Towards the development of a standardized performance evaluation framework for eye gaze estimation systems in consumer platforms

Anuradha Kar*, Peter Corcoran*

* Center for Cognitive, Connected & Computational Imaging, College of Engineering & Informatics, NUI Galway, Galway, Ireland

Email: {a.kar2, peter.corcoran}@nuigalway.ie

***Abstract*— There is a need to standardize the performance of eye gaze estimation (EGE) methods in various platforms for human computer interaction (HCI). Because of lack of consistent schemes or protocols for summative evaluation of EGE systems, performance results in this field can neither be compared nor reproduced with any consistency. In contemporary literature, gaze tracking accuracy is measured under non-identical sets of conditions, with variable metrics and most results do not report the impact of system meta-parameters that significantly affect tracking performances. In this work, the diverse nature of these research outcomes and system parameters which affect gaze tracking in different platforms is investigated and their error contributions are estimated quantitatively. Then the concept and development of a performance evaluation framework is proposed- that can define design criteria and benchmark quality measures for the eye gaze research community.**

I. INTRODUCTION

The prolific and interdisciplinary nature of research on eye gaze estimation in the past decades has resulted in the development of a wide range of techniques and applications [1-4]. Most of the recent developments in this field comprise of passive video-based gaze tracking methods which estimate the gaze direction or the point of gaze by capturing and processing images of the eye region in natural light or under active near infrared (NIR) illumination. Eye gaze based applications are also implemented in a variety of consumer platforms such as desktop setups (for text entry, attention analysis, gaze based passwords), head-mounted equipment (gaming, virtual and augmented reality), automotive systems (driver alertness monitoring) and handheld devices (gaze based scrolling, content navigation) [5-15].

In spite of the large volume of research on EGE techniques over the years, currently there are no unified standard schemes to evaluate the performance of gaze tracking algorithms or setups in research across various platforms. Few works report the impact of system parameters such as camera and screen resolution, viewing angle and platform movements- that play significant roles in determining the ultimate accuracy of a gaze tracking system. Also there is no agreement on the metrics of accuracy across various methods. Therefore it is difficult to state with certainty if the results of a recent design would perform better than conventional ones and under what conditions certain research claims are valid.

The aim of the work presented in this paper is to address the challenges of standardization in the development and performance evaluation of gaze based HCI systems. First a literature review is provided where it is described how the implementation and outcomes in eye gaze research differ-to the extent that they are neither comparable nor reproducible. Then several sources of error including head pose variations, viewing angle, camera resolution are identified and their impact is quantified experimentally. Finally the development of a framework for evaluating the performance of EGE setups is outlined. Such a concept can be used to provide reliable estimates of error and standardized accuracy scores to users and designers of gaze based systems.

II. LITERATURE REVIEW

A comprehensive literature review on the development of EGE techniques and applications for different consumer platforms was made and the diversity in the algorithms, research outcomes, setups, sources of errors and accuracy measures for different platforms is presented below.

A. Eye gaze tracking algorithms in literature

These fall broadly into the following classes [1]: 1) Feature based: These use single or multiple near infrared (NIR) cameras and LEDs to produce glints on the cornea and use the vector between the pupil center and glint positions to track the eye gaze. These are further subdivided as: (i) regression based methods that use a polynomial regression function to map the vector to gaze coordinates on screen (ii) 3D model based methods that use a geometrical model of the eye to develop accurate ray tracing to estimate gaze direction. Feature based methods typically have good accuracy (0.5 to 1 degree) but need elaborate hardware. 2) Appearance based: These utilize the shape and appearance features of the eye region to train a model which is then used to match with captured eye images for estimating gaze. These methods need simple hardware but have lower accuracy [40].

B. Eye gaze use cases in consumer platforms

Use of eye gaze information has found applications in a variety of user platforms –some of them are described along with their recent and popular applications.

978-1-5090-1897-0/16/$31.00 ©2016 IEEE

a) Desktop based: Majority of eye gaze tracking methods have been developed for desktop based systems [30-32] wherein one or more cameras with NIR LEDs are used to track a user's eye gaze on the computer screen-with the user seated in front of the computer. Some methods allow free head movement whereas some require a fixed user head position for accurate tracking. Typical tracking methods- like PCCR (pupil center corneal reflection) or appearance models are used for gaze based applications ranging from assistive technologies for the physically impaired, understanding the development of psychiatric disorders, e-learning, studying consumer attention patterns, e-commerce and web-design [2].

b) TV panels: Gaze controlled intelligent TVs have recently been introduced [33-36] that uses PCCR techniques for gaze localization. In these systems, a user can select and navigate menus and switch channels by looking at icons shown on the TV's display.

c) Automotive: Interactive driver support systems have been developed based on actively tracking driver gaze or blinking patterns to evaluate driver vigilance and drowsiness levels [37-39]. The setups employ machine vision algorithms, appearance models or PCCR techniques using cameras and NIR light sources mounted on the car's dashboard.

d) Handheld devices: Methods for detecting user gaze location and patterns have been developed for smartphone and tablets [40-42] to activate functions such as locking and unlocking devices, interactive displays, dimming backlights or suspending sensors based on user attention. The operating principles rely on PCCR methods, appearance models or ellipse fitting techniques using IR light sources and the device front camera.

e) Head-mounted setups: These usually comprise of two or more cameras mounted on a head unit – one facing the user and recording the eye gaze (called the 'eye camera'), while the other facing outward, recording the scene (called 'scene camera'). These setups allow free head movements. Major applications include virtual and augmented reality, observing user attention, psychoanalysis and occulo-motor movements [43-45]. In Table I the characteristics of setups for gaze estimation in various use cases are summarized.

TABLE I
FEATURES OF EGE SYSTEMS IN VARIOUS USE CASES

EGE use cases	Distance between eye and tracker	Viewing angle (deg)	Screen size (inch)
Desktop	30-50 cm	~40	14-17
TV panels	2-5 m	60-120	26-70
Automotive	50 cm	40-60	--
Handheld devices	20-40 cm	5-12	5-10
Head mounted	2-5 cm	55-75	--

C. Sources of errors in EGE systems

A variety of error sources affect gaze tracking accuracy in different platforms. These include: a) Head-pose changes- which alters the eye-socket geometry, user field of view, causes LED glints on cornea-surface to disappear [8]. b) Camera resolution- affects amount of eye details and contrast in the eye image captured by the EGE setup c) Display properties and viewing angle-includes effect of size and pixel resolution of the screen where gaze is tracked and distance of the user from setup d) Platform movements–may cause variable position, orientation and jitter in the setup e) Changes in illumination- affects eye feature detection, causes additional glints to appear on cornea surface. f) Human eye limitations-the eyes can fixate as accurately as 10 minutes of visual angle (0.16 degree) which sets the accuracy limit of gaze tracking. High frequency eye movements or saccades lead to motion blur and missing gaze points if the frame-rate of the gaze estimation camera is less than 100 Hz [29].

However the problem is that only a few of these factors are reported in literature while the impact of the others is not considered at all. As an illustrative example, the effect of head pose variations are considered in [8-10] whereas impact of factors such as camera quality, display size and resolution, user viewing angle and platform movements are inadequately characterized and therefore it is not possible to know as to how an algorithm would perform under their influence. The occurrence of the various error sources in different user platforms are tabulated below.

TABLE II
SOURCES OF ERRORS IN EGE SYSTEMS

Error Sources	Head Mounted	Desktop	Dynamic
Head pose	--	X	X
Camera resolution	X	X	X
Display properties	--	X	X
Viewing angle	--	X	X
Platform motion	--	--	X
Illumination changes	--	X	X
Human eye limitation	X	X	X

D. Diversity in research outcomes in EGE literature

The metrics used to represent performance scores of gaze estimation algorithms for different CE platforms were studied and it was found that a large number of reported results cannot be compared in any meaningful way as the system performance is reported in varied formats. Some papers [16-20] report tracking accuracy in degrees while others report them as gaze recognition rates [21-24]. Again there are results where accuracy is reported in heterogeneous formats-e.g. relative pixel/distance shifts or rate of correct detections which have relevance solely to the procedure stated in the paper [25-28]. Apart from this, wide variations in setup configurations are noted in terms of the number of cameras and light sources used- varying between 1 to 4 cameras and 1 to 16 LEDs. It is also observed that head mounted setups appear to have a lower mean error (~1.08 degree) compared to desktop based (~1.87 degree) or dynamic platforms (~4.8 degree). However each platform is affected by several error sources and unique operating conditions as described in the previous sections, and therefore the absolute validity of these results are uncertain.

III. QUANTIFYING EFFECTS OF ERROR SOURCES IN EGE

A set of eye tracking experiments were conducted to simulate and observe the impact of various error sources mentioned above. These include studying head pose tolerance levels and effects of viewing angle, display characteristics and camera resolution on gaze tracking accuracy. For the setup a commercial eye tracker that works with a laptop or desktop system and has a specified gaze tracking accuracy of 0.5 degree was used. Eye tracking data is collected from users who were asked to sit in front of the eye tracker connected to a computer as shown in Fig. 1 and an eye calibration routine for the user is run first. The experimental flowchart is shown in Fig3 where each experiment in general comprises of recording the eye gaze coordinates from the tracker when the user gazes and clicks on fixed points on the screen. The gaze error in degrees is calculated from the shift between click and gaze locations on the screen as shown in Fig 2 Further details of the setup and experiments are described below.

A. Head pose tolerance

In this experiment a user is asked to sit in front of the tracker with their head initially positioned frontally. A video camera is used to capture the position of the head and the head pose in roll, pitch and yaw angles is obtained from an appearance model as shown in Fig 4. For each position gaze accuracy data is recorded. Then the user is asked to turn their head to specific fixed positions (in roll pitch yaw angles) and perform the gaze tracking experiment described above. Gaze accuracy measures corresponding to various head pose angles are recorded. The results are presented in Fig.5 below. It is observed that the gaze tracking accuracy significantly decreases as the head moves slightly beyond 20 degrees in either directions of roll, pitch and yaw angles.

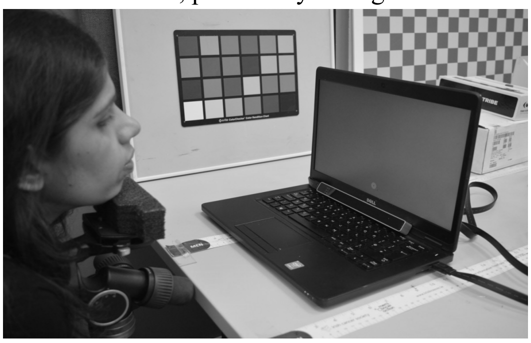

Fig.1 Experimental setup with the user, laptop and eye-tracker with the user head fixed with a chin-rest

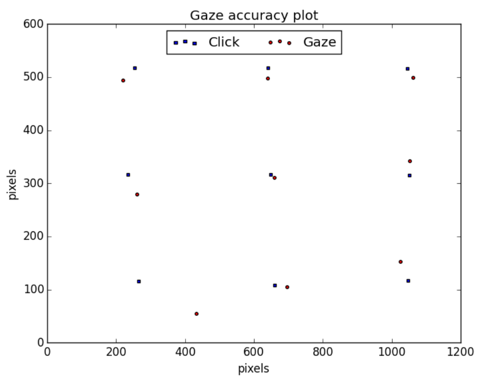

Fig2. Output from an eye tracking experiment

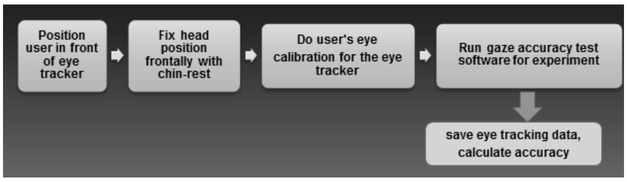

Fig.3 Experimental flowchart

Fig.4 Estimating head pose angles using a face model

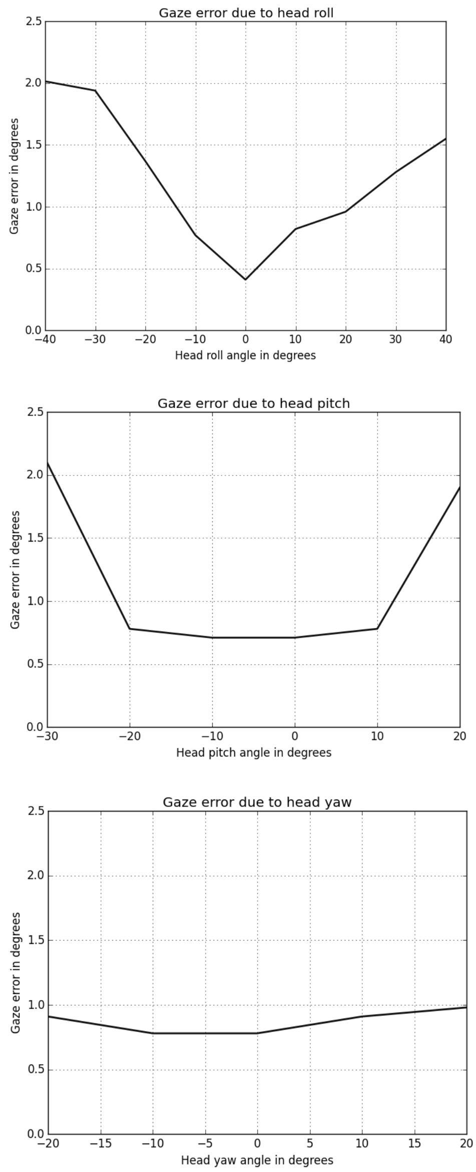

Fig5 Variation of gaze angular error with head pose angles

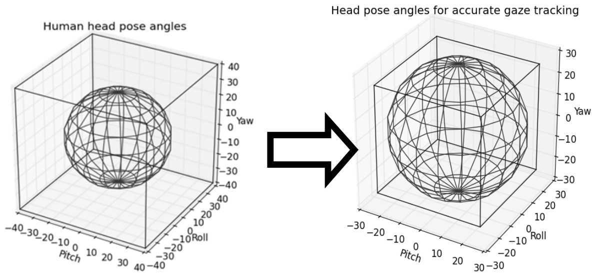

Fig6 Head pose tolerance limits of the eye tracker

The human head has considerable degree of flexibility in the order of nearly 40 degrees of angular movement. However the above results imply that for reliable eye tracking results with the given tracker, the user needs to keep their head position limited within an angular "box"-equivalent to ~20 degrees about the central position in 3 directions as shown in Fig. 6 or the gaze tracking performance will drop below acceptable levels.

B. Viewing angle

For this experiment, the users were positioned at successively increasing distances from the tracker - computer screen setup (40 cm to 100 cm in 10 cm interval) and the gaze tracking accuracy data were recorded for each user position for fixed frontal head poses and plotted in Fig 7 below. The user viewing angle decreases as the user moves away from the screen as shown in Fig. 7(top). It is observed that for the tracker, the accuracy decreases rapidly with an increased viewing angle. However it is also found that the tracker errors are high when the user is too close to the tracker-that is the user to tracker-screen setup is less than 40 cm.

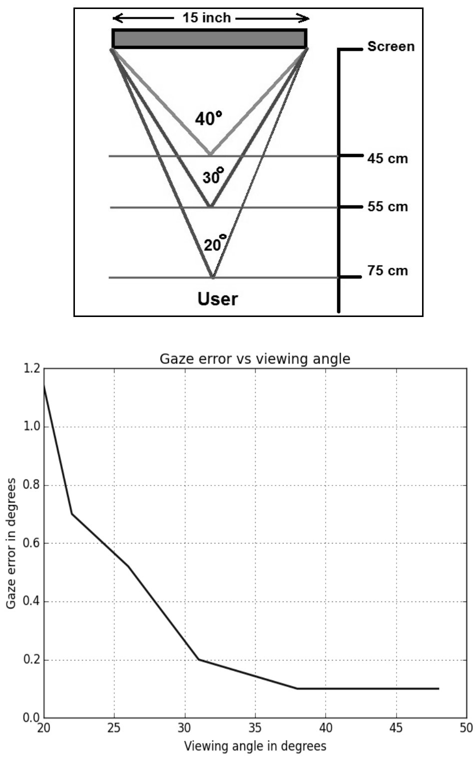

Fig7. Top: Variation of user viewing angle with distance from screen. Bottom: Variation of gaze accuracy as a function of user distance.

C. Display size and resolution

The eye tracking experiment was run on displays of various sizes (9, 11, 13.5, 15 inches) and resolutions (800x600, 1024x768, 1280x768, 1366x768) and the corresponding gaze tracking errors are shown below.

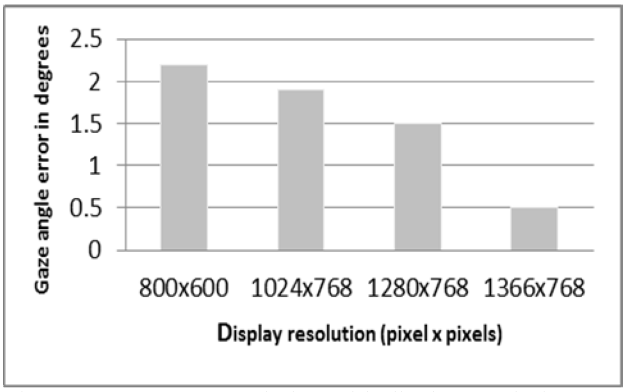

Fig.8 Gaze accuracy variation with display resolution

D. Camera resolution

Every gaze tracking device has one or more cameras to capture images of the eye region and the camera resolution directly affects the amount of eye details in the captured image and hence the accuracy of gaze tracking [29].To study the effect of camera resolution on gaze estimation errors, eye images were captured using camera resolutions starting from 1.3 to 24 megapixels for different distances between the camera and user under constant illumination as shown in Fig.9.

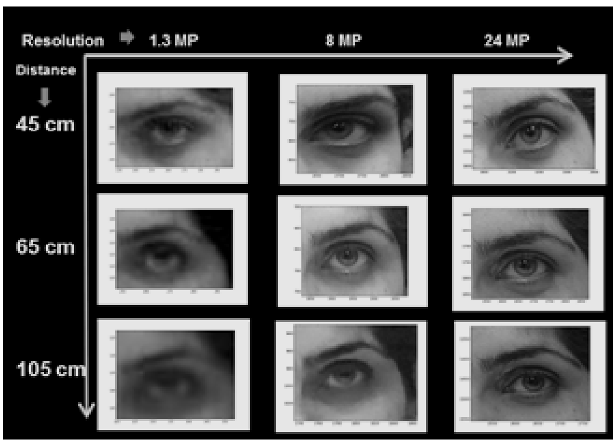

Fig9 Eye images taken with different camera resolutions at different distances from the user

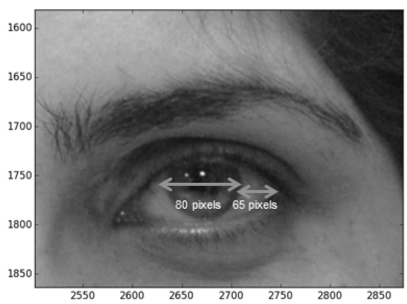

Fig10. Details of the eye socket region in an eye image

The eye details in an image are estimated as the eye socket width in pixels as shown in Fig. 10 above. The eye-details as a function of distance from the camera are estimated and plotted in Fig.11 & 12. It is observed that for a camera with low resolution the eye details are inherently low and the image quality reduces drastically as the eyes move away from the camera. With a high quality camera the eye details remains nearly constant with distance. To observe how the camera quality might affect eye gaze tracking- a pupil detection algorithm on cropped eye images was applied using circular Hough transform (Fig.13). The shift in positions between the actual and detected pupil centers is estimated as pupil detection errors. The plot of errors as a function of camera resolution is shown in Fig14. It is seen that pupil detection errors reduce with better camera quality-but levels off around 8 MP which could be an optimal camera resolution that can be used to build reliable remote eye trackers operating in the horizontal range of 1meter.

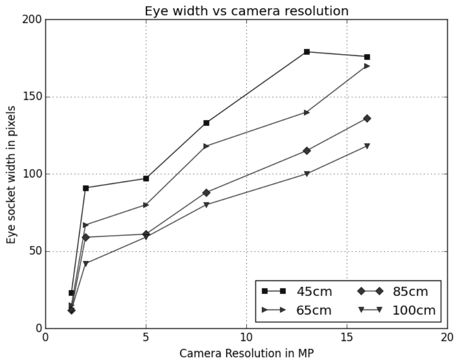

Fig11 Variation of eye details with camera resolution.

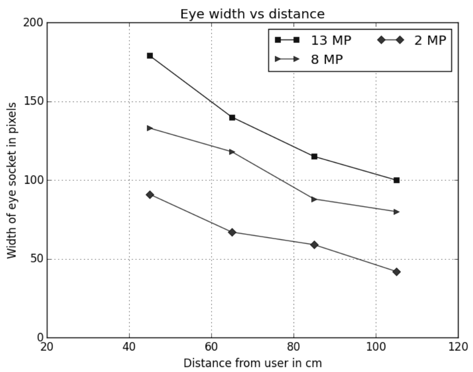

Fig12 Variation of eye details with varying distance between camera and user for different camera resolutions

IV. DISCUSSIONS AND FUTURE WORK

In this paper, the diverse range of algorithms, performance metrics and system specifications of gaze based HCI systems was investigated. Then a set of experiments were designed and implemented to estimate the impact of various sources of errors that affect EGE performance. The major contribution of our study was identifying two main issues that affect realistic performance evaluation of conventional gaze based systems. Firstly there is a lack of protocols for reporting research outcomes and accuracy scores in literature.

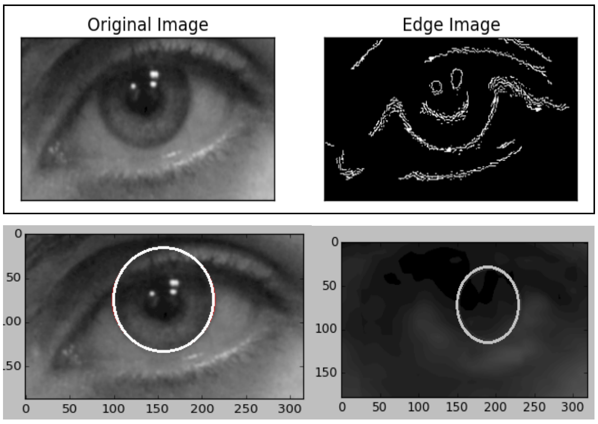

Fig.13 Top: Applying pupil detection on eye images. Bottom: Pupil detection in good and poor quality images

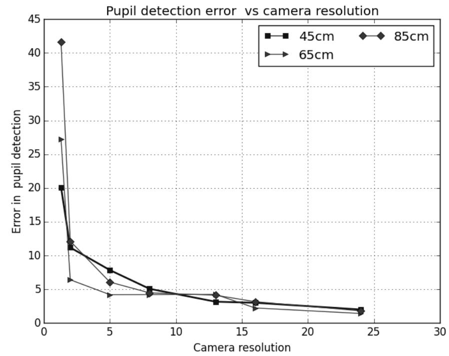

Fig14. Pupil detection errors as a function of camera resolution and distance from the user.

Secondly, there are several parameters in EGE system development that potentially affect tracking performance, but their impacts are very sparsely evaluated and reported by researchers in this field. Our studies on various error sources showed that error contribution from these parameters are significant and it can be concluded that the reported results on gaze tracking techniques that are measured under controlled setup conditions do not give a correct estimate about the overall accuracy of the system.

With this background we are designing a unified framework that would provide standardized performance scores of gaze tracking systems in different platforms while taking into consideration the multiple error sources that may affect such a system. The framework is expected to operate through multiple control experiments wherein each error source is evaluated separately. It may be extended to estimate the summative error in the system as a measure of its true performance. Through such a framework it will be possible to evaluate and compare the performance of multiple algorithms while measuring specific gaze behavior/actions (e.g. pursuit/fixation), under uniform experimental conditions and identify performance bottlenecks of different algorithms when they perform under identical circumstances. The ultimate aim is to provide more reliable performance measures and design guidelines that will facilitate both the future users and designers of advanced EGE systems.

ACKNOWLEDGMENT

The research work presented here was funded under the Strategic Partnership Program of Science Foundation Ireland (SFI) and co-funded FotoNation Ltd. Project ID: 13/SPP/I2868 on "*Next Generation Imaging for Smartphone and Embedded Platforms*".

REFERENCES

[1] D. W. Hansen and Q. Ji, "In the eye of the beholder: a survey of models for eyes and gaze.," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, 2010.

[2] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Comput. Vis. Image Underst.*, vol. 98, no. 1, pp. 4–24, 2005.

[3] G. Vasilescu, "Springer Series on Signals and Communication Technology Signals and Communication Technology," *Electron. Noise Interf. Signals*, 2005.

[4] F. Song, X. Tan, S. Chen, and Z. H. Zhou, "A literature survey on robust and efficient eye localization in real-life scenarios," *Pattern Recognit.*, vol. 46, no. 12, pp. 3157–3173, 2013.

[5] P. M. Corcoran, F. Nanu, S. Petrescu and P. Bigioi, "Real-time eye gaze tracking for gaming design and consumer electronics systems," in *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 347-355, May 2012

[6] A. Bulling, F. Alt, and A. Schmidt, "Increasing the security of gaze-based cued-recall graphical passwords using saliency masks," *Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst. - CHI '12*, p. 3011, 2012.

[7] S. H. Lee, J. Y. Lee, and J. S. Choi, "Design and implementation of an interactive HMD for wearable AR system," *2011 17th Korea-Japan Jt. Work. Front. Comput. Vision, FCV 2011*, 2011.

[8] A. T. Duchowski, E. Medlin, N. Cournia, A. Gramopadhye, B. Melloy, and S. Nair, "3D eye movement analysis for VR visual inspection training," *Proc. Symp. Eye Track. Res. Appl. - ETRA '02*, p. 103, 2002.

[9] I. H. Choi and Y. G. Kim, "Head pose and gaze direction tracking for detecting a drowsy driver," *Appl. Math. Inf. Sci.*, vol. 9, no. 2, pp. 505–512, 2015.

[10] A. Doshi and M. M. Trivedi, "On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 453–462, 2009.

[11] H. Skovsgaard, J. Hansen, and E. Møllenbach, "Gaze Tracking Through Smartphones," *Work. Gaze Interact. Post-WIMP World CHI*, 2013.

[12] H. Drewes, A. De Luca, and A. Schmidt, "Eye-gaze interaction for mobile phones," *Mobil. '07 Proc. 4th Int. Conf. Mob. Technol. Appl. Syst. 1st Int. Symp. Comput. Hum. Interact. Mob. Technol.*, vol. 07, pp. 364–371, 2007.

[13] H. C. Lee, D. T. Luong, C. W. Cho, E. C. Lee, and K. R. Park, "Gaze tracking system at a distance for controlling IPTV," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2577–2583, 2010.

[14] E. Wood and A. Bulling, "Eyetab: Model-based gaze estimation on unmodified tablet computers," *Etra*, pp. 3–6, 2014.

[15] M. F. Land, "Eye movements and the control of actions in everyday life," *Prog. Retin. Eye Res.*, vol. 25, no. 3, pp. 296–324, 2006.

[16] a Meyer, M. Böhme, T. Martinetz, and E. Barth, "A single-camera remote eye tracker," *Percept. Interact. Technol.*, vol. 4021, pp. 208–211, 2006.

[17] S. J. Baek, K. A. Choi, C. Ma, Y. H. Kim, and S. J. Ko, "Eyeball model-based iris center localization for visible image-based eye-gaze tracking systems," *IEEE Trans. Consum. Electron.*, vol. 59, no. 2, pp. 415–421, 2013.

[18] C. C. Lai, S. W. Shih, and Y. P. Hung, "Hybrid method for 3-D gaze tracking using glint and contour features," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 1, pp. 24–37, 2015.

[19] A. Tsukada, M. Shino, M. Devyver, and T. Kanade, "Illumination-free gaze estimation method for first-person vision wearable device," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 2084–2091, 2011.

[20] Zhiwei Zhu and Qiang Ji, "Novel Eye Gaze Tracking Techniques Under Natural Head Movement," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 12, pp. 2246–2260, 2007.

[21] P. Smith, M. Shah, and N. da Vitoria Lobo, "Determining driver visual attention with one camera," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 4, pp. 205–218, 2003.

[22] L. M. Bergasa and J. Nuevo, "Real-time system for monitoring driver vigilance," *IEEE Int. Symp. Ind. Electron.*, vol. III, no. 1, pp. 1303–1308, 2005.

[23] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, 2007.

[24] A. Tawari and M. M. Trivedi, "Robust and continuous estimation of driver gaze zone by dynamic analysis of multiple face videos," *IEEE Intell. Veh. Symp. Proc.*, no. Iv, pp. 344–349, 2014.

[25] C. Hennessey, B. Noureddin, and P. Lawrence, "A single camera eye-gaze tracking system with free head motion," *Measurement*, vol. 1, no. March, pp. 27–29, 2006.

[26] V. Vaitukaitis and A. Bulling, "Eye gesture recognition on portable devices," *Proc. 2012 ACM Conf. Ubiquitous Comput. - UbiComp '12*, p. 711, 2012.

[27] Y.-W. Sun, C.-K. Chiang, and S.-H. Lai, "Integrating eye tracking and motion sensor on mobile phone for interactive 3D display," vol. 8856, p. 88560F, 2013.

[28] T. Kocejko, J. Ruminski, J. Wtorek, and B. Martin, "Eye tracking within near-to-eye display," *Proc. - 2015 8th Int. Conf. Hum. Syst. Interact. HSI 2015*, no. 1, pp. 166–172, 2015.

[29] W. W. Abbott and a a Faisal, "Ultra-low-cost 3D gaze estimation: an intuitive high information throughput compliment to direct brain–machine interfaces," *J. Neural Eng.*, vol. 9, no. 4, p. 046016, 2012.

[30] D. LeBlanc, A. Forget, and R. Biddle, "Guessing click-based graphical passwords by eye tracking," *PST 2010 2010 8th Int. Conf. Privacy, Secur. Trust*, pp. 197–204, 2010.

[31] H. Drewes, R. Atterer, and A. Schmidt, "Detailed monitoring of user's gaze and interaction to improve future e-learning," *Univers. Access Hum. Comput. Interact. Ambient Interact. Proc. 4th Int. Conf. Univers. Access Hum. Comput. Interact. UAHCI2007, held as Part HCI Int. 2007*, pp. 802–811, 2007.

[32] S. Milekic, "Using Eye- and Gaze-tracking to Interact with a Visual Display," *Electron. Vis. Arts*, no. Yarbus, pp. 42–48, 2012.

[33] Y. L. Chen, C. Y. Chiang, C. W. Yu, W. C. Sun, and S. M. Yuan, "Real-time eye tracking and event identification techniques for smart TV applications," *Dig. Tech. Pap. - IEEE Int. Conf. Consum. Electron.*, pp. 63–64, 2014.

[34] D. T. Nguyen, K. Y. Shin, W. O. Lee, Y. G. Kim, K. W. Kim, H. G. Hong, K. R. Park, C. Oh, H. Lee, and Y. Jeong, "Gaze detection based on head pose estimation in smart TV," *Int. Conf. ICT Converg.*, pp. 283–288, 2013.

[35] D. M. Small, W. Y. Sanchez, M. J. Hickey, and G. C. Gobe, "Multiphoton fluorescence microscopy of the live kidney in health and disease Multiphoton fluorescence microscopy of the live kidney in health and disease," *J. Biomed. Opt.*, vol. 19, no. 2, p. 020901, 2014.

[36] H. C. Lee, D. T. Luong, C. W. Cho, E. C. Lee, and K. R. Park, "Gaze tracking system at a distance for controlling IPTV," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2577–2583, 2010.

[37] J. H. Oh and N. Kwak, "Recognition of a Driver's gaze for vehicle headlamp control," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2008–2017, 2012.

[38] P. Jiménez, L. M. Bergasa, J. Nuevo, N. Hernández, and I. G. Daza, "Gaze Fixation System for the Evaluation of Driver Distractions Induced by IVIS," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–12, 2012.

[39] X. Sun, L. Xu, and J. Yang, "Driver fatigue alarm based on eye detection and gaze estimation," *Proc. SPIE*, vol. 6786, no. 1, pp. 678612–678612–6, 2007.

[40] A. Kar, S. Bazrafkan, C. C. Ostache and P. Corcoran, "Eye-gaze systems - An analysis of error sources and potential accuracy in consumer electronics use cases," *2016 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, 2016, pp. 319-320.