Case Studies of See-Through Augmentation in Mixed Reality Project

Kiyohide Satoh, Toshikazu Ohshima, Hiroyuki Yamamoto,
and Hideyuki Tamura

Mixed Reality Systems Laboratory Inc.
6-145, Hanasaki-cho, Nishi-ku, Yokohama 220-0022, Japan
E-mail: {ksato, ohshima, ymmt, tamura }@mr-system.com
URL: http://www.mr-system.com/

Abstract

This paper introduces two case studies of augmented reality (AR) system, which use see-through HMDs. The first case is a collaborative AR system called AR^{2}Hockey that requires real-time interactive operations, moderate registration accuracy, and a relatively small registration area. The players can hit a virtual puck with physical mallets in a shared physical game field. The other case study is the MR Living Room where participants can visually simulate the location of virtual furniture and articles in the physically half-equipped living room. The registration is more crucial in this case because of the requirement of visual simulation. As well as the system configurations of both systems, the details of registration algorithms implemented are described.

Keywords: Augmented reality, mixed reality, collaboration, visual simulation, registration, see-through HMD

1 Introduction

Most of VR systems we have experienced in this decade made it possible for participants to interact with virtual environments which are totally synthesized in computers. As everyone knows, the reality in synthetic world is limited in its nature. Because of this limitation, people tend to positively utilize rich information in the real world.

We have been participating in the Research Project on Mixed Reality (MR) whose target is to develop the technology merging the real world and the virtual world seamlessly. Mixed Reality Systems Laboratory Inc. was established to conduct the project in January 1997 and is planned to be extended to March 2001. The research topics include the display equipment, such as the HMD (Head Mounted Display) and the 3-D display without eyeglasses, and the software technologies such as the registration of both worlds and the system architecture.

MR [1] is a concept that covers "augmented reality (AR)" and "augmented virtuality (AV)." Although both AR and AV handle the physical world and virtual one simultaneously, AR is based on the physical world and AV is constructed on the virtual space. There is no clear distinction, however, between AR and AV, and MR is a "virtual continuum." The geometrical registration between the physical space and the cyberspace is a common problem in MR. The ways to combine both spaces photometrically, however, are slightly different according to applications. This paper introduces two case studies of AR system with see-through HMDs and the registration algorithms developed in the first one and half years of the MR project.

The first case is a collaborative AR system that requires real-time interactive operations, moderate registration accuracy, and a relatively small registration area. The collaborative AR allows multiple participants to simultaneously share a physical space surrounding them and a virtual space, visually registered with the physical one. They can also communicate with each other through the mixed space. AR^2Hockey (Augmented Reality AiR Hockey)[2] is described as a case study of the collaborative AR system where players can hit a virtual puck with physical mallets in a shared physical game field. Image quality is not a serious factor in this case.

The other case study is the MR Living Room where participants can visually simulate the location of virtual furniture and articles in the physically half-equipped living room. In this case, we have to render many objects photo-realistically and support a wide registration area so that the participants can walk around and make visual simulation. That is, the registration between the physical space and the virtual space is more crucial in this case than in the AR^2Hockey. A different type of head-tracker and optical see-through HMDs with wider field of views, compared to the former case, are equipped in this system.

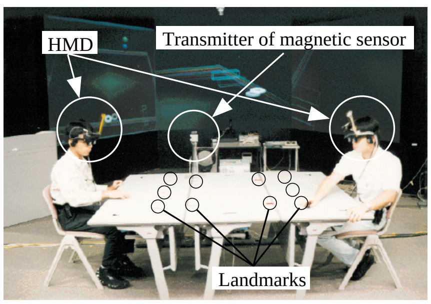
For each system, an appropriate video-rate registration algorithm is implemented with head-trackers and video cameras attached to see-through HMDs. Especially in the MR Living Room, we have proposed a new registration framework in AR systems.

2 Collaborative AR

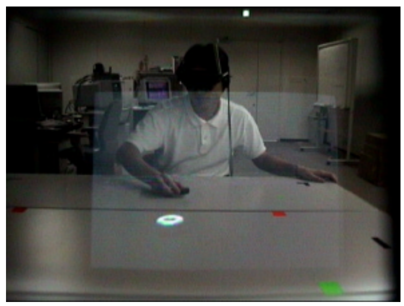
AR researches have so far been made mainly on single-user applications [3][4][5]. New application fields, especially in the field of human communication, will appear if multiple participants can share a physical space and if we can seamlessly offer a virtual space into the shared physical space[6]. For example, it becomes possible for multiple people to collaborate for object design in the physical space while exchanging their ideas through virtual objects[7].

We have developed the AR^{2}Hockey system [2] as a case study of the collaborative AR for human communications. This section describes the configuration of the AR^{2}Hockey. Air hockey is a game in which two players hit a puck with mallets on a table and shoot it into goals. In our AR^{2}Hockey, the puck is in a virtual space. This simple application challenges to the following problems of the collaborative AR. Firstly, more than two persons share a single physical space and a virtual space. Secondly, since the puck moves fast, the response time becomes severe and the synchronization problem should be solved. Thirdly, since the virtual puck is hit by the physical mallets, the positional error and the time lag must be minimized.

Figure 1(a) shows the scene of playing the AR^{2}Hockey and (b) is an image seen through the HMD while the system is operating.



(a) Playing scene



(b) Augmented view (optical see-thru)

Figure 1: Playing scene of AR^{2}Hockey

2.1 Hardware components

HMDs

This system uses optical see-through HMDs[8]. The HMD contains an LCD of 180,000 pixels and two prisms for each eye. One prism is used to lead images displayed on the LCD to the eye. This prism has two off-axial reflective surfaces. To correct the off-axial aberrations, the aspherical surface without rotational symmetry is used in this prism. Attaching the compensation prism to the outside of the prism, good see-through view is achieved. This HMD gives us 34 degrees of horizontal view angle and 22.5 degrees of vertical one.

Trackers

The HMD uses a magnetic sensor (Polhemus' Fastrak) to measure the player's viewpoint. Since this positional and orientational sensor dose not have enough accuracy to produce images without notable displacement, we have placed a small color CCD camera (ELMO) having 45 degrees of view angle near the right eye position of the HMD. This camera detects landmarks in the physical space in order to compensate the error of the magnetic sensor. See Section 4.1 for the details of registration algorithm.

Each player holds a mallet as a physical device to hit a virtual puck. The mallets are simple devices having infrared LEDs. The position of the mallets are tracked on the image captured by an infrared CCD camera set directly above the table. In our AR^{2}Hockey, movement of the mallets is constrained on a two dimensional plane, but 3D tracking may be required depending on applications.

Computers

The AR^{2}Hockey system uses three SGI O2s for processing two video images from the cameras attached to the HMDs and a mallet-tracking process. In addition, one SGI ONYX2 (with eight CPUs and three InfiniteReality graphic pipelines) handles head-tracking, image and sound rendering, and the total system control. All the computers in the system proceed processing while communicating with each other over the Ethernet network.

By making only one super-graphic workstation to process all the rendering, images given to the participants can completely be synchronized and solve the synchronization problem. Currently, this system distributes other processes to three computers over the Ethernet, however, it is ideal to do all processes by only one graphics workstation since the time lag on the network communication may be a problem.

2.2 Process flow

Figure 2 shows the process flow of this system. Duplicated blocks in the figure represents blocks prepared separately for the two players. As shown in the figure, the process is composed of six types of sub-processes and one main process. Three sub-processes, landmark-tracking, registration and rendering are invoked for each player.

The four tracking processes that drive input devices work asynchronously. This means that they proceed independently from the main process and parallel to each other. On the other hand, the registration, the space management and the rendering processes are synchronous to the main process. By configuring the system in this way, it becomes possible to reduce the effect caused by the difference of the sampling rates of head-trackers, video capturing rate and the rendering rate. This effect directly influences the time lag of the system.

Head-tracking Process

The positions of the players' heads are measured by the magnetic sensors attached to the HMDs. The sampling rate is about 50 Hz depending on the characteristics of the Fastrak system. Note that two sensors for two HMDs are used in the system. The head-tracking process receives the head positions at this rate and sends the data to the registration process.

Landmark-tracking Process

An image from a player's viewpoint is taken from the CCD camera mounted on the HMD. From the image, the system extracts a position of a landmark on the table. The landmark position is sent to the registration process at 30 Hz. However, there is a latency of about 40 ms to capture and process an image, the data sent to the registration process delays at that interval.

Ten small square-shaped landmarks are placed on the table as shown in Fig.1(a). Each landmark has one of two colors, red or green, and the color specifies which players of the two the landmark is used for.

A landmark is extracted by a simple image processing. The process decides that a point is a landmark when there is a point which has an intensity of the predefined color over a certain threshold near the landmark position at the previous frame. If there is not such a point, the system scans the entire image and detects a point having the nearest color intensity.

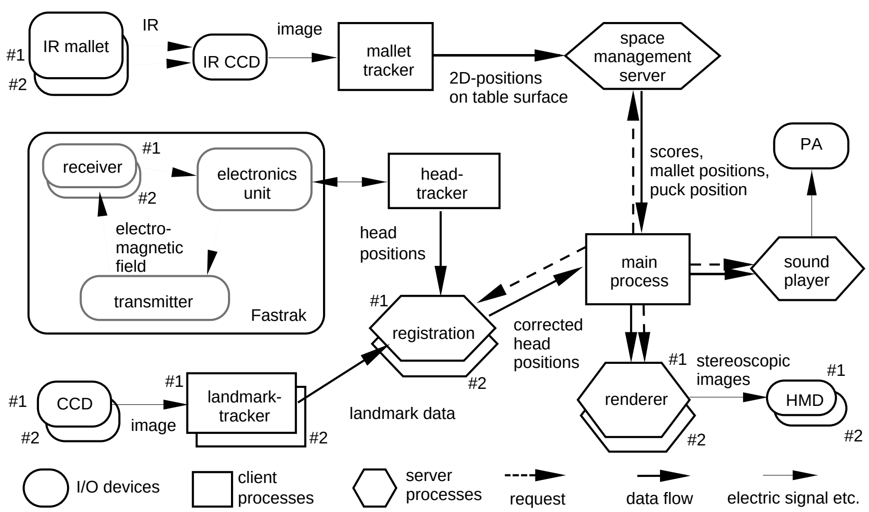


Figure 2: Process diagram of AR^{2}Hockey

Registration Process

This process makes registration in response to the request from the main process based on the stored head-tracking and landmark-tracking data. Then it sends corrected head position and orientation data to the main process.

As described, the latest data stored in this process does not correspond to the current status of the physical world. There is some time delay. Thus the registration process records the timestamps that indicate when the data is updated by the tracking processes. Based on the timestamps and data itself, this process first predicts the current head position, orientation, and landmark position by the second order liner prediction. The registration algorithm is applied to the predicted data. See the Section 4.1 for the registration algorithm.

Mallet-tracking Process

This process measures 2D positions of mallets on the table. It is implemented by a simple image processing as the landmark tracker and sends mallet position data to the space management process at 30 Hz with a latency of about 40 ms.

Space Management Process

This process manages the state of the game such as puck position, speed and a score and updates the game status by predicting recent mallet positions based on the stored mallet data in response to the request from the main process. This process also generates sound effects when it updates the game status.

Rendering Process

This process is synchronous to the main process and generates a set of stereoscopic images. The process obtains the required data from the main process and displays rendered images to the HMD.

Main Process

This is the process which coordinates, controls the other processes and creates the system highly cohesive. The process requests corrected head position and orientation data from the registration process, and the game state from the space management process at exact rate of 36 Hz. Then it sends these data to the rendering process for rendering.

3 Visual Simulation with AR

In the AR^{2}Hockey system, we have studied mainly on the static and dynamic registrations, that is, positional misalignment and time lag, while taking a game requiring quick motion as a subject of research. “MR Living Room” is another experimental AR system for the interior simulation. This is developed using the knowledge from the AR^{2}Hockey project while taking the technical problem related to the image quality consistency into consideration. This section describes the outline of this project.

The MR Living Room has a 2.8 m $\times$ 4.3 m floor of wooden flooring staff half-equipped with a few pieces of furniture and articles. In this space, two participants with see-through HMDs can experience virtual interior simulation such as selecting and placing furniture. Figure 3 shows the inside of the experiment space. Virtual furniture and articles are merged into this physical space and presented in real-time on the HMDs. Figure 5 shows a see-through image and an augmented image.

Figure 3: Space of MR Living Room

Figure 4: See-through HMD

(a) Scene in the living room (b) Augmented view (video see-thru)

Figure 5: Visual simulation in MR Living Room

HMDs

The HMD in this system utilizes two TFT LCDs of 920,000 pixels as the display unit and can present a set of stereo color images in the 640×480 VGA resolution. Though the optical system is basically same as the HMD used in the AR^{2}Hockey, this HMD has 51 degrees of horizontal view angle and 37 degrees of vertical one with a modified prism having a very little distortion. The transparency rate of the real image is set to a higher value to achieve more realistic see-through feeling.

The HMD has a color CCD camera and an optional visor, which shuts out the light from the physical space. When the visor is put on the HMD, the video see-through configuration is realized by displaying the image captured by the color camera. A head-tracker and an infrared camera are also mounted for the registration. Figure 4 shows this HMD.

Trackers

A hybrid tracker (InterSense's IS-600) is equipped in the MR Living Room. This tracker uses a gyro-sensor for the orientation measurement and a ultrasonic sensor for the position measurement. The gyro-sensor is suitable for this application where the participants walk around. However, it shows apparent drift in the heading direction (yaw) because of the measurement error accumulation.

In order to eliminate this kind of measurement error using image information, infrared LEDs are placed in the living room as the landmarks and observed by a small infrared CCD camera (ELMO's ME411R) mounted on the HMDs. Since the color information is quite important in the application such as the interior simulation, it is inappropriate to assign a certain color to the landmarks. Therefore, the infrared landmarks are used. The positions of these landmarks on the image are used in the registration.

Computers

This system consists of three graphic workstations (GWS) and two PCs. Two GWSs (SGI ONYX2 and O2) among them are used for image generation, one for each participant. Ideally, two GWSs of ONYX2 class are necessary to present a stereoscopic image to each participant. One O2 is used, however, for one participant because of the limitation of available equipment in our laboratory. Thus one participant whose HMD is driven by ONYX2 can see stereoscopic images with optical see-through, but the other can see monoscopic images with video see-through. A separate GWS (SGI O2) is used for the virtual interior object manipulation and operated by an operator. Two PCs each of which is equipped with one image processing hardware (HITACHI IP5005) are used for landmark-tracking for each participant.

Process Configuration

Figure 6 shows the block diagram of this system. The Virtual Interior Object Management Process is the server process which comprehensively manages all information related to the virtual interior objects placed in the virtual space. The operator adds, moves, or deletes the virtual interior objects through the Virtual Interior Object Manipulation process.

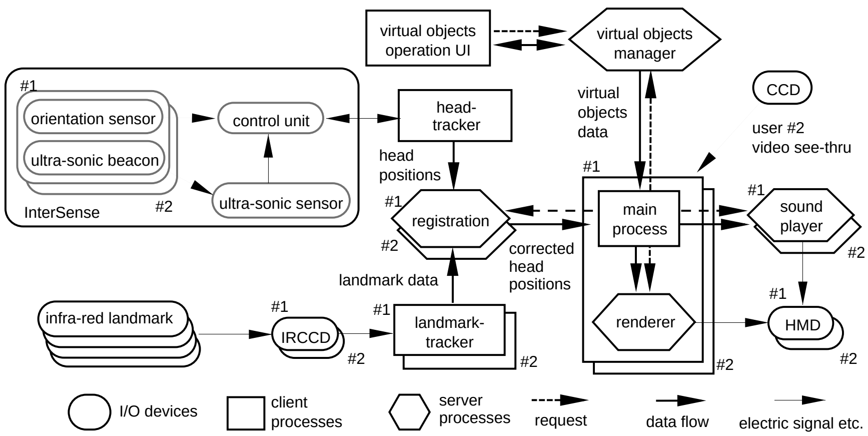


Figure 6: Process diagram of MR Living Room

4 Static Registration Algorithm

For the AR systems, the registration between a physical space and a cyberspace is essential. This section describes the registration algorithms implemented in the above systems. The discussion below assumes the perspective projection. All the inner camera parameters are already known and an image is captured by an ideal capturing system without any distortion.

4.1 Registration with one landmark

Correcting positional error using only one landmark is a simple, fast, and effective method for the registration. Thus we chose this method in the AR^{2}Hockey system and implemented a registration algorithm based on the method proposed by Bajura et al.[9] since the processing speed is crucial for this application.

Figure 7 shows the basic theory to correct positional error using one landmark. In the figure, let C, I, and Q be the camera position, the image plane, and the landmark position in the physical 3D space, respectively. For those C and I, the landmark Q will be projected and detected on the

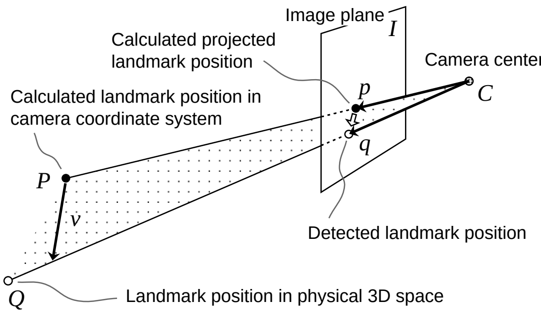


Figure 7: Registration with one landmark

image at the point q where the line connecting C and Q intersects the image plane I. On the other hand, the landmark position P in the camera coordinate system and its projected position p on the image are calculated based on the data from the 3D head-tracker. These P and p can be thought as the landmark in the virtual space and its corresponding image position. Ideally, point Q and P coincide in the 3D space. That is, the projected position q and p coincide on the image plane. This is, however, usually not true because of the head-tracker's error.

The correction is done by translating the virtual space coordinates so that the calculated projected position of the landmark p' on the image coincides with the point q. This is done by translating objects in the virtual space by

$$v = n(\overrightarrow{Cq} - \overrightarrow{Cp}) \tag{1}$$

where n is a scale factor derived by $n = |CP|/|Cp|$.

Since this method only registers the positions on 2D image plane, the three dimensional positions may not coincide correctly even after the registration. This method, however, is still effective if the sensor error is not so much. Moreover, the calculation cost is inexpensive and suitable to the real-time process required to the system such as the AR^{2}Hockey.

While the collaborative operation requires broader registration area, this registration is limited to only a small area. That is because this method requires a landmark always be observed in the captured images. Placing the multiple landmarks in the physical space will make it possible to broaden the registration area. In this case, how to identify the detected landmark becomes a new problem. Our method uses the information from the head-tracker as a guide to identify the landmark.

Suppose multiple landmarks $Q_i (i = 1 \cdots N)$ are placed in the physical space and their world coordinates are already known. Here N is the number of landmarks. For Q_i, let the point p_i denotes the calculated projected image coordinate of the landmark derived from the head-tracker. When a landmark is detected at an image coordinate of q, we can simply choose the landmark Q_k as the target landmark which minimizes the following evaluation values e_i:

$$e_i = |p_i q_i| \quad (i = 1 \cdots N) \tag{2}$$

After the landmark is decided in this way, the same method as described above can be used as it is.

4.2 Registration with depth information

The registration algorithm described in the previous section is very simple, fast and effective in the AR^{2}Hockey system. The method, however, is not appropriate for the MR Living Room because the visual simulation application requires higher registration accuracy than what the AR^{2}Hockey requires. In addition, the type of available information for registration will change while the participants walk around in the space. Thus, we have developed a framework which uniformly handles the various information available for the registration. The following describes this framework.

Assume that a landmark Q_i in the 3D space, whose known world coordinate is $Q_{Wi} = (X_{Wi}, Y_{Wi}, Z_{Wi}, 1)^T$, is projected onto the image coordinate $q_i = (x_i, y_i)$. This projection can expressed by the 3×4 matrix $\mathbf{C}$ as;

$$U_i = \begin{pmatrix} x_i h_i \\ y_i h_i \\ h_i \end{pmatrix} = \mathbf{C} \cdot Q_{Wi} \tag{3}$$

How to obtain this matrix $\mathbf{C}$ is the problem of estimating camera parameter, that is, also the problem of registration. Normally, it is necessary to detect more than 6 landmarks, which are not on the same plane, to get the matrix $\mathbf{C}$ by liner calculation. In our registration algorithm, even when only less than 6 landmarks are detected, it is possible to calculate the camera parameter matrix $\mathbf{C}$ by utilizing the output of the head-tracker.

(1) With four (or more) landmarks

It is known that the parameter h_i in the Eq.(3) is proportional to the depth value Z_{Ci} of the landmark Q_i in the camera coordinate system[10]. If we can measure the depth values for the detected landmarks, three equations are given for each single landmark by substituting h_i in the Eq.(3) with Z_{Ci}. If four or more landmarks, which are not on the same plane, are detected, the Eq.(3) can be expresses as

$$\mathbf{U} = \mathbf{C} \cdot \mathbf{W}. \tag{4}$$

where $\mathbf{U} = (U_1, U_2, U_3, U_4, \cdots)$ and $\mathbf{W} = (Q_{W1}, Q_{W2}, Q_{W3}, Q_{W4}, \cdots)$. Then the matrix $\mathbf{C}$ can be calculated from the following equation:

$$\mathbf{C} = \mathbf{U} \cdot \mathbf{W} \cdot (\mathbf{W} \cdot \mathbf{W}^T)^{-1}. \tag{5}$$

In case of the four landmarks, it is simply,

$$\mathbf{C} = \mathbf{U} \cdot \mathbf{W}^{-1}. \tag{6}$$

Since the matrix $\mathbf{W}^{-1}$ (or $\mathbf{W} \cdot (\mathbf{W} \cdot \mathbf{W}^T)^{-1}$) is obtained from known world coordinates of the landmarks, it can be calculated beforehand. Thus, the problem of estimating the camera parameter $\mathbf{C}$ becomes the problem how to obtain the matrix $\mathbf{U}$, that is, the image coordinates of four (or more) landmarks and their depth values Z_{Ci}.

Here, the main idea of our registration algorithm is to lead the depth information of detected landmarks from the rough information of the camera position and orientation obtained from the head-tracker. By using viewing matrix $\hat{\mathbf{M}}$ (expressed as 4×4 matrix to translate from the world coordinate system to the camera coordinate system) obtained from the head-tracker, the landmark $\hat{Q}_{Ci} = (\hat{X}_{Ci}, \hat{Y}_{Ci}, \hat{Z}_{Ci}, 1)^T$ in the camera coordinate system can be estimated. We use the element $\hat{Z}_{Ci}$ as the depth information of the landmark Q_i to get the matrix $\mathbf{U}$.

(2) With three landmarks

It is possible to obtain the camera parameter $\mathbf{C}$ by using only three landmarks Q_1, Q_2, Q_3. Let assume that all Z values of landmarks are zero in the world coordinate system and the coordinate is expressed as $Q'_{Wi} = (X_{Wi}, Y_{Wi}, 1)^T$. Then the projection of landmarks can be expressed by the 3×3 matrix $\mathbf{C}'$ as the following equation.

$$U_i = \mathbf{C}' \cdot Q'_{Wi} \tag{7}$$

Here matrix $\mathbf{C}'$ is the subset of the matrix $\mathbf{C}$ which omits its third column (factors related to the Z coordinate) and it is known that the matrix $\mathbf{C}$ can be derived from the matrix $\mathbf{C}'$[11].

If three landmarks are detected, the matrix $\mathbf{C}'$ can be obtained from the equation below.

$$\mathbf{C}' = \mathbf{U}' \cdot \mathbf{W}'^{-1} \tag{8}$$

where $\mathbf{U}' = (U_1, U_2, U_3)$ and $\mathbf{W}' = (Q'_{W1}, Q'_{W2}, Q'_{W3})$. Thus, the problem of estimating the camera parameter $\mathbf{C}$ becomes the problem how to obtain the matrix $\mathbf{U}'$, that is, the image coordinates of three landmarks and their depth values. As same as the case of using four landmarks, the depth values could be obtained from the head-tracker.

Note that the landmarks are not required to be placed on an actual $Z = 0$ plane because the translation matrix $\mathbf{N}$ (4×4) from the plane including the landmarks to the $Z = 0$ plane always exists.

This algorithm compensates the error of the camera parameter $\hat{\mathbf{C}}$ obtained from the head-tracker to eliminate the positional errors on the image for the detected three landmarks.

(3) With two landmarks

In case that two landmarks Q_1 and Q_2 are detected, the camera parameter $\mathbf{C}$ could be estimated in the same manner as stated above by setting an imaginary landmark Q_3.

Suppose that the third (imaginary) landmark Q_3 is at the world coordinate Q_{W3} not on the line $\overline{Q_{W1}Q_{W2}}$. The depth values of landmarks Q_1, Q_2, and Q_3 is obtained from the head-tracker. The image coordinate $\hat{q}_3 = (\hat{x}_3, \hat{y}_3)$ of the imaginary landmark Q_3 is also calculated by the output of the head-tracker. We can define the matrix $\mathbf{U}'$ by using these values.

The matrix $\mathbf{C}$ are compensated to eliminate positional errors for two detected landmarks from the camera parameter obtained from the head-tracker.

(4) With one landmark

It is possible to compensate the positional error on a landmark by setting two imaginary landmarks in the same manner as in the case of two landmarks. Note that, however, if only one landmark is detected, the compensation of the positional error on a landmark is accomplished by the method described in Section 4.1.

(5) Other cases

Many other methods besides the ways described above are applicable to this framework, since it is only necessary to get the depth information of the detected landmarks. For example, if a pair of stereo cameras are mounted on the HMD, the depth information is obtained by making correspondence between the landmarks detected on the stereo images. If the 3D head-tracker is available in addition to the stereo cameras, the camera parameter, that is, registration, is obtained with at least one fiducials in this framework.

5 Discussions and Future Studies

Two choices are available for the augmentation of physical world with virtual one: video-based and optical-based. A video see-through HMD works using a closed-view HMD and one or two cameras attached to the HMD. The video from the cameras is combined with computer generated images and displayed on the HMD. This configuration is often used in the applications where accurate registration is necessary. This is because 1) digitized video images are available for additional registration methods, 2) delay, brightness, and contrast between the two spaces can be easily matched in the video see-through[3].

On the other hand, an optical see-through HMD uses optical combiners so that users can see the physical world through glasses and simultaneously look at an image displayed on the HMD monitor. Since the physical world is seen directly, there is no time delay to see it. In addition, the resolving power of the physical space is only limited by the resolution of human fovea. This type, however, can not avoid the time lag between the physical world and the virtual one.

In collaborative AR, interactions are not limited to physical to virtual and vice versa. Physical to physical interaction between participants is also important. In that case, scenes of physical space for all the participants should be synchronized and the time delays should be minimized in order to match the kinesthetic and visual systems. Thus we believe that the optical see-through HMD is more suitable for the AR^{2}Hockey than the video see-through one.

In order to provide seamless feeling in the visual simulation process of the MR Living Room, the difference of image qualities between the physical space and the virtual space should be minimized. For that purpose, the video see-through would be better than the optical see-through. Thus we have adopted a system configuration in which both see-through methods can be used at the same time in order to examine optimal displaying and registration methods.

Many other problems should be solved to make AR systems generic and practical. As to the system configuration of AR, as you see in Fig.2 and Fig.6, there are common processes in the two AR systems. They are the head-tracking process, the landmark-tracking process, the registration process, and the rendering process. They are core for AR system and independent of the application. The other processes, including the main process, should be rebuilt according to the application requirement. This is the merit of the distributed AR system configuration we chose to the case studies. We have to confirm that this core system actually works through the many case studies.

The registration is another problem. Sensor fusion with physical 3-D sensors and tracking landmarks on the captured images is an optimal solution for the smart registration for the time. Our registration framework described in Section 4.2 can handles the various situation uniformly. More robust algorithms, however, are necessary for the practical applications where the registration area is relatively wide. In addition, prediction of the head movement might be effective to decrease the time lag. Smart researchers in computer vision and sensory technology area could be aware of that AR is a treasure house of problems that they can make the most use of their knowledge.

The study on human factor of AR system is another important subject.

The new version of the AR^2Hockey system, which has slightly different system configurations, was installed at ACM SIGGRAPH'98. More than 2,000 people played during the conference and most participants reported that they played the game in much the same way as the actual game. During the installation, we changed the augmentation methods and collect the survey from the participants as well as the scores and playing times. The results of analysis will be reported at another opportunity.

Acknowledgments

The authors would like to thank the people of Mixed Reality Systems Laboratory Inc. for their cooperative works and useful discussions.

References

[1] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual display," *IEICE Trans. Inf. & Sys.*, Vol.E77-D, No.12, pp.1321-1329, 1994.

[2] T. Ohshima et al., "AR^2Hockey: a case study of collaborative augmented reality," *Proc. VRAIS '98*, pp.268-275, 1998.

[3] R. Azuma, "A survey of augmented reality," *Presence*, Vol.6, No.4, pp.355-385, 1997.

[4] S. Feiner et al., "Knowledge-based augmented reality," *Comm. ACM*, Vol.36, No.7, pp.52-62, 1993.

[5] T. P. Caudell and D. W. Mizell, "Augmented reality, an application of heads-up display technology to manual manufacturing processes," *Proc. the Hawaii Int. Conf. on System Sciences*, pp.659-669, 1992.

[6] M. Billinghurst et al., "Shared space: collaborative information spaces," *Proc. HCI International '97*, pp.7-10, 1997.

[7] J. Rekimoto, "TransVision: a hand-held augmented reality system for collaborative design," *Proc. VSMM '96*, pp.85-90, 1996.

[8] H. Hoshi et al., "Off-axial HMD optical system consisting of aspherical surfaces without rotational symmetry," *SPIE*, Vol.2653, pp.234-242, 1996.

[9] M. Bajura and U. Neumann, "Dynamic registration correction in video-based augmented reality systems," *Computer Graphics and Applications*, No.15, Vol.5, pp.52-60, 1995.

[10] J. P. Mellor, "Realtime camera calibration for enhanced reality visualization," *Proc. CVRMed '95*, pp.471-475, 1995.

[11] Y. Nakazawa et al., "Simple digital mixing of real and CG moving image sequences for post-production of augmented reality," *Proc. ICIP '97*, Vol.2, 1997.