Effect of Wide FOV and Image Stabilization on Spatial Perception for View Sharing System

Daisuke Kondo[1], Kyo Hattori[2], Keitaro Kurosaki[1], [1]Hiroki Kawasaki[1], Yuki Hashimoto[1], Tomoko Yonemura[1], Hiroyuki Iizuka[1], Hideyuki Ando[1] and Taro Maeda[1]

[1] Graduate School of Information Science and Technology, Osaka University
[2] Faculty of Engineering, Osaka University

ABSTRACT

We have proposed a view-sharing system where two distant people can share their views by mixing or exchanging images captured by head-mounted cameras with HMD. Our final goal of the system is to share spatial perception and to transmit skills from one to another. To realize such a transmission, there are two problems such as a narrow field of view (FOV) of HMD and shaky images due to inconsistent motions between two. To overcome these problems, we propose wide-FOV video-see-through HMD and stabilizer function of images to maintain spatial perception. Our proposed method is evaluated by the spatial recognition task under CG environment.

KEYWORDS: Remote cooperative work, view sharing, spatial perception

INDEX TERMS: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - Artificial, augmented, and virtual realities

1 INTRODUCTION

By extending the rotobot-human telexistence[1] technology to human-human situation, we are developing an environment where a skilled person, who actually exists at a different place, can work with high quality on the ground instead of non-skilled person. The skilled person feels as if he exists at the place and work there. The non-skilled person can show high-quality performances with the skilled person's help. In order to realize such a telexistence environment in human interactions, remote communication technologies exploiting sense-motion sharing is proposed.

In the field of computer supported cooperative work, human-human collaborative interactions between two distant people has been studied. However, most studies only pay attention to how visual information should be presented or shared without considering the viewpoints. For example, to simplify collaboration and increase effectiveness, Kuzuoka et al. [2][3] and Fussel et al. [4] developed similar shared-view video support systems where a worker's view is captured by a head-mounted camera, but the captured images are displayed to an instructor on a monitor on a desk. The instructor can see the worker's view, but the perspectives are not consistent. The inconsistent perspectives require the user (instructor) to consciously interpret from the displayed images such situations as the worker's movements and the spatial relation among objects and the worker. On the other hand, a consistent perspective induces an immersive experience and the feeling of presence, and users can unconsciously understand the surrounding environments.

We have a developed view-sharing system to share the first-person perspectives between remote two people[5]. The system consists of a head-mounted display and cameras, which realize video-see-though. The users wearing the video-see-through HMD can see his own and the partner's view and also send his own view to the partner. The sharing system has been applied to a skill transmission and learning task. However, there are two following problems in the system that may make the performance of the system for sharing senses and motions less.

One is that the displayed view swings regardless of his own motion because the view is captured on the basis of the partner's motion. The discrepancy between motion and visual information causes breakdown of user's embodiment, which makes it difficult to maintain proper spatial perception such as positional relationship between an object and a world, a recognition of posture with which the partner see the view, the relationship between his own and a world. This problem can be overcome by giving the sender's both motion and visual information to the receiver to maintain the coherence between them in the receiver's situation. By realizing an interface device that supports the receiver to follow the sender's motion and that displays the flow of images in response to the receiver's actual motion, the embodiment or coherency between motion and visual information can be maintained and the proper spatial perception can be sent to the receiver.

Another problem is a narrow view angle. The conventional video-see-through HMDs do not equip enough view angles compared with our naked eyes. It is shown that wider view angles improve the immersiveness to the view and abilities of spatial perception and searching in a space[6][7]. A new design for video-see-through HMD with wider view angles resolve lack of the visual information in the view-sharing system.

2-1 Yamadaoka, Suita, Osaka, Japan
{kondo, iizuka, hide, t_maeda}@ist.osaka-u.ac.jp

20th International Conference on Artificial Reality and Telexistence (ICAT2010), 1-3 December 2010, Adelaide, Australia
ISBN: 978-4-904490-03-7 C3450

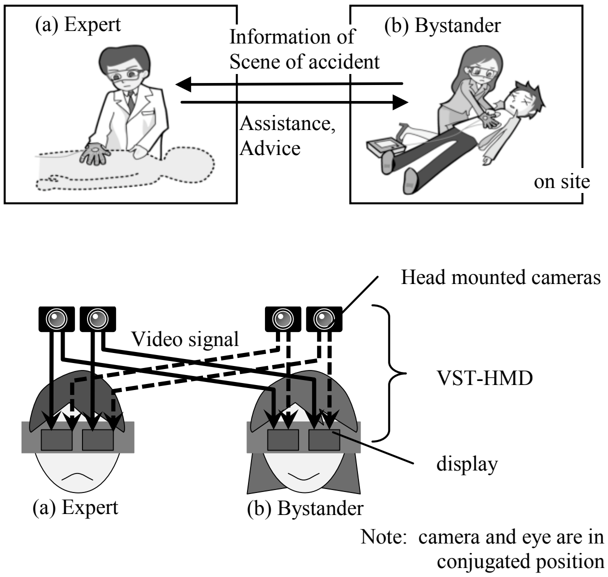


Figure 1. The concept of view-sharing system.

In this paper, we propose a spatial stabilizer of flow of views on a video-see-through HMD with wide view angle to overcome these two problems. The effect of the spatial stabilizer and view angles are tested and evaluated by the performance of communication of spatial perception from one to another.

2 VIEW SHARING SYSTEM

In this section, system arrangement and construction will be described. Here, we assume the situation such as scene of an accident for example. Our concept is that the bystander performs the emergency treatment according to the expert's direction transmitted from other place.

As shown in figure 1, two users wear the video-see-through HMDs at the different places. The left user (a) is a doctor who can direct the emergency treatment, and another (b) bystander is layperson who actually executes it under the instruction of expert's direction.

The expert, who has knowledge of techniques for the emergency treatment, is not on site, uses the view sharing system to understand conditions of the injured person and the environment, and then the expert directs the bystander. The bystander achieves cardio pulmonary resuscitation, for example, the cardiac massage. Bystander can execute the cardiac massage with precision by receiving the instruction of cardiac massage; the posture, the position to put hands, timing, etc. in real time.

The purpose of this system is to transmit the vision and motion between the users. For expert user, in order to provide the bystander user with precise direction, sharing first person view is important. Because sharing first person view helps expert to understand situation in the site of bystander's place with accuracy, such as spatial relationship of bystander and injured person, furthermore the notice what is within arms' reach are required.

The exact correspondence of viewpoints is required for our purpose. In our previous research, using VST-HMD with the precise correspondence of viewpoints improved the work performance and spatial perception.

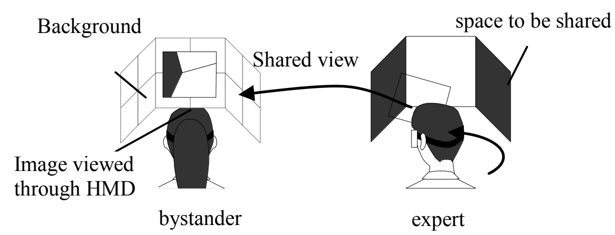
2.1 The procedure of view sharing

The image that captured by the expert's camera is transmitted to the bystander's HMD, and the bystander user can share the expert's viewpoint. The spatial perception is strongly related with ones the head posture. Therefore, correspondence of head postures between the bystander and expert is required. As shown in figure 2(a), the expert turns his head left to instruct the bystander to turn left, and the images are send to the bystander's eyes. Here, the image seen by bystander is not corresponds to the bystander's head posture, therefore the bystander may not understand the spatial information from the image with accuracy. The bystander has to follow the movement of shared image, to correspond to her head orientation. Finally, in figure 2(b), the bystander turns her neck left.

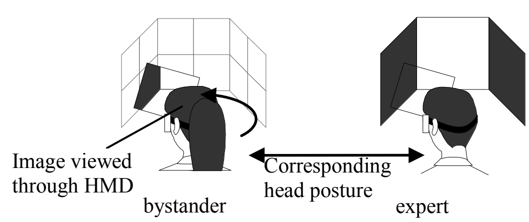
2.2 Tracking support by guide markers

In order to follow the other user's head motions, a Tracking support using guide markers is introduced. In figure 3, two markers are in front of the user, (A) is fixed in front of the user, it means posture of user's own head. Another one (B) is target marker, and it means partner's current head posture. Therefore, the positional relationship between center marker (A) and target marker (B) reflects the positional relationship between the user and the partner.

The user has to follow the target marker by moving head. When the both marker are aligned, the user and partner's motion are successfully corresponding.



(a) The expert moved and the image is send to the bystander's eyes.



(b) The bystander moves following the expert, and the bystander's orientation corresponds to expert's. It means the expert's direction is transmitted to the bystander.

Figure 2. The overview of the view sharing and motion tracking.

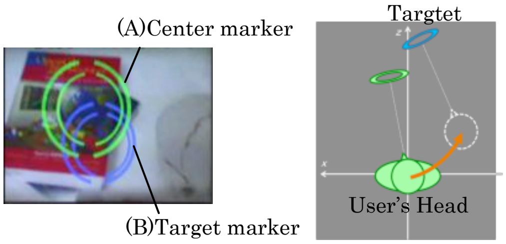


Figure 3. Positional relationships of guide markers.

3 Proposed Method

3.1 Target tracking in the other user's view

It is difficult for the user to track the marker in the partner's view, because during the users are moving, the image provided to the user do not respond the head motion. Ordinary, one's view as the background and its optical flow is an important cue of one's faculty of orientation during head motion. In this system, the background that is not correspond the user's head motion makes the user feels the scene is wobble or disoriented. Finally, the user's head movement is disturbed. The shaky images due to inconsistent motions between two must be eliminated.

3.2 Spatial Stabilization

If the user can track with precision to the other user's vision corresponds to the user's head motion, the user can recognize the other user's view as own background naturally. But the perfectly corresponding is difficult. The background image that does not correspond the user's head motion gives the user a disturbance in task. the shaky images due to inconsistent motions between two must be eliminated. In this section, the method for eliminate and stabilize the background will be introduced.

The stabilization technique is sometimes used in the field of tele-robotics[8] where the camera image attached on the mobile robot is need to be stabilized.

In following section, we introduce the two methods named "Image stabilizer" and "Timeline stabilizer."

3.2.1 Image stabilizer

The function named "image stabilizer" that transforms the image to generate the nearest image to the correct one. The function slides, warps and rotates the partner's camera image according to the relative position and orientation between the user and the partner in real time. The concept image is shown in figure 4. Image stabilization is implemented using OpenGL functions, such as texture mapping. In actuality, the image stabilizer is implemented as the transformation of texture coordinates by following matrix,

$$\left(F \times M^{-1} \times W \times T^{-1}\right)^{-1} \quad (1)$$

Here, M,F,W,T are 4x4 matrix. M is the user's eye's position and orientation, W is the partner's, and F is projection matrix expresses camera parameters, T is the translation matrix means the distance to assumed attention subject. the process is performed for the right and left eyes individually.

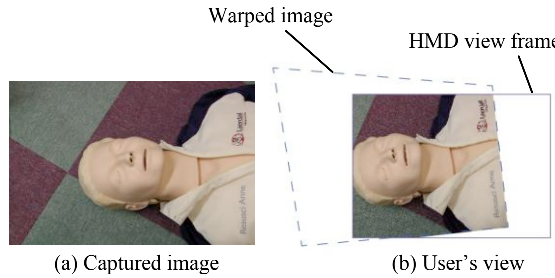


Figure 4. Method of image stabilization.

3.2.2 Timeline Stabilizer

Previous image stabilize function is a pseudo elimination of the image shaking derives from the users' head rotation, but not positional movement.

"Timeline stabilizer" method supports pseudo compensation of positional movement of user's viewpoint. The function stocks the another person's view in the memory as a series of images, and the system select one image that is taken from the position nearest from the user's current viewpoint. The timeline stabilization is effective when the user (bystander) follows the partner (expert) on the similar path.

Timeline stabilization is implemented using OpenGL, the stereo pair images of another person's camera is to be memorize on the graphics hardware, and rerated motion data is stored with it. The images of latest four seconds are stored, and older one are deleted sequentially, namely 240 pairs of images are kept in the video memory. Second, the user's current position is measured, and then the system finds the nearest one from the stored position data, and a pair images linked to the position is selected. Third, a spatial stabilization filter is processed on the selected pair of image, and finally, the image is shown to the user.

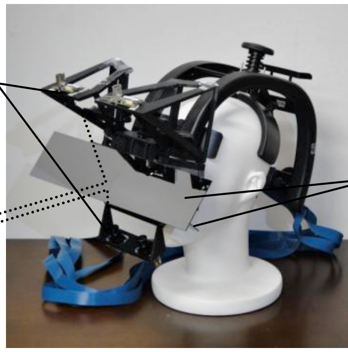
3.3 Wide field of view VST-HMD

To solve the second problem; the narrow FOV, the VST-HMD with wider FOV is discussed. The conventional VST-HMD we have developed is based on eMagin z-800 head mounted display that has horizontal 32 degrees and vertical 24 degrees field of view.

We are developing a wide field of view video see-through HMD. It is based on the Sensics piSight head mounted display in figure 5(a), and the cameras positioned around the display unit. The optical design is shown in figure 5(b). The piSight we used is binocular stereoscopic, wide FOV HMD, the display consists of 6 LCD panels and has 2400x1200 pixels(including overwrapped areas) and provides approximate horizontal 90 deg., and vertical 45 deg. field of view for each eyes.

The cameras are Firefly MV (Point Grey Research), the cameras are positioned optical conjugate point. The eye and the camera are located optical conjugate point, thus the camera takes the image placing the user's eye position. And the view angle of camera is horizontal 90 deg., and vertical 45 deg., covering the field of view of piSight.

(a) piSight HMD



(b) piSight with wide FOV camera unit

Figure 5. The wide FOV VST-HMD.

4 EXPERIMENT

Experiments have been performed to evaluate our proposed method. In the experiments, we measure accuracies of spatial perception in a situation where a subject sees the partner's view and has to understand the partner's space. It means that the view basically flows in response to the partner's motion. In such a situation, it is evaluated how accurate subjects can perceive the relative position between objects placed in a virtual space. The objects are drawn by CG. The marker in the view is displayed to show the partner's head motion and the subjects follow the marker and understand the space in the flow of images.

The following conditions are examined in terms of view angles and use of the spatial stabilizer. There are two conditions for the view angles such narrow and wide. The view angles are shown in Table 1. The narrow view angle is set according to the conventional video-see-through HMDs and the wide view angle is based on the pi-Sight.

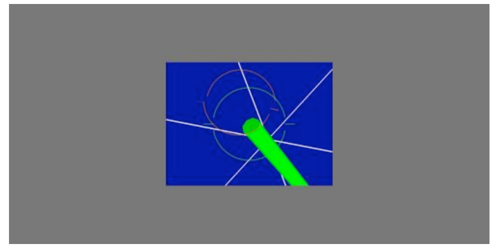
Table1: View angle condition(per one eye).

	Narrow condition	Wide condition
Display FOV	Horizontal 32° Vertical 24°	Horizontal 90° Vertical 45°
Camera FOV	Horizontal 40° Vertical 30°	Horizontal 90° Vertical 45°

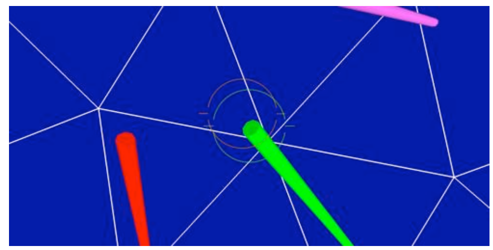
There are two conditions for the spatial stabilizer such as on and off. The stabilizer makes the flows of images stable by calculating the images captured by the head-mounted camera in the virtual space.

The camera FOV is 40° x 30° in narrow condition, and 90° x 45° in wide condition. When the condition is narrow and stabilizer-off, only 32° x 24° area of FOV is displayed on HMD.

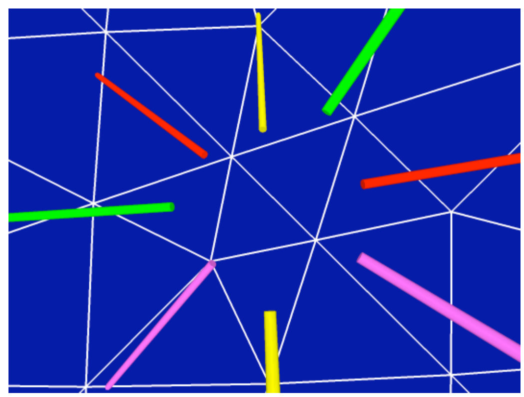
In order to control the conditions, the only piSight is used for the head-mounted display in the all conditions. In the narrow conditions, the display area is reduced to the narrow view angles. The examples of the views are shown in Fig. 6.



(a) Narrow view angle.



(b) Wide view angle.



(c) Whole virtual space
(this scene is not shown to the subjects.)

Figure 6. The virtual space represented to the user.

Throughout the experiments, Polhemus electro-magnetic position and rotation sensor, Fastrak, is used to detect the head position and motion of subject. In stabilizer-use condition, this head motion information is used by the image stabilizer and the timeline stabilizer to make the images stable.

4.1 Task

There are 4 pairs of colored cylinders in the virtual space. The different pairs have different colors, i.e. 4 kinds of colors. The cylinders are placed annularly. The only single pair of them is placed coaxially as shown in Figure 6(c). It is yellow pair in the example. The subjects are required to answer which colored pair is placed in coaxially after following the maker which represents the partner's head motion. The partner's motion is generated as looking-around motion in which all cylinders are passed in the images at least once and the average speed is set to 25 degrees/sec. All trajectories are different. The background is triangle mesh pattern of "geodesic dome". The pattern is designed as a homogeneous pattern that doesn't have absolute horizontal or vertical cues, but gives the user a feel of optical flow to understand the partner's head orientation

4.2 Results

In the experiment, four young male adults participate in the experiment as subjects. We performed experiments 10 times for each condition. It means that the total number of experiments is 40. The orders of conditions are randomly determined. The percentage of right answers was recorded as a performance. The results are shown in Figure 7.

Two way ANOVA (alpha-level = 0.05) shows statistically significant main effects of view angle ($F(1,3)=32.0$, $p<0.05$) and image stabilization($F(1,3)=14.4$, $p<0.05$). The interaction between view angle and image stabilization was not statistically significant. Namely, the results show that wide view angle and our proposed stabilizer increase the rate of right answers in the task.

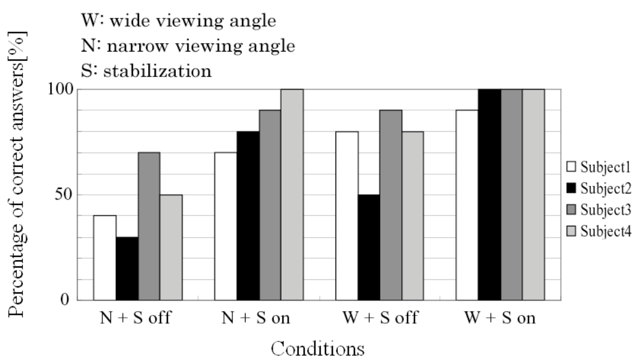


Figure 7. Experimental result.

5 Discussion

Our results show that the wide view angle improves the spatial perception in a situation where the images flow in response to the partner's motion. The reason of the improvement would be that the subject can see the cylinders relatively longer in the wide view angle condition and can see the wide area of the background, which makes it easier to understand the relative positional relationship between the objects and the background. Another reason will be that the wide area of the background helps to predict the partner's head motion from the flow of images.

The stabilizer could also improve the spatial perception. Even in the wider view angle condition, without the stabilizer, the subjects tend to mis-perceive the position of objects since the discrepancy between flow of images and their own motion occurs. When using the stabilizer, the coherence between motion and flow of images is maintained even if the subjects could not follow the partner's motion marker. Thus, it is considered that the subjects can correctly answers in the task.

In addition, it was shown that the stabilizer function and the wider view angle are compatible and produced the best performance. Even if the view angle of HMD is narrow, the stabilizer has effect to the spatial perception for view sharing system. This could be an important design principle to develop the view-sharing system.

6 Conclusion

In this paper, we proposed image stabilizer function which makes the sender's image flow stable for the receiver. The image flow is controlled according to the receiver's head motion. We investigated the effect of our proposed stabilizer function and view angles in the spatial perception task sent by the partner. From the results of our experiments, it was shown that the view angle is crucial to perceive the positional relationships between objects in the partner's image flow and that the stabilizer function could compensate for the gap of ability of spatial perception between narrow and wide view angles.

For the future work, the number of subjects needs to increase to clarify the significant difference between conditions. The effect of our proposed method in real space also needs to be shown.

Acknowledgments

This research was supported by JST, CREST. Research Collaboration and "Global COE (Centers of Excellence) Program" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] Susumu Tachi: Tele-existence - Toward Virtual Existence in Real and/or Virtual Worlds, Proceedings of the International Conference on Artificial Reality and Tele-existence (ICAT '91), pp.85-94, July 1991.

[2] H. Kuzuoka, Spatial Workspace Collaboration: a SharedView video support system for remote collaboration capability, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 533-540, 1992.

[3] N. Sakata, T. Kurata, T. Kato, M. Kourogi, and H. Kuzuoka, WACL:Supporting Telecommunications Using Wearable Active Camera with Laser Pointer, 7th IEEE International Symposium on Wearable Computers, pp. 53-56, 2003.

[4] S. R. Fussell, L. D. Setlock, and R. E. Kraut, Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vol. 5, No. 1, 513-520, 2003.

[5] Hiroki Kawasaki, Hiroyuki Iizuk, Shin Okamoto, Hideyuki Ando, Maeda Taro, Collaboration and Skill Transmission by First-Person Perspective View Sharing System, 19th IEEE International Symposium on Robot and Human Interactive Communication, September 2010. (in press)

[6] Arthur, K. W., Effects of Field of View on Performance with Head-mounted Displays, ISBN:0-599-73372-1, University of North Carolina at Chapel Hill Doctoral Thesis, 2000.

[7] Tao Ni, Doug A. Bowman, Jian Chen, Increased display size and resolution improve task performance in Information-Rich Virtual Environments, Proceedings of Graphics Interface 2006, pp.139-146, 2006.

[8] N.Shiroma, J.Kobayashi, E.Oyama, Compact image stabilization system for small-sized humanoid, 2008 IEEE International Conference on Robotics and Biomimetic, pp. 149-154, Feb. 2009.