Combining multiple face recognition systems using Fisher's linear discriminant

Thomas Heseltine[*], Nick Pears, Jim Austin
Department of Computer Science, The University of York, United Kingdom

ABSTRACT

The application of image processing as a pre-processing step to methods of face recognition can significantly improve recognition accuracy. However, different image processing techniques provide different advantages, enhancing specific features or normalising certain capture conditions. We introduce a new method of isolating these useful qualities from a range of image subspaces using Fisher's linear discriminant and combining them to create a more effective image subspace, utilising the advantages offered by numerous image processing techniques and ultimately reducing recognition error. Systems are evaluated by performing up to 258,840 verification operations on a large test set of images presenting typical difficulties when performing recognition. Results are presented in the form of error rate curves, showing false acceptance rate (FAR) vs. false rejection rate (FRR), generated by varying a decision threshold applied to the euclidean distance metric performed in combined face space.

Keywords: Face Recognition, Eigenface, Fisherface, Linear Discriminant

1. INTRODUCTION

It has been shown that the application of image processing techniques as a pre-processing step to methods of face recognition, such as the eigenface and fisherface methods, can significantly improve recognition accuracy[1, 2]. Such image processing techniques work on several principles, such as reducing noise, enhancing features or normalising environmental conditions. Therefore, each technique provides unique advantages, specifically suited to different conditions. For example, colour normalisation techniques[3, 4] may aid recognition by making such features as skin-tone and hair colour consistent despite the effect of lighting conditions. Another system may incorporate edge detection filters, focusing purely on facial structure, while a third may blur an image, reducing inaccuracies introduced by the feature alignment stage. Unfortunately, often incorporated with these beneficial characteristics are surplus side effects, which can actually degrade system performance: normalising colour and removing the effect of lighting conditions will reduce the geometric information encapsulated within the facial surface shading; edge-detection or gradient based filters preserve structural cues, but remove skin-tone information.

In this paper we analyse and evaluate a range of face recognition systems, each utilising a different image processing technique, in an attempt to identify and isolate the advantages offered by each system. Focusing on appearance based methods of face recognition we propose a means of selecting and extracting components from the image subspace produced by each system, such that they may be combined into a unified face space. We apply this method of combination to the eigenface approach[5, 6] and fisherface approach[7]. The benefit of using multiple eigenspaces has previously been examined by Pentland et al[8], in which specialist eigenspaces were constructed for various facial orientations and local facial regions, from which cumulative match scores were able to reduce error rates. Our approach differs in that we extract and combine individual dimensions, creating a single unified face space.

In section 2 we begin with a brief explanation of the eigenface and fisherface methods. We describe the database of face images used for testing and training in section 3, which are then analysed in section 4, discussing the image processing techniques evaluated in previous work[1, 2], the rationale for combining multiple systems and the criteria used to identify the most discriminatory components of each system. The algorithm used for combining these components is then described in section 5. After applying this combination process to the eigenface and fisherface methods, we compare the effectiveness of the resultant face space combinations with the optimum systems of previous work[2]. The evaluation

---

[*] tom.heseltine@cs.york.ac.uk; www.cs.york.ac.uk/~tomh

procedure is described in section 6, by which we perform verification operations on a large test set of facial images that present typical difficulties when performing recognition, such as variations in illumination direction and facial expression. We present the results in the form of error rate curves in section 7, generated by varying a decision threshold in the verification operations.

2. THE EIGENFACE AND FISHERFACE METHODS

In this section we give a brief explanation of the eigenface and fisherface methods of face recognition, while referring the reader to Turk and Pentland[5, 6] and Belhumeur et al[7] for more detailed explanations. Both approaches work on the same principle of analysing the image space of a given training set of face images Γ_{ni}, of c different people, attempting to reduce image space dimensionality down to the most discriminating components. This is accomplished by computing eigenvectors of one or more scatter matrices (equation 1) using standard linear methods, ultimately producing subspace projection matrices, U_{ef} and U_{ff}, of the top $c\text{-}1$ components with the highest eigenvalues for the eigenface and fisherface systems respectively.

$$\textit{Training Set} = \{X_1, X_2, \ldots X_c\}$$

$$\textit{where } X_n = \{\Gamma_{n1}, \Gamma_{n2}, \Gamma_{n3}, \ldots\}$$

$$\Psi = \frac{1}{\sum_{m=1}^{c}|X_m|}\sum_{n=1}^{c}\sum_{i=1}^{|X_n|}\Gamma_{ni}$$

$$\Psi_n = \frac{1}{|X_n|}\sum_{i=1}^{|X_n|}\Gamma_{ni}$$

$$S_C = \tfrac{1}{c}\sum_{n=1}^{c}(\Gamma_{n1} - \Psi)(\Gamma_{n1} - \Psi)^T$$

$$S_T = \sum_{n=1}^{c}\sum_{i=1}^{|X_n|}(\Gamma_{ni} - \Psi)(\Gamma_{ni} - \Psi)^T$$

$$S_B = \sum_{n=1}^{c}|X_n|(\Psi_n - \Psi)(\Psi_n - \Psi)^T$$

$$S_W = \sum_{n=1}^{c}\sum_{i=1}^{|X_n|}(\Gamma_{ni} - \Psi_n)(\Gamma_{ni} - \Psi_n)^T \quad . \qquad (1)$$

The two approaches differ in the scatter matrices from which the eigenvectors are calculated. The eigenface method applies principal component analysis (PCA) using the covariance matrix S_C, constructed from single examples of each person in the training set. Whereas the fisherface method is able to take advantage of multiple examples of each person, minimising within-class scatter (S_W), yet maximising between-class scatter (S_B). In addition, the fisherface approach applies PCA to the total scatter matrix S_T, producing a preliminary projection matrix U_{pca}, used to reduce the dimensionality of the scatter matrices S_B and S_W, ensuring they are non-singular, before computing the eigenvectors (U_{fld}) of the reduced scatter matrix ratio.

$$U_{ef} = \arg\max_U\left(\left|U^T S_c U\right|\right) \qquad U_{pca} = \arg\max_U\left(\left|U^T S_T U\right|\right)$$

$$U_{fld} = \arg\max_U\left(\frac{\left|U^T U_{pca}^T S_B U_{pca} U\right|}{\left|U^T U_{pca}^T S_W U_{pca} U\right|}\right) \qquad U_{ff} = U_{fld}\,U_{pca} \quad . \qquad (2)$$

Finally, the matrix U_{ff} is calculated as shown in equation 2, such that it will project a face image into a reduced image space of $c\text{-}1$ dimensions. Once projection matrices have been constructed, they are used to reduce the dimensionality of a given face image Γ (5330 element vector) down to a $c\text{-}1$ element vector ω, termed a face-key, as shown in equation 3.

$$\omega_k = u_k^T(\Gamma - \Psi) \qquad \textit{for } k = 1 \ldots c\text{-}1 \quad . \qquad (3)$$

These face-key vectors contain the coefficients of the respective projection matrix eigenvectors, referred to as eigenfaces (Fig. 1) and fisherfaces (Fig. 2). Face-keys are compared using the Euclidean distance measure, followed by a threshold to determine an acceptance/rejection decision as shown in equation 4.

$$\varepsilon = \left\|\omega_q - \omega_g\right\| \qquad (\varepsilon \le \textit{threshold} \Rightarrow \textit{accept}) \wedge (\varepsilon > \textit{threshold} \Rightarrow \textit{reject}) \quad . \qquad (4)$$

Fig. 1. The average face (*left*) and first five eigenfaces (*right*) computed with no image pre-processing.

Fig. 2. The first five fisherfaces, defining a face space with no image pre-processing.

3. THE TEST DATABASE

We conduct experiments using a database of 960 bitmap images of 120 individuals (60 male, 60 female) of various race and age, extracted from the AR Face Database provided by Martinez and Benavente[9]. From this database we take a training set of 240 images (60 people under a range of lighting conditions and facial expressions), used to compute the scatter matrices described in section 2 and ultimately produce the face space projection matrices. The remaining 720 images (60 people, 12 images each) are then separated into two disjoint sets of equal size (test set A and test set B). We use test set A to analyse the face-key variance throughout face space, calculate discriminant weightings (see section 4) and compute the optimum face space combinations. This leaves test set B as an unseen set of data to evaluate the final combined system. The six examples shown in Table 1 were repeated on two days, making up the 12 images of each subject in the test sets. All images are pre-aligned with the eye centres 25 pixels apart, before being cropped to a width and height of 65 and 82 pixels respectively.

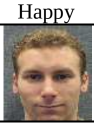
Lighting	Natural	From left	From right	Left & right	Natural	Natural
Expression	Neutral	Neutral	Neutral	Neutral	Happy	Angry
Example						

Table 1. Image capture conditions included in the database training and test sets.

4. ANALYSIS OF FACE RECOGNITION SYSTEMS

In this section we analyse image subspaces produced when various image pre-processing techniques are applied to both the eigenface and fisherface methods. We begin by providing the results obtained in previous work[2], shown in Fig. 3, showing the range of error rates produced when using various image processing techniques. Continuing this line of research we persist with these same image processing techniques, referring the reader to Heseltine et al[1, 2] for implementation details, while in this paper we focus on the effect and methodologies of combining multiple systems, rather than the image processing techniques themselves.

Fig. 3 clearly shows that the choice of image processing technique has a significant effect on the performance of both the eigenface and fisherface approaches, with detail enhancement filters providing the lowest equal error rates (EER, the error when FAR equals FRR) when used in conjunction with the fisherface approach. However, we find it surprising that some image processing techniques give such poor performance, especially when designed specifically to compensate for conditions known to be a source of error in face recognition systems[10]. For example, we see that intensity normalisation increases error rates for fisherface-based systems, despite being the optimum image processing technique for eigenface-based recognition. Hence, it is apparent that this processing technique is able to preserve discriminatory information, while normalising lighting effects, yet is unsuitable for fisherface-based recognition. We now carry out further investigation into the discriminating ability of each face recognition system by applying Fisher's Linear Discriminant (FLD), as used by Gordon[11] to analyse 3D face features, to individual components (single dimensions) of each face space. Focusing on a single face space dimension we calculate the discriminant d, describing the discriminating power of that dimension, between c people in test set A.

$$d = \frac{\sum_{i=1}^{c}(m_i - m)^2}{\sum_{i=1}^{c}\frac{1}{|\Phi_i|}\sum_{x\in\Phi_i}(x - m_i)^2} \ . \tag{5}$$

Where m is the mean value of that dimension in the face-keys of test set A, m_i the within-class mean of class i and Φ_i the set of vector elements taken from the face-keys of class i.

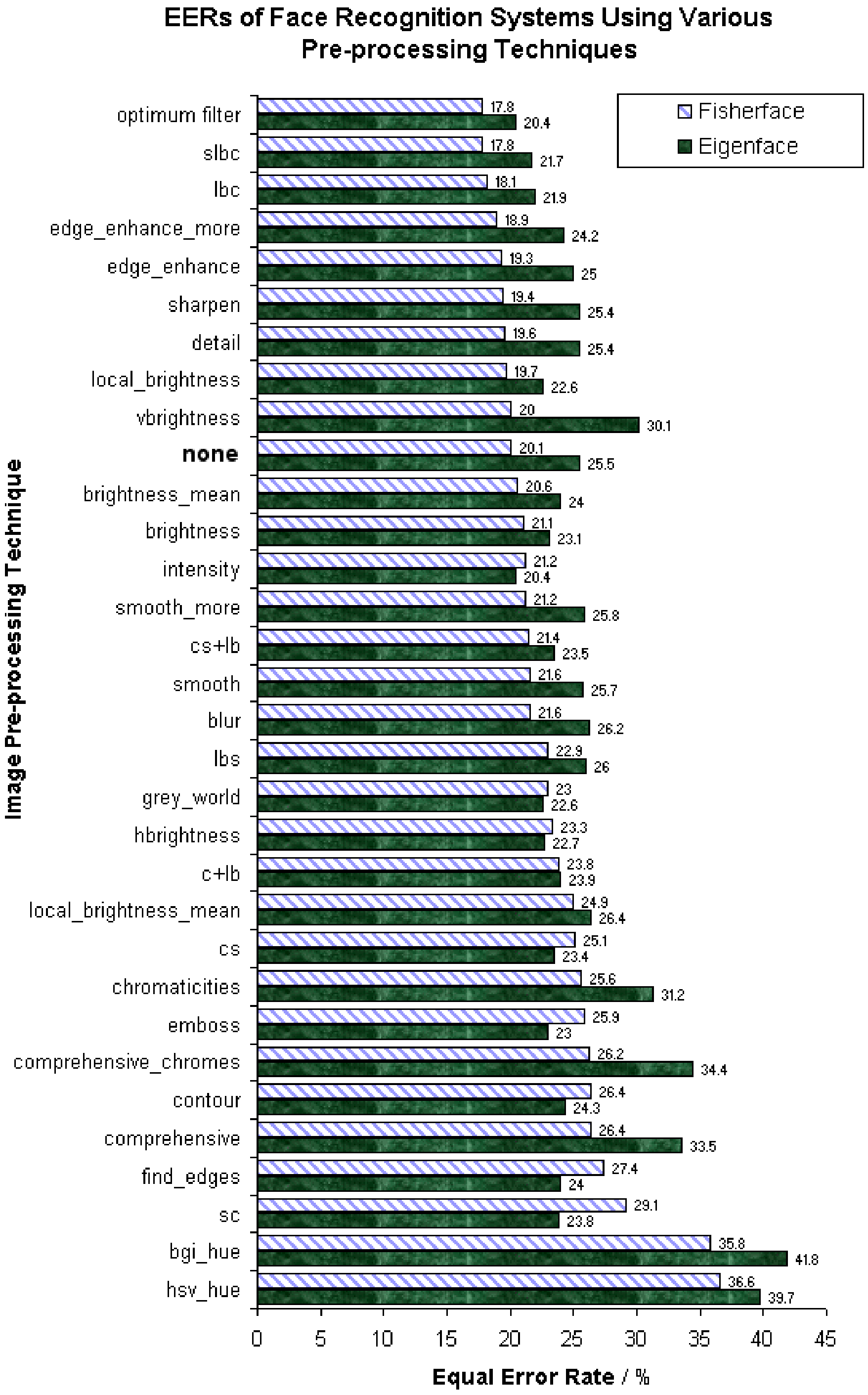


Fig. 3. EERs of eigenface and fisherface systems using a range of image processing techniques as calculated by Heseltine et al[2].

Applying equation 5 to each dimension of the eigenface face space (using no image pre-processing), provides a set of discriminant values as shown in Fig. 4. Looking at the range of discriminant values, we note that the higher discriminants appear at the lower end of the face space. This is exactly as we would expect, showing that the order of principal components, in terms of eigenvalues, is related to that dimensions discriminating ability.

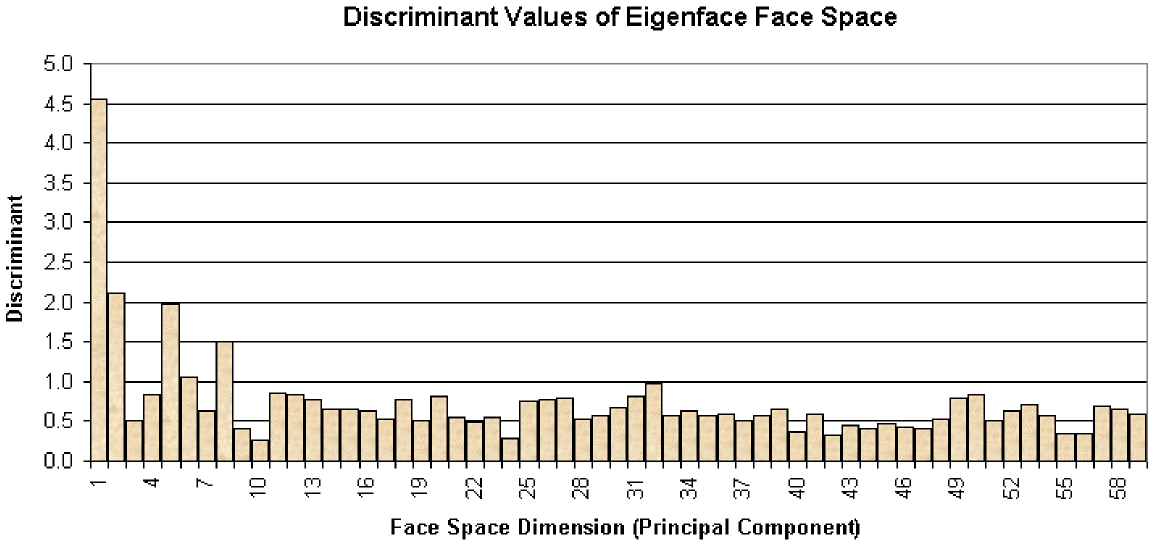


Fig. 4. Discriminant values of the eigenface face space dimensions using no image pre-processing.

However, it can also be seen that certain dimensions produce very low discriminant values. In Fig. 4 we see that the third principal component in particular has a very low discriminating ability, despite its relatively high eigenvalue. This highlights some problems in the eigenface training method, in that the third dimension obviously represents something of high variance in the training set that has little or no use in discriminating between different people. In other words, it is a feature of the environmental capture conditions. Applying equation 5 to each dimension in the assortment of fisherface systems, we see similar results to those of the eigenface systems, with a wide range of discriminant values across the different image processing techniques. Fig. 5 shows the top ten dimensions with the highest discriminant value from each fisherface system.

It is clear that although some image processing techniques do not perform well in the face recognition tests, producing high EERs as seen in Fig. 3, some of their face-key components do contain highly discriminatory information. We hypothesise that the reason for these highly discriminating anomalies, in an otherwise ineffective subspace, is that a certain image processing technique may be particularly suited to a single discriminating factor, such as skin tone or hair colour, but is not effective when used as a more general classifier. Therefore, if we were able to isolate these few useful qualities from the more specialised image subspaces, they could be used to make a positive contribution to a generally more effective face space, reducing error rates further. For example, *grey_world* pre-processing results in a particularly high EER (23.0%), yet we see that two dimensions of this face space have discriminant values significantly greater than any dimension from the optimum fisherface system (using *slbc* image pre-processing). Therefore, it is not unreasonable to assume (given that *grey_world* normalises colour and *slbc* enhances edges) that if these two dimensions were extracted and combined with the existing *slbc* face space, a further reduction in error may occur, due to the additional discriminatory information being introduced.

In order to combine multiple dimensions from a range of face spaces, we require some criterion to decide which dimensions to combine. It is not enough to rely purely on the discriminant value itself, as this only gives us an indication of the discriminating ability of that dimension alone, without any indication of whether the inclusion of this dimension would benefit the existing set of dimensions. If an existing face space already provides a certain amount of discriminatory ability, it would be of little benefit (or could even be detrimental) if we were to introduce an additional dimension describing a feature already present within the existing set, unless it was of a discriminant significantly high as to provide a valued contribution. Ideally we would use the EER as this criterion, such that a new dimension would be incorporated into any existing system if it resulted in a reduced EER. However, such an approach is problematic in that the time taken to process a complete verification evaluation for all dimension combinations would be unfeasible, unless we used a particularly small test set, in which case we run the risk of over-training: only selecting those dimensions particularly suited to the small amount of test data.

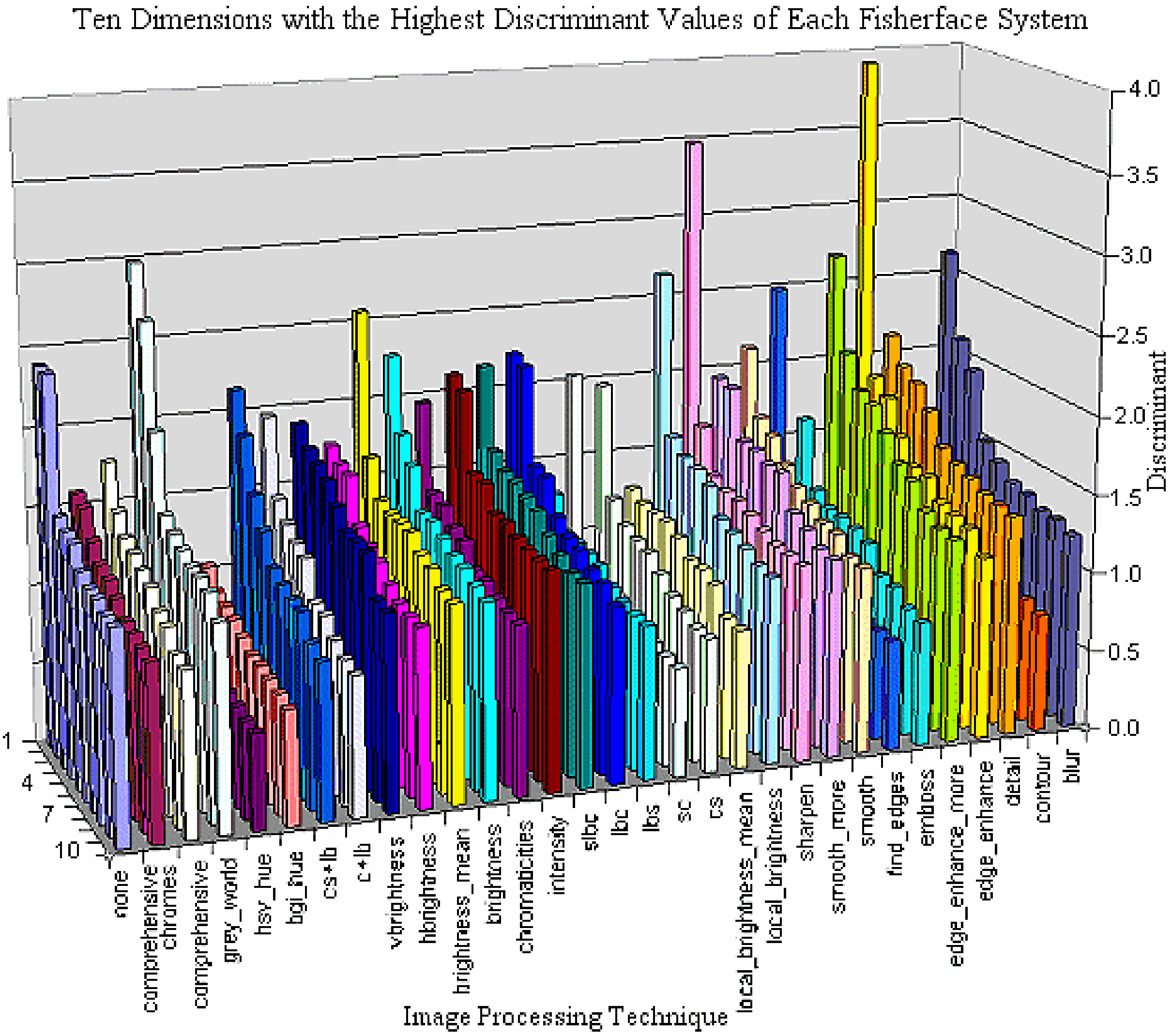


Fig. 5. Ten greatest discriminant values of dimensions from fisherface face spaces using a range of image pre-processing techniques.

What we require is some method of providing an indication of how effective a given combination of face space dimensions is likely to be, without the need of processing a complete evaluation of all verification operations. An obvious solution, already used to analyse individual face space dimensions is that of FLD, which with a small amount of adaptation can be applied to whole face-key vectors, rather than individual vector elements, providing a global discriminant value *d* for the entire face space,

$$d = \frac{\sum_{i=1}^{c} \left\| \overline{\omega}_i - \overline{\omega} \right\|}{\sum_{i=1}^{c} \frac{1}{|\Phi_i|} \sum_{\omega \in \Phi_i} \left\| \omega - \overline{\omega}_i \right\|} \; . \tag{6}$$

where ω is the face-key of some single or combined face space, to which we apply the recognition distance metric (equation 4) to the average $\overline{\omega}$ and class average $\overline{\omega}_i$. Applying equation 6 to each fisherface system shown in Fig. 3 and comparing the result with their respective EERs, it becomes evident there is some correlation between this global discriminant value and the effectiveness of a face recognition system, as seen in Fig. 6.

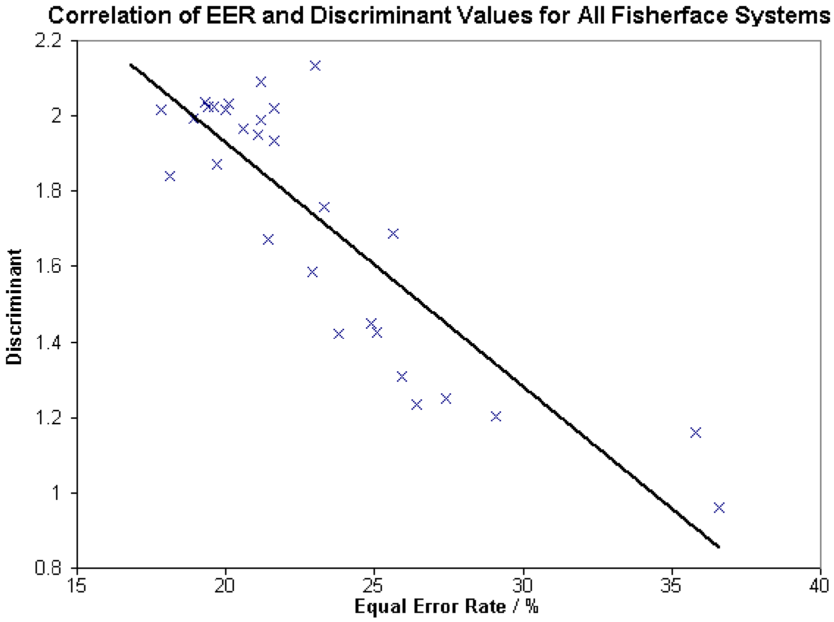


Fig. 6. Scatter graph showing the correlation between the global discriminant value and EER of fisherface systems.

5. COMBINING SYSTEMS

In this section we describe how the analysis methods discussed in section 4 are used to combine multiple face recognition systems. Firstly, we need to address the problem of prioritising face space dimensions. Because the average magnitude and deviation of face-key vectors from a range of systems are likely to differ by some orders of magnitude, certain dimensions will have a greater influence than others, even if the discriminating abilities are evenly matched. To compensate for this effect, we normalise moments by dividing each face-key element by its within-class standard deviation. However, in normalising these dimensions we have also removed any prioritisation, such that all face space components are considered equal. Although not a problem when applied to a single face space, when combining multiple dimensions we would ideally wish to give greater precedence to the more reliable components. Otherwise the situation is likely to arise when a large number of less discriminating (but still useful) dimensions begin to outweigh the fewer more discriminating ones, diminishing their influence on the verification operation and hence increasing error rates. In section 4 we showed how FLD could be used to measure the discriminating ability of a single dimension from any given face space. We now apply this discriminant value d (equation 5) as a weighting for each face space dimension, prioritising those dimensions with the highest discriminating ability.

With this weighting scheme applied to all face-keys produced by each system, we can begin to combine dimensions into a single unified face space. The criterion required for a new dimension to be introduced to an existing face space is a resultant increase in the global discriminant, calculated using equation 6. However, as can be seen from Fig. 6 this method can only provide a rough indication of system effectiveness and if we were to build up the combination from a single dimension, we may achieve a greater discriminant but not necessarily the lowest EER. Therefore, in order to provide the combination with the best head start, we initialise the dimension set with the optimum face space achieved so far (*intensity* and *slbc* for eigenface and fisherface systems respectively). Beginning with this small preliminary set of dimensions (the face space of the optimum eigenface or fisherface system), we then iteratively test each additional dimension from other face spaces for combination with the existing dimension set as shown in Table 2.

The result is a new face space consisting of the dimensions taken from the original optimum system, plus a selection of additional dimensions from other systems. Each new dimension will have increased the global discriminant, such that the final combination has a significantly higher discriminant value and will therefore also have reduced the EER when evaluated on test set B.

```
Combined face space = face space dimensions of current optimum system
Calculate global FLD of combined face space
For each face space system:
    For each dimension of face space system:
        Concatenate new dimension onto combined face space
        Calculate global FLD of combined face space
        If global FLD has not increased:
            Remove new dimension from combined face space
Save combined face space ready for evaluation
```

Table 2 Combination algorithm used to select and combine dimensions from multiple face spaces.

6. THE TEST PROCEDURE

The effectiveness of the face recognition systems is evaluated by means of error rate curves (FRR vs. FAR) generated by performing a large number of verification operations on the database test sets. The images in the test set are verified against every other image, producing a distance value using equation 4. No image is compared with itself and each pair is compared only once (the relationship is symmetric). This provides 64,620 verification operations when performed on all images in test set B or 258,840 operations if both test sets A and B are combined. After calculating the distance values for each comparison, a threshold is applied in order to derive the rejection/acceptance decision for each image pair. FAR is calculated as the percentage of acceptance decisions when images of different people are compared and FRR is the percentage of rejection decisions when images of the same person are compared. By varying the threshold we produce a set of FRR FAR plots, forming the error rate curve, as shown in Fig.8.

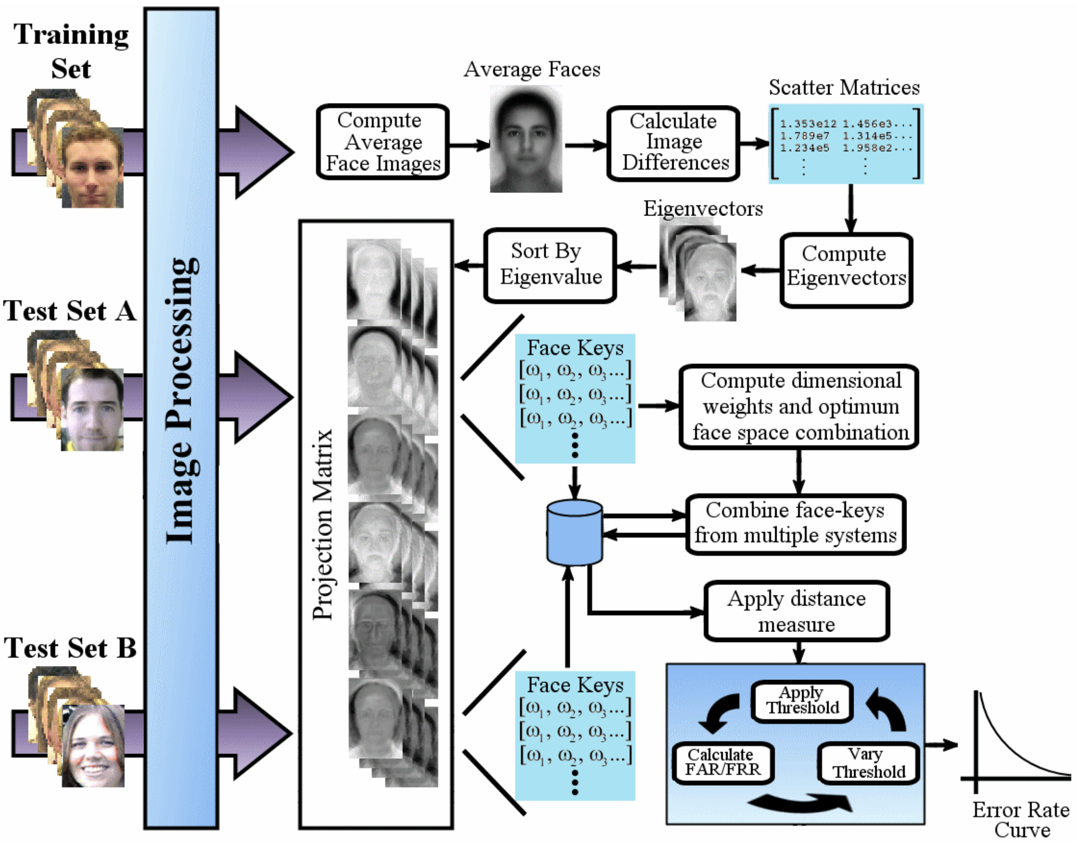


Fig. 7. Flow chart of face recognition evaluation procedure.

7. RESULTS

In this section we present results obtained from evaluating the optimum single systems and combined face recognition systems formed using the eigenface and fisherface methods. The results are presented in the form of error rate curves (FAR vs. FRR) generated using the procedure described in section 6, taking the EER as a single comparative value.

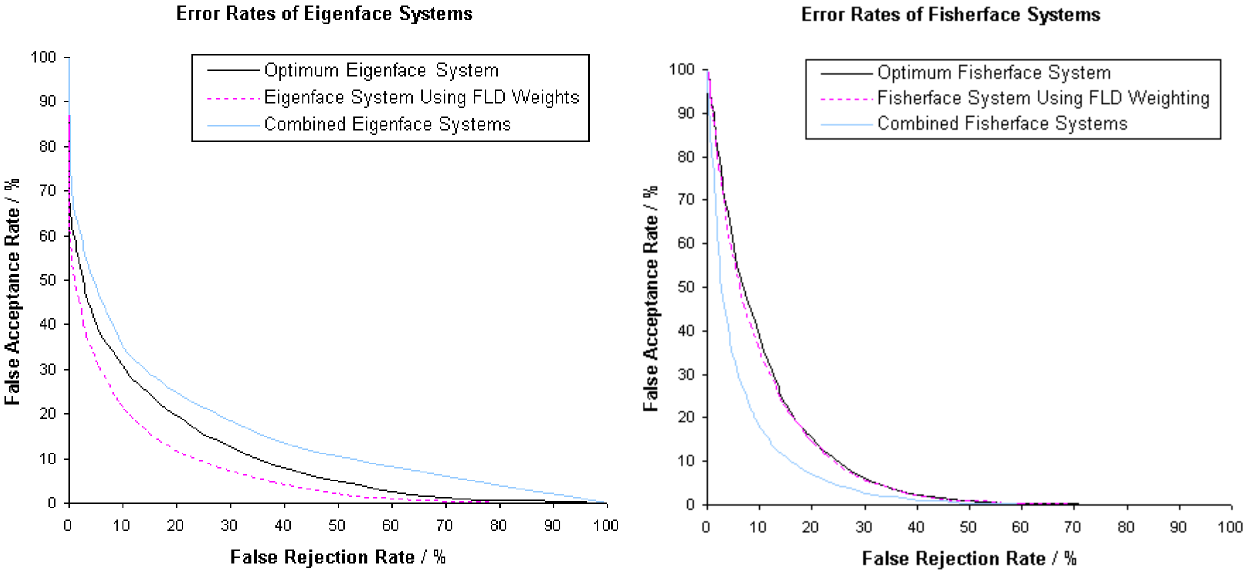
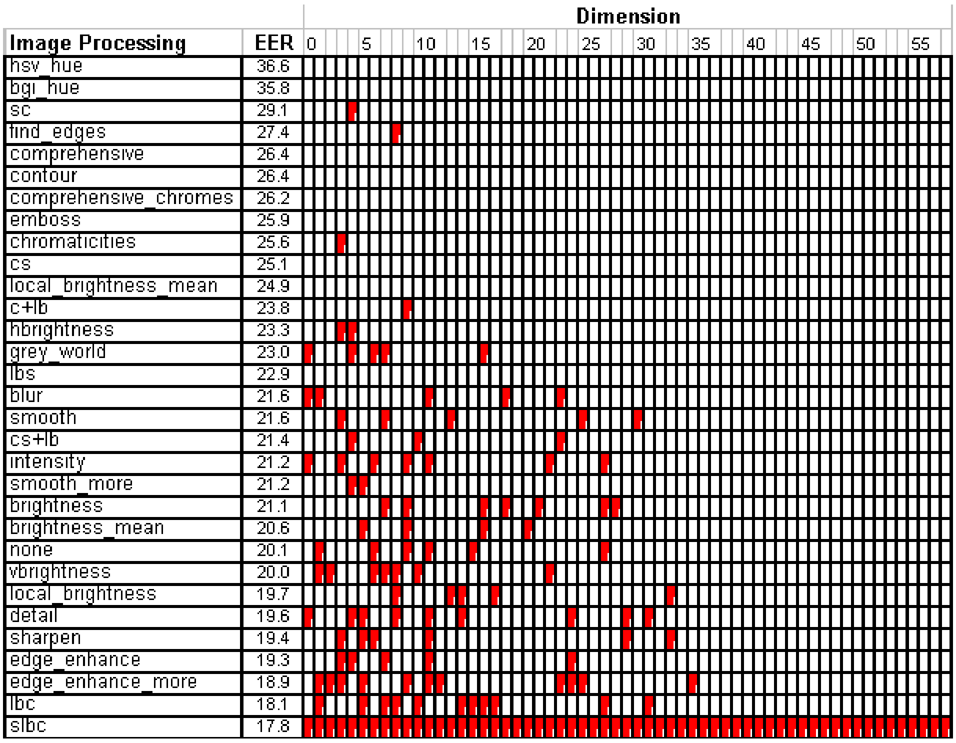


Fig. 8. Error rate curves of single optimum, weighted and combined eigenface (*left*) and fisherface (*right*) systems, produced when applied to test set B.

Fig. 8 (left) shows the error rates obtained using the eigenface approach, when applied to test set B (the previously unseen test set). We see that applying the optimum eigenface system (incorporating the best image pre-processing technique as described by Heseltine et al[2]) to test set B, produces an EER of 19.7%. A significant improvement is witnessed when the discriminant values (calculated using test set A) are applied as a weighting scheme (described in section 5), prioritising the most discriminating principal components, reducing the EER to 15.4%. With this weighting scheme in place for all eigenface systems, we then apply the combination algorithm in Table 2, producing the third error rate curve in Fig. 8. Unfortunately, this combined system substantially increases the error rate, resulting in a final EER of 22.8%.

Fig. 8 (right) shows the results obtained when performing the same evaluation experiments using the fisherface approach. The initial EER, using the optimum image pre-processing technique is 17.9%. Weighting the components according to discriminant values, unlike the eigenface system, has very little effect on system performance and although provides marginal improvement at some points along the error curve, actually results in the same EER of 17.9%. However, combining the weighted dimensions, from all fisherface systems, produces a significant error reduction to 13.0% EER. Displaying the face space dimensions selected for inclusion in this final combined fisherface system in Fig. 9 shows that those systems with lower EERs generally provide more highly discriminating dimensions for inclusion in the final system than systems with higher EERs. It is also evident that dimensions with higher eigenvalues provide the most discriminating information, as expected. However, it is interesting to note that even a few of the least effective systems provide some contribution to the final combined face space.

Having evaluated the initial baseline systems and combined systems on the unseen test set B, demonstrating the improvement gained by combining multiple fisherface dimensions, we now explore how these results vary when the images used to compute the optimum combination are also present in the evaluation test set. This experiment is analogous to training the face recognition system on the database (or gallery set) of known people, which are then compared to newly acquired (unseen) images.



		Dimension																																																								
Image Processing	EER	0					5					10					15					20					25					30					35					40					45					50					55			
hsv_hue	36.6																																																										
bgi_hue	35.8																																																										
sc	29.1					■																																																						
find_edges	27.4									■																																																	
comprehensive	26.4																																																										
contour	26.4																																																										
comprehensive_chromes	26.2																																																										
emboss	25.9																																																										
chromaticities	25.6				■																																																							
cs	25.1																																																										
local_brightness_mean	24.9																																																										
c+lb	23.8										■																																																
hbrightness	23.3				■	■																																																						
grey_world	23.0	■				■		■	■									■																																										
lbs	22.9																																																										
blur	21.6	■	■										■							■					■																																		
smooth	21.6				■				■						■													■					■																												
cs+lb	21.4					■						■													■																																		
intensity	21.2	■			■			■			■		■											■					■																																
smooth_more	21.2					■	■																																																					
brightness	21.1								■		■							■		■			■						■	■																														
brightness_mean	20.6						■				■							■				■																																						
none	20.1		■					■			■		■				■												■																															
vbrightness	20.0		■	■				■	■	■		■												■																																			
local_brightness	19.7									■					■	■			■																■																									
detail	19.6	■				■	■			■			■			■										■					■		■																											
sharpen	19.4				■		■	■					■																		■				■																									
edge_enhance	19.3				■	■			■				■													■																																		
edge_enhance_more	18.9		■	■	■		■				■		■	■											■	■	■										■																							
lbc	18.1		■				■		■	■		■				■	■	■	■										■				■																											
slbc	17.8	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Fig. 9. Face space dimensions included (■) in the final combined fisherface face space used to produce the results shown in Fig. 8.

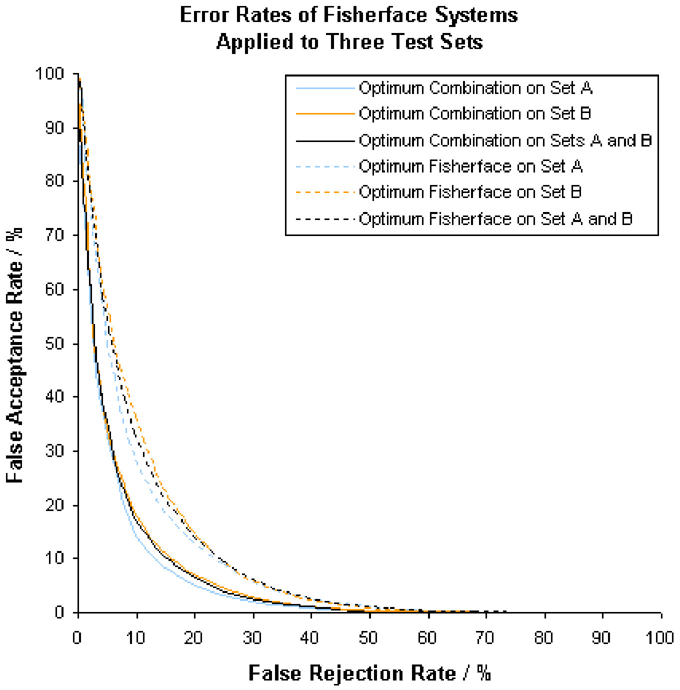


Fig. 10. Error rate curves of optimum single and combined fisherface systems applied to three test sets.

Fig. 10 shows the results obtained when the optimum fisherface system and combined fisherface system are applied to test set A (used to construct the combined system), test set B (the unseen test set) and the full test set (all images from sets A and B). We see that the combined system does produce lower error rates (11.7% EER) when applied to images used to construct the combination, as would be expected. However, we also see that test set A produces better results (16.6% EER) than test set B for the single fisherface system, suggesting that it is actually a slightly easier test set anyway. Performing the evaluation on the larger set, providing 258,840 verification operations, the error rate drops slightly to 12.8% EER, showing that a small improvement is introduced if some test images are available for training, as well as suggesting that the method scales well, considering the large increase in image comparisons. The distinct separation between the error curves of the single optimum fisherface system and those using the combined system further enforces the fact that combining multiple face space dimensions provides a substantial advantage over individual systems that use a single image processing technique.

8. CONCLUSION

We have highlighted the importance of using image processing as a pre-processing step to two well-known methods of face recognition and discussed the possibilities of combining face space dimensions of multiple systems in an attempt to utilise the advantages offered by numerous image processing techniques. Using FLD as an analysis tool we have confirmed the hypothesis that although an image subspace may not perform well when used for recognition, it may harbour highly discriminatory components that could complement other more superior systems and hence we have shown the potential to improve recognition error rates by combining multiple dimensions from a range of face recognition systems.

Using this method of FLD analysis, we have overcome two problems presented when combining face recognition systems. Firstly, using the respective discriminant values to weight face space dimensions, according to their discriminating ability, has allowed us to combine multiple dimensions yet maintain a bias towards those that present the most distinguishing features. We have shown this weighting scheme to be highly beneficial when used with the optimum eigenface system, reducing the EER from 19.7% to 15.4%, but have little influence on the effectiveness of individual fisherface systems. Secondly, applying FLD to entire face-keys and comparing these global discriminant values with the EERs of existing systems has demonstrated how this analysis can be used to provide a rough indication of the effectiveness of a given face recognition system, requiring significantly less processing time than completing a full set of verification operations on the same data. Using this global discriminant as criteria for selecting face space dimensions as potential sources of additional discriminatory information to an existing system has enabled an iterative approach of appending new face space dimensions to existing combinations, increasing the global discriminant value and hence improving the performance of the combined face recognition system.

Testing this method of combination on face space representations produced using the eigenface approach has shown it to be ineffective for this method of face recognition, increasing the EER significantly from 15.4% to 22.8%. However, applying the same combination process to fisherface systems has shown that combining multiple face spaces can improve system performance substantially, reducing the EER from 17.9% down to 13.0%.

This key difference between the eigenface and fisherface approaches is of particular interest and at first thought perhaps quite surprising. In order to understand this phenomenon, we must consider the method of combination used. We have created a criterion in which face space dimensions are incorporated into an existing set if they increase the discriminant of the combined face space. This criterion only takes into account the discriminating ability of the new dimension when compared with the level of discrimination already achieved within the existing combination. It does not allow for the inter-dependency of facial features or the possibility that features represented in the additional dimension may already be present in the existing face space. For example, consider a combined face space, in which its current set of dimensions encapsulates such features as the nose base width, bridge curvature and nose length. Now suppose we identify a new dimension for inclusion in the face space, representing the more general feature of nose shape, which due to its high between-class variance will increases the global discriminant. However, this new dimension represents a feature that is largely dependent on those already represented in the face space. Therefore the discriminatory information available in this new dimension is predominantly redundant, meaning that the only real contribution to the combined face space is the additional noise of within-class variance.

This reasoning begins to uncover the grounds for failing to successfully combine multiple eigenface systems. The eigenface approach creates a face space that maximises image distribution, but uses no examples of within-class variance, therefore doing nothing to reduce noise or environmental features. Any dimension combined with an existing face space not only introduces the primary discriminating feature (which may have been present beforehand anyway) but also incorporates substantial within-class variance. Adini et al[10] have shown that differences due to lighting conditions and facial expression are greater than the differences between images of different people, suggesting that the noise introduced when combining dimensions will be more diverse and cumulative than the discriminating features, which will often reoccur and hence be redundant. The fisherface approach differs in its ability to formulate face space such that within-class variation is minimised, hence reducing environmental influence, allowing multiple dimensions to combine with relatively little increase in noise content. Therefore, even if the dimension contribution is redundant, little or no degradation is introduced.

The criterion used to select dimensions is obviously an important factor in the combination process. In this paper we have developed a method of using FLD to predict system effectiveness, which due to its short processing time allows many combinations to be tested in a relatively small amount of time, yet we see from Fig. 6 that the system with the greatest discriminant value does not necessarily have the lowest EER. Therefore it is highly likely that other face space combinations exist that will produce a lower EER than the best combination presented in this paper. Such a face space combination could easily be found if a more accurate representation of system effectiveness was used in the combination selection criteria. One obvious choice is the EER itself. Although this would take an extremely long time to process, once the dimensions have been identified, the combined face space projection matrix can be stored for latter use and providing the training set is sufficiently large and varied, re-training and re-combining would not be required.

Previous work had shown that image processing improves the fisherface method of face recognition from an EER of 20.1% using no image processing, to 17.8% using the optimum processing technique. We have extended this line of research to show that creating a face space combination, incorporating multiple fisherface systems reduces the EER further, down to 12.8% when tested on a large amount of data presenting typical difficulties when performing recognition. Evaluating this system at its fundamental level, using 258,840 verification operations between face images, demonstrates that combining multiple face space dimensions improves the effectiveness of the core face recognition engine. We have not applied any additional heuristics, typically incorporated into fully functional commercial and industrial systems. For example, we have not experimented with different distance metrics, multiple facial alignments, optimising crop regions or storing multiple gallery images. All of which are known to improve error rates and can easily be applied to the combined face recognition systems presented in this paper. With these additional measures in place, it is likely that the improvements made to the core recognition engine will bring the error rates of fully functional commercial and industrial systems substantially closer to those required for the application scenarios in mind.

REFERENCES

1. T. Heseltine, N. Pears, J. Austin, "Evaluation of image pre-processing techniques for eigenface based face recognition", *Proc. of the 2nd Int. Conf. on Image and Graphics*, SPIE, Vol. 4875, pp. 677-685, 2002.
2. T. Heseltine, N. Pears, J. Austin, "Face Recognition: A Comparison of Appearance-based Approaches", *Proc. of the 7th Int. Conf. on Digital Image Computing: Techniques and Applications*, Vol. 1, pp. 59-69, 2003.
3. G. Finlayson, G. Schaefer, "Hue that is Invariant to Brightness and Gamma", *In Proc. of the British Machine Vision Conference*, Session 3: Colour & Systems, 2001.
4. G.D. Finlayson, B. Schiele, J. L. Crowley, "Comprehensive Colour Image Normalisation", *Proc. of the European Conference on Computer Vision*, LNCS 1406, Springer, pp. 475-490, 1998.
5. M. Turk, A. Pentland, "Eignefaces for Recognition", *Journal of Cognitive Neuroscience*, **Vol. 3**, pp. 72-86, 1991.
6. M. Turk, A. Pentland, "Face Recognition Using Eignefaces", *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.
7. P. Belhumeur, J. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Face Recognition using class specific linear projection", *Proc. of the European Conference on Computer Vision*, pp. 45-58, 1996.
8. A. Pentland, B. Moghaddom, T. Starner, "View-Based and Modular Eigenfaces for Face Recognition", *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1994.
9. A.M. Martinez, R. Benavente, *The AR Face Database*, CVC Technical Report 24, 1998.
10. Y. Adini, Y. Moses, S. Ullman, "Face Recognition: the Problem of Compensating for Changes in Illumination Direction", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19 No. 7, pp. 721-732, 1997.
11. 11 G. Gordon, "Face Recognition Based on Depth and Curvature Features", *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 108-110, 1992.