Multi Local Feature Selection Using Genetic Algorithm For Face Identification

Dzulkifli Mohamad dzulkifli@utm.my

Falkuti Sains Komputer dan Sistem Maklumat
Universiti Teknologi Malaysia
Skudai,81310, Johor, Malaysia

Abstract

Face recognition is a biometric authentication method that has become more significant and relevant in recent years. It is becoming a more mature technology that has been employed in many large scale systems such as Visa Information System, surveillance access control and multimedia search engine. Generally, there are three categories of approaches for recognition, namely global facial feature, local facial feature and hybrid feature. Although the global facial-based feature approach is the most researched area, this approach is still plagued with many difficulties and drawbacks due to factors such as face orientation, illumination, and the presence of foreign objects. This paper presents an improved offline face recognition algorithm based on a multi-local feature selection approach for grayscale images. The approach taken in this work consists of five stages, namely face detection, facial feature (eyes, nose and mouth) extraction, moment generation, facial feature classification and face identification. Subsequently, these stages were applied to 3065 images from three distinct facial databases, namely ORL, Yale and AR. The experimental results obtained have shown that recognition rates of more than 89% have been achieved as compared to other global-based features and local facial-based feature approaches. The results also revealed that the technique is robust and invariant to translation, orientation, and scaling.

Keywords: Face Recognition, Facial Feature Extraction, Localization, Neural Network, Genetic Algorithm (GA)

1. INTRODUCTION

Face recognition is one of the physiological biometric technologies which exploit the unique features on the human face. Although face recognition may seem an easy task for human, but machine recognition is a much more daunting task [1]. The difficulties due to pose, present or absent of structural components, occlusion, image orientation, facial expression and imaging conditions [2]. For the last two decades, there has been growing interest in machine recognition of faces due to its potential applications, such as film processing, user authentication, access control system, law enforcement, etc. Typically face recognition system should include four stages. The first stage involves detecting human face area from images, i.e. detect and locate face. The second stage requires extraction of a suitable representation of the face region. The third stage classifies the facial image based on the representation obtained in the previous stage. Finally, compares facial image against database (gallery) and reports a match.

To design a high accuracy recognition system, the choice of feature extractor is very crucial. In general, feature extraction methods can be divided into two categories: face based and constituent based. The face based approach uses raw pixel information or features extracted from the whole image which as a representation of face. Therefore face based method uses global information instead of local information. Principal Component Analysis (PCA) is a typical and successful face based method. Turk and Pentland developed a face recognition system using PCA in 1991 [3]. In 1997, Belhumeur et. al. proposed Fisherface technique based on Linear Discriminant Analysis (LDA) to overcome the difficulty cause by illumination variation [4]. Haddadnia et. al. introduced a new method for face recognition using Pseudo Zernike Moment Invariants (PZMI) as features and Radial Basis Function (RBF) neural network as the classifier [5], [6], [7]. Since the global information of an image are used to determine the feature elements, information that are irrelevant to facial region such as shoulders, hair and background may contribute to creation of erroneous feature vectors that can affect the face recognition results. Furthermore, due to the variation of facial expression, orientation and illumination direction, single feature is usually not enough to represent human face. So the performance of this approach is quite limited.

The second one is the constituent based approaches are based on relationship between extracting structural facial features, such as eyes, mouth, nose, etc. The constituent approaches deal with local information instead of global information. Therefore constituent based method can provides flexibility in dealing facial features, such as eyes and mouth and not affected by irrelevant information in an image. Yuille et. al. use Deformable Templates to extract facial features [8]. These are flexible templates constructed with a priori knowledge of the shape and size of the different features [9]. The templates can change their size and shape so that they can match properly. These methods work well in detection of the eyes and mouth, despite variations in tilt, scale and rotation of head. However modeling of the nose and eyebrow was always a difficult task [8], [9]. Additionally it cannot deals with complicated background settings. Moreover the computation of template matching is very time consuming. In 1999, Lin et. al. presented an automatic facial feature extraction using Genetic Algorithm (GA) [10]. In 2002, Yen et. al. proposed a novel method using GA to detect human facial features from images with a complex background without imposing any constraints [11]. The normal process of searching for the features is computationally expensive; therefore GA is used as a search algorithm [11]. Genetic algorithm possesses the following feature that make them better suited that traditional search algorithm [12]. Comparing to face based approach, constituent based approach provide flexibility in dealing facial features, such as eyes and mouth and not affected by irrelevant information in an image; therefore constituent based approach is selected as a solution in this paper.

In the literature [13] and [14], the combination of an ensemble of classifiers has been proposed to achieve image classification systems with higher performance in comparison with the best performance achievable employing a single classifier. In Multiple Classifier System [15], different structures for combining classifier systems can be grouped in three configurations. In the first group, the classifier systems are connected in cascade to create pipeline structure. In the second group, the classifier systems are used in parallel and their outputs are combined named it parallel structure. Lastly the hybrid structure is a combination of the pipeline and parallel structure.

So, this paper proposes a human face recognition system that can be designed based on hybrid structural classifier system. The intended scheme actually is designed to have evolutionary recognition results by gathering available information and extracting facial features from input images. In this paper, Pseudo Zernike Moment Invariant (PZMI) has been used as a feature domain to extract features from facial parts. Radial Basis Function (RBF) neural network is used as the classifier in the proposed method. RBF neural network is chosen due to their simple topological structure, their locally tuned neurons and their ability to have a fast learning algorithm in comparison with the multi-layer feed forward neural network [16], [19].

The organization of the paper is structured as follow. Face parts localization using GA, moment generation using PZMI, facial feature classification using RBF, multi local feature selection, experimental results and conclusion.

2. PROPOSED METHOD

2.1 Facial Parts Localization using GA

This is a face segmentation and facial feature extraction process [11], which gathers the sub-regions of right eye, left eye, mouth and nose using GA. All the images captured were head and shoulder images and in a frontal view.

2.1.1 Genetic Algorithm

GA is a powerful search and optimization algorithm, which are based on the theory of natural evolution. In GA, each solution for the problem is called a chromosome and consists of a linear list of codes. The GA sets up a group of imaginary lives having a string of codes for a chromosome on the computer. The GA evolves the group of imaginary lives (referred to as population), and gets and almost optimum solution for the problem. The GA uses three basic operators to evolve the population: selection, crossover, and mutation.

2.1.2 Face Segmentation

The face segmentation process is proceeded under the assumption that human face region can be approximated by an ellipsoid [17]. Therefore each chromosome in the population during the evolutionary search has five parameters genes, the centre of the ellipse (x and y), x directional radius (r_x), y directional radius (r_y) and the angle (ϴ). Figure 1 shows the chromosome for face segmentation.

x-8bits	y-8bits	r_x-8bits	r_y-8bits	ϴ-7bits

FIGURE 1: Chromosome for Face Segmentation

The fitness of the chromosome is defined by the number of edge pixels in the approximated ellipse like face to the actual number of pixels in the actual ellipse. The ratio is large when both ellipses overlap perfectly.

2.1.3 Facial Feature Extraction

After the process of face segmentation, segmented image is fed into facial feature extraction process. The facial feature extraction is based on horizontal edge density distribution [11]. The horizontal edge map of the image from segmented image is obtained in order to extract facial features. In this method, rectangle templates of different sizes for different facial features are used. The sizes of the templates for different features are decided according to general knowledge of the size of the features. Here, both the eye and eyebrow are contained in the same rectangle template.

In order to make the search process less computational expensive, face is divided into sub-regions as shown in Figure 2. The right eye is in the region E_r, left eye in the region E_l, and region M contains the mouth. The nose region N can be obtained once the eyes and mouth are located.

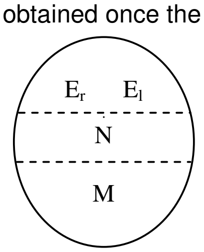

FIGURE 2: Sub-regions of the face

GA is used in the process of facial feature extraction to search for the global maximum point when the template best matches the feature. The chromosome for face feature extraction shown in Figure 3.

x-direction (7 bits)	y-direction (7 bits)

FIGURE 3: Chromosome for face feature extraction

The chromosome represents the position of the feature in the *x* and *y* direction. The fitness is evaluated in terms of the density of the template. The best template is selected when the fitness is maximized. The fitness, *F* is shown below,

$$F = \frac{1}{m \times n} \sum_{x=1}^{m} \sum_{y=1}^{n} T(x, y)$$

(2.1)

$$\text{where} \begin{cases} T(x, y) = 1 & \text{if the pixel is white} \\ T(x, y) = 0 & \text{if the pixel is black} \end{cases},$$

and *T* is the template, (*x,y*) are the coordinates of the template, and $m \times n$ is the size of the template.

2.2 Moment Generation using PZMI

PZMI is an orthogonal moment that is shift, rotation and scale invariant and very robust in the presence of noise. PZMI is been used for generating feature vector elements. Pseudo Zernike polynomials are well known and widely used in the analysis of optical systems. Pseudo Zernike polynomials are orthogonal set of complex-valued polynomials, V_{nm} defined as [5], [6], [7], [16]:

$$V_{nm}(x, y) = R_{nm}(x, y) \exp(jm \tan^{-1}(\frac{y}{x}))$$

(2.2)

where $x^2 + y^2 \leq 1, \quad n \geq 0, \quad |m| \leq n$ and Radial polynomial R_{nm} are defined as:

$$R_{nm}(x, y) = \sum_{s=0}^{n-|m|} D_{n,|m|,s} (x^2 + y^2)^{\frac{n-s}{2}}$$

(2.3)

where:

$$D_{n,|m|,s} = (-1)^s \frac{(2n+1-s)!}{s!(n-|m|-s)!(n-|m|-s+1)!}$$

(2.4)

The PZMI can be computed by the scale invariant central moments $CM_{p,q}$ and the radial geometric moments $RM_{p,q}$ as follows:

$$\begin{aligned} PZMI_{nm} = {} & \frac{n+1}{\pi} \sum_{(n-m-s)\text{even}, s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{k} \sum_{b=0}^{m} \binom{k}{a} \binom{m}{b} \\ & (-j)^b CM_{2k+m-2a-b, 2a+b} \\ & + \frac{n+1}{\pi} \sum_{(n-m-s)\text{odd}, s=0}^{n-|m|} D_{n,|m|,s} \sum_{a=0}^{d} \sum_{b=0}^{m} \binom{d}{a} \binom{m}{b} \\ & (-j)^b RM_{2d+m-2a-b, 2a+b} \end{aligned}$$

(2.5)

where k =(n-s-m)/2, d=(n-s-m)/2, $CM_{p,q}$ is the central moments and $RM_{p,q}$ is the Radial moments are as follow:

$$CM_{p,q} = \frac{\mu_{pq}}{M_{00}^{(p+q+2)/2}}$$
(2.6)

$$RM_{p,q} = \frac{\sum_x \sum_y f(x, y)(\tilde{x}^2 + \tilde{y}^2)^{1/2} \tilde{x}^p \tilde{y}^q}{M_{00}^{(p+q+2)/2}}$$
(2.7)

where $\tilde{x} = x - x_0$, $\tilde{y} = y - y_0$ and M_{pq}, μ_{pq} and x_0, y_0 are defined as follow:

$$M_{pq} = \sum_x \sum_y f(x, y)\, x^p\, y^q$$
(2.8)

$$\mu_{pq} = \sum_x \sum_y f(x, y)(x - x_0)^p (y - y_0)^q$$
(2.9)

$$x_0 = M_{10} / M_{00}$$
(2.10)

$$y_0 = M_{01} / M_{00}$$
(2.11)

2.3 Facial Feature Classification Using RBF

RBF neural network has been found to be very attractive for many engineering problem because [18], [19]:

(i)
They are universal approximators, (ii)They have a very compact topology and (iii) Their learning speed is very fast because of their locally tuned neurons.

Therefore the RBF neural network serve as an excellent candidate for pattern applications and attempts have been carried out to make the learning process in this type of classification faster then normally required for the multi-layer feed forward neural networks [19]. In this paper, RBF neural network is used as classifier in face recognition system.

2.3.1 RBF Neural Network Structure

Figure 4 shows the basic structure of RBF neural networks.

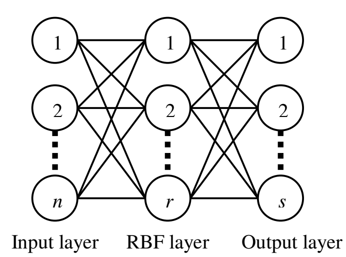

FIGURE 4: RBF Neural Network Structure

The input layer of the neural network is a set of *n* unit, which accept the elements of an *n*-dimensional input feature vector. The input units are fully connected to the hidden layer *r* hidden units. Connections between the input and hidden layers have unit weights and, as a result, do not have to be trained. The goal of the hidden layer is to cluster the data and reduce its dimensionality. In this structure the hidden units are referred to as the RBF units. The RBF units are also fully connected to the output layer. The output layer supplies the response of the neural network to activation pattern applied to the input layer. The transformation from the input space to the RBF-unit space is nonlinear (nonlinear activation function), whereas the transformation from the RBF-unit space to the output space is linear (linear activation function). The RBF neural network is a class of neural network where the activation function of the hidden units is determined by the distance between the input vector and a prototype vector. The activation function of the RBF units is expressed as follow [7], [18], [20]:

$$R_i(x) = R_i\left(\frac{\|x - c_i\|}{\sigma_i}\right), \qquad i = 1,2,\ldots,r$$

(2.12)

where x is an n-dimensional input feature vector, c_i is an n-dimensional vector called the centre of the RBF unit, σ_i is the width of the RBF unit, and r is the number of the RBF units. Typically the activation function of the RBF units is chosen as a Gaussian function with mean vector c_i and variance vector σ_i as follow:

$$R_i(x) = \exp\left(-\frac{\|x - c_i\|^2}{\sigma_i^2}\right)$$

(2.13)

Note that σ_i^2 represents the diagonal entries of the covariance matrix of the Gaussian function. The output units are linear and the response of the jth output unit for input x is:

$$y_j(x) = b(j) + \sum_{i=1}^{r} R_i(x) w_2(i,j)$$

(2.14)

where $w_2(i, j)$ is the connection weight of the ith RBF unit to the jth output node, and $b(j)$ is the bias of the jth output. The bias is omitted in this network in order to reduce the neural network complexity [5], [19], [20]. Therefore:

$$y_j(x) = \sum_{i=1}^{r} R_i(x) \times w_2(i,j)$$

(2.15)

2.4 Multi Local Feature Selection

The layout of multi local feature selection has been shown in Figure 5. In the first step, facial parts localization process is done, so the exact location of the facial parts regions is localized. Secondly, sub-image of each facial parts will be created, which contain only relevant information of facial parts, such as eyes, nose, mouth, etc. Next in third stage, each of the facial parts is extracted in parallel from the derived sub-image. The fourth stage is the process of classification, which classify the facial features. Finally the last stage combines the outputs of each neural network classifier to construct the recognition.

3. EXPERIMENTAL RESULTS

To validate the effectiveness of the algorithm, a simple experiment was carried out. The human face images were taken using a monochrome CCD camera with a resolution of 768 by 576 pixels. There are also some images from international face database is been used, such as face image from ORL, Yale and AR Database. The total number of 3065 images have been selected from all the database as a test and train images. The GA parameters setting used for both face segmentation and facial feature extraction in the simulation process are shown in Table 1.

	Face segmentation	**Feature extraction**
Population	100	50
Crossover	0.8	0.8
Mutation	0.001	0.001

TABLE 1: GA Parameters

Figure 7 displays the head and shoulder original image before the process of facial parts localization. Figure 8 shows the result after the process of facial parts localization.

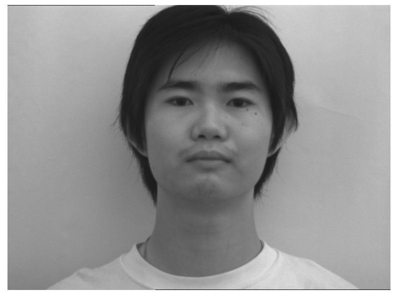

FIGURE 7: Head and shoulder original image

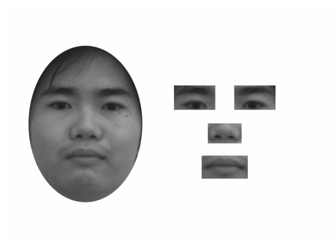

FIGURE 8: Result of Facial Parts Localization

Table 2 shows some of the features extracted by PZMI. Though it may argued that there exists a similar value (or closed to) among different facial features but it never happens for the entire complete set. To investigate the effect of the method of learning on the RBF neural network, three categories of feature vector based on the order (n) of the PZMI have been set (Table 3). The neural network classifier was trained in each category based on the training images.

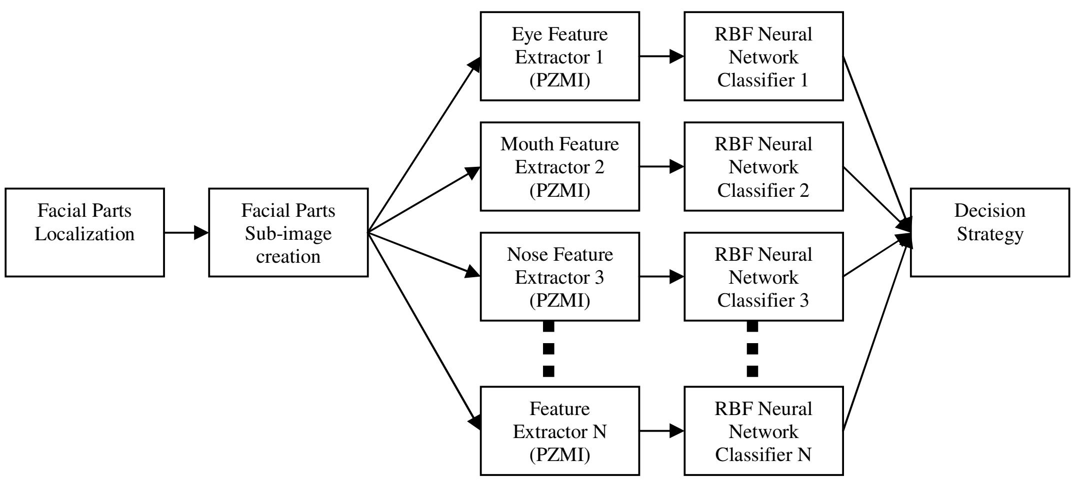

FIGURE 5 : The layout of multi local feature selection

	Person A		Person B	
	Left eye	Right eye	Left eye	Mouth
PZMI $_{9,1}$	0.025769	0.030027	0.027727	0.010727
PZMI $_{9,2}$	0.017139	0.011290	0.012600	0.000254
PZMI $_{9,3}$	0.021621	0.017175	0.024139	0.008444
PZMI $_{9,4}$	0.002486	0.003062	0.006773	0.027121
PZMI $_{9,5}$	0.036770	0.035310	0.030024	0.020046
PZMI $_{9,6}$	0.090679	0.092341	0.091703	0.003933
PZMI $_{9,7}$	0.062495	0.070282	0.075366	0.011679
PZMI $_{9,8}$	0.082933	0.080637	0.083488	0.058776
PZMI $_{9,9}$	0.020375	0.014172	0.022936	0.016866

TABLE 2: Features extracted by PZMI

Category No.	PZMI feature elements
1	n=1, m=0,1 n=2, m=0,1,2 n=3, m=0,1,2,3 n=4, m=0,1,2,3,4 n=5, m=0,1,2,3,4,5 n=6, m=0,1,2,3,4,5,6
2	n=6, m=0,1,2,3,4,5,6 n=7, m=0,1,2,3,4,5,6,7 n=8, m=0,1,2,3,4,5,6,7,8
3	n=9, m=0,1,2,3,4,5,6,7,8,9 n=10, m=0,1,2,3,4,5,6,7,8,9,10

TABLE 3: Feature Vectors Elements based on PZM

The experimental results and the comparison between the previous research works using the same dataset from three distinct facial databases are shown in Table 4. It shows that the overall recognition rate of more than 89% has been achieved by the proposed method. The results also reveal that the proposed technique is robust and invariant to translation, orientation, and scaling.

Database	Eigen	Fisher	EGM	SVM	NN	Proposed method
ORL	80.3%	93.8%	81.5%	95.5%	91.5%	96.5%
Yale	66.7%	77.6%	82.4%	78.2%	74.5%	83.0%
AR	28.7%	89.2%	58.7%	59.7%	76.4%	89.2%
Average	58.6%	86.9%	74.2%	77.8%	80.8%	89.6%

TABLE 4: Recognition rate of experiment

4. CONCLUSION

This paper presented a method for the recognition of human faces in 2-Dimentional digital images using a localization of facial parts information. The combination of an ensemble of classifiers has been used to achieve image classification systems with higher performance in comparison with the best performance achievable employing a single classifier.

5. REFERENCES

1. Chellappa, R.; Wilson, C. L.; and Sirohey, S. "*Human and Machine Recognition of Faces: A Survey*". Proceedings of the IEEE, 83(5): 705-740, 1995
2. Yang, M. H.; Kriegman, D. J.; and Ahuja, N. "*Detecting Faces in Images: A Survey*". IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1): 34-58, 2002
3. Turk, M.; and Pentland, A. "*Face Recognition using Eigenfaces*". In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 586-591. Maui, USA, 1991
4. Belhumeur, P.N.; Hespanha, J.P.; and Kriegman, D. J "*Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection*". IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7): 711-720, 1997
5. Haddadnia, J.; Ahmadi, M.; and Faez,. *"A Hybrid Learning RBF Neural Network for Human Face Recognition with Pseudo Zernike Moment Invariant"*. In Proceedings of the IEEE 2002 International Joint Conference on Neural Networks, 1: 11-16. Honolulu, USA, 2002
6. Haddadnia, J.; Ahmadi, M.; and Faez. *"An Efficient Method for Recognition of Human Faces using Higher Orders Pseudo Zernike Moment Invariant"*. In Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 315-320. Washington, USA. 2002
7. Haddadnia, J.; Faez, K.; and Ahmadi, M. *"An Efficient Human Face Recognition System Using Pseudo Zernike Moment Invariant and Radial Basis Function Neural Network*". International Journal of Pattern Recognition and Artificial Intelligence 17(1): 41-62, 2003
8. Yullie, A. L.; Cohen, D. S.; and Hallinan, P. W. *"Feature Extraction from Faces using Deformable Templates*". In Proceeding of the IEEE International Conference on Pattern Recognition, 104-109. San Diego, USA, 1989
9. Huang, L. C.; and Chen, C. W. "*Human Face Feature Extraction for Face Interpretation and Recognition*". In Proceedings of the IEEE International Conference on Pattern Recognition, 204-207. Hague Netherlands, 1996
10. Lin, C. H.; and Wu, J. L. "*Automatic Facial Feature Extraction by Genetic Algorithms*". IEEE Transactions on Image Processing, 8(6): 834-845, 1999
11. Yen, G. G.; and Nithianandan, N. "*Facial Feature Extraction Using Genetic Algorithm*". In Proceedings of the IEEE 2002 Congress on Evolutionary Computation, 2: 1895-1900. Honolulu, USA, 2002.
12. Tang, K. S.; Man, K. F.; Kwong, S.; and He, Q. "*Genetic Algorithms and Their Applications"*. IEEE Signal Processing Magazine 13(6): 22-37, 1996
13. Giacinto, G.; Roli, F.; and Fumera. "*Unsupervised Learning of Neural Network Ensembles for Image Classification"*. In Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, 3: 155-159. Como Italy, 2000
14. Kittler, J.; Hatef, M.; Duin, R. P. W.; and Matas, J. "*On Combining Classifier"*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20: 226-239, 1998
15. Ho, T. K.; Hull, J. J.; and Srihari, S. N. "*Decision Combination in Multiple Classifier System"*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(1): 66-75. 1994
16. Haddadnia, J.; Faez, K.; and Moallem, P. "*Neural Network Based Face Recognition with Moment Invariants",* In Proceedings of the IEEE International Conference on Image Processing, 1: 1018-1021. Thessaloniki Greece, 2001
17. Yokoo, Y.; and Hagiwara, M. "*Human Face Detection Method using Genetic Algorithm"*. In Proceedings of the IEEE Congress on Evolutionary Computation, 113-118. Nagoya Japan, 1996

18. Yingwei, L.; Sundarajan, N.; and Saratchandran, P. "*Performance Evaluation of A Sequential Minimal Radial Basis Function (RBF) Neural Network Learning Algorithm*". IEEE Transactions on Neural Network, 9(2): 308-318, 1998
19. Zhou, W. "*Verification of the Nonparametric Characteristics of Backpropagation Neural Networks for Image Classification*". IEEE Transactions Geoscience and Remote Sensing, 37(2): 771-779. 1999.
20. Jang, J. -S. R. ANFIS: "*Adaptive-Network-Based Fuzzy Inference System*". IEEE Transactions Systems, Man and Cybernetics, 23(3): 665-684, 1993