Performance Evaluation through KICA and Feature level Fusion for Human Face Recognition

Priya Saha, Mrinal Kanti Bhowmik, Barin Kumar De
Department of Computer Science & Engineering
Department of Physics
Tripura University (A Central University)
Suryamaninagar, Tripura, India
priyasaha.cse@gmail.com, mkb_cse@yahoo.co.in, barin_de@yahoo.com

Debotosh Bhattacharjee, Mita Nasipuri
Department of Computer Science & Engineering
Jadavpur University
Kolkata, India
debotosh@indiatimes.com, mitanasipuri@yahoo.com

Abstract—**This paper concentrates on kernel features of fused images of visual and thermal images in the wavelet domain. Here, Daubechies wavelet transform, termed as db4, is used to get fused coefficients from gradient visual and corresponding thermal image. After applying Gaussian filtering, fusion of the gradient extracted from the gray level visual image and the corresponding gray level thermal image is performed. These fused images are passed to Kernel Independent Component Analysis (KICA) to extract robust and independent low level features from the original image data. KICA is introduced to overcome the complex, nonlinear variations due to illumination changes, view point changes etc. These feature extracted fused images are then classified using support vector machine multiclass method. The experiment is conducted for performance evaluation of the independent face features using KICA on IRIS Thermal/Visual Face Database. Experimental results show that independent face features achieve higher recognition rate in polynomial kernel than linear kernel.**

Keywords-fused image; daubechies wavelet transform; kernel independent component analysis (KICA); support Vector Machine (SVM); classification; IRIS database

I. INTRODUCTION

Face recognition has been extensively spread within the research and development communities in the last decade. The researchers for face recognition have developed a large number of methods. However, these methods give poor performance under variable illumination, pose and facial expression.Visual image based face recognition are used the most at present, but in different lighting condition, visual images lose their quality. To tackle these problems, thermal images are gaining much interest in these days because of their illumination invariant properties. Thermal infrared (IR) images have several advantages over conventional visible images. Most of the light in the mid-to-long wave IR is emitted rather than reflected [1], [2]. Thermal images also have some limitations like facial hair, glasses, or cosmetics. Therefore, to extract the advantages from both types of images, fusion is the most constructive notion as it combines multiple information sources together and produces a more informative representation of the data [3].

Hanif et al. [4] have discussed data fusion of thermal and visual images to overcome the drawbacks present in individual thermal and visual images. Bebis et al. [5] discussed that infrared (IR) imagery offers a promising alternative to visible imagery, due to its relative insensitivity to variations in face appearance caused by illumination changes. Thermal IR images have several shortcomings including that it is opaque to glass. A detailed overview on wavelet transform for image fusion using image decomposition and reconstruction has been put forward by M. K. Bhowmik et al. in [6]. M. K. Bhowmik et al. [7] discussed different wavelet transformations of fused images. Coefficients of discrete wavelet transforms from both optical and thermal images were computed separately and combined. They also used inverse discrete wavelet transformation in order to obtain fused face image. Haar and Daubechies (db2) wavelet transforms have also been used. M. K. Bhowmik et al. [8] have been investigated Quotient based fusion of thermal and visual images, which were passed separately through level-1 and level-2 multi-resolution analyses. This approach, based on a definition of an illumination invariant signature image, facilitates an analytic generation of the image space with varying illumination. Principal Component Analysis (PCA) was used for dimension reduction of quotient fused images and then, those images are classified using a multilayer perceptron (MLP). The fusion of visual and thermal IR images is presented here to enhance the robustness of face recognition. Visual gradient image and corresponding thermal image are fused using db4 wavelet transform which has been described in section II. Then the fused image is used by KICA for feature extraction. These extracted images are classified using SVM. Different researchers have been examined the performance of KICA on face images, which has shown better recognition result than the other linear methods. J. H. Cao et al. [9] investigated face recognition method using KICA and Kernel-based improved PSVM that results better performance over native ICA method. Y. Huang et al. [10] proposed a new gabor based Kernel Independent Component analysis (GKICA) which got higher recognition rate than ICA and KICA.

The contribution of the paper is to provide systematic KICA based method to classify the fused images of visual and infrared images for efficient face recognition system. To cope up with the non-linear variations due to illumination changes, we have introduced gradient based wavelet fusion and nonlinear feature extraction method, KICA in our work. The paper is organized as follows: system overview has been given in Section II, which has demonstrated gradient based image fusion in the wavelet domain, nonlinear extension of independent component analysis, algorithmic steps of KICA and classification using SVM. Section III and IV are all about the discussion of the experimental results using IRIS database. Finally, the conclusion is made in section V.

II. SYSTEM OVERVIEW

The complete system implementation of this work is described with a block diagram, which is shown in Figure 1. At first, 2-D discrete wavelet transform of thermal and gradient visual images has been completed. After that, the fusion of coefficients (i.e. approximate and details) of thermal and visual images have been generated. These methods are described below.

A. Gradient based Feature Level Fusion using Daubechies Wavelet

It is appropriate to use Gaussian filtering [11] to smooth an image as well as generate the gradient images. According to the Tikonov regularization theory, Gaussian smoothing is a good approximation of the optimal polynomials to regularize a signal [12]. Regularization step is performed before the actual computation of partial derivatives of the image intensity and the effects of image noise are greatly reduced while preserving the integrity of the differential. After the completion of applying Gaussian filtering, fusion of the gradient extracted from the gray level visual image and the corresponding gray level thermal image is carried out.

If, I be an image under variable lighting conditions, then Gradient image (GI) of image I is defined as in (1).

$$G = \tan^{-1}\left(\frac{Iy}{Ix}\right), G \in [0, 2\pi] \quad (1)$$

where I_x and I_y are the gradients of image I in the X and Y directions respectively [13]. We have performed the fusion of gradient visual and thermal image fusion because the gradient of the gray levels is partially insensitive to illumination changes (shadows can still produce spurious edges), but the combination with the thermal frequency component improves the robustness to illumination variation.

This visual image has been smoothed first with Gaussian filter to compute the gradient. The smooth image is produced using the convolution type of smoothing. In the next step, smooth image has been computed to find out the gradient in the x and y direction. By taking the inverse tangent of the image gradients in Y and X direction respectively, the gradient images have been generated. The algorithm to generate the gradient images is given below.

Input: Image I. Steps: 1. Creation of Gaussian Filter G, 2. To generate the smooth image I_G by convolving the input image I with Gaussian Filter. I_G = I * G, where, * is the convolution operator, 3. Find the gradients G_X and G_Y of image I_G in X and Y direction respectively, $I_X = I_G * Grad_X(x, y, \sigma)$ and $I_Y = I_G * Grad_Y(x, y, \sigma)$, Where, $Grad_X(x, y, \sigma)$ and $Grad_Y(x, y, \sigma)$ are the derivative of Gaussian Kernel function in X and Y direction respectively and σ is the standard deviation, 4. Calculating the inverse tangent of the image component, Output: Gradient Image I_{Grad}.

Visual gradient image and corresponding thermal image are fused using db4 wavelet transform. In Daubechies wavelet transform termed as db4, the image fusion is based on the combination of wavelet decompositions of the two original images. The fusion algorithm is as follows: the two images are to be processed and resampled to the one with the same size; then, decomposition of images into the sub-images is made using forward wavelet transform having the same resolution at the same levels and different resolution at different levels; information fusion is carried out based on the high-frequency sub-images of decomposed images, and finally, inverse wavelet transform is applied to obtain the resultant image [14]. Next the fused image is used by KICA for feature extraction, which is described in the next section.

B. Kernel Independent Component Analysis

The KICA algorithm that was primarily developed for separating randomly mixed auditory signals [15], is efficient enough as a feature extraction method in face recognition. Kernel Independent Component Analysis (KICA) is a method based on kernel function, which allows data space to be mapped into high dimensional feature space, and also an improved feature extraction method from existing ICA. The concept of kernel in ICA has been introduced in face images due to the failure of linear methods of feature extraction and KICA combines the strength of both kernel and ICA. Linear methods like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA), which are three powerful tools largely used for dimension reduction, and feature extraction in the appearance-based approaches, are inadequate to describe complex non-linear variations due to illumination changes, facial expressions and aging [16], [17]. Non-linear extension of linear methods has been initiated to eliminate these problems. We have brought in the KICA as a feature extraction method since much of the information may be contained in the high order dependences among the pixels in the face image [18]. The steps of KICA are presented here as follows.

1) Centering of Fused Image: Let MS be a mixed signal with m sources and n samples, $MS = (a_{ij})_{m \times n}$, then we have to calculate the mean vector, shift= $\frac{1}{p}\sum_{i=1}^{n} a_{ij} = E\{x\}$ centered image= $MS - shift$,

2) Whitening and Initializing of De-mixing matrix: Now we have to apply orthogonal triangular de-composition on the mean centered image and produce one unitary matrix Q and one upper triangular matrix R. $Q = (\pm u_{ij})_{m \times n}$, $R = (\pm ut_{ij})_{m \times n}$, here the elements below diagonal are zero. Apply singular value decomposition on transpose of upper triangular such that $R' = theta * lamda * V, R' = (\pm ut_{ij})_{m \times n}$ where elements above the diagonal are zero. Here, theta and V are two unitary

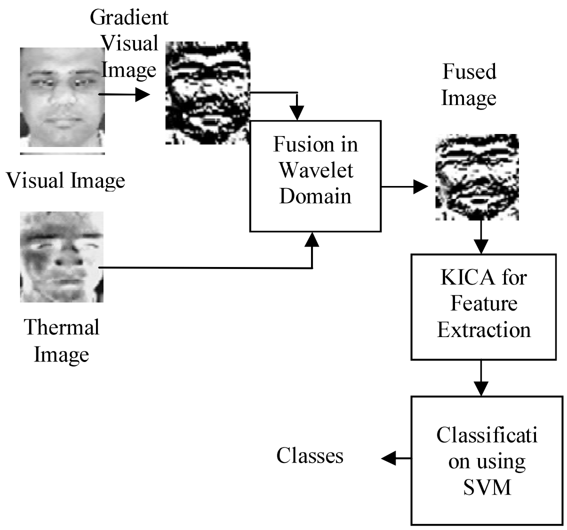

Figure 1. System Overview

matrices and lamda is diagonal matrix with non-negative values in decreasing order D_L = transpose of lamda and then make it diagonal. Whitening signal $ww = square\ root\ of\ n * D_L * theta$, Final whitening signal $ws = ww * MS$.

(a) Memory Initialization for storing de-mixing matrix and Hilbert Space Independent Component (HSIC) at each iteration: De-mixing matrix $XS = \sum_{k=1}^{itra} \sum_{i=j=1}^{n} (x_{ij})^k\ for\ all\ x_{ij} = 0$, itra=maximum number of iteration required for better performance of the algorithm, Hilbert Space Independent Component (HSIC) = $\sum_{j=1}^{itra} x_{ij}$, $for\ all\ j's, i = 1\ and\ x_{ij} = 0$,signal $ws = ww * MS$.

3) Define the width of Gaussian kernel:

$sigmas = \sum_{j=1}^{m} x_{ij}$, $for\ all\ j's, i = 1\ and\ x_{ij} = 1$.

4) Memory utilization by Incomplete cholesky: An incomplete Cholesky factorization is given by a sparse lower triangular matrix K that is in some sense close to L (lower triangular matrix). The corresponding pre-conditioner is KK^*. One popular way to find such a matrix K is to use the algorithm to find the exact Cholesky decomposition, except that any entry is set to zero if the corresponding entry in A is zero. This gives an incomplete Cholesky factorization [19].

5) Independence measures based on RKHS covariance operator: By introducing the problem variables, and describing covariance operator in RKHS. The spectral and Hilbert-Schmidt norms of the covariance operators may be expressed as a function of the kernels of the RKHSs. The Hilbert space $\mathcal{F}$ is an RKHS if at each x $\in \mathcal{X}$, the point evaluation operator $\delta_x : \mathcal{F} \to \mathbb{R}$ which maps $f \in \mathcal{F}$ to $f(x) \in \mathbb{R}$ is a bounded linear functional. The covariance operator $C_{x,y} : \mathcal{G} \to \mathcal{F}$ is defined such that for all $f \in \mathcal{F}$ and g $\in \mathcal{G}$.

$$\langle f, C_{x,y}\ g \rangle_{\mathcal{F}} = E_{x,y}\,[[f(x) - E_x[f(x)]]\ [g(y) - E_y\,[g(y)]]].$$

6) Computation of Euclidean Gradient: Let U be an open subset of $\mathbb{R}^d$, and let $\varepsilon : U \to \mathbb{R}$ be a differentiable function. The Euclidean Gradient of ε is the function $\nabla_{euc}\varepsilon(u) \in \mathbb{R}^d$ such that

$$\varepsilon'(u)v = \langle \nabla_{euc}\varepsilon(u), v \rangle_{euc} \text{ For every } v \in \mathbb{R}^d$$

7) Compute Approximate Hessian: The Hessian matrix is the square matrix of second-order partial derivatives of a function, which describes the local curvature of a function of many variables.

Given the real-valued function $f(x_1, x_2, \ldots\ldots, x_n)$, If all second partial derivatives of f exist, then the Hessian Matrix of f is the matrix

$$H(f)_{i,j}(x) = D_i D_j f(x)$$

Where $x = (x_1, x_2, \ldots\ldots, x_n)$ and D_i is the differentiation operator with respect to the i^{th} argument.

8) Finally check for convergence and find the final De-mixing matrix.

C. Classification using Support Vector Machine

Support Vector Machine is a supervised learning method used for classification and regression which performs classification tasks by constructing optimal separating hyperplanes in multidimensional space [20]. SVM approach was primarily used to classify linearly separable data, but later in the case of linearly inseparable data, kernel methods are introduced to nonlinearly map the input data to a high dimensional space [21]. For classification purpose, we have used multiclass SVM to carry out recognition on face images, and Sequential Minimal Optimization (SMO) algorithm to train the SVM.

Kernels that have been used here are given below:

Linear: $k(x_i, x_j) = (x_i^T x_j + c)$

Polynomial: $k(x_i, x_j) = (1 + x_i^T x_j)^p$

III. Experimental Discussion

A. Database Description

The benchmark database OTCBVS (Object Tracking and Classification Beyond Visible Spectrum) contains both thermal and visual face images under the same situation like pose, illumination etc. It contains different sets of data like OSU Infrared (IR) Pedestrian Database, IRIS Infrared (IR)/Visible Face Database, OSU Color-Infrared (IR) Database, Terravic Facial IR Database, Terravic Weapon IR Database, and CBSR NIR Face Dataset. Among all of these different datasets, IRIS (Imaging, Robotics and Intelligent System) Infrared (IR)/Visible face dataset has only been considered in our experiments. This database contains simultaneously acquired unregistered infrared (IR) and visual face images under varying illuminations, expressions, and poses [22].

B. Image Registration

Image registration is the method of transforming two images into the same coordinate system. It's a basic task in image processing for the purpose of aligning two different images [23]. We have done registration of Visual and Thermal IR images before they are fused. In our experiment, we have taken images from different expressions and illumination types of each person from IRIS database and registered the thermal and visual images separately with respect to the frontal image of each expression and illumination types

IV. EXPERIMENT RESULTS

The whole experiment consists of expression, illumination, and full datasets with 660 images in 20 classes, 748 images with 17 classes and 1626 images with 28 classes respectively. For the sake of experiment, the registered visual and thermal images are then cropped into 50×50 dimensions. The cropped gradient visual and corresponding thermal images are fused using wavelet transformation, which is described previously. After that, all the fused images are taken for feature extraction using KICA. From feature extracted expression dataset, out of 660 images, total numbers of training and testing images are 462 and 198 respectively. We have taken 523 and 225 face images for training and testing purposes respectively from feature extracted illumination face image dataset. 1138 and 488 face images from feature extracted full dataset are taken for training and testing purpose respectively. For training support vector machine, Sequential Minimal Optimization (SMO) algorithm has been used where tolerance is 1.0000e-003 with which KKT conditions are checked, maximum number of iterations of the algorithm is 20,000 and size of the kernel matrix cache is 1000. Linear kernel and Polynomial kernel with degree 3 are used in our SVM kernel functions. The accuracy of a classification process defined as the portion of true positives, and true negatives in the population of all instances, classification accuracy A = (TP+TN) / (TP+TN+ FP+FN). TP=True Positives, TN=True Negatives, FP=False positives and FN=False Negatives. All these values have been taken from the confusion matrix acquired from the classifier performance. "True positives" are active compounds which have correctly been classified as active, and "false positives" are inactive compounds which have wrongly been classified as Active [24]. The average classification accuracies of two SVM kernels are listed in table I. The whole experimentation is divided into 3 parts; one is SVM classification with KICA extracted expression face image dataset, another is SVM classification with KICA extracted illumination face image dataset and the last one is SVM classification with KICA extracted full dataset. Among the two kernels, polynomial kernel produces better accuracy than linear kernel in all three cases. Now considering the average classification accuracy of expression face dataset, SVM with linear kernel generates 91.14% and polynomial kernel is 2.31% more than linear kernel. In illumination dataset, highest average recognition rate is obtained using polynomial kernel. Classification accuracy of linear kernel is 5.01% less than polynomial kernel, which generates 91.81% accuracy in full dataset. So, it is clearly observed that polynomial kernel is more efficient than linear kernel in our experiment. Comparative study in different fusion techniques reveals the improvement of recognition accuracy in our present method. Table II shows the comparative study.

TABLE I. EXPERIMENTAL RESULTS

Dataset	**Recognition Rate**	
	Linear	***Polynomial***
Expression	91.14%	93.28%
Illumination	85.16%	89.00%
Fulldata	86.80%	91.81%

V. CONCLUSION AND FUTURE WORK

This paper formulates Daubechies wavelet (db4) for fusion of gradient optical and infrared image, KICA for feature extraction method and SVM for classification purpose in human face classification. We have used IRIS thermal/visible face database. The visual and thermal face images have their own advantages and disadvantages. Therefore, we have tried to combine the advantages from visual and thermal image. The experimental results show that polynomial kernel gives better classification accuracy than linear kernel. We can conclude that our approach has achieved efficient classification performance for face recognition purpose. We will further investigate different wavelet fusion methods and other classifiers for face recognition.

ACKNOWLEDGMENT

The research has been supported by the grant from DeitY, MCIT, Govt. of India, Vide No. 12(2)/2011-ESD, dated 29/03/2011.

TABLE II. COMPARATIVE STUDY

Methods	**Recognition Rate**	
Present Method (Gradient DB4 fusion+ KICA)	Expression	93.28%
	Illumination	89.00%
	Fulldata	91.81%
Fusion of Thermal and Visual [25]	90.00%	
Abs max selection in DWT [4]	90.31%	
Fusion of Visual and LWIR + PCA [26]	87.87%	
Wavelet Subband + Kernel associative Memory with XM2VTS database [27]	84.00%	

REFERENCES

[1] L. Wolff, D. Socolinsky, and C. Eveland, "Quantitative measurement of illumination invariance for face recognition using thermal infrared imagery," *Proc. CVPR Workshop on Computer Vision Beyond the Visible Spectrum*, Dec. 2001.

[2] C. K. Eveland, Diego A. Socolinsky and L. B. Wolff, "Tracking Human Faces In Infrared Video", *Image and Vision Computing*, vol. 21, No. 7, pp. 579-590, July 2003.

[3] R. Maruthi and K. Sankarasubramanian, "Multi focus image fusion based on the information level in the region of the images," *Journal of theoretical and applied information technology*, 2005.

[4] M. Hanif, U. Ali,"Optimized visual and thermal image fusion for efficient face recognition", IEEE Conf. on Information Fusion, Florence, 2006, p. 1-6.

[5] G. Bebis, A. Gyaourova, S. Singh, and I. Pavlidis. "Face Recognition by Fusing Thermal Infrared and Visible Imagery," *image and vision computing*, Vol. 24, p-727-742, 2006.

[6] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, "Fusion of Daubechies Wavelet Coefficients for Human Face Recognition", *International Journal of Recent Trends in Engineering*, Vol. 3, No.1, pp.137-143, May 2010.

[7] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu, "Fusion of Wavelet Coefficients from Visual and Thermal Face Images for Human Face Recognition – A Comparative Study", *International Journal of Image Processing (IJIP)*, vol. 4, Issue 1, pp.12-23, Mar. 2010.

[8] M. K. Bhowmik, D. Bhattacharjee, M. Nasipuri, D. K. Basu and M. Kundu,"Quotient Based Multiresolution Image Fusion of Thermal and Visual Images Using Daubechies Wavelet Transform for Human Face Recognition", *International Journal of Computer Science Issues (IJCSI)*, Vol. 7, Issue 3, pp. 18-27, May 2010.

[9] J.H. Cao, Y. Z. Ran, Z. J. Xu," A New Kernel Function Based Face Recognition Algorithm", 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE), Changchun, pp. 226 – 230, 2010.

[10] Y. Huang, M. Li, C. Lin, L. Tian," Gabor-based Kernel Independent Component Analysis for Face Recognition", Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), Darmstadt, pp. 376 – 379, 2010.

[11] R. A. Haddad and A. N. Akansu, "A class of fast Gaussian Binomial Filters for speech and image processing," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 39, pp. 723-727, Mar. 1991.

[12] T. Poggio, V. Torre and C. Koch, "Computational vision and regularization theory", *Nature*, Vol. 317, pp. 314-319, Sept. 1985.

[13] T. Zhang, Y. Y. Tang, B. Fang, Z. Shang and X. Liu, "Face recognition under varying illumination using gradient faces", *IEEE Trans. on Image Processing*, vol. 18, No. 11, pp. 2599-2606, 2009.

[14] Z, Shu-long, "Image Fusion using Wavelet Transform", *Symposium on Geospatial Theory, Processing and Applications*, pp.5-9, 2004.

[15] T. Martiriggiano, M. Leo, T. D"Orazio and A. Distante, Face Recognition by Kernel Independent Component Analysis, Lecture Notes in Artificial Intelligence, pp. 55-58, 2005, vol. 3533.

[16] A. Turk, A. Pentland, "Eigenfaces for recognition" *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[17] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[18] M.-H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods", in *Proc. of IEEE Internation Conf. on Face and Gesture Recognition*, 2002, Washington, DC, p. 215–220.

[19] G. H. Golub, C. F. Van Loan, *Matrix Computations*, 3rd ed., John Hopkins University Press, Baltimotre, Md, 1996.

[20] http://www.statsoft.com/textbook/support-vector-machines/

[21] S. N. Sivanandam, S. Sumathi, and S. N. Deepa, *Introduction to Neural Networks using Matlab 6.0*, 2nd Ed.,Tata McGraw-Hill Pub. Co., 2006.

[22] http://www.cse.ohio-state.edu/otcbvs-bench/.

[23] L G. Brown, "A survey of image registration techniques." *ACM Computing Surveys*, vol. 24, no. 4, pp. 325-376, 1992.

[24] A. G. K. Janecek, W. N. Gansterer, M. A. Demel and G. F. Ecker, "On the Relationship Between Feature Selection and Classification Accuracy," in *Proc. of JMLR Workshop and Conference*, 2008, vol. 4, p. 90-105.

[25] A. Singh, G. B. Gyaourva and I. Pavlidis,"Infrared and Visible Image Fusion for Face Recognition," *Proc. SPIE,* 2004, vol.5404, pp. 585-596.

[26] I. Pavlidis, and P. Symosek,"The Imaging Issue in an Automatic Face/Disguise Detecting System," *Proc. of IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, 2000, pp. 15 – 24.

[27] M. Marchand, M. Golea and P. Rujan," A convergence theorem for sequential learning in two-layer perceptrons," *Europhys.Lett,* 1990, 11, pp. 487-492.