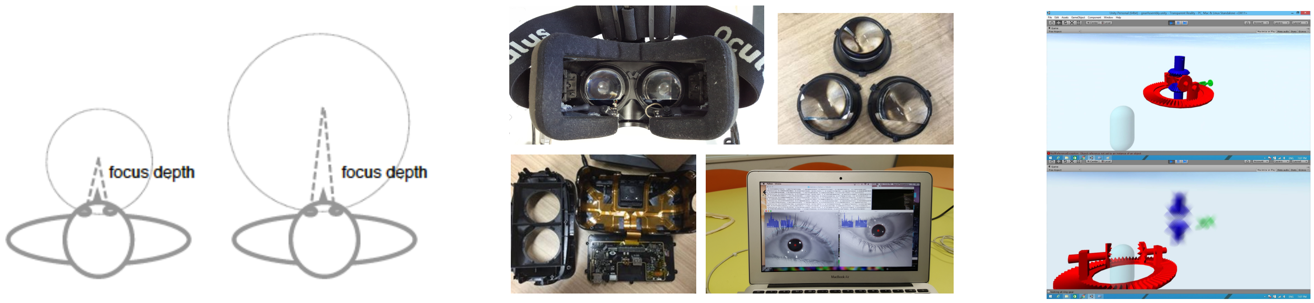
GazeSim: Simulating Foveated Rendering Using Depth in Eye Gaze for VR

Yun Suen Pai*[1], Benjamin Tag[1], Benjamin Outram[1], Noriyasu Vontin[2], Kazunori Sugiura[1], Kai Kunze[1]
[1]Graduate School of Media Design, KEIO University, Yokohama, Japan
[2]Fujitsu Design Ltd., Kawasaki, Japan



(a) (b) (c)

Figure 1: *(a) Focus depth of a user, left is close, right is far. (b) Custom hardware, integrating Pupil Labs tracker into Oculus Rift. (c) Example Application: Gear Assembly, foveated rendering depending on depth of focus of the eye.*

Abstract

We present a novel technique of implementing customized hardware that uses eye gaze focus depth as an input modality for virtual reality applications. By utilizing eye tracking technology, our system can detect the point in depth the viewer focusses on, and therefore promises more natural responses of the eye to stimuli, which will help overcoming VR sickness and nausea. The obtained information for the depth focus of the eye allows the utilization of foveated rendering to keep the computing workload low and create a more natural image that is clear in the focused field, but blurred outside that field.

Keywords: depth of field, eye gaze, foveated rendering, virtual reality

Concepts: **• Computing methodologies ~ Tracking; • Computing methodologies ~ Rendering; • Computing methodologies ~ Sensor devices and platforms**

1 Introduction and Motivation

Since their first appearance, developers of virtual reality (VR) solutions have been constantly working on improving the image quality and therefore the experienced sensation of reality and immersion. Depth of field (DOF) is one of the main factors that contribute to the sense of presence and immersion in an environment. Despite the recently rapidly evolving technology and promising new applications, the used components still consist of plain displays that render images stereoscopically to create sensations of depths. In order to enable the viewer's eyes to actually focus in different depths, rest at focal cues, and therefore behave and react more naturally, different solutions have been developed.

In 2015 researchers at Stanford University have proposed a head mounted display (HMD) solution that contains two layered, transparent LCD panels which. The human eye can use the displays that are placed in differing distances from the viewer to focus at different depths. Nevertheless, users will have to wait for these next generation HMDs to be able to enjoy this greater sensation of reality. [Huang 2015]

The second solution is offered by companies such as SMI and FOVE that use eye tracking within the HMD to monitor where the user is actually looking. The obtained data can help to create a sensation of depth by blurring everything the user is not exactly looking at, called foveated rendering. Besides the creation of a depth sensation, foveated rendering also distinguishes itself through a much smaller demand for computational work. Still, this technology does not allow a natural depth focus of the eyes.

To tackle this problem, we are proposing a solution that uses eye tracking technology to utilize eye-gaze information in order to detect the point in depth the user is focusing at (cf. 1a). The advantage of our system is that a regular one plain panel display HMD and an eye tracker installed in the mount are sufficient to allow the eyes to move, react, and focus more naturally.

2 Our Approach

We modified an Oculus Rift DK2 HMD for the initial prototype of the system. It consists of two infrared (IR) eye trackers by Pupil Labs that are placed in the cavity of the HMD just below the lenses to capture the user's eyes. However, there was initially insufficient space as cameras would be in contact with the user's face which proves to be discomforting due to the heat from the IR cameras. Therefore, the HMD's lenses were modified to provide

*e-mail:yspai1412@gmail.com.

SIGGRAPH '16 Posters, July 24-28, 2016, Anaheim, CA,
ACM 978-1-4503-4371-8/16/07.
http://dx.doi.org/10.1145/2945078.2945153

more space for the cameras without interfering with the user's comfort. The lenses were cut about one third of its total height at the bottom to achieve this. The results are shown in Figure 1b. The placement of the camera has been tested and despite it being not adjustable, it was able to accommodate all of the users for the case study. Figure 1b also shows the eyes being tracked in the VR headset.

Since the pupil trackers do not support VR environments by default, a custom plugin was written with Python for the tracking software that enables the raw data from the trackers to directly stream into Unity via open sound control (OSC). Unity then reads these values directly into the VR environment.

We are using the two normals perpendicular to the iris for both eyes (provided by the pupil software) to perform the depth calculation, detecting the intersection point or the vector that represents the shortest distance between the two normals. The focus depth is accurate for distances between 5 - 25 meters in VR space.

For a first user study we used a calibration system based on the K nearest neighbor (KNN) algorithm that teaches the system to recognize two layers of depth in the VR world. By selecting K =3, KNN calculates the Euclidean distances of the current eye gaze with the trained values to determine the two closest values of K that contains the information of the layer currently being observed. This allows a robust user dependent recognition.

3 Implementation and Future Work

The proposed system offers a new layer of interaction in VR, it has the potential to provide a hands-free experience in VR that preserves the immersion.

Transparent Heads-Up-Displays

For example, this method of selection would be useful for heads-up-display (HUD) based interaction where the user just focuses close to see the HUD, and focuses far to see the main content. This is particularly useful in spectating sports. The user simply needs to focus near for the score board to render clearer, and focusing far allows him or her to continue spectating the sports while the scoreboard fades away. More interesting applications can overlay heat-maps of ball possessions or other information the users view depending on the depth focus.

Item Selection in Assembling Applications

Focus depth can also be used for selecting certain items that are in a 3D gear assembly application (cf. figure 1c). While both hands can be used for drafting, crafting and editing of the item or part, the depth focus in combination with foveated rendering can be utilized for easier selection of items and clearer focus on particular details.

Window to other Worlds

Another application in VR is that it can provide a user with a window to the physical world by mounting a camera on the HMD. By focusing close, the user may switch back to the physical world and blurs the virtual background, while focusing far causes the physical world layer to fade away. This is useful for VR application in collaborative meetings for design spaces, where the user can switch between the VR design and the physical world where a discussion is being held, simply through depth of focus.

Acknowledgements

We would like to thank all our supporters, advisers, that helped with the creation of this first prototype. This research is partially supported by Fujitsu Design Limited Japan and NTT Mirai Net.

References

HUANG, F., CHEN K., AND WETZSTEIN G. 2015. The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-Eye Light Field Displays with Focus Cues. ACM *Transactions on Graphics 33, 5*.

LOUISE, R., MACKENZIEA, K. J., AND WATTA S. J. 2012. Multiple-focal-planes 3D displays: A practical solution to the vergence-accommodation conflict?., *IEEE International Conference on 3D Imaging (IC3D)*.

SHIRO, S., ET AL. 2013. Edge-based depth-fused 3D display. *Digital Holography and Three-Dimensional Imaging*. *Optical Society of America, 2013*.