ADAPTATIVE TECHNIQUES FOR THE HUMAN FACES DETECTION

João Fernando Marar, Danilo Nogueira Costa
*Laboratório de Sistemas Adaptativos e Computação Inteligente (SACI) – Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP) - Campus de Bauru –
Departamento de Computação Bauru – SP – Brasil

Olympio Jose Pinheiro*
Faac – UNESP

Edson Costa de Barros Carvalho Filho*
Cin – UFPE

Keywords: Machine Vision and Image Processing; Adaptive Architecture, Artificial Neural Network, Backpropagation, Face Recognition, Feature Extraction.

Abstract: This paper presents results from an efficient approach to an automatic detection and extraction of human faces from images with any color, texture or objects in background, that consist in find isosceles triangles formed by the eyes and mouth.

1 INTRODUCTION

Recently, there has been a growing interest in automatic identification and in its applicability to diverse types of situations in which personal authentication is necessary.

Systems based on biometric characteristics, such as face, fingerprints, geometry of the hands, iris pattern and others have been studied with attention. Face recognition is a very important of these techniques because through it non-intrusive systems can be created, which means that people can be computationally identified without their knowledge. This way, computers can be an effective tool to search for missing children, suspects or people wanted by the law.

This study presents a system for detection and extraction of faces based on the approach presented in (Lin and Fan, 2001), which consists of finding isosceles triangles in an image, as the mouth and eyes form that geometric figure when linked by lines.

In order for these regions to be determined, the images must be converted into binary images, thus the vertices of the triangles must be found and a rectangle must be cut out around them so that their size can be brought to normal and the area can be fed into a second part of the system that will analyze whether or not it is a real face. Two different approaches are tested here. First, a weighing mask is used to score the region; then, a Backpropagation Artificial Neural Network (ANN)(Cheng *et al*, 2001; Row *et al*, 1996; Row *et al*, 1998), for the analysis to be performed.

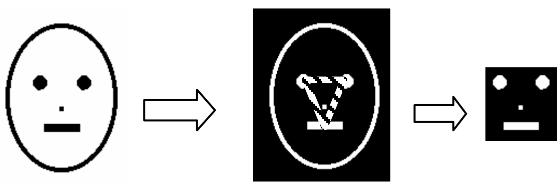


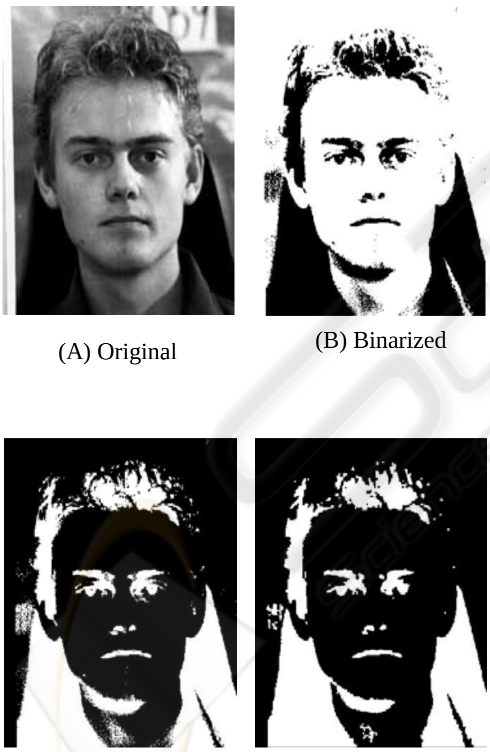
Figure 1: Shows the system basic idea.

Fernando Marar J., Nogueira Costa D., Jose Pinheiro O. and Costa de Barros Carvalho Filho E. (2004).
ADAPTATIVE TECHNIQUES FOR THE HUMAN FACES DETECTION.
In *Proceedings of the Sixth International Conference on Enterprise Information Systems*, pages 465-468
DOI: 10.5220/0002644704650468

2 DETECTION AND SEGMENTATION (1st PART)

First the image was read with the purpose of allocating a matrix in which each cell indicates the level of brightness of the correspondent pixel; then, it is converted into a binary matrix by means of a Threshold parameter T. This stage changes to 1 (white) a brightness level greater than T and to 0 (black), the others (Filho and Neto, 1999). By then, the interest areas – the eyes and the mouth – are shown in black and the skin is shown in white. In order to facilitate the process, the algorithm all 1's is changed to 0's and vice-versa, as shown in Figure 1.

In most of the cases, due to noise and distortion in the input image, the result of the binary transformation can bring a partition image and isolated pixels. Morphologic operations – opening followed by closing – are applied with the purpose of solving or minimizing this problem (Young *et al*, 1998). Figure 2 shows the result of these operations.



(A) Original (B) Binarized

(C) Changing 0 by 1 and 1 by 0 (D) Opening and Closing operation

Figure 2. Image treatment after morphologic operations**.**

After binarization the task is finding the center of three 4-connected components that meet the following characteristics:

I. vertex of an isosceles triangle;
II. distance between the eyes must be 90-100% the distance between the mouth and the central point between the eyes;
III. the triangle base is at the top of the image.

The last restriction does not permit finding upside down faces, but it significantly reduces the number of triangles in each image, thus reducing the processing time to the following stages. These are shown in Table 1.

Table 1: Numbers of triangles in figure 2(D)

With last restriction	Without last restriction
399	769

The opening and closing operations are vital, since it is impossible to determine the triangles without this image treatment. The processing mean time to find the results presented in Table 1 was 4 seconds; on the other hand, 35 hours were insufficient in an attempt at finding the same results to figure 2(C).

3 DETECTION AND SEGMENTATION (2nd PART)

The purpose of this stage is to decide whether a potential face region in an image (the region extracted in the first part of the process) actually contains a face. To perform this verification, two methods were applied. First, a weighing mask was used to bring weight to each region and then Principal Components Analysis (PCA) was used together with Backpropagation ANN. However, prior to that, all regions had to be normalized to the 60 x 60 size by bicubic interpolation, because every potential area needs to present the same amount of information for comparison.

3.1 Mask Generation

The mask was created using 10 images (Figure 3). The first five are pictures of females and the others are pictures of males. All of them were manually segmented, binarized, normalized, morphologically treated (opening and closing) and then the sum of the correspondent cell of each image was stored in the 11^{th} matrix. Finally, that matrix was binarized with another Threshold T, for which values lower than or equal to T were replaced by 0, and the others by 1.

Figure 3: Segmented and binarized faces used to create the mask

The result was improved with T=4. Whereas at lower values the areas of the eyes and mouth become too big, at higher values these areas almost disappear. In both cases, determining the triangles is considerably difficult. Figure 4 shows the constructed mask

Figure 4. Mask image binarized with T =4

3.2 Algorithm for obtaining the weight

The algorithm used to decide whether a potential face contains a real face is based on the idea that the binary image of a face is highly similar to that of the mask.

•Input: Region *R* and mask *M*
•Output: Weight for R
•For all pixels of *R* and *M*
–If the pixel from R and M are white
•Then p = p + 6;
–If the pixel from R and M are black
•Then p = p + 2;
–If the pixel from R is white and from M is black
•Them p = p – 4;
– If the pixel from R is black and from M is white
•Then p = p –2;
•Experimentally: face 3400 <= P >= 6800

3.3 PCA and Artificial Neural Network

In order to obtain better performance, a Backpropagation Network was implemented to analize when a potential face region really contains a face. To make it possible it was necessary to reduce the dimension of the faces. Sixty by sixty pixels give 3600 x 1 dimension vectors, but in an image pixels are highly correlated. Due to redundant information present, the PCA transformation is highly recommended to create a face space that represents all the faces using a small set of components (Campos, 2000; Romdham 1996). In that case, one hundred manually segmented faces (fifty women and fifty men) were used to evaluate the eigenvectors and the eigenvalues. A set of ten eigenvectors with respective ten higher eigenvalues was chosen to create the face space.

A Backpropagation ANN with a 10-input layer, 3 x 2 hidden sigmoid layer and one sigmoid output layer was designed, then it was trained with those one hundred faces used to evaluate the PCA and 40 more non-face random images. All the training set was transformed by face space and normalized before fed in ANN.

3.4 ANN Deciding Algorithm

- Find a potential face in binary image and then extract a correspondent region R in a 8-bit image.
- Project R in a Face space and obtain a 10x1 characteristics vector C and normalize it.
- Feed C in ANN and obtain the weight P.
- If P grater than 0.5 it's a face.

4 RESULTS

Several tests were performed to determine an ideal threshold value for the conversion of the images into binary figures. In a scale from 0 (black) to 1 (white), 0.38 was empirically determined as a good value to most of the images, but to darker images 0.22 was a better value.

An important aspect to be mentioned is that sometimes the system would cut the same face several times into small or big different framings. That is explained by the presence of the eyebrows, blots or noise that "deceive" the program. An algorithm that both verifies if the cut regions are contained into one another and if they share a big intersection area was developed to solve this problem. If that happens, the algorithm chooses the region with the greatest weight, as shown in Figure 5.

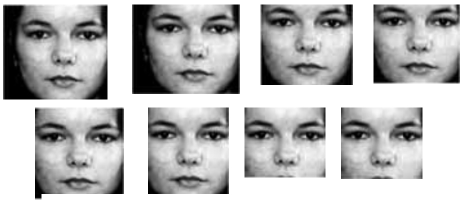


Figure 5: Result without the use of intersection verification

The figure above shows all regions obtained without the intersection verification function. By using that function, the result will be only one figure: the one placed in the first line and the third column.

Confirmation of the efficacy was done through the use of 100 images – 50 male and 50 female – from (PICS, 2003). Each image was used as an input to the developed systems at two different threshold values. The results are shown in Tables 2 and 3.

Table 2: Weighting Mask Results with 2 threshold values

T	0.22	0.38
Processing Mean Time	15.20 seconds	31.62 seconds
Correct detection	81%	48%
False detection	25%	21%

Table 2: ANN Results with 2 threshold values

T	0.22	0.38
Processing Mean Time	12.11 seconds	19.15 seconds
Correct detection	88%	67%
False detection	6%	12%

False detection is characterized by segmentation and high punctuation for a non-face region. In some cases, the correct face was detected and there was false detection to another cut area to the same figure.

The best result for T=0.22 is explained by the low brightness and consequently low contrast of the images in the set. All the images used are at an 8 bit gray scale and 540 x 640 pixels. All tests were performed in an IBM –compatible PC, Athlon 700MHz, 128Mb RAM memory, Windows2000 platform and MATLAB 6 software.

5 CONCLUSION

The system presented in this study is robust and effective, but its efficacy is dependent on optimum threshold value, which depends on illumination levels. That suggests that the system is ideal for use in places where illumination can be controlled, such as shopping malls and airports, because in such places a unique threshold value can be more easily reached for the pictures captured. ANN is faster and reduces significantly false detection..

ACKNOWLEDGMENT

PIBIC/CNPq and Fapesp process 97/13.709-5 financed this work.

REFERENCES

Campos, T. E, *Técnicas de Seleção de Atributos e Classificação para Reconhecimento de Faces*, Msc dissertation, IME- USP, São Paulo, 200.

Cheng,Y.D, O'Toole,A. J, Abdi,H, *Classifying adults' and children faces by sex: computational investigations of subcategorical feature encoding*, Cognitive Science, Elsevier, 2001.

Filho, O. M. and Neto, H. V. *Processamento Digital de Imagens*. Brasport, Rio de Janeiro, 1999.

Kirby,K. and Sirovich, L, *Application of the Karhunen-Loève Procedure for the Characterization of Human Faces*, IEEE Transactions on Pattern AnalysisandMAchine Intelligence,1990.

Lin, C. e Fan, K. *Triangle-based approach to detecion of human face*. Pattern Recogniton Society, Elsevier Science Ltd, 2001.

Psychological Image Collection at Stirling, University of Stirling, Psycology Department, http://pics.psych.stir.ac.uk/, at 02/2003.

Romdham, S, *Face Recognition Using Principal Components Analysis*, http://www.robots.ox.ac.uk/~teo/pca/, at 05/2003

Rowley, H. A.,Baluja,S. and Kanade,T., *Neural Network-Based Face Detection*, IEEE, 1996.

Rowley, H. A.,Baluja,S. and Kanade,T., *Rotation Invariant Neural Network-Based Face Detection*, IEEE, 1998.

Sung, K, and Poggio, T, *Exampled-Based Learning for View-Based Human Face Detection*, IEEE Transactions on Pattern AnalysisandMAchine Intelligence,1998.

Young, I. T., Gerbrands, J. J. e Van Vliet, L. J. *Fundamentals of Image Processing*. Delft University of Tecnology, Netherlands, 1998.