Abstract
In this paper, we propose an interaction system which displays see-through images on the mobile display and that allows a user to interact with virtual objects overlaid on the see-through image using the user’s hand. In this system, the camera which tracks the user’s viewpoint is attached to the front of the mobile display and the depth camera which captures color and depth images of the user’s hand and the background scene is attached to the back of the mobile display. Natural interaction with virtual objects using the user’s hand is realized by displaying images so that the appearance of a space through the mobile display is consistent with that of the real space from the user’s viewpoint. We implemented two applications to the system and showed the usefulness of this system in various AR applications.

Author Keywords
Augmented reality; mobile device; geometric consistency; depth camera

ACM Classification Keywords
H.5.1 [Multimedia Information System]: Artificial, augmented, and virtual realities.

Figure 1. Interaction with a virtual object using the See-through Mobile AR System.
Introduction
Recently, systems which enable a user to interact with virtual objects using the Augmented Reality (AR) technology have been developed. Among them, systems with a Head Mounted Display (HMD) enable a user to walk around in a wide area. However, such systems have a problem that it is bothersome to wear the device on the head.

To solve this problem, AR interaction systems using a mobile device have been developed. In [1], the user puts the user's hand behind a smartphone and the hand seen by the camera is displayed on the screen. The user can manipulate a virtual object displayed on the same screen with the user's hand. However, because of the difference between the viewpoint position of the user and the position of the camera, the appearances of the real space and a space through the screen of the smartphone are inconsistent. This causes two problems: the field of view of the user becomes narrow since the user gazes only at the screen, and the user has a feeling of strangeness since the position of the user's hand in the screen is different from the user's somatic sense.

On the other hand, there has been research to match the appearances of a space through the display of the mobile device and the real space [2, 3, 4]. Since the appearance of a space through the mobile display is consistent with that of the real space, the user can have a wide field of view. However, they have realized only transformation of images according to the user's viewpoint and have not overlaid virtual objects, and interaction with virtual objects using the user's hand has not been realized.

Therefore, we propose a system which consists of a mobile display, a front-mounted camera, and a rear-mounted depth camera, and that enables the user to interact with virtual objects. The system executes face tracking and obtains the viewpoint position of the user. The system obtains a three-dimensional scene by the depth camera on the back of the mobile display, and projects it on the display as if the display were transparent from the viewpoint position. Since the appearances of the space through the mobile display and the real space are consistent, it seems that the spaces inside and outside the display are connected and the user can have a wide field of view. Moreover, since the user can see the user's hand in the correct position through the display, the user does not have a feeling of strangeness. The position of the user's hand is obtained using a depth camera and the user can interact with virtual objects in a three-dimensional space. Thus, natural interaction is realized even through the screen of the mobile display.

System Configuration
Figure 2 shows the appearance of the system. The system consists of three components: a mobile display, a camera on the top of the front of the mobile display for face tracking, and a depth camera on the center of the back of the mobile display for obtaining the depth and color images of the user's hand and the background scene. The depth camera can detect the range from 15 centimeters to one meter. We also used a PC for computation and image rendering.
Generation of See-through Images
To make the appearance of a space through the mobile display consistent with that of the real space, the system needs to obtain the viewpoint position of the user. The system obtains the two-dimensional viewpoint position and the distance between the user’s both eyes in the image by face tracking. In this system, the position of the user’s dominant eye is regarded as the viewpoint position. The system obtains the three-dimensional viewpoint position of the user by computing the depth of the face from the real distance between the user’s eyes. The computed viewpoint position is used to render the images. As illustrated in Figure 3, the three-dimensional scene which is obtained by the depth camera is projected on the display according to the viewpoint position of the user.

Figure 3. Generation of see-through images.

An image of face tracking is shown in Figure 4A. At this time, the image displayed on the mobile display is obtained by drawing planes with adjacent points which are obtained by mapping the smoothed image of the depth image shown in Figure 4B and the color image shown in Figure 4C captured by the depth camera, and it is shown in Figure 4D.

The images from different viewpoint positions by using the proposed method are shown in Figure 5. They show that the images displayed on the mobile display are consistent with the real space wherever the viewpoint positions of the user are.

Figure 4. A: face tracking image, B: depth image, C: color image, D: generated image from the user’s viewpoint

Figure 5. Images from the user’s viewpoint.
Bare-hand Interaction using a Depth Camera

The user interacts with virtual objects in the three-dimensional space using the point cloud data obtained by the depth camera. A virtual object is approximated to a sphere, and if there are points inside the sphere, the distances from the all points in the area to the surface of the sphere are computed. Then, the virtual object is moved the mean of the distance. Figure 6 shows an example of overlaying a virtual object on the image. The user can touch the ball overlaid on the image of the real space using the user’s hand.

Figure 6. Overlaying a virtual object.

Applications

We implemented two applications to the proposed system. The first one is the bowling game shown in Figure 7A. The user pushes the ball which is overlaid on the display and the ball hits the three pins in the back. Depending on the position where the user touches, rolling direction of the ball changes. The second one is AR Catalog shown in Figure 7B. In this application, when the user pushes a button overlaid around the virtual car, the color of the car changes to the color of the button. Since the car is rotating in a horizontal direction, the user can confirm the coloring of the car from various angles. Moreover, when the user pushes the circle button in the upper right, the car stops rotating and the user can manipulate the car with the user’s hand as shown in Figure 7C. As seen above, it is possible to create entertainment and practical applications using the proposed system.

References


Figure 7. A: Bowling game, pushing the ball to hit the pins. B: Pushing the virtual color buttons for changing the color of the car. C: Manipulating a virtual car.