

An Accurate Eye Center Localization Method for Low Resolution Color Imagery

Evangelos Skodras and Nikolaos Fakotakis

*Artificial Intelligence Group, Wire Communications Laboratory,
Department of Electrical and Computer Engineering, University of Patras, Patras, Greece
{evskodras, fakotaki}@upatras.gr*

Abstract—The development of eye localization and gaze tracking systems without the constraint of dedicated hardware, has lately motivated the scientific community. A lot of passive methods have been proposed; however their precision is limited and significantly reduced compared to the commercial, hardware-based, eye tracking devices. In this paper we introduce an automatic, non-intrusive method for precise eye center localization in low resolution images. To this end, the proposed system utilizes radial symmetry transform in both the original eye image and a novel eye map image derived from color information. Experimental results demonstrate great accuracy and robustness even in challenging cases, achieving also significant improvement over existing methods.

Keywords—eye localization, non-intrusive method, radial symmetry transform, eye map

I. INTRODUCTION

Being the most salient landmarks in the human face, eyes and their movements play a key role in expressing a person's cognitive and affective state, interest and attention. Accurate eye center localization and gaze tracking constitutes an integral part of a wide variety of applications including face alignment and normalization, monitoring of drivers attention and vigilance, visual attention analysis (e.g. for marketing purposes) and interactive gaze-based interfaces for disabled people.

Although many commercial products for eye detection and tracking are available on the market, they all require dedicated, high-priced hardware. The most common approaches use active infrared (IR) illumination, to obtain accurate eye location through corneal reflection [1]. The use of non-intrusive, computer vision techniques for the task of eye localization has bloomed in the last decade, as they can be incorporated in many applications where the use of extra dedicated hardware is impracticable.

Despite active research, eye center localization with high precision remains a very challenging task; eyes present great variability in shape and color depending on eye state, iris direction, head pose and ethnicity. Occlusions caused by hair, glasses or shadows make localization even more difficult. The localization process becomes even more challenging when dealing with low resolution images derived from inexpensive imaging devices (e.g. webcams, pinhole cameras or mobile devices).

Over the last decades, a great number of techniques have been employed for the task of eye detection [2]. These

can be coarsely divided into two major categories, namely *appearance-based* and *feature-based* methods.

Appearance-based methods, also known as holistic or image-based methods, incorporate eye knowledge implicitly by using the intensity distribution or filter responses of the eye area and its surroundings to train a system using example datasets [3-5]. They generally require a large amount of training data and powerful non-linear algorithms in order to learn the high variability of eyes. Although appearance-based methods can achieve remarkably high accuracy in detecting the eye area, they fail to provide an accurate detection of the eye center.

Feature-based methods make explicit use of the a priori eye knowledge in order to derive features such as shape, geometry, color and symmetry. Detailed modeling of the eye shape parametric models and complex shape-based methods have been employed [6], but despite the accuracy of these methods, they are computationally demanding, require high-resolution images and a close to the eye initialization. A number of methods have been employed in order to exploit the circular shape of the eyes, including filter responses [7], the Hough transform [8] and isophote curvatures [9]. Color has been mostly used in order to distinguish the eye region from the rest of the skin by building eye maps [10], and more rarely to accurately detect eye centers. Finally, symmetry operators have been largely investigated for the purpose of automated eye detection [11,12].

In this paper, a fast, fully automatic method for accurate eye center localization even in low resolution images, is presented. It is based on a synergy of color and radial symmetry to precisely and robustly detect eye centers. The main novelty of our method lies in the construction of a color based eye map, which emphasizes the iris area and in the synergy of the radial symmetry on both the original eye region and the eye map.

II. PROPOSED METHOD

The overview of the proposed system is illustrated in Figure 1. Once a face is detected in a given image, regions containing the eyes are defined. Color information is used to build an eye map which emphasizes the iris area. Then, a radial symmetry transform is applied both to the eye map and the original eye image. The cumulative result of the transforms indicates the precise positions of the eye centers.

A detailed description of the different stages of the algorithm follows below.

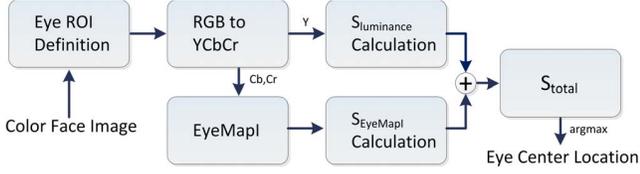


Figure 1. Block diagram of the proposed system

Face detection is carried out using the real time face detector proposed by Viola and Jones [13]. It uses boosted cascade detectors based on Haar features and represents the state-of-the-art method for face detection. Within each detected face, a *Region of Interest (ROI)* containing each of the eyes is defined heuristically based on face geometry. The dimensions of the ROIs are calculated so that they contain the whole eye areas even when reaching the detection limits of the face detector, regarding the in-plane and out-of-lane rotations [13]. The width and height of the eye regions are determined as $EyeRegionWidth = FaceWidth/3$ and $EyeRegionHeight = FaceHeight/4$. The proposed procedure is then applied to the cropped eye ROIs in order to localize the exact positions of the eye centers.

A. Eye Map Construction

With the goal of building an eye map that optimally distinguishes the eye area from the skin area, we transform the *RGB* images into the *YCbCr* color space. *YCbCr* was adopted because it is a perceptually uniform color space, which performs very well in separating the luminance from the chrominance component, as well as in modeling the skin regions [10]. These attributes prove to be very useful in order to build reliable eye maps, mitigating the effects of uneven illumination, shadows and other intensity variations.

The eye map derived from the chrominance components of the *YCbCr* color space is based on the observation that eye areas present high *Cb* values and low *Cr* values. The eye map is constructed as follows:

$$EyeMapC = \frac{1}{3} \{ (Cb)^2 + (\overline{Cr})^2 + (Cb/Cr) \} \quad (1)$$

where *Cb* and *Cr* are normalized to range $[0, 1]$ and \overline{Cr} denotes the complement of *Cr* (i.e. $1 - Cr$). This formula emphasizes pixels with high *Cb* values and low *Cr* values in order to efficiently separate skin pixels from non-skin pixels (eyes and eyebrows). Large values on the eye map are observed at the position of the irises, where the color difference from the skin pixels is maximized. Based on the observation that irises comprise the darker pixels of the eye area in the luminance component, and the brighter pixels in the eye map, gray-scale morphological operations

(dilation/erosion) can further emphasize the iris area. Gray-scale dilation and erosion with circular structural elements is used to build a new eye map:

$$EyeMapI = \frac{Y \oplus B1}{EyeMapC \ominus B2} \quad (2)$$

where \oplus denotes gray-scale dilation of the luminance component with the structuring element *B1* and \ominus denotes gray-scale erosion with the structural component *B2*. *B1* and *B2* are flat circular structuring elements with radius half of the iris radius for *B1* and equal to the iris radius for *B2*. The estimation of the radius of the iris is based on the eye ROI size and is defined as $Radius = EyeRegionWidth/10$. Deviations in the estimation of the iris radius do not have a noticeable influence in the outcome.

B. Radial Symmetry Transform

Symmetry constitutes one of the primary properties of the eyes and can be exploited both in the original eye image and in the eye map constructed. In our approach a fast and highly efficient radial symmetry transform is used, proposed in [14]. The transform is a gradient-based interest operator that works by considering the contribution each pixel makes to the symmetry of pixels around it.

The parameters that have to be defined for this process are the radial strictness parameter α and the set of radii *N* which define the range of radially symmetric features to be detected. A low radial strictness parameter ($\alpha = 1$) proves to be the more judicious choice as it also gives emphasis to non-radially symmetric features. Apart from the experimental evidence for this, an intuitive explanation can be also provided, as the radial shape of the iris largely depends on the eye state, but always preserves some level of symmetry. The set of radii $N = [n_{min}, n_{max}]$ is estimated based on the size of the face detected. The minimum radii is defined as $n_{min} = EyeRegionWidth/15$ and the maximum as $n_{max} = EyeRegionWidth/4$. In order to speed up calculations, it is possible to use a sparser set of non-continuous values instead. The result constitutes a very good approximation to the output obtained if all the continuous ranges were considered.

The final transformed image is calculated by adding the individual results of applying the radial transform to the luminance component of the eye image and to the calculated eye map (as shown in Figure 1). The position of the maximum value indicates the localized eye center:

$$(x_l, y_l) = argmax(S_{luminance} + S_{eyemap}) \quad (3)$$

where (x_l, y_l) denotes the estimated eye center coordinates and *S* the radial symmetry transforms of the luminance component and the eyemap (eq.2) respectively.



Figure 2. Eye center localization results on (a) GTAV database and (b) Caltech database. Yellow dots indicate the detected eye centers

III. EXPERIMENTAL SETUP

The performance of the proposed algorithm was tested on frontal and near-frontal face images in two publicly available databases, the GTAV face database [15] and the Caltech face database [16].

The GTAV face database contains low resolution (240 x 320 pixels) color images. It includes a total of 44 people with 27 pictures per person, which correspond to different poses, occlusions and facial expressions, under different illuminations. Due to these conditions, the GTAV database is considered as a challenging dataset. Of all the images in the database, 713 images (of all 44 persons) were meaningful for the evaluation of our algorithm: Images with extreme poses where one of the eyes was completely occluded (which also caused the frontal face detector to fail) or both eyes were completely hidden, were excluded.

The Caltech face database is used in order to test how well our algorithm performs on images of higher resolution (896 x 592 pixels). It contains 450 face images of 27 people under various facial expressions and lighting conditions. The cases where the face detector failed to correctly find the positions of the faces were excluded (15 images). Both databases were manually annotated in order to define the ground truth for the left and right eye centers.

In order to evaluate the accuracy of the proposed eye center localization algorithm, the *normalized error* (e) is used. The *normalized error* measure is specified as the Euclidean distance between each located eye center (x_l, y_l) and the corresponding ground truth (x_{gt}, y_{gt}), divided by the interocular distance (distance between eyeball centers - calculated as the Euclidean distance between the manually labeled left and right eye centers). The accuracy of the algorithm is expressed as the number of eye localizations that fall below the assigned error threshold, divided by the total number of them. The thresholds used for our tests (which are also the most commonly used) are $e \leq 0.25$

which roughly corresponds to the distance between the eye center and the eye corners, $e \leq 0.1$ which corresponds to the range of the iris and $e \leq 0.05$ which corresponds to the pupil area and its nearby region (precise localization).

IV. EXPERIMENTAL RESULTS

The evaluation of the proposed method yields robust and precise localization of eye centers, as shown in Figure 2. We observe that the method deals successfully with challenging illumination conditions, out-of-plane rotations, presence of glasses and partial occlusions by hair, reflections or shadows. It is also unaffected by the eye state, as long as the eyes are not completely obscured by the eyelids. The proposed method fails to extract accurate results only in cases when the eyes are closed and in extreme cases of uneven illumination, reflections and occlusions (Figure 3).



Figure 3. Examples of failures of the proposed system

The results of the proposed method are compared against state-of-the-art methods for accurate eye center detection in literature. The radial symmetry transform is used in [11] for fine localization of the eye centers, in coarse eye regions defined in advance. Although the parameters of the transform and the thresholds for defining the accuracy are not explicitly defined, a simulation using a wide set of ranges yields the upper detection boundary when using this approach. The method of Valenti and Gevers [9] achieves

a considerable improvement in accurately locating eye centers, when compared to the state-of-the-art methods. Their implementation of the method is evaluated in both databases to test its efficiency.

Tables I, II provide supporting evidence of the superior performance of the proposed algorithm in both databases. We can observe from Table I that in the low-resolution and difficult images of GTAV face database the proposed method achieves significant improvement in performance over the rest of the approaches examined, especially for high-precision localization ($d_n \leq 0.05$). The results remain significantly higher for detection within the iris area ($d_n \leq 0.1$), and are almost equal when providing a coarse detection of the eye area ($d_n \leq 0.25$).

Table I

ACCURACY VS. NORMALISED ERROR OF DIFFERENT METHODS FOR GTAV FACE DATABASE

Method	Accuracy		
	$d_n \leq 0.05$	$d_n \leq 0.1$	$d_n \leq 0.25$
Proposed Method	87.94 %	94.26 %	97.45 %
Yang et al.[11]	68.72 %	81.56 %	96.45 %
Valenti et al[9].	56.54 %	84.69 %	97.92 %

Table II presents the accuracy of the proposed method in the images of Caltech face database. In this case, due to higher resolution and absence of occlusions, the proposed method achieves almost perfect localization in every case.

Table II

ACCURACY VS. NORMALISED ERROR OF DIFFERENT METHODS FOR CALTECH FACE DATABASE

Method	Accuracy		
	$d_n \leq 0.05$	$d_n \leq 0.1$	$d_n \leq 0.25$
Proposed Method	99.43 %	99.54 %	99.66 %
Yang et al.[11]	98.16 %	98.39 %	99.08 %
Valenti et al[9].	89.81 %	98.59 %	99.88 %

One additional feature of the proposed algorithm is its low computational complexity. The most computationally demanding part is the radial symmetry transform, which presents relatively low complexity, depending linearly on the size of the image i.e. $O(KN)$, where K is the number of pixels and N is the radii (range) of the local neighborhood [14]. Using a Matlab implementation on a 2.53GHz Intel i5 (single core implementation), the system was able to process a 240 x 320 pixel image in 0.12 sec. With a proper C or hardware implementation, requirements of real time even for higher resolution images can be met.

V. CONCLUSION

In this paper we proposed a new, fully automatic method for precise eye center localization exploiting radial symmetry and color information. An extensive evaluation of the proposed approach was performed mostly on low resolution images containing different cases of challenging conditions, and also on higher resolution images. Experimental results reported high accuracy rates, outperforming existing methods, especially in low-resolution images. This advantage

permits the method to be used in applications utilizing low-cost image capturing devices like a simple webcam. Future efforts will be directed towards using the synergy of temporal information, for eye center tracking in video sequences. A proper real-time implementation will also allow the system to be incorporated in practical applications.

REFERENCES

- [1] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Pupil detection and tracking using multiple light sources," *Image and Vision Computing*, vol. 18, pp. 331-335, 2000.
- [2] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, pp. 478-500, 2010.
- [3] P. Campadelli, R. Lanzarotti, and G. Lipori, "Precise eye localization through a general-to-specific model definition," in *BMVC Edinburgh, UK*, pp. 187-196, 2006.
- [4] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *International Conference on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, WA, USA, pp. 84-91, 1994.
- [5] M. Castrillon, O. Dniz, C. Guerra, and M. Hernandez, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," *Journal of Visual Communication and Image Representation*, vol. 18, pp. 130-140, 2007.
- [6] X. Xie, R. Sudhakar, and H. Zhuang, "On improving eye feature extraction using deformable templates," *Pattern Recognition*, vol. 27, pp. 791-799, 1994.
- [7] S. A Sirohey and A. Rosenfeld, "Eye detection in a face image using linear and nonlinear filters," *Pattern Recognition*, vol. 34, pp. 1367-1391, 2001.
- [8] M. Dobes, J. Martinek, D. Skoupil, Z. Dobesova, and J. Pospisil, "Human eye localization using the modified Hough transform," *Optik-International Journal for Light and Electron Optics*, vol. 117, pp. 468-473, 2006.
- [9] R. Valenti and T. Gevers, "Accurate eye center location and tracking using isophote curvature," in *International Conference on Computer Vision and Pattern Recognition (CVPR'08)*, Amsterdam, the Netherlands, pp. 1-8, 2008.
- [10] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 696-706, 2002.
- [11] P. Yang, B. Du, S. Shan, and W. Gao, "A novel pupil localization method based on gaboreye model and radial symmetry operator," in *International Conference on Image Processing (ICIP'04)*, Beijing, China, Vol. 61, pp. 67-70 2004.
- [12] L. Bai, L. Shen, and Y. Wang, "A novel eye location algorithm based on radial symmetry transform," in *International Conference on Pattern Recognition (ICPR'06)*, Hong Kong, China, pp. 511-514, 2006.
- [13] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, pp. 137-154, 2004.
- [14] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, pp. 959-973, 2003.
- [15] Caltech face database (Faces 1999), available at <http://www.vision.caltech.edu/html-files/archive.html>
- [16] F. Tarrs, and A. Rama, GTAV Face Database, available at <http://gps-tsc.upc.es/GTAV/ResearchAreas/UPCFaceDatabase/GTAVFaceDatabase.htm>