The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Asynchronous Visual Information Sharing System with Image Stabilization

Naoji Shiroma and Eimei Oyama

***Abstract*— With advances in wearable computing, wearable virtual reality and wearable robotics have become popular areas of research. In the present study, we consider the case in which an expert in first aid treatment at a remote critical care center receives an image from an assistant/cooperator who is with a patient, i.e., the sharing of visual information between an expert and an assistant. A head-mounted camera worn by the assistant is used to capture visual information related to the patient and his/her surroundings. The expert directs the assistant to provide the desired visual information. Information of the status of the patient, the surrounding environment, and the situation can be obtained from the shared visual information. The assistant should make observations according to the directions of the expert. However, these persons are not required to be tightly linked/synchronized at all times, but only when necessary. There are primarily two behaviors in visual information sharing, namely, searching and staring behaviors. The proposed asynchronous visual information sharing system with image stabilization can separate these two behaviors and can reduce the stress of both the expert and the assistant in visual information sharing. This system consists of three components: image stabilization, virtual viewing direction operation, and relative orientation display. We experimentally evaluate the performance of virtual viewing direction operation, image stabilization, and asynchronous visual information sharing of the proposed system.**

I. INTRODUCTION

With recent advances in wearable computing, wearable virtual reality (VR) and wearable robotics are popular areas of research [1]. Maeda *et al.* developed the Parasitic Humanoid (PH) system. Even though the PH system does not include any actuators, it can guide the wearer of the PH system to behave in desirable ways by galvanic vestibular stimulation (GVS), a wearable moment display, and augmented reality (AR) technology using a head-mounted display (HMD) that incorporates cameras [2][3][4][5]. We have proposed and developed the Wearable Behavior Navigation System (WBNS), which is based on augmented reality (AR) technology, to realize general human behavior navigation [6].

Sharing information with others would be useful in several situations, and this sharing should be performed without generating discomfort or stress. For example when an individual who is in a steady state is watching a shared image that is trembling, a contradiction would exist between the received visual information and the somesthetic senses of the individual, which would cause motion sickness [7]. This type of contradiction between visual information and somesthetic senses can be reduced by image stabilization [8].

In the present study, we consider the case in which an expert in first aid treatment at a remote critical care center receives an image from an assistant/cooperator who is near a patient. In other words, this case involves the sharing of visual information between an expert and an assistant. A head-mounted camera worn by the assistant is used to capture visual information related to the patient. The expert directs the assistant to provide the visual information that is of interest to the expert. Information on the status of the patient, the environment, and the overall situation can be obtained from the shared visual information.

Sharing of visual information between an expert and an assistant is usually synchronized. In other words, the same image is provided to both the expert and the assistant. Since visual information is captured by a head-mounted camera worn by the assistant, the expert should direct the observations of the assistant. The expert and the assistant can be said to be tightly linked through visual information. The visual information received by the expert and the assistant is constrained because of this tight link and both the expert and the assistant would suffer from stress. This linkage can cause problems. For example, it is difficult for the expert to intentionally change the viewing direction because the viewing direction can be changed by the assistant. It is difficult for the assistant to quickly change the viewing direction as requested by the expert. Limitations in information processing and communication would also cause a time delay in remote instruction. The assistant would be exhausted because he/she would have to follow the viewing directions given by the expert at all times.

Although, in normal visual information sharing, the assistant should perform observation based on directions given by the expert, the expert and the assistant need not be tightly linked/synchronized at all times, but only when necessary. Visual information sharing consists of primarily two behaviors, namely, searching and staring behaviors. In ordinary visual information sharing, these two behaviors are not separated. Even though the expert and the assistant should look in the same viewing direction during staring behavior, it is not necessary for both the expert and the assistant look in the same viewing direction during searching behavior. In other words, synchronization is necessary for staring behavior but not for searching behavior.

We have developed an asynchronous visual information sharing system that can separate searching and staring behaviors during visual information sharing. Image stabilization,

This work was supported by CREST of JST.

N. Shiroma is with the Faculty of Intelligent Systems Engineering, Ibaraki University, Hitachi, Ibaraki 316-8511, Japan naoji@mx.ibaraki.ac.jp

E. Oyama is with Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8568, Japan eimei.oyama@aist.go.jp

virtual viewing direction operation, and relative orientation display are the three primary components of the proposed system. Finally, we experimentally evaluate the performance of the virtual viewing direction operation, the image stabilization, and the asynchronous visual information sharing of the proposed system.

The remainder of the present paper is organized as follows. In Section II, we describe the asynchronous visual information sharing system and the wearable behavior navigation system. In Section III, the system configuration is presented. Section IV describes experiments to evaluate the proposed system. Finally, the paper is summarized in Section V.

II. Asynchronous visual information sharing system

Performing general human behavior as directed by an expert is expected to be realized as wearable computing advances. Although ordinary human behaviors generally do not require navigation by/assistance of other individuals, in the case of administering first aid treatment, for example, an individual with an injured/ill person can provide first aid treatment with the aid of an expert.

In order to realize general human behavior navigation, we have proposed and developed the Wearable Behavior Navigation System (WBNS), which is based on augmented reality (AR) technology [6]. The camera on the HMD worn by the assistant/cooperator (head-mounted camera) can capture the visual information provided by the assistant. The image acquired by the camera is sent to an expert at a remote site, such as an emergency room or a critical care center. The expert can determine the status of the injured/ill person using the sensory information provided by the assistant. One important technology incorporated into the WBNS is the display system, which superimposes a computer generated image of the expert over the real images captured by the cameras on the HMD of the assistant. The expert and the assistant see both the computer generated image of the expert and the real camera image of the assistant on the screens of the HMDs. The assistant can then directly mimic the movements of the expert. The expert can induce the assistant to move in a prescribed manner by having the assistant mimic the movements of the expert. A head-mounted camera plays an important roll in human behavior navigation, and image stabilization systems [8] will be useful for visual information sharing because the camera moves and trembles in response to the movements of the assistant.

There are primarily two types of behavior in visual information sharing, namely, searching and staring behaviors. Searching behavior involves searching for an interesting region in the surroundings, and staring behavior involves looking closely at an interesting region in the surroundings. Although it is better that the viewing directions of both the expert and the assistant are synchronized during staring behavior, the viewing directions need not be synchronized during searching behavior. In other words, during searching behavior, the assistant does not need to track the viewing direction requested by the expert. Since ordinary visual information sharing cannot separate these two behaviors and the assistant must track the viewing direction requested by the expert at all times, both the expert and the assistant would be under considerable and constant stress. The proposed asynchronous visual information sharing system can separate these two behaviors and thereby reduce the stress involved in visual information sharing. Since, in the proposed system, the expert and the assistant can change their viewing directions independently during searching behavior, the expert can conduct searching behavior quickly and intentionally in a less stressful manner. Once the expert finds an interesting region in the surroundings, the relative orientation display in the proposed system can be used to synchronize the viewing directions of the expert and the assistant for staring behavior. In the proposed system, the assistant need only track the viewing direction requested by the expert in this synchronization.

The developed asynchronous visual information sharing system consists primarily of three components, image stabilization, virtual viewing direction operation, and relative orientation display. The assistant can change his/her viewing direction with less effect on the expert due to the image stabilization. The expert can change his/her viewing direction without any effect on the assistant by the virtual viewing direction operation. The expert and the assistant can understand the relative orientation differences between their viewing directions by the relative orientation display. In this section, we describe the three components of the proposed system.

A. Image stabilization

We incorporate an image stabilization method [8] in order to enable the assistant to change his/her viewing direction with less effect on the expert. Image stabilization consists primarily of two components: camera motion detection and motion compensation. As shown in **Fig. 1**, even when a camera is rotated about the camera center, an object that moves inside the image can be seen as long as the object is within the field of view of the camera. A small rectangular region around the object is selected in the image, and this region is shifted to follow the movement of the object when the camera is rotated. This image stabilization method is realized by sphere mapping of camera images and shifting of a selected rectangular region according to orientation information of a camera for rotational motion compensation.

1) Wide-viewing-angle image: In the proposed system, we used a single camera with a fish-eye lens. The wide-viewing-angle image obtained from the fish-eye-lens camera enables significant rotation of the camera. Although images obtained from the fish-eye-lens camera cover a wide area, there is significant distortion in the image, as shown in **Fig. 2**. This distortion is eliminated using sphere mapping, as explained in the next section.

2) Sphere mapping: The distorted image captured by the fish-eye-lens camera, which is obtained from the viewing position, is mapped onto the inside plane of the sphere

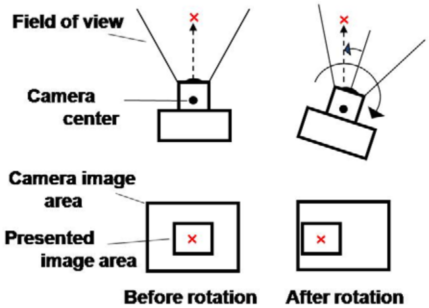


Fig. 1. Image stabilization method.

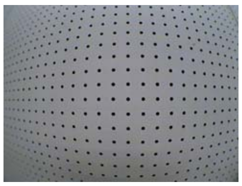


Fig. 2. Significant distortion occurs in images captured by a fish-eye-lens camera, especially around the image boundaries.

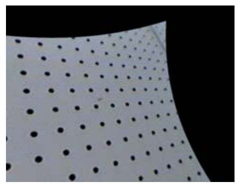


Fig. 3. A corner of an image captured by a fish-eye-lens camera after distortion elimination.

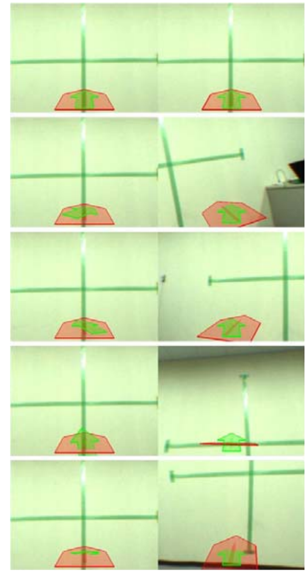
Fig. 4. Sphere mapping.

according to a fish-eye-lens model. We can obtain a non-distorted image, as shown in **Fig. 3**, by viewing an image mapped onto the sphere from the viewing position. In the present paper, as shown in **Fig. 4**, the information of each pixel of the image of a fish-eye-lens camera is projected onto the inner plane of the sphere. The fish-eye lens used in the proposed system can be modeled by equidistant projection [9]. We used OpenGL functions for image mapping and image extraction.

3) Image stabilization process: A small rectangular region around a point of gaze is selected from a recovered non-distorted image, and the selected region is shifted according to camera rotation for camera motion compensation. An inner image of the region is used as a stabilized image. In our implementation of image stabilization, the selected rectangular region around a point of gaze is fixed, and the sphere is rotated according to the detected camera orientation instead of shifting the region.

B. Virtual viewing direction operation

We incorporate the virtual viewing direction operation into the proposed system in order to enable the expert to change his/her viewing direction without affecting the assistant.



(a) Assistant view (b) Expert view

Fig. 5. Examples of the virtual viewing direction operation.

The virtual viewing direction operation is implemented by moving the stabilized gaze intentionally with the selected rectangular region. We use a fish-eye lens with a 180-deg FOV. The horizontal and vertical FOVs of an extracted rectangular region are approximately 60 deg and 45 deg, respectively. The virtual pan and tilt ranges are approximately $\pm$30 deg and $\pm$45 deg, respectively, and the virtual roll range is $\pm$180 deg within the FOV of the camera image. **Fig. 5** shows examples of the virtual viewing direction operation. The images in the left-hand column (assistant view) are the extracted original camera images for the assistant, and the images in the right-hand column (expert view) are the images for which the viewing directions are changed by the virtual viewing direction operation. The viewing direction can be changed by the expert moving his/her head, the orientation of which can be measured by a 3D motion sensor on the HMD of the expert. The top row shows the images for the initial view direction. The images in the second through fifth rows are the images obtained for the pan and tilt movements of the virtual gaze. Even though the camera of the assistant and the assistant view are fixed, the expert can change the viewing direction intentionally and can obtain visual information of the intended area.

C. Relative orientation display

The expert and the assistant can understand the relative orientation differences between their viewing directions by means of the relative orientation display. This information is also useful in determining the limitation of the FOV of the camera image. The relative orientation display shows the relative orientation between the expert and the assistant by superimposing the information graphically onto the provided images, as shown in **Fig. 5**. The superimposed arrows on the images represent the relative 3D orientation. The thin arrow indicates the relative orientation of the expert with

Fig. 6. NEC/TOKIN 3D motion sensor.

Fig. 7. CCD USB 2.0 camera Chameleon with a fish-eye lens.

Fig. 8. Head-mounted displays (left: expert, right: assistant).

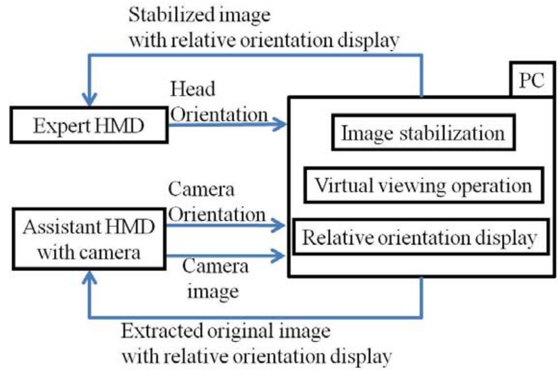


Fig. 9. System configuration.

respect to the assistant, and the fat arrow indicates the relative orientation of the assistant with respect to the expert. It is important to show the self-centered relative orientation for each person. Thus, the fat arrow, which, in the assistant view, represents the orientation of the assistant with respect to the expert, remains fixed, and only the thin arrow, which represents the orientation of the expert with respect to the assistant, moves according to the orientation of the expert (**Fig. 5** (a)). The same is true for the expert view, but in this case, the thin arrow, which represents the orientation of the expert with respect to the assistant, remains fixed, and only the fat arrow, which represents the orientation of the assistant with respect to the expert, moves according to the orientation of the assistant (**Fig. 5** (b)).

III. System configuration

A. Sensors

The NEC/TOKIN MDP-A3U9S-DK, shown in **Fig. 6**, is used for orientation detection. The Point Grey Research CCD USB 2.0 camera, Chameleon CMLN-13S2C, with a fish-eye lens, FUJINON YV2.2×1.4A-2, is used as the camera (**Fig. 7**). The horizontal field of view of the camera is approximately 180 deg.

B. Head-mounted displays

Two eMagin Z800 3DVISOR displays are used as head-mounted displays (**Fig. 8**). The display on the left is for the expert, and the display on the right is for the assistant. Each HMD has a 3D motion sensor to detect orientation. The assistant HMD has a camera with a fish-eye lens.

The system configuration is shown in **Fig. 9**. The figure shows the expert HMD, the assistant HMD, and the PC used in the system. The three-dimensional head orientation of the expert as detected by the 3D motion sensor is sent to the PC from the expert HMD. The three-dimensional head-mounted camera orientation and captured images are sent to the PC from the assistant HMD. The information is processed at the PC. The extracted original image augmented with the assistant-centered relative orientation between the viewing directions of the expert and the assistant is displayed on the assistant HMD (**Fig. 5** (a)), and the stabilized image augmented with the expert-centered relative orientation between the viewing directions of the expert and the assistant is displayed on the expert HMD (**Fig. 5** (b)). The superimposed relative orientation represented by the thin arrow (for the expert) and the fat arrow (for the assistant) on an image for the assistant HMD can be chosen to be displayed or to not be displayed by the expert, and the appearance of the arrows on the screen of the assistant HMD can be used as a trigger for synchronization, so that when the assistant sees the arrows, he/she tries to look in the direction requested by the expert.

The head-mounted camera orientation is changed by the action of the assistant for synchronization. The advantage of a head-mounted camera is that the camera can move according to the movement of the assistant and can provide almost the same image that the assistant is viewing. The expert can directly select a viewing direction by the virtual viewing direction operation. The assistant can use the relative orientation display to synchronize with the expert (look in the same direction as the expert).

The size of the extracted original and stabilized images is 640×480 pixels (45 deg vertical field of view), and the size of captured original images captured by the fish-eye-lens camera is 1280×960 pixels. The PC has a 2.53-GHz Intel Core2 Extreme CPU Q9300 (3 GB main memory with NVIDIA GeForce GTS 160M graphic chip).

IV. Experiments

Fig. 10 shows the experimental setup. In the figure, the assistant is wearing the HMD with the camera and the other individual is the expert. The assistant is looking at the cross on the wall. We conducted three experiments. Experiments 1 and 2 were conducted in order to evaluate the virtual viewing direction operation, and Experiment 3 was conducted in order to evaluate asynchronous visual information sharing.

In the synchronization experiments, the relative orientation display is not initially displayed on the screen of the assistant HMD, and the relative orientation display is displayed on the assistant screen at the request of the expert when the expert wants the assistant to look in the same direction.

(a) Expert and assistant with HMDs. (b) Assistant looking at a cross.

Fig. 10. Experimental setup.

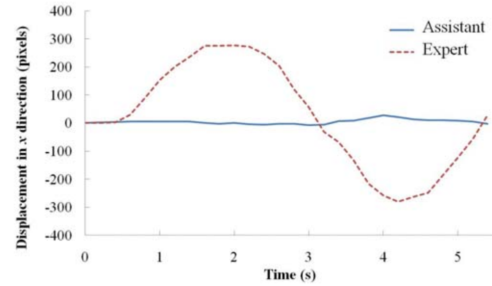


Fig. 11. Displacement in the *x*-direction for Experiment 1.

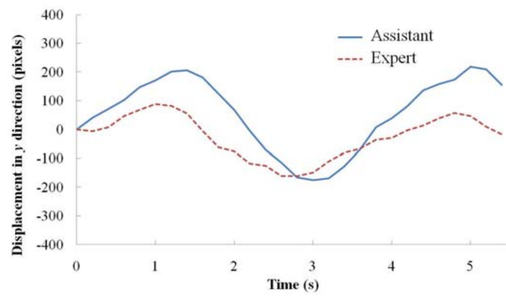


Fig. 12. Displacement in the *y*-direction for Experiment 1.

The assistant uses the appearance of the relative orientation display as a trigger for synchronization and uses the relative orientation display to follow the direction in which the expert is looking. This response to the trigger and the following motion for the viewing direction of the expert by the assistant cause a time delay in responding to the expert.

A. Evaluation of virtual viewing direction operation

In the experiments, the expert and the assistant cyclically move their heads in orthogonally different directions to demonstrate that the expert can change his/her viewing direction intentionally while suppressing the effects of the head movement of the assistant. In Experiment 1, the assistant performs a head tilting motion, while the expert performs a head panning motion by moving their respective gazes along the appropriate lines of the image of the cross. Experiment 2 is similar to Experiment 1, except that the moving directions of the heads of the expert and the assistant are reversed.

The results of Experiment 1 are shown in **Figs. 11** and **12** which are data plots of the displacement of the center of the cross in the *x*- and *y*- directions, respectively. The solid lines indicate the displacement in the images viewed by the assistant, and the dashed lines indicate the displacement in the images viewed by the expert. **Fig. 11** indicates that the expert can intentionally change his/her viewing direction by a head panning motion. The standard deviations of the assistant and the expert graphs in **Fig. 12** are 127.3 pixels and 75.5 pixels, respectively, and the effects of the head tilting motion of the assistant are suppressed in the images viewed by the expert.

The results of Experiment 2 are shown in **Figs. 13** and **14** which are data plots of displacements of the center of the cross in the *x*- and *y*- directions, respectively. The solid lines indicate the displacement in the images viewed by the assistant, and the dashed lines indicate the displacement in the images viewed by the expert. **Fig. 14** indicates that the expert can intentionally change his/her viewing direction by a head tilting motion. The standard deviations of the assistant and the expert graphs in **Fig. 13** are 159.0 pixels and 40.4 pixels, respectively, and the effects of the head panning motion of the assistant are suppressed in the images viewed by the expert. **Figs. 12** and **13** reveal a number of effects caused by the assistant motion. This is because the camera motion generated by the panning and tilting motions of the assistant cause not only camera rotation but also camera translation. The effects of the assistant motion are suppressed by the proposed method.

B. Evaluation of asynchronous visual information sharing

Experiments are conducted to evaluate asynchronous visual information sharing. We consider the task in which the expert looks around the environment surrounding the assistant and directs the assistant to look at a point of interest in the environment.

In the experiments, we compare two cases, namely, synchronous and asynchronous viewing direction tracking. In Experiment 3, the task consists of searching and staring behaviors executed by both synchronous and asynchronous tracking of the viewing direction of the expert in order to determine the effectiveness of task execution by separating visual information sharing into searching behavior and staring behavior. In synchronous tracking, the expert and the assistant have tightly linked visual information. In other words, there is no separation of visual information sharing, and the expert can only see the images that the assistant is viewing. The expert directs the assistant to look at a desired region using the relative orientation display. In this case, the assistant tries to align his/her viewing direction

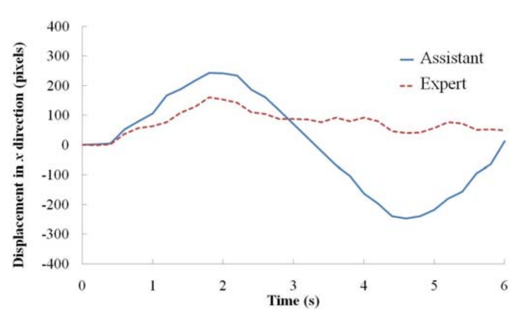


Fig. 13. Displacement in the *x*-direction for Experiment 2.

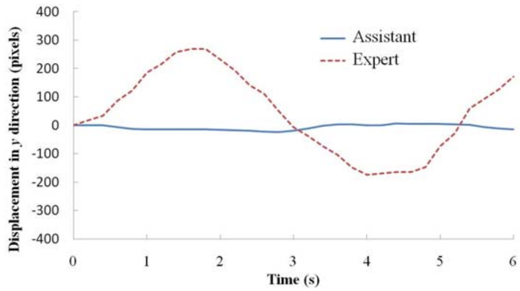


Fig. 14. Displacement in the *y*-direction for Experiment 2.

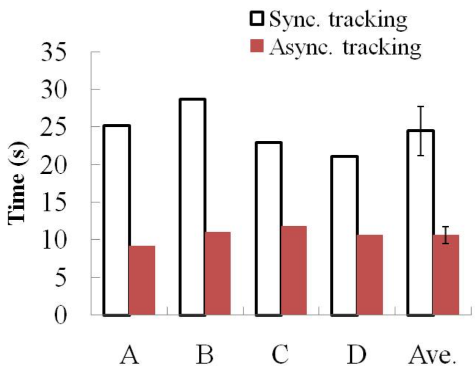


Fig. 15. Results of synchronous and asynchronous tracking.

with that of the expert at all times using the information obtained from the relative orientation display, and the expert can only change his/her viewing direction when the assistant changes his/her viewing direction. In asynchronous tracking, the expert can select the viewing direction intentionally without considering the assistant. In other words, there is separation of visual information sharing. First, the relative orientation is not shown on the images. Once the expert finds an interesting region in the surroundings, synchronization of viewing directions of the expert and the assistant for staring behavior is performed by presenting their relative orientations. In this case, the assistant need only align his/her viewing direction with that of the expert when the relative orientation between the assistant and the expert is shown on the images, and the assistant can concentrate on whatever he/she want to for the rest of the time. In the experiment, random combinations of panning and tilting motions were performed as a searching behavior, and after the searching behavior, one end of the cross is randomly selected as the focus of the staring behavior.

The results of the experiment are shown in **Fig. 15**. In the figure, the light and dark bins represent the amounts of time required for task execution under synchronous and asynchronous tracking, respectively. The results for four subjects and their average are shown. The error bars in the figure represent the standard deviation. Based on **Fig. 15**, the asynchronous tracking enables the task to be performed quickly, as compared to the synchronous tracking. The reason for the time differences stems primarily from the amount of time required for the searching behavior. The searching behavior is highly influenced by the tracking performance of the assistant as requested by the expert during synchronous tracking, whereas the expert can change his/her viewing direction at will during asynchronous tracking. Not only does the asynchronous tracking have better performance but it also reduces the stress on both the expert and the assistant because there is no tight link/synchronization between the expert and the assistant, and they can change their viewing directions at will most of the time, except when the expert wants the assistant to look at a region of interest. This advantage of the proposed system is significant when the expert and the assistant are performing a lengthy task. The experimental results demonstrate the effectiveness of the proposed asynchronous visual information sharing system.

V. Conclusions

We have developed an asynchronous visual information sharing system with image stabilization. The proposed system consists primarily of three components: image stabilization, virtual viewing direction operation, and relative orientation display. The proposed system uses a 3D motion sensor for camera rotation and detection of the head motion of the expert. Experimental results demonstrate the effectiveness of the proposed system, and less stressful and more efficient visual information sharing between the expert and the assistant is realized. A wider FOV image obtained by multiple cameras or by pan-tilt camera actuation will increase the applicable area of the proposed system, and we intend to implement this improvement in the proposed system in the future.

References

[1] M. Hirose, K. Hirota, T. Ogi, H. Yano, N. Kakehi, M. Saito and M. Nakashige, ”HapticGEAR: The development of a wearable force display system,” *IEEE Virtual Reality (IEEE VR)*, pp.123–129, 2001.

[2] T. Maeda, H. Ando, M. Sugimoto, J. Watanabe and T. Miki, ”Wearable robotics as a behavioral interface -the study of the parasitic humanoid,” *Proc. of 6th International Symposium on Wearable Computers*, pp.145–151, 2002.

[3] H. Ando, M. Sugimoto and T. Maeda, ”Wearable moment display device for nonverbal communications,” *IEICE Trans. Inf. & Syst.*, Vol. E87-D, No. 6, pp.1354–1360, 2004.

[4] T. Maeda, H. Ando and M. Sugimoto, ”Virtual acceleration with galvanic vestibular stimulation in a virtual reality environment,” *Proc. of IEEE VR 2005*, pp.289–290, 2005.

[5] T. Amemiya and T. Maeda, ”Asymmetric oscillation distorts the perceived heaviness of handheld objects,” *IEEE Transactions on Haptics archive*, Vol. 1, No. 1, pp.9–18, 2008.

[6] E. Oyama, N. Watanabe, H. Mikado, H. Araoka, J. Uchida, T. Omori, K. Shinoda, I. Noda, N. Shiroma, A. Agah, K. Hamada, T. Yonemura, H. Ando, D. Kondo and T. Maeda, ”A Study on Wearable Behavior Navigation System -Development of Simple Parasitic Humanoid System -”, *Proc. of IEEE International Conference on Robotics and Automation*, Anchorage, USA, 2010.

[7] R. Kurazume and S. Hirose, ”Development of image stabilization system for remote operation of walking robots,” *Proc. of IEEE International Conference on Robotics and Automation*, pp.1856–1861, 2000.

[8] N. Shiroma, J. Kobayashi and E. Oyama, ”Compact Image Stabilization System for Small-sized Humanoid,” *Proc. of IEEE International Conference on Robotics and Biomimetics*, Bangkok, Thailand February 2009.

[9] Nikon Corporation, available from http://imaging.nikon.com/products/imaging/technology/nikkor/n06_e.htm, (accessed 2010-07-07).