Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor

Tran Son Hai  
Informatics Technology Department, University of Pedagogy, Ho Chi Minh City, Vietnam, member of IACSIT  
Email: haits@hcmup.edu.vn

Le Hoang Thai  
Computer Science Department, University of Science, Ho Chi Minh City, Vietnam  
Email: lhthai@fit.hcmus.edu.vn

Nguyen Thanh Thuy  
University of Technology, Ha Noi City, Vietnam  
Email: nguyenthanhthuy@vnu.edu.vn

Abstract — Facial Expression is a key component in evaluating a person's feelings, intentions and characteristics. Facial Expression is an important part of human-computer interaction and has the potential to play an equal important role in human-computer interaction. The aim of this paper is bring together two areas in which are Artificial Neural Network (ANN) and K-Nearest Neighbor (K-NN) applying for facial expression classification. We propose the ANN_KNN model using ANN and K-NN classifier. ICA is used to extract facial features. The ratio feature is the input of K-NN classifier. We apply ANN_KNN model for seven basic facial expression classifications (anger, fear, surprise, sad, happy, disgust and neutral) on JAFEE database. The classifying precision 92.38% has been shown the feasibility of our proposal model.

Index Terms — Facial Expression Classification, Artificial Neural Network (ANN), K-Nearest Neighbor (K-NN), Independent Component Analysis (ICA)

I. INTRODUCTION

A facial expression is an expression of an emotional state can be seen, cognitive activities, intent, character of a person [1]. Mehrabian [2] showed that while emotional communication, 55% of the message conveyed by facial expression, while only 7% by linguistic language and 38% by intonation. Facial Expressions are an important component of human-computer communication. There are many approaches for facial expression classification problem such as: using AdaBoost, K-NN, Support Vector Machine (SVM) and Artificial Neural Network (ANN).

AdaBoost is a weak classifier and can be used in conjunctions with another machine learning algorithm to improve their performance. The k-nearest neighbor (k-NN) is a common tool in image classification, but its sequential implementation is slowly and requires the high calculating costs because of the large representation space of images.

SVM applies for pattern classification even with large representation space. In this approach, we need to define the hyper-planes for pattern classification. The number of hyper-plans is ratio to the number of classes. Unseen elements can be identified by calculating posterior probabilities in 1-vs-1/1-vs-rest binary SVMs (Platt's method) or in multiclass SVMs [3].

ANN will be trained with the patterns to find the weight collection for the classification process. ANN overcomes this weakness of SVM of using suitable threshold in the classification for outside pattern. If the patterns do not belong any in L given classes, ANN identify and report results to the outside given classes. Thus, ANN has been applied for facial expression classification by many researchers [4, 10, 11, 12, and 13].

Besides, some fusion model using multi soft computing techniques for classifying. For example, ANN_SVM [4] has been brought together Artificial Neural Network and Support Vector Machine applying for facial expression classification.

The aim of this paper is bring together two areas in which are Artificial Neural Network (ANN) and K-Nearest Neighbor (K-NN) applying for facial expression classification. Thus we propose a model for Facial Expression Classification using ANN and KNN below:

In the Fig. 1 model, we use two classifiers ANN and K-NN. The input of ANN classifier is ICA feature extraction. The input of KNN classifier is the geometrical features and their ratios. Finally, we combine the conclusions of ANN classifying and K-NN classifying to give the final classifying result.

Firstly, Facial Expression is classified by ANN. In second phase, we calculate width of left eye, width of right eye, width of mouth, distance from left eyebrow to right eyebrow, height of left eye, height of right eye, height of mouth, distance from left eye to left eyebrow, distance from right eye to right eyebrow. After that, we use k-mean to classify Facial Expression base ratio of distance between them.
II. LOCAL FEATURE EXTRACTION

Local feature analysis in facial expression is very significant for facial expression analysis. Canny is used to detect local region such as: left-right eyes, mouth, left-right eyebrows [5]. In first phase, we use ICA to present local features in small presenting space. Our local feature extraction below:

After facial image had equalized histogram, it would be standardized to 30x30 size and map to feature space:

\[ x_{face} = (x_1, x_2, \ldots, x_{900}) \]

\[ y_{face} = (y_1, y_2, \ldots, y_K) \]  

Where K: the number dimension of feature space. The above vector \( y_{face} \) is the global feature vector. Detecting and extracting eyebrows, eyes and mouth features, we get 5 local feature vectors \( (y_{left\_eyebrow}, y_{left\_eye}, y_{right\_eye}, y_{right\_eyebrow}, y_{mouth}) \) below:

\[ x_{left\_eyebrow} = (x_1, x_2, \ldots, x_{450})^T \]
\[ x_{right\_eyebrow} = (x_1, x_2, \ldots, x_{450})^T \]
\[ x_{left\_eye} = (x_1, x_2, \ldots, x_{450})^T \]
\[ x_{right\_eye} = (x_1, x_2, \ldots, x_{450})^T \]
\[ x_{mouth} = (x_1, x_2, \ldots, x_{800})^T \]  

Combining local and global feature vector:

\[ y_{comp} = \begin{bmatrix} y_{face} \\ y_{left\_eyebrow} \\ y_{right\_eyebrow} \\ y_{left\_eye} \\ y_{right\_eye} \\ y_{mouth} \end{bmatrix} \]  

Fig. 1. Facial Expression Classification Process

Fig. 2. Facial Feature Extraction Process

Fig. 3. Eyebrow, eye and mouth detection

The local features are projected to K dimensions feature space. There are 5 local feature vectors below:

\[ y_{left\_eyebrow} = (y_1, y_2, \ldots, y_K)^T \]
\[ y_{right\_eyebrow} = (y_1, y_2, \ldots, y_K)^T \]
\[ y_{left\_eye} = (y_1, y_2, \ldots, y_K)^T \]
\[ y_{right\_eye} = (y_1, y_2, \ldots, y_K)^T \]
\[ y_{mouth} = (y_1, y_2, \ldots, y_K) \]
Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor

III. FACIAL EXPRESSION CLASSIFICATION USING ANN AND K-NN

A. Independent Component Analysis for feature extraction

Global and local features are extracted by ICA [6] in order to reduce the number dimensions of feature space:

\[ Y_{\text{ICA}} = ICA(Y_{\text{Face}}) \]

\[ Y_{\text{ICA}} = ICA(Y_{\text{Left Eyebrow}}) \]

\[ Y_{\text{ICA}} = ICA(Y_{\text{Right Eyebrow}}) \]

\[ Y_{\text{ICA}} = ICA(Y_{\text{Left Eye}}) \]

\[ Y_{\text{ICA}} = ICA(Y_{\text{Right Eye}}) \]

B. Artificial Neural Network for facial expression classification

Multi-Layer Perceptron (MLP) [7] is a function

\[ \hat{y} = MLP(x,W), \]

where

\[ x = (x_1, x_2, \ldots, x_n) \]

\[ \hat{y} = (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_m) \]

W is the set of parameters \[ \{w_{ij}^L, w_{li}^L\}, \forall i,j,L \]

The MLP uses the algorithm of Gradient Back-Propagation for training to update \( W \) and transfer:

\[ f(t) = \frac{1}{1 + e^{-t}} \]  \hspace{1cm} (6)

Number of input layer neurons: \( n = 200 \) input nodes corresponding to the total dimension of five feature vectors in \( V \) set. In the input layer (\( L = 0 \)):

\[ y_i^0 = x_i, \text{where } i = 1..200 \]

Number of hidden layers: 1. Number of hidden layer neurons: will be identified based on experimental result. In the hidden layer (\( L = 1 \)):

\[ \hat{y}_j^1 = \sum_j y_j^{L-1} w_{ij}^L + w_{i0}^L \text{ and } y_j^1 = f(\hat{y}_j^1) \]  \hspace{1cm} (7)

Number of output layer neurons: \( m = 7 \) output nodes corresponding to seven basic facial expression analyses: anger, fear, surprise, sad, happy, disgust and neutral. The first output node gives the probability assessment belong anger. In the output layer (\( L = 2 \)):

\[ \hat{y}_l^2 = \hat{y}_l, \text{where } l = 1..7 \]  \hspace{1cm} (8)

Table 1. Output node corresponding to anger, fear, surprise, sad, happy, disgust and neutral

<table>
<thead>
<tr>
<th>Feeling</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>Y1</td>
</tr>
<tr>
<td>Fear</td>
<td>Y2</td>
</tr>
<tr>
<td>Surprise</td>
<td>Y3</td>
</tr>
<tr>
<td>Sad</td>
<td>Y4</td>
</tr>
<tr>
<td>Happy</td>
<td>Y5</td>
</tr>
<tr>
<td>Disgust</td>
<td>Y6</td>
</tr>
<tr>
<td>Neutral</td>
<td>Y7</td>
</tr>
</tbody>
</table>
C. K-NN Classification using Ratios local features

Facial ratios features: The primary facial features are located to compute the ratios for facial expression classification. Nine ratios are calculated for facial face database comprising anger, fear, surprise, sad, happy, disgust and neutral. Figure 7 gives the ratios of L_eyebrow to R_eyebrow distance, width of L_eye, width of R_eye, width of mouth, L_eyebrow to L_eye distance, R_eyebrow to R_eye distance, height of L_eye, height of R_eye and height of mouth for feature extraction.

\[
\text{ratio}_1 = \frac{\text{L_eyebrow to R_eyebrow distance}}{\text{width of L_eye}} \\
\text{ratio}_2 = \frac{\text{L_eyebrow to R_eyebrow distance}}{\text{width of R_eye}} \\
\text{ratio}_3 = \frac{\text{L_eyebrow to R_eyebrow distance}}{\text{height of mouth}} \\
\text{ratio}_4 = \frac{\text{height of L_eye}}{\text{height of mouth}} \\
\text{ratio}_5 = \frac{\text{height of R_eye}}{\text{height of mouth}} \\
\text{ratio}_6 = \frac{\text{height of L_eye}}{\text{height of mouth}} \\
\text{ratio}_7 = \frac{\text{height of R_eye}}{\text{height of mouth}} \\
\text{ratio}_8 = \frac{\text{height of L_eye}}{\text{height of mouth}} \\
\text{ratio}_9 = \frac{\text{height of R_eye}}{\text{height of mouth}} \tag{9}
\]

The ratios feature is the input of K-NN classifying method using Mahalanobis distance measure.

D. Fusion of ANN and KNN for Facial Expression Classification

The output of ANN classifier:

\[
y_{\text{ANN}}(y_{\text{ANN}_1}, y_{\text{ANN}_2}, y_{\text{ANN}_3}, y_{\text{ANN}_4}, y_{\text{ANN}_5}, y_{\text{ANN}_6}, y_{\text{ANN}_7})
\]

The output of K-NN classifier:

\[
y_{\text{KNN}}(y_{\text{KNN}_1}, y_{\text{KNN}_2}, y_{\text{KNN}_3}, y_{\text{KNN}_4}, y_{\text{KNN}_5}, y_{\text{KNN}_6}, y_{\text{KNN}_7})
\]

The final output of integrated classifier is computed by using minimum function below:

\[
y_{\text{final}} = \min (1-y_{\text{ANN}}, y_{\text{KNN}}) \\
y_{\text{final}_l} = \min (1-y_{\text{ANN}_l}, y_{\text{KNN}_l}) \quad (10)
\]

Where, l=1..7 corresponding to seven basic facial expression analyses: anger, fear, surprise, sad, happy, disgust and neutral.

Table 2. Output node corresponding to seven basic facial expressions

<table>
<thead>
<tr>
<th>Feeling</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>(y_{\text{final}_1})</td>
</tr>
<tr>
<td>Fear</td>
<td>(y_{\text{final}_2})</td>
</tr>
<tr>
<td>Surprise</td>
<td>(y_{\text{final}_3})</td>
</tr>
<tr>
<td>Sad</td>
<td>(y_{\text{final}_4})</td>
</tr>
<tr>
<td>Happy</td>
<td>(y_{\text{final}_5})</td>
</tr>
<tr>
<td>Disgust</td>
<td>(y_{\text{final}_6})</td>
</tr>
<tr>
<td>Neutral</td>
<td>(y_{\text{final}_7})</td>
</tr>
</tbody>
</table>
In the above Table, $Y_{\text{final,1}}$ represents the ability of the facial image in the class ‘anger’. $Y_{\text{final,2}}$ represents the ability of the facial image in the class ‘fear’. $Y_{\text{final,1}}$ represents the ability of the facial image in the class ‘neutral’.

IV. EXPERIMENTAL RESULTS

We apply our proposal method for seven basic facial expressions on JAFEE database consisting 213 images posed by 10 Japanese female models. The result of classification sees the table below:

<table>
<thead>
<tr>
<th>Feeling</th>
<th>ICA_ANN</th>
<th>Ratios_KNN</th>
<th>ICA_ANNN Ratios_KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger</td>
<td>93.33%</td>
<td>93.33%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Fear</td>
<td>86.67%</td>
<td>86.67%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Surprise</td>
<td>93.33%</td>
<td>86.67%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Sad</td>
<td>93.33%</td>
<td>93.33%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Happy</td>
<td>93.33%</td>
<td>86.67%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Disgust</td>
<td>86.67%</td>
<td>93.33%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Neutral</td>
<td>93.33%</td>
<td>93.33%</td>
<td>93.33%</td>
</tr>
<tr>
<td>Average</td>
<td>91.43%</td>
<td>90.48%</td>
<td>92.38%</td>
</tr>
</tbody>
</table>

In the above Table, $Y_{\text{final,1}}$ represents the ability of the facial image in the class ‘anger’. $Y_{\text{final,2}}$ represents the ability of the facial image in the class ‘fear’. $Y_{\text{final,1}}$ represents the ability of the facial image in the class ‘neutral’.

V. CONCLUSION

In this paper, we propose a model for facial expression classification. The facial image is extracted to six feature vectors (one global feature representing the whole face and five local feature vectors representing the eyebrow, eye and mouth of the face). All feature vectors prossessed by ICA are the input of ANN classifier. The distance ratios of the local region of face are the input of K-NN classifier. The minimum function used to combine the output of ANN and K-NN classifier. Our proposed model ANN_KNN uses ANN and KNN with the suitable feature for classifying.

We apply ANN_KNN model for seven basic facial expression classification (anger, fear, surprise, sad, happy, disgust and neutral) on JAFEE database. The classifying precision is 92.38%. This experimental result has been showed the feasibility of our proposal model.

REFERENCES

Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor


How to cite this paper: Tran Son Hai, Le Hoang Thai, Nguyen Thanh Thuy, “Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor”, International Journal of Information Technology and Computer Science (IJITCS), vol. 7, no. 3, pp. 27-32, 2015. DOI: 10.5815/ijitcs.2015.03.04

Authors’ profiles

Tran Son Hai is a member of IACSIT and received B.S degree and M.S degree in Ho Chi Minh University of Natural Sciences, Vietnam in 2003 and 2007. From 2007-2010, he has been a lecturer at Faculty of Mathematics and Computer Science in University of Pedagogy, Ho Chi Minh city, Vietnam. Since 2010, he has been the dean of Information System department of Informatics Technology Faculty and a member of Science committee of Informatics Technology Faculty. His research interests include soft computing pattern recognition, and computer vision. Mr. Tran Son Hai is co-author of eight papers in the international conferences and national conferences.

Prof. Dr. Le Hoang Thai received B.S degree and M.S degree in Computer Science from Hanoi University of Technology, Vietnam, in 1995 and 1997. He received Ph.D. degree in Computer Science from Ho Chi Minh University of Sciences, Vietnam, in 2004. Since 1999, he has been a lecturer at Faculty of Information Technology, Ho Chi Minh University of Natural Sciences, Vietnam. His research interests include soft computing pattern recognition, image processing, biometric and computer vision. Dr. Le Hoang Thai is co-author over thirty five papers in international journals and international conferences.

Prof. PhD. Nguyen Thanh Thuy received B.S degree in Mathematics, and Ph.D. degree in Computer Science from Hanoi University of Technology, Vietnam, in 1982 and 1987. He has been the professor of Vietnam since 2010. Now he is a Vice Rector of VNU University of Engineering and Technology, Ha Noi city, Vietnam. He majors in Machine Learning, Intelligence Computing and Computer Science.