ELSEVIER

Multiresolution face recognition

Hazim Kemal Ekenel*, Bülent Sankur

Electrical and Electronic Engineering Department, Boğaziçi University, Bebek, İstanbul, Turkey

Abstract

In this paper the contribution of multiresolution analysis to the face recognition performance is examined. We refer to the paradigm that in classification tasks, the use of multiple observations and their judicious fusion at the data, feature or decision level improves the correct decision performance. In our proposed method, prior to the subspace projection operation like principal or independent component analysis, we employ multiresolution analysis to decompose the image into its subbands. Our aim is to search for the subbands that are insensitive to the variations in expression and in illumination. The classification performance is improved by fusing the information coming from the subbands that attain individually high correct recognition rates. The proposed algorithm is tested on face images that differ in expression or illumination separately, obtained from CMU PIE, FERET and Yale databases. Significant performance gains are attained, especially against illumination perturbations.
© 2004 Elsevier B.V. All rights reserved.

1. Introduction

Face recognition problem has become one of the most relevant research areas in pattern recognition. Face recognition debts its popularity to its potential application areas, ranging from human computer interaction to authentication and surveillance.

The holistic or appearance-based approach has been gaining popularity vis-à-vis anthropometrical feature-based approach in face recognition [1]. In the holistic approach, all the pixels in the entire face image are taken as a single signal, and processed to extract the relevant features for classification. Most of the appearance-based face recognition algorithms perform some kind of subspace analysis in the image space to extract the relevant feature vectors. The most widely used subspace analysis tools are the principal component analysis (PCA) [2], linear discriminant analysis (LDA) [3] and a blind source separation technique, called independent component analysis (ICA) [4]. All face recognition algorithms, however, witness a performance drop whenever face appearances are subject to variations by

factors such as occlusion, illumination, expression, pose, accessories and aging. In fact, often these factors lead to intra-individual variability of face images, to the extent that they can be larger than the inter-individual variability [5].

In this study, we apply multiresolution techniques in order to mitigate the loss of classification performance due to changes in facial appearance. We design experiments specifically to investigate the gain in robustness against illumination and facial expression changes. The underlying idea in the use of the multiresolution analysis is firstly, to obtain multiple evidences from the same face, and search for those components that are less sensitive to intrinsic deformations due to expression or due to extrinsic factors, like illumination. Secondly, our approach follows the paradigm of fusion that utilizes multiple evidences. Although at first sight, these evidences can appear somewhat redundant and may contain less information, their judicious combination can prove often to be superior for classification.

The most popular multiresolution analysis technique is the wavelet transform. Therefore in this study we use the 2D discrete wavelet transform in order to extract multiple subband face images. These subband images contain coarse approximations of the face as well as horizontal, vertical

___

* Corresponding author.

E-mail address: ekenelha@boun.edu.tr (H. Kemal Ekenel).

and diagonal details of faces at various scales. Subsequently, we extract PCA or ICA features from these subbands. We exploit these multiple channels by fusing their information for improved recognition. We have compared three fusion approaches, namely, fusion at the subband data level, fusion at the ICA/PCA feature level, and finally, fusion of the classifier decisions at the subband channel level. The main contribution of the paper is thus to search for most discriminative set of wavelet channels, and to construct face recognition schemes using fusion techniques at different levels of data processing.

Discrete wavelet transform has been used in various studies on face recognition [6–10]. In [6], three-level wavelet transform is performed to decompose the original image into its subbands, on which the PCA is applied. The experiments on Yale database show that third level diagonal details attain highest correct recognition rate. A wavelet transform-based speaker identification system in a tele-conferencing environment is proposed in [7]. In this algorithm a three-level wavelet decomposition is performed. The scaling components at each level as well as the original image are used for classification. The classifier used in this study is a kind of neural network with one-class-in-one-network structure, that is, each subnet is trained separately and there is one subnet per individual. Wavelet packet analysis-based face recognition system is proposed in [8]. The original image is decomposed into its subbands by using two-level wavelet packet decomposition. Out of the 16 subbands, a 21-dimensional feature vector is obtained consisting of variances of 15 detail subbands and three mean values and three variances calculated from different parts of the approximation subband. From this 21 components, only the most meaningful components are selected resulting in a final feature vector size of 11. Bhattacharya distance between these statistical features is used to classify faces. In [9], three-level wavelet decomposition is performed and the resulting approximation subbands at each level are concatenated to produce a new data vector on which PCA is applied. Radial basis functions are used as the classifier of the system. Discriminant waveletfaces approach is proposed in [10]. In this study, third level approximation resulting from three level wavelet decomposition, called the wavelet-face, is used as the input of the LDA. For classification, presented nearest feature plane (NFP) and nearest feature space (NFS) classifiers are examined. Different from these previous studies, we put into evidence the contribution of wavelet subbands to combat; specifically, illumination and expression factors, and we investigate the interplay of subband information fusion styles, choice of metrics and of features. In other words, the thrust of the paper is to explore how the discriminatory ICA and PCA features can be de-sensitized or rendered more invariant to the effects of illumination and expression via the judicious selection of subbands and via fusion at various levels.

The paper is organized as follows. In Section 2, multiresolution analysis is briefly reviewed. Subspace analysis tools (PCA, ICA) and fusion techniques used in the study are explained in Sections 3 and 4, respectively. In Section 5 experimental results against expression and illumination variations are presented separately. Finally, in Section 6 conclusions are given.

2. Multiresolution analysis

Multiresolution methods provide powerful signal analysis tools, which are widely used in feature extraction, image compression and denoising applications. Wavelet decomposition is the most widely used multiresolution technique in image processing. Images have typically locally varying statistics that result from different combinations of abrupt features like edges, of textured regions and of relatively low-contrast homogeneous regions. While such variability and spatial nonstationarity defies any single statistical characterization, the multiresolution components are more easily handled. Wavelet transform can be performed for every scale and translation, resulting in continuous wavelet transform (CWT), or only at multiples of scale and translation intervals, resulting in discrete wavelet transform (DWT). Since, CWT provides redundant information and requires a lot of computation, generally DWT is preferred.

The two-dimensional wavelet transform is performed by consecutively applying one-dimensional wavelet transform to the rows and columns of the two-dimensional data. In Fig. 1, the tree representation of one level, two-dimensional wavelet decomposition is shown. In this figure, G denotes high-pass filtering and H denotes low-pass filtering, while $\downarrow 2$, represents downsampling by a factor of 2. In the final stage of the decomposition we have four $N/2 \times N/2$ resolution subband images: A_1, the scaling component containing global low-pass information, and three wavelet components, H_1, V_1, D_1, corresponding, respectively, to the horizontal, vertical and diagonal details. We can perpetuate this decomposition, either pursuing the same pattern along the scaling component, or obtaining the full-blown tree, or achieve some intermediate tree where 'interesting' branches
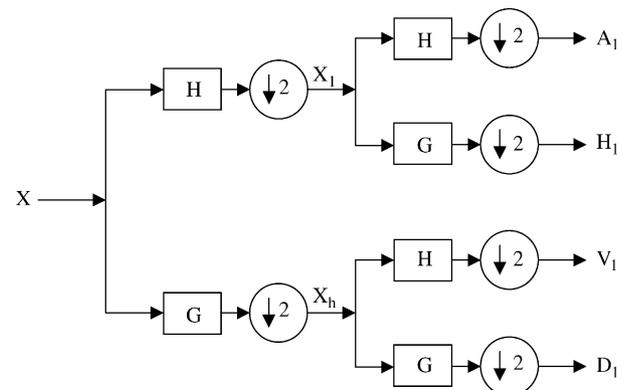
Fig. 1. Tree representation of one-level 2D wavelet decomposition.
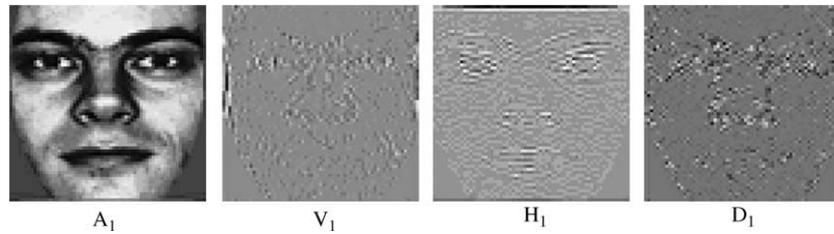
Fig. 2. Sample one-level wavelet decomposed image.

are grown using, for example, some projection pursuit scheme [11]. In Fig. 2, one-level wavelet decomposition of a face image is shown.

In Fig. 3 the schematics of the wavelet decomposition used in this study is shown. The letters in the figure serve to differentiate the scaling component or the orientations of the wavelet components, while the accompanying numbers denote the level of decomposition. If the subbands are obtained by decomposing the original image or any of the scaling components, then they are represented with single letter. If however, a subband is derived by decomposing one of the detail subbands, then these are denoted with two letters, where the first letter indicates the parent subband and the second letter denotes the orientation of the child.

In the first level, a 128×128 original face image is decomposed and four 64×64 pixels resolution subband images—A_1, H_1, V_1 and D_1—are obtained. The H_1, V_1, and D_1 components are not further decomposed, because we found their classification performance figures to be very low. Consequently we proceed to decompose only A_1, yielding four 32×32 subband images—A_2, H_2, V_2 and D_2. In the third level, we decompose all components, A_2, H_2, V_2 and D_2, producing 16 16×16 subband images. In summary, we obtain 24 different subband images from the original face image and input them into the classification scheme.

3. Subspace analysis

An $m \times n$ resolution face image can be considered as a point in an $N = m \times n$ dimensional image space. For example, a 128×128 face image corresponds to a point in 16,384-dimensional huge feature space. On the other hand, face images are very similar, and therefore highly correlated. It follows than that they can be represented in a much lower dimensional feature subspace. PCA and ICA are the two popular methods to descend to such face subspaces.

3.1. Principal component analysis (PCA)

Principal component analysis (PCA) is based on the second-order statistics of the input image, which tries to attain an optimal representation that minimizes the reconstruction error in a least-squares sense. Eigenvectors of the covariance matrix of the face images constitute the eigenfaces. The dimensionality of the face feature space is reduced by selecting only the eigenvectors possessing largest eigenvalues. Once the new face space is constructed, when a test image arrives, it is projected onto this face space to yield the feature vector—the representation coefficients in the constructed face space. The classifier decides for the identity of the individual, according to a similarity score between the test image's feature vector and the PCA feature vectors of the individuals in the database.

3.2. Independent component analysis (ICA)

Independent component analysis (ICA) can be seen as a tool, based on higher order statistics, for extracting independent sources from an observed mixture, where neither the mixing matrix nor the distribution of the sources
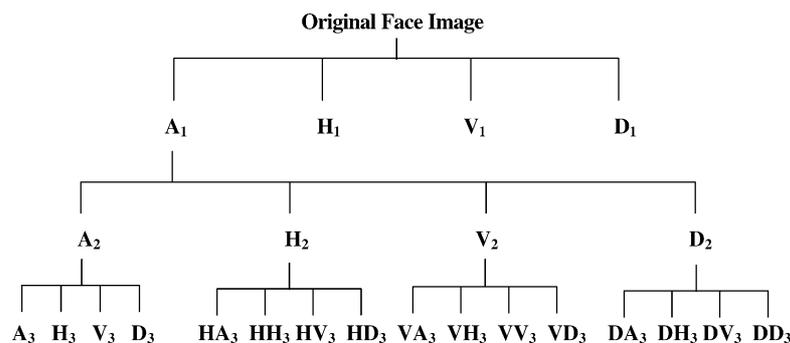
Fig. 3. Wavelet decomposition tree used in the study.
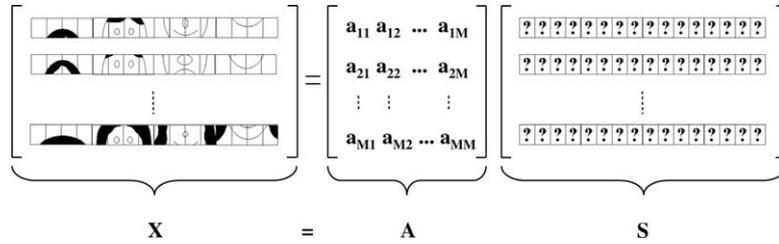
Fig. 4. First face recognition architecture of ICA (ICA1). **X**: observation, **A**: rows of mixing matrix, representation coefficients, **S**: statistically independent basis images.
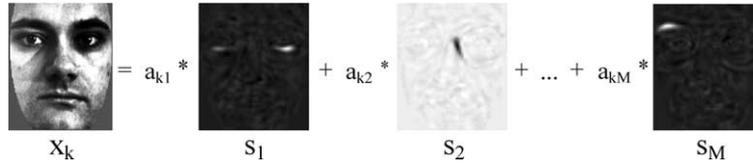
Fig. 5. Face representation using ICA1.

are known. The system model of ICA is given as

$$\mathbf{X} = \mathbf{AS}$$

where **A** denotes the mixing matrix, **S** denotes the source matrix containing statistically independent source vectors in its rows, and **X** denotes the observation matrix containing the 'linear mixtures' in its rows. The un-mixing matrix **W** is found by minimizing or maximizing some objective function, such as likelihood ratio, network entropy, mutual information or Kullback–Leibler divergence [12].

The separation matrix, **W**, under ideal conditions, is the inverse of the mixing matrix **A**

$$\mathbf{Y} = \mathbf{WX} \text{ and } \mathbf{W} = \mathbf{A}^{-1} \text{ and } \mathbf{Y} \approx \mathbf{S}$$

In the context of face recognition, the use of ICA features was first proposed in [4], where two different approaches were presented. In the first approach (called ICA1 architecture), the face images are assumed to be a linear mixture of an unknown set of statistically independent source images. Therefore, in this architecture,
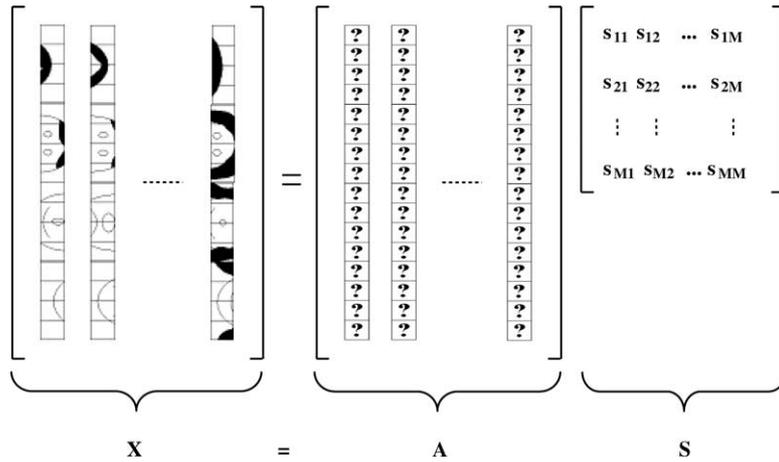
Fig. 6. Second face recognition architecture of ICA (ICA2). **X**, observations; **A**, columns of mixing matrix; **S**, statistically independent representation coefficients.
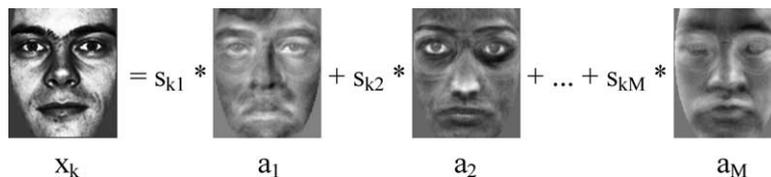
Fig. 7. Face representation using ICA2.
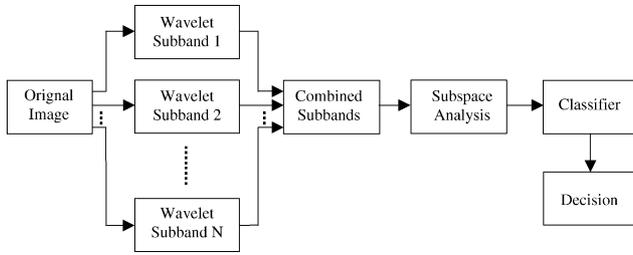
Fig. 8. Block diagram of the data fusion scheme.

the lexicographically ordered face images constitutes the rows of the observation matrix **X**, the statistically independent basis images constitutes the rows of the source matrix **S**, and the representation coefficients constitutes the rows of the mixing matrix **A** (Fig. 4). The source images obtained in this architecture are spatially local and sparse in nature (Fig. 5).

In the second approach (called ICA2 architecture), the representation (weighting) coefficients are assumed to be statistically independent. Therefore, in this architecture, the lexicographically ordered face images constitute the columns of the observation matrix **X**, the statistically independent representation or weighting coefficients constitute the columns of the source matrix **S**, and the basis images constitutes the columns of the mixing matrix **A** (Fig. 6). In this second architecture, while mixing coefficient vectors are independent, source images tend to have global face appearances, similar to the case of PCA (Fig. 7).

4. Fusion

The outcomes from the various wavelet channels are fused to achieve possibly higher correct recognition rates. We investigated three schemes, namely, fusing raw pixel values of the subbands, fusing ICA/PCA feature vectors extracted from the subbands, and fusing the classification decisions of the subbands.

4.1. Data fusion

In data fusion, lexicographically ordered pixels of the subband images are concatenated to construct a new data vector. Following this operation, the subspace projection
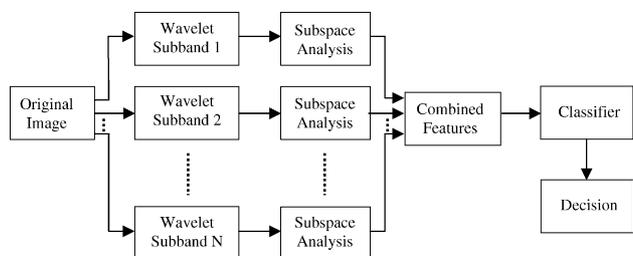
Fig. 9. Block diagram of the feature fusion scheme.
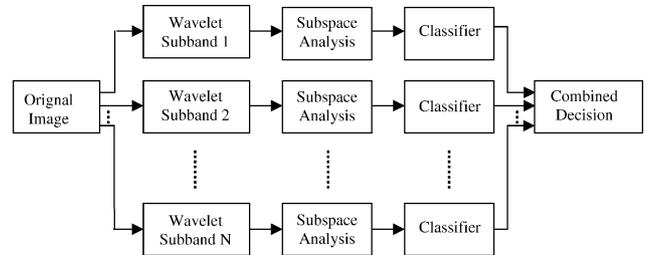
Fig. 10. Block diagram of decision fusion scheme.

and feature extraction are performed on the combined data vectors (Fig. 8).

4.2. Feature fusion

In feature fusion, subspace analysis tools are performed on each subband, and then the extracted feature vectors are concatenated to construct a new feature vector to be used for classification (Fig. 9).

4.3. Decision fusion

In decision fusion, face classification is run separately in each subband. According to the distance values between the test face feature vector and feature vectors of the individuals in the database, a confidence measure is calculated for each classifier. If we have K images in the database and if we define the distance between two feature vectors by the function $d(\cdot)$, then the confidence score, c_i, of a classifier's decision for ith class, is proportional to:

$$c_i = \frac{\sum_{k=1}^{K} d(x_{\text{test}}, x_{\text{database},k})}{d(x_{\text{test}}, x_{\text{database},i})}$$

The final decision is made through these confidence values by using sum rule, product rule or maximum rule [13]. Note that there are other fusion techniques based on the training with decision patterns of experts. For example, each subband can be considered to be a 'face recognizer expert' and their decisions could be fused via a neural

Fig. 11. Sample face images containing changes in expression—first row from CMU PIE, second row from FERET.

Table 1
Correct recognition rates of successful subband images against changes in expression

	PCA-120			ICA1-120			ICA2-120		
	L1	L2	CC	L1	L2	CC	L1	L2	CC
Original	92.33	90.33	92	93.33	90.33	92	87	87.67	96
A_1	93	91	91.67	93.67	91	91.67	90	88.67	95.33
A_2	94	93	93	93	93	93	90.33	92	96
A_3	94	91.67	91.67	94.33	91.67	91.67	90.33	93.33	95

network. In our case, the limitations in the training data precluded this approach (Fig. 10).

5. Experiments

Two separate experiments are conducted to test the advantage of the wavelet-based face recognition scheme. In the first experiment, the subbands that are potentially insensitive to changes in expression are searched, whereas in the second experiment the subbands that are insensitive to variations in illumination are searched. In both the experiments, feature vectors are extracted from the subband images via PCA, ICA1 and ICA2. The FastICA algorithm [12] is used to perform ICA. Daubechies 4 wavelet is used in the study. The Daubechies wavelets, with their compact support and orthonormal nature, are one of the most widely used wavelet families [14]. Besides this, in [6] it is shown that Daubechies 4 wavelet performs best in terms of computation time and recognition performance with respect to other order Daubechies wavelets, and other well-known wavelets such as biorthogonal, Symlets and Lemarie.

We used the nearest neighborhood classifier in our study. We evaluated comparatively three different distance metrics, namely, the $L1$ norm, the $L2$ norm, and the normalized correlation coefficient, defined as follows

$$L1: \ d = \sum_{m=1}^{M} |f_{\text{training},m} - f_{\text{test},m}|$$

$$L2: \ d = \left(\sum_{m=1}^{M} |f_{\text{training},m} - f_{\text{test},m}|^2 \right)^{1/2}$$

$$CC: \ d = \frac{f_{\text{training}} f_{\text{test}}}{\|f_{\text{training}}\| \|f_{\text{test}}\|}$$

where $f_{\text{training}, m}$ is the mth ($m=1,\ldots,M$) component of the training feature vector, and similarly for $f_{\text{test}, m}$.

5.1. Experiments with changes in expression

The experimental data we used to test the performances of subbands against expression changes consists of 600 face images of 150 individuals (four images per individual) (Fig. 11). The images were chosen from CMU PIE [15] and FERET [16] databases. Two hundred and seventy-two of these images belong to CMU PIE database, and remaining 328 images belong to FERET database, fafb image set. Facial expression changes in the images occur due to smiling, blinking or talking in CMU PIE database and due to so-called alternative expressions in FERET database. It would be desirable to have a more extensive set of expressions that cover the whole gamut of human emotions. Nonetheless, these experiments allow us to show the proof of concept, that is that subbands and fusion bring in improvements in the recognition performance. All the face images are aligned with respect to the manually detected eye coordinates, scaled to 128×128 pixels resolution, and histogram equalized. For each individual in the set, two of their images are used for training, and the remaining two images are used for testing purposes. The images that contain neutral facial expression are put in the training set. For recognition, 120-dimensional feature vectors, conserving 91.55% of the energy, are used (Fig. 11).

In Table 1, the correct recognition rates from selected 'successful' subbands are given. Note that under expression change, only the scaling components A_1, A_2, A_3 are selected, and in fact, none of the detail bands qualifies in the recognition competition. For faces subject to expression changes only, we found out that performing PCA or ICA1 on scaling components slightly increases the correct recognition rate. On the other hand, in ICA2, no improvement is observed. With PCA and ICA1 features, the $L1$ norm gives the best results, whereas for ICA2, the best

Table 2
Correct recognition percentages using fusion techniques for faces having changes in expression

	Best performing subband	Data fusion	Feature fusion	Decision fusion— sum rule	Decision fusion— product rule	Decision fusion— max. rule
PCA	94.00	94.00	93.67	94.00	93.67	94.67
ICA1	94.33	94.33	94.00	94.67	94.67	94.33
ICA2	96.00	95.67	96.33	96.33	96.33	96.67

Fig. 12. Sample face images containing variations in illumination—first row from CMU PIE, second row from Yale.

results are obtained with normalized correlation measure. ICA2 proves overall superior to ICA1 and PCA in all subbands, its best being realized on the A_2 component.

We next tried to fuse the information contained in the selected subbands, A_1, A_2 and A_3, at the data level, the feature level and the decision level as documented in Table 2. For each feature type we took into consideration the distance metric with which they performed best. Therefore $L1$ norm is used for PCA and ICA1, and CC is used for ICA2. In data and decision fusions 120-dimensional feature vectors are used. In feature fusion, 360-dimensional feature vectors are constructed by concatenating the individual feature vectors of the subbands. The best result is obtained with ICA2 by using the decision fusion, based on the maximum rule principle. Since the individual subbands already have relatively high recognition rates, only a small improvement in the performance is achieved. If we take the performance of the PCA on the original image as a reference, the correct recognition rates increases from 92.33 to 96.67, thus an overall 4.33% improvement is achieved. On the other hand, if we take as a reference the performance of ICA2 on the original sized image, the improvement is a meagre 0.67%. Since the performance of the original 128×128 image is on a par with those of the lower-resolution (64×64, 32×32, 16×16) versions,

a computational advantage would accrue by selecting, let us say, the much smaller 32×32 images (A_2 component).

5.2. Variations in illumination

The experimental data we used to test the performance of subbands against illumination variations consists of 332 face images of 83 individuals (four images per individual) (Fig. 12). The images were chosen from CMU PIE [15] and Yale databases. Two hundred and seventy-two of these images belong to CMU PIE database and remaining 60 images belong to Yale database. Illumination variations in the images occur due to the intensity and direction of the light. All the face images are aligned with respect to the manually detected eye coordinates, scaled to 128×128 pixels resolution and histogram equalized. For each individual in the set, two of their images that contain frontal illumination with different amounts of light are used for training, and the remaining two images that contain illumination from sides are used for testing purposes. Eighty-dimensional feature vectors, which conserve 92.74% of the energy, are used (Fig. 12).

The correct classification rates of subbands selected on the basis of their success are given in Table 3. It is interesting to observe that the horizontal detail subbands (H_2, H_3, HH_3) attain higher correct classification rates as compared to the scaling components. In this respect, PCA and ICA1 features extracted from the three horizontal detail images lead to better results than the ICA2 features. The normalized correlation measure, CC, has proved to be superior in all of the three feature extraction methods. In contrast, ICA2 features performed better when extracted from the scaling components (A_1, A_2, A_3) as well as the original face image, though the recognition rate remained overall inferior to those attained with the horizontal components H_2, H_3, HH_3. As can be observed from Table 3, a significant performance improvement, of the order of 40%, is achieved by using horizontal details of wavelet subbands. If we again take as a reference the performance of the PCA on the original image (54.82%

Table 3

Correct recognition rates of sample subband images in the presence of illumination variations (first five rows correspond to the successful subband images)

	PCA-80			ICA1-80			ICA2-80		
	$L1$	$L2$	CC	$L1$	$L2$	CC	$L1$	$L2$	CC
Original	54.82	52.41	51.20	51.81	52.41	51.20	57.83	56.02	66.87
A_1	56.02	53.01	51.20	51.20	53.01	51.20	59.64	56.63	67.47
A_2	57.83	53.01	51.20	51.81	53.01	51.20	56.02	60.24	63.86
H_2	34.94	33.13	71.69	33.73	33.13	71.69	34.94	30.12	71.08
A_3	57.23	51.81	50.60	51.81	51.81	50.60	57.83	60.24	60.24
H_3	65.06	64.46	72.29	67.47	64.46	72.29	58.43	59.64	62.65
HH_3	45.18	45.78	68.67	45.78	44.58	69.28	40.36	42.17	60.24
V_2	21.08	21.08	37.95	23.49	21.08	37.95	18.67	18.67	48.80
D_2	6.02	8.43	36.75	8.43	8.43	36.75	6.63	6.02	39.16
V_3	43.37	43.98	45.78	43.37	43.98	45.78	42.17	44.58	50
D_3	33.73	34.94	56.02	36.14	34.94	56.63	34.34	32.53	49.40

Table 4
Correct recognition rates of fusion techniques against variations in illumination

	Best performing subband	Data fusion	Feature fusion	Decision fusion— *sum rule*	Decision fusion— *product rule*	Decision fusion— *max. rule*
PCA	72.29	75.90	77.11	77.11	77.11	75.30
ICA1	72.29	75.90	75.90	77.11	77.11	75.90
ICA2	71.08	77.11	72.89	75.90	75.30	77.71

score), the correct recognition rate increases to 72.29, by using only a 16×16 third-level horizontal subband component. It is quite interesting that the vertical and diagonal details bring in no improvement at any level of fusion. In Table 3, the last four rows are reserved for the vertical and diagonal subband components on two successive scales, where one can observe the poor performance with these components.

Next, we carried out fusion experiments, with fusions realized at the data level, feature level and decision level. In all fusion experiments, we used the correlation coefficient, CC, since it was overall the best performing distance metric. The subbands involved in the fusion were the three components, H_2, H_3 and HH_3 using PCA and ICA1 features, and the six components A_1, A_2, H_2, A_3, H_3 and HH_3 using ICA2 features. These subbands were selected on the basis of their performance in single-band experiments. Note that since the energy levels of the scaling components and of the horizontal detail components are very different, for the data fusion experiments, they must be rendered commensurate. In other words, to prevent the low-frequency subbands containing higher energy from dominating the horizontal details, energy normalization is applied by scaling each component by its respective standard deviation.

In the data fusion, the concatenated subband coefficients form longer vectors. For example, when the A_1, A_2, H_2, A_3, H_3 and HH_3 data are fused, the resulting vector becomes $4096 + 1024 + 1024 + 256 + 256 + 64 = 6656$ dimensional. These longer concatenated vectors are, however, still reduced to dimension 80, as shown in Fig. 8, after PCA projection (recall that ICA scheme has a PCA preprocessing stage.) In the feature fusion case, on the other hand, 240-dimensional feature vectors are formed in PCA and ICA1, which are constructed by concatenating the feature vectors of H_2, H_3 and HH_3 subbands. In ICA2, 480-dimensional feature vectors result from concatenation of the features of the A_1, A_2, H_2, A_3, H_3 and HH_3. Finally, in the case of decision fusion, the confidence scores from the different channels (three channels in the case of PCA and ICA1 and six channels in the case of ICA2) are combined to a final decision by using sum rule, product rule and maximum rule [13].

Table 4 shows the improvements in correct recognition rate achievable as a result of fusion schemes. All features (PCA, ICA1, ICA2) seem to benefit from fusion, whether it is data, feature or decision fusion. The first column of Table 4 gives the best attainable 'pre-fusion' score for

a comparison. The highest classification performance increase is obtained with ICA2 using decision fusion based on the maximum rule principle. The recognition performance amounts to 77.71 in a database where faces are subject to lighting variations. The contribution of fusion techniques in this case is a 5.42% improvement vis-à-vis the single best performing subband and feature combination (PCA, ICA1, H_3).

6. Conclusions

In this study, we searched for the frequency subbands that qualify as being insensitive to expression differences and illumination variations on faces. Briefly, it was observed that the frequency subbands containing coarse approximation of the images are successful against expression differences, whereas the subbands containing horizontal details are successful against illumination variations.

Since the recognition performance is not in the first place very adversely affected by changes in facial expression, the performance improvement brought about by the multi-resolution analysis and/or fusion remains unimpressive. One interesting observation made is that the performance with the original image and scaling components at various levels, namely, the A_1, A_2 and A_3 components, respectively, at the 128×128, 64×64, 32×32 and 16×16 resolutions remain almost the same, hence low-resolution versions should be preferred for computational simplicity.

The search for alternative wavelet channels for faces in the presence of illumination variations proves much more effective. The horizontal wavelet components, obtained after either one or two stage of low-pass filtering, were found to be the channels yielding the highest performance scores. One can conjecture that horizontal wavelet removes any horizontal illumination pattern, e.g. one cheek darker, the other lighter. Alternatively, one can conjecture that the horizontal details emphasize the left–right asymmetry on faces, and the facial asymmetry was shown to be a very good feature in face recognition [17]. Fusing a number of channels further improves the highest performance achieved by a single channel, by another 5.42%. The Tables 5–9 summarize the recognition performance improvements instrumented by the choice of metrics, choice of wavelet components and choice of fusion scheme. The comparison reference is the original face image with PCA features and

Table 5
Improvement due to choice of features and metrics vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
3.67	12.05

Both improvements are attained with the max-rule fusion of the ICA2 features of the A_1, A_2, A_3, bands in lieu of the original images (compare first rows of Tables 1 and 3).

Table 6
Improvement due to wavelet decomposition vis-à-vis the full-scale performance

	Facial expression change (%)	Illumination change (%)
PCA	1.66	17.47
ICA1	1.00	19.88
ICA2	–	4.21

Compare corresponding columns in Tables 1 and 3: for example, PCA on original image performance is 54.82, while PCA on the H_3 band performance is 72.79 resulting in 17.47% improvement.

Table 7
Improvement due to fusion vis-à-vis the best subband component

	Facial expression change (%)	Illumination change (%)
PCA	0.67	4.82
ICA1	0.33	4.82
ICA2	0.67	6.63

Table 8
Overall improvement vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
4.33	22.89

Table 9
Decrease in the error rate vis-à-vis the PCA-L1 performance

Facial expression change (%)	Illumination change (%)
56.45	50.66

$L1$ metric (referred to, simply as PCA-L1), as this is the most commonly occurring technique in the literature.

In conclusion, it can be said that natural expression changes, i.e. smiling, blinking, talking, do not cause severe performance reduction in face recognition. The attained correct recognition rates are already relatively high; therefore, multiresolution analysis and fusion provide a small improvement. On the other hand, it is observed that recognition of faces, subject to illumination changes is a more sensitive task. Utilizing multiresolution analysis and fusion is quite effective in combating the detrimental effects of illumination variations. For example, it would be intriguing to consider the illumination variations coupled with expression variations. However, not only the effects of illumination variations are more dominant vis-à-vis

the expression variations, but also we have found that the subbands useful in combating illumination variations subsumes already subbands effective for expression variations.

Similar studies can be conducted for other facial factors, of aging, accessories or pose, provided adequate databases become available. In such relatively more complicated tasks, where more intra-class variability and/or more number of classes can be encountered, we believe the multiresolution scheme can also be beneficial.

References

[1] R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (10) (1993) 1042–1052.

[2] M. Turk, A. Pentland, Eigenfaces for recognition, Journal of Cognitive Science 1991; 71–86.

[3] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman, Eigenfaces vs. fisherfaces: recognition using class specific linear projection, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 711–720.

[4] M.S. Bartlett, H.M. Lades, T.J. Sejnowski, Independent component representations for face recognition, Proceedings of the SPIE Conference on Human Vision and Electronic Imaging III, San Jose, CA, 1998 pp. 528–539.

[5] Y. Moses, Y. Adini, S. Ullman, Face recognition: the problem of compensating for changes in illumination direction, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (7) (1997) 721–732.

[6] G.C. Feng, P.C. Yuen, D.Q. Dai, Human face recognition using PCA on wavelet subband, SPIE Journal of Electronic Imaging 9 (2) (2000) 226–233.

[7] J. Tang, R. Nakatsu, S. Kawato, J. Ohya, A wavelet-transform based asker identification system for smart multi-point tele-conferences, Journal of the Visualization Society of Japan 20 (1) (2000) 303–306.

[8] C. Garcia, G. Zikos, G. Tziritas, Wavelet packet analysis for face recognition, Image and Vision Computing 18 (4) (2000) 289–297.

[9] B. Li, Y. Liu, When eigenfaces are combined with wavelets, International Journal of Knowledge-Based Systems 15 (5/6) (2002) 343–347.

[10] J.T. Chien, C.C. Wu, Discriminant waveletfaces and nearest feature classifiers for face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (12) (2002) 1644–1649.

[11] P.J. Phillips, Matching pursuit filters applied to face identification, IEEE Transactions on Image Processing 7 (8) (1998) 1150–1164.

[12] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural Networks 13 (2000) 411–430.

[13] J. Kittler, M. Hatef, R.P. Duin, J.G. Matas, On combining classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (3) (1998) 226–239.

[14] I. Daubechies, Ten lectures on wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM Press, Philadelphia, 1992.

[15] T. Sim, S. Baker, M. Bsat, The CMU pose, illumination, and expression (PIE) database, Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, May 2002.

[16] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (10) (2000) 1090–1104.

[17] Y. Liu, J. Palmer, A quantified study of facial asymmetry in 3D faces, Proceedings of the 2003 IEEE International Workshop on Analysis and Modeling of Faces and Gestures, in conjunction with the 2003 International Conference of Computer Vision (ICCV '03), October 2003.