

A Practical Multi-viewer Tabletop Autostereoscopic Display

Gu Ye[‡]

Andrei State[§]

Henry Fuchs[¶]

Department of Computer Science
University of North Carolina at Chapel Hill

ABSTRACT

This paper introduces a multi-user autostereoscopic tabletop display and its associated real-time rendering methods. Tabletop displays that support both multiple viewers and autostereoscopy have been extremely difficult to construct. Our new system is inspired by the “Random Hole Display” design [11] that modified the pattern of openings in a barrier mounted in front of a flat panel display from thin slits to a dense pattern of tiny, pseudo-randomly placed holes. This allows viewers anywhere in front of the display to see a different subset of the display’s native pixels through the random-hole screen. However, a fraction of the visible pixels will be observable by more than a single viewer. Thus the main challenge is handling these “conflicting” pixels, which ideally must show different colors to each viewer. We introduce several solutions to this problem and describe in detail the current method of choice, a combination of color blending and approximate error diffusion, performing in real time in our GPU-based implementation. The easily reproducible design uses a pattern film barrier affixed to the display by means of a transparent polycarbonate layer spacer. We use a commercial optical tracker for viewers’ locations and synthesize the appropriate image (or a stereoscopic image pair) for each viewer. The system supports graceful degradation with increasing number of simultaneous views, and graceful improvement as the number of views decreases.

Keywords: autostereoscopic display, 3D display, tabletop display, mixed reality, multi-user display, collaborative display, random hole barrier, parallax barrier.

Index Terms: I.3.1 [Computer Graphics]: Hardware Architecture—Three-dimensional displays;

1 INTRODUCTION AND MOTIVATION

Autostereoscopic displays provide stereo perception without users having to wear special glasses. A wide variety of these displays have been devised over the past several decades, but with the possible exception of advertising, where their novelty attracts the attention of potential customers, they have not been even modestly adopted. We believe that the major reason for this lack of success is the high price/performance ratio of such devices: the capable autostereo displays are prohibitively expensive, while the affordable ones have very limited capabilities.

The work described here is based on the recently introduced “Random Hole Display” concept [11], which it expands into a practically usable multi-user autostereo system, as we shall demonstrate. We hope that such displays will eventually be used in many scenarios. In this work, we provide a glimpse into possibilities for local multi-user collaboration on tabletop geometric data (e.g. as

Figure 1 shows); beyond that, we expect future applicability to teleconferencing and telemedicine, where autostereoscopy could deliver imagery that contains a complete set of depth cues for multiple users, giving each user the sense of being co-located with all the other participants.



Figure 1: Mixed reality scenario using tabletop autostereoscopic display: Virtual teapot and real cup. The two photos are taken at the same time by two cameras at different positions. Two views are generated simultaneously by the tabletop RHD.

[‡]e-mail: guye@cs.unc.edu

[§]e-mail: andrei@cs.unc.edu

[¶]e-mail: fuchs@cs.unc.edu

2 BACKGROUND AND CONTRIBUTION

At present, three broad classes of autostereo display technologies exist (in the most inclusive sense of autostereo). They are holographic, volumetric, and parallax-based. Both dynamic holographic displays [5] and volumetric displays [3] can generate stereo imagery for multiple users. However, traditional volumetric displays

lack the ability to support occlusion.

Recent work [6] solves the occlusion problem through the introduction of a spinning anisotropic mirror that provides full-surround distinct views in the horizontal direction; however, while it could provide single-view motion parallax in the vertical direction by means of viewpoint tracking, it could not possibly offer multi-user motion parallax (multiple distinct views) along the vertical direction.

Parallax-based displays, based on barriers or lenticular lens sheets, provide a relatively simple and inexpensive solution for autostereoscopy. They produce a stereoscopic effect by spatially subdividing the display surface. In parallax barrier displays, the barrier consists of an optically opaque film with holes or slits that allow light to pass through. The barrier sheet is placed at a small distance in front of the display panel, usually on the order of a few millimeters. At a certain distance from the display, the left and right eyes of a viewer, when looking through the same slit or hole in the barrier, each see a different area of pixels of the display surface, thus enabling the display to present different images to the viewer's left and right eyes. Head tracking is needed to support dynamic view positions; this requirement is a limitation of this type of display. One of the more impressive examples for parallax barrier systems is the Varrier display [14], consisting of 35 LCD panels with film barriers. The visibility of the display pixels is determined by tracking the user's eye positions, and the correct pixels are illuminated for each eye.

Another method that uses a variable barrier pattern is Perlin's NTU autostereoscopic display [12]. This display contains an active light-blocking shutter that changes in response to the tracked user's head position. Similarly, the Dynallax display [13] uses an active LCD barrier to present up to 4 distinct views. The combination of tracking and dynamic barrier allows these systems to maximize the use of the display pixels, improving perceived resolution and brightness. The main drawback of a dynamic barrier is that the barrier LCD panel significantly reduces the overall brightness and contrast ratio of the display.

Lenticular barrier displays operate in a similar fashion, but with a series of long lenses called lenticules instead of barrier slits. These lenticules are narrow, cylindrical lenses with convex-planar cross-section and are packed to form a sheet that is affixed to the front of the display. Each lenticule directs the light from a particular subpixel in a certain direction. The lenticular sheet transmits significantly more light than a slit/barrier arrangement, leading to a brighter display, but at the expense of focus degradation. As a result, the effective viewing distance is limited to a modest range. Most displays in this category support a small number of views. The MERL 3D TV system [10] supports 16 views, deploying 16 projectors. The lenticular sheet can be replaced by a sheet of holographic lens elements, such as in some systems by Holografika [1, 2], which support 64 views in the horizontal direction, using 64 projectors. With so many views, the (horizontal) viewing zone becomes sufficiently large for untracked viewers to continually obtain distinct correct imagery as they move across the front of the display. However, its high display performance comes at a very high price since all 64 views are continually generated and displayed by means of 64 projectors. Lenticular displays provide parallax only in the fixed horizontal direction, which makes them unsuitable for tabletop scenarios.

Integral imaging display uses 2D array of lenslets to generate full-parallax autostereo images. The lenslets distribute the outgoing pixels across two angular dimensions, providing support for distinct views for both horizontal and vertical viewer positions in front of the display. Such displays can, in principle, be used for tabletop applications. Their major advantage over our random-hole-barrier based approach is that no tracking system is needed. However, the limited viewing angle [16] and low resolution make these displays

hard to use in tabletop scenarios. For example, the integral imaging display constructed at the University of Tokyo [9], uses a multiple projector system with 9 XGA projectors, producing a display of 2872×2150 pixels across $241\text{mm} \times 181\text{mm}$ (about 300 DPI), and its lens pitch is 1.016mm; therefore, the effective resolution for each view is no more than 240×197 pixels. Each lens of this display generates 12×12 views and its viewing area (expressed as an angle measured at the center of the display) is about $\pm 8^\circ$ in the horizontal direction and $\pm 7^\circ$ in the vertical direction. Such a viewing area is probably too limited for tabletop applications since viewers often stand next to the table and a large viewing angle is expected. In order to increase the viewing area, the integral imaging display needs to generate more views at each lens; for example, if the number of views generated by each lens were increased to 24×24 , then lens array size would decrease to 120×99 . Furthermore, it is quite challenging to manufacture lenslets such that they support large viewing angles [7]. In comparison, our display does not rigidly assign pixels to viewing directions, making the effective resolution depend on the number of viewers instead of on the range of viewing zones. For our system, the effective resolution is currently $1/13.2$ of the underlying 2560×1600 display panel, which is approximately 705×440 (of course, with additional views there are conflicting pixels; a more detailed discussion and evaluation will follow in Sections 6.2 and 7.2.). And the viewing angle is much more flexible (e.g. the photographs in Figure 8 were taken with a viewing angle of 73°).

Our concept is vaguely similar to the Illusion Hole [8] which provides stereo images for three or more users in tabletop applications, but since these displays require stereo glasses to separate the two (for left eye and right eye) views of each user, they cannot be considered autostereoscopic.

The "Random Hole Display" (RHD) concept [11] proposes a non-uniform barrier, using a Poisson-disk distributed pseudo-random hole pattern. When multiple viewers are present, some pixels are visible to more than one viewer, but due to the non-uniform hole pattern, the RHD is able to diffuse the errors resulting from these conflicts across the image, turning them into pseudo-random white noise. However, there are several problems that interfere with the use of RHDs in practical applications:

1. The rendering method does not handle the subpixel coverage issue (explained below) and does not contain methods for optimal image quality, nor does it deploy hardware-accelerated rendering.
2. Only four fixed viewer positions were calibrated, no user tracking is used to support dynamic viewing positions. The fixed-position calibration method does not apply to dynamic viewpoint situations.
3. The inter-reflection between barrier and display generates significant interference, which visibly degrades the final imagery.

In this paper, we demonstrate our enhanced random hole barrier display, a significant improvement over the initial proof-of-concept demonstration [11], and practically usable as a multi-viewer tabletop autostereoscopic display. The major contributions of our work can be summarized as follows:

1. Hardware-accelerated rendering algorithm, delivers optimal image quality and minimizes noise.
2. Novel calibration method integrated with the viewer tracking system, provides support for multiple stereoscopic users with dynamic viewpoints and arbitrary parallax directions.
3. Development of sample applications and informal evaluation of the complete system for multi-user collaborative tabletop scenarios.

- Quantitative analysis of the display’s performance when supporting one or more users.

To the best of our knowledge, this is the first practical multi-viewer full-color autostereo display supporting tabletop applications.

The next section describes the hardware configuration of our prototype, including our approach for reducing inter-reflections. In Section 4, we describe our hardware-accelerated rendering algorithm. We then discuss our calibration methodology in Section 5 and present first application results in Section 6. Section 7 contains quantitative measures of our display’s performance and is followed by conclusions and an outlook on future opportunities.

3 SYSTEM CONFIGURATION

We developed our RHD system based on a single high-density LCD display panel. We have selected the highest-resolution commercially available and reasonably priced LCD for this purpose (30-inch diagonal, 2560×1600 pixels, $0.25mm$ pixel size, brand: LG). This display appears to be of the in-plane switching (IPS) type, since it has very large horizontal and vertical viewing angles, which makes it ideally suited for tabletop mounting. In contrast, so called twisted-nematic (TN) panels, which are widely used for low-cost consumer-grade LCD displays, have a limited vertical viewing angle and exhibit color inversion when viewed from below.

The random-hole mask was printed on a thin polyester film at 2500 dpi using a Hitachi Linotronic printer in a commercial facility. The mask film is attached to the display surface through the intermediary of a Lexan™ polycarbonate sheet, which forms the separating layer. The thickness of the sheet is $6.35mm$. The Lexan’s refractive index is slightly above 1.5 and similar to the index of the LCD panel’s built-in transparent cover. The polyester mask film is non-permanently attached to the barrier surface with Karo-syrup based adhesive, so it is relatively easy to adjust if necessary.

Our experiments showed that this design and combination of materials solves the problem of internal reflection between the mask and the display surface [11] in a satisfactory manner.

The display is mounted flat on a table at a height of $85cm$.

We deployed a commercially available and budget-friendly OptiTrack™ infrared optical tracking system with passive markers for user tracking and for overall geometric calibration. In our configuration, the tracking error is within $1cm$.

We used the Poisson disk sampling algorithm [4] to generate the mask pattern. We decided not to align the hole positions with the pixel grid of the display, in order to reduce possible artifacts due to the sub-pixel problem (explained below). Figure 2 shows close-ups of the corner of the barrier pattern, with marks for alignment. The hole density is a parameter that requires tuning. With lower hole density, the probability of a pixel becoming visible to different viewers (through different holes) is reduced, that is, the conflict rate is reduced, resulting in less crosstalk. However, for a given display panel, lower hole density also means lower effective resolution to each viewer, and lower brightness. This problem can be alleviated with higher-density and higher-brightness LCD/LED panels. For the currently available display hardware described above, we chose a hole density of $1/13.2$, meaning that the ratio between the number of display pixels and the number of holes in the mask is 13.2 . This choice is based on our (subjective) evaluation of the quality of synthetic images of a simulated RHD, by balancing the trade-off of single view brightness, effective resolution and the ratio of conflicting pixels.

4 RENDERING

In this section, we introduce the rendering method for our random-hole autostereo display. Figure 3 shows a close-up of an LCD Display. Each pixel contains a red, a green and a blue sub-pixel. This

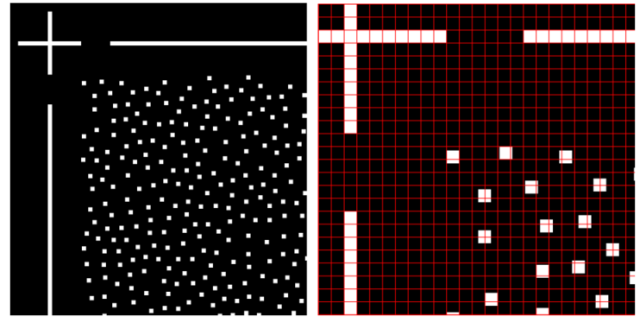


Figure 2: Left: Close-up of a corner of the barrier pattern with randomly distributed transparent squares(white), and margin/corner marks for alignment of the barrier. Right: Further enlargement of the corner area with superimposed pixel grid(red); note that while the alignment marks match the pixel grid, the randomly distributed transparent squares do not.

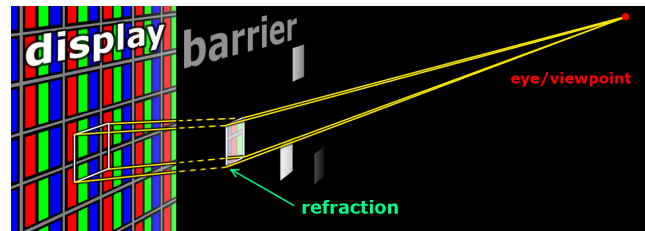


Figure 3: LCD display panel with pixel grid showing red,green and blue sub-pixels, as well as random hole barrier. Determining which part of the pixel grid is visible through each hole from a particular viewpoint requires a projective geometry calculation that takes into account refraction through the separating layers.

configuration is typical for most LCD displays, including the four-megapixel LCD panel we use. From a specific viewing position (corresponding to one of a user’s eyes), a certain area of the pixel grid is visible through each hole in the mask. As Figure 3 shows, the exact shape and location of that area can be calculated by projecting the hole onto the pixel grid, with the viewpoint as projection origin. A number of layers (the barrier film, the Lexan separating layer, and the LCD panel’s own protective cover) are located between the actual pixel grid and the mask, and their refractive effects must be taken into account.

4.1 The subpixel problem

In general, the display area visible through each hole is not aligned with the LCD pixel grid (Figure 3, right), therefore a subpixel-based rendering technique is required in order to produce the correct color for each view. The color of each individual subpixel can be precisely calculated based on the fraction of its visible area, as long as this visible area itself can be precisely calculated. However, this method requires very-high-accuracy eye position tracking; the accuracy is given by:

$$e_{tracking} \leq \frac{(z - g)pe_{color}}{3g},$$

where z is the viewing distance (from the barrier), g is the gap between the barrier and display surface, and p is the pixel width. For example, if the maximum color error is limited to 10%, then the tracking error should be no more than $1.3mm$ when the user is $1m$

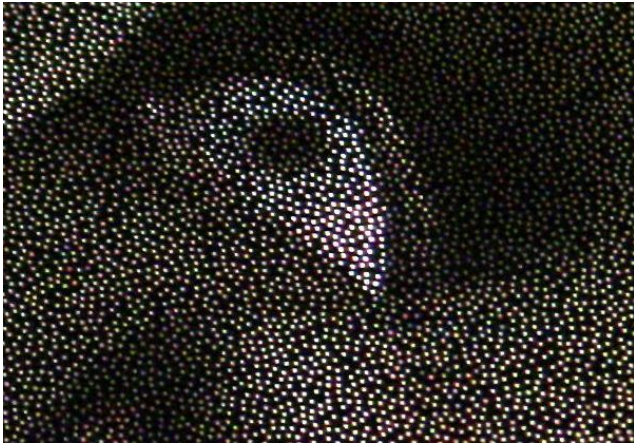


Figure 4: Close-up photo of the tabletop RHD captured by a (tracked) camera shows the randomly distributed holes in the mask and the pre-filtered image.

away from the display. Our tracking system’s accuracy is on the order of $5 - 10mm$. To compensate for the sub-pixel problem as well as for the tracking error, the rendering algorithm fills all four partially visible (through each hole) pixels (see Figure 3, right) with the same desired color for a given viewpoint.

4.2 Image pre-filtering

For the current barrier configuration, the RHD permits no more than 7.6% of the display surface to be visible to each view. For our system, the effective resolution is approximately 705×440 . The RHD actually down-samples the display image, therefore image pre-filtering is required to reduce aliasing. A simple technique for accomplishing this is to generate down-sampled images in the framebuffer. Instead of generating images at full resolution for the whole display grid, filtered images are rendered; their ratio r is given by:

$$r = \sqrt{d},$$

where d is the hole density; $d = 13.2$ for our system. This is equivalent to applying a box filter with kernel size $1/r$ for image pre-filtering. Figure 4 shows close-up imagery provided by the tabletop RHD to a single view point.

4.3 Blending between multiple views

With a random hole barrier, it frequently happens that some display pixels are visible to more than one viewer. We call such pixels conflicting pixels or conflicts. Thanks to the pseudo-random Poisson distribution of the hole pattern, such conflicts are distributed across the entire display. The total number of conflicts varies dynamically and depends on the display configuration and on the number of viewers and their positions.

The single-view rendering algorithm synthesizes the proper image for each single view and does not take conflicts into account. Therefore, we developed a multiple-view blending algorithm that offers a good solution for conflicts and reduces noise and color error.

One simple strategy is to take the average of all visible views’ desired colors [11]. In addition to this, we investigated several additional strategies (e.g. rendering the conflict pixel as black, or randomly choosing the color of one of the conflicting view) for color blending. Any of these will result in some color discrepancy for the conflicting pixels.

In a given view, each hole with a conflicting pixel shows a reconciled color instead of the true color for that view. We developed a special neighborhood error diffusion algorithm to diffuse the error to the neighboring pixels visible to the same viewer, as Figure 5 shows. Our error diffusion method tries to adjust the color of neighboring visible pixels to compensate for the color discrepancy of conflict pixel; using the average color as the blending strategy makes the error diffusion perform more smoothly over all views.

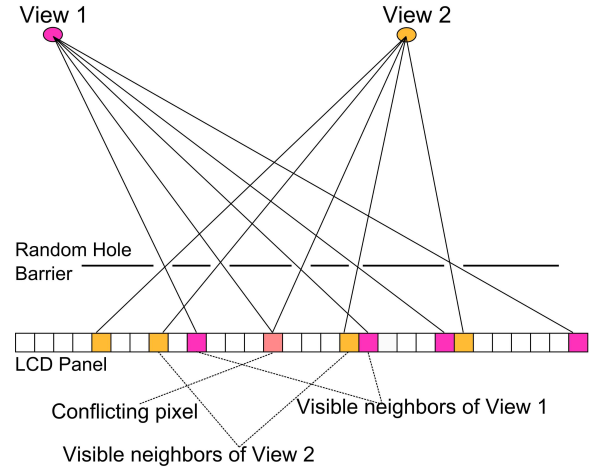


Figure 5: A conflicting pixel and its visible neighboring pixels for two independent views.

4.4 Hardware-accelerated algorithm

We have developed a hardware-accelerated rendering algorithm consisting of four rendering passes. Figure 6 shows the pipeline of the entire rendering algorithm. This rendering algorithm incorporates our a new error diffusion method.

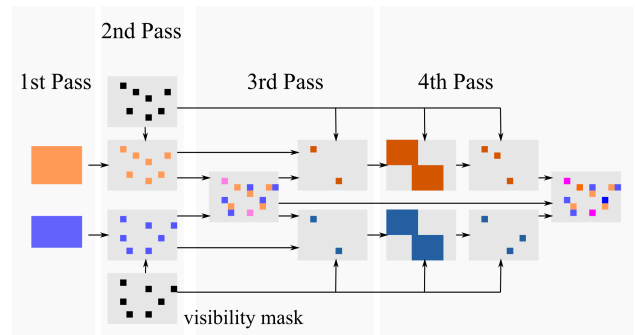


Figure 6: The pipeline of the 4-pass rendering algorithm. (The color of the pattern does not represent the true value of the processing data.) 1st pass: The image of each view is rendered to down-sized frame buffer. 2nd pass: The visibility mask is generated for each view and applied to the original image. 3rd pass: All masked images are combined, the blended color is computed for the conflict pixels, and an error map is generated for each view. 4th pass: The error is diffused to the visible neighboring pixels for each view.

4.4.1 First pass

This pass renders the scene from each viewer’s perspective to the frame buffer, at a resolution which is a fraction r of the native reso-

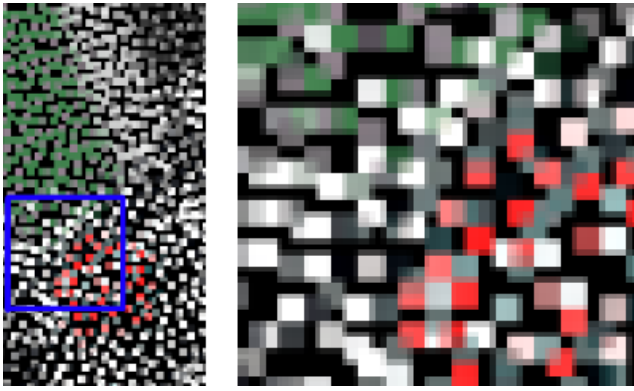


Figure 7: A part of the blended and error-diffused image actually displayed on the LCD panel. Error diffusion makes some gray points near the red pixels appear dark cyan color.

lution of the LCD panel.

4.4.2 Second pass

This pass generates the point images (the subset of visible points in each image), by calculating rays from the viewpoint through the holes (the hole position could be calculated according to calibration, described in Section 5, and to the known position of each hole in the mask.) and onto the display. This visibility mask is recorded in the alpha channel. In order to reduce visual artifacts in subsequent processing, the colors are converted from RGB color space to CIE-LUV color space [18], so that linear color blending and error diffusion can more closely approximate human eye perception.

In our rendering algorithm, we incorporated the following assumptions about the physical properties of the display:

1. The separating layer is of uniform thickness.
2. The size of the holes is tiny compared to the viewing distance.
3. The thickness of the separating layer is small compared to the viewing distance.
4. There is only a single layer of refractive material with a fixed refractive index.

In experimentation, we discovered that the assumption of a single refractive layer does not hold for a large range of different viewing angles to the display. We address this issue with our calibration method.

4.4.3 Third pass

This pass blends the point images of all views into a single texture; for the conflict pixels, we take the average as the blended color.

This pass also calculates the error between the blended color and the true color of the point image for each of the views; this could be calculated in a single pass by using multiple rendering targets. The resulting error map is subsequently used for approximate dithering in the fourth pass (see below).

4.4.4 Fourth pass

This pass diffuses the color error of each pixel of each view to the neighboring visible pixels in the same view. Note that the neighboring visible pixels are not the same as the neighboring pixels on the LCD panel (Figure 5).

We designed a dithering algorithm suitable for GPU-based acceleration. We assume that the number of neighboring pixels in a $w \times w$ area centered at the processing pixel is given by:

$$n_{neighbor} = \frac{w^2 d(z-g)}{z} - 1 \approx w^2 d - 1 \quad (g \ll z)$$

where d is the hole density (1/13.2 in our case), z is the viewing distance, and g is the gap between barrier and display.

We “splat” the error of each pixel onto its visible neighboring pixels. This could be implemented by adding to the blended image a certain level of the mipmap of the error map, masked by the visibility map. In order to make the error diffused to each neighboring pixel equal to e/n , before generating the mipmap, the value of the error map is multiplied by the approximate density d . We use alpha blending to accumulate the mipped and masked error map into the final blended image, and then convert the color from CIE-LUV back into RGB color space. Figure 7 shows a part of a final blended image. The implementation is simple and well suited for GPU acceleration. In our experiment, the post-processing time (including the 2 to 4 rendering passes) of four views, is under 10ms on a desktop machine with Intel i7 CPU and NVIDIA GTX 295 graphics card. Figure 8 shows the image quality improvement of the error diffusion process for the conflict pixels.

4.5 Priority rendering

In some applications, it could be useful to provide certain viewers with better image quality, i.e. less crosstalk. For example, in teleconferencing, currently active speakers could have better image quality for better interaction such as eye contact. Therefore, another strategy to deal with conflict pixels is to choose the color from one of the conflict views randomly, and make the probability of choosing a particular view proportional to the viewer’s current priority. This strategy could result in image degradation for viewers with lower priority.

5 CALIBRATION

User tracking is necessary for dynamic viewpoints. Our system must know the 3D positions of all users’ eyeball centers in order to generate correct imagery. We built a wearable head tracker consisting of a lightweight eyeglass frame and two tracking markers (retro-reflective spheres) positioned at the wearer’s temples. The spheres can be moved forward or back, so as to be aligned with the line connecting the centers of the two eyeballs. The interpupillary distance (IPD) of the user is measured and entered into system, then the position of the eyeballs can be calculated based on the positions of the tracked marker spheres, assuming that eye balls and marker spheres are all collinear.

In addition to the calibration of the tracking system itself, there are other system parameters that require calibration:

- The offset and rotation of the barrier relative to the display, caused by imprecise mounting.
- The refractive index and thickness of the separating layer.

Offset and rotation are calibrated by positioning a high-resolution camera along the axis through the center of the display panel, while presenting uniform green imagery with the current parameters and refining them in a random-walk process until the maximum overall brightness is achieved in the camera image. Since the relative position of each hole on the laser-printed mask (printing resolution 2500 DPI) is precisely known, with the calibrated offsets and rotation, each hole’s position in tracking space can be calculated and used in the second pass of the rendering.

We could also use the same method to calibrate the refractive index and thickness for a given location, with which the system can produce correct imagery with the error-tolerant rendering algorithm described above. However, since we assume that there is only one layer introducing the major refractive effects and we ignore the

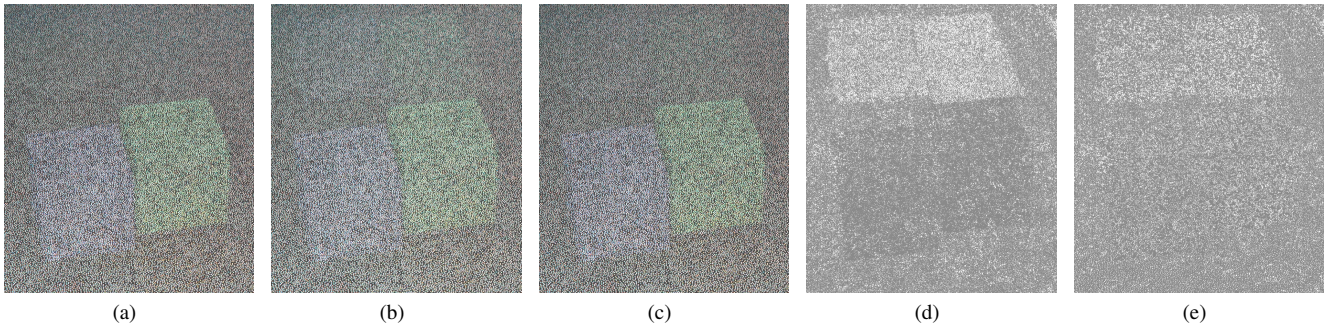


Figure 8: Comparison of actual photos taken by a fixed camera of the Cubes data set, with and without error diffusion. Above the cuboid, crosstalk from another view is noticeably reduced. (a) Single view, therefore no crosstalk (baseline). (b) Two views, without error diffusion. (c) Two views, with error diffusion. (d) Absolute difference between (b) and (a). (e) Absolute difference between (c) and (a). Comparing (d) and (e) shows that error diffusion effectively reduces the structural noise caused by crosstalk from other views. See Section 7.1 for more discussion. (The brightness and contrast of images (a),(b),(c) were adjusted; (d),(e) were enhanced through histogram equalization).

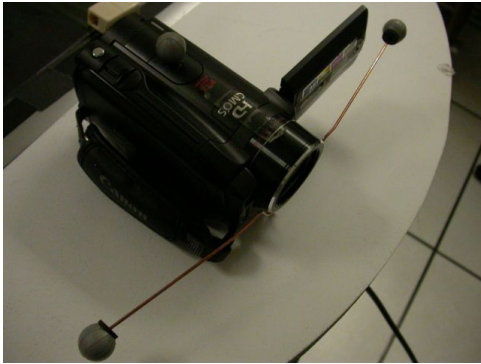


Figure 9: Camcorder with retro-reflective markers mounted onto it.

other layers, the calibrated results at different view positions are not consistent, which is exacerbated by tracking errors. While we plan to work on precise calibration of a more sophisticated physical model of the barrier (e.g. multiple refractive separating layers), for now we use a simple distance weighting algorithm[15] to interpolate the calibrated parameters at different view positions, as given by

$$u(\mathbf{x}) = \sum_{k=0}^N \frac{w_k(\mathbf{x})}{\sum_{i=0}^N w_i(\mathbf{x})} u_k,$$

where u_k is the vector of calibrated parameters sampled at x_k and

$$w_k(\mathbf{x}) = \frac{1}{\|\mathbf{x} - \mathbf{x}_k\|^2}.$$

In practice, for our configuration, only the neighboring sampled points x_k within a sphere of radius $20cm$ are considered in the interpolation. If there is no sample point within this space, the calibration parameter of the nearest sampled point is used. More sophisticated interpolation methods could be used here; they are expected to require fewer calibrated points to get a good calibration result.

In order to perform fast adaptive incremental calibration, we use a high-resolution camera with tracking markers mounted onto it, as Figure 9 shows. The center of projection of the camera lens is pre-calibrated and its spatial position is tracked.

We perform incremental calibration to sample a number of calibrated viewpoints distributed throughout the viewing area of the

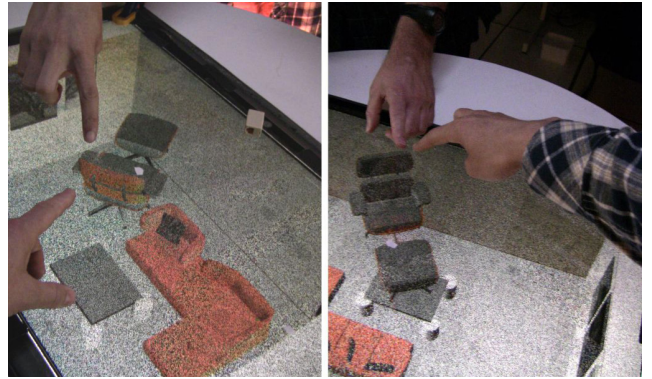


Figure 10: Interior design application. The two photos were taken at the same time by two tracked cameras at different locations. The two views are generated by the tabletop RHD simultaneously. Two users point at the same part of a virtual chair located on the surface of the display; each user sees a perspective correct view.

tabletop RHD display. As the rendering program runs in calibration mode, the camera is positioned at a location where the imagery is not correct according to the current interpolated calibration result. Then a calibration point is created at that location, and the calibration parameters at that point are adjusted until good imagery is obtained. Then the interpolation grid is updated and the camera is moved to next area where the calibration result is not correct. About 40 points are sampled to cover the whole tracking and viewing area in our experiment. Figures 10 and 11 shows the image quality of the calibrated RHD.

6 RESULTS

6.1 Applications

With tracking, real-time rendering, and a calibrated RHD, we implemented several interesting and effective applications for multi-user interactive collaboration in a tabletop environment. These include collaboration on interior design, shown in Figure 10, virtual and real objects manipulation, shown in Figure 1, and sandbox-like 3D cityscape navigation, shown in Figure 11. Figure 11 is a set of stereo pair images, showing the autostereoscopic image quality provided to a single user.

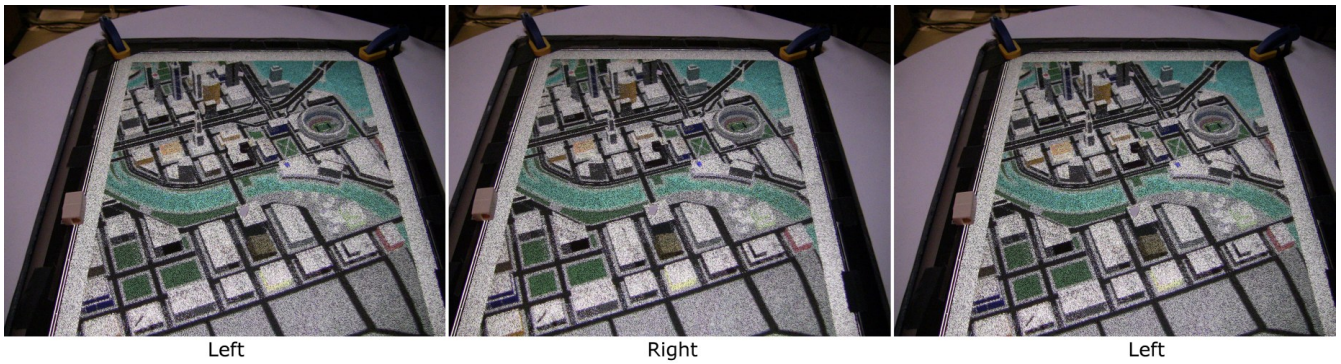


Figure 11: (Fusible) stereo pair generated by the tabletop RHD simultaneously. The uniform image quality shows the general correctness of tracking and calibration.

6.2 Crosstalk

In the configuration mentioned in Section 3, when a single user of height 170cm with an IPD of 6.5cm is standing 65cm away from the center of the tabletop display, 1.82M pixels are visible, of which 1.37M pixels are conflict-free and 0.55M pixels (28.5% of all visible pixels) are visible to both eyes at the same time. However, because of view continuity of the stereoscopically displayed 3D objects, the actual color error among these conflicting pixels is usually small. For example, for the head models scene shown in Figure 13(a), only 47K pixels (2.6% of all visible pixels) have a color error larger than 10% for at least one of the eyes.

When two viewers are present, and if they are positioned at opposite sides of the display, which is considered a worst-case situation in terms of crosstalk in many 3D applications, the total number of visible pixels increases to 2.84M . The number of conflict-free pixels is 50% of all visible pixels; 35% of visible pixels are visible to two eyes, 12% are visible to 3 of the 4 eyes. 380K conflicting pixels (13.4% of visible pixels) have a color error larger than 10% for at least one of the 4 eyes. Figure 13(d) shows this situation. We evaluate such image quality degradation in Section 7.2. Nevertheless, our display is superior to other designs supporting two-dimensional parallax (such as lenslets) since it is capable of dynamically allocating a larger fraction of the display’s native pixels to each of the currently required viewpoints.

7 EVALUATION

To quantitatively analyze the performance of our display, we separately measured the image degradation caused by adding viewpoints, as well as the image quality improvement obtained through our error diffusion technique.

7.1 Image quality improvement through error diffusion

In order to evaluate the effect of our error diffusion method, we set up an experiment for quantitative analysis. We mounted a 10MP DSLR camera on a fixed stable tripod; the camera was controlled by a computer through a USB cable. We repeated the experiment on 4 data sets with different 3D geometry. For each data set, we acquired pictures of the display imagery from the camera in three different scenarios: 1. Rendering single-view imagery from the viewpoint of the camera. 2. Rendering two-view imagery without error diffusion, one from the viewpoint of the camera, the other one from the opposite site of the tabletop display, so as to introduce significant crosstalk. 3. Rendering with the same configuration as in Scenario 2, but with error diffusion. The camera was focused on the surface of the display, and in the resulting digital photographs, each hole of the barrier appears as a blob covering approximately 6×6

camera pixels. All pictures were down-sampled by a factor of 10 before processing. For Scenario 1, there is no crosstalk since only one view is generated, hence it is used as the gold standard to assess the image quality of Scenarios 2 and 3. For convenience, we denote the images of Scenario 1, 2 and 3 as I_1 , I_2 , and I_3 respectively.

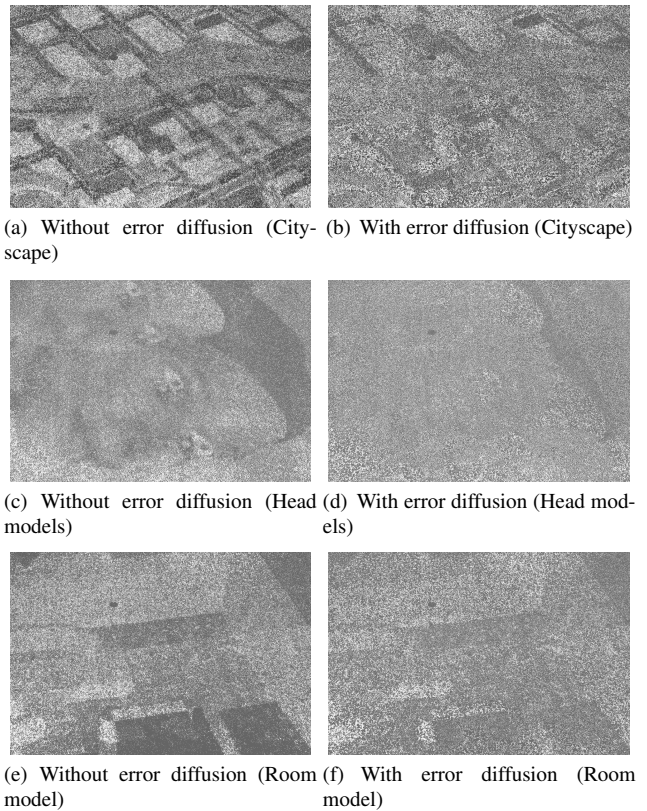


Figure 12: Absolute error differences between one view and two views (enhanced by histogram equalization for visibility). These figures are generated in the same way as Figure 8(d)(e), but with a data set that has higher geometric complexity and more texture details. Note the significant structural noise from the crosstalk of the other view when error diffusion is disabled.

We calculated the brightness difference between I_1 and I_2 , as well as between I_1 and I_3 , and plotted the histograms. Figure 14



Figure 13: Graceful image degradation as the number of views increases.

shows the histogram of the difference between I_1 and I_2 (in red), and the histogram of the difference between I_1 and I_3 (in green). These two histograms are superimposed to emphasize the difference between them. The distribution of the green histogram is slightly left-shifted compared to the red one, which means I_3 has fewer pixels with large errors than I_2 . However, this small shift does not show the major advantage of the error diffusion, since it does not indicate the spatial distribution of the reduced error. The absolute error images shown in Figures 12 and 8(d)(e) visualize this information more clearly (for visibility, the contrast is enhanced through histogram equalization). Comparing the absolute error images with and without error diffusion of each data set, demonstrates that the structural noise caused by crosstalk is effectively reduced by error diffusion.

The error diffusion method was designed based on our (unverified) assumption that with a random hole barrier pattern, the neighboring visible pixels of a conflicting pixel are less likely to be in conflict. Error diffusion does not improve image quality if most of the pixels are in conflict. In order to keep the percentage of conflicting pixels within tolerable bounds, the hole density must be reduced as the number of views increases. As mentioned in Section 3, the current barrier density of $1/13.2$ was designed for a small number of views, taking all the trade-off factors into consideration.

7.2 Image degradation caused by additional viewpoints

We conducted a similar experiment as in Section 7.1 to evaluate the image degradation as the number of views increases from one view to six views. We chose the locations of the 6 viewpoints to approximate a common usage scenario where three users are present around the tabletop display. For each test case, the display generated imagery for one view to six views, and the camera, fixed at the first viewpoint, captured images as the number of generated views was increased. We therefore expected to observe degradation of the image quality due to crosstalk. We use the single-view image as the gold standard for comparison, and computed the mean absolute error (MAE) comparing to the single view scenario (Table 1 and Figure 15(a)); we also calculated the Structural SIMilarity (SSIM) index [17] for assessment (Table 2 and Figure 15(b)), as an evaluation of perceptual image quality. The SSIM index ranges from 0 to 1, where 1 means two compared images are identical. As shown in the figures, the MAE increases and the SSIM decreases with more views, as expected, since the percentage of conflict pixels increases as well (as discussed in Section 6.2), introducing more crosstalk across different views. In general, such degradation is smooth and graceful in our system.

As mentioned in Section 2, our display's effective resolution is approximately 705×440 or 310k pixels; this is the resolution avail-

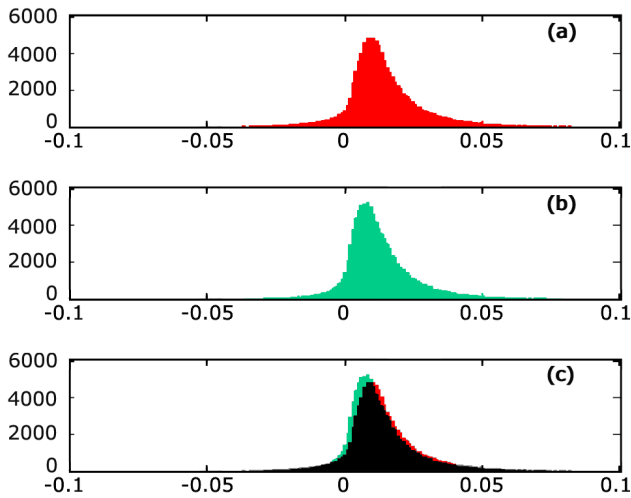


Figure 14: Quantitative comparison of photos of display with/without error diffusion for cityscape data set. (a) Histogram of image error without error diffusion. (b) Histogram of image error with error diffusion. (c) Superimposition of (a) and (b): The distribution of (b) is slightly left shifted, meaning it has fewer pixels in large error bins and more pixels in small error ones. However, this figure does not indicate the difference of the spatial distribution of the error, which is shown more clearly in Figure 12.

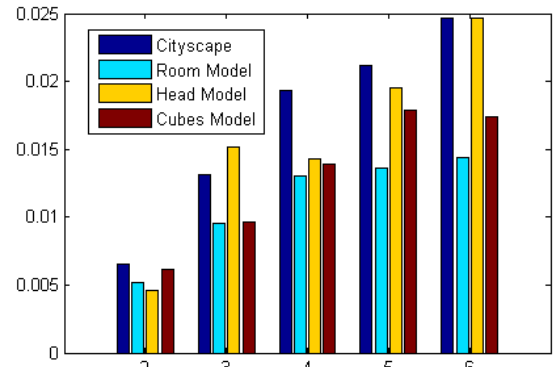
Table 1: Mean Absolute Error compared to single view.

number of views	Cityscape	Head Model	Room Model	Cubes Model
2	0.0065	0.0052	0.0046	0.0062
3	0.0131	0.0095	0.0152	0.0096
4	0.0193	0.0130	0.0143	0.0139
5	0.0212	0.0136	0.0195	0.0179
6	0.0247	0.0144	0.0247	0.0174

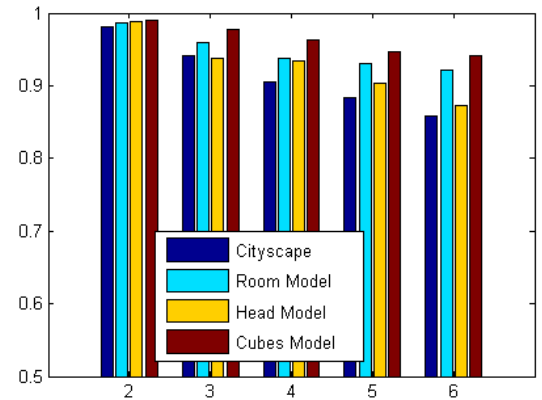
able to a unique viewer, regardless of display complexity (similarly to any conventional computer display, except that our samples are randomly positioned). The maximum viewing angle is at least $\pm 70^\circ$ in each direction. Once additional views must be supported, this effective resolution decreases because of conflict pixels. Furthermore, when viewing imagery with high spatial frequencies, the image quality may decrease more rapidly than with imagery containing large areas that are uniformly colored or have smooth color gradients (such as untextured synthetic imagery); this could explain why the MAE increases differently for different data sets as shown in Figure 15(a). To be more specifically, the room model is relatively simple, with an area of uniformly colored floor, whereas the cityscape has more complex geometry and textured ground, causing the MAE to increase faster. Interestingly, when more than 2 views are present, a portion of the crosstalk from a particular view might occasionally have coincident alignment with the current view. In this situation, such crosstalk could actually reduce the error of current view. This can be the reason why the MAE does not always increase as the number of views increases in some test cases (e.g. comparing 2 views and 3 views of Head model, as well as 5 views and 6 views for the Cubes model).

8 CONCLUSIONS AND FUTURE WORK

We introduced a practical multi-viewer tabletop autostereoscopic display. The selection of construction materials and the design of the separating layer and mask film effectively reduce the



(a) Mean absolute error



(b) SSIM

Figure 15: Graceful image quality degradation as the number of views increases. (a) Mean absolute error. (b) Image quality measured by the Structural SIMilarity (SSIM) index [17].

Table 2: Structural SIMilarity index compared to single view

number of views	Cityscape	Head Model	Room Model	Cubes Model
2	0.9817	0.9862	0.9884	0.9907
3	0.9407	0.9601	0.9384	0.9773
4	0.9050	0.9384	0.9333	0.9628
5	0.8840	0.9300	0.9031	0.9467
6	0.8586	0.9211	0.8735	0.9422

inter-reflection present in the original prototype RHD. The real-time hardware-accelerated rendering algorithm provides interactive stereoscopic imagery for multiple users. It includes a post-process for error diffusion, which is an effective approximation of image dithering over pseudo-randomly distributed points; this process noticeably reduces crosstalk from conflict pixels. We also analyzed the tracking requirements for the RHD, and proposed a sampling-based calibration method for the imperfect, “real-world” tracking system, and for the complex refraction properties of the display.

In the future, we expect to use brighter and higher-density LCD/LED panels or high-resolution multiple projectors system as display surfaces. As the pixel density of the display increases, the absolute hole density of the barrier could be raised as well, so that the barrier mask would eventually become almost imperceptible to

human eyes. Moreover, with the same effective resolution, a higher density display surface would allow the relative hole density of the barrier to be reduced, which could effectively decrease the conflict rate. We are also interested in investigating dynamic random hole barriers, which could potentially increase the effective resolution and provide optimal hole patterns that dynamically minimize conflicts.

ACKNOWLEDGEMENTS

The authors would like to thank Kurtis Keller for his ideas and help on construction of the barrier film as well as of the separating layer, Tao Li for help with tracking system configuration, John Thomas and Herman Towles for help with tabletop display construction, and Andrew Nashel for suggestions on rendering issues. We also thank Stephen M. Pizer and Marc Pollefeys for insightful discussions on comparing image quality. We are also grateful to the anonymous ISMAR 2010 reviewers and submission coordinator for their critique and help on significantly strengthening this work. The research we presented here is based in large part upon prior work supported by the National Science Foundation under Grant No. CNS-0751187.

REFERENCES

- [1] T. Agocs, T. Balogh, T. Forgacs, F. Bettio, E. Gobbetti, and G. Zanetti. A large scale interactive holographic display. In *Proc. IEEE VR 2006 Workshop on Emerging Display Technologies*, Alexandria, VA, USA, 2006.
- [2] T. Balogh, Z. Dobranyi, T. Forgacs, A. Molnar, L. Szloboda, E. Gobbetti, F. Marton, F. Bettio, G. Pintore, G. Zanetti, E. Bouvier, and R. Klein. An interactive multi-user holographic environment. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Emerging technologies*, page 18, New York, NY, USA, 2006. ACM.
- [3] B. Blundell and A. Schwartz. *Volumetric Three-Dimensional Display Systems*. Wiley-IEEE Press, 1999.
- [4] R. Bridson. Fast poisson disk sampling in arbitrary dimensions. In *SIGGRAPH '07: ACM SIGGRAPH 2007 sketches*, page 22, New York, NY, USA, 2007. ACM.
- [5] M. W. Halle, S. A. Benton, M. A. Klug, and J. S. Underkoffler. The ultragram: A generalized holographic stereogram. In *Master's thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology*, pages 142–155, 1991.
- [6] A. Jones, I. McDowall, H. Yamada, M. Bolas, and P. Debevec. Rendering for an interactive 360° light field display. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 40, New York, NY, USA, 2007. ACM.
- [7] Y. Kim, K. Hong, and B. Lee. Recent researches based on integral imaging display method. *3D Research*, 1(1):17–27, 2010.
- [8] Y. Kitamura, T. Konishi, S. Yamamoto, and F. Kishino. Interactive stereoscopic display for three or more users. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, page 240. ACM, 2001.
- [9] H. Liao, M. Iwahara, T. Koike, N. Hata, I. Sakuma, and T. Dohi. Scalable high-resolution integral videography autostereoscopic display with a seamless multiprojection system. *Applied optics*, 44(3):305–315, 2005.
- [10] W. Matusik and H. Pfister. 3d tv: A scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Transactions on Graphics*, 23:814–824, 2004.
- [11] A. Nashel and H. Fuchs. Random hole display: A non-uniform barrier autostereoscopic display. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2009*, pages 1–4, 4-6 2009.
- [12] K. Perlin, S. Paxia, and J. S. Kollin. An autostereoscopic display. In *SIGGRAPH '00: ACM SIGGRAPH 2000 Papers*, pages 319–326, New York, NY, USA, 2000.
- [13] T. Peterka, R. L. Kooima, J. I. Girado, J. Ge, D. J. S. A. Johnson, J. Leigh, J. Schulze, and T. A. Defanti. Dynallax: Solid state dynamic parallax barrier autostereoscopic vr display. In *In Proceedings of IEEE Virtual Reality 2007*, pages 155–162. IEEE, 2007.
- [14] D. J. Sandin, T. Margolis, J. Ge, J. Girado, T. Peterka, and T. A. Defanti. The varrier™ autostereoscopic virtual reality display. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pages 894–903, New York, NY, USA, 2005. ACM.
- [15] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *ACM '68: Proceedings of the 1968 23rd ACM national conference*, pages 517–524, New York, NY, USA, 1968. ACM.
- [16] H. Takahashi, H. Fujinami, and K. Yamada. Wide-viewing-angle three-dimensional display system using hoe lens array. In *Proceedings of the SPIE, the Stereoscopic Displays and Virtual Reality Systems XIII*, volume 6055, pages 453–461, 2006.
- [17] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [18] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae, Second Edition*, chapter 3.3.9, pages 165–166. Wiley-Interscience, 2000.