Low Cost Eye Tracking Technique For Visual Scan Path Detection in the diagnosis of Special Learning Disorders in Children

Sumathi.V, M.Monicasubashini, Himabindu. J
School of Electrical Engineering, VIT University, Vellore-632014, Tamilnadu, India

Abstract— **Eye movements are the key to human thought process. Tracing these movements, with or without the knowledge of the user, as an input can be leveraged to several intelligent and intuitive design interfaces. An attempt is made to develop a computationally efficient and cost effective application for visual scan path analysis by real time eye tracking using a low resolution USB web camera. Eye ball movement can be followed by obtaining the image of the eye using the camera, mounted on the system, interfaced with LabVIEW. The glint- reflection of the light source formed on the pupil in the eye ball acts as the ROI (Region of Interest) and can be located in the real time images of the eye. The glint is tracked in the incoming video stream and monitored to trace the scan path. The LabVIEW VI gives the coordinates of the glint which are stored in the database. As a test application, a scenario is developed for diagnosis of special learning disorders in children.**

Keywords-ROI; Glint; Tracking; Special Learning Disorders (SpLDs)

I. INTRODUCTION

Eye movement tracking has been only used for developing applications for physically handicapped people. With the new technologies and modernization, eye movement tracking is being explored in a broader perspective for integrating with daily applications. The detection of visual scan with or without user's knowledge opens new possibilities in the field of automation and intuitive application development. As progress toward the goal, a web camera-based computer interface makes communication possible by detecting the eye-gaze direction and storing the tracked path coordinates in the database.

Existing methods for eye detection and tracking, as well as gaze analysis that do not rely on infrared lighting typically depend on high-resolution images of eyes, often obtained by pan/tilt/zoom cameras. The system proposed here is designed around an inexpensive visible-light "webcam," and it works with images that have a lower resolution than the images used in previous approaches [1], [2]. These approaches require high-resolution images of the eyes for appearance- or feature based methods that locate, for example, eye corners, eyelids, or contours of the iris. The use of webcam also allows the system to be made available to people at significantly less expense than those requiring specialized hardware. The features of the eye affect the task of tracking eye movements, determining the position of the eye and inferring the point or object regarded. In building a system for this task, it is necessary to simplify and use workable models of the eye. It is important to judge these simplifications and the performance of a system for a given application, and regard every measured characteristic separately.

A. Eye Terminology:

To describe the eye, it is necessary to introduce common terms to refer different parts of its anatomy. Figure 1 depicts a cross-section and a frontal view of the object of interest. The sensing part of the eye is the retina. It lines the back of the eye and contains photoreceptors that relay information about incident light to the brain. A particular area of the retina is the fovea. It is about 2mm in diameter and is the most sensitive area that provides most acute vision. From the anterior (front) of the eye, incident light first passes through the cornea, the first refracting surface on the way to the retina. In the cross-section, it is visible as a slight bulge with a greater curvature than the rest of the eye. The light then passes through the pupil, a circular hole whose size can change to regulate the intensity of incident light. The area surrounding the pupil is the iris. The white area that makes up the most of the visible surface of the eye ball is the sclera. In the front view, a thin area has been marked that connects the cornea to the sclera, the limbus. The "glint" is the reflection of the stationary light source from the cornea.

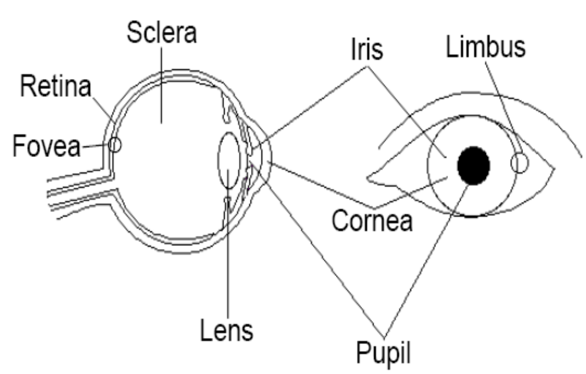


Fig 1 Eye Anatomy [6]

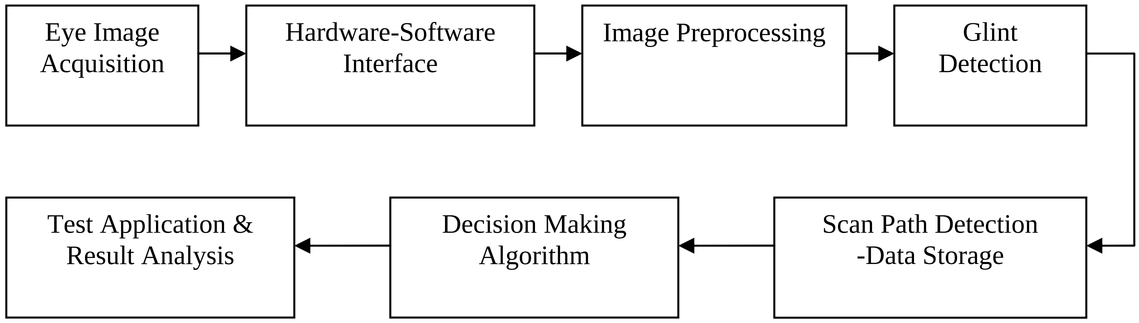


Fig.2: Block Diagram

B. Viewing:

When we are reading, viewing the symbols on a computer screen or regard any other scene, our eyes are constantly moving. Similarly, our eyes move when we are fixating on a point while we are moving our head. The purpose of this movement is to position the image of the object we are recording on the fovea.

II. EYE TRACKING TECHNIQUES BACKGROUND

Recordings of eye movements have been made as early as 1936 [4], but the technique used for the recordings only improved gradually. In particular since the availability of video cameras, the non-intrusiveness of tracking systems has substantially improved. Since the beginning, a great variety of systems have been developed. According to [5] there are four broad categories to classify measurement methodologies for eye movements.

A. Electro-oculography (EOG):

It involves measurement of electrical potential at the skin around the eye which makes the technique a lot more expensive and complex.

B. Scleral contact lens/Search coil methods:

It involves measurement of the position of a special contact lens inserted into the user's eye. The measurement can be mechanical, electromagnetic or visual. This technique can only be performed under the supervision of an expert and can have side effects.

C. Photo-oculography(POG):

It involves photo or video image analysis for the measurement of distinguishable features of the eye under rotation/translation. The efficiency and sensitivity of this process is very low in comparison to the other processes.

D. Video-based pupil-corneal reflection technique:

It involves video image analysis to detect two points of reference, e.g., pupil centre and the corneal reflection of a stationary light source. The eye position is determined based on the difference between these points. As the difference remains relatively constant with "minor head movements" but changes with eye movements, the eye position can be measured relative to the source or a computer screen, respectively. The position within the screen is called ROI (Region of Interest). Most current eye tracking systems are less invasive, by relying only on video images. The majority of these systems belong to category four. The main advantage in comparison to systems in category three is the increased convenience while measuring the ROI.

The corneal and pupil reflection relationship technique uses the reflection of stationary light source from the cornea, the “glint”, and the retro-reflectivity of the eye. The latter causes the retina, which is visible through the pupil, to appear as a bright disk in a captured video image. The relative position of the pupil in respect to the highlight is used for measuring eye movements. As already explained this technique allows good usability due to free head movements. An additional problem is caused by the eye movements themselves. If the glint is not reflected from the cornea but the surrounding eye, the mathematical computations to infer the eye position are complicated. The previously mentioned glint is only the first reflection that a ray of light causes on its way into the eye. Further reflections are caused by the internal structure of the eye.

III. METHOD IMPLEMENTED

The various stages involved in the process are shown in Figure 2.

A. Pre processing of eye image

The purpose of pre processing is to minimize the number of on-pixels in the binary image on which Hough transform is computed. An ideal image for Hough transform will only contain a circle representing the outline of the ‘glint’, however it is difficult to obtain and an attempt is made to remove unwanted pixels without removing the eyeball pixels. The pre processing steps are:

1. Color filtering:

Three channel RGB image is converted to a single channel image for edge detection. It has been found experimentally [3] that green channel image provides best results for edge detection even though difference with other channels is not huge. This is done using IMAQ ExtractSingleColorPlane.vi and setting the color priority as green.

2. Histogram equalization:

The single channel image is equalized to compensate for different lighting conditions. Since eye image contains two regions of comparable areas and distinct intensities, one representing the eyeball and the other white area around it, equalization tends to highlight the edges of the eyeballs.

3. Edge detection:

The green channel image is then subjected to edge detection to get an outline of the eyeball. Since the eyeball pixels have different intensity than the white pixels around it an edge is readily obtained however it is occluded. A 3 by 3 vertical Prewitt operator, Figure 3, is used for this purpose. Horizontal Prewitt operator is also tried but since the outline of the eyes form horizontal edges, extra edges appear in the edge image. Vertical operator returns the vertical edges as all the pixels are not required to obtain the circle outlining the glint. The IMAQ EdgeDetection.vi helps to implement the Prewitt Filter.

−1	−1	−1
0	0	0
1	1	1

−1	0	1
−1	0	1
−1	0	1

Fig 3.Horizontal and Vertical Prewitt Operator

4. Particle Removal:

Particle Analysis is used to locate the “glint” in the eye and eliminates particles that are of no interest. All this is obtained by thresholding using IMAQ Threshold.vi. At first, the image is converted to binary by a proper threshold value. It is important to find the proper threshold value in order to separate the eyes, nose and eye brows from the face. The threshold Td is used to guarantee a minimal spatial difference between the minimum and maximum values when motion is detected. This helps prevent motion that is not an intentional eye movement from being detected. There are many methods to find the threshold value. A manual threshold method is employed for this purpose. The threshold interval is defined by two parameters i.e. the Lower Threshold (>=) and Upper Threshold (<=).

The IMAQ RemoveParticle.vi is included to eliminate or keep particles resistant to a specified number of 3 × 3 erosions. The particles that are kept are exactly of the same shape as those found in the original source image. The connectivity 4/8 specifies the algorithm for particle detection. The connectivity mode directly determines whether an adjacent pixel belongs to the same particle or a different particle. We choose connectivity-4 so that pixels adjacent in the horizontal and vertical directions are considered neighbors to increase the efficiency. The pixel frame is treated as hexagonal during the transformation.

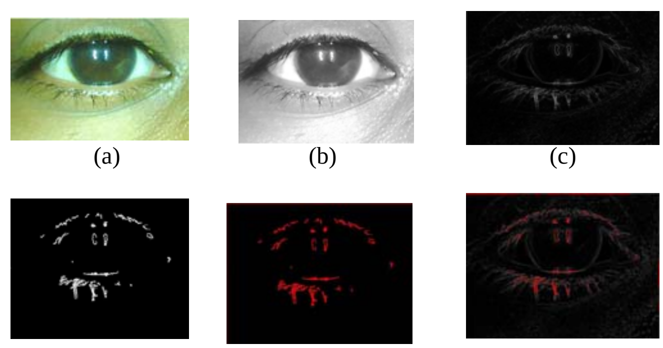
Finally the number of erosions specified is a 3 × 3 erosions applied to the image. The default is 2 but it is set to 1. It reduces the size of an object along its boundary and eliminates isolated points in the image. The erosions eliminate pixels isolated in the background and erode the contour of particles according to the template defined by the structuring element. A high pass filter emphasizes significant variations of the light intensity usually found at the boundary of objects. High pass frequency filter helps to isolate abruptly varying patterns that correspond to sharp edges, details, and noise in the acquired image of the eye.

5. Circular Hough Transform:

Hough transform can be used to extract various features from an image. The transform takes the geometric equation of the feature to be found and then inverts it so that x and y become constants in the equation, thus resulting in an inverse function space. The intersections in this space reveals the constants associated with the equation of the features discovered in the original image. Circular Hough transform is used to find the coordinates of circles in an image. In order to determine the location of the “glint” in the processed image, a range for the radius of the desired circle is specified. Each threshold pixel in the binary image is examined to determine whether it fits the desired circle. The best circle gives the outline of the “glint”. The technique is very powerful since it can fit circles even when a partial outline of a circle is available which is common in an eye image due to the occluded eyeball. Figure 3 shows the extraction of “glint” from eye image using circular Hough transform. The image has been equalized and vertical Prewitt operator has been used to get the approximate outline of the eyeball.

B. Simulated Result

The final image is obtained using the IMAQ CastImage.vi which converts the current image type to Grayscale (SGL) with 32 bits resolution.



(a) (b) (c)

(d) (e) (f)

Fig 4.Image (a) Original ,.After (b) Color Filtering (c) Edge Detection (d) Threshold (e) Particle Removal (f) Cast Image

IV. Real Time Implementation

Hough transform can be used to locate “glint” in each eye image thus determining its’ location. The ROI could be lost due to a blink or a rapid movement of the eyeball or eye itself. The problem can be addressed by constantly monitoring the tracked ROI to determine if the tracking has been lost, in which case Hough transform is used to re initialize the “glint” location. The best feature point for the “glint” is the center of the eyeball and its two outer edges. The relative distance among these three points should remain constant as the eyeball moves within the eye.

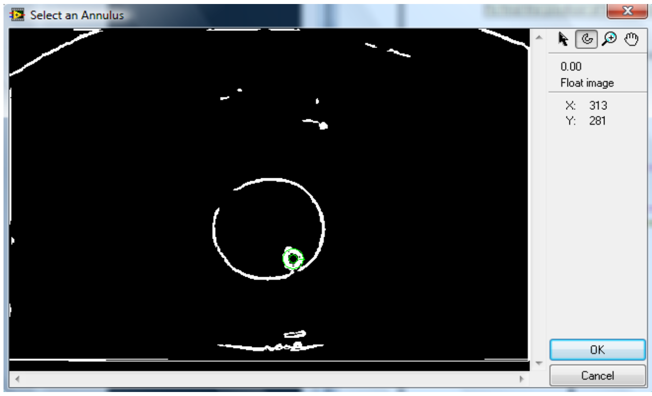


Fig 5 ROI selection with Annulus Tool

After obtaining the processed image of the eye we select the Region of Interest (ROI) – "glint" to help with further processing. This application uses an annulus, Figure 5, to select the ROI as it is independent of the shape chosen. The ROI when selected is converted into a template which has to be searched during the matching phase. The template image contains the data that defines the color template pattern for the matching stage. This data is appended to the input color template image. During the color pattern matching step, the color template descriptor is extracted from the color template image and used to search for the template in the color match image.

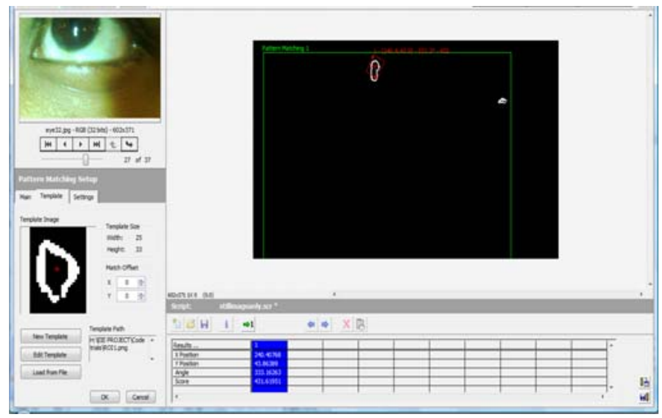
The real time images of the eye, being input at high speed, or the tracking video is processed using the same algorithm mentioned above. After processing the input video, it gives the feature point or the "glint" and helps in continuously monitoring it. The IMAQ Find Pattern 2.vi helps to implement the pattern matching technique in which the input template image having "glint" as the ROI is compared with the real time video of the eye. The rotation invariant technique is to look for the template pattern in the image. The minimum match score, the smallest score so that a match must have to be considered valid can be varied depending on how accurate the user wants the system to be. We get the tracked "glint" specified with a red mark in the output image along with the position where it is found. The position is in terms of the coordinates with respect to a reference plane. These coordinates can be viewed in a graph. The graph shows the dynamic movement of the "glint" as the eye moves. The coordinates of the glint in each frame streamed in, is stored in a database for further analysis.

The data is used to generate the visual scan path by retrieving the stored data. The scan path is formed on the screen with a set of connected lines. Starting from the origin, each line is drawn from the current coordinate position to the incoming coordinate value. This generates the scan pattern. As eye movement leads to the sequence of coordinates, scan pattern is the replication of human visual scan process of any particular scene.

V. Results

The eye tracking in the streaming video input from the webcam is done successfully. The first image snap shot is taken for the template formation for the particular user and environment conditions. The template formed is used for pattern matching to track the glint being streamed in at the rate of 30 frames per second. The coordinate position of the glint is traced on to monitor resolution, Figure 6. A graph layout is drawn to see the movement of point of concern. The cursor moves smoothly with the eye movement when it is interfaced with the mouse.

Fig 6 Eye tracking and Coordinate Recognition



The coordinates of the glint tracked are recorded for every frame in the streaming video. The data is stored in an excel sheet, as shown in Figure 7, for future application specific analysis. The coordinates represent the gaze points of the user at that particular instant. By tracing the coordinates, the visual scan path can be obtained. The tracking speed and accuracy depends on the camera specifications and grabbing speed. The movement replication on the screen is found to be accurate. Head movement brings in change of coordinates. It's a highly sensitive system. Small variations would not hamper the final result.

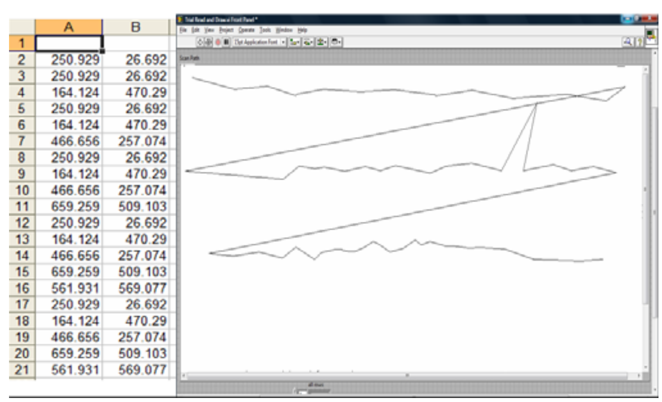
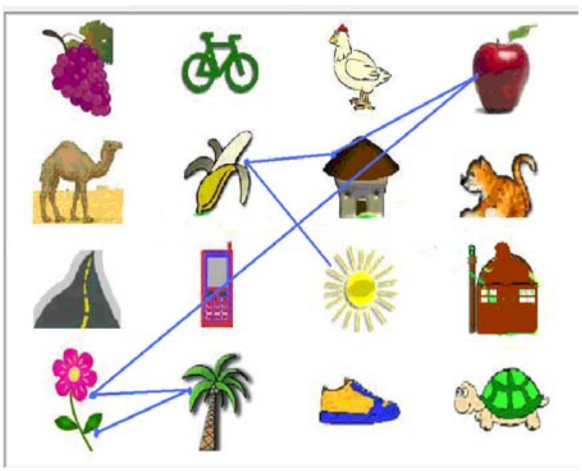


Fig 7 Coordinate Storage and Visual Scan Path

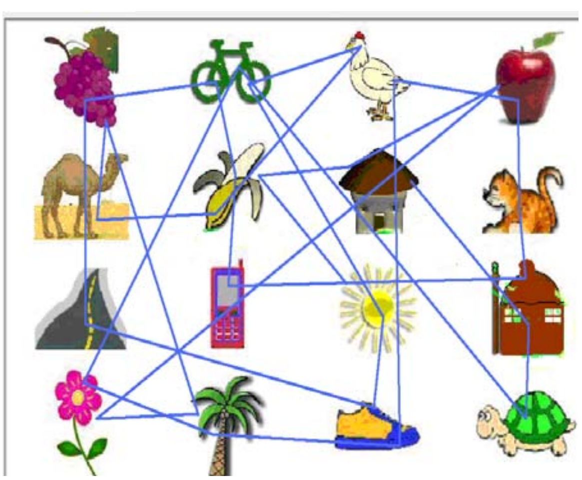
VI. Game for Diagnosis of Special Learning Disorders in Children

Designing interactive learning systems for users with special needs is a challenge to system developers that have an inadequate understanding of the life experiences of people with disabilities. Dyslexia and Attention Deficit Disorders (ADD) are developmental disorders of neurobiological origin. Research has shown that individuals with these SpLDs have more problems in working memory than people without SpLDs and has linked these memory deficits to difficulties in acquiring literacy skills and navigating interactive systems. While many interactive activities designed for such a group, very few have been dwelled for children. It's difficult to find out such a problem in children. Hence with such a game, people can find out if their child is suffering with some problems or not.

This inspired the idea of an interactive game for memory strategies. This game is aimed at teaching children memory strategies and developing their abilities in applying these strategies in learning contexts with identification of the problems faced by children.



(a)



(b)

Fig 8 Visual Search Pattern for a flower of a person (a) without and (b) with Special Learning Disorder

The aim of the eye tracking study is to explore the thought process of a child while searching the object in the grid. A game is designed for searching for specific objects in which we track their eye gaze as they examine a 4x4 grid of elements that includes the object with 15 distracting elements. The eye tracking procedure explained previously has been used to record eye movements of children as they interact with the game. Retrospective Think-Aloud protocol (RTA) is used to explore difficulties in recognitions observed by the child with SpLDs.

Here, as a test prototype, a person is shown the game to search an object. For example, the gazeplots in Figure 8, show a scattered search for people who do not know the object, possibly a child with SpLDs. In contrast, people who know the object, like children without SpLDs, exhibited focused and directed searches. The visual scan path in the first case would be random as the person doesn't know how the item looks like. In the second case, scan path is organized and focused towards the target object. Thus with the help of scan path data, such special disorders can be found out through gaming without any discomfort to children. It is of enormous help, if the child's problem is detected in early stages of life and the possible corrections could be carried out.

VII. CONCLUSION

The visual scan path is successfully generated for different people. The eye movement is tracked without the knowledge of the user making it more users friendly and robust. The usage of webcam makes it an integral modality in daily life. The system is cost effective as the software implementation can be done on any platform without any special hardware or software. The results may not be strong enough to suffice the hypothesis regarding the applications. But an extensive test with a more realistic group of individuals suffering from the disorders would help in deciding upon the accuracy in the system predictions.

REFERENCES

[1] M. Betke and J. Kawai, "Gaze detection via self-organizing gray-scale units," in Proc. Int. Workshop Recog., Anal., Tracking Faces Gestures *Real-Time Syst.*, Kerkyra, Greece, Sep. 1999, pp. 70–76.

[2] Z.-H. Zhou and X. Geng, "Projection functions for eye detection," Pattern Recognition., vol. 37, no. 5, pp. 1049–1056, May 2004.

[3] Yassir Nawaz and Shiladitya Sircar" Real Time Eye Tracking and Blink Detection" Memorial University, St. John's, NF, CANADA A1B 3X5 unpublished.

[4] Arne John Glenstrup and Theo Engell- Nielsen. Eye controlled media: Present and futurestate. http://www.diku.dk/»panic/eyegaze/article.html [Last Accessed Aug. 12, 2005].

[5] Andrew T. Duchowski. *Eye Tracking Methodology. Theory and Practice*.Springer-Verlag, 2003.

[6]EnchantedLearning. http://www.enchantedlearning.com/subjects/anatomy/eye/label/labeleye.shtml

[7] Canny, 1986 a computational approaches to edge detection. IEEE Trans. PAMI, 8:769 - 698.

[8] "IMAQ Vision Concepts Manual" by National Instruments.