ABSTRACT
In this study, we propose a method to present an image that maintains geometric consistency between the actual scene outside the mobile display and the camera image. Thereby, we expect that interaction with the virtual object through the mobile display becomes more intuitive, and the operability is improved. By cameras mounted on the front and back of the mobile display, the user’s face position and the distance to the subject are obtained. Using the information, it is possible to present an image that maintains geometric consistency between the inside and outside of the display depending on the user’s viewpoint.

Categories and Subject Descriptors
H.5.1 [Multimedia Information System]: Artificial, augmented, and virtual realities.

General Terms
Experimentation

Keywords
Augmented reality, mobile device, geometric correction, projective transformation, geometric consistency

1. INTRODUCTION
Mobile devices like smartphones have become widespread, but they have narrow operation area and are hard to use because of its smallness. Therefore, research of shifting the operation area to actual space using the technology of augmented reality (AR) is under active investigation. For example, there is a system which allows the user to manipulate objects that are projected onto the real space using a projector[1]. However, it has a drawback that the objects are restricted by the shape and characteristics of the projection surface. Meanwhile, research of AR using a mobile display has been also widely conducted[2]. However, as a unique problem of mobile AR, it is difficult for the user to recognize the locational relationship between the actual scene and a virtual object which was superimposed on the screen since it does not keep geometric consistency between the inside and outside of the display as seen from the viewpoint of the user. Even though there is a study that solved the problem[3], the user has to wear a device on the head. In this study, we propose a method to present an image on a mobile display that maintains geometric consistency between the inside and outside of the display wherever the user is seeing from.

2. SYSTEM CONFIGURATION
Small USB cameras are attached on the front and back sides of the 10.1-inch mobile display. Figure 1 is the appearance of the system. The camera that is attached on top of the front side captures the user’s face, while the camera that is attached at the center of the back side captures the user’s hand which manipulates the virtual object. The location of the viewpoint and the distance to the hand are needed to present an image that keeps geometric consistency. In this study, we used AR markers to get them. Though the consistency will be kept only in the plane on which the marker is, that would be sufficient for interaction.

3. CORRECTION OF CAMERA IMAGES
M1 is the marker attached to the hand, M2 the marker attached to the face, C1 the back camera and C2 the front camera. Figure 2 shows the positional relationship among the markers and cameras.

By using an internal parameter matrix A of the cameras, a rotation matrix R and translation vector t which are position and orientation information of a marker seen from a camera, the
formula for calculating the projective transformation matrix $H$, which transforms the marker plane coordinates into two-dimensional coordinates of the camera image, is as follows.

$$
egin{pmatrix}
    x' \\
    y' \\
    1
\end{pmatrix} =
A(R, t)
\begin{pmatrix}
    x \\
    y \\
    0
\end{pmatrix} =
A
\begin{pmatrix}
    r_{11} & r_{12} & t_x \\
    r_{21} & r_{22} & t_y \\
    r_{31} & r_{32} & t_z
\end{pmatrix}
\begin{pmatrix}
    x \\
    y \\
    1
\end{pmatrix} =
H
\begin{pmatrix}
    x \\
    y \\
    1
\end{pmatrix}
$$

Using the formula above, $H_{M1}$ is calculated from the rotation matrix $R_{M1}$ and translation vector $t_{M1}$ which are obtained from the marker attached to the hand. Similarly, $H_{M2}$ is calculated from $R_{M2}$ and $t_{M2}$ which are obtained from the marker attached to the head. Moreover, $H_{M1,M2}$ is calculated from $R_{M1}$ and $t_{M1}$ which are position and orientation information of the hand seen from the face, and that are calculated from $R_{C1}$, $t_{C1}$, $R_{C2}$, and $t_{C2}$ using the following formula.

$$
\begin{align*}
R_{M1} & = R_{C1} R_{C2} \\
M_{M1} & = M_{C1} R_{C2} + M_{C2}
\end{align*}
$$

The projective transformation matrix $H$, which transforms the camera image coordinates into the display plane coordinates, is derivable from the following formula.

$$
H_{M1,M2} = H_{M2} H_{C2} H_{M1} H_{C1}
$$

Meanwhile, $H$ is represented as

$$
H = H_{disp} H_{img}
$$

$H_{disp}$ is the projective transformation matrix which transforms the display image coordinates into the display plane coordinates, and $H_{img}$ is the projective transformation matrix which transforms the camera image coordinates into the display image coordinates, and is used for image transformation which keeps geometric consistency in accordance with the user’s viewpoint.

Since the directions of the y-axis and z-axis of the coordinate system of marker coordinate system is different from those of camera coordinate system, coordinate transformation is conducted so that they become the same direction. In addition, since the positions of two cameras are different, coordinate transformation is conducted so that the camera attached to the front side of the display moves to the center of the back side.

4. EXPERIMENTAL RESULTS

We verified whether the geometric consistency between the inside and outside of the mobile display is maintained or not. To show the results clearly, we conducted an experiment using a standing object instead of a hand to which the marker is attached. Figure 3 is the picture of the experiment. Figure 4 shows the results when the user moves his/her viewpoint, and Fig. 5 shows the results when the user moves the display. In both cases, you can see that the user can see the image which is aligned with the actual scene as if the display were transparent like a glass plate.

5. REFERENCES

