ISSN (e): 2250 – 3005 || Vol, 04 || Issue, 11 || November– 2014 ||
International Journal of Computational Engineering Research (IJCER)

Face Detection for identification of people in Images of Internet

Rodolfo Romero Herrera[1,] Francisco Gallegos Funes[2]
[1]Escuela Superior de Computo (ESCOM) IPN,
[2]Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME) IPN

Abstract:
One method for searching the internet faces in images is proposed by using digital processing topological with descriptors. Location in real time with the development of a database that stores addresses of internet downloaded images, in which the search is done by text, but by finding facial image, is achieved.

Face recognition in images of Internet has proved to be a difficult task, because the images vary considerably depending on viewpoint, illumination, expression, pose, accessories, etc. The descriptors for general information: containing low-level descriptors.

Developments on face recognition systems have improved significantly since the first system; image analysis is a topic on which much emphasis is being given in order to identify parameters, visual features in the image that provide environment data that it is represented in the image, but without the intervention of a person.

In this project raises its realization using the method of viola and jones as face descriptor. We can distinguish even different parts of the face such as eyes, eyebrows, nose and mouth.One method for searching faces in image taken from internet intends to use digital processing using topological descriptors. It is located the face in real time.

Keywords: *Face Recognition, viola and jones, identifying people, Internet, descriptor for points.*

I. INTRODUCTION

The system has a key content manager, who also supplies the database so that the main features or image patterns are stored. The system has a key content manager, who also supplies the database so that the main features or image patterns are stored. The query module applies the same algorithms as the administrator to generate the object description and comparing them with those stored in the database using the Mahalanobis distance [1]. The query module sends the results to the block "Result Set". See figure 1.

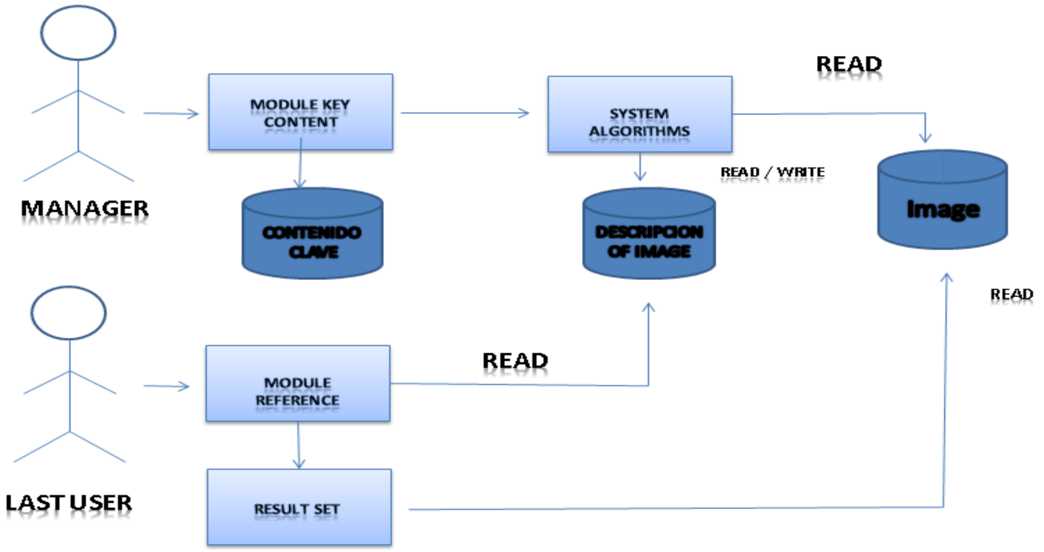


Figure 1. Block diagram system

The shape of an object in an image may consist of a single region or set of regions as shown in the Figure 2:

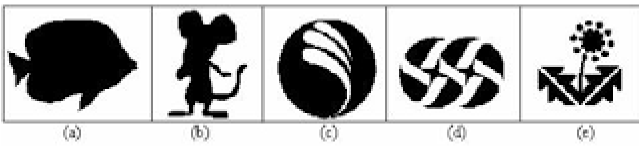


Figure.2 . Examples of images with which they work.

A descriptor can classify them according to the features to be identified, so that we can compare different forms of such images and see if it is the same object or similar objects. The great advantages of this descriptor are its small size and speed [2].

The face recognition is a task that humans perform routinely and effortlessly. So much so that the run almost every day without realizing.The low cost of computers and technological advancement has enabled has a huge interest in automating processes and video image processing. The research worldwide recognition has been motivated by the number of applications that require the identification of the individual.

The problem is based on:

Given an image, the goal of face detection is to identify all image regions which contain a face.

It is difficult to detect faces due to the following parameters [3]:

- Pose
- Image orientation
- Tone of individuals and the background image

II. FACE DETECTOR

The system performs a scan the entire image at different scales for detecting faces in different sizes. See figure 3.The implementation of the detector system consists of two stages:

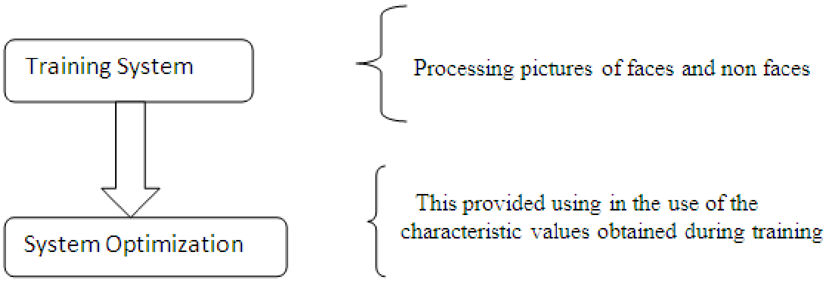


Figure 2 Detector system

The images for training are preprocessed by changing the color image to a grayscale image and then resizing is performed. See Figure 4.

Figure 4 Scaling to gray.

The flow chart of Figure 5 shows the procedure for face detection.

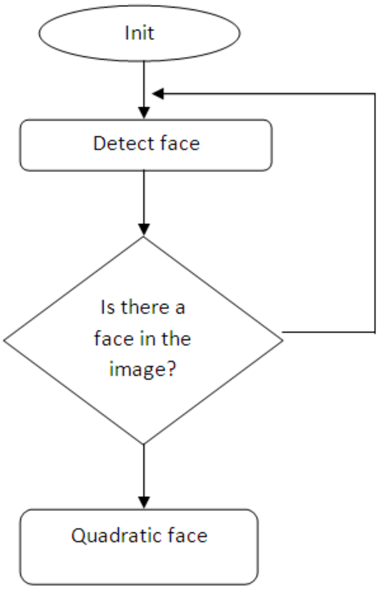


Figure 5 Workflow of face recognition

III. FACE DETECTION AND LOCATION

A technique proposed by Paul Viola and Michael J. Jones was analyzed. [4]. This method is basically the face location by using a group of rectangular features, which operates faster than a pixel based on system [4]. The two, three and four boxes: three kinds of features are used. Said rectangular features (CRs) are 10 different types shown in Figure 6.

How you use these CRsr, is put in a certain position within the image and calculate the difference between the sum of the pixels within the clear part and the dark part of the CR, obtaining an integer value that must exceed a threshold for a finding that is on facial feature is desired. The training phase is to find the CRs in a certain scale and position within a window of a certain size located a characteristic feature of a face, this CR is called classifier. In Figure 7 can be seen a set of classifiers in the classifier where only figure 4d locate a feature of the face.

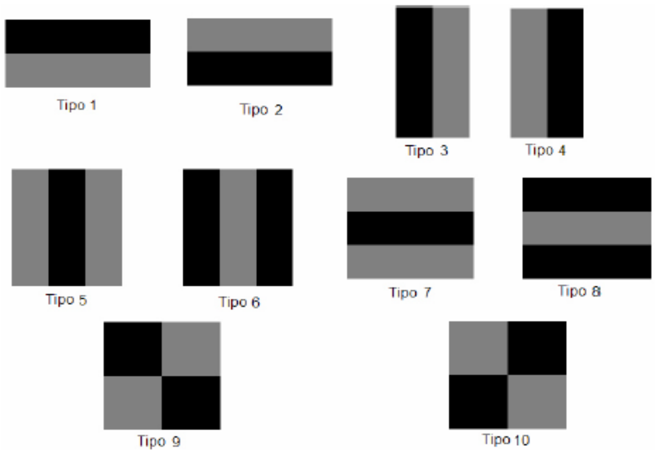


Figure 6 Set of ten CRs used in the process of locating faces.

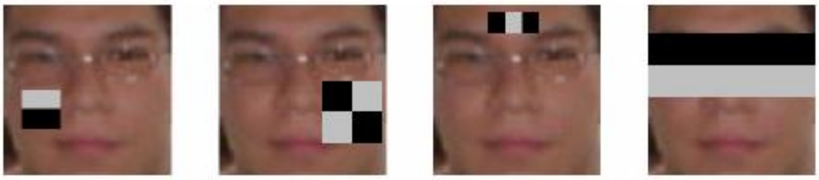


Figure 7. Four rectangular features from a) to d) which only d) is a useful feature that can be considered how classifier.

The method of Viola & Jones is to train or find a set of classifiers, and thresholds, testing each of the classifiers in each position and scale of a set of test images used for this, the learning algorithm AdaBoost [5].This set consists of a group of images in which faces and a standard set of pictures in which there are faces; classifiers with the highest number of correct classification of the faces and the lowest number of false positives obtained are selected.

To calculate the CRs rapidly at various scales an intermediate image representation (integral image) is used. Once obtained the CR is calculated at any scale and location in constant time. The integral image at a position x, y contains the sum of the pixels above and to the left of x, y. Equation 1.

$$ii(x,y) = \sum_{x' \leq x, y \leq y} i(x', y') \tag{1}$$

Where ii (x, y) is the integral image and i (x, y) is the original image (Figure 8). Using the following pair of recurrences:

$$S(x,y) = S\,(x,y-1) + i(x,y) \tag{2}$$
$$ii(x,y) = ii(x-1,y) + s(x,y)$$

Where s (x, y) is the cumulative sum of the values of the pixels in the row, s (x, -1) = 0, and ii (-1, y) = 0), the integral image is calculated in a single pass over the original image.

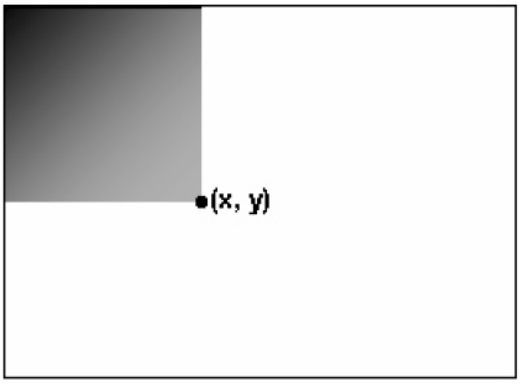


Figure 8 The value of an integral image at a point (x, y) is the sum of all pixels above and to the left.

Using the integral image any amount of a rectangle is calculated using an array of four references (see Figure 9). Also the difference between two sums of rectangles is computed eight references. As the characteristics of two rectangles (types 1, 2, 3 and 4) defined above have adjacent rectangular sums calculated with an arrangement of six references, eight for the characteristics of three rectangles and nine for four rectangles.

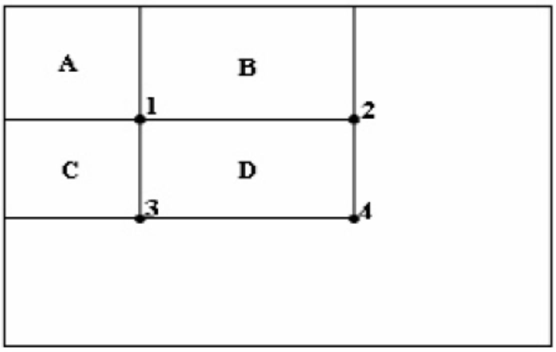


Figure .9 Sum of pixels.

The sum of the pixels within the rectangle D can be calculated with an arrangement of four references. The value of the integral image at the position 1 is the sum of the pixels in the rectangle A. The value at position 2 is A + B, in the 3 position is A + C, in position 4 is A + B + C + D. The sum within D can be calculated as 4 + 1 - (2 + 3).

IV. MAHALANOBIS DISTANCE

In statistics, Mahalanobis distance is a distance measure introduced by Mahalanobis in 1936 Its utility is that it is a way to determine the similarity between two multi-dimensional random variables. It differs from Euclidean distance in that takes into account the correlation between random variables [6].

Distance is called the length of the shortest path between two entities. From a formal point of view, for a set of elements X, distance is defined as any binary function d (a, b) of X xX in $\mathbb{R}$ verifying the following conditions:

No negativity:

$$d(a,b) \geq 0 \; \forall \; a,b \; \in X \tag{3}$$

Symmetry:

$$d(a,b) = d(b,a) \forall \; a,b \; \in X \tag{4}$$

Triangle inequality:

$$d(a,b) \leq d(a,c) + d(c,b) \forall \; a,b \; \in X \tag{5}$$

It called Euclidean distance between two points A (x, y) 11 and B (x, y) 22 to the length of the line segment whose endpoints A and B. It is expressed mathematically as [7]:

$$d = \sqrt{(X_2 - X_1)^2 + (y_2 - y_1)^2} \tag{6}$$

The distance between a point P and a line R is the length of the shortest path connecting point P (x, y) 11 with the line R: Ax + By + C = 0 is expressed mathematically as:

$$d = \frac{AX_1 + B\,y_1 + C}{\sqrt{A^2 + B^2}} \tag{7}$$

The distance between two parallel lines is the length of the shortest path between them and any point of the other. The distance between a point P and L plane is the length of the shortest path between P (x1, y1, z1) point and the plane L=Ax + By + Cz + D. Mathematically it is expressed as:

$$d = \frac{AX_1 + BY_1 + CZ_1 + D}{\sqrt{A^2 + B^2 + C^2}} \tag{8}$$

Formally, the Mahalanobis distance between two random variables with the same probability distribution $\vec{x}$ and $\vec{y}$ with covariance matrix Σ is defined as:

$$d_m(\vec{X},\vec{Y}) = \sqrt{(\vec{X} - \vec{Y})^T \sum^{-1} (\vec{X} - \vec{y})} \tag{9}$$

V. REGIONS

For detection; once having the position of the face within the imageis done an estimate of where the eyebrows are located, nose and mouth. This position can be estimated, due to the method of locating faces, wherein the facial parts near the same position will be obtained. (See Figure 10). Subsequently, processes each of these areas with special image processing for each region [8].

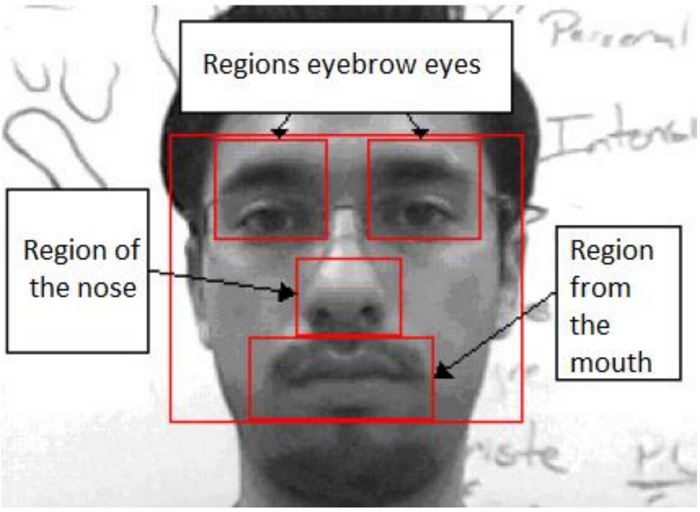


Figure .10 Location eyebrows eyes, nose, and mouth.

VI. Eyebrows Detection

For detection of eyebrows, once estimated position within the facial area in an area of 45 x 40 and the grayscale image as that of Figure 11a proceed to make the following image processing:

1. Equalization image eyebrow area is performed; Figure 11b.
2. The procedure to perform a sum of equalized image to itself; figure 11c
3. Applies an adaptive enhanced with a 3x3 window with values of k1 = 1 and k2 = 1; Figure 11d.
4. Then proceed to perform binarization with a threshold of 40; Figure 11e.
5. Proceeds to convex components labeling; Figure 11f.
6. The largest component of the top of the area is selected and the rest is eliminated, thereby generating the image of the eyebrow; Figure 11g.

Figure 11 Processing eye.

The next step is to locate the points with the Xs highest and lowest brow, obtaining the UA A and B to the right eyebrow (Figure 12) and C and D for the left eyebrow.

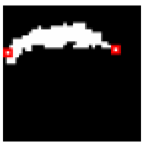


Figure 12 Action Units A and B located on the right eyebrow.

VII. DETECTION OF MOUTH.

For detection of the mouth, as eyebrows, once it has been estimated position in an area of 74x33, and the image in grayscale (Figure 13a) is determined, the image processing that described below:

1. Subtraction is performed between the image and its causes enhanced adaptive; Figure 13b.
2. Subtraction is performed between the image obtained in step 1 and the image of the mouth and is binarized equalized to a threshold of 40 (Figure 13c).
3. Convex component labeling is performed; Figure 13d.
4. The largest component is chosen; Figure 13e.

Figure 13 Getting mouth.

The next step is to locate the points with theXs highest and lowest of the mouth, obtaining action units F and G (Figure 14).

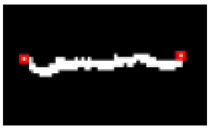


Figure 14 Action Units F and G obtained with the digital image processing of the mouth.

After obtaining the red spots on the images to the mouth, nose, eyebrows and eyes, Mahalanobist distance is applied to determine how closely a face over another, suggesting a threshold. The distance between red dots is calculated and also determines how far (in terms similar) is an image of another.

VIII. ANALYSIS OF RESULTS

Finally in the center point of the nose (UA E) is positioned in the center of the window in the region of the nose. This is because it is only used as a benchmark to measure the distances shown; which are the basis for analysis of the expression of the user.

Figure 15A low Light and without glasses

With the same light that the previous experiment with new lenses were found manner that does not affect the use of the lens system Figure 16.

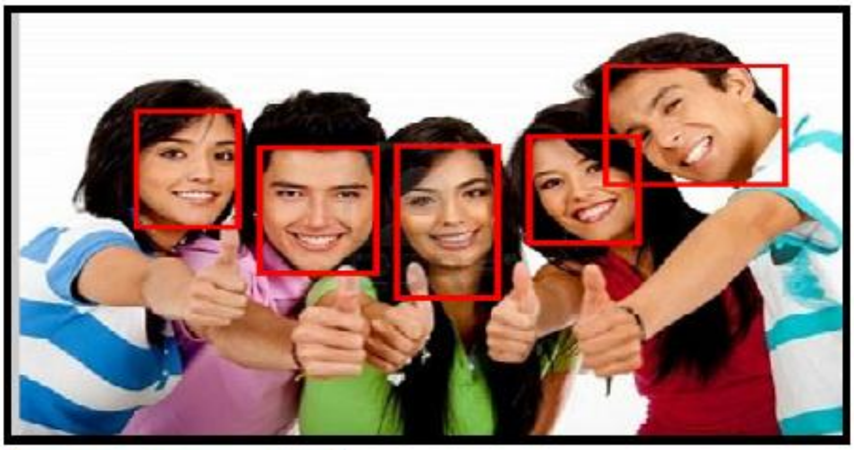
Figure 16- A low Light and with glasses

By using brightness of incandescent without a focus lenses or eyeglasses in figure 16. Figure 17 can be seen giving more points that do not belong to the characteristics of the face.

Figure 17 Incandescent light without glasses

Having that was found Cloudy with a face, an appropriate threshold equal to 2, with or without glasses and with a white fund plain background. Now the question was how the system would affect the brightness which makes testing varying brightness. In conducting the experiment in low light without glasses and the system could find features of the face but showing items that are not part of the features of the face, figure 18.

The faces of the individuals are located and red squares are drawn on their faces. For Figure 14 shows the detection of 5 faces in a picture of Internet.



```
>> Personas

numCaras =

     5
```

Figure 16 Location of five faces.

The method works for the location of one or more face as seen in Figures 19.

Figure 19 Location of a face.

IX. CONCLUSIONS

Face Detection efficiency depends on the lighting that you have when this image was taken. But there was no trouble locating the face even if the person has glasses, moustache, etc.

As for the location of faces in images from the Internet, there is also no problem, even if the image has more than one face.

The location of faces in internet pages without using text, we believe it is important because most search engines do not perform that search this manner; and is especial as well that you have a new option especially for those who cannot accomplish this using a keyboard or text.

The proposed method can be added more easily descriptors and thus enhance the search. Use for example those that are invariant objects, count the number of holes or the Euler number.

ACKNOWLEDGEMENTS

The authors acknowledge the support received from the Instituto Politécnico Nacional and the Escuela Superior de Computo.

REFERENCES

[1]. María Teresa Escobedo Portillo, Jorge A. Salas Plata MendozaP. CH. Mahalanobis y las aplicaciones de su distancia estadística; CULCyT, 2008.

[2]. Lira Chávez, J.. Introducción al tratamiento digital de imágenes. Ciencia de la Computación, InstitutoPolitécnico Nacional, Universidad Nacional Autónoma de México.2002.

[3]. Guevara, M. L., Echeverry Correa, J. D., & Ardila Urueña, W. (2008). Detección de rostros en imágenes digitales usando clasificadores en cascada. Scientia.

[4]. VIOLA, Paul y JONES, Michael "Robust Real-Time Object Detection" En International Conference on Statistical and Computational Theories of Vision – Modeling, Learning, Computing and Sampling (2001, Vancouver, Canada). Memorias. Canada.

[5]. FREUND Yoav, y Robert E., SCHAPIRE. "A decision-theoretic generalization of on-line learning and an application to boosting". En Computational Learning Theory: Eurocolt '95, pp. 23–37. Springer- Verlag, 1995

[6]. Mahalanobis, P. C. .On the generalized distance in statistics.Proceedings of the National Institute of Sciences (Calcutta), 2, 49-55, 1936.

[7]. Danielsson, P. E..Euclidean distance mapping. Computer Graphics and image processing, 14(3), 227-248.1980

[8]. PANTIC, Maja y ROTHKRANTZ, León M. "An Expert System for Recognition of Facial Actions and their Intensity". En Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence (2000, Austin, Texas).Memorias.EUA.ed. AAAI Press / The MIT Press. 2000. pp. 1026 – 1033.A. Blake and M. Isard. Active Contours.Springer, 1998.