

视频编码率失真优化技术研究综述

郭红伟^{1,2}, 朱 策¹, 刘宇洋¹

(1. 电子科技大学信息与通信工程学院, 四川成都 611731; 2. 红河学院工学院, 云南蒙自 661100)

摘 要: 率失真优化技术能有效地提升编码器的压缩性能, 是视频编码领域重要的研究内容. 本文综合论述了基于拉格朗日乘子的率失真优化方法, 从独立率失真优化、依赖率失真优化和码率控制中的比特分配等几方面归纳和分析最新研究进展. 首先, 通过形象的举例解释了约束优化问题通过拉格朗日乘法转换为非约束优化问题, 并根据相关文献分析拉格朗日乘子与量化参数的关系. 其次, 分析了编码过程中存在的率失真依赖性, 并分类讨论近年提出的依赖率失真优化方法. 然后, 简要介绍了率失真依赖关系在码率控制比特分配中的影响. 最后通过分析比较不同率失真优化技术的特征和性能, 提出当前面临的挑战和值得进一步探索的工作.

关键词: 视频编码; 率失真优化; 时域依赖; 空域依赖; 自适应拉格朗日乘子

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2020)05-1018-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2020.05.024

Overview of Rate-Distortion Optimization for Video Coding

GUO Hong-wei^{1,2}, ZHU Ce¹, LIU Yu-yang¹

(1. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China;

2. School of Engineering, Honghe University, Mengzi, Yunnan 661100, China)

Abstract: Rate-distortion optimization (RDO) can effectively improve the compression performance for encoders, which is an important topic in video coding. In this paper, the Lagrange multiplier based RDO method is reviewed comprehensively from several aspects including independent RDO, dependent RDO, bit allocation for rate control, etc. First, an example is employed to explain how the constrained optimization problem is converted into an unconstrained optimization problem based on the Lagrange multiplier method, and the relationship between Lagrange multiplier and quantization parameter is analyzed according to relevant references. Second, the rate-distortion dependencies in the process of video coding are analyzed, and the dependent RDO methods presented in recent years are classified and discussed. Third, the effect of the rate-distortion dependencies on bit allocation for rate control is discussed. Finally, the features and performance of different RDO methods are tabulated for analysis and comparison, and the existing challenges and the research works deserving further exploration are presented.

Key words: video coding; rate-distortion optimization; temporal dependency; spatial dependency; adaptive Lagrange multiplier

1 引言

网络信息时代, 图像、视频已成为信息的主要载体. 未经处理的原始数字视频数据量非常巨大, 为了存储和传输, 必须进行压缩, 即视频编码. 为有效地去除视频数据中的各种冗余, 当前视频编码标准, 如 H. 264/AVC、H. 265/HEVC、AVS、VP9 等, 均采用一个包括预测、变

换、量化和熵编码等多种压缩工具的混合视频编码框架. 率失真优化 (Rate-Distortion Optimization, RDO)^[1] 技术在视频编码中扮演着非常重要的角色, 编码器通过 RDO 为输入视频选择一组最优的编码参数, 具体的编码参数包括编码单元的划分模式、预测模式、变换模式、量化系数等. 但是, 率失真优化技术不属于视频编码标准化的范畴, 编码器可以采用不同的优化方案以达到

收稿日期: 2019-04-23; 修回日期: 2019-11-21; 责任编辑: 孙瑶

基金项目: 国家自然科学基金 (No. U19A2052, No. 61571102); 四川省科学技术厅重点项目 (No. 2018JY0035); 云南省地方本科高校基础研究联合专项资金 (No. 2018FH001-056)

编码性能最优。

率失真理论源于香农在 1959 年发表的“保真度准则下的离散信源编码定理”^[2],是视频编码的理论基础。在编码器中,率失真优化技术贯穿着整个视频编码系统,其涉及到建立参考帧选择方式^[3,4]、设计加权量化矩阵、选择最优运动矢量、编码模式以及编码复杂度优化^[5]等,长期以来,得到广泛研究。视频编码的目标是在有限码率约束下,使压缩视频失真最小。因此,率失真优化另一重要的作用是优化编码资源分配问题。本文从编码资源分配方面概述了视频编码中的率失真优化方法,从独立 RDO、依赖 RDO、以及码率控制中的比特分配等归纳和分析最新研究进展。

2 视频编码中的率失真优化

视频编码过程将输入视频序列划分为许多子任务,可以是一个图像组(Group Of Picture, GOP)、一帧图像或一个基本编码单元(Coding Unit, CU)。特定的视频编码器中,在有限码率约束下如何优化比特资源分配决定了编码视频的率失真性能,该优化目标是在给定码率下最小化编码失真^[6],即

$$\min \sum_{i=1}^N D_i, \quad \text{s. t.} \quad \sum_{i=1}^N R_i < R_c \quad (1)$$

式(1)中: R_i 和 D_i 分别表示编码单元 U_i 的编码码率和对应的失真; N 是编码单元数; R_c 表示总码率预算。采用拉格朗日乘子法,式(1)的约束优化问题可以转换为非约束的最优化问题:

$$\min \{J\}, \quad J = \sum_{i=1}^N D_i + \lambda \sum_{i=1}^N R_i \quad (2)$$

式(2)中: J 称作率失真代价函数; λ 称为拉格朗日乘子,较小的 λ 使失真较小但会消耗较多的比特数,反之则失真较大而消耗比特数较少。式(2)中并没有消耗比特数量的限制,但一个特定的 λ 值得到的解将对应于一个 R_c 值时式(1)的解。为了简化问题描述,假设每个子任务相互独立,即编码单元之间的码率和失真互不相关,则式(2)可以分解为分别最小化每个子任务的率失真代价函数:

$$\min \{J_i = D_i + \lambda_i R_i\} \quad (3)$$

为了能形象地描述用拉格朗日乘子法求解最优化问题,举一个简单的例子:小明是大学一年级学生,第一学期他所能花在历史和电路基础两门课程的学习时间是 30 天,他打算合理分配时间以取得最好的成绩,或者说使两门课程的总失分最小。假设每门课程考试的失分与所投入的学习时间满足图 1 所示的关系,随着投入时间的增加,考试失分减少。一种时间分配策略是学习电路基础用 20 天,学习历史投入 10 天,如图中点 C_1 和 H_1 所示。这样电路基础考试得 80 分,历史考试得 70 分,

两门课总共得 150 分。图中, C_1 点斜率的绝对值是 2 分/天, H_1 点斜率的绝对值是 6 分/天。此时,如果小明对电路基础的投入时间减少 1 天以用于学习历史,那么两门课的总分就能增加 4 分。若持续地减少电路基础的学习时间以增加学习历史的时间,直到两边的斜率相等,如图 1 中的点 C_2 和 H_2 所示,如此的方案即为 30 天约束下,小明能取得最优成绩的学习时间分配方式,这时的总成绩为 $70 + 90 = 160$ 分,点 C_2 和 H_2 斜率绝对值即为 30 天对应的拉格朗日乘子。

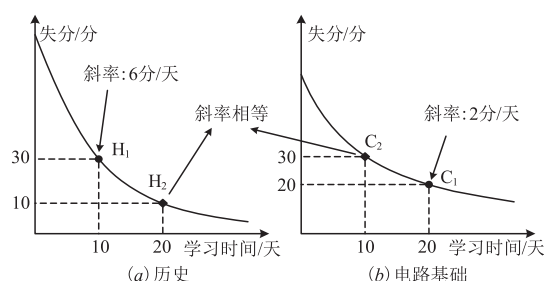


图1 拉格朗日乘子法举例(第一学期)

同理,在视频编码中假设每个编码单元相互独立,则所有编码单元采用相同的拉格朗日乘子最小化各自的率失真代价函数时,可以达到最优的编码性能。因此,确定一个适合的拉格朗日乘子在视频编码中显得非常重要。

3 独立 RDO

为了改善率失真优化在视频编码过程中的可操作性,一种可行的方法是独立地优化每一个编码单元,而不考虑编码单元之间的率失真依赖关系及编码单元内不同模块间的相互影响,被称为独立率失真优化。在视频编码标准 H. 264 联合模型(Joint Model, JM)和 HEVC 测试模型(HEVC test Model, HM)中,率失真优化遍历一个编码单元可选的编码模式,如块划分模式、帧内或帧间预测模式,并根据给定的量化参数(Quantization Parameter, QP)对每一编码模式产生的预测残差进行变换、量化、熵编码,然后计算每一个编码模式的率失真代价,最终具有最小率失真代价的编码模式即为最佳的编码模式。可见,QP 是视频编码中另一重要的参数,QP 值越大则对变换系数的量化尺度越大,相应的编码失真也越大,反之则编码失真较小而编码码率增加。在高码率条件下,可通过率失真理论推导出量化步长 Δ 与 λ 的关系^[1]为

$$\lambda = C \cdot \Delta^2 \quad (4)$$

式(4)中: C 为一个常数; Δ 与 QP 满足一一对应的关系。Thomas Wiegand 等^[7]以 H. 263 的参考软件为平台,通过实验统计也得出了式(4)的 Δ 与 λ 关系,并确定 C 的取值为 0.85。之后,这一关系式被延用至 H. 264 和

HEVC 以及其它标准的视频编码器中,只是 Δ 与 QP 的关系及参数 C 随不同标准有一些相应改变.

3.1 QP- λ 关系的研究

视频编码中,通过引入拉格朗日乘子简化了率失真优化过程,然而基于式(4),QP- λ 关系的拉格朗日乘子确定方法还有需进一步考虑的问题. 首先,这里的拉格朗日乘子只与量化参数 QP 相关,相同 QP 设置下不同的输入视频仍采用同一拉格朗日乘子进行 RDO. 其次,式(4)是基于高码率(对应于采用低 QP 编码)假设推导的,然而实际编码中往往不满足高码率假设. 为此,一些改进的拉格朗日乘子优化算法相继被提出. 文献[8,9]提出基于拉普拉斯分布的 R-Q 模型和 D-Q 模型,进而推导出与每一帧编码图像率失真特性相关的自适应拉格朗日乘子计算方法,编码每一帧采用的 λ 动态地由量化参数 QP 和视频内容确定,而 H. 264 测试软件 JM 中的 QP- λ 关系可以被认为是文献[9]方法简化后的一个特例. 通过分析编码一个宏块的比特数而不是编码量化系数的比特数,文献[10]提出一种内容自适应的拉格朗日乘子计算方法,其对于具有快速运动内容的视频,编码性能改善较大. 文献[11]提出一种改进的拉格朗日乘子运动估计算法,对于帧间 16×16 划分块通过测试最小化失真(对应于 λ 为 0)和最小化码率(对应于 λ 为 ∞)两种极值,避免了较大的运动估计误差. 文献[12]进行大量实验统计后认为现有编码器 JM 和 HM 中的 QP- λ 关系不是最优的,提出通过对已编码帧的失真统计调整待编码 B 帧的 λ 值. 文献[13]采用分类最大化提出一种自适应的拉格朗日乘子选择方法,并应用到编码色度分量的 QP 补偿中.

事实上,这里讨论的 QP- λ 是统计意义上的关系. 从物理含义来说,拉格朗日乘子是编码信号率失真曲线斜率的绝对值,在率失真优化过程中用于调整优化目标,最小化率失真代价函数的编码参数选择过程中,所有的编码参数均由 RDO 决定. 为了得到最优的编码模式,每个基本编码单元的量化参数 QP 同样应该通过尝试不同的取值由具有最小率失真代价的模式确定. 但是多 QP 优化会极大的增加编码复杂度,很难应用于实际的视频编码系统. 文献[14]以 HEVC 的参考软件 HM-6.0 为平台,进行了类似于文献[7]中的实验,使用固定的 λ 值进行多 QP 优化为每个编码树单元(Coding Tree Unit, CTU)确定最优的量化参数,通过统计后得到 QP 与 $\ln(\lambda)$ 呈现非常好的线性关系:

$$QP = 4.2005 \times \ln(\lambda) + 13.7122 \quad (5)$$

式(5)被称为 QP 修正技术,其能为 HEVC 带来 1.4% ~ 1.9% 的码率节省,并且该 QP- λ 关系已被广泛应用到基于 R- λ 模型的码率控制算法^[15-17]. 另外,文献[18]在 AVS 标准下也得到了式(5)的 QP- λ 关系,但由于

AVS 与 HEVC 在量化参数设置上的不同,式中对应的参数值稍有变化.

3.2 考虑显著性的率失真优化

考虑到人类视觉系统(Human Visual System, HVS)对视频场景的感知具有选择性. 直观地,减小场景中人们感兴趣区域(Region-Of-Interest, ROI)的编码失真将能改善编码视频的主观质量^[19],特别是对于移动设备网络带宽较低的视频应用环境^[20]. 为此,基于 ROI 的视频编码优化问题可表述为

$$\min \sum_{i=1}^N W_i D_i, \quad \text{s. t.} \quad \sum_{i=1}^N R_i < R_c \quad (6)$$

式中, W_i 是编码单元 U_i 的权重因子, U_i 的显著性越重要,则 W_i 的值越大. 因此,考虑显著性的编码优化需要事先对输入视频进行感兴趣区域的检测,并确定相应的权重因子,图 2 所示为一个显著性视频编码的框图.

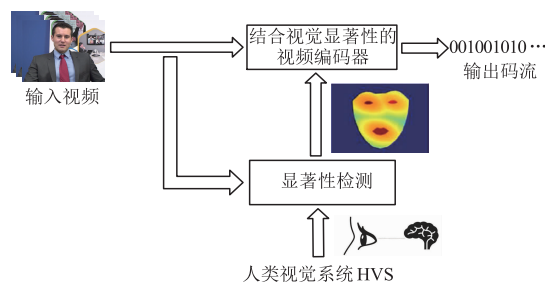


图2 显著性视频编码框图

基于人类视觉系统感知模型的 ROI 检测是显著性视频编码优化的关键问题. ROI 检测主要分为像素域^[21]的方法和压缩域^[22]的方法,前者相比后者更为精确,但像素域 ROI 检测计算复杂度较高. 另一方面,压缩域 ROI 检测可以利用码流中的变换系数、编码模式和运动矢量等信息,因此压缩域的方法计算复杂度相对较低. 显著性视频编码优化可以应用于任何视频编码标准,其不需要改变码流的语法结构. 由于量化参数 QP 直接决定了编码中的量化失真,通常采用调节 QP 的方式实现显著性视频编码优化. 直观地,会话视频中面部区域会被提取为 ROI 区域^[23],利用 ROI 检测生成的编码块权重图,文献[24]在 H. 263 编码器中通过减小面部区域的量化参数以改善低码率下会话视频编码的主观质量. 类似地,文献[25 ~ 27]分别在 H. 264、H. 265 和 VP9 视频编码器中提出基于 ROI 的编码优化. 由于 ROI 中的编码块相比其它区域使用了较小的 QP 编码,常常会在边缘处产生编码伪影从而转移了显著性分布. 为此,文献[28]根据 DCT 域的显著性检测为不同编码宏块分配量化参数,然后引入显著性失真执行率失真优化过程. 另外,针对 H. 265 编码器,文献[29 ~ 32]通过调节 CTU 的拉格朗日乘子进行显著性视频编码优化.

4 依赖 RDO

基于块划分的混合视频编码框架中各种预测编码工具的应用,使基本编码单元内部各模块间和编码单元间广泛存在率失真依赖关系。例如,帧内和帧间像素预测使得编码块间存在空域或时域的失真传播;运动矢量(Motion Vector, MV)预测使空域或时域相邻预测单元(Prediction Unit, PU)运动信息互相影响;熵编码中当前 CU 编码参数将影响相邻 CU 的上下文建模。因此,当前编码单元的编码决策可能影响后续编码单元可达到的最优编码性能。从全局优化的视角看,考虑依赖性的率失真优化能进一步提升编码性能。

同样地,我们通过前面提到的小明分配学习时间来原因来说明考虑依赖性的资源优化:第二学期,小明所能花在模拟电子技术和数字电子技术两门课程的学习时间也是 30 天。但是,由于课程内容之间的相关性,第一学期电路基础的成绩会影响小明第二学期学习后两门课程效率。如图 3 所示,虚曲线和实曲线分别表示当电路基础成绩为 70 分和 80 分后两门课程的考试失分与投入学习时间的关系,显然电路基础的成绩高有利于后两门课程的学习。一种时间分配策略为 (H_2, C_2, A_2, D_2) , 此时第一学期成绩达到最优,整学年的总成绩为 $90 + 70 + 60 + 60 = 280$ 。另一种时间分配策略为 (H_1, C_1, A_1, D_1) , 此时第一学期成绩没有达到最优,但整学年的总成绩为 $70 + 80 + 80 + 80 = 310$ 。可见,考虑到电路基础成绩对第二学期两门课程学习效率的影响,若第一学期投入多一点的时间学习电路基础,最终能使 4 门课程的总成绩更高。第二种分配策略中, C_1 点的斜率绝对值(或拉格朗日乘子)小于 H_1 点的斜率绝对值。也就是说,优化的单元间存在依赖性时,为达到全局最优,对后续影响较大的单元应该使用较小的拉格朗日乘子优化,对应地也分配到相对更多的资源。

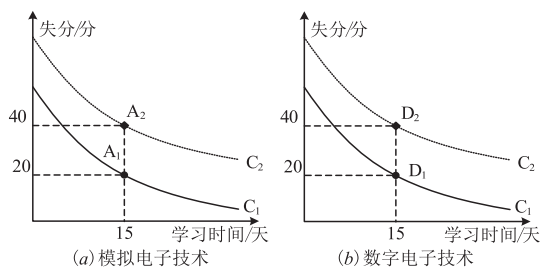


图3 拉格朗日乘子法举例(第二学期)

通过如上优化时间分配举例可知,为进行全局编码优化需要考虑编码单元之间的依赖性。下面分析视频编码中存在的依赖关系并对相关的依赖 RDO 研究进行概述和归纳。

4.1 空域依赖性

帧内预测编码中,当前编码块的像素用其上边界和左边界重建像素预测。H. 264/AVC 的帧内预测有 9 种模式, H. 265/HEVC 中有包括 Planar、DC 和其它 33 个角度共 35 种模式^[33]。其中,DC 模式使用参考像素的平均值预测当前像素块;Planar 模式使用当前块四角的四个像素进行双线性插值得到预测像素;角度模式预测时,首先对参考像素采用 3 抽头的平滑滤波,然后根据不同角度模式得到预测像素值。因此,帧内预测编码存在来自编码块上边界和左边界的失真传播。另外,为了减少语法信息的编码比特数,HEVC 在编码帧内预测模式时,根据当前 PU 周围块的帧内预测方向信息生成最可能模式列表(Most Probable Mode, MPM),使用 MPM 只需要 1~2 个比特表示当前 PU 选择的具体帧内预测方式。在帧内预测编码中,为了减少表示运动信息的比特开销,HEVC 采用了合并(Merge)、跳过(Skip)和高级运动矢量预测(Advanced Motion Vector Predictor, AMVP)模式从相邻 PU 中获得运动信息^[34],这些都使得相邻编码块间存在语法元素相关性。最后,基于上下文的熵编码过程中,二值化语法元素需要考虑其概率分布,当前块的编码参数也可能会影响到相邻块熵编码的上下文建模,即相邻块间存在熵编码依赖关系。综上所述,空域相邻编码块之间可能存在空域失真传播、编码语法元素相关、熵编码相关等的依赖关系。从而,空域内上方和左方编码块的编码决策会影响当前块可达到的编码性能。

由于帧内编码采用多种方式产生预测像素,很难准确表示空域失真传播关系,而语法元素相关和熵编码依赖性也无法进行定量描述。为了实现帧内编码全局优化,一种直接的方式就是对一帧内的所有编码单元进行联合率失真优化(Joint RDO, JRDO),即遍历所有编码单元的全部可选编码参数,采用全局率失真代价最小的参数组合进行编码。然而,随着编码单元数的增加,JRDO 的计算量呈指数提升。文献[35]分析了帧内预测的空域依赖性,然后提出对垂直方向上的两个编码块进行 JRDO,其以数倍增加编码运算复杂度为代价实现了大约 0.7% 的率失真性能提升。随后,文献[36]对文献[35]的 JRDO 方法进行了降低复杂度优化,根据图像内容自适应地采用 JRDO 和独立 RDO 选择编码参数,使运算复杂度降低到原始 HEVC 编码器的 1.4 倍,同时编码性能改善也只剩下大约 0.4%。考虑到只有编码块下边界和右边界的像素被作为参考像素,直观地,提高这些像素的编码质量能减少空域失真传播带来的影响。针对 H. 264 帧内编码,文献[37]通过调整 4×4 像素块的率失真代价运算式以改善边界像素编码质量,最终实现了大约 1.7% 的编码性能提升。类似

地,文献[38]在 HEVC 中提出松弛失真约束和迭代边界失真最小化方法以改善 HEVC 帧内编码的率失真性能.另外,文献[39,40]建立了空域失真传播模型描述边界像素失真与编码块失真关系,提出基于空域失真传播的 RDO 方法.事实上,由于优化帧内编码带来的压缩性能提升有限,而且相应的运算复杂度增加较大,针对空域依赖性的 RDO 研究相对较少.下一小节,我们将讨论视频编码中的时域依赖性及其相关的编码优化方法.

4.2 时域依赖性

帧间编码的运动补偿预测 (Motion Compensated Prediction, MCP) 和运动矢量预测 (Motion Vector Prediction, MVP) 使编码单元的编码决策在时域上产生依赖性. RDO 过程中,这种时域依赖关系对编码性能的影响比空域依赖性产生的影响大得多.为了有效地提高编码性能,已有一些研究关注于利用编码单元间的时域依赖性改进编码优化方法.

HEVC 采用了分级的编码结构 (Hierarchical Coding Structure, HCS), GOP 中的编码帧属于不同的层级 (Layer, L), 以 GOP 大小为 8 的随机接入 (Random Access, RA) 配置为例,图 4 显示了其编码参考结构.直观地,由于高层级的编码帧参考低层级的重建帧,为了使 MCP 获得精确的预测像素,应该提高低层级编码帧的编码质量.事实上,HEVC 参考编码器 HM^[41] 已经利用了这种时域依赖性,其不同层级编码帧的 QP 被设置为 $QP_L = QP_0 + \Delta QP_L$, QP_0 为 I 帧的量化参数,编码帧的层级 L 越小对应的 ΔQP_L 越小.此外,HEVC 采用的 QP 与 λ 运算关系为

$$\lambda = W_L \cdot 2^{(QP-12)/3} \quad (7)$$

式中, W_L 是一个与编码结构和编码帧所属层级相关的权重因子,其同样考虑了 HCS 中的时域依赖性. RA 的默认配置中, W_L 的设置

$$W_L = \begin{cases} 0.3705, & L=0 \\ 0.442, & L=1 \\ 0.3536 \times \text{Clip}(2, 4, (QP-12)/6), & L=2 \text{ or } 3 \\ 0.68 \times \text{Clip}(2, 4, (QP-12)/6), & L=4 \end{cases} \quad (8)$$

式中, $\text{Clip}(\cdot)$ 输出第 3 个参数,并且其输出值被限制在前两个参数之间.式(8)表明层级越低,权重因子取值越小,相应地式(7)计算的拉格朗日乘子越小,其偏向于使编码失真越小.

为了最大限度地提升编码器的率失真性能,一些研究关注于分析编码单元的时域依赖关系,并提出了自适应视频内容的优化方法.这些方法大体上可以分为自适应量化参数级联 (QP Cascading, QPC) 法和自适应拉格朗日乘子选择法,下面对这两类方法进行概述.

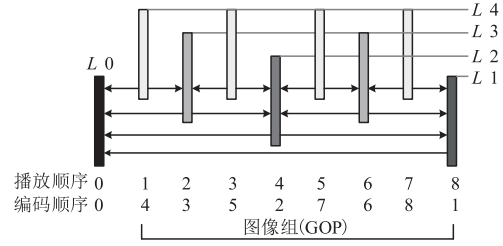


图4 HEVC的随机接入分级编码参考结构

4.2.1 自适应量化参数级联

分级 QP 级联方式较早出现在 H.264 的可伸缩视频编码 (Scalable Video Coding, SVC)^[42] 中,其实现时域可伸缩编码的同时获得了率失真性能提升.随后,文献[43]分析了 SVC 分级 QPC 获得率失真性能改善的原因,并提出根据量化失真传播动态地调整 GOP 中各帧的 QP 补偿值,对内容变化缓慢的视频采用较大的 QP 间隔,对内容变化较快的视频采用较小的 QP 间隔.类似地,文献[44,45]针对 SVC 提出自适应的 QPC 方法,但是它们对时域依赖的度量方式各不相同.对于 HEVC 的分级编码结构,文献[46,47]提出了低延时 (Low Delay, LD) 编码配置下的自适应 QPC 方法,文献[48~54]提出了 RA 配置下的自适应 QPC 方法.以文献[54]为例简要介绍编码帧之间的时域依赖性,文献[54]通过调节参考帧的编码 QP 进行实验,并统计得到图 5 所示的参考帧失真对后续帧编码的影响,图 5(a) 表明预测帧的编码失真随着参考帧失真的增大而增大,图 5(b) 表明预测帧的编码码率基本不随参考帧失真变化而变化.据此,文献[54]得到 RA 参考结构下各层级视频帧的编码失真对整个 GOP 编码失真的影响如式(9)所示:

$$\xi_{Li} = \begin{cases} 1 + \sum_{k=i+1}^4 \mu_{Lk, Li} \cdot \xi_{Lk}, & i = 0, 1, 2, 3 \\ 1, & i = 4 \end{cases} \quad (9)$$

式中, $\mu_{Lk, Li}$ 表示 L_i 层级的参考帧失真对 L_k 层级视频帧的编码失真影响,它是图 5(a) 中直线的斜率.通过 ξ_{Li} 自适应地设置 GOP 中每一帧的 QP 补偿值,文献[54]相比于 HEVC 固定分级 QPC 方法获得了平均 3.5% 的率失真性能提升.

以上 QPC 方法根据帧间依赖性自适应调整 GOP 中每一帧的编码量化参数,其帧内所有 CTU 均采用相同的 QP 编码.事实上,一帧中的不同 CTU 编码失真对后续帧编码的影响不尽相同,因此更精细的自适应量化参数选择可以执行在 CTU 层.文献[55~59]提出 CTU 层的自适应 QPC 优化方法,在文献[58,59]中,采用简化的预编码过程确定 CTU 编码失真的时域传播,进而提出自适应 CTU 层量化参数选择方法,获得了较大的率失真性能提升.

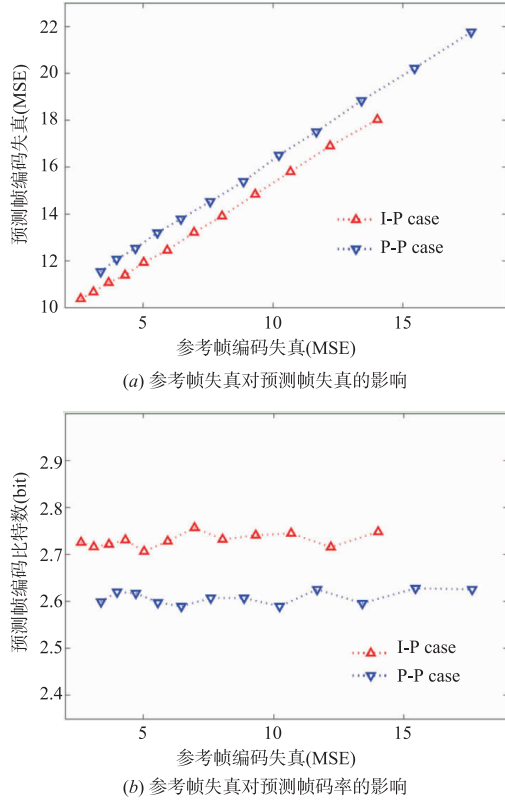


图5 参考帧失真对后续帧编码的影响 (BasketballDrill)

4.2.2 自适应拉格朗日乘子选择

由于编码单元之间的时域依赖关系,对当前编码单元的编码决策会影响后续参考它的编码单元可达到的率失真性能.如前文提到的小明优化分配学习时间的例子所示,为达到全局性能最优,对后续编码影响较大的编码单元应该使用较小的拉格朗日乘子进行编码,那么式(3)的最小化率失真代价问题可更改为

$$\min \left\{ J_i = D_i + \frac{\lambda_g}{1 + \psi_i} R_i \right\} \quad (10)$$

式中, λ_g 称为全局拉格朗日乘子, ψ_i 是第 i 个编码单元对后续编码过程的影响.若第 i 个编码单元中的像素从不被作为参考像素,则其对应的 ψ_i 为零;若第 i 个编码单元中的像素被后续编码单元参考,则 ψ_i 大于零.关于 ψ_i 的定量测量,文献[60]提出了信源失真时域扩散模型, ψ_i 被称为时域传播因子,可定义为

$$\psi_i = \frac{\partial \sum_{k=i+1}^M D_k}{\partial D_i} \quad (11)$$

式中, D_i 是编码单元 U_i 的失真, D_k 是时域上直接或间接参考 U_i 的编码单元失真,这里的 ψ_i 表示编码单元 U_i 的失真变化对后续参考它的编码单元失真的影响.对于 H.264 的 IPPP 参考结构,首先在原始视频中通过前向运动搜索建立基于编码块的时域传播链以确定编码中

被影响的所有后续编码块;然后通过分析信源失真与编码失真的关系估算出时域传播因子.文献[60]的自适应拉格朗日乘子选择方法在 H.264 编码器中获得了平均 13.5% 的率失真性能提升,随后文献[61]和文献[62]把该方法扩展到 HEVC 的 LD-HCS 和 RA-HCS 中同样获得了率失真性能改善.

文献[63]在高码率假设下认为编码失真只与量化步长有关,因而不考虑参考像素失真的时域扩散,其提出了式(12)的最小化率失真代价表达式:

$$\min \{ J_i = D_i + \lambda \cdot R_i + \lambda \cdot \Delta R_{i+1}(D_i) \} \quad (12)$$

式中, $\Delta R_{i+1}(D_i)$ 表示编码单元 U_i 的失真引起后续参考它的编码块的码率增量.根据信源失真与编码失真的关系和高码率下的率失真函数,式(12)最终变换成类似式(10)的形式,如式(13)所示:

$$\min \left\{ J_i = D_i + \frac{\lambda}{1 + 1/(2 \ln(2) \cdot D_{\text{OMCP}})} R_i \right\} \quad (13)$$

式中, D_{OMCP} 是在原始视频中进行前向运动搜索得到的运动补偿预测误差,其值越小表明编码单元 U_i 与后续帧的相关性越强,相应地 U_i 将使用较小的拉格朗日乘子编码,从而为后续编码单元提供较高质量的参考像素.文献[60~62]考虑当前编码单元失真对后续编码单元失真的增量,而文献[63]考虑当前编码单元失真对后续编码单元码率的增量,但最终它们通过拉格朗日乘子都转换成了当前编码单元失真对后续编码单元率失真代价的增量.因此,本质上两种方法都是利用编码单元之间的时域依赖关系自适应地调节拉格朗日乘子从而实现编码性能改善.

针对屏幕内容编码 (Screen Content Coding, SCC)^[64],文献[65]提出加权率失真优化,实质上也是通过自适应调节 CTU 的拉格朗日乘子进行编码.屏幕内容常常包含一些重复的纹理和图案,在 HEVC-SCC 中引入帧内块复制 (Intra Block Copy, IBC) 编码工具实现了类似于 MCP 的帧内像素预测,因此 SCC 中一个 CTU 的编码决策不仅在时域上影响后续帧的编码,也会影响当前帧中参考它的后续 CTU 的编码.在文献[65]中,通过统计一个 CTU 在空域和时域上被参考的次数来估计权重因子,若 CTU 被参考次数越多则对应的权重因子越大,编码时将采用较小的拉格朗日乘子计算率失真代价.类似地,文献[66]通过预分析检测视频中的静态区域和快速变化区域提出自适应拉格朗日乘子优化方法.其它的一些研究要么采用实验统计^[67,68]要么根据理论推导^[69,70]建立时域率失真依赖模型,然后提出了不同的自适应拉格朗日乘子方法.

5 码率控制中的比特分配

视频通信中,若编码器输出码率大于信道带宽将

引起丢帧或客户端播放卡顿,若编码器输出码率小于信道带宽将造成带宽浪费,因为更多的码率可以用于提升编码视频质量.因此,码率控制是视频编码系统不可缺少的重要模块,它控制编码器输出预定的码率,同时尽可能地减小编码视频失真^[71].在设计码率控制算法时,率失真性能是需要考虑的基础问题之一.为此,码率控制的基本问题可以表示为式(14)所示的率失真优化问题^[72],描述为:编码视频消耗的比特数 R 不超过目标比特数 R_i 的情况下,最小化编码视频失真 D .

$$\min D, \quad \text{s. t.} \quad R \leq R_i \quad (14)$$

码率控制算法中,根据对限制码率波动的要求不同可分类为恒定比特率(Constant Bit Rate, CBR)方式、动态比特率(Variable Bit Rate, VBR)方式和平均比特率(Average Bit Rate, ABR)方式.其中,由于每一帧编码复杂度的不同,采用 CBR 方式会导致编码帧质量的激励波动;VBR 方式的码率可以随着图像复杂程度不同而变化,其编码图像质量稳定,且编码效率较高,但较大的码率波动可能会引起缓存器上溢或下溢;ABR 是 VBR 的一种改进方式,码率可随着图像复杂度不同而变化,但在一小段时间范围内其平均码率保持恒定.因此,ABR 被做为 VBR 和 CBR 的一种折衷选择,下面讨论的码率控制方法基本都可以认为是 ABR 方式.一般而言,码率控制过程涉及比特资源分配和比特率控制.比特资源分配是根据目标码率、帧率等信息分级地为一个 GOP、帧和 CTU 分配适当的目标比特数,比特率控制是通过调节编码参数(通常直接调节 QP 和 λ)使编码一个 CTU 或一帧的消耗比特数等于其预算比特数.为了使编码器输出较为平滑的码率,通常平均地为每一个 GOP 分配目标比特数;而为了改善编码器失真性能,帧级的比特分配除了考虑当前 GOP 剩余预算比特数外,还应该考虑编码结构和视频内容特性等信息;若使用 CTU 级的比特率控制则需要进一步地根据当前帧的剩余预算比特数为每一个待编码的 CTU 分配目标比特.

每一代视频编码标准发布后,都会有许多针对该标准的码率控制方法被提出.本节通过讨论几个针对当前视频编码标准 HEVC 的码率控制方法,以说明编码单元间的率失真依赖性在比特资源分配中的影响.文献[73]提出了 UR-Q(Unified Rate-Quantization)模型码率控制方法,其帧级比特预算仅根据当前 GOP 剩余比特数、未编码帧数和缓存器状态等进行分配,尽管实现了码率控制,但由于不合理的比特资源分配和不够精确的码率模型导致严重的率失真性能损失,其相比于不开启码率控制的 HEVC 基准编码器的压缩性能降低了 11%.文献[74]分析了编码量化系数和编码语法元素的码率与 QP 的关系,提出一个更精确的 R-Q 模型用于帧级码率控制,虽然获得比文献[73]更高的码率精

度,其仍然严重的降低编码器压缩性能.视频编码中,拉格朗日乘子 λ 对应于率失真曲线上某一点的切线,理论上每一个 λ 取值唯一地对应着一组码率和失真.基于此,文献[15]提出了 R- λ 模型码率控制方法,包括等比率比特分配和分级比特分配两种帧层比特分配方案,如式(15)所示:

$$T_{\text{Pic}} = \frac{T_{\text{GOP}} - \text{Coded}_{\text{GOP}}}{\sum_{\text{AllNotCodedPics}} \omega_{\text{Pic}}} \times \omega_{\text{PicCurr}} \quad (15)$$

式中, T_{GOP} 和 T_{Pic} 分别是一个 GOP 和一帧的目标比特数; $\text{Coded}_{\text{GOP}}$ 是当前 GOP 中已消耗的比特数; ω 表示每一帧的比特分配权重.等比率比特分配方案中,每一帧的比特分配权重都是 1;分级比特分配方案中,根据编码参考结构为 GOP 中不同层级的帧设置固定的 ω ,其类似于式(8)中的 λ 权重因子设置,在一定程度上考虑了编码帧之间的率失真依赖关系.文献[15]的分级比特分配方案减小了码率控制引起的率失真性能损失,与基准编码器相比,其压缩性能降低约为 6%.为了获得更为准确的码率模型以实现精确的比特率控制和率失真性能优化,文献[75~77]中的码率控制方法采用了二次编码.例如文献[77]仅采用 16×16 的 CU 划分进行一次预编码,根据实际编码信息估计每个 CTU 的 R- λ 模型,一定程度地提高了比特率控制精度和率失真性能.

正如自适应量化参数级联和自适应拉格朗日乘子选择能改善编码性能,一些改进的码率控制方法^[78~83]根据时域依赖关系自适应地分配比特资源改善了率失真性能.文献[79]在 λ 域的码率模型和失真模型分析框架下提出自适应比特分配方法,提高了 R- λ 模型码率控制算法的码率精度和率失真性能.文献[81,82]在 ρ 域建立失真时域依赖模型,提出了自适应的帧层比特分配方法,其中文献[81]使用 R-Q 模型进行码率控制,文献[82]用 R- λ 模型进行码率控制,它们均获得了较好的率失真性能.基于全局率失真性能最优的假设,文献[83]利用帧级失真时域传播和编码帧的率失真特性建立了 GOP 比特预算约束下的最优帧级比特分配公式,并采用迭代泰勒级数展开法^[84]求解实现帧层的码率控制,其率失真性能相比不开启码率控制的 HEVC 基准编码器提升了 3.5%.另外,文献[85]把文献[61]的时域失真传播因子结合到 CTU 层的码率控制算法设计中,其提出的 CTU 层比特分配优化目标是在当前帧比特预算约束下最小化编码视频的失真,而不是最小化当前帧的失真,最终实现了 5.1% 的率失真性能提升.

6 讨论

率失真优化技术极大地改善了混合视频编码器的压缩性能,但是当前技术远未达到全局率失真优化,编

码器仍然有进一步提升压缩性能的空间. 另外, 下一代视频编码标准在联合视频专家组 (Joint Video Experts Team, JVET) 第 10 次会议后被正式命名为多功能视频编码 (Versatile Video Coding, VVC)^[86], 当前仍在制定过程中. 由于率失真优化和码率控制均不属于编码标准规定的内容, 关于 VVC 的前期研究主要集中在编码工具的改进方面^[87,88], 预计 2020 年 VVC 标准发布正式版

本后也将会有许多针对该标准的率失真优化和码率控制技术被提出. 表 1 列出了前文提及的部分编码优化方法, 以此呈现不同技术带来的率失真性能变化, 表中最后一列的 R-D 性能指文献中算法相比于基准编码器的 BD-rate 平均值, 此外所列算法运算复杂度均不大于基准编码器的 2 倍.

表 1 不同编码优化方法的特征和性能

序号	第一作者, 年	预处理	调节 QP	调节 λ	考虑时域依赖	考虑空域依赖	码率控制	R-D 性能 (BD-rate)
1	Bin Li ^[14] , 2013	-	✓	-	-	-	-	-1.7%
2	Bin Li ^[15] , 2014	-	✓	✓	✓	-	✓	6.5%
3	Maxime Bichon ^[36] , 2018	-	-	-	-	✓	-	-0.5%
4	Jing He ^[54] , 2018	✓	✓	✓	✓	-	-	-3.5%
5	Michael Ropert ^[58] , 2017	✓	✓	-	✓	-	-	-10%
6	Tianwu Yang ^[60] , 2012	✓	✓	✓	✓	-	-	-13.5%
7	Yanbo Gao ^[61] , 2017	✓	✓	✓	✓	-	-	-3.4%
8	Yanbo Gao ^[62] , 2019	✓	✓	✓	✓	-	-	-4.5%
9	Shuai Li ^[63] , 2016	✓	✓	✓	✓	-	-	-4.0%
10	Wei Xiao ^[65] , 2018	✓	✓	✓	✓	✓	-	-10%
11	Hyomin Choi ^[73] , 2013	-	✓	✓	-	-	✓	11%
12	Li Li ^[79] , 2018	-	✓	✓	✓	-	✓	0.7%
13	Shanshe Wang ^[81] , 2013	-	✓	✓	✓	-	✓	-1.6%
14	Hongwei Guo ^[83] , 2019	-	✓	✓	✓	-	✓	-3.5%
15	Hongwei Guo ^[85] , 2019	✓	✓	✓	✓	-	✓	-5.1%

在不显著增加编码器运算复杂度的情况下, 通过优化编码决策过程实现率失真性能改善具有重要的现实意义, 它可以节省带宽和设备存储空间, 从而降低运营成本. 近年来针对当前视频编码器的率失真优化技术已经取得了一定进展, 但仍面临很多挑战:

(1) 理论上全局率失真优化需要对整个视频序列的编码参数进行联合优化, 但是当前的计算能力不可能进行如此高复杂度的编码过程. 深入分析编码决策过程中的前后依赖关系, 结合编码单元之间的空域和时域依赖性建立可操作的全局率失真优化模型是极具挑战的任务.

(2) 文献^[58,60,65]的优化方法实现了高达 10% 的率失真性能提升, 它们对应的基准编码器分别是 H.264 参考软件 JM、HEVC 工业编码器 x265 和 HEVC-SCC 参考软件 SCM, 这些编码器本身具有较大的压缩性能优化提升空间. 在较高运算复杂度下探索当前视频编码系统压缩性能极限是比较有意义的研究课题.

(3) 编码单元的依赖关系中, 时域依赖性强于空域依赖性, 参考像素失真对预测块的失真影响大于对码率的影响. 从理论上建立定量度量空域和时域依赖性的方法, 并针对具体编码结构设计可执行的测量操作还需要进一步研究.

(4) 量化参数 QP 和拉格朗日乘子 λ 是决定编码码率和失真的主要参数, 当前 QP 与 λ 在统计意义上的数学关系式并不总是最优的. 进一步研究根据视频内容特性和率失真依赖性自适应地调整 λ 与 QP 的数学关系式能更好地优化视频压缩性能.

(5) 下一代视频编码标准 VVC、AVS3 以及 AV1 正在制定过程中, 更多编码工具和新技术的引入将使下一代视频编码器具有更高的压缩性能和新的特性. 如何把当前的率失真优化技术较好地应用到新一代视频编码器中是即将开展的研究工作.

7 结束语

率失真优化是视频编码领域重要的研究内容, 贯穿于编码单元的模式选择、编码资源分配、码率控制等编码过程. 视频编码标准仅规定了编码器使用的编码工具和输出码流的句法结构, 编码器可以采用不同的率失真优化方案以期实现更高的压缩性能. 本文综述了视频编码率失真优化技术的研究进展, 通过形象的举例阐述了基于拉格朗日乘子的独立率失真优化和依赖率失真优化, 分析了视频编码中 λ 与 QP 的关系, 简要介绍了编码单元之间的率失真依赖性, 并分类归纳各种率失真优化技术. 最后, 展示了近年出现的部分

率失真优化方法的特征和性能,并提出率失真优化技术所面临的挑战和值得进一步探索的研究内容。

参考文献

- [1] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression [J]. IEEE Signal Processing Magazine, 1998, 15(6): 74–90.
- [2] SHANNON C E. Coding theorems for a discrete source with a fidelity criterion [J]. Institute of Radio Engineers, International Convention Record, 1959, 7(7): 142–163.
- [3] 杨天武, 彭强, 邓云. 一种基于率失真优化的关键参考帧选择算法 [J]. 电子学报, 2006, 34(7): 1241–1245.
YANG T-W, PENG Q, DENG Y. A novel key frame reference picture selection algorithm based on rate-distortion optimization [J]. Acta Electronica Sinica, 2006, 34(7): 1241–1245. (in Chinese)
- [4] LI H Q, LI B, XU J Z. Rate-distortion optimized reference picture management for high efficiency video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1844–1857.
- [5] 陆寄远, 朝红阳, 黄承慧. 计算能力可伸缩的运动估计率失真优化 [J]. 电子学报, 2014, 42(8): 1495–1502.
LU J-Y, CHAO H-Y, HUANG C-H. Rate distortion optimization of complexity scalable motion estimation [J]. Acta Electronica Sinica, 2014, 42(8): 1495–1502. (in Chinese)
- [6] ORTEGA A, RAMCHANDRAN K. Rate-distortion methods for image and video compression [J]. IEEE Signal Processing Magazine, 1998, 15(6): 23–50.
- [7] WIEGAND T, GIROD B. Lagrange multiplier selection in hybrid video coder control [A]. Proceedings of International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2001. 542–545.
- [8] LI X, OERTEL N, KAUP A. Adaptive Lagrange multiplier selection for intra-frame video coding [A]. Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS) [C]. Piscataway: IEEE, 2007. 3643–3646.
- [9] LI X, OERTEL N, HUTTER A, et al. Laplace distribution based lagrangian rate distortion optimization for hybrid video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(2): 193–205.
- [10] ZHANG J, YI X Q, LING N, et al. Context adaptive lagrange multiplier (CALM) for rate-distortion optimal motion estimation in video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(6): 820–828.
- [11] GONZALEZ-DE-SUSO J L, JIMENEZ-MORENO A, et al. Improved method to select the Lagrange multiplier for rate-distortion based motion estimation in video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(3): 452–464.
- [12] ZHANG F, BULL D R. Rate-distortion optimization using adaptive Lagrange multipliers [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29(10): 3121–3131.
- [13] AN C H, NGUYEN T Q. Adaptive Lagrange multiplier selection using classification-maximization and its application to chroma qp offset decision [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(6): 783–791.
- [14] LI B, XU J Z, ZHANG D, et al. QP refinement according to lagrange multiplier for high efficiency video coding [A]. Proceedings of International Symposium on Circuits and Systems [C]. Piscataway: IEEE, 2013. 477–480.
- [15] LI B, LI H Q, LI L, et al. λ -domain rate control algorithm for high efficiency video coding [J]. IEEE Transactions on Image Processing, 2014, 23(9): 3841–3854.
- [16] 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.
- [17] LI L, LI B, LIU D, et al. λ -domain rate control algorithm for HEVC scalable extension [J]. IEEE Transactions on Multimedia, 2016, 18(10): 2023–2039.
- [18] LI B, XU J, LI H. Refining QP to improve coding efficiency in AVS [A]. Proceedings of IEEE International Conference on Image Processing [C]. Piscataway: IEEE, 2013. 1528–1531.
- [19] 张方, 吴成柯, 肖嵩. 一种基于 MPEG-4 的兴趣区域视频编码新方法 [J]. 电子学报, 2005, 33(4): 634–637.
ZHANG F, WU C K, XIAO S. A new regions of interest video coding method based on MPEG-4 [J]. Acta Electronica Sinica, 2005, 33(4): 634–637. (in Chinese)
- [20] LI Y, TAO X M, LU J. Hybrid model-and-object-based real-time conversational video coding [J]. Signal Processing-Image Communication, 2015, 35(2015): 9–19.
- [21] HU P, SHUAI B, LIU J, et al. Deep level sets for salient object detection [A]. Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition [C]. Piscataway: IEEE, 2017. 540–549.
- [22] XU M, JIANG L, SUN X Y, et al. Learning to detect video saliency with HEVC features [J]. IEEE Transactions on Image Processing, 2017, 26(1): 369–385.
- [23] XIONG B, FAN X J, ZHU C, et al. Face region based conversational video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(7): 917–931.
- [24] YANG X K, LIN W S, LU Z K, et al. Rate control for videophone using local perceptual cues [J]. IEEE Trans-

- actions on Circuits and Systems for Video Technology, 2005, 15(4):496–507.
- [25] LIU Y, LI Z G, SOH Y C. Region-of-interest based resource allocation for conversational video communication of H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(1):134–139.
- [26] XU M, DENG X, LI S X, et al. Region-of-Interest based conversational HEVC coding with hierarchical perception model of face [J]. IEEE Journal of Selected Topics in Signal Processing, 2014, 8(3):475–489.
- [27] 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(0):13677–13688.
- [28] HADIZADEH H, BAJIC I V. Saliency-aware video compression [J]. IEEE Transactions on Image Processing, 2014, 23(1):19–33.
- [29] LI S, XU M, DENG X, et al. Weight-based $R-\lambda$ rate control for perceptual HEVC coding on conversational videos [J]. Signal Processing-Image Communication, 2015, 38(2015):127–140.
- [30] 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.
- [31] LI S X, XU M, et al. Closed-form optimization on saliency-guided image compression for HEVC-MSP [J]. IEEE Transactions on Multimedia, 2018, 20(1):155–170.
- [32] ZENG H, NGAN K N, WANG M. Perceptual adaptive Lagrangian multiplier for high efficiency video coding [A]. Proceedings of Picture Coding Symposium [C]. Piscataway: IEEE, 2013:69–72.
- [33] LAINEMA J, BOSSEN F, et al. Intra coding of the HEVC standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12):1792–1801.
- [34] LIN J L, CHEN Y W, HUANG Y W, et al. Motion vector coding in the HEVC standard [J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6):957–968.
- [35] BICHON M, TANOU J L, et al. Inter-block dependencies consideration for intra coding in H. 264/AVC and HEVC standards [A]. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Piscataway: IEEE, 2017. 1537–1541.
- [36] BICHON M, LE TANOU J, ROPERT M, et al. Low complexity joint RDO of prediction units couples for HEVC intra coding [A]. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing [C]. Piscataway: IEEE, 2018. 1733–1737.
- [37] YOU J, CHOI C, JEONG J. Modified rate distortion optimization using inter-block dependence for H. 264/AVC intra coding [J]. IEEE Transactions on Consumer Electronics, 2008, 54(3):1383–1388.
- [38] SUN L, AU O C, DAI W, et al. Modified distortion redistribution problem for high efficiency video coding (HEVC) [A]. Proceedings of IEEE 14th International Workshop on Multimedia Signal Processing (MMSP) [C]. Piscataway: IEEE, 2012. 278–282.
- [39] PANG C, AU O C, ZOU F, et al. Optimal distortion redistribution in block-based image coding using successive convex optimization [A]. Proceedings of IEEE International Conference on Multimedia and Expo (ICME) [C]. Piscataway: IEEE, 2011. 1–5.
- [40] WU Q B, XIONG J, LUO B, et al. A novel joint rate distortion optimization scheme for intra prediction coding in H. 264/AVC [J]. IEEE Transactions on Information and Systems, 2014, E97-D(4):989–992.
- [41] KIM I-K, MCCANN K, SUGIMOTO K, et al. High Efficiency Video Coding (HEVC) Test Model 11 (HM11) Encoder Description [R]. Document JCTVC-M1002, Incheon, KR, 2013.
- [42] SCHWARZ H, MARPE D, WIEGAND T. Overview of the scalable video coding extension of the H. 264/AVC standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(9):1103–1120.
- [43] LI X A, AMON P, HUTTER A, et al. Model based analysis for quantization parameter cascading in hierarchical video coding [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2009. 3765–3768.
- [44] WAN W X, CHEN Y, WANG Y K, et al. Efficient hierarchical inter picture coding for H. 264/AVC baseline profile [A]. Proceedings of Picture Coding Symposium (PCS) [C]. Piscataway: IEEE, 2009. 1–4.
- [45] LI X, AMON P, HUTTER A, et al. Adaptive quantization parameter cascading for hierarchical video coding [A]. Proceedings of IEEE International Symposium on Circuits and Systems [C]. Piscataway: IEEE, 2010. 4197–4200.
- [46] XU Y, LI Q, LI X, et al. Efficient QP cascading in H. 265/HEVC low-delay prediction [A]. Proceedings of IEEE International Conference on Multimedia & Expo Workshops [C]. Piscataway: IEEE, 2017. 1–6.
- [47] AMER H, YANG E H. Low-delay HEVC adaptive quantization parameter selection through temporal propagation length estimation [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2018. 211–215.
- [48] PAPADOPOULOS M A, ZHANG F, et al. An adaptive QP offset determination method for HEVC [A]. Proceed-

- ings of IEEE International Conference on Image Processing [C]. Piscataway: IEEE, 2016. 4220 – 4224.
- [49] YANG Y, WAN S, GONG Y, et al. Adaptive quantization parameter cascading for random-access prediction in H. 265/HEVC based on dependent R-D models [A]. Proceedings of IEEE International Conference on Image Processing [C]. Piscataway: IEEE, 2016. 4235 – 4239.
- [50] ZHAO T S, WANG Z, CHEN C W. Adaptive quantization parameter cascading in HEVC hierarchical coding [J]. IEEE Transactions on Image Processing, 2016, 25 (7): 2997 – 3009.
- [51] 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.
- [52] YANG K, WAN S, GONG Y, et al. Content adaptive quantization parameter cascading for random-access structure in HEVC [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2017. 2498 – 2502.
- [53] ZHOU Y M, WANG H Y, TIAN L, et al. Temporal correlation based hierarchical quantization parameter determination for HEVC video coding [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2017. 2478 – 2482.
- [54] HE J, YANG E H, et al. Adaptive quantization parameter selection for H. 265/HEVC by employing inter-frame dependency [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28 (12): 3424 – 3436.
- [55] WANG M H, NGAN K N, et al. Improved block level adaptive quantization for high efficiency video coding [A]. Proceedings of IEEE International Symposium on Circuits and Systems [C]. Piscataway: IEEE, 2015. 509 – 512.
- [56] YIN H B, CAI H, FAN M T, et al. Quantization parameter cascading for video coding: leveraging a new temporal distortion propagation model [J]. Signal Image and Video Processing, 2017, 11 (5): 801 – 808.
- [57] XIANG G Q, JIA H Z, YANG M Y, et al. A novel adaptive quantization method for video coding [J]. Multimedia Tools and Applications, 2018, 77 (12): 14817 – 14840.
- [58] ROPERT M, TANOU J L, BICHON M, et al. R-D spatio-temporal adaptive quantization based on temporal distortion backpropagation in HEVC [A]. Proceedings of IEEE International Workshop on Multimedia Signal Processing (MMSP) [C]. Piscataway: IEEE, 2017. 1 – 6.
- [59] BICHON M, TANOU J L, ROPERT M, et al. Temporal adaptive quantization using accurate estimations of inter and skip probabilities [A]. Proceedings of Picture Coding Symposium [C]. Piscataway: IEEE, 2018. 81 – 85.
- [60] YANG T, ZHU C, FAN X, et al. Source distortion temporal propagation model for motion compensated video coding optimization [A]. Proceedings of International Conference on Multimedia & Expo (ICME) [C]. Piscataway: IEEE, 2012. 85 – 90.
- [61] GAO Y, ZHU C, LI S, et al. Temporal dependent rate-distortion optimization for low-delay hierarchical video coding [J]. IEEE Transactions on Image Processing, 2017, 26 (9): 4457 – 4470.
- [62] GAO Y, ZHU C, et al. Source distortion temporal propagation analysis for random-access hierarchical video coding optimization [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29 (2): 546 – 559.
- [63] LI S, ZHU C, GAO Y B, et al. Lagrangian multiplier adaptation for rate-distortion optimization with inter-frame dependency [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26 (1): 117 – 129.
- [64] XU J Z, JOSHI R, COHEN R A. Overview of the emerging HEVC screen content coding extension [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26 (1): 50 – 62.
- [65] XIAO W, LI B, XU J, et al. Weighted rate-distortion optimization for screen content coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 28 (2): 499 – 512.
- [66] GONZALEZ-DE-SUSO J L, MARTINEZ-ENRIQUEZ E, DIAZ-DE-MARIA F. Adaptive Lagrange multiplier estimation algorithm in HEVC [J]. Signal Processing-Image Communication, 2017, 56 (2017): 40 – 51.
- [67] YANG K F, WAN S, GONG Y C, et al. An efficient Lagrangian multiplier selection method based on temporal dependency for rate-distortion optimization in H. 265/HEVC [J]. Signal Processing-Image Communication, 2017, 57 (2017): 68 – 75.
- [68] ZHANG F, BULL D R. An adaptive Lagrange multiplier determination method for rate-distortion optimisation in hybrid video codecs [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2015. 671 – 675.
- [69] WANG X, SONG L, LUO Z, et al. Lagrangian method based rate-distortion optimization revisited for dependent video coding [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Piscataway: IEEE, 2017. 3021 – 3025.
- [70] DE ABREU A, CHEUNG G, FROSSARD P, et al. Optimal Lagrange multipliers for dependent rate allocation in video coding [J]. Signal Processing-Image Communication, 2018, 63 (2018): 113 – 124.

- [71] 崔子冠,朱秀昌,干宗良. H. 264 视频编码率失真优化和码率控制技术研究进展[J]. 电子学报, 2013, 41(12): 2443 – 2450.
CUI Z G, ZHU X C, GAN Z L. Advances in rate distortion optimization and rate control techniques for h. 264 video coding[J]. Acta Electronica Sinica, 2013, 41(12): 2443 – 2450. (in Chinese)
- [72] CHEN Z Z, NGAN K N. Recent advances in rate control for video coding[J]. Signal Processing-Image Communication, 2007, 22(2007): 19 – 38.
- [73] CHOI H, YOO J, NAM J, et al. Pixel-wise unified rate-quantization model for multi-level rate control[J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 7(6): 1112 – 1123.
- [74] LEE B, KIM M, NGUYEN T Q. A frame-level rate control scheme based on texture and nontexture rate models for high efficiency video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(3): 465 – 479.
- [75] WANG S, REHMAN A, ZENG K, et al. SSIM-motivated two-pass VBR coding for HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(10): 2189 – 2203.
- [76] ZUPANCIC I, NACCARI M, MRAK M, et al. Two-pass rate control for improved quality of experience in UHDTV delivery[J]. IEEE Journal of Selected Topics in Signal Processing, 2017, 11(1): 167 – 179.
- [77] WEN J T, FANG M Y, TANG M H, et al. R-lambda model based improved rate control for HEVC with pre-encoding[A]. Proceedings of Data Compression Conference (DCC)[C]. Piscataway: IEEE, 2015. 53 – 62.
- [78] LIU M H, REN P, XIANG Z. Frame-level bit allocation for hierarchical coding of H. 265/HEVC considering dependent rate-distortion characteristics[J]. Signal Image and Video Processing, 2016, 10(8): 1457 – 1463.
- [79] LI L, LI B, LI H, 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.
- [80] FIENGO A, CHIERCHIA G, CAGNAZZO M, et al. Rate allocation in predictive video coding using a convex optimization framework[J]. IEEE Transactions on Image Processing, 2017, 26(1): 479 – 489.
- [81] WANG S S, MA S W, 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.
- [82] Gao W, Kwong S, Yuan H, et al. DCT coefficient distribution modeling and quality dependency analysis based frame-level bit allocation for HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 139 – 153.
- [83] GUO H W, ZHU C, LI S X, et al. Optimal bit allocation at frame level for rate control in HEVC[J]. IEEE Transactions on Broadcasting, 2019, 65(2): 270 – 281.
- [84] LI S, XU M, WANG Z, et al. Optimal bit allocation for CTU level rate control in HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27(11): 2409 – 2424.
- [85] GUO H W, ZHU C, XU M, et al. Inter-block dependency based CTU level rate control for HEVC[J]. IEEE Transactions on Broadcasting, 2020, 66(1): 113 – 126.
- [86] WIEN M, BOYCE J M, STOCKHAMMER T, et al. Standardization status of immersive video coding[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2019, 9(1): 5 – 17.
- [87] GARRIDO M J, PESCADOR F, CHAVARRÍAS M, et al. A 2-D multiple transform processor for the versatile video coding standard[J]. IEEE Transactions on Consumer Electronics, 2019, 65(3): 274 – 283.
- [88] YANG H, SHEN L, DONG X, et al. Low complexity CTU partition structure decision and fast intra mode decision for versatile video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019, doi: 10.1109/TCSVT.2019.2904198.

作者简介



郭红伟 男, 1980 年 2 月出生, 云南金平人. 电子科技大学在读博士生、红河学院副教授、CCF 高级会员. 2002 年和 2010 年在云南大学获理学学士和工学硕士学位. 主要研究方向为视频编码与通信.

E-mail: ghw202@foxmail.com



朱 策(通信作者) 男, 1969 年 9 月出生, 四川自贡人. 教授、博士生导师、IEEE Fellow. 1989 年在四川大学获理学学士学位, 1992 年和 1994 年在东南大学获工学硕士和工学博士学位. 主要研究方向为视频编码与通信、视频分析和处理.

E-mail: eczhu@uestc.edu.cn