Eye Movement based Human Recognition using Scan Path Signals

Vivek Srivastava
Department of Computer Science & Engineering
Rama University, Uttar Pradesh, Kanpur, India
viveksrivastavakash@gmail.com

***Abstract*—This paper is devoted for human recognition using scan path signals in eye movement analysis. Velocity and dispersion threshold based fixation identification algorithms are employed for processing the raw scan path signals in eye movement matrices. A new hybrid intelligent model is deployed for classification over data retrieved from scan-path signals. Experimental results demonstrate the endeavor of eye movement signals as an effective biometric trait. This paper also demonstrates the relative comparison of the two fixation identification techniques combined with hybrid intelligent model.**

Keywords- Eye movement, scan-path, fixation, saccade, hybrid intelligence, biometrics.

I. INTRODUCTION

Human recognition is an important biometric application useful in the areas of security, law enforcement, criminal justice and other related commercial applications. There are a wide range of biometric traits such as face, iris, ears, gait, retina etc. which are being used for biometric recognition in recent past. However, some of these traits are easier to reproduce and causes a gap in identification accuracy. Therefore it is necessary that biometric traits should not be easy to reproduce. Eye movements characteristics depend on the brain activity and extra-ocular muscles properties that make it highly counterfeit resistant. Therefore, it is not possible to accurately replicate the eye movement of any individual. The researches in eye movement as biometric trait is at very initial stage and require more attention to work with advance computational intelligence techniques besides the conventional statistical techniques used in [1], [2], [6] for recognition. In this paper, we use a new hybrid intelligent model (HIM) for classification which combines evolutionary fuzzy clustering (EFC) and syndicate neural networks (SNN). For feature extraction from the raw scan path signals, we adopt velocity threshold (I-VT) and dispersion threshold (I-DT) identification algorithms [3]. Schematic diagram of the proposed methodology is presented in fig. 1 which involves three basic processes: fixation identification from raw eye motion signals; derivation of eye movement matrices from fixations / saccades groups and finally classification scheme which employs hybrid intelligence model.

On the other hand, Hybrid intelligence approaches are being popular for various applications [4] [7] because of their combined learning capability. The proposed HIM consists of two components: cluster allocation and adaptation. Cluster allocation component is responsible for allocation of training data into number of clusters generated by evolutionary fuzzy clustering (EFC), which decides the basis for network structure selection in adaptation component. Adaptation performs training & subsequent generalization using an ensemble of neural networks. The EFC is used to make the neural network (NN) system more efficient in terms of incorporating better flexibility in defining different classes. It performs pre-classification task and following this distribution NN works as final classifier. The EFC is also responsible for searching the optimal number of clusters which decides the number of NNs in ensemble. EFC yields better clustering over conventional fuzzy clustering as it provides optimal partition matrix. Moreover, this novel fusion yields a new hybrid intelligent model (HIM) which is quite efficient and performance rich in comparison to fuzzy clustering or conventional neural networks or their combination, especially in solving many complex pattern recognition problems having overlapping, inexact and ill-defined boundaries.

II. PRE-RPOCESSING

In order to find biometric feature matrices, we process the eye movement scan-path signals. Basically, scan-path refers to the special path formed by fixations and saccades. Saccades occur when eye rotates quickly between fixations and fixation occurs when eye is held in relatively stable position on an object [2]. It is necessary to derive fixations and saccades from raw scan-path signals for extracting the specific biometric features. Raw scan path signals include time stamp, the reading of scan-paths in horizontal and in vertical direction, stimuli in horizontal and vertical directions with data validity flag 1 for valid and 0 for invalid data [2]. In order to process the raw signals and to extract distinguishable features, we adopt two well known fixation identification algorithms, namely velocity threshold identification (I-VT) and dispersion threshold identification (I-DT). Basically fixation identification algorithms are employed to show the statistical description of eye motion behavior [3]. Fixation identification is the crucial aspect of the eye motion analysis because poor fixation identification scheme can lead to too many or very less fixation groups which can further affect the eye movement analysis. I-VT separates fixation and saccade points based on their point to point velocities [3]. I-DT employs a window and duration, dispersion threshold for obtaining saccades and fixations.

Eye movement raw signals can be represented by Z and described as following:

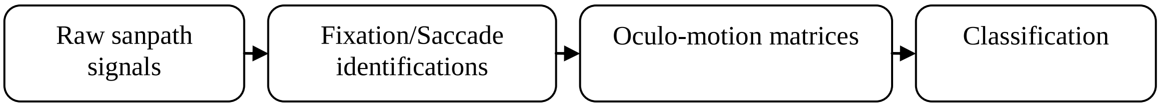

Figure 1. Schematic diagram of proposed methodology

$$Z=\begin{bmatrix} t_1, x_1, & y_1 \\ t_2, x_2, & y_2 \\ ::::::::: & :: \\ t_i, x_i, & y_i \\ ::::::::: & :: \end{bmatrix}$$

Where t_i is the time stamp for sequential eye position, x_i is the horizontal eye position in the degrees of visual angle and y_i is the vertical eye position in the degrees of visual angle. For derivation of feature vector, we have selected only three attributes t_i, x_i and y_i from raw scan path signals in Z.

In I-VT fixation algorithm, the velocity v_i for each point can be calculated as the Euclidian norm which is defined as:

$$v_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \qquad (1)$$

Hence the augmented matrix Z' including velocities at each point is

$$Z'=\begin{bmatrix} t_1, x_1, & y_1, & v_1 \\ t_2, x_2, & y_2, & v_2 \\ ::::::::: & ::::::::: & ::::: \\ t_i, x_i, & y_i, & v_i \\ ::::::::: & ::::::::: & :::::: \end{bmatrix}$$

From the above augmented matrix Z', we can obtain a fixation group matrix M= $\forall v_i < \varphi \ [fg]$ having all velocities less than the velocity threshold φ. Rest elements which are grater than threshold are termed as saccades. Fixation group matrix contains number of fixation groups fg each fixation group contains k consecutive elements as i^{th} fixation group contains:

$$(fg)_i = \begin{bmatrix} t_{i+1}, x_{i+1}, y_{i+1}, v_{i+1} \\ t_{i+2}, x_{i+2}, y_{i+2}, v_{i+2} \\ :::::::::\ :::::::::\ :::::::: \\ t_{i+k}, x_{i+k}, y_{i+k}, v_{i+k} \end{bmatrix}$$

From each fixation group, fixation vector [3] is derived as $(fixv)_j = [x_j, y_j, t_j, d_j]$ where x_j and y_j are the centroid of horizontal and vertical eye positions in a fixation group respectively. Hence, x_j and y_j is defined as:

$$x_j = (x_1 + x_2 + ...x_k)/k \ ,$$
$$y_j = (y_1 + y_2 + ...y_k)/k$$

t_j is the time stamp which indicates the first time stamp point of a fixation group and d_j is the duration of a fixation group. Hence, t_j and d_j is defined as:

$$t_j = t_{i+1}$$

$$d_j = \|t_{i+k} - t_i\|$$

It is worth to mention here that number of fixation group is the number of fixations. Based on the fixations and saccades, feature vectors are computed.

On the other hand, I-DT algorithm involves a window which works as following:

$$\| \ t_1\, t_2 \quad ...t_{100} \| \ t_{101} \quad t_{102} .... \quad t_{60,000}$$
$$\Downarrow (DuT = 100), \quad D > DT$$

where D=[max(x)-min(x)]+[max(y)-min(y)]

No fixation move rightwards

$$t_1 \| \ t_2 \quad ...t_{101} \| \ t_{102} \quad t_{103} .... \quad t_{60,000}$$
$$\Downarrow move \quad rightwards$$

$$t_1 \| \ t_3 \quad ...t_{102} \| \ t_{103} \quad t_{104} .... \quad t_{60,000}$$
$$\Downarrow \begin{array}{l} D < DT \\ Identified \ as \quad fixation \end{array}$$

I-DT identifies fixations as group of consecutive points within a particular dispersion or maximum separation [3]. I-DT involves two parameters viz duration threshold (DuT) and dispersion threshold (DT); while I-VT involves only single parameter viz velocity threshold (VT). A value below the threshold represents fixation. These threshold values can be estimated from exploratory analysis of scan path data.

After getting fixations, we select 5 eye movement metrics as fixation count (FC), average fixation duration (AFD), average horizontal, vertical and vectorial saccade amplitude (AHSA, AVRSA and AVCSA) [2] which are defined as following:

Fixation count: It is measured as total number of fixations contained within the scanpath.

Average fixation duration: It indicates the amount of time a subject spends interpreting an object [8]. It is measured as sum of fixation durations over fixation count.

Let *FC*= Number of fixations, *d* is the fixation duration and *S*=Number of saccades then

$$AFD = \frac{\sum_{i=1}^{FC} d_i}{FC} \tag{2}$$

Average vectorial saccade amplitude: Average saccade amplitude is considered as biometric feature under assumption that differences in amplitude may be apparent between subjects. It is measured as sum of vectorial saccade amplitudes over the total number of saccades as:

$$AVCSA = \frac{\sum_{i=1}^{S} \sqrt{x_i^{\,2} + y_i^{\,2}}}{S} \tag{3}$$

Average horizontal saccade amplitude: It is measured as sum of horizontal saccade amplitudes greater than 0.5 degree over the total number of horizontal saccades with amplitude greater than 0.5 degree as:

$$AHSA = \frac{\forall x_i > 0.5 \sum_{i=1}^{k} x_i}{S_{x_i > 0.5}} \tag{4}$$

Average vertical saccade amplitude: It is measured as sum of vertical saccade amplitudes greater than 0.5 degree over the total number of vertical saccades with amplitude greater than 0.5 degree as:

$$AVRSA = \frac{\forall y_i > 0.5 \sum_{i=1}^{k} y_i}{S_{y_i > 0.5}} \tag{5}$$

Feature vector for eye movement characteristics is obtained as:

$$fv = [FC, AFD, AHSA, AVRSA, AVCSA]$$

Now, HIM takes these feature vectors as inputs for classification. Cluster allocation component of HIM uses EFC which incorporates the evolutionary search strategy for finding the optimal partitioning generated by different runs of fuzzy c-means (FCM) clustering. Here we use Minkowski distance measure instead of Euclidian distance for providing flexibility in shapes of clusters because Euclidian distance restricts cluster shapes to spherical only which may not be true for different datasets. The objective function of the EFC is following-

$$J = \sum_{k=1}^{N} \sum_{i=1}^{C} (\mu_{ik})^m \, dis^2 (fv_{k,} O_i) \tag{6}$$

Where μ_{ik} is membership grade, *dis* is Minkowski distance between input $fv_{k,}$ and prototype O_i. N is the number of subjects and C is the number of clusters. The fitness function *'f'* is a criterion to determine the best partitioning in evolutionary search, which is selected same as in [5] and its value is inversely proportional to the Xie Beni index (XB), defined as follows:

$$f = \frac{N \min_{i,j} \left\| fv_j - O_i \right\|^2}{\sum_{i=1}^{C} \sum_{j=1}^{N} \mu_{ij}^{\;m} \left\| fv_j - O_i \right\|^2} \tag{7}$$

Let $U^{(0)}, U^{(1)}, ...U^{(l)}$ be the *l* partition matrix generated by the *l* runs of this algorithm. Best *U* is selected based on highest value of the fitness function *'f'*, which then generate the new off-springs by choosing this *U* as parent. Thus repetitive execution of EFC produces best partitioning among the various populations generated by different runs. In conventional FCM clustering, the size of each cluster varies with number of members. In order to avoid this variability and to cope up with the associated syndicate neural networks of the developed model, we need to obtain a uniform structure of all clusters. Therefore, cluster allocation component then allocates fixed number of elements in each cluster, for which fixed number of top membership grade elements are selected as cluster members (CM).

A three layer syndicate neural networks (SNN) is considered in adaptation component of HIM. Back propogation learning algorithm is used for training of SNN. The number of output neurons in a SNN is same as the number of CM, while the number of SNNs is equal to the number of clusters. Training patterns are entered in each SNN for learning. For

TABLE I. EXAMPLE OF A FIXATION GROUP

Time stamp	**X degree**	**Y degree**	**Velocity**
8.0000	17.4361	11.1618	0.0080
9.0000	17.4361	11.1699	0.0028
10.0000	17.4342	11.1719	0.0055
11.0000	17.4380	11.1679	0.0038
12.0000	17.4342	11.1679	0.0083

TABLE II. EXAMPLE OF FIXATION VECTOR

x	**y**	**t**	**d**
-1.8401	-1.6321	0	23
-1.8134	-1.594	26	56
-1.8095	-1.6259	85	17
-1.7919	-1.5176	104	2
-1.7922	-1.6031	108	8

testing, feature vectors of unknown eye movement metrics are fed into the SNN. Let *Op* (M_1), *Op* (M_2)...*Op* (M_C), be the maximum output of SNNs $M_1, M_2 ...... M_C$ respectively. Let $\phi = \overset{C}{\underset{i=1}{Max}}(Op_{x_k}(M_i))$; where $Op_{x_k}(M_i)$ is the maximum outcome of the ith SNN corresponding to the $fv_{k,}$ pattern. A pattern is identified by corresponding member of cluster for which maximum value of *Φ* is obtained. Thus, eye movement metrics are classified by HIM.

III. EXPERIMENTAL EVALUATION

We have used EMBD V1.0 dataset III [2] in which raw scan path signals are provided for 32 subjects. Out of 32, 29 subjects performed 4 recordings and 3 subjects performed 2 recordings. Each recording contain nearly 60,000 raw scan-path signals. After applying I-VT and I-DT, we get 241 and 234 fixations respectively. Table I shows an example of fixation group after applying I-VT at velocity threshold=0.01. Five eye movement metrics described in previous section have been taken as 5 attributes representing a subject. For 29 subjects, we select 2 recordings for training and rest 2 for testing; while for 3 subjects, single recording is used for training and rest single is for testing. Table II shows the 5 fixation vector of first sample of first person. The performance of the system is measured in terms of testing accuracy which is defined as:

Testing accuracy= 100 × (number of correct matches/total number of samples in test set)

Table III shows the various parameters of the HIM which are obtained for best accuracy achieved. From this, it is clear that I-DT+HIM requires more number of clusters, members & learning cycles in comparison with I-VT+HIM. It has been observed that combining I-DT with HIM gives good result as compared with I-VT+HIM. But the draw back of I-DT is that it takes more time in processing the raw signals & in training compared with I-VT due to more number of threshold & HIM parameters involvement. Table 3 shows the performance of these two techniques with selected parameter values. Fig. 2 indicates effect on accuracy on varying number of clusters of HIM. Up to certain number of clusters, the performance of the system increases; after that it starts decreasing. In experiments it is also observed that best accuracy could be obtained when C and CM are nearly comparable. A maximum 63% accuracy has been obtained by combining I-DT with HIM that demonstrates the capability of eye movement scan-path signals as biometric

TABLE III. VARIOUS PARAMETERS OF HIM

Technique	**No. of cluster**	**CM**	**MSE**	**Learning cycles**
I-VT+HIM	6	6	0.007	2000
I-DT+HIM	8	7	0.003	2500

TABLE IV. RESULTS

metho d	**Testing Accurac y**	**FAR**	**FRR**	**VT**	**Du T**	**DT**
IVT+ HIM	59%	41%	41%	0.01	-	-
IDT+ HIM	63%	37%	37%		100	0.05

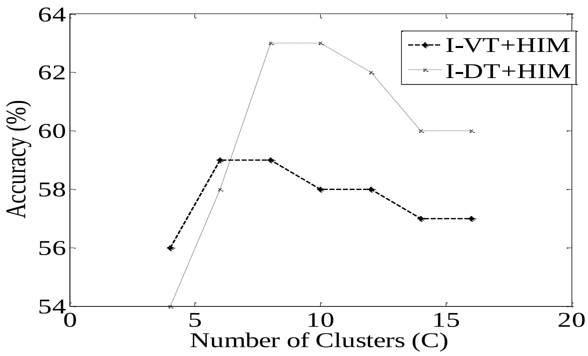

Figure 2. Plot betwwen accuracy and number of clusters

features. For computation of average accuracy, we have used five cross validation. The average accuracy achieved is 58.2 and 62.3% for I-VT+HIM and I-DT+HIM respectively. The false rejection rate (FRR) is observed 41% and 37% for I-VT+HIM and I-DT+HIM respectively. For computation of false acceptance rate (FAR), we have used 2 different instances of 32 subjects. The FAR is 41% and 37% for I-VT+HIM and I-DT+HIM respectively. Since FAR and FRR are equal, therefore equal error rate(EER) is also 41% and 37% for I-VT+HIM and I-DT+HIM respectively.

IV. Conclusion

This paper has presented the potential applicability of the eye movement scan path signals for biometric analysis. It also presented the novel application of proposed hybrid intelligent model 'HIM'. Combination of I-DT with HIM produces more robust results as compared to IVT with HIM. It can be achieved in the cost of slightly more complexity of I-DT+HIM than I-VT+HIM in terms of number of clusters, cluster members and learning cycles. In future work, one can employ learning and visual search characteristics of human brain and can add more eye movement metrics for identification purposes.

ACKNOWLEDGEMENT

Authors would like to thanks to Dr Oleg Komogrtsev and Dr Corey Holland for providing EMDB V1.0 dataset and valuable suggestion in pre-processing the dataset.

REFERENCES

[1] Kasprowski, P., "Human Identification using eye movements", 2004.

[2] Carry Holland & Oleg V. Komogortsev, "Biometric Identification via Eye Movement Scan-paths in Reading", IEEE International joint conference on biometrics (IJCB), pp. 1-8, 2011.

[3] Salvucci D D & Goldberg, "Identifying fixations and saccades in eye tracking protocols*", In proceedings of the 2000 symposium on Eye Tracking Research & Applications*, pp 71-78, New York: ACM press.

[4] J P S Catalao, H. M. I. Pousinho, V. M. F. Mendes, "Hybrid intelligent approach for short term wind power forecasting in portugal", IET renew. Power gener., vol. 5, issue 3, pp. 251-257, 2011.

[5] W sheng, G Howells, M Fairhurst, F Deravi, "Template free biometric key generation by means of fuzzy genetic clustering", IEEE Trans. on information forensic and security, vol. 3, no 2, pp. 183-191, June 2008.

[6] Roman Bednarik, Tomi Kinnunen, Andrei Mihaila, and Pasi Fränti, "Eye-Movements as a Biometric" LNCS 3540, pp. 780 – 789, 2005.

[7] V Srivastava, B K Tripathi, V K Pathak, "Evolutionary fuzzy clustering and functional modular neural network based human recognition, Neural computing and application ,Springer 2012, DOI 10.1007/s00521-012-0973-7

[8] Goldberg J H and Kotval H C, "Computer interface evaluation using eye movements: methods and constructs", international journal of industrial ergonomics, 24 (6): 631-645, 1999.