Implementations toward Interactive Glasses-free Tabletop 3D Display

Shunsuke Yoshida*
National Institute of Information and Communications Technology (NICT)

1 Introduction

We have proposed a glasses-free tabletop 3D display that employs a hollow conical screen (anisotropic diffuser) and tiny circularly arranged projectors installed underneath the table [Yoshida et al. 2010]. The scheme resembles a horizontal-parallax-only 3D display based on light field reproduction, but it provides horizontal parallax in a circular direction. Both eyes of each viewer can see the individual aspects of 3D objects from any angle within 360° around the table (the latest prototype called *fvision* covers 150° of the viewing area by using 120 projectors). Although our display has allowed 3D images to be placed beside real objects as on an ordinary table, the next challenge is to establish an interactive tabletop environment for diverse collaborative activities. In this paper, we study sensing methods designed for our natural tabletop interaction environment and show several experimental implementations.

2 Our Approach

Our system is inspired by the goal of providing a natural glasses-free 3D display that allows users to perform ordinary tabletop tasks. Obstructive devices are removed from the tabletop surface, and 3D images are displayed on an empty and flat tabletop surface as if they were there. For this purpose, untethered and marker-less sensors are clearly desirable. Using our image-display principle, optical sensing methods can be employed with the table while hiding the system's presence, unlike other tabletop 3D displays. For example, the systems using a fly-eye lens, such as integral photography, or mechanically rotating disk optics must place their components on or over the tabletop area, significantly obscuring the view over the tabletop area from behind the screen. Other choices would be to install non-optical sensors under the table or optical sensors on or above the tabletop.

In contrast, our tabletop 3D display effectively deploys optical sensing methods that use a camera in addition to non-optical sensors. Our display's tabletop surface is simply covered by a semi-transparent plate (Fig. 1). Here, multiple viewers actually observe the side surface of a hollow area. Since the screen's upper base is open, the camera placed at the bottom of the hollow and facing the overhead area can see gestures around the displayed 3D objects. In practice, an optical sensor should use an invisible wavelength to avoid interference with the light of the 3D images. In our experimental implementation, the IR motion sensor placed inside the hollow could recognize several gestures and manipulate the posture of the displayed 3D object.

Additional optical sensors installed around the table can capture the viewers' attributes. For example, RGB-D sensors can detect their faces and locations, and these are used for adaptive vertical-perspective correction (Fig. 2). Our 3D display ensures horizontal motion parallax without eye tracking and gently leads the viewers' viewpoints around the desired viewing area by assuming a seated condition. Consequently, real-time perspective correction is not necessarily required for showcase-like usage, but adaptive adjustment of eye levels for the multiple viewers is preferable. Our 3D display is based on the light-field reproduction method, and it must generate multi-perspective images according to the projectors used. For computing viewpoint location per pixel, the following techniques are implemented. Any number of faces is detected by the RGB-D sensor, and the i^{th} viewer's viewpoint is computed as the midpoint of both eyes. It is converted into azimuth θ_i (the angle from the frontal direction where the viewer is located) and 2D parameter $\boldsymbol{D}_i$ (horizontal and vertical distances from original ring-shaped viewing area $\boldsymbol{R}$). Ideally, our 3D display does not limit the number of viewers because the appropriate horizontal parallax is always provided in any direction, but each viewer's body must dominate some angular region θ_α physically. Therefore, the ring-shaped viewing area around θ_i should be computed as $\boldsymbol{R}(\theta) = \boldsymbol{R} + \boldsymbol{D}_i$, where $|\theta - \theta_i| < \theta_\alpha/2$.

Figure 1: *Projector array (dots seen from opened hatch) and hollow conical screen covered by semi-transparent plate.*

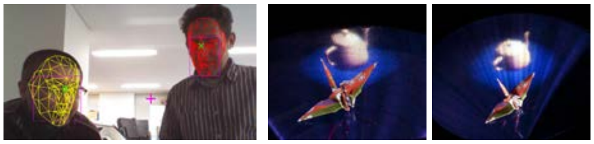

Figure 2: *Adaptive vertical-perspective correction. Results of face detection and 3D images for two different viewing heights.*

To avoid conflicts when several viewers are positioned in the same region, the $\boldsymbol{D}_i$ of the viewer nearest to the table is given priority. For the other regions, the original $\boldsymbol{R}$ is employed.

The system uses a master PC and five node PCs to compute over one hundred multi-perspective images in real time. The sensors are peripherals of the master PC, and sensing data are distributed to the node sequentially. Then, each node generates 24 multi-perspective images in real time [Yoshida 2013]. In pixel shading, every traveling direction of the pixels seeks a corresponding viewing area from $\boldsymbol{R}(\theta)$. To omit redundant multiple computations of $\boldsymbol{R}(\theta)$, it should be computed once a frame and quantized like a 1D texture in advance. — Part of this research was supported by JST CREST.

References

YOSHIDA, S. ET AL. 2010. Prototyping of Glasses-free Table-style 3D Display for Tabletop Tasks, *SID 2010*, 211–214.

YOSHIDA, S. 2013. Real-time Rendering of Multi-perspective Images for a Glasses-free Tabletop 3D Display. *3DTV-CON 2013*, 1–4.

---

*e-mail: shun@nict.go.jp

SIGGRAPH 2014, August 10 – 14, 2014, Vancouver, British Columbia, Canada.

ACM 978-1-4503-2958-3/14/08