

An Appearance based Method for Eye Gaze Tracking

¹S.V. Sheela and ²P.A. Vijaya
¹BMS College of Engineering, Bangalore,
²Malnad College of Engineering, Hassan, India

Abstract: Problem statement: Gaze estimation systems compute the direction of eye gaze based on observed eye movements. The need for gaze-contingent applications is the basis of the current research work. The gaze pointing systems is a substitute for the existing input devices. **Approach:** The gaze tracking methods are either feature based or appearance based. In this study, an appearance based approach for gaze tracking is proposed based on Run Length Coding (RLC). The experiment was conducted considering transitional changes and the class-intervals in iris pixels. The image acquisition begins from the center of the screen in anticlockwise direction. The center of the screen was the pivot point. **Results:** Using RLC, the recognition rate of 95% was achieved. The image analysis in different directions determines the gaze point. The directions was determined with respect to the pivot point. **Conclusion:** The proposed system provides a robust, less computational gaze tracking method using web camera.

Key words: Region growing, run length coding, gaze point, transitional changes, class-intervals

INTRODUCTION

Eyes provide reliable and prominent features for communication using gaze enabled interfaces. The gaze tracking system captures intention of a person on the screen. The gaze point determines the direction of where an individual is looking at as shown in Fig. 1. Eye movements are categorized into fixations and saccades. A fixation occurs when focus of eye is on a particular object. The movements of our eyes from one fixation to another are known as saccades. As the human eye scans over the scene or image, the focus shifts about 25 times per second, to take in the disparate parts in its field of vision. The movements and information combine to form a cohesive vision of the scene. Analysis of the fixations and saccades are important for visual behavior. The major challenges are due to illumination, variability in position, faster saccades and eye blinks (Hansen and Ji, 2010). Existing eye gaze tracking systems are confined to controlled environments. The usability of the system under natural environments needs improvement (Zhu and Ji, 2004). Some methods require strict calibration procedure prior to gaze tracking. The accuracy of gaze tracking system depends on size of eye's visual field, range of eyeball rotation, diameter of the fovea and radius of the eyeball.

The existing gaze tracking techniques are broadly classified into intrusive and non-intrusive. The intrusive techniques require attachments around the eye to

determine the gaze. These include search coils, electro-oculography, contact lens and head mounted devices. Non-intrusive techniques use video cameras under infrared or natural light sources. The non-intrusive or video based techniques are classified into Appearance based and Feature based techniques. Appearance based techniques use the image contents as to map directly to the screen coordinates (Hansen and Ji, 2010). These methods require several significant calibration points to infer the gaze direction from the images. The analysis of the images at calibration points is important for gaze estimation. Explicit camera calibration is eliminated. In a morphable model developed by Rikert and Jones (1998), the texture for a set of prototype images is mapped to the reference image based on shape transformation. Neural network is used for training the prototypes with parameters for shape and texture of the eye region. Betke and Kawai (1999) determine gaze direction by gray scale unit images. During calibration, gray scale units are created in an elliptic pattern to form model images. Learning process use self-organizing map. The comparison of the pupil positions in model image with various regions in the trial image using correlation coefficient determines the gaze direction. The method by Tan *et al.* (2002) use appearance manifold for gaze estimation. A set of sample eye images with varying parameters and pose represent a continuous set of points called appearance manifold in the high dimensional space.

Corresponding Author: S.V. Sheela, BMS College of Engineering, Bangalore, India

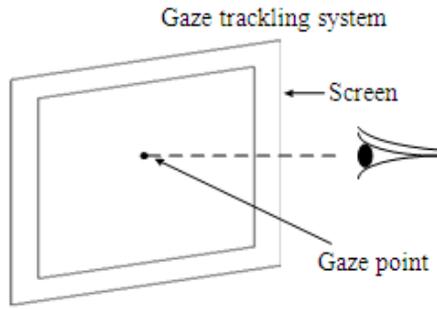


Fig. 1: Gaze tracking system

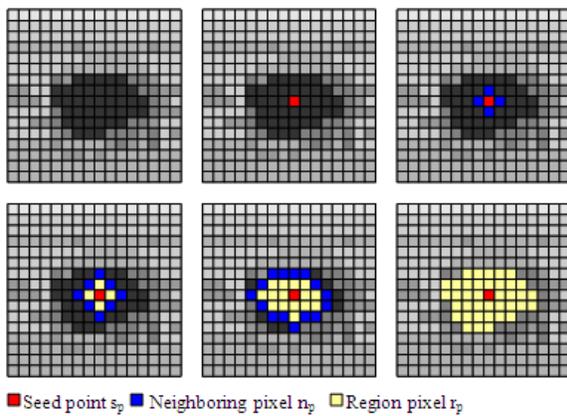


Fig. 2: Region growing

For a test image, the set of closest manifold points is determined by interpolation based on least squares criteria. In the method by Hansen *et al.* (2002), the shape and pixel intensity information of eye corners and pupil position is obtained by active appearance model. The mapping function is based on Gaussian process interpolation method considering mean value. Yoo *et al.* (2002) determined the gaze point on the screen based on the glint positions. The cross ratios in the image are mapped to the monitor screen to obtain the coordinates of the gaze point. The appearance models are used for tracking smaller eye movements compared to the size of the object. In feature based methods, the gaze estimation requires as prerequisite, a set of features. The pupil corneal reflection or the pupil-glint vector is the most common feature used in feature based techniques (Morimoto and Mimica, 2005; Zhu *et al.*, 2006; Baluja and Pomerleau, 1993). The local gaze features include pupil and limbus position, iris center, eye corner, inner eye boundary and sclera region. The global gaze features are face skin color interpupillary distance, ratio between average intensity, shapes, sizes of both the pupil and orientation of pupil ellipse with respect to face pose (Zhu and Ji, 2004;

Khosravi and Safabakhsha, 2008). The gaze mapping functions determine the screen coordinates. The mapping is an analytical function of either linear or second order polynomial. In the method by Kim and Ramakrishna (1999), the displacement of the iris center is based on linear approximation. The method by Zhu and Yang use eye corner-iris center vector as input for gaze angle calculation (Zhu and Yang, 2002). Interpolation is used to determine the gaze direction. Some nonlinear mapping functions use neural network, support vector machines and radial basis functions. The gaze detection system by Kiat and Ranganath (2004) use two Radial Basis Function Neural Network (RBFNN) to determine x and y coordinates of the gaze point on the screen. The pupil and glint parameters are used to train the RBFNN. In the method by Zhu and Ji (2004), six input parameters, the pupil glint displacement along x and y direction, ratio of major to minor axes of the ellipse that fits the pupil, pupil ellipse orientation and glint image coordinates are used for training the generalized regression neural networks. An extension of the work was developed using Support Vector Regression (SVR) to determine the gaze coordinates (Zhu *et al.*, 2006). In this study, an appearance based model is presented for eye gaze tracking. The features are the pupil and iris pixels in the captured eye images. The gaze direction is estimated based on RLC. The sequence of transitional changes and class intervals remain unique for each gaze direction.

Prologue:

Region growing: The entire image region is represented by S. Region based segmentation partitions S into n subregions, S_1, S_2, \dots, S_n , such that (a) The regions cover the whole image, $\bigcup_{i=1}^n S_i = S$ (b) S is a connected region, $i=1,2,\dots,n$. (c) The two regions are disjoint, $S_i \cap S_j = \phi, \forall i \text{ and } j, i \neq j$ where ϕ is the null set. (d) A region satisfies the logical predicate $\text{Pr}(S_i)$ defined over the points in set S. (e) $\text{Pr}(S_i \cup S_j) = \text{FALSE}$ for any adjacent region S_i and S_j .

Region growing groups pixels or subregions into larger regions based on predefined criteria. The pixel aggregation starts with a set of seed points Abas, (2010). The seed mark each of the objects to be segmented. Regions are iteratively grown by appending to each seed points those unallocated neighboring pixels that have similar properties. The region growing is shown in Fig. 2.

The distance between a pixel's intensity value and the region's mean, dist is used as a measure of similarity. The pixel with the smallest difference measured this way is allocated to the respective region.

This process continues until all pixels are allocated to a region. It stops when the intensity distance between region mean and new pixel become larger than a certain threshold value. This value is the region's Maximum Intensity Distance (MID). The algorithm for region growing is given in Algorithm 1.

Algorithm 1: Region Growing

```

Let S be the image region with intensity values I.
The seed point,  $s_p$  and MID values are initialized.
The q neighbor locations are determined and the
neighboring pixel  $n_p$  is added to list L. For instance
if q = 4, 4 neighboring pixels are added to L.
The mean value of the region,  $m(S)$  is computed.
The distance between pixel's intensity value and
region's mean is computed,  $dist=I(S)-m(S)$ .
For each pixel,  $n_p$  in L
    if  $n_p \notin R$  and  $dist < MID$ 
         $n_p$  is added to L
         $n_p$  is assigned as region pixel  $r_p$ ,  $r_p \in R$ 
        The new mean value of the region
        is calculated
    end
end
    
```

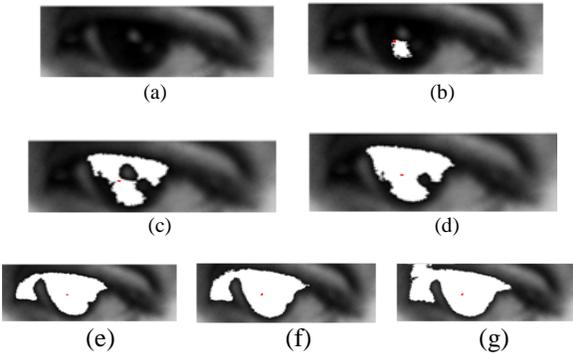


Fig. 3: (a) Eye image; Region growing for (b) MID = 0.01 (c) MID = 0.02 (d) MID = 0.03 (e) MID = 0.04 (f) MID = 0.05 (g) MID = 0.06

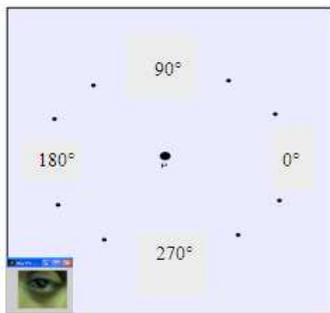


Fig. 4: Screen displaying directions

The segmentation results are dependent on the choice of seeds. Seed point selection is based on some user criterion like pixels in a certain gray-level range, texture, color and shape. The initial region begins as the exact location of these seeds. The MID determine the condition for region growing. The region growing for MID values ranging from 0.01-0.06 with seed point is shown in Fig. 3. It is observed that the optimum value of MID for region growing needs to be selected on trial and error basis. Even a small change in the value does not provide the complete region.

Run length coding: RLC provides compact representation of a binary image. The sequence of repeated intensity values is represented as a single value and count. The representation is useful for images which contains runs of data. The long sequence of same value is replaced by a two values. The intensity values $v_1, v_2, v_3, \dots, v_n$ are mapped to pairs $(a_1, r_1), (a_2, r_2), \dots, (a_n, r_n)$ where a_i represent image intensity and r_i represent runs of pixels (Gonzalez and Woods, 2005). The algorithm for RLC is given in Algorithm 2.

Algorithm 2: Run Length Coding

```

Let seq be the sequence of intensity values of a binary
image.
    k=1;
    for i=1 to length(seq)
        if (seq(i)==1 and seq(i+1)==0) or
            (seq(i)==0 and seq(i+1)==1)
            a(k) ← seq(i)
            d(k) ← i, k=k+1
        end
    end
    r(1)=d(1)
    for j=2 to length(d)
        r(j)=d(j)-d(j-1)
    end
    
```

For the sequence of binary image $seq = \{1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\}$ the following values are obtained. $a = \{1\ 0\ 1\ 0\}$, $d = \{4, 10, 12, 15\}$ and $r = \{4, 6, 2, 3\}$, where the intensity values are represented by \bar{a} and cumulative count is represented by d . The vector r represents run length values with respect to \bar{a} . The RLC algorithm applied on iris images returns the number of 1s in each row. The number of 1s corresponds to the iris pixels, i_p in the segmented image.

Related work: The gaze estimation algorithm based on pupil-corneal reflection and second order polynomial calibration function was proposed by Morimoto and Mimica (2005). An average error was achieved for the

entire screen. Yamazoe *et al.* (2008) developed single-camera-based gaze estimation algorithm. The method consists of facial feature detection, eye model estimation and gaze estimation. An average estimation error is obtained. The method proposed by Lee and Park (2009) used head-mounted display environment. The method used virtual eyeball model by analyzing 3D structure of the eyeball. A head based approach developed by Kaminski and Knaan (2009) determines 3D face orientation from the two glints and bottom point of nose. The gaze detection involves estimating the center of cornea. The method by Ohno *et al.* (2002) use pupil and centroid of the Purkinje image as input to the gaze detection based on eyeball model. The model determines two parameters, center of cornea curvature and center of pupil in the camera coordinate system. The method by Wang *et al.* (2003) estimates gaze using iris contours. Eye gaze is determined as the line joining the eyeball center and iris center in the eye model. A calibration free method by Shih and Liu (2004) estimates gaze direction directly by the orientation of the Line of Sight. In the study proposed by Park (2007), the pupil center and six boundary points of pupil are used for gaze estimation. The gaze vector is obtained by the average of six gaze vectors, where each vector is computed by the cross product of the pupil boundary points.

Proposed system:

Image acquisition: The system is initiated with image acquisition:

- vobj = vi(adn, dvid) (1a)
- pv(vobj, h) (1b)
- conf(vobj) (1c)
- fscreen(scr) (1d)

The video input object vobj is created to aid communication between the system and image acquisition device using (1a) where adn is the adapter and dvid indicates device identification. The video data for vobj is previewed using (1b) for positioning the eye. h contains the information of the image such as video resolution and the number of bands. The video resolution is 640×480. The capturing of the image is event driven. The trigger configuration and the number of frames per trigger is initialized using (1c). The directions screen depicted with degrees on the monitor is displayed in fullscreen mode with a resolution of 1280×800 and a small preview window of size 150×150 on the left bottom side of the screen using (1d). This is shown in Fig. 4.

The image capture begins with acquiring eye images pointing to center direction. This is the pivot point. The image capture continues in anticlockwise direction from 0° at an offset of 30°. The degrees in the screen are depicted as D_i , $i = 1, 2, \dots, 13$ such that $P, 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ, 180^\circ, 210^\circ, 240^\circ, 270^\circ, 300^\circ, 330^\circ$ maps to $D_1, D_2, D_3, D_4, D_5, D_6, D_7, D_8, D_9, D_{10}, D_{11}, D_{12}, D_{13}$. The directions D_1, D_2, D_5, D_8 and D_{11} are reference directions. Gaze directions $D_3, D_4, D_6, D_7, D_9, D_{10}, D_{12}$ and D_{13} are non-reference directions. The non-reference gaze directions are estimated with respect to reference directions. The images in each direction are captured as given in the Algorithm 3.

Algorithm 3: Capture of eye and screen images

Let N be the number of eye images to be captured.

Repeat for $i=1$ to N

The execution of vobj is initiated.

init(vobj)

The image acquisition from vobj is activated. An event is triggered for the capture of eye images.

act(vobj)

The eye images as seen in the preview are captured and buffered.

buffer eye(i)=cap_eye(vobj)

The screen dimensions, width (wd) and height (hgt) are possessed to determine the rectangular coordinates.

[wd,hgt]=scr_size(FC); rec [0 0 wd hgt]

The image is created by reading the pixels from the screen with respect to coordinates given by rec. The image data is buffered to the input stream.

Buffer_screen(i)=cap_scr(FC)

end

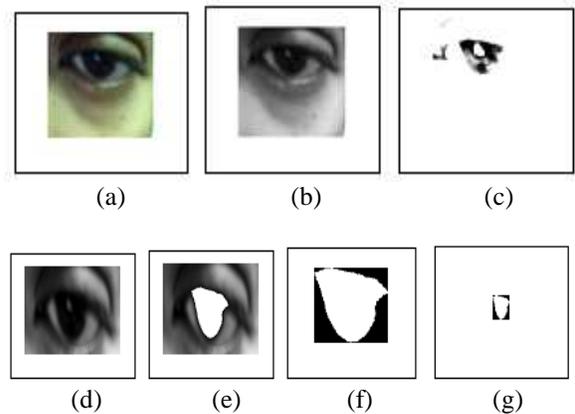


Fig. 5: Preprocessing stages

The center of screen is denoted by pivot P. It is the initial focus point and all other directions are determined with respect to this center point. The changes in the shape of the iris are observed. The system does not require exclusive calibration. With horizontal head movements parallel to screen, the position of the iris with respect to sclera of eye do not change remarkably as shown in the Fig. 5. The primary position of eye is defined anatomically by head and eye planes. Photographic and video analyses show that the primary position of the eyes is a natural constant position in alert normal humans (Jompel and Shi, 1992). If axis of horizontal rotation of head and screen plane are parallel, the algorithm is independent of initial head position.

Preprocessing: The eye images are captured using web camera. The RGB eye image is shown in Fig. 5a. The captured 640×480 sized eye images $Img(x,y)$ are converted to grayscale as shown in Fig. 5b. Filtering is performed using a suitable mask to extract the exact position of the eye using (2a) as shown in Fig. 5c. The mask is of different sizes such as 6×6, 24×6 and 42×6 of 1s. Each subject has a unique mask. The image is binarized and the maximum value of x- coordinate is determined using (2b). The rectangular coordinates for the exact eye position are given by (2c). The image is cropped using (2d). The filter response determines the bounding box for the eye region as shown in Fig. 5d:

$$Img(x, y) = F(Img(x, y), M) \text{ where } M \text{ is the mask} \quad (2a)$$

$$\max_x = \max(Img(x)) \quad (2b)$$

$$R = [0, \max_x - 120, \text{ncols}, 230] \quad (2c)$$

$$I(x, y) = C(Img(x, y), R) \quad (2d)$$

An approximation of pupil center is determined by considering the summation of intensity values. The image Img is scanned row wise to obtain the summation of intensity values s_x using (3a). The minimum of the summation value m_x gives the x-coordinate of the darkest point in the image using (3b). Similarly, the summation of intensity values column wise is s_y and the minimum summation value m_y provides the y-coordinate using (3c) and (3d). The (m_x, m_y) .

values correspond to the point in the pupil area with minimum intensity value:

$$s_x = \sum I(x), x = 1, \dots, \text{nrows} \quad (3a)$$

$$m_x = \min(s_x) \quad (3b)$$

$$\text{sum}_y = \sum I(y), y = 1, \dots, \text{ncols} \quad (3c)$$

$$m_y = \min(s_y) \quad (3d)$$

The value $s_p = (m_x, m_y)$ represents the seedpoint to grow the iris region. Region growing is performed on the eye image as shown in Fig. 5e. Growing terminates when dist value exceeds MID using (4a) and (4b). In the proposed work, the optimum value of MID is 0.05:

$$\text{dist} = I(x, y) - \text{mean}(I(x, y)) \quad (4a)$$

$$R(x, y) \Leftarrow I(x, y) \text{ if } \text{dist} < \text{MID} \quad (4b)$$

The reflections formed in the eye due to illumination are eliminated and smoothing is performed to define the contour of the iris region. The boundary of the iris is determined and iris region is extracted as shown in Fig. 5f. The MID determine the condition for region growing. The segmented iris is normalized to the size 25×25 as shown in Fig. 5i(g).

RLC based on transitional changes: The result of preprocessing is the iris region. The image of the iris region is resized to 50×100. The binary image is scanned row wise and each row is given as input to RLC algorithm. The run length algorithm returns the count of the number of 1s and 0s row wise. The sequences are '1 0', '0 1 0', '0', '1' and '0 1'. Table 1 and 2 show similar coding values for direction pointing to P, D_1 of a subject. Table 3-4 show similar runlength values for direction 90°, D_5 of a subject. The coding values for D_1 and D_5 are different for a subject.

The run length coding algorithm returns the count of runs of 0s and 1s. The first run length of zeros (FRZ) is considered. FRZ gives the count of first occurrence of the sequence of 0s. The FRZ values are similar for each direction. Figure 6 shows FRZ plot for direction D_5 .

Table 1: Run length of sample 1 for direction D_1

Sample 1					
r_1	r_2	r_3	r_4	r_5	
...
99	1				
99	1				
21	2	1	75		1
40	59	1			
44	55	1			
47	5	1	46		1
65	34	1			
47	4	18	30		1
75	24	1			
78	21	1			
...

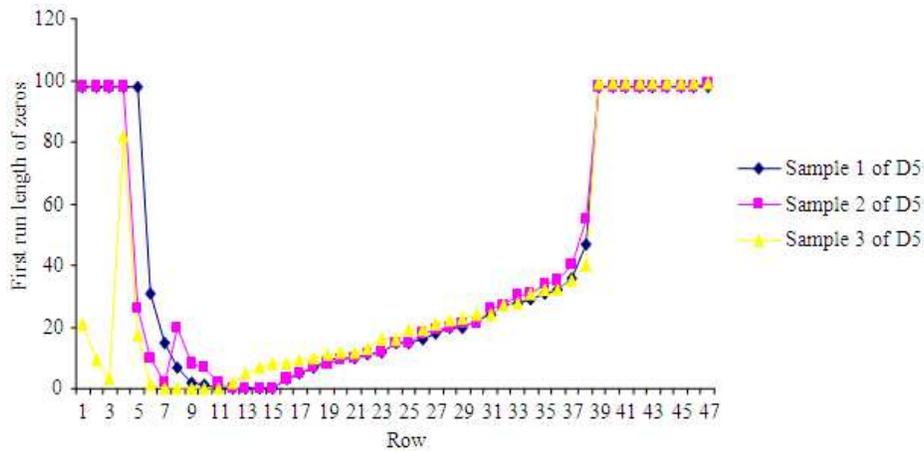


Fig. 6: FRZ for samples in direction D5

Table 2: Run length of Sample 2 for Direction D₁

Sample 2				
r ₁	r ₂	r ₃	r ₄	r ₅
...
99	1			
97	3			
24	1	2	72	1
42	57	1		
43	56	1		
42	57	1	40	1
64	35	1		
54	3	9	33	1
73	26	1		
79	20	1		
...

Table 4: Run length of Sample 2 for Direction D₅

Sample 2			
r ₁	r ₂	r ₃	r ₄
...
1	99		
1	99		
1	9	52	38
73	26	1	
89	10	1	
87	11	2	
97	3		
100			
100			
100			
...

Table 3: Run length of Sample 1 for Direction D₅

Sample 1			
r ₁	r ₂	r ₃	r ₄
...
1	98	1	
3	97		
1	2	51	46
79	20	1	
91	8	1	
92	7	1	
97	2	1	
100			
100			
100			
...

It was observed that for row numbers in the range [20-40], there is variation in the values of FRZ. The corresponding values are shown in Table 5. The variations are due to changes in position of upper eyelid and visible region of iris. The range of values is promising to identify the direction of gaze.

The transitional changes are considered in this range. The change over from 1 to 0 is a negative transition represented by '-1'. Similarly, the change over from 0 to 1 is a positive transition indicated by '+1'. The intensity values and transitional changes for direction D₁ is listed in Table 6. The transitional changes for a D₅ direction are shown in Table 7. It is evident that it is different from D₁. Similar analysis is made for other directions. The count of positive transitions is considered for different gaze directions. The proposed system considers gazing pivot irrespective of initial head position. In this experiment it is evident that, all the other gaze directions can be identified with respect to pivot using difference in positive transition threshold values.

Figure 8 depicts the plot of FRZ in different gaze directions. It was observed that for row numbers in the range [20-40], there is variation in the values of FRZ. The variations are due to changes in position of upper eyelid and visible region of iris. The range of values is promising to identify the direction of gaze. Figure 7 shows the plot of FRZ for directions, D₁, D₅, D₈ and D₂ in the range [1-50] and [20-40].

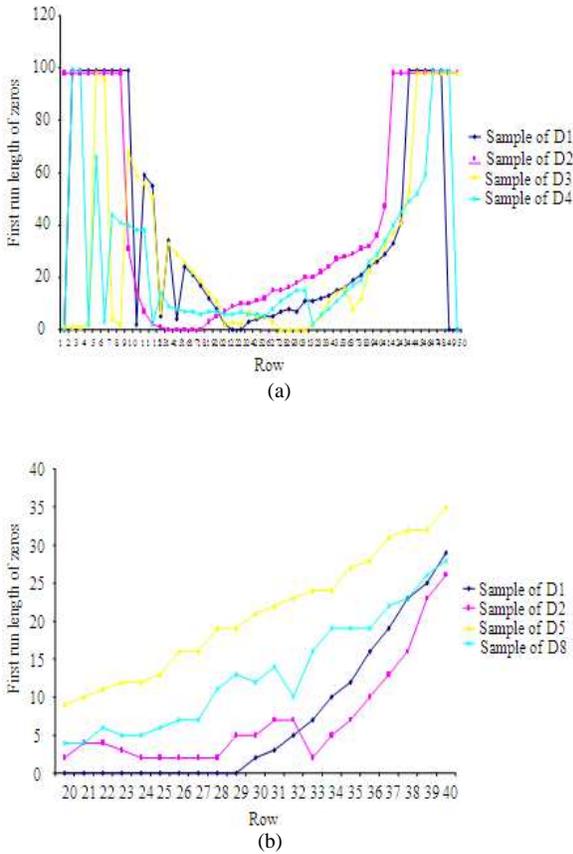


Fig. 7: (a) FRZ for rows in the range [1-50] (b) FRZ for rows in the range [20-40]

Table 5: FRZ values for gaze directions

Row	D ₁	D ₅	D ₈	D ₂
20	0	9	4	2
21	0	10	4	4
22	0	11	6	4
23	0	12	5	3
24	0	12	5	2
25	0	13	6	2
26	0	16	7	2
27	0	16	7	2
28	0	19	11	2
29	0	19	13	5
30	2	21	12	5
31	3	22	14	7
32	5	23	10	7
33	7	24	16	2
34	10	24	19	7
35	12	27	19	10
36	16	28	19	10
37	19	31	22	13
38	23	32	23	16
39	25	32	26	23
40	29	35	28	26

The significant information for shape analysis is given by the number of 1s. The shape of the iris depends on the bent in the eyelid and variations in the iris pixels for each direction. The variations are indicated in the binarised image in terms of growth or shrinkage of iris pixels i_p . The shape of the iris is different for different directions. The number of 1s from the RLC algorithm is different for each direction in each row. The eye images in different directions are shown in Fig. 8.

The count of i_p is considered. The class-interval is defined to determine the variations in the i_p . The iris pixels are assigned to specific class intervals. The class intervals considered are, 1-5, 6-10, 11-15, 16-20 and 21-25. The intervals are denoted by the grades A, B, C, D and E respectively. The sequence of grades forms a unique pattern for each direction. For instance, the sequence {BEEEEEEEEEDDDDDDDDDCCCCBB} represents direction of eye gaze pointing to D₁. The eye samples pointing to the same direction acquire similar Sequence Of Grades (SOG). The grades for direction D₁ and D₅ are given in Table 8 and 9. The grades are shown for two samples of a subject.

The sequence signifies the shape of the iris. In direction D₁, the iris pixels are more concentrated towards the initial few rows and decreases for the other rows. This gives a bulged appearance of the iris. D₁ denotes center of the screen. As the gaze moves to the direction D₅ which denotes 90°, the iris shape becomes elongated and appears lean. The iris pixels are uniformly distributed for most of the rows in the eye image. Similar analysis is done for other directions. The SOG is unique for each direction. There is variation in the grades for iris pixels in the boundary of the class intervals leading to different SOG for the same direction. This is highlighted in the Table 8-9. In order to attain equal weights for i_p , the summation of iris pixels is considered segment wise in horizontal and vertical directions. The horizontal segment, h_s is the summation of iris pixels every 5 rows using (5). The vertical segment, v_s corresponds to summation of iris pixels every 5 columns using (6):

$$h_s = \sum_b^{b+5} i_p(b) \tag{5}$$

Where:
 $b = 1, 5, 10, 15, 20$
 $i = 1, 2, 3, 4, 5$

$$v_s = \sum_c^{c+5} i_p(c) \tag{6}$$

Where:
 $c = 1, 5, 10, 15, 20$
 $j = 1, 2, 3, 4, 5$

Table 6: Transitional changes for direction D₁

Row	a ₁	a ₂	a ₃	a ₄	Transitional changes
20	1	0	1		{-1,+1}
21	1	0	1		{-1,+1}
22	0	1			{+1}
23	0	1			{+1}
24	0	1			{+1}
25	0	1			{+1}
26	0	1			{+1}
27	0	1			{+1}
28	0	1			{+1}
29	0	1			{+1}
30	0	1			{+1}
31	0	1			{+1}
32	0	1			{+1}
33	0	1	0	1	{+1, -1, +1}
34	0	1	0	1	{+1, -1, +1}
35	0	1	0	1	{+1, -1, +1}
36	0	1	0	1	{+1, -1, +1}
37	0	1	0	1	{+1, -1, +1}
38	0	1	0	1	{+1, -1, +1}
39	0	1	0	1	{+1, -1, +1}
40	0	1	0	1	{+1, -1, +1}

Table 7: Transitional changes for direction D₅

Row	a ₁	a ₂	a ₃	a ₄	a ₅	Transitional changes
20	1	0	1			{-1,+1}
21	1	0	1			{-1,+1}
22	1	0	1	0	1	{-1, +1, -1, +1}
23	1	0	1	0	1	{-1, +1, -1, +1}
24	1	0	1	0	1	{-1, +1, -1, +1}
25	1	0	1	0	1	{-1, +1, -1, +1}
26	1	0	1	0	1	{-1, +1, -1, +1}
27	1	0	1	0	1	{-1, +1, -1, +1}
28	1	0	1	0	1	{-1, +1, -1, +1}
29	1	0	1	0	1	{-1, +1, -1, +1}
30	1	0	1	0	1	{-1, +1, -1, +1}
31	1	0	1	0	1	{-1, +1, -1, +1}
32	1	0	1	0	1	{-1, +1, -1, +1}
33	1	0	1	0	1	{-1, +1, -1, +1}
34	1	0	1	0	1	{-1, +1, -1, +1}
35	1	0	1	0	1	{-1, +1, -1, +1}
36	1	0	1	0	1	{-1, +1, -1, +1}
37	1	0	1	0	1	{-1, +1, -1, +1}
38	1	0	1	0	1	{-1, +1, -1, +1}
39	1	0	1	0	1	{-1, +1, -1, +1}
40	1	0	1			{-1, +1}

MATERIALS AND METHODS

The Smart Infocomm digital webcam is used in the experiments for image acquisition. The USB color webcam captures 30 frames per second with a resolution of 640×480. The data format is RGB24. The focal distance is 3cm with 62° view angle.

The class-intervals are defined as 1-25, 26-50, 51-75, 76-100 and 101-125. The grades are assigned as A, B, C, D and E for horizontal direction and F, G, H, I and J for vertical direction. Each horizontal and vertical segment is assigned a grade.

Table 8: SOG for direction D₁

Row	Sample 1		Sample 2	
	i _p Count	Grade	i _p Count	Grade
1	7	B	6	B
2	22	E	22	E
3	25	E	25	E
4	25	E	25	E
5	24	E	24	E
6	24	E	24	E
7	24	E	24	E
8	24	E	23	E
9	23	E	22	E
10	22	E	21	E
11	22	D	20	D
12	20	D	20	D
13	20	D	19	D
14	19	D	19	D
15	19	D	18	D
16	18	D	18	D
17	17	D	17	D
18	16	D	17	D
19	16	D	16	D
20	15	C	15	C
21	14	C	15	C
22	13	C	13	C
23	11	C	11	C
24	9	B	6	B
25	6	B	4	B

Table 9: SOG for direction D₅

Row	Sample 1		Sample 2	
	i _p Count	Grade	i _p Count	Grade
1	11	C	12	C
2	15	C	14	C
3	22	E	20	D
4	21	E	21	E
5	19	C	19	D
6	18	D	19	D
7	18	D	19	D
8	17	D	19	D
9	16	D	18	D
10	16	D	18	D
11	16	D	18	D
12	17	D	18	D
13	17	D	17	D
14	16	D	16	D
15	16	D	16	D
16	14	C	16	D
17	13	C	15	C
18	12	C	14	C
19	11	C	12	C
20	10	B	10	B
21	10	B	10	B
22	9	B	10	B
23	7	B	9	B
24	6	B	7	B
25	4	A	4	A

RESULTS AND DISCUSSION

The SOG form similar patterns for same direction. Table 10-11 show SOG for directions D₁ and D₅ for two samples of a subject. Similar analysis has been made for all other directions. The segment-wise SOG values determine the gaze direction. The correct recognition rate of 95% was achieved.

Table 10: Segment-wise SOG for direction D₁

Row	Sample 1		Sample 2	
	\sum_i	Grade	\sum_i	Grade
hs ₁	72	C	75	C
hs ₂	116	E	116	E
hs ₃	106	E	107	E
hs ₄	86	D	87	D
hs ₅	50	B	49	B
vs ₁	57	H	61	H
vs ₂	100	J	102	J
vs ₃	83	I	87	I
vs ₄	110	J	107	J
vs ₅	50	G	50	G

Table 11: Segment-wise SOG for direction D₅

Row	Sample 1		Sample 2	
	\sum_i	Grade	\sum_i	Grade
hs ₁	95	D	89	D
hs ₂	103	E	101	E
hs ₃	91	D	91	D
hs ₄	79	D	73	D
hs ₅	46	B	43	B
vs ₁	53	H	49	H
vs ₂	113	J	110	J
vs ₃	123	J	120	J
vs ₄	96	I	93	I
vs ₅	28	G	26	G

CONCLUSION

A simple gaze tracking technique using webcam has been developed. The value of MID = 0.05 is determined as optimum value for region growing. The variation in eyelid position and shape of the iris determine the gaze direction. For each gaze direction, a unique SOG is generated. Further, the distance between camera and the screen can be varied to analyze the gaze directions. The distance between direction D₁ and other directions can be investigated to fix the radius of vision with respect to pivot point.

REFERENCES

Abas, K.H. and O. Ono, 2010. Implementation of multi-centroid moment invariants in thermal-based face identification system. *Am. J. Applied Sci.*, 7: 283-289. DOI: 10.3844/AJASSP.2010.283.289

Baluja, S. and D. Pomerleau, 1993. Non-Intrusive Gaze Tracking Using Artificial Neural Networks. Carnegie Mellon University. <https://www.aai.org/Papers/Symposia/Fall/1993/F S-93-04/FS93-04-032.pdf>

Gonzalez, R.C. and R.E. Woods, 2005. *Digital Image Processing*. 2nd Edn., Prentice Hall of India, New Delhi, ISBN: 8120327586, pp: 793.

Hansen, D.W. and Q. Ji, 2010. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Trans. Patt. Anal. Mach. Intell.*, 32: 478-500. DOI: 10.1109/TPAMI.2009.30; PMID: 20075473

Hansen, D.W., J.P. Hansen, M. Nielsen, A.S. Johansen and M.B. Stegmann, 2002. Eye typing using Markov and Active Appearance Models. *Proceeding of the IEEE Workshop on Applications of Computer Vision*, pp: 132-136. DOI: 10.1109/ACV.2002.1182170

Jompel, R.S. and D.X. Shi, 1992. The primary position of the eyes, the resetting saccade, and the transverse visual head plane. *Head movements around the cervical joints*. *Invest. Ophthalmol. Visual Sci.*, 33: 2501-2510. PMID: 1634348

Kaminski, J.Y. and D. Knaan, 2009. Single image face orientation and gaze detection. *Mach. Vision Appl.*, 21: 85-98. DOI: 10.1007/S00138-008-0143-1

Khosravi, M.H. and R. Safabakhsha, 2008. Human eye sclera detection and tracking using a modified time-adaptive self-organizing map. *Patt. Recog.*, 41: 2571-2593. DOI: 10.1016/J.PATCOG.2008.01.012

Kiat, L.C. and S. Ranganath, 2004. One-time calibration eye gaze detection system. *IEEE Int. Conf. Image Proc.*, 2: 873-876. DOI: 10.1109/ICIP.2004.1419438

Kim, K.N. and R.S. Ramakrishna, 1999. Vision-Based Eye-Gaze Tracking for Human Computer Interface. *Proceeding of the IEEE International Conference on Systems*, Oct. 12-15, IEEE Xplore Press, Tokyo, Japan, pp: 324-329. DOI: 10.1109/ICSMC.1999.825279

Lee, E.C. and K.R. Park, 2009. A robust eye gaze tracking method based on a virtual eyeball model. *Mach. Vision Appl.*, 20: 319-337. DOI: 10.1007/S00138-008-0129-Z

Morimoto, C. and M. Mimica, 2005. Eye gaze tracking techniques for interactive applications. *Comput. Vision Image Understand.*, 98: 4-24. DOI: 10.1016/J.CVIU.2004.07.010

Ohno, T., N. Mukawa and A. Yoshikawa, 2002. Freegaze: A gaze tracking system for everyday gaze interaction. *Proceedings of the Symposium on Eye Tracking Research and Applications*, (ETRA '02), ACM New York, NY, USA, pp: 125-132. DOI: 10.1145/507072.507098; PMCID: 2391256

Park, K.R., 2007. A Real Time Gaze Position Estimation Method based on a 3-D Eye Model. *IEEE Trans. Syst. Man Cybernet.*, 37: 199-212. DOI: 10.1109/TSMCB.2006.883426

- Rikert, T.D. and M.J. Jones, 1998. Gaze estimation using morphable models. Proceedings of the 3rd IEEE International Conference on Automatic Face Gesture Recog., Apr. 14-16, IEEE Xplore Press, Nara, Japan, pp: 436-441. DOI: 10.1109/AFGR.1998.670987
- Shih, S.W. and J. Liu, 2004. A Novel Approach to 3-D Gaze Tracking Using Stereo Cameras. IEEE Trans. Syst. Man Cybernet., 34: 234-245. DOI: 10.1109/TSMCB.2003.811128
- Tan, K.H., D.J. Kriegman and N. Ahuja, 2002. Appearance-based eye gaze estimation. Proceedings of the 6th IEEE Workshop on Applications of Computer Vision, Dec. 3-4, IEEE Xplore Press, USA pp: 191-195. DOI: 10.1109/ACV.2002.1182180
- Wang, J.G., E. Sung and R. Venkateswarlu, 2003. Eye Gaze Estimation From a Single Image of One Eye. Proceedings of the 9th IEEE International Conference on Computer Vision, Oct. 13-16, IEEE Xplore Press, Nice, France, pp: 136-143. DOI: 10.1109/ICCV.2003.1238328
- Yamazoe, H., A. Utsumi, T. Yonezawa and S. Abe, 2008. Remote and Head-Motion-Free Gaze Tracking for Real Environments with Automated Head-Eye Model Calibrations. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, June, 23-28, IEEE Xplore Press, Anchorage, AK., pp: 1-6. DOI: 10.1109/CVPRW.2008.4563184
- Yoo, D.H., J.H. Kim, B.R. Lee and M.J. Chung, 2002. Non-Contact Eye Gaze Tracking System by Mapping of Corneal Reflections. Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition, May, 20-21, IEEE Xplore Press, Washington, DC, USA., pp: 94-99. DOI: 10.1109/AFGR.2002.1004139
- Zhu, J. and J. Yang, 2002. Subpixel Eye Gaze Tracking. Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition, May, 20-21, IEEE Xplore Press, Washington, DC, USA., pp: 124-129. DOI: 10.1109/AFGR.2002.1004144
- Zhu, Z. and Q. Ji, 2004. Eye and gaze tracking for interactive graphic display. Machine Vision Appl., 15: 139-148. DOI: 10.1007/S00138-004-0139-4
- Zhu, Z., Q. Ji and K.P. Bennett, 2006. Nonlinear eye gaze mapping function estimation via support vector regression. Proceedings of the 18th IEEE International Conference on Pattern Recognition, Aug. 20-24, IEEE Xplore Press, Hong Kong, 1132-1135. DOI: 10.1109/ICPR.2006.864