Interfaces for Humans: Natural Interaction, Tangible Data, and Beyond

Moderator
Michael Harris
NCR Human Interface Technology Center

Panelists
Bill Buxton
Alias | Wavefront Inc. and Silicon Graphics, Inc.

William T. Freeman
Mitsubishi Electric Research Laboratory

Mark Lucente
IBM Research

Hiroshi Ishii
Massachusetts Institute of Technology

Michael J. Sinclair
Microsoft Research

Michael Harris
NCR Human Interface Technology Center
mh@mindspring.com
www.ncrhitc.com

Huge advances in interface modalities are evident and imminent. We demonstrate and explore some of the most interesting, promising, and clever of these, as well as their integration into powerful multimodal systems.

When users talk about computers, they usually describe the interfaces, because, for most users, the interface is the system. As Bill Buxton says, "The most powerful force in shaping people's mental model of the nature of the beast is that which they see, feel, and hear." It seemed to take forever for toggle-switch panels to evolve into today's WIMPs, and both are still visual/motor-based controls; in fact, switch panels were probably more haptically satisfying! "Keyboards only work for people who know the Roman alphabet. In 20 years, people will laugh at us for calling that technology," says Mark Lucente.

Now, thanks to exponential increases in commonly available computer power and versatility (and concomitant cost decreases), significant progress in interface modalities and their affordances can be perceived. In this panel, we emphasize demonstrable and practical stuff; we have hardware to monkey with, ideas to ponder and try.

This is a gadget-intensive topic, and we present gadgets galore. Input devices that can tell systems where users are looking, the gestures they are making, the direction and content of their sounds and speech, and what and how they are touching. Display devices that image directly onto the retina, high-resolution miniature LCDs, spatial sound generators. Some of these innovative transducers operate not just noninvasively but invisibly. "No one should ever have to see a computer. The complexity should be soaked into the world around you," says Lucente.

While humans are adept at sensory integration and data fusion, computers are far less so. It is clear (and probably has been since GLOWFLOW in 1968) that multimodal interaction is a seminal goal, and that achieving it is a formidable challenge. Now that computational power seems be catching up with algorithmic understanding, the panelists can report and discuss exciting progress in this area.

Our panelists have decades of experience in interface design. Their perspectives are theoretical and pragmatic, incremental and radical; their work is elegantly inspiring and often delightfully unconventional. All were considered visionaries, but now their visions are achievable, and even industry is paying attention. They are seasoned practitioners with their own viewpoints. All are articulate, and none is shy; the Q&A and discussion periods promise to be stimulating.

Interfaces to newborn technology are usually "close to the machine." Early automobiles had spark-advance levers, mixture adjustments, hand throttles, choke controls. As automobiles have evolved, their affordances have moved "closer to the user:" speed, stop, reverse. We're tracking a similar evolution in human-computer interaction (HCI) space. Perhaps interfaces are finally growing up?

A video game can be controlled by real-time gesture sensing and recognition. (William T. Freeman)

User Domains and Appropriate Design

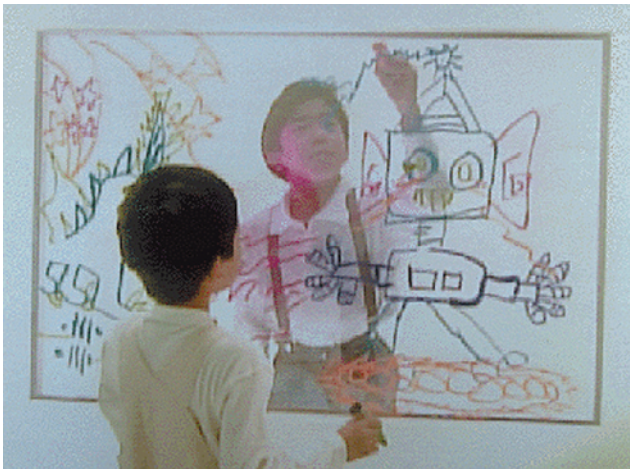
Bill Buxton
Alias | Wavefront Inc. and Silicon Graphics, Inc.
buxton@aw.sgi.com
www.dgp.toronto.edu/people/BillBuxton/billbuxton.html

When people are asked to "draw a computer," about 80 percent of the time they draw the I/O devices. This says two key things:

1 The most powerful force in shaping people's mental model of the nature of the beast is that which they see, feel, and hear.

2 This same shaping influence is an accident of history, and hence a candidate for change.

So, in designing a system, I can design its physical manifestation in such a way that its affordances conjure up the mental model that I am trying to encourage. And the better I understand the application domain, the skills of the intended user, and the context (physical and social) where the system is to be used, the more appropriate the mental model I can develop. And, consequently, the more appropriate the affordances and design of the system. But this flies directly in the face of how we design systems today. Today we follow the "Cuisinart," "superappliance" approach to design, which more or less says that the same basic type of box suits all types of users. But different users and tasks may well require (often radically) different approaches to what constitutes "a computer," because the two key components - the display and the input transducer(s) - are affected. I attempt to show how this particular approach to design affects the design of computer graphics systems for animators and industrial designers, as reflected in some of our research at Alias | Wavefront and SGI, in the form of live demonstrations and video examples.



Clearboard-2 seamlessly integrates groupware and videoconferencing technologies. (Hiroshi Ishii)

Computers Looking at People

William T. Freeman
MERL, a Mitsubishi Electric Research Laboratory
freeman@merl.com
www.merl.com/people/freeman/index.html

Computers can be used to interpret users' movements, gestures, and glances. Fundamental visual measurements include tracking, shape recognition, and motion analysis. For interactive graphics applications, these algorithms need to be robust and fast, and they need to run on inexpensive hardware. Fortunately, interactive applications can make the computer-vision problem easier. They can constrain the possible visual interpretations and provide visual feedback to let users adjust their inputs. I present several vision algorithms for interactive graphics and various vision-controlled graphics applications that use them: vision-based computer games, a hand-signal recognition system, and a television set controlled by hand gestures. Some of these applications can employ a special artificial retina chip for image detection or pre-processing.

Tangible Media

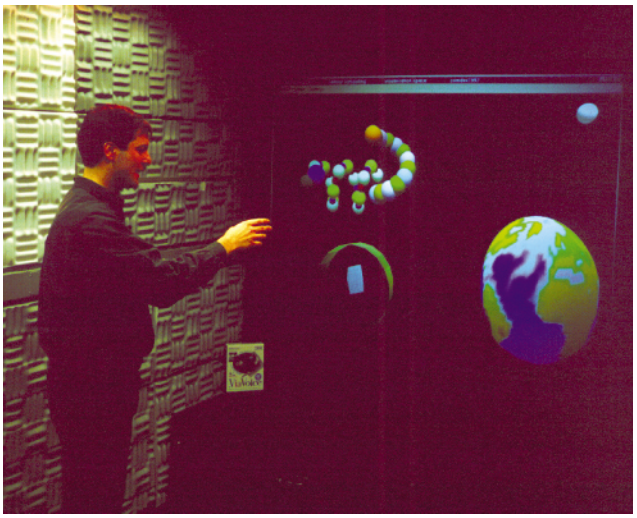
Hiroshi Ishii
Massachusetts Institute of Technology
ishii@media.mit.edu
ishii.www.media.mit.edu/people/ishii/

Eyes are in charge, but hands are under-employed. We live between two realms: our physical environment and cyberspace. Despite our dual citizenship, the absence of seamless couplings between these parallel existences leaves a great divide between the worlds of bits and atoms. "Tangible Bits" explores seamless interfaces among people, digital information, and everyday physical environments to go beyond eye-centric graphical user interfaces. We are designing "tangible user interfaces" that employ augmented physical objects, instruments, surfaces, and spaces as media to bridge virtual and physical worlds. We are making bits physically accessible through graspable objects as well as through ambient media in an augmented space. These interfaces emphasize both visually-intensive "hands-on" foreground interactions and background perception at the periphery of human senses through ambient light, sound, airflow, and water movement.

Natural Interaction

Mark Lucente
IBM Research
lucente@alum.mit.edu
www.research.ibm.com/natural/

Humans communicate using speech, gesture, and body motion, yet today's computers do not use this valuable information. Instead, computers force users to sit at a typewriter keyboard, stare at a TV-like display, and learn an endless set of arcane commands, all of which often leads to frustration, inefficiencies and disuse. We have created DreamSpace, a system that enables natural interaction through an intuitive yet richly interactive interface that "hears" users' voice commands and "sees" their gestures and body positions. Interactions are natural, more like human-to-human interactions. This information system understands the user, and - just as important - other users understand. Users are free to focus on virtual objects and information and understanding and thinking, with minimal constraints and distractions by the computer, which is present only as wall-sized 3D images and sounds (no keyboard, mouse, wires, wands, etc.) The multimodal input interface combines voice (IBM ViaVoice speech recognition), body tracking (machine-vision image processing), and understanding (context, and small amounts of learning). DreamSpace is essentially a smart room that employs a deviceless natural multimodal interface built on these emerging technologies and combined with ever-cheaper computing power. Future natural interfaces will allow information and communication anywhere, anytime, any way the user wants it - in the office, home, car, kitchen, design studio, school, and amusement park.

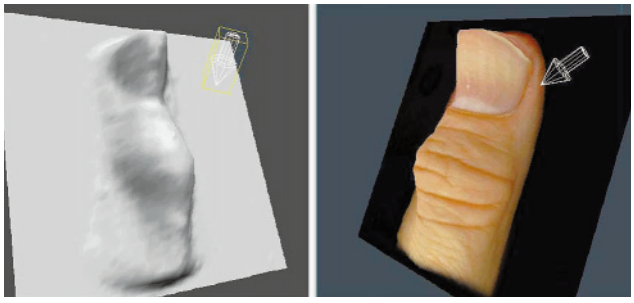


DreamSpace optimizes ease of use, enjoyment, and the organization and understanding of information. (Mark Lucente)

HCI Through Creative Plagiarization

Michael J. Sinclair
Microsoft Research
sinclair@microsoft.com
www.oip.gatech.edu/imtc/html/mike.html

To find creative HCI solutions that are both usable and affordable for a potential volume market, we can look at existing well-executed engineering efforts and discover a multitude of existing embodiments waiting to be re-purposed. From a steady diet of engagements with sponsors looking for low-cost commercializable outcomes, we have learned to investigate and exploit appliances from unrelated fields. As a result, we can demonstrate high-fidelity, low-cost solutions for 3D tactile feedback, digital panoramic photography, medical instrumentation, and low-cost 3D digitization. Significant gains can be realized through creative re-purposing. So: "Don't shade your eyes - plagiarize."



The Haptic Lens acquires 3D surfaces in real time, while applying a known force on the object or body part. (Michael J. Sinclair)