# PARALLEL GABOR PCA WITH FUSION OF SVM SCORES FOR FACE VERIFICATION

Ángel Serrano, Cristina Conde, Isaac Martín de Diego, Enrique Cabello
[1]*Face Recognition & Artificial Vision Group, Universidad Rey Juan Carlos*
*Camino del Molino, s/n, Fuenlabrada (Madrid), E-28943, Spain*
*angel.serrano@urjc.es, cristina.conde@urjc.es, isaac.martin@urjc.es, enrique.cabello@urjc.es*

Li Bai, Linlin Shen
*School of Computer Science & IT, University of Nottingham, Jubilee Campus, Nottingham, NG8 1BB, United Kingdom*
*bai@cs.nott.ac.uk, lls@cs.nott.ac.uk*

Keywords: Biometrics, face verification, face database, Gabor wavelet, principal component analysis, support vector machine, data fusion.

Abstract: Here we present a novel fusion technique for support vector machine (SVM) scores, obtained after a dimension reduction with a principal component analysis algorithm (PCA) for Gabor features applied to face verification. A total of 40 wavelets (5 frequencies, 8 orientations) have been convolved with public domain FRAV2D face database (109 subjects), with 4 frontal images with neutral expression per person for the SVM training and 4 different kinds of tests, each with 4 images per person, considering frontal views with neutral expression, gestures, occlusions and changes of illumination. Each set of wavelet-convolved images is considered in parallel or independently for the PCA and the SVM classification. A final fusion is performed taking into account all the SVM scores for the 40 wavelets. The proposed algorithm improves the Equal Error Rate for the occlusion experiment compared to a Downsampled Gabor PCA method and obtains similar EERs in the other experiments with fewer coefficients after the PCA dimension reduction stage.

## 1 INTRODUCTION

Face biometrics is receiving more and more attention, mainly because it is user-friendly and privacy-respectful and it does not require a direct collaboration from the users, unlike fingerprint or iris recognition.

Different approaches for face recognition have been applied in the last years. On the one hand, holistic methods use information of the face as a whole, such as principal component analysis (PCA) (Turk & Pentland, 1991), linear discriminant analysis (LDA) (Belhumeur et al., 1997), independent component analysis (ICA) (Bartlett et al., 2002), etc.

On the other one, feature-based methods use information from specific locations of the face (eyes, nose, mouth, for example) or distance and angle measurements between facial features. Some of these methods make use of Gabor wavelets (Daugman, 1985), such as dynamic link architecture (Lades et al., 1993) or elastic bunch graph matching (Wiskott et al., 1997).

Gabor wavelets have received much interest as they resemble the sensibility of the eye cells in mammals, seen as a complex plane wave with a Gaussian envelope. Most researchers follow the standard definition used by Wiskott et al. (1997),

$$\psi(\vec{r}) = \frac{k_\nu^2}{\sigma^2}\exp\left(-\frac{k_\nu^2\|\vec{r}\|^2}{2\sigma^2}\right)\left[\exp\left(i\vec{k}_{\mu\nu}\cdot\vec{r}\right)-\exp\left(-\frac{\sigma^2}{2}\right)\right] \quad (1)$$

with 5 frequencies and 8 orientations (Figure 1):

where $\vec{r} = (x, y)$, $\sigma$ is taken as a fixed value of $2\pi$, the wave vector is defined as $\vec{k}_{\mu\nu} = k_\nu(\cos\varphi_\mu, \sin\varphi_\mu)$ with module equal to $k_\nu = 2^{(-(\nu+2)/2)}\pi$ and orientation $\varphi_\mu = \mu\pi/8$ radians. The range of parameters $\mu$ and $\nu$ is $0 \le \mu \le 7$ and $0 \le \nu \le 4$, respectively. Therefore $\nu$ determines the wavelet frequency and $\mu$ defines its orientation.

Some previous works have combined Gabor wavelets with PCA and/or LDA (see Shen & Bai,

[1] http://frav.escet.urjc.es

2006, for a complete review on Gabor wavelets applied to face recognition). For example, Chung et al. (1999) use the Gabor wavelet responses over a set of 12 fiducial points as input to a PCA algorithm, yielding a vector of 480 components per person. They claim to improve the recognition rate up to a 19% with this method compared to a raw PCA.

Liu et al. (2002) vectorize the Gabor responses over the FERET database (Phillips et al., 2000) and then apply a downsampling by a factor of 64. Their Gabor-based enhanced Fisher linear discriminant model outperforms Gabor PCA or Gabor fisherfaces, although they perform a downsampling to obtain a low-dimensional feature representation of the images.

Zhang et al. (2004) suggest taking into account raw gray level images and their Gabor features in a multi-layered fusion method comprising PCA and LDA in the representation level, while using the sum and the product rules in the confidence level. They obtain better results for data fusion in the representation compared to confidence level, so they state that the fusion should be performed as early as possible in the recognition process.

In Fan et al. (2004), the fusion of Gabor wavelets convolutions after a downsampling process is explored with a null space-based LDA method (NLDA). The combination is made by considering the responses to all the wavelets of the same orientation and different rules of fusion, such as product, sum, maximum or minimum. Their best face recognition with Gabor+NLDA is 96.86% using a sum combination and after filtering the Gabor responses and keeping only the 75% of the pixels. As a comparison, their raw NLDA method obtains only a recognition rate of 92.26%.

On the other hand, in Qin et al. (2005) the responses of Gabor wavelets over a set of key points in a face picture are fed into an SVM classifier (Vapnik, 1995). These responses are concatenated into a high dimensional feature vector (of size 34355), which is then downsampled by a factor of 64. They obtain the best recognition accuracy with a linear kernel for gender classification using the FERET database and with a RBF kernel for face recognition using AT&T database.

In all previous works, the huge dimensionality that occurs when applying Gabor wavelets has been tackled (1) downsampling the size of the images (Zhang et al., 2004), (2) considering the Gabor responses over a reduced number of points (Chung et al., 1999), or (3) downsampling the convolution results (Liu et al., 2002, Fan et al., 2004). Strategies (2) and (3) have also been applied together (Qin et al., 2005). These methods suffer from a loss of information because of this downsampling. We propose a novel method that combines non-downsampled Gabor features with a PCA and an SVM for face verification.

The following paper is organized as follows: In Section 2 we describe the face database used in this work, FRAV2D. Section 3 explains the method proposed in this paper, the so-called parallel Gabor PCA, and revises other standard algorithms. The results and their discussion can be read in Section 4. Finally the conclusions are to be found in Section 5.

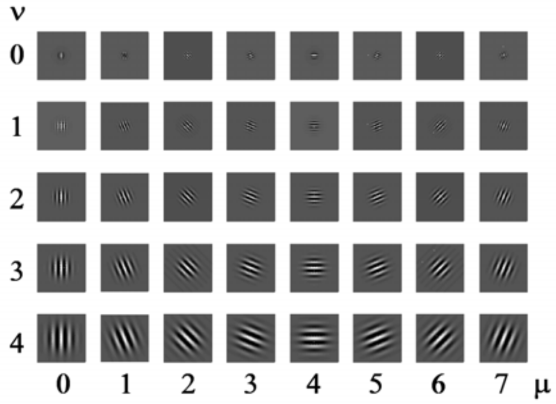


Figure 1: Real part of the set of 40 Gabor wavelets ordered by frequency (ν) and orientation (μ).

## 2 FRAV2D FACE DATABASE

We used the public domain FRAV2D face database, which comprises 109 people, mainly 18 to 40 years old (FRAV2D, 2004). There are 32 images per person, which is not usual in other databases. It was obtained during a year's time with volunteers (students and lecturers) at the Universidad Rey Juan Carlos of Madrid (Spain). Each image is a 240×340 colour picture obtained with a CCD video-camera. The face of the subject occupies most of the image.

The images were obtained under many different acquisition conditions, changing only one parameter between two shots to measure the effect of every factor in the verification process. The distribution of images is like this (Figure 2): 12 frontal views with neutral expression (diffuse illumination), 4 images with a 15º turn to the right, 4 images with a 30º turn to the right, 4 images with face gestures (smiles, winks, expression of surprise, etc.), 4 images with occluded face features, and finally, 4 frontal views with neutral expression (zenithal illumination).

Figure 2: Examples of images from FRAV2D database (from left to right, top to bottom: neutral expression with diffuse illumination, gestures, occlusion and neutral expression with zenithal illumination).

By means of a manual process, the position of the eyes was found in every image in order to normalize the face in size and tilt. Then the corrected images were cropped to a 128×128 size and converted into grey scale, with the eyes occupying the same positions in all of them. For the images with occlusions, only the right eye is visible. In this case, the image was cropped so that this eye is located at the same position as in the other images, but no correction in size and tilt was applied. In every case, a histogram equalization was performed to correct variations in illumination.

## 3 DESIGN OF THE EXPERIMENTS

As can be seen in Table 1, we have divided the database into two disjoint groups, one for training (gallery set) and the other one for tests (test set). On the one hand, as a training set we have considered four random frontal images with neutral expression for every person. On the other one, we have performed several test sets, yielding a total of four experiments, which allows us to study the influence of the different types of images of the database. Test 1 takes images with neutral expression, which are different to those considered in the gallery set. Test 2 comprises images with gestures, such as smiles, open mouths, winks, etc. Test 3 tackles images with occlusions, while test 4 considers images with changes of illumination. In every test, four images per person were taken into account.

As follows, we describe the method proposed in this paper (parallel Gabor PCA) and three baseline algorithms (PCA, feature-based Gabor PCA and downsampled Gabor PCA), used for comparison.

Table 1: Specification of our experiments.

| Exper-iment | Images/person in gallery set | Images/person in test set |
|---|---|---|
| 1 | 4 (neutral expression) | 4 (neutral expression) |
| 2 | | 4 (gestures) |
| 3 | | 4 (occlusions) |
| 4 | | 4 (illumination) |

### 3.1 Parallel Gabor PCA

We propose applying a PCA after the convolution of 40 Gabor wavelets with the images in the database, similarly to Liu et al. (2002), Shen et al. (2004), Fan et al. (2004) and Qin et al. (2005). The main difference is that we do not perform a fusion of downsampled Gabor features before a PCA, as these authors do. On the contrary, we suggest carrying out in parallel a PCA and an SVM classification for each wavelet frequency and orientation, followed by a final fusion of SVM scores.

As we take into account 40 Gabor wavelets, in all we have 40 PCAs, each of them is performed over the 128×128 wavelet-convolved images turned into vectors of size 16384×1. Therefore no downsampling is applied to the convolutions, so no loss of information is produced.

Specifically, for each of the 40 Gabor wavelets with orientation $\mu$ and frequency $\nu$, we have followed these steps, divided into training phase and test phase (Figure 3):

1. The training phase begins with the convolution of the wavelet with all the images in the gallery database.
2. Then we perform a dimension reduction process with a PCA after turning the convolutions into column vectors. The projection matrix generated with the eigenvectors corresponding to the highest eigenvalues is applied to the results of the convolutions, so the projection coefficients for each image are computed. The number of coefficients is the “dimensionality”.
3. For each person in the gallery, a different SVM classifier is trained using these projection coefficients. In every case, the images of that person are considered as genuine cases, while everybody else’s images are considered as impostors. A face model for each person is computed.
4. The test phase starts with the convolution of the wavelet with all the images in the test database. Then the results are projected onto the PCA

framework using the projection matrix computed in step 2.

5. Finally we classify the projection coefficients of the test images with the SVM for every person using the corresponding face model. As a result every SVM produces a set of scores.

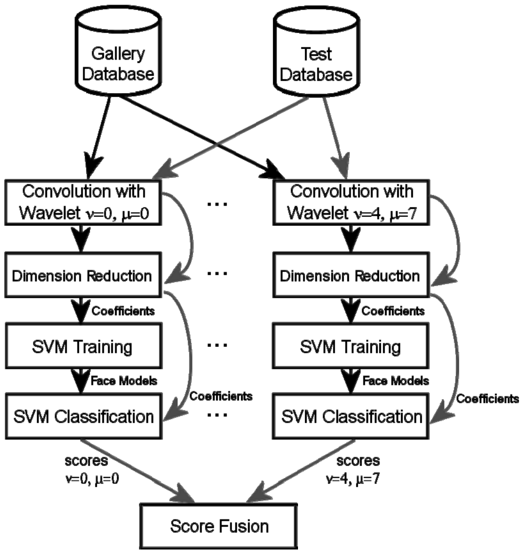


Figure 3: A schematic view of our algorithm. The black arrows correspond to the sequence for the training phase, where the gallery database is used to train the SVMs and generate the face model for every person. The grey arrows belong to the test phase, where the test database is used to perform the SVM classification and the fusion of scores.

With all the SVM scores obtained for each of the 40 wavelets, a fusion process was carried out by computing an element-wise mean average of the scores. An overall receiver operating characteristic curve (ROC) was computed with the fused data, which allowed us to calculate the equal-error rate (EER) at which the false rejection rate equals the false acceptation rate. The lower this EER is, the more reliable the verification process will be.

## 3.2 PCA

The first baseline method considered was a PCA with the usual algorithm (Turk & Pentland, 1991), followed by an SVM classifier (Vapnik, 1995).

The gallery set was projected onto the reduced PCA framework, where the number of projection coefficients is the "dimensionality". After training an SVM for every person, the test set was also projected onto the PCA framework. The SVMs produced a set of scores that allowed computing an overall ROC curve and the corresponding EER.

## 3.3 Feature-based Gabor PCA

Following Chung et al. (1999), we also considered the convolution of the images with a set of 40 Gabor wavelets evaluated at 14 manually-selected fiducial points located at the face features. A column vector of 560 components was created for each person. For the images in the occlusion set, only 8 fiducial points could be selected, as the other ones were hidden by the subject's hand. All the column vectors were fed to a standard PCA. After the dimension reduction, an SVM classifier was trained for each person, as in the previous sections.

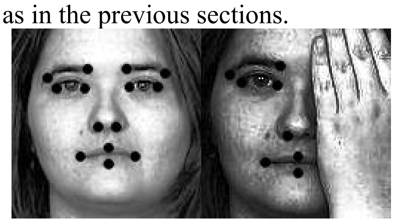

Figure 4: Image with the 14 fiducial points considered for an image with all the face features visible (left). For images with occlusions, only 8 points were taken into account (right).

## 3.4 Downsampled Gabor PCA

Instead of keeping only the Gabor responses over a set of fiducial points, as a third baseline method we also considered all the pixels in the image in the same way as Liu et al. (2002), Shen et al. (2004), Fan et al. (2004) and Qin et al. (2005), among others. The results of the 40 convolutions over the 128×128 images were fused to generate an augmented vector of size 655360×1. Because of this huge size, a downsampling by a factor of 16 was performed, so that the feature vectors were reduced to 40960 components. These vectors were used for a PCA dimension reduction and an SVM classification, as in the previous sections.

Table 2: Best EER (%) and optimal dimensionality (in brackets) for every method and experiment.

| **Exper-iment** | **PCA** | **Featured-based Gabor PCA** | **Down-sampled Gabor PCA** | **Parallel Gabor PCA** |
|---|---|---|---|---|
| 1 | 1.83 (60) | 0.46 (140) | 0.23 (200) | 0.23 (50) |
| 2 | 10.55 (200) | 11.93 (190) | 5.50 (160) | 5.96 (120) |
| 3 | 34.00 (180) | 30.96 (140) | 25.09 (160) | 22.40 (180) |
| 4 | 3.70 (180) | 1.78 (190) | 0.42 (180) | 0.46 (140) |

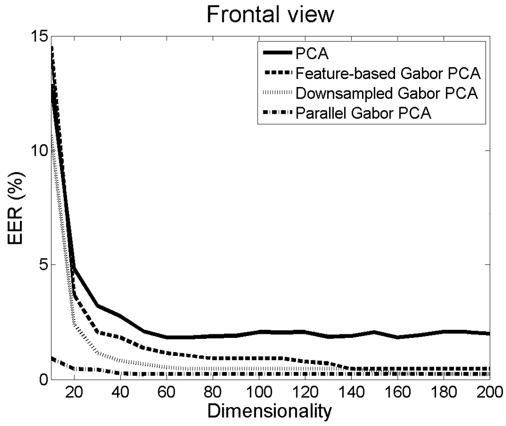


Figure 5: EER as a function of dimensionality for experiment 1 (frontal views with neutral expression).

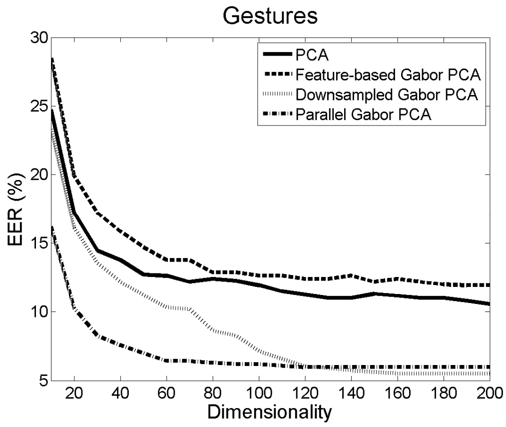


Figure 6: EER as a function of dimensionality for experiment 2 (images with gestures).

# 4 RESULTS AND DISCUSSION

For every experiment in Table 1 and the four methods considered in Section 3, we computed the ROC curve and the corresponding EER for a set of dimensionalities ranging between 10 and 200 coefficients (Figures 5 – 8). Table 2 summarizes the results and shows the best EER and the dimensionality for which it was obtained for every experiment and method.

On the one hand, for experiment 1 (frontal images with neutral expression) both the downsampled Gabor PCA and the parallel Gabor PCA obtain the best EER (0.23%), although the latter needs fewer coefficients than the former in the PCA phase (200 vs. 50). The worst results are obtained for a standard PCA, which can only achieve an EER of 1.83% with a dimensionality of 60.

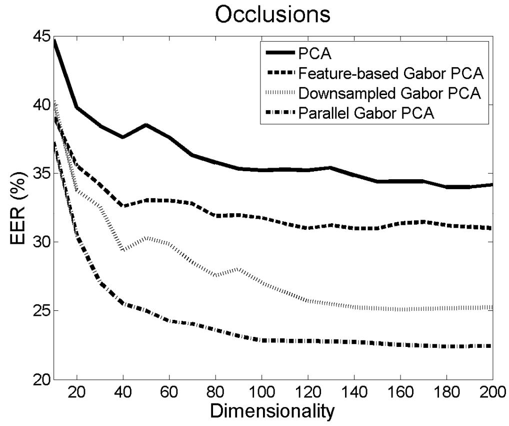


Figure 7: EER as a function of dimensionality for experiment 3 (images with occlusions).

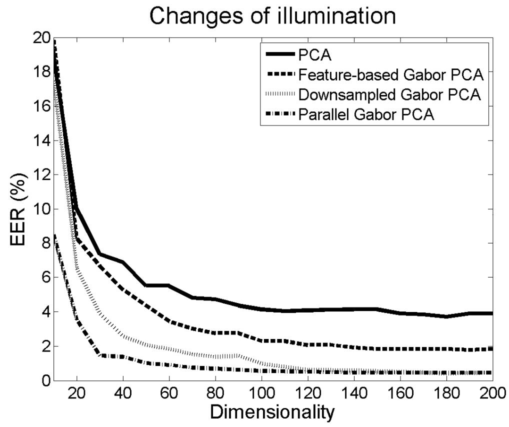


Figure 8: EER as a function of dimensionality for experiment 4 (images with changes of illumination).

In experiment 2 (images with gestures), an EER of 5.50% is obtained with the downsampled Gabor PCA (160 coefficients), although the EER for parallel Gabor PCA is only slightly worse (5.96%), but can be obtained with a lower dimensionality. In this case the worst EER corresponds to the feature-based Gabor PCA (11.93%) with 190 coefficients.

In experiment 3 (images with occlusions), parallel Gabor PCA outperforms the other methods with an EER of 22.40% (180 coefficients). As a comparison, the worst EER corresponds to the standard PCA (34.00%), also with a dimensionality of 180.

Finally in experiment 4 (images with changes of illumination), the downsampled Gabor PCA and the parallel Gabor PCA obtain the best results with EERs of 0.42% and 0.46%, respectively, but again the latter needs fewer coefficients (140 vs. 180). The highest EER corresponds to PCA (3.70%) with a dimensionality of 180.

Figures 5 – 8 also show that for parallel Gabor PCA the EER drops drastically as the dimensionality increases and it stabilizes quickly to its lowest value, always for few coefficients. Moreover, although the downsampled Gabor PCA obtains better EERs in certain experiments compared to our method, this one achieves similar EERs with fewer coefficients.

Unlike Zhang et al. (2005), our results also show that the data fusion can be performed at the score level instead of the feature representation level.

Finally, as a drawback to the proposed algorithm, the bigger computational load has to be taken into account. The PCA computation and the SVM training and classification have to be repeated 40 times, each for every Gabor wavelet. As a future work, it would be interesting to implement this algorithm in a parallel architecture in order to tackle each wavelet concurrently.

## 5 CONCLUSIONS

A novel method for face verification based on the fusion of SVM scores has been proposed. The experiments have been performed with the public domain FRAV2D database (109 subjects), with frontal views images with neutral expression, gestures, occlusions and changes of illumination.

Four algorithms were compared: standard PCA, feature-based Gabor PCA, downsampled Gabor PCA and parallel Gabor PCA (proposed here). Our method has obtained the best EER in experiments 1 (neutral expression) and 3 (occlusions), while the downsampled Gabor PCA achieves the best results in the others. In these cases, the parallel Gabor PCA obtains similar EERs with a lower dimensionality.

## ACKNOWLEDGEMENTS

This work has been carried out by financial support of the Universidad Rey Juan Carlos, Madrid (Spain), under the program of mobility for teaching staff. Special thanks have to be given to Ian Dryden, from the School of Mathematical Sciences of the University of Nottingham (United Kingdom), for his warm help and interesting discussions.

## REFERENCES

Bartlett, M.S., Movellan, J.R., Sejnowski, T.J., 2002. Face recognition by independent component analysis. *IEEE Transactions on Neural Networks*, Vol. 13, No. 6, pp. 1450-1464.

Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J, 1997. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE PAMI*, Vol. 19, No. 7, pp. 711-720.

Chung, K.-C., Kee, S.C., Kim, S.R., 1999. Face Recognition using Principal Component Analysis of Gabor Filter Responses. *International Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-Time Systems*, pp. 53-57.

Daugman, J.G., 1985. Uncertainty relation for resolution in space, spatial-frequency and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*: *Optics Image Science and Vision*, Vol. 2, Issue 7, pp. 1160-1169.

Fan, W., Wang, Y., Liu, W., Tan, T., 2004. Combining Null Space-Based Gabor Features for Face Recognition, *17th International Conference on Pattern Recognition ICPR'04*, Vol. 1, pp. 330-333.

FRAV2D Face Database Website, 2004. http://frav.escet.urjc.es/ databases/FRAV2D/

Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P., Konen, W., 1993. Distortion invariant object recognition in the Dynamic Link Architecture, *IEEE Transactions on Computers*, Vol. 42, No. 3, pp. 300-311.

Liu, C.J., Wechsler, H., 2002. Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition, *IEEE Transactions on Image Processing*, Vol. 11, Issue 4, pp. 467-476.

Philips, P.J., Moon, H., Rauss, P.J., Rizvi, S, 2000. The FERET evaluation methodology for face recognition algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1090-1104.

Qin, J., He, Z.-S., 2005. A SVM face recognition method based on Gabor-featured key points, *4th International Conference on Machine Learning and Cybernetics*, Vol. 8, pp. 5144-5149.

Shen, L., Bai, L., 2004, Face recognition based on Gabor features using kernel methods, *6th IEEE Conference on Face and Gesture Recognition*, pp. 170-175

Shen, L., Bai, L., 2006, A review on Gabor wavelets for face recognition. *Pattern Analysis & Applications*, Vol. 9, Issues 2-3, pp. 273-292.

Turk, M., Pentland, A., 1991. Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, Vol. 3, Issue 1, pp. 71-86.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory, Springer Verlag.

Wiskott, L., Fellous, J.M., Kruger, N., von der Malsburg, C., 1997. Face recognition by Elastic Bunch Graph Matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, Issue 7, pp. 775-779.

Zhang, W., Shan, S., Gao, W., Chang, Y., Cao, B., Yang, P., 2004. Information Fusion in Face Identification, *17th International Conference on Pattern Recognition ICPR'04*, Vol. 3, pp. 950-953.