Paper Summary & Critique

Paper Title	OmniLive: Super-Resolution Enhanced 360° Video Live Streaming for Mobile Devices
DOI	10.1145/3581791.3596851
Name	Sahil Pattni
Concordia ID	40216177
Course	COMP 691
Lecture	5

Summary

In this paper, the authors introduce a novel method to upscale real-time 360° video quality via a flexible deep neural network. The system - OmniLive - is capable of video super-resolution on such real-time video data while maintaining a speed of 30 frames per second. OmniLive achieves a fast inference speed by implementing anytime inference, where the DNN has multiple exists throughout the architecture, instead of just a single exit at the end of a propogation. Additionally, given that the model is intended to run on a variety of mobile devices, OmniLive also implements a neural architecture search (NAS) system, that allows it to determine the optimal neural architecture for the given hardware.

Strengths

1. OmniLive addresses the limitations of live streaming 360° video given the current network bandwidths by improving the video quality at the client side.
2. Achieves an inference time of ≤ 33 milliseconds per frame.
3. Is able to adapt and optimize model performance via NAS for various hardware conditions.
    - For example, the NAS determined that an architecture based on a bidirectional recurrent neural network was the most optimal for the Galaxy S22, an architecture that may have been too computationally heavy to run on other mobile devices.
4.

Weaknesses

- OmniLive's real-time inference speed is limited to process 30 frames per second, which is impressive in isolation. However, given the nature of real-time 360° video, a 60 frames per second (16 ms) processing rate would be more desirable. OmniLive was able to achieve a rate as low as 10.97 at the first exit, with a caveat being that the PSNR was comparable to the worst-performing state-of-the-art model they benchmarked against (EVSRNet).

Applicability to Practice

There are still limitations to streaming high-quality omni-directional video, key of them being the bandwidth limitations of current network environments. OmniLive shows promise as a solution to such a problem, transferring the responsibility of up-scaling to client devices. As mobile devices get more powerful, the solution presented in this paper could help bolster the quality of experience for real-time 360° video streaming.

Notes for Improvement

The authors note that sending an entire omni-directional video requires significant network overhead, which could be reduced by only transmitting the view port window that the viewer is predicted to be looking at (with margins for error).The authors note that certain solutions such as Flare [1], more recent developments such as the saliency-driven predictions by Shibo Wang et al. [2] may offer an increase in performance that may make such a strategy feasible.

References

[1] Feng Qian et al. "Flare: Practical Viewport-Adaptive 360-Degree Video Streaming for Mobile Devices". In: *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. Association for Computing Machinery, 2018. DOI: 10.1145/3241539.3241565.

[2] Shibo Wang et al. "Robust Saliency-Driven Quality Adaptation for Mobile 360-Degree Video Streaming". In: *IEEE Transactions on Mobile Computing* (2023), pp. 1–18. DOI: 10.1109/TMC.2023.3235103.