Volumetric Video Live Streaming (Final Report)

Zoraiz Khan
University of Calgary
Calgary, AB, Canada

ABSTRACT

In an era dominated by virtual platforms, this paper explores the extensive applications of volumetric video streaming, envisioning immersive scenarios in education, healthcare, and entertainment. Despite the great potential, challenges like prohibitive costs and technical intricacies impede widespread adoption. The research strives to contribute an end-to-end system, simplifying volumetric video streaming and creating an accessible, user-friendly, and open-source platform. Beyond a technical solution, the project aspires to offer an open-source platform tailored for real-time volumetric video streaming, addressing usability, knowledge dissemination, and economic challenges. The main hurdle lies in the lack of user-friendly platforms, coupled with knowledge gaps and economic barriers, making the technology less inclusive. The research aims to democratize access, fostering innovation in this dynamic field. Looking ahead, the project outlines a plan emphasizing the design of a Docker Container-based system, prioritizing modularity and adaptability. Central to our contribution is the introduction of a novel benchmarking approach within the system, expanding evaluation beyond traditional metrics to include qualitative exploration of the user experience. This benchmarking component promises a more comprehensive understanding of volumetric video streaming performance, enhancing real-time monitoring and facilitating continuous refinement by incorporating testing and feedback loops. The ultimate goal is of this research paper is to contribute to volumetric video exploration and to pave the way for future research directions, fostering curiosity, innovation, and exploration in this innovative sphere.

ACM Reference Format:
Zoraiz Khan. 2024. Volumetric Video Live Streaming (Final Report). In . ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The notable rise in our dependence on virtual platforms has reshaped the dynamics of human interaction. Envision a realm where the applications of volumetric video streaming extend far beyond our current understanding. In educational settings, students might immerse themselves in historical events or explore architectural wonders through virtual experiences. In the realm of healthcare, volumetric video streaming emerges as a pivotal tool for remote surgical training, providing aspiring surgeons with real-time observation of procedures. Furthermore, the entertainment industry could leverage from this technology in storytelling by harnessing the interactive potential of this technology, breaking free from traditional formats. These diverse scenarios illuminate the versatility of volumetric video streaming, showcasing its revolutionary potential across various of fields.

The objective of this research is to contribute to the establishment of an end-to-end system that simplifies the streaming of volumetric videos. This will be achieved by developing a system that sets up a reproducible, easy-to-configure and benchmarkable platform that is tailored for both researchers and enthusiasts. The anticipated outcome of this collaborative system will empower users to navigate the volumetric video landscape more seamlessly.

This research aims to address the challenge of enhancing virtual interactions through the medium of live streaming volumetric video. Volumetric video, with its unique ability to capture and render three-dimensional spaces, offers the potential to narrow the gap between physical and virtual encounters. However, the current landscape of volumetric video platforms grapples with challenges such as prohibitive costs, technical complexities, and latency issues. These hurdles act as substantial roadblocks, impeding the widespread adoption of this technology, both within academic research and mainstream usage.

In response to these challenges, this project aspires to deliver more than just a technical solution; it seeks to provide an open-source platform that is accessible, user-friendly, and specifically tailored for real-time or near-real-time volumetric video streaming. The primary goal is to spark curiosity, foster innovation, and encourage exploration in the domains of volumetric video. This report provides a thorough overview of the substantial efforts dedicated to developing a platform for live volumetric video streaming. The exploration encompasses a literature review, preliminary solutions, and planned methodologies, shedding light on related works, along with a discussion on the current status of the system, its potential, and envisioned use cases.

2 BACKGROUND AND MOTIVATION

2.1 Background

Volumetric video technology has emerged as a captivating medium, revolutionizing the way we experience visual content by offering immersive three-dimensional reproductions of scenes. Unlike traditional video formats, which are limited to two dimensions, volumetric video leverages advanced imaging techniques such as arrays of cameras and depth-sensing devices to capture comprehensive volumetric representations of real-world objects and environments (Figure 1). This evolution in technology has unlocked new possibilities for content creation, storytelling, and user engagement.

At its core, volumetric video works by capturing not just the surface appearance of objects but also their spatial depth information,

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

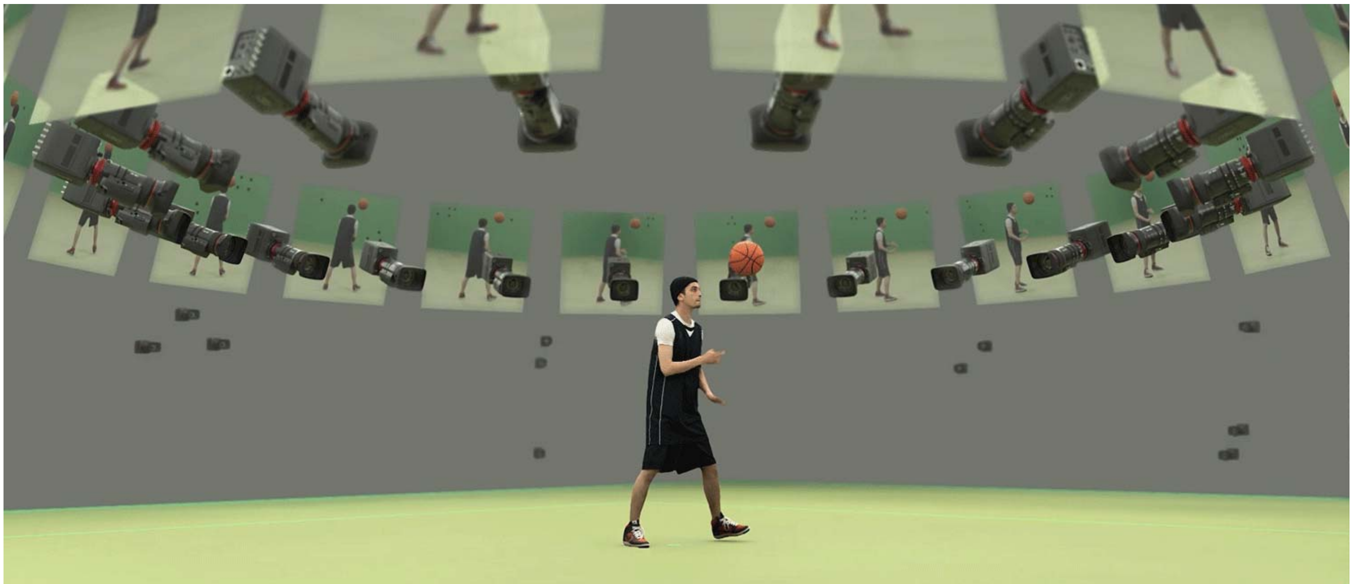


Figure 1: Volumetric video streaming system image capturing array of cameras. Adapted from [1].

allowing viewers to explore scenes from different angles and perspectives. This is achieved through the simultaneous recording of multiple viewpoints using an array of synchronized cameras, combined with depth-sensing technologies like LiDAR or structured light scanners. These captured viewpoints are then processed and integrated to create a three-dimensional representation of the scene, commonly referred to as a "point cloud" or "volumetric mesh."

However, despite its potential, the current state of volumetric video technology presents challenges, particularly in the realm of end-to-end streaming. The absence of user-friendly platforms and standardized workflows complicates the process of capturing, processing, and delivering volumetric content in real-time. This complexity extends beyond technical intricacies, encompassing usability issues that hinder widespread adoption and limit the realization of the technology's full potential.

To address these challenges, there is a growing need for comprehensive platforms that streamline the entire volumetric video production pipeline, from capture to delivery. By providing intuitive tools, automated processes, and standardized protocols, such platforms can democratize access to volumetric video technology, empowering creators and developers to unleash their creativity and explore new frontiers in immersive storytelling.

2.2 Motivation and Challenges

The motivation behind this research stems from the desire to seamlessly integrate a user-friendly volumetric live-streaming video system into existing platforms. Recognizing the game-changing capabilities of volumetric video, there is a drive to simplify experimentation and enhance reproducibility within the volumetric video domain.

However, the path to achieving seamless integration is fraught with challenges, rooted in the inherent complexities of volumetric video technology. At its core, volumetric video involves capturing and rendering three-dimensional representations of real-world objects and environments, requiring sophisticated imaging techniques and computational algorithms. The process of capturing volumetric data involves synchronizing multiple cameras or depth-sensing devices to capture spatial and temporal information simultaneously, leading to massive data volumes and computational requirements.

The main challenge revolves around the absence of user-friendly platforms for volumetric streaming. Like previously mentioned, this is partially due to the technology becoming more complex as it evolves. It is crucial to democratize access because this not only facilitates the work of researchers but also empowers enthusiasts eager to explore and create volumetric videos.

Moreover, knowledge gaps and usability issues compound the challenge, raising questions about the inclusivity of the technology. Economically, the cost associated with volumetric video platforms acts as a bottleneck for widespread adoption. The high costs and technical complexities, makes this technology inaccessible especially to the average enthusiast who lacks significant financial resources. In short, the challenges encountered extend beyond technical intricacies, raising fundamental questions about accessibility, knowledge dissemination, and the economic considerations tied to the technology's widespread use.

These usability challenges act as a significant impediment for both researchers and enthusiasts, creating barriers to entry. Addressing these multifaceted challenges requires a nuanced approach, which will be pivotal for democratizing access to volumetric video streaming and fostering innovation in this dynamic field.

3 RELATED WORK

The landscape of live volumetric video streaming has witnessed notable advancements, each presenting distinct approaches and contributions. MetaStream [2] introduces a system designed on low-cost commercial platforms, achieving close to 30FPS on Wi-Fi networks. The paper addresses critical challenges, including camera calibration, resource-constrained smart cameras, point cloud synthesis for multi-camera setups, and reduction of transmission delays. By incorporating dynamic camera calibration, selective segmentation, efficient point cloud synthesis, and foveated rendering on MR devices, MetaStream significantly reduces end-to-end latency by up to 31.7%, simultaneously enhancing visual quality by up to 12.5% when benchmarked against state-of-the-art systems.

In a parallel endeavor, GROOT [4] takes a mobile-centric approach to volumetric video streaming, introducing a novel PD-Tree data structure. This innovation enables real-time streaming at 30fps on mobile devices, addressing the bottleneck of decoding through the independence of leaf nodes. GROOT's suite of techniques, such as color information compression and filtering of 3D points, results in faster frame rates for high-fidelity volumetric videos on mobile platforms compared to prior methods. The paper not only contributes a performant system but also offers insights into memory usage, computation for decoding, and data-size reduction techniques.

Addressing a gap in virtual communication tools, the third paper [3] presents a novel approach to live volumetric video streaming, focusing on authenticity and immersive interaction. Leveraging the capabilities of Wi-Fi 6 networks, the research emphasizes an open-source platform's development for widespread accessibility. By integrating Microsoft Kinect with emerging network technologies, the paper democratically places volumetric video within reach. Acknowledging challenges in processing speed, the study suggests avenues for future research, particularly in optimizing heavy-data 3D applications and improving multi-camera support. The conceptual contribution lies in the creation of a common research platform, facilitating global collaboration and paving the way for artistic and entrepreneurial expressions in the evolving landscape of live, immersive, virtual interactions.

These papers address the crucial aspect of performance in live volumetric video streaming. They present advancements achieved in terms of frame rates, latency reduction, and visual quality improvements. Despite their commonalities, each paper introduces distinguishing characteristics. These include unique approaches to performance metrics prioritization, target platforms, and conceptual contributions. At the same time they also acknowledge and address common challenges in the field, such as camera calibration, resource constraints, transmission delays, and processing speed. While the papers focus on specific technologies and platforms, there is potential for implementing standards or protocols that promote interoperability. The aim of our research will Creating a common framework that allows different systems to work seamlessly together could be a valuable addition.

In addition to the comprehensive exploration of performance aspects in live volumetric video streaming, our research aims to contribute to the field by creating a common framework that fosters interoperability among different systems. In order to address challenges mentioned in these papers such as transmission delays and processing speeds, our novel approach incorporates a specialized benchmarking module, poised to provide valuable insights and pave the way for overcoming these obstacles.

4 PLATFORM DESIGN

The culmination of the presented papers on live volumetric video streaming have undeniably marked significant strides in the domain of volumetric video technology. As we chart the course for future exploration and development, several untapped avenues emerge, showcasing the novel potential for advancing this transformative field. At its core, the project aims to introduce a Docker Container-based framework, positioned to reshape the landscape of volumetric video exploration.

4.1 Novelty Contribution

Expanding on the ideas proposed in existing paper [3], the central concept revolves around creating an environment that is both platform-independent and easily reproducible. The modular design philosophy serves not only as a foundational principle but as a driver for facilitating seamless integration of future innovations. This approach seeks to overcome the constraints of current volumetric video systems, establishing an ecosystem conducive to continuous evolution.

This research introduces a groundbreaking approach to evaluating volumetric video streaming quality by incorporating a specialized system benchmark component within a separate container. While traditional video metrics focus on quantitative measures like frames per second (FPS), bitrate, and stalls, this novel addition expands the scope to include qualitative exploration, aiming to uncover the subtleties shaping the volumetric video experience.

4.1.1 Why it is Necessary.

- **Comprehensive Understanding**: By integrating a system benchmark component, this research ensures a more holistic evaluation of the volumetric video streaming system. It goes beyond surface-level metrics to provide deeper insights into performance and user experience.
- **Real-time Monitoring**: The system benchmark component facilitates real-time monitoring of key metrics such as FPS, high, low, and average frame load time. This continuous monitoring allows for immediate identification of performance issues and enables prompt adjustments to optimize the streaming experience.
- **Modular Design Consistency**: Keeping the benchmark component in a separate container aligns with the overarching design principle of modularity. This modular approach ensures platform independence and facilitates seamless integration with the main volumetric video streaming application.
- **Scalability and Flexibility**: The inclusion of a system benchmark component enhances the scalability and flexibility of the volumetric video streaming platform. As the system evolves and adapts to new technologies, the benchmarking module can be easily updated and expanded to accommodate emerging requirements.
- **Validation and Optimization**: The system benchmark component serves as a valuable tool for validating the performance of the volumetric video streaming system under various conditions. By providing comprehensive metrics, it enables researchers and developers to identify performance

bottlenecks and optimize system parameters for an enhanced user experience.

This approach represents a paradigm shift in the evaluation of volumetric video streaming quality. While existing methodologies primarily focus on quantitative measures, this research introduces a qualitative exploration that delves deeper into the nuances of the user experience. By incorporating a system benchmark component within a separate container, the research not only enhances the comprehensiveness of the evaluation process but also demonstrates a commitment to modularity and innovation in platform design. This novel contribution not only advances the field of volumetric video streaming but also sets a new standard for performance evaluation in multimedia applications.

4.2 Specifications

The forthcoming stage involves the design and conceptualization of a Docker Container-based system, emphasizing modularity for future adaptability. Subsequently, the implementation phase will unfold, focusing on the identification and analysis of specialized metrics that transcend routine video assessments, diving into the nuanced and qualitative aspects of volumetric video streaming quality.

At the heart of the volumetric video streaming system is the **Docker-Compose Interface** (Figure 2), serving as the orchestrator and facilitator of communication among the various components. This module is responsible for managing the deployment and configuration of Docker containers, allowing for a responsive and efficient coordination of resources. It acts as the gateway for transmitting configuration settings, parameters, and status updates, thereby playing an important role in maintaining the system's adaptability.

The **Reci Module's** (Figure 2) purpose is to capture the raw data. It collaborates closely with the Docker-Compose Interface and it translates the configuration parameters into an orchestrated deployment of containers for data capture from cameras or sensors. This module not only records video data but potentially generates metadata, providing valuable contextual information. Receiving input from the Broadcasting Module, the **Broadcasting Module** (Figure 2) simultaneously engages in the real-time streaming process, converting the recorded data into a format suitable for live consumption.

The **System Benchmarking Component** (Figure 2) operate as the analytical backbone of the system, evaluating the performance metrics derived from the Streaming Module. This module provides valuable insights into the system's efficiency, enabling continuous improvement and optimization. By establishing a dynamic feedback loop, it contributes to the refinement of the overall system.

The research will embrace an iterative approach, incorporating testing and feedback loops to refine and optimize volumetric videos based on emerging findings. The ultimate objective of this report is to synthesize the research journey, provide insights, and give significant contributions. Following this, the implementation plan and a thoughtful discussion will unravel the results.

5 PLATFORM IMPLEMENTATION

Platform implementation represents the crucial phase where the theoretical blueprint of the volumetric video streaming system is transformed into a concrete and functional software framework. Building upon the design principles and specifications outlined earlier, this phase involves the meticulous development, integration, and testing of individual modules and components. By translating abstract concepts into executable code, the platform implementation stage lays the foundation for realizing the envisioned system's capabilities in a real-world context. Through iterative refinement and validation, this process ensures that the platform meets its objectives of modularity, scalability, and usability, setting the stage for practical application and experimentation in volumetric video streaming.

5.1 Recording Module

The recording module within the volumetric video streaming system utilizes the Kinect sensor as the primary capture device. The Kinect sensor captures both RGB color data and depth information at a consistent rate of 30 frames per second (FPS). This data is transmitted in real-time via a USB 2.0 connection to the recording module, where it undergoes further processing. Upon receiving the RGB and depth data, the recording module utilizes Open3D [5], a versatile library for 3D data processing, to transform the raw data into Polygon File Format (PLY) files. These PLY files contain three-dimensional point cloud data along with color attributes, making them suitable for volumetric video representation. These processed PLY files are stored within a Docker container, ensuring efficient management and accessibility. By encapsulating the files within a containerized environment, the recording module maintains modularity and platform independence, allowing seamless integration with other system components.

To expedite development and preserve modularity, the recording module leverages stored PLY files from an existing project. The decision to leverage stored PLY files from an existing project within the recording module is a strategic choice aimed at expediting development while maintaining modularity. This approach capitalizes on the wealth of data already available within the system, reducing the need for real-time data capture and processing during the development phase. Approximately 200 PLY files are readily available within the Docker container, and a single file loader approach is employed to mimic real-time data capture. This approach streamlines development without compromising flexibility. Each PLY file is grouped into sequences of 30 frames, representing one second of volumetric video content. This grouping ensures synchronization and coherence between frames, facilitating smooth playback for end-users.

In essence, the processing of captured data within the recording module is a complex yet essential aspect of the volumetric video streaming pipeline.

5.2 Broadcasting Module

The broadcasting module operates within a separate Docker container, adhering to the modular design principles of the volumetric video streaming system. By encapsulating the broadcasting functionality within a containerized environment, the module ensures

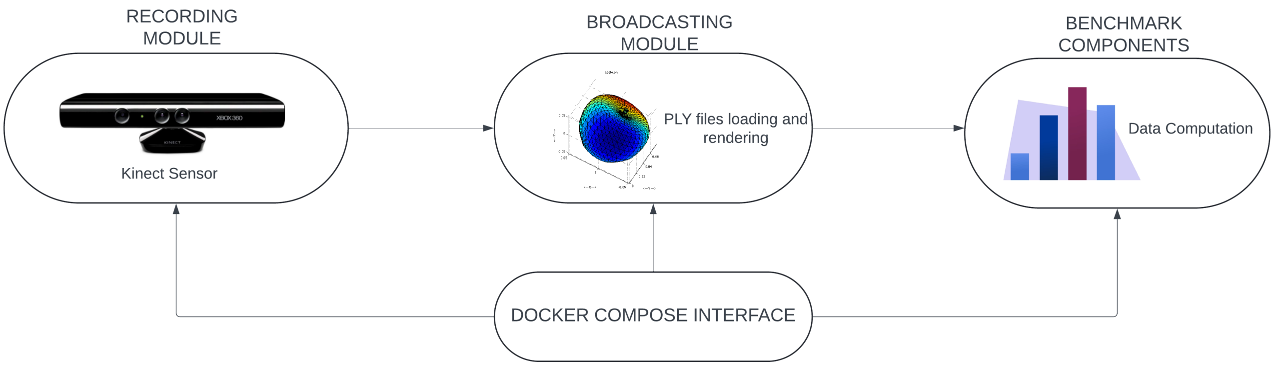


Figure 2: Conceptual volumetric video streaming system architecture diagram.

Figure 3: A still frame captured by the Kinect sensor. Adapted from [3].

platform independence, ease of deployment, and seamless integration with other system components. This module utilizes the Meta Quest 3 headset for rendering volumetric video content, chosen for its accessibility and feature-rich VR capabilities. With its widespread availability and user-friendly design, the Meta Quest 3 headset offers an ideal platform for showcasing volumetric video content to a broad audience.

In the broadcasting module, real-time rendering and broadcasting processes are initiated upon receiving volumetric video data in the form of PLY files. While the frame rate of the video stream typically hovers around 25 frames per second (fps), translating to an average frame duration of 40 milliseconds per frame, the loading of PLY files often outpaces the broadcasting rate. Consequently, some frames may be loaded before they are scheduled for broadcasting. To address this disparity between frame loading and broadcasting timing, the broadcasting module implements a broadcasting queue mechanism.

```
# Rendering and broadcasting processes initiated
# upon receiving PLY files
render_and_broadcast(PlyFile ply):
    FRAME_RATE = 30  # fps

    load_frame(ply)

    while (broadcast_queue is not empty):
        frame = retrieve_from_queue()
        broadcast(frame)
        remove_from_queue(frame)

# Broadcasting queue mechanism implementation
def add_to_queue(PlyFile ply):
    broadcasting_queue.enqueue(ply)

def load_frame(PlyFile ply):
    add_to_queue(ply)

def retrieve_from_queue():
    return broadcasting_queue.dequeue()

def remove_from_queue(PlyFile ply):
    broadcasting_queue.remove(ply)
```

Figure 4: Psuedo-code for broadcasting module.

As each frame is loaded, it is added to a broadcasting queue, ensuring that frames are broadcasted in the correct sequence and at the appropriate timing. This queue management system optimizes the broadcasting process, minimizing latency and ensuring smooth transmission of volumetric video content (Figure 4).

Additionally, the broadcasting module collects raw data related to the streaming process, capturing crucial information about frame load times, broadcasting delays, and transmission metrics. This raw data serves as valuable input for the benchmark module, facilitating performance evaluation and optimization of the streaming system. The collected data, while comprehensive, remains unprocessed within the broadcasting module. It serves as a raw feed of streaming performance metrics, ready to be analyzed and interpreted by the benchmark module once the broadcast is complete. This separation of data collection and processing ensures that the benchmark module can perform detailed analysis and optimization based on the raw data provided by the broadcasting module (Figure 4).

5.3 Benchmark Module

The benchmark module operates within its dedicated Docker container, adhering to the modular architecture of the volumetric video streaming system. Its primary function is to analyze and optimize the performance of the streaming system based on raw data collected during the broadcasting process. Upon initiation, the benchmark module establishes a unidirectional communication channel with the broadcasting module to receive raw streaming performance data. This direct channel ensures efficient data transfer and minimizes overhead, streamlining the benchmarking process.

Once the benchmark module receives the raw data from the broadcasting module, it initiates processing to compute essential frame-related metrics. Some metrics include the average frame rate and average frame load time, which are crucial indicators for assessing render latency and overall streaming performance. In addition to frame-related metrics, the benchmark module computes the frequency of stuttering, a key parameter indicative of streaming quality. This calculation involves comparing the average frame rate with individual frame render times, enabling the identification of instances where frame delivery falls below the desired threshold. On top of that, the loading status of each file is also tracked. By monitoring the loading process, the module provides insights into the success rate of file loading operations and detects any instances of duplicate file loading. This functionality serves as a valuable diagnostic tool, prompting developers to address loading. Additionally, the benchmark module captures the number of points in the point cloud for each loaded file. This metric provides valuable information about the complexity and density of the volumetric data being processed. Understanding the point cloud density is crucial for optimizing resource allocation and system performance, as it directly impacts the computational load and memory requirements during rendering. By monitoring the number of points in the point cloud per file, we are able to identify patterns in data complexity and anticipate potential performance bottlenecks.

The modular design of the benchmark module facilitates rapid iteration and debugging of the streaming system. By isolating the benchmarking functionality into a separate container, developers can make targeted changes to the broadcasting module during debugging. This approach eliminates the need to restart the entire system, streamlining the development and optimization process. Furthermore, the benchmark module serves as a decision support tool, providing actionable insights derived from the analyzed data. Developers can leverage these insights to fine-tune system parameters, optimize resource utilization, and enhance overall streaming performance.

6 EVALUATION

The evaluation of the current system was conducted utilizing the benchmarking module, which provided valuable insights into performance metrics and system behavior. Through this analysis, opportunities for refinement and optimization were identified, guiding strategic adjustments aimed at enhancing the overall performance and user experience of the volumetric video streaming system.

Upon evaluation of the original software a point of inefficiency which was identified which was impacting the loading and rendering processes of PLY files. The issue observed was the repetitive loading of PLY files, leading to redundant resource utilization and potential performance degradation. This occurred when the total number of frames exceeded the actual number of PLY files available, resulting in multiple loading instances for certain files.

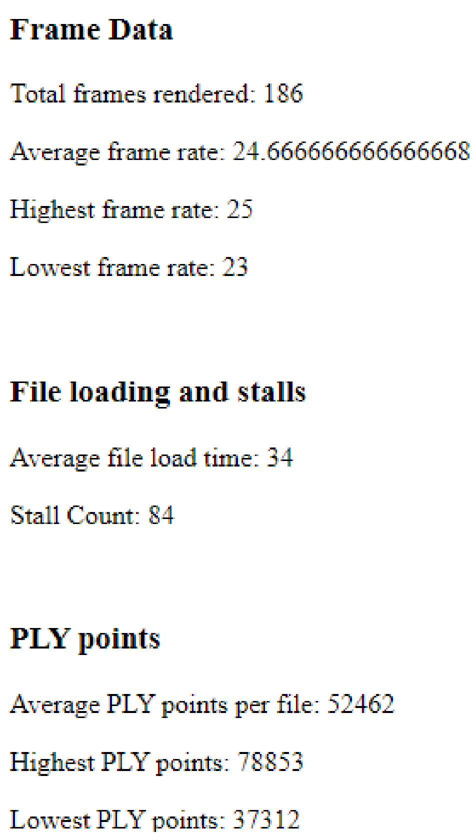
Furthermore, instances of frame loading timeouts were encountered during the evaluation process. These timeouts occurred due to the average frame load time surpassing the maximum threshold of 33.33 milliseconds. With an average frame load time of approximately 32 milliseconds, certain PLY files took slightly longer to load, triggering multiple frame loading timeout warnings. To address this issue, adjustments were made to optimize the loading process and minimize the occurrence of timeouts.

Additionally, the benchmarking module provided valuable insights into the discrepancy between the desired and actual frame rates of the software. While the expected frame rate was set at 30 frames per second (fps), the actual observed frame rate ranged between 24 and 25 fps. Further analysis revealed that each frame took approximately 40 milliseconds (Figure 2) from loading into buffer to rendering, exceeding the expected frame render time of 33.33 milliseconds. Surprisingly, the average frame load time calculated by the benchmarking module was 32 milliseconds, indicating that the remaining 8 milliseconds were utilized by the picture rendering function.

To optimize performance and enhance rendering efficiency, it became evident that adjustments were required in the picture rendering function to reduce the time taken for frame rendering. These optimizations aim to align the observed frame rate more closely with the desired frame rate, improving overall system responsiveness and user experience.

The benchmarking module also facilitated the calculation of the stall rate or stutter by comparing individual frame load times with the average frame load time. The analysis revealed a relatively high stall rate of around 80 frames, which is almost half of the total frames rendered (Figure 2), indicating interruptions or delays in the streaming process. Further investigation using the benchmarking tool highlighted significant variations in the number of points within the PLY files (Figure 2). This variability resulted in considerable differences in file sizes, subsequently affecting individual load times.

The observed variations in file sizes underscored the need for consistency in data density within the PLY files to mitigate loading delays and minimize stutter during streaming. Addressing this issue requires optimization of the data acquisition process to ensure uniformity in point cloud density across files. By reducing variability in file sizes, the system can achieve more consistent frame load times and improve overall streaming performance.



Frame Data

Total frames rendered: 186

Average frame rate: 24.666666666666668

Highest frame rate: 25

Lowest frame rate: 23

File loading and stalls

Average file load time: 34

Stall Count: 84

PLY points

Average PLY points per file: 52462

Highest PLY points: 78853

Lowest PLY points: 37312

Figure 5: Resulting output from the benchmark module.

7 FUTURE POSSIBILITIES

As we venture further into the domain of volumetric video streaming, it becomes increasingly apparent that there are numerous opportunities for exploration and enhancement within the existing framework. Building upon the groundwork laid out in our current work, in this section, we outline several potential avenues for future work, ranging from the integration of advanced technologies to the optimization of user experience and system performance. Each of these directions represents an opportunity to address existing challenges, unlock new capabilities, and ultimately contribute to the evolution of volumetric video streaming as a transformative technology.

Integration of Machine Learning Algorithms: Explore the integration of machine learning algorithms within the benchmark module to automate the identification of performance bottlenecks and suggest optimization strategies. This could involve training models on historical performance data to predict potential issues before they arise.

Dynamic Resource Allocation: Investigate techniques for dynamic resource allocation based on real-time performance metrics gathered by the benchmark module. This could involve dynamically adjusting processing power, memory allocation, or network bandwidth to optimize performance under varying workloads.

Expanded Metrics: To enhance the optimization of volumetric video streaming systems, it's crucial to broaden the range of performance metrics. This expansion should include monitoring memory usage, CPU and GPU utilization. These metrics offer valuable indicators of system health and performance bottlenecks. Additionally, tracking network throughput, latency, and packet loss facilitates real-time monitoring of data transmission quality.

8 CONCLUSION

In this paper we introduce a novel approach to evaluating live volumetric video streaming quality through the incorporation of a specialized benchmarking module within a separate Docker container. Unlike traditional video metrics that focus solely on quantitative measures such as frames per second (FPS) and bitrate, this pioneering methodology expands the evaluation scope to include qualitative exploration, delving into the nuanced aspects of the user experience. By integrating a benchmarking component, our research ensures a comprehensive understanding of volumetric video streaming performance, facilitating real-time monitoring, modular design consistency, scalability, flexibility, and validation.

In terms of design, our project adopts a modular and platform-independent Docker Container-based framework, emphasizing scalability and adaptability. The architecture, illustrated in (Figure 2), orchestrates communication among various components, including the Broadcasting Module, Streaming Module, and System Benchmarking Component. This design philosophy fosters seamless integration of future innovations while overcoming the constraints of existing volumetric video systems.

Moving to implementation, meticulous development, integration, and testing of individual modules lay the foundation for the envisioned system's capabilities in a real-world context. Leveraging the Kinect sensor and Open3D library, the Recording Module transforms raw data into Polygon File Format (PLY) files for volumetric representation. Meanwhile, the Broadcasting Module utilizes the Meta Quest 3 headset for real-time rendering and broadcasting.

Our findings using the Benchmark Module analyzes raw streaming performance data, computing essential frame-related metrics and providing valuable insights for optimization. Findings from the evaluation phase underscore opportunities for refinement and optimization in the volumetric video streaming system. Identified inefficiencies, such as repetitive loading of PLY files and frame loading timeouts, prompt strategic adjustments to enhance performance. Moreover, discrepancies between desired and actual frame rates necessitate optimizations in the picture rendering function to improve system responsiveness. Insights from the benchmarking module also highlight the importance of data consistency in mitigating loading delays and minimizing stutter during streaming.

This research contributes to advancing the field of volumetric video streaming by introducing a comprehensive evaluation framework, facilitating seamless integration, and optimizing system performance. Through iterative refinement and validation, our platform aims to empower users to navigate the volumetric video landscape more seamlessly, unlocking new possibilities for immersive storytelling, remote training, and interactive experiences across various fields.

REFERENCES

[1] Canon. 2023. https://global.canon/en/technology/volumetric-video2023.html.

[2] Yongjie Guan, Xueyu Hou, Nan Wu, Bo Han, and Tao Han. 2023. MetaStream: Live Volumetric Content Capture, Creation, Delivery, and Rendering in Real Time. In *The 29th Annual International Conference on Mobile Computing and Networking (ACM MobiCom '23)*. Association for Computing Machinery, Madrid, Spain, 15. https://doi.org/10.1145/3570361.3592530

[3] Kyle Jorgensen. 2024. Towards Live Volumetric Video Streaming. In *Proceedings of CPSC 601.38 (Conference'23)*. ACM, New York, NY, USA, 8. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[4] Kyungjin Lee, Juheon Yi, Youngki Lee, Sunghyun Choi, and Young Min Kim. 2020. GROOT: a real-time streaming system of high-fidelity volumetric videos. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking (MobiCom '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3372224.3419214

[5] Jaesik Park Qian-Yi Zhou and Vladlen Koltun. 2023. Open3D: A Modern Library for 3D Data Processing. https://www.open3d.org/wordpress/wp-content/paper.pdf