

Neuro-fuzzy model for estimating race and gender from geometric distances of human face across pose

K Nanaa¹, M N A Rahman², M Rizon³, F S Mohamad⁴, M Mamat⁵

^{1,2,4,5}Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus, 22200 Besut, Terengganu, Malaysia

³Faculty of Visual Communication Design, National Academy of Arts Culture and Heritage, 50480Kuala Lumpur, Malaysia.

Email: ps0340@putra.unisza.edu.my

Abstract. Classifying human face based on race and gender is a vital process in face recognition. It contributes to an index database and eases 3D synthesis of the human face. Identifying race and gender based on intrinsic factor is problematic, which is more fitting to utilizing nonlinear model for estimating process. In this paper, we aim to estimate race and gender in varied head pose. For this purpose, we collect dataset from PICS and CAS-PEAL databases, detect the landmarks and rotate them to the frontal pose. After geometric distances are calculated, all of distance values will be normalized. Implementation is carried out by using Neural Network Model and Fuzzy Logic Model. These models are combined by using Adaptive Neuro-Fuzzy Model. The experimental results showed that the optimization of address fuzzy membership. Model gives a better assessment rate and found that estimating race contributing to a more accurate gender assessment.

1. Introduction

With advances in face recognition technology, there is a need to further expand the discussion of various sub-areas. These discussions contribute to bridging the gap to reach an advanced recognition system. The main challenges facing recognition systems are the pose, illumination, and occlusion. Furthermore, face recognition among huge database is considered an additional incrementally challenge. Although face recognition seeks to identify the corresponding image, providing a facial classification of subjects facilitates identification process. In another word, the classification of facial images across human groups contributes to the indexing of databases. Thus, the search for the target is narrowed during the identification process. In addition, classification process is useful in the process of constructing 3D face model based on a reference depth map. This is due to the conclusion that depth information does not differ much within the same race [1, 2]. Moreover, gender and race classification can provide a demographic classification system.

To create a system that classifies human groups through images, it first requires detection and representation characteristics of face. This is usually done by linear methods such as Principal Component Analysis (PCA) or Linear Discriminate Analysis (LDA). These methods also contribute to process of identification. They compare characteristics between given image and training samples in order to determine the nearest sample. The facial characteristics are given by shape and appearance. The comparison is usually done using measurement criteria such as Mahalanobis and Support Vector Machines (SVM). However, appearance characteristics play a large role in the process of assessing



gender in adults. In contrast, the importance of the shape characteristics is taken into consideration with regard to subjects under the age of 18 years [3].

Combination of linear models was experimented in order to propose a model for classifying the population of Malaysia [4]. The database was created for both genders of Malay, Chinese, and Indians. The experiment was reported to be 90% based on classification rate in face images and it rises to 95.5% in head images. It is also concluded that using linear models in classifying gender provides more accurate results than the classifying race. Generally, non-linear classification methods may be more useful.

Neural networks are utilized to classify a texture data of race of FERET database represented by SVM [5]. The classification rate is reported as 92%, probabilistic Boosting Tree (PBT) is utilized to classify the gender of races: Mongoloid, Caucasian and African [6]. The experimental results show that PBT inclines to classify Mongoloid patterns as females while African patterns as males. This is due to their difference decision boundary of linear model in each race. Furthermore, the classifying gender by PBT on each race separately gave higher accuracy reaches up to 93.71%.

In general, the psychological process of face recognition by the human brain is still somewhat ambiguous. Moreover, humans are less likely to identify individuals across other races [7]. Thus, people rely on facial recognition algorithms according to their race. Therefore, it is useful to index the database in order to get people out of their race. Thus, the identification strategy is enhanced by using algorithms that are more appropriate for each race [8].

In this paper, we utilized non-linear model to classify the gender and race of Asian and Caucasian subjects. The used model is a combination of fuzzy logic and neural network models. Fuzzy logic is among the techniques of artificial intelligence used to indicate truth by inception of Membership Functions MFs. Neural networks are clustering nodes in several layers that are affected by each other Simulates biological neural system. The Adaptive Network based Fuzzy Model (ANFM) is compelled by data, which is known as data driven. It utilizes neural network technique for the arrangement of capacity guess issues. Presumption of the race and gender is achieved with only geometric information. A set of the database is defined as a training set. Then it is clustered into a numeric pattern. Fuzzy rules and memberships are optimally generated to designate the data set. Finally, ANFM approximate proper model for new pattern based on the generated rules.

2. Data Set Preparation

We build data set by combining CAS-PEAL [9] of Asian individuals and PICS [10] of Caucasian individuals. From each database, we chose 600 images under gender and race equality with different three poses. We split the data set into three sub sets; training set, validation set and test set by the ratio 2:1:1. In ANFM training phrase, training set is repeatedly used to adapt the weights, biases and rules. Validation data is also repeatedly used to estimate the non-training performance error of candidate designs. In addition, it is used to indicate and rebuild structures of ANFM design or stop training according the validation error. Test data is used once and only once on the best design to obtain an unbiased estimate for the predicted error of unseen non-training. All types of data set are considered pairs of input and output.

Input data set is a matrix of vectors that reflect geometric distances between landmarks. All landmarks are detected by using Restricted Shape and Abbreviated Appearance Model [11]. It gives more accurate landmark detection because using composite cross-correlation methods [12, 13] in matching term. Based on the landmarks, each vector consists of 12 values; 5 values are vertical distances and 7 values are horizontal distances as shown in Figure 1. To overcome the obstacle of face difference size or distance between the camera and subject during capturing, we have identified a unit of measurement for the normalization of values through it. The unit of measurement is a vertical distance between the left corner of left eye and the left corner of lips γ_0 . Consequently, all input vector values reflect the aspect ratio geometric distances to the unit of measurement as shown in Equation (1).

$$\gamma'_{vi} = \frac{\gamma_{vi}}{\gamma_0} \quad (1)$$

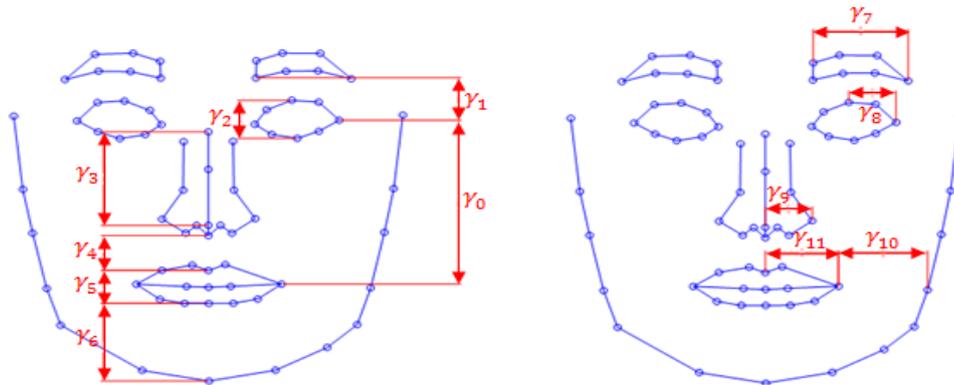


Figure 1. Input data of ANFM from the geometric distances between landmarks.

While horizontal distance between landmarks is changeable through pose, it is useful to mathematically deduce distance value for either frontal or profile pose. Equation (2), Equation (3) and Equation (4) describe the distance through pose θ depending on the value in both frontal and profile pose.

$$\gamma^{P'_{vi}} = \frac{\gamma'_{vi}}{\gamma_0 \sin \theta} - \frac{\gamma^{F'_{vi}}}{\tan \theta} \tag{2}$$

$$\gamma^{F'_{vi}} = \frac{\gamma'_{vi}}{\gamma_0 \cos \theta} - \gamma^{P'_{vi}} \tan \theta \tag{3}$$

$$\gamma'_{vi} = \gamma^{P'_{vi}} \sin \theta + \gamma^{F'_{vi}} \cos \theta \tag{4}$$

3. Forwarding Data through Fuzzy interface system Structure

ANFM is proposed to combine several Fuzzy interface systems (FISs). FIS is structured on several layers. Each layer contains one or more nodes. We utilize FIS for valuating race based on frontal image *FIS_RF*. Other entities of FISs are designed based on the same number of layers but differ in the number of nodes. As shown in Figure 2, there are five layers consist of *FIS_RF*. First layer is input layer with 12 nodes, which is feeding by input data set. Membership function is applied on input data in the second layer *inputmf*. There are five fuzzy sets for each input node. The rule used to explain membership of an input value ($input_i$) with fuzzy set (A_i) is given by Equation (5) whereas $f(A_i, input_i)$ is a MF.

$$Rule : \text{If } input_i \text{ is } A_i \text{ and then } inputmf_i \text{ is } f(A_i, input_i) \tag{5}$$

To simplify, we represented the membership by real values using the premise parameter set as in Equation (6). The parameters set $a_{i,j}$ and $b_{i,j}$ are changeable through training in order to optimize the fuzzy set.

$$inputmf_{i,j} = e^{-\left(\frac{input_i - a_{i,j}}{b_{i,j}}\right)^2} \tag{6}$$

Third layer is considered as rule layer. It has fixed number of nodes correspond to the number of data set. It calculates the strength of the rules for each membership. It is given by Equation (7).

$$Rule_j = \prod_i inputmf_{i,j} \tag{7}$$

The ratio of rules are calculated in the fourth layer *outputmf*. This layer has the same number of node in previous layer. Each node $outputmf_j$ figures the ratio as shown in Equation (8).

$$outputmf_j = \frac{Rule_i}{\sum_j Rule_j} \tag{8}$$

The last layer has only one node represents the output. This node is connected with the nodes in previous layer by weighted links. The overall output of FIS is given by Equation (9).

$$output = \sum_i w_i \times outputmf_j \quad (9)$$

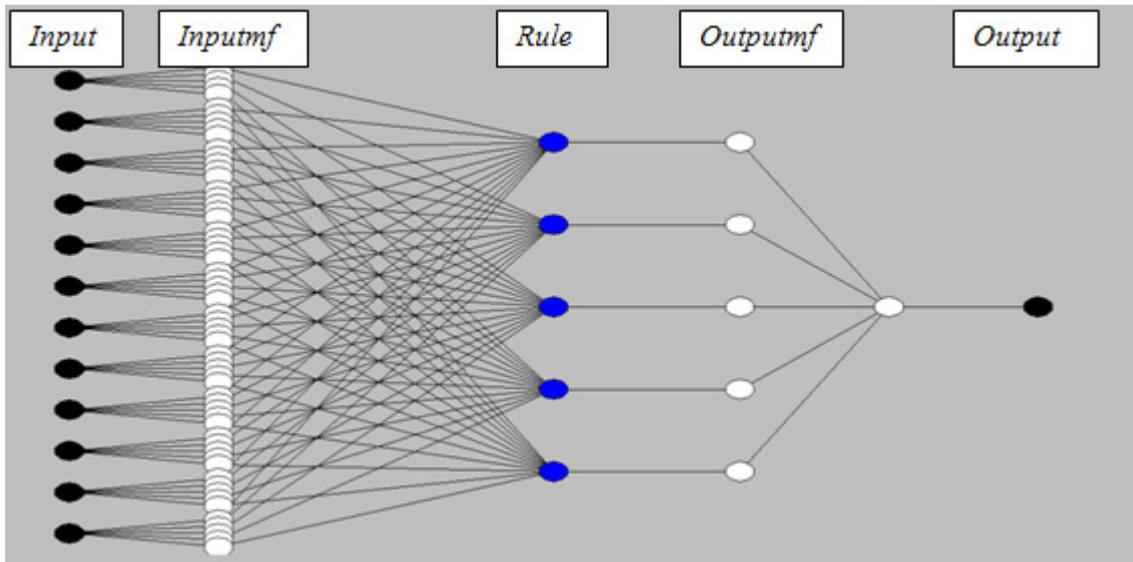


Figure 2. The structure of FIS.

4. Updating Weights and Memberships of FIS

The overall output function of FIS could be represented by a linear combination. In fact, the parameters of the function are the weights and premise parameters; w_i , $a_{i,j}$ and $b_{i,j}$. In training phase, the squared error between the returned output and target is calculated and back propagated. The weights are updated by Equation (10) and Equation (11).

$$\Delta(w_i) = w_i \times (output - target)^2 \quad (10)$$

$$new(w_i) = old(w_i) + \Delta(w_i) \quad (11)$$

On other hand, gradient descent method is efficiently used in order to update the premise parameters $a_{i,j}$ and $b_{i,j}$.

5. Performance of FIS

Prudent representing of the target plays an import role to optimizing FIS, in addition to Input data set and training rules. There are four groups of target, covering race and gender. We built FIS_{RG_F} which are provided by frontal geomantic distances in the input layer. The evaluation of FISs in Figure 3 shows red stars to represent the output and blue dot represents a target sample. The target indicates 1 for female Caucasian, 0.67 for male Caucasian, 0.33 for female Asian and 0 for male Asian. As shown in Figure 3.a, the test set is not accurate especially for Asian gender.

The neuro-fuzzy model FIS_{RG_F} could be divided into two models, FIS_{R_F} to estimate race and FIS_{G_F} to estimate gender. Caucasians are indicated by 1 and Asians are indicated by 0 through the target of FIS_{R_F} . Females are indicated by 1 and Males are indicated by 0 through targets of FIS_{G_F} . As shown in Figure 3.band Figure 3.c, race is easily estimated while gender still has unsatisfactory results. Therefore, the idea is to rebuild FIS_{G_F} in addition to providing the input layer with the estimated race target, which is returned by FIS_{R_F} output. The FIS is indicated by $\overline{FIS_{G_F}}$ and it test results as shown in Figure 3.d.

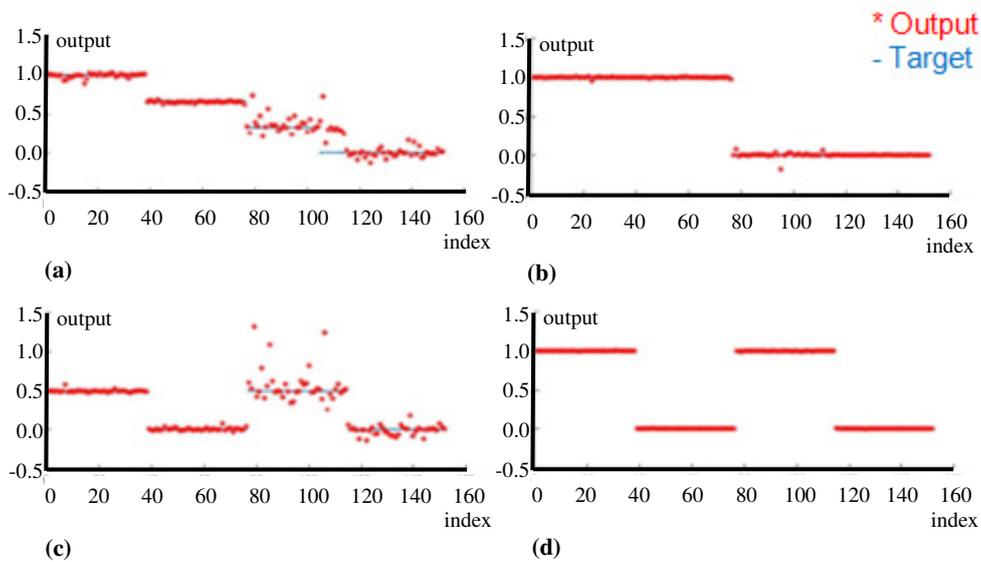


Figure 3. The result of test by training various FISs for estimating race and gender.

Horizontal geometric distances are changeable dramatically between the frontal and profile images of the same subject. Because of that, it is helpful to dedicate FIS to deal with the front pose and the others to deal with the profile pose. Table 1 demonstrates RMS error after training each FIS200 epochs. Results are explaining RMS on training, test, and validation set. We note that the best results are given by the models, which designed to estimate only the race. However, for estimating the gender more accurately, we need a prior result of estimating the race.

Table 1. RMS of various FIS for estimating race and gender after 200 epochs.

FIS	Input data	Output data	RMS		
			Training	Validation	Test
FIS _{RG_F}	Frontal	Race + Gender	0.005690	0.083657	0.095541
FIS _{R_F}	Frontal	Race	0.000489	0.009874	0.018080
FIS _{G_F}	Frontal	Gender	0.019786	0.175846	0.236010
$\overline{\text{FIS}}_{G_F}$	Frontal+FIS _{R_F}	Gender	0.000185	0.001047	0.001279
FIS _{RG_P}	Profile	Race + Gender	0.003857	0.068742	0.086523
FIS _{R_P}	Profile	Race	0.000353	0.007589	0.014856
FIS _{G_P}	Profile	Gender	0.013846	0.097854	0.202587
$\overline{\text{FIS}}_{G_P}$	Profile+FIS _{R_P}	Gender	0.000154	0.000856	0.001021

Finally, the proposed ANFM structure consists of four sub-FIS model as shown in Figure 4. Two sub FIS are designed to deal with frontal shape to predict the race FIS_{R_F} and the gender $\overline{\text{FIS}}_{G_F}$. Other two sub models are designed to deal with profile shape to predict the race FIS_{R_P} and the gender $\overline{\text{FIS}}_{G_P}$. The input of FISs responsible for gender prediction is improved by providing the results of race assessment.

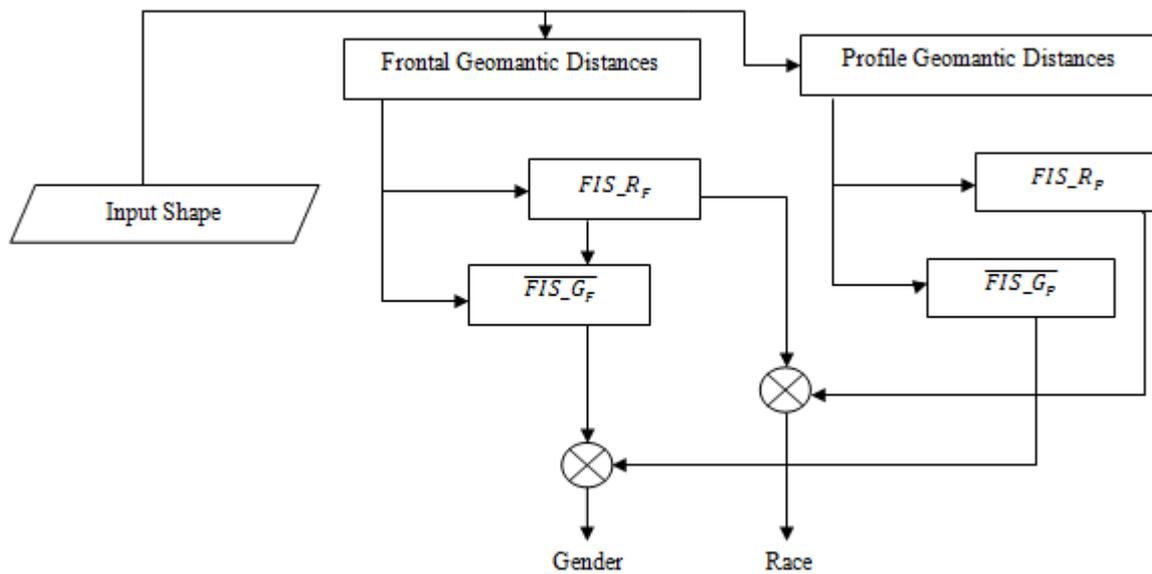


Figure 4. The diagram of the proposed FIS model for estimating race and gender.

6. Discussion and Conclusion

This paper investigates the estimation of race and gender from geometric distances of face across pose face. Based on detecting shape of face images, 12 distances are taken from facial landmarks. Then Frontal and profile distances are calculated in order to provide an input of classification system. The classification process is achieved by ANFM. The used dataset is 600 images taken from CAS and PICS of Asian and Caucasian individuals. After observing several structures of FISs, we concluded that race estimation was easier than gender estimation. For a more accurate gender assessment, we need prior information about the individual's race. The proposed ANFM structure was created based on this conclusion. In this way, we were able to obtain a correct estimate of 100% on the samples of the test and the validation.

However, the previous results were different from some previous study. The present study [4] found that it is easier to classify gender than race on Malaysian dataset. The dataset included samples from three ethnic groups, Malay, Chinese and Indian. In order to compare our results with the study [4], we conducted a study experiment on our database. The rate of classification increased to 97% in the test samples. Furthermore, the race was classified more accurately. The reason for is that gaps between races in our dataset is greater than ethnical groups in Malaysian dataset. Moreover, the use of ANFM is greater in susceptibility than linear models as described in the study [4].

Overall, the results pave way for future work on how to identify races by increasing the size and variety of training samples. Furthermore, it shows the importance of advanced techniques in classify mixed or convergent races.

References

- [1] Heo J and Savvides M. 2012 Gender and ethnicity specific generic elastic models from a single 2D image for novel 2D pose face synthesis and recognition. *IEEE transactions on pattern analysis and machine intelligence*, 34 (12), 2341-2350.
- [2] Moeini A, Faez K, and Moeini H 2015 Real-world gender classification via local Gabor binary pattern and three-dimensional face reconstruction by generic elastic model *IET Image Processing*, 9(8), 690-698.
- [3] Dantcheva A and Brémond F 2017 Gender estimation based on smile-dynamics. *IEEE Transactions on Information Forensics and Security*, 12(3), 719-729.
- [4] Suit C C 2011 *Gender and ethnic classification of Malaysian face images for face recognition* (Doctoral dissertation, Multimedia University (Malaysia)).

- [5] Hsu C W and Lin C J 2002 A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, 13(2), 415-425.
- [6] Gao W and Ai H 2009 Face gender classification on consumer images in a multiethnic environment *International Conference on Biometrics* Springer, Berlin, Heidelberg 169-178
- [7] Marsh B U, Pezdek K and Ozery D H 2016 The cross-race effect in face recognition memory by bicultural individuals *Acta psychologica*, 169, 38-44.
- [8] Sun G, Zhang G, Yang Y, Bentin S and Zhao L 2014 Mapping the time course of other-race face classification advantage: a cross-race ERP study *Brain topography*, 27(5), 663-671.
- [9] Hancock P 2008 Psychological image collection at stirling (pics). *Web address: <http://pics.psych.stir.ac.uk>*.
- [10] Gao W, Cao B, Shan S, Chen X, Zhou D, Zhang X and Zhao D 2008 The CAS-PEAL large-scale Chinese face database and baseline evaluations *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(1), 149-161.
- [11] Nanaa K, Rahman M N A, and Rizon, M 2017 Restricted Shape and Abbreviated Appearance based Correlation Model for Detecting Facial Landmarks *4th International Conference on Innovations in Engineering, Technology, Computers and Industrial Applications IETCIA-2017*.
- [12] Nanaa K, Rizon M, Rahman M N A, Almejrad A, Aziz A Z A and Mohamed S B 2013 Eye detection using composite cross-correlation *American Journal of Applied Sciences*, 10(11), 1448
- [13] Nanaa K, Rizon M and Rahman M N A 2014 Eye detection Using Composite Cross-Correlation form Face Images in Varied Illumination *The International Conference on Artificial Life and Robotics*.