International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 1 (2016) pp 650-655
© Research India Publications. http://www.ripublication.com

Improving the Classification Accuracy of Emotion Recognition using Facial Expressions

Ch. Satyananda Reddy
Andhra University, Visakhapatnam,
Andhra Pradesh, India.
E-mail: satyanandau@yahoo.com

T. Srinivas
GITAM University, Visakhapatnam,
Andhra Pradesh, India.
E-mail: tadisrinivas_reddy@hotmail.com

Abstract: The science of image processing helps to recognize the human gesture for general life applications. Human gestures can be identified by observing the different movements of eyes, mouth, nose and hands. The face is a rich source of information about human behavior. The proposed method of facial expression recognition system is based on PCA and Neural Networks, to recognize the facial expression from a well captured image by means of extracting the features of face. This paper presents classification accuracy of neural network with principal component analysis (PCA) for feature selections in emotion recognition using different facial expressions. Dimensionality reduction of a feature set is a common preprocessing step used for pattern recognition and classification applications. From the experimental results it is concluded that, neural networks with PCA is effective in emotion recognition using facial expressions, in which it is attained a recognition rate of approximately 85% when testing six emotions on benchmark image data set.

Keywords: Emotion recognition, Feature selection, Facial Expression, Neural Network, PCA.

Introduction

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate expression of their emotions and intentions. The face can express emotion sooner than people verbalize or even realize their feelings. The need for reliable recognition and identification of facial expressions is obvious. Mehrabian [1] pointed out that 7% of human communication information is communicated by linguistic language (verbal part), 38% by paralanguage (vocal part) and 55% by facial expression. Therefore facial expressions are the most important information for emotions perception in face to face communication. For classifying facial expressions into different categories, it is necessary to extract important facial features which contribute in identifying proper and particular expressions.

Recognition and classification of human facial expression by computer is an important issue to develop automatic facial expression recognition system in vision community. Further facial expressions can be ambiguous. They have several possible interpretations. Facial expression recognition deals with classification of facial motion and facial feature deformation in to abstract classes that are purely based on visual information. In this paper, it is proposed a computational model of facial expression recognition, which is fast, reasonably simple, and accurate. The proposed approaches have advantages over the other face recognition schemes in its speed and simplicity, learning capacity and relative insensitivity to small or gradual changes in the face image.

The rest of this paper is organized as follows: Section 2 briefly describes the related work done for developing automatic facial expression recognition system. Section 3 gives the overview of the proposed framework and the functionality of each of the components used. Section 4 describes the experimental design and the dataset used. Section 5 summarizes the experimental results. The final section concludes that neural networks with PCA is effective in emotion recognition using facial expressions and gives the scope for the future work.

Related Work

In the recent past, the research of developing automatic facial expression recognition systems has attracted a lot of attention from many different fields as in [2], [3]. Conventional methods extract features of facial organs, such as eyes and a mouth, in gray or color images of front faces and recognize the expressions from changes in their shapes or their geometrical relationships by different facial expressions as in [4], [5], [6]. However, estimation of their precise positions and shape attributes in real images is difficult, because of the wide variety of the face features, skin color/brightness, illumination conditions and geometrical variations such as head orientations. As a result, many of the systems need human assistance such as attaching marks on the subject's face or specifying windows covering each organ in the image. Neural networks seem promising for recognizing facial expressions as in [7], [3], but the methods using Neural Networks assumes locations of facial organs are to be provided as its input. It is demonstrated to estimate the movement of muscles from optical flow to recognize facial expression as in [9], [10] with its success depends on the reliability of optical flow estimation from image sequences. Its accurate estimation seems difficult because of the complexity of facial images. Moreover, the method should compensate the flow vectors for the head movements for which estimation is not easy. These systems work under many restricted conditions.

The success of facial expressions recognition system depends heavily on how well the movement of the key features points, like eyeballs and mouth corner, are tracked on the human face. To eliminate such restrictions and to take the advantage of similarity measure between face and facial expression, in [11] the authors have proposed a method to recognize facial expressions from the whole face, rather than from changes in the shape of the facial organs such as eyes and a mouth, or their geometrical relationships. In other words, the expressions can be recognized without extracting the individual facial features. The idea is similar as the face identification method proposed in [12] and method used in [13], but the characteristics of the problem domain are quite different. In out proposed method it was designed to recognize the expression of an unknown subject from a single front view of his/her face.

Proposed Method

The proposed method of facial expression recognition system consists of four components: image preprocessing, component analysis or feature selection, classification and expression recognition. Image preprocessing consists of scaling and image rendering to prepare the face for expression recognition. The image of the face will be taken as input and the preprocessing techniques were applied, which will convert the image into the desired resolution, size and color. For feature extraction the PCA algorithm is used as in [14]. The PCA algorithm will generate the Eigenfaces for each of the image and through these Eigenfaces; the system will generate the Eigenvectors. For the classification, neural network with back propagation training algorithm is used in [15]. The process of expression recognition involves processing images by extracting the facial features, and then using an algorithm to identify the expressions made based on the movements of the feature made. The trained database consists of the extracted features of the face using the PCA. These extracted features have some known meaning for different expressions as shown in Table 1 [16]. These extracted features were compared and with the help of neural network and the expressions are recognized.

A. Psychological Basis for Recognizing Facial Expression

Table 1 summarizes the result of Ekman and Friesen as in [16], on the universal cues for recognizing the six principal emotions. These cues describe the peak of each expression and thus they provide a human interpretation of the static appearance of the facial feature. For Example: A description such as “Brows are raised” means that the human interpretation of the location of the brows relative to the other facial features indicates that they are not in neutral state but higher than usual. The viewer uses many cues to deduce such information from the image. Unfortunately the performance of humans in arriving such descriptions is far better than what can be currently achieved by computers if only static images are considered. These descriptions seem rather instinctive to human but are quite difficult to translate into computational procedures.

Table 1: The cues for facial expression as suggested by Ekman and Friesen

Emotion	Observed Facial Cues
Surprise	Brows Raised(curved and high) Skin Below brows stretched Horizontal across Forehead Eyelids opened and more of the white of the eye is visible Jaw drops open without tension or stretching of the mouth
Fear	Brows raised and drawn together Forehead wrinkles drawn to the center Upper eyelid is raised and lower eyelid is drawn up Mouth is open Lips are slightly tensed or stretched and drawn back
Disgust	Upper lip is raised Lower lip is raised and pushed up to upper lip or is lowered and slightly protruding Nose is wrinkled Cheeks are raised Brows are lowered
Anger	Brows lowered and drawn together Vertical lines appeared between brows Eyes have a hard stare and may have a bulging appearance Lips are either place firmly together with corner straight or down or are open, tensed in squarest step.
Happiness	Corners of lips are drawn back and up Mouth may or may not be parted with teeth exposed or not Cheeks are raised Lower eyelids show wrinkles below it, and may be raised but not tensed
Sadness	Inner corners of the eyes are drawn up Skin below the eye is triangulated, with inner corner up. Upper lid inner corner is raised Corner of the lips is drawn up or lip is trembling

The basic human expressions are Neutral, Happy, Sad, Disgust, Surprise and Anger, Fear is shown in Fig.1.

Figure 1: Basic Human Expressions

B. Image Pre-processing

It is performed pre-processing on the images used to train and test our algorithms as follows:

1. Images are scaled and cropped to a fixed size (150 x 120) keeping the eyes in all images aligned

2. The image is histogram equalized using the mean histogram of all the training images to make it invariant to lighting, skin color etc.

3. A fixed oval mask is applied to the image to extract face region. This serves to eliminate the background, hair, ears and other extraneous features in the image which provide no information about facial expression.

This approach works reasonably well in capturing expression-relevant facial information across all databases. Examples of pre-processed images from the dataset are shown in Fig.2 below.

Figure 2: Preprocessed images with mask

C. Feature Extraction

Gabor Filter Representations:

Gabor filters are often used in image processing and are based on physiological studies of the human visual cortex as in [17]. The use of Gabor filtered facial images has been shown to result in improved accuracy for facial expression recognition as in [18], [8], [3]. One approach to using these filters is to generate a bank of filters across multiple spatial frequencies and orientations. The filtered outputs are then concatenated, and down-sampling or PCA is often used to reduce dimensionality. It is used the similar approach as in [8] that provides competitive results, and use the L1 norm of each of the Gabor bank features for a given image (Fig.3). Our Gabor bank contains filters at 5 spatially varying frequencies and 8 orientations.

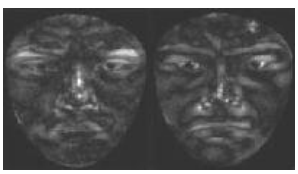

Figure 3: L1 Norm of Gabor bank features

Facial Component Extraction:

The feature vectors suffer from high dimensionality, which can cause over-fitting during classification. One approach to reducing the dimension of the feature vectors is to apply principal component analysis. Principal Component Analysis (PCA) is a statistical technique used for dimension reduction and recognition and is widely used for facial feature extraction and recognition. PCA is known as Eigen Space Projection which is based on linearly Projection the image space to a low dimension feature space that is known as Eigen space.

Many PCA-based face-recognition systems have also been developed in the last decade [19]. In this paper PCA is used along with neural network for more efficient results. When using the PCA, there is no need to calculate the facial features like lips, cheeks, etc. But rather, the whole face is considered as the principal component for facial expression recognition. In this paper first Eigenfaces are calculated for each of the different expressed image. After calculating the Eigenfaces of each image, the eigenvector will be calculated. With these Eigenvectors, a threshold value will be calculated for each of the facial expression.

Calculating Eigenfaces:

The face library entries are normalized. Eigenfaces are calculated from the training set and stored. An individual face can be represented exactly in terms of a linear combination of eigenfaces. The face can also be approximated using only the best M eigenfaces, which have the largest eigenvalues. It accounts for the most variance within the set of face images. Best M eigenfaces span an M-dimensional subspace which is called the "face space" of all possible images. For calculating the eigenface PCA algorithm was used. Let a face image I (x, y) be a two-dimensional N x N array.

An image may also be considered as a vector of dimension N2, so that a typical image of size 92 x 112 becomes a vector of dimension 10,304, or equivalently a point in 10,304-dimensional space. An ensemble of images, then, maps to a collection of points in this huge space. Images of faces, being similar in overall configuration, will not be randomly distributed in this huge image space and thus can be described by a relatively low dimensional subspace. The main idea of the principal component analysis (or Karhunen- Loeve expansion) is to find the vectors that best account for the distribution of face images within the entire image space. A set of original images is shown in Fig.4.

Figure 4: Original Images

These vectors define the subspace of face images, which we call "face space". Each vector is of length N2, describes an N x N image, and is a linear combination of the original face images. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face images, and because they are face-like in appearance, we refer to them as "eigenfaces".

Let the training set of face images be Γ1, Γ2, Γ3... ΓM then the average of the set is defined by

$$\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n \quad (1)$$

Each face differs from the average by the vector

$$\Phi i = \Gamma i - \Psi \quad (2)$$

An example training set is shown in Figure 5, with the average face Ψ. This set of very large vectors is then subject to principal component analysis, which seeks a set of M orthonormal vectors, un , which best describes the distribution of the data. The k^{th} vector, u_k , is chosen such that

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}\left(u_k^T \Phi_n\right)^2 \quad (3)$$

is a maximum, subject to

$$u_l u_k = \delta_{lk} = \begin{cases} 1, & if\ l = k \\ 0, & otherwise \end{cases}$$

The vectors u_k and scalar λ_k are the eigenvectors and eigen values, respectively of the covariance matrix

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n \Phi_n^T = A\,A^T \quad (4)$$

where the matrix A = [Φ1, Φ2,........ΦM] The covariance matrix C, however is N2 x N2 real symmetric matrix, and determining the N2 eigenvectors and eigen values is an intractable task for typical image sizes. It is needed a computationally feasible method to find these eigenvectors. If the number of data points in the image space is less than the dimension of the space (M < N2), there will be only M-1, rather than N2, meaningful eigenvectors. The remaining eigenvectors will have associated eigenvalues of zero. This can be solved for the N2 dimensional eigenvectors in this case by first solving the eigenvectors of an M x M matrix such as solving 16 x 16 matrix rather than a 10,304 x 10,304 matrix and then, taking appropriate linear combinations of the face images Φi.

Consider the eigenvectors vi of ATA such that

$$A\,A^T A v_i = \mu_i v_i \quad (5)$$

Pre multiplying both sides by A, we have

$$A\,A^T A v_i = \mu_i\ A v_i \quad (6)$$

from which we see that Av_i are the eigenvectors of $C = AA^T$.
Following these analysis, we construct the M x M matrix L = A^TA, where Lnm = Φm T Φn, and find the M eigenvectors, v_i , of L. These vectors determine linear combinations of the M training set face images to form the eigenfaces u_i.

$$\mu_i = \sum_{k=1}^{M} v_{lk}\,\Phi_k \quad (7)$$

where i = 1,2 ... M

With this analysis, the calculations are greatly reduced, from the order of the number of pixels in the images (N2) to the order of the number of images in the training set (M). In practice, the training set of face images will be relatively small (M << N2), and the calculations become quite manageable. The associated eigenvalues allow us to rank the eigenvectors according to their usefulness in characterizing the variation among the images. The set of Eigenfaces for the original images were shown in Fig.5.

Figure 5: Eigenfaces for Original Images

Training and Classification:
The training and classification is done using the Back propagation Neural Network [15]. During training, the network is trained to associate outputs with input patterns. When the network is trained, it identifies the input pattern and tries to output the associated output pattern. In order to train a neural network to perform some task, it must adjust the weights of each unit in such a way that the error between the desired output and the actual output is reduced. This process requires that the neural network to compute the error derivative of the weights (EW). In other words, it must calculate how the error changes as each weight is increased or decreased slightly. The back propagation algorithm is the most widely used method for determining the EW. The power of neural networks is realized when a pattern of tokens, during testing, is given as an input and it identifies the matching pattern it has already learned during training.

The Back Propagation Neural Network is designed based on the facial components extracted as above. The neurons in the layer are fully interconnected with weight. The training in Back propagation Neural Network involves three stages (Fig.6). The feed forward of the input training pattern, then calculation and back propagation of associated error and then adjustment of weights, to detect the facial expression in the image. The number of hidden layers can be made accordingly by experience, the more number of layers the more is training time.

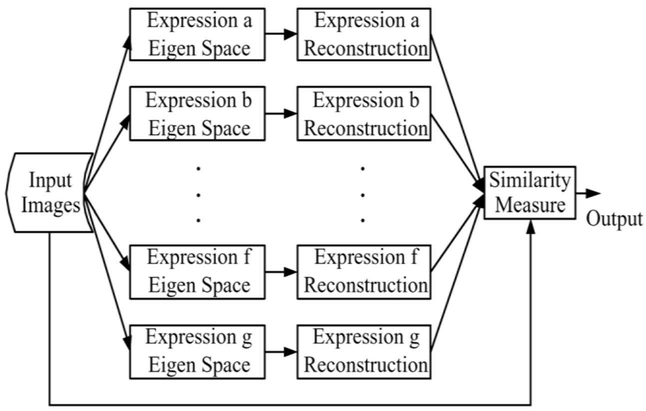

Figure 6: Classification procedures of PCA reconstruction

Recognition:
After the completion of training, the network is ready to recognize expression presented at its input. For recognizing the expression of a face two options are provided. If user wants to recognize the gesture of existing image, then it can be loaded from memory. And other option is to capture the live image. Images are captured from the web cam and preprocess the image to feed as an input.

Dataset and Experimental Design
JAFFE database of facial expression images was used (available at http://www.kasrl.org/ja_e.html). Ten expressors posed 3 or 4 examples of each of the six basic facial expressions (happiness, sadness, surprise, anger, disgust, and fear) and a neutral face for a total of 213 images of facial expressions. For simplicity of experimental design only Japanese female expressors were used.
For emotion recognition it is used neural networks and PCA-based dimensionality reduction. Algorithm for facial expression recognition classifies the given image into one of the seven basic facial expression categories (happiness, sadness, fear, surprise, anger, disgust and neutral). PCA is used for dimensionality reduction in input data. It retains those characteristics of the data set that contributes most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Because, low-order components contain the"most important" aspects of the data. The extracted feature vectors in the reduced space are used to train the supervised neural network classifier. This approach does not require the detection of any reference point or node grid.
The expression classifier was first tested using a set of images of expressions posed by ten Japanese female's expressor initials: KA, KL, KM, KR, MK, NA, NM, TM, UY and YM. Each expresser posed three or four examples of each of the six fundamental facial expressions and a neutral face. The image set was partitioned into ten segments, each corresponding to one expressor. Two facial expression images of each expression of each subject were randomly selected as training samples, while the remaining samples were used as test data. Not all expressions were equally well recognized by the system.

Our simulation experiment results show that neural networks is active in emotion recognition using facial expressions, and it is achieved a recognition rate of approximately 85% when testing six emotions. It is not so convenient to compare categorization performance, because the problem that posed expressions is not always pure examples of a single expression category. It is important to realize that expression is never pure expressions of one emotion, but always admixtures of different emotions. The expression labels on the images in JAFFE database just represent the predominant expression in that image - the expression that the subject was asked to pose.

Result Analysis
The proposed method was applied for recognition of six basic facial expressions on JAFEE database consisting 213 images posed by 10 Japanese female models. The training was carried out (with maximum 4000 epochs of training) with the learning rate λ in 0.1, and the number of hidden nodes in 2 to identify the optimal configuration. Based on the above optimal MLP_FEA configuration, it is conducted the training with error = 10^{-2} and obtained the result below in Table 2.

Table 2 : Facial Expression Classification

Feeling	Correct Classification	Classification Accuracy %
Anger	9/10	90
Fear	8/10	80
Surprise	9/10	90
Sadness	9/10	90
Joy	8/10	80
Disgust	9/10	90
Neutral	9/10	90

The average facial expression classification of the proposed method is 85.7%. It is compared the proposed method with Rapid Facial Expression Classification Using Artificial Neural Network [20], Facial Expression Classification Using Multi Artificial Neural Network [21] in the same JAFFE database (See Table 3).

Table 3: Comparison of Classification Accuracy

Method	Classification Accuracy %
Rapid Facial Expression Classification Using ANN	73.3 %
Facial Expression Classification Using Multi ANN	83.0 %
Proposed System	**85.7 %**

Conclusions and Future Work
The proposed method improves the Classification Accuracy than Rapid Facial Expression Classification Using Artificial

Neural Networks and Facial Expression Classification Using Multi Artificial Neural Network (only used ANN). Experimental results shows that, with recognition rate of approximately 85.7% when testing six emotions on benchmark image data set, the neural networks with PCA is effective in emotion recognition using facial expressions.
This research could help in future works, like capturing non-static images in real time and simultaneously analyzing these images according to affective computing techniques. By making these analyses some of the user's emotional states could be seen like joy, fear, angry, and with these probable results, assistants and computer optimizers could help users in the most different applications. These conclusions and recommendations will be tested on larger data sets using various classification algorithms in the near future.

References

[1] Mehrabian, A., Silent messages: Implicit communication of emotions and attitudes, 2nd edition, Wadsworth Pub Co; June 1980.
[2] Pantic, M. and L.J.M. Rothkrantz., "Automatic analysis of facial expressions: The state of the art", IEEE Transactions on Pattern Analysis and Machine Intelligence, 22: 1424-1445, 2000.
[3] B. Fasel, Juergen Luettin, "Automatic facial expression analysis: a survey", The Journal of Pattern Recognition, Elsevier Science, vol. 36, no. 1, pp. 259-275, January 2002.
[4] Ioannou, S., G. Caridakis, K. Karpouzis and S. Kollias, "Robust feature detection for facial expression recognition", EURASIP Journal of Image Video Proc., 29081-29081, 2007.
[5] Feng, X. B. Lv, Z. Li and J. Zhang, "A novel feature extraction method for facial expression Recognition", Proceedings of the Joint Conference on Information Sciences, (JCIS' 06), Kaohsiung, Taiwan, ROC, pp: 8-11, 2006.
[6] Terzopoulos, D. and K. Waters, "Analysis and synthesis of facial image sequences using physical and anatomical models". IEEE Transactions on Pattern Analysis and Machine Intelligence ,l5: 569-579. 1993.
[7] Kobayashi, H. and F. Hara, "The recognition of basic facial expressions by neural network", Proceedings of the IEEE International Joint Conference on Neural Networks, Nov. 18-21, IEEE Explore Press, pp: 460-466, 1991.
[8] L. Ma and K. Khorasani, "Facial Expression Recognition Using Constructive Feed forward Neural Networks", IEEE Transactions On Systems, Man, And Cybernetics—part B: Cybernetics, Vol. 34, No. 3, June 2004.
[9] Fleet, D.J., M.J. Black and A.D. Jepson,. "Design and use of linear models for image motion analysis",.Int. J. Comput. Vision, 36: 171-193, 2000.
[10] Yacoob, Y. and L.S. Davis, "Computing spatiotemporal representations of human faces", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jun. 21-23, IEEE Xplore Press, Seattle, WA., pp: 70-75. 1994.
[11] Murthy, G.R.S. and R.S. Jadon, "Recognizing facial expressions using eigenspaces". Proceeding of IEEE International Conference on Computational Intelligence and Multimedia Applications, Dec. 13-15, IEEE Xplore Press, Sivakasi, Tamilnadu, India, pp: 201-207, 2007.
[12] Turk, M. and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, 3: 71-86. 1991.
[13] Frank, C. and E. Noth, "Automatic pixel selection for optimizing facial expression recognition using eigenfaces", Pattern Recognition., 2781: 378-385, 2003.
[14] Irene Kotsia and Ioannis Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines"- IEEE Transactions On Image Processing, Vol. 16, No. 1, January 2007.
[15] Bartlett, M.S.; Littlewort, G.; Frank, M.; Lainscsek, C.; Fasel, I.; Movellan, J., "Recognizing facial expression: machine learning and application to spontaneous behavior," Computer Vision and Pattern Recognition, 2005. IEEE Computer Society Conference on , vol.2, no., pp. 568-573 June 2005.
[16] Amit Kumar, Prashant Sharma, Shishir Kumar, "Face Recognition using Neural Network and Eigen values with Distinct Block Processing"- IJCTT,Vol.1,2010.
[17] R.Q.Feitosa ,M.M.B.Vellasco, "Facial Expression Classification Using RBF AND Back-Propagation Neural Networks", Proceedings of the 4th World Multiconference on Systemics, Cybernetics and Informatics (SCI'2000), 2000.
[18] P.Ekman and W.Friesen, The Facial Action Coding System, Consulting psychologists Press, SanFrancisco, CA, 1978.
[19] K. Susheel, Shitala,Vijay, R C Tripathi, "Real Time Face Recognition Using Adaboost Improved Fast PCA Algorithm", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.2, No.3, 45-58 , July 2011.
[20] Nathan Cantelmo "Rapid Facial Expression Classification Using Artificial Neural Networks", Northwestern University. [Online] Available: http://steadystone.com/research/ ml07/ncantelmo_final.doc, 2007.
[21] L. H. Thai, T. S. Hai, "Facial Expression Classification Based on Multi Artificial Neural Network", Proc. International conference on Advance Computing and Applications, Vietnam, pp. 125-133, 2010.