AR Tabletop Interface using a Head-Mounted Projector

Yusuke Kemmoku*  
Saitama University

Takashi Komuro†  
Saitama University

Abstract

In this paper, we propose a tabletop interface in which a user wears a projector with a depth camera on his or her head and can perform touch operations on an image projected on a flat surface. By using the head-mounted projector, images are always projected in front of the user in the direction of the user’s gaze. By changing the image to be projected based on the user’s head movement, this interface realizes a large effective screen size. The system superimposes an image on the flat surface by performing plane detection, placing the image on the detected plane, performing perspective projection to obtain a 2D image, and projecting the 2D image using the projector. Registration between the real world and the image is performed by estimating the user’s head pose using the detected plane information. Furthermore, touch input is recognized by detecting the user’s finger on the plane using the depth camera. We implemented some application examples into the system to demonstrate the usefulness of the proposed interface.

Keywords: Augmented reality, peephole interaction, projector-camera system, depth camera.

Index Terms: Human-centered computing [Human computer interaction (HCI)]: interaction paradigms—Mixed / Augmented reality

1 Introduction

In recent years, mobile devices, such as smartphones and tablets, and systems using large displays and keyboards, such as PCs, have become widely used. While mobile devices have the advantage of portability, one problem is their small operating areas because of their small screens. In contrast, while systems using a large display and a keyboard have good operability, it is difficult to carry them around. Therefore, there is demand for systems that can realize both good portability and large operating areas.

As related work, there have been studies of user interfaces (UIs) using projectors [4, 6, 14, 18]. By using a projector, it is possible to display an image on a large display area. A user can interact with the image by touching an image projected on a plane. However, to increase the size of the display area, there is a constraint in terms of the installation locations, namely, that the projection plane and the projector should be separated by a large distance.

To solve this problem, there have been studies of UIs using a handheld projector [3, 9]. By changing the image to be projected according to the projector’s position and direction, it is possible to extend the display area. However, this approach has some problems; for example, one hand is busily occupied holding the projector, and images are not always projected in front of the user in his or her gaze direction.

On the other hand, there have been studies of UIs using a head-mounted projector (HMP) to constantly display images in front of the user in his or her gaze direction [2, 7, 8, 10, 13, 16, 19]. In these studies, it is possible to present images in front of the user in his or her gaze direction because a projector is worn on the head. However, there are no systems in which an HMP is used for the purpose of projecting images in a wide area.

Therefore, we propose a tabletop interface that projects images on a flat surface such as a table by using an HMP and that realizes a large effective display area by changing the image to be projected according to the direction of the user’s head.

The system obtains a 3D scene using a depth camera attached to the projector and performs plane detection. In the virtual world, the system places an image to be superimposed on the detected plane and transforms the plane into a 2D image by perspective projection. Then, in the real world, the system projects the 2D image on the surface. Registration between the real world and the superimposed image is performed by estimating the user’s head pose using the detected plane information. Furthermore, touch input is recognized by detecting a finger on the surface using the depth camera.

Another method for constantly displaying images in front of the user in his or her gaze direction is to use a head-mounted display (HMD) [5, 12]. By wearing an HMD, the user can interact with virtual objects displayed on a plane or in the air in a real space. However, HMDs have the disadvantages that the user has to wear the device on his or her head, and some people are reluctant to do so because of the appearance. In addition, the user’s eyes are covered by the device, hindering their vision, which may cause a feeling of anxiety.

Also in the case where a projector is worn on the head, there are similar problems to wearing an HMD, but the user’s eyes are not covered by the device. Moreover, eye fatigue can be suppressed because a mismatch between the convergence angle and focusing of the eyes does not occur, which is another problem with HMDs. On the other hand, HMDs give a superior sense of immersion and ensure confidentiality since the operations carried out by the user are not seen by someone else.

2 AR Tabletop Interface using a Head-Mounted Projector

In this paper, we propose a tabletop interface in which a user wears a projector with a depth camera on his or her head and can perform touch operations on an image projected on a flat surface. The user sits in a chair, wearing the system on his or her head, and the system projects images on a flat surface, such as a table. This system has two features. First, since the projector worn on the user’s head projects images in front of the user in his or her gaze direction, the images cover a large part of the user’s central area of vision, as shown in Figure 1. Second, the image to be projected is changed according to the user’s head position so as to fix the superimposed images on the surface. Thus, the system makes the flat surface act like a touch panel using the depth camera and realizes a large effective display area. Also, to enable interaction with the images superimposed on the flat surface, this system detects touch input by the user.

2.1 System

We created an experimental system using a mobile projector and a depth camera, as shown in Figure 2. The mobile projector was a Koolertron Mini Micro 640×480 HD DLP Home Theater Projector. The depth camera was a SoftKinetic DepthSense 325, which can acquire a 320×240 pixel depth map at 60 fps and can be used in a
15–100 cm range. Also, in this system, the depth camera and the projector were connected to a desktop PC (CPU: Intel Core i7-4790 3.6 GHz, RAM: 8 GB), which was used for calculations.

2.2 Superimposing images

The system superimposes an image that seems to have the proper size and shape on a flat surface. First, the system obtains a 3D scene in the real world by using the depth camera (Figure 3(a)) and performs plane detection from the obtained 3D scene. Second, the system replicates, in a virtual world, the detected plane (Figure 3(b)). Then, the system places the image to be superimposed on the plane and transforms the plane into a 2D image by perspective projection (Figure 3(c)). When the system performs perspective projection, the viewing angle is set to the projector’s viewing angle. Finally, the system projects the 2D image on the flat surface in the real world (Figure 3(d)). By using this method, the system can superimpose an image on the flat surface.

Figure 4 shows the result of superimposing an image on a flat surface using the proposed method. Figure 4 (a) shows the positional relationship between the projector and the flat surface. Figure 4 (b) shows the result of projection without using the proposed method. Figure 4 (c) shows the result of projection using the proposed method. In (c), we find that the image is superimposed on the flat surface while keeping the same size and shape, even when the system is tilted.

2.3 Registration

To fix the image on a flat surface, registration between the real world and the image is performed. To perform registration without using markers, algorithms such as ICP [1], which performs registration between two 3D point clouds, and PTAM [11], which finds correspondence of the feature points between frames, are widely used. However, in the proposed system, we assume a flat surface with little texture, and therefore, we cannot use these algorithms. Also, there is a registration method using an inertial measurement unit (IMU), but careful calibration between devices is required for accurate registration. Therefore, the proposed system performs registration between the real world and images by estimating the user’s head pose using only information of the plane detected from the 3D scene.

In the proposed system, the user is seated in a chair, and the images are projected on a flat surface such as a table. Therefore, we assume that the head movement has only two rotational degrees of freedom (DOF): pan and tilt. The 2-DOF pose can be calculated using detected plane information.

The system calculates rotation angles around the x-axis and the y-axis, $\alpha$ and $\beta$, that make a vector $(0, 1, 0)$ correspond to a normal vector $(a, b, c)$ of the detected plane.
Then, the system calculates the relative head pose at a certain time $k$ from the first plane detection. A rotation matrix $R_k$ of the head pose at a time $k$ from the initial pose is calculated using the rotation matrices.

The camera undergoes translation when the head is rotated because the rotation center of the head does not correspond to the center of the camera. Figure 5 shows translation of the camera from the rotation center of the head. Therefore, the camera position, $t_k$ from the rotation center of the head when the head rotates by $R_k$ is calculated.

Using $R_k$ and $t_k$, the transformation matrix $T_k$ from the head coordinate system with the origin at the rotation center of the head to the camera coordinate system at $k$ is calculated. Using $T_k$ and $T_0$, which is the transformation to the camera coordinate system at time 0, when the system detects a plane for the first time, the virtual object position is calculated.

![Figure 5: Camera translation by head rotation.](image)

Figure 6 shows a user demonstrating the system. The user was seated, and the images projected on the table in front of the user were taken from behind. Figure 7 shows images projected on the table with the user’s head moving using the methods involving superimposing images and registration.

![Figure 6: A user demonstrating the system.](image)

![Figure 7: Images displayed when the user’s head moved, taken from behind the head.](image)

### 2.4 Touch detection

To enable interaction with the images superimposed on the flat surface, the system detects touch input when the user’s finger touches the surface. Thereby, the system allows the projection surface to act like a touch panel.

There are methods for detecting contact with a surface using a depth camera. A touch detection method using a depth image obtained from a depth camera has been proposed [17]. In this method, background subtraction of depth images is used. However, there is a problem that it is difficult to detect the user’s fingertips because the depth of the flat surface and that of the fingertips at the time of touch are close. To solve this problem, a touch recognition method using an IR image to detect the edge of the fingertips has been proposed [15]. However, these touch recognition methods using background subtraction cannot be applied to our system because the depth camera is not fixed, and the background images change from frame to frame. Therefore, we extended these existing touch recognition methods to the case where the depth camera is not fixed.

Instead of performing background subtraction, our method compares the average depth in the region inside the fingertip and that near outside the fingertip. First, the system obtains an edge image from the IR image which is captured by the depth camera, and fills inside a fingertip by applying closing to the edge image. Then, the fingertip region is extracted by finding a white pixel by scanning the closing image from upper left. Finally, the system detects touch input when the difference between the average depth in the region inside the fingertip and that near outside the fingertip is lower than a threshold. Figure 8 shows the system recognizing touch input. A user is touching a virtual button that is projected by the projector.

![Figure 8: Touching a virtual button.](image)

### 3 Applications

We implemented some application examples into the system to demonstrate the usefulness of the proposed interface.

Figure 9 (a) shows the main screen of a GUI. Thumbnails of application windows are placed in the screen, and a user can enlarge the windows to the full screen by touching them. In this demonstration, the user can use applications such as a photo viewer, a document viewer and a map viewer.

Figure 9 (b) shows an image of the photo viewer (left) and the user operating the application (right). The user can see an entire group of photos by moving the user’s head to look around the screen. The user can also see the detail of each photo due to a large effective screen size. The user can turn pages of the album and go back to the main screen by touching buttons in a control panel which is located at the bottom of the screen.

Figure 9 (c) shows the document viewer. The user can see both an entire page of a document and the details of contents. The user can turn pages by touching buttons in a control panel.

Figure 9 (d) shows the map viewer. When the user touches a part of a wide-area map displayed in the screen, the detailed map around the touched position with information of buildings is displayed. The photo of a building is displayed by touching it.
of wearing the device. As an earphone-shaped one, which would reduce the disadvantages it should be possible to manufacture a more compact system such to verify the effectiveness of the proposed interface.

The system superimposes an image on the flat surface by performing plane detection, placing the image on the detected plane, performing perspective projection to obtain a 2D image, and projecting the 2D image using the projector. Registration between the real world and the image is performed by estimating the user’s head pose using the detected plane information. Furthermore, touch input is recognized by detecting the user’s finger on the plane using the depth camera. We implemented some application examples into the system to demonstrate the usefulness of the proposed interface.

The system superimposes an image on the flat surface by performing plane detection, placing the image on the detected plane, performing perspective projection to obtain a 2D image, and projecting the 2D image using the projector. Registration between the real world and the image is performed by estimating the user’s head pose using the detected plane information. Furthermore, touch input is recognized by detecting the user’s finger on the plane using the depth camera. We implemented some application examples into the system to demonstrate the usefulness of the proposed interface.

Future work includes recognition of multi-touch input which enables various interactions such as pinching in and out. In addition, usability evaluation involving a number of participants is required to verify the effectiveness of the proposed interface.

In the future, with the miniaturization of projectors and cameras, it should be possible to manufacture a more compact system such as an earphone-shaped one, which would reduce the disadvantages of wearing the device.

4 Conclusion

In this paper, we proposed a tabletop interface that projects images on a flat surface, such as a table, by using a projector mounted on the user’s head and that realizes a large effective display area by changing the projected image according to direction of the user’s head.

Figure 9: Application examples. (a) main screen of a GUI, (b) photo viewer, (c) document viewer, (d) map viewer.

4 Conclusion

In this paper, we proposed a tabletop interface that projects images on a flat surface, such as a table, by using a projector mounted on the user’s head and that realizes a large effective display area by changing the projected image according to direction of the user’s head.

The system superimposes an image on the flat surface by performing plane detection, placing the image on the detected plane, performing perspective projection to obtain a 2D image, and projecting the 2D image using the projector. Registration between the real world and the image is performed by estimating the user’s head pose using the detected plane information. Furthermore, touch input is recognized by detecting the user’s finger on the plane using the depth camera. We implemented some application examples into the system to demonstrate the usefulness of the proposed interface.

Future work includes recognition of multi-touch input which enables various interactions such as pinching in and out. In addition, usability evaluation involving a number of participants is required to verify the effectiveness of the proposed interface.

In the future, with the miniaturization of projectors and cameras, it should be possible to manufacture a more compact system such as an earphone-shaped one, which would reduce the disadvantages of wearing the device.

REFERENCES


