

一种基于 TMN8 模型的 H. 264 码率控制方法

李娜 王中元 何政 傅佑铭 常军

(武汉大学计算机学院国家多媒体软件工程技术研究中心 武汉 430072)

摘要 针对 TMN8 码率控制模型存在的一些局限客观上限制了其在实时视频通信中性能的问题,提出了3个方面的改进措施:目标比特率计算、图像中心加权、分层次码率控制。同时,针对 H. 264 码率控制中存在的码率控制与 RDO 运动估计间互为因果的矛盾,提出一种两遍运动估计的解决方案,并充分利用预运动估计阶段产生的附带信息来加速和优化编码过程。最终实现了一种基于 TMN8 模型的精确的 H. 264 码率控制方法。实验结果和在 SIP 视频会议系统中的具体应用都表明,本方法对提高实时通信系统的视频品质和 QoS 均起到了积极效果。

关键词 码率控制, H. 264, TMN8

中图分类号 TN919.8 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.6.025

Novel H. 264 Rate Control Method Based on TMN8 Model

LI Na WANG Zhong-yuan HE Zheng FU You-ming CHANG Jun

(National Multimedia Software Engineering Research Center, Computer School, Wuhan University, Wuhan 430072, China)

Abstract To address the drawbacks of TMN8 rate control model in real-time video communications in practical scenarios, this paper proposed three technical modifications; target bit rate computation, center-oriented perceptual weighting, and layered rate control. Meanwhile, a two-pass motion estimation approach was also established to deal with the dilemma of cause and effect between RDO estimation and rate control, in which the motion information generated by pre-estimation stage is further used to speed up and optimize formal encoding process. A high-precision H. 264 rate control method employing TMN8 model was ultimately implemented on the basis of the presented improved techniques and the two-pass approach. Simulation experimental results and practical applications in SIP video conference system demonstrate that this method does play a positive effect on the visual experience of real-time video.

Keywords Rate control, H. 264, TMN8 model

1 引言

近年来视频压缩技术的应用范围日益广泛,不仅在传统的广播电视、视频会议、视频监控等业务中得到普遍应用,而且也支撑了个人及时通信、社交媒体、数码影像档案等新兴应用的发展。在视频通信中,压缩视频需要在带宽需求不一致且变化的网络上传输,为了充分利用所提供的网络资源,并保证用户获得最优的感觉质量,需要引入码率控制机制,以获得在质量和带宽利用上的最佳均衡。在视频编码器的设计中,码率控制一直是非常开放而又关键的技术,不同的码率控制策略不会影响到系统的互通互联,但是其控制策略的好坏则直接影响到压缩视频的品质和对现有网络带宽的有效利用。迄今为止,被国际视频编码标准采纳的公开的码率控制算法主要有 MPEG2 的 TM5 模型^[1]、H. 263 的 TMN8 模型^[2,3]、MPEG4 的 SRC 模型^[4],以及 H. 264 的码率控制技术 JVT-G012^[5]。2013年,ITU-T 发布了最新的视频编码标准 H. 265/HEVC^[6],经过适应性的移植, JCTVC-H0213^[7] 将 H. 264 码

率控制技术应用到 H. 265 中。

在 H. 264 视频编码器中,率失真优化(RDO)作为提高编码效率的一个基本功能而存在。无论 RDO 运动估计还是编码模式选择,量化参数(QP)都是拉格朗日率失真优化公式中的必要参数。现有的码率控制算法都以图像块的复杂性估计为依据,而图像复杂度是运动估计步骤的附带计算结果,由此造成了码率控制和率失真优化间互为因果的矛盾局面。为了得到码率控制过程所用的画面活动复杂度参数, JVT-G012^[5] 方法采用线性预测模型通过上一帧已知的复杂度来估计当前帧未知的复杂度,然而对每个宏块(MB)执行线性预测消耗了大量的运算时间,在高清视频编码中代价更大。其次,对于诸如视频会议多点控制器中的解码还原视频,由于编码过程损失了部分高频信息,致使帧间宏块差异变小,大量宏块被编码器判别为跳过编码模式,过多跳过编码的宏块损害了线性预测器回归分析的精度和稳定性,最终导致预测的宏块复杂度严重偏离其真实的数值,从而影响码率控制的精度。

TMN8 码率控制模型是公认的数学基础最为牢固的模

到稿日期:2014-07-08 返修日期:2014-10-10 本文受国家自然科学基金(61170023,61303114,61172174,61302111,U1404618)资助。

李娜(1974-),女,副研究员,主要研究方向为档案影像处理;王中元(1972-),男,副教授,主要研究方向为多媒体通信, E-mail: wzy_hope@163.com(通信作者)。何政(1976-),男,讲师,主要研究方向为多媒体信息系统;傅佑铭(1979-),男,讲师,主要研究方向为网络通信;常军(1972-),男,讲师,主要研究方向为信息检索。

型,享有极高的恒定码率(CBR)控制精度。但是,将 TMN8 应用于 H. 264 时,不仅面临上述码率控制和 RDO 运动估计间互为因果矛盾的问题,而且鉴于 H. 264 采取的是周期量化方式, TMN8 模型产生的量化步长不能直接用于 H. 264 量化器。同时, TMN8 模型在码流缓冲区管理、感知失真度计算及量化参数控制等方面存在的一些局限也客观限制了其在实时应用中的表现,例如, TMN8 模型假定每次从编码缓冲区移走的数据等于每帧的平均比特数。为此,本文提出帧级目标比特数计算、图像中心加权感知失真计算、分层次码率控制等改进措施来提升 TMN8 模型的性能,并且引入预运动估计和量化步长转换步骤,将 TMN8 模型成功应用到 H. 264 的码率控制中。通过重用预运动估计的运动信息,加快正式运动估计的搜索效率,从而将运算复杂度控制在一个合理的水平。离线实验和 SIP 视频会议系统的在线测试结果都表明,本文方法对提高系统的实时视频通信品质和 QoS 都有明显效果。

2 TMN8 模型的改进

2.1 帧级目标比特数计算

在典型的控制算法中,每帧目标比特数的预算和跳帧处理都是基于码流缓冲区的填充度。码流缓冲区是介于信源编码器和信道传输模块间的内存缓冲区。缓冲区的输入为每帧实际编码的比特数,输出为网络传输模块发送的比特数。

在现有的 TMN8 模型中,从缓冲区读取的数据被视为每帧的平均比特数,即目标码率除以目标帧率。实际上,对于有 QoS 控制的传输,在帧编码间隔内实际发送出去的比特数不一定等于每帧的平均比特数,例如对于有拥塞控制的 IP 网络,发送出去的 UDP 包在网络好时大于均值,在网络差时小于均值。因此,根据目标码率和前后两次读取缓冲区的实际采样间隔来计算发送的比特数更为合理,即 $B_t = RT$, 其中 B_t 为比特数, R 为目标码率, T 为时间间隔。

目标编码比特率和帧率是任何码率控制模型的两个重要的输入参数,它们在模型初始化时就设置好。但是真实环境中的采样帧率可能是变化的,当设置的目标帧率明显高出实际采样帧率时,编码器输出的实际码率往往小于目标码率,从而造成可利用带宽的浪费。深入分析发现,问题的根源是现有的 TMN8 码流缓冲区的填充深度不允许出现下溢。因此,笔者引入虚拟缓冲区的概念,允许缓冲区的填充度为“负”值,也即下溢。最终,修订的缓冲区填充度的更新方法为:

$$W = \max(W + B' - B, U) \quad (1)$$

其中, B' 为上一帧编码的实际比特数, U 为下溢门限,默认为帧平均比特数的负数。表 1 的实验数据是在 300 帧 CIF 格式 foreman 序列下计算得到的,它反映了“负”缓冲区概念在提高比特率控制精度上的作用。

表 1 实际和目标码率比较

采样帧率 (fps)	目标帧率 (fps)	目标码率 (kbps)	实际码率(kbps)	
			改进前	改进后
15	25	512	400	502
15	20	384	315	378
15	12	128	130	132

2.2 图像中心加权

基于感兴趣区域的码率控制是近年来的研究热点^[8-10]。

感兴趣区域可以是客观视觉上比较关心的某个具体对象或对象的某一部分,从人眼视觉生理模型角度看,也可以是主观上比较敏感的视觉区域。感兴趣区域编码是在传统的码率控制模块的基础上,加上感兴趣区域先验知识或视觉生理加权信息的前馈输入参数。对于可视电话、视频会议这类以人为为主的交互式应用,画面的主要活动区域集中在图像中心的人头肩部部位,人眼视觉系统对这部分也最敏感。因此,如果在码率控制模型中引入金字塔型感知加权,让加权系数从图像中心到四周满足金字塔型分布,则能实现对画面中心码率资源分配的偏好,从而提升中心画面的编码质量。按照这一思路,本文提出如下的感知权重计算公式:

$$\omega_k = A_1 \left(1 - \frac{|r-R|}{R}\right) \left(1 - \frac{|c-C|}{C}\right) + A_2 \quad (2)$$

其中, A_1 为常数 1.5, A_2 为常数 0.1, R 为图像行宏块数的一半, C 为列宏块数的一半, r 为宏块行号, c 为列号。应用到 TMN8 模型中,将模型的失真加权因子 α_k 修正为 $\alpha_k' = \alpha_k \omega_k$ 。



(a) 无中心加权



(b) 有中心加权

图 1 图像中心加权视觉效果比较

图 1(a) 是没有加权的图像,图 1(b) 是加权后的图像。明显看出,图 1(b) 中心部位的人脸图像看起来更清晰,基本没有出现图 1(a) 那样的编码块效应。

2.3 分层次码率控制

量化参数 QP 的编码会占用一定的比特资源,如果每个 MB 都各自携带一个自己的 QP ,总计起来将是一笔不小的开销。在总比特预算不变的情况下,对 MB 头信息的编码占用了分配给 DCT 系数的比特数,从而引起 DCT 系数的描述精度下降,在低码率下影响尤为明显。因此,在低码率视频编码中有必要降低 QP 的调整粒度,为整帧图像或条带分配一个共同的 QP ,而非为每个 MB 都赋予一个 QP 。更重要的是,视频编码标准通常规定相邻宏块间量化步长的差值不超

过一个范围,如 H. 263、MPEG4 视频编码协议规定相邻宏块间的差值范围为 $[-2, 2]$,如果每帧初始量化步长取前一帧的最后一个值(TMN8 模型规定),当出现场景切换或图像不同部位画面活动变化大时,当前帧或条带的 QP 很难在几个宏块的跨度内适应新的基准值。

通过分级自适应量化参数可以解决上述问题,将码率控制分为两个层次:帧、条带级别以及宏块级别。先根据帧、条带一级的码率控制算法估算出一个较粗的量化步长基准值,将它作为每帧、条带中 TMN8 模型量化步长的初始值,然后根据目标码率大小,采用 TMN8 模型进行宏块级别的细致调整。

对于具体实现,本文采用可伸缩速率控制(SRC)算法^[4]来估计帧、条带层次的量化参数。SRC 算法假定编码器的失真函数由以下模型表示:

$$B = X_1 \times S \times Q^{-1} + X_2 \times S \times Q^{-2} \quad (3)$$

其中, B 是编码比特数, S 是编码复杂度(绝对差值的均值), Q 是量化参数, X_1 和 X_2 是模型参数。模型参数根据当前帧或条带的编码结果逐次更新。

在实际应用中,我们注意到一个现象:使用固定量化参数 $QP=28$ 量化时,对于一般运动程度的视频数据,bpp(bits per pixel)值约为 0.1,对应到 CIF 格式实时编码的码率约为 250kbps,此时视频质量已经足够清晰。因此,为减少一些量化参数离异点(outlier)对编码质量的负面影响,有必要对量化步长的上限进行限制,根据目标 bpp 为 QP 设置最大值,一般情况下这个最大值要比编码标准允许的量程小很多。同时,为防止时域上视觉质量出现大的波动,进一步对前后两帧间量化参数的变化范围进行限定,将当前帧的量化参数(含帧、条带和宏块 3 个级别)限制在上一个同类型帧的平均量化参数的一个较小的变动范围内。

3 TMN8 在 H. 264 中的实现

3.1 总体方案

H. 264 采用的是率失真运动估计技术,运动估计时用到了量化参数 QP,而码率控制算法对 QP 的计算又依赖于运动估计模块得到的宏块复杂度指标,从而造成运动估计和码率控制互相依赖的局面。为了解决这一矛盾,本文采用两遍运动估计的方案,如图 2 所示,在码率控制之前,先执行一遍预运动估计以提供码率控制步骤所需要的图像活动复杂度,在码率控制步骤之后,再利用码率控制模块输出的真实的 QP 参数,执行正式的 RDO 运动估计。预运动估计除了为码率控制输送图像活动复杂度信息外,还带来两个附加好处:首先,可以将已经得到的初步的运动矢量信息传递给正式运动估计模块,加速正式运动估计阶段的搜索效率;其次,图像活动复杂度信息是一种可靠的判断场景切换的依据。

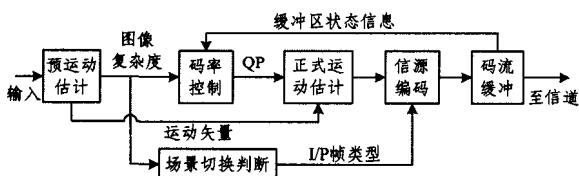


图 2 基于 TMN8 的 H. 264 码率控制总体方案

下面结合图 2 所示的总体方案原理图,对本方案部分技

术要点作进一步的描述,其中码率控制这部分放到下一节详细讨论。

预运动估计。预运动估计阶段计算 RDO 代价的 QP 取上一帧的平均 QP,而且为了节省预运动估计的计算开销,只执行 16×16 的帧间运动搜索和 16×16 的帧内预测。帧间运动搜索在当前输入帧和前一重建帧之间进行,帧内预测在当前输入帧中进行。

正式运动估计。将预运动估计阶段产生的运动矢量作为正式运动估计菱形搜索策略的起始点,提高正式运动估计的搜索效率,抵消预运动估计引起的额外的计算代价,从而将总体运算复杂度控制在一个合理的水平。

场景切换自适应 I 帧编码。设置一个场景切换门限,如果图像总体活度复杂度超过这个门限,就作为 I 帧编码;否则,依然作为 P 帧编码,插入 I 帧的强度通过改变门限值来调整。既然增加了主动 I 帧插入机制,视频编码器所要求的周期性 I 帧刷新的间隔可以扩大,从而可减少周期性 I 帧的数量,提高编码器的压缩效率和降低周期性 I 帧引起的瞬时突发码流。

帧内宏块刷新。在场景切换自适应 I 帧编码策略中,如果场景长时间没有出现切换,帧内编码刷新的频率达不到控制帧间预测误差累积的要求,反而会降低编码质量。为此,本文进一步提出宏块顺序刷新的补充机制,每次刷新若干空间上连续的宏块,刷新位置按光栅扫描顺序逐帧移动。一次刷新的宏块数目按照总宏块数除以 I 帧间隔确定。例如,如果周期性 I 帧间隔拟设定为 5000 帧,对于 1080P 视频,则折算为每帧刷新的宏块数目为 $\lceil \frac{1920 \times 1080}{16 \times 16} / 5000 \rceil = 2$ 。

3.2 码率控制

H. 264 采取的是周期量化策略,不同于以前的 H. 263 和 MPEG4 标准,在 H. 264 中,量化器不是直接用量化参数 QP 来量化 DCT 系数,而是将其映射为量化步长 $Qstep$,然后用量化步长来量化。因此,将 TMN8 用于 H. 264 码率控制需要进行 QP 和 $Qstep$ 间的转换。量化步长 $Qstep$ 到量化参数 QP 的转换公式为:

$$QP = 6 \log Qstep / \log 2 + 4 \quad (4)$$

反之,QP 到量化步长 $Qstep$ 的转化公式为:

$$Qstep = 2^{(QP-4)/6} \quad (5)$$

结合上文的改进措施和 TMN8 模型,本文提出的一种 H. 264 码率控制方法的执行步骤如下:

(1) 初始化

设置目标码率 R 和目标帧率 F 。 M 为跳帧门限,默认为 $M=2R/F$; D 为延迟门限,默认为 $D=0.5R/F$ 。缓冲区填充度 W 初始化为 0。根据 TMN8 算法内容,初始化其它模型参数。

(2) 估算目标比特数

每编码一帧,根据式(1)更新编码缓冲区。在对当前帧编码时,首先检查 W ,如果 $W > M$,则跳过当前帧的编码,同时 $W = W - B_i$;否则每帧的目标比特数为 $B = (R/F) - \Delta$,其中

$$\Delta = \begin{cases} W/F, & W > D \\ W-D, & \text{otherwise} \end{cases}$$

(3) 计算宏块运动复杂度

①对于第 k 个 MB,运动复杂度指标 $\sigma_k = \frac{SAD_k}{256}$, SAD_k 为预运动估计得到的 SAD 值。

②失真权重参数

$$\alpha_k = \begin{cases} 2 \frac{B}{16^2 N} (1 - \sigma_k) + \sigma_k, & \frac{B}{16^2 N} < 0.5 \\ 1, & \text{otherwise} \end{cases}$$

其中, N 为 MB 的总数目。进一步结合式(2)的金字塔感知加权模型,将 α_k 修正为 $\alpha_k = \alpha_k w_k$,且令 $S = \sum_{k=1}^N \alpha_k \sigma_k$ 。

(4)帧级或条带级码率控制

用 2.3 节式(3)描述的 SRC 模型计算帧和条带一级的量化参数,此时编码复杂度等于帧或条带内全部 MB 复杂度之和。计算结果作为帧或条带的量化参数,如果目标码率过低,则不再执行步骤(5)和(6)的宏块级码率控制;否则,将 TMN8 模型的 QP 设置成帧或条带的量化参数,且继续执行步骤(5)和(6)。

(5)计算宏块量化参数

$$TMN8 \text{ 模型给出宏块的量化步长为 } Qstep = \sqrt{\frac{256k \delta_k S}{L \sigma_k}}$$

这里 K, L, S 均为模型参数,每编码一个 MB,模型参数更新一次。在得到 $Qstep$ 后,执行如下后处理步骤:

① 根据式(4)将 $Qstep$ 转化为量化参数 QP。

② 将 QP 限制在上一帧平均量化参数的 ± 2 范围内,限制后的 QP 即为该 MB 最终的量化参数。

(6)TMN8 模型更新

根据式(5)将限制范围后的 QP 重新还原 $Qstep$,以便用 $Qstep/2$ 更新 TMN8 模型参数。

需要指出的是,上面的描述没有给出 TMN8 的所有技术细节,只是侧重于用第 2 节研究的方法对 TMN8 改造的部分,完整的 TMN8 计算过程可参考其技术文档^[2]。

4 实验结果和分析

4.1 实验环境

测试设备为 HP 4411s 笔记本电脑,配置 Intel(R) Core (TM) 双核 CPU,主频分别为 2.10GHz, 2.09GHz,内存 2.96GB,硬盘 230GB,操作系统为 Windows XP。

本文测试的 H.264 编码器采用经过高度优化的 X264 开

源版本^[11],运行在 Baseline 档次。除 X264 中自带的 ABR (average bit rate) 码率控制方法外,本文将 TMN8 方法和 JVT-G012 方法加入到 X264 中,本实验将比较这 3 种码率控制方法的性能。

测试分为离线测试和在线测试,前者运行独立的 X264 编码器,测试视频源为文件测试序列,所用的测试序列的描述如表 2 所列;后者将 X264 编码器集成到 H.323/SIP 视频会议系统,在 IP 网络传输环境下实时测试,现场采集室内场景下的实时视频。

表 2 离线测试的视频序列

名称	分辨率	帧数	帧率(fps)	场景描述	原始或解码视频
Foreman	CIF	300	25	个人画面	原始
Shake	720p	288	25	摄像头前急剧晃动的物体	解码
Street	1080p	634	25	街道交通场面	解码
Home	1080p	676	25	办公室场景	原始

4.2 离线测试

离线测试主要是用客观的方法衡量不同方法的码率控制精度、视频质量和处理效率。客观质量用 PSNR 度量,运算效率以每秒编码的帧数 (fps) 衡量。为准确地评估码率控制精度,不仅要考察实际输出的平均码率是否与设置的目标码率相符合,而且还要比较每帧编码比特数的平稳性,本文用帧字节数偏差来评估,偏差越小,表明码率控制越平滑,瞬时抖动越小。帧字节数偏差 σ 的计算如下:

$$\sigma = \frac{\sum_{i=1}^N |B_i - \bar{B}|}{N} \quad (6)$$

其中, N 为编码的帧数, B_i 为每帧编码实际产生的字节数, \bar{B} 为平均字节数。

表 3 显示了 ABR、JVT-G012 和 TMN8 3 种方法的实验结果。从帧字节数偏差指标可以看出, TMN8 几乎将每帧编码的比特数控制在一个恒定的水平,每帧的波动远远小于 ABR,也小于 G012,而且 PSNR 反映的客观质量没有下降,与 G012 比,还略有提升。三者之间, G012 尽管能给出中等的码率控制精度,但在 720P 以上分辨率时,其码率控制的运算代价显著增加,导致整体编码效率急剧下降,从而限制了实际应用。因此,在下面的在线评测中,只对 ABR 和 TMN8 作进一步比较。

表 3 离线测试结果

序列	设置码率 (kbps)	实际码率 (kbps)			帧字节数偏差			平均 PSNR (dB)			运算帧率 (fps)		
		ABR	G012	TMN8	ABR	G012	TMN8	ABR	G012	TMN8	ABR	G012	TMN8
Foreman	384	380	380	376	326	61	38	36.51	36.38	36.46	194	178	190
Shake	1000	936	716	1000	1042	1475	357	45.16	41.32	47.29	30	19	27
Street	2000	1989	2101	2015	1644	576	75	42.78	42.76	42.71	18	7	16
Home	2000	1991	1930	1989	1465	516	79	46.07	45.75	46.16	16	7	16

4.3 在线测试

在线测试主要评估不同码率控制方法下压缩视频的主观视觉感受,如连贯性、视觉品质、码率和帧率的统计数据是否稳定等。鉴于 JVT-G012 不能实时处理 720P 以上分辨率的视频,即无法满足实时评测的要求,本部分只对 ABR 和 TMN8 的效果进行比较。实验为两个 SIP 软终端对通,其中一个集成的是 ABR 算法的 X264,另一个集成了 TMN8 算法的 X264,两个终端的摄像头对着同一场景摄像。

表 4 列出了 SIP 视频会议终端实时传输的效果,视频分辨率为 720P,传输带宽为 1Mbps,帧率设置为 25fps。可以看

出, TMN8 能够将码率严格控制在设置的目标码率内。由于码率控制准确,视频编码器很少启动跳帧,因此 TMN8 方法下的帧率稳定,画面连贯性和流畅度明显优于 ABR。相反, ABR 在画面运动幅度增大时,码率溢出明显,在跳帧模式的作用下,帧率迅速下降,顿挫感增加。二者的主观视觉质量基本相当,在动态时 TMN8 的块效应略明显,但从动态画面恢复到清晰的过程较 ABR 快很多,且源于 TMN8 的量化步长能够做到宏块单位自适应,其背景画质更细腻。

(下转第 150 页)

structural division probability[J]. Chinese journal of computers, 2014, 37(5): 1206-1214

[13] 徐迪威. 大数据与科技管理[J]. 科技管理研究, 2013(24): 216
Xu D W. Big Data and Technology Management[J]. Technology Management Research, 2013(24): 216

[14] 董杨慧, 谢友宁. 大数据视野下的数据泄露与安全管理[J]. 情报杂志, 2014, 33(11): 154-158
Dong Y H, Xie Y N. Data Disclose and Secure Management in Gig Data Vision[J]. Information Magazine, 2014, 33(11): 154-158

[15] 朱建明, 宋彪, 黄启发. 基于系统动力学的网络安全攻防演化博弈模型[J]. 通信学报, 2014, 35(1): 54-60

Zhu J M, Song B, Huang Q F. Evolution Game Model of Offense-defense for Network Security Based on System Dynamics [J]. Journal on Communications, 2014, 35(1): 54-60

[16] 黄启发, 朱建明, 宋彪, 等. 社交网络用户隐私保护的博弈模型[J]. 计算机科学, 2014, 41(10): 184-189

Huang Q F, Zhu J M, Song B, et al. game model of user's privacy-preserving in social networks[J]. Computer Science, 2014, 41(10): 184-189

(上接第 114 页)

表 4 实时通信视频性能指标比较

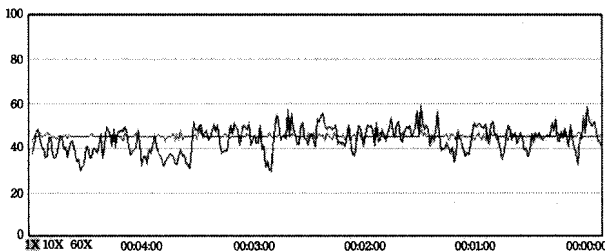
方法	码率	帧率	CPU 占用	图像质量	画面连贯性
ABR	画面动态时候会超过较多, 码率统计曲线波动大	不稳定, 在 15~25fps 范围频繁跳动, 动态时帧率急剧下降	~50%	块效应少一些, 但细节欠清晰, 动态画面恢复到清晰的过程较 TMN8 慢	运动时流畅性下降, 且偶有停顿
TMN8	基本没有超过设定带宽, 码率统计曲线平滑	稳定在 24~25fps	相当 (~50%)	块效应略多, 细节清晰, 背景细腻, 恢复快	保持流畅, 无停顿现象

此外, 本文特别地用专业的流量测试软件 Netmon 统计了带宽的使用情况, 结果如图 3 所示。可以看出, TMN8 的码率波动曲线明显较 ABR 平滑, 基本上没有出现 ABR 那种瞬时码率上升很高的情形。众所周知, 在带宽受限且得不到 QoS 保障的公网上, 瞬时的突发性码流是引起网络丢包的一个最主要的因素, 因此, TMN8 特别适合于公网视频通信。

精度和实时视频通信的品质。由于 H. 265 采用了与 H. 264 一致的率失真编码框架, 本文技术有望被推广到 H. 265 的码率控制中, 下一步我们将在 H. 265 环境下开展测试验证。

参考文献

- [1] ISO/IEC JTC1/SC29/WG11/93-225b, MPEG-2 Test Model 5 [S]. Test Model Editing Committee, 1993
- [2] ITU-T Video Coding Experts Group. Video Codec Test Model, Near-Term, Version 8 [S]. Portland, 1997
- [3] Jordi R C, Lei S. Rate control in DCT video coding for low-delay communications [J]. IEEE Trans. on Circuits and Syst. for Video Technol., 1999, 9(12): 172-185
- [4] Lee H J, Chiang T H, Zhang Y Q. Scalable rate control for MPEG-4 video [J]. IEEE Trans. on Circuits and Syst. for Video Technol., 2000, 10(6): 878-894
- [5] JVT-G012-r1, Adaptive basic unit layer rate control for JVT [S]. Pattaya II, Thailand, 2003
- [6] ITU-T Recommendation H. 265 & ISO/IEC HEVC, High Efficiency Video Coding [S]. 2013
- [7] Choi H, Nam J, Yoo J, et al. JCTVC-H0213, Rate control based on unified RQ model for HEVC [S]. San José, 2012
- [8] Grois D, Hadar O. Complexity-aware adaptive bit-rate control with dynamic ROI pre-processing for scalable video coding [C]// IEEE Int. Conf. on Multimedia and Expo(ICME). 2011: 1-4
- [9] Qu T S, Huang Y H, Chen H H. SSIM-based perceptual rate control for video coding [J]. IEEE Trans. on Circuits and Syst. for Video Technol., 2011, 21(5): 682-691
- [10] Wu Guan Lin, Fu Yu Jie, Huang Sheng, et al. Perceptual quality-regulable video coding system with region-based rate control scheme [J]. IEEE Trans. on Image Process, 2013, 22(6): 2247-2258
- [11] X264 Source Code[OL]. <http://developers.videolan.org/x264.html>



上下波动较大的线表示 ABR, 上下波动较小的线表示 TMN8

图 3 码率波动曲线

结束语 本文对 TMN8 模型在 H. 264 码率控制中的改进和应用展开研究。文章首先从实际应用环境和 TMN8 文档中描述的仿真环境的差异出发, 对 TMN8 模型提出 3 点完善措施, 包括帧级目标比特数计算、图像金字塔感知加权、分层次量化步长控制。每个改进措施都有不同的针对性: 引入“负”缓冲概念是为了在实际采样帧率达不到目标帧率的情况下, 更好地利用网络带宽资源; 采用金字塔加权是考虑到在交互式视频应用中, 人眼对位于图像中心的头肩部画面更敏感的缘故; 分层次控制量化步长则是为了更好地保障视频画面的流畅性, 因为从主观感受来看, 帧率高但不流畅的视频反而不如帧率稍低但流畅的视频的视觉感受好。针对 TMN8 在 H. 264 中应用面临的 RDO 运动估计和码率控制互为因果的矛盾, 本文提出了一种两遍运动估计的方案, 充分利用预运动估计阶段的附带信息来加速和优化编码过程。本文最终实现了一种基于 TMN8 模型的 H. 264 码率控制方法。离线和在线测试结果都表明本文方法显著改善了 H. 264 码率控制的