EXPERIMENTS IN HUMAN FACE DETECTION AND RECOGNITION

Dorian Cojocaru, Raluca Vidican, Bogdan Prună

University of Craiova, Faculty of Automation, Computers and Electronics, Mechatronics Department,

Abstract: This paper presents an application regarding face recognition. The first task, the detection, is accomplished using color information. The second task, the identification of a detected face, uses the component analysis method called eigenfaces approch. The experiment is implemented in a Matlab aplication which consists in two image processing steps. First the person's head is separated from the background and then the recognition is performed by comparing characteristics of the face to those of known individuals.

Keywords: computer vision, image processing, face detection, face recognition.

1. INTRODUCTION

In the last thirty years the detection and recognition of human faces approches become an important issue, as a result of the development of many applications such as security sistems, credit card verification and crimal recognition. Therefore more and more research has been done in this area. Trying to understand how faces are processed by human brain it was obvios that a computational model of face detection and recognition is very complex and difficult that's why we used only frontal views.

In order to recognize a person in an image, it is first necessary to find the face of each person in that image so that face detection is the first step in the process of face recognition (Mikolajczyk et al., 2001). For the detection of facial regions in color images, several techniques have been proposed so far, using texture, shape and color information. The detection of faces using skin color method should be a good choice due to the fact that color is the most discriminating feature of a facial region. This method first realizes the detection of skin colored regions from a colored image which are then classified as face or non-face images. In the recognition process the output image obtained from the detection is used as an input. Tacking in consideration its speed, simplicity and learning capability, a principal component analysis, that operates using information theory concepts, is a desirable recognition method. The eigenfaces approach searches a computational model that best describes a face, by extracting the most relevant information contained in that face.

2. FACE DETECTION

One of the simplest methods to locate faces in images is to look for oval shaped regions of skin color. Many researchers have developed this technique including Hunke (1994), Yang and Waibel (1996), Wang and Chang (1997), Jones and Rehg (1998), Choudhury *et al*. (1999), Chang and Robles (2000).

1.1 Skin color method

The first step of this method consists in using a skin model in order to detect the skin regions from the backgroung. The next step consists in locating frontal face regions from all skin regions. Within the first step a chromatic map is realized so that it shows regions close to the skin colors. This map is used to transform a color image into a gray scale image. The next operation is the segmentation of the gray image, separating therefore the skin regions from the others.

Those skin regions are then compared with a template face. If the resemblance in close enough the segmented skin region is considered a frontal face image. The goal is to obtain an output image that contains only the head of the subject.

Fig. 1. The original colored image and the result of the face detection part.

The skin color model must ajuste people different kinds of skin and different lighting conditions (Yang and Waibel, 1996). Due to the fact that RGB images contain in addition whit color information also lighting information (luminance (Cai *et al.*)) it is necessary to eliminate those lighting effects by projection in the chromatic color space. The normalization process that defines the chromatic colors (Wyszecki and Styles, 1982; Gong and Sakauchi, 1995) is shown below:

$$r = R / (R+G+B)$$
$$b = B / (R+G+B) \qquad (1)$$

Color green is redundant after the normalization because $r+g+b = 1$.

For creating a skin model in chromatic color space 17 samples from different people images were cropped and then low-pass filtered for noise reduction. Skin color distribution in the chromatic space can be fit in to a Gaussian model by $N(m,C)$:

$$m=E\{x\} \text{ where } x=(r\ b)^T$$
$$C=E\{(x-m)(x-m)^T\} \qquad (2)$$

The Gaussian model allows the likelihood value computation, determinating the regions in the original image that are most likely a skin regions:

$$P(r,b) = \exp[-0.5(x-m)^T C^{-1}(x-m)] \qquad (3)$$

where : $x=(r,b)^T$

To perform the skin segmentation the original color image is transformed into a chromatic image. The likelihood value is calculated for every pixel of the chromatic image, each value coresponding to a gray value of the output gray scale image, the skin likelihood image. The skin segmentated image is obtained by using an adaptive thresholding of the skin likelihood image. After the optimal threshold value is determined by finding the point when the change regarding the number of segmentated images is minimum, the gray scale image is converted to binary image. White segments show skin regions and the black areas show non-skin regions.

Fig. 2: The skin likelihood image and the binary image.

The selected skin areas can be, not only faces, but also any other skin area, so it is necessary to label and evaluate each one of them. It was considered that a face is a closed area that contains at least one hole inside it. The number of holes, *H*, is calculated using Euler number (Ramesh *et al.*, 1995):

$$E = C - H \qquad (3)$$

where *C*, representing the number of connected components, is equal to 1 because we analyze one segmented region at time.

$$H = 1 - E \qquad (4)$$

When the region that contains at least one hole is found, it's necessary to study it by calculating some characteristics such as: the area A, the center of mass, the orientation angle, the width, the height, a ratio of height to width (Cojocaru 2002).

$$\bar{x} = \frac{1}{A} \cdot \sum_{i=1}^{N} \sum_{j=1}^{M} j \cdot I[i,j]$$
$$\bar{y} = \frac{1}{A} \cdot \sum_{i=1}^{N} \sum_{j=1}^{M} i \cdot I[i,j] \qquad (5)$$
$$\Theta = \frac{1}{2} a\tan \frac{b}{a-c} \qquad (6)$$

where:

$$a = \sum_{i=1}^{N} \sum_{j=1}^{M} (x'_{ij})^2 \cdot I[i,j]$$
$$b = 2 \cdot \sum_{i=1}^{N} \sum_{j=1}^{M} (x'_{ij}) \cdot (x'_{ij}) \cdot I[i,j]$$
$$c = \sum_{i=1}^{N} \sum_{j=1}^{M} (y'_{ij})^2 \cdot I[i,j]$$
$$x' = x - \bar{x},\ y' = x - \bar{y} \qquad (7)$$

I is a two-dimensional NxM array representing the region.

Since faces normally have a ratio of height to width about 1, this parameter can be used to establish if the segmented area is a face.

Now, with all this parameters calculated, the template matching can be performed. The template face, an average of 16 different faces taken from men and

women without glasses or facial hair, is determined if the selected segmented skin region is really a face.
To perform the match between the template and the skin region, first the holes from this area have been filled. The template face is resized according to the measurements taken earlier on the segmented image. Based on the height and width, the template face is resized to the dimensions of the segmented skin region so that it can later be overlaped the original image. The template face is rotated to the same theta angle as the segmented region. The center of mass of the segmented skin region is used to place the template face directly in the center of the segmented image. For knowing if the template fits inside the region we compute a correlation value that must be at least 0.6.

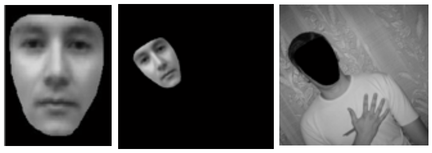


Fig. 3. The template face, the template match and the image after template matching.

Once the algorithm successfully determines that the segmented area is a human face, a rectangle surrounds the face. The system's output is an image that contains only the face.

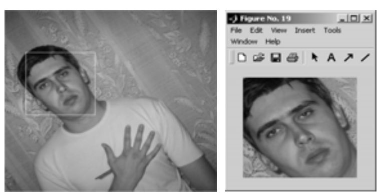


Fig. 4. The detection and the output image

1.2 Results and Discussion

The application was implemented in Matlab Release 13 using the Matlab Image Processing Toolbox. It runs on any type of personal computer that allows the installation of this program. The color input image with a 640x480 pixels resolution is captured with a Canon PowerShot S30 camera.

The color skin approch has among advantages also disadvantages. Some of the advantages are it's insensibility to changes in head scale and orientation and it's speed. The main disadvantage consists in it's sensibility to different ilumination.

In the following examples the first picture is the original color image, the second is the skin likelihood image, the third is the binary image, the fourth is the image after template matching and the fifth is the face detection image.

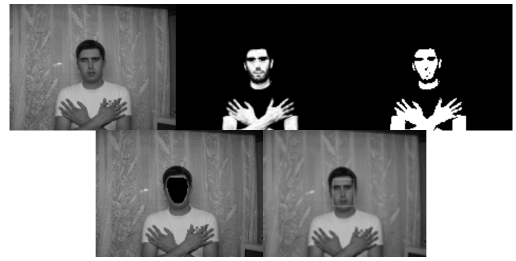


Fig. 5. The detection using neon (white) light.

Fig. 6. The detection using (yellow) houselight.

As a result of our experiments we concluded that using ilumination sources that produce white light, the sistem has better results. This is possible taking in consideration the importance of samples from the image set used in the extraction of the color skin model. In this case those samples were taken from pictures captured with white ilumination conditions. As it is shown in figure 6 the image taken with yellow ilumination, not only misinterpretats some areas as faces but also does not detect the face from the image. One of the situations we often came across in our experiments are images with right and left profile views. Therefore when we calculate the mass center of this regions we observ that this removes from the center of nose (the ideal situation). This removal generates a lightly displaced face detection from the real one. Figure 7 is an example of face detection that also includes the neck because the center of the skin region is located on the left cheek. Another problem interferes in the detection process when the template orientation is incorrect, meaning that theta angle has a wrong value. Another important aspect is the width/height ratio. There are situations when some other objects with the same skin color and containing holes are wrongly detected as faces.

Fig. 7. The detection using daylight and a wrong estimation of the mass center.

Fig. 8. The detection using daylight and a wrong orientation angle

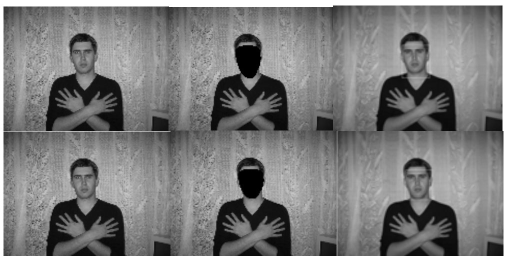


Fig. 9. The detection using the same input image and different ratios of height to width (1.5 and 1.4).

Fig. 10. Example of wrong detection.

Fig. 11. Examples of an accurate detection.

2. FACE RECOGNITION

After the face detection process was done, the face identification can be performed. Many methods of faces recognition were developed and most of them are based on individual features such as the eyes, nose, mouth, and head outline. Those approches involve complex computations and there are very difficult to implement. A more intuitive, flexibile and simple method is the eigenfaces approch propesed by Turk and Pentland (1991).

Face recognition is a type of pattern recognition task where the pattern contains face descriptors. The answer of the system must be a classification "known face n" or "unknown", after comparing it with stored known individuals. It is also desirable to have a system that has the ability of learning to recognize unknown faces.

2.1 Eigenfaces approch

This method consists in extractig and encoding the most important information from a face image and then compare it with a database of models encoded similarly (Turk and Pentland, 1991). Mathematicaly, the eigenvectors of the covariance matrix of a small set of characteristic pictures are sought. These eigenvectors are called eigenfaces due to their resemblence of face images. Recognition is performed by assigning weight vectors to face images, according to their contributions to the face space spanned by the eigenfaces (Atalay, 1996).

The first step of this approach is the formation of a face library that contains face images of known subjects. An image database of 40 individuals including 10 images for each person, with different expressions and illuminations, is created. Form this library a training set: $\Theta_1,\ \Theta_2, \Theta_3, \Theta_4 \ldots \Theta_K$ is choosed.

Fig. 12. Some subjects from the face library

Then, the average face of the set is computed by:

$$\Psi = \frac{1}{K} \cdot \sum_{n=1}^{K} \Theta_n \qquad (8)$$

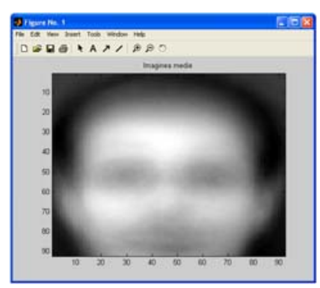


Fig. 13. The main image rezulted from the aplication.

Every image differs from the mean by a vector:

$$\Phi_i = \Theta_i - \Psi \qquad (9)$$

The covariance matrix of the data is thus defined as:

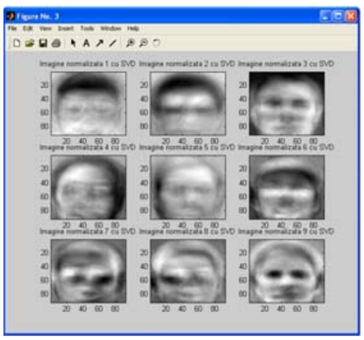
$$C = \frac{1}{K} \cdot \sum_{n=1}^{K} \Phi_n \cdot \Phi_n{}^t = A \cdot A^t \qquad (10)$$

where A is a column-wise concatenation $A=[\Phi_1\ \Phi_2\ ....\ \Phi_K\ ]$. C has dimension MxN x MxN where M is the width of the image and N is the height. The size of this matrix is enormous, but since we only sum up a finite number of image vectors K, the rank of this matrix can not exceed K-1. We note that if v_i are the eigenvectors of $L=A^T A$:

$$A^t A\mathbf{v}_i=\mu_i\mathbf{v}_i \qquad (11)$$

where μ_i are the eigenvalues, then $A\ v_i$ are the eigen-vectors of $A\ A^T = C$ as we see by multiplying on the left by A the previous equation:

$$AA^t A\mathbf{v}_i= \mu_i A\mathbf{v}_i \qquad (12)$$



So defining $\boldsymbol{u}_l$ the eigenvectors of C we have:

$$\boldsymbol{u}_l = A\mathbf{v}_i \qquad (13)$$

With this analysis, the calculations are greatly reduced and become quite manageable.

Fig. 14 Nine of the eigenfaces calculated from the input images.

The next step is to calculate and store a feature vector, for each member of the face database, according to:

$$\Omega^T= [\omega_1\ \omega_2\ .....\ \omega_{M'}] \qquad (14)$$

where ω_i are the weights.

For approximately reconstruct a face image it is necessary to use the feature vector and the eigenfaces as:

$$\Phi\ ' = \Psi + \Phi_f \qquad (15)$$

where $\Phi_f = \sum_{i=1}^{K} \omega_i u_i$ is the projected image.

For the identification of each new face image, the feature vector is calculated and compared with the stored feature vectors of the face library members. The application returns three images, that are most similar to the original imge, and a message. If the subject is in the image database, the message contains the name of the subject, otherwise the message is „The image was not recognized ”.

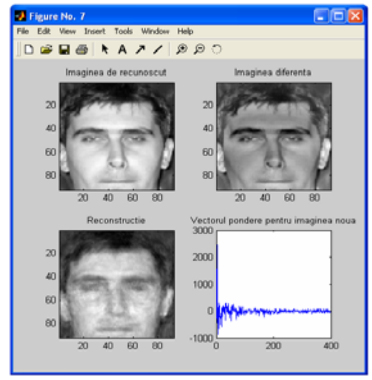


Fig. 15. The reconstruction made by the aplication including the original image, the difference image, the reconstruction of the original image and the plot of weights vector.

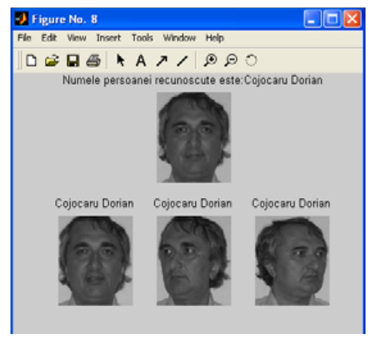


Fig. 16. The result of the regonition process.

2.2 Results and Discussion

The application was implemented also in Matlab Release 13 using the Matlab Image Processing Toolbox. The color input image can be the output of the detection application or it can be an image that contains only the head of the induvidual, also captured with the Canon PowerShot S30 camera. First we made some attempts of reconstructing faces for observing the response of the application when the input image is a new image or an image from the database.

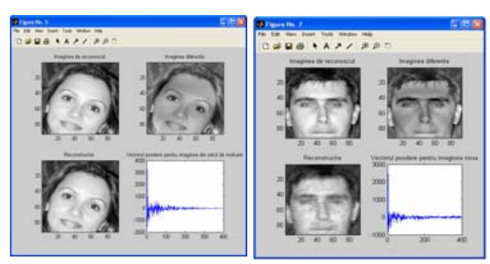


Fig. 17. Examples of reconstruction (first image is from the library and the second is a new image)

The reconstruction using a face from the database is better than the other one even if the first face image is rotated. So we can conlude that the application is insensitive to changes in head orientation.

We also tested the recognition process when features were occluded (sunglasses, beards) and we concluded that the presence of small details is not a problem for the system.

Fig. 18. The accurate recognition of a face with sunglasses.

In figure 18 the top picture is the input image and those below are three similar faces form the database. All images from the library are taken with a black background. After we have changed this aspect the rezults were gratifying.

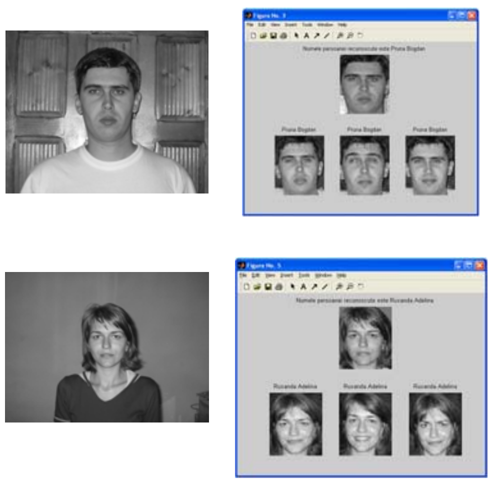


Fig. 19. The original image and the result of recognition.

There are situations when conditions of taken pictures are very different from the initial ones and the result of the recognition is compromised. In the following figure the subject has another hairline, the background is green, and the face is shining because of sunlight.

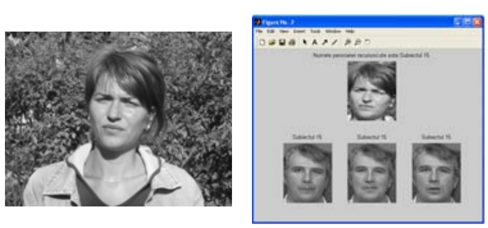


Fig. 20. The original image and the result of recognition.

3. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

The detection program effectively finds instances of frontal faces in color images. From our tests, it is approximately 80% accurate. The program works very well with images with only one person and a solid background. When the background colors are confound with skin region colors, the program encounters some difficulties. Experimental results have shown that changes in illumination and the presence of small details did not cause a major problem to the system but the application is sensitive to scale head.

Beside the recognition the eigenfaces approch may also be used to obtain other features such as sex, race, age and expression (Cottrell and Matcalfe).

The approche's acccuracy can be improved using one of the two aspects: scanner and camera support (because the current system acquires face images only from files located on magnetic mediums) and increasing it's speed.

REFERENCES

Atalay, I. (1996). *Face Recognition Using Eigenfaces*, M. Sc. Thesis, Istanbul Technical University.

Cai, J., A. Goshtasby and C. Yu. *Detecting Human Faces in Color Images*, Wright State University, U. of Illinois.

Cojocaru, D. (2002). *Achiziţia, prelucrarea şi recunoaşterea imaginilor*, Editura Universitaria, Craiova, ISBN 973-8043-146-6.

Cottrell, G. W. and J. Metcalfe, *EMPATH: Face, gender and emotion recognition using holons*, In R.P. Lippman, J. Moody, and D.S. Touretzky (Eds.), *Advances în neural information processing systems*, San Mateo, CA: Morgan Kaufmann.

Gong, Y. and M. Sakauchi (1995). Detection of regions matching specified chromatic features, *Computer Vision and Image Understanding*, **vol. 61**, no. 2, pp 263 – 269.

Mikolajczyk, K., R. Choudhury and C. Schmid (2001). *Face detection in a video sequence - a temporal approach*, Computer Vision and Pattern Recognition

Ramesh, R., R. Kasturi and B. Schunck (1995). *Machine Vision*, pp 31 - 51, McGraw Hill, New York.

Turk, M., and A. Pentland (1991). Eigenfaces for recognition, *Journal of Cognitive Neuroscience*, **Vol. 3**, pp. 71-86,

Wyszecki, G. and W.S. Styles (1982). *Color Science: Concepts and Methods, Quantitative Data and Formulae*, second edition, John Wiley & Sons, New York.

Yang, J and A .Waibel (1996). *A Real-Time Face Tracker*, CMU CS Technical Report.