International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

Automatic Cursor Movement with Eye Gaze Tracking

Priyanka Jadhav[1], Rupali Jagtap[2]

[1, 2]Department of E&TC, Annasaheb Dange COE, Ashta, Shivaji University, Sangli, India

Abstract: ***Most of the eye tracking devices is based on computer vision imaging systems. This paper presents a complete system for eye tracking through avoiding the restrictions that are caused due to the head movements. The Viola-Jones algorithm, in combination with Kalman Filtering, is utilized for detecting and tracking the face. The Viola –Jones-based eye tracker that is employed here offers eye tracking with high speed and accuracy, irrespective of the resolution change. Testing of the algorithm in MATLAB demonstrates the supremacy of the proposed eye gaze tracking approach in controlling the mouse pointer.***

Keywords: eye gaze tracking, Viola-Jones algorithm, Kalman Filtering, eye gaze detection

1. Introduction

Eye gaze serves as a new modality, which aids human computer interactive in an increasingly intuitive and effective way. Controlling the mouse cursor using eye gaze seems to be a low-cost approach and it involves an eye tracker and a head tracker that are attached to a head mount. The eye tracker is based on the images that are recorded by a modified webcam to acquire the eye movements. These eye movements are then mapped to a computer screen to position a mouse cursor accordingly. Eye trackers have been extensively used in psychology, cognitive linguistics and product design. The study of eye movements began in the 1800s. In 1879, Louie ´ Emile Javal has observed that during a reading task, the eye movements do not involve smooth sweeping of eyes along the text as previously understood. Instead, it involves a series of rapid stop and goes motion patterns. Eye trackers differ in the degrees of freedom which they can track. Simple eye trackers report only the direction of the gaze relatively to the head (EOG and systems rigidly mounted on the head) or for a fixed position of the eyeball (systems which require a head fixation). Systems that are more sophisticated allow free head movements in front of a stationary system. Such systems do some kind of (implicit) head tracking. In addition, wearable eye trackers for use in 3D virtual worlds have to report the direction of the gaze in space and not only relatively to the head. Due to advancement in processing power of modern computers, the 1980s also signaled the genesis of eye tracking in human computer interaction. The application domain was primarily targeted towards physically challenged users. However, recently eye tracking has been extensively used to evaluate the design of interfaces. This provides a solid platform to test the ease of use of a computer interface and helps quantify their intuitiveness. Eye tracking technology has also been a useful tool to evaluate the utility of websites to communicate information effectively. It is also becoming popular in human computer interaction (HCI) where scan paths are utilized to build gaze contingent displays, also known as gaze-based interfaces. Online advertising is another field where eye tracking technology will have a significant impact in the near future. Whatever the application may be, the system for tracking the eyes should exhibit increased level of robustness, in addition to being non-intrusive and inexpensive. There are three different methods to track the motion of the eyes. The most direct method is the fixation of a sensor to the eye. The big advantage of such a method is the high accuracy and the nearly unlimited resolution in time. For this reason, medical and psychological research uses this method. The second method is electrooculography (EOG), where the sensors attached at the skin around the eyes measure an electric field. The big advantage of the method is its ability to detect of eye movements, even when the eye is closed. Both the methods, which are explained so far, are obtrusive and are not suited well for interaction by gaze. The third and the most preferred method for eye-gaze interaction is the video. The central part of this method is a video camera connected to a computer for real-time image processing. The image processing takes the pictures delivered from the camera and detects the eye as well as the pupil to calculate the gaze's direction. The big advantage of video-based eye tracking is the unobtrusiveness. Consequently, it is the method of choice for building eye-gaze interfaces for human-computer interaction.

2. Overview of Conventional Eye Tracking Systems

The first eye-tracking result we present reveals a relationship between eye movements and task-performance. The first observation we make about the scan paths is that in both cases, novice and expert, there is a noticeable organization in the patterns. (Users are not wandering their eyes all over the user interface with no recognizable pattern.) This is a result of the authors' providing the necessary stress on the visual attention system through a combination of dispersion in the UI and the dynamic nature of the task. In the absence of this foveal stress, we would have expected to see greater randomness in the pattern and accumulations of fixations unrelated to task performance. The second observation we make is that there are distinguishable differences in the scan paths associated with the level of expertise. Because this PILOT study was able to produce data that consisted of both (1) organized eye movements during the completion of a challenging task and (2) a good variation in task-performance scores from different users, we were successful in searching for relationships between these two variables (eye movement, task-performance). It is precisely this zone of eye movement pattern and performance results that an EPC verification experiment must generate if it is to be capable of detecting connections between task-performance and eye-tracking measures.

In the LINE study, a novel eye-tracking metric is constructed by discarding data (specifically, the authors retain only the vertical dimension of the eye-tracking data) with the goal of transforming the complex scan path metric into a simpler format. Importantly, not only did this technique provide an effective method for characterizing eye movement behavior, but also specific patterns emerged which were found to be associated with better task performance.

Further, we explore an eye-tracking study which had the original research goal of measuring the impact on task performance brought about by changes in the user interface and to attempt to relate these results to eye-tracking measures. As an eye-tracking study that explicitly focused on the relationships between user interface, task performance, and eye-tracking measures, the NEWS study was clearly of interest to this survey. Unfortunately, the NEWS study had deficiencies in its experimental design, which may explain the weak connections observed between the eye-tracking measures and user performance. These defects in the experimental design included both the simultaneous use of multiple information formats (e.g., audio and video) – which inadvertently reduced foveal stress – as well as poor control of the presentation of visual stimuli – which interfered with the collection of eye-tracking data. These shortcomings in the NEWS study were especially disappointing because its interface/task combination had strong similarities to our motivating example—so, if the experimental design had been better, we would have been very interested in the results. Nevertheless, we will use the negative aspects of this study as examples of what to avoid in any future EPC verification experiments that we might perform.

3. Proposed Methodology

The proposed eye tracking system depends on Viola-Jones algorithm and the various steps involved are Eye Gaze Detection, Eye Tracking, Key frame extraction and Cursor pointing, in order.

3.1 Eye Gaze Detection

The initial step of any eye tracking system is to detect the eye gaze in an exact way. This paper utilizes Viola Jones algorithm, which is followed by Kalman filtering, to perform eye gaze detection. Fig. 1 portrays the procedure for the proposed eye detection.

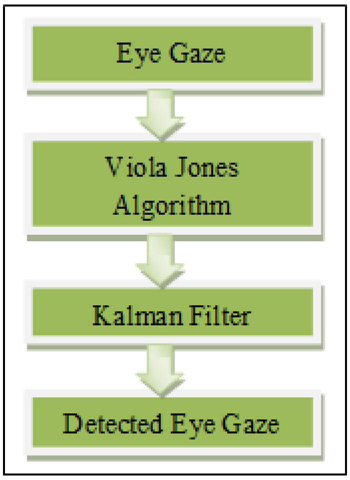


Figure 1: Eye Detection Process

The Viola-Jones algorithm involves four stages, namely, Haar Feature Selection, Creation of an Integral Image, Adaboost Training and Kalman filtering. Haar Feature Selection and Creation of an Integral Image: The Viola-Jones algorithm uses Haar-like features that are a scalar product between the image and some Haar-like templates. Let I and P denote a same-sized image and a pattern, respectively. Then, the feature associated with the pattern of eye gaze can be defined as:

$$\sum_{1\le i\le N}\sum_{1\le j\le N} I(i,j)\, I_{P(i,j)\,is\,white} - \sum_{1\le i\le N}\sum_{1\le j\le N} I(i,j)\, I_{P(i,j)\,is\,black} \qquad (1)$$

To compensate the effect of different lighting conditions, all the images should be mean and variance normalized in advance. The derived features are assumed to hold all the information needed to characterize a face. Since faces are large and regular by nature, the use of Haar-like patterns seems justified. However, another crucial element that lets this set of features take precedence is the integral image, which allows calculation at a very low computational cost. Instead of summing up all the pixels inside a rectangular window, this technique mirrors the use of cumulative distribution functions. The integral image II of I is expressed as:

$$II(i,j) = \begin{cases} \sum_{1\le s\le i}\sum_{1\le t\le j} I(s,t), & 1\le i\le N \text{ and } 1\le j\le N \\ 0, & otherwise \end{cases} \qquad (2)$$

Eq. (2) can also be stated as:

$$\sum_{N_1\le i\le N_2}\sum_{N_3\le j\le N_4} I(i,j) = II(N_2,N_4) - II(N_2,N_3-1) - II(N_2-1,N_4) + II(N_1-1,N_3-1) \qquad (3)$$

Eq. (3) holds for all $N_1 \le N_2$ and $N_3 \le N_4$. As a result, computing an image's rectangular local sum requires at most four elementary operations, when provided with its integral image. Moreover, the integral image itself can be obtained in linear time by setting $N_1 = N_2$ and $N_3 = N_4$ as in Eq. (4).

$$I(N_1,N_3) = II(N_1,N_3) - II(N_1,N_3-1) - II(N_1-1,N_3) + II(N_1-1,N_3-1) \qquad (4)$$

Feature Selection with Adaboost: The best features are selected using Adaboost learning. Given the probabilistic weights $w \in R_+$ assigned to a training set made up of n observation-label pairs (x_i, y_i), Adaboost aims to iteratively drive down an upper bound of the empirical loss under mild technical conditions.

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391

$$\sum_{i=1}^{n} w_i 1_{yi \neq f(x_i)} \quad (5)$$

The building block of the Viola-Jones face detector is a decision stump, or a depth one decision tree, parametrized by a feature $f \in \{1,....d\}$, a threshold, $t \in R$ and a toggle $T \in \{-1,1\}$. Given an observation $x \in R^d$, a decision stump h predicts its label using the following rule:

$$h(x) = (1_{\pi fx \geq t} - 1_{\pi fx < t})T = (1_{\pi fx \geq t} - 1_{\pi fx < t})1_{T=1} + (1_{\pi fx < t} - 1_{\pi fx \geq t})1_{T=-1} \in \{-1,1\} \quad (6)$$

Where, πfx is the feature vector of $f-th$ coordinate. By adjusting individual example weights, Adaboost makes more effort to learn harder examples and adds more decision stumps in the process. Intuitively, in the final voting, a stump ht with lower empirical loss is rewarded big, when a T-member committee (vote-based classifier) assigns an example according to the following equation.

$$f^T(.) = Sign[\sum_{t=1}^{T} \alpha_t h_t(.)] \quad (7)$$

For notational simplicity, we denote the empirical loss as:

$$\sum_{i=1}^{n} w_i(1) 1_{y_i \sum_{t=1}^{T} \alpha_t h_t(x_i) \leq 0} = P(f^T(X) \neq Y) \quad (8)$$

Where, (X,Y) is a random couple distributed according to the probability P defined by the weight, $w_i(1)$, $1 \leq i \leq n$ set when the training starts. As the empirical loss goes to zero with T, so do both false positive $P(f^T(X)=1|Y=-1)$ and false negative rates $P(f^T(X)=-1|Y=1)$, owing to

$P(f^T(X) \neq Y) = P(f^T(X)=1|Y=-1) + P(f^T(X)=-1|Y=1)$ must tend to 1 (9)

Thus, the size T of the trained committee depends on the targeted false positive and false negative rates.

Kalman filtering: Kalman filtering, also known as linear quadratic estimation (LQE), is an algorithm that works in a two-step process. In the prediction step, the Kalman filter produces estimates of the current state variables, along with their uncertainties. Once the outcome of the next measurement (necessarily corrupted with some amount of error, including random noise) is observed, these estimates are updated using a weighted average with more weight being given to estimates with higher certainty. The algorithm is recursive. It can run in real time, using only the present input measurements as well as the previously calculated state and its uncertainty matrix; no additional past information is required. In order to use the Kalman filter to estimate the internal state of a process, given only a sequence of noisy observations, one must model the process in accordance with the framework of the Kalman filter. This means specifying the following matrices: F_k, the state-transition model; H_k, the observation model; Q_k, the covariance of the process noise; R_k, the covariance of the observation noise; and sometimes $P(f^T(X) \neq Y)$, the control-input model at each time-step k, as described below. The Kalman filter model assumes that the true state at time k is evolved from the state at $(k-1)$ according to Eqn. (10).

$$x_k = F_k x_{k-1} + P(f^T(X) \neq Y) u_k + w_k \quad (10)$$

Where, F_k is the state transition model which is applied to the previous state x_{k-1}; B_k is the control-input model that is applied to the control vector u_k; w_k is the process noise that is assumed to be drawn from a zero mean multivariate normal distribution with covariance Q_k.

$$W_k \sim N(0, Q_k) \quad (11)$$

At time k, an observation (or measurement) z_k of the true state x_k is made according to $z_k = H_k x_k + v_k$. Here, H_k is the observation model that maps the true state space into the observed space and v_k is the observation noise that is assumed to be a zero mean Gaussian white noise with covariance.

3.2 Eye Tracking

Once the eye gaze is detected, the eye should be tracked to identify the movement of the eye. In the detected eye ball, two processes are applied and they are:

- Bounding Box
- Centroid calculation

In geometry, the minimum or smallest bounding or enclosing box for a point set (S) in N dimensions is the box with the smallest measure (area, volume, or hyper volume in higher dimensions) within which all the points lie. For any object in n-dimensional space, its centroid is the mean position of all the points in all of the coordinate directions. For the two coordinate system with (x_1, y_1) and (x_2, y_2), the centroid is calculated using the formula

$$C = \left(\frac{x_1 + x_2}{2}, \frac{x_1 + x_2}{2}\right) \quad (12)$$

3.3 Key Frame Extraction

All the key frames are extracted from the detected eye using the Euclidean distance. The distance between the coordinates of the two detected eye are estimated. For the two coordinate system with (x_1, y_1) and (x_2, y_2), the distance is calculated using the formula,

$$d = \sqrt{(x_1 - x_2)^2 - (y_1 - y_2)^2} \quad (13)$$

3.4. Cursor Pointing

A cursor is an indicator that is used to show the current position for user interaction on a computer monitor or other display device that will respond to input from a text input or pointing device. Once the key frame is extracted, the cursor is caused to move in the direction of the eye.

4. Result and Discussion

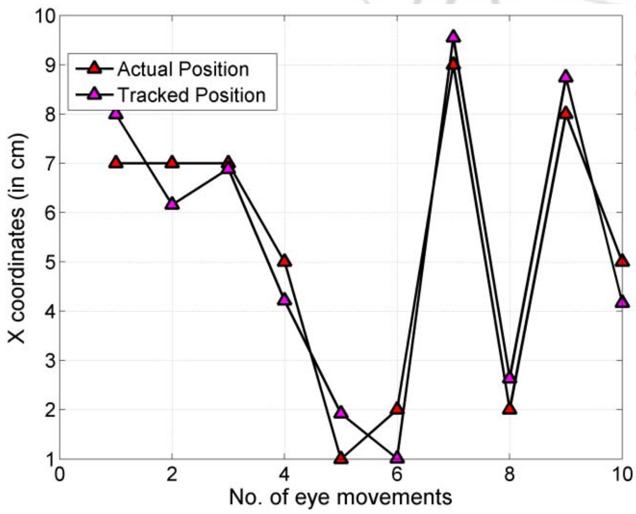
4.1. Experimental Setup

The procedure for Eye graze tracking is done in MATLAB R 2015a. The performance analysis of the proposed system in terms of accuracy and error is discussed in the next section.

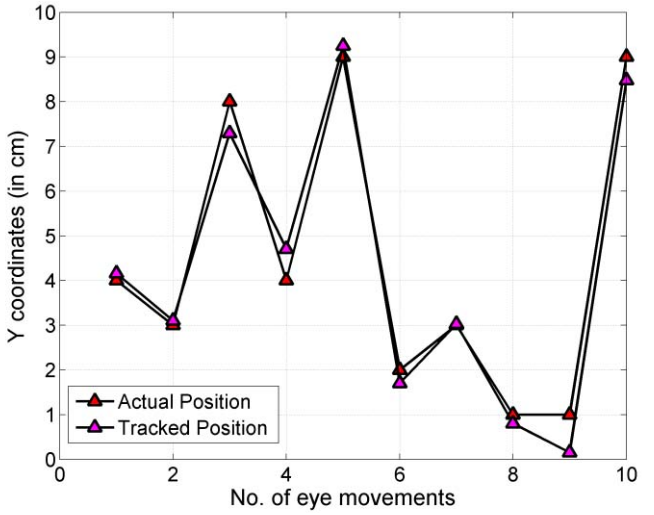
4.2. Performance Analysis

Efficiency of the algorithms: Though the identified tracking features are the ones that are available in the image without any spectral or other kind of processing, it is still hard to detect them from the user's image. The strip detection is a simple search for a known geometrical pattern and hence, requires a very little computational effort to reach the edge points. The detected edges, as discussed earlier, are the true bottom corners of the strip in the user's image. However, the detected eyeball centers may not be the exact centers at times. The algorithms are dependent on the distribution of gray levels, which in turn depend upon the ambient light variation in the field of view of the CCD camera. This variation can result in multiple light distribution patterns over users face resulting in a complex grabbed image in term of gray level variation. Non-uniform distribution of light causes shadows and partially darkened images. The lamp over the monitor does help to provide a uniform pattern, but some variations in light still exist due to the shadows of the surrounding objects. The deviation of the eye position in different coordinates is shown in Fig. 2.

In Fig. 2 (a), the actual position and the tracked position of the eye corresponding to X-coordinate is given. It is experimented for 10 movements of eye. In the first movement, the distance between the actual position and the tracked position is slightly high. In the third point, the same position is obtained by the tracked location. Till the 10 movements, the distance between the actual and the tracked position of eye is varied.



(a)



(b)

Figure 2: Deviation of eye position (a) X-coordinates (b) Y-coordinates

In Fig. 2 (b), it represents the actual and tracked position of eye movement corresponding to Y-coordinate. Here, it obtains almost same position in both actual and tracked one. Some points may get varied in both positions, till the 10 movements of eye. Computation of Mean and Standard Deviation: Table I portrays the computed mean and variance in the eye gaze tracking procedure. In the table, the mean error and standard deviation of the right pupil, right upper eyelid, right lower eyelid, left pupil, left upper eyelid and left lower eyelid was given. It completely computes the error occurred in this method.

Table 1: Mean and standard deviations in the computation of features

Sl. No	*Feature*	*Mean Error*	*Std. deviation*
1	Right pupil center X value	1.57	0.70
2	Right pupil center Y value	1.75	0.71
3	Right upper eyelid displacement	1.7	0.75
4	Right lower eyelid displacement	1.25	0.64
5	Left pupil center X value	1.32	0.58
6	Left pupil center Y value	1.67	0.70
7	Left upper eyelid displacement	1.60	0.68
8	Left lower eyelid displacement	1.30	0.59

5. Conclusion

In this paper we have presented Viola-Jones algorithm for improving the active contour method. We have shown increased performance by using the prior knowledge that the iris is darker than its surroundings. This prevents the algorithm from fitting to the sclera. Also, a novel approach to eye tracking that is based on Viola-Jones algorithm has been initialized by a simple heuristic. This has enabled the algorithm to overcome rapid eye movements. The active contour method handles these by broadening the state distribution and thus, recovering the fit in a few frames. Furthermore, the accuracy has been increased by fitting to the pupil rather than iris. This is particularly the case, when a part of the iris is occluded. It has been shown that the Viola-Jones algorithm is accurate, independent of the resolution and it is very fast for low resolution images. Further, we have demonstrated that it is feasible to estimate gaze from a single generic camera. This opens for a multitude of new applications in human computer interfaces. By the estimation of gaze dynamics, we may detect the emotional states too. Emotion synthesis is feasible through the appearance model, which can simulate faces with emotional expressions and given gaze.

References

[1] Ohno T, Mukawa N., and Yoshikawa A, ―Freegaze A Gaze Tracking System for Everyday Gaze Interaction", Proc. Eye Tracking Research Applications Symp., pp. 125-132, 2002

[2] Xiaohui Yang, Jiande Sun, Ju Liu, Jinyu Chu, Wei Liu, and Yuling Gao, ―A gaze tracking scheme for eye-based intelligent control", proceedings of the 8th World Congress on Intelligent Control and Automation, Jinan, China, 2010

[3] Beymer, D. and Flickner, M., ―Eye gaze tracking using an active stereo head", Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), Volume 2, pp. 451–458, 2003

[4] Brolly, X.L.C. and Mulligan, J.B, ―Implicit calibration of a remote gaze tracker", Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '04), Volume 8, pp. 134, 2004

[5] Morimoto C. H., Amir A, and Flickner M, ―Detecting eye position and gaze from a single camera and 2 light sources," in Proceedings of the International Conference on Pattern Recognition, 2002

[6] Hansen D. W., Agustin J. S. and Villanueva A, ―Homography Normalization for Robust Gaze Estimation in Uncalibrated setups", Proc. of the Symposium on Eye Tracking Research and Applications, pp. 13-20, 2010

[7] Iannizzotto L. and La Rosa F., ―Competitive combination of Multiple Eye detection and Tracking Techniques", IEEE Transactions on Industrial Electronics, 58(8): 3151-3159, 2011

[8] Cerrolaza, J. J., Villanueva, A., and Caveza, R, ―Study of polynomial mapping functions in videooculography eye trackers", ACM Trans. Comput.-Hum. Interact. 19, Article 10, 25, 2012

[9] Morimoto C. and Mimica M, ―Eye gaze tracking techniques for interactive applications", Computer Vision and Image Understanding, 98(1): pp. 4-24, 2005

[10] Baluja S. and Pomerleau D, ―Non-intrusive gaze tracking using artificial neural networks", In Proceedings of Advances in Neural Information Processing Systems, volume 6, pages 753–760, 1994

[11] Demjen E., Abosi V., and Tomori, ―Eye Tracking Using Artificial Neural Networks for Human Computer Interaction", Physiological Research, 2011

[12] Javier Orozco, F. Xavier Roca and Jordi Gonzalez, ―Real time gaze tracking with appearance based models", Machine Vision and Applications, 20(6), 353-364, 2009

[13] Lu F., Sugano Y., Takahiro O. and Sato Y., ―A head pose-free approach for appearance-based gaze estimation", in Proc. of the 22nd British Machine Vision Conf, 2011

[14] Sheela S.V.and Vijaya P.A. ―Mapping Functions in Gaze Tracking", International Journal of Computer Applications 26(3):36-42, 2011

[15] Rikert T.D. and Jones M.J. ―Gaze estimation using morphable models", IEEE International Conference on Automatic Face and Gesture Recognition, 436-441, 1998

[16] Yang X-H, Sun J-D, Liu J., Li X-C, Yang C-X and Liu W., ―A Remote Gaze Tracking System Using Gray-Distribution-Based Video Processing", Journal of Biomedical Engg: Applns., Basis & Communications, 24(3): pp.217-227, 2012

[17] Bartels, M. and S.P. Marshall, ―Eye tracking insights into cognitive modeling", in Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, ACM: San Diego, California, 2006

[18] Bednarik, R. and M. Tukiainen, ―Temporal eye-tracking data: evolution of debugging strategies with multiple representations", in Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, ACM: Savannah, Georgia. p. 99-102, 2008

[19] Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., Wickens, C., ―Comparison of expert and novice scan behaviors during VFR flight", in 11th International Symposium on Aviation Psychology. 2001.

[20] Uwano, H., et al., ―Analyzing individual performance of source code review using reviewers' eye movement", in Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, ACM: San Diego, California, 2006

[21] Josephson, S. and M.E. Holmes, ―Clutter or content?: how on-screen enhancements affect how TV viewers scan and what they learn", in Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, ACM: San Diego, California, 2006

Author Profile

Priyanka Jadhav received the B.E degree in Electronics & Telecommunication Engineering in 2014 from Shivaji University. Currently P.G. student at Annasaheb Dange College of Engineering and Technology (ADCET), Ashta.

Rupali Jagtap received the B.E. (E&TC) and M.E. from Shivaji University in 1998 and 2010 respectively. Currently working as Assistant Professor at Department Electronics & Telecommunication Engineering, ADCET , Ashta. I have teaching experience of 15 years.