



Image complexity analysis with scanpath identification using remote gaze estimation model

Mohsina Ishrat¹  • Pawanesh Abrol¹

Received: 16 February 2018 / Revised: 20 May 2020 / Accepted: 27 May 2020 /

Published online: 20 June 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Analysis of gaze points has been a vital tool for understanding varied human behavioral pattern and underlying psychological processing. Gaze points are analyzed generally in terms of two events of fixations and saccades that are collectively termed as scanpath. Scanpath could potentially establish correlation between visual scenery and human cognitive tendencies. Scanpath has been analyzed for different domains that include visual perception, usability, memory, visual search or low level attributes like color, illumination and edges in an image. Visual search is one prominent area that examines scanpath of subjects while a target object is searched in a given set of images. Visual search explores behavioral tendencies of subjects with respect to image complexity. Complexity of an image is governed by spatial, frequency and color information present in the image. Scanpath based image complexity analysis determines human visual behavior that could lead to development of interactive and intelligent systems. There are several sophisticated eye tracking devices and associated algorithms for recording and classification of scanpath. However, in the present scenario when the chances of viral infections (COVID-19) from known and unknown sources are high, it is very important that the contact less methods and models be designed. In addition, even though the devices acquire and process eye movement data with fair accuracy but are intrusive and costly. The objective of current research work is to establish the complexity of the given set of images while target objects are searched and to present analysis of gaze search pattern. To achieve these objectives a remote gaze estimation and analysis model has been proposed for scanpath identification and analysis. The model is an alternate option for gaze point tracking and scanpath analysis that is non intrusive and low cost. The gaze points are tracked remotely as against sophisticated wearable eye tracking devices available in the market. The model employs easily available softwares and hardware

✉ Mohsina Ishrat
mohsina.ishrat@gmail.com

Pawanesh Abrol
pawanesh.abrol@gmail.com

¹ Department of Computer Science & IT, University of Jammu (J&K), Jammu, India

devices. In the current work, complexity is derived on the basis of analysis of fixation and saccade gaze points. Based on the results generated by the proposed model, influence on subjects due to external stimuli is studied. The set of images chosen, act as external stimuli for the subjects during visual search. In order to statistically analyze scanpath for different subjects, certain scanpath parameters have been identified. The model maps and classifies eye movement gaze points into fixations and saccades and generates data for identified parameters. For eye detection and subsequent iris detection voila jones and circular hough transform (CHT) algorithms have been used. Identification by dispersion threshold (I-DT) is implemented for scanpath identification. The algorithms are customized for better iris and scanpath detection. Algorithms are developed for gaze screen mapping and classification of fixations and saccades. The experimentation has been carried on different subjects. Variations during visual search have been observed and analyzed. The present model requires no contact of human subject with any equipment including eye tracking devices, screen or computing devices.

Keywords Gaze · Fixation · Saccade · Voila Jones · CHT · Visual search

1 Introduction

Movement of gaze is a non verbal communication medium that could reveal information about human perception of visual scene. Eye gaze is a means of human behavioral analysis through fixations and saccades. A gaze point is a spatial positional value of eye movement on a screen [34], while as cluster of gaze points that approximately occupy a position on a screen within a threshold value for a certain time period is a fixation [55]. Saccades are rapid movement of gaze points between fixations [11]. It has been noticed that the gaze points may vary at the rate of 100 deg/s with spatial dispersion of half to one degree of visual angle and 100-200 ms of duration. Such cluster of gaze points may be classified as fixations. Furthermore, spatial dispersion along horizontal and vertical directions could also be fixed through the exploratory analysis of data. Gaze points not qualifying as fixations are designated as saccades [8, 55]. Fixation and saccades gaze points are collectively called as scanpath [36] and are the most commonly analyzed eye movement events. Figure 1 presents a sequence of processes that are required for scanpath analysis. The process of scanpath analysis requires detection of eye segment from the features present in the facial region. Detection of segment of eye is followed by mapping of location of eyes on the visual scene that the eyes are viewing at a particular instant of time i.e. gaze point. The process is termed as gaze screen mapping. Detection of gaze points is followed by identification of scanpath in terms of fixations and saccades. The detection of components of scanpath is dependent on focus of viewer on the visual scene. The more focused viewing would generate more fixations and lesser will have more saccades.

Eye detection is one of the facial feature detection essential for various research applications that include expression analysis, psychological and medical applications [21]. With the

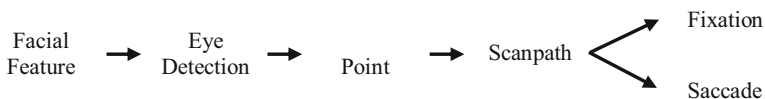


Fig. 1 Process of scanpath for gaze analysis

technological advancements remote gaze estimation has been increasingly investigated in comparison to head mounted equipments for capture of gaze. Many algorithms and techniques have been proposed for digital video analysis to estimate gaze for interactive systems, advertising, driving and analytic purposes [24, 31]. Remote gaze estimation has an obvious advantage of being non intrusive, cost effective and more comfortable as it does not require anything to be attached around eyes. The only necessary requirement is a capturing device at a certain distance from the subject and a gaze point processing hardware and software. However, the research study on scanpath indicates that scanpath analysis has mostly been carried out by sophisticated wearable eye tracking devices. The prominent eye tracking devices include Eyelink, SMI RED and Tobii [3, 16, 26, 27, 39–41, 49, 50, 53] that generate scanpath with fair accuracy but are non ubiquitous, intrusive and costly. The scanpath analysis using these eye trackers is a part of varied research and usability applications. Reading effect on eye movements [3, 49], impact of website viewing [16], individual perception of the visual environment [39, 41], marketing [50], driving [27] and visual search [26] are some of the application of eye movement analysis. Impact of bottom up features like color [41, 48], luminance [4] and sound [10] on eye movements has also been studied that point towards the expansion of research analysis around eye movement data. Amongst the application areas that employ head mounted eye trackers, visual search is prominent. During visual search, a given target object is to be searched among a group of distracters. The distraction is governed by presence of heterogeneity in structural content of image, variation in image chromaticity, image saliency and knowledge base of searcher [35, 52]. The presence of distracters directly effects complexity of an image. Complexity of an image is a multidimensional function of spatial, frequency and color variation present in an image. The presence of color, edges, intensity variation, quantity of objects and clutter defines amount of complexity present in an image [9]. Analysis of visual search with respect to complexity present in visual image has resulted in better understanding of human behavior. Analyzing search strategies while visualizing images of different complexities, can improve the understanding of underlying cognitive processing, which in turn leads to development of artificial intelligent systems and gaze based interactive systems [14, 58].

The objective of this paper is to establish scene complexity and common behavior of subjects based on scanpath analysis. The analysis of scanpath has been presented for fixation and saccades using an alternate remote gaze estimation and analysis model. It is cost effective with readily available hardware and software. The model presented is ubiquitous and non intrusive. The non intrusive systems that do not require physical touch by the humans assume very high significance in the present context. The operation of remote non intrusive systems rules out any chance of transmission of viral infection like COVID-19 through human or surface touch and also achieves the objective of social distancing (<https://www.mygov.in/covid-19>). The model performs mapping and classification of gaze points for the purpose of analysis and has been evaluated using an experiment of visual search. The gaze is captured remotely using image capturing device. External illumination which is mostly IR illumination has not been used that may in certain cases irritate the eyes [24, 34]. The authors have implemented different techniques and algorithms that include voila jones for eye detection and circular hough transform (CHT) for iris detection. Identification by dispersion threshold (I-DT) algorithm is implemented for scanpath identification. CHT and IDT are customized for effective iris and scanpath identification. Algorithms are developed for gaze screen mapping and fixation-saccade classification as will be discussed later in the paper. The results are based on scanpath data generated by a specified target search on two visual images by ten different subjects. The scene complexity of one image is designated higher than the other based on the results of analyzed parameters. The eye movement analysis is carried on the basis of different

selected parameters: number of fixations, maximum saccade length, total saccade length, maximum fixation duration, total fixation duration and total search time. The variations in terms of total time duration, number of fixations, saccade length have been analyzed. Many experiments of visual search have used measures of eye movement that include fixation duration, number of fixation and distribution of saccades, for analysis of eye movements [5, 17, 18, 60]. The parameters are thus chosen based on the study of analytical work of visual search [38, 58]. An analysis of gaze path traversed by two subjects among ten subjects is presented. Gaze path analysis points towards individual behavior of subjects and scanning pattern of subjects. The paper has been organized as follows. In Section 2 and Section 3 gaze movement analysis algorithms and proposed gaze analysis model have been discussed respectively. Section 4 discusses about results generated. Finally conclusion is presented in Section 5 followed by the references used in the paper.

2 Gaze analysis algorithms

In this section a discussion about different algorithms implemented for evaluation of scanpath is put forth. Many of standard algorithms like Voila Jones and CHT are implemented. The algorithms, however, are not sufficient for generation of values for identified parameters. A set of custom algorithms and specification of constraints are implemented for the generation of parametric data by the proposed model. A discussion on algorithms of Voila Jones, CHT and I-DT, custom algorithms and specification of constraints required for the proposed model is put forth.

2.1 Eye region detection

Many of the algorithms first identify facial region followed by detection of eye region and features around eyes [30, 62]. Remote eye detection methods are roughly classified into four groups: feature based, appearance based, template matching and hybrid methods [24, 44, 47]. Feature based methods identify different characteristics of eye region. Intensity, shape and gradient information is used to identify eyes in addition to iris, corner points, pupil shape, color and texture information [1, 59]. Appearance based methods learn structural information of the eye region [20, 56] while in template matching the template of eye region is correlated with the search area for detecting of eyes and hybrid methods use combination of shape, intensity, color and gradient for eye detection [12, 22, 25]. The appearance based techniques have been emerging as efficient approaches in detection of facial features. In the proposed approach voila jones has been implemented for detection of eye region, which is an appearance based object detection approach [15]. The algorithm can be trained for any object detection and tracking and is a robust and efficient method of eye detection as well. The algorithm works on image representation called ‘integral image’. The integral image facilitates faster calculation of statistical image features [29, 46]. The algorithm works by combining the discriminatory power of weak classifiers to develop a strong classifier. Weak classifiers identify similarity patterns based on rectangular features called ‘Haar features’. Samples that are wrongly classified in iteration have their weights increased in subsequent iterations, so that total sum of classification error is minimized. The classifier is of the form of.

$$F_M(x) = \sum_{m=1}^M E_m f_m(x) \quad (1)$$

Expanding Eq. (1)

$$F_M(x) = E_1f_1(x) + E_2f_2(x) + E_3f_3(x) + \dots + E_mf_m(x) \quad (2)$$

In Eq. (2) the linear combination of weak classifiers also called as learners generates a strong classifier. E_m denotes the set of coefficients such that sum of training error is minimized. Each of f_m is a weak learner that takes object 'x' as input and returns the class of input object. $F_M(x)$ is the final strong classifier response on object classification. The cascaded structure in voila jones accelerates average time of detection. In cascading negatives, i.e. the features that cannot contribute in the detection of eye region are rejected in the early stage of classification, since each weak classifier represents a stage. Only the features that pass through each of the classifier contribute in the final detection of eye region and are termed as positives [19, 46]. In the proposed model, once eye region is detected, the region is localized and segmented and then iris needs to be located for further analysis.

2.2 Region localization and segmentation

Since head movement is restricted in the experimental setup through chinrest, eye detection algorithm is not applied for each frame of the video. The implementation is required only when the subject's moves out of the detection range. Based on the output of eye detection algorithm, the region around eyes is localized and the algorithms in the pipeline track the localized region. The localized region is segmented for implementation of iris detection algorithm. Image segmentation is based on different parameters of region, clustering, edges, intensity and color [37, 45]. In this work, segmentation is region based dependant since a region of an image is isolated. The region segmentation is based upon the outcome of localization. The process of segmentation is very important for reduction in approximation error of eye region and iris centers for subsequent video frames after eye and iris is detected.

2.3 Iris detection

Detection of iris has been an important feature for various biometric applications [7, 61]. In the proposed work, iris coordinates are detected and tracked which are further classified into fixations and saccades. For iris detection, circular object detection algorithm, circular hough transform (CHT) has been implemented [33]. The basic rule is that line perpendicular to edge point of a circle cross center of the circle. The line segments follow a parametric form, i.e.

$$x = a + r * \cos\theta$$

$$y = b + r * \sin\theta$$

Where a, b are the coordinates of the center and r is the radius. The points (x,y) are the high probability circles present in the image for radius r along horizontal and vertical direction. Angular dimension θ ranges from 0 to 360 degrees. CHT uses 3D array with two dimensions representing coordinates of circle and one corresponding to radius [57]. Maintaining array for each of the pixel value takes a toll on performance of the algorithm. In the proposed model certain constraints have been introduced to detect seemingly circular objects, iris, using CHT. Algorithm of CHT is an efficient

algorithm for detection of circular objects but in the proposed model mere implementation of CHT is not efficient for iris detection. Two constraints (3) and (4) are introduced for accurate detection of iris.

$$rmin \leq r \leq rmax \quad (3)$$

$$(Tx_L, Ty_L) \leq |(x_L, y_L) - (x_R, y_R)| \leq (Tx_R, Ty_R) \quad (4)$$

Using constraint I (Eq. 3) the range of radius is fixed between minimum ($rmin$) and maximum ($rmax$) numeric value that a radius of an iris can have. The minimum value that a radius of a circle can have as detected by CHT, is fixed as $rmin$ and the upper numeric value that a CHT detected radius of a circle can have is fixed as $rmax$. The radius is fixed on exploration of the range of iris radius for the ten subjects chosen. Constraint II (Eq. 4) greatly enhances accuracy of iris detection, wherein in the minimum (Tx_L, Ty_L) and maximum (Tx_R, Ty_R) threshold values of distances along horizontal and vertical direction between iris centers of left and right eye is set. The difference between left eye center and right eye center along horizontal and vertical direction has to be between the values set in minimum (Tx_L, Ty_L) and maximum (Tx_R, Ty_R). The maximum and minimum value of coordinate difference between left and right iris center from the analysis of eye movement data defines threshold values for left iris and right iris center. Fig. 2 presents comparative difference that application of constraints I and II, have on efficient identification of iris. Figure 2a is the result on application of CHT on eye region. As shown in Fig. 2a, multiple small and large circles are detected instead of left and right iris center for left eye and right eye respectively. Figure 2b is the snapshot of resultant eye region on application of CHT and constraint I and II. The two frames of same eye region in Fig. 2a and 2b show the effectiveness of iris detection on application of CHT and CHT in addition to two constraints respectively.

On successful iris detection, the model is able to trace eye movements that can later be classified as fixations or saccades but before scanpath detection and later analysis, one of the important algorithmic processes is mapping of detected iris center position onto the location on screen where the subject is looking at an instant of time.

2.4 Gaze screen mapping

Many gaze estimation algorithms have been proposed by the researchers for gaze screen mapping. Most of the algorithms work on remote, non intrusive gaze estimation and tracking. Broadly, these algorithms are classified into three categories on the basis of feature, model structure and appearance [24]. Feature based gaze estimation methods rely

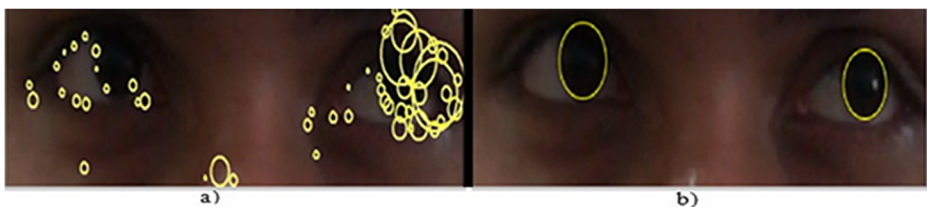


Fig. 2 Comparative iris detection accuracy. **a** without constraints using only CHT, **(b)** identification of iris after application of 2 step constraints and CHT

on local features peculiar to eyes. Pupil, iris, eye contours, corneal reflection, eye corner and iris coordinate relation, are some of the vectored features for gaze estimation [55, 63]. In contrast, model based methods use 3D geometric eye model to determine gaze location and are more robust for varied variation in facial region [54]. The appearance based, unlike other categories, learns mapping function from eye image directly onto the screen coordinates [51]. In the proposed approach, direct eye coordinate- screen mapping is calculated based on the values obtained during calibration phase. The method for gaze estimation is implemented since the screen is divided into sixteen equally spaced cells i.e. into 4*4 grid area. During calibration phase subjects are directed to look at four positions on screen, $g_{(x1,y1)}$, $g_{(x2,y1)}$, $g_{(x3,y1)}$, $g_{(x1,y2)}$ (Fig. 3). Once the value of iris coordinates are known, while looking at four positions, all other mapping values for twelve screen positions can be known by different combination of four known values. As for instance knowing value of $g_{(x1,y2)}$ and $g_{(x2,y1)}$ all the values in row second of screen grid can be known due to equally spaced cell positions. Based on the iris coordinate on these four positions along horizontal and vertical direction, the expected coordinate values for other twelve positions on screen are calculated and stored in a matrix called gaze mapping matrix (GMM) as shown in Fig. 3. As the iris coordinate values during subsequent image viewing for the purpose of target search are calculated the values are mapped and subtracted from each of the value in GMM. Based on the minimum value obtained from Eq. (5), each gaze position on screen $S_{(x,y)}$ is calculated and mapped. Equation (5) is the gaze mapping function for calculation of gaze points onto the screen.

$$S_{(X,Y)} = \min(g_{(x,y)} - [GMM]) \tag{5}$$

$S_{(x,y)}$ is the actual gaze position on screen during target object search. The $g_{(x,y)}$ values are the detected gaze points. GMM is an $n*n$ gaze mapping matrix. The underlying reason for Eq. (5) is that any value of $g_{(x,y)}$ can be mapped to the scene viewed based on the minimum difference of gaze point values with respect to those calculated in GMM.

The inherent advantage of dividing screen into 4* 4 grid is that calibration procedure is reduced from $n*n$ points to just n points and the remaining $((n*n)-n)$ points are generated from the gaze mapping matrix (Fig. 3). The method implemented for gaze screen mapping is simplistic in approach than any other approaches as implemented by different researchers. The simplicity is achieved since the screen visualized is divided into equal sized cells. However the method has a disadvantage of mapping detected iris coordinate values to one of the sixteen cell regions than onto

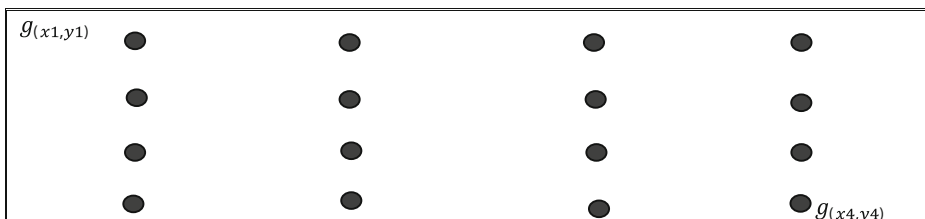


Fig. 3 Gaze mapping 4*4 screen matrix

a specific object in that cell region. The shortcoming has to be worked upon subsequently.

After the iris detection and gaze estimation, it is essential to identify and classify the fixations and the saccades.

2.5 Scanpath detection

The algorithms especially in software setup of eye tracking devices classify fixation and saccades based on spatial dispersion, velocity, acceleration difference and motion thresholds [28, 43]. Various common algorithms like I-DT, IVT, IHMM, IMST and IKF are available that parse eye movement events. One of the prominent algorithm implemented by researchers is I-DT, a dispersion threshold algorithm [2, 43], that classifies scanpath components based on threshold set on spatial dispersion and time duration. The algorithm exploits intrinsic property of fixation and saccade in terms of difference of time duration (100-200 ms for fixation and 80-150 ms for saccades) and spatial coverage ($\frac{1}{2}$ to 1 degree of visual angle for fixation and higher values for saccade). In the proposed model, constraints for spatial dispersion and temporal dimension similar to I-DT are implemented, as discussed in classification section below.

2.6 Classification

The classification of scanpath into fixations and saccades is achieved by exploration of fundamental difference between fixations and saccades. The difference is in velocity, spatial dispersion and temporal dimension. In the work presented, two threshold conditions on two properties associated with eye movements are set. The thresholds set are conceptually similar to constraints set by I-DT algorithm. The threshold is set on spatial region covered by the gaze points and for the time duration difference between gaze points.

2.7 Fixations

The model proposed identifies fixations by setting a threshold on spatial dispersion and temporal dimension of gaze points in the scanpath. Mathematically, if (x_i, y_i, t_i) is the triad of values at any point i , with iris center (x, y) at any instant of time ' t ', then for the set of gaze points to be a fixations, the following condition must hold:

$$|x_b - x_i| \leq d_t \quad (6)$$

$$|y_b - y_i| \leq d_t \quad (7)$$

$$t_i \geq t_c \tag{8}$$

(x_b, y_b) are the chosen horizontal, vertical gaze coordinate values that had satisfied conditions (6–8). (x_i, y_i) pair represents current gaze point that needs to be checked for inclusion in the fixation cluster. ' d_i' ' is the threshold set on spatial region coverage that should hold, while verifying Eqs. (6) and (7). The value t_c is the upper cap of time duration for identification of fixation. The gaze point satisfying all three conditions is designated as fixation. The centroid for each of the fixation is calculated as:

$$x_{min} = \min(x_b, x_i) \tag{9}$$

$$x_{max} = \max(x_b, x_i) \tag{10}$$

$$y_{min} = \min(y_b, y_i) \tag{11}$$

$$y_{max} = \max(y_b, y_i) \tag{12}$$

$$f_i = \left\{ \frac{x_{min} + x_{max}}{2}, \frac{y_{min} + y_{max}}{2} \right\} \tag{13}$$

Here, f_i is the fixation centroid for each of the gaze points cluster labeled as fixation. Equations (9) and (11) find minimum and maximum gaze value along horizontal direction among all the gaze points in each identified fixations respectively. Similarly Eqs. (10) and (12) find minimum and maximum coordinate values from a range that are termed as single fixation along vertical directions. The obtained minimum and maximum values along horizontal and vertical direction are summed and then divided by two to obtain central point of each fixation amidst the range of coordinate positions obtained for a particular fixation that have value less than specified spatial and temporal threshold [2]. The constraint is applied on all gaze points to isolate fixations.

2.8 Saccade

The random gaze points in a scanpath that do not fixate at a particular position during visual search are saccades. Based on the definition of saccades, the gaze points that do not satisfy conditions (6)–(8) are classified as saccade. A matrix based on the definition of saccades is

Table 1 Saccade matrix representation

F_1	S_1F_1	S_2F_1	S_3F_1	...	S_mF_1	F_2
F_2	S_1F_2	S_2F_2	...	S_mF_2	[0]	F_3
⋮						
F_{n-1}	S_1F_{n-1}	...	S_mF_{n-1}	[0]	[0]	F_n
F_n	S_1F_n	...				

created for storage of saccades i.e. gaze point (x, y) coordinates that occur between fixation gaze points [2, 11, 36, 55]. The value of saccades at any instant, is a matrix where the dimension are dependent on number of gaze points from fixation F_i to F_{i+1} for number of columns, whereas number of rows in the matrix are dependent on total number of fixations present in the eye movement data. The number of rows should be equal to $(1 - N_f)$ for F_n fixations, but the gaze points after n th fixations are not lost and treated as saccades, as is indicated in the last entry of Table 1. N_f is the total number of fixations.

At any instant the matrix might look of the form as shown in Table 1. The matrix records each of the positional value between any two fixation positions. The matrix represents only one of the possible instances between F_i to F_n fixations in the eye movement data. Each of the $S_i F_i$ entry represents spatial saccade values from fixation F_i to F_{i+1} . The column dimension is equal to the highest number of saccadic positions traversed between the fixations. The fixations with lesser saccadic length have [0] as entries in the $m [i][j]$, where m is the saccade matrix and i, j are the number of fixations and saccade length indices respectively. With classification of each of the scanpath component, different vectors holding the values of identified parameters are generated. The selected parameters are number of fixations (nf), total saccade length (tsl), maximum saccade length (msl), maximum fixation duration (mfd), total fixation duration (tfd) and total time (tt).

3 Proposed gaze analysis model

The objective of the current research work is to analyze scanpath of subjects while performing visual search. The motive is to establish scene complexity based on analysis of scanpath components. The analysis is performed using the proposed model based on identified scanpath parameters. The analysis of eye movement and gaze points in the preliminary stage requires capturing of eye movements while a scene is being viewed.

Each of the frames of video is separately processed in a sequence. Eye region is detected from each frame using eye detection algorithm of voila jones. The region is localized and segmented with the help of information about region coordinates acquired during implementation of voila jones. From the detected eye region, iris center coordinates are detected using constraint bound CHT as discussed already. The process is applied across the video frames. During the calibration phase, the positional value of four positions is known by directing each of the subject's eye position on to them. The rest of the positions are derived from the coordinate values of four captured positions. While the subjects view stimulus, the eye coordinate values are mapped onto the screen positional values to derive gaze positions. The classification algorithm is implemented to differentiate between fixation and saccades. During the analysis process, the positional values and number of fixations and saccades are stored in a vector based approach and accordingly, duration of fixation and total time of search are calculated based on the values stored. Figure 4 provides a summarized view of the process sequence undertaken. The proposed approach has been divided into different processing steps. The frames of the captured video are read through conditional indexing. Each of the frames is an input that is to be processed by subsequent processes. For the initial frames, eye detection algorithm is implemented and once the eye detection algorithm returns success, the region is localized and segmented for further processing. Detection of eyes patch is not enough for generation of gaze points and subsequently scanpath and hence iris detection algorithm is implemented for detection of centre of eyes that are then mapped on to screen as a part of 'gaze

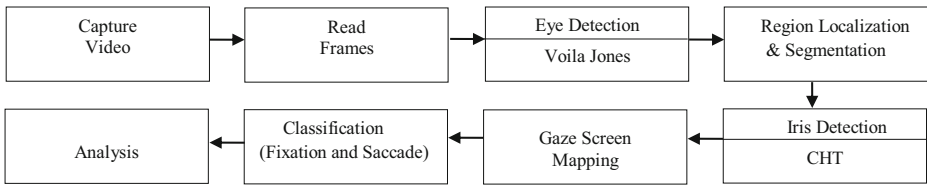


Fig. 4 Steps of the proposed eye movement analysis model

screen mapping’ procedure. The algorithms are discussed already in Section 2. The next process is to classify gaze points into eye movement event of fixation and saccades. For both the events, vectors are generated to hold values of selected parameters.

A digital camera (SONY NEX-5) with resolution of 4592* 3056 pixels, 14.2 megapixels with frame rate of 25fps is placed on a tripod at a height of 120 cm. The subjects are seated at a distance of 50 cm from the camera, with a chinrest at a height of 110 cm. The optimal person camera distance is stated to be between 40 and 70 cm [23, 32]. In the present experimental setup, fair resolution of video frames is obtained at 50 cm. The screen size is 158*252 cm and it is at a distance of 250 cm from seating position of a subject as shown in Fig. 5. Images are displayed on the screen through projector connected to the computing device. The first image to be displayed is for calibration purpose for gaze screen mapping. The stimulus image 1 is displayed followed by image 2 and subjects viewing the stimuli are directed to search for specified objects in the respective images designated as $s1$ and $s2$. The dimensions of $s1$ is 893*669 pixels and of $s2$ is 561*373 pixels which is roughly around 23.62*17.70 cm² and 14.84*9.86 cm² respectively. The images are resized to 25.4*19.05cm² as power point slides and displayed as 158*252 cm² images via projector on the screen. The first slide of the presentation is displayed for the calibration purpose. The calibration area has been divided into a 4*4 grid having a total of 16 calibration points. Typically, calibration points range between one to 25 points, with optimum being nine calibration points. However, it has been observed that the increase of calibration does not increase the accuracy of gaze tracking systems [30].

The stimuli images $s1$ and $s2$ are shown in Fig. 6, first the calibration slide is displayed where the subjects had to fixate at some of the positions in the image for duration of two to three seconds while the video is being recorded by the camera.

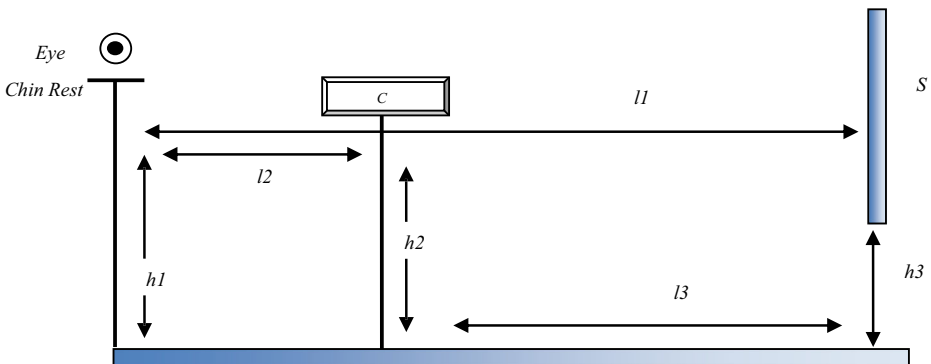


Fig. 5 Layout of experimental setup of proposed research work. *Gaze distance ($l1$) = 250 cm, * person camera distance ($l2$) = 50 cm, *screen- camera distance ($l3$) = 200 cm. * $h1$ = 110, * $h2$ = 120 cm, * $h3$ = 100 cm. *S-screen(158*252 cm²)

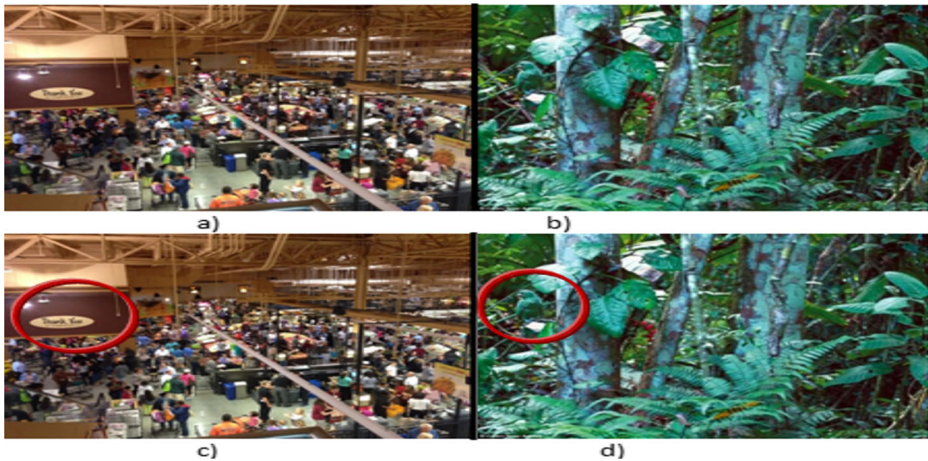


Fig. 6 Image as displayed to subjects in a sequence. **a** first stimulus to find ‘Thank You’ board in the image. **c** second stimulus to find ‘parrot’ in the all green image. **c** and **d** the red circled regions indicate objects to be found

Due to gaze mapping strategy discussed earlier, the calibration procedure is considerably shortened. Once the calibration step is completed, the subjects are directed to close their eyes. During the period when eyes are closed, $s1$ with ‘Thank You’ board in it, is displayed on screen. Each of the subjects has to search for the board in the image. Once the search is completed, the eye closure procedure is repeated till $s2$ is displayed on screen. In $s2$, the subjects have to search for the ‘parrot’ in the image. The end of visual search is marked by the closure of eyes and serves as checkpoints during video analysis for captured videos. Both the searched objects reside in cell 5 of the calibrated slide (Fig. 3). The system requires no physical touch between subjects and each of the subject remains at a distance of approximately 250cms from the screen.

As mentioned above, for the analysis, a set of parameters have been identified. The parameters identified for analysis are number of fixations (nf), total saccade length (tsl), maximum saccade length (msl), maximum fixation duration (mfd), total fixation duration (tfd) and total time (tt). For each of the parameter a vector based approach is implemented, wherein each fixation is recorded during each of the stimulus viewing. The total length of the vector is same as the total number of fixations. From each of the row entry in the saccade matrix, saccade length can be deduced. Fixation duration is obtained while timing each of the fixations. Total time taken is equal to total number of gaze coordinates recorded during each stimulus viewing.

The next segment focuses on the evaluation of experimental results generated after a visual search is carried out in the specified setup with the implementation of the proposed model.

4 Results and analysis

In this section, a discussion on scanpath components and scene complexity in terms of six identified parameters is presented. Gaze path traversed by two subjects during $s1$ and $s2$ visual search is analyzed. Complexity of a scene is considered as a decisive factor affecting eye movement events in visual search [5]. Complexity or difficulty of a scene is inversely

proportional to the distinguishable objects present in the scene image [58]. The authors based on visual inspection, hypothesize that variation in color is seemingly less in *s2*. Less color variation seems to be a factor contributing to image complexity of *s2* in comparison to *s1*. Therefore, identification of object like ‘parrot’ might get difficult and less conspicuous in second visualized image (*s2*). It has been reported by different researchers that number of fixation increases and fixation duration decreases during difficult scene processing. The increased fixation duration convey more objects are analyzed during that duration. Less complex scenes allow more number of objects to be identified in a single fixation [42, 58]. Also the spatial saccadic distribution is biased toward objects looking similar to the target object that is to be found [5, 58].

The present experiment evaluates fixations and saccades on six identified parameters as mentioned in Section 3. In Table 2, the values derived for each of the subjects for all the parameters is tabulated. Number of fixations (*nf*) indicates total of fixations recorded while subjects search for given targets in *s1* and *s2*. The results indicate *nf* for *s2* is mostly higher than *s1*. Only one of the subjects (*subject5*) highlighted in the table entries of *nf* show higher number of fixations for *s1*. The reason for variation in behavior of one of the subjects need further exploration as authors cannot be conclusive for one odd behavior with lesser target searches. Further *subject 5* has shown exception in behavior for other chosen parameters as well that needs further exploratory study. The fixation with maximum time duration (*mfd*) among a set of fixations during each target search is obtained. The parameter *mfd* is observed for all the subjects. Maximum duration of fixation (*mfd*) is exhibited by majority of subjects during *s1* object search than in *s2* object search. The parameter *mfd* is majorly higher for *s1* than *s2*, as only three subjects among ten subjects have value otherwise, as is highlighted in the table. The gaze coordinates for each saccade movement between fixations is recorded. Among the saccade paths traversed between fixations, maximum of all the saccade paths recorded between fixations is designated as maximum saccade length (*mssl*). The parameter *mssl* is calculated during both object searches for all subjects. During the analysis of maximum saccade length (*mssl*), nothing conclusive can be inferred as the results generated for the parameter have less difference of values for *s1* and *s2*. Another parameter total time (*tt*) is the recorded showing total duration of time spent by subjects during target search in *s1* and *s2*. The value for total time parameter (*tt*) is more during target object search in *s2* than in *s1*,

Table 2 Value of selected parameters namely number of fixations (*nf*), maximum fixation duration (*mfd*), maximum saccade length (*mssl*), total saccade length (*tsl*), total fixation duration (*tfd*), total time (*tt*) during scanpath analysis for two images *s1* and *s2* employed for visual search

Subjects	s1						s2					
	nf	mfd (ms)	mssl	tsl	tfd (sec)	tt (sec)	nf	mfd (ms)	mssl	tsl	tfd (sec)	tt (sec)
Subject1	28	1520	4	36	13	22	36	1000	19	56	15	25
Subject2	02	2320	4	05	05	07	05	1960	04	10	04	29
Subject3	04	2160	2	05	05	07	04	2920	01	04	07	07
Subject4	03	2320	1	03	04	04	07	1120	02	08	04	05
Subject5	12	1120	4	21	06	10	10	760	03	17	04	08
Subject6	03	4520	1	03	06	06	05	560	01	05	02	04
Subject7	12	800	6	18	07	10	19	1480	05	27	10	12
Subject8	07	1760	1	07	06	06	08	1360	02	09	05	07
Subject9	08	560	3	12	03	04	08	800	02	10	03	04
Subject10	05	1000	4	06	03	09	09	960	03	08	05	13

except for two subjects. The variation for two subjects can be attributed to the understanding of the experiment by the subjects. The parameter total saccade length (*tsl*) specifies total of all the gaze points generated during saccadic movement by each of the subjects while target object is searched in *s1* and *s2*. It has been observed that *tsl* for *s2* is generally more than *s1*. However, three subjects among ten show lower value for target object search in *s2*. The results obtained could indicate longer saccade length for difficult scenes in majority of cases. The total fixation duration (*tfd*) parameter is the total time duration of all the fixations during both visual searches. It is found that although *mfd* is higher for first visual search, the same cannot be concluded for *tfd*. The results for *tfd* do not give out any definite results, as the values obtained are divided within the subjects without any clear majority.

It has been already established that increase in number of fixations is an indication of lesser number of identifiable objects in a scene. Since objects are not easily identifiable in an image, fixation duration decreases and saccade movement increases as put forth by different researchers. In the present work, scanpath analysis based on different parameters indicated *nf* being majorly higher for *s2* than *s1*. The parameter *mfd* generally showed lesser values in *s2* than in *s1*. The data analysis showed mostly longer *tsl* for *s2* than *s1*. Having higher values for *nf*, *tsl* and *tt* for *s2* target search than for *s1* and also lower values of *mfd* for *s2* target object search, indicates that *s2* is more complex during visual search than *s1*. The results of two parameters *tfd* and *mstl* have been inconclusive and thereby need further analysis. Also, in the present work, two images are considered as stimulus conditions to understand human visual behavior during object search and put forth the resultant findings. The authors suggest increasing number of images for visual search and number of subjects as a part of future work. Increase in number of subjects and number of stimuli images would present a better understanding of behavioral pattern of subjects in terms of different parameters. The resultant value of the parameters is an indicator for similarity in behavioral pattern observed between subjects in terms of target search.

Figure 7 presents a plot of gaze coordinates while subjects search for target objects in *s1* and *s2*. The graph represents 4 s of searching by *subject3* and *subject4*. The subjects are chosen as there are no incorrect gaze coordinate detections in the video clips of the respective subjects. The gaze points plotted are of the left eye movement of subjects as the left eye movements render least incorrect detections. The reason for incorrect detections has been majorly due to tilted placement of camera due to which right eye of subjects is not correctly captured at many instances. The horizontal axis ‘time frame’ signifies the frames of the captured video as the gaze coordinate values are recorded. The vertical axis plots the gaze coordinates during target object search in *s1* and *s2* by *subject3* and *subject4*. Figure 7a and c depict horizontal movement of gaze during *s1* and *s2* object search for *subject3* and *subject4* respectively. Figure 7b and d show vertical movement of gaze during search in *s1* and *s2* for *subject3* and *subject4* respectively. The variations in the vertical axis values for *subject3* and *subject4* are primarily due to seating position of subjects while on chinrest. A slight shift on chinrest changes the reference coordinate values especially along horizontal direction. Each of the ten videos captured are processed separately through a process explained in Sections 2 and 3. It is not necessarily required to have similar frame of reference for analysis of scanpath.

It has been observed that variation along vertical direction is less while searching for target objects in *s1* and *s2* as shown in Fig. 7b, d. In contrast the search pattern along horizontal direction for *subject3* and *subject4* shows more variation during visual search in *s1* and *s2* (Fig. 7a, c). Although the objects to be searched in *s1* and *s2* are placed at almost same spot, cell 5 of calibration setup (Fig. 3), yet the path traversed by the subjects are different during the two

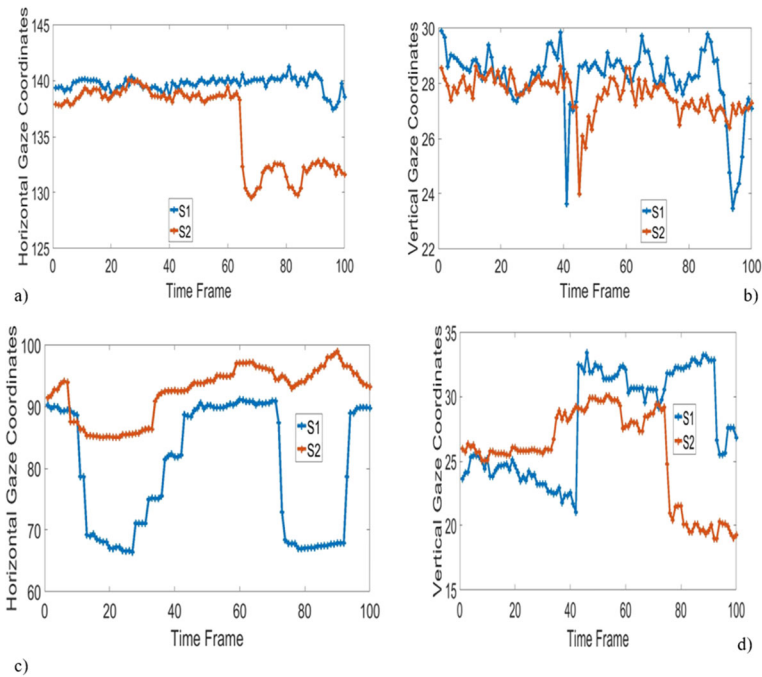


Fig. 7 Gaze path (4 s) along horizontal and vertical direction for two subjects. **a** Horizontal direction gaze positions for subject3 during *s1* and *s2* visual search. **b** Vertical direction gaze positions for subject3 during *s1* and *s2* visual search. **c** Gaze positions for subject4 during *s1* and *s2* visual search along horizontal direction. **d** Gaze positions for subject3 during *s1* and *s2* visual search along vertical direction

object searches. The difference in searching pattern is also observed between the subjects. The gaze movement of *subject3* shows less variation along horizontal direction during *s1* object search and is almost similar during *s2* object search for a certain period of time after which the searching pattern shows divergence. Although the difference in variation along vertical direction for *s1* and *s2* is less prominent (*subject3*), individually the variation is more prominent along vertical direction in comparison to horizontal direction variation. For *subject4* variation during *s1* visual search is higher along vertical and horizontal direction and in contrast *subject3* exhibits less variation of gaze movement along horizontal direction during *s1* and *s2* object search. The difference in movement along vertical direction for *s1* and *s2* is more prominent unlike *subject3* gaze movements.

The analysis of gaze movement for *subject3* and *subject4* suggests that movement of eye gaze is more inclined towards scanning upper to lower areas of visual area (vertical movement) than searching left to right areas of screen (horizontal movement). The contrast of gaze movements between subjects points towards individualistic behavior of subjects while searching for target objects in a given scene. The difference in movement of gaze during *s1* and *s2* object search is more prominent for *subject4*, while *subject3* exhibits less difference in gaze movements. The results do indicate variation of searching pattern between subjects, variation within same subject on two different stimuli and percentage variation in movements along vertical and horizontal direction, still the authors believe in increasing the amount and variation in data before any conclusion could be drawn.

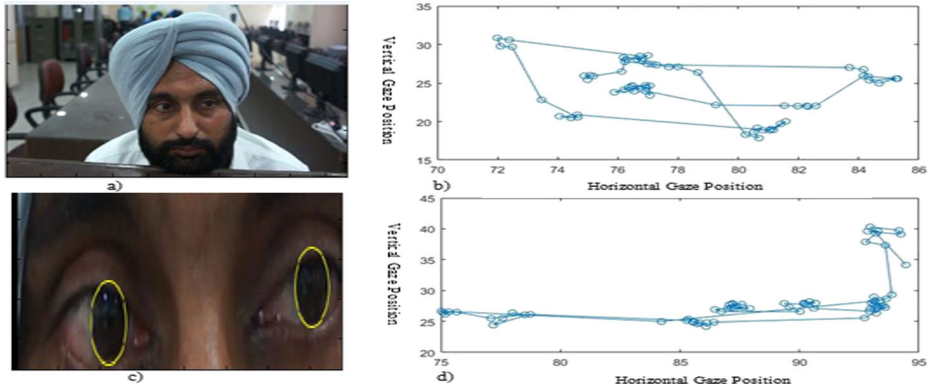


Fig. 8 Interface for eye movement analysis. **a** Video capturing during visual search) Gaze search path during *s1*. **c** Detection of iris. **d** Gaze path traced during *s2* object search

Figure 8 presents a snapshot of interface created for gaze estimation and scanpath analysis. Figure 8d, d shows gaze path traversed by one of the subjects while searching for the target objects in *s1* and *s2* respectively. The interface is divided into four windows. In the first window subject is displayed, the second window focuses on capturing the movement of eyes. The next two subjects window display gaze path traversal during *s1* object search (Fig. 8b) and gaze path traversal during *s2* object search (Fig. 8d). The snapshot is of one of subjects while the captured video is processed. The approach presented is able to detect the scanpath traversed by subjects during visual search.

The image complexity in the present work is established by parametric analysis of the identified parameters. On evaluation of parameters identified, one can infer the complexity of the image visualized for target object search. The basis of this conclusion is based on the results of previous research analysis of the parameters. The property that the parameters *nf*, *mst*, *tsl*, *tfid*, *mfd* and *tt* should exhibit if one image is complex than the other is reviewed and correlated with the current findings. The parameters chosen also indicate towards the commonality of behaviors between subjects while performing visual search. The results indicate most of the subjects displayed similar trend for the parameters that enabled in conclusion of common behavior in terms of parametric data. However, the analysis of gaze path traversal by two subjects points towards individualistic behavior that needs further exploration. The analysis of scanpath data based on different parameters and understanding of gaze path traversal could be useful for development of interactive tools and behavioral analysis studies.

5 Conclusion

In this paper, scanpath analysis is carried out to understand complexity of given set of images. Image complexity is multidimensional in terms of spatial, frequency and color properties. An experimental evaluation on one of the scanpath analysis application i.e. visual search is presented. The evaluation is based on different scanpath parameters i.e. number of fixation, fixation duration, saccade length and total search time for each of the object search. Although results are obtained for each parameter suggesting scene complexity effecting eye movements but the outcome is based only on two images acting as stimuli. Also the evaluation of gaze path traversed points towards individualistic behavior of subjects and needs further analysis.

The experimental evaluation is to be further extended by including more stimulus conditions and subjects in the future work.

The gaze path and scanpath analysis is presented using remote gaze estimation and analysis model. Different standard and custom algorithms are implemented. An algorithm for gaze mapping and scanpath classification is presented. Adding constraints to iris detection algorithm, the accuracy of the algorithm has been considerably enhanced. Gaze mapping function has been able to reduce calibration points from 16 sixteen to only four points. The classification of saccade and fixation has been based on the spatial and temporal thresholds. The eye movement data obtained is low considering the eye movement data captured by any standard eye tracker. The approach has also not considered impact of background lighting, subject movement and any sort of noise. With a camera at 25fps only, there could be a definite loss of eye movement data, however, there are number of budget friendly cameras with higher fps, which could be used instead for better data capturing. Visual search is only one eye movement analysis application area, the same approach can be applied for any field of eye movement analysis. The work presented experimentally evaluates a low cost, remote based and non intrusive approach for eye movement analysis. The model presented effectively detects eyes, iris and estimates gaze before classifying scanpath into fixation and saccades for the purpose of scanpath analysis.

In the present, times when epidemics like COVID-19 that transmits rapidly among the population through touch or via close contact, systems operating remotely are the need of the hour. Models of the similar pattern as are presented in this present paper could encourage further design of contactless, interactive and analytical systems.

References

1. Abbasi M, Khosravi R (2019) A robust and accurate particle filter-based pupil detection method for big datasets of eye video. *J Grid Comput* 18:1–21
2. Andersson R, Larsson L, Holmqvist K, Stridh M, Nyström M (2017) One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behav Res Methods* 49:616–637
3. Argilés M, Cardona G, Pérez-Cabrè E, Rodríguez M (2015) Blink rate and incomplete blinks in six different controlled hard-copy and electronic reading conditions. *Invest Ophthalmol Vis Sci* 56:6679–6685
4. Benedetto S, Carbone A, Drai-Zerbib V, Pedrotti M, Baccino T (2014) Effects of luminance and illuminance on visual fatigue and arousal during digital reading. *Comp Human Behav* 41:112–119
5. Bertera JH, Rayner K (2000) Eye movements and the span of the effective visual stimulus in visual search. *Percept Psychophys* 62:576–585
6. Bignaut PJ, Wium DJ (2013) The effect of mapping function on the accuracy of a video-based eye tracker. *Proceedings of the Eye Tracking South Africa Conference (ETSA)* 29-31
7. Bowyer KW, Hollingsworth K, Flynn PJ (2008) Image understanding for iris biometrics: a survey. *Comput Vis Image Underst* 110(2):281–307
8. Çöltekin A, Fabrikant SI, Lacayo M (2010) Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *Int J Geogr Inf Sci* 24:1559–1575
9. Corchs SE, Ciocca G, Bricolo E, Gasparini F (2016) Predicting complexity perception of real world images. *PLoS One* 11(6):1–22
10. Coutot A, Guyader N, Ionescu G, Caplier A (2012) Influence of soundtrack on eye movements during video exploration. *J Eye Mov Res* 5(4):1–10
11. Coutot A, Hsiao JH, Chan AB (2018) Scanpath modeling and classification with hidden Markov models. *Behav Res Methods* 50:362–379
12. Cristinacce D, Cootes TF (2006) Feature detection and tracking with constrained local models. In *Bmvc* 1: 1–10
13. Duchowski AT (2003) *Eye tracking methodology: theory and practice*. Springer, New York
14. Eckstein MP (2011) Visual search: a retrospective. *J Vis* 11:1–36

15. El Kaddouhi S, Saaidi A, Abarkam M (2017) Eye detection based on the Viola-Jones method and corners points. *Multimed Tools Appl* 76:23077–23097
16. Eraslan S, Yesilada Y, Harper S (2016) Scanpath trend analysis on web pages: clustering eye tracking scanpaths. *ACM Trans Web (TWEB)* 10:1–35
17. Findlay JM (2004) Eye scanning and visual search. In: Henderson JM, Ferreira F (eds) *The interface of language, vision, and action: eye movements and the visual world*. Psychology Press, New York, pp 135–159
18. Findlay JM, Gilchrist ID (1998) Eye guidance and visual search. In: *Eye guidance in reading, driving and scene perception* pp. Elsevier, Oxford, pp 295–312
19. Friedman J, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann Stat* 28(2):337–407
20. Ghazali KH, Jadin MS, Jie M, Xiao R (2015) Novel automatic eye detection and tracking algorithm. *Opt Lasers Eng* 67:49–56
21. Gou C, Wu Y, Wang K, Wang K, Wang FY, Ji Q (2017) A joint cascaded framework for simultaneous eye detection and eye state estimation. *Pattern Recogn* 67:23–31
22. Grauman K, Betke M, Gips J, Bradski GR (2001) Communication via eye blinks-detection and duration analysis in real time. In *CVPR, Proceedings of the IEEE Computer Society Conference* 1:1-1 IEEE
23. Guestrin ED, Eizenman M (2007) Remote point-of-gaze estimation with free head movements requiring a single-point calibration. In *29th Annual International Conference, EMBS, IEEE* 4556–4560
24. Hansen DW, Ji Q (2010) In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Trans Pattern Anal Mach Intell* 32:478–500
25. Hansen DW, Hansen JP, Nielsen M, Johansen AS, Stegmann MB (2002) Eye typing using Markov and active appearance models. In *Applications of computer vision, proceedings, sixth IEEE workshop* 132-136, IEEE
26. Hayes TR, Henderson JM (2019) Scene semantics involuntarily guide attention during visual search. *Psychon Bull Rev* 26:1683–1689
27. He S, Liang B, Pan G, Wang F, Cui L (2017) Influence of dynamic highway tunnel lighting environment on driving safety based on eye movement parameters of the driver. *Tunn Undergr Space Technol* 67:52–60
28. Ishrat M, Abrol P (2016) Effect of external stimuli on eye movement for eye gaze applications. *Int J Comput Syst (IJCS)* 3:3587–3591
29. Ishrat M, Abrol P (2017) Eye movement analysis in the context of external stimuli effect. In *International Conference on informatics, Health & Technology (ICIHT)* (pp. 1-6). IEEE
30. Kim H, Jo J, Toh KA, Kim J (2017) Eye detection in a facial image under pose variation based on multi-scale iris shape feature. *Image Vis Comput* 57:147–164
31. Laddi A, Prakash NR (2019) Eye gaze tracking based directional control interface for interactive applications. *Multimed Tools Appl* 78:31215–31230
32. Liang K, Chahir Y, Molina M, Tijus C, Jouen F (2013) Appearance-based gaze tracking with spectral clustering and semi-supervised gaussian process regression. In *Proceedings of the 2013 Conference on Eye Tracking South Africa* 17-23, ACM
33. Liu H, Qian Y, Lin S (2010) Detecting persons using Hough circle transform in SurveillanceVideo. *VISAPP* 2:267–270
34. Morimoto CH, Mimica MRM (2005) Eye gaze tracking techniques for interactive applications. *Comput Vis Image Underst* 98:4–24
35. Nagy AL, Thomas G (2003) Distractor heterogeneity, attention, and color in visual search. *Vis Res* 43: 1541–1552
36. Noton D, Stark L (1971) Scanpaths in eye movements during pattern perception. *Science* 171:308–3011
37. Pal NR, Pal SK (1993) A review on image segmentation techniques. *Pattern Recogn* 26:1277–1294
38. Pomplun M (2006) Saccadic selectivity in complex visual search displays. *Vis Res* 46:1886–1900
39. Rubo M, Gamer M (2018) Social content and emotional valence modulate gaze fixations in dynamic scenes. *Sci Rep* 8:1–11
40. SensoMotoric Instruments GmbH (2015) <http://www.smivision.com/>. Accessed Nov 2017
41. Shahrbabaki ST (2015) Contribution of colour in guiding visual attention and in a computational model of visual salience. Dissertation, Université Grenoble Alpes
42. Shen J, Reingold EM, Pomplun M (2000) Distractor ratio influences patterns of eye movements during visual search. *Perception* 29:241–250
43. Smeets JBJ, Hooge ITC (2003) Nature of variability in saccades. *J Neurophysiol* 90:12–20
44. Song F, Tan X, Chen S, Zhou ZH (2013) A literature survey on robust and efficient eye localization in real-life scenarios. *Pattern Recogn* 46:3157–3173
45. Viergever MA, Maintz JA, Klein S, Murphy K, Staring M, Pluim JP (2016) A survey of medical image registration—under review. *Med Image Anal* 33:140–144

46. Viola P, Michael JJ (2001) Rapid object detection using a boosted Cascade of simple features. In Proc IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1:511–518
47. Wan M, Li M, Yang G, Gai S, Jin Z (2014) Feature extraction using two-dimensional maximum embedding difference. *Inf Sci* 274:55–69
48. Wang D, Li J, An X (2016) What's the role color play in saliency detection. In Computer and Communications (ICCC), 2nd IEEE International Conference 657–661
49. Wang Y, Zhong X, Zhang Y, Tu Y, Wang L, Chen Y, Zhang C, Zhou W (2017) Visual fatigue following long-term visual display terminal work under different light sources. *Light Res Technol* 49:1034–1051
50. Wedel M, Pieters R (2017) A review of eye-tracking research in marketing. In review of marketing research. Routledge, Abingdon, pp 123–147 Please check provided publisher's location if appropriate. →
51. Williams O, Blake A, Cipolla R (2006) Sparse and semi-supervised visual mapping with the δ 3gp. In Proc. CVPR, New York, pp. 230–237
52. Wolfe JM, Horowitz TS (2017) Five factors that guide attention in visual search. *Nat Hum Behav* 1:1–8
53. Xue J, Quan C, Li C, Yue J, Zhang C (2017) A crucial temporal accuracy test of combining EEG and Tobii eye tracker. *Medicine* 96(13):1–11
54. Yamazoe H, Utsumi A, Yonezawa T, Abe S (2008) Remote gaze estimation with a single camera based on facial-feature tracking without special calibration actions. In Proc. ETRA, Savannah, pp. 245–250
55. Yarbus AL (1967) Eye movements and vision. Plenum, New York
56. Yi D, Lei Z, Li SZ (2011) A robust eye localization method for low quality face images. In Biometrics (IJCB), International Joint Conference 1–6, IEEE
57. Yip RK, Tam PK, Leung DN (1992) Modification of Hough transform for circles and ellipses detection using a 2-dimensional array. *Pattern Recogn* 25:1007–1022
58. Young AH, Hulleman J (2013) Eye movements reveal how task difficulty moulds visual search. *J Exp Psychol Hum Percept Perform* 39:168–190
59. Yu M, Tang X, Lin Y, Schmidt D, Wang X, Guo Y, Liang B (2018) An eye detection method based on convolutional neural networks and support vector machines. *Intell Data Anal* 22:345–362
60. Zelinsky GJ, Sheinberg DL (1997) Eye movements during parallel-serial visual search. *J Exp Psychol Hum Percept Perform* 23:244–262
61. Zhang M, He Z, Zhang H, Tan T, Sun Z (2019) Toward practical remote iris recognition: a boosting based framework. *Neurocomputing* 330:238–252
62. Zhu Z, Ji Q (2005) Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Comput Vis Image Underst* 98:124–154
63. Zhu J, Yang J (2002) Subpixel eye gaze tracking. In Automatic face and gesture recognition 5th IEEE international conference, pp. 131–136, IEEE

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Mohsina Ishrat is a Research Scholar in the Department of Computer Science and IT, University of Jammu. She has completed her master's in computer applications from University of Kashmir. She has been working in the area of Human computer interactions. She has presented papers in different Science Conferences.



Pawanesh Abrol Dr. Pawanesh Abrol is Professor at the Department of Computer Science and IT, University of Jammu, Jammu. He has done his PhD in Computer Science from University of Jammu. He also holds the degree of MBA (HRM). Dr. Abrol has more than twenty years of teaching and research experience at post graduate level. He has received INSA visiting Fellowship in 2007 to visit IIT Kanpur for research. Dr. Abrol has more than fifty research publications in various reputed national and international proceedings and journals. His research interests include aura based texture analysis, image forgery & authentication and eye gaze technologies. Dr. Abrol also remained Head, Department of Remote Sensing & GIS, University of Jammu. He has been the member of various committees and academic bodies. Presently, Dr. Abrol is holding the charge of Director, Reasi Campus, University of Jammu at Reasi. Dr. Abrol is Fellow of Institution of Engineers (IE), Kolkatta and Institution of Electronics and Telecommunication Engineers (IETE) New Delhi.