

Tangible Tele-Meeting in Tangible Space Initiative

Joong-Jae Lee*, Hyun-Jin Lee**, Mun-Ho Jeong[†], SeongWon Jeong***
and Bum-Jae You[§]

Abstract – Tangible Space Initiative (TSI) is a new framework that can provide a more natural and intuitive Human Computer Interface for users. This is composed of three cooperative components: a Tangible Interface, Responsive Cyber Space, and Tangible Agent. In this paper we present a Tangible Tele-Meeting system in TSI, which allows people to communicate with each other without any spatial limitation. In addition, we introduce a method for registering a Tangible Avatar with a Tangible Agent. The suggested method is based on relative pose estimation between the user and the Tangible Agent. Experimental results show that the user can experience an interaction environment that is more natural and intelligent than that provided by conventional tele-meeting systems.

Keywords: Tangible Tele-Meeting (TTM), Tangible Space Initiative (TSI), Tangible Space (TS), Tangible Interface (TI), Responsive Cyber Space (RCS), Tangible Agent (TA), Tangible Avatar (TAV), Registration

1. Introduction

Massive progress in internet and digital technologies has allowed us to experience new paradigms that were previously unimaginable. Among them, virtual space or cyberspace is capable of creating new value and has been expanded as a concept supplementary to the real world. In most cases, Virtual Reality (VR) is used with a Virtual Environment (VE), an artificial environment that is generated and maintained by a computer [1]. Although it enables a user to interact with a computer-generated environment and feel the effects of being placed in another location, it has some elementary problems, including latency, mismatching sensory information, and update frequency. Tangible space (TS) is a technology to overcome the spatial restriction of the real world and to extend the living space of humans and provide can feel seamless integration between real space and cyberspace. TSI is a new initiative to explore novel integration framework for next generation Human Computer Interaction (HCI) issues. TSI is inspired by comprehensive areas of research such as Augmented Reality (AR), Telerobotics, Telepresence, Ubiquitous computing, and HCI.

• *Augmented Reality*

AR is a combination of real world and virtual objects employed to provide a natural interface through which virtual objects appear to coexist with the real world [2]. Diverse applications for AR includes computer-aided surgery, special effects for TV or film, tour guidance services, industrial plant maintenance, interior design, medical education, and so on [3, 4].

• *Ubiquitous computing*

Ubiquitous computing is a paradigm for human interaction with computers in conjunction with virtually every object and activity we do. One of the attractions of ubiquitous computing is context awareness which deals with linking changes of computer systems and supports computer usage in varying physical environments. Furthermore, this technology actively appears in mobile context-aware applications [5, 6].

• *Telepresence*

Telepresence allows a user to feel as if he or she were present. The distinction between telepresence and virtual reality is that a user feels a sense of presence in the former. For example, B. Lei and Hendriks presented a system, VIRTUE (Virtual Team User Environment), which integrates a real and virtual world environment with high-quality 3-D telepresence [7]. One of the aims of this system is to allow three-way telepresence videoconferencing, providing distinctive presence features and experience for the conference participants in contrast with the traditional videoconferencing system.

• *Telerobotics*

Telerobotics is a combination of telepresence and

[†] Corresponding Author: School of Robotics, Kwangwoon University, Korea. (mhjeong@kw.re.kr)

* Center of Human-centered Interaction for Coexistence (CHIC), Korea. (arbitlee@chic.re.kr)

** Advanced R&D Group, Security Solution Division, SAMSUNG TECHWIN Co. Ltd., Korea. (hyunjin7740.lee@samsung.co.kr)

*** Department of Product Design and Manufacturing Engineering, Seoul National University of Science & Technology. (swjeong@seoultech.ac.kr)

§ Interaction and Robotics Research Center, Korea Institute of Science and Technology. (ybj@kist.re.kr)

Received: September 12, 2013; Accepted: December 9, 2013

teleoperation, and involves the control of robots from a remote location. In other words, a telerobotic system provides a framework for connecting humans and robots to reproduce operator actions at a distance. There are numerous research fields for telerobotics such as remote manipulation in space investigation [8], underwater services, telesurgery, minimum invasive surgery, etc.

• **Human Computer Interaction**

HCI is a research area of interaction between users and computers. For instance, eye tracking is a visual attentive interface that focuses on human attention as the crucial input to computers. It could be used to continuously update eye movement to identify what a person is looking at [9].

Several attempts have been proposed in the literature to develop a system that allows people to communicate with each other without any spatial limitations.

TELEPORT is a prototype of an immersive teleconferencing environment that enables participants to meet face-to-face virtually [10]. A drawback of this system is to provide communication between real humans and only one person in a location can have his or her own viewpoint.

Tangible Bits is an attempt to bridge the gap between cyberspace and the physical environment by making digital information tangible [11]. However, current HCI research including Tangible Bits focuses primarily on how to enable users to feel real sensations only for the information stored in computer memory in advance and/or generated by cyberspace during operation. The main difference between Tangible Space and Tangible Bits including AR research is that we continuously augment a newly explored tele-existed real world.

Tachi proposed a teleoperation system wherein an operator can experience the feeling of existing in a surrogate robot [12]. While this system enables the operator to feel mutual telexistence, it is still unnatural to human beings because the appearance of the surrogate robot and the human operator is different.

TS is a technology to overcome the limitations described above and to extend the living space of human beings in order to provide seamless integration between real space and cyberspace.

We designed a scenario framework of TSI experience, a Tangible Tele-Meeting system (TTM), which allows people to join and enjoy the human to human remote communication naturally. When communicating with the others using the TTM, the user can utilize multi-modal senses such as voice, the sense of sight, and the sense of touch as well as surroundings. Since the TTM successfully overcomes some restrictions of virtual reality system, it can be more natural and effective than existing meeting systems mentioned before. The user wears only see-through-HMD and sees a Tangible Avatar which looks like a human in appearance. In addition, the user can even touch the body of a TA with the help of a humanoid robot.

In this paper, we describe a new frame work, TSI, which

can provide a more natural and intuitive Human Computer Interface. In this framework, we present a TTM system that can provide more natural and effective communication between a human user and a physical avatar. We also introduce a method for registering a TAV with a TA, which is based on relative pose estimation between the user and the TA.

2. Tangible Space Initiative

TSI is a novel integration framework that enables a user to interact with TS that is located in the physical interaction environment. TS is defined as a space in which people can feel seamless integration between the real world and cyberspace. A user may feel immerse in the TS as if he or she were there, even for a newly explored real world. Moreover, it enables users to touch and manipulate objects generated by cyberspace. The TSI involves a variety of technologies such as virtual reality, augmented or mixed reality, robotics, image processing, intelligent control, multimedia database and artificial intelligence. The common goal of these technologies is to reduce the hindrances between human and computer generated virtual environments.

Three major technologies for the TS are TI Technology, RCS Technology, and TA Technology, as shown Fig. 1.

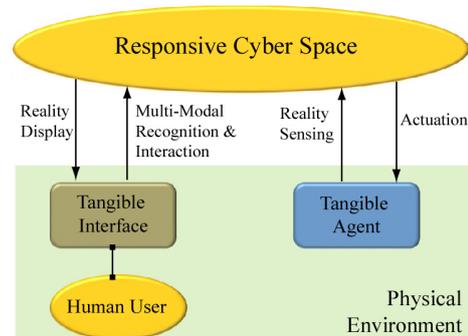


Fig. 1. Conceptual diagram of TS

2.1 Tangible interface

TI allows users to manipulate virtual objects and synchronize the real world and the virtual environment. Accordingly, the TI should provide reality and naturalness to users. For reality and naturalness of a visual interface, it is necessary to provide the feeling of immersion and natural interaction with displayed objects. The interactive reality-imaging interface provides these aspects by incorporating an immersive large-scale display and intelligent interaction channels. In TSI, a tiled display system is used as the immersive large-scale display, which incorporates vision-based techniques. In contrast to other previous tilted display systems, the interactive reality-imaging interface exploits an independent warping hardware between image generating computers and

projectors, which warps images according to screen shapes and user interaction. Furthermore, the TI includes haptic and/or the Perceptual User Interface (PUI) between users and responsive cyber space as the interaction channels. As shown in Fig. 1, the PUI is characterized by interaction techniques that combine multi-modal computer I/O devices and machine perception of human behavior (e.g. face and gesture recognition).

2.2 Responsive cyber space

RCS comprises all the virtual objects included in the virtual environment and their behaviors. RCS is responsible to render and display virtual or virtual-real visual / haptic feedback images to the user by TI. RCS returns a response directly or indirectly when receiving the user input from TI. The direct response may influence changes in the visual/haptic image while the indirect response can affect changes in the physical agent of the virtual environment using TA's actuation. Furthermore, RCS is able to monitor the state or situation of a user and his or her surrounding environment. Since RCS is proactively responding to the user's need and providing the required information, the user may not explicitly request his/her need to the system. Thus, the responsiveness of RCS depends on the context of different activities and is triggered only when the corresponding context is obtained from the user.

2.3 Tangible agent

TA is an interface between RCS and the real world. TA is an agent capable of (1) sensing the real world and augmenting the sensed environments to the predefined cyberspace and/or giving information to generate a new user cyberspace, and (2) navigating the real world to sense it and/or interacting with the real world to do the user-specified job or gathering tactile information.

The definition of TA is similar to that of Autonomous Agent (AA), which is defined as a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to affect what it senses in the future.

Typical characteristics of an AA are as follows: (1) to perceive and interpret sensor data, (2) reflect events in their environment, and (3) to take actions for achieving given goals. Here, a common characteristic to both TA and AA is reactivity, which means sensing and acting capability. However, TA is clearly distinguished from other AA in the sense that the TA should provide its own reactive function while providing users with seamless integration.

3. Tangible Tele-Meeting

TTM is one of the attractive applications of TS technology.

The goal of this application is to eliminate spatial limitations in the real world and support effective communication between the real world and cyberspace. Hence, once the versatility of TS technology is available, two people in different remote locations will be able to meet live, share situational information, and interact in a lively manner with each other in the given situation.

As mentioned before, the main idea of TTM is to provide more natural and effective communication between a human user and a physical avatar using AR technologies. Additionally, TTM uses a humanoid robot as a TAV, which enables users to experience more realistic interactions. In this paper, we describe the development of a real time HMD (Head Mounting Device) display of avatar image registration for the TA.

Fig. 2 shows the overall architecture of TTM. The configuration of TTM is composed of two different locations: a local location (Location A) and a remote location (Location B).

Fig. 3 shows how users interact with each other via TTM in more detail. The user in the location A can see another user through the HMD which displays 3D video avatar images for a participant in location B. In particular, we use marker-based visual tracking and real-time 3D motion information in order to align the avatar images with the TA. Further details of avatar image registration for the TA are presented in Section 4. Meanwhile, the user in location B is

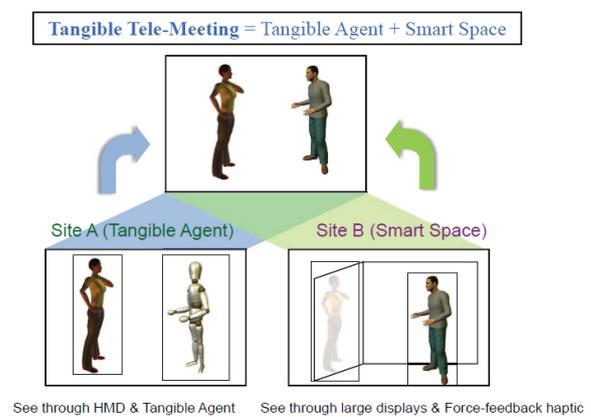


Fig. 2. Overall Architecture of TTM

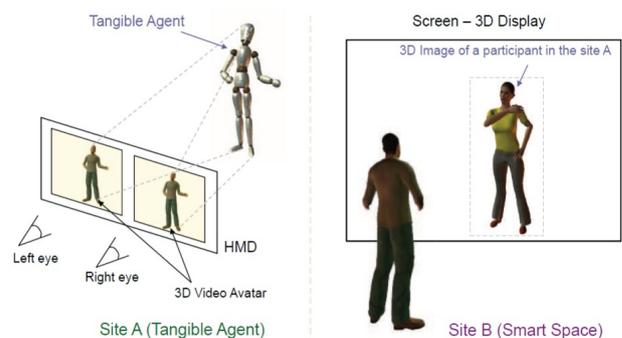


Fig. 3. Interaction between users via TTM

Table 1. Information for TTM between Location A and Location B

Direction	Information	Device & Object
Location A →	Scene model information	HMD
	Image/Force of the user	Tangible Avatar
Location B →	Virtual camera information	Network Camera Indoor GPS
	3D video avatar images for a user	Large Display
Location B →	Real-time 3D motion information	Haptic Glove
	Force information for a user	Motion Capture Network Camera

able to see 3D images of location A projected on the large display system and touch the participant or the objects in location A using a haptic glove.

The detailed descriptions for Location A and Location B are as follows:

• **Location A**

A human user in local physical space (Location A) wears the HMD and meets another user in location B through a display system. The TAV which takes the place of physical contact with the remote user, appears in front of the local user and conveys physical senses of touch such as handshaking, hugging, and touching. However, the local user who wears the HMD only sees the appearance of the remote user by means of the display system at this time. The results of a real-time human motion analysis using 3D cameras equipped at the TA and its surrounding environment are reflected on the display system of the remote location.

• **Location B**

In the remote physical space (Location B), a remote user who wears the HMD on his or her head and haptic devices sensible to physical contact meets the other user in location A through a large display system. With the surrounding environment equipped with a 3D camera system and a motion capture system, human motions are recognized in real-time and then transferred to the TAV or the display system in the remote location.

Table 1 lists the information for TTM transmitted between Location A and Location B.

4. Registration of a Tangible Avatar and a Tangible Agent

As noted earlier, we focus on the development of a real time HMD display of avatar image registration for a TA among various technologies to implement TTM. In AR, the registration process is defined as the precise alignment and synchronization of virtual objects with the real world [13].

Similarly, from the perspective of TTM, the targets to be properly registered are the TAV and the TA. The result of the registration is tightly linked to the quality of a user’s

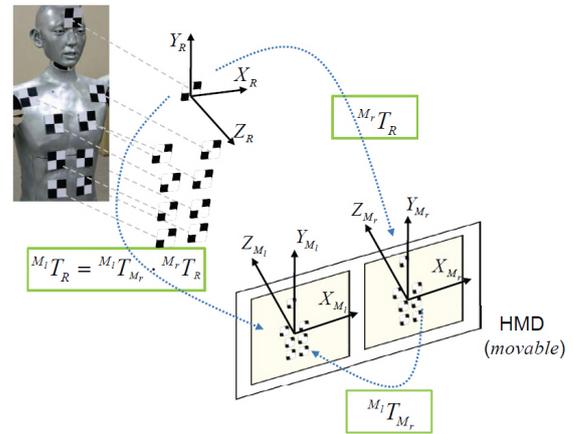


Fig. 4. Pose estimation with relevant coordinate frames and transformations

experience for TTM. When the user sees the TAV from different views or the TA is moved according to 3D motion information, registration should be performed on the TAV seamlessly. For this purpose, robust pose estimation is necessary to achieve the accuracy required for augmenting the real world. Pose estimation is the process of estimating the rigid transformation that aligns the model frame with the world reference frame. In this work, the relative pose is an estimate of the translation and rotation of the model (TA) to the camera frame (HMD).

Fig. 4 shows the relevant coordinate frames and transformation for estimating the relative pose between the TA and the HMD.

The pose estimation is composed of two steps: initial pose estimation and pose tracking. The first step is initial pose estimation, which is required to initialize the pose tracking step. In order to obtain a priori knowledge of the TA whose pose and motion is to be estimated, we compute initial 3D model coordinates using stereo triangulation [14]. To compute the initial estimate of relative pose, we detect marker positions attached on the TA and then compute the initial estimate of pose between the TA and the HMD using Tsai’s calibration method [15].

When markers have been identified on the 2D image, we can estimate the HMD’s pose with respect to the markers on the TA. We use a method to automatically detect calibration markers on a chessboard [16]. This approach involves the use of two symmetric properties of a chessboard pattern related to geometric and brightness distribution and two concentric circles as a probe. In particular, it is comparatively robust to the change of pose, scale, and illumination. Thus, we adopt this approach towards solving robust marker detection. After completing the initial pose estimation, the pose tracking can be started through the initial pose passed to the IEKF. Initialization is also performed at the start of pose tracking or whenever loss of tracking occurs.

However, since detecting markers on every frame is computationally intractable in real-time, it is generally not

utilized in practical applications. Therefore, we employ an iterative extended Kalman filter (IEKF) to track the marker's position and orientation relative to the HMD over time, which provides robust predictions and estimates of pose. As known already, the EKF is an extension of the standard Kalman filter to a non-linear system, and addresses the problem of optimally estimating the state of the system. The IEKF iteratively estimates the state of the system to handle the significant error caused when the measurement function is highly nonlinear. The Kalman filter estimates the system state by using a form of feedback control. As such, the tasks of the Kalman filter fall into two phases: prediction step and update (correction) step. The prediction step is responsible for projecting forward the current state and obtaining a priori estimates of the state and measurement for the next time frame. The update step is responsible for feedback. That is, it incorporates an actual measurement into the a priori estimate to obtain an improved a posteriori estimate. The improved estimate in turn feeds into the prediction step and the prediction-update cycle is repeated. In the IEKF framework, the measurement process is implemented using an optical flow based on the iterative Lucas-Kanade method in pyramids [17]. This method provides the ability to find individual marker positions with sub-pixel accuracy.

$${}^G T_C = {}^G T_A \cdot {}^A T_R \cdot {}^R T_H \cdot {}^H T_C$$

where

$$\left\{ \begin{array}{l} {}^b T_a : \text{Transformation from a to b} \\ G : \text{Global coordinate frame} \\ A : \text{Avatar coordinate frame} \\ R : \text{Tangible Agent coordinate frame} \\ H : \text{see-through-HMD coordinate frame} \\ C : \text{Camera coordinate frame} \end{array} \right. \quad (1)$$

Once the relative pose estimate is computed by pose tracking, we can render a virtual object (TAV) onto real object (TA) images from given the viewpoint. For this purpose, five coordinate frames are defined as shown in Fig. 5: (1) see-through HMD coordinate frame in the real

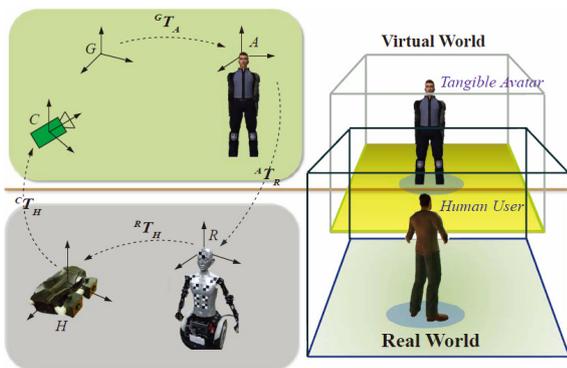


Fig. 5. Transformations for registering a TAV and TA

world, (2) TA (robot) coordinate frame in the real world, (3) camera coordinate frame in the virtual world, (4) global coordinate frame for the virtual world, and (5) TAV (virtual object) coordinate frame in the virtual world. The overall registration procedure can then be represented by the chain product of successive transformation matrices in Eq. (1). In particular, the 3D relative pose estimate from visual measurements is given by the transformation between the HMD coordinate and the TA coordinate frame. Finally, the TAV images are rendered to augment the TA along with the estimated pose over time.

Table 2 summarizes the overall procedure of registering the TAV and the TA.

Table 2. Complete algorithm for registering the TAV and the TA

We describe the following registration algorithm in the IEKF based framework:
1. Initialization
(a) Detect chessboard marker positions [16].
(b) Estimate an initial estimate of pose based on Tsai's calibration method [15].
(c) Compute initial 3D model coordinates using stereo triangulation [12].
2. Registration (Tracking and Rendering) in the IEKF framework
(a) Feature measurement – track marker positions based on optical flow techniques [17].
(b) Compute the change in pose for given marker positions.
(c) Check the total projection error for all marker positions.
If the error is greater than the threshold value, start the initialization procedure again.
(e) Update the current pose estimate and render the TAV onto the TA image

5. Experimental Results and Discussion

In this section, we present the experimental setup and discuss the results. The experiments were carried out on a Pentium 3.0 GHz with 1Gbyte RAM. As described in Sec.3, the TTM system is composed of two locations: a local location (Location **A**) and a remote location (Location **B**). We show that a user can experience an interaction environment in location **A** in which the user wears a HMD and meets another user in Location **B** through the TAV.

Fig. 6 shows the experimental setup for TTM. Here, a participant can experience more realistic interaction through a binocular HMD equipped with a stereo camera, as shown in Fig. 6(a). The TAV in Fig. 6(b) is a virtual (graphic) object that represents the effect of reaction, eye gaze, and body gestures of the other participant in a remote location. In particular, we show that a participant experiences a sense of presence in the real world in which he/she and the robot shake hands with each other. As shown in Fig. 6(c), the TA is a two-wheeled humanoid robot with 23 degrees of freedom (6DOF for each arm, 2DOF for neck 4DOF for each hand, and 1DOF for waist). Furthermore, 9 markers attached on the upper body of the robot are used to estimate the relative pose between the



Fig. 6. Experimental setup for TTM (Location A): (a) See-through HMD; (b) TAV; (c) TA; (d) Example of interaction with the participant and the TA

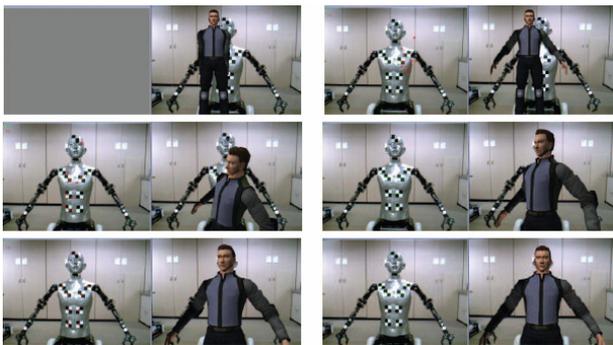


Fig. 7. Initial pose estimation of the TA to be tracked

user and the TA. Fig. 6(d) finally shows an example of interaction with the human user and the TA.

Fig. 7 shows the process of initialization for registering the TAV and the TA. When the user sees the TAV from different views or the TA is moved freely, registration should be performed on the TAV seamlessly. For this purpose, it should be noted that a good estimate of the initial pose has a negligible effect on the accuracy of pose estimation. As shown in Fig. 7, the initial pose of the TAV is refined over time and is deemed to have converged to the desired position according to suitable criteria.

Fig. 8 depicts the augmented scenes in the greeting sequence in which the TAV is registered to the TA along with the estimated pose. The registration method should provide spatially seamless augmented space where the participant, wearing a HMD, is located. In practice, the upper body of the TAV is naturally augmented to spatially registered locations of the TA. However, there are some seams in the hand regions of the TAV because the robot's



Fig. 8. Selected frames from augmented scenes projected on HMD (*greeting sequence*)

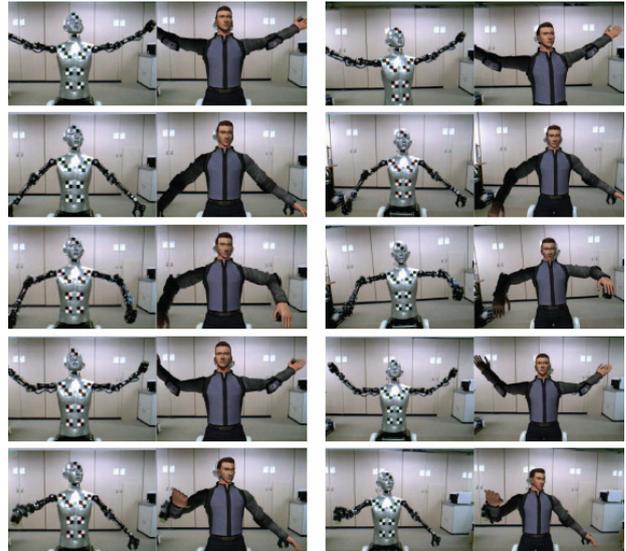


Fig. 9. Selected frames from the various gesture sequences. Both cases are considered: the camera is stationary (left column) and moving (right column).

hands are quite different compared to those of the human user.

We performed the registration process on the various gestures sequences in Fig. 9. Both cases are considered: the camera is stationary and moving. In the case of a moving camera in those sequences, we can identify various camera motions, such as zooming in/out, panning, tilting, and swing. It should be noted that the participant wearing the HMD can move freely in space. Considering the problem of a moving camera is therefore more practical than addressing only the case of a stationary camera.

As mentioned before, Fig. 9 shows the results of the registration process for each sequence. In both cases, the results show that our method is useful for registering the virtual (TAV) and physical (TA) objects. Notably, the registration process yields satisfactory results in spite of camera motions.

Figs. 10 and 11 respectively show the observed translation and orientation of the TA under stationary and moving camera motions. These sequences involved the loss of visual markers caused by rapid motion of the TA. The loss of visual markers is introduced between frames 581 and 832 and then the tracking is failed. In addition, registration errors cause misalignment between the TAV

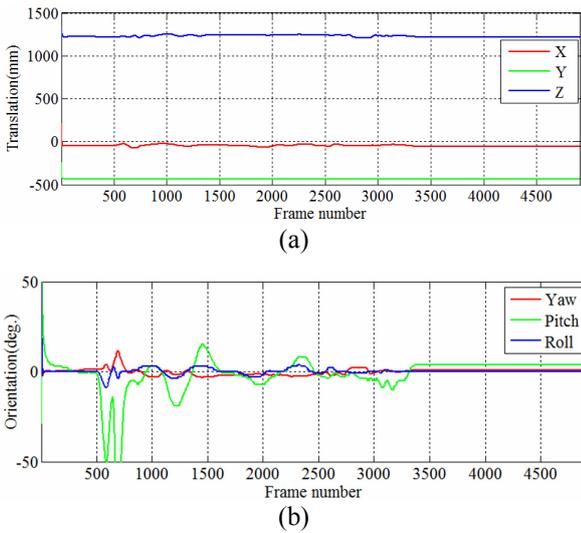


Fig. 10. Stationary camera case: observed translation and orientation of the tracked TA: (a) translation; (b) orientation

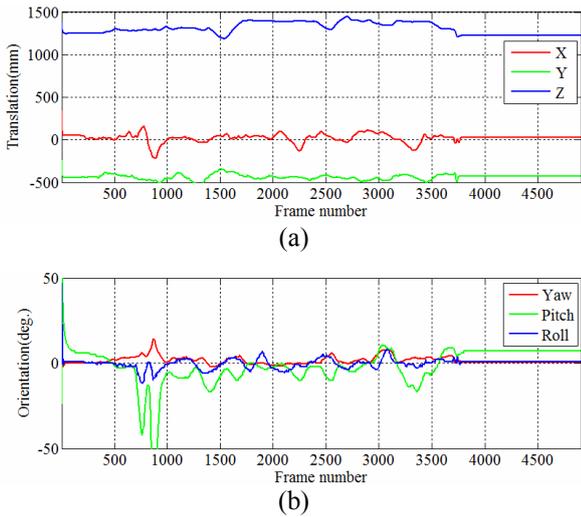


Fig. 11. Moving camera case: observed translation and orientation of the tracked TA: (a) translation; (b) orientation



Fig. 12. Failure of tracking markers due to rapid motion of the TA

and the TA, as shown in Fig. 12.

Below are some quantitative aspects of the experimental results.

The plot in Fig. 13 shows the observed and estimated x-coordinates of selected marker points for each frame in the

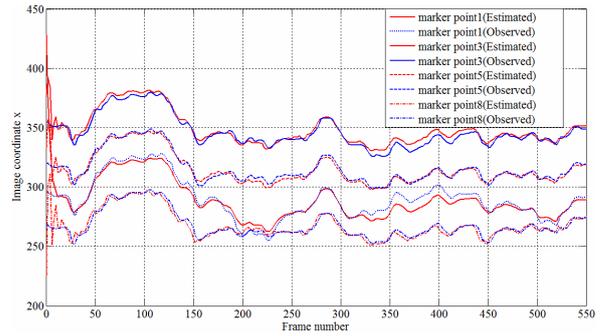


Fig. 13. Moving camera case: the observed and estimated x-coordinate of selected marker points (1, 3, 5, and 8).

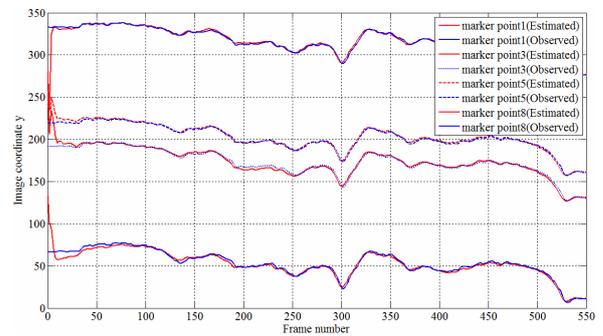


Fig. 14. Moving camera case: the observed and estimated y-coordinate of selected marker points (1, 3, 5, and 8).

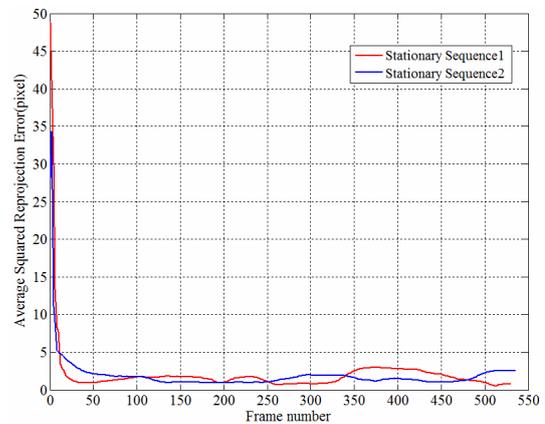


Fig. 15. Stationary camera case: average squared reprojection errors for each frame in the sequences

sequence. Although marker detection and association errors occasionally arise, the majority of points do not perturb the accuracy of the pose estimation.

Since the movement in y-direction is comparatively small, the accuracy of the y-coordinates estimated is more stable, as shown Fig. 14.

In the experiment, the accuracy of the registration process is evaluated by measuring the average squared reprojection error. The reprojection error is a geometric error corresponding to the image distance between a

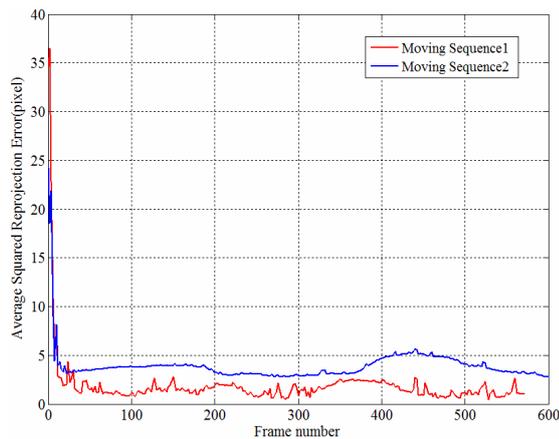


Fig. 16. Moving camera case: average squared reprojection errors for each frame in the sequences



Fig. 17. An example of a tangible interaction between the human user and the TAV (*handshaking*)

projected point and a measured one [18].

Fig. 15 shows that the average squared reprojection errors quickly converge to reasonable values. As expected, in the moving camera, the error value is a little bigger than one in the stationary camera, as shown Fig. 16. This might be introduced by rapid motion or change in direction of motion between frames 400 and 500 in the first sequence.

Despite these distractions, the experimental results demonstrate that the proposed method maintains the accuracy of the registration process.

Fig. 17 finally shows an example of a tangible interaction between the user and the TAV. This is a handshaking motion involving physical contact with the remote user. Ultimately, it enables users to experience more realistic interactions.

6. Conclusion

In this paper we described a new frame work, Tangible Space Initiative, which can provide more natural and intuitive Human Computer Interface. To achieve this goal, TSI incorporates comprehensive technologies such as Augmented Reality, Telerobotics, Telepresence, Ubiquitous computing, HCI, etc. Additionally, we presented a Tangible Tele-Meeting system that allows people to communicate with each other without any spatial limitation. We finally described a method for registering the Tangible Avatar with

the Tangible Agent, which is based on relative pose estimation between the participant and the Tangible Agent. Results from the experiments demonstrate that the participant can experience an interactive environment, which is more natural and intelligent than that provided by conventional tele-meeting systems.

Acknowledgements

This work is supported by the Research Grant of Kwangwoon University in 2012.

References

- [1] UVA User Interface Group, "Alice: Rapid prototyping for virtual reality," *IEEE Computer Graphics and Applications*, vol. 15, no. 3, pp. 8-11, 1995.
- [2] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Transactions on Information Systems*, E77-D (12), 1321-1329, 1994.
- [3] Wagner, D., Langlotz, T., and Schmalstieg, D., "Robust and unobtrusive marker tracking on mobile phones," *Mixed and Augmented Reality, ISMAR 2008*, pp. 121-124, 2008.
- [4] Aaron Kotranza, Benjamin Lok, Adeline Deladisma, Carla M. Pugh, and D. Scott Lind, "Mixed Reality Humans: Evaluating Behavior, Usability, and Acceptability," *IEEE Transaction on Visualization and Computer Graphics*, vol. 15, no. 3, pp. 369-382, 2009.
- [5] Anind K. Dey, Gregory D. Abowd, and Daniel Salber, "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototype of Context-Aware Applications," *Human-Computer Interaction*, vol. 16, pp. 97-166, 2001.
- [6] T. Selker, "Visual attentive interfaces," *BT Technology Journal*, vol.22, no.4, pp.146-150, 2004.
- [7] L.-Q Xu, B Lei and E Hendriks, "Computer Vision for a 3-D Visualisation and Telepresence Collaborative Working Environment," *BT Technology Journal*, vol. 20, no. 1, pp 64-74, 2002.
- [8] Dang Quoc Khanh and Young-Soo Suh, "Mobile Robot Destination Generation by Tracking a Remote Controller Using a Vision-aided Inertial Navigation Algorithm," *Journal of Electrical Engineering & Technology*, vol. 8, no. 3, pp. 613- 620, 2013.
- [9] Michael D. Fleetwood and Michael D. Byrne, "Modeling the Visual Search of Displays: A Revised ACT-R Model of Icon Search Based on Eye-Tracking Data," *Human-Computer Interaction*, vol. 21, no. 2, pp. 153- 197, 2006.
- [10] C J Breiteneder, S J Gibbs, and C Arapis, "TELEPORT - An Augmented Reality teleconferencing Environment," *Proceedings of 3th Euro-graphics Workshop on Virtual Environments*, pp. 41-49, 1996.
- [11] H. Ishii, and B Ullmer, "Tangible Bits: Towards

Seamless Interfaces between People, Bits and Atoms,” *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 234-241, 1997.

- [12] S. Tachi, “Recent Advance in Telexistence-Retro-reflective Projection Technology based Mutual Telexistence,” *International Symposium on Measurement and Control in Robotics*, 2003.
- [13] Ronald T. Azuma, “Making Direct Manipulation Work in Virtual Reality,” *SIGGRAPH 97 Course Notes*. August 1997.
- [14] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, 2002.
- [15] R. Y. Tsai, “A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses,” *IEEE Journal of Robotics and Automation*, vol. RA(3), no. 4, pp. 323-244, 1987.
- [16] Jong-Eun Ha, “Automatic detection of calibration markers on a chessboard,” *Optical Engineering*, vol. 46, no. 10, 2007.
- [17] Jean Y Bouguet, “Pyramidal Implementation of the Lucas Kanade Feature Tracker: Description of the algorithm,” *Intel Corporation Microprocessor Research Lab*, 2002.
- [18] Richard Hartley and Andrew Zisserman, *Multiple View Geometry in computer vision*, Cambridge University Press, 2003.



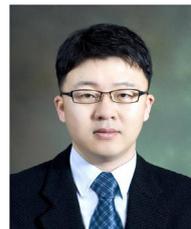
Joong-Jae Lee received the B.S, M.S, Ph.D. degrees in Computer Science from Soongsil University, Korea, in 1999, 2001, and 2005, respectively. He is a senior researcher in Center for Human-centered Interaction for Co-existence. His research interests include computer / robot vision, ADAS, augmented reality, and human robot interaction.



Hyun-Jin Lee received the M.S. degree from Han-Yang University, Seoul, Korea in 2008, in Electronics and Computer Engineering. In addition, he was student researcher in Center for Cognitive Robotics Research, Korea Institute of Science and Technology (KIST), Seoul, Korea. In September 2008, he joined the Samsung Techwin Security Solution Advanced R&D Group, Seongnam, Korea as a member of Researcher. His research interests include the joint area of 3D computer vision, computer graphics, and multimedia application, and motion detect and motion tracking for surveillance system.



Mun-Ho Jeong the B.S. and M.S. degrees from Korea Advanced Institute and Science and Technology (KAIST), Korea and his Ph.D. from the Department of Mechanical Engineering for Computer controlled Systems, Osaka University, Japan. He has been a professor at the School of Robotics in Kwangwoon University since 2010. His major interests are robot vision and intelligence, image processing, and human-robot interaction.



SeongWon Jeong is a professor at the department of Product Design and Manufacturing Engineering in Seoul National University of Science & Technology. His research interests include robot design, emotional design, and human-robot interaction. He is a regular member of Korean Society of Design Science (KSDS), Korean Society for Emotion & Sensibility (KSES). He had worked as a product designer for the major company in Korea, Samsung Heavy Industries for several years. He received the B.S. and Ph.D. degrees from the department of Industrial Design in Korea Advanced Institute and Science and Technology (KAIST).



Bum-Jae You received the B.S. degree in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1985 and the M.S. and Ph.D. degrees in electrical and electronic engineering from the Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1987 and 1991, respectively. From 1991 to 1994, he was the Head of the Robotics Division, Turbo-Tek Company, Ltd. In 1994, he joined the Korea Institute of Science and Technology (KIST), Seoul. From 2004 to 2011, he was the Director of the Center for Cognitive Robotics Research. Since 2010, he has been a Director with the Center of Human-centered Interaction for Coexistence and a Principal Research Scientist of Interaction and Robotics Research Center, KIST. His research interests include coexistent interaction, humanoids, vision-based robotics, real-time computer vision, and digital signal processor applications.