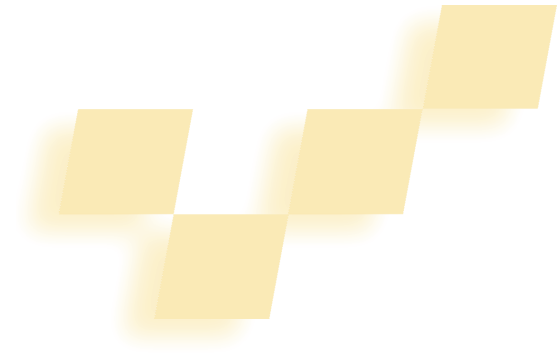


Perceptual Multimodal Interfaces



L. Miguel Encarnação
The imedia Academy

Lawrence J. Hettinger
Northrop Grumman Information Technologies

Computer graphics from its beginning has aimed to facilitate the dialogue between the human user and the computing system. Although the very meaning of dialogue—Merriam-Webster's definition is "a conversation between two or more persons; also: a similar exchange between a person and something else (as a computer)"—implies an equal share of participation from both sides of a conversation, the state of the art in interactive systems is far from providing similar opportunities to man and machine for information exchange and collaboration. Moreover, it seems as though the computer graphics research community has focused increasingly on image synthesis and animation, treating dialogue and interaction as the difficult stepchild that has to be recognized yet receives little attention.

Don't be a WIMP

While researchers have successfully applied various approaches to data and information visualization to diverse application areas and although there exists a plethora of commercially available products and solutions, with respect to the human-machine interface we're still limited to the decades-old Windows, icons, menus, pointers (WIMP) paradigm. We restrict ourselves to a few input and output modalities that make little use of the perceptive capacity of human users on the one hand, and impede their communicative capabilities on the other hand. This limitation, in turn, has hampered the introduction of interactive computer graphics to many areas of daily life that potentially represent important application areas as well as potent markets that could satisfy the needs of several neglected (in this regard) groups of society. Just imagine the challenge a surgeon faces with her task to save a life, with her hands occupied by a complex procedure, and her voice obstructed by a mask. Or the challenges of elderly people attempting to continue their participation in modern society by following the hyperlinks of a Web site from their home computer, when they more often than not are visually impaired and might lack some motor skill precision.

The vision—a work of many

To implement a vision of the computing system as a human-like partner, mimicking aspects of interhuman interaction requires—to a much higher degree than any other research agenda in the computer graphics community—the close collaboration of researchers from various disciplines, such as computer science, design, human factors, psychology, sociology, and artificial intelligence. Such collaboration emerges as an insurmountable challenge at first sight, pushing the availability of viable solutions even further away. This is because such interdisciplinary collaborative efforts can hardly be dictated by the need to evolve over time, the first and foremost hurdle is for a mutual understanding and acceptance between the different disciplines.

Yet there is hope. In recent years, a variety of groups have evolved that attempt to conduct research in certain aspects of post-WIMP user interface design. Under buzzwords such as perceptual computing, attentive interfaces, intelligent interfaces, and affective computing, researchers from different areas try to understand the nature and mechanisms of interhuman communication and map it to human-computer interaction. They accomplish this by developing technologies to enable the computing system to track and understand the human user's conscious and unconscious behavior, in conjunction with their data input. They also study technologies and methodologies that, upon successful application, would allow the computing system to provide adequate feedback, effective information presentation, and valuable decision aids while sending signals to as many senses of the human user as she can perceptively and cognitively process.

About this issue

This special issue showcases some of the approaches that could potentially push the development of post-WIMP user interfaces a significant step forward toward human-centered and anthropomorphic human-computer interaction. Initially we had hoped to actually find contributions that depict the whole spectrum, from analysis, to design, implementation, and evalua-

tion of such advanced interfaces. Yet the complexity of the problem, the limited amount of pages available for publication, as well as the immaturity of many critical hardware components led to a more pragmatic approach to selecting articles for this issue: showcasing specific important components of what we consider perceptual multimodal interfaces.

Bentley et al. describe their approach of supporting system awareness with respect to the user state by employing face tracking and observing the user's behavior in specific widgets. Realizing the need for robust and customizable application design, the authors present a component-based architecture based on perceptual user interface widgets for creating presence applications.

Takács and Kiss present a system that employs a virtual human interface based on advanced facial modeling and animation techniques, which uses artificial intelligence methods to adaptively and intelligently respond to the human user, who is sensed by the system using multiple sensory modalities. Application areas range from information and knowledge access to health care and rehabilitation.

Focusing more on educational applications, Harless et al. present a similar concept when discussing their voice-activated multimedia model, which employs natural language speech for one-on-one, face-to-face dialogue and lets a user conduct a virtual interview with the representation of a real person whose digitized video images are stored on a personal computer.

In the context of universal accessibility, Blenkhorn et al. present and discuss the architecture, evolution, and evaluation of screen magnifiers. Some of these perceptual interfaces for visually impaired users actually have multimodal characteristics by producing visual as well as audio outputs.

Finally, Sharon Oviatt presents an excellent survey on robust multimodal interface design, which doesn't limit itself to the average user working in controlled and static environments, but rather supports challenging field environments, mobile use, and diverse applications. ■

Acknowledgments

We thank everyone who helped make this issue possible. First, we thank the authors who submitted 11 high-quality and exciting articles for this issue. Choosing only four was difficult and cannot satisfy the breadth

of this topic. Second, we thank the multidisciplinary group of reviewers who helped with the selection. Finally, we thank the staff at *IEEE CG&A* for their immense help in preparing this special issue.



L. Miguel Encarnação is the president of imedia—The ICPNM Academy, and is its program director for research and development in interactive digital media technologies. He also holds an adjunct professorship position in computer science with the University of Rhode Island. Previously he worked as a senior scientist and head of the Human Media Technologies department of Fraunhofer CRCG, where he was responsible for research in mixed reality interface technologies, computer-aided education and instruction (CAE/CAI), and advanced distributive learning and training technologies. Encarnação has a BS and MS in computer science from Technische Universität Darmstadt, Germany, and a PhD in computer science from the University of Tübingen, Germany.



Lawrence J. Hettinger has been directly involved in multidisciplinary research in advanced human-machine interface concepts for more than 20 years, including research and development of advanced control and display concepts for NASA, all branches of the US armed forces, and several NATO nations, as well as for various medical organizations. He currently is responsible for Northrop Grumman's human-systems integration research and evaluation activities for advanced human-machine system development projects for the US Navy. Hettinger has a BA in psychology from the University of New Hampshire, and an MA and PhD, both in psychology, from Ohio State University.

Readers may contact L. Miguel Encarnação at imedia, 400 Westminister St., Providence, RI 02903; me@imedia-academy.org