

Facial Expression Recognition Via Using ICA And PCA Technique

Rajneesh Singla¹ and Gurmukh Singh¹

¹Assistant professor, UIET, Panjab University, Chandigarh
eng.singla85@gmail.com

¹Assistant professor, UIET, Panjab University, Chandigarh
gurmukh86@gmail.com

Abstract: Facial expressions plays an important role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces; it may also be used in behavioural science and in clinical practice. Facial expressions, and other gestures, convey non-verbal communication cues in face-to-face interactions. This paper depicts an idea regarding detecting an unknown human face from input imagery and recognise his/her current mood. The objective of this paper is that psychological state giving information about some disorders helpful with diagnosis of discouragement, mentally ill or foolishness. In this paper we use PCA and ICA algorithm to extract facial features. The algorithms used in this paper are computationally efficient to resolve errors. In this paper we have accepted five different moods to be recognized are: Joy, Fear, Contempt, Sad, Disgust and Astonished. The main part of this paper is an emotional database which will contain images of faces, their corresponding Action Units and their labels. The contribution of this database to the problem stated above is that it can be used by systems in order to recognize emotional facial expressions given one of the database data i.e. action units' combination. Experimental results show ICA is a real effective facial expression recognition method and the recognition rate based on ICA is greater than based on PCA.

Keywords - Feature Extraction, Facial Expression Detection, Principal Component Analysis (PCA), Independent component analysis (ICA)

1. INTRODUCTON

Facial expression recognition is short for computer Automatic Facial Expression Recognition is an important part of the artificial psychology theory and computer vision research. It refers to the feature extraction of facial expressions using the computer to be understood, according to human cognition and ways of thinking, classified, and then analyzes the discriminate emotions, such as revulsion, surprise, anger, fear, happy, miserable and so on. With the rapid

development of computer vision technology, facial expression recognition has also been obtained more and more attention, and there are a lot of applications such as human-computer interaction, cognitive science and psychology and so on. After decades of continuous research and development, facial expression recognition technology has achieved very important achievements [1]. Face detection, expression feature extraction, facial expression classification and other aspects have a lot of mature and effective algorithms. And they have promoted related disciplines to be rapid development. However, we still see that the field still has a lot of problems to be solved. While humans have a strong ability to identify facial expressions, but the computer achieve quite easy. The main reason is because: First, the human face is a flexible body rather than rigid, hard to facial muscle movement to establish a consistent expression model. Different people different because of age, race, face and facial features, showing the same kind of expression, there will be a considerable difference. This increased the difficulty of expression modeling. Secondly, the expression of the face is a dynamic process, before and after a great deal of relevance in the process. If only extract the expression of a particular time for facial expression recognition, it will lose a lot of information. Finally, expression feature extraction is very difficult because of a wide variety of facial expressions and subtle changes in complex. Facial expression recognition is an attractive field which is related to image processing, pattern recognition, movement tracking, physiology, psychology and so on.

In everyday life such prototypic expressions occur relatively infrequently. Instead, emotion more often is communicated by subtle changes in one or a few discrete facial features, such as tightening of lips in anger or obliquely lowering the lips corner in sadness. Change in isolated, especially in the area of the areas of eyebrows or eyelids, is typical of paralinguistic displays; for instance raising the brows signals greeting. To capture such subtlety of human emotion and paralinguistic, automated fine grained facial expressions is needed.

II. RELATED WORK

Lee *et.al* [6] approached a method of expression-invariant face recognition that transforms input face image with an arbitrary expression into its corresponding neutral facial expression image. To achieve expression-invariance, first extract the facial feature vector from the input image using AAM. Next, transform the input facial feature vector into its corresponding neutral facial expression vector using direct or indirect facial expression transformation. Finally, perform the expression-invariant face recognition by distance-based matching

techniques nearest neighbour classifier, linear discriminant analysis (LDA) and generalized discriminant analysis (GDA). Mandeep Kaur *et. al.* [7] paper presents a new idea for detecting an unknown human face in input imagery and recognizing his/her facial expression. The objective of this project is to implement highly intelligent machines or robots that are mind implemented. Fasel *et. al.* [5] fulfills the recognition of facial action units, i.e., the subtle change of facial expressions, and emotion-specified expressions. The optimum facial feature extraction algorithm, Canny Edge Detector, is applied to localize face images, and a hierarchical clustering-based scheme reinforces the search region of extracted highly textured facial clusters. Some proposed a method based on Fisher's Linear Discriminant and produces well separated classes in a low-dimensional subspace, even under severe variation in lighting and facial expressions. The Eigen face technique, another method based on linearly projecting the image space to a low dimensional subspace, has similar computational requirements. Yet, extensive experimental results demonstrate that the proposed "Fisher face" method has error rates that are lower than those of the Eigen face technique for tests on the Harvard and Yale Face Databases. Bartlett *et. al.* [2] explores and compares techniques for automatically recognizing facial actions in sequences of images. These techniques include analysis of facial motion through estimation of optical flow; holistic spatial analysis, such as independent component analysis, local feature analysis, and linear discriminant analysis; and methods based on the outputs of local filters, such as Gabor wavelet representations and local principal components.

III. FACIAL EXPRESSION DATABASE

The Database used in my research paper for facial mood detection is Real Time Database. This Database contains (52 images of which 6 facial expressions including neutral images). The Database contains different images of an individual which represents different moods according to different situations. For the implementation the Database contains 48 coloured face images of individual. There are 4 images per subject, and these 4 images are respectively, under the following different facial expressions or configuration. In this implementation, all images are resized to a uniform dimension of 256 x 256. Following Figure shows the database images considered for face Expression recognition could be used.



Figure 3.1 Samples of the database used for training the recognition system.

IV. DIMENSION REDUCTION TECHNIQUES

In statistics, dimension reduction is the process of reducing the number of random variables under consideration, and can be divided into feature selection and feature extraction.

(i) Feature Selection

Feature selection approaches try to find a subset of the original variables (also called features or attributes). Two strategies are filter (e.g. information gain) and wrapper (e.g. search guided by the accuracy) approaches. See also combinatorial optimization problems. In some cases, data analysis such as regression or classification can be done in the reduced space more accurately than in the original space. It is mainly considered as first step in face recognition process.

(ii) Feature Extraction

Feature extraction transforms the data in the high dimensional space to a space of fewer dimensions. The data transformation may be linear, as in principal component analysis (PCA), but many nonlinear dimensionality reduction techniques also exist.

The main linear technique for dimensionality reduction, principal component analysis, performs a linear mapping of the data to a lower dimensional space in such a way that the variance of the data in the low-dimensional representation is maximized. In practice, the correlation matrix of the data is constructed and the eigenvectors on this matrix are computed. The eigenvectors that correspond to the largest eigen values (the principal components) can now be used to reconstruct a large fraction of the variance of the original data. Moreover, the first few eigenvectors can often be interpreted in terms of the large-scale

physical behavior of the system. The original space (with dimension of the number of points) has been reduced (with data loss, but hopefully retaining the most important variance) to the space spanned by a few eigenvectors.

Principal component analysis can be employed in a nonlinear way by means of the kernel trick. The resulting technique is capable of constructing nonlinear mappings that maximize the variance in the data. The resulting technique is entitled Kernel PCA. Other prominent nonlinear techniques include manifold learning techniques such as locally linear embedding (LLE), Hessian LLE, Laplacian eigenmaps, and LTSA. These techniques construct a low-dimensional data representation using a cost function that retains local properties of the data, and can be viewed as defining a graph-based kernel for Kernel PCA. More recently, techniques have been proposed that, instead of defining a fixed kernel, try to learn the kernel using semi definite programming. The most prominent example of such a technique is maximum variance unfolding (MVU). The central idea of MVU is to exactly preserve all pair wise distances between nearest neighbors (in the inner product space), while maximizing the distances between points that are not nearest neighbors.

An alternative approach to neighborhood preservation is through the minimization of a cost function that measures differences between distances in the input and output spaces. Important examples of such techniques include classical multidimensional scaling (which is identical to PCA), Isomap (which uses geodesic distances in the data space), diffusion maps (which uses diffusion distances in the data space), t-SNE (which minimizes the divergence between distributions over pairs of points), and curvilinear component analysis.

A different approach to nonlinear dimensionality reduction is through the use of auto encoders, a special kind of feed forward neural networks with a bottle-neck hidden layer. The training of deep encoders is typically performed using a greedy layer-wise pre-training (e.g., using a stack of Restricted Boltzmann machines) that is followed by a finetuning stage based on back propagation.

A. PRINCIPAL COMPONENT ANALYSIS:

Principal component analysis (PCA) is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal

components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (uncorrelated with) the preceding components. Principal components are guaranteed to be independent only if the data set is jointly normally distributed. PCA is sensitive to the relative scaling of the original variables. Depending on the field of application, it is also named the discrete Karhunen–Loève transform (KLT), the Hostelling transform or proper orthogonal decomposition (POD).

PCA was invented in 1901 by Karl Pearson. Now it is mostly used as a tool in exploratory data analysis and for making predictive models. PCA can be done by eigenvalue decomposition of a data covariance matrix or singular value decomposition of a data matrix, usually after mean centering the data for each attribute. The results of a PCA are usually discussed in terms of component scores (the transformed variable values corresponding to a particular case in the data) and loadings (the weight by which each standardized original variable should be multiplied to get the component score).

PCA is the simplest of the true eigenvector-based multivariate analyses. Often, its operation can be thought of as revealing the internal structure of the data in a way which best explains the variance in the data. If a multivariate dataset is visualized as a set of coordinates in a high dimensional data space (1 axis per variable), PCA can supply the user with a lower-dimensional picture, a "shadow" of this object when viewed from its (in some sense) most informative viewpoint. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced.

PCA is closely related to factor analysis; indeed, some statistical packages (such as Stata) deliberately conflate the two techniques. True factor analysis makes different assumptions about the underlying structure and solves eigenvectors of a slightly different matrix. Fisherface Algorithm: Fisher's linear discriminant are methods used in statistics, pattern recognition and machine learning to find a linear combination of features which characterize or separate two or more classes of objects or events. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification. In the other two methods however, the dependent variable is a numerical quantity, while for LDA it is a categorical variable (i.e. the class

label). Logistic regression and probit regression are more similar to LDA, as they also explain a categorical variable. These other methods are preferable in applications where it is not reasonable to assume that the independent variables are normally distributed, which is a fundamental assumption of the LDA method.

B. INDEPENDENT COMPONENT ANALYSIS (ICA)

Independent component analysis (ICA) is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals.

ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed nongaussian and mutually independent, and they are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA.

ICA is superficially related to principal component analysis and factor analysis. ICA is a much more powerful technique, however, capable of finding the underlying factors or sources when these classic methods fail completely.

The data analyzed by ICA could originate from many different kinds of application fields, including digital images, document databases, economic indicators and psychometric measurements. In many cases, the measurements are given as a set of parallel signals or time series; the term blind source separation is used to characterize this problem. Typical examples are mixtures of simultaneous speech signals that have been picked up by several microphones, brain waves recorded by multiple sensors, interfering radio signals arriving at a mobile phone, or parallel time series obtained from some industrial process

V. EXPERIMENT

We have experimented on The Real Time Database. Database contains (52 images of which 6 facial expressions including neutral images). The Database contains different images of an individual which represents different moods according to different situations. For the implementation the Database contains 52 coloured face images of individual. There are 4 images per subject, and these 4 images are respectively, under the

following different facial expressions or configuration. In this implementation, all images are resized to a uniform dimension of 256 x 256. Vigorous experimentation is done by selecting proper number of epochs, number of runs, step size on randomize data set to generalize the problem. Input image forms the first state for the face recognition module. To this module a face image is passed as an input for the system. The input image samples are considered of non-uniform illumination effects, variable facial expressions, and face image with glasses. In second phase of operation the face image passed is transformed to operational compatible format, where the face Image is resized to uniform dimension; the data type of the image sample is transformed to double precision and passed for Feature extraction. In Feature extraction unit runs both ICA and PCA algorithms for the computations of face for extraction. These features are passed to classifier which calculates the minimum Euclidean distance from the neutral image and the image having minimum distance is selected for output. For the implementation of the proposed recognition architecture the database samples are trained for the knowledge creation for classification. During training phase when a new facial image is added to the system the features are calculated and aligned for the dataset formation. Comparing the weights of the test face with the known weights of the database is found by calculating the norm of the differences between the test and known set of weights, such that a minimum difference between any pair would symbolize the closest match.

A. METHODOLOGY

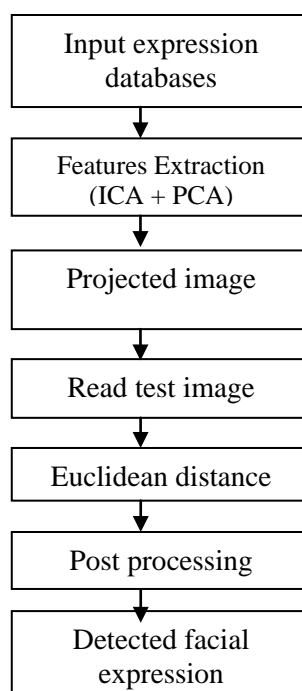


Figure 5 (a) Proposed system for facial expression based emotion detection

VI. RESULTS AND DISCUSSION

The optimally design PCA and ICA tested on the training set. The results obtained are brilliant. The recognition rate for all five principal moods namely Sad, Contempt, Fear, Joy, Revulsion and Astonished along with Neutral is obtained which is more than previous existing techniques. Finally the network is tested on real time dataset with excellent recognition rate.

Table 1. Recognition Rates for Various Facial Expressions on Test Images using PCA and ICA

Facial expression	Recognition Rate using PCA and ICA
Happy	92
Revulsion	88
Surprise	92
Angry	97
Sad	94

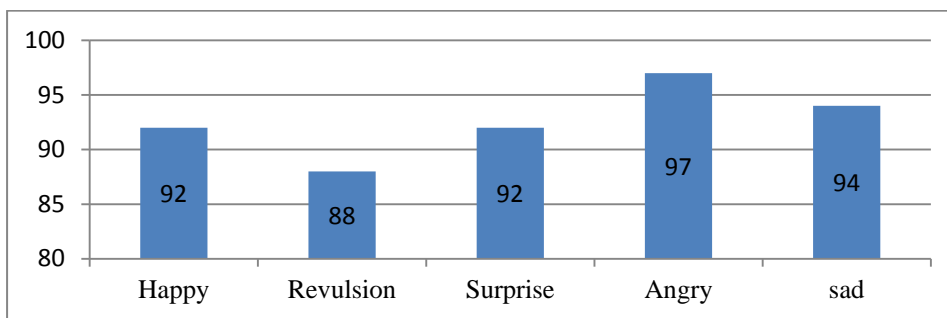


Figure 6(a) Recognition Rates of Various Facial Expressions represented on bar Chart

VII. CONCLUSIONS

The aim of this research paper was to explore the area of facial expression. A wide variety of image processing techniques was developed to meet the facial expression recognition system requirements. However, there are still many challenges and problems to solve in such

systems, especially in the area of their performance and applicability improvement. In this research paper we proposed PCA and ICA methods for dimension reduction of different types of facial expressions. The proposed algorithm is successfully implemented on Real time database. Experiments results show that algorithm can effectively recognize different facial expressions by indentifying different features.

REFERENCES

1. Bartlett, M. A., Ekman, P., Hager, J. C., and Sejnowski T., (1999) “Measuring facial expressions by computer image analysis”, *Psychophysiology*, Vol. 36, No. 2, pp.253–263.
2. Bartlett, M. S., Donato, G., Ekman, P., Hager, J. C., Sejnowski, T.J., (1999) "Classifying Facial Actions", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, No.10, pp. 974-989.
3. Cohn, J., Kanade, T., Lien, J., (2000) “Detection, tracking and classification of action units in facial expression”, *Journal of Robotics and Autonomous Systems*, Vol. 31, pp. 131–146.
4. Shang-Hung Lin, (2000) "An Introduction to Face Recognition Technology", *IC Media Corporation*, Vol. 3, No. 1, pp. 65-77.
5. Fasel, B. and Luetin, J. (2003)” Automatic facial expression analysis:” A survey. *Pattern Recognition*, Vol. 36, pp.259–275.
6. Lee, H. S., Kim, D., (2008), “Expression-invariant face recognition by facial expression transformations”, *Journal of Pattern Recognition*, Volume39, Issue 13, pp. 1797-1805.
7. Mandeep Kaur, Rajeev Vashisht., (2010) “Recognition of Facial Expression with Principal component Analysis and Singular Value Decomposition”,*International Journal of Computer Applications*, pp. 735–740.
8. Pinar Dursun, Murat Emul, Faruk Gençoz., (2010) "A Review of the Literature on Emotional Facial Expression and Its Nature", *Middle East Technical University (METU)*, pp207-215.
9. Rajneesh Singla, Simpel Jindal., (2011). “Comparitive study of dimension reduction techniques for mood detection”, *Research Cell: An International Journal of Engineering Sciences* Vol.4, pp.45-61

10. RajneeshSingla, Gurmukh Singh., (2015). “Audio Visual mood detection system via using various dimension reduction techniques. Vol. 2 Issue 12, pp. 83-92.

11. Simpel Jindal and Rajneesh Singla., (2011). “A new approach for mood detection via using principal component analysis and fisherface algorithm” *Journal of Global Research in Computer Science*, Vol. 2, No. 7, pp. 88-92