



研究与开发

基于显著特征的 HEVC 帧内率失真优化

李金龙, 陈芬, 彭宗举, 蒋刚毅, 郁梅

(宁波大学信息科学与工程学院, 浙江 宁波 315211)

摘要: 引入感知因素的 RDO 算法和感知编码可以提高编码视频主观质量, 但也产生了明显的率失真性能损失。针对上述问题, 提出了一种基于显著特征的 RDO 方法。首先探索了 HEVC 编码树单元的码率分布和显著特征之间的相关性, 提出显著码率先验模型并确定了模型的初始参数; 然后建立了码率差异最小化的率失真代价模型, 并根据该模型推导得到自适应拉格朗日乘子计算方法; 最后设计了模型参数更新规则。实验结果表明, 算法在率失真性能基本不变的情况下, 提高了视频的主观质量。

关键词: 高性能视频编码; 率失真优化; 拉格朗日乘子; 显著特性

中图分类号: TN919.81

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017250

Rate distortion optimization based on saliency characteristic for HEVC intra coding

LI Jinlong, CHEN Fen, PENG Zongju, JIANG Gangyi, YU Mei

College of Information Science and Engineering, Ningbo University, Ningbo 315211, China

Abstract: Although the perception based rate distortion optimization (RDO) and coding algorithms can improve subjective quality, they significantly decrease rate distortion performance. For the above questions, a saliency-based RDO method was proposed. Firstly, according to the relation between rate distribution and saliency features of coding tree unit in HEVC, a priori saliency-rate model was designed. Furthermore, the initial parameters of the model were determined. Secondly, the rate difference minimization rate distortion model was proposed, and then a calculation method of the adaptive Lagrangian multiplier was derived from the model. Finally, update mechanism of model parameters was designed. Experimental results show that the proposed algorithm can improve the subjective quality while it maintains rate distortion performance.

Key words: high efficiency video coding, rate distortion optimization, Lagrangian multiplier, saliency characteristic

收稿日期: 2017-04-17; 修回日期: 2017-08-18

基金项目: 国家自然科学基金重点国际(地区)合作研究项目(No.61620106012); 国家自然科学基金资助项目(No.61771269); 国家自然科学基金联合基金(NSFC-广东联合基金)资助项目(No.U1301257); 浙江省自然科学基金资助项目(No.LY16F010002, No.LY15F010005, No.LY17F010005)

Foundation Items: The National Natural Science Foundation of Key International(Regional) Cooperative Research Project of China (No.61620106012), The National Natural Science Foundation of China (No.61771269), The National Natural Science Foundation of China (NSFC-Guangdong Joint Foundation) (No.U1301257), Zhejiang Provincial Natural Science Foundation of China (No.LY16F010002, No.LY15F010005, No.LY17F010005)

1 引言

随着人们对视频服务的需求日益增长,传统视频编码标准 H.264/AVC 已无法应对新的视频编码要求^[1]。高性能视频编码 (high efficiency video coding, HEVC)^[2]中使用了更先进的编码技术,具有更高的压缩性能^[3]。率失真优化 (rate distortion optimization, RDO) 技术是 HEVC 的关键技术之一,它能够实现码率和失真的平衡,从而获得最优的编码性能。HEVC 中以传统信号处理方法如绝对误差和、平方和误差 (sum of the squared error, SSE) 等度量编码失真。这些方法简单易行,但与人类主观视觉感受并不完全吻合^[4]。在 HEVC 中结合视觉感知因素,可以进一步提高压缩视频的主观质量。有学者结合人类视觉因素提出了感知编码 (perceptual video coding, PVC) 方法。Bae 等人^[5]结合 HEVC 的编码树结构特点,建立局部失真检测概率模型,用以计算任意尺度 DCT 变换块的失真可见度,然后抑制 DCT 系数幅值以滤除重建视频中部分内容,从而达到提高压缩效率的目的。Wei 等人^[6]直接根据编码单元 (code unit, CU) 的显著值调整其量化参数 (quantization parameter, QP),在非显著区域中减少码率以提升编码性能。

在 RDO 中引入感知因素是提高 HEVC 编码性能的另外一种方式。RDO 通过为编码器确定最佳的编码参数来获得最优的率失真性能。拉格朗日乘子 λ 的求取是 RDO 的关键,它可将码率和失真相互约束关系转为不受限的极值问题^[7]。HEVC 沿用 H.264/AVC 中的 λ - q 模型^[8]和 QP 模型^[9]来解决 λ 的求取问题。近年来,关于 RDO 的研究有了更多进展。Li 等人^[10]研究了 QP 与 λ 的关系后提出了一种多 QP 编码方式,获得了更好的率失真性能。Li 等人^[11]根据 CU 间的依赖性自适应地调整 λ ,提高了率失真性能。Huang 等人^[12]提出了一种更有效的分析方法,直接解决了熵编码的 RDO

问题,并通过离线训练建立查询表,实现了高效快速的基于 RDO 的量化。这些方法^[10-12]仍然以传统的方法度量失真,因此编码输出的视频不能反映人类视觉感知特性^[4]。

目前,有学者将感知因素引入 RDO 中,提出了基于感知的率失真优化 (perceptual RDO, PRDO) 方法。Huang 等人^[13]提出了一种结合结构相似度指标 (structural similarity index, SSIM) 的 RDO 方法,在 SSIM 相同情况下进一步降低了码率。Yeo 等人^[14]利用 YUV 分量的 1-SSIM 值加权来度量失真,Wang 等人^[15]用 SSIM 的倒数来度量失真,他们分别给出了 λ 的计算方法并建立了 PRDO 模型。Zhu 等人^[16]提出了一种结合对比度和运动掩蔽效应的空间质量指标作为失真度量方法,据此建立了相应的 PRDO 模型。Yang 等人^[17]根据视频的时空域特征调整 λ 值,提出相应的 PRDO 算法。Cen 等人^[18]以最小变换单元为基本单位建立了基于 MS-SSIM 的失真度量模型,并推导了 λ 的计算方法。这些引入了感知因素的 PVC^[4-6]或 RDO 方法^[12-18]在相同视觉质量下获得了更好的压缩效率。然而,这些方法也存在计算复杂高、失真衡量指标含义模糊以及视频质量难以定量研究等问题,如参考文献[5]和参考文献[16]中需要相应的主观评价指标和主观打分的方式来确定编码视频的质量。同时,完全摒弃传统客观评价方法使得编码视频客观率失真性能下降,从而导致针对视频压缩性能的传统研究方法失去作用。Zhao 等人^[19]提出了一种结合 SSIM 和 SSE 的 RDO 方法,该方法在 PRDO 方法中结合传统失真度量方法减少了客观率失真性能的下降,然而其客观率失真性能的损失仍然无法忽略。

针对上述问题,本文提出了一种结合显著特征的 RDO 方法,在客观率失真性能基本不受影响的情况下,引入显著特征对编码树单元 (coding tree unit, CTU) 的 λ 进行自适应调整,进一步提高其主观质量。信息最大化注意 (attention based



on information maximization, AIM) 模型^[20]是一种单纯从数学计算角度建立的视觉注意模型, 其建立过程中以自然图像的眼动数据作为训练样本, 具有基于自信息熵的严格数学定义, 因此该指标含义清晰, 也能够在一定程度上反映人眼的感知特征。鉴于上述分析, 本文的显著计算模型采用 AIM 模型, 首先分析了 CTU 的原始码率分布和显著特征的吻合度, 通过拟合方法建立显著码率先验模型, 并根据拟合结果确定模型的初始参数值。然后以修正 CTU 码率分布为目的, 结合显著码率先验模型建立码率差异最小化的 RDO 模型, 在此基础上推导得到自适应 λ 乘子的计算方法。最后建立了相应的模型参数更新机制, 实现模型参数的自适应修正。实验结果表明, 本文提出的算法在客观率失真性能基本没有损失前提下, 提高了重建视频的主观质量, 在低码率配置下提升幅度更为明显。

2 引入显著特征的率失真优化

本文算法整体框架如图 1 所示, 包括预编码、码率差异最小化 RDO、编码和参数更新等步骤。引入预编码的目的是进行显著特征和率失真分析。本文仅对帧内编码帧 (intra-frame, I-Frame) 进行预编码, 采用 AIM 模型计算显著图。先验数据是指原始编码平台编码的输出。后文将详细阐述显著先验模型、码率差异最小化 RDO 模型和参数更新修正这 3 个过程。

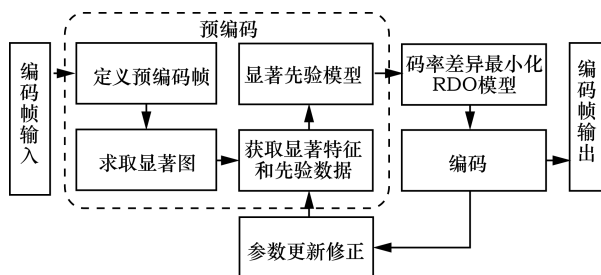


图 1 本文算法整体框架

2.1 显著码率先验模型

本文选取 Bruce 等人^[20]提出基于香农信息

论的 AIM 模型作为显著计算方法, 该模型计算了图像局部相对于周围区域为眼动点的概率, 计算概率的自信息熵作为显著值。AIM 的求取主要分为 3 个步骤: 首先利用眼动仪对大量自然图像进行测试, 获得图片中的眼动点, 并且对眼动点处图像块进行独立成分分析, 得到一系列基函数; 然后利用基函数计算目标图片块的基系数, 并对基系数进行概率密度估计; 最后计算每个图像块的联合概率密度和自信息熵作为显著度。

求取编码帧的显著图之后, 探索显著特征与原始编码码率的相关性, 通过拟合的方法建立了显著码率模型, 并根据拟合结果确定模型初始参数值。具体过程为: 在固定 QP 情况下, 以 CTU 为单位统计编码码率; 将显著图的边界填充到 CTU 尺寸的整数倍, 计算 64×64 像素块的显著均值并进行归一化得到 s_i , 填充边界像素去除计数; 使用统计方法拟合显著值和原始编码码率, 得到显著码率先验模型为:

$$r_i = a \times s_i^b + 1 - a \quad (1)$$

其中, r_i 为第 i 个 CTU 的归一化先验码率。

原始编码帧中 CTU 的编码结果一般难以符合显著分布, 通过修正 CTU 的编码过程可使得码率分布更符合显著分布。为了尽量减少需要编码修正的 CTU, 采用最小二乘法拟合的方法建立式 (1), 并将其定义为先验模型。偏离先验模型的 CTU 将会得到修正, 符合先验模型则不会修正。为使拟合模型更合理, 通过大量的先期实验对编码结果进行分析, 考虑了模型的顽健性和复杂程度, 选取幂函数模型进行拟合, 通过常数项将 r_i 限定在 $(0, 1]$ 的范围内, 以实现归一化。

在 HEVC 测试模型 (HEVC test model, HM) 中进行编码测试, 得到各 CTU 的码率。然后分析原始编码码率和先验码率的关系, 结果见表 1。相关系数 R^2 表示原始编码方法的码率与先验码

率的吻合度。由表 1 可以看出, 大部分序列的相关性较小。这种现象反映了原始编码平台没有考虑到视频内容的显著性分布。本文的 RDO 算法中, 相关系数 R^2 作为反馈来对编码进行修正。 R^2 的值越接近 1, 实际编码过程的修正幅度越小; R^2 的值越接近 0, 实际编码过程的修正幅度越大。

表 1 原始编码码率和先验码率的关系

序列	帧编号	相关系数 R^2	a	b
ChinaSpeed	1	0.54	0.99	2.47
	2	0.56	1.00	2.39
	3	0.56	1.00	2.39
FourPeople	1	0.42	0.98	2.65
	2	0.42	0.98	2.58
	3	0.42	0.98	2.64
Johnny	1	0.62	0.96	2.63
	2	0.65	0.96	2.76
	3	0.65	0.98	2.60
BasketballDrill	1	0.18	0.93	2.65
	2	0.18	0.93	2.65
	3	0.18	0.93	2.58
均值	-	0.45	0.97	2.60

在表 1 中, a 和 b 为模型拟合后的参数值, 参数 b 影响了先验模型曲线的斜率变化, 通过修正 b 值可以使先验数据均匀分布在曲线两侧, 使得偏离先验模型的 CTU 得到对称的编码修正; 参数 a 使得曲线上限为 1, 令 a 的值接近且小于 1, 保证 r_i 取值范围为(0,1]。由于参数 a 和 b 需要自适应地修正, 表 1 中列出的参数 a 和 b 的值仅作为初始参数值。

2.2 码率差异最小化率失真模型

在显著码率先验模型基础上定义码率差异最小化率失真模型, 以显著码率先验模型为参考, 利用 RDO 技术重新调整编码帧的码率分布, 使之更符合显著分布。码率差异是指 CTU 的先验码率和修正后码率的差异, 码率差异最小化率失真代价函数定义为:

$$J_i = D_i + \lambda' R_i + c \times \Delta R_i \quad (2)$$

其中, c 为加权系数, ΔR_i 为显著特征因子, 计算式为:

$$\Delta R_i = |R_i - R_{si}| \quad (3)$$

其含义是先验码率 R_{si} 和修正后码率 R_i 之差的绝对值, R_{si} 定义为:

$$R_{si} = r_i \times R_{\max} \quad (4)$$

其中, $R_{\max}$ 为预编码帧中的最大的原始码率。

在当前 CTU 编码之前, 实际码率是无法获取的, 令 R_i 为变量, 则式 (3) 转为:

$$\Delta R_i = \begin{cases} R_i - R_{si}, & R_i \geq R_{si} \\ R_{si} - R_i, & R_i < R_{si} \end{cases} \quad (5)$$

式 (5) 的显著因子转化为阈值问题, 对两种情况各自进行简化合并, 求取拉格朗日乘子。满足高斯信源分布的率失真函数一般有如下关系^[21]:

$$R = \frac{1}{2} \text{lb}(\frac{\delta^2}{D}) \quad (6)$$

结合式 (3) ~ 式 (6) 可得:

$$\Delta R_i = \begin{cases} \text{lb}(\frac{\delta_i}{\delta_k^r}) + \frac{1}{2} \text{lb} D_i - \frac{1}{2} r_i \cdot \text{lb} D_k, & R_i \geq R_{si} \\ \text{lb}(\frac{\delta_i}{\delta_k^r}) + \frac{1}{2} r_i \cdot \text{lb} D_k - \frac{1}{2} \text{lb} D_i, & R_i < R_{si} \end{cases} \quad (7)$$

其中, k 是先验数据中失真最大的 CTU 的索引。对函数 $\text{lb}x$ 在 $x_0=2$ 处泰勒展开, 保留一阶项得到:

$$\Delta R_i = \begin{cases} \text{lb}(\frac{\delta_i}{\delta_k^r}) + \frac{D_i - r_i \cdot D_k}{4 \ln 2} + \frac{1 - \ln 2}{2 \ln 2} (1 - r_i), & R_i \geq R_{si} \\ \text{lb}(\frac{\delta_k}{\delta_i^r}) + \frac{r_i \cdot D_k - D_i}{4 \ln 2} + \frac{1 - \ln 2}{2 \ln 2} (r_i - 1), & R_i < R_{si} \end{cases} \quad (8)$$

将式 (8) 代入式 (2), 等式两边对 R_i 求导, 并且已知:

$$\lambda_{\text{HM}} = -\frac{dD_i}{dR_i} \quad (9)$$

综合式 (1)、式 (2)、式 (8) 和式 (9), 得



到拉格朗日自适应因子为:

$$\lambda' = \begin{cases} \lambda_{\text{HM}} \left[\frac{4 \ln 2 + ac(1 - s_i^b)}{4 \ln 2} \right], & R_i \geq R_{si} \\ \lambda_{\text{HM}} \left[\frac{4 \ln 2 + ac(s_i^b - 1)}{4 \ln 2} \right], & R_i < R_{si} \end{cases} \quad (10)$$

式(10)表明,本文基于显著特征的RDO模型最终转化为自适应的拉格朗日乘子模型,模型中包含一个阈值问题,显著码率先验模型求得的先验码率 R_{si} 为阈值。

2.3 模型的参数分析

式(1)定义为先验模型,通过表1得到的拟合参数仅为其初始参数值。结合式(10)可知,通过式(1)输出的先验码率 R_{si} 直接影响了 λ 的调整方式。定义阈值分布参数 ε 为:

$$\varepsilon = \frac{N_{R_i > R_{si}}}{N_{\text{CTU}}} \quad (11)$$

其中, N_{CTU} 为预编码帧中CTU总数目, $N_{R_i > R_{si}}$ 为先验数据中的码率大于预测码率的CTU个数, ε 的实际含义为式(10)中落在阈值一侧的比例。

为保证模型参数得到自适应修正,令 $\varepsilon=0.5$,使得式(10)中的阈值范围总是各占一半的比例,从而使得分布在曲线两侧的CTU得到对称修正,减少编码修正对客观率失真性能的影响。式(1)中参数 b 的值对阈值分布影响最为明显,对参数 b 进行修正以满足 ε 的取值要求。算法中设计了以下参数修正步骤。

步骤1 根据表1列出的拟合结果,设置参数 b 的初始值 $b_0=2.6$ 并令 $b=b_0$,参数 $a=0.97$,其中参数 a 不做修正。

步骤2 将 b 和 a 代入式(1)并结合预编码的码率分布计算 ε 的值。根据 $\varepsilon-0.5$ 的符号对 b 的值进行修正,修正步长 $\text{Step}_b=0.01$ 。若 $\varepsilon-0.5>0$,则令 $b'=b-\text{Step}_b$;否则令 $b'=b+\text{Step}_b$ 。

步骤3 令 $b=b'$,将 b 和 a 重新代入式(1)计算 ε 的值。当满足条件 $|\varepsilon-0.05|<\text{Step}_b$ 时,结束参

数修正;否则重复步骤2。为避免修正过程出现发散或震荡情况,设置总循环次数 N_{for} ,当循环次数达到 N_{for} 时,强制结束修正过程, $N_{\text{for}}=b_0/\text{Step}_b$ 。

式(2)中引入的参数 c 经实验验证取 $c=1$ 。

3 实验结果及分析

为了验证本文算法的有效性,在HEVC参考平台HM16.0上实现并测试算法。所有测试序列来自JCT-VC推荐的公共测试条件(CTC)^[22]。编码配置为全帧内编码,分别测试了37、32、27和22共4组QP值,失真度量参数包括峰值信噪比(peak signal to noise ratio, PSNR)、多尺度结构相似度(multi-scale SSIM, MSSSIM)^[23]以及信息加权结构相似度(information content weighting SSIM, IWSSIM)^[24]。其中PSNR是传统的客观失真度量指标,MSSSIM和IWSSIM是与人类视觉主观感受具有较好相关性的评价指标。以编码平台HM为参考,分别计算3种失真评价指标下的码率节省和失真变化,以此代表不同失真指标下算法的率失真性能。为证明阈值分布参数 ε 的有效性,设置仅第一帧进行参数修正,后续帧使用第一帧所得模型,实验结果见表2。BD-BR为在相同质量情况下的码率变化,BD-PSNR、BD-MSSSIM和BD-IWSSIM分别反映在相同码率的情况下PSNR、MSSSIM和IWSSIM的变化^[25]。

由表2可知,本文RDO算法对编码视频的客观率失真性能基本没有影响,BD-PSNR平均仅下降了0.00276dB,主观率失真性能均得到一定程度的提升。在码率一致的情况下,MSSSIM值和IWSSIM值分别提升了0.00013和0.00012;MSSSIM值和IWSSIM值均一致情况下平均节约1.126%和0.628%的码率。其中BQTerrace和ChinaSpeed的主观特性提升比较明显,前者是因为场景变化缓慢,而后者是因为视频内容的显著区域突出。序列Kimono和PeopleOnStreet的主客观率失真性能均未出现明显提升,且整体性能提

表 2 3 种失真指标下的率失真性能与 HM16.0 的对比

序列	PSNR		MSSSIM		IWSSIM	
	BD-BR	BD-PSNR/dB	BD-BR	BD-MSSSIM	BD-BR	BD-IWSSIM
BasketballDrill	0.979 62%	-0.046 23	-0.284 14%	-0.000 03	-0.984 15%	0.000 03
BQTerrace	0.698 06%	-0.042 12	-0.781 17%	0.000 12	-1.111 59%	0.000 25
ChinaSpeed	0.461 73%	-0.043 85	-3.317 88%	0.000 44	-1.780 13%	0.000 32
Johnny	-1.121 25%	0.090 23	-2.098 99%	0.000 18	-0.652 26%	0.000 22
Kimono	-0.871 95%	0.033 08	0.330 44%	-0.000 04	0.726 85%	-0.000 08
PeopleOnStreet	0.112 59%	-0.007 65	-0.605 39%	0.000 10	0.035 37%	-0.000 02
均值	0.043 13%	-0.002 76	-1.126 19%	0.000 13	-0.627 65%	0.000 12

升不明显,原因是后续帧编码时模型参数无法得到及时修正。设置所有编码帧都进行模型参数修正,实验结果见表 3。

由表 3 可知,在参数及时修正情况下算法性能整体得到提升。客观率失真性能平均只下降了 0.019 01 dB。基于 MSSSIM 和 IWSSIM 的主观率失真性能均得到明显提升,码率一致时 MSSSIM 值和 IWSSIM 值分别提升了 0.000 7 和 0.001 00, MSSSIM 值和 IWSSIM 值均一致时平均节省约 3.87% 和 4.76% 的码率。其中 ChinaSpeed、FourPeople 和 BasketballDrill 等序列表现较好,是因为背景相对复杂,客观率失真性能较差,通过修正码率分布可有效提高主观质量。Johnny 序列视频内容前景显著背景简单的特点,具有较高的客观率失真性能,已经具有较好的视觉质量,主

观质量难以进一步提升; BasketballPass 序列在第 40 帧后开始发生剧烈的场景变化,主观质量难以得到提升。表 3 相对表 2 所有序列指标均有明显提升,证明了参数阈值分布参数 ε 的有效性。

图 2 为采用本文方法和平台 HM 对序列 BasketballPass 编码时,基于 IWSSIM 指标的率失真性能对比。由图 2 可知,高码率配置下压缩失真较小,算法对主观质量的提升有限。这是因为较高客观质量的视频也能够提供良好的视觉效果。低码率配置下压缩失真较大,通过调整码率分布可提高主观质量。

为了进一步说明本文算法与 PRDO 和 PVC 算法的区别,本文使用参考文献[5]和参考文献[16]中提供的部分数据进行对比。对于参考文献[5],则采用低时延 P 帧编码配置, I 帧定义为预编码,

表 3 参数实时修正时 3 种失真指标下本文算法与 HM16.0 的率失真性能对比

序列	PSNR		MSSSIM		IWSSIM	
	BD-BR	BD-PSNR/dB	BD-BR	BD-MSSSIM	BD-BR	BD-IWSSIM
BasketballDrill	1.230 00%	-0.057 75	-4.995 28%	0.000 85	-5.891 22%	0.000 91
BasketballPass	-0.010 00%	0.002 88	-1.771 78%	0.000 38	-2.781 62%	0.000 54
BQMall	0.253 55%	-0.017 00	-3.857 11%	0.000 62	-4.294 17%	0.000 65
BQTerrace	0.246 83%	-0.015 15	-4.112 38%	0.000 84	-4.163 53%	0.000 79
ChinaSpeed	0.480 53%	-0.045 32	-4.688 16%	0.000 93	-6.332 35%	0.002 09
FourPeople	0.375 45%	-0.021 53	-6.138 77%	0.001 53	-8.102 50%	0.002 52
Johnny	-0.099 46%	0.004 62	-2.170 08%	0.000 19	-2.310 52%	0.000 36
Kimono	0.025 86%	-0.000 87	-2.067 25%	0.000 18	-3.861 06%	0.000 46
RaceHorseC	0.288 20%	-0.021 01	-4.994 85%	0.000 81	-5.117 75%	0.000 69
均值	0.310 11%	-0.019 01	-3.866 18%	0.000 70	-4.761 63%	0.001 00

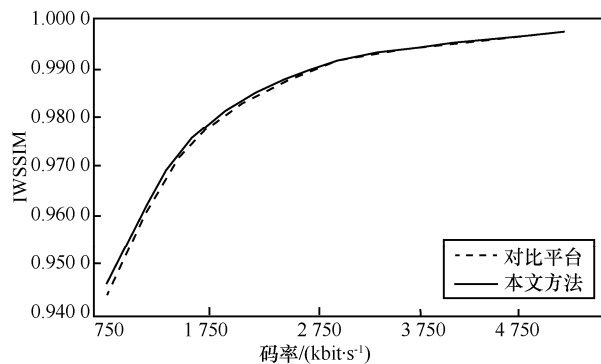


图2 序列 BasketballPass 的 IWSSIM 率失真性能对比

获取模型参数, P 帧使用预编码帧所得模型。对于参考文献[16], 采用低时延 B 帧编码配置, 其余与前者一致。由于本文所提取的显著特征为空域特征, 因此将其直接使用在帧间编码会有性能损失。

表4中的BD-BR和BD-PSNR是以参考文献[5]

中提供的数据作为参考计算而来。由表4可知, 以典型的PVC算法为参考, 本文的算法具有更高的客观率失真性能, 说明参考文献[5]在获得较高的主观率失真性能时损失了较多的客观率失真性能。对于序列中如BQTerrace和RaceHorseC的客观率失真性能反而出现下降, 这是因为本文算法应用于帧间编码时, 后续P帧编码过程中模型参数无法及时修正, 导致客观率失真性能下降。

表5为本文方法与参考文献[16]的客观率失真性能对比, 对比数据由参考文献[16]提供, $\Delta PSNR$ 为本文编码输出与对比数据中的PSNR之差。由表4可知, 在码率基本一致情况下, 本文的RDO算法能够保持更高的客观率失真性能。部分序列如BasketballDrill、BasketballDrive和BQTerrace客观质量处于与参考文献[16]相当水

表4 本文方法与参考文献[5]方法的PSNR率失真性能比较

序列	QP	参考文献[5]方法		本文方法		BD-BR	BD-PSNR/dB
		PSNR/dB	码率/(kbit·s ⁻¹)	PSNR/dB	码率/(kbit·s ⁻¹)		
BQTerrace	22	35.70	16 788	38.50	59 218	10.55%	-0.20
	27	34.77	5 273	34.93	7 260		
	32	33.37	1 918	33.01	1 778		
	37	31.36	760	30.78	800		
Cactus	22	37.35	12 483	38.43	21 384	-7.56%	0.15
	27	36.21	5 367	36.34	5 628		
	32	34.47	2 551	34.38	2 314		
	37	32.25	1 267	32.07	1 093		
ParkSence	22	38.14	6 280	39.72	7 596	-19.79%	0.64
	27	36.43	2 979	37.00	3 009		
	32	34.13	1 323	34.50	1 227		
	37	31.59	576	31.95	508		
BQMall	22	37.61	3 292	39.46	4 055	-10.13%	0.44
	27	36.25	1 758	36.70	1 711		
	32	34.09	895	33.77	787		
	37	31.37	457	30.89	385		
PartySence	22	33.91	5 793	38.19	6 403	-13.19%	0.82
	27	32.97	3 111	34.06	3 142		
	32	30.8	1 484	30.65	1 358		
	37	27.8	643	27.28	552		
RaceHorseC	22	36.59	3 732	39.56	4 853	5.68%	-0.03
	27	35.27	2 008	35.96	2 076		
	32	32.93	985	32.04	953		
	37	30.16	457	29.07	417		
均值		—	—	—	—	-5.74%	0.30

表 5 本文方法与参考文献[16]方法的 PSNR 率失真性能比较

序列	参考文献[16]方法		本文方法		$\Delta PSNR/dB$
	码率/(kbit·s ⁻¹)	PSNR/dB	码率/(kbit·s ⁻¹)	PSNR/dB	
BasketballPass	360.02	32.77	345.36	32.88	0.11
BQSquare	256.19	29.03	252.59	29.36	0.33
RaceHorseD	166.03	29.66	164.78	29.89	0.23
RaceHorseC	537.08	30.57	522.34	30.68	0.11
BasketballDrill	703.93	33.24	704.24	33.19	-0.05
BasketballDrive	1 499.78	32.81	1 465.97	32.87	0.06
Cactus	1 200.12	31.72	1 269.28	31.93	0.21
BQTerrace	599.92	30.12	592.98	30.13	0.01
均值	—	—	—	—	0.07

平, 这是因为本文算法应用于帧间编码时, 后续 B 帧编码过程中模型参数无法及时修正, 会出现一定程度的客观率失真性能下降。表 4 和表 5 的对比结果表明, 本文算法能使编码器具有更高传统客观率失真性能。

图 3 和图 4 为本文算法与参考平台 HM 实际编码结果的局部对比, 编码参数 QP 设置为 37。对于图 3 和图 4, 本文算法与 HM 对比分别只有 0.04 dB 和 0.11 dB 的 PSNR 下降, 而基于 IWSSIM 的率失真性能都有提升。在相对比较重要的人脸

区域中, 图 3 中 HM16.0 的编码结果出现明显的大方块效应, 而本文的编码结果则避免了这种方块效应, 且左上角处的细节保留的比 HM16.0 要多; 图 4 中 HM16.0 的编码结果存在较多的色差斑块, 而本文算法的编码结果则避免了这种现象。实际编码对比结果表明, 本文算法能够提高编码帧显著区域的视觉效果, 且对传统的客观率失真性能影响很少。



图 3 序列 RaceHorseC 的编码结果对比



图 4 序列 FourPeople 的编码结果对比



4 结束语

针对目前 HEVC 中典型的 PRDO 或 PVC 算法会产生明显的客观率失真性能损失这一问题, 本文在客观率失真指标的框架下通过引入显著特征建立了新的 RDO 方法。本文算法与传统 RDO 算法相比, 在客观率失真性能基本不变情况下提高了压缩视频的主观视觉特性, 在低码率配置下效果更为明显。对比其他典型 PRDO 算法和 PVC 算法, 客观率失真性能基本不受影响, 因而针对视频压缩失真的传统研究方法仍然是有效的。后续工作主要有两个方面, 其一是进一步探索在客观率失真性能不变情况下主观质量提升幅度的上限, 尤其是在低码率情况下, 提升视频主观特性极具意义; 其二是将引入的空域特性推广到时域特性, 从而更好地应用到帧间编码。

参考文献:

- [1] 赵翠, 张诚, 欧阳峰. 有线电视网络的发展现状与趋势[J]. 电信科学, 2017, 33(6): 146-153.
ZHAO C, ZHANG C, OUYANG F. Development status and trends of cable TV networks[J]. Telecommunications Science, 2017, 33(6): 146-153.
- [2] 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.
- [3] 姚英彪, 李晓娟. 基于图像空间相关性与纹理的 HEVC 块划分快速算法[J]. 电信科学, 2015, 31(1): 38-46.
YAO Y B, LI X J. Fast block partitioning algorithm for HEVC based on spatial correlation and image texture[J]. Telecommunications Science, 2015, 31(1): 38-46.
- [4] 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.
- [5] BAE S H, KIM J, KIM M. HEVC-based perceptually adaptive video coding using a DCT-based local distortion detection probability model[J]. IEEE Transactions on Image Processing, 2016, 25(7): 3343-3357.
- [6] WEI H, ZHOU X, ZHOU W, et al. Visual saliency based perceptual video coding in HEVC[C]//IEEE International Symposium on Circuits and Systems, May 22-25, 2016, Montreal, QC, Canada. New Jersey: IEEE Press, 2016: 2547-2550.
- [7] SHOHAM Y, GERSHO A. Efficient bit allocation for an arbitrary set of quantizers[J]. IEEE Transactions on Acoustics Speech and Signal Processing, 1988, 36(9): 1445-1453.
- [8] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H.264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [9] YANG E H, YU X. Rate distortion optimization for H.264 inter-frame coding: a general framework and algorithms[J]. IEEE Transactions on Image Processing, 2007, 16(7): 1774-1784.
- [10] LI B. QP determination by lambda value[C]//Document JCTVC-I0426, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 9th Meeting, April 27-May 7, 2012, Geneva, Switzerland. [S.l.:s.n.], 2012.
- [11] LI S, ZHU C, GAO Y. Lagrangian multiplier adaptation for rate-distortion optimization with inter-frame dependency[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 117-129.
- [12] HUANG T Y, CHEN H H. Efficient quantization based on rate-distortion optimization for video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(6): 1099-1106.
- [13] 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.
- [14] YEO C, TAN H L, TAN Y H. On rate distortion optimization using SSIM[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(7): 1170-1181.
- [15] WANG S, 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.
- [16] ZHU X, HONG W, XU H, et al. Spatial quality index based rate perceptual-distortion optimization for video coding[J]. Journal of Visual Communication and Image Representation, 2016, 38(C): 423-432.
- [17] YANG A, ZENG H, CHEN J, et al. Perceptual feature guided rate distortion optimization for high efficiency video coding[J]. Multidimensional Systems and Signal Processing, 2016: 1-18.
- [18] CEN F, LU Q, XU W. SSIM based rate-distortion optimization for intra-only coding in HEVC[C]//IEEE International Conference on Consumer Electronics, January 10-13, 2014, Las Vegas, NV, USA. New Jersey: IEEE Press, 2014: 17-18.
- [19] ZHAO P, LIU Y, LIU J, et al. Perceptual rate-distortion optimization for H.264/AVC video coding from both signal and vision perspectives[J]. Multimedia Tools and Applications, 2016, 75(5): 1-20.

- [20] BRUCE N D B. Saliency, attention and visual search: an information theoretic approach[J]. Journal of Vision, 2009, 9(3): 1-24.
- [21] GISH H, PIERCE J. Asymptotically efficient quantizing[J]. IEEE Transactions on Information Theory, 1968, 14(5): 676-683.
- [22] BOSSEN F. Common test conditions and software reference configurations[C]//Document JCTVC-G1200, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 7th Meeting, November 21-30, 2011, Geneva, Switzerland. [S.l.:s.n.], 2011.
- [23] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]// Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, November 9-12, 2003, Pacific Grove, CA, USA. New Jersey: IEEE Press, 2003:1398-1402.
- [24] WANG Z, LI Q. Information content weighting for perceptual image quality assessment[J]. IEEE Transactions on Image Processing, 2011, 20(5): 1185-1198.
- [25] BJONTEGAARD G. Calculation of average PSNR differences between RD-curves[C]//Austin: 13th Video Coding Experts Group (VCEG) Meeting, April 2-4, 2001, Austin, TX, USA. [S. l.: s.n.], 2001.

[作者简介]



李金龙(1993-), 男, 宁波大学信息科学与工程学院硕士生, 主要研究方向为视频编码。



陈芬(1973-), 女, 博士, 宁波大学信息科学与工程学院副教授, 主要研究方向为三维视频信号处理与编码。



彭宗举(1973-), 男, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。



蒋刚毅(1964-), 男, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。



郁梅(1968-), 女, 博士, 宁波大学信息科学与工程学院教授, 主要研究方向为三维视频信号处理与编码。