Play4D: Accelerated and Interactive Free-viewpoint Video Streaming for Virtual Reality and Light Field Displays

Jonghyun Kim, NVIDIA, Santa Clara, USA, jonghyunk@nvidia.com
Michael Stengel, NVIDIA, Santa Clara, USA, mstengel@nvidia.com
Amrita Mazumdar, NVIDIA, Santa Clara, USA, amritam@nvidia.com
Tianye Li, NVIDIA, Santa Clara, USA, tianyel@nvidia.com
Cheng Sun, NVIDIA, Santa Clara, USA, chengs@nvidia.com
David Luebke, NVIDIA, Santa Clara, USA, dluebke@nvidia.com
Shalini De Mello, NVIDIA, Santa Clara, USA, shalinig@nvidia.com

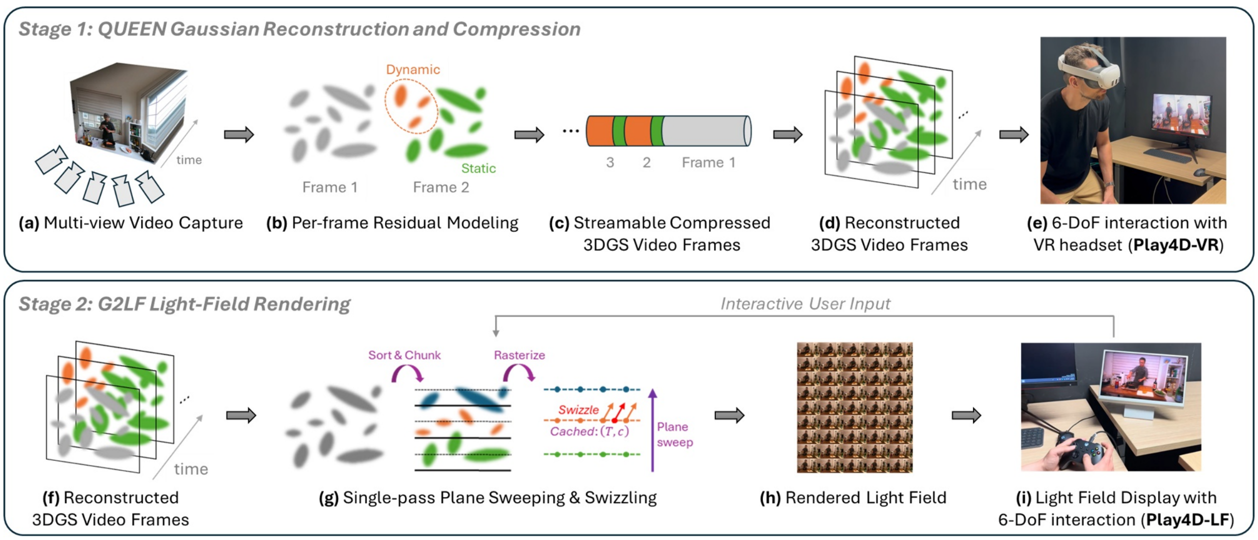

Figure 1: Overview of Play4D system architecture. Play4D powers two interactive demos: Play4D-VR, which streams compressed 4DGS to a VR headset for photorealistic 6-DoF exploration, and Play4D-LF, which renders 45+ views on LF displays for multi-user free-viewpoint playback. Stage 1 (top): The QUEEN algorithm compresses 4DGS sequences by modeling per-frame residuals, freezing static Gaussians and streaming dynamic updates for fast reconstruction and high compression. Stage 2 (bottom): G2LF performs LF rendering with single-pass plane sweeping and swizzle blending, synthesizing quilt images at 200+ FPS to enable smooth interactive playback.

ABSTRACT

We present Play4D, an accelerated and interactive free-viewpoint video (FVV) streaming pipeline for next-generation light field (LF) and virtual reality (VR) displays. Play4D integrates 4D Gaussian Splatting (4DGS) reconstruction, compression and streaming with an efficient radiance field rendering algorithm to enable live 6-DoF user interaction with photorealistic dynamic scenes. Our demo showcases two complementary experiences: Play4D-LF, a walk-up station offering interactive multi-view playback on a LF display; and Play4D-VR, a headset-based experience that enables full volumetric navigation through 4DGS content.

KEYWORDS

Free-Viewpoint Video, Light Field Display, 4D Gaussian Splatting

SA Emerging Technologies '25, December 15-18, 2025, Hong Kong, Hong Kong

ACM ISBN 979-8-4007-2133-5/2025/12.
https://doi.org/10.1145/3757373.3763771

ACM Reference Format:
Jonghyun Kim, Michael Stengel, Amrita Mazumdar, Tianye Li, Cheng Sun, David Luebke, and Shalini De Mello. 2025. Play4D: Accelerated and Interactive Free-viewpoint Video Streaming for Virtual Reality and Light Field Displays. In *SIGGRAPH Asia 2025 Emerging Technologies (SA Emerging Technologies '25), December 15-18, 2025*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3757373.3763771

1 INTRODUCTION

Free-viewpoint video (FVV) provides an immersive 4D viewing experience, enabling users to explore reconstructed 4D (3D + time) scenes with six degrees of freedom (6-DoF: translation + rotation) [Broxton et al. 2019; Collet et al. 2015; Kanade et al. 1997; Orts-Escolano et al. 2016; Overbeck et al. 2018; Taguchi et al. 2009; Wu et al. 2024]. Recent advances in neural scene representations, especially 3D Gaussian splatting (3DGS) [Kerbl et al. 2023], enable high-fidelity dynamic scene reconstruction from sparse multi-view inputs while significantly improving real-time rendering performance, making it well-suited for immersive content. Rendering these reconstructions on VR or light field (LF) displays lets users naturally observe scenes from multiple viewpoints with binocular disparity and motion parallax. LF displays further support multi-user 4D playback without wearable devices. This capability enables novel applications in sports, entertainment, telepresence, industrial monitoring, and robotic teleoperation.

However, creating interactive FVV using 4DGS for VR and LF displays introduces several challenges in reconstruction, transmission, and rendering. First, reconstructing and transmitting 4DGS for each frame independently leads to high latency and bandwidth consumption, ignoring temporal redundancy in dynamic scenes [Luiten et al. 2024]. Second, rendering 3DGS frames to LF displays requires generating 45 or more precisely angled views, causing further computational inefficiency [Kim et al. 2025]. Recent approaches improve LF multi-view rendering [Chen et al. 2022; Ji et al. 2025; Yang et al. 2024], but none achieve real-time performance or support 3DGS.

To address these challenges, we present Play4D, an accelerated and interactive end-to-end 4D reconstruction and playback pipeline. Play4D integrates two of our recent innovations: QUEEN [Girish et al. 2024], a streamable framework for compressing and transmitting FVVs, and G2LF [Kim et al. 2025], a real-time LF renderer optimized for 3DGS. Together, these methods enable low-latency streaming and real-time 6-DoF interaction with FVVs/4D videos, allowing users to freely change viewpoint, focus, and field-of-view during playback (Fig. 1). Our demo showcases a complete immersive pipeline from 4D reconstruction to compression, streaming, and real-time rendering for VR and LF displays.

QUEEN reconstructs compressed, streamable 4DGS from multi-view video. For the first frame, QUEEN takes input images from a multi-view capture setup (Fig. 1a) and reconstructs a traditional 3DGS reconstruction [Kerbl et al. 2023]. For subsequent frames, QUEEN computes residual differences in Gaussian attributes (position, scale, color, opacity) relative to the previous frame. This residual modeling achieves high compression and accelerates reconstruction. QUEEN reduces spatio-temporal redundancy by using a learned gating mechanism (Fig. 1b), which sparsifies static Gaussian attributes (green), and quantizing dynamic Gaussian attributes (orange) into compressed codes.

QUEEN performs Gaussian reconstruction on the server and serializes per-frame residuals as lightweight, time-ordered packets for low-latency streaming to many clients (Fig. 1c). On the client side, the QUEEN decoder incrementally reconstructs the 4D scene frame-by-frame, in real time, from the compressed packets (Fig. 1d). Compared to naive per-frame reconstruction (25 MB, 5 min/frame), QUEEN significantly compresses residuals (0.7 MB in 8 sec/frame), measured on a single NVIDIA RTX 3090 GPU [Girish et al. 2024]. Our demo setup uses an RTX 5090, which achieves higher rendering performance. After training, QUEEN renders frames at 300+ FPS, enabling practical, real-time FVV streaming.

While QUEEN enables efficient streaming, conventional 2D displays cannot deliver immersive 4D experiences. To address this, we introduce G2LF, a real-time LF renderer for 3DGS scenes, optimized for multi-view interactive playback on OpenGL-based systems. G2LF achieves real-time performance on LF displays using a single-pass plane-sweeping algorithm combined with swizzle blending. Instead of rendering each of the 45+ quilt views independently, the system approximates the visible 3D volume using a sequence of forward-sweeping planes. In our GL-based implementation, Gaussians are sorted by depth once per frame and divided into fixed-size chunks, minimizing CPU-GPU synchronization overhead while maintaining efficient memory usage.

Each chunk is rasterized into intermediate color and transmittance buffers, which are reused across multiple neighboring views via swizzle blending. By caching Gaussian attributes on the GPU and reusing shading results, G2LF reconstructs all quilt views without redundant sampling, significantly reducing computation (Fig. 1g). Additionally, we incorporate an adaptive anti-aliasing filter that adjusts Gaussian footprint sizes per plane, preserving rendering fidelity across different quilt resolutions and viewing angles.

On a single RTX 5090 GPU, the GL-based renderer achieves 200+ FPS for a 45-view, 512×910 quilt, enabling interactive, low-latency 6-DoF playback on LF displays while maintaining both temporal and angular consistency. This makes G2LF a core enabler of Play4D-LF (Fig. 1i), supporting smooth photorealistic visualization on commercial LF displays. Additional details are in [Kim et al. 2025].

2 DEMO

Our demo consists of two experiences: **Play4D-LF** and **Play4D-VR**. In both demos, participants explore photorealistic volumetric videos decoded in real-time from compressed QUEEN reconstructions. The Play4D-LF station (Fig. 1i) offers a walk-up experience with interactive 6-DoF FVV on a 16” LF display, accompanied by a 65” LF display showing an explanatory 3D slide deck. Visitors can intuitively explore content with correct binocular disparity and motion parallax, without wearing headsets. The Play4D-VR station (Fig. 1e) provides a fully immersive experience using a VR headset.

3 CONCLUSION

We introduce Play4D, an accelerated and interactive pipeline for streaming FVV to VR and LF displays. Play4D-LF offers attendees a glasses-free 4D experience, enabling them to explore 4DGS content in real-time using a LF display. Play4D-VR provides a fully immersive 6-DoF experience, allowing users to freely navigate dynamic volumetric scenes with a standalone headset. This unique experience can only be fully appreciated through hands-on interaction with the demo. Play4D sets a new standard in capturing, preserving, and reconstructing real-world 4D events, opening unprecedented possibilities in telepresence and interactive communication. We believe Play4D is a significant step towards the vision of a "4D streaming services," ultimately enabling the dream of broadcasting live immersive 4D events to millions of users.

REFERENCES

Michael Broxton, Jay Busch, Jason Dourgarian, Matthew DuVall, Daniel Erickson, Dan Evangelakos, John Flynn, Ryan Overbeck, Matt Whalen, and Paul Debevec. 2019. A low cost multi-camera array for panoramic light field video capture. In *SIGGRAPH Asia 2019 Posters*. 1–2.

Shuo Chen, Binbin Yan, Xinzhu Sang, Duo Chen, Peng Wang, Zeyuan Yang, Xiao Guo, and Chongli Zhong. 2022. Fast virtual view synthesis for an 8k 3d light-field display based on cutoff-nerf and 3d voxel rendering. *Optics Express* 30, 24 (2022), 44201–44217.

Alvaro Collet, Ming Chuang, Pat Sweeney, Don Gillett, Dennis Evseev, David Calabrese, Hugues Hoppe, Adam Kirk, and Steve Sullivan. 2015. High-quality streamable free-viewpoint video. *ACM Trans. Graph.* 34, 4, Article 69 (July 2015), 13 pages. https://doi.org/10.1145/2766945

Sharath Girish, Tianye Li, Amrita Mazumdar, Abhinav Shrivastava, David Luebke, and Shalini De Mello. 2024. QUEEN: QUantized Efficient ENcoding for Streaming Free-viewpoint Videos. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*. https://openreview.net/forum?id=7xhwE7VH4S

Luyu Ji, Xinzhu Sang, Shujun Xing, Xunbo Yu, Binbin Yan, and Jiahui Yang. 2025. Text-driven light-field content editing for three-dimensional light-field display based on Gaussian splatting. *Optics Express* 33, 1 (2025), 954–971.

T. Kanade, P. Rander, and P.J. Narayanan. 1997. Virtualized reality: constructing virtual worlds from real scenes. *IEEE MultiMedia* 4, 1 (1997), 34–47. https://doi.org/10.1109/93.580394

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.* 42, 4 (2023), 139–1.

Jonghyun Kim, Cheng Sun, Michael Stengel, Matthew Chan, Andrew Russell, Jaehyun Jung, Wil Braithwaite, Shalini De Mello, and David Luebke. 2025. Real-time 3D Visualization of Radiance Fields on Light Field Displays. https://doi.org/10.48550/arXiv.2508.18540 arXiv:2508.18540 [cs.GR] Submitted 25 August 2025.

Jonathon Luiten, Georgios Kopanas, Bastian Leibe, and Deva Ramanan. 2024. Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis. In *3DV*.

Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. 2016. Holoportation: Virtual 3D Teleportation in Real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 741–754. https://doi.org/10.1145/2984511.2984517

Ryan S Overbeck, Daniel Erickson, Daniel Evangelakos, Matt Pharr, and Paul Debevec. 2018. A system for acquiring, processing, and rendering panoramic light field stills for virtual reality. *ACM Transactions on Graphics (TOG)* 37, 6 (2018), 1–15.

Yuichi Taguchi, Takafumi Koike, Keita Takahashi, and Takeshi Naemura. 2009. TransCAIP: A live 3D TV system using a camera array and an integral photography display with interactive control of viewing parameters. *IEEE Transactions on Visualization and Computer Graphics* 15, 5 (2009), 841–852.

Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Xinggang Wang. 2024. 4d gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 20310–20320.

Zongyuan Yang, Baolin Liu, Yingde Song, Lan Yi, Yongping Xiong, Zhaohe Zhang, and Xunbo Yu. 2024. DirectL: Efficient Radiance Fields Rendering for 3D Light Field Displays. *ACM Transactions on Graphics (TOG)* 43, 6 (2024), 1–19.