EYE TRACKING APPLICATIONS

F. BARBUCEANU[1] **C. ANTONYA**[1]

Abstract: *As computer applications keep growing in complexity, human-computer interaction must keep the pace accordingly. Being the most important human sense, eye sight is considered a powerful new solution for enhancing communication with computers. Recent progress in eye tracking research outlines several applicability opportunities in the fields of virtual environment, augmented reality, computer navigation, internet usability tests and psychological evaluation. Sound recognition modules, haptic input devices and eye movements' information shape together a set of powerful tools for enabling natural interaction within immersive virtual environments. This paper is a review of existing eye tracking applications in the field of virtual reality and human-computer interaction.*

Key words: *eye tracking, gaze vector, peripheral blur, depth-of-field blur.*

1. Introduction

In their early stage, eye movements investigations were being carried out using intrusive methods such as scleral search coil surrounded by magnetic fields. This method consists in pairs of contact lens applied directly on the subject's eye. The coil is embedded in the contact lens and is passed by electric currents, generating perturbations inside the neighbouring magnetic field, depending on how the orientation of the ocular globe changes [18]. The main disadvantage of this method is its highly invasive characteristics. Although the system is able to detect the real orientation of the eye in 3D space very accurately, it is not able to sense head translations due to the uniformity of magnetic field. Secondly, if the head rotates, the ocular globe changes its centre of coordinates leading to wrong results.

Another method extensively used in the 70's was Electrooculography (EOG). Using electrodes applied around the ocular globe, the system could sense differences of potential above and below the eye. The main advantage is the capability of detecting very fast movements due to high sample frequencies which can be employed. The above-mentioned disadvantages also apply to this method [1].

Commission Internationale de l'Eclairage (CIE) and Communication by Gaze Interaction (COGAIN) network of excellence initiated a Technical Committee within the CIE, with the purpose of setting a standard for eye trackers. The resulting study of the first meeting held in November 2007 outlines the kinds of light sources used in various image-based eye tracking devices, and possible hazards of exposure to nearly infrared light [14].

The advantages of using infrared light instead of natural light is that it is comfortable for the user since it is not visible, and preserves image quality as it is not influenced by any variation of the visible light. An infrared camera will only sense the infrared light reflected by objects illuminated with an infrared LED (Figure 1). Thus, unwanted influence from the surrounding

---

[1] Dept. of Product Design and Robotics, *Transilvania* University of Braşov.

visible light will not perturb the image.

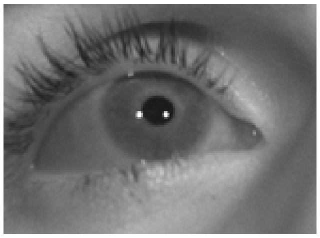

Fig. 1. *Eye illuminated by two IR light sources* [14]

The skin Penetration of Radiation within a range from 100 to 3000 nm is minimum for IR-A light (760-1400 nm). Effects of IR light depend on the distance between the source and the eyes, which is different for various types of eye trackers [14].

2. Eye Tracking: State-Of-The-Art

A common difficulty that most eye tracking systems face is represented by head movement. A first solution to this problem was the introduction of Head Mounted Systems, which are designed to carry the equipment on a helmet attached to the scull.

The next-generation remote eye tracking devices eliminate the inconvenience of wearing the helmet. They allow the user to move within a reasonable range. Video-oculography (VOG) remote gaze tracking systems consist of several IR light sources and a camera. There are cases when the camera and the light sources are built together in the same device - the so-called Time of Flight (TOF) cameras. LEDs can be positioned either on or off the optical axis. When the LED is on the optical axis, the pupil is bright because the light reflects back to the camera, following the same path. When the light source is off the optical axis, the pupil appears as black. The number of light sources can vary. By also considering corneal reflections, the estimation of gaze direction will be improved: the more reflection points, the better quality. With most of the systems, the interval of infrared light wavelength ranges between 820 - 950 nm [14]. The new and very efficient 'Image Difference' method for pupil detection uses pulse light synchronized with the video system's frequency. It is based on two light sources, one on the axis and the other off the axis, both with alternative working intervals. For each acquired video frame, one of the two sources is on, the other is off. In the former frame the pupil appears bright, whilst in the latter the pupil appears black. This leads to a high contrast, making pupil detection more accurate, by comparing the differences between two consecutive frames. The frequency should be high enough so that eye positions in consecutive frames differ very little.

3. Eye Tracking Applications

3.1. Automotive Industry

Data analysis of driver's eye movements can be very helpful for attention, sleepiness and drowsiness measurements while driving. It is well-known the relationship between eye movements and attention, both sharing the same area on the brain cortex.

On the road, the driver is not focused on driving all the time. When using in-car devices, for example, it typically takes about 1-2 seconds or even longer (depending on the traffic conditions and road geometry) to get enough information from a specific device. The amount of time is important as feedback information for designers of in-car devices.

The sleepiness of the driver can be detected by analyzing either blink duration and amplitude or the level of gaze activity.

Swedish National Road and Transport Research Institute (VTI) developed a series of driving simulators. In 2004, the third release, Simulator III, was introduced (see Figure 2).

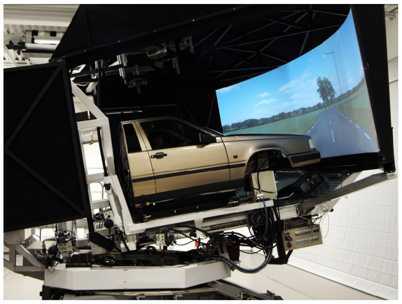

Fig. 2. *Simulator III driving simulator* [1]

The system allows a lateral movement of ±3.75 m and an acceleration of up to 0.8 g. It can be pitched, rolled and rotated up to 90°.

Fig. 3 *Smart Eye Pro eye tracking system cameras* [1]

Driver's eye movements are analyzed using the Smart Eye Pro image-based eye tracking system. This is a remote, non-intrusive eye tracking system with 2 to 6 video cameras spread horizontally in front of the driver (Figure 3).

Using a chessboard, each of these cameras is calibrated for positioning correlation with all the others. The system analyzes head position, gaze direction, quality measures etc., at a frequency of 60 Hz, which in some situations makes it difficult to measure blink amplitude. Blink duration is typically between 10-100 ms, whereas the refreshment rate of the eye tracking system is 17 ms. The accuracy of the system is maximum with straight forward gazing and decreases with large-angle displacements, since the cameras are mounted in front of the driver. Also, the accuracy diminishes when the user is looking downwards, glances sideways or sits at too large a distance from the cameras.

A set of measurements is collected, mainly from ISO and SAE standards, which are further interpreted by the researchers from VTI [1].

While driving, eye movements and steering are tightly linked. When drivers are dealing with a road bend they tend to look at the inner curvature just before beginning the steering procedure. In the experiment conducted in [21], they compared the benefits of different eye movement patterns as the visibility of the inner curvature was getting diminished. The results show that those users who looked towards the inner curvature, even though it was not visible, had a better performance than those who chose not to look in that direction, as the target became invisible. This implies that eye movements improve coordinated moves, even if the gathering of visual information is obstructed, confirming that visually guided manual tracking improves the accuracy when eye movements match steering movements. The oculomotor controller assists the neural system in the steering process, focussing the eye-ball to the target and deducing the required value of the steering angle [21].

3.2. Gaze Vector of Avatar

Guessing avatar's gaze in a virtual environment is another applicability of eye tracking (Figure 4). In a series of experiments in which a subject was projected into a virtual environment, eye movements and orientation were recorded with a head-mounted eye tracker and then reproduced through an avatar's eyes.

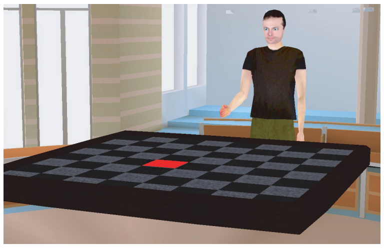

Fig. 4. *Avatar in virtual environment looking at certain chessboard squares* [1]

In this experiment, the Immersive Projection Technology (IPT) was used, which allows users to be projected into a virtual environment where they can meet other users possibly located in different geographical areas. The purpose was to see if the users are able to determine avatar's gaze point. Knowing the gaze point of another user's avatar is an important interactional enhancement. With only head tracking and without eye tracking, the users were able to determine only 1.8 out of 9 objects. When combining head and eye tracking, the users had a success rate of 8.8 out of 9 objects [1].

The second experiment tested how eye convergence (Figure 5) influence gaze coordination.

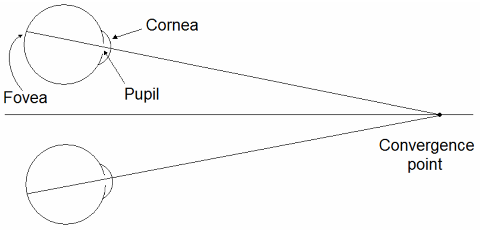

Fig. 5. *Convergence point*

The users reported that for the nearest 5 positions from the avatar, the intuition of gaze point was better with eye convergence.

Another experiment was conducted to asses whether stereoscopic vision improves somewhat gaze perception, but the results did not display any significant differences between mono and stereo mode [15].

With the help of gaze detection, robots can be programmed to react on natural, human-like communication mechanisms. In [19] a robot was able to detect gaze direction using an active robotic vision system which scans user's eyes and detects the direction of focus. When the user is staring at a certain object, the vision system detects that object according to gaze direction, then picks it up and hands it over to the user. Robots with this kind of human-robot communication methods can be successfully employed in various daily tasks, especially for disabled people.

Gaze vector calculation within immersive virtual environments is also possible using a portable head-mounted eye tracking device. The length of the gaze vector is equal to the distance between user's head and the convergence point (Figure 5). The calibration is made through sequential fixations on predetermined positions in virtual 3D space. The application is well suited for large objects selection [5].

3.3. Depth-of-Field Blur

Depth-of-Field Blur Effect for first person navigation in virtual environments consists in blurring parts of an image depending on the gaze vector. There are two types of blurring: in depth (Depth-of-Field - DOF) (Figure 6), which blurs objects in the back and front of the gaze point, and Peripheral Blur, which blurs objects located near the peripheral area of sight.

The focal distance is the distance from the eyes to the gaze point. The most efficient way of measuring this distance is by using eye tracking. The image near the gaze point is perceived as sharp. The peripheral blur effect simulates the decreasing of sharpness towards the periphery of human vision field. This effect grows progressively with the distance from the gaze point [8]. Experiments have proved

Fig. 6. *Depth-of-Field blur*

that users prefer the DOF effect only when using an eye tracker. Also, the feeling of immersion, entertainment and the depth of virtual environment are improved when using a DOF effect. Camera motion was also tested. It was designed to simulate human walking motion with sinusoidal displacement on each of the three coordinate axes. The results also indicate an enhancement if using eye tracking; the users feel the same improvement in terms of fun and a sensation of immersion into the virtual world [9].

3.4. Level of Detail (LOD)

Level of Detail in Gaze-Contingent Displays (GCD) can be adjusted according to user's focus point. There are two approaches regarding LOD. One is geometrical, whereas the other relies on image-based techniques. GCDs are susceptible to some sort of delays caused, in different measures, by the eye tracker response time and by the amount of time needed to degrade the image. The second delay might be reduced if using programmable hardware equipments [2].

3.5. Desktop Navigation

As a stand-alone navigation tool, eye tracking has shown more than satisfactory results recently, being successfully employed in a commercial monitor model equipped with a video camera and two illuminators below the screen [12]. Tobii 1750 eye tracker represents a good solution for page up/down scrolling, off screen gaze actuated buttons, dwell and micro dwell action triggering, as the results of experiments presented in [11] conclude.

When the users are looking downwards or upwards and go beyond a certain threshold, the scrolling speed changes accordingly.

The overall results of experiments in [11] outline their option for gaze-enhanced scrolling techniques, in preference to a classical, manual input approach.

4. Gaze Prediction

Electrical signals within our body are perturbed by noise. Thus, the brain must do some estimation. For example, if we swiftly point with our finger to a specific point, we will not always hit the target very accurately, but nevertheless will get within a certain range. Also, we do not know precisely what the position of our hand is, unless we look at it. The brain makes estimations in order to make decisions. Bayesian decision theory is a good representation for how prior knowledge and experience should be combined in the scope of choosing the optimal solution [10].

Recent studies have shown that the relationship between eye movements and visual perception is determinant. It depends on the task given to the subject and the thoughts of his/her mind, but these high-level models are far too complicated to be applied in reliable applications. Low level models were developed instead, based on simple tasks and bottom-up primitive stimuli like texture contrast, variance of pixels, intensities of colour, corners, crosses etc. The framework elaborated in [17] reveals how bottom-up heuristics are significantly good gaze predictors.

In multiple object tracking, the users tend not to focus on each object, but on some areas from where the attention is distributed in a relatively similar way. In the experiment conducted in [7], the users where not assigned any task. When multiple moving points are displayed on the screen, they tend to focus on the centre of the triangle determined by the 3 moving points. The results suggest that users tend to associate these points with a virtual object, thus they concentrate on the centre of that object. Some studies reveal that grouping target points into a virtual object improves multiple object tracking [22].

In another experiment [13], where the users were concerned with a specific task, their eye movements had a different pattern. They had to focus on potential object collision in an air-traffic control application. As a consequence, they focused more on the points that were in a potential collision situation.

The results of experiments [13], [20] [22] imply that objects can be tracked even without focusing directly on them and that mind can map objects even if they are not within eyesight, improving thus interaction with complex surrounding environments.

In the experiment presented in [3], the influence of eye-hand coordination on the anticipatory eye movement pattern was studied. The experiment consists in grasping rectangles on the screen using the mouse, in a specified manner. Users must either grasp the objects themselves or just observe other users or the computer completing the task. The active condition corresponds to the situation in which the users had to complete the task themselves, while the passive condition corresponds to the situation in which they had to observe other users or the computer completing the task. In passive conditions the movements of other users are recorded and reproduced by the computer. Furthermore, the subject is not informed on this fact, and it is induced to him that the other user is generating the movements.

The results reveal that anticipatory eye movements persist in active conditions, regardless of the direct contact between hand and manipulated object. This aspect can be related to target setting. Focusing on the target, the only task that the neural system should accomplish is to shorten the distance between object and target, while the object appears on the fovea area of the retina and can be tracked.

In [3] is suggested that the main factor which influences gaze behaviour when moving things is the level of understanding of the phenomenon causing the events.

In another experiment [4], eye movements were monitored while following a 3D path with a binocular head mounted eye tracking system. The purpose was to analyze eye-hand coordination while tracing a 3D path and tracking an object moving in 3D space. The results have shown that eye movement surpasses hand movement with about 20-40 ms for azimuth and elevation, and with about 250 ms for in-depth movements while tracking a moving object. When tracing 3D paths, the advance of eye movement was more consistent: about 220 ms for azimuth and elevation and 390 ms for in-depth tracking. These differences of values between in-depth and directional movement leading time confirm the idea that vergence and saccades are controlled by different neural parts with different response time.

If a human could be perceived as a machine with a large number of mental states, classified according to the level of abstraction, it would be possible to group them in sequences corresponding to a certain goal. High-level abstraction (long-term) states like, for example, passing a car on a highway would be composed of low level (short-term) states like pushing the gas pedal or keeping the lane. If these states could be identified and classified, they could lead to good action prediction results and eventually could enhance the implementations of assistance systems for car drivers [16]. Generally, visual information from the environment is used for momentary task and, if needed for subsequent tasks, it will be stored for later use. Unpredictable information is more likely to be preserved, since it contrasts with learned regularities [6].

5. Conclusion

Measurements of eye movement can serve in transferring very useful information from man to machine. Depending on the complexity of that information, machines might need to implement specialized cognitive architectures in order to be able to understand and utilize it accordingly. The higher the level of abstraction of the information encrypted in eye movements' data, the higher the complexity of the algorithms employed in its decoding.

Acknowledgement

This work has been supported by the Romanian National University Research Council, under Grant INCOGNITO, Exploratory Research Projects 608/2009.

References

1. Bergstrand, M.: *Automatic Analysis of Eye Tracker Data.* VTI notat 12A, VTI, Linkoping, 2008.
2. Duchowski, A.T., Çöltekin, A.: *Foveated Gaze-Contingent Displays for Peripheral LOD Management, 3D Visualization, and Stereo Imaging*. In: ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), New York 3(4), article 24, 2007.
3. Gesierich, B., Bruzzo, A., et al.: *Human Gaze Behaviour during Action Execution and Observation*. In: Acta Psychological **128** (2008) No. 2, p. 324-330.
4. Gielen, C., Dijkstra, T., et al.: *Coordination of Gaze and Hand Movements for Tracking and Tracing in 3D*. In: Cortex **45** (2008) No. 3, p. 340-355.
5. Haffegee, A., Alexandrov, V., et al.: *Eye Tracking and Gaze Vector Calculation within Immersive Virtual Environments*. In: Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology, ACM New York, p. 225-226.
6. Hayhoe, M.M., Ballard, D.H., et al.: *Vision in Natural and Virtual Environments*. In: Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, ACM, p. 7-13.
7. Hilda, M.F., Adriane, E.S.: *Eye Movements During Multiple Object Tracking: Where Do Participants Look?* In: Cognition **108** (2008) No. 1, p. 201-209.
8. Hillaire, S., Lécuyer, A., et al.: *Depth-Of-Field Blur Effects for First-Person Navigation in Virtual Environments*. In: IEEE Computer Graphics and Application **28** (2007) No. 6, p. 47-55.
9. Hillaire, S., Lécuyer, A., et al.: *Using an Eye-Tracking System to Improve Depth-of-Field Blur Effects and Camera Motions in Virtual Environments*. In: Proceedings of IEEE Virtual Reality Conference 2008, Reno, p. 47-51.
10. Kording, K.P., Wolpert, D.M.: *Bayesian Decision Theory in Sensorimotor Control.* In: Trends in Cognitive Sciences **10** (2006) No. 7, p. 319-326.
11. Kumar, M., Winograd, T.: *Gazeenhanced Scrolling Techniques*. In: Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology, Newport, 2007, p. 213-216.

12. Kumar, M., Winograd, T.: *GUIDe: Gaze-Enhanced UI Design*. In: CHI'07 Extended Abstracts on Human Factors in Computing Systems, San Jose, 2007, p. 1977-1982.
13. Landry, S.J., Sheridan, T.B., et al.: *A Methodology for Studying Cognitive Groupings in a Target-Tracking Task*. In: IEEE Transactions on Intelligent Transportation Systems **2** (2001) No. 2, p. 92-100.
14. Mulvey, F., Villanueva, A., et al.: *Exploration of Safety Issues in Eye Tracking*. In: Deliverable 5.4, Communication by Gaze Interaction (COGAIN), IST-2003-511598, 2008.
15. Murray, N., Roberts, D., et al.: *An Assessment of Eye-Gaze Potential within Immersive Virtual Environments*. In: ACM Transactions on Multimedia Computing, Communi-cations, and Applications (TOMCC AP) **3** (2007) No. 2.
16. Pentnald, A., Liu, A.: *Modelling and Prediction of Human Behaviour*. In: Neural Computation **11** (1995), p. 229-242.
17. Peters, R.J., Itti, L.: *Applying Computational Tools to Predict Gaze Direction in Interactive Visual Environments*. In: ACM Transactions on Applied Perception **5** (2008) No. 2, Article 9.
18. Ronsse, R., White, O., Lefèvre, P.: *Computation of Gaze Orientation under Unrestrained Head Movements*. In: Journal of Neuroscience Methods **159** (2007) No. 1, p. 158-169.
19. Rowel, A., Zelinsky, A.: *Intuitive Human-Robot Interaction Through Active 3D Gaze Tracking*. In: Springer Tracts in Advanced Robotics, Springer **15** (2005), p. 172-181.
20. Wexler, M., van Boxtel, J.J.: *Depth Perception by the Active Observer*. In: Trends in Cognitive Science **9** (2005) No. 9, p. 431-438.
21. Wilson, M., Stephenson, S., et al.: *Eye Movements Coordinated with Steering Benefit Performance Even When Vision is Denied*. In: Experimental Brain Research **197** (2006) No. 3, p. 397-412.
22. Yantis, S.: *Multielement Visual Tracking: Attention and Perceptual Organization*. In: Cognitive Psychology **24** (1992) No. 3, p. 295-340.