

# 基于 Lagrange 的 H. 264 率失真编码优化算法<sup>\*</sup>

李洛<sup>1</sup> 张剑<sup>2</sup>

(广东轻工职业技术学院 广州 510300)<sup>1</sup> (广东纺织技术学院 佛山 528041)<sup>2</sup>

**摘要** 为了进一步提高编码速度,促进视频编码技术在多媒体通信领域的实际应用,本文对率失真视频编码的关键算法进行了深入的研究并提出了其优化方法,并从信息论理论出发,着重围绕图像压缩的理论,利用率失真特征数学模型、拉格朗日乘法,对目前新一代的 H. 264 视频标准进行编码算法模式选择,试图从最根本的理论角度探讨信源编码,以便得到更高的压缩比、更好的图像质量和更快的编码速度。最后,利用本文提出的编码模式优化算法,对 H. 264 的 JM61e(JUT Test Model)系统测试模型的性能进行全面测试。

**关键词** 率失真函数,拉格朗日乘法,信息论,编码器, H. 264 标准

## The Algorithm of H. 264 Constrained Distortion Coding Optimization Based on Lagrange

LI Luo<sup>1</sup> ZHANG Jian<sup>2</sup>

(Guangdong Industry Technical College, Guangzhou 510275)<sup>1</sup> (Guangdong Textile Polytechnic, Foshan 528041)<sup>2</sup>

**Abstract** Based on the thorough study of key algorithms of constrained distortion video codec, several optimization methods are proposed in this paper to enhance the coding speed and promote the applications of video coding in actual practice. Entropy theory and image compression theory adopted, the model selection on coding algorithm of the new standards of video coding which named H. 264 is proposed based on constrained distortion feature mathematical model and lagrange multipliers. For achieving the higher compression ratio, the better picture quality and quicker coding speed, we probe into the ultimate theory on source coding. In the last of this paper, a total performance test is performed on the M61e(JUT Test Model) system test model with the method of coding model selection we proposed.

**Keywords** Rate-Distortion function, Lagrange multipliers, Entropy theory, Coder, H. 264 standard

## 1 引言

JVT 小组在 H. 264 编解码软件中提供的多解码器模拟信道误码下的编码模式率失真判决算法(MD),模拟了多种信道失真情况,在相同的误码概率下编码器启用了较多的帧内刷新宏块<sup>[1]</sup>,因而编码器输出码率较大,且运算量也相当大。Zhang 等<sup>[2]</sup>提出的判决算法(ROPE),把所有的传输误码失真和编码失真全部放在一起计算,可以获得较好的视频质量,但是计算太过繁琐,难以实时实现,并且对率失真判决函数中  $\lambda$  的取值不适合运用在 H. 264 编码中。本文在 H. 264 新标准的测试软件中使用了拉格朗日率失真优化策略,拉格朗日方法虽然能够找出给出一个或数个最好性能的编码参数集,并使这些参数用到视频编码器以获得最佳的率失真性能。但是,它是基于使用每种图像块尺寸和每种预测模式而产生的参差及其传输的码率。这样,模式选择可以取得优化的率失真性能,但这是以提高运算复杂度为代价的。另外,在此优化操作是对下面拉格朗日函数的最小化: $J = SATD + \lambda R$ ,其中, SATD 为经过哈德曼变换的  $4 \times 4$  块的预测参差绝对值总和<sup>[3]</sup>。对于所有帧内、帧间宏块编码模式及多参考帧的选择都通过对拉各朗日函数的最小化来实现。而通常视频标准只是包括解码规范,模式选择的技术研究是属于编码端的范畴,所以并没有列在 H. 264 标准之内。另外,编码参数组合的数目极大,计算复杂性常常比编码算法本身还高,只能应用在这些非实时编码中。因此,我们有必要采用一种视频抗误码编码模式,提前判断某些编码参数优劣,使用率失真特征模

型来估计最佳工作点等简化算法,来保证恢复图像的质量,构造适合于信道误码扩散估计和失真计算模型,有助于今后无线和 IP 视频通信产品的研制和开发。

## 2 率失真函数理论

为了达到在目标数据率的条件下使传输的视频信号的失真最小,在编码比特率和图像失真之间必需选择一个恰当的折衷,这是香农(Shannon)的信息论中的率失真理论问题<sup>[4]</sup>,它揭示了编码的数据率与失真度之间的关系,首次对信息的概念进行了科学的表述并给出了量度信息的方法,信息论中的无损信源编码定理和率失真理论是整个信源编码理论的基础。

**定义 1** 假设在有损压缩中原始图像是连续的,重构图像  $g(x, y)$  与原始图像  $f(x, y)$  之间的差别为失真度,则重构图像与原始图像之间总存在量化均方误差,为:

$$D = E[f(x, y) - g(x, y)]^2 = \frac{1}{N^2} \int_0^{N-1} \int_0^{N-1} [f(x, y) - g(x, y)]^2 dx dy \quad (1)$$

式中,  $D$  为失真量,  $N$  为信源长序列,从统计学上考虑,  $N$  趋于无穷大。

从香农信息编码定理知,只要在大于信息熵值的条件下编码,是可以达到信息无损译码的,但是其压缩比就会比较小,在实际应用的许多场合是允许有一定程度失真的,只要失真度小于某一允许范围,便可以获得更低的编码比特率。

率失真理论讨论的主要问题是:在允许一定程度失真的

<sup>\*</sup> 本课题得到全国教育科学十五规划重点课题(NO: AYA010034)和国家自然科学基金(NO. 69735020)资助。李洛 副教授,研究方向:计算机多媒体技术, Web 技术;张剑 副教授,高级工程师,博士,研究方向:计算机多媒体视频压缩算法及其实现。

条件下,能够把信源的信息压缩到什么程度。如图 1 所示,给出点对点通信系统的基本框图。

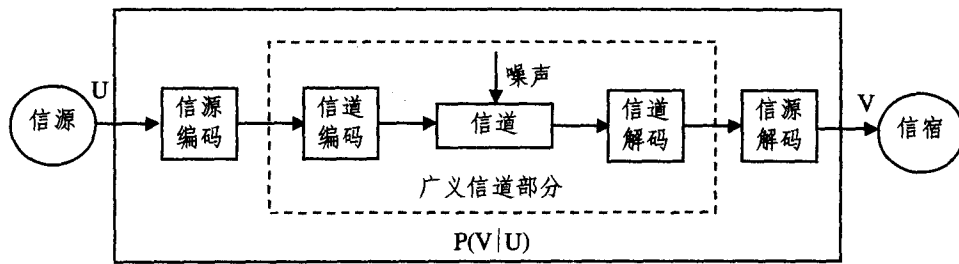


图 1 点对点通信系统框图

**定义 2** 假设信道中没有噪声产生,信息的失真仅与信源编码有关,其概率空间  $[U, P]$  为离散无记忆信源,  $U = \{u_1, u_2, \dots, u_k\}$ ,  $P = \{p(u_1), p(u_2), \dots, p(u_k)\}$ , 接受端变量  $V = \{v_1, v_2, \dots, v_l\}$ , 则平均失真, 为:

$$\bar{D} = \sum_{i=1}^k \sum_{j=1}^l p(u_i) p(v_j | u_i) d(u_i, v_j) \quad (2)$$

式中,  $d(u_i, v_j)$  为失真函数, 常见的失真函数有绝对值失真、平均失真等。

由式(2)可以看出, 如果信源的概率分布给定之后, 决定平均失真的仅仅是信源到信宿的转移概率。有了平均失真的表达式, 就可以定义率失真函数如下式:

$$R(D) = \min_{p_{ij} \in P_D} I(U, V) \quad (3)$$

式中,  $P(D)$  表示能使平均失真  $\bar{d} < D$  的所有转移概率的集合, 它是关于从  $U$  变换到  $V$  的互信息的及值问题,  $R(D)$  是在失真限于  $D$  以及  $p(x)$  已知的条件下的互信息的极小值,  $I(U, V)$  为接收端获得的平均信息量, 它的最小值仅与  $D$  有关。

可以看出, 率失真函数给出了在满足限失真准则(平均失真不大于  $D$ )的条件下, 信源必须传输的信息率  $R$  的下限值。当每个信源符号的比特数低于  $R(D)$  时, 不存在能使失真小于  $D$  的编码方法; 而给定允许的最大失真  $D$  时, 也没有一种编码方法可以是每个信源符号的比特数低于  $R(D)$ 。一般情况,  $R(D)$  是很难计算的, 只有在一些特殊情况下可以得到  $R(D)$  函数的表达式。另外, 根据不同的应用, 其最大允许失真度是不相同的。失真度越小, 编码比特率越大, 反之, 编码比特率越小<sup>[4]</sup>。

若定义最大允许失真度为  $D^*$ , 则其对应的编码比特率的下限  $R(D^*)$  是  $D^*$  的单调递减函数。称为率失真函数。率失真函数的定义表明: 一个具有率失真函数  $R(D^*)$  的信源, 倘若  $R < R(D^*)$ , 则不存在任何编码会使失真度小于  $D^*$ 。或者说, 没有任何一种编码方法可以使平均编码比特率小于  $R(D^*)$ , 而又使失真度小于  $D^*$  的。所以, 率失真函数提供了一把衡量实际所采用的编码方法效率是高或低的尺子, 由此可见, 率失真理论给我们提供了达到最佳效果的编码原则。

### 3 拉格朗日乘法

拉格朗日乘法转化为一个无约束的求极小值问题, 即:

$$\min_M \{D(X, M) + \lambda R(X, M)\} \quad (4)$$

其中, 失真  $D(X, M)$  和数据速率  $R(X, M)$  的计算分别如式(5)和式(6)所示。

$$D(X, M) = \sum_{i=1}^N D(X_i, M) \quad (5)$$

$$R(X, M) = \sum_{i=1}^N R(X_i, M) \quad (6)$$

失真的衡量方法有多种, 通常采用均方误差(MSE)或者其等价形式平方误差和(SSE)与峰值信噪比(PSNR)。如下式(7)、式(8)和式(9)分别给出宏块  $X_i$  在衡量方法下的失真:

$$D_{SSE}(X_i, M) = \sum_{s \in X_i} (Y(s) - \hat{Y}_M(s))^2 \quad (7)$$

$$D_{MSE}(X_i, M) = \frac{1}{\Omega} \sum_{s \in X_i} (Y(s) - \hat{Y}_M(s))^2 \quad (8)$$

$$PSNR(X_i, M) = 10 \log_{10} \frac{(255)^2}{D_{MSE}(X_i, M)} \quad (9)$$

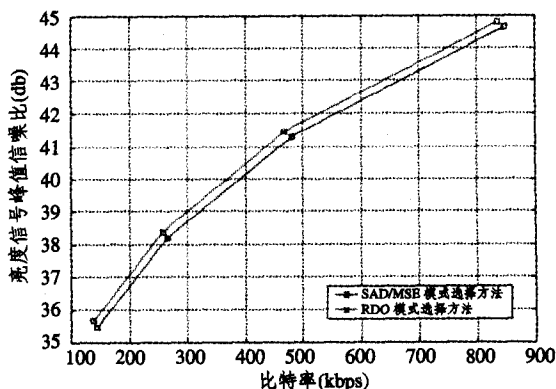
式中,  $\Omega$  为宏块中的像素个数,  $Y(s)$  和  $\hat{Y}_M(s)$  分别为点  $s$  编码前和按照模式  $M$  编解码后的重建值。此外, 另一种误差衡量函数为绝对误差和(SAD)函数, 它的计算方法如式(10)所示。

$$D_{SAD}(X_i, M) = \sum_{s \in X_i} |Y(s) - \hat{Y}_M(s)| \quad (10)$$

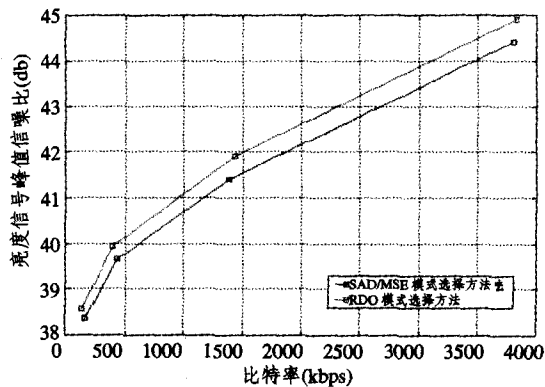
相对而言, SAD 具备计算简单, 直观性好的优点。对于宏块  $X_i$  在模式  $M$  下数据速率, 则需要编码器按照模式  $M$  对宏块集  $X$  中的所有宏块全部进行编码后才能得到。对于宏块  $X_i$  在模式  $M$  下数据速率, 则需要编码器按照模式  $M$  对宏块集  $X$  中的所有宏块全部进行编码后才能得到。然而, 由于各宏块之间的编码过程存在着复杂的相互影响关系, 求这个最优解并不是很容易的一件事情。本文拟给出最简单情况下求解方法, 即忽略各宏块编码过程的相互影响。

### 4 率失真编码模式选择

由上式(1)和率失真理论知: 最简单的模式选择方法就是以宏块内各点的值与其预测值的均方误差或平均绝对误差(INTRA 模式下可以认为预测值为 0)为依据, 选择均方误差或平均绝对误差最小的编码模式作为当前宏块的编码模式。在编码模式较少的情况下, 这种模式选择方法尚能满足编码器性能的要求, 但是, 随着编码模式的增加, 该方法开始严重地限制编码器整体性能的提高, 如图 2 所示。



(a) foreman 序列(QCIF)



(b) miss American 序列(CIF)

图2 H.264/AVC 中两种模式选择方法性能比较

为此,在失真不小于给定失真  $D$  的条件下,尽可能降低比特率。基于这一思想,将提出了一种新的模式选择方法,即采用率失真优化的模式选择方法。

定义3 假设  $X = \{X_1, X_2, \dots, X_N\}$  是图像中  $N$  个宏块组成的一个宏块集,  $X$  中的每一个宏块有  $K$  种编码模式,  $K$  种编码模式由集合  $I = \{I_1, I_2, \dots, I_K\}$  给出;若  $M_i \in I$  为宏块  $X_i$  采用的编码模式,宏块集  $X$  采用的编码模式为  $M = \{M_1, M_2, \dots, M_N\} \in I^N$ 。再根据限失真编码的准则,为宏块集  $X$  选择合适的编码模式  $M$  的问题成为一个有约束的非线性规划问题,则:

$$\min_M \{D(X, M)\} \quad \text{s. t.} \quad R(X, M) \leq R_c \quad (11)$$

式中,  $D(X, M)$  和  $R(X, M)$  分别表示宏块集  $X$  在编码模式  $M$  下的整体失真和数据速率,  $R_c$  为给定的速率限制。这个问题可以通过上面拉格朗日乘子法转化为一个无约束的求极小值来解决。

由上述式(4)知,若给定一个  $\lambda$ ,便可以得到一个最优解  $M_{opt}$ ,而这个  $M_{opt}$  也就是式(11)最优解。

定义4 假设  $D(X_i, M) = D(X_i, M_i)$ ,  $R(X_i, M) = R(X_i, M_i)$  成立,则式(11)和(4)简化为:

$$\min(D) \quad \text{s. t.} \quad R \leq R_c \quad (12)$$

$$\min(J) \quad \text{其中} \quad J = D + \lambda R \quad (13)$$

式中,变量  $D$  和变量  $R$  表示当前宏块在各种编码模式下的失真和数据速率,模式选择的基本单位为宏块。为了使式(12)中的  $J$  取到最小值,令  $\frac{\partial J}{\partial D} = 0$ ,得  $1 + \lambda \frac{dR}{dD} = 0$ ,则:

$$R(D) = k \ln\left(\frac{\sigma^2}{D}\right) \quad (14)$$

那么,由式(13)可以近似获得量化过程中失真与比特率的关系曲线,式中,  $\sigma^2$  为信源的方差,  $k$  是一个常数,取决于信源的概率密度函数。综述上式和 H.264 的编解码标准文[2]可知,满足 Lagrange 条件极值的乘积因子  $\lambda$  的值为:  $\lambda = \frac{D}{k}$ 。在

H.264 草案对于无误码环境中的  $\lambda$  取值为  $\lambda_0 = 0.85 \times 2^Q$ , 这里  $Q$  为宏块量化参数。

定义5 假设信源为均匀分布,则均匀量化的误差函数可以表示为:  $D = \frac{\Delta^2}{12}$ , 由此,可以得到  $\lambda$  和量化间隔  $\Delta$  的关系式为:  $\lambda = c \cdot \Delta^2$

Sullivan 在 JM6le 测试模型实验中给出的  $c$  值为 0.2125<sup>[3]</sup>, 这是一个比较合适的值,  $c$  值对编码器性能的影响如图3所示。

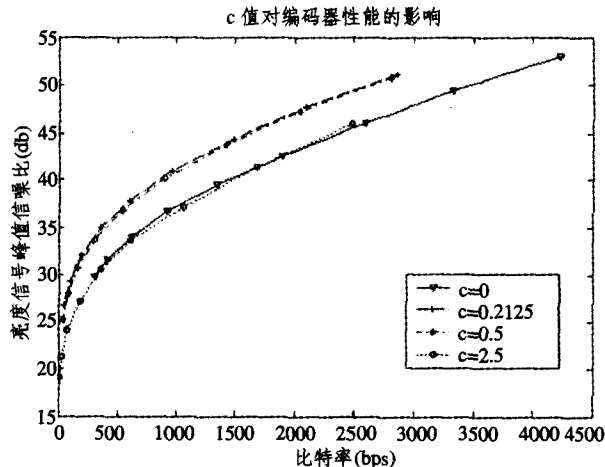


图3 H.264 编码测试模型(JM6le)中  $c$  的取值对编码性能的影响

由此可知,当  $c = 0.2125$  时,编码器的率失真性能达到最佳。适当调整  $c$  的取值,可以应用不同的环境。当  $c = 0$ ,从失真的角度来进行编码模式选择时,编码器的性能将急剧下降,这是因为大部分宏块都采用了失真较小、比特率庞大的帧内编码模式进行编码。另外,上述的率失真优化方法不仅可以应用于编码模式的选择过程,而且可以应用于最新的 H.264 标准技术运动预测的搜索过程<sup>[5]</sup>,即将误差的因素和信息率的因素结合起来作为选择最佳运动矢量的依据  $\min\{J_{motion}\}$ , 这可以避免采用一些幅度大,但对编码效率提高不是很明显的运动矢量,从而在总体上改善编码器的性能。其中,  $J_{motion} = SAD + \lambda_{motion} R_{motion}$ ,  $R_{motion}$  为对应运动矢量编码需要的比特数。在这样的依据下,可以避免采用一些幅度大,但对编码效率提高不是很明显的运动矢量,从而在总体上改善编码器的性能<sup>[6]</sup>。

### 5 实验测试与分析

信道测试环境实验,典型无线信道比特误码率在  $10^{-5} \sim 10^{-3}$  之间,本为采用 ITU 推荐的 3GPP/3GPP2 无线网络的 BER 分别为  $5 \times 10^{-4}$  和  $2 \times 10^{-4}$ , 基于 RTP/UDP/IP 协议的视频流传输通用环境<sup>[7]</sup>,以视频码流通过信道传输模拟后解码端重建图像的 PSNR (peak signal-to-noise ratio) 作为衡量标准,模拟 3G 中的传输数据分组协议,使用信道移动速度分别为 3 和 50km/h 的两个比特误码图样文件 wcdma.64kb.3kph.5e.04.bin 和 wcdma.64kb.50kph.2e.04.bin 来模拟实际信道比特误码,视频测试序列使用 foreman, hall monitor, car phone 和 silent4 种 QCIF 测试序列<sup>[8]</sup>。对多解码器模拟信道误码下的编码模式率失真判决算法 MD (MultiDecoder)、H.26L 草案提供的无误码率失真编码 NE 和本文的率失真编码模式优化选择方法测试,相应的实验结果如图4所示。

根据实验结果图可见,NE 和 MD 算法对不同测试序列性能差异很大,如图4(a)、(c)中 foreman 和 carphone 序列,相同码率下 NE 算法 PSNR 明显高于 MD 算法;但图4(b)、(d)中 hallmonitor 和 silent 序列的实验结果却相反。4 种测试序列的实验结果均表明,本文算法在相同码率下的 PSNR 值明显优于其他两种算法,其编码程序在信噪比降低小于 0.05db, 比特率增加不超过 1% 的前提下,整体编码时间节约 10% ~ 20%。

(下转第 93 页)

表 3  $\bar{A}_{MRGP}$  vs.  $\bar{A}_{CTMC}$

alpha( $\alpha$ )	$\bar{A}_{MRGP}$	$\bar{A}_{CTMC}$	$\eta$
0.50000	0.000815	0.001086	33.25153%
0.60000	0.000574	0.000768	33.79790%
0.70000	0.000426	0.000571	34.03755%
0.80000	0.000328	0.000442	34.75609%
0.90000	0.000261	0.000352	34.86590%
1.00000	0.000212	0.000287	35.37735%
1.10000	0.000127	0.000172	35.43307%
1.20000	0.000110	0.000149	35.45454%

**结束语** 本文对面向用户的普适计算系统可用性进行了定量分析。首先针对普适计算模式下用户行为的特点建立了双扩展用户模型,从用户请求层次分析了普适计算系统的可用性。基于此,建立了普适计算系统可用性的马尔可夫重生过程模型,模型取消了用户态内系统态不变的约束(即重生点间状态不变的约束),并且模型允许使用任意分布来刻画用户行为,因此该模型可以更加准确地描述用户请求可用性的变化。基于随机过程的方法对普适计算系统可用性进行了分析,给出了对应瞬态和稳态行为的度量方法,并以此为基础定义了普适计算系统的请求可用性度量。最后给出了实验分

析,对影响相关因素进行了讨论,结果表明该度量模型和方法能够较为恰当地评估普适计算系统可用性,为进一步的研究提供参考。

**参考文献**

- 1 Xie W, Sun H, Cao Y, Trivedi K S. Modeling of user perceived webserver availability. ICC2003, Anchorage, Alaska, May 2003, 1796~1800
- 2 Wang Dazhi Trivedi K S. Modeling user-perceived service availability. In: Intl. Service Availability Symposium (ISAS'05), 2005
- 3 Govindan R, Reddy A. An analysis of internet inter-domain topology and route stability. In: Proc. of the INFOCOM1997, 1997, 850~857
- 4 Kalyanakrishnan M, Iyer R K, Patel J U. Reliability of internet hosts: A case study from the end user's perspective. Computer Networks, 1999, 31:45~57
- 5 Kaaniche M, Kanoun K, Martinello. M. A user-perceived availability evaluation of a Web based travel agency. In: Proc. of the 2003 Intl. Conf. on Dependable Systems and Networks (DSN'03), pages 709-718, San Francisco, California, June 2003. 709~718
- 6 李春江, 李东升, 肖依, 杨学军. 计算网络应用可用性的度量模型. 计算机研究与发展, 2003, 40(12): 1705~1709
- 7 Choi H, Kulkarni V G, Trivedi K S. Markov regenerative stochastic petri net. Performance Evaluation, 1994, 20: 335~357
- 8 Deng S. Empirical model of WWW document arrivals at access link. ICC1996, 1996. 1797~1802

(上接第 51 页)

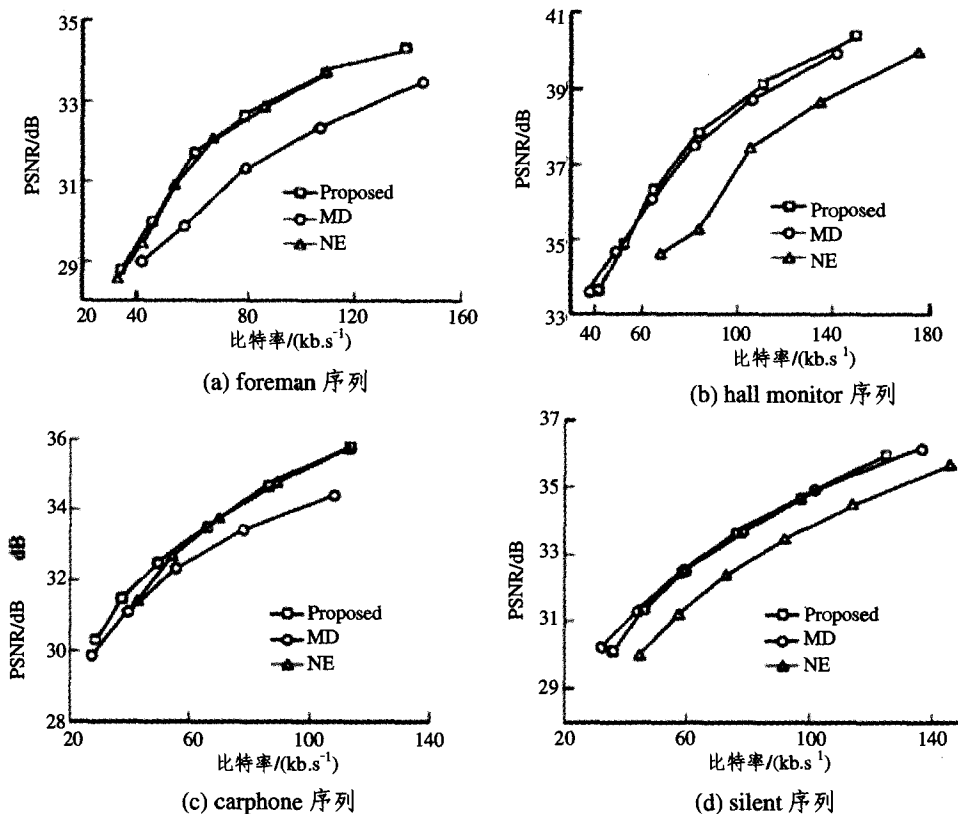


图 4 无线信道比特误码率在  $10^{-5} \sim 10^{-3}$  之间的相关率失真算法性能比较

**参考文献**

- 1 Wiegand T. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H. 264 | ISO/IEC 14496-10 AVC), JVT-G050, March 2003
- 2 Wenger S. H. 264 Complexity Analysis according to VCEG-L36 section 2.1.4, document VCEG-M23, ITU-T Video Coding Experts Group (VCEG) Meeting, 2-4 April, 2001
- 3 Schäfer R, Wiegand T. the Emerging H. 264/AVC Standard. EBUT echnical Review, Jan. 2003
- 4 傅祖芸著. 信息论——基础理论与应用. 电子工业出版社, 2001. 9

- 5 张剑. H. 264 标准的量化 16 位算法. 计算机科学, 2004, 11
- 6 Editor's Proposed Draft Text Modifications for Joint Video Specification (ITU-T Rec. H. 264 ISO/IEC 14496-10 AVC), Draft 7 5th Meeting: Geneva, Switzerland, 9-17 October, 2002
- 7 Zhang jian. The Computation Model of Code Error Distortion Based on the Rate-Distortion. Theory, international conference on embedded software and systems(ICESS), 2005, 12
- 8 Zhang jian. Research on the Computation Model of Macro block Coding Based on the Rate-Distortion Optimization, International Symposium on Communications and Information Technologies 2005 (ISCIT 2005), 2005. 10