Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Face Recognition: A Comparison of Appearance-Based Approaches

Thomas Heseltine[1], Nick Pears, Jim Austin, Zezhi Chen

Advanced Computer Architectures Group, Department of Computer Science,
The University of York, York, England
[1]tom.heseltine@cs.york.ac.uk
http://www.cs.york.ac.uk/~tomh

Abstract. We investigate the effect of image processing techniques when applied as a pre-processing step to three methods of face recognition: the direct correlation method, the eigenface method and fisherface method. Effectiveness is evaluated by comparing false acceptance rates, false rejection rates and equal error rates calculated from over 250,000 verification operations on a large test set of facial images, which present typical difficulties when attempting recognition, such as strong variations in lighting conditions and changes in facial expression. We identify some key advantages and determine the best image processing technique for each face recognition method.

1 Introduction

Despite significant advances in face recognition technology, it has yet to be put to wide use in commerce or industry, primarily because the error rates are still too high for many of the applications in mind. These problems stem from the fact that existing systems are highly sensitive to environmental factors during image capture, such as variations in facial orientation, expression and lighting conditions. In this paper we attempt to address these issues by use of image pre-processing techniques, focusing on three face recognition methods, all coming under the general heading of appearance-based approaches: direct correlation; the eigenface method and the fisherface method.

We begin with brief explanations of each face recognition method (section 2, 3 and 4), followed by a performance comparison of each system (section 5) with no image pre-processing (the ëbaseline systemsí). In Section 6 we outline a range of image pre-processing techniques, which may improve the baseline systems.

By applying the face recognition methods to a substantial database of facial images (described in Section 7), producing graphs of FAR (False Acceptance Rate) against FRR (False Rejection Rate), from which the EER (equal error rate) is taken as a single comparative value, we compare the recognition accuracy across the full range of systems (Section 9).

2 The Direct Correlation Method

The direct correlation method of face recognition (also referred to as template matching by Brunelli and Poggio [1]) involves the direct comparison of pixel intensity values taken from facial images. We convert bitmap images of 65 by 82 pixels into a vector of 5330 elements, describing a point within a 5330 dimensional image space. By measuring the distance between these points, we gain an indication of image similarity. Similar images are located close together within the image space, while dissimilar images are spaced far apart. Extending this idea to faces, calculating the Euclidean distance *d*, between two facial image vectors (often referred to as the query image *q*, and gallery image *g*), we get an indication of similarity. A threshold is then applied to make the final verification decision.

$$d = \|q - g\| \qquad (d \le threshold \Rightarrow accept) \wedge (d > threshold \Rightarrow reject)\ . \tag{1}$$

3 The Eigenface Method

In this section we give a brief explanation of the eigenface method of face recognition, while referring the reader to Turk and Pentland [2, 3] for more detailed explanations.

We compute the covariance matrix C, of facial images from a set of M (60) training images: $\{\Gamma_1, \Gamma_2, \Gamma_3,$ Ö $\Gamma_M\}$,

$$C = \frac{1}{M}\sum_{n=1}^{M} \Phi_n \Phi_n^T = AA^T \qquad A = [\Phi_1 \Phi_2 \Phi_3 \ldots \Phi_M] \quad \Phi_n = \Gamma_n - \Psi \quad \Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n\ . \tag{2}$$

The eigenvectors and eigenvalues of this covariance matrix are calculated using standard linear methods and the M\` eigenvectors with the highest eigenvalues chosen to formulate the projection matrix *u*. For the sake of consistency with the fisherface method, we use the first 59 principal components when testing the eigenface method.

Fig. 1. The average face and first five eigenfaces computed with no image pre-processing

A face-key ω (image vector projected into face space) can then be produced by the following equation.

$$\omega_k = u_k^T(\Gamma - \Psi) \qquad for\ k = 1\ Ö\ M`\ . \tag{3}$$

These face-keys (vectors of 59 principal component coefficients) can than be compared using the Euclidean distance measure as with the direct correlation method (see equation 1).

4 The Fisherface Method

The fisherface method of face recognition as described by Belhumeur et al [4] uses both principal component analysis and linear discriminant analysis to produce a subspace projection matrix, similar to that used in the eigenface method. However, the fisherface method is able to take advantage of ëwithin-classí information, minimising variation within each class, yet still maximising class separation.

To accomplish this we expand the training set to contain multiple images of each person, providing examples of how a personís face may change from one image to another due to variations in lighting conditions, facial expression and even small changes in orientation. We define the training set as,

$$\textbf{Training Set} = \{\underbrace{\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4,}_{X_1} \underbrace{\Gamma_5, \Gamma_6, \Gamma_7, \Gamma_8,}_{X_2} \underbrace{\Gamma_9, \Gamma_{10},}_{X_3} \underbrace{\Gamma_{11}, \Gamma_{12},}_{X_4} \underbrace{\Gamma_{13}, \ldots \Gamma_M}_{X_c}\} \quad (4)$$

Where Γ_i is a facial image and the training set is partitioned into c classes, such that all the images in each class X_i are of the same person and no single person is present in more than one class.

We begin by computing three scatter matrices, representing the within-class (S_w), between-class (S_b) and total (S_t) distribution of the training set throughout image space.

$$S_T = \sum_{n=1}^{M} (\Gamma_n - \Psi)(\Gamma_n - \Psi)^T \qquad S_B = \sum_{i=1}^{c} |X_i|(\Psi_i - \Psi)(\Psi_i - \Psi)^T \qquad S_W = \sum_{i=1}^{c} \sum_{\Gamma_k \in X_i} (\Gamma_k - \Psi_i)(\Gamma_k - \Psi_i)^T \quad (5)$$

Where $\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$, is the average image vector of the entire training set, and $\Psi_i = \frac{1}{|X_i|}\sum_{\Gamma_i \in X_i}\Gamma_i$, the average of each individual class X_i (person). By performing PCA on the total scatter matrix S_t, and taking the top $M\text{-}c$ principal components, we produce a projection matrix U_{pca}, which is used to reduce the dimensionality of the within-class scatter matrix, ensuring it is non-singular, before computing the top $c\text{-}1$ (in our case 59) eigenvectors of the reduced scatter matrices, U_{fld} as shown below.

$$U_{fld} = \arg\max_{U} \left(\frac{\left| U^T U_{pca}^T S_B U_{pca} U \right|}{\left| U^T U_{pca}^T S_W U_{pca} U \right|} \right). \quad (6)$$

Finally, the matrix U_{ff} is calculated as shown in equation 7, such that it will project a facial image into a reduced image space of $c\text{-}1$ dimensions, in which the between-class scatter is maximised for all c classes, while the within-class scatter is minimised for each class X_i.

$$U_{ff} = U_{fld} U_{pca} \,. \quad (7)$$

Once the U_{ff} matrix has been constructed it is used in much the same way as the projection matrix in the eigenface system (see equation 3), reducing the dimensionality of the image vectors from 5330 to just 59 ($c\text{-}1$) elements. Again, like the eigenface system, the components of the projection matrix can be viewed as images, referred to as fisherfaces.

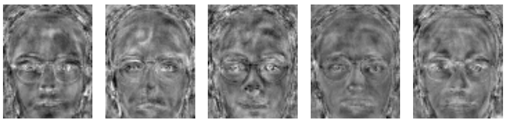


Fig. 2. The first five fisherfaces, defining a face space with no image pre-processing

5 Baseline Results

We begin with some preliminary experimentation; comparing the three baseline systems (direct correlation, eigenface and fisherface) with no image pre-processing, applied to the same database used by Heseltine, Pears and Austin in previous eigenface experiments [5]. The results clearly show that the fisherface method has a significantly higher EER than the eigenface and direct correlation methods (Fig. 3). This seems contradictory to other investigations [4, 6], which identify the fisherface method as the superior system.

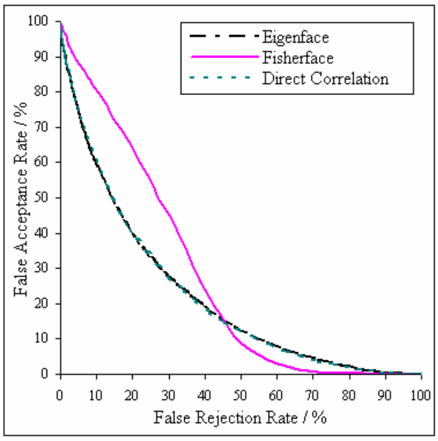


Fig. 3. Error rates of face recognition methods when applied to a test set with partially occluded images

However, looking at the error rate curve, we see that a steep drop in FAR occurs just as the FRR increases above 40%, to the point where the fisherface system outperforms both the eigenface and direct correlation methods. Also, applying the fisherface system to the training set achieves perfect classification.

This suggests that within the test set there are some images for which the fisherface system is not suitable, but as the threshold decreases, effectively discounting these images from the statistics, we see an improvement in performance. After removing the partially occluded images from the test set, it becomes obvious that these

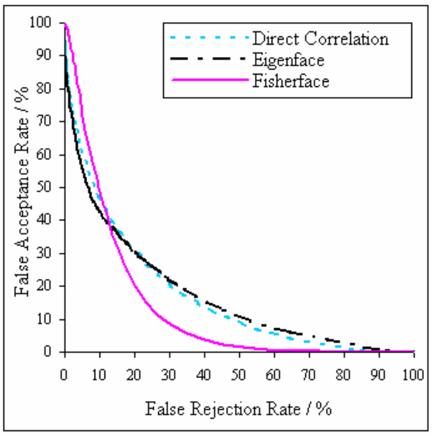


Fig. 4. Error rates of face recognition methods when applied to a test set without partial occlusion

images were the causes of the high EER. The fisherface method is clearly the most accurate, with an EER of 20%, especially if an application is required with a low false acceptance rate.

It is also evident that there is no significant difference between the accuracy of the eigenface and direct correlation methods, with EERs of 25.5% and 25.1% respectively. However, each system does have other advantages, in that the eigenface method requires much less processing time per verification, whereas direct correlation takes more time per verification, but does not required a training phase.

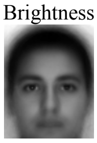
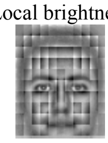
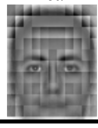
6 Image Pre-processing

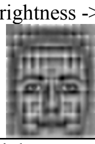
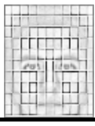
It has been shown that introducing an image pre-processing step to the eigenface method of face recognition can significantly reduce error rates [5]. We now continue this line of investigation; by applying the same pre-processing techniques to the fisherface and direct correlation methods, prior to training and testing each method.

The image pre-processing techniques fall into four main categories: colour normalisation, statistical methods, convolution filters and combinations of these methods. A summary of these techniques (described in more detail by Heseltine et al [5]) is given in table 1.

Table 1. Brief descriptions of image pre-processing techniques, with examples of the average face and equations and pixel template kernels given where appropriate

Colour Normalisation Techniques			
Comprehensive 	Comprehensive colour normalization as described by Finlayson[8], invariant to lighting geometry and colour. The method involves the repetition of intensity normalisation and grey world normalisation, until a stable state is reached.		
Chromaticities 	Summation of the R, G components of colour intensity normalisation.	Comprehensive chromes 	Summation of the R, G components of comprehensive normalisation.
Grey world 	Grey world normalisation. $(r_{norm}, g_{norm}, b_{norm}) = \left(\frac{Nr}{r_1+r_2+...+r_N}, \frac{Ng}{g_1+g_2+...+g_N}, \frac{Nb}{b_1+b_2+...+b_N}\right)$	Bgi hue 	Brightness and gamma invariant hue, introduced by Finlayson and Schaefer [7]. $H = \tan^{-1}\frac{\log(r)-\log(g)}{\log(r)+\log(g)-2\log(b)}$
Intensity 	Intensity normalisation. $(r_{norm}, g_{norm}, b_{norm}) = \left(\frac{r}{r+g+b}, \frac{g}{r+g+b}, \frac{b}{r+g+b}\right)$	Hsv hue 	Standard hue definition. $H = \cos^{-1}\left(\frac{\frac{1}{2}[(r-g)+(r-b)]}{\sqrt{(r-g)(r-g)+(r-b)(g-b)}}\right)$

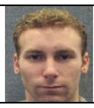
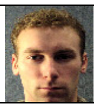
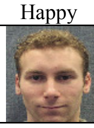
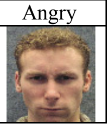
<table>
<tr><th colspan="4">Statistical Methods</th></tr>
<tr><td>Brightness
</td><td>Global transformation of brightness, such that intensity moments are normalised.</td><td>Local brightness
</td><td>Application of brightness method to individual local regions of the image.</td></tr>
<tr><td>Brightness mean
</td><td>Global transformation of brightness, such that the mean becomes a constant specified value.</td><td>Vertical brightness
</td><td>Application of brightness method to individual columns of pixels.</td></tr>
<tr><td>Horizontal brightness
</td><td>Application of brightness method to individual rows of pixels.</td><td>Local brightness mean
</td><td>Transformation of brightness, such that the mean becomes a constant specified value within local regions of the image.</td></tr>
<tr><th colspan="4">Convolution Filters</th></tr>
<tr><td>Smooth
</td><td>Standard low-pass filtering using a 3x3 pixel template.
1 1 1
1 5 1
1 1 1</td><td>Find edges
</td><td>Edge detection followed by segmentation by application of a threshold.
-1 -1 -1
-1 8 -1
-1 -1 -1</td></tr>
<tr><td>Smooth more
</td><td>Smooth filtering with a larger 5x5 pixel neighbourhood.
1 1 1 1 1
1 5 5 5 1
1 5 44 5 1
1 5 5 5 1
1 1 1 1 1</td><td>Blur
</td><td>An extreme blurring effect.
1 1 1 1 1
1 0 0 0 1
1 0 0 0 1
1 0 0 0 1
1 1 1 1 1</td></tr>
<tr><td>Contour
</td><td>Edge detection by application of a 3x3 template.
-1 -1 -1
-1 8 -1
-1 -1 -1</td><td>Detail
</td><td>Enhance areas of high contrast.
0 -1 0
-1 10 -1
0 -1 0</td></tr>
<tr><td>Edge
</td><td>Enhances the edges of an image.
-1 -1 -1
-1 10 -1
-1 -1 -1</td><td>Sharpen
</td><td>Reduces the blur in the image.
-2 -2 -2
-2 32 -2
-2 -2 -2</td></tr>
<tr><td>Edge more
</td><td>Another edge enhancement filter.
-1 -1 -1
-1 9 -1
-1 -1 -1</td><td>Emboss
</td><td>A stylising filter that enhances edges with a shadow casting affect.
-1 0 0
0 1 0
0 0 0</td></tr>
</table>

Method Combinations			
Contour -> Smooth 	Contour filtering followed by smoothing.	Contour + Local brightness 	The summation of the resulting images from the Contour filter and the Local Brightness transformation.
Smooth->Contour 	Smoothing followed by contour filtering.	Local brightness ->Smooth 	Local brightness transformation followed by smoothing.
C->S + LB 	Contour filtering followed by smoothing, summed with the Local Brightness transformation.	Local brightness -> Contour 	Local brightness transformation followed by contour filtering.
S->LB->C 	Smoothing followed by the Local Brightness transformation and Contour filtering.		

7 The Face 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]. The database is separated into two disjoint sets: i) The training set, containing 240 images of 60 people under a range of lighting conditions and facial expressions; ii) the test set containing 720 images (60 people of various gender, race and age, 12 images each). The six examples shown in table 2 were repeated on two days, making up the 12 images of each subject in the test set. All the images are pre-aligned with the centres of the eyes 25 pixels apart. Each image is cropped to a width and height of 65 and 82 pixels respectively.

Table 2. Image capture conditions included in the database test set.

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

8 Test Procedure

Effectiveness of the face recognition methods is evaluated using error rate curves (FRR against FAR) for the verification operation. The 720 images in the test set are compared with every other image using one of the face recognition methods, producing a distance value using equation 1. No image is compared with itself and each pair is compared only once (the relationship is symmetric). A threshold is applied in order to derive the rejection/acceptance decision. Hence, each FRR (percentage of incorrect rejections), and FAR (percentage of incorrect acceptances) pair is calculated from 258,840 verification operations. By varying the threshold we produce a set of FRR FAR plots, forming the error rate curve, as shown in fig. 5. We then take the EER (point at which FRR equals FAR) as a single comparative value.

9 Results

Having tested the full range of image pre-processing techniques, we present the EERs in fig. 6, identifying optimum image processing techniques for each of the three face recognition methods. Both the direct correlation and eigenface methods perform best when used with intensity normalisation, achieving an EER of 18.0% and 20.4% respectively. The fisherface method achieves the lowest EER of 17.8%, when used with the ìslbcî pre-processing technique.

We also see that only a slight improvement is gained by the fisherface method, from 20.1% to 17.8% EER, whereas direct correlation has a much more significant improvement, from 25.1% down to 18.0%. In fact, when using the optimum image pre-processing technique the fisherface method is only marginally better than direct correlation, although it still maintains the advantage of a reduced processing time, due to the shorter length of the projected image vectors.

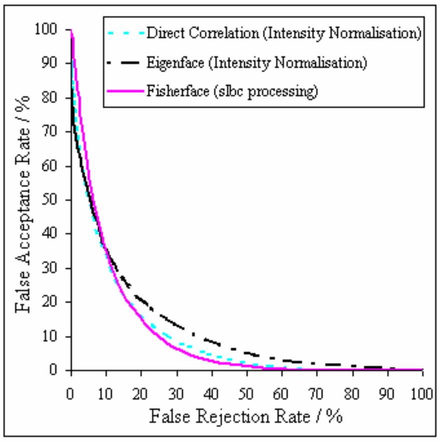


Fig. 5. Error rates of face recognition methods using optimum image pre-processing techniques

10 Conclusion

Initial comparison of the baseline systems produced results that are contradictory to other experiments carried out on the eigenface and fisherface methods [4, 6]. Further investigation identified that the training set used for the fisherface method did not include sufficient examples of all conditions represented in the test data. In order for the fisherface method to perform recognition effectively, it is vital that the training set is an adequate representation of the real application data. If such training data is not available, or the real world image capture conditions cannot be predicted, the eigenface and direct correlation methods are a better alternative. However, providing a suitable training set is available, the fisherface method has significantly lower error rates (20.1%) than both the eigenface (25.5%) and direct correlation methods (25.1%), which are comparable in terms of recognition accuracy. However, with image vectors of 5330 elements, the processing time and storage requirements of the direct correlation method are significantly higher than the eigenface method, which uses vectors of only 59 elements.

We have shown that the use of image pre-processing is able to significantly improve all three methods of face recognition, reducing the EER of the eigenface, fisherface and direct correlation methods by 2.3, 5.1 and 7.1 respectively. However, it has also become apparent that different image pre-processing techniques affect each method of face recognition differently. Although some image processing techniques are typically detrimental (blurring, smoothing, hue representations and comprehensive normalisation) and others are generally beneficial (slbc, sharpen, detail, edge enhance) to recognition, there are also techniques that will decrease error rates for some methods while increasing error rates for others. The most prominent example of this is intensity normalisation, which is evidently the best technique for both direct correlation and eigenface methods, yet increases the EER for the fisherface method.

Taking the optimum image pre-processing technique shows that the fisherface method has the lowest EER (17.8%), yet its lead over the other two methods is considerably reduced. In this case, although much more computationally efficient, it is only marginally better than direct correlation (EER 18.0%), but still maintains a significant improvement over the eigenface method (EER 20.4%).

Further experimentation is required in order to identify which specific features are enhanced by which pre-processing method and in what circumstances a given pre-processing method is most effective. In addition, it may be the case that using a different number of principal components will reduce error rates further, but this may also be dependent on the pre-processing method used.

References

1. Brunelli, R., Poggio, T.: Face Recognition: Features versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (1993) 1042-1052
2. Turk, M., Pentland, A.: Eignefaces for Recognition. Journal of Cognitive Neuroscience, Vol. 3, (1991) 72-86
3. Turk, M., Pentland, A.: Face Recognition Using Eignefaces. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (1991) 586-591
4. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. Fisherfaces: Face Recognition using class specific linear projection. In Proc. ECCV, (1996) 45-58
5. Heseltine, T., Pears, N., Austin, J.: Evaluation of image pre-processing techniques for eigenface based face recognition. In Proc. of the Second International Conference on Image and Graphics, SPIE vol. 4875, (2002) 677-685
6 Marcialis, G., Roli, F.: Fusion of LDA and PCA for Face Recognition. Department of Electrical and Electronic Engineering, University of Cagliari, Piazza díArmi
7. Finlayson, G., Schaefer, G.: Hue that is Invariant to Brightness and Gamma. BMVC01, Session 3: Colour & Systems, (2001)
8. Finlayson, G., Schiele, B., Crowley, J.: Comprehensive Colour Image Normalisation. In Proc. ECCV '98, LNCS 1406, Springer, (1998) 475-490
9. Martinez, A., Benavente, R.: The AR Face Database. CVC Technical Report #24, (1998)

Appendix: Results Table

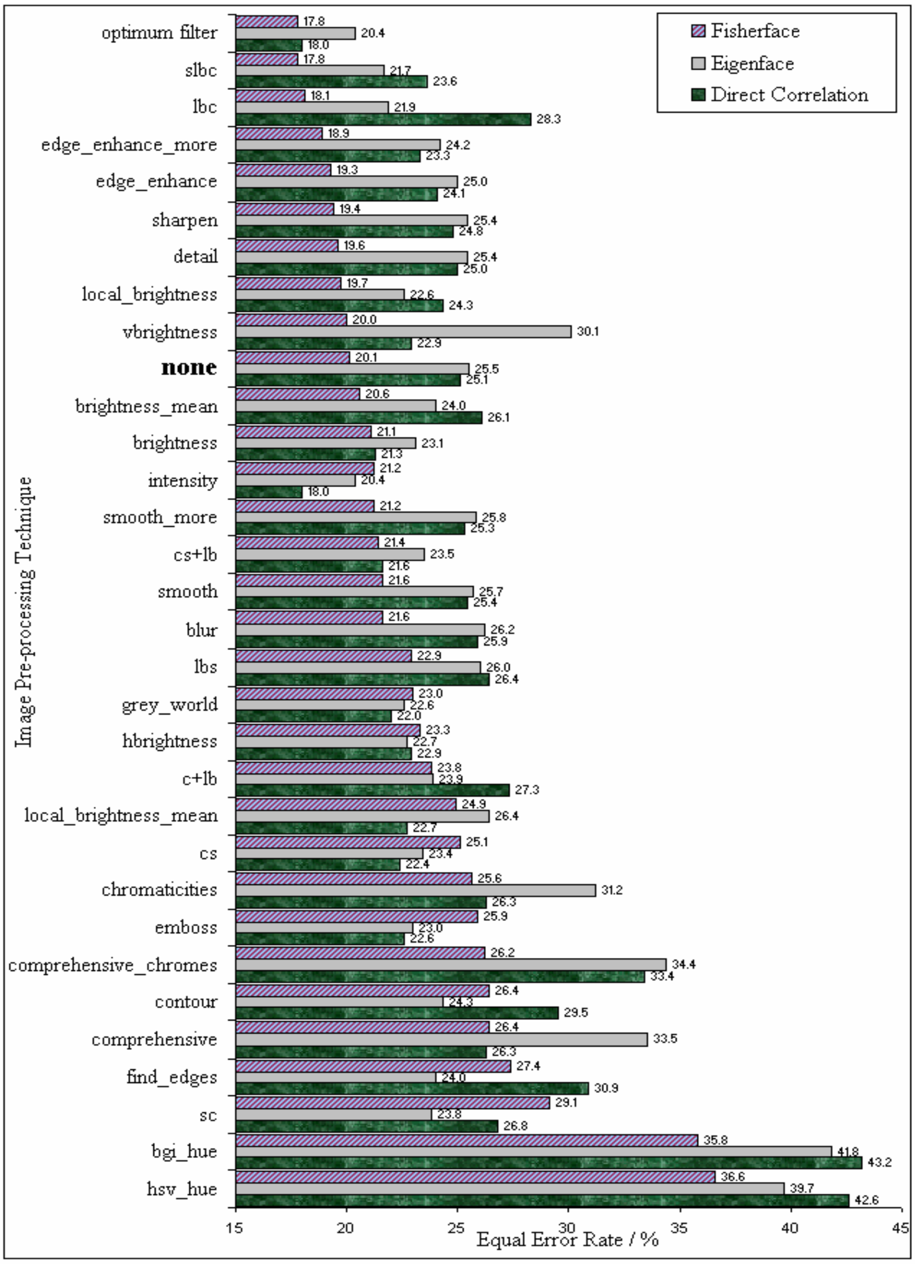


Fig. 6. Equal Error Rates of face recognition methods used with a range of image pre-processing techniques