Click/Talk/Touch/Look/Think Here: User Interface with Virtual Space

Tim Baldwin, Alexandros Christodoulou, Cary Gillenwater, Nicholas Johnson, Amit Kumar, Gary Marchionini, Brian Moynihan, Gyorgy Polczer, Derek Rodriguez, Joshua Purvis, & Jeff VanDrimmelen

University of North Carolina at Chapel Hill

INTRODUCTION

The most critical bottlenecks in information flow are human input and output (I/O). These bottlenecks are due to a combination of physiology, cognition, and technological prosthetics and are strongly exacerbated when the information flows are mediated by or with information technology. As people interact with each other or with information systems, the actions taken and the resulting information flows are outputs from the initiator's perspective and inputs from the receiver's perspective. This paper provides an overview of the I/O problem space by examining different theoretical models that are or have been considered in Human Computer Interaction (HCI) as well as summarizing different kinds of techniques and devices that are in use or in development to facilitate human information interaction in cyberspace.

People sense the natural world and listen, read, and view information in the built world at differential rates ranging from a few bits per second to millions of bits per second depending on the perceptual organ. People move, talk, and write at relatively slow rates but we have created tools to change the rates. This paper presents an overview of different input devices organized by the degree to which people consciously control the devices (explicit vs implicit), considers some of the advantages and limitations of these devices and trends toward using multiple devices to facilitate natural human-computer interaction.

TASK MAPPING THEORIES

The main impetus of HCI is the study of how humans interact with technological systems. Originating as a field of study in the 1980s, this discipline initially situated its theoretical framework within cognitive psychology (Seow, 2005). Drawing from Shannon and Weaver’s (1949) Information Theory, HCI sought to create/improve computer systems interfacing with humans. During these early years time tasks were mapped to devices and software, but not necessarily reflective of real life experiences. The initial focus was on the systems themselves and how they functioned, or did not. Much of interface design was predicated on a “top-down” approach that first focused on design objectives and only later considered the user's interactions (Foley, van Dam, Feiner, & Hughes, 1990). Due to the initial reliance on cognitive theory in early HCI research, the user was largely overlooked based upon the belief that people’s thought processes were quantifiable and thus predictable (Suchman, 1987, as cited in Shneiderman & Plaisant, 2005); however, this theoretical model has subsequently given way to more user- and human-centered approaches to design that recognize that user actions and experiences are situated in a time and place, i.e. a context, and are receptive to this context. In turn, the I/O paradigm has become viewed as less a binary system and more of a complex one where interaction is now determined to be a *dialogue* between the human and the computer, all of which is situated in a context of social, cultural, and organizational significance (Kaptelinin & Hardy, 2003, emphasis added). This shift away from cognitive theory has sparked much debate within HCI as to what theory (if any) should underpin the design of interfaces (Castel, 2002). The considerations are now both philosophically and materially what should be considered when mapping “tasks” and “interactions” from the physical plane to the informational.

Linguistic and Metaphor Theories: A Different Foundation?

Much of I/O research and design follows what resembles a linguistics framework, using words such as “user-computer dialogue”, “language”, “rules and vocabulary”, and so forth (Foley , et al., 1990, pp. 393-394). This theoretical perspective allows for technically rich concepts to be translatable to a common abstracted conception, i.e. language. For example, Foley , et al. utilize this framework when they outline the three types of designs involved in interfacing. These are the functional design, also called *semantic design*, the sequencing design, also known as the *syntactic design*, and the binding design or *lexical design*.

Language also plays an associative function within I/O interface design; often I/O is transformed into a metaphor to help translate the concepts and functions of interfacing into terms that the user can connect with more easily (Hamilton, 2000). One common example is the metaphor of the “desktop” for the operating system interface and the use of “menus”, “folders”, and “windows” for program and data interfacing. Kay (1988) and Goldberg (1988) found that designers have also been able to utilize metaphors to assist them with conceptualizing interfaces (as cited in Hamilton, 2000). This metaphor theory could be considered foundational with regards to I/O, because it has played such a significant role in the nomenclature of interface systems, one that was critical, especially in the beginning of design when there was no known existing over-reaching conceptual framework.

However, metaphor theory has its issues. At the time of use, metaphors are conceptualized by the user, and thus are based strongly on personal cognitive associations (Hamilton, 2000). These associations contribute to expectations of what an interface will or will not do. Initial research has indicated that when the metaphor or association is violated, it can create cognitive dissonance leading to the user not trusting the interface. However, this research also indicates that this may be less of an issue as users become more accustomed to interface systems. Furthermore, as users have become more accustomed to commonly used interface systems, HCI has begun to pursue the possibility of interfaces "disappearing", i.e. of the user being unaware of their presence (Scoari, 2001, Schiphorst, 2007). Metaphor theory may be contributing to this disappearance as the metaphors are becoming less figurative and more literal, e.g. data is no longer data, but files.

Revisiting Cognitive Theory

Much of HCI's early theory was built on cognitive science (Sutcliffe, 2000, Kaptelinin & Nardi, 2003). Information processing was to guide the design of I/O interfaces. However, with a greater recognition that humans are more complex than originally conceived, cognitive theory has come to be challenged as insufficient to explain HCI. Furthermore, it has been argued that cognitive theory has scalability issues, based on both the complexity of human relationships with the computer and the complexity of modern technological systems. Barnard and May (1999) have argued that one way to deal with the scalability issues of cognitive theory is to have a system of theoretical models that would address different aspects of users' cognition and interactive systems design (as cited in Sutcliffe, 2000). Citing this complexity as well, Scolari (2001) has argued that one theory is not enough, proposing a semio-cognitive theory to accommodate the complexities underlying current HCI.

Noting that semiotics is not only about the science of signs, but the "sense production and interpretation" as well, Scolari (2001) connected semiotics and cognitive theory together (p. 85). Perception to Scolari was to be based on the relationship between the designers and the users. The area between was where the production of meaning was to take place. Kaptelinin & Hardy (2003) echoed this argument when they stated that technology research has moved beyond the realm of information processing and into the social, cultural, and organizational contexts of people's lives. Additionally, in a panel discussion on "second wave theories", panel participants added other possibilities of HCI theory based on these contexts.

Language/action theory is focused on the linguistic interactions of people in a conversation, or "loop" that ultimately works toward completion of a task (Kaptelinin & Hardy, 2003, p. 693). This differs from the previous semio-cognitive theory, in the sense that it accounts for the connection between speaking and human cognition via actions of the user. However, this theory only focuses on the space between the person and the computer where the dialogue or action exists. Activity theory on the other hand, incorporates tenets of language/action theory, but extends them to unite "the interpersonal aspects of acting through speaking with a focus on material objects" (Kaptelinin & Hardy, 2003, p. 693). In short, activity theory makes the user an active member in the design of the interface (Shneiderman & Plaisant, 2005). In contrast to activity theory, distributed cognition theory moves away from the individual by considering cognitive processes that are "enacted in interaction with the social and material environment" (Kaptelinin & Hardy, 2003, p. 693). Distributed cognition theory also asserts that knowledge is stored in other places besides the individual (Shnedierman & Plaisant, 2005). Furthermore, Bodeker (2006) posits that we now are on the threshold of a third wave of theory that focuses specifically on the cultural aspects via aesthetic theory, expansion of the cognitive to the emotional realm, and experiential theory.

Ultimately, the central issue in these post-cognitive theories is the user and the user's context, what Shneiderman and Plaisant, (2005) call "context-of-use theories" (p. 95). As the context expands to incorporate more aspects of a user's life beyond say the workplace, the need for tailorability of the interface by the user, something Foley , et al. called for in the 1980's, has come back to the forefront.

EXPLICIT AND IMPLICIT INPUT

There are many ways to think about how people interact with computers, depending on whether human physical action, hardware device, work tasks, or mental requirements are emphasized. In this paper we emphasize the degree to which people consciously control the input device.

Explicit Input

This section provides overviews of standard input devices such as keyboards and mice, speech and other explicit forms of interaction.Human generated sound, gesture, gaze, and brain activity are considered along with the different devices that facilitate these actions.

Keyboard, Mouse/Pointer, Touchpads, and Stylus

The most widely used explicit input devices are the mouse and keyboard. These devices have undergone some interesting developments in the past 20 years in an effort to increase the efficiency and ease of input. However, they are still a long way from achieving the goal of 'natural' interaction where the input device becomes transparent to the user (Schiphorst, 2007).

Over the last 20 years, most keyboards have utilized a variation of the QWERTY layout. Only recently have keyboard developers begun to move away from this traditional format and have begun tolook at different layouts including multi-display layouts and digital. Some examples of alternative keyboard layouts are the Frogpad keyboard that uses only one hand to allow for greater mobility, the Optimus Maximus keyboard that includes a completely customizable display for each key, and the One-Key Keyboard that can be worn on the user's wrist (Kim , et al., 2006). Furthermore, digital keyboards are also an attractive area of research, becausewithout hardware limitations developers are able to completely customize a layout and make it as big or small as they want. Many digital keyboards still use a variation of the QWERTY layout, but there are some researchers developing different layouts such as the logo based keyboards for people with motor impairments (Norte & Lobo, 2007).

Since Engelbart's initial conception in 1963, mice have also been redesigned and augmented. The most common augmentation is additional buttons, but there have also been efforts to create tactile mice that vibrate under certain conditions, 3D mice that tilt using a rounded bottom, two-ball mice, two-handed mice that allow for additional input, and most recently pressure sensitive mice (Cechanowicz , et al., 2007, p. 1386).

Another input device that is becoming more common in the market is the stylus. The stylus is a pen type device, typically associated with PDAs and Tablet PCs, which offers an alternative and potentially more natural way of interacting with data. Although the technology for the tablet PC has been around for decades, it was not until 2002 when Microsoft introduced it to consumers that it gained popularity (French, 2007, p. 84). Currently it is used in many fields such as the arts, music, and sciences, which benefit from direct interaction. Digital keyboards can be easily accessed and documents can be 'inked' like you would with a normal pen and paper. Ultimately, with styli we are getting closer to natural interaction with a computer.

Perhaps the most intriguing and promising technology of all of the implicit input devices is the touch and multi-touch display. These devices allow direct manipulation of digital objects with the touch of a finger and are the ultimate synthesis of the tools mentioned above. With a multi-touch display users can touch a screen in many different spots just like you would with a stylus. Furthermore, a user can also bring up a digital keyboard in any type of layout they want.

Speech/Sound

Human produced sounds, especially speech, are also innovative possibilities for human-computer interaction. We focus here on speech recognition, the crucial technology to support speech interfaces. Computer speech processing can improve work as well as daily activities for different types of populations. For example, Automatic Speech Recognition can benefit Medical and law professionals by providing instantaneous transcription and playback services Furthermore, automated call centers handle daily plane ticket booking requests offering travelers instant access to schedules and other travel services. Companies such as Microsoft, Google and Yahoo are working on software to allow drivers to talk to their car or Internet users to talk to their search engines and then listen to their search results. Computer speech processing can also make life easier for visually impaired populations, recovering patients or even psychiatric or psychological patients. Nevertheless, the widespread adoption of speech technology is not yet possible because of challenges related to factors both internal and external to speakers.

The main challenges of speech recognition systems are speaker independence and error rate (Benzeghiba et al, 2007; Baker, 2006; Deng & Huang, 2004). Speaker independent speech recognition error has been steadily dropping since 1988 (Deng & Huang, 2004); nevertheless, the most successful commercial program, *Naturally Speaking*, still requires lengthy speaker dependent training, and internal factors related to the speaker such as age, sex, pitch, rate of speech, and emotional state pose challenges to the template matching models that speech recognizers use. For example, if the recognizer is trained on an older person, the recognition rate will drop when a younger person uses the system. Speaker external factors such as external noise, pronunciation variation, as well as conversational setting pose additional challenges.

Compilation of large spoken linguistic corpora such as the *Switchboard* and the *Fisher* provide an expanding amount of data for the purpose of training speech recognizers. Computation linguists can use the corpora to calculate the probability of word occurrence depending on the previous and following context . Results from these studies can provide word recognizers with probability estimates of upcoming information given the already recognized word sequence. Therefore, speech recognizers are acquiring constraint satisfaction methods similar to the ones used by the human language processing mechanism.

Recent methods also take advantage of visual cues in order to increase the robustness of the recognizer. There is a close relationship between the shape of the mouth and the phonetic sequence a speaker is about to produce (Chen, 2001; Potamianos, 2006). For example, when speakers produce a stop consonant, the lips are always closed in order to block the airflow from the mouth. However, when the speaker is producing a vowel the lips are always open in order to allow for air to escape the mouth. Researchers are beginning to build viseme inventories that are similar to phonemic inventories of each language, which will help reduce error rates especially under noisy environments.

Parallel to speech recognition, researchers are starting to explore non-speech recognition. The idea is to take advantage of speaker elicited sounds that are non=linguistic in order to control basic computer commands. A recent study explored the possibility of using a humming sound with different pitch contours in order to control arrow cursors (Sporka, 2006) . The researchers found that controlling the cursor is more efficient with a humming sound as compared to speech input when under pressure. This type of a sound input can provide additional computer handling capabilities to patients with different types of speech or motor deficiencies.

Developments in speech synthesis are geared towards making synthesized speech sound more natural. Currently most synthesizers use formant synthesis, which replicates the formants associated with each phoneme (Baker, 2006). However computationally efficient this method might be, it still lacks in naturalness. Concatenative synthesis is computationally heavier but can reach higher more desirable output. In this method a serious of prerecorded sounds or words are combined together. Even though speech corpora are becoming more readily available, these methods require an even larger number of corpora in order to acquire a sufficient vocabulary to replicate a large number of phonetic sequences.

The ultimate goal of computer speech processing is to create an I/O device that fools the speaker into thinking the computer is another human conversational partner. To this effect, the move towards bimodal (sound and vision) computer speech processing can have dramatic effects by increasing robustness (Chen, 2001; Massaro, 2000; Potamianos, 2006). Even more, when combined with the development of conversational agents, intelligibility of synthesized speech can increase drastically (Massaro, 2000). Maybe an annual “who is my interlocutor” competition should be established, similar to the annual Turing test of the artificial intelligence community. It seems though this is still a distant possibility because semantic and pragmatic knowledge will have to be built into the speech processors, a process which is under slow progress (Benzhegiba , et al. 2007; Deng & Huang 2004).

Gesture

Gesture technology research aims to expand the capabilities of HCI by providing a new interface for controlling computers using hand posture and movement. The input channel to the computer can consist of one or more cameras, one or two special gloves or a hybrid mix of both. A large number of research projects are active in this area and several excellent surveys (LaViola, 2004; Mitra & Acharya, 2007; Jaimes & Sebe, 2005; Cerney & Vance, 2005) exist to introduce the field. An active and promising subset of the current research looks at making use of gesture technologies in the medical field. We focus on medical applications here.

Hospital operating rooms present a wide range of issues for surgeons who rely on computer interfaces to perform surgeries. In general, computers and their peripheral equipment are hard to keep sterile, the conditions are cramped and awkward for traditional mouse and keyboard combinations, and there is a large degree of background noise (Grätzel, Grange & Baur, 2004). One common solution is an awkward and error-prone two-step interaction between the surgeon and an assistant manning a computer displaying critical data. Replacing the assistant with a vision based system removes the extra step and places the surgeon directly in charge of the interaction with the system.
In one study (Grätzel, Grange & Baur, 2004) researchers replaced simple mouse interactions with gestures. In a second study with richer image data, Wachs , et al., (2005 & 2007) captured both hand posture and gestures for a richer set of controls. In twenty trials, each consisting of four tasks, they reported highly accurate results from 95% - 100% for each task.

However, sterilization is not just an issue in the operating room. Gesture based input can also provide a richer and dynamic interaction with complex visual medical data in other medical environments as well. For instance, some researchers are seeking ways of eliminating contact between humans and computer equipment for basic patient record review (Feied , et al., 2006) while others have developed a prototype system with a gloved based gesture and hand posture input to work with radiological data (Tani , et al., 2007).

In a similar vein, other studies are working on a new way to sculpt 3D medical data (Da Rosa Junior , et al., 2007). Interestingly, more than one channel of new input was explored in this experiment, employing a combination of gesture and blown air. Extending this concept, multimodal systems accept data from two or more input channels, for instance voice and gesture or gesture and the keyboard. The advantage of the multimodal approach is the extension of current interaction models rather than replacing one form of input with another (Jaimes & Sebe, 2005).

Augmented reality is another interaction model that can be paired with gesture-based technologies to create powerful new tools to interact with computers. By overlaying physical objects with virtual data a new kind of workspace is created with physical tools creating only virtual information. For example, a prototype system developed by Fischer & Bartz (2005) combines an augmented reality interface with gesture input tools and infrared-tagged surgical tools. Particularly interesting is their use of a virtual menu system that can be positioned and used close to the physical object being manipulated.

Although we typically think of gestures as hand based input, any tracked portion of the body or even the body as a whole can be considered as an input channel. Touching on the possibilities of new technologies to open new worlds to the disabled, for example Pei , et al. (2007) prototyped an intelligent wheelchair. Using a camera based system that tracked the angle of the head they were able to provide directional input for a motorized hands-free wheelchair.

The medical profession is excellent ground for practical uses of evolving gesture technologies. The field's needs are somewhat basic, and even today's technologies are sufficient to provide relatively low cost and high accuracy systems.

Gaze

Tracking eye movements is probably the oldest method to determine the path of a person's attention. The study of eye movement dates back to the early 1900s when eye-tracking methods were primitive. The location of eye fixations was quite invasive, involving direct contact with the cornea of the eye, while the head of the participant was fixed or motionless. With the advent of sophisticated cameras, eye-tracking techniques have come a long way. Today a participant's gaze is tracked using corneal reflections and the head is freed from all constraints.

The dawn of the computer age has pushed researchers to scrutinize a variety of eye tracking methods to address challenges faced by the users in human-computer interaction. This technology could be of a great value, finding answers that would help the optimization of computers as tools. Early work in this area focused on disabled users and flight simulators in an attempt to understand the gaze patterns of people using a system. With the infiltration of computers and related devices like PDAs into our lives, as well as increased use of Internet, e-mail, and videoconferencing, there is an increased interest in using eye tracking for usability studies and to develop advanced user interfaces using gaze as an input device.

Gesture, speech, and gaze have been studied as passive non-command inputs by researchers in the field of human computer interaction. Eye movements form an ideal illustration of a non-command based, passive way of input. To successfully accomplish this, a careful examination of the gaze interface design is warranted to avoid problems related to methods of engaging and disengaging the system. Like the other passive, non-command inputs, eye movements are habitually non-intentional; this poses a great interpretation challenge to make the user- experience a pleasant and effective one. It is often referred to as the "Midas Touch" problem i.e. unnecessary activation of commands every time a user looks at something (Jacob, 1990., Tanriverdi & Jacob, 2000). This also leads to another challenge with gaze interface as the users would have to familiarize themselves to operating devices by making appropriate eye movements.

Regardless of the above challenges there are various reasons and advantages that encourage HCI researchers to explore gaze as an input device. There are some situations that do not allow the user to make use of his or her hands, for example disabilities such as motor disorders like paraplegia. There are also several work situations where the hands are engaged, e.g. in a hospital operating room or a dental operatory (*Grätzelet al, 2004*). Gaze input interface could make the use of a computer interface a lot easier in daily functions; for example icon selection, moving an object (visually dragging and dropping the object), scrolling, maneuvering menu commands etc. Another potential advantage of gaze input is that eye movements are much quicker than other parts of the body and this would be excellent in accomplishing tasks at a faster rate. Gaze can also reduce fatigue and potential injury (e.g. carpal tunnel syndrome) caused due to using other input devices like a keyboard ormouse (Zhai, Morimoto, 1999).

Gaze input also presents a benefit of ease of operation, as with advanced gaze interfaces one may need little or no formal training to maneuver the objects around on the screen. The eye is also much faster than any high-speed cursor-positioning tool. This is evident after interpreting several eye tracking results that clearly reveal user's center of attention. The user changes his or her focus of attention and then directs the cursor to the new focus when pointing with a commonly used device such as a mouse,. However, with gaze as an input, every change of focus will be available as a pointing command to the computer. Thiscan also becomea disadvantage as well, as lack of control of eye movement is often an unintentional act. It is relatively hard to control eye position consciously and accurately at all times. The eyes continually zoom from one point to another and it is not desirable for each such move to initiate a computer command. It may be a challenge to operate user interface widgets like hyperlinks and scrollbars (*Jacob, 1991*), however, when combined with other input modalities, these side effects may be mitigated.Additionally, the 'dynamic' nature of gaze makes it impossible to discriminate when the users are giving a command by their gaze from times where they are just glancing around and resting their gaze. This makes it difficult for a system to predict when to engage gaze as an input device and when to disengage (*Neilsen, 1993*). Many eye movements such as closing the eyes or using blinks as an indication are unacceptable, because it digresses from the natural eye movements if done consciously. It would be very hard for the user to have an eye movement-based dialogue by thinking about when to blink or close the eye. Existing input devices have built-in buttons that serve as efficient commands for a system, but the gaze interfaces currently lack a parallel analogue of built-in buttons.

For the present, eye tracking is useful as a tool to help understand human behavior. Current eye tracking technology is evolving, however it lacks practicality compared to the existing manual input devices; further research is desired to refine the gaze interface design. Another line of research may be to seek answers in a direction where gaze is not an explicit command but an implicit one, where the computer observes the user's gaze and derives appropriate conclusions of the user's interest.

Brain-Computer Interfaces

As impressive as input devices controlled by a simple touch, sound or gaze are, researchers are pushing the limits of human input even further by completely removing physical movement from the equation. These researchers envision brain-computer (or brain-machine) interfaces as the next step in human input into cyberspace and are beginning to create systems that allow users to control computers with a mere thought.

Serious research into brain-computer interfaces first began in the 1960s, but progress was slow. Technology limited the research on two fronts: first, capturing strong, relatively noiseless brain waves was a difficult (and often invasive and intercranial) task, and second, processing any signals that did happen to get captured was painfully slow with early electronics (Thorpe, Oorschot & Somayaji, 2005). Progress in the field began to pick up when Chapin et. al. (1999) successfully trained rats to "position a robot arm to obtain water by pressing a lever" via a brain-machine interface, thus providing the first solid demonstration that brain waves could be converted into a workable input method (Lebedev & Nicolelis, 2006). Brain-computer interfaces have progressed rapidly since that demonstration, enabled by non-invasive brain wave capture methods like electroencephalograms and fast processors to analyze the signals in real time (Lebedev & Nicolelis, 2006).

Despite the surge in brain computer-interface research in the past 10 years, current input devices are still relatively crude compared to some of their physical counterparts. Systems considered advanced can, for instance, allow users to "type" at a rate of about five letters per minute (by monitoring spikes in brain activity in response to rare stimuli) or give users the ability to "steer a cursor by oscillatory brain activity into one of two or four possible targets" (Krepki et. al., 2007). Other systems offer slightly more fine grained control, giving users the ability to direct mobile robots in a physical environment (Millan, 2006) or control a basic, Pac-Man-like computer game (Krepki et. al., 2007). While all of these systems operate in real time, they also all require significant extra hardware such as hat-like "brain caps" to capture the brain activity needed to drive the interface.

Even with current limitations of brain-computer interfaces, work is being done to develop them into commercially viable products. Much of the work being done in this area comes from private firms who hope to market the finished systems to users who have lost physical control of their bodies--users for whom *any* autonomous control of a computer would be an improvement. Cyberkinetics Neurotechnology Systems, for instance, is currently running a pilot clinical trial of its "BrainGate Neural Interface System" (Cyberkinetics Neurotechnology Systems, 2007). The system is driven by a small sensor implanted directly onto the motor cortex of a patient lacking the ability to control a computer via normal input devices. The sensor passes brain wave activity to an analyzer, which then translates the signals into "cursor movements, offering the user an alternate 'BrainGate pathway' to control a computer with thought, just as individuals who have the ability to move their hands use a mouse."[

Researchers are working to make brain-computer interfaces useful input devices for those without physical disabilities as well. Gaming researchers, for instance, are investigating the usefulness of such interfaces for play. Not only are these researchers beginning to develop systems where users can control their in-game avatar through thought alone, they are also taking steps toward "affective gaming," where games can react and adapt to users' emotions--whether the user is conscious of such emotions or not (Nijholt & Tan, 2007). Brain-computer interface developers are also exploring ways to authenticate users through brain activity, such as "pass-thoughts," passwords composed of seemingly random brain activity unique to each individual when "transmitting a thought" (Thorpe et. al., 2005).

Any such future applications of brain-computer interfaces, though, will likely remain unacceptable to most users until means of detecting brain signals less obvious than "brain caps" and less intrusive than implanted sensors become available. Strides are being made in this area as well. Recently, researchers at Tufts University, for instance, demonstrated a system that measured brain waves using functional near-infrared spectroscopy (Hirshfield et. al., 2007). Brain waves were detected using small external sensors, held in place by a relatively svelte headband, to detect "hemoglobin concentration and tissue oxygenation" in the brain. The researchers noted especially that the "equipment places no unreasonable restrictions on a subject using an interactive system" and was both portable and capable of wireless signal transmission.

Implicit I/O

Biometrics and Sensors On/In Body

In addition to explicit input methods that draw on the intentional use of body and mind, there are also a number of implicit methods that make use of somewhat less conscious body processes. Biometrics, the use of biological data for input and identification, employs various means of sensing the body, providing a wide array of forums for human-computer interaction. For instance, studies have explored the use of hands, ears, fingerprints, gait, body odor, ear, skin reflectance, lip motion and DNA as ways of uniquely identifying individuals. Iris and retinal scans, face and voice recognition, and even thermograms have been used for similar purposes (Bolle, 2004; Pankanti et al, 2000). Other studies have explored the monitoring of heart rate and galvanic skin response for purposes, ranging from medical and psychological uses to employment in commercial and entertainment contexts (Moore & Umang, 2003; Isbister et al, 2006; Sakurazawa et al, 2003). In a number of ways, then, biometrics has opened up new avenues for human input into cyberspace.

One area where biometrics has made a significant impact is the field of security. As Pankanti et al (2000) point out, using the body to ensure positive identification has a number of advantages over more commonly used methods. After all, the body is not easily forged, stolen or forgotten as a set of keys or a simple password. Though fingerprint analysis has been fairly common for decades, biometric identification has moved beyond its initial stage, where it was predominantly used in criminology, to increased commercial usage. Identification by biometrics has been hailed as a good solution for an increasingly mobile populous, with companies like *Clear* providing iris scans and fingerprint analysis for a fee in exchange for expedited passage through airport security. However, implementation in this field faces a number of challenges. Some of these challenges, like cost and the lack of biometric standards, are rapidly being overcome. However, the possibility for inaccuracy or various forms of attacks on the technology also pose ongoing difficulties, though the co-ordination of multiple biometrics can help alleviate both problems. Perhaps the most intractable battle going forth will take place over issues of privacy, as disagreements continue to separate advocates of biometric identification and those who worry that it will lead to an erosion of personal freedoms (Bolle, 2004).

Another arena where biometrics has attracted a great deal of interest is the medical field. For instance, biometrics has been profitably used to train patients to monitor and control their own vital signs. This monitoring gives valuable feedback that can help patients quantitatively see the effects that various factors can have on their bodies, from breathing to exercise. The potential of such feedback for improving physiology has found wide recognition in the medical community, with studies exploring its use as a memory aid, in telemedicine, and in the treatment of terminal diseases, as well as helping to cure obesity, anxiety, addiction and a variety of other medical problems (Papazoglou , et al., 2006, Sharry , et al., 2003). Psychiatry has also gained from advances in the field, with biometrics alternately used in the treatment of patients suffering from post-traumatic stress disorder, disassociative disorders, depression as well as other problems (Simeon , et al., 2007, Szeszko , et al., 2005). Researchers have increasingly sought to recognize methods, such as skin conductance, that might accurately reflect and even quantify affective states.

The use of biometrics to measure emotional response has drawn the attention of commercial interests as well. One researcher patented a process for monitoring the stress level of customers at their computer so that companies and their software will be better able to respond to their needs. Another patent evaluates whether or not a price is set to the correct level by gauging customer satisfaction with galvanic skin response. Video game companies like the Wild Divine Project integrate biofeedback into their products, creating games that mix parts game play and meditation, demanding an unusual measure of physiological input on the part of their users. And the growth in biometrics is already spurring an industry that provides ever more accurate and non-intrusive measuring devices, such as watches and shoes that contain sensors to study skin conductance, or earrings that measure blood volume pulse (Beckhaus & Kruijff, 2004).

Because of its variety and scope, biometrics has impacted a diverse variety of fields, from medicine to education and from security to gaming. Whether its application is subtle, as in user studies that lead to new interfaces for a web site, or obvious, like an iris scan that help users pass through security, it is a field that will likely see continued growth in the coming years. For good or ill, its impact on daily life is likely to be significant.

Sensors In the Environment

In the late 1980s, Mark Weiser and his colleagues at Xerox PARC (Weiser, 1991;Weiser, Gold, and Brown, 1999) envisioned a future for computing in which computers would “disappear” and become an integral part of the human environment. Interactions with “ubiquitous computing” environments would then be more natural and demand less attention. As computers disappeared into the background, participant-users would be less distracted by the mechanics of computing and focus more intently on the task at hand. Since 1991, advances in mobile computing and sensor technologies have supported research and development of smart environments to support the activities of everyday life. At the heart of this movement is the concept that interactions with computers in the environment should be natural and implicit. Support of implicit interactions can be supported through biometrics (discussed above) or through sensors in the environment. This review discusses the attributes required of sensor-based systems in support of implicit input and research challenges in the field.

For decades, man-made sensors connected to actuators have aided us by automating simple tasks. Pressure mats that activate automatic doors and motion detectors that turn on the lights when we pull into the driveway are two examples. The simple sensor-actuator combination can be taken to a new level by adding a bit of "machine intelligence." A license plate reader at a gated parking lot uses video technologies to "monitor" the license plates of automobiles that enter the gate paired with recognition algorithms and a database of clients' license plates. When a match is found, the driver is allowed to proceed into the parking lot. Demonstrating a stimulus-response activity, these sensors can be considered somewhat primitive, but they possess two important traits of implicit interaction and invisibility.

Beyond Stimulus-Response: The Smart Environment

Cook and Das (2005) define smart environments as "a small world where all kinds of smart devices are continuously working to make inhabitants' lives more comfortable (p. 3). Common "smart environment" applications include the "automated home," the efficient office, and smart vehicles. These "smart environment" applications are made up of networks of sensors, which areconnected "intelligent" devices that have decision-making capabilities. Early sensors like the ActiveBadge described by Want , et al. (1992) used infrared technologies to relay a user's location information to a network of receivers throughout an office building. Today's Radio Frequency Identification (RFID) based systems rely on similar techniques. Schmidt (2000, as cited in Loke ,2007) identifies a wide variety of environmental sensor types including those that detect light and vision, audio, movement, location, and proximity. Engineering challenges include networking technologies and standards, operating systems for embedded sensor technologies, and software for inferring meaning from gathered data and making decisions (Cook & Das, 2005).

Challenges to sensor-laden environments

Challenges to developers of applications of sensor networks to support human pursuits at work, play, and daily life include the need to support implicit interactions and be context-aware, as well as the need for information fusion techniques to support data integration and eventual decision making.

Implicit Interaction. Most of our interactions with computer interfaces are explicit and intentional. For instance the act of writing and sending an email requires both a focus on the computer screen, the keyboard, and a mouse and then secondly the intentional interaction with the email application through typing and clicking the send button. Implicit interactions with computing systems require neither a shift of focus from our primary task nor conscious interactions with the system. Schmidt, Kranz, and Holleis (2005) define implicit interaction as "the interaction of a human with the environment and with artifacts, which is aimed to accomplish a goal. Within this process the system acquires implicit input from the user and may present implicit output to the user" (p. 148). Schmidt , et al. further define implicit input as "actions and behaviors of humans that are done to achieve a goal and are not primarily regarded as interaction with a computer, but captured, recognized and interpreted by a computer system as input" (p. 148). Abowd and Mynatt (2005) describe implicit input as "natural interactions with the physical environment [that] provide sufficient input to a variety of attendant services, without any further user intervention" (p.155). Examples of implicit input include geographic positioning systems embedded in vehicles which track vehicle movement without the driver's knowledge, a smart home's use of motion sensors to track individuals in the home, or pen-based input systems which extend the natural act of writing by hand to digital input. Sensors provide powerful ways to detect state changes in people or environments; however, interpreting these changes requires rules or intelligence rooted in context.

Context-Awareness. A major challenge to researchers attempting to integrate implicit input from environmental sensors with explicit input or environmental data is constructing context-awareness. Dey (2001) defines context as "any information that can be used to characterize the situation of an entity" where an "entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves" (p. 5). While early sensor-based assistive applications tracked the location of a user as context (Want, 1992), Abowd and Mynatt (2000) note that other dimensions of a situation are equally relevant when constructing context-aware applications including the presence of other users, what the user is doing, why the user is performing the task, and the passage of time during a given activity. Context-aware systems should also be aware of past history and be potentially able to learn what is typical for a given situation. Consider an application to monitor the homebound, which alerts a loved one when unusual behavior is observed. Systems should be aware of what is typical-- such as "patient *A* usually sleeps through the night, but patient *B* typically stays up late and sleeps late"-- before sending an alert (Helal, Mann, and Lee, 2005, p. 366).

These requirements place extraordinary demands on sensor-based systems, needing complex processing to infer context before making a decision and triggering an action. Dey and Mankoff (2005) make the salient point that because applications may not be able to construct context from sensed data with 100% reliability, context ambiguity arises where the system's interpretation of a user's actions are not aligned with the user's intent. Dey and Mankoff built a system to disambiguate context through mediation with the user to ensure the system does not take improper actions based on misinterpretation of ambiguous context. Addtionally, Kawsar, , et al. (2005) created and evaluated a smart environment application that demonstrated implicit interaction and context-awareness through sentient artifacts. In their application, everyday objects like chairs, toothbrushes, and bathroom mirrors were connected to sensors to detect *state of use.* For instance their AwareMirror application includes a sensing mirror as well as an RFID- and accelerometer-enhanced toothbrush. The sensing mirror detects a person's presence and displays pertinent information about the weather and commuting conditions to the user only when they are brushing their teeth. The system infers that when the user is brushing her teeth, she is preparing to leave for work and the display of weather information makes sense in that context.

Information Fusion. Wireless sensor networks (WSNs) used for military purposes, remote sensing, and monitoring must gather, integrate, and make inferences from multiple sources of data. Information fusion techniques have been developed to aid in this process. Developers of smart environment and ubiquitous computing applications face similar challenges in gathering, interpreting, and making sense of multiple sources of implicit and explicit input from human activities and behaviors. Information fusion techniques hold promise to support these needs.

Wireless sensor networks consist of energy efficient embedded microcontrollers gathering data, often of different types, and low-power wireless communications technologies which support transmitting the sensed data to central servers for storage, processing, interpretation, or decision. Distributed short-range sensor networks are capable of detecting local phenomena up close, continuously, often with real-time transmission to a central server allowing the monitoring of phenomena that would otherwise be invisible. Wireless sensor networks in the environment can be entirely passive or augmented with actuators that can alter the network's configuration or affect the environment itself (Elson & Estrin, 2004, p.7). Applications of wireless sensor networks include target tracking, environmental monitoring, robotics, military applications, (Nakamura , et al., 2007, p. 6) and smart environments.

Information Fusion in its broadest definition "encompasses the theory, techniques, and tools intended to leverage the synergy in the information collected by multiple sources with the purpose of making the resulting decision or action measurably better than if the individual information sources were used independently" (Nakamura , et al., 2007, p.4). Information fusion techniques support the three critical properties of wireless sensor networks: cooperation, redundancy, and complementarity (Nakamura , et al., 2007, p. 5).

Cooperative information fusion is the process of combining two independent measures of a phenomenon producing more complex and realistic data. *Redundancy* is supported in several ways. Individual sensors may fail due to hostile environmental conditions, power failure, or natural obstructions so the coverage of sensors in a network must overlap to provide redundancy. Should multiple sensors detect identical measures at a given location, that measure can be given greater confidence. This processing is typically handled in-network to reduce the amount of data returned to the central server and conserving scarce energy required for wireless communications. Information fusion techniques are also used to detect when a sensor fails and can support network recovery to fill in the gaps left by the failed sensor. *Complementarity* addresses how the different sensor inputs influence interpretation. On occasion aspects of a phenomenon may be more richly monitored using multiple senses (light and temperature or movement and sound). So a sensor network may of necessity consist of sensors gathering different types of sense data and communicating those data to the central server. To achieve complementarity, however, information fusion techniques must be used to integrate these multiple sense data and subsequent inference or decision making as needed (Nakamura , et al., 2007; Elson & Estrin, 2004).

The sophistication of a wireless sensor network can be classified by its inputs and outputs. At their most fundamental, WSNs can accept raw data as their input and then generate data, possibly scrubbed or refined, as an output. WSNs may also convert raw data to an abstraction (such as shape, texture, or position). At their highest levels of sophistication, WSNs can ingest raw data or abstractions and use inference techniques to support decisions (Nakamura , et al., 2007; Dasarathy, 1997). Fusion techniques are required to support each of these input-output paths. *Inference* "refers to the transition from one likely true proposition to another, whose truth is believed to result from the previous one" (Nakamura , et al., 2007, p.9). Nakamura , et al. report that typical methods are based on probability theory (Bayesian Inference, Dempster-Shafer inference, Fuzzy Logic), supervised learning techniques (Neural Networks), abductive reasoning, and abstraction (Semantic information Fusion) (p. 9 - 15).

We are beginning to see applications of information fusion techniques for integrating implicit input from multiple sensors in HCI research. Lewis and Powers (2004) demonstrate research evaluating an audio-visual speech recognition system. The concept is to improve recognition by combining explicit input (speech) and implicit input in the form of visible speech gestures or visemes through the use of redundant and complementarity fusion techniques. Paleari and Lisetti (2006) present a framework for accurate machine interpretation of affective or emotional state through detection and fusion of paralanguage (voice tone, voice volume, expressions, body language). Bernardin and

Stiefelhagen (2007) developed and evaluated an unobtrusive smart environment application for tracking and detecting individuals in a room by fusing person tracking, speech recognition, and facial identification data. The work of Kawsar , et al. (2005) using sentient artifacts also relied on information fusion techniques.

While sensors have been used to detect environmental conditions in remote sensing and military applications, the use of environmental sensors for purposes of human input to computing systems is still in its naissance. Sensing the data in some ways is the easy part. The difficult problems of building systems able to construct context, infer meaning from gathered data, and make decisions that are useful and unobtrusive are far more difficult to resolve.

MULTIPLE INPUTS: COORDINATED vs INDEPENDENT

How These Tools and Styles Empower

As noted in several of the sections above, input devices are often combined with other devices to disambiguate intention or provide alternatives under special conditions such as human disability or severe environmental constraints. In addition to these advantages of combining multiple devices and techniques, combinations may add new capabilities, empowering people to perform beyond their natural capabilities. By their very nature, computer interfaces are tools of empowerment, as they serve as the sole means of interaction between the user and the mess of bits and micro-circuitry that constitute cyberspace. In doing so, they act as an enabler, granting access to an entirely separate environment in which the user would not otherwise be able to participate. Thus any interface technology, regardless of how poorly conceived it may be, can be considered an empowering technology for the user.

However, this proclamation is derived from what is only a cursory look at the tools and methods employed in developing interface systems for human-computer interaction. Looking at the technology more closely, one can denote specific traits of any number of systems that help showcase their ability to empower the user. While it is true that early computers operated quite sufficiently using punch cards and print-outs, one cannot deny that the introduction of devices such as the keyboard, mouse, and monitor has greatly increased the ease with which we interact in cyberspace, as well as our productivity whilst there. Yet even the freedom afforded by these tools is not without its limitations. Keyboards only track two types of motion: presses and releases. The same holds true for mice, though they add new degrees of motion. Monitors manage to feed merely one degree of perception. This leaves only a limited range of input and output that can be successfully communicated between the user and the system.

This is where other technologies come into play. Advanced technologies such as audio tools (speakers and microphones), force-feedback devices, and gesture and brainwave pattern recognition serve to accentuate the interaction between the user and the device, expanding the range of input far beyond simple presses and releases. These elements work in tandem with other systems to allow for more complex command mappings and bindings, as well as richer feedback from the machine, both of which promise to enhance the experience of the user by giving them more modes of communication with cyberspace. Rather than simply viewing a static print out of an interaction instance, the experience is enhanced with motion, color, and sound. Instead of taking single mouse-clicks as user input, a system can queue and process complex gestures, allowing for thoughts and ideas to be communicated between both the human and the computer with an enormous amount of detail. Greater detail leads to more options, and more options mean more control on the part of the user.

Gajos et al (2002) noted that the goal in making computing a pervasive part of our culture was to "change fundamentally the interaction between humans and computers to one in which computers adapt to human forms of discourse, rather than the other way around.". This is precisely what the tools discussed in this paper aim to do. By adapting the input methods of the machine to address the limitations of human discourse – voice, speech, sound, etc – we are empowering the user. Users have a choice in how they would prefer to interact with the computer system, and as such, they have power.

Perhaps the best example of the empowering nature of these technologies exists in their implementation for special-needs users. While standard interfaces such as the keyboard, mouse, and monitor have created a means through which the average person can interact with computing systems reasonably well, they are still of limited use by those with disabilities, particularly in the case of individuals with impaired motor or vision capabilities. For such users, having multiple interface options increases the likelihood that they will have a productive experience within the operating environment, by removing the limitations that might be associated with any one particular interface. Paralysis of the arms removes what might be considered the most efficient means of interaction with a computer. With speech-enabled computing, however, many of the limitations associated with the previous example are circumvented, thus bringing a new degree of control to an entirely new audience. Thus, the opportunities available to special-needs users are increased, which can have numerous positive side effects, such as increased learning capabilities in the case of students and greater range of functionality in the workplace (Ryba 1995).

Such tools are not limited to use by special-needs individuals, but rather anyone whose physical capacities are restricted in some way, for example, the temporarily injured, or those who are occupying one or more input sources with other challenges. Consider a scenario wherein a student is taking notes with a pen and paper, but at the same time, is able to navigate the document she is reading with optical input. Or, one in which a user is able to surf the web while performing another task - such as cooking or cleaning - using their voice. This opens up a whole new avenue of functionality for the user, by allowing them to still interact in cyberspace in a meaningful way, while still participating in other activities.

What these new input mechanisms are in effect doing is increasing human potential when interacting in computer environments, by providing new and varied means of communication between the user and the system. This effectively increases the ease with which a larger range of users can successfully function in our technologically dominated society.

SIDE EFFECTS: FATIQUE, ADDICTION, and INTERRUPTION

Although electronic devices such as computers have greatly improved many of our everyday tasks, human interactions with computers are also creating various negative side effects. Many of these effects were unknown or rare before the frequent use of these devices. Some of these side effects are physical, like repetitive stress injury and fatigue, and some are psychological like addiction and interruption.

Modern engineering has made tremendous advances in human-centered interface design and ergonomics, but even with the best design, spending many hours interacting with computers can take a toll on its users. Even with alternative input devices, tiredness and fatigue still arise. With the more traditional interfaces like keyboards, mice, touch pads, track points, joysticks and game controllers, fatigue is mostly localized in the fingers, hands and arms of the user. The element of fatigue is also present with alternative inputs, but it can take different forms. For example, using speech recognition systems, gaze or brainwave readers still fatigue the user's voice, eyes and minds. Because interactions lead to fatigue even with the best ergonomic interface design, many engineers strive to limit these side effects as much as possible (Zhai, 2004). Additionally, people are learning to use good work practices (e.g., breaks, exercises, furniture selection) to minimize these effects, which are a growing feature of public health campaigns.

Psychological side effects, however, are less easily controlled. While interface designers may be able to limit fatigue, there is little they can do to limit addiction. Addiction to technology use in general or to some specific device, software or environment has reached never before seen heights. Although some studies claim that “computer addiction” is really only an outlet “to counteract other deficiencies” (Griffiths, 2000), sadly, people die because of their psychological ties to their devices and games. Of course, the concept of addiction is not a new one, but technology addiction shows that even empowering technologies can have serious side effects.

Interruption is another psychological effect of existing and emerging technologies such as interactive devices (McFarlane, 1998). In our busy, fast-paced lives, we are constantly communicating, searching or simply entertaining ourselves by using all kinds of technological interfaces. Mobile devices and desktop applications are often the most disruptive, because phone calls and instant messages often interrupt other actions. These interruptions are becoming part of our everyday multitasking routines, and some of them may enhance productivity while others impede it. Interruption-guided multitasking allows us to respond to events more rapidly, but it also can prove disruptive to the original task that was being performed. Balancing these effects, using them for measurable effectiveness and allowing for controlled interruptions are areas that need to be explored even further.

CONCLUSION

This review considers techniques for input and some of the challenges the techniques and accompanying devices bring. Just as Jacob , et al., predicted almost 20 years ago, the trend is towards multiple devices and multiple styles. The challenges of mapping the increasingly powerful detections of human action to the underlying intentions remain the key research challenge for HCI researchers.

REFERENCES

Abowd, G. D. & Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Trans. ACM Transactions on Computer-Human Interactions.* 7, 1 (Mar. 2000), 29-58.

Abowd, G.D. & Mynatt, E. D. (2005). Designing for the human experience in smart environments in Cook, D.J, & Das, S.K. (eds.) *Smart Environments: Technologies, Protocols, and Applications.* Hoboken, NJ: John Wiley and Sons, (p. 153 - 174).

Baker, J. M., & Pugliese, P. J. (2006). From pipedreams to products, and promise! Proceedings of the 2006 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, New York, New York. 257-260.

Beckhaus, S. & Kruijff, E. (2004). Unconventional human computer interfaces. International Conference on Computer Graphics and Interactive Techniques. ACM SIGGRAPH 2004. Article No. 18. ACM Press: New York.

Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouvet, D., et al. (2007). Automatic speech recognition and speech variability: A review. Speech Communication, 49(10), 763-786.

Betts, B., Burlingame, D., Fischer, G., Foley, J., Green, M., Kasik, D., Kerr, S., Olsen, D., & Thomas, J. (1987). Goals and objectives for user interface software. *Computer Graphics*, *21*(2), 73-78.

Bernardin, K. & Stiefelhagen, R. (2007). Audio-visual multi-person tracking and identification for smart environments. In *Proceedings of the 15th international Conference on Multimedia* (Augsburg, Germany, September 25 - 29, 2007). MULTIMEDIA '07. ACM, New York, NY, 661-670.

Bodker, S. (2006). When second wave HCI meets third wave challenges. *NordiCHI 2006: Changing Roles*, *14-18*, 1-8.

Bolt, R. A. (1982). Eyes at the interface. In *Human Factors in Computer Systems. Gaithersburg, Maryland: ACM.*

Castel, F. (2002). Theory, theory on the wall.... *Communications of the ACM*, *45*(12), 25-26.

Bolle, Ruud. (2004). Guide to biometrics. New York: Springer.

Cechanowicz, Jared, Pourang Irani, & Sriram Subramanian. (2007). Augmenting the mouse with pressure sensitive input. Proceedings of the SIGCHI conference on Human factors in computing systems.

Cerney, M. & Vance, J.M. (2005). Gesture recognition in virtual environments: A review and framework for future development. Iowa State University Human Computer Interaction Technical Report.

Chapin, J.K., Moxon, K.A., Markowitz, R.S., & Nicolelis, M.L. (1999). Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature Neuroscience*, 2, 664–670.

Chen, T. (2001). Audiovisual speech processing. Signal Processing Magazine, IEEE, 18(1), 9

Cook, D.J, & Das, S.K. (eds.) (2005). *Smart Environments: Technologies, Protocols, and Applications*. Hoboken, NJ: John Wiley and Sons.

Cyberkinetics Neurotechnology Systems, Inc. (2007). BrainGate neural interface system. Retrieved 9/26, 2007, from http://www.cyberkineticsinc.com/content/medicalproducts/braingate.jsp

Da Rosa Junior, R., Marcus, F., Trevisan, D., Nedel, L. & Freitas, C. (2007). A multimodal medical sculptor. Lecture Notes in Computer Science 4663, 637-640.

Dasarathy, B. V. (1997). Sensor fusion potential exploitation-innovative architectures and illustrative applications. Proceedings of the IEEE 85, 1 (January), 24–38.

Deng, L., & Huang, X. (2004). Challenges in adopting speech recognition. Communications of the ACM, 47(1), 69-75.

Dey, A. (2001). Understanding and using context. *Personal Ubiquitous Computing*. 5, 1 (Jan. 2001), 4-7.

Dey, A. K. & Mankoff, J. (2005). Designing mediation for context-aware applications. *ACM Transactions on Computer-Human Interactions*. 12, 1 (Mar. 2005), 53-80.

ElAarag, H., & Schindler, L. (2006). A speech recognition and synthesis tool. ACM-SE 44: Proceedings of the 44th Annual Southeast Regional Conference, Melbourne, Florida. 45-49.

Elson, J. and Estrin, D. (2004). Sensor networks: a bridge to the physical world in Raghavendra, C. S., Sivalingam, K.M., and Znati, T. (eds.). (2004). *Wireless Sensor Networks*. Boston, MA : Kluwer Academic Publishers, p. 3 – 20.

Feied, C., Gillam, M., Wachs, J., Handler, J., Stern, H. & Smith, M. (2006). A real time gesture interface for hands-free control of electronic medical records. *AMIA 2006 Symposium Proceedings*. 920.

Fischer J. & Bartz D. (2005). Utilizing image guided surgery for user interaction in medical augmented reality. Technical Report WSI-2005-04, WSI/GRIS, University of Tbingen.

Foley, J., van Dam, A., Feiner, S., & Hughes, J. (1990). *Computer Graphics: Principles and Practice 2nd Ed.*. Reading, MA: Addison-Wesley Publishing Company, Inc..

French, Jean H. (2007). Beyond the tablet PC: using the tablet PC in a collaborative learning environment. J. Comput. Small Coll 23, no. 2: 84-89.

Frogpad. http://frogpad.com/ (accessed November 24, 2007).

Gajos, K., Fox, H., & Shrobe, H. (2002). End user empowerment in human centered pervasive computing. *Proceedings of Pervasive 2002*, 1-7.

Gomes, L. (2007). After years of effort, voice recognition is starting to work. Wall Street Journal - Eastern Edition, 249(8), B1-B1.

Grätzel, C., Grange, S., & Baur, C. (2004). A Non-contact mouse for surgeon-computer interaction. *Technology and Health Care*, 12, 45-257.

Griffiths, M. (2000). *Does Internet and Computer "Addiction" Exist? Some Case Study Evidence.* CyberPsychology & Behavior. 3(2): 211-218.

Hamilton, A. (2000). HCI metaphors framed in a wide-ranging look at metaphoric discourse. *ACM Journal of Computer Documentation*, *24*(4), 237-253.

Helal, A., Mann, W.C., & Lee, C. (2005). Assistive environments for individuals with special needs in Cook, D.J. and Das, S.K. (eds.) *Smart environments: technologies, protocols, and applications*. Hoboken, NJ : John Wiley. p. 361-383.

Hirshfield, L.M., Girouard, A., Solovey, E.T., Jacob, R.J.K., Sassaroli, A., Tong, Y., & Fantini, S. (2007). Human computer interaction and brain measurement using functional near-infrared spectroscopy. Poster session presented at the 20th annual ACM symposium on user interface software and technology, UIST '07, Newport, R.I.

Hutchinson, T., White, K.P., Martin, W., Reichert, K., & Frey, L. (1989). Human-computer interaction using eye-gaze input. *IEEE Transactions on systems, man, and cybernetics*. 19(6), 1527-1534.

Isbister, Katherine, Hook, Kristina, Sharp, Michael & Laaksolahti, Jarmo (2006). The sensual evaluation instrument: developing an affective evaluation tool. *Proceedings of the SIGCHI conference on Human Factors in computing systems*. Association for Computing Machinery Press: New York, p. 1163-1172.

Jacob, R.J.K. (1990). What you look at is what you get: Eye movement-based interaction techniques. *CHI 90 Proceedings*, 4, 11-18.

1. Jacob, R.J.K. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Trans. Information Systems*, 9, 2, 152-169.

Jaimes, A., & Sebe, N. (2005). Multimodal human computer interaction: A survey. In *IEEE International Workshop on Human Computer Interaction* in conjunction with ICCV, Beijing (China).

Kawsar, F., Fujinami, K., & Nakajima, T. (2005). Augmenting everyday life with sentient artefacts. In *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient intelligence: innovative Context-Aware Services: Usages and Technologies* (Grenoble, France, October 12 - 14, 2005). sOc-EUSAI '05, vol. 121. ACM, New York, NY, 141-146.

Kaptelinin, V., & Nardi, B. (2003). Post-cognitivist HCI: second wave theories. *CHI 2003: New Horizons*, 692-693.

Kim, Seoktae, Minjung Sohn, Jinhee Pak, & Woohun Lee. (2006). One-key keyboard: a very small QWERTY keyboard supporting text entry for wearable computing. Proceedings of the 20th conference of the computer-human interaction special interest group (CHISIG) of Australia on Computer-human interaction: design: activities, artefacts and environments.

Krepki, R., Curio, G., Blankertz, B., & Muller, K. (2007). Berlin Brain-Computer Interface-The HCI communication channel for discovery. *International Journal of Human-Computers Studies*, *65*(5), 460-477.

LaViola, J., (1999). A survey of hand posture and gesture recognition techniques and technology. Technical Report CS-99-11, Department of Computer Science, Brown University.

Lebedev, M., & Nicolelis, M. L. (2006). Brain–machine interfaces: past, present and future. *Trends in Neurosciences*, *29*(9), 536-546.

Lee, S. K., William Buxton, & K. C. Smith. (1985). A multi-touch three-dimensional touch-sensitive tablet. *Proceedings of the SIGCHI conference on Human factors in computing systems*.

Loke, S. (2007). *Context-Aware pervasive systems: Architectures for a New Breed of Applications*. New York: Auerbach Publications, Taylor and Francis Group.

Lewis, T. W., & Powers, D. M. W. (2002). Audio-visual speech recognition using red exclusion and neural networks. ACSC '02: Proceedings of the Twenty-Fifth Australasian Conference on Computer Science, Melbourne, Victoria, Australia. 149-156.

Lewis, T. W. & Powers, D. M. (2004). Sensor fusion weighting measures in Audio-Visual Speech Recognition. In *Proceedings of the 27th Australasian Conference on Computer Science - Volume 26* (Dunedin, New Zealand). Estivill-Castro, Ed. ACM International Conference Proceeding Series, vol. 56. Australian Computer Society, Darlinghurst, Australia, 305-314.

Massaro, D. W., Cohen, M. M., Beskow, J., & Cole, R. A. (2000). Developing and evaluating conversational agents., 287-318.

McFarlane, D. C. (1998). *Interruption of People in Human-Computer Interaction*. Dissertation submitted to The School of Engineering and Applied Science of The George Washington University.

Millan, J. d. R. (2006). Brain-computer interaction. UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology, Montreux, Switzerland. 277-278.

Microsoft Surface. http://www.microsoft.com/surface/ (accessed November 24, 2007).

Mitra, S. & Acharya, T. (2007,). Gesture recognition: A survey. Systems, man and cybernetics, part c: Applications and reviews, *IEEE Transactions* on , vol.37, no.3, pp.311-324.

Moore, M. & Dua U. (2003). A galvanic skin response interface for people with severe motor disabilities. ACM SIGACCESS Conference on Assistive Technologies, *Proceedings of the 6th International ACM SIGACCESS conference on Computers and Accessibility*. Association for Computing Machinery Press: New York, p. 48-54.

Nakamura, E.F., Loureiro, A.F., & Frery. (2007). Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys*, 39(3),Article 9, p. 1-55.

Nielsen, J., (1993). Non command user interfaces. *Communications of the ACM*, 36(4), 83-99.

Nijholt, A., & Tan, D. (2007). Playing with your brain: brain-computer interfaces and games. *ACE '07: Proceedings of the international conference on Advances in computer entertainment technology*, Salzburg, Austria. 305-306.

Norte, S.& Fernando, G.. (2007). A virtual logo keyboard for people with motor disabilities. *Proceedings of the 12th annual SIGCSE conference on Innovation and technology in computer science education*.

Optimus Maximus Keyboard. http://www.artlebedev.com/everything/optimus/ (accessed November 24, 2007).

Pankanti, S., Bolle, R. & Jain, A. (2000) Guest editors' introduction: biometrics-the future of identification, *Computer*, *33*, p. 46-49.

Paleari, M. & Lisetti, C. L. (2006). Toward multimodal fusion of affective cues. In *Proceedings of the 1st ACM international Workshop on Human-Centered Multimedia* (Santa Barbara, California, USA, October 27 - 27, 2006). HCM '06. ACM, New York, NY, 99-108.

Papazoglou, , et al. (2006). Evaluation of a multisensor armband in estimating energy expenditure in obese individuals. *Obesity*. 14:2217-2223.

Pei J., Huosheng H., Tao L., & Kui Y. (2007). Head gesture recognition for hands-free control of an intelligent wheelchair. *Industrial Robot: An International Journal*, vol 34, no. 1, pp.60-68.

Potamianos, G. (2006). Audio-visual speech processing: Progress and challenges. VisHCI '06: Proceedings of the HCSNet Workshop on use of Vision in Human-Computer Interaction, Canberra, Australia. 5-5.

Ryba, K. & Selby, L. (1995). Computers empower students with special needs. *Educational Leadership*. 82-84.

Sakurazawa, , et al. (2003). A Ccomputer game using galvanic skin response. A galvanic skin response interface for people with severe motor disabilities. ACM SIGACCESS Conference on Assistive Technologies, *Proceedings of the 2nd International ACM SIGACCESS conference on Computers and Accessibility*. Association for Computing Machinery Press: New York, p. 1-3.

Scolari, C. (2001). Towards a semio-cognitive theory of human-computer interaction. *CHI 2001*, 31 March- 5 April, 85-86.

Schmidt, A., Kranz, M., & Holleis, P. (2005). Interacting with the ubiquitous computer: towards embedding interaction. In *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient intelligence: innovative Context-Aware Services: Usages and Technologies* (Grenoble, France, October 12 - 14, 2005). sOc-EUSAI '05, vol. 121. ACM, New York, NY, 147-152.

Seow, S. (2005). Information theoretic models of HCI: a comparison of the Hick-Hyman Law and Fitts' Law. *Human-Computer Interaction*, *20*, 315-352.

Sharry, John, McDermott, Matt & Condron, Jim (2003) 'Relax to Win' - Treating children with anxiety problems with a biofeedback video game.

Shneiderman, B., & Plaisant, C. (2005). *Designing the User Interface 4th Ed. Strategies for Effective Human-Computer Interaction*. Boston: Pearson Education, Inc..

Simeon, D. , et al. (2007). Hypothalamic-pituitary-adrenal axis function in dissociative disorders, post-traumatic stress disorder, and healthy volunteers. *Biological Psychiatry*, *61*, (8), p. 966-973.

Sporka, A. J., Kurniawan, S. H., Mahmud, M., & Slav\'{\i}k, P. (2006). Non-speech input and speech recognition for real-time control of computer games. Assets '06: Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility, Portland, Oregon, USA. 213-220.

Srinivasan, S., & Vergo, J. (1998). Object oriented reuse: Experience in developing a framework for speech recognition applications. ICSE '98: Proceedings of the 20th International Conference on Software Engineering, Kyoto, Japan. 322-330.

Sutcliffe, A. (2000). On the effective use and reuse of HCI knowledge. *ACM Transactions on Computer-Human Interaction*, *7*(2), 197-221.

Szeszko, PR, , et al. (2005). White matter abnormalities in obsessive-compulsive disorder: A diffusion tensor imaging study. *Archives of General Psychiatry*. 62:782-790.

Tani, B., Rafael, M., & Von Wangenheim, A. (2007). A gesture interface for radiological workstations. *IEEE International Symposium on Computer-Based Medical Systems*, 27-32.

Thorpe, J., Oorschot, P. C. v., & Somayaji, A. (2005). Pass-thoughts: authenticating with our minds. NSPW '05: Proceedings of the 2005 workshop on New security paradigms, Lake Arrowhead, California. 45-56.

Wachs, J., Stern, H., Edan, Y., Gillam, M., Feied, C., Smith, M. & Handler, J. (2005). A real-timehand gesture interface for medical visualization applications. In *The 10th Online World Conference on Soft Computing in Industrial Applications*.

Wachs, J., Stern, H., Edan, Y., Gillam, M., Feied, C., Smith, M. & Handler, J. (2007). Gestix: a doctor-computer sterile gesture interface for dynamic environments. Soft Computing in Industrial Applications: Recent and Emerging Methods and Techniques: 30-39. Springer-Verlag, NY.

Want, R., Hopper, A., Falcão, V., & Gibbons, J. (1992). The active badge location system. *ACM Trans. Inf. Syst.* 10, 1 (Jan. 1992), 91-102.

Weiser, M. (1991). The computer for the 21st century. *Scientific American 265*, 3 (Sept.), 94–104.

Weiser, M., Gold, R., & Brown, J.S. (1991). The origins of ubiquitous computing research at PARC in the late 1980s. *Pervasive Computing*. 38(4), 1999.

Zhai, S. & Morimoto, C.S. (1999). Manual and gaze input cascaded (mAGIC) pointing. *CHI* 99, 5, 15-20.

Zhai, S. (2004). *The Computer Mouse and Related Input Devices*. Berkshire Encyclopedia of Human-Computer Interaction.