

摘要

现如今，视频逐渐成为不可或缺的信息载体。视频质量的飞跃式提升、数据量的爆炸式增长使视频存储、传输面临着严峻的挑战，亟需研究更高效的视频编码方法。近年来，以数据驱动为核心的深度学习技术飞速发展，为视频处理相关领域带来了深刻变革。面向视频压缩效率提升的难题，基于深度学习的视频编码方法不断涌现，展现了巨大的性能优势。然而现有方法面临着规则固化的问题，在模式选择、运动表示等方面仍然存在着局限性。在此背景下，本文充分发挥边信息在视频编码中信息传递的作用，利用模式决策、运动特征等边信息来提升深度学习视频编码方法的性能，从而实现更高的视频压缩效率。围绕着边信息的生成、压缩与融合，本文提出基于边信息的深度学习视频编码方法。从信号重建增强和信号冗余去除两方面展开研究，主要创新点包括以下三个方面：

第一，针对深度学习环路滤波规则固化的问题，提出了基于自适应率失真决策的多模型环路滤波，引入模式决策信息提升重构质量。考虑到不同颜色分量之间的相关性，融合残差学习策略与跨颜色分量优化方法，提出高效的嵌套残差环路滤波网络。针对不同的编码场景，采用集成式学习策略，根据量化步长、预测方法等指标训练多组模型，提升深度学习环路滤波的泛化能力。针对帧级、块级两阶段率失真决策的不足，提出自适应的率失真决策方法，以模式决策信息作为边信息来提升网络模型的适用度。为了实现模式决策信息的紧凑表示，基于决策粒度的概率实现可变长编码。在时间复杂度层面，提出并行推理框架加速网络模型推理，根据帧间预测的规律设计快速算法。所提方法从网络模型设计、率失真决策方法、推理加速多方面入手提升深度学习滤波方法的有效性、可行性，可以有效抑制块效应、振铃效应等压缩噪声，提升重构序列的主客观质量。所提方法被 AVS 标准采纳，贡献了超过 8% 的编码性能增益。

第二，针对单向帧间预测场景中运动表示冗余的问题，提出了基于运动分解的渐进式帧间预测编码，引入运动修正信息实现高效帧间压缩。结合运动差分编码思想，将连续的运动信息分解为内在运动信息和运动修正信息两部分。利用循环神经网络感知和提取过去一段时间内时空域的变化规律，获取时域上下文作为内在运动信息进行时域外插，实现第一步时域转换。为了进一步提升帧间预测质量，以运动修正信息作为边信息对第一步时域转换进行增强，建立从粗到精的渐进式帧间预测方法。针对运动修正信息的高效编码，采用一阶段条件编码范式，结合非线性变换、条件上下文熵模型进行紧凑表示。所提方法可以实现高效帧间预测，提升全神经网络视频编码方法压缩效率，被 MPAA-EEV 标准采纳，在 MS-SSIM 指标相比于 VVC 可以实现约 7% 的编

码性能增益。

第三，针对单一时域假设在复杂运动场景中的不足，提出了基于特征域多假设融合的帧间预测编码，引入特征索引信息提升帧间预测有效性。通过模板注意力机制建模多假设预测，以特征索引信息为边信息，按照索引匹配、相关性排序、加权求和的方式实现特征域的多假设预测。从时域、空域两个角度提供多个内容相似的预测假设，通过特征相似度隐式导出多假设系数，提升帧间预测准确性。为了实现特征索引信息的高效压缩，采用一阶段无监督学习框架，以高度相关的键值特征作为时域先验，结合空域先验、超先验上下文提高熵模型的概率分布建模能力。本文所提方法可以灵活应用于单向、双向帧间预测场景，提升帧间压缩效率，具有很强的扩展性和自适应性。所提方法在多个全神经网络视频编码方法上验证，相比于单一时域假设，可以实现超过 15% 的编码性能增益。

综上所述，本文提出了基于边信息的深度学习视频编码方法，按照从局部率失真优化到联合率失真优化的研究思路，提出了基于自适应率失真决策的多模型环路滤波，基于运动分解的渐进式帧间预测编码，基于特征域多假设融合的帧间预测编码。本文所提方法在新一代视频编码方法的基础上进一步提升了压缩效率，并且部分获得标准化应用，本文为深度学习技术与传统编码方法的结合提供了新思路、新方法。

关键词：视频编码，环路滤波，帧间预测，边信息，深度学习