

基于 SSIM 的 HEVC 帧内编码率失真优化

朱天之, 郁 梅, 蒋刚毅*, 陈 芬, 邵 枫, 彭宗举

(宁波大学 信息科学与工程学院, 浙江 宁波 315211)

摘要:率失真优化(RDO)是视频编码压缩中的关键技术之一。传统的率失真优化技术使用误差平方和(SSE)或绝对差和(SAD)度量失真,不能获得较好的视觉感知质量。本文针对新一代视频编码标准高效视频编码(HEVC)帧内编码 RDO,提出了一种基于结构相似度(SSIM)度量失真的 RDO 方法。首先,根据 HEVC 编码树形单元(CTU)结构设计了基于 SSIM 的失真计算方法;然后,提出了一种基于量化器推导的码率-量化步长($R-\Delta$)模型和一种基于统计分析的失真-量化参数($D_{SSIM}-QP$)模型,用于求取拉格朗日乘数;最终,使用多 QP 优化方法求取模型参数。实验表明,相对于 HEVC 传统的 RDO,针对全 I 帧(即所有帧都为帧内编码)固定 QP 编码和多 QP 优化编码,在相同的 SSIM 条件下,码率分别平均降低 8.4% 和 13.9% 左右,同时编码复杂度分别平均增加约 3% 和 2%。

关键词:视频编码; 高效视频编码(HEVC); 率失真优化(RDO); 结构相似度(SSIM)

中图分类号:TN919.8 **文献标识码:**A **文章编号:**1005-0086(2014)12-2362-09

SSIM-based rate distortion optimization for HEVC intra coding

ZHU Tian-zhi, YU Mei, JIANG Gang-yi*, CHEN Fen, SHAO Feng, PENG Zong-ju

(Faculty of Information Science and Engineering, Ningbo University, Ningbo 315211, China)

Abstract:Rate distortion optimization (RDO) is one of the key techniques in the field of video coding. Either sum of square error (SSE) or sum of absolute difference (SAD) is widely adopted to evaluate distortion in conventional RDO, which fails to achieve high consistency with human visual perception. Recently, the structural similarity (SSIM) index has been attracting more and more attention as a perceptual visual quality metric for optimizing video codecs. Meanwhile, high efficiency video coding (HEVC) as the emerging video coding standard creates many new opportunities and challenges. In this paper, we propose a method of structural similarity index based RDO for HEVC intra coding. Firstly, an SSIM-based distortion computational method is designed according to the coding tree unit (CTU) structure of HEVC. Secondly, a quantizer-derived rate-quantization step ($R-\Delta$) model and a statistical analysis-based distortion-quantization parameter ($D_{SSIM}-QP$) model are presented for Lagrange multiplier calculation. Finally, the model parameters are estimated based on multi-QP optimization. Extensive experimental results show that compared with traditional RDO approach of HEVC, for the same SSIM score, the proposed approach can achieve an average rate reduction of about 8.4% and 13.9% with 3% and 2% encoding time increasement, respectively for fixed QP and multi-QP optimizations on all intra coding configure.

Key words:video coding; high efficiency video coding (HEVC); rate distortion optimization (RDO); structural similarity (SSIM)

1 引 言

由 ISO/IEC 运动图像专家组和 ITU-T 视频

编码专家组联合组建的视频编码小组(JCT-VC), 于 2013 年 1 月正式发布了新一代视频编码标准——高性能视频编码(HEVC)^[1], 相比于 H.

* E-mail:jianggangyi@126.com

收稿日期:2014-06-12 修订日期:2014-10-16

基金项目:国家自然科学基金(U1301257,61171163,61271270,61271021,61311140262)资助项目

264/AVC,其压缩性能有了显著的提升。

视频压缩的目标是在任何给定的码率下,尽可能地获得高的视觉质量。为此,率失真优化(RDO)技术被用来指导编码器选择最优编码模式。传统的RDO技术使用误差平方和(SSE)和绝对差和(SAD)度量失真。然而,SSE和SAD不能很好体现主观质量^[2],难以很好地指导编码器获得最优的视觉质量。由于结构相似度(SSIM)^[3,4]计算简单,并能较好地体现主观质量,被广泛应用于失真度量。Mai等^[5]率先提出了一种基于SSIM的H.264/AVC帧内编码RDO方法,给出了拉格朗日乘数的计算方法,但获得性能的提升有限;Huang等^[6]针对H.264/AVC编码的RDO,使用1-SSIM作为失真度量,提出用传统RDO对关键帧进行二次编码,从而求取拉格朗日乘数^[6];杨春玲等^[7]认为图像平坦区域可以直接使用SSE度量失真,提出了联合SSE和SSIM的RDO方法;Wang等^[8]根据残差变换系数的统计特性,提出了一个部分参考(RR)的SSIM-QP模型以及一个简化的码率与量化参数(R-QP模型,并以此计算拉格朗日乘数和估计模型参数;后期又对该算法进行了扩展,提出了基于运动信息以及局部对比度加权的宏块层拉格朗日乘数自适应调整的方法^[9];Wang等^[10]还提出一种基于变换域SSIM的残差系数归一化方法,应用于H.264/AVC的RDO^[10];随后,Rehman等^[11]将该方法运用到了HEVC编码中,但性能提升并不显著^[11];Yeo等^[12]使用1/SSIM度量失真,并提出一种用SSE估计SSIM的方法;同时Yeo等^[13]又将该方法推广到了HEVC中,所提升的性能不明显,且随着QP的增加,使用SSE对SSIM的估计精度会逐渐下降;Qi等^[14]在Yeo等方法的基础上结合运动信息,提出了一种针对HEVC编码标准的基于SSIM的RDO方法;此外,也有将SSIM应用于音视频编码标准(AVS)的帧间模式选择中,对静止块较多的序列,表现出较优的率失真性能^[15]。

综上所述,针对HEVC编码标准使用SSIM进行RDO的研究还比较少,且直接将H.264的RDO方法应用于HEVC并不能获得较好的率失真性能。本文针对HEVC编码新的技术特性,提出了一种以SSIM作为失真度量的RDO新方法。先根据HEVC编码树形单元(CTU)结构设计了基于SSIM失真计算方法;然后提出了一种基于量化器分析的码率-量化步长(R-Δ)模型和一个基于统计分析的失真-量化参数($D_{SSIM-QP}$)模型,并用此计算拉格朗日乘数。实验结果表明,在相同码率下,解码重建帧的视觉质量有较明显的提高。

2 基于SSIM的RDO

2.1 SSIM失真计算

假设 x 和 y 分别表示原始图像块和失真图像块,其SSIM定义为^[3]

$$SSIM = \left(\frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \right) \left(\frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right) \quad (1)$$

其中: μ_x 和 μ_y 分别表示 x 和 y 图像块的均值; σ_x 和 σ_y 分别表示 x 和 y 图像块的方差; σ_{xy} 是 x 和 y 的协方差; c_1 和 c_2 为调节常数。SSIM的取值范围属于 $[0,1]$,且值越接近1,表示图像块质量越好。本文使用1-SSIM作为失真度量。

当引入SSIM度量失真时,为了使编码器递归划分编码单元(CU)过程正确,失真的计算必须满足“可加性准则”。这是由“单一字符失真”(如SSE,SAD)过渡到“非单一字符”(如SSIM)失真所带来的问题,即所谓可加性准则,是指对于相同内容的图像,其失真按照不同窗口大小计算时满足加性相等。如图1所示,以HEVC中LCU(largest coding unit)的失真计算为例,对于一个给定的LCU,其按照原始尺寸(64×64)计算的失真必须严格等于按照子CU尺寸计算的失真之和。

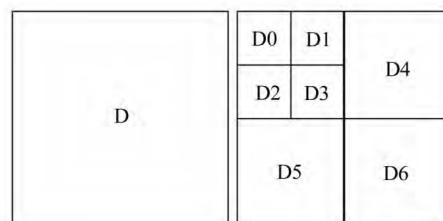


图1 可加性准则

Fig. 1 The rule of additivity

考虑到HEVC中变换单元(TU)的最小尺寸为 4×4 ,所以计算SSIM的窗口必须设置为 4×4 。此外,为了满足可加性准则,采用无重叠滑动窗,根据文献^[16],采用无重叠滑动窗计算SSIM,其对图像质量的预测准确度并不会明显下降。对于某一编码块,其失真可根据公式

$$D_{SSIM} = \sum_{i=0}^N (1 - SSIM_i) \times c \quad (2)$$

计算。其中: N 是窗口个数; c 为一缩放因子。图2给出了一个 8×8 图像块的CU失真计算示例。

2.2 基于量化器分析的R-Δ模型

引入SSIM度量失真后,率失真代价计算公式重新定义为

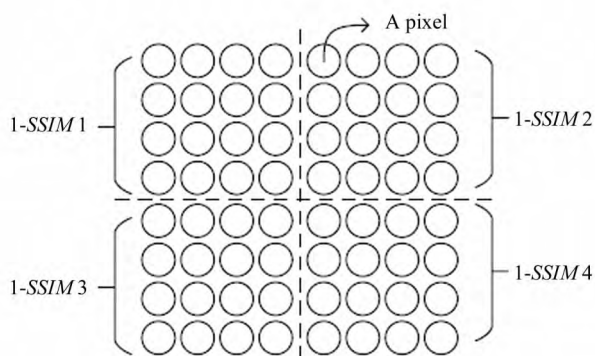


图2 8×8图像块的CU失真计算示例

Fig. 2 The example of distortion calculation for 8×8 CU

$$J = D_{SSIM} + \lambda R \quad (3)$$

为了最小化率失真代价,等式两边分别对码率 R_{SSIM} 求导,并令其等于0,有

$$\lambda = -\frac{dD_{SSIM}}{dR} = -\frac{dD_{SSIM}/d\Delta}{dR/d\Delta} \quad (4)$$

其中, Δ 为量化步长。由于拉格朗日乘数 λ 和量化步长 Δ 有着密切的关系,而编码失真主要是来自量化失真,即由量化器负责。事实上,HEVC 采用的是均匀标量量化。若将量化实轴分成 M 个不相交的区间,区间的大小即为量化步长 Δ ,那么信源符号经量化重建为 $q\Delta, q=0, \pm 1, \pm 2, \pm 3 \dots$ 。在对实际图像编码的情况下,其量化级是有限的,即输入的符号集是一个有限长集合。假设令 L 表示有效量化范围, $L < \infty$ 。则有

$$\Delta = \frac{L}{M} = \frac{L}{2^b} \quad (5)$$

其中, b 表示每个符号所需的比特数。由此可得

$$R' = Nb = N \log_2 \left(\frac{L}{\Delta} \right) \quad (6)$$

其中: N 为当前编码块符号个数(像素个数); R' 表示量化器的输出比特率。假设 R 是经过熵编码后的输出比特率, X 表示当前编码块信源, \hat{X} 表示量化后的信号。那么

$$R \approx NH(\hat{X}) = -N \sum_{q=-\infty}^{\infty} P_q \log_2 P_q \quad (7)$$

其中, P_q 为某一信源符号量化为 q 的概率。假设信源概率密度为 $f_x(x)$, 当 Δ 较小时, $f_x(x)$ 在每个区间内近似为常数,有 $f_x(x) \approx f_x(q\Delta)$, $P_q \approx f_x(q\Delta)\Delta$, 带入公式(7)得

$$R \approx -N \sum_{q=-\infty}^{\infty} [f_x(q\Delta)\Delta] \log_2 [f_x(q\Delta)\Delta] \approx$$

$$-N \left(\int_{-\infty}^{\infty} f_x(x) \log_2 f_x(x) dx + \int_{-\infty}^{\infty} f_x(x) \log_2 \Delta dx \right) = \quad (8)$$

$$N(H(X) - \log_2 \Delta) = N \log_2 \left(\frac{2^{H(X)}}{\Delta} \right)$$

式(6)和(8)左右两边分别对量化步长 Δ 求导,有

$$\frac{dR}{d\Delta} \approx \frac{dR'}{d\Delta} \quad (9)$$

其中, R' 是量化器的输出比特率。由此可见,量化器输出比特率 R' 和熵编码输出比特率 R 相对于 Δ 的变化率近似相等。对于求取拉格朗日乘数只需要知道 R 对 Δ 的导数,因此 R - Δ 模型可用式(6)近似表示。于是有

$$\lambda = -\alpha \frac{dD_{SSIM}/d\Delta}{dR'/d\Delta} \quad (10)$$

其中, α 为调节因子,保证式(10)等式两边相等。需要说明的是,以上为了便于分析,使用的是量化步长 Δ ,而最终编码时使用的是量化参数 QP 。由于 HEVC 的 Δ 和 QP 存在^[17]

$$\Delta = 2^{(QP-4)/6} \quad (11)$$

因此,根据式(6)、(10)和(11)得

$$\lambda = \alpha' \cdot 2^{(QP-4)/6} \left(\frac{dD_{SSIM}}{dQP} \right) \quad (12)$$

其中, α' 为与信源相关的系数。

2.3 基于统计分析的 D_{SSIM} - QP 模型

由于很难得到 D_{SSIM} 和 QP 的封闭解^[18], 因此通过统计分析得到 D_{SSIM} - QP 模型。根据大量实验结果可以发现, D_{SSIM} 和 QP 满足

$$D_{SSIM} = \beta e^{\gamma QP} \quad (13)$$

其中, β 和 γ 为与信源相关的常数。图 3(a)显示了序列 BQMall (832×480), KristenAndSara (1280×720) 和 FourPeople (1280×720) 的 D_{SSIM} 和 QP 的关系图。图中的 6 个数据点由实际编码得到, 3 条曲线为采用最小二乘法按照式(13)拟合得到。从图可以发现, 式(13)可以较准确地反映出实际编码中失真 D_{SSIM} 值与 QP 之间的关系, 且拟合中的参数 β 和 γ 的取值与信源相关。图 3(b)给出了所有序列的拟合线性相关系数 R^2 , R^2 值越接近 1 表示拟合结果越准确。可以发现, 所有序列拟合的线性相关系数 R^2 都在 0.98 以上, 且平均值高达 0.9938。因此能够认为, 指数模型可较精准地描述 D_{SSIM} 和 QP 的关系。

2.4 参数求取

由于 1-SSIM 的值在 0~1 之间, 因此使用缩放因子 c 把根据式(2)求取的失真放大, 最后转化为无

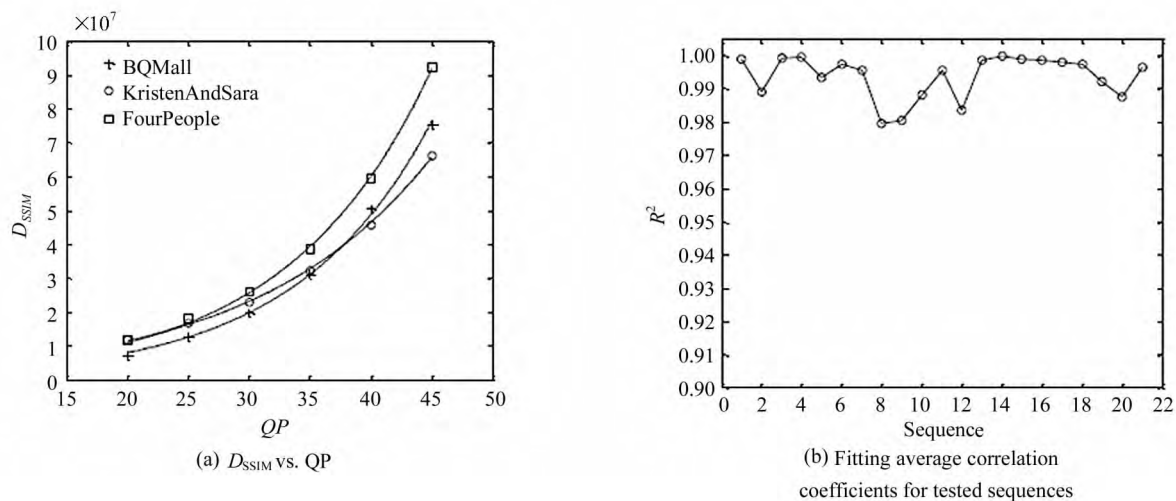


图 3 D_{SSIM} -QP 统计模型
Fig. 3 The statistical model of D_{SSIM} -QP

符号整型参与计算。实验发现,当取 $c=10^2$ 时,相当于只取原始(1-SSIM)值的两位小数位,此时精度损失太大;而当 $c=10^4$ 时,编码器对纹理信息过于敏感,导致编码器总是趋向于使用较小的 CU 分割,使得编码重建帧在时域上连续性较差,会出现闪烁的现象;当参数 $c=10^3$,闪烁现象得以克服,并且 CU 分割结果符合一般规律,即纹理复杂的区域划分为较小的块,平坦区域划分为较大的块。因此,本文最终取 $c=10^3$ 。

根据式(12)和(13),最终有
$$\lambda = \alpha' \cdot 2^{(QP-4)/6} d(\beta e^{\gamma QP})/dQP \quad (14)$$

考虑到除了 QP,拉格朗日乘数 λ 的求取还和其他编码参数有关,如帧类型、编码结构等,而这些因素不易被量化。针对 HEVC 帧内编码,这里提出用调节因子 ω 综合考虑信源参数及其他因素的影响。因此,拉格朗日乘数 λ 可表示为

$$\lambda = \omega \cdot 2^{(QP-4)/6} (e^{\gamma QP}) \quad (15)$$

为了求取该式中涉及的两个参数 ω 和 γ ,使用多 QP 优化的方法,并设优化目标为

$$\arg \min \{D_{SSIM}(QP_i) + \lambda R(QP_i)\}, QP_i = QP \quad (16)$$

首先每帧分别使用 $\lambda = \{2, 4, 8, 16, 32, 64\}$ 编码;然后让编码器遍历所有的 $QP = \{1, 2, \dots, 51\}$;最后根据率失真代价,选取对应代价最小的 QP。实验选取了分辨率为 WQVGA、WVGA、720p、1 080p 以及 4K 的 5 个测试序列。针对全 I 帧(即所有帧都为帧内编码)编码测试了 100 帧,把被选择频率最高的 QP 作为最优 QP,见表 1。图 4 给出了部分序列 λ -

QP 的拟合曲线。从图可以发现,式(15)可以很好的反应各序列 λ 与 QP 之间的关系,且不同序列得到的曲线重合度较高。因此,为了简化编码过程,可以利用所有序列不同 λ 及对应的最优 QP 的平均值,使用式(15)进行拟合,获取参数 ω 和 γ 用于计算最终拉

表 1 λ 对应的最优 QP
Tab. 1 The optimal QPs corresponding to λ

Sequence	Lagrange multiplier					
	2	4	8	16	32	64
BlowingBubble	22	26	32	37	42	47
BQMall	23	28	32	37	41	47
KristenAndSara	22	26	31	36	41	47
BQTerrace	21	24	32	37	42	46
PeopleOnStreet	21	25	30	36	40	48
Average	21.8	25.8	31.4	36.6	41.2	47.0

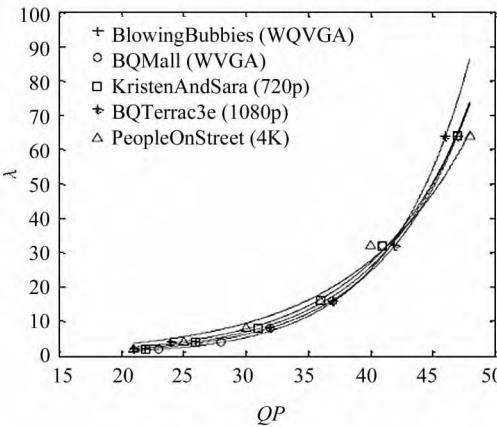


图 4 λ -QP 拟合曲线
Fig. 4 The fitted curves of λ -QP

格朗日乘数。根据拟合结果,得到参数 $\omega=0.2373, \gamma=0.0135$ 。

3 实验结果与分析

为了验证本文算法的率失真性能,在 HEVC 参考软件 HM 13.0^[17] 上实现并测试。所有的测试序列来自 JCT-VC 推荐的公共测试条件 (CTC)^[19], YCbCr 为 4:2:0 格式, 8 bit 深度。从 Class A~F 中每类分别选取 2 个测试序列,共 12 个。Class A~E 分别为 4K、1 080p、WVGA、WQVGA 和 720p 分辨率的测试序列,Class F 主要为几种不同分辨率的屏幕内容测试序列。其他具体编码配置见表 2。

表 2 编码配置参数

Tab. 2 Parameter setting for encoding

Parameter	Setting
Profile	Main
Encoding structure	All intra
Frames	200
QP	22, 27, 32, 37
RDOQ/RDOQTS/SAO	off

使用亮度 Y 分量的 $SSIM_Y$ 以及由 Y;Cb;Cr 3 个分量加权的 $SSIM_w$ 两种方法衡量解码帧质量。加权方法为

$$SSIM_w = \omega_Y SSIM_Y + \omega_{Cb} SSIM_{Cb} + \omega_{Cr} SSIM_{Cr} \quad (17)$$

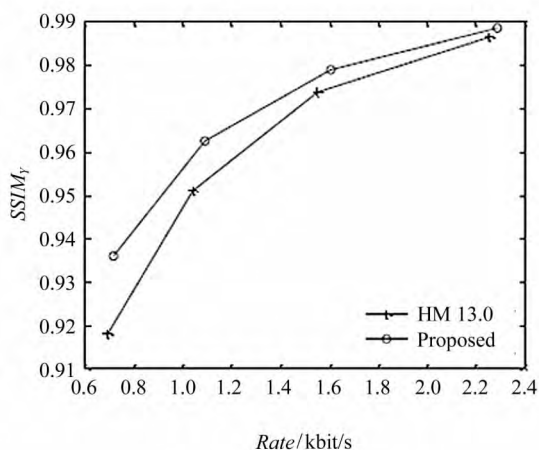
其中: ω_Y 取 0.8; ω_{Cb} 和 ω_{Cr} 分别取 0.1^[10]。此外,使用 BD-Rate(bjontegaard delta bit rate)^[20] 衡量本文算法的率失真性能。BD-Rate 表示相同 SSIM 条件下码率变化的百分比。

首先测试了固定 QP 编码的率失真性能。固定 QP 编码是指每帧中所有 CU 全部使用相同 QP 编码。以 HM 13.0 作为对比,本文算法的率失真性能见表 3。在全 I 帧编码下, $(BD-Rate)_Y$ 和 $(BD-Rate)_w$ 平均分别降低 7.5% 和 7.0%。图 5 给出两个序列的 $SSIM_Y$ -Rate 率失真曲线图。

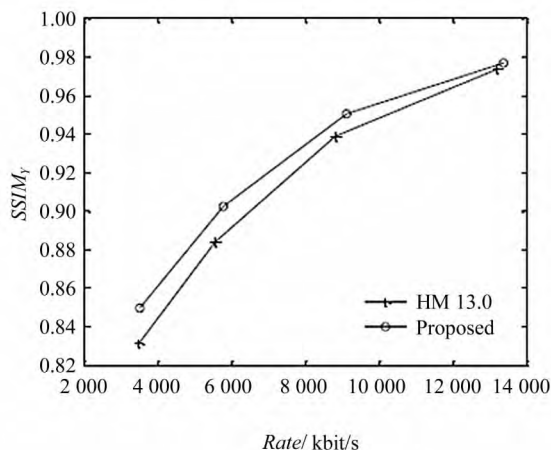
表 3 固定 QP 编码的 BD-Rate

Tab. 3 BD-Rate results with fixed QP Unit: %

Sequence	$(BD-Rate)_Y$	$(BD-Rate)_w$
Traffic	-6.5	-5.8
PeopleOnStreet	-7.6	-7.0
Cactus	-6.1	-5.3
BQTerrace	-7.6	-7.9
BasketballDrill	-8.7	-6.7
PartyScene	-8.5	-8.7
BQSquare	-10.4	-10.0
BlowingBubbles	-7.9	-7.9
FourPeople	-8.2	-7.9
KristenAndSara	-7.9	-7.5
ChinaSpeed	-12.5	-11.2
SlideShow	-9.1	-8.0
Average	-8.4	-7.8



(a) ChinaSpeed



(b) BQSquare

图 5 $SSIM_Y$ -Rate 率失真曲线(固定 QP)

Fig. 5 $SSIM_Y$ vs. rate plots with fixed QP encoding

然后,测试了多 QP 优化编码的率失真性能。HEVC 支持每个 CU 使用不同 QP 编码,在 HM 13.0 平台中通过选项“-dqd(delta QP depth)”开启,深度范围为 0~3,例如:“-dqd 0”表示以 LCU(64×64) 为单位可以使用不同 QP,“-dqd 1”表示以 32×32 的 CU 为单位可以使用不同 QP,以此类推。多 QP 优化是指编码器在给定的一个 QP 范围内,选取率失真代价最小的 QP 作为最终编码量化参数。本文测试的 QP 范围为{QP-2, QP-1, QP, QP+1, QP+2}, 在 HM 13.0 平台中通过选项“-d 2”指定。这种多 QP 优化方法复杂度较高,但是率失真性能也较好,常用于对编码复杂度要求不高同时对率失真性能要求苛刻的场合。在编码器同时开启 dqd 的条件下,

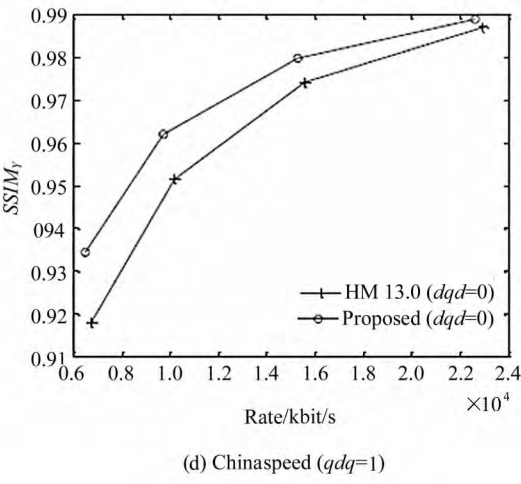
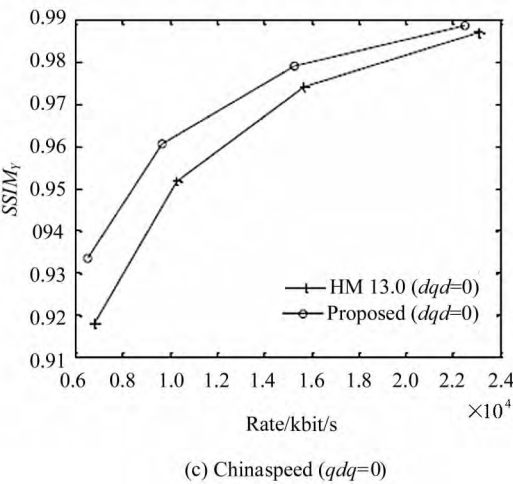
对比 HM 13.0,本文算法的率失真性能见表 4。

由表 4 可以看出,当 HM 的 dqd 选项开启后, $(BD-Rate)_Y$ 和 $(BD-Rate)_w$ 平均分别降低 11.6~13.1%和 9.7~11.7%。图 6 给出了 ChinaSpeed 序列在 dqd 取 0~3 时的 $SSIM_Y-Rate$ 率失真曲线。可以看出,QP 越大,本文算法的性能越好。

为了进一步验证本文算法的主观性能,图 7 给出了当 dqd=3 时 BQSquare 序列第 1 帧的编码结果。图的第 2 行分别给出了框中对应内容的放大图。可以看出,本文算法的结果对结构轮廓保留的比较好,如地板的轮廓和伞的轮廓。尤其是地板的轮廓, HM 13.0 编码重建的结果几乎全部丢失,而本文算法的地板轮廓还清晰可见。

表 4 多 QP 优化编码的 BD-Rate
Tab. 4 BD-Rate results with multi-QP optimization Unit: %

Sequence	$(BD-Rate)_Y$				$(BD-Rate)_w$			
	dqd0	dqd1	dqd2	dqd3	dqd0	dqd1	dqd2	dqd3
Traffic	-9.1	-10.0	-10.5	-10.0	-7.6	-8.4	-9.3	-9.3
PeopleOnStreet	-11.6	-12.5	-13.1	-13.0	-9.1	-10.0	-11.0	-11.2
Cactus	-9.5	-10.3	-10.9	-10.7	-8.1	-8.9	-9.6	-9.7
BQTerrace	-14.0	-15.3	-16.4	-17.0	-11.9	-13.1	-14.4	-15.5
BasketballDrill	-14.7	-15.4	-15.7	-15.1	-9.8	-10.3	-11.0	-11.0
PartyScene	-11.6	-12.3	-13.6	-14.6	-10.8	-11.5	-13.0	-14.3
BQSquare	-15.6	-16.9	-18.6	-20.0	-13.6	-14.8	-16.1	-17.9
BlowingBubbles	-10.8	-12.1	-13.2	-13.7	-9.7	-10.7	-12.1	-13.0
FourPeople	-11.0	-11.7	-12.1	-12.1	-10.1	-11.1	-11.5	-11.8
KristenAndSara	-13.2	-13.6	-13.5	-13.1	-10.7	-11.4	-11.7	-11.7
ChinaSpeed	-17.8	-18.7	-19.3	-19.5	-15.2	-16.1	-16.7	-17.3
SlideShow	-15.8	-15.9	-16.2	-15.2	-13.3	-13.1	-13.6	-13.0
Average	-12.9	-13.7	-14.4	-14.5	-10.8	-11.6	-12.5	-13.0



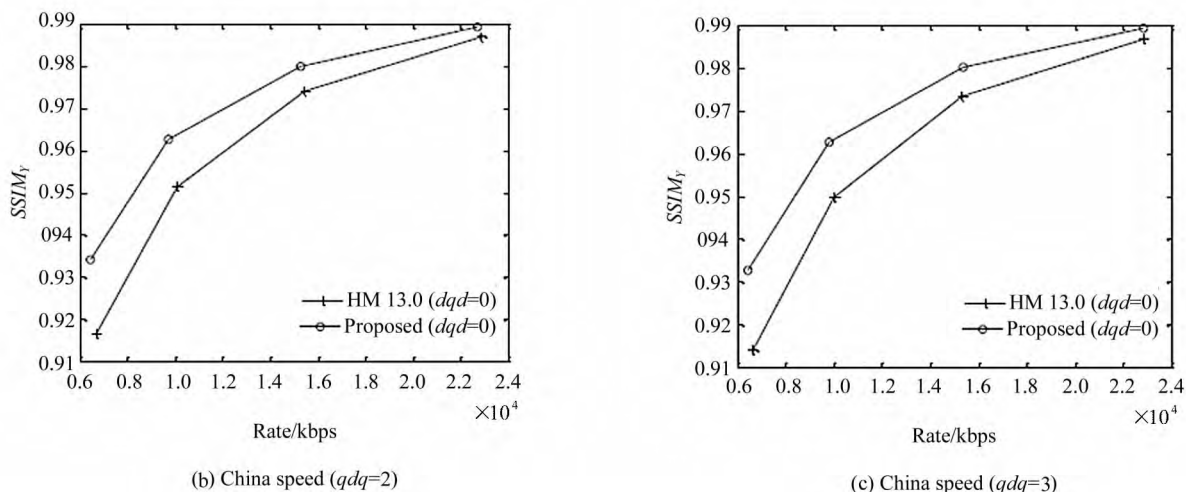
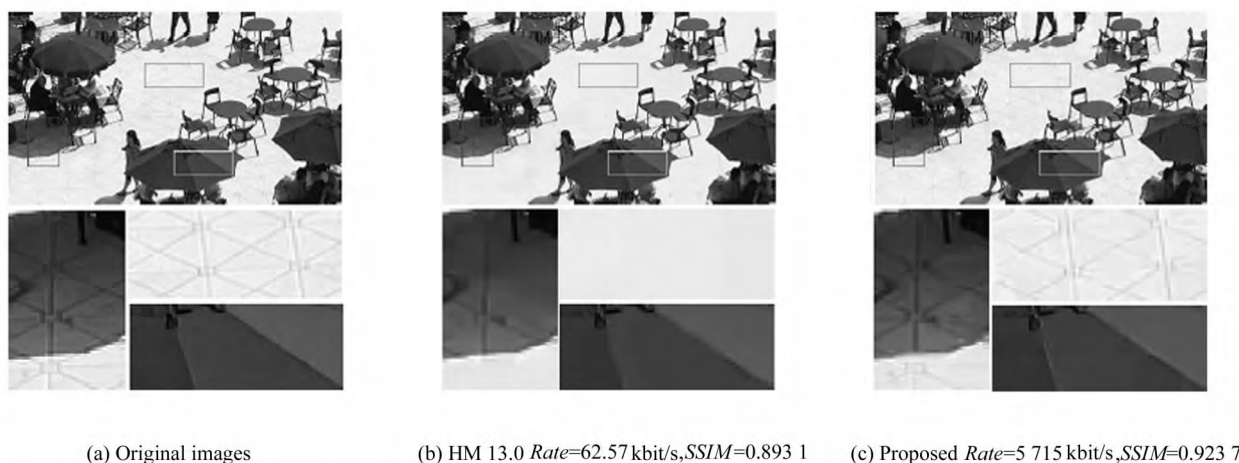
图6 $SSIM_Y$ -Rate 率失真曲线(dqd 开)Fig. 6 $SSIM_Y$ vs. rate plots with multi-QP encoding (dqd on)

图7 本文算法对比 HM 13.0 主观质量

Fig. 7 Subjective quality comparison HM 13.0 and the proposed method

基于感知的失真度量方法很难被应用到视频编码中的最重要的一个原因,是复杂度过高。SSIM 的计算复杂度大约是 PSNR 的 5.9 倍左右^[16]。由于本文计算 SSIM 采用的是无重叠滑窗,因此其计算复杂度大幅下降。此外,HEVC 需要对每个 LCU 遍历分割,在计算原始图像的均值或方差时,每个像素点会重复参与累加计算,这是完全没有必要的。因此在编码一个 LCU 时,先将原始图像按照 4×4 大小计算累加和,并存成一个索引表;然后编码遍历到当前 CU 时,只需要按照索引取出以 4×4 为单位的累加和再进行计算。这种方法可以进一步提高 SSIM 的计算速度。使用编码时间的百分比衡量编码复杂度 C ,即

$$C = \frac{T_{\text{proposed}}}{T_{\text{HM13.0}}} 100\% \quad (18)$$

其中, T_{proposed} 为本文算法编码时间; $T_{\text{HM13.0}}$ 为 HM 13.0 在同环境下测得的编码时间。实验结果见表 5。可见,本文算法在固定 QP 编码和多 QP 优化编码两种测试条件下,复杂度分别平均增加了 3% 和 2%。

为了对比其他使用 SSIM 进行 RDO 的算法,本文算法分别在 Yeo 等算法^[13]和 Qi 等算法^[14]相同的配置条件下进行实现。表 6 给出了本文算法在全 I 帧配置下对比 Qi 等算法的率失真结果。由于 Yeo 等算法只给出了随机访问(RA)配置下的测试结果,因此,本文简单的将本文算法应用到帧间编码,和 Yeo 等算法一致。本文测试了 RA 配置下多 QP 优化的结果,见表 7。可以看出,尽管是简单的应用到帧间编码,本文算法也表现出优异的性能。

表 5 本文算法对比 HM 13.0 编码复杂度

Tab. 5 Encoding complexity results of comparison with HM 13.0

Unit: %

Sequence	Fix QP	Encoding complexity			
		$dqd0$	$dqd1$	$dqd2$	$dqd3$
Traffic	102	102	95	96	98
PeopleOnStreet	102	103	103	103	103
Cactus	104	102	105	103	103
BQTerrace	102	101	102	102	103
BasketballDrill	105	104	104	104	98
PartyScene	103	103	102	102	102
BQSquare	105	102	104	101	103
BlowingBubbles	103	102	104	101	103
FourPeople	104	103	105	104	102
KristenAndSara	104	104	105	104	104
ChinaSpeed	103	103	102	103	103
SlideShow	103	104	103	104	105
Average	103	103	103	102	102

表 6 本文算法对比 Qi 等算法的 BD -RateTab. 6 BD -Rate results of comparison with the method proposed by Qi et al.

Unit: %

Sequence	Proposed	Qi et al. ^[13]
Kimono	-3.0	-0.8
ParkScene	-5.3	-1.6
Cactus	-6.1	-2.2
BasketballDrill	-8.7	-1.7
BQMall	-9.0	-2.3
RaceHorses	-6.7	-2.1
FourPeople	-8.2	-3.7
Johnny	-5.6	-2.7
Average	-6.6	-2.1

表 7 本文算法对比 Yeo 等算法的 BD -RateTab. 7 BD -Rate results of comparison with the method proposed by Yeo et al.

Unit: %

Sequence	Proposed				Yeo et al. ^[13]			
	$dqd0$	$dqd1$	$dqd2$	$dqd3$	$dqd0$	$dqd1$	$dqd2$	$dqd3$
Traffic	-5.7	-5.3	-5.2	-4.4	-9.6	-11.0	-10.6	-9.6
PeopleOnStreet	-10.6	-11.4	-11.2	-10.4	-5.8	-7.5	-7.9	-7.0
Cactus	-5.5	-5.7	-4.9	-4.4	-0.6	-0.9	-0.5	0.0
BQTerrace	-13.0	-13.6	-13.9	-13.6	-3.5	-6.2	-7.3	-7.2
BasketballDrill	-8.7	-8.6	-9.2	-8.6	-20.0	-19.9	-19.1	-18.1
PartyScene	-3.8	-5.1	-6.0	-7.3	-2.3	-5.5	-6.8	-5.3
BQSquare	-16.4	-17.8	-19.9	-20.7	-10.7	-17.9	-21.7	-21.9
BlowingBubbles	-3.6	-4.8	-6.7	-5.2	0.2	-2.8	-5.7	-4.0
ChinaSpeed	-17.0	-17.1	-17.2	-17.5	-1.4	-2.4	-3.1	-3.2
SlideShow	-16.2	-16.3	-17.2	-14.8	-5.5	-7.5	-7.9	-7.8
Average	-10.0	-10.6	-11.1	-10.7	-5.9	-8.2	-9.1	-8.4

4 结 论

针对基于 SSE 失真的传统 RDO 无法得到最优的视觉感知质量问题,本文将基于 SSIM 的失真度量用于 HEVC 帧内编码 RDO 中,提出了一种针对 HEVC CTU 使用 SSIM 度量失真的具体计算方法。首先分析了由量化器负责的 R- Δ 模型;然后根据统计分析,提出了一个 D_{SSIM} -QP 模型;最后使用多 QP 优化的方法求取模型的最终参数。实验表明,与传统的 HEVC RDO 相比,在相同的码率下,本文算法可获得更好的主观视觉质量,同时编码复杂度平均只增加了 3% 左右。进一步的研究工作是如何将本文算法扩展到 HEVC 帧间率失真优化中。

参考文献:

- [1] Sullivan G J, Ohm J, Han W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.
- [2] Wang Z, Bovik A C. Mean squared error; love it or leave it? A new look at signal fidelity measures[J]. IEEE Signal Processing Magazine, 2009, 26(1): 98-117.
- [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] JIN Xin, JIANG Gang-yi, CHEN Fen, et al. Adaptive image quality assessment method based on structural similarity[J]. Journal of Optoelectronics • Laser, 2014, 25(2): 378-385.
靳鑫, 蒋刚毅, 陈芬, 等. 基于结构相似度的自适应图像质量评价[J]. 光电子·激光, 2014, 25(2): 378-385.
- [5] MAI Zhi-yi, YANG Chun-ling, PO Lai-man, et al. A new rate-distortion optimization using structural information in H. 264 I-frame encoder; Advanced concepts for intelligent vision systems, Berlin: Springer, 2005, 435-441.
- [6] HUANG Yi-hsin, OU Tao-sheng, SU Po-yen, 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.
- [7] YANG Chun-ling, 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.
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.
- [8] Wang S Q, Rehman A, Wang Z, et al. Rate-SSIM optimization for video coding[A]. Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)[C]. 2011: 833-836.
- [9] Wang S Q, Rehman A, Wang Z, et al. SSIM-motivated rate distortion optimization for video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(4): 516-529.
- [10] Wang S Q, Rehman A, Wang Zhou, et al. Perceptual video coding based on SSIM-inspired divisive normalization[J]. IEEE Transactions on Image Processing, 2013, 22(4): 1418-1429.
- [11] Rehman A, Wang Zhou. SSIM-inspired perceptual video coding for HEVC[A]. Proc. of IEEE International Conference on Multimedia and Expo (ICME)[C]. 2012, 497-502.
- [12] YEO Chuo-hao, TAN Hui-li, TAN Yih-han. On rate distortion optimization using SSIM[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(7): 1170-1181.
- [13] YEO Chuo-hao, TAN Hui-li, TAN Yih-han. SSIM-based adaptive quantization in HEVC[A]. Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)[C]. 2013, 1690-1694.
- [14] QI Ji, LI Xiao-yu, SU Fan, et al. Efficient rate-distortion optimization for HEVC using SSIM and motion homogeneity[A]. Proc. of Picture Coding Symposium (PCS), 2013, 217-220.
- [15] CHEN Yun-shan, GAO Hui-bin, SU Wan-xin, et al. Application of structural similarity in inter mode selection for AVS[J]. Journal of Optoelectronics • Laser, 2011, 22(3): 435-439.
陈云善, 高慧斌, 苏宛新, 等. 结构相似度在 AVS 帧间模式选择中的应用[J]. 光电子·激光, 2011, 22(3): 435-439.
- [16] Zhao T, Zeng K, Rehman A, et al. On the use of SSIM in HEVC[A]. Proc. of IEEE Asilomar Conference on Signals, Systems and Computers (ACSSC)[C]. 2013, 1107-1111.
- [17] Kim I K, McCann K, Sugimoto K, et al. High efficiency video Coding (HEVC) test model 13 (HM 13) encoder description[A]. Proc. of 15th Joint Collaborative Team on Video Coding (JCT-VC) Meeting, JCTVC-O1002[C]. 2013.
- [18] Channappayya S S, Bovik A C, Heath R W, et al. Rate bounds on SSIM index of quantized images[J]. IEEE Transactions on Image Processing, 2008, 17(9): 1624-1639.
- [19] Bossen F. Common test conditions and software reference configurations[A]. Proc. of 12th Joint Collaborative Team on Video Coding (JCT-VC) Meeting, JCTVC-L1100[C]. 2013.
- [20] Bjontegaard G. Calculation of average PSNR differences between RD curves[A]. Proc. of 13th Video Coding Experts Group (VCEG) Meeting, VCEG-M33[C]. 2001.

作者简介:

蒋刚毅 (1964—), 男, 浙江绍兴人, 教授, 博士生导师, 主要从事多媒体信息传输与信息安全、图像处理与视频信号编码方法的研究。