

基于最佳量化的序列图像编码转换码率控制策略¹

李晓辉 吴小培

(安徽大学计算智能与信息处理教育部重点实验室 合肥 230039)

摘 要 该文分析了编码转换缓冲区的状态, 导出了编码转换缓冲区为防止解码器缓冲区下溢和上溢应满足的条件, 建立了序列图像编码转换模型. 并根据编码转换缓冲器的状态和信道速率, 为待编码帧在图像层上预分配目的序列图像编码比特数, 使用 DCT 系数分布特性来表征图像特性. 继而为帧内每一具体宏块选定最佳量化因子, 提出了基于最佳量化的码率控制策略. 模拟实验表明, 该码率控制策略能有效地减少、避免缓冲区出现上、下溢的情况, 使输出码率趋于稳定, 提高了重建序列图像的信噪比.

关键词 序列图像, 编码转换, 最佳量化, 码率控制

中图分类号 TN919.8

1 引 言

由于通信网络的带宽不同, 因而对应着不同的传输码率, 这就使得不同带宽网络之间的接入问题成为信息交换时应首先考虑的重要环节. 当已压缩编码的序列图像信号由较高传输速率的信道接入到较低速率的信道时, 将会导致传输码流与信道失配, 从而引起网络拥塞, 使得序列图像信号的传输无法正常进行. 因此, 需要采用编码转换器对传输码流进行码率转换, 以保证序列图像码流在网络间的正确传输, 为不同网络的用户提供不同服务质量的视频服务^[1,2].

2 编码转换缓冲区分析

编码转换的瞬时压缩率通常不能随编码器的瞬时压缩率同步变化, 因此在编码转换器中需要一个小的编码转换缓冲区来平滑编码转换器输出的误差^[3,4]. 序列图像编码转换模型如图 1 所示. 设编码缓冲区的大小为 M , 编码转换器的缓冲区大小为 M_T . R_1 和 R_2 分别是编码转换器输入和输出端的恒定传输码率. 编码后的可变比特率 (VBR) 码流 $r(t)$ 进入编码缓冲区, 然后变成码率为 R_1 的恒定比特率 (CBR) 码流后, 再传送到编码转换器, 通过编码转换器进行编码转换下降为 R_2 的 CBR 码流. 该码流经解码缓冲区输出为 VBR 码流以供解码.

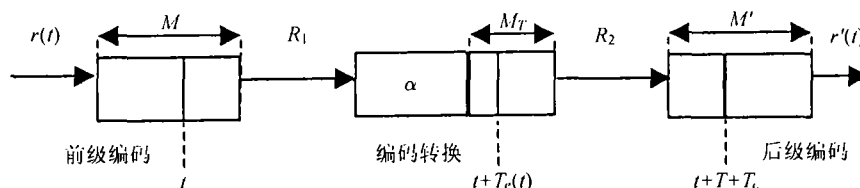


图 1 序列图像编码转换模型

编码转换压缩比 α 定义为

$$\alpha = R_2 / R_1 \quad (1)$$

后级解码缓冲区的大小 M' 为

$$M' = \alpha M + M_T \quad (2)$$

¹ 2001-10-10 收到, 2002-04-12 改回
安徽省自然科学基金资助 (03042209)

编码转换后的 VBR 输出 r' 为

$$r'(t) = \alpha r(t) \quad (3)$$

设 $M_e(t)$, $M_t[t + T_e(t)]$ 以及 $M_d(t + T + T_c)$ 分别为 t , $t + T_e(t)$ 以及 $t + T + T_c$ 时刻解码器缓冲区的占用量。宏块编码后在 t 时刻进入前级编码器缓冲区, 等待 $M_e(t)/R_1$ 秒后被传送到编码转换器缓冲区; 在编码转换器缓冲区等待 $M_t[t + T_e(t)]/R_2$ 秒后被传送到后级解码器缓冲区; 又等待 $M_d[t + T + T_c]/R_2$ 秒后被解码。因此可推导出宏块从编码到解码的总延迟 T_{tot} 为

$$\begin{aligned} T_{\text{tot}} &= M_e(t)/R_1 + M_t[t + T_e(t)]/R_2 + M_d(t + T + T_c)/R_2 \\ &= \left[\int_0^t r(\tau) d\tau - R_1 t \right] / R_1 + \left\{ \int_0^t r'(\tau) d\tau - R_2 [t + T_e(t)] \right\} / R_2 \\ &\quad + \left[R_2 (t + T + T_c) - \int_0^t r'(\tau) d\tau \right] / R_2 \end{aligned} \quad (4)$$

它的简化形式为

$$T_{\text{tot}} = T_e(t) + \frac{1}{R_2} \int_0^t r'(\tau) d\tau - t - T_e(t) + T + T_c - \frac{1}{R_2} \int_0^t r'(\tau) d\tau = T + T_c \quad (5)$$

由 (5) 式可见, 总延迟是由初始解码器给出的常量。由于包括了不同 CBR 的码流, 3 个缓冲区的占用量大小不是常量。

为防止解码器缓冲区的下溢或上溢, 宏块编码后经过前级编码器缓冲区的延迟时间与经过编码转换器缓冲区的延迟时间以及经过后级解码器缓冲区的延迟时间之和 (即系统总延迟 T_{tot}), 应等于解码器缓冲区 M' 的抽空时间。

$$M_e(t)/R_1 + M_t[t + T_e(t)]/R_2 + M_d(t + T + T_c)/R_2 = M'/R_2 \quad (6)$$

经公式推导得解码器缓冲区不会下溢的条件为

$$M_t(t) \leq M_T \quad (7)$$

为防止解码器缓冲区上溢, 充分条件是在任一时刻满足

$$M_t(t) \geq 0 \quad (8)$$

编码转换以宏块为单位进行处理, 一个宏块无论何时被接收到, 它都立即被编码转换, 然后被送往输出缓冲区。如编码器和编码转换缓冲区没有下溢或上溢, 那么解码器缓冲区也不会发生上溢或下溢。其基本思想是通过增加编码器和解码器的总延迟, 为此在编码转换器中用一个缓冲区, 它起到轮流防止上溢或下溢的作用。

3 编码转换码率控制策略

编码转换缓冲区控制的主要手段是改变编码转换器中编码器的量化系数, 从而改变编码转换的输出码率。编码转换缓冲区控制方法就是要合理改变编码转换的输出码率, 使编码转换缓冲区达到动态平衡^[5]。

本文提出的码率控制策略是通过前向和后向控制相结合的途径来达到码率控制目的的^[6]。它首先在总体上对每帧图像的编码比特数进行预分配, 此过程称为图像层码率控制。然后在对每帧图像具体编码时, 又根据当前图像的具体特征分别对待, 为帧内不同特征的宏块选用不同的

量化因子, 更为合理地在帧内分配比特数, 从而使得在编码转换输出码流趋于稳定的同时, 编码转换后图像的主观质量也有所提高, 尤其是在重要的细节部分。该过程称为宏块层码率控制。

图像层码率控制实际上是根据信道速率和缓冲器状态, 在图像层上为每一帧图像确定目的编码比特数 B , 并由此确定图像层的参考量化因子, 这是一种宏观预分配。

在宏块层为帧内每个宏块确定具体的量化步长时, 应充分考虑图像本身的特征。本文中用 DCT 系数的数学分布来表征图像特征^[7]。DCT 系数反映了相邻两帧的变化率, D 愈大, 说明相邻帧的相关性愈弱, 编码产生的比特位愈多。为了实现码率控制的目的, 应适当增加量化步长 Q_i 的值。对一宏块编码需要的比特位可用 (9) 式计算

$$B_i = A(Z\sigma_i^2/Q_i^2 + C) \approx A(ZE_i^2/Q_i^2 + C) \quad (9)$$

式中 B_i 是对第 i 个宏块进行编码时产生的比特数, A 是一个宏块内的像素数, E_i 为当前宏块 DCT 系数的数学期望, Q_i 为当前宏块的量化系数, σ_i^2 是第 i 宏块 DCT 系数的标准偏差, Z 是一个常数, C 为实际编码中组织成帧、编码运动矢量、帧头信息以及句法结构所占用的比特数。在第 i 个宏块中, 用量化步长为 Q_i 的量化器对 DCT 系数进行量化, 势必引入一定的失真。对当前编码帧进行量化产生的失真 σ_q^2 为

$$\sigma_q^2 = \frac{1}{N} \sum_{i=1}^N \frac{9}{2} \sigma_i Q_i^2 \quad (10)$$

式中 N 是一帧内的宏块数目。为了实现在保证失真最小的情况下为帧内每一宏块确定量化系数 Q_i^* , 所有宏块产生的比特数 $\sum_{i=1}^N B_i$ 应该等于在图像层上分配的目的比特数 B 。因此有下式成立。

$$\left. \begin{aligned} Q_1^*, \dots, Q_N^* &= \operatorname{argmin}_{Q_1, \dots, Q_N} \frac{1}{N} \sum_{i=1}^N \frac{9}{2} \sigma_i Q_i^2 \\ \sum_{i=1}^N B_i &= B \end{aligned} \right\} \quad (11)$$

(11) 式表示在 $\sum_{i=1}^N B_i = B$ 时保证 (10) 式最小, 即失真最小。采用 Lagrange 定理, 定义一个 Lagrange 常数 λ , λ^* 是最佳值。对 (11) 式做相应变换, 得

$$\begin{aligned} Q_1^*, \dots, Q_N^*, \lambda^* &= \operatorname{argmin}_{Q_1^*, \dots, Q_N^*, \lambda} \frac{1}{N} \sum_{i=1}^N \frac{9}{2} \sigma_i Q_i^2 + \lambda \left[\sum_{i=1}^N B_i - B \right] \\ &= \operatorname{argmin}_{Q_1^*, \dots, Q_N^*, \lambda} \frac{1}{N} \sum_{i=1}^N \frac{9}{2} \sigma_i Q_i^2 + \lambda \left[\sum_{i=1}^N A(K \frac{\sigma_i^2}{Q_i^2} + C) - B \right] \end{aligned} \quad (12)$$

对上式中 Q_i , λ 求偏导, 并令偏导表达式为 0, 得到一个方程组。解此方程组, 最终得到最佳量化系数 Q_i^* 的表达式为

$$Q_i^* = \sqrt{\frac{AZ}{B - ANZ} \sigma_i^{\frac{1}{2}} \sum_{k=1}^N \sigma_k^{\frac{3}{2}}} \quad (13)$$

此式是本文提出的码率控制技术的核心。该式表明： Q_i^* 随宏块标准偏差 σ_i^2 的增加而增加，随着可供编码的目的比特数 B 的增大而减少。

4 实验结果与分析

本文对具有中等运动的 Foreman 测试序列和具有高速运动和少量场景切换的 Coastguard 测试序列进行了计算机模拟。测试序列为准公共中间格式 (QCIF) 格式，每个像素用 8bit 表示。帧频为 30fps，以 64kbps 的速率编码，通过编码转换器编码转换为 32kbps。初始量化步长 $QP=15$ ，编码转换缓冲区容量 $B_T=6400\text{bit}$ 。

采用最佳量化码率控制策略的缓冲区占用量 (实线所示) 和采用 H.261 码率控制策略的缓冲区占用量 (虚线所示) 如图 2 所示。

由图 2(a) 可见，H.261 码率控制策略在当图像变化时，缓冲区占用量波动较大；对于细节丰富的图像帧，将引起编码转换缓冲区占用量的增加。而采用最佳量化码率控制策略时，随着图像的变化，编码转换缓冲区占用量波动较小。使得即使出现细节丰富的图像帧，编码转换缓冲区占用量也较小。尤其在图像序列出现高速运动和少量场景切换时，如图 2(b) 所示，采用 H.261 码率控制策略的编码转换缓冲区占用量出现了较大的波动，且由于图像细节的丰富使得编码转换缓冲区占用量较大；而采用最佳量化码率控制策略的编码转换缓冲区占用量仍能保持较为稳定，即使对于细节丰富的图像帧，也能保持编码转换缓冲区占用量较小。

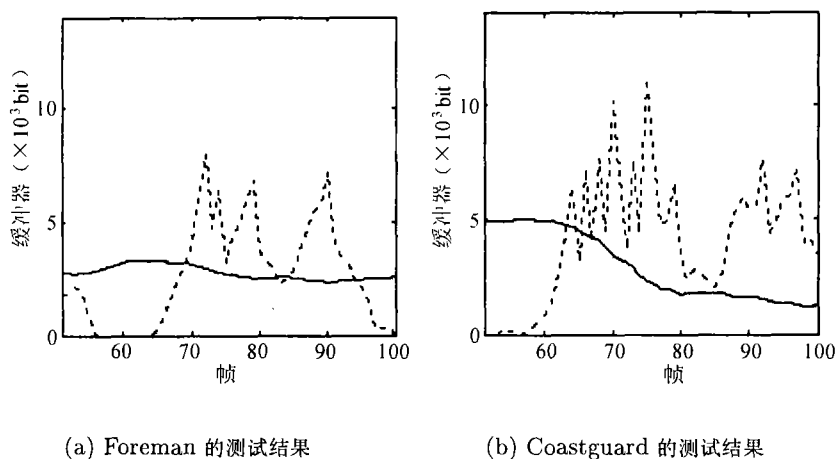


图 2 不同控制策略下缓冲区占用量的比较

采用最佳量化码率控制策略的重建图像峰值信噪比 (实线所示) 与采用 H.261 码率控制策略的重建图像峰值信噪比 (虚线所示) 如图 3 所示。

从图 3 可见，采用 H.261 码率控制策略的 PSNR 低于采用最佳量化编码转换缓冲区码率控制策略的 PSNR。当出现高速运动或场景切换时，采用最佳量化编码转换缓冲区码率控制策略能获得较高的 PSNR。

Coastguard 和 Foreman 测试序列 50~100 帧的重建图像平均 PSNR 的结果如表 1 所示。

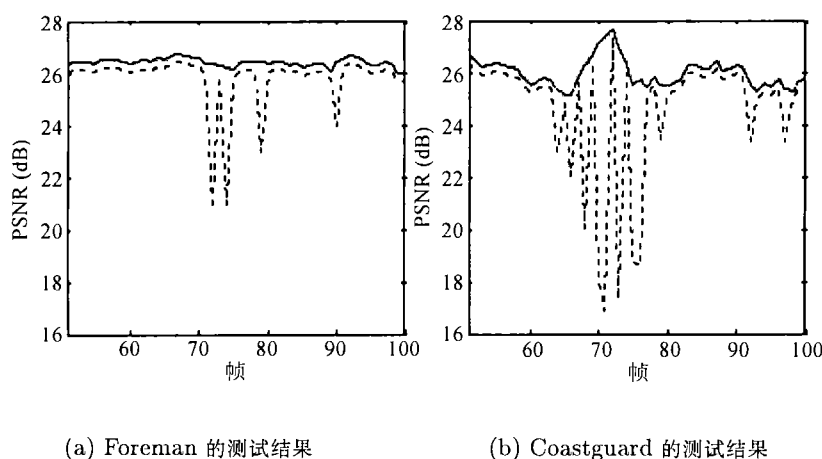


图 3 不同控制策略下重建图像峰值信噪比的比较

表 1 两种码率控制策略下平均峰值信噪比的比较

测试序列	峰值信噪比 (dB)	
	H.261 码率控制策略	最佳量化码率控制策略
Foreman	26.34	26.86
Coastguard	25.26	26.19

从表 1 可知, 采用最佳量化码率控制策略得到的 PSNR 较 H.261 使用的码率控制策略得到的平均 PSNR 要高出 0.5~0.9dB 左右, 尤其是在采用活动性大的 Coastguard 测试序列的情况下, 重建图像的 PSNR 有明显的改善。

不同控制策略下测试序列 Foreman 的重建图像质量的比较如图 4 所示。

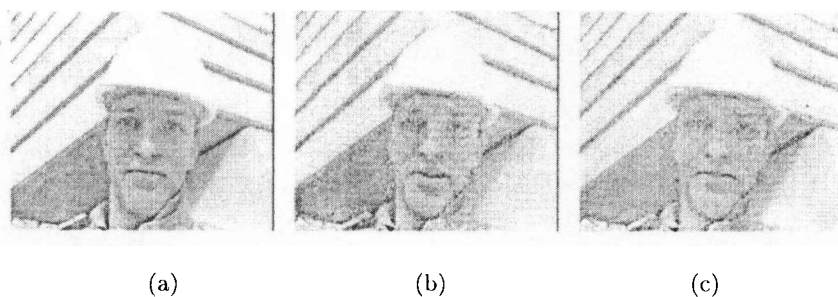


图 4 不同控制策略下重建图像的画面质量的比较

图 4(a) 为 Foreman 测试序列的 90 帧的原始图像。图 4(b) 为使用 H.261 码率控制策略得到的重建图像, 实际上使用的是其前一帧 (第 89 帧) 的图像来代替当前帧。图 4(c) 为使用本文的最佳量化编码转换缓冲区码率控制策略得到的重建图像, 较图 4(b) 更接近于原始图像。

5 结 束 语

编码转换器可以改变序列图像信号的码率, 从而保证序列图像码流在不同网络间的正确传输, 以适应带宽不同的各种接入网, 为不同网络的用户提供不同服务质量的视频服务。采用基于最佳量化的码率控制策略能很好地反映图像的特征, 为量化因子的调节提供较好的判据。由此得到的编码转换输出码流和编码转换缓冲区占用量较稳定, 可较好地防止出现缓冲区上、下溢。由该方法得到的重建图像信噪比较高, 尤其在图像序列出现高速运动或场景切换时, 重建图像的信噪比能得到明显的改善。

参 考 文 献

- [1] S. M. Lei, T. C. Chen, M. T. Sun, Video bridging based on H.261 standard, IEEE Trans. on Circuits and Syst. Video Technol., 1994, 4(4), 425-437.
- [2] B. Christopoulos, Transcoder architectures for video coding, IEEE Trans. on Consumer Electronics, 1998, 4(1), 88-98.
- [3] W. Ding, B. Liu, Rate control of MPEG video coding and recoding by rate-quantization modeling, IEEE Trans. on Circuits and Syst. Video Technol., 1996, 6(1), 12-19.
- [4] S. Zammit, G. Carpenter, Low complexity, low delay H.261/MPEG-1 transcoders, in Proc. Int. Picture Coding Symp., Melbourne, Australia, Mar. 1996, 599-603.
- [5] B. Tao, B. W. Dickinson, H. A. Peterson, Adaptive model-driven bit allocation for MPEG video coding, IEEE Trans. on Circuits and Syst. Video Technol., 2000, 10(1), 147-157.
- [6] T. Chiang, Y. Q. Zhang, A new rate control scheme using quadratic rate distortion model, IEEE Trans. on Circuits and Syst. Video Technol., 1997, 7(1), 246-250.
- [7] I. M. Pao, M. T. Sun, Modeling DCT coefficients for fast video encoding, IEEE Trans. on Circuits and Syst. Video Technol., 1999, 9(4), 608-616.

RATE CONTROL STRATEGY IN SEQUENCE IMAGE
TRANSCODING BASED ON THE QUANTIZER OPTIMIZATION

Li Xiaohui Wu Xiaopei

*(The Key Lab of Intelligent Computing and Signal Processing,
Ministry of Education, Anhui Univ., Hefei 230039, China)*

Abstract In this paper, the status of transcoder buffer is analyzed. The conditions that have to be met by both the encoder and transcoder buffers for preventing the decoder buffer from underflowing or overflowing are derived also. A model of sequence image transcoding is established. A new rate control strategy based on the quantizer optimization is proposed, and the target number of bits firstly on picture layer based on the status of the buffer and the channel rate is allocated. Then the characteristics of the image by the distributed characteristics of the DCT coefficients are denoted and a best quantization parameter for every macroblock in a frame is selected. The experiment shows that this rate control strategy can effectively reduce or avoid the overflow and underflow of the encoder buffer, and the output bit rates tend to stabilization. The Peak Signal-to-Noise Ratio(PSNR) of the reconstructed image is raised.

Key words Sequence image, Transcoding, Quantizer optimization, Rate control

李晓辉: 男, 1961 年生, 教授, 主要研究方向为图像编码和多媒体通信.

吴小培: 男, 1966 年生, 教授, 主要研究方向为信号处理理论及应用.