

Transcendent Communication: Location-Based Guidance for Large-Scale Public Spaces

Hideyuki Nakanishi¹, Satoshi Koizumi², Toru Ishida^{1,2}, Hideaki Ito¹

¹Department of Social Informatics, Kyoto University

²JST CREST Digital City Project

Kyoto 606-8501, JAPAN

nakanishi@i.kyoto-u.ac.jp, satoshi@digitalcity.jst.go.jp, ishida@i.kyoto-u.ac.jp, ito@kuis.kyoto-u.ac.jp

Abstract

Many studies have been conducted on supporting communication in home and office spaces, but relatively few studies have explored supporting communication in large-scale public spaces, despite the importance of such environments in our daily lives. We propose a transcendent means of communication as an emerging style in this pervasive computing era: a system that allows administrative staff to effectively help visitors in large-scale public spaces. The visitors' context is used to provide a bird's-eye view of a simulated public space for the staff to grasp the situation and point at a particular location within the view to indicate the visitors they intend to address. The results of an experiment showed synergic effects between the bird's-eye view and the first-person one in determining the spatial movements of people. In indoor and outdoor large-scale public spaces, a central railway station and a park, we installed our prototypes and learned the implications of its use.

Categories & Subject Descriptors: H.5.3. [Information Interfaces and Presentation (e.g., HCI)]: Group and Organization Interfaces - *Computer-supported cooperative work*.

General Terms: Design, Human Factors

Keywords: Transcendent communication, simulated space, public space, bird's-eye view, visual communication, station, park.

INTRODUCTION

Our living space consists of home, office and public spaces. The studies on visual communication have predominantly focused on the first two spaces. Consequently, the primary issue of these studies is how to use computer network technologies for connecting distributed spaces [3][6][14][20], and a recent additional issue is how to use the technologies for enhancing collocated spaces [11]. We tackle a third but

increasingly important issue—how to use the technologies to secure and enrich large-scale public spaces. Such spaces have characteristic participants: administrative staff and visitors passing through. Every visitor is a candidate for on-demand guidance. Their efficient communication is the key to comfortable large-scale public spaces. Our research goal is to create a new communication style to connect the staff with the visitors.

Our approach is to combine the monitoring and announcement facilities seamlessly in large-scale public spaces. Monitoring facilities are usually used only for surveillance. We propose a technique to use the monitored view of a public space for guidance that is basically one-way announcements but can also be bidirectional communication. The view can be used to grasp the situation and determine the appropriate spot and content of guidance. If the addressed visitors could be specified and limited by controlling the volume of each speaker or using some special devices to make local audio channels, it would be possible to devise a system that gives a location-based guidance with custom information for each visitor. In contrast to a conventional one-way uniform announcement, location-based guidance would enable the staff of a public space to give more site-specific services to the visitors.

For location-based guidance, the single global view of a large-scale public space is better than a collection of fragmented views from surveillance cameras. Recently, 3D graphics hardware is becoming inexpensive, and visual simulation technologies are rapidly advancing. Since a visually simulated public space can provide a more flexible and consistent view than surveillance cameras, we adopted the bird's-eye view of a simulated space as the monitoring interface. By using this interface, the staff can indicate the spot of communication by pointing at a particular location in the view. Since this is a novel communication style that is different from any conventional visual communication, we named it *transcendent communication*.

The simulated space has to be rendered based on the context information retrieved from the real world. The minimum information that is necessary for transcendent communication is the positions of visitors. The positions can show the spatial movements of visitors and specify the expected communication participants. Thus, our prototypes of transcendent communication interfaces display human fig-

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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2004, April 24–29, 2004, Vienna, Austria.

Copyright 2004 ACM 1-58113-702-8/04/0004...\$5.00.

ures representing the positions. The prototypes can be a pointable map to establish the audio communication channels with visitors. In the Digital City project [12], we installed the prototypes in two real-world large-scale public spaces.

An indoor example of a large-scale public space is a central railway station. Crisis management is vital for securing lives in such indoor large-scale public space as a central railway station. It is known that leadership is influential in the safety of escaping crowds [19][29]. Transcendent communication has considerable significance for emergency evacuation since it brings a distributed fashion to evacuation guidance announcements and makes it possible to lead multiple groups separately at once. Such usual facilities as surveillance cameras and announcement speakers have little capability for distributed leading. Kyoto Station in Kyoto City, Japan is a central station where the number of visitors per day is more than 300,000. Thus, we developed our first prototype as an evacuation guidance system and installed it in the station to confirm the usefulness of transcendent communication. This trial can also give the visual simulators of emergency situations [10] an additional value as an evacuation guidance system.

An outdoor example of a large-scale public space is a park. Environmental education is increasingly important in school education [7]. Using computer network technologies in environmental education is a highly practical concept [28]. Kyoto University's Experimental Forest was a good testbed because it is like a park and it's also a facility of the university. Thus, we developed our second prototype as an environmental education support system and installed it in the forest.

RELATED WORKS

FTF-oriented Studies

Both transcendent and visual communications are supported by rich communication media that can transmit awareness information. However, the design goals of the media are different from each other.

Since visual communication is one of the most frequently tackled problems in HCI and CSCW, many approaches have been proposed. Those studies proposed various designs and technologies but share the same goal, which is the reproduction of face-to-face (FTF) communication environments. For example, media space research tried to connect distributed office spaces [3]. Telepresence and shared workspace research explored a way to integrate distributed deskwork spaces [6][14]. Spatial workspace collaboration research dealt with spatially configured workspaces [20]. CVE research proposed using virtual environments as virtual workspaces [2]. We agree that the elimination of the geographical barrier of FTF communication should be one of the ultimate goals of computer network technologies. This kind of efforts still continues [16]. However, instead of sticking to this traditional goal, we are

pursuing a new communication style that is inherently supported by the technologies and will engender an emerging application field.

Transcendent communication exploits the recent advances in wireless communication infrastructures and sensor devices. Wireless communication infrastructures have already been widely adopted. GPS, Radio Frequency Identification (RFID) tags, vision sensors, magnetic sensors, ultrasonic sensors and other sensor devices are available at much lower prices than ever. These devices can be used to construct perceptual user interfaces [26], which can continuously track the real world. These technological advances make it possible to go beyond the FTF-oriented designs of visual communication.

Ubicomp and MR

Transcendent communication is driven by the real-world context. Both Ubiquitous Computing (Ubicomp) and Mixed Reality (MR) studies proposed context-driven communication, too. However, these studies seem to aim at the augmentation of FTF communication rather than the redesign of human communication.

Both context-aware communication [21] and transcendent communication need the context information of people to establish communication channels among them. However, the user of the information is different. In context-aware communication, the system uses the information to automatically decide which devices, e.g. a speaker and a microphone, should be activated in what manner. A caller still specifies the ID, which is a name, of a callee to begin a conversation. On the other hand, a caller uses the context information in transcendent communication. The information is visualized in the simulated space and the caller can begin a conversation by specifying the location of a callee in the space.

Both social MR [5][8] and transcendent communication provide desktop users with a 2D or 3D map that displays the positions of mobile users. However, that map plays a different role in forming social interaction. In a social MR system, the map brings the spatial metaphor of FTF-like proxemics to a desktop-mobile interaction. This means that both desktop and mobile users become avatars in the same MR space. On the other hand, the map provides the context information of mobile users for desktop users in transcendent communication. The map enables a desktop user to glance over the real world where mobile users exist.

DESIGN OF TRANSCENDENT COMMUNICATION

In transcendent communication, a user watches the bird's-eye view of the real world to grasp its situation and point at a particular location within the view to indicate a person or people to talk to. Figure 1 is a conceptual drawing. In the figure, a user looks down at the real world and talks to selected people.

We argue that transcendent communication is different from conventional visual communication that aims at the

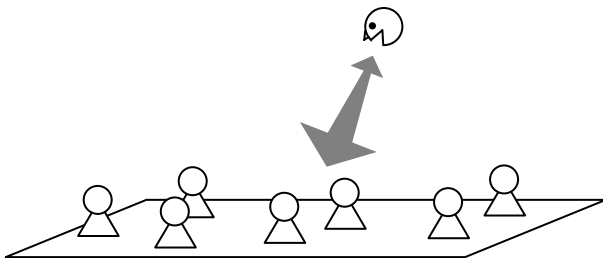


Figure 1. Conceptual transcendent communication

reproduction or augmentation of FTF communication. To show this difference, we compare FTF, visual, and transcendent communication in two aspects as illustrated in Figure 2.

The first aspect is the state of communication participants. In visual communication, participants can be geographically distributed, while they are collocated in FTF communication. In transcendent communication, it is not a significant distinction for participants to be distributed or not. The important distinction is the two layers of participants. A participant who looks down at the real world is a participant in the *transcendent mode*, while a real-world participant is in the *immanent mode*. Both participants are generally distributed but could also be collocated.

The second aspect is the type of communication space. In visual communication, a virtual space provides such social cues as proxemics and eye contact [24]. These cues are necessary to establish FTF-like communication. In FTF communication, a real space inherently mediates the cues. In transcendent communication, those cues are not the triggers of communication. The situation represented by the simulated space determines to whom a transcendent participant should talk. Transcendent communication depends on a surrounding environment of immanent participants rather than each individual of them.

In conventional visual communication, participants should be basically reciprocal. Privacy is an important issue and intrusiveness should be avoided [4][30]. In transcendent communication, the asymmetry of participants allows a transcendent participant to be an intrusive observer. This characteristic is effective when a small number of online experts like the staff of a public space have to help many real-world people like the visitors.

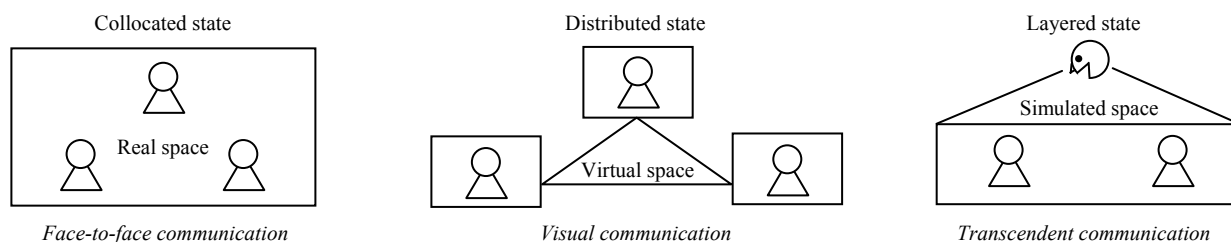


Figure 2. Comparison of FTF, visual, and transcendent communication

EVALUATION

Goal

We conducted a laboratory experiment to understand transcendent communication before designing prototypes. Transcendent communication interfaces provide the ability to grasp a situation through the bird's-eye view and interact with people who have first-person views. This design distinguishes the interfaces from conventional visual communication interfaces that provide only a subjective view, which is usually a first-person view. Therefore, we tried to confirm this distinctive nature in the experiment.

We investigated how much the combination of the bird's-eye observation and the first-person experience could contribute to the understanding of the crowd behavior. We used a 3D visual simulator *FreeWalk* [23] and a scenario description language Q [13] to construct an evacuation simulation. As shown in Figure 3, transcendent subjects observed the simulation and immanent subjects participated in it as avatars.

Hypothesis

Our intuition is that a bird's-eye view is more effective in understanding crowd behavior as well as navigation [9] than a first-person view. However, both views may have different efficacies. To derive their synergic effects, we tested a combination of them in both orders. We compared four groups: experiencing a first-person view (FP group); observing a bird's-eye view (BE group); experiencing a first-person view before observing a bird's-eye view (FP-BE group); and observing a bird's-eye view before experiencing a first-person view (BE-FP group). The subjects are 96 college students. They are divided into the four groups. Six subjects participated in the simulation at once. So, four simulations were conducted in each group.

Measure

The previous experiment [29] gave us a gauge to measure subjects' understandings of crowd behavior. This study demonstrated how the following two group leading methods cause different crowd behaviors.

Follow-direction method: The leaders point their arms at the exit and shout out, "the exit is over there!" to indicate the direction. They begin escaping after all evacuees go out.

Follow-me method: To a few of the nearest evacuees, the leaders whisper, "follow me" and proceed to the exit. This behavior forms a flow toward the exit.

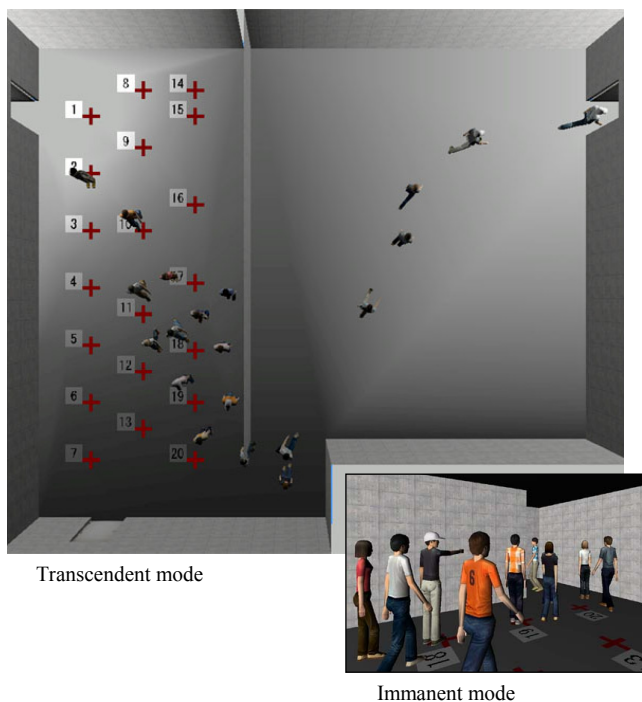


Figure 3. Evacuation Simulation System

The simulation was constructed based on this study. At the beginning of the simulation, everyone was in the left part of the room, which was divided into left and right parts by the center wall as shown in Figure 3. The four leaders had to lead the sixteen evacuees to the correct exit at the right part and prevent them from going out through the incorrect exit at the left part. In the FP simulations, six evacuees were subjects and the others were software agents. In the BE simulations, all evacuees and leaders were agents.

In the experiment, subjects observed and experienced the two different crowd behaviors caused by the two methods. We used the caused behaviors as questions and the causal methods as answers. In a quiz including 17 questions, subjects read each description of crowd behavior and chose one of the two methods. They answered the quiz before and after the experiment. We used a t-test to find significant differences between the scores of pre- and post-quizzes. A significant difference meant that the subject could learn the asked nature of crowd behavior through his or her observation and experience.

Results

Table 1 summarizes the results of the t-test on nine questions. Since no group could learn the other eight questions, they are omitted. Even though the results depend on the design of the quiz, it seems clear that a bird's-eye observation was necessary to grasp the crowd behavior. The FP group could not learn the questions from no. 3 to 9 that were related to the behaviors of evacuees. However, a first-person experience is not worthless. It is interesting that the BE-FP group could learn the questions no. 6 and 7 that could not be learned by the BE and FP-BE groups. The questions seem to be related to the dense nature of crowd behavior. This result implies that the background knowledge of the overall behavior enabled subjects to infer the gathering behaviors from the denseness they felt in their first-person experiences.

The conclusive result is that a bird's-eye view is effective in understanding the spatial movements of many people, and this understanding can convert the feeling created by a first-person view into further understandings. It can be said that communication between transcendent and immanent participants is meaningful to understand the situation.

PROTOTYPES

Based on the results that showed the value of transcendent communication, we designed prototypes. We deployed two prototypes in the indoor and outdoor large-scale public spaces to demonstrate the applicability of the transcendent communication model.

Central Railway Station – Indoor Large-scale Public Space

Figure 4 is a snapshot of our evacuation guidance system and the escaping passengers on a station platform with their mobile phones held. You can see a pointing person who stands in front of a large-scale touch screen. Suppose that this person is a leader working at an emergency operations center. The screen displays the bird's-eye view of the simulated station drawn by FreeWalk. The station is a static model while human figures are animated pedestrians whose walking paths are synchronized with those of the escaping passengers.

In this snapshot, the leader is pointing at a human figure, which represents one of the passengers. When the screen detects this pointing operation, the audio connection han-

Table 1. Summary of the results of the quiz (one-sided paired t-test)

No.	Question (the answer of all items is the follow-me method.)	FP	BE	FP-BE	BE-FP
1	Leaders are the first to escape.	4.3***	2.2*	2.3*	4.0***
2	Leaders do not observe evacuees.	2.8**	4.4***	4.0***	4.2***
3	Leaders escape like evacuees.	1.6	2.2*	1.9*	2.9**
4	One's escape behavior is caused by others' escape behavior.	1.2	2.1*	3.3**	2.9**
5	Nobody prevents evacuees from going to the incorrect exit.	1.6	4.9***	3.7***	4.5***
6	Evacuees follow other evacuees.	1.3	1.0	0.7	2.1*
7	Evacuees form a group.	1.6	0.5	1.2	1.9*
8	Leaders and evacuees escape together.	0.7	2.0*	0.2	3.4**
9	Evacuees try to behave the same as other evacuees.	0.9	2.5*	1.5	0.9

*p<.05, **p<.01, ***p<.001 (df=23)

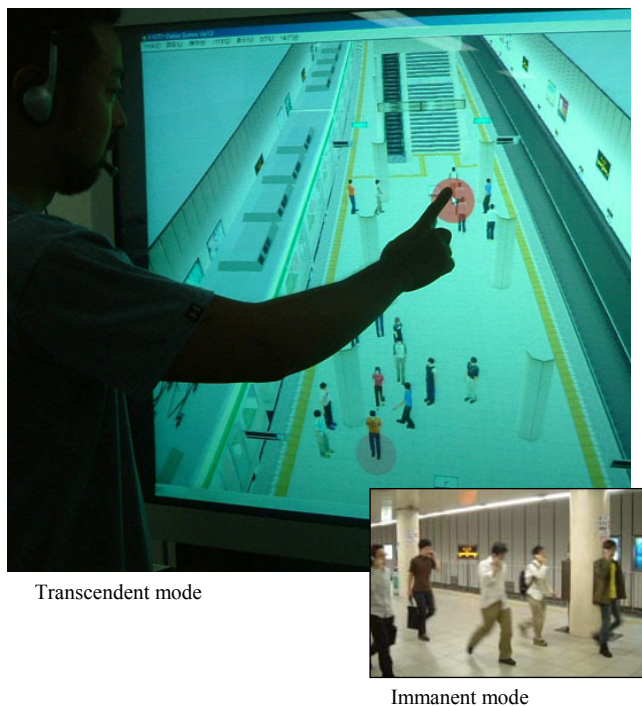


Figure 4. Evacuation Guidance System

When the leader of the system immediately activates the connection between the leader's headset and the passenger's mobile phone. This trick is possible because the headset is connected to the PC that is equipped with a special interface card. The card can control audio connections between the PC and several telephone lines. The simple coupling between pointing operation and audio activation makes it easy for the leader to begin and close a leading instruction.

The leader can talk to anybody and also any group. For example, a dragging operation can be used to indicate a rectangular area that includes a group of human figures. Furthermore, a pointing operation can be easily modified for placing a circular area of a certain size to indicate the figures. When the leader chooses a group, the audio connection handler activates several connections simultaneously. It is also possible that a group of leaders collaborate in leading a large number of crowds. Each leader can observe his or her own bird's-eye view of the same simulated station because FreeWalk has a multi-user capability. The leaders may talk to the same person simultaneously. Since a mobile phone cannot establish multiple audio connections at once, the audio connection handler plays the role of a mixer that compounds the voice streams of the leaders.

We also designed a smarter interaction, which is closer to the conceptual drawing of Figure 1. In Figure 5, a user wears an eye-tracking device and indicates an addressed person by gazing instead of touching. In our experience of experimental use, the coupling between gazing and talking gives a much more seamless feeling than that between touching and talking. However, that seamless feeling disappears when a user tries to indicate a group. The reason is that a user tends to control his or her eyeballs rather than



Figure 5. Seamless coupling between gazing and talking

gaze when he or she places a circular area to include the intended human figures. The center of a circular area is usually a blank spot. It is reported that gazing at a blank spot is less natural than gazing at a figure [15].

Another difference between the two interaction techniques is the number of pointing devices. Gaze is a single device while a large-scale touch screen enables a user to use two devices, i.e. his or her two hands. Even though the screen can detect that a single spot is touched at once, the two hands are more efficient than a hand or a gaze.

In the preliminary analysis of transcendent communication, we found that the leader used objective and subjective ways of guidance appropriately. In the objective way, the leader indicated the destination by a landmark, e.g. a stairs. The leader also indicated the relative direction, e.g. forward and backward directions, based on the subjective viewpoint of the passenger.

Vision sensor network

We installed a vision sensor network as a positioning system in Kyoto Station. We installed 12 sensors in the concourse area and 16 sensors on the platform. In Figure 6, the black dots on the floor plan show their positions and the white circles on the pictures show how they are installed. The vision sensor network can track passengers between the platform and the ticket gate.

In Figure 7(a), you can see a CCD camera and a reflector with a special shape [22]. If we could expand the field of view (FOV) of each camera, we could reduce the number of required cameras. However, a widened FOV causes minus (barrel) distortion in the images taken by conventional cameras. The reflector of our vision sensor can eliminate such distortion. The shape of the reflector can tailor a plane that perpendicularly intersects the optical axis of the camera to be projected perspective to the camera plane. As shown in Figure 7(b), this optical contrivance makes it possible to have a large FOV without distortion. Figure 7(c) is a screenshot of the simulated passengers synchronized with retrieved positions.

The installed facilities include 28 vision sensors, 7 quad processors, 7 PCs for image processing, and a PC for tra-

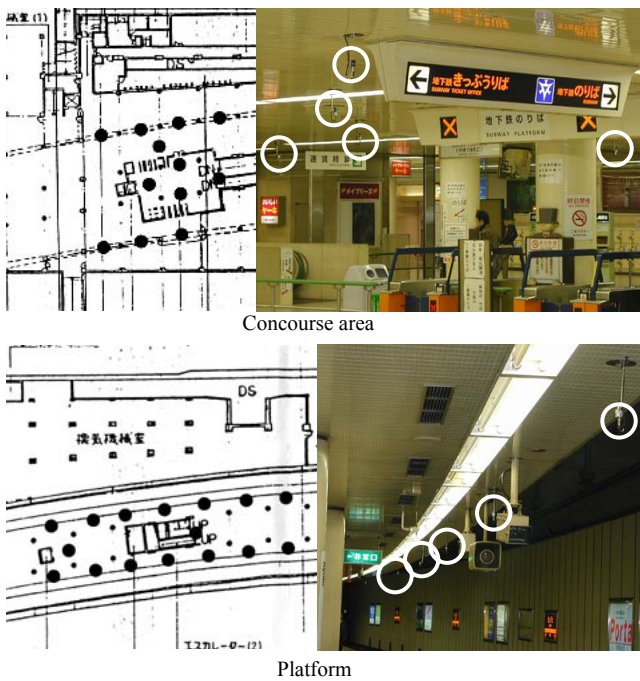


Figure 6. Vision sensor network in the railway station

jectory detection. The method of processing images is as follows. First, a quad processor assembles the images captured by four sensors into one video image and sends it to the image-processing PC. Next, the PC extracts the regions of moving objects by the background subtraction technique and sends the results to the trajectory-detection PC. Finally, the PC detects the positions of the moving objects based on such geographical knowledge as the positions of the cameras, the occlusion edges in the views of the cameras, and the boundaries of movable areas.

A vision sensor network can work well in indoor environments and track the more fine-grained positions of people than GPS. Furthermore, that is a device-free approach since people do not need to carry any devices to be tracked. This characteristic also means that the simulated station can visualize all escaping crowds, including people who are not explicitly the users of our system. Our vision sensor network is currently used to only track movements due to the limited resolution of the cameras. In the future, it may become possible for higher-resolution cameras to capture

richer context information to make a simulated space more realistic and informative.

3D visual simulator

We connected the 3D visual simulator FreeWalk with the vision sensor network through the Internet. The positions captured by the sensors are transmitted to FreeWalk every millisecond. FreeWalk uses physical and social rules to synchronize robustly the movements of human figures and passengers. Based on the captured positions, FreeWalk generates the next step of the corresponding human figure. This next step is modified according to the social rules described in Q language. Examples of rules are such flocking behaviors as following others and keeping a fixed distance from others [27] and such cultural behaviors as forming a line to go through a ticket gate and forming a circle to have a conversation [18]. Then, the next step is modified again based on the pedestrian model to avoid collision with others, walls, or pillars [25]. Finally, gait animation is generated based on the hybrid algorithm of kinematics and dynamics [31].

We constructed the 3D model of Kyoto Station. FreeWalk needs the model as a graphical model to draw a simulated space and also as a geometric model to simulate crowd behaviors. To reduce the building cost and fill these requirements, we combined pictures taken by digital cameras with a simple geometric model based on the floor plan. The simple model also helps to reduce the workload in the collision detection calculation between human figures and the station model.

Audio connection handler

Since almost everybody is always carrying a mobile phone, we used it as a device for an escaping passenger to communicate with a leader. Even though it is well known that a wireless telephone facility easily becomes unusable in emergency situations, this facility is currently the most practical device for our system. We developed the audio connection handler to control telephone lines between our system and mobile phones. FreeWalk informs the handler which human figure is pointed at. The handler knows which line corresponds to the figure and activates the line to establish an audio communication channel.

A question that arises is how to know the correspondence

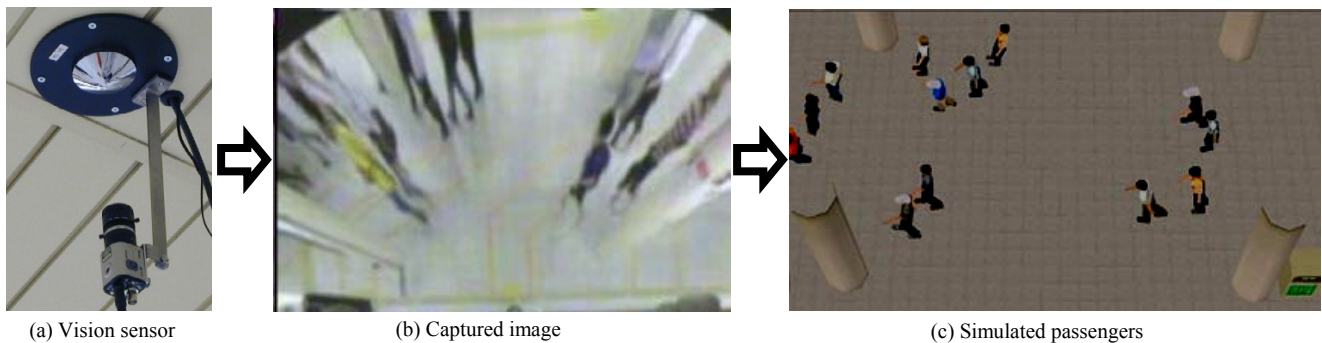


Figure 7. Visualization of tracked passengers on a station platform

between telephone lines and human figures. There are several possible techniques to resolve this registration problem. Currently, our system uses a manual calibration method in which an experimenter assigns a telephone number to each human figure. A better method is for everyone to call the system at an appointed spot so that the system can know which figure is the current caller. Obviously, both methods are not practical. A straightforward improvement is to use such detection methods as face detection, voice detection, and so on. However, those methods have technological and social difficulties. It is a practical solution to attach a tag or a sensor to each passenger's mobile phone so that he or she is registered automatically. The array of ultra-directional speakers and microphones on the ceiling may be a more promising solution because that array makes mobile phones unnecessary.

Park – Outdoor Large-scale Public Space

The implementation of our first prototype needs some state-of-the-art technologies, e.g. a vision sensor network and a 3D visual simulator. There is a study of using similar configuration for simulating the real world [17]. However, transcendent communication can be supported by using widely spread technologies, e.g. GPS and a map. Our second prototype is an example of that approach.

Figure 8 is pictures of the environmental education support system [1] and the learners in a park. In an environmental education class, a remote teacher can use this system to monitor learners walking around a park. The background picture of the system is a simple map of the park. The flag icons on the map represent plants that should be learned in the class. The cap icons move around based on the positions captured by the GPS receivers attached to the PDAs carried by the learners. Those icons inform the teacher about the positional relationship between the learners and the plants. This information enables the teacher to know what he or she should talk about to instruct a learner.

In the system, the combination of GPS and a map works well because the rough positions of the learners are sufficient to know who is approaching which plant. We installed five wireless LAN access points in the experimental forest for the system to gather the positions. The same tool set of GPS and a map can be applied to urban areas. In an urban area, it is not necessary to prepare our own wireless LAN facilities because such wireless data communication services as Personal Handy-phone System (PHS) and hot spot locations are available.

Currently, the audio connection handler is not embedded in this system. We tested only the monitoring capability. The test of full functionality is our future work.

CONCLUSION

We created transcendent communication interfaces for large-scale public spaces. These interfaces operate between administrative staff and visitors. Our interfaces provide the staff the ability to grasp the behaviors of visitors in a public

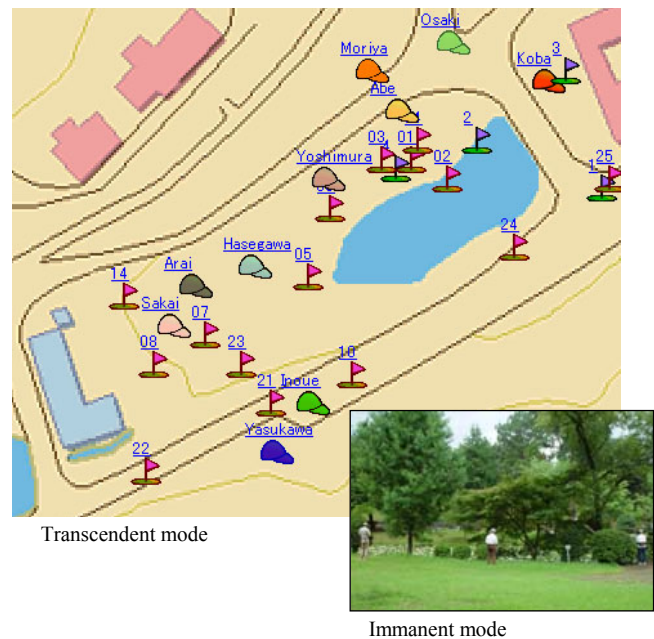


Figure 8. Environmental Education Support System (by Abe and Sakai)

space in order to communicate with them and give location-based guidance. We confirmed this ability in a statistical analysis of synergic learning effects between the transcendent and immanent modes of participation in a crowd behavior simulation. We also confirmed that in an experimental use of indoor and outdoor prototypes of the interfaces installed in the real-world large-scale public spaces. Those prototypes show examples of both experimental and practical implementations. Our future work includes the design of an appropriate use of spatial language in transcendent communication.

ACKNOWLEDGMENTS

We thank Kyoto City Municipal Transportation and General Planning Bureaus for their cooperation, Tetsuro Sakai and Mitsutoshi Abe for providing the environmental education support system, Hiroshi Ishiguro for deploying the vision sensor network, Toshio Sugiman, Shigeyuki Okazaki, Ken Tsutsuguchi for constructing the evacuation simulation, Reiko Hishiyama, Tomoyuki Kawasoe, Toyokazu Itakura, CRC solutions, Mathematical system, and CAD center for developing the evacuation simulation and guidance systems.

REFERENCES

- [1] Abe, M., Yoshimura, T., Hasegawa, N., Osaki, T., Yasukawa, N., Koba, K., Moriya, K. and Sakai, T. Development and Evaluation of a Support System for Forest Education, *Journal of Forest Research*, (submitted).
- [2] Benford, S., Greenhalgh, C., Rodden, T. and Pycock, J. Collaborative Virtual Environments. *CACM*, 44(7), (2001), 79-85.

- [3] Bly, S.A., Harrison, S.R. and Irwin, S. Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment *CACM*, 36(1), (1993), 28-47.
- [4] Borning, A. and Travers, M. Two Approaches to Casual Interaction over Computer and Video Networks. *Proc. CHI91*, (1991), 13-19.
- [5] Brown, B., MacColl, I., Chalmers, M., Galani, A., Randell, C. and Steed, A. Lessons From The Lighthouse: Collaboration In A Shared Mixed Reality System. *Proc. CHI2003*, (2003), 577-584.
- [6] Buxton, W. Telepresence: Integrating Shared Task and Person Spaces. *Proc. Graphics Interface 92*, (1992), 123-129.
- [7] Dresner, M. Teachers in the Woods: Monitoring Forest Biodiversity. *Journal of Environmental Education*. 34(1), (2002), 26-31.
- [8] Flintham, M., Anastasi, R., Benford, S., Hemmings T., Crabtree, A., Greenhalgh, C., Rodden, T., Tandavanitj, N., Adams, M. and Row-Farr, Ju. Where On-Line Meets On-The-Streets: Experiences With Mobile Mixed Reality Games. *Proc. CHI2003*, (2003), 569-576.
- [9] Fukatsu, S., Kitamura, Y., Masaki, T. and Kishino, F. Intuitive Control of "Bird's Eye" Overview Images for Navigation in an Enormous Virtual Environment. *Proc. VRST98*, (1998), 67-76.
- [10] Hareesh, P.V., Shibano, N., Nakanishi, H., Kashiwagi, M. and Sawada, K. Evacuation Simulation: Visualization Using Virtual Humans in a Distributed Multi-User Immersive VR System. *Proc. VSMM2000*, (2000), 283-289.
- [11] Huang, E.M. and Mynatt, E.D. Semi-Public Displays for Small, Co-located Groups. *Proc. CHI2003*, (2003), 49-56.
- [12] Ishida, T. Digital City Kyoto: Social Information Infrastructure for Everyday Life. *CACM*, 45(7), (2002), 76-81.
- [13] Ishida, T. *Q*: A Scenario Description Language for Interactive Agents. *IEEE Computer*, 35(11), (2002), 54-59.
- [14] Ishii, H., Kobayashi, M. and Arita, K. Iterative Design of Seamless Collaboration Media, *CACM*, 37(8), (1994), 83-97.
- [15] Jacob, R.J.K. What You Look at is What You Get: Eye Movement-Based Interaction Techniques. *Proc. CHI90*, (1990), 11-18.
- [16] Jancke, G., Venolia, G.D., Grudin, J., Cadiz, J.J. and Gupta, A. Linking Public Spaces: Technical and Social Issues. *Proc. CHI2001*, (2001), 530-537.
- [17] Kelly, P.H., Katkere, A., Kuramura, D.Y., Moezzi, S. and Chatterjee, S. An Architecture for Multiple Perspective Interactive Video, *Proc. Multimedia95*, (1995), 201-212.
- [18] Kendon, A. Spatial Organization in Social Encounters: the F-formation System. A. Kendon, Ed., *Conducting Interaction: Patterns of Behavior in Focused Encounters*, Cambridge University Press, (1990), 209-237.
- [19] Klein, A.L. Changes in Leadership Appraisal as a Function of the Stress of a Simulated Panic Situation. *Journal of Personality and Social Psychology*, 34(6), (1976), 1143-1154.
- [20] Kuzuoka H. Spatial Workspace Collaboration: a SharedView Video Support System for Remote Collaboration Capability. *Proc. CHI92*, (1992), 533-540.
- [21] Nagel, K., Kidd, C.D., O'Connell, T., Dey, A.K. and Abowd, G.D. The Family Intercom: Developing a Context-Aware Audio Communication System. G.D. Abowd, B. Brumitt, S.A.N. Shafer Ed., *Ubicomp2001, LNCS2201*, Springer-Verlag, (2001), 176-183.
- [22] Nakamura, T. and Ishiguro, H. Automatic 2D Map Construction using a Special Catadioptric Sensor. *Proc. IEEE/RSJ IROS2002*, (2002), 196-201.
- [23] Nakanishi, H. FreeWalk: A Social Interaction Platform for Group Behavior in a Virtual Space. *IJHCS* (to appear).
- [24] Okada, K., Maeda, F., Ichikawa, Y. and Matsushita, Y. Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design. *Proc. CSCW94*, (1994), 385-393.
- [25] Okazaki, S. and Matsushita, S. A Study of Simulation Model for Pedestrian Movement with Evacuation and Queuing. *Proc. International Conference on Engineering for Crowd Safety*, (1993), 271-280.
- [26] Pentland, A. Perceptual Intelligence. *CACM*, 43(3), (2000), 35-44.
- [27] Reynolds, C.W. Flocks, Herds, and Schools: A Distributed Behavioral Model. *Proc. SIGGRAPH87*, (1987), 25-34.
- [28] Rohwedder, W.J. *Using Computers in Environmental Education: Interactive Multimedia and On-Line Learning*. Kendall/Hunt Publishing Company, (1994).
- [29] Sugiman, T. and Misumi, J. Development of a New Evacuation Method for Emergencies: Control of Collective Behavior by Emergent Small Groups. *Journal of Applied Psychology*, 73(1), (1998), 3-10.
- [30] Tang, J.C. and Rua, M. Montage: Providing Teleproximity for Distributed Groups. *Proc. CHI94*, (1994), 37-43.
- [31] Tsutsuguchi, K., Shimada, S., Suenaga, Y. Sonehara, N. and Ohtsuka, S. Human Walking Animation based on Foot Reaction Force in the Three-dimensional Virtual World. *Journal of Visualization and Computer Animation*, 11(1), (2000), 3-16.