CONTENTS

www.vtucs.com

ABSTRACT

Is it possible to create a computer, which can interact with us as we interact each other? For example imagine in a fine morning you walk on to your computer room and switch on your computer, and then it tells you "Hey friend, good morning you seem to be a bad mood today. And then it opens your mail box and shows you some of the mails and tries to cheer you. It seems to be a fiction, but it will be the life lead by "BLUE EYES" in the very near future.

The basic idea behind this technology is to give the computer the human power. We all have some perceptual abilities. That is we can understand each others feelings. For example we can understand ones emotional state by analyzing his facial expression. If we add these perceptual abilities of human to computers would enable computers to work together with human beings as intimate partners. The "BLUE EYES" technology aims at creating computational machines that have perceptual and sensory ability like those of human beings.

www.vtucs.com

1. INTRODUCTION

Imagine yourself in a world where humans interact with computers. You are sitting in front of your personal computer that can listen, talk, or even scream aloud. It has the ability to gather information about you and interact with you through special techniques like facial recognition, speech recognition, etc. It can even understand your emotions at the touch of the mouse. It verifies your identity, feels your presents, and starts interacting with you. You ask the computer to dial to your friend at his office. It realizes the urgency of the situation through the mouse, dials your friend at his office, and establishes a connection.

Human cognition depends primarily on the ability to perceive, interpret, and integrate audio-visuals and sensoring information. Adding extraordinary perceptual abilities to computers would enable computers to work together with human beings as intimate partners. Researchers are attempting to add more capabilities to computers that will allow them to interact like humans, recognize human presents, talk, listen, or even guess their feelings.

The BLUE EYES technology aims at creating computational machines that have perceptual and sensory ability like those of human beings. It uses non-obtrusige sensing method, employing most modern video cameras and microphones to identifies the users actions through the use of imparted sensory abilities. The machine can understand what a user wants, where he is looking at, and even realize his physical or emotional states.

www.vtucs.com

2. EMOTION MOUSE

One goal of human computer interaction (HCI) is to make an adaptive, smart computer system. This type of project could possibly include gesture recognition, facial recognition, eye tracking, speech recognition, etc. Another non-invasive way to obtain information about a person is through touch. People use their computers to obtain, store and manipulate data using their computer. In order to start creating smart computers, the computer must start gaining information about the user. Our proposed method for gaining user information through touch is via a computer input device, the mouse. From the physiological data obtained from the user, an emotional state may be determined which would then be related to the task the user is currently doing on the computer. Over a period of time, a user model will be built in order to gain a sense of the user's personality. The scope of the project is to have the computer adapt to the user in order to create a better working environment where the user is more productive. The first steps towards realizing this goal are described here.

2.1 EMOTION AND COMPUTING

Rosalind Picard (1997) describes why emotions are important to the computing community. There are two aspects of affective computing: giving the computer the ability to detect emotions and giving the computer the ability to express emotions. Not only are emotions crucial for rational decision making as Picard describes, but emotion detection is an important step to an adaptive computer system. An adaptive, smart computer system has been driving our efforts to detect a person’s emotional state. An important element of incorporating emotion into computing is for productivity for a computer user. A study (Dryer & Horowitz, 1997) has shown that people with personalities that are similar or complement each other collaborate well. Dryer (1999) has also shown that people view their computer as having a personality. For these reasons, it is important to develop computers which can work well with its user.

www.vtucs.com

By matching a person’s emotional state and the context of the expressed emotion, over a period of time the person’s personality is being exhibited. Therefore, by giving the computer a longitudinal understanding of the emotional state of its user, the computer could adapt a working style which fits with its user’s personality. The result of this collaboration could increase productivity for the user. One way of gaining information from a user non-intrusively is by video. Cameras have been used to detect a person’s emotional state (Johnson, 1999). We have explored gaining information through touch. One obvious place to put sensors is on the mouse. Through observing normal computer usage (creating and editing documents and surfing the web), people spend approximately 1/3 of their total computer time touching their input device. Because of the incredible amount of time spent touching an input device, we will explore the possibility of detecting emotion through touch.

2.2 THEORY

Based on Paul Ekman’s facial expression work, we see a correlation between a person’s emotional state and a person’s physiological measurements. Selected works from Ekman and others on measuring facial behaviors describe Ekman’s Facial Action Coding System (Ekman and Rosenberg, 1997). One of his experiments involved participants attached to devices to record certain measurements including pulse, galvanic skin response (GSR), temperature, somatic movement and blood pressure. He then recorded the measurements as the participants were instructed to mimic facial expressions which corresponded to the six basic emotions. He defined the six basic emotions as anger, fear, sadness, disgust, joy and surprise. From this work, Dryer (1993) determined how physiological measures could be used to distinguish various emotional states.

www.vtucs.com

Six participants were trained to exhibit the facial expressions of the six basic emotions. While each participant exhibited these expressions, the physiological changes associated with affect were assessed. The measures taken were GSR, heart rate, skin temperature and general somatic activity (GSA). These data were then subject to two analyses. For the first analysis, a multidimensional scaling (MDS) procedure was used to determine the dimensionality of the data. This analysis suggested that the physiological similarities and dissimilarities of the six emotional states fit within a four dimensional model. For the second analysis, a discriminant function analysis was used to determine the mathematic functions that would distinguish the six emotional states. This analysis suggested that all four physiological variables made significant, nonredundant contributions to the functions that distinguish the six states. Moreover, these analyses indicate that these four physiological measures are sufficient to determine reliably a person's specific emotional state. Because of our need to incorporate these measurements into a small, non-intrusive form, we will explore taking these measurements from the hand. The amount of conductivity of the skin is best taken from the fingers. However, the other measures may not be as obvious or robust. We hypothesize that changes in the temperature of the finger are reliable for prediction of emotion. We also hypothesize the GSA can be measured by change in movement in the computer mouse. Our efforts to develop a robust pulse meter are not discussed here.

2.3 EXPERIMENTAL DESIGN

An experiment was designed to test the above hypotheses. The four physiological readings measured were heart rate, temperature, GSR and somatic movement. The heart rate was measured through a commercially available chest strap sensor. The temperature was measured with a thermocouple attached to a digital multimeter (DMM). The GSR was also measured with a DMM. The somatic movement was measured by recording the computer mouse movements.

www.vtucs.com

2.3.1 Method

Six people participated in this study (3 male, 3 female). The experiment was within subject design and order of presentation was counter-balanced across participants.

2.3.2 Procedure

Participants were asked to sit in front of the computer and hold the temperature and GSR sensors in their left hand hold the mouse with their right hand and wore the chest sensor. The resting (baseline) measurements were recorded for five minutes and then the participant was instructed to act out one emotion for five minutes. The emotions consisted of: anger, fear, sadness, disgust, happiness and surprise. The only instruction for acting out the emotion was to show the emotion in their facial expressions.

2.3.3 Results

The data for each subject consisted of scores for four physiological assessments [GSA, GSR, pulse, and skin temperature, for each of the six emotions (anger, disgust, fear, happiness, sadness, and surprise)] across the five minute baseline and test sessions. GSA data was sampled 80 times per second, GSR and temperature were reported approximately 3-4 times per second and pulse was recorded as a beat was detected, approximately 1 time per second. We first calculated the mean score for each of the baseline and test sessions. To account for individual variance in physiology, we calculated the difference between the baseline and test scores. Scores that differed by more than one and a half standard deviations from the mean were treated as missing. By this criterion, twelve score were removed from the analysis. The remaining data are described in Table 1.

www.vtucs.com

Table 1: Difference Scores.

		Anger	Disgust	Fear	Happiness	Sadness	Surprise
GSA	Mean	-0.66	-1.15	-2.02	.22	0.14	-.1.28
	Std. Dev.	1.87	1.02	0.23	1.60	2.44	1.16
GSR	Mean	-41209	-53206	-61160	-38999	-417990	-41242
	Std. Dev.	63934	8949	47297	46650	586309	24824
Pulse	Mean	2.56	2.07	3.28	2.40	4.83	2.84
	Std. Dev.	1.41	2.73	2.10	2.33	2.91	3.18
Temp	Mean	1.36	1.79	3.76	1.79	2.89	3.26
	Std. Dev.	3.75	2.66	3.81	3.72	4.99	0.90

In order to determine whether our measures of physiology could discriminate among the six different emotions, the data were analyzed with a discriminant function analysis. The four physiological difference scores were the discriminating variables and the six emotions were the discriminated groups. The variables were entered into the equation simultaneously, and four canonical discriminant functions were calculated. A Wilks' Lambda test of these four functions was marginally statistically significant; for lambda = .192, chi-square (20) = 29.748, $p < .075$. The functions are shown in Table 2

Table 2: Standardized Discriminant Function Coefficients.

	Function			
	1	2	3	4
GSA	0.593	-0.926	0.674	0.033
GSR	-0.664	0.957	0.350	0.583
Pulse	1.006	0.484	0.026	0.846
Temp.	1.277	0.405	0.423	-0.293

The unstandardized canonical discriminant functions evaluated at group means are shown in Table 3. Function 1 is defined by sadness and fear at one end and anger and surprise at the other. Function 2 has fear and disgust at one end and sadness at the other. Function 3 has happiness at one end and surprise at the other. Function 4 has disgust and anger at one end and surprise at the other. Table 3:

Table 3: Functions at Group Centroids.

EMOTION	Function			
	1	2	3	4
anger	-1.166	-0.052	-0.108	0.137
fear	1.360	1.704	-0.046	-0.093
sadness	2.168	-0.546	-0.096	-0.006
disgust	-0.048	0.340	0.079	0.184
happiness	-0.428	-0.184	0.269	-0.075
surprise	-1.674	-0.111	-0.247	-0.189

To determine the effectiveness of these functions, we used them to predict the group membership for each set of physiological data. As shown in Table 4, two-thirds of the cases were successfully classified

Table 4: Classification Results.

		Predicted Group Membership						Total
	EMOTION	Anger	Fear	sadness	disgust	happine	surprise	
Original	anger	2	0	0	0	2	1	5
	fear	0	2	0	0	0	0	2
	sadness	0	0	4	0	1	0	5
	disgust	0	1	0	1	1	0	3
	happiness	1	0	0	0	5	0	6
	surprise	0	0	0	0	1	2	3

The results show the theory behind the Emotion mouse work is fundamentally sound. The physiological measurements were correlated to emotions using a correlation model. The correlation model is derived from a calibration process in which a baseline attribute-to emotion correlation is rendered based on statistical analysis of calibration signals generated by users having emotions that are measured or otherwise known at calibration time. Now that we have proven the method, the next step is to improve the hardware. Instead of using cumbersome multimeters to gather information about the user, it will be better to use smaller and less intrusive units. We plan to improve our infrared

www...cs.co

pulse detector which can be placed inside the body of the mouse. Also, a framework for the user modeling needs to be develop in order to correctly handle all of the information after it has been gathered. There are other possible applications for the Emotion technology other than just increased productivity for a desktop computer user. Other domains such as entertainment, health and the communications and the automobile industry could find this technology useful for other purposes.

www.vtucs.com

3. MANUAL AND GAZE INPUT CASCADED (MAGIC) POINTING

This work explores a new direction in utilizing eye gaze for computer input. Gaze tracking has long been considered as an alternative or potentially superior pointing method for computer input. We believe that many fundamental limitations exist with traditional gaze pointing. In particular, it is unnatural to overload a perceptual channel such as vision with a motor control task. We therefore propose an alternative approach, dubbed MAGIC (Manual And Gaze Input Cascaded) pointing. With such an approach, pointing appears to the user to be a manual task, used for fine manipulation and selection. However, a large portion of the cursor movement is eliminated by warping the cursor to the eye gaze area, which encompasses the target. Two specific MAGIC pointing techniques, one conservative and one liberal, were designed, analyzed, and implemented with an eye tracker we developed. They were then tested in a pilot study. This early stage exploration showed that the MAGIC pointing techniques might offer many advantages, including reduced physical effort and fatigue as compared to traditional manual pointing, greater accuracy and naturalness than traditional gaze pointing, and possibly faster speed than manual pointing. The pros and cons of the two techniques are discussed in light of both performance data and subjective reports.

In our view, there are two fundamental shortcomings to the existing gaze pointing techniques, regardless of the maturity of eye tracking technology. First, given the one-degree size of the fovea and the subconscious jittery motions that the eyes constantly produce, eye gaze is not precise enough to operate UI widgets such as scrollbars, hyperlinks, and slider handles *In Proc. CHI'99: ACM Conference on Human Factors in Computing Systems. 246-253, Pittsburgh, 15-20 May1999 Copyright ACM 1999 0-201-48559-1/99/05...$5.00* on today's GUI interfaces. At a 25-inch viewing distance to the screen, one degree of arc corresponds to 0.44 in, which is twice the size of a typical scroll bar and much greater than the size of a typical character.

www.vtucs.com

Second, and perhaps more importantly, the eye, as one of our primary perceptual devices, has not evolved to be a control organ. Sometimes its movements are voluntarily controlled while at other times it is driven by external events. With the target selection by dwell time method, considered more natural than selection by blinking [7], one has to be conscious of where one looks and how long one looks at an object. If one does not look at a target continuously for a set threshold (e.g., 200 ms), the target will not be successfully selected. On the other hand, if one stares at an object for more than the set threshold, the object will be selected, regardless of the user's intention. In some cases there is not an adverse effect to a false target selection. Other times it can be annoying and counter-productive (such as unintended jumps to a web page). Furthermore, dwell time can only substitute for one mouse click. There are often two steps to target activation. A single click selects the target (e.g., an application icon) and a double click (or a different physical button click) opens the icon (e.g., launches an application). To perform both steps with dwell time is even more difficult. In short, to load the visual perception channel with a motor control task seems fundamentally at odds with users' natural mental model in which the eye searches for and takes in information and the hand produces output that manipulates external objects. Other than for disabled users, who have no alternative, using eye gaze for practical pointing does not appear to be very promising.

Are there interaction techniques that utilize eye movement to assist the control task but do not force the user to be overly conscious of his eye movement? We wanted to design a technique in which pointing and selection remained primarily a manual control task but were also aided by gaze tracking. Our key idea is to use gaze to dynamically redefine (warp) the "home" position of the pointing cursor to be at the vicinity of the target, which was presumably what the user was looking at, thereby effectively reducing the cursor movement amplitude needed for target selection.

Once the cursor position had been redefined, the user would need to only make a small movement to, and click on, the target with a regular manual input device. In other words, we wanted to achieve Manual And Gaze Input Cascaded (MAGIC) pointing, or

www.vtucs.com

Manual Acquisition with Gaze Initiated Cursor. There are many different ways of designing a MAGIC pointing technique. Critical to its effectiveness is the identification of the target the user intends to acquire. We have designed two MAGIC pointing techniques, one liberal and the other conservative in terms of target identification and cursor placement. The liberal approach is to warp the cursor to every new object the user looks at (See Figure 1).

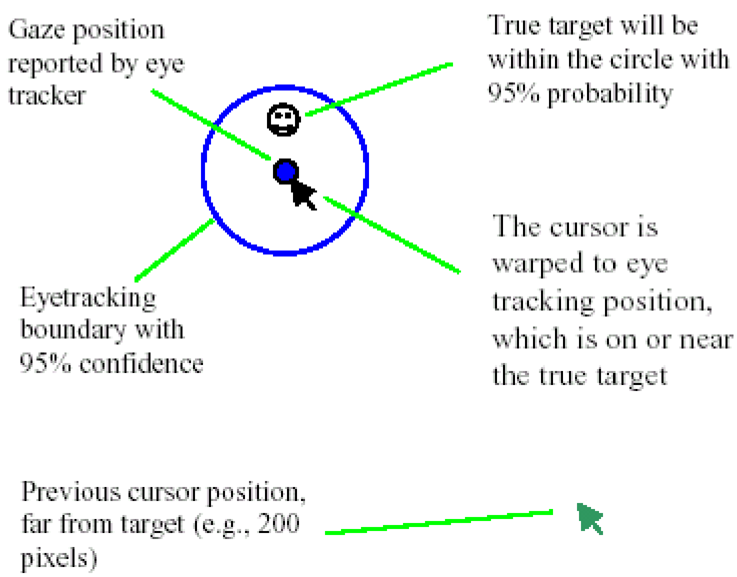

Figure 1. The liberal MAGIC pointing technique: cursor is placed in the vicinity of a target that the user fixates on.

The user can then take control of the cursor by hand near (or on) the target, or ignore it and search for the next target. Operationally, a new object is defined by sufficient distance (e.g., 120 pixels) from the current cursor position, unless the cursor is in a controlled motion by hand. Since there is a 120-pixel threshold, the cursor will not be warped when the user does continuous manipulation such as drawing. Note that this MAGIC pointing technique is different from traditional eye gaze control, where the user uses his eye to point at targets either without a cursor or with a cursor that constantly follows the jittery eye gaze motion.

The liberal approach may appear "pro-active," since the cursor waits readily in the vicinity of or on every potential target. The user may move the cursor once he decides to acquire the target he is looking at. On the other hand, the user may also feel that the cursor is over-active when he is merely looking at a target, although he may gradually adapt to ignore this behavior. The more conservative MAGIC pointing technique we have explored does not warp a cursor to a target until the manual input device has been actuated. Once the manual input device has been actuated, the cursor is warped to the gaze area reported by the eye tracker. This area should be on or in the vicinity of the target. The user would then steer the cursor annually towards the target to complete the target acquisition. As illustrated in Figure 2, to minimize directional uncertainty after the cursor appears in the conservative technique, we introduced an "intelligent" bias. Instead of being placed at the center of the gaze area, the cursor position is offset to the intersection of the manual actuation vector and the boundary f the gaze area. This means that once warped, the cursor is likely to appear in motion towards the target, regardless of how the user actually actuated the manual input device. We hoped that with the intelligent bias the user would not have to Gaze position reported by eye tracker Eye tracking boundary with 95% confidence True target will be within the circle with 95% probability. The cursor is warped to eye tracking position, which is on or near the true target Previous cursor position, far from target (e.g., 200 pixels) Figure 1.

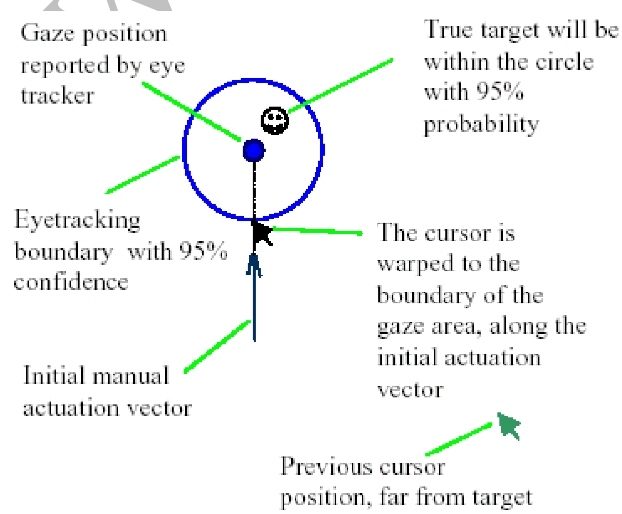

Figure 2. The conservative MAGIC pointing technique with "intelligent offset"

www.vtucs.com

The liberal MAGIC pointing technique: cursor is placed in the vicinity of a target that the user fixates on. Actuate input device, observe the cursor position and decide in which direction to steer the cursor. The cost to this method is the increased manual movement amplitude. Figure 2. The conservative MAGIC pointing technique with "intelligent offset" To initiate a pointing trial, there are two strategies available to the user. One is to follow "virtual inertia:" move from the cursor's current position towards the new target the user is looking at. This is likely the strategy the user will employ, due to the way the user interacts with today's interface. The alternative strategy, which may be more advantageous but takes time to learn, is to ignore the previous cursor position and make a motion which is most convenient and least effortful to the user for a given input device.

The goal of the conservative MAGIC pointing method is the following. Once the user looks at a target and moves the input device, the cursor will appear "out of the blue" in motion towards the target, on the side of the target opposite to the initial actuation vector. In comparison to the liberal approach, this conservative approach has both pros and cons. While with this technique the cursor would never be over-active and jump to a place the user does not intend to acquire, it may require more hand-eye coordination effort. Both the liberal and the conservative MAGIC pointing techniques offer the following *potential* advantages:

1. Reduction of manual stress and fatigue, since the cross screen long-distance cursor movement is eliminated from manual control.
2. Practical accuracy level. In comparison to traditional pure gaze pointing whose accuracy is fundamentally limited by the nature of eye movement, the MAGIC pointing techniques let the hand complete the pointing task, so they can be as accurate as any other manual input techniques.
3. A more natural mental model for the user. The user does not have to be aware of the role of the eye gaze. To the user, pointing continues to be a manual task, with a cursor conveniently appearing where it needs to be.

www.vtucs.com

4. Speed. Since the need for large magnitude pointing operations is less than with pure manual cursor control, it is possible that MAGIC pointing will be faster than pure manual pointing.
5. Improved subjective speed and ease-of-use. Since the manual pointing amplitude is smaller, the user may perceive the MAGIC pointing system to operate faster and more pleasantly than pure manual control, even if it operates at the same speed or more slowly.

The fourth point wants further discussion. According to the well accepted Fitts' Law, manual pointing time is logarithmically proportional to the A/W ratio, where A is the movement distance and W is the target size. In other words, targets which are smaller or farther away take longer to acquire.

For MAGIC pointing, since the target size remains the same but the cursor movement distance is shortened, the pointing time can hence be reduced. It is less clear if eye gaze control follows Fitts' Law. In Ware and Mikaelian's study, selection time was shown to be logarithmically proportional to target distance, thereby conforming to Fitts' Law. To the contrary, Silbert and Jacob [9] found that trial completion time with eye tracking input increases little with distance, therefore defying Fitts' Law. In addition to problems with today's eye tracking systems, such as delay, error, and inconvenience, there may also be many potential human factor disadvantages to the MAGIC pointing techniques we have proposed, including the following:

1. With the more liberal MAGIC pointing technique, the cursor warping can be overactive at times, since the cursor moves to the new gaze location whenever the eye gaze moves more than a set distance (e.g., 120 pixels) away from the cursor. This could be particularly distracting when the user is trying to read. It is possible to introduce additional constraint according to the context. For example, when the user's eye appears to follow a text reading pattern, MAGIC pointing can be automatically suppressed.

www.vtucs.com

2. With the more conservative MAGIC pointing technique, the uncertainty of the exact location at which the cursor might appear may force the user, especially a novice, to adopt a cumbersome strategy: take a touch (use the manual input device to activate the cursor), wait (for the cursor to appear), and move (the cursor to the target manually). Such a strategy may prolong the target acquisition time. The user may have to learn a novel hand-eye coordination pattern to be efficient with this technique. Gaze position reported by eye tracker Eye tracking boundary with 95% confidence True target will be within the circle with 95% probability The cursor is warped to the boundary of the gaze area, along the initial actuation vector Previous cursor position, far from target Initial manual actuation vector

3. With pure manual pointing techniques, the user, knowing the current cursor location, could conceivably perform his motor acts in parallel to visual search. Motor action may start as soon as the user's gaze settles on a target. With MAGIC pointing techniques, the motor action computation (decision) cannot start until the cursor appears. This may negate the time saving gained from the MAGIC pointing technique's reduction of movement amplitude. Clearly, experimental (implementation and empirical) work is needed to validate, refine, or invent alternative MAGIC pointing techniques.

3.1 IMPLEMENTATION

We took two engineering efforts to implement the MAGIC pointing techniques. One was to design and implement an eye tracking system and the other was to implement MAGIC pointing techniques at the operating systems level, so that the techniques can work with all software applications beyond "demonstration" software.

www.vtucs.com

3.2 THE IBM ALMADEN EYE TRACKER

Since the goal of this work is to explore MAGIC pointing as a user interface technique, we started out by purchasing a commercial eye tracker (ASL Model 5000) after a market survey. In comparison to the system reported in early studies (e.g. [7]), this system is much more compact and reliable. However, we felt that it was still not robust enough for a variety of people with different eye characteristics, such as pupil brightness and correction glasses. We hence chose to develop and use our own eye tracking system [10]. Available commercial systems, such as those made by ISCAN Incorporated, LC Technologies, and Applied Science Laboratories (ASL), rely on a single light source that is positioned either off the camera axis in the case of the ISCANETL-400 systems, or on-axis in the case of the LCT and the ASL E504 systems. Illumination from an off-axis source (or ambient illumination) generates a dark pupil image.

When the light source is placed on-axis with the camera optical axis, the camera is able to detect the light reflected from the interior of the eye, and the image of the pupil appears bright (see Figure 3).

This effect is often seen as the red-eye in flash photographs when the flash is close to the camera lens.

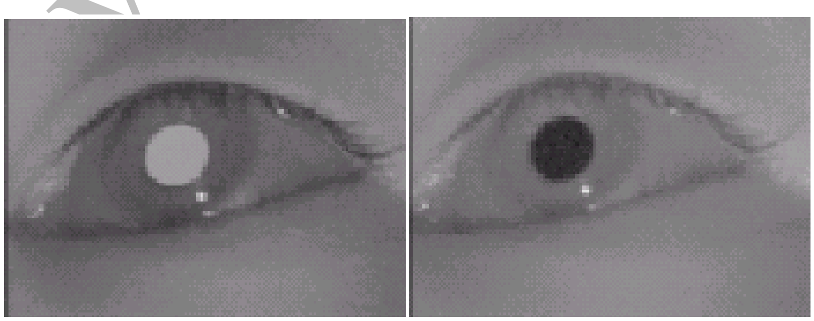

Figure 3. Bright (left) and dark (right) pupil images resulting from on- and off-axis illumination. The glints, or corneal reflections, from the on- and off-axis light sources can be easily identified as the bright points in the iris.

www.vtucs.com

Bright (left) and dark (right) pupil images resulting from on- and off-axis illumination. The glints, or corneal reflections, from the on- and off-axis light sources can be easily identified as the bright points in the iris. The Almaden system uses two near infrared (IR) time multiplexed light sources, composed of two sets of IR LED's, which were synchronized with the camera frame rate. One light source is placed very close to the camera's optical axis and is synchronized with the even frames. Odd frames are synchronized with the second light source, positioned off axis. The two light sources are calibrated to provide approximately equivalent whole-scene illumination. Pupil detection is realized by means of subtracting the dark pupil image from the bright pupil image. After thresholding the difference, the largest connected component is identified as the pupil. This technique significantly increases the robustness and reliability of the eye tracking system. After implementing our system with satisfactory results, we discovered that similar pupil detection schemes had been independently developed by Tomonoetal and Ebisawa and Satoh.

It is unfortunate that such a method has not been used in the commercial systems. We recommend that future eye tracking product designers consider such an approach. Once the pupil has been detected, the corneal reflection (the glint reflected from the surface of the cornea due to one of the light sources) is determined from the dark pupil image. The reflection is then used to estimate the user's point of gaze in terms of the screen coordinates where the user is looking at. The estimation of the user's gaze requires an initial calibration procedure, similar to that required by commercial eye trackers. Our system operates at 30 frames per second on a Pentium II 333 MHz machine running Windows NT. It can work with any PCI frame grabber compatible with Video for Windows.

www.vtucs.com

3.3 IMPLIMENTING MAGIC POINTING

We programmed the two MAGIC pointing techniques on a Windows NT system. The techniques work independently from the applications. The MAGIC pointing program takes data from both the manual input device (of any type, such as a mouse) and the eye tracking system running either on the same machine or on another machine connected via serial port. Raw data from an eye tracker cannot be directly used for gaze-based interaction, due to noise from image processing, eye movement jitters, and samples taken during *saccade* (ballistic eye movement) periods. We experimented with various filtering techniques and found the most effective filter in our case is similar to that described in [7]. The goal of filter design in general is to make the best compromise between preserving signal bandwidth and eliminating unwanted noise. In the case of eye tracking, as Jacob argued, eye information relevant to interaction lies in the *fixations*. The key is to select fixation points with minimal delay. Samples collected during a saccade are unwanted and should be avoided. In designing our algorithm for picking points of fixation, we considered our tracking system speed (30 Hz), and that the MAGIC pointing techniques utilize gaze information only once for each new target, probably immediately after a saccade. Our filtering algorithm was designed to pick a fixation with minimum delay by means of selecting two adjacent points over two samples.

3.4 EXPERIMENT

Empirical studies are relatively rare in eye tracking-based interaction research, although they are particularly needed in this field. Human behavior and processes at the perceptual motor level often do not conform to conscious-level reasoning. One usually cannot correctly describe how to make a turn on a bicycle. Hypotheses on novel interaction techniques can only be validated by empirical data. However, it is also particularly difficult to conduct empirical research on gaze-based interaction techniques, due to the complexity of eye movement and the lack of reliability in eye tracking

www.vtucs.com

equipment. Satisfactory results only come when "everything is going right." When results are not as expected, it is difficult to find the true reason among many possible reasons: Is it because a subject's particular eye property fooled the eye tracker? Was there a calibration error? Or random noise in the imaging system? Or is the hypothesis in fact invalid? We are still at a very early stage of exploring the MAGIC pointing techniques. More refined or even very different techniques may be designed in the future. We are by no means ready to conduct the definitive empirical studies on MAGIC pointing. However, we also feel that it is important to subject our work to empirical evaluations early so that quantitative observations can be made and fed back to the iterative design-evaluation-design cycle. We therefore decided to conduct a small-scale pilot study to take an initial peek at the use of MAGIC pointing, however unrefined.

3.5 EXPERIMENTAL DESIGN

The two MAGIC pointing techniques described earlier were put to test using a set of parameters such as the filter's temporal and spatial thresholds, the minimum cursor warping distance, and the amount of "intelligent bias" (subjectively selected by the authors without extensive user testing). Ultimately the MAGIC pointing techniques should be evaluated with an array of manual input devices, against both pure manual and pure gaze-operated pointing methods.

Since this is an early pilot study, we decided to limit ourselves to one manual input device. A standard mouse was first considered to be the manual input device in the experiment. However, it was soon realized not to be the most suitable device for MAGIC pointing, especially when a user decides to use the push-upwards strategy with the intelligent offset. Because in such a case the user always moves in one direction, the mouse tends to be moved off the pad, forcing the user adjust the mouse position, often during a pointing trial. We hence decided to use a miniature isometric pointing stick (IBM Track Point IV, commercially used in the IBM ThinkPad 600 and 770 series notebook computers). Another device suitable for MAGIC pointing is a touchpad: the

www.vtucs.com

user can choose one convenient gesture and to take advantage of the intelligent offset. The experimental task was essentially a Fitts' pointing task. Subjects were asked to point and click at targets appearing in random order. If the subject clicked off-target, a miss was logged but the trial continued until a target was clicked. An extra trial was added to make up for the missed trial. Only trials with no misses were collected for time performance analyses. Subjects were asked to complete the task as quickly as possible and as accurately as possible. To serve as a motivator, a $20 cash prize was set for the subject with the shortest mean session completion time with any technique.

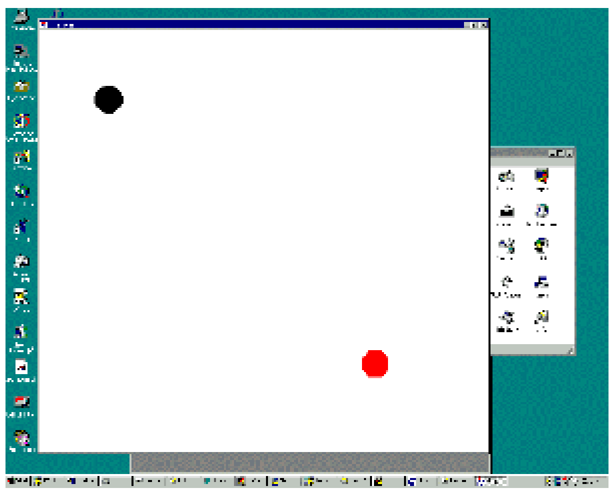

Figure 4. Experimental task: point at paired targets

The task was presented on a 20 inch CRT color monitor, with a 15 by 11 inch viewable area set at resolution of 1280 by 1024 pixels. Subjects sat from the screen at a distance of 25 inches. The following factors were manipulated in the experiments:

- two target sizes: 20 pixels (0.23 in or 0.53 degree of viewing angle at 25 in distance) and 60 pixels in diameter (0.7 in, 1.61 degree)
- three target distances: 200 pixels (2.34 in, 5.37 degree), 500 pixels (5.85 in, 13.37 degree), and 800 pixels (9.38 in, 21.24 degree)
- three pointing directions: horizontal, vertical and diagonal

www om

A within-subject design was used. Each subject performed the task with all three techniques: (1) Standard, pure manual pointing with no gaze tracking (No Gaze); (2) The conservative MAGIC pointing method with intelligent offset (Gaze1); (3) The liberal MAGIC pointing method (Gaze2). Nine subjects, seven male and two female, completed the experiment. The order of techniques was balanced by a Latin square pattern. Seven subjects were experienced Track Point users, while two had little or no experience. With each technique, a 36-trial practice session was first given, during which subjects were encouraged to explore and to find the most suitable strategies (aggressive, gentle, etc.). The practice session was followed by two data collection sessions. Although our eye tracking system allows head motion, at least for those users who do not wear glasses, we decided to use a chin rest to minimize instrumental error.

3.6 EXPERIMENTAL RESULTS

Given the pilot nature and the small scale of the experiment, we expected the statistical power of the results to be on the weaker side. In other words, while the significant effects revealed are important, suggestive trends that are statistically non-significant are still worth noting for future research. First, we found that subjects' trial completion time significantly varied with techniques: $F(2, 16) = 6.36$, $p < 0.01$.

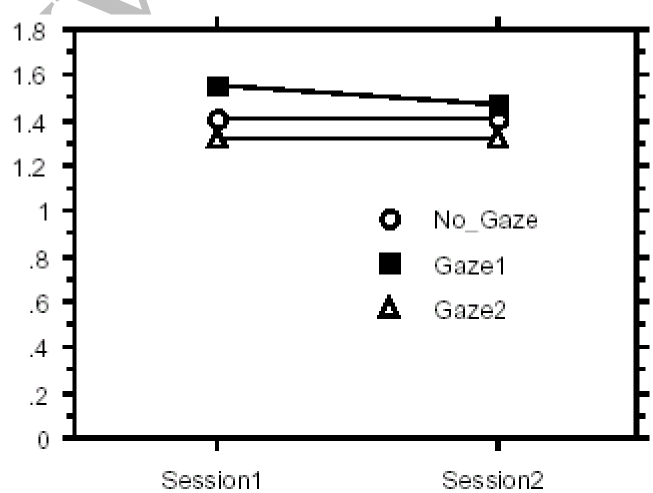

Figure 5. Mean completion time (sec) vs. experiment session

www.vtucs.com

The total average completion time was 1.4 seconds with the standard manual control technique 1.52 seconds with the conservative MAGIC pointing technique (Gaze1), and 1.33 seconds with the liberal MAGIC pointing technique (Gaze2). Note that the Gaze1

Technique had the greatest improvement from the first to the second experiment session, suggesting the possibility of matching the performance of the other two techniques with further practice.

As expected, target size significantly influenced pointing time: $F(1,8) = 178$, $p < 0.001$. This was true for both the manual and the two MAGIC pointing techniques (Figure 6).

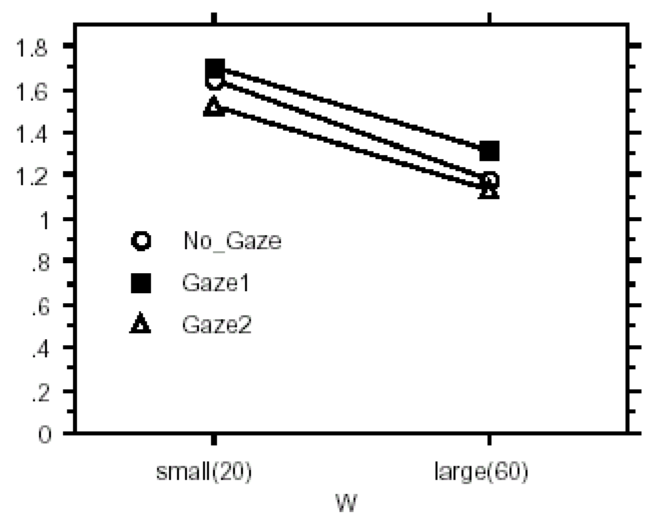

Figure 6. Mean completion time (sec) vs. target size (pixels)

Pointing amplitude also significantly affected completion time: $F(2, 8) = 97.5$, $p < 0.001$. However, the amount of influence varied with the technique used, as indicated by the significant interaction between technique and amplitude: $F(4, 32) = 7.5$, $p < 0.001$ (Figure 7).

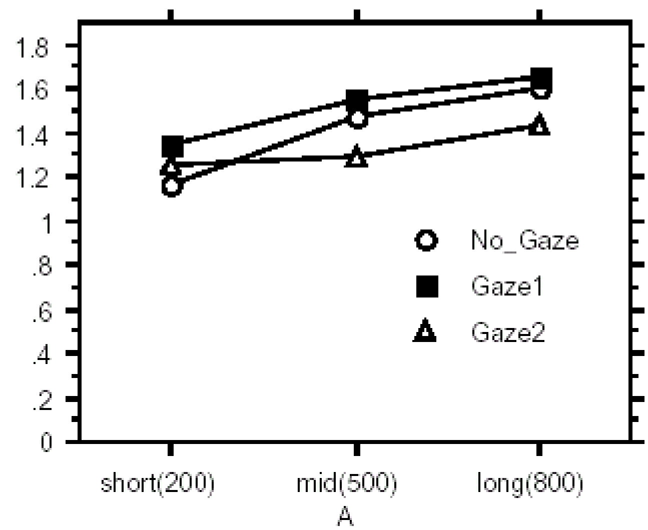

Figure 7. Mean completion time (sec) vs. pointing amplitude (pixels)

As pointing amplitude increased from 200 pixels to 500 pixels and then to 800 pixels, subjects' completion time with the No Gaze condition increased in a non-linear, logarithmic-like pace as Fitts' Law predicts. This is less true with the two MAGIC pointing techniques, particularly the Gaze2 condition, which is definitely not logarithmic. Nonetheless, completion time with the MAGIC pointing techniques did increase as target distance increased. This is intriguing because in MAGIC pointing techniques, the manual control portion of the movement should be the distance from the warped cursor position to the true target. Such distance depends on eye tracking system accuracy, which is unrelated to the previous cursor position.

In short, while completion time and target distance with the MAGIC pointing techniques did not completely follow Fitts' Law, they were not completely independent either. Indeed, when we lump target size and target distance according to the Fitts' Law

Index of Difficulty $ID = \log 2(A/W + 1)$ [15],

www.vtucs.com

We see a similar phenomenon. For the No Gaze condition:

$$T = 0.28 + 0.31\ ID\ (r^2=0.912)$$

The particular settings of our experiment were very different from those typically reported in a Fitts' Law experiment: to simulate more realistic tasks we used circular targets distributed in varied directions in a randomly shuffled order, instead of two vertical bars displaced only in the horizontal dimension. We also used an isometric pointing stick, not a mouse. Considering these factors, the above equation is reasonable. The index of performance (*IP*) was 3.2 bits per second, in comparison to the 4.5 bits per second in a typical setting (repeated mouse clicks on two vertical bars) [16].

For the Gaze1 condition:

$$T = 0.8 + 0.22\ ID\ (r^2=0.716)$$

$$IP = 4.55 \text{ bits per second}$$

For Gaze2:

$$T = 0.6 + 0.21\ ID\ (r^2=0.804)$$

$$IP = 4.76 \text{ bits per second}$$

Note that the data from the two MAGIC pointing techniques fit the Fitts' Law model relatively poorly (as expected), although the indices of performance (4.55 and 4.76 bps) were much higher than the manual condition (3.2 bps).

Finally, Figure 8 shows that the angle at which the targets were presented had little influence on trial completion time: $F(2, 16) = 1.57$, N.S.

www.vtucs.com

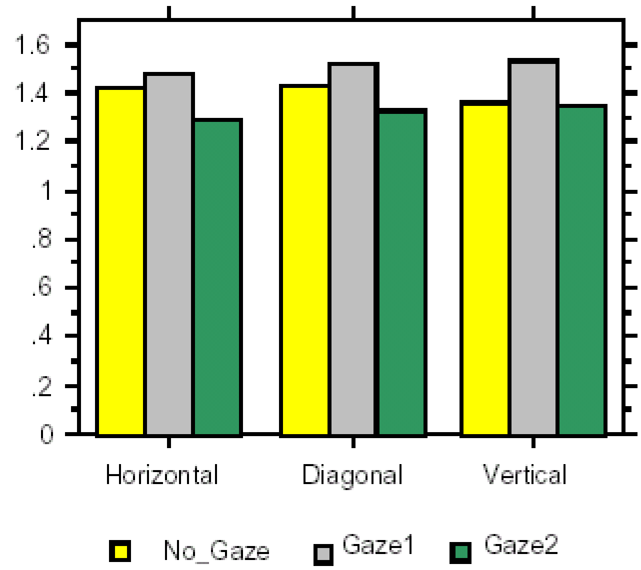

Figure 8. Mean completion time (sec) vs. target angle (degrees)

The number of misses (clicked off target) was also analyzed. The only significant factor to the number of misses is target size: $F(1,8) = 15.6$, $p < 0.01$. Users tended to have more misses with small targets. More importantly, subjects made no more misses with the MAGIC pointing techniques than with the pure manual technique (No_Gaze – 8.2 %, Gaze1 –7%, Gaze2 – 7.5%).

www.vtucs.com

4. ARTIFICIAL INTELLIGENT SPEECH RECOGNITION

It is important to consider the environment in which the speech recognition system has to work. The grammar used by the speaker and accepted by the system, noise level, noise type, position of the microphone, and speed and manner of the user's speech are some factors that may affect the quality of speech recognition .When you dial the telephone number of a big company, you are likely to hear the sonorous voice of a cultured lady who responds to your call with great courtesy saying "Welcome to company X. Please give me the extension number you want". You pronounce the extension number, your name, and the name of person you want to contact. If the called person accepts the call, the connection is given quickly. This is artificial intelligence where an automatic call-handling system is used without employing any telephone operator.

4.1THE TECHNOLOGY

Artificial intelligence (AI) involves two basic ideas. First, it involves studying the thought processes of human beings. Second, it deals with representing those processes via machines (like computers, robots, etc). AI is behavior of a machine, which, if performed by a human being, would be called intelligent. It makes machines smarter and more useful, and is less expensive than natural intelligence. Natural language processing (NLP) refers to artificial intelligence methods of communicating with a computer in a natural language like English. The main objective of a NLP program is to understand input and initiate action. The input words are scanned and matched against internally stored known words. Identification of a key word causes some action to be taken. In this way, one can communicate with the computer in one's language. No special commands or computer language are required. There is no need to enter programs in a special language for creating software.

www.vtucs.com

4.2 SPEECH RECOGNITION

The user speaks to the computer through a microphone, which, in used; a simple system may contain a minimum of three filters. The more the number of filters used, the higher the probability of accurate recognition. Presently, switched capacitor digital filters are used because these can be custom-built in integrated circuit form. These are smaller and cheaper than active filters using operational amplifiers. The filter output is then fed to the ADC to translate the analogue signal into digital word. The ADC samples the filter outputs many times a second. Each sample represents different amplitude of the signal .Evenly spaced vertical lines represent the amplitude of the audio filter output at the instant of sampling. Each value is then converted to a binary number proportional to the amplitude of the sample. A central processor unit (CPU) controls the input circuits that are fed by the ADCS. A large RAM (random access memory) stores all the digital values in a buffer area. This digital information, representing the spoken word, is now accessed by the CPU to process it further. The normal speech has a frequency range of 200 Hz to 7 kHz. Recognizing a telephone call is more difficult as it has bandwidth limitation of 300 Hz to3.3 kHz.

As explained earlier, the spoken words are processed by the filters and ADCs. The binary representation of each of these words becomes a template or standard, against which the future words are compared. These templates are stored in the memory. Once the storing process is completed, the system can go into its active mode and is capable of identifying spoken words. As each word is spoken, it is converted into binary equivalent and stored in RAM. The computer then starts searching and compares the binary input pattern with the templates. t is to be noted that even if the same speaker talks the same text, there are always slight variations in amplitude or loudness of the signal, pitch, frequency difference, time gap, etc. Due to this reason, there is never a perfect match between the template and binary input word. The pattern matching process therefore uses statistical techniques and is designed to look for the best fit.

www.vtucs.com

The values of binary input words are subtracted from the corresponding values in the templates. If both the values are same, the difference is zero and there is perfect match. If not, the subtraction produces some difference or error. The smaller the error, the better the match. When the best match occurs, the word is identified and displayed on the screen or used in some other manner. The search process takes a considerable amount of time, as the CPU has to make many comparisons before recognition occurs. This necessitates use of very high-speed processors. A large RAM is also required as even though a spoken word may last only a few hundred milliseconds, but the same is translated into many thousands of digital words. It is important to note that alignment of words and templates are to be matched correctly in time, before computing the similarity score. This process, termed as dynamic time warping, recognizes that different speakers pronounce the same words at different speeds as well as elongate different parts of the same word. This is important for the speaker-independent recognizers.

4.3 APPLICATIONS

One of the main benefits of speech recognition system is that it lets user do other works simultaneously. The user can concentrate on observation and manual operations, and still control the machinery by voice input commands. Another major application of speech processing is in military operations. Voice control of weapons is an example. With reliable speech recognition equipment, pilots can give commands and information to the computers by simply speaking into their microphones—they don't have to use their hands for this purpose. Another good example is a radiologist scanning hundreds of X-rays, ultrasonograms, CT scans and simultaneously dictating conclusions to a speech recognition system connected to word processors. The radiologist can focus his attention on the images rather than writing the text. Voice recognition could also be used on computers for making airline and hotel reservations. A user requires simply stating his needs, to make reservation, cancel a reservation, or making enquiries about schedule.

www.vtucs.com

5. THE SIMPLE USER INTERST TRACKER (SUITOR)

Computers would have been much more powerful, had they gained perceptual and sensory abilities of the living beings on the earth. What needs to be developed is an intimate relationship between the computer and the humans. And the Simple User Interest Tracker (SUITOR) is a revolutionary approach in this direction.

By observing the Webpage a netizen is browsing, the SUITOR can help by fetching more information at his desktop. By simply noticing where the user's eyes focus on the computer screen, the SUITOR can be more precise in determining his topic of interest. It can even deliver relevant information to a handheld device. The success lies in how much the suitor can be intimate to the user. IBM's BlueEyes research project began with a simple question, according to Myron Flickner, a manager in Almaden's USER group: Can we exploit nonverbal cues to create more effective user interfaces?

One such cue is gaze—the direction in which a person is looking. Flickner and his colleagues have created some new techniques for tracking a person's eyes and have incorporated this gaze-tracking technology into two prototypes. One, called SUITOR (Simple User Interest Tracker), fills a scrolling ticker on a computer screen with information related to the user's current task. SUITOR knows where you are looking, what applications you are running, and what Web pages you may be browsing. "If I'm reading a Web page about IBM, for instance," says Paul Maglio, the Almaden cognitive scientist who invented SUITOR, "the system presents the latest stock price or business news stories that could affect IBM. If I read the headline off the ticker, it pops up the story in a browser window. If I start to read the story, it adds related stories to the ticker. That's the whole idea of an attentive system—one that attends to what you are doing, typing, reading, so that it can attend to your information needs."

www.vtucs.com

6. CONCLUSION

The nineties witnessed quantum leaps interface designing for improved man machine interactions. The BLUE EYES technology ensures a convenient way of simplifying the life by providing more delicate and user friendly facilities in computing devices. Now that we have proven the method, the next step is to improve the hardware. Instead of using cumbersome modules to gather information about the user, it will be better to use smaller and less intrusive units. The day is not far when this technology will push its way into your house hold, making you more lazy. It may even reach your hand held mobile device. Any way this is only a technological forecast.

www.vtucs.com

7. BIBILIOGRAPHY

[1] Ekman, P. and Rosenberg, E. (Eds.) (1997). What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS*)*. OxfordUniversity Press: New York.

[2] Dryer, D.C. (1993). Multidimensional and Discriminant Function Analyses of Affective StateData. Stanford University, unpublished manuscript.

[3] Dryer, D.C. (1999). Getting personal with computers: How to design personalities for agents.

[4] Applied Artificial Intelligence Dryer, D.C., and Horowitz, L.M. (1997). When do opposites attract? Interpersonal Complementarity versus similarity. Journal of Personality and Social Psychology Johnson, R.C. (1999). Computer Program Recognizes Facial Expressions. EE Times www.eetimes.com, April 5.

[5] Picard, R. (1997). Affective Computing. MIT Press: Cambridge.

www.vtucs.com