

AVS3 视频编码 关键技术及应用

Key Technologies and Applications of AVS3 Video Coding Standard

张嘉琪 /ZHANG Jiaqi¹, 雷萌 /LEI Meng², 马思伟 /MA Siwei^{2,3}

(1. 中国科学院计算技术研究所, 中国 北京 100086;

2. 北京大学, 中国 北京 100871;

3. 北京大学信息技术科创中心, 中国 绍兴 312300)

(1. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100086, China;

2. Peking University, Beijing 100871, China;

3. Information Technology R&D Innovation Center of Peking University, Shaoxing 312300, China)



摘要: 超高清 (UHD) 视频能为用户带来质量更高、沉浸感更强的视觉体验, 但高带宽成本限制了其推广和应用。为解决超高清视频传输和存储的难题, 中国数字音视频编解码技术标准 (AVS) 工作组制定了新一代视频编码标准——AVS3, 并在超高清产业化应用方面取得重要进展。介绍了 AVS3 视频编码关键技术, 以及其与 AVS2、多功能视频编码 (VVC)、开放媒体联盟视频 (AV1) 等标准的性能对比情况。

关键词: 视频编码; AVS3; 超高清

Abstract: Ultra-high definition (UHD) videos can provide users a higher quality and more immersive visual experiences. However, the application of UHD is limited by high bandwidth cost. To solve the transmission and storage problem of UHD, China Audio and Video Coding Standard (AVS) workgroup established a new generation of video coding standard—AVS3. Currently, AVS3 has made a great contribution to the development of UHD industries in China. Key technologies of AVS3 are described, and a comprehensive comparison with AVS2, versatile video coding (VVC) and alliance for open media video 1 (AV1) is conducted.

Keywords: video coding; AVS3; UHD

视觉是人类获取信息的重要来源, 视频承载了海量非结构化视觉信息, 是应用最广泛的多媒体数据格式, 它与人们的生活息息相关, 是人类获取信息的重要途径之一。目前, 互联网 70% 以上的流量来自于图片和视频, 并且这个比例仍在持续攀升^[1], 视频已成为网络上体量最大的数据格式。据统计^[1], 2017 年标清和高清视频内容约各占视频流量的一半; 2019 年标清内容的占比约下降到 1/3, 高清内容

成为主流, 而超高清内容的占比正在逐步攀升; 预计到 2022 年, 超高清内容的占比约提升到 1/4。

超高清视频具有更高的空间和时间分辨率、更广的色域和更宽的动态范围, 是继视频数字化、高清化之后的新一轮重大技术革新。视频技术从高清向超高清的演进, 不仅引发了内容制播、芯片制造、网络传输等产业链各环节的升级换代, 而且驱动了广播电视、安防监控、智能交通等以视频为核心的行业服务转型。自 2018 年起, 中国超高清视频产业已达万亿元级别。预计到 2022 年, 中国超高清视

频产业总体规模将超过 4 万亿元^[2]。

成倍增长的数据量给超高清视频的高效传输和存储带来了巨大的挑战。以 8K、10 bit、120 帧/秒的 YUV (一种颜色编码方法) 420 格式的超高清视频为例, 其原始数据的码率会达到约 88.99 Gbit/s。若采用第 2 代数字音视频编解码技术标准 (AVS2) / 高效视频编码 (HEVC)^[3-4] 标准对原始数据进行压缩, 压缩码率约 310 Mbit/s, 带宽传输压力极大。因此, 超高清视频应用迫切需要更加高效的压缩技术。

为解决超高清视频带宽需求大、存储难等问题, 中国 AVS 工作组率先

基金项目: 国家自然科学基金 (61961130392); 北京大学百度基金 (2019BD003)

展开了具有自主知识产权的、针对超高清视频的视频编码标准的制定工作。在 2017 年 12 月举行的会议中, AVS 工作组决定开展面向超高清视频应用的新一代数字视频编码标准(以下简称 AVS3)的制定工作。AVS3 的制定工作分为两个阶段:第 1 阶段(基准档次)是从 2018 年 3 月到 2019 年 6 月,制定面向复杂度优先的应用,其性能相较于 AVS2 提升 30%;第 2 阶段(增强档次)是从 2019 年 6 月到 2021 年 12 月,目标是编码效率比 AVS2 提升 1 倍以上,同时编码性能超越同时代的其他国际标准。2020 年 5 月 13 日, AVS3 基准档次标准正式获批并被颁布为团体标准。

1 AVS3 视频编码关键技术

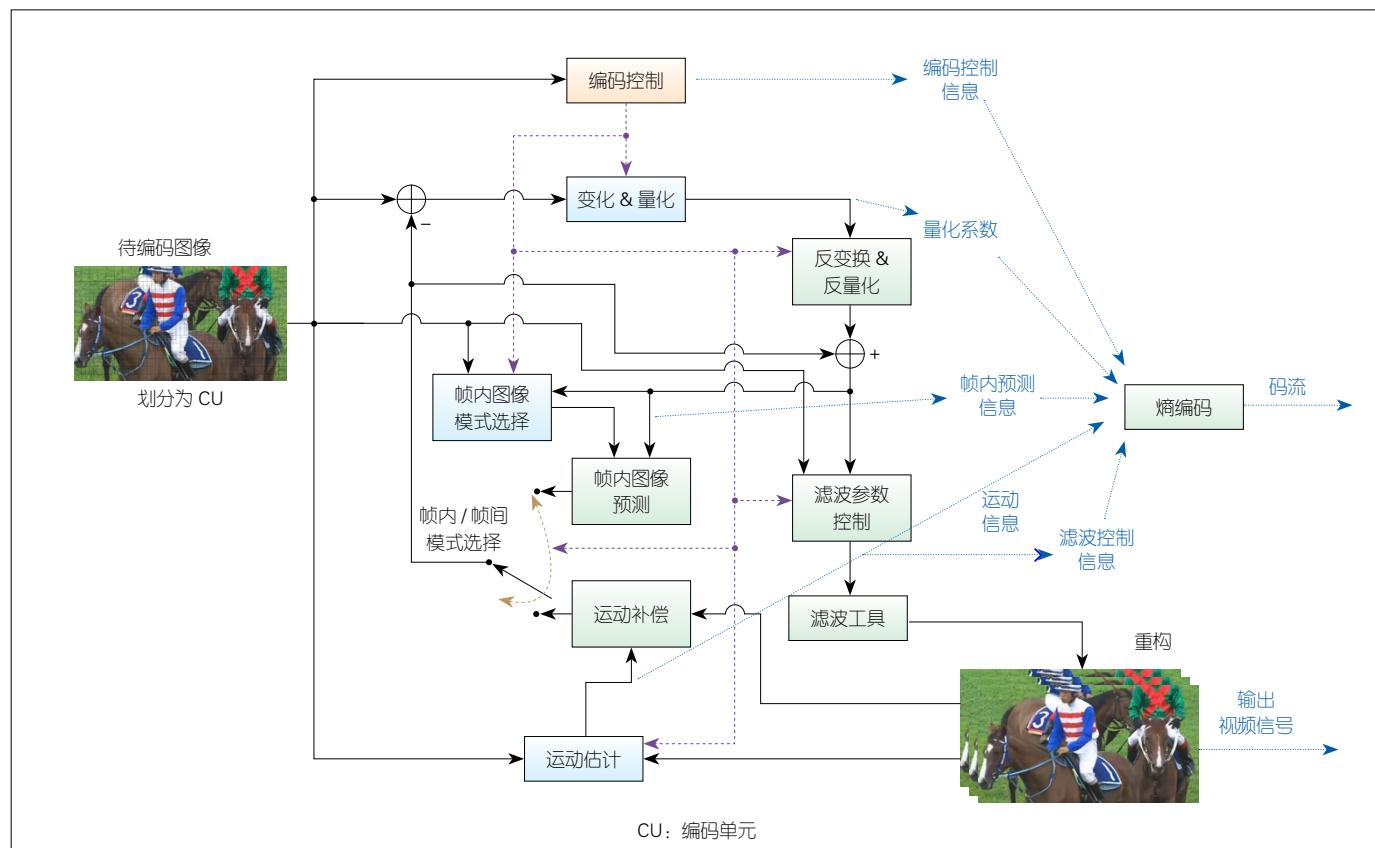
AVS 工作组自 2002 年成立以来,

一直致力于制定高压缩率和友好专利政策的视频编码标准。经历了 19 年的发展, AVS 工作组已经制定从 AVS1 到 AVS3 这 3 代视频编码标准。面向超高清视频应用, AVS3 沿用了基于块的预测变换混合编码框架,具体如图 1 所示。AVS3 包括块划分、帧内预测、帧间预测、变换量化、熵编码、环路滤波等模块。相较于 AVS2, AVS3 在保留部分编码工具的同时,针对不同模块引入了一些新的编码工具^[5],并采用了更灵活的块划分结构、更精细的预测模式、更具适应性的变换核,实现了约 30% 的码率节省,显著提升了编码效率。

1.1 块划分

如图 2(a) 所示, AVS2 采用了基于四叉树(QT)的递归划分编码框

架,每个编码单元(CU)的尺寸都是方形且允许被进一步划分为不同形状的预测单元(PU)。为提升划分的灵活性, AVS3 引入了基于四叉、二叉(QTBT)和扩展四叉树(EQT)的划分方式,如图 2(b)。QTBT 加 EQT 的划分方式允许出现非方形编码单元,编码单元是后续预测、变换和量化的基础,非方形划分更加符合纹理精细和为了便于硬件实现, AVS3 采用了局部分离树(LST)。LST 技术为了避免色度出现边长等于 2 像素的变换块,在亮度块划分时,如果亮度块出现边长等于 4 像素的边,则仅对亮度块划分,无须对色度块划分。为提高硬件流水处理效率, AVS3 对一些小块添加了模式限制。当块大小满足限制后,该节点及其划分得到的编码块的编码模式只能全部选择同一种预测模式,



▲图 1 基于块的混合编码框架

如帧间预测或帧内预测。

1.2 帧间预测

帧间预测工具可以分为3类：一类是针对跳过模式和直接模式候选项的扩充，一类是差分运动矢量（MVD）编码，最后一类则是基于子块的运动补偿。

跳过模式和直接模式是一项使用相邻编码块的运动矢量（MV）进行预测的高效编码技术。AVS2中的跳过模式和直接模式候选项只有4个相邻模式和1个时域模式，对图像非相邻结构性和纹理多变性的区域编码效率不高。AVS3引入了基于历史运动矢量的预测（HMVP）和高级运动矢量表达（UMVE）等技术。HMVP利用非局部

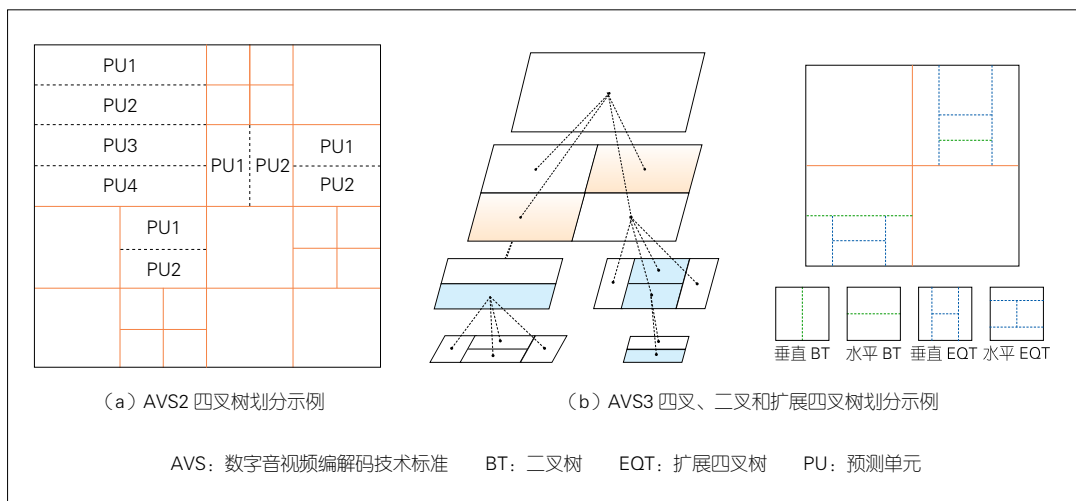
相似性的原理获取更多非相邻的运动矢量候选，如图3（a）所示。HMVP通过动态更新运动候选矢量列表，保留了与当前块运动相关性最高的候选项，提高了跳过模式和直接模式、处理非局部相似性运动的能力。UMVE通过对跳过模式和直接模式候选项加入运动矢量偏移，对运动矢量进行更精细的表达，可以更好地消除视频场景中因剧烈运动而带来的匹配误差。

自适应运动矢量精度（AMVR）和扩展运动矢量精度（EMVR）的引入提升了MVD的编码效率。在AVS2中，运动矢量精度只有1/4像素和1/2像素，且无法灵活选择。AVS3中的AMVR使用了1/4、1/2、1、2、4像素精度的运动矢量，根据视频内容自适应地选

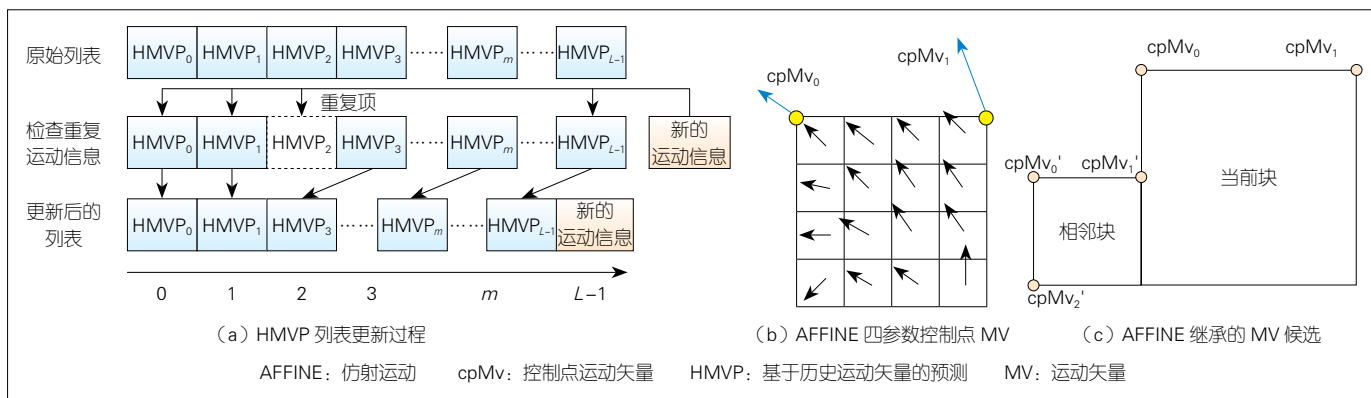
择预测精度，提高了帧间预测在不同区域的适应性。EMVR提供了不同的运动搜索起始点，扩大了运动矢量的搜索空间，有效提升了运动估计的准确性。

双向光流（BIO）^[6]、仿射运动（AFFINE）和解码端运动矢量修正（DMVR）^[7]等技术采用基于子块的运动补偿，提高了帧间预测准确度。基于物体运动轨迹是平滑的这一假设，BIO通过最小化每个子块的前向和后向预测样本之间的差异来计算运动细化差，然后使用运动细化差来调整每个子块的预测样本值。如图3（b）所示，AFFINE根据仿射变换模型，利用2个（四参数）或3个（六参数）控制点的运动矢量导出当前编码块的运动矢

量场。AFFINE运动模型相对于AVS2中的平移运动模型，可以有效提升具有缩放、旋转、透视和其他不规则运动等性能的视频序列编码。DMVR将编码区域划分为若干个不重叠的子块，以初始MV为起始位置，使用最小化均方误差的模板匹配方法对当前MV进行偏移，进一步修正双向



▲图2 AVS2、AVS3 块划分示例



▲图3 HMVP 和 AFFINE 示意图

预测样本值。

1.3 帧内预测

帧内预测方面的新技术包括帧内预测模式扩展(EIPM)、预测像素滤波、跨分量预测等。

EIPM^[8]扩展了帧内预测的角度,如图4(a)所示。帧内预测模式从33种扩展到66种,包括62种角度模式和4种特殊模式,提高了对方向性纹理的预测能力,可以适应纹理丰富的超高清视频内容。

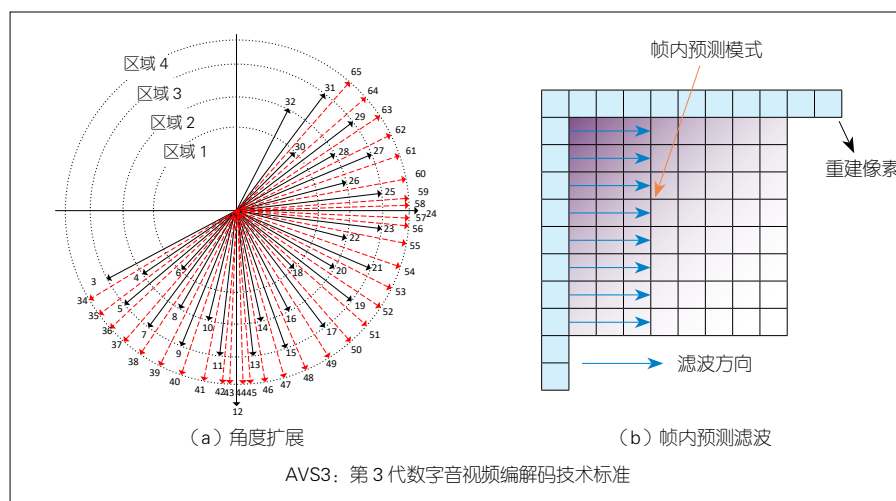
帧内预测滤波包含分像素插值滤波和预测像素值滤波。多组滤波(MIPF)根据块内像素点的个数和所在位置^[9],使用4组不同的插值滤波器生成预测像素。多组滤波适用于不同的颜色分量和像素平滑程度,在复杂度极低的情况下,取得了可观的性能增益。MIPF得到预测像素后,还可以对预测像素进行帧内预测滤波(IPF)。IPF使用高斯平滑滤波器,根据参考像素、预测模式和与参考像素的距离对预测像素做进一步的修正,如图4(b)所示。跨分量预测是指在色度预测编码过程中,通过两步预测模式(TSCPM)对色度进行预测编码。其原理是假定亮

度和色度分量之间线性相关,通过最小二乘法求解对应线性回归的参数,在求得参数后,使用亮度重构像素以精细重建对应位置的色度像素,在色度上取得了显著的增益。

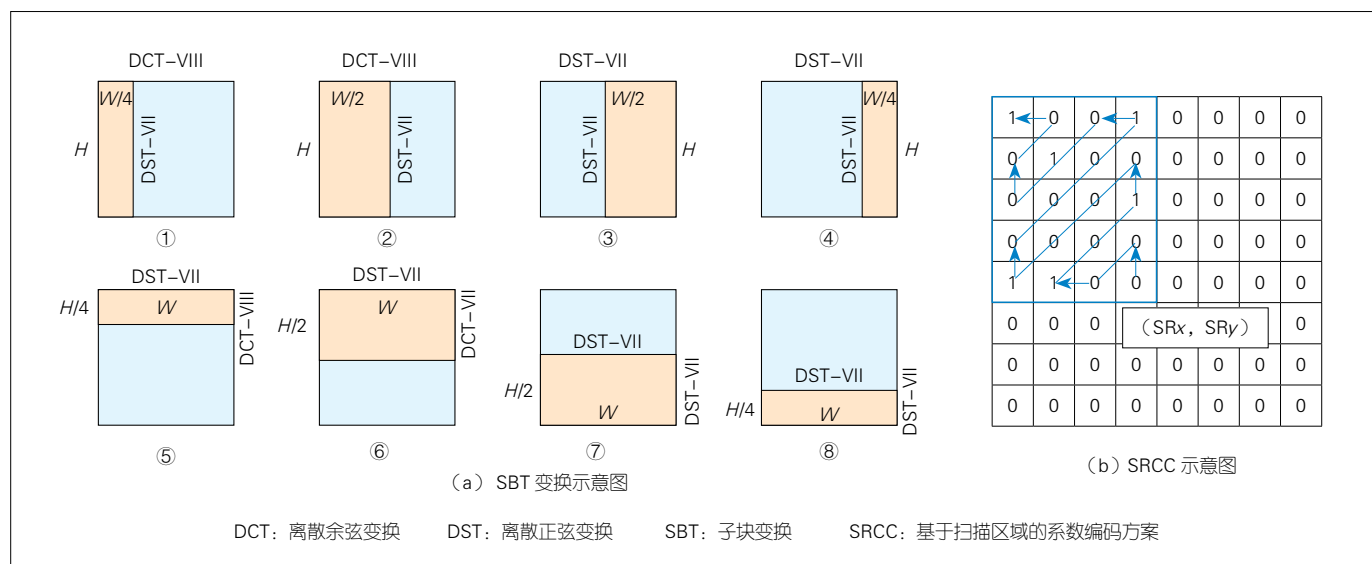
1.4 变换与量化

变换可以集中能量,利于熵编码进行系数压缩。离散余弦变换(DCT)具有很好的去相关能力^[10],且由于其对称性有利于软硬件实现,因此能够在视频压缩领域得到广泛的应用。在上二代视频编码标准中,DCT-II作为

主要应用的变换核,适用于均匀分布的残差变换,但缺乏处理不均匀残差分布的能力。在AVS3中,隐则变换(IST)和子块变换(SBT)引入了新的变换核DST-VII和DCT-VIII,能够聚集不均匀分布残差的能量。IST^[11]通过量化块中偶数系数个数的奇偶性隐式地导出变换核的类型,在提高变换灵活性的同时,没有引入额外的比特消耗。基于帧间预测残差分布的局部性,SBT把预测残差分布的位置限制在残差块的1/2或者1/4区域,如图5(a)所示,从而降低变换系数的局部分量,



▲图4 AVS3 帧内编码工具



▲图5 变换与系数编码

并减少了全零块的编码代价,提高了压缩性能。

在系数编码中,AVS3采用了一种基于扫描区域的系数编码方案(SRCC)^[12]。SRCC使用参数(SR_x , SR_y)控制量化系数非零的区域。为了达到码率和失真之间的平衡以及提高系数编码的灵活性,SRCC使用率失真优化选择最优扫描区域。在扫描编码区域内的非零系数时,SRCC采取了从右下到左上的反Z形扫描方式,如图5(b)所示;非零系数采用了分层编码,不同层级使用多套上下文,根据系数在扫描区域的位置和扫描区域的面积确定上下文模型。精确的上下文建模显著提升了压缩效率。

1.5 基于卷积神经网络的环路滤波

为了探索神经网络在编码标准中的可实现性,AVS3工作组设立了智能编码专题小组,对基于卷积神经网络的环路滤波(CNNLF)^[13]进行了深入研究。CNNLF能够代替传统的去块(Deblock)滤波和样本自适应偏移(SAO)滤波,并取得了6%左右的性能增益。

CNNLF使用神经网络探索视频信号之间的非线性关系和变化规律,对视频信号的全局信息和局部关系进行了联合建模。得益于海量的训练数据和算力的提升,CNNLF的网络泛化能力要远高于传统滤波方式。CNNLF训练时以残差块为单位,加速了网络收敛过程,并且设置不同量化参数(QP)段为训练单元,

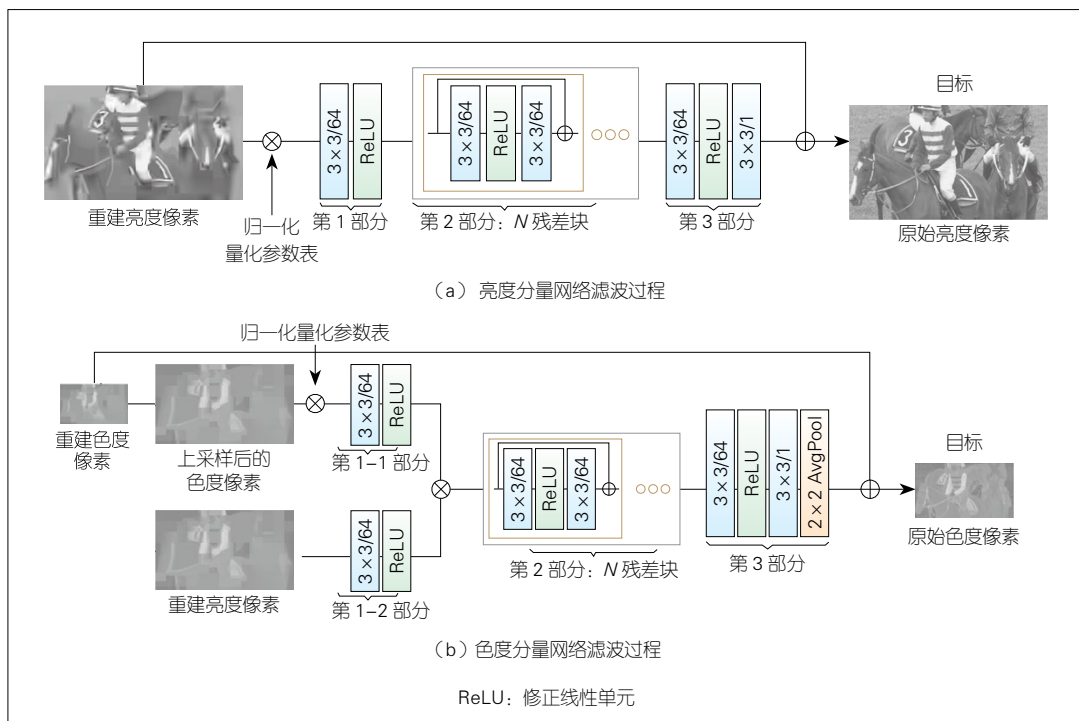
增强了网络对QP的泛化能力。如图6(a)所示,CNNLF的网络由全局残差、残差块、卷积层和激活层组成,采用亮度、色度分量分离训练的方式,且亮度分量指导色度分量滤波,进一步提升色度分量的重建质量,如图6(b)所示。

1.6 性能对比

本文中,我们首先对AVS3与AVS2进行了性能对比,测试时使用的参考软件版本分别为参考设计模型19.5(RD 19.5)、高性能平台4.0(HPM 4.0)和高性能平台9.0(HPM 9.0),其中HPM 4.0和HPM 9.0分别用于测试AVS3第1阶段和第2阶段。测试配置为随机访问(RA)配置,测试结果见表1。可以看出AVS3第1阶段相比AVS2平均可以获得24%的性能提升,且对4K分辨率序列的提升更为明显,达到了平均25%的性能提升。AVS3第2阶段是基于第1阶段的进一步推进。相比第1阶段,AVS3第

2阶段实现了约8%的性能提升;相比AVS2,实现了平均31%的性能提升,同时各分辨率序列的性能提升较为均衡,在部分4K序列上可以达到超40%的性能提升。此外,我们还测试了AVS3采用了CNNLF后的性能,如表1最右侧所示可以再获得近3%的性能提升。

我们还将AVS3和VVC^[14]、开放媒体联盟视频标准(AV1)^[15]进行了对比。我们选取了5个2K序列及6个4K序列进行测试,测试平台分别为通用编码测试平台(VTM 10.0)、开放媒体联盟视频标准测试平台(AOM-Oct)、HPM 4.0和HPM 9.0。如表2所示,VVC、AVS3和AV1相较于HEVC,在客观性能上都有较大的提升,尤其是VVC的性能提升最为显著,平均达到了40%;其他各标准中,AV1平均提升了25.5%的性能,对于AVS3,第1阶段和第2阶段分别达到了平均23%和30%的性能提升。综合来看,VVC、AV1和AVS3在超高清序列方



▲图6 基于卷积神经网络的环路滤波过程

面都表现出了优异的性能,达到了平均 25% 及以上的性能提升。VVC 和 AVS3 更是达到了超 30% 的性能提升,个别序列能达到 40% 的性能提升。

2 AVS3 超高清产业应用

随着超高清、全景视频等应用的高速发展,8K 超高清,乃至 16K、32K 等更高分辨率的视频内容将进一步流行。2019 年,中国发布的《超高清视频产业行动计划(2019—2022)》明确指出超高清视频将成为未来视频产业的重要发展方向。

AVS3 标准的颁布显著加速了超高清产业链的升级革新。为了缩短标准制定和成果落地的时间,AVS3 工作组在标准制定过程中,采用了分档制定与芯片集成技术协同研发的推进方式,同步推进全产业链应用开源合作。2019 年 6 月,AVS3 第 1 阶段基准档完成;2019 年 9 月,在阿姆斯特丹举办的第五十届荷兰广播电视设备展览会上,海思发布了首个基于 AVS3 标准的 8K 端到端解决方案,同时推出了全球首颗基于 AVS3 标准的支持 8K 分辨率、120 帧的超高清解码芯片 Hi3796CV300,如图 7(a)所示;随后,北京大学、北京博雅睿视科技有限公司和英特尔合作推出了 SVT-AVS3 8K 实时编码器,并搭建了 8K 端到端实时编解码系统,如图 7(b)所示。北京大学深圳研究生院开发了支持 AVS3 标准,8K 分辨率、60 帧实时解码器 uAVS3d。2020 年 5 月,当虹科技 AVS3 8K 超高清编码器和上海海思 AVS3 8K 超高清解码板完成了 AVS3+5G+8K 全国直播首测,主要测试在 5G 链路下的 8K 超高清节目直播传输应用。近期,中央广播电视总台启动“5G+4K/8K 超高清制播示范平台”项目,其中包括搭建 AVS2/AVS3 标准超高清电影院直播系统以及 5G 和超

▼表 1 AVS3 与 AVS2 在通用测试条件下的性能对比

序列	HPM 4.0			HPM 9.0			HPM 9.0-Mod AI 4.0		
	Y/%	U/%	V/%	Y/%	U/%	V/%	Y/%	U/%	V/%
通甲 4K	-23.9	-27.5	-29.6	-32.5	-36.8	-37.8	-36.0	-35.3	-38.8
通乙 1 080P	-23.7	-28.0	-29.6	-32.6	-35.5	-36.4	-35.3	-37.1	-34.9
通丙 720P	-23.0	-27.2	-26.5	-28.4	-30.2	-28.1	-32.7	-38.9	-40.5

AVS: 数字音视频编解码技术标准 HPM: 高性能平台 U、V: 统称为色度分量 Y: 亮度分量

注: 测试配置为随机访问, 均与设计模型 RD19.5 对比。

▼表 2 VVC、AV1、AVS3 和 HEVC 的性能对比

测试软件	随机访问编码配置		
	Y/%	U/%	V/%
VTM10.0	-38.17	-39.81	-39.71
AOM-Oct	-23.09	-33.69	-31.84
HPM 4.0	-23.20	-19.19	-18.66
HPM 9.0	-31.16	-27.87	-26.73

AOM-Oct: 开放媒体联盟视频标准测试平台

AV1: 开放媒体联盟视频标准

AVS3: 第 3 代数字音视频编解码技术标准

HEVC: 高效视频编码

HPM: 高性能平台

U、V: 统称为色度分量

VTM: 通用编码测试平台

VVC: 多功能视频编码

Y: 亮度分量



▲图 7 基于 AVS3 的 8K 应用实例

高清相关的测试体系。中央广播电视总台会将 AVS3 8K 超高清现场直播运用在 2022 年北京冬季奥运会中。

3 结束语

本文简要介绍了新一代视频编码标准 AVS3 的关键技术和 AVS3 超高清应用情况。与 AVS2 视频编码标准相比, AVS3 编码效率显著提升。AVS3 标准在技术创新、专利政策与生态建设方面已有全面的布局,为中国 8K 超高清

视频产业的发展奠定了坚实的基础。可以预见的是,随着 5G 的快速发展和超高清时代的来临, AVS3 标准前景广阔,将获得更广泛的应用。

参考文献

- [1] FORECAST G. Cisco visual networking index: global mobile data traffic forecast update, 2017–2022 [R]. 2017
- [2] 工信部等三部门联合印发《超高清视频产业发展行动计划(2019–2022 年)》[EB/OL]. [2020–12–22]. http://www.gov.cn/gongbao/content/2019/content_5419224.htm

- [3] MA S W, HUANG T J, READER C, et al. AVS2 ? making video coding smarter [standards in a nutshell] [J]. IEEE signal processing magazine, 2015, 32(2): 172–183. DOI:10.1109/msp.2014.2371951
- [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. DOI:10.1109/tcsvt.2012.2221191
- [5] ZHANG J Q, JIA C M, LEI M, et al. Recent development of AVS video coding standard: AVS3 [C]//2019 Picture Coding Symposium (PCS). Ningbo, China: IEEE, 2019: 1–5. DOI:10.1109/pcs48520.2019.8954503
- [6] 王凡, 欧阳晓, 吕卓逸, 等. CE: BIO 双向光流 [C]// 数字视频编解码技术标准化工作组第六十九次会议. 成都, 中国: AVS 工作组, 2019
- [7] 徐巍伟, 赵寅, 杨海涛. CE3.1: 简化 DMVR 方案 [C]// 数字视频编解码技术标准化工作组第七十次会议. 海口, 中国: AVS 工作组, 2019
- [8] 雷萌, 罗法蕾, 王苕社, 等. CE2-related: 帧内角度模式扩展 [C]// 数字视频编解码技术标准化工作组第七十次会议. 海口, 中国: AVS 工作组, 2019
- [9] 王英彬, 许晓中, 李一鸣, 等. CE1-related: 帧内预测参考像素滤波设计方法 [C]// 数字视频编解码技术标准化工作组第七十次会议. 海口, 中国: AVS 工作组, 2019
- [10] NUSSBAUMER H J. The fast Fourier transform [M]//Fast fourier transform and convolution algorithms. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981: 80–111
- [11] 张玉槐, 张凯, 张莉, 等. 帧内自适应变换 [C]// 数字视频编解码技术标准化工作组第六十九次会议. 成都, 中国: AVS 工作组, 2019
- [12] 王凡, 欧阳晓, 吕卓逸, 等. SRCC 基于扫描区域的系数编码 [C]// 数字视频编解码技术标准化工作组第七十次会议. 海口, 中国: AVS 工作组, 2019
- [13] 林凯, 贾川民, 赵政辉, 等. CE: 基于残差网络的神经网络滤波 [C]// 数字视频编解码技术标准化工作组视频组 2020 年 1 月加会. 北京, 中国: AVS 工作组, 2020
- [14] BROSS B, CHEN J, LIU S, et al. Versatile video coding (Draft 10): ITU-T and ISO/IEC JVET-S2001 [S]. 2020
- [15] CHEN Y, MURHERJEE D, HAN J N, et al. An overview of core coding tools in the AV1 video codec [C]//2018 Picture Coding Symposium (PCS). San Francisco, CA, USA: IEEE, 2018: 41–45. DOI:10.1109/pcs.2018.8456249

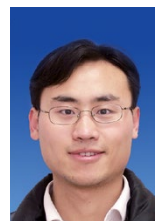
作者简介



张嘉琪, 中国科学院计算技术研究所读博士研究生; 主要研究方向为视频编码及处理; 所提多项标准提案已被 AVS3、VVC 等标准采纳。



雷萌, 北京大学信息科学技术学院在读博士研究生; 主要研究方向为视频编码及处理; 发表论文 2 篇。



马思伟, 北京大学信息科学技术学院教授、博士生导师, 国家杰出青年科学基金获得者, 现担任 AVS 视频组组长; 主要研究方向为视频编码及处理; 曾主持“863”计划、科技支撑计划、国家自然科学基金重点项目等多项国家级课题; 曾获国家技术发明奖二等奖、国家科学技术进步奖二等奖、中国电子学会特等奖等奖励; 已发表 SCI 论文 70 余篇, 已获得授权发明专利 50 余项。