Designing Emergency Guidance in a Social Interaction Platform

Hideyuki Nakanishi and Toru Ishida

Department of Social Informatics, Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan
`{nakanishi, ishida}@i.kyoto-u.ac.jp`
`http://www.lab7.kuis.kyoto-u.ac.jp/`

Abstract. Future computing systems interact with a large number of users moving around buildings and streets. In this paper, we propose an example of such systems and how to evaluate ubicomp systems equipped with a large-scale physical environment that includes a large number of people inside. In our emergency guidance system, off-site guiding staff monitors a crowded large-scale public space to understand its situation, and instruct on-site guiding staff how to guide crowds effectively. Our system tracks and synthesizes the public space to enable the off-site staff to grasp it, and support communication between the on-site and the off-site staffs. Because it is not affordable to use the physical public space and a lot of human subjects to evaluate the system, we used our social interaction platform to simulate our guidance system. We could successfully construct simulations, in which the crowds are replaced with social agents in the virtual public space.

1 Introduction

Development of ubiquitous computing (ubicomp) systems compels us to explore their design space in physical environments. Physical part of ubicomp systems is usually much larger than that of desktop computing systems. In terms of ubicomp systems, 1) offices, home, and classrooms have been a major development field [1]; 2) navigation in shopping malls is one of promising applications [4]; and 3) streets can be a playfield of mixed reality games [7]. In this regard, designing ubicomp systems is very different from designing desktop computing systems, most of which can be accomplished in digital environments.

This nature of ubicomp systems interferes seriously with their design process since it is expensive to test them in an actual physical environment. We need many people and an experimental field to evaluate a system. Consequently, development of ubicomp systems tends to give birth to a compromise in order to make it feasible. For instance, 1) context information was simulated to compensate for lack of sensor devices [16]; 2) a university building was used to evaluate a navigation system for shopping malls [4]; and 3) entertainment was selected as an application domain to take advantage of errors produced by unreliable communication and positioning systems [7].

T. Ishida, L. Gasser, and H. Nakashima (Eds.): MMAS 2004, LNAI 3446, pp. 283–294, 2005.

Designing an ubicomp system becomes extremely difficult if it is supposed to be used in indoor large-scale environments such as tall buildings, central railway stations, and airports. GPS is unavailable in indoor situations and wide-area positioning sensors are not very affordable. Because indoor large-scale environments may contain a large amount of people inside, a lot of human subjects are needed to analyze the effectiveness of the system especially when interaction between the system and each person generates significant side effects on others' behavior. An example of such massively multiagent ubicomp systems is our emergency guidance system explained in next section.

To support the design process of massively multiagent ubicomp systems, we have created methods to simulate a situation where those systems interact with many people. In these methods, human observers are substituted for sensing functionality of a system, physical environments are changed into virtual environments, and most subjects are replaced with social agents, which are software agents equipped with social interaction capability [17]. Our social interaction platform introduced in the third section enables these replacements. We discuss the replacing methods in the fourth section.

2 Emergency Guidance System

In large-scale public spaces like central railway stations, appropriate guidance for crowd control is critical because a vast amount of people visits. There are two ways to guide visitors. Staff in a control room provides overall guidance for all visitors through announcement speakers while on-site staff working in a public space gives location-based guidance for each visitor. At the present time these two guidance do not seem to be tied together very much, and collaboration between the off-site and the on-site staffs has not been extensively considered.

We propose an efficient means of crowd control, in which the off-site and the on-site staffs cooperate with each other. In this guidance illustrated in Figure 1, the off-site staff monitors an overall public space to find an unsafe group of crowd, and then instruct the nearby on-site staff how to guide the crowd effectively. According to this instruction the on-site staff tries to modify the crowd's behavior. Since current facilities, which are surveillance cameras and announcement speakers, cannot support such operations very well, technologically advanced systems are necessary. Our emergency guidance system is an asymmetric communication environment which connects people who can remotely perceive an entire situation of a large-scale physical space and people who exist in the space [18]. This system is an example of massively multiagent ubicomp systems.

Figure 2 depicts the first prototype of the guidance system installed in Kyoto Station, which is a central railway station where more than 300,000 passengers visit per day. The movements of the people on a station's platform are tracked by using a vision sensor network. We attached 12 sensors to the concourse area and 16 sensors to the platform as shown in Figure 3's floor plan, on which the black dots show the sensors' positions. The movement data are transmitted to the control room where the off-site staff is viewing the synthesized platform to find a location where their remote help is needed. They can see a bird's-eye view of the platform, which cannot be taken directly through cameras attached to the ceiling of the platform. When they notice a dangerous spot, they point at human figures which correspond to the on-site staff around the spot so as to establish vocal communication channels between their headsets and the on-site staff's mobile phones. This trick is possible because their phone numbers are registered at the system beforehand.

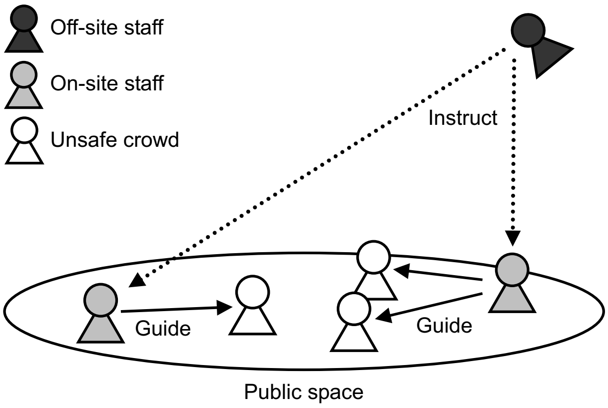


Fig. 1. Collaboration in Emergency Guidance

It is hard to guess what happens when the guidance system is used in actual emergency situation, because many people behave autonomously and complex interaction occurs. We need to carry out a lot of experiments in the station, but that is faced with several problems. Even though the installation of the sensor network is a result of our considerable efforts, its covered area is merely a very small part of the station, and its accuracy is not sufficient for earnest evaluation of the system's performance. Furthermore, it is enormously difficult to conduct an experiment with many subjects to evaluate the systems' effectiveness in such a public space. Simulating techniques which can solve these problems are very helpful in designing this kind of systems.

In the case of massively multiagent ubicomp systems, human-computer interaction is only a part of all interactions. Human-computer interaction between a system and users can cause face-to-face interaction between users and other people who are not using the system. And moreover, this interaction can affect everyone's behavior in the same physical space. The emergency guidance system deals with the following three interactions: 1) interaction between the system (namely the off-site staff) and on-site staff, 2) interaction between the on-site staff and the crowd, and 3) interaction among the crowd. The social interaction platform described in the next section makes it possible to construct and execute a simulation of these interactions.

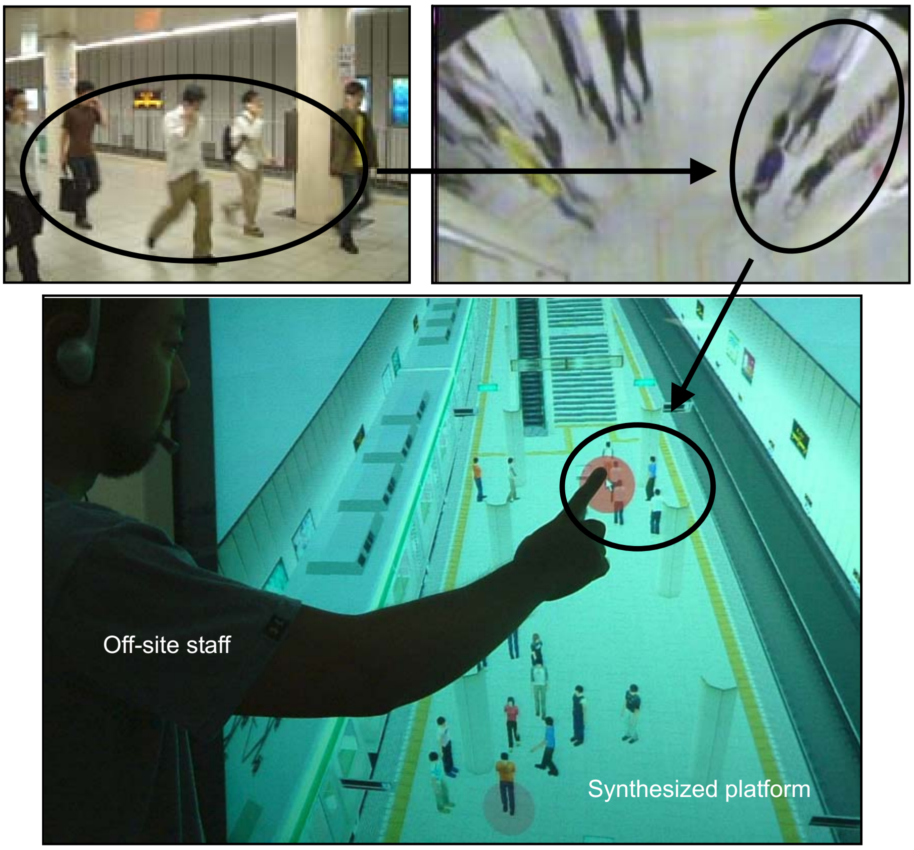


Fig. 2. Emergency Guidance System

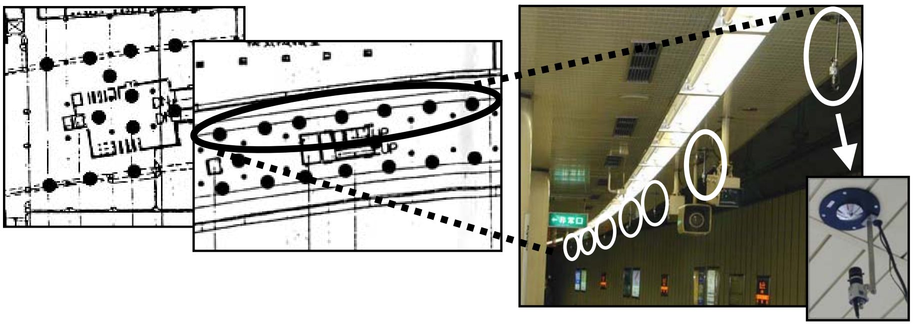


Fig. 3. Vision Sensors in Kyoto Station

3 Social Interaction Platform

There are various technologies which can be utilized for simulating social interaction in a physical space, e.g. embodied conversational agents [5], collaborative virtual environments [3], virtual cities [14], and multiagent-based simulations [8]. Our social interaction platform called "FreeWalk/*Q*" is an integration of these technologies [19]. As the result of that integration, people's avatars and social agents can socially interact with one another in a shared virtual space. The platform is a combination of the two components: "FreeWalk" and "*Q*." *Q* is a scenario description language, in which we can describe interaction scenarios as explained below [11]. FreeWalk is an interaction platform which mingles agents and avatars based on the interaction model detailed later.

An interaction scenario is a definition of a social role. It is a collection of scenes, each of which includes a set of interaction rules that define the agent's social behavior in the situation. Each rule is a couple made up of a conditional *cue* and the consequent series of *actions*. A short example is presented below. Cues start with a question mark and actions start with an exclamation point. In `scene1`, the agent says "Hello" and its state switches to `scene2` when it hears someone say "Hello." But if the agent observes someone wave his/her hand in the same scene, its state switches to `scene3`. The same cue yields different actions in different scenes. In `scene2`, the agent approaches the person and responds with "Yes, may I help you?" when it hears "Hello" again.

```
(defscenario reception
  (scene1
   ((?hear "Hello" :from $x)
    (!speak "Hello" :to $x)
    (go scene2))
   ((?observe :gesture "wave")
    (go scene3)))
  (scene2
   ((?hear "Hello" :from $x)
    (!walk :to $x)
    (!speak "Yes, may I help you?" :to $x)))
  (scene3 ...
```

In the interaction model of FreeWalk, each action is mapped to a modification of the walking, gestural, and speech parameters of the agent. Other agents' cues can inspect these action parameters if such perception parameters as visual and hearing powers of those agents allow them to perceive the action parameters. In this manner, the virtual space of FreeWalk transmits verbal cues and such nonverbal cues as interpersonal distance [9], gaze direction [13], and pauses in a conversation [6]. These cues influence each agent's next action and finally invoke group behaviors. For example, an agent's walking direction may change and this change may form a flow of other following agents [23] when the agent observes a pointing gesture produced by the gestural parameters of another agent.

In the platform FreeWalk/*Q*, people's avatars and social agents can share the same virtual space, the same interaction model, and the same interaction scenario. Agents are controlled through the application program interface (API). API functions are called to evaluate cues and actions described in the current scene of the assigned scenario. On the other hand, avatars are controlled through the user interface (UI). The action parameters are modified based on how the input devices are controlled. The perception parameters determine the distance of the far clipping plane in the view displayed on the screen, and the sound volume.

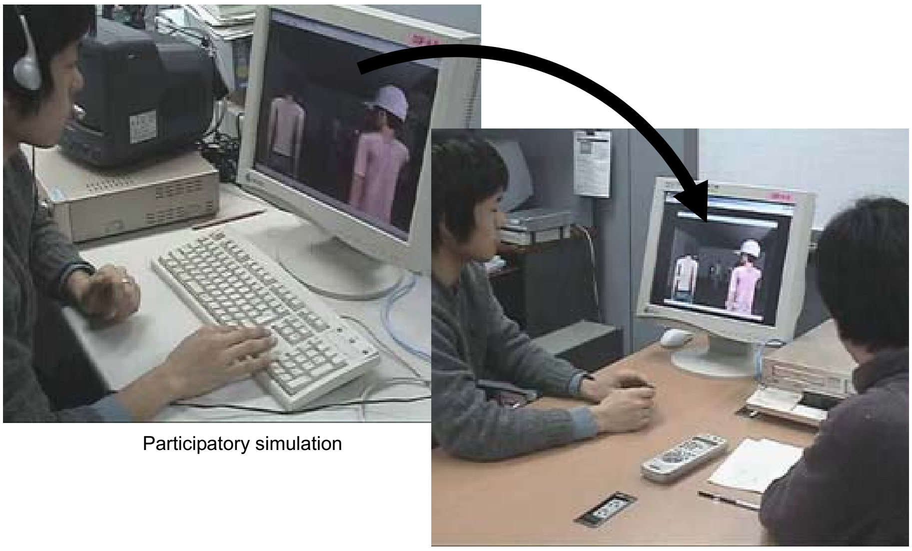
Based on the framework of “Society-centered Design” [12], we have created methods to simulate the emergency guidance system on the platform. The platform is useful for constructing a simulated collaborative guidance as well as executing it. When we construct the simulation, the platform becomes an environment for experiments from which we can retrieve interaction models and scenarios of social agents that act as the crowd. In the society-centered design framework, this construction phase corresponds to the “participatory simulations,” which are multiagent-based simulations that include avatars. When we execute the simulation, the platform not only produces the simulated crowd but also functions as a tool to track subjects playing a role of the on-site staff. In the framework, this execution stage corresponds to the “augmented experiments,” which are real-world experiments augmented by simulated users.

4 Simulating Methods

4.1 Describing Interaction Among Agents

To replace human subjects who play the role of an escaping crowd with extra agents, we need to simulate humanlike decision-making capability to form a group behavior. First, we describe an interaction scenario of the group behavior to construct its simulation as realistic as possible. Next, we conduct an experiment in which subjects take part in the simulation to experience the group behavior. Our social interaction platform can support this participatory simulation. After that experiment, we ask the subjects what decisions they have made throughout the virtual group behavior. In this interview, we show the subject a replay of his/her recorded first-person view which was displayed on the screen so that he/she can easily recall and answer what he/she was doing every moment. For example, we ask what they were paying attention to, what they were thinking about, and what they were trying to do per second. Finally, we can improve the scenario based on the interview result.

To simulate the emergency guidance system, we need a simulation of following behavior of the crowd. Previous studies gave us a basis for describing an interaction scenario of the following crowd [25, 15]. In the experiment, subjects experienced the described behavior and we interviewed them as shown in Figure 4. We successfully obtained deeper decision-making rules of the following crowd. For example, one of the prepared rules was “if you find a leading person, you follow his/her instruction.” But, we found this was imperfect. The obtained rule is “if you find a leading person and observe others are following him/her, you follow his/her instruction.”



Participatory simulation

Interview

Fig. 4. Interviews after Participatory Simulations

When we model a group behavior based on recorded materials of past real-world events [10], we can observe what happened actually in the event but have to make enormous efforts to interview people there. In a physical-space simulation [25], we can interview the subjects but it is difficult to record their first-person views without distracting them who are being engaged in the group behavior. In virtual-space simulations, it is very easy to record exactly the same views and replay it in the interviews.

4.2 Modeling Interaction with Humans

In terms of agent-human interaction, we have focused on designing humanlike response in non-conversational interaction because we suppose that conversational interaction would be a territory of human subjects. In the emergency guidance system, interaction between the off-site and the on-site staffs is conversational, but in the other interactions in which the crowd participates, non-conversational interaction is very influential as described below. Hence, subjects are in charge of the staff and social agents constitute the crowd in the simulations.

A kind of nonverbal communication is much related to non-conversational interaction. It has been reported that nonverbal communication has several different functions [21]. Some functions are related to conversational interaction. A typical one of them is to provide information. For instance, nods and frowns transmit the meanings of yes and no. Another conversational function is to coordinate interaction, e.g. beginning a conversation and taking a turn. There are also non-conversational functions, which strongly influence human communication and relationships. For example, people use nonverbal cues to bolster inter-

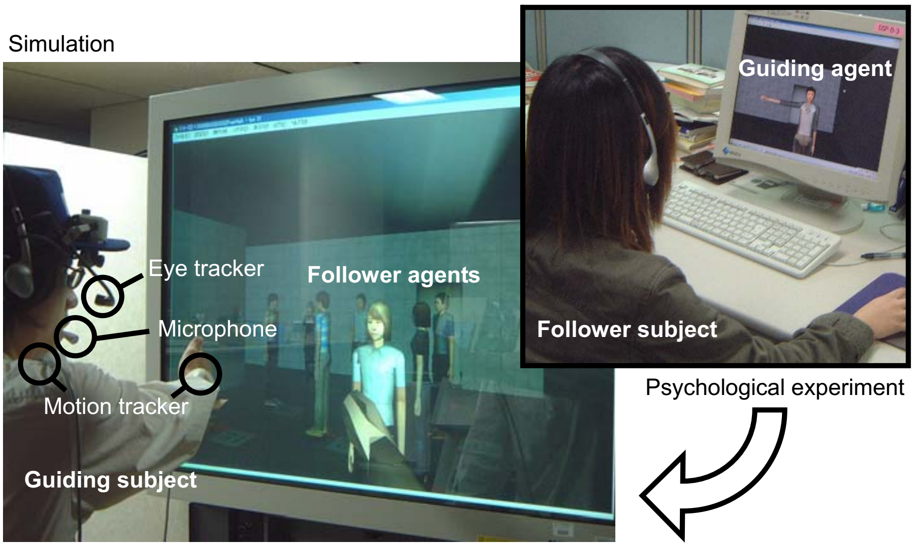


Fig. 5. Role-reversal Experiments

personal connections, such as building intimacy by increasing gaze and moving closer to the other person [26].

When the staff tries to guide the crowd in an emergency guidance, they usually shout something like, “the exit is over there!” The influencing power of this verbal cue shifts according to the nonverbal cues which accompany the words [28]. If the nonverbal cues do not indicate enough implicit influence that is trustworthiness, the crowd will not follow the staff’s verbal instructions.

We created a method to develop social agents that can perceive implicit influences as humans do [20]. In this method, human perception mechanism is analyzed in psychological experiments, and then the mechanism becomes embedded in agents. At first, qualitative model to interpret nonverbal cues that provide implicit influences is fabricated based on social scientific knowledge of human responses to other humans. Next, the model is quantified by using the results of the experiments in which subjects respond to agents that emit the varied nonverbal cues as stimuli. Finally, we construct agents that can respond to multi-modal input data which compose the nonverbal cues. In the experiments, subjects play the role of the agents that perceive implicit influences in a simulation and the agents play the role of humans who try to make the influences. Thus, the method is called role-reversal experiments, which is presented in Figure 5. Findings from past media studies [22] enable the transition from the first step to the second step. Statistical analysis enables the transition from the second step to the third step. The social interaction platform is useful for making agents in both the second and the third steps.

A self-confident person speaks loud and fast [24]. A self-confident person tends to use gesture more [27]. Extended gaze into another’s eyes is the strongest cue to provide the impression of self-confidence [21]. If the staff speaks loudly and quickly, points to the destination clearly, and makes steady eye contact, you will probably follow their instructions. We can develop this qualitative model into a quantitative model according to the result of an experiment to observe how four cues–gaze ratios, gesture sizes, voice volumes, and speech speeds–contribute to trustworthiness. Using the questionnaire data, we are able to find the threshold of each cue's positive effect and the relative influence of that cue. We can use ANOVA analysis to determine which level's effect is significantly stronger or weaker than another level's effect and how independent the cues are. If each cue is independent and their influences are linear, we can also use multiple linear regression analysis to make a formula that calculates the degree of trustworthiness based on each cue's input data. In Figure 5, you can see an example of the input interface that is a combination of motion tracking sensors (detecting gesture sizes), an eye tracking sensor (detecting gaze ratios), and a microphone (detecting voice volumes and speech speeds).

4.3 Tracking Without Sensors

The installation of sensors in an experimental physical environment is always problematic. Our idea is that someone (may be an experimenter) becomes the wizard who provides the system context information instead of the sensors that would be attached to the environment. Please suppose that the system is connected to a virtual space that is a copy of the physical environment. In the same or a remote place, the wizard observes what the subject is doing. By this observation, the wizard controls the avatar which corresponds to the subject in order to make it do the same thing in the virtual space. Lastly, the ubicomp system obtains the subject's context information from the virtual space.

In the simulation of the guidance system, the movements of the subjects who play a role of on-site staff are converted into the avatars' movements so that the guidance system can track the subjects without using a positioning sensor network. As a result of this tracking, human figures in the synthesized public space can walk along the same route as that of the subjects. The tracking also enables the subjects to see the simulated crowd on their see-through face-mounted displays or other display devices since positional relationships between the subjects and the crowd can be calculated.

In the experiment to simulate interaction between the off-site and the on-site staffs at Kyoto Station, the subject who played a role of off-site staff was viewing the virtual station synthesized by the guidance system running on the PC in our laboratory. This laboratory is about three kilometers distant from the station where the subject who was playing a role of on-site staff was being tracked by the wizard person. As you can see in the left picture of the Figure 6, the wizard was carrying a laptop PC which sends the avatar's position to the PC in the laboratory through wireless networks. The subject in the laboratory could understand the location of the on-site subject and could talk with him/her by phone.

There are two ways for the wizard to observe a subject walking around in a large-scale physical space. If the wizard can have a bird's-eye view of the

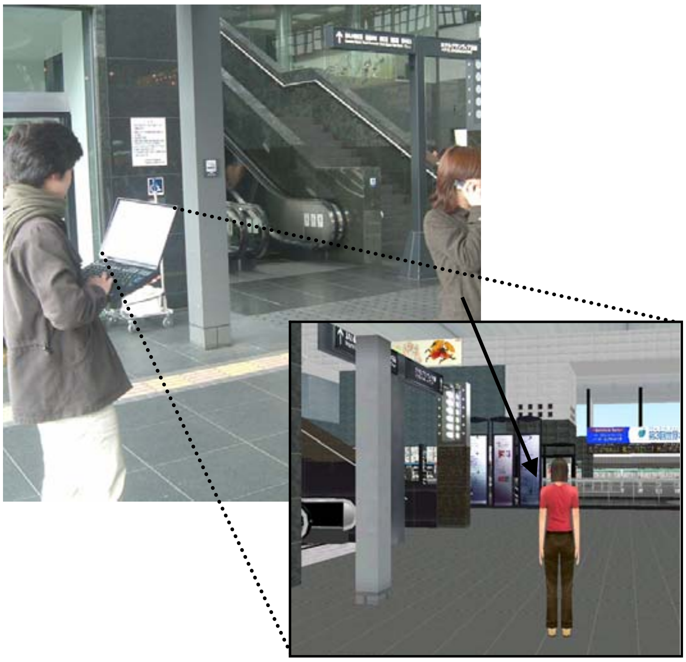


Fig. 6. Tracking by Observation

whole space and look down the subject from some higher location, pointing at his/her position on a 2D map would be better than controlling the avatar in a 3D space. However, there was not any good location to keep watch over the subject continuously because Kyoto Station has a complex structure, which fosters many blind spots. In such an environment, the other way in which the wizard keeps following the subject is an acceptable option. In this case, a 3D navigating interface is a good solution. In the experiment, the wizard could make use of the third-person viewpoint located behind the avatar very well since it was possible to move the avatar precisely just by keeping coincidence between the displayed view and his sight as given in Figure 6. This way of observation is more effective also in acquisition of such detailed context as gestural movements.

5 Conclusion

We could successfully simulate interactions necessary to evaluate the emergency guidance system, which is the example of massively multiagent ubicomp systems. We have tackled scalability issues about smart environments and people inside them [2]. We showed that virtual space can be a participatory simulation platform of smart environments.

We have focused on the case that many people are moving around and interacting with each other in a large-scale crowded smart environment. In that situation, the three different interactions occur. We proposed the three methods, each of which simulates each kind of interaction. Our future work includes unification of the methods.

Acknowledgement. We express our thanks to the Municipal Transportation Bureau and General Planning Bureau of Kyoto city for their cooperation. This work would have been impossible without the invaluable participation of Hideaki Ito, Shinya Shimizu, Tomoyuki Kawasoe, and Toyokazu Itakura. We thank Shigeyuki Okazaki, Toshio Sugiman, Ken Tsutsuguchi, Satoshi Koizumi, CRC Solutions, Mathematical Systems, and CAD Center for their support in the development of the guidance system and the social interaction platform. The platform is available at `http://www.lab7.kuis.kyoto-u.ac.jp/freewalk/` and `http://www.digitalcity.jst.go.jp/Q/`.

References

1. Abowd, G. D., Atkeson, C. G., Bobick, A. F., Essa, I. A., MacIntyre, B., Mynatt, E. D. and Starner, T. E. Living Laboratories: The Future Computing Environments Group at the Georgia Institute of Technology. ACM CHI2000 extended abstracts, 215-216, 2000.
2. Abowd, G. D. and Mynatt, E. D. Charting Past, Present, and Future Research in Ubiquitous Computing. ACM TOCHI, 7(1), 29-58, 2000.
3. Benford, S., Greenhalgh, C., Rodden, T. and Pycock, J. Collaborative Virtual Environments. CACM, 44(7), 79-85, 2001.
4. Bohnenberger, T., Jameson, A., Kruger, A. and Butz, A. Location-Aware Shopping Assistance: Evaluation of a Decision-Theoretic Approach. Mobile HCI 2002, LNCS 2411, 155-169, 2002.
5. Cassell, J., Sullivan, J., Prevost, S. and Churchill, E. Embodied Conversational Agents. MIT Press, 2000.
6. Clark, H. H. Using Language, Cambridge University Press, 1996.
7. Crabtree, A., Benford, S., Rodden, T., Greenhalgh, C., Flintham, M., Anastasi, R., Drozd, A., Adams, M., Row-Farr, J., Tandavanitj, N. and Steed, A. Orchestrating a Mixed Reality Game 'On the Ground'. ACM CHI2004, 391-398, 2004.
8. Drogoul, A., Vanbergue, D. and Meurisse, T. Multi-Agent Based Simulation: Where are the Agents? MABS2002, LNCS 2581, 1-15, 2002.
9. Hall, E. T. The Hidden Dimension. Doubleday, 1966.
10. Helbing, D., Farkas, I. J. and Vicsek, T. Simulating Dynamical Features of Escape Panic. Nature, 407(6803), 487-490, 2000.
11. Ishida, T. *Q*: A Scenario Description Language for Interactive Agents. IEEE Computer, 35(11), 54-59, 2002.
12. Ishida. T. Society-Centered Design for Socially Embedded Multiagent Systems. CIA2004, LNCS 3191, 16-29, 2004.
13. Kendon, A. Spatial Organization in Social Encounters: the F-formation System. A. Kendon, Ed., Conducting Interaction: Patterns of Behavior in Focused Encounters, Cambridge University Press, 209-237, 1990.
14. Linturi, R., Koivunen, M. and Sulkanen, J. Helsinki Arena 2000 - Augmenting a Real City to a Virtual One. Digital Cities, LNCS 1765, 83-96. 2000.

15. Murakami, Y., Ishida, T., Kawasoe, T. and Hishiyama, R. Scenario Description for Multi-agent Simulation. AAMAS2003, 369-376, 2003.
16. Nagel, K., Kidd, C. D., O'Connell, T., Dey, A. K. and Abowd, G. D. The Family Intercom: Developing a Context-Aware Audio Communication System. Ubicomp2001, LNCS 2201, 176-183, 2001.
17. Nakanishi, H., Nakazawa, S., Ishida, T., Takanashi, K. and Isbister, K. Can Software Agents Influence Human Relations? - Balance Theory in Agent-mediated Communities -. AAMAS2003, 717-724. 2003.
18. Nakanishi, H., Koizumi, S., Ishida, T. and Ito, H. Transcendent Communication: Location-Based Guidance for Large-Scale Public Spaces. ACM CHI2004, 655-662, 2004.
19. Nakanishi, H. and Ishida, T. FreeWalk/Q: Social Interaction Platform in Virtual Space. ACM VRST2004, 97-104, 2004.
20. Nakanishi, H., Shimizu, S. and Isbister, K. Sensitizing Social Agents for Virtual Training. Applied Artificial Intelligence, 19, 2005. (to appear)
21. Patterson, M. L. Nonverbal Behavior: a Functional Perspective, Springer-Verlag, 1983.
22. Reeves, B. and Nass, C. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, 1996.
23. Reynolds, C. W. Flocks, Herds, and Schools: A Distributed Behavioral Model. ACM SIGGRAPH87, 25-34, 1987.
24. Scherer, K., London, H. and Wolf, J. The Voice of Confidence: Paralinguistic Cues and Audience Evaluation. Journal of Research in Personality. 7, 31-44, 1973.
25. Sugiman T. and Misumi J. Development of a New Evacuation Method for Emergencies: Control of Collective Behavior by Emergent Small Groups. Journal of Applied Psychology, 73(1), 3-10, 1988.
26. Tepper, D. T. and Haase, R. F. Verbal and Nonverbal Communication of Facilitative Conditions. C. E. Hill Ed., Helping Skills: The Empirical Foundation, 211-223, 1978/2001.
27. Timney, B. and London, H. Body Language Concomitants of Persuasiveness and Persuasibility in Dyadic Interaction. International Journal of Group Tensions, 3, 48-67, 1973.
28. Zimbardo, P. G. and Leippe, M. R. The Psychology of Attitude Change and Social Influence. McGraw-Hill, 1991.