Connecting Digital and Physical Cities

Toru Ishida[1], Hiroshi Ishiguro[2], and Hideyuki Nakanishi[1]

[1]Department of Social Informatics, Kyoto University
[1]JST CREST Digital City Project
[2] Department of Computer and Communication Sciences, Wakayama University

Abstract. As a platform for community networks, public information spaces that mirror the city metaphor are being developed around the world. The aim of *digital cities* is to pursue a future information space for everyday urban life, unlike the creation of new businesses which is the current obsession of the Internet. We started the basic research project called "*Universal Design of Digital City*," a five year project established in 2000, a part of the Core Research for Evolutional Science and Technology (CREST) run by the Japan Science and Technology Corporation (JST). The objective of this project is to construct digital cities as the infrastructure that encourages the participation of all people, including the disabled and the aged. We will develop basic technologies for the universal design, focusing on '*sending information*,' '*receiving information*,' and '*participation*.' This paper introduces some of various experiments such as crisis management, environmental learning, and shopping street navigation. Digital cities are not imaginary since they correspond to the physical urban spaces in which we live. Basic technologies including perceptual information infrastructure and social agents are being developed for connecting digital and physical cities.

1. Why Digital Cities

The notion of digital cities can be defined as follows: *digital cities will collect and organize the digital information of the corresponding cities, and provide a public information space for people living in and visiting them* [4]. Digital cities have been developed all over the world, and can be connected to each other via the Internet, just as physical cities are connected by surface and air transport systems.

Why do regional information spaces attract people given that we are in the era of globalization? We realize that the Internet has fostered global businesses, but at the same time, it enables us to create rich information spaces for everyday life. While the Internet makes businesses global, life is inherently local. Business requires homogeneous settings to allow global and fair competition, while daily life can remain heterogeneous reflecting our different cultural backgrounds. If differences exist in business, standard protocols are needed to over come them, but we do not need any standard for social interaction. If the differences are significant, we will turn to the Internet for cross-cultural communication support.

M. Tanabe, P. van den Besselaar, T. Ishida (Eds.): Digital Cities, LNCS 2362, pp. 246-256, 2002.

With the development of the Internet, economies of scale have driven business activities to become global. Even small companies can participate in the worldwide business network, since the Internet significantly reduces search and negotiation costs to find partners and markets. The Internet makes commercial transactions much easier in most business areas. On the other hand, globalization is not so frequent in everyday life. Although the Internet has widened the information channels available, it cannot physically move people. Statistics show that people still spend their income for housing, shopping dining and so on where they live. Everyday life substantially remains local even though the Internet offers international reach.

Table 1. Two Extremes in Internet Use

Business	Everyday Life
Global	Local
Market (Competitive)	Community (Collaborative)
Homogeneous	Heterogeneous
Standard Protocol	Cross-Cultural Communication

Table 1 shows two extremes of Internet use. The table does not mean that the two types of usages are always disjoint, various combinations are common in Internet use. The motivation for studying digital cities is to shift our view of Internet use from one side (business) to another (everyday life). Both perspectives should be combined to build a public information space in which people can participate and interact, and to create consensus for the various problems.

2. What Is Real?

Digital cities are for supporting our everyday life. It is natural for people living in a city to create a public information space that corresponds to the physical city. The question is, however, whether or not it represents an efficient usage of the Internet. The Internet is often viewed as a new continent. For example, the Internet yields the possibility of building a virtual mall comprising a huge number of shops that cannot exist in any physical city. The locality of digital cities might constrain the possibilities inherent in the Internet.

To discuss this issue, we examine an interesting example found in the shopping street community in Kyoto. The community is formed by 3,000 shops in Kyoto. They jointly started a site that enables customers to make electronic account settlements by debit and credit cards. Purchasing requests from within or beyond the local community are processed electronically and goods are delivered by logistics companies. As a result, all shopping streets in Kyoto now appear on the Internet. Thousands of shops are already offering services under this framework. At first glance, this seems to be just another form of global virtual mall like Yahoo. However, its business model is totally different.

Global virtual malls are often called platform businesses. Providers of the platform offer reliable and trustful places, and invite suppliers as well as customers. It does not matter whether or not the suppliers have a presence in the physical world. The Kyoto shop alliance, on the other hand, represents real world entities. Credibility has been established through the long history of physical Kyoto. The question is how this local mall can compete with huge existing global virtual malls. Economies of scale are often observed in the Internet. It must be hard for small malls to achieve success because fewer services are offered. It seems that local malls, which are hard to scale up, cannot compete with global malls. This, however, is not the entire story, if consumers are not interested in just buying goods but in the city as a whole. In this case, a digital city, which creates a whole city in the Internet, can be a solution to support local malls.

We tend to think that a city constructed in cyberspace is fictional. But what is a real city? Is a physical city real? Let us take a modern company as an example. A physical company is composed of buildings and staff. However, it is almost impossible to understand its activities by observing only its physical space. Without reading e-mails or checking project WEB pages, activities of the company are hard to grasp. The company’s real activities are lie in both digital and physical space. We think “digital” and “physical” make things “real.” We think digital cities are not imaginary cities existing only in cyberspace, but complement the corresponding physical cities, and provide an information center for actual urban communities. As in modern companies, digital activities will become an essential part of urban life in the near future.

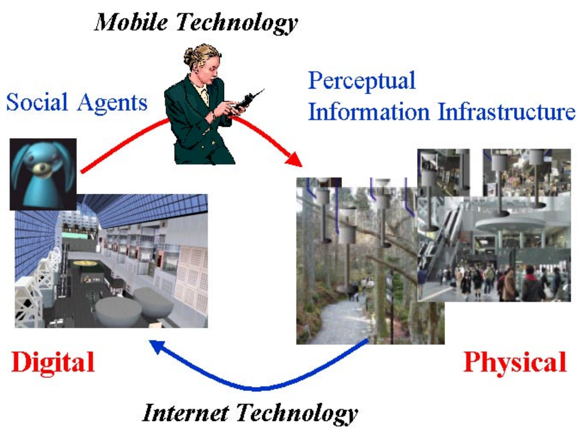

Fig. 1. Connecting Digital and Physical Cities

In the rest of this paper, we describe two different types of technologies currently under research to connect digital and physical cities. One is called the *perceptual information infrastructure*, which gives computer networks the ability to selectively obtain information from the physical space. The other is software called *social agents*, which can play a social role in human communities [3]. Figure 1 illustrates one approach to using these two technologies for integrating digital and physical cities.

3. Perceptual Information Infrastructure

The perceptual information infrastructure bridges the digital and physical cities by using computer vision technologies [6]. Omni-directional vision sensors embedded in the environment monitor human activities, build photo-realistic virtual spaces, and recognizing human behaviors. The perceptual information infrastructure offers new vistas to the Internet. Figure 2 shows a newly developed omni-directional vision sensor [5]. This camera takes 360-degree images. The images are reflected from the curved surface mounted just above the upward facing CCD camera. A black needle is mounted in the center of the camera. This needle absorbs all reflections from within the glass and makes it possible to get a clear picture. The entire camera can be downsized to just a few centimeters. As an example, a omni-directional sensor yielded the middle and right images in Figure 2. The omni-directional sensor is convenient for observing a town and communicating with multiple persons. The development of omni-directional sensors triggered the study of the perceptual information infrastructure. Various application systems for supporting people to send and receive information are under development.

Fig. 2. Omni-Directional Vision Sensor [5]

Sending information becomes easier with the omni-directional sensor. With a normal camera, it is necessary to turn the camera to the direction desired when taking a picture. This operation must be performed by users, but is not easy for aged or handicapped people. Otherwise, more expensive and complex machine control is needed. The benefit of the omni-directional sensor is that such control becomes unnecessary. Software has been developed that allows the desired section to be extracted from the 360-degree picture. For example, consider the case in which a handicapped or aged person wants to send their own picture to close relatives. It is not easy for them to adjust the orientation of the camera and speak to it accurately. With the omni-directional sensor, once it has been positioned, all they have to do is to speak to it, the software extracts the proper picture.

The omni-directional sensor also makes receiving information easier, because a virtual space can be created easily from 360-degree photos. When a series of video streams are available, a 3D virtual space that permits walk through can be created automatically. In detail, the system smoothly shifts the pictures to match the user's movements. If we place multiple omni-directional sensors in a physical space, users can walk though the live videos taken by those sensors. This is a new alternative to building 3D computer models. It enables us to create historical and natural objects, such as temples and forests. The methodology, called *Town Digitizing*, forms a real-time virtual space that permits walk-through. The images captured by a series of omni-directional vision sensors are automatically fused based on the visitor's "virtual" location to reproduce what the user would see in a corresponding physical city. Although Town Digitizing requires a high-speed computer network, the rapid progress of the Internet will relax this problem in near future.

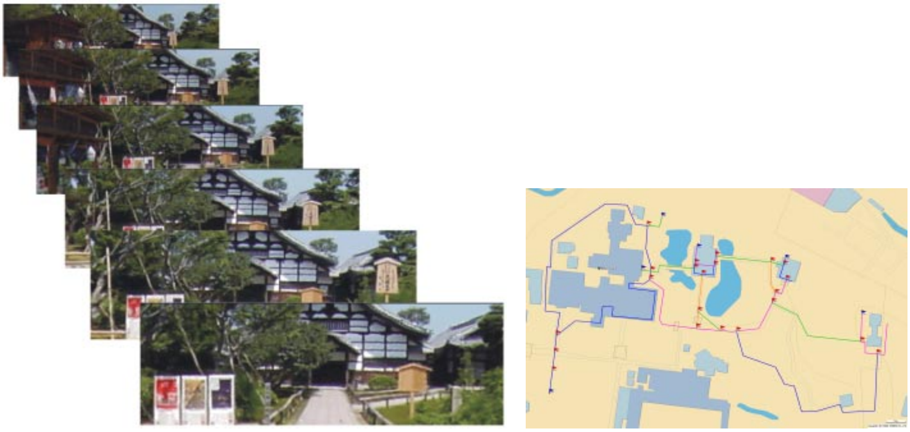

Fig. 3. Town Digitizing in Kodaiji

Figure 3 shows a map of Kodaiji (temple) in Kyoto. The photos in Figure 3 are a typical image sequence of users moving in a virtual space. As shown in Figure 3, Town Digitizing is especially suitable for representing environments with many historical or natural objects [7]. Such environments are often far from the city center and are not easy, especially for aged people, to access. People can share a historical or natural environment through the virtual spaces built *by Town Digitizing*.

The omni-directional vision sensor is not only for building virtual spaces but can also be used for recognizing human activities. We develop software to track walking humans, record their trajectories, and recognize their behavior. One merit of this multiple camera system is that the viewing area of the cameras can overlap. That is, the sensor network can cover a wide area with a small number of cameras and offer stereo views. This sensor network can be used in many situations. It can monitor and support the activities of aged people in hospitals and streets. It can realize secure monitoring systems in public places, such as stations and banks.

As described above, the omni-directional sensor and the network remove a barrier between the physical and digital spaces and make sending and receiving information easier, even for the aged and handicapped people.

4. Social Agents

Digital cities have two functions: to collect and reorganize digital information of corresponding cities, and to provide a public information space for people living in /visiting them. The idea of human-like software to support activities in digital cities seems unrealistic. However, research into artificial intelligence and social psychology is bringing us close to realizing this goal.

To connect residents and visitors, one trial for developing social interaction in digital cities uses avatars in the 3D space. Figure 4 shows a view of avatars controlled by humans walking in the 3D virtual space called *FreeWalk* [8]. This technology allows a number of avatars to walk around the digital city in real-time. By making links between the avatars and the people walking in the corresponding physical city, we can realize communication between digital tourists and people walking in the city. As walking motion can be generated by the user's machine via a WEB browser plug-in, only the walking position/velocity and direction need be downloaded. Thus, a large number of avatars can be created rapidly in real-time. Aside from the "known" avatars, adding a virtual population will activate the digital city and make it more attractive. In addition, a scenario description language called *Q* that can control the behavior of thousands of social agents is currently under development. Agents autonomously behave under the social constraints given by the scenarios, and dynamically respond to other avatars (humans) and agents (software).

We are trying to create a group of social agents that support the various activities of participants in a digital city. We expect social agents to show their ability as coordinators for building and maintaining online communities. The playgrounds of social agents include virtual and mobile environments. In a virtual environment, social agents interact with avatars (humans) visiting the same world. In a mobile environment, social agents appear on PDAs or wearable computers to interact with people living in the physical world. Social agents can enter into communities in both environments. The service industry is a good example. Social agents can troll through vast information repositories of products and services, communicate with customers, and work twenty-four hours a day.

Fig. 4. Avatars and Agents in a 3D Virtual Kyoto

Evacuating people safely in the case of disaster is another example of social agents playing an important role both in digital and physical cities. In evacuation simulation, the possible roles of social agents include pedestrians, station employees, salesclerks and so on. Realistic evacuation simulations can be realized by having pedestrian agents act as people running around to escape. Such simulations would help a crisis management center to accumulate experiences and to make correct decisions, since social agents would exhibit the mistakes typical of people. In real evacuation, social agents would appear on the user's PDA to provide appropriate instructions. For success of simulation, social agents have to be trusted by people. Functional accuracy to make correct decisions is not enough to be trustworthy agents. Technologies to create trustworthy agents is to be explored.

To understand the nature of social agents, we are performing a series of social psychological experiments. Effects of introducing social agents in human communities have not been well investigated. We are currently conducting experiments to see how social agents support human communication and influence human relations. Social agents can act as go-betweens among people who have different social identities such as inhabitants and visitors, young and aged, and so on. In our experiment on cross-cultural communication between Japanese and American students, the agents influenced not only the impressions of agents but also the impressions of conversation partners and the stereotypes of nationalities [2]. For example, if the agent encouraged students to discuss political problems, the Japanese students became as talkative as the Americans.

Another experiment to observe the influence of agents on human-human relations shows that social agents have the ability to affect human relations. We first observed that if an agent agrees/disagrees with a human's opinion, the human develops positive or negative feelings toward the agent. Is balance theory [1] works in this situation? If so, when a human has a positive or negative feeling toward the agent and thinks that the partner has the same feeling to the agent, the human tends to have a positive feeling toward the partner. Contrary, if a human thinks that the partner has a different feeling toward the agent, the human tends to have a negative feeling toward the partner. Figure 5 explains this situation. We observed that the agent that lies, i.e. has exhibits different responses to different people, can create discord between people. On one hand, if there is less communication between the two people, it is easier for the agent to control their relationship: the agents may become unsafe entities if the amount of communication between users is insufficient. On the other hand, we found that people could use agents safely in their community if the communication among people is adequate while using the agents. Results of this experiment demonstrate the ability and the limitations of agents to influence human relations. In future, we may need an ethics law governing the development of agents.

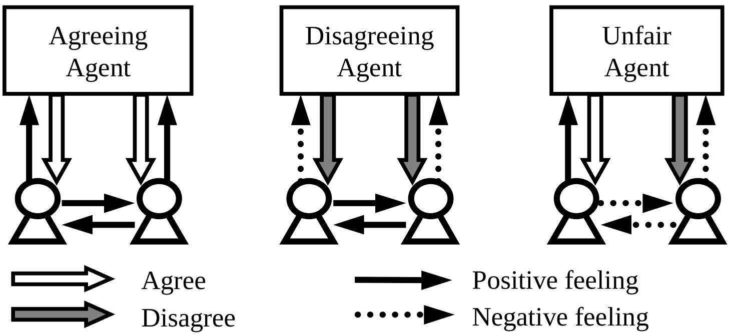

Fig. 5. Balance Theory for Agent and Two People

We are currently studying conversational agents that learn; they improve their social behavior automatically. It is still difficult to develop general-purpose agents that can communicate on all topics with people satisfactorily. However, by carefully selecting the application domains there is a hope that we can create task-specific agents. The Wizard of Oz method has a framework that makes agents learn conversation gradually. In the first stage, agents suggest speech candidates for a wizard (a human) behind them to select the most appropriate one. If no appropriate speech is suggested, the wizard enters dialog directly. Users speaking to those agents do not notice the presence of the wizard behind the agents. The first stage can be implemented easily, but the agents never become intelligent in this stage. In the second stage, a learning capability should be added to the agents. The agents memorize conversations, create finite state automata, and generalize them to keep the automata tractable. By narrowing down tasks, the agents would offer quite good speech candidates to the wizard after training. They would finally converse without help of the wizard. Such conversational agents can be applied to various tasks. The agents will be useful for shopping, navigation, and supporting the aged and handicapped.

5. Applications

Connecting digital and physical cities will stimulate the creation of various applications. As part of our digital city project, we advanced three pilot applications.

The first application involves evacuation simulations in a three dimensional virtual Kyoto station; this pilot application project links computer scientists, architects, and social psychologists. The simulation is planned to be performed with 100 people connected via the Internet, and 1000 social agents controlled by a given scenario. In previous research into simulating human behavior in crisis situations, people were displayed as small circles in two-dimensional flows, a technique common in hydrodynamic simulations. In a real crisis situation, directions from officers and speech among people will greatly influence the result. Unfortunately, the 2D simulation fails to well reproduce the social interaction. The 3D virtual cities now make it possible to recreate this interaction. The scenario is as follows.

> In 200X, an evacuation simulation involving Kyoto railway station will be conducted in 3D virtual Kyoto with 1000 social agents. 100 citizens will participate via the Internet. The data is to be collected and analyzed for planning real evacuations. In 200Y, an earthquake occurs and a fire starts in Kyoto station. The situation in the physical space is captured by omni-directional sensors and sent to the control center through wireless networks. The movement of people as a group is grasped and displayed on the screen in the center. Based on the results of previous evacuation simulations, appropriate directions are sent to the peoples' mobile terminals, and conversational agents start to guide their users.

The evacuation simulation in the virtual Kyoto station brings up the following issues. What can we learn from the simulation results in the virtual space? Even if the simulation indicates heavy casualties, it is not clear what this means. People may think the virtual space simulation is just like a video game. We thus started research from two different aspects. The first aspect is to understand how people behave differently in the physical space from the virtual space. People are given the same task in both spaces (for instance, leave via a given exit) and all behaviors recorded by eye cameras are analyzed. We are interested in how difficult it is to move in a virtual space without peripheral vision, which is inherent in the physical space. The second aspect is to understand the difference between human-human and human-software interaction. People's reactions must be different according to whether the directions were given by humans or software officers, but different in what way? These issues must be analyzed carefully in controlled experiments before conducting a large-scale simulation.

The second application involves navigating aged people through shopping streets using the perceptual information infrastructure and social agents.

> An aged person is looking for a shop on Shijo Street. There is a map, however, it does not well correspond to the real town. The omni-directional sensors detect the person's position, orientation, and behavior. All information is sent to the control center through wireless networks. The information is combined to understand the situation of the street including that of the aged person. His view is reproduced using the virtual street, and landmarks appropriate for him are extracted. A conversational agent then directs him by using the extracted landmarks. An image extracted from the virtual street is displayed on his portable terminal to indicate the landmarks.

For the third application, the same technologies are applied to environmental learning in a university forest.

> Parents and children are provided with studies on the environment in a forest of Kyoto University. Live videos, taken by pre-installed omni-directional sensors, are sent to the education center via a broadband fiber network. A virtual space is reproduced from the live videos. An instructor (or a social agent) of environmental learning in the center can know what is going on in the forest in real time. A participant may send photos of plants to the center to ask the instructor for information. The instructor then moves to the participant's location in the virtual forest, talks to the participant, and sends the requested information to the participant's portable terminal.

What is common in these three experiments is that the perceptual information infrastructure and social agents combine physical and digital cities to support people not only when they sit in front of their PCs but also when they move in the physical world.

6. Cross-Cultural Communication in Digital Cities

Unlike the conventional telephone network, there is no centralized control mechanism in the Internet. Even the power of governments cannot stop the flow of information. Anyone such as a politician, scholar, businessperson, and student has equal access to the enormous store of information. Instead of huge hierarchical organizations, a knowledge network of small organizations or individuals covers the world.

Unlike conventional mass media, the Internet allows us to directly contact thousands of vivid instances. While mass media filters collected information, refines and edit them, the Internet is a huge depository simply accumulating original instances. For example, children who want to be music conductors can find plenty of information not only about famous conductors but also about other children in foreign countries with the same dreams.

There are two reasons for discussing cross-cultural communication in the context of digital cities. The first is that, in any city in the world, the citizens are becoming more diverse. A public information space can play a major part in bridging the cultures. The second reason is that a digital city will represent the real city on the Internet. We have 40,000,000 visitors per year in Kyoto, but we expect more people will visit Kyoto via the Internet. Therefore, cross-cultural communication must be supported, even though the main purpose of digital cities is to support the everyday life of local residents. A local information space and a platform for cross-cultural communication are two sides of the same coin.

We should be aware that the ratio of English web pages worldwide has been decreasing rapidly. As the Internet will be applied to everyday life more and more, the ratio of English web pages will keep on decreasing. It is obvious that the need for cross-cultural communication will increase significantly. Without linguistic and cultural support tools, it is not possible for people to create productive interaction. Connecting digital and physical cities will, together with the latest machine translation technologies, increase the opportunity to participate in cross-cultural communication.

Acknowledgement. The author wishes to thank, Scott Brave, Satoshi Koizumi, Clifford Nass, Satoshi Nakazawa, Katsuya Takanashi, Mohan Trivedi, and Ken Tsutsuguchi for their collaborative work, and many other colleagues for helpful discussions.

References

1. F. Heider, The Psychology of Interpersonal Relations, Wiley, 1958.
2. K. Isbister, H. Nakanishi, T. Ishida and C. Nass, “Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space,” *CHI-2000*, pp. 57-64, 2000.
3. K. Isbister, “A Warm Cyber-Welcome: Using an Agent-led Group Tour to Introduce Visitors to Kyoto,” *Digital Cities: Experiences, Technologies and Future Perspectives*, Springer-Verlag, pp. 391-400, 2000.
4. T. Ishida and K. Isbister Eds., *Digital Cities: Experiences, Technologies and Future Perspectives*, Lecture Notes in Computer Science, 1765, Springer-Verlag, 2000.
5. H. Ishiguro, “Development Low-Cost and Compact Omnidirectional Vision Sensors,” In Ryad Benosman and Sing Bing Kang Eds., *Panoramic Vision: Sensors, Theory and Applications*, Springer-Verlag, Berlin, 2001.
6. H. Ishiguro, “Distributed Vision System: A Perceptual Information Infrastructure for Robot Navigation,” *IJCAI-97*, pp. 36-41, 1997.
7. S. Koizumi, G. Dai and H. Ishiguro, “Town Digitizing for Building an Image-Based Cyber Space,” *Digital Cities II*, Springer-Verlag, 2002.
8. 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. H. Rheingold, *The Virtual Community*, Addison-Wesley, 1993.