International Journal of Engineering Sciences & Research Technology

(A Peer Reviewed Online Journal)
Impact Factor: 5.164

IJESRT

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

EYE GAZE PECULIARITIES DETECTION IN CHILDREN WITH AUTISM USING A HEAD-FREE CAM

Vijayaprabhuvel Rajavel, Goverdhenan B, Abishek V. Gomathinayagam
Electronics and Communication Engineering, Velammal Engineering College, Chennai, India

ABSTRACT

Eye tracking studies are useful to understand human behaviour and reactions to visual stimuli. To conduct experiments in natural environments it is common to use mobile or wearable eye trackers. To ensure these systems do not interfere with the natural behaviour of the subject during the experiment, they should be comfortable and be able to collect information about the subject's point of gaze for long periods of time. Most existing mobile eye gaze trackers are costly and complex. Furthermore, they partially obstruct the visual field of the subject by placing the eye camera directly in front of the eye. In this project, a real time monitoring system for social interaction between young children with Autism spectral disorder (ASD) using a non-head mounted camera has been proposed. Here, we deal with children with autism where when the tutor interact with the children either verbally or through facial expression, the children fail to pay attention to him/her due to lack of concentration. So in order to find where the child is paying attention, we adopt eye gaze detection technique with which we will be able to find where the child is seeing. Here we use circular Hough transform and vision cascade object detector system with which the accuracy of eye gaze detection has been improved. Once we have found out, it will be easy to make the child learn things as we can teach in his stream of interest by conducting this study for many a number of times, so as to confirm his area of interest.

KEYWORDS: Eye gaze detection, Autism, Hough transform, Vision Cascade Object detection, ASD

1. INTRODUCTION

Considerable research effort has been dedicated to exploring how well children with autistic spectrum disorders infer eye gaze direction from the face of an actor. Here we combine task performance (accuracy to correctly label a target item) and eye movement information ('where' the participant fixates when completing the task) to understand more about the components involved in completing eye direction detection tasks.

Autism is a neuro-developmental disorder characterized by qualitative impairments of social communication, accompanied by unusual repetitive or stereotyped behaviours. One of the earliest behavioural indicators of the disorder is a deficit in the development of joint visual attention; the ability to share attention with other people in a co-ordinated manner. Early deficits of joint attention compromise subsequent opportunities for the development of social cognition plausibly contributing to the impaired social skills of individuals with autism. Clifford and Dissanayake (2008) studied home videos of infants later diagnosed with autism, identifying poor quality and timing of eye contact even during the first year of life. Similarly, 2-year–old infants with autism show less frequent joint attention behaviours. Thus sharing attention and being able to identify the direction of another persons' gaze, is a core problem for individuals with autism. Understanding the underlying mechanisms of such deficits can be particularly valuable in the design of training and intervention programs. The current research emphasises the value of tracking eye movements for revealing components of task performance in domains known to be of difficulty.

Understanding the anatomy of the eye is essential to effectively detect the eye and its members. Three primary components of the eye relevant to tracking gaze include the cornea, pupil, and lens, as shown in figure 1.1. The

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

image enters the eye as light rays through the cornea and pupil. The inverted light image is projected onto the back of the eyeball, or retina, which contains two types of light-sensitive cells, or photoreceptors, rods and cones. Photoreceptors convert the light rays into neural signals that are transferred to the visual cortex of the brain via the optic nerve. Light rays are interpreted by the visual cortex of the brain, providing the sense of sight.

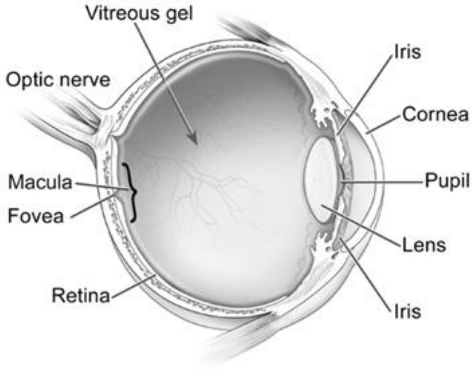

Figure 1.1 Cross Section of Eye (Side View) Showing Major Parts

Silvia et al. [1] reports on the study of gazes, conducted on children with pervasive developmental disorders (PDD), by using a novel head-mounted eye-tracking device called the WearCam. They are able to monitor naturalistic interactions between the children and adults. The study involved a group of 3 to 11 year-old children (n=13) with PDD compared to a group of typically developing (TD) children (n=13) between 2 and 6-years old. We found significant differences between the two groups, in terms of the proportion and the frequency of episodes of directly looking at faces during the whole set of experiments. They also conducted a differentiated analysis, in two social conditions, of the gaze patterns directed to an adult's face when the adult addressed the child either verbally or through facial expression of emotion. We observe that children with PDD show a marked tendency to look more at the face of the adult when she makes facial expressions rather than when she speaks. This paper employs a wearable head mounted camera which is not a suitable device for a child with ASD. Basilio et al. [2] system suffers from the uncompatability of a head mounted wearcam for a child with PDD conditions such as ASD. Bengoechea et al. [3] Active Appearance Model (AAM) detector needs a previous training stage can represent a drawback of this method compared to the one based on Harris detector stage

2. MATERIALS AND METHODS

Proposed Methodology

In the existing method, a head mounted device called Wear-cam is used to monitor the interactions. But a head mounted wear-cam is practically unsuitable for children with ASD due to their uncommon natural behaviour. Here Support Vector Regression algorithm has been implemented.

With further advances in computing technology, eye tracking has become far less intrusive with video-based eye tracking systems. Earlier video-based eye trackers required manual detection of the pupil in each frame, which took hours of analysis for even short video clips. Computer vision and image processing techniques allows this process to be automated which is not the case in existing method . However, depending on algorithm performance, some precise systems would also take hours to process video clips.

In the proposed system, we use vision cascade technique, putative matching and circular Hough transform in order to improve the accuracy of the eye gaze, which has not been implemented in the existing method where they use simple ROI technique and cropping. Also, no transform is used for the eye gaze detection whereas here we use circular hough transform in order to detect the iris of the eye which gives us a clear image of the eye direction ,so that we will be able to find where the child is seeing quite accurately. In existing method, ROI was used to detect the eye portion from the face, here we use vision cascade technique which detects the eye pair from the video to frame converted image using centroid. This helps to enhance the perfect detection and cropping of eye which improves the percentage of accuracy.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Putative matching technique is adopted on the tutor side which helps in object detection in a more precise manner. This helps to extract the inline and outline features of the object so that it can match exactly with the object which the child sees. Another great improvement is that canny edge detection is used here, which provides the gaussian filter, finding intensity gradient of the image, non-maximum suppression and threshold that provides accurate edge detection. All these algorithms help to enhance the accuracy of the project so that we can get a precise output.

Eye tracking algorithms, like other computer vision tasks, uses image processing techniques to analyse and understand images. Image processing techniques enhance images and extract useable data for the computer vision algorithms.

Module Description

There are two major modules – one on the child side and other at the tutor side. Let us now have a deep insight into both the modules and their operation.

Child side Module

A web camera is used on the child side to detect the image of the child. Using video to frame converter, the snapshot of the current frame is taken as reference. The image is captured and the eye pair is detected using circular Hough transform and circular peaks are used to detect the outline of the eye and vision cascade object detector. Next the gray-scale converter is used to convert the coloured image to black and white which will help to detect the circle quite easily. In canny edge detection threshold technology is used to detect the edge of the eye accurately. Dilating feature is used by which the noise is reduced . With the help of eye cart technology the eye portion is clearly detected and the direction of the eye is determined. With this we can say whether the child is seeing front, left, right or top. We need to conduct this for quite a number of times in order to ensure where the child is seeing. This helps to determine the child's line of interest with which he/she will be able to grasp things easily.

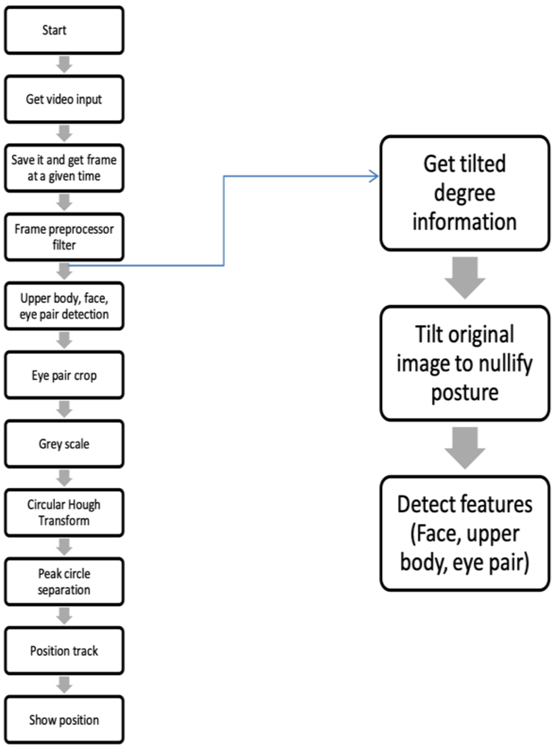

Figure 2.1 Child side module

Tutor side Module

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

Another web camera is used on tutor side to detect the position of the toy and tutor actions. The child side program is synced with the tutor side program in order to determine where the child is seeing. In the tutor side, the tutor is surrounded by four different toys spaced at a certain distance between each other. Each toy is detected separately from all four and their internal and external features are deduced with the help of putative matching method. Also montage is used to match the features of the given object to that of the original image which consists of all four objects. Hence the object that the child sees can be made out easily from the whole set of objects with which the child's area of interest can be determined. So when the tutor side program is synced with the child side the object the child sees will be the output. The position of the object will be changed in order to ensure which object the child sees in particular. So this helps us to make sure that the child sees the particular object alone always.

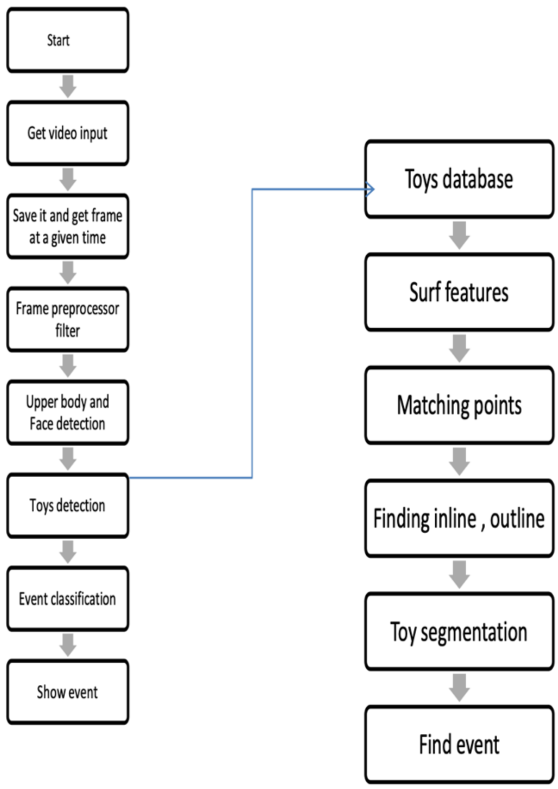

Figure 2.2 Tutor side module

Techniques and Algorithms

Vision Cascade Object Detector

The main characteristics of Viola–Jones algorithm which makes it a good detection algorithm are robust, real time and face detection with very high detection rate (true-positive rate) & very low false-positive rate always.
The goal is to distinguish faces from non-faces. The algorithm has mainly 4 stages: Haar Features Selection, Creating Integral Image, Ada boost Training algorithm and Cascaded Classifiers.

The features employed by the detection framework universally involve the sums of image pixels within rectangular areas. As such, they bear some resemblance to Haar basis functions, which have been used previously in the realm of image-based object detection. However, since the features used by Viola and Jones all rely on more than one rectangular area, they are generally more complex. The figure at right illustrates the four different types of features used in the framework. The value of any given feature is always simply the sum of the pixels within clear rect angles subtracted from the sum of the pixels within shaded rectangles. As is to be expected, rectangular features of this sort are rather primitive when compared to alternatives like steerable filters. Although they are sensitive to vertical and horizontal features, their feedback is considerably coarser. The cascade object detector uses the Viola-Jones algorithm to detect people's faces, noses, eyes, mouth, or upper body.

Circular Hough Transform

The child sits in front of the web camera. The camera is attached to monitor employs specially designed software to read the movement of child's eye so that it will help the tutor to know where the child is actually seeing. Under this detection method we use a unique algorithm called circular hough transform which helps to determine the position of the pupil. This algorithm is quite effective an accurate when compared to the existing method as it clearly detects the eye region from the rest of the image.

Hough transform is a standard image analysis tool for finding curves that can be defined in a parametrical form such as lines and circles. The circular Hough transform can be employed to deduce the radius and centre coordinates of the pupil's region. Hough transform can be applied to detect the presence of a circular shape in a given image. It is used to detect any shape or to locate the pupil in the eye.

Canny Edge Detection

Canny edge detection algorithm is one of the most strictly defined methods that provide good and reliable detection. Owing to its optimality to meet with the three criteria for edge detection and the simplicity of process for implementation, it becomes one of the most popular algorithms for edge detection.
The Process of Canny edge detection algorithm can be broken down to 4 different steps:

1. Apply Gaussian filter to smooth the image in order to remove the noise
2. Find the intensity gradients of the image
3. Apply non-maximum suppression to get rid of spurious response to edge detection
4. Apply threshold to determine potential edges

Segmentation

After preprocessing the next step is segmentation. Segmentation means, separate the objects from the background. The aim of image segmentation algorithms is to partition the image in to perceptually similar regions. Every segmentation algorithm addresses two problems, the criteria for a good partition and the method for achieving efficient partitioning. In the literature survey it has been discussed various segmentation techniques that are relevant to object tracking. They are, Mean shift clustering, and image segmentation using Graph-cuts (Normalized cuts) and Active contours.

Foreground segmentation is the process of dividing a scene into two classes; one is foregrounding another one is background. The background is the region such as roads, buildings and furniture. While the background is fixed, its appearance can be expected to change over time due to factors such as changing weather or lighting conditions.

The foreground any element of the scene that is moving or expected to move and some foreground elements may actually be stationary for long periods of time such as parked cars, which may be stationary for hours at a time. It is also possible that some elements of background may actually move, such as trees moving in a breeze.
The main approaches to locating foreground objects within in the surveillance system is Background modelling or subtraction-incoming pixels compare to a background model to determine if they are foreground or background.

Object Detection

Algorithms, such as object detection, use the extract data to shape an understanding of the image. Fitting a group of features for an object together forms the model of the object. Features are extracted from the image and fitted to the model to detect the object. Viola-Jones object detection, a popular object detection method, describes the model of the object using Haar-like features.

Haar-like features were adapted from Haar wavelets to create a real-time face detector. However, the robustness of this algorithm allows it to be trained for a variety of objects. A Haar-like feature consists of a rectangle of specific pixel size with 2 to4 rectangular region. These features are placed over the objects to be calculated. The pixels values are added in the light region and subtracted in the dark region to calculate a feature value. The features for the object will vary in size and position resulting in a large quantity of features available to choose.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

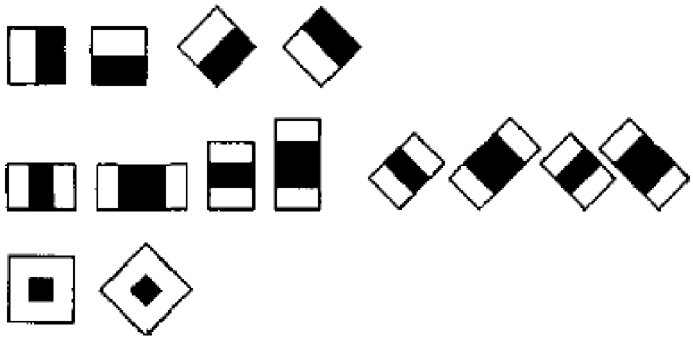

Figure 2.3 Haar-like Features

Putative Matching

Putative matches of the feature points in both images are computed by using a correlation measure for points in one image with features in the other image. Only features within a small window are considered to limit computation time. Mutually best matches are retained.

If **x** is a point in one image and **x'** a point in another image, then **x'Fx** = 0 ,where **F** is fundamental matrix. To compute **F** completely we begin by using edge detector to find interest points in an image.

A minimal sample is selected from the putative matches from which a tentative **F** is calculated. The process is iterated until a sufficient number of samples have been taken. The **F** which fits the matches becomes the input for the next step.

Once an initial **F** has been computed, more matches can be found by searching along epipolar lines. A non-linear minimization is used to fit an **F** to a large number of points.

Object Tracking

The aim of an object tracker is to generate the trajectory of an object over time by locating its position in every frame of the video. But tracking has two definition one is in literally it is locating a moving object or multiple object over a period of time using a camera. Another one in technically tracking is the problem of estimating the trajectory or path of an object in the image plane as it moves around a scene. The tasks of detecting the object and establishing a correspondence between the object instances across frames can either be performed separately or jointly. In the first case, possible object region in every frame is obtained by means of an object detection algorithm, and then the tracker corresponds objects across frames. In the latter case, the object region and correspondence is jointly estimated by iteratively updating object location and region information obtained from previous frames. There are different methods of tracking.

Point is tracking- Tracking can be formulated as the correspondence of detecting objects represented by points across frames. Point tracking can be divided into two broad categories, i.e. Deterministic approach and Statistical approach. Objects detected in consecutive frames are represented by points, and the association of the points is based on the previous object state which can include object position and motion.

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

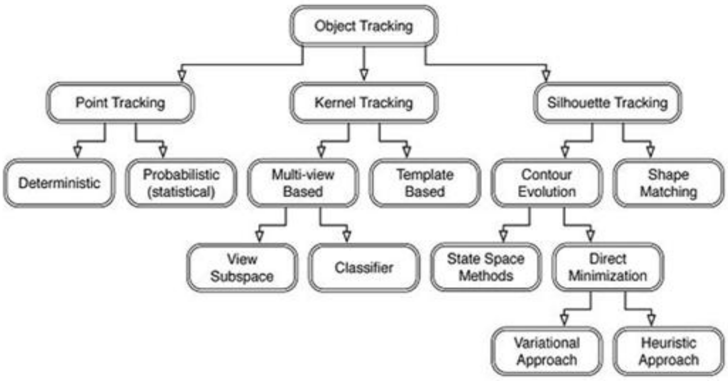

Figure 2.4 Different tracking categories

It plays a vital role to select a proper feature in tracking. So feature selection is closely related to the object representation. For example, color is used as a feature for histogram based appearance representations, while for contour-based representation, object edges are usually used as features. Generally, many tracking algorithms use a combination of these features. The details of common visual features are as follows:

- Color: Color of an object is influenced by two factors. They are Spectral power distribution of the illuminant and Surface reflectance properties of the object. Different color models are RGB, L*u*v and L*a*b used to represent color.
- Edges: Edge detection is used to identify strong changes in image intensities generated by object boundary. Edges are less sensitive to illumination changes compared to color features. Most popular edge detection approach is Canny Edge detector.
- Optical Flow: It is a dense field of displacement vector which defines the translation of each pixel in a region. It is computed using the brightness constraint, which assumes brightness constancy of corresponding pixels in consecutive frames. Optical Flow is commonly used as a feature in motion based segmentation and tracking application.
- Texture: Texture is a measure of the intensity variation of a surface which quantifies properties such as smoothness and regularity. It requires a processing step to generate the descriptors. There are various texture descriptors: Gray-Level Co-occurrence Matrices, loss texture measures, wavelets, and steerable pyramids.

Mostly features are chosen manually by the user depending on the application. The problem of automatic feature selection has received significant attention in the pattern recognition community. Automatic feature selection methods can be divided into, Filter Methods and Wrapper Methods. Filter methods try to select the features based on general criteria, whereas Wrapper methods selects the features based on the usefulness of the features in a specific problem domain.

3. RESULTS AND DISCUSSION

The results and discussion may be combined into a common section or obtainable separately. They may also be The experiment was conducted on five children and the eye gaze peculiarities were detected under a normal teaching environment setup of ASD children. A moderate amount of light is necessary to provide a quality image to process gaze estimation. The distance between the monitor and the child was kept to be around 30 cm. The corresponding iterations of x and y centre and the position of the eye are calibrated as F- Front, L – Left and R – Right

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

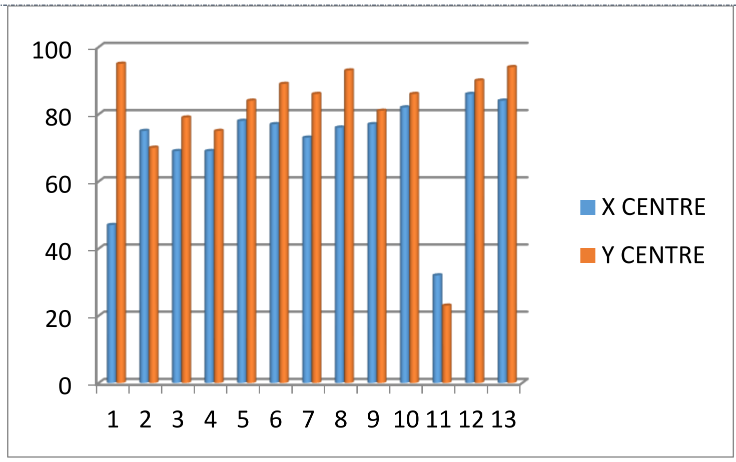

Fig 3.1- Plot of X Centre AND Y Centre

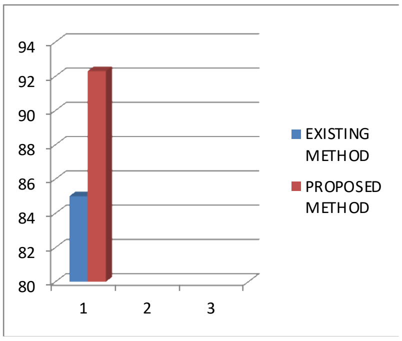

Fig 3.2- Accuracy of both methods

Table 3.1 Comparison table for accuracy of both methods

	Existing method	Proposed Method
Accuracy	85%	92.3%

4. CONCLUSION

In this project, we detect the eye position of the child using circular hough transform which enhances the accuracy of the output. Thus, the position of the child's eye can be easily determined. Also we can take the snapshot of the child when he sits straight as we implement a video kind of coverage. Canny edge detection and putative matching are used in order to improve the precision of the output. The canny edge detection helps in detecting the edges of the eye so that the particular eye region alone can be cropped from the rest of the image. Putative matching is used to read the inline and outline features of a particular object. This helps in matching the object which the child sees with the original background using montage feature.

The vision cascade method uses Viola Jones algorithm to track the object. The result of collection of many images shows that the efficiency is 92.3% compared to previous existing methods. In future work, the usage of accelerometer will help to detect the child's face even when his head is tilted up to 45 degrees . When a higher resolution camera used, the efficiency can be improved.

REFERENCES

1. Silvia Magrelli, Basilio Noris, Patrick Jermann, Franc¸ois Ansermet, Franc¸ois Hentsch, Jacqueline Nadel,and Aude G. Billard, "A Wearable Camera Detects Gaze Peculiarities during Social Interactions in Young Children with Pervasive Developmental Disorders" in IEEE 2013
2. Basilio Noris, Jean-Baptiste Keller and Aude Billard, "A Wearable Gaze Tracking System for Children in Unconstrained Environments"
3. Jose Javier Bengoechea, Arantxa Villanueva, Rafael Cabeza, "Evaluation of accurate eye corner detection methods for gaze estimation"
4. Aun Kei Hong , "A Novel approach to a wearable eye tracker using region-based gaze estimation"
5. Stephen Bialkowski, Zeb Johnson, Kathryn Morgan, Xiaojun Qi, and Donald H. Cooley, "A Non-Intrusive Approach to Gaze Estimation"
6. H.R. Chennamma, Xiaohui Yuan, "A Survey on eye gaze tracking techniques"
7. Arantxa Villanueva, Juan J. Cerrolaza and Rafael Cabeza, "Geometry Issues of Gaze Estimation"
8. Radoslaw Mantiuk, Michal Kowalik, Adam Nowosielski, and Bartosz Bazyluk, "Do-It-Yourself Eye Tracker: Low-Cost Pupil-Based Eye Tracker for Computer Graphics Applications"
9. Onur Ferhat, "Eye-Tracking with Webcam-Based Setups: Implementation of a Real-Time System and an Analysis of Factors Affecting Performance"

ISSN: 2277-9655
Impact Factor: 5.164
CODEN: IJESS7

10. De Luna, Paolo & Mustafar, M. Faiz & Rainer, Gregor, "A MATLAB-based eye tracking control system using non-invasive helmet head restraint in the macaque" in Journal of neuroscience methods. 235. 10.1016/j.jneumeth.2014.05.033.