

An Rate Control Algorithm for Low Delay Video Communication of H.264/AVC

HE Shuqian¹, WANG Jianxin¹, ZENG Qingsong²

1. School of Information Science and Technology Dept. name of organization, Hainan Normal University, Haikou, China

2. School of Information & Technology, Guangzhou PanYu Polytechnic, Guangzhou, China

e-mail: heshuqian05@126.com, pypzengqs@gmail.com

Abstract: In order to reduce the computational complexity and improve the efficiency of H.264/AVC encoder for low delay communication, A new simpler and efficient R-1/Q model and D-Q model was proposed based on the spatio-temporal correlation and the frame content of complexity with an accurate mean-absolute-difference (MAD) prediction model. Based on the new Rate-Distortion model, the proposed algorithm is able to adaptively optimize macroblock layer rate control so that the overall coding efficiency is improved. Experimental results show that the proposed algorithm increased the code efficiency, got more stable image quality and achieves a bit rate closer to the target rate.

Keywords: rate-distortion; rate control; H.264/AVC; quantification; video encoding template

一种面向 H.264 的低延时视频通信码率控制方法

何书前¹, 王建新¹, 曾青松²

1.海南师范大学信息科学技术学院, 海口, 中国, 571158

2. 广州番禺职业技术学院信息工程学院, 广州, 中国, 511483

e-mail: heshuqian05@126.com, pypzengqs@gmail.com

【摘要】为了降低 H.264/AVC 编码器低延时通信中的计算复杂度和提高编码效率, 结合视频序列的时空相关性, 利用精确平均绝对误差 MAD 预测模型作为帧的内容复杂度, 提出了一种简单高效的 R-1/Q 和 D-Q 模型, 基于该率失真模型, 提出了一种新的自适应的码率控制算法。实验结果表明, 该算法获得更高的编码效率, 图像质量更加平稳, 更准确的控制输出码率。

【关键词】率失真; 码率控制; H.264/AVC; 量化; 编码

1 引言

随着多媒体服务的迅速发展和网络技术不断的更新换代, 人们对实时视频通信的要求越来越高, 因此, 研究有效的低延时视频压缩技术已经成为了视频编码领域的重要研究热点问题, 在实现视频编码码率与信道传输带宽相匹配的同时, 防止缓冲器上溢下溢的条件下, 使得接收视频质量最优化。H.264/AVC 是国际电信联盟电信标准化组织 (ITU-T)、国际标准化组织 (ISO) 和国际电工委员会 (IEC) 联合研究并发布的最新的国际视频压缩标准, 与以前的视频压缩标准相比, 其高效性和网络友好性得到了非常明显的改善, 适合于低延时的实时视频应用。在低延时视频通信中, 为了解决常带宽信道与视频压缩输出的变比特信源之间的匹配问题, 同时, 为了增加视频编码的效率, 码率控制起到了非常

重要的作用。

对于低延时视频通信的应用, 如视频会议, 视频监控应用, 通常需要非常小的缓冲器容量。然而, 在实际应用中, 小的缓冲器容量容易导致上溢和下溢。在这种情况下, 码率控制中的率失真优化模型精确性要求非常高, 同时, 编码器参数控制要及时有效。因此, 低复杂的率失真模型和准确的量化参数选择是码率控制研究需要解决的重要问题。相对于以前的视频压缩标准 MPEG-1, MPEG-2, H.261 和 H.263, H.264 采用了率失真最优化 (RDO) 决定运动信息, 利用经典的率失真最优化理论, 采用拉格朗日乘子法在运动信息和残差参数之间找到最佳比特分配解; 另外, H.264 也采用拉格朗日方法得到最优的运动估计和模式判决^{[1][2]}, 也就是说, 在 H.264 中, 量化参数 Q 不仅与残差参数比特有关, 而

且与运动信息比特也相关，所以以往传统的码率控制方法不再适用。

为了方便码率控制，现有码率控制方法不注重模型的精确性，大部分的率失真模型主要研究码率和量化步长的关系，并在视频压缩标准中使用，如 MPEG-4 中采用的二次模型^[3]。为了有效的解决精确性和码率控制问题，^[4]提出了一种高效的高码率环境下的 R-D 模型，可适用于各种视频编码标准。针对 H.264 的特殊情况，^{[5][6]}提出了基于该编码环境下的率失真模型，^{[7][8]}进一步改进了率失真模型在 H.264 下的应用。在这些模型中，T.Wiegand 等^[3]提出的率失真模型最有效，该模型根据熵限制条件下的矢量量化来拟合高码率下的码率曲线，其率模型为： $R(D) = a \log_2(\frac{b}{D})$ ，其中 a 和 b 为常数。

假设源信号均匀的分布在量化区间内，则失真模型为： $D = \frac{(2Q)^2}{12} = \frac{Q^2}{3}$ ，其中 Q 为量化参数。通过拉格朗日乘法把该约束问题转化为非约束问题，得 $J = D + \lambda R$ ，代入求得到最优的高码率环境下的拉格朗日参数： $\lambda = -\frac{dR}{dD} = c \cdot Q^2$ 。

在率失真模型中，针对当前编码量化参数与所需 MAD 值之间存在因果矛盾问题，需要根据以前的已编码的 MAD 值来预测当前 MAD 值来解决该先后矛盾问题。因此，准确的 MAD 值预测决定了率失真模型的准确性。Liu^[9]在原 H.264 标准的线性 MAD 模型基础上，发展了一种结合时空相关性的预测模型，能够准确的预测当前编码 MAD 值，避免了视频内容波动情况下不准确的预测问题。其形式如下：时域 MAD 预测：

$$MAD_{pred,adapt}[i+1] = MAD_{pred,temp}[i+1] \quad (1)$$

和空域 MAD 预测：

$$MAD_{pred,adapt}[i+1] = MAD_{pred,spat}[i+1] \quad (2)$$

其中， $MAD_{pred,spat}[i] = Z_1[i] \times MAD_{rough}[i] + Z_1[i]$ ， $MAD_{rough}[i]$ 为当前原始帧与前一重构帧之间的差值； $MAD_{pred,temp}[i] = Y_1[i] \times MAD_{actual}[i-1] + Y_2[i]$ ， $MAD_{actual}[i-1]$ 为前一帧相同位置宏块的实际 MAD 值。最后利用 SAD 作为判决准则，判定预测 MAD 值与实际的 MAD 值之间的时空相关性，从而确定最终的预测 MAD。

2 基于 H.264 率失真模型

在 H.264 码率控制算法中采用了二次失真模型 D-Q

$$\text{和线性率模型 R-1/Q: } \begin{cases} R[i] = S \frac{X_1[i]}{Q[i]} + X_2[i] \\ D[i] = \frac{Q^2[i]}{\varepsilon} \end{cases} \quad (3),$$

其中，式中 R[i] 为分配给第 i 帧的比特率，D[i] 为第 i 帧的失真，S 为帧的复杂度参数，Q[i] 为第 i 帧的量化步长， $X_1[i]$ 和 $X_2[i]$ 为 R-1/Q 模型的参数值，随着每一帧图像编码自适应的更新， $\varepsilon=3$ 为常量；在该率失真模型中，^[4]已经验证了比特率 R 与量化参数的倒数存在线性关系，其中，量化步长直接决定帧的输出码率和失真，输出码率与量化步长成反比关系；因为每帧图像的复杂度不一样，R-1/Q 之间的线性关系也不一样，这里采用 MAD 来表示帧的复杂度参数，即 $S = MAD[i]$ ，其中： $MAD[i] = \sum_{NM} |f[n,m] - f_p[n,m]|$ ， $f[n,m]$ 为原始视频帧 (n,m) 处的像素值， $f_p[n,m]$ 为预测视频帧 (n,m) 处的像素值，[N,M] 为图像帧的大小。于是，码率的控制可以自适应的根据视频内容来决定控制参数，不同的复杂度的视频序列，控制参数不一样，可以得到准确的率预测模型；另外，视频编码的失真主要来源于残差量化失真，DCT 变换之后的残差编码一般服从拉普拉斯分布，得到的失真模型是非常复杂的数学表达式，为了简化，我们采用二次曲线拟合该失真曲线，从而得到了该失真模型

$D[i] = \frac{Q^2[i]}{\varepsilon}$ ，其中，量化步长与失真 D[i] 之间成反比关系。根据香农的率失真理论，在给定的码率条件下，使得失真最小，采用拉格朗日乘法得到：

$$J = D + \lambda R, \text{ 求 } \arg \min_{Q[i]} J, \text{ 求导得:}$$

$$\begin{cases} \frac{\partial J}{\partial Q} = \frac{2}{\varepsilon} Q - \lambda MAD[i] \frac{X_1[i]}{Q^2} = 0 & (4) \\ \frac{\partial J}{\partial \lambda} = MAD[i] \frac{X_1[i]}{Q} + X_2[i] - R = 0 & (5) \end{cases}$$

$$\text{由 (4) 得: } \frac{2}{\varepsilon} Q = \lambda MAD[i] \frac{X_1[i]}{Q^2} \Rightarrow Q =$$

$$\sqrt[3]{\frac{\lambda \cdot MAD[i] \cdot X_1[i]}{2}} \quad (4)$$

$$\text{将 (6) 代入 (5) 得: } \frac{MAD[i] X_1[i]}{\sqrt[3]{\varepsilon \lambda S X_1[i]}} + X_2[i] - R = 0$$

$$\Rightarrow \lambda = \frac{MAD[i]^2 X_1^2[i]}{(R - X_2[i])^3 \varepsilon} \quad (7)$$

$$\text{将 (7) 代入 (6) 得: } Q = \frac{1}{\sqrt[3]{2\varepsilon}} \frac{MAD[i] X_1[i]}{R - X_2[i]}.$$

资助信息：海南省自然科学基金项目 (No.807063,807062) (the Province Natural Science Foundation of Hainan under Grant No.807063 807062); 海南省教育厅高等学校科学研究项目 (No. HJ2009-99)

3 码率控制算法

本文沿用了 JVT-G012^[10]码率控制算法的分配策略,对于 I 帧和 GOP 中的第一个 P 帧,采用固定的量化参数进行编码处理;对于其他 P 帧,则根据目标码率和视频信源的实际编码特性为每个宏块逐个计算最优的量化参数。码率控制算法的步骤如下:

3.1 初始化

对于 GOP 中第一个 I 帧和第一个 P 帧,采用固定的量化参数 Q_0^t 进行编码, t 为帧类型 I, P, B 帧,对第一帧进行编码之后,初始化模型参数: $X_1[i]$ 和 $X_2[i]$ 。

3.2 帧层目标比特分配

目标比特: $R_{sum}[i] = (1 - \alpha) \times R_b[i] + \alpha \times R_r[i]$, α 一般取 0.75, 其中, R_b 涉及到目标比特与缓冲器关系,包括了目标缓冲大小,实际缓冲数据量,帧率与带宽;其表达式为: $R_b[i] = \frac{R_w[i]}{F} + \beta \times (B_t[i-1] - B_c[i])$, $B_c[i-1]$ 为前一时刻缓冲器实际的数据量, $B_t[i]$ 为当前时刻缓冲器目标数据量, F 为帧率, $R_w[i]$ 为当前时刻可获得的带宽, β 为常量,一般取 0.5。同时,帧层目标比特也与 GOP 中剩余比特数量有关,即 $R_r[i] = \frac{R_{rs}[i]}{N_r[i]}$, R_{rs} 为剩余比特量, N_r 为在 GOP 中未编码的剩余帧数。

3.3 确定帧层量化参数

由公式 (1) 得到帧层量化值为: $Q[i] = \frac{MAD[i]X_1[i]}{R_{sum}[i] - X_2[i]}$ 。为了保持视频序列的视觉质量平滑, Q 值进一步控制在一定范围内 $Q[i] = \min\{\bar{Q}[i-1] + 2, \max\{\bar{Q}[i-1] - 2, Q[i]\}\}$ 。

3.4 宏块级码率控制

$$(1) \text{ 宏块层目标比特: } T[i] = \frac{MAD[i]}{\sum_{i=1}^N MAD[i] \cdot w_i} R_{sum},$$

其中 $MAD[i]$ 为宏块 MB 的 MAD 值, w_i 为每个宏块的权重因子, 这里我们取 $w_i=1$ 。

(2) MB 层 MAD 预测: 由式 (1)、(2) 预测得到 MB 层的 MAD 值。

(3) 计算第 i 个 MB 的量化参数 $Q[i]$:

$$Q[i] = \frac{1}{\sqrt[3]{2}} \frac{MAD[i]X_1[i]}{T[i] - X_2[i]}$$

Q 值的选取应该与前一 MB 的值相关, 因此:

$Q[i] = \min\{Q[i-1] + 1, \max\{Q[i-1] - 1, Q[i]\}\}$; 另外, 为了保证时域上视觉质量的平滑, 限定量化参数值: $Q[i] = \min\{\bar{Q} + 3, \max\{\bar{Q} - 3, Q[i]\}\}$; 最后, 量化参数值应该在标准限制范围内: $0 < Q[i] \leq 51$ 。

(4) 更新 MB 级 R-Q 模型 $X_1[i]$ 和 $X_2[i]$ 和 MAD 预测模型参数 $Z_1[i]$ 和 $Z_2[i]$ 。

3.5 更新模型参数

更新帧级 R-Q 模型 $X_1[i]$ 和 $X_2[i]$ 和 MAD 预测模型参数 $Z_1[i]$ 和 $Z_2[i]$ 。

4 实验分析

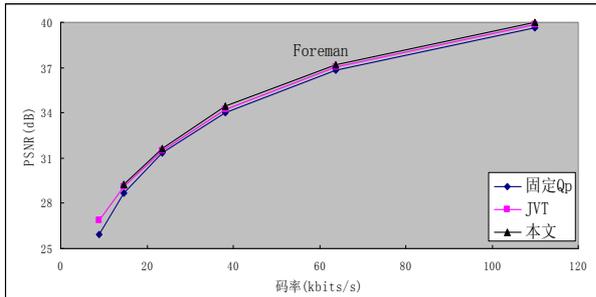
以上理论上推导得出了相应的算法, 这一节验证本文方法的性能。仿真环境采用 H.264 最新的软件参考平台 JM15.1, 在该环境下实现了以上码率控制算法, 并与该环境下实现的宏块层码率控制方法进行比较; 在相同的编码条件下, 分别对比了以下性能指标: 编码效率, 峰值信噪比, 实际码率输出几个重要的参数。

仿真实验参数设置分别为: 编码帧数为 100 帧, 帧率为 15 帧/秒, 参考帧为 1 帧, 搜索范围为 16, 采用 CAVLC 编码, 使用率失真优化, 编码结构为 IPP, 运动精度为 1/4 像素。限于篇幅, 文章给出了 Silent, News, Foreman 和 Mobile 等 4 种代表不同运动激烈程度的典型视频序列的实验数据, I 帧和 P 帧使用固定的初始量化参数进行编码处理, 其中, 目标码率的设定采用固定量化参数的编码码率输出数据, 初始量化参数一般采用固定量化参数值减 4 的量化参数。

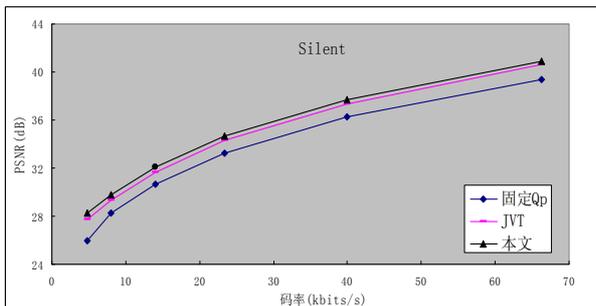
图 1 给出了固定量化参数、JVT 和新算法的 4 个典型视频的编码效率曲线, 其中横坐标为码率, 纵坐标为视频质量 PSNR 值。从图中可以看出, 4 个不同类型的视频序列得到的结果一致, 新算法的曲线均位于其他两条曲线的上方。在相同的码率输出的情况下, 相比固定量化参数编码, 采用新算法码率控制的 PSNR 值的增益分别为: Foreman: 0.52dB, Mobile: 0.51 dB, News: 1.33 dB, Silent: 1.09 dB, 相比 JVT 自适应码率控制方法, 同样得到了最高 1.33dB 的增益。

图 2(a)(b)给出了新算法与固定量化参数、JVT 算法的 Silent 视频序列的 PSNR 值比较和输出码率比较, 其中, 横坐标为帧序号, 纵坐标分别为 PSNR 和输出码率。可以明显的看出, 新算法的平均 PSNR 有显著

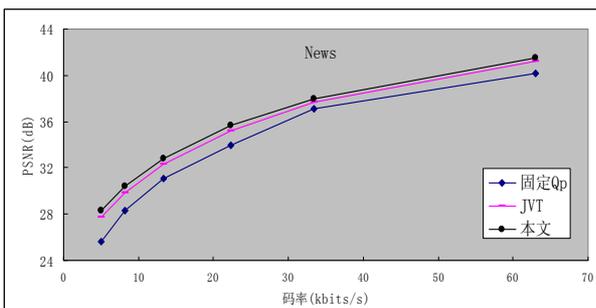
提高,与 JVT 码率控制方法相比,最高达 0.48dB。同时,如图 2 (b) 所示,编码码率输出更接近目标码率输出,比其他两种方式更平滑稳定。



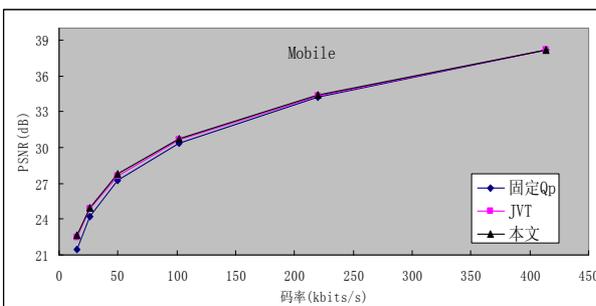
(a) Foreman:QCIF



(b) Silent: QCIF



(c) News: QCIF

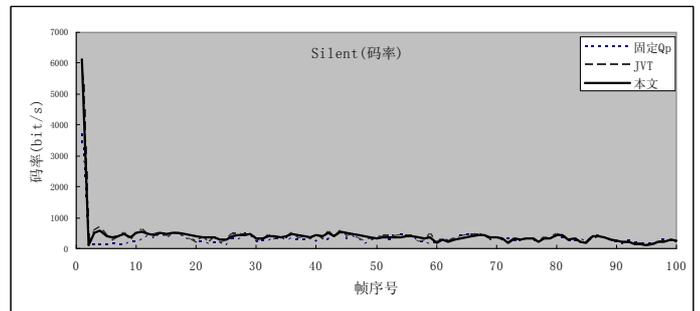


(d) Mobile:QCIF

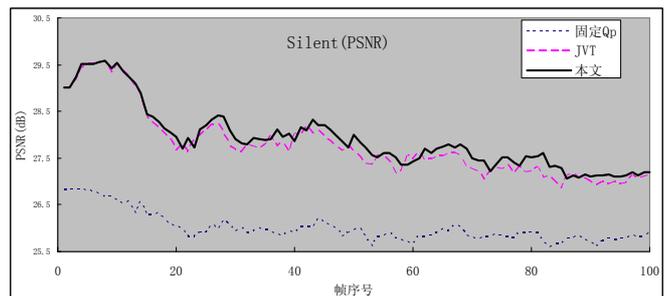
Figure 1. Curve : Rate-distortion of video sequences
图 1. 视频序列率失真曲线

5 结论

本文分析了视频压缩标准 H.264 的特性和现有率



(a) Comparison of frame-by-frame PSNR



(b) Comparison of frame-by-frame number of coding bits

Figure 2. Performance comparison of silent

图 2. Silent 编码输出效果比较

失真模型不精确性,以及低延时视频通信的码率控制问题,结合自适应 MAD 预测和线性 R-Q 模型,得出了新的量化参数控制模型。实验结果表明,该方法比 JVT 码率控制方法性能更优越,率失真模型更加精确,得到的码率控制算法更加灵活,相同输出码率下,得到的视频质量 PSNR 值更高,提高了编码效率,同时,新的码率控制方法得到的输出比特率更加平滑,适合与低延时实时视频通信应用。

致谢

感谢海南省自然科学基金项目(No.807063,807062)和海南省教育厅研究项目(No. HJ2009-99)的资助。

References (参考文献)

- [1] T.Wiegand, G.J.Sullivan, and A.Luthra, Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264---ISO/IEC 14496-10 AVC), Doc. JVT-G050r1. Geneva, Switzerland, 2003.
- [2] T.Wiegand, G.J.Sullivan, G.Bjontegaard, and A.Luthra, Overview of the H.264/AVC video coding standard[J].IEEE Transaction on

- Circuits and Systems for Video Technology, 2003, 13(7), P560-576.
- [3] T.Wiegand, H.Schwarz, A.Joch, F.Lossentini, and G.J.Sullivan, Rate constrained coder control and comparison of video coding standards[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2003, 13(7),P688-703.
- [4] Z.He and S.K.Mitra, A unified rate-distortion analysis framework for transform coding[J]. IEEE Transaction on Circuits and Systems for Video Technology,2001,11(12),P1221-1236.
- [5] Z.G.Li,F.Pan,K.P.Lim,G.Feng,X.Lin,and S.Rahardja. Adaptive basic unit layer rate-control for JVT[S],presented at the 7th JVT Meeting,Pattaya II JVT-G012-rl Thailand, 2003.
- [6] S.W.Ma,W.Gao,and Y.Lu, Rate-distortion analysis for H.264/AVC video coding and its application to rate control[J].IEEE Transaction Circuits and Systems for Video Technology, 2005,15(12):1533-1544.
- [7] Shang Shulin,Du Qingxiu,Lu Hanqing.A Low-Complexity Macroblock Level Rate Control for H.264[J]. Chinese Journal of Computers,2006,29(6),P914-919 (Ch).
- [8] Yuan Wu,Lin Shouxun,Niu zhenDong,Luo Hianyong,Zhang Yongdong, Efficient Rate Control Scheme for H.264/ AVC [J].Chinese Journal of Computers,2008,31(2),P329-339(Ch).
- [9] Yang Liu,Zhengguo Li and Yeng Chai Soh, A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard[J]. IEEE Transaction on Circuits and Systems for Video Technology,2007,17(1),P68-78.