

The Inspection of Very Large Images by Eye-gaze Control

Nicholas Adams*, Mark Witkowski and Robert Spence

Department of Electrical and Electronic Engineering

Imperial College London

Exhibition Road

London SW7 2BT

+44 (0)20 7594 6259

*n8vision@gmail.com; {m.witkowski, r.spence}@imperial.ac.uk

ABSTRACT

The increasing availability and accuracy of eye gaze detection equipment has encouraged its use for both investigation and control. In this paper we present novel methods for navigating and inspecting extremely large images solely or primarily using eye gaze control. We investigate the relative advantages and comparative properties of four related methods: Stare-to-Zoom (STZ), in which control of the image position and resolution level is determined solely by the user's gaze position on the screen; Head-to-Zoom (HTZ) and Dual-to-Zoom (DTZ), in which gaze control is augmented by head or mouse actions; and Mouse-to-Zoom (MTZ), using conventional mouse input as an experimental control.

The need to inspect large images occurs in many disciplines, such as mapping, medicine, astronomy and surveillance. Here we consider the inspection of very large aerial images, of which Google Earth is both an example and the one employed in our study. We perform comparative search and navigation tasks with each of the methods described, and record user opinions using the Swedish User-Viewer Presence Questionnaire. We conclude that, while gaze methods are effective for image navigation, they, as yet, lag behind more conventional methods and interaction designers may well consider combining these techniques for greatest effect.

KEYWORDS: *User interaction studies, Visual interaction, Eye-gaze control, Image space navigation.*

1. INTRODUCTION

Recent advances in technology ([6]) have improved our ability to detect and record a user's eye-gaze behaviour, and especially to do so with diminishing discomfort to the user. As a consequence the range and number of applications of this technology have increased rapidly. Two classes of application for eye-gaze detection can be identified. One, which has been of interest for many decades, exploits its investigative potential. Pirolli *et al* [19], for example, employed gaze detection to identify the manner in which users examine web pages, and Cooper *et*

al [5] have been able to associate image recognition and user preferences to the nature of eye-gaze trajectories. Other examples include the study of advanced interface design ([11]) and the manner in which visual search is conducted ([4]; [24]).

The other class of application addresses the potential of eye-gaze to control. Many schemes have been proposed, for example, in which the use of eye-gaze replaces or augments human motor processes in circumstances where the use of eye-gaze, alone or with augmentation, can go some way to ameliorating limitations on motor processes experienced by people with disabilities. Gaze control has been established for both disabled and able-bodied users for data input (e.g. [15], [9]), display inspection (e.g. [22]) and spatial navigation (e.g. [3]). In this paper we address another potential application for gaze control – the inspection of large images where gaze controls both zoom and pan.

1.1 Large Images

There are many situations in which very large images must be viewed in the execution of a variety of tasks, at levels of granularity ranging from an overview mode to a study of fine detail. They include the viewing of medical images to identify pathological anomalies (e.g. [16]; [23]), the inspection of large maps such as Google Earth¹ and NASA's World Wind² for purposes such as search and aerial surveillance, or the search of astronomical images for a variety of phenomena. Our study is directly concerned with earth images, specifically Google Earth, though we expect our results and conclusions to be relevant to large image inspection in general.



Figure 1: The gaze control system in use

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AVI'08, May 28-30, 2008, Napoli, Italy.

Copyright 2008 ACM, ISBN 1-978-60558-141-5.

¹ <http://earth.google.com/>

² <http://worldwind.arc.nasa.gov/>

Traditionally, the navigation (i.e., pan and zoom) of large images has been achieved by well-established means of interaction such as mouse or tracker ball control. In this paper we explore the use of eye-gaze – alone and in conjunction with other forms of interaction – to control the actions of panning and zooming in the context of exploring a large image in pursuit of a variety of tasks.

1.2 Related work

Most investigations addressing the use of eye-gaze to zoom into an image have been concerned with the activity known as gaze contingent zooming, and for data-rate reduction [7]. The automatic gaze-controlled expansion of a localised area of a display can help to overcome inaccuracies in gaze detection as well as to enhance the readability of small areas of text on a crowded display ([12]). An interesting extension of these investigations considers local stretching of an area identified by gaze, using a technique known as the bifocal display ([21]). Stretching can either be discrete ([8]) or continuous ([2], [19]). A word of caution, however, was sounded by Gutwin [10] who pointed out that continuous magnification actually slows down focus window targeting: it does so because “*the magnification lens makes windows appear to move in the direction opposite to pointer movement*”. Apart from these and similar studies, attention appears to have been confined to situations in which part of a display is at one or the other of two zoom levels, with the possible modification that a localised pre-determined stretching can take place.

By contrast, our study addresses the potential for gaze to control movement of an image through multiple zoom levels ranging, for example, from a view of the entire Earth to a view of a London street. Another word of caution was expressed by Zhai *et al* [25] and is, in a sense, fully acknowledged by our study. Zhai *et al* remarked that “*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....*”. The danger to which Zhai draws attention prompts any investigation of gaze control to compare the use of eye gaze alone with eye-gaze employed to augment other interaction modalities.

1.3 Google Earth

The size of some images that must be inspected can be enormous. For example, were the Earth to be imaged at one square metre over its entire surface, the resulting image would have the equivalent of about 5×10^{14} pixels. The Google Earth “image” is not yet at this resolution, but is still an impressive size and, moreover, as it is freely available on demand, it was chosen as the image for study. An additional advantage is that the data is convenient, it being both familiar and potentially intuitively navigated by anyone with a rudimentary knowledge of geography.

1.4 Goal of the investigation

Rather than study in some detail one selected means of eye gaze controlled inspection we elected to devise what we felt were a number of promising approaches and then compare them with methods in which eye gaze control was either not used or served to augment another interaction modality. In this sense the investigation was exploratory

with the primary aim of providing useful guidance to designers considering similar applications.

2. METHODS OF GAZE CONTROL

In common with other investigators (e.g., [8]), the study we report offers a comparison of eye-gaze on its own (STZ) with two augmented systems (HTZ and DTZ) and with a solely mouse-based system (MTZ) as control.

2.1 Equipment

The system design and investigations described here used LC Technologies (www.eyegaze.com) eye-gaze position monitoring equipment. Gaze position on screen is determined by comparison of the larger retinal (“pupil”) reflection, and small corneal reflection (figure 2, centre) from an axially mounted infra-red source on the eye-imaging camera mounted beneath the screen (figure 1 and figure 2, left). The eye image is available in a relatively small volume (approx. 100 mm^3) centred about 70 cm from the screen. The user’s eye must remain within this volume for the system to operate. Movements towards or away from the screen cause de-focusing (figure 2, right), which is detected by the system and may, within limits, be used to calculate screen to eye distance. The system requires a brief calibration procedure prior to use by each new user. Accuracy is quoted as 1° (about 15 screen pixels); gaze position readings are made 60 times a second.

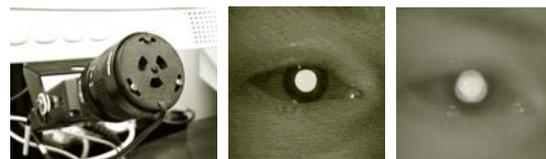


Figure 2: Eye camera (detail) and eye-images

2.2 Stare-to-Zoom (STZ)

In the STZ method all control of pan and zoom is by gaze position and timing. The overall strategy is illustrated in figure 3. The screen is divided into a central zoom region surrounded by a pan region. The extent of the pan region (100 pixels top and bottom, 150 pixels left and right, on a 1024×768 screen) was established empirically, allowing the user sufficient screen space to achieve uninterrupted panning. No zooming takes place while gaze is in the pan region.

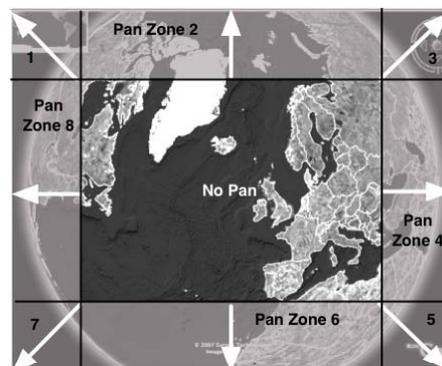


Figure 3: Screen panning regions

Sustained gaze in the central zoom region causes the image to zoom inwards. Normal saccades and fixations in the central region do not cause zooming, so the image may be

inspected in the usual way. However, extended stationary gaze (>420 ms) within the central region causes the image to zoom inwards at a comfortable rate. Zooming continues while the point of gaze remains stationary, as determined by a running calculation of the standard deviation of the screen gaze position. If gaze is fixed within the zoom region but offset from the centre, zooming is also accompanied by panning towards the screen centre. Once the feature of interest is at the centre of the screen, and while gaze is sustained on that feature, zooming continues uninterrupted until maximum resolution is obtained. Zooming outwards is achieved by glancing directly at the camera fixed to the base of the screen (Figure 1).

2.3 Head-to-Zoom (HTZ)

Figure 4 shows the essence of the HTZ technique. HTZ mixes eye-gaze controlled panning with head movement initiated zooming. It was suggested by the intuitive action of leaning forward to examine detail and leaning back to gain an overview (*c.f.* [14]). Small movements (about ± 40 mm) of the user's head, detected by the eye-gaze equipment employed, control zoom direction and rate. Pan is controlled by eye gaze fixation: movement of gaze away from the centre of the display causes movement of the image in the appropriate direction, as previously described for STZ.

Zooming is initiated by the system's calculation of eye to screen distance based on de-focusing (section 2.1). To assist the user, and provide fine control of the zooming rate, a non-linear transfer function was adopted, following comments from users during a pilot study [1] prior to the main evaluation reported here. This is illustrated in the insert (figure 4, top). A narrow "dead-band" where no zooming takes place allows for some positioning error by the user, after which slow zooming occurs. This becomes more rapid as the head moves further from the central point (trace III). Unfortunately, the system returns erratic distance estimates once the focus limits are reached and the user must learn to keep within the operating space.

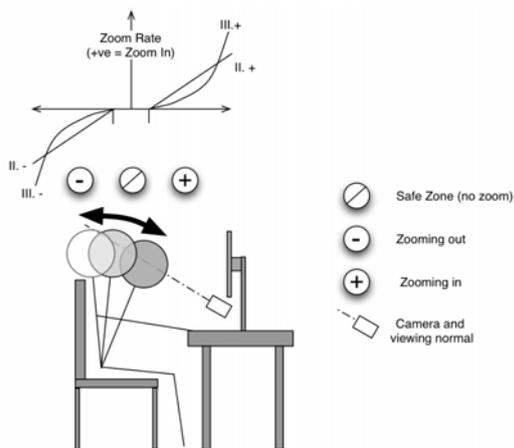


Figure 4: The HTZ zooming mechanism

2.4 Dual-to-Zoom (DTZ)

Following comments as to the care required to use the HTZ method during our pilot investigations, we additionally implemented a Dual-to-Zoom (DTZ) system, which combines gaze position panning input (as STZ and HTZ) with manual zooming using a mouse. The user

clicks the left mouse button to zoom in, and the right button to zoom out.

2.5 Mouse-to-Zoom (MTZ)

To provide some means of comparison with which to establish the benefit of gaze control a fourth method of controlling pan and zoom was implemented that did not use gaze control. As with dual control, left and right mouse buttons were used to initiate zooming in and out respectively, and mouse position to control image panning. It should be noted that this method is distinct from the standard mouse control of Google Earth, and it uses the same program control strategies as the other methods.

2.6 Design Issues

The eye-gaze software (supplied) and Google Earth run on a single computer. Control of Google Earth is achieved by a combination of the Google Earth COM API³ and emulation of keyboard-strokes and mouse clicks; direct view control through the API having been found to be too slow for this type of real time application.

The use of gaze control for cursor movement in this manner necessitates a filter to remove natural eye "jitter", which is highly disruptive to the viewer experience. We created a hybrid filter with a moving average component to stabilise high frequency movements during fixations and intentional looking, and a high-pass component to maintain responsiveness during rapid saccadic movements. Because Google Earth has an inherent centring motion during zooming, we also devised a tracking method that compensated for apparent movement across the screen during extended zooming operations in STZ mode (note also [17] and [18]).

An eye "icon" can be displayed on screen (top-right corner) to assist the user with their head positioning relative to the camera, although the system, by and large, provides its own feedback in terms of pan and zoom. In the trials described here this icon appeared only when tracking was lost, in conjunction with an audible warning, to assist the user to rapidly regain control of the interface.

3. EVALUATION METHOD

Having developed the four methods of pan and zoom control (HTZ, STZ, DTZ and MTZ) described previously, we performed a number of evaluations using Google Earth to assess and compare these methods. In the first evaluation, a search task, we were concerned to evaluate how effectively a user might use each of the methods to locate and identify known targets embedded within the larger image space. The search test also serves to confirm that each of the methods is capable of supporting this important class of browse activity, but also that each of the methods does not cause inadvertent gaze control actions (i.e. panning or zooming) that might interfere with the task. The metric for this task would be how many distinctive objects (in this case London buses) could be located within a test period of 90 seconds. We also monitored zooming activity during this task.

In the second evaluation, a navigation (or tracking control) task, users were asked to zoom directly into a specific location on the earth's surface starting from a "global" view. The purpose of this task was to evaluate and compare how effectively each of the four methods

³ <http://earth.google.com/comapi/>

provided a fine degree of control. The primary metric for this task was taken as the deviation of the gaze or control point from the optimal (“shortest” or “direct”) path between starting and ending screen images.

In order to perform these tests effectively, users were given a period of time with each of the methods they were to use to practice and browse the earth image freely (the “browse” period). Users were allowed to take the time they required for this familiarisation period. The time they took was noted. Each subject used one of the three gaze-controlled methods (STZ, HTZ or DTZ) and, as a control comparison, the mouse-only method (MTZ). Every participant completed a “subjective” questionnaire about their experiences using each method.

3.1 Familiarisation Activity

Each participant was allowed to use the selected method for an unspecified time (“browsing”) to gain familiarity with the control method, being asked to activate a stop key when they were ready to continue with other aspects of the experiment. Actually the time period was interrupted after 180 seconds, but the experimenter restarted the period if requested. The length time the participant elected to use the method was noted as an indication of the time required to become confident with the current method of control (figure 8).

3.2 Navigation/tracking Task

In this task users were asked to perform a straight line navigation from an image of the whole earth to a specific point on the earth’s surface by continuous zooming and, if required, corrective panning and zooming actions. Each subject observed an “ideal” pre-programmed navigation from a fully zoomed out view of the Earth to a specific geographical location. Following a repeat of this demonstration the subject was asked to undertake the same navigation task using the selected method. The task was repeated three times to detect any learning or improvement through practice.

Figure 5 shows the starting and required end locations for the task. Two easily recognised end locations were selected; the southern tip of the island of Sicily at the toe of Italy, and the northern most point of Madagascar, off the eastern coast of continental Africa. The starting location varied, but was always within 10° of the target location, which was therefore fully visible. To ensure that the desired target was maintained, and to reduce the cognitive load required to remember and identify the target, a small (constant sized) red dot was drawn over the target location at all times during the test.

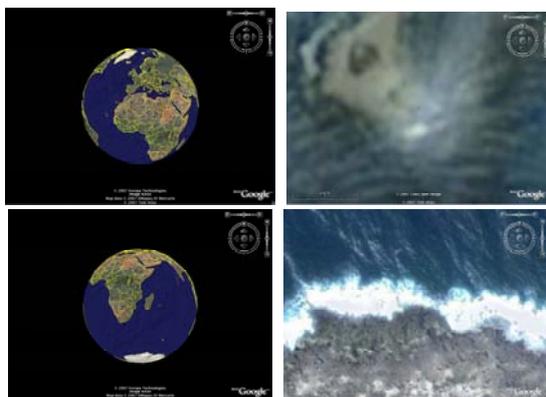


Figure 5: Start and end conditions for navigate task

Data was captured regarding eye gaze and the trajectory followed during execution of the task. The primary measure of performance is taken as the offset between gaze (or mouse) controlled cursor and the position on screen of the target. This is recorded at each time step (16.667 ms, 60 Hz) and measured in screen pixels (0.27 mm). An ideal control strategy would overlap the cursor and target to achieve optimal rates of zooming without the need to pan. Representative offset traces for each of the methods are shown in figure 11. Figure 10 shows the radial offset for a single instance. The time take to complete each instance of the task was recorded (figure 9).

3.3 Search Task

Another method of evaluating the potential of gaze for navigation involved search for a specific type of target. Participants were asked to search for as many London buses as possible within a given time constraint (90 seconds). Two alternate starting locations in central London were selected; a part of Regent Street (figure 6, left), and a part of the Strand just east of Charing Cross railway station (figure 6, right). Each starting image contains five buses, and there are many others in the surrounding area of both starting images (although they are not evenly distributed). Participants were able to zoom in and out or pan around to locate new buses, and were asked to press a clearly indicated key each time they identified a new bus. Participants had all lived in London for a significant time, but were reminded that some buses have white roofs! Again, appropriate data was captured, including the number of buses located, and the amount of zooming in relation to other activities. As with the navigation task, each subject used one of the gaze-controlled methods and, for comparison, MTZ.



Figure 6: The two search task start screens

3.4 Subjective Feedback

A third evaluation elicited the opinions of subjects regarding usability and acceptability. Each of the four methods was evaluated with a “subjective” questionnaire (figure 7), designed to give the experimenters insight into how the users found the experience of using the different methods of control. The 13 questions were designed to provide insight into four aspects of the methods’ usability: (a) “presence” (Q1, Q2, Q4, and Q5), the degree to which the user considered themselves immersed in the task; (b) “enjoyment” (Q3, Q6); (c) “sickness” (Q8, Q9, Q10, Q11, Q12), the degree to which they experienced adverse or nauseous sensations while using the system; and (d) “external awareness” (Q7 and Q13), the degree to which the users focussed their attention on the task in hand. Answers, given on a Likert-like numeric scale (shown 1 – 10 in figure 7), were combined in each category and analysed as a whole. The questionnaire used is derived from the Swedish User-Viewer Presence Questionnaire ([13]). Results were analysed with three non-parametric Wilcoxon tests.

Q1: To what extent did you think that the things you did and saw happened naturally and without much mental effort?
not at all 1 2 3 4 5 6 7 8 9 10 extremely

Q2: How natural was the interaction with Google Earth?
not at all natural 1 2 3 4 5 6 7 8 9 10 extremely natural

Q3: To what extent did you find Google Earth fascinating?
not at all fascinating 1 2 3 4 5 6 7 8 9 10 extremely fascinating

Q4: To what extent did you feel you were present in Google Earth?
not at all present 1 2 3 4 5 6 7 8 9 10 extremely present

Q5: How involved were you in the experience?
not at all involved 1 2 3 4 5 6 7 8 9 10 extremely involved

Q6: To what extent did you think it was enjoyable to interact in Google Earth?
not at all enjoyable 1 2 3 4 5 6 7 8 9 10 extremely enjoyable

Q7: To what extent did you focus your attention on the situation, rather than on other things?
not at all 1 2 3 4 5 6 7 8 9 10 extremely

Q8: I felt nauseous
not at all nauseous 1 2 3 4 5 6 7 8 9 10 extremely nauseous

Q9: My eyes felt strained
not at all strained 1 2 3 4 5 6 7 8 9 10 extremely strained

Q10: I had a headache
not at all 1 2 3 4 5 6 7 8 9 10 extremely

Q11: I had problems concentrating
not at all 1 2 3 4 5 6 7 8 9 10 extremely

Q12: I felt unpleasant
not at all 1 2 3 4 5 6 7 8 9 10 extremely

Q13: To what extent were you aware of things happening around you, outside Google Earth?
not at all aware 1 2 3 4 5 6 7 8 9 10 extremely aware

Figure 7: The subjective evaluation questionnaire

4. EXPERIMENTAL PROCEDURE

We asked 32 volunteer participants primarily drawn from the student population (9 female, 23 male, avg. age 24.6 years) to conduct a familiarising browse session, a navigation task and a search task using one of the three gaze control strategies (STZ, HTZ or DTZ), and an equivalent control session using the MTZ method. Participants were also asked to complete the questionnaire relating to their subjective experiences directly after using each of the two methods.

Each experimental session was conducted according to a pre-prepared script to ensure that the conditions under which the measurements were made were as constant as possible, although the experimenter responded to participant questions as necessary. The interaction method (MTZ vs. gaze method) and the two tasks (navigation and search) were counterbalanced. The schedule of activities is as follows:

1) **Introduction:** The experimenter settles the participant, obtains consent, and explains the reasons for the experiment. The experimenter introduces Google Earth and briefly explains that two methods of control will be used, one following the other.

2) **Set-up:** For the STZ, HTZ or DTZ methods the eye-gaze system requires calibration (section 2.1). MTZ does not require calibration. The participant is asked to complete the calibration routine, in which the gaze follows a dot through five screen locations. At this point the participant is asked to keep their head still for the duration of the experiment due to the limited operating volume of the equipment, and the role of the eye indicator (section 2.6) is explained.

3) **Browse:** Using the selected method, the participant is invited to browse with the system until they are ready to continue the experiment. The participant might try to locate the University site, or their home. An automatic timeout sounded at 180 seconds, but the experimenter would continue this activity if requested.

4a) **Search:** The search task was performed once using the selected starting point for a period of 90 seconds, at which time the system ceased operating.

4b) **Navigation:** The selected navigation task is demonstrated twice and then the participant is asked to “navigate as quickly as you can to the point you just saw, when you are done, please say ‘OK’”. Data recording is automatic.

The order of steps 4a and 4b are determined by the order in which participants used the mouse or eye-gaze method, according to the experiment schedule.

5) **Questionnaire:** The participant is handed the printed questionnaire sheet and asked to choose a value for each of the questions relative to the method they have just used, which the experimenter records.

Steps 2-5 are repeated with the second method. Next, the participant is asked to select the method they preferred.

At the conclusion of the experiment the participant is asked to compare the methods used and to make any additional comments they wished, which were recorded by the experimenter. Finally the participants are asked not to share details of the experiment with others and thanked for their help. No reward was made. The complete procedure took approximately 25 minutes per participant and the sessions were conducted over a period of three consecutive days.

5. RESULTS

This section presents the results of the investigations together with some preliminary analysis. Every participant undertook the MTZ method and sufficient results were obtained to have at least 10 instances each of the STZ, HTZ and DTZ methods. In two cases it proved impossible to achieve calibration with the eye gaze equipment, and these results were discarded.

5.1 Familiarisation Activity

Figure 8 summarises the elective time taken by the participant for each of the four methods.

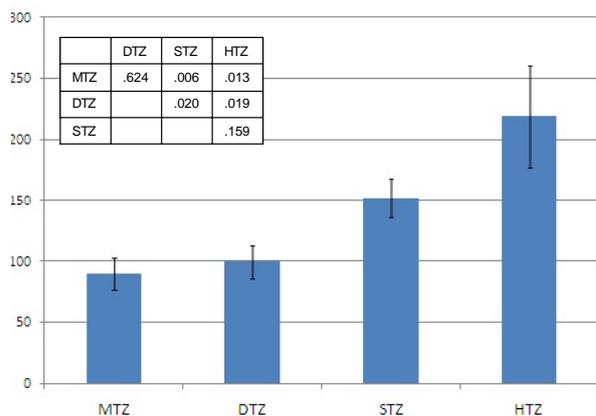


Figure 8: Mean browse time (s)

We note that users spent substantially less time to familiarise themselves with the MTZ and DTZ methods than with STZ and HTZ. We surmise that this is due to the

greater familiarity these technically aware participants will have for the mouse based methods, and, in particular, the novelty value associated with the HTZ method, which appears to take more time to get used to. This effect is apparent in later results also. Analysis of variance between means (inset, figure 8, and also figures 9, 10, 13 and 14) using a heteroscedastic *t*-test indicates, within the limits and applicability of this analysis, that there is no significant difference (at the 95% level, two tail) between MTZ-DTZ ($p = .624$) and between STZ-HTZ ($p = .159$), but significant differences between the other combinations.

5.2 Navigation Task

Data from the navigation task is analysed both in terms of the overall time taken to complete the task (figure 9), and in terms of the overall offset between the target point and the gaze/mouse controlled cursor place (figures 10, 11 and 12). Mean time to complete indicates that MTZ, DTZ and STZ are closely matched⁴, and show little variation between successive trials. MTZ, in particular, demonstrates little variability between participants. However, HTZ shows both a marked increase in mean time to complete and variability in standard deviation (error bars) between trials, but indicates considerable decrease in time between successive trials. It is clear that users found this method less intuitive than the others, and it is tempting to surmise that the rapid improvement is a reflection that users were able to quickly adapt to the requirement to hold the head still in the correct place to achieve smooth and constant zooming along the desired path. Analysis (as section 5.1) shows no significant variation between any of the final attempt times to complete, except between MTZ-DTZ ($p = .004$).

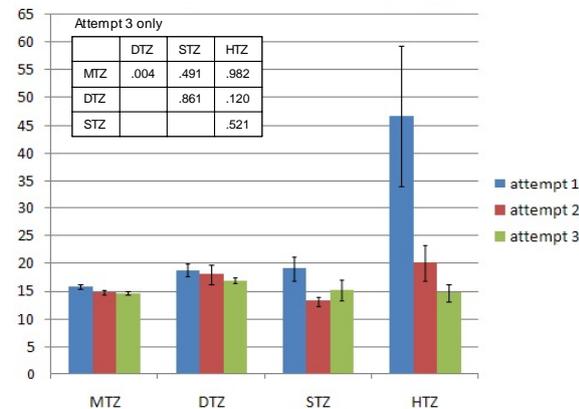


Figure 9: Navigation – Mean time to complete (s)

Figure 10 shows the mean offset from target for each of the four methods. The measure is Euclidean distance (in screen pixels) between target and controlled cursor point. Readings were taken 60 times a second. It may be seen that the MTZ method allows for precise control through positioning of the pointer manually with the mouse. The DTZ and STZ methods are broadly comparable, but less effective. The HTZ method again shows a substantial improvement over the three attempts, although again the final attempt is comparable to DTZ and STZ.

Figure 11 shows four individual traces for each of the tasks. These thumbnails are only intended to convey an

⁴ The MTZ and DTZ times may be overestimated. Some users reported that the zoom rate appeared slower with MTZ and DTZ, and this was later confirmed to be so.

impression of the effect, but they are selected to be representative of their type. In the typical MTZ offset trace (figure 11, top left) there is a small initial spike, corrected as the user quickly corrects the initial offset. Tracking is good under manual control until the very last stage when the image becomes highly magnified and offset to one side, which is immediately corrected. In the DTZ trace (top, right) note several peaks during the second attempt. These appear to be due to the user's gaze falling to one side, necessitating a corrective action to rotate the earth image, involving an element of overshoot. Figure 12 shows a radial plot (i.e. target at centre, gaze as offset) of the same trace, apparently confirming the overshoot hypothesis. The STZ plot (figure 11 bottom, left) indicates continued good control, whereas the HTZ (bottom, right) clearly illustrates the difficulty in control of the first attempt, and the rapid improvement in the second and third attempts in both time and accuracy. Note the MTZ third attempt mean is significantly (DTZ) or marginally different (STZ, HTZ) from the others.

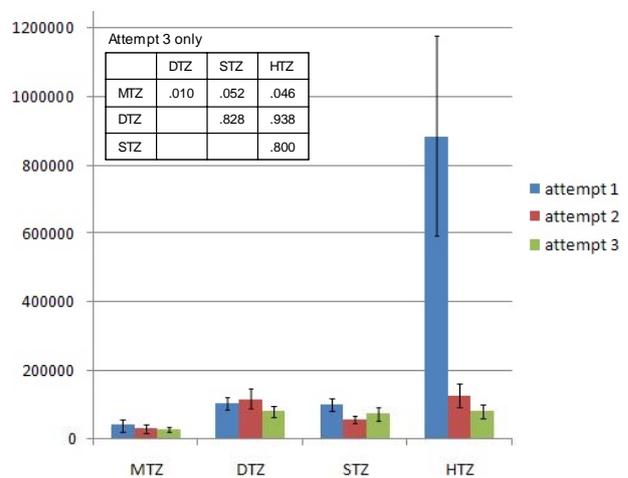


Figure 10: Navigation – Mean total offset from target (pixel distance)

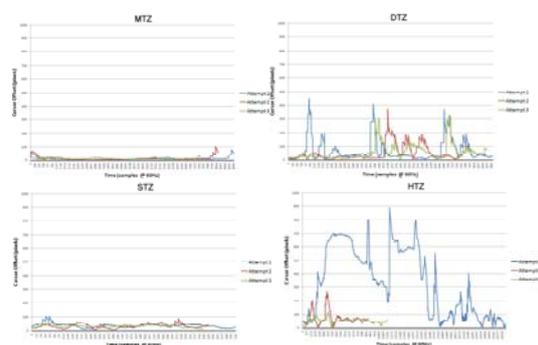


Figure 11: Navigation - Gaze point offset traces

5.3 Search Task

Figure 13 shows the average number of London bus targets found during the search task. The ordering of this result appears to confirm the previous findings, that MTZ offers the highest level of control, followed by DTZ, then STZ, with HTZ proving to be the least effective.

It is perhaps interesting to note from figure 14 that users consistently minimised their use of zooming (expressed as

a percentage of total time) when using the STZ method, perhaps indicating that the zoom strategy inherent in this method was less effective than the panning component.

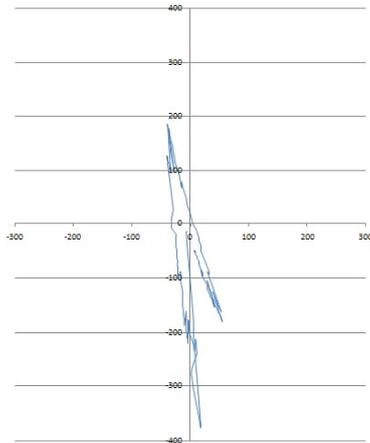


Figure 12: Navigation – Gaze point offset trace

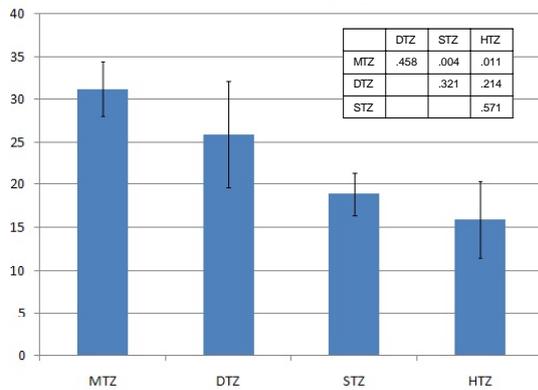


Figure 13: Search – Mean number of targets found

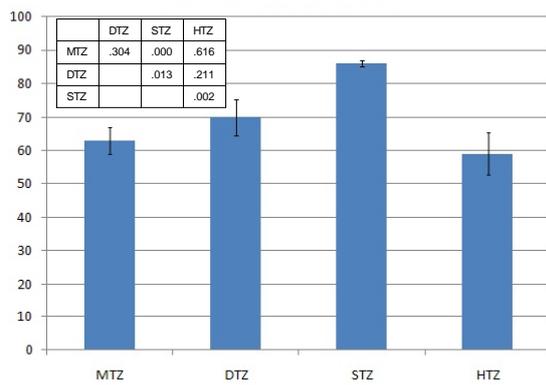


Figure 14: Search – Mean ‘non-zooming’ times (% total)

5.4 Subjective Results

When completing the Swedish User-Viewer Presence Questionnaire, participants reported on four measures: presence, enjoyment, external awareness and sickness. The questionnaire responses are summarised in Table 1.

The differences between users’ subjective experience in the three eye gaze treatments are revealed by using the non-gaze based MTZ treatment as a reference category. Three non-parametric Wilcoxon tests were performed by pairing the four dependent measures as reported for the mouse (MTZ) with the measures reported for each of the eye gaze methods.

When using the DTZ method, as compared to using MTZ, participants’ reports on the measure of presence were non-significant ($Z=-1.90$, ns). Reports on enjoyment were higher when using the DTZ method as compared to MTZ, although this result was only marginally significant ($Z=-1.30$, $p=0.06$). Finally, after interacting with the DTZ eye gaze method, participants’ reports on sickness tended to be higher than those given for MTZ but this result was marginally significant ($Z=-1.40$, $p=0.08$).

When compared to using the MTZ, participants’ reports on presence and enjoyment for the HTZ method were non significant. Conversely, reports on the measure of sickness were higher for the HTZ eye gaze method than the equivalent MTZ treatment ($Z=-2.52$, $p<0.05$).

The results yielded in the paired test for STZ and MTZ were identical to that found in the HTZ method: users’ reports on presence and enjoyment did not differ in the two conditions, while the sickness measure was higher in the STZ eye gaze method as opposed to the MTZ treatment ($Z=-2.31$, $p<0.05$).

7. SUMMARY AND CONCLUSIONS

We have devised several methods for allowing users to browse very large image spaces using either eye-gaze control as a sole method of input, or gaze control combined with other input modalities. We used the publicly available Google Earth image data set and application, as representative of a massive continuous image space, to perform a series of studies to evaluate the effectiveness of these methods relative to a mouse only method.

We were encouraged to find that each method was effective in traversing the image space, although none of the gaze based methods proved as efficient as the more conventional mouse based input. We were also encouraged by the generally positive comments from the test user group, admittedly young and technically aware, who were largely supportive and interested by the possibilities these methods offer. Although our test sample was smaller than we would have liked, we were pleased to note that there were no clear differences or disadvantages to these methods in relation to our “presence”, “enjoyment” and “external awareness” criteria. However, each of the gaze methods scored poorly on the “sickness” criteria, and this

Table 1: Summary of subjective responses

	DTZ (N=9)		HTZ (N=10)		STZ (N=11)		MTZ (N=30)	
	<i>M</i>	<i>S.D.</i>	<i>M</i>	<i>S.D.</i>	<i>M</i>	<i>S.D.</i>	<i>M</i>	<i>S.D.</i>
Presence	28.89	8.01	24.40	5.74	27.09	6.25	25.40	5.92
Enjoyment	16.89	3.52	13.60	5.56	15.45	4.08	13.73	3.91
External awareness	12.89	2.47	12.10	2.56	10.64	2.80	11.67	2.86
Sickness	12.00	4.15	16.50	8.85	12.00	3.82	7.90	4.11

is a cause for concern. We believe, however, that these are in part due to the relatively short period of familiarisation, and to some limitations inherent in the equipment (notably with the zoom range for HTZ), which might be overcome by expected advances in technology.

The naturalistic search task gives greater variability in results compared to a controlled artificial task, but is more consistent with the study aims. We also note that the methods are more effective in some tasks than others, this requires further investigation. We would also like to undertake a longer study to determine the effects of user familiarisation with each of the novel control methods.

Although effective in its own right, and echoing Sibert & Jacob's [20] view that "*Eye gaze interaction is a useful source of additional input and should be considered when designing interfaces in the future*", we suspect that eye-gaze control will also serve well as a way of augmenting more conventional input methods as indicated by our HTZ and DTZ methods. It has great potential to make interfaces more responsive and better able to anticipate the intentions of increasingly sophisticated interface users.

Assuming modest improvements in eye-gaze measurement technology and techniques – as well as greater availability – we are encouraged that both "hands-free" methods (STZ and HTZ) offer a viable image control and search method, notably for those with severe motor disability, but also for those who routinely monitor and search large image spaces and wish to use their hands for other tasks, such as data entry. Clearly such navigation methods might equally be applied to scanning and traversing three dimensional image sets, such as tomographic scans or architectural designs, and this remains a future task for investigation.

8. ACKNOWLEDGEMENTS

The authors would like to express their thanks to Asimina Vasalou (Imperial College London) and Päivi ("Curly") Majaranta (University of Tampere, Finland) for their generous assistance during the development of this paper.

9. REFERENCES

- [1] Adams, N., Witkowski, M. and Spence, R. (2007) The Exploration of Large Image Spaces by Gaze Control, Proc. COGAIN-07, 78-81.
- [2] Ashmore, M., Duchowski, A. T., & Shoemaker, G. (2005) Efficient eye pointing with a fisheye lens, Proc. Graphics Interface GI-2005, 203-210.
- [3] Bates, R. and Istance, H. (2005) Fly Where You Look: Enhancing Gaze Based Interaction in 3D Environments, Proc. COGAIN-05, 30-32.
- [4] Bertera, J.H. and Rayner, K. (2000) Eye Movements and the Span of the Effective Stimulus in Visual Search, *Perception & Psychophysics*, **62**(3), 576-585.
- [5] Cooper, K., de Bruijn, O., Spence, R. and Witkowski, M. (2006) A Comparison on Static and Moving Presentation Modes for Image Collections, Proc. AVI-06, 381-388.
- [6] Duchowski, A.T. (2003) *Eye Tracking Methodology: Theory & Practice*. Springer-Verlag, London, UK.
- [7] Duchowski, A.T., Cournia, N. and Murphy, H. (2004) Gaze-Contingent Displays: A Review, *Cyberpsychology & Behavior*, **7**(6), 621-634.
- [8] Fono, D. & Vertegaal, R. (2005) EyeWindows: evaluation of eye-controlled zooming windows for focus selection. Proc. CHI-05, 151-160.
- [9] Gips, J. and Olivieri, P. (1996) EagleEyes: An Eye Control System for Persons with Disabilities, Proc. 11th Int. Conf on Technology and Persons with Disabilities, 13pp.
- [10] Gutwin, C. (2002) Improving Focus Targeting in Interactive Fisheye Views, Proc. CHI-02, 267-274.
- [11] Jacob, R.J.K. (1995) Eye Tracking in Advanced Interface Design, in: Barfield, W. and Furness, T.A. (eds.) *Virtual Environments and Advanced Interface Design*, New York:Oxford University Press, 258-288.
- [12] Kumar, M., Paepcke, A. and Winograd, T. (2007) EyePoint: Practical Pointing and Selection Using Gaze and Keyboard, Proc. CHI-07, 421-430.
- [13] Larsson, P., Västfjäll, D., and Kleiner, M. (2001). The Actor-observer Effect in Virtual Reality Presentations. *CyberPsychology and Behavior*, **4**(2), 239-246.
- [14] Lepinski, G.J. and Vertegaal, R. (2007) Using Face Position for Low Cost Input, Long Range and Oculomotor Impaired Users, Proc. COGAIN-07, 71-73.
- [15] Majaranta, P. and Riih , K.J. (2002) Twenty Years of Eye Typing: Systems and Design Issues, Proc. ETRA-02, 15-22.
- [16] Mello-Thomas, C. (2003) Perception of Breast Cancer: Eye-Position Analysis of Mammogram Interpretation, *Acad. Radiol.*, **10**, 4-12.
- [17] Miniotas, D. and Špakov, O. (2004) An Algorithm to Counteract Eye Jitter in Gaze-Controlled Interfaces. *Information Technology and Control*, **1**(30), 65–68.
- [18] Miniotas, D., Špakov, O. and Scott MacKenzie, I. (2004) Eye Gaze Interaction with Expanding Targets, Proc. CHI-04, 1255–1258.
- [19] Pirolli, P., Card, S.K. and van der Wege, M.M. (2000) The Effect of Information Scent on Searching Information Visualizations of Large Tree Structures, Proc. AVI-00, 161-172.
- [20] Sibert, L.E. and Jacob, R.J.K. (2000) Evaluation of Eye Gaze interaction, Proc. CHI-00, 281-288.
- [21] Spence, R. and Apperley, M.D. (1982): Data Base Navigation: An Office Environment for the Professional. *Behaviour and Information Technology*, **1**(1), 43-54.
- [22] Starker, I. and Bolt, R.A. (1990) A Gaze Responsive Self-Disclosing Display, Proc. CHI-90, 3-9.
- [23] Tiersma, E.S.M., Peters, A.A.W., Mooij, H.A. and Fleuren, G.J. (2003) Visualising Scanning Patterns of Pathologists in the Grading of Cervical Intraepithelial Neoplasia, *J. Clin. Pathol.*, **56**, 677-680.
- [24] Zelinsky, G.J. and Sheinberg, D.L. (1997) Eye Movements During Parallel-Serial Visual Search, *J. Exp. Psychol.: Human Perception and Performance*, **23**(1), 244-262.
- [25] Zhai, S., Morimoto, C. and Ihde, S. (1999) Manual and Gaze Input Cascaded (MAGIC) Pointing, Proc. CHI-99, 246-253