

LiveSphere: Sharing the Surrounding Visual Environment for Immersive Experience in Remote Collaboration

Shohei Nagai

The University of Tokyo
Tokyo, Japan
shohei.nagai14@gmail.com

Shunichi Kasahara

Sony Computer Science
Laboratory
Tokyo, Japan
kasahara@csl.sony.co.jp

Jun Rekimoto

The University of Tokyo
Sony Computer Science
Laboratory
Tokyo, Japan
rekimoto@acm.org

ABSTRACT

Sharing an immersive experience enhances situational awareness, enabling effective collaboration between persons in different places. The development of a head-worn camera enables us to capture the first-person view and share it as the personal experience. However, the limited viewing angle of such cameras prevents a remote viewer from obtaining a complete surrounding situation. In this paper, we propose a system called LiveSphere, which presents the entire surrounding visual environment using a head-worn system with multiple cameras. The image-stabilizing algorithm compensates image motion caused by the wearer's head movement. This enables the remote viewer to look around the environment independently from the wearer's head direction. We developed a prototype to examine how sharing the surrounding visual environment improves collaboration between persons in different places.

Author Keywords

Experience sharing; spherical image; wearable computing; remote collaboration; visual environment;

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Multimedia Information Systems

INTRODUCTION

Sharing an immersive experience in real-time is a highly effective solution for collaboration between persons in different places. Such experience makes it easy to understand the situation, and provide effective guidance as if they exist in the same place. Particularly, Fussel et al. [2] suggested that sharing “visual space” could facilitate effective collaboration. Previous studies explore various approaches to share visual space utilizing video

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee, provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

TEI '15, January 16 - 19 2015, Stanford, CA, USA
Copyright 2015 ACM 978-1-4503-3305-4/15/01...\$15.00
<http://dx.doi.org/10.1145/2677199.2680549>



Figure 1. The LiveSphere headgear with six cameras captures the entire surrounding scene of the wearer, without hampering the wearer's activities.

streaming with cameras [e.g. 7]. The cameras capture parts of the visual space and share them to a person in remote place.

Head-mount cameras such as a GoPro [3] allow us to capture a first-person view and share that experience. In previous studies of remote collaboration, using wearable cameras is one of the primary approaches for sharing visual information. Such wearable camera can share the first-person view of an individual's environment, enabling collaboration with that individual from a remote location. That first-person view is useful for understanding the situation and providing effective guidance and instruction.

On the other hand, sharing a first-person view may introduce several issues. When the wearer moves their head and changes direction, the captured image tracks their field of vision. Often times, this movement can cause intense shaking, which may also cause motion sickness [1]. During remote collaboration, the remote viewer needs to fix their vision on specific objects, when necessary. This allows the remote viewer to understand the wearer's environment and give appropriate guidance. However, the object in question can easily go out of perspective depending on the head movement of the wearer, making it more difficult for the viewer to keep up with the situational context.

To solve these problems, JackIn [4], which is one of our research projects, proposed an approach by which a series of first person images are spatially merged to create a virtual out-of-body view. Using this technique, a remote

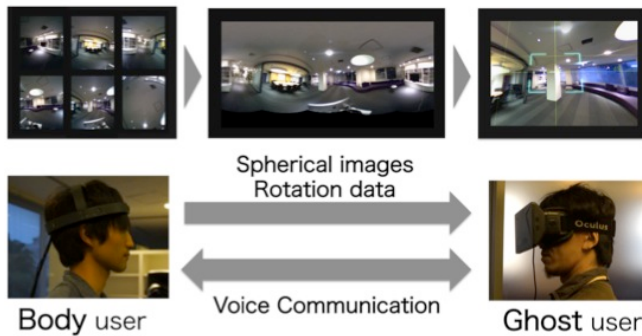


Figure 2. LiveSphere system configuration. A Body user wears the headgear capturing the surrounding visual environment and transmits it. A Ghost user can look around the environment independently from Body’s head direction. The direction of Body’s head is overlaid in Ghost’s perspective.

viewer with the appropriate instructional expertise can observe the scene independently from the wearer’s head direction. However, as it relies on the assumption that the wearer pans the environment to capture images for merging, it is not always robust.

In this paper, we propose LiveSphere, a novel remote collaboration system that shares the visual environment surrounding a wearer with a viewer in a remote location using a spherical image. The image rotation caused by a change in the wearer’s head movement is compensated by rotation data. We gather this data from a nine-axis sensor recorded simultaneously with the images. That provides the remote viewer an immersive experience as if present in the real environment, and looked around independently from the wearer’s head direction.

In the remainder of this paper, we use the term “Body” to represent a person who performs activity in the real environment, and “Ghost” for the person who observes the shared environment and gives guidance and instruction.

SHARING SURROUNDING VISUAL ENVIRONMENT

To provide Ghost with an immersive experience and realize effective collaboration, LiveSphere shares the surrounding visual environment. Sharing the surrounding visual environment consists of the following:

- Provide perspective ranging over a fixed scope.
- Provide independent perspective from Body’s head rotation.

Perspective over fixed scope

To understand the Body’s situation, Ghost needs to observe the environment from different angles and pay attention to everything related to the collaborative task. However, the perspective captured by one wearable camera is limited in fixed scope. Conventional studies have used different approaches to remove this limitation. For example, Fussell et al. [2] utilized the combination of a wearable camera and a scene-oriented stationary camera. Kashiwabara et al. [5]

used an actuator to change the direction of the camera to capture wider area.

LiveSphere captures almost the entire surrounding environment visible from Body’s position using six cameras directed toward the outside. This enables Ghost to look around and observe everything in the surrounding environment as if they were present. In addition, this enables Ghost to look around and observe everything around Body over the limitation of the single fixed-scope camera. That enhances Ghost’s awareness of the situation surrounding Body.

Independent Perspective from Body’s Head Motion

For a deeper understanding of Body’s location and situation, it is important that Ghost can look around. In addition, if there is an object that should be paid attention to, Ghost can focus on it for a certain period. In the conventional wearable camera system, Body moves and turns their head of their own volition, with the resulting view dependent on head direction. Therefore, the shared perspective is suddenly turned around, and Ghost may lose perspective on an object of focus. If Ghost wants to focus on a specific object, Body must maintain their head direction toward it, which is not practical in a real activity. LiveSphere removes this limitation by combining the images captured from the surrounding visual environment with Body’s head rotation. The shared images’ rotation is compensated with the rotation data. That provides Ghost with perspective, which is independent from Body’s head motion. This also resolves any motion sickness caused by shaky images.

Interaction through sharing visual environment

In sharing surrounding visual environment, LiveSphere can realize various applications, which are difficult in a conventional remote collaboration system. The following are possible applications on LiveSphere.

Disaster rescue

We can use LiveSphere in situations that require persons to perform an activity in a dangerous place. For example, in the case of disaster rescue, Body actually enters a disaster-stricken area (which might be dangerous) to search victims. Conversely, Ghost observes the area remotely from the perspective of Body’s position using the shared surrounding visual environment. Ghost can guide Body to assist in searching for the victims. In this manner, Ghost can remain in a safe place and focus more on searching. Body, who is in dangerous place, may concentrate on moving their body.

Navigation to specific object

LiveSphere is also useful in daily cases where Body looks for a specific object in an unfamiliar place that Ghost knows well. If Body and Ghost communicate only with voice, it takes a lot of time for Body to share location details and for Ghost to communicate where Body should search. Conversely, if Body shares the surrounding visual

environment with Ghost, Ghost gains a more intuitive understanding of where Body is located. This enables effective guidance using directional expressions such as “It’s on your right,” and deictic expressions such as “It is on that shelf.”

Virtual sports experience

We can utilize LiveSphere in entertainment applications. For example, Body is a soccer player and Ghost as a spectator, shares the visual environment surrounding the player on the ground. Ghost can experience it as if they were the soccer player and played the game. Ghost can enjoy the game from literally any different viewpoint. This enables elderly or handicapped individuals who cannot play sports to enjoy them. It is also possible for Ghost to make decisions on how to move and play on the game by giving directions to Body. This can induce the feeling of participation, which makes sports more enjoyable.

PROTOTYPE DESIGN REQUIREMENT

We developed a LiveSphere prototype to demonstrate how a shared surrounding visual environment changed the performance of remote collaboration. In designing the LiveSphere prototype, besides sharing the surrounding visual environment, we listed the following:

Show Body’s Situation

For a remote collaboration system where persons in different places work together, it is important to support mutual awareness [3]. To provide effective guidance and instruction, Ghost should be able to understand how Body sees the environment in addition to what they are focused on, and in this situation, what assistance is needed to achieve a task. LiveSphere should present the information, which shows Body’s perspective to understand above things. This also enables Ghost to use directional and deictic expressions, which give Body a more intuitive understanding of the guidance and instruction from Ghost.

Shape of Capturing Device

Wearability is an important consideration for designing LiveSphere’s capturing device. Some devices record spherical images, such as 360Heros [11]. However, it might be cumbersome in situations where Body moves their body and head dynamically due to the device’s weight and high center of balance. In one of our projects where gymnasts wore 360Heros to record the spherical view from their perspective, they lost balance, and some fell when landing.

The height of the cameras is also important consideration. The captured environment should be close to Body’s viewpoint. As such, the cameras need to be located near the height of Body’s eyes. One reason for this is coverage of the captured visual environment. If Body wears cameras on Body’s head, the top of Body’s head obstructs the captured images. This causes Ghost to lose track of specific objects that Body can watch. The other reason is gaze matching. In

human interaction such as face-to-face conversation and eye contact, sharing the correct scene expression depends on proximity of the gap between the camera and the wearer’s eyes [6].

PROTOTYPE SYSTEM CONFIGURATION

The LiveSphere system consists of the following three parts (Figure 2): In the first part, the visual environment surrounding Body is captured and transmitted to Ghost in remote location. In the second part, Ghost is presented with the shared images as the surrounding visual environment. They can observe the shared environment as if present. In the third part, according to the shared information, Ghost and Body communicate through voice.

Capturing Visual Surrounding Environment

In LiveSphere, Body wears a head-mount device called the headgear to capture the surrounding environment (Figure 1). The headgear has six wide-angle cameras. It is designed to affix the cameras lower than the top of Body’s head, close to Body’s eye level. This lowers the center of balance and prevents the headgear from hampering any activities (Figure 3). This also brings the captured images close to the Body’s actual viewpoint. Each of the cameras has approximately 135° horizontal angle, and 122° vertical angle of view. The cameras cover approximately 93% of the surrounding visual environment. Such wide-angle cameras tend to have barrel-shaped image distortion. Therefore, we calibrate the images captured using Zhang’s method [10]. We synthesize these images into one spherical image through equirectangular projection using a GPU, which is streamed to Ghost via network in a real-time. The headgear also has nine degrees of freedom sensor to record the Body’s direction of the head. We use this data for stabilizing the rotation of Body’s head.

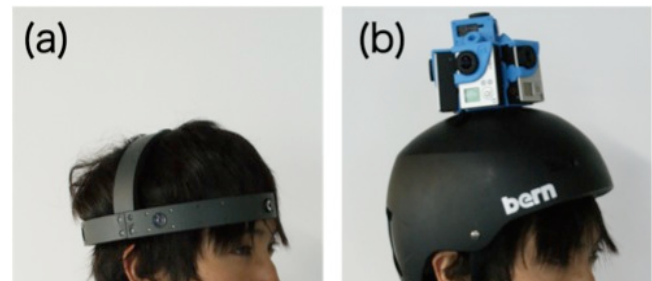


Figure 3. (a) The LiveSphere headgear, (b) 360Heros. The headgear is small and light weight. The center of balance is lower than existing devices, which capture the surrounding scene without hampering activities, and cover the area Body can see from his/her viewpoint.

Present the Shared Environment

We use Oculus Rift [8] to present the transmitted images of the visual environment surrounding Body (Figure 2). Oculus Rift is a head-mounted display, which can detect the user’s head rotation and render scenes depending on it. The shared image, which was projected through Equirectangular

projection on Body side, is mapped to sphere, and Oculus rift shows the perspective from the center of the sphere. When the Body moves their head, the image might be turn around. We compensate this turning by rotating sphere with the same rotation angle as the Body's rotation, which is recorded simultaneously with the images using the 9-axis sensor (Figure 4). Therefore, the user can look around in the shared environment independently from Body's head rotation as if present (Figure 5).

In the display shown by Oculus Rift, the direction of Body's head is also provided using an indicator, which is overlaid on Ghost's perspective (Figure 2, 5). This makes it easy for Ghost to understand the Body's situation, and give effective guidance.

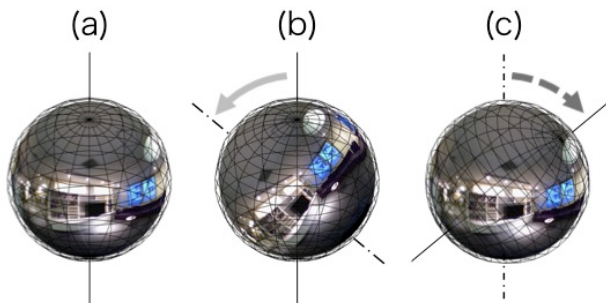


Figure 4. (a) The shared spherical image mapped on a sphere (b) Body's head rotation turns the Ghost's perspective. (c) The image rotation is compensated by turning the mapped sphere with the rotation data recorded using 9-axis sensor.

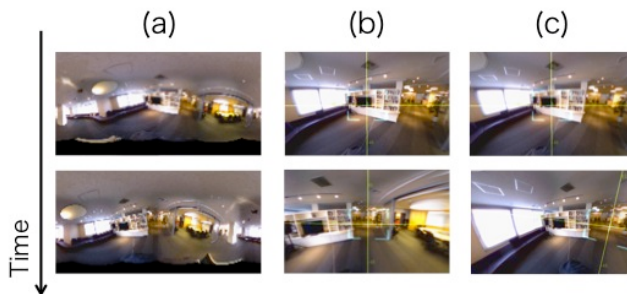


Figure 5. (a) Spherical image. (b) Unstabilized view. (c) Stabilized view. Ghost is able to look around the shared environment independently from Body's head rotation.

Guidance

Voice communication is available on LiveSphere. In the shared visual environment surrounding Body, Ghost can understand Body's situation without the constraint of Body's head direction. Ghost can even provide information about objects in directions unrelated to Body's field of vision. Because Body and Ghost share the surrounding environment, they can communicate with each other more efficiently using deictic expressions and directional expressions.

CONCLUSION AND FUTURE WORK

In this paper, we present a new approach that shares the surrounding visual environment using a spherical image, and provides immersive experiences with a person in remote location for effective collaboration. We developed the LiveSphere prototype and demonstrated its conceptual feasibility. As future work, we explore a more effective way for individuals to interact other than voice communication. JackIn utilized AR annotation by see-through head-mounted display to navigate Body to a specific point. The virtual force is also a possible approach. Rekimoto [9] proposed a small, lightweight actuator, which provides virtual force. This actuator may easily be embedded in the headgear, and does not hamper Body's head movement. Moreover, we are implementing real-time image processing for stabilization because stabilization only by motion sensor may have difficulty in synchronizing with image and sensing, which can cause jitters in the video sequence.

REFERENCES

1. Benson, Alan J. Motion sickness. *Medical aspects of harsh environments* 2 (2002), 1048-1083.
2. Fussell, S.R., Setlock, L.D., and Kraut, R.E. Effects of Head-Mounted and Scene-Oriented Video Systems on Remote Collaboration on Physical Tasks. In *Proc. CHI '03*, ACM Press (2003), 513-520.
3. GoPro Official Website, <http://gopro.com/>
4. Kasahara, S. and Rekimoto, J. JackIn: Integrating the First Person View with Out-of- Body Vision Generation for Human-Human Augmentation. In *Proc. AH'14*, ACM Press (2014), 1-8.
5. Kashiwabara, T., Osawa, H., Shinozawa, K., and Imai, M. TEROOS: A Wearable Avatar to Enhance Joint Activities. In *Proc. CHI'12*, ACM Press (2012), 2001-2004.
6. Kondo, K., Mukaigawa, Y., and Yagi, Y. Wearable Imaging System for Capturing Omnidirectional Movies from a First-person Perspective. In *Proc. VRST '09*, ACM Press (2009), 11-18.
7. Kuzuoka, H., Spatial Workspace Collaboration: A SharedView Video Support System for Remote Collaboration Capability, In *Proc. CHI '92*, ACM Press (1992), 533-540.
8. Oculus Rift Official Website, <http://www.oculusvr.com/>
9. Rekimoto, J. Traxion : A Tactile Interaction Device with Virtual Force Sensation. In *Proc. UIST'13*, ACM Press (2013), 427-431.
10. Zhang, Z. A Flexible New Technique for Camera Calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 22.11 (2000). 1330-1334.
11. 360Heros Official Website, <http://www.360heros.com>