

Pepper's Cone: An Inexpensive Do-It-Yourself 3D Display

Xuan Luo¹

xuanluo@cs.washington.edu

¹University of Washington
Seattle, WA, USA

Jason Lawrence²

jdlaw@google.com

²Google Inc.
Seattle, WA, USA

Steven M. Seitz^{1,2}

seitz@cs.washington.edu

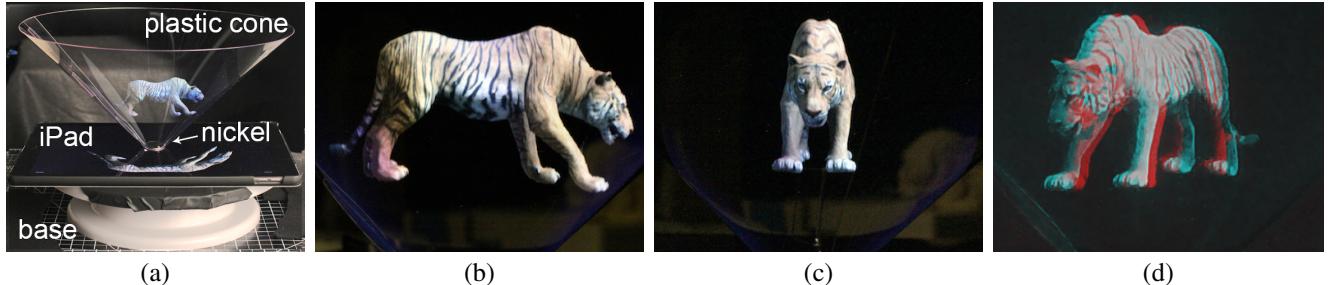


Figure 1. Our 3D display consists of (a) an iPad, a thin hollow plastic cone, and a rotatable base. The nickel at the base provides stability. (b-c) As the user rotates the display, the system renders a perspective-correct image for their point of view that gives a convincing impression of a 3D object suspended inside the cone. This provides a simple way of interactively examining a 3D scene for a fraction of the cost of alternative volumetric or light field displays and doesn't require the use of special glasses. (d) The system can be extended to produce correct binocular cues by incorporating stereoscopic rendering and glasses.

ABSTRACT

This paper describes a simple 3D display that can be built from a tablet computer and a plastic sheet folded into a cone. This display allows viewing a three-dimensional object from any direction over a 360-degree path of travel without the use of special glasses. Inspired by the classic Pepper's Ghost illusion, our approach uses a curved transparent surface to reflect the image displayed on a 2D display. By properly pre-distorting the displayed image our system can produce a perspective-correct image to the viewer that appears to be suspended inside the reflector. We use the gyroscope integrated into modern tablets to adjust the rendered image based on the relative orientation of the viewer. Our particular reflector geometry was determined by analyzing the optical performance and stereo-compatibility of a space of rotationally-symmetric conic surfaces. We present several prototypes along with side-by-side comparisons with reference images.

ACM Classification Keywords

H.5.1. [Information interfaces and presentation]: Multimedia Information Systems – Artificial, augmented, and virtual realities; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism – Virtual reality

Author Keywords

3D Displays; Pepper's Ghost; Computer Graphics.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UIST 2017, October 22–25, 2017, Quebec City, QC, Canada

© 2017 ACM. ISBN 978-1-4503-4981-9/17/10...\$15.00

DOI: <https://doi.org/10.1145/3126594.3126602>

INTRODUCTION

People have long been fascinated by the promise of true “holographic” or “volumetric” displays. Indeed, reproducing the appearance of an object in a way that allows viewing it from any direction, without the aid of a head-mounted display or special glasses, is the most natural way to convey its three-dimensional shape. However, delivering this type of experience is challenging. There are currently no consumer displays that achieve this type of interaction and existing glasses-free volumetric and light field displays involve complex and expensive setups that still fall short of this lofty goal.

We introduce a very simple “3D” display based on the *Pepper's Ghost illusion*, a 19th century stage technique named after John Pepper [6]. This effect involves projecting a person or object from a hidden room or screen onto a physical stage through the use of a large half-silvered mirror or other reflective surface. This technique has found a recent resurgence in popularity [1, 23] and is used for high-profile “live” stage performances [8, 12].

While this illusion traditionally involves a planar reflector, we introduce a variant based on a curved reflector and implementation on a gyro-enabled commodity tablet PC. This enables the user to interactively view the displayed object over a full 360-degree rotation and produces the sensation that it is suspended above the tablet. Supporting curved reflectors requires a calibration and pre-distortion process to ensure that the image as viewed through the display is perspective-correct. We call our display *Pepper's Cone*.

Shown in Figure 1, our display consists of a truncated plastic cone resting on a conventional flat 2D display. When viewed from the side, the user sees a distorted reflection of the display that appears to be located *inside* the reflector. By pre-

warping the image on the flat display in order to compensate for the distortion introduced by the reflector, we can deliver a perspective-correct image at a desired viewpoint. Finally, by querying the tablet's gyroscope, we allow the user to smoothly and interactively view a three-dimensional scene from any direction over a full 360-degree path by physically rotating the display in front of them.

To arrive at the cone shape, we analyzed the space of rotationally-symmetric conic surfaces (cones, paraboloids, ellipsoids, hyperboloids). Specifically, we developed quantitative measures of the desirable optical properties of the reflector surface and performed an optimization to find the best reflector for a specific 2D display. In the case of an iPad Pro, we recommend using a 51° cone.

We also performed a study of binocular stereopsis cues for this class of displays. One of the biggest challenges with 3D displays is achieving proper stereopsis; indeed, we have found that curved reflectors can produce images at each eye that are uncomfortable or impossible to fuse. To this end, we included a measure of stereo divergence in our analysis, measured as the degree of vertical parallax between the user's two eyes, and seek reflector shapes that minimize this effect. The Pepper's Cone produces a fuseable image over most of its surface, with the perceived object (based on binocular cues) appearing to lie just inside the cone.

While our display is stereo-compatible, there remains a mismatch between binocular and monocular cues, which can reduce the effect of depth perception. We show that this remaining conflict can be eliminated by introducing stereo glasses and projecting separate images to each eye. Keeping with our theme of an inexpensive do-it-yourself setup, we demonstrate a prototype that uses paper Red-Blue anaglyph glasses. We call this modification *binocular correction*.

Our 3D display can be built in an afternoon for a tiny investment beyond the tablet computer. We have built 3D displays using reflectors as common as a plastic cup (Sec. 6.2). It can be used for a number of applications in education, medicine, art, product exhibition, telepresence, and virtual assistants. This paper describes how to build a Pepper's Cone display, analyzes the design space and introduces binocular correction. We show results on both synthetic and captured objects.

PRIOR WORK

This paper is related to a number of fields ranging from volumetric and light field displays to head-mounted virtual reality (VR) and augmented reality (AR) systems. Although our simple 3D display does not support the full range of head motion, it provides a compelling "look around" 360-degree viewing experience at a fraction of the cost and complexity of alternative systems without requiring a head mount or special glasses.

Volumetric displays achieve a 3D volume of individually addressable display elements through various means. One approach is to synchronize a projector or other conventional display with a 2D screen or mirror that is swept over a 3D volume at high speed [27, 28]. Due to persistence of vision, the user perceives a stationary 3D entity. Another approach is to use a projector to selectively illuminate a volume composed

of many small optical scatterers, like fog, falling drops of water [3], scatterers produced through laser etching [25], etc. Although volumetric displays provide correct stereo and parallax cues and can accommodate multiple simultaneous viewers, few can reproduce proper occlusions and view-dependent effects like specular highlights. They also often suffer from low resolution due to the higher bandwidth required to refresh a volume of display elements at an acceptable frame rate. In contrast, our setup leverages the high resolution of modern tablet displays and can accurately reproduce view-dependent effects (e.g., the shiny surface of the car in Figure 13).

Holographic displays use laser interference to reproduce a continuous light field. Researchers have developed holographic displays using a variety of materials and optical designs [21, 18, 29, 14, 33, 5]. Nevertheless, widely available high resolution and high frame rate holographic displays are still years away.

Head-mounted displays can achieve a stunning sense of immersion [31, 7]. Only recently has the necessary combination of low-latency displays and mobile high-bandwidth GPUs enabled large-scale consumer VR efforts such as the *Oculus Rift*, *HTC Vive*, and *Google Cardboard*. Augmented reality (AR) displays such as the *Microsoft Hololens* overlay imagery onto the physical world through the use of micro projectors and mirrors or transparent waveguides. Although our display does not achieve the same expansive viewing volume as many of these systems, it also doesn't require the use of head mounts or special glasses, providing a more natural user experience.

Autostereoscopic and light field displays are able to present a correct left/right stereoscopic image pair to a viewer over some viewing range without the need of special glasses [9]. Bonding a spherical lenslet array [24], cylindrical lenticular array [22], or parallax barrier [15] onto a conventional high-resolution 2D display is a popular approach. Another option is to combine multiple projectors using a reflective or transmissive screen that has a very narrow scattering profile [2, 22, 17]. Xia et al. [37] use light field generation to achieve a 360-degree surround viewable volumetric display with proper occlusion. Jones et al. [16] combine a fast spinning slanted anisotropic mirror with a synchronized projector to reproduce a light field that can be viewed from any angle.

An important benefit of many of these methods is that they show the proper image to the viewer's left and right eye. Some can also simultaneously accommodate multiple viewers. On the other hand, these displays are significantly more expensive and complex to create than our simple "do-it-yourself" approach. Further, we show how our display can achieve correct stereoscopic views by incorporating traditional stereoscopic rendering methods (Section 5).

More recently, researchers have explored stacked arrangements of printed transparencies [35] and planar LCD panels [19, 36]. Although these systems can theoretically reconstruct a light field with proper focus cues, they do not allow examining an object from all sides, and current prototypes are either limited to static scenes or are more complex and expensive than ours.

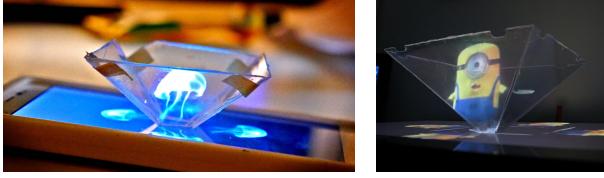


Figure 2. Two do-it-yourself "holograms" with pyramid reflectors. These prior systems do not perform the necessary distortion to achieve perspective-correct views and use non-ideal reflector shapes that have undesirable seams and are difficult to fuse at oblique viewing angles.

Pepper’s Ghost displays refer to the use of a centuries old stage technique that causes a displayed object to appear to float in air. Our work was inspired by the commercial Holus display [13] along with other related “do-it-yourself hologram” videos on YouTube¹ (Figure 2). These systems employ a Pepper’s Ghost technique using a tablet computer and a four-sided pyramid. Although the effect is surprisingly compelling, such systems have a number of drawbacks due to their use of planar mirrors. First, the seam along the edge of each face disrupts the viewing experience and increases sensitivity to calibration errors and viewing position. Second, the 3D illusion often fails when viewed from oblique angles since, for such cases, it is normally impossible to approximate correct stereoscopic cues.

Movie-Maps [20] and Anamorphicons [30] deploy curved mirrors placed on a flat display to achieve a similar anamorphic effect. The Movie-Maps system displays distorted panoramic images that are meant to be viewed through a conical mirror. Anamorphicons uses a cylindrical mirror in conjunction with a tablet computer in a way that allows rotating the mirror itself in order to rotate the displayed object. However, neither of these systems properly pre-distort the images on the flat display to account for the distortion from the curved reflector nor do they consider the same type of semi-transparent display surfaces that we explore. Another paper closely related to ours is the Virtual Showcase [4], which consists of either a curved or polyhedral reflector placed on top of a large tabletop display. It achieves a 3D effect through the combination of a Pepper’s Ghost illusion, the use of stereo shutter glasses, and a head-tracker. One important limitation of this system is that it requires the user to wear a head-mounted motion tracker and is also a rather complicated and expensive setup. In contrast, our display relies on easily sourced materials and can be assembled in only a few minutes. Our system also achieves a smooth viewing experience during object rotations without the use of glasses or head mounts. Additionally, to the best of our knowledge, our analysis of the optimal reflector shape and the conditions under which comfortable binocular fusion is achieved for these types of displays is original.

THE PEPPER’S CONE 3D DISPLAY

In this section we review the major hardware and software components of our display along with key design trade-offs. Later sections analyze the reflector design space and present experimental results and a preliminary user study.

¹<https://www.youtube.com/watch?v=7YWTtCsvgvg>
<https://www.youtube.com/watch?v=nh91P1Rs1EU>

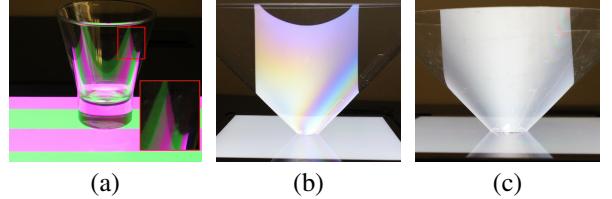


Figure 3. (a) Double reflections caused by a reflector that is too thick (4.2mm). Thin-film interference effects are visible in thinner materials such as (b) 0.5mm PETG plastic and (c) 0.17mm PETG plastic.

Hardware

As illustrated in Figure 1(a), our display consists of a tablet computer, a reflector, and a base that lets the user easily rotate the entire assembly. Although there are potentially many embodiments that meet these specifications, we made a number of observations about important aspects of this design space.

Tablet

Our display includes a tablet computer that integrates a high-resolution display, dedicated real-time graphics hardware, and a low-latency gyroscope. Unsurprisingly, a tablet with a large high-resolution display is desirable as this achieves better overall image quality and creates a larger display volume. Besides, a rendering system that can generate high-quality images at least 30fps is recommended, which seems to be readily obtainable in modern tablet computers using GPUs. Finally, our display uses the built-in gyroscope to track the user’s vantage point. The latency of the gyroscope is another important factor as this directly impacts end-to-end system lag (Section 6.1).

Reflector

The reflector can potentially have a wide range of shapes, sizes and material properties. We explored a number of different options, some of which are shown in Figures 1 and 14, and make a number of useful observations.

Although many different reflector geometries could be used, we focus on surfaces of revolution, as they provide a consistent shape under object rotations, and greatly simplify the calibration and rendering process (Section 3.2) since the geometric relationship between each pixel in the tablet’s display and its corresponding reflection direction is independent of the orientation of the display with respect to the viewer. However, a curved reflector can produce images that are uncomfortable to fuse. So in Section 4, we evaluate the *comfortable viewing region* of all rotationally-symmetric conic reflectors and find ideal reflector shapes for different display sizes.

The optical performance of the reflector also depends on its material properties. The surface should be very *smooth* and *shiny* to produce a sharp and smooth reflection of the display. It should also have uniform reflectance across wavelengths to prevent color artifacts and undesirable intensity modulation. In our experience, commonly found polyethylene terephthalate (PETG) plastic works well.

The thickness of the reflector is another important factor in the system design. As shown in Figure 3, reflectors that are too

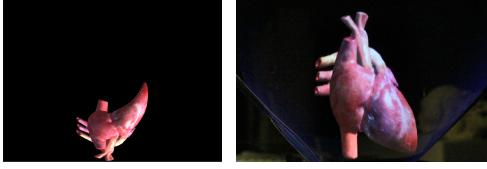


Figure 4. Left: A pre-warped image that is displayed on the tablet. Right: The distortion caused by the reflection from the cone produces a perspective-correct image to the viewer.

thick result in visible double reflections and those that are very thin have noticeable color artifacts due to wave interference. We also experimented with solid opaque reflectors, which would avoid these issues altogether, but found that they failed to achieve the desired effect of making the 3D object appear to be floating above the tablet, *inside* the reflector; transparent reflectors provided a much stronger illusion.

Distortion Calibration

Our goal is to produce a perspective-correct image to someone who is viewing our display from the side. Because the reflector is curved it will show a distorted reflection of the tablet's display. Therefore, some type of calibration is required to determine how to warp the image that is shown on the tablet to compensate for this distortion (Figure 4).

Problem statement

To help describe the calibration problem and our proposed solution, consider the coordinate system in Figure 5. We assume the tablet display is centered at the origin and that the viewer is located at $\hat{v} = (0, y, z)$, a distance y above the plane of the tablet's display and a distance z away from the center of the tablet. The user may rotate the display surface by an angle θ around the y -axis. Finally, imagine a virtual camera is located at the viewer's position as this provides a convenient way of parameterizing the cone of rays that are visible to the viewer. Specifically, the ray through a point $\hat{p} = (u, v)$ on the image plane of this virtual camera is redirected by the reflector and intersects the tablet display at the point $\hat{q} = (s, t)$ or, equivalently, the viewer would observe the tablet pixel at \hat{p} along this viewing ray.

Consider the optical transfer function $\mathbf{e} : \hat{p} \rightarrow \hat{q}$ that maps points on the virtual image plane of the virtual camera to points on the tablet display. The goal of our calibration is to compute the inverse of this function $\mathbf{f} = \mathbf{e}^{-1}$, which we refer

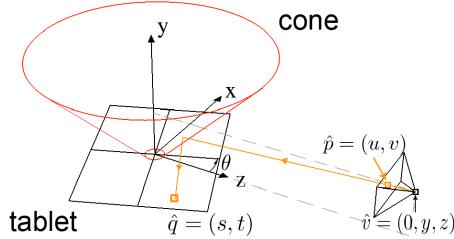


Figure 5. The arrangement of the tablet display and reflector along with the coordinate system used throughout the paper. The origin of this coordinate system is located in the center of the tablet display with its positive y -axis pointing up along the central axis of the reflector and its positive z -axis pointing toward the viewer.

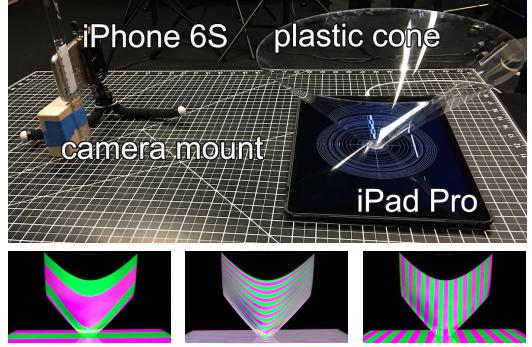


Figure 6. Our calibration setup consists of (top) a Pepper's Cone display and a calibration camera placed at the desired viewing location. We record a video while (bottom) a series of green and magenta striped patterns are displayed on the tablet that assign each pixel a unique binary code [38]. This allows establishing a sparse set of correspondences between pixels in the calibration camera and pixels in the tablet display.

to as the *distortion map*. Note that this distortion map allows “displaying” a target image $I(u, v)$ on the virtual image plane by displaying the warped image $I(\mathbf{f}(s, t))$ on the tablet display.

We make two important assumptions about the distortion map to make its estimation tractable. First, we assume that both \mathbf{e} and \mathbf{f} exist. This would not be the case for setups that cause multiple tablet locations to be superimposed along the same viewing ray (e.g., a thick reflector that shows double reflections as in Figure 3). We ignore these situations. Second, we assume the reflector geometry can be constructed by sweeping a 1D curve around the y -axis. This allows us to estimate the distortion map \mathbf{f} for only a single rotation angle ($\theta = 0^\circ$) and then trivially compute \mathbf{f} for any rotation angle by simply composing \mathbf{f} with a 2D rotation by θ around the y -axis: $I(\mathbf{f}(R_\theta(s, t)))$. If the reflector does not have this property, then \mathbf{f} would need to be estimated over a dense set of rotation angles, complicating the calibration process.

Estimation

We estimate the distortion map \mathbf{f} in three steps. First, we recover samples of \mathbf{e} over a discrete domain using a method similar to environment matting [38]. Second, we compute the inverse of these samples to produce a sparse reconstruction of \mathbf{f} . Finally, we perform a data interpolation and smoothing step to fill in any holes.

As shown in Figure 6, we place a calibration camera at \hat{v} and record a video as a series of alternating horizontal and vertical magenta and green striped images are shown on the tablet display. This assigns each tablet pixel a unique temporal sequence of colors. We start by calculating a binary mask that indicates which pixels in the calibration camera observe the tablet display by applying a simple intensity threshold to the observed temporal sequence. Similar to Zongker et al. [38], we apply an *open* and then *close* morphological operator using a 10×10 box filter to clean this mask. Let P denote the set of pixels in the calibration camera that survive this masking step. We then decode the sequence of green and magenta colors observed at each pixel in P to obtain discrete samples of \mathbf{e} .

As illustrated in Figure 7(a), we simply splat the sparse set of samples of \mathbf{e} computed above into our reconstructed distortion map, \mathbf{f} . Specifically, let $\mathcal{Q} = \{\mathbf{e}(\hat{p}) \mid \hat{p} \in P\}$ be the set of tablet pixels for which we observe the mapping function \mathbf{e} . We may trivially invert these samples to obtain a scattered reconstruction of $\mathbf{f}(\hat{q}) = \mathbf{e}^{-1}(\hat{q}) \forall \hat{q} \in \mathcal{Q}$.

The final step in our estimation process is to interpolate and smooth the scattered reconstruction of \mathbf{f} . To achieve this, we convolve the sparse samples using a wide 2D Gaussian kernel. Specifically, we reconstruct values of $\mathbf{f}(\hat{q})$ at a regular sampling resolution of $M \times N$ by convolving the set of scattered samples, \mathcal{Q} , with an isotropic 2D Gaussian kernel of size $k \times k$ with standard deviation σ :

$$\mathbf{f}(\hat{q}) = \frac{\sum_{\hat{q} \in \mathcal{Q}} w(\hat{p}, \hat{q}) \mathbf{f}(\hat{q})}{\sum_{\hat{q} \in \mathcal{Q}} w(\hat{p}, \hat{q})}, \quad w(\hat{p}, \hat{q}) = \mathcal{N}(\|\hat{p} - \hat{q}\|_2^2 \mid \sigma). \quad (1)$$

Figure 7(b) shows the resulting distortion map and Figure 7(c) shows the effect of warping an image according to this map.

Ray tracing

When an accurate analytical model of the reflector shape is available, another option is to replace the camera-based estimation process described above with ray tracing plus interpolation. This can be used as an alternative for computing correspondences between viewing rays and pixels on the primary display. We investigated this approach, but found the environment matting method above offers a more general strategy that can handle non-ideal reflector shapes.

Interactive Rendering

Rendering images for a Pepper’s Cone display is straightforward. We first render a standard perspective image that we would like to present to the viewer using a conventional real-time graphics stack, $I(\hat{p})$, and then warp this image in a final rendering pass using a fragment shader that applies the distortion map: $I(\mathbf{f}(R_\theta(\hat{q})))$, as discussed in Section 3.2.

In order to allow the viewer to interactively study the scene from any direction, we query the tablet’s gyroscope to determine its current heading θ and render a view of the 3D scene from this viewpoint. As with other virtual and augmented reality displays, maintaining the lowest possible latency between the moment when the user rotates the display and when the rendered image is updated, called *motion-to-photon* lag, is critically important to achieve the illusion of smoothly rotating a 3D object that is suspended inside the reflector. There are four main sources of latency in the system: 1) the time it takes before motion is detected by the gyroscope; 2) the time

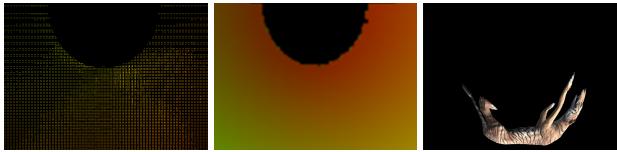


Figure 7. (a) Result of reconstructing the portion of the distortion map, f , that can be decoded directly from the calibration sequence shown in Figure 6. Note the many holes. (b) We convolve these scattered samples with a wide 2D Gaussian kernel to compute a dense hole-free distortion map. (c) An image warped according to this distortion map.

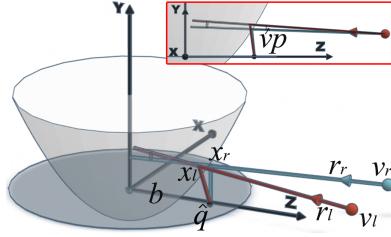


Figure 8. Left and right viewing rays of a single point on the primary display surface. Note that these rays are not guaranteed to intersect for curved reflectors (closeup looking along x-axis). We use the vertical parallax, which is equal to the vertical angle between these two rays, to characterize the performance of different reflector shapes.

required to predict the rotation angle from the gyroscope’s raw data; 3) the time required to render this new viewpoint; and 4) the time required to update the tablet’s display. We observed that the desired illusion would fail if the end-to-end lag is greater than 80ms. This value is consistent with related studies on minimum tolerable lag in virtual-reality and augmented-reality displays [7]. Section 6 discusses our specific prototype in more detail and includes an analysis of lag.

REFLECTOR SHAPE AND BINOCULAR FUSION

An important contribution of this paper is a systematic analysis of different reflector shapes on the basis of how comfortably a user can “fuse” the images seen by their left and right eyes. As we will show, our analysis indicates an optimal configuration is achieved with a cone. The size of the cone and the angle it makes with its axis are functions of the size of the primary display and viewer position.

Ideally, every displayed point should be observed along a single ray from the left and right eye, respectively, that intersect at a 3D position located inside the reflector. In general for curved reflectors, however, the left and right eyes observe slightly different curved virtual images, leading to some amount of stereo divergence.

Specifically, consider a single pixel \hat{q} on the primary display and assume the user’s left and right eyes are located at $v_l = (-IPD/2, y_0, z_0)$ and $v_r = (IPD/2, y_0, z_0)$, where IPD denotes their interpupillary distance (Figure 8). The image of \hat{q} is observed along a ray from each eye, $r_l(\hat{q})$ and $r_r(\hat{q})$, that correspond to the specular reflections of \hat{q} . Let us denote the points on the reflector surface where this specular reflection occurs as $x_l(\hat{q})$ and $x_r(\hat{q})$, respectively.

Note that in general $r_l(\hat{q})$ and $r_r(\hat{q})$ are not guaranteed to intersect. We consider the “vertical parallax,” which is equal to the vertical angle between these two rays:

$$vp(\hat{q}) = \arccos \left(\frac{r_l(\hat{q})_{yz} \cdot r_r(\hat{q})_{yz}}{\|r_l(\hat{q})_{yz}\| \cdot \|r_r(\hat{q})_{yz}\|} \right),$$

where r_l and r_r are the directions of the two rays from the eyes with a slight abuse of notation and $(\cdot)_{yz}$ is the component of the vector in the y-z plane.

Prior work has shown that the human visual system can tolerate vertical parallax up to 0.5 degrees before they experience difficulty fusing the left and right stereo images [26]. We

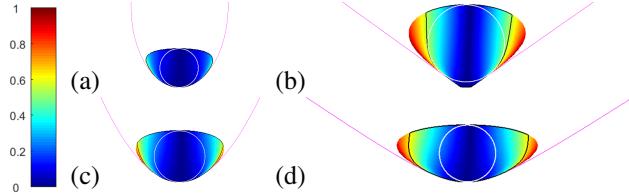


Figure 9. Example vertical parallax and comfortable viewing region of (a) elliptic ($e = 0.8, p = 7$), (c) parabolic ($e = 1, p = 7$) and (d) hyperbolic ($e = 1.7, p = 7$) (b) conical ($e = 1.7, p = 0$) reflectors (the rest of the parameters are set as the 27" monitor settings in Sec. 4.1). The black curve and white circle enclose the region where $vp \leq 0.5$ and its maximum inscribed circle, or the *comfortable viewing region*, respectively. The magenta curve shows the contour of the reflector.

denote as *stereo-compatible* viewing conditions those that achieve this level of vertical parallax. In the following sections, we consider an important class of reflector geometries and find optimal points in this space that achieve the largest stereo-compatible viewing area.

Reflector Shape Optimization

We consider reflector shapes from the family of rotationally-symmetric conic surfaces [32]. By this, we mean the profile of the reflector is a conic section. They can be parametrized by eccentricity, e , and focus, p :

$$S(x, y, z) = x^2 + z^2 - (e^2 - 1)y^2 - 2py + p^2 = 0. \quad (2)$$

This corresponds to a cone reflector when $p = 0$, and an elliptic ($e < 1$), parabolic ($e = 1$) and hyperbolic ($e > 1$) reflector when $p > 0$. We chose this particular class of surfaces because they cover a wide range of rotationally symmetric shapes that can be built rather easily from available materials. They are also widely used in imaging and astronomical applications [32].

To make the cone self-supporting, we truncate the base to a circle of radius $b = 1.25\text{cm}$. Mathematically, this corresponds to shifting the reflector down along the y -axis so that the base is at $y = 0$ with radius b by using reflector $S(x, y + \frac{-p + \sqrt{e^2(b^2+p^2)-b^2}}{e^2-1}, z) = 0$. Additionally, to simplify our analysis, we model the primary display as being circular, with a radius of R . We also assume the user is located roughly an arm's length away from the display, $z_0 = 50\text{cm}$, and use an eye height of $y_0 = R/2$, which centers their view with the reflector. We assume an IPD of 6.35cm.

For each value of e and p , we compute the vertical parallax over a dense rectangular grid of eye rays. Figure 9 shows example visualizations of the vertical parallax for one type of conical, elliptic, parabolic, and hyperbolic reflector. Note that the amount of vertical parallax generally increases towards the edge of the reflector where it has a more grazing orientation from the viewer's perspective. Also note that each reflector gives rise to its own *comfortable viewing region*, which we define as the field of view of the largest inscribed circle with a maximum vertical parallax of 0.5 degrees, denoted by $FOV(e, p)$.

We computed the size of this viewing region over a dense set of values of $e \in [0.1, 0.2, 0.3, \dots, 4]$ and $p \in [0, 0.5, 1, 1.5, \dots, 20]$.

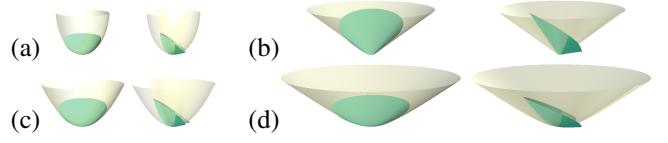


Figure 10. Visualization of the optical surface (green) for an (a) elliptic ($e = 0.8, p = 7$), (c) parabolic ($e = 1, p = 7$), (d) hyperbolic ($e = 1.7, p = 7$) and (b) conical ($e = 1.7, p = 0$) reflector. For each shape we show images from the viewer's perspective (left) and another viewpoint (right).

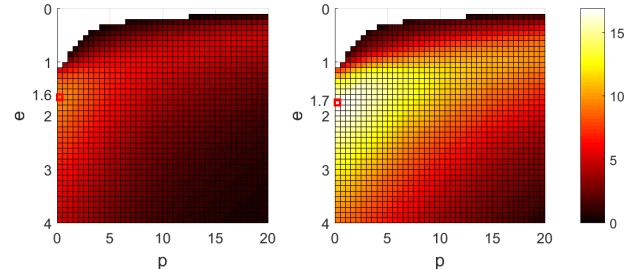


Figure 11. Visualization of $FOV(e, p)$ (in degree) for (left) a 12.9" iPad Pro and (right) a 27" 16:9 display. The reflectors that maximize the comfortable viewing region are indicated with small red boxes and detailed in the paper.

As shown in Figure 11, we observed that $FOV_{max}(p) = \max_e FOV(e, p)$ decreases monotonically as p increases and that for each p there is a value of e that maximizes $FOV(e, p)$.

Based on this analysis, we found that values of $e = 1.6, p = 0$, or a 51° cone, gives the largest comfortable viewing region with a FOV of 9.6° in the case of a 12.9" iPad Pro ($R=9.8\text{cm}$). For a 27" 16:9 display ($R=17\text{cm}$), we found that a reflector with $e = 1.7, p = 0$, or a 54° cone, would maximize the comfortable viewing region with a $FOV(e, p) = 16.8^\circ$ (Figure 11).

Optical Surface

Although our system is monocularly calibrated, the fact that the images seen by the two eyes are stereo-compatible implies that the viewer perceives the displayed contents as being located on a specific 3D surface. This surface, which we call the *optical surface* is defined by the intersection points of corresponding rays from the left and right eyes. Due to the presence of small vertical parallax, we compute the *closest* points of intersection. Figure 10 shows optical surfaces for four types of quadric surfaces. Note that the optical surface is generally *inside* the reflector surface (the part of the optical surface that sticks out of the reflector is almost invisible from the viewer's perspective); hence the object is seen as floating inside the reflector.

Note the optical surface is *independent* of the particular object being visualized. Hence, the binocular depth cues may differ from the monocular cues, which depend on the shape of the rendered object. In practice, we find that this mismatch is rather subtle, with the monocular cues serving to provide fine-scale geometrical details, whereas the binocular cues serve to place the object at an absolute position inside the reflector. More analysis into the effect of this mismatch is an interesting topic of future work.

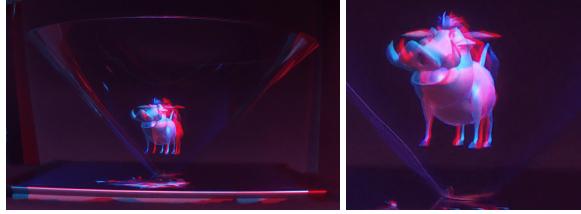


Figure 12. Stereoscopic image of our prototype display showing the Pumbaa model with binocular correction enabled. Please use red/cyan stereo glasses to view this figure and the high-resolution ones in the auxiliary material.

BINOCULAR CORRECTION

Although the methods above can produce a compelling illusion of a 3D object suspended inside a conical reflector, the perceived shape will not *exactly* correspond to the input 3D model because of the mismatch of monocular and binocular cues. However, our display can easily be extended to address this limitation by use of stereoscopic rendering and glasses.

We implemented a version of our display that presents correct binocular images to the viewer, shown in Figure 12. This involved performing the calibration process described in Section 3.2 separately for each left/right eye location, and maintaining two distortion maps. Inside the main render loop, an image is generated for each eye using these distortion maps and then multiplexed as a classic red/cyan anaglyph. The viewer observes correct binocular cues with matching red/cyan stereo glasses.

RESULTS

Our prototype display is shown in Figure 1(a). It consists of an iPad Pro tablet computer and a thin hollow truncated cone. The iPad has a 2732×2048 resolution LED-backlit LCD display that is 12.9 inches along its diagonal. The cone is constructed from a single PETG plastic sheet that is 0.5mm thick. The height of the cone is 141.4mm and it makes an angle of 45 degrees with its central axis (close to the optimal 51° but easier to construct).

The diameter of the opening that rests on the tablet is 16mm. We taped a US nickel to the bottom of the cone to help maintain its balance and increase friction. For the rotatable base we used a simple “lazy susan” designed for decorating cakes that we bought online. Excluding the tablet computer, the total cost of this setup is \$27 and required only 5 minutes to assemble from pre-cut pieces.

Our prototype display is calibrated with viewing location at $y=8\text{cm}$ and $z=43\text{cm}$ (Figure 5). The resolution of our iPhone 6S calibration camera is $M \times N = 1920 \times 1080$. We used values of $\sigma = 7$ and $k = 43$ for the scattered data interpolation step (Equation 1). Before interpolation, the set of sparse samples Q includes only 0.83% of the pixels on the tablet and we manage to reconstruct a smooth mapping \mathbf{f} that covers 68.13% of the pixels. We use a total of 24 calibration images for the green-and-magenta binary code sequence. The total distortion map construction and processing time required 42s on a Macbook Pro with a 2.7 GHz Intel Core i5 CPU and 8 GB of main memory.

We implemented our rendering application using the Unity game engine with a custom OpenGL ES fragment shader that performs final distortion rendering pass.

Lag

The iPad Pro display uses a variable refresh rate which can cycle down from its maximum of 60 Hz to 30 Hz when the image is mostly static. This corresponds to a lower bound of 33.3ms motion-to-photon lag in the worst case and 8.3ms on average once the display begins to move. We use the Extended Kalman Filter (EKF) [10] that is implemented in the Google Cardboard SDK [11] to predict θ from the raw gyroscope data. In particular, it predicts the rotation angle ahead of time according to previous predictions and the most recent gyroscope heading. We measured the overall motion-to-photon latency of our display using a 240fps video camera and found it to be roughly 157ms when the display first starts to move. This number falls to roughly 17ms after the display is moving. This significant decrease in latency is due to the increased refresh rate and the benefit of more accurate predictions. In all, we found our prototype display was able to achieve the illusion of smooth continuous motion without noticeable lag or drift. Please see the accompany video for results.

Applications

Shown in Figure 13, we evaluated our prototype using a variety of static and dynamic 3D models. In each row a reference image is shown next to an image of our display at the same viewpoint for comparison. The remaining three images in each row show our display at other viewpoints.

The top four rows are static objects. The car and the cartoon house are synthetic 3D models. Note the specular highlights on the surface of the shiny car. These types of view-dependent effects are difficult to reproduce with typical volumetric displays (Section 2). The woolly mammoth skeleton and the kylix (ancient Greek cup) in the third and fourth rows are 3D scans of archaeological artifacts [34]. The tiger, beating heart and turtle are all animated models. Some of these include audio which we found added to the realism of the experience. We also experimented with showing captured 3D content on our display. The rings in the 8th row were captured by photographing a real object as it was rotated on a turntable. This result suggests a simple end-to-end capture and display pipeline for physical objects enabled by this work.

Our prototype display is able to reproduce the overall appearance of the input model quite well without any obvious geometric distortions. There are some noticeable color differences that result from both the color shift of the capture camera and light interference. We believe this could be reduced by picking better material and thickness or performing a specific radiometric calibration. The high resolution of our display allows viewing small details such as the individual blood vessels in the heart model. These results also demonstrate possible applications of our display in product exhibition, decoration, archaeology and cultural heritage preservation, medicine, education, virtual assistants, games and animation.

While the best illusion is obtained with custom shaped reflectors, we’ve also achieved compelling results with a range of



Figure 13. Some static and dynamic models that we used to evaluate our display and illustrate different application domains. Left to right: Reference image; image of our Pepper's Cone display at same viewpoint for comparison; our display at three other viewpoints. Top to bottom: car, cartoon house, woolly mammoth skeleton, kylix, animated tiger, beating heart with sound, talking cartoon avatar with sound, rings. Please see the accompanying video for dynamic versions of these results.

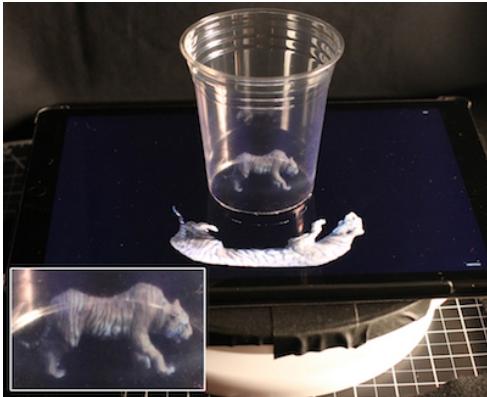


Figure 14. A Pepper’s Cone display built from an ordinary plastic cup with its base removed.

everyday objects like cups, vases, and other household items. Here we show a Pepper’s Cone built from an ordinary plastic cup with its base removed.

PRELIMINARY USER FEEDBACK

We collected feedback for both our glasses-free system (monocular system) and the system with binocular correction (binocular system) from 12 users: 7 male, 5 female, 23-40 years old. Four out of twelve were doing research in AR/VR.

The primary goal of this study was to evaluate the quality of the illusion, namely that 1) the object appears to be floating in the center of the cone, and 2) that the object appears three-dimensional. Users were first asked to rate how believable the illusion of a real object floating inside the cone is for both monocular and binocular systems, on a scale of 1: poor / looks fake, 2: unsure, 3: fair, 4: good, 5: excellent / looks like a real physical object. Both monocular and binocular systems had a median rating of 4, and means of 4.17 (monocular) and 4.0 (binocular). In other words, the monocular system does a surprisingly good job of producing the illusion of depth, matching or even slightly outperforming the binocular system.

All participants reported that the displayed object appears to be inside the cone. And even after comparing with the binocular system, eight out of twelve reported that the rotation axis of the displayed object is well centered and several commented that they could not distinguish between monocular and binocular. Only four participants reported noticing a slight misalignment in the monocular system—recall that the optical surface leans towards the front surface of the cone.

Users were also asked to rate how compelling the illusion of the displayed object being three-dimensional is for both systems on a scale of 1: poor, 2: unsure, 3: fair, 4: good, 5: excellent. Both the monocular and binocular systems had a median rating of 4, and means of 3.92 (monocular) and 4.17 (binocular). Interestingly, the binocular system, which produces correct stereo images was only marginally better, demonstrating the effectiveness of the monocular system. We attribute this to the strong visibility cues achieved with the semi-transparent reflector.

Multiple participants remarked that our system achieves a high-level of realism with such a simple setup and without glasses. For example, “It’s a very easy way to get the illusion of having a real object in front of you - definitely more convincing than 3D glasses with a display. The rotation adds a lot.” All but one user preferred the monocular system due to its comfort (no glasses), convenience, and the fact that it avoids the color artifacts caused by anaglyph stereo glasses, while achieving a comparable 3D effect. Note that using polarizing or shutter glasses could mitigate the color artifacts.

When asked if they would prefer having either of these systems in their home, nine answered yes to the monocular system. Three answered maybe, but wanted to see more evidence of practical uses and applications.

In terms of limitations, several participants reported disliking the color artifacts of the binocular system caused by the anaglyph glasses. A couple of participants remarked that the connecting seam of the cone is distracting and one user suggested adding some type of interaction. Finally, five participants complained that the system works properly at only a single calibrated head position and encouraged us to lift that requirement.

CONCLUSION AND FUTURE WORK

We introduced *Pepper’s Cone*, a simple inexpensive do-it-yourself display that provides a fun and compelling way to view a 3D scene. Inspired by the classic Pepper’s Ghost illusion, our setup consists of a thin hollow plastic cone placed directly on top of a tablet computer display. We pre-warp images on the tablet based on its gyroscope heading so that when viewed from the side, the rendered 3D scene appears to be suspended inside the cone. The illusion is quite convincing, as illustrated by the results presented in the paper and accompanying video. Compared to existing 3d displays, our setup is far simpler and less expensive and does not require the use of special glasses.

Our design is based on a thorough analysis of reflector geometries, to determine the optimal shape for a given 2D display. We also present the first analysis of stereo-compatibility for Pepper’s Ghost displays and find reflector shapes that enable comfortable stereo viewing with an apparent depth behind the reflector, providing the illusion that the object is floating inside the reflector. When stereo glasses are available, we present a binocular correction technique that brings the monocular and binocular cues into agreement. However, our preliminary user feedback shows that the monocular system can achieve similar compelling 3D illusion while being more comfortable and convenient.

Our display is not suitable for bright outdoor environments. As shown in Figure 15, we studied the performance of our setup in different lighting conditions. As with other augmented-reality displays, environments with less ambient lighting and uniform dark areas directly behind the display provide the best viewing conditions. However, we found that the iPad Pro was bright enough to work in typical office environments quite well.

Another limitation of our display is that it assumes a specific viewing position. However, small displacements from

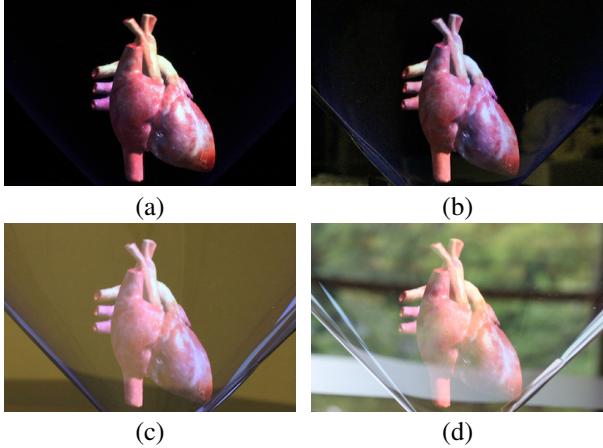


Figure 15. We studied the performance of our prototype Pepper’s Cone display in different lighting environments. As expected, this type of display performs best in (a,b) dark environments, although it is still usable in (c) moderately bright spaces like inside a building or in an office environment. The performance degrades significantly in (d) very bright environments with direct sunlight.

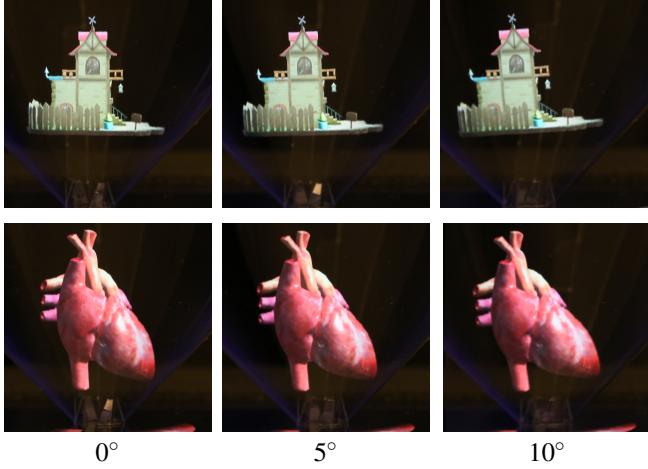


Figure 16. Demonstration of the distortion in our display at non-ideal viewing locations. All images were captured at the calibration position of $y = 8\text{cm}$ and $z = 50\text{cm}$, but with 0, 5 and 10 degree angular displacements along the x-axis. The object appears to tilt when viewed off-axis.

this calibrated position do not rapidly degrade the experience. Specifically, the displayed image appears to tilt when the viewing position is offset along the x-axis by more than 5 degrees (Figure 16). Displacements along the y- and z-axes are far less noticeable. The supplemental video demonstrates these distortions as well. We found that users are able to find an acceptable viewing position with no guidance. One avenue of future work is to add camera-based head-tracking to automatically account for the actual viewing position. This would allow the user to freely walk around our display, which would be better in certain applications like a virtual greeter at the entrance of a building. Finally, we would like to extend our system to accommodate more than one viewer at a time. Our rotational symmetric reflector geometry is particularly well suited for this purpose.

ACKNOWLEDGEMENTS

This work was supported by funding from the National Science Foundation grant IIS-1538618, Intel, Google, and the UW Animation Research Labs. We also wish to thank Dan Goldman for early input and all the participants of our user study for their time and participation.

REFERENCES

1. Arena3D. 2015. Arena3D Industrial Illusion. <http://arena3d.com/>. (2015).
2. Tibor Balogh, Péter Tamás Kovács, and Zoltán Megyesi. 2007. Holovizio 3D display system. In *Proceedings of the First International Conference on Immersive Telecommunications*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 19.
3. Peter C. Barnum, Srinivasa G. Narasimhan, and Takeo Kanade. 2010. A Multi-layered Display with Water Drops. *ACM Transactions on Graphics* 29, 4 (2010), 76:1–76:7.
4. Oliver Bimber, Bernd Fröhlich, Dieter Schmalstieg, and L Miguel Encarnaçao. 2006. The virtual showcase. In *ACM SIGGRAPH 2006 Courses*. ACM, 9.
5. P-A Blanche, A Bablumian, R Voorakaranam, C Christenson, W Lin, T Gu, D Flores, P Wang, W-Y Hsieh, M Kathaperumal, and others. 2010. Holographic three-dimensional telepresence using large-area photorefractive polymer. *Nature* 468, 7320 (2010), 80–83.
6. Jeremy Brooker. 2007. The Polytechnic Ghost. *Early Popular Visual Culture* 5, 2 (2007), 189–206.
7. Grigore Burdea and Philippe Coiffet. 2003. Virtual reality technology. *Presence: Teleoperators and virtual environments* 12, 6 (2003), 663–664.
8. Dimensional Studios on going. Dimensional Studios. <http://www.eyeliner3d.com/>. (on going).
9. Neil A. Dodgson. 2005. Autostereoscopic 3D Displays. *Computer* 38, 8 (2005), 31–36.
10. Garry A Einicke and Langford B White. 1999. Robust extended Kalman filtering. *IEEE Transactions on Signal Processing* 47, 9 (1999), 2596–2599.
11. Google on going. Google Cardboard SDK for Unity. <https://developers.google.com/cardboard/unity/>. (on going).
12. Hologram USA on going. Hologram USA. <http://www.hologramusa.com/>. (on going).
13. Holus. 2015. Holus. <http://hplustech.com/>. (2015).
14. Michael Huebschman, Bala Munjaluri, and Harold Garner. 2003. Dynamic holographic 3-D image projection. *Optics Express* 11, 5 (2003), 437–445.
15. Haruo Isono, Minoru Yasuda, and Hideaki Sasazawa. 1993. Autostereoscopic 3-D display using LCD-generated parallax barrier. *Electronics and Communications in Japan (Part II: Electronics)* 76, 7 (1993), 77–84.

16. Andrew Jones, Ian McDowall, Hideshi Yamada, Mark Bolas, and Paul Debevec. 2007. An Interactive 360° Light Field Display. In *ACM SIGGRAPH 2007 Emerging Technologies*.
17. Andrew Jones, Jonas Unger, Koki Nagano, Jay Busch, Xueming Yu, Hsuan-Yueh Peng, Oleg Alexander, and Paul Debevec. 2014. Creating a Life-sized Automultiscopic Morgan Spurlock for CNNs "Inside Man". In *ACM SIGGRAPH 2014 Talks*.
18. Yoshihiro Kajiki, Hiroshi Yoshikawa, and Toshio Honda. 1996. 3D display with focused light array. *Proceedings of SPIE* 2652 (1996), 106–116.
19. Douglas Lanman, Gordon Wetzstein, Matthew Hirsch, Wolfgang Heidrich, and Ramesh Raskar. 2011. Polarization Fields: Dynamic Light Field Display Using Multi-layer LCDs. *ACM Transactions on Graphics* 30, 6 (2011), 186:1–186:10.
20. Andrew Lippman. 1980. Movie-maps: An application of the optical videodisc to computer graphics. In *AcM Siggraph Computer Graphics*, Vol. 14. ACM, 32–42.
21. Keiichi Maeno, Naoki Fukaya, Osamu Nishikawa, Koki Sato, and Toshio Honda. 1996. Electro-holographic display using 15mega pixels LCD. *Proceedings of SPIE* 2652 (1996), 15–23.
22. Wojciech Matusik and Hanspeter Pfister. 2004. 3D TV: A Scalable System for Real-time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes. *ACM Transactions on Graphics* 23, 3 (2004), 814–824.
23. Musion Das Hologram Ltd. since 1996. Musion. <http://musion.com/>. (since 1996).
24. Susumu Nakajima, Kozo Nakamura, Ken Masamune, Ichiro Sakuma, and Takeyoshi Dohi. 2001. Three-dimensional medical imaging display with computer-generated integral photography. *Computerized Medical Imaging and Graphics* 25, 3 (2001), 235–241.
25. S. K. Nayar and V. N. Anand. 2007. 3D Display Using Passive Optical Scatterers. *Computer* 40, 7 (2007), 54–63.
26. Damin Qin, Mamoru Takamatsu, and Yoshio Nakashima. 2004. Measurement for the Panum's fusional area in retinal fovea using a three-dimension display device. *Journal of Light & Visual Environment* 28, 3 (2004), 126–131.
27. Eric G Rawson. 1969. Vibrating varifocal mirrors for 3-D imaging. *IEEE Spectrum* 6, 9 (1969), 37–43.
28. L. Sawalha, M. P. Tull, M. B. Gately, J. J. Sluss, M. Yearly, and R. D. Barnes. 2012. A Large 3D Swept-Volume Video Display. *Journal of Display Technology* 8, 5 (2012), 256–268.
29. Maurice Stanley, Patrick B Conway, Stuart D Coomber, John C Jones, Dave C Scattergood, Christopher W Slinger, Robert W Bannister, Carl V Brown, William A Crossland, and Adrian RL Travis. 2000. Novel electro-optic modulator system for the production of dynamic images from giga-pixel computer-generated holograms. In *Electronic Imaging*. International Society for Optics and Photonics, 13–22.
30. Chihiro Suga and Itiro Siio. 2011. Anamorphicons: An extended display with a cylindrical mirror. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 242–243.
31. Ivan E. Sutherland. 1968. A Head-mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I))*. ACM, New York, NY, USA, 757–764. DOI: <http://dx.doi.org/10.1145/1476589.1476686>
32. Rahul Swaminathan, Michael D Grossberg, and Shree K Nayar. 2006. Non-single viewpoint catadioptric cameras: Geometry and analysis. *International Journal of Computer Vision* 66, 3 (2006), 211–229.
33. Savaş Tay, P-A Blanche, R Voorakaranam, AV Tunc, W Lin, S Rokutanda, T Gu, D Flores, P Wang, G Li, and others. 2008. An updatable holographic three-dimensional display. *Nature* 451, 7179 (2008), 694–698.
34. The Smithsonian Institute 2014. Smithsonian X 3D. <http://3d.si.edu/>. (2014).
35. G. Wetzstein, D. Lanman, W. Heidrich, and R. Raskar. 2011. Layered 3D: Tomographic Image Synthesis for Attenuation-based Light Field and High Dynamic Range Displays. *ACM Trans. Graph.* 30, 4 (2011).
36. Gordon Wetzstein, Douglas Lanman, Matthew Hirsch, and Ramesh Raskar. 2012. Tensor Displays: Compressive Light Field Synthesis Using Multilayer Displays with Directional Backlighting. *ACM Transactions on Graphics* 31, 4 (2012), 80:1–80:11.
37. Xinxing Xia, Xu Liu, Haifeng Li, Zhenrong Zheng, Han Wang, Yifan Peng, and Weidong Shen. 2013. A 360-degree floating 3D display based on light field regeneration. *Optics Express* 21, 9 (2013), 11237–11247.
38. Douglas E Zongker, Dawn M Werner, Brian Curless, and David H Salesin. 1999. Environment matting and compositing. In *Proceedings of ACM SIGGRAPH*. 205–214.