Hideyuki Nakanishi
Department of Social Informatics
Kyoto University, Japan

Design and Analysis of Social Interaction in Virtual Meeting Space

Doctoral Thesis Series of Ishida Laboratory
Department of Social Informatics
Kyoto University

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my sincere gratitude to my thesis supervisor, Professor Toru Ishida at Kyoto University, for helpful guidance, invaluable advice and worthy discussions. I would also like to express my thanks to my other reading committee members at Kyoto University, Professor Haruo Hayashi and Professor Tetsuro Sakai for their valuable comments. Professor Hayashi, my vice adviser, also gave me lots of useful suggestions.

I must express my special thanks to the cooperation of Professor Clifford Nass at Stanford University and Doctor Katherine Isbister at Finali Corporation, who played an important role in the design and the evaluation of the social agent.

I received really a lot of support in the preparation of the network and the equipments used in the cross-cultural experiment. NTT GEMNet project provided us the international dedicated line. Doctor Fumio Hattori and Tomoaki Tanaka at NTT Software made strenuous efforts kindly to connect the dedicated line with the Stanford campus network, and prepare the equipments in Stanford side. Takeshi Oguro, Kaoru Hiramatsu, and Jun-ichi Akahani at NTT Communication Science Laboratories helped us to prepare the network connection and the equipments in Kyoto side. Assistant Professor Kenji Fujikawa and Associate Professor Sumio Okabe at Kyoto University maintained the connection between the dedicated line and the campus network of Kyoto University.

I would like to thank the members of the Media Center of Tohwa University, and Professor Edmund H. Durfee of the University of Michigan for their help in conducting the intranet/Internet experiments. I would also like to thank Eva Jettmar and other students at Stanford University for their assistance in conducting the cross-cultural experiment.

I am grateful to Professor Ken-ichi Okada at Keio University, another vice adviser, for his valuable advice and discussions.

I appreciate the constructive discussions and the kindly advices from Associate Professor Hiroshi Ishiguro at Wakayama University, Associate Professor Toshikazu Nishimura at Ritsumeikan University, and Assistant Professor Hirofumi Yamaki at Kyoto University.

I am deeply indebted to my colleagues of the Community Computing Project at Professor Ishida's research group, Chikara Yoshida, Hiroki Nakano, Masayuki Okamoto, Minoru Nakata and Satoshi Nakazawa, for their devotion to the project, hard work and precious discussions. Chikara Yoshida and Hiroki Nakano were also my colleagues of Kyoto University Microcomputer Club (KMC), where I had learned and experienced many things especially about videogame programming.

I would like to extend many thanks to all the members and the alumni at Professor Ishida's research group for their help and fruitful discussions. They were very kind and helpful especially in conducting experiments.

Thanks are also due to all subjects that attended my experiments. Without their participation, I would not have been able to advance my research.

At last, I sincerely thank my parents and other family members for their generous support.

Abstract

In this thesis, we propose new approaches to support daily conversations via computer network. First, we designed a meeting environment where social interaction is casual and relaxed. Then, we designed a character acting as an in-between of people.

Videoconferencing systems usually make a user see and hear all the other users. This style of interaction fits formal meetings where everyone discusses the same topic together. However, that interface cannot effectively support casual meetings like daily conversations. While some studies tried to design special functions that extend videoconferencing systems to support casual interaction, we combined simply three-dimensional virtual spaces with video-mediated communication.

Casual meetings play an important role in workplaces as well as in everyday life. Casual meetings are characterized by spontaneous conversations and meetings with many people. People accidentally form groups to have multiple conversations in the same place. Telephone-like protocol to start interaction cannot support spontaneous conversations. Videoconferencing systems usually display the videos of all the meeting participants at once. That interface prevents participants from forming subgroups. We developed a virtual meeting space *FreeWalk* that provides a common virtual space for spontaneous interaction. Users can change their positions to select whom they observe and listen to, and to form conversation groups.

FreeWalk provides a three-dimensional virtual space where everyone can meet and talk with each other. In the space, each user is embodied as a object on which his/her video image is pasted, has a location and a view direction. Users can approach someone to begin talking, and watch what others are doing from a distance. Conversations may occur when walking users encounter accidentally. Since voice volume attenuates in proportion to the distance between the sender and the receiver, users can form multiple groups, each of which is separately located in the same space. Users can smoothly join the conversation that attracts their interest, since they can guess the topic by listening to the conversation beforehand. In FreeWalk, the spatial positions determine how

to transmit the video and audio data. This simple mechanism makes all the above features that are necessary to support casual meetings.

This mechanism is also used to minimize the network traffic. In the FreeWalk system, video data are transmitted among clients corresponding to users. Each user's video is sent to the only clients that need it to draw their screen. Furthermore, the size of the video image is adjusted beforehand according to the receiver's view.

We evaluated our idea by comparing communication in the FreeWalk space with video-mediated and face-to-face communications. As a result of the analysis, we have categorized the effects of a three-dimensional virtual space into two types. The first type is to make video-mediated communication similar to face-to-face one. This type of the effects is found in the frequency of chat and the behavior of participants. The second type is peculiar to virtual spaces. This type of effects equalizes the amount of utterances for each participant, increases the number of turns, and sometimes stimulates participants in moving around to have free conversations. The freedom of virtual spaces seems to enable participants to relax, and that atmosphere may stimulate participants into talking easily.

The difference of participants' moves in the FreeWalk meetings gives important implications. Participants seldom moved after forming a circle when they were given the common topics to discuss together. However, participants formed multiple groups to greet and chat with others when they did not have any common topics. Virtual spaces seem to support the special style of casual communication as well as make video-mediated communication more casual. These results shows that FreeWalk supports forming multiple conversation groups and spontaneous conversations.

As described above, virtual meeting spaces have a potential ability to support everyday casual interactions. However, this ability causes a problem. Since it is easy to arrive at a virtual meeting space from many entry points, it is hard for visitors to assume much about one another's backgrounds. Virtual meeting spaces usually provide little social context to find a common topic to talk about. To eliminate this difficulty, we developed a *social agent* supporting human-human

interaction in virtual meeting spaces. This agent mimics a party host, and tries to find a common topic for two meeting participants whose conversation has lagged.

The agent tracks audio from a two-participants conversation, to look for a silence that means an awkward pause. When the agent detects a pause, it approaches the participants to begin the topic-suggestion cycle. In the cycle, the agent conducts a series of yes/no questions to both participants to draw shared or conflicted points. After the cycle, the agent recommends what they should talk about. This interaction provides a chance to start a conversation.

Since the meeting for the first time among strangers from different national cultures has the least social context, we evaluated the agent's ability to assist in the cross-cultural first time meeting between Japanese and Americans. We designed two kinds of agents to introduce culturally common safe or unsafe topics to conversation pairs, through a series of question and answer. In the experiment, the safe agent had positive effects for American students. Meanwhile, it had negative effects for Japanese students, but it made them think their partner was more similar to themselves. In the condition with unsafe agent, both Japanese and American students thought their conversations were more interesting, and Japanese students acted more American. As a result, we found that provocative topics are useful, an agent adaptive to each participant is good, and an agent's presence affects participants' style of behavior. This experiment demonstrated the agent's effectiveness, and raised interesting considerations for further development.

The conclusion is that our communication environment and communication helper are useful for daily social interactions. The tight combination of three-dimensional virtual space and video-mediated communication enables forming multiple conversation groups and spontaneous conversations. Autonomous social characters can diminish the demerit of the low social context in virtual meeting spaces.

Chapter 1

Introduction

1.1 Objectives

Most computer systems for collaborative work provide videoconferencing tools for planned meetings to discuss business topics. However, meetings aren't always planned or business related. Casual meetings such as chatting during a coffee break or in a hallway occur daily. They maintain human relationships, and also play an important role in collaboration. We believe that conventional videoconferencing systems, which simply transmit videos and voices like video phone, cannot support casual meetings. Our goal of this study is to propose innovative approaches to support casual meetings.

The first approach is to use three-dimensional virtual spaces as a social communication environment to facilitate casual meetings. *FreeWalk* lets people meet casually in virtual spaces such as a park or a lobby [Nakanishi96, Nakanishi98a, Nakanishi98c, Nakanishi99]. FreeWalk is just like a videoconferencing system that assigns a location and a view direction to a video window of each user. Users control their video windows in a virtual space. The following list describes the inherent features of casual meetings and how FreeWalk can support them.

- Spontaneous conversations

In conventional videoconferencing systems such as *Office Mermaid* [Watabe90], participants turn on the system when they start a meeting. The system displays the videos of all participants on their workstations, which hinders free conversation. The system also lists the participants before the meeting starts, thereby prohibiting accidental encounters with other participants.

Several videoconferencing systems have tried to extend their functions to support spontaneous conversations. *Cruiser* randomly selects some of the participants and displays their videos to other participants to simulate accidental encounters in a hallway [Root88].

In contrast, FreeWalk's approach provides a common virtual space for spontaneous conversations wherein participants can move and meet by themselves. It does not promote any system-directed encounters. The participants' videos are displayed on screens only when their embodiments meet.

- Meetings with many people

  In meetings such as parties, several tens of participants simultaneously exist in the same space. People gather physically to maximize the possibility of having spontaneous conversations. In the situation like that, it is almost impossible to use videoconferencing systems, since they try to display the videos of all participants at once. Plus, even if it were possible, it would be very hard for users to comprehend the situation. It is very troublesome for users to select whom they observe. Furthermore, they may have to adjust each level of awareness such as the size of the video image, for each person they observe. Even if users do not mind doing that, it is almost impossible to have symmetrical awareness among participants.

  Vrooms has virtual rooms, one of which is the window containing video images of participants [Borning91]. This design can divide all participants into several subgroups. Sharing the same room

  with others means that the user is observed by them as well as he/she can observe them. In this system, it is easy for users to form conversation groups. However, the user in a virtual room cannot observe other users in another room.

  In FreeWalk, participants can freely change their locations and view directions to watch other participants in the virtual space. They can walk around before they talk to someone else. They do not need to see the videos of all participants at once. They can select whom they observe and listen to with keeping symmetrical awareness among them in a simple manner.

Many systems use a three-dimensional virtual space as a multiuser environment. *DIVE*, a multiuser platform, lets people create, modify, and remove objects dynamically [Hagsand96]. This system has a script language to define autonomic behaviors of objects. Another multiuser virtual environment, *Diamond Park*, has a park, a village, and an open-air cafe [Waters97]. *Community Place* integrates Virtual Reality Modeling Language (VRML) and has an online chat forum [Lea97]. *InterSpace* [Sugawara94] supports audio and video communication for the experimental service CyberCampus, which features distance learning and online shopping. *Valentine* [Honda97] uses a three-dimensional virtual space to simulate a office room shared with several workers. These systems aim to construct realistic virtual worlds containing many kinds of virtual objects such as mountains, buildings, workplaces, artifacts, and so on.

We designed a simple mechanism to support casual meetings. The FreeWalk space allows people to change their positions dynamically during the meetings. The role of three-dimensional virtual space in our system is similar to the spatial model of interaction in *Massive*, a VR-based conferencing system with text and audio communication [Greenhalgh95]. However, the spatial position is much more tightly combined with video-mediated communication in FreeWalk. The implementation of FreeWalk, as well as the design, is light. We used videogame technologies instead of customized devices [Nishimura98].

Social interaction in three-dimensional virtual meeting space has not been studied sufficiently. Some earlier studies tried to compare face-to-face communication with video-mediated communication [O'Conaill93]. Various characteristics of conventional video communication became clear through those studies. However, the characteristics of the communication aided by a three-dimensional virtual space remained unclear. We conducted an experiment to find the characteristics of social interaction in FreeWalk [Nakanishi98b, Nakanishi2000].

The second approach to facilitate casual meetings is employ an autonomous social character as a coordinator of meetings. Virtual meeting spaces make it easy to have casual meetings between strangers from across town, or even across the world. However, virtual spaces provide little social context for users to assume social identities of others, since people can gather in the space from everywhere. People have a difficulty in starting and advancing their conversation when they do not know what they should talk about.

To diminish this difficulty, we developed a social agent that is an embodied character in virtual spaces [Isbister2000a, Isbister2000b]. Our agent acts as a coordinator to maintain social interaction among people. Our agent conducts simple question and answer so that people whose conversation is faltering can find a common topic to talk about. Since the meeting for the first time among strangers from different national cultures has the least social context, we applied our agent to the cross-cultural first meetings held in FreeWalk space.

These our efforts were conducted as a part of the *community computing project*, which aims to establish the design and the methodology to develop tools supporting everyday activities by forming a community through global computer networks [Ishida98a, Ishida98b]. FreeWalk is one of the products developed in this project [Ishida96, Ishida97].

The first product of the project, *Socia* is the function to schedule meetings on videoconferencing systems [Ishida94, Yamaki96a]. In Socia, the agents, the proxies of users, negotiate the schedule of the meeting with one another. In the conventional way, the agent can control the user's calendar and make the final decision. Socia introduced the notion of *non-committed scheduling*, which continues the negotiation until the all users agree to the suggestion of the beginning time of their meeting. FreeWalk provides users a continuous environment to meet instead of scheduling the meetings.

Socia and FreeWalk focus on desktop conferencing. Meanwhile, another activity of the community computing project tried to apply mobile computing technology to community support [Nishibe98]. This trial distributed a hundred personal digital assistants to the participants of an international conference, and provided diverse community services such as electronic mail, bulletin board, delivering information related to the participants, the conference and the sightseeing, and meeting arrangement. This experiment gave implications about how to use the emerging mobile devices for community support.

Another emergent device is a large-scale screen. The project tried to apply this device to community support. *Silhouettel* is a digital wall that augments real-world encounters [Okamoto98, Okamoto99]. The purpose of this system is very similar to our social agent that acts as a coordinator to maintain social interaction among people in FreeWalk space. But, this system maintains interaction in the real world, not in virtual worlds. The second difference is that this system is ambient while our social agent directly interacts with users. Silhouettel screen displays the augmented silhouettes of people standing in front of the screen. The silhouette images, the personal information of users, and the Web document related to their personal information are visualized in the graph representation method [Kitamura96]. The system lets users gathering in the same place know what is the characteristic of each user, and what is the commonality of them.

1.2 Outline of the Thesis

This thesis consists of seven chapters, including this chapter as the introduction.

Chapter 2 is dedicated to introduce the background of this thesis to show why a new approach is necessary to support daily conversation in computer networks. Related studies are divided into several categories to summarize conventional approaches.

Chapter 3 introduces a new approach to support casual meetings in everyday life. We developed FreeWalk that provides a three-dimensional virtual meeting space where everybody can meet and talk. Most precedent research attempted to design special functions for initiating spontaneous conversations. Instead of those functions, FreeWalk has a mechanism to administrate spatial positions of meeting participants. It is discussed how spatial positions make accidental simultaneous conversations possible.

Chapter 4 presents several techniques used in the implementation of FreeWalk. The system architecture consisting of server and client processes, and the data transmission mechanisms among these processes are depicted to show how FreeWalk system works. Videogame technologies introduced to the implementation of FreeWalk are explained. This chapter described the performance evaluation of FreeWalk to confirm the effectiveness of the system implementation. The results taken from the three experiments, laboratory, intranet and the Internet experiments, are discussed.

In Chapter 5, the experiment to observe the effects of three-dimensional virtual meeting spaces on social interactions is described. We conducted an experiment to evaluate the design of FreeWalk. In the experiment, communication in a virtual meeting space is compared with communications in a videoconferencing system and a face-to-face meeting. Our hypothesis was that the FreeWalk space is more similar to face-to-face environment than videoconferencing tools. Two types of effects were found in the analysis of the collected data. One is to make video-mediated communication approximating to face-to-face one, and another is peculiar to three-dimensional virtual space. Interesting behaviors observed in a FreeWalk meeting are also reported.

In Chapter 6, a social agent supporting casual meetings in virtual

spaces is presented. We believe that social agents can eliminate the inherent difficulty of virtual meeting spaces. The concept of social agent and the unique design of our agent are introduced. Since the first time meeting of people who have different cultural backgrounds seems to be most difficult situation to initiate conversations, we tested out agent in the cross-cultural meetings held on the network between Kyoto University and Stanford University. The design of the experiment and the results of the analysis are described. Based on the results, interesting considerations for further development are discussed.

Chapter 7 concludes the thesis summarizing the result obtained through this research and the prospect of the future research.

Chapter 2

Background

2.1 Casual Meeting

Casual meetings occur in workplaces as well as in everyday life. This sort of communication is called *informal communication* in contrast to the formal aspect of scheduled communication in meeting rooms. This section summarizes studies about supporting informal communication.

Informal interaction is like conversations occurring around a coffee pot, near a mail box, or in a hallway. In such interaction, the purpose of interaction, the length of interaction, and the degree of each person's involvement in interaction are not determined beforehand, but are negotiated through the interaction itself in a subtle and dynamic manner [Borning91]. Most of formal communication in organizations at workplaces, is planned beforehand, and conducted in formal style. Informal communication is more spontaneous, frequent, expressive, and interactive [Fish93]. Informal communication occupies 31 percent of workplace activities, occurs mostly between two people, continues very shortly, and does usually not include greetings and farewells[Whittacker94].

Informal communication is important for organizations for accomplishing tasks, transmitting the culture and knowledge of an organization, and keeping loyalty and willingness of the members. It provides the members of an organization the flexibility to treat uncertain and ambiguous issues, goals and decisions. In spite of such importance of informal communication, people neither pursue it nor make an effort to do it, but only make the most of their opportunities to do it [Fish93].

Informal communication is always caused by physical proximity, and occur effectively in people physically gathered, since people have more opportunities to find possible conversation partners [Whittacker94]. Physically colocated people have many opportunities to contact and communicate with each other, because of frequent chance encounters and easy access to other people [Fish93].

Physical proximity has effects to facilitate communication. Continuous linkage of voice and video channels is one way to give these effects to online communication. It is important to support coordination and negotiation process to begin interaction so that people can know the good time to contact others [Tang94]. To minimize these efforts required to start conversations, and provide an environment for frequent interactions, it is a possible solution to open voice and video channels continuously [Whittacker94]. If visual channels are available, people can recognize the conversation partners, the conversation topic and the time to begin talking simultaneously, and then people start interaction smoothly. Consequently, the probability of spontaneous interaction increases [Fish93].

When people communicate remotely, it is desirable that people can behave just as they are in face-to-face situation. It is useful for people to notice who observes them, and whether someone whom they like to talk, engages in a conversation by eavesdropping on it [Whittacker94]. Fluid group activities include such behaviors as distinguishing a particular speaker from many speakers, attracting others' attention, having a side conversation [Bly93].

Sufficient understanding of the characteristics of informal communication occurring in face-to-face environments is necessary to design communication environments. Most of interactions involve two people, and 88 percent of them are interrupted by the attendance of the third

party. Such an interruption should be supported by communication environments [Whittacker94].

It is hard to support informal communication in conventional remote communication environments. In the telephone model, to which most desktop conferencing systems and video phone adhere, the called person has to accept the call to start interaction. Since video phone requires a deliberated action, calling the other, to use it, and implies the beginning and the end of a conversation, video phone can be used for neither keeping peripheral awareness nor spontaneous interaction [Bly93]. In the telephone model, the caller side cannot know the appropriate time to call, and the called side cannot give feedback to indicate whether it is good time except accepting the call. The telephone model is too restricted, and that makes it burdensome to call the other to have a conversation [Tang94].

From the above description, a new communication environment must be designed to support informal communication. Some studies trying to do it are described below.

Bly *et al.* observed the long-term use of *media space*, which is a voice and video network controlled by computer [Bly93]. Media space was used for awareness support and informal interaction regularly and frequently. Looking at the control panel of media space, which continuously displays who is connected to whom, was similar to glancing a party room to grasp ongoing conversations. The panel enabling users to recognize who is talking to whom, who is eavesdropping on it and who is available, contributed to keeping the sense of a group.

Borning and Travers developed two systems, one is *Polyscope* and another is *Vrooms* [Borning91]. Polyscope supports awareness, and Vrooms supports casual communication.

In Polyscope system, video images captured in rooms of other members are arranged in a window on the computer screen of each user so that users can continuously observe others. The observer can select the observed members and the frequency of video capturing. Each user decides how much his/her information is disclosed, and chooses the level of feedback he/she receives. The system provides three levels of feedback, which are nothing, the names of all members who observes the user, and the video images of all observers. If the user set the video-symmetry flag, his/her video image can be seen by only users who disclose their video images. Before doing an experiment, people had commented that feedback and symmetry are important. However, both were actually seldom used during the experiment.

The metaphor of Vrooms is virtual rooms, one of which is the window containing video images as Polyscope system displays. This metaphor removes the complexity as the interface of Polyscope. Sharing the same vroom with others means that the user is observed as well as he/she can observe the others, everyone has a choice to start a conversation, and conversations occurring in the vroom are mostly casual. In the vroom, when a user moves his/her video image near the video image of the other user, voice and video channels are established between them. Since the frame picture surrounding the two images of conversation pair appears, everyone sharing the same room can recognize the ongoing conversations.

By extending Polyscope, Dourish and Bly developed *Portholes*, which informs the situations of remote work groups [Dourish92]. In this system, each user's computer screen displays video images sent from cameras attached to users' offices and public spaces per several minutes. When a video image is clicked, a button to start an e-mail program appears. Portholes provided opportunities to watch remote colleagues as well as local colleagues, and increased remote communication. Remote informal spontaneous communication had not been seen before people started using Portholes.

The overview model of Portholes, which arranges video images on one's own screen, enables users to grasp continuously what is going on in various places the way they watch camera images of surveillance systems. However, it may be difficult to use systems based on that model among loosely tied people, since users do not always aware of being observed by someone [Tang94].

Fish *et al.* developed a videoconferencing system called *Cruiser*, which has several special functions that are *Glance* and *Autocruise* [Fish93]. Glance opens a video channel for a second between the caller and the called sides. Autocruise, which imitates the situation that a walking person passes by someone accidentally in a hallway, generates a call automatically between two users selected randomly. Glance was used a little, however, Autocruise did not work effectively, and displeased users. The new version of Autocruise, which select two users communicating with each other frequently in the past usage of Cruiser, did still not work well and was disliked. This shows the manner of starting conversation causes the problem. Sudden appearances within enough short distance to have a conversation, provide people a chance to begin talking, and also make them disturbed.

Tang and Rua developed *Montage*, according to the hallway model imitating the behavior that one walks to another's office to look in at his/her room [Tang94]. In this model, anyone can glance and hear the inside of the room shortly without the explicit permission, if the door is open. In Montage, when a user specifies a person to be glanced, a duplex video channel is opened immediately without waiting the acceptance of the person. Then, on each user's workstation screen, a small video frame fades in with the sound effect to represent the approaching behavior. If both does not do anything within several seconds after this, the video channel is automatically disconnected. If either presses the voice button, a voice channel is opened. If the visit button is pushed, they have a meeting with larger video frames. As a result of an experiment, the voice button was pressed in one third of all the glances, and one third of that voice group also pressed the visit button. Conversations in the glances were shorter than those in the visit meetings. This indicates that Montage could support instant communication and the coordination to begin interaction.

Obata and Sasaki developed *OfficeWalker*, whose interaction model is the extended hallway model that has a part of the characteristic of the overview model [Obata98]. This model has two phases of the sense of distance, which are the public and private places. When users like to talk to a person, they enters the public space before starting a conversation in the private place of the person. While users stay at the public space, they can see the distant video images of the neighbors sharing the place, and the neighbors can see them, too. Since the neighbors do not know whom the visitor likes to talk to, they can easily ignore the visitor in this phase. In the public place, unintended interactions may occur among the visitors and the neighbors while the visitors are waiting for the targeted person to be available. In an experiment, the problem of intrusiveness was diminished, however, spontaneous interactions did not occur.

Examples described above tried to overcome the shortcomings of conventional videoconferencing systems and video phone. Several research aimed to support informal interaction in text-based communication instead of communication mediated by voice and video.

Isaacs *et al.* developed a distributed presentation system *Forum*, in which the presenter and the audience can attend the presentation in front of their own workstations [Isaacs94]. This system has an awareness support function, which displays the list of participants when users press the button. Users can browse personal informations of other audiences, and send a short text message to a particular participant. In the presentations in this system, one forth of the audience sent messages. The questionnaire result shows this message function was useful for interactions with others. However, this function could not create the same degree of shared experience as that in face-to-face presentations. For example, participants sometimes behaved in informal manner, which is appropriate for presentations with a small number of audience. The reason is that the window listing participants did not provide enough sense of the audience group.

Erickson *et al.* developed a text-based communication tool *Babble*, which displays graphic representations of participants "social proxy" to provide the state of conversation synchronously [Erickson99]. In the interface of "social proxy" each participant is represented as a small

circle showing the existence and the activity of the participant. The circles of active participants who post and read messages frequently are located near the center of a ring frame displayed on a screen. This representation lets users know how many participants have a conversation and how actively they are talking. The graphic representation is more effective than a simple list of participants in helping users to grasp a situation intuitively. In an experiment, the effects of "social proxy" caused spontaneous interactions. The moving circles attracted people's attention, and caused various interactions, which are from greetings to asking questions about work, on the chat tool. The circles triggered even calling the other participant on the phone and going to the other's room.

Brothers *et al.* added a function generating mailing lists automatically and immediately to a simple electronic bulletin board, which did not have the functionality to make subgroups and was used by all the members of a community [Brothers92]. This additional program is called *topics*. In the model of topics, every article has a potential theme of discussion and a potential subgroup interested in the article. In the model, one's joining an article represents his/her interest in its theme. When users finish reading an article, topics asks them about their intention of joining it before they begin reading the next one. If they choose to join the article, they are added to the mailing list of its subgroup, and then an e-mail system sends them the copy of discussion that have already happened. If a newcomer to an article is the first person to join it, the new mailing list for the subgroup of the article is generated. As a result of an experiment, the discussions occurred in the subgroups were significantly more informal than normal posted articles.

2.2 Media Space

As described in the previous section, some inherent characteristics of face-to-face communication must be considered to design systems supporting information communication. Several previous studies tried to prove the benefits of media space, and extend media space according to the nature of face-to-face communication. We explain these studies in this section.

Media space is a voice and video network controlled by computer to support collaborative work. One of media space systems *CAVECAT* is a combination of computer network and individual terminals connected by a voice and video network. Each terminal consists of a computer, a monitor, a camera, a microphone and a speaker [Mantei91]. Similarly, another media space *RAVE* is a set of nodes, each of which consists of a monitor, a camera, a microphone and a speaker. The connections of nodes are controlled by a computer [Gaver92].

Almost all media space systems use not only voice but also video as communication channels. Video-mediated communication is considered superior to communication mediated by only voice, because video can carry nonverbal information that is a part of face-to-face interaction [Grayson98]. Conversation is not a simple exchange of words, but includes various nonverbal communications as gaze, posture, gesture and facial expression [Sellen95]. However, the effectiveness of video channels is not enough clear.

Sellen compared four-person conversations in face-to-face environment with those in voice only conferencing system, and three kinds of videoconferencing systems [Sellen95]. As a result, turn-taking behavior in conversations was not affected by the existence of video channels. Hight quality voice channels were enough to have conversations. However, the opinions of participants showed that visual information was important and valuable for having conversations. An effective design that solves such issues as interpersonal distance, feedback and controlling visual information, may overcome the shortcomings of existing videoconferencing systems. Interpersonal distance is effective to follow multiple conversation threads in parallel. Appropriate feedback is needed to be designed to make users be aware of others' attention to them. In the case that the number of simultaneous visible participants displayed on a monitor is limited in controlling visual information, invisible participants make communication unnatural, and automatic selection of visible participants deprives users of their freedom.

Veinott *et al.* conducted an experiment to see how useful video channels are for everyday conversations [Veinott99]. Voice only condition was compared with voice and video condition. The task was that one subject taught another subject the route on a map. Two language groups of subjects did that task. One was the group of pairs, of which both were native English speakers. In another group, a subject's native language of each pair is different from each other, and that was not English. As a result of the experiment, there was no effect of video channels in the English group, but the non-English group accomplished the task more efficiently with the help of video channels. The video channels might enable the non-English subjects to use gesture when they could not find fit words to express something, and watch the partner's facial expression to estimate the degree of the understanding and to proceed to explain appropriately.

In the experience of Mantei *et al.* in using a media space system, the size of video image affected how influential a participant is in a conversation, and how a participant is recognized by other participants [Mantei91]. Participants displayed largely seemed to be influential in discussions. Conversely, participants displayed small seemed to sit far, and not to be very influential in conversations. Inappropriate sizes of video images made people's perception of others more personal or more impersonal in conversations. This is consistent with the social psychological theory saying that interpersonal distances reflect human relationships. Since video images affect interpersonal recognition and interaction, it is important for participants to adjust how they watch others' images and others watch their images.

Media space should be designed to support various cues used implicitly in face-to-face communication [Mantei91]. An example is a request of communication implied by physical co-presence in the same room. Proximity that is a distance between conversation partners, is a important element of nonverbal communication in face-to-face interaction. Gayson and Coventry analyzed proximity through video communication [Grayson98]. In comparison between short distance condition in which only the face was projected on a screen, with long distance condition in which the upper half of the body was projected, there was no significant difference in subjective impressions. However, a little influence was found in the structure of conversations. The reason of small influence may be that only video was the transmitter of proximity in the experiment while face-to-face proximity is multimodal information.

Generally, media space lacks the directions of voice and video sources. In media space, users feel that voice comes out of thin air, not from the speaker. This often prevents users from identifying speakers when they are communicating with multiple partners [Mantei91].

O'Conaill *et al.* compared face-to-face communications with communications in the ISDN videoconferencing system, and in the *LIVE-NET* videoconferencing system [O'Conaill93]. In the ISDN system, voice communication is delayed and half-duplex, and the quality of video communication is low. In the LIVE-NET system, voice communication is full-duplex and not delayed, the quality of broadband video is high. As a result of the analysis, communication in the ISDN system was more formal than face-to-face communication. The result shows that the LIVE-NET communication was more similar to face-to-face one than the ISDN one was, however, that was formal, too. Face-to-face environment differs from such videoconferencing systems as the ISDN system and the LIVE-NET system in the existence of the directions of voice and video, since voice and video come from speakers in face-to-face environment. The lack of direction of voice and video forced participants to transfer the floor explicitly by using words in videoconferencing systems. That might make conversations formal. The implementation of directional voice and video, as well as the increase of bandwidth and the evolution of compression technology, may improve videoconferencing systems.

Okada *et al.* developed three-person videoconferencing system *MA-*

JIC, in which voice and video have direction [Okada94]. Each individual terminal consists of a half-transparent large curved screen, on which two life-size persons can be projected, and two sets of video projector, camera, directional microphone and speaker. Since each camera and microphone are placed behind of the part of a screen projecting the life-size image of each participant, eye contact and the identification of speakers is easy. A participant can recognize the eye contact even between the other two participants. However, the heads of participants should not be moved away from the fixed positions to keep eye contact. It is problem that the scale of the whole equipment becomes large because each user needs a large screen.

One of crucial issues of media space is the protection of privacy. Since the sudden appearance of others in a user's office does not give enough feedback to him/her in media space, it is important for a user to notice the state that him/herself's office is visible to others, and how to prohibit this [Mantei91].

Heath and Luff observed people using media space, and found that they were often unaware of others' turning to face them [Heath92]. People have an ability to pay attention continuously to the behaviors of others who are in the same place. This forms the foundation of privacy in public spaces. Voice and video communication technologies threatens this foundation. The behavior that is watching others, loses its strong influence on others, and its power to facilitate others' behavior. Consequently, people can stare at their colleagues without being noticed by them. The ability of large-size monitors to increase visual attention to others' behavior was also tested. As a result, it is found that large monitors could increase attention to others' gaze and overall behaviors that cause large changes such as appearing on a monitor. However, people became to be sensitive to behaviors and changes occurring inside the environments of others. This prevented people from easily grasping only important changes by peripheral vision.

Gaver *et al.* developed *RAVE*, which provides several kinds of connections such as *sweep*, *glance*, *office share* and *vphone* [Gaver92]. The sweep and glance are one-way video connections to peek at others. The office share and vphone are reciprocal voice and video connections to interact with one another. The establishment of a connection should be noticed by the observed side to protect privacy without losing functionality. In RAVE, a different sound used to represent each kind of a connection. For example, a ring represents the vphone, a door's opening and closing sound is used for the glance, and a footstep corresponds with the sweep. These sounds mimicking actual ones transmits information intuitively without using words that disturb users, and remove the necessity to establish a symmetrical connection that bothers the other side.

2.3 Virtual Environment

The previous section mentioned that one of drawbacks of media space is the lack of directions of video and audio. Directional voice and video may make it easy to identify speakers, imply addressed persons, smooth turn-taking, and be useful for fluid communication. Virtual environments can deal with these directional aspects naturally. Virtual environments for multiuser communication are called *CVE (Collaborative Virtual Environment)*, or *DVE (Distributed Virtual Environment)*.

CVE is spaces generated by computers, which contain various representations such as data program and users [Benford96]. In the spaces, users have their graphical representations, control their view points, and interact with each other. The essence of CVE is that a shared space provides a consistent shared spatial framework that can deal with relations of positions and angles of different entities effectively. Since each user's view point is placed to each user's embodiment independently, a user can guess what other users pay attention to. CVE aims to provide continuous integrated context by displaying collaborators and the information of them in the same space. CVE has a wide range of potential application domains such as training, visualization, simulation.

DVE is a system that enables geographically distributed users to

interact with each other in real-time [Waters97]. It is possible to scale up a DVE system to support thousands of users. In videoconferencing systems, a user sees other users displayed on his/her monitor. In DVE systems, users enter into a virtual world to gather and see others' activities. In virtual worlds, users are represented as graphics called avatars, and the users control them. Users can engage in interaction and computer simulation simultaneously. Even unreal worlds can be constructed on DVE systems

The spatial model like CVE and DVE, does not imitate real worlds perfectly. The advantages of the spatial model are providing familiar and continuous collaborative environments, supporting awareness that is what others are doing, translating the motion of users in a shared space into accidental encounters, and making it possible to construct usable and learnable collaborative environments based on the natural understandings of physical worlds [Benford96]. The spatial model is becoming a popular design of collaborative environments because of the hypothesis that our behavior patterns are formed according to spatial elements of the real world, and we can transfer our behavior patterns directly from our daily life to a virtual space, which design is similar to the real world. However, our behavior is regulated by the sense of place, not by the structure of space. That sense is a standard of appropriate behavior in cultures and communities. It is dangerous to confuse the notion of place with that of space [Harrison96].

Many test-based chat tools combined with virtual environments have been developed. There are some research evaluating the effectiveness of virtual spaces in these tools.

Kauppinen *et al.* observed communications in virtual environments where users can have voice-mediated or text-based conversations with animated avatars [Kauppien98]. They found that users had similar habits to those of the real world, such as greetings, farewells, forming groups, and spatial positions representing relationships. The proportion of groups forming a circle to all conversation groups was fifty percent. Facing others meant the intention to join their conversation, and taking interest in that. It was observed that people approached to the chat occurred at a distance in the similar way to that of the real world. Users seldom used animated avatars to represent gestures and facial expressions. They used text-based facial symbols frequently instead.

Smith *et al.* analyzed communication in the chat tool *V-Chat* that has three-dimensional virtual spaces and avatar representations [Smith2000]. The V-Chat system supports connections with IRC clients, and provides IRC users a pure text-based interface without any graphics. As a result of analysis, it is found that spatial interactions occurred, which are very similar to those of the real world. Users tended to stand close to their conversation partners, and face them. The proportion of the messages including the name of an addressed user was less in V-Chat users than that in IRC users. This shows that positions of avatars can provide nonverbal information to indicate the addressed user like face-to-face communication. It is also found that frequent users were unlikely to use gesture animations of avatars.

2.4 Social Agent

There are various definitions of the word 'agent' [Ishida95]. We define social agent as a autonomous character interacting with people socially.

Under a complex dynamic environment, autonomous agent acts autonomously to accomplish its goal or task based on the information its sensors provide [Maes95]. When the environment is a physical world, the agent is embodied as a robot. It is embodied as a animated character in a two-dimensional or three-dimensional simulation environment. Software agent and interface agent that act inside computer network are not embodied.

Almost all interactive programs are designed for interaction with a single user, are regarded as a function to answer questions, and never talk actively [Takeuchi95]. In everyday conversations, there is no special role like the chairperson, the floor is controlled by eye contact, facial

expression, and gesture. The computer engaging in social interaction should be able to interact with more than two people, act based on a perceived situation, and join in interaction actively.

Two studies about autonomous agents interacting with human beings are described below.

Cassell *et al.* developed the human-like embodied agent *Rea* that can control intonation, gaze, gesture and facial expression in conversations like human beings, and respond to these visual and auditory cues [Cassell99]. The graphics of Rea is displayed on a projector screen. A camera to catch the positions of users' heads and hands is placed on the top of that screen. Rea interacts with a user in mixed initiative manner, accept a user's interrupting during Rea's speaking, and even read a user's intention to speak in his/her gesture. When Rea cannot catch what a user says, it tries to recover the conversation by asking voluntarily him/her to repeat again. Users liked Rea's function to control turn-taking. However, the function to recover conversations did not work well. Embodied conversation agent such as Rea, is a logical extension of conversation metaphor in human-computer interaction.

Maes *et al.* developed a virtual world *ALIVE* in which a user can interact with autonomous agent that is an animated character, by using the gesture of his/her whole body [Maes95]. To a large screen, the system projects a synthesized picture that is a three-dimensional virtual space overlaid with the picture of a user retrieved from the image captured by a video camera. One of ALIVE's goals is to show an actual example of more emotional virtual environments that involve interaction with animated characters. The lessons in this system is that the design of interaction, in which the user engages, is more important than fancy graphics in order to design an attractive and immersive environment.

In the studies described above, agents are embodied as animated character of human beings or animals. However, it may be unnecessary for a social agent to be represented as a humanoid character, or even to be embodied. Some researchers studied the effects of the embodiments of social agents.

Sproull *et al.* analyzed how the difference in the interface asking the questions of a psychological test affected subjects' answers [Sproull96]. In the experiment, a graphical human face and a text-based interface was compared. Social psychological theories tell that the existence of others make a person doing a task more aroused and present him/herself in a more positive light. The results showed that the interface of a human face made people more aroused and present themselves more positively than text-based interface. People responded to the human face interface as it was a real person.

Takeuchi and Naito developed a conversational system that can control synthesized facial expression, and interact with multiple users simultaneously [Takeuchi95]. This system control facial expression, gaze, and head movement based on the perceived situation of the external environment through a video camera. In an experiment, in which the computer gave hints about a card game nonverbally to each player, the face was compared with the three-dimensional graphic of an arrow that can give players the hints similarly without facial expression. As a result, the face was more enjoyable for subjects, but less useful. The face drew the subjects' gaze. The subjects responded to some kinds of facial expressions unconsciously. People seem to try to read subtle signals in a humanoid interface, and respond to those.

Kiesler *et al.* analyzed the effect of different interfaces to the degree of cooperation in two-person prisoner's dilemma [Kiesler96]. Four kinds of partners that were a human partner, computer-generated human voice and face, computer-generated human voice, and text, were tested. As a result, it was found that people were more cooperative if they had a simple discussion about the next choice regardless of the kind of the partner. The order of easiness in breaking the promise with the partner was computer voice and face, computer voice, text, and a human partner. This meant that the more human-like computer was more likely to be disliked. In the same task, Parise *et al.* tested, a human face agent, a realistic dog agent, a cartoon dog agent, and a real

human communicating with the subject through a videoconferencing system [Parise96]. The results showed that both dog agents impressed the subjects as cuter, more believable, and more likable than both of the real human and the human agent. However, the probability of cooperation was higher in the two human conditions than in the two dog conditions.

Lester *et al.* developed a pedagogical agent that is an animated character advising users in tutorial software [Lester97]. Several kinds of agents that differ in the variety of the advises, were used by the students of a middle school in a classroom. As a result, every kind of agent was regarded as very useful, believable and enjoyable. The agent providing more various advises was perceived more positively by the students, and increased the efficiency of their learning more.

The findings mentioned above shows that the humanized agent is appropriate for collaborations that require interaction through facial expression, and the agent should be embodied as an animated character if it has to be liked and enjoyed by the users.

Chapter 3

Interaction Design for Virtual Meeting Space

In this chapter we describe FreeWalk's design of social interaction in the three-dimensional community common area. We also discuss how FreeWalk supports casual meetings.

3.1 Three-dimensional Community Common Area

Figure 3.1 shows an image of a FreeWalk screen. FreeWalk provides a three-dimensional community common area where people can meet. Participants move and turn freely in the space using their mouse (just as in a videogame). Locations and view directions of participants in the space determine which pictures and voices get transmitted.

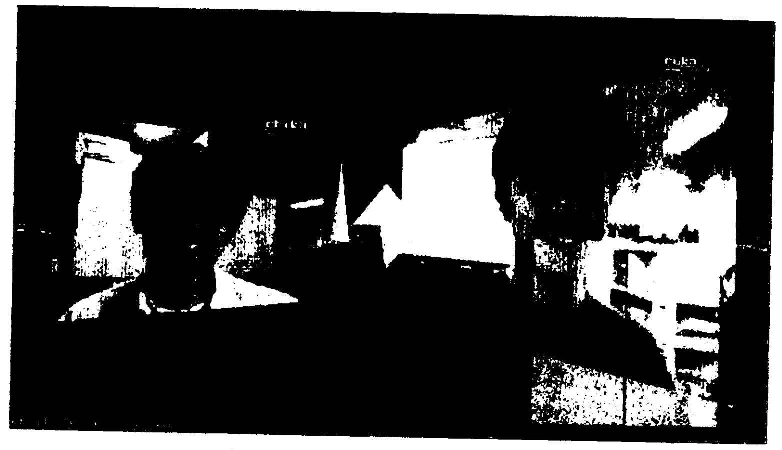
In this three-dimensional space, a pyramid of three-dimensional polygons represents each participant. The system maps live video of each participant on one rectangular plane of the pyramid, and the participant's viewpoint lies at the center of this rectangle. The view of the community common area from a participant's particular viewpoint appears in the FreeWalk screen. Figure 3.2a shows participant A's view when participants B and C are located as shown in Figure 3.2b.

Figure 3.1: FreeWalk: Three-dimensional Virtual Meeting Space

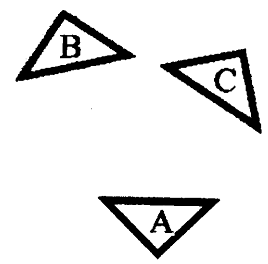
Participants standing far away in the three-dimensional environment appear smaller and those closer appear larger. FreeWalk doesn't display participants located beyond a predefined distance. The system also transfers voices under the same policy–that is, voice volume changes in proportion to the distance between sender and receiver. See Figure 3.3. Moreover, a participant hears others' voices in stereo so that he/she can easily recognize the speaker.

3.2 Simulating Casual Meetings

In FreeWalk, meetings can start with an accidental encounter. Figure 3.4 shows an example of an accidental encounter, where the user finds others on the radar screen displayed at the right bottom corner of the screen (Figure 3.4a), watches them to find out what they're talking



(a) Participant A's view



(b) Map of B and C's locations

Figure 3.2: Participant's View of Community Common Area

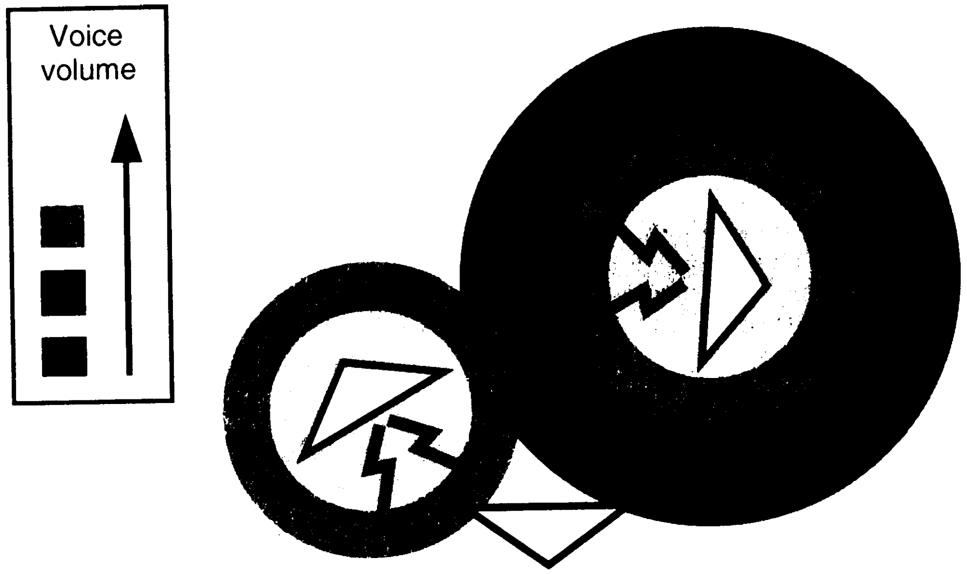


Figure 3.3: Voice Transmission

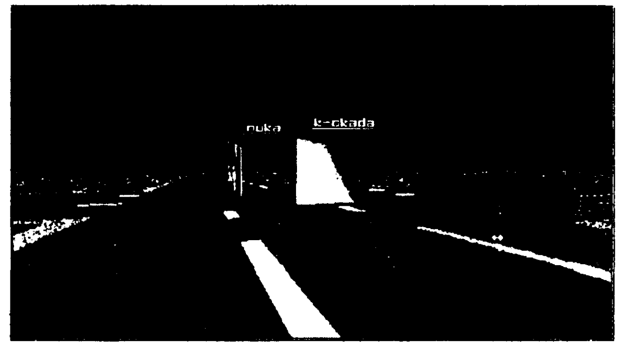
about (Figure 3.4b), then joins them (Figure 3.4c).

Since distance attenuates voice, a participant must approach the others in order to talk to them. On the other hand, not only can the participants in the conversation hear the speaker's voice, but also 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.

3.3 Organizing Meeting Groups

Desktop videoconferencing systems provide various functions to support the organizational behavior of participants, such as speaker selec-

(a) Finding others on the radar screen



(b) Watching a talking pair

(c) Joining their conversation

Figure 3.4: Accidental Encounter

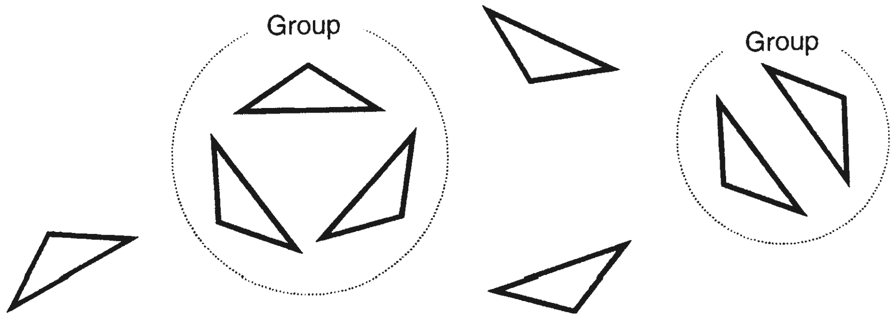


Figure 3.5: Meeting Organization

tion. 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 three-dimensional space that promotes a casual feeling in communication.

People form a group by standing close to each other to engage in conversation. Figure 3.5 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.

3.4 Using Nonverbal Signals

Space provides a context for interaction [Duck98]. Spatial positioning is a nonverbal cue, which serves to communicate liking and disliking and attraction to a relationship. The orientation of a body and eye contact are used to start, sustain, and end interactions. These nonverbal signals

are important in casual meetings since they smooth out and regulate social behavior. In the virtual space of FreeWalk, users can partially use these signals.

Interaction is controlled by a behavior that is changing body orientation, and communication becomes easy to do if this behavior is judged correctly [Cranch71]. The behavior that is turning directions of eyes, a head and a body is based on the structure of a human body. Turning behavior reflects emotional attitude toward others. The orientation of a pyramid represents the body orientation of a user in FreeWalk.

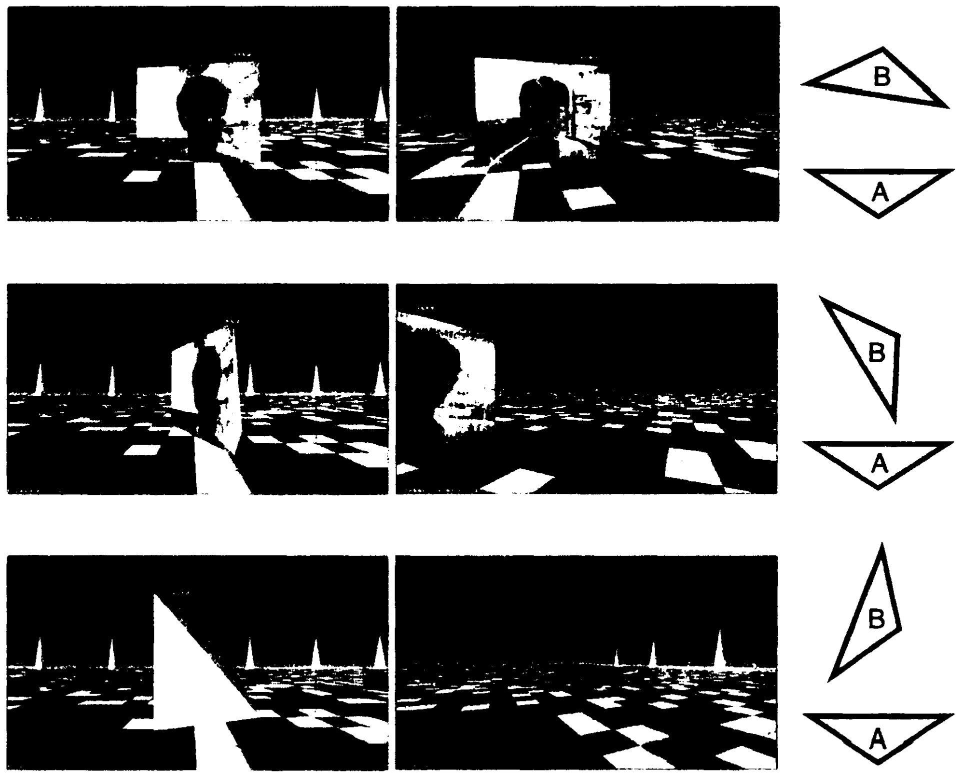
Since the participants' locations and view directions reflect a pyramid orientation, each participant can grasp the locations and view directions of other participants, and observe what other people are doing from a distance. Participants can also observe others around them by turning their body. Figure 3.6 shows the view changes of participants A and B as participant B changes his direction in front of A.

Since the volume and direction of one's voice is determined by its position, participants can easily identify the others in the same conversation group. Comprehending the correspondence between each face and each voice is not as difficult as that in conventional videoconferencing systems. The speaker can assume that the listeners recognize who is speaking. Furthermore, the speaker can turn to face the person he/she is talking to. These natures may transmit nonverbal information that makes it unnecessary to call the name of the person whom you talk to before you begin speaking.

3.5 Summary

In this chapter, we proposed the new design of a communication environment for casual meetings. FreeWalk is a virtual meeting space where everybody can meet and talk with each other.

In the space, each user is embodied as a three-dimensional pyramid object on which his/her live video is pasted. Each embodiment has a location and a view direction, and moves around freely. Users can



(a) A's view of B (b) B's view of A (c) A map of the views

Figure 3.6: Changes of Participants' Views

approach someone to begin talking, watch what others are doing from a distance. Conversations may occur when walking users encounter accidentally. Since voice volume attenuates in proportion to the distance between sender and receiver, users can form multiple meeting groups, each of which is separately located in the same space. The voice of a user can be heard by anyone in the neighborhood. Users can smoothly join the conversation that attracts their interest, since they can guess the conversation topic by listening to the conversation beforehand. Since the volume and direction of one's voice is determined by its position, and the speaker can turn to face the person he/she is talking to, the speaker can assume that the listeners recognize who is speaking.

In FreeWalk, video and audio channels transmit nonverbal information such as facial expressions and paralanguage in the same way of videoconferencing systems, and spatial positions of participants make accidental simultaneous conversations possible. The spatial model is tightly combined with the video-mediated communication. The meeting style is more fluid and more dynamic than that in videoconferencing systems.

Chapter 4

Implementation of Virtual Meeting Space

In this chapter we discuss FreeWalk's system design and implementation.

4.1 System Configuration

The FreeWalk system consists of a community server and clients, each of which includes vision and voice processes. Figure 4.1 illustrates the interaction between the community server and clients.

When participants move in the three-dimensional 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 three-dimensional community common area. The server finally sends the list back to each client for screen updating. Since only control information is transmitted 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

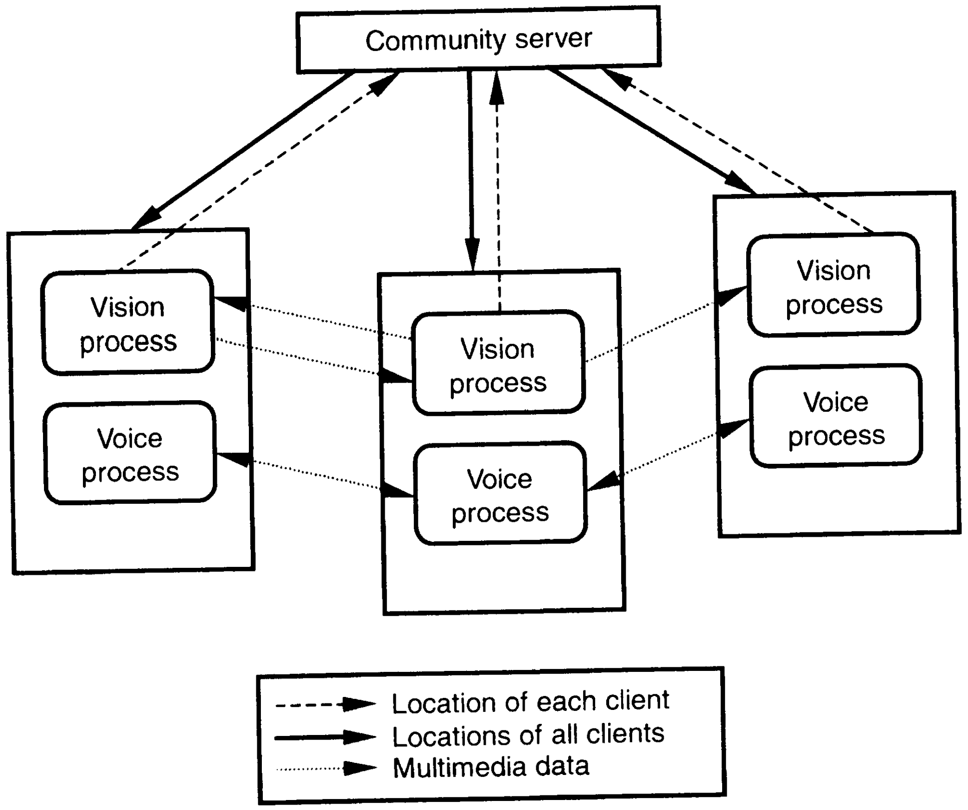


Figure 4.1: FreeWalk System Configuration

pictures from other clients, the vision system redraws the display based on the information in the list and the received pictures.

Because each client cannot see all the other clients, it's not necessary for each one to send its picture to all the others. Similarly, each client does not 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 size of the picture to the size the receiver needs.

- The client sends its picture to others who can see the client.

Figure 4.2 shows an example of a video transmission 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 do not send voice data to those clients located too far away to hear the participants' voices.

4.2 Applying Videogame User-interface

Some videogames provide three-dimensional multiuser environments where users can control their characters. The virtual spaces of those videogames enable players to control their characters freely. So, videogames and FreeWalk can share the ideas about user-interface design. We believe that the following features of videogame can be introduced into FreeWalk.

- Displaying global situations in a three-dimensional space

  In videogames, since the move of characters is fast, and they are often time-bounded, most of them provide facilities for users to grasp global situations.

- Running on a low cost machines

  Though some of arcade games need special expensive input devices, most of videogames, especially game machines for home use (known as "Nintendo" machines) or personal computers, do not. People can play them only with the low-priced general-purpose machines with a joy-pad, joystick, mouse or keyboard. Moreover, most of these devices are the standard attachments to such machines.

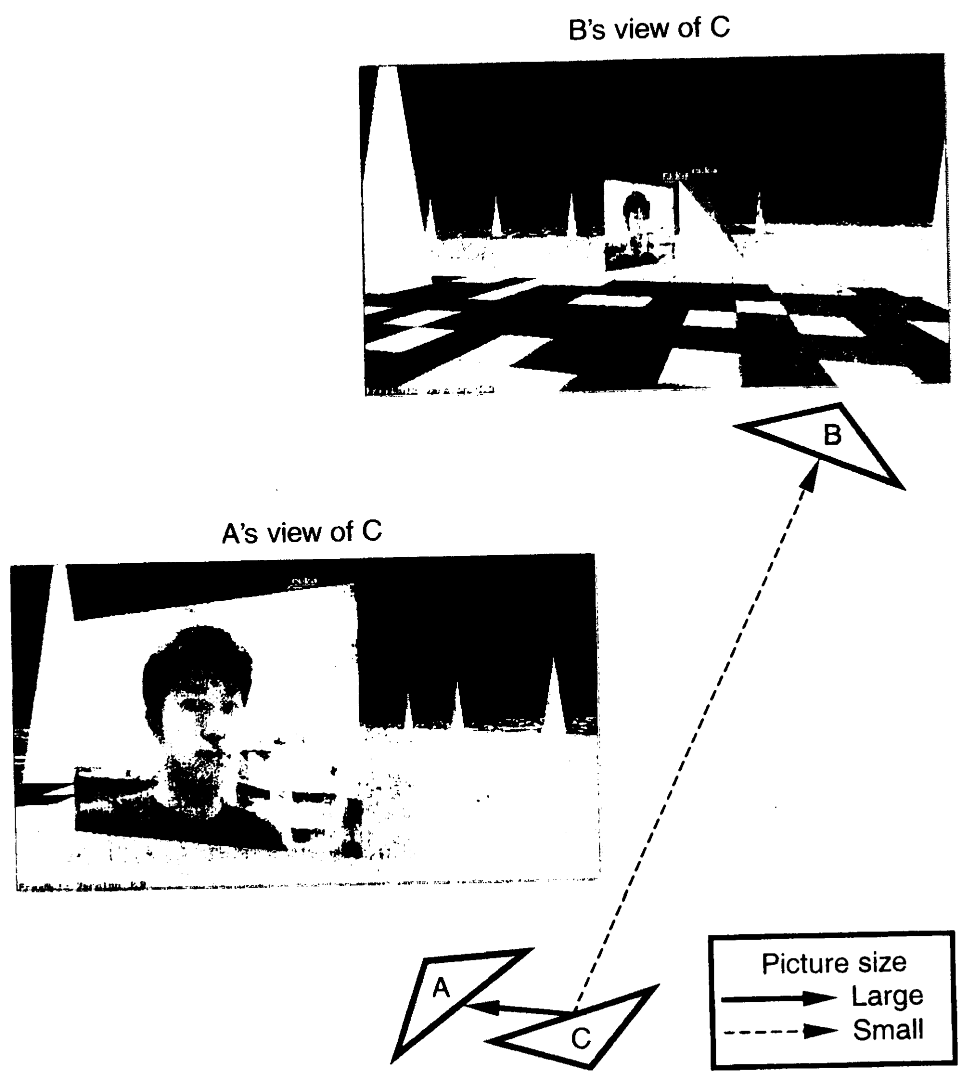


Figure 4.2: Video Transfer among Clients

We applied the user-interface of such as action, shooting and racing games to the design and implementation of FreeWalk. FreeWalk imitates the user-interface of videogame in controlling the user's character and in grasping the surrounding situation.

4.2.1 Freely Walking in the Three-dimensional Space

Since the standard input device for a home game machine is joystick with several buttons, almost all videogames are designed for this device. Since one of the standard devices of workstations is a mouse, the motion of users' pyramids is controlled by X and Y valuators of a mouse pointer during its left button is down. A user controls the orientation of the pyramid by a mouse so that the moving/turning speed is proportional to the distance between the mouse pointer and the center of the FreeWalk screen. Since users can easily control the speed of moving/turning, they can run when the target is in a distance and slow down as the target becomes closer.

4.2.2 Grasping Situations in the Three-dimensional Space

Since the view angle in CRT is much narrower than that of human eyes, it is hard to grasp surrounding situations. Moreover, a human in the real world can easily look around by turning his/her head, but cannot do the same thing in the three-dimensional space. Though virtual reality systems can simulate this by a head-mounted display (HMD), widely used machines are not equipped with it. In videogames, therefore, additional auxiliary indicators and viewpoint switching functions are introduced to help users to grasp their situations. From this observation, we implemented the following functions in FreeWalk.

- Radar Screen

A radar-like screen indicates the simplified view of surroundings, including locations of characters. Figure 4.3 shows the FreeWalk radar screen. The radar screen can also indicate the volume of people's voice so that user can roughly know the activities within the groups.

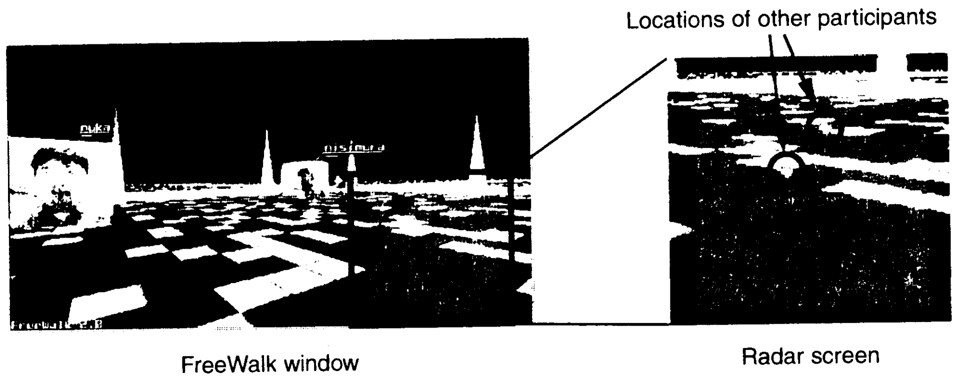


Figure 4.3: Radar Screen

- Viewpoint Switching

  In some situation, the bird's-eye view is more suitable for grasping the situation. The viewpoint switching function allows users to use multiple viewpoints to select an appropriate view. Figure 4.4 shows the bird's-eye view from the back of the user's character. This view enables the user to watch both the user's character and his/her surroundings. As a result, the user can have a better view of geometric relations among participants, and thus move easily than using the normal view.

4.3 Sharing Web Documents

In real life, people often watch magazines or TV together to make their conversation richer in topics. FreeWalk has a function, which enables

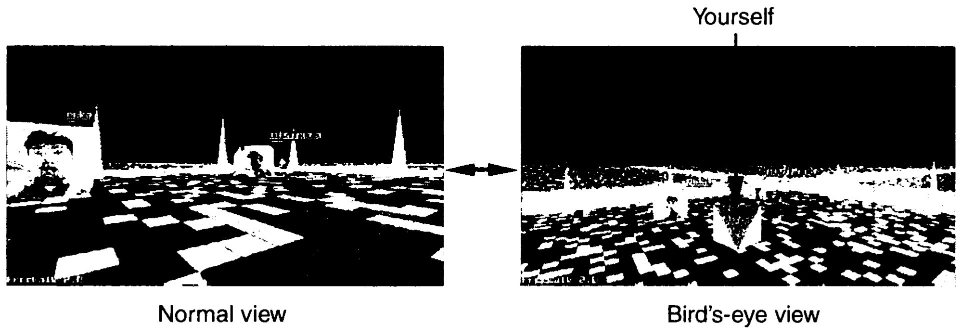


Figure 4.4: Switched Viewpoint

participants to generate topics by watching common information resources. A participant uses this function to make others' Web browsers display the same Web document by transferring URL of his/her Web browser to others. Participants can watch the same Web document to have a conversation in three-dimensional virtual space with a rich stock of topics.

Figure 4.5 shows the route of URL transfer. In this figure, user A transfers the URL, which indicates the document in the user A's browser, to user B by pointing to user B drawn in user A's FreeWalk screen. As a result, user B's browser displays the same document as user A's browser displays. Another possible way is to set up three-dimensional objects corresponding to blackboards or bulletin boards in a three-dimensional virtual space. But, by reducing the number of three-dimensional objects to a minimum, the speed of three-dimensional drawing is kept high in FreeWalk so that participants can move smooth in a three-dimensional virtual space. Therefore, we did not choose this approach.

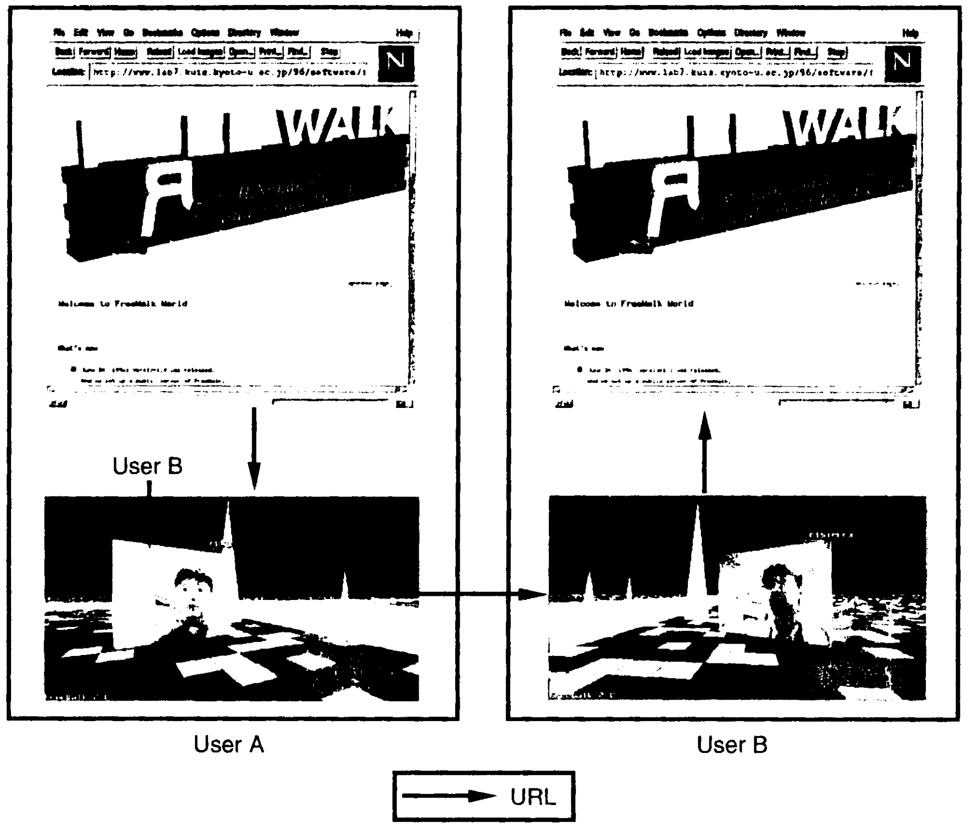


Figure 4.5: URL Transfer

4.4 Using Large Screen Interface

We implemented the FreeWalk system on a room environment as well as on a desktop environment. We used a special room with a large-scale projector screen connected to a graphics workstation. Figure 4.6 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

Figure 4.6: Virtual Space on a Large Screen

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.

4.5 Preliminary Experiments

4.5.1 Experiments in Laboratories

We organized six clients in different rooms of our department building and validated our implementation policy. The major results we obtained are as follows.

- Each participant could move according to his/her own will. The six people formed several groups from time to time. People reported that they could share the same space without confusion.

Various behaviors were noted, such as approaching a pair of participants talking with each other to eavesdrop on their conversation, and chasing a moving participant while calling him/her to stop. Most of the participants enjoyed the experience due in part to its relaxed atmosphere.

- Most users felt that its user-interface is similar to videogames, intuitively understandable, and easier than other videoconferencing systems. There were a few people who found it difficult to control the moving speed using the distance between the mouse pointer and the center of the FreeWalk screen. For this type of user, we added a mode where a user specifies only the direction of movement by the keyboard while the speed is set constant. Since the radar screen covers a wide area, it is not easy to distinguish adjacent participants. We are planning to make the range variable and customizable by users.

4.5.2 Intranet Experiments

We experienced an intranet meeting with FreeWalk in the event called Open Campus (the campus was open to public) held in Tohwa University. The visitors of the event joined FreeWalk meetings without any scheduling beforehand. The meeting continued about six hours and a maximum of 13 users participated simultaneously. As a result of analyzing the log data, interesting behaviors of users in the virtual space were found as follows.

- Most people move around the center landmark of the space

  Few participants tried to go far from the center. As population density around the center became high, network traffic exploded. This is because multimedia data of many participants are transmitted to each client though he/she does not talk to most of them.

- A group of people moved together

It was often observed that a couple of people moved together to a long distance, but seldom more than three. Some users reported that they wanted to ride a bus, because it is hard to move together.

- Some people wandered from a group to group

  In the latter half of the meeting, a number of participants who moved around decreased. Moving participants then wander from a group to group. It is very often observed that a couple of people face each other.

4.5.3 Internet Experiments

We also conducted a preliminary experiment to verify whether our implementation of FreeWalk is competent to communicate through Internet. In this experiment, four users at Kyoto University in Japan and one user at University of Michigan in the United States joined the community server of FreeWalk at Kyoto University.

The frame rate of three-dimensional drawing was about 10 frames per second, the same as in the previous intranet experiments. Though the delay of the live video was longer than that of an intranet, it was inconspicuous and did not affect the control of the player's character much. The users reported they could hold a meeting as good as through an intranet. Sometimes the bandwidth between Japan and the United States forced us to lower the video frame rate to 4 frames per second. However, the users was still able to find others smiling through the live video.

The delay of audio was inconspicuous, too. However, the audio of the user at the United State sometimes became intermittent and the other users were unable to catch what he said while he could clearly hear the voice from Japan. Farther experiments found that the loss of UDP (Internet User Datagram Protocol) packet transmitting audio data caused the intermittent audio. The market-based approach to control quality-of-service [Yamaki96b, Yamaki98a, Yamaki98b] can be one solution.

4.6 Summary

In this chapter, The system architecture, the data transmission mechanisms, applied videogame technologies, the document sharing function, the optional large screen interface, and the performance evaluation are explained.

The FreeWalk system consists of a community server controlling information about locations and view directions of users, and clients, each of which corresponds to each user. Video and voice data are transmitted among clients.

To minimize the network traffic, each user's video is not sent to all the other clients, but is sent to the only clients that need it to draw their screens. Furthermore, the size of the video image is adjusted beforehand according to the receiver's view. Similarly, each user's voice data is sent to the only clients that stand enough close to the user to listen to his/her voice.

Videogame technologies are introduced to the implementation of FreeWalk. FreeWalk has a radar screen to help users to grasp the situation in the space, and a pseudo-joystick function for controlling position smoothly. These techniques are cost-effective.

FreeWalk has a function to enable meeting participants to see a common Web document so that they can use the document as the resource of conversion topics. The large screen version of FreeWalk was developed to make a whole room involved in a virtual space. These additional features are useful to extend virtual meeting spaces.

In the experiment inside our laboratory, six users could freely control their positions to form conversation groups. In the experiment in Tohwa University, thirteen users could gather in the space. We observed various behaviors such as moving around the landmark of the space, moving together, and wandering from a group to group. The Internet experiment between Kyoto University and University of Michigan showed that the implementation enables users to communicate through the Internet. However, several technical problems were caused by network congestion.

Chapter 5

Interaction Analysis for Virtual Meeting Space

In this chapter we show the characteristics of three-dimensional communication compared with face-to-face (FTF) and conventional video-mediated communications.

5.1 Comparison of Communication Environments

There are many functional differences between three-dimensional virtual meeting spaces and conventional videoconferencing systems. We compare these two kinds of communication environments to show the inherent advantages of virtual meeting spaces for casual meetings. We took Silicon Graphics' InPerson (see http://www.sgi.com/software/inperson/) as the example of a conventional videoconferencing system and FreeWalk as the example of a virtual meeting space. Table 5.1 shows the functional differences between these two environments.

- Process of joining

  In FreeWalk, the process used to join a meeting is just to enter the three-dimensional virtual space. Each user selects which virtual

Table 5.1: Functional Comparison

	FreeWalk	InPerson
Process of joining	Enter a 3D virtual space voluntarily	Called by someone who has already joined
Maximum number of participants	Unspecified (practically, 20 or so)	7
Occurrence of conversation	Caused by participants' approach of their own accord	Caused by turning on the system by a coordinator
Meeting group	Multiple groups	Single group

space to enter when he/she starts up the system. The conversation protocol of InPerson inherits that of telephones: in order to hold a meeting between two persons, one should call the other via InPerson. If one wants to join the meeting, he/she needs to be called by someone who has already joined it. A newcomer cannot join an InPerson meeting freely.

- Maximum number of participants

  The maximum number of participants is seven in InPerson. This limitation results from the size of workstation displays. On the other hand, FreeWalk does not limit the number of participants, though if the number exceeds 20, the performance of the system becomes intolerable given the current condition of computer networks.

- Occurrence of conversation

  In FreeWalk, a conversation may be started by an accidental encounter while the participants are walking around the three-dimensional virtual space. A conversation is started by participants' contact of their own accord. In the case of InPerson, conversation is started when the coordinator of a meeting turns on the system and contacts all participants.

- Meeting group

  In FreeWalk, participants approach one another to organize a meeting group. Participants can form multiple meeting groups simultaneously. In the case of InPerson, however, participants always form a single meeting group since everyone faces the others and hears the voices of the others.

5.2 Hypotheses

We believe that three-dimensional environments are more effective for casual communication than conventional video environment as follows.

- Participants using a conventional videoconferencing system tend to be strained and their conversations do not proceed smoothly. This is because all their faces are always displayed and the system keeps everyone facing the others. A three-dimensional virtual space eliminates this strain by giving them locations and view directions.

- It is impossible to reproduce communication with moves like real life communication in a conventional video environment. A three-dimensional virtual space reproduces communication with moves by enabling participants to move freely.

Sellen compared communication in two video conferencing systems, *Hydra* and *Picture-in-a-Picture (PIP)*, and in the FTF environment [Sellen95]. 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.

Bowers *et al.* studied how the movement of avatars coordinated with conversation in a virtual environment [Bowers96]. The results showed that the avatars' moves transferred the initiative of conversation. In three-dimensional 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.

5.3 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 three-dimensional communications (see Figure 5.1). We set up seven SGI O2 workstations connected by a 100-Mbps Ethernet for the video environment (InPerson) and the three-dimensional environment (FreeWalk). The meetings in the three environments consisted of three tasks as follows.

1. Agreeing on a group travel destination (Task 1)

   This was a decision-making task. We made the participants decide where they would travel a month later. They were asked to pretend to be friends in high school days. Also, they did not have many chances to meet after they left the high school.

2. Discussing social problems (Task 2)

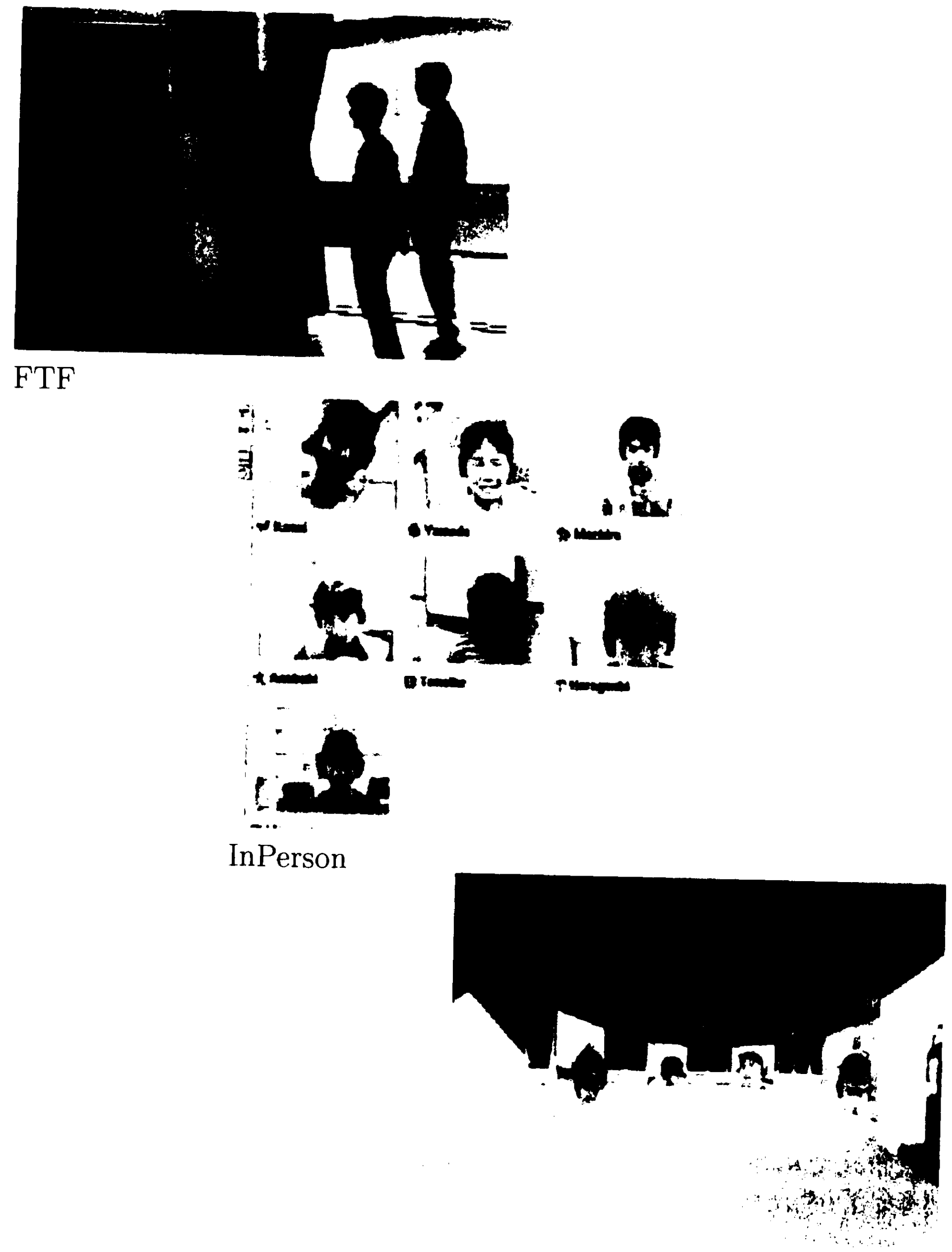


Figure 5.1: Three Different Environments for Conversation

This task was to shape ideas. They were asked to pretend that they attend the same lecture and had to hand in reports.

3. Conversing freely (Task 3)

   Participants had conversation without any guidelines.

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 twenty 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.

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 three-dimensional virtual space during meetings. The FreeWalk community server stores system logs in which it records locations and orientations of participants in a three-dimensional virtual space. We made a tool called *SimWalk* to analyze participants' moves. SimWalk draws lines along the participants' moves and connecting their locations in sequence. It also can reproduce participants' moves by drawing animated triangles, each of which represents the location and view direction of each participant. Those triangles blink to indicate utterances of participants.

5.4 Results

In this section we present the results of analysis of participants' conversations and moves.

5.4.1 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 a 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 5.2 shows the 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 is characterized as:

   FreeWalk $>$ FTF $\approx$ 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 5.2 summarizes the standard deviations of utterance. It provides the following ranking of environments for each task:

   Task 1 FTF $>$ InPerson $>$ FreeWalk

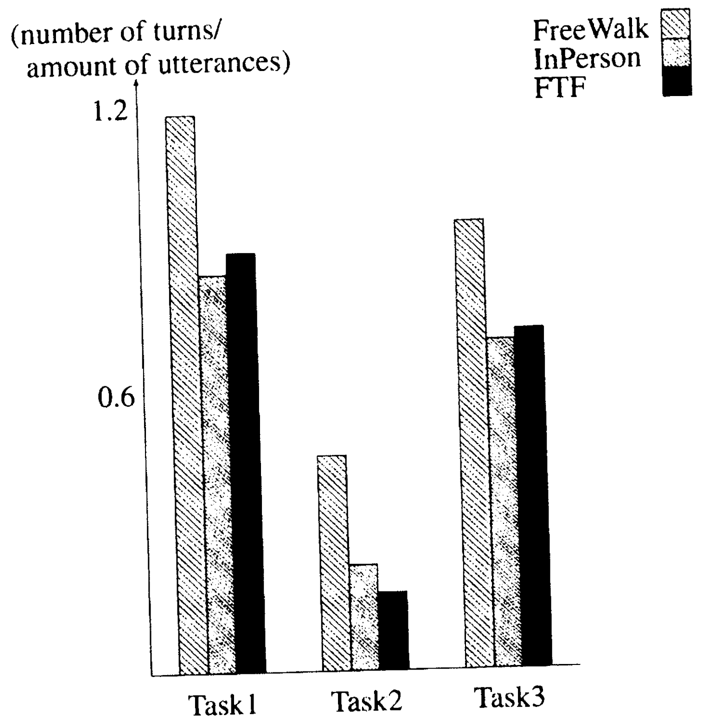


Figure 5.2: Frequency of Turns

Task 2 FTF $>$ InPerson $\approx$ FreeWalk

Task 3 InPerson $>$ FTF $\approx$ 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 does not contribute to accomplishing the task. Figure 5.3 shows the occurrence of chat in Task 1 and Task 2 in each environment. In this figure, the horizontal axis represents time, and each mark represents the occurrence of chat. You can see that chat occurred

Table 5.2: 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

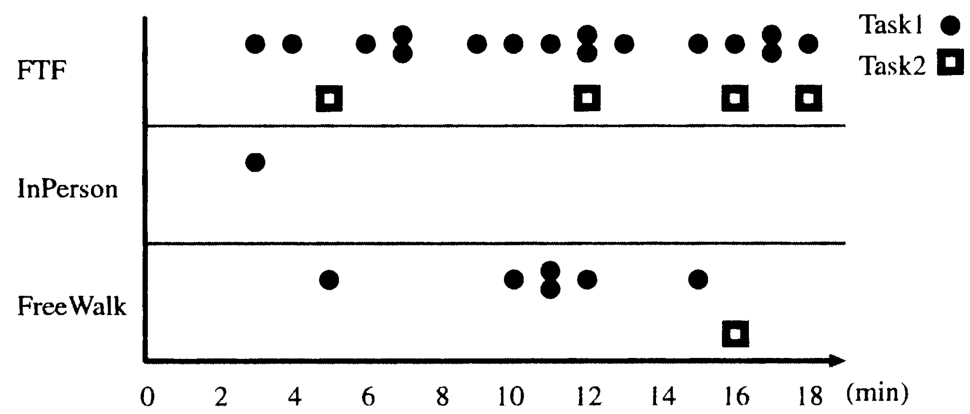


Figure 5.3: Occurrence of Chat

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.

5.4.2 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 5.4 shows participants' moves during a 15-minute period in each FreeWalk meeting. 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 three-dimensional virtual space in Task 3. In Task 3—free conversation—we observed the following behaviors:

1. Moving in a three-dimensional virtual space

   At the beginning of the task, participants moved actively. For example, they moved to the edge of the three-dimensional virtual space and rushed toward others. The occurrence of conversation was scarce. Figure 5.5 presents snapshots of SimWalk, which reproduces the participant's moves.

2. Facing each other to greet

   In the middle of the task, participants faced one another frequently to greet. The lengths of conversations were short. We noted that some participants blamed others for approaching them when they tried to whisper to each other. You can see the participants greeting each other in Figure 5.6.

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. This situation is represented in Figure 5.7.

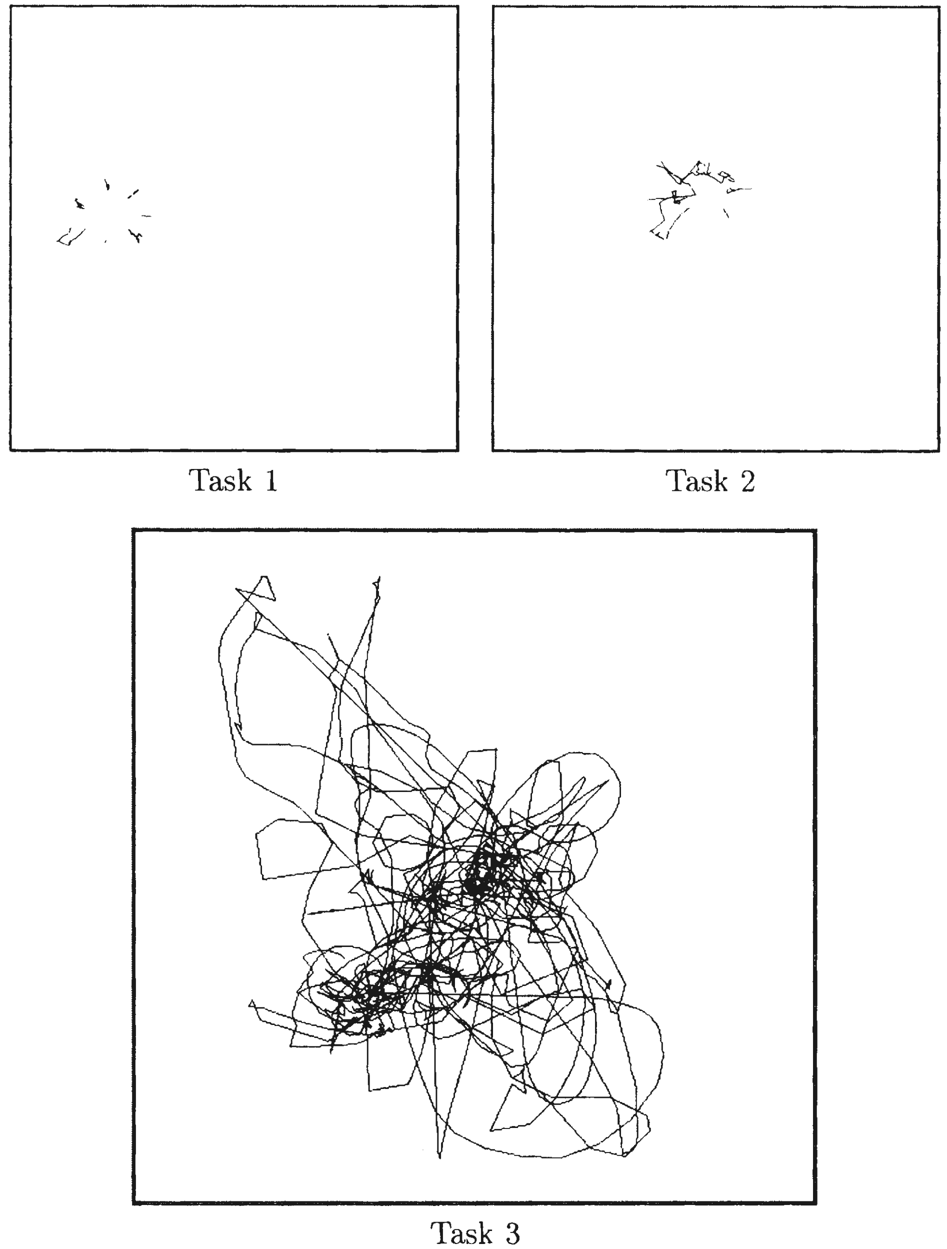


Figure 5.4: Pattern of Moves in a Three-dimensional Virtual Space

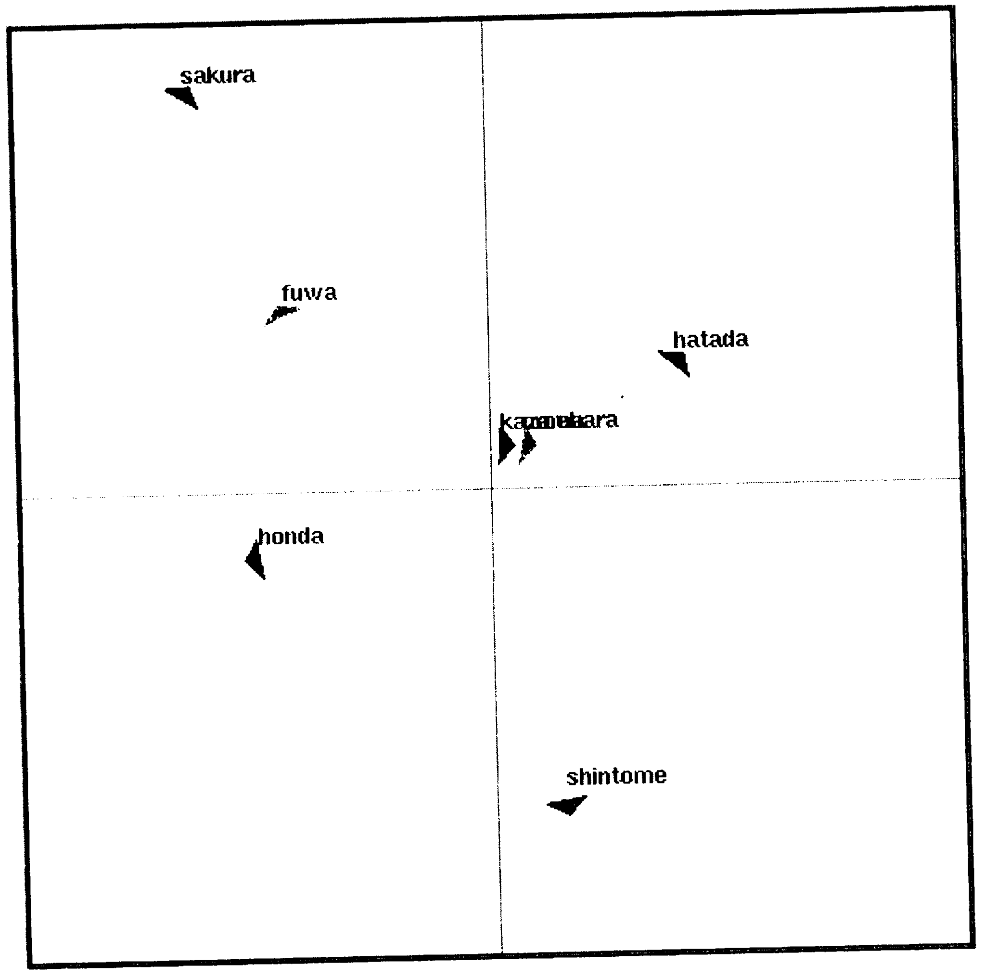


Figure 5.5: Moving in a Three-dimensional Virtual Space

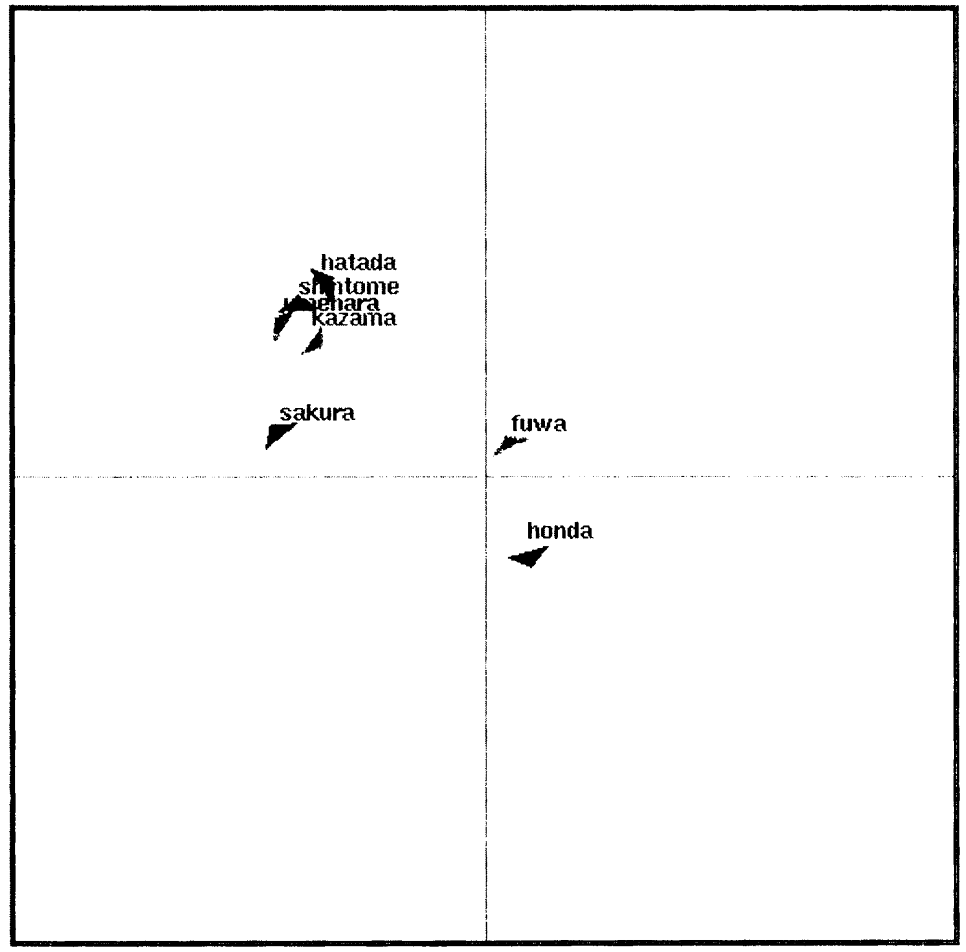


Figure 5.6: Facing Each Other to Greet

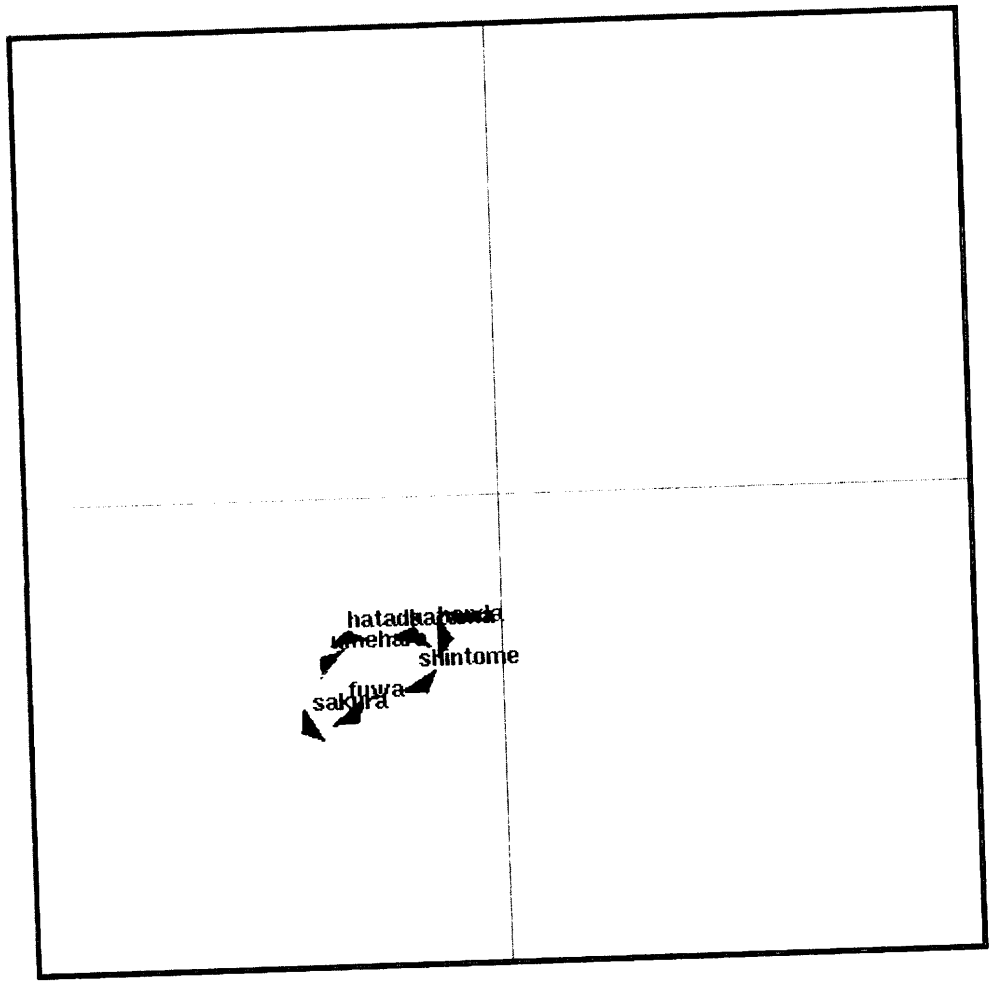


Figure 5.7: Gathering to Start Conversation

5.5 Discussion

As a result of our analysis, we categorized the effects of a three-dimensional virtual space into two types. In the first type, we observed that three-dimensional communication resembles FTF communication. Two primary characteristics exist: frequency of chat and behavior of participants. The second category, however, remains peculiar to three-dimensional 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.

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

The difference of participants' moves in the three FreeWalk meetings gives us important implications about the advantages of virtual meeting spaces. In Task 1 and Task 2, participants seldom moved after forming a circle, since they have common topics to discuss together. In this case, virtual spaces have no advantage except assigning positions, which provide directional voices and videos. Such kind of meetings as Task 1 or Task 2 can be held in conventional videoconferencing systems. However, participants formed multiple groups at once to greet and chat with others in Task 3, since they did not have any common topics to discuss. They cannot have conversations like that in conventional videoconferencing systems. Therefore, virtual spaces have an incomparable advantage in free conversations. Virtual spaces can support the special style of casual communication as well as make video-mediated communication more casual.

5.6 Summary

In this chapter, we discussed the comparison of communications in a three-dimensional virtual meeting space (FreeWalk) with communications in a conventional desktop conferencing system (InPerson) and face-to-face meetings.

In InPerson, one calls the other to hold a meeting through the telephone-like protocol. To hold a multiparty meeting, the coordinator of a meeting has to call all participants. This means that a newcomer needs to be called by someone who has already joined the meeting. Moreover, the maximum number of participants is strictly limited, since video frames of all participants should be displayed on a workstation screen. This interface also makes participants form a single meeting group to discuss common topics. FreeWalk differs from InPerson in these ways. In FreeWalk, one enters the three-dimensional virtual space to join the meeting, and may start a conversation by an accidental encounter. FreeWalk does not limit the number of participants, and not prevent participants from forming multiple meeting groups at once.

Subjects of the experiment accomplished three tasks: agreement for the destination of group travel, discussion about social problems, and free conversation. We chose these tasks to examine various types of communication comprehensively. For each task, we told subjects to organize three groups of seven people. Each group was assigned to one of the three environments.

As results of the analysis, we have categorized the effects of a three-dimensional virtual space into two types. The first type is to make video communication resemble face-to-face one. This type of the effects is found in the frequency of chat and the behavior of participants. The second type is peculiar to three-dimensional virtual space. Three-dimensional virtual space equalizes the amount of utterances for each participant more than the other environments, increases the number of turns, and sometimes stimulates participants in moving around to real-

ize free conversation. The freedom of three-dimensional virtual spaces seems to enable participants to relax, and that atmosphere may stimulate participants into talking easily. The difference of participants' moves in the FreeWalk meetings tells that virtual spaces can support the flexible style of casual meetings.

Chapter 6

Social Agent for Virtual Meeting Space

In this chapter, we present a social agent that is an embodied character assisting social interaction in a virtual meeting space.

6.1 Social Agent

Virtual meeting space enables people to meet accidentally and to have multiple conversations simultaneously. There are many virtual meeting spaces such as Community Place [Lea97], InterSpace [Sugawara94], Diamond Park [Waters97], DIVE [Hagsand96], Massive [Greenhalgh95] and CU-SeeMe VR [Han96]. Virtual meeting spaces make it easy to have casual meetings between strangers from across town, or even across the world. Virtual meeting spaces usually provide little socially meaningful context to use as a basis for finding common ground with each other. Since it is easy to arrive at a virtual meeting space from many entry points, it is often hard for visitors to assume much about one another's cultural backgrounds, group memberships, and other aspects of social identity. People need this sort of common context in order to build new human relationships [Clark96].

We believe social agents could provide ongoing, in-context help in forming social relationships and building common ground between visitors to virtual meeting spaces. We developed a social agent playing a role of party host in a virtual space. This agent was applied to support cross-cultural communication in our experiment.

6.1.1 Related Work

Previous studies have discussed and demonstrated some benefits of interface agents in one-on-one task settings, such as taking an educational tutorial [Lester97], going on a tour [Isbister99], or looking at real estate [Cassell99]. Lester *et al.* notes that the presence of an agent can lead to a strong positive effect on students' perception of their learning experience [Lester97]. Cassell *et al.* discusses the value of an embodied conversation partner with the proper human verbal and nonverbal communication skills [Cassell99]. However, these findings concern task-support agents interacting with a single user.

There are projects, which have created agent-based social support through text-based conversation. Julia [Foner97] plays a role of a guide in virtual worlds of MUD; the Extempo bartender agent converses with chat visitors, and is designed to enhance the social atmosphere [Isbister97]; and there are bots for Web sites that answer questions and direct visitors in a friendly way (e.g. http://www.artificial-life.com). However, these agents are designed to engage in one-on-one social interactions, rather than facilitating human-human interaction. There are few studies about agents, which interact with multiple people. Takeuchi and Naito proposed a synthesized facial display as a conversational partner in multiparty meetings [Takeuchi95].

The social agent we developed differs from the agents described above, which support specific tasks or play a role of a conversation partner. Our agent aims to work as an in-between of human-human interaction. Another example of agent facilitating human-human interaction provides information to support thought process [Nishimoto99].

Our agent is designed to conduct simple question and answer so that people whose conversation is faltering can find a common topic to

talk about. Another possible solution for such an awkward situation is providing an information search tool to find a common topic based on the retrieved data about the social identities of conversation partners. However, that tool does not help the process to start a conversation. There is a gap between finding topics and beginning conversations. Through question and answer, people can share one another's answer to the same question. That is an opportunity to start a conversation based on the answers. Furthermore, it may be invasive for the participants' privacy to collect personal information about conversation partners.

6.1.2 Cross-Cultural Communication

For testing our agent, we focused on an extreme case of low social context in a virtual meeting space: strangers from different national cultures, meeting for the first time. Even when people can use a common language with reasonable fluency, they do not necessarily have a common context for their conversation. Different cultures have different notions of how to begin and develop conversations. What is a safe topic that is unlikely to harm the conversation and destroy the relationship in one culture, may be very unsafe in another culture. For example, in some cultures it is appropriate to ask about family members right away; whereas in other cultures this is private [Hall90, Clark96]. Since it is very hard to establish a common ground in this sort of meetings, we thought we could find the clear effect of our agent's assistance in conversations.

We focused on conversations between Japanese and Americans. These two national groups are known to have very different interaction styles and cultural norms [Hall90], and so we felt this was a good test case.

6.2 Design of the Agent

In a virtual meeting space of FreeWalk, our agent basically acts in the same way of a busy party host looking for clues that the guests' conversations are going badly. The agent tracks audio from a two-person conversation, to look for longer silences that will trigger its conversation aid. Pauses are a powerful cue for what is happening in a conversation [Clark96]. When the agent finds the pause, it approaches to the conversation pair. The agent then directs a series of yes/no questions to both conversation partners in turn, and uses their answers to guide its suggestion for a new topic to talk about. Then the agent retreats until it is needed again.

6.2.1 Nonverbal Communication Abilities

In virtual spaces, our social agent is embodied the same way of users (see Figure 6.1). This allowed us to take advantage of nonverbal cues in designing the agent, such as a spatial position and direction for turning to face users, and animated pictures to present facial expressions and gestures.

The agent approaches to the conversation pair to direct yes/no questions when it detects an awkward silence in their conversation. After concluding a suggestion cycle, the agent departs from the conversation zone, and wandering at a distance, until it detects another awkward silence. This makes it easy for the conversation pair to know whether the agent joins their discussion [Hall66]. The agent orients its face toward the conversation partner it is addressing so that the pair can intuitively recognize whom the agent asks.

On the rectangle of the agent's embodiment, an animated dog is pasted. We chose a dog because we wanted users to think of the agent as subservient, friendly, and reasonably socially intelligent [Reeves96]. The agent has a set of animations for asking questions, reacting to affirmative or negative responses, and making suggestions. Each of them corresponds to a phase in the process of question and an-

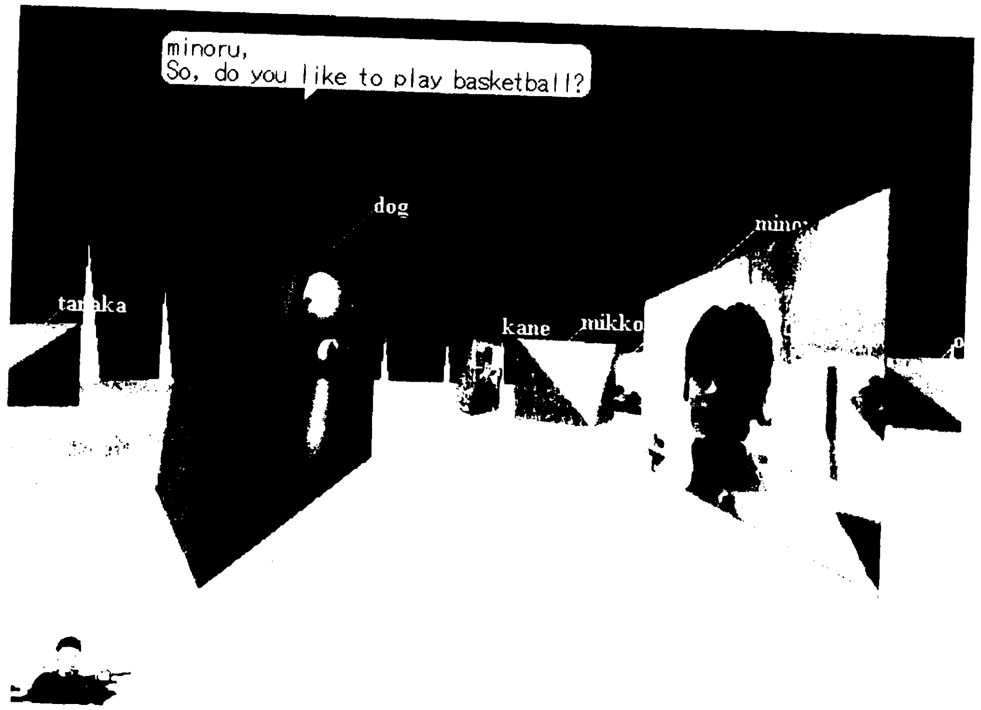


Figure 6.1: Social Agent in a Virtual Meeting Space

swer. We crafted these animations as a supplement to the agent's speech [Cassell99].

6.2.2 Topic-Suggestion Mechanisms

Silence-detection

The agent decides there is silence when the sum of the voice volumes of both participants is below a fixed threshold value. When the agent detects a silence that lasts for more than a certain period of time, it decides the participants are in an awkward pause.

Positioning

The agent decides how to position itself, based on the location and orientation of each participant. The agent turns toward the participant that it is currently addressing. If the participants move while the agent is talking, the agent adjusts its location and orientation. The agent tries to pick a place where it can be seen well by both people, but also tries to avoid blocking the view between them. If it's hard to find an optimal position, the agent will stand so that it can at least be seen by the participant to whom it is addressing the question.

State-transitions

The agent has three states, which are idling, approaching, and talking. When idling, the agent strolls at the corner of the virtual space, further away than the normal conversation zone [Hall66]. When the agent detects an awkward pause in the participants' conversation, it begins an approach. Upon reaching the participants, the agent goes into the talking state. However, if the participants start talking again before the agent reaches them, it stops the approach and goes back to idling. This behavior is strikingly similar to the actions of a hesitant subordinate trying to approach a superior, who is engaged in a conversation with another dominant person. The agent will also remain in idling state if the participants are standing far apart from each other (out of conversation range), or are not facing each other. If the participants turn away from each other during the agent's approach, or while it is talking, it will return to idling state, as well.

6.2.3 Topic Knowledge

We gathered safe and unsafe topics for the first time meeting, using a Web survey, which university students from Japan and the United States filled out. We used the collected pool of topics to select common safe and unsafe topics for people from both countries. From these topics, we crafted a set of questions that the agent could ask in the

question and answer process. Safe topics included: movies, music, the weather, sports, and what you did yesterday. Unsafe topics included: money, politics, and religion. A sample safe question: "Is the weather nice where you are right now?" A sample unsafe question: "So, do you think it is alright for a country to fish for and eat whales?"

6.2.4 Conversation Model and Interface

The user interface for communicating with the agent is very easy to learn. The agent presents questions to the participants in a text-balloon above its head. We did not use synthesized voice because we were afraid that unnatural utterance may affect participants, and participants may fail to catch what the agent says. The participant indicates 'yes' or 'no' by clicking the mouse on his/her answer displayed under the question in the text-balloon. We did not use natural language as an input interface to prevent participants from expecting too much intelligence of the agent, since they might be frustrated by not smooth conversation with the agent. Both participants see all questions, but only the addressed person sees the Yes/No options. When the person answers the question, his/her answer is displayed in a text-balloon above his/her own embodiment (see Figure 6.2).

Each topic has a tree structure, with nodes that are: a question for a participant, possible answers by participants, agent's reply to each answer, and flags indicating whether the agent will address its next question to the other person or to the same person (see Figure 6.3). Topics were designed to draw participants into a dialogue, so each turn is tailored for this purpose. Basically, the agent asks both participants the same question to draw shared or conflicted points from the interaction. The cycle always concludes with a recommendation for how the participants could make use of the particular topic area, given their own answers to the agent.

When the agent approaches to start a cycle, it selects a topic from its repertoire of safe (or unsafe) topics randomly, out of those that have not yet been used. Then it randomly chooses one of the two partic-

Figure 6.2: Conversation from Both Participant's Point-of-view

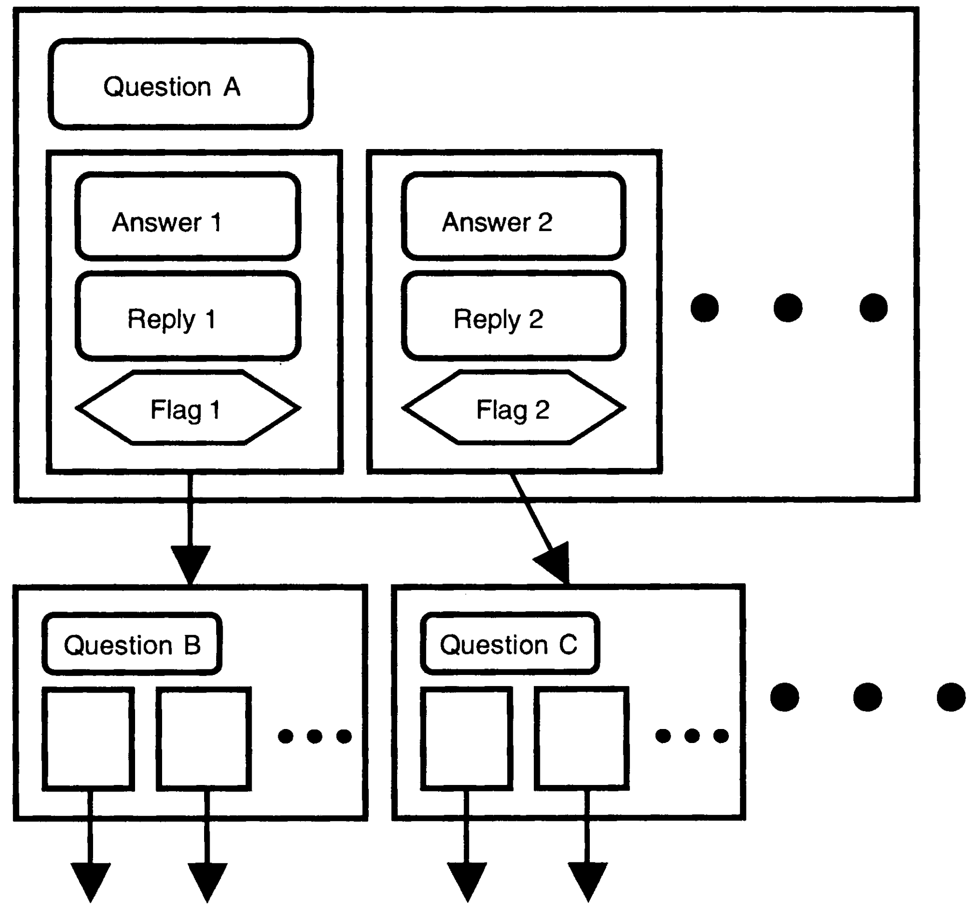


Figure 6.3: Tree Structure of Topic Data

ipants as the target for the first question. Let's call this person A. When A answers, the agent replies to A's answer. Based on what A answered, the agent then chooses a follow-up question. This question might be directed at A or at B. If it is directed at B, the agent turns to B to pose the question. When B answers, the agent makes a general comment that is meant to guide the participants into using this topic. This general comment is selected based upon the previous answers from the participants. Figure 6.2 shows a part of this cycle from both participant's point-of-view. In this figure, (1) person A is asked the first question (2) and responds, (3) then the agent comments. (4) Next person B is asked a question. As we described above, the agent faces the person it is addressing.

After making this comment, the agent departs. If at any time a user does not respond to the agent's question, the agent will wait for an interval, and then go back into idling mode, without trying to continue its question cycle. This makes it clear that the agent is a in-between only to lead a question cycle. Therefore, participants can intuitively understand they do not need to include the agent in their conversation.

6.3 Cross-Cultural Communication Experiment

6.3.1 Hypotheses

We focused on the relation between the effects of the agent on social interaction and the difference in the cultural tendency of the topics the agent provided. Our initial expectations were:

1. The safe-topic agent would create a more satisfying experience, than if there were no agent. Participants would feel they were more similar, would be happier with the interaction and conversation partner, and would form more positive impressions of one another's nationality.

Figure 6.4: Set-up for the Experiment (Stanford Side)

2. The unsafe-topic agent would make people uncomfortable, but might lead to a more meaningful and interesting conversation than the safe-topic agent.

6.3.2 Procedure

The study was a collaboration among NTT, Kyoto University and Stanford University. We used a 1.5-Mbps dedicated line to connect both PCs in the two universities. The two research teams used chat software to communicate while running the study. We set up a PC with a camera and headset at each location (see Figure 6.4).

We designed a three-condition experiment using pairs of students who were located in the United States and in Japan. Pairs either interacted one-on-one, or had the help of the safe-topic or unsafe-topic agent. We divided the twenty-minutes conversation session into five segments, and forced the agent to display a topic within each four-minutes segment. The agent looked for an awkward pause during a minute in each time segment. The agent introduced topics immediately if it could not find a pause. Thus, in the safe-agent condition, the agent introduced all five safe topics in random order. In the unsafe-agent condition, the agent introduced all five unsafe topics in random order.

Each research team recruited students for the study. The Stanford students were a part of an undergraduate class, which required participation in experiments for credit. The Japanese students were undergraduates from Kyoto University and other nearby universities, who were paid for their participation. Because the study would be held in English, we screened Japanese students and selected those who scored at a reasonably high level on English proficiency tests. Both sets of students were screened for a high level of familiarity with one another's culture, and those with high experience were not asked to participate. In total we had ninety participating students. Students were assigned randomly to same-gender pairs. Each pair was randomly assigned to one of the three conditions.

Students were told that they would be testing out a new communication environment with a student from the other country. They were asked to talk about anything they liked. They were trained in how to use the system, then left alone to talk for twenty minutes. We made video recordings of all sessions by capturing what was on the screen in the Kyoto side onto videotapes. After their conversation, participants filled out a survey presented on a Web browser.

We prepared the survey in their native language. Questionnaire items of the survey were translated and then reverse-translated for accuracy. The questionnaire included questions about the conversation, their conversation partner, and the agent (in agent conditions). We also asked them to make assessments of themselves, and the typical person of both participants' cultures on some commonly used stereotypic adjectives. We ended up with data from forty-five Japanese students, and forty-three American students (because we could not get the questionnaire results from two subjects in the unsafe agent condition), for our analysis.

6.3.3 Results

Safe Agent versus No Agent

We summarized the values of only the questionnaire items that had significant differences.

American Reaction

The safe agent had positive effects for American participants as we expected (see Table 6.1; all items on an 8-point scale, 8 highest):

- *opinion of their own behavior higher*
  They rated themselves as more confident, less domineering, and less restrained in the safe agent condition.
- *opinion of partner higher*
  They rated their partner as significantly more trustworthy in the safe agent condition.
- *opinion of the typical Japanese person higher*
  The safe agent condition had a positive effect on impression of typical Japanese people. Those in the safe agent condition rated the typical Japanese person as more creative and more friendly. However, they rated the typical Japanese person as less emotionally expressive. (Americans typically stereotype Japanese people as less creative, less friendly, and less emotionally expressive.)

Japanese reaction

The Japanese participants had a different response to the safe agent's presence. The safe agent did not improve their experience. However, it made them think their partner was more like themselves, as we expected (see Table 6.2).

Table 6.1: Summary of T-test Comparisons of American Students' Ratings, Safe Agent versus No Agent

variable	mean (safe agent)	mean (no agent)	t-value (n=13)
confident	6.46	5.54	-2.33**
domineering	4.00	4.92	2.03*
restrained	3.61	5.00	2.52**
partner trustworthy	6.54	5.91	-2.46**
Japanese creative	5.38	4.54	-2.06*
Japanese friendly	5.92	5.23	-2.08**
Japanese emotionally expressive	3.15	4.23	2.75***

*p=.05, **p<.05, ***p=.01

- *opinion of the experience lower*
  Japanese in the safe agent condition rated the experience as less safe and more uncomfortable. They were less interested in continuing such a conversation, and were less satisfied afterward.
- *opinion of their own behavior lower*
  They also rated themselves in a more negative light in the safe agent condition. In the safe agent condition, participants rated themselves as more evasive and quieter than participants did in the no agent condition.
- *opinion of their partner mixed*
  Their ratings of their American partners were mixed. In the safe agent condition, participants found their partners more talkative and more effusive, and less engaging. (Japanese

Table 6.2: Summary of T-test Comparisons of Japanese Students' Ratings, Safe Agent versus No Agent

variable	mean (safe agent)	mean (no agent)	t-value (n=15)
unsafe	3.29	2.24	-2.05*
uncomfortable	5.14	2.71	-3.9*****
desire to continue	4.86	7.07	3.55****
satisfying	4.79	6.14	2.32**
self evasive	5.86	4.71	-2.09**
self quieter	4.68	3.36	-2.08**
partner talkative	5.00	4.07	-2.06**
partner effusive	2.21	1.50	-2.06**
partner engaging	6.00	6.78	2.47**
partner typically American	5.22	6.14	2.26**
partner similar to self	5.28	4.21	-2.20**
Americans competitive	6.57	5.00	-2.40**
Americans domineering	6.07	3.36	-4.44******
Americans selfish	6.14	4.93	-2.26**
Americans effusive	6.79	5.86	-2.21**

*p=.05, **p<.05, ***p=.01, ****p<.01, *****p=.001, ******p<.001

tend to stereotype Americans as talkative and emotionally effusive.) Yet, they rated their partners as less typically American and more similar to themselves.

- *opinion of the typical American person lower*

  The safe agent condition seemed to exacerbate negative views of Americans for Japanese participants. In the safe agent condition, they rated the typical American as more competitive, more domineering, more selfish, and more effusive than those in the no agent condition. (All of these are stereotypical American traits, from the Japanese point of view.)

We cannot be sure why the two groups had such different reactions. One reason may be that the agent's questions were implemented in English. It's possible that Japanese subjects felt it was a two-against-one situation. This might explain why they disliked the interaction, even though it seemed to make them rate their partner as more similar to themselves. Another reason may be that Japanese subjects disliked the sudden interruptions by the agent that failed to find an awkward pause. Most of Japanese subjects seemed to be interested in talking with Americans.

Safe Agent versus Unsafe Agent

- *Awkward isn't necessarily bad*

  As we had expected, the unsafe agent made things more awkward, but also more interesting. We counted awkward pauses, by observing the videotapes, and found a higher number of awkward pauses in the unsafe agent condition versus the safe agent condition (Means = 4.34(unsafe) and 3.09(safe), $t(56)=-3.06$, $p<.01$).

  Despite the higher level of awkwardness in these conversations, both Japanese and American participants found the conversation that included the unsafe agent more interesting. Americans rated

the unsafe agent interaction more interesting; Japanese rated the unsafe agent experience more desirable to continue and more comfortable (see Tables 6.3 and 6.4).

- *American partner seemed worse in the unsafe agent condition*

  Japanese participants rated their partner as less similar to themselves, less considerate, more domineering, less friendly, and less talkative in the unsafe agent condition. These rankings suggest that the unsafe agent led to more negative impressions of the partner, for Japanese participants.

- *Unsafe topics made Japanese act more American*

  Japanese rated themselves as less evasive, less restrained, less self-abasing, and less team-oriented in the unsafe agent condition. These are all stereotypically American traits, for Japanese. It seems they thought they acted more American than those in the safe agent condition. Americans rated their partner as more similar to themselves in the unsafe agent condition. This seems to corroborate the Japanese self-ratings.

- *Safe/unsafe topic choice affected stereotyping in complicated ways*

  Japanese participants in the unsafe agent condition thought the typical American was less domineering. This conflicts with their ranking of their own partner.

  American participants rated the typical Japanese in conflicting ways: after the unsafe agent condition, they thought the typical Japanese person was more emotionally expressive, more outgoing, and more talkative; but also more evasive and quieter.

- *Safe/unsafe agents were perceived differently by Japanese and Americans*

  The two groups differed in their impressions of the safe and unsafe agents. The Americans formed the intended impression: they

Table 6.3: Summary of T-test Comparisons of Japanese Students' Ratings, Safe Agent versus Unsafe Agent

variable	mean (safe agent)	mean (unsafe agent)	t-value (n=15)
desire to continue	4.86	6.21	-2.00 ^
uncomfortable	5.14	3.57	2.41**
self evasive	5.86	4.64	2.03*
self restrained	5.43	3.79	2.37**
self self-abasing	5.54	4.07	2.33**
self team-oriented	4.00	2.64	2.34**
partner similar to self	5.29	3.64	2.58**
partner considerate	7.31	5.93	3.02****
partner domineering	1.14	2.07	-2.43**
partner friendly	7.29	6.14	2.31**
partner talkative	5.00	3.79	2.14**
Americans domineering	6.07	5.00	2.26**
agent nice	3.43	5.29	-2.25**
agent competent	4.29	5.57	-2.04*
agent typically Japanese	4.50	3.43	2.71**
agent talkative	5.61	4.36	2.66***
agent nationalistic	1.43	3.29	-2.87***

*p=.05, **p<.05, ***p=.01, ****p<.01, ^ p=.056

Table 6.4: Summary of T-test Comparisons of American Students' Ratings, Safe Agent versus Unsafe Agent

variable	mean (safe agent)	mean (unsafe agent)	t-value (n=13)
interesting	5.85	6.77	-2.18**
partner similar to self	3.31	4.77	-2.55**
Japanese emotionally expressive	3.15	4.15	-2.16**
Japanese outgoing	4.08	4.77	-2.04*
Japanese talkative	3.77	4.85	-2.30**
Japanese evasive	3.85	4.85	-2.39**
Japanese quiet	6.00	4.38	2.82***
agent blunt	4.69	7.36	-3.84*****
agent domineering	3.38	5.25	-2.31**
agent restrained	3.15	1.92	2.52**
agent friendly	5.46	4.08	2.03*
agent typically American	6.62	4.92	2.40**

*p=.05, **p<.05, ***p=.01, *****p=.001

rated the unsafe agent's topics as less appropriate, thought it acted more blunt, more domineering, less restrained, and less friendly. They also said it was less typically American, distancing it from their own in-group's behavior. The Japanese thought that the unsafe agent was nicer and more competent than the safe agent. They rated the unsafe agent as less typically Japanese, and as less talkative. They found the unsafe agent more nationalistic, probably because it brought up more political topics than the safe agent.

6.4 Lessons Learned

Provocative help can be good

Our evaluation suggested that a communication assistant can be helpful both when it offers safe topics to talk about, and when it steers the conversation in less safe directions. In fact, the Japanese participants seemed to prefer the unsafe topic agent, and both groups found it more interesting than the safe topic agent.

User-adaptation would make the agent more effective

The two cultural groups had very different impressions of the same agent behaviors, and reacted in different ways. For example, behavior that was perceived as blunt and unfriendly by Americans was seen as nice and competent by Japanese. We believe we created a more American identity for our agent by delivering its topic help in English. It may be better to choose the topics the agent delivers and the language it uses based on participants' information retrieved from their personal Web sites.

Agent behavior may shift user behavior

Both the Japanese and American participants noted that Japanese seemed to act more American in the unsafe agent condition. This result indicates that it may be possible to mold user behavior with the choices one makes about how the agent behaves and what it talks.

6.5 Summary

Since it is easy to arrive at a virtual meeting space from many entry points, it is hard for visitors to assume much about one another's cultural backgrounds. Virtual meeting spaces usually provide little social context to find a common ground for communication with each other. To eliminate this difficulty, we developed a social agent supporting human-human communication in virtual meeting spaces. This chapter described the social agent mimicking a party host, and trying to find a common topic for two meeting participants whose conversation has lagged.

Previous studies mainly focused on the autonomous agents to assist users in specific tasks, or to engage in one-on-one interaction as a conversation partner. Our social agent aims to work as an in-between of human-human interaction in casual meetings. The agent provides opportunities to start a conversation.

We selected the cross-cultural first time meeting as a good test case for the agent, since it seems to be very hard to establish a common ground in such a situation. We focused on conversations between Japanese and American, since they are known to have very different interaction styles and cultural norms.

In the FreeWalk space, our agent acts in the same way of a busy party host looking for clues that the guests' conversations are going badly. The agent tracks audio from a two-person conversation, to look for a longer silence that means an awkward pause. When the agent detects a pause, it begins an approach. Upon reaching the participants, the agent goes into the topic-suggestion cycle. However, if the participants start talking again before the agent reaches them, it stops the approach. In the suggestion cycle, the agent conducts a series of yes/no questions to both participants to draw shared or conflicted points. After the cycle, the agent makes a general comment based on the previous answers from the participants to recommend how they could make use of the particular topic area. After making this comment, the agent departs from the conversation zone, and wandering at a distance, until it detects another awkward silence.

In virtual spaces, our social agent is embodied the same way of users. Such moving behaviors of the agent as approaching and retreating makes it intuitive for the conversation pair to understand the agent is an in-between only to lead a question cycle, and they do not need to include the agent in their conversation. The agent turns toward the participant it is currently addressing so that the pair can intuitively recognize who is asked.

On the rectangle of the agent's embodiment, an animated dog is pasted. We wanted users to think of the agent as subservient, friendly, and reasonably socially intelligent. The agent has a set of animations, each of which corresponds to a phase in the process of question and answer. These animations are a supplement to the agent's speech.

We gathered safe and unsafe topics for the first time meeting by using a Web survey, to craft a set of questions that the agent could ask in question and answer. Each topic has a tree structure, with nodes that are: a question, possible answers, agent's reply to each answer, and so on.

We performed an experimental evaluation of the agent's ability to assist in cross-cultural communication between Japanese undergraduates and American undergraduates. We designed two kinds of agents to introduce culturally common safe or unsafe topics to conversation pairs, through a series of question and answer.

In the experiment, the safe agent had positive effects for American students. On the other hand, it had negative effects for Japanese students, but simultaneously it made them think their partner was more similar to themselves. In the unsafe agent condition, both Japanese and American students thought their conversations were more interesting,

and Japanese students acted more American.

As a result, we found that provocative topics are useful, an agent adaptive to participants is good, and an agent's presence affects participants' style of behavior. This experiment demonstrated the agent's effectiveness, and raised interesting considerations for further development.

Chapter 7

Conclusion

7.1 Contributions

In this thesis, we proposed new approaches to support daily conversations via computer network. First, we designed a virtual meeting space *FreeWalk* where social interaction is more casual and relaxed than telephone-like environments. Next, we designed a social agent that is a character acting as an in-between of people to diminish the demerit of the low social context in virtual meeting spaces. We analyzed the effects of the virtual meeting space and the social agent on social interaction. In this section, we summarize the design and analysis of social interaction in FreeWalk and with the social agent.

- *Design of social interaction in virtual meeting space*

  In FreeWalk, video and audio channels transmit nonverbal information such as facial expressions and paralanguage in the same way of other videoconferencing systems, and spatial positions of participants make accidental simultaneous conversations possible.

  The spatial communication model is tightly combined with the video-mediated communication. Each user is embodied as a three-dimensional pyramid object, on which his/her live video is pasted. Each embodiment has a location and a view direction, and moves around freely. The sizes of others' video images change according to the distance and angle of them. Voice volume attenuates in proportion to the distance between sender and receiver. The voice of a user can be heard by anyone in the neighborhood. Users hear others' voice in stereo. The volume and direction of one's voice is determined by its position.

As a result of the design described above, FreeWalk enables many familiar behaviors seen in casual meetings. Before approaching someone to begin talking, users can watch what others are doing from a distance. Conversations may occur when walking users encounter accidentally. Users can form multiple meeting groups, each of which is separately located in the same space. Users can smoothly join the conversation that attracts their interest, since they can guess the conversation topic by listening to the conversation beforehand. Since the speaker can turn to face a person he/she is talking to, the speaker can assume that the listener recognizes who is speaking. All of these behaviors are almost impossible in normal videoconferencing environments.

Although virtual meeting spaces has significant advantages as we explained, the easy access of virtual spaces makes it hard for visitors to assume much about one another's social identities in order to begin talking to. Our social agent mimics a party host to solve this problem.

The agent tries to find a common topic for two meeting participants whose conversation has lagged. The agent tracks audio from a two-person conversation, to look for a longer silence that means an awkward pause. When the agent detects a pause, it begins an approach. Upon reaching the participants, the agent goes into the topic-suggestion cycle. In the suggestion cycle, the agent conducts a series of yes/no questions to both participants to draw shared or conflicted points. After the cycle, the agent makes a general comment based on the previous answers from

the participants to recommend how they could make use of the particular topic area. After making this comment, the agent departs from the conversation zone, and wandering at a distance, until it detects another awkward silence.

This sort of agents can be embedded in other communication environments. However, some behavioral designs of our social agent depend on the nature of virtual meeting spaces. The agent is embodied the same way of users in virtual spaces. The moving behaviors of the agent such as approaching and retreating makes it intuitive for the conversation pair to understand the agent is an in-between only to lead a question cycle, and they do not need to include the agent in their conversation. The agent turns toward the participant it is currently addressing so that the pair can intuitively recognize who is asked. On the rectangle of the agent's embodiment, an animated dog is pasted. We wanted users to think of the agent as subservient, friendly, and reasonably socially intelligent. The agent has a set of animations, each of which corresponds to a phase in the process of question and answer. These animations are a supplement to the agent's speech.

Previous studies mainly focused on the agents to assist users in specific tasks, or to engage in one-on-one interaction as a conversation partner. Our social agent aims to work as an in-between of human-human interaction in casual meetings.

- *Analysis of social interaction in virtual meeting space*

  We implemented FreeWalk and the social agent to see whether they make communication more casual, more relaxed, easier, more meaningful. We conducted two experiments to test FreeWalk and the social agent.

  First, we compared communication in FreeWalk meetings with communications in video-mediated (SGI's InPerson) and face-to-face meetings. As results of the analysis, we have categorized the effects of a three-dimensional virtual meeting space into two types. The first type is to make video communication resemble face-to-face one. This type of the effects is found in the frequency of chat and the behavior of participants. The second type is peculiar to three-dimensional virtual space. Three-dimensional virtual space equalizes the amount of utterances for each participant more than the other environments, increases the number of turns, and sometimes stimulates participants in moving around to realize free conversation. The freedom of three-dimensional virtual spaces seems to enable participants to relax, and that atmosphere may stimulate participants into talking easily.

The difference of participants' moves in the FreeWalk meetings tells that virtual spaces can support the flexible style of casual meetings. Participants seldom moved after forming a circle when they have common topics to discuss together. However, they formed simultaneously multiple groups to greet and chat with others when they did not have any common topics to discuss. In InPerson, the maximum number of participants is strictly limited, since the video frames of all participants should be displayed on a workstation screen. This interface also makes participants form a single meeting group to discuss common topics. FreeWalk does not limit the number of participants, and not prevent participants from forming multiple meeting groups at once.

Next, we performed an experimental evaluation of the agent's ability to assist in cross-cultural communication between Japanese undergraduates and American undergraduates. We designed two kinds of agents to introduce culturally common safe or unsafe topics to conversation pairs, through a series of question and answer. We gathered safe and unsafe topics for the first time meeting by using a Web survey, to craft a set of questions that the agent could ask in the question and answer. We selected the cross-cultural first time meeting as a test case for the agent, since it seems to be very hard to establish a common ground in such a situation. We focused on conversations between Japanese and

American, since they are known to have very different interaction styles and cultural norms.

In the experiment, the safe agent had positive effects for American students. On the other hand, it had negative effects for Japanese students, but simultaneously it made them think their partner was more similar to themselves. In the unsafe agent condition, both Japanese and American students thought their conversations were more interesting, and Japanese students acted more American. As a result, We found that provocative topics are useful, an agent adaptive to participants is good, and an agent's presence affects participants' style of behavior. This experiment demonstrated the agent's effectiveness, and raised interesting considerations for further development.

Computer network is increasingly being vital for everyday life. We believe that the combination of our communication environment and communication helper is useful for supporting daily activities.

7.2 Future Directions

- *Comparison of different modalities in virtual meeting spaces*

  Some important issues remain in the design of virtual meeting spaces. One of them is whether we should use video as a communication channel. As we summarized related research in Chapter 2, the advantage of video channel is not very strong. The weak merit of video seems to be easily overwhelmed by such demerits as consuming large network bandwidth, requiring cameras and invading privacy. Most of three-dimensional graphical chat tools use the avatar representation instead of live video. As described in Chapter 2, it is reported that people behave spatially even in a text-based chat tool, if they can control their avatars in a virtual space.

Another designing choice is using Text-To-Speech (TTS) instead of users' voices. Voice-mediated communication is often uncomfortable when you share the same real room with others, because speaking to the computer display is not popular as speaking to a person over the telephone. Moreover, you have to put on something like a microphone or a headset to record your voice. For FreeWalk interaction, it is a crucial mechanism that the volumes of voices attenuate according to the distance. Text-based channels do not match that mechanism, but it is not a problem to use TTS-voice. One possible way is every participant uses TTS, another is only participants who do not like to speak use TTS. Conversations between a real voice user and a TTS user may be interesting situation.

- *Statistical analysis of the effects of three-dimensional virtual meeting space*

  Our experiment to compare FreeWalk with a normal videoconferencing system and FTF environment is not statistical. It is desirable to conduct a statistical experiment to find the effect of three-dimensional virtual meeting spaces more clearly. However, it seems to be very hard to test the dynamic behavior to form conversation subgroups in experiments like that. To do that, an enough large number of subjects should join a meeting so that they can form subgroups. Since subjects should accustom themselves to control their embodiments until they learn the social meanings of spatial positions, a long-term experiment may be better. It is reasonable to test only the turning behavior, and the directional voice and video, because changing the view direction is much easier to learn than moving freely.

- *Comparison of the different representations of the social agent*

  We pasted an animated dog on the pyramid, that is FreeWalk embodiment of the social agent, to suppose it is regarded as the subservient and friendly. However, we do not know what happens

  if we use an animated bird instead. The still picture of a dog may differ from an animated dog. We can paste anything such as a photographic human picture, the recorded movie of a real person, and so on. It may be strange but beneficial to paste the face of meeting participants, with whom the agent interacts [Nass98]. The FreeWalk's embodiment that is shaped like a pyramid, can be replaced with a three-dimensional animal or human avatar. These differences in the representation may affect users' responses to the social support of the agent.

- *Development of an intelligent topic-selection mechanism*

  Our prototype of social agent picks up a topic to deliver in a completely random fashion. The agent tracks participants' voices to only detect awkward pauses indicating lagged conversation. The limitation of the ability of this method is obvious. Voice recognition technology can be used to retrieve more information from participants' voices. The keywords included in participants' utterances give us hints as to what they are talking about. Based on those keywords, the agent can provide the talking pair appropriate topics, which do not break down their conversation. Pattern matching techniques may be useful to process the recognized sentences [Ball99].

  The current prototype can speak only sentences prepared beforehand. The order of presented sentences must be a path in the tree structure of topic data. This makes the conversation style of the agent very functional. The agent should generate its script dynamically based on recognized keywords in the participants' conversation and the behavior of participants [Loyall97].

- *Study how social agents are perceived*

  In our experiment, the agent affected the participants' impressions of their conversations and partners. The agent affected even the participants' behaviors. It is interesting to study how strong is the social presence of social agents. The agent may be just an alternative user interface, but may be a social entity, which is independent from any particular users.

It is known that social rules can be applied to the interaction between people and computer agents [Nass94]. However, it is unknown whether the agent can have social relationships with people. The difference between the human relationships and the relationships between the agent and people is another question.

- *Another application of virtual meeting spaces and social agents*

We developed the virtual meeting space and the social agent to support the inherent characteristics of casual meetings. However, the potential of these systems is not limited only within the support of casual interaction. For example, they can become online training environments where social interaction is vital. The social agent can pretend to be a leader or a teacher in virtual training spaces [Rickel97].

We are developing the new version of FreeWalk and the social agent as the three-dimensional interface of digital cities. The proposed concept of digital cities is a social information infrastructure for urban everyday life including shopping, business, transportation, education, welfare and so on [Ishida99, Ishida2000, Ishida2001]. Digital cities integrate both Web documents and real-time data retrieved from sensors placed around the city. Digital cities visualize those through two-dimensional and three-dimensional views of the city, and make these views usable in the social interaction among the citizens and the visitors.

Since FreeWalk is originally designed as a pure communication environment, we are extending FreeWalk to contain the three-dimensional model of the real city. The strong correspondence of the virtual city in FreeWalk space with the real city is the most important design policy. This correspondence enables users to simulate dynamic multiuser activities, which are difficult to be

performed in the real city. We plan to develop an online environment for a fire and an earthquake drill, in which the citizens of Kyoto city can participate from their home through Internet. The correspondence is also useful for the seamless interaction between mobile users walking in the real city and desktop users navigating in the virtual city. The mobile interface of FreeWalk is a challenging issue because the screen of a mobile device is usually small. The techniques used in the meeting arrangement service provided in the mobile community support experiment [Nishibe98] may be useful.

Bibliography

[Ball99] G. Ball, "Mixing Scripted Interaction with Task-oriented Language Processing in a Conversational Interface," *International Conference on Intelligent User Interfaces (IUI-99)*, pp. 101–103, 1999.

[Benford96] S. Benford, C. Brown, G. Reynard and C. Greenhalgh, "Shared Spaces: Transportation, Artificiality, and Spatiality," *International Conference of Computer Supported Cooperative Work (CSCW-96)*, pp. 77–86, 1996.

[Bly93] S. A. Bly, S. R. Harrison and S. Irwin, "Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment," *Communications of the ACM*, Vol. 36, No. 1, pp. 28–47, 1993.

[Borning91] A. Borning and M. Travers, "Two Approaches to Casual Interaction over Computer and Video Networks," *International Conference on Human Factors in Computing Systems (CHI-91)*, pp. 13–19, 1991.

[Bowers96] J. Bowers, J. Pycock and J. O'Brien, "Talk and Embodiment in Collaborative Virtual Environments," *International Conference on Human Factors in Computing Systems (CHI-96)*, pp. 58–65, 1996.

[Brothers92] L. Brothers, J. Hollan, J. Nielsen, S. Stornetta, S. Abney, G. Furnas and M. Littman, "Supporting Informal Communication via Ephemeral Interest Groups," *International Conference* *of Computer Supported Cooperative Work (CSCW-92)*, pp. 84–90, 1992.

[Cassell99] J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjálmsson and H. Yan, "Embodiment in Conversational Interfaces: Rea," *International Conference on Human Factors in Computing Systems (CHI-99)*, pp.520–527, 1999.

[Clark96] H. H. Clark, *Using Language*, Cambridge University Press, 1996.

[Cranch71] M. V. Cranch, "The Role of Orienting Behavior in Human Interaction," *Behavior and Environment, International Symposium held at the 1968 meeting of the American Association for the Advancement of Science*, pp. 217–237, 1971.

[Dourish92] P. Dourish and S. Bly, "Portholes: Supporting Awareness in a Distributed Work Group," *International Conference on Human Factors in Computing Systems (CHI-92)*, pp. 541–547, 1992.

[Duck98] S. Duck, *Human Relationships, Third Edition*, SAGE Publications, London, 1998.

[Erickson99] T. Erickson, D. N. Smith, W. A. Kellogg, M. Laff, J. T. Richards and E. Bradner, "Socially Translucent Systems: Social Proxies, Persistent Conversation, and the Design of "Babble"," *International Conference on Human Factors in Computing Systems (CHI-99)*, pp. 72–79, 1999.

[Fish93] R. S. Fish, R. E. Kraut, R. W. Root and R. E. Rice, "Video as a Technology for Informal Communication," *Communications of the ACM*, Vol. 36, No. 1, pp. 48–61, 1993.

[Foner97] L. Foner, "Entertaining Agents: A Sociological Case Study," *International Conference on Autonomous Agents (AGENTS-97)*, pp. 122–129, 1997.

[Gaver92] W. Gaver, T. Moran, A. MacLean, L. Lövstrand, P. Dourish, K. Carter and W. Buxton, "Realizing a Video Environment: Europarc's RAVE System," *International Conference on Human Factors in Computing Systems (CHI-92)*, pp. 27–35, 1992.

[Grayson98] D. Grayson and L. Coventry, "The Effects of Visual Proxemic Information in Video Mediated Communication," *SIGCHI Bulletin*, Vol. 30, No. 3, pp. 30–39, 1998.

[Greenhalgh95] C. Greenhalgh and S. Benford, "Massive: A Collaborative Virtual Environment for Teleconferencing," *ACM Transactions on Computer-Human Interaction*, Vol. 2, No. 3, pp. 239–261, 1995.

[Hagsand96] O. Hagsand, "Interactive Multiuser VEs in the DIVE System," *IEEE MultiMedia*, Vol. 3, No. 1, pp. 30–39, 1996.

[Hall66] E. T. Hall, *The Hidden Dimension*, Doubleday, 1966.

[Hall90] E. T. Hall and M. R. Hall, *Hidden Differences: Doing Business with the Japanese, Reprint edition*, Anchor Books, 1990.

[Han96] J. Han and B. Smith, "CU-SeeMe VR Immersive Desktop Teleconferencing," *International Conference on Multimedia*, pp. 199–207, 1996.

[Harrison96] S. Harrison and P. Dourish, "Re-Place-ing Space: The Roles of Place and Space in Collaborative Systems," *International Conference of Computer Supported Cooperative Work (CSCW-96)*, pp. 67–76, 1996.

[Heath92] C. Heath and P. Luff, "Media Space and Communicative Symmetries: Preliminary Observations of Video-Mediated Interaction," *Human-Computer Interaction*, Vol. 7, Issue. 3, pp. 315–346, 1992.

[Honda97] S. Honda, H. Tomioka, T. Kimura, T. Oosawa, K. Okada and Y. Matsushita, "Valentine: An Environment for Home Office Worker Providing Informal Communication and Personal Space," *International Conference on Supporting Group Work (GROUP-97)*, pp. 368–375, 1997.

[Isaacs94] E. A. Isaacs, T. Morris and T. K. Rodriguez, "A Forum for Supporting Interactive Presentations to Distributed Audiences," *International Conference of Computer Supported Cooperative Work (CSCW-94)*, pp. 405–416, 1994.

[Isbister97] K. Isbister and B. Hayes-Roth, "Social Implications of Using Synthetic Characters," *Workshop on Animated Interface Agents in IJCAI-97*, pp. 19–20, 1997.

[Isbister99] K. Isbister and P. Doyle, "Touring Machines: Guide Agents for Sharing Stories about Digital Places," *Workshop on Narrative and Artificial Intelligence in AAAI-99*, 1999.

[Isbister2000a] K. Isbister, H. Nakanishi, T. Ishida and C. Nass, "Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space," *International Conference on Human Factors in Computing Systems (CHI-2000)*, pp. 57–64, 2000.

[Isbister2000b] K. Isbister and H. Nakanishi, "Helper Agent: A Chat Assistant for Cross-cultural Conversations," *NTT REVIEW*, Vol. 12, No. 2, pp. 55–59, 2000.

[Ishida94] T. Ishida, "Bridging Humans via Agent Networks," *International Workshop on Distributed Artificial Intelligence*, pp. 419–429, 1994.

[Ishida95] T. Ishida, "Discussion on Agents," *Journal of the Japanese Society for Artificial Intelligence*, Vol. 10, No. 5, pp. 663–667, 1995 (in Japanese).

[Ishida96] T. Ishida, H. Yamaki, H. Nakanishi and T. Nishimura, "Casual Meetings in a Network," *International Symposium on Cooperative Database Systems for Advanced Applications*, 1996.

[Ishida97] T. Ishida, "Towards Communityware," *New Generation Computing*, Vol. 16, No. 1, pp. 5–21, 1998 (Invited Paper) (also appeared in *International Conference and Exhibition on the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM-97)*, pp. 7–21, 1997 (Invited Talk)).

[Ishida98a] T. Ishida Ed., *Community Computing: Collaboration over Global Information Networks*, John Wiley and Sons, 1998.

[Ishida98b] T. Ishida Ed., *Community Computing and Support Systems*, Lecture Notes in Computer Science 1519, Springer-Verlag, 1998.

[Ishida99] T. Ishida, J. Akahani, K. Hiramatsu, K. Isbister, S. Lisowski, H. Nakanishi, M. Okamoto, Y. Miyazaki and K. Tsutsuguchi, "Digital City Kyoto: Towards A Social Information Infrastructure," M. Klusch, O. Shehory, G. Weiss Eds., *Cooperative Information Agents III, Lecture Notes in Artificial Intelligence*, Vol. 1652, pp. 23-35, Springer-Verlag, 1999.

[Ishida2000] T. Ishida and K. Isbister Eds., *Digital Cities: Experiences, Technologies and Future Perspectives*, Lecture Notes in Computer Science, 1765, Springer-Verlag, 2000.

[Ishida2001] T. Ishida, "Digital City Kyoto: Social Information Infrastructure for Everyday Life," *Communications of the ACM (CACM)*, 2001 (to appear).

[Kauppien98] K. Kauppinen, A. Kivimäki, T. Era and M. Robinson, "Producing Identity in Collaborative Virtual Environments," *ACM Symposium on Virtual Reality Software and Technology (VRST-98)*, pp. 35–42. 1998.

[Kiesler96] S. Kiesler, L. Sproull and K. Waters, "A Prisoner's Dilemma Experiment on Cooperation With People and Human-Like Computers," *Journal of Personality and Social Psychology*, Vol. 70, No. 1, pp. 47–65, 1996.

[Kitamura96] Y. Kitamura, H. Nakanishi, T. Nozaki, T. Miura and T. Ishida, "MetaViewer and MetaCommander: Applying WWW Tools to Genome Informatics," *Workshop on Genome Informatics*, pp. 137–146, 1996.

[Lea97] R. Lea, Y. Honda, K. Matsuda and S. Matsuda, "Community Place: Architecture and Performance," *Symposium on Virtual Reality Modeling Language (VRML-97)*, pp. 41–50, 1997.

[Lester97] J. C. Lester, S. A. Converse, S. E. Kahler, S. T. Barlow, B. A. Stone and R. S. Bhogal, "The Persona Effect: Affective Impact of Animated Pedagogical Agents," *International Conference on Human Factors in Computing Systems (CHI-97)*, pp. 359–366, 1997.

[Loyall97] A. B. Loyall and J. Bates, "Personality-rich Believable Agents That Use Language," *International Conference on Autonomous Agents (AGENTS-97)*, pp. 106–113, 1997.

[Maes95] P. Maes, "Artificial Life Meets Entertainment: Lifelike Autonomous Agents," *Communications of the ACM*, Vol. 38, No. 11, pp. 108–114, 1995.

[Mantei91] M. M. Mantei, R. M. Baecher, A. J. Sellen, W. A. S. Buxton, T. Milligan and B. Wellman, "Experiences in the Use of a Media Space," *International Conference on Human Factors in Computing Systems (CHI-91)*, pp. 203–208, 1991.

[Nakanishi96] H. Nakanishi, C. Yoshida, T. Nishimura and T. Ishida, "FreeWalk: Supporting Casual Meetings in a Network," *International Conference on Computer Supported Cooperative Work (CSCW-96)*, pp. 308–314, 1996.

[Nakanishi98a] H. Nakanishi, C. Yoshida, T. Nishimura and T. Ishida, "FreeWalk: Supporting Informal Communication in a Three Dimensional Virtual Space," *Transactions of IPSJ (The Information Processing Society of Japan)*, Vol. 39, No. 5, pp. 1356–1364, 1998 (in Japanese).

[Nakanishi98b] H. Nakanishi, T. Nishimura and T. Ishida, "Effects of a Three-dimensional Virtual Space in Desktop Conferences," *Transactions of IPSJ (The Information Processing Society of Japan)*, Vol. 39, No. 10, pp. 2770–2777, 1998 (in Japanese).

[Nakanishi98c] H. Nakanishi, C. Yoshida, T. Nishimura and T. Ishida, "FreeWalk: A Three-Dimensional Meeting-Place for Communities," T. Ishida Ed., *Community Computing: Collaboration over Global Information Networks*, John Wiley and Sons, pp. 55–89, 1998.

[Nakanishi99] H. Nakanishi, C. Yoshida, T. Nishimura and T. Ishida, "FreeWalk: A 3D Virtual Space for Casual Meetings," *IEEE MultiMedia*, Vol. 6, No. 2, pp. 20–28, 1999.

[Nakanishi2000] H. Nakanishi, C. Yoshida, T. Nishimura and T. Ishida, "FreeWalk: Shared Virtual Space for Casual Meetings," B. Furht Ed., *Handbook of Internet Computing*, CRC Press, pp. 227–247, 2000.

[Nass94] C. Nass and S. Steuer and E. R. Tauber, "Computers are Social Actors," *International Conference on Human Factors in Computing Systems (CHI-94)*, pp. 72–78, 1994.

[Nass98] C. Nass, E. Kim and E. Lee, "When My Face is the Interface: An Experimental Comparison of Interacting with One's Own Face or Someone Else's Face," *International Conference on Human Factors in Computing Systems (CHI-98)*, pp. 148–154, 1998.

[Nishibe98] Y. Nishibe, H. Waki, I. Morihara, F. Hattori, T. Ishida, T. Nishimura, H. Yamaki, T. Komura, N. Itoh, T. Gotoh, T. Nishida, H. Takeda, A. Sawada, H. Maeda, M. Kajihara, H. Adachi, "Mobile Digital Assistants for Community Support," *AI Magazine*, Vol. 19, No. 2, pp. 31-49, 1998.

[Nishimoto99] Kazushi Nishimoto, Kenji Mase and Ryohei Nakatsu, "How an Autonomous Information Retrieval Agent Affects Divergent Thinking by a Group," *Journal of the Japanese Society for Artificial Intelligence*, Vol. 14, No. 1, pp. 58–70, 1999 (in Japanese).

[Nishimura98] T. Nishimura, H. Nakanishi, C. Yoshida and T. Ishida, "Applying Videogame Technologies to Video Conferencing Systems," *ACM Symposium on Applied Computing (SAC-98) Multimedia Systems Track*, pp. 471–476, 1998.

[Obata98] A. Obata and K. Sasaki, "OfficeWalker: A Virtual Visiting System Based on Proxemics," *International Conference on Computer Supported Cooperative Work (CSCW-98)*, pp. 1–10, 1998.

[O'Conaill93] B. O'Conaill, S. Whittaker and S. Wilbur, "Conversations Over Video Conferences: An Evaluation of the Spoken Aspects of Video-Mediated Communication," *Human-Computer Interaction*, Vol. 8, Issue. 4, pp. 389–428, 1993.

[Okada94] K. Okada, F. Maeda, Y. Ichikawa and Y. Matsushita, "Multiparty Videoconferencing at Virtual Social Distance: MAJIC Design," *International Conference of Computer Supported Cooperative Work (CSCW-94)*, pp. 385–393, 1994.

[Okamoto98] M. Okamoto, H. Nakanishi, T. Nishimura and T. Ishida, "Silhouettell: Awareness Support for Real-World Encounter," T. Ishida Ed., *Community Computing and Support Systems*, Lecture Notes in Computer Science 1519, Springer-Verlag, pp. 316–329, 1998.

[Okamoto99] M. Okamoto, K. Isbister, H. Nakanishi and T. Ishida, "Supporting Cross-Cultural Communication in Real-World Encounters," *International Conference on Human-Computer Interaction (HCI-99)*, Vol. 2, pp. 442–446, 1999.

[Parise96] S. Parise, S. Kiesler, L. Sproull and K. Waters, "My Partner is a Real Dog: Cooperation with Social Agents," *International Conference of Computer Supported Cooperative Work (CSCW-96)*, pp. 399–408, 1996.

[Reeves96] B. Reeves and C. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, Cambridge University Press, 1996.

[Rickel97] J. Rickel and W. L. Johnson, "Integrating Pedagogical Capabilities in a Virtual Environment Agent," *International Conference on Autonomous Agents (AGENTS-97)*, pp. 30–38, 1997.

[Root88] R. W. Root, "Design of a Multimedia Vehicle for Social Browsing," *International Conference on Computer Supported Cooperative Work (CSCW-88)*, pp. 25–38, 1988.

[Sellen95] A. J. Sellen, "Remote Conversations: The Effects of Mediating Talk With Technology," *Human-Computer Interaction*, Vol. 10, Issue. 4, pp. 401–444, 1995.

[Smith2000] M. A. Smith, S. D. Farnham and S. M. Drucker, "The Social Life of Small Graphical Chat Spaces," *International Conference on Human Factors in Computing Systems (CHI-2000)*, pp. 462–469, 2000.

[Sproull96] L. Sproull, M. Subramani, S. Kiesler, J. H. Walker and K. Waters, "When the Interface Is a Face," *Human-Computer Interaction*, Vol. 11, Issue. 2, pp. 97–124, 1996.

[Sugawara94] S. Sugawara, G. Suzuki, Y. Nagashima, M. Matsuura, H. Tanigawa and M. Moriuchi, "InterSpace: Networked Virtual World for Visual Communication," *IEICE (Institute of Electronics, Information, and Communication Engineers) Transactions on Information and Systems*, Vol. E77-D, No. 12, pp. 1344–1349, 1994.

[Takeuchi95] A. Takeuchi and T. Naito, "Situated Facial Displays: Towards Social Interaction," *International Conference on Human Factors in Computing Systems (CHI-95)*, pp. 450–455, 1995.

[Tang94] J. C. Tang and M. Rua, "Montage: Providing Teleproximity for Distributed Groups," *International Conference on Human Factors in Computing Systems (CHI-94)*, pp. 37–43, 1994.

[Veinott99] E. S. Veinott, J. Olson, G. M. Olson and X. Fu, "Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other," *International Conference on Human Factors in Computing Systems (CHI-99)*, pp. 302-309, 1999.

[Watabe90] K. Watabe, S. Sakata, K. Maeno, H. Fukuoka and T. Ohmori, "Distributed Multiparty Desktop Conference System Mermaid: Platform for Groupware," *International Conference on Computer Supported Cooperative Work (CSCW-90)*, pp. 27–38, 1990.

[Waters97] R. C. Waters and J. W. Barrus, "The Rise of Shared Virtual Environments," *IEEE Spectrum*, Vol. 34, No. 3, pp. 20–25, 1997.

[Whittacker94] S. Whittacker, D. Frohlich and O. Daly-Jones, "Informal Workplace Communication: What Is It Like and How Might We Support It?," *International Conference on Human Factors in Computing Systems (CHI-94)*, pp. 131–137, 1994.

[Yamaki96a] H. Yamaki, M. Kajihara, G. Tanaka, T. Nishimura, H. Ishiguro and T. Ishida, "Socia: Non-Committed Meeting Scheduling with Desktop Vision Agents," *International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM-96)*, pp. 727–742, 1996.

[Yamaki96b] H. Yamaki, M. P. Wellman and T. Ishida, "A Market-Based Approach to Allocating QoS for Multimedia Applications," *International Conference on Multi-Agent Systems (ICMAS-96)*, pp. 385–392, 1996.

[Yamaki98a] H. Yamaki, M. P. Wellman and T. Ishida, "Market-Based QoS Control for Incorporating Community Preferences," T. Ishida Ed., *Community Computing: Collaboration over Global Information Networks*, John Wiley and Sons, pp. 91–126, 1998.

[Yamaki98b] H. Yamaki, Y. Yamauchi and T. Ishida "Implementation Issues on Market-Based QoS Control," *International Conference on Multi-Agent Systems (ICMAS-98)*, pp. 357–364, 1998.

Publications

Major Publications

Journals

1. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: A 3D Virtual Space for Casual Meetings," *IEEE MultiMedia*, Vol. 6, No. 2, pp. 20–28, 1999.

2. Hideyuki Nakanishi, Toshikazu Nishimura and Toru Ishida, "Effects of a Three-dimensional Virtual Space in Desktop Conferences," *Transactions of IPSJ (The Information Processing Society of Japan)*, Vol. 39, No. 10, pp. 2770–2777, 1998 (in Japanese).

3. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: Supporting Informal Communication in a Three Dimensional Virtual Space," *Transactions of IPSJ (The Information Processing Society of Japan)*, Vol. 39, No. 5, pp. 1356–1364, 1998 (in Japanese).

Conferences

1. Katherine Isbister, Hideyuki Nakanishi, Toru Ishida and Cliff Nass, "Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space," *International Conference on Human Factors in Computing Systems (CHI-2000)*, pp.

57–64, 2000.

2. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: Supporting Casual Meetings in a Network," *International Conference on Computer Supported Cooperative Work (CSCW-96)*, pp. 308–314, 1996.

Chapters in Books

1. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: Shared Virtual Space for Casual Meetings," Borko Furht Ed., *Handbook of Internet Computing*, CRC Press, pp. 227–247, 2000.

2. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: A Three-Dimensional Meeting-Place for Communities," Toru Ishida Ed., *Community Computing: Collaboration over Global Information Networks*, John Wiley and Sons, pp. 55–89, 1998.

Other Publications

Awarded Papers

1. Hideyuki Nakanishi, Katherine Isbister, Toru Ishida and Clifford Nass, "Designing a Helper Agent for Communication in a Virtual Space," *IPSJ (The Information Processing Society of Japan) Interaction2000 Workshop*, pp. 107–114, 2000 (in Japanese). Outstanding Paper Award.

Workshops

1. Satoshi Nakazawa, Hideyuki Nakanishi and Toru Ishida, "A Virtual Space Agent that Facilitates Conversation," *IPSJ (The Information Processing Society of Japan) DiCoMo Symposium*, pp. 19–24, 2000 (in Japanese).

2. Toru Ishida, Jun-ichi Akahani, Kaoru Hiramatsu, Katherine Isbister, Stefan Lisowski, Hideyuki Nakanishi, Masayuki Okamoto, Yasuhiko Miyazaki and Ken Tsutsuguchi, "Digital City Kyoto: Towards A Social Information Infrastructure," M. Klusch, O. Shehory, G. Weiss Eds., *Cooperative Information Agents III, Lecture Notes in Artificial Intelligence*, Vol. 1652, pp. 23-35, Springer-Verlag, 1999.

3. Toshikazu Nishimura, Hideyuki Nakanishi, Chikara Yoshida and Toru Ishida, "Applying Videogame Technologies to Video Conferencing Systems," *ACM Symposium on Applied Computing (SAC-98) Multimedia Systems Track*, pp. 471–476, 1998.

4. Hideyuki Nakanishi, Toshikazu Nishimura and Toru Ishida, "Interaction Analysis in a Three-Dimensional Meeting Space," *International Workshop on Agents in CommunityWare (ACW-98)*, 1998.

5. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "Using the Casual Meeting System on the Internet," *IPSJ (The Information Processing Society of Japan) DiCoMo Workshop*, pp. 305–310, 1997 (in Japanese).

6. Toru Ishida, Hirofumi Yamaki, Hideyuki Nakanishi and Toshikazu Nishimura, "Casual Meetings in a Network," *International Symposium on Cooperative Database Systems for Advanced Applications*, 1996.

7. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: Supporting Casual Meetings in a Network," *Visual Proceedings of SIGGRAPH96*, pp. 120, 1996.

8. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: A Promenade in Information Networks," *IEICE (The Institute of Electronics, Information and Communication Engineers) Data Engineering Workshop (DEWS96)*, pp. 31–36, 1996 (in Japanese).

Articles

1. Katherine Isbister and Hideyuki Nakanishi, "Helper Agent: an Assistant for Human-Human Interaction in a Virtual Meeting Space," *NTT R & D*, Vol. 49, No. 2, pp. 96–101, 2000 (in Japanese).

2. Katherine Isbister and Hideyuki Nakanishi, "Helper Agent: A Chat Assistant for Cross-cultural Conversations," *NTT REVIEW*, Vol. 12, No. 2, pp. 55–59, 2000.

Conventions

1. Hideyuki Nakanishi, Toshikazu Nishimura and Toru Ishida, "Communication Analysis in a Three-Dimensional Meeting Space," *IPSJ SIG Notes*, 98-GW-28-1, pp. 1–6, 1998 (in Japanese).

2. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: Casual Meetings in a Network," *Technical Report of IEICE*, IN97-33, 1997 (in Japanese). Invited Talk.

3. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "Experiences and Functional Extensions on the Casual Meeting System FreeWalk," *The 54th National Convention IPSJ*, 7R-6, 1997 (in Japanese).

4. Chikara Yoshida, Hideyuki Nakanishi, Toshikazu Nishimura and Toru Ishida, "Applying Videogame Technologies to Desktop Conferencing Systems," *IPSJ SIG Notes*, 97-HI-70-4, pp. 25–32, 1997 (in Japanese).

5. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "Using 3-D Virtual Space to Support Encounters," *Annual Conference of JSAI*, pp. 633–636, 1996 (in Japanese).

6. Hideyuki Nakanishi, Chikara Yoshida, Toshikazu Nishimura and Toru Ishida, "FreeWalk: An Open Network Meeting System," *The 52th National Convention IPSJ*, 3X-4, 1996 (in Japanese).