

## 视频主观观测实验启发的 HEVC 感知帧内码率控制

公衍超<sup>1,2</sup>, 王玲<sup>1</sup>, 刘颖<sup>1,2</sup>, 杨楷芳<sup>3</sup>, 林庆帆<sup>1,4</sup>, 王富平<sup>1,2</sup>

(1. 西安邮电大学通信与信息工程学院, 陕西 西安 710121; 2. 无线通信与信息处理技术国际联合研究中心, 陕西 西安 710121;  
3. 陕西师范大学计算机科学学院, 陕西 西安 710119; 4. 新加坡 Maschine Technology 有限公司, 新加坡 787820)

**摘要:** 针对高效视频编码 (HEVC) 采用的码率控制方法感知率失真性能较低的问题, 提出了一种有效提高视频感知率失真性能的帧内码率控制算法。首先, 度量视频的空时域复杂度, 并结合空时域复杂度设计视频主观观测实验。其次, 基于视频主观观测实验结果, 结合人类视觉系统相关感知特性, 构建能够有效衡量人眼对视频不同区域感知差异性的视频内容空时域感知敏感因子。最后, 将空时域感知敏感因子运用到帧内图像最大编码单元层目标比特分配中, 实现感知帧内码率控制。实验结果表明, 在比特估计准确度和感知率失真性能 2 个核心指标上, 所提算法均明显优于 HEVC 采用的码率控制算法。

**关键词:** 高效视频编码; 码率控制; 人类视觉感知; 目标比特分配

**中图分类号:** TN911.7

**文献标识码:** A

**DOI:** 10.11959/j.issn.1000-436x.2021146

## HEVC perceptual intra-frame rate control inspired by video subjective observation experiment

GONG Yanchao<sup>1,2</sup>, WANG Ling<sup>1</sup>, LIU Ying<sup>1,2</sup>, YANG Kaifang<sup>3</sup>, LIN Qingfan<sup>1,4</sup>, WANG Fuping<sup>1,2</sup>

1. School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

2. International Joint Research Center for Wireless Communication and Information Processing Technology, Xi'an 710121, China

3. School of Computer Science, Shaanxi Normal University, Xi'an 710119, China

4. Maschine Technology Pte. Ltd., 787820, Singapore

**Abstract:** In view of the low perceptual rate-distortion performance of the rate control method used in high efficiency video coding (HEVC), an intra-frame rate control algorithm to effectively improve the perceptual rate-distortion performance of video was proposed. First, the spatiotemporal complexity of videos was measured, and the video subjective observation experiment was designed combined with the spatiotemporal complexity. Secondly, based on the experimental results of video subjective observation, combined with the relevant human vision system perception characteristics, the spatiotemporal perceptual sensitivity factors of video content which could effectively measure the perception difference of human eyes for different regions of video were constructed. Finally, the spatiotemporal perceptual sensitivity factors were applied to the target bit allocation of the largest coding unit layer in the intra-frame image, then the perceptual intra-frame rate control was realized. Experimental results show that the proposed algorithm is superior to the rate control algorithm adopted by HEVC in terms of the bit estimation accuracy and the perceptual rate-distortion performance.

**Keywords:** high efficiency video coding, rate control, human vision perception, target bit allocation

收稿日期: 2021-02-18; 修回日期: 2021-05-06

通信作者: 杨楷芳, kfyang@snnu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61801381, No.61801282); 陕西省自然科学基金资助项目 (No.2020JQ-426); 陕西省国际科技合作与交流计划基金资助项目 (No.2018KW-003); 陕西省教育厅科研计划基金资助项目 (No.18JK0708); 西安邮电大学研究生联合培养工作站重点基金资助项目 (No.YJGJ201902)

**Foundation Items:** The National Natural Science Foundation of China (No.61801381, No.61801282), The Project of Natural Science Basic Research in Shaanxi Province (No.2020JQ-426), The Project of International Scientific and Technological Cooperation and Exchange in Shaanxi Province (No.2018KW-003), The Project of Education Department of Shaanxi Province (No.18JK0708), The Key Project of Postgraduate Joint Training Workstation of Xi'an University of Posts and Telecommunications (No.YJGJ201902)

## 1 引言

视频作为重要的信息传输载体, 已经被广泛应用于人们的日常生活。在实时的视频通信系统中, 传输信道的带宽通常是有限且时变的, 如何在满足信道带宽限制的条件下尽可能提高视频的质量, 是一个被广泛关注的研究问题。为了解决这一问题, 码率控制技术<sup>[1]</sup>应运而生。码率控制技术的基本原理是综合考虑编码视频内容特性、信道带宽大小、缓存器状态等因素, 给编码视频选择一组最优的量化参数 (QP, quantization parameter)。使用这组 QP 编码视频后所产生的码率最接近信道带宽大小的同时, 其对应的重建视频的质量最高。

随着硬件技术、信息处理技术等的高速发展, 在各类视频系统中, 高清视频及其设备已经被普及。为了提高高清视频的编码效率, 高效视频编码标准 (HEVC, high efficiency video coding) 被广泛应用<sup>[2]</sup>。在 HEVC 标准化过程中, 研究者先后提出了基于码率与 QP 二次模型的码率控制算法、基于  $\rho$  域的码率控制算法和基于  $\lambda$  域的码率控制算法<sup>[3-4]</sup>, 其中,  $\rho$  表示预测变换量化后取值为 0 的系数的百分比,  $\lambda$  表示拉格朗日因子。由于基于  $\lambda$  域的码率控制算法具有更加高效的码率控制性能, 因此其被 HEVC 采用<sup>[5]</sup>。

文献[1,5]从拉格朗日率失真优化角度分析指出, 相比于传统使用的 QP,  $\lambda$  与码率的关系更加密切, 通过调整  $\lambda$  可以获得更加精确的码率值。基于此, 文献[1,5]提出了基于  $\lambda$  域的码率控制算法, 其算法过程主要包括目标比特分配及 QP 确定两大部分, 即由分配的目标码率获得  $\lambda$ , 再由  $\lambda$  得到最终编码的 QP。文献[6-8]针对文献[1,5]中的一些算法细节进行了进一步的改进。针对帧内图像, 文献[6]提出了一种基于绝对变换误差和 (SATD, sum of absolute transformed difference)<sup>[9]</sup>的最大编码单元 (LCU, largest coding unit) 级码率控制算法。该算法将哈达玛变换后的 SATD 作为复杂度衡量测度, 根据 LCU 的 SATD 占剩余未编码 LCU 的 SATD 的比重为每个 LCU 分配合适的目标比特。文献[7]考虑图像所处的时间层和视频内容特性的影响, 提出了一种在编码过程中自适应更新计算图像目标比特分配权重的方法。为了在码率控制过程中获得更好的编码率失真性能, 文献[7]构建了失真与  $\lambda$  的关系模型, 并进一步结合已有的码率与  $\lambda$  的关系模型, 从率失真优化的角度出发针对 LCU 级的码率控制

提出了考虑视频内容特性的最优比特分配方法。目前, 文献[1,5-8]提出的算法由于高效的编码性能已经被 HEVC 采用并集成到最新的测试模型 HM<sup>[10]</sup>中。

但是上述算法<sup>[1,5-8]</sup>都是在优化编码视频的客观率失真性能, 没有考虑人类视觉系统 (HVS, human visual system) 感知特性对于编码重建视频质量的影响。在视频通信系统中, 重建视频的最终接收端通常是人眼。所以理论上, 考虑 HVS 相关感知特性, 以优化重建视频感知率失真性能为目标的感知码率控制算法可以获得更好的编码性能及用户感知体验。因此, 针对 HEVC 标准, 一些考虑 HVS 感知特性的感知码率控制算法被相继提出<sup>[11-18]</sup>。

文献[11]把 LCU 分成背景 LCU 和前景 LCU, 并根据分类信息调整编码 QP 以提高重建视频的感知质量, 但是, 在性能测试时还是以峰值信噪比 (PSNR, peak signal to noise ratio) 为质量测度。面向感兴趣区域 (RoI, region of interest) 视频编码应用, 文献[12]提出了一种基于离散傅里叶变换系数模型和径向基函数神经网络的码率控制算法, 使用均方误差 (MSE, mean square error) 作为质量测度用于率失真优化过程指导  $\lambda$  与 QP 的选择。文献[13]采用梯度衡量图像中不同编码树单元 (CTU, coding tree unit) 的复杂度, 并在帧内图像 CTU 码率分配及 QP 确定中考虑 CTU 的梯度信息, 最终为不同复杂度的 CTU 分配优化的 QP, 但是在性能测试时也是以 PSNR 为测度衡量人眼的感知质量。目前, 大量研究已经表明, PSNR 或 MSE 等不考虑 HVS 感知特性的质量测度与人眼的感知质量一致性较差<sup>[14-16]</sup>。所以文献[11-13]中报道的感知率失真性能的提升还有待商榷。

文献[14-18]使用了一些考虑 HVS 相关感知特性的质量测度以提高码率控制的感知率失真性能。针对帧内图像, 文献[14]以结构相似度 (SSIM, structural similarity) 为质量测度, 提出了一种 CTU 级的目标比特分配方法。同样以 SSIM 为失真测度, 文献[15]首先提出了 LCU 目标比特分配权重计算方法, 并创建了与 SSIM 匹配的新的码率与  $\lambda$  的关系模型用于指导  $\lambda$  与 QP 的计算。面向帧内图像, 文献[16]采用感知加权均方误差 (PWMSE, perceptually weighted mean squared error) 代替 MSE, 建立了码率与感知加权失真的关系模型, 然后基于此率失真模型应用拉格朗日率失真优化理论推导得

到感知加权因子,并最终利用感知加权因子指导 CTU 级的目标比特分配。文献[17]提出了一种基于感知敏感测度 (PSM, perceptual sensitivity measurement) 的面向 HEVC 帧间图像的感知码率控制方法,图像中的每个 CTU 都对应一个 PSM, PSM 的取值与 CTU 的空时复杂度及其对应的 MSE 线性相关, PSM 的大小反映了人眼对于当前 CTU 的感知敏感程度,将 PSM 应用于目标比特分配中以提高码率控制的感知率失真性能。文献[18]采用一种经典的面向图像的考虑亮度掩盖与纹理掩盖的像素域恰可感知失真 (JND, just noticeable distortion) 模型<sup>[19]</sup>作为失真的基本衡量,然后基于此 JND 模型构建了相应的率失真模型,并将其应用于 HEVC 码率控制中,以提高编码视频的感知率失真性能。但是由于包含了像素域 JND 的计算过程及基于 KKT (Karush-Kuhn-Tucker) 条件的最优  $\lambda$  集合搜索方法等步骤,文献[18]提出方法的复杂度比较高。

相比于 PSNR、MSE 等测度,SSIM、PWMSE、PSM、JND 考虑了部分 HVS 感知特性,所以其与人眼感知质量的一致性相对较好。PSNR 与 MSE 是以像素为基本单元进行计算的。但人眼在实际观测视频质量时,不可能精确地观测到像素点的差异,而是以图像中的区域或者物体为基本单元感知视频的质量。SSIM 是以图像区域为基本单元进行计算的。PWMSE、PSM、JND 在创建时考虑了 HVS 的部分感知特性,例如掩盖效应。但是 HVS 的质量感知是一个高度复杂的过程,目前对其运行机理、感知特性等方面的生理心理学研究还处于初始阶段<sup>[20]</sup>。所以仅模型化部分 HVS 感知特性的 SSIM、PWMSE、PSM、JND 等测度与人眼实际感知质量还是有一定的差距<sup>[20-23]</sup>。另外,目前对于 JND 的研究主要集中在对第一个 JND 等级 (JND1) 的研究<sup>[19-20, 24-25]</sup>。JND1 只能反映感知有损对应的失真阈值,要进一步衡量感知有损以上的失真等级,得到一个完整的失真衡量测度,则需要创建多 JND 等级模型<sup>[20-23]</sup>。多 JND 等级是目前理论上最接近人眼感知的失真测度。但是受限于 HVS 的相关基础研究,目前对多 JND 等级的研究进展缓慢。综上,虽然相比于 HEVC 原始码率控制算法,采用 SSIM、PWMSE、PSM、JND 等作为失真测度基于拉格朗日率失真优化理论思路提出的码率控制算法<sup>[14-18]</sup>在一定程度上提升了码率控制的感知率失真性能,但其提升的空间有限。

在设计感知视频编码算法时,除了上述提到的

采用某一具体感知失真测度基于拉格朗日率失真优化的思路外,还有一种切实可行的思路是设计视频主观观测实验<sup>[20]</sup>,即针对具体感知问题,设计匹配的主观观测实验,然后在主观观测结果的指导下设计相关的感知算法。此思路需要设计合适的视频主观观测实验,在获得人眼感知准确结果的同时有效规避了质量测度问题。本文采用设计视频主观观测实验的思路,考虑视频内容特性及其对应的感知特性,提出了一种面向 HEVC 帧内图像的 LCU 级感知码率控制算法。

## 2 所提算法

本节对所提算法进行详细说明。首先描述衡量视频空时域复杂度的测度,然后结合此测度说明视频主观观测实验的设计及视频内容空时域感知敏感因子的构建,最后说明空时域感知敏感因子在帧内图像码率控制中的应用,并引出所提算法的详细流程。

### 2.1 视频内容空时域复杂度衡量

作为一个多维的数字信号,数字视频的内容特性通常可以从空域和时域 2 个角度进行分析。空域复杂度和时域复杂度通常又被分别称为纹理复杂度和运动复杂度。

使用像素亮度分量的梯度衡量视频图像的纹理复杂度。梯度越大表明视频图像纹理越复杂。图像的梯度  $T$  为

$$T = \frac{1}{wh} \sum_{i=1}^{w-1} \sum_{j=1}^{h-1} (|x(i+1, j) - x(i, j)| + |x(i, j+1) - x(i, j)|) \quad (1)$$

其中,  $w$  和  $h$  分别表示图像的宽度和高度;  $x(i, j)$  表示当前图像中坐标为  $(i, j)$  的像素的亮度分量值,图像最左上角的像素为起始像素,其坐标为  $(1, 1)$ 。

采用帧差衡量视频的运动复杂度,帧差越大表明视频内容运动越快。图像的帧差  $D$  为

$$D = \frac{1}{wh} \sum_{i=1}^w \sum_{j=1}^h |x(i, j) - x'(i, j)| \quad (2)$$

其中,  $x'(i, j)$  表示前一帧图像中坐标为  $(i, j)$  的像素的亮度分量值。

### 2.2 纹理感知敏感因子

#### 2.2.1 视频主观观测实验设计

目前常见的视频应用包括广播电视、视频监

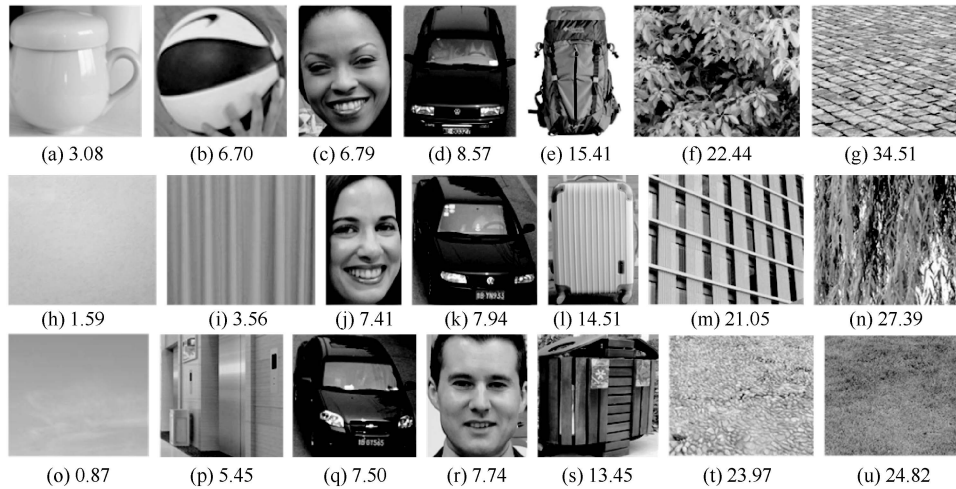


图 1 测试图像

控、视频会议、视频网络点播等。按照人眼对于视频内容的感兴趣程度，视频中的内容通常被划分为人、车、物三类<sup>[26-27]</sup>。所以本文在设计主观观测实验时，选择的测试图像也覆盖这三类内容，并且其空时域复杂度也对应一个比较广泛的范围。如图 1 所示，选择的测试图像都是从标准视频中裁剪出的包含单一内容的图像，每幅图像的梯度值在图像下方给出。测试图像共分三组，图 1(a)~图 1(g)为第一组，图 1(h)~图 1(n)为第二组，图 1(o)~图 1(u)为第三组，每组图像都包括人、车、物三类内容，且都包含纹理简单到复杂的内容。

对于每一组图像，将该组的 7 幅图像随机地按照从上到下的顺序等间隔排列到显示器的最左端。然后让这些图像以相同的运动速度水平向右匀速运动。显示器的分辨率为 3 840 像素×2 160 像素。同时，为了尽可能地消除显示器背景色彩对人类视觉注意力的影响，将显示器的背景色设置为中灰色<sup>[28]</sup>。其他的实验室环境严格按照文献<sup>[28]</sup>中的规定设置。

根据标准规定<sup>[28]</sup>，选择 30 位对视频图像处理领域不熟悉非专家类型人员作为测试者，包括 18 名男性和 12 名女性。这些测试者都通过了必要的视觉感知能力测试，包括采用 Snellen 氏 E 字视力表的视敏度检测、采用石原氏色盲检测图的彩色视觉检测。采用单激励重复播放 (SSMR, single stimulus with multiple repetitions) 方法播放每组图像。测试者观看图像的距离大约是图像高度的 4 倍<sup>[29]</sup>。第一次播放某组图像时，测试者只需观测，播放完毕后显

示器屏幕上内容全部消失，保持 3 s 后第二次播放该组图像。在第二次播放图像时，要求测试者按照约定的规则对该组所有图像按照其感兴趣程度进行评级打分。

本文实验中使用的评级打分准则如图 2 所示。定义感兴趣程度得分 (SoI, score of interest) 衡量人眼对于图像内容的感兴趣程度，一共分为 5 个等级，对每组图像的评级打分按照下面步骤顺序执行。步骤 1，选出 7 幅图像中最感兴趣 (对应 SoI=5) 和最不感兴趣 (对应 SoI=1) 的图像；步骤 2，在剩余的图像中选出中等感兴趣 (对应 SoI=3) 的图像；步骤 3，将剩余图像与中等感兴趣图像进行比较，感兴趣程度较高的为 SoI=4 的图像，其他图像为 SoI=2 的图像。

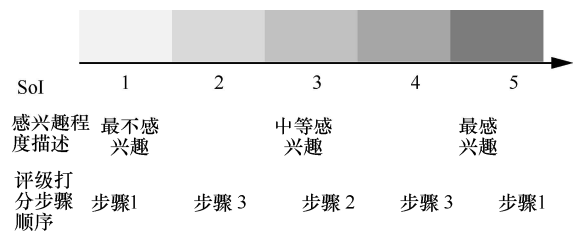


图 2 SoI 测度

考虑测试者个体先验知识的差异性 & 个体观测认真程度的差异性对于评级打分的影响，为了得到更加准确的实验数据，首先采用 95% 置信区间的方法对观测者评价打分数据进行筛选处理。具体处理过程如下。

计算所有测试者对第  $g$  幅图像评级打分的均值  $\bar{v}_g$  为

$$\bar{v}_g = \frac{1}{L} \sum_{r=1}^L v_{r,g} \quad (3)$$

其中,  $L$  是测试者人数, 本文实验中  $L=30$ ;  $v_{r,g}$  是第  $r$  个测试者对第  $g$  幅图像的评级打分值, 即 SoI 值。

所有测试者对第  $g$  幅图像评级打分的标准差  $S_g$  为

$$S_g = \sqrt{\frac{\sum_{r=1}^L (\bar{v}_g - v_{r,g})^2}{L-1}} \quad (4)$$

$$\text{若 } v_{r,g} < \left( \bar{v}_g - 1.96 \frac{S_g}{\sqrt{L}} \right) \text{ 或 } v_{r,g} > \left( \bar{v}_g + 1.96 \frac{S_g}{\sqrt{L}} \right),$$

则舍弃第  $r$  个测试者对第  $g$  幅图像的评级打分值。

按照式(5)重新计算第  $g$  幅图像对应的筛选后所有评级打分值的平均值  $\bar{v}'_g$  为

$$\bar{v}'_g = \frac{1}{L'} \sum_{r=1}^{L'} v'_{r,g} \quad (5)$$

其中,  $v'_{r,g}$  表示经过数据筛选后有效的第  $r$  个测试者对第  $g$  幅图像的评级打分值,  $L'$  表示第  $g$  幅图像经过筛选后的有效测试者样本数。

### 2.2.2 实验数据分析与建模

采用式(5)计算的评级打分平均值衡量人眼对不同纹理复杂度图像内容的感兴趣程度。为了后续内容描述一致性, 这里重新定义一个变量, 即纹理感知敏感因子  $P_T$ 。对于一幅测试图像, 其对应的  $P_T$  值等于式(5)计算的评级打分平均值, 即用  $P_T$  衡量人眼对不同纹理复杂度图像内容的感兴趣程度。图 3 给出了所有测试图像的梯度  $T$  与纹理感知敏感因子  $P_T$  的关系。

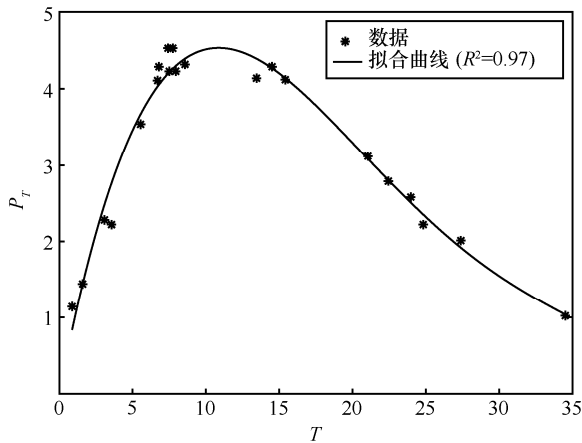


图 3  $T$  与  $P_T$  的关系

对图 3 中的数据采用最小二乘法进行拟合, 拟合精度用相关系数的平方  $R^2$  衡量<sup>[3, 29]</sup>。拟合曲线模型如式(6)所示。

$$P_T = a_1 T^4 + a_2 T^3 + a_3 T^2 + a_4 T + a_5 \quad (6)$$

其中,  $a_1 \sim a_5$  是模型参数, 取值分别为  $-1.6189 \times 10^{-5}$ 、 $0.0018$ 、 $-0.0726$ 、 $1.0084$ 、 $0.0115$ 。

由图 3 可以看出, 相比于纹理简单和纹理复杂的图像内容, 人眼对中等纹理复杂度的图像内容具有更高的感知敏感度, 这一实验结果与 HVS 的相关感知特性是相符的<sup>[30-31]</sup>。纹理简单的图像内容包含的信息量较少, 由于 HVS 选择注意力机制的作用, 这类区域很难吸引人眼的关注。纹理复杂的图像内容虽然包含较多的信息量, 但由于纹理掩盖效应的影响, 人眼很难分辨其细节信息, 因此感知敏感度会降低。相比于前两类图像内容, 中等纹理复杂度的图像内容包含了较丰富的信息量, 易引起人眼的关注, 同时其对应的纹理掩盖效应作用微弱, 人眼易分辨图像内容细节, 因此其对应的感知敏感程度最高。

## 2.3 运动感知敏感因子

### 2.3.1 视频主观观测实验设计

选择图 1(f)、图 1(i)、图 1(r)所示的 3 幅纹理复杂度具有显著差别的图像作为测试图像。对每一幅测试图像执行以下操作: 首先复制当前测试图像, 生成另外 6 幅内容完全相同的图像, 并将这 7 幅图像按照从上到下的顺序等间隔排列到显示器的最左端; 然后让这 7 幅图像以随机确定的运动速度水平向右匀速运动。本节实验中共设定了由慢到快 7 个运动速度等级, 对应于 7 幅图像, 其运动速度快慢由帧差  $D$  衡量。

主观观测实验涉及的其他内容, 如硬件环境、测试者选择、视频播放方法、评级打分规则、测试者评价打分数据筛选处理等与 2.2.1 节描述相同。最后, 定义运动感知敏感因子  $P_D$  衡量人眼对于不同运动等级图像内容的感兴趣程度。图像的  $P_D$  值等于图像对应的经过筛选后所有评级打分值的平均值。

### 2.3.2 实验数据分析与建模

所有测试图像的帧差  $D$  与运动感知敏感因子  $P_D$  的关系如图 4 所示。其关系可以用式(7)所示的多项式关系准确拟合。

$$P_D = b_1 D^4 + b_2 D^3 + b_3 D^2 + b_4 D + b_5 \quad (7)$$

其中， $b_1 \sim b_5$  是模型参数，取值分别为  $-1.464 \times 10^{-8}$ 、 $8.9013 \times 10^{-6}$ 、 $-0.002$ 、 $0.1556$ 、 $0.8673$ 。

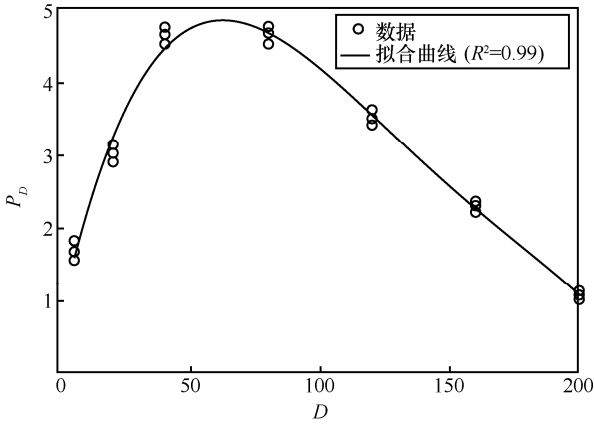


图 4 D 与 P<sub>D</sub> 的关系

由图 4 可以看出，人眼对不同纹理复杂度的图像具有相似的运动感知敏感特性，即相比于运动慢和运动快的图像内容，人眼对于运动速度中等的图像内容更加敏感。这一现象也与 HVS 的相关感知特性一致<sup>[32-33]</sup>。受 HVS 选择注意力机制的影响，人眼容易首先关注视频中运动速度中等或快的图像内容，而易忽略运动速度慢或静止的图像内容。另外，运动快的图像内容虽然易吸引人眼的关注，但是受运动掩盖效应的影响，人眼很难观察到图像内容的细节信息，因此其对应的感知敏感程度会显著降低。而运动速度中等的图像内容不仅易吸引人眼关注，并且运动掩盖效应的作用微弱，所以其对应的感知敏感程度最高。

#### 2.4 空时域感知敏感因子

人眼在实际观测视频时，会同时对视频区域的纹理及运动特性进行观察感知，所以在这一过程中，视频区域的纹理特性和运动特性并不是孤立地作用于人眼的感知，而是相互融合地影响人眼的感知。这就决定了人眼对于视频区域的感知敏感程度不是纹理感知敏感因子  $P_T$  和运动感知敏感因子  $P_D$  的简单累加，而应考虑这 2 个因素在融合影响过程中所产生的重叠效应<sup>[31]</sup>。

文献[31]在研究多种掩盖效应对视频感知质量影响时通过建立非线性加法模型来消除多影响因素间的重叠效应。借鉴这一研究，本文使用如式(8)所示的非线性加法模型构建视频内容空时域感知敏感因子  $P$ 。

$$P = P_D + P_T - \mu \min(P_D, P_T) \quad (8)$$

其中， $\min()$  表示取最小值函数； $\mu$  表示重叠效应系数， $0 < \mu < 1$ ， $\mu$  值越大，表示重叠效应越显著。所提算法中， $\mu$  设置为 0.2。

式(8)构建的空时域感知敏感因子具有普适性，适用于监控视频、会议视频及其他自然视频。第 3 节的实验结果也验证了空时域感知敏感因子的有效性和普适性。

#### 2.5 所提算法的框架

与传统的基于  $\lambda$  域的码率控制算法<sup>[3]</sup>相同，所提算法也主要分为目标比特分配和 QP 确定两部分。所提算法具体步骤如图 5 所示。

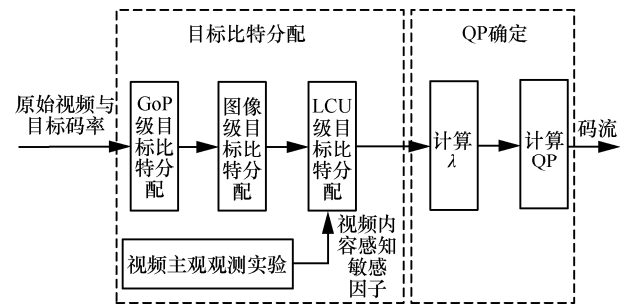


图 5 所提算法具体步骤

##### 1) GoP 级目标比特分配

视频在编码前，首先根据图像在视频中的显示顺序将其划分到不同的图像组 (GoP, group of picture)。GoP 是码率控制目标比特分配过程中最大的处理单元。所提算法按照式(9)进行 GoP 级的目标比特分配。

$$R_c = \text{round} \left( \frac{N_c \left( R_{c,\text{left}} - R_{\text{PicAvg}} \left( N_{c,\text{left}} - \min(N_{c,\text{left}}, S_{\text{GoP}}) \right) \right)}{\min(N_{c,\text{left}}, S_{\text{GoP}})} \right) \quad (9)$$

其中， $R_c$  表示视频中第  $c$  个 GoP 分配的目标比特， $\text{round}()$  表示四舍五入取整函数， $N_c$  表示 GoP 的大小， $R_{c,\text{left}}$  表示编码视频剩余的目标比特数， $N_{c,\text{left}}$  表示视频中剩余未编码的图像数， $S_{\text{GoP}}$  表示 GoP 级滑动窗口的大小，用于保证重建视频质量的平滑性， $S_{\text{GoP}}$  取固定值  $40$ <sup>[1]</sup>， $R_{\text{PicAvg}}$  表示视频每幅图像对应的平均目标比特数，计算式为

$$R_{\text{PicAvg}} = \text{round} \left( \frac{R_{\text{tar}}}{F} \right) \quad (10)$$

其中， $R_{\text{tar}}$  表示目标码率， $F$  表示视频的帧率。

### 2) 图像级目标比特分配

针对帧内图像提出的码率控制算法通常采用全帧内结构进行算法性能测试, 对于全帧内编码结构, 其 GoP 默认为 1。所以, 第  $c$  个 GoP 第  $m$  帧的目标比特数  $R_{c,m}$  为

$$R_{c,m} = \text{round} \left( \alpha \left( \frac{4C_{c,m}}{R_{c,m,\text{left}}} \right)^\beta R_{c,m,\text{left}} \right) \quad (11)$$

其中,  $C_{c,m}$  为第  $c$  个 GoP 第  $m$  帧经过哈达玛变换后的 SATD;  $m=1$ ;  $\alpha$ 、 $\beta$  为模型参数, 当  $40R_{c,m,\text{left}} < wh$  时,  $\alpha=0.25$ , 否则,  $\alpha=0.3$ ,  $\beta=0.5582$ ;  $R_{c,m,\text{left}}$  为视频中未编码的每帧平均剩余比特数, 计算式为

$$R_{c,m,\text{left}} = \text{round} \left( \frac{R_{c,\text{left}}}{N_{c,\text{left}}} \right) \quad (12)$$

### 3) LCU 级目标比特分配

在编码第  $c$  个 GoP 第  $m$  帧的所有 LCU 之前, 对其所有的 LCU 先分配一个基础目标比特。第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 的基础目标比特  $R_{c,m,n}$  为

$$R_{c,m,n} = \text{round} \left( R_{c,m} \frac{P_{c,m,n} C_{c,m,n}}{\sum_{n=1}^{N_{\text{LCU}}} (P_{c,m,n} C_{c,m,n})} \right) \quad (13)$$

其中,  $P_{c,m,n}$  为第  $c$  个 GoP 第  $m$  帧中第  $n$  个 LCU 的感知敏感因子, 其计算模型如式(8)所示;  $C_{c,m,n}$  为第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 经过哈达玛变换后的 SATD;  $N_{\text{LCU}}$  为一幅图像中 LCU 的总数。

对第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 进行编码时, 根据当前帧实际已经编码的数据, 对该 LCU 按照式(14)分配其对应的实际目标比特  $R'_{c,m,n}$ 。

$$R'_{c,m,n} = \text{round} \left( \frac{R_{c,m,\text{left}} + \frac{\sum_{k=1}^{n-1} (R_{c,m,k} - A_{c,m,k})}{\min(S_{\text{LCU}}, N_{c,m,\text{left}})} P_{c,m,n} C_{c,m,n}}{\sum_{u=n}^{N_{\text{LCU}}} (P_{c,m,u} C_{c,m,u})} \right) \quad (14)$$

其中,  $R_{c,m,\text{left}}$  为第  $c$  个 GoP 第  $m$  帧剩余的目标比特数;  $R_{c,m,k}$  和  $A_{c,m,k}$  分别为第  $c$  个 GoP 第  $m$  帧第  $k$  个

LCU 分配的基础目标比特和实际编码的比特, 其中  $1 \leq k \leq n-1$ ;  $N_{c,m,\text{left}}$  为第  $c$  个 GoP 第  $m$  帧剩余未编码的 LCU 数目;  $S_{\text{LCU}}$  为 LCU 级滑动窗口大小, 其取固定值  $4^{[6]}$ 。

根据式(13)和式(14)可以看出, 对当前 LCU 分配目标比特时考虑了当前 LCU 感知敏感因子和其他 LCU 感知敏感因子的影响。相比于当前帧其他 LCU 的感知敏感因子, 如果当前 LCU 的感知敏感因子较大的话, 则认为人眼对其感知较敏感, 其对于整帧图像质量的影响较大, 则对其分配更多的实际目标比特, 优先保证其编码质量。

### 4) QP 确定

图像级 QP 按照文献[6]中的策略确定。第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 的拉格朗日因子  $\lambda_{c,m,n}$  为<sup>[6]</sup>

$$\lambda_{c,m,n} = \frac{\chi}{256} \left( \frac{(C_{c,m,n})^\phi}{R'_{c,m,n} (N_{c,m,n})^{\phi-1}} \right)^\phi \quad (15)$$

其中,  $\chi$  和  $\phi$  为模型参数, 其初始值分别为 6.7542 和 1.7860, 其后取值随视频内容特性更新;  $\phi=1.2517^{[6]}$ ;  $N_{c,m,n}$  为第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 的像素数。

第  $c$  个 GoP 第  $m$  帧第  $n$  个 LCU 的量化参数  $Q_{c,m,n}$  为<sup>[34]</sup>

$$Q_{c,m,n} = \text{round} (4.2005 \ln(\lambda_{c,m,n}) + 13.7122) \quad (16)$$

为了减少重建视频中的质量波动, 根据式(17)和式(18)对  $Q_{c,m,n}$  的取值进一步调整。

$$Q_{c,m,n-1} - 3 \leq Q_{c,m,n} \leq Q_{c,m,n-1} + 3 \quad (17)$$

$$Q_{c,m} - 5 \leq Q_{c,m,n} \leq Q_{c,m} + 5 \quad (18)$$

其中,  $Q_{c,m}$  是第  $c$  个 GoP 第  $m$  帧的 QP。

特别说明, 文献[6]提出的如式(15)所示的计算帧内 LCU 拉格朗日因子的模型和文献[34]提出的如式(16)所示的 QP 与  $\lambda$  的关系模型, 由于高效的编码性能被 HM 原始码率控制算法及一些其他的码率控制算法<sup>[1, 3, 5]</sup>所采用。本文所提算法仍然默认采用式(15)和式(16)所示的模型计算 QP。采用其他的模型计算 QP 理论上也是可行的, 但需要重新设计及调整码率分配部分与 QP 确定部分涉及的模型或者参数, 使其匹配, 从而获得高效的编码性能。

## 2.6 所提算法的流程

所提算法的完整流程描述如下。

**步骤 1** 对于待编码的视频图像, 根据式(1)和式(2)分别计算图像中每个 LCU 对应的梯度值和帧差值。

**步骤 2** 根据式(6)~式(8), 计算得到每个 LCU 对应的空时域感知敏感因子  $P_{c,m,n}$ 。

**步骤 3** 根据式(9)~式(14)完成 GoP 级、图像级及 LCU 级的目标比特分配过程。

**步骤 4** 采用文献[6]的策略计算图像级量化参数  $Q_{c,m}$ , 然后根据式(15)~式(18)得到每个 LCU 对应的拉格朗日因子  $\lambda_{c,m,n}$  和量化参数  $Q_{c,m,n}$ 。

**步骤 5** 使用  $Q_{c,m,n}$  编码当前 LCU。循环执行步骤 1~步骤 5, 直到整个视频编码完成。

### 3 实验结果

#### 3.1 实验设置

目前文献[6]提出的码率控制算法已经被 HEVC 采用, 并被集成到测试模型 HM16.0 中(简称 HM16.0 算法)。针对帧内图像, 文献[3-4,13]分别从不同角度提出了优化的码率控制算法。本节将详细比较所提算法和 HM16.0 算法、文献[3]算法、文献[4]算法、文献[13]算法的编码性能。

选择如表 1 所示的 13 个包含不同分辨率和内容特性的视频作为测试视频。首先, 将 HM16.0 中的码率控制功能关闭, 将全帧内编码结构配置文件 `encoder_intra_main.cfg`<sup>[35]</sup> 中的参数 `QP` 分别设定为 34、37、40、42 的情况下编码每一个测试视频, 从而获得每一个测试视频对应的 4 个目标码率。然后在已经获得的 4 个目标码率下, 分别用不同的算法编码每一个测试视频, 获得最终的编码结果。

对于码率控制算法来说, 比特估计准确度及率失真性能是 2 个核心指标<sup>[1,3,36]</sup>。同时, 算法的复杂度也是需要关注的指标。所以下面依次从比特估计准确度、率失真性能、算法复杂度 3 个方面分析所提算法的性能。

#### 3.2 比特估计准确度

采用被广泛使用的如式(19)所示的比特误差 (BE, bit error)<sup>[1,3]</sup> 衡量码率控制算法的比特估计准确度。

$$BE = \frac{|R_{act} - R_{tar}|}{R_{tar}} \times 100\% \quad (19)$$

其中,  $R_{act}$  为视频实际编码码率, 单位为 kbit/s。

BE 计算结果如表 1 所示, 其中每个视频的 BE 为其对应的 4 个目标码率下获得的 4 个 BE 的平均

值。由表 1 的数据易看出, 对于大部分视频, 所提算法都可以获得较小的 BE 值。另外, 对于所有的视频, HM16.0 算法、文献[3]算法、文献[4]算法、文献[13]算法及所提算法的 BE 平均值分别为 0.005 9%、0.005 5%、0.002 1%、0.002 7% 和 0.002 0%。由此可见, 所提算法的比特估计更加准确。

表 1 不同码率控制算法的 BE

序列	HM16.0 算法	文献[3] 算法	文献[4] 算法	文献[13] 算法	所提 算法
BasketballDrill	0.008 5%	0.008 3%	0.001 4%	0.005 6%	0.005 6%
BQMall	0.001 5%	0.001 7%	0.000 3%	0.002 4%	0.001 2%
ParkScene	0.000 7%	0.000 4%	0.001 1%	0.000 7%	0.000 5%
Johnny	0.007 3%	0.004 3%	0.001 4%	0.005 1%	0.000 5%
KristenAndSara	0.002 6%	0.002 1%	0.001 2%	0.005 5%	0.008 3%
Vidyo1	0.004 3%	0.007 8%	0.000 6%	0.001 2%	0.000 7%
Vidyo4	0.000 5%	0.003 6%	0.001 5%	0.004 2%	0.000 4%
Intersection	0.008 7%	0.004 8%	0.001 2%	0.002 3%	0.001 1%
Mainroad	0.014 6%	0.002 1%	0.001 1%	0.001 0%	0.000 9%
Metro	0.012 4%	0.012 1%	0.009 2%	0.000 5%	0.002 6%
Runners	0.000 7%	0.000 7%	0.001 1%	0.000 7%	0.002 0%
Corridor1	0.005 4%	0.015 3%	0.005 0%	0.002 0%	0.001 0%
Corridor2	0.009 9%	0.007 8%	0.002 2%	0.003 9%	0.001 6%
平均	0.005 9%	0.005 5%	0.002 1%	0.002 7%	0.002 0%

#### 3.3 率失真性能

本文采用被广泛使用的 Bjøntegaard 德尔塔码率 (BD-rate)<sup>[37]</sup> 衡量码率控制算法的率失真性能。在计算不同算法的 BD-rate 时, 将 HM16.0 算法设定为基准算法, 失真测度使用平均主观得分 (MOS, mean opinion score) 五级失真测度<sup>[28]</sup>。将原始视频作为基准视频, 分别获得 HM16.0 算法、文献[3]算法、文献[4]算法、文献[13]算法和所提算法对应的重建视频的 MOS 值。

表 2 给出了不同码率控制算法的 BD-rate。对于所有测试视频, 文献[3]算法、文献[4]算法、文献[13]算法和所提算法对应的 BD-rate 平均值分别为 -3.764 8%、-0.279 6%、-7.045 3%、-16.133 6%。即相比于 HM16.0 算法, 文献[3]算法、文献[4]算法、文献[13]算法和所提算法可以在获得同等质量的重建视频前提下, 码率分别平均减少 3.764 8%、0.279 6%、7.045 3%、16.133 6%。所以, 所提算法的率失真性能更好。这里特别说明, 由于在计算 BD-rate 时, HM16.0 算法被选为基准算法, 因此对于每个测试视频, HM16.0 算法对应的 BD-rate 值都为 0。考虑

表 2 的空间有限, HM16.0 算法对应的 BD-rate 值不在表 2 中呈现。

表 2 不同码率控制算法的 BD-rate

序列	文献[3]算法	文献[4]算法	文献[13]算法	所提算法
BasketballDrill	-5.831 3%	2.228 2%	-10.398 5%	-14.437 5%
BQMall	0.989 3%	5.729 4%	-4.267 4%	-11.859 2%
ParkScene	-2.702 7%	-1.367 6%	-8.749 2%	-12.059 7%
Johnny	-6.088 6%	-3.306 6%	-9.238 8%	-17.364 4%
KristenAndSara	-2.578 4%	-3.818 0%	-5.401 1%	-23.322 2%
Vidyo1	-4.817 5%	-2.899 6%	-6.579 8%	-15.795 0%
Vidyo4	-5.450 9%	-3.165 3%	-6.311 9%	-17.130 8%
Intersection	-3.470 8%	-1.563 6%	3.580 3%	-20.574 6%
Mainroad	-3.470 8%	0.986 0%	-10.076 4%	-17.781 3%
Metro	-4.523 0%	3.296 0%	-8.782 5%	-12.545 1%
Runners	-4.472 1%	5.956 9%	-9.845 1%	-16.792 3%
Corridor1	-3.405 8%	-2.162 8%	-7.176 5%	-13.281 8%
Corridor2	-3.119 4%	-3.547 4%	-8.342 5%	-16.792 3%
平均	-3.764 8%	-0.279 6%	-7.045 3%	-16.133 6%

另外, 相比于其他视频, 所提算法对于背景内容几乎不动而前景内容运动的视频更加有效。这类视频中最典型的为监控视频和会议视频。对于这类视频, 大部分人在观看时容易被前景运动的内容所吸引。相应地, 在主观测试时大部分测试者的测试结果都比较一致, 即他们会更加关注前景运动区域的质量, 前景运动区域的质量提升对整个视频编码的率失真性能影响很大。所提算法能够有效提高前景区域的感知质量。但是, 对于其他类型的视频, 例如整幅图像内容都在运动的视频, 不同测试者在观看这类视频时观测的结果(例如感兴趣的区域)容易产生分歧。所以对于这类视频, 比较难以提出为大部分测试者都认可的感知模型或者算法。

图 6~图 8 分别给出了 Johnny、Intersection、Corridor2 这 3 个序列的率失真性能曲线, 以直观地说明不同码率控制算法的率失真性能。

由图 6~图 8 可以得到以下 2 个结论。首先, 所提算法的率失真性能曲线始终位于其他算法率失真曲线的上方, 即在同等码率下, 所提算法可以获得更高感知质量的重建视频, 所提算法的率失真性能更高。其次, 在不同码率段下, 所提算法的感知质量提升有较大差异。在中等码率下, 所提算法对应的感知质量提升较明显, 相应地, 在高码率和低码率下感知质量提升相对较小。

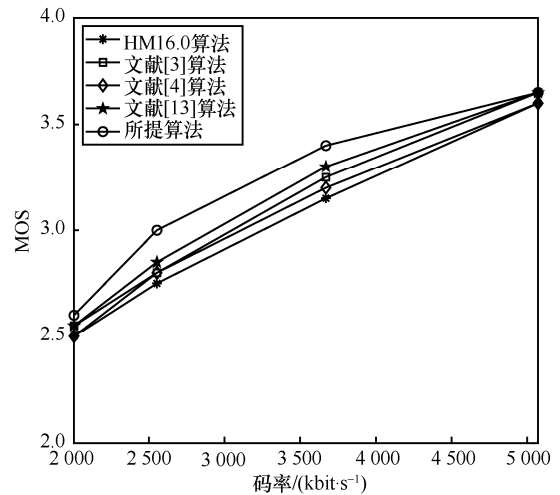


图 6 Johnny 序列率失真曲线

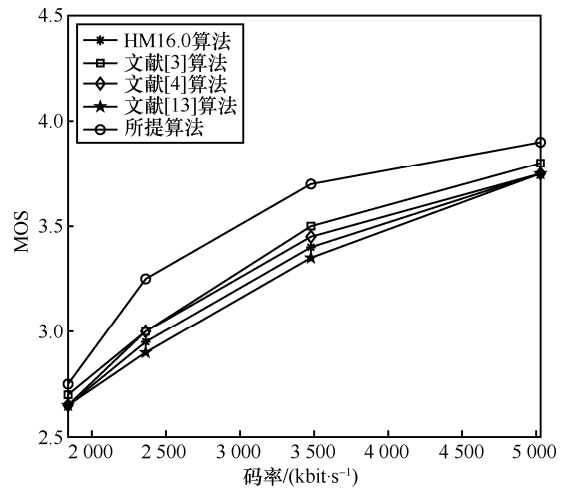


图 7 Intersection 序列率失真曲线

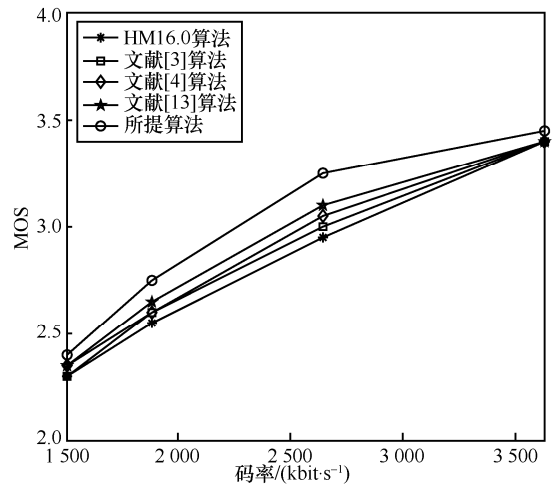


图 8 Corridor2 序列率失真曲线

在高目标码率下, HM16.0 算法对应的重建视频感知质量较好, 其与原始视频之间的质量差距较小, 感知质量可提升的空间很小。同时, 韦伯-费

希纳定律<sup>[38]</sup>也决定了此时只有所提算法重建视频比 HM16.0 算法重建视频有较大的质量提升，人眼才能够感知到质量的变化。上述原因共同决定了在高码率下很难提升重建视频的感知质量。在低目标码率下，由于目标比特资源很有限，这时从感知敏感度较低区域给感知敏感度高的区域搬移的少量码率不足以使感知敏感度高的区域产生明显的感知质量提升。

图 9~图 11 分别给出了 KristenAndSara 序列

第 172 帧、Intersection 序列第 144 帧和 BQMall 序列第 36 帧，来直观展示不同算法的重建图像质量。图 9 中可以明显看出，相比于其他算法的重建图像，所提算法对应重建图像中右边女士的面部区域，例如牙齿、鼻子、眼睛、嘴唇等区域，更加清晰，更加接近于原始图像中的相关内容。图 10(b)~图 10(e) 中重建图像右下角轿车的车牌信息已经模糊不清甚至完全丢失，而图 10(f)中所提算法重建图像对应的车牌区域仍然可以看清楚部分车牌字符，例如



(a) 原始图像 (b) HM16.0 算法重建图像 (c) 文献[3]算法重建图像 (d) 文献[4]算法重建图 (e) 文献[13]算法重建图像 (f) 所提算法重建图像

图 9 KristenAndSara 序列第 172 帧图像及其局部区域放大图



(a) 原始图像 (b) HM16.0 算法重建图像 (c) 文献[3]算法重建图像 (d) 文献[4]算法重建图 (e) 文献[13]算法重建图像 (f) 提出算法重建图像

图 10 Intersection 序列第 144 帧图像及局部区域放大图



(a) 原始图像 (b) HM16.0 算法重建图像 (c) 文献[3]算法重建图像 (d) 文献[4]算法重建图 (e) 文献[13]算法重建图像 (f) 所提算法重建图像

图 11 BQMall 序列第 36 帧图像及局部区域放大图

第 4 个字符“2”和第 5 个字符“6”。由图 11 也可以明显看出,相比于其他算法的重建图像,图 11(f)中所提算法重建图像中右下角坐着的女士的面部区域细节更清晰。

### 3.4 算法复杂度

采用编码时间衡量算法的复杂度<sup>[3]</sup>。算法运行的主要软硬件环境为 Windows 10 操作系统,处理器为 Intel(R) Core(TM) i7-6700 CPU (3.40 GHz),内存 8.0 GB。具体地,采用式(20)所示的  $\Delta T$  衡量各码率控制方法的复杂度。

$$\Delta T = \frac{T_{\text{pro}} - T_{\text{ori}}}{T_{\text{ori}}} \times 100\% \quad (20)$$

其中,  $T_{\text{ori}}$  表示 HM16.0 算法对应的编码时间,  $T_{\text{pro}}$  表示文献[3-4,13]算法或所提算法对应的编码时间。

表 3 给出了不同码率控制算法的  $\Delta T$ 。对于所有的测试视频,文献[3]算法、文献[4]算法、文献[13]算法以及所提算法对应的  $\Delta T$  平均值分别为 -4.777%、0.004%、0.055%和 0.078%。由此可见,与 HM16.0 算法相比,所提算法的编码复杂度增加是非常小的。特别说明,由于在计算  $\Delta T$  时, HM16.0 算法被选为基准算法,因此对于每个测试视频, HM16.0 算法对应的  $\Delta T$  值都为 0。

表 3 不同码率控制算法的  $\Delta T$

序列	文献[3]算法	文献[4]算法	文献[13]算法	所提算法
BasketballDrill	1.247%	0.830%	0.312%	0.480%
BQMall	-1.214%	-0.699%	-3.337%	-0.045%
ParkScene	-2.600%	0.634%	-1.106%	-0.134%
Johnny	-4.977%	0.014%	0.925%	1.110%
KristenAndSara	-4.578%	-0.466%	0.544%	0.552%
Vidyo1	-5.992%	-0.006%	0.813%	-1.033%
Vidyo4	-4.636%	-0.496%	1.375%	-0.937%
Intersection	-6.964%	0.384%	0.917%	-1.347%
Mainroad	-7.163%	0.320%	-0.405%	0.088%
Metro	-7.101%	-0.826%	0.271%	0.363%
Runners	-6.252%	1.072%	-0.266%	-0.938%
Corridor1	-4.723%	-0.796%	-0.255%	0.215%
Corridor2	-7.150%	0.092%	0.922%	2.641%
平均	-4.777%	0.004%	0.055%	0.078%

## 4 结束语

本文提出了一种有效提高码率控制准确度及感知率失真性能的 HEVC 感知帧内码率控制算法。所提算法通过设计匹配的视频主观观测实验,在主观观测实验结果的指导下,构建了准确反映人眼对于视频纹理及运动感知敏感程度的空时域感知敏感因子,并将空时域感知敏感因子应用于 LCU 级比特分配过程中,完成整个码率控制过程。算法为中等纹理复杂度及中等运动速度等人眼感知敏感的区域分配更多比特,从而提升了整个视频的感知质量。实验结果证明,在码率控制准确度及感知率失真性能方面,所提算法优于 HM16.0 算法。本文的研究进一步证明了基于设计视频主观观测实验的思路研究感知码率控制算法的有效性。本文的研究思路可被进一步拓展到针对 HEVC 帧间图像和针对其他视频应用或其他视频类型的码率控制算法研究中。

### 参考文献:

- [1] LI B, LI H Q, LI L, et al.  $\lambda$  domain rate control algorithm for high efficiency video coding[J]. IEEE Transactions on Image Processing, 2014, 23(9):3841-3854.
- [2] SULLIVAN G J, OHM J R, 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] GONG Y C, WAN S, YANG K F, et al. Temporal-layer-motivated lambda domain picture level rate control for random-access configuration in H.265/HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(1):156-170.
- [4] ZHAO F, KU C W, XIANG G, et al. A novel quality enhanced low complexity rate control algorithm for HEVC[C]//2020 IEEE International Conference on Visual Communications and Image Processing. Piscataway: IEEE Press, 2020: 278-280.
- [5] LI B, LI H Q, LI L, et al. Rate control by R-lambda model for HEVC: JCTVC-K0103[C]//Proceedings of the 11th Joint Collaborative Team on Video Coding Meeting. New York: ACM Press, 2012: 1-11.
- [6] WANG X, KARCZEWICZ M. Intra frame rate control based SATD: JCT-VC M0257[C]//Proceedings of the 13th Joint Collaborative Team on Video Coding Meeting. New York: ACM Press, 2013: 1-5.
- [7] LI B, LI H Q, LI L, et al. Adaptive bit allocation for R-lambda model rate control in HM: JCTVC-M0036[C]//Proceedings of the 13th Joint Collaborative Team on Video Coding Meeting. New York: ACM Press, 2013: 1-7.
- [8] LI L, LI B, LI H Q, et al.  $\lambda$ -domain optimal bit allocation algorithm

- for high efficiency video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018,28(1): 130-142.
- [9] 李维, 赵凡, 张二虎, 等. 基于SATD的H.265/HEVC拉格朗日因子选择算法[J]. *通信学报*, 2017, 38(6): 68-74.
- LI W, ZHAO F, ZHANG E H, et al. Lagrange multiplier selection based on SATD for H.265/HEVC[J]. *Journal on Communications*, 2017, 38(6): 68-74.
- [10] MCCANN K, ROSEWARNE C, BROSS B, et al. High efficiency video coding (HEVC) test model 16 (HM 16) encoder description: JCTVC-R1002[C]//*Proceedings of the 18th Joint Collaborative Team on Video Coding Meeting*. New York: ACM Press, 2014: 1-59.
- [11] LI X J, ABUDOULIKEMU Y. Background-foreground information based bit allocation algorithm for surveillance video on high efficiency video coding (HEVC)[C]//2016 IEEE International Conference on Visual Communications and Image Processing. Piscataway: IEEE Press, 2016: 1-4.
- [12] ZHANG Z W, JING T, HAN J N, et al. A new rate control scheme for video coding based on region of interest[J]. *IEEE Access*, 2017, 5:13677-13688.
- [13] WANG M H, NGAN K N, LI H L. An efficient frame-content based intra frame rate control for high efficiency video coding[J]. *IEEE Signal Processing Letters*, 2015, 22(7):896-900.
- [14] GAO W, KWONG S, ZHOU Y, et al. SSIM-based game theory approach for rate-distortion optimized intra frame CTU-level bit allocation[J]. *IEEE Transactions on Multimedia*, 2016, 18(6):988-999.
- [15] ZHAO H L, XIE W, ZHANG Y C, et al. An SSIM-motivated LCU-level rate control algorithm for HEVC[C]//2013 Picture Coding Symposium. Piscataway: IEEE Press, 2013: 85-88.
- [16] LIU X Y, ZHANG Y, ZHU L W, et al. Perception-based CTU level bit allocation for intra high efficiency video coding[J]. *IEEE Access*, 2019, 7:154959-154970.
- [17] ZENG H Q, YANG A S, NGAN K N, et al. Perceptual sensitivity-based rate control method for high efficiency video coding[J]. *Multimedia Tools and Applications*, 2016, 75(17):10383-10396.
- [18] ZHOU M L, WEI X K, KWONG S, et al. Just noticeable distortion-based perceptual rate control in HEVC[J]. *IEEE Transactions on Image Processing*, 2020, 29(6): 7603-7614.
- [19] CHOU C H, LI Y C. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1995, 5(6):467-476.
- [20] WU H R, REIBMAN A R, LIN W S, et al. Perceptual visual signal compression and transmission[J]. *Proceedings of the IEEE*, 2013, 101(9):2025-2043.
- [21] YANG K F, WAN S, WU H R, et al. Detection and estimation of supra-threshold distortion levels of pictures based on just-noticeable difference[C]//2016 IEEE International Conference on Visual Communications and Image Processing. Piscataway: IEEE Press, 2016: 1-4.
- [22] WU H R, LIN W S, NGAN K N. Rate-perceptual-distortion optimization (RpDO) based picture coding-issues and challenges[C]//2014 International Conference on Digital Signal Processing. Piscataway: IEEE Press, 2014: 777-782.
- [23] WANG H Q, KATSAVOUNIDIS I, ZHOU J T, et al. VideoSet: a large-scale compressed video quality dataset based on JND measurement[J]. *Journal of Visual Communication and Image Representation*, 2017, 46(7):292-302.
- [24] WU J J, LI L D, DONG W S, et al. Enhanced just noticeable difference model for images with pattern complexity[J]. *IEEE Transactions on Image Processing*, 2017, 26(6): 2682-2693.
- [25] WANG S Q, MA L, FANG Y M, et al. Just noticeable difference estimation for screen content images[J]. *IEEE Transactions on Image Processing*, 2016, 25(8): 3838-3851.
- [26] 张尧尧. 基于视频结构化的侦查应用[J]. *法制与社会*, 2020, 19: 75-76.
- ZHAGN Y Y. Investigation application based on video structure[J]. *Legal System and Society*, 2020, 19: 75-76.
- [27] 孙展明, 尹伟中. 论视频侦查的几个基本问题[J]. *公安教育*, 2015,12:32-36.
- SUN Z M, YIN W Z. On several basic issues of video investigation[J]. *Public Security Education*, 2015, 12: 32-36.
- [28] ITU-R. ITU-R BT.500-13-2012: Methodology for the subjective assessment of the quality of television pictures[S]. 2012.
- [29] GONG Y C, WAN S, YANG K F, et al. Rate-distortion- optimization-based quantization parameter cascading technique for random-access configuration in H.265/HEVC[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(6): 1304-1312.
- [30] LEE J S, EBRAHIMI T. Perceptual video compression: a survey[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2012, 6(6):684-697.
- [31] YANG X K, LIN W S, LU Z K, et al. Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 15(6):742-752.
- [32] YAMADA K E. Some structural features of the fovea centralis in the human retina[J]. *Arch Ophthalmol*, 1969, 82(2):151-159.
- [33] ONG P E, YANG X K, LIN W S, et al. Perceptual quality and objective quality measurements of compressed videos[J]. *Journal of Visual Communication and Image Representation*, 2006, 17(4):717-737.
- [34] LI B, ZHANG D, LI H Q, et al. QP determination by lambda value JCTVC-I0426[C]//*Proceedings of the 9th Joint Collaborative Team on Video Coding Meeting*. New York: ACM Press, 2012: 1-10.

- [35] BOSSEN F. Common test conditions and software reference configurations[C]//Proceedings of the 11th Joint Collaborative Team on Video Coding Meeting. New York: ACM Press, 2012: 1-4.
- [36] 李维, 杨付正, 任鹏. 考虑视频内容的 H.265/HEVC 帧层码率分配算法[J]. 通信学报, 2015, 36(9):76-81.  
LI W, YANG F Z, REN P. H.265/HEVC frame layer rate allocation algorithm considering video content[J]. Journal on Communications, 2015, 36(9):76-81.
- [37] PATEUXS B, JUNG J. An excel add-in for computing Bjøntegaard metric and its evolution: VCEG-AE07[C]//Proceedings of the Video Coding Experts Group Meeting. New York: ACM Press, 2007: 1-10.
- [38] DABEER O, CHAUDHURI S. Analysis of an adaptive sampler based on weber's law[J]. IEEE Transactions on Signal Processing, 2011, 59(4):1868-1878.



刘颖 (1972- ), 女, 陕西户县人, 博士, 西安邮电大学教授、博士生导师, 主要研究方向为图像检索、图像识别与理解。



杨楷芳 (1987- ), 女, 山西阳泉人, 博士, 陕西师范大学副教授、硕士生导师, 主要研究方向为视频编码、视频质量评价。

[作者简介]



公衍超 (1986- ), 男, 山东蒙阴人, 博士, 西安邮电大学讲师、硕士生导师, 主要研究方向为视频编码、视频质量评价。



林庆帆 (1969- ), 男, 福建晋江人, 博士, 西安邮电大学特聘教授, 主要研究方向为视频编码、图像清晰化及人工智能。



王玲 (1994- ), 女, 河南光山人, 西安邮电大学硕士生, 主要研究方向为视频编码。



王富平 (1987- ), 男, 陕西咸阳人, 博士, 西安邮电大学讲师、硕士生导师, 主要研究方向为图像特征提取。