

CLASSROOM IN THE ERA OF UBIQUITOUS COMPUTING SMART CLASSROOM

CHANGHAO JIANG, YUANCHUN SHI, GUANGYOU XU AND WEIKAI XIE

*Institute of Human Computer Interaction and Media Integration
Computer Science Department, Tsinghua University
Beijing, 100084, P.R.China*

*E-mail: jiangch@media.cs.tsinghua.edu.cn, shiyc@tsinghua.edu.cn,
xgy-dcs@mail.tsinghua.edu.cn, xwk@media.cs.tsinghua.edu.cn*

Abstract: This paper first presents four essential characteristics of futuristic classroom in the upcoming era of ubiquitous computing: natural user interface, automatic capture of class events and experience, context-awareness and proactive service, collaborative work support. Then it elaborates the details in the design and implementation of the ongoing Smart Classroom project. Finally, it concludes by some self-evaluation of the project's present accomplishment and description of its future research directions.

Keywords: Ubiquitous Computing, Intelligent Environment, Multimodal Human- Computer Interaction, Smart Classroom

1 Introduction: From UBICOMP to Smart Classroom

Desktop and laptop have been the center of human-computer interaction since the late of last century. As is a typical situation of human's dialogue with computer that a single user sits in front of a screen with keyboard and pointing device, interacting with a collection of applications [Winograd 1999]. In this model, people often feel that the cumbersome lifeless box is only approachable through complex jargon that has nothing to do with the tasks for which they actually use computers. Too much of their attention is distracted from the real job to the box. Deeper contemplation on valuable matured technologies tells us: the most profound technologies are those that disappear, which means they weave themselves into the fabric of everyday life until they are indistinguishable from it [Weiser 1991]. We use them everyday, everywhere even without notice of them. Based on this point of view, computer is far from becoming part of our life.

Mark Weiser first initiated the notion of Ubiquitous Computing (UBICOMP) at Xerox PARC [Weiser 1993], which envisioned in the upcoming future, ubiquitous interconnected computing devices could be accessed everywhere and used effortlessly, unobtrusively even without people's notice of them, just as electricity or telephones of today. This inspiring view of prospect has been accepted and spread so fast and widely that in a short time of a few years, many ambitious projects have been proposed and carried on to welcome the advent of UBICOMP. There are a bunch of branch research fields under the banner of UBICOMP, such as Mobile Computing, Wearable Computing and also Intelligent Environment, etc. The focus of this paper, Smart Classroom, belongs to the field of Intelligent Environment. But what is Intelligent Environment? In our point of view, we define it as an augmented spacious environment populated with many sensors, actuators and computing devices. These components are interwoven and integrated into a distributed computing system which is able to perceive its

context through sensors, to execute intelligent logic on computing devices and serve its occupants by actuators. (In some research projects, Intelligent Environment is also referred as Interactive Space, Smart Space etc.)

In researches of Intelligent Environment, there are several relevant and challenging issues need to be solved, such as the interconnection of computing devices on many different scales, the handling of various mobility problems caused by user's movement, and network protocol, software infrastructure, application substrates, user interfaces issues etc. Although many projects have been conducted in the name of Intelligent Environment, they have different emphases. Some focus on the integration of different sensing modalities [Coen 1999, HAL 2000], some aim at the adaptability of Intelligent Environment to user's preference [Mozer 1999], some are interested in automatic capture of events and rich interactions that occurs in an Intelligent Environment [eClass 2000, Adowd 2000], and some target at facilitating the collaboration of multi-user multi-device within a technology-rich environment [Interactive Workspace 2000, Fox 2000]. We can easily enumerate several other ongoing Intelligent Environment projects with different specializations, such as Georgia Tech's Aware Home [Aware Home 1999, Kidd 1999], IBM Research's DreamSpace [DreamSpace], Microsoft Research's EasyLiving [EasyLiving] etc. Our institute developed special interest in exploring the impact of ubiquitous computing to education. This leads to the project of Smart Classroom. The Smart Classroom is a physical experimental environment, which integrates multimodal human computer interface with CSCW modules collaborating through inter-agent communication language to provide a smart space for lecturer's natural use of computer to give class to distance learning students.

In the rest of this paper, we're going to first present our views of futuristic classroom in UBICOMP. And then toward the ideal model of classroom, which sounds a little utopian, we'll explain the idea and focus of our exploration. Later some details in the design and implementation of our present work will be illustrated. We'll conclude by a short description of our future goals.

2 What Should Classroom in The Era of UBICOMP Be Like?

Michael H. Coen from MIT Artificial Intelligence Lab said, "Predicting the future is notoriously difficult". [Coen 1999] Yes, we're not able to prescribe what future would be, but we're able to create toward what we think it would be. In our point of view, the following features are essential to a smart classroom in the era of UBICOMP, and will serve as the guidelines in our ongoing Smart Classroom project.

We have generalized four characteristics of futuristic classroom, which are: natural user interface, automatic capture of class events and experience, context-awareness and proactive service, collaborative work support.

2.1 Natural user interface

As Mark Weiser has observed "Applications are of course the whole point of ubiquitous computing". In accordance with this essence of UBICOMP, it is necessary for a smart classroom to free its occupant's attention to computer. To rescue people's energy from irrelevant interaction with computer to the intentioned goal, allowing user's interaction with computer as naturally as possible is vital. In such a new paradigm of human-computer interaction, people input information into computer in their most familiar and accustomed ways like voice, gesture, eye-gaze, expressions etc. Auxiliary input devices like keyboard, mouse, are not necessary. In the reverse side, computer tends to serve people like an intelligent assistant. It utilizes technologies like projector display, voice synthesis, avatar, etc. This is what we call natural user interface.

To get a clearer image, suppose a lecturer in the Smart Classroom conducting the class by voice. "Let's go to chapter two". Computer recognizes phonetic command and projects the wanted courseware of chapter two on display. Lecturer also uses hand gesture as a virtual mouse to annotate on the projected electronic board. Through combination of eye gaze (or finger pointing) and voice command, lecturer can zoomed in the image of an area in the projector to give emphasized explanation on a specific topic.

2.2 *Automatic capture of class events and experience*

This is what eClass project of Gatech called "automated capture, integration and access" problem. We use computer in classroom not only to improve the quality of teaching activity, but also to augment its capability, which was impracticable traditionally. The automatic capture of class event and experience belongs to such capabilities. It's not just record of video and audio in the environment, which is common in traditional distance learning-television broadcasting program. It includes the record of group collaboration, multimedia events, multiple channels of human computer interaction, etc, all the events and experience that happened in the environment. The captured events and experience should be assembled into a kind of multimedia compound document. People can recreate the class experience by play the recorded multimedia compound document, and also can search a specific event or query knowledge within the compound document. This technology provides lecture content to students who are unable to attend the class in person, as well as to those who wish to review the materials later.

For example, suppose a lecturer giving a class on Artificial Intelligence in a Smart Classroom. All the audio, video information, lecturer's annotation events, student's question events, Smart Classroom's controlling of lights, slides, etc, are recorded into a multimedia compound document. When a student wants to review the knowledge of Alpha-Beta Pruning algorithm, he can just query about it through his laptop computer, and rewind to the previous talk on it for a quick review and then comes back. After the class, students can also replay the document to recreate the class experience.

2.3 *Context-awareness and proactive service*

What is context-awareness? According to Dey & Abowd (from Gatech 1999), "context is any information that can be used to characterize the situation of an entity, where an entity can be a person, place, physical or computational object", "context-awareness is to perceive the context by system so as to provide task-relevant information and/or services to a user, wherever they may be". Which means the Intelligent Environment can understand user's intention not only based on audio-visual inputs, but also based on its situational information. Proactive service means to serve the user without his request. Proactive service is based on the Intelligent Environment's capability of Context-awareness. This model of service is disparate from traditional human-computer interaction paradigm, in which computer respond to human's explicit command. In the Intelligent Environment, the computer remembers the past, recognize the present, and predicate the future. It reasons human's intention through analysis of all the information from accumulated knowledge base. Then it tries to serve its occupant's proactively with the reasoned intention of its user.

For example, the lecturer is explaining a formula displayed on the electronic board. When the lecturer points at it and starts to talk about it, the computer understands that the lecturer is going to attract students' attention to the specific area of display. Then it zooms in the area containing that formula on the display without the need of lecturer's commanding "Zoom in this region". Another example, when the lecturer wants to have a student named Wang to give his

opinion on a topic, he points at the student and says, “Wang, would you please say something about what you think of this problem?” The computer then automatically focuses the video camera and microphone array on Wang and filters out the noise emitted from other spaces

2.4 Collaborative work support

Class is essentially a collaborative procedure involving multiple participants. In the environment of ubiquitous computing, the introduction of many interconnected computing devices and wide area network support enables us to extend beyond the space boundaries imposed by traditional classrooms. With this technological advance, collaborations of multi-user and multi-device can be possible. And the support for collaboration is becoming a requisite of a smart classroom. The collaborative work support of a Smart Classroom can be categorized into two classes. One is the collaboration of multiple attendants within the Smart Classroom holding various computing devices like, pen-based devices, handheld devices and wearable computer etc. The other is the collaboration of remote participants and local attendants. The demand for collaboration support is so obvious that many commonly observed tasks in a classroom, such as group discussion, evolve the collaboration among multiple persons. There are some projects specialized in enabling and exploiting smart classroom’s collaboration support, such as collaborative note-taking in both Gatech’s eClass [eClass 2000] and Stanford’s Interactive Workspaces [Interactive Workspaces 2000].

3 The Focus of Smart Classroom

Smart Classroom is a big project, every above-mentioned aspect of it is challenging and a long-term effort. Our institute has been committing itself to the research on multimodal human computer interfaces, CSCW in wide area network, and also multimedia integration. Based on our existing research results we have been investigating Smart Classroom’s following features: natural user interface, automatic capture of classroom events and experience, and collaborative work support. So in initial phase of our project, we focus on applying our previous research achievements to realize an experimental environment. We have set up the physical experimental environment to demonstrate our idea and focus. In this Smart Classroom, we mainly aim at the following features: conducting lessons by means of gesture and voice command, capturing class events and operation such as manipulation on courseware, video-audio streams of class, etc, admission control of students using all kinds of mobile computing devices. In order to give clearer image of our research, an elaboration of the physical experimental environment layout and its user-experience scenario are given as following.

3.1 The layout of Smart Classroom

Our Smart Classroom is physically built in a separate room of our lab. Several video cameras, microphones are installed in it to sense human’s gesture, motion and utterance. According to UBICOMP’s characteristic of invisibility, we deliberately removed all the computers out of sight. Two wall-sized projector displays are mounted on two vertically crossed walls. According to their purposes, they are called “Media Board” and “Student Board” separately. The Media Board is used for lecturer’s use as a blackboard, on which prepared electronic courseware and lecturers’ annotation are displayed. The Student Board is used for displaying the status and information of remote students, who are part of the class via Internet. The classroom is divided in to two areas, complying with the real world classroom’s model. One is the teaching area, where is close to the two boards and usually dominated by lecturer. The

other is the audience area, where is the place for local students. Why are both remote students and local students supported in this room? The answer is simple, that we're complying with the philosophy of Natural and Augmented. Natural means we'll obey real-world model of classroom as much as possible to provide lecturer and students the feeling of reality and familiarity, which leads to the existence of local students. Augmented means we'll try to extend beyond the limitation imposed by the incapability of traditional technology, which is the reason for remote student.

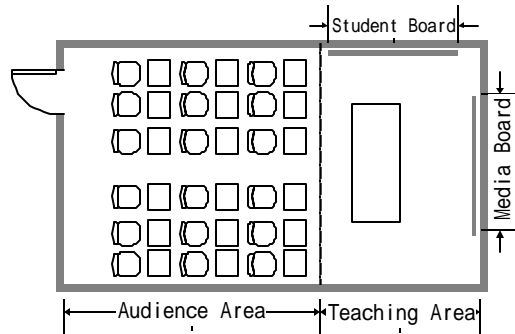


Figure 1 Layout of Smart Classroom

In Smart Classroom, users' rights to use the room are mapped to their identification. There is an audio-visual identification module for identifying the users in this room and authorizing control right to lecturer. With the help of visual motion-tracking module, Smart Classroom can be aware of its occupants' places in the room. Once a user identified as lecturer entering the teaching area, he is authorized to control the Smart Classroom by voice and gesture command. Lecturer can use hand-gesture as a virtual mouse on the Media Board to annotate, or add/move objects on the electronic board. He can also command linking in courseware, perform operations like scrolling pages, removing objects, granting speech right, etc by voice. All the lecturer's operation on the courseware and audio-video information captured in the Smart Classroom are automatically recorded and integrated to a multimedia compound document. The recorded information is simultaneously broadcasted to remote Students via Internet synchronously. Through software's application layer trans-coding and adaptive reliable multicast transport, remote students are promised to join the class with devices varied in computational power and display resolution through heterogeneous network varied in quality of service.

3.2 A typical user experience scenario

The following is a typical user-experience scenario happened within the Smart Classroom. Multiple persons enter the room through the door. At the door, there is an audio-visual identification module identifying the entering person's identity through facial and voice identification. If the person is identified as lecturer, he is granted the control right of the Smart Classroom. The visual motion-track module tracks the lecturer's motion in the room. Once he steps into the teaching area, he will be able to use gesture and voice command to exploit the Smart Classroom to give lessons. Persons in the Smart Classroom other than lecturer are deemed as local students. When the lecturer is in the teaching area, he can start the class by just saying, "Now let's start our class." The Smart Classroom then launches necessary modules such as Virtual Mouse agent, Same View agent (which will be talked about later). Lecturer loads

prepared electronic courseware by utterance like, “Go to Chapter 1 of Multimedia course”. The HTML-based courseware is then projected on the wall display. Lecturer can use hand-motion to stimulate the Virtual Mouse agent to annotate on the electronic board. Several type of hand gestures are assigned corresponding semantic meanings, which cause several operations like highlighting, annotating, adding pictures, remove object, executing links, scrolling pages etc, on the electronic board. Lecturer can also grant speech right to remote students by finger pointing or voice command, like “Li, please give us your opinion”. On the Student Board, remote students’ photos and some information as name, role, speech right etc are displayed. When a remote student requests for floor, his icon on the Student board twinkles. Once the lecturer grants the floor to a specific remote student, his video and audio streams are synchronously played both in the Smart Classroom and on other remote students’ computers.

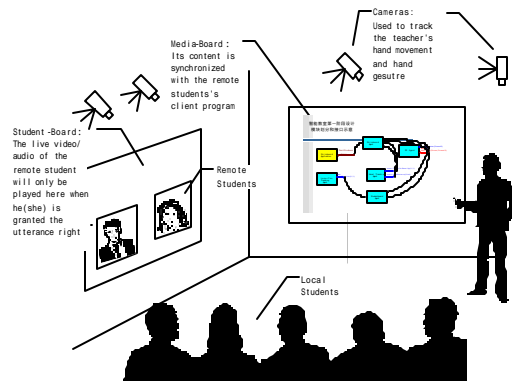


Figure 2 Typical scenario in Smart Classroom

4 Details of Smart Classroom’s Design and Implementation

The Smart Classroom is essentially a distributed parallel computing environment, in which many distributed software/hardware modules collaborate to accomplish specific jobs. Software infrastructure is the enabling technology to provide facilities for software components’ collaboration. There are some alternative solutions to software infrastructure, such as Distributed Component-Oriented Model, like EJB, CORBA, DCOM, et c, and Multi-Agent Systems (MAS). In the context of Intelligent Environment, Multi-Agent System is more competent than Distributed Component-Oriented Model due to the following reasons: higher encapsulation level, faster evolution from design to implementation, easier development and debugging and most importantly, more accordant to the need of dynamic reconfiguration and loose-coupling.

4.1 4.1 Software platform--OAA

In current stage, instead of developing our own Multi-Agent System, we choose to use SRI’s famous open source MAS product, Open Agent Architecture (OAA) [OAA]. There’re already many successful multimodal human-computer interaction projects built on OAA. Its delegating computing model also fits well in our need of software infrastructure. In OAA’s delegating computing model, the network of distributed software modules is conceptualized as a dynamic community of agents, where multiple agents contribute services to the community.

When external services or information are required by a given agent, instead of calling a known subroutine or asking a specific agent to perform a task, the agent submits a high-level expression describing the needs and attributes of the request to a specialized *Facilitator agent*. The Facilitator agent will make decisions about which agents are available and capable of handling sub-parts of the request, and will manage all agent interactions required to handle the complex query. Such a distributed agent architecture allows the construction of systems that are more flexible and adaptable than distributed object frameworks. Individual agents can be dynamically added to the community, extending the functionality that the agent community can provide as a whole. The agent system is also able to adapt to available resources in a way that hard-coded distributed objects systems can't.

4.2 Five dedicated agents in Smart Classroom

In the schematic figure of our Smart Classroom there're five dedicated agents (except for the Facilitator agent of OAA). The Facial-voice identification agent is in charge of the Smart Classroom's login identification and authentication. When a person entering the room, he is required to place his face into a specific zone of a video camera's capture range, and speak a login word. The vision-part of the agent identifies the person by searching in a pre-trained user library, and the voice-part authenticates the identified person by voice-based speaker recognition.

The motion-tracking agent is a computer vision-based agent. There's a pan-tilt video camera mounted on the upper side of the front wall, monitoring the whole range of the room. The motion-tracking agent receives video stream input from the camera and tracks the lecturer's position and movements in the room. When lecturer enters /leaves the teaching area, motion-tracking agent will signal the corresponding events to the agent society.

The voice command support of Smart Classroom is realized by a speech recognition agent, which can perform speaker independent and continuous voice recognition. We use IBM's simplified Chinese version of ViaVoice SDK to wrap the voice recognition agent. The agent receives digitized signals from a wireless microphone, which is carried by the lecturer. And then recognizes its command within a dynamically loaded vocabulary set. Once a recognizable command is reached, the voice recognition agent dispatches that command to the agent society. Other corresponding agent is responsible for the execution of that command.

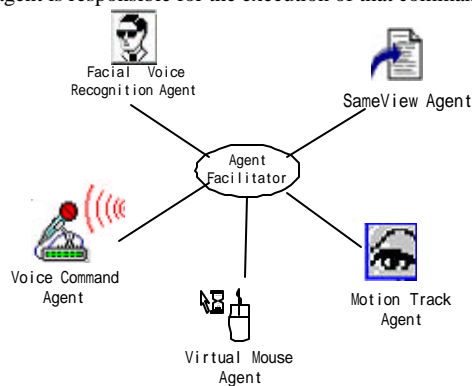


Figure 3. Five dedicated agents in OAA model

The Virtual Mouse agent is used for handling hand-gesture, which stimulates the mouse events and shortcut command to activate operations on the playing courseware. It's also a vision-based agent. There are two video cameras specialized for virtual mouse event recognition. One is installed on top of the screen, the other is mounted on the ceiling of the room. Through detecting and analyzing 3D movements of hand, gestures can be recognized. The virtual mouse agent then dispatches the recognized mouse event or shortcut command to the agent society.

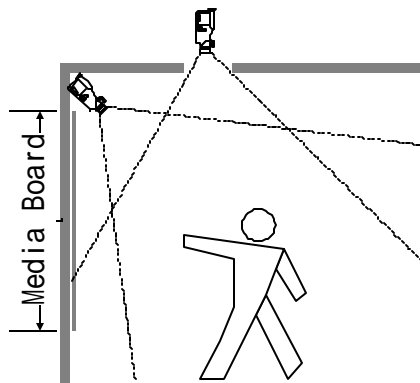


Figure 4 How virtual mouse agent works

SameView agent plays a core role in our Smart Classroom's pedagogical scenario. It is based on a legacy desktop application, namely SameView [Pei 1999, Liao 2000, Tan 2000], which is developed by media group of our institute. The purpose of SameView is a software for supporting multimedia based group discussion whose members are spaciouly distributed and connected by heterogeneous networks. SameView has the following features: a shared MediaBoard (multimedia extensions to traditional electronic whiteboard), adaptive multimedia contents trans-coding according to terminal's network Qos and computing power, adaptive reliable multicast in wide range of heterogeneous networks, live capture of video/audio streams and multimedia events into self-defined multimedia compound document, post-edit and playback of the captured multimedia compound document, self-equipped authoring tools for courseware-editin g.

In our Smart Classroom, we recur to the SameView's desktop version code as much as possible. We only revised some of its input/output user interface, such as adding a separate Student Board for display of remote students information and status by exploiting dual display adapter card support of Microsoft Windows 98/2000, projecting the Media Board in full-screen mode to remove the vestige s of desktop software with Windows-style menu, toolbar, title bar, etc. The most crucial reformation to SameView is the wrapping of it as an autonomous agent in Smart Classroom's agent community, which enables it to receive user's natural input from other dedicated agents like voice command recognition agent and virtual mouse agent, and then behave interactively.

5 Conclusion: Future Goals for Smart Classroom

Our current stage Smart Classroom is a primitive prototype of futuristic classrooms, which attempts to embody some of its distinguishing features like natural user interface, capture of class events and collaborative support. It is still far from a real Smart Classroom. Its resolutions to some key problems in Intelligent Environment are simple, intuitive and somewhat application-specific. Although many research issues need to be addressed in order to realize genuine Smart Classroom, we stride forward the first step toward the ambitious goal. In the near future, we'll make efforts in enhancing the Smart Classroom in some of the following aspects.

Add more modalities and applications. We'll try to equip some more modalities of human inputs like vision-based tracker, embedded microphone array and various distributed sensors to sense human's context. And progress in the sensing technologies needs to be matched by progress in applications that use sensed information. Application is one of the key driving forces of technical advance. We'll conceive more realistic and useful scenarios in the Smart Classroom and also cooperate with different research groups whose application projects have high potential to take advantage of the capability of Smart Classroom. We believe the rise in the amount and sort of applications will enable generalization of Smart Classroom's design and implementation.

Add a brain. The current design implementation of the Smart Classroom focus on the human's natural input to the computing environment, the next step is to move to a higher level and to give it the ability to understand. It is not just to utilize multimodal interface, but also to add-on context-aware intelligence. The classroom should be able to reason human's intention through analysis of all the gathered inputs and proactively serve its occupants. There is some research studies in reasoning human's intention based on predefined grammars [Johnston 1998] or probabilistic statistical model. Each of them has innate weakness. We'll try to explore the combination of them.

Add multi-user interaction. In the current stage Smart Classroom, there is only one user (lecturer) naturally interacting with the Intelligent Environment Other attendants are just observers or listeners and are not able to exploit the fascinating features of natural interaction. Because a class is bound to have multiple participants, to make a qualified Smart Classroom, we need to enhance the classroom's support for multi-user interaction. In our next step, we'll empower the classroom with capability to track and identify more than one user dynamically, and enable Smart Classroom's in-place service to every user in the room.

Reference

1. [Winograd 1999]. Winograd, Terry. Toward a Human-Centered Interaction Architecture. 1999. <http://www.graphics.stanford.edu/projects/iwork/papers/humcent/index.html>
2. [Weiser 1991]. Weiser, Mark. The Computer for the 21st Century. Scientific American. pp. 94-10, September 1991. <http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html>
3. [Weiser 1993]. Weiser, Mark. Ubiquitous Computing. IEEE Computer "Hot Topics", October 1993. <http://www.ubiq.com/hypertext/weiser/UbiCompHotTopics.html>
4. [Weiser 1994]. Mark Weiser. The world is not a desktop. Interactions, pages 7-8, January 1994. <http://www.ubiq.com/hypertext/weiser/ACMInteractions2.html>
5. [Coen 1999]. Coen, Michael. The Future Of Human-Computer Interaction or How I learned to stop worrying and love My Intelligent Room. IEEE Intelligent Systems. March/April. 1999.
6. [HAL 2000] MIT AI Lab HAL Project (previous Intelligent Room project), 2000. <http://www.ai.mit.edu/projects/hal>

7. [Mozer 1999]. Mozer, Michael C. An intelligent environment must be adaptive. IEEE Intelligent Systems. Mar/Apr. 1999
8. [eClass 2000] Georgia Tech, eClass Project (previous Classroom 2000) 2000 <http://www.cc.gatech.edu/fce/eClass/>
9. [Abowd 2000]. Abowd, Gregory. Classroom 2000: An experiment with the instrumentation of a living educational environment. IBM Systems Journal, Vol. 38. No.4
10. [Interactive Workspaces 2000]. Stanford Interactive Workspaces Project. <http://graphics.stanford.edu/projects/iwork>
11. [Fox 2000] Fox, Armando, et al. Integrating Information Appliances into an Interactive Workspace, IEEE CG&A, May/June 2000
12. [Aware Home 1999]. Georgia Tech, Aware Home Project,1999 <http://www.cc.gatech.edu/fce/house>
13. [Kidd 1999]. Kidd, Cory D., Robert J. Orr, Gregory D. Abowd, et al. The Aware Home: A Living Laboratory for Ubiquitous Computing Research. In the Proceedings of the Second International Workshop on Cooperative Buildings - [CoBuild'99](#), Position paper, October 1999.
14. [DreamSpace] IBM Research. <http://www.research.ibm.com/natural/dreamspace/index.html>
15. [EasyLiving] Microsoft Research. <http://www.research.microsoft.com/vision>
16. [OAA]. SRI. <http://www.ai.sri.com/~oaa>
17. [Johnston 1998]. Johnston, Michael. Unification-based multimodal parsing. In the Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics (COLINGACL 98), August 98, ACL Press, 624-630.
18. [Ren 2000]. Ren, Haibing et al. Spatio-temporal appearance modeling and recognition of continuous dynamic hand gestures. Chinese Journal of Computers (in Chinese), 1999. Vol 23, No. 8, Aug.2000, pp.824-828.
19. [Pei 1999]. Pei, Yunzhang, Liu, Yan, Shi, Yuanchun, Xu, Guangyou. Totally Ordered Reliable Multicast for Whiteboard Application. In proceedings of the 4th International Workshop on CSCW in Design, Paris, France, 1999.
20. [Tan 2000]. Tan, Kun, Shi, Yuanchun, Xu, Guangyou. A practical semantic reliable multicast architecture. In proceedings of the third international conference on multimodal interfaces, BeiJing, China, 2000
21. [Liao 2000]. Liao, Chunyuan, Shi, Yuanchun, Xu, Guangyou. AMTM – An Adaptive Multimedia Transport Model. In proceeding of SPIE International Symposia on Voice, Video and Data Communication, Boston, Nov. 5-8, 2000.