

Research Article

A Gabor-Block-Based Kernel Discriminative Common Vector Approach Using Cosine Kernels for Human Face Recognition

Arindam Kar,¹ Debotosh Bhattacharjee,² Dipak Kumar Basu,² Mita Nasipuri,²
and Mahantapas Kundu²

¹Indian Statistical Institute, Kolkata 700108, India

²Department of Computer Science and Engineering, Jadavpur University, Kolkata 700032, India

Correspondence should be addressed to Arindam Kar, kgparindamkar@gmail.com

Received 14 March 2012; Revised 16 July 2012; Accepted 13 August 2012

Academic Editor: Samanwoy Ghosh-Dastidar

Copyright © 2012 Arindam Kar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper a nonlinear Gabor Wavelet Transform (GWT) discriminant feature extraction approach for enhanced face recognition is proposed. Firstly, the low-energized blocks from Gabor wavelet transformed images are extracted. Secondly, the nonlinear discriminating features are analyzed and extracted from the selected low-energized blocks by the generalized Kernel Discriminative Common Vector (KDCV) method. The KDCV method is extended to include cosine kernel function in the discriminating method. The KDCV with the cosine kernels is then applied on the extracted low-energized discriminating feature vectors to obtain the real component of a complex quantity for face recognition. In order to derive positive kernel discriminative vectors, we apply only those kernel discriminative eigenvectors that are associated with nonzero eigenvalues. The feasibility of the low-energized Gabor-block-based generalized KDCV method with cosine kernel function models has been successfully tested for classification using the L_1 , L_2 distance measures; and the cosine similarity measure on both frontal and pose-angled face recognition. Experimental results on the FRAV2D and the FERET database demonstrate the effectiveness of this new approach.

1. Introduction

Face authentication has gained considerable attention in the near past through the increasing need for access verification systems using several modalities like voice, face image, fingerprints, pin codes, and so forth. Such systems are used for the verification of a user's identity on the Internet, when using automated banking system, or when entering into a secured building, and so on. The Gabor wavelet transformation (GWT) models well the receptive field profiles of the cortical simple cells and also has the properties of multiscale and multidirectional filtering. These properties are in accordance with the characteristics of human vision [1–3]. Further, the discriminant analysis is an effective image feature extraction and recognition technique as they allow the extraction of discriminative features, reduce dimensionality, and consume less computing time [4, 5]. In our previous work [6], we combined the GWT and Bayesian principal component analysis (PCA) techniques and presented a GWT-Bayesian PCA face recognition method which outperforms some

conventional linear discriminating methods. As an extension of linear discriminant technique, the kernel based nonlinear discriminant analysis technique has now been widely applied to the field of pattern recognition. Baudat and Anouar [7] developed a commonly used generalized discriminant analysis (GDA) method for nonlinear discrimination. Jing et al. [8] put forward a Kernel Discriminative Common Vectors (KDCVs) method. In this paper we develop blockbased GWT KDCV and propose a block-based low-energized nonlinear GWT discriminant feature extraction for enhanced face recognition. As the high energized blocks of GWT image generally have larger nonlinear discriminability values. Then the nonlinear discriminant features are extracted from the selected low-energized block of GWT image by presenting a new generalized KDCV method is then extended to include cosine kernel model which extracts the nonlinear discriminating features from the selected blocks to get the best recognition result. These features are finally used for classification using three different classifiers. The experimental results demonstrate the effectiveness of this new approach.

In this paper a novel method is proposed based on selecting low-energized blocks of Gabor wavelet responses as feature points, which contain discriminate facial feature information, instead of using predefined graph nodes as in elastic graph matching (EGM) [9], which reduces representative capability of Gabor wavelets. This corresponds to enhancement of edges for eyes, mouth, nose, which are supposed to be the most important points of a face; hence the algorithm allows these facial features to keep overall face information along with local characteristics.

The remainder of this paper is organized as follows. Section 2 describes the derivation of low-energized blocks from the GWT images. Section 3 details the generalized KDCV method with cosine kernel function for enhanced face recognition. Section 4 shows the performance of the proposed method on the face recognition by applying it on the datasets from the FERET [10], and FRAV2D [11] face databases, and by comparing it with some of the previous KDCV methods and we conclude our paper in Section 5.

2. 2D Gabor Wavelets

Gabor wavelets are used in image analysis because of their biological relevance and computational properties [12, 13]. The Gabor transform is suitable for analyzing gradually changing data such as the face, iris, and eyebrow images. The Gabor filter used here has the following general form:

$$\varphi_{\mu,\nu}(z) = \frac{|k_{\mu,\nu}|^2}{\sigma^2} e^{-\|k_{\mu,\nu}\|^2 \|z\|^2 / 2\sigma^2} \left[e^{ik_{\mu,\nu}z} - e^{-(\sigma^2/2)} \right], \quad (1)$$

where μ and ν define the orientation and scale of Gabor kernels, respectively, $z = (x, y)$ is the variable in spatial domain, $\|\cdot\|$ denotes the norm operator, and $k_{\mu,\nu}$ is the frequency vector which determines the scale and orientation of Gabor kernels, $k_{\mu,\nu} = k_\nu e^{i\phi_\mu}$ where $k_\nu = k_{\max}/f^\nu$ and $k_{\max} = \pi/2$, $\phi_\mu = \pi\mu/8$, $\mu = 0, \dots, 7$, where f is the spacing factor. Here Gabor wavelets at five different scales, $\nu \in \{0, \dots, 4\}$ and eight orientations $\mu \in \{0, \dots, 7\}$ are chosen. The term $e^{-(\sigma^2/2)}$ is subtracted from (1) in order to make the kernel DC-free, thus becoming insensitive to illumination. The magnitude of the convolution outputs is indicated as $O_{\mu,\nu}(z)$. The kernels exhibit strong characteristics of spatial locality and orientation selectivity, making them a suitable choice for image feature extraction when one's goal is to derive local and discriminating features for (face) classification.

2.1. Gabor-Based Feature Representation. The Gabor wavelet representation of an image is the convolution of the image with a family of Gabor kernels as defined in (1). Let $I(x, y)$ be the gray level distribution of an image, the convolution output of image I and a Gabor kernel $\varphi_{\mu,\nu}$ is defined as

$$O_{\mu,\nu}(z) = I(z) * \varphi_{\mu,\nu}(z), \quad (2)$$

where $z = (x, y)$, and $*$ denotes the convolution operator.

Applying the convolution theorem, convolution outputs are derived from (2) via the Fast Fourier Transform (FFT):

$$\begin{aligned} \mathfrak{F}\{O_{\mu,\nu}(z)\} &= \mathfrak{F}\{I(z)\} \mathfrak{F}\{\varphi_{\mu,\nu}(z)\}, \\ O_{\mu,\nu}(z) &= \mathfrak{F}^{-1}\{\mathfrak{F}\{I(z)\} \mathfrak{F}\{\varphi_{\mu,\nu}(z)\}\}, \end{aligned} \quad (3)$$

Here \mathfrak{F} , \mathfrak{F}^{-1} denote the forward and inverse discrete Fourier transforms, respectively. The outputs $O_{\mu,\nu}(z)$, where $\nu \in \{0, \dots, 4\}$, $\{\mu \in \{0, \dots, 7\}\}$, consist of different local, scale, and orientation features in both real and imaginary parts in the specific locality as described later in Section 2.2. The magnitude of $O_{\mu,\nu}(z)$ is defined as modulus of $O_{\mu,\nu}(z)$, that is, $\|O_{\mu,\nu}(z)\|$.

2.2. Low-Energized Block Based GWT Feature Extraction.

It is to be noted that we considered the magnitude of $O_{\mu,\nu}(z)$, but did not use the phase, which is consistent with the application of Gabor representations [14, 15]. As the outputs ($O_{\mu,\nu}(z) : \mu \in \{0, \dots, 4\}, \nu \in \{0, \dots, 7\}$) consist of 40 different local scale and orientation features, the dimensionality of the Gabor transformed image space is very high. So the following technique is applied for the extraction of low-energized discriminability feature vector χ_k from the convolution outputs. The method for the extraction of low-energized block based features from the GWT image is explained in Algorithm 1.

Algorithm 1. Consider

Step 1. Find the convolution outputs of the original image with all the Gabor kernels. As the convolution outputs contain complex values, so replace each pixel value of the convolution output by its modulus and the resultant image is termed as G_I , where $I = 1, 2, \dots, k$, $k = \text{total Gabor kernels}$.

Step 2. Obtain the final single Gabor transformed image $I_{GF} = \sum_{I=1}^k G_I$, $k = \text{no. of Gabor kernels}$.

Step 3. Compute the overall mean (\bar{g}) of the final Gabor transformed image I_{GF} as, $\bar{g} = (1/(m \times n)) \sum_{x,y} I_{GF}(x, y)$, where $m \times n$ is the size of image.

Step 4. Divide the final Gabor transformed image I_{GF} into windows of size $\omega \times \omega$. Thus the total number of windows, $l = \lfloor m/\omega \rfloor \times \lfloor n/\omega \rfloor$.

Step 5. For each window w_i , if minimum (w_i) $\leq \bar{g}$, then extract a block B_i of size $c \times c$ from w_i , with centre pixel as the minimum (w_i). The value of c must be odd integer and less than $\omega/2$.

Step 6. For each window w_i , if minimum (w_i) $\leq \bar{g}$, and there does not exist block B_i of size $c \times c$ from w_i as mentioned in Step 5, with centre pixel as the minimum (w_i), then create a block B_i of size $c \times c$, by considering the unavailable pixel values as \bar{g} .

Step 7. For each window w_i , if minimum (w_i) $> \bar{g}$, then create a pseudo block B_i of size $c \times c$ with all elements as \bar{g} .

Step 8. Extract feature vector f_i from each block B_i in a systematic order, where f_i contains all elements of the block B_i .

Step 9. Concatenate all the feature vectors f_i , $i = 1, 2, \dots, l$ to obtain the final feature vector χ , which is the final extracted low-energized feature vector. This extracted feature vector χ encompasses the low valued discriminable elements of the Gabor transformed image, and the size of this feature vector is [(Total no. of blocks) \times (size of the block)] which is much lower in dimension in comparison to the original image (dimension: $m \times n$) and the GWT image (dimension: $k \times m \times n$).

Thus this augmented Gabor feature vector encompasses most of the discriminable feature elements of the Gabor wavelet representation set, $S = (O_{\mu,v}(z) : \mu \in \{0, \dots, 4\}, v \in \{0, \dots, 7\})$. The window size $\omega \times \omega$ is one of the important features of the above algorithm, as it must be chosen small enough to capture most of the important features and large enough to avoid redundancy. Since it is observed that there are some windows each of whose minimum value is not less than the overall mean, so Step 7 is applied in order not to get stuck on a local minimum.

In the experiments we took a window and a block of size 7×7 and 3×3 , respectively, to extract the low-energized feature vector. Thus the extracted facial features can be compared locally, instead of using a general structure, allowing us to make a decision from the parts of the face.

3. Generalized Kernel Discriminative Common Vector (KDCV) Method

Sometimes the discriminative common vectors are not distinct in the original sample space. In such cases one can map the original sample space to a higher-dimensional space F , where the new discriminative common vectors in the mapped space are distinct from one another. This is because a mapping, $\Phi : R^N \rightarrow F$, $x \rightarrow \phi(x)$, can map two vectors that are linearly dependent in the original sample space onto two vectors that are linearly independent in F . As the mapped space can have arbitrarily large, possibly infinite, dimensionality, hence it is reasonable to use the DCV method.

Let $\Phi = \{\Phi(x_1^1), \Phi(x_2^1) \dots \Phi(x_{N_1}^1), \Phi(x_1^2) \dots \Phi(x_{N_c}^c)\}$ represent the matrix whose columns are the transformed training samples in F . Here c is the number of training classes; the i th class contains N_i samples. The within-class scatter matrix S_W^Φ , the between-class scatter matrix S_B^Φ , and the total scatter matrix S_T^Φ in F are given by

$$\begin{aligned} s_W^\Phi &= (\phi - \phi G)(\phi - \phi G)^T; \\ s_B^\Phi &= (\phi U - \phi L)(\phi U - \phi L)^T; \\ s_T^\Phi &= (\phi - J_M \phi)(\phi - J_M \phi)^T = S_B^\Phi + S_W^\Phi, \end{aligned} \quad (4)$$

where μ^Φ is the mean of all samples, and μ_i^Φ is the mean of samples of the i th class in F . Here $G = \text{diag}(G_1, G_2, \dots, G_c) \in$

$R^{M \times M}$ is a block-diagonal matrix and each $G_i \in R^{N_i \times N_i}$ is a matrix with all its elements equal to $1/N_i$; $U = \text{diag}(u_1, u_2, \dots, u_c)$ is a block-diagonal matrix and each $u_i \in R^{N_i \times 1}$ is a vector with all its elements equal to $1/N_i$; $L = \text{diag}(l_1, l_2, \dots, l_c) \in R^{M \times C}$ is a block-diagonal matrix and each $l_i \in R^{M \times 1}$ is a vector with the entries $\sqrt{N_i}/M$; $J_M \in R^{M \times M}$ is a matrix with entries $1/\sqrt{M}$. The aim of the DCV algorithm is to acquire the optimal projection transform W in the null space of S_W [16]:

$$J(W) = \arg \max_{W^T S_W W = 0} |W^T S_B W| = \arg \max_{W^T S_W W = 0} |W^T S_T W|. \quad (5)$$

The approach for computing this optimal projection vector is as follows.

Step 10. Project the training set samples onto the range $R(S_T^\Phi)$ of S_T^Φ through the Kernel PCA.

Step 11. Find vectors V that span the null space of \tilde{S}_W^Φ .

Step 12. Remove the null space of $V^T \tilde{S}_B^\Phi V$ if it exists.

Step 13. Obtain the final projection matrix W , which will then be $W = (\Phi - \Phi 1_M) U \Lambda^{-1/2} V L$, where Λ is the diagonal matrix with nonzero eigenvalues, U , the associated matrix of normalized eigenvectors, and V is the basis for the null space of \tilde{S}_W^Φ , here there are at most $(C - 1)$ projection vectors.

Let the common vector be $\Phi(x_{\text{com}}^i)$, then each of the feature vectors can be written as $\Phi(x_m^i) = \Phi(x_{\text{com}}^i) + \Phi(x_{\text{diff}}^i)$, where $\Phi(x_{\text{com}}^i) \in V^\perp$, and $\Phi(x_{m,\text{diff}}^i) \in V$.

Here $\Phi(x_{\text{com}}^i)$ and $\Phi(x_{m,\text{diff}}^i)$ represent the common and different parts of $\Phi(x_m^i)$ separately. It has been proved by Gülmezoğlu et al. [17] that for all samples of the i th class, their common vector parts are same. The common vector can be written as $\Phi(x_{\text{com}}^i) = \Phi(x_m^i) - \Phi(x_{m,\text{diff}}^i)$.

Thus, a set of common vectors for is obtained as:

$$Q = \{\Phi(x_{\text{com}}^1), \Phi(x_{\text{com}}^2) \dots \Phi(x_{\text{com}}^c)\}. \quad (6)$$

Compute the optimal projection transform. Let $\tilde{S}_{\text{com}}^\Phi$ denote the total scatter matrix of Q . W_{com} is composed of the eigenvectors corresponding to the positive eigenvalues of $\tilde{S}_{\text{com}}^\Phi$. W_{com} is designed to satisfy the criteria:

$$J(W) = \arg \max_{W^T \tilde{S}_W W = 0} W^T \tilde{S}_W W \quad (7)$$

$$J(W) = \arg \max_{W^T \tilde{S}_W W = 0} W^T S_{\text{com}}^\Phi W$$

W_{diff} is calculated from the different vectors. W_{diff} is composed of the eigenvectors corresponding to the positive eigenvalues of $\tilde{S}_{T,\text{diff}}^\Phi$. The optimal projection transform W is obtained as

$$W = W_{\text{diff}} + W_{\text{com}}. \quad (8)$$

Thus for each sample $\Phi(x)$ in the kernel space using the generalized nonlinear KDCV method, we construct W and

then extract the kernel discriminative common and different vector Y_{com} and Y_{diff} . Then, $Y_{\text{com}} = W_{\text{com}}^T \Phi(x)$, and $Y_{\text{diff}} = W_{\text{diff}}^T \Phi(x)$. Finally,

$$Y = Y_{\text{diff}} + Y_{\text{com}}. \quad (9)$$

Thus we obtain a new sample set Y corresponding to X . This sample set Y is used for image classification. All mathematical properties of the linear DCV are carried over to the kernel DCV method with the modifications that are applied to the mapped samples, $\Phi(x_m^i)$, $i = 1, 2, \dots, c$, where $c = 1, 2, \dots, N_i$. After performing the feature extraction, all training set samples of each class typically give rise to a single distinct discriminative common vector.

3.1. KDCV Approach Using Cosine Kernel Function. Let $\chi_1, \chi_2, \dots, \chi_n \in \mathbb{R}^N$ be the data in the input space, and Φ be a nonlinear mapping between the input space and the feature space; $\Phi : \mathbb{R}^N \rightarrow F$. Generally three classes of kernel functions are used for nonlinear mapping: (a) the polynomial kernels, (b) the Radial Basis Function (RBF) kernels, and (c) the sigmoid kernels [18].

The RBF kernels, are also known as isotropic stationary kernels, are defined by $\Phi : [0, \infty) \rightarrow \mathbb{R}$ such that $k(x, x') = \Phi(\|x - x'\|)$, where $x, x' \in \chi \subset \mathbb{R}$, and $\|\cdot\|$ is the norm operator. Normally a Gaussian function is preferred as the RBF, in most of the RBF kernels in pattern classification applications. The Gaussian function for RBF kernels is given by $k(x, x') = \exp(-\|x - x'\|/\sigma^2)$. But the globally used RBF kernels yield dense Gram matrices, which can be highly ill-conditioned for large datasets.

So in this work the cosine kernel function is considered as the kernel function Φ , defined by

$$\Phi(x, y) = \frac{\pi}{4} \cos\left(\frac{\pi(x \cdot y)}{2}\right). \quad (10)$$

This result can be expressed in terms of the angle θ between the inputs: $\theta = \cos^{-1}(\Phi(x, y)/\sqrt{(\Phi(x, x))\Phi(y, y)})$. This shows that this kernel has a dependence on the angle between the inputs.

As a practical matter, we note that cosine kernels do not have any continuous tuning parameters (such as kernel width in RBF kernels), which can be laborious to set by cross validation.

Large margin classifiers are known to be sensitive to the way features are scaled [19]. Therefore it is essential to normalize either the data or the kernel itself. The recognition accuracy can severely degrade if the data is not normalized [19]. Normalization can be performed at the level of the input features or at the level of the kernel. It is often beneficial to scale all features to a common range, for example, by standardizing the data. An alternative way to normalize is to convert each feature vector into a unit vector. If the data is explicitly represented as vectors one can normalize the data by dividing each vector by its norm such that $\|x\| = 1$, after normalization. Here normalization is performed at the level of the kernel, that is, normalizing in feature-space, leading to $\phi(x) = 1$ (or equivalently that $k(x, x) = 1$). This is accomplished by using the cosine kernel which normalizes a kernel

$k(x, x')$ to $k_{\text{cosine}}(x, x') = k(x, x')/\sqrt{k(x, x) \cdot k(x', x')}$. Normalizing data to unit vectors reduces the dimensionality of the data by one since the data is projected to the unit sphere.

In order to derive positive kernel nonlinear discriminative features (9), we consider only those eigenvectors that are associated with positive eigenvalues.

4. Similarity Measures and Classification

Finally the lower-dimensional, low-energized extracted feature vector of the GWT image is used as the input data instead of the whole image in the proposed method to derive the kernel discriminative feature vector, W , using (8). Let M'_k be the mean of the training samples for class w_k , where $k = 1, 2, \dots, l$ where l is the number of classes. The classifier then applies, the nearest neighbor (to the mean) rule for classification using the similarity (distance) measure δ :

$$\delta(Y, M'_j) = \min_k(\delta(Y, M'_k)) \implies Y \in w_j. \quad (11)$$

The low-energized KDCV vector Y_L is classified to that class of the closest mean M'_k using the similarity measure δ . The similarity measures used here are L_1 distance measure, δ_{L_1} , L_2 distance measure, δ_{L_2} , and the cosine similarity measure, δ_{cos} , which are defined as

$$\delta_{L_1} = \sum_i |X_i - Y_i|, \quad (12)$$

$$\delta_{L_2} = (X - Y)^T (X - Y), \quad (13)$$

$$\delta_{\text{cos}} = \frac{-X^T Y}{\|X\| \|Y\|}, \quad (14)$$

where T is the transpose operator and $\|\cdot\|$ denotes the norm operator. Note that the cosine similarity measure includes a minus sign in (14) because the nearest neighbour (to the mean) rule (11) applies minimum (distance) measure rather than maximum.

4.1. Experiments of the Proposed Method on Frontal and Pose-Angled Images for Face Recognition. This section assesses the performance of the low-energized Gabor-block-based KDCV method for both frontal and pose-angled face recognition. The effectiveness of the low-energized block based KDCV method is successfully tested on FRAV2D and FERET databases. For frontal face recognition, the data set is taken from the FRAV2D database, which consists of 1100 frontal face images corresponding to 100 individuals. The images are acquired, with partially occluded face features and different facial expressions. For pose-angled face recognition, the data set taken from the FERET database contains 2200 images of 200 individuals with different facial expressions and poses. Further studies have been made on the FERET dataset using the standard protocols, that is, the Fa, Fb, DupI, and DupII set to assess the performance of the proposed method.

4.1.1. FRAV2D Face Database. The FRAV2D face database, employed in the experiment, consists of 1100 colour face

TABLE 1: Average recognition results using FRAV2D database.

Method	Recognition accuracy (%)		Average recognition rates (%)
	No. of training samples		
	3	4	
GWT	85.5	89.5	87.5
KDCV	79.5	82	80.75
GWT-LDA	88.3	90.33	89.33
GWT-KDCV (RBF)	87	90	88.5
GWT-KDCV (Cosine)	88	90	89
GWT-KDCV (RBF, low-energized)	90	92.5	91.25
GWT-KDCV (Cosine, low-energized)	96.125	97.75	96.9

TABLE 2: Specificity and sensitivity measure of the FRAV2D dataset.

Total no. of classes = 100, total no. of images = 1800			
Individual belonging to a particular class			
Using first 3 images of an individual as training images			
		Positive	Negative
FRAV2D test	Positive	$T_P = 759$	$F_P = 7$
	Negative	$F_N = 41$	$T_N = 693$
		Sensitivity = $T_P/(T_P + F_N) \approx 94.875\%$	Specificity = $T_N/(F_P + T_N) = 99.0\%$
Using first 4 images of an individual as training images			
		Positive	Negative
FRAV2D test	Positive	$T_P = 678$	$F_P = 2$
	Negative	$F_N = 22$	$T_N = 698$
		Sensitivity = $T_P/(T_P + F_N) \approx 96.85\%$	Specificity = $T_N/(F_P + T_N) \approx 99.7\%$

So considering the first 4 images in Figures 2(a)–2(d) of a particular individual for training the achieved rates are as follows.

False positive rate = $F_P/(F_P + T_N) = 1 - \text{Specificity} = .3\%$.

False negative rate = $F_N/(T_P + F_N) = 1 - \text{Sensitivity} = 3.15\%$.

Accuracy = $(T_P + T_N)/(T_P + T_N + F_P + F_N) \approx 98.3\%$.

So considering the first 3 images in Figures 2(a)–2(c) of a particular individual for training the achieved rates are as follows.

False positive rate = $F_P/(F_P + T_N) = 1 - \text{Specificity} = 1\%$.

False negative rate = $F_N/(T_P + F_N) = 1 - \text{Sensitivity} = 5.125\%$.

Accuracy = $(T_P + T_N)/(T_P + T_N + F_P + F_N) \approx 96.9\%$.

images of 100 individuals, 11 images of each individual are taken, including frontal views of faces with different facial expressions, under different lighting conditions. All colour images are transformed into gray images and are scaled to a size of $m \times n$ here (92×112) is used. The details of the images are as follows: Figure 1(a) regular facial status; Figures 1(b) and 1(c) are images with a 15° turn with respect to the camera axis; Figures 1(d) and 1(e) are images with a 30° turn with respect to the camera axis. Figures 1(f) and 1(g) are with gestures, such as smiles, open mouth, winks, laughs; Figures 1(h) and 1(i) are images with occluded face features; Figures 1(j) and 1(k) are images with change of illumination. Figure 3 shows all samples of one individual.

4.1.2. Specificity and Sensitivity Measures for the FRAV2D Dataset. To measure the sensitivity and specificity [23] the dataset from the FRAV2D database is prepared in the following manner. For each individual a single class is constituted with 18 images. Thus a total of 100 classes are obtained, from the dataset of 1100 images of 100 individuals.

Out of the 18 images in each class, 11 images are of a particular individual, and 7 images are of other individuals taken by permutation as shown in Figure 2. Using this dataset the true positive (T_P); false positive (F_P); true negative (T_N); false negative (F_N) are measured. From the 11 images of the particular individual, at first the first 4 images Figures 2(a)–2(d), then the first 3 images Figures 2(a)–2(c) of a particular individual are selected as training samples and the remaining images of the particular individual are used as positive testing samples. The negative testing is done using the images of the other individuals. Figure 2 shows all sample images of one class of the dataset used from FRAV2D database.

4.1.3. FERET Face Database. The FERET database, employed in the experiment here, contains 2,200 facial images corresponding to 200 individuals with each individual contributing 11 images. The images in this database were captured under various illuminations, which display a variety of facial expressions and poses. As the images include the background and the body chest region, so each image is cropped to

TABLE 3: Average recognition results using FERET database.

Method	Recognition rates (%)		Average recognition rates (%)
	No. of training samples		
	3	4	
GWT	79.65	82.5	81.1
KDCV	69.5	75	72.25
GWT-LDA	82.76	85.33	84.1
GWT-KDCV (RBF)	81	82.50	81.75
GWT-KDCV (Cosine)	83	85	84.5
GWT-KDCV (RBF, low-energized)	88.5	91.5	90
GWT-KDCV (Cosine, low-energized)	92.5	95.75	94.1

TABLE 4: Specificity and sensitivity measure of the FERET dataset.

Total no. of classes = 200, Total no. of images = 3600			
Individual belonging to a particular class			
Using first 3 images of an individual as training images			
		Positive	Negative
FERET test	Positive	$T_P = 1492$	$F_P = 16$
	Negative	$F_N = 108$	$T_N = 1384$
		Sensitivity = $T_P/(T_P + F_N) \approx 93.25\%$	Specificity = $T_N/(F_P + T_N) \approx 98.8\%$
Using first 4 images of an individual as training images			
		Positive	Negative
FERET test	Positive	$T_P = 1316$	$F_P = 10$
	Negative	$F_N = 84$	$T_N = 1390$
		Sensitivity = $T_P/(T_P + F_N) \approx 94\%$	Specificity = $T_N/(F_P + T_N) \approx 99.28\%$

So considering the first 4 images in Figures 4(a)–4(d) of a particular individual for training the achieved rates are:

False positive rate = $F_P/(F_P + T_N) = 1 - \text{Specificity} = .72\%$.

False negative rate = $F_N/(T_P + F_N) = 1 - \text{Sensitivity} = 6\%$.

Accuracy = $(T_P + T_N)/(T_P + T_N + F_P + F_N) \approx 96.6\%$.

So considering in the first 3 images Figures 4(a)–4(c) of a particular individual for training the achieved rates are:

False positive rate = $F_P/(F_P + T_N) = 1 - \text{Specificity} = 1.2\%$.

False negative rate = $F_N/(T_P + F_N) = 1 - \text{Sensitivity} = 6.75\%$.

Accuracy = $(T_P + T_N)/(T_P + T_N + F_P + F_N) \approx 96.1\%$.

exclude those, and are transformed into gray images and is scaled to $m \times n$ here (92×112) is used. Figure 3 shows all samples of one subject. The details of the images are as follows: Figure 3(a) regular facial status; Figure 3(b) $+15^\circ$ pose angle; Figure 3(c) -15° pose angle; Figure 3(d) $+25^\circ$ pose angle; Figure 3(e) -25° pose angle; Figure 3(f) $+40^\circ$ pose angle; Figure 3(g) -40° pose angle; Figure 3(h) $+60^\circ$ pose angle; Figure 3(i) -60° pose angle; Figure 3(j) alternative expression; Figure 3(k) different illumination.

First 4 images of each individual, that is, Figures 3(a)–3(d) are regarded as training samples. The remainders are regarded as testing samples. After that 3 images of each individual, that is, Figures 3(a)–3(c) are regarded as training samples.

4.1.4. Specificity and Sensitivity Measure for the FERET Dataset. To measure the sensitivity and specificity, the dataset from the FERET database is prepared in the following manner. For each individual a single class is constituted with 18 images. Thus a total of 200 classes are obtained, from the dataset of 2200 images of 200 individuals. Out of

these 18 images in each class, 11 images are of a particular individual, and 7 images are of other individuals taken using permutation as shown in Figure 4. Similarly using this dataset from the FERET dataset the specificity and sensitivity are being measured. From the 11 images of the particular individual, at first the first 4 images Figures 4(a)–4(d), then the first 3 images Figures 4(a)–4(c) of a particular individual are selected as training samples and the remaining images of the particular individual are used as positive testing samples. The negative testing is done using the images of the other individuals. Figure 4 shows all sample images of one class of the dataset used from FERET database.

4.2. Further Evaluation on the FERET Face Dataset. We further use the FERET face database for testing our proposed method, as it is one of the most widely used databases to evaluate face recognition algorithms [20]. FERET contains a gallery set, Fa, and four testing sets: Fb, Fc, DupI, and Dup II. In the Fa set, there are 1196 images and contains one image per individual. The Fb set has 1195 images of people with

TABLE 5: Recognition results of different algorithms on the FERET probe sets.

Method	FERET probe sets			
	Fb (%)	Fc (%)	Dup I (%)	Dup II (%)
Phillips et al. [20]	96	82	59	52
Local Gabor binary pattern histogram sequence, [21]	98	97	74	71
Grassmann registration manifolds for face recognition, [22]	98	98	80	84
Low-energized Gabor-block-based KDCV with RBF kernels (Gauss) using cosine measure.	96	97	79	81
Low-energized Gabor-block-based KDCV with cosine kernels using L_1 measure.	95	96	79	80
Low-energized Gabor-block-based KDCV with cosine kernels using L_2 measure.	96	97	86	82
Low-energized Gabor-block-based KDCV with cosine kernels using cosine measure.	98	98	89	84

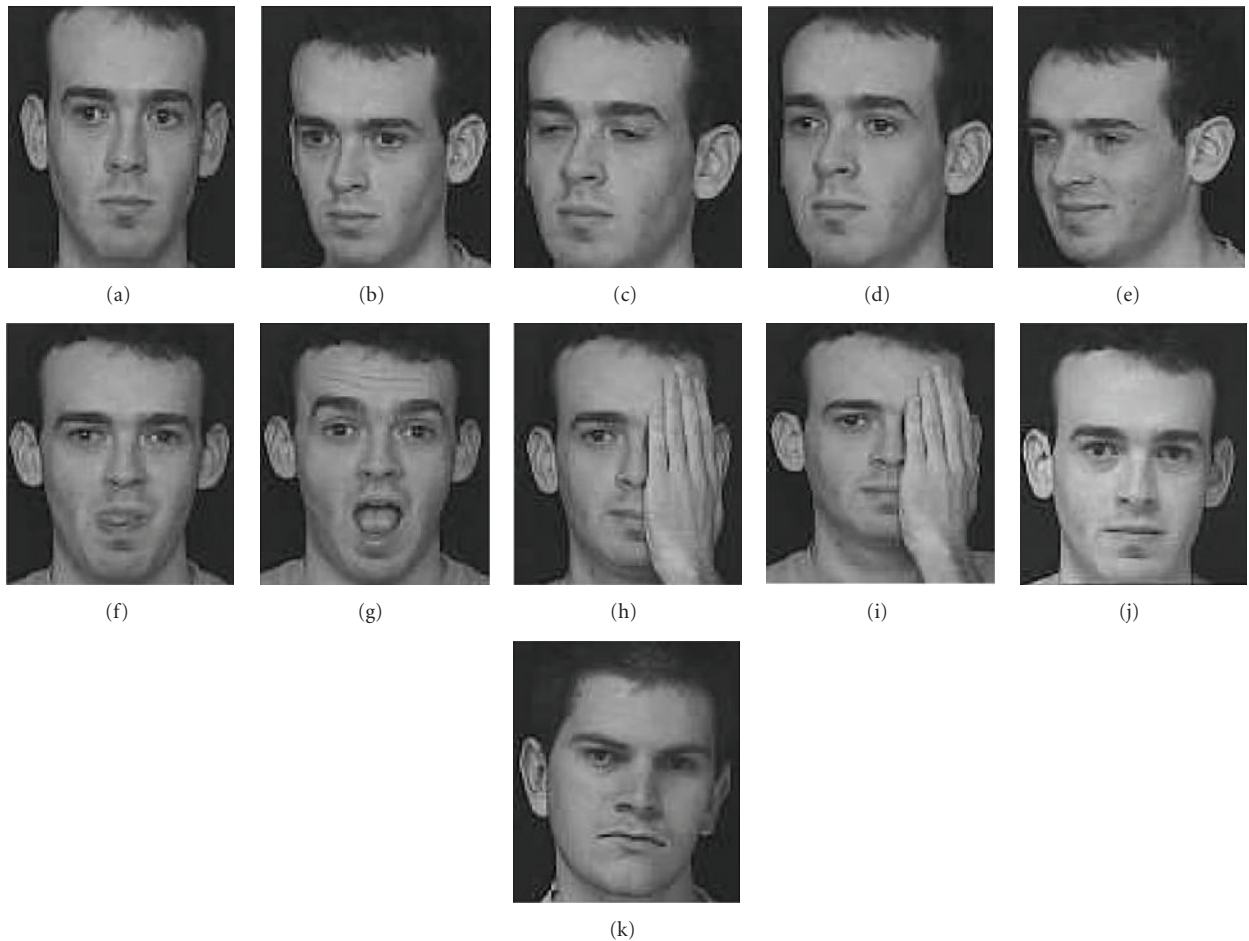


FIGURE 1: Demonstration images of an individual from the FRAV2D database.

different gestures. In the Fc set, there are 194 images under different lighting conditions. The DupI set has 722 images of pictures taken between 0 and 34 months of difference with those taken for Fa. The DupII set contains 234 images acquired at least 18 months after the Fa set. DupII is a subset of DupI. Figure 4 shows samples from FERET face database. All images are aligned and cropped to 112×92 according to [24].

The extracted low-energized Gabor feature vector is considered as input to a trained KDCV with cosine kernel

and its output is compared with a gallery set using the L_1 , L_2 , and cosine similarity measure. The recognition rates of different methods on the FERET probe sets are shown in Table 5. The results are compared with the most recent state-of-the-art with the FERET database. Our results with the FERET database are equivalent (with a difference of $\pm 1\%$) to the most recent works on the FERET dataset. Note that the methods described in [21, 22] use the Gabor wavelets to generate their feature vectors. As the Gabor wavelets have a much higher algorithmic complexity, so the overall

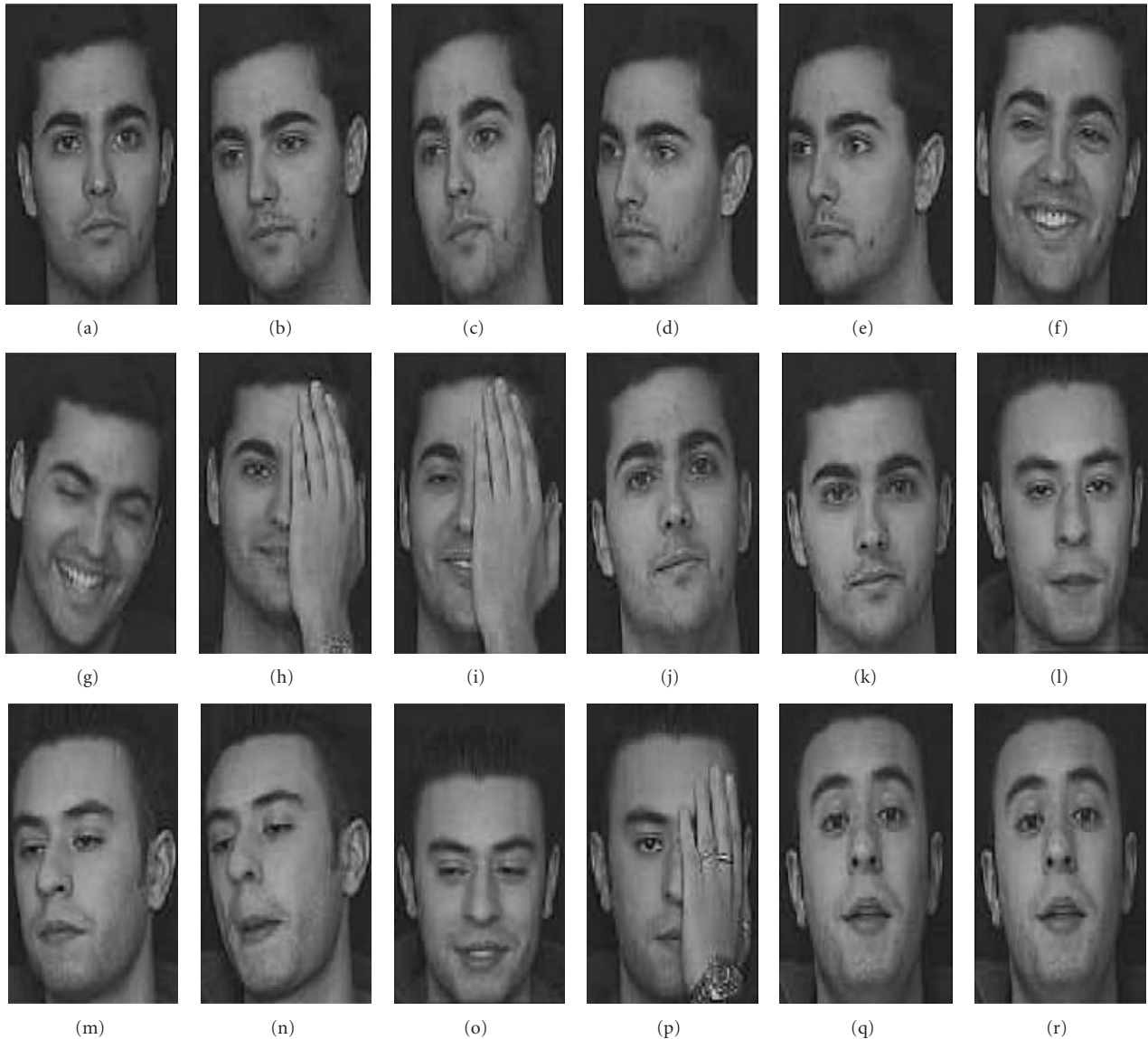


FIGURE 2: Demonstration images of one class from the FRAV2D database.

computing cost is very high. On the other hand, our block-based low-energized Gabor feature is a very low-dimensional feature vector which reduces the algorithmic complexity. Also with the use of cosine kernel as the kernel function in the KDCV makes the proposed method quite fast and more suitable to real applications.

Furthermore, we compare the face recognition performance of our proposed low-energized GWT-KDCV method using cosine kernels, with some other well-known methods like generalized discriminant analysis (GDA) method [7], (EGM) [9], Discrete Cosine transformation (DCT) and linear discriminant analysis (LDA), DCT-LDA method [4], DCT-GDA [23], GWT-LDA, DCT-KDCV method [25], and the Gabor fusion KDCV method [8]. Classification results obtained from the proposed method are comparable or even better in some cases than above-mentioned methods. Also the cosine similarity measure is more suitable for classifying

the nonlinear real KDCV features shown in the tables, Tables 1, 2, 3, 4, 5, and 6.

As the proposed method performs best with the cosine similarity classifier, so the specificity rate of the proposed method is evaluated for the FERET and FRAV2D dataset using the cosine similarity measure shown in Figure 7.

4.3. Results. Experiments conducted using the low-energized block based Gabor KDCV method, with three different similarity measures on both the FERET and FRAV2D databases are shown in Figures 5 and 6. Considering, only the low dimensional, low-energized features of GWT image, greatly improves the computing speed of nonlinear discriminant method. The results of recognition accuracy (in terms of sensitivity) versus dimensionality reduction (number of features) and the cumulative match curves using the three

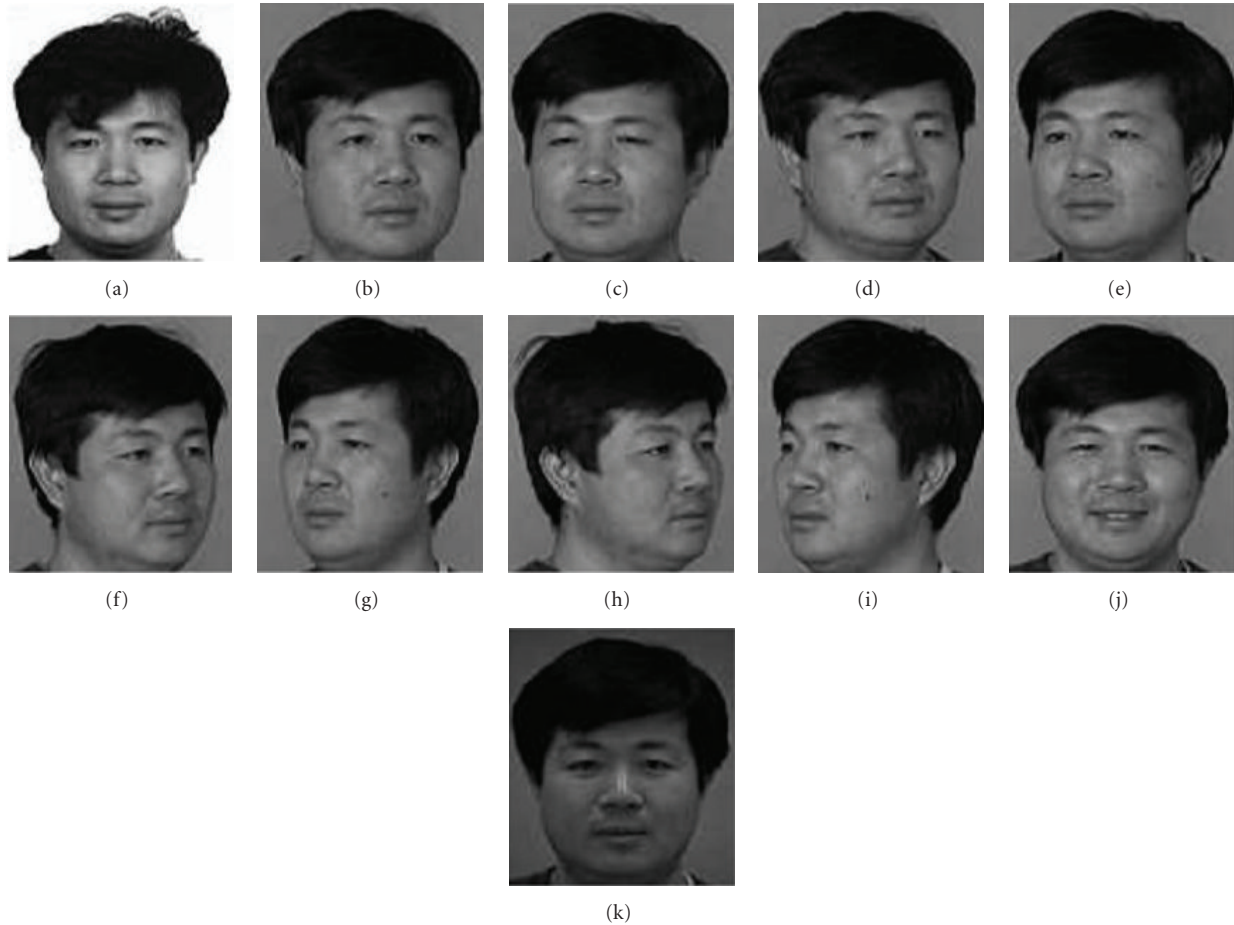


FIGURE 3: Demonstration images of one individual from the FERET database.

TABLE 6: Comparison of recognition accuracy of various methods with the proposed method.

Method	Highest recognition accuracy
GDA	78.04%
Elastic graph matching (EGM) [9]	80.00%
DCT-LDA	80.87%
GWT	81.1%
DCT-GDA	82.84%
GWT-LDA	84.1%
DCT-KDCV	85.13%
Gabor fusion KDCV	91.22%
Proposed approach	94.25%

different similarity measures are shown in Figures 5 and 6. From the results on both FERET and FRAV2D database it is seen that the cosine similarity measure performs the best, followed in order by the L_2 and the L_1 measure.

Figures 5 and 6 indicate that the proposed method performs well with a lower dimension as well. These results show that there is certain robustness to age and illumination. Our results indicate that

- (i) the low-energized block-based Gabor features with KDCV approach greatly enhance the face recognition performance as well as reduce the dimensionality of the feature space when compared with the Gabor features as shown in Tables 1 and 3. For example, the similarity measure improves the face recognition accuracy by almost 10% using only the few low-energized Gabor features with improved discriminative power compared to the original Gabor features as shown in Tables 1 and 3.
- (ii) The proposed method further enhances face recognition with the use of cosine kernel along with the cosine similarity measure.

Experimental result indicates that the use of cosine kernels in the KDCV further increases the discriminative power of the feature vector extracted from the low-energized block of GWT image and hence, is an effective feature extraction approach, performing better way to extract more effective discriminating features than the GDA. The extracted low-energized feature vector by the proposed Algorithm 1 of Section 2.2 enhances the face recognition performance in presence of occlusions. Experimentally it has been observed that this method is less time consuming than the EGM and



FIGURE 4: Demonstration images of one class from the FERET dataset.

other well-known algorithms [4, 21, 22, 24, 25]. (a) Our results show that the proposed method greatly enhances recognition performance, (b) reduces the dimensionality of the feature space, and (c) the Cosine similarity distance classifier further increases the face recognition accuracy, as it calculates the angle between two vectors and is not affected by their magnitude.

5. Conclusion

This paper introduces a novel block-based GWT generalized KDCV method using the cosine kernel for frontal and pose-angled face recognition. As cosine kernel function is used here, so there is no need of data normalization and the parameter tuning can be avoided as in the case

of (RBF) Gaussian kernels. Also the derived low dimensional low-energized features are characterized by spatial frequency, locality, and orientation selectivity to cope with the variations due to illumination and facial expression changes as a property of Gabor kernels. Such characteristics produce salient local features, such as the eyes, nose, and mouth, that are suitable for face recognition. The KDCV method extended with cosine kernel is then applied on these extracted feature vectors to finally obtain only the real nonlinear discriminating kernel feature vector with improved discriminative power, containing salient facial features that are compared locally instead of a general structure, and hence allows to make a decision from the different parts of a face and thus maximizes the benefit of applying the idea of “recognition by parts.” So the method performs well in presence of occlusions (e.g., sunglasses, scarf

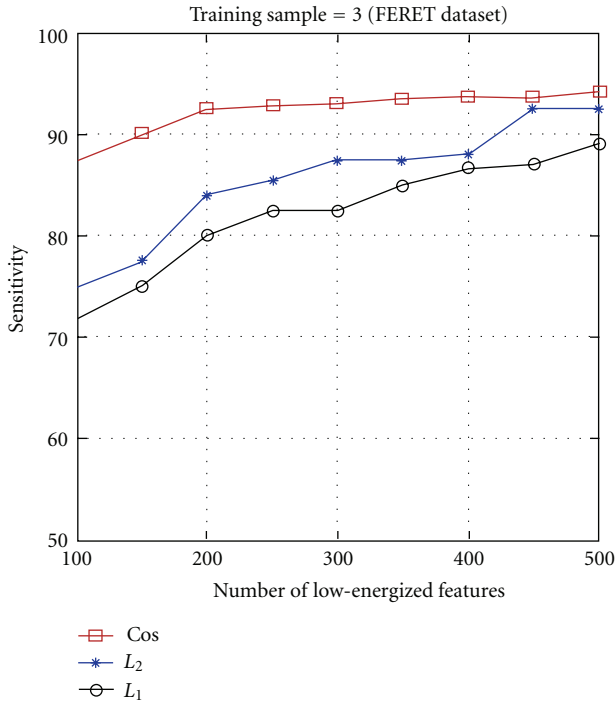


FIGURE 5: Face recognition performance of the proposed method using cosine kernel function and considering the first 3 images as training set on the FERET dataset, with the three different similarity measures: cos (cosine similarity measure), L_2 (L_2 distance measure), L_1 (L_1 distance measure).

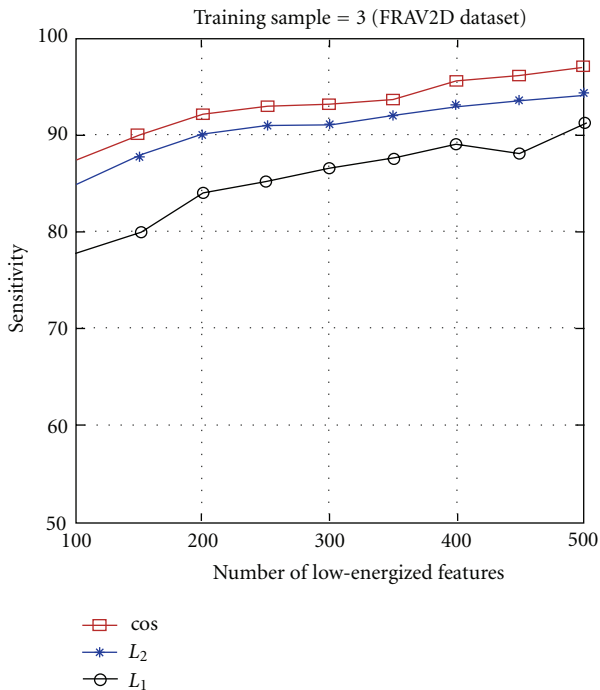


FIGURE 6: Face recognition performance of the proposed method using cosine kernel function and considering the first 3 images as training set on the FRAVD2D dataset, with the three different similarity measures: cos (cosine similarity measure), L_2 (L_2 distance measure), L_1 (L_1 distance measure).

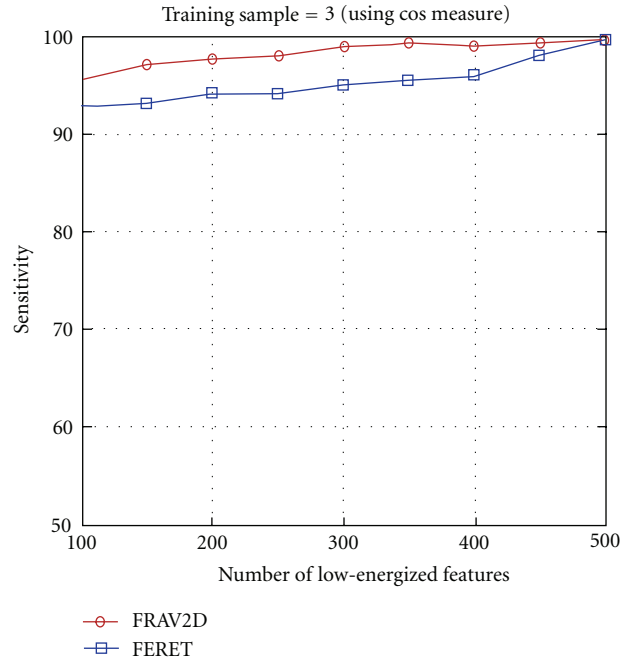


FIGURE 7: Negative recognition performance of the proposed method using cosine similarity distance measure on the FRAVD2 and FERET dataset.

etc.) that is when there are sunglasses or any other obstacles the algorithm compares face in terms of mouth, nose, and other features rather than the eyes.

Acknowledgments

The authors are thankful to a major project entitled “Design and Development of Facial Thermogram Technology for Biometric Security System,” funded by University Grants Commission (UGC), India and “DST-PURSE Programme” and CMATER and SRUVM project at Department of Computer Science and Engineering, Jadavpur University, India for providing necessary infrastructure to conduct experiments relating to this work.

References

- [1] O. Ayinde and Y. H. Yang, “Face recognition approach based on rank correlation of Gabor-filtered images,” *Pattern Recognition*, vol. 35, no. 6, pp. 1275–1289, 2002.
- [2] C. Liu and H. Wechsler, “Evolutionary pursuit and its application to face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 570–582, 2000.
- [3] C. Liu and H. Wechsler, “Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition,” *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.
- [4] X. Y. Jing and D. Zhang, “A face and palmprint recognition approach based on discriminant DCT feature extraction,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 34, no. 6, pp. 2405–2415, 2004.

- [5] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: a survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–740, 1995.
- [6] A. Kar, D. Bhattacharjee, M. Nasipuri, D. Kumar Basu, and M. Kundu, "Classification of high-energized gabor responses using bayesian PCA for human face recognition," *International Journal of Recent Trends in Engineering*, pp. 106–110, 2009.
- [7] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.
- [8] X. Y. Jing, Y. F. Yao, D. Zhang, J. Y. Yang, and M. Li, "Face and palmprint pixel level fusion and Kernel DCV-RBF classifier for small sample biometric recognition," *Pattern Recognition*, vol. 40, no. 11, pp. 3209–3224, 2007.
- [9] L. Wiskott, J. M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by elastic graph matching," in *Intelligent Biometric Techniques in Fingerprint and Face Recognition*, chapter 11, pp. 355–396, CRC Press, New York, NY, USA, 1999.
- [10] "The Face recognition technology (FERET) database," <http://www.itl.nist.gov/iad/humanid/feret>.
- [11] "Face Recognition and Artificial Vision group FRAV2D face database," <http://www.frav.es/>.
- [12] T. S. Lee, "Image representation using 2d gabor wavelets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, pp. 959–971, 1996.
- [13] D. Gabor, "Theory of Communication," *Journal of the Institution of Electrical Engineers*, vol. 93, pp. 429–459, 1949.
- [14] G. Donate, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974–989, 1999.
- [15] M. Lades, J. C. Vorbrueggen, J. Buhmann et al., "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on Computers*, vol. 42, no. 3, pp. 300–311, 1993.
- [16] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative common vectors for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 4–13, 2005.
- [17] M. B. Gülmezoğlu, V. Dzhafarov, and A. Barkana, "The common vector approach and its relation to principal component analysis," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 655–662, 2001.
- [18] B. Scholkopf and A. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*, MIT Press, Cambridge, Mass, USA, 2002.
- [19] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," 2001, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- [20] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing*, vol. 16, no. 5, pp. 295–306, 1998.
- [21] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV'05)*, pp. 786–791, October 2005.
- [22] Y. M. Lui and J. R. Beveridge, "Grassmann registration manifolds for face recognition," in *Proceedings of the European Conference on Computer Vision*, pp. 44–57, 2008.
- [23] <http://www.lifescience.com/bioinformatics/sensitivity-specificity-accuracy-and-relationship-between-them>.
- [24] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2617–2628, 2007.
- [25] S. Li, Y. F. Yao, X. Y. Jing, Z. L. Shao, D. Zhang, and J. Y. Yang, "Nonlinear DCT discriminant feature extraction with generalized KDCV for face recognition," in *Proceedings of the 2nd International Symposium on Intelligent Information Technology Application (IITA'08)*, pp. 338–341, December 2008.




Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

