FreeWalk: A 3D Virtual Space for Casual Meetings

Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura, and Toru Ishida
Kyoto University

A meeting environment for casual communication in a networked community, FreeWalk provides a 3D common area where everyone can meet and talk freely. FreeWalk represents participants as 3D polygon pyramids, on which their live video is mapped. Voice volume remains proportional to the distance between sender and receiver. For evaluation, we compared communications in FreeWalk to a conventional desktop videoconferencing system and a face-to-face meeting.

Most computer systems for collaborative work provide desktop videoconferencing tools for business meetings. However, meetings aren't always formal or business related. Casual meetings such as chatting during a coffee break or in a hallway occur daily. They also play an important role in collaboration. We believe that conventional desktop videoconferencing systems, which multicast pictures and voices, can't support casual meetings.

We aim to support everyday activities by forming a community through computer networks. Our product FreeWalk,[1] a social environment for communication, lets people meet casually in common three-dimensional (3D) virtual spaces such as a park or a lobby. The following list describes the inherent features of casual meetings and how FreeWalk can support them.

1. *Casual meetings*. In conventional desktop videoconferencing systems such as Office Mermaid,[2] participants turn on the system when they start a meeting. When in operation, the system displays the faces of all participants on their workstations, which hinders free conversation. The system lists the participants before the meeting starts, thereby prohibiting accidental encounters with other participants.

   Several desktop videoconferencing systems have tried to extend their functions to support casual meetings. Cruiser[3] randomly selects some of the participants and displays their faces to other participants to simulate accidental encounters. In contrast, FreeWalk's approach provides a common virtual space for casual meetings wherein participants can move and meet by themselves. It doesn't promote any system-directed encounters. The participants' faces display on screen only when the bodies of their avatars meet.

2. *Meetings with many people*. In meetings such as parties, several tens of participants simultaneously exist in the same space. In such cases, it's almost impossible to use desktop videoconferencing systems, since they try to display the faces of all participants at once. Plus, even if it were possible, it would be very hard for users to comprehend the situation.

   In FreeWalk, participants can freely change their locations and view directions. For example, they can wander around before they talk to someone else. They can also watch other participants.

Many systems realize a 3D shared virtual space. The Distibuted Interactive Virtual Environment (DIVE),[4] a multiuser platform, lets people create, modify, and remove objects dynamically. This system has a script language to define autonomic behaviors of objects. Another multiuser virtual environment, Diamond Park,[5] has a park, a village, and an open-air cafe. In addition, Community Place[6] integrates Virtual Reality Modeling Language (VRML) and has an online chat forum. InterSpace[7] supports audio and video communication for the experimental service CyberCampus, which features distance learning and online shopping. These systems aim to construct realistic virtual worlds containing many kinds of virtual objects such as mountains, oceans, buildings, artifacts, and so on.

We implemented a basic system to support casual meetings in a 3D virtual space that represents dynamic changes in people's locations during casual meetings. The role of 3D space in our system resembles the spatial model of interaction in Massive,[8] a VR-based conferencing system with text and audio communication.

Since researchers haven't sufficiently investigated social interactions in 3D virtual space, we conducted an experiment to determine the characteristics of interactions in FreeWalk.

Interaction design

Here we describe FreeWalk's design for interaction of the 3D community common area. We also discuss how FreeWalk supports casual group meetings.

1070-986X/99/$10.00 © 1999 IEEE

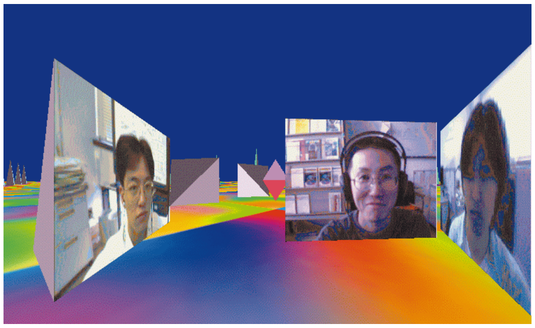


Figure 1. FreeWalk window.

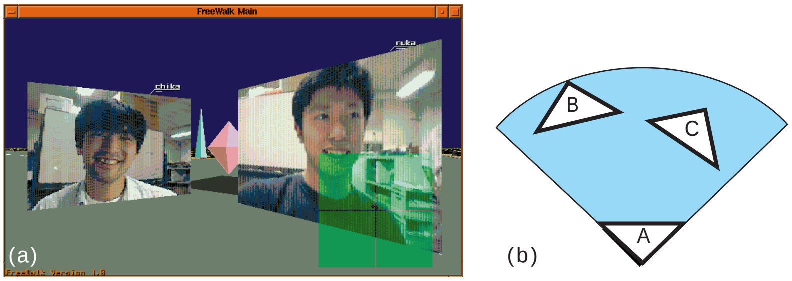


Figure 2. (a) Participant A's view of the community common area. (b) Map of B and C's locations.

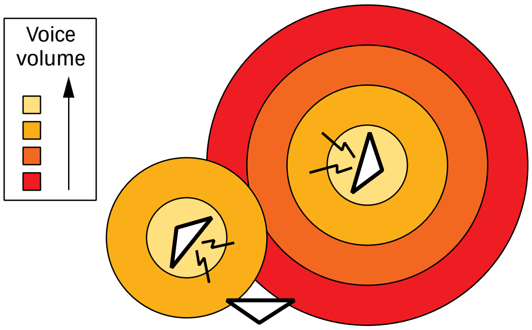


Figure 3. Voice transfer.

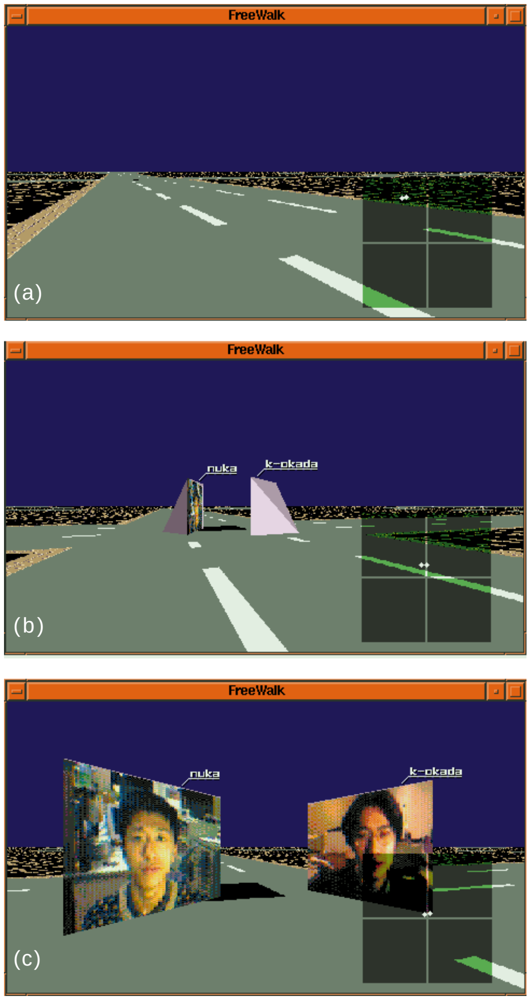


Figure 4. Accidental encounter. (a) Finding others on the radar screen. (b) Watching a talking pair. (c) Joining their conversation.

3D community common area

Figure 1 shows an image of a FreeWalk window. FreeWalk provides a 3D community common area where people can meet. Participants move and turn freely in the space using their mouse (just as in a video game). Locations and view directions of participants in the space determine which pictures and voices get transmitted.

In this 3D space, a pyramid of 3D polygons represents each participant. The system maps live video of each participant on one rectangular plane of the pyramid, and the participant's viewpoint lies at the center of this rectangle. The view of the community common area from a participant's particular viewpoint appears in the FreeWalk window. Figure 2a shows participant A's view when participants B and C are located as shown in Figure 2b.

Participants standing far away in the 3D environment appear smaller and those closer appear larger. FreeWalk doesn't display participants located beyond a predefined distance. The system also transfers voices under the same policy—that is, voice volume changes in proportion to the distance between sender and receiver. See Figure 3.

Simulating casual meetings

In FreeWalk, meetings can start with an accidental encounter. Figure 4 shows an example of an accidental encounter, where the user finds others on the radar screen displayed at the right bottom corner of the window (Figure 4a), watches them to find out what they're talking about (Figure 4b), then joins them (Figure 4c).

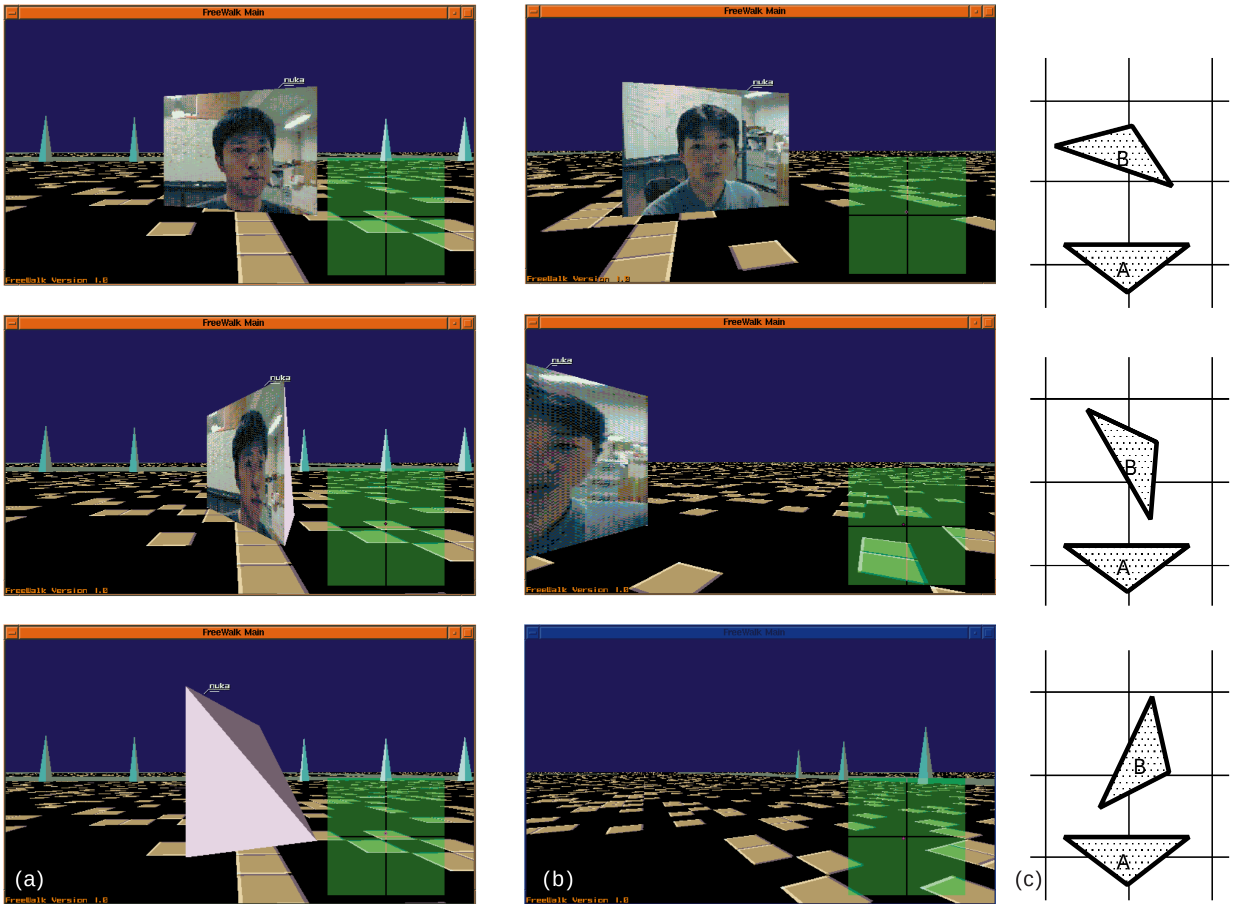


Figure 5. Changes of participants' view. (a) A's view of B, (b) B's view of A, and (c) a map of the views.

Since the participants' locations and view directions reflect a pyramid orientation, each participant can observe the distances or directions of other participants and what other people are doing from a distance. Participants can also observe others around them by turning their head. Figure 5 shows the view changes of participants A and B as participant B changes his direction in front of A.

Since distance attenuates voice, a participant must approach the others to talk to them. On the other hand, not only can the participants in the conversation hear the speaker's voice, but anyone in the neighborhood can listen. This mechanism forces people to combine actions and conversations in the space. People can smoothly join the conversation that attracts their interest, since they can guess the subject by listening to the conversation beforehand. People can exit a conversation by leaving a group and join a conversation by approaching another group.

Organizing meeting groups

Desktop videoconferencing systems provide various functions to support the organizational behavior of participants, such as speaker selection. Although these functions let participants manage multiple conversation threads in parallel, they also damage the freedom we're aiming for. FreeWalk doesn't take this approach. Instead, it uses a common 3D space that promotes a casual feeling in communication.

People form a group by standing close to each other to engage in conversation. Figure 6 shows this situation. Since voice volume attenuates in proportion to the distance between sender and receiver, people can have a confidential conversation by keeping away from others. If groups have enough distance between them, people in one group can't hear people in other groups. Therefore, participants can form separate meeting groups and not bother each other. This feature makes FreeWalk an effective tool for holding a party.

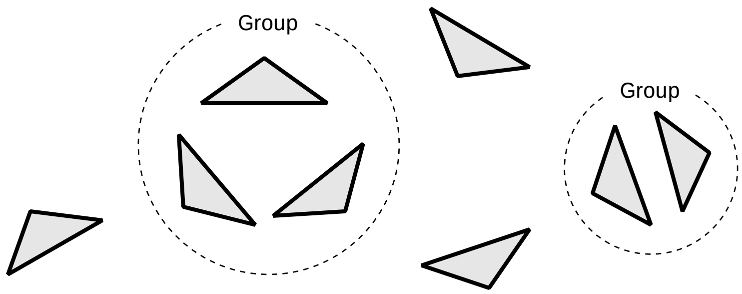


Figure 6. Meeting organization.

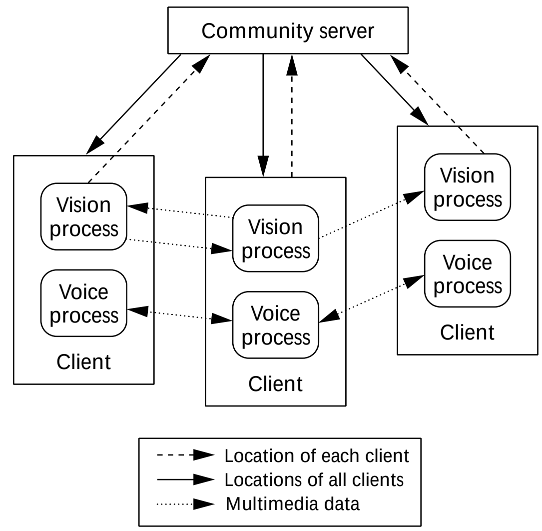


Figure 7. FreeWalk system configuration.

System design

Here we discuss FreeWalk's system design and implementation.

System configuration

The FreeWalk system consists of a community server and clients, each of which includes vision and voice processes. Figure 7 illustrates the interaction between the community server and clients.

When participants move in the 3D space using their mouse, the corresponding client calculates the new location and orientation, and sends them to the community server. The server then compiles this information into a list of client locations in the 3D community common area. The server finally sends the list back to each client for screen updating. Since only control information transfers between the server and the clients, the community server can efficiently maintain a global view of the ongoing activities in the community common area.

When a client receives the list of other clients, the client's vision system sends its owner's picture to the other clients. On receiving pictures from other clients, the vision system redraws the display based on the information in the list and the pictures received.

Because each client can't see all clients, it's not necessary for each one to send its picture to all others. Similarly, each client doesn't have to send full-size pictures to clients far away. FreeWalk uses these facts to optimize the bandwidth of video communication as follows:

- The sender adjusts the picture's size to the size the receiver needs.
- The client sends its picture to others who can see the client.

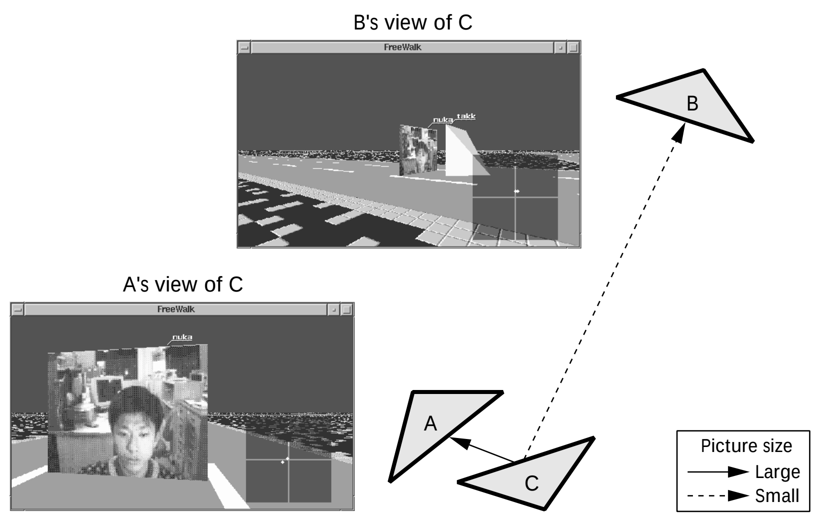


Figure 8. Video transfer among clients.

Figure 8 shows an example of a video transfer in FreeWalk. Since client A lies near client C, client C sends a large picture to client A. In contrast, client C sends a small picture to client B, because it's located far away.

Voice communication occurs in the same manner. FreeWalk clients don't send voice data to those clients located too far away to hear the participants' voices.

Using a large screen

We implemented the FreeWalk system on an immersive environment as well as on a desktop environment. We used a special room with a large-scale projector screen connected to a graph-

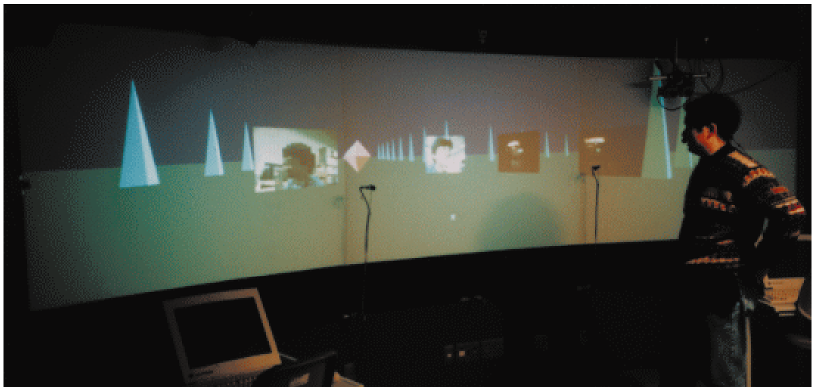


Figure 9. Virtual space on a large screen.

Figure 10. Three different environments for conversation: (a) FTF, (b) InPerson, and (c) FreeWalk.

ics workstation. Figure 9 shows a virtual space displayed on the large screen in the room.

Several people can simultaneously view the large virtual space displayed on the screen and talk to other people moving within that space. People using a desktop environment see the room represented as a larger pyramid. A large live video of the room visible in the space makes it easy to include the room and its participants in the virtual space.

Interaction analysis

Some earlier studies tried to compare communication aided by desktop videoconferencing systems to face-to-face (FTF) communication. Various characteristics of conventional video communication became clear through those studies.[9] However, the characteristics of the communication aided by a desktop videoconferencing system with a 3D virtual space remained unclear. In this section we show the characteristics of 3D communication compared to FTF and conventional video communications. We used Silicon Graphics' InPerson (see http://www.sgi.com/Products/software/InPerson/) as the conventional video environment and FreeWalk as the desktop videoconferencing system with a 3D virtual space.

Sellen compared communication in two video conferencing systems, Hydra and Picture-in-a-Picture (PIP), and in the FTF environment.[10] She found no differences among the three environments for conversation in terms of *turns* (transferring the initiative of speech), even though previous studies showed that more turns occurred in the FTF environment than in the videoconferencing environment. We expected that the number of turns might increase in casual meetings, so we analyzed the number of turns in our experiment.

In another study, Bowers investigated how the movement of avatars coordinated with conversation in a virtual environment.[11] Results showed that the avatars' moves transferred the initiative of conversation. In 3D and FTF environments, the moves of people relate to their communication skills. In our experiment, we analyzed the moves of people in meetings.

Additionally, we counted the number of occurrences of chat and calculated the standard deviation of utterance. We thought a casual atmosphere might stimulate the occurrence of chat and change the amount of utterance of each participant.

Design of experiment

Twenty-one undergraduate students participated in our one-day experiment. We prepared three environments for conversation to compare FTF, conventional video, and 3D communications (see Figure 10). We set up seven SGI O2 workstations connected by a 100-Mbps Ethernet for the video environment (InPerson) and the 3D environment (FreeWalk).

The meetings in the three environments consisted of three tasks: agreeing on a group travel destination, discussing social problems, and conversing freely. We chose these tasks to examine various types of communication comprehensively. For each task, we told participants to organize three groups of seven people.

Thus nine types of meetings took place. Each meeting lasted for 20 minutes. We didn't choose any chairpersons of the meetings in advance. Before performing the three tasks, the participants introduced themselves in each group so that they could memorize each other's faces and voices. They also practiced operating FreeWalk. The independent variables of this experiment were the differences between the environments and the tasks.

Data collection

We collected experimental data using videotape recordings. During the FreeWalk and InPerson meetings, we recorded the screen images of the workstations on videotape recorders. In FTF meetings, we recorded the scenes on 8-mm video. We reviewed the videotape pictures to record the start and end times of participants' utterances to create conversation records. In addition, we collected the system logs of FreeWalk to find the pattern of moves in the 3D virtual space during meetings. The FreeWalk community server stores system logs in which it records locations of participants in a 3D virtual space. We analyzed these system logs by drawing lines along the participants' moves and connecting their locations in sequence.

Results

The analysis results of participants' conversations and moves follow.

Conversation. We organized the analysis results of the conversations into number of turns, standard deviation of utterance, and occurrence of chat.

1. *Number of turns*. This value represents the number of events. Each event transfers the initiative of talking from one person to another. The turn occurs when someone starts talking immediately after or while another talks. We didn't count cases in which someone stopped talking and started talking again after a brief silence.

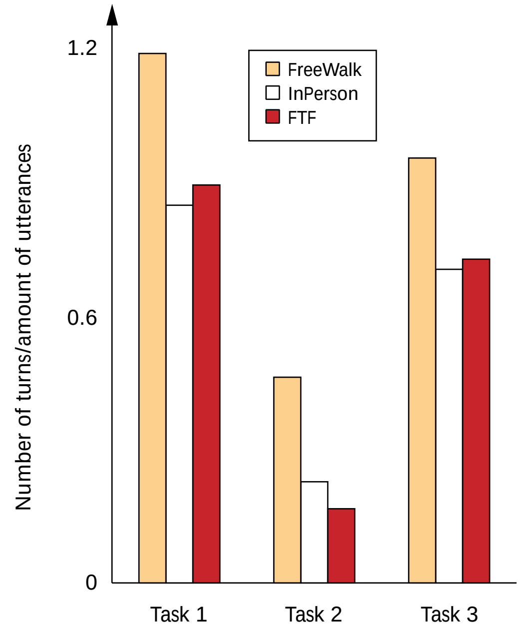


Figure 11. Frequency of turns.

Table 1. Standard deviation of utterance.

	Task 1	Task 2	Task 3
FTF	13.93	19.19	14.07
InPerson	12.31	15.97	17.25
FreeWalk	9.28	15.45	13.45

Figure 11 shows the clear relation between the frequency of turns and environments. The frequency of turns equals the number of turns divided by the amount of utterances. The rankings of contributions of environments to the number of turns follow:

FreeWalk > FTF ≈ InPerson

The effect of the difference in environments showed that FreeWalk activated turns more often than InPerson and FTF.

2. *Standard deviation of utterance*. This value represents the standard deviation of the ratio of the total time of utterances of each participant to the total time of all utterances of all participants. Table 1 summarizes the standard deviations of utterance. It also provides the following ranking of environments for each task:

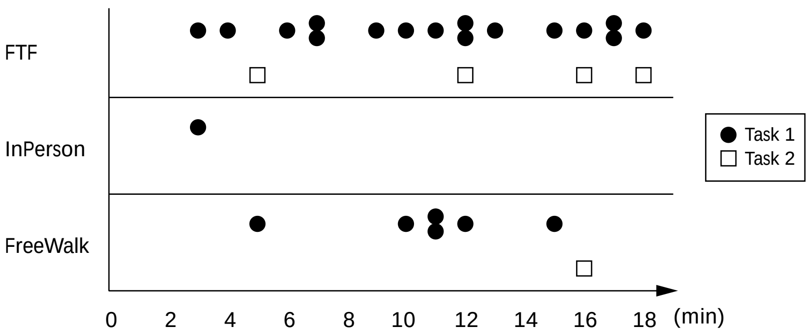


Figure 12. Occurrence of chat.

Task 1 FTF > InPerson > FreeWalk
Task 2 FTF > InPerson ≈ FreeWalk
Task 3 InPerson > FTF ≈ FreeWalk

Interestingly, the deviation remained the smallest in FreeWalk for all tasks. This means that the amount of utterances of each participant became equalized in FreeWalk.

3. *Occurrence of chat*. This value represents starting a conversation that doesn't contribute to accomplishing the task. Figure 12 shows the occurrence of chat in Task 1 and Task 2 in each environment. In Figure 12, the horizontal axis represents time, and each mark represents the occurrence of chat. Figure 12 shows that chat occurred more actively in FTF than in FreeWalk, while it seldom occurred in InPerson. The rankings of the contributions of environments to the occurrence of chat follow:

FTF > FreeWalk > InPerson

In FreeWalk, the atmosphere among participants might have been relaxed since they formed a circle to have a conversation, while in InPerson everyone faced the others.

Participants' moves. In FTF meetings, participants seldom moved after forming a circle to have a conversation. During InPerson meetings, everyone faced the others on the screen.

Figure 13 shows participants' moves during a 15-minute period in FreeWalk meetings. In Task 1 and Task 2, they seldom moved after forming a circle as in FTF. Unlike the other two tasks, they moved actively around the 3D virtual space in Task 3, as Figure 13c shows. In Task 3—free conversation—we observed the following behaviors:

1. *Moving in a 3D virtual space*. At the beginning of the task, participants moved actively. For example, they moved to the edge of the 3D virtual space and rushed toward others.

2. *Facing each other to greet*. In the middle of the task, participants faced one another frequently to greet. We noted that some participants blamed others for approaching them when they tried to whisper to each other.

3. *Gathering to start conversation*. Toward the end of the task, all participants gathered to converse. We noted that a certain participant ran about trying to escape from the meeting place since he was unwilling to talk, while another participant looked for someone else who had gone elsewhere.

Discussion

As a result of our analysis, we categorized the effects of a 3D virtual space into two types. In the first type, we observed that 3D communication resembles FTF communication. Two primary characteristics exist: frequency of chat and behavior of participants. The second category, however, remains peculiar to 3D virtual spaces. These environments equalize the amount of utterances for each participant more than the other environments, increase the number of turns, and sometimes stimulate participants to move around to converse freely.

The results show the effectiveness of a 3D virtual space in casual meetings. The freedom of 3D virtual space lets participants enjoy their conversation, and its relaxed atmosphere stimulates participants into initiating conversations. On the other hand, participants having a meeting in a 3D virtual space tend to concentrate less than in the other environments. MM

References

1. H. Nakanishi et al., "FreeWalk: Supporting Casual Meetings in a Network," *Proc. Int'l Conf. Computer Supported Cooperative Work (CSCW 96)*, ACM Press, New York, 1996, pp. 308-314.
2. K. Watabe et al., "Distributed Multiparty Desktop Conference System Mermaid: Platform for Groupware," *Proc. Int'l Conf. Computer Supported*

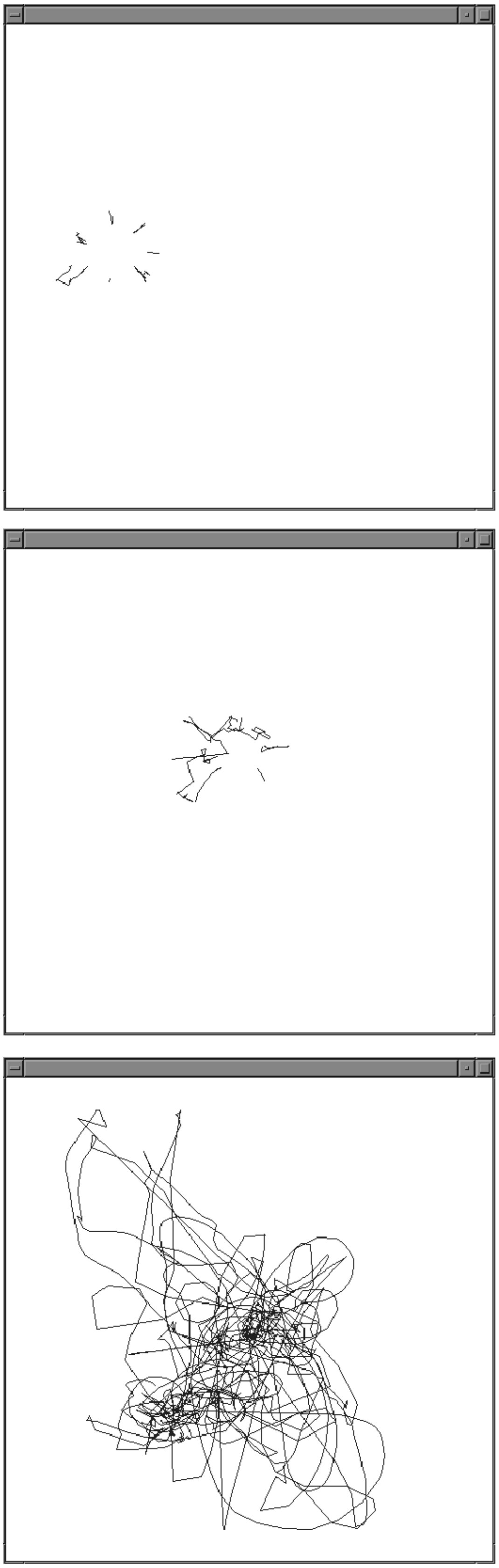
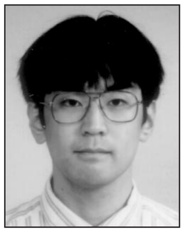


Figure 13. Pattern of moves in a 3D virtual space.

Cooperative Work (CSCW 90), ACM Press, New York, 1990, pp. 27-38.

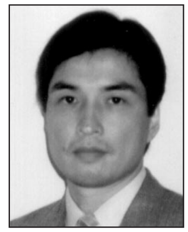
3. R.W. Root, “Design of a Multimedia Vehicle for Social Browsing,” *Proc. Int’l Conf. Computer Supported Cooperative Work (CSCW 88)*, ACM Press, New York, 1988, pp. 25-38.
4. O. Hagsand, “Interactive Multiuser VEs in the DIVE System,” *IEEE MultiMedia*, Vol. 3, No. 1, Spring 1996, pp. 30-39.
5. R.C. Waters and J.W. Barrus, “The Rise of Shared Virtual Environments,” *IEEE Spectrum*, Vol. 34, No. 3, 1997, pp. 20-25.
6. R. Lea et al., “Community Place: Architecture and Performance,” *Proc. Symp. Virtual Reality Modeling Language (VRML 97)*, ACM Press, New York, 1997, pp. 41-50.
7. S. Sugawara et al., “InterSpace: Networked Virtual World for Visual Communication,” *IEICE (The Institute of Electronics, Information, and Communication Engineers) Trans. on Information and Systems*, Vol. E77-D, No. 12, 1994, pp. 1344-1349.
8. C. Greenhalgh and S. Benford, “Massive: A Collaborative Virtual Environment for Teleconferencing,” *ACM Trans. on Computer-Human Interaction*, Vol. 2, No. 3, Sept. 1995, pp. 239-261.
9. J. Galegher and R.E. Kraut, “Computer-Mediated Communication and Collaborative Writing: Media Influence and Adaptation to Communication Constraints,” *Proc. Int’l Conf. Computer Supported Cooperative Work (CSCW 92)*, ACM Press, New York, 1992, pp. 155-162.
10. A.J. Sellen, “Speech Patterns in Video-Mediated Conversations,” *Proc. Int’l Conf. Human Factors in Computing Systems (CHI 92)*, ACM Press, New York, 1992, pp. 49-59.
11. J. Bowers, J. Pycock, and J. O’Brien, “Talk and Embodiment in Collaborative Virtual Environments,” *Proc. Int’l Conf. Human Factors in Computing Systems (CHI 96)*, ACM Press, New York, 1996, pp. 58-65.

Hideyuki Nakanishi is a PhD student in the Department of Social Informatics at Kyoto University, Japan, and a Research Fellow of the Japan Society for the Promotion of Science. He is currently working on community computing, a research field for creating methodologies, mechanisms, and tools that support social interaction via global computer networks. He received his BE and ME from Kyoto University in 1996 and 1998, respectively.



Chikara Yoshida is an engineer in the Multimedia Development Center at Matsushita Electric Industrial, Osaka, Japan. He is currently working on an object-oriented network language for consumer electric appliances. He received his BE and ME from Kyoto University in 1995 and 1997, respectively.

Toshikazu Nishimura is currently an associate professor in the Department of Computer Science at Ritsumeikan University, Japan. His research interests include computer-mediated communication and mobile computing systems. He received his BE, ME, and PhD in information science from Kyoto University in 1990, 1992, and 1998, respectively.



Toru Ishida is currently a professor in the Department of Social Informatics, Kyoto University, Japan. He has been working on multiagent systems and recently initiated the study of community computing and support systems with his colleagues. He received an ME and a PhD in information science from Kyoto University in 1978 and 1989, respectively.

Readers may contact Nakanishi at Dept. of Social Informatics, Kyoto University, Yoshida-Honmachi, Sayko-ku, Kyoto 606-8501, Japan, e-mail nuka@kuis.kyoto-u.ac.jp.