Manual And Gaze Input Cascaded (MAGIC) Pointing

Shumin Zhai **Carlos Morimoto** **Steven Ihde**

IBM Almaden Research Center
650 Harry Road
San Jose, CA 95120 USA
+1 408 927 1112
{zhai, morimoto, ihde}@almaden.ibm.com

ABSTRACT

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.

Keywords

Gaze, eye, computer input, eye tracking, gaze tracking, pointing, multi-modal interface, Fitts' law, computer vision.

INTRODUCTION

Using the eyes as a source of input in "advanced user interfaces" has long been a topic of interest to the HCI field [1] [2] [3] [4]. Reports on eye tracking frequently appear not only in the research literature, but also in the popular press, such as the July 1996 issue of Byte magazine [5]. One of the basic goals that numerous researchers have attempted to achieve is to operate the user interface through eye gaze, with pointing (target acquisition) as the core element. There are many compelling reasons to motivate such a goal, including the following:

1. There are situations that prohibit the use of the hands, such as when the user's hands are disabled or continuously occupied with other tasks.
2. The eye can move very quickly in comparison to other parts of the body. Furthermore, as many researchers have long argued [3] [6], target acquisition usually requires the user to look at the target first, before actuating cursor control. Theoretically this means that if the eye gaze can be tracked and effectively used, no other input method can act as quickly. Increasing the speed of user input to the computer has long been an interest of HCI research.
3. Reducing fatigue and potential injury caused by operating keyboard and pointing devices is also an important concern in the user interface field. Repetitive stress injury affects an increasing number of computer users. Most users are not concerned with RSI until serious problems occur. Utilizing eye gaze movement to replace or reduce the amount of stress to the hand can be beneficial.

Clearly, to replace "what you see (and click on) is what you get" with "what you look at is what you get" [4] [6] has captivating appeal. However, the design and implementation of eye gaze-based computer input has been faced with two types of challenges. One is eye tracking technology itself, which will be briefly discussed in the Implementation section of the paper. The other challenge is the human factor issues involved in utilizing eye movement for computer input. Jacob [7] eloquently discussed many of these issues with insightful observations.

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

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.

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.

MAGIC POINTING

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 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). 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

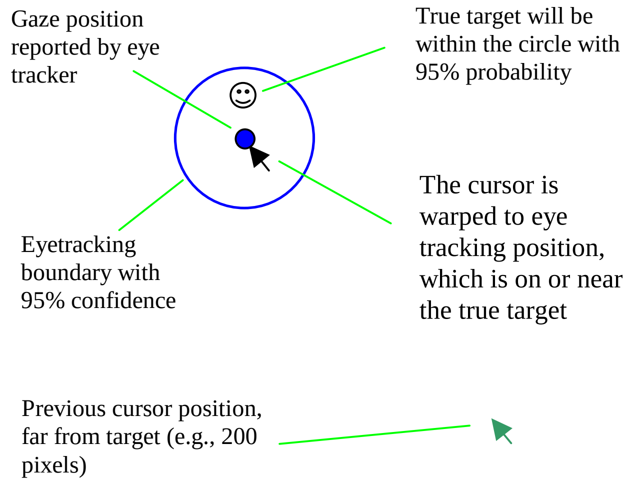

Figure 1. The liberal MAGIC pointing technique: cursor is placed in the vicinity of a target that the user fixates on.

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 [7] or with a cursor [3] 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 manually 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 of 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 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.

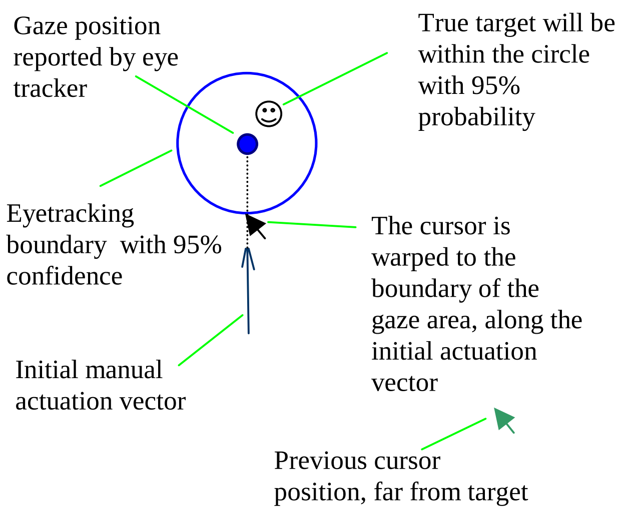

Figure 2. The conservative MAGIC pointing technique with "intelligent offset"

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. For example, on a small touchpad, the user may find it convenient to make an upward stroke with the index finger, causing the cursor to appear below the target.

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.
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 warrants further discussion. According to the well accepted Fitts' Law [8], 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 [3], 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.
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.

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.

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.

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 ISCAN ETL-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 [11] [12] (see Figure 3). This effect is often seen as the red-eye in flash photographs when the flash is close to the camera lens.

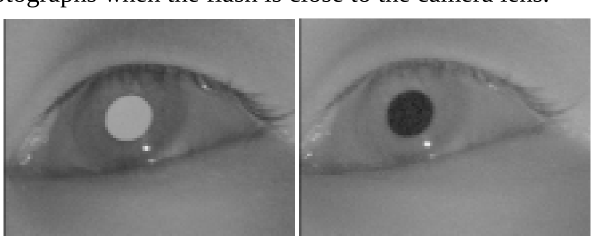

Figure 3. Bright (left) and dark (right) pupil images resulting from on- and off-axis illumination. The glints, or corneal reflections, from the on- and off-axis light sources can be easily identified as the bright points in the iris.

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 Tomono et al [13] and Ebisawa and Satoh [14]. 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.

Implementing 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 can not 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.

EXPERIMENT

Empirical studies, such as [3], 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 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.

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 (in the case of large targets suitable for gaze pointing). 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 TrackPoint IV, commercially used in the IBM Thinkpad 600 and 770 series notebook computers). Another device suitable for MAGIC pointing is a touchpad: the 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.

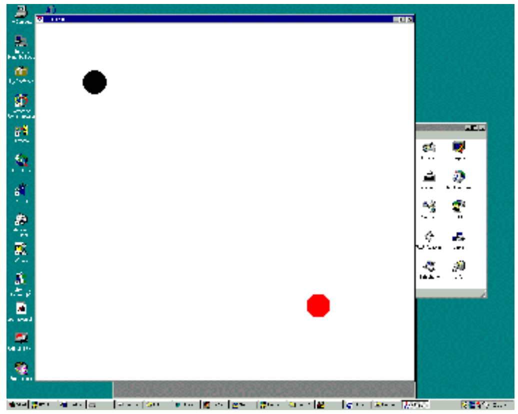

Figure 4. Experimental task: point at paired targets

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

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 TrackPoint 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.

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: $F(2, 16) = 6.36$, $p < 0.01$. The total average completion time was 1.4 seconds with the standard manual control technique (No_Gaze in Figure 5), 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.

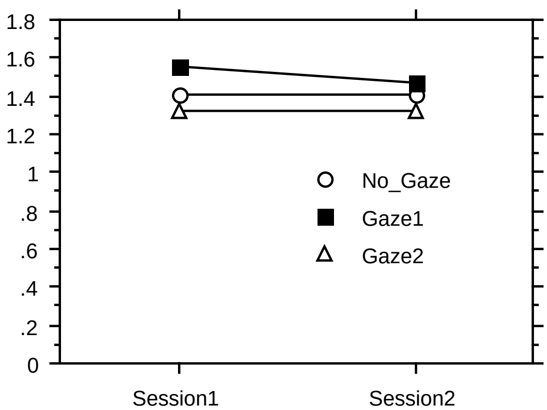

Figure 5. Mean completion time (sec) vs. experiment session

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).

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). 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.

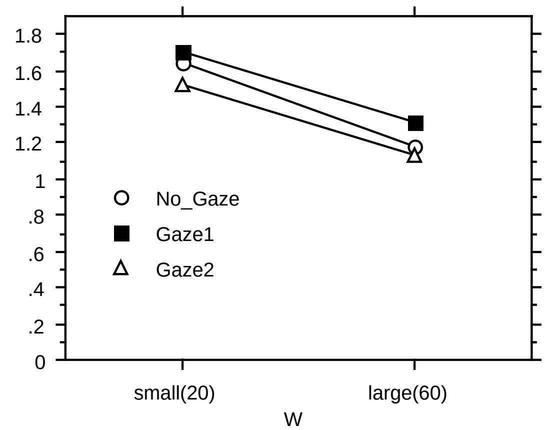

Figure 6. Mean completion time (sec) vs. target size (pixels)

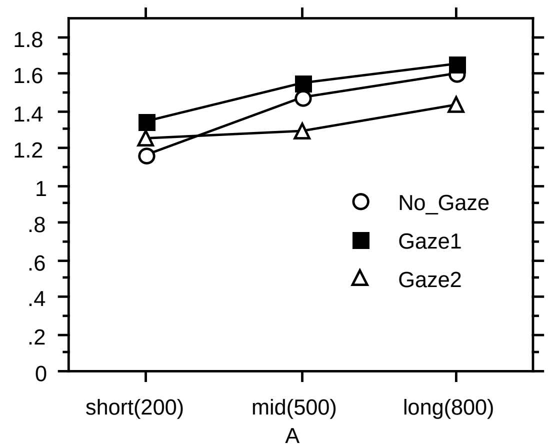

Figure 7. Mean completion time (sec) vs. pointing amplitude (pixels)

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], we see a similar phenomonon.

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.

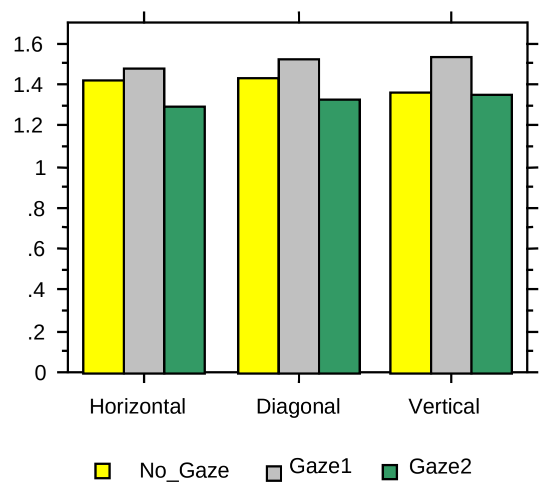

Figure 8. Mean completion time (sec) vs. target angle (degrees)

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%).

DISCUSSIONS AND CONCLUSIONS

The performance data from this pilot study shows both promises and shortcomings with the very first implementation of MAGIC pointing techniques. First, the MAGIC pointing approach actually worked. All subjects were able to operate the two novel techniques with minimal instruction. By the end of the experiment, subjects had less than 10 minutes of exposure to each technique, but were able to perform at a speed similar to their manual control skills. In the second session of the experiment, on average, subjects using the liberal MAGIC pointing technique performed slightly faster (6.8%) and those using the conservative technique slightly slower (4.3%) than those using pure manual pointing (1.41 seconds). The US$20 cash prize was claimed by a subject whose shortest mean completion time was 1.03 second, achieved with the Gaze2 technique. The closest runner up was 1.05 second, also achieved with the Gaze2 technique. Although some users performed in fact slower with the new techniques, subjectively they tended to feel faster with MAGIC pointing techniques. On a –5 (most unfavorable) to +5 (most favorable) scale, subjects gave an average rating of 1.5 (spread from -1 to +3) to the Gaze1 technique and 3.5 (from 2 to 4.5) to the Gaze2 technique. The overall positive reaction from the users could be due to any of the following factors: 1) The novel experience, which may or may not be fundamentally beneficial; 2) the reduced physical effort. Users might have liked the fact that a big chunk of the physical task was done automatically. Some subjects were disappointed after the MAGIC pointing sessions when they realized that the cursor would no longer move to the vicinity of the target "by itself."

The targets used in the experiment varied from small (0.53 degree) to large (1.6 degree), resembling realistic targets in practice. Notably, the traditional gaze pointing technique works well only for large targets (2.0 by 1.6 degree in [3] and deteriorated rapidly when target was smaller than 1.5 degree.

The reduced fatigue from pure manual pointing is self-evident, simply because less cursor movement is needed.

On the other hand, the speed advantage, when there was one, was not obvious. It is undoubtedly possible to improve the performance of the MAGIC pointing techniques. First, many aspects of the proposed techniques can be refined, including optimizing the parameters in the gaze system's filter and in the MAGIC pointing techniques themselves. The input device transfer function was designed to accommodate both large and small cursor movements. It is possible to optimize the transfer function for MAGIC pointing techniques. Second, the engineering aspects of the eye tracking system may also be improved. Many subjects commented that the eye tracker performance varied over time, probably due to their head motions during the session. In order to achieve the best results, we turned off the camera's servo mode and used a chin-rest. Some subjects did not stay steady in the chin rest as asked. Some subjects also noticed the delay in the tracking system, which depended on how quickly a pair of samples was detected, which in turn depends on noise in the system. In the ideal case, the delay can be as small as one sampling period (33 ms). Other times it may take several samples to find a pair of adjacent points.

In summary, the pros and cons of the two techniques were demonstrated both in the performance data and in subjects' comments. The conservative MAGIC pointing method was truly "conservative." Its average speed was slower than the "liberal" and the manual technique, although such a difference tended to shrink with practice. Some subjects commented that the conservative technique required more effort to coordinate the timing of eye-hand cooperation.

Others found it less distracting and more "discreet" than they found the liberal technique. Some also pointed out that it took them several trials to get used to the conservative technique, specifically the uncertainty of not knowing exactly where the cursor would appear. Interestingly, the intelligent offset, designed to reduce the directional uncertainty, was not unnoticed by some users who pointed out that the conservative technique had greater "tracking error": the cursor was farther from the target. Clearly we need to further test the conservative technique without the offset.

Overall subjects liked the liberal technique better for its responsiveness. This may change in a more realistic setting where pointing is mixed with other tasks, in which case the more discreet conservative technique may become more favorable.

Based on the results of this pilot experiment, we are refining the proposed MAGIC technique. Alternative techniques may also be designed in future research.

The IBM Almaden Gaze tracker described in the paper points to the rapid improvement in eye tracking technology. The price (and size) of commercial eye tracking equipment has dropped significantly in the last decade, from over US$100k to around US$20k. Our system hardware cost was around US$2000 (US$200 for camera and US$1500 for frame grabber), in addition to the computer (which also ran the applications used by subjects). As computer power and the price of cameras and video processing hardware continue to exponentially improve, it is conceivable that in the future mainstream computers will all be equipped with technology similar to that which we used in this experiment. Such a prospect calls for continued, in-depth research on eye-based interaction techniques. This work attempts to serve as one stepping stone in this process.

ACKNOWLEDGMENTS

This study was conducted as part of the IBM Blue Eyes project, led by Myron Flickner, who provided us great support. Barton Smith developed the Fitts' Law testing program used in the experiment. Dragutin Petkovic was a constant source of support and inspiration. Dave Koons, Rob Barrett, and Arnon Amir contributed to brainstorming discussions; and Johnny Accot provided much-needed help with data processing and analysis.

REFERENCES

1. Levine, J.L., *An Eye-Controlled Computer*, . 1981, IBM TJ Watson Research Center: Yorktown Heights, New York.

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

3. Ware, C. and H.H. Mikaelian. *An evaluation of an eye tracker as a device for computer input*. in *CHI+GI: ACM Conference on Human Factors in Computing Systems and Graphics Interface*. 1987. Toronto.

4. Jacob, R.J.K. *What You Look At is What You Get: Eye Movement-Based Interaction Techniques*. in *CHI'90: ACM Conference on Human Factors in Computing Systems*. 1990: Addison-Wesley/ACM Press.

5. Joch, A., *What Pupils Teach Computers.* Byte, 1996(July): p. 99-100.

6. Jacob, R.J.K., *The Use of Eye Movements in Human-Computer Interaction Techniques: What You Look At is What You Get.* ACM Transactions on Information Systems,, 1991. **vol. 9**(no. 3): p. 152-169.

7. Jacob, R.J.K., *Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces*, in *Advances in Human-Computer Interaction*, H.R. Hartson and D. Hix, Editors. 1993, Ablex Publishing Co.,: Norwood, N.J. p. 151-190.

8. Fitts, P.M., *The information capacity of the human motor system in controlling the amplitude of movement.* Journal of Experimental Psychology, 1954. **47**: p. 381-391.

9. Silbert, L. and R. Jacob, *The Advantage of Eye Gaze Interaction*, . 1996: Unpublished manuscript.

10. Morimoto, C., *et al.*, *Pupil detection and tracking using multiple light sources*, . 1998, IBM Almaden Research Center: San Jose.

11. Young, L. and D. Sheena, *Methods & designs: Survey of eye movement recording methods.* Behavioral Research Methods & Instrumentation, 1975. **7**(5): p. 397--429.

12. Hutchinson, T.E., *et al.*, *Human-computer interaction using eye-gaze input.* IEEE Transactions on Systems, Man, and Cybernetics, 1989. **19**(Nov.Dec): p. 1527--1533.

13. Tomono, A., I. Muneo, and Y. Kobayashi. *A TV camera system which extracts feature points for non-contact eye movement detection*. i*SPIE vol. 1194. Optics, Illumination, and Image Sensing for Machine Vision IV*. 1989.

14. Ebisawa, Y. and S. Satoh. *Effectiveness of pupil area detection technique using two light sources and image difference method. 15th Annual Int. Conf. of the IEEE Eng. in Medicine and Biology Society*. 1993. San Diego, CA.

15. MacKenzie, I.S., *Fitts' law as a research and design tool in human computer interaction.* Human Computer Interaction, 1992. **7**: p. 91-139.

16. MacKenzie, I.S., A. Sellen, and W. Buxton. *A comparison of input devices in elemental* pointing *and dragging tasks*. in *CHI'91: ACM Conference on Human Factors in Computing Systems*. 1991. New Orleans, Louisiana.