I. Cover Page

Ubiquitous Smart Spaces

Topic Area: Smart Spaces
Gregory Abowd, Chris Atkeson, Irfan Essa
404 894 6856, 404 894 0673 (Fax)
abowd@cc.gatech,edu, cga@cc.gatech.edu, irfan@cc.gatech.edu
Georgia Institute of Technology, College of Computing
801 Atlantic Drive, Atlanta, GA 30332-0280

II.   Ubiquitous Smart Spaces

II.1   Innovative Capability Envisioned

The next revolutionary advance in smart spaces research is to create and experiment with ubiquitous smart spaces. Current research has created isolated smart rooms and enhanced individuals. This situation is analogous to having one telephone in each small town. The power of the telephone and the way the telephone system was used changed dramatically when there were telephones in almost every room and individuals could carry portable telephones. Similarly, users of smart spaces won't have to delay, interrupt, or restructure their activities to take advantage of a central smart room facility if every space is smart. The visionary application that will motivate and drive a coordinated effort by the research community is to create ubiquitous smart spaces: demonstrations of smart spaces that encompass entire working communities, and cover all aspects of each participant's life.

We specifically propose creating a ubiquitous smart space that encompasses the College of Computing at Georgia Tech. The details of this proposal are given in [1] . This demonstration will provide a place to explore new user interface and usability concepts, as well as evaluate systems and networking infrastructure. This system will be challenging to build because of its large scale and because of the need for coverage of all aspects of user's lives. The system will extend into the homes and cars of users, as well as including personal mobile components.

II.2   Applications and Benefits

There is growing interest in using computing technologies to build systems that support our daily lives more effectively. Examples of such systems are smart offices, classrooms, and homes that allow computers to monitor what is happening in the environment, to figure out how to most effectively assist in the tasks being performed, and to have sufficient output devices to assist effectively. For example, a classroom might help students take more effective notes by capturing much of the details of a lecture, freeing the student to pay more attention to the content of the class and simply annotating or indexing into the record of the experience for later review. A meeting room could record the events and discussion in a meeting to allow others who did not attend the meeting to efficiently find out what happened. An office could remind its user of relevant information, or signal urgent messages. A machine room could provide documentation for operators or visiting repair technicians. A home could babysit children or allow elderly to remain self-sufficient. A building could supplement the capabilities of physically or mentally challenged individuals through visual assistance, movement assistance such as opening and closing doors, and providing and receiving information in alternate forms such as lip-reading and sign-language recognition. These systems can be built into the environment as well as being hand-held or even worn as part of our clothing.

Our ubiquitous smart space will provide several specific types of assistance for users:

- **Capturing everyday experiences:** The system will monitor and capture what happens in an environment.
- **Access to information:** "Remind me how to repair the interference problem on this piece of equipment." The system can assist users in rapidly accessing their own and others past experience, and reference material such as manuals and textbooks. This access will be ubiquitously available and will make effective use of all display functionality available in the space for interactive visualization of the information.

- **Communication and collaboration support:** Users will be able to see if someone is available for a virtual 'chat' with minimal effort. The system will monitor whether someone is interruptible at any given moment, and be able to provide information as to the location and status of other users, if they so desire. The system will provide subtle cues as to the status of incoming communications, and also monitor the urgency of messages or attempts to establish communication. We will provide a much more lightweight and flexible form of remote conferencing and long distance collaboration than the current use of special equipment in prepared locations.

- **Natural interfaces:** These systems will have perceptual abilities that will allow for effective, noninvasive, and ubiquitous interaction with the users. Cameras, microphones and touch sensors will be installed to track users and allow for interaction with the system using gesture, speech and touch. Users will be able to access information and manipulate it without being restricted to using a keyboard and/or a mouse.

- **Environmental awareness:** One particular type of environmental augmentation is to provide subtle cues, or peripheral awareness, that would allow one to know what activity is underway in another space. For example, a colleague might want to monitor activity around the coffee cart to determine when would be a good time to join other colleagues for a casual conversation. Similarly the environment can share information with the users if a remote user is interested in joining the conversation and/or the number of users participating remotely. Such aware environments will go beyond serving as surveillance devices and allow for more intelligent systems that could detect a pot boiling over on the stove or a person in need of medical assistance.

- **Automatic receptionist and tour guide:** The system will identify visitors, ask them who they want to see, and guide them. We have used early prototypes of our ubiquitous interface to provide tours and guidance to visitors during our open houses or visitor days. In addition, universities have a high turnover rate. Every fall we replace a quarter of our undergraduate population. Our system could help guide these newcomers.

- **Training:** The system will train users in various procedures, such as loading new paper into or repairing the copier, how to use all the facilities in an electronic classroom, or simply filling out forms and regular maintenance and repair of equipment.

Our goal is to support these types of tasks in offices, classrooms, common spaces, and hallways of the Georgia Tech College of Computing. We also intend to instrument the homes and cars of participants in this community. As part of this goal, the interaction with the computer must appear coherent to the user, even though the user is moving from location to location, the interaction is interrupted by other activities, and there are many other users interacting with the system at the same time in the near vicinity. The College of Computing is currently spread over space in four buildings, with the furthest building a mile away. How can our ubiquitous interface help combine physically distributed spaces so that communities (formal and informal, active and passive) still thrive?

Our major research strength is the development of smart spaces in which the environment monitors and captures what happens in that environment, and makes the information available for later access. We have used classrooms and meeting rooms to prototype smart space designs. The Classroom 2000 project [9] has as its goal to help students, teachers, and meeting participants function more effectively by augmenting a single room with the capabilities to automatically take notes on behalf of the participants [3, 8, 4]. The interface includes large-scale electronic whiteboards, portable electronic notebooks, microphones, and video cameras. Its purpose is to allow a class or meeting to proceed

as normal, with the end result being an automatically generated record that allows a participant to review the entire experience or efficiently browse and navigate to points of interest. By viewing the classroom or meeting experience as a multimedia authoring session, we simplify content generation for educational software, training, and information dissemination.

Our CyberGuide [20, 21, 5] and CyberDesk [27, 12] projects explore moving the interface to the user by providing context-aware applications. We have also defined the topic of dual augmentation of physical and virtual worlds and demonstrated this concept through the Domisilica project [22, 23, 24].

II.3   Major Technical Obstacles and Our Approaches

In this section we outline the *major technical obstacles* that we will encounter in our research and our approaches to overcome these.

Exploring new paradigms for human-computer interaction through ubiquitous computing requires full-scale prototypes. This has been a serious problem for research on ubiquitous smart spaces, which require large areas to be instrumented. Current state of the art is to instrument a single room or a dedicated area in an office environment. Like the telephone system or the road network, the services provided in this prototype interface must be available throughout a user's workspace in order for us to fully reveal how it will be used in the future and to identify the scientific issues raised. Consider how computer use changed as we went from using terminal rooms to laptops. We are working on developing a ubiquitous smart space that will encompass an entire community, the College of Computing at Georgia Tech.

The greatest technical obstacle to ubiquitous smart spaces is developing natural user interfaces and interaction paradigms. Our approach is to explore two types of interfaces. The first type of interface attempts to be as invisible as possible. As computation becomes more pervasive in the user's environment, the intent is to make the actual interface less obtrusive, to the point where the physical interface becomes transparent to the user. User's activities are changed or disrupted as little as possible. The second approach attempts to create the illusion that there are additional human-like observers and actors in the environment, and the actions of the smart space system can be explained and understood in human terms.

We need to make significant progress in empowering people, places, and objects. This progress must be balanced to obtain potential synergies. We are working on both instrumented rooms [9] and mobile and wearable systems [11].

Reasonable access methods must be provided to the ongoing and stored information flood. How do we support and augment our human ability to segment, structure, index, link, and navigate captured experiences? One approach we are using is to apply statistical methods for natural language understanding [6]. Statistical approaches are usually based on matching word frequencies (vocabularies) in written or spoken text. The key insight in many statistical approaches to language is that "concepts" can be represented by vocabularies. Using vocabularies enables concepts to emerge from the data, and allows us to avoid the tedious work of categorization by hand. Pieces of text which have similar vocabularies are probably about related topics, and are probably relevant to each other. Statistical approaches can: answer more general queries about captured information in a ubiquitous computing system, including trying to identify the concept the user is searching for rather than the word; automatically segment and structure captured material; automatically link captured material to other material such as textbooks, briefing materials, or web pages; and automatically find keywords or make summaries or abstracts, to aid human navigation of large amounts of material.

Often critical information from the actual experience is lost, making it difficult to understand what was captured ("Which of the symbols in this equation was the professor referring to when she said 'take this term and multiply it by that term to complete the proof.' *"?)* Computational perception attempts to understand the activity and the actions of people in an environment ("what is happening?"). This requires the development of algorithms for recognizing who is in the environment, and what they are doing. Users can be identified by face recognition and/or speaker identification ("who is in the space?"). After identifying the user(s), we pursue automatic interpretation of gaze and body orientation, speech attributes, facial expressions, gesture, and body movements, from the information contained in the video and audio streams. We expect the combined information from these visual and aural modalities will allow understanding, coding, and re-synthesis of signals in many situations where information from only one of the modalities would not be sufficient for understanding. This level of information would be very useful in determining the focus of attention of the user and/or the audience in an environment such as a classroom. Additionally, this information will aid in the interpretation of the communicative aspects of speech, gesture and body language, allowing for a more natural interaction between the user and the system. This type of work will build on our previous work on machine perception for intelligent interaction [10]. Specific papers addressing this work are [15, 14, 16, 13, 17, 7, 25, 26, 19].

System design and management techniques must be developed for smart spaces built out of thousands of interacting systems. Much of our research has focused on rapid prototyping of ubiquitous computing applications to support everyday activities. Because we have forced ourselves to use our own systems on an everyday basis, we have been faced with significant software engineering issues. We have identified three general functional characteristics of a number of ubiquitous computing applications —capture, integration, and access; context-aware computing; and dual augmentation [22, 23]. To support rapid prototyping, we need to provide toolkit-level support for each of these functional themes. We are developing a general capture, integration, and access framework to support a wide variety of applications.

Smart spaces need to be aware of user context. Interoperability is a common research theme for the engineering of complex distributed systems. Solutions to this general problem can be found in networking and systems-layer research. Programming in the large has benefited from dynamic integration to allow for flexible negotiation between system components at run-time. Our CyberDesk project shows the importance and usefulness of dynamic integration of computational services [27, 12]. Dynamic integration is beneficial to the programmer because there is no need to know in advance all possible contexts of use when a service is designed. It is beneficial to the user because it allows for continuous switching and selection of different computational services. As we move to a larger ubiquitous computing testbed, the variety of different services that students and faculty will want at their fingertips will require more attention to facilitation of dynamic integration.

Smart spaces need to support a wide range of interfaces. A challenge for ubiquitous computing is to develop applications that can automatically adjust, or scale, to very different devices. This application interoperability problem is referred to as the scalable interfaces problem [2], which we are addressing in a number of projects [18].

Yet another consequence of interoperability requirements is that a working ubiquitous computing system consists of a heterogeneous collection of devices with numerous failure modes. Designing a system to be robust under the inevitable threat of failures is a challenge we are currently addressing.

Section III lists key citations to our past and ongoing work in this area.

4

III. Additional Information

References

[1] Gregory Abowd, Chris Atkeson, Irfan Essa, and Kishore Ramachandran. Building a large-scale ubiquitous computing interface, 1998. A proposal submitted to the National Science Foundation Major Research Instrumentation Program, `www.cc.gatech.edu/fce/pubs/BaLSUCI.pdf`.

[2] Gregory D. Abowd. Ubiquitous computing: Research themes and open issues from an applications perspective. Technical Report GIT-GVU 96-24, GVU Center, Georgia Institute of Technology, October 1996. Available from `www.cc.gatech.edu/fce/publications.html`.

[3] Gregory D. Abowd, Christopher G. Atkeson, Jason Brotherton, Tommy Enqvist, Paul Gulley, and Johan LeMon. Investigating the capture, integration and access problem of ubiquitous computing in an educational setting. In *Proceedings of the 1998 conference on Human Factors in Computing Systems — CHI'98*, 1998. `www.cc.gatech.edu/fce/c2000/pubs/`.

[4] Gregory D. Abowd, Christopher G. Atkeson, Ami Feinstein, Cindy Hmelo, Rob Kooper, Sue Long, Nick Sawhney, and Mikiya Tani. Teaching and learning as multimedia authoring: The Classroom 2000 project. In *Proceedings of the ACM Conference on Multimedia — Multimedia'96*, pages 187–198, 1996. Available from `www.cc.gatech.edu/fce/c2000/pubs/`.

[5] Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. Cyberguide: A mobile context-aware tour guide. *ACM Wireless Networks*, 3, 1997.

[6] Gary Boone. Concept features in Re:Agent, an intelligent email agent. In *The Second International Conference on Autonomous Agents (Agents '98)*, 1998. `www.cc.gatech.edu/grads/b/Gary.N.Boone/papers/reagent.ps.Z`.

[7] M. Brand and I. Essa. Causal analysis for visual gesture understanding. In *Proceedings of AAAI Fall Symposium on Computational Models for Integrating Language and Vision*, 1995. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[8] Jason A. Brotherton and Gregory D. Abowd. Rooms take note! Room takes notes! In *Proceedings of the 1998 Spring AAAI Symposium on Intelligent Environments*, 1998. `www.cc.gatech.edu/fce/c2000/pubs/`.

[9] Classroom 2000. `www.cc.gatech.edu/fce/c2000`.

[10] Computational Perception Laboratory at Georgia Tech. `www.cc.gatech.edu/gvu/perception/`.

[11] Cyberguide. `www.cc.gatech.edu/fce/cyberguide`.

[12] Anind Dey, Gregory D. Abowd, and Andy Wood. Cyberdesk: A framework for providing self-integrating context-aware services. In *Proceedings of the 1998 Intelligent User Interfaces Conference — IUI'98*, pages 47–54, 1998. `www.cc.gatech.edu/fce/cyberdesk`.

[13] I. Essa, S. Basu, T. Darrell, and A. Pentland. Modeling, tracking and interactive animation of faces and heads using input from video. In *Proceedings of Computer Animation Conference 1996*, pages 68–79. IEEE Computer Society, June 1996. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[14] I. Essa, T. Darrell, and A. Pentland. Tracking facial motion. In *Proceedings of the Workshop on Motion of Nonrigid and Articulated Objects*, pages 36–42. IEEE Computer Society, 1994. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[15] I. Essa and A. Pentland. A vision system for observing and extracting facial action parameters. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 76–83. IEEE Computer Society, 1994. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[16] I. Essa and A. Pentland. Facial expression recognition using a dynamic model and motion energy. In *Proceedings of the International Conference on Computer Vision*, pages 360–367. IEEE Computer Society, Cambridge, MA, 1995. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[17] I. Essa and A. Pentland. Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 19(7):757–763, 1997. An earlier version available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[18] Future Computing Environments. `www.cc.gatech.edu/fce/cyberguide`.

[19] A. Gardner and I. Essa. Prosody analysis for speaker affect determination. In *Proceedings of Perceptual User Interfaces Workshop (In Conjunction with UIST 1997 Conference)*, Banff, Canada, October 1997. Available from `www.cc.gatech.edu/fac/irfan/pubs/`.

[20] Sue Long, Dietmar Aust, Gregory D. Abowd, and Christopher G. Atkeson. Rapid prototyping of mobile context-aware applications: The cyberguide case study. In *Proceedings of the 1996 conference on Human Factors in Computing Systems — CHI'96*, pages 293–294, 1996. Available from `www.cc.gatech.edu/fce/publications.html`.

[21] Sue Long, Rob Kooper, Gregory D. Abowd, and Christopher G. Atkeson. Rapid prototyping of mobile context-aware applications: The cyberguide case study. In *Proceedings of the 2nd Annual International Conference on Mobile Computing and Networking*, November 1996. `www.cc.gatech.edu/fce/publications.html`.

[22] Jennifer Mankoff and Gregory D. Abowd. Domisilica: Providing ubiquitous access to the home. Technical Report GIT-GVU-97-17, GVU Center, Georgia Institute of Technology, September 1997. Available from `www.cc.gatech.edu/fce/publications.html`.

[23] Jennifer Mankoff, Jonathan Somers, and Gregory D. Abowd. Bringing people and places together. In *Proceedings of the 1998 Spring AAAI Symposium on Intelligent Environments*, 1998. Available from `www.cc.gatech.edu/fce/publications.html`.

[24] Jennifer Mankoff, Jonathan Somers, and Gregory D. Abowd. Bringing people and places together with dual augmentation. Paper submitted for review to conference on Cooperative Virtual Environments, 1998. Available from `www.cc.gatech.edu/fce/publications.html`.

[25] A. Pentland. Smart rooms. *Scientific American*, 274(4):68–76, April 1996. Available at `http://www.sciam.com/0496issue/0496pentland.html`.

[26] A. Pentland, I. Essa, T. Darrell, A. Azarbayejani, and S. Sclaroff. Visually guided animation. In N. Thalmann and D. Thalmann, editors, *Interactive Computer Animation*, pages 143–164. Prentice Hall, 1996.

[27] Andy Wood, Anind Dey, and Gregory D. Abowd. Cyberdesk: Automated integration of desktop and network services. In *Proceedings of the 1997 conference on Human Factors in Computing Systems — CHI'97*, pages 552–553, 1997. `www.cc.gatech.edu/fce/cyberdesk`.