

基于视点质量一致性的多视点视频码率控制

熊 婕, 刘 峰, 范晶晶, 徐 琼

(南京邮电大学 图像处理与图像通信江苏省重点实验室, 江苏 南京 210003)

摘 要:多视点视频是未来视频技术的重要发展方向,而码率控制是影响多视点视频传输质量的重要因素。基于多视点视频编码统计特性,提出了一种面向多视点视频编码的比特分配与码率控制算法。该比特分配算法分为三层:GGOP层、GOP层和帧层,其中重点研究了GOP层比特分配,即视点间的比特分配策略,利用拉格朗日乘子法,提出了一种基于视点质量一致性优化的视点间比特分配模型。实验结果表明,与固定量化参数的控制策略相比,本算法能在精确控制输出码率的前提下,使得各视点视频质量更加均匀。

关键词:多视点视频编码;码率控制;比特分配;质量一致性

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2011)07-0063-04

Rate Control Algorithm for Multi-View Video Based on Consistent View Quality

XIONG Jie, LIU Feng, FAN Jing-jing, XU Qiong

(Image Processing & Image Communication Key Lab of Jiangsu Province, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract: Multi-view video is a very important developing direction in video technologies in future, and rate control is one of the key factors influencing the multi-view video transmission quality. Based on the statistical features of the multi-view video coding, an efficient rate control algorithm for multi-view video coding is proposed. The rate control algorithm contains three layers: GGOP layer, GOP layer and frame layer. The bit allocation strategy based on Lagrange optimum algorithm among views is emphasized and a bit allocation model for consistent view quality is proposed. Compared with the fixed QP control strategy, the experiment results show that the proposed algorithm can efficiently control the bit rate and obtain more consistent views.

Key words: multi-view video coding; rate control; bit allocation; consistent view quality

0 引 言

多视点视频(Multi-view Video)是一种新型的具有立体感和交互操作功能的视频,通过在场景中放置多台摄像机,记录多个视点数据,提供给用户以视点选择和场景漫游的交互能力,在自由视点视频、虚拟视点合成、自由视点电视、多视点视频点播等数字娱乐领域都有着非常广泛的应用前景^[1,2]。但是随着视点的增加,视频数据量也急剧增大,使得视频数据的存储和传输变得十分困难,有必要对多视点视频进行有效的压缩编码,而码率控制是多视点视频压缩编码中一项非常重要的技术,任何视频压缩标准离开码率控制其应用都会受到限制。

联合视频专家组(Joint Video Team, JVT)长期致力于视频编码的标准化工作,目前已经提出了多视点视频编码参考模型(Joint Multi-view Video Coding, JMVC)^[3]。但是该模型本身并没有给出有效的码率控制策略,而是简单地地为每一个时间分解层指定一个固定的量化参数。针对JMVC模型的测试序列仿真结果表明,经固定量化参数编码后的视点间质量波动较大。

在多视点视频编码中,峰值信噪比(Peak Signal Noise Ratio, PSNR)和视频突变(Jerkiness)是评价视频质量的重要参考指标,也直接影响用户的视觉感受。例如用户在观看时改变视点,若视点质量波动太大,对视频感知质量有较大的负面影响。因此多视点视频编码的码率控制除了需要在视点内进行比特分配,还需要在各个视点间进行合理的比特分配,以保持各视点质量的一致性。文中针对多视点视频编码的特性,提出了一种基于视点质量一致性优化的多视点视频码率控制算法。

收稿日期:2010-12-13;修回日期:2011-03-20

基金项目:国家自然科学基金(61071166, 61001152);江苏省高校自然科学基金项目(09KJB510015)

作者简介:熊 婕(1987-),女,硕士,研究方向为图像处理与视频编码;刘 峰,教授,研究方向为图像处理与视频编码。

1 多视点视频编码结构

由于视频编码质量、码率控制算法与视频编码结构有着非常密切的联系,文中首先对多视点视频编码结构简单介绍如下。作为多视点视频编解码的研究平台,多视点视频编码的参考模型 JMVC 采用了一种由德国 HHI 实验室提出的基于分级 B 帧(Hierarchical B Prediction, HBP)预测结构,该预测结构在时间方向上,各视点内采用分层 B 帧预测结构去除时间上的冗余,在视点方向上,各视点间采用 IBP 预测结构去除视点间的冗余。由于 HBP 充分采用视点内和视点间的相关性消除空间和时间上的冗余,因而取得了较高的编码效率^[4-6]。图 1 给出了 JMVC 编码预测结构示意图,其中 S_n 表示视点, T_n 表示时间,水平方向箭头表示时间预测关系,垂直方向箭头表示视点间预测关系。图中 GOP(Group of Pictures)长度为 8,多个视点在同一段时间内的 GOP 所构成的单元称为 GGOP(Group of GOP)。

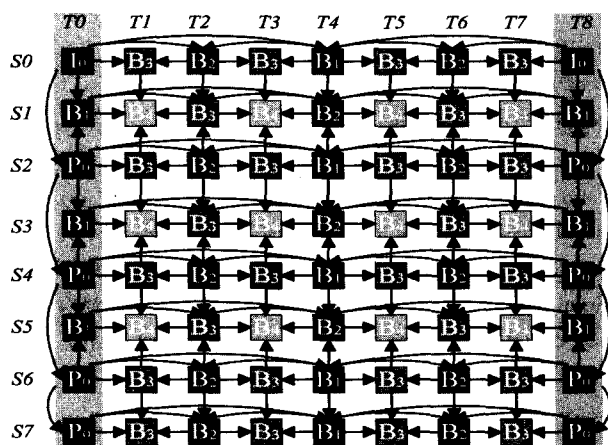


图 1 多视点视频编码结构

2 码率控制算法

码率控制算法主要包括三部分:(1)分配目标比特数;(2)根据目标比特数执行率失真函数计算量化参数(Quantization Parameter, QP);(3)更新率失真模型的各个参数。为了得到精确的目标比特,文中将目标比特的分配细分为三层:GGOP 层、GOP 层和帧层。其中,GGOP 层的比特分配采用均分的思想;在 GOP 层将介绍一种基于拉格朗日优化算法的比特分配模型;帧层目标比特则由时域分级因子和剩余比特共同决定。

2.1 GGOP 层比特分配

GGOP 的目标比特数由当前给定目标比特率、帧率及 GOP 的长度决定,比特分配公式为:

$$R_{\text{GGOP}}(i) = \frac{B_r}{F_r} \times N_{\text{view}} \times N_{\text{GOP}} - [B_0 - B_c(i-1)] \quad (1)$$

其中: B_r 表示目标比特率; F_r 表示帧率; N_{view} 表示编码视点数; N_{GOP} 表示 GOP 长度; B_0 表示缓冲区初始值; $B_c(i-1)$ 表示编码完第 $i-1$ 个 GGOP 后缓冲占有度。从(1)式可知,目标比特率越大,视点数越大, GOP 长度越长,帧率越小,分配到每一个 GGOP 的目标比特数就越大。

2.2 GOP 层比特分配

为了保持各视点编码质量的一致性,例如为了保持相邻两视点 (i, j) 编码图像质量一致,要求在编码后失真 $D(i)$ 和 $D(j)$ 必须尽可能接近,同时还要考虑到编码后输出码率 R 必须满足信道速率约束条件。而通常经量化后数据服从拉普拉斯分布,即率失真函数为:

$$R = \log_2 \left(\frac{\pi \sigma^2}{2eD} \right) \quad (2)$$

其中 σ 是当前信号的均方差。

这是一个典型的优化问题。其目标函数和受限条件分别如(3)式和(4)式所示。

目标函数为:

$$\min \left| \frac{D(i)}{D(j)} - 1 \right| \quad (3)$$

受限条件:

$$R_{\text{GOP}} \leq \frac{B_r}{F_r} \times N_{\text{GOP}} \quad (4)$$

对于上述优化问题,可以用拉格朗日乘法求解,首先引入一个辅助函数 L , 定义为:

$$L = \left| \frac{D(i)}{D(j)} - 1 \right| + \lambda_i \left(R_{\text{GOP}}(i) - \frac{B_r}{F_r} \times N_{\text{GOP}} \right) \quad (5)$$

其中 λ_i 为拉格朗日乘子,将(2)式率失真函数代入辅助函数 L 中,得:

$$L = \left| \frac{\sigma_i^2}{\sigma_j^2} 2^{-\left[R_{\text{cor}}(i) - R_{\text{cor}}(j) \right]} - 1 \right| + \lambda_i \left[R_{\text{GOP}}(i) - \frac{B_r}{F_r} \times N_{\text{GOP}} \right] \quad (6)$$

其中 i 为当前编码 GOP 所在视点, j 为其前向参考视点。接下来使用求极值法求解上述问题的最优解 $R_{\text{GOP}}(i)$, 即对式(6)中的 $R_{\text{GOP}}(i)$ 求偏导,然后置为零,得:

$$R_{\text{GOP}}(i) = R_{\text{GOP}}(j) + \log_2 \frac{\sigma_i^2}{\sigma_j^2} - \log_2 \frac{\lambda_i}{\ln 2} \quad (7)$$

当缓存 buffer 中比特数较多时,那么就应该适时减少编码输出比特数;相反,缓存 buffer 中比特数较少时,那么就应该适时增加编码输出比特数,即输出比特数与缓存 buffer 中比特数成线性关系^[7],因此(7)式中的拉格朗日乘子 λ 可定义为:

$$\lambda_k = \alpha \cdot \frac{B_{\text{size}} - B_k}{B_{\text{size}} + B_k} \quad (8)$$

最后,将(8)式中的 λ 代入(7)式中,并将参考视点编码后的实际码率 $A_{\text{GOP}}(j)$ 替代 $R_{\text{GOP}}(j)$;整理后便得到基于拉格朗日乘子法比特分配模型:

$$R_{\text{GOP}}(i, k) = A_{\text{GOP}}(i, l) + \log_2 \frac{\sigma_k^2}{\sigma_l^2} + K(i) \cdot \log_2 \frac{B_{\text{size}} - B_k}{B_{\text{size}} + B_k} \quad (9)$$

其中, $K(i)$ 为模型参数。(9)式中的第一项表示参考视点的实际比特数。第二项是根据视频内容复杂度所做的调整量,如果当前编码视点的内容复杂度大于参考视点的内容复杂度,则调整量为正数,否则为负数。第三项是根据当前缓存的充盈度所做的调整量。第一个编码视点的 GOP 层比特数等于 CGOP 层比特数均分给每个视点的比特数。

2.3 Frame 层比特分配

上述多视点视频编码中在时间上采用的是分层 B 帧去除冗余,而不同时间分层 B 帧的 R-D 特性是不同的,例如:与较低时间分层 B 帧相比,要获得相同编码质量,较高时间分层 B 帧需要的编码比特数较少,根据这个特性,先计算每个时间分层的权重因子,然后由时间分层权重因子来估计帧层目标比特数,权重因子的定义如式(11)所示。时间分层权重因子由图像编码复杂度决定^[8]。首先介绍帧编码复杂度、时间分层的编码复杂度和 GOP 的编码复杂度的计算方法,分别如下:

$$\begin{cases} X_F = QP \times \text{Bits} \\ X_{TL}(k) = \sum_{i=0}^{N_n(k)-1} X_F(k, i) \\ X_{\text{GOP}} = N_I X_F^I + N_P X_F^P + \sum_k N_B(k) X_F^B(k) \end{cases} \quad (10)$$

其中 N_I 、 N_P 和 $N_B(k)$ 分别为一个 GOP 中 I 帧数、P 帧数和第 k 层 B 帧数, X_F 、 X_{TL} 和 X_{GOP} 分别为帧编码复杂度,时间分层编码复杂度和 GOP 编码复杂度。由时间分层编码复杂度和 GOP 编码复杂度便可推导出时间分层权重因子:

$$\begin{cases} w_{\text{Key}} = \frac{X_{TL}(0)}{X_{\text{GOP}}} \\ w_B(k) = \frac{X_{TL}(k)}{X_{\text{GOP}} - \sum_{i=0}^{k-1} X_{TL}(i)} \end{cases} \quad (11)$$

其中 w_{Key} 、 $w_B(k)$ 分别为关键帧的权重因子和第 k 级 B 帧的权重因子。

帧层目标比特便根据剩余比特数、每个分层剩余 B 帧数和时间分层权重因子计算,如(12)式所示:

$$\begin{cases} R_F^{\text{Key}} = w_{\text{Key}} R_{\text{GOP}} \\ R_F^B(k, i) = \frac{w_B(k) R_{\text{GOP}}(i)}{N_B(k, i)} \end{cases} \quad (12)$$

其中 $R_{\text{GOP}}(i)$ 为编码完第 i-1 帧后剩余比特数, $N_B(k, i)$ 为编码完第 i-1 帧后第 k 层剩余 B 帧数。

2.4 率失真函数

目前已经有许多学者,针对不同的编码结构提出了很多率失真模型,如二次率失真模型^[9]、线性率失真模型^[10,11]以及很多改进的率失真模型^[12-14]。为了得到适合上述多视点视频编码结构的率失真模型,对多个多视点视频序列进行测试,发现不同类型的帧或不同时间分层率失真统计特性有很大的区别。因此文中对不同的帧类型或不同时间分层运用不同的率失真模型,具体策略为:I 帧采用二次率失真模型,B 帧采用线性的率失真模型。其中不同层次 B 帧采用不同的率失真模型;而最底层 B 帧的量化参数则简单地由次底层插值得到^[15]。

编码后更新率失真模型中的模型参数,I 帧的模型参数在相邻若干个 GOP 的各自帧类型范围内更新;B 帧模型参数除了在相邻 GOP 间的相应层次的帧间更新外,还需要在本层内帧间更新。

3 实验结果

在多视点视频编码参考软件 JMVC3.0 基础上,实现了文中基于视点质量一致性优化的码率控制算法,并与 JMVC 固定量化参数控制策略进行了对比测试实验,本实验中采用两种多视点视频序列,包括内容复杂度较高的 Ballroom (640x480) 序列和运动复杂度较高的 Race1 (640x480) 序列。

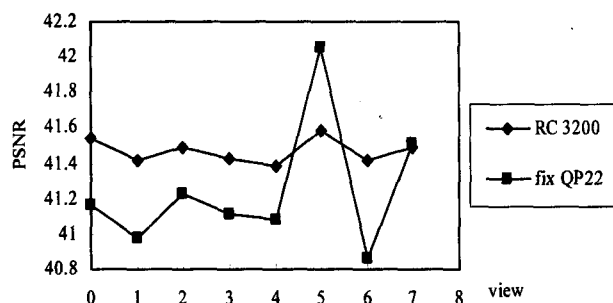
表 1 为文中所提码率控制算法和固定量化参数策略的仿真结果和比较。表中给出了测试序列采用上述两种不同策略编码后的平均 PSNR、实际输出码率 (Bit rate) 及其与给定码率的误差 (Inaccuracy)。由表可知,与固定量化参数策略相比,采用文中提出的码率控制后,Ballroom 和 Race1 序列的平均 PSNR 均有所提高,尤其是在低比特率下。与此同时,文中的码率控制算法有效地控制了输出码率,实际输出码率与给定码率的误差不超过 0.8%。

表 1 仿真结果和比较

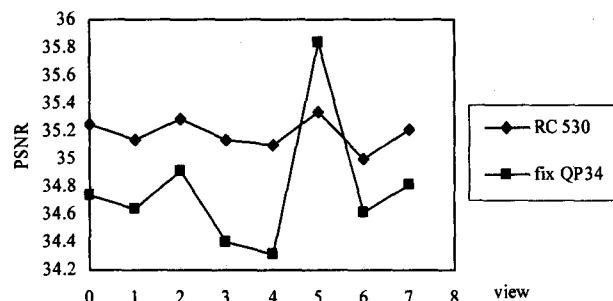
序列	固定量化参数策略			文中码率控制算法			Δ PSNR (dB)
	QP	Bit rate (kbps)	PSNR (dB)	Bit rate (kbps)	PSNR (dB)	Inaccuracy (%)	
Ballroom	22	3198.3	41.249	3192.0	41.472	0.197	0.223
	34	531.5	34.780	527.4	35.183	0.771	0.403
Race1	22	3203.2	42.341	3199.5	42.648	0.116	0.306
	34	560.3	35.405	558.6	35.852	0.303	0.447

图 2 和图 3 为 Ballroom 序列和 Race1 序列在不同码率下编码后每个视点的 PSNR。图中横坐标为八个视点,纵坐标为每个视点的平均 PSNR。由图可知,采用固定量化参数编码后视点质量波动较大,而采用文

中的码率控制算法编码后视点质量较一致。

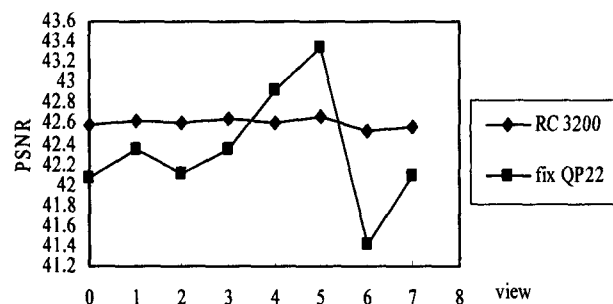


(a) ballroom 序列 (平均码率为 3200kps)

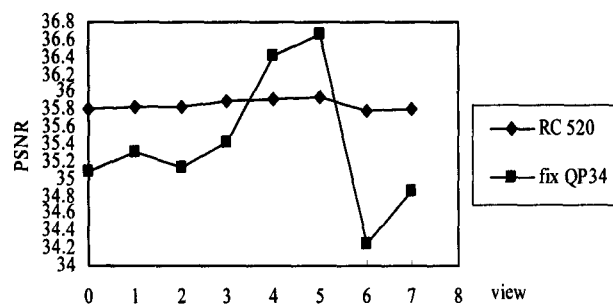


(b) ballroom 序列 (平均码率为 530kps)

图 2 Ballroom 序列在不同码率下编码后每个视点的 PSNR



(a) race1 序列 (平均码率为 3200kps)



(b) race1 序列 (平均码率为 520kps)

图 3 Race1 序列在不同码率下编码后每个视点的 PSNR

4 结束语

针对多视点视频编码,文中提出了基于视点质量一致性优化的多视点视频码率控制算法,其中视点质量一致性优化问题的关键在于视点间的比特分配算法。文中利用拉格朗日乘子法,并根据当前编码视点的前向参考视点实际输出比特数、当前编码视点图像

复杂度和当前缓存充盈度,推导出视点间的比特分配模型。实验结果验证了该码率控制算法不仅使得实际输出码率更接近给定目标码率,而且使得各视点编码后的图像质量更为一致。

参考文献:

- [1] Tanimoto M. Overview of Free Viewpoint Television[J]. Signal Processing: Image Communication, 2006, 21(6): 454-461.
- [2] 霍俊彦,常义林,李明,等. 多视点视频编码的研究现状及其展望[J]. 通信学报, 2010, 31(5): 113-121.
- [3] Joint Video Team of ITU-T VCEG and ISO/IEC MPEG. WD 1 Reference software for MVC (JMVC4.0) [S]. Geneva: [s. n.], 2008.
- [4] 毕厚杰. 新一代视频压缩编码标准-H. 264/AVC [M]. 北京: 人民邮电出版社, 2005: 14-22.
- [5] Li Xiaoming, Zhao Debin, Ma Siwei, et al. Fast Disparity and Motion Estimation Based on Correlations for Multiview Video Coding[J]. IEEE Transactions on Consumer Electronics, 2008, 54(4): 2037-2044.
- [6] Ma Lini, Pan Feng. Efficient Compression of Multi-view Video Using Hierarchical B Pictures [C] // International Conference on Multimedia and Ubiquitous Engineering. [s. l.]: [s. n.], 2008: 118-121.
- [7] Choi J, Park D. A Stable Feedback Control of the Buffer State Using the Controlled Lagrange Multiplier Method [J]. IEEE Transaction on Image Processing, 1994, 3(9): 546-557.
- [8] Xu L, Gao W, Ji X, et al. Rate control for hierarchical B-picture coding with scaling-factor [C] // IEEE International Symposium on Circuit and System. [s. l.]: [s. n.], 2007: 49-52.
- [9] Lee H J, Chiang T, Zhang Y Q. Scalable Rate Control for MPEG-4 Video [J]. IEEE Transaction on Circuits System for Video Technology, 2009, 10(9): 878-894.
- [10] 黄晓平,杨黎东,种衍文,等. 一种基于 ρ 域的新码率模型 [J]. 计算机工程与应用, 2010(3): 154-158.
- [11] He Z, Mitra S K. A linear source model and a unified rate control algorithm for DCT video coding [J]. IEEE Trans. on Circuits and System for Video Technology, 2002, 12(11): 970-982.
- [12] 王科技,康志伟,刘新焕,等. 基于 Cauchy 分布的 H. 264/AVC 码率控制优化算法 [J]. 计算机应用, 2010(9): 2444-2448.
- [13] Zhu Zhongjie, Liang Feng, Jiang Ganyi, et al. Bit allocation and rate control algorithm for stereo video coding [J]. Journal on Communication, 2007, 28(7): 15-21.
- [14] 刘光,房胜,尉希娟. 可伸缩视频编码的码率控制技术 [J]. 计算机技术与发展, 2010, 20(10): 26-29.
- [15] 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(7): 1175-1184.