Coexistent Space: Toward Seamless Integration of Real, Virtual, and Remote Worlds for 4D+ Interpersonal Interaction and Collaboration

Bum-Jae You^{*1,2}, Jounghuem R. Kwon¹, Sang-Hun Nam¹, Jung-Jea Lee¹, Kwang-Kyu Lee¹ and Kiwon Yeom¹

¹Center of Human-centered Interaction for Coexistence

²Korea Institute of Science and Technology

Abstract

Three worlds are integral to our daily life: the real world, virtual world, and remote world. In the paper, there is proposed coexistent space where networked users can communicate, interact, and collaborate together by exchanging 4D+ sensation, human intension, and emotion. The 4D+ sensation includes 3D vision, 3D sound, force and torque, touch, and movements. The coexistent space is generated by seamless integration of real, virtual, and remote worlds while networked users experience the feeling of coexistence through 4D+ bi-directional interaction. Initial software framework and experimental results for interaction between multiple remote users are shown successfully.

CR Categories: I.3.3 [Computer Graphics]: Three-Dimensional Graphics and Realism—Display Algorithms I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Radiosity;

Keywords: radiosity, global illumination, constant time

1 Introduction

Futurologists predict the boundaries between human and machine and/or between reality and virtual reality will collapse by 2020, and the intelligence of machines will gradually begin to transcend human intelligence by 2030. The era of trans-human, replacement of human body parts with bionics, will come by 2040. In particular, information and communication technology in conjunction with software technology will provide the basis for technology fusion and innovation.

Three worlds are integral to our daily life: the real world, virtual world, and remote world. The real world is the space where we are physically present, such as the home or office. In this world there are many artifacts like networked printers, PCs, service robots, TVs, home appliances, prosthetics, motorized video cameras, etc. The virtual world refers to the cyberspace of computer systems such as those provided by computer games, PC user environments, cyber second lives, social network spaces like Facebook, Twitter, and KakaoTalk, etc. Finally, the remote world physically exists, but one cannot be there. For example, one cannot be in New York at this instant if you're living in Seoul. In the remote world there are also many artifacts as in the real world.

There have been proposed various concepts to connect the real world with the remote world [Lanier 2001]. Tele-immersion in

multi-user virtual environments (MUVE) was proposed for communication between remote users in [Lanier 2001] by using a pair of polarizing glasses for stereo vision effect, two walls of cubicle dissolved into windows, a camera system, a head-tracking device, and a pointing device to handle virtual objects. Immersive group-to-group telepresence was introduced by using 3D video wall, mobile Kinect cameras and a head tracker [Stephan Beck and Froehlich 2013]. Two groups were interacting by pointing and gesturing. In [Henry Fuchs and Bazin 2014], the concept of immersive 3D telepresence is proposed in geographically remote rooms, virtually joined as if co-located and separated by seemingly transparent walls. 3D reconstruction using multiple RGB-D cameras and 3D display with rear projection for eye contact were introduced together.

In the paper, there is proposed coexistent space where networked users can communicate, interact, and collaborate together by exchanging 4D+ sensation, human intension, and emotion. The 4D+ sensation includes 3D vision, 3D sound, force and torque, touch, and movements. The coexistent space is generated by seamless integration of real, virtual, and remote worlds while networked users experience the feeling of coexistence through 4D+ bi-directional interaction. Initial software framework and experimental results for coexistent space are developed. The paper is organized as follows. Chapter 2 introduces the concept of coexistent space while representative applications are proposed in chapter 3. Initial experimental results are shown in chapter 4 and the paper is concluded in chapter 5.

2 Coexistent Space

Coexistent Space is a place where networked users can experience the virtual and remote worlds as if they are real worlds and the users are located in the same place. The users get the feeling of coexistence by transferring and sharing 4D+ sensation, human intension and emotion in real-time while coexistent space not only allows users to see the physical properties of real-world environments, but also promotes them to perceive, operate, manipulate, and interact with actual and virtual objects with no physical restriction. Networked users will actively participate in producing and sharing their ideas, information, sensations, emotions, behavioral intentions and experiences with each other. Ultimately, the users will efficiently collaborate together in the space and become collectively intelligent.

The sense of coexistence is categorized into two types of human subjective experience: presence and flow. Presence is a psychological state where users experience virtual objects as actual objects. Sensory-based immersive experience can promote three types of presence: self, social and spatial presence. Self-presence is a psychological state in which users experience virtual self/selves as actual self while social presence is a psychological state in which users experience virtual social actuals as actual social actors. Spatial presence is a psychological state in which users experience virtual physical and spatial objects as actual physical and social objects. Flow is another human experience that explains users sub-

*ybj@kist.re.kr, ybj@chic.re.kr

jective feeling of immersion, often characterized by control over, attention to, and intrinsic involvement in their own action in particular. It is important to note that, unlike the feeling of presence, flow is only experienced when users have a high level of skills to successfully operate and manipulate the tasks of challenge. In this view point, flow is an action-based immersive experience and it is proactive and goal-oriented. When in the flow state, users become absorbed in their activity with total concentration: their awareness is narrowed to the activity itself; they lose self-consciousness, and they feel in control of their environment. The concept has been applied to both mediated and non-mediated environments where users interact with actual or virtual objects. It is expected that advanced interaction solutions, such as direct interaction using his/her hands and wearable haptic devices for arms and fingers, will provide the better feeling of immersion, flow.

From a technological perspective, three major technologies are required to develop coexistent space. First is a core technology that enables the coexistence of human and intelligent machines in real and remote worlds. It includes the development of wearable multi-modal human interface and capturing devices for whole body including fingers, tele-presence, bionic human interface, and intelligence for learning human behaviors and sensation. The purpose of these systems is to transfer human intention into artifacts in the real world, virtual world and remote world and to transmit real sensation back to the human. When this is achieved, humans can handle and feel virtual objects, and interact together like in the real world. A second core technology involves overcoming spatial and time constraints in the real world to enable the coexistence of human and virtual worlds. It includes technologies for constructing three-dimensional human avatars and mirror worlds that can handle objects interactively, merging real, virtual and remote worlds into an integrated world like real world, augmenting the mirror world with a real (or remote) world, and allowing direct interaction between users and the mirror world by using users hand. The mirror world is a mixture of the real (or remote) world and the virtual world and is a space in which a human avatar can navigate in lieu of humans. Multiple users can meet to collaborate or interact in the mirror world. The third core technology enables emotional interaction and innovative interaction for coexistence. People can exchange social emotions and information with artifacts in real, virtual and remote worlds through innovative means like emotion recognition without contacts with users. Social emotion can be detected by bio-signals from the human body or other visible clues which signal a change of emotions. In this way, emotional interaction with artifacts or other worlds will become possible.

3 Scenarios for Interaction and Collaboration in Coexistent Space

Coexistent space can be generated by 3D display systems using multiple 3D TVs and/or 3D projectors, AR glasses like google glass and see-thru head-mount displays like Oculus Rift, and telepresence robots equipped with 3D display. 3D hologram may be a best solution but it is not feasible to generate real-size human images. So, there are planned three scenarios that shape the future of society by the space.

3.1 HMD-based Coexistent Space for Learning and Education

One person in real world is wearing a HMD and haptic devices in his arm/hand while other four persons, who come from remote sites, are using the same devices, as shown in Fig. 1. The person in real world can see other four persons (avatars) and virtual information through his HMD while five persons are standing around a

virtual table. The teacher, who has a red necktie, makes his lecture on chemical, physics and mathematics by using air writing interface and by providing virtual information to students. Students can manipulate virtual information together using their hands, give and take some objects and information, feel the compliance of objects, and handshake together.

Figure 1 shows a chemistry class to understand the process to make water. During the typical course users look at a newcomer (avatar) which is virtually created from remote user site, talk and shake hands. They select, operate, disassemble, assemble, and manage a molecular of water with haptic feedback.

This platform utilizes coexistence space to provide a natural environment for face-to-face collaboration of teachers and students. The main advantage is that students directly work in 3D space and they actually manage and comprehend complex spatial problems of molecular formula better and faster than traditional methods. In this example, it is required to obtain real time 3D human images (avatar) from real world user to display the composed scene maintaining correct registration of real and virtual worlds. Each member is aware of each other created 3D avatar, as the others physical presence is guaranteed. This coexistence sense is the key feature of this application scenario. In addition, direct manipulation and dynamic interaction with virtual 3D objects using tangible devices are possible.

In public learning and education in each country, many under-represented regions lack support for the hands-on experience that is important in the learning process. HMD-based coexistent space will make effective education more precise and expand the capacity of learning services since the travel distance between teacher and students can be eliminated.

Figure 1: HMD-based Coexistent Space

3.2 3D Display-based Coexistent Space for Business

One person in real world is standing in front of a 3D display system wearing a stereoscopic glass and haptic devices in his arm/hand while other two persons, who come from remote sites, are using the same devices, as shown in Fig. 2. Three persons are standing around a virtual table. The user in real world can register 4D+ models of objects by using a high-speed modeling device for capturing 3D shape, texture, and haptic characteristics, like compliance and material characteristics, of real objects. Users can manipulate virtual information together using their hands, give and take some objects and information, disassemble and assemble a virtual object, and handle deformable objects. An eye contact system using stereo RGB-D cameras and full-HD color cameras is under development to generate a front image of each user.

In business, it is expected that designers at coexistent space can be an enriched their collaborated works. All participants in this system can touch, grasp, and feel the virtual model with worn haptic interfaces. The shared participation of captured and virtual content in a physics simulation enables realistic physical experiences. It is expected that coexistent space will provide a natural experience and realistic physical interaction between people, objects, and places even if they are geographically distributed.

This new telecommunication medium which combines aspects of 3D stereoscopic video conferencing with haptic collaborations, aims to provide alternative solution for business and daily life. Individuals in several locations are able to meet together at a place in coexistent space and share digital information through the immersive collaborative environment.

Figure 2: 3D Display-based Coexistent Space

3.3 Telepresence Robot-based Coexistent Space for Medical Examination

The above two platforms for realization of coexistence space, feasible interaction space between participants is limited to the front of the display or the space covered by its motion capture system. In order to overcome such spatial limit and extend potential interaction fields horizontally without bound, the third platform is proposed as a system which consists of a mobile telepresence robot and a master station as shown in Fig. 3. The system supports bidirectional communication of 3D video and 3D audio information between local (the robot) and remote sites (the master station) with eye contact. In additions, the system aims for the enhancement of coexistence feeling between participants by embodying various devices and functionalities for more natural interaction.

In an aging society, remote medical examination by doctors or nurses will be possible through telepresence robot with live images, sounds, and direct touch of sore spots in the human body. This platform will make medical examinations more precise and expand the capacity of medical services since the travel distance between doctor and patient can be eliminated. An additional advantage is that an elderly individual who cannot move arms or hands can control and manipulate home appliances through a coexistent interface based on sensation.

4 Experimental Results

This section introduces initial experimental results, and shows why coexistent space is indeed a step-change from current interactive telecommunication technologies. There have been done two experiments for human avatar-based remote collaboration and real-time

Figure 3: Telepresence Robot-based Coexistent Space

3D human image-based real-time collaboration.

There is developed a software framework including high-speed network module, multi-user collaboration framework module, and real-time rendering engine supporting physics. For reliable data transmission of real-time images, GPU-based H.264 codec for full-HD images from a video camera is developed while multi-channel data transmission module is developed to reduce transmission time and to minimize the latency by the transmission. The image data is transmitted as a 3D video data combined as RGB color images compressed and raw depth images. In this case, a standard H.264 video codec provided by NVidia is employed to compress and decompress Full-HD color images in real-time (up to 30fps). We get the final transmission performance of 20Hz for a full-HD color image, a VGA depth image, stereo sound, and other control data. The full-HD color image is captured by Flea 3 from Point Grey Inc. while the depth image is obtained from ASUS Xtion Pro Live RGB-D camera. Two cameras are combined mechanically and synchronized by using a trigger signal from the RGB-D camera.

Figure 4: Two users in site A and site B are pushing the ball together in coexistent space.

In order to synchronize various kinds of data, such as video and sound, in realizing scalable networks with quality of service guarantees, a timestamp-based synchronization mechanism, which involved a session manager covering from intra-media synchronization to point-to-multipoint synchronization except for inter-sensor synchronization, is adopted. This session manager supports various functions that inter-media synchronization, rate adaptation, dynamic media encoding, playing out buffer management, and so on. It is noted that the data consistency can be seen differently depending on the degree of motion and the number of scene changes. As

expected, rapid motions of camera and object without synchronization eventually lead to higher inconsistency at remote sites.

Figure 5: Desktop 3D display using two 3D TVs.

Based on the network framework, the first version of a remote collaboration framework called Spatium, including physics-based rendering engine, is developed to construct the coexistent space. Real-time video and sound is transferred and shared by users in remote sites. Fig. 4 shows an RGB-D image of coexistent space generated for three users while two users are pushing a ball in the space. Two desktop 3D displays are shown in Fig 5

4.1 Human avatar-based remote collaboration

The first experiment is done using large 3D wall displays, whose sizes are 165 and 135, for remote collaboration to carry a virtual box together for realizing the concept of coexistent space as shown in Fig. 6. Two remote users use the system to collaborate with each other through their bare hands and virtual avatar. Cooperation between users and improvement of spatial perception are treated mainly. One site is composed of a 165-inch stereoscopic wall display combined with passive polarized glasses and a projector with keystone effect correction to have rectangular projection on real floor. The display system has been used as a gate for a local user to meet someone in his own place with the enhancement of spatial sense. Users motion and pose are captured by an ASUS Xtion Pro Live RGB-D camera and each user wears data glove. Users share virtual objects and manipulate them together with bare hands. Vibration feedback is offered once there is a contact between the hand and the virtual object.

Figure 6: Large 3D display-based remote collaboration system.

4.2 Human Image-based remote collaboration

Two desktop 3D display systems shown in Fig. 5 are used as networked hardware platforms and the developed software framework, Spatium, is installed to generate coexistent space for collaborative interaction. The desktop system consists of two 3D TVs and a camera module including a full-HD color camera and ASUS Xtion Pro Live RGB-D camera. The image and depth data are converted into a 3D mesh and displayed in the desktop stereoscopic display. In order to provide the better feeling of coexistence by improving the reality of 3D images, the full-HD color texture is merged with the 3D mesh as shown in Fig. 7.

Figure 7: The comparison of 3D image resolution.

As shown in Fig. 4, three systems transfer various sensor data and combine them in the coexistence space. Two users in site A and B played with a virtual ball based on physics interaction. Site C displays real humans, virtual ball and background together.

4.3 Telepresence robot-based 3D interaction

There is developed a mobile robot-based 3D interaction platform for realizing coexistent space as shown in Fig. 8. Now, an additional medium that goes beyond conventional verbal communication is provided, in order to draw more attentions and raise involvement in mutual interaction from participants. A 3D projector module combined with a RGB-D camera and a full-HD color camera is developed for more natural interaction between remote users by augmenting virtual information onto real objects or environments. It is mounted on 2-DoF (yaw-tilt) neck placed at both sides, respectively. It is used to track human hands and objects such that bare hand interactions, mutual interactive drawing, and texture mapping on a 3D object are envisioned.

Next, the needs for exploration using the telepresence robot may arise even in the middle of interactions involving both hands of the participant on the master station. This necessitates the development of non-hand held type of user interface that can control the mobility of the telepresence robot intuitively. To this end, a rudder pedal type user interface is developed with two actuated joints. Each joint of the pedal is directly mapped to the translational and rotational motion of differential type of mobile platform, respectively, such that intuitive control is ensured. Further, user can feel the obstacles on the way of telepresence robot through kinesthetic feedback. Patterned array of miniaturized vibration motors is being developed which will be used for delivering the texture of the road surface or the study of nonverbal communication between users.

5 Conclusion

The coexistent space is proposed as a space where networked users can communicate, interact, and collaborate together by exchange-

Figure 8: Overview of the coexistent system with the master station (left) and telepresence robot (right): 1. camera/projector module, 2/8. 3D display, 3. rudder pedal type user interface, 4. head tracking system, 5. side-camera module, 6. stereo speaker system, 7. differential drive type mobile platform.

ing 4D+ sensation, human intension, and emotion. Initial software framework and interaction-and-collaboration platforms are under developed to provide the feeling of coexistence to networked users. Initial results show its strong possibility hopefully as a future communication and interaction platform. The future research may be focused on the development of wearable human interface and interaction technologies to enhance the ‘Flow’.

Acknowledgements

This work was supported by the Global Frontier R&D Program on Human-centered Interaction for Coexistence funded by a National Research Foundation of Korea grant funded by the ministry of Science, ICT, and Future Planning in Korean Government (NRFM1AXA003-2011-0028374).

References

- HENRY FUCHS, A. S., AND BAZIN, J.-C. 2014. Immersive 3d telepresence. *IEEE magazine on Computer* 47, 7 (July), 46–52.
- LANIER, J. 2001. *Scientific American* 284, 4, 68–75.
- STEPHAN BECK, ANDRE KUNERT, A. K., AND FROEHLICH, B. 2013. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics* 19, 4, 616–625.