IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-12 E-ISSN: 2347-2693

Expression Invariant Face Recognition System based on Topographic Independent Component Analysis and Inner Product Classifier

Aruna Bhat

Department of Electrical Engineering, IIT Delhi, New Delhi, India

**Corresponding Author: abigit06@yahoo.com*

Available online at: www.ijcseonline.org

Received: 07/Nov/2017, Revised: 22/Nov/2017, Accepted: 14/Dec/2017, Published: 31/Dec/2017

Abstract- A technique for expression invariant face recognition using topographic modelling approach for feature extraction and Inner Product Classifier for performing classification of the faces is proposed. The topographic analysis which treats the image as a 3D surface and labels each pixel by its terrain features is used as the base for feature extraction. Based on this concept, the Topographic Independent Component Analysis (TICA) has been used to obtain the independent components such that the dependence of two components is approximated by their proximity in the topographic representation. The components that are not close to each other in the topography are independent. TICA is an extension of Independent Component Analysis for which a model needs to be developed that represents the correlation of energies for components that are close in the topographic grid. This methodology was used to extract such features from the face that are independent in terms of topography and thus invariant to changes in expression to a large extent. The feature vectors thus generated were input to the Inner Product Classifier (IPC) which considers the errors between the training and the test image features bases on triangular or t-norms. Triangular norms highlight the errors and determine a margin between them. Inner product between the aggregated training features vector and t-norm of the error vectors should be the least for the test feature vectors so as to match with the training feature vectors. The training feature vectors with the least inner product or margin give the identity of the test feature vector. Application of an effective feature extraction technique based on topographically independent components, and its combination to a classifier that works on the principle of minimization of error between the features by emphasising a margin between them, yields an efficient design for an expression invariant face recognition system.

Keywords: Topographic Independent Component Analysis, Terrain Features, Correlation of Energies, Frank t-norm, Inner Product Classifier

I. Introduction

The area of face recognition has achieved only limited applicability in real world applications so far because it suffers from the variations in the face of a person owing to various factors like pose, illumination, expressions, emotions, ageing and facial distraction like spectacles, hair growth and makeup. Therefore the efficacy of such a system can be increased only by making it invariant to these changes. The existing procedures need to be improved further and new methods have to be developed to ensure the invariance of the system against pose, lighting conditions, expressions, emotions etc. Development of novice procedures which will aid in boosting the performance of face recognition with marked robustness is important so that the system can become practically usable for authentication purposes. It will also help in forming a foundation of the layout for unobtrusive access control systems for protecting the high security zones. Humans are known to have a remarkable capability of recognizing people accurately through their face. Therefore utilization of face as a biometric trait has lately become a very popular area of research. Face is one of the most unobtrusive biometric authentication modality that can be effectively used for recognition as well as surveillance. However it has been a major challenge in all these years to design a face recognition methodology nearly as efficient as that of human ability to precisely recognize individuals even in presence of the variations in the face due to aforementioned factors. These variations impact the appearance of face to a great extent which leads to many changes in the face image making it difficult for a face recognition system to correctly authenticate a person.

Changing expressions are known to create a lot of variations in the face which makes it very hard to recognize the faces accurately. Various models for face recognition have been proposed so far each with a different approach to make a system robust and efficient. The technique of constructing Eigen Faces, proposed by M.A.Turk & A.P.Pentland [1] [2] has been a founding methodology for robust face recognition. Extensive work has followed ever since to analyse the facial expressions and understand their impact on facial image [3]

[4]. Ekman and Rosenberg proposed the facial action coding system (FACS) [5] which captured the variations of face features under facial expressions from a single image frame. A total of forty six Action Units (AUs) were defined in this system, which comprised of twelve AUs for upper face, eighteen AUs for lower face and rest being the combination of different AUs. It related the expressions to the muscles contraction in face. Face recognition on the basis of the contribution of anatomical changes in muscles towards changes in facial expressions was thereafter studies extensively. Ang et al. [6] analysed the activities of the muscles of face to recognize the emotions. The emotions were captured from the facial muscles through electromyogram Sensors (EMG) signals. Takami et al. [7] proposed a method of using quasi-muscles to quantify the facial expressions. Pentland , Moghaddam and Starner [8] proposed a view based modular Eigen space approach for Face Recognition. I. Essa and A.P.Pentland designed a method of classifying facial expressions by analysing multiple image frames in a video sequence [9]. Zhang et al [10] proposed an approach to find features using Active Appearance Model (AAM). Vretos et al [11] also used Appearance based approach and SVM classifier with a classification accuracy of around 90%. Kuilenburg [12] used Delauny Triangulation for feature extraction and PCA based BPNN for classification. Kotsia and Pitas [13] used facial feature geometry for facial expression recognition. Jun et al [14] proposed a methodology for face expression recognition using wavelet energy distribution features and Neural Networks. More work has been done in this field using different approaches of machine learning [15] [16], however the efficiency and performance of such systems still needs to be ameliorated to a much larger extent so as to use an expression invariant face recognition system actively in biometric based access control.

This paper presents a method based on the topographic analysis which treats the image as a 3D surface and labels each pixel by its terrain features is used as the base for feature extraction. Based on this concept, the Topographic Independent Component Analysis (TICA) [19] has been used to obtain the independent components such that the dependence of two components is approximated by their proximity in the topographic representation. The components that are not close to each other in the topography are independent. TICA is an extension of Independent Component Analysis for which a model needs to be developed that represents the correlation of energies for components that are close in the topographic grid. This methodology was used to extract such features from the face that are independent in terms of topography and thus invariant to changes in expression to a large extent. We have used the Inner Product Classifier (IPC) [21] which is based on the errors between the training features and the test image features using triangular or t-norms. The feature vectors thus generated from TICA were input to the Inner Product Classifier. An aggregate of the two training features is taken and fused with their errors using Frank t-norm. A minimum of two errors can thus be obtained. The inner product between the aggregated training features vector and t-norm of the error vectors should be lowest for the test feature vectors so as to match with the training feature vectors. Thus the triangular norms aid in highlighting the errors and demarcate a margin between them. The difference between the highest and the lowest inner products provides the largest margin. The training feature vectors with the lowest inner product signifying the margin decide the best match for the test feature vector. Using IPC classifier with topographically independent features based on TICA was found to yield higher success rate for expression invariant face recognition with lesser computational time and cost.

The rest of this paper is organized as follows: Section 2 provides a detailed description of the proposed methodology for robust expression invariant face recognition based on Topographic Independent Component Analysis and Inner Product Classifier. Experimental results have been elaborated in Section 3 followed by Section 4 which provides the conclusion. Section 5 elaborates on the possible future scope of the methodology in designing practically usable sturdy face recognition systems.

II. METHODOLOGY

A technique for expression invariant face recognition using topographic modelling approach for feature extraction and Inner Product Classifier for performing classification of the faces is proposed.

Feature Extraction using Topographic Independent Component Analysis

The topographic analysis [17] [18], which treats the image as a 3D surface and labels each pixel by its terrain features is used as the base for feature extraction. Extending this concept, the Topographic Independent Component Analysis (TICA) [19] has been used to obtain the independent components such that the dependence of two components is approximated by their proximity in the topographic representation. The components that are not close to each other in the topography are independent.

Correlation of Energies is used as a measure of dependency. We say that s_i and s_j are close in the topography if the below condition holds true for the covariance between the two components:

$$\mathrm{Cov}\,(s_i^2, s_j^2) = E\{ s_i^2\, s_j^2\} - E\{ s_i^2\}\, E\{ s_j^2\} \neq 0 \quad (1)$$

TICA is an extension of Independent Component Analysis (ICA) proposed by Jutten and Herault [20]. ICA is a statistical model where the observed data is expressed as a linear transformation of the latent variables that are non-gaussian and mutually independent. It gives an arbitrarily chosen subset of independent components, corresponding to a local minimum of the objective function.

A classical version of the model can be expressed as:

$$x = As \tag{2}$$

where,
$x = (x_1, x_2, \ldots, x_n)^T$ is the vector of observed random variables,
$s = (s_1, s_2, \ldots, s_n)^T$ is the vector of the independent components,
A is an unknown constant matrix, called the mixing matrix.

Using observations of x, the matrix A and independent components need to be estimated.

For TICA, a model [19] needs to be developed that represents the correlation of energies for components that are close in the topographic grid. For the model described above, we need to define the joint density of s so that it expresses the topography.

First we define a neighbourhood function h(i, j), that expresses the proximity between the i^{th} and j^{th} components. This neighbourhood function is defined as a monotonically decreasing function of some distance measure.

$$h(i, j) = \begin{cases} 0, & if\ |i-j| \geq y \\ 0, & otherwise \end{cases} \tag{3}$$

where, y is the neighbourhood width.

As per the principle of topography, the variances corresponding to proximal components should be positively correlated, and the variances of the components which are not close should be independent, at least approximately.

Therefore using the topographic relation h(i,j), the variances can be described as:

$$\sigma_i = \phi\left(\sum_{k=1}^{n} h(i,k)u_k\right) \tag{4}$$

Here ϕ represents scalar non linearity and u_k are the higher order independent components that are used to generate the variances.

If z_i is a random variable having same distribution as s_i, then s_i can be expressed as:

$$s_i = \phi\left(\sum_{k=1}^{n} h(i,k)u_k\right) * z_i \tag{5}$$

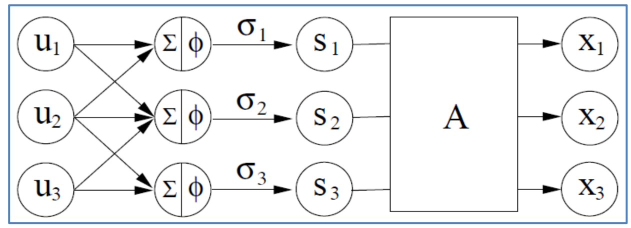

Figure 1. Topographic Independent Component Analysis

The above methodology was used to extract such features from the face that are independent in terms of topography and thus invariant to changes in expression to a large extent. TICA was applied over the testing and the training data sets for feature extraction. The feature vectors thus generated were input to the Inner Product Classifier (IPC) as described in detail in the next sub-section.

2.2. Classification using Inner Product Classifier

The Inner Product Classifier (IPC) [21] is used to perform classification. The features determined from the training database are used to train the IP classifier. It basically works around the errors computed between the features from the training image set and the test image set. The feature error is based on the triangular norms or the t-norms such that an aggregate of the training set of features and their subsequent fusion of errors produces the classifying margin.
Triangular norm is a function which may be defined as follows:

$$f: [0, 1] \times [0, 1] \rightarrow [0, 1]$$

Given two fuzzy variables x and y, the t-norm fundamentally generalizes their logical conjunction over the interval [0, 1]. Here these fuzzy variables are nothing but the two comparable feature vectors.

The t-norm must conform to the properties of identity, commutativity, associativity and monotonicity as described below:

1. Condition of identity, where the number 1 acts as identity element,

   $$f(x, 1) = x \tag{6}$$

2. Condition of commutativity such that,

   $$f(x, y) = f(y, x) \tag{7}$$

3. Condition of associativity such that,

$$f(x, f(y, w)) = f(f(x, y), w) \quad (8)$$

4. Condition of monotonicity:

$$f(x, y) \le f(w, z) \text{ if } x \le w \text{ and } y \le z \quad (9)$$

The classifier design is founded upon the application of the above described t-norms in highlighting the errors for determining a margin between categories. One can calculate the minimum of two errors by the fusion of triangular norms.

The inner product computed between aggregated training feature vectors and the triangular norm of the error vectors has to be the least such that the test feature vectors would find its closest age group class.

Difference between the largest and the smallest inner products would produce the highest margin. The training feature vectors having smallest inner product denoting margin direct us towards the correct class of the test feature vector.

We may note that the errors being positive, the margin is also directed towards the positive side of the demarcating hyperplane.

Let T_M be the sample size of the training images per subject, where u = 1, 2,….., T_M and T_N be the total number of features for each one of such sample, where v = 1, 2,….., T_N.

We need to compute the error between the training and test feature vector samples which may be represented as:

F_r : the feature vectors for the training samples

F_s : the feature vectors for the test samples.

Error E_{uv} between the u^{th} training sample of each user and the test sample may be calculated as:

$$E_u(v) = |F_r(u, v) - F_s(v)| \quad (10)$$

This is followed by fusion of the errors of the u^{th} and the k^{th} training samples by Frank t-norm denoted by $t_{Frank,}$

$$t_{Frank} = \log_p[\, 1 + \frac{(p^a-1)(p^b-1)}{(p^b-1)}\,] \quad (11)$$

such that p = 2.

The classifier developed on this line can also use other variants of triangular norms depending on the application. The use of Frank t-norm for the classification based on texture for the problem at hand was observed to produce favourable results by significantly highlighting the errors for the eventual development of margin.

$$E_{uk}(v) = t_{Frank}(E_u(v), E_k(v)) \text{ for } u \neq k \quad (12)$$

All the combinations of training sample errors have been taken into account for correctly estimating the minimum value of $E_{uk}(v)$.

For u^{th} and the k^{th} training samples, the average feature value is computed as:

$$X_{uk}(v) = \frac{F_r(u,v) + F_r(k,v)}{2} \quad (13)$$

Post fusing, these error vectors represent the support vectors. The corresponding weight is depicted through the average feature vectors.

Collectively they contribute towards development of the demarcating hyper plane by means of their inner product.

Template Matching:
To find a match for a test subject in the database calls for determining the proximity of the test sample to various training samples with respect to the classifying hyperplane. Therefore the inner product of $E_{uk}(v)$ and $X_{uk}(v)$ needs to be computed.

The lowest inner product retrieved aids in finding the best match for the test samples based on the highest proximity of the same to the separating hyperplane.

Inner product for $E_{uk}(v)$ and $X_{uk}(v)$ may be calculated as:

$$P_{uk}(a) = \sum_{v=1}^{N} E_{uk}(v) X_{uk}(v) = \langle E_{uk}, X_{uk} \rangle \text{ for } u \neq k \quad (14)$$

The minimum of $P_{uk}(a)$ is the measure of identity associated with the subject 'a'. It should be minimal so as to find the best match associated with the test sample.

III. EXPERIMENTAL RESULTS

Application of an effective feature extraction technique based on topographically independent components, and its combination to a classifier that works on the principle of minimization of error between the features by emphasising a margin between them, yields an efficient design for an expression invariant face recognition system.

The design was tested over open source data sets like JAFFE database [22] and Cohn Kanade database [23] to evaluate the performance of the proposed methodology. JAFFE database contains various face images of different subjects with varying expressions and emotions. The database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects.

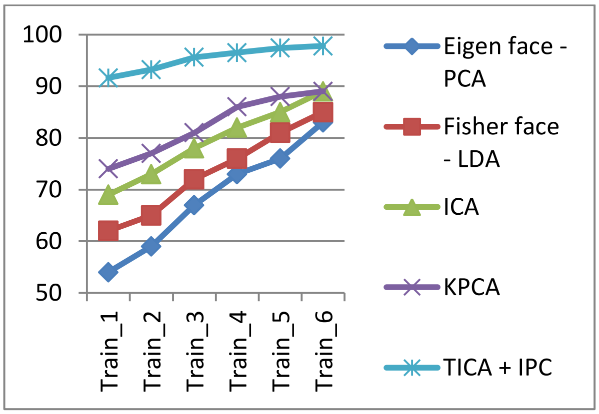

Figure 2. Recognition Accuracy w.r.t number of samples per subject in training data set

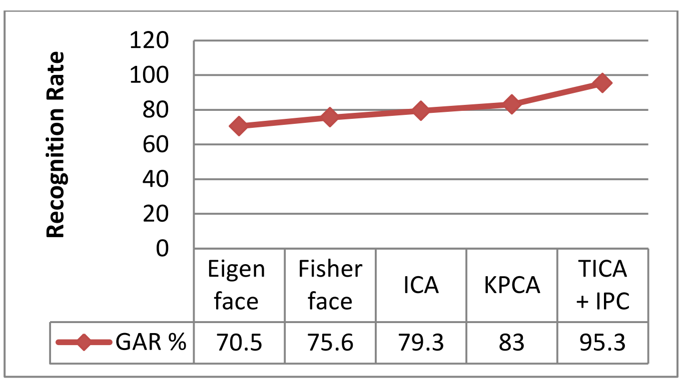

Figure 3. Comparison of Recognition Rate in terms of Genuine Acceptance Rate (GAR %)

The proposed technique has been compared with various well known standard face recognition approaches like Eigen face method which is based on Principal Component Analysis, Linear Discriminant based Fisher face technique, Independent Component Analysis method and Kernel PCA methodology with varying size of the training data set for each test subject. The performance of the proposed system was reasonably consistent and stable across the changing training data set size.

The proposed technique not only requires lesser training effort, but is also computationally much more efficient and reliable.

The results hence produced support the utility of TICA and IPC for developing a robust expression invariant face recognition system.

IV. Conclusions

Using Inner Product Classifier with topographically independent features based on TICA produced higher success rate for expression invariant face recognition with lesser computational time and cost. The performance was remarkably stable even with varying size of the training data set.

The uniqueness of the said methodology lies in using such a technique for generating the feature vector which captures the features that are least susceptible to be altered by changing expressions and emotions. To add towards the efficacy of the system, the robust features so detected are fed to a unique classifier which is based on the principle of minimization of error between the features by emphasising a margin between them. The system thus developed was observed to yield reasonably high recognition accuracy for expression invariant face recognition.

V. Future Scope

The design can be further improvised by considering correlated topographic analysis (CTA) technique and evaluate the system for any significant increase in the overall performance. CTA estimates non-Gaussian components and their ordering (topography) based on the assumption that components nearby on the topographic arrangement have both linear and energy correlations, while far-away components are statistically independent.

The area of face recognition also faces hurdles due to many other unavoidable changes in the face like illumination, pose, facial distractions, age etc. The proposed methodology may be evaluated for its applicability in alleviating the impact of such variations so that a single universally robust system can be developed which may then be used in practical dynamic environment as well.

References

[1]. M.Turk & A.Pentland, (1991) "*Eigenfaces for Recognition*", Journal of Cognitive Neuroscience, vol.3, no.1, pp. 71-86, 1991a.

[2]. M. Turk & A. Pentland, (1991) "*Face Recognition Using Eigenfaces*", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586-591.

[3]. Rothkrantz M., "*Automatic Analysis of Facial Expressions: The State of the Art,*" IEEE Transaction Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424-1445, 2000.

[4]. Fasel B. and Luettin J., "*Automatic Facial Expression Analysis: a Survey,*" IEEE Pattern Recognition, vol. 36, no. 1, pp. 259-275, 2003.

[5]. E. Mary Shyla , Dr.M.Punithavalli, "*Hybrid Facial Color Component Feature Identification Using Bayesian Classifier*", International Journal of Scientific Research in Computer Science and Engineering, Vol.1, Issue.3, pp.14-21, 2013.

[6]. Ang L., Belen E., Bernardo R., Boongaling E., Briones G., and Corone J., "*Facial Expression Recognition through Pattern Analysis of Facial Muscle Movements Utilizing Electromyogram Sensors,*" in Proceedings of IEEE TENCON, vol. 3, pp. 600-603, 2004.

[7]. A. Sur , S. Sarkar , K. Sarkar , "*An Approach towards Face Counting System using Image Processing Techniques*", International Journal of Computer Sciences and Engineering, Vol.5, Issue.2, pp.34-37, 2017.

[8]. Pentland A., Moghaddam B., and Starner T., "*View-Based and Modular Eigenspaces for Face Recognition,*" in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp.84-91, 1994.

[9]. I. Essa and A.P.*Pentland. Coding, analysis, interpretation and recognition of facial expressions*. Pattern Analysis and Machine Intelligence, IEEE Transactions, 19(7):757–763, Jul 1997.

[10]. Zhang Y. and Martinez A., "*Recognition of Expression Variant Faces Using Weighted Subspaces,"* in Proceedings of the 17th International Conference on Pattern Recognition, vol. 3, pp. 149-152, 2004.

[11]. Vretos N., Nikolaidis N., and Pitas I., "*A Model-Based Facial Expression Recognition Algorithm using Principal Component Analysis,"* in Proceedings of the 16th IEEE International Conference on image Processing, Cairo, pp.3301-3304, 2009.

[12]. Kuilenburg H., Wiering M., and Uyl M., *"A Model Based Method for Automatic Facial Expression Recognition*," Springer Verlag Proceedings of the ECML, vol. 54, pp. 194-205, 2005.

[13]. Kotsia I. and Pitas I., "*Facial Expression Recognition in Image Sequences using Geometric Deformation Features and Support Vector Machines*" IEEE Transaction on Image Processing, vol. 16, no. 1, pp. 172-187, 2007.

[14]. Jun C., liang W., Guang X., and Jiang X., "*Facial Expression Recognition based on Wavelet Energy Distribution features and Neural Network Ensemble*," in Proceedings of Global Congress on Intelligent Systems, Xiamen, vol. 2, pp. 122-126, 2009.

[15]. Bettadapura V., *"Face Expression Recognition and Analysis: The State of the Art,*" in Proceedings of IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 22, pp.1424-1445, 2002.

[16]. Bartlett M., Littlewort G., Lainscsek C., Fasel I., and Movellan J., "*Machine Learning Methods for Fully Automatic Recognition of Facial Expressions and Facial Actions,*" in Proceedings of IEEE International Conference on Systems, Man and Cybernetics, vol. 1, pp. 592-597, 2004.

[17]. L. Wang, T. Pavlidis, *Direct gray-scale extraction of features for character recognition*, IEEE Trans. PAMI 15 (1993) 1053–1067.

[18]. O. Trier, T. Taxt, A. Jain, *Ata capture from maps based on gray scale topographic analysis*, in: The Third International Conference on Document Analysis and Recognition, Montreal, Canada, 1995.

[19]. Aapo Hyvarinen, Patrik O. Hoyer, and Mika Inki, *"Topographic Independent Component Analysis*", Neural Computation 13(7):1527-1558 (July, 2001).

[20]. Jutten, C. and Herault, J. (1991). *Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture*. Signal Processing, 24:1–10

[21]. Mamta, Madasu Hanmandlu, "*Robust ear based authentication using Local Principal Independent Components*", Expert Systems with Applications 40 (2013) 6478–6490 (Elsevier).

[22]. Michael J. Lyons, Shigeru Akamatsu, Miyuki Kamachi & Jiro Gyoba *JAFFE database*, *Coding Facial Expressions with Gabor Wavelets* Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16 1998, Nara Japan, IEEE Computer Society, pp. 200-205.

[23]. T. Kanade, J. Cohn, Y. Tian, "*Comprehensive database for facial expression analysis*", in: IEEE 4th International Conference on FGR, France, 2000.

Author's Profile

The author is a research scholar at the Department of Electrical Engineering, Indian Institute of Technology, Hauz Khas, Delhi.