Fusing Facial Features for Face Recognition

Jamal Ahmad Dargham, Ali Chekima, Ervin Gubin Moung

Universiti Malaysia Sabah

Abstract — **Face recognition is an important biometric method because of its potential applications in many fields, such as access control, surveillance, and human-computer interaction. In this paper, a face recognition system that fuses the outputs of three face recognition systems based on Gabor jets is presented. The first system uses the magnitude, the second uses the phase, and the third uses the phase-weighted magnitude of the jets. The jets are generated from facial landmarks selected using three selection methods. It was found out that fusing the facial features gives better recognition rate than either facial feature used individually regardless of the landmark selection method.**

Keywords—**Gabor filter; face recognition; bunch graph; image processing; wavelet**

I. INTRODUCTION

Face recognition approaches can be divided into three groups [2]; global, local, and hybrid approaches. In global based methods the face image is represented as a low dimension vector by being projected into a linear subspace [1][2]. The advantages of global based methods are: their simple applicability, easy computation, and their general function. However, the limitation of global based methods is that they do not detect the differences in faces local regions and as such are not capable of extracting the local or 'topological' structures of the face. In local based approaches, the geometric features such as the position of eyes, nose, mouth, eyebrows, measurements of width of eyes, are used to represent a face [2][3][4]. There are several ways on how to select local features to represent a face, for example; manual feature selection by positioning nodes on fiducially points (e.g, eyes, and nose), and automatic feature selection. Hybrid methods are a combination of global and local approaches.

The bunch graph method is a local approach that works by first locating a landmark on a face, then convolving a sub-image around each landmark with a group of Gabor filters.

This produces a jet from each landmark. These jets will be used for face recognition by computing and comparing similarity scores between jets of two different images. Wiskott et al. introduced a face recognition method called the Elastic Bunch Graph Method [3] and compared the EBGM with several face recognition methods on the FERET and Bochum image databases in different face poses. Their system achieved 98% recognition rate for frontal images. Bolme [4] also used Elastic Bunch Graph Method but he only used one training image per person and the jets were computed from manually selected training images landmarks. These jets were used to find new jet from new image using a displacement estimation method to locate the node on the new image. These new jets are then added to the existing jets database. By using the automatically obtained jets for recognition task an 89.8% recognition rate was reported on the FERET database. Sigari and Fathy [5] proposed a new method for optimizing the EBGM algorithm. Genetic algorithm was used to select the best wavelength of the Gabor wavelet. They had tested the proposed method on the frontal FERET face database and achieved 91% recognition rate. In this paper, a face recognition system that fuse facial features extracted using Gabor wavelet is presented. In section 2 the theory of Gabor wavelet method will be presented while in section 3 the application of bunch graph method to extract facial feature is presented. Section 4 describes the proposed system, while in section 5 the experimental results are discussed before the paper concludes in section 6.

II. GABOR WAVELET TRANSFORM

Gabor wavelet is the fundamental features extraction tool in the bunch graph method. Two dimensional Gabor wavelets shown in (1) were used to extract features from landmarks by convolving the wavelet on the landmarks of the faces. The wavelet has a real and imaginary component representing orthogonal directions. These two parts can be formed into a complex number or used individually. The magnitude and phase of the image content at a particular wavelet's frequency can be computed from the complex number given in (1)

$$g(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \quad (1)$$

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

λ specifies the wavelength of the cosine (or sine) wave. Wavelets with a large wavelength will respond to gradual changes in intensity in the image. Wavelets with short wavelengths will respond to sharp edges and bars.

θ specifies the orientation of the wavelet. This parameter rotates the wavelet about its centre. The orientation of the wavelets dictates the angle of the edges or bars for which the wavelet will respond.

ψ specifies the phase of the sinusoid. Typically, Gabor wavelets are based on a sine or cosine wave. Cosine wavelets are thought to be the real part of the wavelet and the sine wavelets are thought to be the imaginary part of the wavelet. Therefore, a convolution with both phases produces a complex coefficient. The mathematical foundation of the algorithm requires a complex coefficient based on two wavelets that have a phase offset of $\pi/2$.

σ specifies the radius of the Gaussian. The size of the Gaussian is sometimes referred to as the wavelet's basis of support. The Gaussian size determines the amount of the image that effects convolution. In theory, the entire image should effect the convolution; however, as the convolution moves further from the center of the Gaussian, the remaining computation becomes negligible. This parameter is usually proportional to the wavelength, such that wavelets of different size and frequency are scaled versions of each other.

γ specifies the aspect ratio of the Gaussian. Most wavelets tested with the algorithm use an aspect ratio of 1.

The value of the parameters used in this paper are the same as those used by Wiskott in [3], which give 40 Gabor wavelets with different frequencies and orientations.
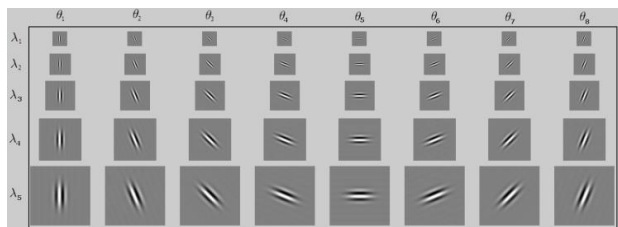

Fig. 1. The real part of the 2D Gabor wavelet mask with different wavelength and orientation.

Convolving the same landmark with many Gabor wavelet configurations produces a collection of Gabor coefficients called jets. Each Gabor coefficient has a real and imaginary component. The magnitude and phase of the image's content at a particular wavelet's frequencies can be computed from the complex number. Let J be a complex number Gabor coefficient, the magnitude. $J_{magnitude}$ and the phase angle ϕ of J are given as in (2) and (3) respectively.

$$J_{magnitude} = \sqrt{J_{real}^2 + J_{imaginary}^2} \tag{2}$$

$$\phi = \cos^{-1}\left(\frac{J_{real}}{J_{magnitude}}\right) = \sin^{-1}\left(\frac{J_{imaginary}}{J_{magnitude}}\right) \tag{3}$$

III. BUNCH GRAPH METHOD

A. Selecting Facial Features

A face image is represented as a bunch graph. A bunch graph is a collection of jets for an image. Fig. 2(a) shows the landmarks that were selected as point of interest to be convolved with a group of Gabor wavelets. An example of a convolution of a Gabor wavelet at the chin of a person is shown in Fig. 2(b). Face images are zero padded for the convolutions where the wavelet exceeds the image dimensions, which normally occur near the edge of the image.
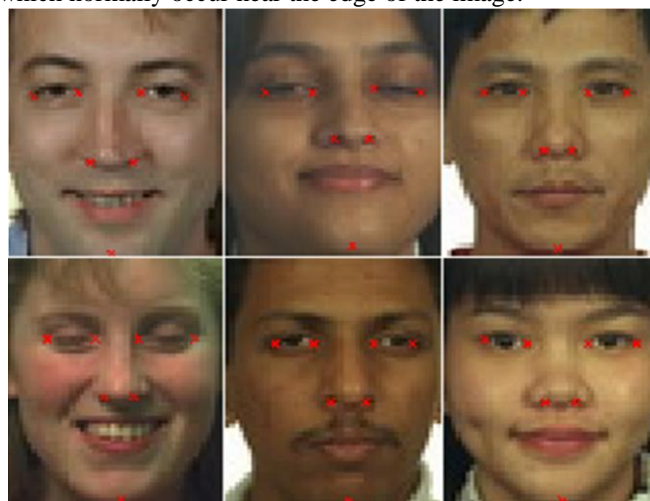
(a)
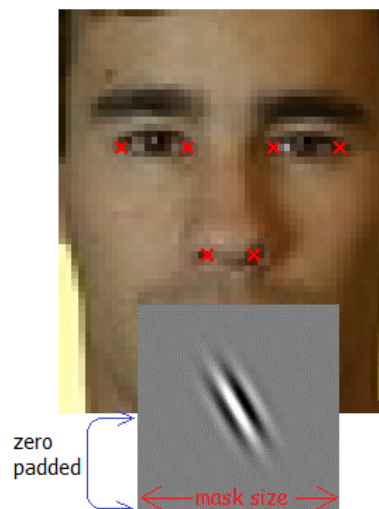
(b)

Fig. 2. (a) FERET face images with the seven landmarks selected (b) convolution of a Gabor kernel at the chin. Face images are zero padded for the convolutions where the wavelet exceeds the image dimensions.

B. Jet Extraction and Bunch Graph Creation

The convolution process produces a matrix having the same dimension as the Gabor wavelet dimension. According to [7], when the mask size of the wavelet comes closer to image size, the recognition performance increases. In this paper, the mask size was set 51 x 51 dimensions. Assuming that matrix **A** contains the complex Gabor wavelet coefficients for one landmark given by a single wavelet from a given image. All matrices **A** for a given landmark given by the 4o wavelets are concatenated into a single vector. A collection of the concatenated version of matrix **A** for one landmark is called a jet. Thus, assuming matrix **B** represent the Jet then, **B** = {**A**$_{J1}$, **A**$_{J2}$, ..., **A**$_{J40}$} contains the entire Gabor coefficient for one landmark. A bunch graph for an image is a collection of jets. Let matrix **C** represent a bunch graph, then Matrix **C** = {**B**$_{N1}$, **B**$_{N2}$, ..., **B**$_{N7}$} will be used for similarity score calculation between images.

IV. PROPOSED SYSTEM

A. Face Recognition System Block Diagram
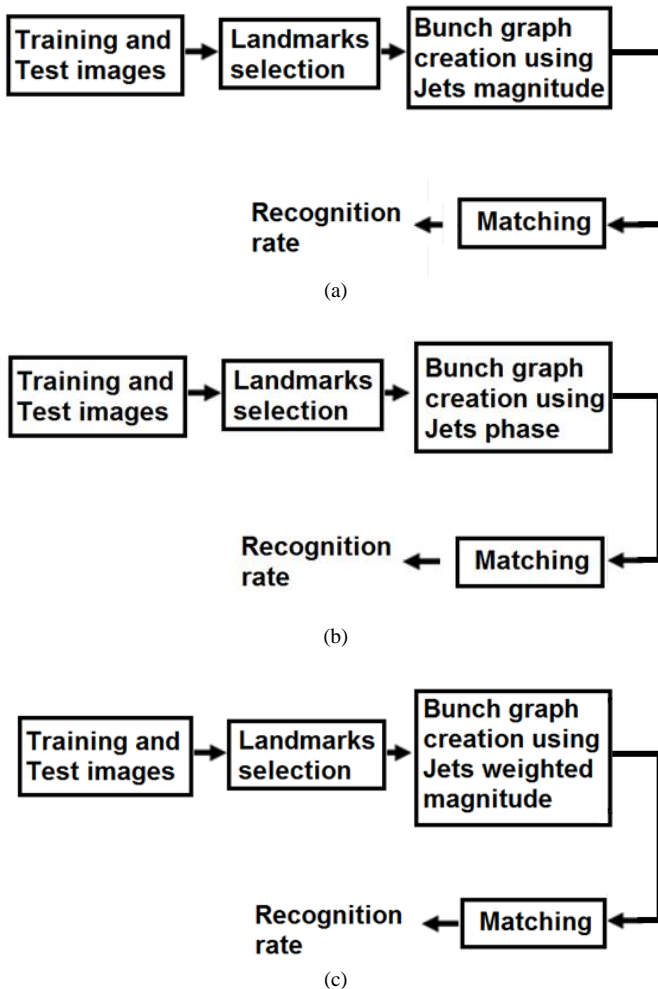

(a)

(b)

(c)

Fig. 3. Bunch graph face recognition system (a) magnitude only (b) phase only (c) weighted magnitude

Fig. 3 shows the block diagram of Gabor based face recognition system. The seven landmarks as shown in Fig.2(a) selected from face images are convolved with group of Gabor wavelets. Jets from each landmark were then collected together to create a bunch graph as face representation and will be use for the matching task. Three systems will be tested.
1) System A uses the jets magnitude information only
2) System B uses the jets phase information only
3) System C uses jets magnitude weighted by similarity of the phase between two different jets

B. Landmark Selection

The landmark selection for training images was done manually. For testing image, three method of landmark selection were conducted.
1) The first method is by manually selecting landmark on the testing image.
2) The second method is by using the mean coordinate from all training image landmark coordinates as shown in (4)

$$mean\_coord = \frac{1}{M}\sum_{i=1}^{M}\{(x,y)_{i1},(x,y)_{i2},...,(x,y)_{iN}\} \quad (4)$$

3) The third method is by using the mode coordinate from all training image landmark coordinates as in (5)
$$mode\_coord=\{mode(x,y)_1,mode(x,y)_2,...,mode(x,y)_N\} \quad (5)$$

where M = total of training image, and N = total landmark.

C. Similarity Score

For bunch graph similarity measurement, three similarity measurements are considered [4];

$$S_m(B,B') = \frac{\sum_{i=1}^{G}J_iJ'_i}{\sqrt{\sum_{i=1}^{G}J_i^2\sum_{i=1}^{G}J'_i^2}} \quad (6)$$

$$S_\phi(B,B') = \frac{\sum_{i=1}^{G}\phi_i\phi'_i}{\sqrt{\sum_{i=1}^{G}\phi_i^2\sum_{i=1}^{G}\phi'_i^2}} \quad (7)$$

$$S_p(B,B') = \frac{\sum_{i=1}^{G}J_iJ'_i\cos(\phi_i-\phi'_i)}{\sqrt{\sum_{i=1}^{G}J_i^2\sum_{i=1}^{G}J'_i^2}} \quad (8)$$

Where G is number of wavelet coefficients in a jet, J_i is the magnitude of the jet and ϕ_i is the phase angle. **B** and **B'** are the jets for two different images. Equation (6) computes jet

similarity score using jet magnitude (System A), (7) computes jet similarity score using jet phase (system B), while (8) use magnitude weighted by similarity of the phase angle to compute jet similarity score (System C). To compute the similarity score between two bunch graphs, (9) was used and N is total number of landmarks.

$$S_{bunch}(C,C') = \frac{1}{N}\sum_{i=1}^{N} S(B,B') \tag{8}$$

D. Matching

For the matching task, if the score $S_{bunch}(C,C')$ produced by (9), between the bunch graphs of a test image y and an image x in the training database is larger than a given threshold t, then images y and x are assumed to be of the same person. The scores produced by equation (8) were normalized so that $0 \le S_{bunch}(C,C') \le 1$, and the threshold t value can be tuned between 0 and 1. To measure the performance of the individual system, several performance metrics are used. These are:

i. For Recall Test
  a. **Correct Classification**. If a test image y_i is correctly matched to an image x_i of the same person in the training database.
  b. **False Acceptance.** If test image y_i is incorrectly matched with image x_j, where i and j are not the same person
  c. **False Rejection.** If image y_i is of a person i in the training database is rejected by the system.
ii. For Reject Test
  a. **Correct Classification**. If y_i, from the unknown test database is rejected by the system
  b. **False Acceptance**. If image y_i is accepted by the system.
iii. **Equal Correct Rate (ECR).** Recall correct classification is equal to reject correct classification.
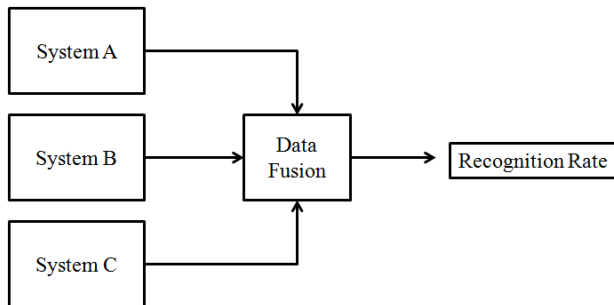
E. Data Fusion

Fig. 4. Block diagram of the fusion system

Fig. 4 shows the block diagram of the fusion of systems A, B, and C, mentioned in Section 4(A). The fusion decision stage is a module that consists of several rules.
1) For Recall
  • If both systems give correct matching, then correct match is found
  • If one system give correct matching and the other system give wrong matching or not found, then correct match is found
  • If both systems give wrong matching, then the fusion system give wrong matching
  • If one system gives wrong matching and the other system give not found, then the fusion system give wrong matching
  • If both system give not found, then the fusion system give not found
2) For Reject
  • If both system correctly reject image from unknown test database, then the fusion system give correct reject
  • If one system correctly reject image from unknown test database and the other system accept unknown test image, then the fusion system give correct reject
  • If both system accept image from unknown test database, then the fusion system give false acceptance

The fusion decision rules can be summarize as an OR operator as shown in Table I, Table II, Table III, and Table IV.

TABLE I
FUSION DECISION RULES

System A	System B	Fusion System output
0	0	0
1	0	1
0	1	1
1	1	1

TABLE II
FUSION DECISION RULES

System A	System C	Fusion System output
0	0	0
1	0	1
0	1	1
1	1	1

TABLE III
FUSION DECISION RULES

System A	System B	Fusion System output
0	0	0
1	0	1
0	1	1
1	1	1

TABLE IV
FUSION DECISION RULES

System A	System B	System C	Fusion System Output
0	0	0	0
1	0	0	1
0	1	0	1
0	0	1	1
1	1	0	1
1	0	1	1
0	1	1	1
1	1	1	1

The definition of the *0* and *1* result for both Recall and Reject test are as follow;

1) Definition for Recall test
   - 0 = Match not found
   - 1 = Correct Match found
2) Definition for Reject test
   - 0 = False Acceptance
   - 1 = Correct Reject

F. Probabilistic OR Rules

A modified OR, Probabilistic OR, is proposed. The rules of this OR gate takes into account confidence score of each individual system during the fusion stage. Table V shows the summary of the Probabilistic OR Rules.

TABLE V
PROBABILISTIC OR RULES

System A	System C	Fusion System output
0	0	0
1	0	$CS_A > CS_C \rightarrow 1$ $CS_A < CS_C \rightarrow 0$
0	1	$CS_A > CS_C \rightarrow 1$ $CS_A < CS_C \rightarrow 0$
1	1	1

If all individual system gives no match found, then the fusion system output give no match found result. The same applies if all individual system gives match found, then the fusion system output give match found result. However, when one system gives a match is found while the other system gives a match not found, then the output will be the state of the system having the highest confidence score. The confidence score is the modulus of the similarity score between test and matched training image, minus the score threshold of the individual system as shown in (10).

$$CS = \|S - t\| \qquad (9)$$

CS is confidence score, *S* is similarity score between test image and the matched training image, and *t* is the score threshold of the individual system.

G. Face Database

A total of 500 images with frontal face of a person were selected from the FERET database. They represent 200 different individuals. 100 individuals are used for training & testing, and the other 100 different individuals are used for testing only. All the 500 selected FERET images were cropped to get only the desired face part of a person (from forehead to the chin). All images are adjusted so that both eyes coordinates of an individual are aligned in the same horizontal line and the dimension for each image is set to 60 x 60 pixels. Three images per individual will be used for training. Two testing databases were created. The first database, Known Test

Database, has 100 images of the 100 persons in the training database. This database will be used to test the Recall capability of the face recognition system. The second database, Unknown Test Database, has also 100 images of 100 different persons. This database will be used to test the Rejection capability of the system. Fig. 5 shows the example of the normalized face image and Fig. 6 shows the FERET Face database tree chart used for experiments.

Fig. 5. Examples of the selected FERET face images are cropped from forehead to chin, eyes coordinates are aligned and images are converted into gray scale format.
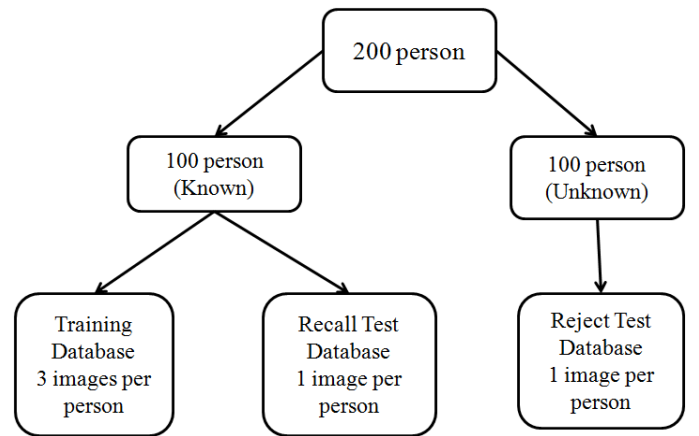

Fig. 6. FERET Face database chart used for experiments

V. RESULTS AND DISCUSSION

As stated earlier, the range of the similarity score can be between 0 and 1. The threshold also can be tuned so that the performance of the system can either have high correct recall with high false acceptance rate for application such as boarder monitoring or high correct rejection rate for unknown persons for application such as access control. For this work, the threshold tuning parameter was set so that each system has equal correct recall rate and correct rejection rate. Three landmark selection criteria were tested and three systems were considered.
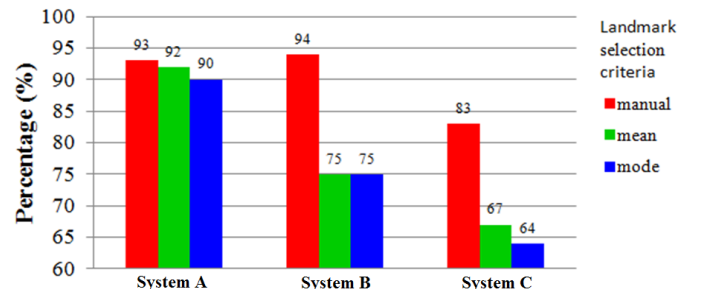

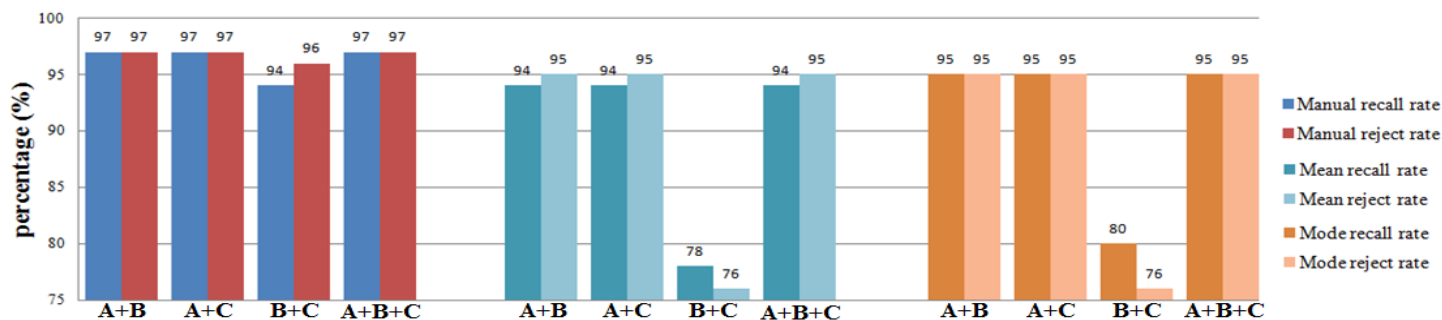Fig. 7. Recognition rate using magnitude, phase, and magnitude with phase

Fig. 8. Recognition rate for data fusion. The '+' sign means two or more systems were OR'ed.

Fig. 7 shows the performance of the system individually.

System A uses the jets magnitude, System B uses the jets phase, and System C uses the jets weighted magnitude. The manual landmark selection method outperforms the mean and mode selection methods for all three systems. Comparing the two automatic selection methods (mean and mode), the mean outperforms the mode selection criteria for all three systems. Comparing the performance of the individual system, system A outperforms the other systems in general except system B which gives slightly better result for the manual selection method.

Fig. 8 shows that the recognition rates for the fusion of all possible combination of two or three systems. In general, the fusion of two systems or more give better performance than a single system alone. In addition, the fusion reduces the effect of the landmark selection method. The result in Fig. 8 shows that fusion of magnitude and phase gives the best performance (system A and system B), thus only the fusion of magnitude and phase features of Gabor wavelet will be used for Probabilistic OR rules experiment.

Fig. 9 shows the result of data fusion using Probabilistic OR rules. Fusion system that uses the manual landmark selection outperforms fusion system that uses the mean and mode landmark selection by 15% approximately, while the performance between mean and mode selection more or less the same.

Comparing the Probabilistic OR rules result and the original OR rules result, the Probabilistic OR rules perform worst than the original OR rules regardless of the landmark selection method. When comparing the Probabilistic OR result with the individual system, the Probabilistic OR fusion based system outperforms all the individual system when using manual landmark selection method. However for the automatic landmark selection method (mean and mode), the Probabilistic OR fusion based system outperformed by System A, but outperform both systems B and C.
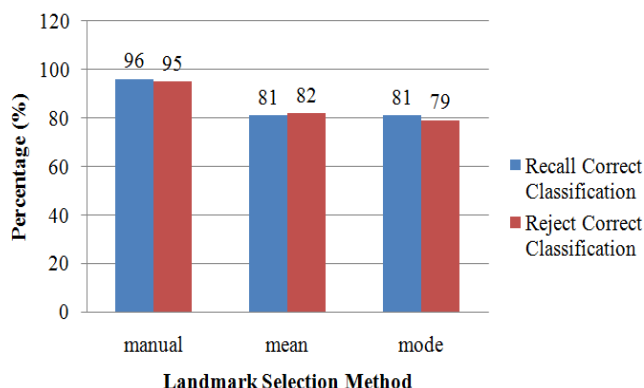

Fig. 9. Recognition rate for data fusion of magnitude and phase (A+B) using the Probabilistic OR rules.

The performance of our system is also compared with several methods that are based on bunch graph methods and use the same database as shown in Table VI. Our system performs better than both systems reported in [4] and [5] but lower than [3]. This may be due to the fact that [3] uses a precise jets extraction instead of just manually selecting a node on a face, thus creating a very detailed face graph with high precision as well designing the system specifically for in-class recognition task.

TABLE VI
COMPARISON OF SEVERAL EBGM-BASED FACE RECOGNITION METHODS ON FERET DATABASE.

Methods	Recognition Rate
Elastic Bunch Graph Method [3]	98%
EBGM (automatic facial feature selection) [4]	89.8%
Gabor wavelength selection based on Genetic Algorithm [5]	91%
Our proposed method (Original OR rules)	
Mean facial feature coordinate selection	94% (recall), 95% (reject)
Mode facial feature coordinate selection	95% (recall), 95% (reject)

VI. CONCLUSION

In this paper, a system that fuses the outputs of three systems is presented. These systems are based on the bunch graph method but one use magnitude of the jets only while the second one use the phase only, and last one use the magnitude weighted with phase. Three methods for selecting the landmarks where the jets are generated are used. It was found that selection method did not significantly affect the

performance of the fused system. However, the manual selection gives the highest recognition rate followed by the mean and mode methods. It was also found that the output of the fusion system using the OR rules gives higher recognition rate than all system individually. We have also introduced a fusion stage based on Probabilistic OR rules. However, it was found that Probabilistic OR rules perform worst than the original OR rules.

APPENDIX

TABLE XI
GABOR WAVELET PARAMETERS, WISKOTT [3]

Parameter	Symbol	Values
Orientation	θ	$\{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8, 5\pi/8, 6\pi/8, 7\pi/8\}$
Wavelength	λ	$\{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$
Phase	ϕ	$\{0, \pi/2\}$
Gaussian Radius	σ	$\sigma = \gamma$
Aspect Ratio	γ	1

VII. REFERENCES

[1] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces", *In Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

[2] Zhao, W., Chellappa, R., Phillips, P. J., Rosenfeld, A., Face recognition: A literature survey, *ACM Computing Surverys (CSUR),* V. 35, Issue 4, pp. 399-458, 2003.

[3] L. Wiskott, J.-M. Fellous, N. Kruger and C. Von Der Malsburg, "Face Recognition by Elastic Bunch Graph Matching", *In Intelligent Biometric Techniques in Fingerprint and Face Recognition*, Chapter 11, pp. 355-396, 1999.

[4] David Bolme. *Elastic bunch graph matching*. Master's thesis, Colorado State University, Summer 2003.

[5] Mohamad Hoseyn Sigari and Mahmood Fathy, "Best wavelength selection for Gabor wavelet using GA for EBGM algorithm", *Machine Vision, ICMV 2007*, Islamabad, pp. 35 - 39. 28-29 Dec. 2007.

[6] L. Wiskott, J.M. Fellous, N. Kruger, C.V.D Malsburg. "Face Recognition by Elastic Bunch Graph Matching". *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 775-779. July 1997.

[7] Berk Gökberk. *Feature Based Pose Invariant Face Recognition* . Master's thesis, Bogazici University, 2001.

J. Dargham, he received his B.Sc. in Control Systems Engineering from Iraq and his M.Sc. in Control System Engineering (UMIST) from Malaysia. He received his PhD from Universiti Malaysia Sabah (UMS). He is holding senior lecture position at Universiti Malaysia Sabah (UMS) and was the head of the Computer Engineering Program from 2006 till 2011. His research interests include Pattern Recognition, Medical Imaging, Biometrics, and Artificial Intelligence. He has published more than 70 papers in refereed journals, conferences, book chapters and research reports. (E-mail: jamalad@ums.edu.my).

A.Chekima, he received his BEngg in Electronics from Ecole Nationale Polytechnique of Algiers in 1976 and his Msc and Phd both in Electrical Engineering from Rensselaer Polytechnic Institute Troy, New York, in 1979 and 1984 respectively. He joined the Electronics Department at the Ecole Nationale Polytechnique in 1984, where he was Chairman of the Scientific Committee of the Department as well as in charge of the Postgraduate Program while teaching at both graduate and undergraduate levels. He was member of several scientific committees at the national level. He has been working as an Associate Processor at the School of Engineering and Information Technology at Universiti Malaysia Sabah since October 1996. His research interests include Source Coding, Antennas, Signal Processing, Pattern Recognition, Medical Imaging, Biometrics, Data Compression, Artificial Intelligence and Data Mining. He has published more than 120 papers in refereed journals, conferences, book chapters and research reports. (E-mail: chekima@ums.edu.my).

E.G. Moung, he received the B.Sc. degree in Computer Engineering from Universiti Malaysia Sabah, Malaysia 2008. He has been working as a research assistant at Universiti Malaysia Sabah, Malaysia. His present research interests include the biometric and image processing. (E-mail: menirva.com@gmail.com).