THE IMPORTANCE OF THE COLOR INFORMATION IN FACE RECOGNITION [1]

L. Torres   J. Y. Reutter   L. Lorente

luis @gps.tsc.upc.es   jean-yves.reutter@epfl.ch   alorente@gps.tsc.upc.es

Polytechnic University of Catalonia
Barcelona, Spain

ABSTRACT

A common feature found in practically all technical approaches proposed for face recognition is the use of only the luminance information associated to the face image. One may wonder if this is due to the low importance of the color information in face recognition or due to other less technical reasons such as the no availability of color image database. Motivated by this reasoning, we have performed a variety of tests using a global eigen approach developed previously [1], which has been modified to cope with the color information. Our results show that the use of the color information embedded in a eigen approach, can improve the recognition rate when compared to the same scheme which uses only the luminance information.

1. INTRODUCTION

Face recognition has been object of much interest in the last years [2] [3]. In addition to many other applications, there is an increasing interest to specify standardized descriptions of various types of multimedia information. This description will be associated with the content itself, to allow fast and efficient searching for material that is of interest to the user. This effort is being conducted within the activities of the new standard MPEG-7 (Multimedia Content Description Interface) [4].

Different approaches have been proposed for face recognition. Among the most promising, let us mention eigenface, elastic matching, and neural nets [5]. The eigenface approach, also known by Principal Component Analysis (PCA), has been object of much effort and has drawn the attention of many researchers due to its excellent results [6].

In . PCA, the recognition system is based on the representation of the face images using the so called *eigenfaces*. A global approach using eigenfaces has been provided in [1]. However, to the best of our knowledge, none of these approaches use the color information.

In the following, we explain how the scheme developed in [1] has made use of the color information. Results are provided that seem to indicate that color information expressed in certain color spaces may be useful in face recognition.

2. GLOBAL EIGEN APPROACH

Our system for face recognition is based on the representation of facial images using eigenfaces [6]. A companion paper which deals with face recognition in video sequences is presented in [10]. We represent every face image as a vector and express it in an orthogonal basis, computed from the training images set. Suppose that we have N training vectors $\vec{x}_1, \vec{x}_2, ..., \vec{x}_N$.

The covariance matrix C can be estimated by

$$C \cong \frac{1}{N} \sum_{k}^{N} (\vec{x}_k - \vec{\mu})(\vec{x}_k - \vec{\mu})^T$$

where $\vec{\mu}$ is the estimation of the training vectors expectation.

We are now able to compute the N-1 eigenvectors $\vec{e}_i$ of this matrix and sort them by descending module of their eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{N-1}$. These eigenvectors are usually called eigenfaces because they look like faces. Once the corresponding eigenfaces are computed, they are used to represent the training and the test faces to be identified. This is done by a projection on the eigenfaces, using a scalar product. Any training image can be obtained by a linear combination of the eigenfaces:

$$\vec{x} = \vec{\mu} + \sum_{i=1}^{N-1} \hat{x}_i \vec{e}_i$$

where $\hat{x}_i = \vec{x}_i \cdot \vec{e}_i$.

The first eigenfaces contain the most information, in the RMS sense, so that we can express a test vector $\vec{y}$ in the eigenspace, to express it in term of the principal components of the training vectors:

$$\vec{y} = \vec{\mu} + \sum_{i=1}^{N-1} \hat{x}_i \vec{e}_i \; .$$

The training and test vectors can be rewritten

$$\vec{\hat{x}} = \begin{pmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_{N-1} \end{pmatrix} \text{ and } \vec{\hat{y}} = \begin{pmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_{N-1} \end{pmatrix}$$

Then the recognition is performed using the maximum likelihood principle by a distance computation. The selected training image is the one which have the minimal distance under the eigenbasis:

$$k_0 = \arg \min_{1 \le k \le K} d(\vec{\hat{x}}_k, \vec{\hat{y}}) \; .$$

Our system uses the Mahalanobis distance, defined as

$$d(\vec{\hat{x}}, \vec{\hat{y}}) = \sqrt{\sum_{i=1}^{N-1} \frac{(\hat{y}_i - \hat{x}_i)^2}{\lambda_i}}$$

To improve the recognition stage, in addition to the eigenfaces, we have also incorporated to the overall scheme some other eigenfeatures. As in [7], we have extended the eigenface concept to certain parts of the face: both eyes (left and right eigeneyes), the nose (eigennoses) and the mouth (eigenmouth).In the recognition stage, the corresponding overall eigenfaces, right and left eigensides, right and left eigeneyes, eigennoses and eigenmouths are found for the training and test faces. The normalized Mahalanobis distance is applied between every eigenfeature of the test and training faces.

This scheme is applied for each color component separately. For every color component, we obtain a set of eigenfeatures. Figure 1 shows the structure of the face recognition system.
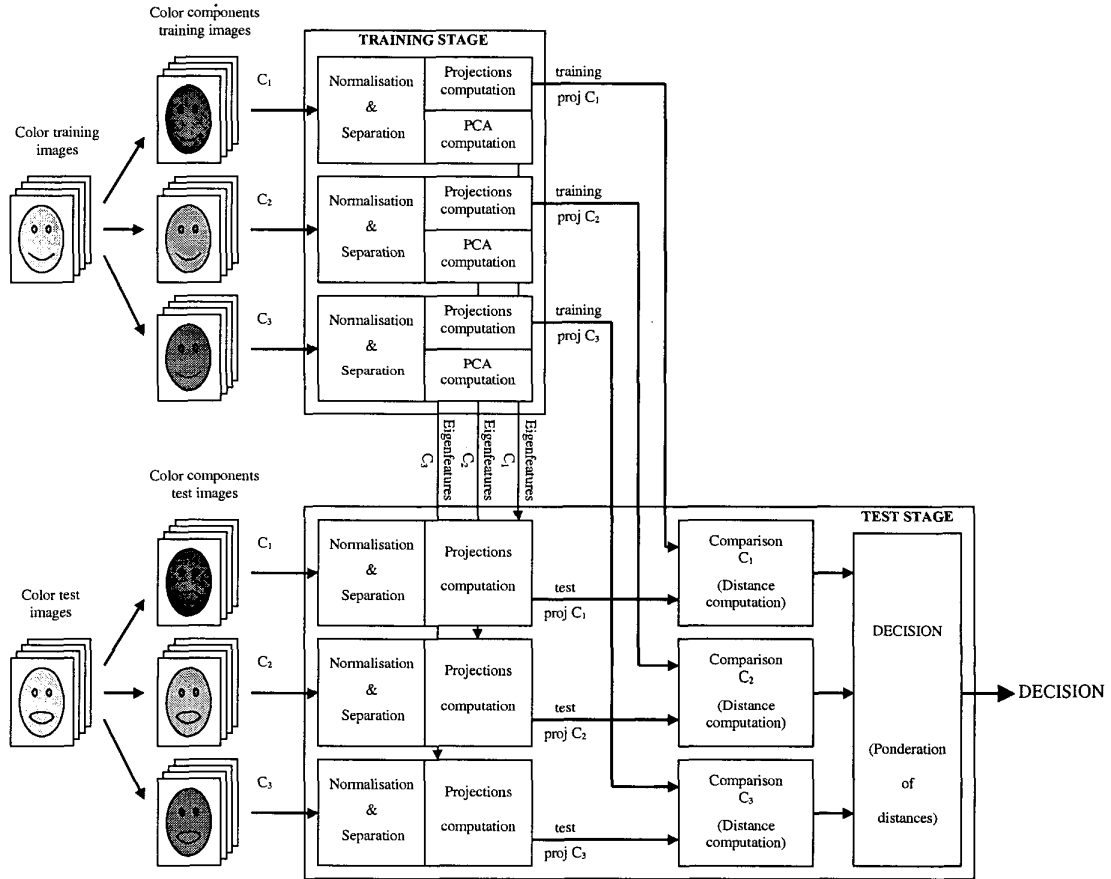

Figure 1. Scheme of the face recognition system using the color information.

The global distance between the test and the training image is the weighted sum of the different distances in such a way that the contribution of each eigenfeature to the recognition stage is the same, and making good use of each color component. Weighting factors may be altered to give more importance to certain features than others. This is important for the color information as different weights can be given to the different color components as well as to the different eigenfeatures, thus emphasizing each one appropriately.

In order to minimize the differences due to changes in size, expression and orientation we have applied first a morphing technique to every image. The morphing process gives an image face in which the facial feature points have been moved to standard predefined positions. This is equivalent to a normalization of the shape of the faces. These standard positions correspond to a frontal view with neutral expression and a given fixed size. The morphing technique is based on texture mapping, a widely known technique in computer graphics [8]. The image of a face is split into triangular polygons (based on the Candide model [9]) whose vertices are characteristic facial points.

The texture of each original triangle is mapped to the triangle defined by the standard points. The points used in this synthesis stage are obtained from the Candide model that has been previously matched to the face. Although some shape information is lost during this morphing stage, the recognition system greatly improves because the resulting images are normalized in size, expression, and above all in orientation. The morphing is applied to all the faces to be identified, independently of their degree of variation in expression or orientation.

Our original face recognition system uses a facial model with 44 points and 78 polygons (triangles), based on the Candide model. The location of the facial points was done manually. The use of the face recognition system with MPEG-1 video sequences (accepted by the MPEG-7 evaluation group) implies that the system must be able to work with small face images. In addition, the final face recognition system must be fully automatic or semi-automatic, a difficult task with small face images with a significant noise level. For all these reasons we have developed a simplified image face model that only needs to know the location of four main facial points (center of both eyes, nose and mouth). The automatic location of these points is an affordable task even with small images. At the present stage of the development the points are located manually. An automatic location algorithm is being developed with the eigenfeature technology.

3. RESULTS

Given the increasing interest on MPEG-7 video indexing activities, we have tested the general eigen scheme on the test video sequences accepted in the MPEG-7 evaluation

group. The training system contains 61 images of different persons (one per individual). For the test stage, 59 images from the same persons that are in the training set are used, but with different views. The global eigen scheme has been applied to each color component independently. Comparisons have been performed between the eigen scheme with only luminance information Y, the eigen scheme using the Y,U,V color space, the R,G,B color space and the H,S,V color space.

When using appropriated weights for the eigenfeatures and the color components, the system provides 3 types of results: Wrong recognition in all cases, Wrong recognition when using only the luminance but correct when using the color information, and correct recognition in all cases. Figure 2 shows these 3 types of results:

1. The system provides a wrong recognition with or without the color information.
2. The system provides a wrong result using only the luminance, but a correct result using the color information.
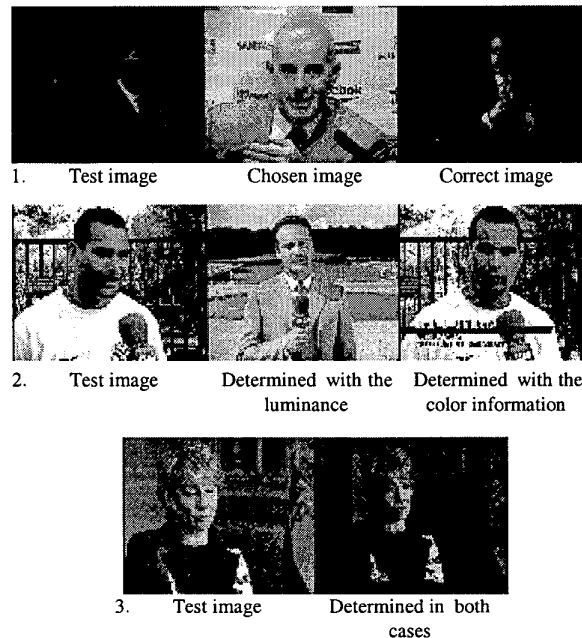3. The system provides a correct result with or without the color information.

1.   Test image   Chosen image   Correct image

2.   Test image   Determined with the   Determined with the
                         luminance   color information

3.   Test image   Determined in both
                             cases

Figure 2. The three types of recognition results.

The percentage of recognition success using only the luminance component is 50 out of 59 images (84.75%). The percentage of success using simultaneously Y,U and V is 52 out of 59 images (88.14%). Using the same face images but grouped in a different way provides a similar improvement

when using the color information in the YUV color space. 48 of 59 recognized faces (81.36%) when using only the luminance and 51 out of 59 images (86.44%) when using the color information.

The same percentage of recognized faces is obtained with the color information expressed in the H, S, V color space. But in this case, only two color components have been used. The H component (H) has been discarded for the recognition because it diminishes the recognition rate.

The color information expressed in the RGB color space has not provided any improvement of the recognition rate. Using R, G or B independently provides the same results as using R,G and B simultaneously and exactly the same results as using the luminance (Y) only.

Color components	Recognition rate
Y	84.75%
R G B	84.75%
Y U V	88.14%
S V	88.14%

Table 1. Comparison of the recognition results when using diferent color components

These results are summarized in table 1. They show that the used color space does have an importance for face recognition. In addition, a color space with a separation of the luminance and chrominance informations tends to provide better results than a color space with these informations mixed (as RGB). Moreover, some color components seem to be more appropriate for face recognition. A glance at the images obtained for the different color components (figure 3) helps to understand the reason. The images show what are the most important color components for face recognition. In the YUV color space, the Y component (luminance) is the most important information. The components of the RGB color space provide almost the same information, for the human eye and for face recognition. Finally, the hue component (H) of the HSV color space looks very strange. This is due to the non-linearity of the color space transformation. This component has to be discarded for face recognition.

The analysis of the detailed results obtained with these color spaces shows that the recognized faces are not exactly the same, although the recognition rate is the same. Figure 4 shows the importance of the color space. In the first case, the color information is useful, but it has to be expressed in the YUV color space. In the second one, the system provides a correct result with the color information expressed in the HSV color space. This shows that it could be possible to obtain better results by using a custom color space.
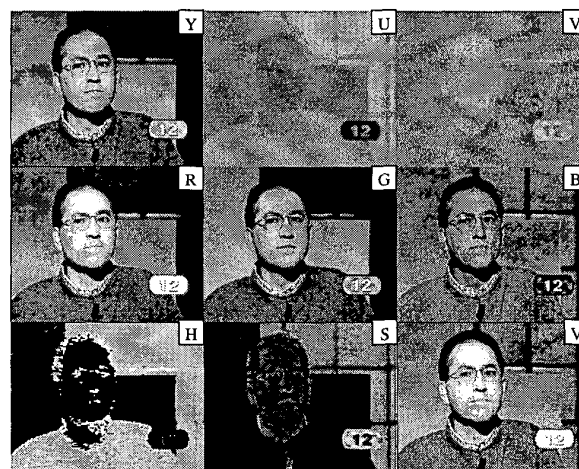

Figure 3. A face image, expressed in different color spaces.

This shows that by adding the color information the recognition rate increases. The general low performance of the system (regardless of using the color information or not) is mainly due to the difficult images used (i.e. non-supervised training and testing of the system) and the morphing that uses only 4 points.
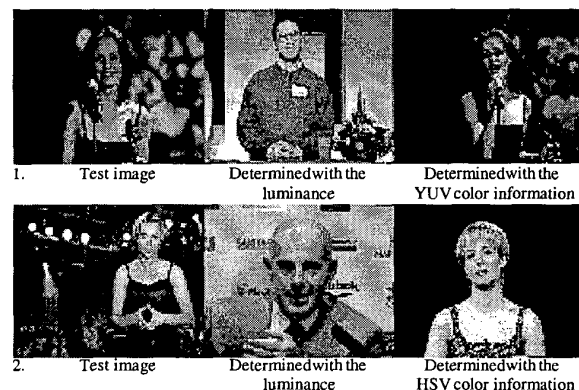

Figure 4. Examples of matches, depending of the used color space.

REFERENCES

[1] L. Lorente, L. Torres, "*A global eigen approach for face recognition*", International Workshop on Very Low Bit-rate Video Coding, Urbana, Illinois, October 8-9, 1998.

[2] *Special Issue on Face and gesture recognition*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 19, No. 7, July 1997.

[3] R. Chellappa, C. L. Wilson, S. Sirohey, *"Human and machine recognition of faces: a survey"*, Proceedings of the IEEE, Volume 83, No. 5, pp. 705-740, May 1995.

[4] ISO/IEC JTC1/SC29/WG11. MPEG Requirements Group. *"MPEG-7: Context and Objectives"*, Doc. ISO/MPEG N2460, October 1998 / Atlantic City, USA.

[5] J. Zhang, Y. Yan, M. Lades, *"Face recognition: eigenface, elastic matching, and neural nets"*, Proceedings of the IEEE, Vol. 85, No. 9, pp. 1423-1435, September 1997.

[6] M. A. Turk, A. P. Pentland, *"Face recognition using eigenfaces"*, Proceedings of the IEEE Computer Society Conf. on Computer Vision and Patter Recognition, pp. 586-591, Maui, Hawaii 1991.

[7] A. Pentland, B. Moghaddam, T. Starner, *"View-based and modular eigenspaces for face recognition"*, MIT Media Laboratory Perceptual Computing Section, Technical Report 245, 1994.

[8] D. Rowland, D. Perret, D. Burt, K. Lee and S. Akamatsu, *"Transforming facial images in 2 and 3-D"*, Imagina 97 Conferences - Actes / Proceedings, Feb. 1997, pp. 159-175.

[9] *"The Candide software package"*, Image Coding Group, Linköping University, 1994.

[10] L. Lorente, L. Torres, "International Conference on Image Processing", ICIP 99, Kobe, Japan, October 99.