An Immersive Surface for 3D Interactions

Yusuke Takeuchi
Interaction Technology Lab.
University of Tokyo
7-3-1 Hongo, Bunkyo-ku,
Tokyo, Japan
yusuke@itl.t.u-tokyo.ac.jp

Masanori Sugimoto
Interaction Technology Lab.
University of Tokyo
7-3-1 Hongo, Bunkyo-ku,
Tokyo, Japan
sugi@itl.t.u-tokyo.ac.jp

ITS'12, November 11–14, 2012, Cambridge, Massachusetts, USA.
ACM 978-1-4503-1209-7/12/11.

Abstract
This paper proposes a new tabletop interface that enables a user to visualize projected objects as if they existed on the tabletop surface. It uses head tracking, without the need for any specialized head-mounted hardware, displays, or markers. Nowadays, many interactive tabletop interfaces support interactions above the surface because this is more intuitive. In these 3D interactions, users should be able to gauge the size and height of the projected virtual objects. We evaluate our system quantitatively via a 3D interaction task, by comparing it with a standard tabletop system.

Author Keywords
Kinect; 3D Interaction; 3D Visualization; Immersive Tabletop; Head Tracking

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User Interfaces - Interaction styles.

Introduction
Nowadays, interaction methods in tabletop interfaces have evolved from multitouch interactions to interaction above the surface [2, 1] because of enabled higher degrees-of-freedom (DOF) input. However, although users can interact with a projection above the surface, they

Figure 1: System configuration

cannot interact with the 3D models directly because the 3D scene is just projected 2D screen.

In our system, the user can see the 3D models as if they existed on the surface, without the need for any specialized hardware or displays, by means of user head tracking. With this technique, the user can interact with 3D virtual models above the surface.We describe the system's implementation, application scenarios, and future expectations in this paper. Finally, we have evaluated our system empirically, via a pilot study that compares it with other tabletop systems.

Related Works

Many 3D interaction methods are realized by using depth sensors to detect the user's hand and any real objects. [2] and [4] support interaction above the surface by detecting the user's finger tips using depth camera. However, the user cannot know the height of a virtual object because of its 2D surface projection via a fixed program camera. Our research is focused on creating a full 3D experience.

Virtual reality systems provide users with realistic 3D experiences. Many early systems relied on head-mounted video, mobile displays, specialized displays or using markers (cf [3]). We have focused on avoiding any head-mounted equipment or specialized display. Depth-touch [1] is an interface that supports 3D interaction and changing perspective views by movements of the user's head position. However, changing perspective is achieved just by moving the camera position in program. As discussed below, the tabletop interface should take into account the surface inclination.

System Overview

3D Visualization Flow

In this research, the user can see projected objects as if they existed on the surface (referred to as 3D visualization in this paper). 3D visualization is enabled by head tracking, and we use Microsoft's Kinect for this purpose. The proposed tabletop interface has a straightforward configuration(Fig. 1.) We need only a Kinect device and a projection screen of known size. We do not need any specialized displays or head-mounted displays to achieve 3D visualization. Fig. 2 gives an overview of our system.

First, each set of coordinates (Kinect, program, and real world) has to be unified into one coordinate system. We also have to prepare 3D models for our program. We used the XNA Game Studio 4.0 platform for our 3D scene visualization and the JiglibX as physics engine. We generated a plane of the same size as the projection screen and applied the physical engine to it in XNA.

Second, the user's head position is tracked by the Kinect and the XNA camera position is synchronized with the user's head position. Besides, the Kinect distinguishes the user from other people in field of view of kinect. In this way, we can obtain the specific user's-view image (compare Fig. 2 (A) and Fig. 2 (B)). If the camera in XNA image (Fig. 2 B) is projected on the tabletop surface, the user's view is as shown in Fig. 2 (A). The user cannot see the projected object as if it existed on the surface. If (Fig. 2 C) is projected on the surface, the user's view can then be as shown in Fig. 2 (D). We calculate the relation between each pixel in the projector plane and the XNA camera plane, extract color data from the XNA camera's image, and set the color data at the pixel positions in every frame. This produces the correct image (Fig. 2 C) that should be projected on the surface.

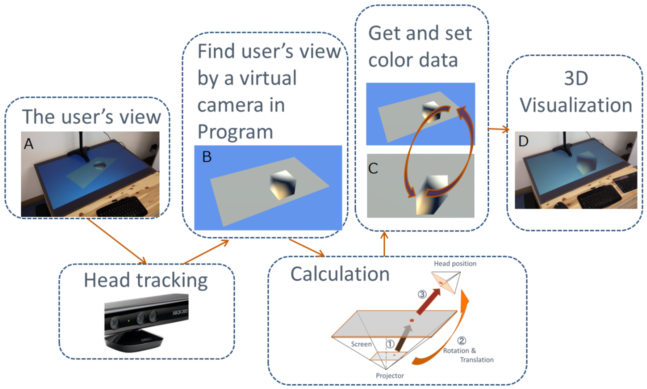

Figure 2: System overview

Figure 3: 3D interaction above the tabletop surface

Morphing Projection
$u = (x, y)$ in projection-plane coordinates is translated to program coordinates $U = (X, Y, Z)$ by a multiplier P: $U = Z \cdot P \cdot [u, 1]$.

P is usually called a projection matrix, where

$$P = \begin{pmatrix} 1/f_x & 0 & -p_x/f_x \\ 0 & 1/f_y & -p_y/f_y \\ 0 & 0 & 1 \end{pmatrix}.$$

Here, f_x, f_y are given by the ratios of the projection width in projection coordinates to the plane width in program coordinates, and p_x, p_y specify the center of projection. Z is the height of the plane in program coordinates. Let F be a representation of a view matrix in the program's camera and let $u' = (x', y')$ be the program's camera plane coordinates. Then, u' is given by a translation from u:

$u' = F \cdot Z \cdot P \cdot [u, 1]$.

In this way, we can find the relation between the projector and the XNA camera for each pixel. We then obtain and set color data for each pixel in real time. Finally, a correct projection will be specified.

3D Interaction
3D visualization assumes that the tabletop supports 3D interaction above the surface. At this stage, our tabletop interface supports simple 3D interactions involving collisions between the user's hand and virtual physical objects (see Fig. 3). In fact, a physical object is synchronized with the user's hand position, as detected by the Kinect, enabling the user to interact with the virtual objects. As mentioned below, we are planning an additional 3D interaction method.

Pilot Study
We conducted a pilot study to evaluate the performance of the system as an interactive tabletop and to obtain user feedback. In this section, we describe the procedure for this study and its results.

We prepared three types of projection (ProjectionA, Projection B, ProjectionC) using the same surface unit. Projection A is a fixed projection system, as used for many tabletop interfaces able to interact above the surface (e.g., [2].) In this system, the camera in program is fixed in position. Projection B is a projection that camera in program is just synchronized with the user's head position, as used for many interfaces with head tracking (e.g., [1].) Projection C is our 3D visualization system. We compared these three types of tabletop projection and evaluated their performance. We invited eight participants (five male, three female) between the ages of 23 and 30. The participants were required to hit some virtual objects their hands (Fig. 3) and score the performance of each system from 1 to 5 in each of the following respects.

- Direct interaction: How directly did the user feel he/she are touching the object?
- Reality: How well did the user feel immersed in the virtual world?
- Ease of use: How easily could the user manipulate the system?

Results
Fig. 4 shows the mean score for all three systems. A non-parametric statistical test (Friedman test) did not prove significant differences among the systems. However, our proposed system outperformed the other two systems in all the criteria. Many users complained that they could not feel intuitive A and B projection because the user cannot know projected objects's height or interact with them correctly.

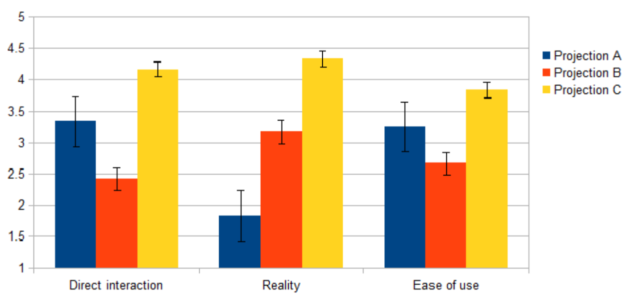

Figure 4: Result of the pilot study.Left: Direct Interaction. Middle: Reality. Right: Ease of use.

Discussion and Future Work

We have designed our first novel tabletop prototype and the system has many interaction possibilities. One of its unique strengths is the ability for users to immerse themselves in a mix of real and virtual worlds. This interplay will lead to interaction gaming possibilities or editing CG. Our system currently is not designed for multiple users because of the 3D visualization. However, we are planning to enable the system to multiplay by connecting two Kinects. Multiple users can share in and interact with a single virtual 3D scene, viewed from different perspectives, at each surface in real time. We are also planning to make the system mobile and implement additional interaction method. We have already tried to detect 3D positions of the user's finger tips by Kinect and going to implement this system.

References

[1] Hrvoje, B., and Andrew, D, W. *DepthTouch: Using Depth-Sensing Camera to Enable Freehand Interactions On and Above the Interactive Surface.* IEEE Work-shop on Tabletops and Interactive Surfaces (11-14), 2008.

[2] Otmar, H., Shahram, I., Andrew, D, W., Steve, H., Armando, G.-M., and Andreas, B. *Interaction in the Air: Adding Further Depth to Interactive Tbletops.* UIST (139-148), 2009.

[3] Papagiannakis, G., and Singh, G. *A survey of mobile and wireless technologies for augmented reality systems.* Comput. Animat. Virtual Worlds 19(3-22), 2008.

[4] Prachyabrued, M., and Borst, C, W. *Dropping the ball: Releasing a virtual grasp.* ACM VRST (59-66), 2011.