Independent Component Analysis (ICA) of fused Wavelet Coefficients of thermal and visual images for human face recognition

Article *in* Proceedings of SPIE - The International Society for Optical Engineering · May 2011

DOI: 10.1117/12.884455

4 authors, including:

MRINAL KANTI BHOWMIK
Tripura University

Debotosh Bhattacharjee
Jadavpur University

Mita Nasipuri
Jadavpur University

Independent Component Analysis (ICA) of fused Wavelet Coefficients of thermal and visual images for human face recognition

Mrinal K. Bhowmik*[a], Debotosh Bhattacharjee[b], Dipak K. Basu[#b], Mita Nasipuri[b]
[a]Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799130, Tripura, India;
[b]Department of Computer Science and Engineering, Jadavpur University, Kolkata- 700032, India;
[#]AICTE Emeritus Fellow

ABSTRACT

In this paper, an image fusion technique based on weighted average of Daubechies wavelet transform (db2) coefficients from visual face image and their corresponding thermal images have been presented. Further, a comparative study has been conducted for dimensionality reduction based on Principal Component Analysis (PCA) and Independent Component Analysis (ICA). Fused images thus obtained are classified using a multi-layer perceptron (MLP). For experiments IRIS Thermal/Visual Face Database has been used. Experimental results show that the performance of ICA architecture-I is better than the other two approaches i.e. PCA and ICA-II. The average success rate for PCA, ICA-I and ICA-II are 91.13%, 94.44% and 89.72% respectively. However, approaches presented here achieves maximum success rate of 100% in some cases, especially in case of varying illumination.

Keywords: face recognition, Principal Component Analysis (PCA), Independent Component Analysis (ICA), dimensionality reduction, feature extraction, classification, neural network

1. INTRODUCTION

In the field of face recognition different techniques have been developed on the basis of feature extraction to solve the problems like expression, illumination and disguise. I. Dagher et al[2] have used the fast incremental principle non-Gaussian direction analysis algorithm, called IPCA-ICA. This algorithm computes the principle components of a sequence of image vectors incrementally without estimating the covariance matrix i.e. covariance free and at the same time transforming these principal components to the independent directions that maximize the non-Gaussianity of the source. The major difference between the IPCA_ICA algorithm and the PCA_ICA batch algorithm is the real time sequential process. IPCA_ICA needs less memory than PCA_ICA as it doesn't need to calculate the eigenfaces. They applied this algorithm into three different visual face datasets which contained the data of expression and illumination variations with changes in pose, but they have not used any thermal face images or any fused face image of visual and its corresponding thermal face image. Three databases are randomly partitioned to check the performance of the algorithm. From the ORL dataset they have selected data of up-right and frontal pose with varying lighting condition. They have also selected the data of wide range of pose form the UMIST database and from the Yale database, they picked the data of varying expression like sad, sleepy and surprise etc. Although in[2] they focused on the feature extraction techniques and they have found that IPCA_ICA gives better performance than other feature extraction methods like DCV, PCA, Batch PCA_ICA and Fisherfaces and average success rate is 93.66% of the three different databases. M. S. Bartlett et al[3] explored one such generalization by Bell and Sejnowski's ICA algorithm, which was derived from the principle of optimal information transfer through sigmoidal neurons. In their case, ICA was performed on face images in the FERET database, which contained only frontal faces images under two different architectures. The first architecture found spatially local basis images for the faces. The second architecture produced a factorial face code.

Both ICA representations were superior to representations based on PCA for recognizing faces across days and changes in expression of visual faces only. A classifier that combines the two ICA representations gave the best performance.

*mkb_cse@yahoo.co.in; phone 0 94361 29933; fax 0381 237 4802

Independent Component Analyses, Wavelets, Neural Networks, Biosystems, and Nanoengineering IX,
edited by Harold Szu, Liyi Dai, Proc. of SPIE Vol. 8058, 80581H · © 2011 SPIE
CCC code: 0277-786X/11/$18 · doi: 10.1117/12.884455

They have done the experiments on the basis of number of ICs of an image and achieved best performance results. J. Kim et al[1] proposed an effective part-based local representation method named locally salient ICA (LS-ICA) method for face recognition that is robust to local distortion and partial occlusion. The LS-ICA method only employs locally salient information from important facial parts in order to maximize the benefit of applying the idea of "recognition by parts". For the study they have used three different databases: The FERET database used for the data of different lighting and facial expression, AT & T for comparison under occlusion, and AR database for changes in expression due to local distortion and occlusion. They have reported that LS-ICA shown better result than PCA, ICA-I and II, LFA (Local Feature Analysis) and LNMF (Local Nonnegative Matrix Factorization). H. K. Ekenel et al[12], proposed a method, prior to the subspace projection operation like PCA and ICA which employs multiresolution analysis to decompose the image into its subbands. Their aim was to search for the subbands that are insensitive to the variations in expressions and illuminations. This algorithm is tested on face images that differ in expression or illumination separately, obtained from CMU PIE, FERET and Yale database. They have found that, 96% is the best performing subbands using ICA architecture 2. Y. Z. Goh et al[13] proposed a wavelet based illumination invariant algorithm based on the illumination-reflectance model. The proposed method aims to remove illumination component. This can be done by setting the coefficients in wavelet approximation subband to zero values. These algorithms are tested on Yale B database and CMU PIE database. Although they have used two databases but they have chosen only the frontal face images varying different illumination and they have shown 10.83% as a lowest error rate. All the method discussed in this section the technique used in the[2] as compare with the data they have used for recognition with different pose varying with expression and illumination have shown better performance than others i.e. the IPCA_ICA algorithm.

Data Fusion for Biometric Verification with the consideration of multimodal biometric system and their applicability to access control, authentication and data privacy applications have been discussed by R. A. Wasniowski[19]. As face recognition system is closely related to individual identity and security system, so it needs high degree of accuracy. The motivation behind using fusion technique is to overcome the drawbacks of both visual and thermal face recognition individually. It is relatively easier to extract and locate facial features in visual images and visual cameras are also less expensive[20]. But it results in poor performance with illumination variations, facial expressions, viewing directions or poses and disguises such as facial hair, glasses, or cosmetics. To solve the above mentioned problems of visual images we can use thermal image as alternate. Recently, thermal images have acquired significant importance over visual images. Unlike of using the visible spectrum, recognition of faces using different multi-spectral imaging modalities, particularly infrared (IR) imaging sensors[21, 22, 23, 24, 25, 26] have become an area of growing interest. Thermal image recognition can work well in detecting disguised faces and also can handle the situations when there is no control over illumination[10]. Visual imagery fails in identifying twins. But it is known that even identical twins have different thermal patterns. So using thermal imagery we can easily identify the twins. As fusion method combines the information of both visual and thermal imagery, it is an alternative solution to that condition when individual visual or thermal face recognition fails.

In this paper, the investigation on filter based fusion of thermal and visual face images to protect the private data using ICA feature extraction method is presented. The reasons behind use of ICA in comparison to PCA are:

(i) ICA has better architecture than PCA, which can easily understand where the data is mainly concentrated in n-dimensional space;

(ii) In case of PCA it reconstructs the data on the orthogonal basis, but on the other hand it is not necessary for ICA; and

(iii) ICA can easily reconstruct the data even in the presence of noise[5].

In this work ICA has been implemented using FastICA algorithm using fixed point algorithm as given by A. Hyvarinen[5, 6]. The ICA is used to estimate the independent characterization of human faces. Independent Component Analysis has two different architecture namely Architecture-I and Architecture-II. In Architecture-I (called as ICA-I) N principal component eigenvectors are loaded into rows of data matrix, and then run ICA on data matrix. But in case of Architecture-II (called as ICA-II), principal component coefficients of data matrix are loaded and then run ICA on. The organization of the rest of this paper is as follows: in section 2, the overview of the system is discussed, in this section general idea on wavelet filters and the feature extraction methods are discussed. In section 3 experimental results with different data and discussions on them are given. Finally, section 4 concludes this work.

2. SYSTEM OVERVIEW

The complete system implementation in this work is described with a block diagram, which is shown in Fig. 1. All the steps used in this paper are shown in the Fig. 1. In the first step, Discrete Wavelet Transform (DWT) is used for decomposition of both the thermal and visual images up to level two using wavelet filter namely Daubechies Coefficient (db2). The decomposition in the wavelet domain is dependent on the dimension of the input image. After decomposition we can separate the high frequency and low frequency information of an image. In this experiment data used for DWT decomposition have been shown in the Fig. 2. Decomposition is a process by which the image is partially divided into its subbands in order to search the subbands those are insensitive to the variations in expressions and illuminations. After the decomposition of both the thermal and visual images at the size of level-2, weighted averages are computed and then fused images are reconstructed. After those two different subspace projection techniques namely PCA and ICA are employed, which are shown in the Fig. 1. After that projected data have been passed through a neural network for classification. A multilayer perceptron has been used for classification.

Two-dimensional wavelet decomposition[12] has been performed to the rows and columns of the two-dimensional data using wavelet filters as shown in Fig. 2. At the first level an 80 x 100 pixel size image has been taken for decomposition. The result of first level decomposition generates four 41 x 51 resolution subband images as an approximate and detailed coefficient and other two diagonals as horizontal and vertical. In the second and final stage of the decomposition subband images of size 22 x 27 pixels were achieved as in Fig. 2.

2.1 Wavelet Filter based decomposition and reconstruction structure

Reason behind the use of wavelet decomposition and reconstruction is to separate the high frequency and low frequency information of an image and generate fused image of high and low frequency sub-bands separately in wavelet transform. Wavelet transforms are multi-resolution image decomposition tool that provide a variety of channels representing the image feature by different frequency subbands at multi-scale. It is a famous technique for analyzing signals. From Fig. 1,

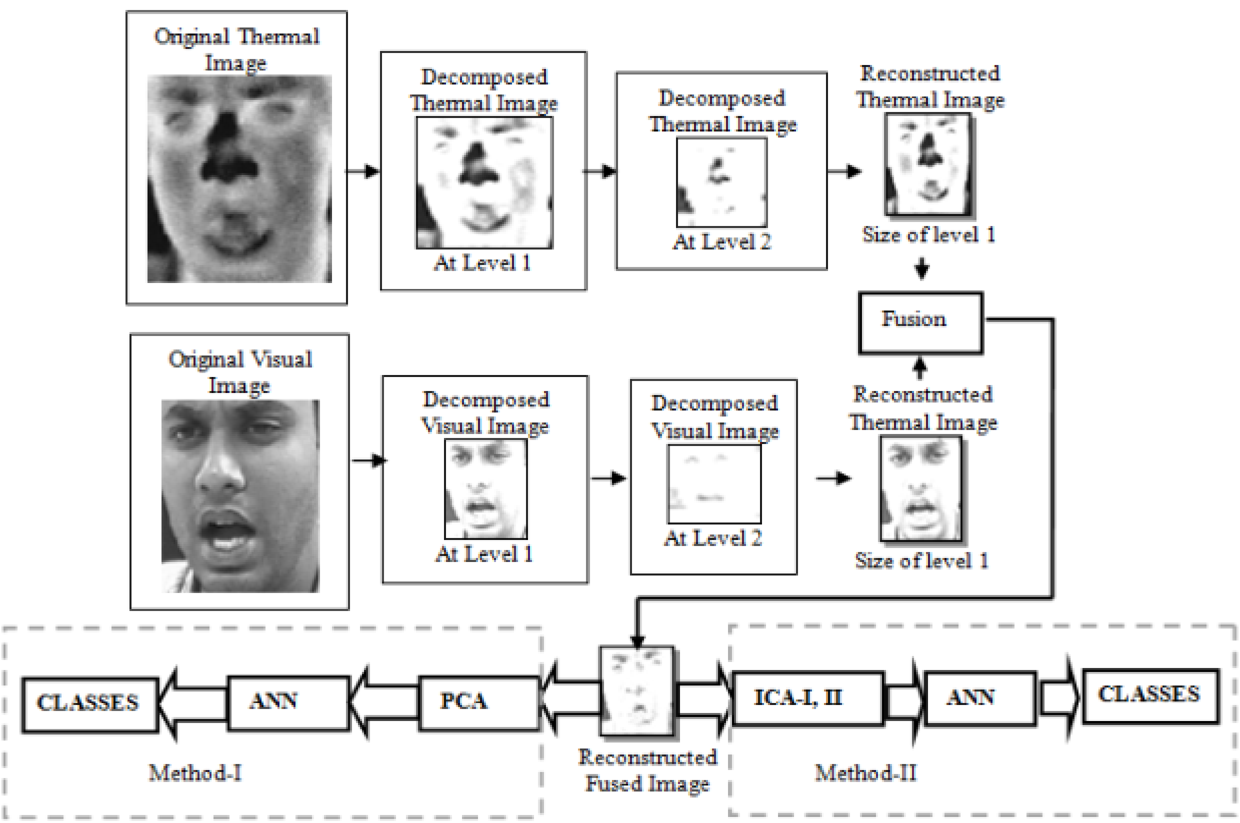

Figure 1. Block diagram of the system.

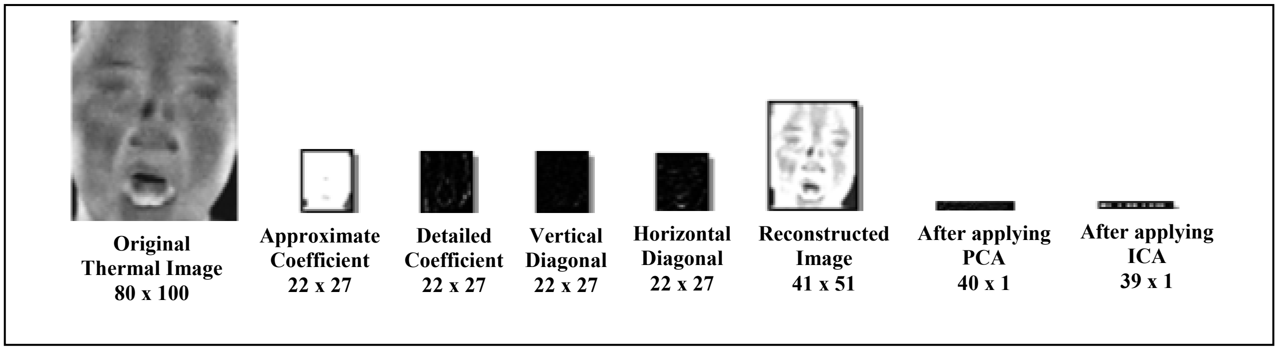

Figure 2. Sample of two-level decomposition image with their corresponding PCA and ICA images.

one can easily justify that before generating the fused image, decomposition of all thermal and visual images have. been completed using decomposition low-pass (Lo_D) and high-pass (Hi_D) filter by down-sampling (keep one row out of two and one column out of two) process of images.

2.2 Image fusion in wavelet domain

In this paper, the process of fusion algorithm is as follows: (i) the two images to be processed and resample to the one with the same size; and (ii) they are respectively decomposed into the sub-images using forward wavelet transform, which have the same resolution at the same levels and different resolution among different levels; and (iii) fusion is performed based on the high-frequency sub-images of decomposed images; and finally the result image is obtained using inverse wavelet transform.

Let $A(x, y)$ and $B(x, y)$ be the images to be fused, the decomposed low-frequency (l) sub-images of $A(x, y)$ and $B(x, y)$ be respectively $lA_J(x, y)$ and $lB_J(x, y)$ (J is the parameter of resolution). The decomposed high-frequency (h) sub-images of $A(x, y)$ and $B(x, y)$ be respectively $hA_j^k(x, y)$ and $hB_j^k(x, y)$ (j is the parameter of resolution and $j = 1, 2, \dots, J$ for every j, k=1, 2, 3). Then the fused high- frequency sub-images $F_j^k(x, y)$ are: if $hA_j^k(x, y) > hB_j^k(x, y)$ then $F_j^k(x, y) = hA_j^k(x, y)$ and if $hA_j^k(x, y) < hB_j^k(x, y)$ then $F_j^k(x, y) = hB_j^k(x, y)$; and the fused low-frequency sub images $F_j(x, y)$ are as follows.

$$F_J(x, y) = k1 \cdot lA_J(x, y) + k2 \cdot lB_J(x, y) \tag{1}$$

In Eq. (1), $k1$ and $k2$ are given parameters, if image B is fused into A, then $k1 > k2$ and if A is fused into B, then $k1 < k2$. Now $F_J(x, y)$ and $F_j^k(x, y)$ are used to reconstruct and generate the fused image $F'(x, y)$ which contains high frequency as well as low frequency information of $A(x, y)$ and $B(x, y)$ [14, 15, 16, 17].

2.3 Dimensionality reduction

Face images are very similar, and therefore highly correlated. So that they can be represented in a much lower dimensional feature subspace. For this study two different dimensionality reduction method is applied here, namely principal component analysis (PCA) and independent component analysis (ICA)**.**

2.4 Principal Component Analysis

Principal component analysis (PCA) is based on the second-order statistics of the input image, which tries to attain an optimal representation that minimizes the reconstruction error in a least-squares sense. Eigenvectors of the covariance matrix of the face images constitute the eigenfaces. The dimensionality of the face feature space is reduced by selecting only the eigenvectors possessing significantly large eigenvalues. Once the new face space is constructed, when a test image arrives, it is projected onto this face space to yield the feature vector – the representation coefficients in the constructed face space. The classifier decides for the identity of the individual, according to a similarity score between the test image's feature vector and the PCA feature vectors of the individuals in the database[12].

2.5 Independent Component Analysis

Independent component analysis (ICA) has many applications in data analysis, source separation and feature extraction. Here, Independent Component Analysis (ICA) is implemented for dimensionality reduction of face images. Independent Component Analysis (ICA) is a technique which is mainly used for subspace projection and which is projects data from high dimensional to low-dimensional space[3, 5, 6, 7]. Independent Component Analysis (ICA) is the generalization of the Principal component analysis (PCA), which decorrelates the high-order statistics in addition to second-order moments.

To rigorously define ICA[4], we can use a statistical "latent variables" model. Assume that we observe n linear mixtures $x_1, \cdots, x_n$ of n independent components

$$x_j = a_{j1}s_1 + a_{j2}s_s + \cdots + a_{jn}s_n \text{ for all j} \tag{2}$$

It is convenient to use vector-matrix notation instead of the sums like in Eq. (2). Let us denote by x the random vector whose elements are the mixtures $x_1, \cdots, x_n$, and likewise by s the random vector with elements $s_1, \cdots, s_n$. Let us denote by A the matrix with elements a_{ij}. All vectors are understood as column vectors; thus x^T, or the transpose of x, is a row vector. Using this vector-matrix notation, the above mixing model is written as

$$x = As \tag{3}$$

Sometimes we need the columns of matrix A; denoting them by a_j the model can also be written as

$$x = \sum_{i=1}^{n} a_i s_i \tag{4}$$

The statistical model in Eq. (4) is called independent component analysis, or ICA model.

2.6 Two ICA architectures used in experiment

There are atleast two ways by which ICA can be applied on the face images for recognition[3], in this experiment a matrix X which contains the fused image database. In the matrix X all the images are representing as a row vector. In this approach images are random variables and pixels are trails. This same approach is used by Bell and Sejnowski for sound source separation[8] and this approach is called as independent component analysis architecture I (i.e. ICA-I). In second architecture, the transpose of matrix X is done. In this architecture the organization of face database of that images are in the column of X. In this approach, pixels are random variables and images are trails. This approach is inspired by Bell and Sejnowski's work on the IC's (Independent Components) of natural images[9] and this approach of ICA is called independent component analysis architecture II (i.e. ICA-II).

2.7 Implementing ICA I & ICA II by FastICA

A number of popular ICA algorithms exist. These include FastICA[5, 7], Infomax[8, 17], Comon's algorithm[4], and KernalICA[18]. The success of ICA for a given data set may depend crucially on performing some application-dependent pre-processing steps. The FastICA is used here for implementing the two architectures of independent component analysis. The FastICA algorithm has many advantages over other independent component analysis (ICA) algorithms.The FastICA algorithm and the underlying contrast functions have a number of desirable properties when compared with existing methods for ICA. The FastICA algorithm directly calculates independent components (ICs). The independent components can be estimated one by one, which is roughly equivalent to doing projection pursuit. This is useful in exploratory data analysis, and decreases the computational load of the method in cases where only some of the independent components need to be estimated. The FastICA has most of the advantages of neural algorithms: It is parallel, distributed, and computationally simple and requires little memory space.

2.8 Classifier

Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyze. The Back propagation learning algorithm is one of the most historical developments in Neural Networks. It has reawakened the scientific and engineering community to the modeling and processing of many quantitative phenomena using neural networks. This learning algorithm is applied to multilayer feed forward networks consisting of processing elements with continuous differentiable activation functions. Such networks associated with the back propagation learning algorithm are also called back propagation networks.

3. EXPERIMENT RESULTS AND DISCUSSION

This work has been simulated using MATLAB 7.9.0.529 (R2009b) in a machine of the configuration 2.13GHz Intel Xeon Quad Core Processor and 16384.00MB of Physical Memory. We analyze the performance of our algorithms using the IRIS thermal / visual face database.

3.1 IRIS thermal / visual face database

In this database, all the thermal and visible unregistered face images are taken under variable illuminations, expressions, and poses. The actual size of the images is 320 x 240 pixels (for both visual and thermal). 176-250 images per person, 11 images per rotation (poses for each expression and each illumination). Total 30 classes are present in that database and the size of the database is 1.83 GB[11].

The detailed overview of the training and testing processes is presented here. To verify the efficiency of the algorithms and to compare them on the basis of experimental results same dataset is used in training and testing. Two different datasets are prepared against expressions and illuminations changes from IRIS thermal/visible face database. Total 220 thermal and 220 visual faces of 10 different classes (i.e. 22 thermal and 22 visual images per class) are randomly selected. These images have been first separated into two groups namely training and testing set (i.e. each set contains 110 thermal and 110 visual images of 10 different classes). Second, the decomposition and reconstruction processes are applied on the images of training and testing set using orthogonal wavelet filters. Third, fused images of training and testing set are generated using Daubechies wavelet coefficient (db2). The whole process of decomposition, reconstruction, and fusion is discussed in section 2. Fourth, two different algorithms for dimensionality reduction methods are used here. First one is Principal Component Analysis (PCA) called as method-I and in the second one, two different architectures of Independent Component Analysis (ICA) are shown, called as method-II. In two different ways, these two dimensionality reduction methods are applied on the face images. In case of PCA (i.e. method-I) all the training set images are first plotted into a single matrix, where each row represents an image. Then PCA is applied on each row of the matrix and calculated reduced size of each image is 40 x 1 (original size of the images is 40 x 50 pixels). This dimension reduction process is applied on the training set images only.

But in case of the method-II (i.e. two architectures of ICA), the dimensionality reduction process is applied in a different way. The ICA algorithm is first applied on the face images and then plotted into a single data matrix. The calculation of dimension of images using method-II is mainly dependent on the dataset. That way calculation of dimension for each class image is varying with the changes of data with varying illumination and changes with expression. In Table 2, the calculated reduced size of each class images using FastICA algorithm with their required steps to calculate the covariance is given. From this table, one can easily justify that reduced size of class images varying and number of iterations (1000 is the highest value) is also changes. The highest picked eigenvalues of classes 1 to 10 are 0.1, 0.23, 0.04, 0.23, 0.27, 0.26, 0.35, 0.24, 0.52 and 0.25 respectively; and the minimum of picked eigenvalue is equal to the dimension of the input vector and the largest eigenvalue is equal to 1 and eigenvalues changes with the changes in dataset. Finally, training images is learned with the help of a classifier and in the next section different experimental results was presented.

3.2 Experimental results with changes in expression

Among 220 fused images 110 (i.e. 11 images of 10 classes) images are used to train the network and rest 110 (i.e. 11 images of 10 classes) images are kept for checking the efficiency of system (i.e. by testing). Out of 110 images 66 fused images (i.e. only 6 classes out of 10 classes) are subjected to expression only. After the completion of training, 66 fused images (i.e. of 6 different classes) from testing set are taken for recognition (data of changes expression only). In Table 1, recognition rates using method I (PCA) and method II (ICA) on the selected 'successful' subbands images of expression are given. Among 6 classes, the highest recognition rate is 91% for 4 different classes; class–1, class–3, class–4 and for class–5 using Method I i.e. by PCA. The achieved rate for class-1 is the maximum recognition rate which means it is 9% more than result achieved for the data fusion using method-I (i.e. PCA) and the recognition rates for two classes (class-3 and class-5) by using ICA-II architecture is also the maximum recognition rate i.e. 100%.

3.3 Experimental results with changes in illumination

Among 10 different classes, 4 classes (i.e. 44 images) are having the data which change in illumination with normal facial expression. In Table 2, the entire recognition rates based on data of illumination using method I as a PCA and method II using ICA-I and ICA-II architecture are shown. In this case also the maximum recognition rate i.e. 100% is achieved for two different classes (i.e. class-2 and class-6) using PCA and the maximum recognition rate is achieved using method-II (i.e. ICA architecture-I) for 3 different classes (class-2, class-6, and class-10). But for ICA architecture-II the maximum rate is also achieved for only one class (class-7).

Table 1. Experimental results changes with expression.

Classes used	No. of Training Images	No. of Testing Images	PCA	ICA-I	ICA-II
Class – 1	11	11	91%	100%	82%
Class – 3	11	11	91%	91%	100%
Class – 4	11	11	91%	91%	73%
Class – 5	11	11	91%	91%	100%
Class – 8	11	11	73%	82%	82%
Class – 9	11	11	82%	91%	91%

Table 2. Experimental results changes with illumination.

Classes used	No. of Training Images	No. of Testing Images	PCA	ICA-I	ICA-II
Class – 2	11	11	100%	100%	91%
Class – 6	11	11	100%	100%	91%
Class – 7	11	11	91%	91%	100%
Class – 10	11	11	91%	100%	82%

3.4 Comparison of results between two methods

After the analysis of success rates based on data which changes with expression and illumination, it can easily be justified that method-II (i.e. Independent Component Analysis Architecture-I) shows better performance than other two methods i.e. PCA and ICA architecture-II. In case of data with changes in expression, the average recognition rate among all the classes is 86.755% for method-I i.e. by PCA and 91.135% by ICA-I which is more than 4% from PCA and 2% from ICA architecture-II. On the other hand data changes in illumination the average recognition rate is 95.5% using method-I (i.e. PCA), 91% using ICA architecture-II and 97.75% on ICA architecture-I which is more than 2% from method-I and over 6% from ICA architecture II. The overall recognition rates of two methods are 91.13% on method-I and 94.44% on ICA-I and 89.72% using ICA-II architecture.

3.5 Experiment results with changes in expression and illumination

In this section, three experiments have been conducted using PCA (expt. 1), using ICA-I (expt. 2) and using ICA-II (expt. 3), on all the 10 classes together, i.e., on the combined dataset of different expression and illumination. Here, both the training (110 fused images) and testing (110 fused images) sets contain both types (different expression and illumination) of images. In the first experiment using PCA all the 110 images are recognized i.e., 100% recognition rate has been achieved while using PCA as the feature extraction technique. While second experiment using ICA-I, 100% recognition rate has been achieved for three classes; class–1, class–4 and class–10 and the average recognition rate is 90%. For third experiment using ICA-II, class–3, class–8 and class–10 gives 100% recognition rate and again a rate of 90% is achieved on an average. Table 3 shows all the recognition rates achieved for the 10 classes on the three experiment using three different feature extraction methods, which indicates that the first experiment using PCA performs better and recognizes all the 110 images used for testing, while the other two experiments using ICA-I and ICA-II, fails to recognize all the testing images and in both cases gives 10% less average recognition rate in comparison to the experiment using PCA.

4. CONCLUSION

In this paper, a novel approach to recognize the faces of varying changes in expression and illumination is presented using data level information fusion. Here, fusion has been done on Daubechies Wavelet Transform. For this study, the IRIS thermal/visible face database is used and which contains the faces having different expression with illumination. After completion of fusion, the two different dimensionality reduction algorithms were used i.e., PCA and ICA (ICA-I & ICA-II). Those projected fused images are classified using a Multilayer Perceptron. Both the systems have achieved a maximum recognition rate of 100% on varying illumination condition as well as change of expression.

Table 3. Experimental results of changes in expression and illumination.

Classes used	No. of Training Images	No. of Testing Images	PCA	ICA-I	ICA-II
Class – 1	11	11	100%	100%	91%
Class – 2	11	11	100%	82%	91%
Class – 3	11	11	100%	82%	100%
Class – 4	11	11	100%	100%	91%
Class – 5	11	11	100%	91%	91%
Class – 6	11	11	100%	91%	73%
Class – 7	11	11	100%	91%	82%
Class – 8	11	11	100%	82%	100%
Class – 9	11	11	100%	82%	82%
Class – 10	11	11	100%	100%	100%

ACKNOWLEDGMENT

Authors are thankful to a major project entitled "Design and Development of Facial Thermogram Technology for Biometric Security System," at Department of Computer Science and Engineering, Jadavpur University, India funded by University Grants Commission(UGC), India for providing necessary infrastructure to conduct experiments relating to this work.

REFERENCES

[1] Kim, J., Choi, J., Yi, J. and Turk, M., "Effective representation using ICA for face recognition robust to local distortion and partial occlusion," IEEE Trans. PAMI. 27(12), 1977-1981 (Dec. 2005).
[2] Dagher, I. and Nachar, R., "Face recognition using IPCA-ICA algorithm," IEEE. Tran. PAMI. 28(6), 996-1000 (June 2006).
[3] Barlett, M. S., Movellan, J. R. and Sejnowski, T. J., "Face recognition by Independent Component Analysis," IEEE Trans. Neural Networks, 13(6), 1450-1464 (2002).
[4] Comon, P., "Independent Component Analysis, a new concept," Signal Processing, 36(3), 287-314 (1994).
[5] Hyvarinen, A., "Fast and robust fixed-point algorithms for Independent Component Analysis," IEEE Trans. on Neural Networks, 10(3) (May 1999).
[6] Hyvarinen, A. and Oja, E., "Independent Component Analysis: algorithms and applications," Neural Networks Research Centre, Helsinki University of Technology.
[7] Hyvarinen, A., Karhunen, J. and Oja, E., "Independent Component Analysis," John Wiley & Sons, (2001).
[8] Bell, A. J. and Sejnowski, T. J., "An information-maximization approach to blind separation and blind decondolution", Neural Compute, 7(6), 1129-1159 (1995).
[9] Bell, A. J. and Sejnowski, T. J., "The independent components of natural scenes are edge filters," Vision Res., 37(23), 3327-3338 (1997).
[10] Heo, J., Kong, S., Abidi, B. and Abidi, M., "Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition," IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2004, Washington, D.C., 94-99 (July 2004).
[11] http://www.cse.ohio-state.edu/otcbvs-bench/Data/02/download.html.
[12] Ekenel H. K. and Sankur, B., "Multiresolution face recognition," Image and Vision Computing, vol. 23, 469-477 (2005).
[13] Goh, Y. Z., Teoh, A. B. J. and Gog, M. K. O., "Wavelet based illumination invariant preprocessing in face recognition," Proc. of the 2008 Congress on Image and Signal Processing, vol. 3, IEEE Computer Society, 421-425 (2008).
[14] Mallat, S., "A theory for multiresolution signal decomposition: the wavelet representation," IEEE PAMI, 11(7), 674-693 (1989).
[15] Chui, C. K., "An introduction to wavelets," Academic Press, Inc., America, (1992).
[16] Daubechies, I., "Orthonormal bases of compactly supported wavelets, communication on pure and applied mathematics," vol. XLI, 909-996 (1988).
[17] Lee, T. W., Girolami, M. and Sejnowski, T. J., "Independent Component Analysis using an extended Infomax algorithm for mixed Sub-Gaussian and Super-Gaussian sources," Neural Computation, 11(2), 417-441 (1999).
[18] Bach, F. R. and Jordan, M.I., "Kernel Independent Component Analysis," J. Machine Learning Research, vol. 3, 1-48 (Mar. 2002).
[19] Wasniowski, R. A., "Using data fusion for biometric verification," World Academy of Science, Engineering and Technology 5 (2005).
[20] Kong, S. G., Heo, J., Abidi, B.R., Paik, J. and Abidi, M. A., "Recent advances in visual and infrared face recognition - A review," Computer Vision and Image Understanding, 97(1), 103 – 135 (Jan. 2005).
[21] Socolinsky, D. A. and Selinger, A., "A comparative analysis of face recognition performance with visible and thermal infrared imagery," Proc. Int. Conf. on Patt. Recog., Quebec, vol. 4, 217-222 (2002).
[22] Yoshitomi, Y., Miyaura, T., Tomita, S. and S. Kimura, "Face identification using thermal image processing," Proc. IEEE Int. Workshop on Robot and Human Communication, 374-379 (1997).
[23] Prokoski, F., "History, current status, and future of infrared identification," Proc. IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications, 514 (2000).

[24] Selinger, A. and Socolinsky, D. A., "Appearance-based facial recognition using visible and thermal imagery: a comparative study," Technical Report, Equinox Corporation (2001).
[25] Wolff, L. B., Socolinsky, D. A. and Eveland, C. K., "Face recognition in the thermal infrared," Technical Report, Equinox Corporation (2002).
[26] Wolff, L. B., Socolinsky, D. A. and Eveland, C. K., "Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery," Proc. CVPR Workshop on Computer Vision Beyond the Visible Spectrum (2001).