An Occlusion-Capable Optical See-through Head Mount Display for Supporting Co-located Collaboration

[*]Kiyoshi Kiyokawa, [**]Mark Billinghurst, [+]Bruce Campbell, [**]Eric Woods

[*]*Cybermedia Center, Osaka University, Japan*
[**]*Human Interface Technology Laboratory New Zealand, New Zealand*
[+]*Human Interface Technology Laboratory, University of Washington, USA*

kiyo@cmc.osaka-u.ac.jp, mark.billinghurst@hitlabnz.org,
bdc@hitl.washington.edu, eric.woods@hitlabnz.org

Abstract

An ideal augmented reality (AR) display for multi-user co-located collaboration should have following three features: 1) Any virtual object should be able to be shown at any arbitrary position, e.g. a user can see a virtual object in front of other users' faces. 2) Correct occlusion of virtual and real objects should be supported. 3) The real world should be naturally and clearly visible, which is important for face-to-face conversation. We have been developing an optical see-through display, ELMO (Enhanced see-through display using an LCD panel for Mutual Occlusion), that satisfies these three requirements. While previous prototype systems were not practical due to their size and weight, we have come up with an improved optics design which has reduced size and is lightweight enough to wear. In this paper, the characteristics of typical multi-user three-dimensional displays are summarized and the design details of the latest optics are then described. Finally, a collaborative AR application employing the new display and its user experience are explained.

1. Introduction

Our goal is to build an augmented reality (AR) display that is suitable for multi-user co-located collaboration. We believe that such a display should satisfy the following three requirements.

- Users can see non-digitized real collaborators as well as rest of the real world. This provides superior awareness of collaborators and allows users to share non-verbal communication cues naturally. In a face-to-face collaboration, perception of the partner at the same time as the task entities ensures smooth communication, yet few current AR systems take users' mutual visibility into account.
- Virtual objects can be shown anywhere in the working space. For example, virtual objects can be floating in the air, they can be pointed at or touched, and they can be shown on any physical background. These features expand the number of possible application domains, and placing a shared object between users' faces encourages user interaction [10].
- Virtual objects can be transparent or opaque, as demanded by the task. Occlusion is one of the strongest depth cues. In order for a virtual object in a real environment to be convincing, it should cover and/or be covered by any real objects in the scene properly in accordance with their spatial relationship.

Several 3-D collaborative environments have recently been developed. There are a number of different display possibilities for the three-dimensional displays used by co-located users. Each approach has advantages and disadvantages as shown in Table 1.

Table 1. Characteristics of typical display systems.

	Eye	Real scene	Location of virtual objects	Opaque virtual objects on a real scene	Practical number of users
Volumetric Display	Naked	Real	Display's position	Difficult	Unlimited
Projection-based Display	Stereo glasses	Real	Limited area	Impossible	2 to 4 users
Video see-through Display	HMD	Video image	Unlimited	Easy	Unlimited
Conventional Optical see-through Display	HMD	Real	Unlimited	Impossible	Unlimited
Occlusion-capable Optical see-through Display	**HMD**	**Real**	**Unlimited**	**Easy**	**Unlimited**

Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)
0-7695-2006-5/03 $17.00 © 2003 **IEEE**

One straightforward way to have a three-dimensional virtual image, which is observable by multiple users from different viewpoints, is to make a physical three-dimensional display, or a volumetric display. Mechanical devices are often used to create true volumetric displays [1][13]. However, with this approach, virtual objects can only be shown within the volume of the display body. Besides, an opaque image is hard to show as a volumetric display when making use of an afterimage effect.

A second approach is by using a time- or space-divided projection interface. The traditional stereo view application using a pair of stereo glasses and a projection screen was originally intended to be viewed by a single observer. Some recent systems, however, do support independent viewpoints for more than two users at a time [2][6]. However, a projection-based approach inevitably needs a physical screen as a background on which to view a virtual scene, and cannot show virtual images in front of any other physical objects. This means users cannot see a virtual object on their hand, for example.

In contrast, AR interfaces using head mounted displays (HMDs) allow virtual objects to be shown at arbitrary locations, such as in front of a partner's face [7][11][12]. Another advantage of HMD-based interfaces is that they have the potential capability of showing correct occlusion phenomena between virtual and real scenes. That is, with a video see-through head mount display and proper depth information of a real scene, virtual objects can cover further real objects, and be covered by closer real objects [4][5]. However, the video see-through approach degrades the quality of the real scene dramatically and inevitably introduces a system delay. Users often feel the captured real image is something like a television or video game image and the sense of presence is severely damaged. On the other hand, an optical see-through display keeps the unpixelated, intrinsic quality of the real scene. However, in optical see-through displays, virtual images often appear as semi-transparent ghost images due to light loss in the optical combiner.

We have been developing an optical see-through display that is capable of showing opaque virtual objects [8]. Our design introduces a secondary LCD panel to selectively block the real scene on a pixel basis. One advantage is that the virtual image can be kept in focus regardless of the distance in the real scene the user is looking. This is because the secondary LCD panel is placed between two convex lenses of the same focal length. Although Sony Corp. also has already proposed this idea [3] and Tatham proposed a similar light-blocking mechanism [14], no functioning system has been built other than ours.

Our previous prototype systems were not practical due to their size and weight. We have come up with improved optics that make it possible to make the display light and small enough to wear without any viewpoint offsets.

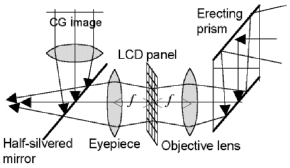

Figure 1. ELMO-1.

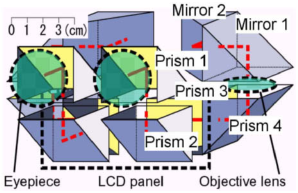

Figure 2. ELMO-2.

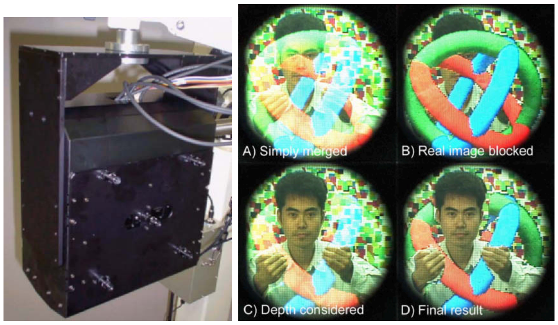

Figure 3. ELMO-3 (left) and its views (right).

2. Previous displays and their problems

Our first prototype, ELMO-1 (Figure 1), was about 30cm in depth and the virtual viewpoint was shifted from that of a naked eye largely to both vertical and horizontal direction offsets.

The second prototype, ELMO-2 (Figure 2), was about 15cm in depth and still had a large (21cm) viewpoint offset in the Z-direction. Finally, the third prototype, ELMO-3, was combined with a five-camera real-time stereovision system (Figure 3). Four images captured through ELMO-3 are shown in Figure 3. Figure 3(A) is a typical 'ghost' image without occlusion. Figure 3(B) shows opaque virtual images without depth information. In this case, all virtual objects occlude the real scene. Figure 3(C) shows the scene with semi-transparent virtual objects maintaining depth information of the real scene. Figure 3(D) shows opaque virtual objects with correct occlusion attributes.

There were two main issues in these prototypes that remained unsatisfactory, the *optical layout* and the *masking LCD panel*. In the prototypes, viewpoint offsets and dimensions remained large. For example, ELMO-3 had a longitudinal viewpoint offset of 21cm which meant the entire real world looked 21cm closer to the user. It also weighed as heavy as 15kg because of the steel frame and heavy glass materials. Another issue was the secondary masking LCD panel for blocking the real scene. We used a normal TN type VGA LCD panel of 10.4 inches with a response time of 300ms. Since our approach depends on pixel density, we could use only about 120x120 pixels for masking, which was unacceptably low. We definitely needed to find new LCD panels of higher resolution and faster response.

3. Designing ELMO-4 optics

An optical see-through display should keep a view of as much of the real scene as possible. When a large viewpoint offset exists, users have some difficulty in performing a hand-eye coordination task and seeing partners in a collaborative task. So, eliminating viewpoint offsets is necessary for designing the ELMO-4 optics, as it is especially intended for co-located multi-user collaboration.

3.1. Design for 1X magnification

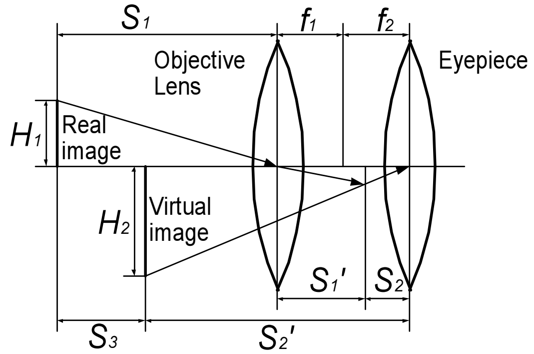

Figure 4. Relationships among an objective lens and an eyepiece.

To design new optics that are free from viewpoint offsets, we first examine a condition of focal lengths of convex lenses so that both the lateral (transversal) and longitudinal magnifications should be one. To simplify, we use a thin lens model, but a thick lens model leads to the same conclusion. Our approach uses at least two convex lenses and a masking LCD panel which is placed at an image plane between the lenses. In Figure 4, a lateral magnification M is represented as,

$$M = H_2/H_1 = (S_1'/S_1)(S_2'/S_2) = 1 \quad \ldots(1)$$

From a common Gaussian form of the lens equation,

$$1/f_1 = 1/S_1 + 1/S_1' \quad \ldots(2)$$

$$1/f_2 = 1/S_2 - 1/S_2' \quad \ldots(3)$$

where f_1 and f_2 are the focal lengths of the objective lens and the eyepiece, respectively. Figure 4 also shows a relationship,

$$f_1 + f_2 = S_1' + S_2 \quad \ldots(4)$$

Then we represent S_1, S_1', S_2' using S_2.

$$S_1' = f_1 + f_2 - S_2 \quad \ldots(4')$$

$$S_1 = (f_1 (f_1 + f_2 - S_2))/(f_2 - S_2) \quad \ldots(2')$$

$$S_2' = (S_2 f_2)/(f_2 - S_2) \quad \ldots(3')$$

Finally the equation (1) is described as,

$$(f_1 + f_2 - S_2) \text{ x } (f_2 - S_2)/(f_1 (f_1 + f_2 - S_2)) \text{ x } (S_2 f_2)/(f_2 - S_2) \text{ x } (1/S_2) = f_2 / f_1 = 1 \quad \ldots(1')$$

Therefore, f_1 and f_2 should be identical for a lateral magnification to be one.

To examine a condition for a longitudinal magnification, we calculate S_3. S_3.is represented as,

$$\begin{aligned} S_3 &= (S_1 + S_1' + S_2) - S_2' \\ &= (f_1(f_1+f_2-S_2))/(f_2-S_2) + (f_1+f_2-S_2) + S_2 - (S_2f_2)/(f_2-S_2) \\ &= (f_1+f_2)(f_1+f_2-2S_2)/(f_2-S_2) \\ &= 2(f_1+f_2) + (f_1+f_2)(f_1-f_2)/(f_2-S_2) \quad \ldots(5) \end{aligned}$$

When S_3.is constant (independent of S_x), the longitudinal magnification is one. From (5), S_3 becomes constant as,

$$S_3 = 4f_1 \text{ when } f_1 = f_2 \quad \ldots(5')$$

In general, the entire optics should be symmetrical along the optical axis, with any even number of lenses, for 1X magnification. For example, inner foci $f_1 = f_2 = f_{in}$ and outer foci f_{out} can be different as shown in figure 5.

3.2. Design for offset free view

The optics shown in Figure 5 introduce a viewpoint offset $2(f_{out} + f_{in})$ along the optical axis. To eliminate this offset, the virtual viewpoint should be located at the real viewpoint.

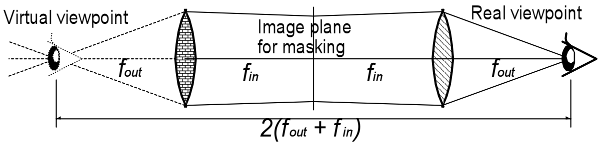

Figure 5. Real and virtual viewpoints.

Two example layouts of the objective lens and the eyepiece are shown in Figure 6. In this figure, the two inner focal points a and b should also be at a same point. This means the optical path should be folded in order to connect a and b. As shown in the figure, at least two reflectors are needed, one for the entrance and the other for the exit. The larger the gap h between the two reflectors becomes, the larger the entrance reflector is needed to keep the field of view. To minimize the entrance reflector and to maximize the field of view, we use a two-sided reflector to make h zero.

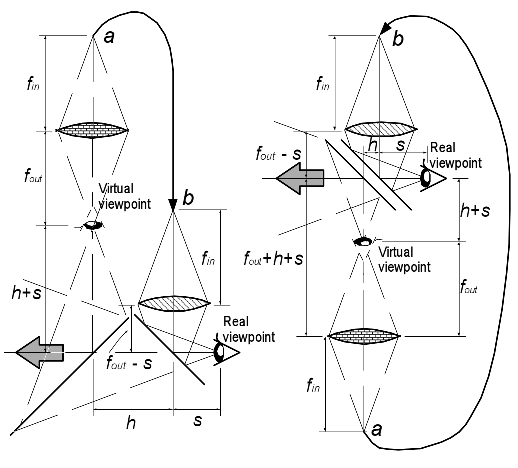

Figure 6. Example layouts.

3.3. Design for erecting image

Figure 7 shows a naïve example of a folded optical layout. A masking LCD panel should be placed at the midpoint C of an optical path AD. A color display is placed at B which is conjugate to C (AB = AC = CD = f_{in}). This layout, however, inverts the real view. There are a number of optical ways to erect an image. Several types of prisms such as Abbe-Konig and Pechan prisms provide full image reversal without shifting the optical axis. However, an erecting prism generally introduces a long optical path. This is disadvantageous because a conjugate part AB for the color display becomes unnecessarily bulky.

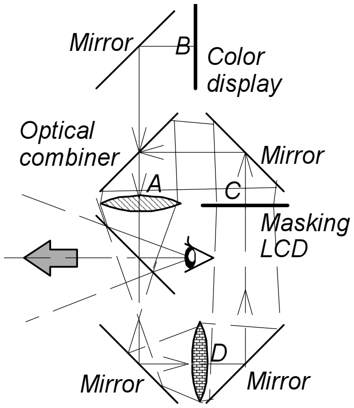

Figure 7. A naïve loop layout.

We then introduced another pair of lenses for erecting the image. This not only simplifies the optical structure for a color display, but also avoids introducing the heavy weight of a prism. Figure 8 shows a simplified illustration of our final ELMO-4 design. Note that there are two virtual viewpoints *V1* and *V2*, and two candidate locations for both a color display (*C1*, *C2*) and a masking LCD (*M1*, *M2*). The locations *C1* and *M1* are selected in our design, as they are close to each other and to their counterparts of the other eye. This selection is instrumental in making the circuit box small.

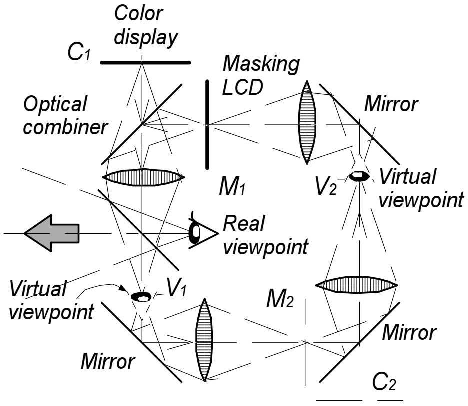

Figure 8. A simplified ELMO-4 optics.

4. Technical innovations of ELMO-4

4.1. Optics design

Figure 9 shows a front view of the ELMO-4 optics, which is composed of twin ring-shaped optical loops. Each loop consists of two pairs of convex lenses, one for erecting the image, and the other for blocking the real view. The inter-pupil distance is adjustable between 5.5cm and 7cm. The optics specification is summarized in Table 2.

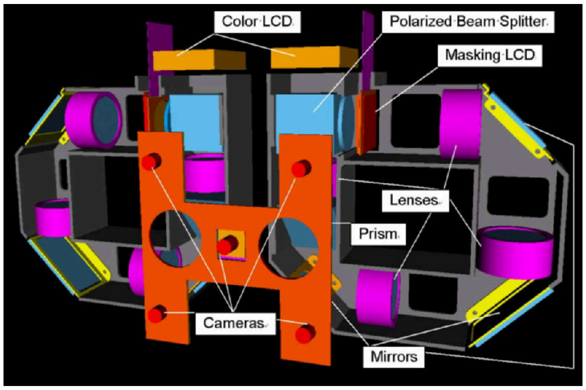

Figure 9. ELMO-4 optics design of.

For the right ring in Figure 9, outside light rays first enter into the loop through the circular hole of the camera base, then are bounced off the first mirror (backside of the eyepiece prism) downwards. These light rays are then erected through the second and third mirrors and two convex lenses, and are reflected on the fourth mirror at the upper right corner. They are then selectively blocked on a pixel basis by the masking LCD, combined with a color CG image by the beam splitter, then finally bounced off the eyepiece prism to the user's left eye. Significant advantages in the ELMO-4 optics are summarized as follows:

Offset free: Viewpoint offsets are completely eliminated. The horizontal field of view is kept at 30 degrees. This means that a real scene viewed through the optics is identical to a normal view of the real scene without the optics. This was accomplished not only by the unique ring-shaped design but also by using prisms that have a high refractive index of 1.83.

Thin body: The ring-shaped design also contributes to reducing the thickness of the optics. The new optics is as thin as the lens diameter (30mm) in the Z-direction, which is thinner than one fourth of that of ELMO-3. As the center of gravity is shifted to the user's head by 8cm, the moment of head rotation is also greatly reduced.

Lightweight: The entire optical components (lenses, prisms, and mirrors) for each ring weigh 200g (about 1/10 of that of ELMO-3). The total weight including the optics, aluminum frame, cameras and LCD modules is about 1.1kg. We also needed to attach a counter balance weight at the back of the head to create a stable head mount. Thus the total weight including a headpiece and the counter balance is about 2.0kg (about 1/3 of that of ELMO-3). Although we could have used lighter materials, such as plastic for prisms and frames, ELMO-4 is much lighter than previous prototypes. This was mainly accomplished by reducing the number of prisms from six to two, and by reducing the diameter of lenses from 38mm to 30mm.

Peripheral vision: ELMO-2 and ELMO-3 did not support peripheral vision. The ring-shaped design makes it possible for the observer to see the real scene directly through the holes at the center of the rings. This means peripheral vision is partially available.

Bright view: The ELMO approach inherently attenuates the light intensity of a real scene due to the transparencies of the beam splitters and the masking LCD panel. For example, the composite transparency of ELMO-3 was only about 10%. While previous displays used half prisms as a beam splitter, ELMO-4 introduced polarized beam splitters, so as to maximize the brightness of the real scene. As a result, composite transparency of ELMO-4 is improved to about 22%.

4.2. LCD module

For the LCD panels for both masking and color CG images, we chose a 1.5 inch QVGA (320x240) LCD module of ultra high resolution and ultra fast response from Hunet Co. in Japan. Advantages over LCD modules used in previous displays include:

Quick Response: Hunet's sequential color LCD module has a response time of as short as 2ms. By using separate RGB backlights, each single pixel can produce full color at 60 frames per second. Note that the masking LCD we used in the previous displays had a response time of as slow as 300ms.

Table 2. Specification of ELMO-4 optics.

Convex lens	
Focal length	42 [mm]
Effective aperture	30 [mm]
Center thickness	15.2 [mm]
Weight	28 [g]
Transparency	98 [%] (lambda = 550nm)
Prism	
Window shape	30 x 30 [mm]
Reflectance index	1.834
Weight	60 [g]
Transparency	98 [%] (lambda = 550nm)
Effective Field of View	> 30 [deg]

High Resolution: The time-sequential color system does not require RGB sub-pixels, so the pixel resolution can be easily increased. The module we chose has a dot pitch of 0.126mm, and is as dense as 260 pixels per inch.

High Contrast: Previous masking LCD modules had a contrast ratio of 10 to 1. Due to low contrast, the masking LCD was unable to block the incoming light completely. On the other hand, the new module's contrast ratio is 100 to 1. Now, the display can show opaque objects in desired color much more clearly.

Pixel Quality: Previous displays used different types of LCD modules for masking and color images. The new display uses identical LCD modules which are accurately calibrated at a pixel level. As a result, image quality is greatly improved.

4.3. Stereo cameras

The ELMO-4 uses Komatsu's FZ-930 real-time stereovision system with five small cameras. This vision system is capable of acquiring a 280x240 pixel depth map of a real scene in real-time. This is necessary for correct mutual occlusion representation. The latency of the depth map calculation is 33ms. We use small and lightweight 1/3 inch CMOS color cameras (Netcam Vision, CB-1030EMN) for the head mount. Focusing on a desktop multi-user AR collaboration, the detectable depth range is set to 30cm to 160cm with a baseline of 40mm. This configuration should be enough for viewing from the observer's own hand at one's reach, to a partner's body sitting at the opposite side of a table. Figure 10 shows an actual image of ELMO-4 with the stereo cameras in use.

As the CMOS camera does not take a sync signal, we needed a set of frame synchronizers (IMAGENICS, FS-5000). Note that even though we use frame synchronizers in order to be accepted by FZ-930 hardware, those five cameras still capture images at slightly different times, causing depth error. The depth map error is easily noticeable, especially when the observer moves his or her head quickly. However, the depth sensing system works fairly well in many situations.

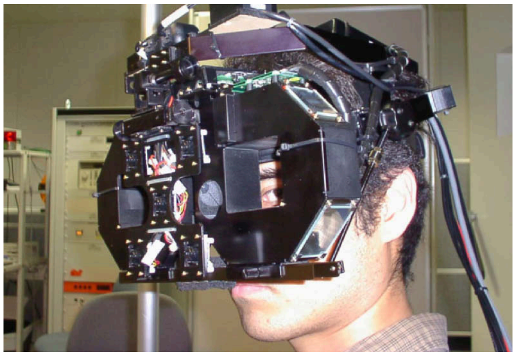

Figure 10. Close-up of ELMO-4 in use.

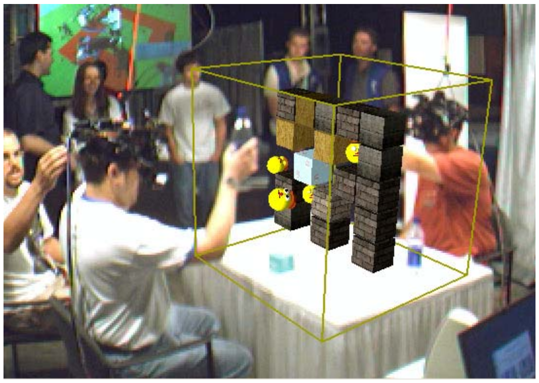

Figure 11. Collaborative AR Game.

5. An application: Collaborative AR Game

5.1. Features of the collaborative environment

We built two sets of ELMO-4 displays and implemented a simple collaborative AR game as an example application. Figure 11 shows a composite snapshot of the environment, which was exhibited at SIGGRAPH 2002 [9]. Players can enjoy a kind of 'breakout' game (explained in detail in the next section). The objective of the implementation was to demonstrate the unique capability of the ELMO display in a collaborative setup, and to assess the feasibility of the display through significant user feedback.

As discussed in Section 1, a collaborative ELMO setup will provide a unique multi-user AR environment, which has following three key features:

First, users can see virtual objects flying around them. Placing a virtual object at an arbitrary position in a multi-user setup is only possible in a head mounted display.

Second, users will see their hands cover and be covered by a virtual object properly, according to their spatial relationship.

Finally, users will notice the real world is naturally and clearly visible as it is supposed to be. This will help them feel co-existent in the real world not in a digitized video world.

5.2. Game design

Two players experience the game at a time. They sit at a table facing each other, each wearing an ELMO-4 as shown in Figure 11. Each ELMO-4 hangs from the ceiling by a rubber band to compensate for the heavy weight. The inter-pupil distance is fixed to 65mm and head motions are tracked by 3-D sensors (InterSense IS-600Mk2).

When users run the application they first see a virtual wall composed of 5 x 5 textured bricks and two yellow balls with faces wearing cowboy hats. Their goal is to remove all the bricks by hand swatting or yellow ball collisions. Since the HMD senses real-time depth information of the real scene, the system can roughly detect collisions between virtual and real objects. This enables players to push and rotate virtual objects with their bare hands. Similarly, virtual yellow balls are bounced off the real objects, including players' hands and the table surface (transparent virtual walls are used to fasten the balls within a certain volume on the table). When a real object or a yellow ball hits a virtual brick, it highlights and makes a sound. A virtual brick is destroyed with an explosive sound when it is hit three times. When all the bricks are gone, a new set of bricks with different textures appear. Each player can either interfere with or help out the partner by changing the moving direction of yellow balls, for example.

The virtual scene is first rendered in a traditional 'ghost' optical see-through mode for fifteen seconds, followed by the occlusion-capable optical see-through mode for fifteen seconds. This endless half-a-minute loop and the intuitive interaction design help players understand the advantages of our HMD. Figure 12 (left)

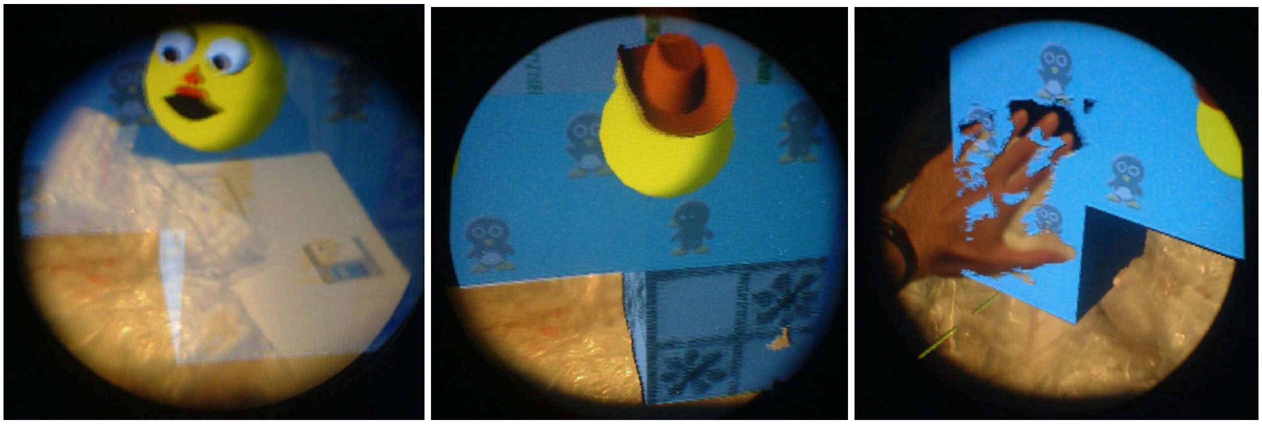

Figure 12. Images seen through ELMO-4 in the collaborative AR game.
Transparent objects (left), opaque objects (center), and a bare hand interaction (right) are shown.

shows a few virtual bricks and a yellow ball in a transparent mode. Note that the table and other objects such as a floppy disk are visible through the virtual bricks. Figure 12 (center) shows the same objects in an opaque mode. Now the penguin and flower textures on the bricks are clearly seen and no real objects are seen through the bricks. Figure 12 (right) shows an example of bare hand interaction. The observer's own left hand is covering further virtual objects. Note that some pixels of the bricks incorrectly remain or fade due to an inaccurate depth map.

5.3. Observations

Over 2,000 people experienced the game during the five-day exhibition. Most people enjoyed the interaction very much. They were eager to continue the game and reluctant to pass the HMD to the next pair of challengers. They shook and waved their arms frantically and laughed as they played the game. However, some people did not enjoy the environment as much. This was partly because of the bulky and uncomfortable HMDs, and the misaligned inter-pupil distance (users said things like "I can see the image only through my right eye!"). Due to the minimum distance of detectable depth range (30cm), they needed to stretch out their arms to the objects, which seemed awkward for some people. In such cases, some players held a real object to help them reach further. Figure 13 shows a child holding and shaking an empty water bottle.

Figure 13. A child playing the game.

5.4. User feedback

After the game players were asked to answer questions about the experience. A total of 615 players filled out a simple questionnaire that asked the following seven questions with answers on a scale of 1 (disagree) to 7 (agree).

Q1. The HMD was comfortable.
Q2. I could see the real world clearly.
Q3. I could see the virtual models clearly.
Q4. The virtual models were more real with the occlusion feature
Q5. I could tell more easily where the virtual models were with the occlusion feature.
Q6. I could easily touch the virtual models.
Q7. The game was fun.

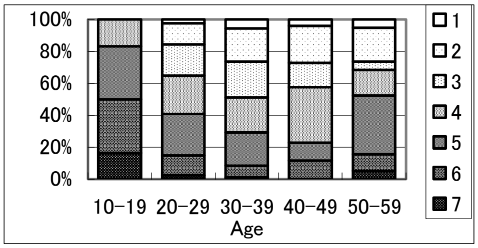

Figure 14. HMD comfort (Q1).

Players ranged in age from 14 to 58 years old with an average age of 30.8 years. 19.8% were female players. Figure 14 shows the results of Q1. The average score was 3.8/7.0. The ratios of positive (5, 6, 7), neutral (4) and negative (1, 2, 3) answers dramatically changed according to age. More than 80% of teenage players felt ELMO-4 was comfortable, while nearly half of players in their thirties felt it uncomfortable. This might be because the headpiece was too small for middle-aged adults.

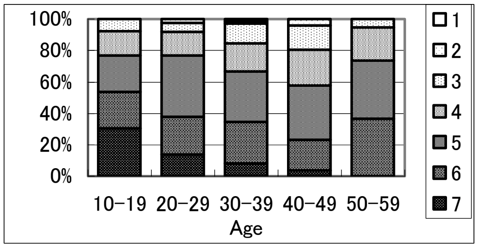

Figure 15. Real world visibility (Q2).

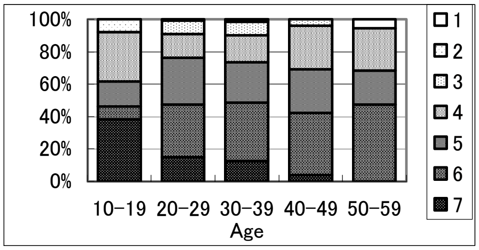

Figure 16. Virtual world visibility (Q3).

Figures 15 and 16 show the results of Q2 and Q3, respectively. The average scores for Q2 and Q3 were 4.9/7.0 and 5.2/7.0. About 70% of players could see the real and virtual worlds clearly. 11% and 7% of them were not able to get clear views, probably because of a misaligned inter-pupil distance. If we had adjusted the inter-pupil distance for each player, these scores should have been much higher.

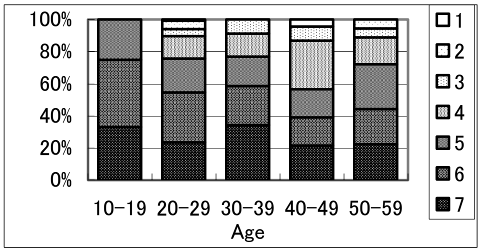

Figure 17. Occlusion effects for perceived reality (Q4).

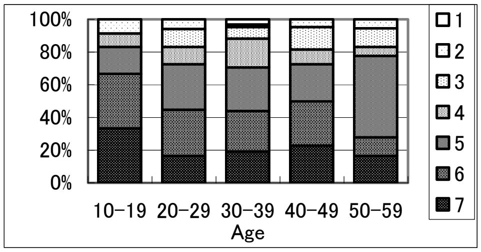

Figure 18. Occlusion effects for spatial recognition (Q5).

Figures 17 and 18 show the results of Q4 and Q5, respectively. Average scores for Q4 and Q5 were 5.3/7.0 and 5.1/7.0. On average, more than 75% players felt the virtual objects were more real and they could tell their locations more easily when the occlusion feature was provided. More than 82% of the players who gave positive answers for Q2 and Q3 also gave positive answers for Q4 and Q5. This suggests that if a player has clear views of real and virtual scenes, then he or she is more likely to appreciate the effectiveness of the occlusion feature.

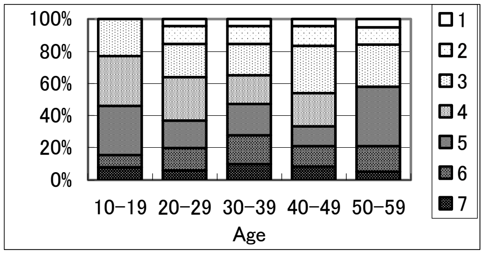

Figure 19. Ease of touch (Q6).

Figure 19 shows the result of Q6. The average score was 4.0/7.0. Less than half (44%) of the players felt it was easy to touch the virtual objects and many felt it difficult or even impossible. This is because of the long minimum length of the detectable distance of the rangefinder, and the inaccurate depth map. This should be improved by using synchronized cameras with a wider field of view.

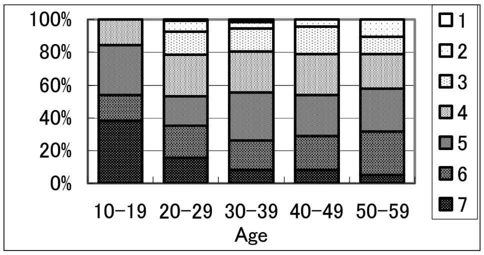

Figure 20. Game amusement (Q7).

Figure 20 shows the results of Q7. The average score was 4.6/7.0. In general, the game amused players at a 'so-so' level. It is interesting that teenagers enjoyed the game much more than other ages (5.8/7.0 in average).

Many players commented that it was fun to play a multi-user game. Most pairs who had known the partners in advance talked to each other very often, while those pairs who had met for the first time talked far less often.

6. Conclusions

In this paper we describe a newly developed optical see-through head mounted display ELMO-4, which has light and small optics without any viewpoint offset. The ELMO-4 is suitable for multi-user co-located AR collaborative systems because of three key features. *First,* since it is a head mount display, any virtual object can be shown at any arbitrary position, including in front of other users' faces. *Second,* either opaque or transparent objects can be shown as demanded by requirements of collaboration. *Finally,* users can see their partner clearly and naturally in an optical way.

A collaborative AR game was implemented using two sets of ELMO-4 headsets and experienced by over 2,000 players. Observations and user feedback showed that the occlusion feature enhanced a sense of presence of virtual objects for more than 75% of people, even when the inter-pupil distance was fixed to the standard 65mm. Further analysis indicated that almost all people would appreciate the occlusion feature, if the optics were adjusted to each person. This simple application demonstrated that our HMD is applicable to collaborative AR systems, and that collaborative AR systems are attractive for many people.

Future studies include conducting rigorous experiments on the effect of occlusion feature and developing practical collaborative AR systems. We will also research improved methods for depth sensing and interaction.

Acknowledgements

We would like to thank Hiroyuki Ohno of Communications Research Laboratory who gave us the opportunity for the study. We also would like to thank Topcon Co., Hunet Inc. and Komatsu Co. for their thorough technical support.

References

[1] Actuality System's Website: http://www.actuality-systems.com/, 2001.

[2] Bimber, O., Frohlich, B., Schmalstieg, D., and Encarnacao, L.M., "The Virtual Showcase," IEEE Computer Graphics & Applications, Vol.21, No.6, pp.48-55, 2001.

[3] Japanese Unexamined Patent Publication 1992-1717, 1992.

[4] Kanbara, M., Okuma, T., Takemura, H., and Yokoya, N., "A Stereoscopic Video See-through Augmented Reality System Based on Real-time Vision-based Registration," Proc. of the IEEE VR 2000, pp.255-262, 2000.

[5] Kato, H., and Billinghurst, M., "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System," Proc. of the IEEE & ACM IWAR '99, pp.85-94, 1999.

[6] Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F., "Interactive Stereoscopic Display for Three or More Users," In ACM SIGGRAPH 2001, pp.231-239, 2001.

[7] Kiyokawa, K, Takemura, H., Yokoya, N., "SeamlessDesign for 3D Object Creation," IEEE MultiMedia, Vol.7, No.1, pp.22-33, 2000.

[8] Kiyokawa, K., Kurata, Y., and Ohno, H., "An Optical See-through Display for Mutual Occlusion with a Real-time Stereo Vision System," Elsevier Computer & Graphics, Special Issue on "Mixed Realities - Beyond Conventions," Vol.25, No.5, pp.2765-779, 2001.

[9] Kiyokawa, K., Kurata, Y., and Ohno, H., "Occlusive Optical See-through Displays in a Collaborative Setup," Proceedings of the ACM SIGGRAPH 2002, Conference Abstracts and Applications (Sketch) , San Antonio, p.74, 2002.

[10] Kiyokawa, K., Billinghurst, M., Hayes, S. E., Gupta, A., Sannohe, Y., and Kato, H., "Communications Behaviors of Co-located Users in Collaborative AR Interfaces," Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality 2002 (ISMAR 2002), pp.139-148, 2002.

[11] Ohshima, T., Satoh, K., Yamamoto, H., and Tamura, H., "AR2 Hockey System: A Collaborative Mixed Reality System," Trans. VRSJ, Vol.3, No.2, pp.55-60, 1998.

[12] Schmalsteig, D., Fuhrmann, A., Szalavari, Z., Gervautz, M., "Studierstube - An Environment for Collaboration in Augmented Reality," Proceedings of CVE '96 Workshop, 1996.

[13] Soltan, P., Trias, J., Dahlke, W., Lasher, M., McDonald, M., "Laser-Based 3D Volumetric Display System: Second Generation," In Interactive Technology and the New Paradigm for Technology, IOP Press, pp.349-358. 1995.

[14] Tatham, E. W., "Getting the Best of Both Real and Virtual Worlds," Communications of the ACM, Vol.42, No.9, pp.96-98, 1999.