Review on Issues of Eye Gaze Tracking Systems for Human Computer Interaction

Rasoul Banaeeyan
Department of Multimedia
University Putra Malaysia
Serdang, Selangor, Malaysia
banaeeyan@ComputerVision.ir

Maliheh Bahari
Department of Communication
University Putra Malaysia
Serdang, Selangor, Malaysia
maliheh.bahari@gmail.com

Abstract—Eye Gaze Tracking (EGT) allows us to estimate the direction of eye gaze and the Point of Regard (PoR) on the screen. This technique has been successfully used in different disciplines, but more technological improvements are still required to make EGT appropriate to be applied in wider range of real world applications or research tools. In this paper a review on recent issues associated with EGTs is conducted and presented. Four issues are: accuracy, comfortability, response time, and cost; but accuracy and comfortability are the most two challenging problems which are discussed here. At the end, based on the conducted survey, new directions for future researches are stated.

Keywords—Eye Gaze Tracking (EGT); Human Computer Interaction (HCI); Review Paper

I. INTRODUCTION

In the last few years there has been a growing interest in Human Computer Interaction (HCI) systems [1-7] due to need of users to interact with computers, and generally computerized devices, in a comfortable and friendly way. Eye Gaze Tracker (EGT) systems are a very effective method for establishing HCI applications which detect the point where user is looking at on the screen. General components of gaze trackers are demonstrated in Fig.1.

EGTs are, structurally, categorized in two main groups of (i) intrusive and (ii) non-intrusive techniques Fig.2. Intrusive techniques [8]-[12] require users to attaché devices to their eyes and heads such as contact lenses and cameras in order to perform eye gaze detection; while, non-intrusive [13]-[22] techniques offer a comfortable and user friendly way of gaze detection by processing the images of detected eye achieved via a digital camera.

Recently, considerable attention has been paid to non-intrusive techniques in view of the fact that they are cost-effective, easy to setup, easy to use, and with more range of applications in wide range of disciplines [23-28] such as cognitive studies, medical research, human factors, computer usability, fatigue detection, vehicle simulator, virtual reality, etc.

In general all EGT applications are divided into two main categories as described in subsections A and B (Fig. 3).

Fig 1. General Components of eye gaze tracker

Fig 2. Types of EGT techniques

Fig 3. Applications of EGT

A. User Interfaces

Eye gaze provides a very efficient method of pointing. We use it all the time in interaction with other humans. Eye-tracking technology enables us to use our gaze to interact with computers and machines. It's fast, and natural. Some examples of this category are (i) EGT in computer interaction: eye tracking enables fantastic new experiences in games and other programs, fundamentally changing the way we communicate, play and work, (ii) EGT in accessibility solution: where most users will benefit from an interface where eye gaze is used as a complement to other controls, a single-mode interface where eye gaze is the only
control method has been designed for users with mobility impairments, (iii) EGT in transportation: Eye tracking is also becoming part of other human-machine interfaces. In cars and trucks, it’s being integrated with advanced driver-assistance systems to improve safety and the user experience, and (iv) specialized interfaces: Like hands-free screen interaction in hospital environments. Industrial workstations controlled by gaze and gestures. Or process support for operators based on their attention span.

B. Human Behavior Analysis

The human brain automatically directs the eye to information it is processing, so by observing what a person is looking at, we can see what information their brain is processing. Market research companies and major advertisers use it to optimize print and TV ads. Product companies use it to optimize interaction design. Web companies use it to optimize online user experiences. And universities use it for research in psychology, neurology and medicine.

Applicable for different real world systems, an EGT should be [29]: (i) accurate, precise to minutes of arcs, (ii) be reliable, have constant, reputable behavior; (iii) be robust, should work under different conditions, (iv) be non-intrusive, (v) allow for free head motion, and (vi) have real time response.

Enhancing accuracy of EGTs both on the vertical and horizontal axes is an important issue for researches and comprehensive studies have been carried out with this respect. Most effective and applicable EGTs are based on the Pupil Corneal Reflection (PCR) [30,31]. However, most of the previous studies do not take to account the problem of finding the expected Point of Regard (POR) within the resulted gaze region (small circle on the screen) after EGT process.

II. ISSUES ASSOCIATED BY EGT SYSTEMS

In this section, four main issues raised by EGT systems are explained (Fig.4) and the most two important are discussed in subsections A and B.

According to conducted survey on recent researches over last twelve years (2003-2015) it is inferred that the most frequently addressed issues associated with EGTs are (i) accuracy, (ii) comfortability, (iii) response time, and (iv) costs as in Fig.4.

Response time refers to time duration between looking at a particular point and detection of coordinates of that point (X and Y) by EGT system. Costs in EGT system are divided in two categories of (i) computational and (ii) financial. Computational refer to the complexity of gaze tracking algorithm and financial refers to type and number of devices and equipment required to design and perform the EGT system.

Concepts of accuracy and comfortability are discussed in detail under subsections A and B.

A. Accuracy

The term accuracy in EGTs mainly refers to horizontal and vertical accuracy of detected gaze on the screen (X and Y coordinates), which is the main aim of each EGT application. Many researches have been carried out on increasing accuracy [32]–[37]; for instance [32] achieved different enhanced accuracies for seven users as states in Table 1.

<table>
<thead>
<tr>
<th>User</th>
<th>Horizontal Accuracy</th>
<th>Vertical Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.86°</td>
<td>1.02°</td>
</tr>
<tr>
<td>2</td>
<td>1.87°</td>
<td>2.35°</td>
</tr>
<tr>
<td>3</td>
<td>1.92°</td>
<td>2.42°</td>
</tr>
<tr>
<td>4</td>
<td>2.11°</td>
<td>1.68°</td>
</tr>
<tr>
<td>5</td>
<td>1.32°</td>
<td>1.72°</td>
</tr>
<tr>
<td>6</td>
<td>2.04°</td>
<td>2.68°</td>
</tr>
<tr>
<td>7</td>
<td>1.74°</td>
<td>1.92°</td>
</tr>
</tbody>
</table>

Table 1. Accuracies of Zhang EGT for seven subjects

Then, accuracy is compared with other significant EGT techniques [38]–[43] as summarized in Table 2.

<table>
<thead>
<tr>
<th>Method</th>
<th>Average Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Shih and Liu 2004)</td>
<td>0.86°</td>
</tr>
<tr>
<td>(Zhu and Ji 2007)</td>
<td>1.92°</td>
</tr>
<tr>
<td>(Morimoto and Mimica 2005)</td>
<td>1.21°</td>
</tr>
<tr>
<td>(Villanueva and Cabeza 2008)</td>
<td>1.32°</td>
</tr>
<tr>
<td>(Guestrin and Eizenman 2006)</td>
<td>2.04°</td>
</tr>
<tr>
<td>(P. Zhang et al.)</td>
<td>1.74°</td>
</tr>
</tbody>
</table>

Table 2. Accuracy of different EGT techniques

Although Shih and Liu [38] achieved an accuracy error of less than 1° for their presented gaze tracker by employing multiple cameras and light sources like the method in [41], but it is an expensive setup for an EGT system. Technique introduced in [39] by Zhu and Zhiwei, presented an average accuracy of 1.6° for their gaze tracker under natural head movement; resulted accuracies for seven subjects are sated in Table 3.
Beymer in [42] reduced the error of detected gaze point to 0.6° by introducing a 3D eye tracking system but it requires use of two stereo systems at the same time and is not calibration free meaning user have to look at nine different points on the screen for initialization process.

Accuracy reported in the research conducted by Zhiwei Zhu in [44] is a variable dependent on distance of user to display (Table 4); but like other techniques this one also has calibration procedure and user has to sit for this step.

### Table 4. Comparison EGT accuracy reported by Zhiwei Zhu

<table>
<thead>
<tr>
<th>User</th>
<th>Horizontal Accuracy (mm)</th>
<th>Vertical Accuracy (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.68</td>
<td>0.83</td>
</tr>
<tr>
<td>2</td>
<td>1.31</td>
<td>1.41</td>
</tr>
<tr>
<td>3</td>
<td>1.54</td>
<td>1.90</td>
</tr>
<tr>
<td>4</td>
<td>1.73</td>
<td>2.34</td>
</tr>
</tbody>
</table>

Accuracy reported in the research conducted by Zhiwei Zhu in [44] is a variable dependent on distance of user to display (Table 4); but like other techniques this one also has calibration procedure and user has to sit for this step.

### B. Comfortability

The term comfortability in EGTs mainly refers to ‘how easily user can interact with EGT system’ which is associated with (i) type of gaze tracker, intrusive or non-intrusive, (ii) calibration procedure, and (iii) freedom of head movements. Taking to account each one of three required characteristics, different scholars are discussed at the following.

1. Intrusive or Non-intrusive

Many intrusive techniques [8]-[12] have been implemented during the last dictate but due to their uncomfortable setup design the use of them has dramatically decreased as demonstrated in Fig.3; from this chart it also can be inferred that non-intrusive techniques [13]-[22] have been in attention. The main reason for not using intrusive techniques is that they require user to attach devices to their head or eyes or electronic lances during EGT procedure unlike non-intrusive methods that mostly rely on computer vision techniques and use not-attached cameras.

### Table 5. Different calibration errors

<table>
<thead>
<tr>
<th>Technique</th>
<th>Error in Pixel</th>
<th>Error in Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear calibration method</td>
<td>35.1</td>
<td>1.27°</td>
</tr>
<tr>
<td>Second order polynomial calibration method (9 points)</td>
<td>26.8</td>
<td>0.97°</td>
</tr>
<tr>
<td>Second order polynomial calibration method (25 points)</td>
<td>24.2</td>
<td>0.88°</td>
</tr>
<tr>
<td>Interpolation</td>
<td>29.9</td>
<td>1.08°</td>
</tr>
<tr>
<td>Model based mapping</td>
<td>25.9</td>
<td>0.94°</td>
</tr>
</tbody>
</table>

Based on Table 5 it can be inferred that the biggest point of gaze error is related to linear calibration model and interpolation.
3. Freedom of Head Movement

Not all non-intrusive EGTs allow user to have free head movement while using the system. Some of them ask users to keep their heads stationary [47]–[53] in order to reduce accuracy errors or avoid repeating calibration process while some of them let user a free/limited range of head movement [54]–[58].

The importance of handling issue of head movement during EGT process is clarified by a comparison presented by Ruian in [59], Table 6.

Although average error may seem acceptable in all cases presented in Table 6, but the number of resources used is not cost effective from both financial and technical points of view.

<table>
<thead>
<tr>
<th>User</th>
<th>Average Error in Pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hold the Head Still</td>
</tr>
<tr>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>6.67</td>
</tr>
<tr>
<td>2</td>
<td>6.53</td>
</tr>
<tr>
<td>3</td>
<td>7.12</td>
</tr>
<tr>
<td>4</td>
<td>6.96</td>
</tr>
<tr>
<td>5</td>
<td>9.89</td>
</tr>
</tbody>
</table>

Table 6. Comparison of average errors in free and not free head movement EGT system

III. CONCLUSION AND FUTURE DIRECTIONS

EGT systems are a good way of human computer interaction (HCI) and there is a demand to develop them in a way addressing four discussed issues (Fig.4). In this paper a review is conducted on different EGT system.

Four frequently addressed problems associated with gaze trackers are discussed and analyzed which are stated in term of priority in Table 7.

<table>
<thead>
<tr>
<th>Importance Level</th>
<th>Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Accuracy</td>
</tr>
<tr>
<td>High</td>
<td>Comfortability</td>
</tr>
<tr>
<td>Medium</td>
<td>Response Time</td>
</tr>
<tr>
<td>High</td>
<td>Costs</td>
</tr>
</tbody>
</table>

Table 7. Four main issues associated by EGT systems

While mentioned problems were resulted in successful development of EGT systems, they are significantly potential for further improvements.

Reliably detecting eyes from images under unexpected conditions of head poses and face orientations, variable lighting, remains problematic.

Although different scholars achieved a very good accuracy of less than 1° for existing EGT systems but they are working under a limited distance of maximum 340.26 mm and they mainly employ more than one camera/light source for implementation which is financially expensive.

Finally, it appears that an integrated approach using several available attributes is the promising direction for further development.

REFERENCES

[10] Liu, Ruian, Shijiu Jin, and Xiaorang Wu. 2007. "REAL TIME AUTO-FOCUS ALGORITHM FOR EYE GAZE TRACKING SYSTEM State Key Laboratory of Precision Measuring Technology and Instruments , Tianjin University , Tianjin", 742–45.
[17] Xia, Dongshi, and Zongcai Ruan. 2007. "IR Image Based Eye Gaze Estimation." Eighth ACIS International Conference on


