

# 基于可分级编码基本层的码率控制方法

宋春霞 熊红凯 余松煜 宋利

(上海交通大学 图像通信与信息处理研究所, 上海 200240)

**摘要** 在  $\rho$  域的码率控制中, 由于编码比特  $R$  和一幅图像的变换量化系数中的零的百分比之间的关系起着非常重要的作用, 因此为了更好地进行码率控制, 首先研究了线性失真函数在可分级视频编码 (scalable video coding, SVC) 中基本层上的可适用性, 然后分别对模型参数进行统计, 结果表明, 统计曲线显示了在同一个时间子带上的帧有着相近的线性模型参数。结合以上特征, 通过最小二乘法实现了线性模型在 SVC 基本层上的码率控制应用, 并且在 JSVM1.0 上实现了该算法。实验表明, 该码率控制算法很有效, 与 TM5 算法相比, 峰值信噪比 (Peak signal noise ratio, PSNR) 增加。

**关键词** 码率控制 可分级编码 视频压缩

**中图分类号:** TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2006)11-1631-05

## Rate Control Scheme for Base layer in Scalable Video Coding

SONG Chun-xia, XIONG Hong-kai, YU Song-yu, SONG Li

(Image communication and signal processing graduate school, Shanghai Jiao Tong University, Shanghai 200240)

**Abstract** The relationship of the coding bit rate  $R$  and  $\rho$ , the percentage of zeros in a frame, plays an important role in  $\rho$  domain source modeling rate control. The R-D model is verified for Scalable Video Coding (SVC) at first. Then we studied the characteristic of the  $\rho$ - $R$  curve of pictures when encoding various test sequences with scalable video coding. We find that the model parameter of the pictures at the same temporal level have a close value. Then an adaptive estimation scheme to estimate this model parameter is provided. Combining with its encoding feature, we propose an effective rate control scheme for the base layer of the scalable codec. We implement our algorithm in JSVM 1.0. The experimental results demonstrate that the output bit stream of encoder is controlled accurately.

**Keywords** rate control, scalable video coding (SVC), video compression

## 1 引言

可分级编码是视频编码中达到最新技术发展水平的编码技术之一, 其主要由 3 个层次的可伸缩性技术组成: 空间域可伸缩性、时间域可伸缩性和信噪比 (signal noise ratio, SNR) 可伸缩性。空间域可伸缩性编码提供了空间分辨率不同的码流, 构成了基本层和增强层, 其中增强层可以有 multiple 层。如果在每一个空间层上, 先对码流进行运动补偿时间滤波 (motion compensated temporal filtering, MCTF), 以产生时间域的基本层和若干增强层码流, 然后编码, 则

可将视频帧基本层重建为最低质量帧, 而质量增强层是对基本层的补充, 它是先重复减小的量化步长, 然后应用修正的算术编码 (context-adaptive binary arithmetic coding) 进行熵编码生成的。

通过等级 B 帧或者运动补偿时间滤波可形成时间域的可伸缩性, 而基本层与 H264/AVC 的 Main Profile 的兼容性就是在时间域的基本层上可以同时采用等级 B 帧的技术, 与 MCTF 相比, 等级 B 帧省去了更新的步骤, 只剩下单纯的预测编码结构, 由于这个结构可以用 H264/AVC 的语法来表示, 所以就天然的与 H264/AVC 兼容。

在空间基本层上 (如图 1 所示), 第 1 帧与其他

收稿日期: 2006-06-20; 改回日期: 2006-08-03

第一作者简介: 宋春霞 (1982 ~ ), 女, 2004 年获西安电子科技大学学士学位, 现为上海交通大学图像通信所硕士研究生。主要研究方向为信源编解码。E-mail: cxsong@sjtu.edu.cn

层相同,都是独立编码的 IDR (instantaneous decoding refresh) 帧,其后的帧都是应用了等级 B 帧的概念,即都是以“B...BP”“B...BI”为一个图像组 (group of pictures, GOP) 进行编码的。每个 GOP 内,图像的编码顺序为“AB<sub>1</sub>B<sub>2</sub>B<sub>3</sub>B<sub>4</sub>B<sub>5</sub>B<sub>6</sub>B<sub>7</sub>B<sub>8</sub>I/P<sub>0</sub>B<sub>1</sub>B<sub>2</sub>B<sub>3</sub>B<sub>4</sub>B<sub>5</sub>B<sub>6</sub>B<sub>7</sub>B<sub>8</sub>I/P<sub>0</sub>”,其中 A、I、P 都是关键帧。第 1 级 B 帧标志为 B<sub>1</sub> 只能用周围的关键帧来做运动补偿预测,第 i (i > 1) 级 B 帧标志为 B<sub>i</sub> 既可以用周围的关键帧来预测,也可以用同一个 GOP 内的 B<sub>j</sub> (j < i) 帧运动补偿预测。

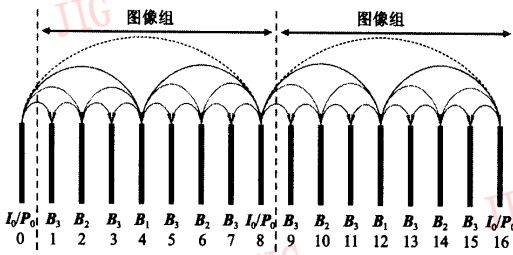


图 1 等级编码结构,4 个时间层, GOP 为 8, 每个 B 帧都有两个参考帧做预测

Fig. 1 Dyadic hierarchical coding structure with 4 temporal levels and a GOP size of 8. Each B picture is predicted by using 2 reference pictures

当前的 SVC (scalable video coding) 系统中,每个 SNR 层上用固定的量化参数 (quantize parameter, QP) 对图像进行量化编码,在基本层上设置一个量化参数初始值 QP<sub>0</sub>,相邻的两个层之间的量化参数相差一个常数,也就是说,量化时用的量化参数都取决于 QP<sub>0</sub>。如果想达到某个目标码率,则必须不断调整 QP<sub>0</sub>,以满足信道的需要,并且基本层信息非常重要,因为它是整个 SVC 结构中最基础部分,而且空间增强层会用到基本层上的信息进行预测,所以对整个编码系统也有着很大的影响。因此,为了防止缓存上溢或下溢,同时为了保证在缓存变化平缓的前提下,取得更好的视频编码质量,码率控制成为 SVC 系统中的重要研究课题之一。

文献[1~3]中的信源模型均采用量化步长 Q 来表述率失真关系,而文献[4,5]则发现 DCT 系数量化后零系数的百分比 (以 ρ 表示) 与输出码率关系较 Q 更为密切,设 ρ 随 Q 单调增加,则由此可建立起二者之间的一一对应关系,这样算法就给出一种简单有效的线性 R-ρ 模型,即先由 r 来预测输出码率,再将 ρ 转化为 Q 进行量化编码,该模型可适用于 MPEG-2、

H. 263 和 MPEG-4 等多种视频压缩编码标准。

R 和 ρ 的关系在码率控制体系中起着至关重要的作用。以前大部分的研究都集中在 I 帧和 P 帧上,认为有 B 帧在内的码率控制是非常困难的,因为 B 帧包含了各种各样的模式,比如说 direct、bidirect、前向、后向等,而且一般来说,B 帧的头信息和运动向量比 I 帧和 P 帧都大。本文针对 SVC 基本层,研究了怎样准确有效地估计 ρ 域模型中的参数,并在此基础上提出了帧层的码率控制算法。

## 2 等级 B 帧的线性率失真模型参数研究

在视频编码中,由于“0”占有很重要的地位,因此通常需进行变换,使大部分系数变为零后再进行量化,零的百分比定义为 ρ,在 QP 和 ρ 之间建立起一一对应的关系,因此可以把 R(QP) 和 D(QP) 映射到 ρ 域,变成 R(ρ) 和 D(ρ)。文献[4,5]已证明了 I/P 帧的 R-ρ 呈线性关系。下面用量化参数 0 到 51 来对不同序列中的同一等级 B 帧进行量化,即得到零的百分比 ρ 和比特率 R (单位: bpp) 的关系,需要声明的是,这里的比特数包含头信息。本文给出了 Bus 和 Football 标准测试序列中两帧不同图像的 ρ-R 曲线 (图 2)。

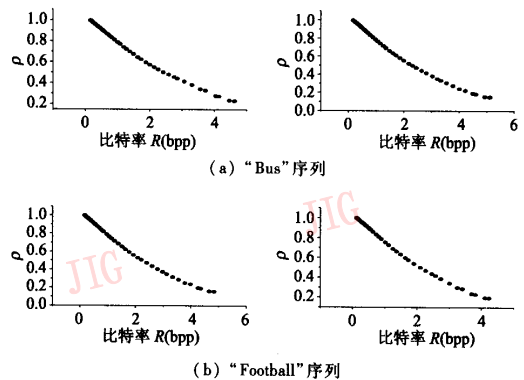


图 2 等级 B 帧中 ρ 与 R 的关系

Fig. 2 ρ - R (bpp) relationship for hierarchical B pictures

由图 2 可见从每个采样图像中观察到的 ρ-R 几乎为一条直线,因此可以表示为下式所示的线性关系:

$$\rho = \omega \cdot R + \psi \tag{1}$$

ω 和 ψ 为待定参数,同时注意到因为包括头信息的缘故,当 ρ 为 1 时,R 并不为 0,因此,针对等级 B 帧需要研究待定参数的特征,运行 MPEG SVC 参考软

件 JSVM1.0<sup>[6]</sup>,把视频测试序列“Bus”,“Football”,“Foreman”和“Mobile”中所有的等级 B 帧都用 0~51 量化参数分别进行量化,先得到如图 2 所示的曲线,然后进行线性拟合,求出拟合后直线的斜率和截距如图 3 和图 4 所示。等级 B 帧的编码顺序为  $B_1, B_2, B_2, B_3, B_3, B_3, \dots$ 。

从图 3 中可以看出,各个序列的斜率变化各不

相同,“Bus”和“Mobile”序列的斜率与“Football”和“Foreman”序列的斜率相比,变化平缓些。同时,注意到在同一 GOP 内,同一时间子带内的等级 B 帧的模型参数值几乎一样,而且不同 GOP 中的大部分在同一个时间域子带上帧的斜率和截距都有相近的值。因每个像素所占比特数在低码率时是一个小于 1 的数,所以截距对  $\rho$  的影响比斜率大。

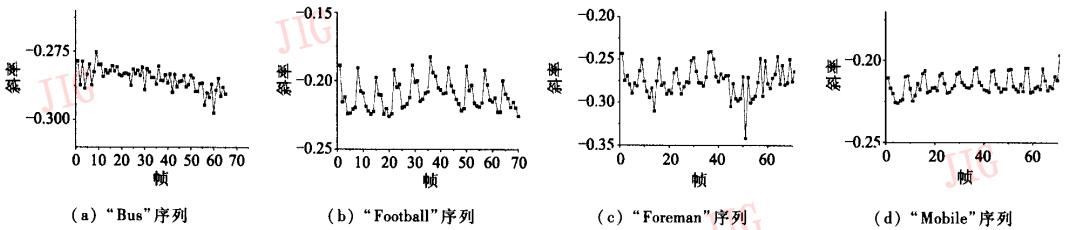


图 3 各视频序列斜率变化曲线

Fig. 3 Curves of the slope for each video sequence

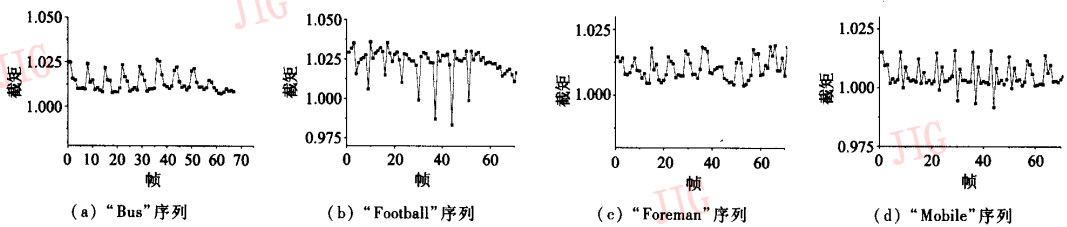


图 4 各视频序列截距变化曲线

Fig. 4 Curves of the intercept for each video sequence

### 3 码率控制算法

根据前述对等级 B 帧的线性率失真函数的验证以及线性模型中的参数特征分析,本文对 SVC 基本层提出了一种帧级的码率控制算法。

由于在 SVC 的基本层上,每个 GOP 中只有一个 I 帧或 P 帧,其编码过程与 H.264 中的 I、P 帧相同,所以对于这些帧的码率控制策略可参考文献[7]。现以恒定码率编码为例,等级 B 帧的码率控制算法步骤如下:

(1) 对位于整个序列的每个时间子带上的前两个等级 B 帧进行量化时用的量化参数是固定的,则转至步骤(4);除此之外,转至步骤(2)。

(2) 在以固定比特编码一帧图像时,首先用式(1)预测每一帧在量化后需要的零的百分比  $\rho(R_t)$ ,  $R_t$  是时间子带  $t$  上可选择的码率。其中的  $\omega$  和  $\psi$

用的是同一个时间子带的上一帧编码后在步骤(4)中更新后的  $\omega$  和  $\psi$ 。

(3) 计算  $QP$ 。等级 B 帧在编码时有 intra 和 residual 两种模式,其中 intra 模式需要进行模式 (INTRA\_4×4, INTRA\_8×8 和 INTRA16×16) 选择,而 residual 模式则不需要这个过程,其大部分的宏块都是用这个模式编码的。为简单起见,可在对每一帧图像进行编码前,通过对标志为 residual 模式编码的宏块进行预变换来得到系数分布,并由此来预测整帧图像中的系数分布。根据步骤 2 中得到的  $\rho$  值大小,由式(2)表示的  $\rho$ - $QP$  关系中计算出  $QP$ :

$$\rho(QP) = \frac{1}{M} \sum_{i,j} \sum_{|x_{i,j}| < T(i,j,QP)} D_{i,j}(x_{i,j}) \quad (2)$$

$M$  为当前帧 residual 模式编码的宏块中总的系数个数,  $x_{i,j}$  是在  $4 \times 4$  矩阵中位于  $(i, j)$  处的系数,  $T(i, j, QP)$  为参考文献[7]中提出的系数门限值,在固定的量化参数下,若  $x_{i,j} < T(i, j, QP)$ ,则表示  $x_{i,j}$  用这

个量化参数量化后变为零。

(4)对线性率失真函数中的参数进行更新。由于利用第 2 节中阐述的在同一个时间子带上模型参数有相似的特征,因此可用最小二乘法对率失真参数进行预测。在编完每一帧后,就可以得到实际的编码比特率  $R(\text{bpp})$  和零的百分比  $\rho$ ,为了使下式最小

$$\min \left[ \sum (\rho - \hat{\rho})^2 \right] \tag{3}$$

其中,  $\hat{\rho}$  为率失真函数输入为  $R$  时,零的百分比的预测值,应更新每个时间子带上的  $\omega$  和  $\psi$ :

$$\theta = \frac{\sum_{i=1}^n (R_i - \bar{R})(\rho_i - \bar{\rho})}{\sum_{i=1}^n (R_i - \bar{R})^2} \tag{4}$$

$$\psi = \bar{\rho} - \omega \bar{R} \tag{5}$$

为了使计算简单并且在参数趋势变化时能够动态地进行跟踪,可采用滑动窗技术,滑动窗的长度为  $n$ ,窗内记录了该子带内最新  $n$  帧的实际  $R$  和  $\rho$  值,  $\bar{R}$  和  $\bar{\rho}$  分别是窗内  $R$  和  $\rho$  的平均值,而且为了减少更新的次数,还设置了更新的门限:

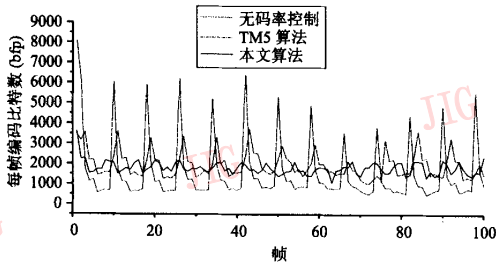
$$\text{If } ((\rho - \hat{\rho}) > \beta) \{ \text{updating} \}$$

$\beta$  是实验值。重复步骤(1)~步骤(5),至编码结束。

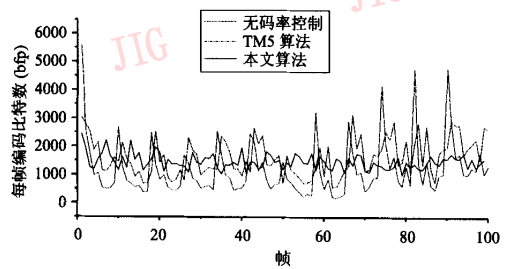
### 4 实验结果

基于软件 JSVM1.0,对本文给出的程序进行了实验,实验中使用了多个 QCIF 格式的标准视频测试序列,这里仅给出“Foreman”和“Bus”序列的实验结果,码率分别为 22.5kbps 和 27kbps,帧频 15fps,GOP 大小为 8,时间域分为 3 级,每级的码率相同,每个时间子带的第 1 和第 2 个等级 B 帧的量化参数分别设置为 46 和 40,滑动窗的长度为 20, $\beta$  为 0.002。

为了公平比较,在同样条件下对 JVT-F086<sup>[8]</sup> 的 TM5 算法进行测试。从图 5 可见,本文算法与 TM5 算法相比,所产生的实际码率更加接近目标码率。图 6 给出了加本文算法和 TM5 算法每一帧的峰值信噪比(peak signal noise ratiar, PSNR)对比,由图 6 可见,本文算法编码图像的峰值信噪比变化比较平缓,不仅大大减少了图像的质量波动,而且有较高的主观视觉质量,其峰值信噪比平均提高了 0.8dB 左右。



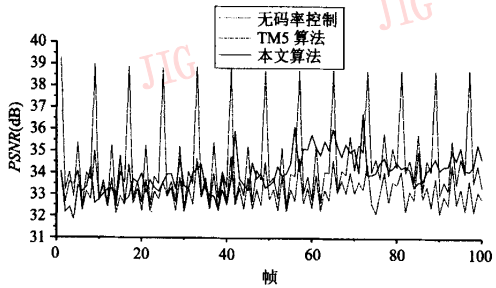
(a) “Bus”序列



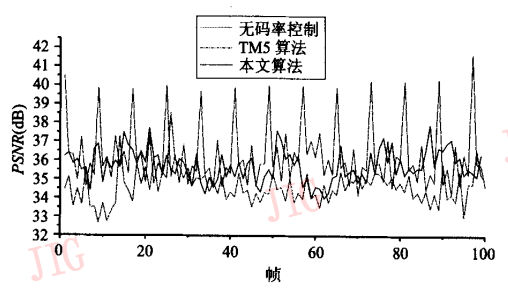
(b) “Foreman”序列

图 5 “Bus”及“Foreman”视频序列每帧编码比特数

Fig.5 Coding bits per frame of “Bus” and “Foreman”



(a) “Bus”序列



(b) “Foreman”序列

图 6 “Bus”及“Foreman”视频序列的峰值信噪比

Fig.6 PSNR of each frame in “Bus” and “Foreman”

## 5 结 论

本文不仅验证了在可分级视频编码中的等级 B 帧的  $\rho$  域率失真模型符合线性关系,并且研究了该模型的参数特征,同时利用在同一个时间子带上的图像有很近的参数值这一特征,把视频图像按照时间域可伸缩性分组分别进行处理,提出了帧层的码率控制算法。实验结果证明本算法对 SVC 的 SNR 基本层非常有效。在可分级编码中,通常用位平面技术实现 FGS (fine granular scalable),因为每加一个位平面就等于减少原来量化参数的一半,相应的失真应减少为原失真的四分之一,因此质量增强层的 PSNR 变化也比较平滑。

### 参考文献 (References)

- 1 Ding W, Liu B. Rate control of MPEG video coding and recording by rate-quantization modeling[J]. IEEE Transactions on Circuits System Video Technology, 1996, 6(1): 12 ~ 20.
- 2 Chiang T, Zhang Y Q. A new rate control scheme using quadratic rate distortion model [J]. IEEE Transactions on Circuits System Video Technology, 1997, 7(1): 246 ~ 250.
- 3 Ribas Corbera J, Lei S. Rate control in DCT video coding for low-delay communications [J]. IEEE Transactions on Circuits System Video Technology, 1999, 9(1): 172 ~ 185.
- 4 He Zhi-hai, Kim Y K, Mitra S K. Low-delay rate control for DCT video coding via  $\rho$ -domain source modeling [J]. IEEE Transactions on Circuits System Video Technology, 2001, 11(8): 928 ~ 940.
- 5 He Zhi-hai, Mitra S K. A linear source model and a unified rate control algorithm for DCT video coding [J]. IEEE Transactions on Circuits System Video Technology, 2002, 12(11): 970 ~ 982.
- 6 Julien Reichel, Mathias Wien, Heiko Schwarz. JSVM 1.0 Software [S]. ISO/IEC JTC1/SC29/WG11/N6900, Hong Kong, China, Jan. 2005.
- 7 Shin Il-Hong, Lee Yung-Lyul, Park Hyun Wook. Rate control using linear rate- $\rho$  model for H. 264 [J]. Signal Processing: Image Communication, 2004, 19: 341 ~ 352.
- 8 Ma S, Gao W, Lu Y. Improved rate control algorithm [A]. In: Joint Video Team of ISO/IEC and ITU 6<sup>th</sup> meeting [C], JVT-F086, Awaji, Island, Japan 2002.