

# 改进的降低空间分辨率视频编码转换算法

李晓辉

(安徽大学电子科学与技术学院计算智能与信息处理教育部重点实验室 合肥 230039)

**摘要:** 该文提出一种改进的降低空间分辨率视频编码转换算法。根据漂移误差与帧间编码块的运动活动性的关系,提出了一种自适应帧内刷新方法。该方法根据目标比特率和实际比特率的差值,动态地调整阈值,从而有效地减少了帧内编码的比特率,保持了对差错的鲁棒性,限制了差错的时间传播。同时,将率失真函数映射为线性函数,以较少计算复杂度,提出了一种线性速率控制策略。仿真结果表明:该文提出的方法有效地平滑了缓冲器的输出,同时峰值信噪比也有所提高。

**关键词:** 视频编码转换; 空间分辨率; 自适应帧内刷新; 速率控制

中图分类号: TN919.81

文献标识码: A

文章编号: 1009-5896(2007)08-1900-03

## An Improved Algorithm for Reduced Spatial Resolution Video Transcoding

Li Xiao-hui

(The Key Laboratory of Intelligent Computing and Signal Processing, Ministry of Education, School of Electronic Science and Technology, Anhui University, Hefei 230039, China)

**Abstract:** An algorithm of improved reduced spatial resolution video transcoding is proposed in this paper. An adaptive intra-refresh scheme is presented based on the relation between drift error and motion activity of inter-coded blocks. The thresholds are dynamically adjusted according to the difference between the target bit rate and actual bit rate. This method can reduce the bit rate of intra coded frames while still maintaining error robustness and limiting temporal propagation of errors. Besides, a linear rate control strategy is also present. In order to reduce computational complexity, the rate-distortion function can be mapped into a linear function. The simulation results show that the buffer fullness tend to be stabilized and the PSNR of the reconstructed image is raised.

**Key words:** Video transcoding; Spatial resolution; Adaptive intra-refresh; Rate control

### 1 引言

随着多媒体网络数量的增加,视频编码转换方法的研究已成为多媒体通信领域的热点之一。视频编码转换是一种将已压缩的视频码流转换成另一种视频码流的处理过程。由于通信网络的带宽不同,因而对应着不同的传输码率,这就使得不同带宽的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。当已压缩编码的视频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,使得视频信号的传输无法正常进行<sup>[1,2]</sup>。因此,需要采用视频编码转换器对传输码流进行码率转换,以保证视频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。

由于受到信道带宽(无线网络)或终端显示能力的限制,为了获得所需的低速率码率,就必须对视频信号进行高倍压缩。这时,如采用原有的空间分辨率对视频信号进行编码转换,将造成图像质量的严重下降。因此,在视频编码转换中需插入下抽样滤波器,以降低视频信号的空间分辨率,从而

改善图像质量<sup>[3]</sup>。

### 2 漂移误差分析

在视频编解码系统中,由于解码器运动补偿环的误差积累,当解码运动补偿帧间图像时,会产生漂移误差,从而引起图像质量的连续下降<sup>[4]</sup>。帧内编码块具有不受漂移影响的特点,因而采用将帧间编码块转换成帧内编码块的帧内刷新方法,可以阻止漂移的传播。帧内刷新降低分辨率编码转换结构框图如图 1 所示。

设  $e_{in}^{(n)}$  为当前帧原分辨率差值输入信号,  $g_{out}^{(n)}$  为当前帧降低分辨率差值输出信号。则有

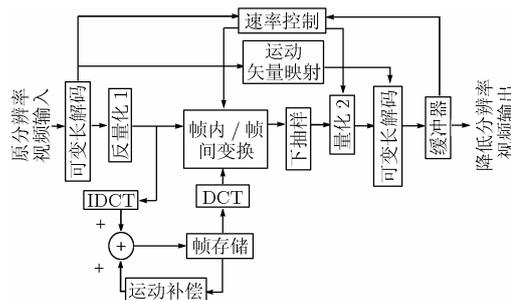


图 1 帧内刷新降低分辨率编码转换结构框图

$$g_{\text{out}}^{(n)} = D_s(e_{\text{in}}^{(n)}) \quad (1)$$

式中  $D_s$  表示下抽样。设  $M_1$  表示原分辨率运动补偿， $M_2$  表示降低分辨率运动补偿， $x$  表示原分辨率图像信号， $y$  表示降低分辨率图像信号。对于 I 帧，有  $y_{\text{in}}^{(n-1)} = D_s(x_{\text{in}}^{(n-1)})$ 。则漂移误差  $d$  为

$$\begin{aligned} d &= D_s(M_1(x_{\text{in}}^{(n-1)})) - M_2(y_{\text{out}}^{(n-1)}) \\ &= [D_s(M_1(x_{\text{in}}^{(n-1)})) - M_2(y_{\text{in}}^{(n-1)})] + [M_2(y_{\text{in}}^{(n-1)}) - M_2(y_{\text{out}}^{(n-1)})] \\ &= [D_s(M_1(x_{\text{in}}^{(n-1)})) - M_2(D_s(x_{\text{in}}^{(n-1)}))] + [M_2(y_{\text{in}}^{(n-1)}) - y_{\text{out}}^{(n-1)}] \end{aligned} \quad (2)$$

令  $d_1 = M_2(y_{\text{in}}^{(n-1)} - y_{\text{out}}^{(n-1)})$ ，表示用于运动补偿的参考图像中的误差，这是一种在编码转换中常见的漂移误差。令  $d_2 = D_s(M_1(x_{\text{in}}^{(n-1)})) - M_2(D_s(x_{\text{in}}^{(n-1)}))$ ，这是一种由于运动补偿和下抽样的非交换性引起的漂移误差，是降低分辨率编码转换特有的。影响  $d_2$  的原因主要有两个：运动矢量映射和下抽样。在由原分辨率到降低分辨率的运动矢量映射中，由于限制了编码精度，所以需对运动矢量进行舍位。在压缩域中，当下抽样为一个低空间分辨率时，为了避免交叠块滤波，通常采用块约束。这两种过程虽然减少了计算复杂度，但也引入了误差，会在预测和差值之间产生进一步的失配，且随着连续预测图像的增加而增加。

### 3 适应帧内刷新

由于帧内编码模式的压缩率较低，因此当大量的宏块需要帧内编码时，视频编码器的输出比特率将会大幅度地增加。由此可见，帧内编码模式会产生突发性的比特，从而引起传输延迟，造成网络的拥塞。如果图像中的运动区域由于差错而遭到破坏，由此引起的质量下降将会随时间扩散，造成长时间的质量损伤，这种影响直到下一个帧内刷新到来时才能结束。为了减少帧内编码的比特率，并保持对差错的鲁棒性，同时限制差错的时间传播，可采用自适应帧内刷新方法。

通常，漂移误差与帧内编码块的运动活动性有关。因此，如果一组(4个)宏块中的运动矢量绝对差总和(SAD)大于某个阈值  $T_h$ ，自适应帧内刷新方法就判定这组宏块为帧内编码。阈值的初始值  $T_0$  由运动矢量绝对差总和的简单线性关系决定。

$$\text{SAD} = \sum_i \sum_j |\mathbf{V}_i - \mathbf{V}_j| \quad (3)$$

$$T_0 = \alpha \cdot \text{SAD} \quad (4)$$

式中  $\mathbf{V}_i, \mathbf{V}_j$  为运动矢量， $\alpha$  由训练序列给定。在每一帧编码后，根据目标比特率  $R_T$  和实际比特率  $R$  的差值动态地调整阈值。如果差值为正，表明目标比特率高于实际比特率，阈值设置低了；相反，如果差值为负，则表明阈值设置高了；即

$$T_h = \begin{cases} T_h + \Delta T, & R_T - R \geq 0 \\ T_h - \Delta T, & R_T - R < 0 \end{cases} \quad (5)$$

式中  $\Delta T$  为一修正量。由于采用帧内刷新会使帧内编码块的数目增加，由此产生的比特率增加必须在速率控制中加以解决。

### 4 速率控制

一般来说，率失真函数  $R(Q)$  定义在  $Q$  域。为了提高信源模型的准确性，率失真函数  $R(Q)$  的表示式通常非常复杂。

在视频变换编码中，变换系数中零的数量具有较为重要的意义，零的个数越多，表示压缩比越大。设  $\varphi$  为已量化的变换系数中零的个数， $\varphi$  随  $Q$  单调增加，为一一对应的映射关系<sup>[5]</sup>。因此，从数学的角度来说，编码比特率  $R$  也是  $\varphi$  的函数，表示为  $R(\varphi)$ 。对于典型的变换编码， $R(\varphi)$  为一线性函数：

$$R(\varphi) = \lambda(1 - \varphi) \quad (6)$$

式中  $\lambda$  为一常数。可见，变换编码模型中只有参数  $\lambda$  与图像内容直接有关。

设量化器的步长为  $Q$ ，则对于帧内宏块和帧间宏块，量化器的死区阈值分别为  $Q$  和  $1.25Q$ 。设帧内和帧间 DCT 系数的分布分别为  $p_1(x)$  和  $p_2(x)$ ，则对应零的个数  $\varphi$  为

$$\varphi(Q) = \frac{1}{N} \left[ \sum_{|x| \leq Q} p_1(x) + \sum_{|x| \leq 1.25Q} p_2(x) \right] \quad (7)$$

式中  $N$  表示当前帧中 DCT 系数的总数。对于基于 DCT 的视频编码，其 DCT 系数具有 Laplace 分布<sup>[6]</sup>。Laplace 分布的变换系数为

$$p_i(x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad (8)$$

定义失真度量为  $D(x, \hat{x})$ ，有

$$D(x, \hat{x}) = |x - \hat{x}| \quad (9)$$

式中  $x$  为输入信号， $\hat{x}$  为量化器的输出。根据香农信源编码理论，假设  $D$  为允许的失真，则有

$$R(D) = \log_2 \left( \frac{1}{\lambda D} \right) \quad (10)$$

对应的失真为

$$D(Q) = 2 \int_0^{0.5Q} p_1(x) x dx + 2 \sum_{i=1}^{\infty} \int_{(i-0.5)Q}^{(i+0.5)Q} p(x) |x - iQ| dx \quad (11)$$

将式(8)代入式(11)，得

$$D(Q) = \frac{1}{\lambda} \left[ 1 + \frac{e^{-\lambda Q}}{1 - e^{-\lambda Q}} (2 - e^{-0.5\lambda Q} - e^{0.5\lambda Q}) - e^{-0.5\lambda Q} \right] \quad (12)$$

DCT 系数中零的个数为

$$\varphi = \int_{-0.5Q}^{0.5Q} \frac{\lambda}{2} e^{-\lambda|x|} dx = 1 - e^{-0.5\lambda Q} \quad (13)$$

将变量  $Q$  改为  $\varphi$  后，式(12)变为

$$D(\varphi) = \frac{1 - (1 - \varphi)}{\lambda[1 + (1 - \varphi)]} \quad (14)$$

联立式(10)和式(14)，得  $\varphi$  域的率失真函数  $R(\varphi)$ ：

$$R(\varphi) = \log_2 \frac{1 + (1 - \varphi)}{1 - (1 - \varphi)} \quad (15)$$

将式(15)按泰勒展开，得

$$R(\varphi) = 2 \log_2 e(1 - \varphi) + O([1 - \varphi]^3) \quad (16)$$

由实验可知,  $\varphi$  的平均值大于 70%, 即  $(1 - \varphi)$  小于 0.3, 则式(16)中的非线性项  $(1 - \varphi)^3$  小于 0.027。因此, 式(16)中非线性项  $[1 - \varphi]^3$  同线性项相比, 可以忽略不计。因此从理论上讲,  $R(\varphi)$  近似为线性函数。

为了估计出率失真函数, 必须首先精确地估计出  $\lambda$  值。设  $m$  为当前帧中已编码的宏块数,  $R_m$  为编码  $m$  个宏块所需的比特数,  $\varphi_m$  为宏块中零的个数。根据式(6), 有

$$\lambda = \frac{R_m}{1 - \varphi_m} \quad (17)$$

编码当前帧所需的比特为

$$R = R_C - \Delta R \quad (18)$$

$$\Delta R = \begin{cases} W, & W > Z \cdot M \\ (W - Z \cdot M)F, & \text{其它} \end{cases} \quad (19)$$

式中  $R_C$  是信道速率,  $F$  是帧频。  $\Delta R$  是一个修正量,  $W$  为当前缓冲器中的比特数,  $Z$  经多次实验测量可取一常量<sup>[7]</sup>。  $M$  为缓冲器不产生上溢的最大值。通常  $\Delta R$  的值很小, 在  $R_C$  附近。

确定量化参数  $Q$  的算法如下:

(1)初始化 在第 1 个宏块编码前, 设  $m = 0$ ,  $R_m = 0$ ,  $\varphi_m = 0$ 。对帧内和帧间编码宏块的 DCT 系数, 分别计算其分布  $p_1(x)$  和  $p_2(x)$ 。设  $\lambda$  为标准视频序列的平均值, 取  $\lambda = 0.75$ 。

(2)确定量化参数 根据式(6), 量化余下宏块所产生的零的个数为

$$\varphi = 1 - \frac{R - R_m}{\lambda} \quad (20)$$

根据  $\varphi$  和  $Q$  之间一一对应的映射图, 可以确定量化参数  $Q$ 。当前宏块根据量化参数  $Q$  进行量化和编码。

(3)刷新 设  $\varphi_0$  和  $R_0$  分别为当前宏块产生的零的个数和比特数, 并且有:  $\varphi = \varphi + \varphi_0$ ,  $R_m = R_m + R_0$ 。根据式(17)刷新  $\lambda$  值。如果是帧内宏块, 则从分布系数  $p_1(x)$  中减去当前宏块 DCT 系数中的频率分量; 如果是帧间宏块, 则从分布系数  $p_2(x)$  中减去当前宏块 DCT 系数中的频率分量。

(4)循环 对下一个宏块重复步骤(2), 步骤(3), 直到当前帧的所有宏块都被编码。

## 5 仿真结果

本文对 Foreman 测试序列进行计算机仿真实验。测试序列为 CIF 格式, 帧频为 15fps, 以 256kb/s 的速率编码。采用 2:1 下抽样滤波器, 通过降低分辨率编码转换器编码转换为 QCIF 格式, 32kb/s。取初始量化参数  $Q = 15$ , 缓冲器容量为 3200bit。采用本文提出的方法与帧内刷新结构 TMNS 速率控制策略的缓冲器占有率比较和峰值信噪比分别如图 2 和图 3 所示。

由图 2 可见, 采用本文提出的方法得到的编码输出缓冲器占有率(实线)较帧内刷新结构 TMNS 速率控制策略得到结果(虚线)稳定。

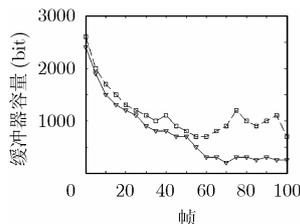


图 2 缓冲器占有率比较

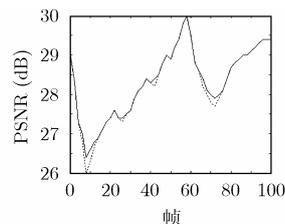


图 3 峰值信噪比比较

由图 3 可知, 采用本文提出的方法得到的信噪比(实线)较采用帧内刷新结构 TMNS 速率控制策略得到的峰值信噪比(虚线)要高, 尤其在场景切换时, 峰值信噪比得到了明显的改善。

## 6 结束语

本文采用联合自适应帧内刷新和速率控制方法, 提出了一种改进的降低分辨率编码转换算法。该方法可以有效地减少帧内编码的比特率, 保持对差错的鲁棒性, 并限制差错的时间传播。同时将率失真函数映射到  $\varphi$  域, 使之具有线性特性, 从而减少了计算复杂度。在对当前帧进行速率控制时, 由于没有利用前一帧的信息和统计特性, 因此在场景切换时不会引起图像质量的下降。

## 参考文献

- [1] Christopoulos B. Transcoder architectures for video coding. *IEEE Trans. on Consumer Electronics*, 1998, 44(1): 88-98.
- [2] Iwasaki O, Uenoyama T, and Ando A, et al. Video transcoding technology for wireless communication systems. In Proc. IEEE Int. Conf. VTC, Boston (U.S.A.), 2000: 1577-1580.
- [3] Shanableh T and Gganbari M, et al. Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats. *IEEE Trans. on Multimedia*, 2000, 2(2): 101-110.
- [4] Yin P, Vetro A, and Liu B, et al. Drift compensation for reduced spatial resolution transcoding. *IEEE Trans. on Circuit and System for Video Technology*, 2002, 12(11): 1009-1019.
- [5] He Z, Kim Y K, and Mitra S K, et al. Low delay rate control for DCT video coding via  $\rho$ -domain source modeling. *IEEE Trans. on Circuit and System for Video Technology*, 2001, 11(8): 928-940.
- [6] He Z and Mitra S K, et al. A line source model and a unified rate control algorithm for DCT video coding. *IEEE Trans. on Circuit and System for Video Technology*, 2002, 12(11): 970-982.
- [7] Li Xiaohui. Video transcoding for wireless mobile Internet access. In Proc. Asia-Pacific Optical and Wireless Communication Conference and Exhibition (APOC), Wuhan (China), SPIE, 2003, 5284: 498-503.

李晓辉: 男, 1961 年生, 教授, 博士, 研究方向为视频编码与多媒体通信。