Augmenting Reality: A new paradigm for interacting with computers

Wendy E. Mackay

Laboratoire de Recherche en Informatique
Bâtiment 490
Université de Paris-Sud
91 405 ORSAY Cedex - FRANCE
FAX : +33 (1) 69 41 65 86
mackay@lri.fr

Mackay, W. (March 1996). *Réalité Augmentée : le meilleur des deux mondes*. La Recherche, Numéro Spécial 284: *L'ordinateur au doigt et à l'œil.*

Biography

Wendy Mackay received her Ph.D. from the Massachusetts Institute of Technology in Management of Technological Innovation. She has managed both software development and research groups at Digital Equipment and most recently managed the multimedia research group at Rank Xerox EuroPARC.

Abstract

A revolution in computer interface design is changing the way we think about computers. Rather than typing on a keyboard and watching a television monitor, Augmented Reality lets people use familiar, everyday objects in ordinary ways. The difference is that these objects *also* provide a link into a computer network. Doctors can examine patients while viewing superimposed medical images; children can program their own LEGO constructions; and construction engineers can use ordinary paper engineering drawings to make live video connections to colleagues far away. Rather than immersing people in an artifically-created virtual world, the goal is to augment everyday objects in the physical world by enhancing them with a wealth of digital information and communication capabilities.

Introduction

Computers are everywhere: in the past several decades they have transformed our work and our lives. But the conversion from traditional work with physical objects to the use of computers has not been easy. Software designers may omit small, but essential, details from the original system that result in catastrophic failures. Even the act of typing is not benign: repetitive strain injuries (RSI) due to overuse of the keyboard has become the major source of workmen's compensation claims in the United States and causes over 250,000 surgeries per year.

The "paper-less" office has proven to be a myth: not only are office workers still inundated with paper, but they must now handle increasing quantities of electronic information. Worse, they are poorly equiped to bridge the gap between these two overlapping, but separate, worlds. Perhaps it is not surprising that computers have had little or no impact on white collar productivity (Landauer, 1996).

3 Photos: Wendy with Ariel,
Closeup of finger on drawing,
Projected icon on drawing.

The problem is not restricted to office work. Medicine has enjoyed major advances in imaging technology and interactive diagnostic tools. Yet most hospitals still rely on paper and pencil for medical charts next to the patient's bed. A recent study of terminally-ill patients in the U.S. concluded that, despite detailed information in their electronic medical charts, doctors were mostly unaware of their patients' wishes. Despite its importance, the electronic informaton remains disconnected from the physical world.

Augmented Reality seeks to address these kinds of problems by integrating electronic information back into the real world. AR is an umbrella term that includes various approaches with a common goal: to enable people to take advantage of their skills in interacting in the everyday world while benefiting from the power of networked computing.

At Rank Xerox EuroPARC in England, we have been exploring a variety of ways to enhance paper. Rather than simply converting information via scanning, printing, faxing and, of course, copying, the idea is to integrate paper and on-line information. In the EuroCODE ESPRIT project, we used paper engineering drawings as the central interface to a distributed multimedia network. Our goal was to create a smooth transition between the real world of bridges and paper drawings and the virtual world of the computer and its network.

Blueprints that talk to the net

A construction supervisor working on the Storebaelt bridge in Denmark (the world's longest suspension bridge) must deal with thousands of engineering design drawings as he manages and inspects the work of the bridge contractors. Any any one time, though, he only works with a few drawings. He rarely uses the computer on his desk, except occasionally for email mail and writing reports. Instead, he travels constantly from his office to the bridge pylons in the waterway and back to the pre-fabrication sites and administrative offices on shore. He prefers the convenience of paper drawings, especially since he can easily make informal notes and sketch ideas for design changes. Some of these changes appear only on these paper drawings, making them the most accurate record of the actual design of the bridge.

We were interested in giving supervisors improved access to critical information as well as to each other. Rather than designing yet another unused program for the supervisor's desk computer, we decided to use an augmented reality approach. Ariel (Mackay et al. 1993) lets the engineering drawings themselves became the interface to the computer. The supervisor places his drawing in a specified location. Ariel identifies the drawing via a bar-code and uses a tiny video camera to capture hand-written notes and the position of a light pen. Ariel uses a portable video projection panel to display computer images and interactive menus onto the drawing. The supervisor can make hypermedia annotations to his drawing with audio, video or text notes (in addition to the ones written in pencil). He can also send multimedia mail, look at annotations made by his colleagues, and even establish live video connections with people at other Ariel sites, via a media space. (See Michel Beaudouin-Lafon's article in this issue for a description of media spaces and collaborative work.)

The drawing is never "down"; it always works. But the supervisor also has all the added benefits of a distributed multimedia network, with all its capabilities for communication and information exchange.

Kinds of Augmented Reality

My colleagues Pierre Wellner (EuroPARC), Rich Gold (Xerox PARC) and I brought together the first collection of articles about Augmented Reality in a special issue of Communications of the ACM in 1993. We wanted present work that "merges electronic systems into the physical world instead of attempting to replace them." The special issue helped to launch this area of research, illustrating a variety of techniques for blending the computer and the real world.

AR applications use one or more of three basic strategies:

- Augment the user
- Augment the physical object
- Augment the environment around the object and the user

In the first category, the user wears or carries a device, usually on the head or hands, in order to obtain information about physical objects. In the second, the physical object is affected by embedding input, output or computational devices. In the third, neither the user nor the object is affected directly. Instead, independent devices provide and collect information from the surrounding environment.

Augment:	Approach	Technology	Applications
Users	Wear devices on the body	VR helmets Goggles Data gloves	Medicine Field service Presentations
Physical objects	Imbed devices within objects	Intelligent bricks Sensors GPS	Education Office facilities Positioning
Environment surrounding objects and users	Project images and record remotely	Video cameras Graphics tablets Bar code readers Scanners Video Projectors	Office work Film-making Construction Architecture

Each strategy has advantages and disadvantages and the choice depends upon the application. The key is to clearly specify how people interact with physical objects in the real world, identify the problems that additional computational support would address, and then develop the technology to meet the needs of the application.

Augment the user

Virtual Reality technology has changed dramatically from the earliest head-mounted display presented by Ivan Sutherland in 1968. Today's VR applications create computer-generated environments, ranging from sophisticated and realistic flight simulators to abstract and highly imaginative games. Some AR applications have borrowed this technology, with the goal of augmenting users' interactions with the real world rather than immersing them in artificially created ones.

Some applications are designed to let people get information by "seeing through" them. For example, researchers at the University of North Carolina, Chapel Hill, are developing See-Through Head Mounted Displays that merge computer-generated images of live data with real-world images. Ohbuchi and his colleagues let an obstetrician look simultaneously at a pregnant woman and the ultrasound image of her baby inside. A video image of the woman, taken from a camera mounted on the helmet, is merged with a computer-generated ultrasound image that corresponds to the current position of the live image. Holloway, also at UNC, uses the same approach to enable plastic surgeons to plan reconstructive surgery. The surgeon can simultaneously feel the soft tissue of a patient's face and examine a 3D reconstruction of bone data from a CAT scan that is superimposed on the patient's head.

Steve Feiner and his colleagues at Columbia University, New York, are working on a system they call KARMA (Knowledge-Based Augmented Reality for Maintenance

Assistance). Rather than making a technician refer back and forth between static images in a repair manual and the device being fixed, he can look through a half-silvered mirror and see the relevant diagrams superimposed onto the actual device (in this case, a laser printer). The system tracks the viewer and the laser printer components in real-time and calculates how best to present the information.

Both of these applications require tight coupling between the electronic images and particular views of the physical world. "Registering" the real world to enable precise matches is an active area of research.

Photo: KARMA from Steve

Augmenting the object

Another branch of AR involves augmenting physical objects themselves. In the early 1970's, Seymour Papert at MIT created a "floor turtle", actually a small robot, that could be controlled by a child with a computer language called Logo. LEGO/Logo is a direct descendant, allowing children to use Logo to control construtions made with LEGO bricks, motors and gears. Mitch Resnick and his colleagues at the Media Lab has been experimenting with *electronic bricks* that contain simple electronic devices such as sensors (light, sound, touch, proximity), logic devices (and-gates, flip-flops, timers) and action bricks (motors, lights). For example, a child could add a sound sensor to the motor drive of a toy car and use a flip-flop brick to make the car alternately start or stop at any loud noise. Children (and their teachers) have created a variety of whimsical and useful constructions, ranging from an "alarm clock bed" that detects the light in the morning and rattles a toy bed to a "smart" cage that tracks the behavior of the hamster inside.

Photo: LEGO/Logo from Mitch

Rather than augmenting a single object, Mark Weiser and his colleagues at Xerox PARC are interested in augmenting the entire surrounding environment. "Ubiquitous computing" involves a network of hundreds of computers designed to be invisible to the user. They have been actively living in a prototype environment at PARC, with hree different sized prototypes. *PARCTab*s fit in the palm of your hand and act like post-it notes. The notebook-sized version acts like a scratch pad and the wall-sized *Liveboard*, now a Xerox Product, is designed for collaborative use by several people.

A related project by Steve Elrod and others at PARC uses embedded sensors to monitor light, heat and power in the building, both to make the environment more comfortable for the occupants when they are there and to save energy when they are not. My colleagues at EuroPARC, Mik Lamming and William Newman are exploring the use of Active Badges (from Olivetti Research laboratory, England) to support collaborative activities, such as sharing documents, and personal memory, such as triggering reminders of important or upcoming events or remembering people or meetings in the recent past.

Ubiquitous computing will require changes to current networked computing models and the invention of new, low-power computer devices. Someday, post-it notes, whiteboards and identification badges will be themselves and still provide lightweight access to an electronic information network.

Photo: Wendy's active badge

Augmenting the environment

The third type of AR enhances physical environments to support various human activities. In Myron Krueger's early work on Video Place, a computer-controlled animated character would move around a wall-sized screen in response to a person's movements in front of the screen. Another early example was Richard Bolt's "Media room" at MIT's Architecture-Machine Group at (the precursor to the Media Lab). In "Put That There", a person sits in a chair, points at objects that appear on a wall-sized screen and speaks commands that move computer-generated objects to specified locations.

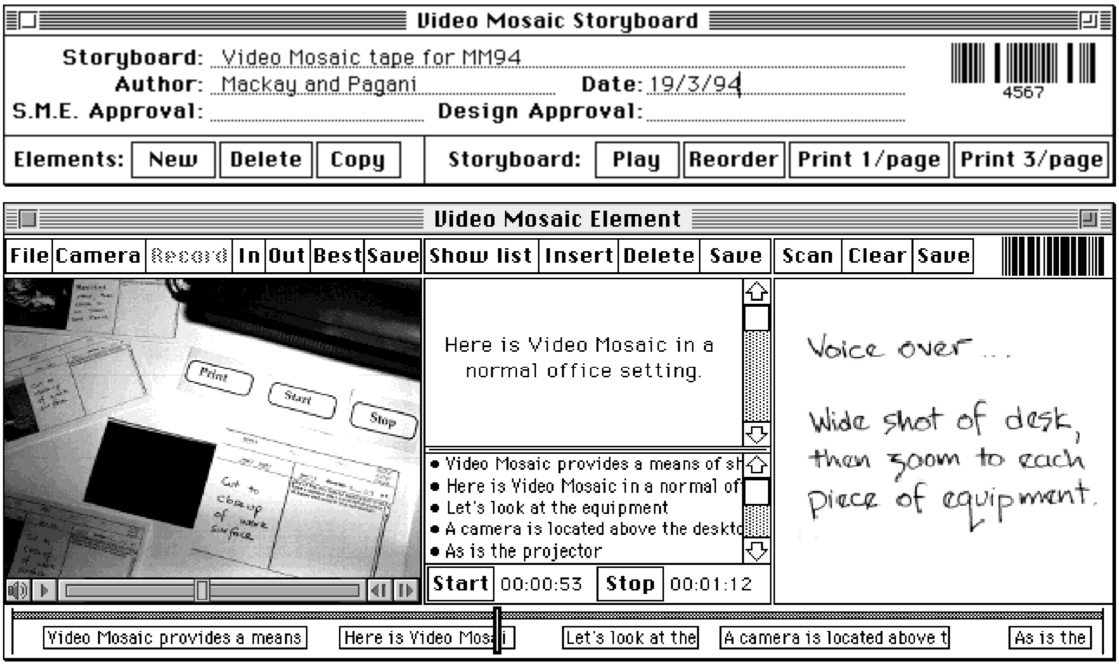
Charade, developed by Thomas Baudel at the Université de Paris-Sud, involves wearing a data glove in order to control the projection of slides and video for a formal presentation. Charade distinguishes between the natural gestures one makes when just talking or describing something and specialized gestures that can be recognized by the system, such as "show the next slide" or "start the video".

Photo: Charade from Thomas

At EuroPARC, we used similar detection and projection techniques to manage interactions with paper. Paper is essentially static: once written or printed, the marks on the paper do not change (except, of course, if changed by hand through erasure or white-out). Yet we are often interested in changing the information on paper along different dimensions. Numbers can be computed: a table summarizing financial data is more useful if we can perform "what-if" speculations on a spreadsheet, as in Wellner's original Digital Desk. Text can be transformed along several dimensions: translated into other languages as in Newman's Marcel (as in Proust), checked for spelling, or analyzed for meaning, grammar or writing style. Carter demonstrated how two-dimensional architectural sketches on paper can be rendered into three dimensions by the computer. Ariel, from the EuroCODE project, lets us add a multimedia dimension to paper by connecting it to a media space and creating sharable multimedia annotations.

Video Mosaic (Mackay and Pagani, 1993) explored another dimension: time. For example, look at the figure below. In this magazine, the image is static. But imagine what it would be like to press "play" with your finger and watch the video clip move? You could see the corresponding subtitles, make additional notes and even rearrange the order of the video or record a new clip.

2 Photos: Video Mosaic,
Closeup of finger on storyboard



We used Ariel to create a Video Mosaic prototype for film-makers, who use paper storyboards to organize and direct the flow of a movie. Video Mosaic lets the user manipulate segments of a storyboard like the one above, reordering, editing and annotating them while still being connected to a sophisticated on-line video editing system. Different media (video, audio, subtitles, notes, and control information) can be linked together or arranged into hypermedia documents that are shared over the network. Video Mosaic merges paper storyboards with an on-line editing system, taking advantage of the best elements of each.

Discussion

When should we try an AR approach? First, determine what the user already does that requires interaction with objects in the real world. Then, determine what a computerized version would be. Sometimes these will be completely separate, and sometimes they will overlap. If it is the case that the user ends up missing important characteristics by using just the non-computerized technique or the computerized technique, AR is a possibility. Then, consider what the integration would involve, not only what problems it would solve, but what problems it might create. Also, think about how the users tasks might change in the new system. Potential industries that could benefit from the AR approach to computing, include: medical, process control, manufacturing, documentation management, air traffic control, education and home uses.

Each AR application must choose the best combination of techniques for detecting information from the real world and presenting electronic information to the user. For example, a number of options are available to track the position of a user's hands. A data glove or special dots on the hands be detected with a sensing device. A video camera and image analysis algorithms can identify location. New technologies, such as the "body net" from Tom Zimmerman at the Media Lab, use a safe, low-voltage electric field next to the skin to create and sense tiny nano-amp currents. In each case, the choice depends upon the nature of the application. If the user is already wearing gloves, a data glove makes sense. If low-resolution pointing is all that is required, the combination of a camera and light pen makes for a lightweight, portable system. If very accurate position information is needed, a graphics tablet may be required.

Similarly, many possibilities are available for displaying information. A doctor can wear a head-mounted helmet and see a video image of the real world mixed with electronic information. Electronic images can be presented over one eye, as in the tiny Private Eye from Reflection Technologies. A video projector can project information directly onto the patient's body. Finally, imagine a kind of transparent, flexible screen that shows critical information when placed over the patient. The choice depends on the specific characteristics of the application: constraints of the user, the object or the environment.

AR applications also present interesting challenges to the user interface designer. For example, when superimposing virtual information onto real objects, how can the user tell what is real and what is not? How can the correspondence between the two be maintained? Actions that work invisibly in each separate world may conflict when the real and the virtual are combined. For example, if a computer menu is projected onto a piece of paper, and then another piece of paper is placed on top of the first paper, the computer project continues to be on top. In a paper world, each piece of paper obscures the ones beneath, giving a clear view of the one on top. On a computer screen, the same thing happens with overlapping windows. But when the paper and electronic information are combined, odd things occur. For Ariel, we created a blank sheet of paper that the computer could detect via a tiny infrared light on the corner. The computer would track this sheet and project pop-up menus onto it, solving the "overlay" problem.

The most innovative aspect of augmented reality is not the technology: it is the objective. Instead of replacing physical objects with a computer, we create systems that allow people to interact with the real world in natural ways and at the same time, benefit from enhanced capabilites from the computer. The future we envision is not a strange world in which we are immersed in "virtual reality". Instead, we see our familiar world, enhanced in numerous, often invisible ways.

Acknowledgments

I am grateful to the many people at MIT, EuroPARC and Xerox PARC who have provided such an inspiring environment in which to work. Thanks are also due to my colleagues on the EuroCODE project, particularly Daniele Pagani, Lone Faber and Petri Launiainen. Finally, thanks to Michel Beaudouin-Lafon for comments on the paper (and some code for Ariel).

References

Feiner, S., MacIntyre, B., and Seligmann, D. Knowledge-Based Augmented Reality. In *Communications of the ACM*, July 1993, Vol. 36, No. 7, pp. 96-97.

Mackay, W., Faber, L., Launianen, P. and Pagani, D. (1993) Design of the High Road Demonstrator,*D4.4*, 30 September 1993, EuroCODE ESPRIT project 6155.

Mackay, W. and Pagani, D. (October 1994). Video Mosaic: Laying out time in a physical space. *Proceedings of Multimedia '94* . San Francisco, CA: ACM.

Mackay, W., Velay, G., Carter, K., Ma, C., and Pagani, D. (1993) Augmenting Reality: Adding Computational Dimensions to Paper. In *Communications of the*

Bajura, M., Fuchs, H. and Ohbuchi, R. (July, 1992) Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *Computer Graphics: Proceedings of SIGGRAPH '92*, Vol. 26(2) , pp. 203-210.

Landauer, T. (1996) *The Trouble with Computers.* MIT Press: Cambridge, MA.

Pagani, D. and Mackay, W. (October 1994). Bringing media spaces into the real world.*Proceedings of ECSCW'93, the European Conference on Computer-Supported Cooperative Work*, Milano, Italy: ACM.

Wellner, P. (1992) The DigitalDesk calculator: Tactile manipulation on a desktop display. In *Proceedings of UIST'92, the ACM Symposium on User Interface Software and Technology*. (Nov 11-13, Hilton Head, SC), ACM, NY.

Resnick, M. Behavior Construction Kits. In *Communications of the ACM*, July 1993, Vol. 36, No. 7, pp. 96-97.

Weiser, M. The Computer for the 21st Century. *Scientific American*.. September, 1991, vol. 265:3.

Additional reading

Brandt, S. *The Media Lab: Inventing the Future at MIT.* (1987) Viking Press: New York.

Rheingold, H. *Virtual Reality.* (1991) Seeker and Warburg: London.

Wellner, P., Mackay, W. and Gold, R. (Editors), (July 1993) Computer Augmented Environments: Back to the Real World. Special Issue of *Communications of the ACM*, July, Vol. 36, No. 7, p 24-26.

Captions

1a. Ariel projects icons onto a paper engineering drawing to indicate text, audio and video annotations. The engineer sets up a live video connection by pointing to the designer's name and then selecting "connect" from a menu projected onto the desk.
1b. Once the paper drawing has been registered (via the barcode), Ariel can use information on the drawing, such as the name of the designer, to set up media space connections.
1c. An icon, showing that a video note is available, is projected onto the drawing.

2. KARMA shows the action of pulling out the paper tray from the laser printer and indicates the resulting change in state.

3: The LEGO/Logo behavior construction kit lets children build complex machines with gears, motors and sensors that they can control with a computer.

4: An Active Badge (developed at Olivetti Research) is used at EuroPARC to exchange documents and provide reminders of important events.

5: Charade uses a data glove to record gestures made by the speaker to control a computer-based presentation.

6a. With Video Mosaic, an editor can work with a paper version of a storyboard and use it to control an on-line version of the video being edited.

6b. The editor can point to the "play" button on the paper storyboard to watch the video play with the corresponding subtitles.