Computer Engineering and Applications Vol. 7, No. 1, February 2018

Faces and eyes Detection in Digital Images Using Cascade Classifiers

Alhamzawi Hussein Ali Mezher

Faculty of Informatics, University of Debrecen, Debrecen, Hungary
Aliov.mezher@gmail.com

ABSTRACT

In this article is presented a way to implementation and detect the face and eyes on digital image, based on Haar-like features extraction and cascade classifier, these techniques used in 100 % and 92% for faces and eyes detection respectively for the best all cases using low processing time, we used cheap equipment in our work (Acer Travel Mate web camera). Open CV library (computer vision library) and Python language used in this work.

Keywords*:* Face Detect, Integral Image, Haar-like Features, Boosting.

1. INTRODUCTION

The human face detection is an image processing and computer vision technique to determine the existence or otherwise of faces in a particular image. Initially the problem of face detection in the recognition systems did not receive the necessary attention and it was assumed that the face had already been detected, it was only in the eighties that the first algorithms emerged, based on heuristic and anthropometric techniques, and in the nineties when the development of face detection algorithms began its growth[1], proposing a great variety of techniques, from basic edge detection algorithms to high-level composite algorithms that use advanced pattern recognition methods.

These detection techniques have been approached from different approaches: Approaches based on facial features or local characteristics, in which we look for certain elements that make up the face, like the eyes, the nose, the mouth, Holistic or image-based approaches, in this case the methods work with the image complete or concrete areas of the same from which can extract characteristics that can represent the object searched, Hybrid approaches, these methods use both local and global information for detection, based on the fact that the perception system human distinguishes both local characteristics and global face [1, 2].

Around the approaches what we mentioned, different works have been proposed, such as those of [3, 4, 5], in which the skin color information is used to perform the detection, obtaining results around 90%, [6] uses neural networks to segment the face, reaching detection rate between 77.9% and 90.3% for different network configurations, [1] uses a Haar base for the extraction of characteristics and AdaBoost for the selection of these and classification, reaching a detection rate of 94.1%, this method proposed by Paul Viola and Michael Jones, It is one of the methods used nowadays because it allowed segmenting multiple faces in an image with low processing times. Around this research other works have been carried out, such as those of [8, 9, 10, 11], They have applied the method to segment faces and have added other stages like the detection of Facial features (eyes and mouth), pose estimation and face tracking. This paper presents the implementation of a frontal faces detection technique as the initial stage of an automatic system for the recognition of emotions from the analysis of movement and deformation of the face.

2. METHODOLOGY

The methodology was based on the proposal in [7]. This is divided into three stages as shown in Figure 1; in the first, a transformation of the image is made generating a new call integral image, in the second stage the extraction of characteristics is carried out using Haar-based filters, and finally, boosting is used for the construction of cascade classifiers.

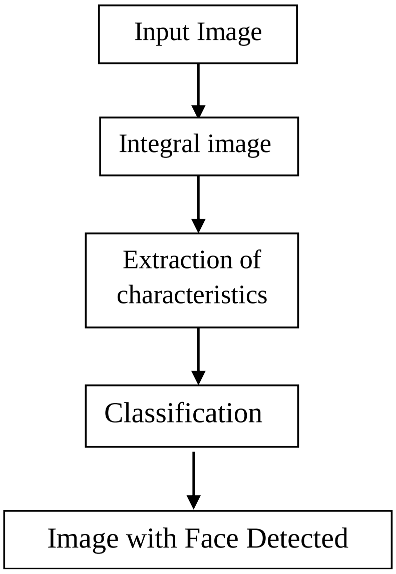

FIGURE 1. System diagram

3. CONCEPTUAL FRAMEWORK

3.1 INTEGRAL IMAGE

This new representation of an image was introduced in [7]. This image allows extracting quickly characteristics at different scales because it does not work directly with the intensity values if not with a cumulative image that is constructed from basic operations.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

FIGURE 2. Integral image

The integral image (see Figure 2), at the location x, y, contains the sum of the pixels from the top left of the image and can be calculated as indicated as follow:

$$II(x, y) = \sum_{\substack{x0 \leq x \\ y0 \leq y}} Im(x`, y`) \qquad (1)$$

where II ***(x, y)*** is the integral image and Im ***(x, y)*** is the original image.

3.2 FEATURE EXTRACTION

In images, the characteristics of each object are extracted by applying certain functions that allow the representation and description of the objects of interest of the image (patterns). The extraction of characteristics is a step in the recognition of patterns in which the measurements or observations are processed to find to tributes that can be used to assign the objects to a certain class [12]. In the methodology followed, the extraction of features is done by applying filters to the image with Haar bases.

These filters can be calculated efficiently on the integral image, they are selective in frequency and spatial orientation, and they allow to be modified in scale and orientation. In Figure 3, some of the filters used for the extraction of characteristics are shown. Filters with Haar bases, performs an encoding of difference of intensities in the image, generating contour features, points and lines, by capturing contrast between regions.

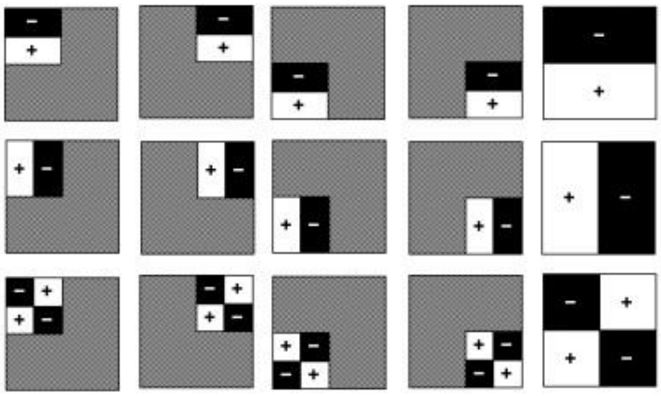

FIGURE 3. Haar filters rotated, moved and with changes in scale

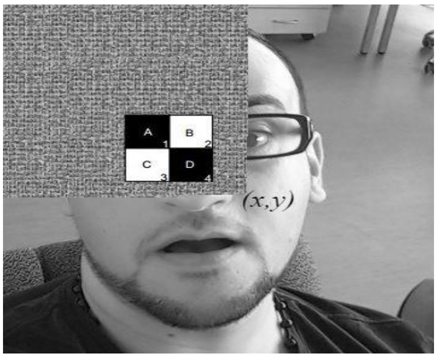

FIGURE 4. Haar filter convolution with an image integral

In Figure 4, The convolution of a Haar filter with the integral image is shown. From this operation, a characteristic can be extracted in a constant time on the integral image by adding and subtracting the values of the vertices for each rectangle. For greater clarity, in the figure the sum of the pixels that form the rectangle D can be calculated as,

$$SumD = (4 + 1) - (2 + 3) \quad (2)$$

where 1, 2, 3, 4, are the values given in the image integrated at such locations.

3.3 CLASSIFICATION

This step in the detection algorithm handles to assign a set of characteristics given to a class with which there is a greater similarity, according to an induced model during training [13]. Boosting was introduced by [13], This is a classification method that combines several basic classifiers to form a single more complex and accurate classifier. The idea is based on the claim that several simple classifiers, each one with a precision slightly higher than a random classification, can be combined to form a classifier of greater precision, as long as it has a sufficient number of training samples. The application of cascade classifiers has allowed getting good results, these are reported in [7], [8], [14]. In Figure 5, a scheme of a cascade classifier is shown.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

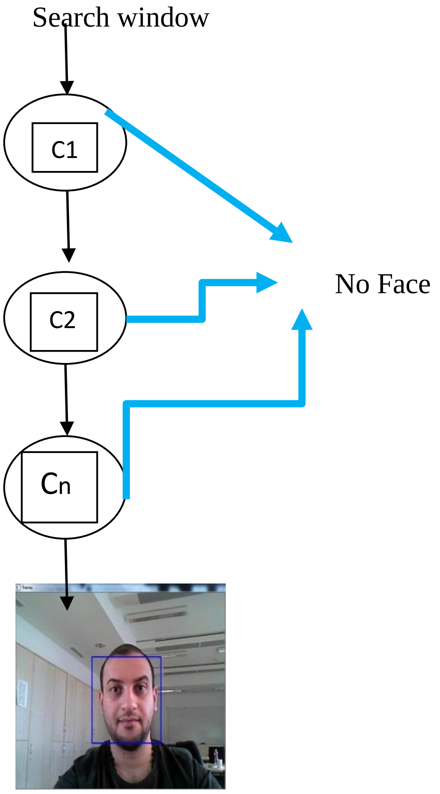

FIGURE 5. A scheme of a cascade classifier

To apply the technique of boosting must first be establish a simple learning algorithm (weak or base classifier), which will be called repeatedly to create various base classifiers. For the training of the base, classifiers are used, in each iteration, a different subset of training samples and a different weight distribution on the training samples [13]. Finally, these base classifiers are combined into a single classifier that is expected to be much more accurate than either of the base classifiers separately. Depending on the base classifiers that are used, the distributions that are used to train them and how to combine them, Different classes of the genetic algorithm of boosting can be created. The boosting algorithm used by Viola and Jones in their work is known as AdaBoost. The algorithm used in the methodology is shown below. This is found in [13].

Given a set of images (x_1,y_1), …. (x_n,y_n), where y_i=0,1 for negative and positive samples respectively.

Initialize the weights $w_{1,I} = \frac{1}{2m}, \frac{1}{2l}$ for y_i=0,1 where m is the number of negative samples and l is the number of positive samples.

For $t = 1,. . . ,T$

Normalize the weights $w_{t,I} \leftarrow \frac{w_{t,i}}{\sum_{i=0}^{n} w_{t,i}}$

Select the best base classifier with respect to error weight: $\epsilon_t = \min \sum_i wi\ |h(x_i,f,P,\Theta)-y_i| f,P,\Theta$

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Define $h_t = h(x, f_t, P_t, \Theta_t)$ where f_t, P_t, Θ_t are used to minimize ϵt

Update Weights: $w_{t+1,i} = w_{t,i}\, \beta^{1-e_i}{}_t$ where $e_i=0$
if the sample xi is correctly classified or $e_i=1$ in another case
with $\beta_t = \frac{\epsilon t}{1-\epsilon t}$
The final robust classifier is:

$C(x)=1 \sum_{t=1}^{T} \alpha t\, ht(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha t$

C(x)=0 in another case

4. TESTS AND RESULTS

The test was performed on 50 images belonging to two people from the Ph.D. room at the University of Debrecen in Hungary. Each photo was in 8-bit JPEG format with a size of 320x240 pixels. The face detection method was applied to each image and the processing time per image was measured. The results are shown below,

TABLE 1.
Results of face detection

	Person 1	Person 2
No image	24	24
No Detection	25	26
No of faults	1	0
% detection	98%	100%
Average detection time	39.24ms	37.44ms

The detection rate is calculated in this article as the ratio of the number of faces detected about the number of used images.

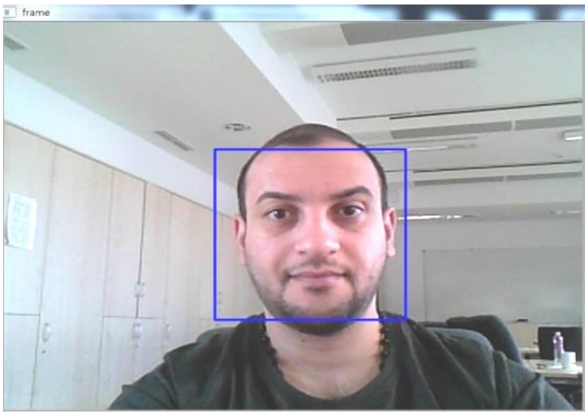

FIGURE 6. Example of face detection for person 1

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

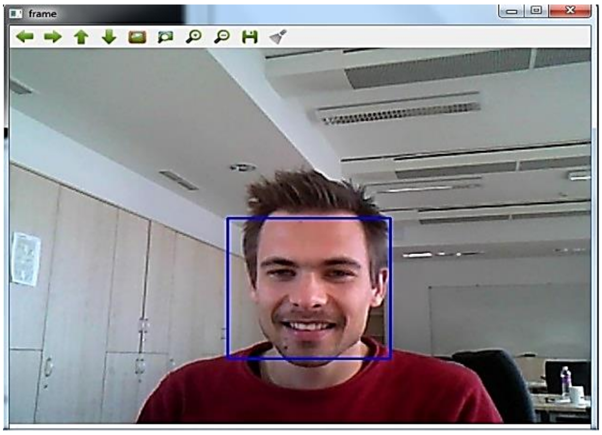

FIGURE 7. Example of face detection for person 2

Given that the project focuses on the recognition of facial expressions and for this is It is necessary to focus the analysis on specific regions of the face like the eyes and mouth, and seeing the versatility of this algorithm and its possible extension to the detection of other regions, this was trained to detect characteristics of the face (eyes). The results obtained for the detection of the eyes are shown below:

TABLE 2.
Eye detection results

	Person1	Person2
No image	24	24
No Detection	23	24
No of faults	2	1
% detection	87%	91%
Average detection time	8.88ms	8.65ms

FIGURE 8. Example of eye detection for person 1

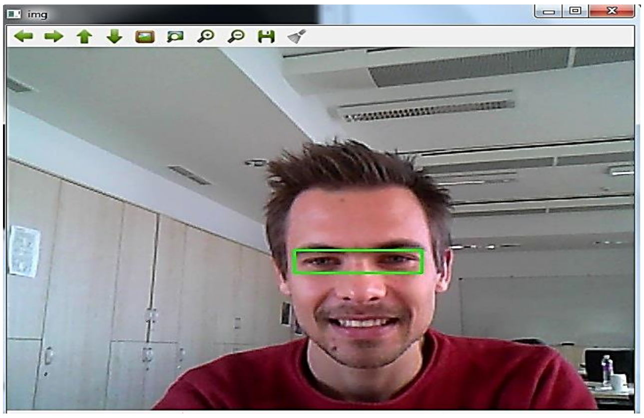

FIGURE 9. Example of eye detection for the person 2

5. CONCLUSIONS

High detection rates and low time processing demonstrate the effectiveness of the combination of boosting techniques with base filters Haar, for face detection and features Facials and give the possibility to take the system to a Real-time application. The reduction of the performance of the method in the detection of eyes is attributed to the algorithm must perform the search on the whole image. This can be improved by applying the method only on the region that has been detected as face, since background information was being downloaded, and also the processing time is decreased.

ACKNOWLEDGMENT

The work of H. AlHAMZAWI was supported by the Stipendium Hungaricum Scholarship.

REFERENCES

[1] Grenier, Ashley Lynn. Conceptual understanding and the use of hand-sketching in mechanical engineering design. *ProQuest*, Pp 57, 2008.
[2] Lapo Governi, Rocco Furferi, Matteo Palai, Yary Volpe, “3D geometry reconstruction from orthographic views: A method based on 3D image processing and data fitting” *Computers in Industry*, Volume 64, Issue 9, Pages 1290-1300, ISSN 0166-3615, December 2013.
[3] http://dx.doi.org/10.1016/j.compind.02.003, 2013.
[4] Passarella, Rossi, Astri Agustina, Sutarno Sutarno, Kemahyanto Exaudi, and Junkani Junkani. "Design Concept of State-Chart Method Application through Robot Motion Equipped with Webcam Features as E-Learning Media for Children.", 115-120, 2013.
[5] Shipp, John I. "LED illumination system for endoscopic cameras." U.S. Patent No. 6,449,006, Sep 2002.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

[6] Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." Automatica 11. 285-296, 23-27, 1975.
[7] Rianmora, Suchada, and Pisut Koomsap. "Recommended slicing positions for adaptive direct slicing by image processing technique." The International Journal of Advanced Manufacturing Technology 46.9-12, 1021-1033, 2010.
[8] Ramirez, Carlos Alejo, Perez, Manuel David. Face detection and analysis of facial expressions.
[9] Lecumberry, R., Federico. Calculation of disparity and segmentation of objects in video sequences. Master's thesis in electrical engineering. University of the Republic Montevideo, Uruguay. August 3, 2005
[10] LI, Yadong, Goshtasby, Ardeshir y Garcia, Oscar. Detecting and tracking human faces in videos. Wright State University, 2000.
[11] Feris, Rogério, Emidio De Campos, Teófilo y Marcondes, César. Detection and tracking of facial features in video sequences. Lecture Notes in Artificial Intelligence, Vol. 1793, pp. 197-206, 2000.
[12] Habili, Nariman. Automatic segmentation of the face and hands in sign language video sequences. Technical report, Adelaide University, Australia, 2001.
[13] Rowley, Henry, Baluja, Shumeet, and Kanade, Takeo. Neural network-based face detection. IEEE, 1998.
[14] Viola, Paul y Jones, Michael. Robust Real-Time Face Detection. International Journal of Computer Vision, Vol.57, No.2, pp 137-154, 2004.
[15] Castrillon, M., Deniz, O., Guerra, C., and Hernandez, M. "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams". Elsevier, Journal Visual Communication and Image Representation, 2007.
[16] Bartlett, Marian Stewart et all. Real time face detection and facial expression recognition: development and applications to human computer interaction. University of California, USA, 2004.
[17] Silapachote, P, Karuppiah, D y Hanson, A. Feature selecting using adaboost for face expression recognition. University of Massachusetts Amherst, USA.
[18] Belaroussi, Rachid and Milgram, Maurice. Face detecting and skin color based tracking: a comparative study. University Pierre an Maire Curie. Paris, Francia, 2007.
[19] IEEE Standar Glossary for Image Processing and Pattern recognition Terminology. Published by the Institute of Electrical and Electronics Engineers, New York, USA, 1990.
[20] Schapire, R and Freund, Y. A decision theoretic generalization of on-line learning and application to boosting. AT&T Bell Laboratories. USA, 1995.
[21] Papageorgiou, C., Oren, M. and Poggio, T. Ageneral frame-work for object detection. Proceedings of International Conference on Computer Vision, Bombay, India, 1998

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)