A Hybrid Fuzzy Approach for Human Eye Gaze Pattern Recognition

Dingyun Zhu[1], B. Sumudu U. Mendis[1], Tom Gedeon[1], Akshay Asthana[2], and Roland Goecke[2]

[1] School of Computer Science
[2] School of Engineering,
The Australian National University, Acton, Canberra, ACT 0200, Australia
dingyun.zhu@anu.edu.au, sumudu.mendis@anu.edu.au,
tom.gedeon@anu.edu.au, aasthana@rsise.anu.edu.au,
roland.goecke@anu.edu.au.

Abstract. Face perception and text reading are two of the most developed visual perceptual skills in humans. Understanding which features in the respective visual patterns make them differ from each other is very important for us to investigate the correlation between human's visual behavior and cognitive processes. We introduce our fuzzy signatures with a Levenberg-Marquardt optimization method based hybrid approach for recognizing the different eye gaze patterns when a human is viewing faces or text documents. Our experimental results show the effectiveness of using this method for the real world case. A further comparison with Support Vector Machines (SVM) also demonstrates that by defining the classification process in a similar way to SVM, our hybrid approach is able to provide a comparable performance but with a more interpretable form of the learned structure.

1 Introduction

Human eyes and their movements are tightly coupled with human cognitive processes, which have been found to be very informative and valuable in various kinds of research areas. Furthermore, previous research has shown that human eye gaze patterns for observing different objects are also quite significant for the understanding of cognitive and decision-making processes.

We have been working on developing effective, efficient and robust approaches to generally provide a clear recognition or unambiguous interpretation of human eye gaze patterns in a variety of settings. In [4], we have successfully shown a sophisticated use of eye gaze information for inference of a user's intention in a game-like interactive task, which effectively eliminates the need of any physical control from the human's side, efficiently improving the communication between the user and the virtual agents.

In this paper, we introduce our hybrid fuzzy approach: hierarchical fuzzy signature construction with Levenberg-Marquardt learning of the generalized Weighted Relevance Aggregation Operator (WRAO) for modeling recognition of human eye gaze patterns between face scanning and text scanning.

M. Köppen et al. (Eds.): ICONIP 2008, Part II, LNCS 5507, pp. 655–662, 2009.

2 Hierarchical Fuzzy Signatures

Hierarchical fuzzy signatures are fuzzy descriptors of real world objects. They represent the objects with the help of sets of available quantities which are arranged in a hierarchical structure expressing interconnectedness and sets of non-homogeneous qualitative measures, which are the interdependencies among the quantities of each set.

The fuzzy signature concept is an effective approach to solve the problem of rule explosion in traditional fuzzy inference systems: constructing characteristic fuzzy structures, modeling the complex structure of the data points (bottom up) in a hierarchical manner [6, 3, 11].

Fuzzy signatures start with a generalized representation of fuzzy sets which are regarded as *Vector Valued Fuzzy Sets (VVFS)* [6]. A Fuzzy Signature is a recursive version of VVFS where each vector can be another VVFS (called a branch) or an atomic value (called a leaf):

$$A : X \rightarrow [a_i]_{i=1}^{k} \tag{1}$$

$$where\ a_i = \begin{cases} [a_{ij}]_{j=1}^{k_i} & ; if\ branch \\ [0,1] & ; if\ leaf \end{cases} \tag{2}$$

Generally, fuzzy signatures result in a much reduced order of complexity, at the cost of slightly more complex aggregation techniques. Unlike conventional rule based hierarchical fuzzy systems, each branch in a fuzzy signature uses a different aggregation function to represent the importance of that branch to its parent, which is a final atomic value called "degree of match". Moreover, fuzzy signatures are different to conventional decision trees as well. They use a bottom up inference mechanism so that even with missing or noisy input data, this structure is still able to find a final result.

The fuzzy signature concept has been successfully applied to a number of applications, such as cooperative robot communication [14], personnel selection models [8], etc. Figure 1 is an example of a fuzzy signature structure which was constructed for a SARS pre-clinical diagnosis system [12].

$$\begin{bmatrix} Fever \begin{bmatrix} 8a.m. \\ 12a.m. \\ 4p.m. \\ 8p.m. \end{bmatrix} \\ BloodPressure \begin{bmatrix} Systolic \\ Diastolic \end{bmatrix} \\ Nausea \\ AbdominalPain \end{bmatrix}$$

Fig. 1. A Fuzzy Signature Example

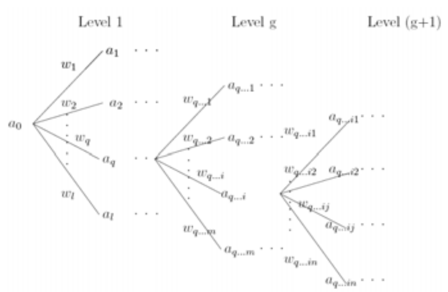

Fig. 2. An Arbitrary Fuzzy Signature Structure

3 Levenberg-Marquardt Learning of WRAO for Fuzzy Signatures

The Weighted Relevance Aggregation Operator (WRAO) [9] is derived from the generalization of the weights and aggregations in Weighted Relevance Aggregation (WRA), which introduces the weighted relevance of each branch to its higher branches of the fuzzy signature structure. In this way, WRAO is able to enhance the accuracy of the results of fuzzy signatures by allowing better adaptation to the meaning of the decision making process [10], and it can help to reduce the number of individual fuzzy signatures by absorbing more patterns into one structure.

The generalized Weighted Relevance Aggregation Operator (WRAO) of an arbitrary branch $a_{q\dots i}$ with n sub-branches, $a_{q\dots i1}$, $a_{q\dots i2}$,..., $a_{q\dots in} \in [0,1]$, and weighted relevancies, $w_{q\dots i1}$, $w_{q\dots i2}$,..., $w_{q\dots in} \in [0,1]$ (see Figure 2), for a fuzzy signature is a function g: $[0,1]^{2n} \to [0,1]$ such that,

$$a_{q\dots i} = \left[\frac{1}{n} \sum_{j=1}^{n} (a_{q\dots ij} \cdot w_{q\dots ij})^{p_{q\dots i}} \right]^{\frac{1}{p_{q\dots i}}} \tag{3}$$

The Levenberg-Marquardt (LM) method is not only a major learning algorithm in neural network training functions, but also a widely used advanced approach that outperforms simple gradient descent and gradient methods for solving most of the optimization based problems. This algorithm is a Sum of Squared Error (SSE) based minimization method that is the function to be minimized is of the following special form [7]:

$$f(s) = \frac{1}{2} \sum_{i=1}^{n} (t_i - s_i)^2 = \frac{1}{2} \parallel \overline{t} - \overline{s} \parallel \tag{4}$$

where $\overline{t}$ stands for the target vector, $\overline{s}$ for the predicted output vector of the fuzzy signature, and $\|\|$ denotes the $2 - norm$. Also, it will be assumed that there are m parameters to be learned and n records in the training data set, such that $n > m$. The next update of the LM is the following equation:

$$\underline{u}[k] = \underline{par}[k] - \underline{par}[k-1] \tag{5}$$

where the vector $\underline{par}[k]$ contains all the parameters to be learned, i.e. all the aggregation factors and weights of WRAO in the equation (3) for the kth iteration. Then the next update of $\underline{u}[k]$ is defined as:

$$\left(J^T[k]\,J[k] + \alpha I\right)\underline{u}[k] = -J^T[k]\,\underline{e}[k] \tag{6}$$

where J stands for the Jacobian matrix of the equation (4), I is the identity matrix of J, and α is a regularization parameter, which control both search direction and the magnitude of the next update $\underline{u}[k]$.

4 Eye Gaze Data Collection

An eye gaze data collecting experiment was conducted. Ten volunteers (Gender: 5 male, 5 female; Occupation: 2 academic staff, 6 postgraduates, 2 undergraduates) from the Australian National University community participated in the study.

Two sets of human face pictures, 20 in each, were selected for the face scanning experiment. Another 5 text only documents with different lengths (minimal half page, maximal one page) were also shown. In the experiment, all the face pictures and documents were demonstrated as full screen scenes on a monitor.

Every participant was firstly asked to view one set of human face pictures with about 5 seconds on each. The second stage of the experiment was to read the 5 text documents to determine which were the most important sentences in each one, no time restriction was imposed for the reading test so the participants could conduct the reading with their usual speed. In the ranking of sentences phase, only 1 participant ranked all the sentences, most participants ranked only 3 sentences so we conclude that our instructions were interpreted as a text scanning task. After that, the face scanning test was performed again with the other set of pictures as the last stage.

There was no time break between any two stages, all the eye movement data was collected by using a Seeingmachines eye-tracking system with FaceLAB software (Version 4.5, 2007) through the entire session of the experiment.

5 Fuzzy Signature Construction for Recognition of Eye Gaze Pattern

Since people tend to concentrate their gaze fixations onto the interesting and informative regions in the scene [13], we further filtered the original collected gaze points into fixations which offered a much easier and more interpretable form for the later data process. In addition, instead of considering all the fixations on each of the test cases (either face scanning or document scanning), we only use the first five fixations from every case. The reason for this is that it is possible to interpret a plausible eye gaze pattern from the early stage of face viewing (as early as the first five fixations) [1]. Moreover, the time period for scanning a document was obviously much longer than viewing a face in the data collecting experiment, so the decision to use only the first five fixations also maintains a more similar pattern for the future structure construction.

To construct the fuzzy signature structure for learning, it is necessary to figure out which essential feature in both of the possible patterns can show the difference for recognition. Figure 3 illustrates the first five fixations for two eye gaze patterns from face viewing as well as text scanning respectively. The two cases are obvious samples and this is actually not the usual source in all the data records we collected in the experiment.

From these two cases, we can easily find the most obvious difference between them is in the geometrical shapes, which shows that compared with face

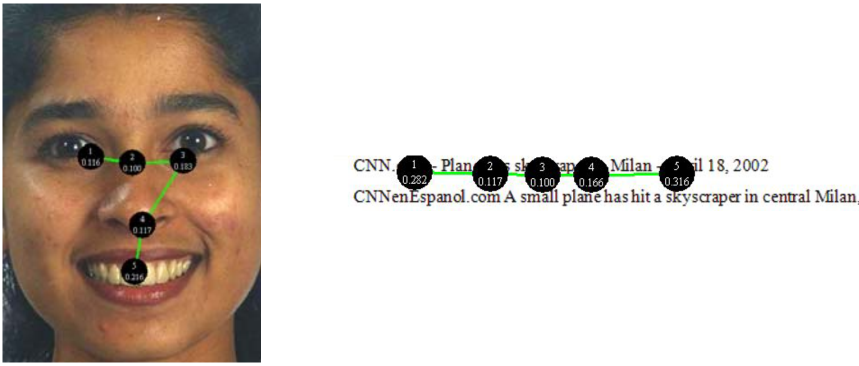

Fig. 3. Two Samples of First Five Fixations Only Eye Gaze Patterns

scanning, participants' gaze fixation locations for text scanning follow a very clear horizontal pattern. On the other hand, although it is still difficult to address a common gaze pattern for the face scanning, the plausible pattern has a much more complicated geometrical shape than the simple form from text scanning, because the informative regions (eyes, nose, mouth and cheeks, etc) in which an observer is interested in a face are not aligned horizontally as are the sentences in a document.

According to the above point, we can further discover that the actual feature in both of the patterns rests on the vertical difference between two fixations which are adjacent on time. Consequently, the constructed fuzzy signatures for the recognition of two patterns can be formed to the structure in Figure 4.

The leaves of each sub-signature in the structure represent the fuzzy value calculated by using the Fuzzy C Mean (FCM) clustering method based on the vertical difference between two adjacent fixations.

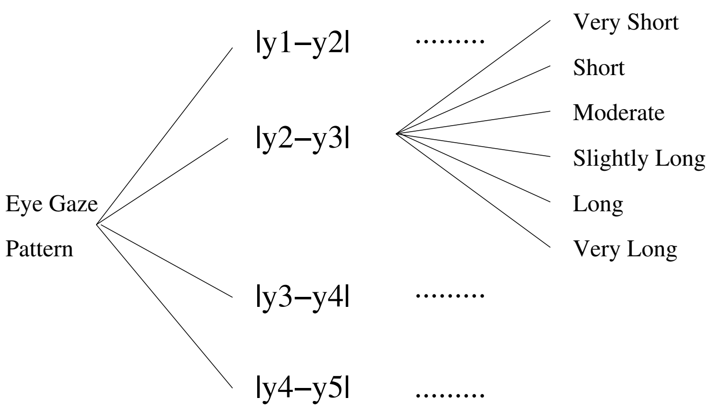

Fig. 4. Fuzzy Signature Structure for Eye Gaze Pattern

Table 1. Fuzzy Signatures Results for Eye gaze Pattern Recognition

Experiment	Mean Squared Error (MSE)	Classification Error (CLE)
Training	0.0605	19.77%
Test	0.0641	20.00%

6 Evaluation and Comparison

According to the constructed fuzzy signature structure for the recognition of these two different patterns, we performed experiments to learn the weights and aggregations by applying the Levenberg-Marquardt optimization method as we explained in the previous section. The following table shows the results of Mean Squared Error (MSE) and Classification Error (CLE) for learning the weights of WRAO for the training and test experiments respectively.

We need to clarify that the way we calculate the Classification Error (CLE) for the fuzzy signature structure is actually based on the degree of difference (d) between the predicted value and the initial desired value. In our eye gaze pattern case, we set three classes according to the degree of difference: *Good* ($|d| \leq 0.2$, not an error), *Bad* ($0.2 < |d| \leq 0.5$, count 0.5 error) and *Very Bad* ($|d| > 0.5$, count 1 error). So the final classification error rate would be the sum of all the error numbers divided by the total number of records in the data set.

Table 1 shows that our hybrid fuzzy approach can perform around 80% accurate predictions for the recognition of different eye gaze patterns between face and text scanning.

Furthermore, in order to have a performance comparison, the Support Vector Machines (SVM) [5] based classifier was chosen to run through the same eye gaze fixation data set. Since the classification problem here is only for the recognition between two different eye gaze patterns, we constructed a simple SVM based classifier using a linear kernel to classify the data of vertical difference between two adjacent fixations as we used in the previous experiment. For our fuzzy signature structure, we also reduced the number of classes from previous three (*Good, Bad and Very Bad*) to two (*Good* $|d| \leq 0.5$ and *Bad* $|d| > 0.5$). Table 2 demonstrates the results of the experiments between fuzzy signatures and SVM.

From the above results we can see for this particular pattern recognition problem, the simple SVM based classifier constructed by using a linear kernel gives highly accurate classification results. Comparatively, by reducing the number of classes for the fuzzy signatures from previous three (*Good, Bad and Very Bad*) to two (*Good* $|d| \leq 0.5$ and *Bad* $|d| > 0.5$), the classification error rate reduces and is comparable to that of SVM.

Table 2. CLE Comparison Between Fuzzy Signatures and SVM

Experiment	FS (3 classes)	FS (2 classes)	SVM (2 classes)
Training	19.77%	3.04%	1.18%
Test	20.00%	5.00%	4.55%

Beyond the results comparison, we also find that the way fuzzy signatures model the classification problem is different from the way SVM models a classification problem. SVM approaches the classification problem through the concept of *margin*, that is equal to the smallest distance between the decision boundary and any of the samples [2]. It computes the maximum margin hyperplane that best defines the decision boundary. The samples that are closest to the decision boundary lie on this maximum margin hyperplane and are know as support vectors. Any other sample point in the data set plays no role in the classification problem and is discarded. Once the decision boundary is known, the new data points can be classified according to which side of the hyperplane it lies. On the other hand, the construction of fuzzy signature is based on the expression of the domain knowledge. Further, both the weight and aggregation learning processes for WRAO offer us a clear view of what exactly produces the results inside the structure, for instance, which aggregation functions are learned for each branch. In addition, the value generated after every aggregation actually represents the degree of match as the importance of the current branch to its parent, which is also useful to discover which sub-branch makes the most contribution and which are not significant factors from the domain for the problem modeling. So fuzzy signatures provide a more interpretable expression of how well the output of the structure approaches the target value or classification.

7 Conclusion

A fuzzy signature with a Levenberg-Marquardt learning based hybrid approach has been introduced for modeling a real world study: recognizing human eye gaze patterns to distinguish between face scanning and text scanning. The constructed structure shows significant performance for the recognition in the experiment results.

From the further discussion of performance comparison with linear SVM, we suggest that our structure is capable of producing a comparable result for the classification, but as an effective approach, it can provide a more interpretable representation signature pattern of a real world object from its construction as well as the learning process.

References

[1] Althoff, R.R., Cohen, N.J.: Eye-movement Based Memory Effect: A Reprocessing Effect in Face Perception. Journal of Experimental Psychology: Learning, Memory, and Cognition 25, 997–1010 (1999)

[2] Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, Heidelberg (2006)

[3] Gedeon, T.D., Koczy, L.T., Wong, K.W., Liu, P.: Effective Fuzzy Systems for Complex Structured Data. In: Proceedings of IASTED International Conference Control and Applications (CA 2001), Banff, Canada, pp. 184–187 (2001)

[4] Gedeon, T.D., Zhu, D., Mendis, B.S.U.: Eye Gaze Assistance for a Game-like Interactive Task. International Journal of Computer Games Technology 2008, Article ID 623752(2008)
[5] Hsu, C., Chang, C., Lin, C.: A Practical Guide to Support Vector Classification. Technical Report, Department of Computer Science, National Taiwan University (2005)
[6] Koczy, L.T., Vamos, T., Biro, G.: 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 1999), Budapest, Hungary, pp. 210–217 (1999)
[7] Marquardt, D.: An Algorithm for Least Squares Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics 11(2), 431–441 (1963)
[8] Mendis, B.S.U., Gedeon, T.D.: A Comparison: Fuzzy Signatures and Choquet Integral. In: Zurada, J.M., Yen, G.G., Wang, J. (eds.) WCCI 2008. LNCS, vol. 5050, pp. 1464–1471. Springer, Heidelberg (2008)
[9] Mendis, B.S.U., Gedeon, T.D., Botzheim, J., Koczy, L.T.: Generalized Weighted Relevance Aggregation Operators for Hierarchical Fuzzy Signatures. In: Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation (CIMCA 2006), Jointly with the International Conference on Intelligent Agents, Web Technologies and International Commerce (IAWTIC 2006), Sydney, Australia, pp. 198–203 (2006)
[10] Mendis, B.S.U., Gedeon, T.D., Koczy, L.T.: Learning Generalized Weighted Relevance Aggregation Operators Using Levenberg-Marquardt Method. In: Proceedings of the 6th International Conference on Hybrid Intelligent System (HIS 2006) and 4th Conference on Neuro-Computing and Evolving Intelligence (NCEI 2006), New Zealand, pp. 34–39 (2006)
[11] Vamos, T., Koczy, L.T., Biro, G.: Fuzzy Signatures in Data Mining. In: Proceedings of Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, Canada, vol. 5, pp. 2841–2846 (2001)
[12] Wong, K.W., Gedeon, T.D., Koczy, L.T.: Construction of Fuzzy Signature from Data: An Example of SARS Pre-clinical Diagnosis System. In: Processings of IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2004), Budapest, Hungary, pp. 1649–1654 (2004)
[13] Yarbus, A.: Eye Movements and Vision. Plenum Press, New York (1967)
[14] Zhu, D., Gedeon, T.: Fuzzy Logic for Cooperative Robot Communication. In: Nguyen, N.T., Jo, G.S., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2008. LNCS, vol. 4953, pp. 401–410. Springer, Heidelberg (2008)