Evaluation of the Average Selection Speed Ratio between an Eye Tracking and a Head Tracking Interaction Interface

Florin Bărbuceanu[1], Mihai Duguleană[1], Stoianovici Vlad[2], and Adrian Nedelcu[2]

[1] Transilvania University of Brasov, Product Design and Robotics Department
[2] Transilvania University of Brasov, Electronics and Computers Department
{florin.barbuceanu,mihai.duguleana,stoianovici.vlad, adrian.nedelcu}@unitbv.ro

Abstract. For a natural interaction, people immersed within a virtual environment (like a CAVE system) use multimodal input devices (i.e. pointing devices, haptic devices, 3D mouse, infrared markers and so on). In the case of physically impaired people who are limited in their ability of moving their hands, it is necessary to use other special input devices in order to be able to perform a natural interaction. For the inference of their preference or interests regarding the surrounding environment, it is possible to take in consideration the movements of their eyes or head. Based on the analysis of eye movements, an assistive high level eye tracking interface can be designed to find the intentions of the users. A natural interaction can also be performed at some extent using head movements. This work is a compared study regarding the promptness of selection between two interaction interfaces, one based on head tracking and the other based on eye tracking. Several experiments have been conducted in order to obtain a selection speed ratio during the process of selecting virtual objects. This parameter is useful in the evaluation of promptness or ergonomics of a certain selection method, provided that eyes focus almost instantly on the objects of interest, long before a selection is completed with any other kind of interaction device (i.e. mouse, pointing wand, infrared markers). For the tests, the tracking of eyes and head movements has been performed with a high speed and highly accurate head mounted eye tracker and a 6 DoF magnetic sensor attached to the head. Direction of gaze is considered with respect to the orientation of head, thus users are free to turn around or move freely during the experiments. The interaction interface based on eye tracking allows the users to make selections just by gazing at objects, while the head tracking method forces the users to turn their heads towards the objects they want to be selected.

Keywords: eye tracking, head tracking, interaction metaphor, interaction interface, virtual reality, virtual environment, CAVE system.

1 Introduction

As vision is one of the most important communication channels, vast research has been conducted lately in the area of eye tracking. Sayings which date from the early

Luis M. Camarinha-Matos (Ed.): DoCEIS 2011, IFIP AICT 349, pp. 181–186, 2011.

ages state that eyes are the window towards mind. It is sometimes facile for relatives, friends or even strangers to guess someone's intentions just by looking at their eyes. Although they are input sensory channels, when they behave according to known gestures it is possible for some meaningful information to be transmitted through. This information is very useful in the implementation of assistive interaction interfaces, especially in the case of severely disabled people [1]. Attentive user interfaces (AUIs) take this information into consideration to infer user's intensions and preferences [2]. For the purpose of an increased degree of self-sufficiency in carrying out daily life activities [3], some heterogeneous environments like inhabited rooms can be controlled by disabled people through communication interfaces based on eye-tracking [4]. The users gradually shift their gaze towards a certain target until their preference is established [5]. The "cascade effect" discovered in 2003 [6] relate the gradual gaze shifts with the interest of users towards solving a given task. Due to this association between saccadic eve movements and interest, a quick determination of user's gaze at any time is essential for a consistent inference of user's interests [7]. The most important advantage of an object selection interface based on eye tracking is the promptness. Based on the fast analysis of eye movements, an attentive interface can make associations between sequentially gazed targets, time spent on each target or the path followed with the gaze, to guess the interest of the users or to determine if they are in a situation of uncertainty [2]. The *iTourist* system is able to analyze user's gaze very quickly and make associations between fixations points over the surface of an electronic map, providing a dynamic flow of information about what he/she appears to be interested in. It reacts as a very attentive humanlike guide, paying attention at all times to what the tourist is looking at [8].

2 Contribution to Sustainability

One of the most inconvenient aspects of an object selection method based on eye tracking is the imprecision of gaze estimation. Gaze position accuracy of the eye tracking system used in our experiments is between 0.5° and 1°. This means that objects located far from the user are more likely to be missed by the estimated direction of gaze, especially if they are small.

If head position and orientation can be tracked, an estimation of the user's gaze can be considered along the orientation of head. Object selection can be performed very precisely in this case if a visual feedback of head orientation is presented to the user. Provided that head movements can be successfully used as an interaction method in a virtual environment, we have conducted a set of experiments to compare the promptness of a head tracking interaction metaphor with respect to a fast eye tracking interaction metaphor. However, the head tracking method is not as fast as the one based on gaze tracking and the time delay between selections made through the two interaction interfaces is rather intuitive. The purpose of this paper is to discuss the results obtained in a series of experiments, regarding the selection speed ratio between the two interaction interfaces mentioned. It can be an instrument in the evaluation of the stress exerted on the user when using head movements to make selections. Any unnecessary load within the interaction metaphor can lead in time to fatigue, especially if the interaction metaphor is complex. Head movements are complex and require more spatial coordination, so in this case it is essential to know the amount of time users spend on the selection procedure.

3 Design of Experiments

Calibration of the eye tracker used in our experiments, ASL H6-HS-BN 6000 Eye-Track model, an accurate and high speed head mounted system, is typically made on a normal desktop screen (Fig. 1), by sequentially gazing at each one of the green points.

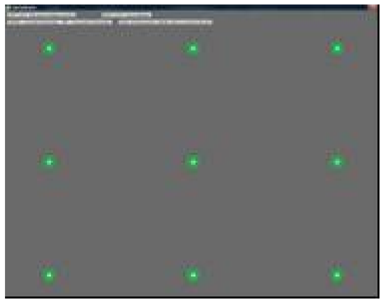


Fig. 1. Eye tracker calibration points on a desktop PC

Our experiments have been conducted on a large projection screen, normally used for visualization of stereoscopic scenes. The nine calibration points were displayed in a similar fashion as on the small desktop screen (Fig. 2), covering the visual field of the user. During the experiments, a black background was chosen for the projection screen in order not to distract the subjects in any way from the task they were assigned to (Fig. 3). On the black background a green square is displayed sequentially in 9 locations on the screen in a random fashion, so that subjects can't anticipate the next location where it will be displayed. Since the accuracy of the selection is not the subject of these experiments, the locations on the screen of the 9 points were chosen at 0.5 m one from another in order to enable a facile discrimination of each gazed objects.

Fig. 2. Eye tracking calibration points displayed on the powerwall

Fig. 3. The green square displayed on the powerwall

The scenario of the tests was very simple; the users had to perform as many selections as possible within 90 seconds. Gaze direction is represented by a bounding box, starting from user's head towards the screen, long and thick enough to collide with the green square. The system detects when the gaze direction of users falls over the green square, by testing the collision between the bounding bar of gaze direction and the green square. In Fig. 4 the frame of the bounding bar is displayed and one could notice that it intersects one of the objects in the virtual environment. In this case the bounding box of the selected object is also drawn to confirm the success of the selection.

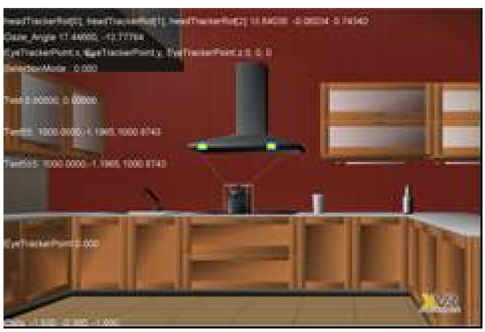


Fig. 4. The frames of the gaze direction bounding box

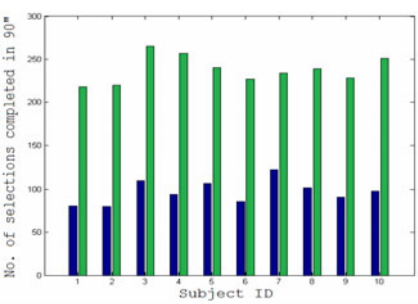
The bounding box of the gaze direction has one of its faces at the corresponding location of subject's head while the orientation is given either by head orientation either by gaze direction, depending on the selection method currently used. Because of this spatial disposition, some users mentioned that the bounding box can be regarded as an extension in the virtual reality of the human body, or as a self-centered pointing device. Head position and orientation are retrieved in real time by a magnetic tracker. A 6 DoF sensor is attached to the helmet (Fig. 5), thus providing a complete freedom of movement to the subjects during the experiments. The software used for visualization is XVR Studio (E**X**treme **V**irtual **R**eality) developed by the VRMedia Spin Off of Scuola Superiore Sant' Anna, Italy. This software architecture has the

Fig. 5. Magnetic sensor attached to the ASL eye tracker's helmet

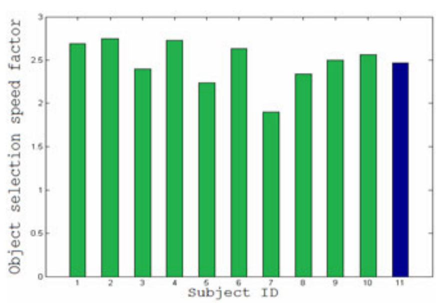
capability of extending its features through external dynamic link libraries (dll). External data can also be injected using socket communication. For head tracking we have used an external connection to a dynamic library written in c++ which retrieves the position and orientation of the head from a magnetic tracker. Data from the eye tracking device was transferred through an UDP port.

4 Discussion of Results

During the experiments, when a selection occurs, the counter for the number of completed selections is increased. This variable is saved in a database, along with time when each selection occurs (minute, seconds and milliseconds). The overall results obtained for each of the 10 subjects are available in Fig. 6. They clearly indicate that object selection speed ratio is considerable superior for the interaction interface based on eye tracking (Fig. 6(a)). In average, this ratio is 2.47 (Fig. 6(b) – the blue bar). The standard deviation of the number of selections completed is 13.6 for the head tracking method and 15.7 for the method based on eye tracking.



(a)



(b)

Fig. 6. Number of selections completed by each subject within 90" (a): green bars – selections made by gaze tracking; blue bars – selections made by head tracking; Selection speed ratio (b): - green bars – selection speed ratio for each subject; blue bar – average selection speed ratio

When designing a natural interaction interface it is essential to have in mind the easiness of the selection procedure, simplification, promptness and user abilities. Our contribution lies in the evaluation of the promptness of a head tracking selection interface, relative to a fast gaze tracking interface. Provided that selections of objects made with an interface based on gaze tracking are faster than any other selection method, values obtained in these experiments can be used as a point of reference for evaluation and comparison of other selection methods, in terms of selection promptness or stress exerted on the user. In the case of the head tracking selection interface, an average 2.47 ratio can be considered as high, because head movements requires complex spatial coordination, forcing the user to be more focused. This delay in combination with the constraint of wearing a sensor on the head, in time can lead to fatigue and discomfort.

Acknowledgments. This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/6/1.5/S/6. Also special thanks to PhD. Eng. Franco Tecchia and Prof. PhD. Eng Marcello Carrozzino from the PERCRO laboratory, Pisa, Italy, for their support during the preparation of the experiments.

References

1. Barea, R., Boquete, L., Mazo, M., López, M.: Wheelchair Guidance Strategies Using EOG. Journal of Intelligent and Robotic Systems 34, 279–299 (2002)
2. Prendinger, H., Hyrskykari, A., et al.: Attentive interfaces for users with disabilities: eye gaze for intention and uncertainty estimation. Univ. Access Inf. Soc. 8, 339–354 (2009)
3. Martens, C., Prenzel, O., Gräser, A.: The Rehabilitation Robots FRIEND-I & II: Daily Life Independency through Semi-Autonomous Task-Execution. In: Rehabilitation Robotics. I-Tech Education Publishing, Vienna (2007)
4. Corno, F., Gale, A., Majaranta, P., Räihä, K.J.: Eye based direct interaction for environmental control in heterogeneous smart environments. Handbook of Ambient Intelligence and Smart Environments 9, 1117–1138 (2010)
5. Bee, N., Prendinger, H., Nakasone, A., André, E., Ishizuka, M.: AutoSelect: What You Want Is What You Get: Real-Time Processing of Visual Attention and Affect. In: André, E., Dybkjær, L., Minker, W., Neumann, H., Weber, M. (eds.) PIT 2006. LNCS (LNAI), vol. 4021, pp. 40–52. Springer, Heidelberg (2006)
6. Shimojo, S., Simion, C., Shimojo, E.: Gaze bias both reflects and influences preference. Nature Neuroscience 6, 1317–1322 (2003)
7. Vertegaal, R., Shell, J.S., Chen, D., Mamuji, A.: Designing for augmented attention: Towards a framework for attentive user interfaces. Computers in Human Behavior 22, 771–789 (2006)
8. Qvarfordt, P., Zhai, S.: Conversing with the user based on eyegaze patterns. In: Proceedings of the ACM CHI 2005 Conference on Human Factors in Computing Systems, pp. 221–230. ACM, New York (2005)