

Face Recognition Using Principal Component Analysis and Wavelet Packet Decomposition

Vytautas PERLIBAKAS

*Image Processing and Analysis Laboratory
Kaunas University of Technology
Studentų 56–305, 51424 Kaunas, Lithuania
e-mail: vperlib@mmlab.ktu.lt*

Received: October 2003

Abstract. In this article we propose a novel Wavelet Packet Decomposition (WPD)-based modification of the classical Principal Component Analysis (PCA)-based face recognition method. The proposed modification allows to use PCA-based face recognition with a large number of training images and perform training much faster than using the traditional PCA-based method. The proposed method was tested with a database containing photographs of 423 persons and achieved 82–89% first one recognition rate. These results are close to that achieved by the classical PCA-based method (83–90%).

Key words: face recognition, PCA, Wavelet Packet Decomposition, WPD.

1. Introduction

Principal Component Analysis (PCA)-based face recognition method was proposed in (Turk, 1991) and became very popular. Using this method we find a subset of principal directions (principal components) in a set of the training faces. Then we project faces into the space of these principal components and get the feature vectors. Face recognition is performed by comparing these feature vectors using different distance measures. Using the PCA-based face recognition method we calculate the eigenvectors and eigenvalues of the covariance matrix of the training data. If this matrix is large, calculation of eigenvectors becomes complicated. In order to solve this problem we can use the decomposition of the covariance matrix (Kirby, 1990), incremental eigenspace learning (Chandrasekaran, 1997; Skocaj, 2002), operations with eigenspaces (Hall, 1999). Also we can choose a small number of representative training images (Moghaddam, 1994), split face images into small pieces (Li, 2002) or use other transforms, e.g., DCT (Hafed, 2001). In this article we propose a novel Wavelet Packet Decomposition (WPD)-based modification of the classical Principal Component Analysis (PCA)-based face recognition method. The proposed method allows to use PCA-based face recognition with a large number of training images. Using the proposed method with large databases the training is performed much faster than using the traditional PCA-based face recognition. Recognition experiments were performed using the database containing photographs of 423 persons. The

experiments showed, that using the proposed method we can achieve 82–89% first one recognition rate. These results are close to that achieved by the classical PCA-based face recognition method (83–90%).

2. PCA-Based Face Recognition

In this section we will present brief description of the PCA-based face recognition method, which details could be found in (Groß, 1994).

Let X_j be N -element one-dimensional image and suppose that we have r such images ($j = 1, \dots, r$). A one-dimensional image-column X from the two-dimensional image (face photography) is formed by scanning all the elements of the two-dimensional image row by row and writing them to the column-vector. Then mean vector, centered data vectors and covariance matrix are calculated:

$$m = \frac{1}{r} \sum_{j=1}^r X_j, \quad d_j = X_j - m, \quad C = \frac{1}{r} \sum_{j=1}^r d_j d_j^T,$$

here $X = (x_1, x_2, \dots, x_N)^T$, $m = (m_1, m_2, \dots, m_N)^T$, $d = (d_1, d_2, \dots, d_N)^T$. Principal axes are found by calculating eigenvectors u_k and eigenvalues λ_k of the covariance matrix C ($Cu_k = \lambda_k u_k$). Because the dimensionality (N^2) of the matrix C is large even for a small images, and computation of eigenvectors using traditional methods is complicated, dimensionality of matrix C is reduced using the decomposition described in (Kirby, 1990) (if the number of training images is smaller than the number of image pixels). Found eigenvectors $u = (u_1, u_2, \dots, u_N)^T$ are normed, sorted in decreasing order according to the corresponding eigenvalues, transposed and arranged to form the row-vectors of the transformation matrix T . Now any data X can be projected into the eigenspace using the following formula:

$$Y = T(X - m), \quad (1)$$

here $X = (x_1, x_2, \dots, x_N)^T$, $Y = (y_1, y_2, \dots, y_r, 0, \dots, 0)^T$.

For projection into the eigenspace we can use not all eigenvectors, but only a few of them, corresponding to the largest eigenvalues (Swets, 1998). When the image is projected into the eigenspace we get its eigenfeature vector $Z = (z_1, z_2, \dots, z_n)^T = (y_1, y_2, \dots, y_n)^T$; here n is the number of features. Recognition is performed by calculating distances ε_i between feature vectors Z_i of the known faces and feature vector Z_{new} of a new unknown face. Then we say that the face with projection Z_{new} belongs to a person $s = \arg \min_i [\varepsilon_i]$, or say that it is unknown if $\varepsilon_s \geq \tau$, here τ – rejection threshold. In order to achieve good recognition results we must choose an appropriate distance measure (Navarrete, 2002). As it was shown in (Perlibakas, 2003), one of the best results with respect to the first one recognition rate are achieved using weighted angle-based distance:

$$d(X, Y) = - \frac{\sum_{i=1}^n z_i x_i y_i}{\sqrt{\sum_{i=1}^m x_i^2 \sum_{i=1}^m y_i^2}}, \quad (2)$$

here X, Y – eigenfeature vectors of length n , $z_i = \sqrt{1/\lambda_i}$, λ_i – corresponding eigenvalues.

3. Wavelet Packet Decomposition

Using the classical wavelet decomposition, the image is decomposed into the approximation and details images, the approximation is then decomposed itself into a second level of approximation and details and so on (Press, 1992). Wavelet Packet Decomposition (WPD) is a generalization of the classical wavelet decomposition and using WPD we decompose both approximations and details into a further level of approximations and details. Theoretical backgrounds of the wavelet transform could be found in (Daubechies, 1992; Strichartz, 1994), comprehensive description of the computerised realisation and source code could be found in (Press, 1992). We will present only the main ideas related to the practical implementation. Using WPD we decompose the two-dimensional initial image A_0^0 (level $l = 0$) into approximation A_0^1 , horizontal details $D_{0,h}^1$, vertical details $D_{0,v}^1$ and diagonal details $D_{0,d}^1$ at level $l = 1$. In order to get decomposition at level l we decompose approximations A_i^{l-1} and details $D_{i,h}^{l-1}$, $D_{i,v}^{l-1}$, $D_{i,d}^{l-1}$ into the following approximations and details:

$$A_i^{l-1} \rightarrow \{A_{4i}^l; D_{4i,h}^l; D_{4i,v}^l; D_{4i,d}^l\}, \quad l > 0, \tag{3}$$

$$D_{i,h}^{l-1} \rightarrow \{A_{4i+1}^l; D_{4i+1,h}^l; D_{4i+1,v}^l; D_{4i+1,d}^l\}, \quad l > 1, \tag{4}$$

$$D_{i,v}^{l-1} \rightarrow \{A_{4i+2}^l; D_{4i+2,h}^l; D_{4i+2,v}^l; D_{4i+2,d}^l\}, \quad l > 1, \tag{5}$$

$$D_{i,d}^{l-1} \rightarrow \{A_{4i+3}^l; D_{4i+3,h}^l; D_{4i+3,v}^l; D_{4i+3,d}^l\}, \quad l > 1, \tag{6}$$

here $i = 0, \dots, (4^{(l-1)} - 1)$. And at level l we have a set of approximations and details $\{A_i^l; D_{i,h}^l; D_{i,v}^l; D_{i,d}^l\}$. In our experiments we decomposed images into two levels using Haar wavelets. The decomposition tree is shown in Fig. 1. Approximations and details are calculated using low-pass and high-pass decomposition filters and dyadic downsampling (Strang, 1996):

$$A_0^0 \rightarrow LoD_{rows} \rightarrow 2 \downarrow 1 \rightarrow LoD_{cols} \rightarrow 1 \downarrow 2 \rightarrow A_0^1, \tag{7}$$

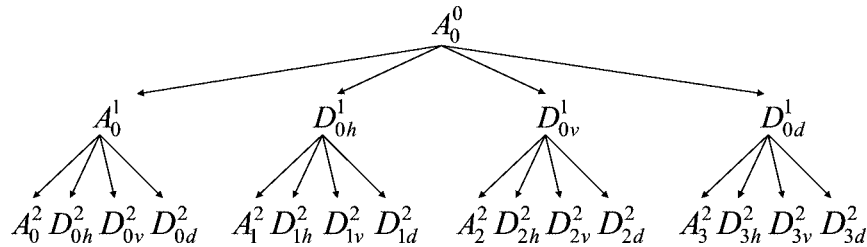


Fig. 1. Wavelet Packet Decomposition tree.

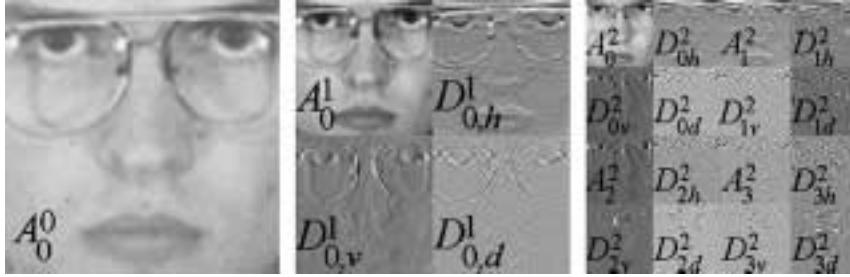


Fig. 2. WPD levels $l = 0$ (initial image), $l = 1$ and $l = 2$.

$$A_0^0 \rightarrow LoD_{rows} \rightarrow 2 \downarrow 1 \rightarrow HiD_{cols} \rightarrow 1 \downarrow 2 \rightarrow D_{0,h}^1, \quad (8)$$

$$A_0^0 \rightarrow HiD_{rows} \rightarrow 2 \downarrow 1 \rightarrow LoD_{cols} \rightarrow 1 \downarrow 2 \rightarrow D_{0,v}^1, \quad (9)$$

$$A_0^0 \rightarrow HiD_{rows} \rightarrow 2 \downarrow 1 \rightarrow HiD_{cols} \rightarrow 1 \downarrow 2 \rightarrow D_{0,d}^1, \quad (10)$$

here A_0^0 – two-dimensional input image, LoD_{rows} , LoD_{cols} , HiD_{rows} , HiD_{cols} – convolutions of rows and columns of the input two-dimensional image with low-pass and high-pass decomposition filters, $2 \downarrow 1$ – dyadic downsampling of the columns and keeping the even indexed columns (if the indexing starts from 1), $1 \downarrow 2$ – dyadic downsampling of the rows and keeping the even indexed rows (if the indexing starts from 1), A_0^1 – approximation at first level, $D_{0,h}^1$, $D_{0,v}^1$, $D_{0,d}^1$ – horizontal, vertical and diagonal details at first level. In order to get decomposition at level l we use the same steps and decompose approximations A_i^{l-1} and details $D_{i,h}^{l-1}$, $D_{i,v}^{l-1}$, $D_{i,d}^{l-1}$. In Fig. 2 we show an example of the WPD.

4. PCA and WPD

Because calculation of the eigenvectors and eigenvalues for a large number of training images is complicated (Hafed, 2001) even if the size of the training images is enough small (Groß, 1994), we propose to decompose initial image into k parts using the Wavelet Packet Decomposition and then perform PCA for k -times with the smaller training images (approximations and details). Time complexity of the eigenvectors calculation is $O(n^3)$, $n = \min(N, m)$, here N – number of image pixels, m – number of training images (Navarrete, 2002). Using the proposed method the time-complexity of the eigenvectors problem is $k \cdot O((N/k)^3) = O(N^3)/k^2$ and is independent from the number of training images. After the training we get k eigenspaces, and in order to perform recognition we must decompose images into k parts, project them into k eigenspaces and get k feature vectors. For face recognition experiments we merged feature vectors into one vector, selected features corresponding to the largest eigenvalues of the merged and sorted eigenvalues vector and calculated the selected distance measure between these merged feature vectors.

5. Experiments and Results

For recognition experiments we collected photographs of 423 persons (2 images per person – 1 for learning and 1 for testing) from 9 databases (Perlibakas, 2003). We manually selected the centers of eyes and lips in order to avoid recognition errors related to incorrectly detected faces. Then we rotated images to make the line connecting eye centers horizontal, resized the images to make the distances between the centers of the eyes equal to 26 pixels, calculated center of the face using the centers of eyes and lips, cropped 64x64 central part of the face, performed histogram equalization on the cropped part of the image. In all the experiments we use the same templates and change only the recognition method and the number (percent) of used features.

At first we measured recognition accuracy of the proposed method and compared it with the classical PCA-based face recognition method. In these experiments $m = 423$, $N = 64 \cdot 64 = 4096$, $k = 16$ ($l = 2$). Because our goal was to increase training speed of the PCA-based method, we also experimentally measured how many times and with what database sizes (m) the proposed WPD+PCA method is faster than the classical PCA-based method.

In order to measure and compare recognition capabilities of the methods, we used Cumulative Match Characteristic (CMC) and Receiver Operating Characteristic (ROC)-based measures, described in (Bromba, 2003): the area above Cumulative Match Characteristic (CMCA), how many images (in percents) must be extracted from the database in order to achieve some cumulative recognition rate (80–100%), Equal Error Rate (EER), the area below Receiver Operating Characteristic (ROCA), first one recognition rate. Graphical representation of the used characteristics is shown in Figs. 3 and 4.

The results of the recognition experiments are presented in Table 1. In order to measure the speed-up of the training process, we performed simulations with different m values and the results are presented in Table 2.

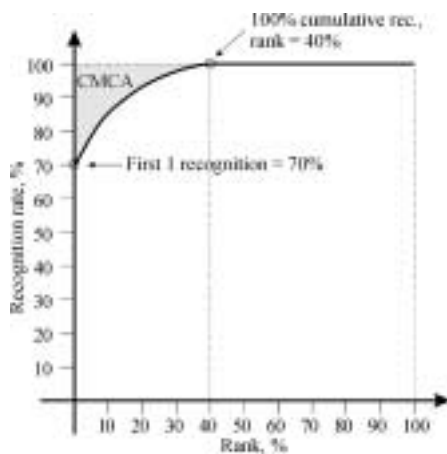


Fig. 3. CMC.

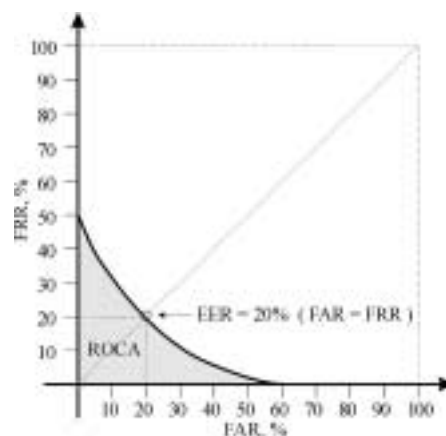


Fig. 4. ROC.

Table 1
Recognition using weighted angle-based distance

Method	Feat. num.	Rank, %				CMCA, [0...10 ⁴]	First 1 rec., %	EER, %	ROCA, [0...10 ⁴]
		85	90	95	100				
PCA (10%)	42	0.5	0.9	2.4	31.0	68.32	84.63	4.49	57.65
WPD+PCA	42	0.5	0.9	3.3	24.8	71.98	82.98	4.49	65.92
PCA (20%)	85	0.2	0.5	1.7	32.9	58.71	87.00	3.78	46.44
WPD+PCA	85	0.2	0.7	2.1	23.9	69.05	85.58	4.49	57.87
PCA (30%)	127	0.2	0.5	1.4	26.0	57.15	88.42	3.31	44.08
WPD+PCA	127	0.2	0.5	1.9	27.7	64.89	86.29	4.02	54.82
PCA (60%)	254	0.2	0.5	0.9	23.4	52.14	89.60	3.07	38.76
WPD+PCA	254	0.2	0.5	1.7	18.0	58.99	86.52	3.55	48.39
PCA (90%)	381	0.2	0.2	0.7	16.3	47.78	90.07	3.07	33.82
WPD+PCA	381	0.2	0.5	1.2	17.7	55.66	86.76	3.55	43.80
WPD+PCA	2000	0.2	0.5	1.2	21.5	50.47	89.13	3.31	37.40
WPD+PCA	3000	0.2	0.5	1.4	21.5	51.86	89.60	3.31	39.05

Table 2
Training speed-up

m	650	700	800	900	1000	1500	2000	2500	3000
Speed-up, times	1.0	1.3	1.9	2.7	3.6	11.5	25.9	50.9	86.3

Now we will compare our face recognition results with other wavelets-based methods. Feng (2000) used Daubechies 4 wavelets and PCA on the classical wavelet decomposition level 4 images. With the database containing photographs of 15 persons (4 images per person – 2 for learning, 2 for training) using PCA they achieved 78.78% recognition, using PCA on wavelet transform level 4 images they achieved 85.45% recognition accuracy. Garcia (2000) used WPD (level 2) and moments. Feature vectors, containing means and variances of the WPD approximations and details, were compared using the Bhattacharyya distance. For experiments they used two databases containing photographs of 200 and 155 persons and achieved 81.9% and 80.5% first one recognition rate. Using WPD+PCA we achieved larger first one recognition rates (82–89%) with larger database, but it must be noted that we selected face positions manually, Garcia (2000) used automatical face localization, Feng (2000) performed recognition without face detection. Our experiments also showed, that classical PCA-based face recognition method achieves 1–4% larger first one recognition rate than the proposed WPD+PCA-based method using the same number of features. With respect to other characteristics (CMCA, ROCA), the classical PCA-based method also performs slightly better than the proposed method using the same number of features. If we use larger number of features (e.g., 2000 and more), the results of the proposed method are very similar to the results of the classical PCA-based method.

The experiments with 64×64 training images and decomposition level $l = 2$ also showed, that if the number of training images is 700–3000, using the proposed method we can perform training 1.3–86.3 times faster than using the classical PCA-based method. If the number of training images is less than 650, training is performed faster using the classical PCA-based method.

6. Conclusions and Future Work

In this article we proposed Wavelet Packet Decomposition (WPD)-based modification of the classical Principal Component Analysis (PCA)-based face recognition method. The proposed modification of the PCA-based face recognition method could be used in practical applications when the number of training images is too large (thousands or more) for traditional PCA-based recognition and the training becomes too slow. The proposed method allows to use PCA-based recognition with large databases, because its training time is independent from the number of training images. Using the proposed method in practice, in order to add or remove training images we decompose them, update covariance matrices and mean vectors and then recalculate the eigenvectors using traditional methods. Recognition is performed by calculating feature vectors and comparing them using the selected distance measure. In order to speed-up recognition we can store feature vectors in the database and recalculate them (reindex the database) depending on the number of added and removed images. The proposed method was tested using the database containing photographs of 423 persons. The experiments showed, that using the proposed method we can achieve 82–89% first one recognition rate and the results are close to that achieved by the classical PCA-based face recognition method (83–90%).

In the future we are going to investigate different wavelet bases and feature selection methods in order to increase recognition accuracy.

References

- Bromba, M. (2003). *Biometrics FAQ*. <http://www.bromba.com/faq/biofaq.htm>.
- Chandrasekaran, S., B.S. Manjunath, Y.F. Wang, J. Winkeler and H. Zhang (1997). An eigenspace update algorithm for image analysis. *Graphical Models and Image Processing*, **5**(59), 321–332.
- Daubechies, I. (1992). *Ten lectures on wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics *61*. Society for Industrial and Applied Mathematics (SIAM).
- Feng, G.C., P.C. Yuen and D.Q. Dai (2000). Human face recognition using PCA on wavelet subband. *Journal of Electronic Imaging* (JEI), **2**(9), 226–233.
- Garcia, C., G. Zikos and G. Tziritas (2000). Wavelet packet analysis for face recognition. *Image and Vision Computing*, **18**, 289–297.
- Groß, M. (1994). *Visual Computing. The Integration of Computer Graphics, Visual Perception and Imaging. Computer Graphics: Systems and Applications*. Springer-Verlag.
- Hall, P., D. Marshall and R. Martin (1999). Adding and subtracting eigenspaces. In *10th British Machine Vision Conference*. pp. 453–462.
- Hafed, Z.M., and M.D. Levine (2001). Face recognition using the discrete cosine transform. *International Journal of Computer Vision*, **43**(3), 167–188.
- Kirby, M., and L. Sirovich (1990). Application of the Karhunen-Loeve expansion for the characterization of human faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (PAMI), **12**(1), 103–108.

- Li, Z., and X. Tang (2002). Eigenface recognition using different training data sizes. In *International Conference on Information Security*. Shanghai, China.
- Moghaddam, B., A. Pentland (1994). Face recognition using view-based and modular eigenspaces, MIT TR-301. In *Automatic Systems for the Identification and Inspection of Humans, SPIE, 2277*.
- Navarrete, P., and J. Ruiz-del-Solar (2002). Comparative study between different eigenspace-based approaches for face recognition. *Advances in Soft Computing (AFSS)*, In *International Conference on Fuzzy Systems, Lecture Notes in Computer Science, 2275*, 178–184.
- Perlibakas, V. (2003). Distance measures for PCA-based face recognition. *Information Technology and Control, 4(29)*, 67–74.
- Press, W.H., W.T. Vetterling, S.A. Tenkolsky and B.P. Flannery (1992). *Numerical Recipes in C. The Art of Scientific Computing*. Second edition. New York, Cambridge University Press.
- Skocaj, D., and A. Leonardis (2002). Incremental approach to robust learning of eigenspaces. In *Vision with Non-Traditional Sensors, 26th Workshop of the Austrian Association for Pattern Recognition*. pp. 111–118.
- Strang, G., and T. Nguyen (1996). *Wavelets and Filter Banks*. Wellesley-Cambridge Press.
- Strichartz, R.S. (1994). Construction of orthonormal wavelets. In J. Benedetto and M. Frazier (Eds.), *Wavelets: Mathematics and Applications*. CRC Press. pp. 23–50.
- Swets, D.L., Y. Pathak and J.J. Weng (1998). An image database system with support for traditional alphanumeric queries and content-based queries by example. *Multimedia Tools and Applications, 7*, 181–212.
- Turk, M., and A. Pentland (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience, 3(1)*, 71–86.

V. Perlibakas is a doctoral student at Image Processing and Analysis Laboratory of the Kaunas University of Technology. Research interests include digital image processing, computer vision, face detection and recognition.

Veido atpažinimas naudojant pagrindinių dedamųjų analizę ir bangelių paketų dekompoziciją

Vytautas PERLIBAKAS

Straipsnyje siūloma pagrindinių dedamųjų analizę pagrįsto veido atpažinimo metodo modifikacija panaudojant bangelių paketų dekompoziciją. Pasiūlytas metodas leidžia pagreitinti pagrindinių dedamųjų analizę pagrįsto veido atpažinimo metodo apmokymą esant dideliame apmokymo vaizdų skaičiui ir išlaikyti panašų atpažinimo tikslumą. Atlikus atpažinimo eksperimentus su 423 asmenų veido vaizdais gautas 82–89% atpažinimo tikslumas.