

8D: Interacting with a Relightable Glasses-Free 3D Display

ACM Student Research Competition Entry

Matthew Hirsch¹

¹MIT Media Lab

75 Amherst St.

Cambridge, MA, USA

{mhirsch,holtzman,raskar}@media.mit.edu

Shahram Izadi²

Henry Holtzman¹

²Microsoft Research UK

7 JJ Thompson Avenue

Cambridge, UK

shahrami@microsoft.com

Ramesh Raskar¹

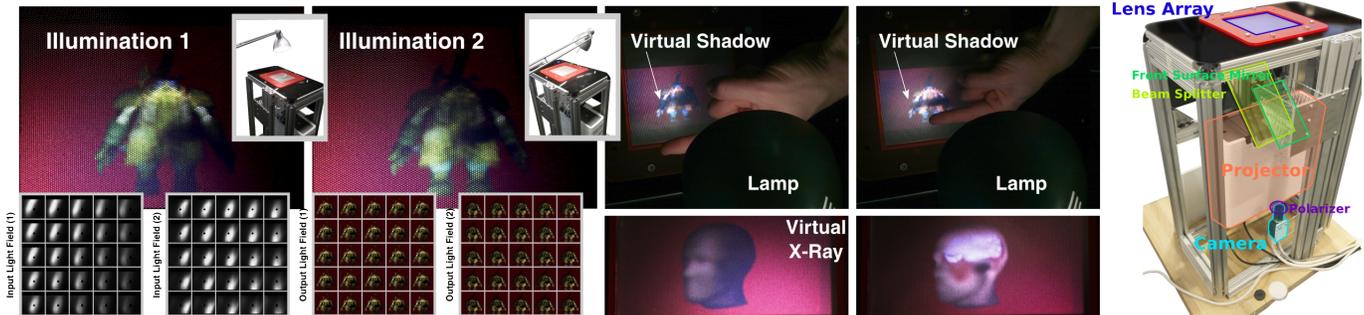


Figure 1. The 8D prototype allows for glasses-free display of 3D content, whilst simultaneously capturing any incident light for re-illumination and interaction. From left to right: A user shines a light source (lamp) at the 3D display, and the rendered 3D model is correctly re-illuminated. Virtual shadows can be cast by placing a finger between the display and light source. This allows any light source to act as an input controller, for example to allow intuitive interaction with a medical data set. 8D works by simultaneously capturing and displaying a 4D light field (as shown bottom left inset).

ABSTRACT

We present an 8-dimensional (8D) display that allows glasses-free viewing of 3D imagery, whilst capturing and reacting to incident environmental and user controlled light sources. We demonstrate two interactive possibilities enabled by our lens-array-based hardware prototype, and real-time GPU-accelerated software pipeline. Additionally, we describe a path to deploying such displays in the future, using current Sensor-in-Pixel (SIP) LCD panels, which physically collocate sensing and display elements.

Author Keywords

3D display; Light Fields; Light-based interaction, relightable

ACM Classification Keywords

H.5.2 Information Interfaces And Presentation: User Interfaces: Input Devices and Strategies

General Terms

Algorithms, Experimentation, Design

INTRODUCTION

Imagine a display that behaves like a window. Glancing through it, viewers perceive a virtual 3D scene with correct parallax, without the need to wear glasses or track the user. Light that passes through the display correctly illuminates both virtual objects on the display and physical objects in the environment. While researchers have considered such displays, or prototyped subsets of these capabilities, we contribute an interactive 8-dimensional (8D) display which simultaneously captures and displays a 4D light field.

We describe the design of our lens-array based, projector-camera 8D display prototype, and GPU-based pipeline for real-time rendering and capture of 4D light fields. Our prototype provides horizontal and vertical parallax as a user moves within the viewing region, without the need for user instrumentation or tracking. Additionally, our display simultaneously captures the incident 4D light field from environmental or user controlled light sources. We demonstrate the use of such a display in interactive scenarios: allowing for realistic relighting of virtual 3D objects, as well as the unique ability to use *any* light source as a user input controller, for example to visualize medical imaging data. With the advent of sensor-in-pixel (SIP) LCD displays, we propose a clear path to implementing thin, portable, 8D displays in the future.

CONTRIBUTIONS

- A real-time, relightable, glasses-free 3D display with horizontal and vertical parallax (8D display)
- Demonstrate two user interaction scenarios made possible by 8D display
 - Relightable objects
 - Non-physical lighting (X-Ray lights)
- Propose architecture for future cheap, portable, light field interaction devices

BACKGROUND

Light fields have long been a valuable tool in rendering [15], and more recently in developing next-generation cameras [18, 21] and displays [22]. They have also seen limited use for HCI, where depth information and gesture can be extracted from a light field captured through a prototype SIP screen and combined mask [8].

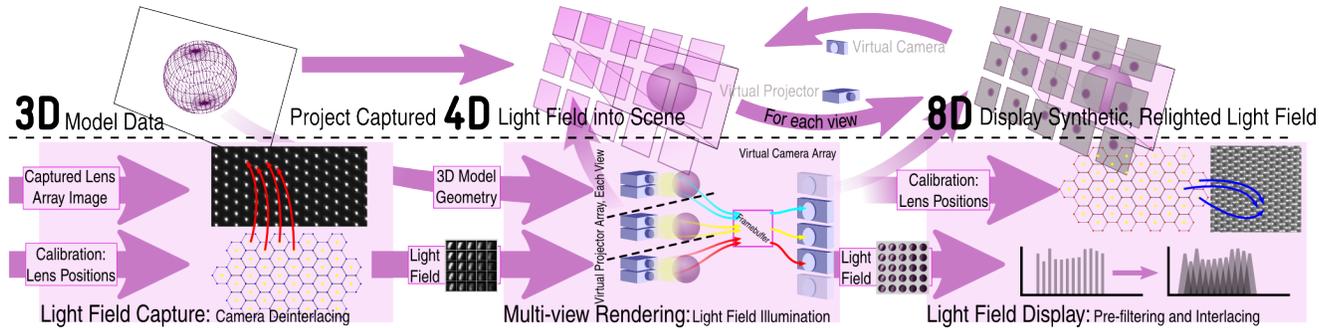


Figure 2. The GPU pipeline and associated data. The pipeline is conceptually divided into three stages: **deinterlacing, light field rendering, and interlacing.** A 4D input light field is combined with a 4D output light field, synthesized from 3D model geometry. From left to right, images captured through the lens array are deinterlaced using lens center positions determined by calibration. Each view of the captured light field is projected onto 3D model data. Output light field views are generated by a virtual camera array, pre-filtered, and interlaced for display on the lens array.

Glasses-free 3D parallax barrier [11] and lens-array [16] displays have existed for over 100 years. Nayar et al. [17] create a lighting sensitive display, though it cannot accurately map shadows and specularities. BRDF displays can simulate flat surfaces with a particular Bi-Directional Reflectance Distribution Function [9]. 6D displays that demonstrate 4D relighting of 2D images have been shown in both active [8] and passive [7] modes. A recently shown 7D display [20] tracks a single light point as input. In a closely related work, Cossairt et al. [4] implement a 7fps 8D display, but focus on rendering illumination effects for a 2D camera, rather than 3D perception for a live viewer. Our work contributes a hardware approach to real-time 8D display that is compatible with emerging display technologies and a new GPU rendering and capture pipeline to make simultaneous, interactive 4D lighting and 4D capture feasible.

Our display offers interesting possibilities for interacting through light. Light pens and widgets have been previously used for interaction [2]. In recent years, lighting widgets have been integrated into tabletop computing systems [14], and novel optics and computer vision have been used for interaction with screens, tables, and physical surfaces over a screen [12, 19, 6]. We are proposing the first interaction system capable of fully capturing and displaying arbitrary light transport within a volume above the display.

IMPLEMENTATION

Hardware

We propose to implement an 8D display by placing an array of microlenses on a SIP LCD screen. Due to the pixel pitch limits of existing SIP hardware, we implement an equivalent projector-camera system (Figure 1). We place a $150\text{mm} \times 150\text{mm}$ hexagonal lens array (Fresnel Tech. #360, 0.5mm pitch) atop a Grafix acetate diffuser onto which we both image and project. This optically simulates the orthographic light field produced by the SIP display.

Our prototype uses a grayscale Point Grey Gazelle 2048 \times 2048, 120fps camera with a 50mm Schneider Xenoplan lens as a sensor. A Sanyo PLV-Z800 1920 \times 1080 projector is modified by shifting lens forward 4mm, allowing it to create a focused 325dpi image, matching the width of the hexagonal lens sheet. The projector and camera share an optical

path through a 40/60 beamsplitter. We prevent cross-talk by multiplexing both through crossed linear polarizers.

Calibration

Our prototype system necessitates a calibration step to align the sample grids of the camera and projector with the real-world coordinates of the lens sheet (Figure 3). To calibrate the camera, a collimated light source is placed above the lens sheet, creating a point below each lens center, which can be located in the camera’s view. A grid of 3^{rd} order polynomial lines are fit to the grid of detected lens centers to reduce the contribution of local intensity variation caused by non-uniformity in the diffuser sheet.

The projector is calibrated using the morié magnifier [10] effect. A hexagonal grid of red bars on a black background is projected at the expected lens center locations. The scale and rotation of the projected image are adjusted until the central view above the lens sheet is solid red.

Real-time GPU Pipeline

The 8-Dimensional nature of our relightable 3D display is apparent within our GPU pipeline implementation. Our main 8D rendering pipeline depicted in Figure 2 and described in this section, generates both the input and output light fields using off-screen rendering to texture arrays. Implemented in HLSL and DirectX 11, the GPU pipeline runs in real-time on an Nvidia GTX 470 GPU. Our draw loop consists of $N \times M$ renderings of our 3D scene, one pass for each of the $N \times M$ light field views shown on our display (Figure 2, Center). Each view is observed using an off-axis oblique camera projection [13], corresponding to the view angle through the lens sheet of our 8D display prototype, and then rendered into a slice of a 2D texture array. We implement a simplified version of the rendering equation, neglecting BRDFs.

$$L_o(\mathbf{x}, \omega) = \int_{\Omega} L_i(\mathbf{x}, \omega')(-\omega' \cdot \mathbf{n})d\omega' \quad (1)$$

where L_i is the measured incident light field, L_o the displayed light field, and ω' the incoming lighting direction. Though in our model local regions are invariant in outgoing light direction, ω , each light field view is generated with a view matrix corresponding to a virtual skewed orthographic camera viewing the scene from ω .

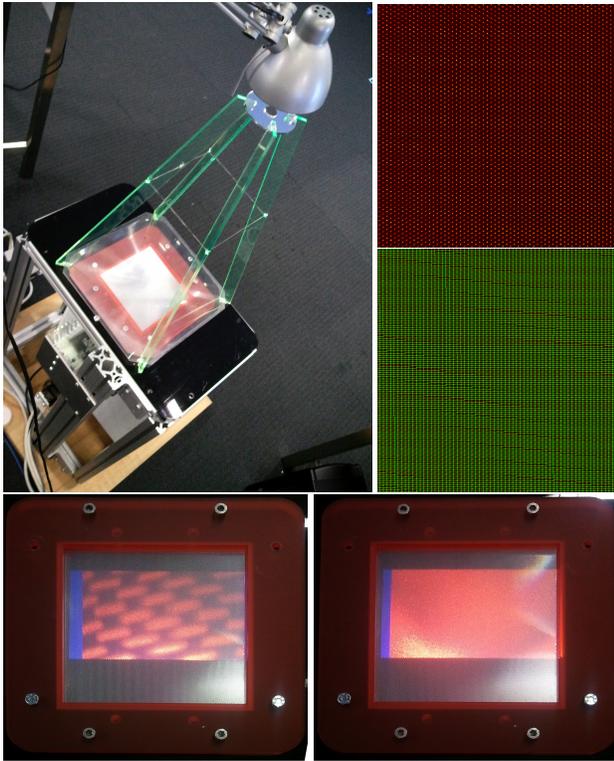


Figure 3. (Left) A Fresnel lens is placed on top of the 8D display prototype to facilitate calibration of the input side. Placing a point light source at the annular mirror disk ensures a columnated beam is emitted from the Fresnel lens. (Top, Right) The camera view of the calibration image. (Middle, Right) A grid of polylines fit to the lens-center grid. The projector is calibrated by creating a repeating pattern that will be aliased by the lens sheet when improperly aligned (Bottom Left) and a solid color when aligned (Bottom Right).

To capture a 4D light field we deinterlace images recorded from the back of the lens array in our GPU pipeline (Figure 2, Left). For each render pass, we project $P \times Q$ captured input views onto the scene using projective texture mapping. In practice, we use 5×5 views for both input and output, as a limited number of texture look-up coordinates can be passed between the shader stages. After the 8D rendering is completed, two additional render passes with associated shader stages implement two 4D spatio-angular filtering operations, and hexagonal interlacing/deinterlacing (Figure 2, Right).

PERCEPTION AND INTERACTION

The importance of lighting in perception has long been recognized in photography and computer graphics. It has been studied in detail with respect to the human visual system and plays a central role in our understanding of the world. Our goal in creating an 8D Display is to take a step towards displays that can produce the convincing illusion of physical reality [1, 5]. A key aspect of this goal will be the ability of these displays to realistically react to incident environmental lighting. Beyond the reproduction of physical reality, it is possible to render non-physical scenes that fulfill interface or interaction goals. Figure 1 shows the 8D display prototype rendering a virtual 3D model. When viewing the display, the model appears to be 3D, with full parallax. In this example,

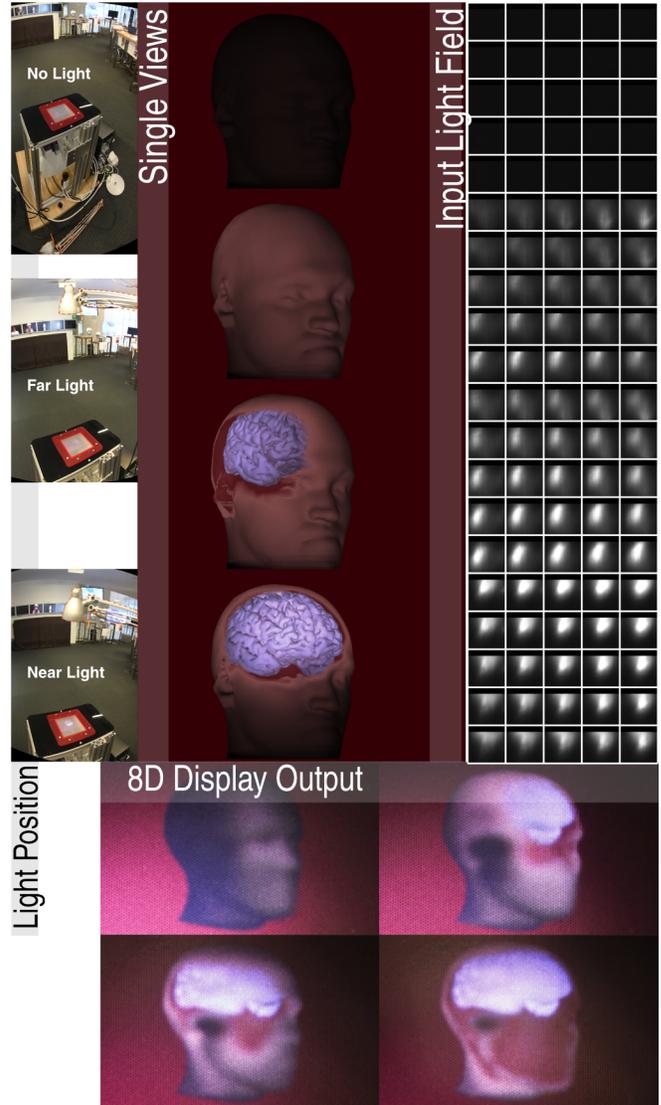


Figure 4. A hand-held light (Left) reveals either a patient's skin and facial features, or brain tissue (Center). This non-physical interpretation of lighting intensity leads to a novel interface for viewing medical data sets. The captured incident light field (Right) is projected onto a head and brain model, modulating the skin transparency by incident light intensity. (Center) One view of the resulting output light field. (Bottom) The x-ray demo under various lighting conditions.

the user also moves a lamp over the display, and the 3D model responds to the incident light and is correctly re-illuminated.

In a second demonstration, depicted in Figure 4, more intense light acts like a virtual x-ray, revealing inner structure in MRI data. In this case, the skin of the patient is visible under lower light conditions, allowing a clinician to visually identify the patient. As the x-ray lamp is brought closer to the screen, the skin layer becomes transparent, revealing the segmented brain imagery.

LIMITATIONS

Our 8D display prototype is subject to optical and computational limitations on its performance. Our prototype supports

7×7 views optically. However, due to limitations of our GPU pipeline, we are able to support only 5×5 views in real time. Though our output images are resampled onto a hexagonal grid in our GPU pipeline in order to accommodate the hexagonal lens array, the approximate equivalent rectilinear resolution of our display is 274×154 per view. With a $3mm$ focal length, the lens array offers a 19° field-of-view.

Sampling theory for automultiscopic displays, which predicts the inherent spatio-angular resolution trade-off shared by all such designs, is characterized by Zwicker et al. [23]. Their frequency domain analysis explains the depth of field exhibited by automultiscopic displays. Following Zwicker, for a display with angular sampling rate Δv , and spatial sampling rate Δt , objects at distance $|z| > \frac{\Delta t}{\Delta v}$ will be blurred. Defining the t plane at the focal point of the lens sheet, and the v plane at unit distance, $\Delta t = 0.547mm$ and $\Delta v = 0.026mm$. Given the above equation for z , objects up to $21mm$ from the display are reproduced without blur. Empirical depth of field characterization shows satisfactory reproduction for objects extending up to $3cm$. This can be improved with increased angular resolution.

Direct reflection and scattering from the lens sheet competes with the light emitted from the screen of 8D display. Anti-reflection coatings can reduce such reflections to less than 1% of incident light, and are common in commercial optics. (E.g., Anti-Reflective Coatings for Acrylic by American Polarizers, Inc.)

These factors limit the spatial resolution of our prototype display currently. For example, in Figure 4 (Center, vs Bottom), details of the face and brain are obscured in our lower spatial resolution display prototype. Even at these low resolutions, our prototype illustrates the potential of such light-based interactive displays.

FUTURE WORK

The examples demonstrated in this work only begin to touch on the possibilities enabled when light transport is modeled by a display in this way. Abundant interactive possibilities include: using an off-the-shelf light source as 6DoF input controller, direct manipulation of physical lights to cause re-lighting of 3D scenes, augmented-reality mirrors, accurately mimicking surfaces with exotic BRDFs, and applying non-realistic physics to real lighting sources. Beyond our demonstrated x-ray example, one could imagine implementing non-euclidean optics, with multiple 8D displays. 8D Displays of sufficient intensity can be used to computationally illuminate objects in the environment, for interaction, user guidance, or advanced shape scanning and acquisition. Modulating rendering properties based on incident light color would also be a powerful interaction metaphor, but our prototype is currently capable only of grayscale input.

Much of the potential impact of this research is predicated on the existence of Sensor-In-Pixel (SIP) LCDs. In recent years LCD manufacturers have introduced semiconductor technologies that combine light sensitive elements into the LCD driver matrix [3]. In combination with collocated, thin, optical capture and display elements, such as those provided

by a SIP LCD, this work suggests a straightforward route to achieving a low-cost, commercially realizable, real-time, 8D display. The 40in diagonal Samsung SUR40 Sensor-In-Pixel display has 1920×1080 resolution, yielding a pixel pitch of $55dpi$. One goal of this work is to inspire manufacturers to increase these numbers.

CONCLUSION

As human capabilities are further enhanced by computation, we are becoming limited by the I/O bandwidth between our brains and external computation. In this paper we have presented a glasses-free 3D display capable of reacting to real-world environmental and user-controlled lighting in real-time. This work paves the way to creating displays that can produce physically convincing illusions that participate optically in the environment in which they are rendered. Our thesis is that such displays will unlock the full power of the human visual system, which itself developed in a vibrant world, with three spatial dimensions, that affords eight-dimensional light transport. Increasing the bandwidth between the human brain and computation will be a key challenge for computer science in the coming decades, with implications far outside of computer science in nearly every human endeavor. The new display and input capabilities afforded by the 8D Display draw on many disciplines within computer science, including graphics and human-computer interfaces – various aspects of the work were presented at ACM SIGGRAPH 2012 and will be shown at ACM SIGCHI 2013 – and diverse fields outside computer science: optics and electronics. We envision a future where interaction with computational systems is perfectly impedance-matched to the capabilities of the human brain and senses. 8D Displays will be an integral part of future interactive systems that hope to fully engage the human visual system, and as such will illuminate areas of future research in computer science.

VIDEO FIGURE

A video demonstrating our prototype 8D Display can be found at <http://web.media.mit.edu/~mhirsch/8D>.

REFERENCES

1. Adelson, E. On seeing stuff: The perception of materials by humans and machines. In *Proc. of the SPIE*, vol. 4299, Citeseer (2001), 1–12.
2. Anderson, R., Knowles, W., and Culley, J. Methods and apparatus involving light pen interaction with a real time display, Apr. 12 1977. US Patent 4,017,680.
3. Brown, C., Kato, H., Maeda, K., and Hadwen, B. A continuous-grain silicon-system lcd with optical input function. *Solid-State Circuits, IEEE Journal of* 42, 12 (dec. 2007), 2904–2912.
4. Cossairt, O., Nayar, S., and Ramamoorthi, R. Light field transfer: global illumination between real and synthetic objects. In *ACM Trans. on Graph.*, vol. 27, ACM (2008), 57.
5. Fleming, R., Dror, R., and Adelson, E. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision* 3, 5 (2003).

6. Follmer, S., Johnson, M., Adelson, E., and Ishii, H. deform: An interactive malleable surface for capturing 2.5d arbitrary objects, tools and touch. In *Proc. 24th ACM symp. on UIST*, ACM (2011), 527–536.
7. Fuchs, M., Raskar, R., Seidel, H.-P., and Lensch, H. P. A. Towards passive 6d reflectance field displays. *ACM Trans. Graph.* 27, 3 (Aug. 2008), 58:1–58:8.
8. Hirsch, M., Lanman, D., Holtzman, H., and Raskar, R. Bidi screen: A thin, depth-sensing lcd for 3d interaction using light fields. In *ACM Trans. Graph.*, vol. 28, ACM (2009), 159.
9. Hullin, M. B., Lensch, H. P. A., Raskar, R., Seidel, H.-P., and Ihrke, I. Dynamic display of brdfs. *Computer Graphics Forum* 30, 2 (2011), 475–483.
10. Hutley, M., Hunt, R., Stevens, R., and Savander, P. The moiré magnifier. *Pure and Applied Optics: Journal of the European Optical Society Part A* 3, 2 (1999), 133.
11. Ives, F. E. Parallax stereogram and process of making same. U.S. Patent 725,567, 1903.
12. Izadi, S., Hodges, S., Taylor, S., Rosenfeld, D., Villar, N., Butler, A., and Westhues, J. Going beyond the display: a surface technology with an electronically switchable diffuser. In *Proc. 21st ACM symp. on UIST*, ACM (2008), 269–278.
13. Lanman, D., and Hirsch, M. Build your own glasses-free 3d display. In *ACM SIGGRAPH 2011 Courses*, ACM (2011), 4.
14. Leitner, J., Powell, J., Brandl, P., Seifried, T., Haller, M., Dorray, B., and To, P. Flux: A tilting multi-touch and pen based surface. In *Proc. 27th intl. conf. extended abstracts on Human factors in computing systems*, ACM (2009), 3211–3216.
15. Levoy, M., and Hanrahan, P. Light field rendering. In *Proc. 23rd conf. on Computer Graphics and Interactive Techniques*, ACM (1996), 31–42.
16. Lippmann, G. Épreuves réversibles donnant la sensation du relief. *Journal of Physics* 7, 4 (1908), 821–825.
17. Nayar, S., Belhumeur, P., and Boult, T. Lighting sensitive display. *ACM Trans. on Graph.* 23, 4 (2004), 963–979.
18. Ng, R., Levoy, M., Brédif, M., Duval, G., Horowitz, M., and Hanrahan, P. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report 2* (2005).
19. Sato, T., Mamiya, H., Koike, H., and Fukuchi, K. Photoelastictouch: Transparent rubbery tangible interface using an lcd and photoelasticity. In *Proc. 22nd ACM symp. on UIST*, ACM (2009), 43–50.
20. Tompkin, J., Muff, S., Jakushevskij, S., McCann, J., Kautz, J., Alexa, M., and Matusik, W. Interactive light field painting. In *ACM SIGGRAPH 2012 E. Tech.*, ACM (2012), 12.
21. Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., and Tumblin, J. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph.* 26, 3 (2007), 69.
22. Wetzstein, G., Lanman, D., Hirsch, M., and Raskar, R. Tensor displays: Compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph.* 31, 4 (2012), 80.
23. Zwicker, M., Matusik, W., Durand, F., and Pfister, H. Antialiasing for automultiscopic 3D displays. In *EGSR* (2006).