ISSN (e): 2250 – 3005 || Vol, 04 || Issue, 10 || October– 2014 ||
International Journal of Computational Engineering Research (IJCER)

Scale Invariant Feature Transform Based Face Recognition from a Single Sample per Person

R.Pavithra[1], Prof. A. Usha Ruby [2], Dr. J. George Chellin Chandran[3]
Dept of PG CSE, CSI College of Engineering, Ketti, The Nilgiris, India[1]
Research scholar, Bharath University, Chennai, India[2]

ABSTRACT

The technological growth has a serious impact on security which has its own significance. The core objective of this project is to extract the facial features using the local appearance based method for the accurate face identification with single sample per class .The face biometric based person identification plays a major role in wide range of applications such as Airport security, Driver's license, Passport, Voting System, Surveillance. This project presents face recognition based on granular computing and robust feature extraction using Scale Invariant Feature Transform (SIFT) approach. The Median filter is used to extract the hybrid features and the pyramids are generated after the face granulation. 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 feature extraction stage, SIFT descriptor utilized to assign the intersecting points which are invariant to natural distortions. This feature is 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 used granulation and feature descriptors has better discriminatory power and recognition accuracy in the process of recognizing different facial appearance.

INDEX TERMS*: Single sample per class, Median Filter, Dog pyramid, Scale Invariant Feature Transform*

I. INTRODUCTION

"Biometrics" means "life measurement" but the term is usually associated with the use of unique physiological characteristics to identify an individual. The application which most people associate with biometrics is security. However, biometric identification has eventually a much broader relevance as computer interface becomes more natural. Knowing the person with whom you are conversing is an important part of human interaction and one expects computers of the future to have the same capabilities. A number of biometric traits have been developed and are used to authenticate the person's identity. The idea is to use the special characteristics of a person to identify him. By using special characteristics we mean the using the features such as face, iris, fingerprint, signature etc, this method of identification based on biometric characteristics is preferred over traditional passwords and PIN based methods for various reasons such as: The person to be identified is required to be physically present at the time-of-identification. Identification based on biometric techniques obviates the need to remember a password or carry a token.

A biometric system is essentially a pattern recognition system which makes a personal identification by determining the authenticity of a specific physiological or behavioral characteristic possessed by the user. Biometric technologies are thus defined as the "automated methods of identifying or authenticating the identity of a living person based on a physiological or behavioral characteristic".
A biometric system can be either an 'identification' system or a 'verification' (authentication) system, which are defined below.

Identification **-** One to Many: Biometrics can be used to determine a person's identity even without his knowledge or consent. For example, scanning a crowd with a camera and using face recognition technology, one can determine matches against a known database.

Verification : One to One: Biometrics can also be used to verify a person's identity. For example, one can grant physical access to a secure area in a building by using finger scans or can grant access to a bank account at an ATM by using retinal scan. The process of receiving and analyzing visual information by the human species is referred to as sight, perception or understanding. Similarly, the process of receiving and analyzing visual information by digital computer is called as digital image processing and scene analysis. Processing of an image includes improvement in its appearance and efficient representation. So the field consists of not only feature extraction, analysis and recognition of images, but also coding, filtering, enhancement and restoration. The entire process of image processing and analysis starts from the receiving of the visual information and ends in giving out of description of the scene. The major components of the image processing system are image sensor, digitizer, processor, display unit and storage unit.

Face Recognition-A Survey : A facial recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a facial database. It is typically used in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems. Face recognition technology is used to extract information from facial images with the help of a face recognition device, without any human interaction. Unlike face detection technology, face recognition technology uses image processing algorithms to recognize, and then compare human facial images with the ones that are stored in the database of face recognition device. Face recognition technology enabled device analyzes the characteristics of overall structure of human face, including width of nose, shape of cheekbones, width between nose and jaw edges, and distance between eyes. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with matching features. Other algorithms normalize a gallery of face images and then compress the face data, only saving the data in the image that is useful for face recognition. A probe image is then compared with the face data. One of the earliest successful systems is based on template matching techniques applied to a set of salient facial features, providing a sort of compressed face representation. Recognition algorithms [1, 3,5, 10, 12] can be divided into two main approaches, geometric, which look at distinguishing features, or photometric, which is a statistical approach that distills an image into values and compares the values with templates to eliminate variances. Popular recognition algorithms include Principal Component Analysis using Eigen faces [1, 3], Linear Discriminate Analysis [5], Elastic Bunch Graph Matching using the Fisher face algorithm, the Hidden Markov model, the Multi linear Subspace Learning using tensor representation.

II. EXISTING SYSTEM

Discriminant analysis methods are tools for face recognition. But this is not be used for the single sample per person scenario because of “with- in subject variability”. This variability is established using images in the generic training set for which more than one sample per person is available. When images are under drastic facial expression variation, the discriminant analysis method can’t be used. LDA (Linear Discriminant Analysis) [5] is one of the discriminant analysis methods. It fails to improve the performance, because the complex distribution of the data set is caused by large intrapersonal variations, which still exist in each cluster of that method. LDA uses a Fisher face [1,5] algorithm. It uses the dataset to store multiple images (with- in variability) of a same person. PCA (Principal Component Analysis) [2] is also not suitable for the single sample per class problem. LBP (Local Binary Pattern) [12] is not only used for face detection but also for face recognition. But LBP is not a robust face finder. 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. 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. There are two popular approaches to face recognition. One approach transforms face images into specific transformation domains. Among the works that appear in the literature are Eigen face, Gabor filters, Fourier Transform, and wavelets. Another approach is to extract principal lines and creases from the face. However, this method is not easy because it is sometimes difficult to extract the line structures that can discriminate every individual well.
From the literature survey, it had been inferred that local matching method will be more efficient comparing to holistic matching method. Hence we can eliminate the high dimensionality problem. It had been found that

multiple samples per class lead to increase space complexity. In order avoid this space complexity problem, single sample per class is considered. In this single sample per class single image of a person is used for matching the face image.

III. PROPOSED SYSTEM

The proposed system use a robust face finder called SIFT(Scale Invariant Feature Transform). By using this feature the face can be fluently detected and recognized. Although only one classifier is trained, and using that frontal, occluded and profile faces are detected.

The proposed system includes Four modules: (1) Image Preprocessing, (2) Feature extraction, (3) face granulation, (4)Face recognition or face identification.

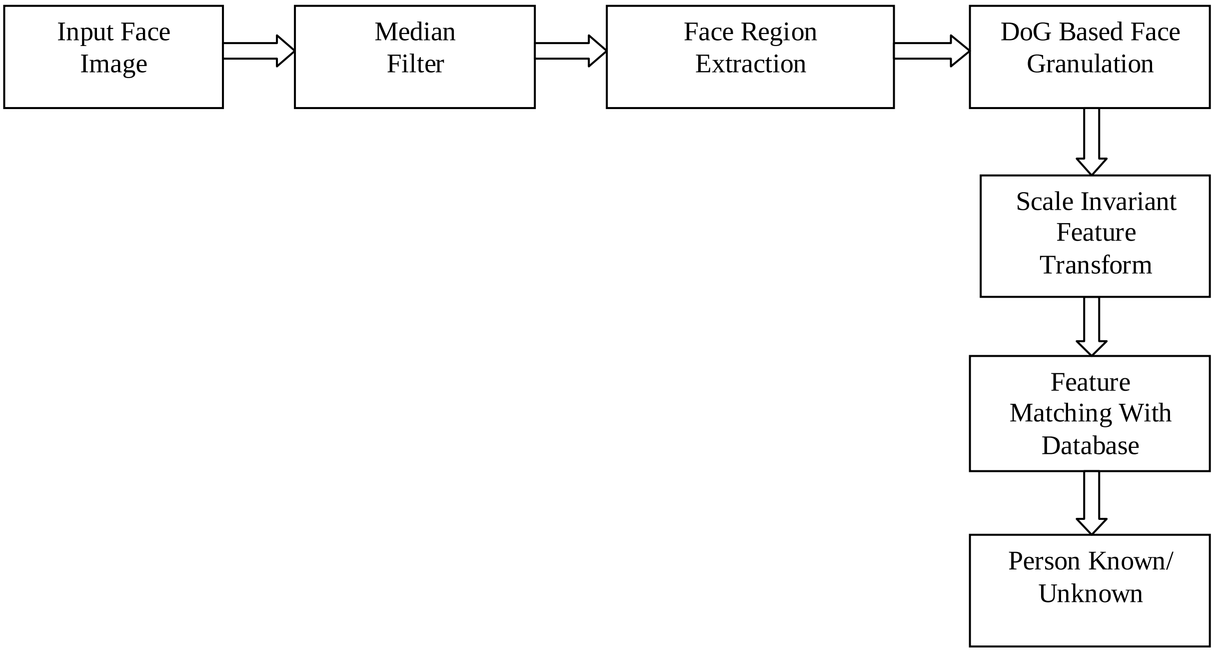

Fig 1 Detailed Diagram for Proposed System

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 face 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 feature extraction stage, SIFT descriptor utilized to assign the intersecting points which are invariant to natural distortions. This feature is useful to distinguish the maximum number of samples accurately and it is matched with already stored original face samples for identification.

Granular Computing & DoG : After digital image has been obtained, the next step deals with preprocessing that image. The key function of preprocessing is to improve the image in ways that increase the chances for success of the other processes. Typically preprocessing deals with techniques for segregating at different resolutions to provide edge information, noise, smoothness and blurriness present in a face image. Subsequently the face granulation is done to generate the Difference of Gaussian (DoG) pyramids to recognize the face by using the granular computing and face 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.

Face Recognition : Recognition is the process that assigns a mark to an object depends on the information provided by its descriptors. Having extracted the hybrid features in the feature extraction step, the Face granulation is done to produce out the Difference of Gaussian (DoG) pyramid formation. The DoG pyramids are produced to analyze the Euclidean distance to discern the face images valiantly

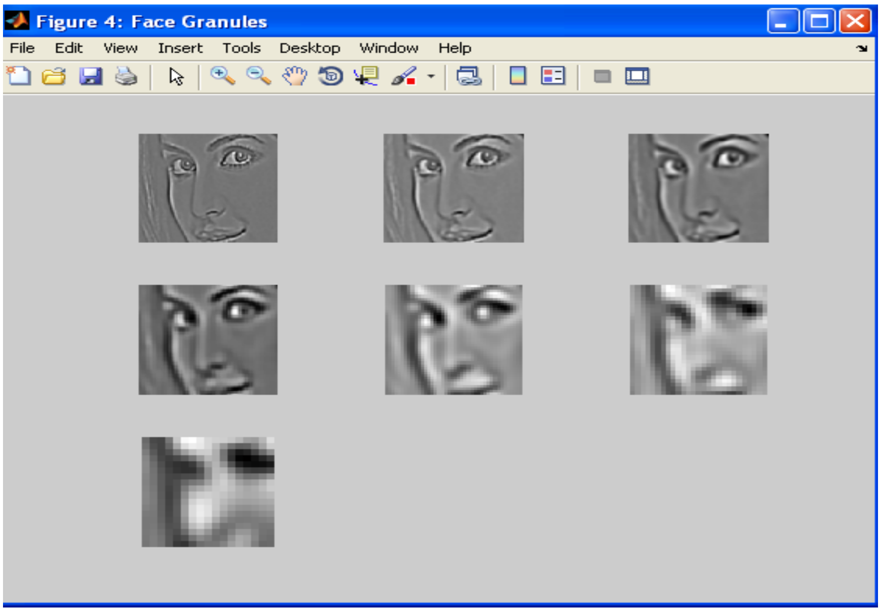

Fig 2 Face Granulation Values

The above Fig 2 gives us the clearest idea of getting out the output values based upon the identification of face vector values along with the different usage of granules computed and taken as successive outcomes. The performance is to be measured based on the comparison with the normal input image and the classified granules computed image.

IV ALGORITHMS

4.1 Algorithm for Difference of Gaussian (DoG)

Difference of Gaussians is a feature enhancement algorithm that involves the subtraction of one blurred version of an original image from another, less blurred version of the original. In the simple case of grayscale images, the blurred images are obtained by convolving the original grayscale images with Gaussian kernels having differing standard deviations. Blurring an image using a Gaussian kernel suppresses only high-frequency spatial information. Subtracting one image from the other preserves spatial information that lies between the ranges of frequencies that are preserved in the two blurred images. Thus, the difference of Gaussians is a band-pass filter that discards all but a handful of spatial frequencies that are present in the original gray scale image.
i)MATHEMATICS OF DIFFERENCE OF GAUSSIANS
Given a m-channels, n-dimensional image

$$I : \{\mathbb{X} \subseteq \mathbb{R}^n\} \rightarrow \{\mathbb{Y} \subseteq \mathbb{R}^m\} \quad \text{--------------(1)}$$

The difference of Gaussians (DoG) of the image I is the function

$$\Gamma_{\sigma1,\sigma2} : \{\mathbb{X} \subseteq \mathbb{R}^n\} \rightarrow \{\mathbb{Z} \subseteq \mathbb{R}\} \quad \text{--------(2)}$$

Obtained by subtracting the image I convolved with the Gaussian of variance σ_2^2 from the image I convolved with a Gaussian of narrower variance σ_1^2, with $\sigma_2 > \sigma_1$. In one dimension, Γ is defined as:

$$\Gamma_{\sigma1,\sigma2}(x) = I * \frac{1}{\sigma_1\sqrt{2\pi}} e^{-(x^2)/(2\sigma_1^2)} - I * \frac{1}{\sigma_2\sqrt{2\pi}} e^{-(x^2)/(2\sigma_2^2)} \quad \text{------------- (3)}$$

and for the centered two-dimensional case :

$$\Gamma_{\sigma,K\sigma}(x,y) = I * \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} - I * \frac{1}{2\pi K^2\sigma^2} e^{-(x^2+y^2)/(2K^2\sigma^2)} \quad \text{------------ (4)}$$

Which is formally equivalent to:

$$\Gamma_{\sigma,K\sigma}(x,y) = I * \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/(2\sigma^2)} - \frac{1}{2\pi K^2\sigma^2} e^{-(x^2+y^2)/(2K^2\sigma^2)} \text{ ---------------- (5)}$$

Which represents an image convoluted to the difference of two Gaussians, which approximates a Mexican Hat function. As a feature enhancement algorithm, the difference of Gaussians can be utilized to increase the visibility of edges and other detail present in a digital image. The difference of Gaussians algorithm removes high frequency detail that often includes random noise, rendering this approach one of the most suitable for processing images with a high degree of noise. Differences of Gaussians have also been used for blob detection in the scale-invariant feature transform.

In fact, the DoG [11] as the difference of two Multivariate normal distribution has always a total null sum and convolving it with a uniform signal generates no response. It approximates well a second derivate of Gaussian (Laplacian of Gaussian) with K~1.6 and the receptive fields of ganglion cells in the retina with K~5. It may easily be used in recursive schemes and is used as an operator in real-time algorithms for blob detection and automatic scale selection.

4.2 Scale Invariant Feature Transform(SIFT) : It generates image features, “key points invariant to image scaling and rotation partially invariant to change in illumination [14].For any object there are many features, interesting points on the object that can be extracted to provide a "feature" description of the object. SIFT image features provide a set of features of an object that are not affected by many of the complications experienced in other methods, such as object scaling and rotation. While allowing for an object to be recognized in a larger image SIFT image features also allow for objects in multiple images of the same location, taken from different positions within the environment, to be recognized. SIFT features are also very resilient to the effects of "noise" in the image.

To aid the extraction of these features the SIFT algorithm applies a 4 stage filtering approach:

i)Scale-Space Extreme Detection : This stage of the filtering attempts to identify those locations and scales that is identifiable from different views of the same object. This can be efficiently achieved using a "scale space" function. Further it has been shown under reasonable assumptions it must be based on the Gaussian function.
The scale space is defined by the function:
$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$
Where * is the convolution operator, $G(x, y, \sigma)$ is a variable-scale Gaussian and $I(x, y)$ is the input image.
Difference of Gaussians is one such technique, locating scale-space extreme, $D(x, y, \sigma)$ by computing the difference between two images, one with scale k times the other. $D(x, y, \sigma)$ is then given by:
$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$
To detect the local maxima and minima of $D(x, y, \sigma)$ each point is compared with its 8 neighbors at the same scale, and its 9 neighbors up and down one scale. If this value is the minimum or maximum of all these points then this point is an extreme.

ii) Key point Localization : This stage attempts to eliminate more points from the list of key points by finding those that have low contrast or are poorly localized on an edge.

It involves three main steps. They are

a) Interpolation of nearby data for accurate position : First, for each candidate key point, interpolation of nearby data is used to accurately determine its position. The initial approach was to just locate each key point at the location and scale of the candidate key point. The new approach calculates the interpolated location of the extreme, which substantially improves matching and stability.

b) Discarding low-contrast key points : To discard the key points with low contrast, the value of the second-order Taylor expansion D(x) is computed at the offset x. If this value is less than 0.03, the candidate key point is discarded.

c) Eliminating edge responses : The DoG function will have strong responses along edges, even if the candidate key point is not robust to small amounts of noise.

For poorly defined peaks in the DoG function, the principal curvature across the edge would be much larger than the principal curvature along it. Finding these principal curvatures amounts to solving for the eigenvalues of the second-order Hessian matrix, H:

$$H = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix}$$

iii)ORIENTATION ASSIGNMENT

This step aims to assign a consistent orientation to the key points based on local image properties. The key point descriptor can then be represented relative to this orientation, achieving invariance to rotation.

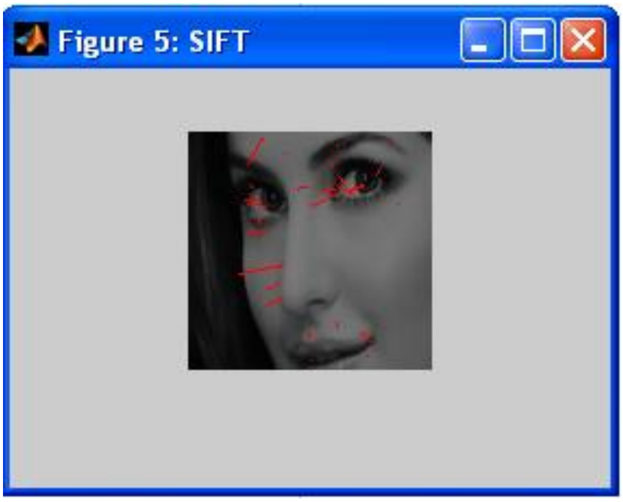

Fig 3 SIFT Calculating Values

iv)Key point Descriptor : The local gradient data, used above, is also used to create key point descriptors. The gradient information is rotated to line up with the orientation of the key point and then weighted by a Gaussian with variance of 1.5 * key point scale. This data is then used to create a set of histograms over a window centered on the key point. Key point descriptors typically uses a set of 16 histograms, aligned in a 4x4 grid, each with 8 orientation bins, one for each of the main compass directions and one for each of the mid-points of these directions. This result in a feature vector containing 128 elements. These resulting vectors are known as SIFT keys and are used in a nearest-neighbors approach to identify possible objects in an image.

V EXPERIMENTAL RESULTS

In this section we have evaluated the performance of Difference of Gaussian and feature extraction based on Scale Invariant Feature Transform. Here we test the proposed approach using FERET dataset for face recognition. The image is cropped and made into 64X64 from middle of location of eyes. Here we have considered the local features such as eyes, nose, mouth and chin and also detects properties of relations (e.g. areas, distances, angles) between the features are used as descriptors for face recognition. We have considered 1sample per person for each individual person. Here we will store each person's image in the FERET database and later which can be used for matching. FERET database is a standard database which is used for storing images. And the resolution of image is 128X128.The results shows that the 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 results produced by using this method provide better discriminatory power for recognizing different facial appearance with accurate results.

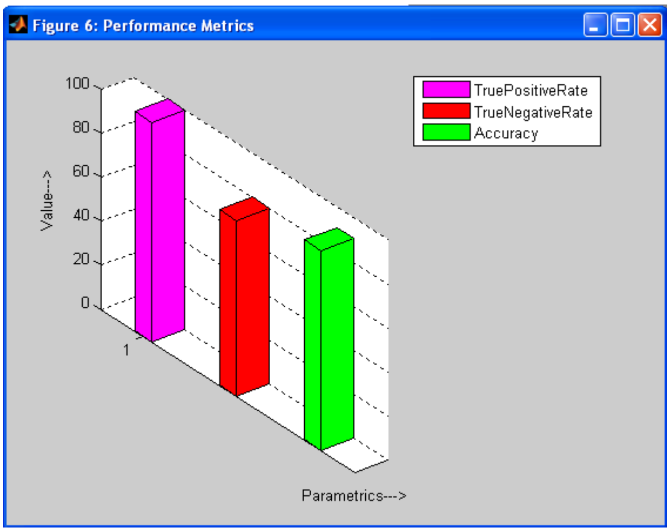

Fig 4 Performance of the Proposed System

VI CONCLUSION AND FUTURE ENHANCEMENT

In this paper, we used local matching method for Face recognition using Difference of Gaussian, Scale Invariant Feature Transform .Here first the facial features are features are segregated at different resolutions to provide edge information, noise, smoothness and blurriness present in a face image using Difference of Gaussian. SIFT descriptor utilized to assign the intersecting points which are invariant to natural distortions .This feature is useful to distinguish the maximum number of samples accurately and it is matched with already stored original face samples for identification. Since we are using single sample, the space and time complexity is reduced and performance has improved. In future, the face recognition system will be enhanced by applying Robust Local Ternary Pattern.

REFERENCES

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

[2] Gunjan Dashore, Dr. V.Cyril Raj Dr.MGR Educational and Research Institute Chennai, Tamilnadu, INDIA," An Efficient Method For Face Recognition Using Principal Component Analysis (PCA)"

[3] Yang, D. Zhang, A. Frangi and J. Yang,"Two dimensional PCA: a new approach to appearance-based face representation and recognition" IEEE Transactions on Pattern Analysis and Machine Intelligence, *vol. 26, pp. 131-137, 2004.*

[4] J. Wu and Z. Zhou, "Face recognition with one training image per person", Pattern Recognition Letters, *vol. 23, no. 14, pp. 1711-1719, 2002.*

[5] Li-Fen Chen!, Hong-Yuan Mark Liao",Ming-Tat Ko", Ja-Chen Lin!, Gwo-Jong Yu," A new LDA-based face recognition system which can solve the small sample size problem."

[6]. Turk M, Pentland A, Neurosci J (1991) Eigenfaces for recognition. J CognNeurosci *3(1):71–86*

[7] Lawrence S, Lee Giles C, Tsoi A, Back A (1997 Face recognition:a convolutional neural-network approach.IEEE Trans Neural Netw *8(1):98–113*

[8] A.Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class", *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 6, pp. 748-763, 2002.*

[9] X. Tan, S. Chen, Z. Zhou and F. Zhang, "Face recognition from a single image per person: A survey", *Pattern Recognition, vol. 39, pp. 1725-1745, 2006.*

[10] 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.*

[11]. Jie Chen, Member, IEEE, Shiguang Shan, Member, IEEE, Chu He, Guoying Zhao, Matti Pietika¨inen, Senior Member, IEEE, Xilin Chen, Senior Member, IEEE, and Wen Gao, Fellow, IEEE,(2010)WLD: A Robust Local Image Descriptor. *IEEE transactions on pattern analysis and machine intelligence, vol. 32, no. 9.*

[12] S. Li, W. Liu, D. Lin and X. Tang, "Nonparametric subspace analysis for face recognition", *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 961-966, 2005.*

[13] R. Brunelli and T. Poggio, "Face recognition: feature versus templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052, 1993.*

[14] Yu-Yao Wang, Zheng-Ming Li, Long Wang, Min Wang, Harbin Institute of Technology Shenzhen Graduate School," A Scale Invariant Feature Transform Based Method"*Journal of Information Hiding and Multimedia Signal Processing* ©2013 *ISSN 2073-4212* Ubiquitous International *Volume 4, Number 2, April 2013.*