A Novel Iterative Approach to Pupil Localization

Ronghang Zhu, Gaoli Sang, Wei Gao, and Qijun Zhao[*]

National Key Laboratory of Fundamental Science on Synthetic Vision,
School of Computer Science, Sichuan University, Chengdu, China
qjzhao@scu.edu.cn
http://vs.scu.edu.cn/

Abstract. This paper proposes a novel method for localizing the center of pupils. Given a face detected in an image, it first empirically initializes the eye regions in the face, and locates the pupils within the eye regions by using an improved isophote curvature based method. It then updates the eye regions according to the detected pupil centers. In the updated eye regions, the pupil centers are also refined. The above process iterates until the detected pupil centers have sufficiently high consistency with the eye regions. Compared with previous methods, the proposed method can better cope with faces with varying pose angles. Evaluation experiments have been done on the public BioID database and a set of self-collected face images which display various pose angles and illumination conditions. The results demonstrate that the proposed method can more accurately locate pupil centers and is robust to illumination and pose variations.

Keywords: Pupil localization, eye localization, face recognition, gaze estimation.

1 Introduction

Faces play an important role in human communication. Looking at a face, people can perceive plenty of information, such as identity, emotion, and gaze (or attention). Among the various components of faces, eyes are believed to be most expressive. For example, in the Viola and Jones' face detector, it has been found that eyes provide important cues for face detection [1]. In face recognition, interocular distance is often used to normalize face images before feature extraction and matching [2]. In gaze estimation, detecting eyes and localizing the positions of pupils are two important steps [3]. Efficient and effective automated eye/pupil localization algorithms are therefore highly demanded for many face related applications, such as face recognition, gaze estimation, and human-computer interfaces.

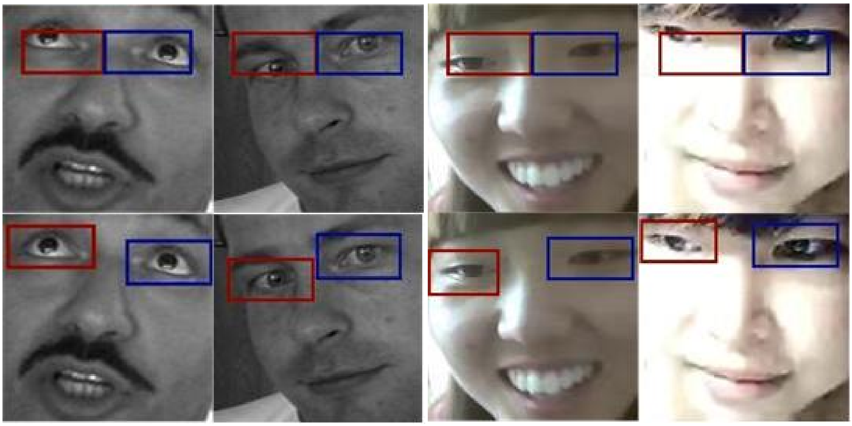
A number of automated eye/pupil localization methods have been proposed during the past decades. These methods can be roughly divided into two categories: model-based and feature-based [5][7][8][9][10][11]. Model-based methods assume statistical

[*] Corresponding author.

Z. Sun et al. (Eds.): CCBR 2014, LNCS 8833, pp. 155–162, 2014.

models of eye/pupil shapes and appearances. A sufficient number of training samples are required to train the models. Localizing eyes/pupils in face images is completed via fitting the learned models to the images. Feature-based methods, on the contrary, directly localize eyes and pupils based on certain properties of them, e.g., symmetry of eyes, circular contours of pupils, and strong intensity contrast in eye regions. Although model-based methods are generally more stable than feature-based methods, the latter are more easily to apply without requiring training/fitting. More detailed review of existing eye/pupil localization methods can be found in [3][4].

Most pupil localization methods require that faces are firstly detected in the input images. They then empirically determine eye regions within the detected faces, and localize pupils in the eye regions. Obviously, accurate localization of eye regions is crucial for pupil localization. In a recent feature-based method proposed by Valenti and Gevers in [5], the eye regions are estimated according to anthropometic relations, and fixed during the process of localizing pupil centers. As a result, the pupil localization accuracy highly depends on the face detection results and the estimated eye regions. When the face has non-frontal poses, the empirically estimated eye regions might severely deviate from the true regions, which leads to inaccurately localized pupil centers. See Fig 1.



(a) (b)

Fig. 1. The empirically determined eye regions (first row) and the eye regions estimated with the proposed method (second row) on some example images from the (a) BioID and (b) SCU-EYE databases

In this paper, we aim to improve the pupil localization accuracy particularly for faces with pose variations. To the end, we propose to iteratively update the estimated eye regions according to the estimated pupil centers, and continuously refine the pupil centers until the pupils and the eye regions have sufficiently high consistency. In this way, the pupils can be more accurately located, even when the faces have large pose angles. The contributions of this paper are three folds:

- Proposing a method to iteratively refine the estimated eye regions and the localized pupil centers.
- Improving the state-of-the-art pupil localization method in [5] by using a better weighted voting scheme in determining the center of pupils, and by employing the proposed iterative approach.
- Constructing a set of face images which display various pose angles and whose ground truth eye regions and pupil centers are manually marked. This dataset will be publicly available upon request for research purposes.

The rest of this paper is organized as follows. Section II introduces in detail our proposed method for localizing the center of pupils. Section III reports our experimental results. Finally, Section IV concludes the paper.

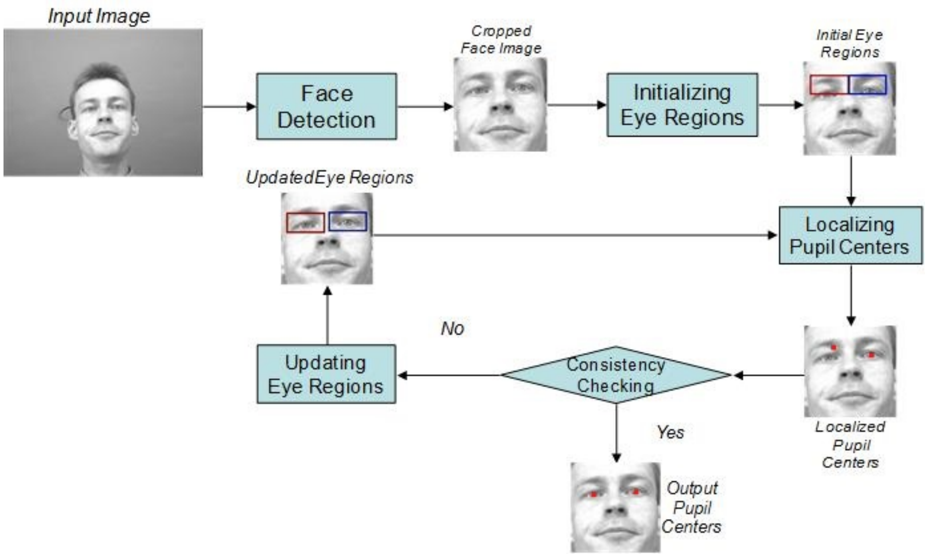


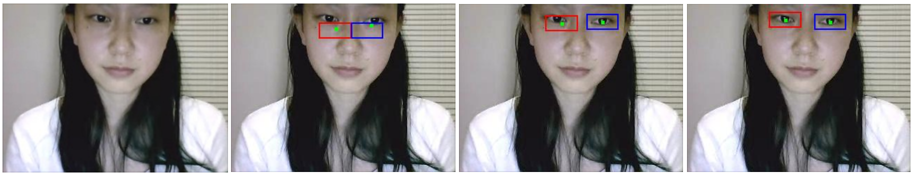
Fig. 2. Block diagram of the proposed method

2 Proposed Method

2.1 Overview

Figure 2 shows the flowchart of the proposed method. As can be seen, it consists of five main steps: face detection, initializing eye regions, localizing pupil centers, consistency checking, and updating eye regions. In this paper, we employ the face detector in [6] to detect the face regions in input images. Once a face region is cropped from the input images, initial eye regions are determined according to anthropometric relations, i.e., the top left corner of the eye region is at 15% height and 10% width of the face region, the right bottom corner of the eye region is at 35% height and 90% width of the face region, and the left and right eyes averagely divide the eye region. Pupil centers are then localized within the eye regions by using an improved isophote curvature based method. The localized pupil centers are then compared with the centers of the eye regions. If the distances between them are smaller than a given threshold (say 4 pixels), the pupils and the eye regions are said to be consistent; otherwise, the eye regions are updated by moving them so that their centers are at the localized pupil centers, and the pupils are re-localized in the updated eye regions. The procedures of localizing pupil centers and updating eye regions are repeated until the pupils are sufficiently consistent with the eye regions.

Compared with previous pupil localization methods, the proposed method has two distinct characteristics. First, the eye regions are iteratively updated. Second, the pupil localization process and the estimation of eye regions are coupled to improve the localization accuracy. Thanks to these characteristics, the proposed method can gradually correct the eye regions when the initially estimated eye regions deviate from the true positions (see Fig.3 for some example). Moreover, the pupils are also more accurately localized as can be seen in Fig. 3. Next, we introduce the improved isophote curvature based pupil localization method.



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

Fig. 3. The iterative localization process of the proposed method. (a) An input image, (b) initial eye regions and the pupil centers localized inside the regions, (c) the result after 2 iterations, and (d) the final result after 3 iterations.

2.2 Isophote Curvature Based Pupil Localization

Given an image I, the isophote in it refers to the curves connecting pixels of equal intensity (e.g., the pixels on the contour of a pupil). The isophote curvature at a pixel is defined as

$$k = -\frac{I_y^2 I_{xx} - 2 I_x I_{xy} I_y + I_x^2 I_{yy}}{(I_x^2 + I_y^2)^{3/2}} \quad (1)$$

where I_x and I_y are, respectively, the 1^{st} order derivatives along x and y axes, I_{xx}, I_{xy}, and I_{yy} are 2^{nd} order derivatives. The sign of isophote curvature is positive if the outer side of the isophote curve is brighter. For example, the pixels on the contours of pupils and irises have positive isophote curvatures, while the pixels on the contours of eye sockets have negative isophote curvatures.

The isophote curvatures can be used to detect circular patterns in images [5]. Mathematically, the displacement of the center of potential circular pattern at the pixel (x, y) can be computed by

$$\begin{cases} \Delta x = \frac{I_x}{\sqrt{I_x^2 + I_y^2}} \cdot \frac{1}{k}, \\ \Delta y = \frac{I_y}{\sqrt{I_x^2 + I_y^2}} \cdot \frac{1}{k} \end{cases} \quad (2)$$

In other words, there is a circle passing the pixel (x, y), whose center is at the position $(x + \Delta x, y + \Delta y)$. Based on this property of isophote curvatures, we can estimate the center of pupils, which are assumed to be circular patterns in the images of eyes.

Specifically, we use the pixels in an eye region to vote for the pupil center of the eye. Each pixel votes for a center of a potential circular pattern which is computed according to the eq. (2). After all the pixels have cast their votes, the circular center that obtains the highest number of votes is taken as the estimated pupil center in the eye region. As pointed by [5], edge pixels are usually more reliable in determining the pupil centers. Hence, it suggests to weight the votes of pixels with the following edgeness measurement at their positions,

$$\omega_e = \sqrt{I_{xx}^2 + 2 \cdot I_{xy}^2 + I_{yy}^2}, \quad (3)$$

In our experiments, we have observed that if only using the above-mentioned edgeness determined weights, the refinement process of pupil centers could become unstable, i.e., the newly estimated positions of pupil centers might quickly move away from the positions estimated last time. When eyeglass frames appear in the eye regions, this phenomenon can be more frequently observed. Considering that distracts like eyeglass frames are usually near the boundary of eye regions, we propose to include another component into the weight as follows,

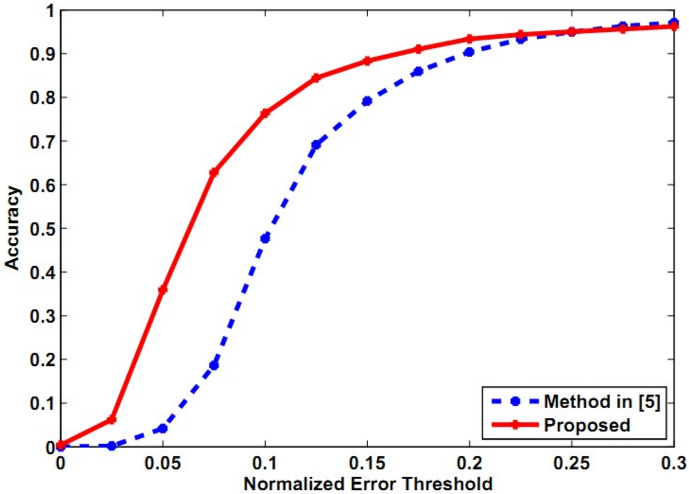
$$\omega = \omega_e \times \omega_d, \quad (4)$$

where ω_d is the reciprocal of the distance from the pixel to the currently estimated pupil center. In this way, the pixels that are far from the estimated pupil centers are assigned with small weights, and the stability of the refinement process can be ensured.

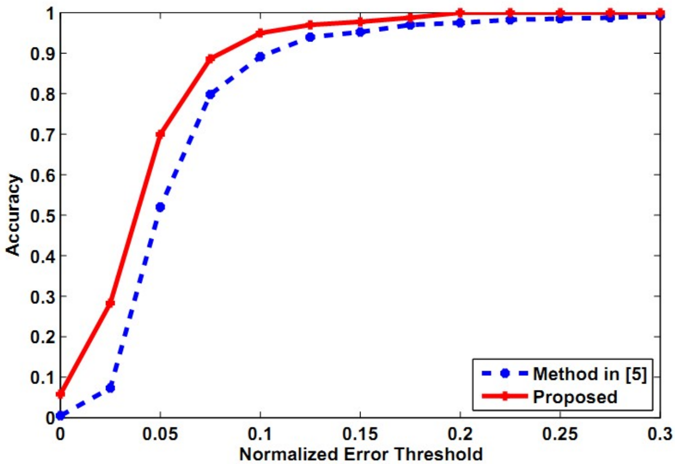
3 Experiments

3.1 Databases and Protocols

Two face databases have been used to evaluate the effectiveness of the proposed method: the BioID face image database and a set of face images collected by ourselves (referred to as the SCU-EYE database). The BioID database contains 1,521 grayscale images of 23 subjects. The image size is 384×288 pixels. The self-collected face image database consists of 396 color images of 16 subjects. The size of these



(a)



(b)

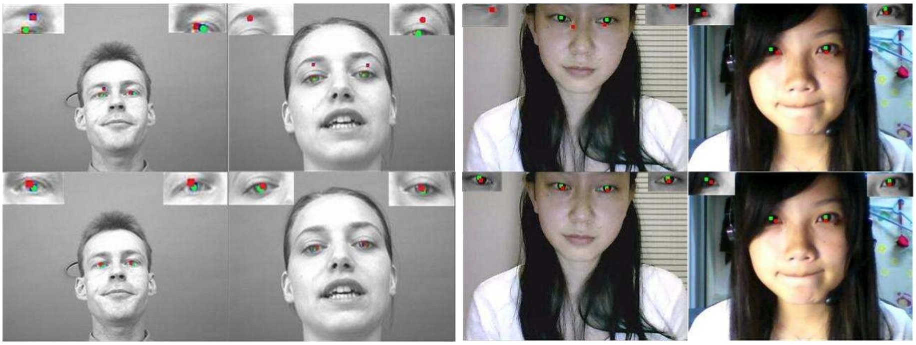
Fig. 4. The pupil localization accuracy on the (a) BioID and (b) SCU-EYE databases

images is 320×240 pixels. They were captured by using web cameras when the subjects were sitting in front of the web cameras and involved in video chats. Note that the subjects were not aware of when the pictures were taken. Consequently, the faces in these images display large variations of scales, pose angles, facial expressions and illumination conditions. The pupil centers in the face images in both databases are manually marked, which are taken as the ground truth.

The faces in these images were automatically detected by using the face detector in [6]. After excluding those images in which the face regions are failed to be correctly detected, we have used 1,198 face images in the BioID database and 396 face images in the SCU-EYE database in our evaluation experiments. We use the normalized error to measure the pupil localization accuracy, which is defined as

$$e = \frac{\max(d_{lfet}, d_{right})}{d_{interocular}}, \quad (5)$$

where $d_{left}(d_{right})$ is the distance between the localized left (right) pupil center and the ground truth left (right) pupil center, and $d_{interocular}$ denotes the distance between the ground truth left and right pupil centers. The pupil localization accuracy at a given normalized error threshold is then computed by dividing the number of face images, for which the pupil localization normalized error is below the threshold, by the total number of testing face images.



(a) (b)

Fig. 5. Results of the state-of-the-art method [5] (first row) and the proposed method (second row) on some typical images in the (a) BioID and (b) SCU-EYE databases. Green dots represent the ground truth pupil centers, and red dots represent the estimated locations.

3.2 Results

Figure 4 shows the pupil localization accuracy at different normalized error thresholds of the proposed method on the two databases. The results of the state-of-the-art method in [5] are also presented for comparison. It can be clearly seen that the proposed method has higher pupil localization accuracy. The localized pupil centers in some typical face images are given in Fig. 5.

4 Conclusions

In this paper, we have presented a novel iterative method for localizing pupil centers. Instead of using a fixed eye region, the proposed method iteratively updates the eye region according to the estimated pupil center, and further refines the pupil center by considering the updated eye region. This coupled iterative procedure of refining eye regions and pupil centers helps to improve the accuracy of pupil localization, particularly for non-frontal faces with varying pose angles. We have evaluated the proposed method on two databases, and the results prove the effectiveness of the proposed method. In our future work, we are going to incorporate the proposed method with more advanced face detection methods, and to combine it with other model-based methods to further improve the accuracy.

Acknowledgements. This work is supported by grants from the National Natural Science Foundation of China (No. 61202161) and the National Key Scientific Instrument and Equipment Development Projects (No. 2013YQ49087904)..

References

1. Viola, P., Jones, M.: Robust Real-Time Face Detection. Internal Journal of Computer Vision 57(2), 137–154 (2004)
2. Li, S.Z., Jain, A.K.: Handbook of Face Recognition, 2nd edn. Springer, London (2011)
3. Hansen, D.W., Ji, Q.: In the Eye of Beholder: A Survey of Models for Eyes and Gaze. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(3), 478–500 (2010)
4. Song, F., Tan, X., Chen, S., Zhou, Z.: A Literature Survey on Robust and Efficient Eye Localization in Real-Life Scenarios. Pattern Recognition 46(12), 3157–3173 (2013)
5. Valenti, R., Gevers, T.: Accurate Eye Center Location through Invariant Isocentric Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(9), 1785–1798 (2012)
6. Jesorsky, O., Kirchbergand, K.J., Frischholz, R.: Robust Face Detection using the Hausdorff Distance. In: Third International Conference on Audio- and Video- Based Biometric Person Authentication, Halmstad, Sweden, pp. 90–95 (2001)
7. Ma, Y., Ding, X., Wang, Z., et al.: Robust Precise Eye Location under Probabilistic Framework. In: Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 339–344 (2004)
8. Kothari, R., Mithchell, J.: L.: Detection of Eye Location in Unconstrained Visual Images. In: International Conference on Image Processing, vol. 3, pp. 519–522 (1996)
9. Asteriadis, S., Nikolaidis, N., Hajdu, A., Pitas, I.: An Eye Detection Algorithm using Pixel to Edge Information. In: Int. Symp. on Control, Commun. and Sign. Proc. (2006)
10. Bai, L., Shen, L., Wang, Y.: A Novel Eye Location Algorithm based on Radial Symmetry Transform. In: ICPR, pp. 511–514 (2006)
11. Jesorsky, O., Kirchberg, K.J., Frischholz, R.W.: Robust Face Detection Using the Hausdorff Distance. In: Bigun, J., Smeraldi, F. (eds.) AVBPA 2001. LNCS, vol. 2091, pp. 90–95. Springer, Heidelberg (2001)