THE EFFECT OF SPATIAL CUES IN AUGMENTED REALITY VIDEO CONFERENCING

Hirokazu Kato[a], Mark Billinghurst[b], Kentaro Morinaga[a], Keihachiro Tachibana[a]

[a]Faculty of Information Sciences, Hiroshima City University
3-4-1, Ozuka-higashi, Asaminami-ku, Hiroshima 731-3194 Japan
[b]Human Interface Technology Laboratory, University of Washington
Box 352-142, Seattle, WA 98195 USA

ABSTRACT

In this paper we introduce an Augmented Reality video conferencing system and show some results of the evaluation. The AR user can see remote users on paper cards in real world. The user can freely put virtual monitors. Then they can use spatial cues to get the communication smoothly. Also we appended a new function that eliminates background scene of desktop users. Evaluation results showed that video images play a important role for both AR and desktop users even if HMD occludes AR user's eyes. Desktop user could identify the conversation relationship from AR user's face angle. Also we expected that the elimination of background scene increases co-presence, but we could not confirm that from the experiments.

1. INTRODUCTION

Recently many communication technologies such as a mobile technology have been developed and we can expect that many kinds of communication systems will be used in near future. The idea to integrate visual channel to audio communication is quite old. However video communication systems have not spread widely yet because they have some technical difficulties. One of the biggest problems was narrow bandwidth for image transmission, but it will be resolved by the advance of wireless or optical communication technologies and compression technologies. In the current situation, system developments with the consideration of human factors would be more important for the video communication. We have been developing a video conferencing system in Augmented Reality (AR) environment (Kato 1999). Remote persons are displayed in virtual monitors on physical markers by using see-through Head Mounted Display (HMD). The advantages of the system are as follows:

1) It is suitable for mobile use because it dose not need traditional displays.
2) It can change the size of display in appearance from hand-on size to actual human size.
3) Virtual monitors can be put freely in the physical workspace.

These mean that the user can use the spatial cues which traditional systems dose not have. Therefore we think our system is effective for recognition of communication relationships and give some influences to co-presence with remote persons. However it also has disadvantage of eyes occluded by HMD. In this paper, we introduce our video conferencing system and shows some results of the evaluation.

2. AUGMENTED REALITY VIDEO CONFERENCING

Figure 1 shows an example of the video conferencing system in which user can use spatial cues. This system uses a monitor and a video camera for each remote person and put them around the table. Then the user can easily get conversation relationships or remote user's attention. However this system needs same number of monitors and cameras as remote persons.

Augmented Reality(AR) technology has a possibility to overcome such a problem. In AR environment the use can see virtual objects in the real space by using HMD (Azuma 1997). That is, the user put each remote person as a virtual object on the table as he likes. So the user can get the conferencing environment where remote persons are sitting around the table.

We developed an AR video conferencing system as shown in Figure 2. The user can see remote persons on paper markers through the HMD. This system dose not use any traditional physical monitors, but assigns a paper marker for each remote person. Also participants can share the virtual whiteboard to collaborate with each other. However this system assumes that there is one AR user and other desktop users. Because AR user can see each desktop user as a virtual object on a paper card, but there are no cameras to capture the image of the AR user. This means that desktop users cannot see AR user's face but share the same view as the AR user's view. This configuration is suitable for remote work assisting (Billinghurst 1999). In this paper we introduce a new AR video conferencing system in which each paper marker has a video camera to capture the user's face image. This system also has a function to get the remote user's background transparent.

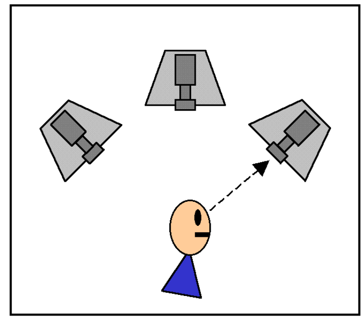
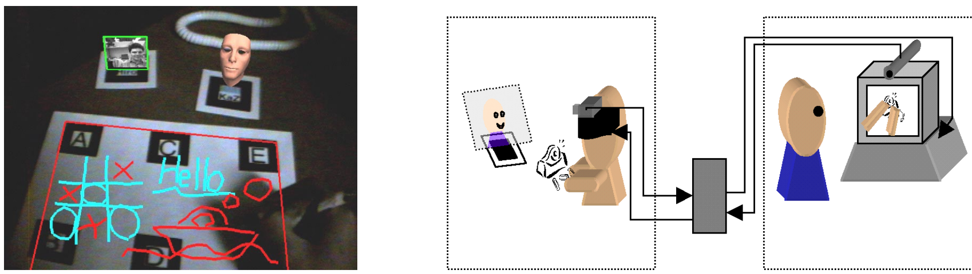


Figure 1. An example of video conferencing system in which the user can use spatial cues.



(a) AR User's view (b) System configuration

Figure 2. An AR video conferencing system we developed. AR user can see desktop users, but desktop users cannot see AR user's face but same view as AR user.

2.1 MARKER BASED AUGMENTED REALITY

In order to display remote person on a paper card as a virtual object, we used free software library called ARToolKit for making Augmented Reality applications (Kato 2000). As shown in Figure 3, ARToolKit detects square markers from an image captured by the camera attached on the HMD. Then it estimates the pose and position of the markers and display virtual objects on those marker coordinates. So we can use a paper card with black square as a virtual monitor by drawing a video stream from a remote computer on the card.

2.2 VIRTUAL MONITOR WITH A CAMERA

As mentioned above, our early system doesn't have a camera on a marker. So desktop user cannot see AR user's face but he can see the same view as AR user while AR user can see the desktop user. We designed this system for remote work assisting and it was useful. However when we assume the general conversation, it is not good that the desktop

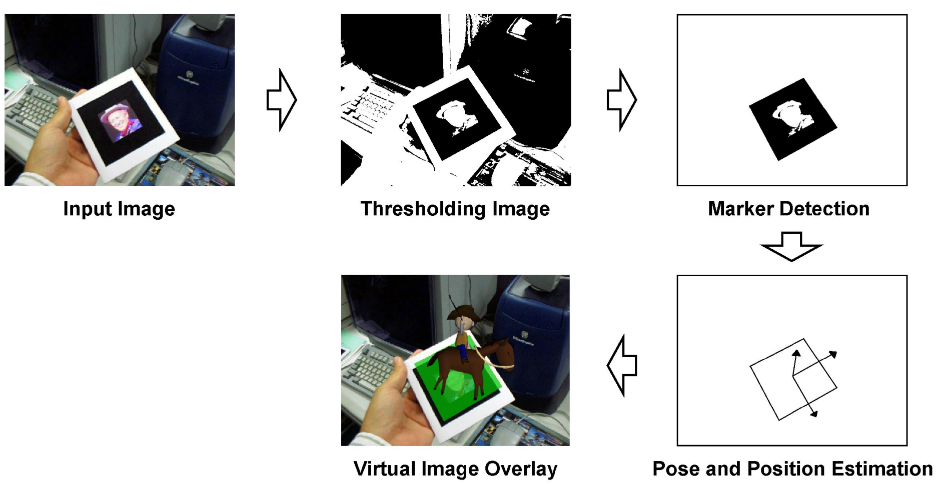


Figure 3. Processing flow of ARToolKit

user cannot see an AR user's face. In order for desktop users to see an AR user's face, a small camera was attached within a paper card as shown in Figure 4.

Suppose that there are two desktop users and one AR user. When the AR user looks at one desktop user, the desktop user can look at the AR user's front face like Figure 5(a). Also the other desktop use can look at the AR user's side face like Figure 5(b). Therefore desktop users can identify if the AR user intend to talk with them. However since AR user wears HMD, desktop users cannot see his eye. This could be a problem on the communication.

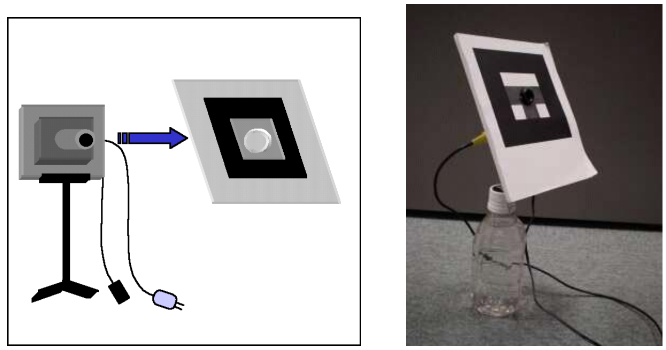
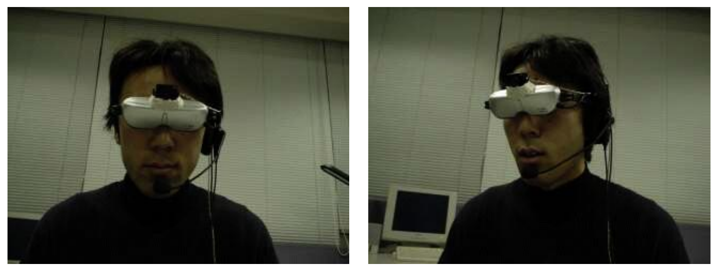


Figure 4. A small camera attached in a card



(a) Front view of the AR user (b) Side view of the AR user

Figure 5. Desktop users can identify if the AR user looks them by his face angle.

2.3 BACKGROUND ELIMINATION

While a remote user is displayed on a virtual monitor, background scene is also displayed on a virtual monitor. This makes a discontinuity between real world and a virtual monitor. If the background scene is eliminated from the virtual monitor, such discontinuity is gone. So we thought elimination of the background increases the co-presence with remote users. Since the camera for desktop user is fixed, background subtraction technique can be used to eliminate the background (Figure 6).

(a) With background scene (b) Without background scene

Figure 6. AR user's view

3. EXPERIMENTS

We had 21 test users evaluate this system. After they discussed with each other by using this system, they answered some questions on the questionnaire. We classified 7 groups of 3 users who know each other well. Evaluated system consists of one AR conferencing interface and two desktop conferencing interfaces. They evaluated it in 3 conditions:

1) AR user: Sound & Video with background
   Desktop users: Sound only
2) AR user: Sound & Video with background
   Desktop users: Sound & Video
3) AR user: Sound & Video without background
   Desktop users: Sound & Video

They had a simple topic of conversation and discussed each other. After each 3 minutes discussion, they changed their environments (AR user and desktop user) and totally they had 9 minutes discussion.

Test users rated 5 questions on the questionnaire for the desktop user situation:

I) Easiness of talk
II) Easiness of awareness of remote person
III) Clearness of conversation relationship
IV) Easiness of the transmission of intention
V) Co-presence

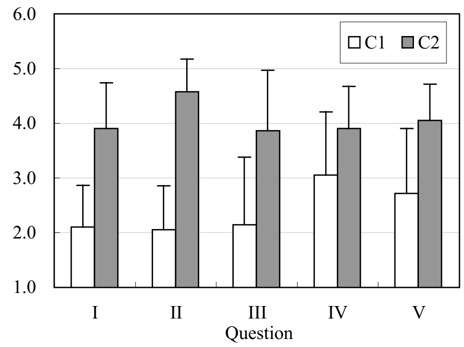


Figure 7. Desktop user's ratings in condition 1 and 2

Figure 7 shows average scores and standard deviations for each question in condition 1 and 2. In all questions condition 2 is significantly better than condition 1. This means that video stream played an important role for the conversation while the AR user wore the HMD. They could understand the conversation relationships by the AR user's face angle.

For the AR user situation, question 6 were appended:

VI) Satisfaction of video quality on the virtual monitor

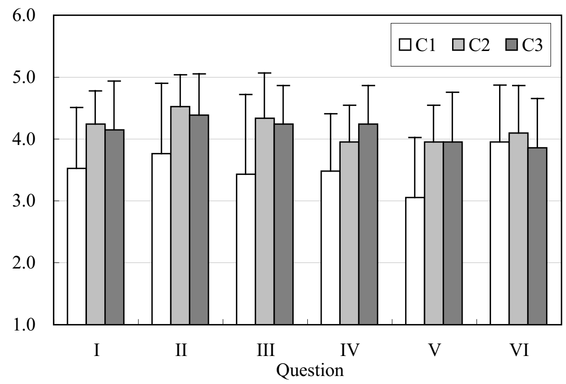


Figure 8. AR user's ratings

Test users rated all 6 questions in 3 conditions as shown in Figure 8. In comparison with condition 1 and 2, condition 2 is significantly better than condition 1 for question I, II, III and V. In condition 1 and 2, AR users were same situation while desktop users were different condition. However AR user felt difference. In comparison with condition 2 and 3, we expected difference of co-presence. However there aren't significant differences between them in any questions. We think that one of the reasons is the way to display remote users. As can be seen in figure 6, remote users are in the air above the table. This gave users the feel of non-reality. Figure 9 shows another example to display a remote user on the paper marker. In this case users felt co-presence much. Because the remote user was put just on the table stably and they didn't feel the non-reality. We need to evaluate about co-presence again in such configuration.

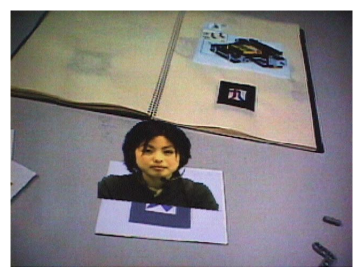


Figure 9. Remote user on the table without background

4. CONCLUSION

In this paper we introduced an Augmented Reality video conferencing system and showed some results of the evaluation. The AR user can see remote users on paper cards in real world. Also we appended a new function that eliminates background scene. Evaluation results showed that video images play an important role for both AR and desktop users even if HMD occludes AR user's eyes. Desktop user could understand the conversation relationship from AR user's face angle. Also we expected the elimination of background scene increase co-presence, but we could not confirm that from the experiments.

REFERENCES

Azuma, R. T. (1997). A Survey of Augmented Reality. Presence: Teleoperators and Virtual Environments, 6, 4, 355-385.

Billinghurst, M., Bee, S., Bowskill, J., Kato, H. (1999). Asymmetries in Collaborative Wearable Interfaces. Proc. of 3rd Int. Sympo. on Wearable Computers, 133-140.

Kato, H., Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System. Proc. of 2nd Int. Workshop on Augmented Reality, 85-94.

Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., Tachibana, K. (2000). Virtual Object Manipulation on a Table-Top AR Environment. Proc. of IEEE and ACM International Symposium on Augmented Reality 2000, 111-119.