

## 摘要

近年来，随着视频直播业务的高速发展，用户对高画质、高流畅度、高平滑性和低延迟的视频直播需求日益增强。为了降低端到端延迟，国际标准化组织 MPEG 针对基于 HTTP 的动态码率自适应传输技术制定了低延迟视频直播扩展标准，通过将视频分段切分为更小的视频分块以降低端到端延迟。目前，在低延迟限制下设计面向客户端的码率决策算法以及面向服务器端的传输码流优化算法已成为视频技术领域最具挑战性的课题。本文深入研究了低延迟视频直播的特性，结合人工智能技术，设计了更智能的码率决策算法以及更高效的传输码流优化算法，显著提升了用户体验质量，保障了从服务器端到客户端的低延迟。本文的创新点主要包括以下三个方面：

第一，提出了一种基于模型预测控制的码率决策算法，为低延迟视频直播提供通用解决方案。针对低延迟视频直播中带宽检测不准确的问题，提出结合视频帧结构信息与应用层数据块到达时间来检测带宽，降低了 88.6% 的检测误差。提出应用深度学习预测带宽分布，为决策算法提供丰富的网络信息。构建了帧级别的传输演进模型与用户体验质量模型，将网络信息和客户端信息输入到模型预测控制器中做出码率决策。在国际通用平台上测试，与已有算法相比，该算法在合成网络轨迹下提升用户体验质量 39.7%，在移动网络轨迹下提升用户体验质量 37.9%。所提算法在更大范围的视频轨迹和网络轨迹下测试均取得了良好的性能，展现出了良好的泛化性。

第二，提出了一种基于深度强化学习的码率决策算法，为低延迟视频直播提供个性化解决方案。深入分析视频分块传输特性，构建帧级视频直播传输仿真环境，并且将低延迟视频直播建模成马尔可夫决策过程。用神经网络表示码率决策策略，通过模仿学习对码率决策智能体进行预训练。针对个性化的用户偏好，应用强化学习继续优化码率决策智能体的参数。通过与仿真环境交互，仅需几十分钟的训练便可获得高质量的码率决策策略，能迅速提供个性化的码率决策算法。在国际通用平台上测试，与已有算法相比，该算法在合成网络轨迹下提升用户体验质量 34.7%，在移动网络轨迹下提升用户体验质量 29.1%。

第三，提出了基于视频内容感知的传输码流优化算法，包括面向低延迟视频直播的码率阶梯优化算法和应对网络与码率不匹配的实时丢帧优化算法。在转码过程，分析了视频内容与帧率、分辨率和编码参数的非线性关系，通过深度学习模型从原始帧中直接提取视频复杂度信息，预测视频分段编码后的视觉质量和大小。进而构建了面向低延迟视频直播的码率阶梯优化问题，通过求解该优化问题得到最优的帧率、分辨率和编码参数组合。在发送过程，分析了每帧视频对整体视觉质量的影响以及帧与帧之

间的依赖关系，获得了视频帧的重要性评分。针对网络与码率不匹配的情况，根据视频帧的重要性构建了实时丢帧优化模型，在传输过程中通过丢帧来提升用户体验。实验证明，在码率相当的前提下，所提码率阶梯优化算法可以提升视觉质量（VMAF）5.17分。所提实时丢帧优化算法显著提升了用户体验质量，明显优于发表时的经典算法。

综上所述，本文从客户端和服务端两个方面入手，研究了面向低延迟视频直播的传输优化问题，提出了基于模型预测控制的码率决策算法、基于深度强化学习的码率决策算法和基于视频内容感知的传输码流优化算法。实验部分对所提算法进行了充分验证，在保障了端到端低延迟的前提下，所提算法显著提升了用户体验质量，为进一步推动基于视频分块传输的低延迟视频直播技术落地应用提供了理论和实践支持。

**关键词：** 视频直播，视频传输，码率自适应，流媒体传输，用户体验质量