

# 基于学习的双向视频压缩方法及系统

## 信息和通信

计算机/人工智能/数据处理和信息技术  
数字广播、电信和光电

## 机会

随着视频内容在流媒体平台、视频会议和存储系统中的消费快速增长，对更高压缩效率的需求日益迫切，以降低带宽和存储成本。传统的视频编解码器（如HEVC和VVC）依赖于手工设计的基于块的算法，可能无法充分利用视频序列中复杂的时空冗余，尤其是在双向（B帧）编码方面。现有的基于学习的双向视频压缩方法通常假设运动均匀分布，并使用固定的分层结构或简单的参考帧插值。然而，真实世界中的视频运动高度复杂且非均匀，导致压缩性能不佳。此外，神经编解码器中的传统熵模型平等对待所有通道，忽略了通道间熵的不均匀分布，限制了压缩效率的进一步提升。这为开发一种更自适应、学习驱动的双向压缩框架创造了重要机遇，该框架能够更好地处理复杂运动模式并优化熵编码。

## 技术

本专利提出了一种新颖的基于学习的双向视频压缩方法，通过两项关键创新提升压缩性能。首先，它采用孪生神经网络从两个双向参考帧（先前和未来重建帧）和当前帧生成单个高质量参考帧，而不是依赖固定插值或分层结构。该孪生网络独立地从帧中提取多层特征堆栈，通过归一化、减法、缩放和L2范数计算特征相似性距离，并将这些距离与参考帧融合，以生成用于条件压缩的优化参考帧。这种方法适应复杂、非均匀的运动。其次，专利提出了一种逆通道熵模型，根据熵分布重新构建通道的编码顺序。它不再平等对待所有通道，而是优先编码熵较低的通道，并将其作为上下文来编码熵较高的通道，从而为高熵通道提供更多参考输入。该模型将潜在表示沿通道维度不均匀地分组，首先编码熵最小的组，并逐步使用已编码组作为条件来编码更高熵的组，从而提高整体压缩效率。

## 优势

- 实现卓越的压缩性能，在PSNR方面平均节省11.29%的比特率，在MS-SSIM方面平均节省54.34%，优于VVC低延迟B配置。
- 通过基于学习的参考帧生成适应真实世界中复杂、非均匀的运动，避免了固定分层结构的局限性。
- 利用通道间不均匀的熵分布提高熵编码效率，使高熵通道能够从更多上下文信息中受益。
- 提供灵活、即插即用的参考帧生成器，可集成到其他基于学习的视频压缩系统中。
- 减少误差传播，并在帧间保持稳定的重建质量，如消融研究所证明。
- 通过在训练中使用不同的 $\lambda$ 值，支持可调节的率失真权衡，实现可扩展的压缩。

备注

IDF:1581

IP状态

已申请专利



技术成熟度等级 (TRL) ?

4

发明人

邝得互教授

王诗淇教授

郭海峰博士

查询: kto@cityu.edu.hk

Follow-on Funding

Develop Concept

Proof Concept

Build Value

## 应用

- 视频流媒体服务（如Netflix、YouTube），在保持高视觉质量的同时降低带宽使用。
- 视频会议平台（如Zoom、Teams），实现高效的实时视频传输。
- 视频存储和归档系统，最小化高分辨率视频内容的存储需求。
- 带宽和计算资源有限的移动和边缘设备。
- 视频点播和直播广播基础设施。
- 监控和安全系统，需要高效的视频压缩以进行存储和传输。

