JSS MAHAVIDYAPEETHA

JSS SCIENCE AND TECHNOLOGY UNIVERSITY

SRI JAYACHAMARAJENDRA COLLEGE OF ENGINEERING

MYSORE - 570 006

Project Report On

"Eye Gaze Tracking"

Submitted by

NAME	USN	Marks-E1	Marks-E2
C.Shravan	01JST18EC028		
Madhusudhan.R	01JST18EC525		
Mohamed Fazal Farhan	01JST18EC055		
Sumantha Ullasa Prabhu	01JST18EC117		

Under the guidance of

Prof. Shashidhar R

Assistant Professor

Department of ECE, SJCE

Table of Content

Acknowledgements

We would first like to thank our supervisor, Prof. Shashidhar R, for his constant guidance, feedback, and support throughout the semester. We have been fortunate enough to have a supervisor who guided us and encouraged us to give our best to the project under difficult circumstances. For providing valuable advice and inspiring us to be creative and innovative with our project. We are indebted to our parents for their continued support and encouragement. We are grateful to them for their help. We would also like to thank all the members of staff at the JSS Science and Technology University for their knowledge and assistance.

Abstract

Recently there has been a growing interest in developing natural interaction between humans and computers. Several studies for human-computer interaction in universal computing are introduced. The vision-based interface technique extracts motion information without any high-cost equipment from an input video image. Thus, a vision-based approach is taken into account as an effective technique to develop human computer interface systems. For vision-based human computer interaction, eye tracking is a hot issue. Eye tracking research is distinguished by the emergency of interactive applications. However, to develop a vision-based multimodal human computer interface system, an eye tracking and their recognition is done. Realtime eye input has been used most frequently for disabled users, who can use only their eyes for input.

List of Figures

Acronyms

AI	Artificial Intelligence
IT	Information Technology
EPOG	Estimated eye point of gaze
HCI	Human Computer Interaction
CV	Computer Vision
ET	Eye tracking
EG	Eye gaze
NP	Numpy (A python package)

CHAPTER 1

1. Introduction

1.1. Overview:

The localization of eye centers has significant importance in many computer vision applications such as human-computer interaction, face recognition, face matching, user attention or gaze estimation. There are several techniques for eye center localization, some of them make use of a head mounted device, others utilize a chin rest to limit head movements. Moreover, active infrared illumination is used to estimate the eye centers accurately through corneal reflections. Although these techniques allow for very accurate predictions of the eye centers and are often employed in commercial eye-gaze trackers, they are uncomfortable and less robust in daylight applications and outdoor scenarios. Therefore, available light methods for eye Centre detection have been proposed.

These methods can roughly be divided into three groups:

1. feature-based methods.
2. Model-based methods
3. hybrid methods.

In this project we describe a **feature-based** approach for eye center localization that can efficiently and accurately locate and track eye centers in low resolution images and videos, e.g., in videos taken with a webcam. We follow a multi-stage scheme that is usually performed for feature-based eye Centre localization, and we make the following contributions: (i) a novel approach for eye Centre localization, which defines the Centre of a (semi-)circular pattern as the location where most of the image gradients intersect. Therefore, we derive a mathematical function that reaches its maximum at the Centre of the circular pattern. By using this mathematical formulation, a fast iterative scheme can be derived. (ii) We incorporate prior knowledge about eye appearance and increase robustness. (iii) We apply simple post processing techniques to reduce problems that arise in the presence of glasses, reflections inside glasses, or prominent eyebrows. The obtained results are extensively compared with state-of-the-art methods for eye Centre localization

1.2. Motivation:

As we all know that there has been tremendous difficulty for specially abled to use new day technology like laptop mobile phone and another personal computer. But the advancement in the technology has brought it to a verge where the computer can be used physically without moving any kind of objects. This recent advancement has motivated us to design and implement a solution with which a personal and industrial computer can be used without contact with appreciably high accuracy. We are creating a software stack which grabs eye gaze, processes it using digital image processing techniques and moves the cursor on the screen relative to position in the eyeball.

We want this solution to be accurate and precise enough for day-to-day use. We also want to reduce the cost of building and maintaining software solutions with the latest robust Technology.

Face is the index of mind and eyes are the window to the soul. Eye movements provide a rich and informative window into a person's thoughts and intentions. Thus the study of eye movement may determine what people are thinking based on where they are looking. Eye tracking is the measurement of eye movement/activity and gaze (point of regard) tracking is the analysis of eye tracking data with respect to the head/visual scene. Researchers of this field often use the terms eye-tracking, gaze-tracking or eye-gaze tracking interchangeably. Eye tracking is mostly used in applications like drowsiness detection [Picot et al. (2010)], diagnosis of various clinical conditions or even iris recognition [Xu et al. (2006)]. But gaze tracking methods can be used in all the ways that we use our eyes, to name a few, eye typing for physically disabled , Cognitive and behavioural therapy visual search [Greene et al. (2001)], marketing/advertising, neuroscience, Psychology and Human Computer Interaction (HCI). Usually the integration of eye and head position is used to compute the location of the gaze in the visual scene. Simple eye trackers report only the direction of the gaze relative to the head (with head-mounted system, electrodes, scleral coils) or for a fixed position of the eyeball (systems which require a head fixation). Such eye tracking systems are referred to as intrusive or invasive systems because some special contact devices are attached to the skin or eye to catch the user's gaze. The systems which do not have any physical contact with the user and the eye tracker apparatus are referred to as non-intrusive systems or remote systems.

1.3. Problem Statement:

Instead of using complicated designed equipment, the proposed system would provide extremely lightweight devices that are more convenient for users to use.

The proposed system is oriented towards the possibility of being used widely, which supports most of the low-cost eye trackers that are affordable for the majority of users.

The proposed system realizes all of the functions of regular input sources, including mouse and keyboard. Users can efficiently interact with computers by only using their eyes.

The proposed system provides a more natural and more convenient communication mechanism for user-computer dialogue and could also avoid annoying users with unwanted responses to their actions.

1.4. Objectives:

- Provide a cheap eye-tracking system.
- To control the cursor of a computer with eyes.
- Allow physically disabled people to use computers.
- To control a computer and communicate with other systems.
- To provide a real-time accurate eye gesture control system.
- To provide a hand free mouse control system.
- To provide a complete generic eye-gesture mouse control system.
- To provide a complete wire free mouse control system.
- Easy to control cursor movement of a mouse.

CHAPTER 2

2. Literature Survey

Paper Title	Method	Merits	Demerits
Shih S.-W., Liu J. A novel approach to 3-D gaze tracking using stereo cameras. *Trans. Sys. Man Cyber. Part B.* 2004;**34**:234–245. doi: 10.1109/TSMCB.2003.811128.	Model based technique	In a model-based technique, the corneal center, which is the point of intersection of visual and optical axes, is considered as an important parameter for gaze estimation. If the corneal curvature is already known, it is possible to determine the corneal center with the help of two light sources and a camera.	For estimation of corneal curvature,anthropomorphic averages are usually adopted due to their simplicity and ease of use. However, if the eye-related parameters are unidentified, at least two cameras and two light sources are required to estimate the corneal center. Several studies, such as used model-based techniques in a fully calibrated arrangement. At a minimum, a single point of calibration is mandatory to estimate the angle between the visual and optical axes..
Hansen D.W., Hansen J.P.,Nielsen M., Johansen A.S., Stegmann M.B. Eye typing using Markov and active appearance models; Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision (WACV 2002); Orlando, FL, USA. 4 December 2002; pp. 132–136.	Single-glint techniques	In initial gaze tracking applications, a single source of IR light was employed to enhance the contrast and consequently produce stable gaze estimates. ● Many single-glint techniques were implicitly based on an erroneous assumption that “the corneal surface is a perfect mirror.”	This assumption inferred that the glint should remain stationary as long as the head position is fixed even when the corneal surface rotates. Therefore, the glint is taken as the origin in glint-centered coordinate systems. In this view, the difference between the pupil center and the glint is utilized to estimate the gaze direction.

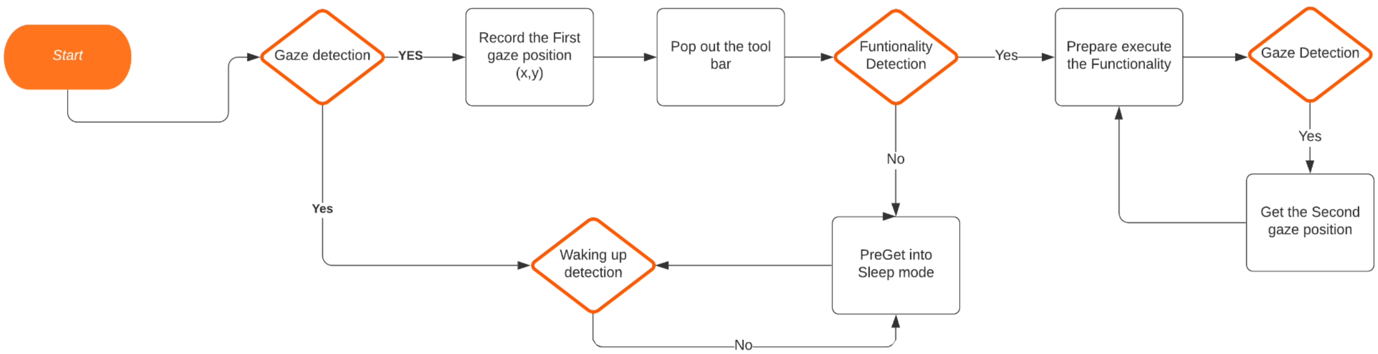
A precise and inexpensive magnetic field search coil system for measuring eye and head movements.	Scleral Search Coils	The advantage of such a method is the high accuracy and the nearly unlimited resolution in time.	It is an invasive method, requiring something to be placed into the eyes. To the best of our knowledge, this method of eye tracking has not been used for HCI by gaze, so far. This method is mostly used in medical and psychological research.
https://www.researchgate.net/publication/327578844_ELECTROOCULOGRAPHY_ANALYSIS_ON_DEVICE_CONTROL_BY_SIGNAL_PROCESSING	ELECTROOCULOGRAPHY	●Sensors are attached at the skin around the eyes to measure if an electric field exists when eyes rotate. ●It is a cheap, easy and invasive method of recording large eye movements. ●By recording small differences in the skin potential around the eye, the position of the eye can be estimated. By carefully placing electrodes, it is possible to separately record horizontal and vertical movements.	●However, the signal can change when there is no eye movement. ●This technique is not well-suited for everyday use, since it requires the close contact of electrodes to the user but is still frequently used by clinicians.
Pérez, A & Córdoba, Mª & Garcia-Dopico, Antonio & Méndez, R & Muñoz, Maria & Pedraza, J.L. & Sánchez, F. (2003). A precise eye-gaze detection and tracking system.	Infrared Oculography	●The light source and sensors can be placed on spherical glasses. Hence it is an invasive method. The infrared oculography has less noise than electro-oculography, but is more sensitive to changes of external light tension.	●The main disadvantage of this method is that it can measure eye movement only for about ±35 degrees along the horizontal axis and ±20 degrees along the vertical axis. .

Table1: Literature Survey

CHAPTER 3

3. Methodology

3.1 Block Diagram:



3.2 Procedure:

To implement the whole project, we will use the Python language for Image Processing, the block diagram of which is shown in Fig. 3.1. To detect the lanes and extract the features of the eye balls in every frame of the video. we will use eye center localization, We incorporate prior knowledge about the eye appearance and increase the robustness. apply simple post processing techniques to reduce problems that arise in the presence of glasses, reflections inside glasses, or prominent eyebrows.

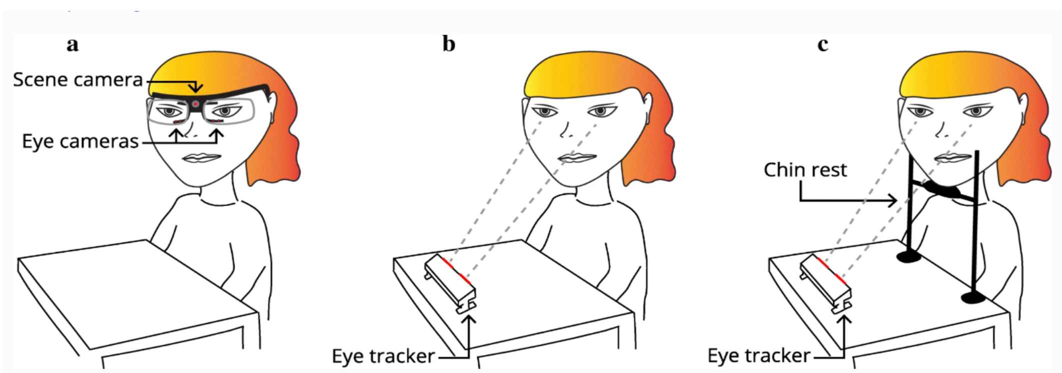


Fig:3.2

An illustration of the different types of eye-tracking setups: an example of a head-free setup containing a wearable eye tracker. b An example of a head-boxed setup containing a remote eye tracker. c An example of a head-restricted setup containing a remote eye tracker and a chin rest

The four main core concepts that we intend to use in this project are:

- Eye center Localization.
- Prior knowledge of facial features.
- Post Processing.

3.3 Algorithm:

Step 1: The image of the user is captured using the laptop/computer using the open source library called OpenCV in Python.

Step 2: The captured image is preprocessed and fed into the library to extract the face from the image. The library that provides this framework is written in C++ and is called dlib.

Step 3: The approximate positions of the eyes are captured from the extracted face using the various facial features.

Step 4: The Fabian-Timm algorithm is applied on the extracted eyes to find the exact position of the iris in both the eyeballs.

Step 5: Two variables are calculated from the relative positions of the eyes as perceived from the camera. The variables are horizontal & vertical ratios.

Step 6: The two ratios calculated are then mapped to the relative coordinates on the screen.

Step 7: The computer's cursor is then moved to the absolute position on the screen as perceived from the camera.

$$\mathbf{c}^* = \arg\max_{\mathbf{c}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(\mathbf{d}_i^T \mathbf{g}_i\right)^2 \right\} ,$$

$$\mathbf{d}_i = \frac{\mathbf{x}_i - \mathbf{c}}{\|\mathbf{x}_i - \mathbf{c}\|_2} , \quad \forall i : \|\mathbf{g}_i\|_2 = 1$$

$$e \leq \frac{1}{d} \max\left(e_1, e_{\mathrm{r}}\right)$$

- ***c*** be a possible centre
- ***gi*** the gradient vector at position xi on the image.
- ***di*** is the normalised displacement vector.

- ***c**** is the optimal center of the circular object in the image which here is the center of the eye.
- As an accuracy measure for the estimated eye centres, we evaluate the normalised error.

CHAPTER 4

4.Result and Discussion

We were quite fascinated with the results that we observed. Initially we implemented an algorithm to detect faces from the captured image. Then we processed the image to extract the face and then further algorithms were applied to extract the facial features. Most specifically the eyes.

We then applied the Fabian Timm Algorithm to find the center of the semicircular objects within the image. Here the image is the extracted part of the eyes from the captured image from the camera.

After finding out the exact location of the center of the eye, we applied another algorithm to find 2 variables namely Horizontal and the Vertical ratios. Where were then mapped on to the screen depending on the relative positions of the user from the camera.

This mapping was then used to determine the location at which the user was looking at and then the cursor was moved to that position using proportional and integral login.

Calibration

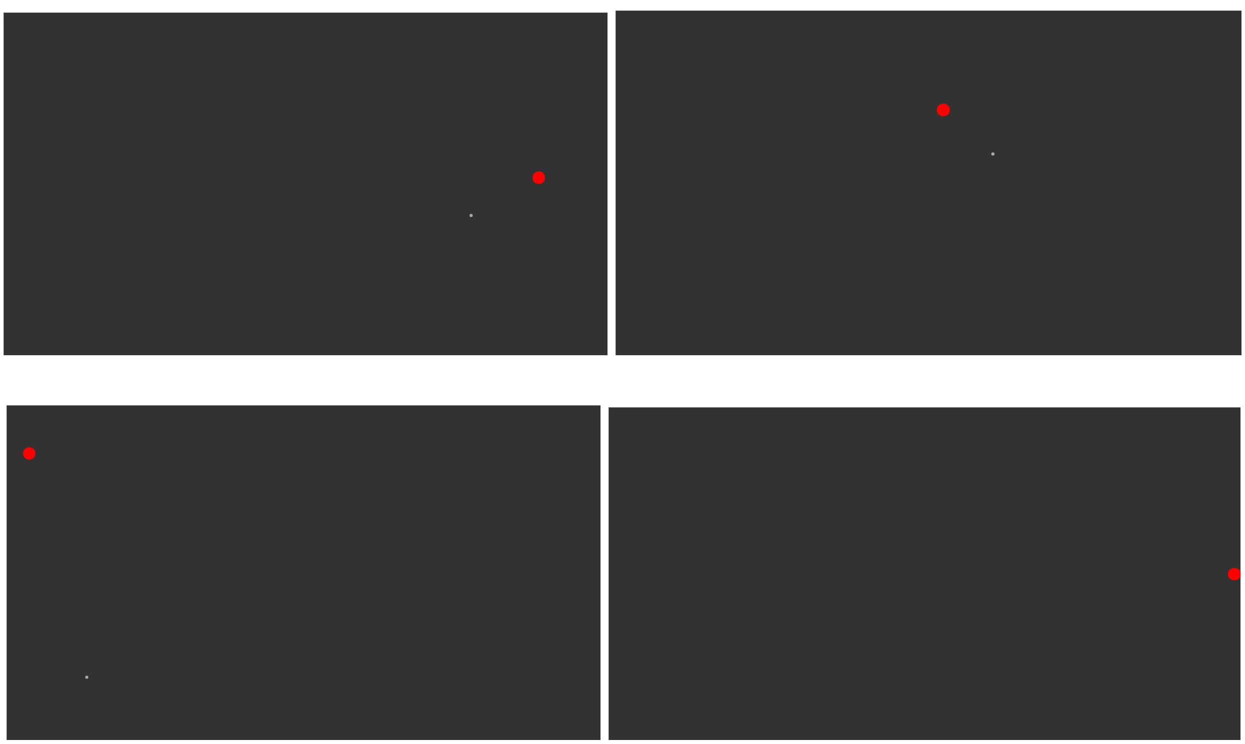


Fig4.1:Calibration Result

Results

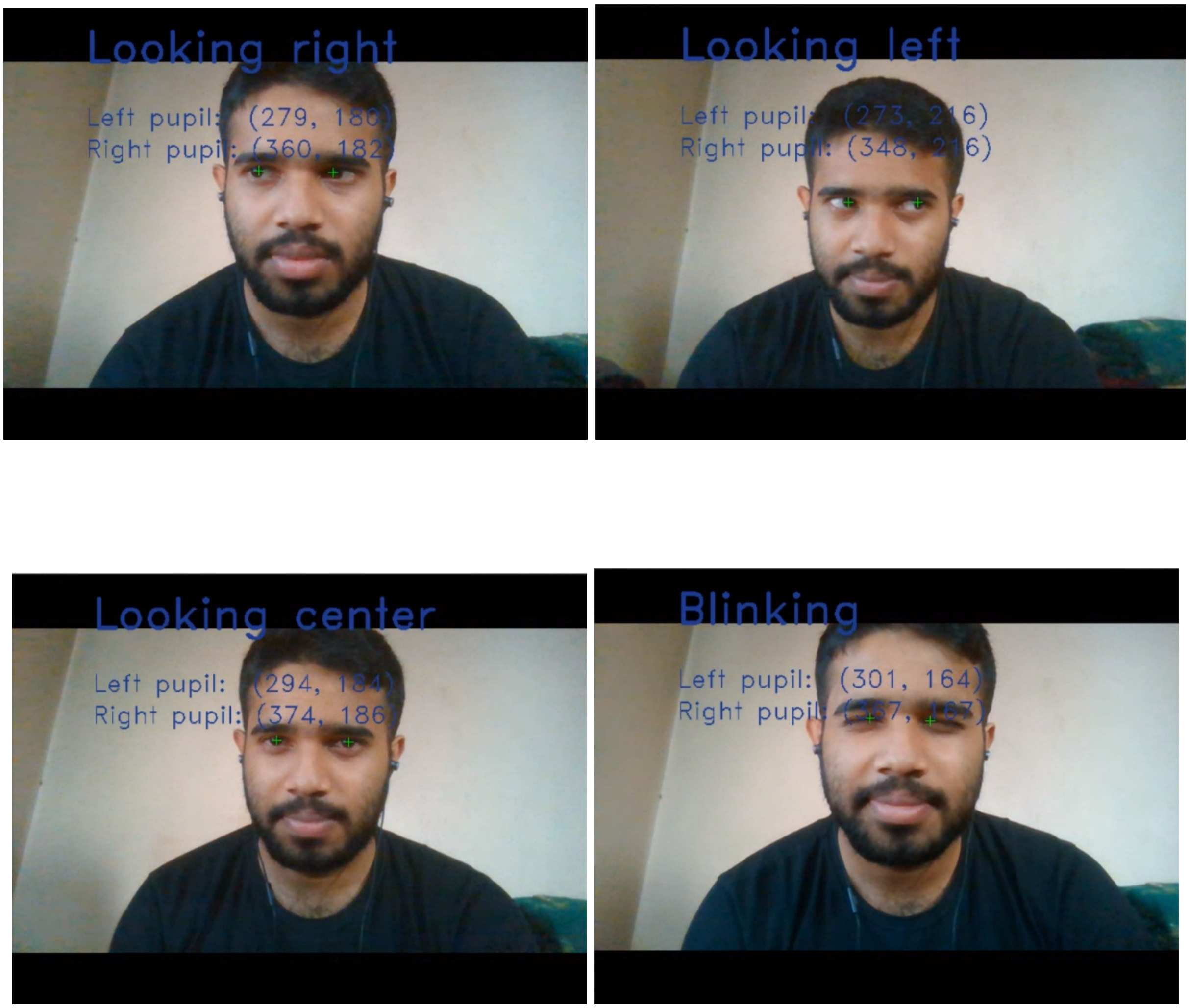


Fig.4: Eye Gaze Tracking Result

CHAPTER 5

5. Conclusion and Future Scope

5.1 Conclusion

We have presented a brief introduction to eye tracking as a technique and shown how the different types of eye trackers can be incorporated into setups used to investigate human interaction. Following, a scoping overview of the eye-tracking literature in human interaction was provided and several studies were reviewed and categorized depending on the amount and type of eye trackers used.

5.2 Future Scope

Since many interesting research questions on social interaction might be answered in real face-to-face interactions, follow-up studies need to validate the present setup in dual mode, using two eye trackers allowing the recording of eye movements of two participants while interacting. A possible application of this setup could be therapeutic sessions in which the gaze behavior of the patient during interaction with the therapist is recorded and might serve as a predictor for the quality of therapeutic alliance or therapy outcome. Moreover, the application of correction algorithms could further improve eye tracking data quality in 3D environments. Finally, a combination of eye tracking with unobtrusive motion tracking used in other setups could allow the assessment of non-verbal behavioral synchrony in social interactions

In this design, the application distance between the user’s head and the screen is fixed during the operation, and the direction of the head pose tries to keep fixed. In future works, a head movement compensation function will be added, and the proposed wearable gaze tracker will be more convenient and friendly for practical uses. To raise the high-precision recognition ability of the pupil location and tracking, the deep-learning model will be updated with online training using the big image databases in the cloud to fit the pupil position for different eye colors and eye textures.

5.3 Software Used:

1. Dlib: Dlib is a general purpose cross-platform software library written in the programming language C++. Its design is heavily influenced by ideas from design by contract and component-based software engineering. Thus it is, first and foremost, a set of independent software components.
   Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. It is used in both

industry and academia in a wide range of domains including robotics, embedded devices, mobile phones, and large high performance computing environments.

2. Python: Python is an interpreted high-level general-purpose programming language. Python's design philosophy emphasizes code readability with its notable use of significant indentation.
3. Open CV: OpenCV is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage then Itseez.
4. Numpy: NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

5.4 References:

1. Shih S.-W., Liu J. A novel approach to 3-D gaze tracking using stereo cameras. Trans. Sys. Man Cyber. Part B. 2004;34:234–245. doi: 10.1109/TSMCB.2003.811128.
2. Hansen D.W., Hansen J.P., Nielsen M., Johansen A.S., Stegmann M.B. Eye typing using Markov and active appearance models; Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision (WACV 2002); Orlando.
3. Electro-Oculography:https://www.researchgate.net/publication/327578844_ELECTROOCULOGRAPHY_ANALYSIS_ON_DEVICE_CONTROL_BY_SIGNAL_PROCESSING.
4. H.J.KasperB.J.M.HessN.Dieringer Brain Research Institute, University of Zürich, Zürich ,Switzerland.
5. Infrared Oculography Pérez, A & Córdoba, Mª & Garcia-Dopico, Antonio & Méndez, R & Muñoz, Maria & Pedraza, J.L. & Sánchez, F. (2003). A precise eye-gaze detection and tracking system.
6. D. Sheena and B. Borah, "Compensation for second-order effects to improve eye position measurements, pp. 257-268, in "Eye movements: Cognition and visual perception'", University of Illinois Press, 1981.
7. S.V. Wass, T.J. Smithand M.H. Johnson, "Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults'", Behav. Res. Methods, Vol. 4, pp 229-250, 2013.
8. O. Geman, H. Costin, "Parkinson's Disease Prediction based on Multistate Markov Models", Int. J. of Computers, Communication and Control, ISSN 1841-9836, 8(4): 525-537, August, 2013.
9. A. Moldoveanu, F. Moldoveanu, V. Asavei, A. Egner and A. Morar, "From HTML to 3DMMO - a roadmap full of Challenges", Adv. In Intelligent Control Systems & Computer Science, AISC 187, I. Dumitrache Ed., Springer-Verlag Berlin Heidelberg, 2013, pp. 379-392.

10. S. Karthikeyan, V. Jagadeesh, R. Shenoy, M. Ecksteinz, and B. Manjunath. From where and how to what we see. In ICCV, 2013. 1
11. A. Khosla, A. S. Raju, A. Torralba, and A. Oliva. Understanding and predicting image memorability at a large scale. In ICCV, 2015. 2, 3
12. A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012. 2, 5, 6, 8
13. F. Lu, T. Okabe, Y. Sugano, and Y. Sato. Learning gaze biases with head motion for head pose-free gaze estimation. Image and Vision Computing, 2014. 2
14. F. Lu, Y. Sugano, T. Okabe, and Y. Sato. Adaptive linear regression for appearance-based gaze estimation. PAMI, 2014. 2
15. P. Majaranta and A. Bulling. Eye tracking and eye-based human– computer interaction. In Advances in Physiological Computing. Springer, 2014. 1
16. C. D. McMurrough, V. Metsis, J. Rich, and F. Makedon. An eye tracking dataset for point of gaze detection. In ETRA, 2012. 2