

# A Multi-resolution Decomposition Method for Human Face Recognition

D Q Dai, P C Yuen and G C Feng  
Department of Computing Studies,  
Hong Kong Baptist University.  
Hong Kong  
e-mail:pcyuen@comp.hkbu.edu.hk

*Abstract*— We present a multi-resolution decomposition based method for human face recognition. The recognition process is carried out by comparing the training image representations for images in a training set and the test image representation for image being tested. The size for both image representations are reduced 64 times from the original images, which enables us to reduce significantly the computational costs. The reduced size training image representation is constructed first by applying a three level wavelet transform on the images. A band pass filtered version of the images are trained by Karhunen-Loève expansion. Although band pass component contains less total variance than the low pass component from wavelet theory, the simulations reveal that the band pass component performs better for the recognition task, which is in sharp contrast to the principal component analysis (PCA) based approach, in which low pass component is used and more efforts are exerted on capturing the total variance. Comparing with PCA based methods, our proposed method provides higher accuracy and lower computation load.

*Keywords*— Human Face Recognition, Wavelet Transform, Karhunen-Loève Expansion, Principal Component Analysis

## I. INTRODUCTION

Identification of human face becomes a very active area of research [1], [14], [21], [24] and many approaches have been proposed [6-24] in last twenty years. Generally speaking, these methods can be classified into two approaches. The first one is based on human facial features and called the local feature approach. The performance of this approach depends heavily on the accuracy in determining facial features that are essential to the specific recognition task in advance. The second approach is based on a global setting. A typical method is derived from the Karhunen-Loève(K-L) expansion, or the principal component analysis (PCA) method. When using PCA based method, the principle is that a collection of face images can be approximately reconstructed by storing a small set of weights (principal components) for each image and a small set of standard images (eigenfaces). For good approximation, the goal is to capture as much

as possible of the total variance of the training set. There are some difficulties encountered. The first is the huge computational cost as this method needs to solve an eigen value problem. If the number of images in a training set is about the size of the image, the computational cost involved is beyond the power of most computers. The second problem is that one needs to determine how many principal components needed to be used. This is usually data dependent and quite difficult to determine in advance. The third one concerns with the hypothesis upon which the PCA method based. The implication that the information is contained by eigenvectors with relatively larger eigenvalues needs to be justified, especially for recognition problems.

In view of the above limitations, a new method, based on a three level wavelet transform and the K-L expansion, is proposed in this paper. We use a two dimensional wavelet transform, which could be derived by taking tensor product of two one dimensional wavelet transform, to extract specific frequency components of images. This process serves as, on one hand, the reduction of the dimension of the images. To reduce the computational cost when implementing PCA based methods, this procedure was performed in [20], where a low pass filter was used so that the result image has size of  $32 \times 32$ . The choice of this number is based on the experiments in [5], [13]. In our method, it is further reduced to  $16 \times 16$  so that the associated eigenvalue problem of an auto-correlation matrix involves matrix of  $256 \times 256$  instead of matrix of  $1024 \times 1024$ . Solving eigen problem for matrix has complexity in cubic proportion to its dimension, there is significant computational benefit. On the other hand, simulations reveal that our selection of a particular frequency component provide higher accuracy than using the low pass filtered images, which is in sharp contrast with the PCA based methods, in which more efforts are exerted on capturing as much as possible of the total variance.

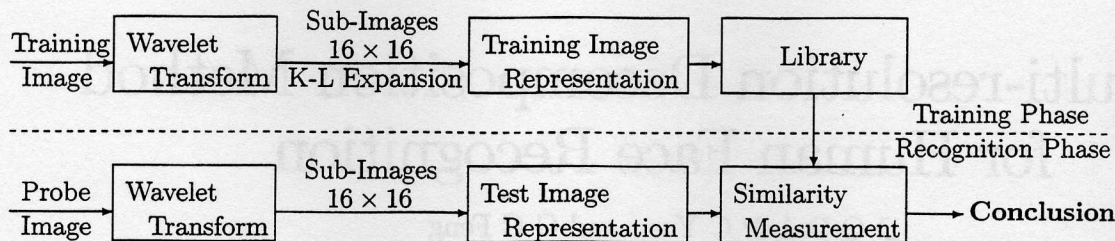


Fig. 1. Block diagram of the proposed system

In the proposed method, a training image representation for each image in the training set and a test image representation for each image being recognized are used. The size is reduced to  $\frac{1}{64}$  of the original image.

## II. PROPOSED METHOD

The block diagram of the proposed system is shown in Figure 1. The proposed system consists of two phases, namely, training phase and recognition phase. In training phase, all images in the domain of interest are represented by a training image representation which are then stored in the reference image library. In recognition phase, when an unknown image is presented to the system, it is then represented by test image representation. A similarity measurement is performed on the test image representation of unknown image and the training image representation of reference images in the library. The similarity score, ranging from zero to one, shows the similarity between the unknown image and reference images. As wavelet transform is a well known technique in pattern recognition, in the following sections, we will concentrate on (1) construction of training image representation, (2) construction of test image representation and (3) similarity measurement.

### A. Training image representation

The procedure of getting the training image representation for an image in a training set is summarized as follows:

- Training set

Choose images to form a training set and subtract each image with mean image to obtain the 'caricature' image.

- 2 - D Wavelet transform

Apply 2 - D three level wavelet transform [3], [9], [23], as shown in figure 2, to each 'caricature' image in the training set.

- Eigenspace

Choose a band, such as band 4 as shown in figure 2, for each transformed image. Notice that the size of band 4 is  $\frac{1}{64}$  of the original. For the chosen

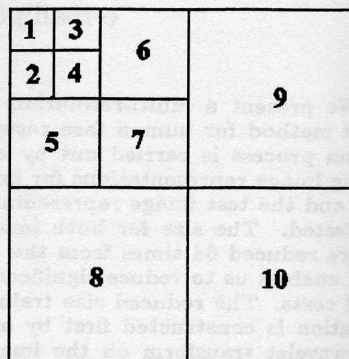


Fig. 2. Three level wavelet decomposition of an image.

band, derive the K-L transform. Only eigenvectors corresponding to nonzero eigenvalues are kept.

- Training image representation

For each wavelet transformed 'caricature' image in the training set, project it onto the eigenspace spanned by eigenvectors corresponding to non zero eigenvalues of the auto-correlation matrix and get the training image representation. Corresponding to the three level wavelet decomposition, the size of its training image representation is  $\frac{1}{64}$  of an image. Figures 3 display five of the images for the two database and their training image representations.

### B. Test image representation

Given an unknown image, by subtracting it with the mean gray level of the training set, we get a test 'caricature' image. Applying the three level two dimensional wavelet transform and extracting a corresponding subband according to the training image representation, this subband will be called the **test image representation** of the test image. Its dimension is  $\frac{1}{64}$  of the original.

### C. Similarity measurement

In all recognition tasks, to measure the difference of two images, a similarity measurement must be introduced. For two images  $A$  and  $B$  with the same dimensions, let  $\vec{A}$  and  $\vec{B}$  be unit vectors formed by row concatenations and normalization. The **similarity measurement**  $\rho(A, B)$  between two images  $A$  and  $B$  is defined by scalar product:  $\rho(A, B) = \vec{A} \cdot \vec{B}$ . It is known that  $-1 \leq \rho(A, B) \leq 1$  and if  $\rho(A, B) = 1$ ,  $A, B$  will be the same.

For convenience of presentation, several notations will be used. Given a training set of  $N$  images  $\{x_1, \dots, x_N\}$ , let  $\{\tilde{x}_1, \dots, \tilde{x}_N\}$  be their training image representations. The self recognition task involves the matrix  $\rho$  of similarity measurement with elements  $\rho(\underline{x}_i, \tilde{x}_j)$ ,  $1 \leq i, j \leq N$ , where  $\{\underline{x}_j, 1 \leq j \leq N\}$  are the test image representation of the original training set. The main diagonal elements of the matrix  $\rho$  are the similarity measurements of the images with oneself in the training set. Its minimum will be denoted by **InMin**.  $\text{InMin} = \min_{1 \leq i \leq N} \rho(\underline{x}_i, \tilde{x}_i)$ . The off diagonal elements of the matrix  $\rho$  measure the similarity of any one image with the rest images in the training set, its maximum will be denoted by **OffMax**.  $\text{OffMax} = \max_{1 \leq i, j \leq N, i \neq j} \rho(\underline{x}_i, \tilde{x}_j)$ . For test images  $\{y_1, \dots, y_M\}$  not in the training set, the maximum of the  $M \times N$  matrix  $\{\rho(\underline{y}_i, \tilde{x}_j), 1 \leq i \leq M, 1 \leq j \leq N\}$  will be denoted by **NotMax**,  $\text{NotMax} = \max_{1 \leq i \leq M, 1 \leq j \leq N} \rho(\underline{y}_i, \tilde{x}_j)$ , where  $\{\underline{y}_i, i = 1, \dots, M\}$  are the test image representations of the test images. In practical recognition, a deformation margin, denoted by **DefMar**

$$\text{DefMar} = \text{InMin} - \text{OffMax}$$

for the difference between **InMin** and **OffMax** is desired. Moreover, upper bounds for **NotMax**, **OffMax** and lower bounds for **InMin** are needed.

### III. RESULTS AND DISCUSSIONS

Two set of images from Brown university[8], [15], [22], [23] and MIT face database[17] are used to evaluate our proposed method and to compare with the PCA based methods. The sizes of the images are respectively  $128 \times 128$  and  $120 \times 128$ . Five of them from each database are shown in Figures 3.

In all tests, the wavelet used is the famous Daubechies wavelet  $D4[3]$ , [9], [23], which is the simplest and most localized of all available wavelets. Its four coefficients are:

$$\begin{aligned} h_0 &= 0.48296291314453 & h_1 &= 0.83651630373781 \\ h_2 &= 0.22414386804201 & h_3 &= -0.12940952255126. \end{aligned}$$

By applying a three level wavelet decomposition, the subbands 1, 2, 3 and 4 in figure 2 are of size  $16 \times 16$  (for the MIT database it is  $15 \times 16$ ). The band 1 is the low pass frequency component of the original image. The bands 2 and 3 measure the changes of the coarsest approximation of the image in vertical and horizontal directions, the band 4 is the highest frequency part among the bands 1, 2, 3, 4. Putting all the bands 1, 2,  $\dots$ , 10 together, they can be ordered as 1, 2 and 3, 4, 5 and 6, 7, 8 and 9, 10 according to their frequency ranges.

In all tests with PCA based methods, the numbers of eigenvectors used are respectively 10, 16, 25 and equal to the number of training images. Instead of using principal components and Euclidean distance as a distance measurement, the similarity measurement described in the previous section is used.

In tables I — III, test images from Brown university are used. The Tables contain data for **InMin**, **OffMax**, **NotMax** and **DefMar** for training set with sizes 50, 143 and 100. A negative value of 'DefMar' means incorrect recognition occurred and a larger positive value of 'DefMar' could lead to an easy decision being made in determining 'accept' and 'reject' thresholds. From table I and table II, if only part of the nonzero eigenvectors is used, negative values of **DefMar** occur. This means at least one person is wrongly recognized as being another. When all of the nonzero eigenvectors (50 and 143 respectively) are used, although the PCA based method could recognize correctly, it is definitely inferior to the proposed method. This can also be easily seen from figure 4. In figure 4, the similarity measurements for self recognition of 143 images are plotted, where the vertical axis indicates similarity measurements. The data in table III records the **NotMax**, which is closely related to **OffMax**. But **NotMax** reports the ability to recognize 'strangers'. In figure 5, the test result with 100 images as training set and the rest 43 as test images is reported. In figures 4(a) and 5(a), the method used all nonzero eigenvalues, from which a higher accuracy of the proposed method can be easily seen.

Since band 4(cf. figure 2) is of size  $15 \times 16 = 240$  for images in the MIT database, this number serves also as the maximum number of the nonzero eigenvectors of related auto-correlation matrix. The performance of the proposed method when the number of images in the training set is reported in table IV, where tests with 256, 300, 330, 360, 400

	PCA with eigenvectors of				Our method
	10	16	25	50	
InMin	0.6606	0.7815	0.8059	0.9861	0.9870
OffMax	0.8872	0.8861	0.8659	0.8034	0.3879
DefMar	-0.2266	-0.1046	-0.0600	0.1827	0.5991

TABLE I

50 IMAGES FROM BROWN UNIVERSITY AS TRAINING SET, SELF RECOGNITION. NEGATIVE DEFMAR MEANS INCORRECT RECOGNITION OCCURRED. LOW DEFMAR INDICATES POOR SEPARATION.

	PCA with eigenvectors of				Our method
	10	16	25	143	
InMin	0.6762	0.7815	0.7869	0.9971	0.9920
OffMax	0.9363	0.9283	0.9237	0.8820	0.4964
DefMar	-0.2601	-0.1468	-0.1368	0.1151	0.4956

TABLE II

143 IMAGES FROM BROWN UNIVERSITY AS TRAINING SET, SELF RECOGNITION. NEGATIVE DEFMAR MEANS INCORRECT RECOGNITION OCCURRED. LOW DEFMAR INDICATES POOR SEPARATION.

	PCA with eigenvectors of				Our method
	10	16	25	100	
NotMax	0.8777	0.8782	0.8776	0.8513	0.4844

TABLE III

100 IMAGES FROM BROWN UNIVERSITY AS TRAINING SET, 43 NOT IN THE TRAINING SET. LARGER NOTMAX MEANS POORER ABILITY TO RECOGNIZE STRANGERS.

and 432 images from MIT as training set are carried out. The task is self recognition. As we are not able to perform tests with PCA based method for image numbers greater than 256, only results for the proposed method is reported. The OffMax is recorded after discarding all the similarity measurements of 27 images of each person and keeping only those with other persons. From Table IV, by InMin, we have very accurate self recognition and by OffMax (less than 0.5), we see that the proposed recognition system has good separation between images in the training set. It can also be seen that the system is stable for large database.

#### IV. CONCLUSION

In this paper, we proposed a new method for human face recognition. In this method, a  $16 \times 16$  wavelet transformed subband with mid-range frequency is used for PCA representation. Experimental results show that the proposed subband method has the following advantages over existing PCA based methods.

\* The proposed method has higher accuracy for

small and large number of training image set.

\* The proposed method only operate on the  $16 \times 16$  resolution image. The computation load in training and recognition stages are greatly reduced.

#### V. ACKNOWLEDGMENT

The authors are grateful to Dr. M. V. Wickerhauser for his kindness to provide us with the images of Brown University.

#### REFERENCES

- [1] R. Chellappa, C. L. Wilson and S. Sirohey, Human and machine recognition of faces: a survey, Proceedings of the IEEE, 83, 5, 705-740, 1995
- [2] D. Beymer and T. Poggio, Face recognition from one example view, Proc. Int'l Conf. Computer Vision, 1995, 500-507.
- [3] I. Daubechies, Ten Lectures on Wavelets, CBMS-NSF series in Applied Mathematics Vol. 61, SIAM Press, Philadelphia, 1992.
- [4] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma and N. Otsu, Face recognition system using local autocorrelations and multiscale integration, IEEE Trans. PAMI, Vol. 18, 10, 1024-1028, 1996
- [5] L. Harmon, The recognition of faces, Scientific American, Vol. 229, 1973, 71-82.
- [6] X. Jia and M. S. Nixon, Extending the feature vector

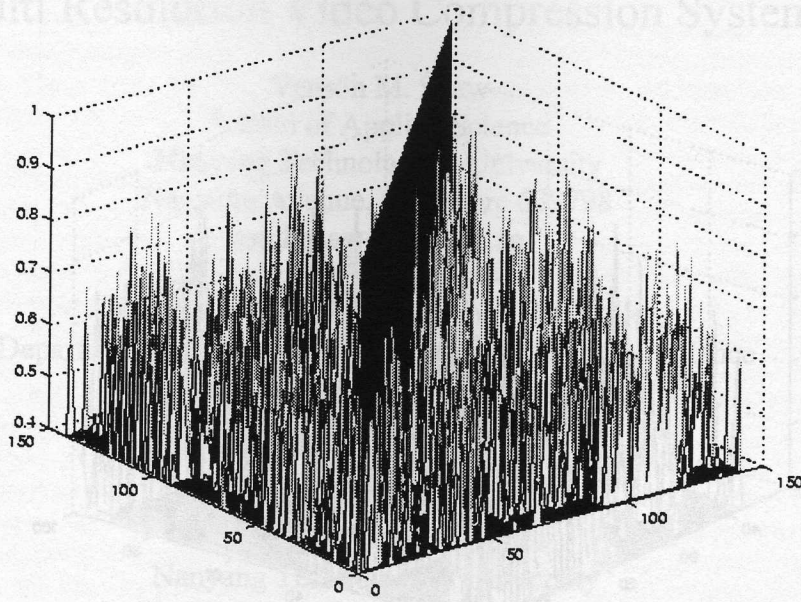
No. of Images	256	300	330	360	400	432
InMin	0.9960	1.0000	1.0000	1.0000	1.0000	1.0000
OffMax	0.3627	0.3689	0.4334	0.4372	0.4444	0.4384
DefMar	0.6333	0.6311	0.5666	0.5628	0.5556	0.5617

TABLE IV

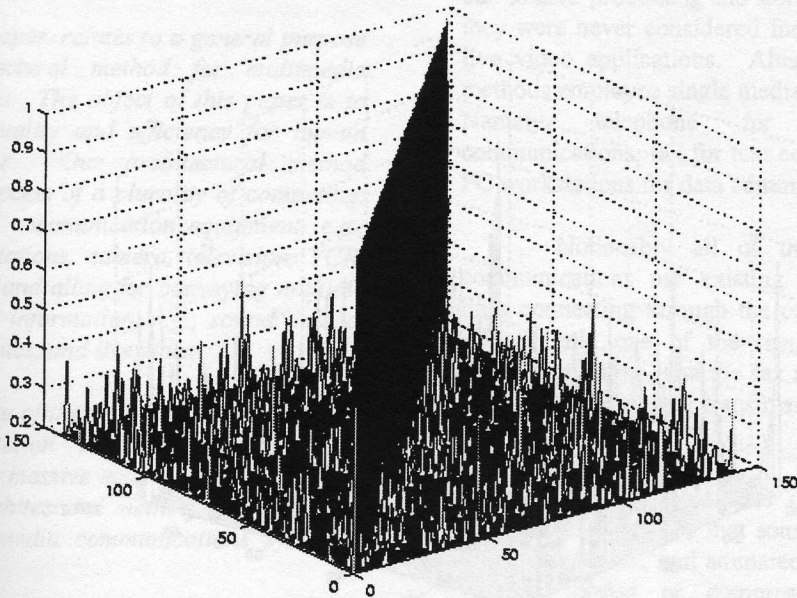
VARIOUS NUMBERS OF IMAGES FROM MIT AS TRAINING SET, THE SELF RECOGNITION. FOR QUITE LARGE DATABASE, THE DEFMAR(LARGER THAN 0.5) IS STABLE.

- for automatic face recognition, IEEE Trans. on PAMI, 17, 12, 1167-1176, 1995
- [7] A. K. Jain and A. Vailaya, Image retrieval using color and shape, Proc. ACCV, Singapore, 1995, Vol. 2, 529-533.
- [8] M. Kirby and L. Sirovich, Application of the Karhunen-Loève procedure for the characterization of human faces, IEEE Trans. PAMI., Vol. 12, 1990, 103-108.
- [9] Y. Meyer, Wavelets: Algorithms and Applications, SIAM Press, Philadelphia, 1993.
- [10] A. Pentland, B. Moghaddam and T. Starner, View-based and modular eigenspaces for face recognition, Proc. IEEE CS Conf. Computer vision and Pattern Recognition, Seattle, June, 1994, 84-91.
- [11] A. J. O'Toole, H. Abdi, K. A. Deffenbacher and D. Valentin, A low-dimensional representation of faces in the higher dimensions of the space, J. Opt. Soc. Am., A, Vol. 10, 1993, 405-411.
- [12] H. A. Rowley, S. Baluja and T. Kanade, Neural network-based face detection, CVPR'96, 203-208, 1996
- [13] A. Samal, Minimum resolution for human face detection and identification, Proceedings of SPIE human vision, visual processing, and digital display, Vol., 1453, 1991, 81-89.
- [14] A. Samal and P. A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: A survey, Pattern Recognition, Vol. 25, 1992, 65-77.
- [15] L. Sirovich and M. Kirby, Low-dimensional procedure for the characterization of human faces, J. Opt. Soc. Am. A, Vol. 4, No. 3, March 1987, 519-524.
- [16] L. Sirovich and C. Sirovich, Low dimensional description of complicated phenomena, Contemp. Mathematics., Vol. 99, 1989, 277-305.
- [17] M. Turk and A. Pentland, Eigenfaces for recognition, J. Cognitive Neuroscience, Vol. 3, March 1991, 71-86.
- [18] E.M.-Tzanakou, E. Uyeda, R. Ray, A Sharma, R. Ramujan and J. Dong, Comparison of neural network algorithm for face recognition, Simulation, 64, 1, 15-27, 1995
- [19] D. L. Swets and J. J. Weng, Using discriminant eigenfeatures for image retrieval, IEEE Trans. PAMI., Vol. 18, No. 8, 1996, 831-836.
- [20] D. Valentin, H. Abdi and A. J. O'Toole, Principal component and neural network analyses of face images: Explorations into the nature of information available for classifying faces by sex, In C. Dowling, F. S. Roberts, P. Theuns, Progress in mathematical psychology, Hillsdale: Erlbaum, (in press, 1996)
- [21] D. Valentin, H. Abdi, A. J. O'Toole and G. W. Cottrell, Connectionist models of face processing: A Survey, Pattern Recognition, Vol. 27, 1994, 1209-1230.
- [22] M. V. Wickerhauser, Large-rank approximate component analysis with wavelets for signal feature discrimination and the inversion of complicated maps, J. Chemical Information and Computer Sciences, Vol. 34, No. 5, 1994, 1036-1046.
- [23] M. V. Wickerhauser, Adapted Wavelet Analysis from Theory to Software, AK Peters, Ltd., Wellesley, Massachusetts, 1994.
- [24] Face recognition: From theory to applications, Nato Advanced Study Institute(ASI) Programme, June 23-July 4, 1997, Stirling, Scotland, UK.



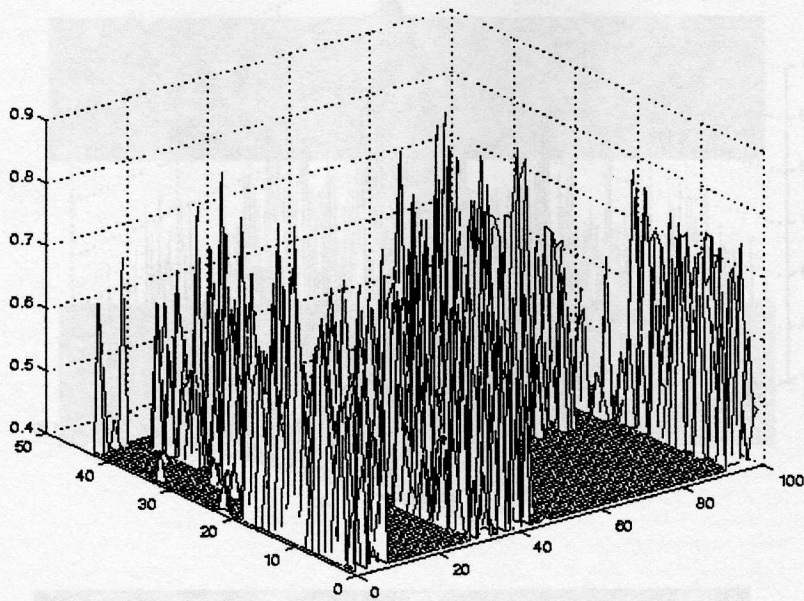


(a)

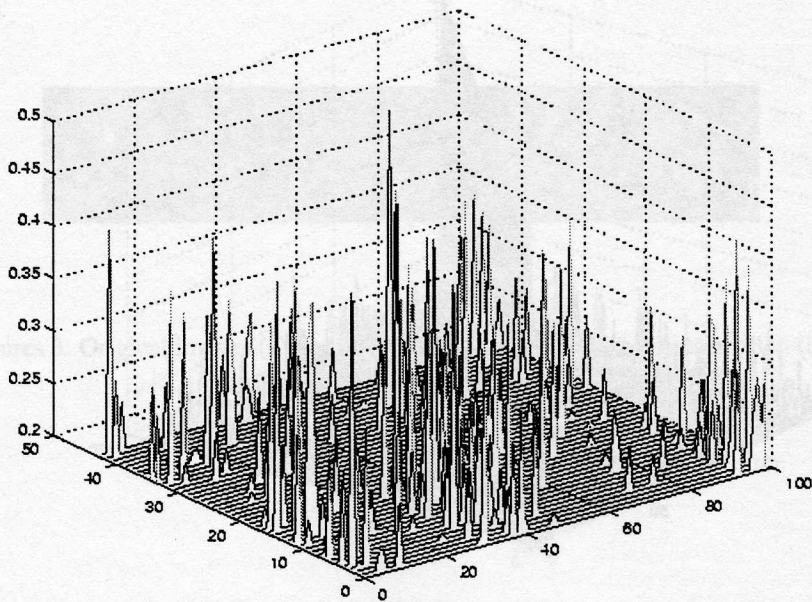


(b)

Figures 4: System performance using training images as testing images.  
(a) Original PCA method and (b) proposed method.



(a)



(b)

Figures 5: System performance using images outside training set.  
(a) original PCA method and (b) proposed method.