

3D Face Recognition Based on Depth and Intensity Gabor Features using Symbolic PCA and AdaBoost

P. S. Hiremath and Manjunatha Hiremath

*Department of Computer Science
Gulbarga University, Gulbarga – 585106
Karnataka, India,
hiremathps53@yahoo.com, manju.gmtl@gmail.com*

Abstract

Traditional 2D face recognition based on optical (intensity or color) images faces many challenges, such as illumination, expression, and pose variation. In fact, the human face generates not only 2D texture information but also 3D shape information. In this paper, the objective is to investigate what contributions depth and intensity information make to the solution of face recognition problem when expression and pose variations are taken into account, and a novel system is proposed for combining depth and intensity information in order to improve face recognition performance. In the proposed approach, local features based on Gabor wavelets are extracted from depth and intensity images, which are obtained from 3D data after fine alignment. Then a novel hierarchical selecting scheme embedded in symbolic principal component analysis (Symbolic PCA) and AdaBoost learning is proposed to select the most effective and most robust features and to construct a strong classifier. Experiments are performed on the three datasets, namely, Texas 3D face database, Bosphorus 3D face database and CASIA 3D face database, which contain face images with complex variations, including expressions, poses and long time lapses between two scans. The experimental results demonstrate the enhanced effectiveness in the performance of the proposed method. Since most of the design processes are performed automatically, the proposed approach leads to a potential prototype design of an automatic face recognition system based on the combination of the depth and intensity information in face images.

Keywords: *3D face recognition, Radon transform, Symbolic PCA, Gabor Filter, AdaBoost*

1. Introduction

Face recognition is one of the most active research areas in the study of pattern recognition and computer vision. Over the past several decades, much work is focused on two-dimensional images. Due to the complexity of the face recognition, it is still difficult to develop a robust automatic face recognition system. The difficulties mainly include the complex variations in many aspects, such as poses, expressions, illuminations, aging and subordinates, of these problems. The pose variations and illuminations commonly influence the accuracy of 2D face recognition. According to evaluations of commercially available and mature prototyped face recognition systems provided by face recognition vendor tests (FRVT), the recognition results under the unconstrained conditions are not satisfactory.

To develop a robust face recognition system, additional information needs to be considered. Two typical solutions are the use of infrared images and the use of 3D images. Infrared images are robust to changes in environmental lighting, but these are too sensitive to changes in environmental temperature. Thus, its use is still limited. Another

solution is to utilize 3D information. With the development of 3D capturing equipments, it has become faster and easier to obtain 3D shape and 2D texture information to represent a real 3D face. Currently, many types of equipment based on active stereo vision are robust to illumination variations. Thus, the 3D shape obtained by such equipment represents the actual information irrespective of lighting. Moreover, the complete transformations between different 3D images can be computed in the image plane, which is very difficult in 2D face recognition. Recently, some research results have illustrated that 3D data have more advantages than traditional 2D data. Usage of 3D data is considered to be a promising way to improve the robustness and accuracy of recognition systems.

Actually, face recognition using 3D information can solve some problems that occur in 2D face recognition. Due to some of the difficulties encountered in 3D face recognition, such as coping with expression variations, the inconvenience of information capture and large computational costs, these problems have been the focus of recent research [10,11].

Face recognition based on 3D information is not a new topic. Studies have been conducted since the end of the last century [1-7]. The representative methods use features extracted from the facial surface to characterize an individual. Lin *et al.*, [8] extracted semi local summation invariant features in a rectangular region surrounding the nose of a 3D facial depth map. Then the similarity between them was computed to determine whether they belonged to the same person. Al-osaimi *et al.*, [9] integrated local and global geometrical cues in a single compact representation for 3D face recognition. In another method, the human face is considered as a 3D surface, and the global difference between two surfaces provides the distinguishability between faces. Brumier *et al.*, [31] constructed some central and lateral profiles to represent an individual and proposed two methods of surface matching and central/lateral profiles to compare two instances. Medioni *et al.*, [32] built a complete and automatic system to perform face authentication by modeling 3D faces using stereo vision and analyzing the distance map between gallery and probe models. Lu *et al.*, [12] used the hybrid iterative closest point (ICP) algorithm to align the reference model with the scanned data and adopted registration errors to distinguish between different faces. Chang *et al.*, [8] divided the whole facial surface into sub regions. The rigid regions around the nose area were matched and combined to perform the recognition. Following the idea of dividing the sub-regions, Faltermoier *et al.*, [6] introduced a system for 3D face recognition based on the fusion of results from a group of regions that had been independently matched. Russ *et al.*, [7] presented an approach enabling a good alignment of 3D face point clouds while preserving face size information. All of these studies illustrated the feasibility of 3D face recognition. However, perhaps due to illumination of data, the above schemes only use the shape features of facial surfaces while ignoring the texture information. Wang *et al.*, [3] described facial feature points by using Gabor filter responses in a 2D domain and point signatures in a 3D domain. Chang *et al.*, [5] evaluated the recognition scheme with different combinations of 2D and 3D information, which showed that the combination of 2D and 3D information was most effective in characterizing an individual face. Cook *et al.*, [34] proposed a new method combining intensity and range images that was insensitive to expression variation based on log gabor templates. Maurer *et al.*, [33] analyzed the performance of the Geometrix Active ID Biometric Identity system, which fused shape and texture information.

The face recognition based on combination of 3D shape information and 2D intensity/color information is a novel approach, which provides an opportunity to improve face recognition performance. In the present paper, the objective is to investigate how depth and intensity information contributes to face recognition when expression and pose variations are taken into account. Further, a robust and accurate face system is designed by selecting and fusing the most effective depth and intensity features. All the processes included in the training and test phases of the proposed approach are fully automated.

2. Materials and Methods

For the purpose of experimentation of the proposed methodology, the face images drawn from the following 3D face databases are considered: (i) Texas 3D face database, (ii) Bosphorus 3D face database, and (iii) CASIA 3D face database.

2.1. Texas 3D Face Database

The Texas 3D Face Recognition (Texas 3DFR) database is a collection of 1149 pairs of facial color and range images of 105 adult human subjects. These images were acquired using a stereo imaging system manufactured by 3Q Technologies (Atlanta, GA) at a very high spatial resolution of 0.32 mm along the x, y, and z dimensions. During each acquisition, the color and range images were captured simultaneously and thus the two are perfectly registered to each other. This large database of two 2D and 3D facial models was acquired at the company Advanced Digital Imaging Research (ADIR), LLC (Friendswood, TX), formerly a subsidiary of Iris International, Inc. (Chatsworth, CA), with assistance from research students and faculty from the Laboratory for Image and Video Engineering (LIVE) at The University of Texas at Austin. This project was sponsored by the Advanced Technology Program of the National Institute of Standards and Technology (NIST).

Texas 3DFRD was created to develop and test 3D face recognition algorithms intended to operate in environments with co-operative subjects, wherein, the faces are imaged in a relatively fixed position and distance from the camera [22-24].

2.2. Bosphorus 3D Face Database

The Bosphorus 3D face database consists of 105 subjects in various poses, expressions and occlusion conditions. The 18 subjects have beard/moustache and the 15 subjects have hair. The majority of the subjects are aged between 25 and 35. There are 60 men and 45 women in total, and most of the subjects are Caucasian. Two types of expressions have been considered in the Bosphorus database. In the first set, the expressions are based on action units. In the second set, facial expressions corresponding to certain emotional expressions are collected. These are: happiness, surprise, fear, sadness, anger and disgust.

The facial data are acquired using Inspeck Mega Capturor II 3D, which is a commercial structured-light based 3D digitizer device. The sensor resolution in x, y & z (depth) dimensions are 0.3mm, 0.3mm and 0.4mm respectively, and colour texture images are high resolution (1600x1200 pixels). It is able to capture a face in less than a second. Subjects were made to sit at a distance of about 1.5 meters away from the 3D digitizer. A 1000W halogen lamp was used in a dark room to obtain homogeneous lighting. However, due to the strong lighting of this lamp and the device's projector, usually specular reflections occur on the face. This does not only affect the texture image of the face but can also cause noise in the 3D data. To prevent it, a special powder which does not change the skin colour is applied to the subject's face. Moreover, during acquisition, each subject wore a band to keep his/her hair above the forehead to prevent hair occlusion, and also to simplify the face segmentation task. The proprietary software of the scanner is used for acquisition and 3D model reconstruction [26].

2.3 CASIA 3D Face Database

CASIA 3D Face Database consists of 4624 scans of 123 persons using the non-contact 3D digitizer, Minolta Vivid 910. During building the database, not only the single variations of poses, but also expressions and illuminations are considered [25].

3. Proposed Methodology

The proposed methodology employs the following: (i) Radon transform, (ii) Gabor wavelets, (iii) Symbolic PCA, and (iv) Adaptive boosting algorithm (AdaBoost) classifier which are described in the following sections.

3.1. Radon Transform

The Radon Transform (RT) is a fundamental tool in many areas. The 3D radon Transform is defined using 1D projections of a 3D object $f(x,y,z)$ where these projections are obtained by integrating $f(x,y,z)$ on a plane, whose orientation can be described by a unit vector $\vec{\alpha}$. Geometrically, the continuous 3D Radon transform maps a function \square^3 into the set of its plane integrals in \square^3 . Given a 3D function $f(\vec{x}) \square f(x,y,z)$ and a plane whose representation is given using the normal $\vec{\alpha}$ and the distance s of the plane from the origin, the 3D continuous Radon Transform of f for this plane is defined by

$$\begin{aligned} \mathfrak{R}f(\vec{\alpha}, s) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\vec{x}) \delta(\vec{x}^T \alpha - s) d\vec{x} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y,z) \delta(x \sin \theta \cos \phi + y \sin \theta \sin \phi + z \cos \theta - s) dx dy dz \end{aligned}$$

where $\vec{x} = [x, y, z]^T$, $\vec{\alpha} = [\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]^T$, and δ is Dirac's delta function defined by $\delta(x) = 0, x \neq 0, \int_{-\infty}^{\infty} \delta(x) dx = 1$. The Radon transform maps the spatial domain (x,y,z) to the domain $(\vec{\alpha}, s)$. The 3D continuous Radon Transform satisfies the 3D Fourier slice theorem.

3.2. 2D Gabor Filter

The 2D Gabor filters of depth and intensity images of faces are used to characterize an individuals' face. The Gabor wavelets represent the properties of spatial localization, orientation selectivity, spatial frequency selectivity and quadrature phase relationships, and these have been experimentally verified to be a good approximation to the response of cortical neurons. The Gabor wavelet based representation of faces has been successfully tested in 2D face recognition. Such representation of an image describes the facial characteristics of both the spatial frequency and spatial relations. The 2D Gabor wavelets are defined as follows:

$$\psi(z) = \frac{k_{\mu,v}^2}{\sigma^2} \exp\left(-\frac{k_{\mu,v}^2 z^2}{2\sigma^2}\right) \left[\exp(ik_{\mu,v}^2 z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (1)$$

where $z = (x, y)$, and μ and ν define the orientation and scale of the Gabor wavelets, respectively. The wave vector $k_{\mu,v}$ is defined as follows:

$$k_{\mu,v} = k_v e^{i\Phi_\mu} \quad (2)$$

where $k_v = k_{\max} / f^\nu$ and $\Phi_\mu = \pi\mu / 8$. The constant k_{\max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain. The Gabor kernels in the above equation are self-similar, since they can be generated from the mother

wavelet by scaling and rotation via the wave vector $k_{\mu,v}$. More scales or rotations can increase the dependencies of neighbor samples.

In the proposed method, Gabor kernels with five different scales $v \in \{0, \dots, 4\}$ and eight orientations $\mu \in \{0, \dots, 7\}$ are used, with the parameters $\sigma = 2\pi$, $k_{\max} = \pi/2$ and $f = \sqrt{2}$. The number of scales and orientations are selected to represent the facial characteristics in terms of spatial locality and orientation selectivity. The Gabor representation of an image, called the Gabor image, is the convolution of the image with Gabor kernels as defined by Eq(1). For each image pixel, it has two Gabor parts: the real part and imaginary part, which are transformed into two kinds of features: Gabor magnitude features and Gabor phase features. Herein, it is proposed to use the Gabor magnitude features to represent the facial features, since the Gabor transformation strongly responds to edges [19-21].

The depth Gabor images are smoother in comparison with the intensity Gabor images due to the fact that the value of the pixels in the depth image changes less than does the value in the intensity images. The smoother depth Gabor image can reduce the influence of noise, but it cannot describe the facial features in detail. This is why the face recognition is performed by combining depth and intensity information [14-16].

3.3. Symbolic Principal Component Analysis (PCA)

Consider the 3D face images $\Gamma_1, \Gamma_2, \dots, \Gamma_n$, each of size $N \times M$, from a 3D face image database. Let $\Omega = \{\Gamma_1, \Gamma_2, \dots, \Gamma_n\}$ be the collection of n face images of the database, which are first order objects. Each object $\Gamma_l \in \Omega$, $l = 1, 2, \dots, n$, is described by a feature vector $(\tilde{Y}_1, \dots, \tilde{Y}_p)$, of length $p = NM$, where each component \tilde{Y}_j , $j = 1, 2, \dots, p$, is a single valued variable representing the values of the 3D face image Γ_l . An image set is a collection of 3D face images of m different subjects; each subject has same number of images but with different facial expressions and illuminations. There are m number of second order objects (face classes) denoted by c_1, c_2, \dots, c_m , each consisting of different individual images $\Gamma_l \in \Omega$. Let $E = \{c_1, c_2, \dots, c_m\}$ and $c_i \subseteq \Omega$, $i = 1, 2, \dots, m$. The feature vector of each face class $c_i \in E$ is described by a vector of p interval variables Y_1, Y_2, \dots, Y_p , and is of length $p = NM$. The interval variable Y_j of face class c_i is declared by $Y_j(c_i) = [\underline{x}_{ij}, \bar{x}_{ij}]$, where \underline{x}_{ij} and \bar{x}_{ij} are minimum and maximum intensity values, respectively, among j^{th} values of all the images of face class c_i . This interval incorporates information of the variability of j^{th} feature inside the i^{th} face class. Let $X(c_i) = (Y_1(c_i), \dots, Y_p(c_i))$. The vector $X(c_i)$ of symbolic variables is recorded for each $c_i \in E$, and can be described by a symbolic data vector which is called as symbolic face : $X(c_i) = (a_{i1}, a_{i2}, \dots, a_{ip})$, where $a_{ij} = Y_j(c_i)$, $j = 1, 2, \dots, p$ [17-18]. Let the m symbolic faces be represented by a $m \times p$ matrix :

$$\underline{X} = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \cdot & \cdot & \cdot \\ a_{m1} & \dots & a_{mp} \end{pmatrix} = (a_{ij})_{m \times p} \quad (3)$$

The symbolic PCA takes as input the matrix \underline{X} containing m symbolic faces pertaining to the given set Ω of images[27]. It is proposed to use centers method, which

essentially applies the conventional PCA method to the centers $x_{ij}^c \in \mathfrak{R}$ of the interval $[\underline{x}_{ij}, \bar{x}_{ij}]$, that is,

$$x_{ij}^c = \frac{x_{ij} + \bar{x}_{ij}}{2} \quad (4)$$

where $j = 1, \dots, p$ and $i = 1, \dots, m$. The $m \times p$ data matrix \underline{X}^c , which contains the centers x_{ij}^c of the intervals α_{ij} for m symbolic faces, is given by:

$$\underline{X}^c = \begin{pmatrix} x_{11}^c & \dots & x_{1p}^c \\ \cdot & \cdot & \cdot \\ x_{m1}^c & \dots & x_{mp}^c \end{pmatrix} \quad (5)$$

The mean vector ψ of \underline{X}^c is defined by $\psi = [\psi_j]$, where $\psi_j = \frac{1}{m} \sum_{i=1}^m x_{ij}^c$, $j = 1, 2, \dots, p$. Each row vector of \underline{X}^c differs from the mean vector ψ by the vector $\Phi_i = (x_{i1}^c, x_{i2}^c, \dots, x_{ip}^c) - \psi$. We define the matrix Φ as $\Phi = [\Phi_1', \Phi_2', \dots, \Phi_m']$. The covariance matrix C is obtained as $C = \Phi' \Phi$. Then, the eigenvalues are calculated and arranged in the order of decreasing magnitude: $\lambda_1 \geq \lambda_2 \geq \dots, \lambda_m \geq 0$. Further, and the corresponding orthonormalized eigenvectors $y_1, y_2, \dots, y_m \in \mathfrak{R}^m$ of the covariance matrix C are obtained. The eigenvectors of the symbolic PCA can be obtained as $V_m = \Phi Y_m$, where $Y_m = (y_1, \dots, y_m)$ is the $m \times m$ matrix with columns y_1, y_2, \dots, y_m and V_m is the $p \times m$ matrix with corresponding eigenvectors v_1, v_2, \dots, v_m , as its columns. The subspace is extracted from the $p \times m$ dimensional space by selecting s number of eigenvectors, which contain maximum variance and are denoted by v_1, v_2, \dots, v_s , corresponding to eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots, \lambda_s$. The weights W_{ik} for i^{th} symbolic face, $i = 1, 2, \dots, m$, are computed as

$$W_{ik} = V_k^T (x_{ij}^c - \psi) \quad (6)$$

where $k = 1, 2, \dots, s$. The weights of i^{th} symbolic face form the feature vector $(w_{i1}, w_{i2}, \dots, w_{is})$ of the i^{th} symbolic face [28,29]. The weights of test image I^{test} are computed by projecting the test image into face subspace of s dimension as :

$$w_k^{test} = v_k^T (I^{test} - \Psi), \quad k = 1, 2, \dots, s. \quad (7)$$

3.4. AdaBoost Classifier

AdaBoost (Adaptive Boosting) is an ensemble learning algorithm that can be used for classification or regression. Although AdaBoost is more resistant to over fitting than many machine learning algorithms, it is often sensitive to noisy data and outliers. AdaBoost is called adaptive, because it uses multiple iterations to generate a single strong learner. AdaBoost creates the strong learner by iteratively adding weak learners. During each round of training, a new weak learner is added to the ensemble and a weighting vector is adjusted to focus on examples that were misclassified in previous rounds.

AdaBoost learning essentially works for a two class classification problem. While face recognition is a multiclass problem, we convert it into one of the two classes using the representation of intra-personal Vs extra-personal classes. Intra-personal examples are obtained by using differences in images of the same person, whereas extra-personal examples are obtained by using differences in images of different people. From the depth Gabor images and intensity Gabor images under each scale and orientation, the effective features in these scale and orientation are selected. After effective depth and intensity Gabor features selected, the cascaded AdaBoost learning procedure is used to create a strong classifier with a cascading structure [30].

3.5. Proposed Method

The proposed methodology comprises the following steps:

- (i) Radon transform is applied to the input depth and intensity images of a 3D face, which yields binary images that are used to crop the facial areas in the corresponding images.
- (ii) Gabor filter are applied to the cropped facial images, which yield Gabor magnitude features as facial feature vectors.
- (iii) Symbolic PCA is applied to the facial Gabor features in order to achieve dimensionality reduction and obtain subsampled feature vectors
- (iv) Lastly, AdaBoost is used to perform face recognition based on subsampled feature vectors.

The Figure 1 shows the overview of the proposed framework.

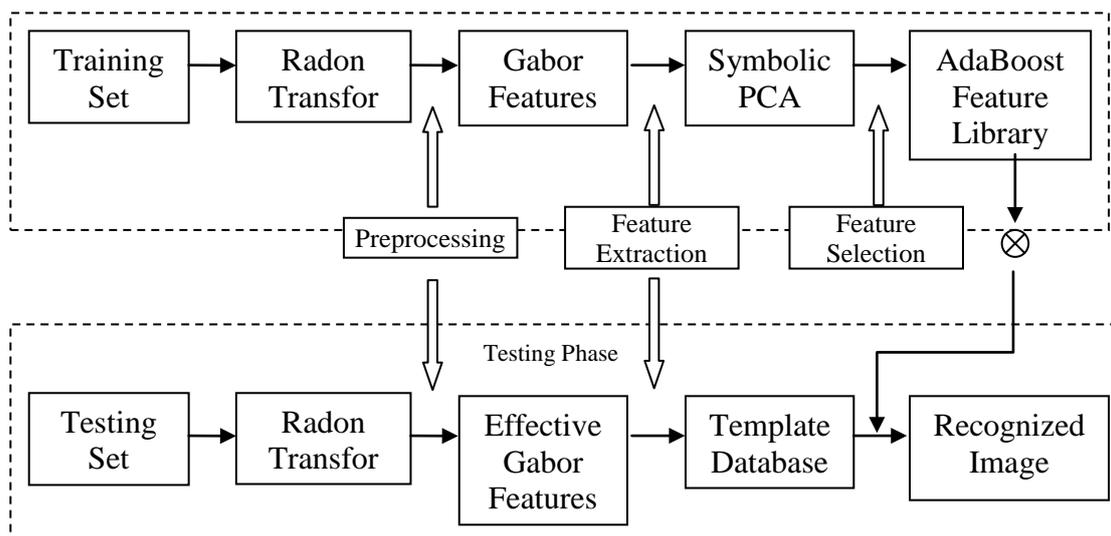


Figure 1. Overview of Proposed Framework

The algorithms of the training phase and the testing phase of the proposed method are given below:

Algorithm 1: Training Phase

1. Input the depth and intensity images of a face from the training set containing M depth images and M intensity images corresponding to M faces.
2. Apply Radon transform, from 0^0 to 180^0 orientations (in steps of h), to the input images yielding the corresponding binary images.

3. Superpose the binary images obtained in the Step 2 on the corresponding input images to obtain the cropped facial images.
4. Repeat the Steps 1 to 3 for all the M facial images in the training set and build the set M_1 of cropped facial images.
5. Apply Gabor wavelet filter on M_1 set to extract Gabor magnitude features.
6. Apply Symbolic PCA to subsample the Gabor feature set to reduce the dimensionality of the feature set. Let P be the reduced dimensionality of feature set.
7. Compute the weights w_1, w_2, \dots, w_p for each training image as its facial features and store these values in the Symbolic PCA feature library of the face database.
8. After effective Symbolic PCA features are selected, the cascaded AdaBoost learning procedure is used to create a strong classifier.

Algorithm 2: Testing Phase

1. Input the depth and intensity images from the probe set containing P images.
2. Apply Radon transform, from 0^0 to 180^0 orientations (in steps of h), to the input images yielding the binary images.
3. Superimpose the binary images on the corresponding input depth and intensity images to obtain the cropped facial images.
4. Apply Gabor wavelet filter on cropped facial images to extract effective Gabor features.
5. Compute the symbolic weights $w_i^{test}, i = 1, 2, \dots, p$, for the extracted Gabor magnitude features by projecting the test image on the Symbolic PCA feature subspace of dimension p .
6. Apply AdaBoost procedure and select the most effective features from redundant feature subspace. Compute the difference with each gallery example and form a difference vector v .
7. Apply Nearest Neighbor scheme to decide which class the test sample belongs to and output the texture face image corresponding to the recognized facial image of the training set.

4. Experimental Results and Discussion

As in typical biometric systems, the proposed method includes two phases: the training phase and the testing phase as illustrated in the Figure 1. The proposed method is implemented using Intel Core 2 Quad processor @ 2.66 GHz machine and MATLAB 7.9. The 4059 images of three databases, namely, Bhosphorus 3D face database, CASIA 3D face database and Texas 3D face database, that are divided into three subsets, that is the training set, the gallery set and probe set. The training set contains 3200 images, corresponding to the 123 subjects with 33 images for each subject. The gallery set contains 100 images from the first images of other subjects. The other 759 images are randomly chosen as probe set from all the three databases.

The training process is performed once on a predefined training set to learn an effective classifier. It includes three important procedures: preprocessing, feature extraction and feature selection. By translating and rotating one input 3D image to align one reference 3D image, face poses and changed positions between the face and the equipment are normalized. This step is fully automatic. According to the aligned images, the normalized depth images and intensity images are obtained. Robust feature representation is very important to the whole system. It is expected that these features are invariant to rotation, scale and illumination. The proposed method uses raw depth and intensity features to describe the individual information using Gabor filters with multiple scales and multiple

orientations. As it is commonly a problem, using multiple channels and multiple types of local features results in a much higher dimensional feature space. A large number of local features can be produced with varying parameters in the position, scale and orientation of the filters. Some of these features are effective and important for the classification task, whereas others may not be so. The proposed framework combines Symbolic PCA and AdaBoost learning to fuse 2D and 3D Gabor features at the “feature” level.

During the testing phase, an input is classified according to the learned classifier in the training phase. By preprocessing, normalized depth image and intensity images are obtained. This procedure is the same as that in the training phase. The feature vector of one probe sample is generated by extracting the corresponding features as in the final cascaded classifier, and its difference with each gallery example forms the difference vector, x . For each vector x , the i^{th} layer of the cascaded classifier returns the similarity measure, S_i . The larger this similarity value, the more this sample belongs to the intra-personal space. If $S_i < 0$, the i^{th} layer rejects the sample. Using the similarities from the multiple layers, we can obtain its total similarity:

$$S = \sum_{i=1}^L S_i$$

where L is the number of layers and S_i is the similarity value from the i^{th} layer. Thus, we can obtain the sample's similarity with each gallery example. Then, the nearest neighbor scheme is used to decide which class the test sample belongs to.

The sample training depth images which are used for the experimentation are shown in the Figure 2, and their corresponding texture images are shown in the Figure 3. The different pose variations in the database are shown in the Figure 4.

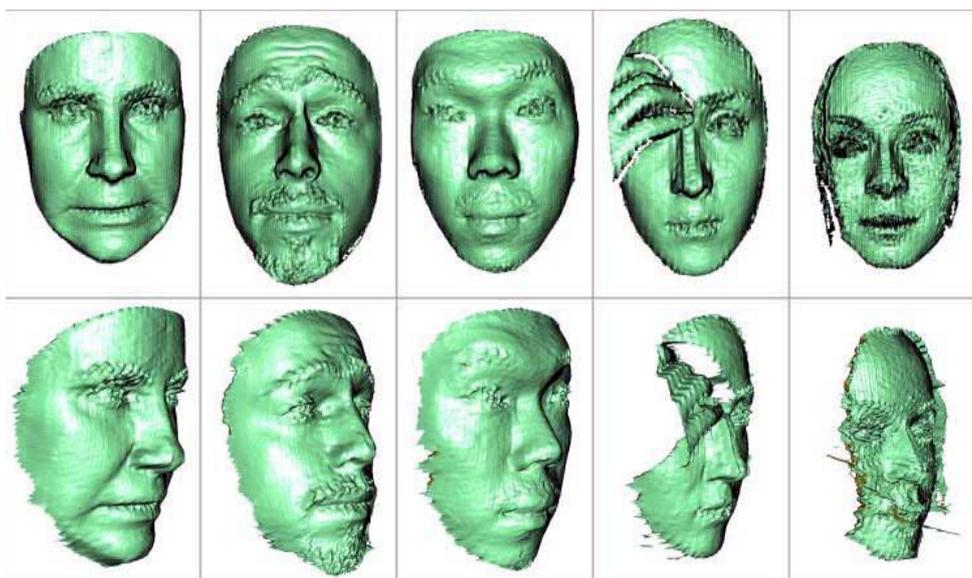


Figure 2. Sample Training Depth Images

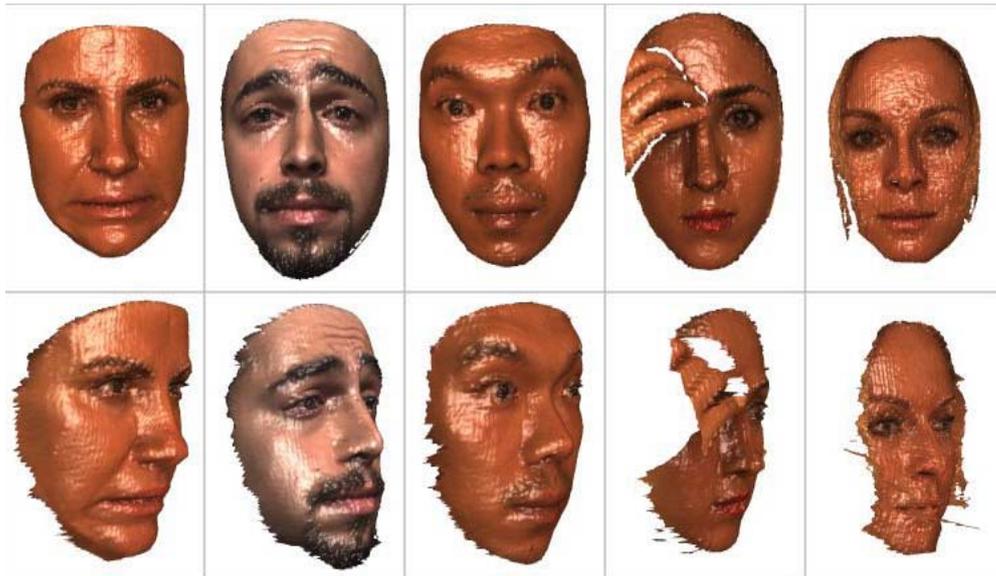


Figure 3. Sample Texture Images of the Figure 2

To demonstrate the performance of the proposed algorithm, comparison of experimental results that have been reported in recent literature, has been done. The Table 1 shows the performance comparison of the proposed method with other methods.

Table 1. The Performance Comparison of Proposed Method with other Methods

Method	Verification rate
Kakadiaris et al. [13]	97.00%
Faltemier et al. [6]	94.90%
Maurer et al. [20]	95.80%
H. usken et al. [19]	97.30%
Mian et al.[35]	99.30%
Proposed Method	99.50%

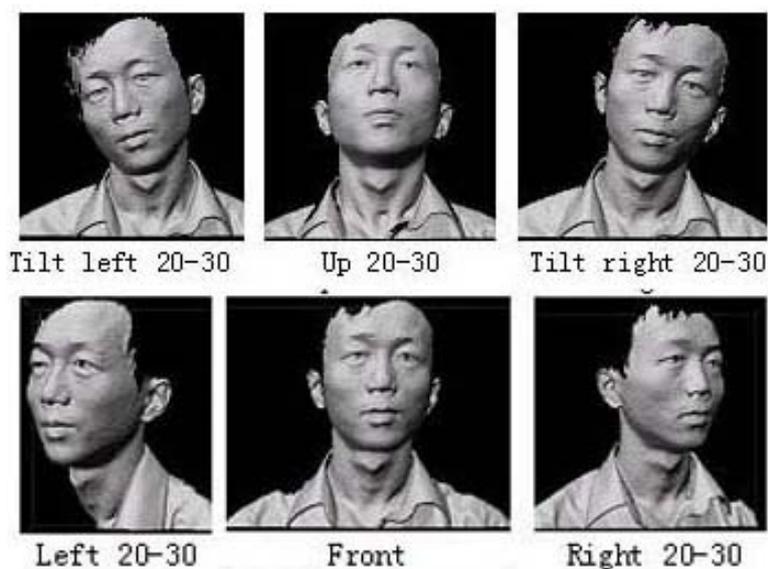


Figure 4. Pose Variations in 3D Face Databases

5. Conclusion

Intensity information is more pertinent than depth information in the context of expression variations; depth information is more pertinent than intensity information in the context of pose variations. Thus, their combination is helpful in improving recognition performance. In this paper, a new scheme is proposed to combine depth and intensity features in order to overcome problems due to expression and pose variations, and thus build a robust and automatic face recognition system. The Gabor features are used, in which the dimensionality of the Gabor features is extremely large, since multiple scales and orientations are adopted. To reduce the large dimensions of Gabor features, we propose a hierarchical selection scheme for selecting effective features by using Symbolic PCA and AdaBoost procedures and the so constructed classifier. This is one attempt to apply statistical learning to fuse 2D and 3D face recognition at the “feature” level. By analyzing our experimental results and comparing those of the existing methods, we demonstrate the promising performance of the proposed scheme. The recognition accuracy can be further improved by considering a larger training set and a better classifier.

Acknowledgement

The authors are grateful to the referees for helpful comments and suggestions. The authors are indebted to the University Grants Commission, New Delhi, for the financial support for this research work under UGC-MRP F.No.39-124/2010 (SR).

References

- [1] W. Zhao, R. Chellappa, P. J. Phillips and A. Rosenfeld, “Face recognition: a literature survey”, *ACM Computing Surveys (CSUR)*, Archive, vol. 35, no. 4, (2003), pp. 399-458.
- [2] K. Bowyer, K. Chang and P. Flynn, “A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition”, *Computer Vision and Image Understanding*, vol. 101, no. 1, (2006), pp. 1-15.
- [3] Y. Wang, C. Chua and Y. Ho, “Facial feature detection and face recognition from 2D and 3D images”, *Pattern Recognition Letters*, vol. 23, (2002), pp. 1191-1202.
- [4] K. I. Chang, K. W. Bowyer and P. J. Flynn, “An evaluation of multi-model 2D+3D biometrics”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, (2005), pp. 619-624.
- [5] K. I. Chang, K. W. Bowyer and P. J. Flynn, “Multiple nose region matching for 3D face recognition under varying facial expression”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, (2006), pp. 1695-1700.
- [6] T. C. Faltemier, K. W. Bowyer and P. J. Flynn, “A region ensemble for 3-D face recognition”, *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, (2008), pp. 62-73.
- [7] T. Russ, C. Boehnen and T. Peters, “3D face recognition using 3D alignment for PCA”, *Proceedings of the CVPR'06*, (2006), pp. 1391-1398.
- [8] W. Lin, K. Wong, N. Boston and Y. Hu, “Fusion of summation invariants in 3D human face recognition”, *Proceedings of the CVPR'06*, (2006), pp. 1369-1376.
- [9] F. R. Al-Osaimi, M. Bennamouna and A. Miana, “Integration of local and global geometrical cues for 3D face recognition”, *Pattern Recognition*, vol. 41, no. 3, (2008), pp. 1030-1040.
- [10] P. J. Phillips, P. J. Flynn, T. Scruggs and K. W. Bowyer, “Overview of the face recognition grand challenge”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2005), pp. 947-954.
- [11] F. Tsalakanidou, S. Malassiotis and M. G. Strintzis, “Face localization and authentication using color and depth images”, *IEEE Transactions on Image Processing*, vol. 14, no. 2, (2005), pp. 152-168.
- [12] X. Lu, A. K. Jain and D. Colbry, “Matching 2.5D face scans to 3D models”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, (2006), pp. 31-43.
- [13] I. A. Kakadiaris, G. Passalis, G. Toderici, N. Murtuza, Y. Lu, N. Karampatziakis and T. Theoharis, “3D face recognition in the presence of facial expressions: an annotated deformable model approach”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, (2007), pp. 640-649.
- [14] T. S. Lee, “Image representation using 2D Gabor wavelets”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 10, (1996), pp. 959-971.
- [15] C. J. 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, (2002), pp. 467-476.

- [16] D. H. Liu, K. M. Lam and L. S. Shen, "Optimal sampling of Gabor features for face recognition", *Pattern Recognition Letters*, vol. 25, (2004), pp. 267-276.
- [17] M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, (1991), pp. 71-86.
- [18] P. N. Belhumur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, (1997), pp. 711-0720.
- [19] M. H. usken, M. Brauckmann, S. Gehlen and C. Malsburg, "Strategies and benefits of fusion of 2D and 3D face recognition", *IEEE Workshop on Face Recognition Grand Challenge Experiments*, (2005), pp. 174-181.
- [20] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West and G. Medioni, "Performance of geometrixActiveID TM 3D face recognition engine on the FRGC data", *IEEE Workshop on Face Recognition Grand Challenge Experiments*, (2005), pp. 154-160.
- [21] J. Cook, C. Mccool, V. Chandran and S. Sridharan, "Combined 2D/3D face recognition using log-Gabor templates", *IEEE Conference on Video and Signal Based Surveillance*, (2006), pp. 83-90.
- [22] S. Gupta, M. K. Markey and A. C. Bovik, "Anthropometric 3D Face Recognition", *International Journal of Computer Vision*, vol. 90, no. 3, (2010), pp. 331-349.
- [23] S. Gupta, K. R. Castleman, M. K. Markey and A. C. Bovik, "Texas 3D Face Recognition Database", *IEEE Southwest Symposium on Image Analysis and Interpretation*, Austin, TX, (2010) May, pp. 97-100.
- [24] S. Gupta, K. R. Castleman, M. K. Markey and A. C. Bovik, "Texas 3D Face RecognitionDatabase", URL:<http://live.ece.utexas.edu/research/texas3dfir/index.htm>.
- [25] C. Xu, Y. Wang, T. Tan and L. Quan, "Automatic 3D Face Recognition Combining Global Geometric Features with Local Shape Variation Information", *Proc. The 6th IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, (CASIA 3D Face Database), (2004), pp. 308-313.
- [26] N. Alyüz, B. Gökberk, H. Dibeklioglu, A. Savran, A. A. Salah, L. Akarun and B. Sankur, "3D Face Recognition Benchmarks on the Bosphorus Database with Focus on Facial Expressions", *The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008)*, Roskilde University, Denmark, (2008) May.
- [27] H. H. Bock and E. Diday, "Analysis of Symbolic Data", Springer Verlag, (2000).
- [28] C. N. Lauro and F. Palumbo, "Principal Component Analysis of Interval Data: a Symbolic Data Analysis Approach", *Computational Statistics*, vol. 15, no. 1, (2000), pp. 73-87.
- [29] P. S. Hiremath and C. J. Prabhakar, "Face Recognition Technique Using Symbolic PCA Method", *PRMI, LNCS 3776*, Springer-Verlang Berlin Hiedelberg, (2005), pp. 266-271.
- [30] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", *Journal of Computer and System Sciences*, vol. 55, (1997), pp. 119-139.
- [31] C. Beumier and M. Acheroy, "Automatic 3D face authentication", *Image and Vision Computing*, vol. 18, no. 4, (2000), pp. 315-321.
- [32] G. Medioni and R. Waupotitsch, "Face modeling and recognition in 3-D", *Proceedings of the AMFG'03*, (2003), pp. 232-233.
- [33] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West and G. Medioni, "Performance of geometrixActiveID TM 3D face recognition engine on the FRGC data", *IEEE Workshop on Face Recognition Grand Challenge Experiments*, (2005), pp. 154-160.
- [34] J. Cook, C. Mccool, V. Chandran and S. Sridharan, "Combined 2D/3D face recognition using log-Gabor templates", *IEEE Conference on Video and Signal Based Surveillance*, (2006), pp. 83-90.
- [35] A. S. Mian, M. Bennamoun and R. Owens, "An efficient multimodal 2D-3D hybrid approach to automatic face recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, (2007), pp. 1927-1943.