Fourth International Conference on Electronics, Communication and Aerospace Technology (ICECA-2020)
IEEE Xplore Part Number: CFP20J88-ART; ISBN: 978-1-7281-6387-1

Feature based Methods for Eye Gaze Tracking

Sheela S V
Member, IEEE
Professor, Dept. of ISE
BMS College of Engineering, Bangalore

Radhika K R
Senior Member, IEEE
Professor, Dept. of ISE
BMS College of Engineering, Bangalore

***Abstract*—Eye gaze has been depicted as one of the means for interaction with computer. The eye tracking systems perform feature extraction and determination on gaze direction. The feature based methods will estimate gaze direction based on the extracted features. The parameters of pupil, iris and eye corners will determine gaze direction. In this paper, two feature based gaze tracking methods are proposed. The first method is based on Fast Fourier Transform (FFT) and the second method is based on number of zeros in Z-domain. The search for template image in successive video frames tracks the pupil. The optical flow extracts the change in position. The velocity component determines displacement of template region, which helps in gaze direction estimation. The Fourier transform is computed by using the magnitude of velocity components, gives the frequency representation for each gaze direction. The number of zeros inside a specific range of circle in Z-domain aids in determining the gaze direction. The recognition rate of 98.5% and 99% are obtained for the methods based on FFT and zeros, respectively.**

Keywords—Fast Fourier Transform; Number of Zeros; Haarcascaded classifier, Kalman prediction; Optical flow.

I. INTRODUCTION

Human attention and focus identified through gaze is a natural means for communication. The spatial and temporal characteristics of eye are useful for real-time monitoring systems [1]. Fixation and saccades represent the visual behavior. Pausing of gaze at certain positions is the fixation and movement from one position to another is referred as saccade. Fixation and saccades together constitute for movements of eye. Significant information for an application is retrieved by fixations. The duration, number and spatial density of fixations are useful parameters to analyze eye movements [2]. Typical fixation period ranges between 200-600 milliseconds [3]. Controlling eye actions require capture of intended and suppression of unintended movements.

Major challenges with capture of images in visible light are spectrum variations, low contrast and increased noise levels. Distance between subject & camera, eye occlusion due to face rotation, eye closure and intrinsic properties of eyes affect the visibility of iris. Pupil center needs to be repeatedly determined due to eye blinks. Incorrect detection of gaze direction is possible with faster saccades [4]. Dwell time determination is important as it is the predefined time to select the target.

Human interaction systems are generally based on keyboard, mouse, touch and voice. Eye contingent systems respond to user based on observed eye movements. Few applications are mentioned here. Interpretation of eye gestures is decoded to obtain spoken words for communication [5]. The application provides a different user interface for people with motor disabilities. Eye and head movements are used to collectively track the visual attention of players in a game [6]. An adaptation of browser capabilities has been designed to access email for users with neurological disorders [7]. The inputs from gaze controls the navigation of teleoperated mobile robot in a human-robot interface [8]. An eye-controlled wheel chair navigation system guides the movement of disabled with feedback mechanism [9]. Visual findings captured at construction sites provide data for effective processing and automation [10]. Usability evaluation based on eye movements is suitable for task performance of interfaces [11]. Several editing options such as selection of text and character typing are integrated with gaze input in addition to normal interaction [12]. Assessment of eye fatigue is measured based on fixation points [13].

In the proposed method, a dynamic approach is initiated by capture of video frames using camera. Preprocessing includes face and eye detection using Haar cascade classifier. The iris and pupil are parametric features. Selection of features increases the accuracy of system [14]. The template is detected and repositioned using Kalman prediction. The optical flow is used to determine the velocity in successive frames. Velocity components computed is the feature vector for direction estimation using FFT and number of zeros. Representation in frequency domain defines certain periodicity and reduces noise. FFT coefficients provide a compressed depiction. Response characteristics and system dynamics are obtained in Z-domain. The stability of system is proportional to number of zeros.

II. PRELIMINARIES

The fundamental concepts of fast Fourier transform and zeros are discussed in this section.

A. Fast Fourier Transform

In FFT, when the zero-frequency component is shifted to the middle of the spectrum, it provides a basis for visual classification. The Fourier coefficients of discrete time signal, $x(n)$ is given as $X(k) = x(n)\, e^{-j2\pi/N}, 0 \le k \le N-1$ where $X(k)$ denotes sequence of N complex valued numbers. The inverse transformation is given as $x(n) = \frac{1}{n} X(k) e^{-j2\pi/N}, 0 \le k \le N-1$. Fourier transform is used to detect definite

periodicity and directionality in the frequency domain [15]. Thresholding high frequency coefficients reduces the effect of noise. FFT provides a compact representation using fixed number of coefficients [16]. The transformation converts the representation from spatial domain to frequency domain [17].

B. Zeros

A zero is the value at time *t* of a continuous function *c(t)* such that the transfer function is zero. Zeros are considered as frequencies where the overall gain of the function is zero. If the order of the numerator polynomial is smaller than the order of the denominator polynomial, it defines a case of proper transfer function. The challenge is to realize transfer function if the order of numerator polynomial is greater than the order of denominator polynomial. Zeros are defined as roots of the numerator.

The system dynamics are represented as zeros, whose location provide response characteristics [18]. They are depicted graphically by plotting their locations on complex s-plane, whose axes denote the real and imaginary parts of the complex variables. The calculation of zeros of polynomials is a classical computational problem [19]. A classical Schur-Cohn problem is encountered for a polynomial with real coefficients. The problem is associated with counting number of zeros inside and outside the circle of specified value, establishing a necessary and sufficient condition for all zeros that lies inside it. These conditions are relevant to the study of the stability of the system [20]. Certain sign variation pattern of the polynomials in the sequence are related to the number of zeros inside and outside the circle of specified value. Polynomials with all zeros inside circle are called stable polynomials [21].

III. Related Work

The non-intrusive methods in gaze tracking capture images of eye using video cameras in natural light or infrared radiation. These techniques are classified into appearance and feature based. Image content is used to estimate gaze direction in appearance based techniques. Significant calibration points, morphable model, manifold, interpolation and cross ratios are the different ways used to determine gaze direction [22]. Gaze estimation in feature based methods use extracted features. Pupil center, corneal reflection, iris center, edges of iris, eye corner, sclera region, inter-pupil distance and orientation of pupil ellipse characterize the features for determination of gaze direction.

A linear approximation technique has been proposed by Kim et. al., using displacement values between iris center and marker [23]. Position of marker and iris radius are the features extracted. The angles of calibration are considered in horizontal and vertical directions using local features such as eye corner and iris center positions in the method by Yang et. al [24]. A non-linear function is devised using pupil and glint parameters [25]. Gaze direction is trained using radial basis function with pupil centers & orientation, pupil area and pupil-glint vectors. A set of pupil geometric features is input for generalized regression neural networks [1]. The parameters correspond to changes in pupil positions. Calibration points are given as input for neural networks to obtain an output layer representing screen coordinates in the method by Weston et. al [26]. Template matching is used to extract iris features and pupil intensity values.

An eye tracker for keyboard application use tree-based selection system [27]. Nearest target item in virtual keyboard is selected using Euclidean distance. A geometrical model in low resolution environment proposed by Martinikorena et. al. considers corneal radius between optical and visual axes [28]. Second degree polynomial is formulated for interpolation model. Projection of pupil center on corresponding visual axis is devised in the method by Su et. al [29]. Pupil center is used to predict epipolar line and visual axis. Extrinsic calibration method by Kang et. al. determines 3D gaze coordinates using head movement and camera configuration [30]. Implementation is based on maximum likelihood estimation.

Gaze tracking systems are designed using convolutional neural network models. The network is trained through multiple abstraction levels [31]. Resnet model used for eye detection and gaze direction determine nine gaze positions and one blink status [32]. Eye patches are trained using VGG16 for real time gaze estimation in natural environments to determine yaw and pitch angles [33]. The network is trained for left and right images separately and a single gaze vector is estimated by Lemley et. al. [34]. A shallow fully convolutional network with large kernel block is developed by Xia et. al., to localize eye center [35]. Lian et. al. proposed a multitask method using Resnet for single-view and cross-view feature fusion networks [36]. The investigation from existing literature relevant to feature based methods are presented in Table I.

TABLE I. FINDINGS FROM EXSITING FEATURE BASED METHODS

Author	Findings
Kim et. al.	Orthogonal projection is used in imaging system by maintaining the displacement between iris center and eye gaze. An adaptive estimation technique requires displacement values and reference model.
Yang et. al.	Gaze regions are determined by gaze angle based on interpolation and local features. Information from single pixel is eliminated.
Kiat et. al.	Pupil center is obtained by ellipse fitting eliminating concave surfaces formed due to glints. Difference image in each frame is obtained resulting in computation cost.
Ji et. al.	Calculation of gaze direction using support vector regression eliminates the need for prior parametric model using linear & polynomial kernels. Data mapping to high dimensional space require regression functions.
Weston et. al.	Pre-processing include image processing such as histogram equalization and downsampling. Head movements lead to inaccurate gaze estimation.

IV. Proposed System

The components of proposed system begin with image acquisition, iris detection, template determination, repositioning of template, velocity computation and determination of gaze direction. The workflow is depicted in Fig. 1 and explained in following subsections.

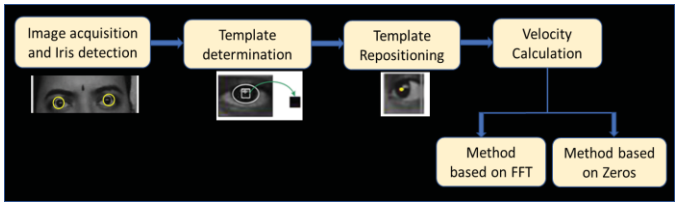

Fig. 1. Workflow of proposed system

A. Image acquisition and Iris detection

Initially, the face image is captured from frames acquired using system built-in camera. The interaction between the computer and image acquisition device is obtained by the video object. Each video object is provided with an adapter and identification number. The information of the image such as video resolution and number of channels are contained in the video object handler. The single channel is used for gray scale image and three channels for RGB image. The capturing of the image is event based by initializing the number of frames per trigger. The depth is used to indicate storage types.

The face and eyes are detected in the frame of capture. The detection of face in a video frame consists of four stages [37]. The first is the determination of integral image obtained by computing summation over image subregions. The second stage extracts Haar features from the integral image. The feature set is generated using rectangular combinations. The edges, lines and center-surround features are identified. A significant set of features is selected. The third stage uses adaboost technique to combine many weak classifiers to create one strong classifier. Weighted combination of weak classifiers results in a strong classifier. The fourth stage is to develop a cascade of classifiers. The face and eyes are detected only if it passes through the entire cascade. The entire image is searched to detect single face, which is the region of interest for gaze tracking by eliminating background.

The eye region is cropped from detected face. Haar features detect pupil, sclera and eye corners. Detection of eye is performed using cascade of adaboost classifiers by eliminating eyebrows, forehead and region surrounding the eyes. Hough gradient method is applied to detect iris. The first order gradient function defined along x and y directions determine center of iris. Template is the specific region surrounding iris center which is the basis for estimation of gaze direction.

B. Template Determination

Template matching is a technique for finding small parts of an image that match the template image. The template is an image patch. The input image used to find the match is the search image. The method is a multiplicative operation of the template against the search image. The template and part of search image is matched using mean values. The method works by sliding the template image patch across the search image for the best match. The criteria for matching is based on the correlation coefficient given in equation (1).

$$res(x,y) = \sum_{x_1,y_1}[tem(x_1,y_1).s\_img(x+x_1,y+y_1)]^2 \quad (1)$$

where *tem* is the template image, *s_img* is the search image and *res* is the resultant image. The x_1 and y_1 are increment values in x and y directions. The value of the resultant image is obtained such that *res(x, y)* is 1 for a perfect match, -1 for a perfect mismatch and 0 for uncorrelated values.

The templates are framed within iris region in the first frame. The first frame corresponds to gaze direction at pivot point in the center of screen. The left eye images are considered for computation. A *LxL* region from the center of the iris circle form the template image. The template determination is shown in Fig. 2(a). A search window of size 2Lx2L is computed from the center of the iris. The value of *L* is 10. Iris and templates are determined from the first frame. The template is matched in the search image for each successive frame captured from the camera as shown in Fig. 2(b).

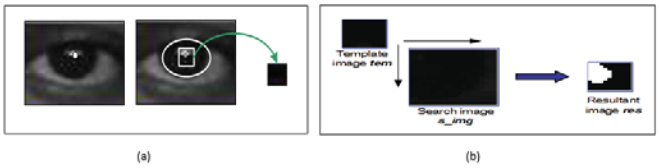

Fig. 2. (a) Template formation (b) Template Matching

The detection of iris using Hough gradient method and determination of templates is not necessary for every frame. The template matching procedure results in tracking iris and pupil, with determination of template in the initial frame. The resultant image displays maximum value for highly correlated points. The location of the best match indicated by global maxima is determined. It is required to update the iris center coordinates for continuous tracking of iris. The template matching continues until the global minima is less than threshold value or search window is within the frame boundaries. The threshold value is set to 1.7. The algorithm for template matching is given in Algorithm 1.

Algorithm 1: Template Matching
Input: Video frames from camera
Output: Template detected in successive frames
Steps:
1. Camera is initialized. The first frame is captured along with video properties, depth and number of channels.
2. Face and eyes are detected using Haar cascade classifiers.
3. Hough gradient method is applied to obtain iris center and radius.
4. Template image, *tem,* and search image, *s_img,* are determined from the center of iris.
5. while((minima value < threshold) | (search image < frame boundaries))
   (i)Successive frames are captured
   (ii)Template matching is performed based on correlation coefficient.
   (iii)Location of maxima value according to correlation values is determined.
   (iv)Iris center coordinates are updated using location of maxima value.
end{while)

C. Template Repositioning

In the scenario to track the object in successive frames of video sequence, the next position may or may not lie within the search region. The next position of the object within search region indicates unpredictable situation. If the next position is outside boundaries of search region, prediction is essential.

Kalman prediction is used to estimate position of iris center in the next frame based on its previous frame position. The system is described in terms of system model S_t and measurement model M_t using $S_{t+1} = AS_t + W_t$; $M_t = HS_t + V_t$, where A is the state transition matrix and H is the measurement matrix used to map the measurements to states. Q is the process variance matrix and R is the measurement variance matrix. Q measures the variance and R measures the error from measurements which accounts for possible changes between t and $t+1$. The W_t denote model noise and V_t denote the measurement noise. The noise term is a Gaussian random variable with zero mean and covariance matrix Q. The probability distribution is given by $pr(W) \approx N(0, Q)$ and $pr(V) \approx N(0, R)$. The filtering operation is performed using equations (2)-(4).

$$K_t = P_{t|t+1}H_t^T(R + H_tP_{t|t-1}H_t^T)^{-1} \quad (2)$$

$$P_{t|t} = (I - K_tH_t)P_{t|t-1} \quad (3)$$

$$S_{t|t} = X_{t|t-1} + K_t(M_t - H_tX_{t|t-1}) \quad (4)$$

where K_t and $P_{t|t}$ are Kalman gain and error covariance matrices at time t. The prediction is calculated using $P_{t+1|t} = AP_{t|t}A^T + Q$;
$S_{t+1|t}=AX_{t|t}$. The state vector is given by $S_t(x_t, y_t, x_t', y_t')$ where x_t' and y_t' are first order derivatives of x_t and y_t, when there is change in pupil position. The matrices, $A = \{1\ 0\ \Delta T\ 0\ ;\ 0\ 1\ 0\ \Delta T\ ;\ 0\ 0\ 1\ 0\ ;\ 0\ 0\ 0\ 1\}$ and $H = \{1\ 0\ 0\ 0\ ;\ 0\ 1\ 0\ 0\}$, where Δ T is the time of frame acquisition. The value of Δ T is set to 0.04 sec.

The pupil center and the template are computed using values based on Kalman prediction. The measurement vector is the position of the pupil center. The initial position of the pupil center is obtained from the Hough gradient method. The predicted pupil center $M_t = (\hat{x_t}, \hat{y_t})$ is determined with initial values. The new measurements are computed based on the predicted values. The Kalman prediction measurements for a sequence of frames are given in Table II and frames are shown in Fig. 3. The sequence of frames is captured for gaze directions from left towards right.

TABLE II. PUPIL CENTER PREDICTION

Frame	1	2	3	4	5	6	7	8	9	10
x	24	23.62	23.95	23.9	23.86	23.81	23.76	23.71	23.67	23.62
y	15	15.19	15.05	15.1	15.15	15.21	15.26	15.31	15.36	15.42

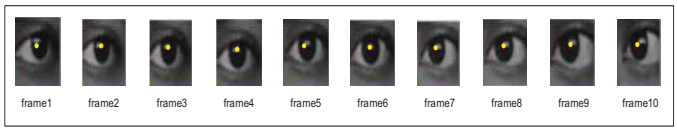

Fig. 3. Sequence of frames for pupil center prediction

D. Velocity Calculation

Every pixel in an image is associated with velocity. The velocity represents the distance a pixel has traversed between previous frame and current frame. The optical flow is used to estimate the motion between two frames. The Lucas-Kanade optical flow algorithm for velocity calculation is based on three assumptions. Firstly, the brightness of a pixel does not change from one frame to another. The second assumption is that the temporal increments are faster compared to motion in the image. The third assumption is that the neighboring points in a scene belong to the same surface [38].

In an image img, img_x and img_y are the spatial derivatives across the first image in x and y directions and img_t is the derivative between images with respect to time. The optical velocity is given by equation (5).

$$img_x . v_x + img_y . v_y + img_t = 0 \quad (5)$$

where v_x is the x component of velocity and v_y is the y component of velocity. Initially, the points of interest are identified and their movement is converted to velocity vectors. The template image, tem is the part in iris region that is tracked. The points of interest are identified to be every alternate pixel in x and y direction. The points of interest are shown in Fig. 4(a).

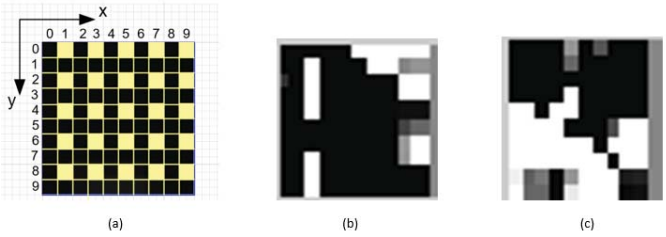

Fig. 4. (a) 10x10 template with points of interest highlighted (b) Horizontal component (c) Vertical component

The velocity is computed at points of interest based on the displacement of template, in successive frames. The horizontal and vertical components of velocity for the selected points of interest in template image are shown in Fig. 4(b)-(c). The magnitude of horizontal and vertical components is computed using $mag_v = \sqrt{{v_x}^2 + {v_y}^2}$. The magnitude, mag_v values is the feature vector for computation in proposed gaze tracking methods.

E. Method based on FFT

The display screen considered is shown in Fig. 5. The magnitude of velocity components mag_v are computed for all the direction points, D_2 to D_{13} with respect to the pivot point D_1. The δ represents intermediate points.

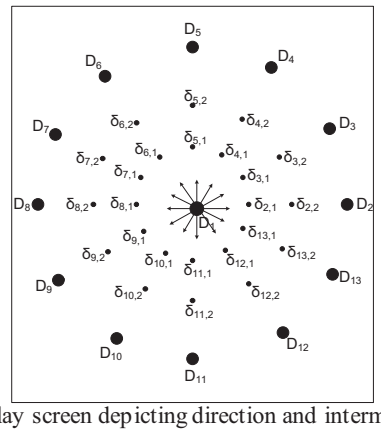

Fig. 5. Display screen depicting direction and intermediate points

The subject gazes at directions D_2 to D_{13} in a specific fashion. The velocity is computed at direction points and intermediate points. For instance, the sequence of points at which the subject looks for direction D_5 is $\{D_5, \delta_{5,2}, \delta_{5,1}, D_1, \delta_{5,1}, \delta_{5,2}, D_5\}$. At each point of gaze, the velocity is computed. The swift movement of eye results in the loss of template and repeated re-computation of template. The intermediate points are important as it allows template to remain within search window. The magnitude value mag_v is normalized and exponent of normalized values is computed. FFT is applied to the exponent values. The real part is extracted. The highest value is eliminated and frequency plots of residuals are considered. The plot for an instance of direction D_2 is shown in Fig. 6(a).

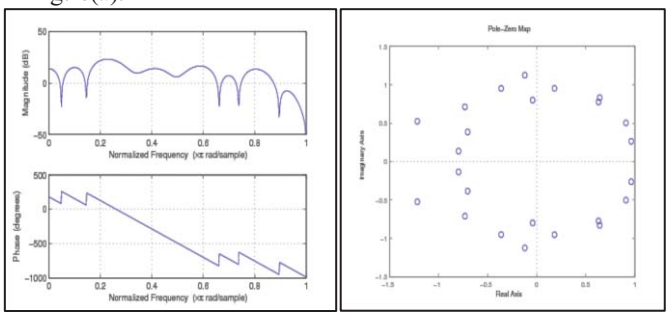

Fig 6. Direction D_2: (a) Frequency plot (b) Zero plot

The average of real part is denoted by avg_r. The value of avg_r is determined for each direction. The average at direction points are considered. The difference in the average value of a direction point and D_1 is computed. For instance, considering the sequence points, $\{D_5, \delta_{5,2}, \delta_{5,1}, D_1, \delta_{5,1}, \delta_{5,2}, D_5\}$, $r_1 = avg_r(D_5) - avg_r(D_1)$ and $r_2 = avg_r(D_1) - avg_r(D_5)$ are computed in the order of sequence. The value r_1 is the difference of avg_r for any direction and the pivot point. The value r_2 is the difference of avg_r for pivot position and any other direction. The ratio $r_d = r_1/r_2$ is determined. The r_d values differ for each gaze direction. A threshold range, Th_{fft} is fixed for each direction. The r_d values for a particular direction fall in that range for a subject. The range is determined during training. The testing sample with r_d value in the range Th_{fft} determine the gaze direction.

F. Method based on Zeros

In order to obtain qualitative insight into the response characteristics of the system, zeros are used. The transfer function is defined for mag_v values. The zeros are obtained and real values are extracted. It is observed that for a specific range, *srange*, the number of zeros, *nz* differ for each gaze direction.

The *nz* values are computed at sequence points similar to the method based on FFT values. The zero plots for an instance of direction D_2 are shown in Fig. 6(b) with *srange* = 0.9. The *srange* values considered are in the range {0.85 - 0.95} with increments of 0.01 for each direction. The *nz* values are determined for different directions that is viewed as the last point in the sequence for a subject with L=10. An instance of five records is shown in Table III.

TABLE III. NUMBER OF ZEROS FOR DIFFERENT DIRECTIONS

srange	D_2	D_3	D_4	D_5	D_6	D_7	D_8
0.88	4	6	10	12	17	19	22
0.89	3	7	10	12	17	19	22
0.9	3	7	11	14	17	20	23
0.91	4	8	11	13	17	20	22
0.92	4	7	12	14	16	19	22

Observing the *nz* values for different gaze directions, the difference is promising for *srange*=0.9. Continuing with the same *srange* value, L values are modified. The values of L considered are 12, 14 and 16. The *nz* values are proportional to the points of interest. The zeros are roots of the numerator in the polynomial. The maximum number of *nz* values is one value lesser than the points of interest. The increase in L value increases the *nz* values. The *nz* values for each direction are different. Further increase in the L value exceeds the size of the iris region. The threshold range is saved as Th_{zero} for each gaze direction with L=14. The value of L is selected by keeping track of different iris radius for various ethnic classes. The value of Th_{zero} is determined during training for a subject. The *nz* value for a test session that lies in the range Th_{zero} is used to determine gaze direction.

V. EXPERIMENT AND RESULTS

The experiment is conducted for 50 subjects. The acquisition is performed using the system's built in camera. The illumination for image capture is 500 lux from a fluorescent light source. The camera captures a sequence of video frames at 30 frames per second with a resolution of 640x480. The implementation has been carried out using Matlab R2020a and OpenCV. During training, the velocity components are computed in five sessions for each direction. In the method based on FFT, r_d values are used to distinguish gaze directions. The r_d value that lie in the range Th_{fft} determine the direction. It is observed that as the gaze changes in an anticlockwise manner for directions above pivotal position, the r_d values decreases. Similar analysis is inferred for directions below the pivotal position. The video sequence of eye gaze starting from any of the direction to D_1 via intermediate points and back to the direction is provided as test input. The testing is carried out in a session per direction for a subject.

In the method based on zeros, five sessions are used during training to determine *nz* values for each direction. The value Th_{zero} is fixed for each direction per subject. The *nz* values are different for each direction and value of *srange*. The *nz* value for a test sample in the range Th_{zero} determine the gaze direction. The results are tabulated in Table IV in terms of Recognition rate, False Rejection Rate (FRR) and False Acceptance Rate (FAR). The best results of 98.5% recognition rate is obtained for direction D_7 in method based on FFT. In the method based on zeros, 99% recognition rate is attained for direction D_5.

TABLE IV. RESULT (IN %) OF THE METHOD BASED ON FFT AND ZEROS

	Method based on FFT			Method based on Zeros		
Direction	Recognition rate	FRR	FAR	Recognition rate	FRR	FAR
D_1	98.2	1	0.8	98.2	1	0.8
D_2	98.1	0.9	1	98.1	0.6	1.3
D_3	98.4	1	0.6	98.2	0.7	1.1
D_4	98.3	0.7	1	98.1	0.9	1
D_5	98.3	0.8	0.9	99	0.5	0.5
D_6	98.2	1	0.8	98.1	1	0.9
D_7	98.5	0.8	0.9	98.3	0.8	0.9
D_8	98.1	1	0.9	98.2	0.8	1
D_9	98.2	0.9	0.9	98.5	1	0.5
D_{10}	98	1	1	98.3	1	0.7
D_{11}	98.4	1	0.6	98.4	0.6	1
D_{12}	98.1	0.9	1	98.2	1	0.8
D_{13}	98.3	0.8	0.9	98.1	1.1	0.8

VI. CONCLUSION AND FUTURE WORK

The feature based gaze tracking based on FFT values and zeros give distinguishing values for each direction. FFT is an approach to measure frequency of parameters in the frame. The peak values are analyzed on subject basis. The measurements which are hard to determine in time domain are accurately represented in the frequency domain. In the method based on zero values, the classification is purely on intrinsic knowledge hidden with representation in the Z-domain. A qualitative insight into the response characteristics of the system is obtained using zeros. The intermediate points in display screen improve the accuracy by minimizing sudden transitions. The template is maintained within search window. The system provides a dynamic model to determine directions for gaze enabled applications. An intuitive and efficient communication through gaze is applied in optical interfaces, driving simulators, gaming, mixed reality systems, advertising, marketing and web usage. Further, gaze interfaces synchronized with authentication or recognition provide basis for security system. On-line authentication/recognition can be established during gazing using tracked images at pivotal position D_1.

REFERENCES

[1] Zhu, Z. and Ji, Q., "Eye and gaze tracking for interactive graphic display", Machine Vision and Applications, Vol. 15, pp. 139–148, 2004.
[2] Piyasirivej, P., "Using eye gaze tracking for the usability evaluation of web sites", IADIS Virtual Multi Conference on Computer Science and Information Systems, pp. 452–459, 2005.
[3] Hennessey, C., Noureddin, B. and Lawrence, P., "Fixation Precision in High-Speed Noncontact Eye-Gaze Tracking", IEEE Transactions on Systems, Man and Cybernetics, pp. 289–298, 2008.
[4] Hansen, D.W. and Ji, Q., "In the Eye of the Beholder: A Survey of Models for Eyes and Gaze", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 478–500, 2010.
[5] Zhang X, Kulkarni H and Morris M.R, "Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities", CHI Conference on Human Factors in Computing Systems, 2017.
[6] Spakov O., Istance, H., Raiha, K-J, Viitanen T. and Siirtola H., "Eye gaze and head gaze in collaborative games", 11th ACM Symposium on Eye Tracking Research and Applications, Article No: 85, pp. 1-9, 2019.
[7] Lupu R.G., Bozomitu R.G., Pasarica A and Rotariu C., "Eye Tracking User Interface for Internet Access used in Assistive Technology", 6th IEEE International Conference on E-Health and Bioengineering, 2017.
[8] Gego D., Carreto C. and Figueiredo L., "Teleoperation of a mobile robot based on eye-gaze tracking", 12th Iberian Conference on Information Systems and Technologies, 2017.
[9] Singer C and Hartmann B., "See-Thru: Towards Minimally Obstructive Eye-Controlled Wheelchair Interfaces", 21st International ACM SIGACESS Conference on Computers and Accessibility, 2019.
[10] Qu T., Staley M. and Sun W., "Eye Gaze Tracking Technology Application in Mobile Computing Platform Input Mechanism Study", http://www.see.eng.osaka-u.ac.jp/seeit/icccbe2016/Proceedings/Full_Papers/198-276.pdf
[11] Groen M. and Noyes J, "Using eye tracking to evaluate usability of user interfaces: Is it warranted?", IFAC Proceedings Volumes, Vol. 43, Issue 13, pp. 489-493, 2010.
[12] Rivu S, Abdrabou Y, Mayer T, Pfeuffer K and Alt F, "GazeButton: Enhancing Buttons with Eye Gaze Interactions", 11th ACM Symposium on Eye Tracking Research and Applications, 2019.
[13] Parisay M, Poullis C and Oertel M.K., "FELiX: Fixation-based Eye Fatigue Load Index A Multi-factor Measure for Gaze-based Interactions", 13th International Conference on Human System Interaction, 2020.
[14] Manoharan S, "Study On Hermitian Graph Wavelets in Feature Detection", Journal of Soft Computing Paradigm (JSCP), Vol. 1, No. 01, pp. 24-32, 2019.
[15] Park, C.H. and P ark, H., "Fingerprint Classification using Fast Fourier Transform and Nonlinear Discriminant Analysis", Pattern Recognition, vol. 38, pp. 495–503, 2005.
[16] Yanikoglu, B. and Kholmatov, A., "On-line Signature Verification Using Fourier Descriptors", EURASIP Journal on Advances in Signal Processing, vol. 2009, pp. 1–13, 2009.
[17] Muthyalampalli, R., "Implementation of Fast Fourier Transform for Image Processing in DirectX 10", September 2009. http://software.intel.com/sites/ billboard/ article-archive/ fast-fourier-transform/.
[18] Fan, L., Wang, S., Wang, H. and Guo, T., "Singular Points Detection based on Zero-Pole Model in Fingerprint Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 929–940. 2008.
[19] Ammar, G.S., Calvetti, D., Gragg, W.B. and Reichel, L., "Polynomial zerofinders based on Szego polynomials", Journal of Computational and Applied Mathematics, vol. 127, pp. 1–16, 2001.
[20] Datta, B.N., "Matrix Equation, Matrix Polynomial and the Number of Zeros of a Polynomial Inside the Unit Circle. Linear and Multilinear Algebra", Vol. 9, pp. 63–68, 1980.
[21] Bistritz, Y., "Zero location with respect to Unit Circle of Discrete-Time Linear System Polynomials", IEEE, pp. 1131–1142. 1984.
[22] Sheela S.V. and Vijaya P.A., "Mapping Functions in Gaze Tracking, International Journal of Computer Applications", Vol. 26, No. 3, 2011.
[23] Kim K-N. and Ramakrishna R.S., "Vision based eye-gaze tracking for human computer interface", IEEE International Conference on Systems, Man, and Cybernetics, pp. 324-329, 1999.
[24] Zhu J. and Yang J, "Subpixel eye gaze tracking", IEEE International Conference on automatic face and gesture recognition, 124-129, 2002.
[25] Kiat L.C. and Ranganath S., "One-time calibration eye gaze detection system", IEEE International Conference on Image Processing, pp. 873-876, 2004.
[26] Weston Sewell and Oleg Komogortsev. "Real-Time Eye Gaze Tracking with an Unmodified Commodity Webcam Employing a Neural Network", ACM Conference on Human Factors in Computing Systems, pp. 1-12, 2010.
[27] Meena Y.K., Cecotti H., Wong-Lin K., Dutta A. and Prasad G., "Toward Optimization of Gaze-Controlled Human–Computer Interaction: Application

to Hindi Virtual Keyboard for Stroke Patients", IEEE Transactions on Neural Systems and Rehabiliation Engineering, Vol. 26, No. 4, pp. 911-922, 2018.
[28] Martinikorena I, Larumbe-Bergera A., Ariz M., Porta S., Cabeza R. and Villanueva A., "Low Cost Gaze Estimation: Knowledge-Based Solutions", IEEE Transactions on Image Processing, Vol. 29, pp. 2328-2343, 2020.
[29] Su D., Li Y-F and Chen H., "Cross-Validated Locally Polynomial Modeling for 2-D/3-D Gaze Tracking with Head-Worn Devices", IEEE Transactions on Industrial Informatics, Vol. 16, No. 1, pp. 510-521, 2020.
[30] Kang M-C, Yoo C-H, Uhm K-H, Lee D-H and Ko S-J., "A Robust Extrinsic Calibration Method for Non-Contact Gaze Tracking in the 3-D Space", IEEE Access, Vol. 6, pp. 48840-48849, 2018.
[31] Sungheetha A and Sharma R., "A Novel CapsNet based Image Reconstruction and Regression Analysis", Journal of Innovative Image Processing (JIIP), Vol. 02, No. 03, 156-164, 2020.
[32] Yiu Y.H., Aboulatta M., Raiser T., Ophey L., Flanagin V.L., Eulenburg P. and Ahmadi S-A, "DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning", Journal of Neuroscience Methods, Vol. 324, pp. 108307-318, 2019.
[33] Fischer T, Chang H.J. and Demiris Y., "RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments", European Conference on Computer Vision – ECCV, Lecture Notes in Computer Science, Vol 11214. Springer, pp. 1-19, 2018.
[34] Lemley J, Kar A, Drimbarean A. and Corcoran P., "Convolutional Neural Network Implementation for Eye-Gaze Estimation on Low-Quality Consumer Imaging Systems", IEEE Transactions on Consumer Electronics, Vol. 65, No. 2, pp. 179- 187, 2019.
[35] Xia Y, Yu H and Wang F-Y, "Accurate and Robust Eye Center Localization via Fully Convolutional Networks", IEEE/CAA Journal of Automata SINICA, Vol. 6, No. 5, pp. 1127-1138, 2019.
[36] Lian D, Hu L, Luo W, Xu Y, Duan L, Yu J. and Gao S., "Multiview Multitask Gaze Estimation with Deep Convolutional Neural Networks", IEEE Transactions on Neural Networks and Learning Systems, Vol. 30, No. 10, pp. 3010-3023, 2019.
[37] Sheela S.V. and Abhinand P., "Iris Detection for Gaze Tracking Using Video Frames", IEEE Advanced Computing Conference, 2015.
[38] Bradski, G. and Kaehler, A., "Learning OpenCV", 1st edn. O'Reilly Media Publisher, 2008.