

考虑全帧间参考依赖性的监控视频 高效视频编码码率控制算法

公衍超^{1,2}, 吕显龙¹, 杨楷芳³, 刘颖^{1,2}, 林庆帆^{1,4}, 王富平^{1,2}

(1. 西安邮电大学通信与信息工程学院, 710121, 西安; 2. 西安邮电大学电子信息现场勘验应用技术公安部重点实验室, 710121, 西安; 3. 陕西师范大学计算机科学学院, 710119, 西安; 4. 矽景公司, 787820, 新加坡)

摘要: 针对高效视频编码(HEVC)标准中的码率控制算法在计算目标码率分配权重时未考虑帧间参考依赖性、编码效率低的问题,提出了一种面向监控视频考虑全帧间参考依赖性的 HEVC 图像级码率控制算法。调研了 HEVC 低时延结构中编码帧与其所有帧间参考帧之间的参考依赖关系,在此基础上构建了适用于监控视频的考虑全帧间参考依赖性的失真模型。利用构建的失真模型,基于拉格朗日率失真优化理论,建立了拉格朗日率失真代价函数。求解此代价函数,获得每幅图像对应的优化的目标码率分配权重,提出了适用于监控视频的图像级码率控制算法。实验结果表明,与 HEVC 测试模型 HM 中使用的平均比率目标码率分配权重的码率控制算法相比,所提算法可以获得更加准确的码率估计,并且所提算法的德尔塔码率(ΔR)可平均降低 9.54%。

关键词: 高效视频编码;码率控制;监控视频;全帧间参考依赖性;目标码率分配权重

中图分类号: TN911.7 **文献标志码:** A

DOI: 10.7652/xjtub202103009 **文章编号:** 0253-987X(2021)03-0072-09



OSID 码

Rate Control for High Efficiency Video Coding Coded Surveillance Videos Considering All of the Inter-Frame Reference Dependency

GONG Yanchao^{1,2}, LÜ Xianlong¹, YANG Kaifang³, LIU Ying^{1,2},
LIM Kengpan^{1,4}, WANG Fuping^{1,2}

(1. School of Communications and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; 2. Key Laboratory of Electronic Information Application Technology for Scene Investigation of Ministry of Public Security, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; 3. School of Computer Science, Shaanxi Normal University, Xi'an 710119, China; 4. Siliconvision Pte. Ltd., 787820, Singapore)

Abstract: To solve the problems that rate control in high efficiency video coding (HEVC) standard does not consider the inter-frame reference dependency when calculating the target bitrate allocation weight, and the coding efficiency is low, a picture level rate control method considering all of the inter-frame reference dependency for HEVC coded surveillance videos is proposed. A new distortion model for surveillance video is constructed following the investigation of the reference dependency between the coding frame and all of its referenced inter-frames in the low-delay structure of HEVC. Adopting the Lagrangian rate-distortion (R-D) optimization theory, the Lagrangian R-D cost function is established according to the constructed distortion model. The optimized bit allocation weights for pictures are solved from the R-D cost function,

收稿日期: 2020-09-25。 作者简介: 公衍超(1986—),男,讲师;杨楷芳(通信作者),女,副教授。 基金项目: 国家自然科学基金资助项目(61801381);陕西省国际科技合作与交流计划资助项目(2018KW-003);陕西省科研计划资助项目(18JK0708)。

网络出版时间: 2020-11-19

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20201119.1025.004.html>

and a picture level rate control algorithm for surveillance videos is obtained. Compared with the rate control algorithm in HEVC test model using equal target bitrate allocation weights, the proposed algorithm achieves a more accurate bitrate estimation. Moreover, the Bjøntegaard delta-rate (ΔR) of the proposed algorithm can be reduced by 9.54% on average.

Keywords: high efficiency video coding; rate control; surveillance video; all of the inter-frame reference dependency; target bitrate allocation weight

近年来,随着网络、多媒体及硬件技术的飞速发展,视频监控作为重要的视频应用之一,在公安安全、智能安防、智能家居等领域得到了广泛的应用。在视频监控系统中,码率控制是一项非常关键的技术,其可以在满足信道传输带宽的限制条件下,获得更高质量的编码重建视频。

当前大部分的视频监控系统都属于高清视频监控视频。高清视频可以提供更加丰富更加清晰的视频内容。针对高清视频,ITU-T 的视频编码专家组和 ISO/IEC 的运动图像专家组联合制定了先进的高效视频编码(HEVC)标准^[1-4]。与之前的 H.264/AVC 标准相比,HEVC 采用了许多新技术^[1,4],大大提高了编码效率。在 HEVC 标准化过程中,根据使用的码率 R 与量化参数 Q 的关系模型 R - Q ,码率控制算法可以分为 3 类:基于二次 R - Q 模型的码率控制算法^[5], ρ 域码率控制算法^[6](ρ 表示量化变换系数中取值为 0 的系数的百分比)和 λ 域码率控制算法^[7](λ 表示拉格朗日因子)。与前两类码率控制算法相比, λ 域码率控制算法的编码效率更高,得到了更加广泛的研究。

λ 域码率控制算法在文献[7]中被首次提出。文献[7]指出,相比于传统的量化参数, λ 与码率的关系更加密切,使用 λ 能更加准确地反映视频编码的码率特性。由此,文献[7]建立了以 λ 为纽带的 R - λ - Q 模型,并基于此模型提出了面向 HEVC 的码率控制算法。 λ 域码率控制算法包含两个主要步骤,即目标码率分配和 Q 的确定。首先,根据信道带宽、缓冲器状态等因素为每个基本单元(图像或最大编码单元)分配合适的目标码率。然后,根据 R - λ - Q 模型为每个基本单元确定优化的 Q 。

近几年,一些针对 λ 域码率控制算法的改进算法被相继提出^[8-13]。文献[8-9]面向视频编码中的帧内图像提出了相应的改进算法。文献[8]考虑并量化了帧内图像编码过程中的失真漂移现象,并将其应用于码率控制过程中,建立了更加准确的 R - λ 模型。文献[9]指出,文献[7]提出的算法没有有效考虑图像的内容复杂度,因此采用梯度表征图像的

内容复杂度,构建了基于梯度的 R - λ 模型,用于提升帧内图像的码率控制性能。

在视频编码中,相比于帧内图像,帧间图像的实际应用更加广泛。面向帧内帧间图像混合的编码结构,文献[10-12]提出了改进的码率控制算法。文献[10]调研了 HEVC 随机接入(RA)结构中时间层图像之间的参考依赖性,并将其应用于指导不同时间层图像目标码率分配和量化参数确定步骤中。文献[11]将恰可感知失真作为失真测度,提出了基于恰可感知失真的感知率失真模型 R - D ,并将此模型应用于传统的码率控制框架中以提高编码视频的主观感知质量。文献[12]指出,文献[7]提出的算法在计算目标码率分配权重时没有有效考虑视频的内容特性,因此进一步补充提出了考虑视频内容特性的失真与 λ 的关系模型 D - λ 。基于 R - λ 与 D - λ 模型,文献[12]创建了能够准确反映不同视频内容编码失真和码率特性的 R - D 模型,并进一步提出了更加优化的目标码率分配权重计算算法。目前的实验证明,文献[7,12]的算法具有高效的码率控制性能,已经被 HEVC 建议使用,并且集成到最新版本的 HEVC 测试模型 HM 中。

在 λ 域码率控制算法中,目标码率分配结果直接影响 R - λ - Q 模型中的码率,显著影响码率控制的性能。在进行目标码率分配时,最核心的是要给每个基本单元计算优化的目标码率分配权重。当前, HM 中集成的码率控制算法包含 3 种目标码率分配权重的计算方式,即:平均比率的目标码率分配权重计算、固定比率的目标码率分配权重计算、自适应比率的目标码率分配权重计算。然而,上述 3 种方式在计算目标码率分配权重时都没有考虑图像间编码参考的依赖性。HEVC 和之前大部分的主流视频编码标准一样,都是采用基于预测、变换、量化、熵编码的混合编码框架。视频在编码时会采用帧内预测与帧间预测技术,以消除视频中的空域和时域冗余信息。这导致编码图像与其参考图像间存在非常强的依赖关系。特别是监控视频,背景区域通常近似静止,只有前景中的目标在动,所以图像间内容的相

关性更大。这最终导致相比于其他类型的视频,监控视频中图像间的编码参考依赖性会更大。理论上而言,在设计视频编码的相关算法时,有效地考虑参考依赖性会显著提高视频编码的效率^[13-14]。

本文针对监控视频,提出了一种考虑编码图像与其所有帧间参考图像参考依赖性的目标码率分配权重计算算法,并基于此提出了一种有效的 HEVC 图像级码率控制算法。实验验证了提出算法的有效性,与 HM 中采用的平均比率目标码率分配权重的码率控制算法相比,德尔塔码率 ΔR 平均降低 9.54%,且码率估计更加准确。

1 低时延编码结构

HM 中支持全帧内、低时延(LD)和 RA 共 3 种编码结构用于通常的视频编码测试及实际应用^[15-16]。

图 1 和图 2 分别是 RA 结构和 LD 结构。在两种结构中,视频的第一幅图像被编码为即时解码刷新图像,剩余的图像被编码为通常的 P 或者 B (GPB)图像或 B 图像。在编码时,视频中的图像会被划分到各个图像组(GOP)。在每个 GOP 中,图像被划分到不同的时间层。在图 1 和图 2 中,蓝色、绿色、黄色标记的图像分别属于第 1、2、3 个时间层。

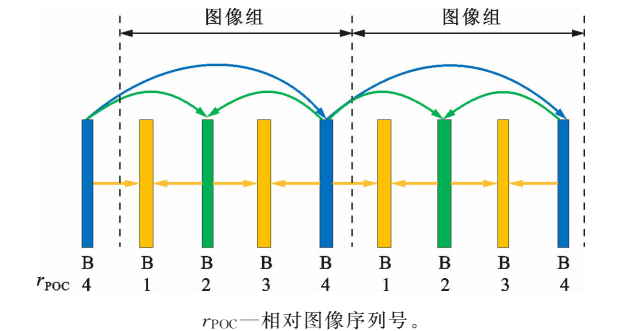


图 1 RA 结构

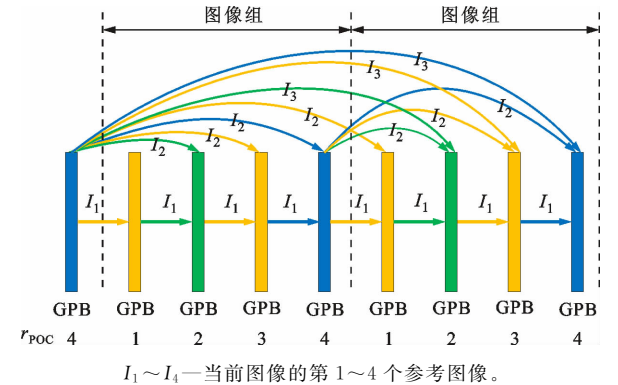


图 2 LD 结构

在图 1 和图 2 中,箭头方向表示参考帧来源。在 RA 结构中,双向预测技术被使用。双向预测技术可以有效地消除视频中的时域冗余信息,提高视频编码的率失真性能,但是也增大了视频编码的结构时延。所以,RA 结构适用于对实时性要求不高的视频应用中,例如视频点播。相对应地,在 LD 结构中为了获得低的结构时延,只有前向预测技术被使用。所以,LD 结构被广泛应用于对实时性要求较高的视频应用中,例如视频会议、视频监控。

HM 使用帧间预测编码图像时会采用多参考图像预测技术,即从多个备选图像中选择最优的匹配信息来源。LD 结构中默认的图像组大小为 4^[16]。HM 中 LD 结构对应的多参考图像来源如表 1 所示,表中 P_{OC} 为图像序列号,视频编码中通常使用 P_{OC} 和 r_{POC} 分别表示图像在视频和 GOP 中的位置。

表 1 HM 中 LD 结构对应的多参考图像来源^[16]

P_{OC}	r_{POC}	I_1	I_2	I_3	I_4
1	1	-1			
2	2	-1	-2		
3	3	-1	-2	-3	
4	4	-1	-2	-3	-4
5	1	-1	-2	-3	-5
6	2	-1	-2	-3	-6
7	3	-1	-2	-3	-7
8	4	-1	-2	-4	-8
9	1	-1	-2	-5	-9
≥ 10	1	-1	-5	-9	-13
	2	-1	-2	-6	-10
	3	-1	-3	-7	-11
	4	-1	-4	-8	-12

P_{OC} 的初始值和最大值分别为 0 和 $N_p - 1$, N_p 表示视频中的图像总数。 r_{POC} 的初始值和最大值分别为 1 和 G_S , G_S 表示 GOP 的大小。表 1 中,用当前图像与其参考图像之间的 P_{OC} 差值指示了参考图像的来源。以第 3 个 GOP 中的第一个图像为例,其 r_{POC} 为 1, P_{OC} 为 x ,则其对应的参考图像序列号是分别是 $x-1$ 、 $x-2$ 、 $x-5$ 、 $x-9$ 。换言之,当前图像的参考图像通常包括与当前图像相邻的前一个已编码的图像和属于最低时间层的其他 3 个最近的已编码图像。特别地,对于视频的前几个图像,当其对应的已编码的图像不足 4 幅时,则参考图像数目为实际已编码图像数目。

2 HM 采用的目标码率分配权重计算

在 HM 码率控制算法完成 GOP 级的目标码率分配后,GOP 对应的目标码率就被确定。紧接着,需要确定 GOP 中的每个图像的目标码率分配权重,以给每个图像分配更加合理的目标码率。通常,重要的图像会被分配更大目标码率分配权重以分配更多的目标码率,从而保证编码质量。

HM 中采用的第一种目标码率分配权重的计算方式是平均比率的目标码率分配权重计算,即 GOP 中所有图像的目标码率分配权重是一样的。这种计算方式显然是不优化的。这是因为视频采用 LD 结构编码时,处在不同时间层的图像的重要性是不一样的。为了进一步提高码率控制性能,HM 也采用

了固定比率的目标码率分配权重计算和自适应比率的目标码率分配权重计算。

固定比率的目标码率分配权重计算首先考虑了图像层的影响。在一个 GOP 中, r_{POC} 为 1、2、3、4 的图像分别处于第 3、2、3、1 个时间层,如图 1 和图 2 所示。对于整个视频的编码率失真性能,通常越低时间层的图像越重要^[13-14],所以越低时间层图像对应的目标码率分配权重应该越大。另外,这种算法也考虑了像素点码率 A 的影响, A 的公式为

$$A = \frac{R_{\text{tar}}}{fwh} \tag{1}$$

式中: R_{tar} 表示编码配置文件中设置的目标码率(b/s); f 、 w 和 h 分别表示视频的帧率、视频图像的宽度和高度; A 表示像素点码率(b/像素)。表 2 为固

表 2 固定比率的目标码率分配权重

r_{POC}	分配权重			
	$A/(b \cdot \text{像素}^{-1}) > 0.2$	$0.2 \geq A/(b \cdot \text{像素}^{-1}) > 0.1$	$0.1 \geq A/(b \cdot \text{像素}^{-1}) > 0.05$	其他
1	2	2	2	2
2	3	3	3	3
3	2	2	2	2
4	6	10	12	14

定比率的目标码率分配权重。

使用固定比率的目标码率分配权重计算算法时,一个视频的所有 GOP 中的图像都按表 2 确定目标码率分配权重。这种算法对于运动慢的视频是比较有效的,可以有效权衡码率控制算法的码率估计准确度、率失真性能和算法复杂度。但是,对于运动快、前后 GOP 图像内容差别较大的视频,固定比率的目标码率分配权重计算算法显然不太适用。相应地,自适应比率的目标码率分配权重计算算法被提出。使用自适应比率的目标码率分配权重计算算法时,视频中的第一个 GOP 的图像仍然按照表 2 确定图像的目标码率分配权重。后续的 GOP 利用前一个 GOP 中最后一个图像的 λ 、预估的 R - λ 模型参数和层级信息动态确定 GOP 中各图像的目标码率分配权重。

通过本节分析可知,视频采用 LD 结构编码时,图像间存在很强的编码参考依赖关系,而 HM 当前采用的目标码率分配权重计算算法虽然考虑了时间层、 A 、 λ 等因素的影响,但忽略了参考依赖关系这一重要因素。因此,本文提出了一种面向视频监控应用的考虑全帧间参考依赖性的目标码率分配权重计算算法。

3 全帧间参考依赖的码率控制算法

本文提出的全帧间参考依赖的码率控制法包括目标码率分配及 Q 的确定两部分。目标码率分配包括 GOP 级的目标码率分配和图像级的目标码率分配。提出算法的 GOP 级的目标码率分配和 Q 确定与 HM 中采用的算法相同,这些算法的详细步骤本文不再详述,具体见参考文献[16]。本文的主要贡献是构建了考虑全参考依赖性的失真模型,并基于此模型提出了优化的图像级目标码率分配权重计算算法。

3.1 考虑全参考依赖性的图像级目标码率分配权重计算

3.1.1 考虑全参考依赖性的失真模型 如第 1 节所述,视频在采用 LD 结构编码时,当前图像与其参考图像之间存在很强的参考依赖关系。当前图像的期望失真为

$$E(D_{i,j}) = \sum_{r=1}^{N_r} P_{i,j,r} D_{i,j}(o_{i,j}^r) \tag{2}$$

式中: $D_{i,j}$ 表示当前第 i 个 GOP 中的第 j 个图像的失真; $E(D_{i,j})$ 表示当前第 i 个 GOP 中的第 j 个图像的期望失真; $P_{i,j,r}$ 表示当前图像的第 r 个参考图像

被选为其最优参考图像的概率; N_r 表示当前图像的参考图像数量, 对于 LD 结构, N_r 的默认取值为 4; $o_{i,j}^r$ 表示当前图像在其第 r 个参考图像的影响下选择的最优编码参数组合; $D_{i,j}(o_{i,j}^r)$ 表示当前图像在其第 r 个参考图像影响下产生的失真。

通过编码实验进一步探索 $D_{i,j}(o_{i,j}^r)$ 的解析表达式。实验的主要软硬件环境为: Windows 10 操作系统, 处理器为 Intel(R) Core(TM) i7-8700 CPU (3.20 GHz), 内存 16.0 GB, 编码器选择 HM16.17, 编码结构为 LD 结构, 主要编码参数设置如表 3 所示。本文后续涉及到的实验环境皆与此实验环境相同。本实验通过调整当前图像和其第一个参考图像的 Q , 并固定其他 3 个参考图像的 Q , 以探索第一个参考图像对于当前图像失真的影响。采用相似的策略依次研究其他 3 个参考图像对于当前图像失真的影响。

表 3 主要编码参数设置

参数	设置
Profile	Main
IntraPeriod	-1
GOPSize	4
FastSearch	1
SearchRange	64
SAO	1
AMP	1

Overbridge 和 Campus 监控视频的 $D_{i,1,1}$ 与 $D_{i,1}(o_{i,1}^1)$ 的关系如图 3 所示。图中, $D_{i,1,1}$ 表示第 i 个 GOP 中第 1 个图像的第 1 个参考图像的失真, $D_{i,1}(o_{i,1}^1)$ 表示第 i 个 GOP 中第 1 个图像在其第 1 个参考图像的影响下产生的失真。失真用像素亮度分量的均方误差表征。采用最小二乘法进行数据拟合, 拟合曲线用实线表示, 拟合精度用相关系数的平方 (r^2) 衡量^[17]。

特别说明, 在 GOP 大小及参考图像数目都默

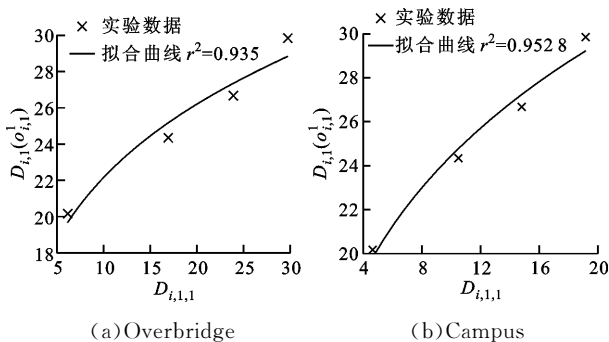


图 3 $D_{i,1}(o_{i,1}^1)$ 与 $D_{i,1,1}$ 的关系

认为 4 的情况下, 一个 GOP 中的所有图像与其对应的各个参考图像的失真关系图一共有 16 幅。由于文章篇幅有限, 图 3 只给了部分图像的结果, 其他图像的拟合曲线也都满足图 3 所示的幂函数形式, 但是幂函数中的两个模型参数取值有差别。最终, 图 3 所示的关系可以表征为

$$D_{i,j}(o_{i,j}^r) = a_{i,j,r} D_{i,j,r}^{b_{i,j,r}} \quad (3)$$

式中 $a_{i,j,r}$ 和 $b_{i,j,r}$ 是模型参数。将式 (3) 代入式 (2), 可得

$$E(D_{i,j}) = \sum_{r=1}^{N_r} P_{i,j,r} a_{i,j,r} D_{i,j,r}^{b_{i,j,r}} \quad (4)$$

3.1.2 目标码率分配权重计算 根据拉格朗日率失真优化^[18], 编码视频总的率失真代价为

$$J = \sum_{i=1}^{N_g} \sum_{j=1}^{G_s} (E(D_{i,j}) + \lambda_{i,j} R_{i,j}) \quad (5)$$

式中: $\lambda_{i,j}$ 和 $R_{i,j}$ 分别表示第 i 个 GOP 中第 j 个图像的拉格朗日因子和码率; N_g 表示视频中 GOP 的数量。

文献[17]的研究表明, 幂函数形式的率失真关系可以准确地描述 HM 编码器的率失真特性, 具体公式为

$$D = CR^{-K} \quad (6)$$

式中 C 和 K 是与序列内容特征相关的模型参数。因此, 拉格朗日因子被进一步求解为

$$\lambda = -\partial D / \partial R = CKR^{-K-1} = \alpha R^\beta \quad (7)$$

式中 α 和 β 是与视频内容特性相关的模型参数。

在第 3.1.1 小节描述的实验中, 不仅可以获得失真数据, 还可以获得相应的码率数据。Overbridge 和 Campus 监控视频的 $R_{i,4,1}$ 与 $R_{i,4}$ 的关系如图 4 所示, $R_{i,4,1}$ 与 $R_{i,4}$ 用像素点码率表征。

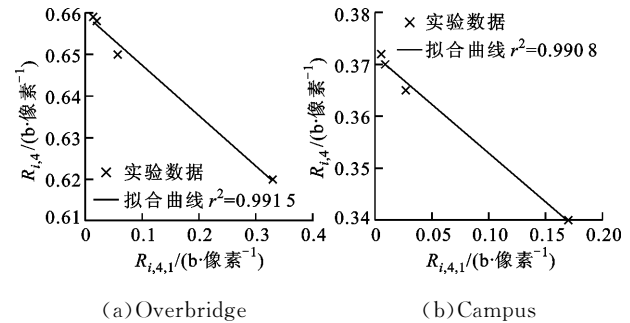


图 4 $R_{i,4,1}$ 与 $R_{i,4}$ 的关系

图 4 所示的关系可表征为

$$R_{i,j,r} = d_{i,j,r} R_{i,j} \quad (8)$$

式中: $R_{i,j,r}$ 是第 i 个 GOP 中第 j 个图像的第 r 个参考图像的码率; $d_{i,j,r}$ 是模型参数。将式 (4) (6) (7)

(8)代入式(5),得到

$$J = \sum_{i=1}^{N_g} \sum_{j=1}^{G_s} \left(\sum_{r=1}^{N_r} (P_{i,j,r} a_{i,j,r} C_{i,j,r}^{b_{i,j,r}} \cdot (d_{i,j,r} R_{i,j})^{-b_{i,j,r} K_{i,j,r}}) + C_{i,j} K_{i,j} R_{i,j}^{-K_{i,j}} \right) \quad (9)$$

将式(9)对 $R_{i,j}$ 求偏导,然后让偏导数 $\partial J / \partial R_{i,j}$ 为 0,从而获得使率失真代价 J 最小的 $R_{i,j}$ 的优化值 $R_{i,j,\text{opt}}$ 。根据牛顿-拉斐逊方法^[19],求得使 $\partial J / \partial R_{i,j} = 0$ 的 $R_{i,j,\text{opt}}$ 的近似解为

$$R_{i,j,\text{opt}} \approx \theta_{i,j} + \left[C_{i,j} \theta_{i,j}^{-K_{i,j}-1} + \sum_{r=1}^{N_r} (P_{i,j,r} a_{i,j,r} b_{i,j,r} K_{i,j,r} C_{i,j,r}^{b_{i,j,r}} d_{i,j,r}^{-b_{i,j,r} K_{i,j,r}} \theta_{i,j}^{-b_{i,j,r} K_{i,j,r}-1} + \theta_{i,j}^{-b_{i,j,r} K_{i,j,r}-1}) \right] \left[C_{i,j} K_{i,j}^2 (K_{i,j} + 1) \theta_{i,j}^{-K_{i,j}-2} + \sum_{r=1}^{N_r} P_{i,j,r} a_{i,j,r} b_{i,j,r} K_{i,j,r} (b_{i,j,r} K_{i,j,r} + 1) C_{i,j,r}^{b_{i,j,r}} d_{i,j,r}^{-b_{i,j,r} K_{i,j,r}-2} \theta_{i,j}^{-b_{i,j,r} K_{i,j,r}-2} \right]^{-1} \quad (10)$$

式中 $\theta_{i,j}$ 为牛顿-拉斐逊参数,公式为

$$\theta_{i,j} = \frac{\chi_{i,j} R_{\text{tar}}}{fwh} \quad (11)$$

其中 $\chi_{i,j}$ 是伸缩因子,具体取值如表 4 所示。

表 4 $\chi_{i,j}$ 取值

$A/(b \cdot \text{像素}^{-1})$	$\chi_{i,j}$			
	$j=1$	$j=2$	$j=3$	$j=4$
$(0.2, +\infty)$	1.0	1.5	1.0	3.0
$(0.1, 0.2]$	1.0	1.5	1.0	4.5
$(0.05, 0.1]$	1.0	1.5	1.0	7.0
$(0, 0.05]$	1.0	1.5	1.0	10.0

式(10)中的牛顿-拉斐逊参数表示图像分配的目标码率初始值,且牛顿-拉斐逊参数越大,给图像分配的目标码率通常就越多。而且,式(11)中的伸缩因子与牛顿-拉斐逊参数是正相关的。所以,图像的伸缩因子越大,给图像分配的目标码率就越大。由于在 LD 结构中,时间层越低的图像对于整个视频编码性能的影响越大^[20],所以在编码时要优先保证低时间层图像的质量,给低时间层图像分配更多的目标码率。所以,表 4 中随着时间层的降低,图像对应的伸缩因子取值是增大的趋势。

一些实验^[16]证明,在目标码率不充裕时,若要保证整个视频的编码效率,则需要优先保证低时间层图像的编码质量,即给低时间层图像分配更大比例的目标码率。当目标码率充裕时,低时间层图像

的编码质量已经很好,给其分配更多的码率对于整个视频编码效率的提升也已经不显著了,这时把更多比例的码率分配给其他更高时间层图像,反而会带来整个视频编码效率的明显提升。所以,表 4 中随着目标码率的增大,时间 1 层图像对应的伸缩因子与所有时间层图像对应的伸缩因子和的比值是减少的,而其他更高时间层图像对应的伸缩因子与所有时间层图像对应的伸缩因子和的比值是增大的。表 4 中的伸缩因子的具体取值是以固定比率目标码率分配权重计算算法中的伸缩因子取值为基础,通过实验调整后确定的优化的经验值。

最终, GOP 中每个图像的目标码率分配权重为

$$\omega_{i,j} = \langle R_{i,j,\text{opt}} / \sum_{j=1}^{G_s} R_{i,j,\text{opt}} \rangle \quad (12)$$

式中 $\langle \rangle$ 表示四舍五入运算。

3.2 全帧间参考依赖的码率控制算法步骤

为了便于描述, HM 中采用的基于平均比率目标码率分配权重的图像级码率控制算法、基于固定比率目标码率分配权重的图像级码率控制算法和基于自适应比率目标码率分配权重的图像级码率控制被分别简称为 EBA-PRC、FRBA-PRC 和 ARBA-PRC。本文提出的码率控制算法的完整步骤如下。

第 1 步 对于给定的监控视频,使用 ARBA-PRC 对前 4 个 GOP 的图像进行编码,获得其编码失真和码率数据。根据式(3)使用最小二乘拟合算法,拟合失真数据获得 $a_{i,j,r}$ 和 $b_{i,j,r}$ 。同理,分别根据式(6)和式(8),采用最小二乘法拟合失真和码率数据得到 $C_{i,j}$ 、 $K_{i,j}$ 、 $C_{i,j,r}$ 、 $K_{i,j,r}$ 、 $d_{i,j,r}$ 。在对相应图像进行编码后, $P_{i,j,r}$ 的准确数据可以直接统计得到。

第 2 步 根据式(10)~(12),计算获得 GOP 中每个图像对应的目标码率分配权重,即 $\omega_{i,j}$ 。

第 3 步 使用第 2 步计算得到的 $\omega_{i,j}$ 对视频后续 GOP 的图像进行码率控制。

4 实验与分析

目前,文献[7,12]提出的码率控制算法由于其高效的编码性能已经被 HEVC 建议使用,并被集成到 HM 中(包括 EBA-PRC、FRBA-PRC 和 ARBA-PRC 算法)。本文算法将与这 3 种算法进行比较。

4.1 实验设置

测试视频包括 6 个不同分辨率和内容特征的标准监控视频,其中: Overbridge、Crossroad、Campus 视频的空间分辨率是 720×576 像素,帧率是 30 帧/s,编码帧数是 300; Lowlight 视频的空间分辨率是

1 280×720 像素,帧率是 30 帧/s,编码帧数是 200; Metro、Runners 视频的空间分辨率是 1 920×1 080 像素,帧率分别是 60 和 30 帧/s,编码帧数分别是 188 和 300。编码器为 HM16.17,编码结构为 LD 结构。首先,将 HM 中的码率控制功能关闭,使用 HM 中固定的 Q 分配方式对每个测试视频进行编码以获得从高到低的 4 个目标码率^[10]。固定的 Q 分配方式中帧内图像的 Q 被分别设定为 22、27、32、37。然后,将 HM 中的码率控制功能打开,在给定的 4 个目标码率下,分别用 EBA-PRC、FRBA-PRC、ARBA-PRC 和本文提出的算法编码每一个测试视频。其他的主要编码参数设置与 HM 配置文件 encoder_lowdelay_P_main^[16]中的设置相同。

4.2 码率估计准确度和率失真性能

码率估计准确度是衡量一个码率控制算法性能关键的基础性指标。只有在码率估计准确的基础上再分析码率控制算法的其他性能才是有意义的。本文采用被广泛使用的 $BE(B_E)^{[7,12]}$ 来衡量码率估计的准确度,公式为

$$B_E = |B_{tar} - B_{act}| / B_{tar} \times 100\% \quad (13)$$

式中 B_{tar} 和 B_{act} 分别表示配置文件中设定的目标码率和码率控制算法的实际码率(kb/s)。

表 5 为各种码率控制算法的 B_E 。 B_E 越小,对应的码率控制算法获得的码率估计误差越小,即码率估计更加准确。可以看出,对于大部分视频,本文算法都可以获得最小的 B_E ,且对于所有的视频,提出算法对应的平均 B_E 也是最小的,为 0.223 2%。所以,相较其他 3 种算法,本文算法的码率估计更加准确。

表 5 不同码率控制算法的 B_E %

序列	EBA-PRC	FRBA-PRC	ARBA-PRC	本文算法
Overbridge	0.271 1	0.262 3	0.270 5	0.224 9
Crossroad	0.115 4	0.118 7	0.121 0	0.119 8
Campus	0.475 4	0.455 9	0.473 0	0.471 3
Lowlight	0.433 8	0.437 0	0.435 5	0.433 6
Metro	0.073 4	0.071 9	0.072 9	0.068 4
Runners	0.034 2	0.030 2	0.030 0	0.021 0
均值	0.233 9	0.229 3	0.233 8	0.223 2

采用被广泛使用的 $\Delta R^{[21]}$ 衡量码率控制算法的率失真性能, ΔR 为负值时表示与基准算法相比,比较算法可以用更少的码率获得同等质量的编码重建视频,即比较算法的率失真性能越好。本文实验在

计算 ΔR 时将 EBA-PRC 算法设定为基准算法,其中视频质量用像素亮度分量的峰值信噪比衡量。

表 6 为 4 种码率控制算法对应的 ΔR 。可以看出:对于所有的视频,本文算法对应的 ΔR 最小,即率失真性能更好;与 EBA-PRC 相比,对于所有视频,本文算法的 ΔR 为-9.54%,即本文算法可以在获得同等质量重建视频的前提下,码率减少 9.54%;FRBA-PRC 算法与 ARBA-PRC 算法的性能近似,FRBA-PRC 可以获得更加准确的码率估计,而 ARBA-PRC 可以获得更高的率失真性能。

表 6 不同码率控制算法的 ΔR %

序列	EBA-PRC	FRBA-PRC	ARBA-PRC	本文算法
Overbridge	0	-10.97	-11.43	-12.40
Crossroad	0	-4.26	-4.18	-4.66
Campus	0	-7.63	-8.49	-8.87
Lowlight	0	9.21	3.77	-3.10
Metro	0	-12.09	-11.77	-12.54
Runners	0	-15.40	-15.37	-15.66
均值	0	-6.86	-7.91	-9.54

4.3 算法复杂度

采用编码时间衡量算法的运行复杂度,公式为

$$\Delta T = (T_{com} - T_{bac}) / T_{bac} \times 100\% \quad (14)$$

式中: T_{bac} 表示 EBA-PRC 算法对应的编码时间; T_{com} 表示其他码率控制算法对应的编码时间。

表 7 为 4 种码率控制算法对应的 ΔT 。可以看出:与 EBA-PRC 算法相比,FRBA-PRC、ARBA-PRC、本文算法的编码时间都更短,且 ΔT 取值近似,对于所有测试视频,平均 ΔT 分别为-0.03%、-0.04%、-0.02%。

表 7 不同码率控制算法的 ΔT %

序列	EBA-PRC	FRBA-PRC	ARBA-PRC	本文算法
Overbridge	0	-0.04	-0.04	-0.02
Crossroad	0	-0.03	-0.04	-0.03
Campus	0	-0.03	-0.03	-0.01
Lowlight	0	-0.03	-0.03	0.001
Metro	0	-0.04	-0.04	-0.04
Runners	0	-0.03	-0.03	-0.02
均值	0	-0.03	-0.04	-0.02

从目标码率分配权重计算的角度分析,EBA-PRC 算法和 FRBA-PRC 算法对应的目标码率分配

权重的计算过程是很简单的,而 ARBA-PRC 算法与本文算法在计算图像的目标码率分配权重时都需要更加复杂的参数更新与优化过程,所以复杂度相对较高。但是,相较于 EBA-PRC 算法,FRBA-PRC、ARBA-PRC、本文算法都会增大高时间层图像的 Q ,且高时间层图像数占视频总图像数的比例更大。一些编码实验^[10]已经证明,使用大的 Q 编码图像会显著降低图像的编码时间。综上所述,虽然 EBA-PRC 算法中简单的目标码率分配权重计算会节约一部分编码时间,但是由于会给高时间层图像选择更大的 Q ,所以总的编码时间是最长的。

5 结 论

在图像级码率控制算法中,优化的图像目标码率分配权重是影响算法性能的关键因素。HEVC 标准中的码率控制算法在计算目标码率分配权重时未考虑帧间参考依赖性,编码效率低。本文针对 HEVC 编码的监控视频,提出了一种有效的考虑全帧间参考依赖性的图像级码率控制算法。本文得出的主要结论及对未来工作的展望如下。

(1)在监控视频 HEVC 码率控制中,考虑当前帧与其所有参考帧之间的参考依赖关系可以构建更加准确地反映编码特性的率失真代价函数,显著提高码率控制算法的性能。

(2)本文的考虑全参考依赖关系的思路也可以被拓展应用到针对 RA 结构的码率控制算法设计中。RA 结构与本文研究的 LD 结构类似,也采用了多参考帧预测技术,但其新增加了后向的参考帧来源。

(3)本文的考虑全参考依赖关系的思路也可以应用到针对其他类型视频的码率控制算法设计中,例如电影电视视频。但是,其他类型的视频通常会包括运动快的视频。针对运动快的视频,本文算法中提出的一些模型的具体形式、模型参数的获得算法及算法步骤的设计都需要进一步进行性能验证或做相应的调整。

参考文献:

- [1] 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.
- [2] 周小鹏, 梁峰, 李冰. 高吞吐率的高效视频编码编

码并行硬件架构设计 [J]. 西安交通大学学报, 2020, 54(7): 180-186.

ZHOU Xiaopeng, LIANG Feng, LI Bing. A high-throughput parallel very-large-scale integration architecture of entropy coding in high efficiency video coding [J]. Journal of Xi'an Jiaotong University, 2020, 54(7): 180-186.

- [3] 李维, 张和仙, 杨付正. 高效率视频编码快速模式判决算法 [J]. 西安交通大学学报, 2013, 47(8): 104-109, 140.

LI Wei, ZHANG Hexian, YANG Fuzheng. A fast mode decision algorithm for high efficiency video coding [J]. Journal of Xi'an Jiaotong University, 2013, 47(8): 104-109, 140.

- [4] LIM Kengpan. The new trend in video coding [J]. Journal of Xi'an University of Posts and Telecommunications, 2013, 18(3): 1-6.

- [5] CHOI H, NAM J, YOO J, et al. Rate control based on unified RQ model for HEVC [EB/OL]. (2012-02-10) [2020-08-01]. <http://phenix.int-evry.fr/jct/JCTVC-H0213.doc>.

- [6] WANG Shanshe, MA Siwei, WANG Shiqi, et al. Rate-GOP based rate control for high efficiency video coding [J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6): 1101-1111.

- [7] LI Bin, LI Houqiang, LI Li, et al. λ domain rate control algorithm for high efficiency video coding [J]. IEEE Transactions on Image Processing, 2014, 23(9): 3841-3854.

- [8] 李维, 任鹏, 赵凡, 等. 高效视频编码中帧内码率控制改进算法 [J]. 西安交通大学学报, 2019, 53(4): 79-84.

LI Wei, REN Peng, ZHAO Fan, et al. A modified method for intra-coding rate control in high efficiency video coding [J]. Journal of Xi'an Jiaotong University, 2019, 53(4): 79-84.

- [9] WANG Miaohui, NGAN K N, LI Hongliang. An efficient frame-content based intra frame rate control for high efficiency video coding [J]. IEEE Signal Processing Letters, 2015, 22(7): 896-900.

- [10] GONG Yanchao, WAN Shuai, YANG Kaifang, 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.

- [11] ZHOU Mingliang, WEI Xuekai, KWONG S, et al. Just noticeable distortion-based perceptual rate control in HEVC [J]. IEEE Transactions on Image Process-

- ing, 2020, 29: 7603-7614.
- [12] LI Li, LI Bin, LI Houqiang, et al. λ -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.
- [13] GONG Yanchao, WAN Shuai, YANG Kaifang, 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.
- [14] GAO Yanbo, ZHU Ce, LI Shuai, et al. Temporally dependent rate-distortion optimization for low-delay hierarchical video coding [J]. IEEE Transactions on Image Processing, 2017, 26(9): 4457-4470.
- [15] BOSSEN F. Common HM test conditions and software reference configurations [EB/OL]. (2013-01-23) [2020-08-01]. <http://phenix.int-evry.fr/jct/JCTVC-L1100.doc>.
- [16] SUEHRING K, SHARMAN K. HM software [EB/OL]. (2014-08-13)[2020-08-01]. <https://hevc.hhi.fraunhofer.de/svn/svnHEVCSoftware/tags>.
- [17] LI B, LI H, LI L, et al. Rate control by R -lambda model for HEVC [EB/OL]. (2012-10-19) [2020-08-01]. <http://phenix.int-evry.fr/jct/JCTVC-K0103.doc>.
- [18] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression [J]. IEEE Signal Processing Magazine, 1998, 15(6): 74-90.
- [19] BEN-ISRAEL A. A Newton-Raphson method for the solution of systems of equations [J]. Journal of Mathematical Analysis and Applications, 1966, 15(2): 243-252.
- [20] YANG Kaifang, GONG Yanchao, MA Miao, et al. An efficient rate-distortion optimization method for low-delay configuration in H. 265/HEVC based on temporal layer rate and distortion dependence [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(4): 1230-1236.
- [21] PATEUX S, JUNG J. An excel add-in for computing Bjøntegaard metric and its evolution [EB/OL]. (2007-01-14) [2020-08-01]. https://www.itu.int/wftp3/av-arch/video-site/0701_Mar/VCEG-AE07.doc.

[本刊相关文献链接]

- 周小鹏,梁峰,李冰.高吞吐率的高效视频编码熵编码并行硬件架构设计.2020,54(7):180-186.[doi:10.7652/xjtuxb202007021]
- 王鹏,孙梦宇,王海燕,等.结合自适应空间权重的改进型时空正则项跟踪算法.2020,54(5):158-169.[doi:10.7652/xjtuxb202005021]
- 马明阳,梅少辉,万帅.采用非线性块稀疏字典选择的视频总结.2019,53(5):142-148.[doi:10.7652/xjtuxb201905019]
- 李维,任鹏,赵凡,等.高效视频编码中帧内码率控制改进算法.2019,53(4):79-84.[doi:10.7652/xjtuxb201904012]
- 袁帅,秦贵和,晏婕.应用残差生成对抗网络的路况视频帧预测模型.2018,52(10):146-152.[doi:10.7652/xjtuxb201810020]
- 赵敏,张为,王鑫,等.时空背景模型下结合多种纹理特征的烟雾检测.2018,52(8):67-73.[doi:10.7652/xjtuxb201808011]
- 王晓芳,齐春.一种运用显著性检测的行为识别方法.2018,52(2):024-29.[doi:10.7652/xjtuxb201802004]
- 杨源,库涛,查宇飞,等.快速多特征金字塔的尺度目标跟踪方法.2016,50(10):49-56.[doi:10.7652/xjtuxb201610008]
- 张泰,张为,刘艳艳.周界视频监控中人员翻越行为检测算法.2016,50(6):47-53.[doi:10.7652/xjtuxb201606008]

(编辑 陶晴)