

## 基于自适应速率模型的低时延视频编码速率控制算法<sup>①</sup>

霍 炎<sup>②\*</sup> 李生红<sup>\*</sup> 荆 涛<sup>\*\*</sup> 冯玉珉<sup>\*\*</sup>

(<sup>\*</sup>上海交通大学电子工程系 上海 200240)

(<sup>\*\*</sup>北京交通大学电子信息工程学院多媒体实验室 北京 100044)

**摘 要** 针对实时视频通信具有的低时延特性,通过研究速率模型形状参数和图像内容之间存在的关系,提出了一种视频图像自适应速率模型,此模型较其它速率模型能够更好地拟合真实码率输出,使得基于该模型速率控制下的输出速率更加平稳。为了避免 TMN8 模型中使用宏块方差方法得到宏块量化步长所引入的计算量,在帧层一次计算得到视频图像的基本量化步长,并在宏块层对其进行简单调整,从而降低了系统的处理时延。仿真实验表明,此文提出的基于新模型的速率控制算法较 TMN8 在输出码率稳定性及系统时延上都具有较好的性能。

**关键词** 平均绝对误差(MAD), 自适应速率模型, 量化步长, 速率控制

## 0 引 言

随着无线视频通信、无线摄像机、网络摄像机等一系列多媒体视讯类消费电子产品的广泛应用,人们在享受视频业务带来的诸多便利的同时,对服务质量提出更高的要求<sup>[1]</sup>,这突出体现在网络视频的实时通信上。网络视频实时通信会受到编码比特率与信道传输能力的限制,为适应视频信源编码变比特率以及信道时变特性的要求,在视频通信的发送端需使用固定大小的缓冲器以及防止缓冲器出现上溢或下溢的速率控制机制。但由于实时通信的缓冲器容量较小,因此对其速率控制的时延要求则更高。

在低时延速率控制方面,人们进行了大量的研究。Jordi 在文献[2]中提出了被 TMN8 方法采用的速率控制算法。该算法从理论分析入手,利用经验速率控制模型和经验失真模型,通过求拉格朗日条件极值的方法得到了一个理论上最优的求取宏块层量化步长的式子,并将该式应用于所有宏块的量化过程。TMN8 的这一算法相较于 TMN7 以及 VM7 有很大的优势,但是却存在着一定的问题。首先,最优量化步长的确定公式是建立在两个经验模型基础上的,由于每一个模型都和实际的情况存在着误差,最后的量化步长将是这些误差的叠加,由此造成编码输出的波动范围较大。其次,在进行宏块层编码时,需要计算每一宏块的方差,计算复杂度高,处理时延

较长。随后 Lam 在文献[3]中详细分析了视频图像离散余弦变换(discrete cosine transform, DCT)系数具有拉普拉斯分布的统计特性,并具此特性在文献[4]中提出基于柯西密度的速率控制算法,但基于统计特性的速率控制算法速率模型复杂,非常不适合实时视频通信应用。于是,为了满足实时视频通信,文献[5]分析了 DCT 交流系数均方和与其变换前的像素域方差的关系,依据 DCT 交流系数的方差和代替像素域的方差的方法,对传统的 TMN8 算法进行了改进,以一定的视频图像量化失真换取了运算量的降低,同时由于该算法依旧采用了 TMN8 的宏块量化步长确定方法,对于量化步长的选取仍具有较高的计算复杂度。而 Tsai 在文献[6]中从提高视频图像质量方面入手,根据宏块编码失真与编码顺序之间的关系,确定由图像的重要性决定图像的编码顺序这一方法,改善了上述的视频质量,但却未能解决计算复杂度的问题。

本文基于以上的分析经验,提出了一种新的速率控制策略。该方法以经过 DCT 变换的带有运动估计的  $P$  帧为研究对象,通过分析图像的速率-量化(rate-quantization, R-Q)曲线和图像的平均绝对误差的关系,得到了一个新的速率模型,并在该模型的基础上提出了新的速率控制的方法。通过与其他模型比较,新的速率模型和实际的 R-Q 曲线具有更好的拟合效果,从而保证了求得量化步长的准确性;

① 国家自然科学基金(60772040, 60772042)资助项目。

② 男,1982 年生,博士;研究方向:多媒体信号处理及网络应用;联系人, E-mail: greathuo@126.com  
(收稿日期:2009-03-23)

同时新的速率控制方法较 TMN8 计算量减小,具有平稳的编码输出。

## 1 图像自适应速率模型

率失真模型本质上是基于香农的率失真理论,它指出了信源在满足一定失真度下可以达到的最小的编码比特数,率失真模型给出了不同的失真情况下信源的压缩极限。但是在速率控制中,率失真函数则是一种实际曲线的回归拟合。由于量化步长在整个压缩编码过程中对传输速率与信源失真都起到重要的影响,因此率失真函数被分成了速率模型和失真模型,本文只考虑速率模型。

现有的速率模型大概分成三类:第一类是基于原始形式的对数模型,第二类是二次模型<sup>[7,9]</sup>,如式

$$R(Q) = A \cdot Q^{-1} + B \cdot Q^{-2} \quad (1)$$

所示。第三类是指数模型<sup>[10,11]</sup>,如式

$$R(Q) = A + B \cdot Q^{-\gamma} \quad (2)$$

所示。其中  $A$  与  $B$  均为模型系数。后两种模型从根本上都是源于对数模型的傅立叶展开,二次模型取的是其中的一次项和二次项,简化了对数模型的复杂度,但是直接计算量化步长仍然十分复杂并且会出现复数结果。指数模型则克服了这一缺点,在

知道目标比特数的情况下可以直接计算量化步长。

本文提出的模型以指数模型为基础,将模型形状参数与图像内容相结合,针对指数模型鲁棒性较差这一问题。

首先分析编码输出  $R$  和图像内容的关系。文献[12]证明了在  $Q$  一定的情况下,速率  $R$  和图像的每像素平均绝对误差(mean absolute deviation, MAD)呈线性关系,因此有

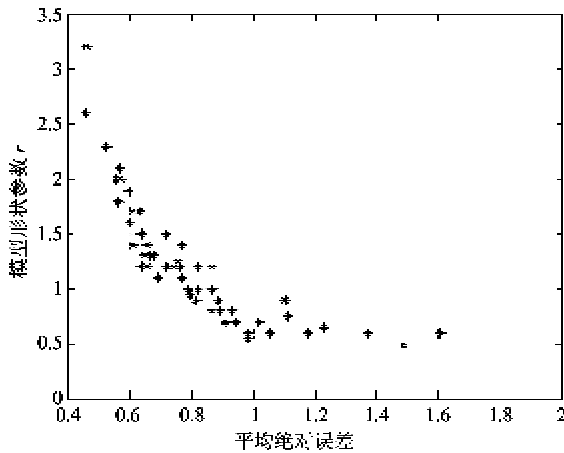
$$R(Q) = A + B \cdot mad \cdot Q^{-\gamma} \quad (3)$$

$$mad = \frac{1}{N} \sum_i \sum_j |p_{i,j}| \quad (4)$$

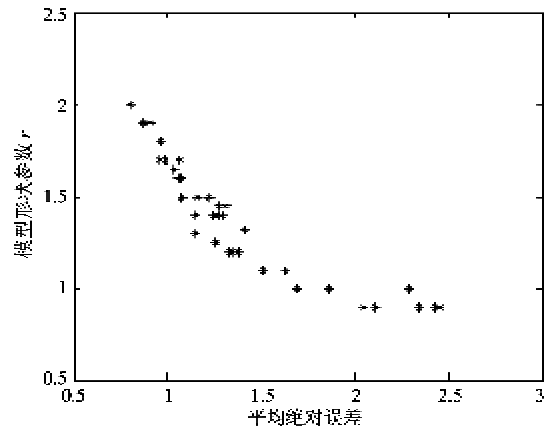
式中,  $A$  与  $B$  均为模型尺度参数,而  $\gamma$  表示模型形状参数,  $mad$  表示平均绝对误差值,  $N$  是编码单元的像素数目,  $p_{i,j}$  是经过运动补偿后的像素差值。下面针对模型的形状参数  $\gamma$  和图像的  $mad$  值之间的关系做进一步分析。

图 1(a)与(b)分别给出了视频序列 *Claire* 和 *Salesman* 的  $(mad, \gamma)$  点集的分布情况。图中的每一个点都代表具有相应  $mad$  值的图像帧。如图所示,在一定的范围内,  $\gamma$  和  $mad$  值近似呈线性关系,利用回归分析可得到回归方程如下:

$$\begin{cases} \gamma = K1 + \bar{K} \cdot mad, & mad \leq MAD \\ \gamma = \Phi, & \text{其它} \end{cases} \quad (5)$$



(a) 视频序列 *claire*



(b) 视频序列 *salesman*

图 1 视频图像  $(mad, \gamma)$  点集的分布

式(5)中:

$$\Phi = \frac{1}{m} \sum_{i=1}^m \gamma(i) \quad (6)$$

$$\bar{K} = \frac{\sum_{j=1}^n (mad(j) - \overline{mad})(\gamma(j) - \bar{\gamma})}{\sum_{j=1}^n (mad(j) - \overline{mad})^2} \quad (7)$$

$$K1 = \bar{\gamma} - \bar{K} \cdot \overline{mad} \quad (8)$$

$MAD$  是当  $\gamma = 1$  时的  $mad$  值,当  $mad < MAD$  时,  $\gamma$  和  $mad$  成线性关系,  $K1$  与  $\bar{K}$  是线性回归系数;当  $mad > MAD$  时,  $\gamma$  在一固定值周围波动,这里将这一固定值表示为  $\Phi$ 。式(6)至式(8)分别用于求出式(5)中的各个参量。

$m$  表示满足  $\gamma < 1$  的  $(mad, \gamma)$  取样点数目,  $n$  为满足  $\gamma \geq 1$  的取样点数目,  $\overline{mad}$  均值  $\overline{mad} = \frac{1}{n} \sum_{j=1}^n mad(j)$ , 平均形状参数  $\bar{\gamma} = \frac{1}{n} \sum_{j=1}^n \gamma(j)$ 。回归方程的准确度取决于回归样点的数目, 数目越多则拟合越准确。

图2和表1分别给出了本文提出的自适应模型、二次模型以及 TMN8 速率模型比较结果。由于自适应模型的形状参数是随图像内容的变化而变化的, 所以它能更好地拟合实际速率曲线的变化曲度, 使得在相同的量化步长下, 得到的编码比特数较其它模型更接近实际编码输出, 如表1所示。

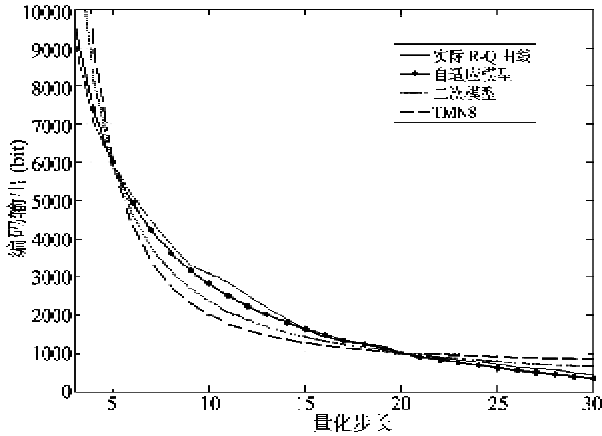


图2 Claire 第30帧图像的 R-Q 曲线

表1 视频序列 Salesman 在不同速率模型、不同量化步长下的平均比特误差比较

| 量化步长    | 5    | 15  | 20  | 25  | 35  | 40  | 45  |
|---------|------|-----|-----|-----|-----|-----|-----|
| 自适应模型   | 1462 | 147 | 147 | 54  | 74  | 100 | 123 |
| 二次模型    | 2204 | 419 | 317 | 131 | 76  | 126 | 183 |
| TMN8 模型 | 2272 | 621 | 465 | 200 | 132 | 244 | 349 |

## 2 基于速率模型的速率控制策略

基于第1节提出的自适应速率模型, 本节给出了一种新的速率控制策略, 其流程如图3所示。

**步骤1:** 为当前图像分配目标比特数  $Tf$ 。

由于算法针对的是  $P$  帧, 图像的内容特性在编码过程中影响较小, 所以在帧层的码率分配上暂不考虑图像的具体内容, 延续使用 TMN8 中的分配策略<sup>[13]</sup>:

$$Tf = \frac{R}{F} - \Delta \quad (9)$$

$$\Delta = \begin{cases} W/F, & W > \alpha \cdot M \\ W - \alpha \cdot M, & \text{其它} \end{cases}$$

$$W = \max(W_{pre} + B' - R, 0)$$

其中,  $Tf$  为当前帧编码目标比特数,  $R$  为信道比特传输速率,  $F$  为帧率,  $W$  为缓冲器中的残留的编码比特数,  $M$  为缓冲器的最大容量,  $W_{pre}$  为在上一帧编码前缓冲器中的比特数,  $B'$  为前一帧的实际的编码输出,  $\alpha$  为常数, 根据不同的  $R$ 、 $F$ 、 $M$  的不同而不同, 一般取 0.1<sup>[13]</sup>。

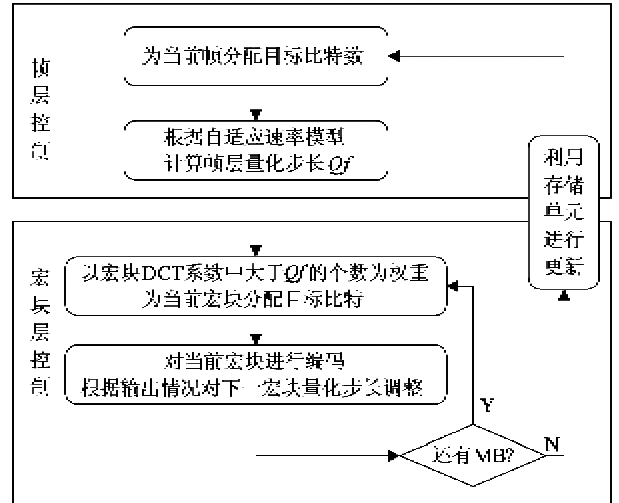


图3 速率控制流程图

**步骤2:** 建立当前帧的 R-Q 模型。

分别根据式(4)与式(5)求出当前帧的  $mad$  值与模型形状参数  $\gamma$ , 并依据式

$$A = \bar{R} - B \cdot mad \cdot \bar{X} \quad (10)$$

$$B = \frac{\sum_i (X^i - \bar{X})(R^i - \bar{R})}{\sum_i (X^i - \bar{X})^2 \cdot mad} \quad (11)$$

计算模型尺度参数  $A$  和  $B$ , 于是可以利用式

$$Qf = \left( \frac{Tf - A}{B \cdot mad} \right)^{-1/\gamma} \quad (12)$$

直接求出满足目标比特数  $Tf$  的量化步长  $Qf$ 。为了方便比较, 给出了 TMN8 中的  $Q$  的求取公式:

$$Q = \sqrt{\frac{M^2 K}{\beta} \frac{\sigma_i}{\alpha_i} \sum_{i=1}^N \alpha_i \sigma_i} \quad (13)$$

式(13)求取的是面向宏块的量化步长, 其中  $M$  为宏块的大小,  $N'$  为剩余的宏块数目,  $\beta$  为前一宏块编码结束后当前帧内剩余的比特数,  $K$  与  $\alpha_i$  为加权系数,  $\sigma_i$  为第  $i$  个宏块的均方差。

**步骤3:** 宏块的比特数分配

在进行宏块层编码时, 如果不从整体上分配每一宏块的目标编码比特, 只依顺序进行编码, 会造成宏块的编码质量随所在编码顺序的增大而降低。针

对这一问题,本文依据宏块的内容,在编码之前对每一宏块的编码比特数进行预分配。由于在编码过程中,直接影响编码输出数量的是 DCT 变换后的系数中大于量化步长的系数的个数  $n^i$  的大小,因此本文以  $n^i$  为权重进行目标比特分配:

$$Tb(i) = \frac{n^i}{\sum_i n^i} \cdot Tf \quad (14)$$

**步骤 4:**对当前宏块进行编码

对于第一个宏块,定义  $Qb = Qf$ 。假如编码比特输出比特数  $|Rb(i) - Tb(i)| > R \cdot TH$  (其中  $R \cdot TH$  是一常数,根据系统编码比特数设置大小),则在下一宏块编码前,进行量化步长的调整:

- 若  $Rb(i) > Tb(i)$ , 则令  $Qb^{i+1} = Qf + \delta$ 。
- 若  $Rb(i) < Tb(i)$ , 则令  $Qb^{i+1} = Qf - \delta$ 。

依据 H.263 的建议,为保证输出图像质量的稳定,相邻宏块之间的量化步长不能超过  $2^{[13]}$ ,本文将  $\delta$  设为 1。

**步骤 5:**模型参数更新过程

当第  $i$  帧编码结束后,计算该帧的实际编码比特数  $Rf$  和平均量化步长  $\overline{Qf}$ ,并对缓冲器残留比特数、模型尺度参数等进行更新。在建立首帧率失真模型时,只通过预编码来获得一组  $(Q^i, R^i)$  样值,并利用式(10)与(11)求得模型尺度参数。式(10)、(11)根据最小二乘估计法获得,其中  $X = Q^{-\gamma}$ 。

由于预编码会引入大量的处理时延,该算法采用建立存储库的方法来避免这一点。图 4 说明图像实际的 R-Q 曲线与图像  $mad$  值相关,当  $mad$  值相近时,实际的 R-Q 曲线则相差很小。

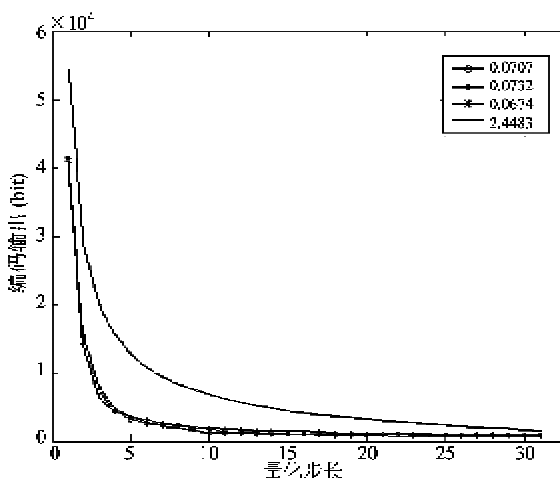


图 4 R-Q 曲线与  $mad$  的关系图

由此本文提出了一种新的方法模型参数进行更

新。当第  $i$  帧的实际的编码比特数  $Rf$  和目标编码比特数  $Tf$  之差小于阈值  $TH$ , 时将第  $i$  帧的  $(mad, Rf, \overline{Qf})$  放入存储单元。这样如果当前帧的  $mad$  和存储单元中的某数据相近时,即可直接使用存储数据由式

$$Qf = \frac{1}{T} \sum_{t=1}^T \overline{Qf}(t) \quad (15)$$

进行  $Qf$  的求得,其中,  $T$  是在存储单元中与当前帧的平均绝对误差差值  $\Delta mad$  满足一定要求的数据组个数。由于这一过程增加了比较的计算量,所以存储单元的容量最大设为 20,当符合要求的数据组超过 20 时,则以先入先出的方式对存储单元进行更新。

### 3 算法性能分析

本文使用 MATLAB 进行仿真分析,通过与 TMN8 策略进行比较,对本文给出的低时延速率控制算法的编码输出平稳性和算法复杂度进行验证。

#### 3.1 编码输出的平稳性分析

速率控制的一个主要的目的就是稳定视频码流的输出以满足给定的带宽条件。图 5 给出了在信道速率为 24kbit/s, 30 帧/s 的帧率下,视频序列 *Claire* 分别在 TMN8 和本文提出的控制策略下的编码输出情况。由于两种策略的量化步长的获得都是基于率失真模型,模型的准确度将直接影响编码输出的波动性,由第 2 节说明了自适应速率模型对于实际编码速率的拟合优于 TMN8,因此图 5 给出了两种速率控制策略的编码输出比较,显然新的速率控制策略的输出稳定性较 TMN8 策略更平稳。

编码输出的稳定性还体现在缓冲器的占有率方面。由于低时延通信中的时延一般要求不超过 100ms<sup>[15]</sup>,所以实验中令  $M = 0.1R$ ,其中  $M$  为缓冲器的最大容量,  $R$  为信道速率。

图 6(a)与(b)分别给出了新的速率控制算法应用于 *Claire* 和 *Salesman* 前 70 帧时,缓冲器的占有率的变化情况。如果占有率大于 1 时,即发生了上溢,这时需要进行跳帧,影响视频的连续性;当占有率降为 0 时说明发生了下溢,这样便会造成信道资源的浪费。从图中可以看出,由于缓冲器的容量较小,所以占有率的变化幅度较大,但是没有上溢的情况发生,所以不会引起跳帧。但是在某个时刻产生了短暂下溢,不过发生的频率较低,不会对信道资源造成较大的浪费。

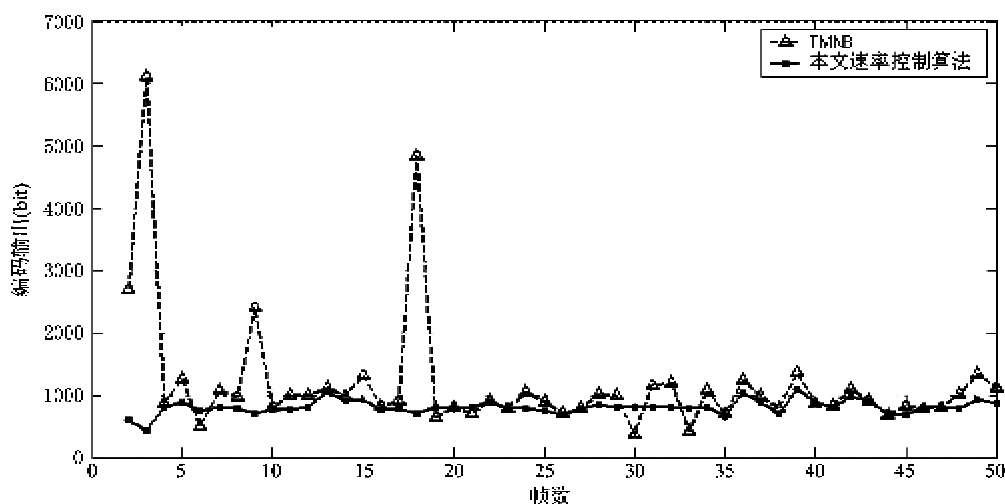


图5 不同模型下的码率输出比较

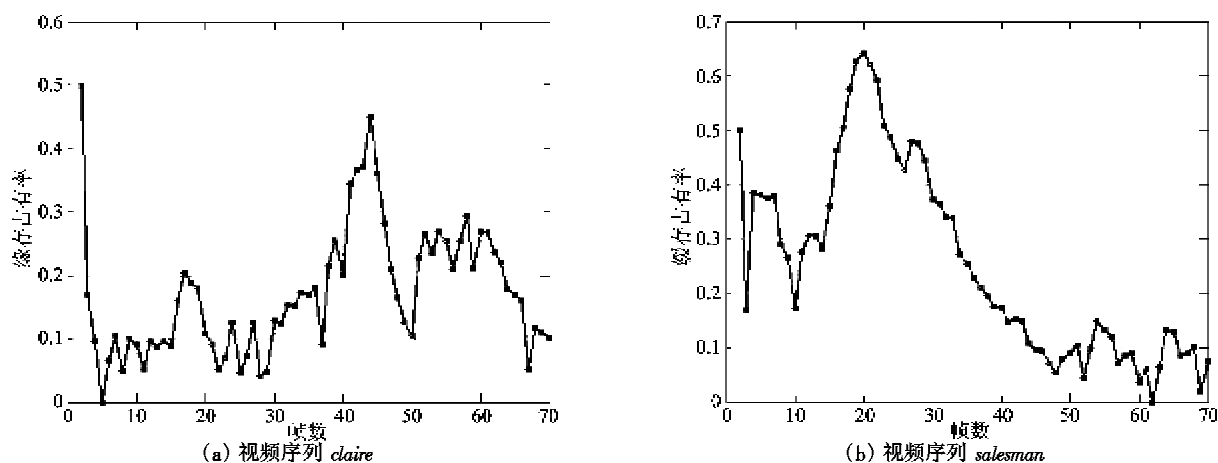


图6 传输编码序列时的缓冲器占有率

### 3.2 算法复杂度与时延分析

时延是实时通信中的一个重要的指标,而算法的复杂度又是造成系统时延的主要因素。在 TMN8 算法中,由于宏块层的量化步长通过统一的公式给出,在对每一宏块进行编码时,可以直接利用式(13)求取。而在新控制算法中,在对每一帧进行编码之前,需要先根据图像的特性确定它的速率模型,然后再根据模型求得帧层的量化步长。从这一方面考虑,新速率控制算法的复杂度较高。但实质上来说在 TMN8 算法中,每一个宏块的量化都要先求得其方差,再利用式(13)求得对应的量化步长,因此含有  $N$  个宏块的视频图像,在使用 TMN8 算法确定编码宏块量化步长时,其复杂度是  $N$  次  $M$  大小宏块方差计算复杂度和  $N$  次式(13)的计算复杂度。

对于本文提出的新速率控制算法,只需完成一次帧级量化步长的确定,其计算量是进行一次式(12)的计算量。而对于算法中根据帧级量化步长

在宏块层进行的简单调整,这一步骤并不是所有宏块都要进行的。即使每个宏块都有进行调整,也仅是进行  $N$  次比较和  $N$  次加法的计算量。由此可以看出,本文算法的计算复杂度要远小于 TMN8 算法的计算复杂度。

此外,本文算法也考虑了模型参数更新时耗费的计算量。由于采用了存储库技术,使模型建立的时间缩短,从而进一步降低了速率控制算法的计算复杂度与时延。表2是针对算法时延特性的分析,

表2 算法的平均时延比较( $s$ )

|              | 视频序列<br><i>salesman</i> | 视频序列<br><i>claire</i> | 视频序列<br><i>coastguard</i> |
|--------------|-------------------------|-----------------------|---------------------------|
| 本文速率<br>控制算法 | 0.9499                  | 0.8447                | 2.3681                    |
| TMN8         | 1.6684                  | 1.4324                | 2.5880                    |

仿真针对三种不同的视频序列前 50 帧图像分别在本文速率控制算法与 TMN8 算法下的平均时延进行比较。如表中所示,新算法的时延明显小于 TMN8。

## 4 结 论

本文提出了一种针对图像内容进行自适应调整的速率-量化模型,并在该模型的基础上给出了新的速率控制算法。该算法以 TMN8 算法的帧层控制为基础,提出利用每一宏块中大于帧层量化步长的 DCT 系数的个数,进行宏块层的目标比特数的分配,并依据实际编码比特数与分配的比特数差值对下一宏块的量化步长进行调整。该方法有效避免了对每一宏块进行的方差计算,减小了计算复杂度和系统时延。仿真实验表明基于自适应速率模型的低时延视频编码速率控制算法较 TMN8 算法来说,具有平稳的编码比特输出、低的复杂度及较小的处理时延。

### 参考文献

- [ 1 ] Artigas X, Tagliasacchi M, Torres L, et al. Analysis of the coset statistics in a distributed video coding scheme. In: Proceedings of the 2nd Workshop on Immersive Communication and Broadcast Systems, Berlin, Germany, 2005. 24-28
- [ 2 ] Ribas-Corbera J, Shawmin L. Rate control in DCT video coding for low-delay communications. *IEEE Trans on Circuits and Systems for Video Technology*, 1999, 9(1): 172-185
- [ 3 ] Lam EY. Analysis of the DCT Coefficient Distributions for Document Coding. *IEEE Signal Processing Letters*, 2004, 11(2): 97-100
- [ 4 ] Kamaci N, Altunbasak Y, Mersereau R M. Frame bit allocation for the H. 264/AVC video coder via Cauchy-density-based rate and distortion models. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 15(8): 994-1006
- [ 5 ] 赵波,吴成柯.一种新的低时延视频编码码率控制算法.计算机学报,2005,28(1):53-59
- [ 6 ] Jyi-Chang T, Chaur-Heh S. Modified TMN8 rate control for low-delay video communications. *IEEE Trans on Circuits and Systems for Video Technology*, 2004, 14(6): 864-868
- [ 7 ] Do-Kyoung K, Mei-Yin S, Kuo C-CJ. Rate control for H.264 video with enhanced rate and distortion models. *IEEE Trans on Circuits and Systems for Video Technology*, 2007, 17(5): 517-529
- [ 8 ] Tihao C, Ya-Qin Z. A new rate control scheme using quadratic rate distortion model. *IEEE Trans on Circuits and Systems for Video Technology*, 1997, 7(1): 246-250
- [ 9 ] Hwangjun S, Kuo C-CJ. A region based H.263 + codec and its rate control for low VBR video. *IEEE Trans on Multimedia*, 2004, 6(3): 489-500
- [ 10 ] Wei D, Bede L. Rate control of MPEG video coding and recording by rate-quantization modeling. *IEEE Trans on Circuits and Systems for Video Technology*, 1996, 6(1): 12-20
- [ 11 ] He Z H, Yong K K, Mitra S K. Low-delay rate control for DCT video coding via  $\rho$ -domain source modeling. *IEEE Trans on Circuits and Systems for Video Technology*, 2001, 11(8): 928-940
- [ 12 ] 万帅,常义林,杨付正等.一种新的视频编码二次率失真模型及其性能分析.电子与信息学报,2007,29(5): 1136-1139
- [ 13 ] Video Expert Group. Video Codec Test Model Near-Term Version 8 (TMN8) Release 0. ITU-T Study Group 16, 1997: Q15-A-59

## An adaptive rate model based rate control algorithm for video coding in low-delay communications

Huo Yan<sup>\* \*\*</sup>, Li Shenghong<sup>\*</sup>, Jing Tao<sup>\*\*</sup>, Feng Yumin<sup>\*\*</sup>

(\* Electronic Engineering of Shanghai Jiaotong University, Shanghai 200240)

(\*\* School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044)

### Abstract

According to the relationship between the shape parameter of the rate model and the image content, a novel rate control algorithm based on an adaptive rate model in video coding for low-delay communications is presented. Compared with other rate models, the model can fit the real coding rate better, and the output rate is more stability by using the rate control scheme based on the model. In order to avoid the computational complexity introduced by the calculation of macroblock (MB) variance in Test Model Near-Term 8 (TMN8), the basic quantization step of video image coding is calculated in frame level and adjusted in the MB level so as to reduce the system processing delay. The results of experiments show that the proposed rate control algorithm based on the adaptive rate model is better than TMN8 in the aspect of stability of output rate and system delay.

**Key words:** mean absolute difference(MAD), adaptive rate model, quantization step, rate control