Selection of Wavelet Subbands Using Genetic Algorithm for Face Recognition

Vinod Pathangay and Sukhendu Das

Visualization and Perception Laboratory
Department of Computer Science and Engineering
Indian Institute of Technology Madras
Chennai - 600 036, India
{vinod@cse, sdas@}iitm.ernet.in

Abstract. In this paper, a novel representation called the *subband face* is proposed for face recognition. The subband face is generated from selected subbands obtained using wavelet decomposition of the original face image. It is surmised that certain subbands contain information that is more significant for discriminating faces than other subbands. The problem of subband selection is cast as a combinatorial optimization problem and genetic algorithm (GA) is used to find the optimum subband combination by maximizing Fisher ratio of the training features. The performance of the GA selected subband face is evaluated using three face databases and compared with other wavelet-based representations.

1 Introduction

The objective of the face recognition task is to determine the identity a person in a given face image. The problem is complicated due to the variations in illumination, pose, facial expression, aging, occlusions such as spectacles, hair, etc. Different techniques proposed for face recognition are surveyed in [1,2,3]. Although the human face is a 3-dimensional structure, only the two-dimensional projection or the 'appearance' of the face is captured by images. In this work, a new representation of the face image based on the wavelet subbands is proposed. The face image is decomposed into subbands using the Discrete Wavelet Transform (DWT) and selected subbands are used to reconstruct the face image. This is based on neurophysiological evidences [4,5] that humans process images similar to multi-resolution, multi-channel processing similar to the wavelet transform. It is therefore surmised that out of all the subbands obtained using DWT, only certain subbands contain the discriminatory information required for face recognition. If such subbands can be identified and the face image is reconstructed from these discriminatory subbands, then the reconstructed *subband face* image can give a better recognition performance. In this paper, GA is used to determine those subbands that contain the discriminatory information for a given population of face images in the presence of illumination variations.

A number of techniques for illumination invariant face recognition have been proposed [6,7,8]. Recently, spherical harmonics based representation has been proposed for illumination invariant recognition [9]. Wavelet-based features have

been used to obtain a representation for face recognition task in [10,11,12,13]. The use of statistical measures (mean, variance) of fully decomposed wavelet subbands or packets as a feature was reported in [10]. Here, 2D-DWT is used to fully decompose the face image and statistical features such as mean and variance are extracted from the wavelet coefficients, and used as a feature vector for representation. The use of wavelet subbands with kernel associative memory for face recognition was reported in [12]. Here, wavelet decomposed faces are used to build an associative memory model for face images of each person and kernel methods are used to exploit higher order relations which cannot be captured by linear transformations. Different wavelet subbands have been combined at different levels (data level, feature level, classifier level and decision level) for face recognition [13]. In [14], the face image was dyadically decomposed and approximation subbands at different levels were suppressed during reconstruction to obtain an *approximation-suppressed subband face*. In this paper, we explore the suppression of different fully decomposed subbands using GA, for selection of the *optimal subband face.*

Genetic algorithms have been used for selecting the optimal subspace in which the projected data gives a higher recognition accuracy [15,16,17,18,19,20]. In [15], GA has been used to find the Optimal Projection Axes (OPA) for face recognition. The OPA is found by searching through all possible rotations in the whitened PCA subspaces. The fitness function used is a weighted sum of class separation and performance accuracy. It is reported that the OPA selected by the GA yields better performance compared to the eigenspace technique. Face recognition using kernel PCA and GA has been reported in [21], where the input data is transformed to higher dimension using a non-linear transfer function (polynomial function) and GA is used to select the optimal subset of the non-linear principal components with the fitness function taken as the recognition performance. An algorithm for Independent Component Analysis (ICA) is used to represent the face image in terms of non-orthogonal bases in order to capture the higher order statistics [18]. Apart from the optimum subspace selection techniques, selection of optimum non-orthogonal wavelet for representing the face image has also been reported [22]. Here, the GA is used to find optimal basis from a combination of frequencies and orientation angles in the 2D Gabor wavelet transform and entropy is used as the fitness function instead of class separation and performance. In this paper, the face image is fully decomposed to obtain different subbands or wavelet packets and GA is used to select an optimum combination of subbands. The optimal subband face is reconstructed from the selected subbands that contain the desired discriminatory information and suppressing those subbands that contain the similarity information.

The rest of the paper is organized as follows. The proposed subband face representation is described in section 2. Section 3 discusses the use of genetic algorithm for selecting optimal subband combination for generating the subband face. In section 4, the proposed representation is evaluated on three different face databases and the recognition results are compared with other similar

techniques. Section 5 concludes the paper and some directions of future work are discussed.

2   Subband Face Representation

A face image of a person contains common (approximation) 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. These similarity information and discriminatory information are segregated in different subbands at different levels of decomposition of the face image. Therefore wavelet decomposition can be used to split the features in a face image into different subbands, with 'approximations' containing the common (smooth) parts of the face and 'details', containing the discriminatory (variations) information. Since level-1 decomposition may not be adequate to effectively isolate these pair of visual features, it is necessary to explore different combination of subbands at higher levels to obtain a suitable isolation.

2.1   Wavelet Decomposition

The wavelet transform [23,24] is expressed as an inner product of a signal $f(x)$ with a family of functions which are translations and dilations of a mother wavelet function $\psi(x)$ and the scaling function $\phi(x)$. This is used to split the signal into approximation and detail. For two-dimensional signal or image, the two-dimensional scaling function $\phi(x,y)$ and three two-dimensional wavelet functions $\psi^H(x,y)$, $\psi^V(x,y)$, $\psi^D(x,y)$ are used to produce the approximation and horizontal, vertical, diagonal details respectively. The two-dimensional discrete wavelet transform (2D-DWT) for an image $f(x,y)$ gives the approximations W_ϕ and the details W_ψ^d

$$W_\phi(s_0,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\phi_{s_0,m,n}(x,y) \tag{1}$$

$$W_\psi^d(s,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\psi_{s,m,n}^d(x,y) \tag{2}$$

where $d = \{H,V,D\}$ for horizontal, vertical and diagonal details, $\phi(x,y)$ is the scaling function and $\psi^H(x,y)$, $\psi^V(x,y)$ and $\psi^D(x,y)$ are the three two-dimensional wavelets, s_0 is the starting scale (taken as zero) and $N = M = 2^S$ such that $s = 0,1,2,...,S-1$ and $m,n = 0,1,2,...,2^s-1$. For level-1 decomposition of $f(x,y)$, $s = 0$. Thus the 2D-DWT uses a family of wavelet functions and its associated scaling function to decompose the original image into four subbands, namely the approximation (A), horizontal (H), vertical (V) and diagonal
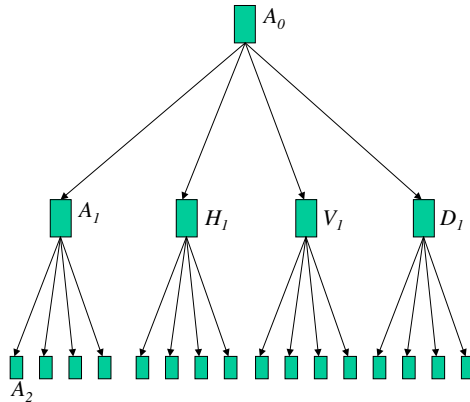
Fig. 1. Wavelet decomposition tree up to level-2. The root node (A_0) represents the original image which is decomposed to A_1, V_1, H_1, D_1. They are in turn decomposed to level-2 subbands which are the leaf nodes.

(D) details. Fig. 1 shows the full level-2 wavelet decomposition tree. The original image A_0 is decomposed to level-1 subbands (A_1, H_1, V_1, D_1) which are in turn decomposed to obtain level-2 subbands. We have used the lifting technique [25] for the 2D-DWT with Haar basis. In the following subsection, the method of reconstructing a subband face using selective wavelet subbands is discussed.

2.2 Reconstruction from Subbands

In order to reconstruct the subband face that contains only the discriminatory information, certain subbands can be selected or retained and others are suppressed during reconstruction. Reconstruction is done using a method similar to inverse DWT. Given the level-1 approximations W_ψ and the details $W_\phi^H, W_\phi^V, W_\phi^D$, the subband image $\hat{f}(x, y)$ is reconstructed using the relation

$$\hat{f}(x,y) = \frac{1}{\sqrt{MN}} \sum_m \sum_n W_\psi(m,n)\psi_{m,n}(x,y)b_A$$

$$+ \frac{1}{\sqrt{MN}} \sum_{d=H,V,D} \sum_m \sum_n W_\phi^d(m,n)\phi_{m,n}^d(x,y)b_d \qquad (3)$$

where b_A, b_H, b_V, b_D are the binary subband selection variables that can be set to zero or one if the subband is to be selected or suppressed respectively. For level-1 subbands, the sequence of four bits $\{b_A, b_H, b_V, b_D\}$ is the *subband code* for generating $\hat{f}(x, y)$. In order to reconstruct the discriminatory subband face from level-l subbands, a 4^l length subband code is to be determined. It can be noted that the number of possible subband combinations at level-l is 2^{4^l}. Therefore at higher levels of decomposition, there are a large number of subband combinations. To

cater for the search in this large search space, the subband selection problem is cast as a combinatorial optimization problem and a genetic algorithm is used to find an optimum subband code. The following section describes the use of GA and the fitness functions used in selecting a suitable subband code that gives a high recognition accuracy.

3   Subband Selection Using Genetic Algorithm

Due to the exponential increase in the subband combinations with higher levels of decomposition, it is not possible to select an optimum combination using an exhaustive search. Therefore genetic algorithm [26] is used for finding an 'optimal' solution or subband code. GA maintains a population of subband combinations or subband codes and iteratively finds a subband code that has the high fitness value among all subband codes of all generations. The subband code is used as the binary chromosome in the GA. Each subband code has an associated fitness value. Fisher ratio of the training features is used as the fitness function. This is the ratio of the between-class scatter and the total within-class scatter of the training features. This is given by

$$F = \frac{\sum\limits_{i} \sum\limits_{r \neq i} ||\boldsymbol{\mu}_i - \boldsymbol{\mu}_r||^2}{\sum\limits_{i} \sum\limits_{j \in C_i} ||\boldsymbol{\mu}_i - \mathbf{x}_j||^2} \tag{4}$$

where i, r are the class indices and j is the feature index, $\boldsymbol{\mu}_i$ mean feature of the class i, $\mathbf{x}_j$ is the jth feature, and C_i is the set of training features of class i. The use of Fisher ratio as the fitness function is motivated by the fact that a large between-class spread and smaller within-class spread will lead to compact clusters with larger separation and thus lesser confusion or overlap between the classes.

The following are the steps involved in the GA used for subband selection:

1. *Initialization*: Start with a random population of subband codes $\mathbf{b}_p$, $p \in [1, P]$ where P is the number of subband codes in the population.
2. *Evaluation*: Evaluate fitness $F(\mathbf{b}_p)$ $\forall p$.
3. *Selection*: Select P_s (where $P_s < P$) fittest subband codes for next generation.
4. *Crossover*: From the fittest subband codes selected in the above step, apply crossover on two randomly chosen samples to generate new subband code. This is repeated to obtain P_c (where $P_c \leq P_s$) new subband codes for the next generation.
5. *Mutation*: Apply mutation to randomly selected sample obtained using crossover to generate a new subband code. This is repeated to obtain P_m (where $P_m \leq P_c$) new subband codes.
6. Repeat steps 2-5 (evaluation, selection, crossover and mutation) for each generation until the fitness converges to a high value.

The number of solutions P for each generation is maintained constant by ensuring that $P_s + P_c + P_m = P$. For applying the *crossover* operator on two parent subband codes selected at random, a crossover position is chosen and the bits on the left of the crossover point from one parent and those on the right from the other parent are copied to the child subband code. Thus the child inherits code segments from both parents. For applying the *mutation* operator on a subband code, a bit value is flipped at a randomly chosen location. In this work, a batch GA which has a constant population for all generations is used. The population size is made ten times the length of the chromosome bitstring. This is to have an increased exploration where there are more free parameters. The values for $\frac{P_s}{P}$, $\frac{P_c}{P}$, $\frac{P_m}{P}$ are set to 0.5, 0.3 and 0.2 respectively. In the following section, the experimental results are discussed.

(a)

(b)

(c)

Fig. 2. Face databases used: (a) Yale (b) ORL and (c) PIE. First row shows the training samples and other rows show the testing samples used for experiments.

4   Experiments

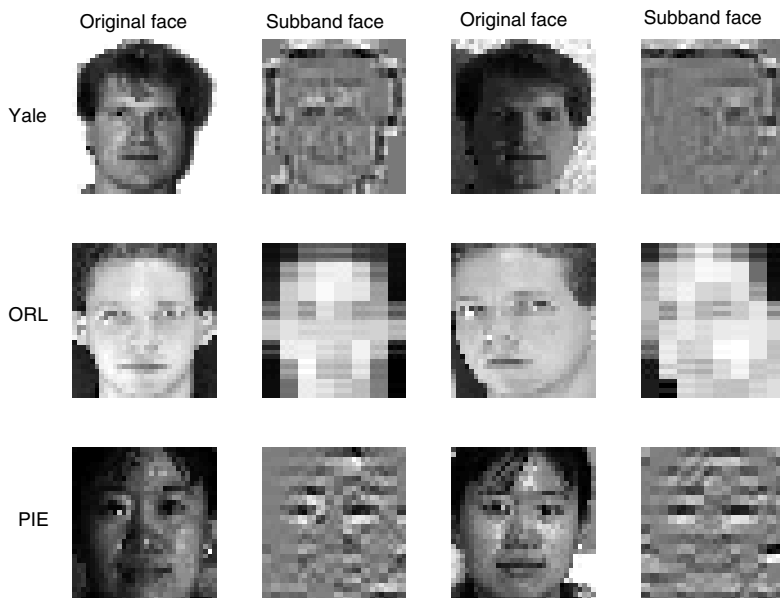In this paper, three face databases are used: Yale, ORL and PIE.

Fig. 3. Original face and the subband face reconstructed from the GA selected subband codes for Yale (top row), ORL (mid row) and PIE (bottom row)

1. The Yale face database (Fig. 2 (a)) has 15 subjects with 11 samples with variations in expression and illumination for each subject. The training sample set contains *rightlight*, *sad*, *noglasses*, *happy* and the testing sample set consists of *centerlight*, *leftlight*, *glasses*, *sleepy*, *surprised*, *wink*, and *normal* samples.
2. The ORL face database (Fig. 2 (b)) has 40 subjects with 10 samples for each. There is no change in illumination but there are variations in pose. Training samples used are images numbered 1 to 4 and the rest are used for testing.
3. A subset of the PIE (Pose Illumination and Expression) face database (Fig. 2 (c)) [27] with 60 subjects with only the frontal poses is used. For each subject, 42 samples (flash from 21 different directions with and without the room lights on) are used. Only four samples with flash numbers 12, 13 for both *lights* and *illum* sets are taken for training and the remaining for testing.

The face region is extracted from the background and resized to 32x32 pixels for generating the subband face. The subband face is obtained using the subband code determined by the GA. Fig. 3 shows the original gray-level face image and the subband face generated for sample face images taken from three databases. Although the subband face does not appear similar to the original gray-level face image, it preforms much better than the original face images because it contains only the discriminatory information that is required for a machine to obtain higher recognition accuracy.
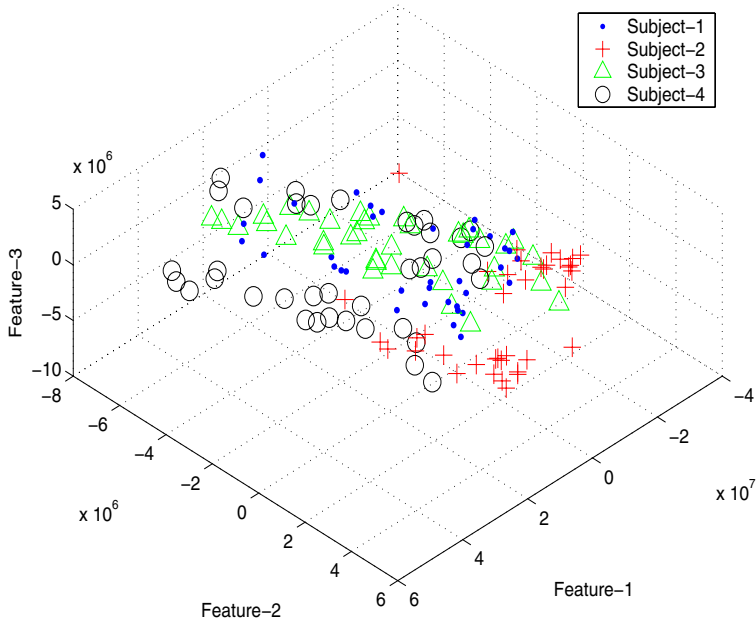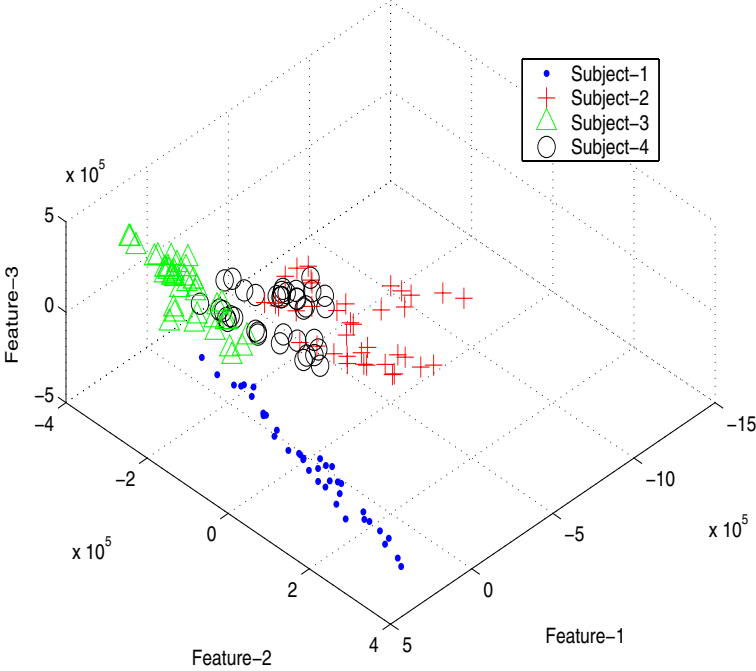
Fig. 4. Features in 3D eigenspace for (a) gray-level and (b) subband faces determined by the GA for four subjects of the PIE database

Table 1. Comparison of the accuracies (%) of face recognition techniques published earlier using wavelet based features. ($A_0 - A_l$ denotes the suppression of approximation A_l during reconstruction of the subband face).

Database	Baseline Eigenface [28]	Approximation suppressed [14]	Best performing subband [13]
Yale	81.9	$(A_0 - A_1)$ 88.6 $(A_0 - A_2)$ 91.4 $(A_0 - A_3)$ 89.5	(D_2) 81.9
ORL	77.1	$(A_0 - A_1)$ 60.0 $(A_0 - A_2)$ 65.0 $(A_0 - A_3)$ 72.5	(A_1) 79.5
PIE	34.3	$(A_0 - A_1)$ 87.1 $(A_0 - A_2)$ 85.3 $(A_0 - A_3)$ 75.4	(H_1) 93.1

Table 2. Accuracy of *subband face* selected by GA using Fisher ratio as the fitness function for PCA features

Database	Decomposition level	Subband code (Hex)	Accuracy (%)
Yale	2	5333	90.5
	3	0354-1040-00CB-0098	87.6
ORL	2	8800	78.5
	3	E888-E084-8004-0000	76.5
PIE	2	5101	88.6
	3	0808-0006-0802-0001	81.1

Fig. 4 (a) shows the projection of the original face images of four subjects of the PIE database in 3D eigenspace. It can be observed that there is a significant overlap of the clusters. Such overlap often leads to misclassification using the nearest-neighbor rule. Fig 4 (b) shows the projection of the subband faces in 3D eigenspace. It can be seen that there is an increase in the separation and compactness of the eigenfeatures of the subband faces generated using the subband code selected by the GA. Therefore there are less chances of misclassification using nearest neighbor rule.

Table 1 shows a comparison of the performance of wavelet-based representations [14,13] against the baseline Eigenface method. Table 2 shows the recognition performance of the GA selected subband face. It can be observed that the performance of the GA selected subband face is comparable to the approximation suppressed subband face and the best performing subbands given in Table 1. It can also be noted that the level-3 subband code found by the GA always gives lesser accuracy compared to level-2. This is due to the sub-optimal determination of the large number of free parameters (64) in level-3 subband code. This is also due to the fact that increase in Fisher ratio of the training

features does not guarantee a higher recognition accuracy of test features. It can also be noted from Table 2 that the subband code selected by the GA is dependent on the database. The subband codes determined for Yale and PIE (5333 and 5101 respectively) are very different form that of ORL (8800). This is because ORL has pose variations that are absent in Yale and PIE. The similarity between the subband codes of Yale and PIE is due to the presence of significant lighting changes in the training sets. There is also a similarity between the subband codes selected by the GA and the best performing subbands [13]. Therefore, it can be generalized to a certain extent that, when there is no pose variation and only lighting changes are present, then the level-2 subband code of the form 5*** (which includes the horizontal and diagonal details of level-1 approximation subband) works best. This eliminates the level-2 approximation (A_2) similar to [14]. However, when there are changes in pose, as in the case of ORL database, the level-2 subband code of the form 88** works best. This is because the subband code 88** retains the smoothed approximations of A_1 and H_1 subbands that are invariant to minor pose changes.

5   Conclusion and Future Work

In this paper, we have proposed the subband face as a representation for face recognition. A genetic algorithm is used to select the optimal choice of subbands by using Fisher ratio as the fitness function. It is shown that the GA is able to find the subband codes that contain only the discriminatory information in the face image of different persons of a given population. The performance of the subbands selected by the GA to other wavelet-based representations are compared. The GA selected subbands are also invariant to lighting changes as shown for the PIE database. The subband face is also scalable as the accuracy is consistent with increasing database size (15 to 60 subjects for Yale and PIE respectively). Future work includes exploring the possibility of using class specific subband codes, and use of single training face image for generating virtual training samples with variations in illumination, pose and scale. The use of subband images can also be used for non-face classes.

References

1. Zhao, W., Chellappa, R., Rosenfeld, A., Phillips, P.: Face recognition: A literature survey (2000)
2. Chellappa, R., Wilson, C.L., Sirohey, S.: Human and machine recognition of faces: A survey. Proc. of the IEEE **83** (1995) 705–740
3. Tolba, A.S., El-Baz, A.H., El-Harby, A.A.: Face recognition: A literature review. Int. Journal of Signal Processing **2** (2005) 88–103
4. Kanwisher, N.: Domain specificity in face perception. Nature neuroscience **3** (2000) 759–776
5. Liu, J., Harris, A., Kanwisher, N.: Stages of processing in face perception: an MEG study. Nature neuroscience **5** (2002) 910–916

6. Gross, R., Baker, S., Matthews, I., Kanade, T.: Face recognition across pose and illumination. In Li, S.Z., Jain, A.K., eds.: Handbook of Face Recognition. Springer-Verlag (2004)

7. Lee, J., Machiraju, R., Pfister, H., Moghaddam, B.: Estimation of 3d faces and illumination from single photographs using a bilinear illumination model. In: Proc. Eurographics Symposium on Rendering (EGSR). (2005)

8. Lee, K.C., Moghaddam, B.: A practical face relighting method for directional lighting normalization. In: Int. Workshop on Analysis and Modeling of Faces and Gestures (AMFG). (2005)

9. Zhang, L., Samaras, D.: Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics. IEEE Tran. on Pattern Analysis and Machine Intelligence **28** (2006) 351–363

10. Garcia, C., Zikos, G., Tziritas, G.: Wavelet packet analysis for face recognition. Image and Vision Computing **18** (2000) 289–297

11. Chien, J.T., Wu, C.C.: Discriminant waveletfaces and nearest feature classifier for face recognition. IEEE Tran. on Pattern Analysis and Machine Intelligence **24** (2002) 1644–1649

12. Zhang, B.L., Zhang, H., Ge, S.S.: Face recognition by applying wavelet subband representation and kernel associative memory. IEEE Tran. on Neural Networks **15** (2004) 166–177

13. Ekenel, H.K., Sankur, B.: Multiresolution face recognition. Image and Vision Computing **23** (2005) 469–477

14. Pathangay, V., Das, S.: Exploring the use of selective wavelet subbands for PCA based face recognition. In: Proc. National Conference on Image Processing (NCIP'05), Bangalore, India (2005) 182–185

15. Liu, C., Weschler, H.: Evolution of optimal projection axes (opa) for face recognition. In: Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition. (1998) 282–287

16. Liu, C., Weschler, H.: Evolutionary pursuit and its application to face recognition. IEEE Tran. on Pattern Analysis and Machine Intelligence **22** (2000) 570–582

17. Zhang, X., Mersereau, R.M., Broun, C.C.: Visual speech feature extraction for improved speech recognition. In: Proc. of the Int. Conf. on Acoustics Speech and Signal Processing, Orlando, Florida, USA (2002) 1993–1996

18. Yi-qiong, X., Bi-Cheng, L., Bo, W.: Face recognition by fast independent component analysis and genetic algorithms. In: Proc. 4th Int. Conf. on Computer and Information Technology. (2004) 194–198

19. Karungaru, S., Fukumi, M., Akamatsu, N.: Face recognition using genetic algorithm based template matching. In: Proc. Int. Symposium on Communications and Information Technologies. (2004) 1252–1257

20. Zhao, Q., Lu, H., Zhang, D.: A fast evolutionary pursuit algorithm based on linearly combining vectors. Pattern Recognition **39** (2006) 310–312

21. Yankun, Z., Chongqing, L.: Face recognition using kernel principal component analysis and genetic algorithms. In: Proc. 12th IEEE Workshop on Neural Networks for Signal Processing. (2002) 337–343

22. Wang, X., Qi, H.: Face recognition using optimal non-orthogonal wavelet basis evaluated by information complexity. In: Proc. 16th Int. Conf. on Pattern Recognition. Volume 1. (2002) 164–167

23. Mallat, S.G.: A theory for multiresolution signal decomposition: The wavelet representation. IEEE Tran. on Pattern Analysis and Machine Intelligence **11** (1989) 674–693

24. Fliege, N.J.: Multirate Digital Signal Processing. John Wiley and Sons, Ltd. (2004)
25. Jensen, A., Cour-Harbo, A.: Ripples in Mathematics: The Discete Wavelet Transform. Springer-Verlag (2003)
26. David E. Goldberg: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley (1989)
27. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression database. IEEE Tran. on Pattern Analysis and Machine Intelligence **25** (2003) 1615–1618
28. Turk, M.A., Pentland, A.: Face recognition using eigenfaces. In: Proc. 11th Int. Conf. Pattern Recognition. (1991) 586–591