Interacting with 3D Model on Tabletop and Mobile Paper Projection

Yusuke Takeuchi[1,*] and Masanori Sugimoto[2]

[1] Department of Electrical Engineering and Information Systems
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
take.yusuke@gmail.com
[2] Department of Computer Science, Hokkaido University
Kita 14, Nishi 9, Kita-ku, Sapporo, Hokkaido, Japan
sugi@ist.hokudai.ac.jp

Abstract. In this paper, we present a system that enables a user to work with a virtual 3D information space on and above a tabletop by combining head-coupled perspective interaction with a mobile paper projection. The mobile paper projection acts as a physical pinhole into the virtual 3D scene. The proposed system is expected to be used for architectural tasks. As it is a work-in-progress study, we describe about the current status of the system and issues to be investigated in our future work.

Keywords: Interaction above a surface, Tabletop platform, Projection on paper, Architecture design task.

1 Introduction

Interaction with virtual 3D models via tabletop surfaces has been investigated and improved by a number of researchers. One major achievement has been the extension of their interaction space from a 2D surface (touch-based interaction) to a 3D space above the surface via intuitive and natural interactions. Extending a tabletop platform to the third dimension should be applicable to many areas, including medicine, education, entertainment, and so on. The system proposed in this paper is expected to be used in the field of architecture design.

An architect develops an idea for a 3D building and presents it to his/her clients. Before presenting the idea, several steps are usually required, such as drawing roughly-sketched blueprints, creating and editing 3D models using CAD software, and building physical architectural models[1]. Much money, time and efforts need to be spent for building architectural models. However, they often have to be modified or even completely rebuilt because of demands from the clients. An architectural model can be used for a range of purposes: to compare its size with surrounding objects, get an idea of how it

---

[*] Currently with IBM Japan.
[1] Personal communications with students in the Department of Architecture, University of Tokyo.

T. Yuizono et al. (Eds.): CollabTech 2014, CCIS 460, pp. 111–118, 2014.

looks from different angles, exhibit it to stakeholders, confirm light and shadow effects, and examine its texture design. Our proposed system is aimed to effectively achieve these various purposes. Two main contributions of this paper are: (1) manipulation techniques for 3D models to be used in architecture design tasks; (2) the integration of 3D visualization, 3D interaction, mobile projection and tabletop systems.

After describing previous research related to this work, we present an overview of our system and its design ideas. Details of the implementation and application scenarios are then described. Finally, we discuss how the system will be extended for collaborative work in our future work.

2 Related Work

2.1 Interaction above a Surface

A 3D object usually has six degrees of freedom (DOF), whereas a multi-touch interface supports only three-DOF manipulations, namely 2D translation on the surface and rotation about the normal vector on the surface. Therefore, multi-touch interaction is not well suited to manipulating 3D objects, which is why the input space for interfaces has been extended recently from a 2D surface to a space above the surface. Hilliges [4] proposes a tabletop system that enables a user to pick up 3D objects based on physics simulations in a 3D scene. The system recognizes a user's hand gesture with a finger-and-thumb circle by processing an IR image. When the user forms the circle with fingers above the surface, the system recognizes it as a 'pickup' gesture, enabling the user to pick up a 3D object. Wilson et al. [11] presents an interactive tabletop system that uses a depth camera to build a height map on the table surface. The height map is used in a driving simulation game that enables players to drive a virtual car over real objects placed on the table. Objects on the surface are captured and reconstructed by the depth camera. The cars and reconstructed 3D objects are controlled by a physics engine, enabling players to build a driving course by folding paper on the surface. In DepthTouch [3], a user can interact with a virtual object with both hands in the 3D space in front of a screen to a multi-touch interface. This work supports head tracking, which enables the user to obtain an immersive experience and to interact in more intuitive manners. By using several calibrated projectors and a depth camera, LightSpace [12] enables several users to interact with digital material, such as pictures, in a large space that includes desks and walls. To enable users to feel immersed in a 3D world, some interactions involve a head-coupled perspective display that renders images on one or more 2D displays with a perspective corrected for the user's view. The idea is based on fish tank virtual reality (FTVR) [10].

Holodesk [5] and Miragetable [2] combine FTVR techniques with 3D interaction. Holodesk is an interactive system combining an optical see-through display and a Kinect camera [1] to create the illusion that users are directly interacting with 3D objects under the screen. In contrast, Miragetable uses a curved projection screen to create a seamless projection on the table surface and the wall. The interaction techniques in both systems are based on 3D reconstruction, user's hand tracking and physics simulation.

2.2 Paper Interfaces

A handheld paper projection screen is a natural and useful interactive interface, and thus has been investigated so far. The mobile projection screen using a paper makes cross-sectional images of projected 3D objects available to a user. Paper Windows [6] presents a prototype windowing environment that simulates the use of digital paper displays. By projecting windows on physical paper with tracking marker, the user can see information, such as web pages or pictures. This work is inspired and motivated through natural manipulations that papers afford and supports a number of interaction methods, which include hold, collocate, collate, flip, rub, staple, point and two-hand pointing. Furthermore, uses can interact with the paper using pens, fingers or other objects by tracking them. Paperlens [8] presents an interactive interface, which uses a handheld paper described as a 'magic lens.' An IR camera and a projector are hung over the user's head. The IR camera tracks the paper in its 6 DOF and IR-reflecting markers are glued to the corners of the rectangular PaperLens to enable detection of the paper's exact rotation and orientation. PaperLens is extended to Tangible Windows [8], which enables to use a paper either as a physical pinhole into a virtual 3D world or as a physical container for part of that world. FlexPad [9] uses Kinect depth data and conducts real-time projection onto 3D deformable surfaces such as an office paper. It can identify hand occlusion based on reflectivity differences from surface materials.

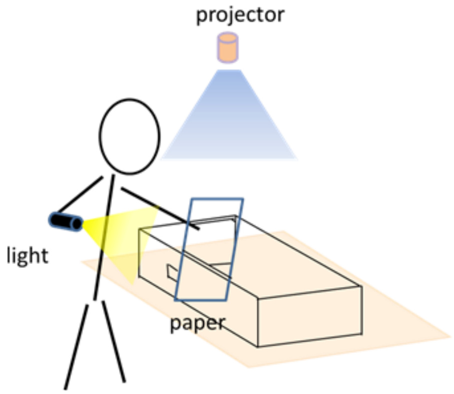

Fig. 1. An overview of the proposed system

3 System Overview

Figure 1 shows an overview of the proposed system. Its tabletop screen shows a 3D model that a user wants to edit or manipulate and its paper screen shows visual or textual information related the projected 3D model on the surface, such as a cross-sectional image or a wire-frame rendering. Each projection provides a user with a correct-perspective view from his/her viewpoint, which is called the head-coupled perspective [10].

The proposed system also supports several interaction techniques that help the user in conducting his/her tasks effectively. Kinect [1] is used to detect a user's head position and capture depth data representing the distance to physical objects such as a user's head or a paper for projection.

By moving the paper, the user can change projected information on its surface, which works as a physical pinhole into a virtual 3D world. Note that the user does not need to wear any devices, such as a head-mounted display or tracking markers. Also, the paper does not require any specific tracking marker - the user can use any hard paper for a physical pinhole window in our system. The proposed system works in the following way: First, all coordinates are unified into one real-world coordinate system by the checkerboard calibration proposed in [13]. After the calibration, 3D models are imported into or created in a 3D scene. Then, Kinect obtains the user's head position and depth data of a paper represented as point clouds, which are transformed from the Kinect coordinate system to the real-world coordinate system. Two Kinect cameras, one directed to the tabletop screen and the other to the paper screen are used and synchronized. When the paper, or the user's head position changes, a correct-perspective image projected on the paper surface is updated immediately. A type of paper projection, such as a cross-sectional image or a wire frame, is chosen by the user depending on a task that he/she is involved.

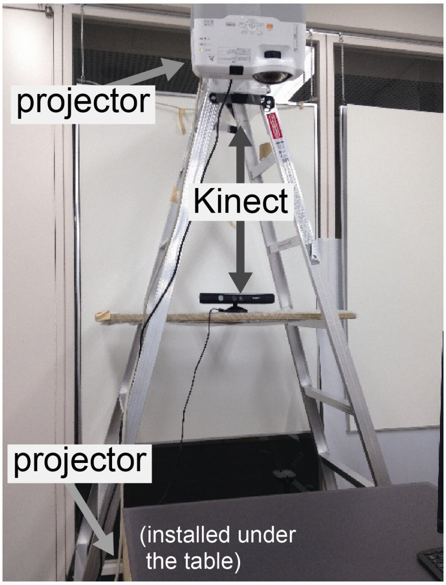

Fig. 2. System configuration: the Kinect acquires the user's current head position and two projectors are used, for the paper projection and the tabletop projection

4 Implementation

4.1 Configuration

Figure 2 shows the configuration of the proposed system. In addition to two Kinect cameras, two projectors, one for the tabletop projection and the other for the paper projection, are used. To avoid generating shadows because of occlusions, the projector for the tabletop projection is mounted under the surface. The Kinect cameras and the projectors are calibrated, with each coordinate system being unified into one real-world coordinate system [13]. The proposed system is currently implemented for single user. However, because of the nature of architecture design tasks, making the system usable for collaborative work by multiple users is inevitable. How the system will be extended is described in Section 5.

4.2 Correct Perspective

To allow a user to experience a 3D real world using our system, a head-coupled perspective is implemented (some VR displays support stereoscopic rendering of 3D scenes, but we do not focus on this aspect here.) Technically, this is achieved by setting the zero-parallax plane in off-axis perspective projections. The position of a camera in the 3D scene is synchronized with the user's head position and off-axis projections are applied. Instead of a stereoscopic image, we use a monoscopic view whose cues involve relative size, texture gradients, linear perspective, occlusion and motion parallax. By walking around the table, the user can feel immersed in a 3D world. The paper projection also supports a head-coupled perspective, enabling the user to work with the paper and the tabletop simultaneously.

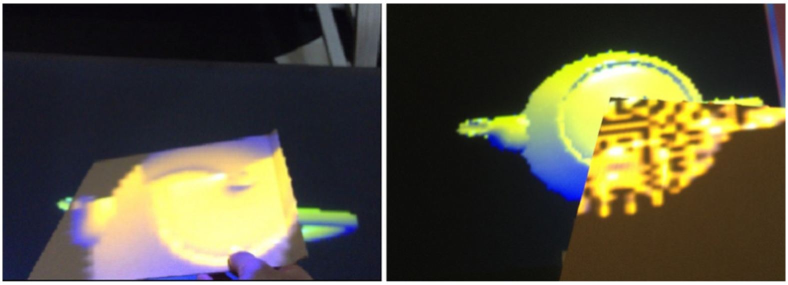

Fig. 3. Left: cross-sectional image, Right: wire-frame rendering

4.3 Paper Projection

A mobile screen using a paper works as a pinhole window into the 3D scene to provide a user with additional information. Just as if the user were actually in the 3D scene, the user can move the paper to an arbitrary position and orientation, thereby obtaining additional information about the projected 3D models on the tabletop screen, such as a cross-sectional image or an alternative rendering (see Figure 3).

These techniques are very useful for architectural tasks. Architects often need to investigate cross-sectional images on arbitrary planes from a variety of viewpoints. However, it is difficult to define plane equations and viewpoints in 3D CAD software. With the proposed system, the user can specify a plane equation and a viewpoint easily and naturally. As the paper has no specific devices attached, such as tracking IR markers, any size and shape of paper for the mobile pinhole window is available. A plane equation of the paper is calculated using its depth data, namely the distances between the paper surface and the Kinect camera. A bounding box is set in advance, enabling the Kinect to ignore areas outside the box, and obtain depth information about the real objects inside the bounding box. Then, the plane equation of the paper is fixed by applying the least square method to point clouds inside the bounding box. To eliminate the user's hand and specify the equation more precisely, points whose distance from the result of the least-squares method exceeds a threshold are eliminated, and the remaining points are processed by the least-squares method again. After iterating this process several times (twice or three times are usually sufficient), we obtain an accurate paper equation. Besides, to avoid overlapping display on the tabletop projection from the paper projection, the region outside the projection of the paper should be black. We prepared a binary image to project exclusively on the object's surface inside the bounding box (see Figure 4, (b)) and then, we subtract the black part in the binary image from the paper projection (see Figure 4).

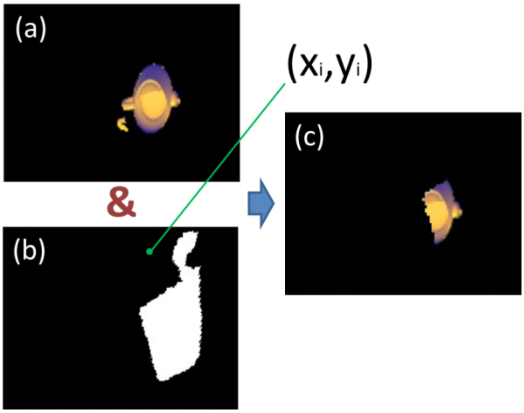

Fig. 4. (a):a paper projection before computing (b):a binary image of the objects above the surface (c):a result image by subtract (b) from (a)

4.4 Interaction Using Mobile Paper

A couple of interaction techniques using the mobile paper have been implemented at the moment. When a user conducting architecture design tasks uses the paper to obtain additional information, he/she performs manipulations to a 3D model so that the paper projection shows its detailed information. Figure 5 shows example interaction techniques implemented in this system ('pick up' and 'release'). These techniques enable the user to manipulate a 3D model such as rotation and translation by tilting and moving the paper, respectively. Even though the user can only rotate the 3D model while the paper's equation remains detectable (if the paper is vertical to the Kinect camera, for example, its plane equation is not computable), in many cases the user can conduct manipulation tasks using the paper.

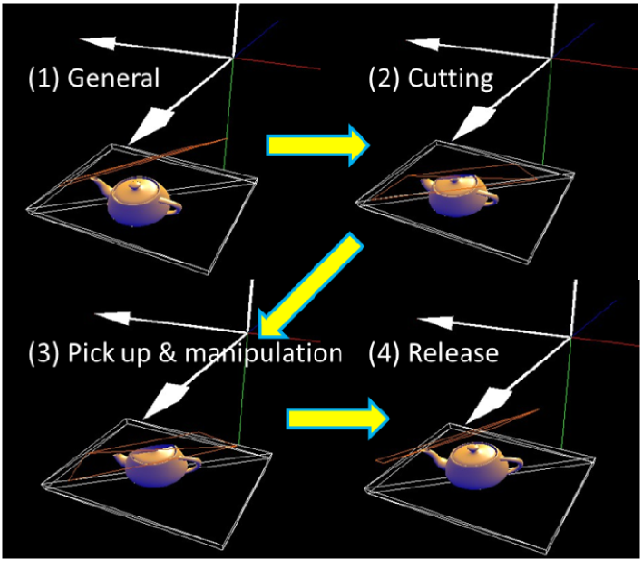

Fig. 5. Interaction using the mobile paper: the 3D model is synchronized with the movement of the paper by picking up mode

5 Future Work and Conclusion

In this paper we present an interactive interface that integrates a tabletop platform with mobile paper projection. The paper projection works as a pinhole window to show additional information and to enable a user to manipulate a 3D model intuitively and naturally.

As it is an ongoing work, many issues remain to be investigated. In architecture design tasks, there are several situations, for example, a designer solely explores his plan by manipulating an architectural model, a designer and developers discuss about construction processes using the model, a designer explains about his plan by showing details of the model, and so on. To extend the system for multiple users, it needs to conduct multi-user head tracking. Also the current tabletop projection may have to be changed to a normal projection so that all the users can see the same image or switched between individual users' correct-perspective projection. Using multiple papers for 3D model manipulation and projection by individual users is not within the scope of our implementation plan, because the need for such functions in architecture design tasks is not clear at the moment. Using a flexible paper as proposed in [9] to show cross-sectional images may make our application scenarios more convincing. We plan to ask professional architectural designers for intensive user studies to evaluate the system.

Acknowledgement. This work has been supported by JSPS Kakenhi Grant Number 25282048.

References

1. Kinect for Windows, http://www.microsoft.com/en-us/kinectforwindows/
2. Benko, H., Jota, R., Wilson, A.: Miragetable: Freehand Interaction on a Projected Augmented Reality Tabletop. In: Proc. of CHI 2012, Austin, TX, pp. 199–208 (2012)
3. Benko, H., Wilson, A.D.: DepthTouch: Using Depth-Sensing Camera to Enable Freehand Interactions On and Above the Interactive Surface, Microsoft Technical Report, MSR-TR-2009-23 (2009)
4. Hilliges, O., Izadi, S., Wilson, A.D., Hodges, S., Garcia-Mendoza, A., Butz, A.: Interactions in the Air: Adding Further Depth to Interactive Tabletops. In: Proc. of UIST 2009, Victoria, Canada, pp. 139–148 (2009)
5. Hilliges, O., Kim, D., Izadi, S., Weiss, M., Wilson, A.: Holodesk: Direct 3D Interactions with a Situated See-through Display. In: Proc. of CHI 2012, Austin, TX, pp. 2421–2430 (2012)
6. Holman, D., Vertegaal, R., Altosaar, M., Troje, N., Johns, D.: Paper Windows: Interaction Techniques for Digital Paper. In: Proc. of CHI 2005, Portland, OR, pp. 591–599 (2005)
7. Spindler, M., Büschel, W., Dachselt, R.: Use Your Head: Tangible Windows for 3D Information Spaces in a Tabletop Environment. In: Proc. of ITS 2012, Cambridge, MA, pp. 245–254 (2012)
8. Spindler, M., Stellmach, S., Dachselt, R.: Paperlens: Advanced Magic Lens Interaction above the Tabletop. In: Proc. of ITS 2009, Banff, Canada, pp. 69–76 (2009)
9. Steimle, J., Jordt, A., Maes, P.: Flexpad: Highly Flexible Bending Interactions for Projected Handheld Displays. In: Proc. of CHI 2013, Paris, France, pp. 237–246 (2013)
10. Ware, C., Arthur, K., Booth, K.S.: Fish Tank Virtual Reality. In: Proc. of CHI 1993, Amsterdam, Netherlands, pp. 37–42 (1993)
11. Wilson, A.D.: Depth-Sensing Video Cameras for 3D. In: Proc. of ITS 2007, Newport, RI, pp. 201–204 (2007)
12. Wilson, A.D., Benko, H.: Combining Multiple Depth Cameras and Projectors for Interactions on, above and between Surfaces. In: Proc. of UIST 2010, New York, NY, pp. 273–282 (2010)
13. Zhang, Z.: A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1330–1334 (2000)