ACTIVATING EVERYDAY OBJECTS

Roy Want, Mark Weiser
Xerox PARC
3333 Coyote Hill
Palo Alto, CA 94304
[want, weiser]@xerox.com

Elizabeth Mynatt (affliation as of 9/15/98)
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280
mynatt@cc.gatech.edu

1 INTRODUCTION

The confluence of technological advances in wireless technology, mobile computing, novel displays, and sensors in addition to the decreasing cost of computing power provides the opportunity for utilizing computational capabilities in an increasing number of specialized, yet networked devices. By integrating and embedding these devices into "smart spaces," we have the opportunity to move computing power from its specialized location as a desktop PC to distributing computational capabilities into the existing infrastructure and tools of everyday offices. In this paper, we discuss the inherent advantages of enhancing everyday objects with computation, describe three key research problems that must be addressed to achieve these advantages and outline our approach in these areas.

People work in the everyday world. We use, when we are at our most effective, those tools and methods with which we have long experience and familiar skill. These include, for many people: pencils, paper, whiteboards, desktops, paperclips, notepads, and the many other objects of everyday office life. Although the WIMP metaphor tries to replicate these items in a virtual space, it fails to capture their ease of use, their flexibility, and their serendipity.

Instead of using a metaphor, can we activate the everyday objects in the world? Yes. Work in Ubiquitous Computing [5][6][7][8], Augmented Reality[1][3], and Tangible Bits [2] has shown the way. Yet work in all three areas has fallen short of what they hoped to achieve. In all of these areas three key research problems show up again and again, and without much progress. Until these problems are solved, real applications will be difficult to achieve. The three key problems:

- Short-range, in-building location. For example, the most widely used research in-building location system (developed by one of us) is the Olivetti active badge system [4], used at Xerox PARC and elsewhere. But its spatial granularity is tens of feet, and its temporal granularity is tens of seconds. These are values ill-coordinated to human scales of fractions of inches and fractions of seconds which is important for many applications the badge system was not designed to support.

- Coordination of active objects into a single UI. For example, a room may contain a digital camera, a PC, a video projector, an active pen-input screen, a microphone, a printer, a scanner, and several palm pilots. To the human in the room, who can see them all, it is obvious that they form a set of resources that can be coordinated to get work done. But the view from inside the microprocessor in each of these devices is of isolated worlds-in-themselves. How can we bring the machine viewpoint in line with the human?

- Coordination of real and virtual objects. For example, a document that is scanned, printed, emailed, stored in different real and virtual files, pasted on a wall, etc. etc. is still the same document from a human point of view. We think of it as "the same". But to our machines most of these versions don't even exist (e.g. the one pasted on the wall), and the rest are not seen as even related. Can we bring the machine's understanding of relationships, even such a fundamental relationship as equality, in line with the human?

The advance we envision is bring the technical elements of everyday technology in line with the human uses of that technology. The result will be a significant increase in the impedance match between machine and human, leading to much more effective work.

2 OPPORTUNITIES

We envision applications within the everyday paper-based work of government and military offices. The individuals in these offices use many sorts of everyday objects in special and specific ways. The placement of items, as well as the content of items, in these offices is crucial for getting the work done.

The measurable challenges are:

- To increase the flow of decisions and accurate information in the office.
- To decrease the number of mislaid decisions, information, or workflow.
- To fit into the everyday office practice, without requiring significant retraining.

To meet these challenges, we intend to enhance and connect interaction with various individual objects in the office-place. By augmenting common objects, such as a whiteboard or notepad, we leverage everyday office practices. By connecting these devices, we create a networked web of information that can be accessed throughout the office in contrast to requiring access at the original input source.

Once this web of information and devices is in place, the user's model of interaction changes from an application-task based model to an information-flow based model. Since many office activities are centered around information management, our new UI paradigms must adopt an information-centric point of view. For example, instead of producing a design specification document for a piece of software, the activity to be supported is managing the design process and the information associated with that process. That information will appear in many places and formats, such as the project plan, within code itself as well as current design specs.

3 TECHNICAL CHALLENGES

In the area of location, there are many technical methods of doing the job. They are all too expensive, or too prone to interference in an office environment. GPS does not have the resolution and is not easily accessible from inside a building. Luckily, there are many possible approaches to resolving the problem. Multiple-camera data fusion, sonic locators, and passive antenna-detuned triangulation are a few of these. The key challenges are low cost, and effective reliability. The standard to beat is the eyes and hands of the person in the office, who rarely fails to reach out and grasp the document once it is in front of him. Our location technology must do as well.

In the area of UI coordination, there are two challenges. First, it is to understand the user's expectations of the linkage of the office objects. That is to say, what is the semantics that the user imputes to the objects. Secondly, it is to deliver on that imputed semantics. This is a challenge that admits many kinds of solutions, with deep challenges in system integration. For instance, when devices with displays are close enough together having them share parts of the same image. In the end, the challenge is one of design: will it all hold together for the user.

In the area of coordinating real and virtual, the technical challenge is labeling. How can the many versions of the "same" thing be labeled as the same, especially if some of them only have physical manifestations? Supposed they all had barcodes? How would the barcodes by applied, when, how would they be read, how would they be looked up, in what database, controlled by who, accessible where and when? Are barcodes the right thing? Should we consider RFID-tags or other electronic tagging technology?

4 TECHNICAL APPROACHES

4.1 Location on a Planar Surface

In an office environment many aspects of our work are focused on the desktop or the whiteboard. If objects in these areas can be identified along with their location, it is possible to enhance work practice by further augmenting the space. We have considered a variety of tagging options for this task. Infrared (IR) [850nm] is a convenient medium for signaling across distances of up to 25 feet. It is invisible to the human eye, consumes a small amount of power and its transducers are the size of electronic diodes (physically very small). The power requirements for infrequent communication using only a small Lithium cell (180mAhr) can result in an application lifetime as long as one year. In addition, many monochrome CCD cameras are sensitive to the near IR band as well as the visible spectrum, thus providing an inexpensive sensing device that can spatially localize an IR source. In this way physical objects can be augmented with IR-micro-tags that have the desired small size and low power-consumption to make them a viable option. Position sensing can be achieved through image processing by inexpensive and easily deployed CCD cameras. This kind of image analysis is considerably simpler than the generalized case of recognizing arbitrary objects. Our belief is that IR is a pragmatic technology with a fast track for building the required location system.

In more detail, assume the IR microtags periodically beacon their ID (there are however a number of creative ways a tag can be triggered). A CCD camera is used to spatially locate the flash in its image, perhaps requiring reference markers within the same scene to calculate the position. We now need to determine the ID of the tag that has just signaled. There are two strategies for extracting the ID.

In-band: the coded flashing of the IR-tag may be tuned so that the frame rate of the camera is synchronized with its transmission clock. In this case, successive frames of a video stream will result in light and dark patches at the same position in the camera's field of view. A relatively lightweight image processing algorithm can therefore extract the encoded data and generate the ID. The system, although slow for data transfer from a single tag, can process multiple tags without data collisions or contention.

Out-of-band: a coded ID transmission is modulated onto the IR carrier at a high frequency (compared to In-band) using techniques that are well known (e.g., IrDA physical layer). This data can be detected by an independent IR-receiver diode, a pre-amp and a decoder in order to recover the ID. The time the data is received needs to be correlated with a unique flash seen by the camera. If a collision occurs, data will be lost in this system. It is possible to reduce a tag transmission time to a minimum and rely on retransmission, randomization and statistics to successfully read a large number of tags in a suitably long period of time.

Devices that operate as described in 1 & 2 above can be built into a single IR tag that is about the diameter of a 1 cent coin and stands only 5 coins thick. Further miniaturization may be possible.

4.2 Integrating Tags with Applications

By using the tagging technology described above, we may consider each object in a scene, and its position relative to every other object, as part of a unified physical UI. By careful interpretation of an image and by the appropriate assignment of function to objects, a computing system set up to coordinate this environment could then initiate the appropriate set of actions.

In some cases it may not be possible to permanently attach an active tag to an object. For instance, in the case of a slim paper document, for much of its life we may not wish to have an IR microtag stapled to it. However, this document can be printed with a coded label (perhaps invisible to the human eye). By placing a suitably modified version of our tag over the label, on an occasion when coordination with the office environment is required, the tag can read data from the label it is now obscuring and use it to modulate the document identity onto its IR signal.

4.3 User Interface Models

We will need new user interface models for working in smart spaces. The goal is to present an integrated set of devices that work across an interconnected web of information. There are three primary strategies for meeting this goal.

Information Appliances: Each device in the smart space that has a perceptible user interface should be designed as a specific information appliance. Each appliance will have its own capabilities that are made apparent in its affordances. Simple physical affordances include visual display space, mobility and the existence of a writing surface. As physical objects and tools (such as a magnifying glass, a briefcase and a safe) have particular capabilities with respect to physical artifacts, information appliances should have particular capabilities with respect to computational artifacts.

Information Flow: More than ever before, we work in an information-rich environment. Our enhanced spaces will be of limited use if the information manipulated in those spaces is only accessible on one particular device. Similarly, the interconnections between pieces of information need to be maintained as the point of access changes. This model of an evolving web or flow requires new methods for capturing, storing and retrieving information.

Unconventional Media: As we work within a flow of information using a variety of devices, we will need to turn to underutilized methods for interacting in these spaces. Output that can be processed with only peripheral attention is useful for maintaining awareness of the state and activity in a smart space. Possible media for peripheral output includes non-speech audio, shadows and tactile stimulation. New physical input methods such as tilt, rotation and pressure will aid in making interaction with individual appliances more natural and intuitive.

4.4 Scenario-Based Design and Implementation

As we achieve demonstrable results in each of these three areas, we believe it is important to integrate these advances into office applications and services. In a possible scenario, the contents of a paper document could be "thrown" to an augmented whiteboard for shared review and editing. Likewise, the whiteboard display could augment the use of display-limited or display-less devices such as a PDA or the telephone. For example, the whiteboard could augment handwritten to-do lists with a record of calls to be returned and appointment reminders. By tracking the location of various pieces of paper, the results of physical activities such as sorting, grouping or prioritizing could be recorded. The results could then be accessed on a number of devices including a PDA, an augmented whiteboard or a desktop computer.

5 REFERENCES

1. Communications of the ACM, Special Issue on Augmented Environments, 36 (7), 1993.

2. Ishii, H. and Ullmer, B. "Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms," in Proceedings of CHI'97, ACM, March 1997. (http://tangible.www.media.mit.edu/groups/tangible/papers/Tangible_Bits_html/index.html)

3. Mynatt, E.D., Back, M., Want, R., and Frederick, R. "Audio Aura: Light-Weight Audio Augmented Reality." Published in the Proceedings of the Tenth ACM Symposium on User Interface Software and Technology (UIST), Banff, Alberta, Canada, October, 1997. (http://www.parc.xerox.com/mynatt/pubs/audio-aura-uist97.ps.Z)

4. Want, R., Hopper, A., Falcao, V. and Gibbons, J., The Active Badge Location System, ACM Transactions on Information Systems. Vol. 10 (1), 1992, pp. 91-102. (ftp://ftp.orl.co.uk/pub/docs/ORL/tr.92.1.ps.Z)

5. Want, R et al. "The Parctab Ubiquitous Computing Experiment" Book Chapter #2, Mobile Computing, Kluwer Publishing, Edited by Tomasz Imielinski Chapter 2. pp45-101, ISBN 0-7923-9697-9, February 1997. (http://www.ubiq.com/parctab/csl9501-abstract.html)

6. Weiser, M. and Brown, J.S. (1995) Designing Calm Technology .(http://www.ubiq.com/weiser/calmtech/calmtech.html)

7. Weiser, M. "Some Computer Science Problems In Ubiquitous Computing." Communication of the ACM. July 1993. (http://www.ubiq.com/hypertext/weiser/UbiCACM.html)

8. Weiser, M. The Computer of the 21st Century. Scientific American 265(3) 1991, pp. 94-104.