

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2021.0528

HEVC 对偶编码单元划分优化算法

刘美琴^{1 2}, 徐晨铭^{1 2}, 姚超^{3 *}, 林春雨^{1 2}, 赵耀^{1 2}

(1. 北京交通大学 信息科学研究所, 北京 100044; 2. 现代信息科学与网络技术北京市重点实验室, 北京 100044;

3. 北京科技大学 计算机与通信工程学院, 北京 100083)

摘 要: 为了解决视频数据量日益增长与用户享受高质量视频体验需求之间的矛盾, HEVC 在 H.264/AVC 标准的基础上通过引入新型的编码结构和算法进一步将编码效率提升了 50%, 但是也极大地提升了编码复杂度。基于此, 提出对偶编码单元 (CU) 划分网络 DualNet, 来降低 HEVC 中帧内编码复杂度。该网络由预测网络和目标网络 2 个部分组成, 其中, 预测网络通过分析图像统计特征实现编码单元划分决策, 从而跳过四叉树的遍历搜索, 提高编码单元划分决策的时间效率; 目标网络基于率失真代价评价和优化决策模型提升编码单元划分性能, 实现模型互补和最优率失真估计。实验结果表明: 与 HEVC 标准对比, 所提算法在实现相近的压缩效果的前提下能够节省 64.06% 的编码时间。

关 键 词: 视频编码; H.265/HEVC; 编码单元 (CU) 划分; 深度学习; 对偶神经网络

中图分类号: TP391

文献标志码: A

文章编号: 1001-5965(2022)08-1383-07

ITU-T 与 ISO/IEC 标准化组织制定了一系列视频编码标准^[1], 如 H.262^[2]、H.264/AVC^[3]、H.265/HEVC^[4]、H.266/VVC^[5] 等。与 H.264/AVC 标准相比, H.265/HEVC 标准提高了帧内预测、帧间预测的性能, 在获得相当解码重建质量的情况下, 可以提升 50% 以上的编码效率。其中, HEVC 使用编码单元 (coding unit, CU) 替代了 H.264/AVC 中的宏块单元 (macro block, MB), CU 结构的划分占据了大量的编码时间, 在提高编码效率的同时, 也急剧增加了编码复杂度, 如基于四叉树的 CU 递归划分方式占用了 HEVC 编码时间的 80%。因此, 如何优化 CU 划分方式、降低 HEVC 的编码复杂度成为了当前的研究热点。

HEVC 将每帧图像划分为编码树单元 (coding tree unit, CTU)^[6], CTU 采用四叉树结构可以

进一步划分为多个 CU。为寻找 CU 的最佳划分方式, HEVC 需遍历 CTU 中所有从 64×64 到 8×8 不同尺度的 CU, 并分别计算率失真代价^[7]。另外, 在确定 CU 最佳划分方式的过程中, 需通过计算编码失真和码率, 拟合最小率失真代价。因此, 在 HEVC 中实现 CU 最优划分所采用的递归搜索算法具有较高的时间复杂度, 需要消耗大量的编码时间。因此, 本文提出基于对偶网络的 CU 划分算法 DualNet。在 CU 划分过程中, 构建基于卷积神经网络 (convolutional neural networks, CNN) 的预测网络, 先提取图像统计性特征, 预测当前编码帧可能的 CU 划分模式; 再构建目标网络, 通过基于率失真函数的强化学习方法决策最优的 CU 划分模式, 反馈训练预测模型, 实现 CU 划分的最优率失真估计, 避免遍历所有 CU 划分模式, 以加

收稿日期: 2021-09-06; 录用日期: 2021-09-17; 网络出版时间: 2021-10-11 16:35

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20211011.1531.001.html

基金项目: 国家自然科学基金 (61972028, 61902022, 62120106009); 中央高校基本科研业务费专项资金 (2019JBM018, FRF-TP-49-015A1)

* 通信作者: E-mail: yaochao@ustb.edu.cn

引用格式: 刘美琴, 徐晨铭, 姚超, 等. HEVC 对偶编码单元划分优化算法 [J]. 北京航空航天大学学报, 2022, 48(8): 1383-1389.
LIU M Q, XU C M, YAO C, et al. Dual coding unit partition optimization algorithm of HEVC [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(8): 1383-1389 (in Chinese).

快 CU 划分速度,提升 HEVC 的编码性能。

综上,本文提出基于率失真约束的 DualNet 对偶优化网络,用于在降低 HEVC 帧内模式 CU 划分决策复杂度的同时提高编码质量;设计 CU 划分阈值动态调整优化策略,实现网络训练的快速优化,用于提高模型对数据及 HEVC 逆行 CU 划分决策的普适性。所提 DualNet 与 HEVC 标准编码软件(HM16.5)相比,编码时间平均降低 64.06%,而平均比特率增量 BD-BR 和平均视频客观质量提高值 BD-PSNR 仅为 2.876% 和 -0.120 dB。

1 相关工作

1.1 HEVC 编码单元划分优化算法

降低 CU 划分方式计算复杂度的方法包括启发式和基于机器学习的方式。在启发式划分算法中,文献[8-13]分别通过改进预测模式、CU 存储数据结构和图像特征优化决策模式提高编码效率。在改进预测模式方面,Zhang 等^[8]采用简化预测模式改进了粗略率失真代价的搜索策略,加快了 CU 的划分速度;Lu 和 Li^[9]提出了利用帧内相邻块预测当前 CU 划分的方法。在改进数据结构方面,Wang 等^[10]将原有四叉树形式改进为扩展四叉树结构,以存储划分结果。在改进图像特征方面,Qing 等^[11]利用图像显著性特征作为 CU 划分的判断标准;Kibeya 等^[12]利用在划分过程中产生的零星图像块处理矩阵特征,以提前判断 CU 的划分性能;朱蕾琦等^[13]利用小波变化后的低频子图做帧内模式选择,以减少 CU 划分过程中的率失真计算次数。在上述工作中,基于图像特征的算法效果最佳,大部分后续研究工作致力于通过分析图像特征来跳过率失真搜索过程。

基于机器学习的划分方式是利用机器学习模型,手动^[14-16]或自动^[17-20]提取图像特征辅助 CU 划分以代替 HEVC 中的 CU 划分方式,减少 CU 划分所需的计算时间。Zhang 等^[14]利用决策树离线训练编码信息,加快编码速度;Kim 和 Park^[15]提出了在线与离线相结合的训练模式,利用贝叶斯决策模型训练损失函数划分 CU;Fu 等^[16]利用支持向量机提取图像特征,离线训练划分模型。上述手动提取特征法难以全面获取图像特征信息,限制了其编码准确度。自动提取图像特征法则利用 CNN,自动挖掘大规模图像数据中与 CU 划分相关的特征。Liu 等^[17]首次提出了利用深度 CNN 辅助 CU 划分决策;Zhang 等^[18]利用深度 CNN 学习图像纹理辅助 CU 划分;易清明等^[19]提出了基于 Inception 模块的 CNN 结构对 CU 划分

进行提前预测;Xu 等^[20]提出了端到端的 CU 划分结构,通过三通道 CNN 综合提取图像特征。上述工作均将 CU 的划分看作二元分类问题,忽略了率失真函数编码标准,间接降低了网络的编码性能,且传统深度 CNN 对全局信息敏感度不高,也会在一定程度上限制 CU 的特征提取精度^[21]。

1.2 对偶网络

对偶网络是一种常用于知识蒸馏和模型强化的深度网络。在知识蒸馏的研究中,文献[22-24]运用了对偶网络将复杂模型转变为相对简单的模型,实现了模型的轻量化。Bae 等^[22]设计了一种多重对偶网络结构进行密集知识转移,提升了图像分类的准确度,并降低了其计算复杂度;Abbasi 等^[23]提出了通用对偶师生网络模型,运用于不同场景;Xiao 等^[24]利用对偶网络将知识转移关系转换为竞争关系,通过无监督方式提升了模型的适应性。基于上述工作,对偶网络也被运用于加强网络模型之间的相关性,提升单网络的训练效果,实现网络模型的互补。Lu 等^[25]提出了基于双域融合的图像质量评估算法,联合频率域和空间域双网络,实现模型互补;Zhou 等^[26]通过对偶网络实现了对磁共振成像的二次重建;Wang 等^[27]利用对偶蒸馏网络,丰富图像重建过程中图像细节并互相监督特征提取过程;苏志雄等^[28]构建了对偶模型,在网络计划模型计算中分别体现机动时间和路差,使网络计划更具针对性和有效性。上述工作均从对偶网络出发,利用并行网络模块解决多维问题,实现模型的互补和强化。

为了简化 HEVC 编码算法中 CU 划分算法并融合率失真函数作为 CU 划分决策的优化函数,本文利用对偶网络结构,构建 CU 划分决策模型,在保证编码性能的前提下提高 HEVC 帧内模式中 CU 的划分效率。其中,对偶网络融合了用于预测 CU 划分的预测网络和用于评价划分结果的目标网络。预测网络使用 CNN 自动提取图像的统计性像素特征,实现像素级 CU 的划分预测,端到端地输出划分的预测结果。目标网络以率失真函数约束损失函数,采用动态规划方式获得最佳划分阈值,优化网络模型参数,提高 CU 划分的精度。

2 编码单元划分对偶网络

本文构建的 CU 划分对偶网络 DualNet 结构如图 1 所示。其中,预测网络用于预测 CU 的划分,目标网络结构与预测网络相似,仅参数更新速度慢于预测网络,采用率失真代价评估预测网络

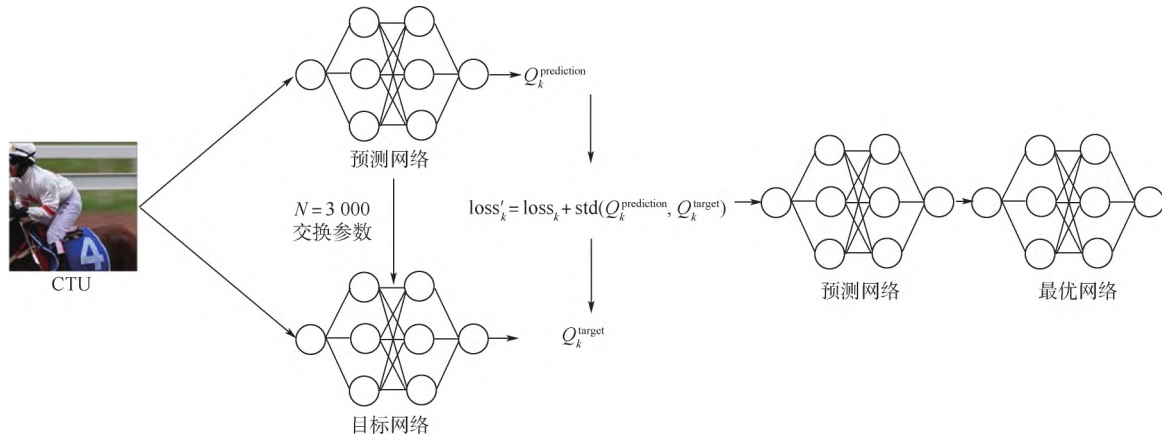


图1 DualNet 结构示意图

Fig.1 Structure of DualNet

的参数更新趋势。对偶网络预测模型评价的标准差将作为损失函数的一部分返回预测网络进行拟合,最终得到具有最优参数的网络模型。

2.1 预测网络结构

预测网络用于替代 HEVC 标准编码软件 (HEVC test model, HM) [29] 中 CU 递归划分过程,在编码过程中读取由 CNN 训练得到的划分预测模型。具体地,采用分层 CU 划分图 (hierarchical CU partition map, HCPM) [20] 存储划分真值和预测值。与 HEVC 编码单元划分中自顶向下逐层计算率失真代价不同,该结构可将 CTU 作为输入,直接得到 $1 + 4 + 16 = 21$ 个划分单元的整体输出。

CU 的划分方式与其图像特征的复杂程度密切相关。因此,预测网络借助 CNN 分析图像特征复杂度,实现 CU 的划分,如图 2 所示。具体地,预测网络首先对输入的 CTU 亮度矩阵进行去均值和降采样操作,加快网络的训练速度;其次,利用由 3 层 CNN 构成的 3 条分支 B_1 、 B_2 、 B_3 提取图像特征,以对应 HCPM 中的 CU 划分层级;然后,不同层级的特征图归并为一个特征向量,在全连接层分别由 3 条支路处理,以符合 HCPM 的 3 层输出模式和量化参数 (quantization parameters, QP) 对 CU 划分的影响;最后,采用图像真值与对应特征向量间的交叉熵损失函数评价特征提取准确度。

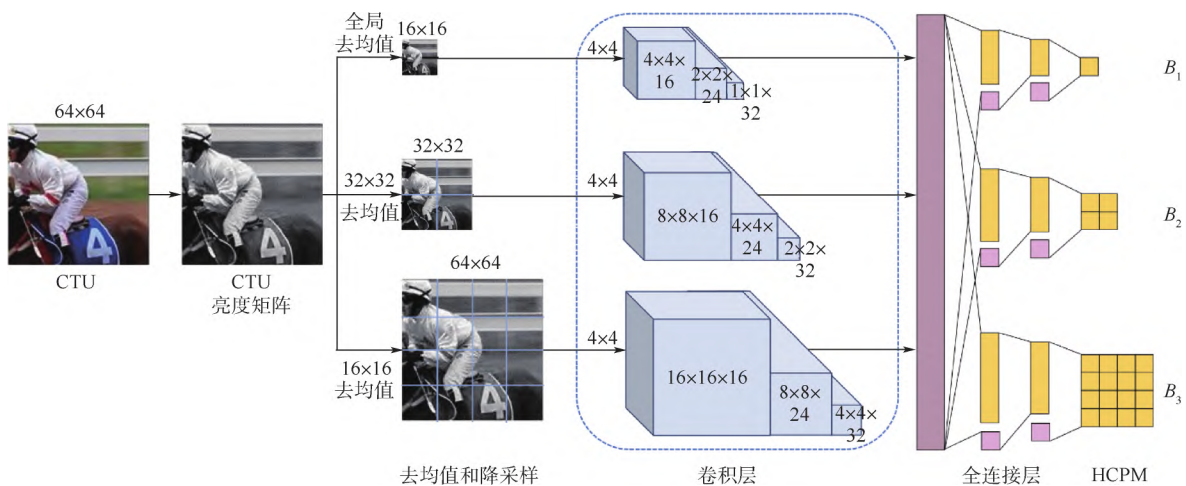


图2 预测网络结构示意图

Fig.2 Structure of prediction network

2.2 对偶编码单元划分网络

总体模型采用对偶网络优化结构,参考率失真代价建立评价体系,对预测网络进行评价和监督,不断优化预测模型参数,其结构如图 3 所示。

本文针对单一 CNN 模型无法全面考虑 CU

划分影响因素的问题,参考 HEVC 中率失真代价原理,设置包括失真代价 Q_{kd} 和时间代价 Q_{kt} 的目标评价体系 Q_k ($k = 64, 32, 16$)。 Q_{kd} 由真值和模型预测结果不相同的情况组成,用 $matrix_k[1][0]$ 与 $matrix_k[0][1]$ 之和表示。具体地,本文采用二

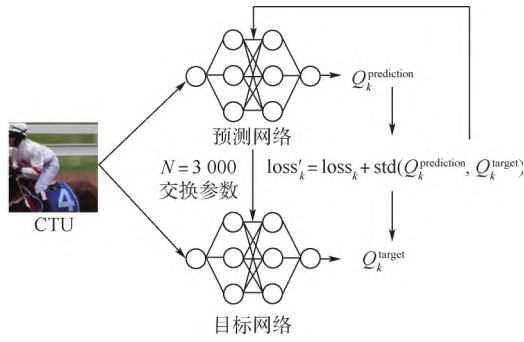


图3 对偶优化网络结构

Fig.3 Structure of dual optimal networks

维数组 $\text{matrix}_k[m][n]$ ($k=64, 32, 16; m=0, 1; n=0, 1$) 记录划分结果。其中 k 代表当前 CU 的大小, m 代表划分真值(值为 0 表示不划分, 1 表示划分), n 代表由预测模型得到的划分结果, 即 $\text{matrix}_k[m][n]$ 记录了 CU 大小为 k 时划分真值为 m 且预测为 n 的结果数量。 Q_{kt} 由预测模型决定划分的情况数之和表示, 即 $\text{matrix}_k[0][1]$ 与 $\text{matrix}_k[1][1]$ 之和。总代价 Q_k 由失真代价和时间代价和表示:

$$Q_k = \alpha Q_{kt} + \beta Q_{kd} \quad (1)$$

式中: α 和 β 分别为 Q_{kt} 和 Q_{kd} 的比例系数。一般而言, 失真代价对编码效果影响大于时间代价。

此外, 由于单一监督信息会导致模型决策朝某一固定的方向发展, 导致最终模型出现负优化或无法适应其他可能出现的情况。因此, 本文建立对偶网络结构平衡图像特征和率失真代价之间的关系, 将原有用于训练预测划分模型的网络视为预测网络, 并采用目标网络监督其训练过程。目标网络的结构与预测网络相同, 在训练迭代的过程中不实时更新参数。对于一组数据, 2 个网络同时对其划分方式输出预测模型, 当预测网络优于目标网络的评价值累计 N 次时, 将预测网络的参数赋给目标网络。该结构有效地对预测网络的训练实现了强监督操作, 避免因损失函数不合理而导致模型的负优化。

为了控制预测网络参数向对偶网络中具有较好效果的网络方向发展, 预测网络只在其评价结果 $Q_k^{\text{prediction}}$ 优于目标网络评价结果 Q_k^{target} 时拟合二者之间的损失。因此, 为了平衡率失真与图像失真, 本文设计的损失函数 loss'_k 由原有卷积损失 loss_k 与 2 个网络评价值的标准差 std 组成, 如下:

$$\text{loss}'_k = \text{loss}_k + \text{std}(Q_k^{\text{prediction}}, Q_k^{\text{target}}) \quad (2)$$

式中: std 表示求解 $Q_k^{\text{prediction}}$ 和 Q_k^{target} 标准差, $k=64, 32, 16$ 。

通常, 二分决策方法取 0.5 作为划分阈值。

DualNet 网络使用率失真影响了决策方向, 因此, DualNet 的划分阈值不能采用 0.5, 而是采用基于率失真的动态阈值策略来计算划分阈值。阈值调整公式如下:

$$Q = \text{matrix}_k[0][1] \cdot \gamma - \text{matrix}_k[1][0] \cdot \delta \quad (3)$$

式中: Q 为最终的划分阈值; γ 和 δ 为比例系数。 $\text{matrix}_k[0][1]$ 对模型预测结果影响更大。

3 实验

3.1 实验数据及设置

1) 实验数据。本文的训练数据采用文献[20]中的 CPH-Intra 数据集^[30], 包括 2 000 张无损图像。每个样本由 CU 的亮度矩阵和一个代表是否划分的二分类标签组成。测试数据使用 5 种不同类别的 JCT-VC^[26] 序列, 序列参数如表 1 所示。

表1 JCT-VC 测试序列参数

Table 1 JCT-VC test sequence parameters

类别	序列名称	分辨率	帧数	帧率/fps
A	Traffic	2 560 × 1 600	150	30
B	BasketballDrive	1 920 × 1 080	500	50
C	BasketballDrill	832 × 480	500	50
D	BasketballPass	416 × 240	500	50
E	Johnny	1 280 × 720	600	60

注: fps 为帧/s。

2) 评价指标。本文使用提案 VCEG-M33 中的 BD-BR 方法^[31], 包含 BD-BR、BD-PSNR 和 ΔT 三个评价指标。其中, BD-BR 表示在视频客观质量相同的情况下, 优化算法相比原始算法的比特率增量; BD-PSNR 表示在相同码率的情况下, 优化算法相比原始算法的视频客观质量提高值; ΔT 表示在编码相同帧数的情况下, 优化算法节省的时间, 计算如下:

$$\Delta T = \frac{T' - T}{T} \times 100\% \quad (4)$$

式中: T 和 T' 分别为 HEVC 标准和优化算法消耗的编码时间。

3) 参数设置。本文采用 Momentum 优化器, 学习率设置为 0.01。式(1)中的关键参数 β 设置为 0.005, α 设置为 0.01; 式(3)中的关键参数 γ 和 δ 分别设置为 0.001 和 0.0005。DualNet 网络训练使用一张 NVIDIA GeForce GTX 1080Ti 显卡, Tensorflow 版本为 1.14, 采用 HM16.5 帧内模式默认参数配置文件 encoder_intra_main.cfg 实现编码过程。

3.2 消融实验

为了评估对偶网络结构和动态阈值的性能,本文完成了4组实验:第1组实验(记为CNN)采用2.1节CNN训练模型,划分阈值设置为 $[0.5, 0.5, 0.5]$;第2组实验(记为Thr-CNN)采用带有动态阈值的CNN训练模型,训练后动态阈值结果为 $[0.49, 0.55, 0.63]$;第3组实验(记为DualNet-E₁)采用2.2节的对偶CNN训练模型,划分阈值设置为 $[0.5, 0.5, 0.5]$;第4组实验(记为DualNet-E₂)采用具有动态阈值的对偶CNN进行训练,训练后动态阈值结果为 $[0.48, 0.55, 0.63]$ 。训练所得模型经过4种QP(QP=22, 27, 32, 37)测试,计算每个序列的前20帧。测试结果使用JCT-VC提供的VCEG-AE07.xls^[29]计算BD-BR和BD-PSNR,消融实验结果如表2所示。

由于本文算法跳过了递归计算率失真损失的过程,以上算法相比HM16.5在大规模降低计算复杂度的情况下,均对视频解码重建质量产生了影响。表2中:BD-BR值均为正,代表HM16.5的编码效率优于上述算法,但其值越低表示码率越小,编码压缩效果越接近HM16.5;BD-PSNR值均为负,代表压缩后视频质量相比HM16.5均有损失,其值越高代表损失越小,视频质量越接近HM16.5。由表2可知,Thr-CNN仅使用动态划分阈值,没有目标网络监督训练,无法确定阈值变化方向,导致在部分数据集下测试效果不佳。DualNet-E₁使用对偶网络结构,预测网络参数模型受目标网络影响,仍然使用0.5作为划分阈值,限制了模型预测效果。采用动态阈值的对偶网络DualNet-E₂的平均BD-PSNR为-0.120 dB、平均BD-BR为2.876%、平均 ΔT 为-64.06%,性能优于CNN、Thr-CNN、DualNet-E₁三个网络,表明该网络不仅适用于不同分辨率的测试序列的CU划分过程,而且获得最佳编码质量的同时,大大节约了HM16.5帧内编码时间。

3.3 实验结果及分析

为了评估本文所提对偶网络对CNN处理CU划分的优化效果,与单神经网络结构相似但无率失真监督学习的PPMAC^[17]和ETH-CNN^[20]算法进行对比,实验结果如表3所示。

从表3可知,本文算法训练得到的模型对大部分数据集的压缩效果和效率优于对比算法,且具有较小的离散程度。由于本文算法将率失真代价作为评价方法,与对比算法中编码效果较优的ETH-CNN相比,BD-PSNR平均提升了0.008 dB, BD-BR降低了0.192%,提高了重建质量。动态

表2 消融实验结果(JCT-VC)

Table 2 Results of ablation study (JCT-VC)

类别	算法	BD-PSNR/dB	BD-BR/%	$\Delta T/\%$
A	CNN	-0.149	2.771	-63.19
	Thr-CNN	-0.133	2.480	-66.79
	DualNet-E ₁	-0.148	2.757	-66.01
	DualNet-E ₂	-0.131	2.429	-63.55
B	CNN	-0.119	4.981	-72.29
	Thr-CNN	-0.094	3.904	-77.10
	DualNet-E ₁	-0.120	4.967	-75.31
	DualNet-E ₂	-0.094	3.941	-74.29
C	CNN	-0.141	2.934	-43.77
	Thr-CNN	-0.134	2.796	-51.98
	DualNet-E ₁	-0.142	2.969	-50.03
	DualNet-E ₂	-0.130	2.738	-47.87
D	CNN	-0.138	2.412	-48.32
	Thr-CNN	-0.116	2.029	-50.53
	DualNet-E ₁	-0.135	2.359	-44.75
	DualNet-E ₂	-0.107	1.853	-57.06
E	CNN	-0.146	3.636	-75.04
	Thr-CNN	-0.136	3.355	-77.51
	DualNet-E ₁	-0.141	3.501	-78.48
	DualNet-E ₂	-0.138	3.421	-77.55
标准差	CNN	0.011	1.016	14.02
	Thr-CNN	0.018	0.735	13.08
	DualNet-E ₁	0.011	1.013	15.01
	DualNet-E ₂	0.019	0.821	12.23
平均值	CNN	-0.139	3.347	-60.52
	Thr-CNN	-0.123	2.913	-64.78
	DualNet-E ₁	-0.137	3.311	-62.92
	DualNet-E ₂	-0.120	2.876	-64.06

表3 编码单元划分对比实验结果

Table 3 Experimental results of CU partition

类别	算法	BD-PSNR/dB	BD-BR/%	$\Delta T/\%$
A	PPMAC	-0.240	4.945	-60.84
	ETH-CNN	-0.125	2.550	-61.01
	DualNet-E ₂	-0.131	2.429	-63.55
B	PPMAC	-0.141	6.018	-69.51
	ETH-CNN	-0.121	4.265	-76.32
	DualNet-E ₂	-0.094	3.941	-74.29
C	PPMAC	-0.538	12.205	-63.58
	ETH-CNN	-0.133	2.863	-52.98
	DualNet-E ₂	-0.130	2.738	-47.87
D	PPMAC	-0.457	8.401	-63.53
	ETH-CNN	-0.106	1.842	-56.42
	DualNet-E ₂	-0.107	1.853	-57.06
E	PPMAC	-0.307	7.956	-66.55
	ETH-CNN	-0.153	3.822	-70.68
	DualNet-E ₂	-0.138	3.421	-77.55
标准差	PPMAC	0.160	2.787	3.32
	ETH-CNN	0.017	0.977	9.78
	DualNet-E ₂	0.019	0.821	12.23
平均值	PPMAC	-0.337	7.905	-64.80
	ETH-CNN	-0.128	3.068	-63.48
	DualNet-E ₂	-0.120	2.876	-64.06

阈值使得 CU 划分模型能在各种情况下做出合理的判断。同时,本文算法比 ETH-CNN 算法平均节省了 0.58% 的编码时间,提升了 HEVC 编码效率。本文算法参考了 HEVC 标准中 CU 划分原则,并利用对偶网络实现监督学习,相比于不使用率失真监督信息和对偶网络的深度模型,算法模型对压缩效果和时间均有优化,验证了对偶 CU 划分网络 DualNet 的有效性。

4 结 论

本文提出了基于对偶网络结构的 CU 划分决策网络模型,分别构建用于决策的预测网络和用于率失真优化的目标网络,并且反馈优化预测网络对 CU 划分的阈值,提高模型的预测准确度。实验结果表明,本文算法可以提高视频编码的质量,降低了编码复杂度。

在后续的研究中,将在此基础上,进一步探索帧内预测模式对 CU 划分的影响,实现帧内模式和帧间模式 CU 划分的优化,以进一步提升 HEVC 的性能。

参考文献 (References)

- [1] LIU D, LI Y, LIN J, et al. Deep learning-based video coding: A review and a case study[J]. ACM Computing Surveys, 2020, 53(1): 1-35.
- [2] TUDOR P. MPEG-2 video compression[J]. Electronics & Communication Engineering Journal, 1995, 7(6): 257-264.
- [3] 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.
- [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] ZHANG Y, ZHAO Y, LIN C, et al. Block partitioning decision based on content complexity for future video coding[C]//International Conference on Image and Graphics. Berlin: Springer, 2019: 70-80.
- [6] GUO H, ZHU C, XU M, et al. Inter-block dependency-based CTU level rate control for HEVC[J]. IEEE Transactions on Broadcasting, 2019, 66(1): 113-126.
- [7] JAMALI M, COULOMBE S. Fast HEVC intra mode decision based on RDO cost prediction[J]. IEEE Transactions on Broadcasting, 2018, 65(1): 109-122.
- [8] ZHANG M, ZHAI X, LIU Z, et al. Fast algorithm for HEVC intra prediction based on adaptive mode decision and early termination of CU partition[C]//2018 Data Compression Conference. Piscataway: IEEE Press, 2018: 434-434.
- [9] LU J, LI Y. Fast algorithm for CU partitioning and mode selection in HEVC intra prediction[C]//2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. Piscataway: IEEE Press, 2019: 1-5.
- [10] WANG M, LI J, ZHANG L, et al. Extended quad-tree partitioning for future video coding[C]//2019 Data Compression Conference. Piscataway: IEEE Press, 2019: 300-309.
- [11] QING A, ZHOU W, WEI H, et al. A fast CU partitioning algorithm in HEVC inter prediction for HD/UHD video[C]//2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Piscataway: IEEE Press, 2016: 1-5.
- [12] KIBEYA H, BELGHITH F, AYED M A B, et al. A fast CU partitioning algorithm based on early detection of zero block quantified transform coefficients for HEVC standard[C]//International Image Processing, Applications and Systems Conference. Piscataway: IEEE Press, 2014: 1-5.
- [13] 朱蕾琦, 张其善, 杨东凯, 等. 改进的帧内帧间模式选择快速算法[J]. 北京航空航天大学学报, 2008, 34(12): 1411-1414.
ZHU L Q, ZHANG Q S, YANG D K, et al. Fast mode selection for intra and inter prediction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(12): 1411-1414 (in Chinese).
- [14] ZHANG D, DUAN X, ZANG D. Decision tree based fast CU partition for HEVC lossless compression of medical image sequences[C]//2017 9th International Conference on Wireless Communications and Signal Processing. Piscataway: IEEE Press, 2017: 1-6.
- [15] KIM H S, PARK R H. Fast CU partitioning algorithm for HEVC using an online-learning-based bayesian decision rule[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 130-138.
- [16] FU B, ZHANG Q, HU J. Fast prediction mode selection and CU partition for HEVC intra coding[J]. IET Image Processing, 2020, 14(9): 1892-1900.
- [17] LIU Z, YU X, GAO Y, et al. CU partition mode decision for HEVC hardwired intra encoder using convolution neural network[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5088-5103.
- [18] ZHANG Y, WANG G, TIAN R, et al. Texture-classification accelerated CNN scheme for fast intra CU partition in HEVC[C]//2019 Data Compression Conference. Piscataway: IEEE Press, 2019: 241-249.
- [19] 易清明, 林成思, 石敏. 利用深度学习的 HEVC 帧内编码单元快速划分算法[J]. 小型微型计算机系统, 2021, 42(2): 368-373.
YI Q M, LIN C S, SHI M. Fast HEVC coding units partitioning algorithm based on deep learning[J]. Journal of Computer Systems, 2021, 42(2): 368-373 (in Chinese).
- [20] XU M, LI T, WANG Z, et al. Reducing complexity of HEVC: A deep learning approach[J]. IEEE Transactions on Image Processing, 2018, 27(10): 5044-5059.
- [21] CHUNG C H, PENG W H, HU J H. HEVC/H.265 coding unit split decision using deep reinforcement learning[C]//2017 International Symposium on Intelligent Signal Processing and Communication Systems. Piscataway: IEEE Press, 2017: 570-

- 575.
- [22] BAE J H , YEO D , YIM J , et al. Densely distilled flow-based knowledge transfer in teacher-student framework for image classification[J]. IEEE Transactions on Image Processing ,2020 , 29: 5698-5710.
- [23] ABBASI S ,HAJABDOLLAHI M ,KARIMI N ,et al. Modeling teacher-student techniques in deep neural networks for knowledge distillation[C] // 2020 International Conference on Machine Vision and Image Processing. Piscataway: IEEE Press , 2020: 1-6.
- [24] XIAO R ,LIU Z ,WU B. Teacher-student competition for unsupervised domain adaptation[C]//2020 25th International Conference on Pattern Recognition. Piscataway: IEEE Press ,2021: 8291-8298.
- [25] LU Y ,LI W ,NING X ,et al. Image quality assessment based on dual domains fusion [C] // 2020 International Conference on High Performance Big Data and Intelligent Systems. Piscataway: IEEE Press ,2020: 1-6.
- [26] ZHOU B ,ZHOU S K. DuDoRNet: Learning a dual-domain recurrent network for fast MRI reconstruction with deep T1 prior[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press ,2020: 4273-4282.
- [27] WANG H ,TIAN Q ,LI L ,et al. Image demoiréing with a dual-domain distilling network [C] //2021 IEEE International Conference on Multimedia and Exposition. Piscataway: IEEE Press , 2021: 1-6.
- [28] 苏志雄 ,李星梅 ,乞建勋. 网络计划中构建对偶网络模型的理论和方法 [J]. 北京航空航天大学学报 ,2012 ,38 (2) : 257-262.
- SU Z X ,LI X M ,QI J X. Theory and method of creating dual network model in network planning[J]. Journal of Beijing University of Aeronautics and Astronautics ,2012 ,38 (2) : 257-262 (in Chinese) .
- [29] HM software [CP/OL]. [2021-08-28]. <https://hevc.hhi.fraunhofer.de/svn/svnHEVCSoftware/tags/HM-16.5/>.
- [30] CPH-Intra[DS/OL]. [2021-08-28]. <https://github.com/Projects/CPH>.
- [31] GRELLERT M ,BAMPI S ,CORREA G ,et al. Learning-based complexity reduction and scaling for HEVC encoders [C] // 2018 IEEE International Conference on Acoustics ,Speech and Signal Processing. Piscataway: IEEE Press ,2018: 1208-1212.

Dual coding unit partition optimization algorithm of HEVC

LIU Meiqin^{1,2} , XU Chenming^{1,2} , YAO Chao^{3,*} , LIN Chunyu^{1,2} , ZHAO Yao^{1,2}

(1. Institute of Information Science , Beijing Jiaotong University , Beijing 100044 , China;

2. Beijing Key Laboratory of Modern Information Science and Network Technology , Beijing 100044 , China;

3. School of Computer and Communication Engineering , University of Science and Technology Beijing , Beijing 100083 , China)

Abstract: To resolve the conflict between the increasing amount of video data and the demand for high-quality video experience , HEVC has boosted the compression performance by 50% based on dramatically increased complexity of H.264/AVC. In this paper , a fast coding unit (CU) partition algorithm is proposed to reduce the computational complexity of HEVC intra coding. To define the partition criteria , we design a convolutional neural network , named dual neural networks (DualNet) . DualNet consists of two subnetworks , a prediction network and a target network. The prediction network is used to determine the partition actions by extracting images statistical features for skipping the traversal search of quadtree and improving the time efficiency of the CU partition. And the target network is to optimize the performance of the CU partition based on rate-distortion for achieving model complementarity. Experimental results show that the proposed algorithms can save 64.06% of the compression time with similar compression performance to HEVC.

Keywords: video coding; H.265/HEVC; coding unit (CU) partition; deep learning; dual neural networks

Received: 2021-09-06; **Accepted:** 2021-09-17; **Published online:** 2021-10-11 16:35

URL: kns.cnki.net/kcms/detail/11.2625.V.20211011.1531.001.html

Foundation items: National Natural Science Foundation of China (61972028 , 61902022 , 62120106009) ; the Fundamental Research Funds for the Central Universities (2019JBM018 , FRF-TP-19-015A1)

* **Corresponding author.** E-mail: yaochao@ustb.edu.cn