Influence of wavelet frequency and orientation in an SVM-based Parallel Gabor PCA face verification system

Ángel Serrano[1], Isaac Martín de Diego[1], Cristina Conde[1], Enrique Cabello[1], Linlin Shen[2], and Li Bai[3]

[1] Face Recognition & Artificial Vision Group, Universidad Rey Juan Carlos, Camino del Molino s/n, Fuenlabrada E-28943 (Madrid), Spain, http://www.frav.es/
{angel.serrano, isaac.martin, cristina.conde, enrique.cabello}@urjc.es
[2] Faculty of Information and Engineering, Shenzhen University, Shenzhen, 518060, China, llshen@szu.edu.cn
[3] School of Computer Science and IT, University of Nottingham, Nottingham, NG8 1BB, United Kingdom, bai@cs.nott.ac.uk

Abstract. We present a face verification system using Parallel Gabor Principal Component Analysis (PGPCA) and fusion of Support Vector Machines (SVM) scores. The algorithm has been tested on two databases: XM2VTS (frontal images with frontal or lateral illumination) and FRAV2D (frontal images with diffuse or zenithal illumination, varying poses and occlusions). Our method outperforms others when fewer PCA coefficients are kept. It also has the lowest equal error rate (EER) in experiments using frontal images with occlusions. We have also studied the influence of wavelet frequency and orientation on the EER in a one-Gabor PCA. The high frequency wavelets are able to extract more discriminant information compared to the low frequency wavelets. Moreover, as a general rule, oblique wavelets produce a lower EER compared to horizontal or vertical wavelets. Results also suggest that the optimal wavelet orientation coincides with the illumination gradient.

Keywords: Face Verification, Gabor Wavelet, Parallel Gabor Principal Component Analysis, Support Vector Machine, Data Fusion.

1 Introduction

Automated face recognition systems are developing rapidly, due to increasing computational capabilities, both in speed and storage, and its ease for use compared to other biometrics where the user collaboration is mandatory [1, 2]. There are a large variety of methods available in the literature for face recognition, such as Principal Component Analysis (PCA) [3] or Linear Discriminant Analysis (LDA) [4]. Some methods make use of Gabor wavelets [5] due to their similarities in behaviour to the human cells in the visual cortex. Following the standard definition given by [6] and [7], a Gabor wavelet is a 2D filter defined as a complex wave with a Gaussian envelope (Figure 1). It can be parameterized by a frequency ν ($0 \le \nu \le 4$) and an orientation μ ($0 \le \mu \le 7$):

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

where $\vec{r} = (x, y)$ and $\sigma = 2\pi$. The wave vector, which determines the direction of the propagation of the wave, is defined as $\vec{k}_{\mu\nu} = k_\nu\left(\cos\varphi_\mu, \sin\varphi_\mu\right)$ with $k_\nu = 2^{(-(\nu+2)/2)}\pi$ and $\varphi_\mu = \mu\pi/8$ radians (with respect to the horizontal axis). $\vec{k}_{\mu\nu}$ is perpendicular to the direction of the wavelet, considered as the wavefronts.

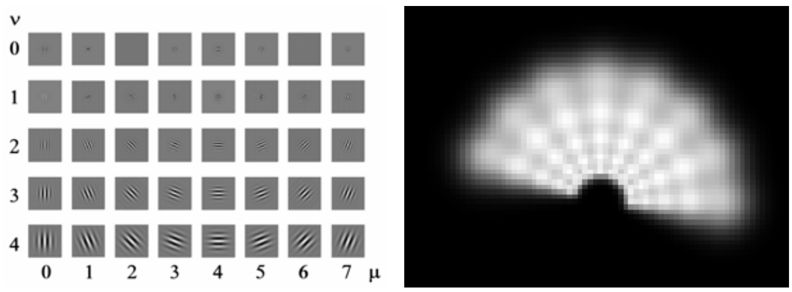

Fig. 1. Left: Bank of 40 Gabor filters, ordered by frequency (ν) and orientation (μ). Right: Response of Gabor filters in the Fourier space (only drawn half space).

The usual strategy for Gabor-based methods consists in convolving the images with the set of 40 filters and then working with the absolute value of the results [8]. As the dimensionality grows by a factor of 40 with these methods, many researchers have tackled this problem by combining Gabor wavelets with a dimension reduction algorithm, such as PCA. Analytic methods consider the Gabor responses computed only over a set of fiducial points, such as eyes, nose and mouth [9]. We shall call these methods "Feature-based Gabor PCA" (FGPCA). Holistic methods take into account the Gabor responses from the whole face image. Due to the huge dimensionality of Gabor features, a downsampling process is usually performed to reduce the dimension by a certain factor (usually 16 or 64) [10, 11, 12, 13]. We shall call these methods "Downsampled Gabor PCA" (DGPCA).

The algorithm proposed here is holistic, but it uses no downsampling process, as all the wavelet convolutions are performed in parallel, i.e., in a multi-channel approach. A final fusion of the results will allow us to evaluate the performance of our method and compare it with others mentioned above.

We also want to explore which of the 40 Gabor wavelets is able to extract the most discriminant features for a face verification problem. Some experiments have been done to study the influence of spatial frequency and orientation of face features [1, 14, 15, 16]. These works suggest that low frequency information can help us distinguish a face from a "non-face", but it is the high frequency information which is needed to tell whether two faces are different. The importance of facial bilateral symmetry as a key element to identify a "beauty face" [17, 18] and its influence in the ability to recognize a face [19, 20] has been also considered in the past.

The remainder of this paper is organized as follows. In Section 2, we describe the face databases used. In Section 3 we explain our algorithm, the Parallel Gabor PCA. The design of our experiments can be found in Section 4. The results and a discussion are in Section 5. Finally, the conclusions are to be found in Section 6.

Fig. 2. Sample images from XM2VTS Database (above) and FRAV2D Database (below).

2 Face databases

2.1 XM2VTS Database

XM2VTS is a multi-modal face database (Figure 2, top) [21] from the University of Surrey, UK, which comprises 2D pictures (frontal and profile views), as well as 3D meshes for 295 people. For our experiments we selected 100 people randomly, each having four frontal pictures taken in three different sessions. The first and the second sessions had frontal diffuse illumination, while in the third one the lighting was lateral (in two images the light came from the left and in the others, it came from the right). We use the set of four images from the first session of every person in the gallery database to train our classifiers. The remaining images from the other sessions are used in the tests to verify the accuracy of our algorithm. A manual process is used to normalize the face images. The images are cropped to 128×128 and converted into grey scale, with the eyes occupying the same locations in all the pictures. Finally a histogram equalization is performed on the images to deal with changes of illumination.

2.2 FRAV2D Database

We have also used the public domain FRAV2D Database (Figure 2, bottom) [22], which is freely available to the scientific community for research purposes. It comprises 109 people, each with 32 images. It includes frontal images with diffuse and zenithal illumination, 15º and 30º head orientations, and images with occlusions. As with XM2VTS Database, the images are normalized to 128×128 manually and histogram equalization is applied on them.

3 Our algorithm

We have developed the so-called Parallel Gabor methods (Figure 3) [23, 24]. The core of the algorithms is a PCA-based dimension reduction process. However, unlike a standard PCA, a set of Gabor wavelet convolutions are applied to the gallery database, in order to extract information of frequency and orientation in the images. Following [7], a set of 40 wavelets (8 orientations and 5 frequencies) are used, so that the overall dimensionality of the problem increases by a factor of 40. Unlike other methods that try to tackle this huge dimensionality by downsampling the feature vectors, the Parallel Gabor methods do not perform any downsampling process at all, but they consider the images convolved with the same wavelet in parallel and independently. After the Gabor-based PCA, a set of person-specific SVM classifiers [25] are trained with the PCA projection coefficients. In this scenario, the images of one person are considered as genuine cases and the remaining ones are impostors. As we work with the Gabor convolutions in parallel, there are 40 different SVMs per person, each corresponding to a wavelet frequency and orientation.

The same steps are applied for the images in the test database. In this case, the PCA projection matrix learnt in the previous step is applied to these images in order to compute the PCA coefficients. These will be fed into the SVM classifiers in order to obtain a set of 40 scores, each one for every wavelet frequency and orientation, which are averaged so as to produce a final score [24]. With the fused scores, an overall equal-error rate (EER), for which the false acceptance rate equals the false rejection rate, is computed in order to characterize the goodness of the method.

In this paper, we compare our algorithm with others such as a standard PCA, a FGPCA (with 14 features, 8 for the occluded images and the 30º-turned images) and a DGPCA with a downsample factor of 16. An alternative one-Gabor PCA method that performs no data fusion has been implemented, in order to find out the influence of the wavelet frequency and orientation on the final EER.

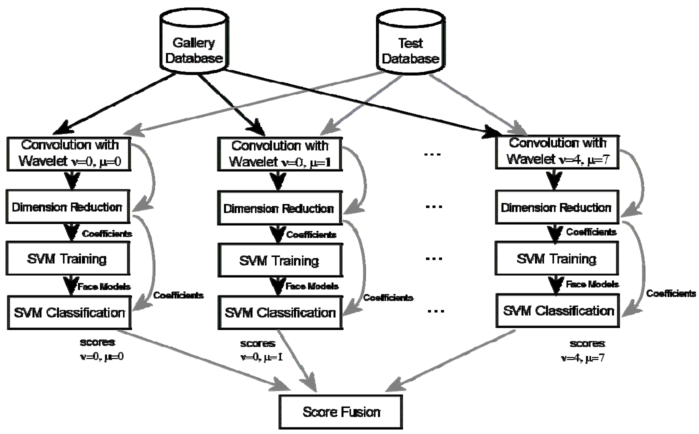

Fig. 3. Outline of our algorithm (PGPCA). Black arrows indicate the SVM training stage and grey arrows show the SVM test phase.

4 Design of Experiments

Six experiments were carried out (Table 1), for which a set of 4 frontal images per person was used to train the SVM classifiers. In every experiment a test with a disjoint set of 4 images per person was completed in order to compute the overall EER of the system, which has to be done by considering the scores for all the person-specific SVMs. Due to the configuration of the XM2VTS Database, only two tests could be performed, although six experiments are available for FRAV2D. As well there is a slight difference of meaning for test 6: While for FRAV2D the light direction changes from frontal diffuse to zenithal (which produces some shadows under the face features, such as the eyebrows, the nose and the mouth), for XM2VTS the illumination changes from frontal diffuse to lateral. This yields a dramatic effect on the images (half face is lit, while the other part is in shadow) and should be taken into account when comparing the results of this experiment for both databases.

Table 1. Specification of our experiments.

Experiment	Images per person in gallery set	Images per person in test set	FRAV2D Database	XM2VTS Database
1	4 (neutral expression)	4 (neutral expression)	✓	✓
2		4 (15º turn)	✓	
3		4 (30º turn)	✓	
4		4 (gestures)	✓	
5		4 (occlusions)	✓	
6		4 (illumination)	✓	✓

5 Results and discussion

5.1 Performance of Parallel Gabor PCA versus other methods

In Figures 4 and 5 we present the EER with respect to the dimensionality, that is, the number of PCA coefficients kept after the dimension reduction, using different Gabor-based methods (FGPCA, DGPCA and PGPCA). A standard PCA has also been included as a reference.

Figure 4 shows the results for the XM2VTS Database. When few eigenvalues (60 – 70) are kept, PGPCA always obtains the lowest error compared to the other methods. This means that PGPCA succeeds even when an important dimension reduction is performed in the PCA stage. However, if we consider a higher dimensionality, DGPCA seems to obtain the lowest EER (1.0%), just slightly better than PGPCA (1.2%), in test 1 (Figure 4, left). In test 6 (Figure 4, right), DGPCA outperforms clearly the other methods with an EER 8.5% (PGPCA can only achieve 12.3%).

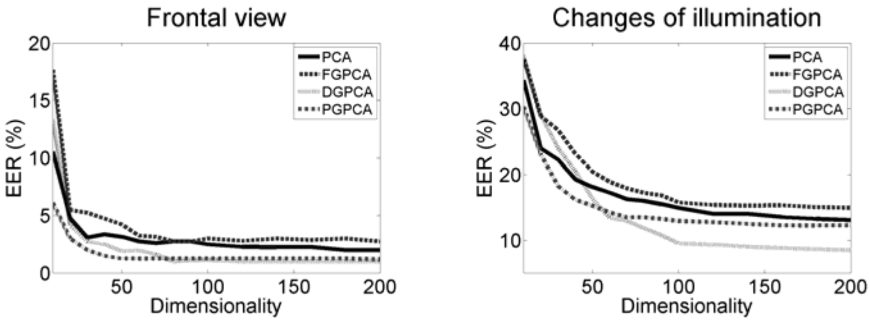

Fig. 4. Evolution of the EER as a function of the dimensionality (number of coefficients kept in the dimension reduction) for XM2TVS Database. From left to right: tests 1 and 6.

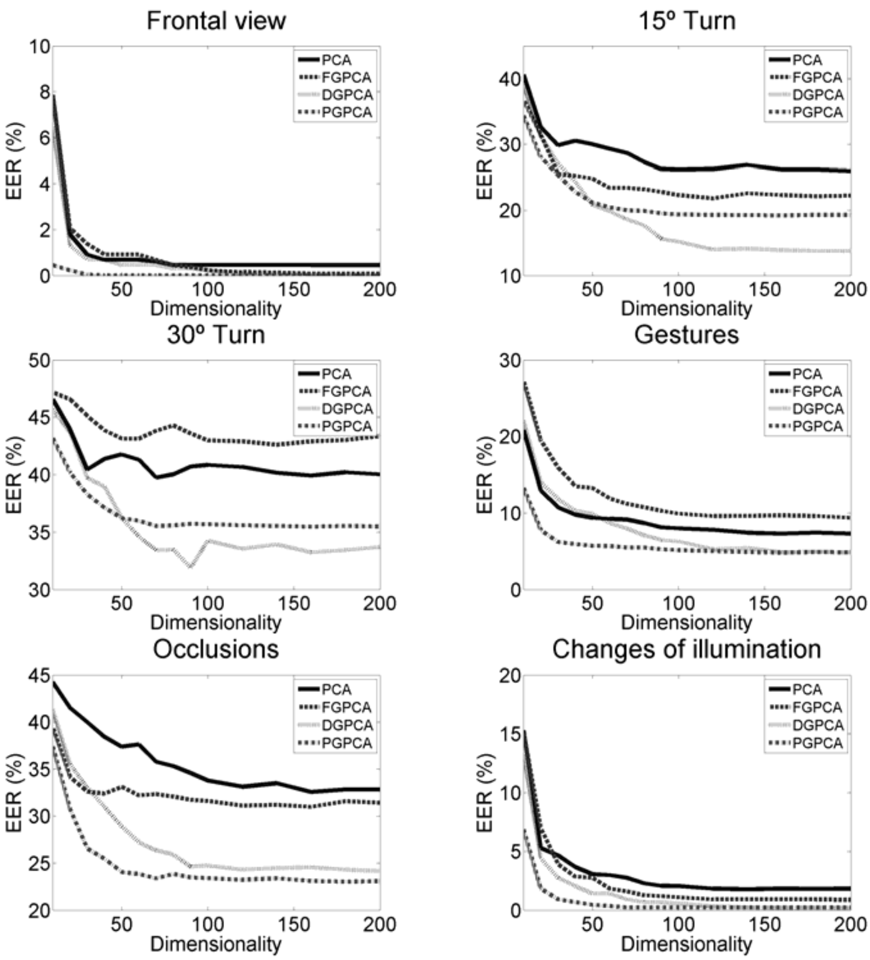

Fig. 5. Evolution of the EER as a function of the dimensionality for FRAV2D Database. From left to right, top to bottom: tests 1 to 6.

In Figure 5 we present the results for the FRAV2D Database. In this case, PGPCA obtains the lowest EER for test 1 (0.00%), test 4 (4.88%) and test 5 (23.04%), beating DGPCA (0.01%, 4.89% and 24.17%, respectively). For test 6, both methods obtain similar results (DGPCA 0.17%, PGPCA 0.23%). However, tests 2 and 3 (turns) show that DGPCA outperforms easily the other methods, included PGPCA (13.76% vs. 19.26% for a 15º head orientation, and 33.26% vs. 35.52% for a 30º head orientation). Therefore, for the FRAV2D Database, PGPCA methods achieves clearly the lowest error in three out of six experiments (tests 1, 4 and 5), although it obtains a slightly worse EER with respect to DGPCA in test 6. On the contrary, DGPCA outperforms PGPCA for tests 2 and 3. The other baseline methods, a standard PCA and FGPCA, always obtain the worst EERs for both databases in all experiments and have been included here only to help the comparison of results.

5.2 Influence of Gabor wavelet frequency and orientation in a one-Gabor PCA

We have carried out another experiment in order to investigate the discriminant capabilities of Gabor wavelets. In this case, all the images in the gallery database are convolved with a unique Gabor wavelet of a certain frequency ν and orientation μ. With no downsampling and after a PCA dimension reduction process, the feature vectors are used to train a set of person-specific SVMs, just like in the previous section. However, the main difference here is that no score fusion is performed. Therefore, we have obtained a set of 40 EERs, each one for every Gabor wavelet, repeated for the six experiments in Table 1. The goal of this section is to learn which wavelet, when considered alone, is able to extract the face features with the highest distinguishing properties.

Figure 6 plots the EER as a function of the wavelet frequency (ν) for all orientations (μ) for both databases. For simplicity only the results for test 1 are shown (the corresponding figures for the other experiments are similar). This figure shows that, as a general rule, the wavelets with a higher frequency (low ν) give a better EER than the wavelets with a lower frequency (high ν), for both databases and all the experiments. This can be easily understood, as the low frequency information allows distinguishing a face from a "non-face", but it is not enough to separate two similar faces. It is the high frequency information which provides the necessary details to tell one face from the other.

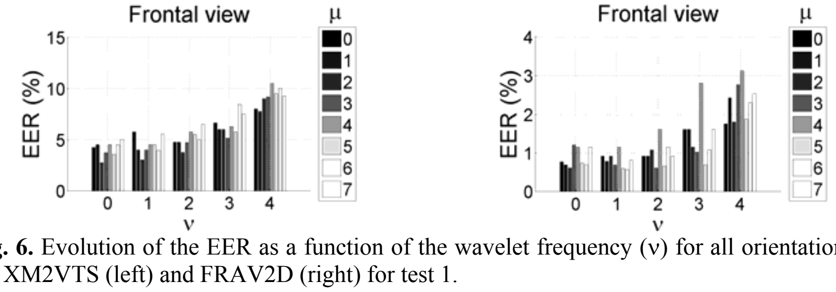

Fig. 6. Evolution of the EER as a function of the wavelet frequency (ν) for all orientations (μ) for XM2VTS (left) and FRAV2D (right) for test 1.

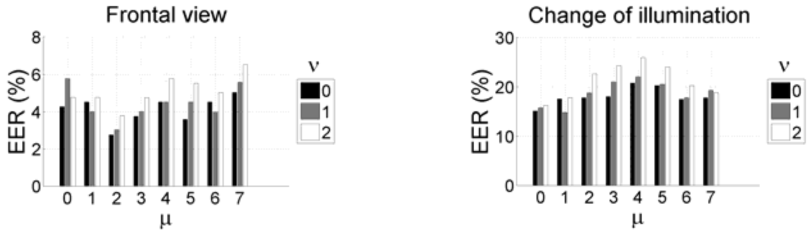

Fig. 7. Evolution of the EER as a function of the wavelet orientation (μ) for all frequencies (ν) for XM2VTS Database (left: test 1, right: test 6).

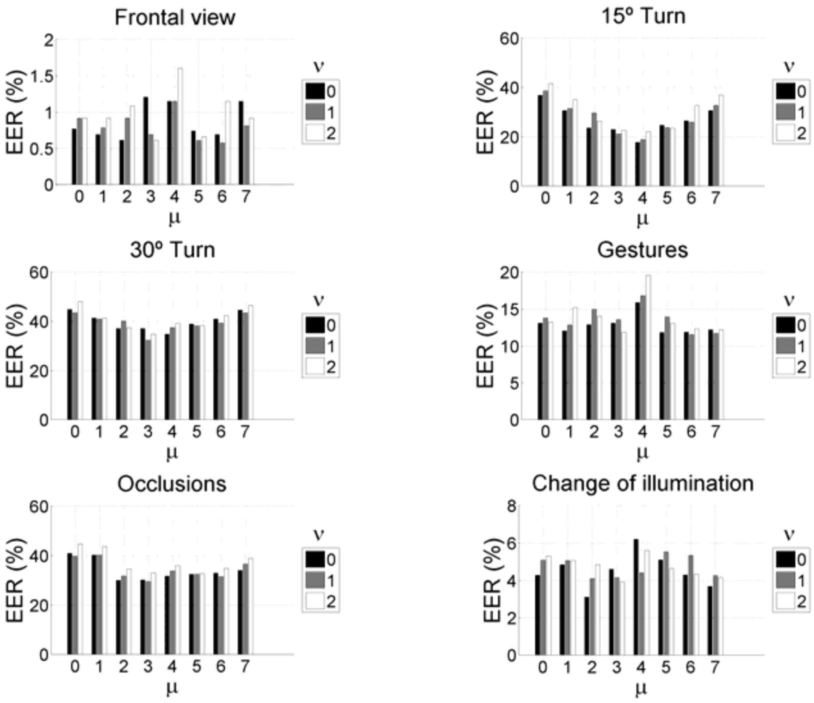

Fig. 8. Evolution of the EER as a function of the wavelet orientation (μ) for all frequencies (ν) for FRAV2D Database (left to right, top to bottom: test 1 to 6).

Figures 7 and 8 show the influence of the wavelet orientation μ. Only the high frequency wavelets have been considered ($0 \leq \nu \leq 2$), as we have seen they are more discriminant. For the XM2VTS Database, the wavelet with the lowest EER is achieved with orientation μ=2 for test 1. Despite the face features can be horizontal (eyebrows, eyes, nostrils, mouth) or vertical (nose), the most influential wavelet extracts information from the lower left to the upper right corner of the image. On the contrary, in test 6, the best wavelet is the one with μ=0, which clearly coincides with the illumination gradient direction. With respect to FRAV2D Database, except for test 2 (with 15º head orientation), oblique wavelets (μ=2, 3, 5, 6) usually have more discriminant power compared to horizontal (μ=4) or vertical wavelets (μ=0).

Another interesting conclusion is that the distribution of the EER as a function of μ is not symmetrical with respect to the central wavelet μ=4, despite the symmetry of a pair of wavelets with parameters μ and 8–μ. The exception is test 2 (images with 15º head orientation). Bearing in mind that the images in the database have been corrected from tilt, this can be understood as evidence that faces are not perfectly symmetrical. Our results seem to agree with those of [19, 20], which state that asymmetrical faces are easier to recognize than their symmetrical counterparts. Specifically, we have seen that some wavelet orientations produce a lower EER compared to the corresponding symmetrical ones, which means that in some cases the left half of the face carries more discriminant information that the right half, or vice versa.

6 Conclusions

We have presented the results of a thorough study of the so-called Parallel Gabor PCA algorithm for XM2VTS and FRAV2D Databases. Our algorithm outperforms other methods, such as PCA, FGPCA and DGPCA, when fewer PCA coefficients are kept. It has also obtained the best EER in three out of six experiments. When it ranked second, the final EER was only slightly worse compared to DGPCA. However, for images with significant head orientation, DGPCA is clearly the most effective.

In a one-Gabor PCA scenario we have seen that the features extracted by the high frequency (0≤v≤2) and oblique orientations (45º – 135º) wavelets are the most discriminant, as they have achieved the lowest EER. The different performance of wavelets and their mirrored equivalents shows that faces are not perfectly symmetrical and that those asymmetries carry more discriminant information. The experiments performed with the images with lateral lighting also show that the optimal wavelet is the one with a wave vector oriented in the illumination direction.

Acknowledgments. Part of this work has been supported by the Universidad Rey Juan Carlos under the financial program of mobility for teaching staff. Special thanks have also to be given to Professor Ian Dryden from the School of Mathematical Sciences of the University of Nottingham, UK, for his interesting comments.

References

1. Chellappa, R., Wilson, C.L., Sirohey, S.: Human and Machine Recognition of Faces: A Survey. Proceedings of the IEEE, vol. 83, Issue 5, pp. 705–740 (1995).
2. Zhao, W., Chellappa, R., Phillips, J., Rosenfeld, A.: Face Recognition: A Literature Survey. ACM Computing Surveys, pp. 399–458 (2003).
3. Turk, M., Pentland, A.: Eigenfaces for Recognition. Journal of Cognitive Neuroscience, vol. 3, issue 1, pp. 71–86 (1991).
4. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. IEEE PAMI, vol. 19, no. 7, pp. 711–720 (1997).
5. Daugman, J.G.: 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 (1985).
6. Lades, M., Vorbrüggen, J.C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R.P., Konen, W.: Distortion invariant object recognition in the Dynamic Link Architecture. IEEE Transactions on Computers, vol. 42, issue 3, pp. 300–311 (1993).
7. Wiskott, L., Fellous, J.M., Kruger, N., von der Malsburg, C.: Face recognition by Elastic Bunch Graph Matching. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, issue 7, pp. 775–779 (1997).
8. Shen, L., Bai, L.: A review on Gabor wavelets for face recognition. Pattern Analysis & Applications, vol. 9, issues 2–3, pp. 273–292 (2006).
9. Chung, K.-C., Kee, S.C., Kim, S.R.: 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 (1999).
10. Liu, C.J., Wechsler, H.: 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 (2002).
11. Shen, L., Bai, L.: Face recognition based on Gabor features using kernel methods. 6th IEEE Conference on Face and Gesture Recognition, pp. 170–175 (2004).
12. Gabor wavelets and general discriminant analysis for face identification and verification, Journal of Image and Vision Computing, 27 (2006) 1758-1767.
13. Qin, J., He, Z.-S.: A SVM face recognition method based on Gabor-featured key points. 4th International Conference on Machine Learning and Cybernetics, vol. 8, pp. 5144–5149 (2005).
14. Gilbert, C., Bakan, P.: Visual Asymmetry in Perception of Faces. Neuropsychologia, vol. 11, issue 3, pp. 355–362 (1973).
15. Rhodes, G.: Perceptual Asymmetries in Face Recognition. Brain and Cognition, vol. 4, issue 2, pp. 197–218 (1985).
16. Mitra, S., Lazar, N.A., Liu, Y.: Understanding the role of facial asymmetry in human face identification. Journal Statistics and Computing, vol. 17, issue 1, pp. 57–70 (2007).
17. Burt, D.M., Perrett, D.I.: Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. Neuropsychologia, vol. 35, issue 5, pp. 685–693 (1997).
18. Fink, B., Neave, N., Manning, J.T., Grammer, K.: Facial symmetry and judgements of attractiveness, health and personality. Personality And Individual Differences, vol. 41, issue 3, pp. 491–499 (2006).
19. Tjan, B.S., Liu, Z.L.: Symmetry impedes symmetry discrimination. Journal of Vision, vol. 5, issue 10, pp. 888–900 (2005).
20. Brady, N., Campbell, M., Flaherty, M.: Perceptual asymmetries are preserved in memory for highly familiar faces of self and friend. Brain and Cognition, vol. 58, issue 3, pp. 334–342 (2005).
21. Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G.: XM2VTS: The Extended M2VTS Database. 2nd International Conference on Audio and Video-based Biometric Person Authentication, pp. 72–77 (1999).
22. FRAV2D Database, 2004. Freely available from: http://www.frav.es/databases/frav2d/
23. Serrano, Á., Conde, C., Martín de Diego, I., Cabello, E., Bai, L., Shen, L.: Parallel Gabor PCA with Fusion of SVM Scores for Face Verification. International Conference on Computer Vision Theory and Applications, pp. 149–154 (2007).
24. Serrano, Á., Martín de Diego, I., Conde, C., Cabello, E., Bai, L., Shen, L.: Fusion of Support Vector Classifiers for Parallel Gabor Methods Applied to Face Verification. Haindl, J. Kittler, and F. Roli (Eds.): MCS 2007. LNCS, vol. 4472, pp. 141–150. Springer, Heidelberg (2007).
25. Vapnik, V.N.: The Nature of Statistical Learning Theory, Springer Verlag (1995).