Chapter1

INTRODUCTION

Facial expressions recognition has an important role in the field of intelligent human-computer interaction, computer animation, surveillance and security, medical diagnosis, law enforcement, and awareness systems etc. Therefore, it has been an active research topic in various fields such as psychology, Cognitive science, human-computer interaction, and pattern recognition. Automatic facial expression analysis is emerging research area from video or images has received much attention in last two decades.

This chapter introduces recent advances in computer recognition of facial expressions. Firstly, we describe the problem space, which includes multiple dimensions: level of description, static versus dynamic expression, facial feature extraction and representation, facial expression recognition, controlled versus uncontrolled data acquisition, correlation with bodily expression, and databases.

FER(face expression recognition) can be thought as an inter-disciplinary problem of image-video processing, pattern recognition, psychology and studies to increase accuracy and speed has been carried out for the last 20 years. An important problem with recognition task is the number of expressions apart from common ones such as anger, joy and fear. Several other messages should be recognized in such a system. A smiling mouth and raised eyebrows meaning "I don’t know” can be given as an example to such messages. There are also expressions commonly used by a specific person but very rare in public. This makes the problem specific for each person and requires an adaptive classification mechanism in interpretation of subject's expression.

In-order to overcome this deriving an effective facial representation from original face images is a vital step for successful facial expression recognition. For the Purpose of Pre-Processing the image Gabor wavelets, are applied to either the whole-face or specific face-regions to extract the appearance changes of the face. Due to their superior performance, the major works on appearance-based methods have focused on using Gabor-wavelet representations. However, it is both time and memory intensive to convolve face images with a bank of Gabor filters to extract multi-scale and multi-orientation coefficients. Further more the proposed work proceeds towards extracting feature set using PCA and SURF and classification technique LBP (local Binary Pattern) which provides improved and better results in comparison to conventional approaches.

.

1.1 Aim of Thesis

We propose a technique for extracting two type of Data Set and there classification with learning mechanism for efficient recognition

- Extract Texture segmentation like feature using Gabor Wavelet.
- Extract Feature for face and apply PCA for dimension reduction of redundant feature vector obtained from Gabor wavelet
- Extract feature for organs like eyes, mouth, nose using the SURF (Speeded

Up Robust Features) method proposes procedures to compute local interest points and descriptors at a higher speed than the SIFT approach.

- Combining these two unique data set and using Local Binary pattern classification Technique to yield improved and efficient recognition rate.

1.2 Related Work

Automatic facial expression recognition has attracted much attention from behavioral scientists since the work of Darwin in1872. Suwa et al made the first attempt to automatically analyze facial expressions from image sequences in 1978. Much progress has been made in the last decade, and a thorough survey

of the exiting work can be found in [1,2]. Here we briefly review some previous work in order to put our work in context.

Automatic facial expression recognition involves two vital aspects: facial representation and classifier design.

Facial representation is to derive a set of features from original face images to effectively represent faces. The optimal features should minimize within-class variations of expressions while maximize betweenclass variations. If inadequate features are used, even the best classifier could fail to achieve accurate recognition. In some existing work, optical flow analysis has been used to model

muscles activities or estimate the displacements of feature points. However, flow estimates are easily disturbed by the non rigid motion and varying lighting, and are sensitive to the inaccuracy of image registration and motion discontinuities.

Facial geometry analysis has been widely exploited in facial representation, where shapes and locations of facial components are extracted to represent the face geometry. For example, Zhanget al. [25] used the geometric positions of 34 fiducial points as facial features to represent facial images. In image sequences, the facial movements can be qualified by measuring the geometrical displacement of facial feature points between the current frame and the initial frame motion.

Another kind of method to represent faces is to model the appearance changes of faces. Holistic spatial analysis including Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) , Independent Component Analysis (ICA) [40] and Gabor wavelet analysis have been applied to either the whole-face or specific face regions to extract the facial appearance changes. Donato et al. [8] explored different techniques to represent face images for facial action recognition, which include PCA, ICA, Local Feature Analysis (LFA), LDA and local schemes such as Gabor-wavelet representation and local principal components. Best performances were obtained using Gabor-wavelet representation and ICA. Due to their superior performance, Gabor-wavelet representations have been widely adopted in face image analysis. However, the computation of Gabor-wavelet representations is both time and memory intensive, 48 _ 48 face image has the high dimensionality. Recently Local Binary Patterns have been introduced as effective appearance features for facial image analysis. We compared LBP features with Gabor features for facial expression recognition, and studied their performance over a range of image .We further presented facial expression manifold learning in the LBP feature space.

Chapter 2

RESEARCH BACKGROUND

In this chapter we will describe about the technologies and algorithms which we will be using for our Methodology, these basic description will help to understand our thesis result

2.1 Gabor Filter

Preprocessing procedure is very important step for facial expression recognition. The ideal output of processing is to obtain pure facial expression images, which have normalized intensity, uniform size and shape. It also should eliminate the effect of illumination and lighting. The preprocessing procedure of our FER system performs the following five steps in converting a TIFF JAFFE image to a normalized pure expression image for feature extraction: 1). detecting facial feature points manually including eyes, nose and mouth; 2). rotating to line up the eye coordinates; 3) locating and cropping the face region using a rectangle according to face model [5] as shown in Fig.1. Suppose the distance between two eyes is *d*, the rectangle will be $2.2d \times 1.8d$; 4). scaling the image to fixed size of 128×96, locating the center position of the two eyes to a fixed position; 5). using a histogram equalization method to eliminate illumination effect. Fig.2 illustrates some examples of pure facial expression images after preprocessing from JAFFE database.

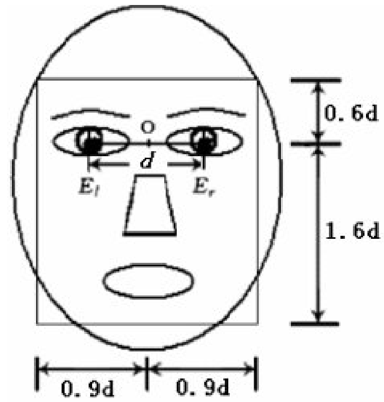


Fig. 1. *Facial model*

Fig. 2. Ex: of facial expression images after from JAFFE database.

2.1.1 Gabor Feature Extraction

The Gabor filters, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells [3][4], have been considered as a very useful tool in computer vision and image analysis due to its optimal localization properties in both spatial analysis and frequency domain

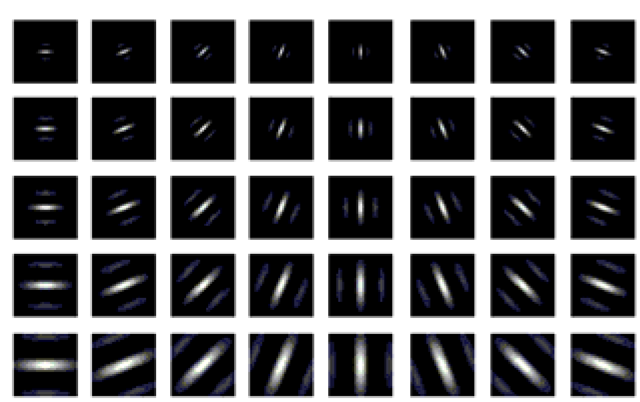


Fig. 3. *The real part of the Gabor filters with fivefrequencies and eight orientations for* $\omega_{max}=\pi/2$*, the row corresponds to different frequency* ω_m*, the column corresponds to different orientation* θ_n*.*

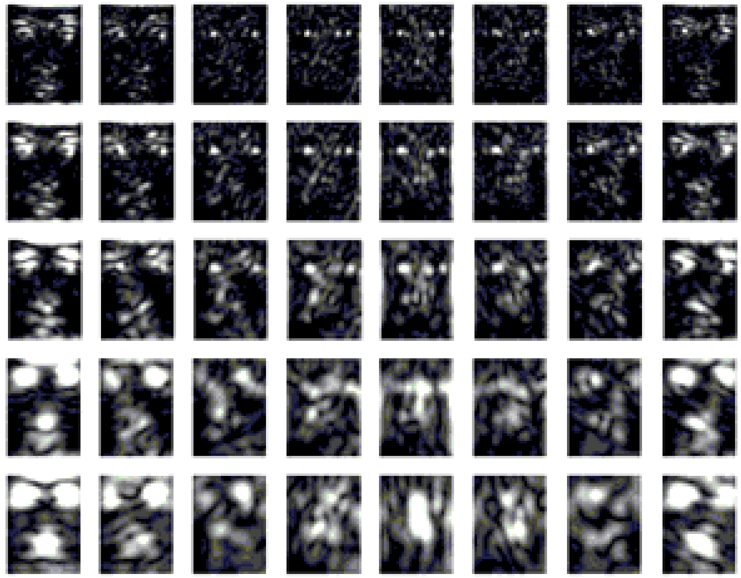


Fig. 4. *The magnitudes of the Gabor feature representation of the first face image in Fig.2*

In the spatial domain, a Gabor filter is a complex exponential modulated by a Gaussian function [4]. The Gabor filter can be defined as follows,

$$\psi(x,y,\varpi,\theta)=\frac{1}{2\pi\sigma^2}e^{-(\frac{x'^2+y'^2}{2\sigma^2})}[e^{i\varpi x'}-e^{-\frac{\varpi^2\sigma^2}{2}}]$$

$$x'=x\cos\theta+y\sin\theta,\ y'=-x\sin\theta+y\cos\theta$$

where (x, y) is the pixel position in the spatial domain, ω the radial center frequency, θ the orientation of Gabor filter, and σ the standard deviation of the round Gaussian function along the x- and y-axes. In addition, the second term of the Gabor filter, , compensates for the DC value because the cosine component has nonzero mean while the sine component has zero mean. According to [4], we set $\sigma \approx \pi/\omega$ to define the relationship between σ and ω in our experiments.

2/22σ□−*e*

In most cases a Gabor filter bank with five frequencies and eight orientations [1][2][18] is used to extract the Gabor feature for face representation. Selecting the maximum frequency $\omega_{max}=\pi/2$, $\omega_m=\omega_{(max)}\times\lambda^{-(m-1)}$, m={1,2,3,4,5}, $2=\lambda$, $\theta_n=(n-1)\pi/8$, n={1,2,…,8}, the real part of the Gabor filters with five frequencies and eight orientations is shown in Fig.3. From Fig.3 it can be seen that the Gabor filters exhibit strong characteristics of spatial locality and orientation selectivity.

2.1.2Gabor Feature Representation

The Gabor feature representation of an image I(x, y) is the convolution of the image with the Gabor filter bank $\psi(x, y, \omega_m, \theta_n)$ as given by:

$$O_{m,n}(x,y) = I(x,y) * \psi(x,y,\varpi_m,\theta_n)$$

where * denotes the convolution operator. The magnitude of the convolution outputs of a sample image (the first image in Fig.2) corresponding to the filter bank in Fig.3 is shown in

Fig.4. In practice, the time for performing Gabor feature extraction is very long and the dimension of Gabor feature vector is prohibitively large. For example, if the size of normalized

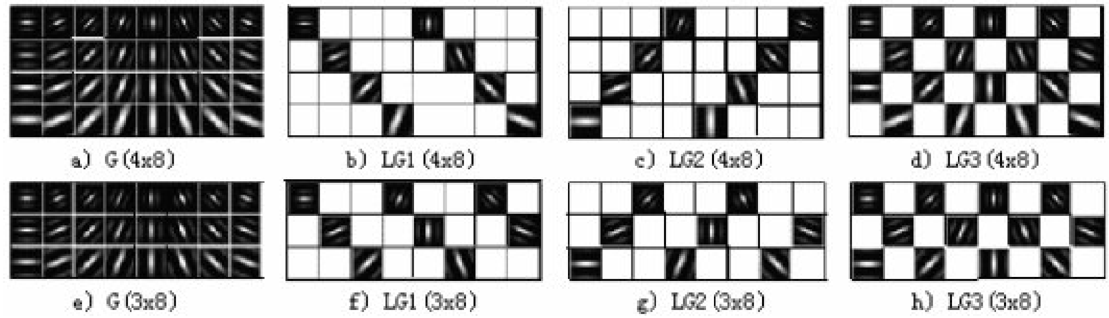


image is 128×96, the dimension of the Gabor feature vector with 40 filters will result in 491520 (128×96×5×8).

2.1.3Local Gabor Filter Bank

It can be seen that the Gabor representations are very similar using the filters with the same orientation, especially using the filters with the two neighboring frequencies, such as the first column in Fig.4. It is found that the Gabor feature vector with all the 40 filters becomes very redundant and correlative. For the global Gabor filter bank with all the *m* frequencies and *n* orientations, we denoted it as G(*m*x*n*). In this paper we proposed a novel local filter bank with part of the entire *m* frequencies and *n* orientations, and we denoted it as LG(*m*x*n*). In order to select few Gabor filters to reduce the dimension and computation without degrading the recognition performance, it should cover all the frequencies and orientations, but only select one frequency for each orientation or increase the interval between the neighboring frequencies with the same orientation. Several global and local Gabor filter banks are shown in Fig.5. The method of selecting the LG1(*m*x*n*) is that the parameter *m* of frequency increases repeatedly from min to max, and the parameter *n* of orientation adds one for each time. The difference of LG2(*m*x*n*) is that the parameter *m* of frequency decreases from max to min. For LG3(*m*x*n*), it is selected with an interval between any two filters.

2.2 PCA (Principal Component Analysis)

Principal Component Analysis is a standard technique used in statistical pattern recognition and signal processing for data reduction and Feature extraction. Principal Component Analysis (PCA) is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically. This is the case when there is a strong correlation between observed variables. The first principal component is the linear combination of the original dimensions that has the maximum variance; the n-th principal component is the linear combination with the highest variance, subject to being orthogonal to the n -1 first principal components.For example, face image from the database with size 112×92 can be considered as a vector of dimension 10,304, or equivalently a point in a 10,304 dimensional space. An ensemble of images maps to a collection of points in this huge space. Images of faces, being similar in overall configuration, will not be randomly distributed in this huge image space and thus can be described by a relatively low dimensional subspace. The main idea of the principle component is to find the vectors that best account for the distribution of face images within the entire image space. These vectors define the subspace of face images, which we call “face space”.

PCA is an information theory approach of coding and decoding face images may give insight into the information content of face images, emphasizing the significant local and global "features". Such features may or may not be directly related to face features such as eyes, nose, lips, and hair.In the language of information theory, we want to extract the relevant information in a face image, encode it as efficiently as possible, and compare one face encoding with a database of models encoded similarly. A simple approach to extracting the information contained in an image of face is to somehow capture the variation in a collection of images, independent of any judgment of features, and use this information to encode and compare individual face images.

These eigenvectors can be thought of as a set of features that together characterize the variation between face images. Each image location contributes more or less of each Eigen vector, so that we can display the eigenvector as a sort of ghostly face which we call an Eigen face. Each individual face can be represented exactly in terms of a linear combination of the Eigen faces. Each face can also be approximated using only the "best" Eigen faces-those that have the largest Eigen values and which therefore account for the most variance within the set of face images. The best M Eigen faces span an M-Dimensional subspace- "face space" – of all possible images.

This approach of expression detection involves the following initialization operations:

- Acquire the intial set of face images (the training set).
- Calculate the eigen faces from the training set, keeping only the M images that correspond to the highest eigen values. These M images define the face space. As new faces are experienced; the eigenfaces can be up-dated or recalculated.
- Calculate the corresponding distribution in M-dimensional weight space for each knownindividual, by projecting his or her face images onto the "face space".

2.3 SURF (Speeded up Robust Feature)

Standard SURF procedure can roughly be divided to three parts: interest point detection, interest point description, and interest point matching. The details of interest point detection and description are described as follows.

2.3.1 Interest point detector and descriptor:

The detector of standard SURF is based on the approximate Hessian matrix. The determinant of the approximate Hessian matrix can represent the blob response at that location and scale of an image. For each point of an image in scale space, if the blob response at this location and scale is a local maximum, this point is denoted as the interest point. For constructing the SURF descriptors, a square region centered on the interest point is extracted and aligned to the dominate orientation. Then the region is spilt up equally into 4 smaller square sub-regions. The wavelet responses within each sub-region are computed and summed up to describe the feature of the interest point. In this scheme, the descriptor vector of each interest point has 128 elements. For face recognition, rotation-invariance is often not necessary, so the upright version of SURF is applied. For the upright version, the SURF descriptors use the reference implementation.

2.3.2 Local descriptor projection:

In the proposed scheme, the dimension of the feature vector is chosen to be 128 because the 64D vector is not good for ending the inline-points which is used in the classification stage. To improve computation efficiency, PCA is applied to modelling the identity of interest points and reducing the dimension of the feature. First, the Eigen space must be built with a training set. The training set is used for collecting the SURF feature vectors. Then, PCA is applied to the scatter matrix of these feature vectors and estimates the projection matrix. The projection matrix is used for projecting the feature vectors to the new feature space as the PCA-SURF descriptors. The process of projection not only reserves the identity of the interest points but also discards un modelled distortions. Besides, the computational complexity is greatly reduced because of the low dimensionality of new feature space. The Fowchart of feature vectors' projection is shown as Figure 2.

2.4 LBP (Local Binary Pattern)

LBP features were originally proposed for texture analysis, which have been recently used in face recognition and facial expression recognition due to its low computation and high discrimination capability. The original BP operator labels the pixel of an image by thresholding the 3X3 neighbourhood of each pixel with the value of the central pixel, and a binary value is assigned to neighbourhood pixel on basis of the following function.

$$f(nh) = \begin{cases} 1, if \ v(nh) \geq v(c) \\ 0 \qquad otherwise \end{cases}$$

Where ? (nh) is gray- scale value of the neighbourhood pixel and ? (c) is gray- scale value of the centre pixel. These neighbourhood bits form a Local Binary Pattern (LBP) corresponding to central pixel. This can be understood from an example. Suppose the values of a pixel and its eight neighbours are as follows.

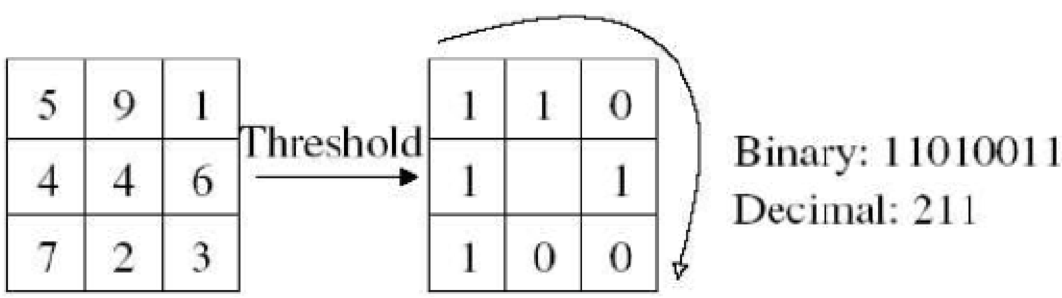


Fig 5. Local Binary Pattern

The derived numbers (called LBP) represent different local patterns like edges, curves, flat regions and spots etc. Using LBP operator the whole image can be transformed to LBP image. An example of LBP image of a facial image is shown in figure 6

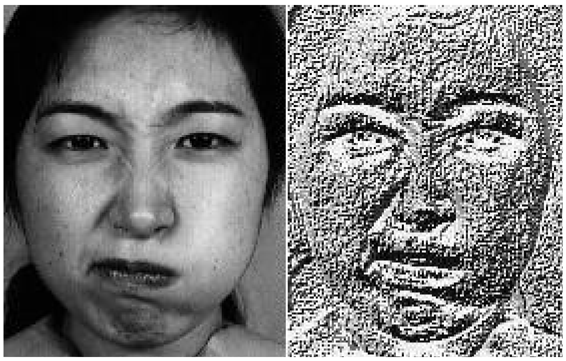


Fig 6. Example of LBP facial Image

The notation LBPP,R is used to denotes the Extended LBP with P pixels and R radius. It has been shown that certain patterns contain more information than others. Therefore it is advantageous to use only those patterns which contain more information. Ojala et el called these patterns as uniform patterns. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions either from 0 to 1 or from 1 to 0. For example, 01111111 is a uniform pattern but 10001101 is not, as it has three bitwise transitions. It is observed that about 90% of the patterns in (8, 1) neighborhood and about 70% of the patterns in (16, 2) neighbourhood are

uniform patterns in texture images [12]. The LBP operator that accumulates only uniform patterns is denoted by LBPU2 P,R. The number of patterns for LBPU2

8,1 is only 59 as compared to number of patterns for standard LBP8,1 , which is 256.

After applying LBP operator to each pixel of an image, the Histogram of LBP operator values is formed as follows.

? ?= ?, (? ? ? ?, ? = ?) ?= 0,1,…. . ? – 1

Where n is the different possible values (Labels) produced by the LBP operator, and

? ? = 1 ?? ? ?? ? ? ? ? or 0 ?? ? ?? ? ? ???

Chapter 3

RESEARCH METHODOLOGY

In this chapter we explain the research methodology followed in this study. We describe the phases involved in the research methodology.

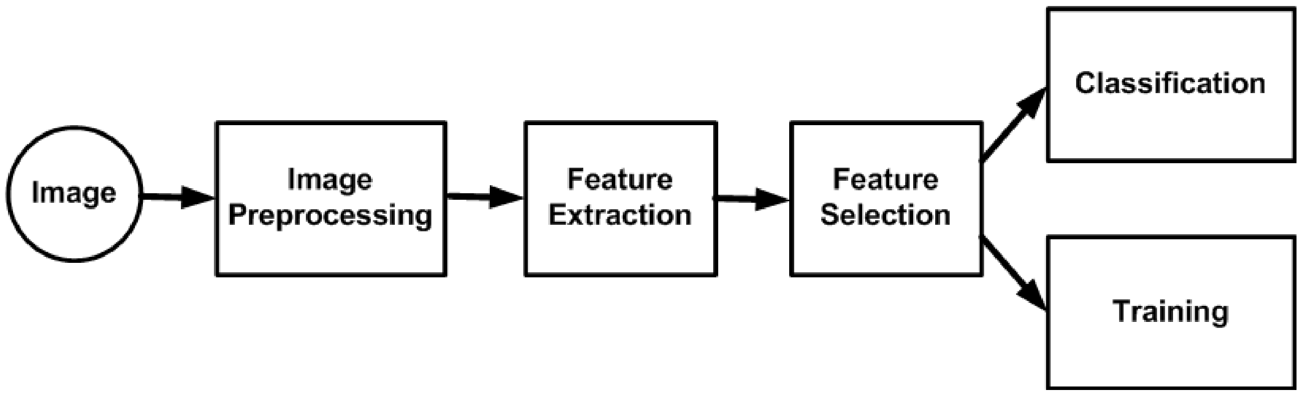


Fig4.1 Basic Module of Face Expression Recognition System

3.1 Image Processing

The general procedure of face recognition can be roughly divided to two parts: feature extraction, and feature classification. In the proposed scheme of face recognition, the algorithm of feature extraction bases on Speeded-Up Robust Features and PCA to establish the local descriptors. In the feature classification stage, the *K*-means algorithm is applied to clustering the local descriptors, and then the local and global similarities are combined to classify the face images. The flowchart of the proposed scheme is illustrated in Figure 1 and three main stages of the proposed scheme (feature extraction, feature clustering, and the classification stage)are briefly described in Section 4.2, Section 4.3, and Section 4.4, respectively.

3.2 Feature Extraction

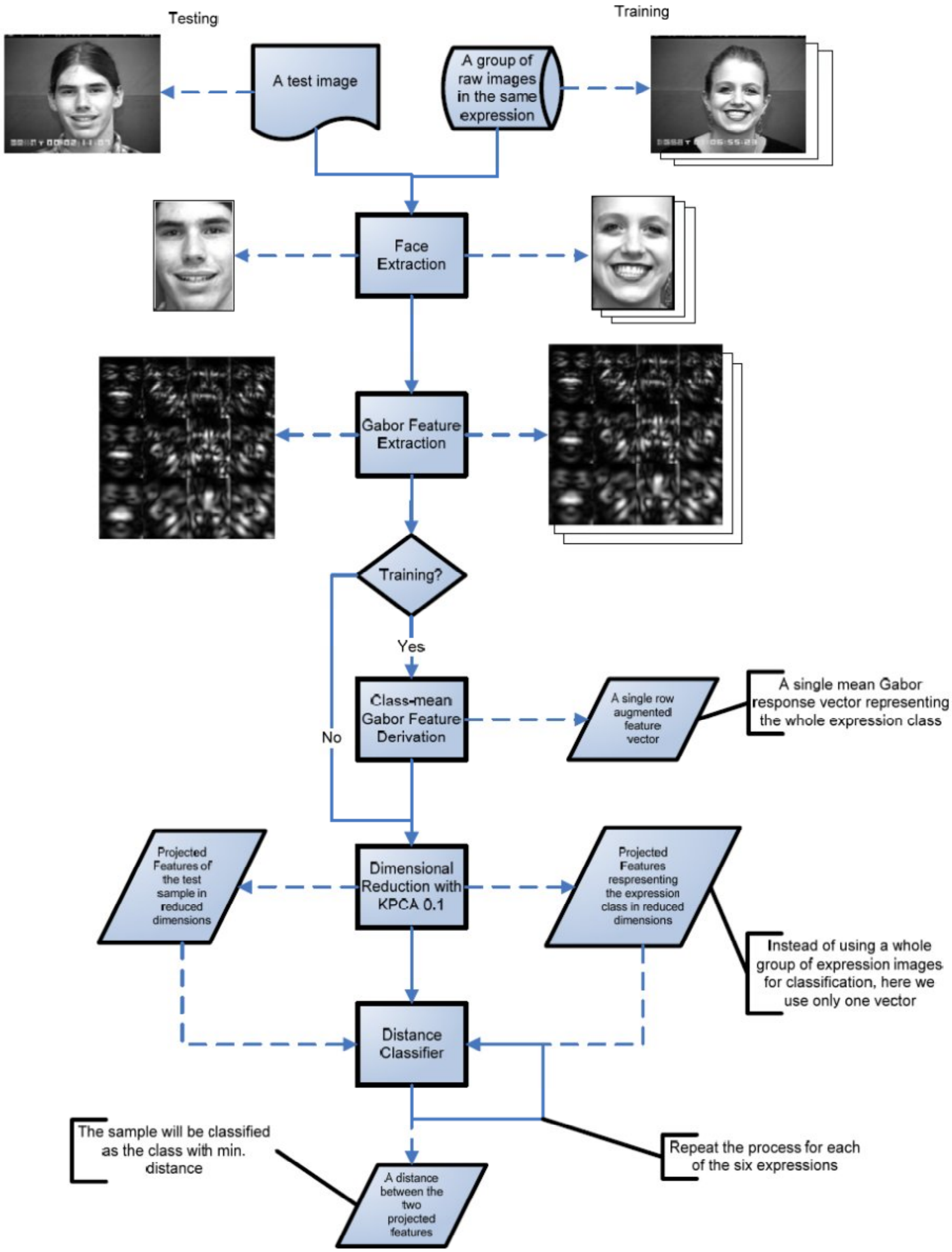


Fig 4.2.2 Feature Extraction using gabor

Gabor filters have been successfully employed in various pattern recognition analyses [34] [14] [8] [11] [35] [16]. Particularly in face and facial expression recognition [24] [36] [37] [7] [15], Gabor filters are used extensively for facial feature extraction. A set of Gabor filters with different frequencies and orientation are used to recover various frequencies and orientations of image textures. The feature outputs from Gabor filters encode the spatial frequencies as well as the orientation activities of various textures in an image. [38] [39] [40].

In most cases, Gabor wavelets of five different scales, $\left\{\frac{\pi}{2}, \frac{\pi}{2\sqrt{2}}, \frac{\pi}{4}, \frac{\pi}{2\sqrt{2}^3}, \frac{\pi}{8}\right\}$, and eight orientations, $\left\{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}, \frac{\pi}{2}, \frac{5\pi}{8}, \frac{3\pi}{4}, \frac{7\pi}{8}\right\}$, are used for face or facial expression recognition.

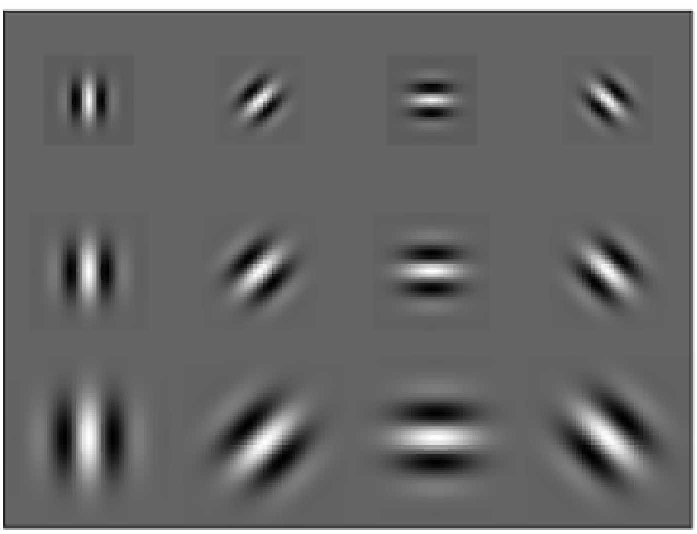


*Figure 4.2.3 S*shows 12 Gabor basis response images in three scales, $\left\{\frac{\pi}{2}, \frac{\pi}{4}, \frac{\pi}{8}\right\}$,and four orientations, $\left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ that were used for extracting facial expression features for the system developed in this thesis.

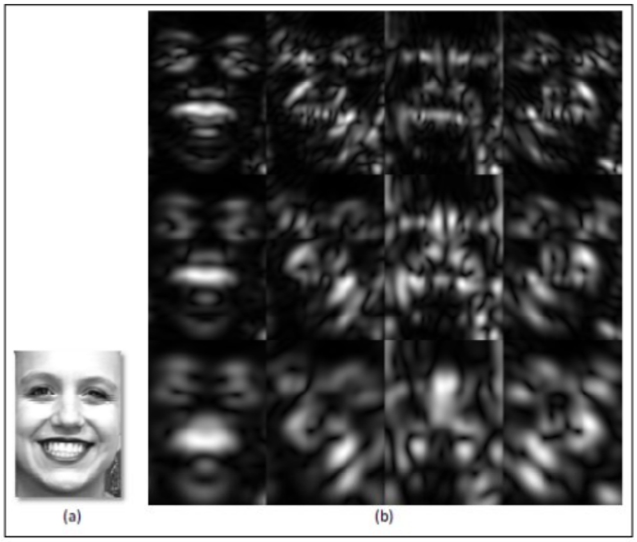


Fig 4.2.4 (a) The original captured face from Cohn-Kanade database. (b) The Gabor magnitude responses of the captured face in (a).

The Gabor representation of an image is a result of the convolution of the image with a bank of Gabor filters as defined by Equation (2.2.3). The convolution of an image I(x,y) and a Gabor kernel function is defined as $\varphi_{\theta_\mu,k_v}(x,y)$

$$P_{\theta_\mu,k_v}(x,y) = I(x,y) * \varphi_{\theta_\mu,k_v}(x,y)$$

where denotes the convolution operator and $P_{\theta_\mu,k_v}(x,y)$ is the convolution result corresponding to the Gabor kernel at orientation and scale. Therefore, the Gabor wavelet representation of the image is formed as a

set Z = $\{P_{\theta_\mu,k_v}(x,y) : \theta_\mu \in$

3.3 Feature Clustering

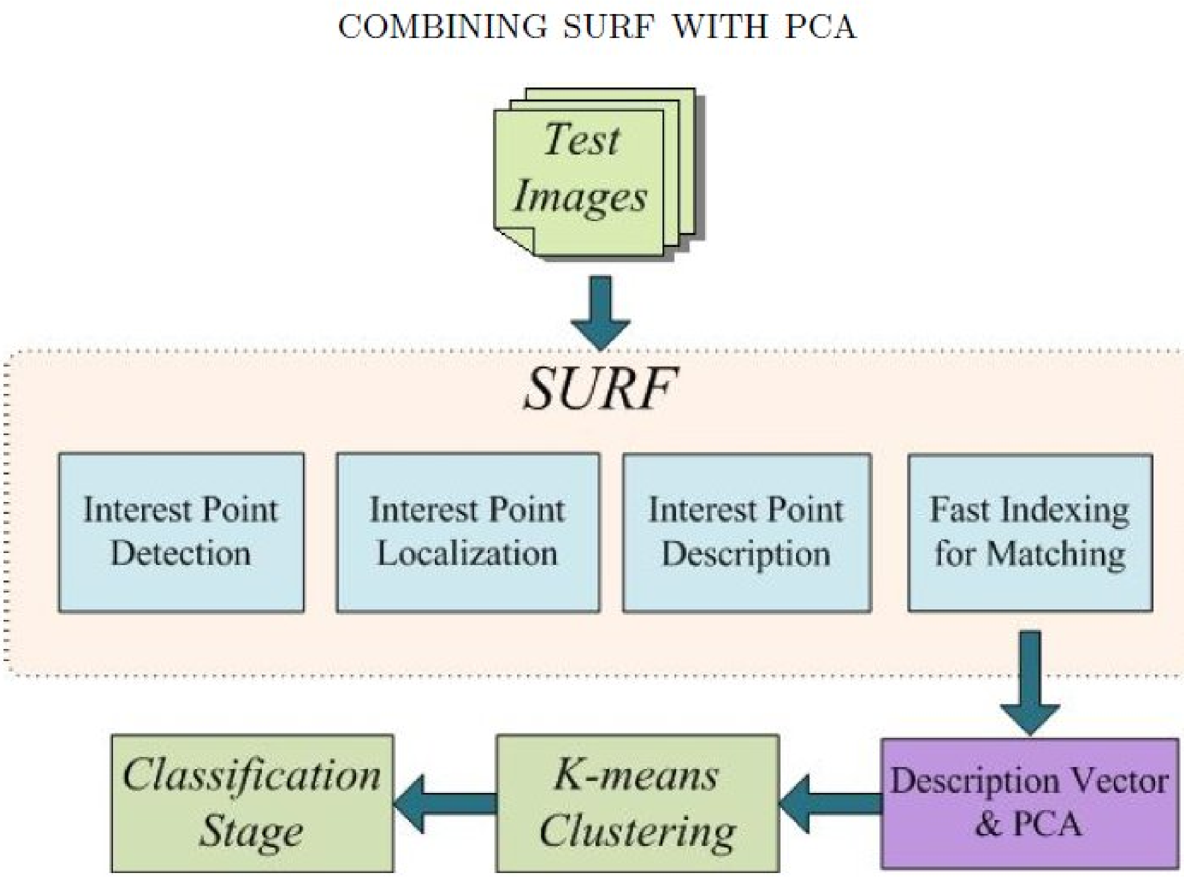


Fig 3.3.1 The complete flowchart of the proposed face recognition scheme

3.3.1 The proposed PCA-based SURF descriptor. Standard SURF procedure [14] can roughly be divided to three parts: interest point detection, interest point description, and interest point matching. The details of interest point detection and description are described as follows. After that, the PCA projection matrix is built and the descriptors are projected to a new space with lower dimension and further utilized to represent the feature of an image.

The detector of standard SURF is based on the approximate Hessian matrix. The determinant of the approximate Hessian matrix can represent the blob response at that location and scale of an image. For each point of an image in scale space, if the blob response at this location and scale is a local maximum, this point is denoted as the interest point. For constructing the SURF descriptors, a square region centred on the interest point is extracted and aligned to the dominate orientation. Then the region is spilt up equally into 4 smaller square sub-regions. The wavelet responses within each sub-region are computed and summed up to describe the feature of the interest point.

In this scheme, the descriptor vector of each interest point has 128 elements. For face recognition, rotation-invariance is often not necessary, so the upright version of SURF is applied. For the upright version, the SURF descriptors use the reference implementation.

In the proposed scheme, the dimension of the feature vector is chosen to be 128 because the 64D vector is not good for finding the inline-points which is used in the classification stage. To improve computation efficiency, PCA is applied to modeling the identity of interest points and reducing the dimension of the feature. First, the eigenspace must be built with a training set. The training set is used for collecting the SURF feature vectors. Then, PCA is applied to the scatter matrix of these feature vectors and estimates the projection matrix.

The projection matrix is used for projecting the feature vectors to the new feature space as the PCA-SURF descriptors. The process of projection not only reserves the identity of the interest points but also discards unmodeled distortions. Besides, the computational complexity is greatly reduced because of the low dimensionality of new feature space. The Flowchart of feature vectors' projection is shown.

The features of the interest points cannot be compared only based on their locations because the positions and number of features are different in the images. The situation of a feature around the eye matching to a feature around the mouth corner will occur if this feature of an image is directly compared with the whole features of another image. Therefore, we divide the image into several sub-regions by *K*-means algorithm and compare the features of each sub-region separately.

K-means clustering algorithm based on the positions of PCA-SURF features is used. Figure shows the flowchart of feature clustering.

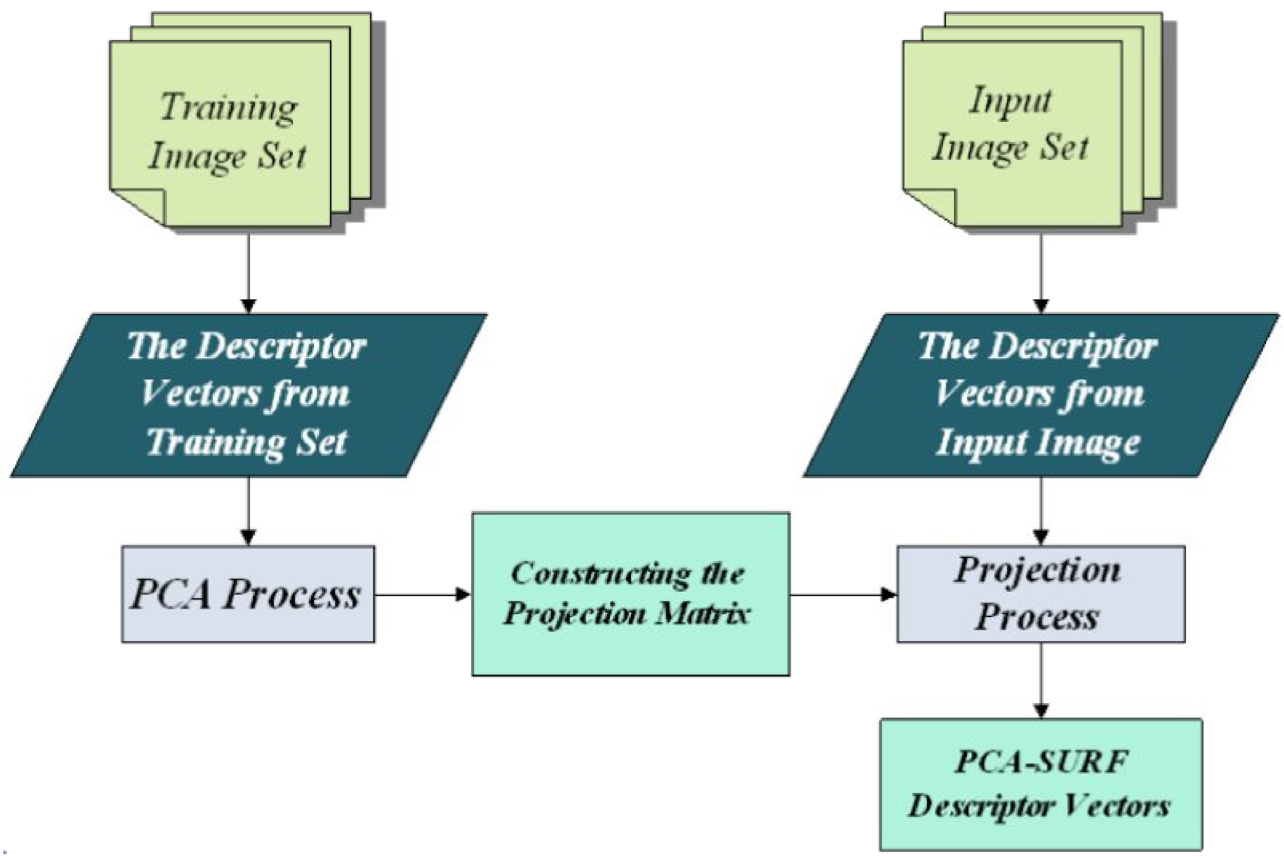


Fig 3.3.2 Flow Chart for Feature Vector Projection

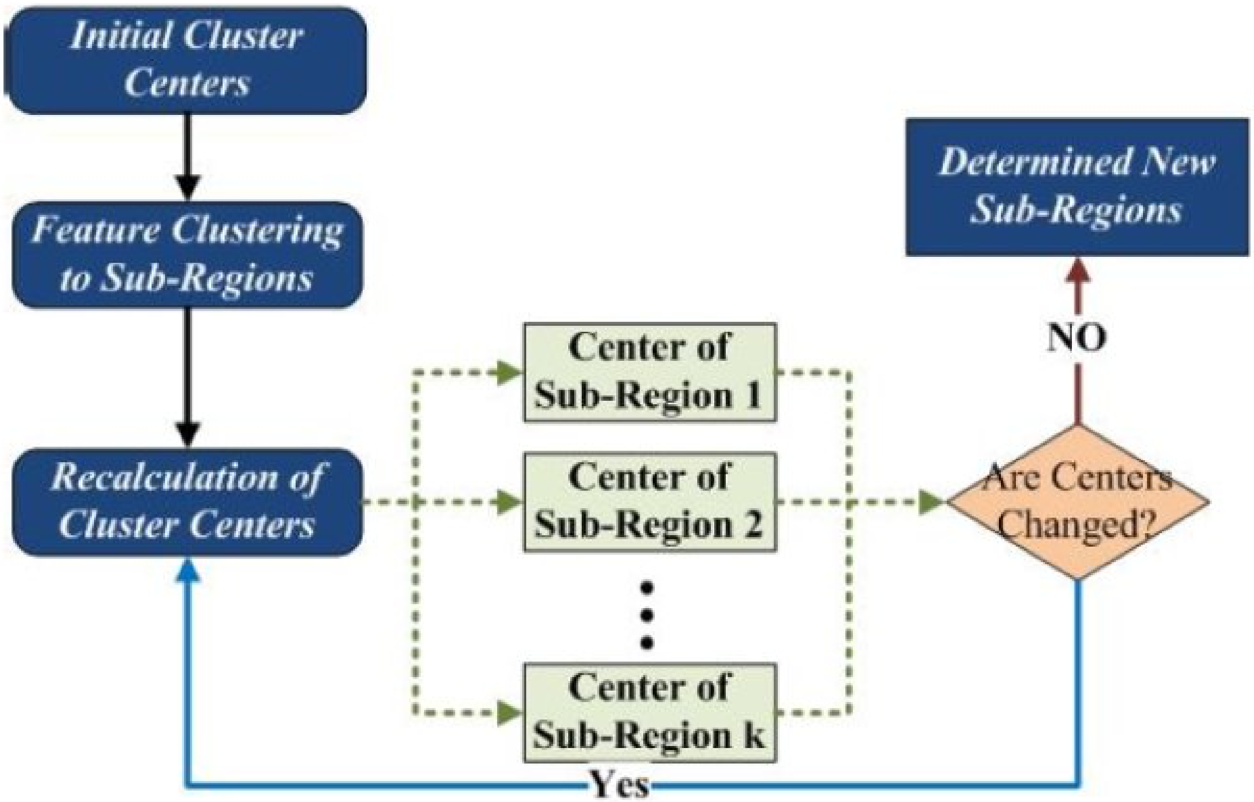


Fig 3.3.3 Flow Chart for Clustering

COMBINING SURF WITH PCA

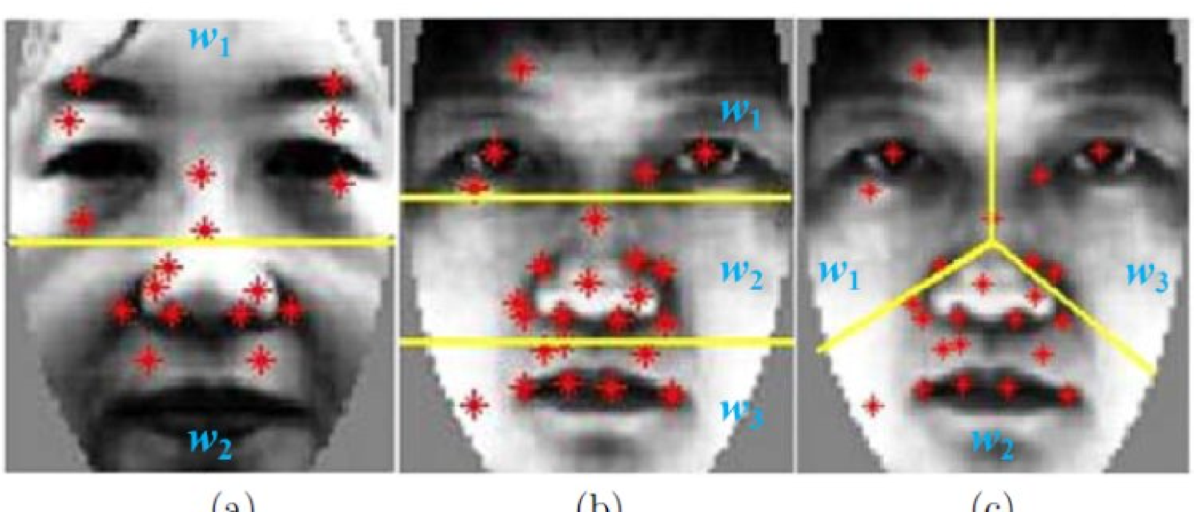


Fig 3.3.4 Face images of different clustering sub-regions. (a) The feature points are divided by the horizontal line. (b) The feature points are clustered to three equal parts. (c) The feature points are clustered and centered on left eye, right eye, and the center of mouth.

sub-regions of face images are determined. When a test image is inputted to the Classification stage, feature points of the test image are classified to the sub-regions based on the cluster centres. In this article, three different approaches are proposed in the feature's clustering algorithm. The sample face images of different clustering approaches are illustrated.

3.4 The classification stage.

The proposed classifying strategy is similar to with a different global similarity and different situations of clustering. As the sub-regions are determined, local similarity and global similarity can be calculated and integrated to classify the face images. However, the complexity of computing similarity is very high. To improve this disadvantage, the fast indexing for matching method [14] is used for filtering the interest points with extreme difference. The flowchart of similarity computation, and the local similarity and global similarity are introduced in *Local similarity and global similarity.* We assume that the feature points of the input image *I* are located in *k* sub-regions and denoted as (1).

$$I = \left(f_1^1, \ldots, f_1^{m1}, f_2^1, \ldots, f_2^{m2}, \ldots, f_k^1, \ldots, f_k^{mk}\right)$$

where the *fjk* means the *j*th feature descriptor in the *k*th sub-region of image *I*.

For each feature in the sub-region of *It*, the _rst _lter uses the fast index for matching method [14] to eliminate the extreme different features and then preserve the useful and similar features in the same sub-region of *Ir* and *It*. Then the local similarity *SL* between a test image *It* and a reference image *Ir* is calculated as the following steps:

Step1. Collect the interest points in each sub-region, and then compute the correlation between each pair of features in *i*th sub-regions of the test and reference image .

$$d(f_{ti}^x, f_{ri}^y) = \frac{(f_{ti}^x, f_{ri}^y)}{\|f_{ti}^x\| \cdot \|f_{ri}^y\|}$$

Step2. Choose the maximal similarity in *i*th sub-region as *Si* by (3). The Steps 1 and 2are illustrated

$$S_i(I_t, I_r) = \max\left(d(f_{ti}^1, f_{ri}^1), \ldots, d(f_{ti}^1, f_{ri}^y), \ldots, d(f_{ti}^x, f_{ri}^y)\right)$$

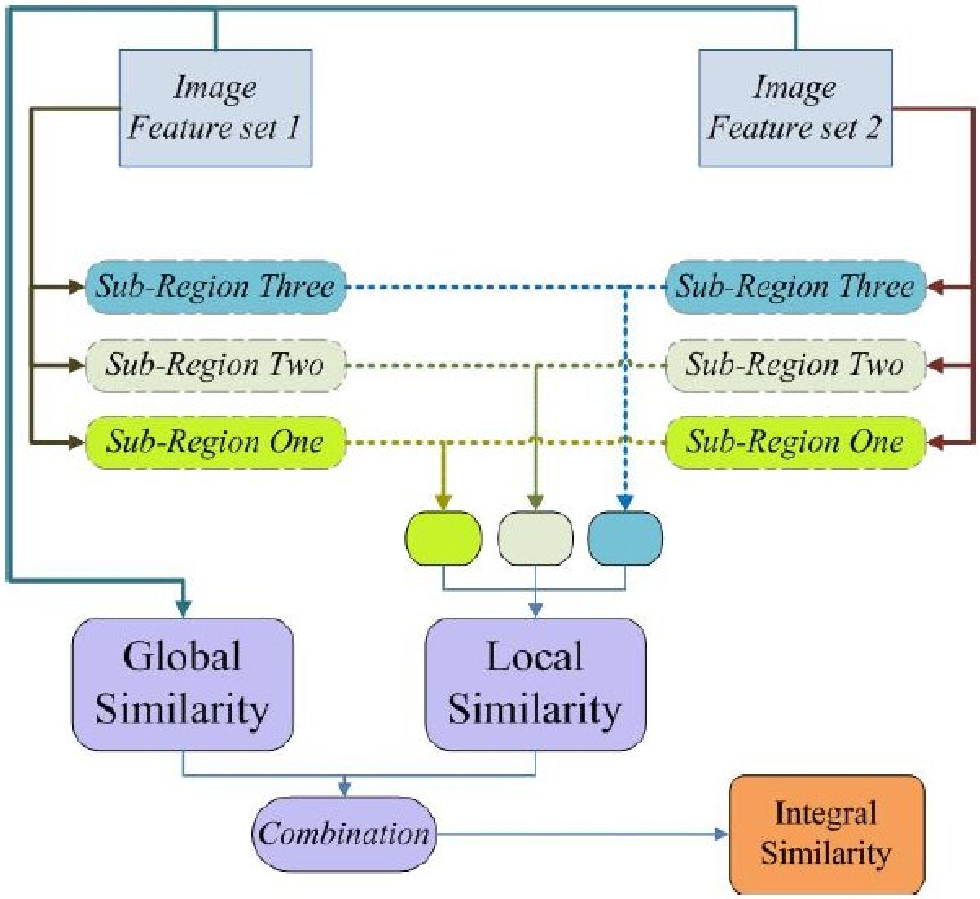


Fig 3.4.1 Flow Chart for similarity Computation

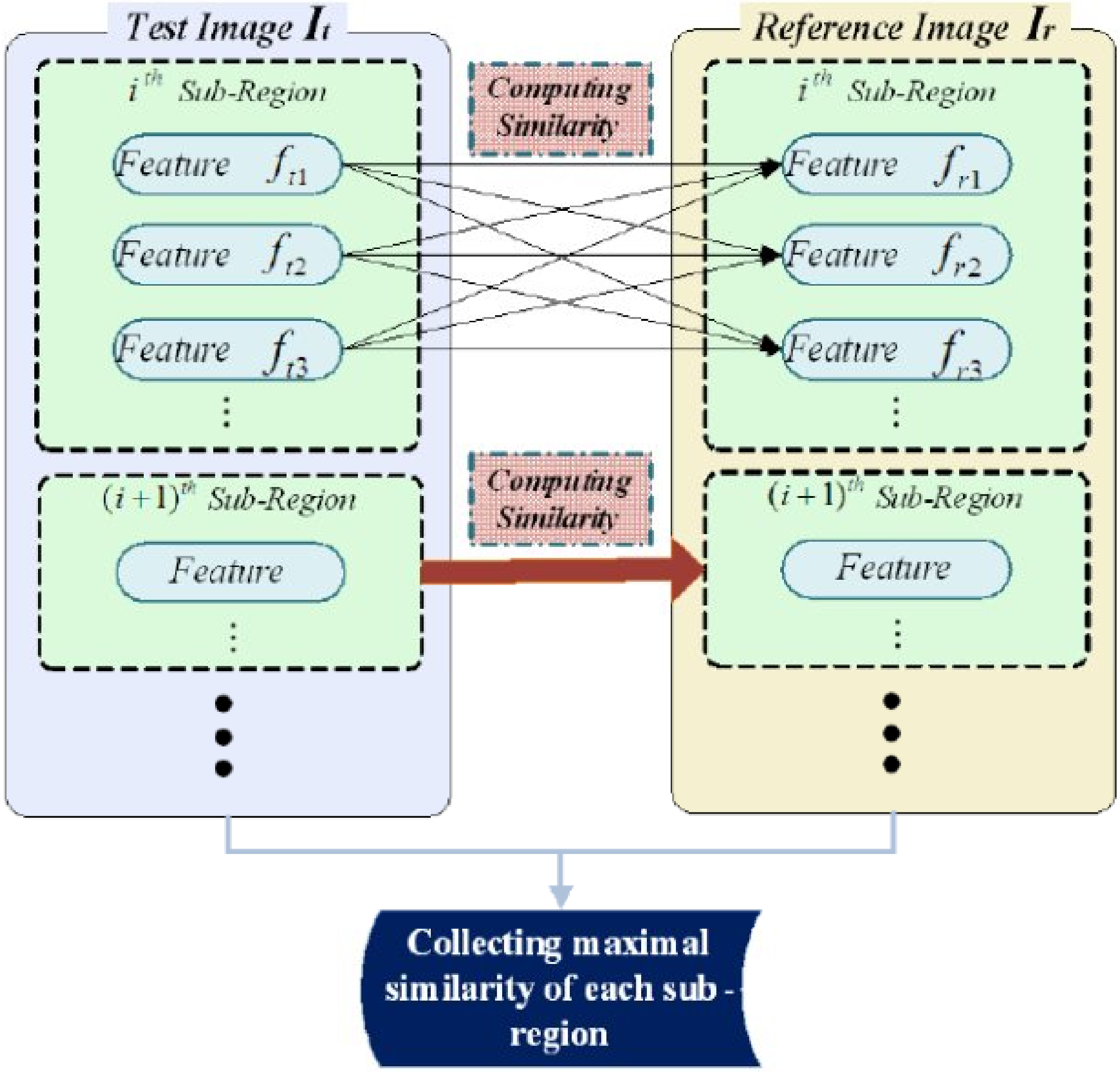


Fig 3.4.2 Computation of similarities in ith sub-region

Step 3. After collecting the maximal similarity from each sub-region of the test and reference image, the local similarity is computed as the average that *Si* multiplies the weight of *i*th sub-region, *wi*, by

$$S_L(I_t, I_r) = \frac{1}{k} \sum_{i=1}^{k} (S_i(I_t, I_r) \times w_i)$$

In steps 1 and 2, *Si* is chosen as the maximum of the similarities rather than the average. Because most matching results are invalid, the similarities often have a small

value such that the average of similarities is reduced. Therefore, choosing the average of the similarities as *Si* may induce the poor discrimination between the sub-regions of *It* and *Ir*. In this article, the global similarity is calculated by combining the inline-point *Sip* and cosine correlation *Sc* which are introduced as follows.

The inline-point similarity is computed as

$$S_{ip} = match(I_t, I_r)$$

where *match*(*It; Ir*) is the number of validly matched features of two images. The feature matching method is the same as [9] and described in the following:

Step1. For each feature *fx t* of image *It*, estimate the Euclidean distances at all features in image *Ir*.

Step2. If the ratio between the nearest and the second nearest distance is less than a threshold, *fx t* is determined to be correctly matched.

The distance ratio is utilized to determine whether the feature is matched because the nearest distance is much shorter than other distances if the feature is correctly matched. The cosine similarity is similar to the local similarity; however, it is not limited to sub regions. The correlations of the matched features between *It* and *Ir* are first computed . Then, the cosine correlation is computed by

$$S_C = \max\left(d(f_t^1, f_r^1), \ldots, d(f_t^1, f_r^y), \ldots, d(f_t^x, f_r^y)\right)$$

where SC means the maximal cosine correlation, and $d(f1t\ ;\ f1r\ )$ is the correlation of the feature between It and Ir. The complete global similarity is the combination of the $Sip$and SC by

$$S_G = S_{ip} \times S_C$$

Finally, the local similarity and global similarity are integrated to avoid wrong classifications caused by the situation that only some local regions of two subjects are very similar. And the final similarity $Sall$ is computed by (and used for face recognition.

$Sall = SG \; x \; SL$

Chapter 4

RESULTS

In oder to evaluate the results the proposed work is executed on Matlab. We have used Data SET of JAFEE (Japanese Female Facial Expression) JAFFE is a very popular database for facial expression recognition, in which in total 213 facial expression images with 10 Japanese women are involved. Each individual has three or four images with seven kinds of facial expressions, including anger, disgust, fear, happy, sadness, surprise and neutral,. Figure 5.1 shows seven expression image examples selected from JAFFE database.

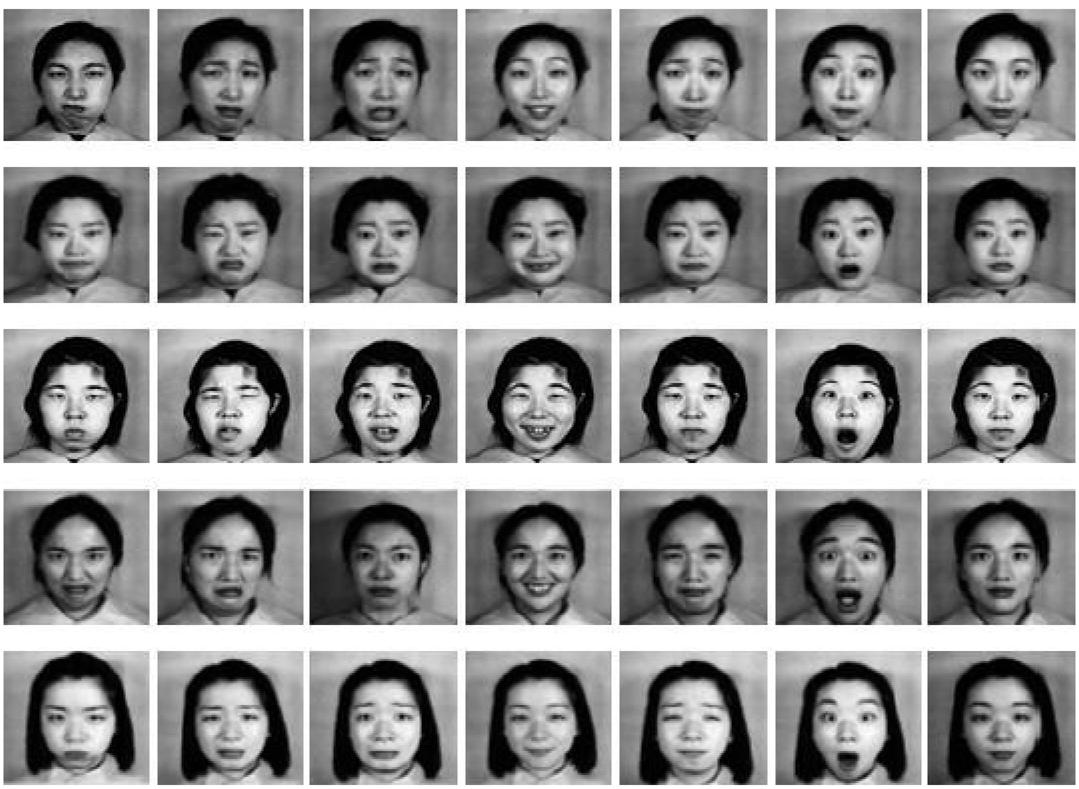


Fig 4.1 Some Sample Images from JAFFE Dataset.

In this section, we perform person-independent as well as person dependent facial expression recognition using LBP features along with template matching as a classifier. Template matching was used in to perform face recognition using LBP-based features: a template is formed for each class of face, and then the nearest neighbor classifier is used to match the test image with the closest template. Here we adopt the template matching for classification of facial expressions.

We adopt two types of template matching, one is the person independent template matching and the other is person dependent.

4.1 Person Independent Template Matching

A template is formed for each class of face expression by averaging the LBP histograms of a particular expression. In the training phase, these seven templates are stored. In the testing phase, a test image is compared with all stored templates. We have selected the Chi square test (X2) as similarity measure.

$$X^2(\mathbf{S}, \mathbf{M}) = \Sigma_i \, (Si-Mi)^2 / (Si+Mi)$$

Where **S** and **M** are two LBP histograms of template and test images respectively. Person independent template matching has achieved the generalization performance of 73.61% for 7 category classification.

	Anger %	Disgust %	Fear %	Happy %	Neutral %	Sad %	Surprise %
Anger	60	10	0	10	10	10	0
Disgust	0	50	10	20	10	10	0
Fear	0	0	54.5	9.1	9.1	9.1	18.2
Happy	0	0	0	91	9	0	0
Neutral	0	0	0	0	90	10	0
Sad	0	0	10	0	10	80	0
Surprise	0	0	10	0	10	0	80

Table1: *Matrix for 7-class Person Independent Facial Expression Recognition.*
Note: Happy, neutral, sad and surprise expressions can be recognized with high accuracy about 80-90%), but anger, fear and disgust are easily confused with other expressions.

4.2 Person Dependent Template Matching

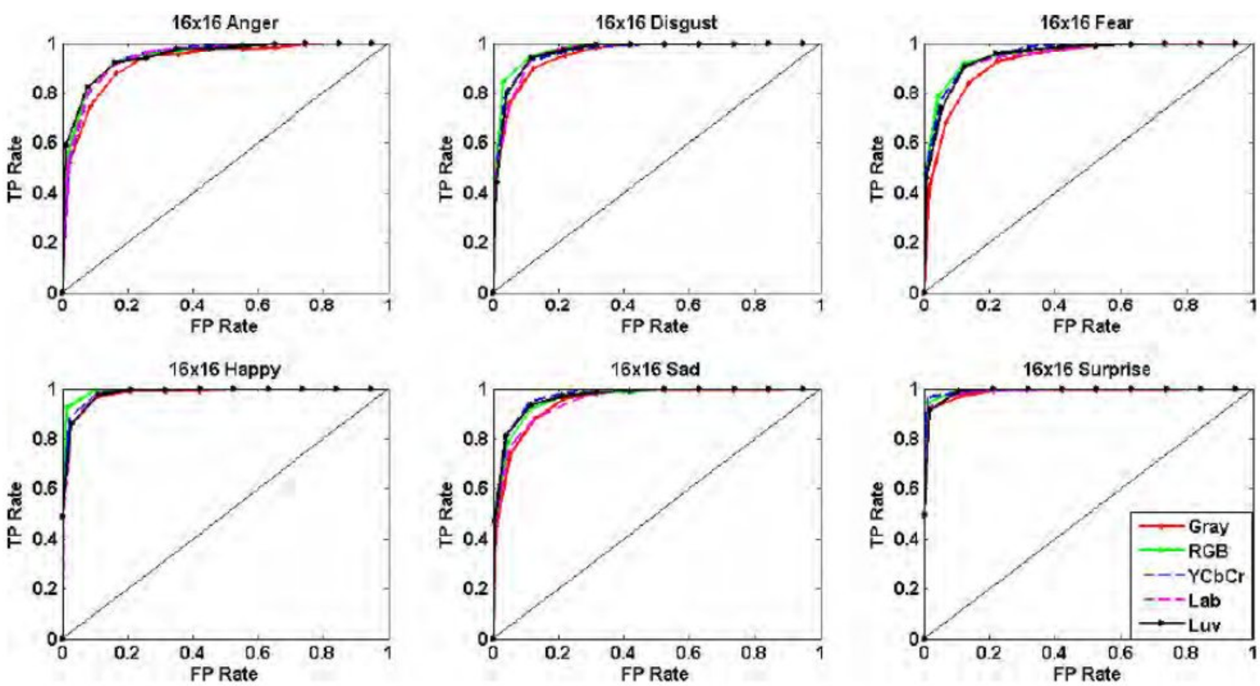
Facial expressions may be expressed differently by different people [15], so low accuracy is achieved in the above method. Therefore, we propose a method that is person dependent. Instead of developing person independent seven templates- one for each expression we propose to form templates that are person dependent. For each person, seven templates are formed- one for each expression- so total of 70 templates for 10 persons are formed. These 70 templates are stored in training phase. In the testing (Recognition) phase a test image is compared with all stored templates. The one with minimum distance is declared as a recognized expression. The person dependent template matching has achieved a very high generalization performance of 94.44% for 7 category classification. The Confusion matrix for person dependent facial expression recognition is shown in Table2.

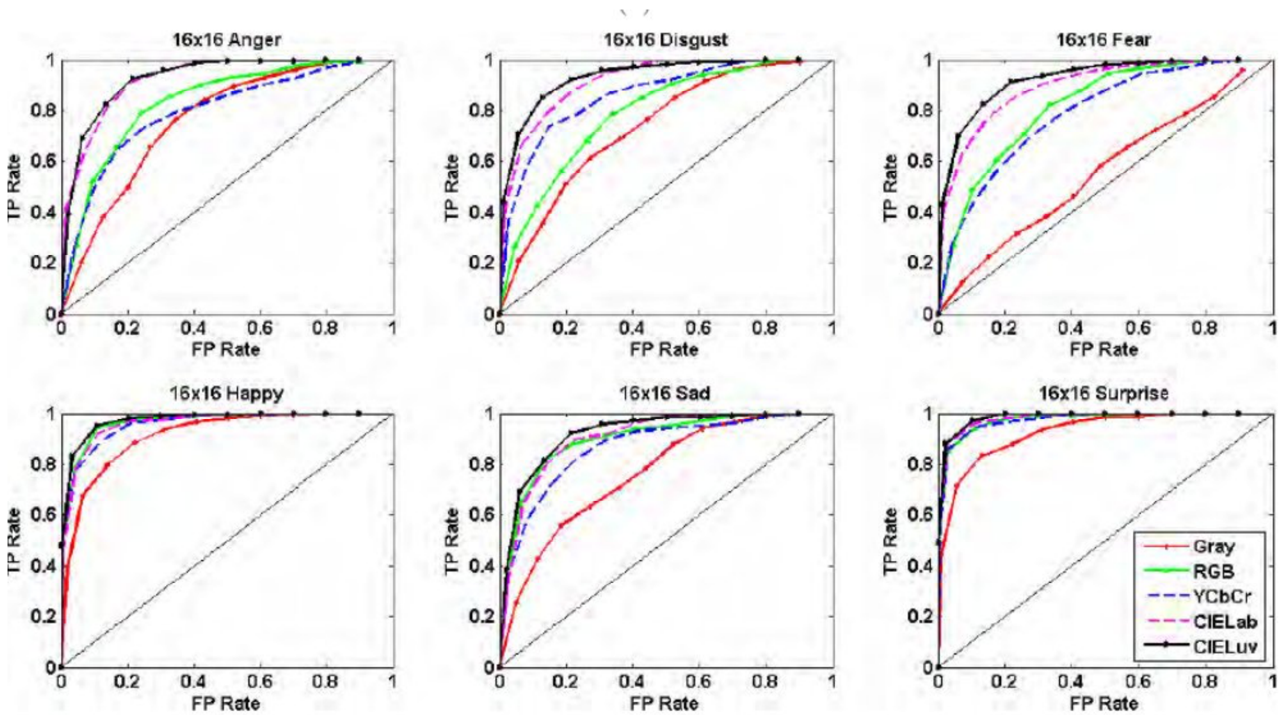
	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	90	0	0	0	0	10	0
Disgust	0	100	0	0	0	0	0
Fear	0	0	90	0	0	10	0
Joy	0	0	0	100	0	0	0
Sad	0	0	0	0	100	0	0
Surprise	0	0	10	0	0	90	0
Neutral	0	0	0	0	0	10	90

Table2: Matrix for 7-class Person Dependent Facial Expression *Recognition*

Another benefit of this approach is that along with the recognition of an expression, it also recognizes with expression belongs to which particular person. The time taken to recognize an expression with our approach is, on an average, 11.5 milli seconds.

Person dependent template matching has achieved an average performance of 94.44% for JAFFE database, and has outperformed other methods as listed





(b)

Fig4.3 ROC Cures of Different Texture of different Expression a) Original image b) Image after preprocessing

Chapter 5

CONCLUSION AND FUTURE WORK

In this paper, we have extracted features based on Local Binary Pattern. As every part of the face does not contribute equally in face expression recognition, we have chosen some important facial parts like sub parts of eyes, nose and mouth. With the templates of extracted facial features, clustering mechanism was used to classify the expression. Experimental results show that the proposed approach is better than approaches that use the whole face image. The proposed method integrates person identity to perform better than conventional expression systems. The proposed person dependent approach achieves higher recognition rates than those of other approaches.

1). Local Gabor filter bank outperforms global Gabor filter bank in the aspects of shortening the time for feature extraction, reducing the high dimensional feature, decreasing the required computation and storage, even achieving better performance in some situations.

2). PCA can significantly reduce the dimensionality of the original feature without loss of much information in the sense of representation, but it may lose important information for discrimination between different classes. When using PCA feature to classify, the L1 distance measure performs better than L2. Illumination normalization is effective for PCA feature to achieve high performance.

3). When using PCA+SURF method, the dimensionality drastically reduced to 6 dimensions and the recognition performance is improved several percent compared with PCA. Experiments show that PCA+SURF feature may partly eliminate the sensitivity of illumination.

4). *K*-means algorithm clusters the features in the face image. Finally, local and global similarities are computed and combined to classify face images. Simulation results show that the performance of the proposed scheme is better than other methods. More precisely, PCA-based SURF local descriptors are more robust than original SURF and SIFT local descriptors to the accessory, expression, and pose variations.

5). Discarding the first one to three PCs will reach to the best performance rates. If removing more than three PCs, it will commonly worsen the results. The best performance, using LG3 (3x8) and PCA+SURF feature by eliminating the first two PCs, was 97.33% in our experiments.

REFRENCES

[1] C. Liu, H. Wechsler, “Independent Component Analysis of Gabor Features for Face recognition”, IEEE Trans. Neural Networks, Vol. 14, 2003, pp.919-928

[2] J. G. Daugman, “Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression”, IEEE Trans. Acoustic, speech and signal processing, Vol. 36, 1988, pp.1169-1179

[3] T. S. Lee, “Image Representation Using 2D Gabor Wavelets”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 18, 1996, pp.959-971

[4] G. Lei, X. H. Li, J. L. Zhou and X. G. Gong, “Geometric feature based facial expression recognition using multiclass support vector machines,” IEEE International Conference on Granular Computing, 2009, GRC '09, pp. 318-321, Aug. 2009.

[5] Zhang B,Shan S,Chen X,Gao W.Histogram of Gabor phase pattern(HGPP):a novel object representation approach for face recognition. IEEE Trans. Image Process.2007,16(1),57- 68.

[7] Praseeda Lekshmi.V, Dr.M.Sasikumar. Analysis of Facial Expression using Gabor and SVM[J].International Journal of Recent Trends in Engineering,Vo1.1,No.2,May 2009.

[8] W. Zhang, S. Shan, X. Chen, and W. Gao. Are Gabor phases really useless for face recognition?. in Proc. Int. Conf. Pattern Recognition,2006,pp.606–609.

[9] X. Tan and B. Triggs, Fusing Gabor and LBP feature sets for kernel based face recognition. Presented at the IEEE Int.Workshop on Analysis and Modeling of Faces and Gestures, 2007.

[10] Michael Lyons,Shigeru Akamatsu.Coding Facial Expressions with Gabor Wavelets. Third IEEE International Conference on Automatic Face and Gesture Recognition,April 14-16 1998, Nara Japan, IEEE Computer Society,pp.200-205.

[11] J. F. Ye, Y. Z. Zhan and S. L. Song, “Facial expression features extraction based on Gabor wavelet transformation” 2004. IEEE International Conference on System, Man and Cybernetics, vol. 3, pp. 2215-2219, Oct. 2004.

[12] Q. Y. Zhao, B. C. Pan, J. J. Pan and Y. Y. Tang, “Facial expression recognition based on fusion of Gabor and LBP features”, 2008. ICWAPR '08. International Conference on Wavelet Analysis and Pattern Recognition, vol. 1, pp. 362-367, Aug. 2008.

[13] A. Koutlas and D. Fotiadis, “A region based methodology for facial expression recognition,” In BIOSIGNALS, vol. 2, pp. 218–223, 2008.

[14] T. Ojala, M. Pietik¨ainen and D. Harwood, “A comparative study of texture measures with classification based on feature distributions”J. PatternRecognition vol. 29, No.1 pp. 51-59,1996.

[15] Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, “Face Recognition with Local Binary Patterns,” M. Lecture Notes in Computer Science, Vol. 3021, pp.469-474,May.2004.

[16] T. Ahonen, A. Hadid, and M. Pietikinen, “Face recognition with local binary patterns,” in *ECCV*, 2004, pp. 469-481.

[17] A. Hadid, M. Pietikinen, and T. Ahonen, “A discriminative feature space for detecting.

[18] M. Bicego, A. Lagorio, E. Grosso and M. Tistarelli, On the use of SIFT features for face authentication, *Proc. of the Conference on Computer Vision and Pattern Recognition Workshop*, p.35, 2006.

[19] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade and B. Lu, Person-speci_c SIFT features for face recognition, *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pp.593-596, 2007.

[20] Y. Ke and R. Sukthankar, PCA-SIFT: A more distinctive representation for local image descriptors, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.2, no.1, pp.506-513, 2004.

[21] K. Mikolajczyk and C. Schmid, A performance evaluation of local descriptors, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.27, no.10, pp.1615-1630, 2005.

[22] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, Speeded-up robust features, *Computer Vision andImage Understanding*, vol.110, no.3, pp.346-359, 2008.