制[D]. 上海: 上海交通大学, 2008.
[18] Benierbah S, Khamadja M. A new technique for quality scalable video coding with H.264 [J]. IEEE Transaction on CSVT, 2005, 15(11): 1332 - 1340.