LOCAL WAVELET ANALYSIS FOR FACE RECOGNITION

Hazım Kemal Ekenel, Rainer Stiefelhagen

Interactive Systems Labs, Computer Science Department, Universität Karlsruhe (TH)
Am Fasanengarten 5, 76131, Karlsruhe, Germany
email: {ekenel, stiefel}@ira.uka.de

ABSTRACT

In this paper we investigate the benefits of using local wavelet analysis to the face recognition problem. We examine two possible approaches to perform local wavelet analysis. In the first approach, discrete wavelet transform is performed on the entire face image and then the transformed image is partitioned into non-overlapping rectangular blocks. In the second approach, as in the JPEG2000 standard, the input face image is first partitioned into non-overlapping rectangular blocks, and then on each block discrete wavelet transformation is performed. Proposed approaches are tested against the occlusion problem using the AR face database and significant improvements are observed in the face recognition performance.

1. INTRODUCTION

Face recognition has been attracting intense research efforts due to its importance both as one of the main building blocks of natural human computer interfaces and as a biometric trait. Moreover, it is expected that the solution of the face recognition problem would provide valuable insights about the other pattern recognition problems.

Holistic approaches have been dominating the face recognition research since the beginning of 1990s [1-4]. However, recently, there has been growing interest on local appearance-based face recognition [5-10]. This growing interest is not surprising, since performing local analysis has many advantages over holistic approaches. For instance, one can select or weight local regions according to their importance in the face classification task. Furthermore, an appearance change in a local region can effect the entire feature vector in a holistic representation scheme, on the other hand it only modifies the feature vector that is extracted from the corresponding local region and the feature vectors that are extracted from the other regions will remain unchanged in a local analysis scheme, causing only a local modification in the overall feature vector.

The local appearance-based approaches can be divided into two groups: The ones that require the use of specific regions [5,6,8] and the ones that simply partition the input face image into blocks without considering any specific regions [7,9,10]. Although using salient regions may seem more reasonable at the first sight, it is a difficult task to detect these regions. Moreover, erroneous detection of these local regions may lead to severe performance drops. Because of these reasons, this study follows the idea of partitioning the input face image into non-overlapping blocks without considering any salient regions.

Discrete wavelet transform has been used in various studies on face recognition [11–16]. In [11], principal component analysis (PCA) is performed on the wavelet subbands that are extracted by applying three-level discrete wavelet transformation on the input face image. In [12], a wavelet transform-based speaker identification system in a teleconferencing environment is proposed. In the proposed system three-level wavelet decomposition is performed on the input face image. The low frequency components at each level as well as the original image are used for classification. A face recognition system based on wavelet packet analysis is proposed in [13]. In [14], PCA is performed on the feature vector that is constructed by concatenating the low frequency subbands obtained at each level of three-level wavelet decomposition. Discriminant waveletfaces approach is proposed in [15]. In this study, the low frequeny subband extracted from the third level wavelet decomposition, called the waveletface, is used as the input of the LDA. In [16], the wavelet subbands that are less sensitive to the expression and illumination variations are searched. Moreover, in this study, the subbands that attain individually high correct recognition rates are fused to improve the performance further. All these proposed discrete wavelet transform-based face recognition algorithms are holistic approaches, none of them utilizes any local information. Different from these previous studies, we put into evidence the contribution of local wavelet analysis to combat, specifically, occlusion problem. We describe two ways to perform local wavelet analysis-based face recognition. We propose an accompanying block selection scheme and we investigate the effect of this block selection scheme on the face recognition performance.

The organization of the paper is as follows. In Section 2, discrete wavelet transform is explained briefly. The idea of face recognition using local wavelet analysis is described in Section 3. Experimental results are presented and discussed in Section 4. Finally, in Section 5 conclusions are given.

2. DISCRETE WAVELET TRANSFORM

Discrete wavelet transform (DWT) is a well-known signal analysis tool, widely used in feature extraction, compression and denoising applications.

The two-dimensional wavelet transform is performed by consecutively applying one-dimensional wavelet transform to the rows and columns of the two-dimensional data. The original image of resolution NxN is first filtered along the rows and downsampled by 2 yielding two NxN/2 images that have high and low frequency contents, respectively. After this decomposition, the wavelet transform is applied to the columns of these NxN/2 resolution images. In the final stage of the decomposition we have four N/2xN/2 resolution subband images: the scaling component containing low-pass global information obtained by low-pass filtering the rows and columns; the horizontal details obtained by low-pass filtering the rows and high-pass filtering the columns; the vertical details obtained by high-pass filtering the rows and low-pass filtering the columns; the diagonal details obtained by high-pass filtering the rows and columns. The decomposition can be continued further by performing same processing steps on the scaling component. In Fig. 1, one level wavelet decomposition of a face image is shown.

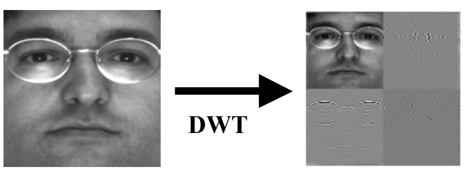

Figure 1. Sample one-level wavelet decomposed image

3. FACE RECOGNITION USING LOCAL WAVELET ANALYSIS

Local wavelet analysis-based face recognition can be performed in two different ways. In the first architecture discrete wavelet transform is performed on the entire face image and then the transformed image is partitioned into non-overlapping rectangular blocks. The important regions for classification are selected with a block selection scheme. The features extracted from each selected block are concatenated to construct the overall feature vector that will be used in classification. In the second architecture, as in the JPEG2000 standard [17], the input face image is first partitioned into non-overlapping rectangular blocks. Then, the block selection is performed to determine on which blocks discrete wavelet transform will be performed. Afterwards, as in the first architecture, the features extracted from each selected block are concatenated to construct the overall feature vector.

3.1 Block Selection

Block selection is done by measuring the similarity between the pixel values of each block in the average training image and the pixel values of the corresponding blocks in the test image. This facilitates adaptive block selection. In other words, for each new test face image, different blocks can be selected.

The average pixel values of the i[th] block of the training images can be calculated as

$$\boldsymbol{b}_{tra,i} = \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{b}_{tra,i,l}\ , \tag{1}$$

where L is the number of images in the training set and $\boldsymbol{b}_{tra,i}$ represents the vector that contains the average pixels values of the i[th] block. This K dimensional vector's mean value can be calculated as

$$m_{tra,i} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{b}_{tra,i,k}\ . \tag{2}$$

By subtracting the vector's mean value from its pixel values, the $\bar{\boldsymbol{b}}_{tra,i}$ is obtained

$$\bar{\boldsymbol{b}}_{tra,i} = \boldsymbol{b}_{tra,i} - m_{tra,i}\ . \tag{3}$$

Similarly, the mean value of the i[th] block in the test image can be calculated as

$$m_{test,i} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{b}_{test,i,k}\ . \tag{4}$$

Again, by subtracting the vector's mean value from its pixel values, the $\bar{\boldsymbol{b}}_{test,i}$ is obtained

$$\bar{\boldsymbol{b}}_{test,i} = \boldsymbol{b}_{test,i} - m_{test,i}\ . \tag{5}$$

Finally, the similarity between the i[th] block in the average training image and the corresponding block in the test image is calculated as

$$s_i = \bar{\boldsymbol{b}}_{tra,i} * \bar{\boldsymbol{b}}_{test,i} / \left\|\bar{\boldsymbol{b}}_{tra,i}\right\| * \left\|\bar{\boldsymbol{b}}_{test,i}\right\|\ .$$

4. EXPERIMENTS

Two separate experiments are conducted to compare the proposed local wavelet analysis-based approach with the baseline holistic approach –eigenfaces. In the first experiment, the face recognition system is trained with face images that have no occlusion and tested against occluded faces from the same recording session. In the second test, again the system is trained with the face images that have no occlusion, but this time tested against occluded faces from a different recording session. The AR face database [18] is used in both of the experiments.

Daubechies 4 wavelet is used to process the face images. The Daubechies wavelets are one of the most widely used wavelet families [19]. They are orthonormal and have compact support. Moreover, in [11] it is shown that Daubechies 4 wavelet performs better in terms of computation time and recognition performance compared to the other well-known wavelets. Two-level discrete wavelet decomposition is performed on the input images. The second level scaling component is used for face recognition.

Nearest neighborhood classifier is used in the study. *L1* norm is used as the distance metric.

4.1 Intra-session Experiments

The face database used in the intra-session experiments consists of 660 face images of 110 individuals that are taken from the first session of the AR face database [18]. Each individual in the derived face database has six images. These images are annotated as “1: neutral expression”, “5: left light on”, “6: right light on”, “11: wearing scarf”, “12: wearing scarf and left light on”, “13: wearing scarf and right light on”. From these six images, the ones with annotations “1, 5, 6” are used for training and the remaining ones with annotations “11, 12, 13” are used for testing. The face images are aligned using the eye center locations and scaled to 128x128 pixels resolution. Sample images can be seen in Fig. 2.

Figure 2. Sample face images from the AR face database.

For the first architecture, two-level discrete wavelet decomposition is performed on the entire face image. The resulting second level scaling component is used for further processing. This 32x32 pixels resolution scaling component is divided into 4x4 pixels blocks, providing 64 non-overlapping blocks. From these blocks the important ones are selected using the scheme described in Section 3.1 and the feature vector is constructed by concatenating them. For the second architecture, the input face image is divided into 16x16 pixels blocks, resulting in 64 non-overlapping blocks. The important blocks are selected from these blocks, again using the scheme presented in Section 3.1. On each selected block two-level discrete wavelet transform is performed. The 4x4 pixels resolution second level scaling components are used as features and they are concatenated to construct the overall feature vector.

The correct classification performances of the presented architectures are shown in Fig. 3. For local wavelet analysis (LWA) architectures, this figure gives their correct recognition rate values for each used number of blocks. For eigenfaces approach, there is a single correct recognition rate. It is the performance value obtained by using 120 dimensional feature vectors. These feature vectors are extracted by projecting the input face images onto the face space that is constructed by the first 120 eigenvectors. It can be observed that both of the proposed local wavelet analysis architectures outperform the holistic baseline –eigenfaces- significantly. For instance the performance values when we only use 10 blocks are: 13.9% for eigenfaces, 46.7% for LWA architecture 1 and 64.9% for LWA architecture 2. The performance value of the eigenfaces may seem very low at first sight, however, these results are similar to the ones obtained in [20], thus verifying the difficulty of face recognition against the occlusion problem.

The second LWA architecture performed better then the first one. This can be due to the effects of neighboring blocks in wavelet filtering in the first architecture. This effect does not exist in the second architecture, because the image is divided into blocks prior to the wavelet analysis and all the blocks are wavelet filtered separately, without using any values from the neigboring blocks at the border regions.

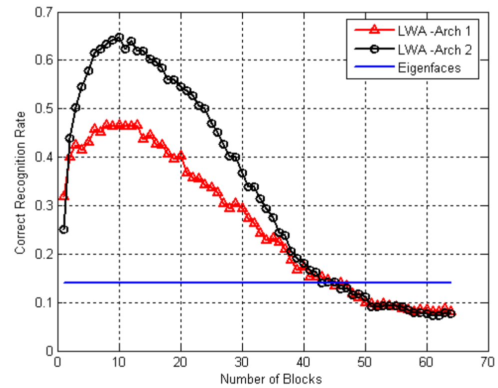

Figure 3. Correct recognition rate versus number of used blocks on the intra-session experiments

The performance increases in both of the LWA architectures at the beginning as the used number of blocks increases. It saturates when the number of used blocks is around 10. As the number of used blocks increases further, that is, as the LWA-based approaches become more holistic, the correct recognition rate decreases severely.

In Fig. 4, average importance order of the blocks is shown. The smaller numbers imply more importance. This order is obtained by sorting the average similarity measures calculated for each test face image. It can be seen that, as one can expect, the blocks located in the upper half of the face image have more importance.

7	16	20	28	34	33	23	1
36	31	15	5	4	13	25	35
10	18	22	19	12	17	8	3
26	29	30	2	9	21	6	24
39	47	37	14	11	40	43	38
51	32	42	62	61	41	27	52
48	64	56	60	57	53	63	45
55	44	49	58	50	59	46	54

Figure 4. Average importance order of the blocks obtained on the intra-session block selection

4.2 Inter-session Experiments

In the inter-session experiments, the training data is the same as the one used in intra-session experiments. The test images are selected from a different session and have the same annotations with the ones in intra-session experiment. The face images are aligned using the eye center locations and scaled to 128x128 pixels resolution.

Fig. 5 plots the correct recognition rates for varying number of used blocks. The observed outcomes are similar to the ones obtained in intra-session experiments. The main difference, one can notice is the worse performance values. This is expected, since in this experiment the occlusion problem is coupled with the time gap between training and testing data, hence causing a more difficult problem.

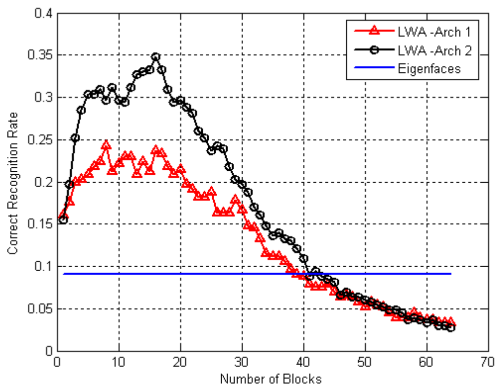

Figure 5. Correct recognition rate versus number of used blocks on the inter-session experiments

The average importance order obtained in this experiment is depicted in Fig. 6. Although there is a change in the order, the blocks located in the upper half of the face image still have more importance. This verifies our block selection scheme.

7	20	19	27	29	31	23	6
36	33	18	2	5	10	26	35
13	14	16	24	11	17	1	3
25	28	30	4	8	21	9	22
46	40	39	15	12	44	37	47
52	34	43	62	61	42	32	49
41	64	54	59	60	56	63	50
55	38	45	58	51	57	48	53

Figure 6. Average importance order of the blocks obtained on the inter-session block selection

5 CONCLUSION

In this paper we presented novel local wavelet analysis-based face recognition approaches. We described two ways to perform LWA-based face recognition and propose an accompanying block selection scheme. We tested the proposed approaches against the occlusion problem. We performed intra-session and inter-session experiments. Significant performance improvements are obtained in both of the experiments using the proposed LWA-based approaches. The second architecture reached higher correct recognition rates. In addition, we also examine the average importance order, and as one can expect, we found out that the blocks located in the upper half of the face image have more importance.

ACKNOWLEDGEMENTS

This work is sponsored by the European Union under the integrated project CHIL, contract number 506909.

REFERENCES

[1] M. Turk, A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Science*, pp. 71–86, 1991.
[2] R. Brunelli, T. Poggio, "Face Recognition: Features versus Templates", *IEEE Trans. on PAMI*, 13(10): 1042-1052, 1993.
[3] J.H.P.N. Belhumeur, D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection", *IEEE Trans. on PAMI*, 19(7):711–720, 1997.
[4] M. Bartlett et al, "Face recognition by independent component analysis", *IEEE Trans. on Neural Networks*, 13(6):1450–1454, 2002.
[5] A. Pentland et al., "View-based and modular eigenspaces for face recognition", *Proceedings of IEEE CVPR*, pp. 84-91, 1994.
[6] B. Heisele et al, "Face recognition: component-based versus global approaches", *Computer Vision and Image Understanding*, 91:6-21, 2003.
[7] R. Gottumukkal, V.K. Asari, "An improved face recognition technique based on modular PCA approach", *Pattern Recognition Letters,* 25(4), 2004.
[8] T. Kim et al, "Component-based LDA Face Description for Image Retrieval and MPEG-7 Standardisation", *Image and Vision Computing*, 23(7):631-642, 2005.
[9] H.K. Ekenel, R. Stiefelhagen, "A Generic Face Representation Approach for Local Appearance-based Face Verification", *IEEE Workshop on FRGC Experiments*, 2005.
[10] H.K. Ekenel, R. Stiefelhagen, "Local Appearance-based Face Recognition Using Discrete Cosine Transform", *EUSIPCO 2005*, Antalya, Turkey, 2005.
[11] G.C. Feng, P.C. Yuen, D.Q. Dai, "Human face recognition using PCA on wavelet subband", *SPIE Journal of Electronic Imaging*, 9(2): 226–233, 2000.
[12] J. Tang et al, "A wavelet-transform based asker identification system for smart multi-point teleconferences", *Journal of the Visualization Society of Japan*, 20 (1):303–306, 2000.
[13] C. Garcia et al, "Wavelet packet analysis for face recognition", *Image and Vision Computing,* 18 (4):289–297, 2000.
[14] B. Li, Y. Liu, "When eigenfaces are combined with wavelets", *International Journal of Knowledge-Based Systems,* 15(5/6):343–347, 2002.
[15] J.T. Chien, C.C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition", *IEEE Trans. on PAMI*, 24(12):1644–1649, 2002.
[16] H.K. Ekenel, B. Sankur, "Multiresolution Face Recognition", *Image and Vision Computing*, 23(5):469-477, 2005.
[17] C. Christopoulos et al, "The JPEG2000 Still Image Coding System: An Overview", *IEEE Trans. on Consumer Electronics*, 46(4):1103-1127, 2000.
[18] A.M. Martinez, R. Benavente, "The AR Face Database*",* *CVC Technical Report* #24, 1998.
[19] I. Daubechies, "Ten lectures on wavelets", *CBMS-NSF Regional Conference Series in Applied Mathematics*, Vol. 61, SIAM Press, Philadelphia, 1992.
[20] R. Gross et al, "Que vadis face recognition?*", Third Workshop on Emprical Evaluation Methods in Computer Vision,* 2001.