Improvement of Wearable View Sharing System for Skill Training

Yuki HASHIMOTO, Daisuke KONDO, Tomoko YONEMURA, Hiroyuki IIZUKA, Hideyuki ANDO and Taro MAEDA

Osaka University / JST CREST

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. Users can share what another user is seeing, and furthermore these users can make their their spatial perception, motion and head movements correspond. Our goal is to transmit non-verbal skills from a skilled person to a non-skilled person by using this view-sharing system. To realize this goal, we have improved this system. This system is light weight, has a wide viewing angle, is stand-alone, and does not require calibration of intraocular distance. These features can facilitate more efficient skill training and expand the sphere of activity for using view sharing systems.

KEYWORDS: view sharing, skill training, parasitic humanoid, wearable system

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

1 INTRODUCTION

By extending robot-human telexistence[1] technology to human-human situations, we are developing an environment where a skilled person, who actually exists at a different place, can work with high efficacy on the ground instead of non-skilled person. The skilled person feels as if he exists at the place and can work there. The non-skilled person can perform with high quality with the skilled person's help. In order to realize such a telexistence environment in human interactions, we are developing remote communication technologies exploiting sense-motion sharing. In this project, we have developed a view sharing system to share first person perspectives between remote two people[2][3]. The system consists of a head mounted display and cameras, which makes possible a video-see-though head mounted display (VST-HMD). The user wearing the VST-HMD can see his own view and the partner's view, and also send his own view to the partner. The view sharing system can be applied to skill transmission and learning tasks. Previously, we considered the effectiveness of our view sharing system in some skill training scenarios [4][5][6][7][8].

Based on previous work, we developed a new view sharing system to improve effectiveness and expand its applications. In this paper, we describe our existing view sharing system. We also show the design and implementation of our new view sharing system.

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

2 VIEW SHARING SYSTEM

To transmit non-verbal skills, body position movement is one of the most important pieces of information. The view sharing system has a potential to transmit body position movement by visual information. Thus, we attempted to transmit three elements of body position movement–visual image, hand position and head motion–in our system. In this chapter, we describe our established system and the method of transmission of these three elements.

2.1 System Construction

The system arrangement and connection of devices are shown in Figure 1 and Table 1. The image taken by the cameras is sent to a PC through IEEE1394 cables, and the PC renders the output images by mixing the input images. The overall latency is 66 milliseconds.

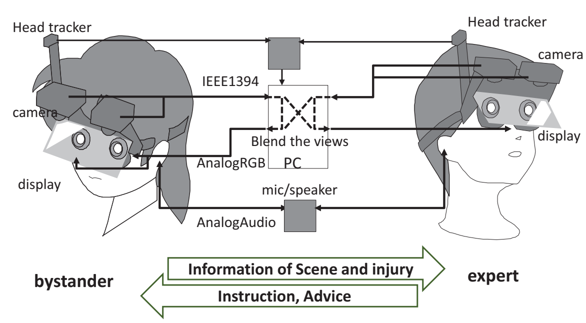

Figure 1. System construction

Table 1. Specification of the system

Display (HMD)	Model	eMagin Z800 3D Visor
	Resolution	800 x 600
	FOV	32° (horizontal) x 24° (vertical)
Camera	Model	Point Grey FireFly MV
	Resolution	752 x 480 (native)
		464 x 348 (effective resolution)
	Connection	IEEE1394a
Motion Tracker	Model	Polhemus LIVERTY (electromagnetic posision/rotation sensor)
	Update rate	240Hz
	Connection	USB2.0
PC	Model	Hewlett-Packard Z800
	CPU	Intel Xeon E5520
	OS	Windows XP SP3 32bit
	Graphics	nVidia QUADRO FX 4700X2

The 21st International Conference on Artificial Reality and Telexistence
November 28-30, 2011, Osaka, Japan
ISSN: 1345-1278

2.2 Visual image transmission

To get an exact first person viewpoint and improve spatial perception, we developed a video see-through HMD (VST-HMD). This VST-HMD has optical conjugation cameras (Figure 2), the user wearing the VST-HMD can see the image taken by the cameras attached on the user's head in real time with little delay. Basically, the user can see his own view as though he were seeing the view directly.

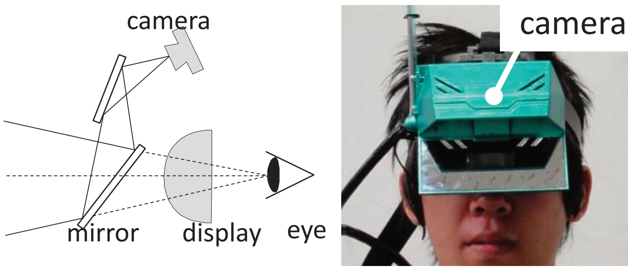

Figure 2. Orthoscopic video see-through HMD

2.3 Hand position transmission

To transmit hand position, image blending is used. In this case, the PC composes the view images from the user and the partner, and the user sees a blended view of him/her and the partner (Figure 3). The user can see the situation around the partner, and furthermore the user can follow the work performed by the other's hands. Because the user can see the partner's hands from a first-person viewpoint, the user can trace easily the action with his hands in real time to mimic the partner's motion.

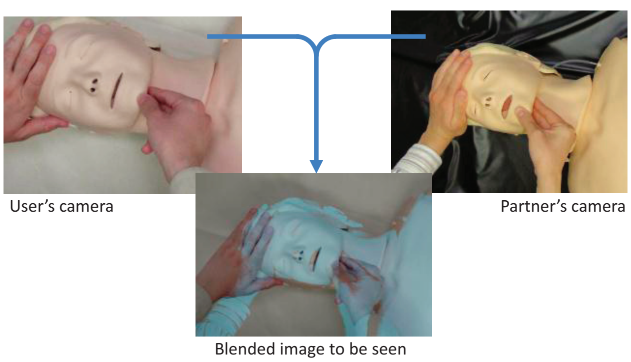

Figure 3. Image Blending

2.4 Head motion transmission

In order to follow the other user's head motions, a target marker is introduced. As shown in Figure 4, the user can see two markers, the green marker (A) is the center marker that is fixed in front of the subject, which indicates the orientation of the user's head. The other blue one (B) is the target marker, which indicates the other person's current head orientation. The marker is in the shape of a ring, with a diameter of 20 cm and a distance from head to the center of the marker of 70 cm. Therefore, the positional relationship between the center marker (A) and the target marker (B) reflects the positional relationship between the user and the partner. The user has to follow the target marker by moving his head. When both markers are aligned, the user and partner's motion are successfully corresponding.

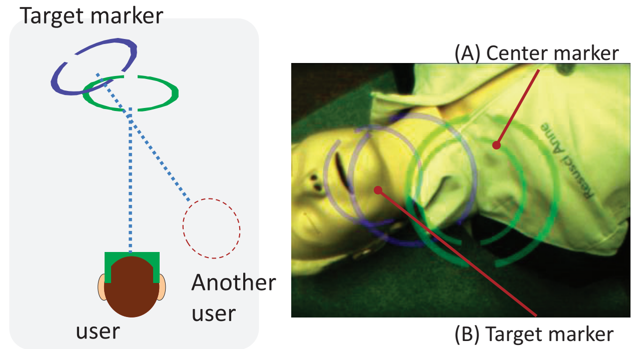

Figure 4. Spatial Relationships of Marker

3 Previous Trials and Problems of our System

3.1 Previous Trials

We have tried to use our system for skill training in the following four kinds of tasks.

1. *Rope work (Figure 5)*
2. *Playing the theremin (Figure 6)*
3. *Juggling (Figure 7)*
4. *Cardiopulmonary resuscitation (CPR) (Figure 8)*

Rope work and playing the theremin require adjustment of hand position. Juggling requires following arm motion over time. CPR requires following position and motion of the body, head and hand.
As a result of these experiments, we confirmed that our view sharing system is effective for skill training[4][5][6][7].

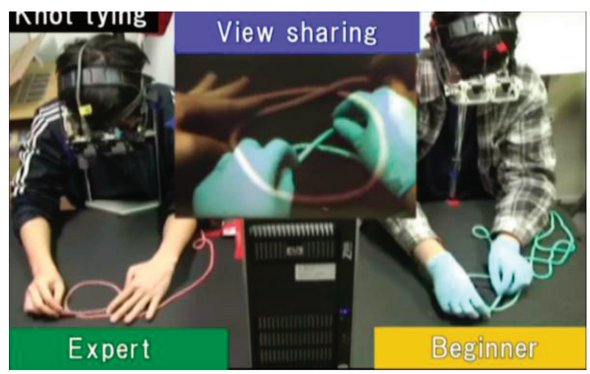

Figure 5. Rope work

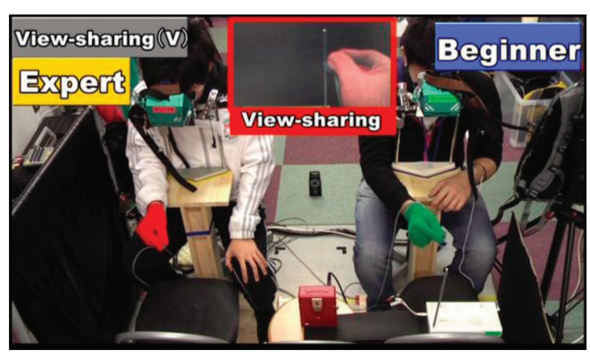

Figure 6. Playing the theremin

Figure 7. Juggling

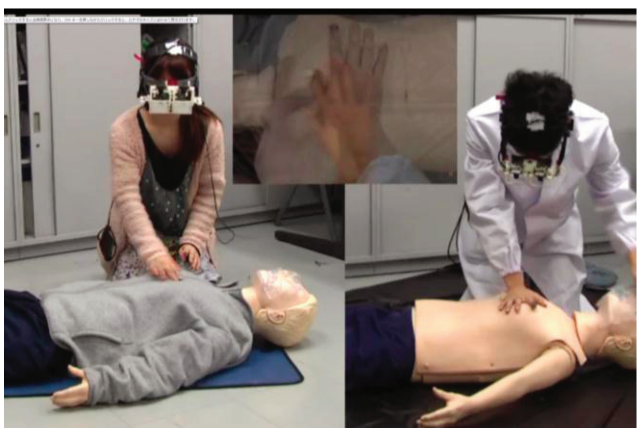

Figure 8. Cardiopulmonary resuscitation

3.2 Purpose

In order to improve the effectiveness of the view sharing system and expand its applications, we proposed a new view sharing system. In this system, we focused on two points. One was to improve the architecture of the VST-HMD to make it easier to wear and to widen the viewing angle. Another one was to make it a stand-alone system to extend users' range of possible actions.

4 Improvements in the New View Sharing System

4.1 System Construction

The system arrangement and connection of devices are shown in Figure 9 and Table 2. In order to realize a stand-alone system, a sensor was changed from an external sensor (Polhemus LIBERTY) to an internal motion sensor (Tokin MDP-A3U9S). The PC was also changed from a desktop PC to a laptop PC. In addition, this system has a battery in its control unit. We simplified connections between the VST-HMD and control unit. Images from cameras and data from motion sensors are sent to the PC through USB2.0. USB2.0 also supplies power to the devices. Therefore, the number of connecting cable between VST-HMD and control unit has been reduced to two (previous system = 7).

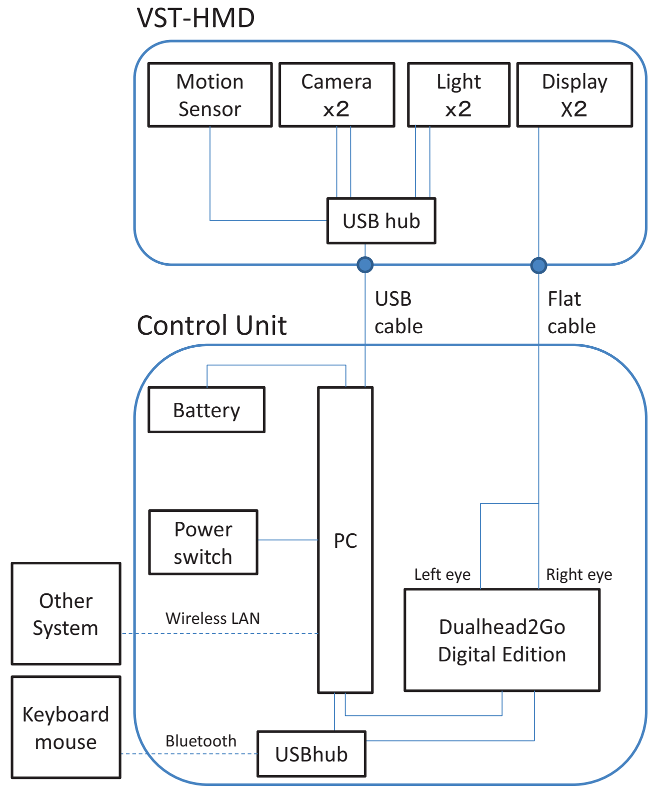

Figure 9. System Construction

Table 2. Specifications of the system

Display	Model	Daeyang FX603
(HMD)	Resolution	800 x 600
	FOV	42° (diagonal)
Camera	Model	Point Grey FireFly MV
	Resolution	752 x 480 (native)
		464 x 348 (effective resolution)
	Connection	USB2.0
Motion	Model	NEC/Tokin MDP-A3U9S
Tracker	Update rate	125Hz
	Connection	USB2.0
PC	Model	Apple Macbook air 11inch
	CPU	Intel Corei7 1.8GHz
	OS	Windows 7 64bit
	Graphics	Intel HD Graphics 3000

4.2 Implementation

4.2.1 Overview

The appearance of the proposed View Sharing System is shown in Figure 10. Weight is drastically lower than the existing system, and there are only two components; the VST-HMD and a vest including the battery, circuits, and laptop PC (Figure 11). Battery life is approximately three hours. For communication, we mainly use wireless networking so that the area of user's actions is not

limited by wires. However, wired networking may be needed in situations which require rapid following.
Therefore, a user can wear the system easily, and expand his/her range with a View Sharing System.

Figure 10. Proposed View Sharing System

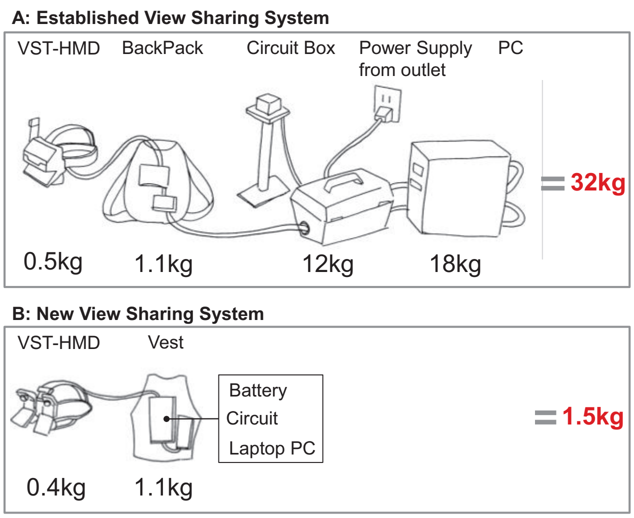

Figure 11. Components and weight

4.2.2 VST-HMD

The VST-HMD part is composed of two cameras (PointGrey FireFly MV), two displays (Daeyang FX603), two mirrors, two lamps and a motion sensor (NEC/Tokin MDP-A3U9S) (Figure 12). The viewpoint of camera and eye maintains an optically conjugated situation. The motion sensor is attached to the right side of the display unit. The display is bonded to goggles. The edges of the goggles are covered with soft rubber so that the eyes and goggles maintain close contact (Figure 13).
This VST-HMD has three advantages compared to the established system. The first is simplicity in wearing the system. The user can wear the proposed system like a pair of goggles, and will not need to adjust the fit after wearing them. The second is that there is no need to adjust the distance between the two displays to a user's interocular distance. Because a display contacts each eye, the viewpoint of each eye and its display is maintained without adjustment (Figure 14, Figure 15). However, the angle of convergence must be adjusted. To solve this problem, we have used software calibration with ARtoolkit. The user sees a marker with the VST-HMD, and we acquire each camera's position and rotation from the marker. The angle of convergence is calculated from this information.

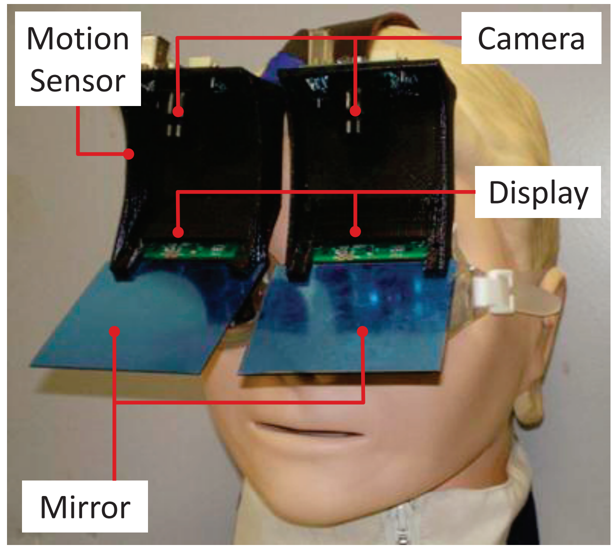

Figure 12. Proposed View Sharing System

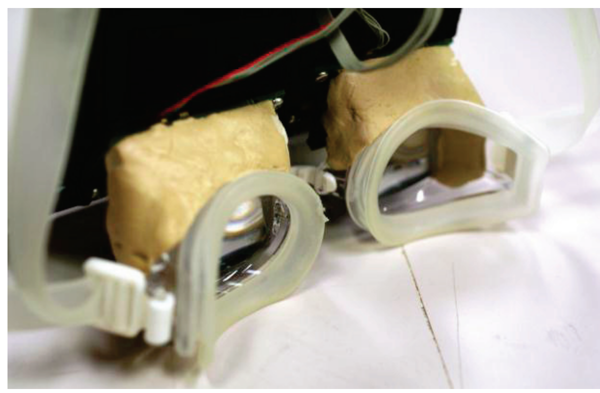

Figure 13. Contact part to eyes

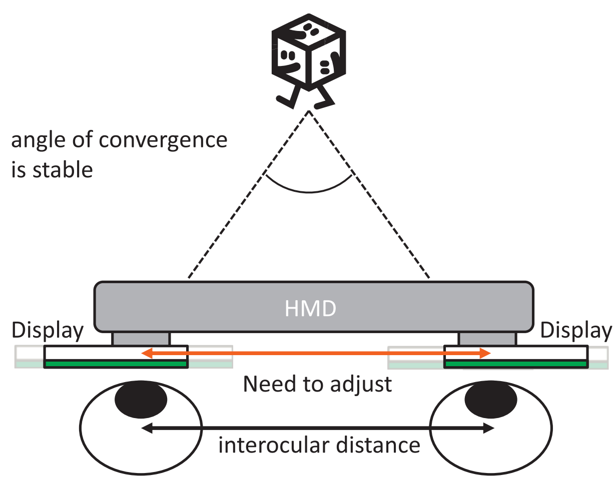

Figure 14. Relationship between interocular distance and angle of convergence in established the VST-HMD

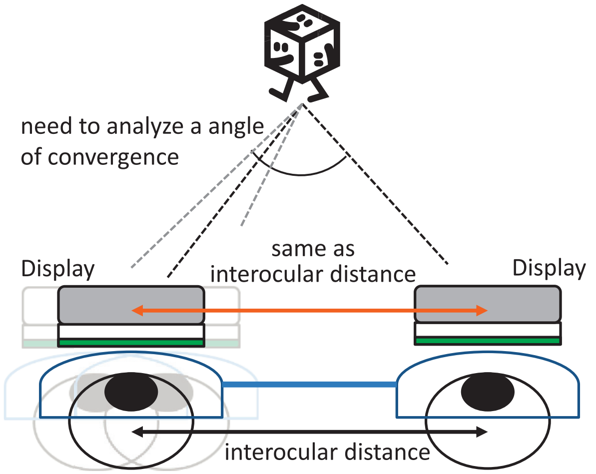

Figure 15. Relationship between a Interocular distance and an angle of convergence in proposed VST-HMD

The third is a wide view angle. The field of view is expanded because the display unit is fitted to the curvature of the face (Figure 16). The field of binocular vision (FOBV) of the proposed VST-HMD has been measured to be about 50° (horizontal) by 24° (vertical). In contrast, the FOBV of established the VST-HMD was approximately 35°(horizontal) by 24°(vertical) (Figure 17). However, the overlapping area between the two views was reduced (approximately 30° horizontal in the established system versus approximately 15° horizontal in the proposed system, distance = 50cm). One of our next steps will be optimization of the overlapping area and viewing angle.

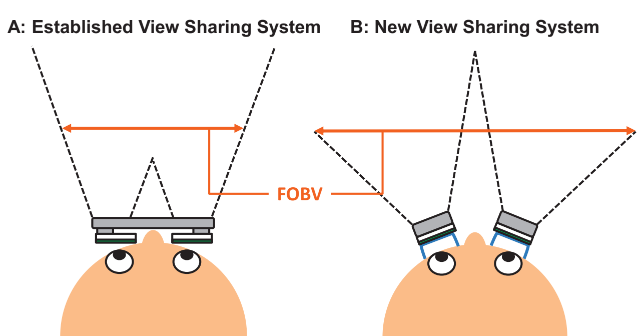

Figure 16. FOBV of VST-HMD

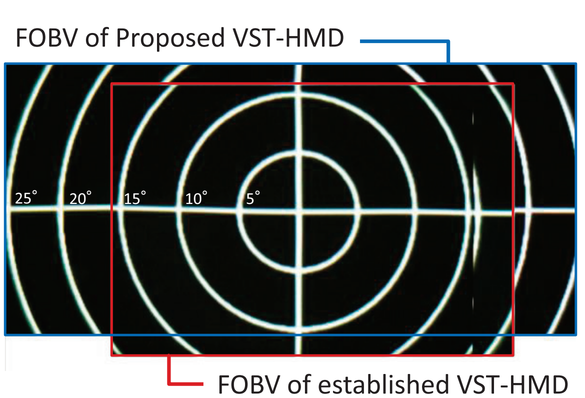

Figure 17. Measurement of FOBV

4.2.3 Control Unit

The control unit is composed of a PC board (Apple Macbook Air 11-inch), a circuit (Matrox Dualhead2Go and USBHub) and a battery (Japan Trust Technology Energizer XP8000). To trim weight, we deconstructed and rebuilt these components. For the laptop PC in particular, we removed the display, internal battery, a keyboard and cut the chassis. As a result, the weight of the PC board was reduced to 280g, making the total weight 800g (without cables) (Figure 18). The control unit is stored in the back pocket of a vest (Figure 19).

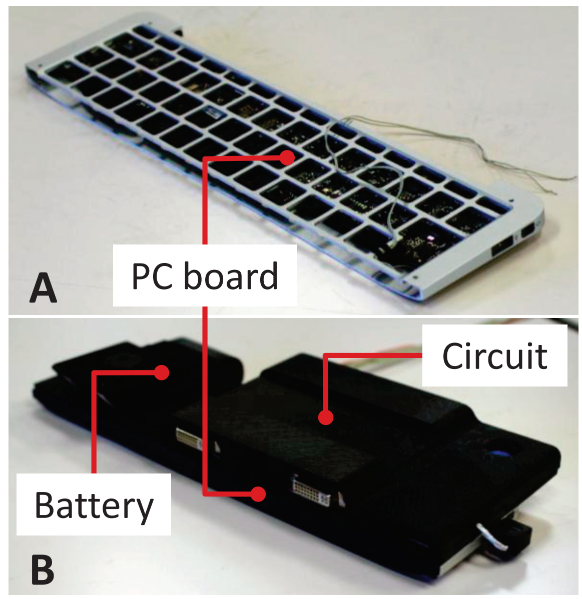

Figure 18. A: PC board, B: Control unit

Figure 19. Circuit unit is stored by back pocket of vest

5 Conclusion

In this paper, we propose a new view sharing system to facilitate more effective skill training and an expanded sphere of activity.

After describing our established view sharing system and previous trials, we explained the concept for a proposed view sharing system and described a prototype.

This system is expected to enable a user to work more efficiently and comfortably. In addition, the system will be able to support tasks which require larger physical movements, such as dance, baseball, or complicated emergency medical procedures.

As a next step, we will solve some issues of the proposed view sharing system, and compare its effectiveness in skill training with the old system. We will also use this system in new applications which require larger body movements.

6 ACKNOWLEDGMENTS

This research was supported by JST, CREST.

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, 1991.

[2] T. Maeda, H. Ando, H. Iizuka, T. Yonemura, D. Kondo and M. Niwa : Parasitic Humanoid: The Wearable Robotics as a Behavioral Assist Interface like Oneness between Horse and Rider, 3rd Augmented Human International Conference, 2011.

[3] D. Kondo, K. Hattori, K. Kurosaki, H. Kawasaki, Y. Hashimoto, T. Yonemura, H. Iizuka, H. Ando and T. Maeda : Effect of Wide FOV and Image Stabilization on Spatial Perception for View Sharing System, Proceedings of 20th International Conference on Artificial Reality and Telexistence, 2010.

[4] K. Kurosaki, K.. Hamada, D. Kondo, H. Iizuka, H. Ando, T. Maeda : Development and Evaluation ofa Zero Eye offset and Orthoscopic Video See-Through Head-Mounted Display, VRSJ the 15th Annual Conference, 2009. (in Japanese)

[5] K. Kurosaki, H. Kawasaki, D. Kondo, H. Iizuka, H. Ando and T. Maeda : Skill Transmission for Hand Positioning Task through View-sharing System, 3rd Augmented Human International Conference, 2011.

[6] H. Kawasaki, H. Iizuka, S. Okamoto, H. Ando, T. Maeda : Collaboration and Skill Transmission by First-person Perspective View Sharing System, 19th IEEE International Symposium in Robot and Human Interactive Communication, 2010.

[7] D. Kondo, K. Kurosaki, H. Iizuka, H. Ando and T. Maeda : View Sharing System for Motion Transmission, 3rd Augmented Human International Conference, 2011.

[8] H. Iizuka, H. Ando and T. Maeda : The Anticipation of Human Behavior Using Parasitic Humanoid, Proceedings of the 13th International Conference on Human-Computer Interaction. Part III: Ubiquitous and Intelligent Interaction, 2009.