Design a Real-Time Eye Tracker

Sara Bilal
Senior Lecturer
Whitieria New Zealand
+64211257123
Sara.Bilal@whitireia.ac.nz

Mohamad Hassrol Bin Mat Hussin
Engineer
IIUM
hassrol14@gmail.com

Rasheed Nassr
Senior Lecturer
University of Kuala Lumpur
rasheed@unikl.edu.my

ABSTRACT

An eye tracker is a device for measuring eye positions and eye movement. During the past decades, various approaches have been presented for eye tracking systems such as helping disabled people in various tasks including communication, writing emails, and drawing, besides other applications such as cognitive studies, electronic games and commercial eye trackers. The eye tracking devices and the algorithms that track the gaze still need an improvement. Many critical problems arise from the fact that the eye tracker system does not retain its accuracy for a long period of time. Besides, the fact that the commercial eye trackers are expensive. The aim of this research work is to develop a typing by gaze system using a PS3 camera. The developed typing by gaze system uses eyes' gaze for typing on a virtual keyboard. As well, the developed system is tested using a built-in laptop camera and a mounted PS3 camera on an eyeglass. First, the system will detect the face from the video stream. Followed by eyes' pupil detection and tracking which has been achieved using template-matching technique and adaptive EigenEye method. In addition, a virtual keyboard has been designed based on a laptop-size keyboard with QWERTY layout. After the eye pupil is tracked, the system will take control of the mouse's cursor movement. Then the eye's gaze will become the pointing device and the selection of the focused key or button will be achieved by using dwell-time of 0.5seconds. The developed eye tacker is an affordable device and it has retained a good typing accuracy either by using the PS3 or the laptop camera. Hence, it can be used for several applications and especially for research purposes.

CCS Concepts

•Computer systems organization ~ Real-time system architecture

Keywords

Eye tracking; Mounted camera; PS3 camera; Virtual keyboard

1. INTRODUCTION

In the recent years, eye-tracking systems turned out to be an important tool to increase the quality of life and enhance the autonomy of persons with disabilities needing alternative input devices [1][2]. An eye-tracker is a device that applies projection patterns and optical sensors to accumulate data about eye position, gaze direction or eye movements with very high accuracy. Most eye-trackers systems are actually based on the fundamental principle of corneal reflection tracking as shown in Figure 1.

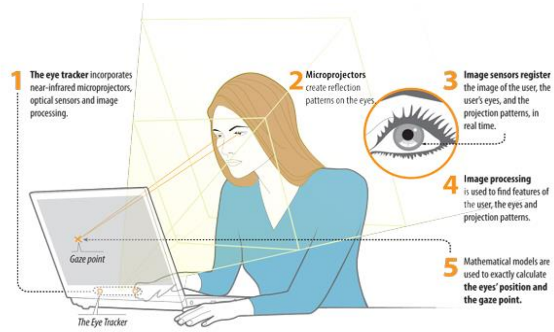


Figure 1: An eye-tracker [2]

A convenient, natural and high-bandwidth source of input is provided and can be attained by the movement of user's eye. Just by tracking the direction of the user's gaze, or the center pupil of the eye, the subject or information about what the user is looking at can increase the communication bandwidth from the user to the computer. However, to track the eye and its movement is not an easy task especially the pupil. Generally, eye movement is divided into two categories which are fixations and saccades. A fixation usually happens when the eye gaze pauses in a certain position. In the eye-tracking situation, fixations are used to denote a starting point for all eye movements. Saccades are when the eye gaze moves to another position. Humans alternate between fixations and saccades. This behaviour corresponds to the scan paths which are used to analyse cognitive intent, interest and importance. The fovea provides the bulk of visual information. The periphery is less informative than the fovea [3]. Figure 2 shows the fovea located in the macula region of the retina. The locations of the scan paths during eye-tracking signify the information that was processed. Fixations normally last for 200 ms when reading text and 350 ms when viewing a scene.

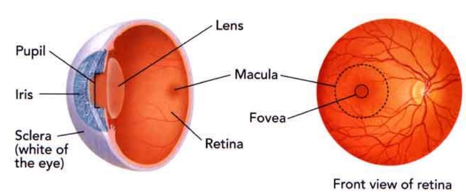


Figure 2: Details part of the eye [3]

In this paper, the existing eye-tracking systems has been surveyed and the methodology to design an affordable eye-tracker is described (see Figure 3). The idea of the system is to mount an infrared camera onto a pair of eyeglass and capture the movement of the pupil using a software. Following that, the developed typing by gaze system was conducted using two systems: 1) a distanced

ICVIP 2018, December 29–31, 2018, Hong Kong, Hong Kong

ACM ISBN 978-1-4503-6613-7/18/12…$15.00
https://doi.org/10.1145/3301506.3301509

camera which requires the detection of the whole face and then locate the eye. 2) A wearable camera device by focusing the camera on the eye.

2. RESEARCH BACKGROUND

2.1 GAZE INPUT

Gaze is naturally used to obtain visual information. For example, gaze location shows the focus of attention. As an input method, gaze has both advantages and disadvantages. It is a natural mode of input as it is easy to focus on items by looking at them. Another advantage is that the target acquisition using gaze is very fast, provided the targets are large enough. However, gaze is not as accurate as the computer mouse. Inaccuracy originates partly from technological reasons and partly from features of the eye [4]. The size of the fovea and the inability of the remote camera to resolve the fovea position restrict the accuracy of the measured point of gaze to about 0.5 degrees, equivalent to a region spanning approximately 15 pixels on a typical display (17 inch display with a resolution of 1,024 × 768 pixels viewed from a distance of 70 cm). One problem is drifting; even if a newly calibrated eye tracking device is accurate at first, with continued use the measured point of gaze drifts away from the actual point of gaze. This is partly due to the technology and partly due to the interaction between head movement and eye movement. Therefore, the practical accuracy is about 0.5–1 degree of drifting corresponds to about 1–1.5 cm on the screen at a normal viewing distance. When we look at things, we fixate (focus) on them, with fixations typically lasting from 200 to 600 ms [4]. For a computer to distinguish whether the user is looking at an object to obtain information or to select it, an interval longer than the typical fixation interval is needed. Stampe and Reingold [5] used a dwell time (an extended look at the object) of 750 ms in their eye typing study. A thousand milliseconds is usually long enough to prevent false selections. For simple tasks, 700 ms or less is enough. Requiring the user to fixate for long intervals is good for preventing false selections (thus, preventing the Midas touch problem), but this is uncomfortable for most users [5]. It is important to note that the use of a dwell time criterion for key selection places an upper limit on eye typing speed. In other words, these is no skilled acquisition system exists that will allow a user to "eye press keys" at a rate faster than 1 / td, where td is the dwell time. If, for example, td = 1,000 ms = 1 s, the upper limit for typing speed is (60 / 1) / 5 = 12 words per minute (wpm) (following the accepted method in computing typing speed of 1 word = 5 inputted characters).

2.2 COMPARISON BETWEEN the TYPING BY GAZE SYSTEMS

Nowadays, there are many systems of typing by gaze such as Tobii PCEye Go [6], The Eye Tribe Tracker [7], Eyegaze Edge Desktop [8] and myGaze Assistive 2 [9]. The Tobii PCEye Go is a peripheral eye tracker that enhances computer accessibility with the speed, power and accuracy of gaze interaction. The device replaces the standard mouse, allowing you to navigate and control a desktop or laptop computer using only your eyes. Tobii PCEye Go system is using an integrated keyboard or built in keyboard to write texts or to enter a web address. Whereas for The Eye Tribe Tracker, the information extracted from person’s face and eyes is used to calculate the location from where a person is looking. The calculation is based on a pair of eye gaze coordinates represented by (x, y) on the screen coordinate system. The Eyegaze Edge Desktop is an eye-operated communication and control system that empowers people with disabilities to communicate and interact with the world. By looking at control keys or cells displayed on a screen, a user can generate speech either by typing a message or selecting pre-programmed phrases. Eyegaze edge systems are used to write books, attend school and enhance the quality of life of people with disabilities all over the world. This system has an adjustable monitor arm with camera bracket and using a high-speed infrared sensitive camera and lens. Eyegaze Edge Desktop system already design their own on-screen keyboard for the typing. In addition, MyGaze Assistive 2 is a gaze control system using computer programs which is affordable. It is well suited for special needs students and their teachers as well as adults who rely on gaze for communication and environment control. The myGaze Eye Tracker relies on a technology by SensoMotoric Instruments (SMI) from Germany, a leading developer of state-of-the-art eye tracking solutions for more than 20 years. This system has less than 50ms system latency, strong robustness and reliable performance. MyGaze Assistive 2 is also an Instant gaze tracking that its calibration is minor. The system will instantly tracks the users gaze. The size of these products is extremely portable, except for the Eyegaze Edge Desktop which is bigger like laptop-size. Table 1 shows a clear view of the comparison however, there are some factors that affect the commercial product such as cost effectiveness, robustness under different environments, real-time application speed, and end-user acceptability.

Table 1: Comparison between existing typing by gaze systems

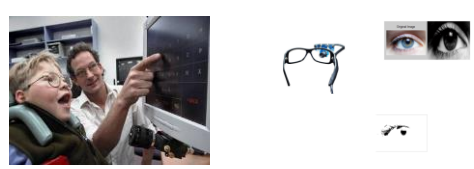
system	Camera	keyboard	Price	Advantage(s)	Limitations
Tobii PCEye Go [6]	Infrared LED	On-screen keyboard integrated	$1,995	- user can wear contact lenses. -Can hit the smallest targets	Expensive
The Eye Tribe Tracker [7]	Infrared LED	On-screen keyboard	$99	-Very cheap -USB 3.0 superspeed	user can’t wear contact lenses
Eyegaze Edge Desktop [8]	Infrared mounted camera	On-screen keyboard	£4,250-$6,446.21	Fastest eye gaze system on the market	-Very expensive -Bigger
myGaze Assistive 2 [9]	Infrared LED	On-screen keyboard	€499-$557.83	-Cheap -Calibration-less mode	Hard to pause interaction

On the other hand, there are several existing free eye tracker software such as IMOTIONS, XLAB [10]. However, their cons is more than their pros. Hence, developing an affordable eye tracker device apart from the computer cost is a demand in the research sector.

3. METHODOLOGY

3.1Eye Tracking Systems

Developing an eye tracker can be based on tracking the eye from a distance or using a wearable camera as illustrated in Figures 3(a) & (b).



(a) **(b)**

Figure 3. (a) Distanced camera [6], (b) Wearable Camera device.

3.1.1 Eye Tracker Using a Wearable Camera

The process of acquiring eye image is done using a PS3 camera that is mounted on a eyeglass. To build this device, the components are listed in Table 2 with their respective price in Ringgit Malaysia (*RM*).

Table 2: Cost summary to build an eye-tracker

No	Components	Quantity	Price per unit (*RM*)	Price (*RM*)
1	PS3 Eye Camera	1	88.00	88.00
2	Eyeglass	1	15.00	15.00
3	IR LED Transmitter	1	1.50	1.50
4	IR LED Receiver	1	1.50	1.50
5	Alligator clips	4	2.60	10.40
6	Pack of wire ties (short)	1	1.50	1.50
7	Pack of wire ties (long)	1	1.90	1.90
8	AAA size battery holder	1	1.00	1.00
9	3 × AAA battery holder	1	1.50	1.50
10	4 × AAA battery holder	1	2.00	2.00
11	9mm Aluminium Wire Gauge	1	15.00	15.00
12	Camera Lens Mount	1	14.00	14.00
13	IR infrared filter gel	1	20.00	20.00
14	Electrical junction connectors	2	13.00	26.00
15	Screws	5	2.20	11.00
			Total	210.30

3.1.1.1. PS3 Camera Configuration

PS3 camera is chosen among all other cameras because it small in size, performs well in variable lighting conditions, has stable USB performance, and capable of high frame rates. It also comes with the resolution of 640×480 pixels which is at 60Hz frequency that gives a solid performance. Also, an eye image that captured by this PS3 camera will give a smoother image since the maximum video frame rate is 120fps. The higher the frame rate, the smoother the screenplay and the image capturing. From the optics perspective, it uses wide angle lens with field of view of 75°. Moreover, the minimum focus distance allowed are 9.84 inches ≈ 25 cm, and the low light adjustment feature also available. PS3 camera is used because even when there is less light, the PS3 camera's infrared light can still provide tracking information with high accuracy because of the presence of the IR light. Basically, the PS3 camera will be dissembled from its body frame and mounted on the eyeglass and the eyeglass's lens will be removed from its frame too. The wire tier is used to secure and tie the camera arm to the eyeglass's frame.

3.1.1.2 Capturing Eye Image Using Wearable Camera

To capture an eye image, the camera is mounted on the eyeglass and is located right in front of the patient's eye so that it can capture the eye image accurately. The camera is selected based on the quality images captured, the frequency and price. There are some cameras which have been used for eye-tracking such as HATCAM, basic web camera and PS3 camera [11]. The distance between the camera and eye must be seriously considered. The distance cannot be too far or too close to the eye. Hence, in this project, the distance of the camera to the eye is fixed to 3.2 cm.

3.1.2 Eye Pupil Detection And Tracking

Most of the image based methods are detecting the eyes' location using features of the eyes which are known as knowledge-based methods, feature-based methods, simple template-matching and appearance-based methods. Another interesting method is "Deformable template-matching" which is based on matching a geometrical eye template on an eye image by minimizing the energy of the geometrical model. To locate and track the eye pupil, the original image of the eye must be converted to greyscale image first. Then, we need to apply two methods that is template matching technique and adaptive EigenEye method.

3.1.2.1 Eye Pupil Tracking Using Adaptive EigenEye Method

EigenEyes is used as core components of the original images. This EigenEyes is the product of decomposition of eyes image by the adaptive EigenEye method. These Eigeneyes have a role as the orthogonal basis vectors of a subspace called eyespace. In this work, we use the Principal Component Analysis (PCA) methods to extract features for eye tracking. PCA is a well-known unsupervised algorithm for linear feature extraction; it is a linear mapping that uses the eigenvectors with the largest eigenvalues. The advantage from this method is the detection process is a simple step throughout the whole algorithm. For PCA, a set of 860 right eye images, 860 left eye images and 320 non eye images have been used. The eye-template image of 40×26 pixels in the training database is defined as the column vector of 1040×1 pixels. The column vector is then reduced to an eigenvector with dimensions by using PCA, which seeks a projection that best represents the original data in a least-squares sense.

3.1.2.2 The localisation of eye centers

There are many commercial products to locate and track the eye gaze such as [6, 7] with are expensive hardware headset. In this research work, webcam gaze tracker which is based on a simple image gradient-based eye center algorithm is used. Timm and Barth, (2011) proposed image gradients-based approach to accurately find the eye centre. Their approach was used here because they have used simple objective function, which consists of dot products. The maximum of this function matches the location where most gradient vectors intersect and thus to the eye's centre. This simple approach, it is invariant to changes in scale, pose, contrast and variations in illumination [12]. Therefore, in an environment with a low resolution images and real time application this approach is good enough to be used for gaze tracking. To determine the eye region fractions, let (x, y) be the upper left corner and W, H the width and height of the detected face. Then, the mean of the right eye centre is located at $(x + 0.3, y + 0)$ and the mean of the left centre is at position $(x + 0.7, y + 0.4)$. The Gaussian blur was used before processing face to smooth noise. The gradient algorithm used in this implementation is as follows: The gradient function creates vectors that always point towards the lighter region. Since the iris is darker than the sclera, the vectors of the iris edge always point out. It means that at the center they will be facing in the same direction as the d vector. This method has simplified tracking a reference point such as eye corner to accurately judge where the user is looking. However; this approach first detect the face then determine the location of the eye's center later; this approach wouldn't help with application that require using camera that is only focused on one eye in order to accurately track eye movement. Even though the previous mentioned approach works, however tracking eyes by using a special camera focused on one eye seems to be more efficient tracking the eyes.

3.1.2.3 Track Single-Eye Pupil Using PS3 Camera

To track an eye and find the center coordinates of pupil, Houghcircles was excluded because it didn't perform well and hence, hue, saturation and val (HSV) filter was used. First, the images are thresholded then a contour finding algorithm was used to find the centroid of all the contours. This gives the center coordinates of the eye pupil. This method was working with high efficiency in real time even when the eye is blinking. In this work, PS3 camera was used. This camera comes with API that can help to run this camera. Then by integrating the previous code with the API of the PS3 camera, we develop a prototype using VC++ 2010 with OpenCV. Sample pictures generated from previous code are presented in the Figure 4.

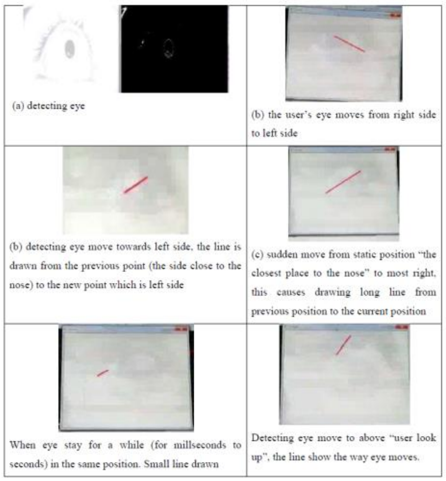


Figure 4: Tracking eye results

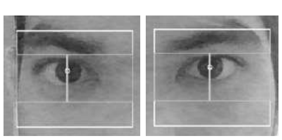
4. Experimental Results

4.1 Eye Pupil Detection and Tracking

First, the camera captured the image or the frames that contain the face and eye pupil. Following that, the image is converted to grayscale and the face is detected by using Haar-like features [13]. Figure 5 shows the result of face detection. The blue rectangle indicates that the face region has been successfully detected.

Figure 5: Face detection.

Next, the captured image also contains the eye pupil. Template-matching technique and EigenEye method has been used to track the eye pupil as shown in Figure 6.



(a) (b)

Figure 6: Pupil detection of (a) Right eye and (b) Left eye.

4.2 Mouse Control

After the tracking process is done, the system will proceed to the mouse control. It will capture the current position of mouse first. Then by using some calculation it will set the position of mouse cursor corresponding to the focus point. The movement of eye will move the mouse cursor. Figure 7 shows the result of scaling up the mouse pointer. Because of the small size of eye pupil region compared to the resolution of desktop, the scale factor or ratio become too big. Thus, it make '1' pixel of eye pupil movement equal to about '45' pixels horizontal movement and '41' pixels vertical movement on the actual screen. Therefore, it will reduce the accuracy of the cursor as the buttons size is about '70' pixels square.

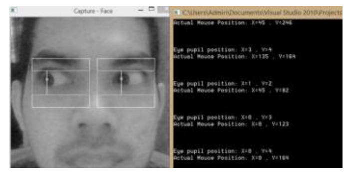


Figure 7: Result of the mouse position around the initial point

4.3 Virtual Keyboard

4.3.1. Virtual Keyboard Design

The virtual keyboard was designed using Visual Basic 2010. The keyboard is using standard-laptop keyboard and QWERTY layout. (See Figure 8)

Figure 8: Virtual keyboard for this gaze typing system.

4.3.1.1 Virtual Keyboard Layout

For the virtual keyboard, when any button is pressed, the system will copy the name or the text of that button and send into textbox. For example, when the user clicks button 'e', the word 'e' will be sent to the textbox (see Figure 9 (a)). If the person needs to use any symbol that is not active at that time which means behind the numeric button, he/she needs to use the "SHIFT" key to activate all the symbols and presses or clicks the desire symbol (see Figure 9(b)). The user also can copy any of the word or sentence by pressing the "COPY" button and paste back to textbox by pressing "PASTE" button (see Figure 9 (c)). If user want to clear everything in the textbox can press "CLEAR" button which will clear everything automatically rather than using "BACKSPACE" button repeatedly (see Figure 9 (d)).

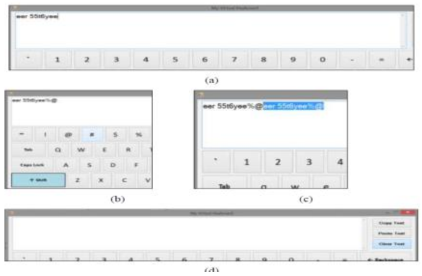


Figure 9: (a) Some buttons are clicked; (b) SHIFT button is clicked to active the symbols and capital letter; (c) COPY and PASTE button are clicked; (d) Textbox is cleared

4.4 Dwell-time Method

Dwell-time method means that the mouse will automatically be clicked when the focus time has reached the desired limit which is 0.5 second. Focus time will start counting when the mouse pointer enters or hovers to any button. The focus time will start new count when the previous count has reached 0.5 second. This developed system also incorporates a sound that will play automatically on the same time with the click. This can increase the typing speed without having to look at the textbox to see whether the character is typed or not because the sound has indicated that the typing is successfully performed.

4.5 FINAL RESULTS

The pupil was extracted to exploit its higher contrast than the background due to the IR components of the natural light. Segmentation is applied to detect the pupil successfully using grey level images and other features' extraction techniques. Following that, the system took control of the mouse cursor. So the users can use their eyes for typing by clicking the buttons on the virtual keyboard. This is achieved by implementing the dwell-time method. The time for an automatic click is 500ms (0.5 second). So the user has to focus on any button at the virtual keyboard. Then the button will be highlighted and automatically send the button's text into textbox or message box. (See Figure 10)

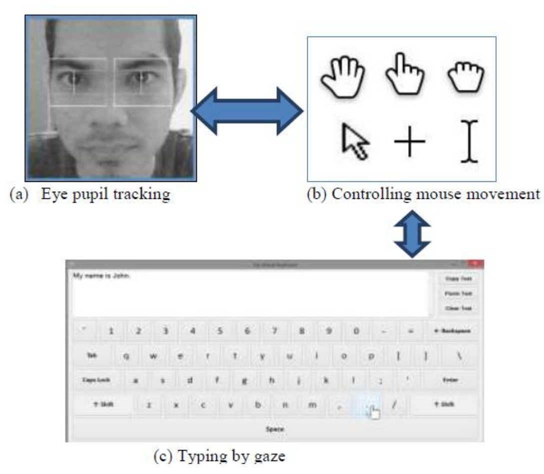


Figure 10: Process of typing using the eye gaze

5. CONCLUSION

Eye tracker is a communication and control system that has been implemented in many applications and it empowers people with disabilities (parallelized people) to communicate and interact with the environment. There are two methods to develop an eye tracker which are based on tracking the eye from a distance or by using a wearable camera. However, by using a wearable camera, the eye tracking system can perform better. The existing eye-trackers are quite costly, therefore the cost of building an eye-tracker using PS3 is affordable. After constructing the eye tracker hardware where an infrared camera has to be mounted onto an eyeglasses and capture the movement of the pupil using VC++ and OpenCV. Following that, the template matching method is used to locate and track the centre of the pupil, where a small patch of dark pixels are used as a template. The search space in each frame could be defined based on the estimates of the pupil location from the previous frames. As a result, this affordable constructed eye tracker can be used to conduct research as well it can help disabled patients in various tasks such as communication, writing emails, drawing and gives them a sense of interaction with the surroundings. In the future, a mobile eye tracker device can be developed using these approaches.

6. REFERENCES

[1] Alan Watt, Fabio Policarp, The Computer Image, Addison Wesley, (1999). Antonio Lanatà,

[2] Lemahieu , W., & Wyns, B. (2011). Low cost eye tracking for human-machine interfacing. Journal of Eye Tracking, Visual Cognition and Emotion nº (8), http://hdl.handle.net/10437/22971

[3] "Age-Related Macular Degeneration" Retrieved March 13, 2016 from http://www.urbaneyemd.com/glossary_details.php?id=77

[4] Jacob, Robert JK. 1995. Eye tracking in advanced interface design. In: Barfield W, Furness TA (eds) Virtual environments and advanced interface design. Oxford University Press, New York, pp 258–288.

[5] Ware C., Mikaelian H.H. 1987. An evaluation of an eye tracker as a device for computer input. In: Proceedings of CHI/GI '87. ACMPress, Cambridge, pp 183–188. DOI 10.1145/29933.275627

[6] "What is eye-tracking". Retrieved April 4, 2014 from http://www.tobii.com/en/about/what-iseye- tracking/

[7] TheEyeTribe. 2015. The eye tribe tracker. Retrieved April 4, 2014 from https://theeyetribe.com/products/

[8] XLab.. 2015. Xlabs, eye gaze and head via the webcam. Retrieved April 4, 2014 from http://xlabsgaze.com/

[9] MyGaze Eye Tracker, Retrieved April 4, 2014 from http://www.mygaze.com/products/mygaze-eye-tracker/

[10] 10 Free Eye Tracking Software Programs. Retrieved 5/2/2018 from https://imotions.com/blog/free-eye-tracking-software/,

[11] Takayuki Ito, Kyle Mcdonald, Golan Levin and students . 2010. Eyewriter collab at Parsons MFADT, (2010). The Eye-Writer 2.0.

[12] Timm, F., & Barth, E. 2011. Accurate Eye Centre Localisation by Means of Gradients. In VISAPP (pp. 125-130). https://doi.org/10.5220/0003326101250130

[13] Viola, P., & Jones, M. J. 2004. Robust real-time face detection. International journal of computer vision, 57(2), 137-54. https://doi.org/10.1023/b:visi.0000013087.49260.fb