SOCIETY FOR SCIENCE AND EDUCATION UNITED KINGDOM

TMLAI TRANSACTIONS ON MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

VOLUME 5 NO 4, 2017
ISSN 2054-7390

Comparative Study of Harris and Active Contour Using Viola-Jones Algorithm for Facial Landmarks Detection

Manal El Rhazi, Arsalane Zarghili, Aicha Majda
Intelligent Systems and Applications Laboratory, Faculty of Science and Technology
Sidi Mohamed Ben Abdellah University, Fez, Morocco
manal.elrhazi@usmba.ac.ma; arsalane.zarghili@usmba.ac.ma; aicha.majda@usmba.ac.ma

ABSTRACT

In this paper, we present a comparative study of two methods: Harris corners detector (H) and Active Contour detector (A.C) using Viola-Jones algorithm (V.J) for facial landmarks (eyes, nose, mouth) corners detection. These methods were implemented on two face databases; ECVP and FEI databases with combination of two methods (V.J + H) and (V.J + A.C). Experimental results showed that (V.J +A.C) gives a higher rate of detection than (V.J + H) method.

Keywords—Harris corner detector; Snake; facial features; Viola and Jones; image processing.

1 Introduction

The detection of facial features (eyes, nose and mouth) is the most important initial step in various facial image interpretation works related to computer vision application, such as: face identification, facial expression recognition, facial features tracking, facial beauty analysis…, and for more analysis accuracy, this detection includes the localization of facial features corners (eyes corners, mouth corners and nose corners).

The problem of features corners detection with accuracy and robustness has been a challenging task and has received a lot of attention in the past decade. So many approaches have been proposed in the literature to extract the facial features from an image and their corners; first for features localization, Viola and Jones method [1] has been widely used in many applications to detect the face and the facial features zones on an image. This method is the most stable and has shown a successful result in several applications, due to its basis which consists of many contributions: the cascade of classifiers trained by Adaboost [2]; for each stage in the cascade, a larger set of features is chosen based on the Adaboost. In addition, Viola and Jones method uses the integral image that helps in the speed feature evaluation [3]. Finally, the Viola and Jones is a powerful method to avoid the detection of regions that do not contain the interest object.

A large margin of approaches has been proposed recently for facial corners detection [4], for example, Harris corner detection [5] which is the most used in that field thanks to its facility of use and its good result of detection dealing with different face position and different level of luminosity. Another example is the active contours (or snakes) introduced by Kass and Witkin [6], they are curves that can deform progressively in order to be close to the outline of an object. This deformation is guided by minimizing an

DOI: 10.14738/tmlai.54.3196
Publication Date: 15th August, 2017
URL: http://dx.doi.org/10.14738/tmlai.54.3196

energy function comprising two terms: Eint an internal energy which helps regulate the outline and an external energy Eext relied to the image and to specific constraints that can be added.

In the present work, we have used the active contour and the Harris corner detector combined with the facial axes that go along the features detected to improve the corners detection.

The purpose of this paper is to compare the rate of detection for both the Harris corner detector and the Active contour (Snake) using the Viola and Jones method, in order to detect the facial landmarks. The experiments were implemented on two face databases: FEI database [7] and ECVP database [8].

The global architecture for corners detection presented in this work is shown in Figure. 1

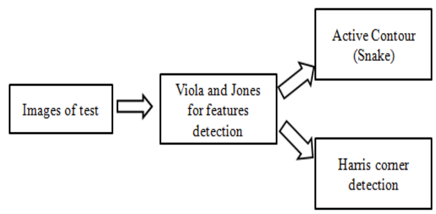

Figure 1. The global architecture for corners detection

This paper is organized as follow: section 2 explains the facial features detection using Viola and Jones, section 3 describes the features corners detection using the two approaches, section 4 will present the experimental results and section 5 contains the conclusion of this paper.

2 Facial features detection

2.1 Face detection

Before we detect the corners, face and facial features must be detected. In this paper we will use the Viola-Jones Adaboost detection method improved by R. Lienhart [9] to detect the face regions in global then eyes, nose and mouth.

Face detection is obtained by delimiting the area of interest by a rectangle based on Haar like features and cascade of classifiers learnt to detect a face zone on an image.

The figure below presents the rectangles using in the detection by the extended Haar like features:

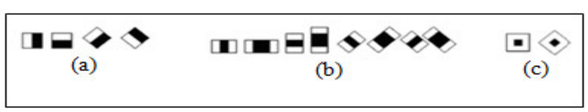

Figure 2. Extended Haar like features: (a) edge features, (b) line features, (c) center surround features

These features are used to calculate the difference between the sum of the white pixels areas and the sum of the black pixels areas.

$$d = \sum (white\text{-}pixels) - \sum (black\text{-}pixels) \quad (1)$$

To delimit the face region, a window of initial size of 24×24 (increased iteratively) scans the input image in all directions to find the face zone. This operation generates many features, which make the use of integral image required for the speed of detection. Then a cascade of classifiers based Adaboost is used to classify the zone as a face or non face depending on the value of its descriptor obtained during the training using images of face and non face.

A cascade classifier consists of several simple classifiers which are applied one after the other on a region of interest in an image, while Boosting means to combine the results obtained by several "weak" classifiers to build one more efficient.

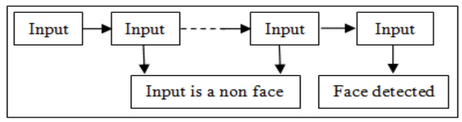

Figure 3. The cascade classifier

The integral image at the pixel (x,y) of an input image is calculated by summing the pixels above and left of the current pixel (x,y):

$$ii(x,y) = \sum_{x' \leq x,\ y' \leq y \leq}(x',y') \quad (2)$$

After this step, the algorithm of face detection return the face region localized (Fig. 4), which will be considered as inputs of the algorithm of facial features detection.

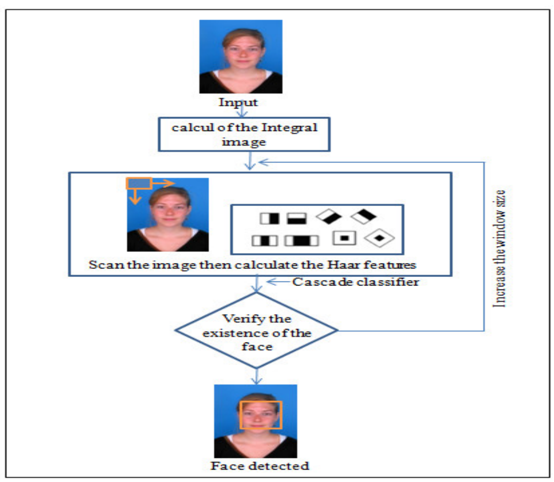

Figure 4. Face detection process

2.2 Facial features detection

To detect facial features, the result face detected in the previous step has been used with different extended Haar like classifier based Adaboost trained using different sets of images. The same steps have been followed for each facial feature (eyes, nose and mouth).

The result of face and features detection by Viola and Jones based Adaboost method, is given in the following figure for each database used.

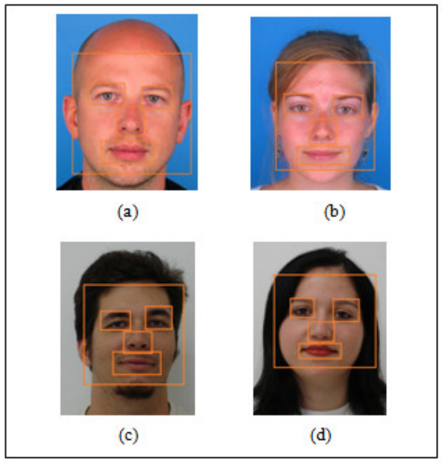

Figure 5. Facial features detection: (a) and (b) from ECVP database, (c) and (d) from FEI database.

3 Features corners detection

3.1 Harris corner detector

Harris [5] is the most popular detector for corners detection, it is based on the research of the zones in an image that has a change in intensity in multiple directions.

To capture the corners, we consider that the intensity change in direction (u,v) is given by:

$$change(u,v) = \sum_x \sum_y [(I(x+u,y+v) - I(x,y))^2] \quad (3)$$

The following matrix form describes the image I area at a given point:

$$E(u,v) = [u,v]M_{xy}[u,v]^t \quad (4)$$

With: $$M_{xy} = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (5)$$

And: $A = \frac{\delta I^2}{\delta x}$, $B = \frac{\delta I^2}{\delta y}$, $c = (\frac{\delta I}{\delta x}\frac{\delta I}{\delta y})$

$\frac{\delta I}{\delta x}$: X Sobel operator.

$\frac{\delta I}{\delta y}$: Y Sobel operator.

To determine the area's type, the eigenvalues are calculated, but since it is computationally expensive, an equivalent calculation of the term "R" is performed:

$$R = det(M) - k(trace(M))^2 \quad (6)$$

With: det(M)= AB – C^2 and Trace(M) = A+B and typically k=0.04.

Depending on the value of R, the result can be interpreted in different ways:

- If R>0 then the area is a corner.
- If R<0 then the area is contour.

If R has a small value, then the area is a uniform zone.

For corners detection, we extract the local maxima and we keep just those with the highest value; which are the areas of corners.

3.2 Active contour (Snake)

An active contour [6] is a closed curve on an image, which can be iteratively deformed due to the external constraint forces and influence of the image forces and pull towards the contours of the segmented object in the image. This deformation is guided by the minimization of an energy function to achieve equilibrium.

This energy is composed of two terms: an internal energy that serve to impose a piecewise smoothness of the snake, and an external energy function that push the snake towards salient image features, and place the snake near to the desired local minimum.

If a snake is defined as a parametric curve **v**(*s*) = (x(*s*), y(*s*)), the energy function is given by:

$$E_{snake} = \int_0^1 E_{internal}\big(V(s)\big) + E_{external}\big(V(s)\big)ds. \tag{7}$$

The external energy can be divided on two energies; Eimage refers to image energy that attracts the snake towards desired features and Econst represents the external constraint energy.

The internal energy and the image energy can be written like:

$$E_{internal} = (\alpha(s)|V_s(s)|^2 + \beta(s)|V_{ss}(s)|^2)/2 \tag{8}$$

$$E_{image} = -|\nabla I(x,y)|^2 \tag{9}$$

Where $|\mathrm{V_s(s)}|^2$ represents the elasticity, and $|\mathrm{V_{ss}(s)}|^2$gives the parametric curve curvature which forms the deforming snake. In the other hand, the parameter $\alpha(\mathrm{s})$ controls the 'tension' and $\beta(\mathrm{s})$ controls the 'rigidity'.

For the image energy, $\nabla \mathrm{I(x,y)}$ denotes the gradient of the image I at (x,y).

4 Experimental results

The experiments were performed on two face databases; the first is Utrecht ECVP [8] Fig. 6, which contains 131 images of 49 men and 20 women, in both a neutral and smile position of each. It was collected at the European Conference on Visual Perception in Utrecht in 2008. The second is the FEI database Fig. 7 [7]. It is a Brazilian face database that contains a set of face images taken at the Artificial Intelligence Laboratory of FEI in São Bernardo do Campo, São Paulo, Brazil. It contains 14 images for each of 200 individuals. All images are taken in an upright frontal position with profile rotation of up to about 180 degrees.

Figure 6. Example of images from ECVP database

Figure 7. Example of images from FEI database

We use these Databases for comparison of different facial landmarks detection methods; V.J+H and V.J+A.C. We detect facial features zones using Viola and Jones method which return the rectangles that enclose the facial features, and then we use each of the Harris and snake to localize the landmarks.

Before the landmarks detection process, we have added one step, which is the detection of facial axes Fig. 8 that goes along the features in order to get the exact location of the corners. As we know, Snake detector is made to detect the contours, so adding this step will help to get the corners of each feature.

We apply the horizontal projection of the gradient image, and we find the maximum that will refer to the eyes axis. To find the nose axis, we apply the vertical projection to the gradient of the image beginning from the eye axis, and the nose axis corresponds to the area which has the minimum skin pixel around the symmetric axis. And finally to find the mouth axis, we do the same projection as eyes, and the axis is the line that has the highest gradient below the nose axis [10].

Figure 8. Features axes detection

4.1 Landmarks detection by Harris corner detector:

This section presents the results obtained from implementation of Harris corner detector on the facial features detected.

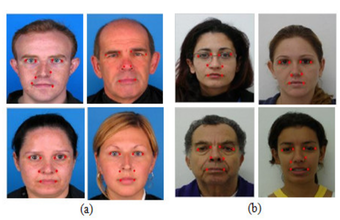

Figure 9. Results of corners detection using Harris: (a) for ECVP face database, (b) for FEI face database

The Harris corner detector detects a set of facial landmarks, among these landmarks, useless points are detected and others which are necessary are missed. Fig.9 shows some examples where the necessary landmarks are presented for the two databases.

4.2 Landmarks detection by Active contour (Snake):

The results obtained using Active contour (Snake) for facial landmarks detection are shown in Figure. 10.

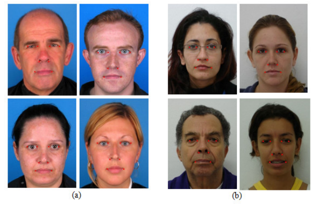

Figure 10. Results of corners detection using Active contour: (a) for ECVP face database, (b) for FEI face database

Facial landmarks detection using active contour (snake) gives very interesting results Fig.10, but the only problem is parameters initialization, which is very difficult.

In the table below, we present the success rate of facial landmarks detection for both Harris and Active contour using our two face databases Utrecht ECVP and FEI database.

Table 1. Success rate of facial features corners detection

Methods / Features	Active Contour		Harris	
	ECVP	FEI	ECVP	FEI
Eyes corners	97 %	94 %	95 %	93 %
Nose corners	89 %	90 %	81 %	88 %
Mouth corners	92 %	91%	87 %	90 %

We can conclude that the detection rate of the facial features (eyes, nose and mouth) corners by the two methods (V.J+H) and (V.J+A.C) is interesting for both methods, mainly the (V.J+A.C) that shows percentages higher than those of (V.J+H) with both ECVP and FEI databases. These results are due to the strengths of the snake detector combined with the facial axes, and the incomplete and non accurate corners detection by Harris.

5 Conclusion

In this paper, we presented two methods for facial landmarks detection using Viola and Jones method .the first combines Viola and Jones with Harris corner detector, and the second combines Viola and Jones method with Active contour (snake). Our experiments were performed on two face databases; Utrecht ECVP and FEI database, and they show that (V.J+A.C) gives very good results and had a higher landmarks detection comparing to (V.J+H).

REFERENCE

[1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001, vol. 1, pp. I–511.

[2] Y. Freund and R. E. Schapire, "A desicion-theoretic generalization of on-line learning and an application to boosting," in European conference on computational learning theory, 1995, pp. 23–37.

[3] D. Vukadinovic and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based boosted classifiers," in 2005 IEEE International Conference on Systems, Man and Cybernetics, 2005, vol. 2, pp. 1692–1698.

[4] C. Schmid, R. Mohr, and C. Bauckhage, “Evaluation of interest point detectors,” Int. J. Comput. Vis., vol. 37, no. 2, pp. 151–172, 2000.

[5] C. Harris and M. Stephens, “A combined corner and edge detector.,” in Alvey vision conference, 1988, vol. 15, p. 50.

[6] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” Int. J. Comput. Vis., vol. 1, no. 4, pp. 321–331, 1988.

[7] D. C. E. Thomaz, “FEI Face Database.” [Online]. Available: http://fei.edu.br/~cet/facedatabase.html. (accessed 17.03.25)

[8] P. Hancock, “Utrecht UCVP database.” [Online]. Available: http://pics.psych.stir.ac.uk/2D_face_sets.html (accessed 17.03.25)

[9] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” in Image Processing. 2002. Proceedings. 2002 International Conference on, 2002, vol. 1, pp. I–900.

[10] F. Abdat, C. Maaoui, and A. Pruski, “Suivi du gradient pour la localisation des caractéristiques faciales dans des images statiques,” in 21° Colloque GRETSI, Troyes, FRA, 11-14 septembre 2007, 2007

Copyright © Society for Science and Education United Kingdom