View-Dependent Light-Field Display that Supports Accommodation Using a Commercially-Available High Pixel Density LCD Panel

Tuotuo Li, Qiong Huang, Santiago Alfaro, Alexey Supikov, and Ronald Azuma
Intel Labs, 2200 Mission College Blvd. Santa Clara, CA, 95054, USA

Abstract
Super multiview displays are light-field displays that support accommodation by steering high densities of directional light into small eyeboxes around the viewer's eyes. Unfortunately, this greatly restricts where the display is viewable. We describe a view-dependent approach that combines custom optics, eye position tracking and a high-density LCD panel to provide full parallax and a larger range of viewing positions. When combined with a directional backlight, this approach provides a 3D volume from which a viewer can see 3D objects with accommodation.

Author Keywords
3D display; super multi view display; vergence accommodation conflict; integral imaging; compute shader.

1. Introduction
One of the biggest challenges in 3D displays is the vergence-accommodation conflict, which significantly contributes to eyestrain. Many different display approaches have been developed to present depth cues correctly to overcome this conflict; a representative but hardly comprehensive list of such approaches include tensor displays, holographic displays, volumetric displays, multi-focal plane displays, and light-field displays (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13). Light-field displays approximate the directional beams of light emitted from real objects. If at least two directional beams from a single 3D point enter the eye pupil, light-field displays can start to support accommodation (14). The most straightforward way to implement a light-field display is integral imaging, but the tradeoff between spatial and angular resolution means it is challenging to simultaneously provide adequate spatial resolution and sufficiently high angular resolutions for accommodation.

Therefore, super multiview approaches aim to provide enough angular resolution to support accommodation with existing display technologies by modifying the basic integral imaging technique to squeeze the angular views into small eyeboxes (15). In turn, this restricts the range of viewpoint locations from which the viewer can see the display, and many super multiview displays limit the viewpoint parallax to small horizontal motions.

This paper describes an approach to extend the range of positions from which a super multiview display can be seen. We designed and built custom microlens arrays (MLAs) with long focal lengths and placed those in front of high pixel per inch (PPI) LCD display panels. We built prototypes demonstrating 72 PPI and 100 PPI spatial resolution light-field displays where the angular view density is 2 by 2 views across a 4mm eye pupil aperture, thus enabling accommodation. To support both horizontal and vertical parallax, we built a custom eye position tracker that computes the viewer's eye positions in real time and dynamically renders the content in real time to adapt to the moving eye positions. We use screen-space ray tracing (16) to achieve real-time rendering rates by minimizing the number of rendered viewpoints. To extend the viewing zone into a 3D viewing volume, these displays can be combined with a directional backlight (5) that enables time-multiplexed sequential illumination of left and right eye views. This paper describes our prototypes that validate this approach, including images taken from the prototype displays.

2. Approach
Our system consists of the following components: the optical subsystem, real-time eye tracking, and real-time rendering.

Optical system: In most basic integral imaging designs, the eyebox and FOV are designed to be large. This spreads the available pixels over a large viewing area, requiring an extremely high pixel density to simultaneously support accommodation and an acceptable spatial resolution. Only a tiny fraction of the pixels is visible at any instant, so most of the pixels are wasted.

Therefore, we designed and manufactured custom MLAs where the lenslets have long focal lengths. These create small eyebox regions with high view density. We steer the eyebox regions dynamically toward the viewer's pupils by tracking the eye positions and updating the rendering in real time.

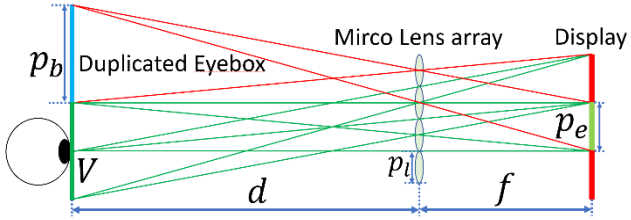

Figure 1. Computing eyebox and elemental image.

Given a pupil location, viewing distance, lens pitch and focal length, the elemental image and eyebox can be computed from **Figure 1**. Assuming viewpoint V is at the center of the eyebox, the ideal principal ray that passes through a lenslet will intersect the center of the elemental image for that lenslet, which means the boundary of that elemental image will be the projection of the lenslet onto the display from the center of the eyebox. The elemental image size is specified by the following equation:

$$\boldsymbol{p_e = p_l * \frac{d+f}{d}} \quad (1)$$

Where p_e is the elemental image size, p_l is the lens pitch. f is the focal length of the lens and d is the viewing distance. The boundary of the eyebox at distance d is the place where the viewer would start to see other elemental images. The size of the eyebox is calculated by the following equation:

$$\boldsymbol{p_b = p_e * \frac{d}{f}} \quad (2)$$

As shown in the above equations, the elemental image size and the eyebox size depend on the viewing distance.

The eyebox must be large enough to ensure a pupil stays within the intended eyebox during reasonable head and eye motions, despite the system latency from the tracking, rendering and display modules. We designed our prototype to provide 12 by 12 views within a 24mm by 24mm eyebox at 500mm viewing distance, which implies a numerical aperture around 22.

Figure 2. Elemental image projected onto viewer's eye. (Image is not at scale).

Figure 2 shows that each elemental image and corresponding lens behave like a mini projector that projects parallel beams from the display pixels into a user's eye. Each pixel inside the elemental image represents a different view. To trigger accommodation, the beams projected from a pixel must have sufficiently high density to ensure that at least 2 by 2 beams appear partially or completely inside the pupil. Since a typical eye pupil is approximately a 4mm diameter circle, we designed the beams to be 2mm by 2mm at the target viewing distance.

We designed a custom MLA with 0.35mm lens pitch and 8mm focal length and had that manufactured. When paired with a 10.1" 4K display panel, the resulting system matches our target parameters and provides 72 PPI spatial resolution, which is the minimum acceptable spatial resolution.

A property of this system is that the eyeboxes repeat themselves. Since LCD panels typically emit diffuse light across a wide angle, the same elemental image can be seen through multiple adjacent lenslets, creating multiple duplicated eyeboxes.

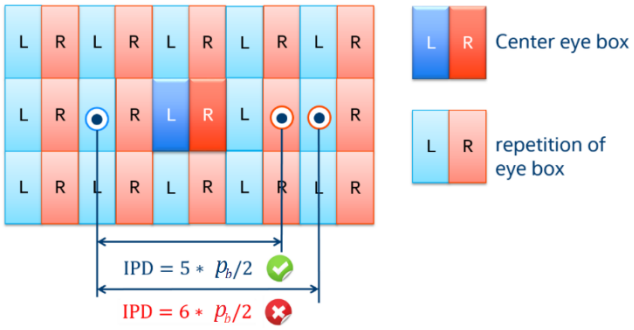

Figure 3. Subdivided eyeboxes (left and right regions).

To provide different imagery to the left and right eyes, we split the eyebox into two horizontal regions, one for the left and the other for the right eye views. The viewer sees the correct imagery when his or her left eye is in the left eyebox region and his or her right eye is in the right eyebox region. **Figure 3** shows this split. However, this requires the viewer to be at particular distances away from the display to see the correct images. The viewer's interpupillary distance (IPD) must be an odd multiple of the half eyebox width ($p_b/2$) to see the correct imagery. If the viewer's IPD is an even multiple of the half eyebox width, then the viewer sees incorrect images. From equations 1 and 2, we can see that the eyebox size is a linear function of the viewing distance. This means that the viewer will see the intended imagery only at certain viewing distances, as illustrated by **Figure 4**.

The solution to this limitation is to add a directional backlight and temporally multiplex the left and right eye imagery. Instead of dividing the eyebox into left and right eye regions that are shown simultaneously, as seen in **Figure 3**, we sequentially display the left eye content in the eyebox, then the right eye content. To ensure that each eye sees only the intended content, a directional backlight steers the light such that the display illuminates only the

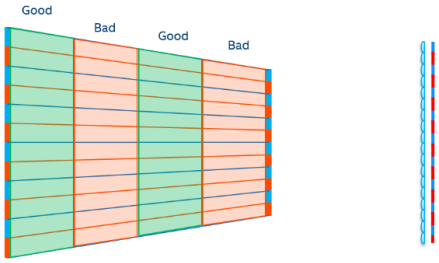

Figure 4. Good and bad viewing distances for a fixed IPD

left eye, then only the right eye, as shown in **Figure 5**. The left and right eye images are displayed and switched at 120 Hz using quad buffered stereo supported graphic hardware in sync with the backlight to preserve visual fusion.

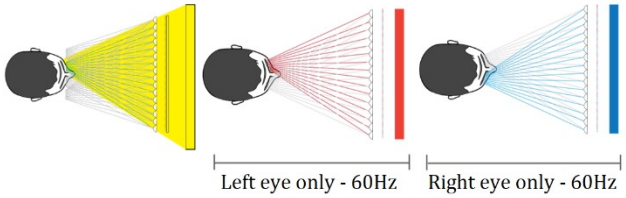

Figure 5. Time Multiplexing with Directional Backlighting.

When combined, all these system components create a glasses-free light-field display that supports accommodation where the viewer can view the display across a large viewing volume. However, this requires a high PPI 120 Hz LCD panel to work with the directional backlight and provide at least 2 by 2 views into the viewer's pupils.

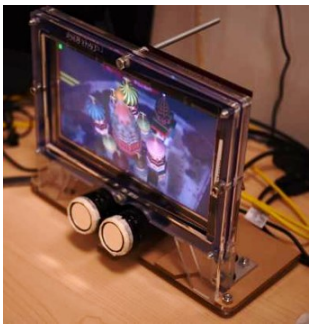

Figure 6. 72 PPI light-field display prototype with stereo cameras for eye position tracking.

Real-time Eye Tracking: We built a custom eye position tracking system that uses two infrared (IR) cameras located beneath the light-field display, facing toward the viewer, as seen in **Figure 6**. The viewer's face is illuminated by IR light to improve the tracking. Our tracking software finds facial features and uses those to detect the user's eyes and outline of each pupil. We then estimate the centroid of each pupil. Computing the 3D location of each pupil center is done through straightforward triangulation using the two stereo camera views. The overall system precision was validated by comparing the computed IPDs against mechanical measurements of a group of subjects, and the difference is no greater than 0.6mm. We tested this tracker with many users of different ages, genders and racial backgrounds.

Real-time Rendering: The most straightforward approach for rendering light-field imagery onto the display from virtual 3D models is to render multiple views with multiple different asymmetric frustums. After rendering the many different subimages, we interleave those into one overall image with the correct elemental images on the display behind the MLA array.

However, in our system each eyebox requires 12 by 12 views. Rendering the same scene 144 times with slightly different view frustums is expensive, particularly for detailed geometries or expensive lighting. Therefore, we instead use a screen-space ray tracing approach (16) to reduce the rendering cost.

In this approach, we render an image at the central viewpoint through a conventional rendering pipeline, storing RGB + depth. The left side of **Figure 7** shows how screen-space ray tracing then generates the light-field images for all other viewpoints. For any pixel *P* on the display panel, the ray direction after passing though a microlens is fixed and can be determined though a calibration procedure. Given the ray direction and the location of *P*, the intersection points between the ray and the near and far clipping planes, which are labelled *A* and *B* in **Figure 7**, are uniquely defined. Points *A* and *B* are then projected onto the RGBD image, and 2D ray marching is performed from the near clipping plane to determine the first intersection with the content, thus determining the RGB value for the pixel associated with that ray. The right side of **Figure 7** shows an example of a rendered scene where all the elemental images have been set through this screen-space ray tracing approach.

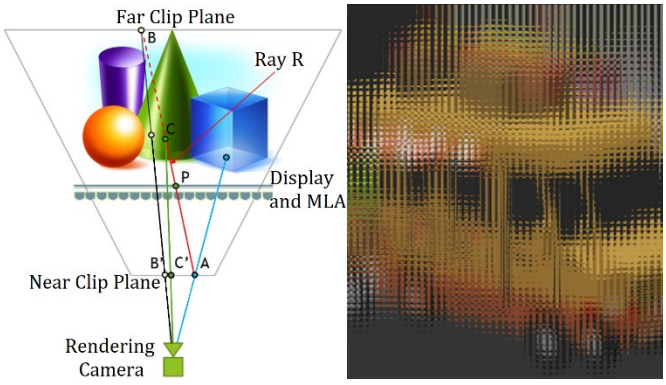

Figure 7. Left: Screen-space rendering. Right: Rendered elemental images on display (as seen without lenslets). Content: Simple Town Lite (3D Warehouse license).

3. Experimental results

To validate our approach, we built three prototypes, using our custom MLAs and off-the-shelf display panels:

1. A real-time view-dependent light-field display system using 0.35mm diameter microlenses and a 10.1" 4K display.
2. A real-time view-dependent light-field display system using 15 tiled 0.25mm diameter microlenses and 4 tiled 5.5" 4K displays.
3. A set of static displays with 0.35mm diameter microlenses placed over 1500 PPI LVT (Light Value Technology) prints in place of a dynamic display.

Integrating system 1 or 2 with a directional backlight requires a high PPI, true 120 Hz LCD display panel, which we have not been able to procure. A directional backlight has been proven to work to produce stereo views with a low PPI, 120Hz display panel and this has been commercially shown by a company called RealD Me (5). Therefore, if we demonstrate that systems 1 and 2 work as real-time light-field displays, we can be confident that the entire approach works once a high PPI, true 120 Hz LCD display panel becomes commercially available.

Systems 1 and 2 achieve real-time rendering, at 100 Hz and 40 Hz respectively, using an NVIDIA GeForce 2080 graphics card. However, the display panels limit the overall update rate to 60 Hz and 30 Hz, respectively. The camera acquisition time is ~12.5ms and the eye tracking system latency is ~20ms. Thus, the total system latency varies from 45-65ms. This is sufficiently low to enable the system to steer the eyeboxes to follow a viewer who moves his/her head slowly.

Figure 8. Captured images from system 1 prototype. Top: Camera focused at near distance. Bottom: Camera focused at far distance. Content: Stockcar racecar (Unity asset store license) and St. Basil's Cathedral (3D Warehouse license)

System 1 is shown in **Figure 6**. **Figure 8** shows two images captured from that display to prove that the display supports accommodation. The top image shows a scene where the camera is focused to the near objects. Note that the text in the physical frame is also in focus. The bottom image shows the situation where the camera is focused to the far objects. Now the far virtual objects are in focus and the text in the physical frame is out of focus. Nothing has changed in the display itself. The only difference is the focal state of the camera.

Figure 9 shows an image from system 2. This also supports accommodation and provides a higher spatial resolution (100 PPI) than system 1 (72 PPI). However, achieving this required tiling four display panels and using 15 separate MLA tiles. This results in noticeable artifacts. The black cross in the center is the space between the displays. We cannot place them seamlessly adjacent to each other. Also, calibrating 15 separate MLA tiles with respect to four display panels proved challenging and resulted in uneven performance across the display.

System 3 uses a high resolution static print in place of the display panel. This represents the performance this approach can achieve when high resolution LCD displays become available in the future. We built three different versions to cover a range of IPDs so that most viewers can see the light field in one of the prototypes. Since the displays are static, each display must be

viewed from a specific viewpoint, and each works for a small range of viewer IPDs. **Figure 10** shows a village scene viewed in the display, where the near buildings at the bottom are in focus but the distant houses at the top of the scene are not in focus.

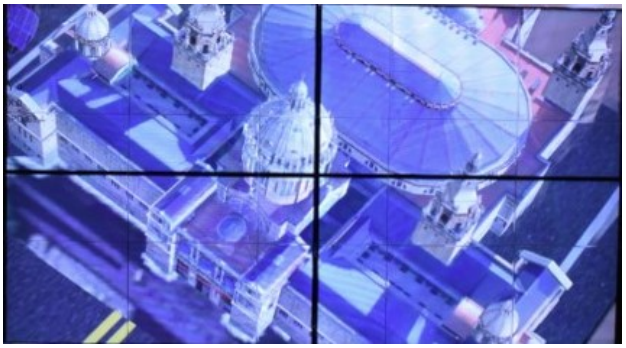

Figure 9. Captured image from system 2. Content: Museu Nacional d'Art Barcelona (3D Warehouse license)

Figure 10. Captured image from system 3. Content: Viking village (Unity asset store license)

4. Discussion

We presented a light-field display approach that provides acceptable spatial resolution, supports accommodation and offers a large viewing volume. The cost is the requirement to track the viewer's eye positions, and the restriction to a single viewer.

Moiré patterns are visible in our prototypes because the MLA and display pixels have different but uniform pitches. To reduce this artifact, we rotated the display with respect to the MLA and also inserted a weak diffuser on top of the display. Display manufacturers can eliminate this problem by using pseudorandom RGB pixel masks or by slightly varying the pixel pitch.

Based on our prototypes and previous work in directional backlights, we have shown that this approach can feasibly achieve a practical light-field display in thin form factors. The core missing component is a true 120 Hz, high PPI LCD display panel. Both high PPI and high refresh rate LCD panels exist, so it is technically possible to manufacture this missing component. Given investment to make this happen, it is possible to demonstrate a fully integrated version of this approach.

5. References

1. Takaaki U, Yasuhiro T. Super multi-view near-eye display to solve vergence–accommodation conflict. Optics Express. 2018; 26(23).
2. Takaki Y. Multi-view 3-D display employing a flat-panel display with slanted pixel arrangement. Society for Information Display. 2010 July; 18(7).
3. Takaki Y. Super multi-view and holographic displays using MEMS devices. Displays. 2014 October; 37.
4. Wetzstein G, Lanman D, Hirsch M, Raskar R. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. ACM Transactions on Graphics. 2012; 31(4).
5. Woodgate G, Robinson M, Harrold J, Ihas B, Ramsey R. Towards Direct-View Accommodative Light Field Displays. SID Symposium. 2018 May; 49(1).
6. Fattal D, Peng Z, Tran T, Vo S, Fiorentino M, Brug J, et al. A multi-directional backlight for a wide-angle, glasses-free three-dimensional display. Nature. 2013 March; 495.
7. Lanman D, Luebke D. Near-Eye Light Field Displays. ACM Transaction on Graphics. 2013 November; 32(6).
8. Maimone A, Georgiou A, Kollin J. Holographic near-eye displays for virtual and augmented reality. ACM Transactions on Graphics. 2017 July; 36(4).
9. Kajiki Y, Yoshikawa H, Honda T. Hologramlike video images by 45-view stereoscopic display. In Stereoscopic Displays and Virtual Reality Systems IV; 1997; San Jose, CA, United States: SPIE. p. 154 -- 166.
10. Takaki Y, Nago N. Multi-projection of lenticular displays to construct a 256-view super multi-view display. Optics Express. 2010 April; 18(9).
11. Häussler R, Gritsai Y, Zschau E, Missbach R, Sahm H, Stock M, et al. Large real-time holographic 3D displays: enabling components and results. Applied Optics. 2017 February; 56(13).
12. Hunter S, Azuma R, Moisant-Thompson J, MacLeod D, Disanjh D. Mid-Air Interaction with a 3D Aerial Display. In Proceedings of SIGGRAPH 2017 Emerging Technologies Installation; 2017; Los Angeles.
13. Narain R, Albert R, Bulbul A, Ward G, Banks M, O'Brien J. Optimal presentation of imagery with focus cues on multi-plane displays. ACM Transactions on Graphics. 2015 August; 34(4).
14. Takaki Y. High-Density Directional Display for Generating Natural Three-Dimensional Images. Proceedings of the IEEE. 2006 March; 94(3).
15. Takaki Y, Tanaka K, Nakamura J. Super multi-view display with a lower resolution flat-panel display. Optics Express. 2011; 19(5).
16. McGuire M, Mara M. Efficient GPU Screen-Space Ray Tracing. Journal of Computer Graphics Techniques. 2014 December; 3(4).