

一种 H. 264 加权预测 MAD 的码率控制算法

胡 旭¹, 武淑红¹, 张 刚²

(1 太原理工大学 计算机科学与技术学院, 山西 太原 030024; 2 太原理工大学 信息学院, 山西 太原 030024)

摘 要: 本文提出一种 H. 264 宏块层加权预测 MAD 值的码率控制算法来提高视频的质量并减少计算的复杂度。在详细分析了时间-空间相关性和物体的运动特征后, 当前宏块的平均绝对差值(MAD)通过前一帧时间相关性模型预测的 MAD 值和由当前帧的空间相关性模型预测的 MAD 值加权得到。所得到的 MAD 值更加准确, 能适应不同类型帧序列。实验结果表明, 与 JVT-G012 相比, 提出的方法在相同条件下码率有所下降, 更加接近目标码率。峰值信噪比(PSNR)平均可以提高 0.22dB。

关键词: 码率控制; 宏块层; 平均绝对差值; 峰值信噪比

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2013)04-0152-04

H. 264 Weighted Prediction Rate Control Algorithm of the MAD

HU Xu¹, WU Shu-hong¹, ZHANG Gang²

(1 College of Computer and Science, Taiyuan University of Technology, Taiyuan 030024, China;

2 College of Information, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: In this paper, the macroblock layer weighted prediction rate control algorithm is proposed for H. 264, which can greatly improve the video quality and reduce the computational complexity. After a thoroughly analysis of the temporal-spatial relationship and the movement of objects, the mean absolute difference (MAD) value of current MB is weighted combination of temporal MAD predicted by previous frame and spatial MAD predicted by current frame. MAD values obtained by this method are more accurate and can be able to adapt to different types of frame sequences. The experimental results show that, compared with JVT-G012, this proposed algorithm can control the rate more accurately and the PSNR can rise up to 0.22dB in average.

Key words: rate control; macroblock layer; Mean Absolute Difference (MAD); PSNR

1 引言

在实际视频传输中, 由于网络带宽、存储空间以及延时等限制, 必须将编码器输出码率的波动限制在一定范围内, 这时需要用合理且有效的码率控制算法, 以获取在质量和带宽利用上的最佳平衡^[1]。

H. 264^[2]作为一种新的视频压缩标准, 和以前的视频标准相比, 可以表现出更好的性能。而这主要是基于 H. 264 编码器采用了更为复杂的方法, 率失真优化(Rate-Distortion Optimization, RDO)^[3]就是一项被采用的技术, 率失真优化可以用于运动估

计和模式选择等方面, 因此 H. 264 比其他的标准有更好的编码效率。但这些特性又使 H. 264 比其他的标准在码率控制上显得更加复杂。其中最主要的问题就是量化参数 QP 的确定和依赖 QP 值的率失真优化之间的数据依赖。为了使当前帧中的率失真优化可以执行, 应该首先确定每一个宏块的量化参数, 而这是通过使用当前帧或宏块的平均绝对差值(Mean Absolute Difference, MAD)计算得到的。但是, 当前帧宏块的 MAD 值只有在率失真优化后才可用。这就是如图 1 所示的“蛋鸡悖论”^[4]。

收稿日期: 2012-07-11; 修回日期: 2012-08-03

基金项目: 国家自然科学基金项目(60772101)

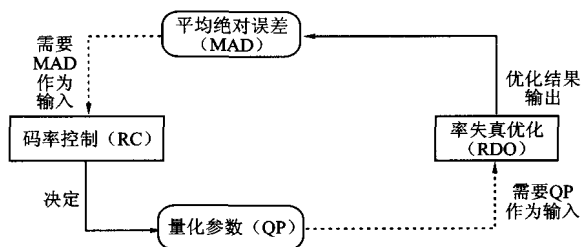


图1 “蛋鸡悖论图”

2 现有码率控制算法

在数字编码标准和应用的方面,许多专家学者对现有的码率控制进行了广泛的研究,并提出许多经典的码率控制算法。主要有MPEG-2的TM5算法^[5],MPEG-4的VM8算法^[6],H.263的TMN8算法^[7]和最新编码标准H.264的JVT-G012^[8]算法等。

现有的码率控制算法主要是基于H.264官方的预测模型,而JVT-G012是较为常用的码率控制算法。在JVT-G012算法中MAD值的预测存在两个问题。当前帧宏块的MAD值是由前一帧相同位置的宏块中预测所得到的。对于一些运动剧烈的序列或发生场景切换的时候,当前的预测模型不能适应MAD值的突然改变,结果码率控制效果不太理想。

因此,很多学者在JVT-G012算法上进行改进。作者^[9]提出一种新的方法,它使用了三种类型的宏块,包括定位宏块、左宏块、和上宏块来预测当前宏块的MAD值。

3 本文改进算法——加权的时间空间预测模型

在详细分析时间—空间的关系和物体运动方向,提出一种加权预测MAD值的码率控制算法来提高MAD值的预测精度并降低计算复杂度。当前宏块的MAD值通过时间相关性模型预测的MAD值(通过前一帧预测)和空间相关性模型预测的MAD值(通过当前帧预测)加权得到,再通过得到的MAD值计算出QP值。

3.1 时间相关性预测模型

当前帧和前一帧存在着一定的联系,这种联系叫做时间相关性。时间相关性模型能够用来预测当前宏块的MAD值。图2中, P_0 为当前宏块在前一帧中相同位置的宏块, P_i 为 P_0 相邻的九个宏块。由于宏块之间的关系与距离的减小成反比。因此,当前宏

块通常和定位宏块 P_0 间有很大的联系,和前一帧其他的宏块 P_i 的联系相对比较小。

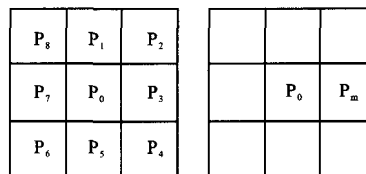


图2 参考帧和当前帧宏块位置关系

基于以上的分析,在本文提出的时间相关性模型中,我们选取两个宏块包括定位宏块 P_0 和最相关宏块 P_m ,采用两个宏块MAD的平均值来获得当前宏块的时间相关性MAD值,具体公式如下:

$$MAD_i = (MAD[i-1, P_0] + MAD[i-1, P_m]) / 2 \quad (1)$$

其中 P_m 是最相关宏块,且属于 P_i ($1 \leq i \leq 8$),图2表示其中的一种情况, P_m 等同于 P_3 ,它表示最相关宏块即为 P_3 。为了选取最相关宏块,本文引入最相关宏块系数 $Q[P_i]$ ($1 \leq i \leq 8$), $Q[P_i]$ 代表前一帧中宏块 P_0 和第宏块 P_i 之间的关系。 $Q[P_i]$ 值越小代表宏块间的关系越大,选择具有最小 $Q[P_i]$ 值的宏块作为最相关的宏块,即 P_m 。具体公式如下:

$$Q[P_m] = \min(MAD[i-1, P_0] - MAD[i-1, P_i]) \quad (1 \leq i \leq 8) \quad (2)$$

3.2 空间相关性预测模型

当前帧宏块之间也存在着空间相关性。尤其是当视频序列剧烈运动或发生场景切换时,时间相关性较弱,空间相关性更加明显。在图3中,由于编码的顺序性,已编码的宏块(阴影部分)位于当前宏块(C_0)的左方或上方,因此可以用已编码的宏块来计算 MAD_{C_0} 。在空间相关性预测模型中,对相关性的影响有两个方面:其一为空间相关性随着距离的增加而减弱,另一个为纹理方向的特性。

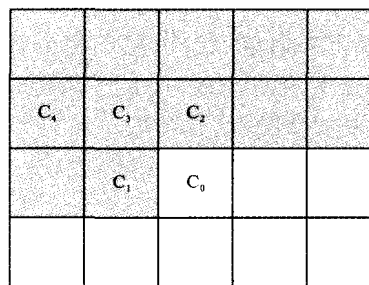


图3 已编码宏块及用于计算模式选择的宏块

基于编码顺序的特性,提出如图4中所示的四种方向预测模式,包括垂直模式、左上模式、水平模式、和右上模式,选取四种模式中的最佳模式来计算

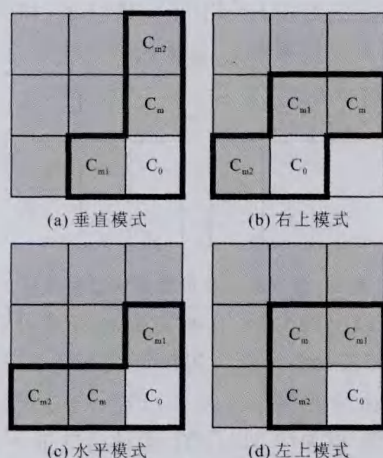


图4 空间四种模式的预测方式

MAD_s . 以上四种模式选择三个相邻的宏块来计算空间相关性的 MAD 值, 其中 C_m 、 C_{m1} 、 C_{m2} 分别代表当前帧中 C_0 的三个相关宏块. 选择三个宏块的原则为和 C_0 临近且加重该方向的权重, 例如垂直方向选择了相邻的 C_m 、 C_{m2} 两个宏块, 并加重垂直方向的权重, 因此选取了 C_{m2} 宏块. 其它模式宏块选择也如此. 当表现为垂直方向的时候, 垂直模式将会被选择作为最佳的预测模式, 其它情况也如此. 为了确定哪一种模式为最佳的预测模式, 通过图 3 中的四个宏块, 包括 $C_1 - C_4$ 计算四种模式的 MAD 距离包括 MAD_v 、 MAD_h 、 MAD_l 、 MAD_r 来选择最佳的预测模式, 具体计算方法如下:

$$MAD_v = |MAD[i, C_1] - MAD[i, C_3]| \quad (3)$$

$$MAD_h = |MAD[i, C_2] - MAD[i, C_3]| \quad (4)$$

$$MAD_r = |MAD[i, C_1] - MAD[i, C_2]| \quad (5)$$

$$MAD_l = |MAD[i, C_1] - MAD[i, C_4]| \quad (6)$$

其中, MAD_v 、 MAD_h 、 MAD_l 、 MAD_r 分别代表垂直模式、水平模式、左上模式、右上模式的 MAD 距离. 选择具有 MAD_v 、 MAD_h 、 MAD_r 、 MAD_l 中最小值的模式作为最佳的预测模式, 然后通过该模式下的三个相关宏块的 MAD 值的平均值来计算当前宏块的空间相关 MAD_s , 具体如下:

$$MAD_s = (MAD_m + MAD_{m1} + MAD_{m2}) / 3 \quad (7)$$

3.3 量化因子的决定

通过时间相关性预测模型和空间相关性预测模型, 可以得到 MAD_t 和 MAD_s , 然后需要计算量化因子 α . 量化因子 α 代表时间-空间相关性, 当序列有巨大运动或较弱的纹理细节的时候, 它的值会减小, 在本文中, 我们引入 Q_t 和 Q_s 来计算 α , 其中 Q_t 表示时间相关性系数, Q_s 表示空间相关性系数, 它

们通过下式来计算:

$$Q_t[j] = |MAD[i-1, C_m] - MAD[i, C_m]| \quad (8)$$

$$Q_s[j] = |MAD[i-1, j] - MAD[i-1, C_m]| \quad (9)$$

式中, C_m 代表图 4 中最佳预测模式中的最相关宏块. 然后使用分段函数来计算 α 的值, 具体公式见公式如下

$$\alpha[j] = \begin{cases} 0.8 & Q_s[j] > 2Q_t[j] \\ 0.3 & Q_s[j] < 0.5Q_t[j] \\ 0.5 & \text{其它} \end{cases} \quad (10)$$

3.4 加权时间-空间预测模式

当前宏块 MAD 值是通过加时间相关性预测模型 MAD 值和空间相关性预测模型 MAD 值加权得到的, 具体公式如下:

$$MAD_p = \alpha \times MAD_t + (1 - \alpha) \times MAD_s \quad (11)$$

4 实验结果和讨论

在 H.264 编码器参考软件 JM11.0 上对本文的改进算法的性能进行了测试. 实验所得到的数据和原始的码率控制算法 JVT-G012 进行比较. 采用宏块作为基本单元, 选取的测试序列纹理和运动复杂度各不相同. 在本试验中, 选取五组运动类型的 QCIF(176 * 144) 格式的序列组成, 包括 akiyo, news, salesman, carphone, foreman, 以上视频序列的图像纹理的细节和运动程度总体上依次由小到大排列. Akiyo 的图像纹理和运动程较为缓慢, 而 carphone 和 foreman 存在较为快速的剧烈运动, 而且 foreman 序列还涉及到场景切换.

实验的测试条件如下: 测试序列选取 IPPP 的编码方式, 搜索范围为 16, 设置帧率为 15f/s, 参考帧数为 2, 熵编码形式采用 CABAC, 编码帧数为 150 帧, 目标码率被设定为 48 Kbps, 64 Kbps, 128 Kbps. 每一种模式中使用相同的参数, 其余的参数为 JM11.0 的默认参数.

从表 1 的实验数据中可以看出本文所提出的码率控制的改进方法和 JVT-G012 算法相比, 测试所得到的码率在很大程度上更加接近目标码率, 所得到的峰值性噪比更加高, 平均提高 0.22dB. 这说明本文的加权预测模型不仅提高了 MAD 值的预测精度而且对视频编码的质量上也有很大的提高.

5 结束语

本文在 JM11.0 测试模型采用的码率控制算法的基础上提出了一种改进的加权预测 MAD 值的码率控制算法. 该算法采用新的时间、空间 MAD 值预

测来提高 MAD 加权预测修正模型来提高 MAD 值 帧类型,在使码率控制更加精准的同时,平均的峰值 预测的精确性. 实验结果表明,该算法能适应不同的 性噪比也获得了较大的提高.

表 1 码率及 PSNR 测试比较

测试序列	目标码率 /Kbps	JVT-G012		本算法		Δ 码率 /Kbps	Δ PSNR /dB
		码率/Kbps	PSNR/dB	码率/Kbps	PSNR/dB		
akiyo	48	48.11	31.73	48.02	31.80	-0.09	+0.07
	64	64.08	33.67	64.03	33.79	-0.05	+0.12
	128	128.21	35.01	128.12	35.36	-0.09	+0.35
news	48	48.17	32.05	48.09	32.41	-0.08	+0.36
	64	64.02	33.41	64.06	33.59	+0.04	+0.18
	128	128.08	35.73	128.06	35.89	-0.02	+0.16
sales- man	48	48.22	33.36	48.15	33.70	-0.07	+0.34
	64	64.09	35.27	64.10	35.36	+0.01	+0.09
	128	128.04	35.44	128.01	35.67	-0.03	+0.23
carphone	48	48.11	31.45	48.03	31.70	-0.08	+0.25
	64	64.18	32.68	64.15	32.85	-0.03	+0.17
	128	128.07	34.37	128.05	34.63	-0.02	+0.26
fore- man	48	48.03	32.73	48.01	32.98	-0.02	+0.25
	64	64.14	34.56	64.17	34.79	+0.03	+0.23
	128	128.09	38.61	128.03	38.80	-0.06	+0.19

参考文献:

[1] 李慧然. H. 26 视频编码标准码率控制算法的优化 [D]. 成都:西南交通大学,2008.

[2] Draft ITU-T recommendation and final draft international standard of joint video specification[S]. 2003.

[3] 毕厚杰. 新一代视频压缩编码标准——H. 264/AVC [M]. 北京:人民邮电出版社,2005.

[4] VETRO A. Transcoding, scalable coding and standardized metadata[C]//Proceedings of the 8th International Workshop on Visual Content Processing and Representation. Berlin; Springer-Verlag, 2003: 15 — 16.

[5] ISO/IEC JTC1/SC29 W1796. MPEG-4 Video-Verification Model, Version 8. 0[S]. Genevese: ISO,1999.

[6] ITU-T TMN 8. Video Codec Test Model, Near-Term Version 8[S]. Genevese: ISO, 1997.

[7] HE Zhihai, KIM Yong Kwan, Sanjit K. Low-delay rate control for DCT video coding via p-domin source modeling[J]. IEEE Trans on CSTV,2001,11(8):813-816.

[8] Li Z, Pan F, Lim K P, et al. Adaptive basic unit layer rate control for JVT[C]//Joint video Team(JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-G012, 7th Meeting ; Pattaya II. Thailand: IEEE, 2003.

[9] 龚道香. H. 264 码率控制中 mad 预测的一种改进算法 [J]. 杭州电子科技大学学报, 2006, 26(5): 125-128.

作者简介:



胡 旭 男, (1984—), 硕士. 研究方向为视频编码.

武淑红 女, (1969—), 硕士生导师. 研究方向为语音编码、视频编码.

张 刚 男, (1953—), 博士生导师, 研究方向为图像信号处理、嵌入式及多媒体技术研究.