

文章编号:2095-6134(2018)03-0425-07

简报

基于感兴趣区域的 HEVC 会话视频容错编码新方案^{*}

和智涛,张 灿[†],陈德元

(中国科学院大学电子电气与通信工程学院,北京 101408)
(2017 年 5 月 4 日收稿;2017 年 5 月 31 日收修改稿)

He Z T, Zhang C, Chen D Y. Region-of-interest based error resilient scheme for HEVC conversational video transmission[J]. Journal of University of Chinese Academy of Sciences, 2018,35(3):425-431.

摘 要 在计算和传输资源受限的智能手机上编码传输实时会话视频一直是引人关注的关键技术。在采用肤色、运动和纹理信息快速提取感兴趣区域并且使用丢包率修正感兴趣区域权重的基础上,提出一种 HEVC 会话视频容错编码方案。仿真结果表明,该方案在丢包率 5% ~ 20% 的网络中编码传输时,会话视频的感兴趣区峰值信噪比 PSNR 比现存的随机帧内刷新方法至少提高 0.7 dB,编码时间减少 20% ~40%。实际比特率符合目标比特率要求。

关键词 会话视频传输;感兴趣区;高效率视频编码;人类视觉系统

中图分类号:TN919.8 **文献标志码:**A **doi:**10.7523/j.issn.2095-6134.2018.03.018

Region-of-interest based error resilient scheme for HEVC conversational video transmission

HE Zhitao, ZHANG Can, CHEN Deyuan
(School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China)

Abstract Conversational video coding and transmission on resource-constrained mobile phones has been a key technology. We propose a novel error resilient scheme for conversational video transmission in HEVC which is based on the fast extraction for region-of-interest (ROI) according to skin color, motion, and texture information and adjustments for the ROI weights using the packet-loss rate. The simulation results show that, compared to the random frame refresh method, the proposed scheme improves PSNR of the ROI by at least 0.7 dB and reduces coding time by 20% - 40% when the video coding and delivery is over the channel with the packet-loss rate of 5% -20% . At the same time the actual bitrate matches the target bitrate.

Keywords conversational video transmission; region of interest; HEVC; human visual system

随着移动通信系统快速发展以及智能手机在 软硬件上取得突破,移动视频会话业务应用越来越

^{*} 国家自然科学基金(61571416,41776204)和中国科学院大学校所合作基金(Y65201GY00)资助
[†] 通信作者,E-mail:czhang@ucas.ac.cn

越普遍。新一代高效率视频编码标准 HEVC (high efficient video coding)^[1] 比 H. 264/AVC 视频编码标准在保持相同质量的情况下提高约一倍的压缩效率,同时也增加了编码的复杂度。由于移动设备电池续航能力的瓶颈受限以及移动信道的复杂性和不可预测性,移动视频会话业务在提高压缩效率的同时,要尽量降低编码的复杂度。因此,HEVC 标准下提高感兴趣区的容错性能和减少编码的复杂度是移动视频会话在智能手机上实时应用的难题。

在现有的改善方法中,一类方法是估计端到端的失真来对视频编码进行率失真优化,增强传输过程中抗差错性能,如通过计算像素一阶矩和二阶矩估计端到端失真的 ROPE 算法^[2],通过在编码端模拟多个信道与解码器估计端到端失真的方法^[3],针对 H. 264/AVC 标准中的端到端失真估计方法^[4-5]以及使用 HEVC 中编码单元的大小估计端到端失真的方法^[6]。这类方法虽然能有效提高视频在进行有差错传输时的容错性能,但需要估计端到端失真,难以满足受限资源的移动设备上实时会话视频的要求。另一类方法是通过特定的模式或者机制增加帧内编码块,抑制传输过程中的错误扩散,如用于 H. 264/AVC 随机帧内刷新方法^[7]、帧内距离系数加权法^[8]等。这类方法虽然没有增加额外编码复杂度,但过多插入帧内模式使得视频压缩效率降低,容易造成视频图像内比特分配和视频质量分布不均,与人眼视觉系统(human visual system, HVS)的特性不匹配。为改进这些不足,另一些编码方案考虑了 HVS,根据视频不同区域的感兴趣程度将视频图像分割为感兴趣区(region of interest, ROI)和非感兴趣区进行自适应码率分配编码。如分别对 ROI 和非 ROI 的进行端到端率失真代价函数加权^[9]的方法,需要估计端到端失真,仍难以满足实时会话视频的要求。文献[10]根据接收端反馈 ROI 内的差错状况进行帧内刷新,增加了系统的延时,影响会话业务的实时性。Meddeb 等^[11]提出一种将 ROI 与 HEVC 的新特性 Tiles 相结合的算法,提高了 ROI 的抗差错性能。然而上述基于感兴趣区域的方法都没有考虑到在不同程度的传输错误发生时人眼对于画面中感兴趣的程度也是不同的。为解决这个不足,本文针对移动会话视频提出一种基于感兴趣区的 HEVC 容错编码方案。该方案联合信道的实时状态和感兴趣区的

兴趣度量构造错误扩散因子,影响感兴趣的编码器的帧内帧间模式决策。由于本文方案采用 ROI 的编码方法,HEVC 标准参考软件 HM^[12]中采用的基于 $R-\lambda$ 模型的码率控制算法^[13]已经不适用于本方案,所以本文同时给出相应的联合信源信道目标比特分配算法。实验结果表明该方案有效提高了会话视频中感兴趣区的错误抵抗能力,并且降低了编码时间复杂度,可用于移动通信网络中的实时会话视频业务。

1 感兴趣区的快速检测与分割

最常见的移动会话视频场景即用户使用移动终端(如移动蜂窝电话、平板电脑等)进行视频聊天,对于此类设备拍摄的视频图像呈现出这样的特点:背景静止基本保持不变或者有快速轻微晃动,人脸位于视频中部区域,脸部有复杂的运动,含有丰富的细节(眼神、口型、神态等)。所以在移动会话视频中,人脸区域是会话者视觉关注的感兴趣区,人脸的检测成为首要的任务。尽管当前人脸识别技术已经取得突破,在准确度和检测速度上都取得了进展,但是仍然不能保证实时编码与传输。并且人脸检测的精确度不是重点,而是需要尽可能快速地检测人脸并用于后面资源消耗更大的编码过程中。针对这个特点本文提出一种简单快速的基于肤色检测和纹理复杂度的方法快速分割出感兴趣区,并给出以像素块为单位的兴趣度量,方便用于后续 HEVC 视频编码和传输。

给定大小为 $N \times N$ 的像素块,使用肤色检测算法^[14]定义肤色因子 s

$$s = \frac{1}{N^2} \sum_{i=0}^N \sum_{j=0}^N p(i, j), \quad (1)$$

式中: $p(i, j)$ 表示像素 (i, j) 的肤色检测结果,若该像素被检测为肤色时 $p(i, j) = 1$, 否则 $p(i, j) = 0$ 。定义像素块的纹理运动复杂因子 t

$$t = \frac{1}{aN^2} \sum_{i=0}^N \sum_{j=0}^N (|G_x(i, j)| + |G_y(i, j)| + |G_t(i, j)|), \quad (2)$$

式中: a 为归一化系数,像素的空间梯度 $G_x(i, j)$, $G_y(i, j)$ 和时间梯度 $G_t(i, j)$ 计算如下:

$$\begin{aligned} G_x(i, j) &= f_i(i, j) - f_i(i-1, j), \\ G_y(i, j) &= f_i(i, j) - f_i(i, j-1), \\ G_t(i, j) &= f_i(i, j) - f_{i-1}(i, j), \end{aligned} \quad (3)$$

式中: $f_{i-1}(i, j)$ 表示该像素在前一帧对应的像素点。定义该像素块的兴趣度量 i_{block}

$$i_{\text{block}} = \alpha s + \beta t, \quad (4)$$

式中 α 和 β 为给定系数。为了将感兴趣区与后续 HEVC 编码过程相匹配,给定 2 个阈值 θ_1 和 θ_2 ,整幅图像依照式(4) 计算的兴趣度量进行阈值化,经过阈值化后对应的兴趣度量 I_{block} 为

$$I_{\text{block}} = \begin{cases} \theta_1, & i_{\text{block}} \geq \theta_1 \\ \theta_2, & \theta_2 \leq i_{\text{block}} < \theta_1 \\ 1, & i_{\text{block}} < \theta_2 \end{cases} \quad (5)$$

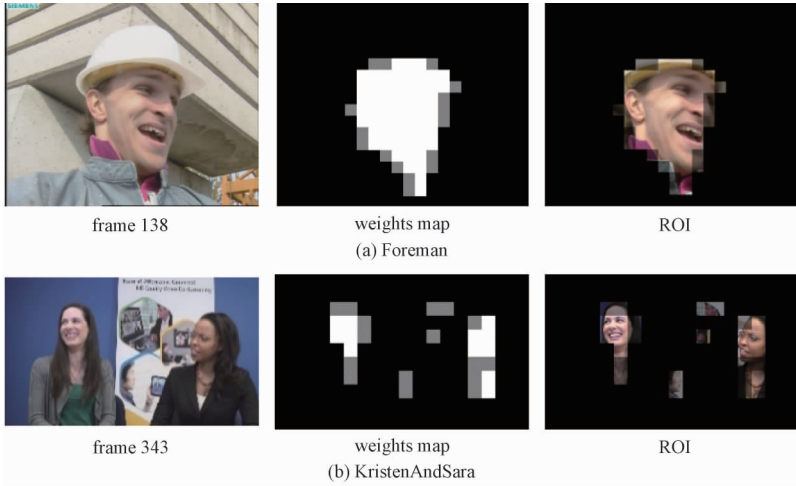


图 1 感兴趣区的分割示例
Fig.1 Examples of extraction for ROI

2 基于感兴趣区的容错编码方案

经过 HEVC 编码的会话视频在容易发生差错的移动网络中传输,会产生错误扩散,严重降低视频传输的质量。为保证接收端视频的主观质量,需要在保证视频编码实时性的同时提高感兴趣区的抗差错性能,本节提出基于感兴趣区的容错编码方案。首先使用信道状况修正感兴趣区的兴趣度量值,然后为不同感兴趣区域内的树形编码单元划分模式进行限定以降低复杂度,接着阐述联合信道的帧内帧间模式选择方法,最后针对方案的特点讨论新的目标比特分配方法以满足本方案的码率控制要求。

2.1 联合信道的感兴趣区兴趣度量

当信道状况变差时,人眼对于感兴趣区内的视频质量会比传输条件正常时更加敏感,因此本文针对不同的感兴趣区的兴趣度量自适应修正,将信道变化状态自适应与兴趣度量的变化匹配:

$$I'_{\text{block}} = \begin{cases} b\theta_1 p, & \text{block} \in \text{ROI}_1 \\ \theta_2, & \text{block} \in \text{ROI}_2 \\ 1, & \text{block} \in \text{NROI} \end{cases} \quad (6)$$

通过式(5),整幅图像被分割成为 3 个不同的区域,分别记作感兴趣区 ROI_1 、感兴趣区 ROI_2 和非感兴趣区 NROI 。图 1 展示使用上述方法对 foreman 序列第 138 帧和 KristenAndSara 序列第 343 帧分割感兴趣区的例子。其中兴趣度量值灰度图中最亮的块为感兴趣区域 ROI_1 ,灰色块为感兴趣区 ROI_2 以及黑色块为非感兴趣区 NROI 。

式中: p 为当前信道的丢包率, $p \neq 0$; b 为给定的错误敏感系数,用于表征丢包对于感兴趣区度量的影响程度。当传输过程中丢包率增加时,人眼对 ROI 区的感兴趣度提升;当信道状态转好时,对 ROI 区的感兴趣度下降。修正后的兴趣度与信道变化状态自适应匹配,用于后续自动调控视频质量与抗差错能力。

2.2 基于感兴趣区的 CTU 划分模式

为提升编码效率,HEVC 引入树形编码单元(coding tree unit, CTU)作为最基本的编码单元。一幅图像被划分为若干个 $L \times L$ 大小的 CTU,每个 CTU 可以当作一个编码单元(coding unit, CU),同时也可以继续分割成 4 个子 CU,此时划分深度为 1,子 CU 可以接着再继续分割,直至到达所允许的最小 CU 尺寸,此时该 CU 具有最大划分深度,形成基于四叉树的循环分层结构。在各个 CU 的编码过程中,又可以进一步将 CU 分割成预测单元(prediction unit, PU)和变换单元(transform unit, TU)以进行帧内帧间模式预测、变换和量化等后续过程。若要获得最优的 CTU 划分方式,则需要遍历所有可能的模式,并取使拉格朗日代价

函数最小的模式为最优模式。尽管能获得最优模式,但是可观的计算量使得该方法不可能用于移动视频实时会话业务中,所以 CTU 的划分与模式选择方法需要在计算复杂度与编码性能之间做出权衡。

考虑到在视频编码中,较大的编码块可以降低编码复杂度,但是较小的编码块又能使得图像的细节预测更加准确。所以为上节分割得出的不同区域设定 CTU 所允许的最大划分深度 depth_{\max} :

$$\text{depth}_{\max} = \begin{cases} 3, & \text{CTU} \in \text{ROI}_1 \\ 2, & \text{CTU} \in \text{ROI}_2 \\ 1, & \text{CTU} \in \text{NROI} \end{cases} \quad (7)$$

根据不同的感兴趣区域设定 CTU 的最大划分深度使得编码过程中减少了很多模式的遍历,降低了计算复杂度,使得编码所需资源耗费减少。

2.3 联合信源信道帧内帧间模式选择方法

HEVC 编码过程中,为得到最优的帧内帧间模式,使用基于拉格朗日的率失真优化方法。通过最小化代价函数(8)得到最优的帧内或帧间预测模式

$$J = D_{\text{mode}} + \lambda_{\text{mode}} R_{\text{mode}}, \quad (8)$$

式中: D_{mode} 为不同模式编码产生的失真, R_{mode} 为不同模式编码所需的比特数, λ_{mode} 为拉格朗日因子。由于视频编码经信道传输后,式(8)中失真 D_{mode} 并没有考虑传输过程中发生差错的影响,为此对失真 D 引入错误因子 ω ,得到新的率失真代价函数

$$J' = \omega D_{\text{mode}} + \lambda R_{\text{mode}}, \quad (9)$$

考虑视频中发生差错的错误敏感程度不同,并且帧内模式会阻止错误扩散,故式(9)中 ω 由当前预测单元的兴趣度量决定:

$$\omega = \begin{cases} 1, & \text{intramode} \\ nI'_{\text{block}}, & \text{intermode} \end{cases}, \quad (10)$$

式中, I'_{block} 在式(6)给出, n 为当前 CU 内所有 PU 的平均帧内距离系数。帧内距离系数的计算方法为:若 PU 最终被决策为帧内模式,则 PU 的帧内距离系数置为 0;若决策为帧间模式,则在对应参考图像的预测块的帧内距离系数进行加 1 作为该 PU 的帧内距离系数。在计算帧内距离系数时,如果参考图像预测块与多个具有不同帧内距离系数的预测块重叠,则根据所占重叠区域比例计算参考图像预测块的平均距离系数^[8]。以 4×4 大小

的块为例,如图 2 所示,阴影部分为参考图像预测块,它的距离系数为 $n = (n_1 \times 3 + n_2 \times 1 + n_3 \times 3 + n_4 \times 9)/16$,则当前编码块的帧内距离系数为 $n + 1$ 。

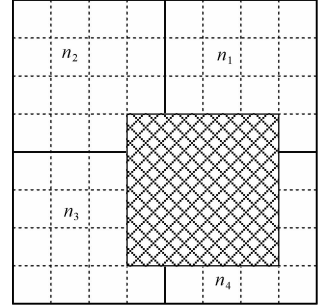


图 2 预测块与不同距离系数块相重叠示例
Fig.2 An example of the prediction unit which overlaps blocks with different weights

2.4 基于感兴趣区的联合信源信道目标比特分配方法

HEVC 标准参考软件 HM16.7^[12]中采用的基于 $R - \lambda$ 模型的码率控制算法^[13],分为两个步骤,首先根据目标码率分配目标比特数,接着确定不同编码单元的量化参数。为了将更多的编码比特分配与上文中分割的感兴趣区域相匹配,并且考虑信道状态对兴趣度量的影响,本文提出基于感兴趣区的图像级目标比特分配和 CTU 级目标比特分配方法代替原先码率控制算法中对应级别的比特分配方法。

2.4.1 图像级的目标比特分配方法

给定图像组(Group of Pictures, GOP)的目标比特 T_{GOP} ,且当前 GOP 内有 N_{GOP} 幅图像,则当前图像 i 分配的比特数 $T_f(i)$ 为

$$T_f(i) = \frac{\omega_f(i)}{\sum_{j=1}^{N_{\text{GOP}}} \omega_f(j)} \cdot (T_{\text{GOP}} - R_{\text{used}}), \quad (11)$$

式中: R_{used} 为当前 GOP 中已经编码的图像生成的比特数, $\omega_f(i)$ 为当前图像 i 在该 GOP 内的所占权重。在比特分配的时候,为了把更多的比特分配给感兴趣区占比例更大的图像,根据式(1)中图像感兴趣度的度量,给出图像 i 的权重 $\omega_f(i)$ 的计算方法

$$\omega_f(i) = \frac{\beta N_{\text{ROI}_1} + N_{\text{ROI}_2}}{N_{\text{CTU}}}, \quad (12)$$

式中: N_{ROI_1} , N_{ROI_2} 和 N_{NROI} 分别为图像 i 中 ROI_1 , ROI_2 和 NROI 内含 CTU 的数目; N_{CTU} 为图像所含的

总 CTU 数; β 可由式(2)算得: $\beta = a\theta_1p/\theta_2$ 。由式(12)可以看到,若图像中感兴趣区所占比例更大的话,那该图像在此 GOP 内会被分配到更多的比特。

2.4.2 CTU 级目标比特分配方法

在计算出当前图像 i 给定的目标比特数 T_i 后,接着对该图像内每个 CTU 确定目标比特数。设当前图像包含有 N_{CTU} 个 CTU,第 m 个待编码的 CTU 的目标比特数 $T_{\text{CTU}}(m)$ 为

$$T_{\text{CTU}}(m) = \frac{\omega_c(m)}{\sum_{k=m}^{N_{\text{CTU}}} \omega_c(k)} \cdot (T_f - R_H - R_{\text{used}}),$$

(13)

式中: R_H 为当前图像的头信息所消耗的比特数; R_{used} 为当前图像中已经编码的 CTU 生成的比特数; $\omega_c(m)$ 为第 m 个 CTU 在该图像内的所占权重,该值可根据式(2)中图像感兴趣度的度量计算出:

$$\omega_c(m) = \frac{I'_{\text{CTU}}(m)}{\left(\sum_{\text{block} \in \text{ROI}_1} I'_{\text{block}} + \sum_{\text{block} \in \text{ROI}_2} I'_{\text{block}} + \sum_{\text{block} \in \text{NROI}} I'_{\text{block}} \right)}$$
$$= \frac{I'_{\text{CTU}}(m)}{N_{\text{ROI}_1} a\theta_1p + N_{\text{ROI}_2} \theta_2 + N_{\text{NROI}}},$$

(14)

式中: $I'_{\text{CTU}}(m)$ 由式(6)给出。从式(14)可以看出,位于感兴趣区内的 CTU 权重更大,分配了更多的比特,并且随着信道丢包的增加,更多的比特会被分配到感兴趣区,非感兴趣区内的 CTU 的比特分配就会减少。

3 实验仿真与结果分析

为验证本文提出的方案的性能,实验参考 HEVC 标准参考软件 HM16.7^[12]。使用 Foreman

和 KristenAndSara 测试序列,分辨率分别为 352×288 和 $1\,280 \times 720$,帧率分别为 30 和 60 fps,用于测试不同分辨率和码率下的算法性能。仿真实验的步骤为:序列使用不同方案编码 150 帧,编码后将输出的比特流按照 $\text{MTU} = 1\,500$ 字节进行 RTP/UDP/IP 打包,进行传输实验丢包仿真。接收端进行拆包后得到损坏的比特流,解码器使用时域错误掩盖方法进行解码后得到重建视频,同一实验重复 100 次。编码器的参数设定基于 HM16.7 提供的标准 encoder_lowdelay_P_main.cfg 配置^[12],使用 6 个 CTU 作为一个 slice。

为比较本文算法和现有算法的性能,本文算法同随机帧内编码算法(RIR)^[7]和使用帧内距离系数加权法(IDW)^[8]进行仿真比较实验。采用 RIR 算法时,帧内刷新率通常设置为当前信道的丢包率 p ;而使用 IDW 算法时,设置加权系数 $K = 5$ 。为公平起见,使用 IDW 方法时采用和本方案相同的感兴趣区分割和 CTU 模式的划分方法。

图 3 为 Foreman 和 KristenAndSara 序列编码经丢包率为 15% 的信道传输后感兴趣区的端到端率失真曲线。从实验结果中看出,在感兴趣区内对于低分辨率的 Foreman 序列来说,本文算法较 IDW 算法提升至少 0.8 dB 的率失真性能,较 RIR 至少提升 1.7 dB 的率失真性能。而对于高分辨率的 KristenAndSara 序列来说本文算法较 IDW 算法至少提升 1.45 dB 的率失真性能,较 RIR 则至少提升 2.4 dB 的率失真性能。实验结果表明在 15% 的丢包率的信道中传输时,本文算法对于不同分辨率的视频来说在感兴趣区内比另外 2 种算法具有更好的抗差错性能,更加符合 HVS 特性。

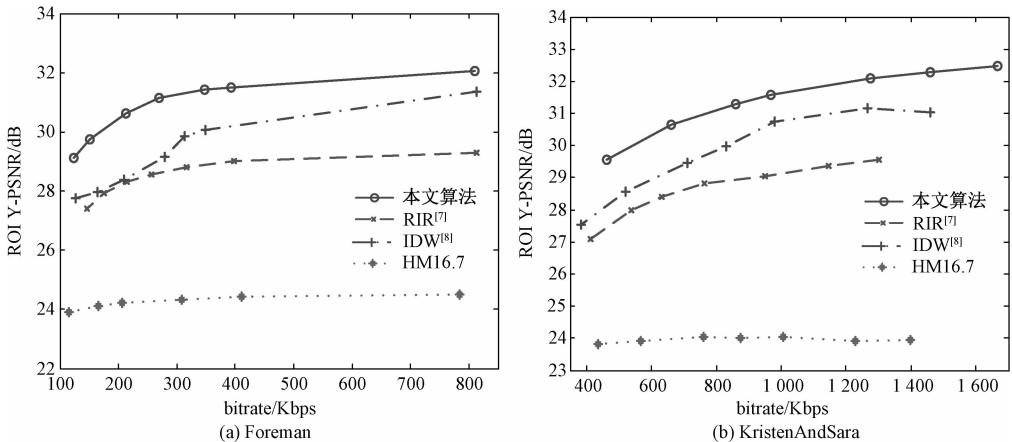


图 3 信道丢包率为 15% 时感兴趣区域端到端率失真曲线

Fig.3 ROI end-to-end rate-distortion curves at the packet-loss rate of 15%

图 4 为将 2 个序列分别以固定比特率进行编码,考察相同比特率在不同丢包率传输时的抗差错性能。图 4(a)为将 Foreman 序列使用不同算法以比特率 150 Kb/s 编码后,通过不同丢包率信道传输后感兴趣区重建质量的对照图。从图中可以看到采用本文算法的感兴趣区视频重建质量比 RIR 方法提高 1.4 ~ 2.49 dB,比 IDW 方法提高 0.74 ~ 0.85 dB。图 4(b)为将 KristenAndSara 序

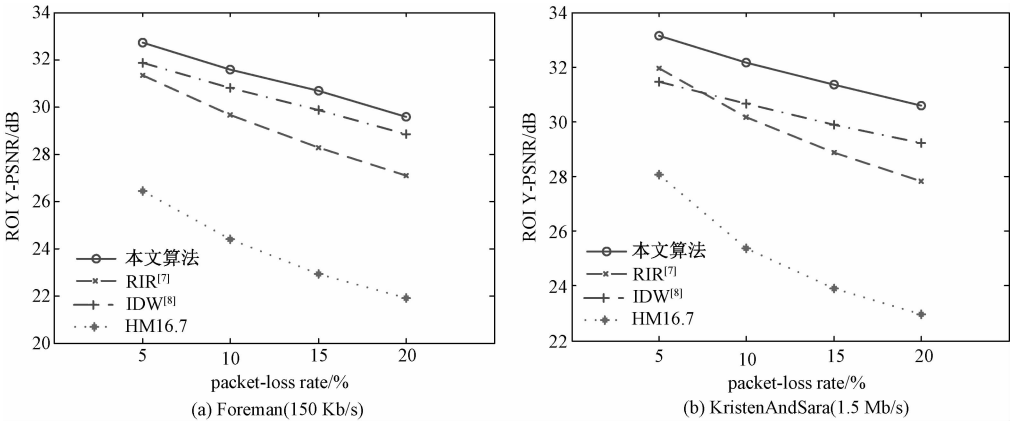


图 4 不同丢包率下以固定比特率编码的算法性能比较

Fig. 4 ROI Y-PSNRs at different packet-loss rates with a fixed bitrate

为了进一步验证本算法与 RIR 和 IDW 的编码计算复杂度,表 1 给出不同算法的编码时间相对于 HM16.7 标准编码时间减少的百分比。从表中可以看到,本文方案采用的 CTU 划分模式方法使得 Foreman 序列编码时间减少 20% 左右,而 KristenAndSara 序列编码时间减少 40% 左右。

表 1 丢包率为 15% 时编码时间复杂度比较

Table 1 Comparison of coding time complexity at the packet-loss rate of 15%

码率/ Kbps	Foreman			KristenAndSara		
	150	300	450	1 000	1 500	2 000
RIR	1. 1	1. 2	1. 0	5. 2	7. 3	4. 3
IDW	23. 7	20. 3	19. 7	40. 7	44. 2	45. 5
本文算法	22. 7	21. 4	23. 1	42. 8	48. 7	43. 0

表 2 给出本文方案中目标比特分配方法性能比较。由于 RIR 和 IDW 算法不涉及码率控制算法,所以本文中目标比特分配方法的性能使用编码实际比特率差错率 R_e 和感兴趣区 ROI 在编码器端重建视频相对于标准 HM16.7 编码增加的 PSNR 来衡量,其中 R_e 计算方法如下:

$$R_e = \frac{|R_{\text{target}} - R_{\text{actual}}|}{R_{\text{target}}} \times 100\% .$$

列使用不同算法以比特率 1.5 Mb/s 编码后,通过不同丢包率信道传输后感兴趣区和非感兴趣区重建质量的对照图。从图中可以看到采用本文算法的感兴趣区视频重建质量比 RIR 方法提高 0.8 ~ 1.59 dB,比 IDW 方法提高 1.35 ~ 1.75 dB。仿真结果说明本文方法在相同比特率的情况下,在不同丢包率的信道传输中均更好地保护了会话视频中的感兴趣区,具有更好的抵抗差错性能。

表 2 目标比特分配方法性能

丢包率/ %	foreman (150 Kb/s)		KristenAndSara (1.5 Mb/s)	
	R_e /%	ROI Δ PSNR/dB	R_e /%	ROI Δ PSNR/dB
5	1. 24	1. 5	3. 85	1. 17
10	0. 4	2. 67	0. 25	2. 32
15	1. 13	2. 75	4. 02	2. 56
20	2. 53	2. 86	1. 42	2. 78

从表 2 可看出,本文方案中的目标比特分配方法取得预期性能。在实际比特率与目标比特率保持基本一致的情况下,将更多比特分配给了感兴趣区域,使得 Foreman 序列的感兴趣区的重建质量提升 1.5 ~ 2.86 dB,而 KristenAndSara 序列则提升 1.17 ~ 2.78 dB。实验结果说明随着信道丢包率的增加,更多的比特被分配到感兴趣区域,感兴趣区的视频质量随之提高。

4 结论

针对会话视频和 HVS 特点,本文提出基于感兴趣区的 HEVC 会话视频容错编码方案。在文中结合人眼特性对会话视频中的人脸区域进行快速分割,并且考虑到信道状况对人眼敏感度的影响修正了人眼兴趣度量,接着根据感兴趣区域简化编码中 CTU 的划分模式,结合信道状况给出基

于感兴趣区的帧内帧间模式选择方法和目标比特分配方法。仿真结果表明,该方案在保持比特率基本不变的同时,能够提升感兴趣区内抗差错性能并且比标准编码节省 20% ~ 40% 的编码时间。下一步的工作,将着重于搭建和实现基于此方案的移动会话视频系统,测试其在实际应用中的性能并加以改善。

参考文献

[1] Sullivan G J, Ohm J, Han 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): 1 649-1 668.

[2] Zhang R, Regunathan S L, Rose K. Video coding with optimal inter/intra-mode switching for packet loss resilience [J]. IEEE Journal on Selected Areas in Communications, 2000, 18(6): 966-976.

[3] Stockhammer T, Hannuksela M M, Wiegand T. H. 264/AVC in wireless environments [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 657-673.

[4] Yuan Z, Wen G, Yan L, et al. Joint source-channel rate-distortion optimization for H. 264 video coding over error-prone networks [J]. IEEE Transactions on Multimedia, 2007, 9 (3): 445-454.

[5] Zhifeng C, Dapeng W. Prediction of transmission distortion for wireless video communication; analysis [J]. IEEE Transactions on Image Processing, 2012, 21(3): 1 123-1 137.

[6] Kulupana G, Talagala D S, Arachchi H K, et al. Error resilience aware motion estimation and mode selection for

HEVC video transmission [C] // 2016 IEEE International Conference on Consumer Electronics (ICCE): IEEE, 2016: 85-86.

[7] Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG. H. 264 reference software JM [CP/OL]. (2016-07-15) [2017-04-20]. <http://iphome.hhi.de/suehring/tml/>.

[8] Nyamweno S, Satyan R, Labeau F. Intra-distance derived weighted distortion for error resilience [C] // 2009 16th IEEE International Conference on Image Processing (ICIP): IEEE, 2009: 1 057-1 060.

[9] Chen Q, Chen Z, Gu X, et al. Attention-based adaptive intra refresh for error-prone video transmission [J]. IEEE Communications Magazine, 2007, 45(1): 52-60.

[10] Maung H M, Aramvith S, Miyana Y. Region-of-interest based error resilient method for HEVC video transmission [C] // 2015 15th International Symposium on Communications and Information Technologies (ISCIT): IEEE, 2015: 241-244.

[11] Meddeb M, Cagnazzo M, Pesquet-Popescu B. ROI-based rate control using tiles for an HEVC encoded video stream over a lossy network [C] // 2015 IEEE International Conference on Image Processing (ICIP): IEEE, 2015: 1 389-1 393.

[12] Joint Collaborative Team on Video Coding (JCT-VC). HEVC reference software HM (HEVC test model) [CP/OL]. (2013-04-01) [2016-07-15]. <https://hevc.hhi.fraunhofer.de/>.

[13] Li B, Li H, Li L, et al. λ domain rate control algorithm for high efficiency video coding [J]. IEEE Transactions on Image Processing, 2014, 23(9): 3 841-3 854.

[14] Hsu R L, Abdel-Mottaleb M, Jain A K. Face detection in color images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 696-706.