Augmented Reality: A New Way of Seeing

Computer scientists are developing systems that can enhance and enrich a user's view of the world

By Steven K. Feiner

New Eyes
AUGMENTED REALITY EYEWEAR

WebViz
Product
Review

LATEST
WEBVIZ
$49

PREVIEW

☆FEATURES☆
JAWS of TERROR
2:00, 4:00, 6:00
CRIMSON STORM
2:30, 4:30, 6:30

THEATER

LOCAL #23

SEE-THROUGH DISPLAY of a city street is shown in this artist's rendition of an augmented-reality system that could be in use by the end of this decade. The display superimposes graphics and text from wireless networks (for example, a bus schedule and a movie preview) over the appropriate objects in the user's view (the bus stop and the movie theater).

Route to
Mr. Christie's Office
ETA 12.5 Minutes

BUS STOP

Local #23
NEXT BUS
30 Seconds

Express #45
NEXT BUS
7.5 Minutes

OFFICE SPACE AVAILABLE
Contact
Megalopolis Realty
cityspace@universal.tech

Gino's Grill

Today's Specials
Sea Bass & Filet Mignon
★★★★★

OFF DUTY

8

73

Incoming Call ... 212-555-2012 (Alpert, Bell & Clemens Brokers)

What will
COMPUTER USER INTERFACES look like
10 years from now?

If we extrapolate from current systems, it's easy to imagine a proliferation of high-resolution displays, ranging from tiny handheld or wrist-worn devices to large screens built into desks, walls and floors. Such displays will doubtless become commonplace. But I and many other computer scientists believe that a fundamentally different kind of user interface known as augmented reality will have a more profound effect on the way in which we develop and interact with future computers.

Augmented reality (AR) refers to computer displays that add virtual information to a user's sensory perceptions. Most AR research focuses on "see-through" devices, usually worn on the head, that overlay graphics and text on the user's view of his or her surroundings. (Virtual information can also be in other sensory forms, such as sound or touch, but this article will concentrate on visual enhancements.) AR systems track the position and orientation of the user's head so that the overlaid material can be aligned with the user's view of the world. Through this process, known as registration, graphics software can place a three-dimensional image of a teacup, for example, on top of a real saucer and keep the virtual cup fixed in that position as the user moves about the room. AR systems employ some of the same hardware technologies used in virtual-reality research, but there's a crucial difference: whereas virtual reality brashly aims to replace the real world, augmented reality respectfully supplements it.

Consider what AR could make routinely possible. A repairperson viewing a broken piece of equipment could see instructions highlighting the parts that need to be inspected. A surgeon could get the equivalent of x-ray vision by observing live ultrasound scans of internal organs that are overlaid on the patient's body. Firefighters could see the layout of a burning building, allowing them to avoid hazards that would otherwise be invisible. Soldiers could see the positions of enemy snipers who had been spotted by unmanned reconnaissance planes. A tourist could glance down a street and see a review of each restaurant on the block. A computer gamer could battle 10-foot-tall aliens while walking to work.

Getting the right information at the right time and the right place is key in all these applications. Personal digital assistants such as the Palm and the Pocket PC can provide timely information using wireless networking and Global Positioning System (GPS) receivers that constantly track the handheld devices. But what makes augmented reality different is *how* the information is presented: not on a separate display but integrated with the user's perceptions. This kind of interface minimizes the extra mental effort that a user has to expend when switching his or her attention back and forth between real-world tasks and a computer screen. In augmented reality, the user's view of the world and the computer interface literally become one.

Although augmented reality may seem like the stuff of science fiction, researchers have been building prototype systems for more than three decades. The first was developed in the 1960s by computer graphics pioneer Ivan Sutherland and his students at Harvard University and the University of Utah. In the 1970s and 1980s a small number of researchers studied augmented reality at institutions such as the U.S. Air Force's Armstrong Laboratory, the NASA Ames Research Center and the University of North Carolina at Chapel Hill. It wasn't until the early 1990s that the term "augmented reality" was coined by scientists at Boeing who were developing an experimental AR system to help workers assemble wiring harnesses. The past decade has seen a flowering of AR research as hardware costs have fallen enough to make the necessary lab equipment affordable. Scientists have gathered at yearly AR conferences since 1998.

Despite the tremendous changes in information technology since Sutherland's groundbreaking work, the key components needed to build an AR system have remained the same: displays, trackers, and graphics computers and software. The performance of all these components has improved significantly in recent years, making it possible to design experimental systems that may soon be developed into commercial products.

Overview/*Augmented Reality*

- Augmented-reality (AR) systems add computer-generated information to a user's sensory perceptions. Whereas virtual reality aims to replace the real world, augmented reality supplements it.
- Most research focuses on "see-through" devices, usually worn on the head, that overlay graphics and text on the user's view of the world.
- Recent technological improvements may soon lead to the introduction of AR systems for surgeons, repairpeople, soldiers, tourists and computer gamers. Eventually the systems may become commonplace.

Seeing Is Believing

BY DEFINITION, the see-through displays in AR systems must be able to present a combination of virtual and real information. Although the displays can be handheld or stationary, they

PAT RAWLINGS/SAIC (*preceding pages*)

OPTICAL SEE-THROUGH DISPLAY

OPTICAL SYSTEMS superimpose computer graphics on the user's view of the world. In this current design, the prisms reflect the graphics on a liquid-crystal display into the user's line of sight yet still allow light from the surrounding world to pass through. A system of sensors and targets keeps track of the position and orientation of the user's head, ensuring that the graphics appear in the correct places. But in present-day optical systems, the graphics cannot completely obscure the objects behind them.
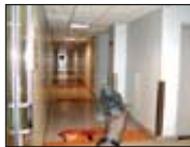
1 View of the real world from a computer gamer's perspective
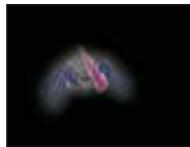
2 Graphics synthesized by the augmented-reality system
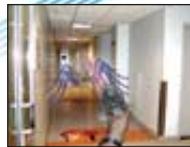
3 Image on optical display with superimposed graphics

LIQUID-CRYSTAL DISPLAY

FREE-FORM SURFACE PRISM

COMPENSATING PRISM

DISPLAY AND TRACKING SENSORS

TRACKING TARGETS

VIRTUAL SPIDER

COMPUTER
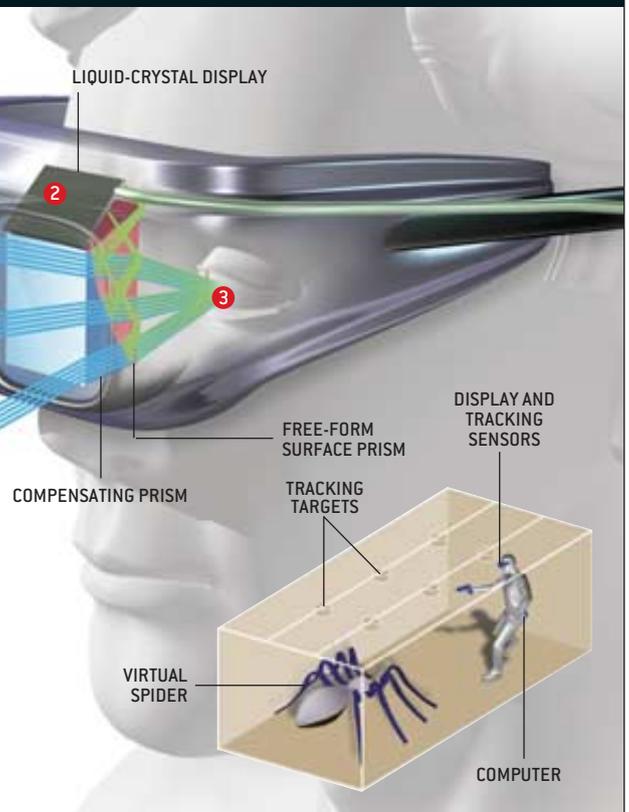
are most often worn on the head. Positioned just in front of the eye, a physically small screen can create a virtually large image. Head-worn displays are typically referred to as head-mounted displays, or HMDs for short. (I've always found it odd, however, that anyone would want to "mount" something on his or her head, so I prefer to call them head-worn displays.)

The devices fall into two categories: optical see-through and video see-through. A simple approach to optical see-through display employs a mirror beam splitter—a half-silvered mirror that both reflects and transmits light. If properly oriented in front of the user's eye, the beam splitter can reflect the image of a computer display into the user's line of sight yet still allow light from the surrounding world to pass through. Such beam splitters, which are called combiners, have long been used in "head-up" displays for fighter-jet pilots (and, more recently, for drivers of luxury cars). Lenses can be placed between the beam splitter and the computer display to focus the image so that it appears at a comfortable viewing distance. If a display and optics are provided for each eye, the view can be in stereo [*see illustration above*].

In contrast, a video see-through display uses video mixing technology, originally developed for television special effects, to combine the image from a head-worn camera with synthesized graphics [*see illustration on next page*]. The merged image is typically presented on an opaque head-worn display. With careful design, the camera can be positioned so that its optical path is close to that of the user's eye; the video image

thus approximates what the user would normally see. As with optical see-through displays, a separate system can be provided for each eye to support stereo vision.

In one method for combining images for video see-through displays, the synthesized graphics are set against a reserved background color. One by one, pixels from the video camera image are matched with the corresponding pixels from the synthesized graphics image. A pixel from the camera image appears in the display when the pixel from the graphics image contains the background color; otherwise the pixel from the graphics image is displayed. Consequently, the synthesized graphics obscure the real objects behind them. Alternatively, a separate channel of information stored with each pixel can indicate the fraction of that pixel that should be determined by the virtual information. This technique allows the display of semitransparent graphics. And if the system can determine the distances of real objects from the viewer, computer graphics algorithms can also create the illusion that the real objects are obscuring virtual objects that are farther away. (Optical see-through displays have this capability as well.)

Each of the approaches to see-through display design has its pluses and minuses. Optical see-through systems allow the user to see the real world with full resolution and field of view. But the overlaid graphics in current optical see-through systems are not opaque and therefore cannot completely obscure the physical objects behind them. As a result, the superimposed text may be hard to read against some backgrounds, and the three-

dimensional graphics may not produce a convincing illusion. Furthermore, although a user focuses physical objects depending on their distance, virtual objects are all focused in the plane of the display. This means that a virtual object that is intended to be at the same position as a physical object may have a geometrically correct projection, yet the user may not be able to view both objects in focus at the same time.

In video see-through systems, virtual objects can fully obscure physical ones and can be combined with them using a rich variety of graphical effects. There is also no discrepancy between how the eye focuses virtual and physical objects, because both are viewed on the same plane. The limitations of current video technology, however, mean that the quality of the visual experience of the real world is significantly decreased, essentially to the level of the synthesized graphics, with everything focusing at the same apparent distance. At present, a video camera and display are no match for the human eye.

The earliest see-through displays devised by Sutherland and his students were cumbersome devices containing cathode-ray tubes and bulky optics. Nowadays researchers use small liquid-crystal displays and advanced optical designs to create systems that weigh mere ounces. More improvements are forthcoming: a company called Microvision, for instance, has recently developed a device that uses low-power lasers to scan images directly on the retina [see "Eye Spy," by Phil Scott; News Scan,

SCIENTIFIC AMERICAN, September 2001]. Some prototype head-worn displays look much like eyeglasses, making them relatively inconspicuous. Another approach involves projecting graphics directly on surfaces in the user's environment.

Keeping Track

A CRUCIAL REQUIREMENT of augmented-reality systems is to correctly match the overlaid graphics with the user's view of the surrounding world. To make that spatial relation possible, the AR system must accurately track the position and orientation of the user's head and employ that information when rendering the graphics. Some AR systems may also require certain moving objects to be tracked; for example, a system that provides visual guidance for a mechanic repairing a jet engine may need to track the positions and orientations of the engine's parts during disassembly. Because the tracking devices typically monitor six parameters for each object—three spatial coordinates (x, y and z) and three orientation angles (pitch, yaw and roll)—they are often called six-degree-of-freedom trackers.

In their prototype AR systems, Sutherland and his colleagues experimented with a mechanical head tracker suspended from the ceiling. They also tried ultrasonic trackers that transmitted acoustic signals to determine the user's position. Since then, researchers have developed improved versions of these technologies, as well as electromagnetic, optical and

VIDEO SEE-THROUGH DISPLAY

VIDEO SYSTEMS mix computer graphics with camera images that approximate what the user would normally see. In this design, light from the surrounding world is captured by a wedge prism and focused on a charge-coupled device that converts the light to digital video signals. The system combines the video with computer graphics and presents the merged images on the liquid-crystal display. Video systems can produce opaque graphics but cannot yet match the resolution and range of the human eye.

LIQUID-CRYSTAL DISPLAY

FREE-FORM SURFACE PRISM

3

WEDGE PRISM

IMAGING LENS

2

COMPUTER (WORN ON HIP)

CHARGE-COUPLED DEVICE

1

DISPLAY AND TRACKING SENSORS

TRACKING TARGETS

COMPUTER

VIRTUAL SPIDER
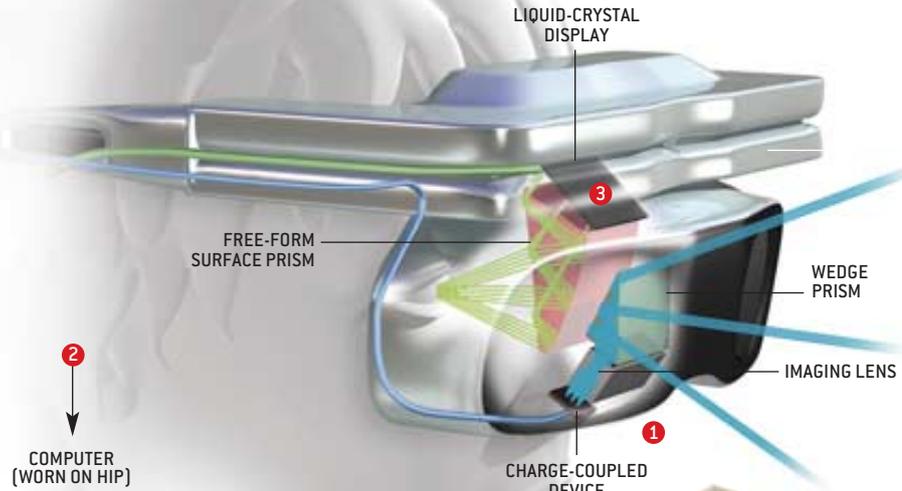
1 View of the real world captured and converted to a video image

2 Computer graphics set against a reserved background color

3 Merged image shown on the liquid-crystal display
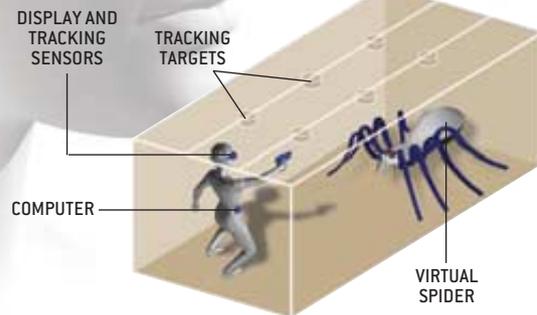
video trackers. Trackers typically have two parts: one worn by the tracked person or object and the other built into the surrounding environment, usually within the same room. In optical trackers, the targets—LEDs or reflectors, for instance—can be attached to the tracked person or object, and an array of optical sensors can be embedded in the room's ceiling. Alternatively, the tracked users can wear the sensors, and the targets can be fixed to the ceiling. By calculating the distance to each visible target, the sensors can determine the user's position and orientation.

In everyday life, people rely on several senses—including what they see, cues from their inner ears and gravity's pull on their bodies—to maintain their bearings. In a similar fashion, "hybrid trackers" draw on several sources of sensory information. For example, the wearer of an AR display can be equipped with inertial sensors (gyroscopes and accelerometers) to record changes in head orientation. Combining this information with data from the optical, video or ultrasonic devices greatly improves the accuracy of the tracking.

But what about AR systems designed for outdoor use? How can you track a person when he or she steps outside the room packed with sensors? The outdoor AR system designed by our lab at Columbia University handles orientation and position tracking separately. Head orientation is determined with a commercially available hybrid tracker that combines gyroscopes and accelerometers with a magnetometer that measures the earth's magnetic field. For position tracking, we take advantage of a high-precision version of the increasingly popular Global Positioning System receiver.

A GPS receiver determines its position by monitoring radio signals from navigation satellites. The accuracy of the inexpensive, handheld receivers that are currently available is quite coarse—the positions can be off by many meters. Users can get better results with a technique known as differential GPS. In this method, the mobile GPS receiver also monitors signals from another GPS receiver and a radio transmitter at a fixed location on the earth. This transmitter broadcasts corrections based on the difference between the stationary GPS antenna's known and computed positions. By using these signals to correct the satellite signals, differential GPS can reduce the margin of error to less than one meter. Our system is able to achieve centimeter-level accuracy by employing real-time kinematic GPS, a more sophisticated form of differential GPS that also compares the phases of the signals at the fixed and mobile receivers.

Unfortunately, GPS is not the ultimate answer to position tracking. The satellite signals are relatively weak and easily blocked by buildings or even foliage. This rules out useful tracking indoors or in places like midtown Manhattan, where rows of tall buildings block most of the sky. We found that GPS tracking works well in the central part of Columbia's campus, which has wide open spaces and relatively low buildings. GPS, however, provides far too few updates per second and is too inaccurate to support the precise overlaying of graphics on nearby objects.

Augmented-reality systems place extraordinarily high demands on the accuracy, resolution, repeatability and speed of tracking technologies. Hardware and software delays introduce a lag between the user's movement and the update of the display. As a result, virtual objects will not remain in their proper positions as the user moves about or turns his or her head. One technique for combating such errors is to equip AR systems with software that makes short-term predictions about the user's future motions by extrapolating from previous movements. And in the long run, hybrid trackers that include computer vision technologies may be able to trigger appropriate graphics overlays when the devices recognize certain objects in the user's view.

Managing Reality

THE PERFORMANCE OF GRAPHICS hardware and software has improved spectacularly in the past few years. In the 1990s our lab had to build its own computers for our outdoor AR systems because no commercially available laptop could produce the fast 3-D graphics that we wanted. In 2001, however, we were finally able to switch to a commercial laptop that had sufficiently powerful graphics chips. In our experimental mobile systems, the laptop is mounted on a backpack. The machine has the advantage of a large built-in display, which we leave open to allow bystanders to see what the overlaid graphics look like alone.

THE AUTHOR

STEVEN K. FEINER is professor of computer science at Columbia University, where he directs the Computer Graphics and User Interfaces Lab. He received a Ph.D. in computer science from Brown University in 1987. In addition to performing research on software and user interfaces for augmented reality, Feiner and his colleagues are developing systems that automate the design and layout of interactive graphics and multimedia presentations in domains ranging from medicine to government databases. The research described in this article was supported in part by the Office of Naval Research and the National Science Foundation.

Part of what makes reality *real* is its constant state of flux. AR software must constantly update the overlaid graphics as the user and visible objects move about. I use the term "environment management" to describe the process of coordinating the presentation of a large number of virtual objects on many displays for many users. Working with Simon J. Julier, Larry J. Rosenblum and others at the Naval Research Laboratory, we are developing a software architecture that addresses this problem. Suppose that we wanted to introduce our lab to a visitor by annotating what he or she sees. This would entail selecting the parts of the lab to annotate, determining the form of the annotations (for instance, labels) and calculating each label's position and size. Our lab has developed prototype software that interactively redesigns the geometry of virtual objects to maintain the desired relations among them and the real objects in the user's view. For example, the software can continually recompute a label's size and position to ensure that it is always visible and that it overlaps only the appropriate object.

It is important to note that a number of useful applications of AR require relatively little graphics power: we already see the real world without having to render it. (In contrast, virtual-reality systems must always create a 3-D setting for the user.) In a system designed for equipment repair, just one simple arrow or highlight box may be enough to show the next step in a complicated maintenance procedure. In any case, for mobile AR to become practical, computers and their power supplies must become small enough to be worn comfortably. I used to suggest that they needed to be the size of a Walkman, but a better target might be the even smaller MP3 player.

The Touring Machine and MARS

WHEREAS MANY AR DESIGNS have focused on developing better trackers and displays, our laboratory has concentrated on the design of the user interface and the software infrastructure. After experimenting with indoor AR systems in the early 1990s, we decided to build our first outdoor system in 1996 to find out how it might help a tourist exploring an unfamiliar environment. We called our initial prototype the Touring Machine (with apologies to Alan M. Turing, whose abstract Turing machine defines what computers are capable of computing). Because we wanted to minimize the constraints imposed by current technology, we combined the best components we could find to create a test bed whose capabilities are as close as we can make them to the more powerful machines we expect in the future. We avoided (as much as possible) practical concerns such as cost, size, weight and power consumption, confident that those problems will be overcome by hardware designers in the coming years. Trading off physical comfort for performance and ease of software development, we have built several generations of prototypes using external-frame backpacks. In general, we refer to these as mobile AR systems (or MARS, for short) [*see left illustration below*].

Our current system uses a Velcro-covered board and straps to hold many of the components: the laptop computer (with its

GLIMPSES OF AUGMENTED REALITY

COLUMBIA UNIVERSITY'S Computer Graphics and User Interfaces Lab built an experimental outdoor system designed to help a tourist explore the university's campus. The laptop on the user's backpack supplies the computer graphics that are superimposed on the optical see-through display. GPS receivers track the user's position.

The lab created a historical documentary that shows three-dimensional images of the Bloomingdale Asylum, the prior occupant of Columbia's campus, at its original location.

A user viewing the documentary can get additional information from a handheld display, which provides an interactive timeline of the Bloomingdale Asylum's history.

MEDICAL APPLICATION was built by researchers at the University of Central Florida. The system overlaid a model of a knee joint on the view of a woman's leg. The researchers tracked the leg's position using infrared LEDs. As the woman bent her knee, the graphics showed how the bones would move.

The OVERLAID INFORMATION will become part of what we expect to see AT WORK AND AT PLAY.
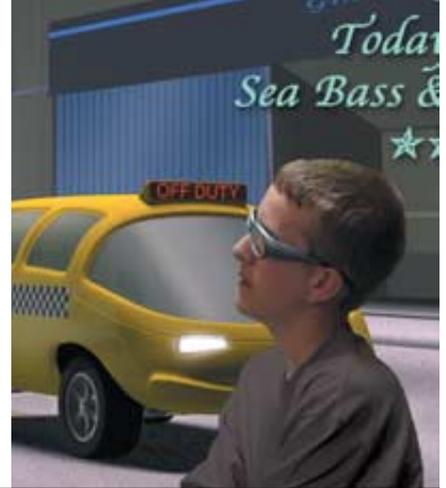
3-D graphics chip set and IEEE 802.11b wireless network card), trackers (a real-time kinematic GPS receiver, a GPS corrections receiver and the interface box for the hybrid orientation tracker), power (batteries and a regulated power supply), and interface boxes for the head-worn display and interaction devices. The total weight is about 11 kilograms (25 pounds). Antennas for the GPS receiver and the GPS corrections receiver are mounted at the top of the backpack frame, and the user wears the head-worn see-through display and its attached orientation tracker sensor. Our MARS prototypes allow users to interact with the display—to scroll, say, through a menu of choices superimposed on the user's view—by manipulating a wireless trackball or touch pad.

From the very beginning, our system has also included a handheld display (with stylus input) to complement the head-worn see-through display. This hybrid user interface offers the benefits of both kinds of interaction: the user can see 3-D graphics on the see-through display and, at the same time, access additional information on the handheld display.

In collaboration with my colleague John Pavlik and his students in Columbia's Graduate School of Journalism, we have explored how our MARS prototypes can embed "situated documentaries" in the surrounding environment. These documentaries narrate historical events that took place in the user's immediate area by overlaying 3-D graphics and sound on what the user sees and hears. Standing at Columbia's sundial and looking through the head-worn display, the user sees virtual flags planted around the campus, each of which represents several sections of the story linked to that flag's location. When the user selects a flag and then chooses one of the sections, it is presented on both the head-worn and the handheld displays.

One of our situated documentaries tells the story of the student demonstrations at Columbia in 1968. If the user chooses one of the virtual flags, the head-worn display presents a narrated set of still images, while the handheld display shows video snippets and provides in-depth information about specific participants and incidents. In our documentary on the prior occupant of Columbia's current campus, the Bloomingdale Asylum, 3-D models of the asylum's buildings (long since demolished) are overlaid at their original locations on the see-through display. Meanwhile the handheld display presents an interactive annotated timeline of the asylum's history. As the user chooses different dates on the timeline, the images of the buildings that existed at those dates fade in and out on the see-through display.

The Killer App?

AS RESEARCHERS CONTINUE to improve the tracking, display and mobile processing components of AR systems, the seamless integration of virtual and sensory information may become not merely possible but commonplace. Some observers have suggested that one of the many potential applications of augmented reality (computer gaming, equipment maintenance, medical imagery and so on) will emerge as the "killer app"—a use so compelling that it would result in mass adoption of the technology. Although specific applications may well be a driving force when commercial AR systems initially become available, I believe that the systems will ultimately become much like telephones and PCs. These familiar devices have no single driving application but rather a host of everyday uses.

The notion of computers being inextricably and transparently incorporated into our daily lives is what computer scientist Mark Weiser termed "ubiquitous computing" more than a decade ago [see "The Computer for the 21st Century," by Mark Weiser; SCIENTIFIC AMERICAN, September 1991]. In a similar way, I believe the overlaid information of AR systems will become part of what we expect to see at work and at play: labels and directions when we don't want to get lost, reminders when we don't want to forget and, perhaps, a favorite cartoon character popping out from the bushes to tell a joke when we want to be amused. When computer user interfaces are potentially everywhere we look, this pervasive mixture of reality and virtuality may become the primary medium for a new generation of artists, designers and storytellers who will craft the future. [SA]

MORE TO EXPLORE

A Survey of Augmented Reality. Ronald T. Azuma in *Presence: Teleoperators and Virtual Environments,* Vol. 6, No. 4, pages 355–385; August 1997. Available at **www.cs.unc.edu/~azuma/ARpresence.pdf**

Recent Advances in Augmented Reality. Ronald T. Azuma, Yohan Baillot, Reinhold Behringer, Steven K. Feiner, Simon Julier and Blair MacIntyre in *IEEE Computer Graphics and Applications,* Vol. 21, No. 6, pages 34–47; November/December 2001. Available at **www.cs.unc.edu/~azuma/cga2001.pdf**

Columbia University's Computer Graphics and User Interfaces Lab is at **www.cs.columbia.edu/graphics/**

A list of relevant publications can be found at **www.cs.columbia.edu/graphics/publications/publications.html**

AR research sites and conferences are listed at **www.augmented-reality.org**

Information on medical applications of augmented reality is at **www.cs.unc.edu/~us/**