Editors: Michael Macedonia and Lawrence Rosenblum

Interface with Angels: The Future of VR and AR Interfaces

Wolfgang Broll and Leonie Schäfer

Fraunhofer Institute for Applied Information Technology, Germany

Tobias Höllerer

Columbia University

Doug Bowman

Virginia Tech

What can virtual reality and augmented reality technologies do for us? Can they help us make better use of computers? Will they ever improve the quality of our lives? What kind of immersive interfaces could reshape the way we communicate with computers? How far along are we on that road, and what remains to be done?

At the IEEE Virtual Reality 2001 conference, we organized a workshop on the future of VR and AR interfaces that addressed these questions. Its main areas of focus were conversational user interfaces, natural interaction, 3D interaction techniques, haptics, augmented and mixed reality, and mobile and wearable interfaces. Here, we want to mix facts with fancy in looking at tomorrow's interface technology. Today's research paves the road for technological possibilities that go beyond the caves and cables, glasses and gloves, and pixels and polygons that still dominate research in VR and AR—possibilities that can assist you in virtually all situations of life.

Guardian angels

Watching an old 2D movie on your projection wall one night, you smile about how the main characters are stumbling through their day. You remember this all too well: Rushing to the office, always late, forgetting to switch off the coffee machine when leaving the house. Being stuck in traffic and bumping your car. Not knowing where a meeting will take place. Not having the right information available for a presentation. Not remembering names or places. Missing important dates such as your best friend's birthday. And most of all, lacking the time and organizational skills to make all your arrangements.

Luckily, times have changed for you. When you are about to leave the next morning, a person appears near the door, "Good morning. You modified some files on your notebook last night. Do you want me to transfer them to your office PC, or will you take the notebook with you?" The conversational interface of your personal guardian angel appears as a holographic projection, moving around. "And by the way, your keys are over here," she exclaims. After you leave, the smart-house environment switches off the coffee machine—no need to consult you for such an obvious task.

When you enter the elevator in your office building, the angel again appears in front of you. This time an inconspicuous retinal display embedded in your eyewear paints her image directly onto your eyeballs. This way the projection is invisible for anyone else. "The meeting is in the conference room on the eleventh floor. Your boss is already there, and your visitors will arrive in about five minutes." Nobody else hears these words, because they come through your wireless microearphones. Knowing your preference for concise visual presentations, the angel software augments your environment with additional information such as the names and affiliations of the other participants. You can position these displays in 3D space using eye or hand movements. Because this doesn't require a conversational interface, the angel stays in the background.

Interacting with a conversational interface is still a little bit cumbersome in situations where you can't use natural speech and gestures, although the eye-blinking interface is working reasonably well for many tasks. Still, you're really looking forward to finally testing out the new thought interface. The big controversy about the safety of brain-wave-activated interfaces has given this technology a lot of publicity.

After your meeting, the angel appears in your office to remind you about tonight's invitation to an old friend's birthday party. She presents you with a selection of true-to-life renderings of possible gifts for direct ordering. You select a gift simply by pointing at it.

Coming from work, you would usually go jogging or stop by the fitness center. Today, however, there isn't much time left before the party. Thus, you decide to do only a short exercise program at home. The display system replaces your drab workout room with a relaxing and stimulating beach scene. Your guardian angel knows your usual workout program and supports you by demonstrating the exercises and giving you feedback. Later, on your way to the party you don't remember where to go, but your angel software has already supplied your car navigation system with all the relevant information. Unfortunately, fully automatic driving still hasn't been approved yet.

The personal user interface gives you pleasant and convenient access to your private information and the electronic services surrounding you, wherever you go. Indeed, the personal virtual interface is so successful that all major providers of electronic services go to great lengths to support it, finally working out a standard for

0272-1716/01/$10.00 © 2001 IEEE

the electronic ether. You sometimes don't know how you managed life without your personal guardian angel.

Back to reality

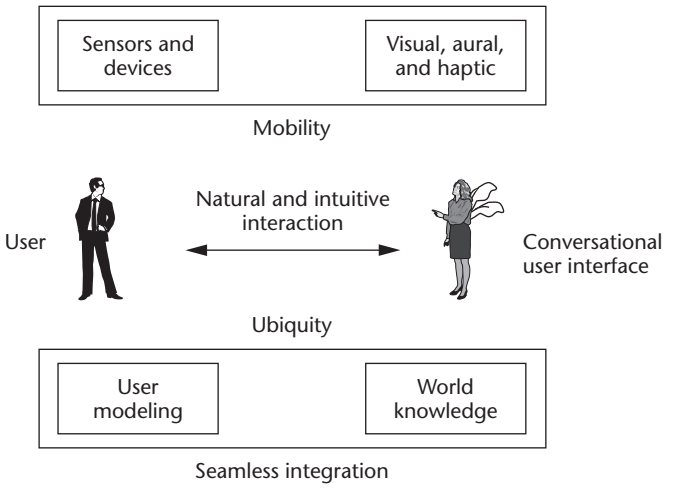
At first glance, this scenario seems similar to other futuristic visions, such as Apple Computer's Personal Assistant. While some aspects are similar, the guardian angel system is distinguished by its use of 3D displays, wearable computers, ubiquitous access, and 3D interaction.

Although real guardian angels aren't easy to get hold of, some of the computer technology needed for such a personal assistant is already available. Other parts exist in the form of research prototypes, but some technological breakthroughs are necessary before we can realize their potential, let alone integrate into our daily routines.

Science fiction literature and Hollywood films such as *Disclosure*, *The Matrix*, *The Thirteenth Floor*, or the Star Trek series have already shown us what an unobtrusive, personal, and expansive 3D interface might look like. In these imaginative works, the virtual worlds often appear undistinguishable from the real world.

Future VR and AR interfaces won't necessarily try to provide a perfect imitation of reality but instead will adapt their display mechanisms to their users' individual requirements. The emergence of these interfaces won't rely on a single technology but will depend on the advances in many areas, including computer graphics, display technology, tracking and recognition devices, natural and intuitive interactions, 3D interaction techniques, mobile and ubiquitous computing, intelligent agents, and conversational user interfaces, to name a few (see Figure 1).

Rosenblum[1] and MacIntyre and Feiner[2] took a look into the future of VR technology and multimedia interfaces, respectively. Brooks, in his IEEE VR 1999 keynote and follow-up survey article,[3] gave a personal assessment of the state of the art in VR at the turn of the millennium. The rest of our article will review research agendas in different AR and VR areas to shed some light on the feasibility of user interfaces of the kind we introduced in our guardian angel scenario.



1 Seamless integration of future AR and VR interfaces, bringing about the guardian angel interface in our example.

2 Augmented roundtable. Intuitive interaction in collaborative working environments.

Display technology

Today's personal head-worn displays are already smaller and provide better resolution than only a few years ago, making them suitable for real-world applications and mobile use (see Figure 2). They already let us project a guardian angel within the user's environment, although the user's perception of the image is limited to a narrow field of view.

Other technologies, such as retinal displays and laser-emitting diodes, are currently under development. They will further enhance the image quality and the field of view while facilitating lightweight models. The miniaturization process will continue, providing affordable, high-resolution personal displays, indistinguishable from regular sunglasses. Additionally, more sophisticated auto-stereoscopic 3D displays that don't require any wearable equipment at all will become available. This includes multiviewpoint displays and true holographic projections. The technology chosen for a particular application or scenario can then be based on the observer's needs or preferences for a personal or a public display. In our scenario, for example, a holographic projection system displays the angel within the user's private environment, and personal displays were used in public areas.

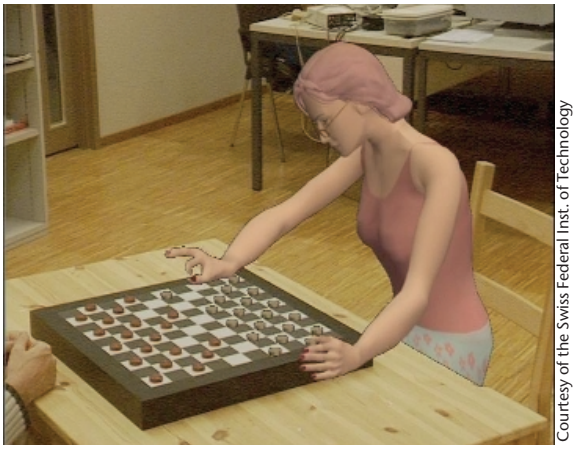
Sensors and devices

Many input devices for VR and AR currently exist, ranging from position and orientation trackers of different kinds to computer-vision-based systems and haptic inter-

© 1999 Columbia University

3 Outdoor mobile augmented reality application, visualizing historic architecture in its original location. This image was captured by a video camera aimed through a see-through head-worn display.

precise sensor data about yourself. Similar to GPS-based navigation systems, future services will be available to announce your exact position within fractions of an inch. Presently, high-resolution GPS-based tracking is used in combination with dead-reckoning for outdoor AR applications (see Figure 3), but general wide-area indoor and outdoor tracking solutions don't exist. Hybrid systems will overcome these limitations and provide the accuracy and speed required for AR systems.



Courtesy of the Swiss Federal Inst. of Technology

4 Autonomous virtual human playing checkers in real time.[4]

faces. Tracking and other sensing devices have become smaller and more accurate over the years, but robustness, general applicability, and the absence of tether-free solutions remain problem areas. The range of most indoor tracking devices is still limited to a small area.

Researchers are currently prototyping computer-vision-based systems, which might emerge as the most flexible and universal tracking devices in the future. However, this will require significant enhancements in performance and robustness. Sensors will be further miniaturized to be less obtrusive. Sensor fusion—for example, among a network of connected camera systems as well as between different types of sensors, such as ultrasonic, gyroscopic, pedometer- or odometer-based, or vital stats—will give an integrated overall account of the user's movements, behavior, and mood. In our example, this provides the angel with the information about the user's whereabouts in the house or office or lets her react to items and persons the user views and augment them with additional information (for example, during the meeting).

Mobility

We already use many mobile devices today, such as notebooks, personal digital assistants, and cellular phones, using a wide range of wireless services. In the future, the combination of sensors, smaller and smaller devices, and ubiquitous electronic services will let you access any data from anywhere and provide you with

Natural and intuitive interaction

Natural and intuitive interfaces play an important role in disseminating VR and AR technology. Today's interfaces are often cumbersome, requiring long training periods for users, and based on heavy or obtrusive equipment. Applications intended for inexperienced users demand a low learning curve and will benefit most from natural and unobtrusive interfaces. Other (more complex) interfaces, on the other hand, can be more efficient for experts. Some natural input modalities, such as gesture and speech input, as well as recognition of facial expressions are already used in some experimental setups today. More robust mechanisms will become available in the future, providing better support for groups of users (see Figure 3). For example, communication with voice and gestures will be feasible, even when user are talking to other people or in crowded places.

Promising research in the area of novel input mechanisms such as brain activity scanning (the thought interface) has already started, although it will take quite a few years until such interfaces will reach a level suitable for real applications. Support for disabled people will probably be the first type of applications where we'll find these types of interfaces. Nevertheless, such interfaces will never completely remove the need for multimodality, at least from the user's point of view.

Conversational user interfaces

Computer graphics and processing power have advanced dramatically over the last 20 years. Assuming these rates of improvements continue in the future, image realism will take great strides toward perfection. Animated characters, complete with synthetic gestures, voice, and facial expressions will become commonplace. Virtual humans (see Figure 4) represent a natural, familiar, and convenient user interface.

Today, we're still far away from providing a realistic conversational interface, but at some point in the future, after we've perfected 3D video recording, there won't be any perceivable difference between a 3D recording of a human and an artificial character. The latter, of course, doesn't only rely on highly sophisticated graphics and animations, but also on the quality of its behavior. Finally, it must perfectly adapt to the individual user (based on ubiquitous access to user modeling informa-

tion) and to the local environment. It must also combine this information with the general world knowledge to create a smart conversational interface resembling a real person. It's an open issue whether we should always be aware that we're communicating with a computer as opposed to a human being.

Seamless integration

Another important aspect comes along with mobility: the continued integration of all electronic data with the network (the Web) and therewith the ability to access and update all types of information anytime, anyplace. Access to a bank account from a wireless application protocol (WAP) capable mobile phone or email communication using your PDA are just the first signs of the dramatic change we'll see within the next several years.

In our example, the smart-house environment receives the information that you left, so it could switch off the coffee machine. The angel knows about the people in the meeting room and can inform you about them ahead of time. This integration of digital information with an advanced 3D user interface will make future use of VR and AR interfaces much more attractive to the average user.

Universal user interfaces

The success of windows, icons, menus, and point-and-click (WIMP) interfaces was because users familiar with one application immediately knew how to use the interface elements of another application. Several approaches pursue a similar universal interface for 3D environments. Others provide individual interfaces for specific applications, taking into account steeper learning curves. Flexible, adaptable, and more universal interfaces are still an open issue. A personal conversational interface such as the guardian angel would be one possible approach to covering a wide range of applications.■

Conclusion

The guardian angel scenario exemplifies how future developments in AR and VR user interfaces might change the way we interact with computers. Although this example is just one of several plausible scenarios, it demonstrates that AR and VR, in combination with user-centered design of their post-WIMP interfaces, can provide increased access, convenience, usability, and efficiency.

Acknowledgments

We thank all the participants of IEEE VR 2001's workshop on the future of VR and AR interfaces for the vivid discussions and their inspiring contributions.

References

1. L. Rosenblum, "Virtual and Augmented Reality 2020," *IEEE Computer Graphics and Applications*, vol. 20, no. 1, Jan./Feb. 2000, pp. 38-39.
2. B. MacIntyre and S. Feiner, "Future Multimedia User Interfaces," *Multimedia Systems*, vol. 4, no. 5, Sept./Oct. 1996, pp. 250-268.
3. F.P. Brooks, Jr., "What's Real About Virtual Reality?," *IEEE Computer Graphics and Applications*, vol. 19, no. 6, Nov./Dec. 1999, pp. 16-27.
4. S. Balcisoy et al., "Interaction Techniques with Virtual Humans in Mixed Environments," *Proc. Int'l Symp. Mixed Reality 01*, Mixed Reality systems Lab, Yokohama, Japan, 2001, pp. 81-86.

Readers may contact Broll at the Fraunhofer Institute for Applied Information Technology (FIT), Schloss Birlinghoven, 53754 Sankt Augustin, Germany; email wolfgang.broll@fit.fraunhofer.de.

Readers may contact the department editors by email at rosenblum@ait.nrl.navy.mil or Michael_Macedonia@stricom.army.mil.