

基于结构相似的 H.264 主观率失真性能改进机制

崔子冠* 朱秀昌

(南京邮电大学图像处理与图像通信重点实验室 南京 210003)

摘 要: H.264 以客观失真作为失真准则进行码率控制(RC)和率失真优化(RDO)模式选择, 无法得到最优的主观质量。该文在作者之前研究成果的基础上将基于结构相似(SSIM)的主观失真用于指导 H.264 基于 RDO 的帧间模式选择, 进一步提出了宏块(MB)层自适应的分析型拉格朗日(Lagrange)乘子来更好地平衡码率和 SSIM 失真。实验结果表明: 在给定目标码率下, 该文算法相比基于客观质量的编码算法及基于 SSIM 的 RC 算法(但未进行基于 SSIM 的 RDO 帧间预测)更有效地编码了图像结构信息, 得到了更好的主观率失真性能和主观图像质量。

关键词: 视频编码; H.264; 率失真优化; Lagrange 乘子; 结构相似性

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2012)02-0433-07

DOI: 10.3724/SP.J.1146.2011.00586

Subjective Rate-distortion Performance Improvement Scheme for H.264 Based on SSIM

Cui Zi-guan Zhu Xiu-chang

(Image Processing and Image Communication Lab,

Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: H.264 takes objective metric as distortion criteria to perform Rate Control (RC) and Rate Distortion Optimization (RDO)-based mode decision, which can not acquire optimal subjective quality. Base on our previous research results, this paper applies Structural SIMilarity (SSIM) based subjective distortion to RDO-based inter mode decision in H.264 video coding, and further proposes an analytic MacroBlock (MB) layer Lagrange multiplier adjustment scheme to adaptively balance the tradeoff between rate and SSIM distortion better. Experimental results show that, at given target bit rate, the proposed method encodes image structural information more effectively, and thus acquires better subjective RDO performance and image quality compared with objective quality based encoding scheme and SSIM-based RC method (without performing SSIM-based RDO inter prediction).

Key words: Video coding; H.264; Rate Distortion Optimization (RDO); Lagrange multiplier; Structural SIMilarity (SSIM)

1 引言

H.264 采用率失真优化(RDO)技术选择最优的编码模式, 在码率与失真(R-D)之间取得了很好的平衡, 极大地提高了编码效率, 其中失真测量对 RDO 性能影响重大。均方误差(MSE)作为一种客观失真计算简单且物理意义明确, 因此传统码率控制(RC)和 RDO 都使用 MSE 作为失真测量^[1], 但 MSE 与主观质量并不相符, 基于 MSE 的 RC 和 RDO 无法得到最优的主观质量。另外求解 RDO 问题时引入的拉格朗日(Lagrange)乘子(λ)仅是量化参数(QP)

的函数^[2], 无法根据图像内容对码率与失真更好地权衡, 也在一定程度上降低了 RDO 性能。因此建立有效的主观质量评价方法并用于指导具体的编码过程, 以及建立内容自适应的 Lagrange 乘子来提高 RDO 性能成为近来视频编码领域的研究热点。

针对上述问题, 已有学者从不同方面展开研究。文献[3]根据人类视觉系统(HVS)对场景中结构信息更加敏感的特性提出了结构相似性(SSIM)测量, 计算出的 SSIM 值与图像主观质量非常一致, 而且计算量较小, 已被引入到 H.264 参考软件^[4]中以评价编码视频的主观质量。文献[5,6]把 SSIM 引入 H.264 基于 RDO 的帧内与帧间预测中, 用基于块的 SSIM 失真或平方误差和(SSE)与 SSIM 联合失真来代替传统的 SSE, 得到了更好的主观质量和更优的主观

2011-06-15 收到, 2011-09-01 改回

国家自然科学基金(60672134, 61071091)和江苏省研究生培养创新工程(CX10B_190Z)资助课题

*通信作者: 崔子冠 czg1982001@163.com

率失真性能。但其率失真代价函数中的 λ 是由实验得到的经验常数,不能根据图像内容自适应调整,其主观 RDO 性能仍有较大的提升空间。使用经验常数作为 λ 是由于很难得到分析型 R-SSIM 失真函数^[7],SSIM 失真与 QP 的关系也不易求得,这是基于 SSIM 的 RDO 相比基于客观质量(MSE)的 RDO 的一个难点问题。文献[8,9]将 SSIM 失真引入 RDO 框架后,发现基于 SSIM 的率失真曲线与基于 MSE 的率失真曲线相邻点处切线的坡度非常相似,使用之前关键帧的率失真模型和前一编码帧的 R-D 数据对来预测当前帧的 λ ,得到了帧层自适应的 λ 。但需要对关键帧用基于 MSE 的 RDO 编码两次来预测当前帧基于 SSIM 的 λ ,计算复杂度较高,得到的 λ 是经验型的没有解析式。文献[10,11]考虑到一帧中不同宏块(MB)也具有不同的率失真特性,对基于 MSE 的 RDO 中的 λ 在 MB 层根据编码复杂度进行自适应调整,有效提高了客观率失真性能,但没有考虑主观特性。文献[12]由实验统计提出了 R-SSIM 的经验模型,进行基于主观质量的基本单元(BU)层码率分配,但 SSIM 失真与 QP 的关系仍不明确。文献[13]提出了 P 帧 SSIM-量化步长(Q_{step})线性关系模型,并结合码率模型进行主观质量最优的 MB 层 RC,但没有将 SSIM 失真用于主观 RDO 中的模式选择,也没有进行基于 SSIM 失真的 λ 调整,因此主观率失真性能还可进一步提高。

本文针对上述问题,在之前研究成果^[13]的基础上,将 SSIM 失真引入 H.264 的 RDO 框架中进行基于主观质量的帧间预测;进一步提出了 MB 层内容自适应的分析型 λ 来更好地调节 R 与 SSIM 失真。实验结果表明,相比基于 MSE 的编码算法^[1]及基于 SSIM 的 RC 算法(但未进行基于 SSIM 的 RDO 帧间预测和 λ 调整)^[13],本文算法更好地编码了图像结构信息,对 R 与 SSIM 失真进行了更好地平衡,进一步提高了主观率失真性能。

2 P 帧 SSIM-量化步长的关系模型

2.1 SSIM 失真测量

文献[3]根据 HVS 对场景中结构信息更加敏感的特性提出了一种全参考的主观质量评价方法,对原始图像 x 和失真图像 y 分别从亮度 $l(x,y)$,对比度 $c(x,y)$ 和结构相似性 $s(x,y)$ 3 个方面进行比较,最后综合得出失真图像的 SSIM 值。计算公式如下:

$$\begin{aligned} \text{SSIM}(x,y) &= l(x,y) \cdot c(x,y) \cdot s(x,y) \\ &= \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \end{aligned} \quad (1)$$

其中 μ_x , μ_y , σ_x , σ_y , σ_{xy} 分别为 x 与 y 各自的亮度

均值、标准差和二者之间的协方差,而 C_1 , C_2 是为了避免当分母接近零时测量值不稳定而定义的小常数。SSIM 测量值具有边界,即 $\text{SSIM}(x,y) \leq 1$,SSIM 值越接近 1,则失真图像的主观质量越好。

2.2 P 帧 SSIM-量化步长线性失真模型

作者在文献[13]中发现 H.264 编码的 P 帧 SSIM 值与 Q_{step} 之间存在很明显的线性关系,提出的 P 帧 SSIM- Q_{step} 线性模型如下:

$$\text{SSIM} = 1 - k \times Q_{step} \quad (2)$$

其中斜率 $k > 0$,代表图像的 SSIM 失真特性。据此可定义 P 帧第 i 个 MB 的 SSIM 失真如下:

$$D_{\text{SSIM},i} = 1 - \text{SSIM}_i = k_i \times Q_{\text{step},i} \quad (3)$$

相邻帧相同位置 MB 具有强时域相关性,因而也具有相似的 SSIM 失真特性。图 1 显示了 Grandma 序列在目标码率为 64 kbps 开启 MB 层 RC 编码的第 5 帧与第 6 帧各 MB 的实际 k_i 值。可见帧内各 MB 的 k_i 值变化非常剧烈,但只要不发生场景变换,相邻帧对应位置 MB 的实际 k_i 值非常相似,且相邻帧各 MB 的 k_i 值的变化趋势非常相似。由此在计算当前 MB 基于 SSIM 的 λ 时,参数 k_i 可由前一帧相同位置 MB 的实际编码结果估计得到,即

$$k_i = k_i^p = \frac{1 - \text{SSIM}_i^p}{Q_{\text{step},i}^p} \quad (4)$$

其中 SSIM_i^p 和 $Q_{\text{step},i}^p$ 分别为前一帧相同位置 MB 实际编码的 SSIM 值和量化步长。

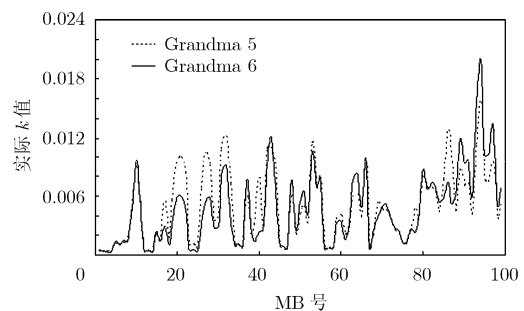


图 1 MB 层 SSIM 失真斜率(Grandma, 64 kbps)

3 基于 SSIM 的 RDO 帧间预测及 MB 层 λ 自适应调整

3.1 传统基于 SSE 的 RDO 模式选择和 λ

传统 RDO 通过引入 λ 选择具有最小客观率失真代价的模式作为最优的编码模式,定义的基于平方误差和(SSE)的代价函数如下:

$$\begin{aligned} J_{\text{SSE}}(s,c,\text{mode} | \text{QP}) &= D_{\text{SSE}}(s,c,\text{mode} | \text{QP}) \\ &\quad + \lambda_{\text{SSE}} \cdot R(s,c,\text{mode} | \text{QP}) \end{aligned} \quad (5)$$

其中 s 与 c 分别表示原始与重建图像块,mode 表示

MB 编码模式, QP 是当前 MB 的量化参数, $R(s, c, \text{mode}|QP)$ 是 MB 编码所需的码率。基于 SSE 的 λ_{SSE} 用于权衡 R 与 SSE 失真, 可由码率和失真模型由解析法求导得到, H.264 中基于 SSE 的码率和失真模型分别如下:

$$R = a \log_2(b/D_{\text{SSE}}) \quad (6)$$

$$D_{\text{SSE}} = Q_{\text{step}}^2/12 \quad (7)$$

其中 a 与 b 是模型参数, Q_{step} 是量化步长。为使代价函数最小, 对式(5)求一阶导数并令导数为零, 代入式(6)与式(7)后可得 λ_{SSE} 的解析式如下:

$$\lambda_{\text{SSE}} = -\frac{dD_{\text{SSE}}}{dR} = c \times Q_{\text{step}}^2 \quad (8)$$

其中 c 在 H.264 参考软件中是经验常数 ($0.85 \times 2^{-8/3}$)。

3.2 基于 SSIM 的 RDO 帧间预测

为了在码率限制下得到最优的感知质量, 本文考虑将 SSIM 失真引入 H.264 基于 RDO 的帧间预测中, 即对 P 帧中任一 MB, 选择最优的帧间模式以使主观率失真性能达到最优。P 帧 MB 的帧间预测分为两步: 运动估计(ME)和模式选择(MD)。ME 是寻找具有最小运动代价的块作为当前分割的最优匹配块, 基于客观失真的运动代价定义如下:

$$\begin{aligned} & \text{MCOST}(s, c, \text{mode}|QP) \\ &= \text{SA}(T)D(s, c, \text{mode}|QP) + \lambda_{\text{MOTION}} \cdot R(\Delta MV) \quad (9) \end{aligned}$$

H.264 亚像素 ME 的性能很高, 是去除帧间冗余的主要手段, 但计算复杂度极高, 是 P 帧编码中最为耗时的部分, 若用 SSIM 失真替换绝对(变换)误差和 $\text{SA}(T)D$ 进行 ME, 计算复杂度将进一步增加, 因为 SSIM 计算的复杂度高于 $\text{SA}(T)D$, 而且 SSIM 失真与运动矢量差值的码率 $R(\Delta MV)$ 的关系不明确, 导致运动代价函数的 λ 不易确定, 因此 ME 时用 SSIM 失真替换 $\text{SA}(T)D$ 将导致计算复杂度剧增而主观 RDO 性能增益相对较小。为此本文算法中的 ME 仍用传统基于 $\text{SA}(T)D$ 的方法, 而将重点放在 ME 后的帧间模式选择中, 定义 MB 帧间模式选择时基于 SSIM 失真的代价函数如下:

$$\begin{aligned} J_{\text{SSIM}}(s, c, \text{mode}|QP) &= D_{\text{SSIM}}(s, c, \text{mode}|QP) \\ &+ \lambda_{\text{SSIM}} \cdot R(s, c, \text{mode}|QP) \quad (10) \end{aligned}$$

其中 λ_{SSIM} 是基于 SSIM 的 RDO 中对码率与 SSIM 失真进行权衡的 Lagrange 乘子。由 2.2 节, P 帧 MB 帧间模式选择时的 SSIM 失真定义为

$$D_{\text{SSIM}}(s, c, \text{mode}|QP) = 1 - \text{SSIM}(s, c) \quad (11)$$

H.264 帧间模式选择的最小单位是 4×4 块, 因此本文 SSIM 失真计算时也是以 4×4 块为单位, 较大块(比如 MB)的 SSIM 值是其中各不重叠 4×4 块 SSIM

值的平均。在 RDO 框架中用 SSIM 失真代替 SSE 以期得到更好的主观质量思想很直接, 但难点在于 λ_{SSIM} 如何确定并自适应地调整, 因为基于 SSE 的码率与失真模型容易得到, 进而可以得到分析型的 λ , 但基于 SSIM 的失真量化及失真码率模型都不易得到解析式, 导致 λ_{SSIM} 难以确定。

3.3 MB 层 λ 的自适应调整

在基于 SSIM 的 RDO 框架中, 如何得到计算简单的分析型 λ_{SSIM} 是一个难点问题, 另外不同帧及不同 MB 都具有不同的率失真特性, 若能在 MB 层根据图像内容自适应地调节 λ_{SSIM} 便可对码率与 SSIM 失真进行更好地平衡, 进一步提高主观率失真性能。这两个问题在作者之前研究成果^[13]的基础上, 可以较容易地解决, 因为已经得到了 SSIM- Q_{step} 的失真模型。

为了使基于 SSIM 失真的代价函数最小, 将式(10)求导并令导数为零, 从而得到

$$\lambda_{\text{SSIM}} = -\frac{dD_{\text{SSIM}}}{dR} = -\frac{dD_{\text{SSIM}}/dQ_{\text{step}}}{dR/dQ_{\text{step}}} \quad (12)$$

由此 λ_{SSIM} 求解分为两部分, 即 $dD_{\text{SSIM}}/dQ_{\text{step}}$ 及 dR/dQ_{step} , 其中 $dD_{\text{SSIM}}/dQ_{\text{step}}$ 表示 SSIM 失真随 Q_{step} 的变化情况, 与具体图像的特性相关, 对 P 帧第 i 个 MB 可由式(3)求导得到, 即

$$dD_{\text{SSIM},i}/dQ_{\text{step},i} = k_i \quad (13)$$

而 dR/dQ_{step} 表示码率随 Q_{step} 的变化情况, 也与具体图像的特性相关, 但与采用的失真属性无关, 因此经典的码率模型均可使用。将式(7)代入码率模型式(6)可得码率与 Q_{step} 的关系:

$$R = a \log_2 \left(\frac{12b}{Q_{\text{step}}^2} \right) \quad (14)$$

由式(14)对 Q_{step} 求导可得下式:

$$\frac{dR}{dQ_{\text{step}}} = \frac{-2a}{\ln 2 \times Q_{\text{step}}} \quad (15)$$

最后将式(13)与式(15)代入式(12), 可求得 P 帧第 i 个 MB 的 λ_{SSIM} 解析式如下:

$$\lambda_{\text{SSIM},i} = \frac{\ln 2}{2a} \times k_i \times Q_{\text{step},i} \quad (16)$$

由式(16)可见, 在相同 Q_{step} 下, $\lambda_{\text{SSIM},i}$ 还与 MB 的 SSIM 失真特性 k_i 线性相关, k_i 较小的 MB 将使用较小的 $\lambda_{\text{SSIM},i}$ 来得到更优的主观率失真性能。但同一帧内不同 MB 的 k_i 变化剧烈, 直接使用 k_i 来调节将导致 $\lambda_{\text{SSIM},i}$ 值不稳定。为此本文通过大量实验来确定式(16)中的常数, 且将 k_i 线性调节到帧中所有 MB 的 k_i 均值附近的一定范围内, 即 $\lambda_{\text{SSIM},i}$ 调节如下:

$$\lambda_{\text{SSIM},i} = 0.000023 \times Q_{\text{step},i} \times k'_i,$$

$$k'_i = \begin{cases} 1 + \alpha \times \frac{k_i - k_f}{k_{\text{max}} - k_f}, & k_i \geq k_f \\ 1 - \alpha \times \frac{k_i - k_f}{k_{\text{min}} - k_f}, & k_i < k_f \end{cases} \quad (17)$$

其中 $k_f, k_{\text{max}}, k_{\text{min}}$ 分别是帧中所有 MB 的 SSIM 失真特性 k_i 的均值, 最大值, 最小值。本文将调节幅度 α 设置为 0.3, 即将各 MB 的 k_i 比率线性调节到 $[0.7k_f, 1.3k_f]$ 范围内。

考虑到相同 Q_{step} 下各 MB 的码率特性也不相同, 若 $\lambda_{\text{SSIM},i}$ 能根据码率特性进行调节, 整帧主观率失真性能将进一步提高。P 帧 MB 的码率特性常由平均绝对差(MAD)来衡量, 如文献[1,10,12,13], 而且由式(12)可见, $\lambda_{\text{SSIM},i}$ 与码率特性成反比, 即 Q_{step} 相同但需要更多码率的图像(MAD 值大)将使用更小的 $\lambda_{\text{SSIM},i}$, 因为这些图像可用相对较小的码率增加来得到很大的失真减小, 从而得到更好的率失真性能, 这一点在文献[2,8,9]中也有类似的结论。因此根据前一帧各 MB 的实际 MAD 值对当前 P 帧第 i 个 MB 的 λ_{SSIM} 进一步调节如下:

$$\lambda_{\text{SSIM},i} = 0.000023 \times Q_{\text{step},i} \times k'_i / \text{MAD}'_i \quad (18)$$

$$\text{MAD}'_i = \begin{cases} 1 + \beta \times \frac{\text{MAD}_i - \text{MAD}_f}{\text{MAD}_{\text{max}} - \text{MAD}_f}, & \text{MAD}_i \geq \text{MAD}_f \\ 1 - \beta \times \frac{\text{MAD}_i - \text{MAD}_f}{\text{MAD}_{\text{min}} - \text{MAD}_f}, & \text{MAD}_i < \text{MAD}_f \end{cases} \quad (19)$$

其中 $\text{MAD}_f, \text{MAD}_{\text{max}}, \text{MAD}_{\text{min}}$ 分别是前一帧所有 MB 实际 MAD 的均值, 最大值, 最小值。同样将调节幅度 β 设置为 0.3, 即将各 MB 的 MAD 比率线性调节到 $[0.7\text{MAD}_f, 1.3\text{MAD}_f]$ 范围内。

3.4 基于 SSIM 的 RDO 帧间预测算法流程

本文在文献[13]的基础上实现基于 SSIM 的 RDO 帧间模式选择及 λ 调整, 其中 I 帧 RC 及 P 帧 RC 的 GOP 层和帧层目标比特分配与文献[1,13]相同。对于 P 帧第 i 个 MB, 在采用文献[13]算法得到最优 $Q_{\text{step},i}$ 后, 其 $\lambda_{\text{SSIM},i}$ 调整及基于 SSIM 的 RDO 帧间模式选择步骤如下:

(1)根据前一帧相应 MB 的实际编码结果, 由式(18)计算当前编码 MB 的 $\lambda_{\text{SSIM},i}$;

(2)为当前 MB 的不同分割块仍根据式(9)进行

基于 SSE 的帧间 ME, 选择运动代价最小的预测块作为最优匹配块;

(3)根据式(10)选择主观率失真代价最小的预测模式作为 MB 最优帧间编码模式。

3.5 计算复杂度分析

本文算法是在文献[13]的基础上增加了 MB 层自适应 λ_{SSIM} 的调整及基于 SSIM 的 RDO 帧间模式选择, λ_{SSIM} 计算非常直接, 只有 RDO 帧间模式选择时 SSIM 失真的计算比 SSE 稍微复杂, 因此本文算法的计算复杂度略高于文献[13], 但都是一次编码算法, 不存在迭代计算, 可以应用于实际编码环境。

4 实验结果与分析

本文算法在 JM10.2^[4]平台上实现, 主要实验参数见表 1。编码结果同时给出了 3 种算法 PSNR 和 SSIM 对比, 即经典的 MB 层 RC 和 RDO 算法(JVT-G012)^[1], 基于 SSIM 的 MB 层最优 RC 算法(未进行基于 SSIM 的 RDO 帧间预测及 λ 调整)^[13]和本文算法。注意 3 种算法中 P 帧 MB 编码时都禁用帧内编码模式, 而 MB 所有帧间编码模式都启用。

图 2 给出了 Students 及 Grandma 序列分别在目标码率为 64 kbps 及 96 kbps 时编码的各帧 SSIM 对比图。可见本文算法相比文献[13]在实施基于 SSIM 的 RDO 帧间模式选择及 λ 调整后得到的绝大多数帧 SSIM 值更高, 表明了本文算法的有效性; 而本文算法相比 JVT-G012^[1]SSIM 提高更加明显, 这是基于 SSIM 的 MB 层最优 RC 和主观 RDO 帧间模式选择及 λ 调整共同作用的结果。

表 2 给出了更多序列在不同目标码率下的实验结果, 包括平均 PSNR, SSIM 值以及实际码率的比较。可见 3 种算法的实际输出码率都非常接近目标码率, 说明这 3 种算法的 RC 准确性很高而且控制能力相当, 这是由于 3 种算法都采用了相似的二次码率量化模型。在平均 PSNR 方面, 文献[13]比文献[1]对绝大多数序列都有不同程度的提高, 这是由于采用了 SSIM 最优的 MB 层目标比特分配和 Q_{step} 求解而非根据复杂度比率进行经验型的比特分配, 但本文算法得到的平均 PSNR 要小于另两种算法, 这是由于采用了基于 SSIM 的 RDO 帧间模式选择而非基于 SSE 的 RDO。而在平均 SSIM 方面, 本文算法相比另两种算法均有明显提高, 说明本文算法更有效地编码了图像结构信息, 获得了更好的主观

表1 主要实验参数

编码方式	格式	帧率	帧数	I 帧周期	搜索范围	ME 精度	参考帧数	熵编码	RDO	RC
IPPP	QCIF	30 fps	100	0	16	1/4	1	CAVLC	on	on

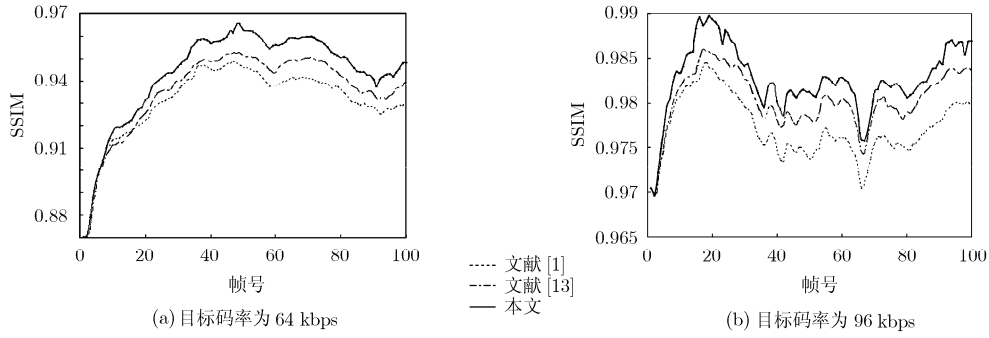


图2 Students 和 Grandma 序列在不同目标码率时编码的各帧 SSIM 对比图

表2 本文算法与文献[1], 文献[13]算法编码性能比较(QCIF)

序列	目标码率(kbps)	PSNR(dB)			SSIM			实际码率(kbps)		
		文献[1]	文献[13]	本文	文献[1]	文献[13]	本文	文献[1]	文献[13]	本文
Container	64	36.74	37.38	35.84	0.9500	0.9548	0.9572	64.12	64.09	64.15
	96	38.73	39.26	37.81	0.9626	0.9676	0.9699	96.13	96.14	96.27
Grandma	64	37.49	38.14	37.73	0.9530	0.9594	0.9638	64.29	64.28	64.22
	96	40.36	40.95	40.52	0.9772	0.9803	0.9832	96.28	96.21	96.41
Mobile	96	25.24	25.37	25.07	0.8601	0.8650	0.8729	96.43	96.27	96.37
	128	26.59	26.70	26.34	0.8953	0.8985	0.9053	128.45	128.39	128.33
Sign_irene	64	33.26	33.51	33.17	0.9071	0.9135	0.9174	64.18	64.10	64.16
	96	35.67	36.03	35.60	0.9470	0.9521	0.9538	96.32	96.18	96.24
Students	64	34.02	34.43	34.16	0.9309	0.9355	0.9443	64.34	64.20	64.22
	96	37.06	37.41	36.87	0.9651	0.9677	0.9718	96.47	96.28	96.38



图3 Students 序列第83帧3种算法编码主观质量对比(64 kbps)

质量, 这是由于本文在文献[13]的基础上又实施了基于 SSIM 的 RDO 帧间模式选择及 MB 层 λ 调整, 比如 Mobile 序列在目标码率为 96 kbps 时本文算法相比文献[13]平均 SSIM 值提高了 0.0079, 而相比文献[1]算法提高了 0.0128。

图 3 给出了 Students 序列目标码率为 64 kbps 时 3 种算法编码的第 83 帧的主观效果图, 可见本文算法编码的图像具有更好的结构信息, 因而具有更好的主观质量。注意人脸部分及左下角人物衣服纹理部分的主观质量提高更加明显。

图 4 给出了不同序列在更多目标码率下编码的

主观率失真性能比较。可见本文算法与文献[13]比 JM 中采用的文献[1]算法在各种目标码率下对所有序列均取得了更好的主观率失真性能。而本文算法相比文献[13]算法的主观率失真性能又有较大提高, 尤其在目标码率较低时主观率失真性能的提高更加明显, 比如 Container 与 Mobile 序列。

表 3 给出了不同序列的总编码时间和不同算法间的编码时间比率。可见本文算法相比文献[1]平均编码时间大约增加了 6%, 说明本文算法带来的计算复杂度增加很小。

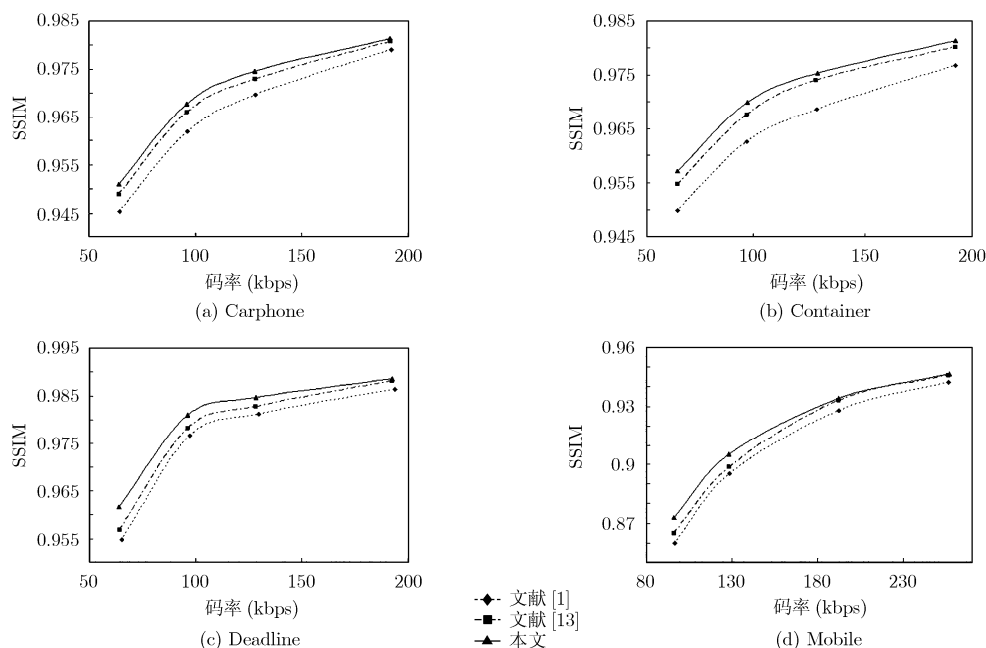


图4 主观率失真性能比较

表3 本文算法与文献[1], 文献[13]算法编码时间比较(QCIF)

序列	目标码率(kbps)	编码时间(s)			编码时间比率	
		文献[1]	文献[13]	本文	文献[13]相比文献[1]	本文相比文献[1]
Container	64	69.682	71.231	71.259	1.02	1.02
	96	66.516	65.784	67.832	0.99	1.02
Grandma	64	71.235	71.187	72.482	1.00	1.02
	96	65.720	67.651	68.451	1.03	1.04
Mobile	96	60.192	62.144	64.573	1.03	1.07
	128	69.045	70.455	73.563	1.02	1.07
Sign_irene	64	61.626	63.791	66.730	1.04	1.08
	96	67.283	69.153	71.012	1.03	1.06
Students	64	58.862	61.857	63.219	1.05	1.07
	96	65.938	67.767	67.674	1.03	1.03
Average		64.621	67.042	68.718	1.04	1.06

5 结束语

针对传统基于客观失真的 RC 和 RDO 无法得到最优的主观质量的问题, 本文在之前研究成果基础上进一步提出了 H.264 基于 SSIM 的 RDO 帧间模式选择算法, 并且提出了一种 MB 层自适应的 λ 调整机制。实验结果证实, 相比基于客观质量的 MB 层 RC 算法以及 SSIM 最优的 MB 层 RC 算法(但未进行基于 SSIM 的 RDO 帧间预测及 λ 调整), 本文算法在 MB 层更有效地权衡了码率和 SSIM 失真, 得到了更高的主观率失真性能, 图像结构信息得到有效保护, 图像主观质量明显提高, 更加符合人的主观感受; 且计算复杂度增加很小, 可用于实际编码环境。如何将 SSIM 用于基于 RDO 的分块整像素

及亚像素 ME 中及如何确定运动代价函数中的 λ 是我们下一步的研究方向。

参考文献

- [1] Li Zheng-guo, Pan Feng, Lim K P, *et al.* Adaptive basic unit layer rate control for JVT[C]. Proceedings of Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG[C]. JVT-G012, 7th Meeting, Pattaya, Thailand, Mar. 2003. http://wftp3.itu.int/av-arch/jvt-site/2003_03_Pataya.
- [2] Wiegand T and Girod B. Lagrange multiplier selection in hybrid video coder control[C]. Proceedings of IEEE International Conference on Image Processing, Thessaloniki, Greece, 2001: 542-545.
- [3] Wang Zhou, Bovik A C, Sheikh H R, *et al.* Image quality assessment: from error visibility to structural similarity[J].

- IEEE Transactions on Image Processing*, 2004, 13(4): 600–612.
- [4] JM Reference Software [CP/OL]. <http://iphome.hhi.de/suehring/tml/download/>. 2008.
- [5] 杨春玲, 肖冬琴. 基于SSE和SSIM的H.264帧内预测模式选择改进算法[J]. 电子与信息学报, 2011, 33(2): 289–294.
Yang Chun-ling and Xiao Dong-qin. Improvements for H.264 intra mode selection based on SSE and SSIM[J]. *Journal of Electronics & Information Technology*, 2011, 33(2): 289–294.
- [6] 杨春玲, 王华兴, 梁荣锟. 基于结构相似的H.264帧间预测改进算法[J]. 计算机学报, 2009, 32(8): 1603–1610.
Yang Chun-ling, Wang Hua-xing, and Liang Rong-kun. Improved inter prediction based on structural similarity for H.264[J]. *Chinese Journal of Computers*, 2009, 32(8): 1603–1610.
- [7] Channappayya S S, Bovik A C, and Heath R W. Rate bounds on SSIM index of quantized images[J]. *IEEE Transactions on Image Processing*, 2008, 17(9): 1624–1639.
- [8] Huang Y H, Ou T S, and Chen H H. Perceptual-based coding mode decision[C]. Proceedings of IEEE International Symposium Circuits and Systems, Paris, France, 2010: 393–396.
- [9] Huang Y H, Ou T S, Su P Y, *et al.* Perceptual rate-distortion optimization using structural similarity index as quality metric[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(11): 1614–1624.
- [10] Wang Miao-hui and Yan Bo. Lagrangian multiplier based joint three-layer rate control for H.264/AVC[J]. *IEEE Signal Processing Letters*, 2009, 16(8): 679–682.
- [11] Zhang Jun, Yi Xiao-quan, Ling N, *et al.* Context adaptive Lagrange multiplier (CALM) for rate-distortion optimal motion estimation in video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(6): 820–828.
- [12] Ou T S, Huang Y H, and Chen H H. SSIM-based perceptual rate control for video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2011, 21(5): 682–691.
- [13] 崔子冠, 朱秀昌. H.264结构相似性最优的宏块层码率控制算法[J]. 电子与信息学报, 2011, 33(6): 1339–1344.
Cui Zi-guan and Zhu Xiu-chang. Structural similarity optimal MB layer rate control algorithm for H.264[J]. *Journal of Electronics & Information Technology*, 2011, 33(6): 1339–1344.

崔子冠：男，1982年生，博士生，研究方向为视频编码与传输。
朱秀昌：男，1947年生，教授，博士生导师，研究方向为图像处理、视频编码。