International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.4 Issue No.2, pp : 83-87 01 Feb2015

A Distinct Approach for Human Face Emotion Detection Using PCA-BFO Method

Banupriya.N[1] , Pushpalatha.M[2]
[1]Department of Computer Science, Padmavani Arts and Science College for Women,Salem-11, Tamilnadu
[2]Department of Computer Science, Padmavani Arts and Science College for Women,Salem-11, Tamilnadu

Abstract:* Research on emotion has increased significantly over the past two decades with many fields contributing includingpsychology,neuroscience,endocrinology,medicine, history, sociology, and even computer science. The numerous theories that attempt to explain the origin, neurobiology, experience, and function of emotions have only fostered more intense research on this topic. Current areas of research in the concept of emotion include the development of materials that stimulate and elicit emotion. Recent discoveries suggest that emotions are intricately linked to other functions such as attention, perception, memory, decision making, and learning. In the field of Human Computer Interaction, audio and visual information are considered to be major indicators. In this paper, We explore a new method PCA-BFO for recognition of human emotional state from audiovisual features. Bacterial Foraging Optimization algorithm has been widely accepted as a global optimization algorithm of current interest for distributed optimization and control. In this paper this algorithm is combined with Principle Component Analysis. The audiovisual feature is used to classify the data into corresponding emotion using PCA-BFO method. Experimental results demonstrate the effectiveness of the proposed system and achieve the good overall recognition.

Keywords: Human Computer Interaction, Emotion, Bacterial Foraging Optimization, Principle Component Analysis.

1. INTRODUCTION

To make Human Computer Interaction (HCI) more natural and friendly, it would be beneficial to give computers the ability to recognize situations the same way a human does. In the field of HCI, audio and visual information are considered to be the two major indicators of human affective state, and thus play very important roles in emotion recognition. In this work, we explore methods by which a computer can recognize human emotion from audiovisual information. Such methods can contribute to human computer communication and to applications such as learning environment, entertainment, customer service, computer games, security/surveillance, and educational software.Certain emotions were associated with distinct facial signals, and these were common to cultures throughout the world. Can be studied but as universally distinguishable. A set of four principal emotions is: happiness, sadness, anger, surprise, neutral, which is the focus of study in this paper. Recently, audiovisual based emotion recognition methods started to draw the attention of the research community. Extracted pitch and energy as audio features, and the motion of eyebrow, eyelid, and cheek as expression features, while that of lips and jaw as the visual speech ones are used.

Face recognition has a number of strengths to recommend it over other biometric modalities in certain circumstances, and corresponding weaknesses that make it an inappropriate choice of biometric for other applications [15]. Face recognition as a biometric derives a number of advantages from being the primary biometric that humans use to recognize one another. Some of the earliest identification tokens, i.e. portraits, use this biometric as an authentication pattern. Furthermore it is well-accepted and easily understood by people, and it is easy for a human operator to arbitrate machine decisions in fact face images are often used as a human verifiable backup to automated fingerprint recognition systems. Because of its prevalence as an institutionalized and accepted guarantor of identity since the advent of photography, there are large legacy systems based on face images such as police records, passports and driving licences that are currently being automated. Video indexing is another example of legacy data for which face recognition, in conjunction with speaker identification is a valuable tool.

The remainder of this paper is organized as follows: Section 2 provides a review of related work. Section 3 present the emotion recognition system and the methodology PCA-BFO. Section 4 We described the experimental result and discussion. As given The paper ends with a discussion on the approach and some conclusion in Section 5.

2. RELATED WORK

As promising results have been obtained in emotion recognition on acted expressions, it is now necessary to move toward modeling naturalistic expressions [1], [2], [3]. In particular, an important challenge is to create systems that can continuously (i.e., over time) monitor and classify affective expressions into either discrete affective states or continuous affective dimensions. Various continuous and dimensional emotion recognition systems have been built using machine learning techniques, such as support vector machines (SVM) [4], [5]. The typical approach is to model each unit of expression (e.g., a video frame, a word) independently and to make it a standard classification problem at frame or word level. The results have been very encouraging [4], [1], [5]. Another interesting approach uses the temporal relationship between different concurrent information to provide a better classification over levels of affective dimensions. Eyben et al. [6] proposed a string-based prediction model and multimodel fusion of verbal and nonverbal behavioral events for the automatic prediction of human affect in a dimensional space. Recently, Nicolaou et al. [7] described a dimensional and continuous prediction method for emotions from naturalistic facial expressions that augments the traditional output-associative relevance vector machine (RVM) regression framework by learning nonlinear input and output dependencies inherent to the affective data.

Mel frequency Cepstral Coefficients (MFCC) is a popular and powerful analytical tool in the field of speech recognition. The purpose of MFCC is to mimic the behavior of human ears by applying cepstral analysis. In this paper, the implementation of MFCC feature extraction follow the same procedure as described in [8]. The MFCCs are computed based on speech frames. However, the lengths of the utterances are different, and thus the total number of coefficients' is different. In order to facilitate classification, the features of each utterance mapped to the feature space should have the same length. Furthermore, with a feature vector of high dimension, the computational cost is high. Usually, in speech recognition, the total number of coefficients' being used is between nine and thirteen. This is because most of the signal energy is compacted in the first few coefficients' due to the properties of the cosine transform. In this work, the first 13 coefficients' as the useful features. We then calculate the mean, median, standard deviation, max, and min of each order of coefficients' as the extracted features, which produce a total number of 65 MFCC features.

To build an emotion recognition system, the extraction of features that can truly represent the universal characteristics of the intended emotion is a challenge. For emotional speech, a good reference model is the human hearing system. Previous works have explored several different types of features. As prosody is believed to be the primary indicator of a speaker's emotional state [9], most of the works adopt prosodic features [10]. However, Mel frequency Cepstral Coefficients (MFCC) and formant frequency are also widely used in speech recognition and some other speech processing applications, and have also been studied for the purpose of emotion recognition [11]. As our goal is to simulate human perception of emotion, and identify possible features that can convey the underlying emotions in speech regardless of the language, speaker, and context.

The collected emotional data usually contain noise due to the background and "hiss" of the recording machine. Generally, the presence of noise will corrupt the signal, and make the feature extraction and classification less accurate. In this work, we perform noise reduction by thresholding the wavelet coefficients' [12]. Leading and trailing edges are then eliminated since they do not provide useful information. To perform spectral analysis for feature extraction, the preprocessed speech signal is segmented into speech frames using a Hamming window of 512 points with 50% overlap.

Formant frequencies are the properties of the vocal tract system. The formant frequency estimation is based on modeling the speech signal as if it were generated by a particular kind of source and filter [13]. To find the best matching system, we use the Linear Prediction method. In order to make the size of the formant frequency features uniform, and achieve compromise between the imitation efficiency of the vocal tract system and dimensionality of the feature space, we take the mean, median, standard deviation, max and min of the first three formant frequencies as the extracted features. In this way, we extract a total number of 15 formant frequency features from each utterance.

3. PCA-BFO METHODOLOGY

A new PCA-BFO method has been proposed by combining Principle Component Analysis with Bacterial foraging optimization.

Architecture

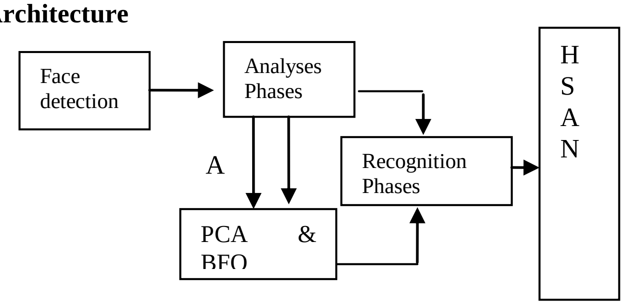


Fig 1.Architecture of PCA-BFO Methods

The fig 1 represents the architecture of the PCA-BFO method. This architecture consists of two phases. In analysis phases the audio & visual features are extracted. The extracted features are subjected to the PCA-BFO method processed the output to recognision phases However as for the result the emotions are categories as Happy, Sad, Anger, Neutral.

The Principle Component Analysis (PCA) is one of the most successful techniques that have been used in image recognition and compression. PCA is a statistical method under the broad title of factor analysis. The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically.

PCA is a common technique for finding patterns in data of high dimension. The jobs which PCA can do are prediction, redundancy removal, feature extraction, data compression, etc. Because PCA is a classical technique which can do something in the linear domain, applications having linear models are suitable, such as signal processing, image processing, system and control theory, communications, etc. The main idea of using PCA for face recognition is to express the large 1-D vector of pixels constructed from 2-D facial image into the compact principle components of the feature space. This can be called eigenspace projection.

Eigenspace is calculated by identifying the eigenvectors of the covariance matrix derived from a set of facial images(vectors).An eigenvector of a linear transformation is a vector that is either left unaffected or simply multiplied by a scale factor after the transformation (the former corresponds to a scale factor of 1). The eigenvalue of a non-zero eigenvector is the scale factor by which it has been multiplied. An eigenvalue of a linear transformation is a factor for which it has a non-zero eigenvector with that factor as its eigenvalue. The eigenspace corresponding to a given eigenvalue of a linear transformation is the vector space of all eigenvectors with that eigenvalue. The three important things that PCA deals with are as follows:

- Eigenvector – Set of features that characterize the variation between face images
- Eigen face – Displaying the eigenvector as ghostly image
- Face Space – Best M eigenfaces span an M-Dimensional subspace.

BFO is an evolutionary optimization technique motivated by the foraging behavior of the Escherichia coli abbreviated as recoil bacteria. The biological aspects of the bacterial foraging strategies and their motile behavior as well as their decision making mechanisms. As a heuristic method, BFO is designed to tackle non gradient optimization problems and to handle complex and non-differentiable objective functions. Searching the hyperspace is performed through three main operations, namely chemo taxis, reproduction and elimination dispersal activities. BFO consists of following steps: chemo taxis, swarming, reproduction, elimination and dispersal.

An e-coli bacterium can move in two different ways alternatively: tumble and run. A tumble is represented by a unit walk with random direction, a unit walk with the same direction as the previous step indicates a run. A chemo tactic process is started by one step of tumble and followed by uncertain steps of run, depending on the variation of the environment.

E-coli bacterium has a specific sensing, actuation and decision-making mechanism. As each bacterium moves, it releases attractant to signal other bacteria to swarm towards it. Meanwhile, each bacterium releases repellent to warn other bacteria to keep a safe distance from it. BFA simulates this social behavior by representing the combined cell to cell attraction and repelling effect. The proposed objective function is given below: The number of threshold level = k = {t1, t2, t3, ... Tk}. Then the objective function h is given by: Where and pi is the histogram value of the ith gray level. The above proposed objective function is a global objective function based on entropy in combination with histogram, and the user can tailor the objective function based on the application. If the number of threshold levels is 2, then the system becomes binary thresholding based on otsumethod. However, the same algorithm can be extended to multilevel thresholding if the value of k is more than 2. Later, the maxima of the selected threshold is optimized by using the BFO algorithm based on chemo taxis with random value of length within limit, random rate of elimination and dispersion of bacteria and random swim and tumbling of bacteria. The random rate of swim, tumbling and rate of elimination and dispersion give a better optimization of the maxima of the threshold level from the given threshold levels. The movement of the ith bacterium is described by ps (f + 1, g; u) = ps (f; g; u) + c (s) x v (f) where ps (f; g; u) is the sth location of the bacterium at the fth chemo tactic, gth reproductive, and uth elimination steps. C(s) is the length of one walking cycle. Here, it is defined as a small constant value. V (f) is the direction angle of the fth chemo tactic step; its default value is set at a range of $[0; 2\pi]$.

Algorithm:PCA-BFO

1. A is an image contains n pixels with gray levels from 0 – l-1.
2. Nt is the maximum no. of thresholds, nt = l-1.
3. T = {tk, k=1,2,3,. . . Nt} is the set of thresholds.
4. S = {x1,x2,. . . Xi} is the no. of particles such that xi indicates particle i, with xij ϵ {0.1}.
5. For j=1,2 …nt, such that, if xij = 1, then the corresponding tk in t has been chosen to be part of the solution proposed by xi.
6. Otherwise, if xij = 0, then the corresponding tk in t is not part of the solution proposed by xi.
7. nt is the no. of thresholds used by the multi-threshold segmentation solution represented by particle, xi, such that
8. The optimized threshold levels can be tested for their performance by evaluating the standard deviation, class variance, psnr and entropy of the thresholded images obtained by proposed algorithm and otsu algorithm.

4. RESULT AND DISCUSSION

Step 1: Data collection

In a simple example, a sample data set used. It contains 2 dimensions to show what the PCA analysis is doing at each step.

Step 2: Subtract the mean

For PCA to work properly, we have to subtract the mean from each of the data dimensions. The mean subtracted is the average across each dimension. So, all the x values have subtracted, and all the y values have subtracted from them. This produces a data set whose mean is zero.

Data =

x	y
2.5	2.4
0.5	0.7
2.2	2.9
1.9	2.2
3.1	3.0
2.3	2.7
2	1.6
1	1.1
1.5	1.6
1.1	0.9

DataAdjust =

x	y
.69	.49
-1.31	-1.21
.39	.99
.09	.29
1.29	1.09
.49	.79
.19	-.31
-.81	-.81
-.31	-.31
-.71	-1.01

Fig 2 Accuracy-chart based on PCA-BFO methodology

Step 3: Calculate the covariance matrix

Since the data is 2-D or Two dimensional, the covariance matrix will be 2 x 2.

$$Cov = \begin{pmatrix} .616555556 & .615444444 \\ .615444444 & .716555556 \end{pmatrix}$$

Since the non-diagonal elements in this covariance matrix are positive, we should expect that both the x and y variable increase together.

Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix

Since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues for this matrix. These are rather important, as they tell us useful information about our data. Here are the eigenvectors and eigenvalues:

$$Eigenvalues = \begin{pmatrix} .0490833989 \\ 1.28402771 \end{pmatrix}$$

$$\text{Eigenvectors} = \begin{pmatrix} -.735178656 & -.677873399 \\ .677873399 & -.735178656 \end{pmatrix}$$

It is important to notice that these eigenvectors are both unit eigenvectors that is their lengths are both 1. This is very important for PCA. Most maths packages, when asked for eigenvectors, will give you unit eigenvectors.

Step 5: Choosing components and forming a feature vector

Here is where the notion of data compression and reduced dimensionality comes in to it. In our example, the eigenvector with the larges eigenvalue was the one that pointed down the middle of the data. It is the most significant relationship between the data dimensions.

In general, once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives you the components in order of significance. Now, if you like, you can decide to ignore the components of lesser significance. You do lose some information, but if the eigenvalues are small, you don't lose much. If you leave out some components, the final data set will have less dimensions than the original. To be precise, if you originally have n dimensions in your data, and so you calculate n eigenvectors and eigenvalues, and then you choose only the first p eigenvectors, then the final data set has only p dimensions.

What needs to be done now is you need to form a feature vector, which is just a fancy name for a matrix of vectors. This is constructed by taking the eigenvectors that you want to keep from the list of eigenvectors, and forming a matrix with these eigenvectors in the columns.

$$\text{FeatureVector} = (\text{eig1eig2eig3......eign})$$

Given our example set of data, and the fact that we have 2 eigenvectors, we have two choices. We can either form a feature vector with both of the eigenvectors:

$$\begin{pmatrix} -.677873399 & -.735178656 \\ -.735178656 & .677873399 \end{pmatrix}$$

or, we can choose to leave out the smaller, less significant component and only have asingle column:

$$\begin{pmatrix} -.677873399 \\ -.735178656 \end{pmatrix}$$

1. **Step 6: Deriving the new data set**

This is the final step in PCA, and is also the easiest. Once we have chosen the components (eigenvectors) that we wish to keep in our data and formed a feature vector, we simply take the transpose of the vector and multiply it on the left of the original data set, transposed.

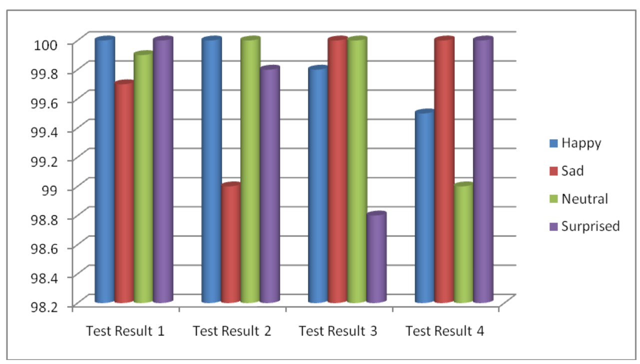
Final Data=Row Feature Vector x Row Data Adjust

where Row Feature Vector is the matrix with the eigenvectors in the columns transposed so that the eigenvectors are now in the rows, with the most significant eigen vector at the top, and Row Data Adjust is the mean-adjusted data transposed, ie. the data items are in each column, with each row holding a separate dimension. Final Data is the final data set, with data items in columns, and dimensions along rows. Table 1 gives the tabulated result of PCA-BFO algorithm.

Table1: Result based on PCA-BFO Algorithm

Proposed(PCA-BFO)	*Happy*	*Sad*	*Neutral*	*Surprised*
Test Result 1	*100*	*99.7*	*99.9*	*100*
Test Result 2	*100*	*99*	*100*	*99.8*
Test Result 3	*99.8*	*100*	*100*	*98.8*
Test Result 4	*99.5*	*100*	*99*	*100*

The graph represent PCA-BFO method for human emotions. Graph represents the result and recognition accuracy after applying the PCA-BFO algorithm the selected feature set (Happy, Sad, Surprised, Anger).



5.CONCLUSION

Emotion recognition processing helps us for extract the Emotion and recognizing emotion extraction. In this paper, a new PCA-BFO method has been proposed which combines both PCA and BFO algorithm for better optimization. This work is at infant stage and worked with sample data set. In future this method could be experimented with real time data, which would perform efficiently.

ACKNOWLEDGEMENT

I thank and unbound sense for the timely help and assistance given by my supervisor **Mrs. M.PUSHPALATHA.,** Her remarkable guidance at every stage of my research work .

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.4 Issue No.2, pp : 83-87 01 Feb2015

References

i. M. S. Bartlett, G. Littlewort, M. G. Frank, C. Lainscsek, I. R. Fasel, and J. R. Movellan, "Automatic recognition of facial actions in spontaneous expressions," J. Multimedia, vol. 1, no. 6, pp. 22–35, 2006.

ii. J. F. Cohn, L. I. Reed, Z. Ambadar, J. Xiao, T. Moriyama, and T. Moriyama, "Automatic analysis and recognition of brow actions and head motion in spontaneous facial behavior." in Proc. SMC, 2004, pp. 610–616.

iii. A. Kleinsmith and N. Bianchi-Berthouze, "Affective body expression perception and recognition: A survey," IEEE Trans. Affective Comput., vol. 99, 2012, prePrint.

iv. B. Schuller, M. Valstar, F. Eyben, G. McKeown, R. Cowie, and M. Pantic, "AVEC 2011-–The first international audio/visual emotion challenge," in Proc. Int. Conf. Affective Comput. Intell. Interaction, Oct. 2011, pp. 415–424.

v. F. Schwenker, S. Scherer, Y. Magdi, and G. Palm, "The GMM-SVM supervector approach for the recognition of the emotional status from speech," in Proc. ICANN, vol. 5768. 2009, pp. 894–903.

vi. F. Eyben, M. W¨ollmer, M. Valstar, H. Gunes, B. Schuller, and M. Pantic, "String-based audiovisual fusion of behavioural events for the assess- ment of dimensional affect," in Proc. FG, 2011, pp. 322–329.

vii. M. Nicolaou, H. Gunes, and M. Pantic, "Output-associative RVM regression for dimensional and continuous emotion prediction," in Proc. FG, 2011.

viii. G. Bailly, C. Benoit, and T. R. Sawallis, Talking Machines: Theories, Models, and Designs. Amsterdam, The Netherlands: Elsevier,

ix. C. Becchetti and L. P. Ricotti, Speech Recognition: Theory and C++ Implementation. Toronto, ON, Canada: Wiley, .

x. F.Dellaert,T.Polzin,andA.Waibel,"Recognizingemotioninspeech," in Proc. 4th Int. Conf. on Spoken Language Processing, Philadelphia, PA, Oct. 1996, vol. 3, pp. 1970–1973.

xi. T. L. Nwe, F. S. Wei, and L. C. De Silva, "Stress classification using subband features," IEICE Trans. Inform. Syst., vol. E86-D, no. 3, pp. 105–116, 2003.

xii. A. Bartlett, V. Evans, I. Frenkel, C. Hobson, and E. Sumera, Dig- ital Hearing Aids [Online]. Available: www.owlnet.rice.edu/~elec301/ Projects01/dig_hear_ aid/ 2004

xiii. Introduction to Computer Programming With MATLAB Department of Phonetics and Linguistics, Univ. College Landon [Online]. Avail- able: www.phon.ucl.ac.uk/courses/spsci/matlab/,

xiv. Tu, J., Zhang, Z., Zeng, Z. and Huang, T.S., Face Localization via Hierarchical Condensation with Fisher Boosting Feature Selection, In Proc. Computer Vision and Pattern Recognition, .

xv. Tao, H. and Huang, T.S., Explanation-based facial motion tracking using a piecewise Bezier volume deformation mode .