

# 不依赖于量化参数的视频编码控制优化技术

袁 武<sup>1)</sup> 牛振东<sup>1)</sup> 林守勋<sup>2)</sup> 林子涵<sup>1)</sup> 谢 辉<sup>3)</sup> 罗海勇<sup>2)</sup>

<sup>1)</sup>(北京理工大学计算机学院 北京 100081)

<sup>2)</sup>(中国科学院计算技术研究所 北京 100190)

<sup>3)</sup>(中国石油长城钻探工程有限公司 北京 100101)

**摘 要** 目前,混合式视频编码技术普遍采用基于量化参数的 Lagrange 乘子优化算法(QP-L)实现预测编码的率失真优化计算.文中则提出一种新的率失真优化算法 Adapt-L,可以根据视频信号的率失真统计特性,自适应地调整 Lagrange 乘子的计算.实验及数据分析表明,在 QP-L 算法的 Lagrange 乘子和视频信号的率失真统计信息之间存在线性关系;由此,提出一个启发式的 Lagrange 乘子计算模型,并对该数学模型的合理性进行理论验证.大量实验证明,Adapt-L 的编码效率优于基于量化参数的优化算法(QP-L).虽然 Adapt-L 相对稍为复杂,但其计算不依赖于量化参数,有利于实现高效的码率控制.

**关键词** 率失真优化;视频编码控制;Lagrange 优化技术

中图法分类号 TP391 DOI号: 10.3724/SP.J.1016.2011.01519

## An Adaptable QP-Independent Approach to Rate-Distortion Optimization for Video Coding

YUAN Wu<sup>1)</sup> NIU Zhen-Dong<sup>1)</sup> LIN Shou-Xun<sup>2)</sup> LIN Zhi-Han<sup>1)</sup> XIE Hui<sup>3)</sup> LUO Hai-Yong<sup>2)</sup>

<sup>1)</sup>(School of Computer Science & Technology, Beijing Institute of Technology, Beijing 100081)

<sup>2)</sup>(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

<sup>3)</sup>(Greatwall Drilling Company Limited, China National Petroleum Corporation, Beijing 100101)

**Abstract** Currently a QP-based Lagrange Multiplier (QP-L) method is widely used to achieve Rate-Distortion optimal joint mode selection and motion estimation in hybrid video coding. In this work, a novel cost function for RDO is presented, which is independent of QP and adaptable to the statistics. Firstly, it is shown there be a linear relationship between  $\lambda$  of the QP-L method and some functionality about statistic for distortion and rate. Thusly, a heuristic model is presented to determine  $\lambda$  with consideration of statistic for distortion and rate, which is after then theoretically verified. It is shown via extensive experiments that the proposed scheme achieves a better coding performance than the traditional QP-based scheme. Moreover, the proposed scheme promises an easy encoding process, due to the fact that it's computation does not involve QP.

**Keywords** rate distortion optimization; hybrid coder control; Lagrange optimization

收稿日期:2009-10-09;最终修改稿收到日期:2011-07-11. 本课题得到国家自然科学基金(60902059,41071252)、国家“八六三”高技术研究发展计划项目基金(2007AA12Z239)和北京理工大学 2008 年优秀青年教师资助计划等项目资助. 袁 武,男,1974 年生,博士,主要研究方向为视频编码、多媒体检索及空间信息检索. E-mail: yuanwu@bit.edu.cn. 牛振东,男,1968 年生,博士,教授,主要研究领域为多媒体信息资源管理、数字图书馆、数字化教育、复杂软件体系结构等. 林守勋,男,1948 年生,博士,研究员,主要研究领域为视频编码、视频内容分析和检索、统计机器翻译. 林子涵,男,1988 年生,硕士研究生,主要研究方向为计算机软件与理论技术. 谢 辉,男,1977 年生,博士,主要研究方向为计算机应用技术. 罗海勇,男,1967 年生,博士,副研究员,主要研究方向为视频处理、普适计算、嵌入式系统.

## 1 引 言

目前,常见的视频压缩系统基本上都采用了有损压缩的混合式(Hybrid)编码架构.在该架构下,视频编码器可以针对各种具有复杂统计特征的视频信号,使用不同的编码工具(及组合)消除各种冗余信息,从而压缩视频数据.其基本处理过程包括预测编码、预测残差 DCT 系数量化及熵编码等环节.在预测编码环节,通过模式选择和预测补偿,消除视频信号的时-空域相关性;在码率控制环节,通过 DCT 系数量化控制预测残差的码率及失真.视频编码控制技术就是研究相关参数(包括编码模式  $MD$ 、运动矢量  $MV$  和量化参数  $QP$ )的计算问题,以期用尽可能少的比特获得高质量的重建图像.

据信息论,视频压缩的码率和失真满足下凸非递减关系.一般地,码率越低则失真越大,反之亦然,码率极低和失真极小二者不可兼得<sup>[1]</sup>.因此,只有在失真与码率之间实现恰当的折衷,才能实现编码控制的优化计算,获得较高的编码效率.考虑到信道带宽对多媒体通信的客观限制,该问题实质上是一个有约束极值求解,其数学表达形式如下:

$$\begin{aligned} \min_{MD \in MD, MV \in MV, QP \in QP} & D(MD, MV, QP) \\ \text{s. t. } & R(MD, MV, QP) \leq R_c \end{aligned} \quad (1)$$

其中,  $MD$ 、 $MV$  和  $QP$  分别表示编码模式、运动矢量和量化参数的取值范围,  $D(MD, MV, QP)$  表示图像失真,  $R(MD, MV, QP)$  表示码率,而  $R_c$  表示信道带宽.

对凸函数有约束极值问题的求解,采用 Lagrange 乘子法,可以转换为对等价的无约束极值问题进行求解,如式(2)所示:

$$\begin{aligned} \min_{MD \in MD, MV \in MV, QP \in QP} & \{D(MD, MV, QP) + \\ & \lambda(R(MD, MV, QP) - R_c)\} \end{aligned} \quad (2)$$

其中,  $\lambda$  表示 Lagrange 乘子.根据式(2)的最小值选择编码工具(及组合),在码率和失真之间实现折衷,很显然,该式能否实现高效的编码效率,关键在于  $\lambda$  值的选取.围绕上述问题,学术界开展了大量的研究工作.

文献[2]指出,在  $R_c$  已知的条件下,可以根据码率-失真的单调性,利用二分法通过多次编码逐渐逼近  $\lambda$  的最佳值,缺点是计算复杂度高,编码延迟大.

文献[3-4]根据率失真理论建立了  $\lambda$  关于量化参数的数学模型  $\lambda^{QP} = cQ^2$  (QP-L 模型),其中  $Q$  为  $QP$  对应的量化步长,  $c$  为某个常数.在计算编码模

式和运动矢量时,可根据代价函数  $J = D + \lambda^{QP}R$  的最小值来选择. QP-L 算法有效地提高了视频压缩的编码效率,在 H. 26L 及 H. 264/AVC 等标准中得到应用<sup>[5-6]</sup>.但是,由于视频信号的率失真统计特征关联诸元在 QP-L 模型中的缺位,其计算结果不可能对任意视频序列都为最优<sup>[7]</sup>.此外, QP-L 模型对量化参数存在依赖性,不利于实现高效的视频编码控制:一般地,量化参数主要用于控制预测残差的码率及失真,其值在预测编码后才能有效计算,而 QP-L 在预测编码计算时就需要预先知道量化参数,二者内在相互抵触,属于“鸡与蛋”逻辑悖论<sup>[8-9]</sup>.

文献[10]对 QP-L 算法进行改进,在运动补偿环节,根据上层或相邻图像块的率失真统计信息修正代价函数.该算法对编码效率的提高较为有限,仅在低复杂度模式启用 ( $RDOptimization = 0$ , 也即关闭率失真优化模式)的情况下较为明显,而对于高复杂度模式基本无用.

文献[11]提出一种基于 Laplace 分布假设的率失真优化模型,数学表达形式较为复杂.该算法根据量化参数及视频信号的率失真统计特征,以图像帧为单位,计算代价函数的 Lagrange 乘子  $\lambda$  值,对不符合 Laplace 分布假设的情况,则修正为 QP-L 模型.该模型能够有效地改进图像运动平缓的视频序列的编码效率;对于图像运动剧烈的视频序列,编码效率和 QP-L 模型相近.

本文目标在于建立一个不依赖于量化参数,能够根据视频信号的率失真统计特征自适应调整的率失真优化模型.本文作者在大量实验的基础上,发现  $\lambda^{QP}$  与视频信号率失真统计特征诸元的一定数学形式满足线性关系,由此建立  $\lambda$  关于视频信号的率失真统计特征的函数模型 (Adapt-L); 并进一步提出一种新的率失真优化算法,可以在图像宏块级,根据视频压缩的实际情况自适应地调整模型参数.大量的实验表明, Adapt-L 算法的编码性能优于 QP-L 算法.虽然 Adapt-L 相对 QP-L 稍为复杂,但其计算过程不引入量化参数,避免了“鸡与蛋”悖论问题,有利于实现高效的视频编码控制.

本文在第 2 节介绍经典的 QP-L 算法;在第 3 节,通过实验与数据分析,表明  $\lambda^{QP}$  与视频信号的率失真统计特征符合线性关系;在第 4 节,建立  $\lambda$  关于视频信号的率失真统计特征的函数模型,并在理论上验证该模型的合理性;在第 5 节,介绍 QP-L 算法的具体实现,在第 6 节介绍对比测试实验和结果分析;最后是总结.

## 2 QP-L 算法

一般地, 视频信号经过预测编码以及 DCT 变换之后, 预测残差 DCT 分量之间的相关性很小(可以认为它们“近似”不相关), 且满足 Laplace 分布规律. 根据文献[1], 如果使用重建图像与原始图像像素值之间的平均绝对值误差(MAD)表示图像失真, 也即是  $D(x, \bar{x}) = |x - \bar{x}|$ , 那么在预测残差的码率与失真之间存在如下函数关系:

$$R = \ln\left(\frac{\sigma}{D}\right) \quad (3)$$

不妨假设码率-失真是可微函数关系, 把式(3)代入到式(2)中, 对式(2)求关于  $R$  的导数, 并置为 0. 经过推导, 可得下式:

$$\lambda = -\frac{\partial D}{\partial R} \quad (4)$$

假设图像失真符合均匀分布, 可以导出失真关于量化步长  $Q$  的函数关系, 如下所示:

$$D = \frac{Q^2}{12} \quad (5)$$

把式(5)代入到式(4)中, 可得 QP-L 模型:

$$\lambda_{\text{mode}}^{\text{QP}} = cQ^2 \quad (6)$$

其中,  $c$  为一个常数, 可用实验方法确定其值(一般用 0.85 的经验值).

在 QP-L 算法中, 根据式(7)最小值来选择编码模式; 对于运动补偿, 则按照式(8)最小值来选择运动矢量.

$$J(s, c, \text{mode} | \text{QP}, \lambda_{\text{mode}}) = \text{SSD}(s, c, \text{mode} | \text{QP}) + \lambda_{\text{mode}}^{\text{QP}} \cdot R(s, c, \text{mode} | \text{QP}) \quad (7)$$

$$J(\mathbf{m}, \lambda_{\text{motion}}) = \text{SA}(T)D(s, c(\mathbf{m})) + \lambda_{\text{motion}}^{\text{QP}} \cdot R(\mathbf{m} - \mathbf{p}) \quad (8)$$

其中,  $\text{SSD}(s, c, \text{mode} | \text{QP})$  是原始像素和重建像素之间的误差平方和,  $\text{SA}(T)D(s, c(\mathbf{m}))$  是预测残差分量的绝对值和. 在式(8)中,  $\lambda_{\text{motion}}^{\text{QP}} = \sqrt{\lambda_{\text{mode}}^{\text{QP}}}$ .

我们知道, 视频压缩码流包括预测残差的编码信息和宏块编码副信息两部分. 对于低码率应用场景, 宏块编码副信息所占码流比例可达 20% ~ 30%, 对视频压缩编码性能的作用不可忽视<sup>[7]</sup>. QP-L 算法在副信息和预测残差之间优化比特分配, 显著地提高了视频编码性能; 加之计算形式也比较简单, 在 H. 263 和 H. 264/AVC 中得到了广泛的应用. 但是, 也正由于模型的过于简单, QP-L 算法存在以下缺陷:

(1) 视频编码控制过程复杂化.

在 QP-L 模型中, QP 是唯一一参变量, 该模型的编码性能取决于量化参数的合理取值. 在 Hybrid 编码结构中, 量化参数主要用于控制预测残差的码率及失真的大小. 为了有效地控制视频压缩码流大小, 需要根据预测残差的率失真统计信息计算量化参数, 但是, 相关统计信息无法先于预测编码得到. 为此, 可以通过多次编码处理逐渐逼近, 或者采用预测技术估计. 对于前者, 计算复杂度非常高, 不适用于实时应用场景; 而对于后者, 预测的准确性是关键因素. 而视频信号属于非平稳信号, 统计特征非常复杂, 难于用简单方式有效地预测.

(2) 不具备自适应性.

视频信号统计特征非常复杂. 在不同的视频序列之间, 统计特征差异较大(在第 7 节可以看到, 不同的视频序列具有不同的码率-失真曲线). QP-L 模型缺乏表征视频信号统计特性的关联量, 只要量化参数相等, 即使信号特征迥异, QP-L 算法对所有的视频序列都使用相同值的  $\lambda$ , 其优化结果不可能对任何视频序列都是最优的<sup>[7,10-11]</sup>, 在编码效率上存在提升空间.

## 3 $\lambda_{\text{mode}}^{\text{QP}}$ 与率失真统计信息之间的函数关系

在 QP-L 模型中,  $\lambda_{\text{mode}}^{\text{QP}}$  的计算只与量化参数相关, 而据文献[4],  $\lambda$  在物理意义上对应着码率-失真曲线的负斜率. 因此,  $\lambda$  的最佳取值应取决于视频信号的率失真特性. 如果把式(3)直接代入式(4), 可以发现  $\lambda$  又是关于失真的函数关系. 由此, 有足够的理由相信,  $\lambda_{\text{mode}}^{\text{QP}}$  与视频信号统计特征之间一定存在某种形式的函数关系.

顾名思义, 率失真优化是在失真和码率两个不同的量纲上进行折衷的技术. 我们知道, 失真和码率是无法进行直接比较的, 但是如果把它们转化为相同量纲, 不是就可以比较了吗? 从考虑当前宏块对视频序列总体率失真性能的综合影响最小的角度出发, 设计了第 1 个实验.

**实验 1.** 假设存在预测编码的其它判断准则, 根据式(9)的最小值选择当前宏块的编码模式:

$$J(\text{mode}) = \text{SSD}^n(\text{mode}) + \lambda^{(1)} \frac{\text{SSD}_{n-1}^{\text{total}}}{R_{n-1}^{\text{total}}} R^n(\text{mode}) \quad (9)$$

其中,  $n$  表示在当前帧内当前宏块序号,  $\lambda^{(1)}$  表示该

判断准则的 Lagrange 乘子,  $SSD_{n-1}^{total} = \sum_1^{n-1} SSD_i$  表示

当前总失真,  $R_{n-1}^{total} = \sum_1^{n-1} R_i$  表示当前总码率.

实验以 QP-L 算法优化结果  $\{SSD_n^{QP-RDO}, R_n^{QP-RDO}\}$  作为测试基准, 如果新判断准则要取得和 QP-L 算法相近的编码性能, 那么对任意优化结果  $\{D_n, R_n\}$ ,  $\lambda^{(1)}$  应该满足以下条件:

$$SSD_n^{QP-RDO} + \lambda^{(1)} \frac{SSD_{n-1}^{total}}{R_{n-1}^{total}} R_n^{QP-RDO} \leq SSD_n + \lambda^{(1)} \frac{SSD_{n-1}^{total}}{R_{n-1}^{total}} R_n \quad (10)$$

等价于下式:

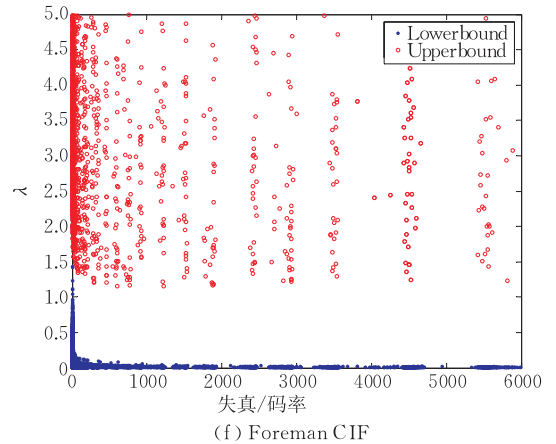
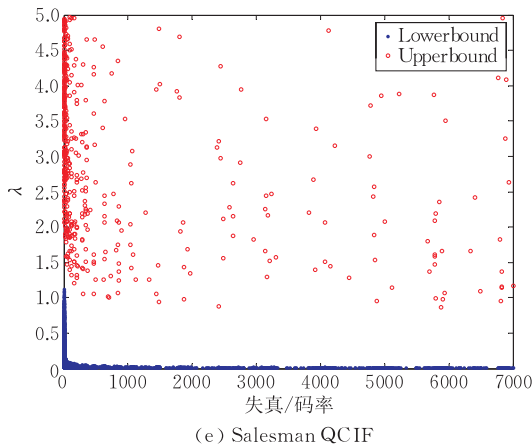
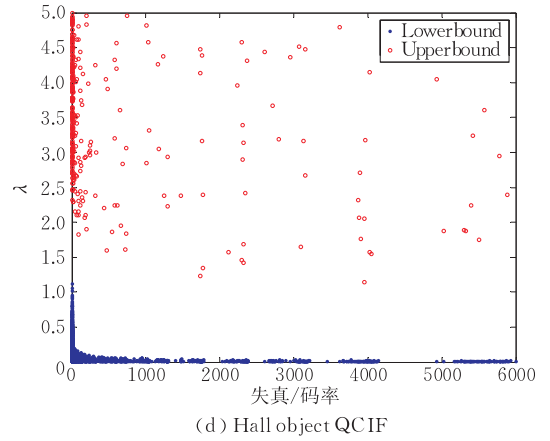
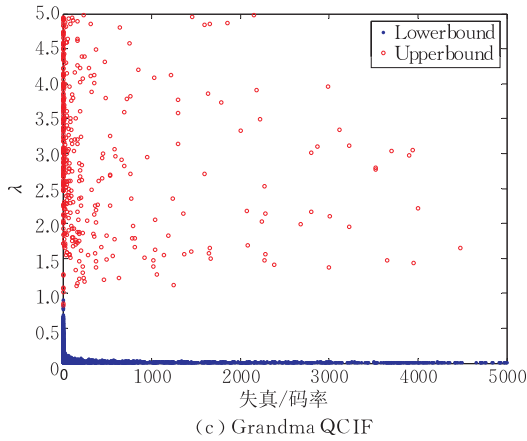
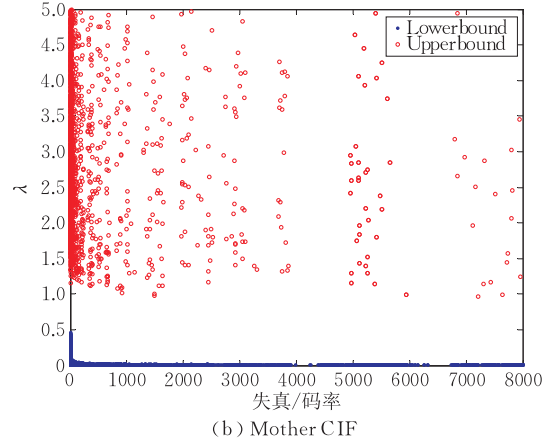
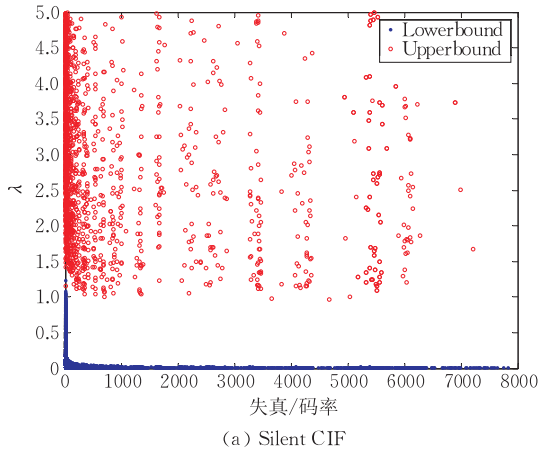


图 1  $\lambda^{(1)}$  可行值分布

$$\begin{cases} \lambda^{(1)} \leq \frac{SSD_n - SSD_n^{QP-RDO}}{SSD_{n-1}^{total}} \frac{R_{n-1}^{total}}{R_n^{QP-RDO} - R_n}, & R_n^{QP-RDO} > R_n \\ \lambda^{(1)} \geq \frac{SSD_n - SSD_n^{QP-RDO}}{SSD_{n-1}^{total}} \frac{R_{n-1}^{total}}{R_n^{QP-RDO} - R_n}, & R_n^{QP-RDO} < R_n \end{cases} \quad (11)$$

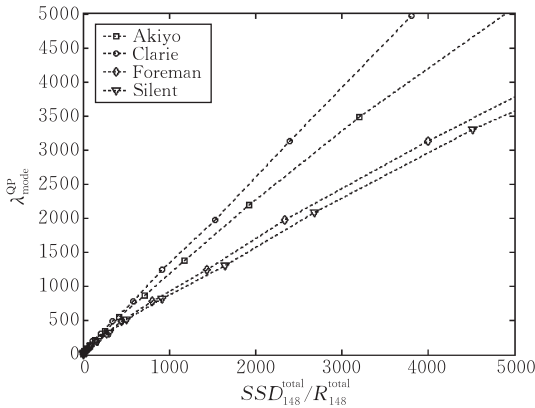
为了确定  $\lambda^{(1)}$  取值范围, 对大量视频序列进行 Intra 模式编码处理, 所用的量化参数分别为 8, 10, 12,  $\dots$ , 50. 针对每个视频序列, 分别统计取值范围的上确界 (Upperbound) 以及下确界 (Lowerbound). 实验数据表明, 在 Lowerbound、Upperbound 之间存在较为明显的分界区域, 如图 1 所示. 注意: 实验数

据绝大多数都满足  $\text{Lowerbound} < \text{Upperbound}$ , 但是, 也会出现少数  $\text{Upperbound} < \text{Lowerbound}$  的情况. 后者意味着无法找到一个  $\lambda^{(1)}$  能够满足式(11), 也就是说其计算结果不可能与 QP-L 算法相同, 但这种情况出现的概率很低, 对分析结论的影响可以忽略.

不妨假设可以用一个函数  $f$  表示  $\text{Lowerbound}$ 、 $\text{Upperbound}$  的分界(根据实验结果, 实际上可以用线性函数甚或一个常数作为  $f$  的表达式). 由此表明  $\lambda_{\text{mode}}^{\text{QP}}$  与视频信号统计特性之间可能存在一定形式的简单数学关系.

上述结论并不适用于 Inter 编码模式. Inter 编码模式拥有更多、更复杂的编码选项, 不同选项之间的码率-失真特性差异较大, 而优化结果仅仅反映单个选项的编码情况. 因此, 无法简单地利用个体的率失真信息  $\{SSD^{\text{lagrange}}, R^{\text{lagrange}}\}$  有效地表征视频信号的统计特征.

**实验 2.** 在图像宏块编码时, 针对编码模式及



运动估计的所有可能选项及其组合, 计算图像失真及

码率均值:  $SSD_n^{\text{average}} = \frac{\sum_{MD, MV} SSD_{MD, MV}}{\|MD, MV\|}$  和  $R_n^{\text{average}} =$

$\frac{\sum_{MD, MV} R_{MD, MV}}{\|MD, MV\|}$ , 其中  $\|MD, MV\|$  表示编码模式和运动估计不同组合的基数. 实验的目的在于针对不同的

量化参数, 分析  $\lambda_{\text{mode}}^{\text{QP}}$  与  $f(SSD_n^{\text{total}}, R_n^{\text{total}})$  之间的函数

关系, 其中  $SSD_n^{\text{total}} = \sum_1^n SSD_i^{\text{average}}$ ,  $R_n^{\text{total}} = \sum_1^n R_i^{\text{average}}$ ,

$f$  表示关于  $SSD_n^{\text{total}}$  和  $R_n^{\text{total}}$  的二元函数关系.  $f$  可能

存在多种不同的数学形式, 经过多次实验, 发现  $\lambda_{\text{mode}}^{\text{QP}}$  与  $f = \frac{SSD_n^{\text{total}}}{R_n^{\text{total}}}$  具有线性关系.

图 2 显示了部分实验数据, 这些数据都是关于同一个宏块(位于第一个  $P$  帧第 5 行 5 列)的. 可以看到

在  $\lambda_{\text{mode}}^{\text{QP}}$  与  $\frac{SSD_n^{\text{total}}}{R_n^{\text{total}}}$  之间存在较为明显的线性关系.

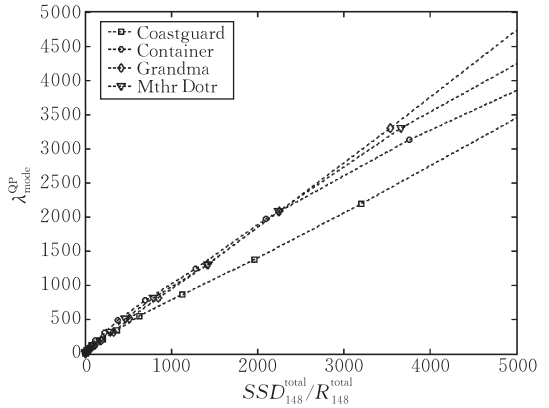


图 2  $\lambda_{\text{mode}}^{\text{QP}}$  与  $\frac{SSD_n^{\text{total}}}{R_n^{\text{total}}}$  的函数关系

实验 1 和实验 2 都表明  $\lambda_{\text{mode}}^{\text{QP}}$  与视频信号统计特性之间存在一定形式的简单数学关系. 由此, 我们可以得到  $\lambda$  关于视频信号统计特性的一个启发性函数模型:

$$\lambda^{(2)} = \frac{aD}{R+b} + c \quad (12)$$

下面, 我们利用率失真理论验证上述公式的合理性.

## 4 理论验证

对式(3)采用泰勒级数展开, 舍弃高阶项, 可以得到 MPEG-4 Q2 码率模型<sup>[14]</sup>:

$$R = X_2 \frac{\sigma^2}{SSD} + X_1 \frac{\sigma}{\sqrt{SSD}} \quad (13)$$

其中,  $X_2$  和  $X_1$  是模型参数, 其值可以用线性回归技术加以确定.

把式(13)代入式(2), 则视频编码控制的率失真优化问题可以如下表述:

$$\min \left\{ SSD + \lambda \left( X_2 \frac{\sigma^2}{SSD} + X_1 \frac{\sigma}{\sqrt{SSD}} \right) \right\} \quad (14)$$

利用 Lagrange 乘子技术, 经过数学推导, 可以得到

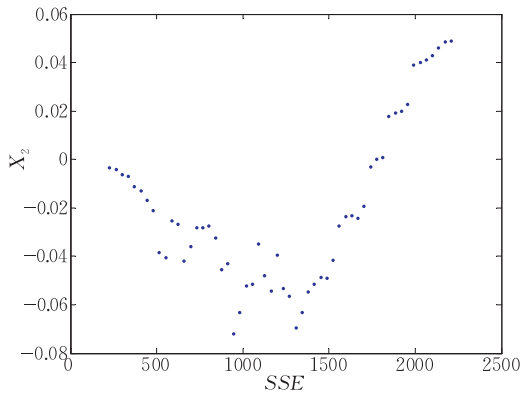
$$\lambda = \frac{2SSD}{2X_2 \frac{\sigma^2}{SSD} + X_1 \frac{\sigma}{\sqrt{SSD}}} = \frac{2SSD}{R + X_2 \frac{\sigma^2}{SSD}} \quad (15)$$

式(15)与式(12)在数学表达形式具有一致性. 在式(15)中,  $X_2$  是唯一的未定量. 下面, 我们将通过实验和数据分析建立  $X_2$  关于预测残差分量平方和 (Sum of Squares Error/SSE) 的数学模型.

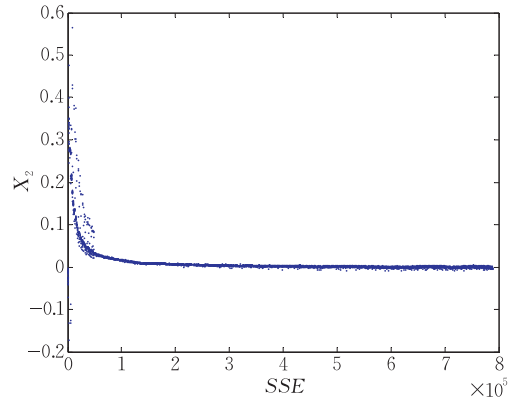
## 5 $X_2$ 数学建模

文献[11]指出,可以使用线性回归技术计算 MPEG-4 Q2 码率模型的参数  $X_2$  和  $X_1$ . 我们在实验中发现,模型参数( $X_2$  和  $X_1$ )与预测残差  $SSE$  之间存在一定的相关性. 为了对这种关系进行量化分析,设计如下的实验:

在宏块编码时,记录所有的关于图像失真、预测



(a)  $SSE < 2500$



(b)  $SSE > 2500$

图 3  $X_2$  关于预测残差  $SSE$  的函数关系

$$X_2 = \begin{cases} \frac{3.94E+09}{SSE^3} - \frac{3.75E+06}{SSE^2} + \frac{1900}{SSE} - 0.00313, & SSE > 3038 \\ 1E-07 \times SSE^2 - 0.0002 \times SSE + 0.041, & \text{其它.} \end{cases} \quad (16)$$

$$\lambda = \begin{cases} \frac{2SSD}{R + \left( \frac{3.94E+09}{SSE^3} - \frac{3.75E+06}{SSE^2} + \frac{1900}{SSE} - 0.00313 \right) \times \frac{SSE}{SSD}}, & SSE > 3038 \\ \frac{2SSD}{R + (1E-07 \times SSE^2 - 0.0002 \times SSE + 0.041) \times \frac{SSE}{SSD}}, & \text{其它} \end{cases} \quad (17)$$

优化算法具体实现如下所述:

首先,分配缓冲区,用于保存当前帧以及前一帧每个宏块的平均失真、预测残差平均码率及平均  $SSE$ .

针对第一个  $I$  帧,采用 QP-L 方法进行优化计算,记录每次编码处理的失真、预测残差码率和  $SSE$ ,并把相关均值保存在缓冲区中;

对于  $P$  帧宏块,在当前帧及前一帧搜索与当前宏块相邻( $|\Delta x + \Delta y| \leq \text{Radius\_Threshold}$ )的已编码宏块,计算各统计信息均值;把所得到的平均失真、平均码率以及平均  $SSE$  代入到式(17)中,计算出  $\lambda$ ;为了避免  $\lambda$  变化太大,可以根据历史平均值  $\lambda_{\text{average}}^{\text{frame}}$  对  $\lambda$  进行限制. 获得的  $\lambda$  直接用于模式选择,

残差码率;然后根据预测残差  $SSE$  及量化参数,对这些数据进行分类统计,计算其平均值;然后针对不同的预测残差  $SSE$ ,利用线性回归技术计算  $X_2$ ;最后,建立  $X_2$  与预测残差  $SSE$  之间的函数关系.

按照上述步骤,进行了大规模的编码实验和数据分析. 最终得到  $X_2$  与预测残差  $SSE$  之间的函数关系,如图 3 所示.

采用分段拟合函数,得到  $X_2$  与预测残余方差的数学表达式:

## 6 率失真优化算法

根据以上理论及实验分析,可以获得预测编码控制代价函数乘子的一个新计算公式,如下所示:

对于运动补偿,采用  $\sqrt{\lambda}$ . 在优化计算时,需要记录每次编码处理的失真、预测残余码率和  $SSE$ ,并把相关均值保存在缓冲区中.

## 7 实验与结果分析

在 H. 264/AVC 参考软件 JM10.0<sup>[12]</sup> 的基础上,按照编码预测控制优化新算法的要求,改写了部分程序. 在算法的具体实现中,宏块搜索半径  $\text{Radius\_Threshold}$  设为 2;在每个图像帧编码结束时,更新历史平均值  $\lambda_{\text{average}}^{\text{frame}}$ ;在计算宏块  $\lambda$  时,按照如下策略对其进行修正:

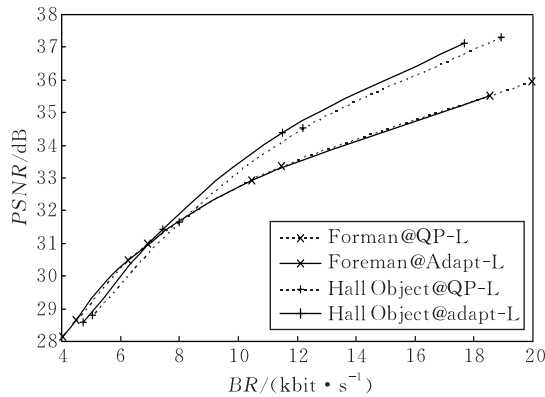
if  $(\lambda > \lambda_{\text{average}}^{\text{frame}} \times 1.25)$ ,  $\lambda = \lambda_{\text{average}}^{\text{frame}} \times 1.25$

else if  $(\lambda < \lambda_{\text{average}}^{\text{frame}} \times 0.75)$ ,  $\lambda = \lambda_{\text{average}}^{\text{frame}} \times 0.75$

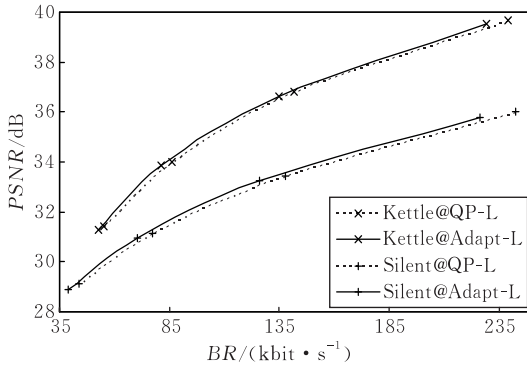
为了客观而全面地评价新优化方法的编码性能, 本文对 Adapt-L 方法和 QP-L 方法进行了大量的对比编码实验, 涵盖了常见的 QCIF 及 CIF 格式的视频测试序列. 限于文章篇幅, 本文只列出部分测试序列的实验结果. QCIF 格式的视频序列包括 Coastguard、Foreman、Grandma、Hall Objects、News 和 Salesman, 其中 Coastguard 帧率为 30 帧/s, Foreman 和 Salesman 帧率为 10 帧/s, Grandma 和 Hall Object 帧率为 15 帧/s, News 帧率为 3.33 帧/s. CIF 格式的视频序列包括 Kettle、Mother、Paris、News

和 Silent, 其中除 Paris 帧率为 7.5 帧/s, 其它均为 30 帧/s. 测试序列 GOP 结构均按照 IPPP 方式组织, 编码总帧数为 50. 视频编码实验采用固定的量化参数进行, 量化参数分别设置为 28, 32, 36 和 40. 其它的编码参数设置情况为: 运动估计为 1/4 像素精度, 参考帧数为 5, Hadamard 变换开关打开, 熵编码模式设为 CABAC, 运动估计搜索范围设为  $\pm 32$ , IntraPeriod 为 0, RestrictSearchRange 为 2.

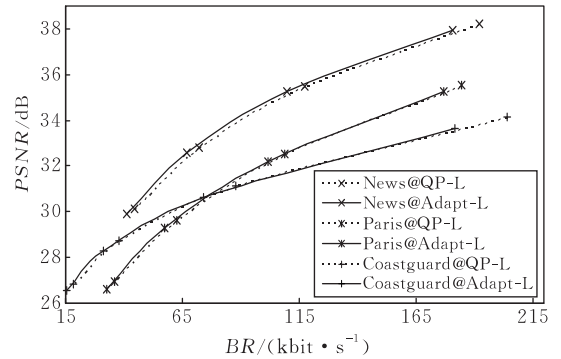
图 4 对 Adapt-L 方法和 QP-L 方法的编码效率进行了比较(相关实验数据列于表 1~2 中). 可见, Adapt-L 方法的编码效率优于 QP-L 方法, 为了客观地评价改善程度, 进一步采用 AVSNR<sup>[13]</sup> 软件计算



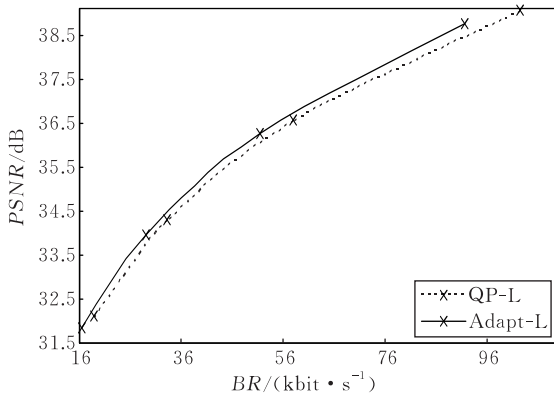
(a) Foreman QCIF, Hall Object QCIF



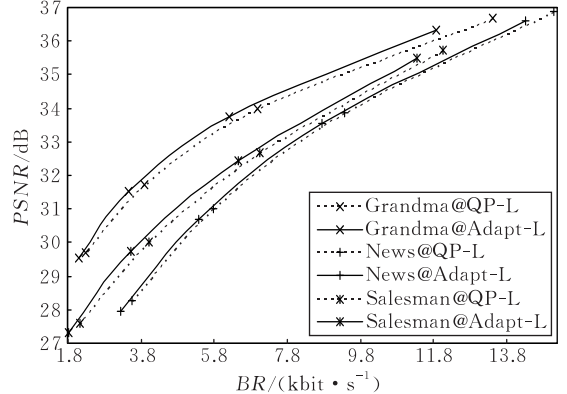
(b) Kettle CIF, Silent CIF



(c) News CIF, Paris CIF, Coastguard QCIF



(d) Mother CIF



(e) Grandma QCIF, News QCIF, Salesman QCIF

图 4 QP-L 和 Adapt-L 编码效率比较

编码效率增益. 对于 QCIF 格式的视频序列, 据计算, Coastguard 平均增加 0.030 dB, Foreman 平均增加 0.006 dB, Grandma 平均增加 0.218 dB, Hall Objects 平均增加 0.214 dB, News 平均增加 0.084 dB, Salesman 平均增加 0.208 dB; 对于 CIF 格式的视频序列, 据计算, Kettle 平均增加 0.118 dB, Mother 平均增加 0.201 dB, Paris 平均增加 0.053 dB, News 平均增加 0.116 dB, Silent 平均增加 0.157 dB.

表 1 QCIF 格式的视频序列实验数据详表

测试序列	优化方法	量化参数				
		28	32	36	40	
Coastguard	QP-L	Y PSNR	34.15	31.14	28.71	26.83
		码率	204.29	88.03	37.80	18.20
	Adapt-L	Y PSNR	33.65	30.62	28.29	26.57
		码率	182.06	74.31	31.34	15.61
Foreman	QP-L	Y PSNR	35.94	33.36	30.99	28.65
		码率	19.98	11.47	6.92	4.47
	Adapt-L	Y PSNR	35.53	32.94	30.48	28.14
		码率	18.54	10.47	6.24	4.02
Grandma	QP-L	Y PSNR	36.67	33.98	31.74	29.71
		码率	13.45	7.01	3.91	2.31
	Adapt-L	Y PSNR	36.34	33.73	31.52	29.55
		码率	11.88	6.23	3.47	2.10
Hall_Object	QP-L	Y PSNR	37.31	34.55	31.64	28.81
		码率	18.92	12.18	7.98	5.04
	Adapt-L	Y PSNR	37.12	34.37	31.43	28.59
		码率	17.68	11.49	7.45	4.72
News	QP-L	Y PSNR	36.88	33.87	31.01	28.26
		码率	15.10	9.37	5.78	3.55
	Adapt-L	Y PSNR	36.61	33.55	30.68	27.95
		码率	14.33	8.77	5.40	3.25
Salesman	QP-L	Y PSNR	35.73	32.66	30.02	27.61
		码率	12.06	7.07	4.02	2.15
	Adapt-L	Y PSNR	35.50	32.42	29.72	27.30
		码率	11.36	6.48	3.54	1.84

表 2 CIF 格式的视频序列实验数据详表

测试序列	优化方法	量化参数				
		28	32	36	40	
Kettle	QP-L	Y PSNR	39.66	36.81	34.02	31.40
		码率	239.07	141.77	86.17	55.24
	Adapt-L	Y PSNR	39.51	36.62	33.85	31.30
		码率	229.46	134.98	81.16	52.94
Mother	QP-L	Y PSNR	39.06	36.57	34.30	32.10
		码率	102.62	57.93	33.22	18.77
	Adapt-L	Y PSNR	38.74	36.27	33.97	31.86
		码率	91.71	51.45	28.99	16.53
Paris	QP-L	Y PSNR	35.54	32.52	29.63	26.97
		码率	184.87	108.83	62.51	36.18
	Adapt-L	Y PSNR	35.25	32.18	29.30	26.61
		码率	176.91	101.53	57.58	32.69
News	QP-L	Y PSNR	38.22	35.50	32.80	30.14
		码率	192.12	117.48	72.04	44.36
	Adapt-L	Y PSNR	37.92	35.24	32.56	29.9
		码率	180.82	110.05	67.05	41.24
Silent	QP-L	Y PSNR	36.00	33.42	31.14	29.13
		码率	242.67	138.00	77.40	44.04
	Adapt-L	Y PSNR	35.80	33.23	30.93	28.88
		码率	226.36	126.16	70.57	38.86

## 8 结束语

本文另辟蹊径, 利用率失真理论建立  $\lambda$  关于视频信号的率失真统计特征的函数关系, 提出了一种新颖的预测编码控制优化方法, 能够根据视频信号的率失真统计特征, 自适应地调整模型参数. 经过大量的实验证明, Adapt-L 方法的编码效率优于 QP-L 方法. 虽然 Adapt-L 方法的计算相对复杂, 由于它不需要引入量化参数, 反而为实现高效的、精确的码率控制创造了可能性.

## 参 考 文 献

- [1] Viterbi A, Omura J. Principles of Digital Communication and Coding. New York: McGraw-Hill Electrical Engineering Series, 1979
- [2] Ramchandran K, Vetterli M. Best wavelet packet bases in a rate-distortion sense. IEEE Transactions on Image Processing, 1993, 2(2): 160-175
- [3] Sullivan G J, Wiegand T. Rate-distortion optimization for video compression. IEEE Signal Processing Magazine, 1998, 15(6): 74-90
- [4] Wiegand T, Girod B. Lagrange multiplier selection in hybrid video coder control//Proceedings of the IEEE International Conference on Image Processing (ICIP'01). Thessaloniki, Greece, 2001, 3: 542-545
- [5] ITU-T/SG16/Q15-D-65. Video Codec Test Model, Near Term, Version 10 (TMN-10), Draft 1. Download via anonymous ftp to: standard.pictel.com/video-sitd9804-Tam/ql5d65.doc, Apr. 1998
- [6] Wiegand T, Sullivan G J. Draft ITU-T Recommendation H. 264 and Final Draft International Standard of Joint Video Specification (ITU-T Recommendation H. 264 | ISO/IEC 14496-10 AVC). Joint Video Team of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q. 6 Doc. JVT-G050, Pattaya, Thailand, Mar. 2003
- [7] Zhang Jun, Yi Xiaoquan, Ling Nam, Shang Weijia. Bit rate distribution analysis for motion estimation in H. 264//Proceedings of the IEEE International Conference on Consumer Electronics. Las Vegas, 2006: 483-484
- [8] Li Z, Pan F, Lim K P, Feng G, Lin X, Rahardja S. Adaptive basic unit layer rate control for JVT. JVT-G012-r1, 7th Meeting. Pattaya II, Thailand, 2003
- [9] Yuan Wu, Lin S X, Zhang Y D, Yuan Wen, Luo H Y. Optimum bit allocation and rate control for H. 264/AVC. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 705-715
- [10] Zhang Jun, Yi Xiaoquan, Ling Nam, Shang Weijia. Context adaptive lagrange multiplier (CALM) for rate-distortion op-

timal motion estimation in video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(6): 820-828

- [11] Li X, Oertel N, Hutter A, Kaup A. Laplace distribution based Lagrangian rate distortion optimization for hybrid video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009, 19(2): 193-205
- [12] JM10.0. [http://iphome.hhi.de/suehring/tml/download/old\\_jm/jm10.zip](http://iphome.hhi.de/suehring/tml/download/old_jm/jm10.zip)

- [13] Bjontegaard G. Calculation of average PSNR differences between RD-curves//*Proceedings of the VCEG-M33, ITU-T Video Coding Experts Group (VCEG) Meeting, Austin, TX, 2001*
- [14] Chiang T, Zhang Y-Q. A new rate control scheme using quadratic rate distortion model. *IEEE Transactions on Circuits and Systems for Video Technology*, 1997, 7(1): 246-250



**YUAN Wu**, born in 1974, Ph. D. . His research interests include video coding, grid computing and content management.

**NIU Zheng-Dong**, born in 1968, Ph. D. , professor. His research interests include multimedia information management, digital library, digital education and software architecture.

**LIN Shou-Xun**, born in 1948, Ph. D. , professor. His research interests include video coding, video content analysis and retrieval, statistic machine translation.

**LIN Zi-Han**, born in 1988, M. S. candidate. His research interests focus on computer application.

**XIE Hui**, born in 1977, Ph. D. . His research interests focus on network management.

**LUO Hai-Yong**, born in 1967, Ph. D. , associate professor. His research interests include covers pervasive computing, wireless localization, multimedia technology and embedded systems.

## Background

Currently, hybrid coding techniques predominate in the literature of video coding, wherein various modes of operations are stipulated in most standards for effective picture difference refinement. As for any lossy coding scheme, there is an inherent tradeoff between rate and distortion. Ideally, the source encoder should balance the quality of restored pictures and the channel capacity. That is to say, to minimize distortion subject to the channel bandwidth constraint  $R_c$ , by adaptively assigning the predefined operations to parts of the encoded pictures on a block-by-block basis primarily via control parameters such as coding mode (MD), motion vector (MV) and quantization parameter (QP).

This constrained problem can be formulated as  $\min_{MD, MV, QP} \sum_{n,k} D_{n,k}$  s. t.  $\sum_{n,k} R_{n,k} \leq R_c$ , and can be converted with the Lagrangian Multiplier Method into an unconstrained form:  $\min_{MD, MV, QP} \sum_{n,k} D_{n,k} + \lambda (\sum_{n,k} R_{n,k} - R_c)$ , where  $\lambda$  can be determined by R-D models. General speaking, the more accurate the R-D models are, the better  $\lambda$  can be achieved.

Currently a QP-based Lagrange Multiplier (QP-L) method is widely used to achieve Rate-Distortion optimal

joint mode selection and motion estimation in hybrid video coding. Due to the fact that  $\lambda = cQ^2$ , the QP-L method is only related to QP, leading to a well-known egg-hen dilemma for rate control. Furthermore, the mathematical formula is too simple and includes no index to video coding complexity, which means that all the video sequences are treated without any differentiation between various rate-distortion characteristics, and therefore, the QP-L method fails to adapt itself to the real world.

In this work, a novel cost function for RDO is presented, which is independent of QP and adaptable to the statistics. Firstly, it is shown there be a linear relationship between  $\lambda$  of the QP-L method and some functionality about statistic for distortion and rate. Thusly, a heuristic model is presented to determine  $\lambda$  with consideration of statistic for distortion and rate, which is after then theoretically verified. It is shown via extensive experiments that the proposed scheme achieves a better coding performance than the traditional QP-based scheme. Moreover, the proposed scheme promises an easy encoding process, due to the fact that its computation does not involve QP.