



DOI: 10.12086/oe.2018.170627

融合视觉感知特性的 HDR 视频 编码率失真优化算法

杨桐¹, 郁梅^{1,2*}, 姜浩^{1,3}, 蒋刚毅^{1,2}¹ 宁波大学信息科学与工程学院, 浙江 宁波 315211;² 南京大学计算机软件新技术国家重点实验室, 江苏 南京 210093;³ 浙江工商职业技术学院智能家电宁波市重点实验室, 浙江 宁波 315012

摘要: 针对高动态范围 (HDR) 视频较之于传统低动态范围 (LDR) 视频所需存储资源和传输带宽急剧增加的问题, 本文提出了一种基于视觉感知特性的 HDR 视频编码的动态率失真优化算法, 以提高高效视频编码 (HEVC) Main 10 编码 HDR 视频的性能。本文通过引入视觉选择性关注信息, 对不同区域采取非均等的失真权重分配策略, 优化常规的失真计算方法; 同时, 为了进一步去除视频中的感知冗余, 融合视频内容的纹理特性自适应调节拉格朗日乘子, 并应用于编码量化器动态调节量化参数, 实现编码比特和失真感知权衡。实验结果表明: 与 HEVC Main 10 相比, 在相同 HDR-VDP 和 PSNR DE 质量指标下, 所提算法平均节省 7.46% 和 6.53% 码率, 最大分别节省 18.52% 和 11.49%, 所提算法在保持视觉质量的前提下能够有效降低码率。

关键词: 高动态范围视频编码; 率失真优化; 人类视觉系统; 视觉显著

中图分类号: TN919.81

文献标志码: A

引用格式: 杨桐, 郁梅, 姜浩, 等. 融合视觉感知特性的 HDR 视频编码率失真优化算法 [J]. 光电工程, 2018, 45 (1): 170627

Visual perception based rate distortion optimization method for high dynamic range video coding

Yang Tong¹, Yu Mei^{1,2*}, Jiang Hao^{1,3}, Jiang Gangyi^{1,2}¹ Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China;² National Key Lab of Software New Technology, Nanjing University, Nanjing, Jiangsu 210093, China;³ Ningbo Key Lab of Intelligent Household Appliances, Zhejiang Business Technology Institute, Ningbo, Zhejiang 315012, China

Abstract: In view of the drastic increase of storage resources and transmission bandwidth requirement for high dynamic range (HDR) video compared to the traditional low dynamic range (LDR) video, we propose a dynamic rate distortion optimization algorithm based on visual perception for HDR Video encoding to improve the performance of high efficiency video coding (HEVC) Main 10 for coding HDR video. With the information of visual selective attention, we design a non-uniform distortion weight distribution strategy to different regions of interest and improve the conventional method of distortion calculation. At the same time, in order to further eliminate the perceptible

收稿日期: 2017-11-14; 收到修改稿日期: 2017-12-20

基金项目: 国家自然科学基金项目 (61671258); 浙江省自然科学基金项目 (LY15F010005)

作者简介: 杨桐 (1993-), 男, 湖北孝感人, 硕士研究生, 主要从事多媒体信号处理、视频编码等方面的研究。

E-mail: youngtong163@163.com

通信作者: 郁梅 (1968-), 女, 江苏无锡人, 博士, 教授, 博士生导师, 主要研究方向为多媒体信号处理、视频压缩与通信、

计算机视觉及其应用。E-mail: yumei2@126.com

redundancy in HDR video coding, the texture characteristics of video content are used to adjust Lagrange multipliers adaptively, which is applied to the encoder to dynamically adjust the quantization parameters to realize reasonably the trade-off between coded bits and distortion perception. The experimental results show that the proposed algorithm can save an average of 7.46% and 6.53% bitrate with the same HDR-visible difference predictor-2.2(HDR-VDP-2.2) and PSNR_DE compared with HEVC Main 10, saving the maximum of 18.52% and 11.49% respectively. The proposed algorithm can effectively reduce the consumption of the overall bitrates and still maintain the visual quality of the reconstructed HDR video.

Keywords: high dynamic range video coding; rate distortion optimization; human visual system; visual saliency

Citation: Yang T, Yu M, Jiang H, et al. Visual perception based rate distortion optimization method for high dynamic range video coding [J]. *Opto-Electronic Engineering*, 2018, 45(1): 170627

1 引言

目前广泛应用的传统数字图像所能表示的动态范围是非常有限的,这使得场景中的高亮度区域由于曝光过度丢失了细节,而黑暗区域由于曝光不足,细节信息也有所缺失^[1]。高动态范围(high dynamic range, HDR)图像可以表示真实场景中高动态范围的亮度信息,与普通低动态(low dynamic range, LDR)图像相比动态范围更大、存储所需位深更多, HDR 图像所能表现的层次更加丰富,从而达到更加逼近现实的光影效果^[2-3]。因此, HDR 内容具有非常大的应用价值,近年来成为电视行业继 3D、自由视点电视和 4K 之后又一个关注的焦点^[3-7]。

由于 HDR 内容动态范围以及位深的增加,导致 HDR 内容所需存储资源和传输带宽急剧增加。然而,目前最新的视频编码标准——高效视频编码(high efficiency video coding, HEVC)是针对 LDR 视频设计的,这也对 HDR 视频编码提出了新的挑战,因此 HDR 视频编码成为了目前 HDR 视频编码传输急需解决的关键技术。为了提高 HDR 视频编码性能, MPEG 已经推出了新的快速标准化工作,以增强 HEVC 用于 HDR 视频编码压缩的性能^[8-9]。与此同时, MPEG 于 2015 年 2 月发起了 CfE(call for evidence)^[10], 希望通过 HEVC 标准的扩展,并添加新的技术来实现 HDR 视频的高效编码压缩。目前, HDR 视频编码方法可以分为两类:向后兼容的 HDR 视频编码^[11-12]和非向后兼容的编码^[13-15]。

向后兼容的 HDR 视频编码解决方案,能够兼容现有的 8 位 LDR 视频编码传输和显示设备,在解码端可以直接解码基本层的码流,用于解码显示 HDR 色调映射操作(tone-mapping operator, TMO)后的 LDR 内容。同样也可以通过解码全部码流信息,通

过基本层的 LDR 内容和增强层的残差信息解码显示 HDR 内容。首先需要通过 TMO 将 HDR 内容映射到 LDR 内容,并计算原始 HDR 内容与逆色调映射后的 HDR 内容的差值,或者原始 HDR 内容和 TMO 后的 LDR 内容的比值,以此作为增强层并与 LDR 内容构成的基本层进行分级编码压缩^[12]。

另一类是基于感知转换函数的非向后兼容的编码方式^[13-15]:利用基于人类视觉系统(human visual system, HVS)感知模型的感知转换函数(perceptual transformation function, PTF),将浮点型的 HDR 数据转换到编码传输所需的整型数据存储格式,最后利用 10 位或更高位深的视频编码器进行压缩。PTF 的目的是为了使用尽量少的编码位深来表示 HDR 图像内容, Miller 等^[16]基于 Barten^[17]的对比度敏感函数(contrast sensitivity function, CSF)建立了感知量化(perceptual quantizer, PQ)模型,在兼容已有的 REC.709 标准的基础上使光电转换函数平滑,在指定位深限制内最大限度保持 HDR 内容信息和感知质量。Yu^[18]提出了基于 HDR 图像自身特性的自适应 PQ 转换函数,并用于 HEVC Main 10 编码 HDR 视频,有效减少了视频编码失真。Zhang 等^[15]利用文献[19]验证的 HDR 图像亮度掩蔽效应,基于编码图像块的亮度均值感知调节量化参数去除感知冗余。Jung 等^[20]基于 CSF 模型得到 10 bits 的自适应亮度调节方法,并改进 LDR 图像的自由能量恰可察觉失真模型,以此优化 HDR 视频编码器中的拉格朗日乘子。然而,上述方法均未考虑人眼对于 HDR 视频内容的视觉注意力和纹理掩蔽特性。

本文算法将 HDR 视频帧的视觉注意力特性引入到失真的计算中,使得对失真的测量更符合人的视觉系统。同时,考虑到 HVS 对于不容易引起观察者注意的平坦区域中的失真也会非常敏感,对于显著性区

域内纹理复杂的区域也能够容忍更多的失真，为了进一步去除视频中的感知冗余，加入了视频内容的纹理特性来自适应地调节拉格朗日乘子。最终建立了一个基于视觉感知的 HDR 视频编码动态拉格朗日率失真模型，并应用于编码量化器动态调节量化参数，从而实现在保证视觉质量的前提下有效降低编码码率。

2 基于视觉感知的 HDR 动态率失真模型

编码器选择一个最小失真的编码模式可以带来最好的视频质量，然而这往往需要很高的编码比特率。如何在有限的编码比特数下，选择一个失真最小的模式是编码中的关键问题。率失真优化的目的是对于给定编码单元，在给定码率的情况下，尽可能降低失真。对于给定编码单元，上述求极值问题可将其转化为：在给定码率的情况下，尽可能降低失真 D ，即 $\min\{D\}$ s.t. $R \leq R_T$ 。其中， R 和 D 分别表示编码所消耗的比特数码率和失真， R_T 是可用于编码的比特数。在视频编码过程中，编码器引入了拉格朗日乘子将其转变为非限制性问題，并根据比较率失真代价值

J 来选择最优的编码参数，使编码器达到最优的率失真性能。

$$\{Para\}_{opt} = \arg \min_{\{Para\}} J, \quad J = D_{HEVC} + \lambda_{HEVC} \cdot R, \quad (1)$$

式中： λ_{HEVC} 表示原始编码器所采用的拉格朗日乘子， D_{HEVC} 使用绝对差和或误差平方和度量失真， $\{Para\}_{opt}$ 表示最佳的编码参数集，包括模式选择、运动估计等。

为了将 HDR 内容的视觉感知特性引入到 HDR 视频编码，从而去除感知冗余。本文设计了一种基于视觉感知的动态率失真模型，以提高 HEVC Main 10 编码 HDR 视频的性能，其模型框架如图 1 所示。首先，对当前输入 HDR 视频帧获取视觉显著图，将 HDR 视频帧的人眼选择性关注的信息引入到失真的计算中；同时，将输入视频帧双边滤波处理，获取图像的细节纹理特征，并局部自适应调节拉格朗日乘子。然后，计算融合视觉感知特性的率失真代价函数，替代原始的率失真代价计算公式，并感知调节量化参数。最终建立了一种基于视觉感知特性的 HDR 视频率失真优化算法，并应用到整个编码过程，包括模式决策、运动估计和率失真优化量化，使得在降低码率的同时

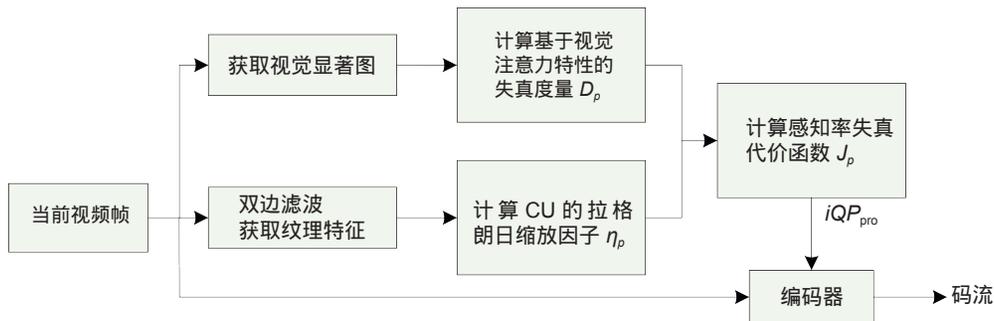


图 1 基于视觉感知的动态率失真优化模型
Fig. 1 Dynamic distortion optimization model based on visual perception

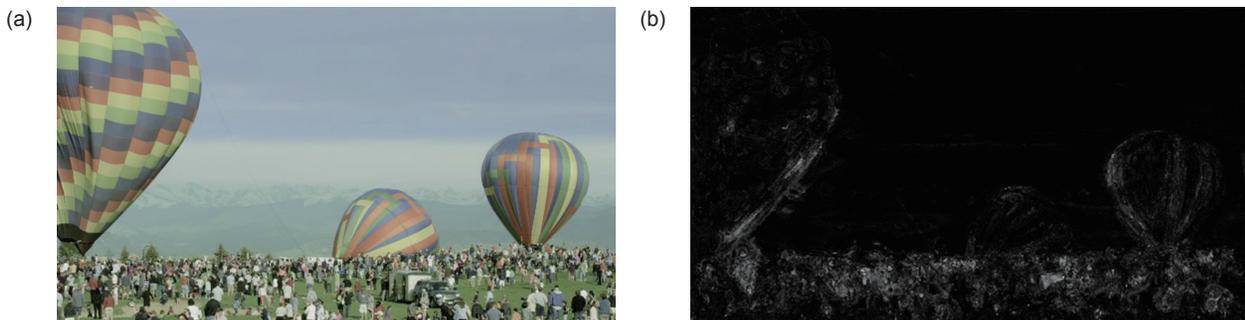


图 2 BalloonFestival 序列 (a) 及其显著图 (b)
Fig. 2 BalloonFestival sequence (a) and its saliency map (b)

保持 HDR 视频质量符合人的视觉感知。

2.1 基于视觉注意力特性的失真计算

所谓的人眼视觉注意力,是指人们在观察周围场景或者图像的时候,会自动有选择性地关注某些特定的目标对象,作为感兴趣或者重要的目标来分析处理。因此,将人眼的视觉注意力机制融合到基于视觉感知的率失真优化算法,对于提升编码性能具有重要的意义。在人的视觉感知系统中,HVS对同一场景图像中不同区域的失真敏感度是不同的,而视觉显著性检测模型可以用来评估图像每个区域可能引起HVS的关注程度。本文采用文献[21]的方法提取HDR视频帧的视觉显著图,文献[21]首先从HDR数据中提取不同特征对应的显著图,然后基于HDR内容眼动数据使用随机森林算法训练模型,将提取的不同特征的显著图融合形成最终的视觉显著图。图2显示了BalloonFestival序列及其使用文献[21]方法得到的显著图,视频帧中每个像素点在显著图中都有一个对应的视觉显著性值,该数值越大代表这个点受到的关注越高,从而HVS对此像素点的失真敏感度也越强。

由于人眼关注的往往不是单个独立的像素点,而是图像上的某个区域。为了便于将视频帧的视觉注意力特性引入到编码单元(coding unit, CU)失真的计算过程,提出基于图像块视觉显著的失真计算权重 w ,

决定图像块在失真计算时的重要程度。与此同时,考虑到HEVC中变换单元的最小尺寸为 4×4 ,故将失真计算权重 w 的计算窗口设置为 4×4 。

$$w = \frac{1}{M} \cdot \sum_{i=1}^M S_i, \quad (2)$$

式中: M 为图像块窗口内像素点总数, S_i 为像素点 i 对应显著图中的视觉显著性值, S_i 越大代表该像素点受到的视觉关注越高,从而HVS对此像素点的失真敏感度也越强。视觉显著性值大的区域表明该区域受到的视觉关注可能性越高,人眼对于对该区域的失真也越敏感,因此在失真计算时分配更大的计算权重,失真计算权重 w 越大。

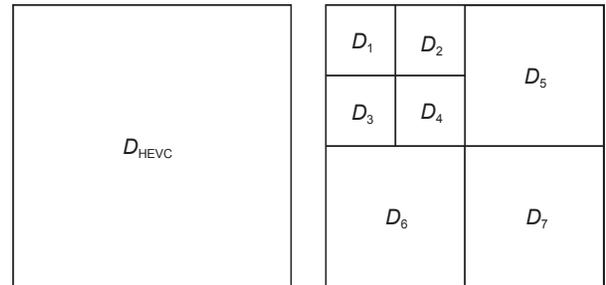


图3 HEVC失真可加性准则
Fig. 3 The rule of additivity for HEVC

HEVC中的原始失真计算满足可加性准则^[22],即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3所示,以HEVC中编码



图4 Market3序列第1帧图像及其细节层。(a)原始图像帧; (b)原始图像帧灰度图; (c)基本层图像; (d)细节层图像
Fig. 4 The first image of Market3 sequence and its detail layer image. (a) Original image frame; (b) Original image frame grayscale; (c) Base layer image; (d) Detail layer image

树单元 (coding tree unit, CTU) 的失真计算为例, 对于一个给定的 CTU, 其按照原始尺寸计算的失真必须严格等于按照子 CU 尺寸计算的失真之和。即当前 CU 的失真 D_{HEVC} 可以分解为 CU 内所有 4×4 大小子图像块的失真之和, 公式表述如下:

$$D_{HEVC} = \sum_{j=1}^N D_{base}(j), \quad (3)$$

式中: $D_{base}(j)$ 表示第 j 个 4×4 大小子块的失真大小, N 为 CU 内 4×4 大小子块的个数。

为了更好地刻画子块在失真计算中的差别性, 在提出的感知率失真优化策略上, CU 内第 j 个 4×4 子块根据其显著程度由式 (2) 计算其失真权重 w_j , 决定该子块在失真计算中的贡献率。采用乘法方式将第 j 个 4×4 子块的失真权重 w_j 与失真 $D_{base}(j)$ 进行融合, 根据式 (3) 将式 (1) 的失真计算修改为如下形式:

$$D_p = \frac{N \cdot \sum_{j=1}^N w_j \cdot D_{base}(j)}{\sum_{j=1}^N w_j} \quad (4)$$

由于区域视觉显著性值较低的子块能够容忍较大的失真, 对于视觉显著性值较低子块分配更小的失真权重, 因此, 视觉显著性值小的 CU 在编码参数选择时, 会优先选择消耗比特率较低的编码参数。相应地, 具有较高视觉显著性值的 CU 对失真更为敏感, 因此 CU 内视觉显著性值大的子块在失真计算时分配更大的失真权重, 从而优先选取编码质量好同时消耗编码比特多的编码参数。

2.2 拉格朗日乘子调节

虽然图像中的平滑区域不太容易引起观察者注意, 但是平滑区域中的失真也会非常容易引起 HVS 察觉; 而对于部分显著性区域中纹理复杂的区域, HVS 也能够容忍更多的失真^[23]。为了降低纹理复杂的区域的码率、保持平滑和结构边缘区域的质量, 从而进一步合理分配有限的码率资源。本文算法采用了边缘保持滤波器分离 HDR 图像的纹理和结构分量, 并引入视频内容的纹理特性, 自适应地调节拉格朗日乘子, 并应用于编码量化器动态调节量化参数。

由于 HVS 对于结构边缘和平滑区域失真的高感知性, HVS 对于结构边缘区域引入的失真比纹理区域失真可见性更高, 所以必须对结构和纹理区域之间进行区分。局部方差常用于区分高频纹理区域和均匀区域, 然而这种方法不能将边缘细分为纹理和结构边

缘区域。为了获取 HDR 图像的纹理区域, 使用双边滤波器^[24]后的细节层可以有效表示 HDR 图像的纹理区域, 并不包含图像结构信息。双边滤波器是非线性的边缘保持滤波器, 滤波系数是由相应位置 $\{p, s\}$ 的高斯滤波器系数, 以及位置 $\{p, s\}$ 对应的当前像素点 I_s 和相近像素点 I_p 的差值得到的高斯系数的乘积构成的。双边滤波器得到的基本层能够很好地保持图像的结构, 主要包含了图像中平滑以及轮廓边缘信息, 像素值 I_s 经双边滤波后的输出值 B_s 为

$$B_s = \frac{\sum_{p \in \Omega} g(p-s) \cdot h(I_s - I_p) \cdot I_p}{\sum_{p \in \Omega} g(p-s) \cdot h(I_s - I_p)}, \quad (5)$$

式中: $g()$ 是空域的高斯函数, $h()$ 是另外一个在像素域的高斯函数, 并可以通过当前像素值 I_s 与滤波后的值 B_s 得到细节层对应的像素值 $|I_s - B_s|$ 。

如图 4 所示, 图 4(a) 为 Maket3 序列第 1 帧原始图像; 图 4(b) 为原始图像帧的灰度图; 图 4(c) 所示的基本层图像。可以看出: 双边滤波器滤波得到的基本层能够有效模糊图像, 同时能够保持清晰的图像边缘, 包含了低频和图像的完整边缘轮廓, 比如滤波后的图 4(c) 中窗子墙壁上的纹理已经被平滑, 但是依然保留了物体的轮廓; 图 4(d) 细节层图像是通过原始图像减去滤波后的图像得到的, 细节层能够完整保留图像中的高频纹理细节, 从细节层图像可以看出: 在图像纹理复杂区域的细节层图像的灰度值较大, 并在图像边缘区域的灰度值较小。当细节层图像中灰度值越大, 表明该区域纹理视觉掩蔽效应明显, HVS 不容易觉察失真。

由于人眼对于平滑没有纹理的区域并不会特别关注, 若仅仅只考虑视觉注意力特性会造成平滑区域失真变得更为严重, 而 HVS 对于平滑区域的失真较为敏感, 故在率失真代价计算过程中同时也应考虑 HDR 图像帧自身的纹理特性。虽然仅仅通过修改失真计算可以达到去除感知冗余的目的, 但是为了更加合理优化率失真代价函数, 需要对其中的拉格朗日乘子作简单的修正, 故引入细节层纹理计算拉格朗日乘子缩放因子 η_p :

$$\eta_p = a + \frac{2 \cdot (1-a)}{1 + \exp\left(-3 \cdot \frac{T_{CU} - T_{frame}}{T_{frame}}\right)}, \quad (6)$$

式中: T_{CU} 、 T_{frame} 分别是当前 CU 和图像帧对应的双

边滤波后细节层灰度均值。 a 为调节参数，参数 a 的取值可以影响最终的编码性能，参数 a 由实验从合理的区间内进行选取。

从式 (6) 可以看出：如果当前 CU 纹理越复杂，HVS 对当前 CU 的编码失真越不敏感，即可以容忍更多的失真，则修正后的拉格朗日乘子缩放因子 η_p 越大。 η_p 值越小 (<1) 表明 CU 平坦光滑，HVS 对于该区域的失真越敏感。然后，根据图像内容自适应调节拉格朗日乘子，并对失真和码率作新的权衡：在纹理复杂的区域拉格朗日乘子相应增大，编码器会优先选择比特消耗少的编码参数；在平坦光滑的区域适当缩小拉格朗日乘子，提高该区域的图像编码质量，以此达到保持图像感知质量不变的前提下节省更多的编码比特。

2.3 量化参数的修正

为了对视频编码率失真过程感知优化，本文利用 HVS 的视觉注意力和纹理掩蔽特性，指导编码量化器最大限度消除视频感知冗余。并结合式 (4) 和式 (6) 修改率失真代价函数，优化后的率失真代价值 J_p 的表述如下：

$$\begin{aligned} \{Para\}_{opt} &= \arg \min_{\{Para\}}, \\ J_p &= D_p + \eta_p \cdot \lambda_{HEVC} \cdot R \quad (7) \end{aligned}$$

编码器通过计算率失真代价函数来选取最优编码参数，在率失真代价函数中，拉格朗日乘子 λ_{HEVC}

与量化参数 (quantization parameters, QP) 之间存在以下关系：

$$\lambda_{HEVC} = \beta \cdot 2^{(QP-12)/3}, \quad (8)$$

式中： β 是与编码图像帧类型相关的常数。因此，利用式 (8) 可以推导得到：

$$QP_{HEVC} = 3 \cdot \log_2 \frac{\lambda_{HEVC}}{\beta} + 12 \quad (9)$$

由式 (9) 可以看出，拉格朗日乘子的变化能够引起量化参数的变化，而 QP 又是编码器中进行量化的依据。令 ΔQP 为原始 QP 的变化量， QP_{HEVC} 、 QP_{pro} 分别为原始 QP 值和感知修正后的 QP 值，由式 (9) 推导可得量化参数变化值为

$$\begin{aligned} \Delta QP &= \underbrace{\left(3 \cdot \log_2 \frac{\eta_p \cdot \lambda_{HEVC}}{\beta} + 12 \right)}_{QP_{pro}} \\ &\quad - \underbrace{\left(3 \cdot \log_2 \frac{\lambda_{HEVC}}{\beta} + 12 \right)}_{QP_{HEVC}} = 3 \cdot \log_2 \eta_p \quad (10) \end{aligned}$$

由于 HEVC 中 QP 值均为整数，因此需对 ΔQP 进行取整操作得到最终的 QP 值 iQP_{pro} ，其中 $floor()$ 为向下取整函数。

$$iQP_{pro} = floor(3 \cdot \log_2 \eta_p + 0.5) + QP_{HEVC} \quad (11)$$

在编码器量化时，可以通过计算其 ΔQP 值，并利用初始 QP 值 QP_{HEVC} 进行简单的增减修改相应变

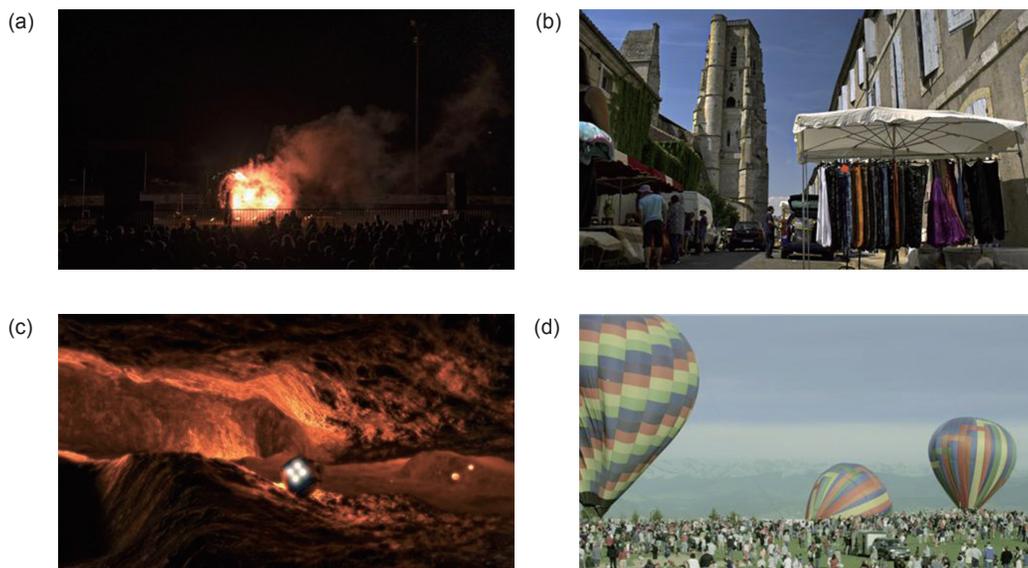


图 5 MPEG AHG 测试材料的 HDR 序列。(a) FireEater2; (b) Market3; (c) Tibul2; (d) Balloon-Festival

Fig. 5 HDR sequence of MPEG AHG test material. (a) FireEater2; (b) Market3; (c) Tibul2; (d) BalloonFestival

换单元的 QP 值 iQP_{pro} ，然后将其发送到解码器。在这种情况下，解码器不需要进行任何对数计算，直接根据修改后的 QP 值 iQP_{pro} 进行反量化和解码。

3 实验结果与分析

为了验证所提算法的有效性，在 HEVC 的参考测试平台 (HM 16.9) Main10 Profile 随机访问 (random access, RA) 配置上进行实验，并使用 CFE [10] 的推荐配置：GOP 大小设置为 8，I 帧周期由视频帧率决定 [10]。实验采用 MPEG AHG(ad hoc group) 的 4 个 HDR 视频测试序列：Market3, Tibul2, BalloonFestival 和 FireEater2，分辨率大小均为 1920×1080 (openEXR 格式)。4 组 QP 设置为 {22, 27, 32, 37}，并与 HEVC Main 10 [10] 比较码率的变化百分比，其最终编码效果使用 Bjøntegaard Delta rate (BD-rate) 来衡量，其中 BD-rate 表示在相同视频质量条件下码率变化的百分比，并通过计算 PSNR_DE [10] 和 HDR-VDP-2.2 [25] 分值衡量视频客观质量。HDR-VDP-2.2 是预测 HDR 图像亮度通道视觉失真差异的全参考质量评价方法，是目前与主观感受最为一致的 HDR 质量评价方法 [26-27]，其中 HDR-VDP-2.2 的 Q 因子越大反映图像质量越好。PSNR_DE 用来衡量重建 HDR 图像和原始 HDR 图像之间颜色失真程度，PSNR_DE 越大 HDR 图像颜色失真程度越小。

3.1 参数 a 的选取

为了确定式 (6) 中参数 a 的取值，实验首先选取了一段合理的区间，由于不可能对区间内所有的值进行遍历，实验比较了区间内 0.4、0.55、0.6、0.7 和 0.85 这 5 组取值对应的编码结果，并从中选取编码性能得到明显提升所对应的取值。图 6 显示了不同测试序列

在不同 a 的取值下对应的 BD-rate 结果，客观质量用 PSNR DE 衡量，其中 BD-rate 为负值代表码率下降、编码性能提升。

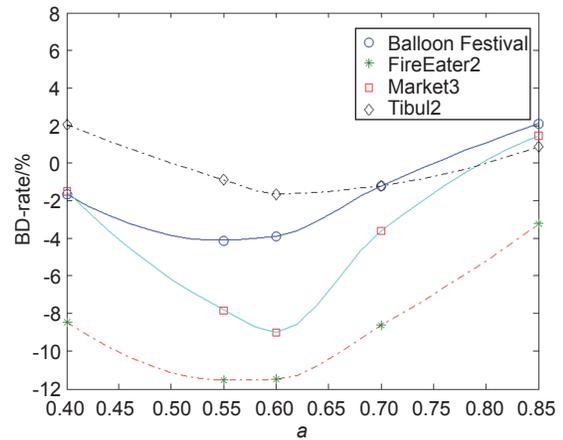


图 6 参数 a 不同取值下测试序列的 BD-rate 结果
Fig. 6 BD-rate results of test sequence under different values of parameter a

从图 6 的实验结果可以看出：当参数 a 的取值小于 0.4 时，拉格朗日乘子变化范围过大，编码重建后的视频块效应明显，导致最终编码性能恶化；当 a 的取值大于 0.85 时，编码过程中拉格朗日乘子动态变化不大，从而导致最终的编码性能提升并不明显。对于 BalloonFestival 和 FireEater2 序列，当 a 取 0.55 时码率节省最多；当 a 取 0.6 时，Market3 和 Tibul2 序列的编码性能提升最为明显。综上所述，当参数 a 的取值为 0.6 时，4 个测试序列的平均 BD-rate 节省最多，因此最终确定参数 a 的取值为 0.6 时较为合理。

3.2 编码效率

表 1 所示为参数 a 取值为 0.6 时测试序列的 BD-rate 结果。从表 1 的实验结果可以看出：与 HEVC 参考测试平台 HM 16.9 Main 10 相比，所提算法在相

表 1 所提算法与对比算法的 BD-rate 结果
Table 1 BD-rate results of the proposed algorithm and comparative algorithm

Sequences	Adaptive PQ [18]		Proposed	
	PSNR_DE	HDR-VDP-2.2	PSNR_DE	HDR-VDP-2.2
BalloonFestival	-3.43	-	-3.88	-18.52
FireEater2	-6.12	-	-11.49	-6.84
Market3	-7.44	-	-9.00	-3.16
Tibul2	-5.04	-	-1.74	-1.30
Average	-5.51	-	-6.53	-7.46

同 HDR-VDP-2.2 的 Q 因子分值下实现了更多的码率节省, 平均节省 7.46%。由于在选取色度分量的 QP 时会参考亮度分量的 QP 值, 所以本文算法对于相同码率下颜色分量的失真也有适当的降低, 在相同 PSNR_DE 客观质量下码率平均下降 6.53%, 同时也优于文献 [18] 的对比算法结果。

Tibul2 序列是由电脑合成产生的, 由于 Tibul2 序列每帧的场景随着机器人的运动而发生变化, 导致每帧运动变化剧烈、时域关联性差; 所提算法会降低部分纹理复杂的 CU 的编码质量, 从而造成时域预测准确性有所降低, 导致最后编码效果提升并不明显。而其他 3 个序列 BalloonFestival、FireEater2 和 Market3 均为自然场景序列, 拍摄时间包含了夜间、白天以及黄昏, 序列整体动态范围较大, 亮暗分明并且纹理细节丰富, 从表 1 的 BD-rate 结果可以看出, 对于这 3 个序列本文算法率失真性能明显优于文献 [18]。FireEater2 序列拍摄时间为晚上, 序列亮度偏暗、整体对比度强, 视觉显著性区域明显, 算法可根据显著区域更加合理分配码率资源。Market3 序列拍摄时间

为光线充足的白天, 细节纹理丰富, 并且相机存在平移运动, 所提算法在保持客观质量的同时, 能够有效降低纹理复杂区域的编码比特。BalloonFestival 序列拍摄的是黄昏下的气球节日, 场景变化缓慢, 整体亮暗分明并且纹理细节丰富, 其编码效果最好; 在相同 PSNR_DE 和 HDR-VDP-2.2 分值下, 与 HM 16.9 相比码率分别降低 3.88% 和 18.52%。

图 7 给出了 4 个测试序列的率失真曲线, 率失真曲线的横坐标均为码率, 图 7(a) 和图 7(b) 的纵坐标为 HDR-VDP2.2 的 Q 因子, 分别为 BalloonFestival 和 Tibul2 序列。图 7(c) 和图 7(d) 纵坐标为 PSNR_DE, 分别为 Market3 和 FireEater2 序列。从图 7 可以看出, 除了 Tibul2 序列编码性能提升不太明显, 所提算法编码性能均明显优化 HM 16.9 Main 10。

图 8 显示了 BalloonFestival 序列第 27 帧图像及其图像帧中红色矩形区域的局部放大图, 图 8(a) 是原始图像帧; 图 8(b) 是 HM 16.9 编码重建得到的, HDR-VDP-2.2 的 Q 因子分值为 53.7123, 编码该帧所消耗比特数为 5280 bits; 图 8(c) 是所提算法编码的重

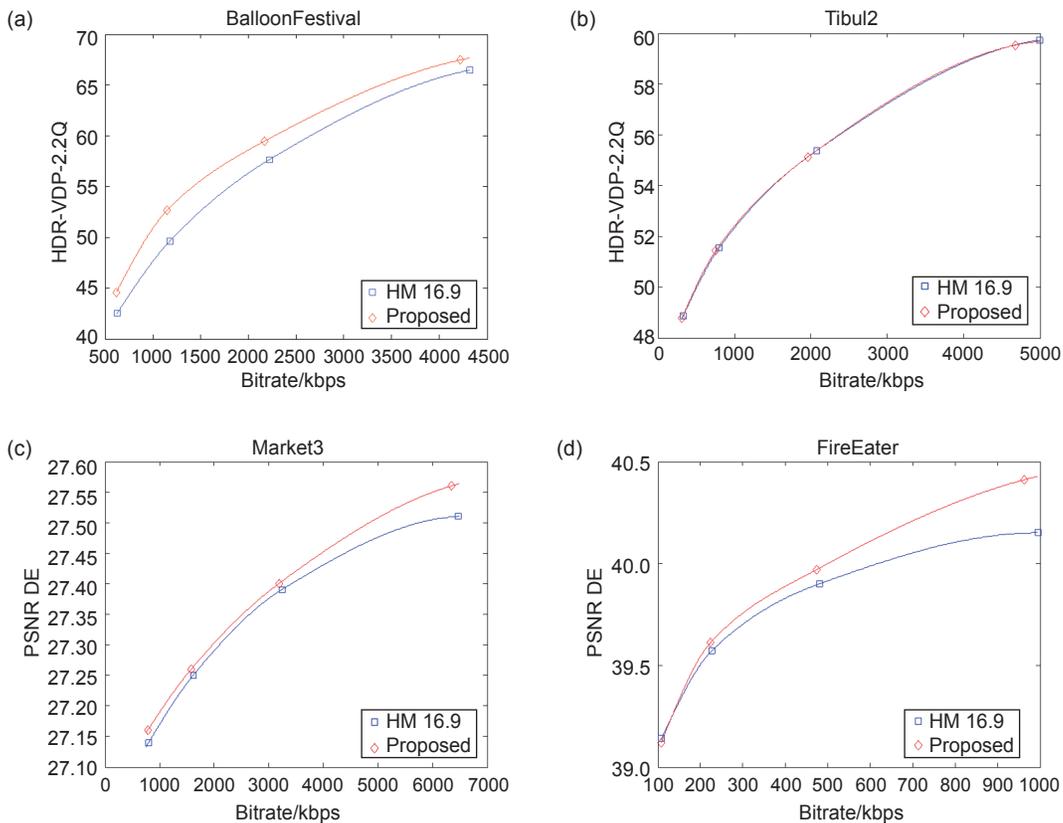


图 7 HM 16.9 和所提算法的率失真曲线比较。(a) BalloonFestival; (b) Tibul2; (c) Market3; (d) FireEater2
 Fig. 7 Comparison of rate-distortion curves between HM 16.9 and the proposed algorithm. (a) BalloonFestival; (b) Tibul2; (c) Market3; (d) FireEater2

建图像, HDR-VDP-2.2 的 Q 因子分值为 53.864, 编码该帧所消耗比特数为 4800 bits。对比可知, 由于所提算法指导量化参数和拉格朗日乘子的自适应变化, 编码比特分配也更加合理, 所以在编码比特消耗减少的同时, 仍能保持重建图像质量。此外, 从局部放大图可以看出, 提出的算法也保留了图像细节和结构信息, 同时也验证了所提算法的有效性。

综上所述, 从实验数据和重建图像局部放大图比较可以看出, 所提算法相比于 HM 16.9 Main 10 在编码比特分配上更加合理, 并且能够保持 HDR 图像的感知质量。尤其对 BalloonFestival 序列, 所提算法显

著提升了编码性能, 与此同时很好地减少了 HDR 图像色度和亮度上的失真。

4 结 论

在视频编码中, 率失真优化技术起着重要的作用。本文依据人眼对于 HDR 内容的视觉注意力选择性关注特性, 将其引入到失真的计算, 使得对失真的测量更符合人的视觉系统。为了进一步去除视频中的感知冗余, 引入了视频内容的纹理特性来自适应地调节拉格朗日乘子。最终建立了一个基于人的视觉感知的 HDR 视频编码动态拉格朗日率失真模型, 并应用

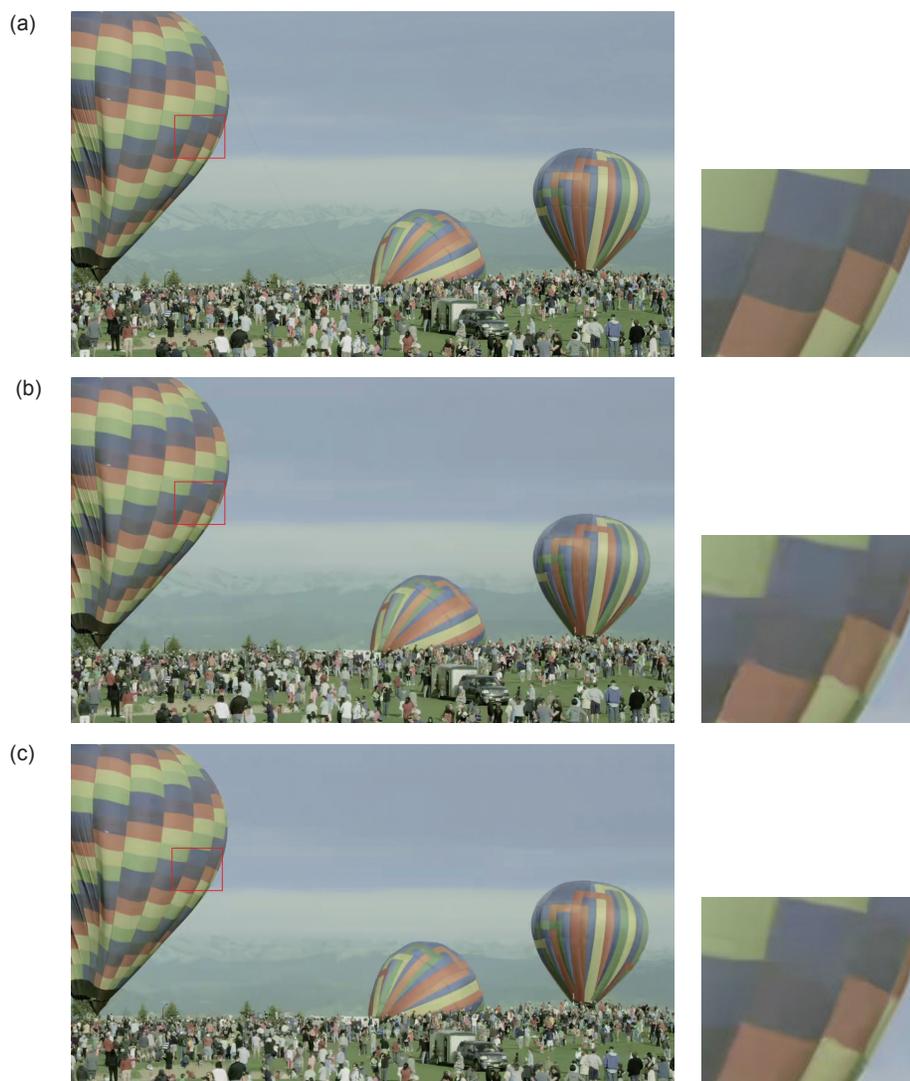


图 8 BalloonFestival 序列第 27 帧图像。(a) 原始第 27 帧图像与局部放大图; (b) HM 16.9 重建图像与局部放大图, $Q = 53.7123$, 5280 bits; (c) 本文算法重建图像与局部放大图, $Q = 53.864$, 4800 bits
Fig. 8 The 27th image of BalloonFestival sequence. (a) The original 27th frame of the image with a partial enlargement; (b) The reconstructed image of HM 16.9 and partial enlargement, $Q = 53.7123$, 5280 bits; (c) The reconstructed image of the proposed algorithm and partial enlargement, $Q = 53.864$, 4800 bits

于编码量化器动态调节量化参数, 以提高 HEVC 压缩 HDR 视频的性能。此外, 与 HM 16.9 Main 10 相比, 所提算法在相同 PSNR_DE 和 HDR-VDP-2.2 分值得下, 码率最大降低 11.49% 和 18.52%, 并对视觉关注区域大的场景和纹理边缘强的序列具有较好的编码效果。

参考文献

- [1] Chalmers A, Debattista K. HDR video past, present and future: a perspective[J]. *Signal Processing: Image Communication*, 2017, **54**: 49–55.
- [2] Hulusic V, Debattista K, Valenzise G, et al. A model of perceived dynamic range for HDR images[J]. *Signal Processing: Image Communication*, 2017, **51**: 26–39.
- [3] Lin Y T, Wang C M, Chen W S, et al. A novel data hiding algorithm for high dynamic range images[J]. *IEEE Transactions on Multimedia*, 2017, **19**(1): 196–211.
- [4] Yang Y, Wang X, Liu Q, et al. A bundled-optimization model of multiview dense depth map synthesis for dynamic scene reconstruction[J]. *Information Sciences*, 2015, **320**: 306–319.
- [5] Yang Y, Liu Q, Liu H, et al. Dense depth image synthesis via energy minimization for three-dimensional video[J]. *Signal Processing*, 2015, **112**: 199–208.
- [6] Yang Y, Deng H P, WU J, et al. Depth map reconstruction and rectification through coding parameters for mobile 3D video system[J]. *Neurocomputing*, 2015, **151**: 663–673.
- [7] Liu Q, Yang Y, Ji R R, et al. Cross-view down/up-sampling method for multiview depth video coding[J]. *IEEE Signal Processing Letters*, 2012, **19**(5): 295–298.
- [8] Francois E, Fogg C, He Y W, et al. High dynamic range and wide color gamut video coding in HEVC: status and potential future enhancements[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, **26**(1): 63–75.
- [9] Kerofsky L, Ye Y, He Y W. Recent developments from MPEG in HDR video compression[C]//*Proceedings of 2016 IEEE International Conference on Image Processing (ICIP)*, 2016: 879–883.
- [10] Luthra A, Francois E, Husak W. Call for evidence (CfE) for HDR and WCG video coding[R]. ISO/IEC JTC1/SC29/WG11 MPEG2015/N15083. Geneva, Switzerland: ISO, 2015.
- [11] Koz A, Dufaux F. Methods for improving the tone mapping for backward compatible high dynamic range image and video coding[J]. *Signal Processing: Image Communication*, 2014, **29**(2): 274–292.
- [12] Mai Z C, Mansour H, Mantiuk R, et al. Optimizing a tone curve for backward-compatible high dynamic range image and video compression[J]. *IEEE Transactions on Image Processing*, 2011, **20**(6): 1558–1571.
- [13] Zhang Y, Reinhard E, Bull D. Perception-based high dynamic range video compression with optimal bit-depth transformation[C]//*Proceedings of the 2011 18th IEEE International Conference on Image Processing (ICIP)*, 2011: 1321–1324.
- [14] Motra A, Thoma H. An adaptive Logluv transform for high dynamic range video compression[C]//*Proceedings of the 17th IEEE International Conference on Image Processing (ICIP)*, 2010: 2061–2064.
- [15] Zhang Y, Naccari M, Agrafiotis D, et al. High dynamic range video compression exploiting luminance masking[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, **26**(5): 950–964.
- [16] Miller S, Nezamabadi M, Daly S. Perceptual signal coding for more efficient usage of bit codes[J]. *SMPTE Motion Imaging Journal*, 2013, **122**(4): 52–59.
- [17] Barten P G J. Formula for the contrast sensitivity of the human eye[J]. *Proceedings of SPIE*, 2004, **5294**: 231–238.
- [18] Yu S T, Jung C, Ke P. Adaptive PQ: adaptive perceptual quantizer for HEVC main 10 profile-based HDR video coding[C]//*Proceedings of 2016 Visual Communications and Image Processing (VCIP)*, 2016: 1–4.
- [19] Zhang Y, Agrafiotis D, Naccari M, et al. Visual masking phenomena with high dynamic range content[C]//*Proceedings of the 20th IEEE International Conference on Image Processing (ICIP)*, 2013: 2284–2288.
- [20] Jung C, Lin Q Z, Yu S T. HEVC encoder optimization for HDR video coding based on perceptual block merging[C]//*Proceedings of 2016 Visual Communications and Image Processing (VCIP)*, 2016: 1–4.
- [21] Banitalebi-Dehkordi A, Dong Y Y, Pourzazd T M, et al. A learning-based visual saliency fusion model for high dynamic range video (LBVS-HDR)[C]//*Proceedings of the 2015 23rd European Signal Processing Conference*, 2015: 1541–1545.
- [22] Sullivan J G, Ohm J, Han J W, 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.
- [23] Zhang H X, Lin S W, Xue P. Improved estimation for just-noticeable visual distortion[J]. *Signal Processing*, 2005, **85**(4): 795–808.
- [24] Durand F, Dorsey J. Fast bilateral filtering for the display of high-dynamic-range images[J]. *ACM Transactions on Graphics (TOG)*, 2002, **21**(3): 257–266.
- [25] Narwaria M, mantiuk R K, Da Silva M P, et al. HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images[J]. *Journal of Electronic Imaging*, 2015, **24**(1): 010501.
- [26] Hanhart P, Bernardo M V, Pereira M, et al. Benchmarking of objective quality metrics for HDR image quality assessment[J]. *EURASIP Journal on Image and Video Processing*, 2015, **2015**: 39.
- [27] Azimi M, Banitalebi A, Dong Y, et al. A survey on the performance of the existing full reference HDR video quality metrics: a new HDR video dataset for quality evaluation purposes[C]//*Int'l. Conf. on Multimedia Signal Processing*, 2014.

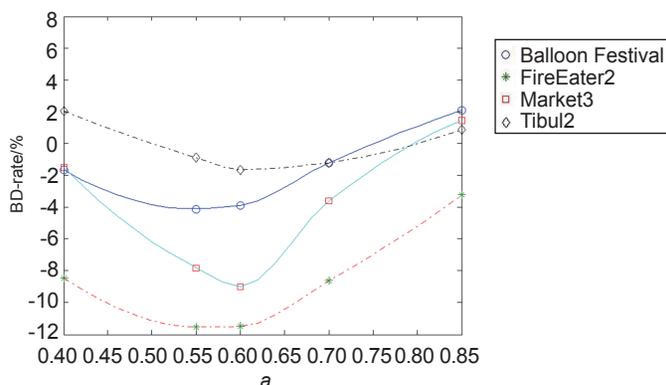
Visual perception based rate distortion optimization method for high dynamic range video coding

Yang Tong¹, Yu Mei^{1,2*}, Jiang Hao^{1,3}, Jiang Gangyi^{1,2}

¹Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China;

²National Key Lab of Software New Technology, Nanjing University, Nanjing, Jiangsu 210093, China;

³Ningbo Key Lab of Intelligent Household Appliances, Zhejiang Business Technology Institute, Ningbo, Zhejiang 315012, China



BD-rate results of test sequence under different values of parameter a

Overview: In view of the drastic increase of storage resources and transmission bandwidth requirement for high dynamic range (HDR) video compared to the traditional low dynamic range (LDR) video, we propose a new dynamic rate distortion optimization algorithm based on visual perception for HDR video encoding to improve the performance of high efficiency video coding (HEVC) Main 10, in which visual attention and texture masking properties of HDR video content are used into HDR video coding. Firstly, the visual saliency map is acquired for the current input HDR video frame. With the information of visual selective attention, we design a non-uniform distortion weight distribution strategy to different regions of interest and improve the conventional method of distortion calculation, which makes the measurement of distortion more in line with human visual system. At the same time, we also take the characteristics of human visual system into account to HDR video coding, such as that human visual system is also very sensitive to distortion in flat areas that are not easily noticeable to the observer, and can tolerate more distortions in areas with complex texture in salient areas. In order to further eliminate the perceived redundancy in HDR video coding, a bilateral filter is used to separate the texture components of the input video frame from which we can extract the texture characteristics to adjust the Lagrange multiplier adaptively. Then, the rate distortion cost function incorporated visual perception is calculated instead of the original rate distortion cost formula, which is applied to the encoder to dynamically adjust the quantization parameters, so as to realize reasonably the trade-off between coded bits and distortion. In the end, the HDR video rate distortion optimization algorithm based on visual perception is established and applied to the whole coding process, including pattern decision, motion estimation and rate-distortion optimization quantization. The proposed algorithm can make it possible to keep the HDR video quality in line with human visual perception while reducing the bitrates. The experimental results show that the proposed algorithm can save an average of 7.46% and 6.53% bitrate with the same HDR-visible Difference Predictor-2.2 (HDR-VDP-2.2) and PSNR_DE compared with HEVC Main 10, saving the maximum of 18.52% and 11.49%, respectively. It can be seen from the experimental results and partial enlargement that the proposed algorithm preserves the image details and structure information well and has good coding effects for scenes with large visual saliency and complex texture. The proposed algorithm is more reasonable in coding bit allocation strategy, which can reduce the consumption of the overall bitrates and still maintain the visual quality of the reconstructed HDR video.

Citation: Yang T, Yu M, Jiang H, *et al.* Visual perception based rate distortion optimization method for high dynamic range video coding[J]. *Opto-Electronic Engineering*, 2018, 45 (1): 170627

Supported by National Natural Science Foundation of China (61671258) and Zhejiang Provincial Natural Science Foundation (LY15F010005)

*E-mail: yumei2@126.com