Beyond WIMP

Andy van Dam
Brown University

In a sense, an interface is a necessary evil. The ideal user interface would let us perform our tasks without being aware of the interface as the intermediary. The longevity and ubiquitousness of the now two-decades old WIMP (windows, icons, menus, point-and-click devices) graphical user interface (GUI) should not mislead us into thinking that it's an ideal interface.

On the positive side, WIMP interfaces are relatively easy to learn and use because they exploit both muscle memory and image recognition. By using a great deal of commonality across applications, they create a de-facto standard and make applications accessible to a wide range of users, including preliterate children.

On the negative side, expert users find pure WIMP GUIs frustratingly slow and thus use keyboard shortcuts. In addition, they don't scale well—GUI bloat accompanies feature bloat. The most serious limitation, however, is that WIMP GUIs were designed for keyboard-plus-mouse desktop computing environments—an environment that takes advantage only of vision and primitive touch. Humans find it very natural to talk, gesture, shift their gaze, and listen all at the same time, but multiple parallel sensory channels are almost impossible in today's WIMP implementations.

The best example of a commercially available non-desktop environment requiring multisensory interfaces is immersive virtual reality (VR). Whether provided by a head-mounted display (HMD) or a CAVE (a cubicle with stereo projections on the walls and floor), VR environments typically employ multimodal interfaces using simultaneous speech and gesture recognition, as well as head (and body) tracking. Unlike WIMP interfaces, such post-WIMP interfaces are characterized by high bandwidth, continuous sampling and processing, and probabilistic decoding and recognition that take into account ("unify") mutually reinforcing input from multiple parallel channels.

New computing environments

The longevity of WIMP GUIs results in large part from the stability of the equally old desktop computing environment. That environment is changing dramatically, however, as computing power continues to explode and new forms of computing devices and environments arise. VR and personal digital assistants (PDAs) with part WIMP, part gesture-based post-WIMP GUIs are just the beginning. Mark Weiser's vision of ubiquitous computing and its variants, called invisible or seamless computing, are becoming a reality as computing power becomes embedded—in smart devices, rooms, and environments, in clothing, and even in our bodies (cochlear implants are here, nano-sensors/effectors and neural implants are being worked on). We will advance from one or a few devices per person to, eventually, swarms of devices per person.

At the same time, software agents will circulate through networks on our behalf. Some of these devices and agents, especially implanted devices, will work largely autonomously, individually, and collectively, without a real UI; others we will control more directly, using a variety of UI techniques. These new devices and computing environments require new UI paradigms and techniques. Designing the interface of the future is a surprisingly difficult task—consider replacing the 2D desktop metaphor with a 3D one (for example, a room or a building) that takes advantage of our natural ability to operate in 3D and our good spatial memory. To date there have been no successful, much less generalizable, 3D replacements for the 2D desktop metaphor.

User interface trends

Post-WIMP interfaces will not only take advantage of more of our senses but also be increasingly based on the way we naturally interact with our environment and with other humans. The WIMP direct-manipulation paradigm foreshadows such natural interaction. Let me focus on five trends that will result in more natural interfaces.

First, future interaction will combine multiple sensory channels and participants. Second, interactions will be high-bandwidth, continuous input modes that may involve, as in VR, body part tracking, gesture and speech recognition (based on natural language understanding), and haptic force input and feedback devices. Third, body-part and gaze tracking and identification technology will become increasingly passive, often

0272-1716/00/$10.00 © 2000 IEEE

video/vision-based, to eliminate intrusive special devices. Fourth, interaction will involve higher-level data and interpretation. Raj Reddy of Carnegie Mellon University uses the term "SILK interfaces" to describe future UI capabilities based on speech, image, and language understanding and driven by knowledge bases. Fifth, while WIMP GUI technology has benefited from the separation and subsequent specialization of interface and application, post-WIMP interfaces must reverse the trend, with the interaction becoming tightly interwoven with application-specific semantic decoding and processing.

WIMP interfaces won't disappear. They'll remain useful, albeit augmented and specialized to specific environments. A continuum already exists from pure WIMP to transitional, such as the Palm Pilot, to post-WIMP, such as VR. But when compute power, storage, and even some form of "intelligence" are embedded in a wide variety of different devices, the one-size-fits-all design philosophy of WIMP interfaces becomes increasingly irrelevant. As Bill Buxton persuasively argues, specialized devices and interfaces are often much more satisfactory than general-purpose ones. To make such customization feasible, we'll need far better user interface prototyping and design tools than we currently have.

Some research issues

Safe, effective, unencumbering, and affordable VR technology: Issues include latency management to avoid cyber sickness, greatly improved spatial and temporal resolution in display and tracking technology, and investigation of the long-term physical and mental effects.

Mixing real and virtual environments: Augmented reality requires greatly improved latency and registration. Current research explores using physical props in conjunction with user interface widgets. We also must understand the human behavioral differences among virtual, physical, and mixed environments.

Understanding and effective use of human capabilities: We need to develop a much better understanding of human perceptual, cognitive, manipulative, and social abilities and create an interdisciplinary UI design discipline based on that understanding.

Integrating direct and indirect control: We need to unify indirect control of agents and embedded devices with direct control of other devices, in a scalable fashion. The huge distributed-systems problem of controlling and integrating thousands of devices for each of hundreds of millions of users has the corresponding problem of providing individual users with a means of controlling and manipulating all their devices, data streams, and interactions with other users.

Rapid prototyping and design tools: Today's graphical GUI builders are purely oriented toward WIMP interfaces and provide no support for non-WIMP interfaces. We need powerful frameworks and a huge library of components to handle non-WIMP interfaces for wildly different components and environments.

Post-WIMP scenarios

Tele-immersive environments: In Henry Fuchs' "Office of the Future" vision, spatially immersive high-resolution images are projected on real surfaces in the office. Synchronized ceiling projectors and cameras use imperceptible structured light to continuously recalibrate the room's contents, including its occupants, to project the right image pixel-by-pixel. Vision-based algorithms then acquire a model of the scene and transmit it to remote sites. The effect will be of a shared, immersive, interaction space that blends appropriate portions of each participant's workspace.

Personal workspaces: Turner Whitted inspired the following description of an ideal personal work environment. I want to perform any task at any place using whatever tools I have available. A simple example is editing a multimedia document using voice in the shower, a tablet computer at the breakfast table, voice in the car, and a smart room with interactive wall-sized displays and passive interaction technologies at the office, while never explicitly opening or closing a new application or document. I want to work seamlessly, under any circumstances, without concern for the location or structure of my documents, while carrying out other routine tasks.

Conclusion

Our goal should be to design user interfaces that match our human perceptual, cognitive, manipulative, and social abilities. We want to interact as naturally with computers and intelligent devices as we communicate and collaborate with each other and as we manipulate our physical environments. Indeed, computers are increasingly being used not just as productivity and amusement engines for individuals, but to facilitate human communication, collaboration, and social interaction. Therefore, I feel we should increasingly focus on human-human interaction, not just human-computer interaction. This change in focus reflects not only changes in the mode of computer use but also the increasingly invisible nature of the computer in our environments. New environments must serve as forcing functions to give us far higher expectations for user interfaces than we have had previously.

Furthermore, the potential of future interfaces for handicapped people, such as paraplegics, autistic children, stroke victims, and even dyslexics, is phenomenal. Thus, despite the technology challenges, the greatest challenge lies in our understanding of human capabilities and how to incorporate that understanding into new design tools, methodologies, and user interfaces.

On a final note, let me observe, as a cynic in the Oscar Wilde sense (a frustrated idealist), that no matter how wonderful future user interfaces will be, it will take far longer for them to gain acceptance than we can imagine. ■

Readers may contact van Dam at avd@cs.brown.edu.