Creating Human Perspective Visual Memory By Eye Tracking

Chanchal R. Agrawal
Computer Engineering
Atharva College of Engineering
Mumbai, India

Suraj Y. Jagtap
Computer Engineering
Atharva College of Engineering
Mumbai, India

Prof. Deepali Maste
Computer Engineering
Atharva College of Engineering
Mumbai, India

Abstract- Visual memory and especially eye tracking systems have been a major subject of research. Creating a human perspective visual memory by eye tracking will help establish a more natural and improved human-computer interface(HCL). A compact head-mounted device is made for the same purpose. Eye tracking system detects pupil and generate focus coordinates using which human perspective vision is created. This artificial human perspective vision is recorded and store forming a visual memory. This system is a real-time system. With its good speed, great accuracy and compact system it results in a great overall product improving human and machine interface.

Keys words- Pupil Detection, Visual Memory, Head-Mounted Device, Human Perspective Vision, Eye Tracking, Human-Computer Interface

I. INTRODUCTION

Demand for more realistic human-computer interface have been a great focus from two to three decades. Rapid evolution in technology make it easier to create compact, simple and non-intrusive eye tracking systems. Eyes being most important facial organ, the more we can create technology using it the more better interface we will have between human and machine. Creating a real time tracking system have been a great achievement in field of technology. The use of eye tracking have a great possibility to improve the quality of everyday human-computer interfaces.

Eye though being a small organ, can communicate with brain at a speed of 10 million bps which is equivalent to Ethernet connection[1]. Many people understood the significance of this organ and have know that creating an artificial device that will work similar to human eye will bring a drastic change in how machine and computer interact.

Creating a system that not only can track one's point-of-gaze in real-time but it also creates a human perspective vision and creating a virtual memory is the main objective of device. Many studies have been done regarding eye tracking techniques & devices and many research are still ongoing to create as natural HCL as we can. This paper is divided into parts so as to our concept more understandable.

(i.)Hardware & its components, to understand how the working of the going. (ii.)Eye tracking. (iii.)Understanding how human perspective works & what method is to follow to create such artificial human perspective visual memory. (iv.)Comparison with existing systems.

II. HARDWARE OF DEVICE

After decades of research and experiments it was now possible to track one's point of gaze. With rapid evolution in eye tracking techniques and devices, it became imperative to make a simple, compact, accurate and non-intrusive device for eye tracking. Several researches where conducted to obtain a head-mounted eye tracking device so it can fulfil all the above criteria and simultaneously it made sure that head movement would not affect the calibration or coordinates of eye tracker. Many researchers done study on it like John M. Franchak[2], Wolfgang Fuhl[3] and Jayson Turner[4] has very well explained the idea of head-mounted eye tracking devices.

With latest innovation we have acquired a head mounted device which is more like a goggle with two cameras, one at the bottom corner of the glass(i.e. eyecam) and other at top centre of the goggle(i.e. worldcam), as shown in Fig.2 The best part of this device is that its hand free and egocentric(means cameras are place at person's perspective level so they have same vision area as person's eyes have).

Working of hardware is explained in following steps:

1. Eye-cam constantly track pupil position to find one's point-of-gaze in real-time. Once position of pupil is tracked it is sent to controller
2. World-cam continue to record surrounding in normal way and send the recording to controller where it gets stored.
3. Controller takes coordinates produce by eye-cam and apply it on the recording sent by world-cam to create artificial human-perspective. Procedure to create is explained in later part().

Controller consist of raspberry pi 3 and other components for befitting connections and efficient working. This controller is where the process of manipulating of film is done in a way that a normal 2D image seems to be human perspective vision. So we can say that Raspberry pi 3 in the controller is brain of the product where all the codes and instructions are stored, both eye-cam and world-cam are the inputs and the video finally recorder with artificial human perspective is the output.

Fig.1. Raspberry pi 3

Fig.2. Pupil Headset

The most interesting and compact head-mounted device we found is a product by Pupil labs. It is called as Pupil-Headset. It is really a great product but the cost is too high for an average person to afford it but we can create similar product at a comparative low cost.

III.   EYE TRACKING

Eye tracking is a process of tracking the movement of the eye and determining where the user is looking[5]. Basically eye tracking find one's point-of-gaze or measures eye motion relative to movement of head. There are two main categories of eye tracking devices, one is remote eye tracking system(also called as desktop or stationary eye tracking) where the eye tracking camera is fixed at a external position in front of the person and other is head-mounted device. Head-mounted is very popular among eye tracking device companies like Pupil Labs[6], Tobii Glasses, SMI Glasses and many more.

Currently there are many algorithms on eye tracking have came into existence from past few decades. Famous among them are ExCuSe[7], Starburst[8], Swirski[9], Pupil Labs[6], SET[10].

In our project, we are using ExCuSe Algorithm for eye tracking. Exclusive Curve Selector, i.e. ExCuSe provides robustness in images, high detection rate and it is best for real world eye-tracking applications. There are two distinct imaging processes used in eye tracking, visible and infrared spectrum imaging. ExCuSe can be used in both imaging processes by applying some given steps. Whether image is a coloured or grayscale, same procedure is followed in both the cases
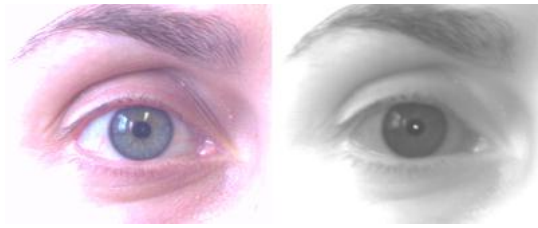

Fig.3. Coloured to grayscale image

In the first step input image under goes normalization(normalization is a process that changes the range of pixel intensity values) and then histogram value of image is calculated(An image histogram is a histogram that acts as a graphical representation of the tonal distribution in a digital image). Only if a peak is found in bright histogram then by doing edge analysis pupil can be found.

For edge analysis, first input image is changed into grayscale image. There are three ways to convert colour image into grayscale image[11]:
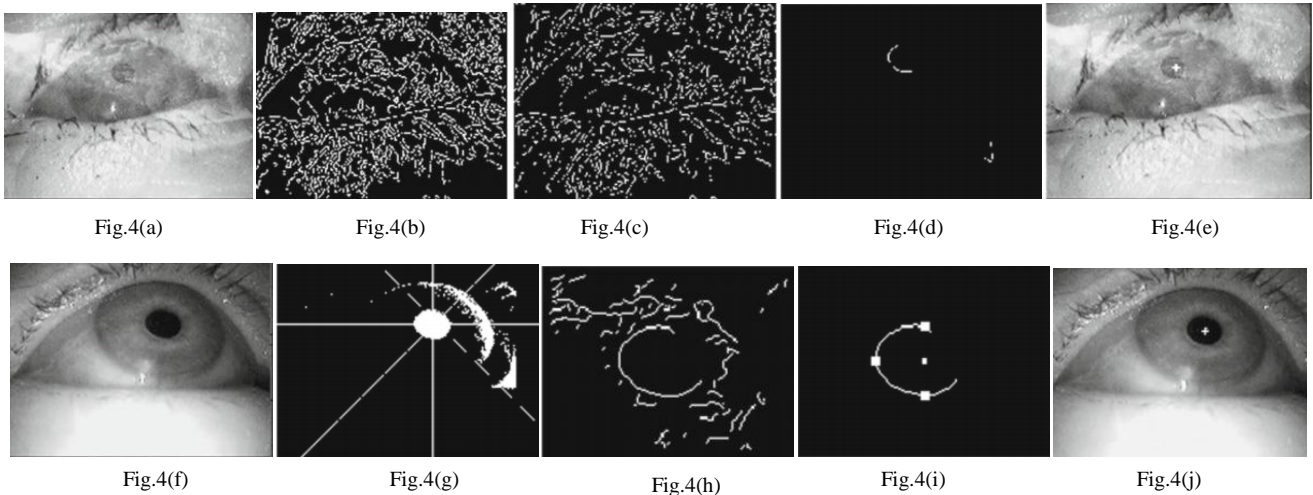
1. The lightness method: Here average of the most prominent and least prominent colours are taken:
$$(max(R, G, B) + min(R, G, B)) / 2 \qquad (1)$$
2. The average method : It simply takes average of all the values:
$$(R + G + B) / 3 \qquad (2)$$
3. The luminosity method:
$$0.21\,R + 0.72\,G + 0.07\,B \qquad (3)$$

Colour image to grayscale image conversion example is shown in Fig.3. This grayscale image undergoes a process called detection technique. Edge detection process is to find the region where there is a sharp change in the intensity. The binary image is passed through edge detection technique in which only if the edges of high intensity are found in the image Most used edge detection technique is Canny edge detection.

Canny Edge detection algorithm[13] is broken down to five different steps:

1. Apply Gaussian filter so as to smooth the image(to remove the noise)
2. Find the intensity gradients of the image. Standard Sobel edge detector is used here.
3. Apply non-maximum suppression to clear false responses.
4. Apply double threshold to find potential edges
5. Track edge by hysteresis: Finalize the edges by suppressing all the other edges that are weak and not connected to strong edges.

In edge analysis, first a grayscale undergoes smoothing where noise is reduced. There are different algorithms to perform smoothing such as Additive smoothing, Box Filter, Kernal smoothing, Gaussian filter and many more

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICIATE - 2017 Conference Proceedings

| Fig.4(a) | Fig.4(b) | Fig.4(c) | Fig.4(d) | Fig.4(e) |

| Fig.4(f) | Fig.4(g) | Fig.4(h) | Fig.4(i) | Fig.4(j) |

ExCuSe algorithm workflow is shown in Fig.4. (**a**) is image of eye with noise and corneal reflection.(**b**) Canny edge filter image.(**c**) Refined using Morphological operators. (**d**) The best edge is fitted by an ellipse and its centre is reported as the pupil centre (**e**).Input image without reflections (**f**). Input image without corneal reflection (**g**)Coarse pupil based on the AIP. (**h**) Canny edge image applied and Pupil position refined. (**i**)Rays are send out of the optimized position to select edges representing the pupil border (white dots on line are ray hits). (**j**) Result of the pupil centre estimation

We are applying Gaussian filter:

$$H_{ij} = \frac{1}{2\pi\sigma^2}\left(-\frac{(i-(k+1))^2+(j-(k+1))^2}{2\sigma^2}\right) \qquad (4)$$

where, $1 \le i, j \le (2k + 1)$

After smoothing, Sobel Edge detection is done to find gradient intensity. Very common operator for doing this is Sobel operator. Sobel operator separates edges in x and y direction.

Sobel Convolution kernels:

Gx:

-1	0	1
-2	0	2
-1	0	1

Gy:

-1	-2	-1
0	0	0
1	2	1

The gradient magnitude is given by:

$$|G| = \sqrt{Gx^2 + Gy^2} \qquad (5)$$

The angle of orientation of the edge (relative to the pixel grid) given by:

$$\theta = \arctan(Gy/Gx) \qquad (6)$$

Non-maximum suppression: It is a technique used for edge thinning. Non-maximum suppression indicates location with the sharpest change of intensity value by suppressing all the gradient values to 0 except the local maximal.

Thresholding is simple process of image segmentation. Thresholding helps to convert grayscale image to binary image by replacing pixels of intensity less than constant T with black pixel and pixels of intensity greater than constant T with white pixels.

Result generated from Canny edge detection technique will give under morphologic operation to remove thin & thick lines to refine the resulting image. Morphological operator are used to perform error correction if any made while following any above process. Basic morphological operations are dilation and erosion. After this mean position of all the connected line is calculate and based on that all straight line are discarded and what is left behind are curved lines. The enclosed mean intensity value is calculated for these curves and the curve with minimum value is consider as pupil curve.

If no bright peak is detected in intensity histogram then ExCuSe determines a coarse pupil position. Then algorithm refines coarse pupil stepwise to find centre. Coarse pupil is estimated on basis of Angular Integral Projection Function (AIPF)[13] on thresholded image. Input image is thresholded and pixels over thresholds are summed up along the row. This summations is done four times by rotating with an angel of 45 degree. Here assumption is made that all the pixels that belong to pupil are surrounded by equally bright or brighter pixels. Once pupil region is found edge detection technique is used and then centre is found mathematically. This centre is final result for eye-cam. The area of vision where human can see detailed image is called focus, whereas the blurred part is peripheral vision and it only works in low light vision and motion detection.

IV.   HUMAN VISION

To create an artificial human perspective vision it is essential for one to understand how human eyes perceive world. This concept has been very well explained in whitepaper of Tobii company[14] Human eye works in a very similar way to camera; light gets reflected from an object and enter into eyes through lens. Lens then

concentrate this light on light-sensitive surface (in case of eyes is called retina) and creates an image.

Main difference between a human vision and image captured by a camera is that camera images are plain

Fig.5. Focus and peripheral vision

images where every point is considered same but in case of image created by human vision is actually divided into three parts.

Human eyes cannot perceive entire scene together. Eyes can see details clearly in a certain limited area of vision and vision after that limited area is gradationally blurs out.

To create a human perspective vision, these two parts of has to be created. Here eye tracking techniques is main here to create this effect. Eye tracking technique finds where a person is looking at a instance. The point where a person is concentrating on is point-of-focus. If the visible area is assumed to be a 2D grid then eye tracking technique helps to find exact coordinates of the point-of-focus in terms of x and y axis. Let us take Fig.6a. as visible area.

In Fig.6b, centre of red circle(i.e. eye of penguin) is the point where a person is focusing at that instance so that point is point-of-focus. A circle of x radius is created with point-of-focus as its centre. This circle is the area where everything is clearly visible. Now as we go away from this circle, image gradually starts blurring. As showed in Fig.6c, as we are going away with certain distance(say y distance) vision starts blurring by some certain value(as shown in Fig., image blurs by 5 or 10%). This burred part is peripheral vision. We have artificially created human perspective vision in Fig.6d.

If we apply this effect on entire recording by keeping the coordinated found by eye tracker as centre and blur image gradually as we move from centre to border of visible portion. One thing to consider that maximum value to blur peripheral should not be more than 75% as human peripheral vision is never completely blurred.
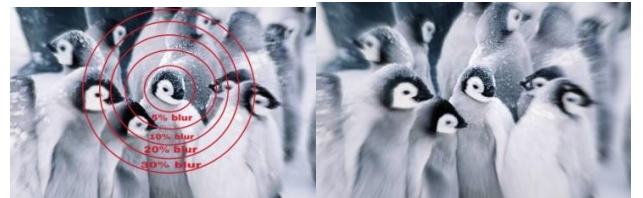

Fig.6(a)

Fig.6(b)

Fig.6(c)

Fig.6(d)

IV. COMPARNG WITH EXISTENCING SYSTEM

Comparing product with Tobii Glasses and SMI eye tracker: Tobii glasses are head-mounted eye tracking device which aims to use eye tracking for learning behaviour of people in different situation. It trackers one's point-of-gaze and records the video but does not create an artificial human-perspective. Similarly even SMI eye tracker, tracks and record a plain video but not as human perspective video. These two products though does not provide human perspective but they do are connected to internet for better usage and accessibility.

V. CONCLUTION

The main purpose of creating an artificial human perspective is to make Human-computer interface so nature that when someone view the recording it should feel as if they are re-living those moments making product more preferable. With further research and development we can connect this device to internet just like Tobii Eye glasses and SMI Glasses which will make this product not only to just track pupil position or to just record human perspective vision but can also be used in face reorganization, by adding microphone we can use it for use call services and search data.

VI. ACKNOWLEDGMENT

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICIATE - 2017 Conference Proceedings

VII. REFERENCE

[1] Koch K, McLean J, Segev R, Freed MA, Berry MJ, Balasubramanian V, Sterling P. How much the eye tells the brain. Current Biology. 2006 Jul 25;16(14):1428-34.

[2] Franchak JM, Kretch KS, Soska KC, Adolph KE. Head-mounted eye tracking: A new method to describe infant looking. Child development. 2011 Nov 1;82(6):1738-50.

[3] Fuhl, W., Tonsen, M., Bulling, A. et al. Machine Vision and Applications (2016) 27: 1275. doi:10.1007/s00138-016-0776-

[4] Turner J, Bulling A, Gellersen H. Extending the visual field of a head-mounted eye tracker for pervasive eye-based interaction. InProceedings of the Symposium on Eye Tracking Research and Applications 2012 Mar 28 (pp. 269-272). ACM.

[5] Hari Singh, Dr. Jaswinder Singh, " Human Eye Tracking and Related Issues," International Journal of Scientific and Research Publications, Volume 2, Issue 9, September 2012 1 ISSN 2250-3153

[6] Kassner, M., Patera, W., Bulling, A.: Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp), pp. 1151–1160 (2014). doi:10.1145/2638728.2641695

[7] Fuhl, Wolfgang, et al. "Excuse: Robust pupil detection in real-world scenarios." International Conference on Computer Analysis of Images and Patterns. Springer International Publishing, 2015.

[8] Li D, Winfield D, Parkhurst DJ. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops 2005 Jun 25 (pp. 79-79). IEEE.

[9] Świrski L, Bulling A, Dodgson N. Robust real-time pupil tracking in highly off-axis images. InProceedings of the Symposium on Eye Tracking Research and Applications 2012 Mar 28 (pp. 173-176). ACM.

[10] Javadi AH, Hakimi Z, Barati M, Walsh V, Tcheang L. Set: a pupil detection method using sinusoidal approximation. Frontiers in neuroengineering. 2015;8.

[11] http://www.johndcook.com/blog/2009/08/24/algorithms-convert-color-grayscale

[12] http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm

[13] Mohammed GJ, Hong BR, Jarjes AA. Accurate pupil features extraction based on new projection function. Computing and Informatics. 2012 Jan 26;29(4):663-80.

[14] Tobii Eye Tracking: An introduction to eye tracking and Tobii Eye Trackers. January 27, 2010