

一种基于 PID 控制器的 H. 264 比特率转码方案

修晓宇, 卓 力, 沈兰荪

(北京工业大学信号与信息处理研究室, 北京 100022)

摘 要: 本文基于 PID 控制器提出了一种高效的 H. 264 比特率转码方案. 该方案采用简化的像素域级联转码结构, 在保证视频重建质量的同时降低了运算的复杂度, 然后在图像层引入 PID 控制器保证了输出比特率与目标比特率一致. 实验结果表明: 本文提出的比特率转码方案具有较高的比特率转换精度, 是一种高效的比特率转码方案.

关键词: 转换编码; 比特率转码; 比特率控制; H. 264

中图分类号: TP919 **文献标识码:** A **文章编号:** 0372-2112 (2006) 06-1062-04

An H. 264 Bit Rate Transcoding Scheme Based on PID Controller

XIU Xiaoyu, ZHUO Li, SHEN Lansun

(Beijing University of Technology, Signal and Information Processing Lab, Beijing 100022, China)

Abstract: An efficient H. 264 bit rate transcoding scheme based on PID controller, is proposed in this paper. In the scheme, simplified cascaded pixel domain transcoding structure is adopted to reduce the computational complexity while remaining the reconstructed video quality. Furthermore, a PID controller is introduced in the picture layer to ensure the output bit rate consistent with the target bit rate. The experimental results show that, the scheme proposed in the paper can achieve high bit rate accuracy, which is an efficient bit rate transcoding scheme.

Key words: transcoding; bit-rate transcoding; rate control; H. 264

1 引言

随着图像压缩编码技术、宽带数字通信技术和多媒体网络技术的发展, 多媒体数据通信已经在广播电视、计算机和网络通信领域获得了广泛的应用. 不同的用户可以从不同的终端、不同的网络获取各式各样的多媒体数据, 因此实现异质网络间通畅的信息交流已经成为需要重点解决的一个课题. 视频码流的转换编码就是对已经压缩编码的码流进行再处理, 使经过处理后的数据满足传输或用户个性化要求^[1,2].

比特率转码是转换编码的一种实现方式. 它是在不改变帧率和空间分辨率的情况下, 降低视频输出比特率. 对于存储类视频业务, 如视频点播 (VoD, Video on Demand), 视频数据往往是压缩后存储在视频服务器端. 由于视频编码器在编码时无法预知信道状况, 因此需要在传输前采用比特率转码技术来适应网络状况的动态变化^[3,4].

总体来说, 比特率转码有两种实现方法: 像素域方法

和压缩域方法^[3]. 在像素域方法中, 最直接的做法就是将标准解码器与标准编码器级联起来, 即首先对编码码流解码, 然后再按新的目标比特率对解码图像进行重新编码. 文献[5]对像素域的比特率转码方案进行了比较全面的研究, 归纳了四种可能的比特率转码结构. 利用了运动补偿的线性运算特点, 对像素域级联结构进行了简化, 提出了一种简化的像素域级联结构. 在压缩域方法中, 一种做法是将输入的编码码流经过变长解码 (VLD) 和反量化得到 DCT 变换系数后, 直接通过重新量化得到新比特率下的压缩码流, 但这种方法的重建图像中存在严重的漂移误差 (Drift error); 另一种做法是在第一种做法的基础上增加对解码 DCT 系数在频率域进行运动补偿, 减小了漂移误差的影响^[3].

本文基于 H. 264 视频编码标准提出了一种高效的比特率转码方案. 该方案使用了简化像素域级联转码结构, 通过运动矢量复用与结构简化大幅度地降低了算法复杂度; 并且采用了基于 GOP 层、图像层和宏块层的比特率控

收稿日期: 2005-01-28; 修回日期: 2005-11-26

基金项目: 国家自然科学基金 (No. 90304001, 60472036); 北京市自然科学基金 (No. 4052007); 北京市教委项目 (KM 200410005022); 通信抗干扰国家重点实验室基金 (No. 51434050105Q70101); 北京市科技新星计划 (No. 2005B08)

制算法, 通过在图像层引入 PID 控制器, 保证输出比特率与目标比特率一致. 实验结果表明: 该方案具有较高的比特率控制精度, 而且比特率控制算法的运算复杂度低, 不会为转码器带来额外的负担, 可以应用于实时处理.

2 基于 H. 264 的比特率转码结构

图 1 是本文使用的 H. 264 比特率转码方案的结构框图. 该结构是像素域级联转码结构的一种改进 (图 1 中虚线框中的部分), 不但将解码器端的运动矢量直接映射到编码端省去了运动估计这一过程, 降低了运算复杂度, 还对原始结构进行了简化, 将原来的两个闭环缩减为现在的一个闭环. 另外, 这种转码结构还通过对转码器中参考帧之间的误差进行补偿来克服漂移误差对视频重建质量的影响^[6]. 为了实现输出比特率的精确控制, 我们在这个结构中增加了相应的 H. 264 比特率控制算法, 具体内容将在下一节中进行讨论.

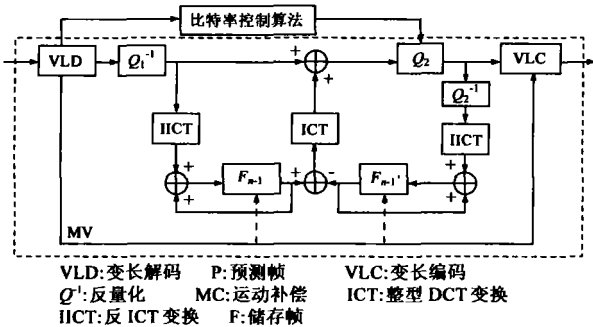


图 1 基于 H.264 的比特率转码结构

3 基于 PID 控制器的比特率控制算法

H. 264 视频编码标准的主要目标是提高压缩效率, 为此采用了多种新的编码技术. 如多种帧内/帧间预测模式、4×4 整型 DCT 变换、基于上下文算术编码 (Context Adaptive Binary Arithmetic Coding, CABAC)、环路滤波等^[6,7]. 与现有的编码标准相比, 其压缩效率提高了 50% 左右. 但在 H. 264 编码标准中并没有提供比特率控制的方法, 因此为了实现对输出比特率的精确控制, 本文将 PID 控制器和 MPEG-2 的 TM5^[8] 比特率控制算法相结合, 提出了一种适用于 H. 264 的比特率控制算法. 该算法从图像组层 (GOP layer)、图像层 (Picture layer) 和宏块层 (Macroblock layer) 三个层次来实现比特率控制.

3.1 图像组层比特率控制

图像组层比特率控制的目的是为一个 GOP 分配所需的比特率. 根据目标比特率和输出帧率, 一个 GOP 所分配的比特率可以表示为^[8]:

$$B = (1 + N_p + N_b) \cdot \frac{B_T}{F} \quad (1)$$

其中 B_T 为目标比特率, F 为输出帧率, N_p 为该 GOP 中 P 帧的数目, N_b 为该 GOP 中 B 帧的数目.

3.2 图像层比特率控制

图像层比特率控制的任務是根据图像组层比特率分配结果 B 为该 GOP 中的每一个待编码图像分配比特率. 为了提高输出比特率的精确性, 我们在图像层引入了一个 PID 控制器, 其结构如图 2 所示. 其中 FBA 环节根据该 GOP 中剩余的比特率为待编码图像分配比特率基值 B_x ($x = I, P, B$), 反馈环节将实际输出比特率 B_R 与目标比特率 B_T 的差值送入 PID 控制器并且输出补偿值 ΔB , 并以此对图像比特率基值 B_x 进行修正.

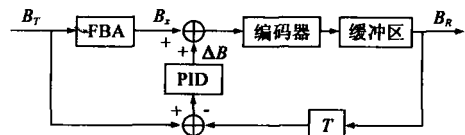


图 2 图像层比特率控制结构

由于比特率分配方法精度不足的原因, 往往会出现实际输出比特率与目标比特率不完全一致的情况, 此时比特率控制算法可以通过 PID 控制器来对 FBA 环节产生的分配误差进行及时的补偿.

3.2.1 FBA 环节

FBA 环节根据该 GOP 中的剩余比特率, 为待编码的图像分配比特率基值. 对于采用不同模式进行编码的帧采用了不同的分配方法计算每帧分配的比特数^[8], 对于 I 帧:

$$B_i = \max \left[\frac{B'}{1 + \frac{N_p X_{ip}}{C_p X_i} + \frac{N_b X_{ib}}{C_b X_i}}, \frac{B_T}{8 \times F} \right] \quad (2)$$

对于 P 帧:

$$B_p = \max \left[\frac{B'}{N_p + \frac{N_b C_p X_b}{C_b X_p}}, \frac{B_T}{8 \times F} \right] \quad (3)$$

对于 B 帧:

$$B_b = \max \left[\frac{B'}{N_b + \frac{N_p C_b X_p}{C_p X_b}}, \frac{B_T}{8 \times F} \right] \quad (4)$$

上式中 B' 为该 GOP 中剩余的比特率; C_i , C_p 和 C_b 是三个比例常数, 通常设置 $C_i = 1$, $C_p = 1$, $C_b = 1.4$; X_i , X_p 和 X_b 分别为 I、P 和 B 帧的复杂度, 其值可以由转码器解码部分解得的该帧图像的平均量化参数 Q_i ($t = i, p, b$) 和所生成比特数 S_i ($t = i, p, b$) 的乘积得到, 即:

$$\begin{aligned} X_i &= S_i \times Q_i \\ t &= i, p, b \end{aligned} \quad (5)$$

3.2.2 PID 控制器

如图 2 所示, 反馈环节将实际的输出比特率输入 T 环节. T 环节将一秒内的输出比特率取算术平均后输出, 这样只有在输出比特率的统计特性发生变化时, 才会输出 ΔB .

校正环节由一个 PID 控制器构成. PID 控制器结构简单, 易于实现, 其基本结构如下式所示^[1,9]:

$$\Delta B = K_p \left[B_E + \frac{1}{K_i} \int B_E dt + K_d \frac{dB_E}{dt} \right] \quad (6)$$

上式中 $B_E = B_T - B_R$ 为目标比特率与实际输出比特率的差值, K_p 、 K_i 和 K_d 分别为 PID 控制器的比例系数、积分系数和微分系数. 式(6)中增大比例系数 K_p 可以提高控制的精度, 但数值过大也会造成系统的不稳定; 增大积分系数 K_i 可以改善系统的稳定特性; 增大微分系数 K_d 可以增强系统的瞬时响应能力, 有助于增强系统的稳定特性.

3.3 宏块层比特率控制

宏块层比特率控制的核心就是为编码图像中的每一个宏块确定量化参数 QP. 首先, 我们计算编码第 j 个宏块时缓冲区的溢出程度^[9], 如下式所示:

$$d_j^i = d_0^i + T_{j-1} - \left(\frac{B_i \times (j-1)}{MB_cnt} \right)$$

或者

$$d_j^p = d_j^p + T_{j-1} - \left(\frac{B_i \times (j-1)}{MB_cnt} \right) \quad (7)$$

或者

$$d_j^b = d_0^b + T_{j-1} - \left(\frac{B_i \times (j-1)}{MB_cnt} \right)$$

上式中 d_j^i 、 d_j^p 和 d_j^b 分别是 I、P 和 B 帧编码第 j 个宏块时缓冲区的溢出程度状态值; d_0^i 、 d_0^p 和 d_0^b 分别是编码 I、P 和 B 帧缓冲区的初始值, 其数值等于编码前一帧最后一个宏块时缓冲区的状态值; T_j 是该帧图像编码到第 j 个宏块时实际输出比特数; MB_cnt 是一帧图像中所包含宏块的个数. 然后, 再计算该宏块的量化参数的基值 QP_j^b , 具体为:

$$QP_j^b = \frac{d_i \times 52}{r}, \quad r = 2 \times \frac{B_r}{F} \quad (8)$$

上式中得到的宏块量化基值最后再被宏块活动性调制得到最终的量化参数 QP_j , 即:

$$QP_j = MB_{act,j} \times QP_j^b \quad (9)$$

宏块活动性 $MB_{act,j}$ 定义为:

$$MB_{act,j} = \frac{2 \times \delta_j + \delta_{avg}}{\delta_j + 2 \times \delta_{avg}} \quad (10)$$

其中 δ_j 为第 j 个宏块的方差; δ_{avg} 为一帧中所有宏块活动性的平均值. 这两个值都可以从转码器解码部分产生的解码图像数据中得到.

4 实验结果

我们基于 H. 264 视频编码标准, 对本文提出的比特率转码方案进行了测试. 原始压缩数据是用标准 H. 264 编码器在 QP=5 情况下编码得到, 采用 IPPP... 的编码结构, 编码帧率为 15fps. 表 1 为采用本文提出的比特率转码方案得到的输出比特率. 从表中可以看出, 平均输出的比特率

可以很好地控制在目标比特率附近, 误差可以控制在 0.1% 以内.

图 3 是目标比特率为 12800bit/s/frame, 输出帧率为 15fps 情况下, Container (QCIF 格式) 序列每帧比特率控制的结果. 其中实线是实际比特输出结果, 虚线是目标比特率. 从图中可以发现, 由于图像层的 PID 控制器使用的是一秒以内输出比特的算术平均值, 在第一个一秒内 PID 控制器并没有工作, 所以在前 15 帧图像中会出现一个比较大输出比特率的波动. 但是随着帧数的增加 PID 控制器开始工作, 输出比特率会稳定在目标比特率的附近.

表 1 不同序列比特率转码后的输出比特率

视频序列	目标比特率 (bits/frame)	输出比特率 (bits/frame)	误差 (bits/frame)	误差精度 (%)
Container (QCIF, 150 帧)	12800	12794.51	5.49	0.04
	9600	9598.56	1.44	0.02
	6400	6401.07	1.07	0.02
Foreman (QCIF, 150 帧)	12800	12786.99	13.01	0.1
	9600	9590.67	9.33	0.09
	6400	6391.04	8.96	0.1
M & D (QCIF, 150 帧)	12800	12809.44	9.44	0.07
	9600	9598.41	1.59	0.02
	6400	6403.85	3.85	0.06

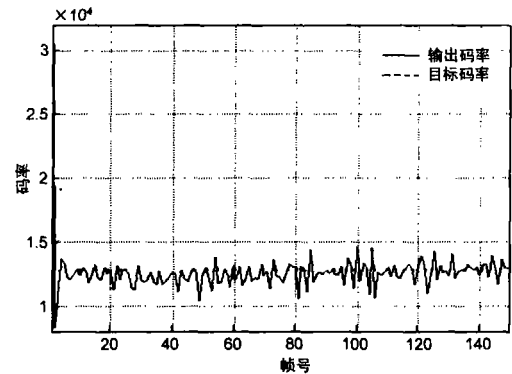


图 3 Container(QCIF 格式) 比特率输出结果

表 2 所示的是在不同目标比特率下, 图像层采用与不采用 PID 控制器两种情况下比特率输出的结果. 通过比较可以发现: 图像层引入 PID 控制器会很大程度地减少输出比特率与目标比特率的偏差, 从而增加了比特率控制算法的稳定性, 提高了输出比特率的精度.

表 2 图像层引入 PID 控制器的结果

视频序列	目标比特率 (bits/frame)	输出比特率 (bits/frame)	
		有 PID 控制器	没有 PID 控制器
QCIF	Akiyo	3199.3	3153.7
	Trevor	3193.8	3090.3
	Silent	3197.5	3117.5
CIF	Mobile	51191.6	50923.5
	Susie	51210.1	52983.9
	Hall	51195.1	51866.2

到目前为止, 学者们还没有就转码效率的评价准则这

一问题达成共识,因此为了说明本文提出比特率转码方案的高效性,我们在相同输出比特率条件下,从视频重建质量和运算复杂度两个方面,将本文提出的比特率转码方案与直接编码方法进行了比较,如表 3 所示。通过比较可以发现,与直接编码相比,本文提出的比特率转码方案虽然在视频恢复质量上会有 1dB 左右的下降,但是却大幅度地节约了运算的复杂度,可以满足实时应用的需求。综合起来看,这种复杂度上的节约相比较质量上的降低而言是值得的,因而本文提出的转码方案具有很高的转码效率。

表 3 本文提出的比特率转码方案与直接编码的性能比较

序列	比特率 (bit/s/Frame)	PSNR ₁ (dB)	PSNR ₂ (dB)	Δ PSNR (dB)	T ₁ (ms)	T ₂ (ms)	(T ₁ -T ₂)/T ₁ (%)	
Foreman	QCIF	6400	35.61	34.32	1.29	208	46	77.9
	CIF	23400	36.78	35.08	1.7	808	187	76.9
Akiyo	QCIF	1400	37.23	36.25	0.98	187	29	84.5
	CIF	5200	38.97	37.79	1.18	728	118	83.8
M & D	QCIF	2700	36.41	35.44	0.97	200	42	79
	CIF	8300	37.15	35.94	1.21	796	161	79.8

PSNR₁: 直接编码后平均每帧视频恢复质量; T₁: 直接编码平均每帧运算复杂度

PSNR₂: 本文提出转码结构的平均每帧视频恢复质量; T₂: 本文提出转码结构的平均每帧运算复杂度

最后需要强调的是:本文提出比特率转码算法中所需要的图像复杂度和宏块活性等参数完全可以从转码过程产生的解码数据中得到,不会对转码器带来额外的负担,因此这种比特率转码结构完全可以应用于实时处理。

5 结论

本文基于 H. 264 视频编码标准,将简化像素域级联转码结构与基于 PID 控制器的比特率控制算法相结合,提出了一种高效的比特率转码方案。该方案的特点可以归纳为:(1) 采纳了简化像素域级联转码结构,与直接将编解码器相级联的转码结构相比,通过运动矢量复用与结构简化大幅度地降低了算法复杂度,可以用于实时处理。(2) 采用了基于 GOP 层、图像层和宏块层三个层次的比特率控制算法,并通过在图像层引入 PID 控制器的办法,保证输出比特率与目标比特率一致。(3) 比特率控制算法中需要的参数都可以从转码中间过程得到,不会对转码器带来额外的负担。实验数据表明,本文提出的比特率转码方案具有较高的比特率转换精度,并且与直接编码相比,该方案在保证相近视频恢复质量前提下,大幅度降低了复杂度,可以应用于实时处理,是一种高效的比特率转码方案。

参考文献:

[1] 沈兰荪,等,著. 视频编码与低速率传输[M]. 北京: 电子工业出版社, 2001. 1- 319.

Shen L S, et al. Video Coding and Low Bit Rate Transmission[M]. Beijing: Publishing House of Electronics Industry, 2001. 1- 319. (in Chinese)

[2] 卓力,沈兰荪,朱青. 视频流关键技术的研究进展[J]. 电子学报, 2002, 30(8): 1213- 1218.

Zhuo L, Shen L S, Zhu Q. Reach advances in key technology of video streaming[J]. Acta Electronica Sinica, 2002, 30(8): 1213- 1218. (in Chinese)

[3] Zhijun Lei. Video Transcoding techniques for wireless video communications[D]. Ottawa: University of Ottawa, 2003.

[4] 褚晶辉,俞斯乐,鲁照华. 视频转换编码及其实现技术的研究[J]. 电子学报, 2004, 10(10): 1678- 1683.

Chu J H, Yu S L, Lu Z H. Research on transcoding technology[J]. Acta Electronica Sinica, 2004, 10(10): 1678- 1683. (in Chinese)

[5] Huifang Sun, Wilson Kwok, R Gary. Method for matching compressed video to ATM networks[J]. IEEE Trans on Circuit and Systems for Video Technology, 1996, 6(2): 191- 199.

[6] Wiegand T, Sullivan G J, Bjntegaard G. An Overview of the H. 264/AVC video coding standard[J]. IEEE Trans on Circuits and Systems for Video Technology, 2003, 13(7): 560- 576.

[7] Joint video team (JVT) of ITU-T VCEG and ISO/IEC MPEG. Final draft international standard of joint video specification (ITU-T Rec H. 264/ISO/IEC 14496-10). JVT-G050, 05. 2003.

[8] ISO/IEC JTC1/SC29/WG11, MPEG Video Test Model 5, Draft, 04, 1993.

[9] 田栋,沈兰荪,段占云. 一种用于视频通信的比特率控制方法[C]. 中国合肥: 第三届全球智能控制与智能自动化大会, 2000. 2512- 2516.

Tian D, Shen L S, Duan Z Y. An bit rate control method used in video communication[C]. He fei: The 3rd World Congress on Intelligent Control and Automation, China, 2000. 2512- 2516. (in Chinese)

作者简介:



修晓宇 男, 1980 年出生于黑龙江省哈尔滨市, 北京工业大学硕士生, 主要研究方向为视频转换编码、传输技术。

E Mail: xiuxiaoyu@emails.bjut.edu.cn