Evaluation of Image Pre-Processing Techniques for Eigenface Based Face Recognition

3 authors, including:

Thomas Heseltine
The University of York
10 PUBLICATIONS **389** CITATIONS

Jim Austin
The University of York
265 PUBLICATIONS **6,551** CITATIONS

Evaluation of image pre-processing techniques for eigenface based face recognition

Thomas Heseltine[1], Nick Pears and Jim Austin

Advanced Computer Architecture Group
Department of Computer Science
The University of York

ABSTRACT

We present a range of image processing techniques as potential pre-processing steps, which attempt to improve the performance of the eigenface method of face recognition. Verification tests are carried out by applying thresholds to gather false acceptance rate (FAR) and false rejection rate (FRR) results from a data set comprised of images that present typical difficulties when attempting recognition, such as strong variations in lighting direction and intensity, partially covered faces and changes in facial expression. Results are compared using the equal error rate (EER), which is the error rate when FAR is equal to FRR. We determine the most successful methods of image processing to be used with eigenface based face recognition, in application areas such as security, surveillance, data compression and archive searching.

Keywords: Face Recognition, Eigenface, Image Pre-processing, Lighting.

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

1 INTRODUCTION

The Eigenface technique for face recognition is an information theory approach based on principal component analysis, as proposed by Turk and Pentland[1, 2]. We take a set of M training images and compute the eigenvectors of their covariance matrix, then select the M` eigenvectors (eigenfaces) with the highest eigenvalues to define an image subspace (face space). By projecting a face-image into face space we obtain a 'face-key' vector of M` dimensions. We define the 'likeness' of any two face-images as the Euclidean distance between their respective 'face-key' vectors.

Using this method, we perform many comparisons between different images of the same face and images of different faces. By applying a range of threshold values to the distance measurements of these comparisons, we obtain false acceptance rates (FAR) and false rejection rates (FRR). The equal error rate (EER) is used as a single measure of the effectiveness of the system and is attained at a specific threshold value.

A problem inherent to information theory approaches is that there are often features within the training data, identified as principal components, which may not necessarily discriminate between faces but between the circumstances under which the images were captured. These may include such factors as lighting direction, intensity and colour, head orientation, image quality and facial expression. For example, suppose some images in the training data were taken with bright sunlight shining on one side of the face. The feature of having one side of the face lighter than the other may be identified as a principle component and hence used to distinguish between different people. This should clearly not be the case.

We introduce a pre-processing step as an addition to the above system, in an attempt to reduce the presence of these unwanted features in both the training and test data. This pre-processing takes place prior to any principal component analysis of training data and before projection of any test images into face space. The result being that the eigenfaces produced are a more accurate representation of the differences between people's faces rather than the differences between environmental factors during image capture. The image processing techniques include a range of standard image filtering algorithms as found in many graphics applications, some image normalisation methods (including those discussed by Finlayson et al[3,4]) and some of our own image processing algorithms.

2 RELATED WORK

The problems caused by changes in lighting conditions have been well researched. Adini, Moses and Ullman suggest that the differences between images of one face under different illumination conditions are greater than the differences between images of different faces under the same illumination conditions[5]. It has been attempted, with some success, to identify and compensate for the effect of lighting conditions in various face recognition systems. Zhao and Chellappa use a generic 3D surface of a face, together with a varying albedo reflectance model and a Lambertian physical reflectance model to compensate for both the lighting and head orientation[6], before applying a recognition system based on linear discriminant analysis. Much research has also been carried out to improve eigenface recognition systems. Cutler has shown that it can successfully be applied to infrared images[7], resulting in a much decreased error rate. This shows that an artificial infrared light source could be used to reduce the effect of external light sources, producing an accurate system for use in security applications such as site access. However, the use of such a light source is not always practical, particularly if the camera is far from the subject.

Pentland, Moghaddam and Starner extended their eigenface system to include multiple viewing angles of a persons face[8], improving the systems performance when applied to faces of various orientations. They also incorporate a modular eigenspace system[8], which significantly improves the overall performance, although it does not tackle the lighting problem directly. Belhumeur, Hespanha and Kreigman use Fisher's linear discriminant to capture the similarities between multiple images in each class (per person)[9], hoping to discount the variations due to lighting from the defined subspace. Their results show a significant improvement over the standard eigenface approach (from 20% to less than 2% error rate). However, as both the eigenface and fisherface methods are both information theory approaches, it's likely that additional image processing could improve both methods. Although Adini, Moses and Ullman point out that there is no image representation that can be completely invariant to lighting conditions, they do show that different representations of images, on which lighting has less of an affect, can significantly reduce the difference between two images of the same face[5].

3 THE FACE DATABASE

We conduct experiments using a database of 960 bitmap images of 120 individuals (60 male, 60 female), extracted from the AR Face Database provided by Martinez and Benavente[10]. We separate the database into two disjoint sets: i) The training set, containing 60 images of different people of various gender, race and age taken under natural lighting conditions with neutral expression; ii) the test set containing 900 images (15 images of 60 people of various gender, race and age). Each of the 15 images were taken under the conditions described in table 1 (examples in Figure 1).

Expressions\covering	Lighting			
	Natural	From the left	From the right	From left & right
Neutral expression	Day 1, Day 2	Day 1, Day 2	Day 1, Day 2	Day 1, Day 2
Happy expression	Day 1, Day 2			
Angry expression	Day 1, Day 2			
Mouth covered	Day 1	Day 1	Day 1	

Table 1. Image capture conditions.

All the images are stored as bitmaps (converted into vectors for PCA). After initial investigations to determine an appropriate resolution, we find there is no significant change in the EER for resolutions of 100 pixels between the eyes down to just 15 pixels. As a compromise between test execution time and some pre-processing techniques possibly working better with higher resolutions, we select 25 pixels distance between the eyes to conduct all further experiments. All images are scaled and rotated such that the centres of the eyes are aligned 25 pixels apart, using our eye detection algorithm (not described here). Each image is cropped to a width and height of 75 and 112 pixels respectively.

Figure 1. Example test set images for a single person (first 6 repeated on two days).

4 DEFINING FACE SPACE

In this section we review the eigenface method of face recognition for completeness. Consider our training set of images of 75 by 112 pixels. These images could be represented as two-dimensional (75 by 112) arrays of pixel intensity data. Similarly, vectors of 8400 (75x112) dimensions could represent the images. Interpreting these vectors as describing a point within 8400 dimensional space means that every possible image (of 75 by 112 pixels) occupies a single point within this image space. What's more, similar images (for example images of faces) should occupy points within a fairly localised region of this image space. Taking this idea a step further, we assume that different images of the same face map to nearby points in image space and images of different faces map to far apart points. Ideally, we wish to extract the region of image space that contains faces, reduce the dimensionality to a practical value, yet maximise the spread of different faces within the image subspace. Here we apply Principal Component Analysis to define a space with the properties mentioned above.

We take the set of M training images (in our case M = 60): $\{\Gamma_1, \Gamma_2, \Gamma_3, \ldots \Gamma_M\}$ and compute the average image $\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$, followed by the difference of each image from the average image $\Phi_n = \Gamma_n - \Psi$. Thus we construct the covariance matrix as,

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T \qquad where$$
$$= AA^T \qquad A = [\Phi_1\Phi_2\Phi_3...\Phi_M] \qquad (1)$$

The eigenvectors and eigenvalues of this covariance matrix are calculated using standard linear methods. These eigenvectors describe a set of axes within the image space, along which there is the most variance in the face images and the corresponding eigenvalues represent the degree of variance along these axes. The M eigenvectors are sorted in order of descending eigenvalues and the M` greatest eigenvectors (in our system M` = 30) are chosen to represent face space. The effect is that we have reduced the dimensionality of the space to M`, yet maintained a high level of variance between face images throughout the image subspace.

Each eigenvector contains 8400 elements (the number of pixels in the original bitmap) and can be displayed as images, as shown in Figure 2.

Figure 2. Average face image and the first 5 eigenfaces defining a face space with no image pre-processing.

Due to the likeness to faces, Turk and Pentland to refer to these vectors as eigenfaces, and the space they define, face space.

5 VERIFICATION OF FACE IMAGES

Once face space has been defined, we can project any image into face space by a simple matrix multiplication:

$$\omega_k = u_k^T (\Gamma - \Psi) \qquad \text{for } k = 1,...M` \qquad (2)$$

where u_k is the kth eigenvector and ω_k is the kth weight in the vector $\Omega^T = [\omega_1, \omega_2, \omega_3, \ldots \omega_{M`}]$. The M` weights represent the contribution of each respective eigenface and by multiplying the eigenfaces by their weight and summing, we can view the face image as mapped into face space (shown in Figure 3).

Figure 3. Test images and their face space projections.

The vector, Ω, is taken as the 'face-key' for a person's image projected into face space. We compare any two 'face-keys' by a simple Euclidean distance measure, $\varepsilon = \|\Omega_a - \Omega_b\|^2$. An acceptance (the two face images match) or rejection (the two images do not match) is determined by applying a threshold. Any comparison producing a distance below the threshold is a match.

To gather results for the False Rejection Rate, each of the 15 images for a single person, is compared with every other image of their face. No image is compared with itself and each pair is only compared once (the relationship is symmetric), giving $n_c = \frac{1}{2} n_p (n_i^2 - n_i) = 6300$ comparisons to test false rejection, where n_p is the number of people and n_i is the number of face images per person.

False acceptance results are gathered using only images of the type "Day1, neutral expression, natural lighting" and "Day 2, neutral expression, natural lighting." Other comparisons are unlikely to produce false acceptances (due to the combination of different lighting, expression and obscuring), resulting in an initial low ERR, hence any effect of image

processing on the more problematic comparisons would be seemingly reduced. Using these images, every person is compared with every other person. This gives 4 comparisons per pair of people, with no person compared to himself and each pair only compared once. Thus, a total of $n_c = 2(n_p^2 - n_p) = 7080$ comparisons are made between different people to test false acceptance, where n_p is the number of people (60).

Hence, each equal error rate is based on 13380 verification attempts, using a set of comparisons under a range of conditions, such that the FRR and the FAR are maximised. For each threshold value we produce a FAR and a FRR. By applying a range of threshold values we produce a range of FAR and FRR pairs that are plotted on a graph. The results for our benchmark system (no pre-processing) can be seen below. The equal error rate can be seen as the point where FAR equals FRR (Figure 4).
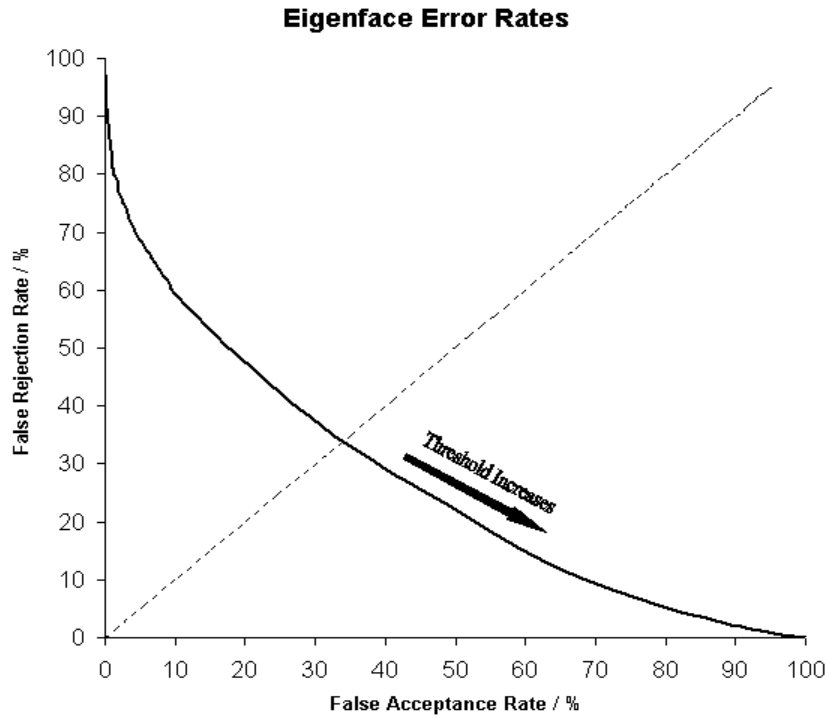
Eigenface Error Rates

Figure 4. Error rates from various thresholds using an eigenface system with no image pre-processing.

6 IMAGE PRE-PROCESSING

We are now in a position to test the effect of image pre-processing on the system performance. A summary of our procedure is given in Figure 5. We present a range of pre-processing techniques, which may affect the EER of the eigenface system, when applied to face images prior to recognition. The image processing methods fall into four main categories: colour normalisation methods, statistical methods, convolution methods and combinations of these methods. The methods are used to produce a single scalar value for each pixel. Examples of these pre-processing methods can be seen in Figure 7.
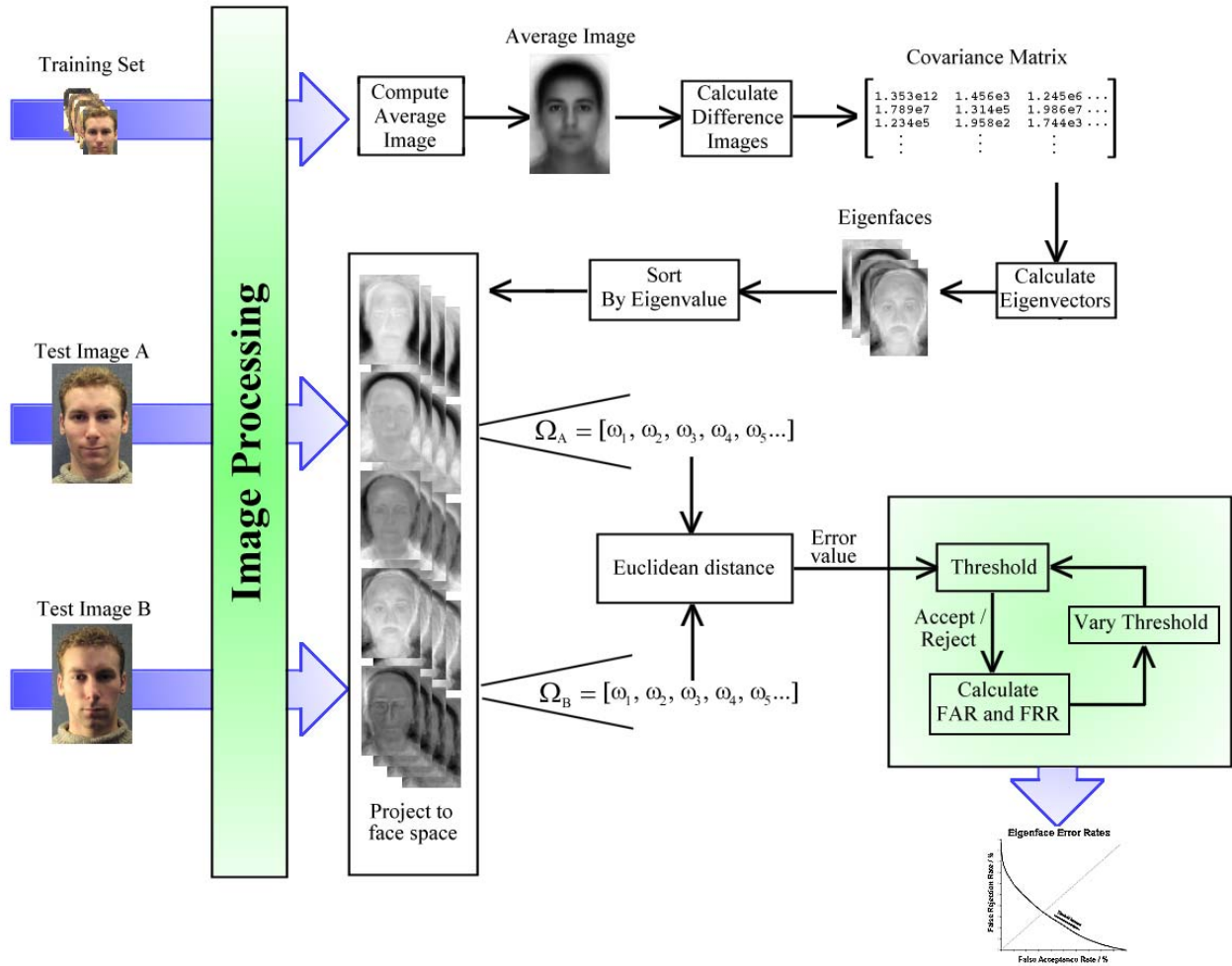
Figure 5. Chart showing the test procedure

6.1 COLOUR NORMALISATION METHODS

6.1.1 INTENSITY: INTENSITY NORMALISATION

We use a well-known image intensity normalisation method, in which we assume that, as the intensity of the lighting source increases by a factor, each RGB component of each pixel in the image is scaled by the same factor. We remove the effect of this intensity factor by dividing by the sum of the three colour components.

$$(r_{norm}, g_{norm}, b_{norm}) = \left(\frac{r}{r+g+b}, \frac{g}{r+g+b}, \frac{b}{r+g+b} \right)$$

(3)

Since the pixels of the resulting image have equal intensity, summing the three colour channels would result in a blank image. Therefore, to create an image with single scalar values for each pixel (as required by our eigenface system) we can either take a single colour channel, or sum just the red and green components (the chromaticities).

6.1.2 GREY WORLD: GREY WORLD NORMALISATION

Here we take a similar approach to the above normalisation, but compensating for the effect of variations in the colour of the light source. Different colours of light cause the RGB colour components of an image to scale apart, by factors α, β and γ respectively, $(r_{new}, g_{new}, b_{new}) = (\alpha r, \beta g, \gamma b)$. Which is normalised using the equation below.

$$(r_{norm}, g_{norm}, b_{norm}) = \left(\frac{Nr}{r_1 + r_2 + ... + r_N}, \frac{Ng}{g_1 + g_2 + ... + g_N}, \frac{Nb}{b_1 + b_2 + ... + b_N} \right)$$

(4)

6.1.3 COMPREHENSIVE: COMPREHENSIVE COLOUR IMAGE NORMALISATION

We use an algorithm proposed by Finlayson[3], which normalises an image for variations in both lighting geometry and illumination colour. The method involves the repetition of intensity normalisation followed by grey world normalisation (as described above), until the resulting image reaches a stable state (i.e. the change in pixel values from one cycle to another is sufficiently small).

6.1.4 HSV HUE: STANDARD DEFINITION OF HUE

The hue of an image is calculated using the standard hue definition, such that each pixel is represented by a single scalar value H.

$$H = \cos^{-1} \left(\frac{\frac{1}{2}[(r-g)+(r-b)]}{\sqrt{(r-g)(r-g)+(r-b)(g-b)}} \right)$$

(5)

6.1.5 BGI HUE: HUE THAT IS INVARIANT TO BRIGHTNESS AND GAMMA

Finlayson and Schaefer introduce a definition of hue that is invariant to brightness (the scaling of each colour channel by a constant factor) and gamma (raising the colour channels to a constant power)[4], which are often caused by variations in scene environment and capture equipment.

$$H = \tan^{-1} \frac{\log(r) - \log(g)}{\log(r) + \log(g) - 2\log(b)}$$

(6)

In summary, the colour normalisation methods used to process the images are shown in table 2.

Intensity	Colour intensity normalisation.
Chromaticities	Summation of the R, G components of colour intensity normalisation.
Grey world	Grey world normalisation.
Comprehensive	Comprehensive colour normalisation.
Comprehensive chromes	Summation of the R, G components of comprehensive normalisation..
Hsv hue	Standard hue definition.
Bgi hue	Brightness and gamma invariant hue.

Table 2. Colour normalisation methods.

6.2 STATISTICAL METHODS

We introduce some statistical methods that apply transformations to the image intensity values in order to make the brightness and contrast constant for all images. The effect is that every image appears to be equally as bright (as a whole) and span across an equal range of brightness.

These statistical methods can be applied in a number of ways, mainly by varying the areas of the image from which the statistics are gathered. It is not necessarily the case that lighting conditions will be the same at all points on the face, as the face itself can cast shadows. Therefore, in order to compensate for the variations in lighting conditions across a single face, we can apply these methods to individual regions of the face. This means that, we are not only compensating for a difference in lighting conditions from one image to another, but also for different lighting conditions from one area of the face to another. In summary, the methods we evaluated are shown in table 3.

Brightness	Global transformation of brightness, such that intensity moments are normalised.
Horizontal brightness	Application of brightness method to individual rows of pixels.
Vertical brightness	Application of brightness method to individual columns of pixels.
Local brightness	Application of brightness method to individual local regions of an image.
Local brightness mean	Transformation of brightness, such that the mean becomes a constant specified value within local regions of the image.

Table 3. Statistical methods used.

6.3 CONVOLUTION METHODS

Convolution methods involve the application of a small template to a window, moved step-by-step, over the original image. These templates can be configured to enhance or suppress features, reduce noise and extract edges. The templates evaluated are described in table 4.

Smooth ($\sigma = 0.788$)	Standard low-pass filtering using a 3x3 pixel template.
Smooth more ($\sigma = 1.028$)	Similar to the above, only with a larger 5x5 pixel neighbourhood.
Blur	An extreme blurring effect.
Edge	Enhances the edges of an image.
Edge more	Same as the above only more so.
Find edges	Segmentation of an image to include only those pixels that lie on edges.
Contour	Similar to Find edges, only more sensitive to changes in contrast.
Detail	Enhance areas of high contrast.
Sharpen	Reduces the blur in the image.
Emboss	A stylise type filter that enhances edges with a shadow casting affect.

Table 4. Convolution methods used.

6.4 METHOD COMBINATIONS

In an attempt to capture the advantages of multiple image processing methods, we combine some of those methods that produce the best improvement in EER, as shown in table 5.

Contour -> Smooth	Contour filtering followed by smoothing.
Smooth->Contour	Smoothing followed by contour filtering.
Local brightness -> Smooth	Local brightness transformation followed by smoothing.

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

Table 5. Method combinations used.

7 RESULTS

We present the results produced by using various image processing methods as a bar chart of EERs (Figure 6). The base-line eigenface system (no image processing) is displayed in the chart as a dark red bar. It can be seen that the majority of image processing methods did produce some improvement to the eigenface system. However, what is surprising is the large increase in error rate produce by some of the colour normalisation methods of image processing, most notably the brightness and gamma invariant hue introduced by Finlayson and Schaefer[4]. We believe that the reduction in effectiveness when using such methods is due to the loss of information during these procedures. Edges become less defined and some of the shading due to geometrical structure is lost (see Figure 7). An increase is also witnessed using the blurring filters. It is therefore not surprising to see that the edge-enhancing methods had a positive impact on the EERs (the find edges and contour filters were particularly effective), as did the statistical methods, which normalise intensity moments (increasing the shading gradient in many areas).

Having identified the most successful image processing method of those evaluated, as normalising intensity moments within local regions of the image, then applying a convolution contour filter, we continue to improve the system by testing different cropping of images to find the optimum for this image processing method, reaching an EER of 22.4% (Figure 8).

8 CONCLUSION

We have shown that the eigenface-based method of face recognition can be significantly improved by means of simple image pre-processing techniques. Without any alterations to the eigenface technique itself, an EER of 22.4% percent can be achieved (a reduction of 11.6%) using a data set containing a majority of extremely difficult images (20% of the images are partially obscured and 40% of the images have extreme lighting conditions).

There are some factors that may be the cause of the remaining 22.4% error, which were not compensated for by the image pre-processing techniques. Firstly, the eye detection algorithm was be no means perfect and although an attempt was made to manually correct any misaligned images, it is clear (from browsing the database) that some images are not aligned well. It would be relatively simple to implement a system in which several small translations and scales of the original image were projected into face space for each recognition attempt, hence compensating for any inaccuracies in the alignment procedure.

Comparable improvements have been witnessed in similar PCA methods of face recognition, such as Pentland et al's modular eigenspace system[8] and in Belhumeur et al's comparison to Fisher faces[9]. It is likely that the image pre-processing methods described could be of similar benefit to these algorithms, and result in a greatly improved face recognition system.

REFERENCES

1. M. Turk, A. Pentland. "Eignefaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 72-86, 1991.

2. M. Turk, A. Pentland. "Face Recognition Using Eignefaces", *In Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

3. G.D. Finlayson, B. Schiele, J. L. Crowley. "Comprehensive Colour Image Normalisation", *Proc. ECCV '98*, LNCS 1406, Springer, pp. 475-490, 1998.

4. G. Finlayson, G. Schaefer. "Hue that is Invariant to Brightness and Gamma", BMVC01, Session 3: Colour & Systems, 2001.

5. 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.

6. W. Zhao, R. Chellappa. "3D Model Enhanced Face Recognition", *In Proc. Int. Conf. Image Processing*, Vancouver, 2000.

7. R. Cutler. "Face Recognition Using Infrared Images and Eigenfaces", citeseer.nj.nec.com/456378.html, 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 (CVPR'94)*, 1994.

9. P. Belhumeur, J. Hespanha, D. Kriegman. "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI19 (7), pp. 711-720, 1997.

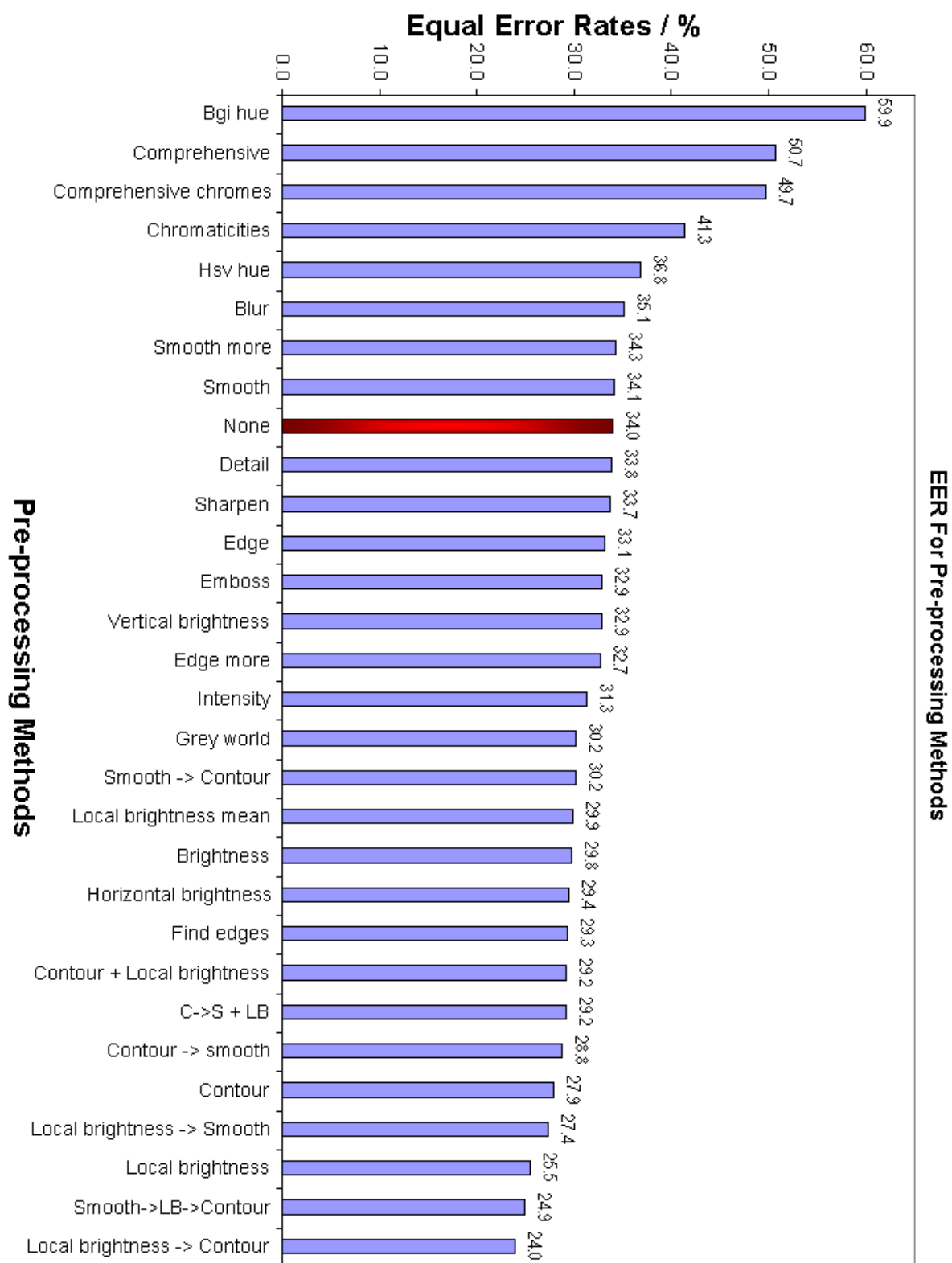10. A.M. Martinez, R. Benavente. "The AR Face Database." CVC Technical Report #24, 1998.

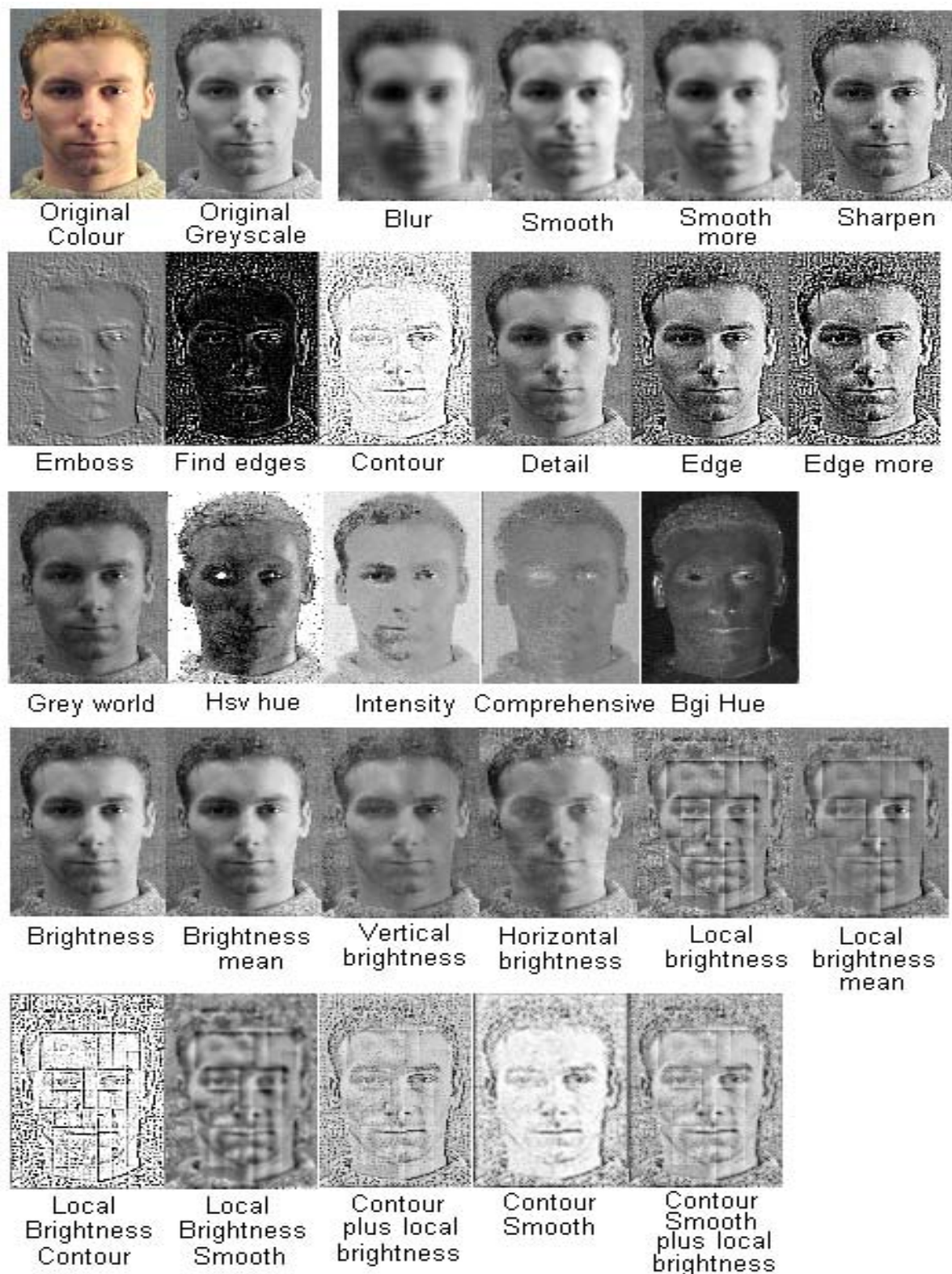**Figure 6. Equal Error Rate Results for various image processing methods.**

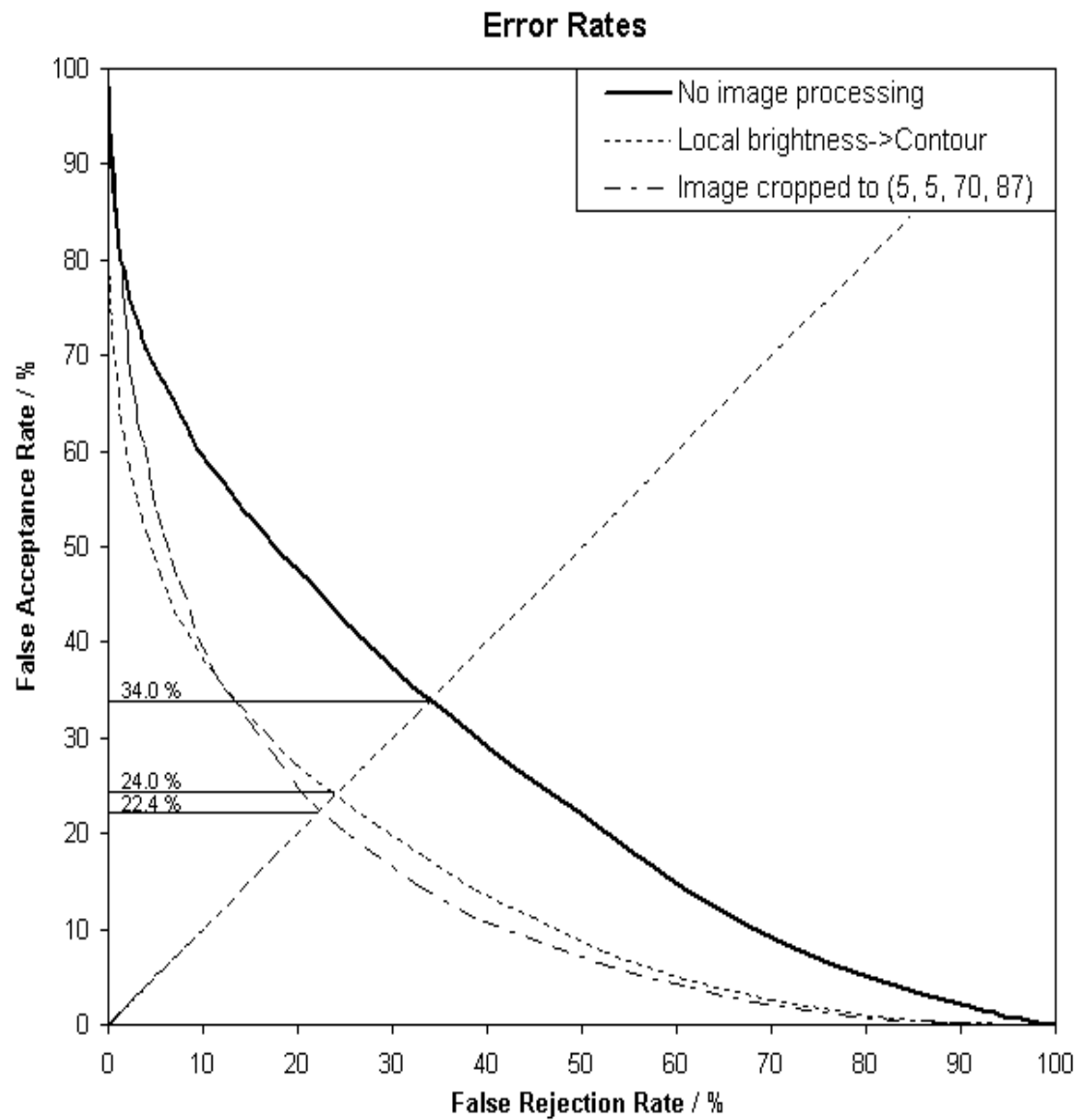**Figure 7. Examples of pre-processed images**

Figure 8. Error rates for base-line system, most successful pre-processing and image crop