Paper Summary & Critique

Paper Title	Towards Optimal Real-time Volumetric Video Streaming: A Rolling Optimization and Deep Reinforcement Learning Based Approach
DOI	10.1109/TCSVT.2023.3277893
Name	Sahil Pattni
Concordia ID	40216177
Course	COMP 691
Lecture	8

Summary

The authors propose a system with two key-components

- A rolling prediction-optimization-transmission (POT) framework to predict both the network bandwidth and the FoV for a rolling window.
- A DRL-based model to perform adaptive transmission to solve the optimization problem of the rolling POT data in real-time.

Strengths

Weaknesses

Applicability to Practice

Notes for Improvement

Paper Title	MetaStream: Live Volumetric Content Capture, Creation, Delivery, and Rendering in Real Time
DOI	10.1145/3570361.3592530
Name	Sahil Pattni
Concordia ID	40216177
Course	COMP 691
Lecture	8

Summary

MetaStream is an end-to-end system that captures, renders and delivers volumetric video: a type of video where the viewer can *explore* the video with six degrees of freedom.

Strengths

- MetaStream's main appeal is its efficiency: despite the high overhead of capturing and streaming volumetric video, the authors of the paper have achieved a streaming rate of 30 frames per second [1].
- MetaStream's Dynamic Camera Calibration (DCC) updates the spatial pose of each camera to be updated as it moves, so that the 3D content can be rendered accurately.
- MetaStream goes to great lengths to reduce the data overhead of streaming, using techniques such as Cross-Camera Point Removal (where common overlapping content between the cameras is removed), Edge-assisted Object Segmentation (where target objects are segmented from complex backgrounds), and Foveated Volumetric Content Rendering (where the level of detail is adjusted based on where the user is looking).

Weaknesses

- The dynamic camera calibration system makes two key assumptions:
  1. The movement speed of each camera must be $\leq 4.2\ m/s$ ($\approx 15\ km/h$). Events with rapid movement (e.g. action sequences or sporting events) may not be able to utilize this camera positional synchronization technique effectively.
  2. The scene should also contain rich feature points.
- MetaStream is not optimized for scenarios where there are multiple receivers.

Applicability to Practice

If the limitations above can be addressed, this system shows promise for a variety of applications; concepts from science-fiction such as holographic video conferencing could become a reality.

Notes for Improvement

The current system requires an initial calibration of the camera's position. With remotely movable cameras (e.g. mounted to / integrated on drones), this process could be automated, having an array of drones map out a target in 3D space.

[1] 30 fps on WiFi networks.