Fly Where You Look: Enhancing Gaze Based Interaction in 3D Environments

Richard Bates & Howell Istance

De Montfort University
The Gateway
Leicester
UK
{richard.bates}{howell.istance}@cogain.org

Mick Donegan & Lisa Oosthuizen

ACE Centre
92 Windmill Road
Oxford
UK
{mick.donegan}{lisa.oosthuizen}@cogain.org

Abstract

This paper examines and seeks to enhance gaze based pointing and interaction in virtual 3D environments. Interaction in 3D virtual environments can offer motor impaired users entertainment, rehabilitation training, collaborative activities with users in remote places, and the opportunity to experience a sense of place afforded by remote locations, but these activities are only feasible if gaze based interaction is both efficient and satisfying. This paper finds that gaze based pointing has lower efficiency and satisfaction than a baseline of hand based pointing in both 2D and 3D environments, and that eye based pointing performance is highly dependent upon target object sizes or distances from the user. This paper then examines work in 2D and shows that enhancing interaction with the ability to 'fly' into target objects in 2D environments greatly increases the performance of gaze based pointing, such that it approaches the performance of a hand mouse. This enhanced ability to 'fly' toward objects is then extended and applied to 3D virtual environments for hand and gaze based pointing. The predictions made from 2D environments were found to be valid for 3D environments, with the ability to 'fly' toward target objects greatly increasing the performance of gaze based pointing such that it rivals the performance of hand based pointing. In addition, initial attempts at 'intelligently' controlling the 'fly' enhancement are then evaluated and found to further increase the performance of gaze based pointing. The paper shows that gaze based pointing in 2D and 3D environments can be effective such that with further improvement, the performance of gaze based pointing may equal that of hand based pointing, enabling users with a motor disability to interact effectively in 2D and especially 3D environments.

1 Introduction

Gaze based interaction in 3D virtual environments has much to offer motor impaired users such as entertainment, rehabilitation training, collaborative activities with users in remote places, and the opportunity to experience a sense of place afforded by remote locations. It is important that, for these users, interaction devices and techniques are provided that are both efficient in utilizing the residual capabilities of motor impaired users and also do not require undue effort or impose undue workload. It has been shown (Donegan 1999) how important it is for the self-image and motivation of motor impaired users to utilize assistive technology both for work and leisure. Without assistive technology, many of these users find it difficult or even impossible to achieve success in these areas. For a significant number of these, eye control offers the potential to achieve success in the most efficient, effective and satisfying way.

Gaze-based pointing devices have been used for inclusion and interaction with 2D graphical environments for some time and can offer access to these environments for people who may have few other choices of interaction device. However, anecdotal evidence suggests that these pointing devices are an inefficient and unsatisfying means of interaction and manipulation when compared to accepted baseline devices such as the desktop hand mouse. This is due to the inaccuracy of eye based pointing making direct pointing within environments difficult for disabled users who use eye gaze devices. Work has suggested that much of this inefficiency and low user satisfaction with eye based pointing can be greatly resolved by the use of supporting 'fly' devices to aid interaction (Bates 2000, Bates & Istance 2002). This paper examines results gained from 2D environments in enabling effective eye based interaction and then extends these findings to enhance gaze based interaction with 3D virtual environments.

2 Previous work

2.1 Gaze Interaction in 2D and 3D Environments

One of the goals of 3D virtual environments has been to enable users to apply the natural skills that they use in the real-world to their interaction with the virtual world. One of the advantages claimed for interaction based on gaze for 2D interaction is that it is more natural than the usual hand-based techniques, which involve a mouse or a joystick (MacKenzie et al. 2001, Jacob 1995, Jacob 1991). Also eye based interaction in 2D contexts has been expected to be more efficient and faster means of pointing than other devices (Edwards 1998, Jacob 1995, Salvucci & Anderson 2000, MacKenzie 1992). Gaze is more direct as it removes the need for a user to first look at a target and then to maneuver a pointer to it by means of a series of coordinated hand movements. Gaze should therefore be highly suited to interacting with 3D virtual environments, with its promise of enhanced naturalness and efficiency if the limitations of eye-based interaction, found in the context of 2D interaction, can be overcome.

An experiment where hand-based pointing, using a joystick, and eye-based pointing were compared for targets of different sizes within a virtual environment has been conducted (Asai et al 2000). Here they reported that pointing was 10 times faster on average across all target sizes for eye-based pointing. They note though that lack of familiarity with the joystick may have been responsible for some of the longer times recorded in this condition. A significant example of the early use of gaze in interacting with 3D virtual worlds was the 'self-disclosing' display (Starker & Bolt 1990). Eye-gaze was used to assign an 'index of interest' to objects in a virtual world presented on a desktop display. The system responded dynamically to changes to the assigned indices of interest by adapting a spoken commentary about the world to reflect either a particular interest in one object, or an interest spread across a group of objects. This technique was subsequently used as a gaze-based interaction technique within a virtual environment (Tanriverdi & Jacob 2000). This entailed that objects in a virtual environment receiving a sufficiently high index of interest increased in size and revealed their internal structure. This is not a 'zoom-in' of the observer towards the object of interest, but rather a 'zoom-out' of the object towards the observer. Tanriverdi and Jacob compared the performance of this technique with that of a hand-based pointing technique for a task that involved selecting objects containing various stimuli in an environment. These targets could either be reached in the hand pointing condition while the user was stationary, or by the user moving 5 to 15 inches forward. A performance advantage in favor of gaze-based pointing over hand-based pointing was found, but only for targets which required the subject to move forward. The gaze pointing condition enabled selection of the same targets without the user having to move. This presumably afforded some advantage to the gaze condition regardless of any inherent difference between eye and hand as a pointing modality.

A similar experiment was conducted (Cournia et al 2003) but used the 'ray-casting' interaction technique for both hand and eye conditions. There was no 'zoom-out' for any target, only a fade to reveal inner structure. They found a significant difference in favor of hand-based pointing for targets at distances where Tanriverdi & Jacob had found advantages for eye-based pointing. These two pieces of work suggest that the benefits of eye-based pointing in virtual environments are contingent on the apparent size of target (assuming targets located further away are smaller) and on the interaction technique used.

2.2 Enhancing Eye-based Pointing in 2D and 3D Environments

Interaction in 3D environments can broadly be characterised as object manipulation, navigation and application control (Hand 1997). Zooming or flying in towards an object can be seen both as a navigation technique and as an object manipulation technique. Temporarily zooming-in on an object of interest to select it means it is easier to select objects with an inaccurate device (Bates & Istance 2002), although it is important to be able to zoom back out to the original position to prevent loss of context and orientation. Without the return to the original position from where the zoom action was initiated, zoom becomes a fly navigation technique ('fly where I look'). 'Intelligent flying' can utilise a similar technique to the 'index of interest' where initiating a 'fly' action assumes the target to be that object with the highest index of interest. Additionally the fly can stop at a reasonable distance in front of the assumed object of interest so that it does not fill too much of the visual field, or indeed, to ensure that the user does not fly straight through the object. Moving the gaze point during the fly can either effect small corrections to the flight path or indicate the intended object to fly is not that which has been assumed by the system.

3 Flying in 2D Environments

3.1 An Experiment in 2D

An experiment was conducted with six users in a complex 2D GUI test environment to measure the effect of providing a basic ‘fly’, enhancement (Bates & Istance 2002). Hand and eye based pointing devices, or hand and eye mice, with and without the fly enhancement were used to manipulate objects of four angular sizes based of the angle the objects subtended from the eye of the user. These ranged from 0.30 to 1.20 degrees visual angle at 60cm from the screen. Interaction typically lasted for 20 minutes and incorporated 150 test tasks for each user. The objective efficiency (based on time and quality of interaction metrics) and subjective user satisfaction (based on ratings of workload, comfort and ease of use) of the manipulation were measured. These metrics were derived from the ISO9241 part 11 standards and the ESPIRIT MUSiC metrics method and previously validated. For this experiment the basic fly enhancement was under the full control of each user via micro switches rather than under intelligent software control in order to determine at what distance of fly into the environment the users opted to stop the fly and start manipulation. The subjective metrics were obtained from a set of rating scales, similar to the NASA TLX.

3.2 2D Interaction Results

3.2.1 Efficiency

Figure 1 shows the task efficiency metrics for the hand mouse, the eye mouse without fly and the eye mouse with the fly enhancement for each of the target size categories. The results for the eye mouse without the fly-enhancement (Figure 1, ‘Normal’ mode) showed near-unusable efficiencies for the smallest object at 18% task efficiency, with efficiency increasing, as expected, with increasing target size. It was notable that even with the largest object size that eye mouse efficiency did not approach the hand mouse efficiency at the smallest target size.

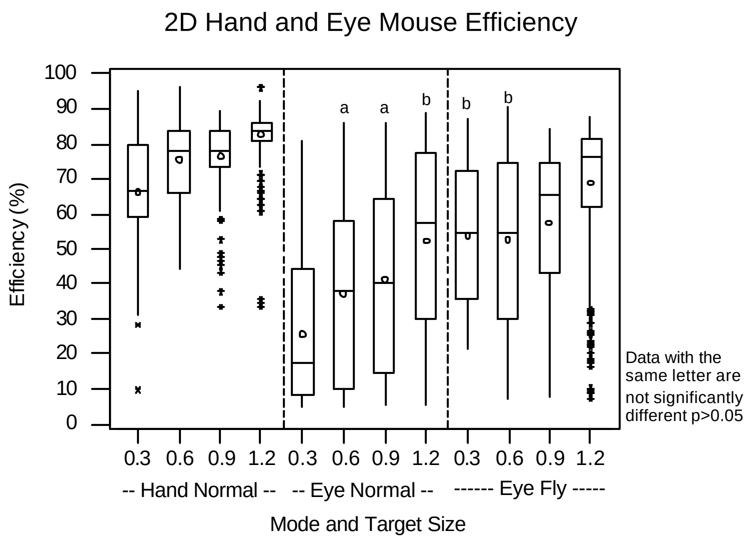

Figure 1: Device Efficiency and the Effect of Object Size And Flight in 2D

Examining the eye mouse enhanced with the ‘fly’ soft device (Figure 1, ‘Fly’ mode) showed dramatic increases in efficiency for all object sizes, with the non-fly performance on the largest target being equaled by the ‘fly’ efficiency on even the smallest target size. Note that this improvement includes the detrimental overhead of manipulating the ‘flying’ during interaction; with the measurement of efficiency including penalties for time spent controlling the fly and errors in controlling the fly distance.

Examining the efficiency results by object size showed a clear relationship between object subtended angle and use of fly, with fly being increasingly used as object angles decreased (Table 1). For eye-based pointing the effective object size after flying was near constant at 1.7 degrees visual angle at 60cm for all objects. These results gave the basis for calculating a 'smart fly' automatic stopping distance in front of an object for the device.

Table 1: Eye Mouse Effective Flown Object Subtended Angles at 2D Fly Stop Point

Original angle subtended	Enhanced eye 'flown' angle
0.3°	1.6°
0.6°	1.6°
0.9°	1.8°
1.2°	1.7°

3.2.2 Satisfaction

Examining the satisfaction ratings (Figure 2) showed high levels of workload for the eye mouse, although with a slight decrease in rating for the enhanced mode. This was encouraging as the decrease was in spite of the overhead of additional workload controlling the fly. Comfort levels remained unchanged, which was to be expected, as the device was essentially little different in comfort to use. There was a slight increase in device ease of use for the fly condition. It is likely that greater efficiencies and larger decreases in workload and increases in comfort and ease of use will arise from an automated 'smart fly' facility where the user does not need to actively control the fly.

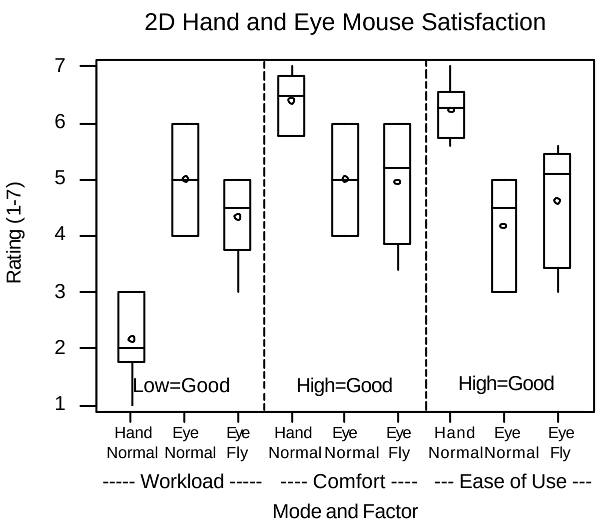

Figure 2: Device Satisfaction and the Effect of Flight in 2D

In conclusion, the provision of a zoom or fly enhancement greatly increases the efficiency of eye-based pointing on a 2D GUI, without adversely affecting subjective ratings of comfort or workload. Furthermore, when the user has full control over the extent of the zoom, it is used such that targets of originally different sizes all subtend approximately 1.70 degrees of visual angle at 60cm.

4 Flying in 3D Environments

4.1 An Experiment in 3D

An experiment was conducted in a virtual environment to examine the extent to which hand-based and eye-based pointing would benefit from a similar fly enhancement to that previously examined with the 2D zoom enhancement. The hand-based and the two eye-based conditions (one with fly enhancement and one without) all used ray casting as the interaction technique, as did Cournia et al. (2003) previously.

Unlike previously reported work, which has used an immersive head mounted display, this experiment was conducted in a reality centre located at De Montfort University, equipped with passive stereoscopic images across an 8-metre wide 1500 cylindrical screen. In all conditions, users were seated 6 meters away from the curved screen. A desk mounted SMI RED eye tracker was used for eye-based pointing, and a desktop mouse was used for the hand based pointing in order to enable direct comparison with the 2D environment (Figure 3 shows a user seated in the environment, with the eye tracker located on the table in front of the user).

Figure 3: Eye Gaze Pointing in the Virtual Environment

4.2 Experimental Task

The task was to select one of a group of virtual students in a virtual lecture theatre. The required target to select was indicated by a hat appearing on the student. Students at the back of the lecture theatre subtended the smallest visual angle in the experiment. The students on the four rows subtended the same four visual angle sizes as in the first 2D experiment. An illustration of the experiment in the VE is shown (Figure 4). Here the upper left frame shows the VE before a fly is invoked. The end of a vector from the user's eye through the VE can be seen as a grey cube in the middle of the picture, with the desired target object signified as the person wearing a hat. The upper right frame shows the subject 'flying' rapidly (12.5m/s) toward the target object. Finally, the lower frame shows the target object being selected with the hat flying off and away from the target. Six users took part in the experiment, with interaction typically lasting for 20 minutes and incorporating 144 test tasks for each user. As before, the objective efficiency and subjective user satisfaction of the manipulation were measured.

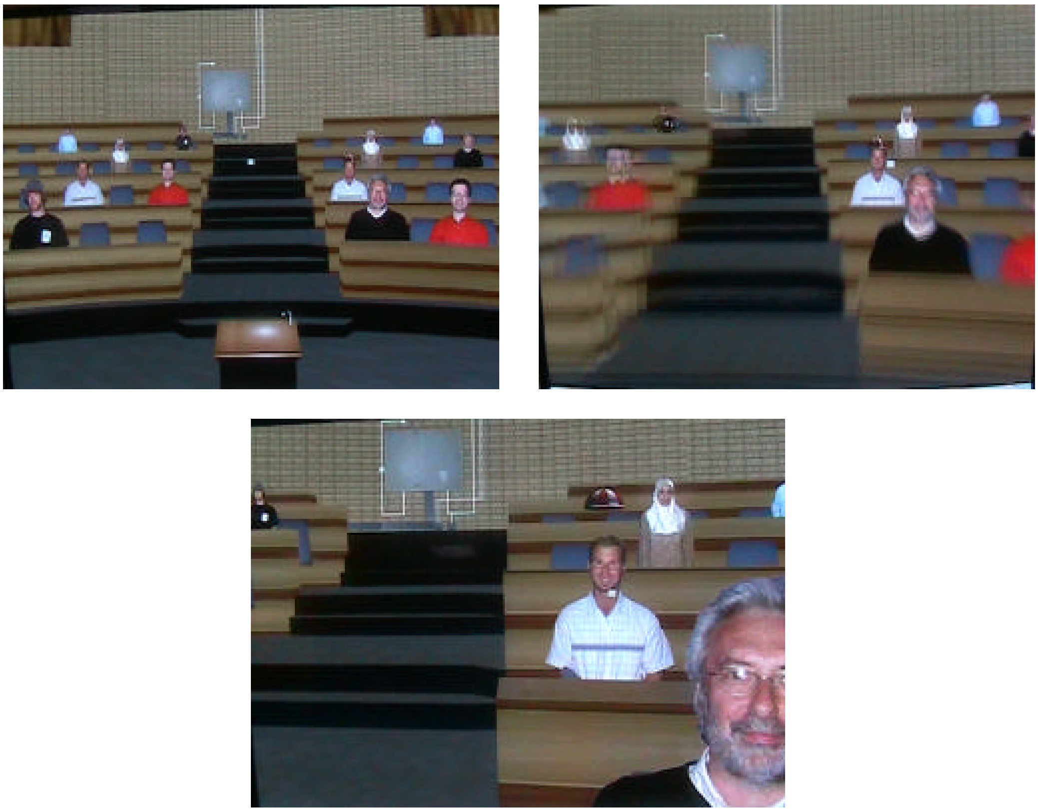

Figure 4: Flight and Object Selection in the Virtual Environment

4.3 3D Interaction Results

4.3.1 Efficiency

Following the procedure of the 2D experiment, the results were broken down by target size visual angle for all devices. For the hand device, objective task efficiency and subjective user satisfaction (Figures 5 and 6, left hand and centre sets of data only) showed that the device was highly dependent on object size, with poor performance for the smaller object sizes. Comparing the hand mouse efficiency in a 2D environment (Figure 1) to that in the 3D environment (Figure 5) showed a marked drop in performance for all object sizes in the 3D environment. Enhancing the hand mouse with the 'fly' soft device (Figure 5, 'Fly' mode) showed large and significant increases in efficiency for the smallest three object sizes, with no significant improvements for the largest object size, where fly was rarely used. The efficiency results for the eye mouse without the fly-enhancement (Figure 6, 'Normal' mode) showed extremely low efficiencies for the smallest objects, with efficiency increasing, as expected from the 2D results, with increasing target size. As with the 2D results, the eye mouse showed lower performance than the hand mouse, although the differences were considerably reduced between the devices in the 3D environment. There is no difference for the smallest target size, but as target size increases hand pointing outperforms eye pointing. This difference is similar to that found by Cournia et al (2003) quoted earlier. With the fly enhancement (Figure 6, 'Fly' mode) the efficiency for all target sizes was increased, with object size now having only a minor effect on efficiency, and the eye mouse achieving near parity with the hand mouse. As before, the efficiency results gave the basis for estimating a 'smart fly' automatic stopping distance in front of an object for the devices (Table 2).

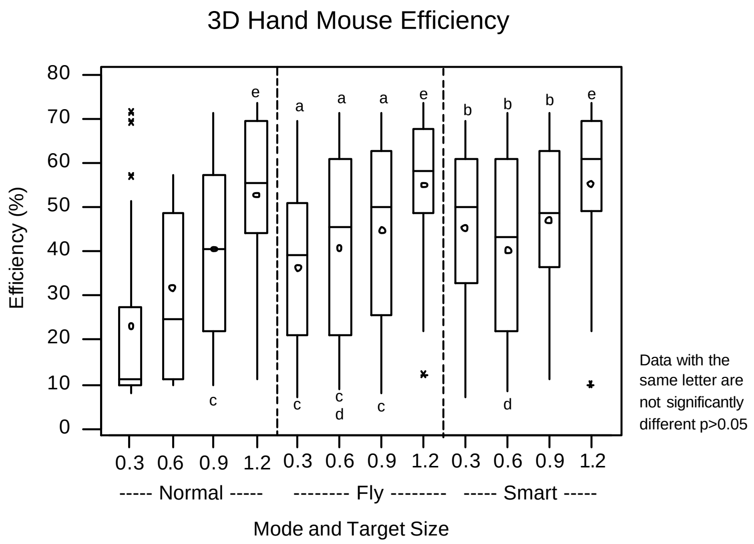

Figure 5: Hand Mouse Efficiency and the Effect of Object Size and Flight in 3D

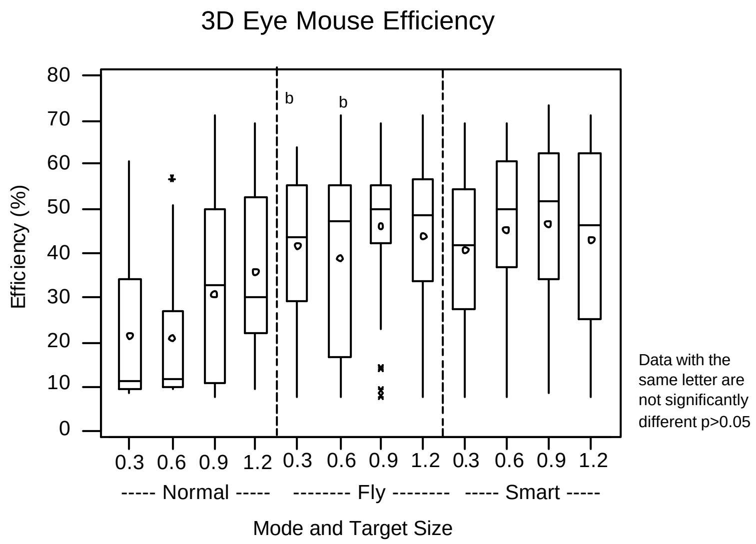

Figure 6: Eye Mouse Efficiency and the Effect of Object Size and Flight in 3D

Table 2: Hand and Eye Mouse Effective Flown Object Subtended Angles at 3D Fly Stop Point

Original angle subtended	Enhanced hand 'flown' angle	Enhanced eye 'flown' angle	Smart hand 'flown' angle	Smart eye 'flown' angle
0.3°	1.3°	2.2°	1.3°	2.1°
0.6°	1.4°	2.6°	1.4°	2.2°
0.9°	1.5°	3.3°	1.5°	2.2°
1.2°	1.9°	3.0°	1.6°	1.9°

4.3.2 Satisfaction

The subjective hand mouse satisfaction ratings (Figure 7) showed increased workload and lower ease of use in the 3D environment compared to the 2D environment. It was notable that the fly enhancement reduced this workload and increased the ease of use in the 3D environment. Examining the subjective eye mouse satisfaction ratings (Figure 8) showed improvement for all ratings for the enhanced mode. As with the hand mouse, it was notable that the fly enhancement reduced the workload and increased the ease of use ratings in the 3D environment.

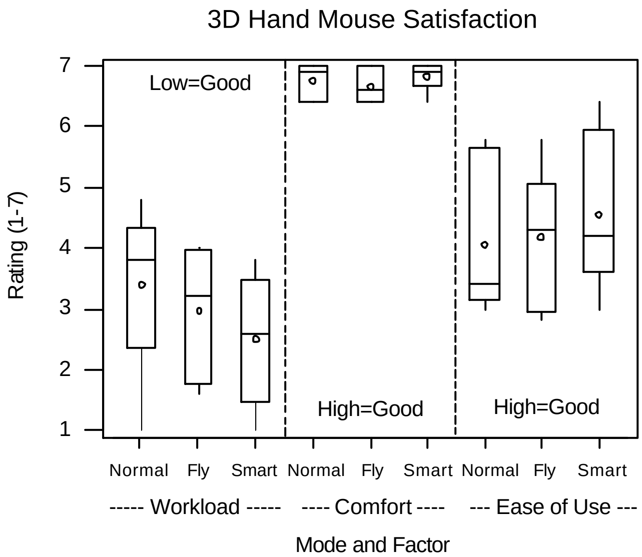

Figure 7: Hand Mouse Satisfaction and the Effect of Object Size and Flight in 3D

Figure 8: Eye Mouse Satisfaction and the Effect of Object Size and Flight in 3D

5. The Benefits of 'Intelligent' Flight

5.1 An Experiment with 'Intelligent' Flying

Based on 3D experimental results, a 'smart' stopping distance was set at a distance where the visual angle of an object would subtend 2.4°; a compromise to give greatest ease of manipulation without overly enlarging objects and potentially disorienting the user, but also flying sufficiently close to give ease of selection. To do this the fly enhancement was modified to 'intelligently' monitor the point of interest in the environment, and the required fly distance toward target objects based on their apparent size, with the fly automatically being stopped when the object subtended a 2.4° visual angle. After manipulation, the user returned to the original starting point by initiating a automatic fly back. This gave the first elements of an 'intelligent' fly, with the fly automatically stopping at an optimal distance, and then returning to the pre-flight position. In the same manner as before, a trial was conducted with the intelligent fly enhancement using the hand and eye mice.

5.2 'Intelligent' Flying Experimental Results

The results for the 'intelligent fly' hand and eye mice were appended to the 'normal' and 'fly' results to aid comparison between the three modes of operation (Figures 5 to 8). Overall, device efficiency was essentially unchanged from the basic 'fly' mode, showing that although test subjects tended to fly closer to objects with the eye mouse (Table 2) than the estimated ideal distance, there were no performance benefits from doing so. In addition, the intelligent fly suggested reduced workload is possible in comparison to the basic fly mode.

6 Conclusions

The work reported has demonstrated that the benefits of enhancing eye pointing by zoom previously demonstrated in 2D interaction are also apparent in 3D interaction. Our results comparing eye and hand pointing in virtual environments accord with those reported by Cournia et al, and show a performance advantage for hand based pointing. Our work shows that this benefit is only apparent for larger target sizes, however, when a 'fly' enhancement is provided the performance levels of eye based pointing increase to a similar level to that of hand based pointing. Our initial attempts to go further and introduce a degree of 'intelligence' have indicated some success. The limited but promising results suggest that more effort is required to add further intelligence to interaction to gain performance benefits. The addition of an intelligent control based on optimal object subtended angles is currently under investigation and is expected to further enhance performance with eye based pointing in 3D environments, to enable the naturalness and efficiency benefits offered by this modality.

7 Future Perspectives

This work has indicated that gaze input in 3D virtual environments can be an effective means of interaction and control. This opens a range of new applications and perspectives for eye based control, where gaze could become the modality of choice in 3-dimensional environments, both virtual and real, for both able-bodied and disabled users. The eye naturally gazes at objects and areas where the user has most interest, and also compensates naturally if those objects move in relation to the user, or the user moves in relation to those objects, making gaze an effective form of pointing. In addition, since the eye exhibits very high angular velocities either between objects of interest, or when tracking an object of interest, it is also a very rapid form of pointing. This makes gaze potentially highly suited to mainstream applications such as gaming and first person target acquisition 'shoot them up' games for users with or without a disability. The natural effectiveness and speed of gaze based pointing also applies to traversal of complex 3D environments, where the eye would naturally maintain fixation on, and automatically maintain movement vector toward, the endpoint object during potentially rapid movement through a complex 3D environment, potentially making gaze the modality of choice for navigation in 3D environments. For example, this has been exploited in wheelchair navigation by gaze (Matsumoto et al. 2001 for example). Such mainstream application of gaze in 3D environments is currently under investigation by the authors, with experimentation in rapid eye based flight *through* virtual 3D environments.

8 Acknowledgements

This work would not have been possible without the support of the European Union and forms part of, and is funded by, the efforts of the European Commission 'Communication by Gaze Interaction' (COGAIN) fp6 Network of Excellence, to which the authors acknowledge their gratitude. The 'COGAIN' {www.cogain.org} Network of Excellence brings together diverse researchers from across Europe with the aim of enabling eye gaze based interaction and communication for people with disabilities. Within the Network, De Montfort University contributes research on eye gaze based direct interaction with graphical 2D and 3D interfaces. The ACE Centre (Aiding Communication in Education) is an independent charity whose role is to promote the use of appropriate technology to help young people with severe communication difficulties. To achieve this, it provides a range of services to benefit these young people and those supporting them including assessments, training, publications and research and development.

9 References

Asai, K., Osawa, N., Takahashi, H., Sugimoto, Y. Y., Yamazaki, S., Samejima, M. & Tanimae, T. (2000) Eye Mark Pointer in Immersive Projection Display, *Proceedings of IEEE Virtual Reality 2000 Conference*, p125.

Bates, R. (2000) Using Soft Devices as Enabling Translators - Standard GUI Interaction For All! *Proceedings of Human Computer Interaction, 2,* pp53-54.

Bates, R. & Istance, H. O. (2002) Zooming interfaces! Enhancing the performance of eye controlled pointing devices, *Proceedings of ASSETS 2002*, The Fifth International ACM SIGCAPH Conference on Assistive Technologies, July 8 - 10, 2002, Edinburgh, Scotland.

Cournia, N., Smith, J. D. & Duchowski, A. T. (2003) Gaze- vs. Hand-Based Pointing in Virtual Environment, in *Proceedings of SIGCHI 2003* (Short Talks & Interactive Posters), April 5-10, 2003, Ft. Lauderdale, FL.

Donegan, M. (1999) Computers and Inclusion - Factors for Success. ACE Centre Advisory Trust.

Edwards, G. (1998) New Software makes eye tracking viable: You can control computers with your eyes, *in Proceedings of CSUN Technology and Persons with Disabilities Conference*, California State University Northridge.

Hand, C. (1997) A Survey of 3D Interaction Techniques, *Computer Graphics Forum*, 16 (5) pp269-281.

Jacob, R. J. K. (1991) The use of eye movements in human-computer interaction techniques: What you look at is what you get, *ACM Transactions on Information Systems 9*, 3, pp152-169.

Jacob, R. J. K. (1995) Eye tracking in advanced interface design, *in Advanced Interface Design and Virtual Environments*, pp258-288, Oxford University Press.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction, 7,* pp91-139.

MacKenzie, I. S., Kauppinen, T. & Silfverberg, M. (2001) Accuracy Measures for Evaluating Computer Pointing Devices, *Proceedings of CHI2001*, ACM Press.

Matsumoto, Y., Ino, T., Ogasawara, T. (2001) Development of Intelligent Wheelchair System with Face and Gaze Based Interface, *Proceedings of 10th IEEE Int. Workshop on Robot and Human Communication (ROMAN 2001),* pp.262-267, 2001.

Salvucci, D. D. & Anderson, J. R. (2000) Intelligent gaze-added interfaces, *Proceedings of CHI 2000*, 2, ACM Press, pp273-280.

Starker, I. & Bolt, R. A. (1990) A Gaze-Responsive Self-Disclosing Display, *Proceedings of CHI '90*, ACM, pp3-9.

Tanriverdi, V. & Jacob, R. J. K. (2000) Interacting with eye movements in virtual environments, Conf. Human Factors in Computing Systems, *Proceedings of SIGCHI conference on Human Factors in computing systems*, The Hague, The Netherlands, pp265 – 272, ACM Press.