



International Journal of Control Theory and Applications

ISSN : 0974-5572

© International Science Press

Volume 9 • Number 43 • 2016

Mind Reading by Analysing Eye Gaze with Webcam Using Fuzzy Logic

Ann Veena Paul^a, Linda Sara Mathew^b and Surekha Mariam Varghese^c

^{a-c}Department of Computer Science and Engineering, M.A College of Engineering, Kothamangalam, Kerala, India. Email: ^aveena.ann@gmail.com; ^blindasaramathew@gmail.com; ^csurekha.var@gmail.com

Abstract: Mind Reading is the act of recognizing the information that is being processed in the user's mind by observing his actions. Researchers have been utilizing eye tracking to study behavior in areas like image scanning, driving, and reading and so on. Medical research supports the fact that there exists a correlation between human brain and eye movement. So the information processed in the brain can be derived by observing a person's gaze. Eye gaze patterns are recorded while the user is looking at a stimulus during eye tracking process. For recording eye movements an off-the-shelf, low-cost web camera in desktops and other smart devices is used. Estimation of gaze location from webcam is not so accurate because of the low resolution of images. To increase accuracy in classification of eye gaze direction a probabilistic neural network is employed. Fixation and saccade are then identified using the Duration-Threshold Identification algorithm. Fuzzy logic based Inference (FIS) system analyses the sequence of fixations and saccades in eye movements to extract the user intentions. To identify the structure of the FIS, eye gaze data is clustered into different clusters using Subtractive Clustering algorithm. The output of FIS will be the score representing the User Interest Level (UIL) on every image that appears on the screen. Thus valuable information regarding user's point of attention can be predicted.

Keywords: Eye gaze tracking, EAC, eye tracking, Gaze direction estimation, Fixation, Saccades, Subtractive Clustering, TSK Fuzzy Inference System.

1. INTRODUCTION

Eyes are the index to one's mind and soul. Hence eye movements provide a rich and informative window into a person's thought and intentions. The study of eye movements helps to determine what people are thinking based on where they are looking. Eye tracking is a technique whereby the position of the eye is used to determine the gaze direction of a person at a given time and also the sequence in which they have moved. Eye-Tracking research has been gaining popularity in Human-Computer Interaction (HCI) and is regarded as a research tool to analyse human behaviour as well. Eye gaze is more natural way to interact with computers than other currently available input devices. While many eye and gaze tracking systems have been developed, many suffer from undesirable drawbacks and restrictions.

Human gaze estimation is a rapidly developing field of research with many useful applications. It is especially useful for applications in human-computer interaction (HCI), such as video conferencing, where the object of the user's focus is enhanced or transmitted in higher quality, and eye-typing, where a human subject selects letters to be typed by looking at a screen display. Eye tracking is useful in a wide range of applications like neuroscience, psychological research, human-computer interactions, medical diagnostic, usability studies, gaze-controlled applications and so on.

Gaze tracking is being used for monitoring driver alertness, where computer system monitors the driver's eyes to ensure that a driver is not drowsy. As the accuracy of gaze tracking and estimation systems improved, this way of interacting with the computer screen may be provided as a substitute means for controlling a computer for those physically or mentally impaired persons who find it difficult to control a mouse.

Human eyes and its movement patterns contains rich information about human cognitive processes, emotions, diseases etc. The nature of eye movements is unique for each. The direction of the eye gaze can express the interests of the user; it is a potential source to locate the current cognitive processes. The patterns in which the eyes move when humans access their memories is known as eye accessing cues (EAC). Eye Accessing Cues gives information about mental process from the direction of eye gaze. The direction of the iris in the socket can give information regarding various cognitive process.

This study is focused on the use of real-time data from human eye gaze movements that can be explored in the area of human-computer interaction. The aim is to provide an assistive technology for professionals who wish to gain deeper insight into gaze interaction. Here we try to estimate the user's attention in front of the computer by observing his eye movements.

2. RELATED WORKS

The human-computer interaction by using the eye gaze as computer input has been studied by several researchers. As the purpose of applications differ, the approaches for realizing eye-computer interaction differ. Monitoring the brain waves by electroencephalography (EEG) [1] and studying users' eye movements (Eye Gaze Tracking) [2] are two main approaches for Human Computer interaction currently under research. Studying eye movements has the direct benefit of demonstrating that, where a user looks reveals what one is looking for. Eye gaze information has a fast response speed compared to other senses. Furthermore eye tracking is less intrusive than EEG. Therefore, gaze tracking is considered as an effective method of interaction between the user and the computer. For the estimation of a user's attention in front of a computer monitor, two key issues to be considered are the identification of directionality and movement patterns of eye gaze. Different methods are used for tracking eye gaze. They include 1. Electro oculography method [3] uses sensors that are attached at the skin around the eyes to measure an electric field that exists when eyes rotate. 2. The infrared oculography method [4] that measuring intensity of infrared light reflected by the sclera. 3. Video Oculography method [5] which is video-based eye tracking work by illuminating the eye with an infrared light source. This light produces a glint on the cornea of the eye and this is used as the reference point for gaze estimation.

For gaze tracking there is Feature-based method that explore the characteristics of the human eye and use them for gaze estimation. *Jixu Chen* proposed a model-based approaches [6] use an explicit geometric model of the eye to estimate 3D gaze direction vector. Interpolation-based approaches methods assume the mapping from image features to gaze co-ordinates. *Kar-Han Tan* proposed appearance-based methods [7] detect and track eyes directly based on the photometric appearance. It use image content to estimate gaze direction by mapping image data to screen coordinates.

The process of fixation and saccade identification is an essential part of eye-movement data analysis. Velocity-based algorithms [8] emphasize the velocity information in the eye-tracking protocols, taking advantage of the fact that fixation points have low velocities and saccade points have high velocities. Dispersion-based algorithms emphasize the dispersion of fixation points, under the assumption that fixation points generally occur near one another. Area-based algorithms identify points within given areas of interest (AOIs) that represent relevant visual targets.

Understanding the different possible identification algorithms will help to facilitate the use and analysis of eye tracking data for both scientific research and the development of eye-based systems and applications.

3. PROPOSED SYSTEM

The overall framework of the proposed system is shown in Figure 1. Different stages of the proposed framework are described below.

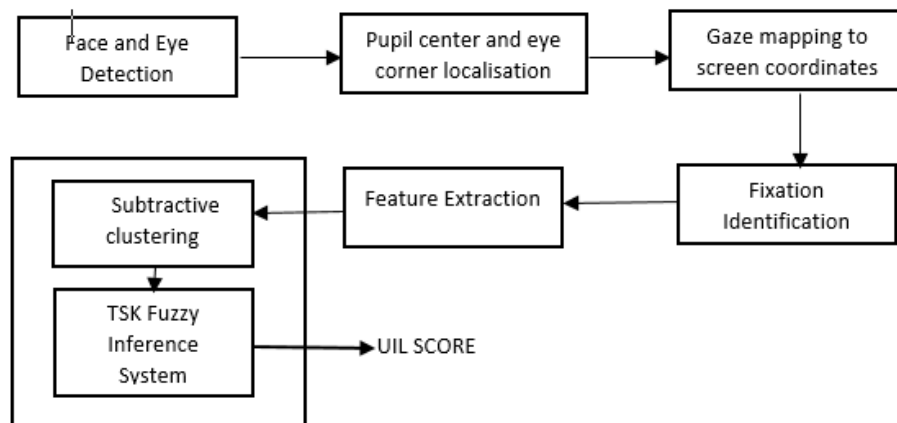


Figure 1: Block diagram of the developed framework

A. Face Detection and Eye Region Localisation

The knowledge of position of face in an image is an essential factor in determining the point of gaze. Detection and tracking of face help in obtaining candidate regions for eye detection. Viola-Jones algorithm [9] is used for face detection as this is fast and invariant to in-plane rotation, with high accuracy. Once face region is localised next stage is to obtain the eye region. The eye region for classification is obtained from the face bounding box returned by the face detector (ROI). The eye regions obtained are rescaled to 42×50 dimensions. Histogram equalisation is applied on each of the channel on ROI to enhance the contrast differences. The resultant image has brightened sclera and darker eye boundary. This is then converted to grayscale image.

B. Eye Gaze Direction Recognition

When a subject gazes at different target areas on the interactive interface, we could collect and extract gaze feature information using a web camera and then achieve gaze direction estimation.

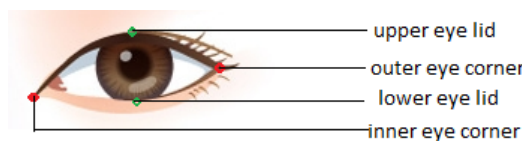


Figure 2: Eye model with all corners labelled

1. **Pupil Center Identification:** The circular Hough transform [10] is employed to deduce the radius and center coordinates of the pupil and iris regions. Radius intervals are defined for pupil circles. Starting from upper left corner of the eyes the circular Hough transform is applied. The votes are calculated in the Hough space for the parameters of circles passing through each edge point. The parameters that have a maximum value corresponds to the centre coordinates.

Hough transform for the pupil center localization uses the basic principal of mathematical expression of circle:

$$(x - a)^2 + (y - b)^2 = r^2 \quad (1)$$

Where (a, b) is the center of the circle and r is the radius of the circle.

Steps to determine the center coordinates (a, b) and the radius r is given below:

Step 1: Create a three-dimensional array (A, B, R) to count the number of points (x, y) corresponding to each circle. A, B is the height and width of the image, R is the smaller value of $A/2$ and $B/2$.

Step 2: Traverse the image (x, y) take each point (x, y) which possibly belong to the circle equation (a, b, r) store in the array (A, B, R)

Step 3: Find the maximum value in the array (A, B, R) , (a, b, r) which is corresponding to the pupil circle equation.

2. **Eye Corner Identification:** The contract between sclera and iris area is used as the basis for eye corner extraction. It uses simple pixel searching algorithm. The eye region image is converted to grey scale image of selected intensity ranges. Some pre-processing like dilation, morphological erode operations and 2D median filtering and noise reduction is applied on it. The edges of this filtered image is then found using Canny edge operator. Then algorithm searches for first and last white pixels in xy axis and they are assumed as eye corners and upper and lower eyelids.
3. **Gaze Mapping:** After finding pupil center and eye corners, the gaze vector can be computed, by calculating horizontal and vertical location differences between the two points. Eye corners are considered as anchor points, as they do not move relative to the face. Gaze coordinates are represented as:

$$\bar{g} = |P_{\text{center}}P_{\text{corner}}| = (X, Y) \quad (2)$$

$$X = |X_{\text{center}} - X_{\text{corner}}|$$

$$Y = |Y_{\text{center}} - Y_{\text{corner}}|$$

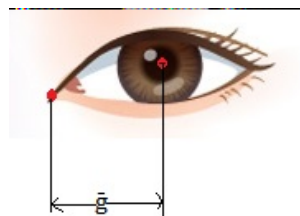


Figure 3: Eye moving trends

As shown in the Figure 3, \bar{g} stands for the distance between inner eye corner and iris center. When gaze direction turns to the left, iris moves to left relative to the eye socket and the position of left or right corner is fixed.

In order to know where a gaze vector is directed, calibration is necessary. A set of points are shown to users, and for each of the point user has to move eyes towards them. Thus we get a set of points on screen and their corresponding eye vectors. The mapping between the eye gaze coordinates to screen coordinates can be represented using the equation:

$$\begin{aligned} X_{\text{screen}} &= \alpha_0 + \alpha_1 X + \alpha_2 Y + \alpha_3 XY \\ Y_{\text{screen}} &= \beta_0 + \beta_1 X + \beta_2 Y + \beta_3 XY \end{aligned} \quad (3)$$

By solving the above equations we can find the best fitting parameters for calculating the screen coordinates given the eye gaze vector. These parameters can be applied to all following eye gaze vectors to determine the screen coordinates.

C. Fixation Identification

Fixation [11] refers to stabilisation of eyes with respect to a target item on screen. The duration of fixation directly relates to a subject's interest because of increased visual attention during fixations.

A saccade is a quick, simultaneous movement of both eyes between two or more phases of fixation in the same direction. A duration threshold algorithm was used for fixation identification. I-VT was chosen for its simplicity, speed and small number of parameters. Identify the screen regions that are mapped by each of the frames. If subsequent frames are mapped to same screen regions for the min Threshold (200 ms) time then mark it as Fixation. All other points are marked as Saccades. Threshold is chosen based on the fact that minimum time of 100-200ms is required to for an eyes to acquire information from a point he sees.

4. MIND READING USING FUZZY INFERENCE SYSTEM

Mind reading is a task of recognising the intensions of the user by using the features extracted from tracking his eye gaze. Fuzzy logic is used to infer the intension by analysing the eye gaze data. For estimating the user's interest eye gaze vector, fixation, time for each fixation and revisit to each fixation area are the essential elements and these are given as input to the TSK based fuzzy inference system developed.

- A. **Fuzzy Signature:** Fuzzy Signatures [12] are fuzzy descriptors of real world objects. They represent the objects with the help of sets of available quantities expressing interdependencies among the quantities of each set. The fuzzy signature is an effective approach to solve the problem of rule explosion in traditional fuzzy inference systems.
- B. **Fuzzy Signature Construction for Eye Gaze Pattern Recognition:** People tend to concentrate their gaze into interesting and informative regions. Original gaze data are further processed into fixations, time spent on each fixation, frequency of a fixation as it is much easier to interpret them. The user's gaze has a longer time of staying and higher frequency appearing in the region which has aroused user's interest.
- C. **Fuzzy Inference System:** Fuzzy Inference System (FIS) is a way of mapping from a given input to an output using fuzzy logic. By using fuzzy rules from expert knowledge the interest level of user on each image is computed. A fuzzy logic based systems is constructed based on the feature vector extracted. The input to it will be the feature vector and output is determined as a score indicating the user's interest level on each image. TSK-FIS is the most commonly used FIS for system modelling. This type of FIS uses the lingual form and mathematical functions for the premise and consequent parts of fuzzy rules of fuzzy system. This structure of TSK-FIS makes it interpretable both for human and machine and adopts flexibility to developed systems.

In order to estimate user's attention to each of the image, a Sugeno Type fuzzy Inference System was built. The metrics used in our case is the gaze vector, time of fixation on the image and the frequency of visit on each of the image.

1. **Takagi-Sugeno-Kang Fuzzy Inference System (TSK-FIS):** It is a fuzzy inference system proposed by Takagi, Sugeno and Kang (TSK) [13] and is the most commonly used system for fuzzy modelling and control. The basic idea of this method is to decompose the input space into fuzzy regions and to approximate the system in every region by a simple model.

A Typical, TSK model consist of IF-Then rules that have the form:

Ri: IF x_1 is $A_{i,1}$ and x_2 is $A_{i,2}$... x_r is $A_{i,r}$

then,
$$Y_i = B_{i,0} + B_{i,1}x_1 + B_{i,2}x_2 + \dots + B_{i,r}x_r \text{ for } i = 1, 2, \dots, L \tag{4}$$

where, L is the number of rules, x_i are the input variables, y_i are local output variables. $A_{i,1}$ are fuzzy sets that are characterized by membership functions $A_{i,1}(x_j)$ and b_{ij} are real valued parameters. TS-FIS consists of inputs, outputs, set of predefined rules and a defuzzification method. Aggregation is employed to combine the outputs of all the rules into a single fuzzy set. The final output of the system is the weighted average of all the rule partial outputs. The overall output of the model is computed by:

$$Y = \frac{\sum_{i=1}^L T_i Y_i}{\sum_{i=1}^L T_i} = \frac{\sum_{i=1}^L T_i (b_{i,0} + b_{i,1}x_1 + \dots + b_{i,r}x_r)}{\sum_{i=1}^L T_i} \tag{5}$$

where, T_i is the firing strength of rule T_i which is defined as:

$$T_i = A_{i,1}(x_1) \times A_{i,1}(x_1) \times \dots \times A_{i,r}(x_r) \tag{6}$$

A. TSK-FIS Structure Identification

Normally to develop a FIS, human experts with the knowledge about inputs and outputs of the system. The model of learning is a 3 step process: first, we gather the data which will be used for learning, then create a learning model based on the data we gathered, and finally adapt a learning model and deploy it. Selection of inputs have high influence on the prediction model.

To reduce the role of human factors in the development process of the FIS's, some algorithms are introduced which automise the process of identifying the FIS structure by Fuzzy Clustering. The idea of fuzzy clustering is to divide the data space into fuzzy clusters, each representing one specific part of the system behavior. After projecting the clusters onto the input space, the antecedent parts of the fuzzy rules can be found. The consequent parts of the rules can then be simple functions. In this way, one cluster corresponds to one rule of the TSK model. Thus by clustering it is possible to generate rules for each cluster and develop corresponding membership function. Fuzzy clustering is done using input-output data. To generate the MF's for the input sets and find the number of rules for our TSK-FIS, subtractive Clustering is used.

1. **Subtractive clustering:** This algorithm was proposed by Chiu [14] and is an improvement of the mountain clustering technique. Subtractive clustering does not depend on initial conditions hence it is consistent in the outputs. It is a one pass algorithm for estimating the number of clusters and the cluster centers in a set of data.

Each data point is considered as a potential cluster center. A measure of potential is associated to each point according to its neighborhood, defined by radius r_a . The point with highest potential, P_1^* is selected to be

the first cluster center. Once the center is selected potential of each pair is decreased according to its distance to x_1^* . The process is repeated until potential is below a threshold.

The steps for subtractive clustering is given below:

Step 1: Consider a collection of n data points, $X = \{x_1, x_2, \dots, x_n\}$

Step 2: Calculate the potential value for each data point to be cluster center

$$P_i = \sum_{j=1}^n e^{-\alpha \|x_i - x_j\|^2}, \alpha = \frac{4}{r_a^2}, r_a \text{ is cluster radius, } \| \quad \| \text{ is Euclidian distance}$$

Step 3: Select data point (X_i^*) with highest potential, P_i^*

Step 4: Revise the potential of each point for next iteration,

$$P_i = P_i - P_i^* e^{-\beta \|x_i - x_j\|^2}, \beta = \frac{4}{r_b^2}, r_b = \eta \times r_a, \eta \text{ is squash factor} = 1.25$$

Step 5: Repeat the steps until potential of all points is below ϵP_i , ϵ is reject ratio = 0.5

Each input attribute and output attribute has as many membership functions as the number of clusters that subtractive clustering algorithm has identified. The number of rules is equal to number of clusters. The number of clusters created by the algorithm determines the optimum number of fuzzy rules.

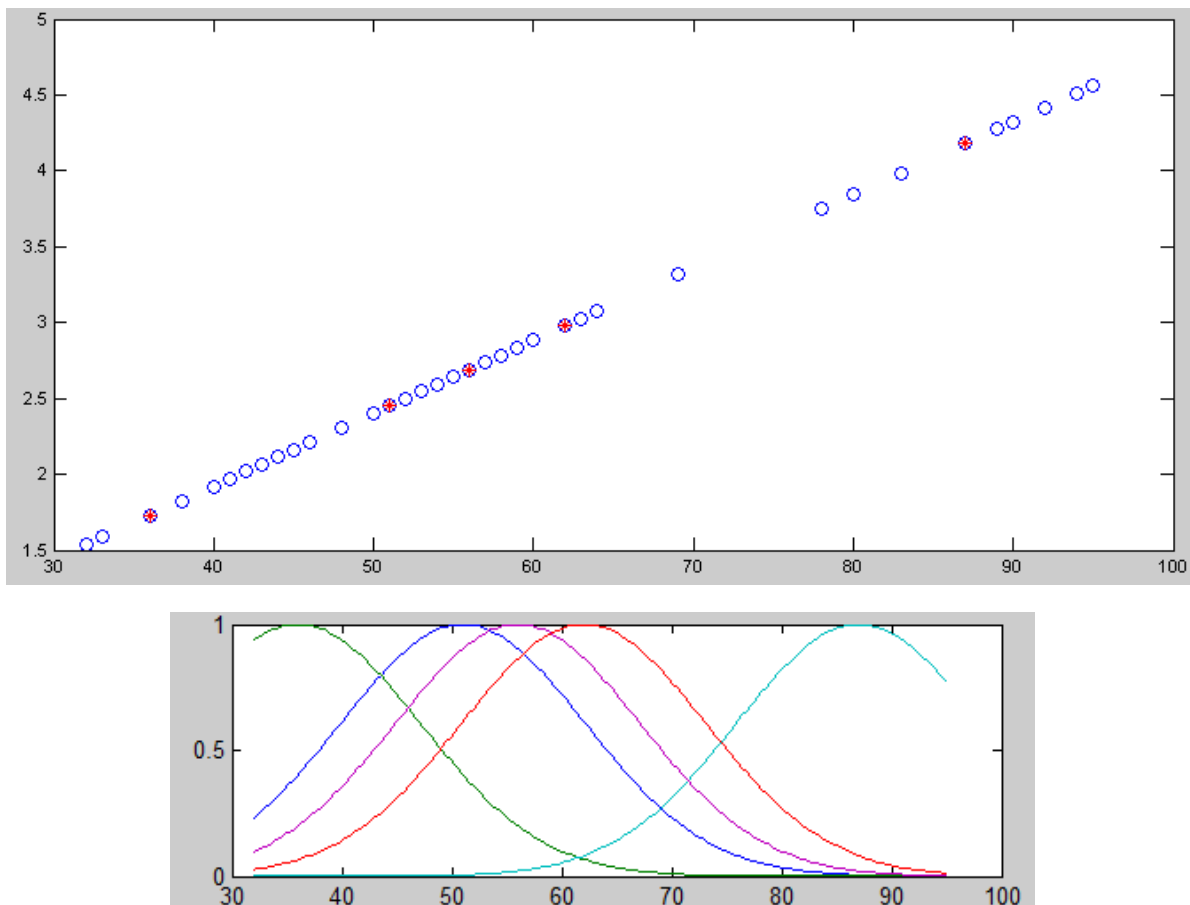


Figure 4: The subtractive clustering. (a) Data pairs and clusters. (b) X_a variable membership functions

The fuzzy inference system has the ability to analyse the gaze movement data. The frequency of gaze appears (F) and the time of gaze stays (T) is the input to the fuzzy system. The system then generates a score for each of the image presented before user. This score gives information about which is the most interesting image for the user. The score is used to predict the interest level user has on each and every image presented to him.

5. CONCLUSION

This paper investigated a system for tracking and analyzing user's eye gaze to evaluate the user's interest in front of the computer monitor was presented. Gaze tracking is used to determine where on the screen a user's attention is directed. The proposed system tracks and records the user's gaze data with a web camera. TSK-Fuzzy inference system constructed analyses the gaze data and generates a score for every observed image. The structure of the TSK-Fuzzy System is identified using Subtractive Clustering. This score tells valuable information regarding user's interest to the presented images. The results of Fuzzy inference system showed that the user's interest is not a crisp value, but it depends on different biometrics. The advantages of the method is that it does not require any special setup in terms of hardware, it runs in real time, it is user independent and can work efficiently under any lighting conditions. Furthermore, it is possible to improve contents and the user experience using these information. Therefore, eye tracking technology can be applied to various fields such as mind reading, driver attention analysis, advertisement effectiveness monitoring, medical research and so on.

REFERENCES

- [1] A.Yazdani, J.-S.Lee, and T.Ebrahimi, "Implicit emotional tagging of multimedia using EEG signals and brain computer interface "in Proc. 1st SIGMM Workshop Social Media, New York, 2009, pp. 81–88, ser. WSM'09, ACM.
- [2] Hutchinson T.E., White K.P., Martin W.N., Reichert K.C., Frey L.A., Human-Computer Interaction Using Eye-Gaze Input. IEEE Transactions on Systems, Man, and Cybernetics, 1989.
- [3] A B Usakli, S Gurkan, E Aloise, G Vecchiato and E Babliloni, "On the Use of Electroculogram for Efficient Human Computer Interfaces", Hindawi Publishing corporation, Computational Intelligence and Neuroscience, pp 1-5, Volume 2010.
- [4] M. W. Johns, A. Tucker, J. R. Chapman, E. K. Crowley, and N. Michael, Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers, Vol. 11, pp. 234–242, 2007.
- [5] Yang X-H, Sun J-D, Liu J., Li X-C, Yang C-X and Liu W., (2012) A Remote Gaze Tracking System Using Gray-Distribution-Based Video Processing, Journal of Biomedical Engg: Aplns., Basis & Communications.
- [6] Jixu Chen, Yan Tong, Wayne Gray, Qiang Ji, A robust 3D eye gaze tracking system using noise reduction, Proceedings of the 2008 symposium on Eye tracking research & applications, March 26-28, 2008, Savannah, Georgia
- [7] Kar Han Tan, Narendra Ahuja, Appearance-based Eye Gaze Estimation, Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, p.191, December 03-04, 2002
- [8] M. A., & Carpenter, P. A. (1984). Using eye fixations to study reading comprehension. New Methods in Reading Comprehension Research.
- [9] M. Kawulok and J. Szymanek, "Precise multi-level face detector for advanced analysis of facial images," Image Processing, IET, Vol. 6, No. 2, pp. 95–103, 2012.
- [10] Ballard DH, "Generalizing the Hough transform to detect arbitrary shapes", In the proceedings of Pattern Recognition, pp. 111–122, 1981
- [11] Salvucci, D. D., Goldberg, J. H.: Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the 2000 Symposium on Eye Tracking Research and Applications, 71–78, NY: ACM Press (2000)

- [12] L. T. K'oczy, T. V'amos, and G. Biro, "Fuzzy signatures," in Proceedings of the 4th Meeting of the Euro Working Group on Fuzzy Sets and the 2nd International Conference on Soft and Intelligent Computing(EUROPUSE-SIC'99), pp. 210–217, Budapest, Hungary, May 1999.
- [13] C.Wagner and H.Hagras, "Toward general type-2 fuzzy logic systems based on zslices,"IEEETrans.FuzzySyst., Vol. 18, No. 4, pp. 637–660, 2010.
- [14] S. L. Chiu, "Fuzzy model identification based on cluster estimation," J. Intell. Fuzzy Syst., Vol. 2, pp. 267–278, 1994

