Rebecca A. Grier, Ph.D.

Next Generation Human-Computer Interfaces

Abstract

The future of the human-computer interface is the future of computers. Technology is advancing exponentially. These advancements combined with greater knowledge of human capabilities and limitations is leading to the next generation of Human-Computer Interaction (HCI), "Reality-Based Interaction." Though a revolution to "Reality-Based Interaction" is still a ways off, we are getting glimpses of the interfaces that will be components of this generation (Speech Control and Virtual Reality) and those of the generation-after-next (Eye Gaze Control, Peripheral Displays, Haptic Displays, Brain Actuated Control). Before discussing these interfaces of the future, it is useful to describe how HCI has progressed, and the status of the current generation HCI.

Introduction

In July 1946, the U.S. Army Ordnance Corps began using ENIAC to solve complex mathematical problems related to ballistic trajectories and atomic energy (Weik, 1961). This was the first use of an electronic digital computer by the U.S. military. In the nearly 60 years since, there have been many generations of computers each one smaller, faster, and more powerful than the previous. As predicted by Moore's Law (Moore, 1965), computer power is doubling every 18 months. As such, computers that are only three years old are considered obsolete. However, the development of human-computer interfaces has not progressed at the same pace; only three generations of HCI have existed in the last 60 years. The first generation was the punch card used by ENIAC and other similar computers. The second generation of HCI was the command based interface (e.g., DOS). The third, and current, generation of HCI is the graphical user interface (GUI) most often instantiated as a WIMP (Window, Icon, Menu, Pointer). Improvements have been made to this generation of HCI, including color displays, scroll mouse, and touch-screens. However, these advances do not replace the GUI. Further, many other enhancements to the human-computer interfaces are just technological improvements (large screen displays, heliodisplays, six DoF joysticks, etc). These technologies are engaging, but have little impact on human performance as compared to the cost of implementation.

In fact the slow development of human-computer interfaces has constrained the total system performance of high-tech systems. The power with which computers can perform in many ways exceeds the capabilities of humans given the current method of interaction. For example, computers can already process and display faster than humans can perceive or react to the output. Additionally, computers can process more information than we can input given current interfaces. Moreover, cognitive resources of humans cannot be focused exclusively on the tasks, because operating current technology imposes a cognitive load on users.

This cognitive load is due to what Norman (1988) has called the "gulf of execution." In other words, users' intentions do not align well with the actions required to achieve the goal in the use of the current generation of HCI. Typing and using a mouse are both artificial actions that require training; neither is how we interact with any other aspects of our environment. We must translate the action we want to perform into the action that will actually accomplish the objective. This translation requires a certain amount of mental resources, thus reducing the amount of resources available for performing the primary task.

Reality-Based Interaction

In order to both reduce the cognitive load of operating a computer and to better utilize the increased processing power of today's computers, we need HCI that takes advantage of all the sensory modalities. Furthermore, these modalities must be used in a way that capitalizes on their strengths. This will reduce the resources required for interaction and increase the number of channels through which humans can receive information. This is the basis for the theory that the next generation of HCI will be "Reality-Based Interaction." In other words, we will be able to interact with computers in a manner that is similar to how we interact with the natural world. As such, the barrier preventing humans from taking advantage of all that computers have to offer is eliminated.

Speech Interface

Speech based control is the interface which has the most interest of the Reality Based Interfaces. It has received considerable attention in both military (e.g., McLaughlin, & Guilliams, 2003) and commercial domains. Often speech or voice controlled interfaces are given high preference marks by operators (McLaughlin, & Guilliams, 2003). This preference exists, because speech is the most natural and direct method of human social communication. It is through speech that humans communicate the largest amount of information, which makes it an excellent candidate for Reality Based Interaction.

However, current technology is not nearly as good at gathering information through speech as humans. Automatic speech recognition systems are poor when used in noisy environments. Additionally, the current technology affords limited flexibility, because of the restrictions of the algorithms that are used for recognition.

That is, if a system is to be used by a large number of people, then the vocabulary of commands that can exist must be relatively small. Conversely, a larger set of commands can be used only if one individual is using the system. Moreover, the individual must spend a substantial amount of time training the system to recognize his/her voice. Further, the individual may not be able to use the system if s/he develops a cold, because of the alterations to his/her voice. Given current technological limitations, speech control should only be used in environments where hands are occupied with other tasks, and then only with short intuitive commands.

Even as speech technology becomes more robust, we cannot forget the old adage "a picture is worth a thousand words." Some information (e.g., complex spatial relationships) is more easily conveyed by drawing with mouse and Smartpad. Other information (e.g., action desired) is conveyed better by gesturing than by speech control (Oviatt, 1999). Therefore, even as the technology matures, speech cannot be the only input device available to the operator. Moreover, it is known that speech changes when other methods of input are available. That is, the input is reduced to the most compact form and the error prone elements of input are not utlized (Oviatt, 1999). Designers therefore must be cognizant of "when" and "how" operators will speak to a machine for speech control to reach its potential.

Virtual Reality

A hallmark of the shift to reality-based interfaces will be the omnipresence of virtual reality. Currently, virtual reality is accomplished by one of two methods. The first is the use of a CAVE (Cave Automatic Virtual Environment), essentially a room in which large screen displays are used. The second method, wearable virtual environment, involves users wearing helmets (tracks movement and presents the virtual environment) and data glove(s) (tracks gestures and hand movements.). Both the helmet and glove are tethered to a computer.

Both systems have shown great benefits for training specific scenarios. The CAVE has two benefits over the helmet based version. First team members can interact more naturally, because the members are not presented virtually. Second, the lack of tethers permit a wider range of motion. Conversely with the wearable virtual environment, the operator is not confined to the space of the CAVE. Further the data gloves allow the user the opportunity to interact with virtual objects. Currently though, there is no haptic feedback provided when the user interacts with these virtual objects. Thus, the fidelity of the experience is reduced.

At this time, Virtual Reality is primarily a vehicle for training. For virtual reality to have benefit as something other than a simulator, it is imperative that we design multimodal interfaces that effectively integrate controls and displays of various types (i.e., Speech, Touch, Gesture Control, Volumetric displays, 3d Audio) into one seamless interface. This integration must allow us to function as we do naturally within the physical world. If we are successful, the promise of Reality Based Interaction will be realized and human performance will not be a limiting factor of total system performance.

Generation-After-Next

Jacob (2004) has recently theorized about the "generation-after-next." This generation of interfaces will go beyond reality based interactions to augmented reality. That is, the majority of the interfaces will capitalize on natural behaviors, but additional interfaces will extend these interactions above and beyond our natural capabilities. Jacob (2004) compares this augmented reality to the way superman interacts with the world (i.e., the same abilities as humans with a few bonuses). In order to effectively integrate augmented reality interfaces, one must have a strong understanding of human capabilities to ensure that the human can effectively use the interface without overloading the reality based interactions.

Eye-Gaze Control

Eye-gaze control is an example of such an interface. It uses eye tracking to determine a user's intended action with a visual display. That is, simply by looking at an item, one can activate that item. Eye-gaze control has been found to be as reliable in selection tasks as mouse based control, with the added benefit of faster response times (Sibert & Jacob, 2000). Participants in Sibert & Jacob's (2000) experiment reported that selecting targets with eye-gaze control felt more like the computer was responding to their intentions, rather than their actions. This is an amazing finding considering that the technology has not yet reached full maturation (Sibert & Jacob, 2000).

That is to say, there are several technological limitations of the systems. Currently, eye-gaze systems require users to either keep their heads very still or wear a cumbersome helmet to track head movements. Furthermore, users must calibrate the system each time it is used. Also, the systems can lose calibration if used for extended periods of time.

In addition to the technological limitations, there are limits as to what the system can be used for given human capabilities. For example, there is no input from eye-gaze control beyond selection that a human can perform naturally. Any additional functionality would require training or the use of another input device, and thus would reduce any gains in performance. However, a system that smartly integrates eye-gaze with speech control could be very useful (Jacob, 2004).

Peripheral Displays

A second HCI technology that capitalizes on augmented reality is the peripheral display. Peripheral displays typically appear to be light sculptures. However, rather than changing in intensity or color, the display changes based on a factor that is selected by the user (e.g., weather conditions, traffic status, etc.). This is information that the user is interested in, but does not need to be alerted to every time it changes. Nor does the user want to spend energy to search for this information. A peripheral display takes advantage of the fact that humans have an extensive awareness of their settings without distracting them from a primary task; the way an audio alert or pop up notification would distract. However, when the user wants the information it is easily obtainable.

Haptic Displays

A third interface that falls into the class of augmented reality is the haptic display. Haptic displays present information via the sensation of touch. By capitalizing on our sense of pressure and vibration in a logical manner, it is plausible that bandwidth to the user can be increased. Rupert (2003) reported that by presenting pressure and vibration through a vest worn by pilots, spatial disorientation was decreased and situational awareness to both flight navigation and threats was increased. Furthermore, the training required to use the haptic display was less than for traditional displays (Rupert, 2003).

This is a clever use of the sense of touch that extends our capabilities. Spatial orientation is normally perceived via gravitational forces, and thus, intuitively translated to touch. Similarly, humans use their sense of touch to allude bumping into objects when other senses are unavailable (e.g., dark room). Thus, the use of a haptic display in this manner is a superior example of what technology can do if based on a thorough understanding of human capabilities and limitations.

Brain Actuated Control

Finally Brain Actuated Control (BAC) is the holy grail of augmented reality. Through BAC, a user would just have to think about an action, for that action to occur. Currently there are several avenues of research towards BAC. These include using (1) EEGs (Event Related Potentials, the MU-rhythym or pattern mapping), (2) Facial Activity, and (3) Neural Implants. Various military and academic research facilities have had success with each of these methods of BAC. However, there are technical limitations with each method. For example, with each, wireless communication between the implants and the device being controlled are not sufficiently reliable. As such, the user must be connected via wires to the device they are controlling. The exception being Helmet Mounted Displays, which can have the sensors for EEGs or Facial activity incorporated into the helmet. Secondly, the systems require extensive training and calibration for use. Thirdly, in regards to EEG based BAC, there is high potential for interference in identifying what is a signal and what is noise.

Conclusion

Though the future of reality based interfaces and augmented reality is foreseeable, it is imperative that engineers, computer scientists, and behavioral scientists work together in the development of HMI. Only by understanding how humans perform tasks can we ensure that we develop technology that supports the human. This includes understanding how we prefer to communicate and obtain diverse types of information in various environments. By not focusing on the human, we may well be developing technology the human must work to support; rather than technology that supports the humans' work.

References

Jacob, R.J.K (2004). Reality-based interaction: Understanding the next generation of user interfaces (preliminary draft). http://www.cs.tufts.edu/~jacob/theory.pdf

McLaughlin, A.B., & Guilliams, N.M. (2003). Evaluating speech interface concepts for air battle managers: From field observations to laboratory simulations. Paper presented at the 50th Meeting of The Department of Defense Human Factors Engineering Technical Advisory Group (DoD HFE TAG) in Phoenix, AZ.

Moore (1965). Cramming more components onto integrated circuits. Electronics, 38.

Norman, D.A. (1988). The Psychology of Everyday Things. Basic Books, USA.

Oviatt, S. (1999). Ten myths of multimodal interaction. Communications of the ACM, 42, 74-81.

Rupert, A. (2003) Tactile displays and their integration with other modalities. Presentation at the 49th meeting of the Department of Defense Human Factors Engineering Technical Advisory Group (DoD HFE TAG) in Augusta, Ga.

Sibert, L.E. & Jacob, R.J.K. (2000). Evaluation of eye gaze interaction. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, The Hague, Netherlands, 281-288.

Weik, M.H. (1961/ Sept. 2004). The ENIAC story. Available at http://ftp.arl.mil/~mike/comphist/eniac-story.html.

Acknowledgement

The author thanks Mila Marchosky, Owen Seely, Brad Collie, Katie Peters, Melissa Weaver, Udo Goff, and Daniel Wallace for all of the valuable information they provided in researching this article.

Rebecca A. Grier, Ph.D. is the principal author and she is a Human Systems Engineer with Aptima, Inc. Primarily she conducts analyses to inform the design of systems for both commercial and military domains. Prior to joining Aptima, Dr. Grier was an Engineering Psychologist in the

Human Systems Integration (HSI) Group at the Naval Surface Warfare Center, Dahlgren, VA. In this position she developed courses to educate the Navy's acquisition workforce about HSI and provided technical support to acquisition programs including writing HSI plans and conducting system evaluations. Rebecca has also worked in the Human Factors Engineering group at SBC Laboratories, Inc. where she evaluated and designed installation wizards, manuals, dashboards, CPE, IVRs, web pages, and voice UIs. Dr. Grier received her Ph.D. and M.A. in Human Factors/ Experimental Psychology from the University of Cincinnati, and a B.S. Honors in Psychology from Loyola University, Chicago.