*International Journal of Computer Applications (0975 – 8887)*
*Volume 59– No.3, December 2012*

# A Neural Network based Facial Expression Recognition using Fisherface

Zaenal Abidin
Department of Mathematics
Semarang State University
Semarang, 50229, Indonesia

Agus Harjoko
Department of Computer Science and Electronics
Gadjah Mada University
Yogyakarta, 55281, Indonesia

## ABSTRACT
Facial expression plays significant role for human beings to communicate their emotions. Automatic facial expression analysis is a flourishing area of research in computer science, and it is also still a challenge. This paper discusses the application of a natural network based facial expression recognition using fisherface. Back propagation neural network is used as a classifier for classifying the expressions. For face portion segmentation and localization, integral projection method is used. The accuracy of system performance have evaluated on a public database "Japanese Female Facial Expression (JAFFE)". The experimental results show the effectiveness of our scheme. The best average recognition rate achieves 89.20%.

## General Terms
Image Processing, Pattern Recognition

## Keywords
Facial Expression, Fisherface, Neural Network, JAFFE.

## 1. INTRODUCTION
Facial expressions provide an important behavioral measure for the study of cognitive activity, intention, and personality; it not only expresses our emotions, but also provides important communicative cues during social interaction. Mehrabian reported that 7% of the meaning of a communication from the words that are spoken, 38% by paralanguage, the way that the words are said, and 55% of message is communicated by facial expression; hence recognition of facial expressions became a major modality in Human-Computer Interaction (HCI) [6],[7].

System of facial expression recognition (FER) consists of three basic components: face detection, feature extraction, and facial expression recognition. Face detection used to classify face and non-face areas. The face detection step provides us with a rectangle head boundary which localize of face. In this work, integral projection method is used to localize and segment of face. The feature extraction is an important key to the whole recognition process. If inadequate features are used, the system of FER could fail to achieve accurate recognition. In this case, fisherface method is proposed to get the feature, and then neural network will be used as classifier to recognize facial expression.

There are many approaches have been proposed for facial expression recognition, namely Gabor feature using a novel local Gabor filter bank [5], hierarchical radial basis function network (HRBFN) based on local features extraction by PCA technique [3], Gabor filter based feature extraction in combination with learning vector quantization (LVQ) [1], Gabor wavelet transform (GWT) and Radial Basis Function (RBF) Network [10], Higher-Order Local Autocorrelation (HLAC) coefficients and Local Binary Pattern (LBP) [9].

In this paper we are focus on facial expression recognition from static images with single person. Six basic facial expressions which correspond to distinct universal emotions: disgust, sadness, happiness, fear, anger, and surprise, plus the neutral expression will be evaluated. Experiments are performed on the Japanese Female Facial Expression (JAFFE) database. The JAFFE database consists of 213 grayscale images of Japanese women posing the six basic expressions plus a neutral one [11]. This database was created with a help of 10 Japanese female models. A sample of images from JAFFE database is shown in Figure 1.

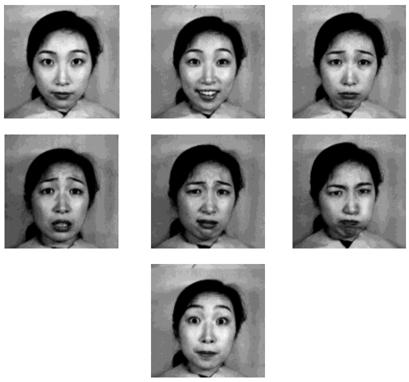

**Fig 1: Sample of static images from JAFFE database**

## 2. IMAGE PRE-PROCESSING
The image pre-processing procedure is a very important phase in the facial expression recognition to obtain images which have normalized intensity, uniform size and shape, and depict only a face expressing certain emotion. The face area of an image is detected using integral projection method.

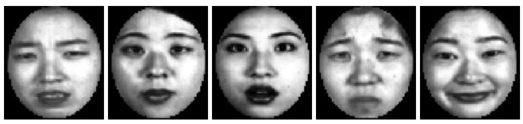

**Fig 2: Sample of images after pre-processing**

The procedure of pre-processing consists of three steps: 1) Uniform size into 33×29 pixels; 2) increase the contrast of images using histogram equalization; 3) masking process by cropping out the four edge corner of face images. Figure 2 shows sample of images after pre-processing.

## 3. INTEGRAL PROJECTION METHOD
Integral projections have already been used in problems of face detection. This method is used to find region or location of object [4],[13]. An integral projection is obtained through the

sum of given set of pixels along a given direction. Suppose the intensity of each pixel $(x, y)$ is $I(x, y)$. The vertical and horizontal integral projection function denoted by $S_y(x)$ and $S_x(y)$ respectively, are discrete and finite 1-D signal given by

$$S_y(x) = \sum_{y=1}^{x} I(x, y) \quad (1)$$

$$S_x(y) = \sum_{x=1}^{y} I(x, y) \quad (2)$$

In this paper, integral projection method used to face detection. The illustration of the use of this method is shown in Figure 3.

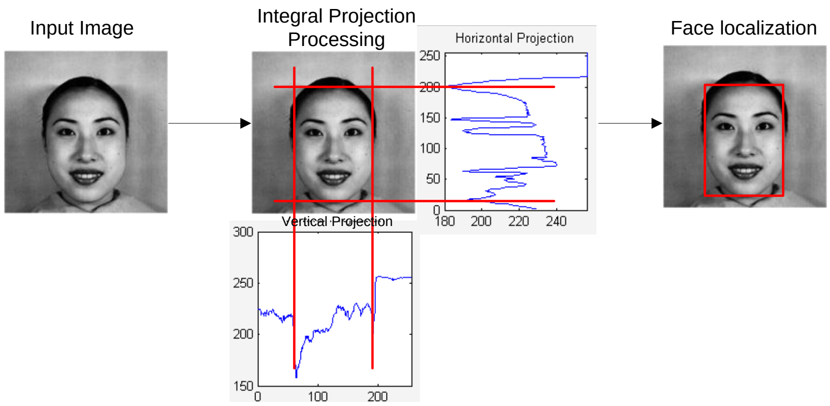


**Fig 3: Illustration of the use of integral projection method for face detection**

## 4. CONSTRUCTION OF FISHERFACE

In this section, we briefly review Fisherface. Fisherface algorithm is developed to determine a matrix that maximizes the ratio of the between-class scatter to the within-class scatter [2]. This method is a well-known technique in classification, discriminant analysis, and it is also commonly used to feature extraction and dimensionality reduction in pattern recognition.

Suppose that the within-class and between-class scatter matrices denoted by $S_W$ and $S_B$ respectively, are defined as follows [12]:

$$S_W = \sum_{i=1}^{C} \sum_{x=x_i} (x - \mu_i)(x - \mu_i)^T \quad (3)$$

where $x$ is the sample of class $X_i$, $\mu_i$ is the mean of class $X_i$ and $C$ is the number of class.

$$S_B = \sum_{i=1}^{N} n_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (4)$$

where $\mu$ represents the mean of all the images and $n_i$ is the number of images in the class $X_i$.

If $S_W$ is non-singular, the project matrix $W_f$ can be chosen as follows:

$$W_f = \arg max_W \frac{|W^T S_B W|}{|W^T S_W W|} \quad (5)$$

The solution to this problem is given by the generalized eigenvalue decomposition:

$$S_B W = S_W W \Lambda \quad (6)$$

where $W$ is the matrix of eigenvectors and $\Lambda$ is a diagonal matrix of corresponding eigenvalues.

The eigenvectors of $W$ associated to non-zero eigenvalues are the Fisherfaces. There is a maximum of $C - 1$ Fisherfaces. This can be readily seen from the definition$S_B$. Note that in our definition, $S_B$ is a combination of $C$ feature vectors. Any $C$ vectors define a subspace of $C - 1$ or fewer dimensions. The equality holds when these vectors are linearly independent from one another.

## 5. LEARNING OF BACKPROPAGATION NEURAL NETWORK

A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationship. A neural network is represented by weighted interconnections between processing elements. These weights represent information being used by the net to solve a problem and they are actually parameters which are defined as the non-linear function performed by the neural network. Backpropagation is the most popular learning algorithms for multi-layer perceptron. The backpropagation algorithm defines a systematic way to update the synaptic weights of multi-layer perceptron (MLP) networks. The supervised learning is based on the gradient descent method, minimizing the global error on the output layer. The learning algorithm is performed in two stages: feed-forward and feed-backward. In the first phase the inputs are propagated through the layers of processing elements, generating an output pattern in response to the input pattern presented. In the second phase, the errors calculated in the output layer are then back propagated to the hidden layers where the synaptic weights are updated to reduce the error. This learning process is repeated until the output error value, for all patterns in the training set, are below a specified value [6].

In this work, the neural network architecture that used to solve the problem consists of three layers. These are an input, a hidden, and an output layer. The number of neurons in input layer is 6 with adding a bias neuron. For the hidden layer, number of neurons was varied to get optimum result. The variations are 12, 60, and 120 neurons plus one bias. The last layer is output layer has 7 neurons in accordance with the amount of target researched. The activation function for hidden layer is sigmoid bipolar, while for output layer is sigmoid, because the output value is expected to be in the range [0, 1]. The output of each neuron is converted into binary.

Learning rate is also varied (0.25 and 0.5). The target of error and maximum epoch are 0.0001 and 1000 respectively. Table 1 provides information about the configuration of neural network output and its interpretation.

**Table 1. Configuration of neural network output and its interpretation**

| Node 1 disgust | Node 2 Fear | Node 3 Surprise | Node 4 Anger | Node 5 Sad | Node 6 Happy | Node 7 Neutral | NN Output |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | Disgust |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | Fear |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | Surprise |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | Anger |
| 0 | 0 | 0 | 0 | 1 | 0 | 0 | Sad |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | Happy |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | Neutral |

# 6. EXPERIMENTAL RESULTS

The proposed method is evaluated in terms of its recognition performance using the JAFFE female facial expression database [11]. Two facial expression images of each expression of each subject were selected as training samples, while the remaining samples were used as test data. We have 140 training images and 73 testing images for each trial. In our experiments, the feed forward neural network is then used to classify the facial expression images. Experiments were performed with Fisherface feature.

Performance evaluation of facial expression recognition system is using two parameters, namely recognition rate and false positive rate. Recognition rate is the ratio between the numbers of successful facial expressions recognized correctly by the total number of existing facial expression image, while false positive rate is the number of positive test results of facial expression images that are not recognized by the expression should be.

Process of face detection on the system was automated. Integral projection method has been used to make face localization. Afterwards, face was cropped and size of image was normalized into 33×29 pixels, then performed histogram equalization, and masked to cover the corners of the image so that the variations that arise in these parts can be reduced.

After registering all training set images, optimal neural network architecture was determined. Table 2 shows the testing result by varying the neural network parameters. Based on the table, it can be seen that the number of neurons in hidden layer influences the training time. The best performance of experiment is 86.85%, and ratio of the result between training set and testing set are 100% and 61.64% respectively. Neural network architecture will perform optimum if the number of neurons hidden layer is 12 neurons, and the learning rate is 0.5. So, the neural network architecture was chosen to be implemented in the facial expression recognition system.

**Table 2. Testing result by varying the neural network parameters**

| Number of neurons hidden layer | $\alpha$ | Result | | | |
|---|---|---|---|---|---|
| | | MSE | Number of Images Successfully Recognized | | Recognition rate (%) |
| | | | Training Set | Testing Set | |
| 12 | 0.25 | 9,98752e-005 | 140 | 42 | 85,45 |
| 60 | 0.25 | 9,92885e-005 | 140 | 43 | 85,92 |
| 120 | 0.25 | 9,84778e-005 | 140 | 39 | 84,04 |
| 12 | 0.5 | 9.99087e-005 | 140 | 45 | 86,85 |
| 60 | 0.5 | 9.91564e-005 | 140 | 36 | 82,63 |
| 120 | 0.5 | 5.46248e-005 | 140 | 38 | 83,57 |

**Table 3. Confusion matrix of facial expression recognition of training set**

| I\O | Neutral | Angry | Happiness | Sadness | Disgust | Fear | Surprise |
|---|---|---|---|---|---|---|---|
| Neutral | **20** | 0 | 0 | 0 | 0 | 0 | 0 |
| Angry | 0 | **20** | 0 | 0 | 0 | 0 | 0 |
| Happiness | 0 | 0 | **20** | 0 | 0 | 0 | 0 |
| Sadness | 0 | 0 | 0 | **20** | 0 | 0 | 0 |
| Disgust | 0 | 0 | 0 | 0 | **20** | 0 | 0 |
| Fear | 0 | 0 | 0 | 0 | 0 | **20** | 0 |
| Surprise | 0 | 0 | 0 | 0 | 0 | 0 | **20** |

The confusion matrices of the performances of experimental result between training and testing set are recorded in Table 3 and 4.

All expression of training set successfully recognized, meanwhile not all expressions of testing set were equally well recognized by the system. The confusion matrix in Table 4 informs that expression of surprise was often misclassified by others. Similarly, disgust and angry expressions were also misclassified. The false positive rate for each expression is summarized in Table 5.

**Table 4. Confusion matrix of facial expression recognition of testing set**

| I\O | Neutral | Angry | Happiness | Sadness | Disgust | Fear | Surprise |
|---|---|---|---|---|---|---|---|
| Neutral | **8** | 0 | 0 | 0 | 0 | 0 | 2 |
| Angry | 1 | **6** | 0 | 0 | 0 | 0 | 2 |
| Happiness | 0 | 2 | **7** | 0 | 0 | 2 | 0 |
| Sadness | 0 | 0 | 0 | **8** | 1 | 2 | 0 |
| Disgust | 1 | 0 | 0 | 1 | **5** | 1 | 0 |
| Fear | 0 | 1 | 0 | 2 | 1 | **7** | 0 |
| Surprise | 3 | 0 | 0 | 0 | 3 | 0 | **4** |

**Table 5. False positive rate for every expression**

| Facial Expression | False positive rate |
|---|---|
| Neutral | 5 |
| Angry | 3 |
| Happiness | 0 |
| Sadness | 3 |
| Disgust | 5 |
| Fear | 5 |
| Surprise | 4 |
| **Total** | **25** |

# 7. CONCLUSIONS

In this paper, a neural network based facial expression recognition using fisherface feature has been designed and implemented. Fisherface construction can be used to extract the feature of facial expression and a neural network is used to recognize the facial expression. To support performance of system, integral projection method has been proposed for face detection. But this method is very sensitive with the illumination and noise. Images with varying background are not appropriate to use this method. The proposed system achieves a recognition rate of 86.85%. Total number of images which have been trained and successfully recognized is 140 images; whereas images were not trained that successfully recognized are 45 images. In fact, facial expressions recognition is a challenging task. The appearance of an expression of person may vary. Furthermore, some expressions are ambiguous, such as sadness, disgust and fear. The difference between such expressions is hard to discriminate. This is a very interesting thing and may be the important clues for the future research to improve performance of recognition.

# REFERENCES

[1] Bashyal, S. and Venayagamoorthy, G.K. 2008. "Recognition of Facial Expressions Using Gabor Wavelets and Learning Vector Quantization", Engineering Application of Artificial Intelligence, Available at http://dx.doi.org/10.1016/j.engappai.2007.11.010, [Retrieved on 2 January 2010].

[2] Belhumeur, P.N. Hespanha, J.P. and Kriegman, D.J. 1997. "Eigenfaces vs. Fisherfaces: Recognition Using Class-Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720.

[3] Daw-Tung Lin. 2006 . "Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network", Journal of Information Science and Engineering, Vol. 22, pp.1033-1046.

[4] García-Mateos, G. Ruiz, A., and Lopez-de-Teruel, P.E. 2002. "Face Detection Using Integral Projection Models", in 9th International Workshop on Structural Syntactic Pattern Recognition, Windsor (Canada).

[5] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, and Jian-Cheng Huang. 2005. "A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA", International Journal of Information Technology, Vol. 11, No. 11, pp. 86-96.

[6] Khandait, S.P., Thool, R.C. and Khandait, P.D. 2011. "Automatic Facial Feature Extraction and Expression Recognition based on Neural Network", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 2, No.1, pp. 113-118.

[7] Kumar, R. Anjum, B. and Ghiday, Y. 2011. "Role of Communication in Management Teaching & Research in the International Context", Research World - Journal of Arts Science & Commerce, Vol. II, No. 1, pp. 70-77.

[8] Lajevardi, S.M. and Hussain, Z.M. 2009. "Local Feature Extraction Methods for Facial Expression Recognition", in 17th European Signal Processing Conference (EUSIPCO 2009) Glasgow, Scotland, pp. 60-64.

[9] Lajevardi, S.M. and Hussain, Z.M. 2009. "Feature Extraction for Facial Expression Recognition based on Hybrid Face Regions", Advances in Electrical and Computer Engineering, Vol. 9, No. 3, pp. 63-67.

[10] Lekshmi, V.P. and Sasikumar, M. 2008. "A Neural Network Based Facial Expression Analysis using Gabor Wavelets", World Academy and Science, Engineering and Technology, Vol. 42, pp. 563-567.

[11] Lyons, M.J. Akamatsu, S. Kamachi, M., and Gyoba, J. 1998. “Coding Facial Expressions with Gabor Wavelets”, in the Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200–205.

[12] Mazanec, J. Melisĕk, M. Oravec, M., and Pavlovičová, J. 2008. “Support Vector Machines, PCA and LDA in Face Recognition”, Journal of Electrical Engineering, Vol. 59, No. 4, pp. 203-209

[13] Zhao, Y., Xiaojun Shen, Georganas, N.D., and Petriu, E.M. 2009. “Part-based PCA for Facial Feature Extraction and Classification”, in IEEE International Workshop on Haptic Audio Visual Environments and Games 2009. Issue date: 7-8 November, pp. 99 – 104.