Robust Face Recognition Technique Based on Granular Computation and Hybrid WLD Spatial Feature Extraction

P.Raghu[1],D.Kavi Bharathy[2],M.Suresh Anand[3],J.Jegan Amarnath[4],J.K.Periasamy[5],A.Kumaresan[6]

[1]Department Of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India.
raghu.chanti2050@gmail.com
[2] Department Of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India.
kavi.bharathy2718@gmail.com
[3] Assistant Professor, Department Of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India. *suresh.anandm@gmail.com*
[4] Associate Professor, Department Of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India. *jeganamarnath@gmail.com*
[5] Associate Professor, Department Of Computer Science and Engineering, Sri Sai Ram Engineering College, Chennai, India. *jkperiasamy@gmail.com*
[6]Professor, Department Of Computer Science and Engineering, SKP Engineering College, Tiruvannamalai, India.
kummaresan@gmail.com

***Abstract*—The paper presents robust face recognition based on granular computation and hybrid spatial features extraction. The face biometric based person identification plays a major role in wide range of applications such as surveillance and online image search. The first stage of recognition starts with face detection module will be used to obtain face images, which have normalized intensity, are uniform in size and shape and depict only the face region. Here granular computing and spatial features will be presented to match face images in various illumination changes. The Gaussian operator generates a sequence of low pass filtered images by iteratively convolving each of the constituent images with a 2-D Gaussian kernel. Then, DOG pyramid will be formed from successive iterations of Gaussian images. By this granulation, facial features are segregated at different resolutions to provide edge information, noise, smoothness, and blurriness present in a face image. In features extraction stage, WLD descriptor represents an image as a histogram of differential excitations and gradient orientations, and has several interesting properties like robustness to noise and illumination changes, elegant detection of edges and powerful image representation. The Gabor filter bank is then used to extract the features from face regions to discriminate the illumination changes. These combined features are useful to distinguish the maximum number of samples accurately and it is matched with already stored original face samples for identification. The simulated results will be shown that used granulation and hybrid spatial features descriptors has better discriminatory power and recognition accuracy in the process of recognizing different facial appearance.**

***Keywords*— face recognition, granular, weber local descriptor, spatial feature**

1. INTRODUCTION

Biometric-based technologies like finger geometry, hand geometry, fingerprints require some voluntary action by the user. However face recognition can be done without any participation on part of the user, since face images can be captured from distance by camera. Based on the way in which the features are extracted, any algorithms are distinguished.

Image normalization refers to pose and illumination changes, whose purpose is to achieve invariance to data capture conditions and to allow biometrics to operate in uncontrolled settings. Multi-view face recognition ,a generative methods, that requires gallery images for each pose and thus cannot handle faces acquired from a quite novel viewpoint, one of the tasks FACE has to handle.

Recognizing object classes in real-world images is a long standing goal in Computer vision. Conceptually, this is challenging due to large appearance variations of object instances belonging to the same class. Additionally, distortions from background clutter, scale, and viewpoint variations can render appearances of even the same object instance to be vastly different. Further challenges arise from interclass similarity in which instances from different classes can appear very similar. Consequently, models for object classes must be flexible enough to accommodate class variability, yet discriminative enough to sieve out true object instances in cluttered images. These seemingly paradoxical requirements of an object class model make recognition difficult. This paper addresses two goals of recognition are image classification and object detection. The task of image classification is to determine if an object class is present in an image, while object detection localizes all instances of that class from an image. Toward these goals, the main contribution in this paper is an approach for object class recognition that employs edge information only. The novelty of our approach is that we represent contours by very simple and generic shape primitives of line segments and ellipses, coupled with a flexible method to learn discriminative primitive combinations. These primitives are complementary in nature, where line segment models straight contour and ellipse models curved contour. We choose an ellipse as it is one of the simplest circular shapes, yet is sufficiently flexible to model curved shapes. These shape primitives possess several attractive properties. First, unlike edge-based descriptors they support abstract and perceptually meaningful reasoning like parallelism and adjacency. Also, unlike contour fragment features, storage demands by these primitives are independent of object size and are efficiently represented with four parameters for a line and five parameters for an ellipse.

In recent studies it is shown that the generic nature of line segments and ellipses affords them an innate ability to represent complex shapes and structures. While individually less distinctive, by combining a number of these primitives, we empower a combination to be sufficiently discriminative. Here, each combination is a two-layer abstraction of primitives: pairs of primitives (termed shape tokens) at the first layer, and a learned number of shape tokens at the second layer. We do not constrain a combination to have a fixed number of shape-tokens, but allow it to automatically and flexibly adapt to an object class. This number influences a combination's ability to represent shapes, where simple shapes favor fewer shape-tokens than complex ones. Consequently, discriminative combinations of varying complexity can be exploited to represent an object class. We learn this combination by exploiting distinguishing shape, geometric, and structural constraints of an object class. Shape constraints describe the visual aspect of shape tokens, while geometric constraints describe its spatial layout (configurations). Structural constraints enforce possible poses/structures of an object by the relationships (e.g., XOR relationship) between shape-tokens.

2. PROBLEM DEFINITION

The identification of objects in an image would probably start with image processing techniques such as noise removal, followed by (low-level) feature extraction to locate lines, regions and possibly areas with certain textures. The clever bit is to interpret collections of these shapes as single objects, e.g. cars on a road, boxes on a conveyor belt or cancerous cells on a microscope slide. One reason this is an AI problem is that an object can appear very different when viewed from different angles or under different lighting. Another problem is deciding what features belong to what object and which are background or shadows etc. The human visual system performs these tasks mostly unconsciously but a computer requires skillful programming and lots of processing power to approach human performance. Manipulating data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically or digitally with a computer.

To digitally process an image, it is first necessary to reduce the image to a series of numbers that can be manipulated by the computer. Each number representing the brightness value of the image at a particular location is called a picture element, or pixel. A typical digitized image may have 512×512 or roughly 250,000 pixels, although much larger images are becoming common. Once the image has been digitized, there are three basic operations that can be performed on it in the computer. For a point operation, a pixel value in the output image depends on a single pixel value in the input image. For local operations, several neighboring pixels in the input image determine the value of an output image pixel. In a global operation, all of the input image pixels contribute to an output image pixel value.

These operations, taken singly or in combination, are the means by which the image is enhanced, restored, or compressed. An image is enhanced when it is modified so that the information it contains is more clearly evident, but enhancement can also include making the image more visually appealing.

3. BACKGROUND FOR RECOGNITION

A. Image Normalization

Image normalization refers to pose and illumination changes, whose purpose is to achieve invariance to data capture conditions and to allow biometrics to operate in uncontrolled settings. Multiview face recognition ,a generative methods, that requires gallery images for each pose and thus cannot handle faces acquired from a quite novel viewpoint, one of the tasks FACE has to handle.

B. Recognition Response Reliability

It is not proper to rely only on distance between the probe and candidate subjects, when one needs to evaluate the reliability of a recognition response. Standard performance measures, such as RR, are useful to compare the performances of different systems, since they measure the overall recognition ability of a system. But they are global in nature and do not provide any clue on the reliability of a single operation. Evaluation of response reliability should rely on the quality of input data and also with galleries consisting of millions of images.

4. THE PROPOSED METHOD

4.1 Face Detection**:**

It is process to extract face regions from input image which has normalized intensity and uniform in size.

The appearance features are extracted from detected face part which describes changes of face such as furrows and wrinkles (skin texture).

In this system model, an executable (.dll-dynamic link library) file is utilized to extract face region.

It is used for face detection process is based on haar like features and adaptive boosting method.

4.2 Face Granulation:

This approach is used to represent the facial information in several parts to extract the features and discriminate presence of variations such as pose, expression and illumination.

To detect face granules, 2D gaussian low pass filter is used to generate difference of gaussian between two successive filtering at each reduced version of image.

At each iteration level, the image will be down sampled to desired size to make difference of gaussian pyramid. These granules are used to provide facial features such as smoothness, edge details and blurriness.

4.3.Difference of Gaussian pyramid creation:

The first stage is to construct a Gaussian "scale space" function from the input image [1]. This is formed by convolution (filtering) of the original image with Gaussian functions of varying widths. The difference of Gaussian (DoG), D(x, y, σ), is calculated as the difference between two filtered images, one with k multiplied by scale of the other.

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

These images, L(x, y, σ), are produced from the convolution of Gaussian functions, G(x, y, kσ), with an input image, I(x, y).

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\{-(x^2 + y^2)/2\sigma^2\}$$

This is the approach we use in the implementation. First, the initial image, I, is convolved with a Gaussian function, $G0$, of width σ 0 . Then we use this blurred image, $L0$, as the first image in the Gaussian pyramid and incrementally convolve it with a Gaussian, Gi, of width $\sigma i$to create the *ith* image in the image pyramid, which is equivalent to the original image filtered with a Gaussian, Gk , of width $k\sigma 0$.

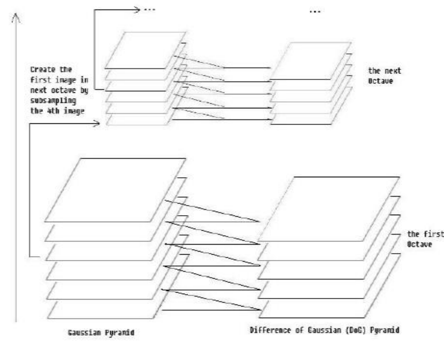

Fig.1 Difference of Gaussian pyramid

4.4 Weber's Local Descriptor:

In this section we give an overview of basic WLD descriptor and its extension. This descriptor represents an image as a histogram of differential excitations and gradient orientations, and has several interesting properties like robustness to noise and illumination changes, elegant detection of edges and powerful image representation.

WLD descriptor is based on Weber's Law. According to this law the ratio of the increment threshold to the background intensity is constant. Inspired by this law, Chen et.al [15] proposed WLD descriptor for texture representation. The computation of WLD descriptor involves three steps i.e. finding differential excitations, gradient orientations and building the histogram.

4.4.1Weber's Law:

Ernst Weber, an experimental psychologist in the 19th century, observed that the ratio of the increment threshold to the background intensity is a constant. This relationship, known since as Weber's Law, can be expressed as:

$$\frac{\Delta I}{I} = k,$$

Where ΔI represents the increment threshold (just noticeable difference for discrimination); I represents the initial stimulus intensity and k signifies that the proportion on the left side of the equation remains constant despite variations in the term. The fraction ΔI/I is known as the Weber fraction. Weber's Law, more simply stated, says that the size of a just noticeable difference (i.e., ΔI) is a constant proportion of the original stimulus value.

4.4.2.Differential Excitation:

We use the intensity differences between its neighbors and a current pixel as the changes of the current pixel. By this means, we hope to find the salient variations within an image to simulate the pattern perception of human beings. Specifically, a differential excitation ξ(xc) of a current pixel xc is computed as illustrated in

Fig. 1. We first calculate the differences between its neighbors and the center point using the filter f00:

$$v_s^{00} = \sum_{i=0}^{p-1}(\Delta x_i) = \sum_{i=0}^{p-1}(x_i - x_c),$$

Where xi (i=0,1,…p-1) denotes the i-th neighbors of xc and p is the number of neighbors. Following hints in Weber's Law, we then compute the ratio of the differences to the intensity of the current point by combining the outputs of the two filters f00 and f01 (whose output 01 s v is the original image in fact):

$$G_{ratio}(x_c) = {}^{v_s^{00}}\!/_{v_s^{01}}.$$

We then employ the arctangent function on G ratio(•):

$$G_{arctan}\left[G_{ratio}(x_c)\right] = \arctan\left[G_{ratio}(x_c)\right].$$

Combining (2), (3) and (4), we have:

$$G_{arctan}\left[G_{ratio}(x_c)\right] = \gamma_s^0 = \arctan\left[\frac{v_s^{00}}{v_s^{01}}\right] = \arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right].$$

So, the differential excitation of the current pixel ξ(xc) is computed as:

$$\xi(x_c) = \arctan\left[\frac{v_s^{00}}{v_s^{01}}\right] = \arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right].$$

Note that ξ(x) may take a minus value if the neighbor intensities are smaller than that of the current pixel. By this means, we attempt to preserve more discriminating information in comparison to using the absolute value of ξ(x). Intuitively, if ξ(x) is positive, it simulates the case that the surroundings are lighter than the current pixel. In contrast, if ξ(x) is negative, it simulates the case that the surroundings are darker than the current pixel.

4.4.3.Gradient Orientation :

Next main component of WLD is gradient orientation. For a pixel the gradient orientation is calculated as follows:

$$\theta(x_c) = \arctan\left[\frac{I_{73}}{I_{51}}\right]$$

Where is the intensity difference of two pixels on the left and right of the current pixel *xc*, and is the intensity difference of two pixels directly below and above the current pixel,

$$\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}].$$

4.5.WLD:

In this part, we describe the two components of WLD: differential excitation (ξ) and orientation (θ). After that we present how to compute a WLD histogram for an input image (or image region).

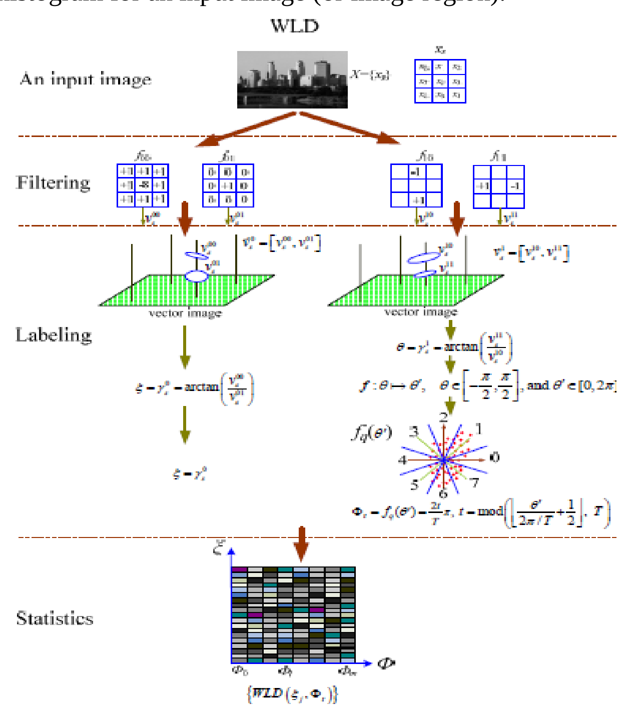

Fig.2 Illustration of computation of WLD descriptor

4.6.System Architecture:

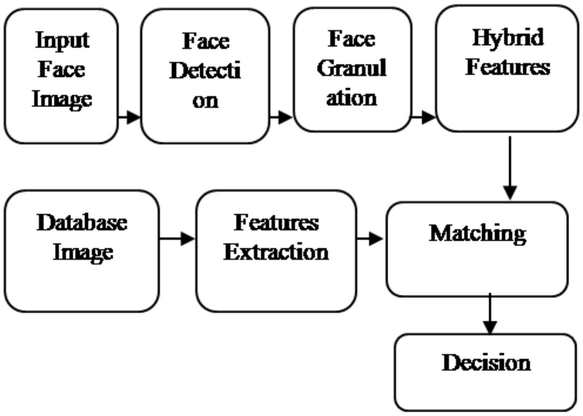

Fig.3 Architecture of face recognition system

4.6.1. Gabor Filter Approach:

The low frequency sub bands of two source images are fused based on selection of appropriate coefficients using Gabor filtering. It is useful to discriminate and characterize the texture of an image through frequency and orientation representation. It uses the Gaussian kernel function modulated by sinusoidal wave to evaluate the filter coefficients for convolving with an image.

The complex Gabor in space domain, here is the formula of a complex Gabor function in space domain

g(x, y) = s(x, y) wr(x, y)

where s(x; y) is a complex sinusoidal, known as the carrier, and wr(x; y) is a 2-D Gaussian-shaped function, known as the envelop. The complex sinusoidal is denotes as follows,

s(x, y) = exp (j (2*pi(u0 x + v0 y) + P))

where (u0, v0) and P denotes the spatial frequency and the phase of the sinusoidal respectively.

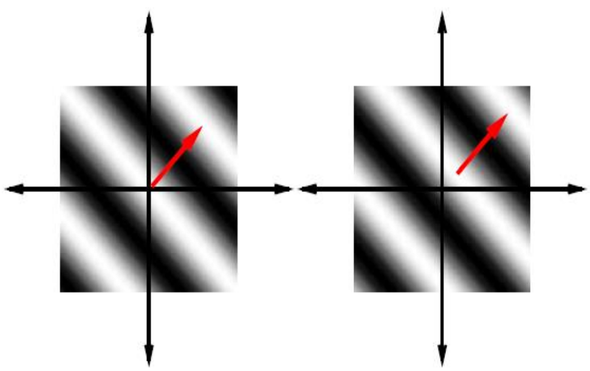

Fig.4 The real and imaginary parts of a complex sinusoidal.

The images are 128 X128 pixels. The parameters are: u0 = v0 = 1=80 cycles/pixel; P = 0 deg.

The real part and the imaginary part of this sinusoidal are

Re (s(x, y)) = cos (2*pi*(u0 x + v0 y) + P)

Im (s(x, y)) = sin (2*pi*(u0 x + v0 y) + P)

The parameters u0 and v0 denotes the spatial frequency of the sinusoidal in Cartesian coordinates. This spatial frequency can also be expressed in polar coordinates as magnitude F0 and direction w0:

$$F_0 = \sqrt{u_0^2 + v_0^2}$$

$$\omega_0 = \tan^{-1}\left(\frac{v_0}{u_0}\right)$$

ie.,

$$u_0 = F_0 \cos \omega_0$$

$$v_0 = F_0 \sin \omega_0$$

Using this representation, the complex sinusoidal is

$$s(x, y) = \exp\left(j\left(2\pi F_0\left(x \cos \omega_0 + y \sin \omega_0\right) + P\right)\right)$$

The Gaussian envelop looks as follows

$$w_r(x, y) = K \exp\left(-\pi\left(a^2\left(x - x_0\right)_r^2 + b^2\left(y - y_0\right)_r^2\right)\right)$$

where (x0; y0) is the peak of the function, a and b are scaling parameters of the Gaussian, and the r subscript stands for a rotation operation3 such that

$$(x - x_0)_r = (x - x_0)\cos\theta + (y - y_0)\sin\theta$$

$$(y - y_0)_r = -(x - x_0)\sin\theta + (y - y_0)\cos\theta$$

Gabor Face:

Fig.6 Gabor face images

5. SIMULATION AND RESULTS:

Texture properties include coarseness, contrast, directionality, line-likeness, regularity and roughness. Texture is one of the most important defining features of an image. It is characterized by the spatial distribution of gray levels in a neighborhood [8]. In order to capture the spatial dependence of gray-level values, which contribute to the perception of texture, a two-dimensional dependence texture analysis matrix is taken into consideration. This two-dimensional matrix is obtained by decoding the image file; jpeg, bmp, etc.

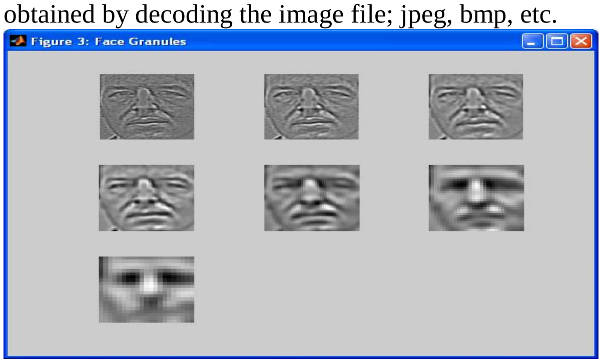

Fig .7 Face Granules

5.1.Methods of Representation:

There are three principal approaches used to describe texture; statistical, structural and spectral. Statistical techniques characterize textures using the statistical properties of the grey levels of the points/pixels comprising a surface image. Typically, these properties are computed using: the grey level co-occurrence matrix of the surface, or the wavelet transformation of the surface. Structural techniques characterize textures as being composed of simple primitive structures called "texels" (or texture elements). These are arranged regularly on a surface according to some surface arrangement rules.

Spectral techniques are based on properties of the Fourier spectrum and describe global periodicity of the grey levels of a surface by identifying high-energy peaks in the Fourier spectrum .

R.M. Haralick, the co-occurrence matrix representation of texture features explores the grey level spatial dependence of texture .

A mathematical definition of the co-occurrence matrix is as follows:

Given a position operator *P(i,j)*, let *A*be an n x n matrix whose element *A[i][j]* is the number of times that points with grey level (intensity) *g[i]* occur, in the position specified by *P*, relative to points with grey level *g[j]*.

Let *C* be the *n* x *n* matrix that is produced by dividing *A* with the total number of point pairs that satisfy *P*. *C[i][j]* is a measure of the joint probability that a pair of points satisfying *P* will have values *g[i]*, *g[j]*.

C is called a co-occurrence matrix defined by *P*.

Examples for the operator *P* are: "*i* above *j*", or "*i* one position to the right and two below *j*", etc.

This can also be illustrated as follows… Let *t* be a translation, then a co-occurrence matrix $C_t$of a region is defined for every grey-level *(a, b)* by [1]:

$$C_t(a,b) = card\{(s,s+t) \in R^2 | A[s] = a, A[s+t] = b\}$$

Here, $C_t(a, b)$ is the number of site-couples, denoted by *(s, s + t)* that are separated by a translation vector *t*, with *a* being the grey-level of *s*, and *b* being the grey-level of *s* + *t*.

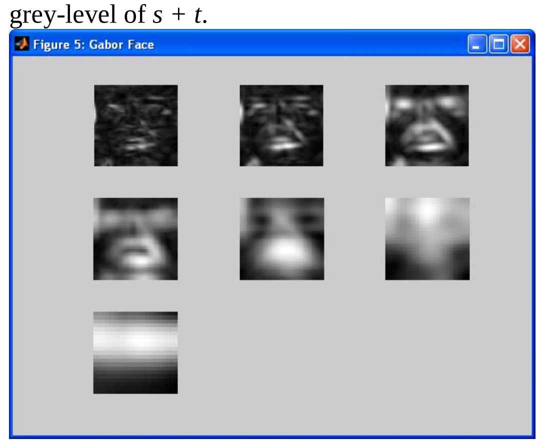

Fig 1.4: Gabor Face

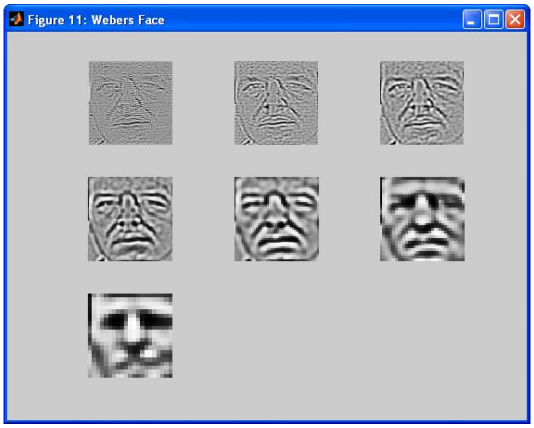

Fig 1.5: Weber's Faces

5.2.FEATURE MATCHING

Euclidean Distance:

Euclidean distance measures the similarity between two different feature vectors using (7).

$$ED = \sqrt{\sum_{j=0}^{J} (FV_{1,j} - FV_{2,j})^2} \quad (7)$$

where J is the length of the feature vector, Fvi is the feature vector for individual i.

6. CONCLUSION

The paper presented the robust human face recognition system based on granular computation and hybrid spatial features extraction. Here granular computing based on the Gaussian operator was used to decompose the image into different scale spaces for effective texture representation. The texture descriptors called Gabor filter bank and Weber's local descriptor was used here to characterize the face appearance. These approaches were well used to identify the illumination changes, intensity distributions characteristics. Here, matching was done between input and original samples using Euclidean distance metrics. Finally the simulated results shows that used methodologies provides better recognition rate with minimum error rate for all samples.

REFERENCES

[1] A.A. Mohammed, R. Minhas, Q.M. Jonathan Wu, and M.A. Sid-Ahmed, "Human face recognition based on Multidimensional PCA and extreme learning machine," Pattern Recognition, vol 44, pp. 2588-2597, 2011.

[2]]Xiaoyang Tan and Bill Triggs, Enhanced Local Texture Feature Sets for Face Recognition under difficult lighting conditions, IEEE transactions on Image Processing, vol.l9, No.6, June 2010

[3] X. Zhao and Y. GAO, "Face recognition across pose: A review," *PatternRecogn.*, vol. 42, no. 11, pp. 2876–2896, 2009.

[4] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition:A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4,pp. 399–458, 2003.

[5] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with localbinary patterns," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 469–481.

[6] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. a, "Robust facerecognition via sparse representation," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[7] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition:A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4,pp. 399–458, 2003.

[8] X. Zhao and Y. GAO, "Face recognition across pose: A review," *PatternRecogn.*, vol. 42, no. 11, pp. 2876–2896, 2009.

[9] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with localbinary patterns," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 469–481.

[10] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust facerecognition via sparse representation," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.[5] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit.Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991

[11] M. Turk and A. Pentland, "Eigenfaces for recognition", J. Cognitive Neuroscience, Vol. 3, 71-86, 1991

[12] D. L. Swets and J. J. Weng, "Using discriminate eigenfeatures for image retrieval", IEEE Trans. PAMI., Vol. 18, No. 8, 831-836, 1996.

[13] A. K. Jain, "Fundamentals of digital image processing", pp.163-175, Prentice Hall, 1989.

[14] B Moghaddam, W Wahid and A pentland, "Beyond eigenfaces: Probabilistic matching for face recognition", Proceeding of face and gesture recognition, pp. 30 –35, 1998.

[15] I. Daubechies, "The wavelet transform time-frequency localization and signal analysis", IEEE Trans. Information Theory, Vol. 36, No. 5, 961- 1005, 1990.

