Steerable Pyramids Feature Based Classification Using Fisher Linear Discriminant for Face Recognition

EL AROUSSI MOHAMED[1]
EL HASSOUNI MOHAMMED[1]2
GHOUZALI SANAA[1]
RZIZA MOHAMMED[1]
ABOUTAJDINE DRISS[1]

[1]GSCM-LRIT, Faculty of Sciences, Mohammed V University-Agdal, Rabat, Morocco
[2]DESTEC, FLSHR Mohammed V University-Agdal, Rabat, Morocco
PO.Box 1014, Rabat, Morocco
moha387@yahoo.fr

Abstract. In this paper, an efficient local appearance feature extraction method based the multi-resolution Steerable Pyramids (SP) transform is proposed in order to further enhance the performance of the well known Fisher Linear Discriminant (FLD) method when applied to face recognition. Each face is described by a subset of band filtered images containing block-based SP coefficients. These coefficients characterize the face texture and a set of simple statistical measures allows us to form compact and meaningful feature vectors. The proposed method is compared with some related feature extraction methods such as Principal component analysis (PCA), as well as Linear Discriminant Analysis, and Fisher Linear Discriminant (FLD), Independent Component Analysis and ICA. Experimental results on ORL, YALE and FERET face databases convince us that the proposed method provides a better representation of the class information and obtains much higher recognition accuracies.

Keywords: Steerable Pyramids, FLD, face recognition, multi-resolution.

(Received April 01, 2009 / Accepted July 04, 2009)

1 Introduction

In the last two decades, several studies have proposed to deploy multi-resolution feature extraction algorithms in face recognition. Among multi-resolution algorithms, the most popular are Discrete Wavelet transform (DWT), Gabor wavelets [19], Contourlet[4] and Curvelet transforms [12, 1]. These approaches have proved to be very successful to capture more discriminant features of face images allowing to achieve good performance and robustness against various challenging conditions such as variations in pose, lighting and expression. Steerable pyramid is another muli-resolution transform similar to the two-dimensional DWT, but with interesting translation- and rotation-invariance properties [2]. Several studies have investigated the discriminating power of steerable pyramid-based features (SP) in various applications including: image denoising [18], textures classification [11], image processing [9, 15] and face hallucination [10].

In this paper, we present a novel face recognition approach based on steerable pyramid decomposition. In order to capture multi-orientation information in face images better, a straightforward solution is calculating derivatives in different directions. Therefore, each face image is described by a subset of band filtered images containing steerable pyramid coefficients which characterize the face textures. We divide the SP sub-bands into small sub-blocks, from which we extract compact and meaningful feature vectors using simple statistical measures. These feature sets are used in order to create templates with different information content for face recognition (SP database). Once done, a Fisherface algorithm is carried out on the SP database in which faces with similar statistics will be grouped together by FLD rules, where the difference between classes is maximized while the difference within classes is minimized. For the purpose of classification, we use the city-block distance. We design experiments specifically to investigate the improvement in robustness against illumination and facial expression changes. Experimental results are presented using images from the FERET, ORL and the YALE databases. The efficiency of our approach is analyzed by comparing the results with those obtained using the well-known subspace reduction based methods PCA, LDA, FLD and ICA.

The remainder of the paper is organized as follows. In Section 2, steerable pyramid transform used in the study is explained. Section 3 describes the computation of the proposed face representation in detail and how to recognize faces based on SP with FLD are presented in Section 4 Experimental results are presented and discussed in Section 5. Finally, in Section 5, conclusions and future recommendations are given.

2 Steerable pyramid Face Representation

In signal processing, a signal can be decomposed into subbands, such as by wavelet transform. The wavelet transform is widely used in many applications including a retrieval system, since the pyramid structure of wavelets responds well to human visual system. However, one drawback of wavelets (orthogonal) is the lack of translation invariance especially in two-dimensional (2-D) signals [7]. To overcome this problem, the steerable pyramid wavelet, a class of arbitrary orientation filters generated by linear combination of a set of basis filters, has been proposed [7]. A face image of a person contains similarity (approximation) information of the face as well as discriminatory (detail) information with respect to faces of all other persons. The discriminatory information is due to structural variations of the face which are acquired as intensity variations at different locations of the face. The location and degree of intensity variations in a face for an individual are unique features which discriminate one person from the rest of the population. Steerable pyramid (SP) decomposition can be used to split the features in a face image into different sub-bands at different levels, with 'approximations' and 'details'. Since a one-level decomposition may not be adequate to effectively isolate these pair of visual features, it is necessary to explore different combination of sub-bands at higher levels to obtain a suitable isolation. In order to capture multi-orientation information in face images better, a straightforward solution is calculating derivatives in all directions, but this method can present high computation cost. Based on the theorem of steerable filter [15], the derivatives of an image in any direction can be interpolated by several basis derivative functions.

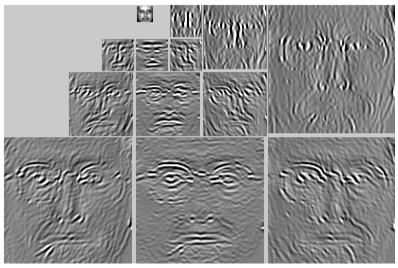

Figure 1: 3-stage & 4-orientation steerable pyramid transform.

Figure 1, shows the analysis and synthesis representation of the steerable pyramid transform. A face image is decomposed into a steerable pyramid by four oriented third-order band-pass basis filters. In the first level, four sub-band images are obtained. In this figure, we can see that each oriented filter is most sensitive to the oriented information (e.g. edges) that is perpendicular to the direction of filter. The steerable pyramid combines facial image's spatial multi-scale features with multi-orientation local features. These features are exactly perceptible by V1 area (the first visual area) of human visual cortex. Therefore it is reasonable that we choose steerable pyramid as local low-level features for face images.

3 Proposed method

3.1 Feature vectors

To generate the image database, each image is decomposed into 3- level and 4-orientation sub-bands. The direct use of S-P coefficients may not extract the most discriminative features as these coefficients contain much redundant and irrelevant information. For an efficient and local representation of the face image, first each S-P sub-band is partitioned into a set of equally-sized blocks in a non-overlapping way. Based on common belief, the statistical measures such as mean, variance and entropy of the energy distribution of the S-P coefficients for each sub-band at each decomposition level can be used to identify a texture. Let $I_{ij}(x, y)$ be the image at the specific block j of sub-band i, the resulting feature vector $\nu_{ij} = \{\mu_{ij}, \sigma_{ij}^2, e_{ij}\}$, where μ_{ij} =mean, σ_{ij}^2 =variance and e_{ij} =entropy.

$$\mu_{ij} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} |I_{ij}(x,y)| \qquad (1)$$

$$\sigma_{ij} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} |I_{ij}(x,y) - \mu_{ij}|^2 \qquad (2)$$

where M and N is the size of $I_{ij}(x,y)$. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy is defined as

$$e_{ij} = -\sum_{p} (p \times \log(p)) \qquad (3)$$

where p contains the histogram counts.

The feature vector of a face is then constructed by concatenating each block measure to one big feature vectors $V = \bigcup_{i=1}^{k} \bigcup_{j=1}^{k_i} \{\nu_{ij}\}$, k is the number of S-P sub-bands and k_i the number of blocks in the ith sub-band. Therefore, we can extract the best features and reduce the size of the data while keeping only the principal discriminant features. Figure 2 shows the overall diagram of the proposed face features extraction.

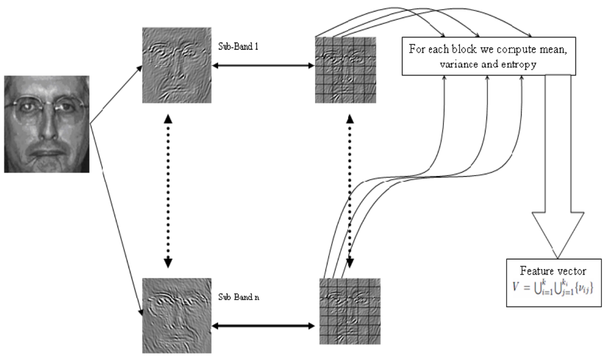

Figure 2: Diagram of the Block-based S-P features extraction process

3.2 Fisher Linear Discriminant FLD

FLD is a popular discriminant criterion that measures the between-class scatter normalized by the within-class scatter [8]. Let $\omega_1, \omega_2, ..., \omega_L$ and $N_1, N2, ..., N_L$ denote the classes and the number of images within each class, respectively. Let $M_1, M_2, ..., M_L$ and M be the means of the classes and the grand mean. The within- and between-class scatter matrices, Σ_ω and Σ_b, are defined as follows:

$$\Sigma_\omega = \sum_{i=1}^{L} P(\omega_i) \varepsilon \{(\gamma^\rho - M_i)(\gamma^\rho - M_i)^t | \omega_i\} \qquad (4)$$

and

$$\Sigma_b = \sum_{i=1}^{L} P(\omega_i)(M_i - M)(M_i - M)^t \qquad (5)$$

where $P(\omega_i)$ is a priori probability, $\Sigma_\omega, \Sigma_b \in R^{m \times m}$, and L denotes the number of classes. FLD derives a projection matrix Ψ that maximizes the ratio $|\Psi^t \Sigma_b \psi| / |\Psi^t \Sigma_\omega \psi|$ [3]. This ratio is maximized when Ψ consists of the eigenvectors of the matrix $\Sigma_\omega^{-l} \Sigma_b$ [16]:

$$\Sigma_\omega^{-l} \Sigma_b \Psi = \Psi \Delta \qquad (6)$$

where $\Psi, \Delta \in R^{m \times m}$ are the eigenvector and eigenvalue matrices of $\Sigma_\omega^{-l} \Sigma_b$, respectively.

Concatenating 2D matrices into 1D vectors leads to very high dimensional nature of image vector, where it is difficult to evaluate the scatter matrices accurately due to its large size. Furthermore, the within-class scatter matrix is always singular, making the direct implementation of FLD algorithm an intractable task. In order to make FLD approach more efficient in this study, we have proceeded the face images locally using a block-based SP feature extraction which renders small the size of the feature vectors.

3.3 classification

Once SP database constructed, a Fisherface algorithm is carried out on the training images taken randomly from the SP database in order to reduce the size of the data while keeping only the principal discriminant features (principal components). Then, the test images are projected on the Eigenspace and the distances to the training images are computed using the city-block distance for the purpose of classification as follows: For two face images p and q, let Vp and Vq representing the corresponding feature vectors, the distance d_{pq} between the two patterns in the feature space is defined as:

$$d_{pq} = \sqrt{\sum_{i=1}^{k} \sum_{j=1}^{k_i} |\nu_{pij} - \nu_{qij}|} \qquad (7)$$

The classification performance evaluation is based on pairwise distance matrix. If there are m training and n test samples, then a distance matrix should be of size $m \times n$, with each column representing the distances from the corresponding test sample to all training samples (classes). The lower the distance, the closer the two samples.

4 Experimental Results

4.1 Databases

To validate the accuracy of the proposed algorithm, we have used three databases: ORL[1], Yale[2] and FERET [14]. The ORL databse contains ten different images of 40 distinct subjects in up-right, frontal position with tolerance for some tilting and rotation of up to 20 degrees. Moreover, the most variation of some image scale is close to 10%. Therefore, it is expected that this is a more difficult database to work with. 5 face images per person are chosen randomly as training images while the remaining 5 images are set as test images. Figure 3 depicts some sample images from the ORL database.

Figure 3: Faces from the ORL Face Database

The Yale face database consists of 15 individuals, where for each individual, there are 11 face images containing variations in illumination and facial expression. From these 11 face images, we use 5 for training, chosen randomly. The remaining 6 images are used for testing. Figure 4 depicts some sample images from the Yale database.

Figure 4: Faces from the YALE Face Database

A subset of FERET face database, fafb image set, containing images of 145 individuals is used in our experiments. In this subset, there are four frontal views of each individual: a neutral expression and a change of expression from one session, and a neutral expression and change of expression from a second session that occurred three weeks after the first. For each of the individual in the set, three of their images are used for training and the remaining is used for testing purposes. Figure 5 depicts some sample images from the FERET database.

Figure 5: Faces from the FERET Face Database

All the images are aligned with respect to the manually detected eye coordinates, scaled to 128×128 pixels resolution. Extensive experiments have been conducted on these databases to determine the optimal block size for features extraction. We have found out that block of 16×16 pixels attains best results.

The experimental data we used to test the performances of sub-bands against expression changes consists of ORL database. To test the performance of SP sub-bands against illumination variations, we used Yale database. The performance has been measured by Cumulative rank error (error when we consider whether the correct identity is among the best n classifier results) [14]. This measure is the most used in literature to show the performance accuracy of different face recognition algorithms and hence is used in this paper for comparison purpose.

4.2 Results and Analysis

In order to assess the efficiency of the proposed technique described above, we carried out a series of experiments using all databases.

For comparison purpose, we use PCA, LDA, FLD, ICA implemented in Statistical Learning Toolbox[3]

Experiment1 In order to establish the reliability of the proposed method we have compared our proposed method(SP-FLD) against well established existing techniques like standard eigenface based methods PCA, LDA, FLD and ICA this section. Table 1 reports the results obtained for ORL,YALE and FERET databases. It is clear from the Table 1 that the proposed method enhances significantly the original PCA, LDA, FLD and ICA algorithms, and this on all databases. For the ORL Database for instance, improvements of 10%, 23.5 and 11% have been obtained for the PCA,LDA and FLD methods, respectively. It is also worth mentioning that significant enhancements have been obtained for different lighting conditions in YALE database since we have

[1]http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.zip

[2]http://cvc.yale.edu/

[3]http://www.mathworks.com/matlabcentral/fileexchange/12333

obtained 97.78% for the proposed method against 90% for PCA, 95.56 for LDA and 92.22% for FLD which mean that SP-FLD is robust against significant variation in illumination and facial details (present in Orl and YALE) as well. Finally for the FERET database the improvement is about 22.07%, 25.52 and 24.83% compared with PCA, LDA and FLD respectively.

The comparative face recognition performance of these three methods is shown in Figure 6, that illustrate the variation of recognition rate compared with selected principal components. We can see from the figure that the SP-FLD method performs better than the PCA, LDA, FLD and ICA method followed by the Eigenfaces method.

Table 1: Experiment results for the proposed method(SP-FLD) and comparison with the original PCA, LDA, FLD and ICA algorithm on ORL, Yale and FERET Databases.

Method	PCA (%)	LDA (%)	FLD (%)	ICA (%)	SP-FLD (%)
ORL	88.5	75	87.5	80.5	**98.5**
YALE	90	95.56	92.22	87.78	**97.78**
FERET	75.86	72.41	73.1	74.48	**97.93**

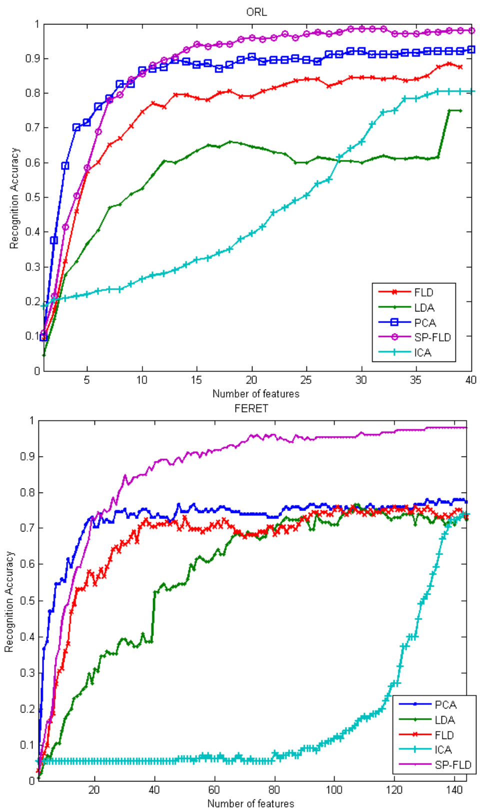

Figure 6: Comparative face recognition performance of the combination of Steerable Pyramids and FLD, PCA, LDA, FLD and ICA on ORL, and FERET Databases

Experiment2 The objective of this experiment is to investigate the effect of the number of images per person available in the ORL and databases on the recognition rates with a view to compare it against some similar algorithms. As illustrated in Figure 7, independently on the database size, the SP-FLD method outperforms the PCA and LDA. Using only three images for training samples the SP-FLD method achieves recognition accuracy of 92% for ORL and 94% for YALE. The SP-FLD method attains better performances using five training images compared to traditional face recognition approaches like PCA and LDA as published in the literature.

Experiment3 In this section, we aim to demonstrate that a combination of the SP Transform with FLD yields improved recognition results when compared against some other well known classifiers available in pattern recognition such as PCA, and LDA. figure 8 shows the results of the same experiments as described in **experiment1** above but with comparison against SP-LDA and SP-PCA methods.

Experiment4 In this section we compare our proposed method against well established existing techniques as presented in [13] like standard eigenface based methods [17], wavelet [6], Curvelet based methods [13], wavelet based kernel Associative Memory (kAM) [20], wavelet based weighted modular PCA [21] and discriminant waveletfaces using Nearest Feature Line (NFL) classifier [5]. Table 2 reports the results obtained for ORL, YALE Database.

The results demonstrate that high recognition accuracy can be achieved using SP-FLD approach for face recognition. Our technique has been found to be robust against significant variation in illumination and facial details (present in ORL and Yale) as well. For all the databases, most prominently for Yale database , the proposed framework shows significant performance improvement over the other schemes. When compared to the best performing method in the table (Curveletface + PCA + LDA[13]), SP-FLD features work the best and show 0.8% and 5.78% gain in accuracy for ORL and YALE respectively.

5 CONCLUSIONS AND FUTURE WORKS

This paper proposes a new approach for face recognition based on exploiting the features of the SP Transform. For each face image, SP is performed to com-

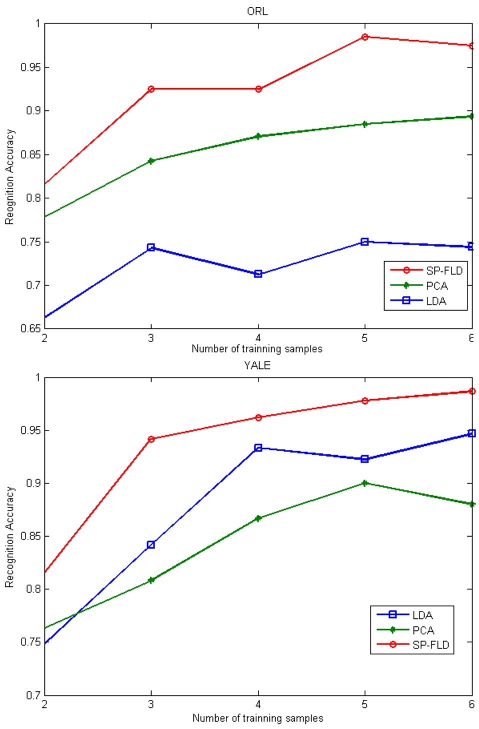

Figure 7: Recognition Accuracy compared with number of training images on ORL and YALE databases

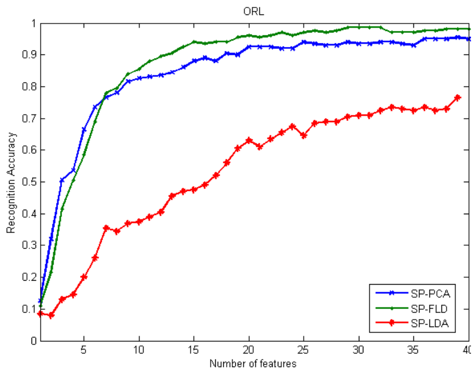

Figure 8: Recognition Accuracy compared with number of training images

pute different sub-bands from which FLD projection are conducted. Different databases (ORL, Yale and FERET) have been used to evaluate the proposed method. The technique introduced in our paper appears to be robust to changes in facial expression as it shows good results for the ORL and FERET databases, and to the lighting variation since good results are obtained for Yale database. SP-FLD transform is able to capture multi-directional features which make it to be very effective in the face recognition. Our future work would include applying feature selection method to extract the most discriminative SP sub-bands and to include different statistical measures giving higher classification performance.

Table 2: Comparative Study.

Method	ORL	YALE
Standard Eigenface [17]	92.2%	76%
Waveletface[6]	92.5%	83.3%
Waveletface + PCA[6]	94.5%	84%
Waveletface + LDA[5]	94.7%	84.6%
Waveletface + Weighted Modular PCA[21]	95%	83.6%
Waveletface + LDA + NFL[5]	95.2%	83.5%
Waveletface + KAM[20]	96.6%	84%
Curveletface + PCA[13]	96.6%	83.9%
Curveletface + PCA + LDA[13]	97.7%	92%
SP-FLD	**98.5%**	**97.78%**

References

[1] Aroussi, M. E., Ghouzali, S., Hassouni, M. E., .Rziza, M., and Aboutajdine, D. Curvelet-based feature extraction with b-lda for face recognition. *The 7th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*, pages 444–448, 2009.

[2] Bamberger, R. H. and Smith, M. J. T. A filter bank for the directional decomposition of images: Theory and design. *IEEE Trans.*, 40(4):882–893, 1992.

[3] Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J. Eigenfaces vs fisherfaces. *IEEE Transaction Pattern Analysis Machine Intelligence*, 19(7):711Ű720, 1997.

[4] Boukabou, W. R. and Bouridane, A. Contourlet-based feature extraction with pca for face recognition. *it In: NASA/ESA Conference on Adaptive Hardware and Systems 2008, IEEE*, 2008.

[5] Chien, J. T. and Wu, C. C. Discriminant waveletfaces and nearest feature classifiers for face recognition. *IEEE Trans. PAMI*, 24(2):1644–1649, 2002.

[6] Feng, G. C., D, D. Q., and Yuen, P. C. Human face recognition using pca on wavelet subband. *Journal of Electronic Imaging,*, 9(2):226–233, 2000.

[7] Freeman, W. T. and Adelson, E. H. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.1991*, 13(9):891–906, 2000.

[8] Fukunaga, K. *Introduction to Statistical Pattern Recognition*. New York: Academic, 2nd ed. edition, 1991.

[9] Huanga, L., Wu, F., Su, C., and Zhuanga, Y. Steerable pyramid-based face hallucination. *Pattern Recognition*, 38:813–824, 2005.

[10] Huanga, L., Wu, F., Su, C., and Zhuanga, Y. Steerable pyramid-based face hallucination. *Pattern Recognition*, 38:813–824, 2005.

[11] Li, S., , and Taylor, J. S. Comparison and fusion of multiresolution features for texture classification. *Pattern Recognition Letters*, 25, 2002.

[12] Mandal, T., Majumdar, A., and Wu, Q. M. J. Face recognition by curvelet based feature extraction. *ICIAR, LNCS 4633*, pages 806–817, 2007.

[13] Mandal, T., Wu, Q. M. J., and Yuan, Y. Curvelet based face recognition via dimension reduction. *Signal Processing*, 2009.

[14] Phillips, P. J., Moon, H., Rizvi, S. A., and Rauss, P. J. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1090–1104, 2000.

[15] Simoncelli, E. P. A rotation-invariant pattern signature. *3rd IEEE Int'l Conf on Image Processing. Laussanne Switzerland, Sept 1996*, 1996.

[16] Swets, D. L. and Weng, J. Using discriminant eigenfeatures for image retrieval. *IEEE Transaction Pattern Analysis Machine Intelligence*, 18:831Ű836, 1996.

[17] Turk, M. and Pentland, A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience,*, 3:71–86, 1991.

[18] Wainwright, M., Simoncelli, E. P., Portilla, J., and Strela, V. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Trans Image Processing*, 12:1338–1351, 2003.

[19] Wiskott, L., Fellous, J. M., Kruger, N., and Malsburg, C. V. D. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, 19:775–779, 1997.

[20] Zhang, B. L., Zhang, H., and Ge, S. S. Face recognition by applying wavelet subband representation and kernel associative memory. *IEEE trans. neural networks,*, 15(1):166–177, 2004.

[21] Zhao, M., L, Z., and Li, P. Face recognition based on wavelet transform weighted modular pca. *Proc. Congress in Image and Signal Processing*, 2008.