

一种新的视频编码二次率失真模型及其性能分析

万 帅^① 常义林^① 杨付正^① 罗 忠^②

^①(西安电子科技大学综合业务网国家重点实验室 西安 710071)

^②(华为技术有限公司 深圳 518057)

摘 要: 基于对视频编码比特率与内容复杂度之间关系的研究, 该文提出了一种新的视频编码二次率失真模型。该模型保留泰勒级数中的常数项, 并通过合理引入内容复杂度, 提高了描述实际信源率失真特性的准确性。大量实验分析表明该模型具有良好的性能, 可广泛应用于各种速率控制算法以及其他使用率失真模型的场合, 以提高视频编码器的率失真性能。

关键词: 视频编码; 二次率失真模型; 方差分析; 速率控制

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2007)05-1136-04

A New Quadratic Rate Distortion Model for Video Coding and Its Performance Analysis

Wan Shuai^① Chang Yi-lin^① Yang Fu-zheng^① Luo Zhong^②

^①(State Key Lab. of Integrated Service Networks, Xidian University, Xi'an 710071, China)

^②(HuaWei Technology Limited Cooperation, Shenzheng 518057, China)

Abstract: Based on the study of the relationship between the bit rate for video coding and the content complexity, this paper presents a new quadratic rate distortion model for video coding, which retains the constant term in Taylor series. Through introducing the content complexity term reasonably, this model provides an accurate description of the rate distortion characteristic of the practical source. The performance of the model is evaluated through analyzing large amount of experimental data. This model can be used in rate control schemes for video coding and other applications where rate distortion model is in demand, by which the rate distortion performance of video coding systems can be improved.

Key words: Video coding; Quadratic rate-distortion model; Analysis of variance; Rate control

1 引言

率失真模型(Rate Distortion Model, RDM)在视频编码及传输技术中有着十分重要的作用。使用精确、恰当的率失真模型优化选择一系列编码参数, 可以使得视频通信系统在给定的速率限制下获得最好的编解码质量, 从而满足信道带宽以及编解码缓冲区等工作条件对编码速率的要求。目前视频编码速率控制及率失真优化中常用的率失真模型有对数模型^[1], ρ 域模型^[2], 二次模型^[3-5], 以及对数模型与二次模型相结合的混合模型^[6,7]等, 其中二次模型由于其简洁的描述形式和较准确的性能获得了广泛的应用^[4,5,8]。本文通过对原有的二次模型进行分析, 指出了其中存在的问题, 并在此基础上提出了一种新的二次模型。根据模型性能分析的结果可见, 本文提出的二次模型能够更为准确地反映编码视频的率失真特性, 与原二次模型相比具有较优的性能。

2 二次率失真模型简介及其存在的问题

Chang 等于 1997 年提出了一种采用二次多项式描述视

频速率失真特性的方法^[3], 假设信源符合拉普拉斯分布, 则有如下对数形式的率失真模型:

$$R(D) = \begin{cases} \ln[1/(\alpha D)], & 0 < D < 1/\alpha \\ 0, & D \geq 1/\alpha \end{cases} \quad (1)$$

其中 R 为编码速率, D 为失真, α 为 Laplace 分布参数, 且 $\alpha > 0$ 。文献[3]对上述率失真模型进行泰勒展开, 然后采用量化步长作为失真的衡量, 则可获得一种率失真函数的二次多项式表达形式, 本文记为二次率失真模型 1:

$$R = a_1 \times Q^{-1} + a_2 \times Q^{-2} \quad (2)$$

其中 R 和 Q 分别是编码比特数和量化步长, a_1 , a_2 是模型参数。该模型在 MPEG-4 速率控制算法中获得了有效应用^[3]。

Lee 等指出二次模型 1 应区分用于纹理信息和非纹理信息编码的比特数^[4], 引入运动预测后原始图像与预测图像之间帧差信号绝对值的平均值(Mean Absolute Difference, MAD)作为图像内容复杂度的度量, 从而给出了描述编码纹理信息所用的比特数与图像内容复杂度及量化步长之间关系的二次模型, 本文记为二次率失真模型 2:

$$\frac{R_T}{MAD} = a_1 \times Q^{-1} + a_2 \times Q^{-2} \quad (3)$$

2005-10-31 收到, 2006-04-20 改回

国家自然科学基金(60172030)和华为科技基金(YJCB2003017MU)资助课题

其中 R_T 为编码纹理信息所用的比特数。二次模型 2 提高了描述信源编码速率与失真、信源内容之间的关系的准确性。目前二次模型 2 已广泛应用于现有视频编码速率控制方案中,如 MPEG-4 的速率控制方案^[4,5], H.264 的速率控制方案^[8]等。

二次率失真模型 1 和模型 2 的推导均是基于相同的泰勒展开过程。根据函数展开成泰勒级数的条件,若函数 $f(x)$ 在 x 的某一邻域内具有直到 $(n+1)$ 阶的导数,则在该邻域内 $f(x)$ 可以展开为 n 阶泰勒级数。然而,根据率失真函数式(1)的定义, $f(x)$ 在 $x=1$ 处即 $D=1/\alpha$ 处不可导,因此在二次率失真模型的推导过程中将 $f(x)$ 在 $x=1$ 处展开成泰勒级数是不合适的。

此外,二次率失真模型 2 的表达形式说明了在量化步长一定的情况下,编码纹理信息所用的比特数 R_T 与编码单元的内容复杂度(以 MAD 衡量)成正比关系。然而通过大量实验我们发现,当量化步长一定时, R_T 与 MAD 之间更接近一种常数项不为零的一般线性关系,而不是一种简单的正比关系。可见,在基于泰勒级数展开的推导过程中,应该保留泰勒级数中的常数项,并根据实验数据合理引入复杂度,使得模型的形式符合当量化步长一定时, R_T 与 MAD 呈一种常数项不为零的一般线性关系。图 1 以 CIF 格式 Claire 序列第 6 帧和 QCIF 格式 Foreman 序列第 34 帧为例给出了各宏块 R_T 与 MAD 的关系图。

由上述理论分析和实验数据可知,已有的二次率失真模型的理论推导过程存在一定的问题。并且,虽然在二次模型 2 引入图像内容复杂度度量是合理的,但是二次模型 2 中 MAD 的引入未能很好符合实际实验观察的结果,在实际应用中需要研究更为准确、合理的二次率失真模型。

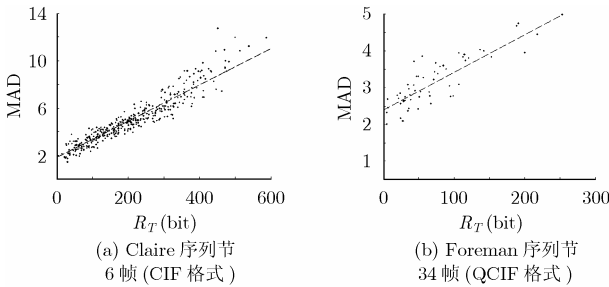


图 1 R_T 与 MAD 关系图(采用 H.264 编码,量化因子 25)

3 新的二次率失真模型

根据上文的分析,若假设信源具有拉普拉斯分布的统计特性,则可以将对数形式的分段函数 $f(x)$ 在 $x = x_0$ ($x_0 > 1$) 处展开成泰勒级数,并代入 $x = 1/(\alpha D)$,从而获得以下表达形式:

$$R(D) = \ln x_0 + \frac{1}{x_0} \left(\frac{1}{\alpha D} - x_0 \right) - \frac{1}{2!(x_0)^2} \left(\frac{1}{\alpha D} - x_0 \right)^2 + R_3(D) \\ = \left(\ln x_0 - 1 - \frac{1}{2!} \right) + \frac{2}{x_0 \alpha} D^{-1} - \frac{1}{2!(\alpha x_0)^2} D^{-2} + R_3(D) \quad (4)$$

其中 $R_3(D)$ 为泰勒级数展开的剩余高次项。同样使用平均量化步长作为失真的测度^[2],并且考虑区分用于编码纹理信息和非纹理信息的比特数,本文初步将率失真方程近似为以下形式的二次模型:

$$R_T = c_0 + c_1 \times Q^{-1} + c_2 \times Q^{-2} \quad (5)$$

其中 R_T 为编码纹理信息所需的比特数, Q 为量化步长, c_0, c_1, c_2 为模型参数。然后根据实验数据的统计特性,在率失真函数中引入图像内容复杂度度量 MAD^[4],则可以获得新的二次率失真模型为

$$\frac{R_T}{\text{MAD}} = \frac{c_0}{\text{MAD}} + c_1 \times Q^{-1} + c_2 \times Q^{-2} \quad (6)$$

本文将式(6)记为二次率失真模型 3。该模型的表达形式说明,当量化步长一定时,编码纹理信息所用的比特数与 MAD 之间是一种线性关系,符合实验观察的结果。该模型各参数的极大似然估计可以通过二元线性回归的方法求得^[9]。

4 模型性能分析

为了进一步说明二次率失真模型 3 更符合实际信源的速率失真特性,对二次率失真模型 2 和模型 3 分别进行性能分析和对比,具体的步骤为:

(1)获得实验数据点 本文采用 H.264 视频编码器(测试模型 JM 7.6)对视频标准测试序列进行编码。为了更好地说明问题,我们对第一帧进行帧内编码,后续帧均采用帧间编码,即编码帧序列的格式是 IPPP...PPP。采用固定的量化因子对整个序列进行编码,用于测试的量化因子的范围是 1-51。对于每个特定的量化因子,分别测量编码每一帧编码亮度信号(Y)纹理信息所用的比特数以及每个宏块的 MAD 值。设第 j 帧编码亮度信号纹理信息所用的比特数为 $R_{j,T}$,可按照下列两式求得该帧平均编码每一宏块亮度信号(Y)纹理信息所用的比特数 $\overline{\text{MR}}_j$,以及该帧平均每个宏块的 MAD,记为 $\overline{\text{MAD}}_j$ 。

$$\overline{\text{MR}}_j = (1/N)R_{j,T} \quad (7)$$

$$\overline{\text{MAD}}_j = \frac{1}{N} \sum_{i=1}^N \text{MAD}_{j,i} \quad (8)$$

其中 N 为一帧图像包含的宏块数, $\text{MAD}_{j,i}$ 为第 j 帧图像中第 i 个宏块的 MAD。可见,对于序列中的第 j 帧图像,可以获得 51 组 $(\overline{\text{MR}}_j(Q_i), \overline{\text{MAD}}_j(Q_i), Q_i, i = 1, \dots, 51)$ 数据。本文对每一序列使用 30 个 P 帧的数据进行统计分析。

(2)模型参数求解和数据预测 对于二次率失真模型 2,使用文献[4]给出的一元线性回归方法求解参数 a_1, a_2 的极大似然估计值 \hat{a}_1, \hat{a}_2 。将各帧的所有数据组 $(\overline{\text{MAD}}_j(Q_i), Q_i)$ 代入二次率失真模型 2,可以预测出编码纹理信息所需比特数的估计值 $\overline{\text{MR}}_{2,j}(Q_i)$ 。

$$\overline{\text{MR}}_{2,j}(Q_i) = \overline{\text{MAD}}_j(Q_i) \times (\hat{a}_1 \times Q_i^{-1} + \hat{a}_2 \times Q_i^{-2}) \quad (9)$$

对于二次率失真模型 3,使用二元线性回归^[9]求解模型参数的极大似然估计值 $\hat{c}_0, \hat{c}_1, \hat{c}_2$ 。将各帧的所有数据组

($\overline{MAD}_j(Q_i)$, Q_i)代入二次率失真模型 3, 预测出编码纹理信息所需的比特数的估计值 $\overline{MR}_{3,j}(Q_i)$ 。

$$\overline{MR}_{3,j} = \hat{c}_0 + \overline{MAD}_j(Q_i) \times (\hat{c}_1 \times Q_i^{-1} + \hat{c}_2 \times Q_i^{-2}) \quad (10)$$

(3)模型性能分析 根据编码纹理信息的比特数的原始数据 MR_j 和利用各率失真模型预测出的数据 $\overline{MR}_{k,j}$ ($k=2,3$) 分别对二次率失真模型 2 和模型 3 进行模型性能分析。首先计算各模型预测的数据与原始数据之间的相关系数, Foreman(QCIF 格式)序列和 Calendar(CIF 格式)序列的结果见表 1。从表 1 可以看出, 与模型 2 相比, 模型 3 预测出的数据与原始数据之间有较大的相关系数(更接近 1), 相关性更好, 可见模型 3 能够更准确地描述实际信源的率失真特性。

表 1 相关系数表

	模型 2	模型 3
Foreman	0.9141	0.9960
Calendar	0.9053	0.9934

另外, 本文还对模型 2 和模型 3 的性能进行了方差分析。在方差分析中, 通常使用 F 比 (F ratio) 作为模型的显著性检验标准^[10]。即 F 比越大, 误差越小, 表示模型和实际情况的匹配越好, 模型越精确。 F 比定义为

$$F = \frac{SS/(p-1)}{RSS/(n-p)} \quad (11)$$

其中 n 为数据点的个数, p 为模型参数的个数, SS 为回归平方和, RSS 为剩余平方和。设 X 为实际数据值, \bar{X} 为实际数据的平均值, \hat{X} 为模型的预测数据值, 则 $SS = \sum (\hat{X} - \bar{X})^2$, $RSS = \sum (X - \hat{X})^2$ 。

对模型 2 和模型 3 可以分别建立表 2 形式的方差分析表^[10]。我们使用多个序列对模型性能进行分析, 采用上文所述的方法获得实验数据点, 则每一序列的方差分析均包含 1530 个数据点。表 3, 表 4 分别给出了 Foreman(QCIF 格式)序列和 Calendar(CIF 格式)序列的方差分析结果, 其它序列可以获得类似的结果。

表 2 方差分析示例

方差	平方和	自由度	均方	F 比
回归	SS	$p-1$	$SS/(p-1)$	F
剩余	RSS	$n-p$	$RSS/(n-p)$	
总计	TSS	$n-1$		

表 3 Foreman 序列的方差分析结果

	方差	平方和	自由度	均方	F 比
率失真模型 2	回归	1.73×10^8	1	1.73×10^8	8.02×10^3
	剩余	3.29×10^7	1528	2.15×10^4	
	总计	2.06×10^8	1529		
率失真模型 3	回归	9.26×10^7	2	4.63×10^7	7.74×10^4
	剩余	9.13×10^5	1527	598.4689	
	总计	9.35×10^7	1529		

表 4 Calendar 序列的方差分析结果

	方差	平方和	自由度	均方	F 比
率失真模型 2	回归	4.10×10^8	1	4.10×10^8	7.09×10^3
	剩余	8.83×10^7	1528	5.78×10^4	
	总计	4.98×10^8	1529		
率失真模型 3	回归	2.08×10^8	2	1.04×10^8	5.75×10^4
	剩余	2.76×10^6	1527	1.81×10^3	
	总计	2.11×10^8	1529		

从各方差分析表可以看出, 与模型 2 相比, 模型 3 预测出的速率和实际速率之间具有较小的方差, 其中回归方差的均方值仅为模型 2 的 25.4%–26.6%, 剩余方差的均方值仅为模型 2 的 2.7%–3.1%。同时, 模型 3 具有较高的 F 比, 通常比模型 2 的 F 比高一个数量级。因此二次率失真模型 3 更为精确, 能够更好地反映实际编码视频信源的率失真性能。

图 2 给出原始速率数据和采用两个模型预测的速率数据与量化因子之间的关系。值得说明的是, 在 H.264 中, 量化因子的范围可达 1–51, 它们和量化步长之间的对应关系与以往的标准不同, 例如量化因子为 10 时对应的量化步长仅为 2, 因此量化因子较小时, 会产生较大的视频编码速率。从图 2 可以直观的看出, 模型 3 具有较好的预测效果, 在整体速率变化区域, 预测的数据点更为接近实际数据点。而模型 2 预测误差相对较大, 尤其是在中高速率(即量化因子较小)预测效果欠佳。

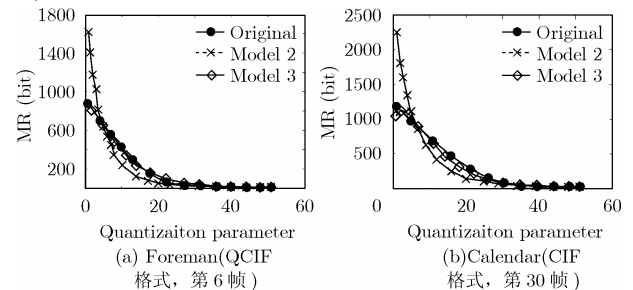


图 2 速率与量化因子关系图

综上所述, 与二次率失真模型 2 相比, 二次率失真模型 3 预测出的数据与原始数据之间有着更好的相关性, 并且 F 比较大, 预测方差较小, 进一步提高了预测效果。因此二次率失真模型 3 的形式能够更准确地反映编码纹理信息所需的比特数与图像内容复杂度(MAD)、量化步长之间的关系, 即能够更准确地描述实际信源的率失真特性。

5 结束语

本文通过理论研究和对实际数据的分析, 提出了一种新的二次率失真模型, 该模型仍然采用多项式的形式, 参数求解简单。模型性能分析的实验数据说明, 与原有的二次率失真模型相比, 本文提出的新的二次模型能够更准确地反映编码纹理信息所需的比特数与图像内容复杂度(MAD)、量化步长之间的关系, 从而进一步提高了描述实际信源的率失真特性的准确性。该模型在实际的视频编码系统有着良好的应

用价值, 可以应用于不同的速率控制算法以及其他使用率失真模型的场合, 以提高视频编码器的率失真性能。

参 考 文 献

- [1] Hang H and Chen J. Source model for transform video coder and its application, – part I: Fundamental theory. *IEEE Trans. on Circuits Syst. on Video Technology*, 1997, 7(2): 287–298.
 - [2] He Z, Kim Y and Mitra S. Low-delay rate control for DCT video coding via ρ -domain source modeling. *IEEE Trans. on Circuits Syst. on Video Technology*, 2001, 11(8): 928–940.
 - [3] Chiang T and Zhang Y. A new rate control scheme using quadratic rate distortion model. *IEEE Trans. on Circuits Syst. on Video Technology*, 1997, 7(1): 246–250.
 - [4] Lee H, Chiang T, and Zhang Y. Scalable rate control for MPEG-4 Video. *IEEE Trans. on Circuits Syst. on Video Technology*, 2000, 10(6): 878–894.
 - [5] Vetro A, Sun H, and Wang Y. MPEG-4 rate control for multiple video objects. *IEEE Trans. on Circuits Syst. on Video Technology*, 1999, 9(1): 186–199.
 - [6] Wei J, Soong B, and Li Z. A new rate-distortion model for video transmission using multiple logarithmic functions. *IEEE Signal Processing Letters*, 2004, 11(8): 694–697.
 - [7] Corbera J and Lei S. Rate control in DCT video coding for low-delay communications. *IEEE Trans. on Circuits Syst. on Video Technology*, 1999, 9(1): 172–185.
 - [8] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG Document JVT-H014. Adaptive Rate Control with HRD Consideration. ftp://standards.polycom.com/imtc_jvtexperts/, May, 2003.
 - [9] 盛骤, 谢式千, 潘承毅. 概率论与数理统计. 北京: 高等教育出版社, 1989, 第 9 章.
 - [10] 中科院数学研究所数理统计组. 回归分析方法. 北京: 科学出版社, 1975, 第 3、4 章.
- 万 帅: 女, 1979 年生, 博士生, 研究方向为视频压缩、速率控制、抗误码等.
- 常义林: 男, 1964 年生, 教授, 研究方向为网络多媒体、网络管理等.
- 杨付正: 男, 1977 年生, 讲师, 研究方向为视频压缩、率失真、质量评估等.
- 罗 忠: 男, 1971 年生, 高级工程师, 研究方向为视频压缩、音频压缩、多媒体等.