Face Recognition Using Wavelet Transform and Non-negative Matrix Factorization

Neo Han Foon, Andrew Teoh Beng Jin, and David Ngo Chek Ling

Faculty of Information Science and Technology (FIST), Multimedia University,
Jalan Ayer Keroh Lama, Bukit Beruang,
75450 Melaka, Malaysia
{hfneo, bjteoh, david.ngo}@mmu.edu.my

Abstract. This paper demonstrates a novel subspace projection technique via Non-Negative Matrix Factorization (NMF) to represent human facial image in low frequency subband, which is able to realize through the wavelet transform. Wavelet transform (WT), is used to reduce the noise and produce a representation in the low frequency domain, and hence making the facial images insensitive to facial expression and small occlusion. After wavelet decomposition, NMF is performed to produce region or part-based representations of the images. Non-negativity is a useful constraint to generate expressiveness in the reconstruction of faces. The simulation results on Essex and ORL database show that the hybrid of NMF and the best wavelet filter will yield better verification rate and shorter training time. The optimum results of 98.5% and 95.5% are obtained from Essex and ORL Database, respectively. These results are compared with our baseline method, Principal Component Analysis (PCA).

1 Introduction

Faces play an important role in the primary focus of attention in social intercourse, and to identify and verify an identity. Human has the ability to recognize thousand of faces learned throughout their lifetime. Therefore, face recognition is one of the most remarkable abilities of human vision. With its extensive and robust application in security systems, identification of criminals, assistance with speech recognition systems, surveillances and user identifications, it has become a significant research area to the physicists and scientists worldwide.

Unlike human beings who have the excellent capability to recognize different faces, machines are still lacking of this aptitude due to its variation in illuminations, complex backgrounds, visual angles, facial expressions and therefore, face recognition has become a complex and challenge task.

In the early study, [1, 2] found that information in low spatial frequency bands have a dominant role in face recognition. Thereafter, [3] shows that low-frequency components contribute to the global description, while the high-frequency components contribute to the finer details required in the identification task. Nastar et al. [4] have

G.I. Webb and Xinghuo Yu (Eds.): AI 2004, LNAI 3339, pp. 192–202, 2004.

observed the relationship between variation in facial appearance and their deformation spectrum. They found that facial expressions and small occlusions affect the high-frequency spectrum whereas changes in pose or scale of a face affect their low frequency spectrum. Only a change in face will affect all frequency components. Bow [5] demonstrates that the quality of the reconstructed image is very good if the image is restored with the lower half-frequency spectrum. The aforesaid show that the low-frequency components are already sufficient for recognition. Therefore, wavelet transform is proposed for facial images decomposition.

The dimensionality of real world objects is inevitably much higher than can be pictured using the three dimensions we can see. Hence, dimensionality reduction is very important to project the facial images from the high-dimensional space onto a lower-dimensional space. [6] have proposed Principal Component Analysis (PCA) to describe face patterns with a lower-dimensional space than the image space. One of the characteristic of PCA is such that a high dimensional vector can be represented by a small number of orthogonal basis vectors, namely principal components. In this paper, PCA is used as a baseline for comparison.

Next stage will be followed by projected the facial images onto the subspace via NMF. NMF was proposed by Lee and Seung in 1997 [7]. NMF performed similar to PCA but it is based on finding a representation of a local space only using additive constraints. NMF is used in various applications including information retrieval [8], summarizing video [9], music transcription [10] and language model adaptation [11]. NMF imposes the non-negativity constraints in learning basis images. The pixel values of resulting images, as well as coefficients for reconstructions, are all non-negative. This technique preserves much of the structure of the original data and guarantees that both the resulting low-dimensional basis and its accompanying weights are non-negative. For this reason, NMF is considered as a procedure for learning a part-based representation.

In this paper, we investigate the performance when WT and NMF are integrated to take the advantages of these two methods in order to achieve an excellent verification rate in identifying the faces. These results are compared with PCA technique which is our baseline. With the adoption of WT, the training time in NMF can also be reduced significantly.

This paper is organized as follows. Section 2 define wavelet transform based features and their basic properties. In section 3, we describe PCA and NMF properties. Integrated framework of WT and NMF is illustrated in section 4. The experimental results are presented in section 5 and lastly, conclusion is discussed in section 6.

2 Wavelet Transform

Wavelet analysis is a common tool for analyzing localized variation of power within a time series. By decomposing a time series into time-frequency space, dominant modes of variability and how these modes vary in time are able to be determined. The wavelet transform are used for numerous studies in geophysics, including analysis of wave aberration, tropical convection and the dispersion of ocean waves.

The wavelet transform can be used to analyze time series that contain non-stationary power at many different frequencies. Assume that one has a time series, x_n, with equal time spacing and n=0…N-1. Also assume that one has a wavelet function, that depends on a non-dimensional "time" parameter. To be "admissible" as a wavelet, this function must have zero mean and be localized in both time and frequency space [12].

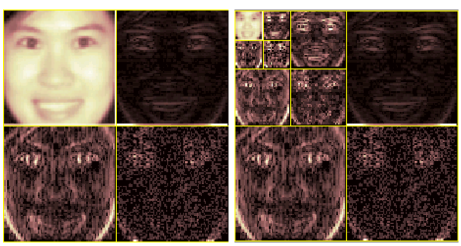
The wavelet function is used generically to refer to either orthogonal or non-orthogonal wavelets. In this paper, only the orthogonal wavelet function is used, namely Haar, Daubechies 5 and 10, Symlet 5 and 10, and Spline bior1.1 and 5.5. The basic idea is to represent any arbitrary function *f(x)* as a superposition of a set of such wavelets or basis functions. The scaling and shifting variables are discretized so that wavelet coefficients can be described by two integers, m and n. Thus, the discrete wavelet transform is given in Equation 1,

$$(W_\psi f(x))(m,n) = \frac{1}{\sqrt{a_o{}^m}} \sum_k x[k]\psi[a_o{}^{-m}n - k] \tag{1}$$

where $x[k]$ is a digital signal with sample index k, and $\psi(n)$ is the mother wavelet.

Two-dimensional wavelet transform leads to a decomposition of approximation coefficients at level j-1 in four components, the approximations at level j, L_j and the details in three orientations (horizontal, vertical and diagonal), $D_{jvertical}$, $D_{jhorizontal}$ and $D_{jdiagonal}$.

Discrete wavelet transform is used to decompose the facial images into a multiresolution representation in order to keep the least coefficients possible without losing useful image information. Fig. 1(a) depicts the decomposition process by applying a two-dimensional haar wavelet transform of a face image in level 1 and Fig. 1(b) depicts three levels wavelet decomposition by applying haar wavelet transform on the low-frequency band sequentially.



a) 1-level wavelet decomposition   b) 3-level wavelet decomposition

Fig. 1. Face image in wavelet subbands (a) 1-level wavelet decomposition (b) 3-level wavelet decomposition

3 PCA and NMF

3.1 Principle Component Analysis (PCA)

PCA is a linear subspace projection technique used to project data from high-dimensionality subspace onto a lower-dimensionality subspace. Let X_j be N-element one-dimensional image and suppose that we have n such images (j=1,…,n). A one-dimensional image column X from the two-dimensional image is formed by scanning all the elements of the two-dimensional image row by row and writing them to the column vector. Then the mean vector, centered data vector and covariance matrix are calculated as in Equation 2,3 and 4.

$$m = \frac{1}{n}\sum_{j=1}^{n} X_j \ , \tag{2}$$

$$d_j = X_j - m \, , \tag{3}$$

$$C = \frac{1}{n}\sum_{j=1}^{n} d_j d_j^{\,T} \tag{4}$$

Here $X = (x_1,\ldots,x_N)^{\mathrm{T}}$, $m=(m_1,\ldots m_N)^{\mathrm{T}}$, $d=(d_1,\ldots d_N)^{\mathrm{T}}$. When calculating the covariance matrix, eigenvectors are sorted by decreasing eigenvalues only taking the most representative ones which correspond to the directions of maximum variance. Once the subspace is fully described by a projection matrix, the classification of a new feature vector is accomplished by projecting and finding the nearest training one using Euclidean distance metric.

3.2 Non-negative Matrix Factorization (NMF)

NMF is a method to obtain a representation of data using non-negativity constraints. These constraints lead to a part-based representation in the image subface because they allow only additive, not subtractive, combinations of original data. This is believed to be compatible with the intuition notion of combining parts to form a whole in an accumulative means, and this is how NMF learns a part-based representation [6]. It is also consistent with the physiological fact that the firing rate is non-negative.

Given an initial database expressed by a n x m matrix X, where each column is an n-dimensional non-negative vector of the original database (m vectors), it is possible to find two new matrices (W and H) in order to approximate the original matrix :

$$X \approx \tilde{X} \equiv WH, \text{where } W \in \Re^{nxr}, H \in \Re^{rxm} \tag{5}$$

We can rewrite the factorization in terms of the columns of X and H as:

$$x_j \approx \tilde{x}_j = Wh_j, \text{where } x_j \in \Re^{m}, h_j \in \Re^{r} \text{for } j = 1,\ldots,n \tag{6}$$

The dimensions of the factorized matrices W and H are n x r and r x m, respectively. Assuming consistent precision, a reduction of storage is obtained whenever r, the number of basis vectors, satisfies $(n+m)r < nm$. Each column of matrix W contains basis vectors while each column of H contains the weights needed to approximate the corresponding column in V using the basis from W.

In order to estimate the factorization matrices, an objective function has to be defined. We have used the square of Euclidean distance between each column of X and its approximation of X=WH subject to this objective function:

$$\Theta_{NMF}(W,H)=\sum_{j=1}^{n}\left\|x_j-Wh_j\right\|^2=\left\|X-WH\right\|^2 \tag{7}$$

This objective function can be related to the likelihood of generating the images in X from the basis W and encoding H. An iterative approach to reach a local minimum of this objective function is given by the following rules [13] :

$$W_{ia} \leftarrow W_{ia}\sum_{\mu}\frac{V_{i\mu}}{(WH)_{i\mu}}H_{a\mu}\ , \tag{8}$$

$$W_{ia} \leftarrow \frac{W_{ia}}{\sum_{j} W_{ja}}\ , \tag{9}$$

$$H_{a\mu} \leftarrow H_{a\mu}\sum_{i} W_{ia}\frac{V_{i\mu}}{(WH)_{i\mu}} \tag{10}$$

Initialization is performed using positive random initial conditions for matrices W and H. Convergence of the process is also ensured. Fig. 2 demonstrates the NMF basis figures extracted from our database. These basis provide a sparse and part-based representation of face images.

In face recognition, NMF is performed where $W = (W^{T} W)^{-1} W^{T}$. In the feature extraction, each training facial images x_i is projected into the linear space as a feature vector $h_i = Wx_i$. This is then used as a training feature point. A testing face image x_t to be classified is represented as $h_t = Wx_t$. Next we classified them using nearest neighborhood classification scheme, Euclidean distance metric. The Euclidean distance between the testing image and each training image, $d(h_t, h_i)$ is calculated. The testing image is classified to the class to which the closest training image belongs.

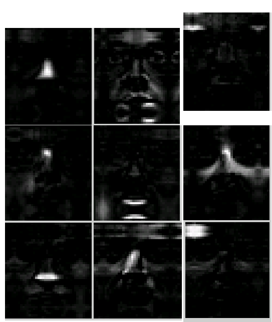


Fig. 2. NMF bases

4 Integrated Framework of Wavelet and Non-negative Matrix Factorization

The integrated framework of wavelet and non-negative matrix factorization (wNMF) produce sparse and part-based images. In addition to that, WT reduce the resolution of the image and decrease the computation load of the feature generation. In this paper, two level of wavelet decomposition is performed on face images. The face image with the low-frequency subband representation, L_l is then subjected to NMF transform as described in Section 3. The block diagram of wNMF feature representation is illustrated in Fig. 3.

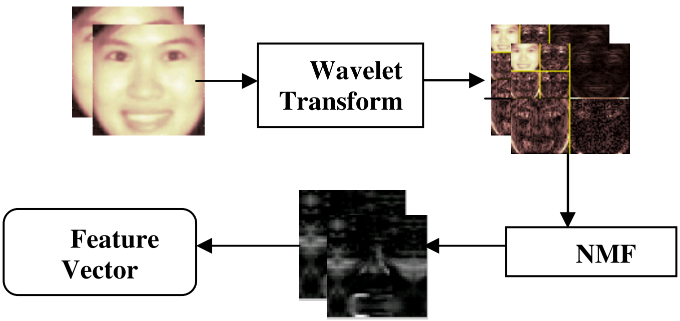


Fig. 3. Block diagram of generating the wNMF features

5 Experimental Results

The experiments were conducted by using *Faces-93* Essex University Face Database (Essex) [14] and Olivetti Research Laboratory (ORL) Database [15]. There are various aspects in the *Faces-93* Essex database which made it appropriate to this experiment. Data capture conditions are subject to photograph at fixed distance from camera, and individuals are asked to speak to produce images of the same individuals with different facial expressions. This database consists of 100 subjects with 10 images for each subject. The set of the 10 images for each subject is randomly partitioned into a training subset of 3 images and a test set of another 5 images. The image size is of 61 x 73 pixels, 256 – level grayscale. The face scale in the images is uniform and there are minor variations in turn, tilt and slant. On the other hand, ORL database contains 40 subjects with 10 images for each subject. The set of the 10 images for each subject is also randomly partitioned into a training subset of 3 images and a test set of another 5 images. The image size is of 92 x 112 pixels, 256 – level grayscale. There are major variations in turn, tilt and slant which we make use of the complexity of this database for our experiments.

For performance evaluation, the error measures of a verification system are False Accept Rate (FAR) and False Reject Rate (FRR) as defined in Equation 11 and 12.

$$\text{FAR}=\frac{\text{Number of accepted imposter claims (FA)}}{\text{Total number of imposter accesses}}\text{x}100\% \tag{11}$$

$$\text{FRR}=\frac{\text{Number of rejected genuine claims (FR)}}{\text{Total number of genuine accesses}}\text{x}100\% \tag{12}$$

A unique measure, Total Success Rate (TSR) is obtained as Equation 13.

$$\text{TSR}=\left(1-\frac{\text{FA + FR}}{\text{Total number of accesses}}\right)\text{x}100\% \tag{13}$$

Another measure, Equal Error Rate (EER) is the datum on the Receiver Operating Characteristics curve where FAR is equal to the FRR.

Firstly, PCA is used to determine the initial face recognition rate. Two different databases were used to test the performance rate. Principal component (pc) is fixed to 100. The result shown that TSR=95.88% with FAR=4.11% and FRR=5% is obtained for Essex database. On the other hand, ORL database achieved TSR=92.31% with FAR=7.69% and FRR=7.5%. This baseline will be used as comparison for further experiments.

Next, an experiment was carried out to verify the performance rate of Non-Negative Factorization (NMF) as shown in Table 1. According to Table 1, different *r,* the number of basis vectors, were chosen to verify the best performance. *r*=10 to *r*=60 were used for verification rate calculation. The optimum verification rate for Essex Database is EER=1.90% with FAR=1.8%, FRR=2% when *r*=40 whilst ORL database attains EER=7.25% with FAR=7%, FRR=7.5% when *r*=20. There is a significant difference in the verification rate of these two databases as ORL database contains more variations to Essex Database, therefore yield poorer verification rate. Nevertheless both are achieving relatively satisfying results.

Table 1. Comparison of Essex and ORL database with NMF method

Database	*r*	FAR(%)	FRR(%)	EER(%)
Essex	10	2.50	2.00	2.25
	20	2.20	2.00	2.10
	30	3.00	3.00	3.00
	40	1.80	2.00	1.90
	50	4.90	5.00	4.95
	60	4.80	5.00	4.90
ORL	10	12.00	12.20	12.10
	20	7.00	7.50	7.25
	30	7.40	7.50	7.45
	40	8.00	7.50	7.75
	50	8.00	8.00	8.00
	60	8.00	8.00	8.00

Another experiment was carried out by using a similar set of r to determine the optimum verification rate when NMF is integrated with multiple wavelet filters with decomposition level 2. The optimum result is recorded in Table 2. Integration of haar filter and NMF achieved the best performance when r=20 whereas integration of Spline bior5.5 filter and NMF is best when r=40 for Essex and ORL database respectively. It can be observed that the optimum r is not big. This indicates that moderately large r of NMF is sufficient to discriminate the different face. Table 3 shows the comparison of PCA, NMF and wNMF for Essex and ORL databases respectively.

Table 2. Comparison of Essex and ORL database with integration of various types of wavelet filters and NMF

Database (Db)	**Filter**	**FAR(%)**	**FRR(%)**	**EER(%)**
Essex	Haar	1.56	2.00	1.58
	Db5	4.30	4.00	4.15
	Db10	5.00	5.00	5.00
	Sym5	4.60	5.00	4.80
	Sym10	7.70	8.00	7.85
	Bior1.1	5.00	5.00	5.00
	Bior5.5	4.00	4.00	4.00
ORL	Haar	5.90	5.00	5.45
	Db5	5.00	5.00	5.00
	Db10	5.00	5.00	5.00
	Sym5	5.10	5.00	5.05
	Sym10	5.00	5.00	5.00
	Bior1.1	5.00	5.00	5.00
	Bior5.5	4.50	5.00	4.75

Table 3. Comparison of PCA, NMF and wNMF for Essex and ORL database

Db	**Method**	**pc/r**	**FAR(%)**	**FRR(%)**	**TSR(%)**
Essex	PCA	100	4.11	5.00	95.90
	NMF	40	1.80	2.00	98.20
	wNMF	20	1.56	2.00	98.50
ORL	PCA	100	7.69	7.50	92.30
	NMF	20	7.00	7.50	93.00
	wNMF	20	4.50	5.00	95.50

Table 4. Training elapse time for Essex and ORL database

Database	WT	Elapse Time (sec)
Essex	Without	534.266
	With	240.094
ORL	Without	398.578
	With	151.547

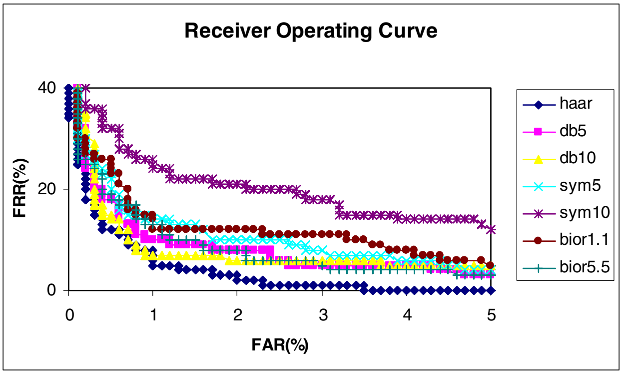


Fig. 4. Integration of various wavelet filters and NMF for Essex database

Table 4 indicates time computation which is implemented on a Pentium 4 2.66 GHz with 256Mb RAM processor. A longer duration is consumed to train the database due to the reason that 1000 iterations are used to update *W* and *H*. The difference between these two databases with the hybrid of NMF and WT is 88.547 seconds.

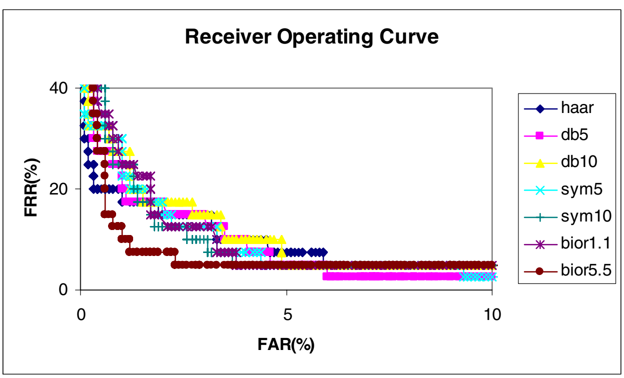


Fig. 5. Integration of various filters and NMF for ORL database

Figure 4 and 5 illustrate the Receiver Operating Characteristics (ROC) curve for the integration of multiple wavelets and NMF for Essex and ORL Database respectively. A ROC curve plots the FRR against FAR at various thresholds. The closer the plot lies to the axis, the better the performance. Thus Figure 4 and 5 reveals that haar filter and spline bior5.5 filter perform the best when the best chosen wavelet is integrated with NMF for both databases.

6 Conclusion

An efficient method for face recognition using Wavelet Transform and Non-Negative Matrix Factorization is presented in this paper. Two databases namely, Essex Database and ORL Database were used throughout the experiments to compare the verification rate. Wavelet analysis produces lower dimension multiresolution representation that alleviates heavy computational load, and also generates noise and minor distortion insusceptible to face wavelet – based template. After wavelet decomposition NMF is performed on the facial images. The results obtained are compared with the PCA technique. The results shown that wavelet and NMF can outperform PCA for better verification rate.

References

1. A.P. Ginsburg, Visual information processing based on spatial filters constrained by biological data, AMRL Technical Report, 1978, pp. 78-129
2. L.D. Harmon, The recognition of faces, Sci. Am. 229 (1973)
3. J. Sergent, Microgenesis of face perception, in: H.D. Ellis, M.A. Jeeves, F. Newcombe, A. Young (Eds.), Aspects of Face Processing, Nijhoff, Dordrecht, 1986
4. C. Nastar, N. Ayache, Frequency-based non-rigid motion analysis, IEEE Trans. Pattern Anal. Mach. Intell. 18 (11) (1996) pp. 1067-1079
5. S. T. Bow, Pattern Recognition and Image Preprocessing, Marcel Dekker, Inc., 1992, pp. 203-274
6. D.D.Lee, H.S.Seung, “Learning the Parts of Obejcts by Non-Negative Matrix Factorization”, Nature, vol. 401, pp. 788-791, 1999
7. M.Turk, A.Pentland, Engenfaces for recognition, J. Cognitive Neurosci. 3(1) 71-86, 1991
8. Baowen Xu, Jianjiang Lu, Gangshi Huang, “A Constrained Non-negative Matrix Factorization in Information Retrieval”, Proc. of The 2003 IEEE International Conference on Information Reuse and Integration (IRI’2003), October 27-29, 2003, Las Vegas, USA
9. M. Cooper, and J. Foote, “Summarizing Video Using Non-negative Similarity Matrix Factorization”, Proc. IEEE Workshop on Multimedia Signal Processing, 2002
10. P.Smaragdis, J.C.Brown, “Non-Negative Matrix Factorization for Polyphonic Music Transcription”, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2003
11. M. Novak, R. Mammone, “Use of Non-negative Matrix Factorization for Language Model Adaptation in a Lecture Transcription Task”, ICASSP, Salt Lake City, UT, USA, 2001
12. A. Laine. J. Fan, “Texture Classification by Wavelet Packet Signatures”, IEEE Trans. Pattern Anal. Machine Intell., Vol. 15, pp. 1186-1191, 1993
13. D.D.Lee, H.S.Seung, “Algorithms for Non-Negative Matrix Factorization”, Proceedings of Neural Information Processing Systems, vol. 13, pp. 556-562, 2001

14. Vision Group of Essex University Face Database http://cswww.essex.ac.uk/mv/allfaces/index.html
15. Olivetti Research Laboratory (ORL) Database http://www.uk.research.att.com/facedatabase.html