Analyzing Interactions between a Pair Out Together Real and Virtual

Ching-Tzun Chang
Department of Computer Science, Graduate School of SIE, University of Tsukuba.
tsubaki@iplab.cs.tsukuba.ac.jp

Shin Takahashi
Faculty of Engineering, Information and Systems, University of Tsukuba
shin@cs.tsukuba.ac.jp

Jiro Tanaka
Faculty of Engineering, Information and Systems, University of Tsukuba
jiro@cs.tsukuba.ac.jp

Abstract

We attempted to find a set of interaction methods to enable indoor and outdoor users to interact and share the "Out Together Feeling," defined as when two people at different locations feel as though they are together. To achieve this, we first designed and implemented system features to share the outdoor user's view with the indoor user and also share their focus to help them look together at the same point. To this end, we designed and performed an experiment to observe the basic elements of communication between people who are really together. Then we performed an experiment in which indoor and outdoor users communicated via a videophone and observed the communication methods of each user as s/he attempted to achieve a given goal. Finally, we performed an experiment to evaluate our prototype system. We also analyzed the subjects' comments and results of the experiment and confirmed that our purpose had been achieved to a certain extent.

1. Introduction

With the use of high-speed mobile networks, it is now possible to provide high-speed and stable mobile communications in outdoor environments. In addition, mobile video communication such as the videophone has become feasible. However, the full potential of mobile video communication has yet to be fully exploited. One reason for this is that most video communications systems developed to date assume mainly face-to-face communication, which is not always helpful for users who may want to focus on other information such as body language and actions.

Mobile video communication has other possibilities. For example, with a videophone, users can shoot and send the video stream of their surroundings to an indoor user. They can share images of the place and talk about it, which is a type of remote video communication that we call the *Out Together Feeling*. This type of communication allows users to feel as if they are together in the same place. To amplify such feelings, it is necessary to realize various actions such as share the focal direction naturally.

In our earlier work, we implemented a system called *WithYou* (Chang et al. AVI2012) to provide the *Out Together Feeling* between an indoor user and an outdoor user. *WithYou* enabled the indoor user to freely peruse the surroundings of the outdoor user, with each user being aware of the direction of the other user's gaze. We also performed an indoor evaluation of *WithYou*, which showed that the system works well in the indoor environment and users could obtain a certain level of *Out Together Feeling*.

Our final goal is to make full use of remote video communication technology and interaction methods to realize the *Out Together Feeling*. As a step toward this goal, we first implemented a basic system based on our early prototype *WithYou* system to communicate between an indoor user and an outdoor user. The indoor user shares the remote environment via a head-mounted display (HMD) with the outdoor user who wears a special device with a pan/tilt/zoom camera mounted on the chest.

To determine the effectiveness for achieving the *Out Together Feeling*, we designed a series of experiments to examine how a videophone call and use of our system differ from the users actually being together in the outside environment.

First, we designed two experiments to examine what types of communication skills and actions people will take when they are actually together in the outside environment and when using a videophone call. In this series of experiments, subjects determined their own mission target, such as to buy something or survey something in which they are interested.

Next, a street evaluation of our system was performed. Subjects used our system to achieve their own mission target. At the end, we collected the results (video recordings and notes) and had participants complete questionnaires to guide future improvements to our system.

The remainder of this paper is organized as follows. In Section 2, we define the *Out Together Feeling* in greater detail and review our basic prototype system and the interface design for realizing the *Out Together Feeling*. Implementation of the system is described in Section 3. We present the design and results of a preliminary evaluation

of our system in Sections 4, 5, and 6, and discuss related work in Section 7. Our conclusions are presented in Section 8.

2. The *WithYou* System

As a step toward realizing the *Out Together Feeling*, we have implemented a system called *WithYou*, which assumes that there are two users: one is outside (outdoor user) and the other is in a room (indoor user) (Figure 1). The indoor user is defined as the person who uses the system to get the *Out Together Feeling* to go outside virtually.

Wearable devices were mounted on the outdoor user's shoulder with a pan-tilt camera and various sensors. Live images from the outdoor user and his/her facing direction are displayed on the indoor user's HMD screen.

When using *WithYou*, the indoor user and outdoor user can communicate and interact by voice, sharing focus and hand gestures. Our system is described in more detail in our previous paper [1].

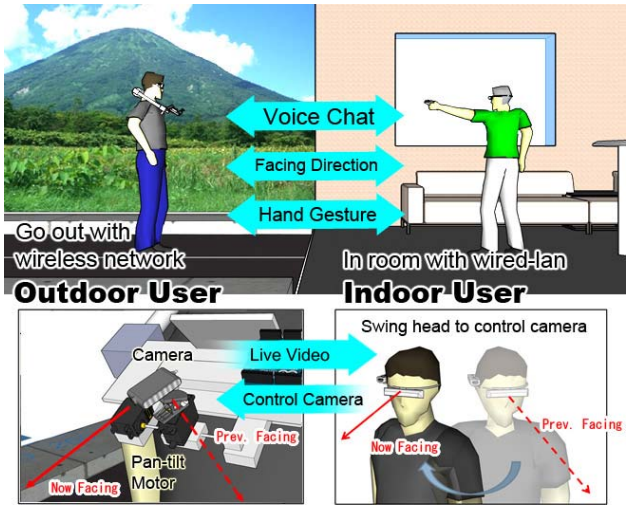

Figure 1. System overview

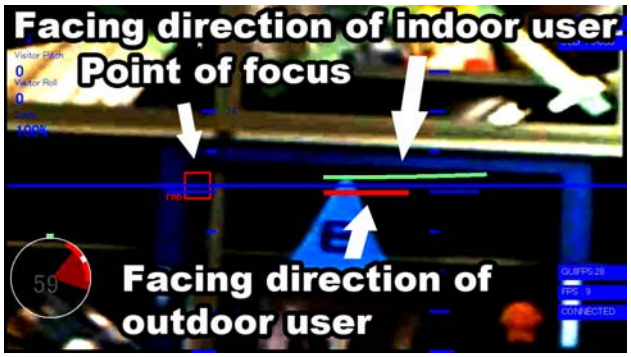

Figure 2. Graphical user interface on the indoor user side

2.1. Free viewing for the indoor user and graphical user interface

A pan-tilt camera is placed on the outdoor user's chest. The direction in which the camera is facing is linked to the direction of the indoor user's head. Thus, the indoor user controls the remote camera by turning his/her head (Figure 1).

The indoor user views the remote camera image with the HMD. A graphical user interface (GUI) is overlaid on the camera image to provide information about the outdoor user. The GUI shows the following information (Figure 2).

1. The indoor user's facing direction (green line);
2. The indoor user's posture (tilt and roll);
3. The outdoor user's facing direction (red line) and focus point (red grid);
4. Camera owner, interaction mode, zoom value, system messages, and status.

2.2. Sharing the focus and joint attention

Sharing the focus of attention means that the users know where the other user is facing. As the pan-tilt camera is mounted on the outdoor user's chest, he/she can know which way the indoor user is facing by observing the direction of the camera. In addition, the facing direction of the outdoor user is displayed on the indoor user's GUI, and the indoor user can know which way the outdoor user is facing by observing the red line (Figure 2) on the HMD screen. When the indoor user and outdoor user focus in the same direction, joint attention mode is activated, and the system will play a sound and show a message to inform both users. Joint attention helps the two users know that they are looking in the same direction.

2.3. Implementation

A wearable device consisting of gyro sensors, a digital compass, a pan-tilt camera, a mono liquid crystal display(LCD), and a hand control device is mounted on the outdoor user's shoulders with a mobile computer placed in a backpack (Figure 3, left). The facing direction (body and head) can be determined by geomagnetic sensors. The LCD shows the outdoor user both the system status and the indoor user's facing direction.

Figure 3. The devices worn by the indoor and outdoor users

To achieve good camera rotation speed and range, we use two high-speed servomotors to make two axes of rotation, with pan-tilt camera rotation ranges of 180° for pan and 130° for tilt. A USB camera (Logicool C910) is mounted on the motor system. The motors are controlled by an embedded computer board (Arduino-mega). In addition, the camera has a built-in digital-zoom function so the indoor user can zoom in/out on remote images by operating a hand controller.

The indoor user wears an HMD with various sensors (Figure 3, right). A digital compass is placed on the indoor user's head to determine the facing direction and to synchronize rotation of the remote camera.

3. Experimental design

To achieve our target, i.e., to determine the best methods of interaction to achieve the *Out Together Feeling*, we designed three experiments to examine the following issues:

1. How people communicate with each other when they actually go outside together;
2. The differences from 1 when people use videophone call features to go out virtually, and how they overcome problems;
3. The advantages and disadvantages of our prototype system, *WithYou*, compared to 1 and 2.

3.1. Experiment A (actually going outside together)

In the first experiment, a pair of subjects went shopping together in the electronics district (Akihabara) of Tokyo, Japan. Subjects were able to choose their own mission target (such as something to buy or survey), which they attempted to achieve during the experiment. Two staff followed them to take video recordings, observe their actions, and record their findings. At the end, the subjects were asked to fill out a questionnaire.

The major purpose of this experiment was to identify the basic communication skills people use when going out together. The results of this experiment also helped us to identify necessary features of our prototype system.

3.2. Experiment B (going outside together virtually by videophone)

In the second experiment, two subjects used a videophone call to go out together virtually. The outdoor user went to the Yobobashi-Camera Akiba shopping center, and the indoor user remained in the rest space of the same shopping center. During the experiment, the mission target was the same as in experiment A but the users had to communicate by videophone call (such as requiring the outdoor user to change direction, etc). The major purpose of this experiment was to observe people's communication skills during the experiment and determine the differences from experiment A.

3.4. Experiment C (street evaluation)

In experiment C, to verify the usage conditions and the effectiveness of our prototype system, we performed an evaluation experiment at Akihabara. During the experiment, subjects chose their own mission target and used the system features freely to achieve their goals. In this experiment, the outdoor user actually went out and moved around, while the indoor user remained inside and went out virtually using our system.

4. Experiment A

4.1. Experimental methods

Four subjects (two pairs) participated in the experiment. During the experiment, subjects were required to have no impression of "doing an experiment" and went shopping together as usual (Figure 4). This experiment was also performed at Akihabara, which was decided taking the subjects' interests and requests into consideration.

Figure 4. Scene from experiment A

4.1.1. Experimental conditions. First, the participants were briefly informed (~5 min) about the flow of the experiment. Subjects were asked to chose their own mission target, which in this experiment was "buy something" or "survey or check something." Then the subjects had 20 min to achieve their mission target, and were able to move freely in the street and into stores.

During the experiment, two staff members followed the subjects during the task; one staff member videoed the subjects, while the other observed and noted the subjects' actions, communication skills, and the methods of interaction between the two subjects. At the end of the task, the two subjects were asked to fill out a questionnaire.

4.2. Experimental results

We confirmed communication skills by analyzing the results of the experiment (i.e., video, notes, and questionnaire). The most important interaction between subjects was "to pick something up and look at it together." The buying target was often an electronic

product; by checking the video and notes, we found that subjects usually took various products in their hands and talked about the specifications or design together.

In addition, we also found that "pointing with a finger" was an important mode of communication between the subjects. The subjects indicated their interest by pointing a finger at a sign or product, which may not be picked up easily. Many other skills were observed in this experiment as outlined in Table 2.

Communication Skill	Frequency
Notice partner's facing then focus to the same direction	10.4Times/10Min
Pointing with a finger	4.6Times/10Min
Notice partner standing still somewhere, and look where the partner is focusing their attention	4.6Times/10Min
Turn around to check that the partner is following	3.9Times/10Min
Pick something up and look at it together	3.3Times/10Min
Describe by gesture	2.6Times/10Min

Table 2. Communication skills observed in experiment A

In Table 2, we analyzed the video recording of the experiment B and indicated of each communication skill. Duo to cannot took full video of each session(ex: in a shop). we calculated the practiced times of each skill per 10 min.

Although we realized that the results could be influenced by the experimental setting, we focused on shopping activities common to everyday life. The results may have been different if the purpose of the activities in the experiment was sightseeing.

The results of the experiment indicated that focus and gesture (pointing) are two important elements for activities when two subjects are out together. Based on these results, our prototype system gave high priority to focus sharing and detection between the indoor and outdoor users. We also implemented a body communication (gesture) feature in our system.

5. Experiment B

5.1. Experimental methods

Four subjects (2 pairs) participated in this experiment, which was performed at a shopping center (Yodobashi Camera Akiba) in Tokyo, Japan. In this experiment, one subject remained in the rest place as the indoor user (Figure 5, left), while the other walked freely (including every floor) around the shopping center as the outdoor user (Figure 5, right).

5.1.1. Experimental conditions. The conditions in this experiment were almost the same as those in experiment A, except that the subjects communicated by videophone. To achieve the mission target, the indoor user could request the outdoor user to change the camera shooting direction and also request that the outdoor user move to a specific product area or floor. During the experiment, the subjects had 15 min to achieve their mission target, also, we took videos of both the indoor user and outdoor user, and one staff member made observations and notes regarding the outdoor user's actions, communication skills, and methods of interaction between the two subjects. At the end of the experiment, the two subjects were asked to fill out a questionnaire.

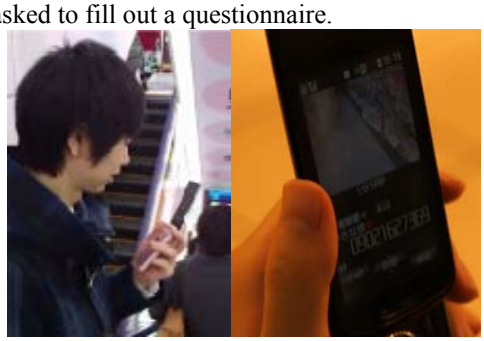

Figure 5. Outdoor user (left) and indoor user (right) in experiment B

5.2. Experiment results

In this experiment, we confirmed many problems when using the videophone to realize the *Out Together Feeling*. We also found that "look freely around the surroundings" was the most important element to achieve this feeling.

We considered it necessary for the indoor user to be able to freely and naturally look around the outdoor user's surroundings. When using a videophone, the shooting direction is controlled entirely by the outdoor user. If the indoor user could control the camera, it would facilitate looking around, which would help in finding places of interest.

During the experiment, we found that the indoor user had difficulty knowing in which direction the outdoor user was moving. We also found that conversation was dominant during the experiment. Although live video images can show the outside environment to the indoor user, the low image quality and low frame rate make the videophone difficult to use for this purpose (e.g., it is difficult to read price tags by checking the live image).

In experiment A, we found that "pointing with a finger" was often used. However, in experiment B, the frequent types of interaction between subjects were those in which the indoor user asked the outdoor user to turn the camera toward a specific direction.

Analysis of the video recording of the experiments indicated that the purposes of these interactions were essentially the same; in experiment B, the outdoor user changed the camera's shooting direction instead of pointing with a finger.

Communication Skill	Frequency
Outdoor user shows product information (price, design, etc.) to the camera	7.5Times/15min
Outdoor user read the price/information for indoor user	4.0Times/15min
Indoor user requests the outdoor user to move somewhere (area or floor)	4.0Times/15min
Outdoor user finding product for indoor user	3.5Times/15min
Indoor user requests the outdoor user to turn the camera toward a specific direction	3.0Times/15min
Outdoor user films something with the camera	3.0Times/15min

Table 3. Interaction methods in experiment B

Question	Indoor user	Outdoor user
I felt a sense of doing something together with my remote partner via the videophone.	3	3
I felt that the videophone is suitable for realizing the *Out Together Feeling*.	2	2
Please rank the videophone image quality (1, low – 5, high)	2	2

Scores were given on a scale from 1 to 5, where 5 indicates strong agreement and 1 indicates strong disagreement.

Table 4. Questionnaire results of experiment B

Comment from subject	Character
Q: During the experiment, what type of interaction did you have with the indoor user? A: Reading the price tag to the indoor user, directing the camera to the product.	Outdoor user
Q: Please give your general impression of this experiment. A: Outdoor user turns camera → indoor user checks the live image and requests the outdoor user to change the facing direction. A: No time to look at the indoor user. It felt no different from a voice call.	Outdoor user
Comment: Due to the low image quality and poor focus, small text on products was difficult to read.	Indoor User

Table 5. Questionnaire results (user comments)

6. Experiment C

6.1. Experimental methods

Four subjects (2 pairs) participated in this experiment in which the outdoor user walked around outside (Figure 6, right), while the indoor user remained in a car (Figure 6, left).

6.1.2. Experimental conditions. First, the participants practiced how to use the system for about 10 min. Then, as in experiments A and B, the indoor user chose a mission target—in this experiment it was "buy or survey something." During the experiment, the subjects had 20 min to achieve their mission target; the outdoor user was able to move freely in the street and into stores. During the experiment, videos were taken of both the indoor user and outdoor user. At the end of the task, the two subjects were asked to fill out a questionnaire.

Figure 6. Indoor and outdoor users' environments in experiment C

6.2. Experimental results

Tables 7 show the results of experiment C.

Question	Indoor user	Outdoor User
I felt a sense of doing something together with my remote partner.	3.5	3.5
I looked around the surroundings freely.	3.5	N/A
I could determine the facing direction of my remote partner.	3.0	N/A
With regard to the system feature "Absolute view for indoor user," I felt that the remote video was stable and that the camera always faced in the required direction.	2.5	N/A
I felt that "Joint attention" was useful and improved activities together with my remote partner.	3.5	3.5

Scores were given on a scale from 1 to 5, where 5 indicates strong agreement and 1 indicates strong disagreement.

Table 7. Questionnaire results

The results of the questionnaire indicated that the indoor and outdoor users can determine each other's facing direction using our system. The outdoor user could determine the indoor user's facing direction by observing the shooting direction of the camera, while the indoor user may check the outdoor user's facing direction from the GUI. Therefore, we concluded that both users could determine each other's focus of interest.

In experiment C, the indoor user often perused the surroundings by controlling the remote camera, which provided the indoor user the freedom to share the outdoor

environment. If the outdoor user approached a store, the indoor user could turn his own head to check the price tag or some goods placed outside the store. One major difference from videophone is that the outdoor user could use both hands, making gestures such as “picking something up and looking at it together” more conveniently than when using the videophone.

6.2.1. Realization of the *Out Together Feeling.* In the questionnaire, the indoor and outdoor users gave average scores of 3.5 points (out of a total of 5 points) for the item “I felt a sense of doing something together with my remote partner.” Therefore, both users experienced the *Out Together Feeling*.

6.2.2. Realization of the gesture “pointing with a finger”. In our system, the indoor user can control the remote camera in two ways: by simply turning the head, and by turning the Wii-remote control in his hand. The indoor user can switch control of the remote camera by pressing a button. Using this feature, the indoor user can point (i.e., change the shooting direction of the camera) by hand.

6.2.3. Enable face-to-face interaction between indoor and outdoor users. At the end of experiment C, we received multiple comments that it may be helpful to allow the outdoor user to see the indoor user by camera. The outdoor user had difficulty in reaching the same level of *Out Together Feeling* as the indoor user. Allowing the outdoor and indoor users to look at each other may further enhance the feeling of sharing an activity together.

7. Related work

Tsumaki et al. [2] proposed and developed a wearable robotic system called Telecommunicator, which allows the local site user to communicate with others at a remote site. Telecommunicator is a wearable robotic device mounted on the user’s shoulders, which consists of a rotatable video camera and a simple arm. The users are divided into the local site user and remote site user; the former wears an HMD and controls the remote camera by turning the head. Live images are displayed on the HMD at the local site.

Kashiwabara et al. [3] developed a system called Teroos, which involves a wearable avatar to enhance the feeling of participation in joint activities between local and remote users. The avatar is remote-controlled by the local user; a pan-tilt camera and a rotatable eye (for virtual expression of the eyes) are mounted on the avatar, to provide a sense of presence for the remote user.

Compared to Telecommunicator and Teroos, our research focused on the interaction method of “facing direction,” and our system makes use of the facing directions of both indoor and outdoor users. The “joint attention” method helps the users focus on the same object together. In our system, the indoor user can view the remote surroundings easily by turning his head and the process does not involve keyboard or mouse control. Such an immersive space involving an HMD provides the indoor user with a greater telepresence.

8. Conclusions and future work

We defined the concept of *Out Together Feeling* and implemented our prototype system called *WithYou* to identify ways to enhance this feeling. To verify this feeling, we designed and performed three experiments to examine what types of interaction (communication skills) people take when they actually go outside together, when they go out together virtually via videophone, and when they go out together virtually using our prototype system (*WithYou*). Based on the experimental results, we attempted to find the best interaction methods to achieve *Out Together Feeling*. We plan to improve our system in future studies based on these results.

Our prototype system, *WithYou*, was highly evaluated by the subjects. However, we also found that at present the outdoor user cannot enjoy the same level of *Out Together Feeling* as the indoor user.

In future studies, to further enhance the *Out Together Feeling*, we plan to implement two new features: a method for allowing “pointing with a finger” and a means of allowing the indoor user and outdoor user to see each other via the camera.

9. References

[1] T. Chang, S. Takahasi, J. Tanaka,"WithYou - A Communication System to Provide Out Together Feeling", Advanced Visual Interfaces 2012 International Working Conference, 2012. pp. 320-323.

[2] Y. Tsumaki, Y. Fujita, A. Kasai, C. Sato, D. N. Nenchev and M. Uchiyama, "Telecommunicator: A Novel Robot System for Human Communications", Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication, 2002, pp. 35-40.

[3] T. Kashiwabara, H. Osawa, K. Shinozawa, and M. Imai, "TEROOS: A Wearable Avatar to Enhance Joint Activities", CHI'12, ACM, Austin, Texas, USA, 2012, pp. 2001-2004.

[4] S. Ota, T. Yukioka, K. Yamazaki, A. Yamazaki, H. Kuzuoka and H. Matsuda, "Remote Instruction and Support Using a Shared-View System with Head Mounted Display (HMD)". Japan Science and Technology Agency(in Japanese), 2000, pp. 1-7.

[5] S. Oyama, H. Kuzuoka, L. Paul and K. Yamazaki, "Supporting the remote instruction by a mobile robot which embodies an instructor", INFORMATION PROCESSING SOCIETY OF JAPAN(in Japanese), 2001, pp. 83-88 .