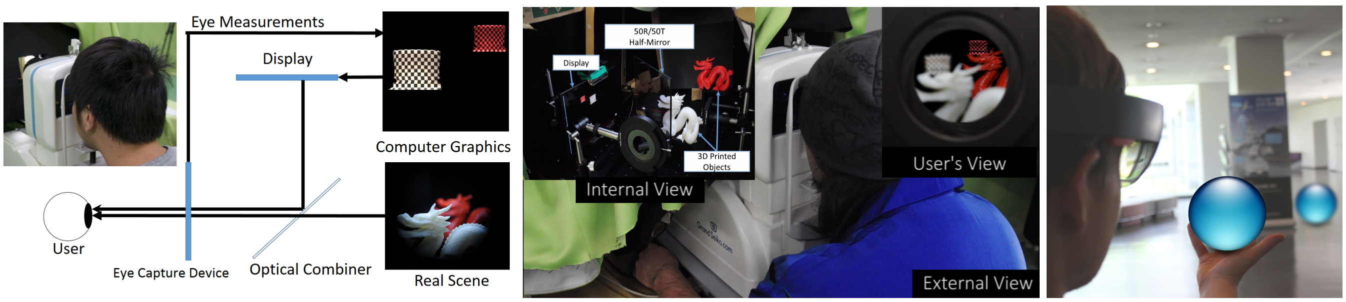
EyeAR: Refocusable Augmented Reality Content through Eye Measurements

Damien Constantine Rompapas * Aitor Rovira * Sei Ikeda† Alexander Plopski * Takafumi Taketomi *
Christian Sandor * Hirokazu Kato *

*Interactive Media Design Laboratory
Nara Institute of Science and Technology

†Mobile Computing Laboratory
Ritsumeikan University



(a) (b) (c)

Figure 1: The goal of our demo is to embed computer graphics (CG) objects into the real world, which are indistinguishable from real objects.(a) Abstract concept of the system presented in [11], (b) Tabletop implementation of EyeAR, (c) A user wearing a HWD with EyeAR integrated.

Abstract

The human visual system always focuses at a distinct depth. Therefore, objects that lie at different depths appear blurred, a phenomenon known as Depth of Field (DoF); as the user's focus depth changes, different objects come in and out of focus. Augmented Reality (AR) is a technology that superimposes computer graphics (CG) images onto a user's view of the real world. A commonly used AR display device is an Optical See-Through Head-Mounted Display (OST-HMD), enabling users to observe the real-world directly, with CG added to it. A common problem in such systems is the mismatch between the DoF properties of the user's eyes and the virtual camera used to generate CG.

In this demonstration, we present an improved version of the system presented in [11] as two implementations: The first as a high quality tabletop system, the second as a component which has been integrated into the Microsoft Hololens [18].

Keywords: Raytracing, Physically Based AR, Augmented Reality, Depth of Field, Optical Defocus

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; I.3.7 [Three-Dimensional Graphics and Realism]: Computer Graphics—Raytracing; I.3.3 [Computing Methodologies]: Picture/Image Generation—Display Methods;

1 Background

In 1965, Ivan Sutherland envisioned the ultimate display [13], a multi-modal display that provides visual and haptic feedback that is indistinguishable from real objects. This vision is similar to the Star Trek Holodeck [17] and has inspired research in the fields of photorealistic Augmented Reality (AR).

Commercial interest in AR has increased significantly over the last few years, paving the way for Optical See-Through Head-Mounted Displays (OST-HMDs) such as the Google Glass [16] and the Microsoft Hololens [18]. One of the largest obstacles to photorealistic AR on OST-HMDs is the mismatch between the Depth of Field (DoF) properties of the user's eyes and the virtual camera used to generate CG. The Human Visual System relies on several depth cues in order to distinguish which objects are closer to us than others. Teittinen [14] and Ware [15] discuss all such depth cues in detail; one of them, accommodation, is closely related to DoF and refers to the eye changing its shape to change its focal length, thus bringing objects at different distances into focus. The goal of our demo is to generate CG objects that behave like real objects during accomodative actions of the user's eyes.

2 Related Work

The distribution of light in space is exhaustively modeled with the plenoptic function [1]. It is at least six-dimensional: three dimensions for the observer's viewpoint, two for the viewing direction, and one for the wave length; its dense sampling and reconstruction is expensive. One assumption that makes either task easier is that the environment is observed—or rendered—from the outside, thus keeping spectral radiance the same along straight lines and effectively removing one dimension from the viewpoint, reducing the representation to a Light Field [6]. One fundamental way of embedding the Light Field of a virtual object into reality is by injecting a Light Field into the user's direct view. Second, one can capture the Light Field of the user's environment, fuse it with the Light Field that we want to embed, and display this combination to the user.

One such display [5] uses an array of micro lenses to re-create a Light Field. However, with this method, there is a trade-off between the achieved spatial resolution and the supported depth range. Multi-plane displays [2] render and display an image across multiple screens which lie on multiple depths to recreate the Light Field. Currently the most accurate method [10] for displaying the con-

*e-mail: {damien.rompapas.dk1, aitor,plopski, takafumi-t, sandor, kato} @is.naist.jp
†e-mail:ikeda.sei.jp@ieee.org

tent on such a display is not real time, whilst real time methods [3] experience temporal artifacts, typically a DoF mismatch. Additionally, MacKenzie and colleagues have shown [8] that a minimum of five focal planes is required to produce an acceptable range of accomodative cues. A stereoscopic multi-focal display was presented by Love and colleagues [7] that uses a set of high-speed lenses to present focused content at four planes at real time response rates. Other approaches for LFDs include voxel displays via flexible membranes [12] to create variable focus image planes; although these displays are capable of covering a larger depth range than our approach, they are only able to cover a limited number of focus planes. Additionally, these LFDs are only capable of producing VR or Video See-Through (VST) AR content and are not suitable for OST-AR applications. Finally, Maimone and colleagues [9] used an array of defocused point lights to create an OST LFD; this display grants a wide field of view and a large accommodation range, but has limitations in terms of reconstruction fidelity. Another refocusable VST-AR approach was presented by Kán and Kaufmann [4]; they obtain the camera's lens parameters and render CG using these camera parameters. The resulting CG's DoF matches the DoF properties of the video in each frame.

What makes this demo unique and special? Our work crucially differs from related research in several ways. Compared to Kan and Kaufmann [4], we created an OST system and that measures the user's eyes instead of taking camera parameters. Compared to LFD approaches, we do not display the full Light Field; instead, we only display the relevant subset. Additionally we created two implementations of this system, one of which can be used as a component of a commodity HMD.

3 Demo Description

We present the participant with two Implementations of our system: The first as a table top setup, the second as a component integrated into a commodity HMD.

In the first implementation, we sit a participant in front of an Autorefractometer and cover one of their eyes with an eye patch. As the participants look inside the box enclosure the Autorefractometer measures the eye. These measurements are then used in combination with a Real-time Raytracer to create a mixed-reality scene with several objects (some of which are virtual) that the participant observes. The rendering quality is improved compared to [11] as we have solved several of the issues in the algorithm.

In our second variant we show participants EyeAR as a component that is integrated into the Microsoft Hololens, allowing them to experience a re-focusable AR scene on a commodity OST-HMD device.

4 Demo Requirements

- *The amount of floor or desktop space needed*:
  - Floor space: 5m x 5m
  - Desktop space: one table(1m x 2m) and two chairs
- *The list of equipment you will bring(as detailed as possible)*:
  - Microsoft Hololens
  - Grand Seiko WAM-5500 Auto Refractometer
  - Desktop computer, Cables, Monitor
  - Netgear R7500 Wireless router
  - Flat-screen display
  - Box-Enclosure
  - 3x 3D Printed Models
- *Any power, socket and outlet needs*:
  - 3x power outlet
  - Total power consumption 1- 1.5 kw
- *Environment requirements*:
  - Please provide one table (1m x 2m) and one chair.
  - We require constant illumination because pupil size depends on light environment.

References

[1] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology, 1991.

[2] K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks. A stereo display prototype with multiple focal distances. In *ACM Transactions on Graphics*, volume 23, pages 804–813, 2004.

[3] F.-C. Huang, D. Luebke, and G. Wetzstein. The light field stereoscope. In *ACM SIGGRAPH Emerging Technologies*, pages 24:1–24:1, 2015.

[4] P. Kán and H. Kaufmann. Physically-based depth of field in augmented reality. In *Eurographics (Short Papers)*, pages 89–92, 2012.

[5] D. Lanman and D. Luebke. Near-eye light field displays. In *ACM SIGGRAPH Emerging Technologies*, volume 32, pages 220:1–220:10, 2013.

[6] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 31–42, 1996.

[7] G. D. Love, D. M. Hoffman, P. J. Hands, J. Gao, A. K. Kirby, and M. S. Banks. High-speed switchable lens enables the development of a volumetric stereoscopic display. *Optics express*, 17(18):15716–15725, 2009.

[8] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, 10(8):22–22, 2010.

[9] A. Maimone, D. Lanman, K. Rathinavel, K. Keller, D. Luebke, and H. Fuchs. Pinlight displays: Wide field of view augmented reality eyeglasses using defocused point light sources. In *ACM SIGGRAPH Emerging Technologies*, page 20, 2014.

[10] R. Narain, R. A. Albert, A. Bulbul, G. J. Ward, M. S. Banks, and J. F. O'Brien. Optimal presentation of imagery with focus cues on multi-plane displays. *ACM Transaction on Graphics*, 34(4):59:1–59:12, July 2015.

[11] D. C. Rompapas, K. Oshima, K. Moser, E. Swan, S. Ikeda, G. Yamamoto, T. Taketomi, C. Sandor, and H. Kato. [DEMO] EyeAR: Physically-Based Depth of Field through Eye Measurements. In *Proceedings of the 14th IEEE International Symposium on Mixed and Augmented Reality*, Fukuoka, Japan, October 2015.

[12] B. T. Schowengerdt and E. J. Seibel. True 3-d scanned voxel displays using single or multiple light sources. *Journal of the Society for Information Display*, 14(2):135–143, 2006.

[13] I. E. Sutherland. The ultimate display. In *Proceedings of the IFIP Congress*, pages 506–508, 1965.

[14] M. Teittinen. Depth cues in the human visual system. *The Encyclopedia of Virtual Environments*, 1, 1993.

[15] C. Ware. *Information visualization: perception for design*. Elsevier, 2012.

[16] Wikipedia. Google Glass — Wikipedia, the free encyclopedia, 2016. https://en.wikipedia.org/wiki/Google_Glass.

[17] Wikipedia. Holodeck — wikipedia, the free encyclopedia, 2016. http://en.wikipedia.org/wiki/Holodeck.

[18] Wikipedia. Windows holographic — wikipedia, the free encyclopedia, 2016. https://en.wikipedia.org/wiki/Windows_Holographic.