



研究与开发

基于统计建模的 VVC 快速码率估计算法

祁伟, 殷海兵, 王鸿奎, 黄晓峰, 牛伟宏

(杭州电子科技大学通信工程学院, 浙江 杭州 310018)

摘要: 为降低新一代通用视频编码 (versatile video coding, VVC) 标准率失真优化过程的编码复杂度, 提出一种基于统计建模的快速码率估计算法。首先, 算法充分考虑依赖性量化 (dependent quantization, DQ) 的量化行为和熵编码中的上下文依赖, 提出可以准确刻画编码过程中上下文状态迁移的码率特征, 初步预估变换单元 (transform unit, TU) 中部分语法元素的码率; 其次, 基于系数分布特性, 定义系数混乱度特征和稀疏度特征来区分系数分布差异带来的码率影响, 并构建 TU 级码率模型; 最后, 算法根据码率构成特性将大尺寸 TU 和小尺寸 TU 分开建模实现更精准的码率预估。通过统计方式对大量样本进行回归训练, 得到最终的线性码率模型, 并应用于 VVC 的模式决策中。实验结果表明, 所提出算法在随机访问 (random access, RA) 配置下, 可以实现 16.289% 的复杂度降低, 而码率变化率 (Bjontegaard delta bit rate, BD-BR) 仅增加 1.567%。

关键词: 码率预估; 通用视频编码; 率失真优化; 回归训练

中图分类号: TN919.81

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022279

Statistical modeling based fast rate estimation algorithm for VVC

QI Wei, YIN Haibing, WANG Hongkui, HUANG Xiaofeng, NIU Weihong

College of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: To reduce the coding complexity of the rate-distortion optimization process of the latest video coding standard versatile video coding (VVC), a fast rate estimation model based on statistical modeling was proposed. Firstly, the quantization behavior in dependent quantization (DQ) and the context dependency in entropy coding were fully considered. Features that could describe context state transition in the coding process were proposed to estimate rate of some syntax elements in a TU preliminarily. Secondly, coefficient chaos and sparsity features were proposed to distinguish the influence of coefficient distribution difference on the rate cost based on the coefficient distribution characteristics which built a TU level rate model. Finally, large-size transform unit (TU) and small-size TU was modeling respectively according to the rate composition character to achieve more accurate rate estimation. A large number of parameters were trained by regression model through statistical methods, and the final linear rate model was obtained which was applied to the mode decision. Experimental results show that the proposed algorithm can achieve 16.289% complexity reduction with 1.567% BD-BR increase for RA configuration.

Key words: rate estimation, VVC, RDO, regression training

收稿日期: 2022-05-13; 修回日期: 2022-10-20

通信作者: 殷海兵, yhb@hdu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61972123, No.62031009); 浙江省尖兵研发攻关计划项目 (No.2022C01068)

Foundation Items: The National Natural Science Foundation of China (No.61972123, No.62031009), Zhejiang Provincial Pioneer Research and Development Project (No.2022C01068)



0 引言

为满足不断增长的视频压缩需求, JVET (Joint Video Exploration Team) 提出新一代通用视频编码^[1], 其中采用诸多新的编码技术, 如多种划分模式、多变换核选择、改进的熵编码^[2]和依赖性量化^[3]等。相比上一代高效视频编码 (high efficiency video coding, HEVC)^[4], VVC 以复杂度增加为代价, 提高约 50% 的视频编码效率。

在 VVC 中模式决策过程基于率失真优化 (rate distortion optimization, RDO)^[5]技术实现。尽管丰富的多模式预测和自适应 DQ 技术的使用极大提升了编码性能, 但是大量的候选模式通过预测、变换、量化和熵编码过程获取码率代价, 这加剧了率失真优化过程的计算复杂度和串行依赖, 为视频编码标准的实际应用带来困难。

因此一些学者针对率失真优化过程中的码率计算复杂度问题展开研究。在 H.264/AVC^[6]中, Sarwer 等^[7]基于上下文自适应变长编码 (context adaptive variable length coding, CAVLC) 特性, 使用变换块系数绝对值之和、高频区域非零系数位置和非零系数数量来预估对应残差块码率; Zhao 等^[8]基于零均值广义高斯分布模型, 使用量化系数加权来预测变换块码率; Tu 等^[9]简化 RDO 流程, 使用变换域中非零系数数量与变换系数和来预估变换块码率。在 H.265/HEVC 中, 一些学者^[10-12]从信息熵角度出发, 统计各语法元素的二元结果, 计算 0 和 1 符号的信息熵, 以此来预估系数码率; Hu 等^[13]和 Huang 等^[14]从硬件友好的角度出发, 使用变换系数和位置信息进行快速的系数码率预估; Sheng 等^[15]基于上下文自适应二进制算术编码 (context-based adaptive binary arithmetic coding, CABAC) 特性, 使用量化系数幅度和位置特征来预估系数块码率; Liu 等^[16]根据上下文分类特性对系数组 (coefficient group, CG) 进行二元分类, 并使用 6 个系数分布特征预估部

分语法元素码率; Sun 等^[17]在变换域建立阈值区分变换系数是否量化为 0, 并使用变换系数幅值和位置信息预估码率; 孟翔等^[18]对不同位置的量化系数进行自适应加权, 并结合位置特征, 实现较为准确的变换块码率预估。然而已有算法均是在 H.264 和 H.265 平台上实现的, 而 VVC 使用多种划分方式和改进的熵编码等, 导致现有算法很难深入描述 VVC 中上下文更新时影响码率变化的根本因素, 因此进行系统精确的码率建模仍面临挑战。

针对以上问题, 本文提出一种基于统计建模的 VVC 快速码率估计算法减少模式决策复杂度。首先, 本文基于 DQ 量化行为和熵编码特性提出可以准确刻画编码过程中上下文状态迁移的码率特征, 初步估计部分语法元素码率; 其次, 本文从系数分布特性出发, 提出系数混乱度特征和稀疏度特征来减小预测误差, 并构建整个变换单元 (transform unit, TU) 级码率模型; 最后, 根据码率构成特性, 将 TU 分为大尺寸和小尺寸建模处理, 可以更准确地预估码率。

1 语法元素编码结构

在模式决策过程中, VVC 通过拉格朗日公式进行率失真代价的计算:

$$J_{RD} = D + \lambda \times R \quad (1)$$

其中, J_{RD} 为最小的率失真代价, D 为原始像素块与重构像素块之间的失真, λ 为拉格朗日因子, R 为编码当前单元的比特消耗。VVC 在对当前变换块进行码率代价计算时会将编码分为两个部分, 一个是对量化系数进行编码, 另一个是对头信息进行编码。通常编码量化系数所需码率占据码率代价主要部分, 因此建立量化系数的高准确性码率模型很有必要。编码器在对量化系数进行编码时, 首先会对当前 TU 中最后一个非零系数的位置信息 (last position, LP) 进行编码。然后将 TU 划分成系数组 (coefficient group, CG), 以反向扫

描顺序对每个 CG 的系数子块标志(coefficient sub block flag, CSBF) 进行编码, CSBF 表示该 CG 内是否存在非零系数。最后对各非零 CG 中的量化系数进行多次扫描, 编码其他语法元素。

第 1 次系数扫描编码时, 有效系数标志(significant coefficient flag, SCF) 表示当前系数是否非零, 系数大于 1 标志(greater than 1, GT1) 表示系数绝对值是否大于 1, 奇偶校验标志(parity, PAR) 表示系数绝对值减去 2 后的奇偶性, 系数大于 3 (greater than 3, GT3) 标志表示当前系数绝对值是否大于 3。第 2 次系数扫描时, 剩余(remaining, REM) 系数标志表示系数绝对值减去 4 的剩余部分。VVC 对第 1 次扫描中已经编码的上下文比特数量有所限制^[2], 如果在扫描位置开始时, 变换块中 SCF、GT1、PAR 和 GT3 编码的比特总数超过一定值, 则在此处终止第 1 次和第 2 次系数扫描, 剩余扫描位置的系数绝对值 $|q|$ 在第 3 次扫描过程中进行旁路编码, 由语法元素剩余系数绝对值标志(dec-abs-level, DEC) 表示。在第 4 次系数扫描时, 系数符号标志(coefficient sign flag, CSF) 表示当前系数是否为正。VVC 语法元素扫描编码图如图 1 所示, 其中闪电标志表示在此处终止第 1 次和第 2 次系数扫描。

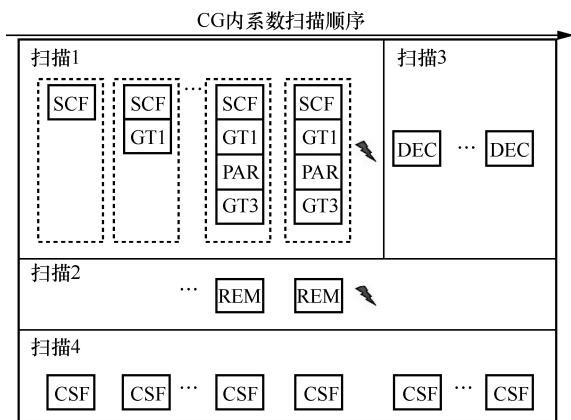


图 1 VVC 语法元素扫描编码图

在 DQ 选择最优量化索引的率失真优化过程和熵编码过程中, 会对语法元素进行复杂的上下

文建模和概率状态更新, 这会带来极大的上下文依赖。此外 DQ 量化技术基于维特比搜索算法进行最优量化路径的选择, 这样可以实现更好的比特节省, 但是其计算复杂度更高且不利于硬件并行化处理。因此如何避免复杂的 DQ 量化和上下文更新进行系统准确的码率建模对于快速的码率估计算法是非常关键的。

2 量化系数码率模型

首先本文使用相对简单的硬决策量化(hard decision quantization, HDQ)^[19]代替复杂度较高的 DQ 量化方式。其次, 为了准确预估量化系数比特, 本文基于 DQ 量化行为和熵编码中的语法元素上下文分类原则开发整体码率特征, 并验证和使用能够表示系数分布特性的混乱度和稀疏度特征, 共同开发了 TU 级码率模型。最后在建模过程中发现 VVC 中存在小于 16 个系数^[2]的 TU, 即 2×2 、 2×4 、 4×2 尺寸的 TU, 这些小尺寸 TU 和大尺寸 TU 的码率构成相差很大, 因此本文将它们分开进行码率建模。

2.1 大尺寸 TU 上下文特征

在现有码率估计算法中, Liu 等^[16]根据上下文分类原则对系数组进行二元分类, 并使用 6 个系数分布特征预估 CG 内语法元素 SCF、GT1 和 GT3 的码率。该算法在 H.265 平台上可以较为准确地预估系数块码率, 但是 VVC 中上下文转换原则发生改变^[2], 使得其码率特征很难深入刻画上下文状态转换时码率变化的根本因素。VVC 局部邻域模板图和频区分布图如图 2 所示。图 2(a)为 VVC 从当前量化系数与邻近系数之间的依赖关系出发提出的局部邻域模板, 其中黑色方块表示待编码系数, 灰色方块表示其局部邻域系数。根据局部统计特性^[20], 语法元素 SCF、PAR、GT1 和 GT2 的上下文概率模型受多种因素影响, 包括有当前系数位置、局部邻域系数的部分重建值之和、局部邻域中非零系数数量和 DQ 状态。局部邻域系



数部分重建值之和 Q 表示如下。

$$Q = \sum_{n=1}^5 (\text{scf}_n + \text{gt1}_n + \text{par}_n + 2 \times \text{gt3}_n) \quad (2)$$

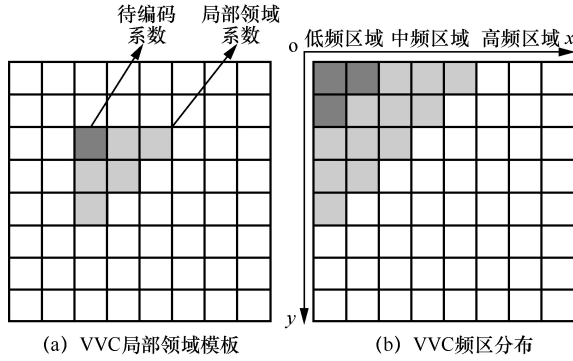


图2 VVC 局部邻域模板图和频区分布图

其中, n 表示当前扫描系数邻域索引, 对应图 2(a) 中的 5 个局部邻域系数。 scf_n 表示索引为 n 的邻域系数语法元素 SCF 的值, 即当该邻域系数非零时, scf_n 为整数 1, 反之则为 0。而 gt1_n 、 par_n 和 gt3_n 也分别对应索引为 n 的邻域系数语法元素 GT1、PAR 和 GT3 的值。尽管在多次系数扫描的上下文编码过程中, VVC 充分利用系数之间的相关性提升了编码性能, 但这会导致更高的编码复杂度和系数依赖。因此本文对上下文状态转换中码率代价变化的深层原因进行分析。

首先, 残差系数经过变换后, 其能量主要集中在中低频区域。在中低频区域比高频区域更容易出现非零系数, 此时编码零系数所需码率也更多。中低频区域分布如图 2(b) 所示, 其中系数坐标和小于 2 和小于 5 时分别属于低频和中频区域。因此本文将 TU 内低频区域的非零系数数量 ε_1 、中频区域的非零系数数量 ε_2 作为衡量频域分量对码率代价影响程度的重要特征。其次在当前扫描系数的局部邻域中, 重建值之和越大、非零系数越多, 则当前系数更倾向于量化为非零, 此时编码零系数所需码率更大。因此本文将 TU 中系数重建值之和 ε_3 与非零系数数量 ε_4 作为对码率代价有重要影响的统计特征。

考虑各个特征对码率代价的影响及模型复杂

度, 本文提出以下线性模型来预估语法元素 SCF、PAR、GT1、GT2 和 CSF 的码率和 R_1 。由于 CSF 语法元素码率代价与 TU 中非零系数数量有直接关系^[4], 因此算法将语法元素 CSF 码率成本考虑在内, 模型定义如下。

$$R_1 = \theta_1 \varepsilon_1 + \theta_2 \varepsilon_2 + \theta_3 \varepsilon_3 + \theta_4 \varepsilon_4 \quad (3)$$

其中, ε_1 、 ε_2 分别为 TU 中横纵坐标之和小于 2 和小于 5 时系数非零的数量, ε_3 表示 TU 中系数的重建值之和, ε_4 表示 TU 中非零系数数量。 $\theta_1 \sim \theta_4$ 为权重系数, 由线性回归训练得到。为提高模型准确度, 算法在区分亮度块和色度块基础上, 按照 TU 尺寸进行分类拟合。特别地, 当变换块为色度块时, 考虑块内量化系数相对较少, 此时不用特征 ε_2 。TU 级别 R_1 与真实码率之间的关系如图 3 所示。实验编码模式为 RA 模式, 训练样本从序列 BasketballPass 中获得, QP 为 32。本文只列出具有代表性的 TU 尺寸即 64、128、256、512 的情况, 对于没有列出的 TU 尺寸, 即 16、32、1 024、2 048 和 4 096 的情况, 其分布大致类似。

2.2 位置参数和系数分布特征

在初步预估语法元素码率和 R_1 后, 为了更清晰地针对其他语法元素开发码率特征, 本文统计了 VVC 中不同 TU 尺寸和不同 QP 下语法元素的码率占比结果, 如图 4 所示。其中 R_s 表示语法元素 SCF、GT1、PAR、GT2 和 CSF 的码率和占总码率的比重, LP 表示语法元素 LP 的码率占比, REM、DEC 和 CSBF 也表示对应语法元素的码率占比情况。从图 4 中可以看出 R_s 码率占比最大, LP 语法元素码率占比相对均匀, 而 REM、DEC 和 CSBF 综合码率占比相对较小。考虑模型的复杂度, 算法只将 LP 语法元素码率特征加入码率模型中。由于 LP 语法元素码率与位置信息相关, 因此本文使用最后一位有效系数的位置特征 R_{LP} 来预估编码 LP 所消耗的码率, 定义如下:

$$R_{LP} = i + j + 1 \quad (4)$$

其中, i 、 j 分别表示最后一位有效系数的横坐标和纵坐标。

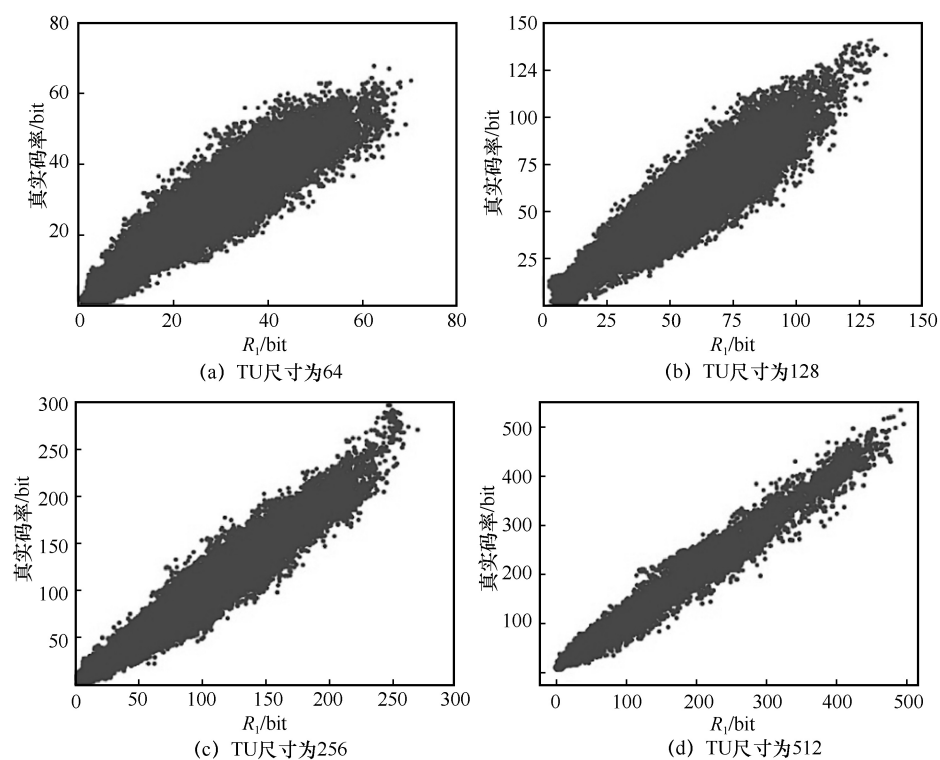
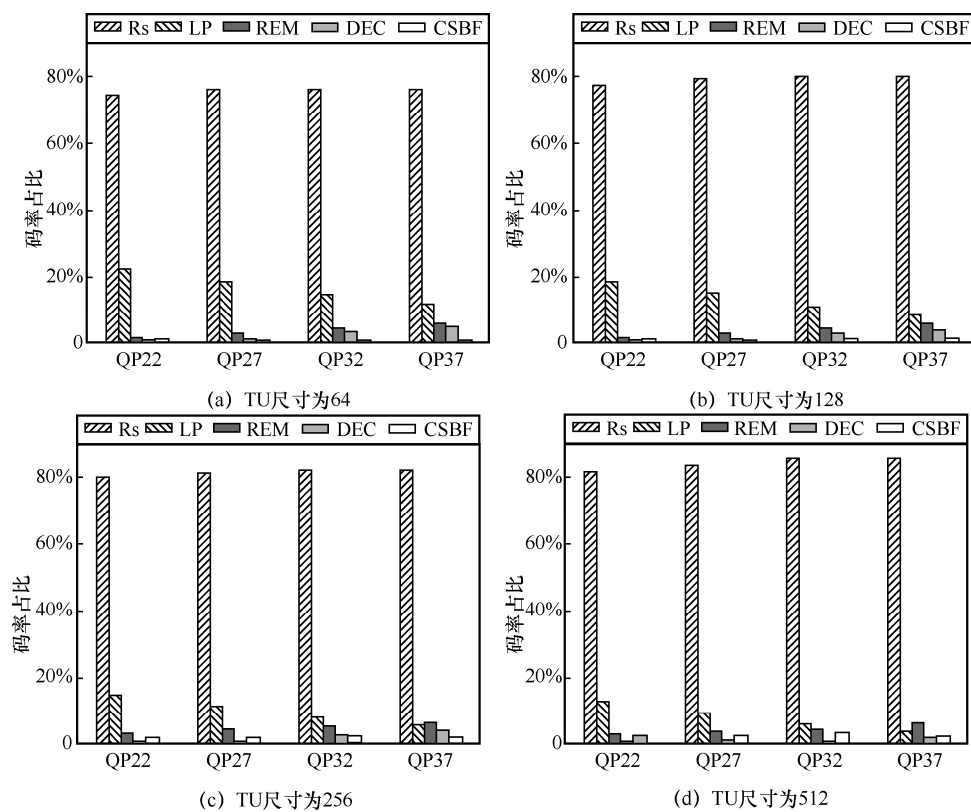
图 3 TU 级别 R_1 与真实码率之间的关系

图 4 不同 TU 尺寸和不同 QP 下语法元素的码率占比结果



为提高模型准确度, 本文将可以表示系数分布特性的特征考虑在内。通常系数的混乱程度和稀疏程度可以表示变换块系数分布特性, 对系数编码的比特产生影响。在系数扫描编码过程中, 零系数与非零系数转换的次数越多, 编码器在最大概率符号 (most probable symbol, MPS) 和最小概率符号 (least probable symbol, LPS) 之间的转换次数也越多, 则系数分布越混乱, 编码变换块所需比特就越多。因此本文定义混乱度特征 λ_1 为 CG 中反向扫描顺序下零系数与非零系数的转换次数。而当 CG 内第一个非零系数与最后一个非零系数之间的零系数数量较少时, 表明非零系数分布更加集中, 此时编码 CG 所需比特就越少。本文将系数稀疏度特征 λ_2 定义为 CG 内第一个与最后一个非零系数之间零系数个数。系数分布特征定义如下。

$$R_s = \sum_{m=1}^{N/16} (\lambda_1 + \lambda_2) \quad (5)$$

其中, N 表示变换单元的尺寸。当 QP 为 37 时, 系数分布特征与真实码率之间的关系如图 5 所示。

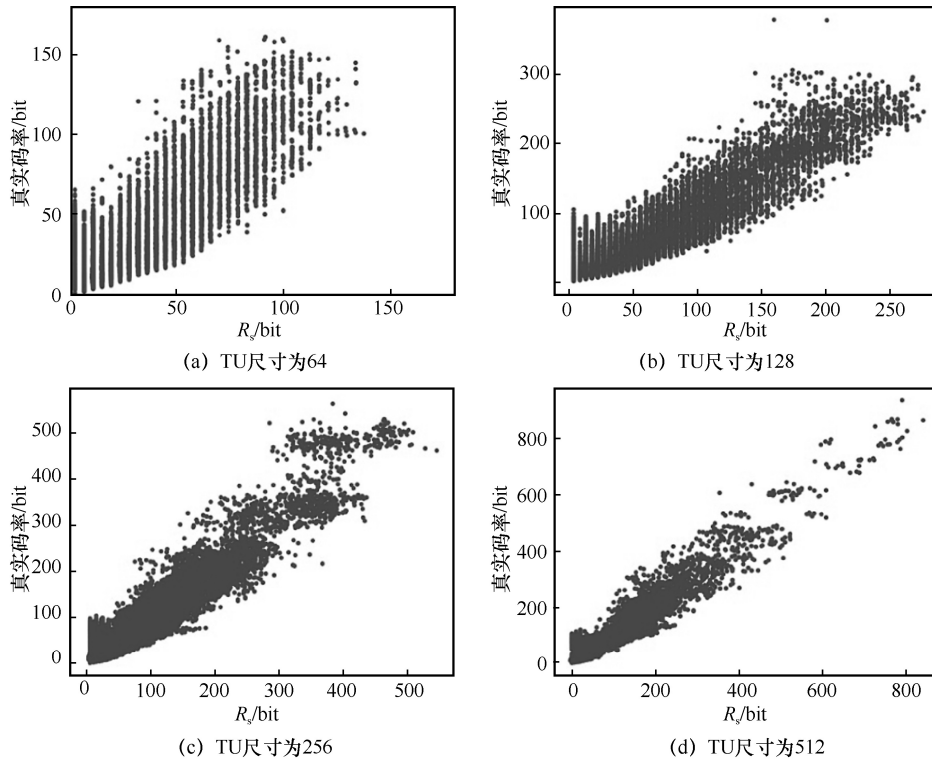


图 5 系数分布特征与真实码率之间的关系

2.3 大尺寸 TU 最终码率模型

经过上述分析, 大尺寸 TU 最终码率模型如下。

$$R_{\text{est}} = u_0 + u_1 R_l + u_2 R_{\text{LP}} + u_3 R_s \quad (6)$$

其中, R_{est} 表示大尺寸 TU 中量化系数的预估码率, R_l 表示语法元素 SCF、GT1、PAR、GT2 和 CSF 的预估码率, R_{LP} 表示为 TU 中最后一位有效系数的位置, R_s 表示 TU 系数分布特征。 $u_0 \sim u_3$ 为模型参数, 由线性回归训练得到。大尺寸 TU 预估码率与真实码率之间的关系如图 6 所示, 即 QP 为 32 时, R_{est} 与真实码率之间的关系, 测试序列为 BasketballPass, 可以看出两者之间有着很强的线性关系。

2.4 小尺寸 TU 最终码率模型

由于小尺寸 TU 系数较少, 处于不同扫描位置的系数对码率代价造成的影响区别较大, 因此本文使用加权量化系数^[18]的方法来描述位置信息对码率代价的不均匀贡献, 并使用全局混乱度特征减小预估码率与真实码率间的误差。

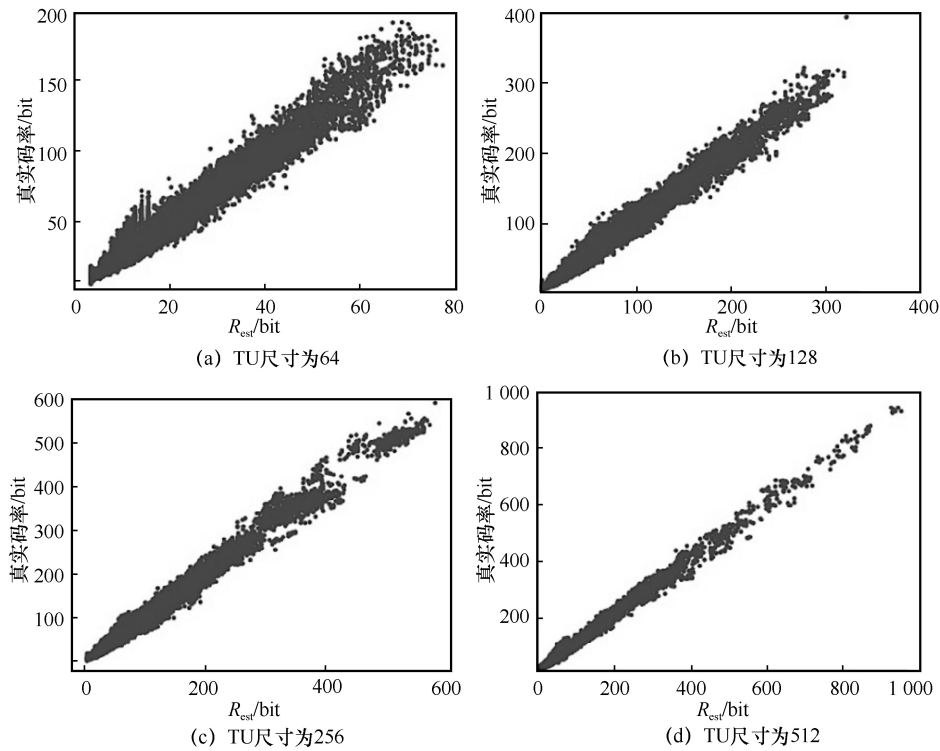


图6 大尺寸 TU 预估码率与真实码率之间的关系

最终模型如下。

$$R_{\text{est}2} = \sum_{i=1}^N w_i l_i + \sum_{m=1}^{N/4} \lambda_1 \quad (7)$$

其中, w_i 为权重系数, l_i 为量化系数, N 为 TU 尺寸。对于 2×2 尺寸的 TU 块, N 为 4; 对于 2×4 和 4×2 尺寸的 TU 块, N 为 8。 λ_1 表示 CG 中零系数和非零系数的翻转次数。小尺寸 TU 预估码率与真实码率之间的关系如图 7 所示, 测试序列为 BasketballPass。

3 实验结果分析

首先为评估第 2 节提出的码率模型精度, 本文对不同情况下预估码率和真实码率的误差进行统计, 并对不同区间误差的概率进行归一化处理, 码率误差的归一化概率如图 8 所示。其中, 方形符号、三角符号和圆点分别表示文献[17]、文献[18]和本文提出的码率预估算法的误差归一化概率图, 需要说明的是, 编码块预估码率会在真实码率附近

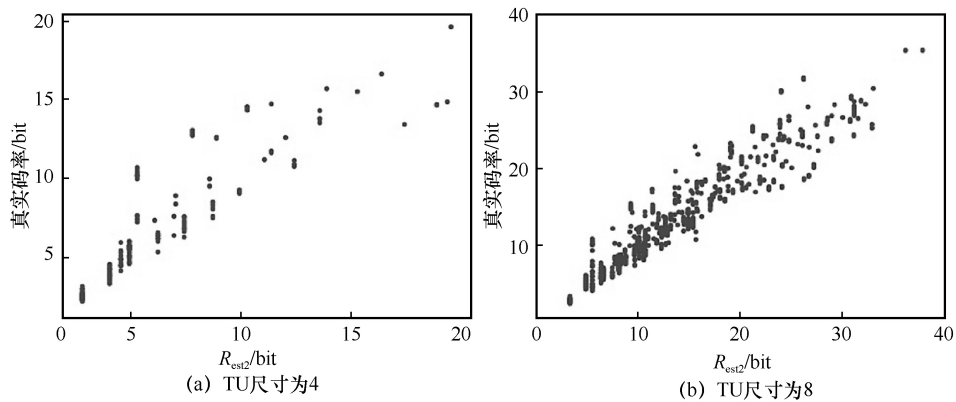


图7 小尺寸 TU 预估码率与真实码率之间的关系



浮动,当预估码率小于真实码率时,对应预测误差为负值。测试序列为纹理细节较多的 BasketballDrillText。可以看出, Sun 等^[17]提出的算法由于特征相对较少,对码率变化的深层原因描述不够精准,导致模型精度不够稳定,误差相对较大。而孟翔等^[18]提出的模型由于 VVC 中上下文分类原则转变导致其算法适用性不强,模型精度相对较低。因此本文提出的码率预估算法相对更加准确,预估误差也更小。

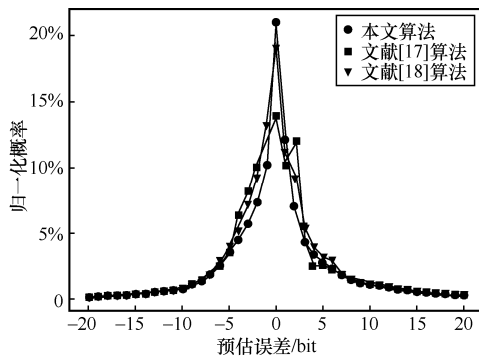


图8 码率误差的归一化概率

其次为评估提出的快速码率失真估计算法在 RD 性能和时间复杂度上的表现,使用 VVC 参考代码 VTM11.0 版本进行实验。实验采用 JVET 通用测试条件 (common test condition, CTC)^[21]进行测试,关闭 RDOQ 和 DQ 功能,关闭变换跳过功能,其余均为默认配置。QP 值取值为{22, 27, 32, 37},使用码率变化率 BD-Rate^[22]作为评价编码性能的指标。为衡量算法时间复杂度变化,本文定义时间节省 ΔT_{RDO} 计算,如下所示。

$$\Delta T_{\text{RDO}} = \frac{T_{\text{VTM_RDO}} - T_{\text{PRO_RDO}}}{T_{\text{VTM_RDO}}} \times 100\% \quad (8)$$

其中, $T_{\text{PRO_RDO}}$ 为本算法进行模式决策过程所消耗的时间, $T_{\text{VTM_RDO}}$ 为原始 VVC 算法进行模式决策过程所消耗的时间, ΔT_{RDO} 为正则表示编码时间的减少。在 Random Access 下模型性能见表 1,其中列出本文所提算法和文献[17-18]所提方法在 VTM11.0 上的实验结果对比情况,其中包括每个类不同序列的编码性能表现和时间节省。

在 Sun 等^[17]的算法中,将 TU 中变换系数幅度,非零系数位置和最后一位非零系数位置作为特征进行码率建模。由于算法码率特征较少,很难深入描述上下文更新时码率变化的根本因素,在 VVC 中造成较大的性能损失。具体地说,对于一些高分辨率视频序列,如 Tango2、Campfire 和 ParkRunning3, Sun 等^[17]的算法 BD-BR 分别上升 1.83%、1.92%、1.65%,而本文算法在这些序列的 BD-BR 分别上升 1.52%、1.75%、1.25%。同时对于一些低分辨率视频序列,如 BasketballPass、RaceHorses 和 BasketballDrillText, Sun 等^[17]的性能损失分别为 2.06%、2.46%、2.59%,本文算法在这些序列的性能损失分别为 1.10%、1.44%、1.57%。这说明对于不同分辨率的视频序列,本文算法皆比 Sun 等^[17]的算法实现更好的编码效率。而在时间复杂度方面,对于高分辨率视频序列,如 Tango2 和 Campfire,时间节省分别为 20.34% 和 17.65%,最大值和最小值相差 2.69%,而本文在高分辨率对应视频序列时间节省分别为 21.36% 和 19.86%,最大值和最小值相差仅 1.5%。同时对于低分辨率视频序列如 BasketballDrillText 和 SlideShow,时间节省分别为 12.32% 和 8.14%,两者相差 4.18%,而本文算法中对对应低分辨率视频序列时间节省分别为 11.03% 和 8.67%,两者相差仅 2.46%。这说明对于不同分辨率的视频序列,而本文算法相比 Sun 等^[17]的算法时间节省更加均匀,算法适用性更好。

在孟翔等^[18]所提算法中,对位于不同位置的量化系数进行自适应加权,并根据上下文分类依据对 CG 进行分组,在 HEVC 中可以实现相对准确的码率预估。但是 VVC 中划分模式的增加和上下文选取原则的改变^[2]使得文献[15]的算法在 VVC 中的效果并不突出。具体地说,对于一些纹理细节较多的视频序列,如 FoodMarket4、BQSquare 和 BasketballDrillText,其 BD-BR 分别上升 2.03%、2.15%、1.68%,而在本文的算法中

表 1 在 Random Access 下模型性能

类别	视频序列	Sun 等 ^[17] 所提算法		孟翔等 ^[18] 所提算法		本文算法	
		BD_BR	ΔT_{RDO}	BD_BR	ΔT_{RDO}	BD_BR	ΔT_{RDO}
A1	Tango2	1.83%	20.34%	1.79%	20.32%	1.52%	21.36%
	FoodMarket4	2.01%	21.11%	2.03%	19.93%	1.94%	20.91%
	Campfire	1.92%	17.65%	1.77%	18.64%	1.75%	19.86%
A2	CatRobot	2.03%	20.02%	1.72%	21.16%	1.64%	21.62%
	DaylightRoad2	2.16%	20.61%	1.64%	19.55%	1.53%	20.64%
	ParkRunning3	1.65%	18.41%	1.38%	18.79%	1.25%	19.62%
B	MarketPlace	2.24%	17.74%	2.15%	17.26%	1.90%	18.31%
	RitualDance	1.92%	19.32%	1.12%	18.20%	1.13%	19.47%
	Cactus	2.79%	16.57%	2.14%	15.44%	2.13%	17.20%
	BasketballDrive	1.87%	17.21%	1.50%	16.72%	1.49%	16.52%
	BQTerrace	2.04%	17.42%	1.96%	17.37%	1.91%	16.78%
C	BasketballDrill	1.81%	16.13%	1.74%	15.29%	1.51%	17.49%
	BQMall	1.94%	15.34%	1.61%	16.53%	1.59%	16.34%
	PartyScene	2.86%	14.65%	2.17%	13.62%	2.24%	15.91%
	RaceHorses	1.98%	15.36%	1.72%	13.87%	1.60%	14.20%
D	BasketballPass	2.06%	15.92%	1.40%	15.52%	1.10%	14.63%
	BQSquare	2.57%	14.87%	2.15%	14.29%	2.02%	14.74%
	BlowingBubbles	1.57%	13.89%	1.22%	14.84%	0.97%	15.24
	RaceHorses	2.46%	13.87%	1.61%	13.66%	1.44%	14.10%
F	BasketballDrillText	2.59%	12.32%	1.68%	10.26%	1.57%	11.03%
	ArenaOfValor	2.16%	11.26%	1.81%	9.87%	1.63%	10.25%
	SlideEditing	1.87%	10.59%	1.19%	9.61%	0.95%	9.74%
	SlideShow	1.74%	8.14%	1.36%	8.19%	1.21%	8.67%
平均值		2.09%	16.032%	1.690%	15.61%	1.567%	16.289%

其性能分别提升 0.09%、0.13%、0.11%。对于背景较为均匀的视频序列，如 RitualDance、BlowingBubbles 和 SlideEditing，其 BD-BR 分别上升 1.12%、1.22%、1.19%，在本文的算法中其性能分别提升 0.01%、0.25%、0.24%。因此对于纹理细节不同的视频序列来说，本文从上下文选取原则角度出发建立的 TU 级码率模型相较于孟翔等^[18]的算法可以实现更加准确的码率预估，性

能损失更小。而在时间复杂度方面，孟翔等^[18]的算法码率模型参数较多，因此 RDO 节省时间较少。具体地说，对于纹理细节较多的视频序列，如 FoodMarket4 和 BasketballDrillText，时间节省分别为 19.93%和 10.26%，而本文算法对应时间节省为 20.91%和 11.03%。而对于背景较为均匀的视频序列，如 RitualDance 和 BlowingBubbles，时间节省分别为 18.2%和 14.84%，而本文对应时间节



省为 19.47%和 15.24%。可以看出无论是在纹理细节较多还是背景相对均匀的视频序列中,本文算法相比于孟翔等^[18]所提算法可以实现更大的时间节省。

4 结束语

本文针对 VVC 率失真优化过程中复杂的码率计算问题,提出一种基于统计建模的快速码率估计算法。算法首先充分考虑依赖性量化的量化行为和熵编码中的上下文依赖,提出一种准确预估部分语法元素码率的初步模型;然后提出混乱度和稀疏度特征描述系数整体分布对码率的影响,并进一步形成 TU 级码率模型;最后根据码率构成特性,算法将大尺寸 TU 和小尺寸 TU 分开建模实现更精准的码率预估。实验结果表明,在 RA 配置下,提出方案在性能基本不变的情况下节省 16.289%的率失真优化时间。

参考文献:

- [1] BROSS B, WANG Y K, YE Y, et al. Overview of the versatile video coding (VVC) standard and its applications[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(10): 3736-3764.
- [2] SCHWARZ H, COBAN M, KARCZEWICZ M, et al. Quantization and entropy coding in the versatile video coding (VVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(10): 3891-3906.
- [3] SCHWARZ H, NGUYEN T, MARPE D, et al. Hybrid video coding with trellis-coded quantization[C]//Proceedings of 2019 Data Compression Conference (DCC). Piscataway: IEEE Press, 2019: 182-191.
- [4] 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.
- [5] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998, 15(6): 74-90.
- [6] WIEGAND T, SULLIVAN G J, BJONTEGAARD G, et al. Overview of the H.264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [7] SARWER M G, PO L M. Fast bit rate estimation for mode decision of H.264/AVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(10): 1402-1407.
- [8] ZHAO X, SUN J, MA S W, et al. Novel statistical modeling, analysis and implementation of rate-distortion estimation for H.264/AVC coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(5): 647-660.
- [9] TU Y K, YANG J F, SUN M T. Efficient rate-distortion estimation for H.264/AVC coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(5): 600-611.
- [10] CHEN W G, WANG X. Fast entropy-based CABAC rate estimation for mode decision in HEVC[J]. SpringerPlus, 2016, 5(1): 1-10.
- [11] CHEN H C, CHANG T S. Fast rate distortion optimization with adaptive context group modeling for HEVC[C]//Proceedings of 2017 IEEE International Symposium on Circuits and Systems. Piscataway: IEEE Press, 2017: 1-4.
- [12] SHARABAYKO M P, PONOMAREV O G. Fast rate estimation for RDO mode decision in HEVC[J]. Entropy, 2014, 16(12): 6667-6685.
- [13] HU L D, SUN H M, ZHOU D J, et al. Hardware-oriented rate-distortion optimization algorithm for HEVC intra-frame encoder[C]//Proceedings of 2015 IEEE International Conference on Multimedia & Expo Workshops. Piscataway: IEEE Press, 2015: 1-6.
- [14] HUANG X F, JIA H Z, CAI B B, et al. Fast algorithms and VLSI architecture design for HEVC intra-mode decision[J]. Journal of Real-Time Image Processing, 2016, 12(2): 285-302.
- [15] SHENG Z, ZHOU D, SUN H, et al. Low-complexity rate-distortion optimization algorithms for HEVC intra prediction[C]//Proceedings of International Conference on Multimedia Modeling. Cham: Springer, 2014: 541-552.
- [16] LIU Z Y, GUO S C, WANG D S. Binary classification based linear rate estimation model for HEVC RDO[C]//Proceedings of 2014 IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2014: 3676-3680.
- [17] SUN H M, ZHOU D J, HU L D, et al. Fast algorithm and VLSI architecture of rate distortion optimization in H.265/HEVC[J]. IEEE Transactions on Multimedia, 2017, 19(11): 2375-2390.
- [18] 孟翔, 殷海兵, 黄晓峰. 基于统计建模的 HEVC 快速率失真估计算法[J]. 电信科学, 2021, 37(1): 58-68.
- [19] MENG X, YIN H B, HUANG X F. Statistical modeling based fast rate distortion estimation algorithm for HEVC[J]. Telecommunications Science, 2021, 37(1): 58-68.
- [20] SCHWARZ H, NGUYEN T, MARPE D, et al. Hybrid video coding with trellis-coded quantization[C]//Proceedings of 2019 Data Compression Conference (DCC). Piscataway: IEEE Press, 2019: 182-191.
- [21] SCHWARZ H, NGUYEN T, MARPE D, et al. Improved quantization and transform coefficient coding for the emerging versatile video coding (VVC) standard[C]//Proceedings of 2019

IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2019: 1183-1187.

[21] BOSSEN F, BOYCE J, LI X, et al. JVET common test conditions and software reference configurations for SDR video[EB]. 2019.

[22] BJONTEGAARD G. Calculation of average PSNR differences between RD-curves [EB]. 2001.

[作者简介]



祁伟（1995—），男，杭州电子科技大学硕士生，主要研究方向为视频编解码。



王鸿奎（1990—），男，博士，杭州电子科技大学讲师，主要研究方向为感知视频编码。



黄晓峰（1988—），男，博士，杭州电子科技大学教授，主要研究方向为感知视频编码。



殷海兵（1974—），男，博士，杭州电子科技大学教授，主要研究方向为数字视频编解码。



牛伟宏（1998—），男，杭州电子科技大学硕士生，主要研究方向为视频编解码。