Paper 60-3 has been designated as a Distinguished Student Paper at Display Week 2025. The full-length version of this paper appears in a Special Section of the *Journal of the Society for Information Display (JSID)* devoted to Display Week 2025 Distinguished Papers. This Special Section will be freely accessible until December 31, 2025 via:

https://sid.onlinelibrary.wiley.com/doi/full/10.1002/jsid.2062

Authors that wish to refer to this work are advised to cite the full-length version by referring to its DOI:

https://doi.org/10.1002/jsid.2062

Real-Time Per-Pixel Predistortion for Head-Tracked Light Field Displays

Tianyu Wu, Anuraag Jajoo*, Hee-Jin Choi**, Benjamin Watson**
*Visual Experience Lab, North Carolina State University, Raleigh NC, USA
**Dept. of Physics, Sejong University, Seoul, Korea

Abstract
The latest light-field displays have improved greatly, but continue to be based on the approximate pinhole model. For every frame, our real-time technique evaluates a full optical model, and then renders an image predistorted at the sub-pixel level to the current pixel-to-eye light flow, reducing cross-talk and increasing viewing angle.

Author Keywords
Light-field display; integral imaging; eye-tracking; real-time ray tracing.

1. Introduction
The integral imaging light-field displays (LFDs) based on the pinhole optical model suffer from low resolution, limited viewing angle and cross-talk. Head-tracking can mitigate some of these issues, especially cross-talk, by identifying and addressing display areas containing it (1,2). Yet because of approximate optical models and rendering methods, existing head-tracked LFDs cannot eliminate cross-talk, especially at display peripheries and when virtual objects are far from the display surface. With more accurate, real-time optical modeling and rendering, our technique significantly reduces cross-talk and improves viewing angle. In continuing our previous work (3), our contributions include: a ray-based GPU algorithm that accurately models subpixel-to-eye light flow for slanted-lens LFDs in real time; a real-time ray-tracing renderer that creates imagery predistorted to match current light flow, reducing cross-talk and increasing viewing angle; and a user study showing that viewers prefer head-tracked 3D imagery made with these algorithms over well-known alternatives.

2. Conventional Optical Modeling for LFDs
Conventional LFDs use an approximate model, in which lenslets are assumed to act like pinholes, with viewers seeing the display panel locations that intersect with the rays from their eyes through the lenslet center.

While this simple pinhole optical model works adequately, viewers often notice distortions, particularly cross-talk. These distortions arise from several factors, including overlapping viewing zones, the angular range of light from lenslets, single-sided lens designs, and vertical refraction. Figure 1 shows how pixel viewing zones can overlap. In effect, viewers see not a single point below each lenslet, but something more like a line segment. Even when viewing zones do not overlap, the angular range of light entering the eye differs from the single-ray assumption of the pinhole models, especially at close viewing distances, causing multiple or partial pixels to be visible under each lenslet. Single-sided lenslets are common in LFDs, and exhibit much more distortion through lens aberration than double-sided lenslets, which are rare due to fabrication challenges and material refractive index mismatches. Optical researchers have attempted to correct these aberrations (4,5,6), either through corrective lens designs, or through predistortion of the image underneath, but are intended for untracked, multi-view LFD applications, and may have limitations for off-lens-axis viewing. Finally, pinhole models for multi-view, integral LFDs neglect vertical refraction, assuming perpendicular viewing for all pixel rows, leading to further distortions, especially at the display peripheries.

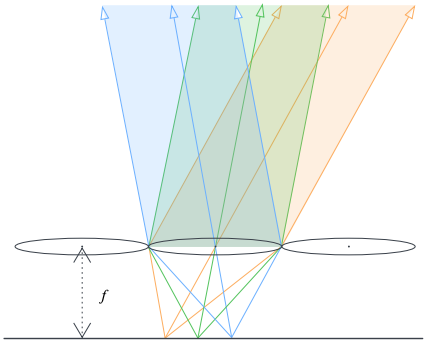


Figure 1. Overlapping viewing zones indicate cross-talk: multiple pixels visible under the same lenslet.

With their knowledge of eye position, recent head-tracked, two-view LFDs (1,2) reduce cross-talk with fewer viewing zones and with more accurate determination of visible pixels. Yet they do not fully eliminate cross-talk, since they do not correct for lens aberration or typically, vertical refraction.

3. Real-Time, Complex Optical Modeling & Predistortion

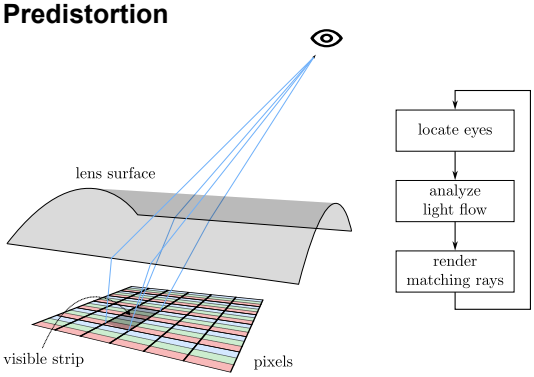


Figure 2. Right, the three stages for generating each frame in our real-time predistortion system. Left, the system's analysis follows light flow from the eye, through lens edges, to the display surface.

To address all of these sources of cross-talk and distortion, our technique combines head-tracking with more accurate optical modeling and matching imagery generated in predistorted form, both performed in real time with parallel GPU computation. As shown in the right of Figure 2, for each pixel row, we model light paths from the eye location to lenslet edges, refract the light paths on the lenslet's surface, and find the intersections between the paths and the boundaries between pixel rows at the bottom of the lens sheet, i.e. the display panel. This defines the visible strip, that is the portion of the pixel row visible under a lenslet. Note that

ISSN 0097-996X/25/5602-0809-$1.00 © 2025

the visible strip varies in width both across each lenslet and across each pixel row, so this model is 2D and must be calculated across the entire display. For slanted-lens LFDs, pixel row boundaries are slanted relative to the lenslet edge, and pixels in the slanted row must be mapped onto physical sub-pixels.

Finding the two light paths that move from a pixel row boundary through a lenslet's edges to the eye is non-trivial. The calculation involves several unknown variables, most importantly the refraction point, and requires solving a set of equations. Such calculation is too heavy for CPU to complete in a timely manner, but given that each set of equations can be solved independently, we can accelerate the computation using GPU parallelization. On an NVIDIA RTX 4080 GPU solving for a 4K screen, this takes well under 10 milliseconds.

Solving these equations requires several display parameters, including the horizontal positions of each lenslet's edges, the vertical position of each display pixel row, the tangent angle of each lenslet's curve at its edges, the lenslet's refractive index, the depth of the lenslets, and the slant angle of the lenslets. We solve for the two vertical positions on each edge of the lenslet through which light traveling from the pixel row boundaries through the lenslet would reach the eye. We also solve for the horizontal positions on the pixel row boundaries that the viewer sees through each lenslet edge. The results determine the visible strips of the display that the ray tracer samples, defined with precision exceeding sub-pixel resolution, as shown in the left of Figure 2.

The resulting trapezoidal visible strips fill the entirety of the display. With real-time ray tracing, for each pixel in each visible strip, we render the color that would reach the eye from the location in which it exits the lenslet, were there in fact no lenslet. In effect, this predistorts the image, common in rendering for lensed displays (7). However, rather than predistorting an already rendered, undistorted image, we directly render a predistorted image.

4. Sub-Pixel Conflict Resolution

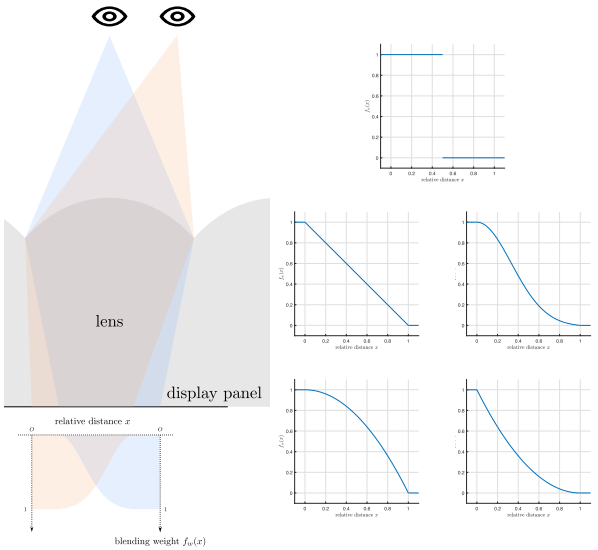


Figure 3. Stereo conflict blending. Upper left, the visible strips of the viewer's two eyes overlap. Lower left, the weights used to blend the colors that should be seen by each eye. Right, alternative weighting functions.

When the visible strips of the viewer's eyes overlap, containing the same pixels, cross-talk results. Due to the physical properties of the lenslets, it cannot be fully eliminated in most setups and viewing positions. But to significantly mitigate the cross-talk, we can blend the colors that would be seen by each eye. We blend with weights that depend on the horizontal relative position of each sub-pixel within the visible strips for both eyes. When a sub-pixel is only visible to one of the eyes, we display the color normally; when it is visible to both eyes, we blend the colors from both views using a function $f_w(x)$, as shown in Figure 3.

Given the location x of a sub-pixel between the two edges (locations 0 and 1) of the overlapping area of the two eyes' visible strips, the function $f_w(x)$ returns a weight between 0 and 1. If the colors for the left and right eyes are C_l and C_r, then the resulting blended color is $\frac{C_l \cdot f_w(x) + C_r \cdot f_w(1-x)}{f_w(x) + f_w(1-x)}$. Different blending functions can achieve various effects such as dividing the overlapping area between the eyes, leaving the overlapping area entirely unused, and blending two views with different strengths.

5. Objective Evaluation

To evaluate our real-time predistortion and its analysis, we display one of two monoscopic test patterns on the LFD: a visible columns (VC) image and a not visible columns (NVC) image. The VC image displays maximum brightness on sub-pixels that should be seen from a viewing position and turns off the sub-pixels that cannot be seen. Through the refraction of the lens sheet, the VC image should appear completely white. The complement of this VC image is the NVC image, it turns off visible sub-pixels and displays maximum brightness on those that cannot be seen, so the perceived image should appear completely black. To capture these images through the lens array, we used a DSLR camera in manual mode. We kept all camera parameters unchanged throughout the evaluation, excepting focal length and focus so that we could capture the best view of the LFD.

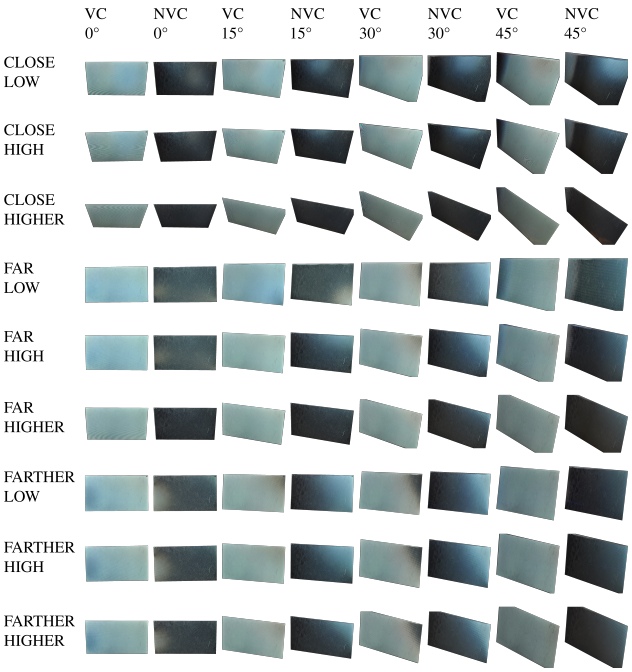


Figure 4. Testing predistortion and light flow analysis from various viewing distances, heights and angles.

Figure 4 shows the results. The contrast between corresponding VC and NVC photos shows that the accuracy holds up well in all perspectives. Areas of display with minor artifacts (e.g. 0° at farther distances) are likely due to imprecise camera position measurement, or unevenness of the prototype display itself. Images from a 45° viewing angle show a larger patch of artifacts near the far edge. At this extreme angle, the light from a pixel may not go through the lenslet directly over it to reach the eye, but rather a neighboring lenslet. In addition, 45° is very close to the theoretical viewing angle limit 50.74°, which is the complement of our display's lenslet edge tangent angle, in this case, the light may pass through two lenslets instead of just one. Our analysis currently models neither of these neighboring or two-lens cases. Finally, we note that a viewing angle of 45° is well beyond the capabilities of our (and indeed many) eye trackers.

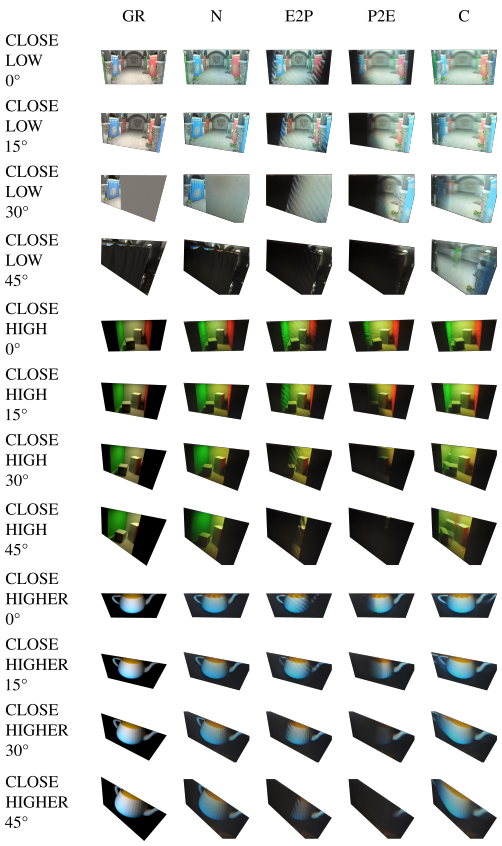


Figure 5. Comparison of our real-time predistortion for LFDs to existing rendering techniques across three scenes. Each of row groups 1-4, 5-8, and 9-12 show a different scene. Each row also shows a different distance, height, and viewing angle. Each column shows a different rendering technique, as described in the text.

Next we compare images generated with our real-time LFD predistortion technique (N) to a lensless, very high-quality "ground truth" (GR), and to images created by three conventional pinhole-based LFD techniques. The first is a head-tracked LFD that determines visible pixels with the ray from the eye to the lenslet center (E2P). The second is an alternative head-tracked LFD that determines the viewing direction of a pixel by the ray from the pixel to the lenslet center (P2E), similar to non-head-tracked multi-view LFDs, but it turns off pixels whose viewing direction rays are too far away from the eye positions. The last is a multi-view LFD (C) that uses head tracking only for determining the image's vertical perspective. Figure 5 shows these techniques across three test scenes including Sponza, Cornell Box, and the Utah Teapot (8).

The images showing our (N) and the E2P techniques contain only one view to avoid the stereo conflicts we discuss in Section 4, easing evaluation (both do support stereo viewing, evaluated in Section 6). The images showing the P2E and C techniques contain many views; they cannot be usefully reduced to single-view. The ground truth imagery (GR) is also single-view, rendered at a much higher resolution using all display pixels, and captured without the lens sheet.

Across all these views and scenes, the images built with our real-time predistortion technique look very similar to the ground truth imagery, and look markedly better than the other LFD techniques. The head-tracked E2P technique is the second best in image quality, but with noticeably unlit areas containing pixels that the pinhole model predicts must be viewed through neighboring (rather than co-located) lenslets. Note that these unlit areas are much larger than in our technique, exposing the approximations of the pinhole model. The multi-view P2E and C techniques have similar quality to one another, with both only containing a fixed set of horizontal views, and producing significantly blurrier images in scenes with greater virtual depth.

6. Subjective Human Evaluation

Table 1. Average rankings of the four LFD techniques by participants considering image sharpness, 3D effect at close and far distances, and overall quality.

LFD techniques	Sharp-ness	3D effect (close)	3D effect (far)	Best overall
N	1.094	1.375	1.625	1.125
E2P	2.406	2.594	2.125	2.750
P2E	2.875	2.875	2.688	2.688
C	3.625	3.156	3.562	3.438

We conducted an experiment to find out which LFD technique viewers prefer. The experiment had one independent variable, the LFD technique: ours (N, with the blending function disabling overlapping areas), E2P, P2E and C. As dependent measures, participants ranked the LFD techniques according to image sharpness, 3D effect at close viewing distance, 3D effect at far viewing distance, and overall. With IRB approval, we recruited 32 participants from NC State University, compensating them either with $4 or course credit. We hypothesized that participants would prefer our technique.

We ran the LFD techniques on a PC equipped with an NVIDIA RTX 4080 GPU capped at 10 Hz for all techniques. A Tobii 4C tracked eye positions, while a prototype 4K display with a slanted lenticular lens array showed stereoscopic imagery. With this equipment, participants moved freely to observe game-like scenes, ranking the four techniques by each of the criteria in sequence. As they ranked by each criteria, participants could cycle across the four LFD techniques, and seven different scenes, by pressing two different keys. The order of the techniques and scenes varied randomly between participants. To support ranking,

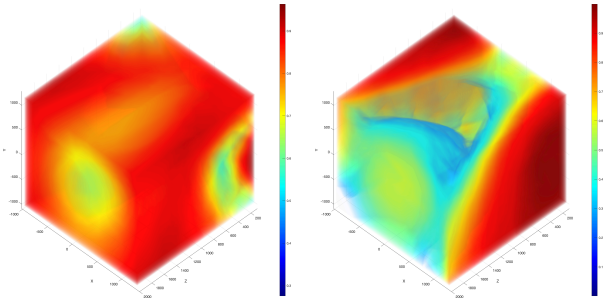
images were labeled with a number indicating the LFD technique. The number labels also varied randomly between participants. Participants did not know which technique we developed. Participants required approximately 20 minutes to complete the experiment.

Table 1 shows our results. Our hypothesis was confirmed: by all four criteria, participants ranked our LFD technique (N) first.

7. Optical Model Simulation

The prototype display hardware we used was designed for untracked, multi-view LFD and may not have the optimal optics for our LFD technique. We therefore conducted a simulation that searched the lens parameter space to find optics that would yield less stereo conflict, while maintaining similar display visibility during typical viewer head movement.

Figure 6 visualizes the simulation results as heatmaps across the display's viewing space, with warm colors indicating higher stereo conflict, and cool colors lower conflict. The X, Y and Z dimensions are spatial and aligned with the display, with the front, left face of the color cube parallel to the display surface and distant from it. The left cube (a) shows the stereo conflict with our current prototype display's optical parameters, which exhibit high conflict rates throughout most of the viewing space, except a cylindrical area along the Z-axis (centered on and perpendicular to the display). This aligns well with our viewing experience. The right cube (b) shows the lowest-conflict result from a coarse search (due to time constraints) of the optical parameter space, which features a larger, low-conflict conical space centered on the Z-axis.



(a) (b)

Figure 6. Stereo conflict visualized across viewing space, with warm colors indicating high conflict, cool colors low conflict. Left (a), conflict with our current display's parameters; right (b) conflict with improved parameters.

Our simulation data suggests that the greatest optical factor affecting the conflict rate is the ratio of the depth of the lens substrate over the axial focal length, with values slightly smaller than 1.0 achieving lower conflict. We believe this is because off-axis viewing decreases focal length, so a smaller ratio brings its average as viewed across the display closer to 1.0, which focuses light on the display panel. We also found that lower refractive indices increase visible display area and lower stereo conflict, but can also reduce viewing space without conflict.

8. Conclusions & Future Work

Conventional and even head-tracked LFDs use approximate optical models that result in significant cross-talk, and limitations in viewing angle. By using a more accurate model that we update per sub-pixel and per frame, we are able to directly render predistorted imagery that overcome these shortcomings.

Our work has at least two limitations. First, our LFD technique can run slowly, dropping to a few Hz when models are complex. Performance will improve when we use the most-current NVIDIA RTX 5000 series and future GPUs. We can realize further speedups by increasing parallelism in our technique, which currently completes analysis for the entire display before beginning to render the predistorted image. Later, we plan to add foveated rendering to improve performance and potentially analysis accuracy as well, which should be relatively straightforward with our ray-traced rendering. Secondly and in the longer term, we plan to make an LFD hardware prototype with improved optics following the simulation results from Section 7.

9. Acknowledgements

We thank Turner Whitted for providing valuable inspiration and advice. This work was supported in part by NSF award IIS-2008590, and by a National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2021R1A2C1011803, Research on gaze-contingent hybrid volumetric display).

10. References

1. Lee Jh, Yanusik I, Choi Y, Kang B, Hwang C, Park J, et al. Automotive augmented reality 3D head-up display based on light-field rendering with eye-tracking. Optics Express. 2020; 28(20): 29788--29804.

2. Aoyama K, Yokoyama K, Yano T, Nakahata Y. 48-5: Eye-sensing Light Field Display for Spatial Reality Reproduction. In SID Symposium Digest of Technical Papers; 2021. p. 669--672.

3. Wu T, Choi HJ, Watson B. P-231: Late-News Poster: Real-Time Analysis and Synthesis of Imagery for Light-Field Displays. In SID Symposium Digest of Technical Papers; 2024. p. 1577--1580.

4. Daniell S. Correction of aberrations in lens-based 3D displays. In Stereoscopic Displays and Virtual Reality Systems XII; 2005. p. 175--185.

5. Yu X, Li H, Sang X, Su X, Gao X, Liu B, et al. Aberration correction based on a pre-correction convolutional neural network for light-field displays. Optics Express. 2021; 29(7): 11009--11020.

6. Zhang Q, Miyanishi Y, Sahin E, Gotchev A. Optical Aberration Analysis of Light Field Displays: A Calibration Approach for Enhanced Performance. In IS and T International Symposium on Electronic Imaging; 2024.

7. Watson BA, Hodges LF. Using texture maps to correct for optical distortion in head-mounted displays. In Proceedings Virtual Reality Annual International Symposium'95; 1995. p. 172--178.

8. McGuire M. Computer Graphics Archive. [Online].; 2017. Available from: https://casual-effects.com/data.