

视频传输中的码率控制技术研究进展

毕迎春¹⁾ 王相海^{1),2)}

¹⁾ (辽宁师范大学计算机与信息技术学院, 大连 116029) ²⁾ (南京大学 计算机软件新技术国家重点实验室, 南京 210093)

摘 要 码率控制技术对提高和平滑解码视频质量起着至关重要的作用。近年来各种视频编码标准和应用都对码率控制技术进行了大量的研究,如基于 H. 261 和 H. 263 的视频会议,基于 MPEG-1 和 MPEG-2 的媒体存储,以及最新的基于 MPEG-4 视频对象编码。首先对码率控制问题进行了基本描述,然后介绍了标准中的经典码率控制技术,接着将目前的码率控制技术归纳为 3 类:基于 DCT 视频编码器的码率控制技术、基于小波视频编码器的码率控制技术和精细粒度可分级视频编码的码率分配算法,并对这三类码率控制方法的发展进行了概括性阐述与比较性研究,讨论了各种算法的基本思想及其优缺点,最后对视频传输中的码率控制技术的未来发展趋势进行了展望。

关键词 视频传输 率失真 码率控制 缓冲器

中图法分类号:TP391 文献标识码:A 文章编号:1006-8961(2008)12-2251-09

Research Progress on Rate Control Technology in Video Transmission

BI Ying-chun¹⁾, WANG Xiang-hai^{1),2)}

¹⁾ (College of Computer and Information Technology, Liaoning Normal University, Dalian 116029)

²⁾ (State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

Abstract Rate control scheme plays an important role in improving and smoothing the decoding video quality. Over the years, rate control techniques have been studied very intensively for various standards and applications, such as video conferencing with H. 261 and H. 263, storage media with MPEG-1 and MPEG-2, and the recent video object coding with MPEG-4. This paper firstly described the basis of rate control problem. And secondly classic rate control technology in the standard was introduced. Thirdly we summed up the present techniques as three kinds: rate control technique based on DCT video coder, rate control technique based on wavelet video coder and rate allocation of Fine Granular Scalable video coding. Then we generally discussed and compared the three kinds of methods including the core idea and the advantages and disadvantages of every algorithm of each kind of techniques. Finally, the state of art and possible new directions of rate control were stated.

Keywords video transmission, rate-distortion, rate control, buffer

1 引言

随着视频编码技术在媒体存储、视频点播、监控系统等诸多领域中越来越多的广泛应用,一系列视频编码标准被提出并应用于实际,像 ISO/IEC 的

MPEG-1/2/4, ITU-T 的 H. 261, H. 263, H. 263 + + 以及 MPEG-4 AVC/H. 264 等,这些标准对码流的语法结构进行了规定,而码率控制的具体算法则大多是开放的,码率控制算法的性能直接影响视频通信的质量^[1],作为提高视频编码图像质量的一项关键技术,码率控制技术的发展已引起人们的广泛关注。

基金项目:国家自然科学基金项目(60372071);辽宁省自然科学基金项目(20032105, 20072156);辽宁省高等学校优秀人才支持计划(RC-04-11);辽宁省教育厅科学技术研究项目(20060486);南京邮电学院图像处理与图像通信江苏省重点实验室开放基金项目(ZK207008)

收稿日期:2007-02-05; **改回日期:**2007-10-23

第一作者简介:毕迎春(1982 ~),女。辽宁师范大学计算机软件与理论专业硕士研究生。主要研究方向为多媒体信息处理、算法分析与设计。E-mail: bych888@163.com

码率控制的作用是保证编码比特的合理分配,以适应网络传输或媒体存储等应用的需求,按照输出比特率是否恒定,码率控制算法可分为 CBR (constant bit rate) 和 VBR (variable bit rate) 两种,当在 CBR 信道下传输压缩码流时,由于压缩后的码流为可变比特流,所以编解码端都要有缓冲器对码流进行“平滑”,编码器端的码率控制算法就是用来确保缓冲器不出现上溢或下溢的。即使在缓冲器可以无限大时,码率控制技术仍然是必须的,因为算法要满足端到端 (end-to-end) 的延迟约束;此外,对于在离线状态下对数据进行编码的存储视频应用中 (如 DVD 或 CD-ROM), 仍需有效控制比特率以防止在实时播放时解码端的缓冲器出现溢出。设计码率控制算法除了能够避免溢出,还能够提供每帧图像更好的视频质量,而且可以避免视频质量的过大波动。这里主要是针对在 CBR 信道下传输压缩码流时的码率控制技术现状的研究,这些码率控制算法经过简单的修改同样适用于 VBR 的传输情况。

2 码率控制问题基本描述

码率控制是视频编码的重要组成部分。一般来讲,它的实现过程包括比特分配、量化系数计算和缓冲器控制等部分^[2],码率控制问题的基本框架如图 1 所示。比特流经常要通过有限带宽的通信信道进行传输,而编码器会产生可变码流,所以有必要在编码器和信道之间设置视频缓冲器来平滑编码过程中的比特流波动,避免缓冲器出现上溢或下溢,设置合适的码率控制算法是必需的。

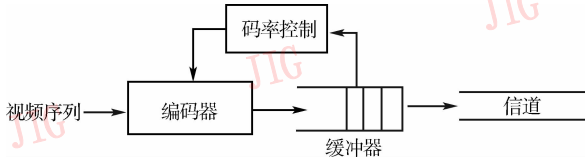


图 1 码率控制问题的基本框架

Fig. 1 Basic framework of rate control problem

码率控制问题的解决总体而言可以归结为对其进行公式化的过程,通常有两种不同的失真准则可以用来度量公式化的码率控制问题^[3],一是最小化 GOP (group of picture) 内平均失真;二是最小化 GOP 内的失真变化。一个 GOP 内共有 N 帧, $\mathbf{q} = (q_1, q_2, \dots, q_N)^T$ 为量化选择,其中 q_i 表示第 i 帧的量化选择;第 i 帧的码率和失真函数分别表示为 $r_i(\mathbf{q})$ 与

$d_i(\mathbf{q})$ 。在给定量化的选择 $\mathbf{q}$ 的情况下,第 i 帧的码率函数 $r_i(\mathbf{q})$ 可以通过计算该帧总的比特数得到,而 $d_i(\mathbf{q})$ 则可以通过计算该帧的均方误差 (MSE) 获得。当该 GOP 已经用量化选择 $\mathbf{q}$ 编码完毕时,在第 i 帧时间间隔内缓冲器占用度表示为

$$\begin{cases} b(i, \mathbf{q}) = \max\left(b(i-1, \mathbf{q}) + r_i(\mathbf{q}) - \frac{R}{F}, 0\right); (i = 1, \dots, N) \\ b(0, \mathbf{q}) = 0 \end{cases}$$

式中, R 为信道码率 (bps), F 为帧率 (fps), $b(0, \mathbf{q})$ 为第 1 帧编码前缓冲器的占用情况,这里使用 $\max$ 函数是为了避免缓冲器出现下溢的情况。采用两种不同失真准则的码率控制问题公式化过程如下:

最小化平均失真 令 $\mathbf{Q} = \{1, 2, \dots, 31\}$ 为可行的量化参数, $b_{\max}$ 为缓冲器最大容量,找出 $\mathbf{q}^* = (q_1^*, q_2^*, \dots, q_N^*)^T$, 其中 $q_i^* \in \mathbf{Q}, i = 1, 2, \dots, N$, 使得

$$\mathbf{q}^* = \arg \min_{\mathbf{q} \in \mathbf{Q}^N} \frac{1}{N} \sum_{i=1}^N d_i(\mathbf{q}) \quad (1)$$

并且满足 $b(i, \mathbf{q}) \leq b_{\max}, i = 1, 2, \dots, N-1, b(N, \mathbf{q}) = 0$, 其中约束 $b(N, \mathbf{q}) = 0$ 强制缓冲器最终的占用度为 0, 以保证每个 GOP 能分配到恒定的比特数,它能够保证快速搜索存储在 CD-ROM 或硬盘上的视频流。当然,如果不要每个 GOP 的比特数相同,那么该约束就可以去掉,以便能够更好地利用缓冲器来提高视频质量。

最小化失真变化 对于每个可行的 $q_i^* \in \mathbf{Q}$ 找到 $\mathbf{q}^*$ 使得

$$\mathbf{q}^* = \arg \min_{\mathbf{q} \in \mathbf{Q}^{N-1}} \sum_{i=2}^N |d_i(q_i) - d_{i-1}(q_{i-1}^*)| \quad (2)$$

满足 $b(i, \mathbf{q}) \leq b_{\max}, i = 1, 2, \dots, N-1$ 。

3 视频编码标准校验模型中的码率控制技术

MPEG-1 的校验模型 SM3 (simulation model 3)^[4] 中提供了一种简单的码率控制方法,其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“充满”程度来确定量化参数 $\mathbf{q}$, 进行码率调整。SM3 基本达到了使编码器输出码流的数率与信道数率相匹配的目的。

TM5 (test model 5)^[5] 码率控制算法是被 MPEG-2 标准采纳的代表性的算法,它的目标比特分配方案考虑了图像内容的复杂程度,在进行码率调整时设置了 3 个虚拟缓冲器,分别用于一种类型图像 (I, P 或 B 帧) 的编码输出,此外,它还考

虑了视觉的空间掩蔽效应,其宏块的量化参数值由码率调整时得到的参考量化参数和图像宏块的活跃程度(由亮度变化的剧烈程度)两个因素最后确定。该算法考虑了人的视觉特性,能在限定的码率下使重建图像具有较高的主观评价质量,但该算法仍存在一些不足,比如不能很好地处理场景切换带来的问题,同一图像中宏块与宏块之间的编码质量不均匀,造成同一图像上图像质量不统一等等。已有不少文献针对以上问题对 TM5 算法进行了改进。

MPEG-4 采用的 VM8 (verification model 8)^[6] 码率控制算法是一个经典的码率控制方法,它使用了可分级码率控制机制 (SRC)。该算法的核心是一个模拟视频编码器 R-D (rate-distortion) 特性的二次函数模型,即 $R = X_1 S Q^{-1} + X_2 S Q^{-2}$, 其中, R 是编码比特数; X_1 、 X_2 为模型参数; S 为编码的复杂度,通常用一帧图像的平均绝对差值 (MAD) 来表示; Q 为量化参数。该算法包括帧级码率控制 (即单视频对象的码率控制)、多视频对象 (M-VOs) 的码率控制和宏块级码率控制。单视频对象 (SVO) 的码率控制算法流程如图 2 所示。多视频对象的码率控制算法是在单视频对象码率控制基础上导出的,它使用了基于颜色容忍度的 HVS (human visual sensitivity of color tolerance) 比特分配表,这样可以对每个对象进行比特分配。SRC 方案可满足不同的码率变化范围 (10kbps ~ 1Mbps)、不同的空间分辨率 (qcif 或 cif)、不同的时间分辨率 (7.5fps ~ 30fps) 以及不同的编码器 (DCT 或 Wavelet) 的需求。

H. 263 标准中采用的 TMN8 (test model near-

term 8)^[7] 码率控制包括两个步骤,一是在帧层根据上一帧编码比特数、编码码率、目标帧率、原始帧率、缓冲器延时和跳帧阈值计算目标比特数;另一个是在宏块层根据 R-D 对数模型进行自适应量化。TMN8 可更精确地控制目标码率,保持缓冲器的稳定性。但 TMN8 没有调整每个宏块的量化参数动态范围,且 TMN8 的码率控制只对 INTER 帧进行,对 INTRA 帧,量化参数设为固定值。TMN11 (test model near-term 11)^[8] 是 H. 263 标准中的另一个经典码率控制算法,该算法在 TMN8 帧级码率控制算法的基础上充分考虑了视频内容和网络带宽大小,能够自适应地调节帧率和宏块级的量化参数,是一种支持 VBR 信道,能够更加有效地在网络上传输压缩视频数据和提高图像的主观质量的码率控制算法,但其计算复杂度较大。

最新的视频编码标准 H. 264 采用的是 JVT-H017 提案中的码率控制算法^[9]。该算法采用了基于分层的码率控制方法,共分 3 个层次:GOP 层码率控制,帧层码率控制和基本单元层的码率控制。GOP 层同 TM5 一样采用固定的目标比特数;帧层的目标比特数由网络带宽、缓存器占用度、缓存器总容量以及剩余比特数共同决定;基本单元层的目标比特数是基于预测的 MAD 值进行分配的。由于在 H. 264 编码算法中采用 RDO (rate distortion optimization) 进行模式选择,只有在 RDO 操作后才能计算出实际的 MAD 值,导致使用 MAD 值预测量化参数方法出现蛋鸡悖论^[10]。目前已有不少学者针对该问题进行深入研究并提出了解决方案^[11-13]。JVT-H017 算法中采用线性模型预测 MAD 值,每处理完一个基本单元要对模型的参数进行更新操作,运算简单但开销较大。

4 基于 DCT 视频编码器的码率控制技术

目前大部分的码率控制技术都是适用于基于 DCT 视频编码器的,而码率控制中的关键问题是估计或模型化视频编码器的码率-失真 (R-D) 行为^[14]。采用率失真技术的目的是防止缓冲器溢出并最大化视频质量。基于 DCT 视频编码器的 R-D 行为由它的码率-量化 (R-Q) 与失真-量化 (D-Q) 函数来描述,R-Q 与 D-Q 函数统称为 R-D 函数。下面根据码率失真函数自变量的不同,将基于 DCT 视频编码器的码率控制技术分为 q 域码率控制技术和 ρ 域码率控制技术分别进行介绍。

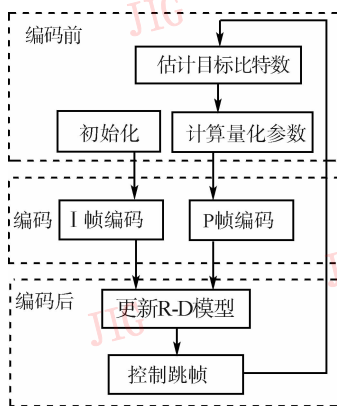


图 2 MPEG-4 SVO 码率控制算法

Fig. 2 MPEG-4 SVO rate control algorithm

4.1 q 域码率控制技术

R-D 模型都是试图找到编码比特率 R 和失真 D 与量化参数 q 相关的函数,对以 q 为自变量的码率和失真函数的研究称为 q 域 R-D 分析。传统的 R-D 分析是在 q 域上执行的。基于 DCT 视频编码器的输出比特率由它的量化参数来控制。基于 R-D 函数,可以通过选择相应的编码器量化参数以获得目标比特率或图像质量。

文献[15]中提出的码率控制算法使用动态规划的方法搜索真正的全局最优方案。该方法复杂度非常高,因为它要在所有量化步长下重复编码源数据以便计算编码器的率失真性能。为降低文献[15]中的计算复杂度,文献[16]中引入一个对数码率-量化函数,其模型参数的更新需要利用重新编码源视频数据所得到的编码统计。文献[17]中提出的指数码率-量化函数要比文献[16]中的更为精确,其相对误差要小一些。但是这两种算法的计算复杂度仍然很高。并且其模型的精确度都还不够,尤其对于低活动性视频序列中的 P 帧和 B 帧。文献[3]中提出的算法在没有明显降低最优性能的前提下,降低了计算 R-D 数据的复杂度。该文中提出了两种降低复杂度的算法:(1)使用模型内插码率与失真特性,用分段三次函数内插码率函数 $r(q)$,用分段线性函数内插失真函数 $d(q)$,并提出了帧间相关性模型;(2)使用过去帧代替当前帧来确定模型。这两种算法与 TM5 相比可以获得 1dB 以上的 PSNR 增益,在场景切换处得到的 PSNR 会更高。第 1 个算法中应用的分段内插函数与文献[16]、[17]中提出的率失真模型相比较,明显降低了误差。它的编码时间大约是 TM5 的 2 倍,但与精确测量 R-D 数据的方法相比复杂度要低得多。所以第 1 个算法适用于编码延迟较长的应用中(如广播视频等),而第 2 个算法的复杂度与 TM5 的复杂度相类似,其 PSNR 性能比第 1 个算法要低些,所以它比较适用于对复杂度或编码延迟有约束限制的交互式视频应用当中。

文献[2]通过对热传导方程的分析发现它与视频编码中的比特率波动的动态行为具有相似性,因而采用偏微分方程来描述视频编码中的比特率波动情况,提出一个新的码率传输方程(RTE),并将 RTE 的解作为 R-D 模型。该模型包括两个变量,即失真 D 和源数据统计特性 M 。与 VM8 相比较,该算法具有更高的控制精度并能获得更加稳定连续的图

片质量。

4.2 ρ 域码率控制技术

文献[18]、[19]发现,在图像和视频变换编码中,量化后的变换系数间零值所占的百分比 ρ 对编码比特率 R 具有关键的作用,尤其是在低比特率的情况下。如果假设变换系数的分布是连续正态分布的, ρ 将随量化参数 q 单调增加,这就意味着 q 和 ρ 具有一一对应的关系,因此, R 和 D 也可以看作是 ρ 的函数,分别记为 $R(\rho)$ 和 $D(\rho)$ 。将 R 和 D 作为 ρ 的函数的分析称为 ρ 域 R-D 分析。此外,文献[18]中引入特性码率曲线和码率曲线分解两个新概念,提出了一个通用的用于视频变换编码的源模型框架。 ρ 域实际的码率曲线 $R(\rho)$ 为两个特性码率曲线的线性结合。利用精确的 $R(\rho)$ 函数,确定出相应的 ρ 值来满足目标比特率,再根据 q 和 ρ 的一对一映射计算出量化值 q 。文献[19]在文献[18]提出的 R-D 函数的基础上,提出了一个基于编码器的码率形状平滑算法。算法使目标图像质量 DT 在较小的变化范围内的变动(如 $[DT-1, DT+1]$),通过调整前后两帧的码率使其相近。利用该平滑算法,编码器的输出比特流既有平滑的码率形状又有着连续的图像质量。

文献[20]提出一个新的线性码率模型, $R(\rho) = \theta(1-\rho)$, 其中 θ 为一个常数。实验结果表明,该模型适用于各种源数据类型,如 I, P, B 帧,可应用在所有视频编码系统中,如 MPEG-2, H. 263 和 MPEG-4 等。文献[21]根据 shannon 的源编码理论,从理论上证明了该线性源模型的正确性,并讨论了模型参数的物理意义。文献[18]~[21]只考虑了码率模型,而没有考虑失真模型,基于此文献[22]提出了一个 ρ 域的失真模型, $D(\rho) = \sigma^2 e^{-\alpha(1-\rho)}$, 其中 α 为常数, σ^2 为图像的方差,结合文献[20]中的线性码率模型应用拉格朗日最优化方法,该文提出了一个最优比特分配机制并将其应用在 MPEG-4 的视频编码中,在不同的视频对象中分配可行的比特数,实验结果表明提出的算法能够精确地控制编码器比特率并明显地提高了视频质量(大约在 1.5dB 以上)。文献[23]中通过实验验证分析得出文献[22]中的 ρ 域码率模型是精确的,而失真模型则不够精确,尤其对于 B 帧,使用该模型的帧级比特分配算法存在着稳定性问题,偶尔会达不到目标比特率。该文中提出一个新的在 ρ 域下的失真模型 $D = \kappa \sigma^2 e^{-2\theta(1-\rho)}$ (其中 κ 为大于零的模型参数),并提出估计模型参数的方法,进而提出一种鲁棒的码率控制算法。实

验结果表明其文中提出的算法缓和了文献[22]中比特分配算法的稳定性问题并能得到更加一致的视觉质量。此外,文献[24]使用基于变长码字表的码率估值方法,构造了线性码率模型,并将该码率模型分别应用于帧级和宏块级码率控制中,该模型比文献[19]中的方法简单,比文献[21]中的模型参数决策规则更加灵活,却可以与它们获得相似的 PSNR 增益。

5 基于小波视频编码器的码率控制技术

小波编码器具有非常优秀的编码性能,在静止图像压缩中已被广泛应用^[25~29]。小波编码器不仅能提供更好的图像主观质量,而且还具有嵌入式特性。嵌入式编码器能提供更好的率失真折衷,它可以在码流的任何点处截取编码比特流而没有明显的可感知的失真。对于嵌入式视频编码器的一个简单码率控制机制就是为每个 I 帧、P 帧和 B 帧分配固定的比特数,然而该简单机制并不能获得最优的 R-D 性能。文献[30]应用指数率失真模型,以最小化编码失真为目标构造了约束最优化问题,采用拉格朗日方法解决构造出的最优化问题,将得到的解作为码率分配模型。文献[31]在文献[30]的基础上考虑了参考帧与预测帧之间的相关性,得出运动补偿后残差帧的方差与参考帧的编码失真具有线性关系的结论,并结合此线性关系提出对嵌入式小波视频编码器的码率控制算法,将码率控制问题转化成对每帧进行比特分配的问题,利用嵌入式小波编码器的率失真性能,应用拉格朗日方法导出有效的码率控制机制。文献[32]在给定每个预测残差帧(PEF)目标码率的情况下,以空间方向树(SOT)为单位,对其进行率失真分析,用拉格朗日最优化方法在 PEF 内对系数进行比特分配,为每个 SOT 选择最优操作点。该方法提高了 CEZW(color embedded zerotree wavelet)的编码性能。文献[33]在分析了误差扩散和帧间依赖关系的基础上,提出了一种兼顾这两个问题的 EAMC(enhanced adaptive motion compensation)算法,该算法的基本思想是:在编码端进行运动补偿时使用的参考帧是一个中间码率 R_{ref} 下的重建帧,当解码端对截断后较高码率的码流进行解码时,通过反馈环路,将解码端的参考帧锁定在 R_{ref} 下,从而不但消除了误差扩散,而且提高了图像质量;当在较低码率下解码时,可以降低误

差扩散的速度。此外,文献[34]在小波分解图像的不同级上采用不同长度的滤波器以消减“振铃”效应,然后重新组织小波系数成小波块,根据每个小波块所含能量的多少分配不同的码率,利用文献[35]中给出的基于 HVS 特性的量化值,通过调整各块的量化因子,逼近各块的目标比特率。该算法解码可做到实时,并能提供基于内容的量化。

6 精细粒度可分级视频编码的码率分配算法

由于网络视频应用的飞速发展和网络带宽的时变特性,快速的码率分配和比特流截断技术对于视频流传输十分重要。以 DCT 为基础的精细的可伸缩性(FGS)视频编码^[36]已经被 MPEG-4 作为流式应用的标准而采纳,FGS 方案能够灵活地适应网络带宽的波动。利用 FGS 编码可产生一个可分级的比特流,其基本层旨在适应最低的网络带宽要求,而 FGS 增强层由于采用了位平面编码技术,可根据需要在任意比特位置截断,并且解码的视频质量与用于解码的比特数成正比。它把增强层可用的比特数平均地分给各帧^[37]。由于这种方法没有考虑各帧图像的率失真特性。使得接收端解码的视频质量存在很大波动。

为了提高 FGS 的编码效率,文献[38]提出了渐进的精细可伸缩性(PFGS)视频编码算法。它在增强层图像编码时使用了前一帧重建的某个增强层图像为参考进行运动补偿,使得运动补偿更有效,从而提高了编码效率。而如果同 FGS 一样使用均匀的码率分配算法,PFGS 相邻帧之间的解码质量会有严重的波动,整个视频序列的平均解码质量也会下降。文献[39]提出了一个简单有效的切分 PFGS 的增强层码流的方法,它建立了 PFGS 增强层的多帧联合的率失真函数,通过 MPEG-4 标准测试序列检验了它的正确性。然后利用联合率失真函数给出了局部最优的码率分配算法,并进行了简化使得它的复杂性极低,非常适合流化视频服务器使用。该算法不仅使得各帧图像的解码质量比较平衡,而且将 PFGS 的编码效率提高了将近 0.15dB。但该方法假设不同帧在相同位平面截取时具有相似的失真,而从 FGS 的率失真曲线看,不同帧即使在相同的位平面下仍然存在较大的失真变化。文献[40]提出了适用于 FGS 视频流的相等图像质量码率分配方法。该方法根据每帧的率失真关系将等质量码率分配问

题公式化为约束最优化问题,使用牛顿搜索方法寻找合适的失真值,确定出每帧增强层中截取比特流的位置。该算法能够确保相邻帧间图像波动小于 0.3dB,牛顿搜索方法的应用加速了码率分配过程,使其更加适用于视频流应用。

此外,文献[41]提出了一种基于视频序列 R-D 特性的 FGS 增强层的优化码率分配算法,以视频总体失真最小和质量波动最小为目标,建立了 FGS 增强层码率分配问题的最优化模型,把对这个问题的解作为码率分配的最优化方案。算法中采用一种线性内插模型来获得图像的 R-D 特性数据,降低了计算复杂度,仿真结果表明该方案在保证解码视频质量恒定和保持视频总体质量最优两方面均收到了良好的效果。

7 展 望

确定合适的编码参数以便在某固定信道码率下获得最优的解码视频质量是人们设计码率控制算法的根本宗旨。尽管已经出现了许多有效的码率控制方案,但人们对于视频编码图像质量的要求却越来越高,使得码率控制方法有待进一步发展和完善,我们认为未来码率控制技术将会在以下方面得到进一步的发展。

(1) 更加精确的率失真模型

码率控制中的关键问题是估计或模型化视频编码器的率失真模型,现有文献中已经提出一些率失真模型^[32-48],但这些模型通常假定信源服从高斯分布或拉普拉斯分布,当实际视频不满足假设条件时,模型的准确性就会受到影响,并进一步导致算法性能的下降。而且有些模型通常只适合某个固定的编码器,对于其他的编码器就不够精确,应用范围很小。所以有必要提出更加精确的率失真模型,使其适合多种视频编码器,并能准确地反映出实际的视频序列的率失真特性。

(2) 对缓冲器更加合理的控制策略

为了防止缓冲器出现上溢或下溢,大量文献中的出发点都是使每帧图像编码结束时缓冲器的占用度趋近缓冲器总容量的一半,MPEG-4 标准中的码率控制部分在对每帧进行目标比特分配时,使其不低于缓冲器容量的 10% 也不超过总容量的 90%,若当前缓冲器的占有度已达到总容量的 80%,则进行跳帧控制。但由于 I 帧产生的比特数要比 P 帧和 B

帧多几倍,将缓冲器占用度设为某一固定值并不是科学的做法,若能根据编码情况对缓冲器进行更加合理的设置,不仅可以有效处理缓冲器溢出问题,同时还能尽量避免跳帧现象的出现。

(3) 对场景切换的处理

在实时视频应用场合,视频序列的复杂度是不断变化的。为了适应视频序列中出现场景切换的情况,出现了动态调整 GOP 大小的方法^[49,50]和检测场景变换的方法^[51,52],对发生场景变换的图像进行特殊处理。但这些方法通常不够准确,并且计算复杂度也很大。因此能够提出更加准确的场景检测方法以及给发生场景变换的图像分配合适的比特数是项非常有意义的工作。

(4) 基于小波视频编码器的码率控制算法

小波编码器具有的优秀特性吸引了不少学者的目光:①提供更好的 R-D 折衷;②提供更令人满意的主观图像质量;③具有嵌入式特性,可以在任意一点截取比特流而没有明显的失真;④不用考虑调整量化系数,只需对各帧合理的分配比特数即可。对小波视频编码器的码率控制算法要比基于 DCT 视频编码器的码率控制方法相对简单一些,更便于调节。相对于 MPEG 的码率控制研究,针对小波视频编码器的码率控制研究数量还非常少,随着小波变换在视频编码和视频信息传输方面的应用,这方面的工作将成为一个研究热点。

(5) 基于视频对象的码率控制算法

自从针对视频对象的视频编码标准 MPEG-4 的提出,已有不少学者对基于对象的码率控制方法进行了研究,提出了一些有效的码率控制算法^[53-56],但这些算法中大部分都是延续单视频对象的方法,并不很精确。如何在视频对象、形状信息及运动向量间合理的分配比特数,对解码图像质量起着至关重要的作用,相信随着基于对象的视频信息的广泛应用,基于视频对象的码率控制算法一定会得到很好的发展。

(6) 面向精细粒度可分级视频编码的码率分配算法

视频编码的最初目标是为了在给定的码率下达到最优的解码质量,然而由于近几年 Internet 上视频服务的增加,使得视频编码的目标除了追求最佳的视频质量外还越来越重视视频编码的可伸缩性。目前已出现了许多有效的视频精细粒度可分级编码方案^[57-60],大量学者都在致力于完善并发展细粒度

可分级编码技术。如何设计适合各种精细粒度可分级视频编码方案的码率控制算法,如何在基本层与增强层间划分比特数,如何在增强层合理地分配比特数,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有待解决的具有重要研究价值的问题。

(7) 低比特率下的实际通信应用的码率控制算法

在通信网络上设计多媒体应用的主要挑战是如何传输最小的多媒体流给用户。视频会议、在线点播等实际通信应用都要求低码率下具有低延迟及低复杂度的码率控制技术。基于拉格朗日最优化的方法^[61~65]已经存在于大量的文献中,但这些算法的复杂度一般都很高。简化这些算法的复杂度,使其适应实时低比特率下的实时通信应用,具有很高的理论和商业价值。

此外,一些文献中的码率控制方法是基于图像内容及人眼视觉特性的^[66,67],这些方法相对率失真的方案要简单,算法易于实现,但并不是很准确,还有待继续改进。

参考文献 (References)

- 1 Zhao Bo, Wu Cheng-ke. A novel rate control algorithm in video coding for low delay communications [J]. Chinese Journal of Computers, 2005, **28**(1): 53 ~ 59. [赵波, 吴成柯. 一种新的低时延视频编码码速率控制算法[J]. 计算机学报, 2005, **28**(1): 53 ~ 59.]
- 2 Zhang Zhong-wei, Liu Gui-zhong, Li Hong-liang, *et al.* A novel PDE-based rate-distortion model for rate control [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, **15**(11): 1354 ~ 1364.
- 3 Lin L J, Ortega A. Bit-rate control using piecewise approximated rate-distortion characteristics [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, **8**(4): 446 ~ 459.
- 4 ISO/IEC JTC1/SC29/WG11, MPEG-1 Simulation Model 3 (SM3) [S].
- 5 ISO/IEC JTC1/SC29/WG11/93-225b, MPEG-2, TestModel5 (TM5) [S].
- 6 ISO/IEC JTC1/SC29/WG11, MPEG-4 Video Verification Model v8.0 Coding of Moving Pictures and Associated Audio MPEG 97/N1796 [S].
- 7 ITU-T/SG15, Video Codec Test Model, Near-Term, TMN8 [S].
- 8 ITU-T TMN11, Video Codec Test Model, Near-Term [S].
- 9 Ma Si-wei, Li Zheng-guo, Wu Feng. Proposed draft of adaptive rate control [A]. In: Proceedings of JVT-H017 8th Meeting [C], Geneva, Switzerland, 2003: 20 ~ 26.
- 10 Li Zheng-guo, Pan Feng, Pang Keng, *et al.* Adaptive basic unit layer rate control for JVT [A]. In: Proceedings of JVT-G012, 7th Meeting [C], Pattava, Thailand, 2003: 7 ~ 14.
- 11 He Zhi-hai. A Unified Approach to Rate-distortion Analysis and Rate Control for Visual Coding and Communication [D]. Santa Barbara, CA: University of California Santa Barbara. 2001.
- 12 JVT of ISO/IEC, MPEG & ITU-T VCEG Document JVT-F086 [S].
- 13 Ma Si-wei, Gao Wen, Gao Peng, *et al.* Rate control for advance video coding (AVC) standard [J]. IEEE International Symposium on Circuits and Systems, 2003, **12**(5): 92-895.
- 14 Haskell B G, Puri A, Netravali A N. Digital Video: An Introduction to MPEG-2 [M]. London, UK: Chapman & Hall, 1997.
- 15 Ortega A, Ramchandran K, Vetterli M. Optimal trellis-based buffered compression and fast approximations [J]. IEEE Transactions on Image Processing, 1994, **3**(1): 26 ~ 40.
- 16 Hang H M, Chen J J. Source model for transform video coder and its application I: Fundamental Theory [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, **7**(2): 287 ~ 298.
- 17 Ding W, Liu B. Rate control of MPEG video coding and recording by rate-quantization modeling [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, **6**(1): 12 ~ 20.
- 18 He Z, Kim Y, Mitra S K. ρ -Domain source modeling and rate control for video coding and transmission [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Salt Lake City, Utah, 2001: 1773 ~ 1776.
- 19 He Z, Kim Y, Mitra S K. Low-delay rate control for DCT video coding via ρ -Domain source modeling [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, **11**(8): 928 ~ 940.
- 20 He Z, Kim Y, Mitra S K. A novel linear source model and a unified rate control algorithm for H. 263/ MPEG-2 /MPEG-4 [EB/OL]. <http://videonet.ece.missouri.edu/papers/conference/icassp012.pdf>, 2001.
- 21 He Z, Mitra S K. A linear source model and a unified rate control algorithm for DCT video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, **12**(11): 970 ~ 982.
- 22 He Z, Mitra S K. Optimum bit allocation and accurate rate control algorithm for video coding via ρ -Domain source modeling [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, **12**(10): 840 ~ 849.
- 23 Nejat Kamaci, Yucel Altunbasak. ρ -Domain rate-distortion optimal rate control for DCT-based video coders [EB/OL]. http://www.ece.gatech.edu/research/labs/MCCL/pubs/dwnlds/kamaci_altunbasak_ICASSP04.pdf. 2004.
- 24 Lee J Y, Park H W. A rate control algorithm for DCT-based video coding using simple rate estimation and linear source model [J]. IEEE Transactions on Signal Processing, 2005, **15**(9): 1077 ~ 1086.
- 25 Shapiro J. Embedded image coding using zerotrees of wavelet coefficients [J]. IEEE Transactions on Signal Processing, 1993, **41**(12): 3445 ~ 3462.

- 26 Taubman D, Zakhor A. Multirate 3-D subband coding of video[J]. IEEE Transactions on Image Processing, 1994, **3**(9): 572 ~ 588.
- 27 Said A, Pearlman W A. A new fast and efficient image codec based on set partitioning in hierarchical trees[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, **6**(3): 243 ~ 250.
- 28 Ramchandran K, Xiong Z, Asai K, *et al.* Adaptive transforms for image coding using spatially-varying wavelet packets [J]. IEEE Transactions on Image Processing, 1996, **5**(7): 1197 ~ 1203.
- 29 Li J, Cheng P Y, Jay Kuo C C. On the improvements of embedded zerotree wavelet (EZW) coding [A]. In Proceedings of SPIE: Visual Communication and Image Processing'95 [C], Taipei, Taiwan, 1995, **2501**: 1490 ~ 1501.
- 30 Cheng P Y, Li J, Jay Kuo C C. Rate-distortion optimized bit rate control scheme for a wavelet video coder [EB/OL]. http://research.microsoft.com/users/jinl/paper_1997/vcip97_rate.PDF, 1997.
- 31 Cheng P Y, Li J, Jay Kuo C C. Rate control for an Embedded Wavelet Video Coder [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, **7**(4): 696 ~ 702.
- 32 Asbun E, Salama P, Delp E J. A rate-distortion approach to wavelet-based encoding of predictive error frames [A]. In: Proceedings of the 2000 IEEE International Conference on Image Processing [C], Vancouver, British Columbia, Canada, 2000: 10 ~ 13.
- 33 Liu Hong-mei, Xiao Zi-mei, Liang Fan, *et al.* Research on rate scalable wavelet video coding algorithm [J]. Journal of Software, 2002, **13**(4): 664 ~ 668. [刘红梅, 肖自美, 梁凡等. 码率可分级小波视频编码算法的研究 [J]. 软件学报, 2002, **13**(4): 664 ~ 668.]
- 34 Liu Jiu-fen, Huang Da-ren. A rate control method based on wavelet transform[J]. Journal of Zhejiang University (Sciences Edition), 2001, **28**(01): 14 ~ 18. [刘九芬, 黄达人. 一种基于 HVS 的码率控制方法[J]. 浙江大学学报(理学版), 2001, **28**(01): 14 ~ 18.]
- 35 Watson A B, Yang G Y, Solomon J A, *et al.* Visibility of wavelet quantization noise [J]. IEEE Transactions on Image Processing, 1997, **6**(8): 1168 ~ 1175.
- 36 ISO/IEC 14496-2/PDAM4. Coding of Audio-Visual Objects, Part-2 Visual, Amendment 4; Streaming Video Profile[S].
- 37 Li Wei-ping. Overview of fine granularity scalability in MPEG-4 video standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, **11**(3): 301 ~ 317.
- 38 Wu F, Li S, Zhang Y Q. A framework for efficient progressive fine granularity scalable video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, **11**(3): 332 ~ 344.
- 39 Wang Qi, Zhao Li, Wu Feng, *et al.* A rate allocation scheme for progressive fine granular scalable video coding[J]. Acta Electronica Sinica, 2002, **30**(2): 205 ~ 209. [王琪, 赵黎, 吴枫等. 可伸缩性视频编码的码率分配算法 [J], 电子学报, 2002, **30**(2): 205 ~ 209.]
- 40 Zhao X J, He Y W, Yang S Q, *et al.* Rate allocation of equal image quality for MPEG-4 FGS video streaming [EB/OL]. <http://amp.ece.cmu.edu/packetvideo2002/papers/32-uwsguasts.pdf>. 2002.
- 41 Shi Cui-zu, Yu Song-yu, Wang Jia. Rate allocation for MPEG-4 FGS video streaming[J]. Computer Simulation, 2004, **21**(6): 46 ~ 55. [史翠竹, 余松煜, 王嘉. FGS 视频流的码率分配算法研究 [J]. 计算机仿真, 2004, **21**(6): 46 ~ 55.]
- 42 Ding W, Liu B. Rate control of MPEG video coding and recording by rate-quantization modeling [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, **6**(1): 12 ~ 20.
- 43 Lin L J, Ortega A, Jay Kuo C C. Rate control using spline-interpolated R-D characteristics [A]. In: Proceedings of Visual Communications and Image Processing 1996 [C], Orlando, FL, 1996: 111 ~ 122.
- 44 Chiang T, Zhang Y Q. A new rate control scheme using quadratic rate distortion model [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, **7**(1): 246 ~ 250.
- 45 Tao B, Peterson H A, Dickinson B W. A rate-quantization model for MPEG encoders [A]. In: Proceedings of International Conference on Image Processing 1997 [C], Santa Barbara, CA USA, 1997: 338 ~ 341.
- 46 Yang K H, Jacquin A, Jayant N S. A normalized rate-distortion model for H.263-compatible codecs and its application to quantizer selection [A]. In: Proceedings of International Conference on Image Processing 1997 [C], Santa Barbara, CA, USA, 1997: 41 ~ 44.
- 47 Ribas-Corbera J, Lei S. Rate control in DCT video coding for low-delay communications [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, **9**(1): 172 ~ 185.
- 48 Sethuraman S, Krishnamurthy R. Model based multi-pass macroblock-level rate control for visually improved video coding [A]. In: Proceedings of Workshop and Exhibition on MPEG-4 [C], San Jose, California, USA, 2001: 59 ~ 62.
- 49 Wang L. Rate control for MPEG video coding [J]. Signal Processing: Image Communication, 2000, **15**: 493 ~ 511.
- 50 Farin D, Mache N, P. H. N. de With. A software-based high-quality MPEG-2 encoder employing scene change detection and adaptive quantization [J]. IEEE Transactions on Consumer Electronics, 2002, **48**(4): 2174 ~ 2193.
- 51 Yeo B, Liu B. Rapid scene analysis on compressed video [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1995, **5**(6): 533 ~ 544.
- 52 Yoneyama A, Nakajima Y, Yanagihara H, *et al.* MPEG encoding algorithm with scene adaptive dynamic GOP structure [A]. In: Proceedings of Ostermann J. Multimedia Signal Processing—1999 IEEE 3rd Workshop on [C], Copenhagen: IEEE Press, 1999: 297 ~ 302.
- 53 ISO/IEC JTC1/SC29/WG11, MPEG 97/M1931. Joint rate control for multiple video objects based on quadratic rate-distortion model [S].
- 54 ISO/IEC JTC1/SC29/WG11, MPEG97/M2554. Multiple-VO rate control and B-VO rate control [S].
- 55 Vetro A, Sun H, Wang Y. Joint shape and texture rate control for MPEG-4 encoders [A]. In: Proceedings of IEEE International

- Conference on Circuits and Systems [C], Monterey, USA, 1998; 285 ~ 288.
- 56 Kondi L P, Melnikov G, Katsaggelos A K. Joint optimal coding of texture and shape [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Thessaloniki, Greece, 2001, 3:94 ~ 97.
- 57 ISO/IEC JTC1/SC29/ WG11, MPEG99/M5552. An all FGS Solution for Hybrid Temporal-SNR Scalability [S].
- 58 ISO/ IEC/ IECJTC/ 29/ WG11, MPEG2000/M6475. Motion-compensati on Based Fine-granular Scalability (MC-FGS) [S].
- 59 Wang Q, Wu F, Li S P, *et al.* Fine-granularity spatially scalable video coding [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [C], Salt Lake City, 2001, 3:1801 ~ 1804.
- 60 M. van der Schaar, Radha H. Adaptive motion-compensation Fine-Granular-Scalability (AMC-FGS) for wireless video [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(6):360 ~ 371.
- 61 Wu S W, Gersho A. Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1991, 1(1):100 ~ 112.
- 62 Choi J, Park D. A stable feedback control of the buffer state using the controlled langrange multiplier method [J]. IEEE Transactions on Image Processing, 1994, 3:546 ~ 558.
- 63 Ramchandran K, Ortega A, Vetterli M. Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders [J]. IEEE Transactions on Image Processing, 1994, 3:533 ~ 545.
- 64 Lin D W, Wang M H, Chen J J. Optimal delayed-coding of video sequences subject to a buffer-size constraint [A]. In: Proceedings of SPIE Visual Communication and Image Processing 1993 [C], Cambridge, MA, USA, 1993:223 ~ 234.
- 65 Lee J, Dickinson B W. Joint optimization of frame type selection and bit allocation for MPEG video encoders [A]. In: Proceedings of International Conference on Image Processing 1994 [C], Austin TX, USA, 1994:962 ~ 966.
- 66 Wang Hui-bai, Zhang Chun-tian. A buffer control strategy based on importance of the image contents [J]. Journal of China Institute of Communications, 2000, 21(8): 21 ~ 26. [王辉柏, 张春田. 一种基于图像内容重要性的缓存器控制策略 [J]. 通信学报, 2000, 21(8): 21 ~ 26.]
- 67 Jay Kuo C C, Leou J J. A new rate control scheme for H.263 video transmission [J]. Singal Processing; Image Communication, 2002, 17(7): 537 ~ 557.