Universität Karlsruhe - Fakultät für Informatik - Bibliothek - Postfach 6980 - 76128 Karlsruhe

Situated Interaction in Ubiquitous Computing

Albrecht Schmidt (Hrsg.)

Interner Bericht 2000-7

Universität Karlsruhe

Fakultät für Informatik

ISSN 1432 - 7864

CHI 2000
THE FUTURE IS HERE

TecO

Preprints

Situated Interaction in Ubiquitous Computing

Workshop at the CHI 2000 Conference on Human Factors
The Hague, The Netherlands,
3. April 2000.

Albrecht Schmidt, TecO
Technical Report, Department of Computer Science, University of Karlsruhe, Germany

Table of contents

Page

Situated Interaction in Ubiquitous Computing

Albrecht Schmidt
Telecooperation Office (TecO)
University of Karlsruhe
Vincenz-Priessnitz-Str. 1
76131 Karlsruhe, Germany
+49 721 6902 29
albrecht@teco.edu

Walter Van de Velde
Starlab Nv/Sa,
Excelsiorlaan 40-42
1930 Zaventem, Belgium
+32 2 721 54 54
wvdv@starlab.net

Gerd Kortuem
Computer Science Department
University of Oregon
Eugene, OR 97403, USA
+1 541 346 1381
kortuem@cs.uoregon.edu

Keywords
Context awareness, information appliances, situated interaction, ubiquitous computing, wearable computing.

INTRODUCTION

The situations in which human-computer interaction takes place are increasingly varied, as computers become highly portable and embedded in everyday environments. Research reported from different communities (Wearable Computing, Mobile Computing, HCI, CSCW, Augmented Reality) indicates that awareness of situations can lead to improvement of human-computer interaction. We propose a workshop at CHI2000 to provide a forum to discuss situational awareness and situated interaction.

With the availability of sensing technologies, such as measuring the surrounding light conditions, the motion of the user, the orientation of a display, users' position relative to an information appliance, the number of users in front of a device, users' emotional state (bio-sensors), etc., this situational context can be captured and used as additional input to the system. The interaction process can benefit from the additional knowledge about the situation [1].

THEMES AND TOPICS

In the design and development of interactive devices that are used in various environments and different situations the classical interaction paradigms as know from desktop applications are often not appropriate. The situation and the surrounding environment provide information that should be taken into account by the applications as additional input. In the workshop we explore the impact of situational awareness on human computer interaction. We will focus especially on the following topics:

- Automated extraction and identification of the situational context from sensory data.
- Adaptation of input and output modalities according to usage situation and recognized requirements [2,3].
- Reduction in the need for explicit user input through automated information capture and reasoning.
- Determination the right time and mode for interrupting the user appropriate to the situation [4].

Situational awareness is especially attractive for two classes of devices, see Fig 1. First ultra-mobile devices, characterized by the fact that these devices are operated and operational while on the move, most notably: wearable computers, smart mobile phone, handheld computers and PDAs. Second for shared stationary devices in common spaces or in changing environments e.g. public displays, meeting rooms, shared offices.

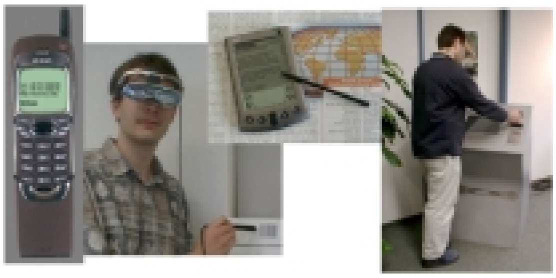


Fig.1: Mobile and stationary devices.

With currently available technology different mechanism are given to detect situational context, namely: sensors for physical, chemical, and biological parameters, tools for capture and analysis of video and audio data.

GOALS

The main goal of the workshop is to develop an understanding of how the situation of use does influence the interaction process. We will provide a forum to share information, results, and ideas on current research in the area of context awareness and situated computing with its respect to human computer interaction. Furthermore we aim to develop new ideas on how to exploit context for improving human computer interaction and to identify possible obstacles on that way. Results will be available at http://www.teco.edu/chi2000ws/.

References

1. Abowd, G.D., Atkeson, C.G, Brotherton, J.A., Enqvist, T., Gully, P., Lemon, J. Investigating the capture, integration and access problem of ubiquitous computing in an educational setting. In *Proceedings of CHI '98,* May, (1998).
2. Kortuem, G., Segall, Z., Bauer, M. Context-Aware Adaptive Wearable Computers as Remote Interface to Intelligent Environments. *Proceedings of the 2nd International Symposium on Wearable Computers*. Pittsburgh. 1998.
3. Schmidt, A., Aidoo, K. A., Takaluoma, A., Tuomela, U., Van Laerhoven, K., Van de Velde, W.. Advanced Interaction in Context. *1st International Symposium on Handheld and Ubiquitous Computing (HUC99)*, Germany, Sept. 1999.
4. Sawhney, N., Schmandt, C. Nomadic Radio: Scaleable and Contextual Notification for Wearable Audio Messaging. *In Proceedings of the 1999 Conference on Human Factors in Computing Systems (CHI '99),* Pittsburgh, PA, May 15-20, 1999.

To appear in the *2000 Conference on Human Factors in Computer Systems (CHI 2000) Extended Abstracts*

Active Environments: Sensing and Responding to Groups of People

Joseph F. McCarthy
Center for Strategic Technology Research
Andersen Consulting
3773 Willow Road
Northbrook, IL 60062 USA
mccarthy@cstar.ac.com

ABSTRACT

Most environments are *passive* – deaf, dumb and blind, unaware of their inhabitants and unable to assist them in a meaningful way. However, with the advent of ubiquitous computing – ever smaller, cheaper and faster computational devices embedded in a growing variety of "smart" objects – it is becoming increasingly possible to create *active* environments: physical spaces that can sense and respond appropriately to the people and activities taking place within them. Most of the early UbiComp applications focus on how *individuals* interact with their environments as they work on *foreground* tasks. In contrast, this paper focuses on how *groups* of people affect and are affected by *background* aspects of their environments.

Keywords

Ubiquitous computing, intelligent environments, social issues, computer supported cooperative work.

INTRODUCTION

Computers are becoming smaller and cheaper, connectivity is expanding in both the wired and unwired domains, and the growth of digital content is outpacing the capabilities of current indexing systems. These trends are ushering in an era of *ubiquitous computing* [5], in which we have computing and communication capabilities available in all kinds of environments and situations beyond the "traditional" model of sitting at a desktop computer workstation. Under this new model of computing, human-computer interaction issues evolve into issues of *inhabitant-environment interaction*. Many of the early applications of ubiquitous computing focus on how these new capabilities will affect an *individual*'s interactions with his or her environment. However, much of our time is spent in shared physical spaces, so it is important to consider how an environment might effectively sense and respond to *groups* of co-located people. This paper raises a number of issues for what might be called "UbiGroup" applications, and describes research that address some of these issues.

Many of the computer applications we are familiar with today are used in the *foreground* of our attention, e.g., checking stock quotes on our web-enabled telephone or using a word processor application on a laptop computer to create a short conference paper. However, as computers become smaller and cheaper, they will increasingly be embedded in a variety of objects that do not typically require, and often do not permit, our focused attention. Such applications will operate in the *background*, at the periphery of our attention; they will affect aspects of our environments, but not necessarily assist us directly with the task(s) at hand.

One example of this type of application is MUSICFX [4], a system embedded in a fitness center environment that is aware of who is working out and what they generally like to listen to, and uses this knowledge to determine the best music to play at any given time. While MUSICFX affects aural aspects of a group environment, another project seeks to affect visual aspects: the Projected Realities proposal [3] calls for pictures or artwork that would be projected on large public displays to reflect the mood of the local population. Finally, Sunset [1] affects both the visual and aural aspects of a group environment, creating a "drive-by interactive drama" in which a large public display shows a vignette – a sequence of pictures accompanied by a soundtrack of "insinuating muzak" – that is influenced by the number of passersby (or loiterers) pressing buttons on their keyfobs and garage door openers.

In general, these and other new environment-affecting applications must be able to sense their contexts, determine the preferences and goals of their inhabitants, and respond appropriately. The remainder of this paper discusses these functions in more detail, and highlights issues that arise in the ways that various applications implement them.

CONTEXT SENSING

The context for the Sunset project consisted of people passing by the Billboard Live club on Hollywood's Sunset Boulevard who were detected by monitoring for transmissions from radio key fobs and garage door openers. The Projected Realities proposal includes a number of potential environmental sensing capabilities, e.g., "ventrovers" that could listen and peak into rooms in an apartment complex and a network of linked security cameras and monitors, although the longer range context was inferred by other means (described in the next section).

Unfortunately, in many contexts, detecting *that* one or more people are present is often not as useful as identifying *which* people are present. Both the Sunset and Projected Realities projects could be extended in interesting ways were they to include the capability to identify the current set of inhabitants. In the fitness center environment affected by MUSICFX, the music is tailored to the preferences of the specific people working out, not to the number of unidentified people. In a workplace environment, temperature and/or lighting levels might respond to the number of people present, but adapting to the preferences of those specific people would likely result in a more hospitable environment [2].

Although some progress has been made in computer vision and speech recognition systems for identifying the faces or voices of different people in a room, many systems, including MUSICFX, that rely upon the accurate identification of different people in a physical space utilize some kind of badge or tagging system. Greater progress in the area of automatic identification is one factor that would enable more widespread deployment of environment-affecting applications, although the privacy concerns of people who might be identified in an increasing number of contexts would have to be addressed.

PREFERENCE AND GOAL DETERMINATION

Once an environment becomes aware of whom its inhabitants are, the next step is to associate preferences or goals with these inhabitants. This might be done *explicitly*, by querying inhabitants about their preferences or goals, or it might be done *implicitly*, by inferring these from the environmental context and/or some observable actions taken by the inhabitants. In either case, these preference or goal determinations might be done in one or more different times and/or places than the environmental context in which the response system is embedded.

Nothing explicit was known about the particular preferences or goals of people passing the Sunset displays. However, it was assumed that passersby would prefer viewing (and participating in) potentially engaging entertainment rather than blank screens. The Projected Realities project collected responses to "cultural probes" that included pictures and or textual material, from which inferences were made about the concerns and activities of the inhabitants of an apartment complex. In MUSICFX, explicit determination of musical preferences was accomplished by asking fitness center members to fill out an electronic enrollment form to specify one of five levels of preference for each of 91 genres of music.

An implicit determination of musical preferences could be accomplished by tracking a person's purchases at an on-line music store or tracking which Internet radio stations a person listens to. However, privacy concerns may dampen people's enthusiasm for participating in such a system. In general, there is a tradeoff between implicitly inferring preferences, which may be inaccurate and perceived as invasive, and explicit requests for preference information, which may be burdensome and imposing (and therefore not used by some/most inhabitants).

RESPONDING APPROPRIATELY

Once an environment has some knowledge (or, at least, presumption) of the preferences and goals of its inhabitants, it can adapt itself in response. Sunset responded to the activity of its passersby by altering the pacing, segues and selection of its vignettes. In Projected Realities, various public displays would respond to the inferred collective mood of the residents by projecting images representing that mood. MUSICFX responds to the musical preferences of the fitness center inhabitants by choosing music that is likely to please them.

CONCLUSION

Two themes common to all these applications is their focus on environmental aspects on the periphery of inhabitants' attention, and their responsiveness to the group of inhabitants rather than single individuals. As computers permeate more aspects of our environments, we expect to see new capabilities for sensing and responding appropriately to inhabitants, which will provide for more engaging, entertaining and hospitable environments in a broader variety of contexts.

REFERENCES

1. Crane, M., MacDonald, D., Minneman, S., Winet, J., and Neuburg, A., Sunset Boulevard: 200 MHz in a 35 mph Zone, exhibited at the *SIGGRAPH '97 Electronic Garden*, http://www.pair.xerox.com/cw/sunset/.
2. Elrod, S., Hall, G., Costanza, R., Dixon, M., and Des Rivieres, J., Responsive Office Environments, *Communications of the ACM* 36(7), 1993, pp. 84-85.
3. Gaver, W., and Dunne, A., Projected Realities: Conceptual Design for Cultural Effect, in *CHI '99*, pp. 600-607.
4. McCarthy, J. F., and Anagnost, T. D., MUSICFX: An Arbiter of Group Preferences for Computer Supported Collaborative Workouts, in *CSCW '98*, pp. 363-372.
5. Weiser, M., The Computer for the Twenty-First Century, *Scientific American*, September 1991, pp. 94-100.

Teaching Contexts to Applications

Kristof Van Laerhoven
Starlab Research NV/SA
St-Michielslaan 47
1040 Brussels
kristof@starlab.net

Abstract

Enhancing applications by adding one or more sensors is not new. Incorporating machine-learning techniques to fuse the data from the sensors into a high-level context description is less obvious. This paper describes an architecture that combines a hierarchy of self-organizing networks and a Markov chain to enable on-line context recognition. Depending on both user and application, the user can teach a context description to the system whenever he or she likes to, as long as the behavior of the sensors is different enough. Finally, consequences and complications of this new approach are discussed.

1 INTRODUCTION

1.1 CONTEXTS AND APPLICATIONS

Some applications enhance their user interface by adding a sensor and using the sensors' value in some simple rule. A typical example is connecting a light sensor to a screen-based device and adjusting the contrast and brightness of the screen according to the value of the light sensor.

Other applications can change their behavior only when the user explicitly tells them to. It is also possible to use user-defined profiles that describe the devices' behavior. Profiles in mobile phones, for example, can be set to make the phone ring very loud outside or on the train, but only vibrate in a meeting. This approach leads to a lot of user-involvement, though: the user first needs to program these profiles, and then these profiles must be set in every context ('in a meeting', 'in the train', …).

The combination of all of the approaches mentioned earlier leads to an *automated profiles selection*: context recognition based on simple sensors sets the behavior of the device (see [1] and [5]). Knowing the context usually leads to being able to improve the application and particularly enhancing the interaction with the user. This approach is far from simple, however: how can a device, equipped with sensors, recognize a context?

1.2 CONTEXT

The notion of context is very broad and incorporates lots of information, not just about the current location, but also about the current activity, or even the inner state of the person describing it. As a consequence, multiple people can describe their contexts in different ways, even if they are in the same location doing the same thing. Someone familiar with a building might know a room as 'classroom 402B', while a visitor would probably describe it as just 'a classroom'.

In addition, the application defines the description of the context as well. Some applications require more location-based contexts, while others need contexts that give more information about the user.

1.3 CONTEXT DESCRIPTION

The simplest method for giving a context description would be to sum up all the values from the sensors to a formatted description, like for instance "movement: (87%, 29%), light: 78%, humidity: 69%, temperature: 50%, …". A simple, rule-based architecture could be used to enhance this description into "moving slowly in a cold, humid, well-lit room".

The architecture described here works the opposite way: the system merges the output from the sensors and maps them to a description given by the user. The description could then be something like "walking in the basement". This way, we deal with the fact that context description depends on both the individual describing it and the application using context perception.

2 ONLINE ADAPTIVE CONTEXT AWARENESS

This section will describe the architecture of the algorithm that does on-line training and recognition.

Instead of just using the raw sensor values as input for the next layer, small pre-processing routines were chosen to enhance the future clustering. For example, instead of just looking at the brightness of the light, it is also possible to look at its frequency, which results in easier distinguishing of several types of artificial light. Taking the standard deviation of the accelerometer values can also give more qualitative information. Other sensors like microphones and infrared sensors have similar mini-transformations from the raw sensor data to one, but usually multiple, values, which are usually called cues or features.

Another advantage of the cues is that that they are sent less frequently to the next layer. The light sensor, for instance, is read a few hundred times per second. The cues from this sensor (light level and frequency) are sent every second. Cues are very significant for a fast, but accurate context recognition system. However, using cues results in a large input dimension, which makes the mapping-algorithm very slow in learning. This difficulty arises when many irrelevant inputs are present and is usually referred to as *the curse of dimensionality* (see [4]).

2.1 SELF-ORGANIZATION

When a rat has learned its location in a labyrinth, certain braincells on the hippocampal cortex respond only when it is in a particular location. Self-organization of neuronal functions seems to exist on very abstract levels (like geographic environments) in the brain. The Kohonen Self-Organizing Map (SOM) has a similar principle: neurons (artificial, this time) are recruited topologically for tasks depending on the (sensory) input. The SOM is also known to handle noisy data relatively well, which makes it a sensible choice for clustering the inputs.

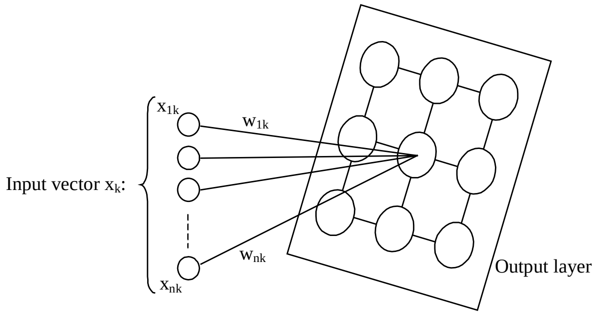


Figure 1: The Kohonen Self-Organizing Map

The Kohonen SOM (1982) [3] is based on earlier work of Willshaw and von der Malsburg [6] (1976), and starts with a competitive network, where basic units 'compete' for a particular kind of input. For every input, one unit is selected to be the winner and can adapt itself a bit more towards this input. More concrete, the winner can adjust an internal weight vector (or codebook vector, or prototype vector) towards the input vector (See Figure 1). Therefore, different sensor inputs result in different neurons being activated on the SOM. It is possible to monitor the activation of the neurons and plot the resulting matrix as a landscape, where different hills ideally represent different contexts (see Figure 2).

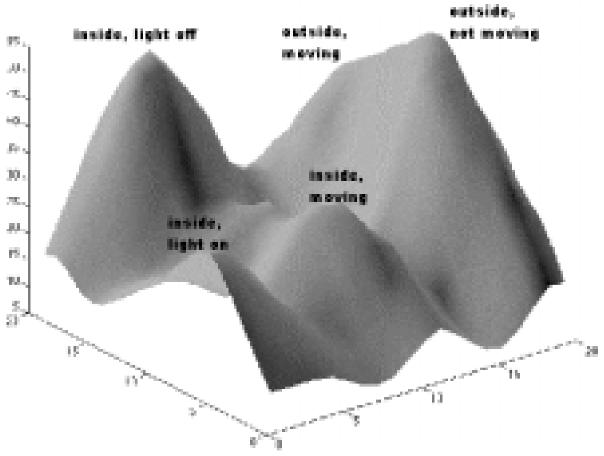


Figure 2: Example of an activity-plot of a SOM.

The traditional algorithm starts out highly adaptive (a large learning rate and huge neighborhood radius) and gradually becomes fixed. After this stage, it is not capable of learning anymore, which poses an obstacle if the system needs to remain adaptive. This is a problem also known as *the stability-plasticity dilemma*. It therefore is necessary to add some mechanism that controls the flexibility of the SOM locally, i.e. on the level of the neurons.

The fixed size of the map is a problem as well, since it leads to a limit of contexts that can be recognized. To deal with this problem, it is possible to use multiple (smaller) SOMs with each a subset of the original input vector instead of one big SOM. Since it is very likely that some contexts will persist much longer than others, the possibility exists that popular (long-lasting) contexts 'overwrite' others, i.e. undo the weight adaptations from the other contexts' signals. It is therefore advantageous to pre-order the weights in a controlled environment before the real use instead of just giving them small random values as is customarily done with the SOM.

2.4 SUPERVISION AND USER BEHAVIOR

The next layer is primarily intended to supervise transitions from one context to another. It uses a probabilistic finite state machine architecture where each context is represented by a state, and transitions are represented by edges between states. The model keeps a probability measure for each transition, so every time a transition occurs, the supervision model can check if this really is likely. If a transition is not really probable, the next state is not entered yet, but a buffer mechanism is initiated so that it does become more likely after several tries in a row. Each transition to a state is thus dependent on the previous state, which makes this model a first-order Markov model. Every state also keeps track of how much time was spend in a particular context, which controls the flexibility of the SOMs: the newer a context, the more flexible and adaptive the map should be.

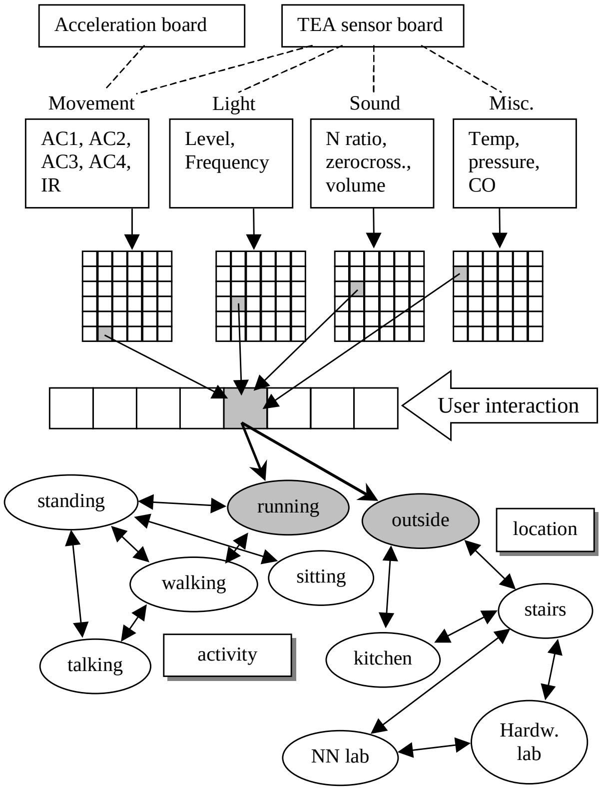


Figure 3: Overall architecture. User interaction is only necessary after the clustering.

The result is that after some time this model generates a graph depicting the behavior of a user with relation to the contexts visited. When the user tends to go from A to B rather than to C, then this will be reflected in the graph's connection strengths. Figure 3 depicts the typical layout of the final architecture.

3 CONCLUSIONS AND FUTURE WORK

But adaptive context perception not just solves problems; it introduces new ones too. For an application (using context-perception) to be effectively user-friendly, it is necessary that the system gets feedback from the user whenever the *user* would like to. These constraints are both hard and challenging from a machine-learning point of view. The combination of unsupervised neural networks and a context model gives promising results, without creating a bulky overhead on the user-computer interaction. However, the performance must be boosted by improving both sensors and cues in both quality and quantity. The experiments up until now used 10 sensors, but we expect to increase this number significantly. Other important issues are placement of sensors, the grouping of sensors for the clustering, and redundancy of sensors to make the system truly robust.

Acknowledgements

The framework of this paper was given by the TEA project [2], which is sponsored by the European Commissions 'Fourth Framework'. Thanks go out to the people from all project-partners: Starlab Research (Belgium), Nokia Mobile Phones (Finland), TecO (Germany) and Omega Generation (Italy).

References

[1] Chen, D., Schmidt, A. & Gellersen, H.-W. (1999) *An Architecture for Multi-Sensor Fusion in Mobile Environments*, In Proceedings International Conference on Information Fusion, Sunnyvale, CA, USA.

[2] Esprit Project 26900. Technology for Enabling Awareness (TEA). http://tea.starlab.net or http://www.omega.it/tea , 1999.

[3] Kohonen, T. (1997) *Self-Organizing Maps*, Springer Springer-Verlag Heidelberg,

[4] Mitchell, T.M.(1997) *Machine Learning*, McGraw-Hill.

[5] Schmidt, A., Aidoo, K.A., Takaluoma, A., Tuomela, U., Van Laerhoven, K. & Van de Velde, W. (1999) *Advanced Interaction in Context*, in H. Gellersen (Ed.) Handheld and Ubiquitous Computing, Lecture Notes in Computer Science No. 1707, Springer-Verlag Heidelberg, p. 89-101.

[6] von der Malsburg, C. (1991*) Self-Organization of Orientation Sensitive Cells in the Striate Cortex*, in G.A. Carpenter and S. Grossberg (Eds.) Pattern Recognition by Self-Organizing Neural Networks, MIT Press, p. 179-205.

Metaphors for Context-Aware Information Access

Joachim Baumann, Peter Coschurba, Uwe Kubach, Alexander Leonhardi

Institute of Parallel and Distributed High-Performance Systems (IPVR),
University of Stuttgart,
Breitwiesenstraße 20-22, 70565 Stuttgart, Germany
E-mail: <firstname>.<lastname>@informatik.uni-stuttgart.de
Phone: +49 711 7816 218, Fax: +49 711 7816 424

Abstract

In this paper we present a taxonomy for categorizing metaphors for context-aware information access. We give an overview of the existing generic metaphors and of implementations, which are classified according to the taxonomy.

1 Introduction

With the increasing pervasiveness of mobile computing devices the need for mobile information access grows continuously. A large variety of mobile information systems already exists, e.g. map / navigation systems and mobile guides (DAVIES 1998, ABOWD ET AL. 1997). In such systems the information access is context-dependent, i.e. the users' interest in a certain information item depends on their current context, e.g. on the geographical position they are located at. Therefore, many mobile information systems consider the user's context when displaying information mainly in order to present only information that really is of interest for the user. Until now, only a few metaphors have been suggested to make such a context-aware mobile information access more intuitive for the users (PASCOE 1997, LEONHARDI ET AL. 1999). However, we think that such metaphors are important to clarify, for the user, which functionality to expect from the system and the fact that some informations are presented and others are omitted.

In this paper we propose a taxonomy, which identifies the most important attributes of current and future metaphors for context-aware information access. Moreover, we use this taxonomy to classify existing metaphors, and provide generalized metaphors for some existing implementations.

The remainder of this paper is structured as follows: In Section 2 we introduce our taxonomy. Section 3 reflects the metaphors and in Section 4 these metaphors are classified according to the taxonomy. Finally, we conclude the paper in Section 5.

2 Taxonomy

By examining the existing metaphors for context-aware information access we have determined distinguishing properties that create a classification space for a taxonomy. We have discovered the following properties:

- *Observer.* A metaphor in this category is able to notice changes in its environment and to modify its internal state in reaction to these changes. An example for such a metaphor is a virtual photoelectric barrier, that notices whenever an object moves over a predefined line of sight.

- *Actor.* A metaphor belonging to this category is able to effect changes in its environment. An actor can e.g. send, in regular intervals, its own location to all objects in its vicinity, thus implementing a radio broadcast metaphor.
- *Container.* A container metaphor is able to hold objects. The scope of included objects (e.g. of the observer or actor category) might be limited to the container's extension. Consider the trash can metaphor: objects can be put into it or can be removed (by emptying the trash can, or by searching it), and the scope of all objects in the trash can is limited to the trash cans' extension.
- *Presenter.* A presenter is able to graphically display information in an augmented reality. A presenter might restrict its display, e.g. to be seen only from a certain direction, or up to a certain distance, or even with differing accuracy of the presented data depending e.g. on the distance. The most complex variant of this type uses a function describing the visibility depending on distance and direction while including visibility aspects of the real world. An example for this category is a virtual information tower metaphor (see Section 3) that might offer less detailed information to observers which are farther away than to those which are near it.
- *Associated*. Many of the metaphors we discuss have an association to their context. The following sub types can be distinguished:
    - *Time dependence*. The metaphor is associated with the time, i.e. changes as time passes. One example is a presenter that changes its display over time, e.g. as an advertisement-board with regularly changing displays. An example for a combination of time dependence and a presenter's visibility restrictions is the lighthouse metaphor. In this metaphor a beacon with a limited range, sending coded information, rotates slowly.
    - *Area of extent*. This sub type distinguishes the different possible areas of extent of a metaphor. It can be either a point (i.e. only a location, but infinitesimal size), an area or a function of the location. An example for a metaphor for which the extent is an area is the desktop metaphor, in which the desktop has a well-defined length and width.
    - *Attached*. Metaphors can be associated with a location or attached to an object. Attachment to another object implies that the attached objects' location depends solely on the location of the object to which it is attached. An example are virtual goods which are attached to a virtual train and are transported by it.
    - *Mobility.* Metaphors can be either static or mobile. Objects which are static are not necessarily immobile, they can e.g. be moved by attaching them to other objects.
- *Limited*. A metaphor of this category limits the access of objects. These metaphors are automatically also in the category *observer.* The access limitation can be of the following types:
    - *User-specific*. Access is only granted to specific users. One example metaphor is a virtual terminal.
    - *Location-specific*. Access is only granted to objects at a specific location or in a particular area.
    - *Time-dependent*. Access is only granted sometimes. An example is a virtual safe deposit box, which is only accessible during work hours.
    - *Complex*. Access is granted only depending on a function that might e.g. take into account time, location, and user.

3 Metaphors and Implementations

3.1 Generic Metaphors

Many existing location-aware systems use some kind of metaphor, very often taken from the real world, to illustrate their concept of interaction with location-aware information. Three of these metaphors are described in this section together with some metaphors which appear in related research areas.

One very important idea is the *Poster* metaphor. A poster has a fixed location (instead of being associated with an object) and is used to provide information to a larger audience. Newer, electronic variants of posters even allow to change the information on the fly and thus provide information which is more up to date. We believe this to be a very useful metaphor for location-aware application. An example for the use of this metaphor is a wall-mounted calendar. This calendar could even show different information depending on the person looking at it. Other applications of this metaphor would be virtual paintings, or virtual windows. A virtual window presents a view of another location; this might not be the actual view of a camera, but information integrated from a multiplicity of sensors in the vicinity.

Another idea that might seem very similar, but in fact has a distinct functionality, is the virtual *Post-It Notes* metaphor. Post-It Notes in the real world are a very useful and versatile tool. Thus to use this metaphor in the realm of location-aware applications is very advantageous. The virtual Post-It contains information and is attached to objects, real or virtual. Hence it can be used to annotate objects. If such an annotated object moves, the Post-It moves with it. Furthermore, the Post-It can be removed from one object to be attached to another. If the object to which the Post-It is attached is no longer present (thrown away or deleted) the Post-It vanishes too. Additionally the information presented by a virtual Post-It might change depending on the person accessing it. Thus Post-It's implement the observer property.

The third metaphor we want to discuss is the metaphor of a *Virtual Information Tower* (VIT). This metaphor is modelled after real world advertising columns. A VIT is assigned to a certain geographical location and has a given area of visibility described by a circle or polygon. Each VIT is a container for a set of posters, which can be structured hierarchically and have their own location and area of visibility inside the area of the VIT.

3.2 Implementations

The *Worldboard* project at the University of Indiana (see WORLDBOARD 2000) is developing a location-aware system to extend the World Wide Web into physical space. Web pages, other digital media or groups of them can be placed as a virtual document at a given location in the real world. These documents are described by the Contextual Media Integration Language (CMIL). Using the CMIL language, information can also be defined as being in the proximity of a certain (mobile or non-mobile) object, which is identified through a unique tag, e.g. a bar-code. Additionally, it is possible to define certain intervals of time, during which a CMIL document is visible. The worldboard implementation implements the poster metaphor.

The concept of *Stick e-Notes*, which has been created at the University of Kent, is described as based on the metaphor of real-world Post-It notes (see PASCOE 1997). A Stick-e note is associated with a certain context, which may consist of a location and additional conditions, e.g. presence of a person or object, time, temperature etc. and of combinations of them. This context is also called a trigger condition. A Stick-e note is invoked (triggered), if the specified condition is met. The Stick-e note architecture does not specify what happens when a note is invoked. For each note this is defined by a specific view class, instead. A view class can be used to present the contents of the note on a display, but it can also trigger an event, thus acting like an observer or actor. It is clear from this description that in our classification the implemented metaphor is not the suggested Post-It metaphor, but the Poster metaphor characterized above (see also Section 4).

At the University of Stuttgart we have realized location-aware information access by using the VIT metaphor (described in LEONHARDI ET AL. 1999). A user equipped with a mobile computer is presented a list with all the VITs which are "visible" from his current location, and can access the contained posters.

Other systems using spatial context contain further metaphors. In his vision of Situated Information Spaces, FITZMAURICE 1993 describes the idea of adding virtual anchor points to an office environment, which can be viewed through a mobile device and represent voice annotations. In Augmented Reality systems (see e.g. STARNER ET AL. 1997) simple *text labels* and other more complex graphical tags can be attached to geographical locations or objects and are overlayed via a half transparent "heads-up" display over the user's view of the real world. These tags are visible when the user looks at the given location or object.

4 Classification

Table 1 holds the classification of the metaphors we have discussed in Section 3,

Metaphor	Observer	Actor	Container	Presenter	Associated	Limited
Poster	Yes[a]	Yes[b]	No	Yes	Yes	Yes[b]
Post-It Notes	Yes	No	No	Yes	Yes	Yes
Virtual Information Towers	Yes	No	Yes	Yes	Yes	Yes

Table 1. Classification of the Metaphors

[a] The window metaphor.
[b] The calendar metaphor.

Table 2 contains the classification of the implementations discussed in Section 3. Here again it is clearly shown that the Stick e-notes implement the poster instead of the Post-It metaphor.

Metaphor	Observer	Actor	Container	Presenter	Associated	Limited
Worldboard	Yes	No	No	Yes	Yes	Yes
Stick e-notes	Yes	Yes[a]	No	Yes[a]	Yes	No
VITs	Yes	No	Yes	Yes	Yes	No
Situated Information Spaces	Yes	No	No	Yes	Yes	No
STARNER ET AL. 1997	Yes	No	No	Yes	Yes	No

Table 2. Classification of the implementations

[a] depends on the specific view class.

5 Conclusion

In this paper we introduced a taxonomy that allows to characterize existing and future metaphors for context-aware information access exhaustively. We also provided an overview of existing metaphors and finally, classified the mentioned metaphors.

We think that our taxonomy can not only be used for classification purposes but also as a means to design a system supporting a context-aware metaphor, since it can be used to identify the basic characteristics of the supported metaphor. These characteristics can then be used to determine the required components of the system.

6 References

ABOWD ET AL. 1997. Abowd, G. D. and Atkeson, C.G. and Hong, J. and Long, S. and Kooper, R. and Pinkerton, M.: Cyberguide: a Mobile Context-Aware Tour Guide. In Wireless Networks, Volume 3, Number 4, pp. 421 - 433, 1997.

DAVIES 1998. Davies, N. and Mitchell, K. and Cheverst, K. and Blair, G.: Developing a Context Sensitive Tourist Guide. in Proceedings of the First Workshop on Human Computer Interaction with Mobile Devices, University of Glasgow, Scotland, U.K., pp. 64 - 68, 1998.

FITZMAURICE 1993. Fitzmaurice, G.W.: Situated Information Spaces and Spatially Aware Palmtop Computers. In Communications of the ACM,Volume 36, Number 7, pp. 39-49, 1993.

LEONHARDI ET AL. 1999. Leonhardi, A., Kubach, U., Rothermel, K. and Fritz, A.: Virtual Information Towers - A Metaphor for Intuitive, Location-Aware Information Access in a Mobile Environment. In Proceedings of the Third International Symposium on Wearable Computers (ISWC'99), San Fransisco, CA, USA, IEEE Press, 1999.

PASCOE 1997. Pascoe, J.: The Stick-e Note Architecture: Extending The Interface Beyond the User. Proceedings of the 1997 International Conference on Intelligent User Interfaces, pp. 261 - 264, 1997.

STARNER ET AL. 1997. Starner, T., Mann, S., Rhodes, B., Levine, J., Healy, J., Kirsch, D., Picard, R. and Pentland, A.: Augmented Reality Through Wearable Computing. In Presence Journal, Volume 6, Part 4, 1999.

WORLDBOARD 2000. Worldboard Project Pages, Indiana University, Indiana, USA,
URL: http://www.worldboard.org, Work in Progress, 2000.

Using Context as a Crystal Ball: Rewards and Pitfalls

Keith Cheverst, Nigel Davies, Keith Mitchell and Christos Efstratiou
Distributed Multimedia Research Group
Department of Computing
Lancaster University
Lancaster, LA14YR, U.K.
+44 (0)1524 594539
{kc, nigel, mitchelk, efstrati}@comp.lancs.ac.uk

1. Introduction

The use of context (both situated and environmental) has significant potential for simplifying the user's understanding of, and interaction with, complex interactive systems. Due to the size of devices appearing on the horizon, e.g. smart phones, the interactive systems that run on these devices are likely too share very limited input/output bandwidth at the interface between the system and the user. Therefore, techniques for simplifying interaction and reducing any input/output bottleneck is certainly desirable in order to serve the notions of ubiquitous computing [Weiser,91] and 'Information appliances' [Norman,99].

This paper discusses some of the potential rewards and pitfalls that can await designers wishing to incorporate context-awareness [Schilit,94][Brown,97] into interactive systems. Many of the issues are described in anecdotal form, based on our experiences developing and evaluating the context-aware GUIDE system [Cheverst,99][Cheverst,00].

To introduce some of the issues that arise concerning context-awareness, consider that popular interactive system: the car, and, more specifically, its braking system.

The Anti-locking Braking System (ABS) is a context-aware system that was introduced as a safety measure in order to reduce braking distance and greatly reduce the potential for a driver causing their car to skid through excessive braking. The context sensed by the antilock braking system includes:

i) Whether the driver is currently trying to brake (i.e. situated context),

ii) Whether or not the wheel is currently 'locked' under braking (i.e. environmental context).

The adaptive element of the system involves detecting when the wheel is locked and then decreasing braking force until the wheel is no longer locked. At this point more braking force is applied provided the situated context is still that of braking.

Before the advent of ABS, drivers were required to develop mental models that took into account the complex interrelationship that exists between braking force and the friction between the car's tyres and the road. The number of accidents each year involving cars skidding illustrates the fact that many drivers miscalculate the aforementioned relationship.

Consideration of the ABS system allows us to identify an agent, acting on behalf of the driver, that reduces the mental and physical demands of driving the car. In effect, the agent takes some control (or power) away from the driver and (providing the driver prefers less rather than more interaction with the car) makes the car easier to drive.

The ABS system enables the driver to form a simplified mental model regarding the cars braking system, i.e. drivers don't need to have such a detailed comprehension of the rules governing 'excessive' braking force and the resultant lack of control. However, if the car's driver is used to a conventional, manual, braking mechanism he or she might have learnt the skill of ‘pumping the brakes’ in order to prevent the car from skidding. Unfortunately, if this skill is employed by the driver of a car with ABS, the two approaches can conflict causing the braking distance to be increased. This example highlights three potential pitfalls that can arise from adapting to context, namely:

i) The problem of failing to reach a stable state [Thimbleby,90].

If both the user and the system attempt to adapt to the current context then it is unlikely that the system will manage to reach a stable state. Under such circumstances the system is likely to appear unpredictable. When designing context-aware systems, it is clearly important to consider the background/expertise of the user, i.e. are they likely to have already formed a mental model for interacting with a similar (non-adaptive) system?

ii) The trade-off between prescription and freedom.

If the driver wanted, for whatever reason, to lock the wheels of the car then the system would prevent him or her from achieving this task.

iii) The user must trust the agent performing adaptation on his or her behalf.

When ABS was first introduced, there was, not altogether surprisingly, some mistrust of the system by drivers. Indeed, the driver who knows the workings of the ABS system is required to trust both the context sensing technology and the intelligence of the agent, i.e. its infallible ability to react appropriately to the context in a failsafe manner.

It is possible to identify three main ways in which context can be used to simplify the user’s interaction with an interactive system:

i) Simplifying/reducing the task specification required from the user in order to achieve his or her desired goals, i.e. reducing the need for input/action by the user.

At one level this can simply mean filling in a required blank, such as the user’s current location, based on information that is sensed by the system. However, at a higher level, it can also involve attempting to pre-empt the user’s current goal in order to reduce his or her task specification (e.g. the ABS system).

ii) Changing the output produced by the system, i.e. the reducing quantity of information that has to be processed by the user or increasing the quality of information presented.

Once again, some reduction in output might be achievable by attempting to pre-empt what output is likely to be required/expected by the user, based on the current context.

iii) Reducing the complexity of rules constituting the user’s mental model of the system.

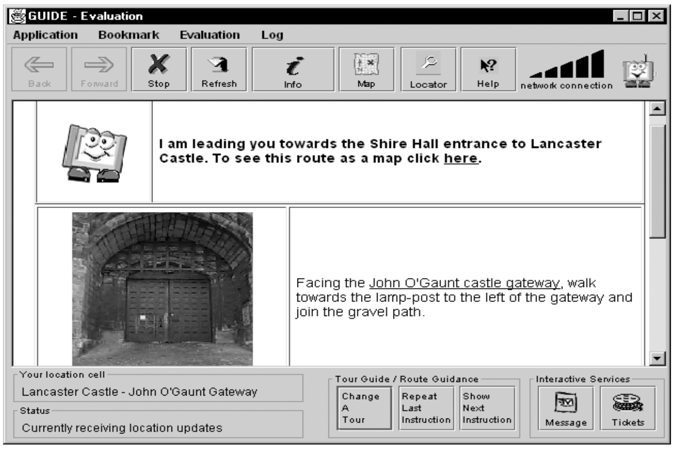
This is generally achieved by some form of intelligent agent that performs some proportion of the required computation on the user’s behalf.

The following section describes and analyses some of the positive and negative experiences of using context gained through our development and evaluation of the GUIDE system.

2 Experiences Developing a Context-aware Tourist GUIDE

The GUIDE system has been developed to provide visitors to the city of Lancaster with information that is tailored to his or her context. The city contains a number of strategically positioned wireless communication cells with a diameter of approximately 300 m depending on the layout of buildings. These communication cells are used for disseminating location information and tourist information to mobile GUIDE units. By carrying a GUIDE unit, visitors will receive up-to-date information about the city's attractions while following a structured tour of the city tailored to their specific requirements.

One of GUIDE's key requirements [Cheverst,00] was that of flexibility. In more detail, visitors using the system should be able to change from one kind of situated interaction to another. So, for example, the visitor should be able to change from following a tour to booking accommodation and then resume their tour. For this reason, we designed a user interface that enables the user to switch between different aspects of functionality in order to meet the demands of his or her current situation (figure 1).



I am leading you towards the Shire Hall entrance to Lancaster Castle. To see this route as a map click here.

Facing the John O'Gaunt castle gateway, walk towards the lamp-post to the left of the gateway and join the gravel path.

Your location cell: Lancaster Castle - John O'Gaunt Gateway

Status: Currently receiving location updates

Figure 1: The user interface presented to a visitor when following a tour.

Following some initial evaluation, a compromise was made to reduce some of the flexibility afforded by the user interface in order to increase ease-of-use. The compromise involved making 'following a tour' the default mode of situated interaction, until the tour has either been completed or aborted.

In order to describe the way in which GUIDE adapts to context the three categories identified in the introduction will be used:

i) **Simplifying/reducing the task specification.**

One type of context utilized by GUIDE is that of the visitor's location and the location of attractions within the city. The system reduces the need for input by assuming that the information required by the visitor is strongly influenced by his or her current location. So, for example, a visitor standing outside Lancaster castle can request the system to 'tell me about the area I am in' as opposed to searching the contents page for 'Information on Lancaster Castle'. However, the initial GUIDE system made the mistake of only allowing visitors to obtain information regarding their current location and so the over-determination [Thimbleby,90] employed by the system to simplify the visitors task was inappropriate. The lack of flexibility in this version frustrated users and so the system was extended to enable visitors to specify their requirements more fully, e.g. by searching for information using a keyword. This last point clearly illustrates the intrinsic problem of trying to successfully pre-empt the goal of the user.

ii) **Changing the output produced by the system**.

In general, the GUIDE system attempts to constrain and tailor information presented to the visitor based on his or her current context. So, for example, when the visitor requests a list of nearby attractions, the list is constrained in such a way that those attractions that are open, and have not already been visited, are placed higher up the list. The assumption is made that the visitor is more likely to be interested in attractions that are open and that have not already been visited. An earlier version of the system constrained the output by removing all closed attractions from the presented list. However, this frustrated some visitors who were interested in visiting the attraction anyway, e.g. to view the architecture of a building. Again, this demonstrates the difficulty of pre-empting the user's goal. A future version of the system will use the visitor's stated interest in architecture to determine whether closed attractions with a clear architectural value are included in the list.

iii) **Reducing the complexity of the user's mental model.**

In GUIDE, the agent that acts on behalf of the user is designed to relieve the user of the onerous task of studying maps and guidebooks in order to devise and follow an interesting tour. In more detail, the agent calculates tours based on a variety of different contexts, such as the visitor's current location, the current time, special opening hours of attractions, the relative positioning of attractions in the city and the preferences of the visitor, e.g. an interest in historic buildings.

While evaluating the GUIDE system we experienced some difficulty capturing the visitor's location context with sufficient accuracy. Of course, the problem with obtaining inaccurate or incorrect contextual information is that the adaptation performed by the system, based on the context, will produce inappropriate results. In the case of GUIDE this meant that when presenting a list of 'nearby attractions' to the visitor, some of the attractions were not always as 'nearby' as might have been expected.

3. Strategies for Building Context-Aware Applications

A number of strategies can be identified for the design of interactive systems that utilize situated and/or environment-based context. The following strategies are based on our analysis of existing context-aware interactive systems, such as ABS, and, in addition, those concerns that were experienced during the development and evaluation of the GUIDE system:

i) When using context to constrain the presentation of information, or to simplify the specification of a task, it is crucial that the adaptation does not inappropriately over-determine the users interaction.

ii) Furthermore, designers need to carefully consider the fundamental trade-off between prescription and freedom/flexibility when deciding how to adapt to context.

iii) When considering adaptation to context, designers should be careful to bear in mind the principal of least astonishment and the need for predictability. Of course, if designed well, then adaptation to context has the potential to increase the integral predictability/consistency of the system. However, as described in the ABS example, the inappropriate transfer of skills can cause difficulties.

From a more technical perspective, the following issues need to be considered when engineering context-aware systems.

i) The sensing technology used for obtaining context needs to be dependable. This means both accurate and available in a timely manner.

ii) The intelligence of the agent responsible for adapting to context needs to be flexible in order to cope with problems obtaining context and the potential for over-riding demands by the user.

Hopefully, by considering these strategies/issues, designers of context-aware interactive systems can avoid many of the potential pitfalls.

4. Conclusion

This paper has considered some of the potential rewards and pitfalls of utilizing situated and/or environment-based context in the design of interactive systems.

In summary, adaptation to context can be used to develop interactive systems that allow users to form a simplified mental model for understanding, and interacting with, the system. In effect, context-aware systems migrate complexity away from the user to some form of intelligent agent.

Context also has the ability to reduce the complexity of a user’s task specification and reduce the quantity (and/or increase the quality) of information output by the system. This ability is crucial when considering the generation of interactive systems on the horizon that will be designed to run on future mobile computing devices. Such devices are likely to share very limited input/output bandwidth at the interface between the user and the device and so maximizing the use of this bandwidth is of paramount importance.

To conclude, providing designers consider the potential hazards, context-awareness could help to bring the ubiquitous information appliance one step closer to reality.

References

[Cheverst,99] Cheverst, K., N. Davies, K. Mitchell, and G.S. Blair. "The Design of an Object Model for a Context-Sensitive Tourist Guide", Computers and Graphics, Vol. 23. No 6, December 1999, pp. 883-891.

[Cheverst,00] Cheverst, K., N. Davies, K. Mitchell, A. Friday and C. Efstratiou. “Developing a Context-aware Electronic Tourist Guide: Some Issues and Experiences, to appear in Proceedings of CHI'00 (Netherlands, 2000), ACM Press.

[Thimbleby,90] Thimbleby, H. User Interface Design. Addison-Wesley Longman Ltd, ISBN 0-201-41618-2. (1990)

[Norman,99] Norman, D.A. The Invisible Computer. MIT Press. ISBN 0-262-14065-9. (1999)

[Weiser,91] M. Weiser, “The Computer for the Twenty-First Century”, Scientific American, 265 (3), September 1991, pp. 66-75.

[Schilit,94] Schilit, B., Adams, N., and Want, R. Context-Aware Computing Applications. Proc. Workshop on Mobile Computing Systems and Applications, Santa Cruz, CA, U.S. (1994).

[Brown,97] Brown, P.J., Bovey, J.D., and Chen, X. Context-aware applications: from the laboratory to the market place. IEEE Personal Communications, Vol. 4, No. 5 (1997) pp 58-64.

Extended abstract, submitted January 28 to CHI2000 Workshop on Situated Interaction in Ubiquitous Computing

Magic Touch: A Simple Object Location Tracking System Enabling the Development of Physical-Virtual Artefacts in Office Environments

Thomas Pederson
Department of Computing Science
Umeå University
SE-90187 Umeå, Sweden
+46 90 786 65 48
top@cs.umu.se

TRACKING OBJECTS

One straight-forward way of getting input from physical user activities in office environments is to track location and location changes of objects. In other words to reduce "activities" to "object change of location". Of course, this simplification implies considerable loss of other kinds of information, including for instance mental state and movements of the user alone, but we believe that it is a good starting point for further improvements. Accepting these limitations, the parameters of interest are:

- Exactly what object is being moved?
- To what new location?

Some Available Technology and Methods

One obvious solution is to let all objects in the environment carry position transmitters, whose signals are received by a motion tracker system. Such an approach gives both parameters accurately and continuously. Drawbacks are that the system is expensive if you want to track many objects, and the identification tags are at least to date fairly large. Another method is to put a camera in the ceiling and to attach visual tags, showing unique graphical patterns, to the objects. These patterns and the location of the tags are interpreted and calculated through analysis of the camera image [9]. In this case the tags are considerably cheaper compared to the other approach since it is possible to print them out on an ordinary printer. Drawbacks include the necessity of free line-of-sight between the camera and the tagged objects and that the tags themselves become fairly large if you want to identify many objects.

A MORE INDIRECT METHOD

While searching for a suitable tracking method we discovered a fundamental fact: objects in office environments don't move by themselves! They move *when they are moved by users' hands*. Put in another way, an object stays where it is until the user grabs it in one or two hands, moves the hand(s) to a new location and drops the object. Based on this insight, we developed another object location tracking method which at least for our purposes probably will be more suitable than the other two mentioned tracking approaches. However, the system is currently under development and evaluation has to be performed in order to substantiate our belief.

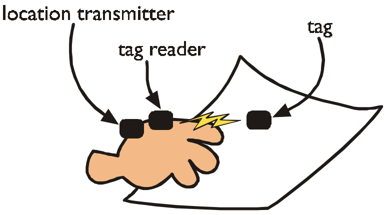


Figure 1: *One of the user's hands holding a tagged paper document.*

As illustrated in Figure 1, the object location tracking system consists of (1) an office environment containing RF/ID-tagged artefacts, (2) wearable wireless tag readers, placed on each of the user's hands, identifying any tagged object the user takes in her/his hand, and (3) a wireless location transmitter always aware of the positions of the user's hands.

The advantage of using RF/ID tags for our kind of application instead of other similar solutions like bar-codes is thoroughly discussed by Want et al. [10].

As soon as the user's hand comes close to a tagged object, it is identified. If the user moves her/his hand and the reader still can read the tag it means that the user has grabbed the tagged artefact. When the tag is no longer readable, the user must have dropped the artefact, and the location of the user's hand at the point when the tag was last readable is regarded as the new location of the artefact.

Having the development of Physical-Virtual Artefacts[1] (PVAs) as an overall goal, the above described object location tracking system can be integrated with a PVA Database Management System (PVA-DBMS) where each artefact instantiation, no matter physical or virtual, has an entry as shown in Figure 2.

---

[1] **Definition:** A physical-virtual artefact is an abstract artefact that (1) is instantiated in both the physical and virtual environment, where (2) these instantiations to a large extent utilize the unique affordances and constraints that the two different environments facilitate, and finally (3) where one instantiation of a specific physical-virtual artefact is easily identified if an equivalent instantiation in the other environment is known [8].

Limitations

With small and powerful enough readers and motion tracking technology we believe that this method would deliver location changes of objects accurately and non-intrusively. However, the system relies on several limiting assumptions. Here are the most evident:

- All objects that are to be tracked have to be tagged. (The proposed system shares this limitation with the other systems however.)
- No tagged object is allowed to change its own position by itself unless it is able to communicate the new position to the system by itself.
- The user does not drop an artefact “in the air”, letting it fall down to its final location. If so, the system will store an incorrect height location.
- The user moves only one artefact at a time. If many tagged artefacts are to be moved at once, some kind of “multi movement mode” has to be entered explicitly.
- The user always moves artefacts directly with the hands and does not use any kind of tool to push or carry the artefacts “from a distance”. This would disable the identification mechanism since the tag reader has a limited reading range.
- Only users that carry the wearable system equipment are allowed to move artefacts in the environment. If non-tracked user hands are active in the environment the system’s artefact location database will become incorrect.
- The position of an object is based on the tag position, that is, one point in space. For small artefacts, it is a reasonable approximation, for larger objects the position approximation error will be more evident.

Some of the limitations mentioned above can be eliminated or at least reduced by the introduction of additional mechanisms, sometimes also involving explicit user activity.

IN SEARCH FOR MEANING

Apart from the fact that the proposed system could serve as a tool for researchers interested in physical behaviour of knowledge workers, *automatised* user modelling could of course be used to enhance the performance of the system itself. One fairly easy way to extract contextual cues would be to consider time and place of physical user activities.

Artefacts that are used (in our simplified version, *moved*) frequently can be distinguished as being generally useful and important for everyday activities. Artefacts that are less frequently used but perhaps often at the same time instant, or often placed close to each other, could be considered as being related to each other. Implicitly acquired knowledge such as this would make it possible for the system to create and maintain a self-organized artefact relationship model, a kind of semantic network. Among other things, this model could help solving interaction ambiguities and/or let the system suggest relevant material that the user might have overlooked. Another feature of this model is that it reduces the need for explicit definitions of artefact relationships and categorisation, which is a task connected to significant cognitive effort [6]. It also opens up for a less predefined and a more individual organisation and interaction style compared to the kind of well-structured dialogue-driven Human-Computer Interaction common today.

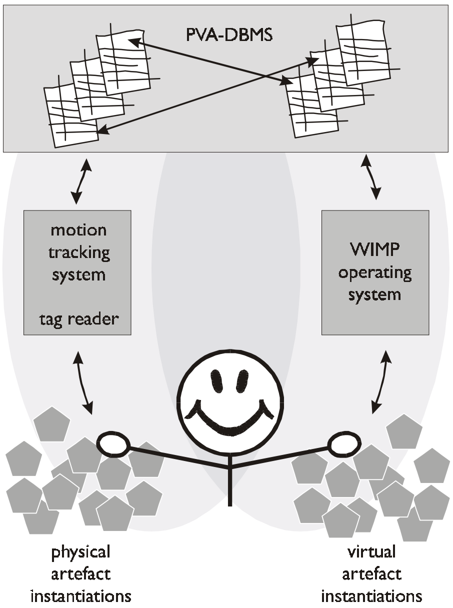


Figure 2: *Magic Touch architecture outline. The PVA-DBMS linking the physical (left) and virtual (right) environments together by keeping track of artefact instantiation changes done by the user in any of the two environments.*

Based on previous activity sequences the system could also try to predict what artefact or what system functionality the user probably would like to get access to in following activities, improving work efficiency.

APPLICATION IDEAS

Having discussed the basic parts and ideas related to the Magic Touch system, here are some examples of possible artefacts and applications:

- A Physical-virtual (PV) search engine enabling search not only for virtual artefact instantiations but also for physical ones. The system points out where in the physical environment the user left it the last time it was moved.
- A PV mail box handling both physical and virtual mail in a similar fashion.
- A PV paper basket handling and synchronising discarding of PV documents.
- Tele-presence. By visualising the PVA-DBMS and enable the access to the visualisation through the Internet, users can visit Virtual Reality versions of their physical offices from any place providing an Internet connection.
- PV containers. Physical instantiations of PV containers (e.g. document folders) can be linked to sets of virtual artefact instantiations and vice versa, making it possible to “keep” both physical and virtual artefact instantiations in the same container, physical or virtual.
- PV Stacks. The user can explicitly define, or the system can implicitly infer (see previous section) stacks or piles of objects that can later be referred to as a physical instantiation of a PV container. Among other things, this could eliminate the necessity of moving one object at a time.
- Active volumes — volumes in physical space that the user explicitly has assigned some “meaning”. The difference compared to the PV stacks and containers is that the defined volume is physically just empty space. The user could for instance define one part of the physical desktop as being a mail outbox or a paper basket. End-user programming would tie appropriate system actions to each active volume in a similar fashion as in Want et al. [10].
- “Magic memory” allowing backtracking of past user actions, and limited UNDO facilities. By just giving the system a point in the past time, it can (at least as a visual-

isation) reconstruct it, showing location and status of each physical and/or virtual objects at that particular time instant.

- "Virtually filled" physical artefact instantiations. If the proposed system is combined with a motion tracked Head Mounted Display (HMD), users can handle blank papers, order them in piles, bookshelves etc. while the actual contents of the papers is projected virtually. In this case, the physical artefact instantiation (the paper) provides the tactile and spatial feedback while the content is provided by the virtual instantiation in the same fashion as a mediaBlock [2]. Benefits would be better system control of artefact status. Drawbacks include more things to wear and decreased visual ergonomy.

CURRENT DEVELOPMENT STATUS

A wearable wireless radio-based (RF/ID) tag reader is being developed at the time of writing. A complete Magic Touch physical-virtual knowledge work environment including a wireless motion tracker and a PVA-DBMS is planned to be set up and developed during spring 2000. Evaluation is planned to take place during autumn 2000.

Hardware Limitations

Current hardware limitations include:

- Limited motion tracking precision. Suitable systems have a precision of around 2-3 cm in xyz directions
- Small RF/ID-based tag readers such as the one used in this project currently have a maximum reading distance of about 2 cm depending on tag types. To work well, a reading distance of 5-10 cm would be desirable. On the other hand, larger reading distance would increase the possibility of confusing tags.

RELATED WORK

Seeing Wellner's DigitalDesk [11] as a starting point, there has been a continuous interest in merging the physical and virtual worlds in office environments and in more specialised settings [1, 5]. Compared to the DigitalDesk, the proposed system Magic Touch covers a whole office rather than a desk. More recent sources of inspiration to the present work have been the research done on Graspable [2], Tangible [4] and Manipulative User Interfaces [3]. One major difference between these systems and the proposed one regards how or rather *where* the physical objects (common terms are "Bricks", "mediaBlocks" or "Phicons") are identified. While in the former systems they are identified by tag readers mounted on a set of designated "physical-virtual docking stations", the number of tag readers needed in Magic Touch is never more than two, both attached as wearable wireless devices on the user's hands. This difference also holds for Want et al. [10] where they, although using RF/ID tags, attach the readers on portable computing devices rather than hands and furthermore don't combine the identification mechanism with motion tracking of these devices as in Magic Touch. The support for searching artefacts in physical environments, which is one of the possible applications of the proposed system, has some similarities with the InfoClip [7] solution. However, InfoClip does not involve any centralised knowledge database for where artefacts are located.

CONCLUSIONS

In this extended abstract we have presented an outline of a system architecture for integrated physical-virtual knowledge work environments [8], Magic Touch. By tracking simple knowledge work actions like the changing of artefacts' locations we believe that we can integrate and enrich the working environment as a whole. By tracking user hands only, and by tagging physical objects, we believe that the system becomes more affordable and powerful than known alternatives for the given application area. However, the system is not yet fully implemented and although it has potential of enabling many interesting new applications, apart from generally integrating physical and virtual environments, the system has to be evaluated in practice.

REFERENCES

1. Arias, E., Eden, H., & Fischer, G. Enhancing Communication, Facilitating Shared Understanding, and Creating Better Artifacts by Integrating Physical and Computational Media in *Proceedings of Designing Interactive Systems (DIS 97)*, ACM Press.
2. Fitzmaurice, G.W., Ishii, H. and Buxton, W. Bricks: Laying the Foundations for Graspable User Interfaces, in *Proceedings of CHI'95*, ACM Press, 442-449.
3. Harrison, B.L., Fishkin, K.P., Gujar, A., Mochon, C., and Want, R. Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces, in *Proceedings of CHI '98*, ACM Press, 17-24.
4. Ishii, H., and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, in *Proceedings of CHI'97*, ACM Press, 234-241
5. Mackay, W. E., Fayard, A.-L., Frobert, L., & Médini, L. Reinventing the Familiar: Exploring an Augmented Reality Design Space for Air Traffic Control, in *Proceedings of CHI'98*, ACM Press, 558-565.
6. Malone, T. (1983) How Do People Organize Their Desks? Implications for the design of office information systems, in *ACM Transactions on Office Information Systems*, 1(1), 99-112, January 1983.
7. Mori, H., Kozawa, T., Sasamoto, E., & Oku, Y. A Computer-Augmented Office Environment: Integrating Virtual and Real World Objects and Behavior, pp1065-1069 in *Proceedings of HCI International'99,* Lawrence Erlbaum, ISBN 0-8058-3392-7.
8. Pederson, T. Physical-Virtual instead of Physical *or* Virtual — Designing Artefacts for Future Knowledge Work Environments, pp1070-1074 in *Proceedings of HCI International'99*, Lawrence Erlbaum, ISBN 0-8058-3392-7.
9. Rekimoto, J. Augmented Reality using the 2D matrix code, in Interactive Systems and Software IV, Kindaikagaku-sha, pp. 199-208, 1996. (In Japanese.)
10. Want, R., Fishkin, K., Gujar, A., & Harrison, B. Bridging Physical and Virtual Worlds with Electronic Tags, *Proceedings of CHI'99*, ACM Press, 370-377.
11. Wellner, P. Interacting With Paper On the DigitalDesk, in *Communications of the ACM 36*, 7 (July 1993).

POSITION PAPER:

Designing a Hierarchy of User Models for Context-Aware Applications

Michael Samulowitz

University of Munich, Dept. of CS
Oettingenstr. 67
D-80538 Munich, Germany
E-Mail: samulo@nm.informatik.uni-muenchen.de

Corporate Technology
Siemens AG
Otto-Hahn-Ring 6
D-81730 Munich
E-Mail: Michael.Samulowitz@mchp.siemens.de

Abstract. *Research in Ubiquitous Computing and Wearable Computing yielded several types of context-aware applications, dramatically changing the way to interact with computers. User Modelling is a key concern for these types of applications in order to accomplish personalization. In most cases each user model is specific to a certain application and opaque to others. System design could be improved if common data structures could be communicated between different user models. Therefore we propose the organisation of user models in a hierarchical structure relating different user models to each other.*

Introduction

Work in Ubiquitous Computing [1], Augmented Reality [2], Smart Rooms [10], Reactive Environments [3] enabled the vision of a *computer-augmented environment*, electronic systems are merged into the physical world to provide computer functionality to everyday objects. The technology should be distributed (ubiquitous), yet invisible, or transparent, since the full potential of the computer can only be realized when the machine itself is hidden from the user. This concept marks a dramatic shift from the status quo in which interaction with the computer interferes with our activities rather than enhancing them. This vision is often enabled by putting sensors in the environment (e.g. rooms).

Another approach to ease human-computer interaction is tackled by work on Wearable Computing [8,9,10]. In opposite to Ubiquitous Computing[1] Wearable Computing does not necessarily require any environmental infrastructure at all. In the purest form, the wearable user would do all detection and sensing on her body [11].

Both the pure Ubiquitous Computing and the pure Wearable Computing paradigms may be applied to context-aware applications. As described in [4], context-aware applications may embed arbitrary context information (e.g. location, collection nearby objects/persons, accessible devices and services) in a flexible way.

---

[1] Ubiquitous Computing = {Augmented Reality, Reactive Environment, Smart Space}

Here, we want to focus on personalization of services taking in account contextual information and *user models*. In general a user model contains the system's assumptions about the user [5]. Assumptions may be constantly refined by monitoring user activities or explicit by user feedback (e.g. [6]). User modelling is common to several contextual applications (list below).

Separating Concerns in User Modelling for Context-Aware Systems

The following list states general examples where user models enhance usability of applications by embedding user related information.

- Message Filtering [7]
- Location-based information services, e.g. [12]
- Service/Resource discovery [13] and selection (Service/Resource Discovery is a key concern given the enormous amount of information potentially available in the environment)
- User-adapted interpretation of context information
- Adaption of Interface Agents
- Environment configuration [10]
- Process Activation/Deactivation in dependence of context, e.g. [14]

The listed examples clarify the difficulty to introduce a single user model for context-aware systems, because eventual personalization covers several areas of use and there is no common model that fits to all of them. For this reason we propose to separate concerns in user modelling for context-aware systems. This would result in multiple user models each being specialized on a certain context of use. Still there is information common to all of them e.g. superior settings regarding security or privacy issues. This information should be propagated to the other user model (modules).

A Hierarchy of User Models

We suggest a hierarchy of user models in order to organize user-centric data in context-aware systems. Figure 1 states an example.

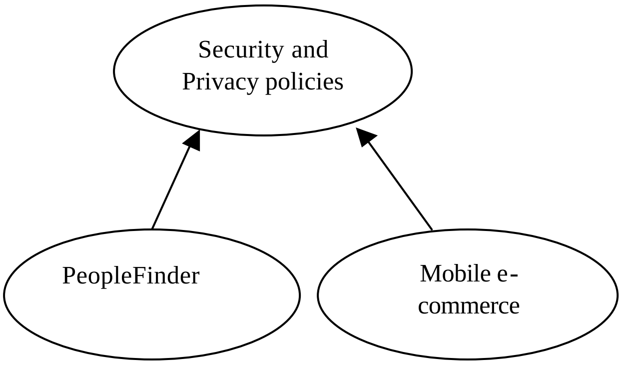


Figure 1

The Figure shows two different user models for two context-aware applications (PeopleFinder and Mobile e-commerce). Both should obey to settings in the user model for security and privacy policies. Hence, if a user changes these policies, information about this should be *propagated* to the user models associated (the PeopleFinder and the Mobile e-commerce application). If the user selects a "high" privacy level he should neither being locatable nor being annoyed by messages that announce offers to buy.

Figure 2 depicts another example for distributing user model data on a hierarchy of entities, each entity representing a certain subset of the user model data. This example takes in account spatial concerns and provides a higher degree of granularity in comparison to the previous example.

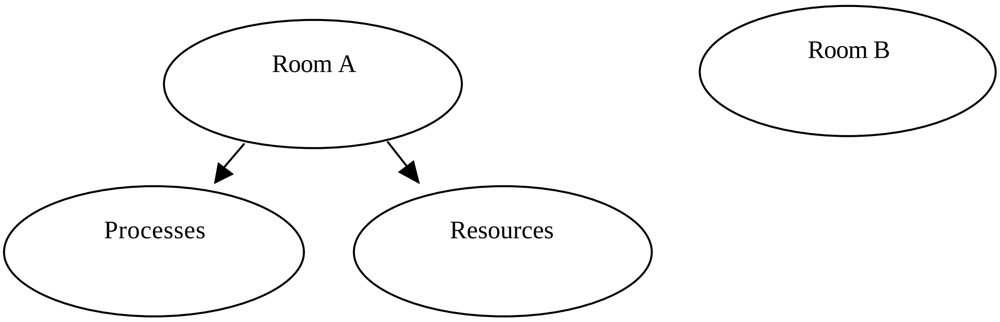


Figure 2

The left side of the figure (Room A) shows the user model for a room; it mainly includes processes which a triggered by certain user actions [14], and references to resources the user normally uses. Now the same user wants to have the same settings ("look & feel") in room B[2] as in room A, he would ask the system to do so. The system would automatically *transfer* all settings related to room A (including the associated entities) to room B. (The set of candidates for transfer operation could be determined by referential integrity constraints.) Transfer is not always trivial; in the given example it is not sure that room B supports everything specified in the user model for room A. This could lead to partial transfer or intelligently adapting functions.

In order to implement this design we could associate an agent with each entity managing it. These agents could co-operatively implement *transfer* and *propagate* operations.

The following points require further studies:

- Interactions schemes to manipulate/browse the hierarchy
- Further Operations
- …

References

1. Weiser, M. Some Computer science issues in ubiquitous computing. *Communications of the ACM* 36, 7 (July 1993), 75-83
2. Buxton, W. Living in augmented reality: Ubiquitous Media and reactive environments. In K. Finn, A. Sellen and S. Wilber, Eds. *Video Mediated Communication*, Erlbaum, NJ. 1997

---

[2] "Room X" means the physical room and the entity modelling related user data for this room. Sorry.

3. Jeremy R. Cooperstock, Sidney S. Fels, William Buxton, and Kenneth C. Smith. Reactive Environments. *Communications of the ACM* 40, 9 (September 1997), 65-75
4. William N. Schilit. A system Architecture for Context-Aware Mobile Computing. PhD Thesis, 1995
5. Alfred Kobsa. User modeling and user-adapted interaction. *Proceedings of the CHI'94 conference companion on Human factors in computing systems*, 1994, 415 - 416
6. Rob Barrett, Paul P. Maglio and Daniel C. Kellem. How to personalize the Web. Conference Proceedings on Human factord in computing systems, 1997, pages 75 – 82
7. Matthew Marx and Chris Schmandt. CLUES: dynamic personalized message filtering. *Proceedings of the ACM 1996 conference on Computer supported cooperative work, 1996, pages 113 – 121*
8. T. Starner, S. Mann and B. Rhodes The MIT wearable computing web page. http://wearables.www.media.mit.edu/projects/wearables.
9. Abowd, G. D., A. K. Dey, R. Orr, and J. Brotherton, "Context-awareness in Wearable and Ubiquitous Computing," GVE, Technical Report GIT-GVU-97-11, May 1997.
10. A. Pentland, "Smart Rooms", Scientific American, April 1993
11. Bradley J. Rhodes, Nelson Minar, Josh Weaver. Wearable Computing Meets Ubiquitous Computing: Reaping the best of both worlds, The Proceedings of The Third International Symposium on Wearable Computers (ISWC '99), San Francisco, CA, October 18-19 1999.
12. B. Rhodes. "The wearable remembrance agent: A system for augmented memory", *Personal Technologies Journal Special Issue on Wearable Computing*, vol. 1, no. 4, pp. 1:218-224, 1997
13. S. Czerwinski, B. Zhao, T. Hodes, A. Joseph, R. H. Katz. An Architecture for a Secure Service Discovery Service. Proceedings of The Fifth ACM/IEEE International Conference on Mobile Computing (MobiCom '99), Seattle, WA, August 1999, pp. 24-35
14. R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Trans. On Info. Sys., vol. 10, no.1, pp. 91-102, January 1992*

Situated interaction in an educational setting.
Position paper

Irene Mavrommati
Computer Technology Institute
Research Unit 3 - Applied Information Systems
http://www.cti.gr/RD3/
Email: irene.mavrommati@cti.gr

Brief description:

At Computer Technology Institute Research Unit 3, we are dealing with education. Within this context we try to include in our research proposals situated interaction for learning activities. We do this by proposing

a) at a lower level the use of different devices for different environments, and
b) At a higher level facilitating the customization of learning by agent technologies.

We would like to see in the future the use of combined networked media (including interactive TV, PDAs, and mobile phones), for education.

Framework: A few observations:

The situation of use of several devices can be adapting to the environment and the situation in which they are used.
As a low-level example I can refer to things such as

- Mobile phones, adapting the volume level according to the environmental noise.
- Mobile domain devices (PDAs / car navigation systems / mobile phones) displaying context related information, such as tourist information for specific landmarks in the proximity, or navigation information in order to find specific points/items (in cities, public spaces, stations, airports, libraries, hospitals).
- Sensing proximity to start up an application or trigger a state of the device.

At a higher level, examples of contextual use are devices or applications that change behavior / input method according to the situation of use or according to what other devices are in the proximity.

- Concepts for interactive filming, whereby by passive interaction (such as pulse measuring), one can determine interest and adapt the dramatic plot accordingly so as to keep the viewers interest within a high level.
- Personal devices that recognize the existence of other devices and use them as input / output (for example a PDA that recognizes the existence of an unused screen in the area, and uses it to display information at a larger layout)
- Mobile phones that recognize the situation (i.e. you are in a meeting) and are able to filter the incoming calls accordingly (i.e. pass you only phones related to that meeting, or urgent ones)

Situated interaction in the service of education

At Computer Technology Institute Research Unit 3, we are dealing with education. We develop applications for education, as well as use our experience and knowledge taken from that area into different context, in order to develop several, other than educational, technologies.
Within this context we try to include in our research proposals the point of situated interaction for learning activities.
We do this by proposing

a) at a lower level the use of different devices for different environments,
b) at a higher level facilitating the customization of learning by agent technologies. Agent technologies can be used to deal with personalization issues as well as situation specific context of use.

We would like to see in the future the use of combined networked media (including interactive TV, PDAs, and mobile phones), for education.

A draft methodology we started with is:
a) Mapping the educational activities in the different spaces,
b) Using the characteristics of each space, for optimizing learning in each environment,
c) Seeing learning as a global activity within all environments.
Then breaking down an educational application scenario to different media, different information appliances, with different interaction methods according to the mood the user is in. The information presented can be tailored according to context of use.

We can derive many alternative scenarios of use, starting by defining the axis that we want to take as reference regarding interaction, user mode, situation of use, environment and environmental conditions

1. Interaction in devices can vary from active to passive, where active is a more focused and conscious action of the user, and passive is depending more on biomentric input, and tailoring the information presented by sensing the environment / area / context of use.
2. Users can be in various different moods ranging from: active to relaxed, energetic to tired, impatient (in a hurry) to patient, e.t.c.
3. Situations can vary from social (collective / public) to individual, from static to mobile, from virtual to real, synchronous to asynchronous, etc.
4. Environments can also vary: public (collective or not), home, mobile (transportation) environments.
5. Environmental conditions can differ: from light to dark, silent to noisy, dry to wet, hot to cold, colorful to dull. Interfaces may reflect those environmental changes in practical ways -such as screens adapting their brightness and contrast to the lighting conditions automatically-, or aesthetic ways -such as interface colors and style being seasonal or reflecting weather changes.

As already said, CTI's interest lies in education: the exploration of information/education appliances, in different context of use, and using situated interaction paradigms.

A student in the learning process moves between different environments: home, university, and transportation. Moreover a student can be moving in both real and virtual environments where learning takes place. Each environment is characterized by different social situations, as well as a different 'mood' that the person is in.
Each space may reflect different learning modes: more attentive - participatory at the university, more focused and concentrated when using the home computer, more passive when watching an educational video on the home TV.
Within each single environment situations can also vary, for example at home one can be in completely different moods: less attentive, doing several tasks simultaneously, or focusing in only one task only; In each setting one can be in a more private or in a more collective environment with others.

The landmark of the different spaces is mapped by a variety of devices: mobile devices (portable audio, mobile phone, wearable, PDA); home devices (TV, audio set, DVD, telephone, PC); Public devices: info-kiosks, phone booths, boards, screens / projections. Etc.

Moreover we can perceive in the future these different spaces having each their own network. Certain areas of the networks are overlapping.
- A personal network can be a network between PDA, mobile phone, audio set, laptop, or any other devices one may be carrying.
- The home network is linking together the devices in the home and distributes their content (stereo set, TV, PC, PDA, Camcorders).
- A university network can be linking public information devices (notice boards), computers, telephones, info kiosks, projection rooms.

We are currently in the preliminary stages of a research concept (concerning informal education) where we try to address customizable, situation dependant learning and integration of it in different platforms. The aim of this presentation is to evoke discussion and thought, about how situated interaction can be used so that learning in particular can benefit from it.

Position Paper for CHI 2000 Workshop No 14: Towards a Deeper Understanding of Task Interruption

Mark Stringer, Marge Eldridge and Mik Lamming.
XRCE Cambridge
61 Regent St
Cambridge
England CB2 1AB

A Future Rich in Context

In the future, if we have not reached this stage already, we will be able to carry around with us many sensors capable of providing a whole slew of contextual information. Is the user tired? Is the user annoyed? Is the user talking to someone? Is the user walking, jogging or driving? It seems conceivable that each of us could soon be carrying a device capable of answering all of these questions. However, what is far from certain is what we can usefully do with this newly-available information. Our particular interest is develop application software which can take such context information gleaned from sensors and use it to support and possibly improve the interruption management strategies of mobile workers.

An Increase in Interruptability

A mobile device which is always on and always with the user is always capable of interrupting her. With the increase in ownership of mobile phones, and now smart phones and wireless PDAs has come a sharp increase in the possible methods of interruption (email/SMS/phone call) and an overall increase in the numbers of hours in every day for which the user is interruptible.

Perhaps we can find some way of using our mobile computing and its increasing access to our context to manage interruptions so that our levels of stress are not increased, and our levels of productivity are not lowered by constant interruption. For example one can easily imagine a situation in which our mobile phone or other mobile computing device is capable of detecting that I am in a face to face conversation with someone and therefore directs all phone calls to my voice mail and waits until the conversation finishes before notifying me of incoming email and other messages. In fact applications similar to this have already been developed [Sawhney and Schmandt, 1999]. This seems to be the archetypal case of how access to context can be used to mediate interruptions. However, further consideration shows just how *unhelpful* such an application might be, depending on the finer detail of the situation. It may be that the interrupting phone call that my mobile phone is shielding me from is from a colleague with a pricing vital to the deal I am discussing. It may be that I am merely chatting with a colleague while killing time. It may be that the silent, thoughtful period after I have left a meeting is exactly the *wrong* time to distract me with a flurry of message notifications.

We may have some intuitions about what effect this significant increase in susceptibility to interruptions has on efficiency and well-being of the mobile worker but there is little concrete research, and what there is offers conflicting evidence. For example, a study has shown that a high percentage of interruptions which occur in an office environment are useful rather than disruptive and significantly affect which tasks are performed [Ziljstra et al., 1999]. This raises the prospect that workers who are away from an office environment may not be being interrupted enough! Another study has shown that the nature of the individual task and the experience of the subject can greatly influence the effect which interruptions have on performance [O'Conaill and Frohlich, 1995]. There may be no general rule which can be applied as to how and when to interrupt a subject without fairly specific knowledge of the task they are trying to accomplish.

It seems to us that a much deeper understanding of how interruptions interact with the supposed "main tasks" of mobile workers, also an understanding of the effective strategies that are already being used by workers (both static and mobile, some of which may be highly task specific) is required before we put forward models for choosing an appropriate interruption time based on context. It also seems to us that, in the case of interruptions, context may well be simply insufficient to decide whether now is a good time to interrupt the user with any specific interruption - the content of the interruption and the nature of the task in hand may also be crucial.

References: -

O'Conaill, Brid and Frohlich, David: Timespace in the Workplace: Dealing with Interruptions. Proceedings of CHI'95, 1995.

Sawhney, Nitin and Schmandt, Chris. Nomadic Radio: Scaleable and Contextual Notification for Wearable Audio Messaging. Proceedings of CHI'99. ACM SIGCHI Conference on Human Factors in Computing Systems, Pittsburgh, May 15-20, 1999.

Ziljstra, Fred R. H., Roe, Robert A., Leonora, Anna B., Krediet, Irene: Temporal Factors in Mental Work: Effects of interrupted activities Journal of Occupational and Organisational Psychology. 1999 Jun; Vol. 72(2): 163-185

Situated Interaction in Art Settings

P. Marti, F. Gabrielli, L. Petroni, F. Pucci

{marti, gabriell, petroni, pucci}@media.unisi.it

University of Siena
Multimedia Communication Laboratory
Via dei Termini 6
I-53100 Siena

Abstract

This paper describes metaphors and design strategies applied to conceive and develop a hand-held location aware electronic guide for museums and cities that includes a number of situated and contextual aware interaction mechanisms. In particular we describe a design approach based on the situated visiting strategies and the contextual space *affordances* in order to support the user involvement in the activity and with the physical environment.

The work reports some of the results achieved within HIPS, a three-year project funded by the European Commission within the I-Cube (I^3) Program. This paper offers a contribution to the design of cognitive artefacts aiming at mediating the emotional involvement with physical and information spaces starting from the understanding of how the situation of use does influence the interaction process.

1. Introduction

The more sensing technology becomes emerging and offers new opportunities to include information about the situation and the context of use, the more the design of applications based on such systems needs new metaphors and interaction paradigms to fit user needs in new contexts of use. This is especially true for everyday leisure activities where the use of artefacts aims at mediating the emotional engagement with physical and information spaces.

A design contribution of the HIPS project[1] is based on understanding how the situation of use and the contextual variables influence the interaction process. This led us to design a cognitive artefact - a hand-held location aware tour guide (Benelli *et al*., 1999) - embedding the situated visiting strategies and the space *affordances* (Norman, 1988), in order to support the user involvement in the activity and with the physical environment.

The context of use for HIPS is a museum or a city. The addressed target of use is broad with heterogeneous needs and "*spaces for desiderata"*, differently from structured workplaces. Therefore the situations of use can be various and idiosyncratic leading the visitors to adjust frequently their goals and objectives during the visiting experience.

2. Situated, contextual and personalised interaction

In designing cognitive artefacts that mediate unstructured and emotionally-driven activities, it is fundamental to adopt a theoretical approach with the central focus on the real and potential activities situated in a specific context of use (Suchman, 1987, Norman, 1993). Indeed actions are always situated in particular social and physical circumstances, therefore the situation is crucial to action's interpretation (Suchman, 1987).

Considering the museum context it becomes clear that individuals generally do not anticipate alternative courses of action, or their consequences, until some courses of action are already under way. Moreover individuals often do not know ahead of time, or with any specificity, what future state they desire to bring about (Suchman, 1987). The visiting experience is a case in which individuals frequently have to "adjust" the way they interact with the environment, depending either on the action carried out or on the produced results. The shift in goal might be produced mainly by two modalities: i) the goal cannot be accomplished (lack of competence or physical constraints); ii) different states of the world are suggested on the basis of the performed activity (incoming information activate alternative patterns of knowledge) (Rizzo *et al*. 1997).

The ways in which individuals try to get control of interaction are contingent and derived from the situated action that they represent (Suchman, 1987).

Inspired by this theoretical approach, we focused our design activities on three main issues:

- **Context:** the context in which the activities occur is composed of natural, material, social and cultural components that affect the course of user' actions and interpretations during the interaction (Norman, 1988). By natural components we intend elements such as the lightening of the room. For instance, the Sala del Mappamondo, the experimental site for

---

[1] HIPS Consortium: University of Siena (I) - Project Coordinator, University of Edinburgh (UK), University College of Dublin (IR); IRST (I), GMD (D), SINTEF (N); ALCATEL-SIETTE (I); CB&J (F).

HIPS contains two famous frescos: the Maestà and the Guidoriccio. During the morning people firstly notice the Maestà that is better lightened by the natural daylight; in the afternoon, the Guidoriccio is preferred since the artificial light brighten this fresco (Gabrielli *et al.* 1999).

- **Situation:** every course of action essentially depends upon its material and social circumstances (Suchman, 1987). For instance, two different situations like a free exploration of a museum or a search for a particular museum content, need to be supported by a flexible tool able to consider different objectives associated to the situations.

- **Personalisation:** "a cognitive artefact is an artificial device designed to maintain, display or operate upon information in order to serve a representational function" (Norman, 1991). Therefore an artefact that mediates the fulfilment of an objective, providing the individuals with the appropriate information at the right time (without pretending to know completely their interests and preferences that could change during a situation of use), can simplify the nature of the activity and, in this way, enhance the overall performance. For instance, a system able to personalise the length and duration of the presentation according to different visiting strategies decreases the need for direct requests of more or less information.

This approach led us to structure the interaction design of HIPS taking into account contextual and situated needs and visiting strategies. The results we reached can be described at two levels:

a) Situation and Context Aware Interaction Mechanisms

The context and its physical, material, social and cultural components orient the exploration of the environment and the meeting with its content: in HIPS we exploited this concept at individual and social level.

Individual Level: when individuals move in a space they are driven both by intentional motivations (personal interests and preferences) and situated strategies but also by the properties of the environment. In the former sense, after ethnographic studies concerning the visiting behaviours, we conceived a "visiting style module" (Marti *et al.*, 1999), that linking the physical movements to the browsing of information space, provides personalised presentations according to the situated visiting strategies. For details about how the module is used for selecting and presenting information to the visitor see Not et al. (1998). In the latter sense, we explored the concept that visitors pathways mostly depend on natural and contextual *affordances* of the space, those properties that are «intrinsically» connected to a particular setting or that depend on the context of use. Hence, we experimented a design solution to augment the physical *affordances* of artworks by means of an auditory information space (Marti *et al.*, 2000).

Social Level: When a human being ties a knot in the handkerchief as a reminder, she constructs the process of memorising by forcing an external object to remembering her of something; *she transforms remembering into an external activity.* According to Vygotsky (1978), human memory is always generated by a social process. At the beginning the social memory is the product of knowledge distribution between the individual and the tools or the other individuals who are involved in a specific activity. Subsequently the social memory undergoes a process of internalisation by which external activities are reconstructed and the knowledge that is acquired through a social process is individualised and can be reused in different contexts.
HIPS is designed to support the process of embodying and transferring knowledge within a social group. A visitor can take a snapshot of a situation of interest (externalisation of knowledge) and then re-use it to suggest a friend to follow a tour, to elaborate on contents, etc. (embodiment and transfer of knowledge).

b) Language, Contents and Reading Styles

Contents are structured as small blocks of information dynamically combined in form of audio presentations (Not *et al*, 1997). This includes different types of contents and integrates contextual features (e.g.: deictics) in order to make more coherent the flow of narration with respect to the interaction of the user with the environment.

Moreover, we realised that the auditory output could be meaningful not only for the information delivery but also to support the physical navigation in the space: taking into account the existing indexical relationship between language and circumstances (Suchman, 1987), we conceived auditory contents that could provide cues about the surrounding physical space. Hence we are currently researching the effectiveness of different reading styles, integration of 3D sounds and music in order to design an auditory viscous space.

In the following, a scenario of use is provided in order to give a more concrete view of the whole system behaviour with respect to the design issues discussed above. Afterwards, we will go more into details describing the specific situated and contextual interaction modalities implemented in HIPS.

An actual scenario in the Museo Civico in Siena

Berthe and Samuel are two Belgian tourists. They move in the centre of the room handling HIPS, a portable guide. Berthe is attracted by the Maestà fresco.

HIPS: (with a male quite and polite voice) "*In front of you there is the Maestà, one of the absolute masterpieces of the Sienese art, depicted by Simone Martini in 1315 ".*

(a small pause, then a new male voice, with a strong Italian accent):*"The Virgin is depicted as Sienese people's protector, and as a symbol of municipal justice: this particular devotion to the Virgin derived from the famous Battle of Montaperti in 1260, when Siena defeated the army of Florence and preserved its freedom.".*

Berthe laughs because the last voice had a strange Italian accent. Then she moves toward Santa Caterina:

HIPS: (with a female first person voice) "*I'm Santa Caterina. I was born in Siena (…).*".

As she enjoys the comment, she takes a snapshot of the situation by pressing the hotspot button on their portable guide. The system continues to provide information about the S. Caterina's life, but Berthe skips this part.

HIPS: (with a male quite and polite voice) *"The portrait of Caterina is set inside a Renaissance-styled shell; it looks like a real sculpture (…)"*.

Berthe and Samuel move to leave the room.

HIPS: (with a male voice and 3D sounds effects) "*Behind you, there is another important fresco of Simone Martini: Guidoriccio da Fogliano*".

So, curious of this artwork, she decides to go back and stops to admire the Guidoriccio. Samuel follows Berthe and after few minutes they sit on a comfortable seat. They would like to visit the Pinacoteca but they don't know where it is. So, Berthe presses the Menu Button: by Find/Museums functionality she queries the system to know where the Pinacoteca is located and how to reach it. They go out the Museo Civico on the way to the Pinacoteca with the HIPS guide in their pocket.

The scenario exemplifies some basic concepts of situated interaction in HIPS.

1) The user is immersed in a rich audio environment. Different reading styles characterise the way in which artworks are described from different perspectives (historical, artistic, anecdotal descriptions).
2) The rhetorical styles are tailored to the context (use of deictic expressions) and to the iconographic contents (artworks representing people are described at the first person, as if the character presents himself/herself).
3) The rhythm of narration (length, duration) is tailored to the visitor's movement (long and detailed descriptions are provided to visitors who move slowly and stop in front of each artwork).
4) Experiential cognition is mediated by a natural input: the physical movement. Reflective cognition is allowed by intentional and context driven interaction (explicit queries to the system).

2.1 NEW INTERACTION METAPHORS THAT INCLUDE SITUATION AND CONTEXT (INDIVIDUAL LEVEL)

Physical spaces are not neutral. They make sense from the very moment we use them. One of the objectives of HIPS is to fill the gap between visitor's navigational strategies and information needs. From a technical point of view, this is realised by continuously monitoring the visitor's movements, thanks to a wireless connection between the portable guide (a PDA) and infrared emitters infrastructure. The museum space is the interface of the system, the physical movement is the main interaction vehicle.

Adaptation of input and output to the situation

(1) The gap between the physical and the information space is bridged by the visitors' behaviours. From ethnographic field studies in artistic exhibitions, Veron and Levasseur (1983) identified four categories of visitors (ant, fish, grasshopper, butterfly) based on their pathways, movements, and time of visit. This classification suggests how to isolate significant variables linked to physical movements and how to relate the physical movements to the browsing of information spaces. Starting from these results, we developed a «visiting style module» that, using an incremental bayesian algorithm, classifies users within the four categories and tailors the delivered information accordingly. In this respect, information can vary for length, duration and details.

(2) Visiting strategies can vary not only with respect to physical paths. In order to access information that is not directly related to a certain position in the space, HIPS provides an off-line browsing function that supports the access to external information. Even if this can resemble an ordinary way to browse an information space, it can be still considered a situated strategy. A goal

shift generated by a new incoming need (e.g. visit another museum containing artworks of the same historical period) requires support for planning a future course of actions.

(3) Deliberate control of the system behaviour is possible through the handling of simple and contextual buttons located on the PDA. These controls change labels and function according to the current task (cancel/confirm choices, stop/play audio comments, stop/more of this kind of information, cancel/select...). The design of these control buttons were inspired by the last generation mobile phones and "game-boy like" video-games.

Environment sensitive UI
Visiting strategies are not sufficient to exploit the idea of the environment as interface. *Affordances* of cultural settings play a central role in shaping the interaction. These include: a) properties that are «intrinsically» connected to a particular setting like the width of the artworks, their position, their artistic importance; b) architectural elements like access points to a room, arches and steps; c) dynamic and contextual configurations of elements present in the space (crowd, lights).
The role of the affordances in attracting the visitor, can be hampered when combined in certain configurations (crowd and bad light conditions often oblige the visitor to skip important artworks). We envisioned the possibility to design audio triggers to attract the visitor's attention. If the user reacts positively moving to the mentioned artwork listening to the description, then the system continues to provide information, otherwise it will just mention the artwork without further elaboration.

2.2 NEW INTERACTION METAPHORS THAT INCLUDE SITUATION AND CONTEXT (SOCIAL LEVEL)

From Vygotsky, we adopt the assumption that the social memory is the product of knowledge distribution between the individual, the tools and the other individuals who are involved in a specific activity. The social memory develops from the externalisation of knowledge through its internalisation and recombination for later use in different activities.
HIPS provides some very basic supports to the development of a social memory in the community of visitors. The externalisation of knowledge is realised by bookmarking a moment of the visit (pressing the "hotspot" button on the PDA, the visitor stores into the system the current position, an image of the artwork, the related description, a personal comment). This knowledge is available for later use to suggest a friend to follow a tour, to elaborate on contents, to plan another tour etc. (embodiment and transfer of knowledge).

2.3 SITUATION-AWARE CONTENT

(1) The audio descriptions in HIPS are segmented in *Macronodes* (Not *et al*., 1999), small blocks of information that are dynamically combined to form an audio presentation. Each of them contains different kinds of contents with explicit deictic reference to the physical position. The flow of narration is made more fluid and harmonised to the context of visit.

(2) The use of different reading styles, the integration of 3D sounds and music are means to design rich audio environments. HIPS aims at creating a sort of "empathic effect" mediated by human voices and immersive information spaces to engage the user in an intense meeting with art.

3. Conclusions

The research domain described in the paper raises challenging issues about the concepts of interaction. People mainly interact with a rich and stimulating environment like a museum for intellectual and aesthetic pleasure. This activity is not structured: visitors move in the physical space guided by their interests, stimulated by the context and adapting their choices to the contingent situations of use. In order to avoid breakdowns in the flow of the activity, the boundary between the physical space and the information space should be seamless. The tool that supports the visit should neither intrude the activity nor, and most critically, require a deliberate cognitive effort to be used. In this way the cognition can flow from the external environment to the interior world of interests and emotions in a transparent way. Such «efficiency» of action is reached when the artefact becomes part of the activity, a sort of invisible aspect of our experiential world.

References

Benelli, G. Bianchi, A. Marti, P. Not, E. Sennati, D. (1999) "HIPS: Hyper-Interaction within Physical Space", Proceedings of IEEE99, Firenze, Italy.
Gabrielli, F. Marti, P. Petroni, L. (1999) "Environment as interface". In Caenepeel, M. Benyon, D., Smith, D. "Community of the Future", Proceedings of the I3 Annual Conference, Siena.
Marti, P. Rizzo, A. Petroni, L. Tozzi, G. Diligenti, M. (1999) "Adapting the museum: a non-intrusive user modelling approach", Proceedings of UM '99, Canada.
Marti, P. Rizzo, A. Gabrielli, F., Petroni, L., Pucci, F. (2000) "Designing emotional artefacts", submitted to AVI 2000.
Norman, D.A. (1988) "The psychology of everyday things", New York: Basic Books.

Norman, D.A. (1993) “Things that make us smart”, Reading, MA: Addison-Wesley.
Norman, D.A. (1991) “Cognitive artifacts” in Carroll J.M. “Designing Interaction”, Cambridge University Press, Cambridge.
Not, E., Petrelli, D. Sarini, M. Stock, O. Strapparava, C. Zancanaro, M. “Hypernavigation in the physical space: adapting presentation to the user and to the situational context”. The New Review of Hypermedia and Multimedia, vol. 4, 1998.
Not, E., Petrelli, D. Sarini, M. Stock, O. Strapparava, C. Zancanaro, M. "Hypernavigation in the physical space: adapting presentation to the user and to the situational context. The New Review of Hypermedia and Multimedia, vol. 4, 1998.
Not, E., Zancanaro, M. “Reusing Information Repositories for Flexibly Generating Adaptive Presentations”. In Proceedings of IEEE International Conference on Information, Intelligence and Systems, Washington DC, November 1-3, 1999.
Rizzo, A., Marchigiani, E;, Andreadis, A. (1997) “Prototyping and evaluation with a Cognitive Walkthrough based on the Norman's model of action”. Designing Interactive, Systems. New York: ACM Press.
Suchman, L. (1987) “Plans and Situated Actions. The problem of human machine communication” Cambridge University Press, Cambridge
Veron E.& Levasseur M. (1983) "Ethnographie de l'exposition", Paris, Bibliothèque publique d'Information, Centre Georges Pompidou.
Vygotsky, (1978) L. “Mind in society” Cambridge MA: Harvard University Press.

Position paper for the CHI 2000 workshop on 'Situated Interaction in Ubiquitous Computing' at the conference in The Hague, Netherlands, April 3rd, 2000.

What About the Situation at the Other End(s)?

On the Multifariousness of Situated Interaction in Ubiquitous Computing

Sara Eriksén
Department of Human Work Science,
University of Karlskrona/Ronneby,
Sweden

This position paper briefly presents a problem area that I have become increasingly aware of through several research projects during the last few years. Basically, it concerns the multi-situated nature of interaction. The 'inter' of 'interaction', after all, would not be relevant if there were not at least two connected instances of situated action involved. In my research, I find that the challenge of the multifariousness of computer-aided interaction to the design of information and communication technology (ICT) is becoming more and more apparent, as computers become more and more ubiquitous. Yet we still lack methods, models and metaphors to really grasp the problem and begin to deal with it effectively in design.

Using ethnographic field methods, including video recording and interaction analysis, I have focused, in my research, on how modern information and communication technology is used in public service one-stop shops. The aim has been to explore how ICT can be designed to successfully support communication, cooperation and 'knowing in action' in the front-office work of public administration.

During the later part of the nineties, as the use of Internet/intranet applications developed and spread, I began to focus on the possibility of integrating public information systems on-line with the continued design and development of advanced ICT for people working within public services. Ideas and inspiration have come partly from the expanding call center business.

What is happening now is, that the diversity of ways in which computer-supported interaction can take place is growing very rapidly, while the design and development of applications to efficiently support and make use of this diversity seems to be slow in the uptake.

How can we represent the multitude of different ways, and combinations of different ways, in which computer-aided interaction can take place today - from both ends/all directions? How can we make efficient use of this variety?

My main focus is on call centers and inter-active public services, but I would really like to share in the experiences of other researchers who have been working with a more direct focus on mobile and ubiquitous computing. What kinds of metaphors are developing with which to grasp the multi-situatedness of inter-action? This is what I would like to learn more about, and discuss, at the workshop.

Bibliography and Background

I am at present a lecturer in the department of Human Work Science, at the University of Karlskrona/Ronneby in southern Sweden. Here, I teach within an interdisciplinary Master's program called *People, Computers and Work* (*MDA* is the Swedish acronym), combining Computer Science and Human Work Science in educating systems developers for the future. During the past years, I have been involved in research projects focusing on the use, design and continual support and development of computer support for public administration in one-stop shops, and the on-going integration of such systems with public electronic information systems (Eriksén, 1998, 1999). These projects were financed by the Swedish Council for Work Life Research. During 1996-98, I also participated in the EC project ATTACH (Advanced Trans-European Telematics Applications for Community Help, UR 1001), in which the University of Karlskrona/Ronneby was a partner.

The University of Karlskrona/Ronneby was founded in 1989. It is a young and small, but rapidly expanding university, with approx. 3,000 students and 330 employees. The main emphasis in both research and teaching is on ***IT in use*** – i.e. on information technology and how it is used.

References

Eriksén, Sara (1998): *Knowing and the Art of IT Management. An inquiry into work practices in one-stop shops*. Ph.D. thesis, Lund, Sweden: Department of Informatics, Lund University.

Eriksén, Sara (1999): *Globalization and 'Genius Loci'; From Intentional Spaces to Purposeful Places*. Working paper for UK-/Nordic workshop held in Ronneby April 15th-16th 1999 on the theme 'Social Theory and ICT'. Website http://www.brunel.ac.uk/research/virtsoc/nordic/eriksen.htm (99-07-15).

Name: Sara Eriksén /Email address: Sara.Eriksen@iar.hk-r.se
Organization: University of Karlskrona/Ronneby
Department: Department of Human Work Science
Address: Soft Center
Zipcode: SE-372 25
City: Ronneby
Country: Sweden
Phone: +46 457 38 55 65
Fax: +46 457 271 25

Innovative User Interfaces that use Mobile Devices at the Same Time as PCs

Proposal to Attend the CHI'2000 Workshop on
Fehler! Textmarke nicht definiert.

Benjamin Bostwick
Brad A. Myers
Rob Miller

Human Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3891
(412) 268-7565
FAX: (412) 268-1266
bb46@andrew.cmu.edu**Fehler! Textmarke nicht definiert.**

There are many signs of the approach of the ubiquitous computing era **Fehler! Textmarke nicht definiert.**. People are carrying Personal Digital Assistants (PDAs), such as Palm Pilots or Windows CE devices, and therefore have computing with them everywhere. "Smart Environments," where computing is embedded in offices and homes, are becoming a reality. Already, most conference rooms and classrooms at Carnegie Mellon University (CMU) have a built-in computer with a projector. And of course, every professor and most students have at least one computer on their desks in their offices. Most companies are similar, since virtually every white-collar worker today uses a computer. Homes are just starting to be set up with embedded computing, and the "Smart Home" has been widely anticipated **Fehler! Textmarke nicht definiert.**.

One aspect of ubiquitous computing that has not been adequately studied is how a user will use *multiple* devices *at the same time*, so the devices work seamlessly together. Most of the research and development about hand-held and mobile computers has focused on how they can be used to *replace* a personal computer (PC) when the PC is not around. The conventional model for PDAs is that the data is "synchronized" with a PC once a day using the supplied cradle, and otherwise the PDA works independently. This will soon change. For example, CMU has installed a Lucent Wavelan wireless network (using the 802.11 protocol) throughout the campus, in a project called "Wireless Andrew" **Fehler! Textmarke nicht definiert.**. Many Windows CE hand-held computers can be connected to this wireless network using a Wavelan PCMCIA card. Other mobile devices will support this protocol soon. Next year, the BlueTooth standard for small device wireless radio communication will finally be available, and most PDAs, cell-phones, and other computerized small devices are expected to support it. Therefore, we expect that connecting the PCs and hand-helds together will no longer be an occasional event for synchronization. **Instead, the devices will frequently be in close, interactive communication.**

The Pebbles research project (**Fehler! Textmarke nicht definiert.**) has been studying the implications of this trend, especially on how the functions, information and user interfaces can be shared across multiple devices in use at the same time. For example, there are many ways that a PDA can serve as a useful adjunct to a personal computer to enhance the interaction with existing desktop applications. New applications may distribute their user interfaces across multiple devices so the user can choose the appropriate device for each part of the interaction. A key focus of our research is that the hand-held computers are used both as output devices and as input devices to *control* the activities on the other computers. The following scenarios illustrate some of the capabilities we are already investigating:

- The presenter of a talk has a laptop where the display is projected onto a large screen. The laptop's powerful processor is needed to control the animations and external applications that are part of the presentation. (Or similarly, the meeting room has a built-in computer with a data projector to control the presentation.) The presenter walks in with a hand-held PDA. The laptop communcates with the PDA, so on the PDA is displayed the current slide's notes. Gestures on the PDA cause the presentation to go forward, backward or skip to a specific slide under discussion. Also on the PDA are custom controls to switch among various other applications on the laptop which the presenter will be demonstrating and discussing. Each member of the audience of the presentation has carried in their mobile hand-held computer, and each attendee sees on their personal hand-held the current slide, which can be kept synchronized with the talk. Each person can also make private notes and annotations on the hand-held. When enabled by the presenter, an audience member's marks on the PDA will display on the main screen for general viewing and discussion. Various techniques are used to coordinate the actions of multiple people who are interacting with the displayed content **Fehler! Textmarke nicht definiert.**.

- When the user is sitting and working at home or in the office, various mobile devices are carried in and laid on the desk: a laptop, a PDA, a cell-phone, etc. Some of the newest hand-helds such as the Palm V and the HP Jornada 430 have rechargeable batteries that are recharged when the device is in its cradle. Therefore, the user is *supposed* to have the device connected to the PC whenever the user is next to the PC, so it might as well be doing something useful. The wireless devices might detect that they are in the context of the main PC, and reconfigure themselves to serve as adjuncts of the PC's applications. We envision having the cradle right next to the PC on the left side of the keyboard, while the mouse is on the right. For example, as the user works, various controls might appear on the screens of the PDA and on the other devices rather than on the desktop computer's screen. Our research has shown that users can use their left hand on the PDA to scroll documents shown on the desktop's screen just as fast or faster than using the regular GUI scroll bars or using devices such as the scroll wheel built into some mice **Fehler! Textmarke nicht definiert.**. The user's custom shortcuts for the laptop applications also appear on the PDA, and the user has memorized their location and can operate them quickly without looking. Information can be easily moved among the devices, and other information is automatically distributed based on predefined user preferences.

- In a "Smart Meeting Room," the main large displays show the shared displays under discussion by the group. Built into in the room are cameras, microphones, and displays, so users can speak, gesture in the air, as as using conventional interaction techniques with keyboards and mice. Individuals enter and leave the room, each carrying their own PDA. While in the meeting room, someone might want more details on an item displayed on a large display. Rather than disrupting the main activities and the main display, the PDA can be pulled out, and pointed at the item on the main display. Then the user can privately have the additional specialized information displayed on their PDA. The display of the information is appropriately adjusted to the limited size of the PDA screen. We are also investigating ways to fluidly move interactions among the modalities, so users can start speaking and gesturing in the air (interpreted by cameras and microphones), and then dynamically switch to handwriting and tapping on the PDA.

- In a "Smart Classroom," the students' hand-held devices should detect what classroom they are in, and immediately configure the hand-held for today's lesson. This might include an in-class test, in which case it should automatically appear on the hand-held. When the instructor starts the test, the network might be configured to prevent students in this class from sending messages to each other, or to access the global world-wide web, but they still must be able to use the network to access the instructor's server to submit the test.

There are many significant research issues involved in bringing these visions to fruition, which we are investigating. We are particularly interested in the appropriate ways to distribute the user interfaces across multiple devices, how to support multiple people interacting with the *same* screen using their various devices as auxiliary input and output devices (which is sometimes called "single-display groupware" **Fehler! Textmarke nicht definiert.**), the automatic creation of appropriate and usable control panels from high-level specifications, and usability issues with multi-device interaction techniques.

The Pebbles research project has made substantial progress by building example applications, releasing them for general use, and formally testing them in usability experiments. Several of our existing applications support meetings where the participants are co-located. All participants' PDAs are in continuous two-way communication with the main computer which is often projected on a screen to serve as the focal point of the discussion. Some of our initial applications use the PDAs as remote mice and keyboards so that everyone in the meeting can control the main computer. The PDA might be used to control a PowerPoint presentation while displaying the slide notes and titles on the PDA, as a shared whiteboard that supports multiple inputs simultaneously, for private side messages via a "chat" program, and to display multiple cursors for pointing and scribbling on arbitrary applications **Fehler! Textmarke nicht definiert.**. We are currently investigating a number of groupware issues, including appropriate floor control mechanisms, and how to fluidly move information between the public and private displays. Another set of applications supports a single person using the PDA as an extra input and output device to enhance desktop applications. The PDA can be used as a scrolling device, as a general-purpose button panel (to create screens of "shortcuts"), as an index page or table of contents for web surfing, and to cut and paste information back and forth from the PDA to the PC. These applications have been downloaded over 15,000 times already, and are available from **Fehler! Textmarke nicht definiert.**.

Benjamin Bostwick of the Pebbles research project would like to participate in this workshop because we are interested in discussing how the devices should adapt to the situation and context, especially in terms of which other computing, input, and output devices are available in the area. If there is a big display on the wall, what parts of the user interface should migrate there? But what if someone else is using the big display for a side discussion? The consideration of context would enhance our work on the use of multiple devices at the same time. Conversely, considering the use of multiple devices simultaneously will expand the range of issues in situational context.

References

[Hills 1999] Hills, A., “Wireless Andrew.” *IEEE Spectrum*, 1999. 36(6)June.

[Levy 1999] Levy, S., “The New Digital Galaxy; A Really Smart House.” *Newsweek*, 1999. (22)May 31. http://newsweek.com/nw-srv/issue/22_99a/printed/us/st/scc122_1.htm.

[Myers 1998] Myers, B.A., Stiel, H., and Gargiulo, R. “Collaboration Using Multiple PDAs Connected to a PC,” in *Proceedings CSCW'98: ACM Conference on Computer-Supported Cooperative Work*. 1998. Seattle, WA: pp. 285-294.

[Myers 2000] Myers, B.A., Lie, K.P.L., and Yang, B.-C.J. “Two-Handed Input Using a PDA And a Mouse,” in *Proceedings CHI'2000: Human Factors in Computing Systems*. April, 2000. The Hague, The Netherlands: To Appear.

[Stewart 1997] Stewart, J.E. “Single Display Groupware,” in *SIGCHI'97 Adjunct Proceedings: Human Factors in Computer Systems*, Extended Abstracts. 1997. Atlanta, GA: pp. 71-72.

[Weiser 1993] Weiser, M., “Some Computer Science Issues in Ubiquitous Computing.” *CACM*, 1993. 36(7): pp. 74-83. July.

Better Living Through Geometry

Barry Brumitt, Steven Shafer
{barry, stevensh}@microsoft.com
Microsoft Research

Mark Weiser described ubiquitous computing as, "invisible, everywhere computing that does not live on a personal device of any sort, but is in the woodwork everywhere."[8] The EasyLiving project at Microsoft Research is focused on those aspects of ubiquitous computing relevant to smart environments, including work in distributed computing, geometric world modeling, computer vision, and user interfaces. Though the need for research in distributed computing, perception, and interfaces is widely recognized, the importance of an explicit geometric world model for enhancing the user's experience of a ubiquitous computing system has not been well-articulated. This paper introduces three scenarios which benefit from geometric awareness, examines three existing mechanisms for providing geometric knowledge, and then describes the EasyLiving Geometric Model. In particular, the focus is on improving the user experience of systems which are comprised of many independent devices and heterogeneous perception technologies.

Introduction

The goal of the EasyLiving research project[5] is the development of a prototype architecture and necessary technologies for intelligent environments. EasyLiving concentrates on applications where interactions with computing can be extended beyond the confines of the current desktop model. Such a computing system should maintain an awareness of its users, understand their physical and functional relationship to I/O devices, respond to voice and gesture commands, and be easily extended. This technology will, for instance, enable a home's resident to make a phone call by simply speaking his intentions from wherever he happens to be. It will allow a user to move from room to room while still maintaining an interactive session with the computer or a particular application. All these tasks require the coordination of many devices for computational activities involving both perception (Where is the user now?) and interaction ("Call Bob.")

In a space populated by many small, networked computing devices, several devices will typically have to work together to perform a particular task. Dynamically collecting a group of smart devices to enable an interaction or to perform a perceptual task requires a shared computational substrate that allows the devices to communicate bits which represent concepts in a shared ontology. For example, three devices might announce to each other over a wireless network: "I am a display device", "I am a DVD Player", and "I am an acoustic speaker." Once these capabilities are known to exist in the same place, the ability for a user to play a movie should be enabled. If a pair of headphones with appropriate capabilities were to enter the fray ("I'm a pair of headphones"), redirecting the sound output to them should become an available option.Where these options are displayed, how the user is informed of these options, and how much autonomy the system has remain open questions.

This paper introduces three scenarios which benefit from geometric awareness, examines three existing mechanisms for providing geometric knowledge, and then describes the EasyLiving Geometric Model. In particular, the focus is on improving the user experience of systems which are comprised of many independent devices and heterogeneous perception technologies.

Scenarios

Location-aware computing services have been proposed for many tasks, including providing driving directions, redirecting phone calls to the phone nearest the recipient, reminding the user of errands appropriate to his location, or even just turning off the lights when the user leaves a room. Most such systems assume a straightforward connection between the sensor providing the position information and the application. For example, the GPS position can be used along with a contact list and address book to deliver a reminder to pick up the dry cleaning when the user is near the store. The problem of performing location-based services in a more general framework, where there are multiple sensors and devices is much more challenging. A further complication is the incorporation of contextual information about the state of the physical world, beyond simply the user's position.

The following three scenarios describe how the user's experience of a ubiquitous computing system is improved by the use of "geometry-aware" system services. Geometry-awareness is distinguished from location awareness in two ways:

- heterogeneous perception technologies, e.g cameras, GPS, and beacons
- an understanding of physical relationships between things in the world, e.g. walls block users visibility

1. Contact with Context: Physical parameters for User Interfaces

Consider the task of contacting someone who is at work, such as to remind him of an important meeting. Which of the many devices in his office (displays, phone, pager, cell phone, computer speaker, stereo, PDA) should be used? One approach would be to use some fixed preference-based scheme, flashing the screen, ringing his phone, and finally paging him. However, there is no point in using a visual signal if he is not looking at the screen, or ringing his phone if he is wearing headphones. A better approach is to understand the location of the person, his physical relationship to devices which are around him, and the various consequences of the current state of the world. In this example, by examining the set of devices which are near the user (phone, pager, screen, speaker), examining the state of the world (facing away from the screen, pager inside briefcase, currently using phone), the system service which delivers the message can select the remaining option (speaker) to get the message to the user in the most expeditious manner.

The message-delivery example above illustrates the ability of geometric knowledge to provide physical parameters for the user interface. While geometric knowledge alone is insufficient to select the best device for a given interaction (context and other world knowledge are helpful), it is a necessary component for reaching the ideal decision.

2. Device Aggregation: Simplified device control

Currently, the PC is the integration point for a cluster of I/O devices which provide the majority of computer/information services. The wired mechanical connection of all the devices (processor, hard drive, display, speakers, mouse, keyboard, etc.) implies that this cluster of devices is intended to work together. What happens, though, when the devices lose this mechanical connection, as is the case with ubiquitous computing? If each device is an independent entity which connects to the network, some system infrastructure must exist to naturally pull together disparate elements to form a usable aggregation of devices.

Imagine your living room, equipped with a panoply of devices, including a couple screens, a wireless keyboard/remote, a sound system, a camera, etc. When you want to initiate a web browsing session, you would have to switch some screen to an appropriate mode using a specialized remote, manually direct the keyboard output to that screen, login as yourself (so you'll have your cookies, preferences, etc.), tell the screen to enter web browsing mode, and manually redirect the audio to come from the PC. However, with geometric awareness, picking up the keyboard could implicitly log you in as yourself, and bring up an appropriate UI on a screen which you can see from your current location, and set up all the other parameters appropriate for your session. Additionally, the ability to dynamically change devices, such as having your session to follow you as you move, is enabled as long as exists some sensor which can provide the current state into a geometric model. Without geometric awareness, all device-to-task coupling must be performed manually, a task whose complexity will grow exponentially with the number of interconnected, networked smart devices.

3. Let there be light: Shared world model

Consider the simple task of turning on a light in an intelligent environment. Here are seven ways this task might be completed:

- Manual: Flip a wall switch.
- Traditional GUI: Use a dialog box with list of lights and buttons for on/off.
- Physically-enhanced GUI: Select from a map of house/room with lamp indicators.
- Direct Speech: Say "Turn on the living room lights."
- Gesture: Make a funny gesture, observed by a camera, indicating a need for light.
- Indirect Speech: Say "I could use more light."
- Implicit Request: Sit down at comfy chair while holding a book.

Some of these examples can be handled by existing technology and do not require any geometric awareness, or, in the case of "Manual", even any computing technology at all. However, if the user and the computing system share some understanding of the physical world the ability to support a wider range of interactions becomes possible.

The "Physically-enhanced GUI" and "Direct Speech" use the least geometric knowledge. They only assume some simple shared map and nomenclature in order to enable requests. This requires the user to understand the building she is in, and the appropriate names of the various locations. Note that some tasks remain tough in this paradigm: how does one precisely name and differentiate all the various lights in one's home? "Gesture" and "Indirect Speech" can use geometric knowledge to turn on the lights with more fidelity. Presumably, the user wants light to appear where she currently is located. By having a model of the location of the lights, and a perception system to track the user, the request can be more accurately fulfilled. The increase in capability of these interactions is directly related to the fidelity of the shared metaphor possessed by the computer and the user. The greater the shared understanding, the more robust the interactions can be.

Finally, notice that only the last option ("Implicit Request") does not involve direct explicit action by the user to request the system to turn on a light. The goal of geometric awareness is not just to provide automatic actions, but more importantly, to improve the shared model between user and computer, resulting in a more intuitive and natural user interface. We expect people to understand the consequences of their actions, and to interpret our implicit requests - why shouldn't we expect the same of our computers?

Three technology-driven notions of geometry

The above scenarios all describe possible ways in which geometric awareness can be used to support better situated interaction in ubiquitous computing. However, most existing systems which provide location information typically couple the mechanism for gaining position information directly to the services which are provided. In a more general ubiquitous computing environment, perception technology would be largely independent from the service - it doesn't matter how the computer knows the relationship between the person and the screen, only that the UI be able take this information into account when displaying information.

This section describes three extant paradigms for providing and utilizing geometric information in ubiquitous computing systems.

1. Latitude, Longitude, Elevation

Many proposals for location-aware systems rely heavily on position information coming directly from a GPS receiver, providing location with a resolution of approximately 30m. For some applications, differential GPS can be used to reduce the error to the sub-meter level; however, this requires significantly more hardware to achieve and constrains the system to operate within range of a differential-corrections transmitter[2]. Typical scenarios supported by such systems include using location to determine driving directions, deliver reminders based on the user's location[3], and record scientific data tagged with location information[4].

This notion of geometry is useful only in outdoor situations that are outside of major urban centers. In cities, the tall buildings can frequently obscure satellite visibility, much as when the antenna is indoors. While this does not preclude the usefulness of GPSs (and latitude/longitude measurements in general) for some scenarios, it implies other technologies for obtaining position information are needed. The above scenarios, for example, would all be impossible to achieve using only GPS as a positioning technology.

2. Beacons, Badges, Transceivers, & Tags

An alternative location-determining technology proposed primarily for indoor use is Active Badges. These systems can provide information about which room a particular tag is in[6], or even the particular position of the tag inside the room[7]. In general, these systems consist of RF, IR or ultrasonic transceivers which can determine the presence and perhaps location of small (usually powered) tags which are attached to objects of interest in the world such as people, phones, printers, computers, etc. These systems represent geometry as a location in a single coordinate frame, such as a map of the building. They require installation of a large number of transceivers throughout the building, and assume that all information of interest will be directly expressed in the single geometric model.

While a complete system to perform Device Aggregation scenario described above could be built by attaching tags to the devices (display, keyboard, speakers, user, etc.) and installing beacons in the rooms, there are two drawbacks. First, all items of interest must be tagged. If you are not wearing your badge (e.g. at home right after getting out of bed) then no position-based services are available to you. Secondly, and more significantly, if other positioning technologies are available, they cannot be readily integrated. Active badge systems are useful for providing positional information, much as is GPS, but current systems lack a general geometric model for expressing arbitrary geometric information.

3. Network Address != Physical Geometry

To avoid the perils perception altogether, one can assume that network or data connectivity is equivalent to co-location. This implies that if two devices can communicate directly (by RF, IR or other "local" transmission method), they are co-located. However, RF transmission (not to mention physical network protocols) can easily span rooms, floors or even buildings. Without some more precise model of geometry, this type of assumption will result in a excessively large set of potentially available devices, many of which may not actually be available or usable for any particular task. All devices are on the network (for some suitably comprehensive definition of network), and yet, the correct selection depends on knowing the physical relationship between the user and these devices. Worse yet, relying upon this alone can destroy the shared metaphor perceived by the user. For example, if a new reading lamp is plugged into an network/power outlet in the kitchen, but is routed around the corner into the den, asking for "lights in the den" to come on will not achieve the expected result. The user could be expected to manually configure the physical location of all devices, but this is a non-negligible burden. Additionally, consider the problem of any RF device relying on connectivity to determine location. If an RF remote control could control multiple devices in the same house, which device should it be activated when a button is pressed? If both devices can receive the signal, then either the user must resort to manual selection or confusion results.

The EasyLiving Geometric Model

The EasyLiving Geometric (EZLGM) model provide a general geometric service for ubiquitous computing, focussing on in-home or in-office tasks in which there are myriad I/O, perception, and computing devices supporting multiple users. The EZLGM is designed to work with multiple perception technologies and abstract the application and its user interface away from perception.

The base item in the EZLGM is an entity. An entity may represent an object or location in the physical world. Measurements are used to define geometric relationships between entities. In particular, a measurement describes the position and orientation of one entity's coordinate frame, expressed in another entity's coordinate frame. For two measurements to involve the same frame, both measurements must have been made by sensors with an implicit understanding of the origin of the frame on the entity; in other words, both must use a particular point and orientation on the entity when making an observation. Since objects in the physical world have some physical extent, this can also be expressed in the geometric model. If one physical object has different components which can be independently measured (e.g. a laptop with both screen and keyboard), then it could be represented as two entities.

Once a set of measurements has been provided to the geometric model the model can be queried for the relationships between entities' frames. The measurements describe an undirected graph, which each vertex as an the frame of an entity, and each edge a description of the (invertible) geometric relationship (including an uncertainty estimate) between the coordinate frames. If at least one path exists between two frames, then the graph can be processed to produce a single geometric relationship between the frames. The replies to these queries are based on previously provided measurements. Since a particular queried relationship may not have been previously directly measured, the response is frequently involves the combination of multiple measurements; uncertainty information is used to accurately merge multiple redundant measurements as needed. When querying about the relationships between entities, it is frequently helpful to be able to refer to the extent of such an entity, such as the field-of-view of a display device.

To get a sense of how this model might be used, consider the Contact with Context scenario above. A process which wants to contact a user might query EZLGM for all devices which have service areas which intersect with the location of the user. It could then look at types and availability to determine the set of devices which might provide the messaging service, and further prune the list by considering the physical constraints, like visibility, etc., in order to reach a set of usable, available, and physically-appropriate devices. Visibility can be checked by examining all entities along the line of sight between the user and the device and ensuring none represent something which would physical block the view. Then, by consulting the users preferences, or by using internal heuristics and other context information, the signal could be sent via the appropriate mechanism. The geometric model provided a way for storing both the devices that could be used, and for aiding in the determination of the appropriate device. Note that no part of this example required any reference to the perception method which provided information about position: it could have been performed via cameras, a badge system, etc.

Conclusions

This paper has described three important scenarios for situated interaction in ubiquitous computing systems, "Contact with Context", "Device Aggregation" and "Extensible Computing." Through those scenarios, three primary benefits of geometric models have been introduced:

- **Physical parameters for UI's:** Device selection and control is performed with knowledge of physical context.
- **Simplified device control:** Device aggregation for a task is performed without requiring step-by-step user action.
- **Shared Metaphor:** User experience is simplified through a common understanding of the physical world shared by system and user

Existing systems which are location-aware provide applications which are tightly coupled to the perception mechanism. For location-awareness to be more generally applicable, abstracting the use from the gathering of such knowledge is essential. The EasyLiving Geometric Model is a first step towards a general sensor abstraction layer.

While geometric knowledge remains challenging to gather, represent, and provide, its inclusion will significantly improve user experience of ubiquitous computing systems

References

1. Dey, A. et al. “A Context-Based Infrastructure for Smart Environments”, Managing Interactions in Smart Environments, Springer-Verlag, 1999, pp. 114-128.
2. “DGPS FAQ”, http://www.navcen.uscg.mil/faq/dgpsfaq.htm
3. Rhodes, B. “The Wearable Remembrance Agent: A system for augmented memory”, Personal Technologies Journal Special Issue on Wearable Computing, Personal Technologies, 1997, Vol. 1, pp. 218-224.
4. Ryan, N.S. et al., “Enhanced Reality Fieldwork: the Context-aware Archaeological Assistant”, in V. Gaffney, M. van Leusen and S. Exxon (eds.) Computer Applications in Archaeology 1997, 1998.
5. Shafer, S. et al.“The New EasyLiving Project at Microsoft Research”,*Proceedings of the 1998 DARPA / NIST Smart Spaces Workshop*, July 1998, pp.127-130.
6. Want, R. et al. “The Active Badge Location System.” ACM Transactions on Information Systems, 10(1):91-102, January 1992.
7. Ward, A. et al. “A New Location Technique for the Active Office”, IEEE Personal Communications, Vol. 4, No. 5, October 1997, pp. 42-47
8. Weiser, M. “Ubiquitous Computing”, http://www.ubiq.com/hypertext/weiser/UbiHome.html

What are a Location's "File" and "Edit" Menus?

Andrew Fano
Center for Strategic Technology Research
Andersen Consulting
afano@cstar.ac.com

Soon we will trade stocks in the park and receive faxes on the beach. While this is all very nice, the real promise of mobile devices lies not in enabling you to drag your desk wherever you go. We go to various places – offices, bowling alleys, airports, Laundromats, restaurants, etc - to do the things that make sense there. Our research is guided by the core idea that mobile devices ought to be about making the things we do at these places easier, rather than trying to be all things to all people. Therefore we are developing mobile device interfaces that enable location-specific tasks. In particular, we are exploring how general mobile devices can be used in a situated context to provide access to the services and resources of a particular location. Broadly speaking, we view the mobile device as a "personal remote control to the world".

Most places tend to have whatever equipment they need to support the activities that occur there. Bowling alleys have ball return machines, doctor's offices have medical equipment, etc. What can a person's general mobile device add that isn't already handled by the location's task-specific equipment? Since people are going to have these increasingly powerful devices anyway, regardless of their value at any one location, it makes sense to see how we can use them to make the things we do at a given location easier. Consider that cars aren't strictly *necessary* to get money from a bank or eat a hamburger at a restaurant. However the fact that most people have cars has led, in many cases, to changes in the way we pursue tasks to take advantage of our cars, as evidenced by drive-through facilities at banks, fast food restaurants, and drug stores among many others.

We are designing applications and services for mobile devices in support of location-specific tasks by exploiting three primary capabilities. First, we see the device as a *persistent, rich channel to an individual.* That is, the device serves as a way to deliver services to the individual. Secondly, mobile devices serve as *user context detectors.* Mobile devices not only contain information about the owner such as personal information, schedule, task-specific data, but also, in the long run, sensors capable of detecting aspects of the user's environment. See Schmidt, et. al, [1999] for an example real-time architecture that derives contextual cues from sensor data. Such sensors might detect features including position, sound, video, temperature, and location contents, among others. These "bottom-up" contextual cues combined with the "top-down" constraints offered by the location's task and the user's input can be used to determine the information to be presented to users and the modality through which it should be presented. Lastly, we see the mobile device serving as a *remote control to the environment.* As mobile devices become as intimate as our wallets, their interfaces will become second nature, and, importantly, feel more familiar than most others. Consequently, we believe that these devices may begin to serve as effective "remote control" interfaces to other devices in the environment. See, Beigl's work on point and click interfaces for control devices in the environment [1999] as an example of work in this direction. While we are also interested in facilitating the use of devices in the environment, we are structuring the interface primarily around the location and its task, rather than on specific devices in the location.

Interfaces for Physical Locations

Designing mobile applications to support location-specific tasks requires an interface that reflects the natural structure of tasks typically performed in various locations. While location-specific tasks will have some functionality that does not generalize, there are, nevertheless, aspects of tasks that do apply across domains that can be used as a starting point to design and structure interfaces. This is much like the "file" and "edit" menus we find on most general computer software applications. While the particular functions a user can invoke necessarily vary from application to application, there are certain general functions that make sense in virtually all applications. Such functions include many of the key features in the File and Edit menus, such as "new", "open", "close", "save", "print", "cut", "paste", "delete", etc. We want to define the analog of these "file" and "edit" menus for locations. That is, we seek to provide a common, intuitive interface to task functions likely to be found across many locations.

What are these common task functions? Consider what tends to happen when we go to places like offices, restaurants, ballparks, bowling alleys, airports, theaters, hotels, train stations, etc. Common steps often include the need to *find the place, park, register, say what we want, buy admission, contact the "host", find any people with whom we will be collaborating on the task at hand, check the readiness of the specific site where we will pursue our task, find that site, find and check the status of any tools we will need, identify who can help us, find a place to buy coffee, do work, and go to the bathroom while waiting for the task to begin, perform the task, cleanup, and get to the next task.*

Naturally we don't intend for mobile devices to actually implement all of these functions any more than, for example, we would expect an orchestra conductor to play all instruments. Instead we want to use the mobile device in the roles described earlier – as a *remote control* to invoke many of these environmental functions and services, as a rich *channel to the user* for delivering information about these services, and as a set of *context sensors* capable of detecting task relevant information necessary to support these services. The device will collaborate with the physical location by doing things as simple as providing a room name or phone number, to functioning as a remote control for a net appliance.

The Prototype

We are currently developing a prototype for a mobile device intended to demonstrate a common core interface for two different task scenarios: a visitor attending a meeting at an Andersen Consulting office, and a customer visiting a store intending to select and purchase a camera. Briefly, in the meeting scenario we are focusing on an out of town Andersen Consulting employee arriving in one of our offices and dealing with issues including registering, coordinating with colleagues, arranging and finding a temporary office, dealing with meeting services requests, finding and using office equipment (e.g. projectors, VCRs, printers), and arranging transportation to leave. The store scenario involves a customer interested in a digital camera arriving at the store, finding the appropriate part of the store, finding the right salesman, getting directions on how to try and use particular models on display, getting third party customer service advice on camera selections, being presented with alternative vendors, third party financing, and third party insurance options, determining compatibility with other equipment. While the two scenarios are clearly different, they share the need to find locations, use local equipment, invoke supporting services, find people, and get help with local resources.

Our prototypes are being designed for both Palm VII devices as well as for Wireless Application Protocol (WAP) enabled phones. In both of these scenarios the device will interact with a server at the location running a "task host" application that is aware of the task to be performed, the resources of the location, and the user. Our lab is equipped with active badge sensors and various tagging and tracking devices that allow us to detect the presence of people in any office or conference room and monitor the whereabouts of people and objects throughout our workplace. The task host application is being designed to exploit the information from these tools and, in part, will make use of earlier awareness systems built on this infrastructure [McCarthy & Meidel, 1999].

The Interface

The interface for our mobile application is being designed around the stages of generic location-based tasks. Within each stage we are identify and enable the core *objects, locations,* and *people,* we work with, and the *actions* we need to perform. At each task stage there is typically a *focus object* (e.g., the presentation we intend to give at a meeting or the person we are going to see / the camera under consideration), and a small number of *supporting objects*, (e.g. overhead projectors, printers, VCRs, executive assistants, AV specialist / other cameras under consideration, accessories, salesmen, customer service reps). These supporting objects play predictable supporting roles (i.e. they provide an alternative or an enabling service).

The actions we want to perform tend to be a function of the focus object and its type. One of our goals in this research is to identify a core set of actions we will want to include as a function of the focus object type. For example, if the focus object is a person we will frequently want to *contact them, invoke their role, send them* *something, pay them, or query them.* If the focus object is a location we will want to *go there, see the current status and contents, reserve it, prepare it, release it.* For physical objects we may want to *buy, configure, send, examine, return, invoke, or release* them. For a service or event we may want to *register, reschedule, cancel, upgrade/downgrade, repeat, undo, start, stop,* and *continue.* More generally, we wish to enable easy access to context sensitive *help*, the *most recent task* pursued by the user at this location (so that, for example, a user can request the same arrangements as last time), and any *changes and exceptions* to the routine since the user's last visit to this location. For example, is someone I normally deal with no longer working here? Are there any new facilities or offerings? And what exceptions to the norm are in effect. Is the store or office closing early? Is someone absent? Is a certain dish not on the menu?

Naturally not all "menu options" are shown at once. The selection is made on the basis of the particular task stage and the services available to support a given focus object. Moreover, not all information is intended to be received by the user through the device. Instead, we are making heavy use of displays available in the environment, but controlled through the mobile device. Once again, the intent is not for the mobile device to become the predominant tool to perform all functions everywhere but rather to see how such devices, with the appropriate interfaces, can complement the facilities and services at a location and ease the way we pursue tasks.

Our work in this area follows a series of projects in which we developed prototype applications for mobile devices to support specific remote tasks. The Avalanche project explores new ways in which technologies such as PDAs, wireless communications, GPS, and onboard computers can be used to provide new services centered around the automobile in support of the tasks that drivers typically engage in. Given how central the automobile is to many of our routine activities – commuting, ferrying children about, shopping and vacationing, among others, the automobile is an ideal theme around which to anchor the delivery of a variety of services in support of these tasks. The Shopper's Eye project focused on using mobile technology in support of *physical* shopping [Fano, 1998]. The intent was not to provide a self contained shopping application, but rather to augment the physical shopping experience of a shopper in a mall. Using a PDA equipped with a global positioning system (GPS) receiver we built an application that maintains a profile of the shopper along with a current shopping list, and, based on the current location of the shopper, presents relevant offers from retailers in close physical proximity to the shopper. More recently, another CSTaR project explored how barcode scanner equipped PDAs could be used to perform live price comparisons within a bookstore [Brody & Gottsman, 1999].

Andrew Fano has been a researcher at Andersen Consulting's Center for Strategic Technology Research (CSTaR) for almost four years. Prior to joining CSTaR, Andrew earned his Ph.D. in Computer Science at the Institute for the Learning Sciences at Northwestern University. His interests include task-based interfaces, situated computing, and intelligent interfaces.

Beigl, M. (1999). Point & Click - Interaction in Smart Environments, International Symposium on Handheld and Ubiquitous Computing (HUC99), Karlsruhe, Germany.

Brody, A. B., and Gottsman, E. J., (1999). Pocket BargainFinder: A Handheld Device for Augmented Commerce, First International Symposium on Handheld and Ubiquitous Computing (HUC '99), Karlsruhe, Germany.

Fano, A. (1998). Shopper's Eye: Using Location-based Filtering for a Shopping Agent in the Physical World. Proceedings of the Second International Conference on Autonomous Agents, 416-421, Minneapolis, MN.

McCarthy, J. F., and Meidel, E. S. (1999). ActiveMap: A Visualization Tool for Location Awareness to Support Informal Interactions, First International Symposium on Handheld and Ubiquitous Computing (HUC '99), Karlsruhe, Germany.

Schmidt, A., Aidoo, K. A., Takaluoma, A., Tuomela, U., Van Laerhoven, K., and Van de Velde, W. (1999). Advanced Interaction in Context. International Symposium on Handheld and Ubiquitous Computing (HUC99), Karlsruhe, Germany.

Composite Device Computing Environment: A Framework for Augmenting the PDA Using Surrounding Resources

Thai-Lai Pham, Georg Schneider, Stuart Goose, Arturo Pizano
Multimedia/Video Technology Department
Siemens Corporate Research, Inc.
755 College Road East
Princeton, NJ 08540 USA
+1 609 734 3344
{pham, gschneider, sgoose, apizano}@scr.siemens.com

Keywords
Personal Digital Assistant (PDA), Handheld Computers, Mobile Computing, Situated Computing, Ubiquitous Computing, Composite Device

INTRODUCTION

The World Wide Web (WWW) continues to enjoy phenomenal growth with the promise of facilitating a digital society. Technology continues to evolve, allowing an increasingly peripatetic society to remain “connected” without any reliance upon wires. As a consequence, mobile computing is a growth area and the focus of much energy. Mobile computing heralds exciting new applications and services for information access, communication and collaboration across a diverse range of environments.

By considering factors such as the user’s identity, profile, location, etc., s*ituated computing* is a methodology for imbuing applications and services with more personal and appropriate behavior. However, much current research is concentrated upon the delivery and presentation of context-aware information. Although using different sensing technologies, the Metronaut system [5] and the Cyberguide project [1] use this approach in guiding visitors to locations on a university campus. Location context is also used by the Active Badge System [7] to determine and track people moving around a building.

MOTIVATION

Contemporary wireless solutions typically include PDAs, or notebooks using cellular modems connecting to wireless networks to access a broad array of IP-based services. In the future it is probable that the devices will change; the networks will change; the protocols will change; and content will change.

Typically, popular mobile devices are sized appropriately to fit conveniently into a pocket. Although it is anticipated that the PDA screen resolution and quality will improve, this key social requirement for the device to be pocket-sized therein imposes a constraint upon the maximum physical size of the small screen display. While the other factors listed are likely to evolve, the physical limitation of the display size is likely to remain constant for a longer period.

At Siemens Corporate Research (SCR), our research focus is to provide mobile users with access to rich multimedia information and services. These observations have led us to investigate approaches for overcoming the inherent display limitations of a PDA. In the Composite Device Computing Environment (CDCE) project the surrounding available computing resources are considered as another facet of situated computing. As such a PDA-centric framework has been built to provide a situation-aware mobile information system. Having acknowledged the limitations of the PDA, the CDCE framework provides mechanisms for seamlessly exploiting and interacting with the available surrounding computing resources (e.g., PCs, workstations, TVs, telephones) to augment the PDA. Based upon the user’s request, the CDCE framework dynamically creates a unified *composite*, or virtual, device composed of an appropriate mix of the surrounding resources. CDCE flexibly combines the positive aspects of mobility with static computing resources in the vicinity. Hence, the CDCE provides at the user’s current location an infrastructure to support a mobile collaborative working environment.

A PDA is used as the primary device through which to requisition information, applications and services. The CDCE framework can offer access to broad range of multimedia services across a multitude of potential output devices. The CDCE framework offers an alternative paradigm for ubiquitous situated computing. The remainder of this paper describes the CDCE framework.

THE COMPOSITE DEVICE COMPUTING ENVIRONMENT

The following medical scenario serves as an illustration of the value of the CDCE framework.

Mobile Healthcare Scenario

In this scenario, the CDCE framework has been deployed within a hospital. Each doctor is equipped with a PDA affording wireless access to the hospital patient information system. When conducting her rounds, the doctor enters the room of the first patient. The doctor wishes to query the patient's medical history, including symptoms, diagnoses, prescriptions and x-rays. The PDA first detects the presence of a TV and a telephone in the room using the infrared interface. The PDA then communicates the doctor's request together with details about the detected devices to the CDCE gateway server. After authorizing the doctor access and verifying a secure connection, the CDCE gateway routes symptoms, diagnoses, and prescriptions directly to the doctor's PDA. As CDCE server is aware of the PDA physical limitations, the x-ray image is transmitted via RF to the TV for viewing. The doctor then uses the infrared capability of his PDA to annotate a region of the x-ray. CDCE then establishes a telephone call to the patient's original doctor for consultation. This arrangement provides a convenient infrastructure for the doctor to access, view, interact and collaborate upon the multimedia information.

This scenario demonstrates the way in which location information is utilized. It demonstrates the use of the PDA as a unique communication and access device. Also, tasks that are not suitable for the PDA to perform are outsourced to more appropriate devices. It illustrates the need for the convergence of wireline and wireless networks to transmit the data as well as to establish a short-range ad-hoc network for device detection. Finally, it stresses the importance of the CDCE gateway to format the information in different ways for the different devices.

The CDCE Framework

The above scenario briefly describes the main components of the CDCE framework. The fundamental idea is to avoid having to use a single PDA to perform all tasks. Instead of tying users to the traditional computing environment, we actively seek to exploit it. This idea is based on the observation that our daily life environment is becoming evermore equipped with electronic and computing hardware. For example, a household with a TV, telephone and PC with Internet connection is becoming commonplace. At the office there is a preponderance of powerful hardware such as workstations, PCs, beamers, printers, high-resolution monitors, etc. All of these devices are potentially available to users of mobile applications and services via CDCE.

The research focus of the CDCE project is to:

- develop scenario specific PDA user interfaces; elicit user requests for mobile applications and services
- dynamically ascertain and reserve the appropriate surrounding computing resources
- collate, process, deliver and display on the appropriate devices(s), either in parallel or in sequence, the requested information
- support multi-modal interaction across the range of available CDCE supported computing resources

A number of desirable design goals for CDCE have been identified:

- Adhere to standard protocols and services where appropriate. The PDA interface is a WWW browser and the user requests are transmitted via HTTP to the CDCE gateway.
- CDCE should not require any proprietary software to be installed on the available computing resources in the surrounding environment. A WWW browser is the only pre-requisite, with additional functionality supported through the browser extension mechanisms.
- CDCE needs to be sufficiently adaptive in order to exploit an ever-changing number and diverse range of available computing resources in the surrounding environment. The CDCE gateway must optimize the presentation of the information and the selection of output devices based on whatever resources are currently available.

From Figure 1 it can be appreciated that the CDCE framework consists of four main elements: the PDA, the CDCE Smart Gateway, the computing resources in the "environment" and the Network Communication Model.

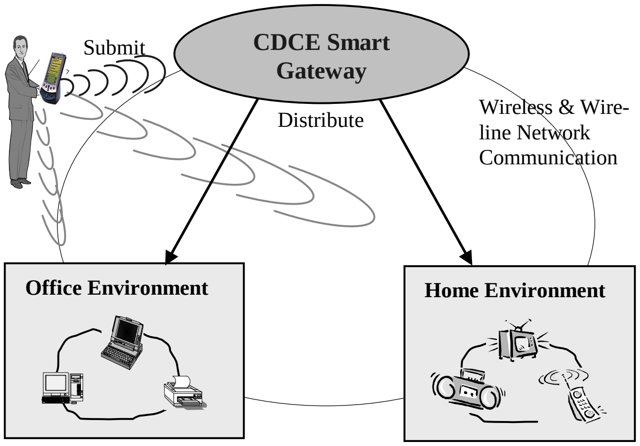


Figure 1: Elements of the CDCE framework.

The PDA detects and reserves available devices in the close vicinity and informs the Smart Gateway. The relatively short existence of the PDA and the continuous introduction of new devices and functionality prevent the definition of standards upon which to rely. Instead, we recognize the initial realization of the CDCE concept requires making decisions about specific computing platform, such as a wireless IR or RF interface and HTTP/HTML or WAP/WML.

The Gateway intelligently organizes, synchronizes and distributes requested information and services for interactive media access. In detail, it fulfills four major tasks:

- it manages the pool of services available to the users. For example, in the medical scenario it enables specific applications based on the doctor's location, identity and privileges.
- it establishes the composite devices based on location-dependent information received from the PDA, predefined knowledge about the Environment, and dynamic information on the current status of the various nodes comprising the composite device.
- it maps the requests issued by the PDA to the applications, and the corresponding output to the appropriate nodes in the virtual device. In the medical scenario this is manifested by the fact that the Gateway will redirect the x-ray image transmission to the TV via RF.
- it performs the dynamic conversions needed to present the information on the selected output node.

The role of the Network Communication Model is to manage the convergence of wireless and wire line networks, as well as the corresponding communication protocols. This is necessary to ensure a seamless device communication and data transmission. In the medical scenario, the network communications model includes the short-range wireless communication used by the PDA to detect surrounding devices; a cellular or wireless network to support the interaction between the PDA and the Gateway.

Finally, the Environment represents the pool of resources available at the user's current location. These resources can vary considerably from location to location. Figure 1 illustrates two possible environments and the typical range of available resources.

CHALLENGES AND OUTLOOK

Two primary research challenges addressed relate to multimedia information management and the network technologies. These topics are elucidated upon in the following sections.

Information Management and Distribution

A key focus of the research for the CDCE concept is the intelligent information distribution. Hence, the following methods are required to adapt the content to multiple output devices with varying capabilities as well as to the changing number of devices:

- *Splitting*: Intelligent content separation. E.g. a user wants to view a video message in an environment where only a PC without sound card and a telephone exist. In this case the CDCE system would split the audio part and redirect it to the available telephone and the video part to the PC.
- *Conversion*: Media conversion techniques, such as text to speech [2], can be offered when no appropriate devices are available
- *Filtering*: Content extraction and delivery of the sub-content, which can be rendered by the output device. E.g. delivery of only the audio part of the video message to a telephone.

Also important is the smart delivery sequence, or in parallel, of information based on both the number and capabilities of the nodes comprising the composite device.

Mobile User Interface

It is clear that the physical constraints of the screen mean that the information to be rendered on the PDA imposes limitations not experienced on a desktop machine.

Interaction with the environment is crucial and, in view of this requirement, we anticipate three alternative modes:

- *Abdication*: In this case the PDA hands over the control to the device. E.g. once an application is started on PC the mouse and keyboard of this PC will act as input devices.
- *Cooperative*: PDA and input devices of the output device can be used to control the application [4]. E.g. a slideshow can be annotated either using the mouse and keyboard of the output device or through a specialized and simplified user interface for the PDA
- *Exclusive*: The only input device is the PDA. This is especially important for output devices where no input facilities are connected (e.g. a TV)

The cooperative and the exclusive modes require specialized user interfaces on the PDA. The functionality that these interfaces must provide depends strongly on the scenario. For example, in one scenario annotation functionality is needed, whereas in another it may not.

Network Architecture

To facilitate a bi-directional flow between *PDA-Environment-Gateway-Environment* elements a suitable abstraction for the communication was sought. The use of the Distributed Component Object Model (DCOM) [3] helped to fulfill these requirements. Security cannot be ignored in the design of distributed information frameworks. It is clearly not tolerable for one CDCE user to monopolize another user's machine when in use. The CDCE gateway is the *only* entity with privileges for remote process invocation. To schedule in advance the resources required some scheme for reservation is also required.

Sensing/Detection Techniques

For the CDCE Gateway to construct a composite device, the PDA must detect and communicate information regarding the available resources. As described in the scenario, IR is a technology suitable for the purpose of device detection. Bluetooth [6] is also another such technology under evaluation for optimizing the process of sensing and detection. The CDCE framework makes the detection and utilization of available devices transparent to the user, thus allowing the user to focus on the actual tasks.

CONLUSION

We have presented a new concept of utilizing location information in order to develop a new class of mobile and ubiquitous services and applications.

Currently, we have successfully developed the first CDCE demonstrator for the “Office Environment”. Our system consists of a PDA running Windows CE and multiple Windows NT workstations equipped with infrared serial interface adapters. IR is used for short-range detection. The PDA and Smart Gateway communication is realized using HTTP over a Cellular Digital Packet Data (CDCP) network. The Smart Gateway and output client communication is current achieved by using a LAN. In the demo system a user can detect the NT workstations as potential output devices with her PDA and request the Smart Gateway to stream multimedia files to one device and to enable Microsoft Exchange Webmail service on another. Distributed Component Object Model (DCOM) [3] enables the CDCE Gateway to remotely invoke processes, without any requirement for proprietary client code. A WWW browser is the only client pre-requisite, with additional functionality supported through the browser extension mechanisms.

REFERENCES

1. Abowd, G. D., et al. Cyberguide: A mobile context-aware tour guide. *ACM Wireless Networks 3,* 1997.
2. Goose, S., Wynblatt, M. and Mollenhauer, H., 1-800-Hypertext: Browsing Hypertext With A Telephone *Proceedings of the ACM International Conference on Hypertext*, June, 1998, pp. 287-288.
3. Grimes, R. Professional DCOM Programming. *Wrox Press,* 1997.
4. Myers, B. A., Stiel, H. and Garggiulo, R., Collaboration Using Multiple PDAs Connected to a PC, *Proceedings CSCW 98: ACM Conference on Computer-Supported Cooperative Work,* November, 1998, pp. 285-294.
5. Smailagic, A., Martin, R., Metronaut: A Wearable Computer with Sensing and Global Communication Capabilities. *Proceeding of IEEE First ISWC*, October 1997, pp. 116-122
6. The Bluetooth Forum. *http://www.bluetooth.com*
7. Want, R., et al. The Active Badge Location System. *ACM Transaction on Information Systems, Vol. 10,* 1992.

"Magic of Today: Tomorrow’ s Technology”

Wearables for Kids

Marilyn Panayi
University of Southern Denmark
Faculty of Science and Engineering
Natural Interactive Systems Laboratory
Odense, Denmark
+4563157304
panayi@nis.sdu.dk

David Roy
University of Southern Denmark
Faculty of Science and Engineering
Natural Interactive Systems Laboratory
Odense, Denmark
+4563157303
roy@nis.sdu.dk

ABSTRACT

Research is presented that illustrates frameworks being developed that involve young children in the process of development of future wearable technologies – A hypercamera - the KidsCam. It is envisioned that such digital technology will become embedded in educational culture forming part of the ‘educational information ecology’ [EIE] and create opportunities for shared reflection on early life experiences. A model of ‘Situated Interaction Design’ [SID] is introduced that is inspired from the domains of educational theory and practice, human computer interaction, cognitive science and interactive performance art.

Keywords

Situated interaction design, children, wearables, educational theory, educational information ecology, interactive performance art, hypercamera.

INTRODUCTION

Children’s willingness to dialogue and play provides a rich opportunity to explore their visions of the future. Research frameworks are being developed that involve young children,
(4-10 years old) in the process of developing future technologies[1].

A model of situated interaction design [SID] is introduced in this paper and has been inspired by and draws upon theories from the domains of educational theory and practice, human computer interaction, cognitive science and interactive performance [1, 2,6, 8,11,13,15,16,17]. The methodologies illustrated also draw upon the authors’ previous work [13].

A Situated Interaction Design model is created in response to the technological brief and the ‘information ecology’. Aspects of the ecology that the model could encompass include needs, desires and expectations within which the technology and potential users are to be located. A brief description of a case study for SID is present for the Today’s Stories project[2]. The case study describes the social-cultural context for technology development, theoretical paradigms and frameworks and introduces the interaction methodologies with reference to both ‘design and educational ecology’ and ‘design and technology integration’.

SITUATED INTERACTION DESIGN

Local habitations, Interactants and New Technology

The local habitations[i] and interactants[ii] with which we are working are children, their teachers and parents from schools in Israel and Denmark. These ‘Open School Communities’ are reference groups that represent cultural ‘educational ecologies’. They

---

[1] European funded Future and Emerging Technologies, Experimental School Environments www.i3net.org/ese

[2] ESE Today’s Stories Project www.stories.starlabs.org

joined the research effort having a pre-commitment to the development of technologies that could support children's reflective thinking and action in future learning environments. They are coming together with researchers, educationalists, psychologists, designers and technologists to develop a wearable technology – A hypercamera - the KidsCam. The interactants, the local habitation, the social dynamics and the technology constitute part of what is termed an 'educational information ecology' [EIE].

Paradigms

A key paradigm that underlies the educational and technology development of this project is that of *reflection* both in terms of the 'reflective practitioner' [16] and Autonomy Oriented Education [AOE]. In Israel [AOE], is the educational philosophy supporting the researcher and school community interaction. It focuses on the development of autonomy, morality and belonging in children [1].

Technology brief and vision

The technological brief of the KidsCam was to create a wearable technology that would allow children to learn from reflecting on their actions and learn from other children's perspectives on their own actions. In time, similar wearable technology could become embedded within these 'educational information ecologies' [EIE] with the aim to support childrens' reflection in their early-life 'experiments in living'.

The KidCam will facilitate capture and document such "reflective experiments in living". Children will build digital portfolios of their day's interesting events. A key pedagogic aim being to support the development of social, communicative and emotional skills of children in the context of their everyday activities.

It is envisioned such technology will have a facilitating role, in that it will complement the discovery of alternative forms of educational interaction and the development of new media. Such phenomena often follow in the wake of the introduction of new technologies and facilitate systems change.

A particularly novel aspect of this technology is that it will support multi-user, multi-perspective [MUMP] interaction via digital artefacts captured by the hyper-camera. A community memory of a group of children will be co-created and evolve through a didactic process of dialogue and reflection. [14]. Such interactive digital artefacts could enhance and also contribute to cross-cultural understanding and critical technology awareness.

The existing local practice, infrastructures and technology use are being documented and conditions for acceptance and success of deploying this new technology in a social, cultural and ethical context are being investigated against the backdrop of the current socio-technic debates [4,5,7].

Cross- Cultural Reference Group Profiles

Staff and children in two countries Denmark and Israel have been contacted and sensitised to the project. There is continuous and ongoing activity to create technology awareness, skill development and pedagogic framing for both staff and children in relation to the Today's Stories project aims.

Four reference groups have been established. Two reference groups are located in the island of Funen, Denmark. The first is an integrated pre-school through eighth grade school with 140 pupils. The school is located about half an hour from the major city of Odense. The second is a 'KidSearcher ™ group established and located at the Natural Interactive Systems Laboratory, on the University of Southern Denmark campus at the Funen International Science Park. In Israel, two reference groups have been established in the suburbs of Tel Aviv. The first is a community school where two integrated pre-school/first grade groups of about thirty children each participate. The second is an elementary school, three first grade classes are involved, each with about twenty-five pupils.

In the case of Israel, the first phase of the Today's Stories' project focus and the majority of contact time has involved working with teachers. By comparison at the Danish site, the majority of contact time has been devoted to working in the classroom with children and staff.

INTERACTION METHODOLOGIES

Interaction methodologies have been designed within this theoretical framework of an 'educational information ecology'. A model of situated interactive design [SID] has been developed and applied to the concept of 'Open school Communities'.

What follows is a brief outline of activities that have taken place at both sites. In Israel the program is currently being applied in an experimental framework in two schools in the Tel Aviv district of Israel, and in four kindergartens. The research team has worked to put in place an "experimental pluralistic" framework [1] for the development of the Today's Stories project. The team works in various settings. Work has focused on the preparation of staff to support the in-class reflective process of children who will be using the proposed technology. This has involved seven teachers, four nursery teachers and one staff member for thirty children [18].

In Denmark the project has been introduced to the whole school staff team. Contact was made with the first four early-years classes of the school 5-10 year-olds (60 children taught by 9 staff). The curriculum is designed so that the children have the opportunity to work in cross year groups approximately 25% of the time [18]. The initial phase of the project has focus on working with children aged 7-10 years old.

INTEGRATING INTERACTION DESIGN AND THE EDUCATIONAL ECOLOGY

Interaction sessions are supported by classroom staff and have focused on developing children's technology awareness. Four key methodologies have been designed based on the authors previous interaction design research and practice. These methodologies have taken the following embodiment and are described as:

1) ***Experimental probe (EP)***
2) ***Community of Enquiry (CoE)***
3) ***Studio Theatre (VT)***
4) ***Smart Things***

Through the concept of the *'Experimental probe'*, we have introduced access to existing baseline technology for capturing digital and analogue material in terms of analogue and digital cameras. During the first phase of the project at the Danish school site, a suite of 'experimental probes' has been launched. This suite comprised of: analogue stills camera, disposable cameras, two digital stills cameras and a digital video camera with associated software including, iMac video editing software.

'Community of Enquiry' is an educational concept and practice developed from the work of Mathew Lipman and his work in teaching philosophy to children [8]. This work has been adapted to create and support the evolution of children's 'critical thinking' skills and dialogue in the context of existing and future technology use from their personal and group perspectives.

'Studio Theatre' has been developed using principles from drama, black theatre and filmmaking. The techniques have been adapted to create on site 'studio' sessions. These have been designed to facilitate children's skills in story-telling, storyboarding, transformation of narrative, prop making, role-play and independent use of video and photographic technology.

'Smart Things' are prototyping and scenario building interaction sessions. Design tasks are set that encourage children to:

a) Imagine the present with the 'new'
b) Imagine the future.

Children make situated imaginary and realistic drawings. They build mock-ups and prototypes of smart things and clothes. They work both in-groups and individually to embody design ideas and are given contexts to 'role play', discuss uses and issues for future 'smart artifacts'. A short video film[iii] has been produced to illustrate the process of SID activity with the two Danish reference groups. The children working in their local environments filmed the majority of video material. Short video clips show

how situated interactive sessions that draw upon pedagogic and user–centered design methodologies were used to:

a) Sensitise the groups
b) Support skills development in base-line technologies
c) Encourage design activity related to the technology development
d) Support technology awareness.

The video is being used to provide a reference and reflection point for staff, children, parents and researchers. In addition it is used to disseminate the interaction methodologies to the research community.

INTEGRATING INTERACTION DESIGN WITH TECHNOLOGY DEVELOPMENT

Outcomes from the interactive classroom sessions and discussions with staff and children have been incorporated into the initial phase of technology development e.g. concept prototypes, hardware specification, software requirements, deployment scenarios. Children have been responsive to requests from the interface design team of the consortium to contribute ideas for the interface 'look and feel'. These ideas have contributed to the concept and prototype development of both the wearable KidsCam and 'Composer Environment [18]. Within the discussions of social responsibility and integration of technology cross-culturally there are ongoing dialogue and incorporation of ethical consideration of technological development, privacy and community [4,5,9,12].

EVALUATION AND TRIALS

Members of staff and students contribute to the ongoing evaluation of existing educational software for story making and editing. Members of the research team and teaching staff are currently exploring how this proposed new technology could support the curriculum.

'Reflection sessions' are planned where children will feedback their experience from using the 'experimental probes' and the KidsCam prototype system. A more in depth interaction analysis [6] is currently underway on samples of video data from the Today's Stories situated interaction design sessions. Scenario based interaction methodologies drawing upon practise from the human computer interaction [3,8] and performance arts arenas [2] are also being developed to support the design of the technology trials of the KidsCam.

ACKNOWLEDGMENTS

Children, staff and parents participating in the Today's Stories project.

REFERENCES

1. Aviram A. *Personal Autonomy and The Flexible School*. International Review of Education 39(5): 1993, 419-433.
2. Boal. A. Rainbow of Desire, Routledge, 1995.
3. Burns. C., Dishman.E., Johnson.B., Verplank.B. "Informance": Min(d)ing Future Contexts for Scenario-Based Interaction Design, BayCHI August Meeting, Tuesday, August 8, 1995. http://www.baychi.org/meetings/archive/0895.html
4. Friedman, B. Human Values and Design of Computer Technology, CSLI Publications and CUP, 1998.
5. Gill, T, ed. Electronic Children: How Children are Responding to the Information Revolution. Concord, MA: Paul and Company, 1996.
6. Hutchins, E. "Cognition in the Wild", The MIT Press, London 1995.
7. Jordan, B. and Henderson, A. Interaction Analysis: Foundations and Practice The Journal of the Learning Sciences (4,1): 39–103, 1995.
8. Laurel. B. Computers As Theatre, Addison-Wesley, 1993.
9. Laurel. B. Technological Humanism and Values-Driven Design, Keynote Address, CHI-98, Los Angeles, California ACM, April, 1998.
10. Lipman, M. Thinking in Education. Cambridge, Cambridge University Press, 1991.
11. Manen, Van, M. Researching Lived Experience: Human Science for an Action Sensitive Pedagogy, 1990.
12. Manen, Van, M., Levering, B. Childhood's Secrets: Intimacy, Privacy, and the Self Reconsidered. New York, NY: Teachers College Press, 1996.
13. Panayi, M., and Roy, D.: "BodyTek: Technology Enhanced Interactive Physical Theatre for People with Cognitive Impairment" in Ryohei Nakatsu, Edward J. Altman, and Claudio Pinhanez (Eds.) Proceedings of ACM 6th International Multimedia Conference,

Workshop on Technologies for Interactive Movies, pp.35-39, 1998.

14. Panayi, M., Van de Velde, W., Roy, D. M., Cakmakci. O., De Paepe, K., Bernsen, N.O. Today's Stories, Hand Held and Ubiquitous Computing, Karsrhule, Germany, 1999
15. Polanyi, M The tacit dimension, Routledge, London, 1967.
16. Schön, D. A. The reflective practitioner. New York: Basic Books, 1983.
17. Schön, D.A. (Ed) The reflective turn: case studies in and on educational practice. New York: Teachers College Press, 1991.
18. Today's Stories Web site. Public documents, reports and publications. Pedagogic and Technology Roadmaps. http://stories.starlab.org/index.html

---

[i] Local Habitation –term used in biological science and ecological studies to refer to the immediate environment. More recently adopted in computing terminology, {see Nardi. B., O'Day V.L. Information Ecologies - Using Technology with Heart, 1999. MIT Press}.

[ii] Interactants – term typically used in theatre/drama context to describe 'action' participants. Extend beyond the traditional term 'actor' to encompass a variety of interactive performances. {See also term 'interactors' Boal. A. Theatre of the Oppressed, Pluto Press, 1979}.

[iii] How Did They Do That? – Wearable Technology, Producers: Panayi & Roy, Media Studio Workshop, Southern Danish University, Denmark, 1999.

A Context / Communication Information Agent

Jason Hong, James Landay
Group for User Interface Research
University of California at Berkeley
{jasonh, landay}@cs.berkeley.edu

The system we envision is a proactive software agent that uses context and human-to-human communication to help find and deliver the right information at the right time. The system constantly searches for information related to the current situation, in order to make it easier to find relevant related information. We call such a system a Context / Communication Information Agent (CIA).

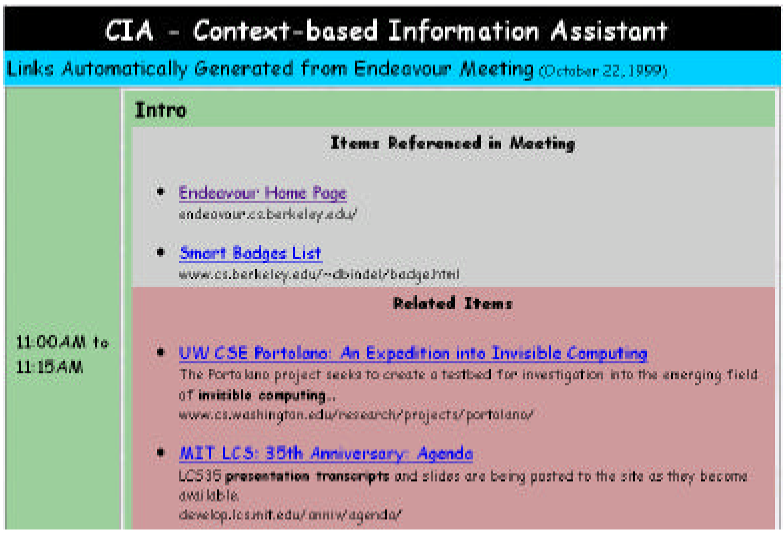
By context, we mean knowing the answers to the “W” questions, such as who is speaking, who else is here, where am I, what calendar event is current, and so on. As an example of how context could be used, suppose that earlier in the day, Francis scribbled down a grocery list. Later, when passing by the grocery store he usually goes to, his PDA beeps, reminding him to buy some food. As he enters, his PDA fetches his handwritten notes for him. As another example, suppose that a person has a weekly meeting to go to, stored as a recurring weekly event on his calendar. When the time for the next meeting takes place, the system could begin retrieving notes, minutes, and action items from last week's meeting, so that he doesn’t have to remember where he saved them.

By human-to-human communication, we mean using microphones, cameras, and other sensors to capture communication between people, such as text, ink, speech, and so on. As an example of how communication could be used to prefetch information, suppose that two people are talking to each other. One person says something along the lines, “There's this interesting paper I just read by some people at Berkeley about user interfaces”, and goes on to describe it more in detail. Using the information that was said, the system could begin searching for potential matches, so that the referenced paper, and possibly related papers, will be there if needed.

What we described above is a process-oriented view, that is it describes how the information is being retrieved. Another way of thinking about it is by the type of information being retrieved. The information being retrieved can be thought of as information a person would have searched for manually; related information the person already knows; serendipitous information the person didn't already know; or completely unrelated and useless information. Our goal is to maximize the first type, information that would have been searched for manually.

However, getting the information is only part of the problem. Just as important is how to present the information in such a manner to support the task, without overly distracting the users. For example, a display of constantly updating results would simply be too disruptive in a meeting.

Before implementing a system, we decided to run a low-fidelity prototype in a meeting situation to explore the domain and to test out some ideas. An audio recording was made of a weekly meeting. After the meeting, one of the authors did searches based on what was said. All of the results were assembled into a web page, organized chronologically and by general topic (see Figure 1). In each topic, the results were grouped by items explicitly referenced during the meeting, and items related to the discussion but never explicitly mentioned.



CIA - Context-based Information Assistant

Links Automatically Generated from Endeavour Meeting (October 22, 1999)

11:00AM to 11:15AM

Intro

Items Referenced in Meeting

- Endeavour Home Page
  endeavour.cs.berkeley.edu/
- Smart Badges List
  www.cs.berkeley.edu/~dbindel/badge.html

Related Items

- UW CSE Portolano: An Expedition into Invisible Computing
  The Portolano project seeks to create a testbed for investigation into the emerging field of **invisible computing**..
  www.cs.washington.edu/research/projects/portolano/
- MIT LCS: 35th Anniversary: Agenda
  LCS35 **presentation transcripts** and slides are being posted to the site as they become available.
  develop.lcs.mit.edu/anniv/agenda/

Figure 1 – Low-fidelity prototype of search results from meeting

Once the results were organized, the meeting participants were asked to look over the results and to fill out a short survey, judging the usefulness of the results as well as the organization scheme. The general results were that people liked the concept a lot, but wanted more useful results, as well as more sophisticated ways of organizing and filtering the results. Furthermore, people were interested in seeing if the system would be useful in a meeting real-time. One serious concern was control of the system: people should be able to turn it on and off when desired.

Next, we built a prototype that takes speech input, processes it through a speech recognizer, and then does web searches based on keywords spotted in the recognized speech. It can be currently thought of as a speech-based interface for web search engines. We are presently in the process of improving the recognized speech, as well as expanding the search to other kinds of information, such as digital libraries.

We are also in the process of investigating several strategies to minimize attention to the agent in a real-time meeting situation. First, we believe that peripheral displays will be useful, that is using secondary monitors and projectors off to the side to display the results. Second, we believe that periodic updates will be more useful than continuous updates, so that people will not have to read constantly changing information. Third, we believe that pre-processing the results to extract the most important headers and text can significantly reduce the amount of reading needed. In addition, there are intriguing directions to explore for asynchronous interaction, such as receiving an email from the agent after a meeting.

In several respects, the CIA as envisioned is similar to Remembrance Agents [1], but moves the focus away from keyboard input and from wearable computers. The CIA is also related to the XLibris system [2], a pen-based portable document reader specifically designed for reading electronic documents. One notable feature in XLibris is implicit linking: highlighting phrases in one document would cause the system to search locally for related documents. Any links found would be presented as a small document icon in the margin next to the highlighted text. Thus, the user never explicitly searches: documents are instead found opportunistically. The key observation is that useful information can be found based on activities one is already doing. The CIA also has a strong relationship with meeting capture systems, such as Classroom 2000 [3] and the data salvaging tools at PARC [4]. A CIA can be thought of as using the same infrastructure as these systems or built on top of these kinds of systems.

References

1. Rhodes, B., and Starner, T. The Remembrance Agent: A Continuously Running Automated Information Retrieval System. In the Proceedings of The First International Conference on The Practical Application of Intelligent Agents and Multi Agent Technology (PAAM '96), London, UK, April 1996. pp. 487-495. http://rhodes.www.media.mit.edu/people/rhodes/research/Papers/remembrance.html

2. Schilit, W.N., Golovchinsky, G., and Price, M. Beyond Paper: Supporting Active Reading with Free Form Digital Ink Annotations. In the Proceedings of CHI '98, 1998. ACM Press.

3. Gregory D. Abowd, Chris Atkeson, Ami Feinstein, Cindy Hmelo, Rob Kooper, Sue Long, Nitin “Nick” Sawhney and Mikiya Tani. Teaching and Learning as Multimedia Authoring: The Classroom 2000 Project. In the Proceedings of the ACM Multimedia'96 Conference, pp. 187-198, November 1996.

4. Moran, T., et al. “I’ll Get That Off the Audio”: A Case Study of Salvaging Multimedia Meeting Records. Proceedings of CHI’97, Atlanta, GA. ACM Press. 1997.

The Memory Glasses: Towards a Wearable, Context Aware, Situation-appropriate Reminder System.

Richard W. DeVaul, Brian Clarkson, Alex “Sandy” Pentland
{rich,clarkson,sandy}@media.mit.edu
MIT Media Laboratory

January 29, 2000

Contents

List of Figures

1 introduction

The Memory Glasses project is an attempt to build a wearable, proactive, context-aware memory aid based on wearable sensors. The primary goal of this project is to produce an effective short-term memory aid and reminder system that requires a minimum of user attention. Secondary goals include the development of more effective wearable computing interfaces and the development of a general framework for building context aware wearable and ubiquitous computing applications.

The function of our system is to deliver reminders to the user in a timely, situation-appropriate way, without requiring intervention on the part of the user beyond the initial request to be reminded. In other words, the system behaves like a reliable human secretary that remembers reminder requests and delivers them under appropriate circumstances. Such a system is qualitatively different from a passive reminder system (such as a paper organizer) which records and structures reminder requests but which can not act to acquire the user's attention.

1.1 categorizing proactive reminder systems

Proactive situation-appropriate reminder systems can be categorized by the complexity of the factors (or context) taken into account in triggering reminder delivery. These factors include (in order of increasing complexity) time, location, and user-environment interaction (user actions not specifically directed towards interacting with the reminder system).

1.1.1 time based reminders

The simplest type of proactive reminder system is essentially an alarm clock, acting to acquire the user's attention and delivering reminders at previously specified times. Most electronic organizers have an alarm-clock feature, in which reminders may be scheduled for delivery at specified times. When the appointed time arrives, a general audible or vibration alert is produced to attract the user's attention; it is then up to the user to acknowledge the alert and receive the specifics of the reminder.

Time based proactive reminder systems are simple, reliable, and limited. Reminders cannot be conditioned on location or the actions of the user. It is only for tasks in which location and action are fixed or do not matter that such reminders work well; the classic example is the bed-side alarm clock which works reliably to alert a user of known location (the bed) and action (sleeping) that it is time to get up. (The user may choose to ignore this reminder, of course, or reschedule through interacting with the "snooze" button.)

1.1.2 location/time based reminders

Location/time based reminder systems are technically more complex, requiring an additional channel of information: the user's location. How this information is acquired and how reliable it is differs from implementation to implementation, with consequences for the reliability and scope of the reminders which can be delivered. Examples of the way location context may be acquired include GPS receivers (outdoors only)[6], tagged environments and portable sensors[9], wearable tags and "smart" environments[11, 5], computer vision[1, 10], and combinations of other types of sensors[4]. Each of these techniques have specific strengths and weaknesses, the most important being the degree of infrastructure dependence.

Being able to condition reminders on location is a powerful addition to a proactive reminder system, enabling location-only conditions (e.g. "next time I'm at the grocery store, remind me to buy milk") as well as sophisticated location-time conditions, such as varying the lead-time of a reminder to go to a meeting depending on current location, time, and probable time of transit.

Location context is broadly useful because it is an important clue to user activity[8]. For example, very different activities take place at home, at the supermarket, and in the office; there is little point in reminding the user to go to a meeting or to go grocery shopping if the user is already engaged in these activities as indicated by the user's current location. However, location context by itself can not differentiate between multiple activities which take place in the same space.

For instance, the comMotion location-based proactive reminder and message delivery system developed by Natalia Marmasse and Chris Schmandt uses GPS to determine location [6]. Differential GPS location measurement has a high degree of accuracy outdoors but does not work indoors. Hence, the comMotion system can identify location at the building level but does not differentiate between being in one part of a building vs. another. This means reminders can be effectively targeted at activities which are segregated by location at the building level (grocery shopping vs. working out, for instance), but not at the room level (eating in the office cafeteria vs. meeting in the conference room).

Likewise, using a Locust-style active IR tag system[9] it is possible to do comparatively fine-grained location determination indoors, but no amount of location accuracy can differentiate between different events which occur in the same location. For instance, determining that one of the authors of this paper is sitting at his office workstation can not, by itself, differentiate between him writing his thesis or playing networked Quake II capture-the-flag. Knowing whether a user is working, resting, or in conversation with some other person has obvious implications for reminder delivery, yet all of these activities may take place at unpredictable times in the same cluttered office.

1.1.3 activity/location/time based reminders

An activity/localtion/time reminder system is able to take into account the user's activity state independent of time and location. This not only allows reminders to be triggered by activity (e.g. "If I'm speaking too loudly, remind me to calm down") it also allows the system to balance overall demands on the user's attention. For instance, such a system might follow the rule that reminders are not to be delivered while the user is is engaged in a high-attention task like talking on the phone, or a high-risk task like crossing the street.

The difficulty is that recognizing user actions (or environmental conditions) generally requires high-bandwidth sensing and sophisticated analysis. Computer vision techniques have been shown to be effective in identifying user location and action [5, 1], but only recently has it become possible to do this type of sensing and analysis in an easily portable or wearable package, independent of outside infrastructure, with reasonable battery life and performance.

The goal of the Memory Glasses project is to produce an activity/location/time based proactive reminder system that is powerful enough to recognize a wide variety of user activities and environmental conditions, is flexible enough to be trained to recognize new user activities and conditions easily and with a minimum of user intervention, is ergonomic enough to be worn while engaged in a wide variety of daily activities, is reliable enough to be trusted, and makes the best possible use of the user's time and attention. In the next section of this document we outline the architecture for the Memory Glasses prototype system as a work in progress.

2 system architecture

The Memory Glasses prototype system is composed of several software and hardware components, some of which have not yet been fully integrated at the time of this writing. Conceptually, the system may be broken down as follows:

1. Hardware
   (a) Sensors (camera and microphone).
   (b) Wearable computing core (CPU/DSP, RAM, storage, power)
   (c) User interaction peripherals
2. Software
   (a) Operating system and device drivers
   (b) Classifier System
   (c) Agent-oriented programming layer
   (d) Memory Glasses application agents

Each of these will be discussed in turn.

2.1 hardware

The current Memory Glasses prototype runs on a repackaged notebook computer using standard sensors and peripherals. Although a repackaged notebook does not make an ideal wearable computer, we decided to employ conventional, easily available hardware rather than construct a special-purpose wearable in order to focus our attention on the Memory Glasses algorithms and interaction model.

2.1.1 sensors

The Memory Glasses context awareness system is based on high-bandwidth context sensing, namely computer vision and computer listening. The sensor data for this system is collected using a low-resolution USB camera and a conventional omnidirectional microphone. In the near future other context sensing hardware, such as a GPS receiver or active-tag reader, will be integrated to increase the number of features the wearable has available for context sensing.

2.1.2 wearable computing core

The current wearable computing core employed by the Memory Glasses prototype system is a Sony Vaio Picturebook PCG-CX1, a Pentium 266 MMX notebook computer. The Picturebook was chosen because of its compact form-factor, light weight, reasonable battery life, and acceptable performance.

2.1.3 user interface peripherals

The current Memory Glasses prototype employs a trackpad-style pointing device with two buttons for user input and headphones for audio output. When packaged, the notebook screen is closed and inaccessible, and is not used for interaction. A software layer allows the user to interact with the system using simple unistroke gestures, and audio clips are played to provide the user interaction feedback.

The MicroOptical clip-on display, a small, light-weight quarter-resolution full-color VGA head mounted display is also used in certain versions of the Memory Glasses system for training and reminder delivery, at the option of the user.

2.1.4 packaging

The wearable version of the Memory Glasses prototype consists of the Picturebook, camera, and microphone packaged into a low-profile, single-strap nylon satchel which is worn bandoleers style across the body. The Picturebook has also been packaged in vest form, with the notebook computer worn in the small of the back and peripherals distributed through the vest using Velcro. Both of these packages have been used for extended periods of time and found to provide reasonable ergonomics and mechanical stability.

2.2 software

The software is the primary focus of the Memory Glasses project. The Memory glasses system software can be divided into three application levels of abstraction plus the operating system.

2.2.1 classifier system

The lowest-level application layer is a signal processing system which converts the raw sensor data into a collection of probabilistic estimates. Conceptually, the classifier system allows the user to train event recognition

functions, or classifiers, to recognize patterns in the sensory data and tag them as specific events. The mechanism by which this time-series recognition occurs is a multi-level HMM grammar which is capable of recognizing patterns at a range of time-scales from seconds to days. A brief description of the classifier system is provided below; for more information see the Vismod technical report TR-519 available online at
`http://vismod.www.media.mit.edu/tech-reports/TR-519/`.

2.2.2 perceptual context classifier

The classifier takes two inputs, the sensor data from camera and microphone, and the label stream from the user or software agents. The goal of the classifier is to extract meaningful features from the sensor data and use these features to detect the events that the user has labeled. The classifier is based on work done by Clarkson [3, 2]. The system overview is as follows:

1. Extract basic features from the sensors at approximately 5Hz. We calculate all spatial moments up to order 2 from the images, 10 equally spaced frequency coefficients from 50Hz to 8000Hz from the audio, including measurements of auditory volume and the amount of speech detected in the environment.

2. These features are collected continually as the user goes through his/her day of activities. All of them together are used to build a World Model by training a Hidden Markov Model (HMM) with the above features. The resulting World Model is really a rough description of the user's surrounding sensory dynamics.

3. Next as the user labels various events and contexts around him/her with the equivalent of a clicker trainer (i.e. impulse labels that don't specify duration), Event Models are built by training more HMMs on the feature sequences surrounding each of the impulse labels.

4. The resulting Event Models are compared with the World Model to recognize these events after the training phase. $L(\mathit{Event\ Model}|\mathit{Observations\ at\ t}) > L(\mathit{World\ Model}|\mathit{Observations\ at\ t})$ indicates a triggering of the event detector (where $L()$ indicates the log likelihood function). Or, equivalently we can define an activation function for each classifier as $A(t) = L(\mathit{Event\ Model}|\mathit{Observations\ at\ t}) - L(\mathit{World}|\mathit{Observations\ at\ t})$.

Results were obtained for the events such as the following:

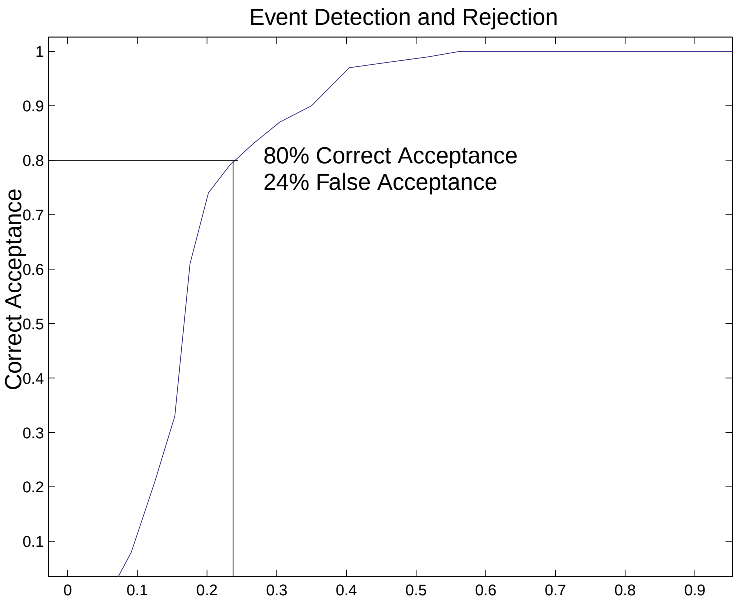


Figure 1: Classifier testing results.

- Entering/Leaving the office
- Entering/Leaving a large common area
- Entering/Leaving the kitchen
- walking down the stairs
- taking the elevator
- participating in a conversation

With the above types of events, after labeling for building a World Model for 2 hours, then labeling for 1 hour and testing for 2 hours, we were able to get the results for detecting and rejecting event occurrences shown in Figure 1. For more additional results on this classifier system please refer to `http://www.media.mit.edu/~clarkson/autodiary/index.html`.

2.2.3 the memory glasses agent layer

The Memory Glasses application is implemented using an agent-oriented programming layer that sits on top of the classifier system. This middle-wear layer is built on Hive[7], a distributed, agent-oriented programming system which is itself based on Java and RMI.

A hive application is a collection of agents work together. Each agent is an independent, modular entity with a separate thread of execution; agents are routinely created, reconfigured, and even killed while an application continues to run. Information is passed between agents through asynchronous events and remote method calls, and events produced by one agent may be received by any number of other agents, either locally or across a network. The environment in which an agent executes is a hive cell, which has a particular location and exposes a specific set of local resources to the agents executing locally on that cell. Agents access these resource through hive shadows, which are analogous to operating system device drivers. For more information, see [7] or `http://hive.media.mit.edu`.

There are a number of advantages to using hive-based agent-oriented programming framework for developing the memory glasses application, including robustness, flexibility, ease of extension, and ease of integration with other wearable and ubiquitous computing resources and applications. The memory glasses application is implemented as a collection of Hive agents, and a hive shadow which provides access to the underlying classifier system. Classifier agents periodically poll the state of the appropriate classifier through the classifier shadow, and produce classifier data events.

Hysteresis/comparator agents interpret classifier data events, and trigger reminder events. Reminder display agents interpret the reminder events and take appropriate action to communicate the content of the reminder to the user. User feedback is provided through classifier training agents, which allow the user to create, label, and train classifiers.

Figure 2 shows two iconic data-flow views of the Memory Glasses application within a standard Hive GUI. The view on the left is a single classifier application employing (in left to right order) a classifier training agent, a classifier agent, a hysteresis/comparator agent, and two reminder display agents (one graphical, the other auditory). The view on the right is a two classifier application in which both classifiers share the same event display agents. In the second view, the raw output one of the classifiers is being directed to graphing agents for visual analysis.

One of the most interesting features of the Hive programming environment is the interaction flexibility afforded by the agents-oriented, event-driven model. For instance, classifier agents and hysteresis/comparator agents work together to create reminder events. It is up to a reminder display agent to interpret how a particular reminder event should be displayed to the user; one type of reminder display agent might produce a

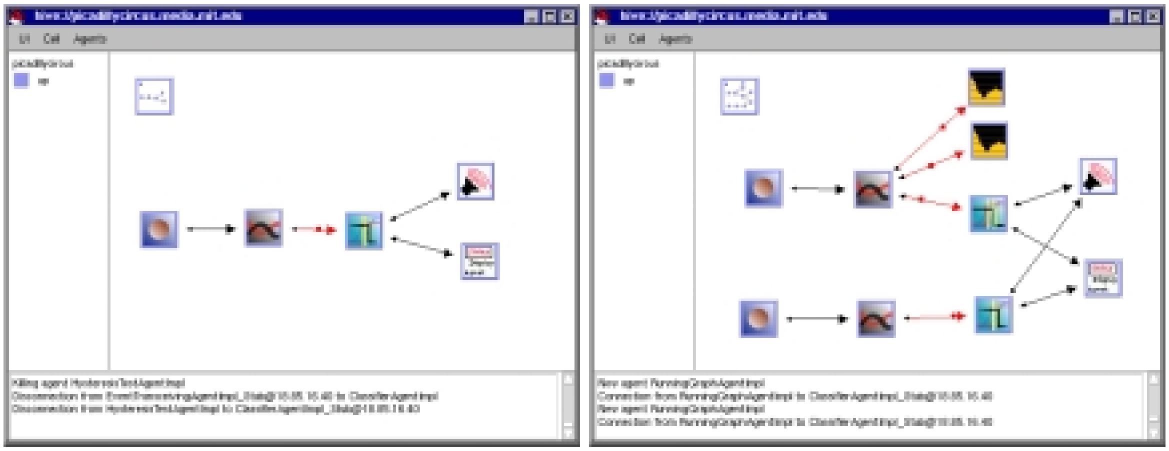


Figure 2: Iconic data-flow views of the Memory Glasses application.

visual interpretation, while another produces an audible one. Yet another might take some other action, such as sending email, displaying a URL in a browser running on a desktop workstation, or killing the Quake server on the office subnet.

2.2.4 software integration

In order to exploit the advantages of the respective operating systems, we developed and tested the classifier system under Windows and the agents-based framework and Memory Glasses application under Linux. We are now in the process of integrating the two parts of the system, and should have a complete system for integrated testing some time in early February. Once the system is integrated, we can work on streamlining the user interaction, building more complex and principled rule systems for reminder delivery, and work to improve the accuracy of the classifier system.

3 preliminary results and conclusions

3.1 results

At the time of this writing, we have a working hardware test platform based on the Sony Vaio Picturebook, as described above. We have a functioning classifier system, with preliminary training results of 80% correct acceptance to 24% false acceptance in the task of identifying six events after a relatively short training period.

We have a working agents-based middle-wear framework and preliminary Memory Glasses application, which is not yet integrated with the functioning classifier system. The primary task to accomplish before we can start full-scale testing of the system is integrating the functioning classifier system with the middle-wear and application layer.

3.2 conclusions

The goal of the Memory Glasses project is to produce a proactive context-aware reminder system that is sufficiently powerful, flexible, ergonomic, reliable, and useful to be a genuine tool for managing the complexity of daily life. Our work to date is only one step in this direction, but if we can accomplish our goals, the technology developed for the Memory Glasses project could be an invaluable aid to ordinary people as well as those with significant short-term memory problems, such as early-stage Alzheimer's patients and many graduate students.

4 Bibliography

References

[1] Hisashi Aoki, Bernt Schiele, and Alex Pentland. Realtime personal positioning system for a wearable computer. In *Digest of Papers. Third International Symposium on Wearable Computers*, pages 37–43. IEEE Computer Society, 1999.

[2] Brian Clarkson and Alex Pentland. Unsupervised clustering of ambulatory audio and video. In *ICASSP'99*, 1999.

[3] Brian P. Clarkson. Recognizing user's context from wearable sensors. Technical Report 519, Vision and Modeling Group, MIT Media Lab, January 2000.

[4] Andrew R. Golding and Neal Lesh. Indoor navigation using a diverse set of cheap, weearable sensors. In *Digest of Papers. Third International Symposium on Wearable Computers*, pages 29–36. IEEE Computer Society, 1999.

[5] Pattie Maes, Bruce Blumberg, Trevor Darrell, and Alex Pentland. The alive system: Full-body interaction with animated autonomous agents. *ACM Multimedia Systems*, 5:105–112, 1997.

[6] Natalia Marmasse. commotion, a context-aware communication system. Master's thesis, Massachusetts Institute of Technology, 1999.

[7] Nelson Minar, Matthew Gray, Oliver Roup, Raffi Krikorian, and Pattie Maes. Hive: Distributed agents for networking things. In *Proceedings of ASA/MA'99, the First International Symposium on Agent Systems and Applications and Third International Symposium on Mobile Agents*, 1999.

[8] Bradley Rhodes. The wearable remembrance agent: A system for augmented memory. *Personal Technologies Journal Special Issue on Wearable Computing*, 1:218–224, 1997.

[9] Thad Starner, Dana Kirsh, and Solomon Assefa. The locust swarm: An environmentally-powered, networkless location and messaging system. In *Digest of Papers. First International Symposium on Wearable Computers*, pages 169–170. IEEE Computer Society, 1997.

[10] Thad Starner, Bernt Schiele, and Alex Pentland. Visual context awareness in wearable computing. In *Digest of Papers. Second International Symposium on Wearable Computers*, pages 50–57. IEEE Computer Society, October 1998.

[11] Roy Want, Andy Hopper, Veronica Falccao, and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, January 1992.

Empirically Based Decision-Theoretic Methods for Situated Interaction

Anthony Jameson*
Department of Computer Science, University of Saarbrücken
P.O. Box 151150, D-66041 Saarbrücken, Federal Republic of Germany
`jameson@cs.uni-sb.de`, http://www.cs.uni-sb.de/users/jameson/
Fax: +49 681 302-4136, Phone: +49 681 302-2474

1 Introduction

1.1 Main Claims

This presentation will argue for three claims about research on situated interaction in ubiquitous computing:

1. It is sometimes necessary or useful to obtain information about the situation of the user ($\mathcal{U}$) by interpreting features of $\mathcal{U}$'s behavior that indirectly reflect $\mathcal{U}$'s situation.
2. It is often useful to conduct thorough empirical research at an early stage in the design of such systems, rather than at a later stage of evaluation.
3. Decision-theoretic methods from artificial intelligence are in many ways well suited to the task of making the inferences and decisions required if a system ($\mathcal{S}$) is to adapt appropriately to the user's situation.

The arguments will be illustrated with examples from research in the project READY that has been conducted since 1996, including some previously unpublished results.

1.2 Scenarios

In the first READY scenario, $\mathcal{U}$ was an automobile driver whose car needed some minor repair while $\mathcal{U}$ was on the road. $\mathcal{U}$ obtained assistance from $\mathcal{S}$ by speech via mobile phone.

In the scenario we're currently developing, $\mathcal{S}$ is a mobile system that is lent to travelers who make use of a large airport (e.g., Frankfurt Airport). $\mathcal{S}$ multimodally answers $\mathcal{U}$'s questions about various aspects of the use of the airport (e.g., how to get to $\mathcal{U}$'s departure gate as quickly as possible).

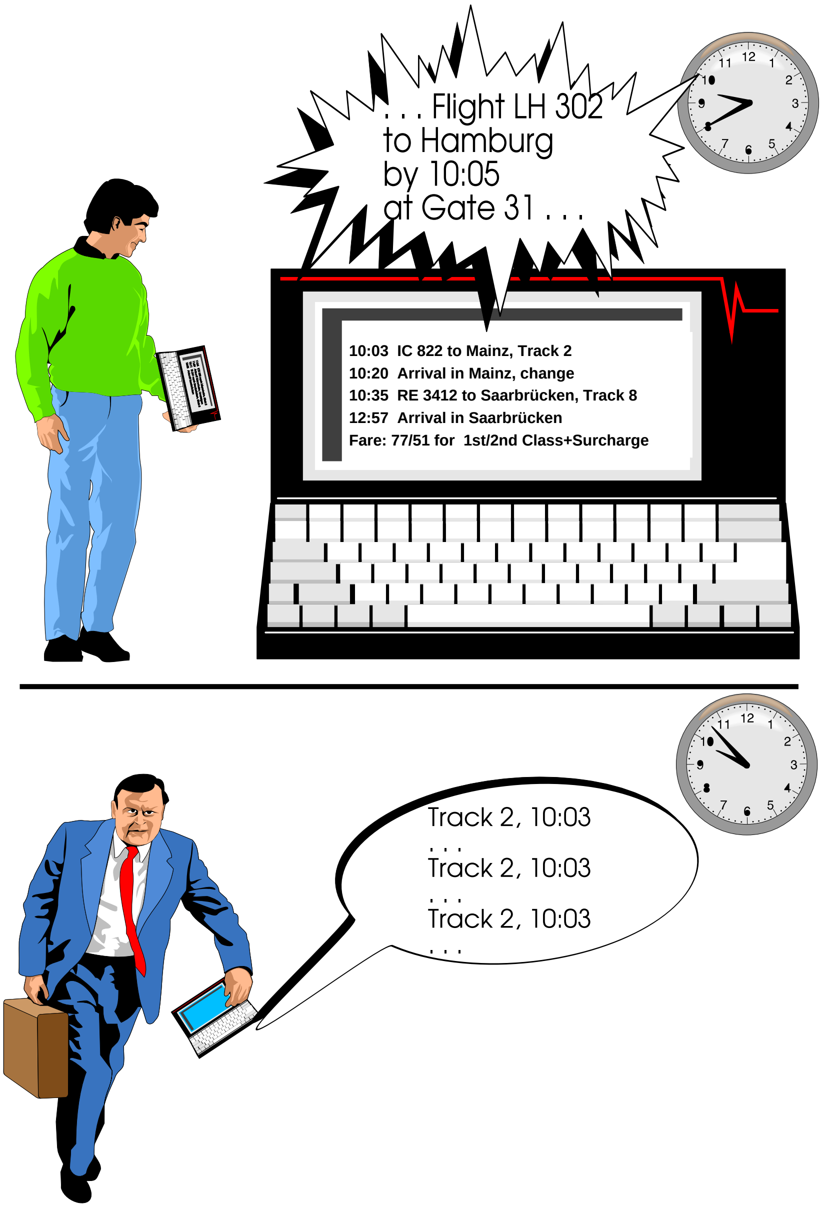
In both scenarios, we have not hooked the system up to actual speech recognizers or synthesizers (although this extension is planned for a later phase). Instead, speech input and output have been simulated. The function of our prototypes is to make inferences and decisions.

The aspects of $\mathcal{U}$'s situation that we have concentrated on are:

1. the extent to which the situation creates *cognitive load* for $\mathcal{U}$, making it harder for $\mathcal{U}$ to perform his tasks and to interact successfully with $\mathcal{S}$; and
2. the extent to which $\mathcal{U}$ is under *time pressure*.

---

*This research is being supported by the German Science Foundation (DFG) in its Collaborative Research Center on Resource-Adaptive Cognitive Processes, SFB 378, Project B2, READY.



10:03 IC 822 to Mainz, Track 2
10:20 Arrival in Mainz, change
10:35 RE 3412 to Saarbrücken, Track 8
12:57 Arrival in Saarbrücken
Fare: 77/51 for 1st/2nd Class+Surcharge

Figure 1. *Illustration of the type of situated interaction aimed at it the* READY *airport scenario.*

2 Using the User’s Behavior as Evidence About the Situation

2.1 Motivation

The most straightforward way for a system to obtain information about $\mathcal{U}$’s situation is through sensors or through data that are more or less directly available to $\mathcal{S}$ (e.g., information about the estimated time of departure of $\mathcal{U}$’s plane). Although READY uses some information of these types, the focus is on analyzing $\mathcal{U}$’s behavior so as to make inferences about the situation and its effects on $\mathcal{U}$.

Up to now we’ve concentrated on the analysis of $\mathcal{U}$’s speech input: How can $\mathcal{S}$ infer from features of $\mathcal{U}$’s speech whether $\mathcal{U}$ is subject to unusual cognitive load or time pressure. We’re now looking into similar questions about other aspects of $\mathcal{U}$’s input, such as $\mathcal{U}$’s use of a scrolling/pointing device.

There are several reasons why it may make sense to focus on $\mathcal{U}$'s behavior in this way:

1. Sometimes more direct evidence about $\mathcal{U}$'s situation is not available. For example, if $\mathcal{U}$ is communicating with speech via a mobile phone, it's unlikely that $\mathcal{S}$ will be able to get information from sensors in $\mathcal{U}$'s environment or on $\mathcal{U}$'s body.
2. Methods for interpreting $\mathcal{U}$'s behavior tend to be more generalizable than those that rely on other sources of information. For example, there are a great many factors that can cause high cognitive load in $\mathcal{U}$. The particular factors may vary from one domain or context to the next, so that it's hard to develop ways of capturing them directly. To the extent to which we're interested in the influence of the situation on $\mathcal{U}$'s psychological state, it may be best to assess it more directly on the basis of $\mathcal{U}$'s behavior.

A difficulty is that $\mathcal{U}$'s behavior often yields only unreliable evidence about $\mathcal{U}$'s current psychological state. But the same difficulty applies when more direct information about the situation is interpreted with a view to inferring $\mathcal{U}$'s psychological state.

2.2 Examples

There already exists a large body of experimental evidence concerning the question of how cognitive load is reflected in features of a speaker's speech (see Berthold & Jameson, 1999, for a brief overview). Some of the more important consequences of cognitive load are an increase in the number and duration of various types of pauses, a slight reduction in the articulation rate, and a higher frequency of sentences that are started and then broken off. We are currently conducting an experiment that yields more directly relevant data about the role of symptoms like these – and also about their dependence on other aspects of the speaker's situation, such as the relative priorities of speed and quality in the production of speech.

3 Performing Empirical Studies at an Early Stage

3.1 Motivation

When one thinks of the role of empirical studies in system design, the first thing that comes to mind is usually the evaluation of prototypes. We believe that, at least in the area of situated computing, more attention should be devoted to studies that create an empirical basis for the system design at an early stage. Without a solid empirical basis, a later evaluation study is likely to reveal that the system doesn't work very well, without giving much indication of how it could be improved.

There are a lot of questions about the consequences of features of the situation which cannot be answered reliably on the basis of a priori considerations.

3.2 Examples

Here are some examples of empirical studies that we've performed, in addition to the experiment mentioned above:

1. We did thinking-aloud studies with firemen who answer emergency calls, to see how they assess a caller's situation on the basis of the evidence that comes over the phone line and how they adapt their own behavior accordingly.
2. We analyzed the transcripts of a field study in which our first usage scenario was simulated, with the role of the system being taken by an experienced auto mechanic. This

analysis revealed, for example, which features of a user's speech occur frequently enough to be potentially useful as evidence (Berthold & Jameson, 1999).

3. We conducted a psychological experiment to see how a help system's instructions should best be adapted to the user's current cognitive load (see Jameson, Großmann-Hutter, March, & Rummer, 2000).

4 Employing Decision-Theoretic Methods From Artificial Intelligence

4.1 Motivation

As was mentioned above, much of the evidence that a system $\mathcal{S}$ can obtain about $\mathcal{U}$'s current situation and/or psychological state is unreliable: Often, it is only on the basis of multiple pieces of evidence that $\mathcal{S}$ can make a useful, though still uncertain, inference. Bayesian networks are a powerful technology for processing this type of evidence. (See Pearl, 1988, for the classic exposition and Jameson, 1996, for an introduction that includes references to many user-adaptive systems.) In particular, dynamic Bayesian networks make it possible to model properties of the situation and the user that change over time (see, e.g., Jameson, Schäfer, Weis, Berthold, & Weyrath, 1999).

Decisions that are made – implicitly or explicitly – by situation-aware systems need to take into account multiple factors and goals, as well as uncertainty about the relevant variables. Decision-making techniques such as *influence diagrams* offer ways of dealing with these complications.

Methods for *decision-theoretic planning* (see, e.g., Boutiller, Dean, & Hanks, 1999) make it possible for a system, when deciding what to do next, to consider how the next few steps in an interaction might proceed. For example, when deciding how to present a route description, $\mathcal{S}$ can consider how likely it is that $\mathcal{U}$ would fail to understand particular parts of the description—and what it would have to do to recover from such a failure.

Finally, decision-theoretic methods include ways of learning user models automatically on the basis of empirical data. It is therefore possible to base a system's inference methods more or less directly on the type of empirical data that was discussed in the previous section (see, e.g., Großmann-Hutter, Jameson, & Wittig, 1999; Jameson et al., 2000).

4.2 Examples

In the workshop presentation, concrete examples of the application of decision-theoretic techniques will be given, as time permits. The emphasis will be not on the technical aspects but on the extent to which these methods constitute useful tools for those who develop systems for situated interaction.

5 Concluding Questions

As a way of opening the discussion, the participants will be asked for their evaluation of the claims just presented. In addition, the question will be raised of how the methods presented here can best be combined with those mainly used by the other workshop participants.

6 References

Note: All of the listed papers that were cowritten by the present author are available via the READY web site: http://w5.cs.uni-sb.de/ ready/.

Berthold, A., & Jameson, A. (1999). Interpreting symptoms of cognitive load in speech input. In J. Kay (Ed.), *UM99, User modeling: Proceedings of the Seventh International Conference* (pp. 235–244). Vienna, New York: Springer Wien New York.

Boutiller, C., Dean, T., & Hanks, S. (1999). Decision-theoretic planning: Structural assumptions and computational leverage. *Journal of Artifical Intelligence*, *11*, 1–94.

Großmann-Hutter, B., Jameson, A., & Wittig, F. (1999). Learning Bayesian networks with hidden variables for user modeling. In *Proceedings of the IJCAI99 Workshop "Learning About Users"*. Stockholm.

Jameson, A. (1996). Numerical uncertainty management in user and student modeling: An overview of systems and issues. *User Modeling and User-Adapted Interaction*, *5*, 193–251.

Jameson, A., Großmann-Hutter, B., March, L., & Rummer, R. (2000). Creating an empirical basis for adaptation decisions. In H. Lieberman (Ed.), *IUI2000: International Conference on Intelligent User Interfaces.* New York: ACM.

Jameson, A., Schäfer, R., Weis, T., Berthold, A., & Weyrath, T. (1999). Making systems sensitive to the user's changing resource limitations. *Knowledge-Based Systems*, *12*, 413–425.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference.* San Mateo, CA: Morgan Kaufmann.

POSITION PAPER

Phil Cohen, David McGee, and Sharon Oviatt
Center for Human-Computer Communication
Oregon Graduate Institute

This workshop presupposes that the device itself is aware of its location and situation of use. Granted that this will be possible soon (if not now). It will be able to adapt its behavior to suit the situation, task, and user. However, just as strongly, the user will adapt his/her usage patterns to suit the situation and location. For example, user studies in our laboratory and others have shown that users of multimodal systems will prefer to speak in some situations, but not others; will prefer to handwrite/draw/gesture in some environments and for some contents, will alter their behavior if one modality is failing (e.g., noisy environment), or if they suspect that they will be unable to perform adequately (e.g., not know how to pronounce a foreign surname).
Likewise, some user responses to the environment will lead to much worse system performance, as when a user adjusts his/her volume and speech characteristics to noisy environments (e.g, Lombard speech). In such environments, multimodal systems will be able to provide both alternative modalities, as well as modalities that when used in combination can overcome the weaknesses of one with the strengths of another.

In the case of our handheld multimodal QuickSet system at OGI (Cohen et al., 1997), we have demonstrated that such mutual disambiguation of modalities (here, speech and pen-based gesture) can enable robust performance in the face of noisy environments, and in response to foreign accents (Oviatt, 1999). We will discuss how mutually compensating multimodal systems can be built, which will be able to adapt to their environment, and to their user's patterns. However, we need to be careful in not adapting too quickly, since many of the environmental factors are transitory (e.g., a plane passing by). Too quick an adaptation could cause the system to be unpredictable.

BACKGROUND:

At OGI, we have built QuickSet, a wireless, collaborative, agent-based multimodal system that is used to control distributed applications, such as control of virtual environments, medical informatics, emergency planning, and military simulations. The system runs on computers ranging from wearables to wall-sized. Recently, it has been extended to support multimodal interaction with paper-based interfaces. Research is underway with Columbia Univ. to develop a multimodal wearable augmented reality system.

References:

Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., and Chen, L., "QuickSet: Multimodal interaction for distributed applications," Proc. of the 5th Int'l ACM Multimedia Conf., ACM Press, New York 1997, 31-40.

Oviatt, S. L., "Mutual disambiguation of recognition errors in a multimodal architecture," Proc. of Human Factors in Computing Systems (CHI'99), ACM Press, New York, 576-583.

Talking wearables exploit context

Sabine Geldof
Vrije Universiteit Brussel - AI laboratory

Pleinlaan 2
B-1000 Brussel
sabine@arti.vub.ac.be

Jacques M.B. Terken
IPO, Center for Research on User-Interaction

Eindhoven
j.m.b.terken@tue.nl

1. Introduction

This paper presents results from the COMRIS1 project in which a wearable, advice giving device is being developed. The main goal is to support a user in an information-rich context. We take the conference situation as an example application, where the user wants to make the best use of her time. Like a parrot on a shoulder, the wearable device whispers hints about interesting people to meet presentations to attend demos and booths to visit. A mechanism of competition for attention (Van de Veiled et. al., 1998) ensures that only appropriate messages reach the user at the right time. Also the formulation of these messages needs to be optimized for communicative effect. The content of the messages is provided by a society of competing software agents and the user's personal agent will annotate it with information about the user's context. We are concerned with the natural language generation (NLG) technology that allows to tune the message form to the user's context in order to enhance communicative effect of the interaction between device and user.

2. Multiple dimensions of context

We claim that wearable devices need to take multiple dimensions of the user's context into account. It is a requirement for all interaction devices but a fortiori for wearable devices, since they are imbedded in the real-world context of the user. Moreover, the use of wearable technology allows for sensing the user's context appropriately, which can be done only indirectly in a conventional desktop environment.

---

[1] Co-habited Mixed Reality Information Spaces is a long-term research project funded by the EC within the framework of I3 (Intelligent Information Interfaces, http://www.i3net.org/)

The 3 dimensions we envisage are discourse context, physical context and user profile context (Geldof, 1999a). Discourse context is a notion inherent to all linguistic interaction (devices): what information has been provided during previous interactions? Clearly, what was said before, influences what follows. Consider an example: 'Penn Sill will give a presentation on Natural Language Generation. *She* proposes to meet you *right after that session*.' This context dimension has been extensively studied in the area of Natural Language Processing (see Dale, 1992 for example on referring expressions). In fact, it is often the only perspective accounted for in Natural Language Processing research concerned with context sensitivity (e.g. Krahmer & Theune, 1998). However, also the physical context of the hearer (and speaker) matters. The following example shows that an utterance sounds more natural if it refers to the hearer's current position in time and space: '*In five minutes*, there will be an interesting demo on context sensing at the booth of the COMRIS project in the demo hall, *just around the corner*.' While most NLG systems haven't paid much attention to this dimension (Maybury, 1991 may count as an exception) probably due to the fact that it is difficult to sense the physical context in a desk-top environment, the advent of wearable systems creates real opportunities for this type of context sensitivity. However, current wearables focus on very low-level physical data sensing like location, motion, lighting environment, body sensing. This paper argues for more abstract physical context sensing. Finally, communication can also be more efficient if the speaker takes into account information about the interest profile of the hearer, as in: 'There will be a presentation on multi-agents systems and scalability, *a topic you're definitely interested in,* this afternoon in Room 4b.' Adaptivity to user characteristics has also received the attention of earlier NLG research (e.g. Paris, 1988).

Assuming that it is an inherent goal of wearable devices to be as least intrusive as possible, i.e. to integrate smoothly in the user's environment, we claim that there is a need for them to take into account multiple dimensions of user context. Talking wearables not only should take into account what they previously mentioned to the user, but also where and when the user is at the moment of utterance and finally what are her interests and priorities. The better a wearable device is able to sense these contextual factors and to adapt its (linguistic) interaction behaviour to these factors, the less intrusive and the more effective it will be. We proposed (Geldof, 1999a) a framework for representing and reasoning with contextual information in view of spoken interaction.

In COMRIS, contextual information is added as an annotation to the input of the text generation component. This means that when available or applicable, contextual annotations are added to the entities to be realised textually. The text generator[2] then takes into account these annotations to choose among alternative ways of realising some entity (Geldof, 1999b). Linguistic (or discourse) context is represented as an integer (labeled 'lcv') indicating the distance in the discourse history, both for the concept and the particular instance of an entity to be mentioned. Physical (or extra-linguistic) context consists of an integer (labeled 'elcv') expressing the spatial or temporal distance of a location or time expression w.r.t. the user's current position. Profile context annotations

---

[2] We use a tool for template-based NLG, TG/2 developed by DFKI and used within COMRIS under a research agreement licence. We'd like to acknowledge Dr. Stephan Busemann for his advice.

(labeled ‘pv’) indicate the degree of interest of the listener in a particular domain entity to be enunciated. The latter information is taken from the user profile which can be acquired directly (e.g. through a form to be filled out when registering for the conference) or indirectly (e.g. using learning and information retrieval techniques). The physical context is acquired through interpretation of sensors on the wearable device, often in combination with beacons. Discourse information is maintained in a dedicated module, updated after each utterance that is generated. An example of contextually annotated input to the text generator and the corresponding output is given:

```
(defparameter *propa2*
           "[(COOP propagandist)
(EVENT [
          (EVTYPE \"presentation\")
          (EV_ID \"s2\")
                      (KEYWORD_LIST   [(FIRST   [(T_NAME  \"speech generation\")   (T_ID
\"657868\")])
                                       (REST  [(FIRST  [(T_NAME  \"discourse markers\")  (T_ID
                          \"202903\")])  (REST [])])]) ])
      (PERSONROLE gives)
      (TAKESPLACE [(ID \"loc-004\")  (NAME \"room 04\")] )
      (START [(DAY 28)  (MONTH 12)  (TIME [(HOUR 11)  (MINUTE 00)])])

      (CONTEXT [(PERSON_PV [(PNUMVAL %pnv)  (P_QVAL %pqval)])
                (PERSON_LCV [(CONCEPT %plcvc)  (ENTITY %plcve)])
                (TOPIC_PV [(FIRST [(NUMBER 3)])
                         (REST [(FIRST [(NUMBER 3)])  (REST [])])])
                (TOPIC_LCV [(FIRST [(CONCEPT 5) (ENTITY 5)])
                                              (REST  [(FIRST  [(CONCEPT  5)  (ENTITY  5)])
[])])])
                (DATE_ELCV .12)
                (LOC_ELCV \"on the same floor\")  ]) ]")
```

[There will be *[an interesting presentation /] on *[speech generation /] and *[discourse markers /] taking place *[in about 10 minutes /] in *[room 04 /] , *[on the same floor /] .///][3]

Without taking into account the context annotations, the output would be:

[There will be *[an interesting presentation /] on *[speech generation /] and *[discourse markers /] taking place on *[December 28th /] at *[11 o'clock /] in *[room 04 /] .

3. Local and global context sensitivity in text generation

In our first experiments with context-sensitive NLG, we focused on how to replace some expressions by the corresponding context-sensitive ones, as in the examples in the previous section. These adaptations are local, i.e. at the lexical level. On the other hand, the context also determines the form of utterances globally, i.e. the sentence structure. Linguistic phenomena of various levels are concerned: word order, morpho-syntactic choices and prosody. In fact, when observing human language production, one can notice

---

[3] This NLG output is annotated for further speech processing: [ ] indicate phrase domains, / // /// phrase boundaries and * prosodic accent.

that we use different constructions to express the same semantic content in varying contexts. Consider the following pairs of allosentences:

Within 5 minutes, Josep Arcos will give a presentation about multi-agent societies.
Josep Arcos will give a presentation about multi-agent societies within 5 minutes.
A presentation by Amanda Huggenkiss will start in 5 minutes.
The presentation by Amanda Huggenkiss will start in 5 minutes.
Penn Sill will give a PRESENTATION[4] on context SENSITIVE natural LANGUAGE generation.
PENN SILL will give a presentation on context sensitive natural language generation.

These variations of the sentence form can be accounted for in terms of information structure theory (Lambrecht, 1994). According to this discourse pragmatics study, a speaker structures the information he wants to convey, in terms of the roles of topic and focus. These terms received various interpretations in different frameworks and depending on whether the viewpoint is cognitive, semantic of pragmatic (Gundel, 1999) but there seems to be a general consensus that in communication, one can distinguish between 'new' and 'given' information. According to Lambrecht, the topic of an utterance is the (more or less established) entity *about which* new information is communicated. The focus is the part of the information that is unpredictable for the hearer at the moment of utterance[5]. These roles, externalised by varying sentence forms, are attributed by the speaker to information elements based on their presumed identifiability and activation in the hearer's mind. In other words, the speaker hypothesises about the attention state of the listener and structures his utterance accordingly in terms of the roles of topic and focus. The mapping between this information structure and the form of sentences is not straightforward but constitutes a very rich phenomenon where many grammatical mechanisms enter into play. Nevertheless, in English, prosodic accentuation usually indicates the role of focus, while definite, unaccented expressions are the preferential marking of topic. Also, Lambrecht distinguishes among 3 sentence structures, according to whether (1) the predicate, (2) an argument or (3) the whole sentence receives the role of focus. He also states that the unmarked, hence most natural sentence construction is the first one, where the topic-role is assigned to an 'activated' discourse referent and the predicate is in focus.

This finding confirms the importance of striving at context-sensitivity in HCI, as we learned in COMRIS. Without bothering about context sensitivity, each message of the parrot contains a maximum of 'new' information, as in (3) the sentence focus construction. Presumable, attending new information requires cognitive effort from the listener. We claim that we can reduce this cognitive effort of the user to attend her parrot's messages by presenting 'new' information in connection with an entity that is somehow activated in her mind. An entity, detected as present in the listener's current context (be it physical, discourse or even profile context) can be supposed identifiable

---

[4] Focused expressions are conventionally represented with UPPERCASE and normally receive prosodic accent, i.e. they are phonetically prominent.

[5] Lambrecht demonstrates that 'unpredictable' is a more appropriate characterisation of the focused information, than the term 'new', which is used in most frameworks. We use single quotes around the term 'new' to indicate this refinement.

and with high potential for activation in her mind, thus a good candidate for the role of topic. This clarifies from a cognitive viewpoint why the linguistic construction of topic-focus (1) is the most natural one. Our strategy for topic-focus assignment in COMRIS is guided by the principle of avoiding the complete sentence to be in focus. The details of the strategy are described elsewhere (Geldof, 2000), but the following scenario illustrates the idea.

Suppose the user just attended a talk on ‘agents and scalability’ by Josep Arcos. A message with the following propositional content happens to be pushed at that time: proximity alert for Dave Roberts, whose main interest is scalability. The default message (without topic-focus strategy) would be:

> DAVE ROBERTS, who shares your interest in SCALEABILITY, is currently in your NEIGHBOURHOOD.

This is a sentence focus construction, where the whole sentence receives the role of focus, i.e. ‘new’ information to be communicated to the user. However, if the wearable device can detect that scalability is very likely to be at the forefront of the user’s mind, due to her recent activity of attending a talk on this subject, it should try to assign this the topic role, as in the following utterance:

> Scalability is also an interest of DAVE ROBERTS, who is currently in your NEIGHBOURHOOD.

The topic assignment is realised by realising ‘scalability’ at the beginning of the utterance and by deaccenting it.

There are indeed a number of morpho-syntactic mechanisms to introduce and realise topics, just as there are alternatives for marking focus. By implementing an NLG strategy that carefully exploits this type of linguistic expressiveness, we can contribute in several ways to enhanced HCI. Note that reducing the number of prosodic accents within an utterance decreases the cognitive effort required from the listener, but also variation in sentence construction makes the interaction more natural and hence better.

4. Requirements for context sensing

The above sketched strategy for making the speech output of a wearable device context-sensitive depends on quite specific information about the context of the user. Especially the global adaptation (i.e. sentence structure) since it is based on assumptions about the attention state of the user. We found out that one way to hypothesise about which entities are at the forefront of the user’s mind, is to detect which activity she is currently involved in and which are the concepts that can be linked to that activity. The use of an ontology of the application domain is very important in that respect. For instance, in the conference situation, events are prime entities and are linked to persons (playing a role in these events) and to keywords (characterising the contents of these events). As a result we can formulate a hypothesis about persons and keywords being at the forefront of a user’s mind depending on the particular activity she is involved in (e.g. a talk given by person X about keyword Y). The latter needs to be determined on the basis of raw context information (i.e. the current location of the user) related to generally available information (i.e. the conference programme). A similar idea is found in (Schmidt et al., 1999). They also distinguish between physical and logical sensors. Other projects reduce the need for symbolic information about context by limiting the adaptive behaviour to non-content related features, presentation modes such as volume, message length or even spatial audio, e.g. Nomadic Radio (Sawhney and Schmandt, 1999). However, we learned in COMRIS that the combination of ‘raw’ context information with symbolic information is a prerequisite when dealing with verbal interaction between devices and humans.

5. Conclusion

In this paper we have advocated the consideration of multiple user context dimensions in order to improve the speech output of a wearable device. Contextual information determines the utterances both locally and globally. With local context sensitivity we mean the use of context-relative expression instead of generally applicable, context-independent formulations. Global context sensitivity in speech output refers to the structure of utterances, determined by the speaker’s hypothesis about the listener’s attention state. Both this hypothesising and the realisation in terms of grammatical categories constitute a very complex phenomenon that received many different interpretations in different frameworks. We don’t claim to have the final word on this issue, but propose a practical (simplified?) framework (based on an extensive linguistic study of observed human language use) in order to experiment with these phenomena, now that wearable devices become available. The acquisition of raw contextual data and its interpretation in combination with domain and application specific knowledge remains a critical issue, not to be neglected. Our contribution to the field of HCI consists in proposing a viable way to capture and interpret contextual information in view of more natural, thus less intrusive verbal interaction with the user.

References

Dale, R. (1992) Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes. Cambridge Mass.: MIT Press

Geldof, S. (1999a) Parrot-talk requires multiple context dimensions. In: Bouquet, P., et.al. (Eds) (Proceedings of the 2nd International Conference on) Modeling and Using Context. LNAI Volume 1688, Heidelberg: Springer Verlag, 467-470

Geldof, S. (1999b) Templates for Wearables in Context. In: Becker, T. and Busemann, S. (Eds) "May I Speak Freely?" Between templates and free choice in natural language generation. Workshop at the 23rd German Annual Conference for Artificial Intelligence (KI'99). DFKI-D-99-01, 48-51

Geldof, S (2000) From Context to Sentence Form. *Submitted* to International Natural Lanugage Generation Conference

Gundel, J. (1999) On different kinds of focus. In Bosh, P. and van der Sandt, R. (Eds.) Focus: Linguistic, Cognitive & Computational Perspectives. Cambridge, Cambridge University Press, 1999. p.293-305

Krahmer, E. and Theune, M. (1998) Context sensitive generation of descriptions. In Proc. of the 5th Intl. Conference on Spoken Language Processing (ICSLP'98) Sydney, Australia 1998, Vol.4, 1151-1154

Lambrecht, K.(1994) Information structure and Sentence Form. New York (NY): Cambridge University Press, 1994

Maybury, M. (1991) Topical, temporal and spatial constraints on linguistic realization. Computational Intelligence, **7**, 266-275

Paris, C. (1988) Tailoring Object Descriptions to a user's level of expertise. Computational Linguistics, Vol. 14, Nr. 3 (Sept. 1988)

Sawhney, N., Schmandt, C. (1999) Nomadic Radio: Scaleable and Contextual Notification for wearable audio messaging. In Proc. of CHI'99, the ACM SIGCHI Conference on Human Factors in Computing Systems. May 15-20, Pittsbrgh, Pennsylvania.

Schmidt, A., K.A. Aidoo, A. Takaluoma, U. Tuomela, K Van Laerhoven, and W. Van de Velde (1999) Advanced Interaction in Context. In H. Gellersen (Ed.) Handheld and Ubiquitous Computing, Lecture Notes in Computer Science No. 1707, Springer-Verlag Heidelberg: 1999, p. 89-101

Van de Velde, W., Geldof, S., Schrooten, R. (1998) Competition for attention. In: Singh, M.P., Rao, A.S., and Wooldridge, M.J. (Eds) Proceedings of ATAL: 4th International Workshop on Agent Theories, Architectures and Languages, LNAI Volume 1365, Heidelberg: Springer Verlag, 282-296

Aware Community Portals: Shared Information Appliances for Transitional Spaces

Nitin Sawhney, Sean Wheeler and Chris Schmandt
Speech Interface Group
MIT Media Lab
20 Ames St., Cambridge, MA
{nitin, swheeler, geek}@media.mit.edu

Abstract

People wish to maintain a level of awareness of timely information and activity of others in a variety of social settings at different times of day. Access to situationally relevant information in the shared environment of the participants provides a means for better exchange, coordination and negotiated order in a community. One approach to enhance community awareness combines sensing of audio/visual context with filtering and display of relevant information in a shared workspace. The challenge is to develop means for lightweight interaction and communication using these 'shared information appliances'.

In this paper, we describe the design of an experimental project that explores such issues within a community environment at the MIT Media Lab. We are investigating peripheral interfaces that can be used in a casual manner in transitional spaces. Such methods can be designed in the context of casual workplace domains, distributed workgroups, and everyday public spaces.

Situated Interaction in Transitional Spaces

To understand the role of shared community appliances and interaction in a public context, we pose a number of questions that may suggest different design approaches. Many such issues are encountered in everyday community appliances which purposely draw attention in public spaces, such as information kiosks in train stations, or an electronic whiteboard in a meeting room. Such appliances tend to be situated in the heart of an area, surrounded by peripheral and transitional spaces, which are typically highly frequented but under-utilized. We are interested in exploring means for enabling brief encounters with contextually relevant information in such transitional spaces.

Why is information about context necessary within a community? One view is that information permits co-ordination and negotiation between community members, clarifying and democratizing their decisions. Another view is that it provides assurances about the social order and one's role in the community. It allows one to stay *in the loop*. Hence transparent public access to situationally relevant information about the community is a desirable goal. The Portholes project [Dourish92] provided distributed awareness via periodic images of others, helping maintain a sense of community. Piazza [Issacs96] supported spontaneous encounters based on shared tasks. We believe shared information in a community space may trigger richer interactions.

Whose needs in the community should the system address? A community is held together partially by spatial proximity, established social orders, and a set shared interests and goals. The question of who belongs in a community could be broadly defined to include its current physical inhabitants, as well as on-line individuals and visitors who may engage in it briefly. Hence shared displays must allow others a level of access, while preserving community standards for privacy and anonymity if desired.

When should a system in transitional space draw attention to itself? An information system that provides continuous alerts or requires active user engagement is distracting. Designing graceful shared systems that coexist with the environment, requires a means to detect when an appropriate interruption is meaningful, based on the context of its participants. [Sawhney 99]

How does interface and modality affect transitional use of space? Many systems tend to utilize high-bandwidth interaction or modalities in a manner that asserts their presence in the environment. We take the converse view that a system, which inhibits the original use of its environment, may have undesirable effects. Hence the presence of the system itself may draw undue attention or distract from the primary social activity of the environment. Lightweight interaction and ambient modalities allow greater cohesion with the environment. Use of human intuitions about social distance [Goffman63], as well as our kinesthetic senses, is underutilized in current systems. We propose using movement and proximity as a *peripheral interface.*

Let's now consider how these questions could be addressed in the design of a shared information appliance for a casual workplace environment such as the 'Garden'.

Social Setting: The Garden Community

The MIT Media Lab is an active research environment and a casual social space inhabited by a diverse group of students and faculty, and frequented by many visitors during the day. The lab has several open spaces for workgroups, and thrives on the interaction among and between such groups to maintain a fluid and social environment. We chose to utilize one such workspace and 'social collective', called the 'Garden' as the primary environment for our exploratory project. The Garden inhabitants are involved in responsive media and interaction research, hence are somewhat more receptive to using prototype HCI systems. The Garden workspace (see figure 1) was recently renovated to further open-up the space and install digital media infrastructure, including computer-controlled lighting, sound, video projections and some cameras (in focused areas). The space consists of clusters of workstations, a large HDTV display, a glass-walled conference room, offices, and an expanded hallway. The Garden hallway faces the entire space and is visible by most people there. This presents a good opportunity to utilize this transitional area for dynamic information display.

Figure 1: The Garden workspace with projected display in the hallway

An existing shared appliance called the Garden-Box, a digital audio jukebox developed at the lab, has been in use here for some time. The system plays MP3 music in the environment, selected via a web-based interface. Shared use of such an appliance requires implicit negotiation among participants regarding the type and volume of music played at different times of day. Continuous use and the kind of music played also serves to define the changing interests of the community and a subtle awareness of others in the space.

We attempted to leverage the infrastructure and negotiated social order within the Garden community to propose an on-going design of shared appliances for display of situated information and awareness patterns.

Aware Community Portals: Design Exploration

Design Themes

We set out to explore a number of key themes and issues for designing a shared media appliance that utilizes short-term awareness of visual activity and long-term community interests, Some of these themes were further developed in an early prototype.

Community-Filtered Information Glances

As a fundamental premise for a shared information appliance, it is necessary to ensure that the information shown is based on community interests. Hence a collaborative mechanism must be provided to allow users to select relevant external and internal data for periodic updates as well as timely display. The framework should be extensible to allow users to create and integrate new information services easily. Finally users should be able to view the status of such 'broadcasts' and add or remove them as needed. Examples of such services would include news, weather, traffic, but emphasis would be placed on group information such as scheduled events, internal messages and informal postings.

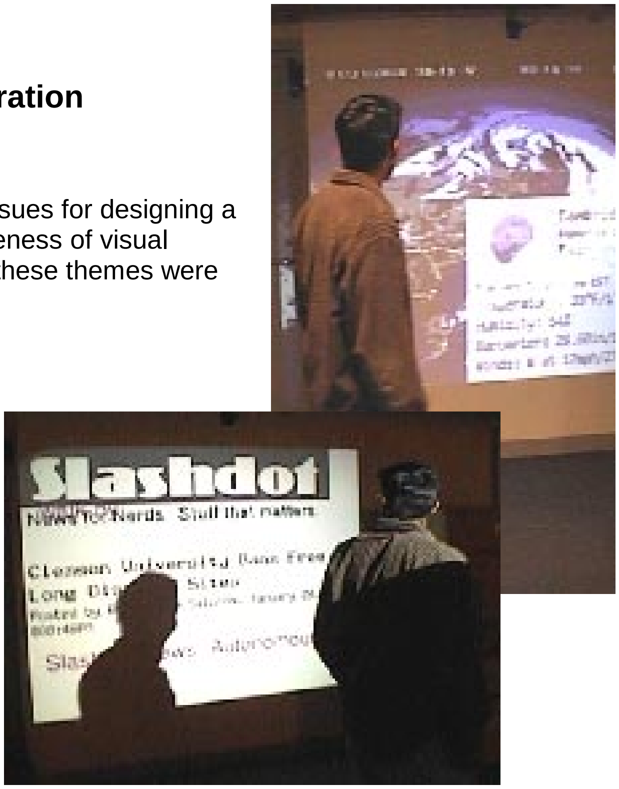


Figure 2: Transitional Interaction - a weather map triggered by the user walking by vs. a news article shown when the user lingers to browse. Notice a potential design element: shadows of prior viewers, cast on the article could be used to gauge community interest in current stories.

Casual and Transitional Interaction

Although selecting services would require greater involvement, the primary model of everyday usage of the portal assumes peripheral and transitional interaction. It is expected that participants maintain subtle awareness of information naturally placed in their environment and be able to engage with the system with minimal effort. Hence a range of unobtrusive presentation and interaction techniques are explored. One example suggests walking by the display and noticing that the meeting scheduled in the conference room has just been postponed. The interaction here is meant to be brief and situationally relevant. The user may then wish to see who's coming or glance at the recent weather or traffic report for possible delays, without much more effort and then simply walk away.

Temporal Awareness Patterns

People remain aware of the activity around them to maintain continuity and retain a sense of the temporal structure of their environment. By capturing the salient traces of activity over time and abstracting into a useful representation, a system can display the regularities and temporal rhythms in the space over time. For example, one may observe the general presence of people shown on a weekly or daily timeline; that allows one to infer the times when most people are around or whether a previously scheduled event was unattended or postponed. Facial glances or shadows of people looking at particular articles on the projected display can be retained to allow one to gauge people's interest in stories (see a simulated example in figure 2). Long-term patterns provide an understanding of the social order of the space and a means to potentially predict the occurrence of activity at specific times of day. Statistical learning of temporal patterns could allow the system to overlay a representation of future activity on a calendar, timeline or other novel temporal maps of community activities.

Exploratory Prototype

The current prototype consists of a projected video display driven by a graphics-rendering engine on an Alpha workstation, live information provided by servers, and active sensing from networked cameras in the environment. This early version serves as platform for design experiments.

Filtering and Display

A content transcoding server (written in Perl) monitors the system's Internet news sources for new information. *Slashdot.com*, a popular technology web-log, was chosen as one news source of potential interest, as many community members visit the site at least once a day. New content is pulled in as an HTML page, parsed, and relevant information is extracted from nodes in the parse tree. The system uses this information to create a graphic rendering for each news story, optimized for the low resolution and other display constraints of the projection system (see figure 2). This system will be extended to handle a variety of information services and a web-based interface will be provided. The portal display engine, written in ISIS [Agamanolis97] (a programming environment for responsive media), renders graphics and text as a video projection, and manages live information queued from servers. Currently the projection periodically shows information such as clock-time, hourly cartoon-strips, MP3 audio titles from music playing in the Garden-Box, and live data (news and weather) from the transcoding server.

Proximity and Movement as an Interface

To maintain a casual and natural interaction, user intent is inferred as they approach the projected display. A phased approach first displays an 'information glance' when new information arrives. When a person is seen *walking-by* the space, a series of images are shown cycling through, depicting the recent stories in memory. If the person stops to *glance* at the display, a preview of the current story (news headlines or weather map) is shown for a short duration. If the person then continues to glance, the system assumes she wishes to *browse* the article in more detail, hence a sequence of related information is shown (see figure 2). After a person leaves, the display gradually fades away. The timing and duration of *information glances* and previews have to be carefully devised and synchronized with movement to provide fluid presentation and interaction without being overtly distracting or prolonging beyond a user's interest.

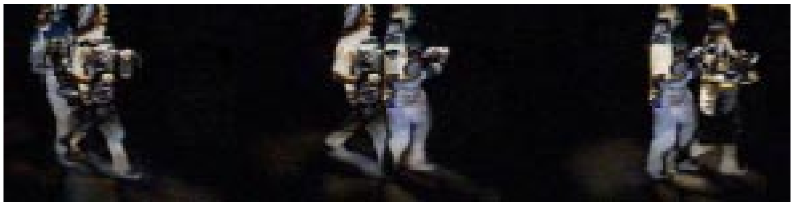


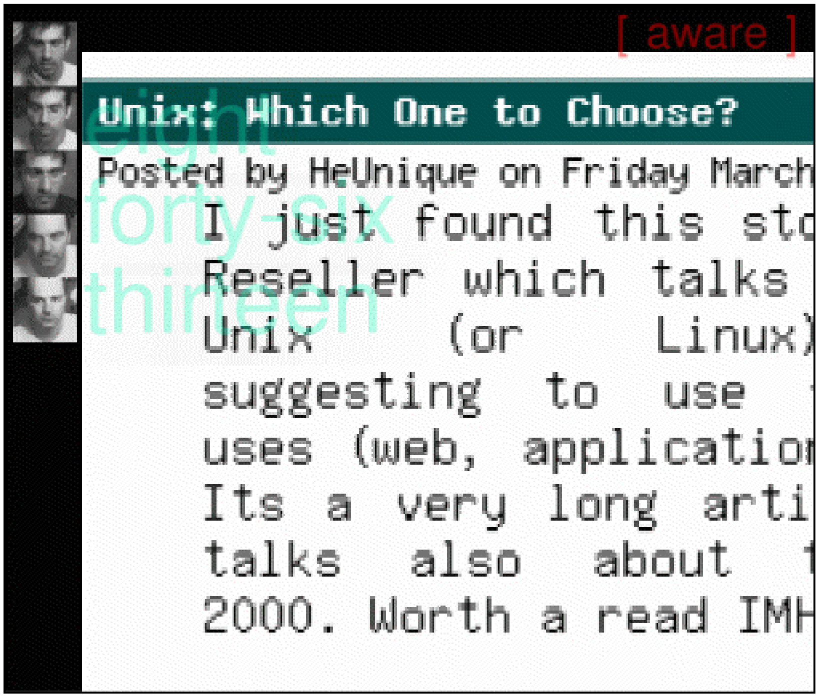
Figure 3: Visual activity detected using image differencing and thresholding. A sequence of such images provides a trace-like representation that could be overlaid on information or abstracted to indicate overall activity patterns.

A networked video camera mounted on top of the portal provides live video via a server to different processes that analyze the video. A video window on the top-right corner of the display shows the view seen by the camera both as a means of providing interaction feedback and assuring people of the purpose of the camera. Movement is detected via image differencing and thresholding techniques in ISIS; these simple techniques operate in real-time (a temporal representation is shown in figure 3). A face-detection system monitors whether a person is looking at the display and triggers a close-up of the current article. When a person continues to maintain a glance (at the camera), the system shows more detail. A bounding box on the person's face indicates the system is tracking their face and indicates that it recognizes a glancing action (see figure 4).

Figure 4: Detecting a person glancing at an article and stepping closer to see more. By gazing at the camera for a longer duration, a person allows the system to capture his face.

Persistence and Awareness

The portal is designed to allow people to maintain persistence of usage over time and provides a form of awareness of the presence of others during the day. One mechanism is to capture people's faces as they browse articles they are interested in. While glancing at an article if people choose to look directly at the camera (to indicate their interest), a bounding-box on the person's face grows gradually (see figure 4) allowing an opportunity to step back if they don't want their face captured. Once captured their face is shown on a timeline in the left corner associated with the article seen (see figure 5). This provides both an awareness of others who have browsed the article (and been around in that physical space recently) as well as a sense of the general popularity of certain articles. It is important to note that the system currently does not recognize the person whose image is captured (maintaining their anonymity to the system), however in the future such recognition may enable personalization of news content or tracking of user's interests (if they choose to do so).



Unix: Which One to Choose?

Posted by HeUnique on Friday March

I just found this sto
Reseller which talks
Unix (or Linux)
suggesting to use
uses (web, applicatio
Its a very long arti
talks also about
2000. Worth a read IMH

Figure 5: A timeline of faces shown on a portion of the large display indicates people's interests in certain articles and also provides an implicit sense of awareness of whose around during the day (most recent person seen at the bottom).

Hence with a simple sequence of *movement* and *proximity*, a range of responsive behavior can be provided along with an implicit form of awareness of others over time. Similar approaches are used in distributed virtual environments to automate communicative social behavior between virtual avatars [Cassell99]. Our view is that such social protocols can be used to provide situation dependent behavior between peripheral information 'actors' and humans in a physical space. Previously stored facial images are used in the Piazza system [Issacs96] to show other users browsing the same document on-line, whereas the portal provides awareness of people's interests and activities in a shared physical space over time. As the system has only recently been up and running, we have not conducted formal evaluations. Informal usage shows that people desire better cues (such as audio/visual or text-prompts) to enable them to understand the different modes of the interface and recognize the range of simple behaviors they can utilize to control the interaction. We hope to share our preliminary observations of the system's usage at the workshop.

Research Directions and Design Implications

Interaction with Mobile Devices

As people carry mobile devices like PDAs in their environment, they may be able to implicitly exchange information between the aware portals in these spaces. This may allow them to post personal messages easily or store updated information (news or changes in schedules) they recently glanced on the display. In this manner, humans act as carriers of contextually relevant information between spaces and communities.

Privacy vs. Awareness

By providing more personal information to a shared portal, one may get more relevant information from the system and allow known others to get better awareness, however there is also a greater potential intrusion of ones privacy [Hudson96]. We need to consider negotiated protocols and interfaces that *perceive* and retain the level of privacy one demands in different situations. Abstract representations of visual activity (such as shadow-views [Hudson96]) and our own experiments with garbled audio techniques point to several design approaches.

Transitional Speech

Although visual sensing provides a simple and unobtrusive form of peripheral interaction, it is not easily extensible for complex tasks. One solution is to provide speech-based interaction when the system detects that the user is engaged in more active use of the display. For example, once the user enters *browse* mode, speech recognition activates to listen for commands and the display provides hints regarding the phrases that can be spoken. Examples may include switching topics directly or requesting specific information. Once the user steps back, recognition could be turned off. We recognize that introducing a spoken modality makes the interaction more directed and less peripheral (especially in a public space), however it may provide rapid access to timely information especially for expert users in a rush. This retains the transitional nature of the interaction as we discussed earlier.

Situated Communication

Most communication devices assume binary and absolute modes for interaction. Telephones can disrupt recipients and require a time-limited conversation between parties rather than an on-going awareness. Video and audio-conferencing permit open connections between participants but typically require complete engagement or attention. Communication technology can now be used to create more graceful means for peripheral awareness and background conversation. The awareness portal may act as a link between distributed spaces, and the proximity of individuals in both spaces may trigger gradual awareness and communication.

Classifying and Representing Salient Activity

One must consider the granularity at which audio and visual events must be segmented and classified for meaningful abstraction of activity information. There are several techniques for establishing context from audio/visual scenes [Clarkson98] as well as extracting high-level activities from motion templates [Bobick97]. There should be an emphasis on modeling and representation of activity derived from both audio and visual features of everyday environments.

Design for Diverse Social Settings

We are currently exploring the use of aware information appliances in a casual workplace scenario. However in public environments such as hospitals, airports and train stations, there will be varying modes of interaction, privacy concerns and benefits of access to timely information or awareness of routines. Awareness provides subtle cues about people’s availability, rhythms, and regularity of activity (or lack thereof) [Mynatt99]. One’s expectations regarding the social and temporal structure of the environment, allows better coordination with others in a workgroup and enhances cognitive well being in a community. Awareness portals placed within elderly communities may provide cognitive support and enhanced mutual awareness within the community or with caretakers and loved ones.

Conclusions

Publicly available information is predominantly accessed through private information appliances. For information of broad community interest, however, it may be more appropriate to integrate the information within the very spaces we occupy [Dourish97]. Providing a platform for shared information access and awareness within a community is a valuable means to support coordination, negotiation and a sense of belonging.

The design of shared information appliances requires consideration of community-driven authoring of content, low-overhead interaction techniques and a framework that takes into account frequent but transitional usage by participants. We have demonstrated an exploratory prototype that provides community-filtered information based on peripheral interaction using visual proximity and motion.

A fluid and situationally sensitive interaction requires appropriate use of modality to the context of use, a protocol for synchronization of information ’actors’ with human activity, and sensitivity to privacy norms in the group. Such systems must have an unobtrusive means for sensing user intention and activity patterns while representing timely information relevant to the community.

Acknowledgements

We wish to thank Stefan Agmanolis and Chris McEniry for supporting us with the ISIS programming environment and their prior work with the projection display.

References

Agamanolis, S., Westner, A., and Bove, Jr. V. M. "Reflection of Presence: Toward More Natural and Responsive Telecollaboration," *Proc. SPIE Multimedia Networks, 3228A,* 1997.

Bobick, A. "Movement, Activity and Action: The Role of Knowledge in the Perception of Motion." *Phil. Trans. R. Soc. Lond.* B (1997) 352, 1257-1265.

Cassell, J. and Vilhjalmsson, H. "Fully Embodied Conversational Avatars: Making Communicative Behaviors Autonomous." Autonomous Agents and Multi-Agent Systems, 2(1), June 1999.

Clarkson, B., Sawhney, N., Pentland, A. "Auditory Context Awareness via Wearable Computing". 1998 Workshop on Perceptual User Interfaces (PUI’98). Nov 4-6, 1998. pp. 37-43.

Dourish, P., and Bly, S. "Portholes: Supporting Awareness in a Distributed Work Group." *Proceedings of CHI ’92.*

Dourish, Paul. "Extending Awareness Beyond Synchronous Collaboration." Position Paper for CHI ’97 Workshop on Awareness in Collaborative Systems, 1997.

Goffman, E. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings.* The Free Press.

Hudson, S. E. and Smith, I. "Techniques for Addressing Fundamental Privacy and Disruption: Tradeoffs in Awareness Support Systems." *Proceedings of CSCW ’96.* pp. 238-247.

Issacs, Ellen, Tang, J.C., Morris, T. "Piazza: A Desktop Environment Supporting Impromptu and Planned Interactions." *Proceedings of CSCW ’96.* pp. 315-324.

Mynatt, E.D., Adler, A., Ito, M., Linde, C., O’Day, V.L. "The Network Communities of SeniorNet." *Proceedings of ECSCW ’99.*

Sawhney, N., and Schmandt, C. 1999. Nomadic Radio: Scaleable and Contextual Notification for Wearable Audio Messaging. *Proceedings of CHI’99*, 96-103.

The role of roomware and sensing technology for supporting narratives in ubiquitous computing environments

Matina Halkia*, Norbert Streitz**

*Starlab Research Laboratories N.V.
Sint Michielslaan 47
Brussels, Belgium
Email: matina@starlab.net

**GMD - German National Research Center for Information Technology
IPSI - Integrated Publication and Information Systems Institute
GMD-IPSI, Dolivostr.15, D-64293 Darmstadt, Germany
Email: streitz@darmstadt.gmd.de

In our view, the study of situated interactions in office environments requires consideration of the following design spaces: mental and social space, information space as well as the physical environment enveloping it (Streitz *et al.*, 1998). It also requires extending context-awareness methods to make accurate and meaningful inferences about human activity in physical space (Halkia & Solari, 1999).

In this workshop, we would like to emphasize the importance of the architectural component in situated interactions and the need to translate the specificity of "place" of physical environments in the information landscape. Buildings in general, and as an example, working environments, have a language of implicit communication, which transparently prompts us to behave accordingly in different types of spaces. In a large public hall with resonant walls we are likely to be quieter than in a room of absorbent cladding. In a corridor we are likely to be moving than standing, in a very important executive office we are likely to be more attentive, careful or formal than in the entrance hall. The cues of physical space whether established by symbol or convention, or communicated by form and material, condition our behavior and movement, and importantly so, our ability to absorb or manipulate information. Physical cues can modify context-awareness and by analyzing action in physical environments we can make useful assumptions about social and organizational context.

We would like to propose introducing narrative and/or cinematic techniques in situated interactions to ensure engaging, cohesive, interactive experiences for people. This may take the form of "scripting of events" in ways that are both predictable and unexpected within a range of possibilities that does not confuse or undermine the user's trust in the system. A useful working term in developing narrative interactive structures is the "memory of the building" whereby the physical envelope is understood as an interface in relation to time rather than as a surveillance or control tool.

Functionality in these information-enhanced, collaborative environments can be achieved by providing so called roomware components (Streitz *et al.*, 1998), i.e. combining familiar interfaces and tools, such as physical objects, chairs, tables, bulletin boards or wall surfaces, pens, key chains with known or emergent computer functionality. Existing examples are CommChairs integrating a chair with a laptop computer, Dynawalls integrating collaborative working methods in writing boards, and InteracTables incorporating intuitive physical gestures in the manipulation of information objects. The Passage mechanism demonstrates how a physical object can be enhanced with information capabilities without obstructing the natural use of the object (Konomi *et al*., 1999).

Designing in context-aware situations requires sensing devices and technologies that can "read" presence and action of people in physical environments as well as meaningfully draw inferences about the dynamic evolution of organizations in time. MUSE is such a system, enabled by a sensor board (TEA board) by Schmidt (*et al*., 1999), which reads light and sound levels as well as detects motion. Using infrared technology we were able to detect presence and location of people and by combining data from the sensor board, and the specific location of events, we were able to make useful distinctions between public and private meetings, importance of office activity, working patterns etc (Halkia & Solari, 1999).

In the future, we will explore how to play out interactive scenarios for social contexts and increased productivity in the workplace in a series of room types and in a continuous, cohesive, and engaging physical space. We will explore traditional design methodologies and their use in hybrid spaces – virtual and real – to ensure simple and effective information manipulation without the drawbacks of information noise or information redundancy. We will also explore non-electronic characteristics of objects and spaces that can be used to tag electronic documents or information objects onto them, therefore assign physical objects to "aware" information without reverting to hardware components.

It is necessary that these aware buildings, integrated furniture and enhanced office aids have a robust and a flexible network infrastructure, which will have to be interoperable, affordable, reconfigurable and safe. It is also important that the network solution will be transparent and will not require drastic physical changes in the building envelope.

References:

1. Halkia, M., Solari, J., (1999). Progress of MUSE (Making User Friendly Spaces) : A context-aware assistant for orientation. In: Streitz, N., Siegel, J., Hartkopf, V., Konomi, S. (Eds.), Cooperative Buildings - Integrating Information, Organizations, and Architecture. Proceedings of the Second International Workshop (CoBuild'99). Springer LNCS 1670. pp. 210-214.
2. Konomi, S., Müller-Tomfelde, C., Streitz, N. (1999). Passage: Physical Transportation of Digital Information in Cooperative Buildings. In: Streitz, N., Siegel, J., Hartkopf, V., Konomi, S. (Eds.), Cooperative Buildings - Integrating Information, Organizations, and Architecture. Proceedings of the Second International Workshop (CoBuild'99). Springer LNCS 1670. pp. 45 - 54.
3. Schmidt, A., Aidoo, K.A., Takaluoma, A., Tuomela, U., Van Laerhoven, K., and Van de Velde, W. (1999). Advanced Interaction in Context. In: H. Gellersen (Ed.) Handheld and Ubiquitous Computing, Lecture Notes in Computer Science No. 1707, p.p89-101. Springer-Verlag Heidelberg.
4. Streitz, N.A., Geißler, J., Holmer, T. (1998). Roomware for Cooperative Buildings: Integrated Design of Architectural Spaces and Information Spaces. In: Streitz, N.A., Konomi, S., Burkhardt, H.-J. (Eds.): Cooperative Buildings - Integrating Information, Organization, and Architecture. Proceedings of CoBuild'98. Springer LNCS 1370, 1998, pp. 4-21.