

Real-time 3D Visualization of Radiance Fields on Light Field Displays

JONGHYUN KIM*, NVIDIA, USA
CHENG SUN*, NVIDIA, USA
MICHAEL STENGEL*, Nvidia, USA
MATTHEW CHAN, nvidia, USA
ANDREW RUSSELL, NVIDIA, USA
JAEHYUN JUNG, NVIDIA, USA
WIL BRAITHWAITE, NVIDIA, USA
DAVID LUEBKE, NVIDIA, USA
SHALINI DE MELLO, NVIDIA, USA

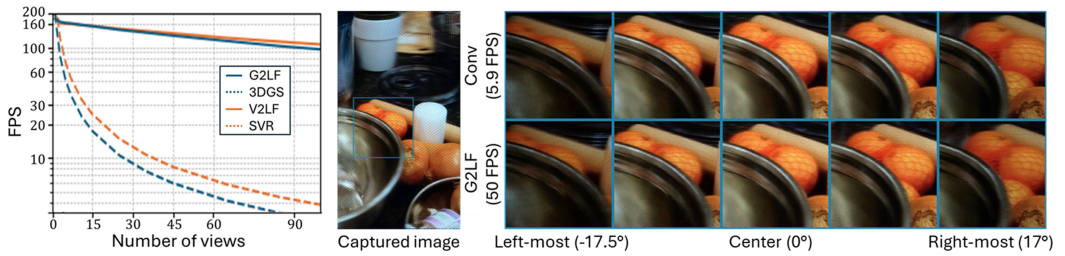


Fig. 1. We present a unified framework for real-time radiance field rendering on light field displays, supporting explicit representations including 3D Gaussians (G2LF) and sparse voxels (V2LF) via single-pass plane sweeping over intermediate multi-plane images (MPI). (Left) Render time comparison across view counts on the MipNeRF-360 dataset [Barron et al. 2022], measured at 512p on a NVIDIA RTX 5090. G2LF and V2LF (solid lines) maintain real-time performance (>60 FPS) beyond 90 views. (Right) Captured results of a 45-view rendered light field shown on a Looking Glass Go display, recorded with a motorized rotational stage, demonstrate high-quality 3D reconstruction across viewpoints. Note that the performance plot uses $N_{\text{chunk}} = 64$, whereas the captured light field on the right uses $N_{\text{chunk}} = 512$ for higher visual quality (see Sec. 4.2).

Radiance fields, including their recent efficient forms such as 3D Gaussian Splatting and Sparse Voxels, have revolutionized photorealistic 3D scene visualization by enabling high-fidelity reconstruction of complex environments, making them a natural match for light field displays. However, integrating these technologies presents significant computational challenges, as light field displays require many high-resolution renderings

*contributed equally to this research

Authors' Contact Information: [Jonghyun Kim](mailto:jonghyunk@nvidia.com), NVIDIA, Santa Clara, USA, jonghyunk@nvidia.com; [Cheng Sun](mailto:chengs@nvidia.com), NVIDIA, Santa Clara, USA, chengs@nvidia.com; [Michael Stengel](mailto:mstengel@nvidia.com), Nvidia, Santa Clara, USA, mstengel@nvidia.com; [Matthew Chan](mailto:matchan@nvidia.com), nvidia, Santa Clara, USA, matchan@nvidia.com; [Andrew Russell](mailto:arusell@nvidia.com), NVIDIA, Santa Clara, USA, arusell@nvidia.com; [Jaehyun Jung](mailto:jaehyunjung@nvidia.com), NVIDIA, Santa Clara, USA, jaehyunjung@nvidia.com; [Wil Braithwaite](mailto:wbraithwaite@nvidia.com), NVIDIA, Santa Clara, USA, wbraithwaite@nvidia.com; [David Luebke](mailto:dluebke@nvidia.com), NVIDIA, Santa Clara, USA, dluebke@nvidia.com; [Shalini De Mello](mailto:shalini@nvidia.com), NVIDIA, Santa Clara, USA, shalini@nvidia.com.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2026 Copyright held by the owner/author(s).

ACM 2577-6193/2026/7-ART56

<https://doi.org/10.1145/3820023>

from slightly shifted viewpoints, while radiance fields rely on computationally intensive volume rendering, which is intractable to achieve real-time speeds even with efficient scene representations. In this paper, we propose a unified and efficient framework for real-time radiance field rendering on light field displays. Rather than re-rendering each view independently, our method converts the input radiance field into shared intermediate sweeping planes that can be efficiently composited into dense light-field views in a single pass. Our method prioritizes shared, non-directional plane caching for real-time performance, trading fine view-dependent color effects for a modest increase in intermediate memory usage. Our framework generalizes across different scene representations without retraining and avoids repeated computation across views. We further demonstrate a real-time interactive application on a Looking Glass display, achieving 200+ FPS at 512p across 45 rendered views and enabling seamless, immersive 3D interactive viewing experiences. On standard benchmarks, our method achieves up to $22\times$ speedup compared to independently rendering each view, while largely preserving image quality.

CCS Concepts: • **Computing methodologies** → **Rendering**; *Appearance and texture representations*; *Shared memory algorithms*.

ACM Reference Format:

Jonghyun Kim, Cheng Sun, Michael Stengel, Matthew Chan, Andrew Russell, Jaehyun Jung, Wil Braithwaite, David Luebke, and Shalini De Mello. 2026. Real-time 3D Visualization of Radiance Fields on Light Field Displays. *Proc. ACM Comput. Graph. Interact. Tech.* 9, 4, Article 56 (July 2026), 19 pages. <https://doi.org/10.1145/3820023>

1 Introduction

Recent advancements in radiance fields have significantly improved both the quantity and quality of 3D content. Radiance fields represent 3D scenes by encoding density and color values across spatial coordinates and view directions, enabling photorealistic reconstructions of complex scenes. Neural Radiance Fields (NeRF) have enabled continuous view synthesis from sparse input images, making it possible to reconstruct complex 3D scenes with high precision [Mildenhall et al. 2021; Müller et al. 2022]. Radiance fields now encompass a variety of representations for improving rendering efficiency for real-time applications, including rasterization-based methods [Kerbl et al. 2023; Sun et al. 2024] and explicit scene structures [Sun et al. 2022; Takikawa et al. 2021] that avoid dense sampling.

Like other 3D content, radiance fields are most effectively visualized using 3D displays. Their ability to represent complex 3D structures aligns naturally with the capabilities of light field displays, which optically reconstruct the light rays of 3D scenes. Recent commercially available light field displays offer high spatial and angular resolution, enabling immersive 3D visualizations ([Leia Inc. 2025], [Sony Electronics Inc. 2025], [Looking Glass Factory Inc. 2026]). These displays provide binocular disparity and motion parallax, leveraging human depth perception to allow users to naturally perceive 3D structures. However, the integration of neural radiance fields with light field displays, introduces significant computational challenges. First, light field displays require high-resolution rendering from many (45+) slightly shifted viewpoints. Second, generating this many views from radiance fields, even with fast rasterization-based methods like 3D Gaussian Splatting [Kerbl et al. 2023] and Sparse Voxels [Sun et al. 2024] remains computationally intractable for real-time applications such as interactive 3D/4D content viewing.

Light field displays fundamentally face substantial computational overhead in rendering due to their unique optical design. Unlike conventional single-view 2D displays, they require the generation of multiple perspective views to reconstruct the full light field, necessitating a dense array of rays projected at precise angles. This significantly increases the computational burden compared to single-view 2D displays. Furthermore, precise optical alignment between the display panel and the lens array is critical; even minor angular or spatial misalignment during manufacturing can lead to incorrect ray-to-subpixel mappings, degrading visual quality (see Section 3 for details). Correcting

these errors requires per-device calibration and often a larger number of rendered views, which further increases the rendering cost.

Radiance fields face inherent computational challenges, particularly due to their reliance on volumetric rendering, where density and color values must be evaluated along each ray. To reduce these costs, recent approaches employ explicit/hybrid data structures and rasterization-based pipelines to accelerate rendering, such as 3D Gaussian Splatting (3DGS) [Kerbl et al. 2023] and Sparse Voxels [Sun et al. 2024]. However, when targeting light field displays, these methods still require rendering many slightly shifted viewpoints, making repeated sampling inefficient and potentially redundant. Addressing this redundancy is essential to enabling real-time radiance field rendering for light field applications.

In this paper, we present a unified framework for real-time radiance field rendering on light field displays. Our framework supports a wide range of explicit radiance field representations—including 3DGS and Sparse Voxels—within a unified rendering architecture. The core of our method is a single-pass of plane sweeping over multi-plane images (MPI) rendered from the input radiance field that enables efficient view synthesis for multiple novel views while largely preserving image quality. We implement this framework in two variants: 3DGS-to-light-field (G2LF) and sparse-voxels-to-light-field (V2LF), corresponding to different input representations. On standard benchmarks, our method achieves up to 22× speedup compared to independently rendering each view, while preserving high image quality. This efficiency comes from rendering and caching shared intermediate planes that are reused across nearby quilt views. The resulting trade-off is that the representation favors view-independent color caching over fine view-dependent effects, while requiring additional memory for the intermediate planes. We further demonstrate the effectiveness of our approach through an interactive application running on a commercial light field display, enabling smooth and immersive real-time 3D visualization of radiance fields. Our OpenGL-based interactive demo achieves 228 FPS (22× faster compared to native quilt rendering, see section 5.3) for 45-view, 512×910 light field images on the BICYCLE scene using a single NVIDIA RTX 5090 graphics card (see supplementary video).

The contributions of this paper are as follows:

- We propose a unified and efficient multi-view rendering framework for real-time rendering of radiance fields on light field displays, supporting multiple explicit (3DGS and Sparse Voxels) representations without retraining them.
- Our method achieves real-time performance (>60 FPS for 90+ views, up to 228 FPS for 45 views) by minimizing redundant sampling through a single-pass plane sweeping strategy.
- We develop an OpenGL-based interactive 3DGS renderer for a commercial light field display, enabling real-time 3D visualization and achieving a 3.6× speedup over our Python/CUDA baseline.

2 Related Work

Radiance field representations. NeRF introduced continuous volumetric scene encoding for novel view synthesis [Mildenhall et al. 2021], prompting numerous efforts to accelerate rendering. Techniques for it include hash-based feature grids [Müller et al. 2022], space decomposition [Chen et al. 2022a; Reiser et al. 2021], and hierarchical training frameworks [Barron et al. 2022; Tancik et al. 2023]. More recently, explicit and hybrid representations such as 3DGS [Kerbl et al. 2023] and voxel-based rasterization [Sun et al. 2024, 2022; Takikawa et al. 2021] have enabled real-time rendering through rasterization-friendly pipelines. While these methods reduce the cost of rendering a single view, they do not directly address the inefficiency of rendering many closely related views,

as required by multi-view displays and remain prohibitively expensive for rendering many views simultaneously.

Depth-guided view synthesis. Classical methods such as depth-image-based rendering (DIBR) [Fehn 2004], layered depth images (LDI) [Broxton et al. 2020; Shade et al. 1998], multi-plane images (MPI) [Zhou et al. 2018], and related homography-based view-synthesis methods synthesize novel views by projecting input images, often with depth information, onto proxy geometry or discrete depth layers. Wizardwongsa et al. propose a fast MPI method with a custom data representation that separates training and render path for specific MPIs containing view-dependent effects from those providing diffuse information [Wizardwongsa et al. 2021]. While our method shares structural similarities with these approaches, it differs in both construction and use: rather than estimating a custom layered representation tailored to a specific display from input images, we render MPIs for dense light-field view synthesis directly from an already reconstructed 3D radiance field. This architecture enables broad compatibility for any explicit and increasingly powerful radiance field representation and allows efficiently sharing rendered information across many closely spaced light-field views, allowing each view to be generated by inexpensive plane reprojection and accumulation instead of repeatedly rendering the original radiance field.

Multi-view rendering. Techniques such as multi-view point splatting [Hübner et al. 2006], single-pass multi-view rendering [Hübner et al. 2007], and unstructured lumigraph rendering [Buehler et al. 2001], dynamically reparameterized light fields on focal surfaces [Isaksen et al. 2000], as well as hardware-level instancing methods [NVIDIA Corporation 2025; Unterguggenberger et al. 2020], reduce rendering cost across nearby views. However, these methods are tailored to traditional 3D content, such as meshes or point clouds, and are difficult to apply to radiance fields, where sharing computation across nearby views remains challenging. Our method addresses this by providing a unified solution that supports both implicit and explicit representations, enabling single-pass rendering of high-quality multi-view outputs.

Adaptive projection tailored to the light field display type can reduce the pixel waste resulting from interleaving for all rasterized content but has been shown to introduce subsampling artifacts without custom anti-aliasing strategies [Fink et al. 2023]. This line of optimization is orthogonal to our strategy and could potentially be coupled with our method for further performance optimization.

Multi-view radiance field rendering. Rendering multiple views from radiance fields is computationally intensive due to high redundancy between adjacent viewpoints. Existing methods typically render each view independently, leading to inefficiencies in multi-view or light field applications [Rabia et al. 2024; Stengel et al. 2023; Tran et al. 2024]. Volume rendering with precomputed ray-to-subpixel mappings [Chen et al. 2022b; Ji et al. 2025; Kim et al. 2026] reduces this overhead but requires per-device calibration for accurate rendering. DirectL [Yang et al. 2024] also targets efficient light-field radiance-field rendering through calibrated ray mappings, but relies on ray-tracing-style sampling, which is less suited to rasterization-based explicit representations. In contrast, our approach is designed for rasterization-based explicit radiance fields: we rasterize each primitive once into shared MPI-like sweeping planes and reuse the cached planes for dense quilt generation. This minimizes redundant sampling across slightly shifted views while providing a unified pipeline for unmodified explicit radiance-field representations.

3 Light Field 3D Display

Multi-view displays are designed to deliver binocular disparities at different viewpoints, providing motion parallax cues for viewer(s). The light rays from each microlens converge to specific viewpoints at the viewing distance (Fig. 2, left). Such view separation is commonly achieved through

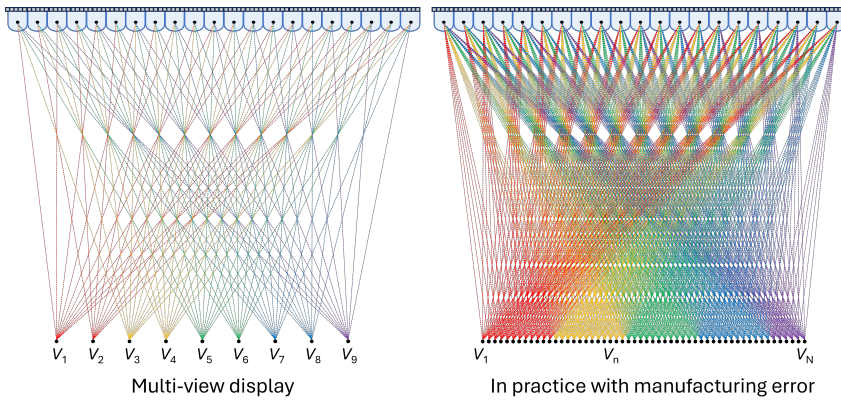


Fig. 2. Ray diagrams of light field displays as viewed from the top. Left: an ideal imperfection-free multi-view display (9-view); Right: a multi-view display with a small manufacturing misalignment (0.00884°) in the lens array. Even with the same nominal 9-view layout, such a small misalignment changes the ray-to-subpixel mapping and breaks the intended view separation. Because calibrated ray directions can vary across panel pixels, more rendered views may be required to densely sample the corrected view mapping. In practice, these imperfections necessitate per-device calibration, which can further increase rendering cost. Here, a 1080p liquid crystal display paired with a non-slanted lens array is assumed for simplicity.

optical elements, including parallax barriers [Lanman et al. 2010], lenticular lenses [Ives 1931], or slanted lenticular lenses [Van Berkel 1999], attached to the display panel. However, each of these different rendered views that are delivered to the viewer via a light field display must all be rendered from the original 3D scene and provided together to the display. The rendering process for these multiple views involves placing perspective cameras in parallel at the designated viewing distance from the 3D scene and rendering from them. After rendering all views, the rendered perspective images are arranged into a 2D view atlas, often called a “quilt” in light field display systems. The quilt is then interleaved into the final panel image according to the display’s optical configuration.

Recent commercially available light field 3D displays often increase angular resolution to achieve smoother motion parallax (i.e., they increase the number of viewpoints). However, this design choice imposes a significant computational burden as significantly more views must be rendered. Additionally, manufacturing imperfections, such as angular and spatial misalignments between the lens array and display panel, further increase computational cost. Even a small angular misalignment, such as a half-subpixel offset (only 0.00884° for a 1080p LCD), can scramble all viewpoints (Fig. 2, right) [Kim et al. 2015]. Correcting these errors requires a per-device calibration process to estimate accurate view-to-subpixel mappings, ensuring geometrically correct 3D reconstruction. Although effective, this calibration further increases the number of rendered views and amplifies computational demands. As a result, modern horizontal motion-parallax light field displays often require 45 or more rendered views, posing significant challenges for real-time, interactive 3D rendering, particularly when dealing with radiance field content.

Recent multi-view rendering methods reduce computation by leveraging ray-to-subpixel mappings for light field displays, but still rely on per-device calibration [Chen et al. 2022b; Ji et al. 2025; Yang et al. 2024]. To address this limitation, we propose a unified framework for radiance field rendering on light field displays. Our method reduces computational redundancy by minimizing

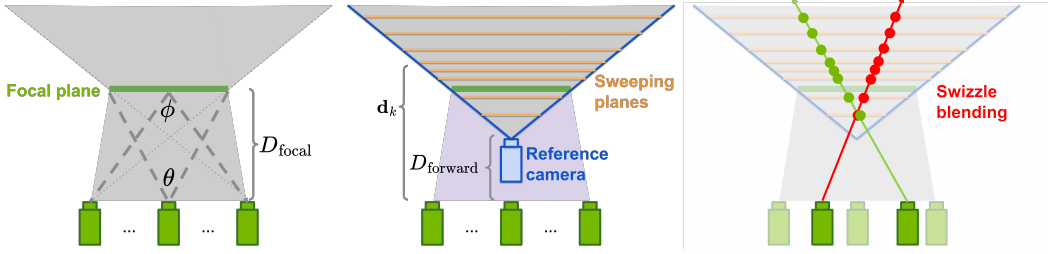


Fig. 3. An illustration of the 3D display parameterization and the sweeping planes. Left: the focal-plane setup, where θ is the base camera's field of view and ϕ is the 3D display's viewing angle. Center: the reference camera is positioned to cover the entire visible volume behind the focal plane, and the sweeping planes are placed at different depths. The purple region denotes a near-field volume excluded from the reference camera frustum, which can be reduced in practice by shifting the reference camera backward by D_{shift} . Right: after rasterizing scene primitives into the sweeping planes, swizzle blending samples the corresponding plane locations for each quilt-view ray and alpha-composites the sampled colors and transmittances into the final light-field quilt. The sampling coordinate is defined in Eq. (6).

repeated sampling across minimally shifted viewpoints, while maintaining overall perceived image quality, and is compatible with various radiance field representations. Detailed algorithmic descriptions follow in the next section.

4 Radiance Fields to Light Field

We first give an overview of our efficient radiance fields to light fields rendering pipeline in Sec. 4.1, and introduce the shared components in Sec. 4.2 and Sec. 4.3. Later, we adapt our algorithm to various state-of-the-art efficient 3D representations commonly used in radiance field reconstruction and novel-view synthesis. The representation-specific adaptations are described in Sec. 4.4 and 4.5.

4.1 Rendering pipeline overview

Given a radiance field and a 3D display viewing setup, our goal is to render a set of $V = (V_x \times V_y)$ perspective viewpoints derived from this setup, called “light field quilts” as illustrated in Fig. 2. A conventional method applies a radiance field renderer V times, separately. However, such a simple approach can slow down rendering up to V times, which prevents interactive experiences even for a 200 FPS single-frame renderer for a common $V=45$ light field display.

Our strategy is to reduce computation across the multiple rendered views by approximating the volume visible to the 3D display by a series of forward-sweeping planes. Each plane represents a disjoint frustum of the original volume. We can then composite the sweeping planes into the V quilt views via an operation that we call *swizzle* blending, instead of repeatedly traversing through the original 3D representation to render each of the multiple views, separately. A single pixel lookup on the planes corresponds to the rendering result of a ray segment from the original radiance field, where the latter is significantly more expensive to compute. Thus the overall rendering time can be largely reduced as long as we can also render the sweeping planes efficiently. Next, we introduce the algorithm for sweeping planes in Sec. 4.2 and *swizzle* blending in Sec. 4.3.

4.2 Sweeping planes rendering

Let's first define the volume of interest of a 3D display setup, which is illustrated in Fig. 3's left panel. We can imagine the 3D display as a “window” (called the focal plane) for viewers to look into the virtual 3D world. The center location of the focal plane is defined by a *base* camera and a

camera-to-plane distance D_{focal} . The size of the focal plane is derived from the base camera's field of view θ_x, θ_y . The viewing angles of the 3D display are defined by ϕ_x, ϕ_y .

To create a series of sweeping planes, we use a perspective *reference* camera that is adjusted to cover the maximum viewing angle, as shown in Fig. 3's center panel. For the horizontal x and the vertical y axes, we compute a forward shift D_{forward} along the z -axis required to cover the display's visible volume and take the maximum over the two axes:

$$D_{\text{forward}} = \max_{k \in \{x, y\}} D_{\text{focal}} \cdot \frac{\tan(0.5\phi_k)}{\tan(0.5\phi_k) + \tan(0.5\theta_k)}, \quad (1)$$

and set the field of view of the reference camera (blue camera in Fig. 3) as:

$$\theta'_{k \in \{x, y\}} = 2 \cdot \arctan\left(\frac{D_{\text{focal}} \cdot \tan(0.5\theta_k)}{D_{\text{focal}} - D_{\text{forward}}}\right). \quad (2)$$

By construction, the reference camera covers the visible volume behind the focal plane. However, some near-field regions between the outermost quilt-view cameras (V) and the focal plane are excluded (purple region in Fig. 3). As a simple workaround, we introduce a hyperparameter D_{shift} that shifts the reference camera backward, i.e., away from the focal plane, such that $D_{\text{forward}} \leftarrow D_{\text{forward}} - D_{\text{shift}}$. This adjustment expands the covered region toward the near-field and reduces the amount of excluded volume in practice.

Finally, the sweeping planes are generated by rendering the scene primitives' chunks or volume chunks into the perspective reference camera. Intuitively, a *chunk* in our context implies the 3D volume between two discrete sweeping plane locations. The rendered sweeping planes are denoted by $\mathbf{C} \in \mathbb{R}^{N_{\text{chunk}} \times (P_s \cdot N_y) \times (P_s \cdot N_x) \times 3}$ for RGB colors and $\mathbf{T} \in \mathbb{R}^{N_{\text{chunk}} \times (P_s \cdot N_y) \times (P_s \cdot N_x)}$ for transmittances, where N_{chunk} is the total number of chunks, $N_x \times N_y$ is the resolution of a single view of the quilt, and P_s is a resolution scaling hyperparameter. We detail how we divide the 3D volume into discrete disjoint chunks and render them for different 3D volumetric representations in Secs. 4.4 and 4.5.

4.3 Swizzle blending

We can now efficiently render the final quilt for a light field display by sub-sampling and alpha composition of these rendered sweeping planes. For clarity, we explicitly parameterize each quilt view by its lateral camera offset and the corresponding principal-point shift before deriving the plane-sampling coordinate. The quilt $\mathbf{Q} \in \mathbb{R}^{V_y \times V_x \times N_y \times N_x \times 3}$ is a 2D array of perspective views formed by laterally shifting the base camera along its horizontal (x) and vertical (y) directions at the same viewing distance, as illustrated in Fig. 3. The lateral camera offsets ($\Delta x, \Delta y$) are distributed uniformly across the viewing window at the viewing distance, and their normalized principal-point shifts (p_x, p_y) are set such that the principal rays from all quilt views converge at the center of the focal plane (see Fig. 3).

Let each quilt view be indexed by row $i \in \{1, \dots, V_y\}$ and column $j \in \{1, \dots, V_x\}$. We define the normalized horizontal and vertical positions of each quilt view within the view grid as

$$s_j^x = \frac{j-1}{V_x-1} - \frac{1}{2}, \quad s_i^y = \frac{i-1}{V_y-1} - \frac{1}{2}. \quad (3)$$

Given the display viewing angles ϕ_x and ϕ_y , the corresponding camera's lateral offsets are:

$$\Delta x_j = 2D_{\text{focal}} s_j^x \tan(0.5\phi_x), \quad \Delta y_i = 2D_{\text{focal}} s_i^y \tan(0.5\phi_y), \quad (4)$$

and their normalized principal-point shifts (p_x, p_y) are:

$$p_j^x = \frac{2s_j^x \tan(0.5\phi_x)}{\tan(0.5\theta_x)}, \quad p_i^y = \frac{2s_i^y \tan(0.5\phi_y)}{\tan(0.5\theta_y)}. \quad (5)$$

Given a normalized horizontal pixel coordinate $u \in [-1, 1]$ in quilt view (i, j) , its projected horizontal coordinate u' on the k -th sweeping plane at distance d_k is

$$u' = \frac{\Delta x_j (D_{\text{focal}} - d_k) / D_{\text{focal}} + d_k \tan(0.5\theta_x) u}{(d_k - D_{\text{forward}}) \tan(0.5\theta'_x)}. \quad (6)$$

where θ'_x denotes the modified horizontal field of view of the reference camera, as defined in Eq. (2). Eq. (6) maps a pixel in quilt view (i, j) to the corresponding sampling location on the k -th sweeping plane in the reference-camera image. In the numerator, the first term accounts for the horizontal shift of quilt view (i, j) due to the camera offset Δx_j , while the second term corresponds to the local pixel offset within that view. We show the horizontal projection for brevity; the vertical coordinate is computed analogously using Δy_i , θ_y , and θ'_y .

We use Eq. (6) to project each quilt-view pixel onto the sweeping planes and sample C and T using either bilinear or nearest-neighbor interpolation from them. The sampled series of colors c_k and transmittance values T_m are blended into a final pixel color with:

$$C = \sum_{k=1}^{N_{\text{chunk}}} \left(\prod_{m=1}^{k-1} T_m \right) c_k. \quad (7)$$

Note that the color c_k here is already weighted by alpha opacity.

In the following, we introduce the representation-specific adaptations of our algorithm.

4.4 G2LF – 3D Gaussians to Light Field

One advantage of a primitive-based 3D volumetric representation, e.g., 3DGS which comprises of a discrete collection of 3D Gaussian primitives, is that we can have a well-distributed rendering computational cost for each volume chunk (Fig. 3's right panel) by balancing sweeping plane positions. Specifically, we apply quantile binning so each volume chunk has similar numbers of primitives. For applying quantile binning with 3DGS, we select Gaussian primitives inside the view frustum of the reference camera and compute a sorted list of the Gaussians according to the reference camera distances. The sorted list is partitioned into N_{chunk} chunks, each with approximately the same number of elements. The plane distance $\mathbf{d}_k$ of the k -th sweeping plane is set to the median Gaussian distance of the k -th chunk.

A Gaussian may overlap neighboring chunks, while our implementation assigns it to a single chunk based on its center depth in the reference camera. Since sorting, chunking, and rasterization all use the same reference camera, this approximation mainly contributes to the discretization error of the sweeping-plane representation rather than introducing an additional ordering ambiguity. In practice, this error is most visible as stack-of-cards artifacts when the chunk resolution is too coarse, and can be reduced by increasing N_{chunk} .

We extend the efficient CUDA rasterizer by [Kerbl et al. 2023] to perform 3D tiling instead of the original 2D tiling with an additional dimension for the chunks. A Gaussian is assigned to a tile by its patch index and chunk index. The Gaussians in each 3D tile are sorted and rendered in parallel, which finally produces the color and transmittance forward-sweeping planes $\mathbf{C}$ and $\mathbf{T}$. A visualization of this rendering pipeline and its pseudocode are shown in Fig. 4 and Algo. 1, respectively.

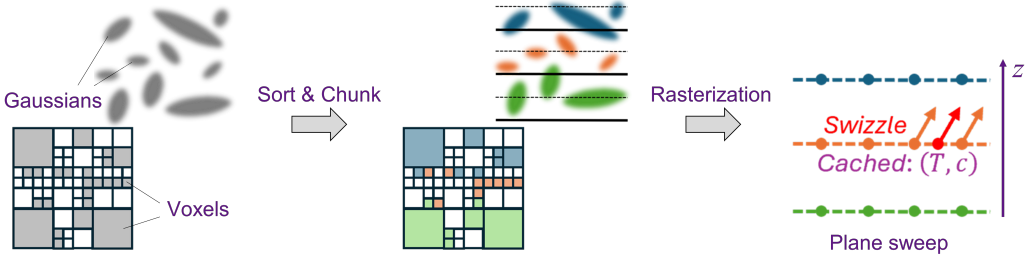


Fig. 4. Illustration of the G2LF and V2LF pipeline. Primitives are sorted along the z-axis of the reference camera and partitioned into depth chunks with similar numbers of primitives. Each chunk is rasterized onto its representative sweeping plane, producing cached color and transmittance planes, which are then composited by *swizzle* blending to form the final quilt.

The original 3DGS implementation uses a hard-coded antialiasing filter in pixel-size unit. Specifically, the variance of the projected 2D Gaussian on the screen space is always dilated by 0.3 pixels. This causes a mismatch in the effective screen-space filter size between conventional 3DGS rendering and our sweeping-plane rendering when the plane resolution scale P_s changes. (Sec. 4.3) Intuitively, P_s alters the resolution/size of each rendered sweeping-plane, which in turn, affects the overall memory utilization of our MPI representation and its overall rendering time. Our solution is to align the filtering strength f based on the pixel size on the focal plane. Specifically, we use an adaptive filtering strength to align our rendering with the conventional rendering by:

$$f' = f \cdot \left(\frac{D_{\text{focal}} \cdot \tan(0.5\theta_x)}{(D_{\text{focal}} - D_{\text{forward}}) \cdot \tan(0.5\theta'_x)} \cdot P_s \right)^2, \quad (8)$$

where $f = 0.3$ is the hard-coded number in the original 3DGS implementation. This equation yields a larger filter strength (in screen space of the reference camera) when the pixel size of the reference camera on the focal plane is smaller than the conventional one.

4.5 V2LF – Sparse Voxels to Light Field

The rendering of Sparse Voxels [Sun et al. 2024] follows the same overall principle as G2LF, since sparse voxels also form a discrete set of primitives. We also extend the CUDA-based sparse voxel rasterizer (SVR) [Sun et al. 2024] in a similar way to how we extended 3DGS’s rasterizer. To handle antialiasing, SVR employs supersampling instead of pre-filtering so our filtering strength calibration Eq. (8) is not applicable to SVR. However, supersampling is still too costly with a large number of chunks (N_{chunk}). As an alternative, we disable supersampling in SVR and apply a low-pass Gaussian filter on the sweeping planes instead, which is implemented in CUDA and performs filtering in-place without allocating extra memory.

5 Results

5.1 Evaluation

Dataset. We use the four indoor and five outdoor unbounded real-world scenes from the Mip-NeRF360 dataset for evaluation. For training we use all the training set cameras and then sample four viewpoints from the test set for evaluating the rendering quality.

Baseline models reproduction. We used 3DGS [Kerbl et al. 2023] and SVR [Sun et al. 2024] as the baseline models. We directly train with their default hyperparameters on the Mip-NeRF360 dataset. The standard perspective novel-view results are (LPIPS 0.215, PSNR 27.53, SSIM 0.816) for 3DGS

Algorithm 1 G2LF/V2LF P - Input Gaussian or Voxel primitives H - Quilt and camera parameters Q, T - Output light field quilt RGB and transmittance buffer

```

1:  $\mathbf{d}' \leftarrow \text{CulledDepth}(P)$  ▷ GS/Voxels z distances to camera
2:  $\mathbf{d} \leftarrow \text{FindQuantile}(\mathbf{d}')$  ▷  $N_{\text{chunk}}$  plane distances
3:  $\{P'_k\} \leftarrow \text{Sort\_and\_Chunk}(P, \mathbf{d})$  ▷ Sort & Chunk GS/Voxels
4: Parallel for all  $k$  do ▷ Rasterize chunks to midplane
5:    $\mathbf{T}_k, \mathbf{C}_k \leftarrow \text{Rasterize}(H, P'_k)$  ▷ Cache grid values
6:  $Q, T \leftarrow \text{Initialize\_buffers}()$ 
7: for  $k$  in  $1 \dots N_{\text{chunk}}$  do ▷ Swizzle blending
8:    $U \leftarrow \text{Quilt2PlaneCoordinate}(H, \mathbf{d}_k)$  ▷ Eq. (6)
9:    $c \leftarrow \text{interpolate}(U, \mathbf{C}_k)$ 
10:   $t \leftarrow \text{interpolate}(U, \mathbf{T}_k)$ 
11:   $Q, T \leftarrow (Q + T \odot c), (T \odot t)$  ▷ Eq. (7), Volume rendering
12: end for

```

Table 1. Light field quilt rendering results comparison on the Mip-NeRF360 dataset using various algorithms and their ablations. The results are averaged on 9 scenes where we sample 4 viewpoints to render light field quilts for each scene (quilt resolution set to $V_x=45, V_y=1, N_x=N_y=512$). 3DGS [Kerbl et al. 2023] and SVR [Sun et al. 2024] are used for the baselines whose rendered quilts are served as the ground truth for G2LF and V2LF, respectively. The FPS indicate the rendering speed of the entire quilts measured on a NVIDIA RTX 5090, and memory usage (MEM) is reported in GB.

Method	FPS↑	MEM↓	LPIPS↓	PSNR↑	SSIM↑
3DGS	5.9	5.8	serve as ground truth		
G2LF ($N_{\text{chunk}}=64$)	127	6.3	0.311	24.89	0.749
G2LF ($N_{\text{chunk}}=64, P_s=2$)	109	7.5	0.129	26.96	0.820
G2LF ($N_{\text{chunk}}=512$)	50	8.5	0.277	26.89	0.831
G2LF ($N_{\text{chunk}}=512, \mathbf{Bi.}$)	24	8.5	0.156	28.78	0.875
G2LF ($N_{\text{chunk}}=512, P_s=2$)	34	16.2	0.084	30.71	0.924
G2LF ($N_{\text{chunk}}=512, P_s=2, \mathbf{Bi.}$)	19	16.2	0.065	33.31	0.950
G2LF ($N_{\text{chunk}}=512, P_s=3, \mathbf{Bi.}$)	15	29.0	0.053	34.77	0.958
SVR	8.2	4.0	serve as ground truth		
V2LF ($N_{\text{chunk}}=64$)	133	4.4	0.344	23.63	0.673
V2LF ($N_{\text{chunk}}=64, P_s=2$)	104	4.9	0.153	26.97	0.793
V2LF ($N_{\text{chunk}}=512$)	48	5.7	0.302	25.57	0.767
V2LF ($N_{\text{chunk}}=512, \mathbf{Bi.}$)	26	5.7	0.182	28.51	0.845
V2LF ($N_{\text{chunk}}=512, P_s=2$)	26	10.2	0.099	31.06	0.908
V2LF ($N_{\text{chunk}}=512, P_s=2, \mathbf{Bi.}$)	17	10.2	0.072	33.87	0.938
V2LF ($N_{\text{chunk}}=512, P_s=3, \mathbf{Bi.}$)	11	17.8	0.055	35.54	0.950

and (LPIPS 0.186, PSNR 27.36, SSIM 0.822) for SVR averaged across all scenes. The proposed G2LF and V2LF algorithms are evaluated against these trained baseline models.

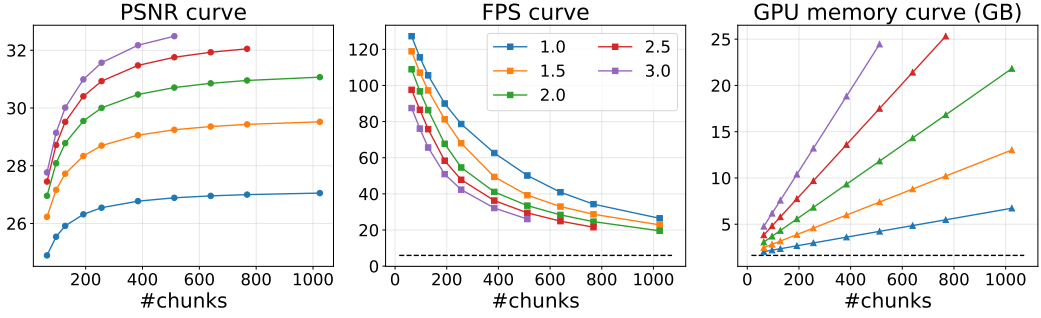


Fig. 5. Performance curves of G2LF under various N_{chunk} (x-axis) and P_s (colors) values. Some curves have fewer data points due to VRAM limits at those configurations. In comparison, the black dash lines are the FPS and required memory for conventional 3DGS rendering. More planes and higher plane resolution both consistently improve quality at the cost of increasing compute and memory usage. The quality (PSNR) improvement is about saturated at around $N_{\text{chunk}}=400$. Higher P_s values keep on improving the quality, but at the cost of greater memory usage.

Results on light field quilt. In Fig. 1 (left), the render FPS obtained for varying number of views is presented using different algorithms. Our G2LF and V2LF algorithms (solid lines) achieve real-time performance (>60 FPS) even for over 90 views, enabling interactive 3D visualizations of scenes on light field displays.

Table 1 presents a quantitative comparison of light field quilts ($V_x=45, V_y=1, N_x=N_y=512$) rendered using various algorithms and their ablative variants. As we do not have the ground truth images for the quilts, we use the rendering results for the quilts from the conventional method which individually renders multiple views to serve as the ground truth for G2LF and V2LF. These results demonstrate that G2LF and V2LF can faithfully render light field quilts while substantially accelerating rendering compared with their conventional counterparts, 3DGS and SVR. The number of chunks N_{chunk} , the sweeping-plane resolution scale P_s , and the interpolation method (nearest-neighbor by default; ‘Bi.’ for bilinear) are the main parameters controlling the speed-quality trade-off. For example, with $N_{\text{chunk}} = 64$, G2LF and V2LF achieve 127 FPS and 133 FPS, respectively, compared with 5.9 FPS and 8.2 FPS for conventional 3DGS and SVR quilt rendering, while maintaining reasonable image fidelity.

Increasing the depth resolution (N_{chunk}) and plane resolution scale (P_s) brings the rendered quilts closer to the baseline quality, at the cost of higher computation and memory usage. As shown in Fig. 5, both parameters consistently improve reconstruction quality, while Fig. 12 confirms that the resulting perspective views remain visually accurate in real-time. In particular, increasing P_s reduces interpolation errors during the swizzle operation. Performance curves in Fig. 5 show many ablative variations of G2LF with different number of chunks and plane resolution scales, both of which consistently mitigate image-quality differences from the slower conventional rendering approach. We also find that using 8-bit unsigned integer values to store C and T leads to similar blending quality compared to using 32-bit floating-point values (see Fig. 7). We therefore use 8 bits by default, which substantially reduces memory usage and improves rendering speed.

We show comprehensive ablation studies of the various individual components of G2LF and discuss their results in the captions—ablating spherical harmonic in Fig. 6, comparing uint8 and fp32 in Fig. 7, showing the effect of the adaptive anti-aliasing filter in Fig. 8, comparing different interpolation algorithms in Fig. 9, and different quilts setup in Fig. 10.

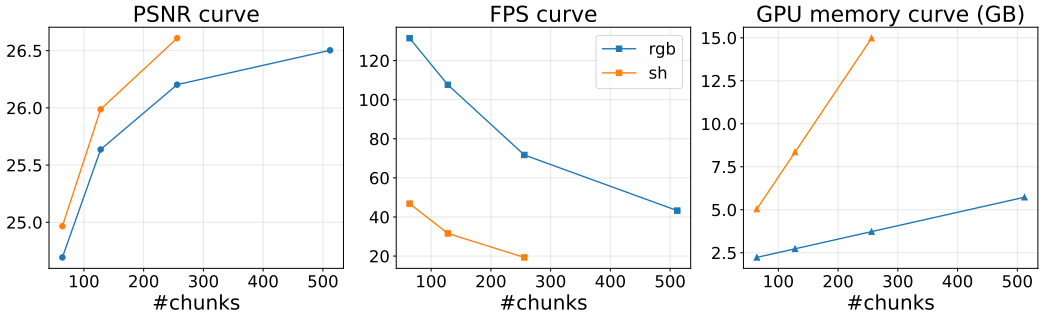


Fig. 6. Comparison of *swizzle* blending with RGB color or spherical harmonic (SH) for more accurate view-dependent colors. The quality improvement (PSNR) is relatively minor compared to using more chunks or higher plane resolution (Fig. 5). However, with spherical harmonics the drop in FPS and the increase in memory usage is significant. We therefore use RGB-alpha *swizzle* blending by default, accepting reduced view-dependent accuracy for off-center views.

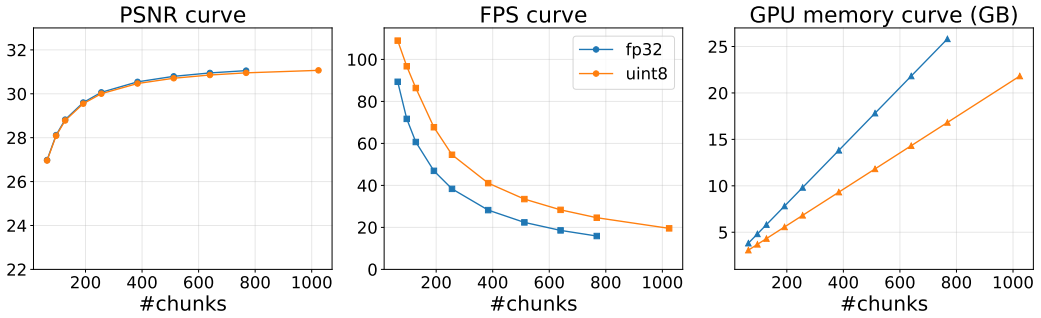


Fig. 7. Comparison of using uint8 and fp32 to store the sweeping planes of G2LF ($P_s=2$). The PSNR is almost identical while the run time and memory usage is largely reduced with uint8. Switching to uint8 also results in being able to use a higher number of chunks without running out of memory, versus fp32.

5.2 Captured Results

We captured the displayed light field results for several algorithm variants using a Looking Glass Go display [Looking Glass Factory Inc. 2026] and through a Google Pixel 9 Pro (f/2.8, 1/50s, ISO 100) [Google Inc. 2026], as shown in Fig. 14. The display was mounted on a motorized rotational stage, and viewpoints were uniformly sampled across the horizontal range (-18.5° to $+17.5^\circ$) at 1.5° intervals. The brightness of the captured images was adjusted linearly for visualization. The video results confirm that our method enables high-quality 3D reconstruction with accurate angular consistency across views (see supplementary video). In practice, artifacts visible in individual rendered views are less noticeable during actual use, as multiple views blend perceptually within the viewer's pupil. See our supplementary materials for more captured results.

5.3 Real-Time Interactive 3D Visualization Demo

We demonstrate the application of our method for dynamic, interactive 3D visualization on a light field display [Kim et al. 2025]. Fig. 11 showcases the demo running on a desktop equipped with a single NVIDIA RTX 5090 and a Looking Glass 16" Light Field Display. In comparison to a traditional

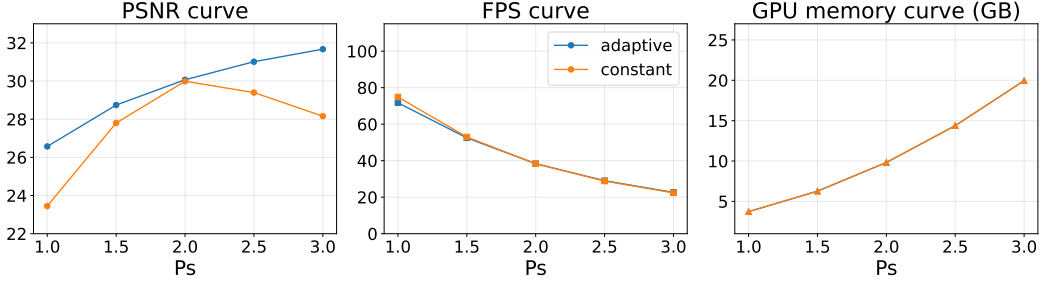


Fig. 8. Ablation of adaptive versus constant anti-aliasing filter strengths in G2LF ($N_{\text{chunk}} = 256$). Using a constant filter strength leads to a screen-space filter mismatch between conventional 3DGS rendering and our sweeping-plane rendering when the plane resolution scale P_s changes. In contrast, the adaptive filter maintains consistent quality improvements with increasing P_s , while having nearly identical FPS and memory usage.

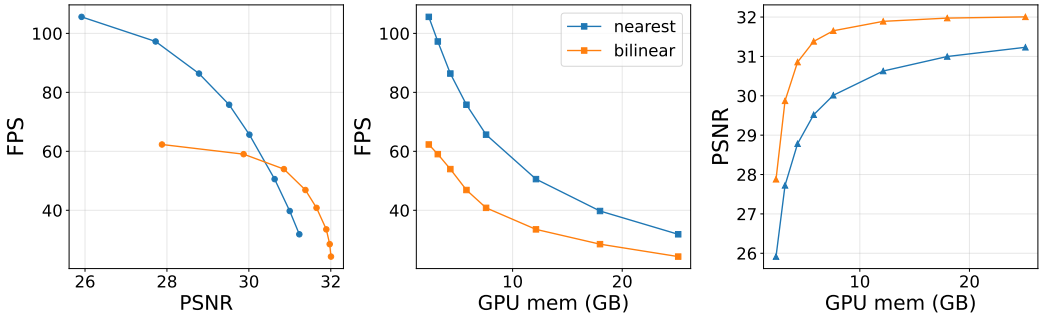


Fig. 9. Comparing bilinear and nearest-neighbor interpolation with *swizzle* blending with $N_{\text{chunk}}=64$ and $P_s \in [1, 6]$. As expected, bilinear interpolation achieves better quality but with slower FPS under the same amount of GPU memory usage comparing to nearest-neighbor interpolation. There is a crossover point in left figure at around 30.5 PSNR. To render in lower quality, nearest-neighbor is faster while bilinear becomes faster for achieving higher PSNR. This is because nearest-neighbor interpolation needs higher P_s to achieve the same quality as bilinear, which becomes the dominant factor for computation cost with high P_s . Theoretically, both versions should achieve the same PSNR with infinite P_s .

2D display, our real-time G2LF implementation allows users to perceive depth directly through binocular cues, providing a more intuitive and immersive interaction with 3DGS content. Our interactive application is implemented in OpenGL/C++ to make use of texture filtering and blending in hardware during the *swizzle* operation. We use 8-bit RGBA framebuffers for both sweeping plane buffers and the 45-views 4K x 4K quilt. Our OpenGL implementation mostly follows our description in Alg. 1. One difference is that we switch the order of quantile computation and sorting to sort all primitives in one pass upfront. From the rendered quilt we apply interlacing provided by the LookingGlass Bridge SDK before sending the frame to the swapchain.

Our optimized OpenGL implementation achieves comparable image quality to the Python/CUDA implementation presented in Tab. 1, but achieves better computational performance due to higher GPU utilization. Compared to the Python/CUDA baseline, the speedup is approximately $3.6\times$ on average across tested scenes ($N_{\text{chunk}}=64, P_s=1$, measured on NVIDIA RTX 5090). Our GL

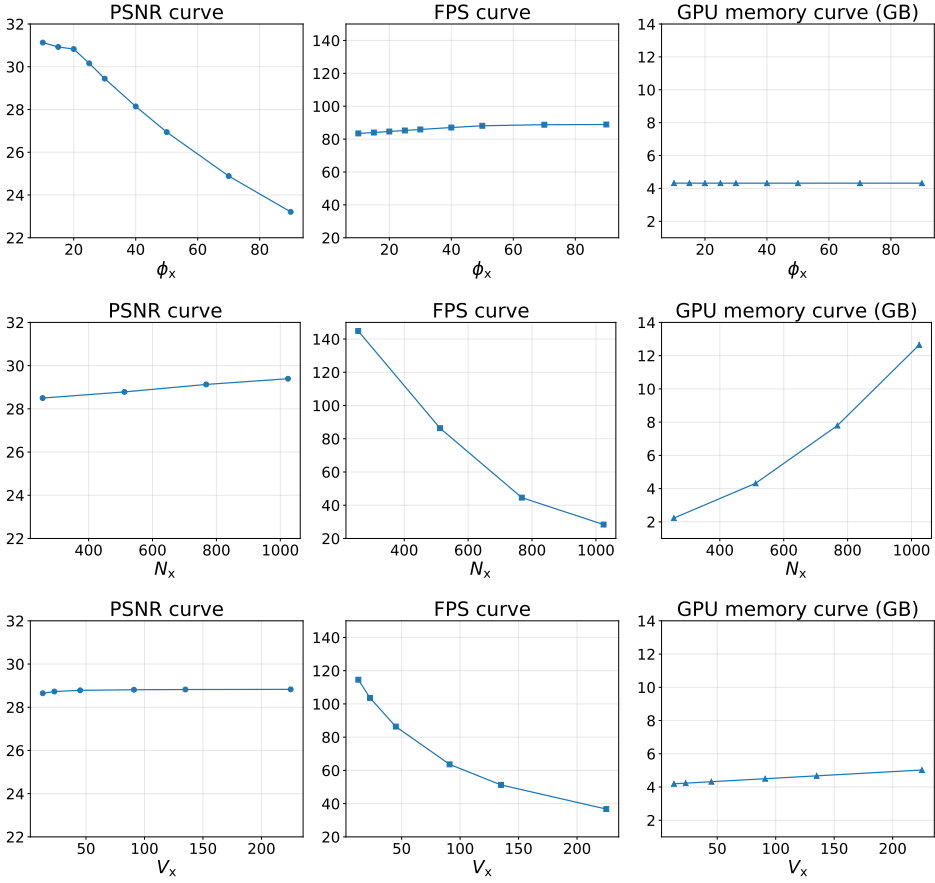


Fig. 10. Results of G2LF under different quilt settings. The default setting is $P_s=2$, $N_{\text{chunk}}=128$, $\phi_x=35$, $\phi_y=0$, $N_x=N_y=512$, $V_x=45$, $V_y=1$, with nearest-neighbor interpolation and uint8 sweeping planes. Top: increasing the viewing angle slightly reduces quality, while FPS and memory usage remain nearly unchanged. Middle: increasing the per-view quilt resolution increases FPS cost and memory usage approximately proportionally, while the quality improvement remains modest. Bottom: increasing the number of quilt views raises runtime and memory usage, while PSNR remains largely similar within the same viewing range.

implementation scales well across varying number of chunks. In our tests $N_{\text{chunk}}=64$ to $N_{\text{chunk}}=128$ provide a sweet spot in terms of performance. Although higher numbers of chunks result in slightly higher image quality metrics the quality increase diminishes when perceived on the light field display (see supplementary video). Lower chunk numbers ($N_{\text{chunk}} < 32$) can result in noticeable discrete depth layers for surfaces close to the camera ("stack-of-cards effect" as shown in Fig. 13) and should be avoided.

6 Discussion

Framework Versatility. Our unified framework supports a wide range of radiance field representations, including all explicit formats like 3D Gaussians and Sparse Voxels, without the need for retraining. Besides static 3D scenes this allows our method to directly support 4D immersive



Fig. 11. Real-time interactive 3D radiance field demonstration on a Looking Glass light field display, enabling users to dynamically change viewpoints and visualize radiance fields in a 3D space [Kim et al. 2025].

content as well. Further, one can apply our rendering algorithm to implicit representations like NeRF variants. To test this, we used the coarse network in NeRF to render a roughly estimated depth map from the base camera view first and quantile the depth points to determine the chunk positions. Despite the effort, we still observe apparent degradation in quality of N2LF comparing to using the conventional quilt rendering with NeRF (LPIPS 0.379, PSNR 23.64, SSIM 0.822 with Nerfacto [Tancik et al. 2023] baseline), mostly because of poor importance sampling. Designing a specific sampling method to address this is out of our current scope and we leave it for future work.

Our algorithm is compatible with all types of light field displays, from 3D displays with head-tracked rendering to integral imaging systems offering vertical parallax. Unlike methods limited to a fixed set of view directions, our slice-based approach enables fast reconstruction from arbitrary viewpoints within the supported viewing angle.

Limitations and Future Work. One of the main limitations of our approach is the increased memory usage due to the intermediate slice-based representation, which is more demanding than direct rendering methods. Future work may explore reducing this memory footprint, for example by employing hardware-accelerated compression or more compact data layouts.

Additionally, blending performance is inherently tied to the number and resolution of slice planes. While more slices enable finer volumetric detail, they also incur higher computational cost. If the number of chunks is too small, the discretized slice representation can introduce stack-of-cards artifacts, particularly in scenes with large depth variation or geometry close to the viewer. This scene-dependent trade-off can be mitigated by increasing the number of chunks at the cost of additional memory and computation.

Our method also trades off view-dependent appearance for rendering efficiency. Because scene appearance is baked into shared sweeping planes, strongly view-dependent effects such as specular highlights and reflections may be less accurate for off-center views than in per-view rendering. However, our spherical-harmonic ablation shows that preserving view-dependent color provides only minor quality improvement while substantially increasing memory usage and reducing rendering speed (Fig. 6). We therefore use RGB sweeping planes by default, which offers a better practical trade-off for real-time light field display rendering.

Since the ultimate output is constrained by the characteristics of physical light field displays (e.g., limited angular or depth resolution), future optimization could benefit from adapting the slice structure to better match the display-specific capabilities and perceptual requirements.

References

- Jonathan T Barron, Ben Mildenhall, Dor Verbin, Pratul P Srinivasan, and Peter Hedman. 2022. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5470–5479.
- Michael J. Broxton, John Flynn, Ryan Overbeck, Daniel Erickson, Peter Hedman, Matthew Duvall, Jason Dourgarian, Jay Busch, Matt Whalen, and Paul Debevec. 2020. Immersive Light Field Video with a Layered Mesh Representation. *ACM Trans. Graph.* 39, 4 (2020), 86:1–86:15. doi:10.1145/3386569.3392485
- Chris Buehler, Michael Bosse, Leonard McMillan, Steven J. Gortler, and Michael F. Cohen. 2001. Unstructured Lumigraph Rendering. In *SIGGRAPH*. ACM, 425–432. doi:10.1145/383259.383309
- Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, and Hao Su. 2022a. TensorRF: Tensorial Radiance Fields. In *ECCV*. 333–350. doi:10.1007/978-3-031-19824-3_21
- Shuo Chen, Binbin Yan, Xinzhu Sang, Duo Chen, Peng Wang, Zeyuan Yang, Xiao Guo, and Chongli Zhong. 2022b. 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.
- Christoph Fehn. 2004. Depth-Image-Based Rendering (DIBR), Compression, and Transmission for a New Approach on 3D-TV. In *SPIE Stereoscopic Displays and Virtual Reality Systems XI*, Vol. 5291. 93–104. doi:10.1117/12.524762
- Laura Fink, Svenja Strobel, Linus Franke, and Marc Stamminger. 2023. Efficient rendering for light field displays using tailored projective mappings. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 6, 1 (2023), 1–17.
- Google Inc. 2026. Google. <https://google.com/> Accessed: March 2026.
- Thomas Hübner, Yanci Zhang, and Renato Pajarola. 2006. Multi-View Point Splatting. In *GRAPHITE*. ACM, 285–294. doi:10.1145/1174429.1174479
- Thomas Hübner, Yanci Zhang, and Renato Pajarola. 2007. Single-Pass Multi-View Rendering. *IADIS Int. J. on Computer Science and Information Systems* 2, 2 (2007), 122–140.
- Aaron Isaksen, Leonard McMillan, and Steven J Gortler. 2000. Dynamically reparameterized light fields. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 297–306.
- Herbert E Ives. 1931. Optical properties of a Lippmann lenticulated sheet. *JOSA* 21, 3 (1931), 171–176.
- 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.
- 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, Chang-Kun Lee, Jong-Young Hong, Changwon Jang, Youngmo Jeong, Jiwoon Yeom, and Byoungho Lee. 2015. Calibration method for the panel-type multi-view display. *Journal of the Optical Society of Korea* 19, 5 (2015), 477–486.
- 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 *Proceedings of the SIGGRAPH Asia 2025 Emerging Technologies*. 1–3.
- Joonsoo Kim, Gugjin Yun, Hyon-Gon Choo, Hyeonju Mun, and Kyeongbo Kong. 2026. 3D Gaussian rasterization for direct rendering of view-multiplexed images on multiview displays. *ETRI Journal* (2026).
- Douglas Lanman, Matthew Hirsch, Yunhee Kim, and Ramesh Raskar. 2010. Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization. *ACM Transactions on Graphics (TOG)* 29, 6 (2010), 1–10.
- Leia Inc. 2025. Leia Official Website. <https://www.leiainc.com/> Accessed: January 2025.
- Looking Glass Factory Inc. 2026. Looking Glass Factory Official Website. <https://lookingglassfactory.com/> Accessed: March 2026.
- Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65, 1 (2021), 99–106.
- Thomas Müller, Alex Evans, Christoph Schied, and Alexander Keller. 2022. Instant neural graphics primitives with a multiresolution hash encoding. *ACM transactions on graphics (TOG)* 41, 4 (2022), 1–15.
- NVIDIA Corporation. 2025. VRWorks - Multi-View Rendering (MVR). <https://developer.nvidia.com/vrworks/graphics/multiview> Accessed: May 2025.
- Sédick Rabia, Guillaume Allain, Rosalie Tremblay, and Simon Thibault. 2024. Orthoscopic elemental image synthesis for 3D light field display using lens design software and real-world captured neural radiance field. *Optics Express* 32, 5 (2024), 7800–7815.

- Christian Reiser, Songyou Peng, Yiyi Liao, and Andreas Geiger. 2021. KiloNeRF: Speeding Up Neural Radiance Fields with Thousands of Tiny MLPs. In *ICCV*. 14335–14345. doi:10.1109/ICCV48922.2021.01408
- Jonathan Shade, Steven Gortler, Li-wei He, and Richard Szeliski. 1998. Layered depth images. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. 231–242.
- Sony Electronics Inc. 2025. Spatial Reality Display by Sony. https://pro.sony/ue_US/products/spatial-reality-displays/3d-professional-images Accessed: January 2025.
- Michael Stengel, Koki Nagano, Chao Liu, Matthew Chan, Alex Trevithick, Shalini De Mello, Jonghyun Kim, and David Luebke. 2023. AI-mediated 3D video conferencing. In *ACM SIGGRAPH 2023 Emerging Technologies*. 1–2.
- Cheng Sun, Jaesung Choe, Charles Loop, Wei-Chiu Ma, and Yu-Chiang Frank Wang. 2024. Sparse Voxels Rasterization: Real-time High-fidelity Radiance Field Rendering. *arXiv preprint arXiv:2412.04459* (2024).
- Cheng Sun, Min Sun, and Hwann-Tzong Chen. 2022. Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5459–5469.
- Towaki Takikawa, Joey Litalien, Kangxue Yin, Karsten Kreis, Charles Loop, Derek Nowrouzezahrai, Alec Jacobson, Morgan McGuire, and Sanja Fidler. 2021. Neural geometric level of detail: Real-time rendering with implicit 3d shapes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11358–11367.
- Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, et al. 2023. Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*. 1–12.
- Phong Tran, Egor Zakharov, Long-Nhat Ho, Anh Tuan Tran, Liwen Hu, and Hao Li. 2024. VODOO 3D: Volumetric Portrait Disentanglement for One-Shot 3D Head Reenactment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10336–10348.
- Johannes Unterwiesingberger, Bernhard Kerbl, Markus Steinberger, Dieter Schmalstieg, and Michael Wimmer. 2020. Fast Multi-View Rendering for Real-Time Applications.. In *EGPGV@ Eurographics/EuroVis*. 13–23.
- Cees Van Berkel. 1999. Image preparation for 3D LCD. In *Stereoscopic Displays and Virtual Reality Systems VI*, Vol. 3639. SPIE, 84–91.
- Suttisak Wizadwongsa, Pakkapon Phongthawee, Jiraphon Yenphraphai, and Supasorn Suwajanakorn. 2021. Nex: Real-time view synthesis with neural basis expansion. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8534–8543.
- 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.
- Tinghui Zhou, Richard Tucker, John Flynn, Graham Fyfe, and Noah Snavely. 2018. Stereo magnification: learning view synthesis using multiplane images. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–12.

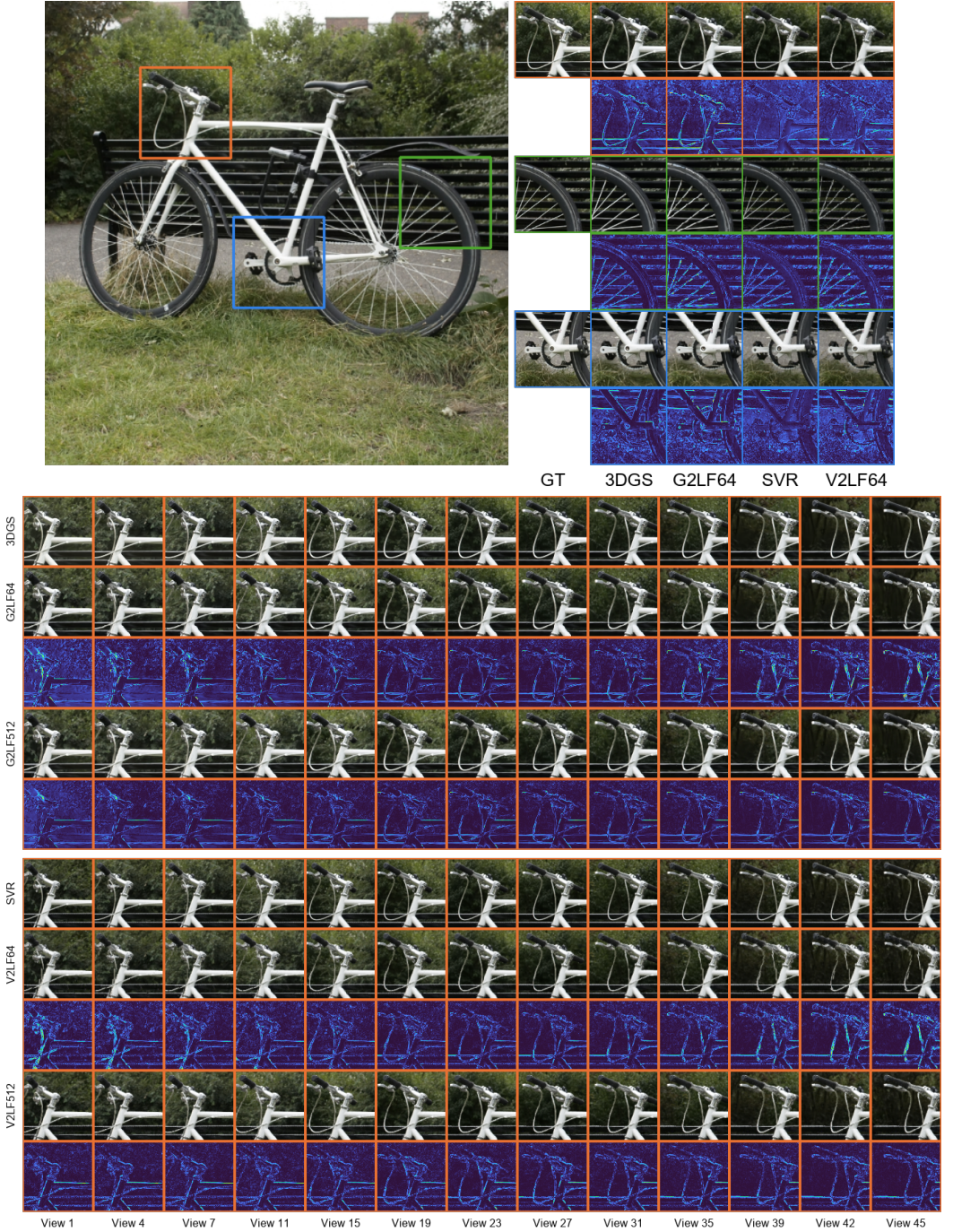


Fig. 12. Evaluation on the rendered BICYCLE scene from the MipNeRF-360 dataset [Barron et al. 2022]. Top: visual comparisons between the ground truth, the baseline models, and our proposed methods, with error maps highlighting per-pixel differences from the reference images. Bottom: rendered light field quilts across viewing angles, where the accompanying error maps show view-dependent rendering differences. The numbers appended to G2LF and V2LF denote the N_{chunk} .

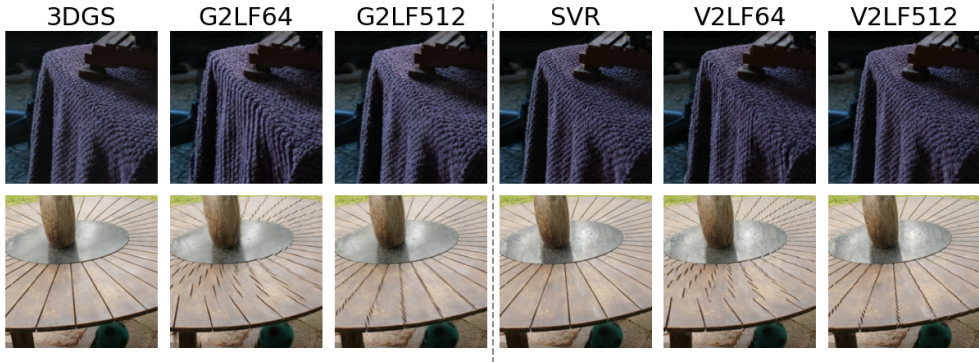


Fig. 13. Close-up comparison of the BONSAI and GARDEN scene from the MipNeRF-360 dataset [Barron et al. 2022] at the (top) left-most (View 45) and (bottom) right-most (View 1) viewing angle. Increasing the number of chunks reduces depth discontinuity errors.

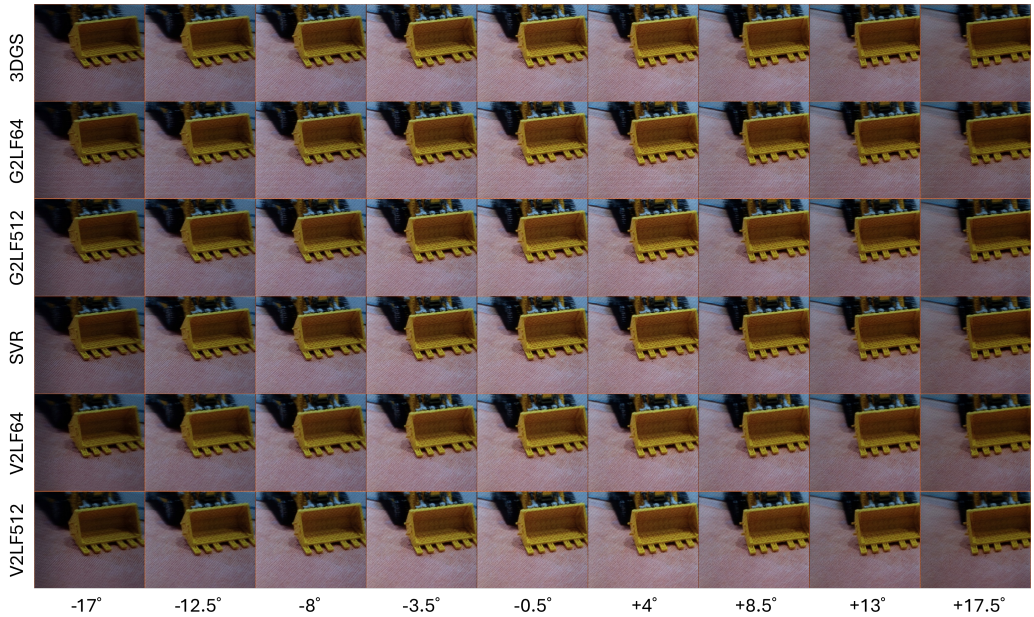


Fig. 14. Captured results of the KITCHEN scene from the MipNeRF-360 dataset [Barron et al. 2022] for multiple algorithm variants on a Looking Glass Go display. All variants produce geometrically consistent light-field rendering and faithful 3D reconstruction under real display conditions. In practice, artifacts visible in individual rendered views are less noticeable on the display, as multiple views perceptually blend within the viewer's pupil. More captured results are provided in the supplementary materials.