Developing a Generic Augmented-Reality Interface

The Tiles system seamlessly blends virtual and physical objects to create a work space that combines the power and flexibility of computing environments with the comfort and familiarity of the traditional workplace.

Ivan Poupyrev
Sony CSL

Desney S. Tan
Carnegie Mellon University

Mark Billinghurst
University of Washington

Hirokazu Kato
Hiroshima City University

Holger Regenbrecht
DaimlerChrysler AG

Nobuji Tetsutani
ATR MIS Labs

An augmented-reality (AR) interface dynamically superimposes interactive computer graphics images onto objects in the real world.[1] Although the technology has come a long way from rendering simple wireframes in the 1960s,[2] AR interface design and interaction space development remain largely unexplored. Researchers and developers have made great advances in display and tracking technologies, but interaction with AR environments has been largely limited to passive viewing or simple browsing of virtual information registered to the real world.[3]

To overcome these limitations, we seek to design an AR interface that provides users with interactivity so rich it would merge the physical space in which we live and work with the virtual space in which we store and interact with digital information. In this single *augmented space,* computer-generated entities would become first-class citizens of the physical environment. We would use these entities just as we use physical objects, selecting and manipulating them with our hands instead of with a special-purpose device such as a mouse or joystick. Interaction would then be intuitive and seamless because we would use the same tools to work with digital and real objects.

Tiles is an AR interface that moves one step closer to this vision. It allows effective spatial composition, layout, and arrangement of digital objects in the physical environment. The system facilitates seamless two-handed, three-dimensional interaction with both virtual and physical objects, without requiring any special-purpose input devices. Unlike some popular AR interfaces that constrain virtual objects to a 2D tabletop surface,[4] Tiles allows full 3D spatial interaction with virtual objects anywhere in the physical workspace. A tangible interface, it lets users work with combinations of digital and conventional tools so that, for example, they can use sticky notes to annotate both physical and virtual objects. Because our approach combines tangible interaction with a 3D AR display, we refer to it as *tangible augmented reality*.

We do not suggest that tangible AR interfaces are perfect for all conceivable applications. Therefore, although our interface techniques can be applied broadly, we ground our design in a real-world application: the rapid prototyping and collaborative evaluation of aircraft instrument panels.

INTERFACE DESIGN SPACE DICHOTOMY

As the "Short History of Augmented Reality Interface Design" sidebar indicates, the AR interface design space can be divided along two orthogonal approaches.

A *3D AR interface* provides spatially seamless augmented space where the user, wearing a head-mounted display, can effectively interact with both

Short History of Augmented Reality Interface Design

In 1965, Ivan Sutherland built the first see-through head-mounted display and used it to show a simple wireframe cube overlaid on the real world, creating the first augmented-reality interface. Developers of early AR interfaces, who followed Sutherland, similarly designed them mostly to view 3D virtual models in real-world contexts for applications such as medicine,[1] machine maintenance,[2] or personal information systems.[3] Although these interfaces provided an intuitive method for viewing 3D data, they typically offered little support for high-level interaction, such as creating or modifying augmented-reality content.

3D AR Interfaces

Recently, researchers have begun to address this deficiency. Kiyoshi Kiyokawa[4] uses a magnetic tracker, while the Studierstube[5] project uses a tracked pen and tablet to select and modify augmented-reality objects. In a 3D AR interface, a head-mounted display lets the user interact with virtual content through different input devices. More traditional input techniques, such as handheld mice[6] and intelligent agents,[7] have also been investigated.

However, in all these cases the user must work with special-purpose input devices, separate from tools used to interact with the real world. This limitation effectively results in two different interfaces—one for the physical workspace and another for the virtual one. Consequently, interaction discontinuities, or seams, disrupt the natural workflow, forcing the user to switch between virtual and real operation modes.

Tangible Interfaces

For more than a decade, researchers have been investigating an alternative approach: computer interfaces based on physical objects. The Digital Desk project[8] used computer-vision techniques to track the position of paper documents and the user's hands on an augmented table. The user could seamlessly arrange and annotate both real paper and virtual documents using the same physical tools—a pen and a finger. Graspable[9] and tangible user interfaces further explore the connection between virtual objects and the physical properties of input devices, using simple wooden blocks to manipulate virtual objects projected on a table's surface, for example.

Most importantly, tangible interfaces allow for seamless interaction, because a single physical device represents each interface function or object. Therefore, users can access interface functions and use traditional tools in the same manner—through manipulation of physical objects.

Information display in tangible interfaces can be a challenge, however. Changing an object's physical properties dynamically is difficult, and these interfaces usually project virtual objects onto 2D surfaces.[10] The users, therefore, cannot pick virtual objects off the augmented surface and manipulate them in 3D space as they would a real object—the system localizes the interaction to an augmented surface and cannot extend beyond it. Given that the tangible interface user cannot seamlessly interact with virtual objects anywhere in space—by, for example, moving a virtual object between augmented and nonaugmented workspaces—the tangible interface introduces a spatial seam into the interaction.

References

1. M. Bajura, H. Fuchs, and R. Ohbuchi, "Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient," *Proc. Siggraph 92*, ACM Press, New York, 1992, pp. 203-210.
2. S. Feiner, B. MacIntyre, and D. Seligmann, "Knowledge-Based Augmented Reality," *Comm. ACM*, vol. 36, no. 7, 1993, pp. 53-62.
3. J. Rekimoto and K. Nagao, "The World through the Computer: Computer Augmented Interaction with Real World Environments," *Proc. UIST 95*, ACM Press, New York, pp. 29-36.
4. K. Kiyokawa, H. Takemura, and N. Yokoya, "Seamless Design for 3D Object Creation," *IEEE MultiMedia*, vol. 7, no. 1, 2000, pp. 22-33.
5. D. Schmalstieg, A. Fuhrmann, and G. Hesina, "Bridging Multiple User Interface Dimensions with Augmented Reality Systems," *Proc. ISAR 2000*, IEEE CS Press, Los Alamitos, Calif., 2000, pp. 20-29.
6. T. Hollerer et al., "Exploring MARS: Developing Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System," *Computers & Graphics*, vol. 23, 1999, pp. 779-785.
7. M. Anabuki et al., "Welbo: An Embodied Conversational Agent Living in Mixed Reality Spaces, *Proc. CHI 2000*, ACM Press, New York, 2000, pp. 10-11.
8. P. Wellner, "Interaction with Paper on the Digital Desk, *Comm. ACM*, vol. 36, no. 7, 1993, pp. 87-96.
9. G. Fitzmaurice and W. Buxton, "An Empirical Evaluation of Graspable User Interfaces: Towards Specialized, Space-Multiplexed Input, *Proc. CHI 97*, ACM Press, New York, 1997, pp. 43-50.
10. J. Rekimoto and M. Saitoh, "Augmented Surfaces: A Spatially Continuous Work Space for Hybrid Computing Environments," *Proc. CHI 99*, ACM Press, New York, 1999, pp. 378-385.

2D and 3D virtual objects anywhere in the working environment. To interact with virtual content, however, the user must rely on special-purpose input devices that would not normally be used in real-world interactions. Thus, *interaction discontinuities* break the workflow, forcing the user to switch between virtual and real interaction modes.

Tangible interfaces, on the other hand, use multiple physical objects tracked on the augmented surface as physical handles or containers for interacting with virtual objects projected onto the surface. Tangible interfaces do not require any special-purpose input devices, and thus they provide an intuitive and seamless interaction with digital *and* physical objects. However, *spatial discontinuities* do break the interaction flow because the interface is localized on augmented surfaces and cannot be extended beyond them. Further, tangible interfaces offer limited support for interacting with 3D virtual objects.

Tracking and Registration in Tiles

An augmented-reality system's fundamental elements include techniques for tracking user position and viewpoint direction, registering virtual objects relative to the physical environment, then rendering and presenting them to the user. We implemented the Tiles system with the ARToolKit software, an open source library for developing computer-vision-based AR applications.

To create the physical tiles, we mark paper cards measuring 15 cm × 15 cm with simple square patterns consisting of a thick black border and unique symbols in the middle. We can use any symbol for identification as long as each symbol is asymmetrical enough to distinguish between the square border's four possible orientations.

The user wears a Sony Glasstron PLMS700 headset. Lightweight and comfortable, the headset provides an 800 × 600-pixel VGA image. A miniature NTSC Toshiba camera with a wide-angle lens attaches to the headset. The system captures the camera's video stream at 640 × 240-pixel resolution to avoid interlacing problems. The image is then scaled back to 640 × 480 pixels using a line-doubling technique.

By tracking rectangular markers of known size, the system can find the relative camera position and orientation in real time and can then correctly render virtual objects on the physical cards, as Figure A shows. Although the wide-angle lens distorts the video image, our tracking techniques are robust enough to correctly track patterns without losing performance. We implemented our current system on a 700-MHz Pentium III PC running Linux, which allows updates at 30 frames per second. For more information on the ARToolKit library, see http://www.hitl.washington.edu/artoolkit/.

① A video frame is captured and converted into a binary image. Marker frames are identified.

② Marker outlines and their transformation T relative to the camera are calculated. Symbols inside of the marker are extracted, normalized, and matched with templates in memory.

③ The appropriate virtual objects are rotated and positioned according to transformation T, rendered into the video frame, and presented to the user in the HMD.

Figure A. The three-step process of mapping virtual objects onto physical tiles so that the user can view them with a head-mounted display (HMD).
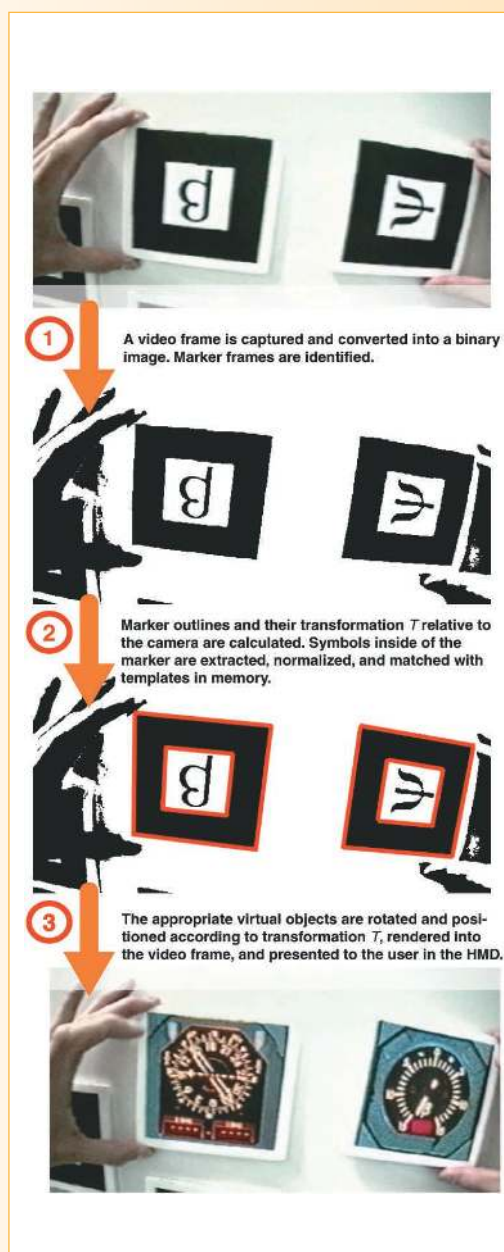
We believe that these opposing approaches also complement each other. In Tiles, we design interfaces that bridge 3D AR and tangible interactions and thus produce a spatially and interactively seamless augmented workspace.

DESIGNING TILES

In Tiles, the user wears a lightweight head-mounted display with a small camera attached, both of which are connected to a computer. The video capture subsystem uses the camera's output to overlay virtual images onto the video in real time as described in the "Tracking and Registration in Tiles" sidebar. The system then shows the resulting augmented view of the real world on the head-mounted display so that the user sees virtual objects embedded in the physical workspace, as Figures 1 and 2 show. Computer-vision tracking techniques determine the 3D position and orientation of marked real objects so that virtual models can be exactly overlaid on them.[5] By manipulating these objects, the user can control the virtual objects associated with them without using any additional input devices.

Design requirements

Although we developed the Tiles system specifically for rapid prototyping and evaluation of air-

craft instrument panels, we believe that this task's requirements have broad application to many common AR interface designs.

Aircraft instrument panel design requires the collaborative efforts of engineers, human factors specialists, electronics designers, and pilots. Designers and engineers always look for new technologies that can reduce the cost of developing the instrument panels without compromising design quality. Ideally, they would like to evaluate instrument prototypes relative to existing instrument panels, without physically rebuilding them. This inherently collaborative design activity involves heavy use of existing physical plans, documents, and tools.

Using observations of current instrument panel design, DASA/EADS Airbus and DaimlerChrysler engineers produced a set of interface requirements.[6] They envisioned an AR interface that would let developers collaboratively outline and lay out a set of virtual aircraft instruments on a board that simulates an airplane cockpit. Designers could easily add and remove virtual instruments from the board using an instrument catalog. After placing the instruments on the board, they could evaluate and rearrange the instruments' position. The interface should also allow use of conventional tools—such as whiteboard markers—with physical schemes and documents so that participants could document problems and solutions.

Basic concepts

We designed the Tiles interface around a set of simple interface principles that produce a generic and consistent AR interface.

A tile is a small cardboard card with a marker. It serves as a physical handle for interacting with virtual objects. Conceptually similar to icons in a graphical user interface, a tile acts as a tangible interface control. Users physically manipulate the corresponding tiles to interact with virtual objects just as they would with real objects. The resulting seamless interface requires no additional input devices to interact with virtual objects.

Although the tiles resemble physical icons, or *phicons*,[3] they exhibit important differences. Phicons propose a close coupling between physical and virtual properties so that their shape and appearance mirror their corresponding virtual object or functionality. The Tiles interface decouples the physical properties of interface controls from the data to create universal and generic data containers that can hold any digital data or no data at all: The tile is a blank object that has no associ-
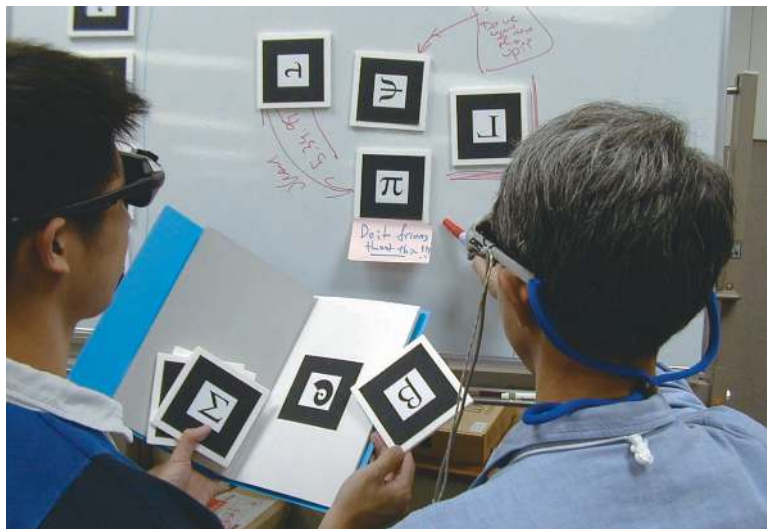
Figure 1. Users collaboratively arrange tangible data containers—tiles—on the whiteboard and use traditional tools to add notes and annotations.
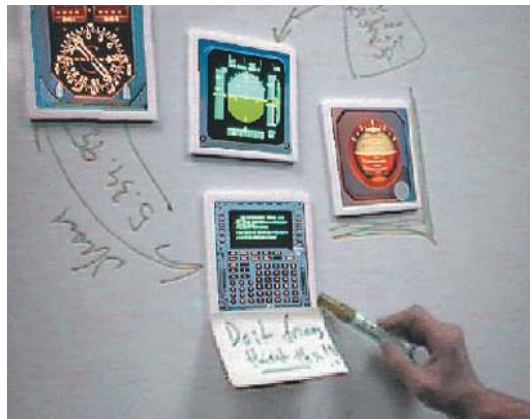

Figure 2. The user, wearing a lightweight head-mounted display with an attached camera, can see real objects and the virtual images registered on tiles.

ated data until the user creates an association at runtime. Hence, techniques for performing basic operations, such as attaching data to tiles, remain the same for all tiles, resulting in a consistent and streamlined user interface.

We use two separate tile classes. *Operation tiles* define the interface's functionality and provide tools to perform operations on data tiles. The Tiles system always attaches animated 3D icons to operation tiles so that the user can identify them. *Data tiles* act as generic data containers. The user can associate or disassociate any virtual objects with data tiles at any time using operator tiles. The user physically manipulates tiles to invoke operations between them, such as controlling the proximity between tiles, their distance from the user, or their orientation. The augmented working space is spatially seamless—aside from cable length, the user has no restriction on interacting with the tiles.

Figure 3. Operation tiles. (a) The printed design of the physical tiles that represent the delete, copy, and help operations, and (b) the virtual icons that represent the same three operations in the augmented-reality interface.
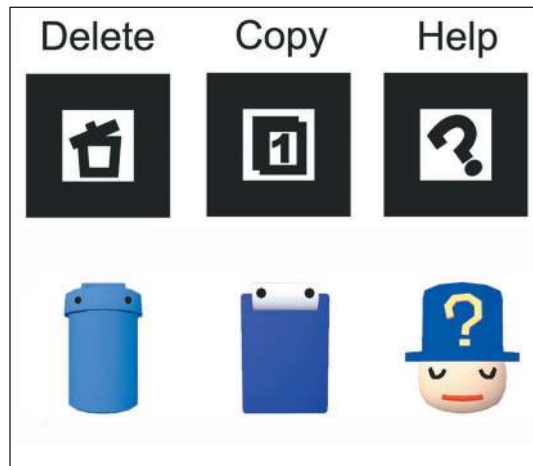
Figure 4. The user cleans a data tile by moving the trash can operator tile next to it.

Figure 5. Copying data from the clipboard to an empty data tile. The user moves the tile close to the virtual clipboard and, after a one-second delay, the virtual instrument slides smoothly onto the data tile.
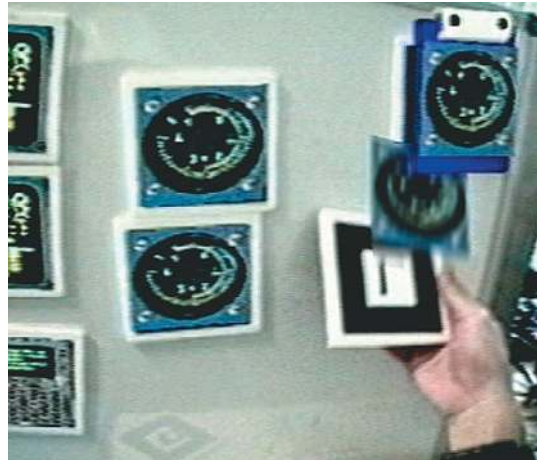
Tiles interface

The Tiles interface consists of a metal whiteboard, a book, and two stacks of magnetic tiles that each measure approximately 15 cm × 15 cm. Sitting in front of the whiteboard, the user wears a lightweight, high-resolution Sony Glasstron head-mounted display with a video camera attached, as Figure 1 shows.

The various tangible interface elements serve different purposes. The whiteboard provides the workspace where users can lay out virtual aircraft instruments. The book serves as a catalog, or menu object, that shows a different virtual instrument model on each page. One tile stack stores blank data containers, which show no content until users copy virtual objects onto them. The remaining tiles function as operator tiles that perform basic operations on the data tiles. Each operation has a unique tile associated with it. Currently supported operations include deletion, copying, and a help function. Each operation tile bears a different 3D virtual icon that shows its function and differentiates it from the data tiles, as Figure 3 shows.

Invoking operations. All tiles can be freely manipulated in space and arranged on the whiteboard. The user simply picks up a tile, examines its contents, and places it on the whiteboard. The user can invoke operations between two tiles by moving them next to each other.

For example, to copy an instrument to a data tile, the user first finds the desired virtual instrument in the menu book, then places any empty data tile next to it. After a one-second delay, a copy of the instrument smoothly slides from the menu page to the tile and can be arranged on the whiteboard. Similarly, the user can remove data from a tile by moving the trash can tile close to the data tiles, which removes the instrument from it, as Figure 4 shows.

Using the same technique, we can implement copy and paste operations using a copy operation tile. The user can copy data from any data tile to the clipboard, then from the clipboard to any number of empty data tiles by moving empty tiles next to the virtual clipboard that has data in it, as Figure 5 shows. The clipboard's current contents can always be seen on the virtual clipboard icon. Users can display as many clipboards as they need—the current implementation has two independent clipboards.

Getting help. The Tiles interface provides a help system that lets the user request assistance without shifting focus from the main task. This approach is more suitable for AR interfaces than traditional desktop help systems, which either distract users with a constant barrage of help messages or interrupt their work by making them search explicitly for help.

The Tiles system provides two help techniques, shown in Figure 6. With Tangible Bubble Help, simply placing the help tile beside the tile the user requires help with brings up a text bubble next to the help icon, as Figure 6a shows. In some cases, however, users only need short reminders, or tips, about a particular tile's functionality. Alternatively, the Tangible ToolTips technique triggers the display of a short text description associated with a tile when the user moves the tile within arm's reach and tilts it more than 30 degrees away from the body, as Figure 6b shows.

Mixing physical and virtual tools. The Tiles interface lets users seamlessly combine conventional physical and virtual tools. For example, the user can physically annotate a virtual aircraft instrument using a standard whiteboard pen or sticky note, as Figure 7 shows.

Multiple users. We designed Tiles with collaboration in mind. Thus, the system lets several users interact in the same augmented workspace. All users can be equipped with head-mounted displays and can directly interact with virtual objects. Alternatively, users who do not wear head-mounted displays can collaborate with immersed users via an additional monitor that presents the augmented-reality view. Because all interface components consist of simple physical objects, both the nonimmersed and immersed user can perform the same authoring tasks.

TILES IN OTHER APPLICATIONS

Our initial user observations showed that developers of tangible AR interfaces must focus on both the interface's physical design and the virtual icons' computer graphics design. Physical component designs can convey additional interface semantics. For example, the physical cards can be designed to snap together like pieces in a jigsaw puzzle, resulting in different functionality profiles depending on their physical configuration.

The interface model and interaction techniques introduced in Tiles can be extended easily to other applications that require AR interfaces. Object modification techniques, for example, can be introduced into Tiles by developing additional operator cards that would let the user dynamically modify objects through scaling, color changes, employing hand gestures, and so on. The interaction techniques we present here will remain applicable, however.

We found that the tight coupling of 3D input and display in a single interface component—the tile—lets users perform complex functions through essentially simple spatial manipulation and physical arrangements of these tangible interface components. Thus, Tiles provides an application-independent interface that could lead to the development of generic AR interface models based on tangible augmented-reality concepts.

Although developing additional interaction techniques would let users apply Tiles to many different application scenarios, in AR environments the user can already switch easily between the AR workspace and a traditional environment. Some tools and techniques better suit augmented
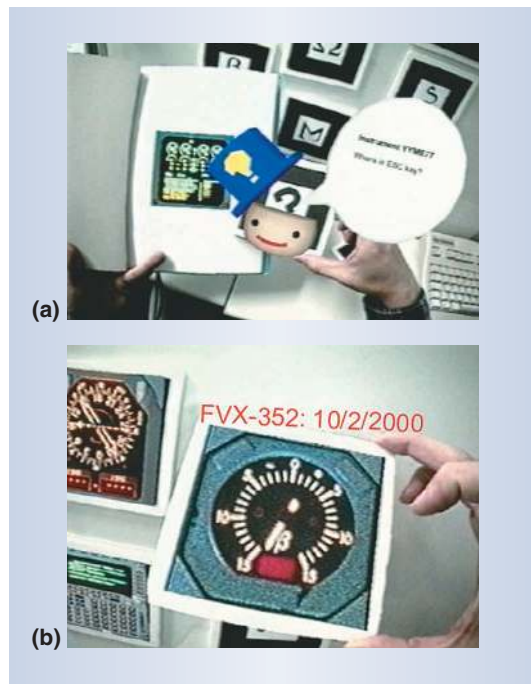
Figure 6. Tangible help in Tiles. (a) To access Tangible Bubble Help, users place the help tile next to the tile they need assistance with, which causes textual annotations to appear within a bubble next to the tile. (b) For less detailed explanations, Tangible ToolTips displays an associated short text description when the user moves the tile closer and tilts it.
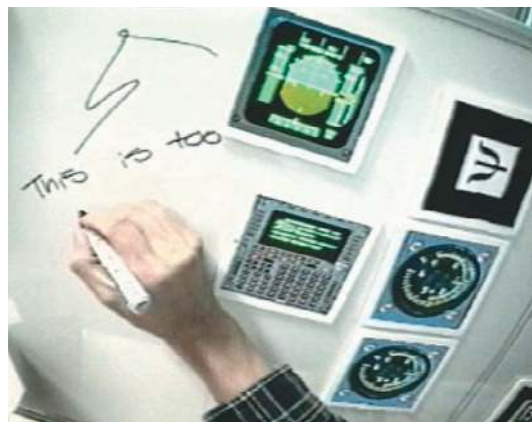
Figure 7. Physically annotating virtual objects in Tiles. Because the printed tiles offer a physical anchor for virtual objects, users can make notes adjacent to them using marking pens and sticky notes.

reality, while others work best in traditional form. Therefore, we believe that development of AR interfaces should not focus on bringing every possible interaction tool and technique into the AR workspace. Instead, it should focus on balancing and distributing features between the AR interface and other interactive media so that they all can be used within a single seamless augmented workspace.

AR interfaces also offer an ad hoc, highly reconfigurable environment. Unlike traditional GUI and 3D VR interfaces, in which the designer determines most of the interface layout in advance, in Tiles users can freely place interface elements anywhere they want: on tables or whiteboards, in boxes and folders, arranged in stacks, or grouped together. The interface configuration and layout of its elements emerges spontaneously as the results of users' work activity, and evolves together with it. How the interface components should be designed for such environments, and whether these systems should be aware of the dynamic changes in their configuration, are important research questions. ■

References

1. R. Azuma, "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, 1997, pp. 355-385.
2. I. Sutherland, "The Ultimate Display," *Proc. Int'l Federation of Information Processing*, Spartan Books, Washington, D.C., 1965, pp. 506-508.
3. H. Ishii and B. Ullmer, "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms," *Proc. Computer-Human Interaction* (CHI 97), ACM Press, New York, 1997, pp. 234-241.
4. B. Ullmer and H. Ishii, "The MetaDesk: Models and Prototypes for Tangible User Interfaces," *Proc. User Interface Software and Technology* (UIST 97), ACM Press, New York, 1997, pp. 223-232.
5. H. Kato and M. Billinghurst, "Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System," *Proc. 2nd Int'l Workshop Augmented Reality*, IEEE CS Press, Los Alamitos, Calif., 1999, pp. 85-94.
6. I. Poupyrev et al., "Tiles: A Mixed Reality Authoring Interface," *Proc. Interact 2001*, IOS Press, Netherlands, 2001, pp. 334-341.

Ivan Poupyrev is an associate researcher in Sony Computer Science Labs' Interaction Lab, Tokyo, where he investigates the interaction implications of computer-augmented environments. Poupyrev received a PhD in computer science from Hiroshima University. Contact him at poup@csl. sony.co.jp.
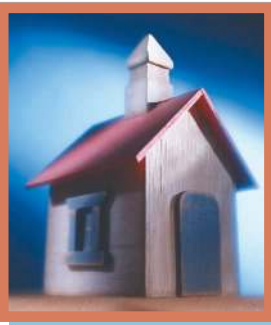
Desney S. Tan is a PhD candidate at Carnegie Mellon University. His research interests include designing human-computer interfaces that augment human cognition, specifically to leverage existing psychology principles on human memory and spatial cognition in exploring multimodal, multidisplay information systems. He received a BS in computer engineering from the University of Notre Dame. Contact him at desney@cs.cmu.edu or see http://www.cs.cmu.edu/~desney.

Mark Billinghurst is a researcher in the Human Interface Technology Laboratory at the University of Washington. His research interests include augmented and virtual reality, conversational computer interfaces, and wearable computers, with his most recent work centering around using augmented reality to enhance face-to-face and remote conferencing. Billinghurst received a PhD in electrical engineering from the University of Washington. Contact him at grof@hitl.washington.edu.

Hirokazu Kato is an associate professor at Faculty of Information Sciences, Hiroshima City University. His research interests include augmented reality, computer vision, pattern recognition, and human-computer interaction. Kato received a PhD in engineering from Osaka University. Contact him at kato@sys.im.hiroshima-cu.ac.jp.

Holger Regenbrecht is a scientist at the Daimler-Chrysler Research Center in Ulm, Germany. His research interests include interfaces for virtual and augmented environments, virtual-reality-aided design, perception of virtual reality, and AR/VR in the automotive and aerospace industry. Regenbrecht received a doctoral degree from the Bauhaus University Weimar, Germany. Contact him at regenbre@ igroup.org.

Nobuji Tetsutani is the head of Department 3 in ATR Media Information Science Laboratories. His research interests include application development for high-speed networks. Tetsutani received a PhD in electrical engineering from Hokkaido University, Hokkaido, Japan. Contact him at tetsutani@ atr.co.jp.