An Eye Detection Algorithm Using Pixel to Edge Information

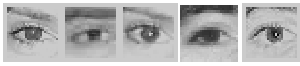
S. Asteriadis, N. Nikolaidis, A. Hajdu, I. Pitas
Department of Informatics, Aristotle University of Thessaloniki, Box 451, 54124
Thessaloniki, Greece
pitas@aiia.csd.auth.gr

Abstract— **We present a new method for eyes localization on a face, based only on geometrical information. A face detector is applied first to detect the bounding box of the face, and the edge map is extracted. A vector is assigned to every pixel, pointing to the closest edge pixel. Length and slope information for these vectors is consequently used to detect and localize the eyes. The method has been tested on two face databases with very good results.**

I. INTRODUCTION

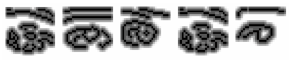
The field of eye detection has been very active in the recent years and a lot of different approaches have been proposed. Z. Zhou and X. Geng [1] use the idea of Generalized Projection Functions (GPF) to locate the eye centers in an approximate area, found using the algorithm of Wu and Zhou [2]. Jesorsky et al [3] use a three stage technique: First, the face is initially localized using the Hausdorff distance [4], second a refinement is performed, taking into account the estimated area of the face and, third, the Multi-Layer Perceptron (MLP) is applied for a more exact localization of the pupils. Cristinacce et al [5] used a multi-stage approach to detect 17 features on the human face, including the eyes. First, a face detector was applied, then the Pairwise Reinforcement of Feature Responses (PRFR) was applied to detect features. Refinement was made using a version of the Active Appearance Model (AAM) search.

A novel algorithm for eye detection on face images, including low resolution ones, is presented in this paper. Pixel intensity information might prove unreliable due to varying lighting conditions. Moreover, in low resolution images, details of the eyes sometimes are lost. Thus, an approach which detects the eyes region on a face, based on geometric information from the eye and the surrounding area is proposed. The method was applied on the XM2VTS [6] and BioID [7] databases and the results were competitive with methods demanding higher resolution images.



(a) Examples of right eye training images

(b) Examples of left eye training images



(c) Examples of length maps for the right eye images

(d) Examples of length maps for the left eye images

(e) Examples of angle maps for the right eye images

(f) Examples of angle maps for the left eye images

Figure 1: Training data

II. EYE REGION DETECTION

A. Method overview

In [8], the standard PCA method was applied on the intensity of facial images to derive the so called eigenfaces for face recognition purposes. Here, the same idea will be used for eye region detection, on data of different nature, though. The Canny edge detector [9], selected because it can be adjusted so that it finds edges of different intensity, is applied. Then, for each pixel, a vector pointing to the closest edge pixel is calculated. The magnitude (length) and the slope (angle) of each vector are the two values assigned to each pixel. Thus, instead of the intensity values, we produce the vector length and angle maps for an image. PCA is then applied on a set of training eye images to derive eigenvectors for these maps. Subsequently, in order to detect the eyes on an image, the length and angle maps of candidate regions are projected on the spaces spanned by the eigenvectors found during training and the similarity of projection weights with those of model eyes is used to declare an eye presence. The steps of the algorithm are described in detail below:

B. Extraction of eigenvectors using Training Data

k=58 right and left eye regions, collected from face images found on the web, were used in the training procedure (Figures 1a,b). The training images were scaled to the dimensions NxM (in our case N=M=25) and the Canny edge detector was applied. The corresponding vector length (Figures 1c,d) and vector angle maps (Figures 1e,f) form two separate NxM matrices which can be alternatively considered as two 1-dimensional vectors of dimension $L=NM$. We normalize these matrices by subtracting, from each of them, the respective average matrix. Then, PCA is applied on the length and angle maps of the training images resulting in 4x58 eigenvectors $U_{Ra,i}$, $U_{Rl,i}$, $U_{La,i}$, $U_{Ll,i}$ ($1<i<58$) that correspond to the angle and length maps of the right and left eye. The dimension of each eigenvector is NM. Despite the fact that the use of the standard PCA on angular data is not very well grounded, due to their periodic nature, the obtained results (see section IV) are very good and much better than using only length information.

C. Eye region detection

Prior to eye detection, a face detector has to be applied on the image. In our approach, the face detection technique proposed in [10, 11, 12] was used. The detected face area is then scaled to a certain dimension. Experiments proved that a 150x90 pixels face region retains all necessary information for locating the eyes on the face with our approach. Subsequently, edge detection was performed.

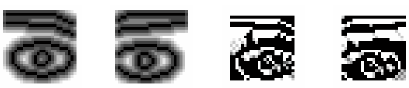
Since edges, related to eyes, are among the most prominent in a face, parameter values of the Canny detector (50 for the high threshold, 25 for the low threshold, and sigma=1 [9]) that detect some of the most significant edges, are used (Figure 2b).

(a) (b) (c) (d)

Figure 2: (a) Detected face, (b) thresholded Canny edge detector output (c) vector magnitude map (d) vector slope map

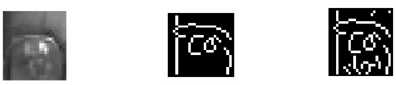
(a) (b)

Figure 3: (a) Right model-eye, (b) left model-eye



(a) (b)

Figure 4: (a):Length maps of model eyes, (b): angle maps of model eyes.



(a) (b) (c)

Figure 5: (a) Eye area of person wearing glasses causing severe reflections, (b) strong edges in the eye area of person wearing glasses (c) strong and weak edges in the same area

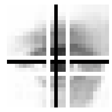


Figure 6: Eye centre localization

A visual representation of the vector length and angle maps of the face can be seen in Figures 2c, d. All the areas of size NxM are examined and the vectors containing the weights that project the length and angle maps Φ_{Ra} , Φ_{Rl}, Φ_{La}, Φ_{Ll} of the area on the corresponding spaces spanned by the eigenvectors calculated during the training stage are found. The projection vector elements $w_{Ra,i}$, $w_{Rl,i}$, $w_{La,i}$, $w_{Ll,i}$ for a given NxM area are found as follows:

$$w_{Ra,i} = U_{Ra,i}^{\,T} \cdot \Phi_{Ra} \quad (1\text{-a})$$

$$w_{Rl,i} = U_{Rl,i}^{\,T} \cdot \Phi_{Rl} \quad (1\text{-b})$$

$$w_{La,i} = U_{La,i}^{\,T} \cdot \Phi_{La} \quad (1\text{-c})$$

$$w_{Ll,i} = U_{Ll,i}^{\,T} \cdot \Phi_{Ll} \quad (1\text{-d})$$

At the end, each area is represented by two k-dimensional vectors, composed by the weights needed to project the NM-dimensional vectors of angles and lengths on the respective k-dimensional space.

To proceed with the detection, artificial eye-models were used (Figure 3). These are of dimensions NxM, where $N=M=25$. The same procedure is followed for each of these two models, as for every area of the face. That is, the vector length map and vector angle map are derived (Figure 4), and the respective projection weights $w_{Ra,Model}$, $w_{Rl,Model}$, $w_{La,Model}$, $w_{Ll,Model}$ are calculated. For each area checked on the face, the projection weights w_{Rai}, w_{Rli}, w_{Lai}, w_{Li}, were compared to those of the model eyes using the L_2 norm:

$$L_{2,R} = \| w_{Ra} - w_{Ra,Model} \| + \| w_{Rl} - w_{Rl,Model} \| \quad (2\text{-a})$$

$$L_{2,L} = \| w_{La} - w_{La,Model} \| + \| w_{Ll} - w_{Ll,Model} \| \quad (2\text{-b})$$

The facial areas with the smallest distance from the model eyes are the ones at which the eyes are located. To make the algorithm faster we utilized the knowledge of the approximate positions of eyes on a face. Thus, we searched for the eyes only in a zone on the upper part of the detected face.

III. EYE CENTER LOCALIZATION

After the area of the eye has been detected, we search in this area for the eye centers. In order to do so, we first apply the Canny edge detector with parameters (13 for the high and 3 for the low threshold, and sigma=1) that result in the detection of even weak edges (Fig.5). This was done in order to handle cases where persons are wearing glasses. In such a case, severe reflections on the glasses might make eye characteristics less visible (Fig.5b).

Looking at the images of the eyes used for training, one can observe that the eyes usually reside at the lower part of the detected area. Experiments showed that a region of the lower 15x25 pixels covers the very area of the eye. The respective edge information in this area comes from the eye itself and not from the eyebrow or glasses skeleton.

Within this area, we search for the three pairs of lines (3 horizontal and 3 vertical) having the most intersections with edges (Figure 6). The intersection of the horizontal line that has the medium number of edge intersections among the three selected horizontal lines with the vertical line with the same characteristic (black lines in Figure 6) was found to give a very good approximation of the eye center. We further refined this result by applying the same search method (using vertical and horizontal lines) around the found point within a smaller area of 10x20 pixels. For even more refined results, the information that the iris of the eye appears to be the darkest area near the point found at the previous step can be used in order to place the eye centre exactly in the middle of it. Furthermore, we can use the position of the eye center found in the right eye area to locate the left eye centre. Due to the similarity between the characteristics of the eyes of a person, the horizontal position of the left eye with respect to the upper and lower boundary of the area should be similar to that of the right eye. Using this observation, we searched for the left eye's centre within a 10x25 region centered, with respect to the vertical dimension, around the vertical position of the right eye in its area.

IV. EXPERIMENTAL RESULTS

A. Description of the test databases

The method has been tested on two databases, the BioID [7] and the extended M2VTS [6]. The BioID database consists of 1521 images of frontal faces. The pictures were taken under various lighting conditions, in a complex background and, thus, this database is considered as the most difficult one for the task of eye detection. Also, the BioID database is more challenging in the sense that it contains tilted and rotated faces. Since, in the BioID database, a lot of faces lie very close to the upper and lower borders of the images, two gray zones were added to the images above and below them, prior to face detection. A similar preprocessing was done in [5]. Furthermore, we applied our method on 600 randomly selected images from the XM2VTS database. The images in this database were taken under controlled lighting conditions with a uniform background, making the task of the face detection step much easier. Both databases contain faces with glasses and, in a few images of the BioID, people have their eyes shut or pose various expressions. Some examples of the two databases can be seen in Figure 7.

B. Results and comparisons with existing methods

Two types of results were obtained: results on eye regions detection and results on eye center localization. For eye region detection, success or failure was declared by visual inspection of the detected regions in all images. Correct detection rates are listed in the columns of Tables I and II denoted as "Eye regions". It is obvious that the obtained rates are very good. One can notice that, in the few images of people with their eyes closed, the eye region detection works with no mistake. This is due to the fact that the length/angle maps of such regions are similar to the ones of an open eye.

For the eye center localization, the criterion of calculating the correct detection rates was the following [3]:

$$m_e = \frac{\max(d_1, d_2)}{s} < T \qquad (3)$$

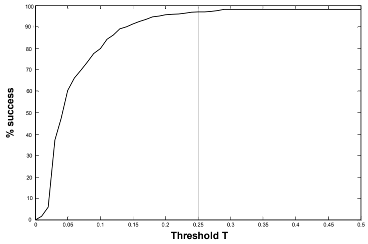
where d_1 and d_2 are the distances between the manually labeled eye centers and the eye centers found by the algorithm, s is the distance between the two manually labeled eye centers and T is the threshold used to declare a successful detection. For T=0.25 the results in Tables I and II (column "Eye centers") were obtained. Furthermore, Figure 8 presents the success rates on both databases for various values of the threshold T. Results in Tables I, II show that glasses do not introduce significant error in comparison to people not wearing glasses. This is because the basic shape of the eye region remains unchanged, even if the eyebrows are not included. In such cases, the geometry of the eyebrows is replaced by the upper part of the skeleton of the glasses. Most failures occur in the cases of people wearing glasses in combination to having their eyes shut. Furthermore, some failures occur in faces with long hair on the forehead. Some results are shown in Figure 9.

(a)

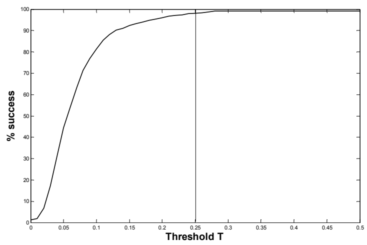
(b)

Figure 7: Examples of BioID (a) and XM2VTS (b) databases

Although the test images were subsampled by a factor of, sometimes, more than 2, the results can be successfully compared to previous methods. Since the face detector had a success of 98.8% on the BioID database, the overall result for *T=0.25* is 0.988x0.986=0.974 (97.4%), while Jesorsky *et al* [3] achieved 91.8% and Z.H. Zhou [1] had a success rate of 94.81%. For *T*=0.1 our method reaches 82.7% on BioID (81.7% with the face detector) while Jesorsky *et al* has 79%. However, the result for the PRFR+AAM method [5] with this strict criterion is 96%.



(a)



(b)

Figure 8: Eye center localization results on (a): XM2VTS and (b) BioID

Figure 9: Some eye localization results

TABLE I: RESULTS ON THE XM2VTS DATABASE

	Eye regions	Eye centers
People without glasses	99.2%	99.2%
People with glasses	96.2%	92.0%
Total	98.2%	96.7%

TABLE II: RESULTS ON THE BIOID DATABASE

	Eye regions	Eye centers
People without glasses	99.5%	99.5%
People with glasses	98.3%	96.5%
People with closed eyes	100%	97.4%
Total	99.1%	98.6%

V. CONCLUSION

A new method for eye detection was introduced. The method uses general geometrical characteristics of the eye and the surrounding area. The method proved to give very good results on images of quite small dimensions and on really natural-like images, such as those included in the BioID database. Despite the fact that the lighting conditions of the BioID database are complex, they did not significantly affect our method. The results are competitive or even better at cases than those of previous methods.

ACKNOWLEDGEMENT

The presented work was developed within VISNET, a European Network of Excellence (http://www.visnet-noe.org), funded under the European Commission IST FP6 program.

REFERENCES

[1] Z.H. Zhou and X. Geng, “Projection functions for eye detection”, *Pattern Recognition 37*, no 5, pp. 1049-1056, 2004
[2] J. Wu, Z.-H.Zhou, “Efficient face candidates selector for face detection”, *Pattern Recognition 36,* no5, pp. 1175-1186, 2003
[3] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, “Robust face detection using the hausdorff distance”, *3rd International Conference on Audio and Video-based Biometric Person Authentication*, Halmstad, Sweden, pp. 90-95, 2001
[4] W. Rucklidge, “Efficient Visual Recognition Using the Hausdorff distance”, vol 1173 of *Lecture notes in computer science*. Springer, 1996
[5] D. Cristinacce, T. Cootes, and I. Scott, “A multi-stage approach to facial feature detection”, *Proc. Of BMVC*, pp. 231-240, London 2004
[6] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre. XM2VTSDB: The extended M2VTS database. *Second International Conference on Audio and Video-based Biometric Person Authentication*, pp. 72-77, March 1999
[7] The BioID face database: [http://www.bioid.com/downloads/facedb/facedatabase.html]
[8] Matthew A. Turk and Alex P. Pentland, “Face Recognition using Eigenfaces”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* pp. 586-591, Maui, Hawaii, 1991.
[9] J. Canny “A Computational Approach to Edge Detection”*, IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 8, no. 6, 1986
[10] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features”, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001
[11] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” *IEEE ICIP*, vol. 1, pp. 900-903, Rochester, NY, 2002
[12] P. Viola and M. Jones, “Robust real-time face detection” *International Journal of Computer Vision 57*, pp. 137-154, 2004