A NEW HIGH PRECISION EYE CENTER LOCALIZATION TECHNIQUE

Nikolaos Poulopoulos and Emmanouil Z. Psarakis

Department of Computer Engineering & Informatics, University of Patras, 26500 Patras, Greece

ABSTRACT

The accurate eye center localization constitutes a challenging problem. In this paper a new technique based on a modified version of Radial Symmetry Transform is proposed. Specifically, computing the components of the transform from different modalities of the original image, we succeeded to improve the accuracy of the eye center localizer. The results we obtained from the application of the proposed method to the most challenging face databases show that our technique outperforms in accuracy most of the state of the art techniques.

Index Terms— Eye Detection, Eye Localization, Eye Tracking, Gaze Tracking

1. INTRODUCTION

Over the last decades many techniques have been proposed in the literature for the eye center detection and its tracking. Eye localization methods, working under different illumination conditions, the presence of occlusions and shadows can be roughly divided into the following two main categories: (i) Feature based methods and (ii) Appearance based methods.

Feature based methods use a priori knowledge to detect candidate eye centers from simple pertinent features based on shape, geometry, color and symmetry. These features are obtained from the application of specific filters on the image and do not require any learning or model fitting techniques. In most cases they are also robust to shape and scale changes. A number of methods have been employed trying to model the eye shape parametrically by matching a deformable template to the image and minimizing an energy based cost function [1],[2] and thus achieving the desired localization. Exploiting the circularity of the iris, Hough transform is another widely used eye localization method [3], [4]. However, its use is constrained only in frontal or near frontal and well illuminated faces in high resolution images. The use of isophote curvatures as a voting scheme for detecting eye locations was proposed by Valenti et al. [5]. However, this method can lead to false detections when the number of features in the eye region is insufficient. This problem was solved in [6] with the use of the SHIFT descriptor and a k-NN based classifier in a machine learning framework. Radial symmetry operators have also been studied and their use, usually in combination with other operators, for the automatic eye detection is proposed [7]. Yang et al. in [8] presented an algorithm for first detecting the eye region with Gabor filters and then localizing the center of the iris with the use of the radial symmetry operator. Finally, color information has been used to distinguish the eye from the skin area. Skodras et al. in [9] proposed a method based on the synergy of color and radial symmetry of the eyes to localize their centers.

In general, appearance based methods employ a prior model of the eye holistic appearance and surrounding structures and try to detect the location of the eyes by fitting the trained model. For this purpose, many machine learning algorithms have been proposed. Specifically, Niu et al. in [10] introduced a two-direction cascaded AdaBoost framework for eye localization, Campadelli et al. in [11] proposed an eye localization technique using a SVM trained on properly selected Haar wavelet coefficients, while techniques based on artificial neural networks, Bayesian models and hidden Markov models (HMM) were proposed in [12, 13], [14] and [15] respectively.

Despite their enhanced accuracy in detecting the eye area, in the case of pose and illumination variations the appearance-based methods fail to locate precisely the eye centers. In addition, a large amount of training data is required to be collected in order the high variability of the eyes to be reliably learned by the algorithms.

The performance, both in accuracy and efficiency, of the technique proposed in [9] by Skodras et al. seems to outperform most of the aforementioned methods. However, the accuracy of this method degrades in the presence of glasses because often the color of the frame of the glasses and the iris are similar and therefore they equally enhanced by the eye map operator.

In this paper, in order to overcome this problem, a modified radial symmetry transform is introduced and a method based on its application on both the original and edge-preserved images is proposed. The proposed method pays attention to the circular shape of the iris and achieves a precise localization of the eye center even in the presence of strong photometric distortions, shadows and occlusions (e.g. glasses).

2. THE PROPOSED METHOD

The proposed method consists of the following steps: firstly, the face is detected and the two eye Regions Of Interest (ROIs) are selected. An edge-preserving filter is applied to enhance the circular shape of the eyes and separate them from the skin. Then, a modified Radial Symmetry Transform (RST) [16] is used to localize the eye centers. Specifically, its magnitude component results from the red color component of the original image while its orientation component from a properly filtered version of the original one. Then, the superposition of their normalized counterparts is used for the accurate identification of the eyes centers. The block diagram of the proposed technique is shown in Figure 1.

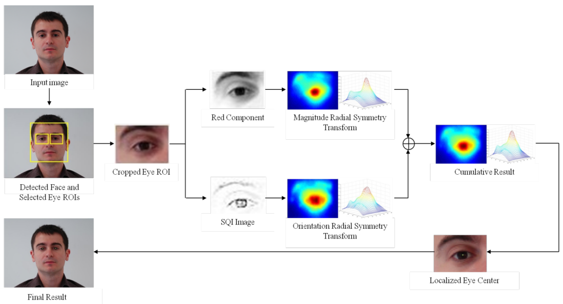

Fig. 1: Overview of the stages of the proposed system

2.1. Face detection and eye ROI selection

Face detection is performed using the real-time face detector proposed by Viola and Jones [17]. This algorithm constitutes the state-of-the-art in the face detection problem. After the determination of the face area FaceWidth $\times$ FaceHeight and based on the face geometry [9], the two eyes ROIs are selected (let us denote them by $\mathbb{R}_L$, $\mathbb{R}_R$ for the Left and Right eye respectively) with the width and height of the eye regions to be determined as follows:

$$\text{EyeRegionWidth} = \frac{\text{FaceWidth}}{3}, \text{ and} \tag{1}$$

$$\text{EyeRegionHeight} = \frac{\text{FaceHeight}}{4}. \tag{2}$$

The proposed method is then applied to the two eyes ROIs, shown in Figure 1, in order to minimize the computation time and increase the accuracy of the detected eyes centers.

2.2. Edge Preserving Filtering

2.2.1. Self Quotient Image

Illumination is considered as one of the main limitations in the eye localization problem. This is because of its influence on the eye discrimination from the shaded skin. A powerful technique achieving invariance to such distortions, is the well known Self Quotient Image (SQI) [18] technique. According to this technique, the SQI is defined as the ratio of the input image and its smoothed version. This method aims to construct a lighting invariant representation of the image which can be applied to both shadowed and no-shadowed regions and effectively removes the shadow for any type of lighting sources.

To this end let $I_\sigma(\mathbf{x})$ be a smoothed version of the image $I(\mathbf{x})$ that results from its convolution with the isotropic Gaussian kernel $G_\sigma(\mathbf{x})$ with subscript σ denoting its standard deviation, that is:

$$I_\sigma(\mathbf{x}) = I(\mathbf{x}) * G_\sigma(\mathbf{x}) \tag{3}$$

where "$*$" denotes the convolutional operator. Then, the Self Quotient Image is defined as follows:

$$Q(\mathbf{x}) = \frac{I(\mathbf{x})}{I_\sigma(\mathbf{x})}. \tag{4}$$

Note that the deviation σ of the Gaussian kernel controls the width of the edges in the shadows free image defined in (4).

As it is expected, because of the division operation in its definition, SQI suffers from noise. To eliminate this undesirable amplification of the noise, in the next paragraph, we are going to propose a denoising scheme.

2.2.2. The Proposed Denoising Scheme

Let as define the following sigmoid function $S : \mathbb{R} \to \mathbb{R}$:

$$S(x) = \frac{1}{1 + e^{-\alpha(x - x_0)}} \tag{5}$$

with x_0 denoting a displacement value that shifts the sigmoid curve and α the growth factor that determines its slope. Then, the following $two - step$ procedure is proposed:

- $\mathcal{S}_1$: *Sigma Correction.* Application of the non-linear function $S(.)$ on the SQ Image defined in (4), i.e.:

  $$T(\mathbf{x}) = S(Q(\mathbf{x})) \tag{6}$$

  with $Q(\mathbf{x})$ denoting the intensity of the SQI at pixel $\mathbf{x}$. From the application of this step, an enhanced version of the original SQI, with the noise suppressed, is obtained.

- $\mathcal{S}_2$: *Gaussian Smoothing.* In order to further smooth the SQ image the use of a convolutional Gaussian kernel is proposed with its deviation σ' controlling the strongness of smoothing effect, i.e.:

  $$Q_f(\mathbf{x}) = T(\mathbf{x}) * G_{\sigma'}(\mathbf{x}). \tag{7}$$

The SQIs resulting from the application of the above mentioned procedure on four (4) face images with strong illumination distortions are shown in Figure 2. As it is evident the SQ images are illumination independent and the removal of the shadows after the filtering makes the eyes to be easily distinguished.

2.3. The Modified Radial Symmetry Transform

Symmetry plays a vital role in the problem of the localization of the eye center. Symmetry is self evident in the eyes. In order to exploit it for solving the eye localization problem we propose a modification of the original RST; namely the computation of its component from different modalities of the

Fig. 2: Face images with strong illumination distortions and the resulting SQIs after the application of the proposed denoising scheme with $x_0 = 80,\ \alpha = 0.02,$ and $\sigma = \sigma' = 1$

original image. The Radial Symmetry Transform proposed in [16] is a voting procedure that highlights the circular shapes. To this end, let us define the following set of radii:

$$\mathbb{W} = \{r \in \mathbb{N} : r_{\min} \leq r \leq r_{\max}\}, \tag{8}$$

where $r_{\min},\ r_{\max}$ are estimated from the under consideration face size as follows [9]:

$$r_{\min} = \max\left\{\frac{\text{FaceWidth}}{60},\ 3\right\},\ r_{\max} = \frac{\text{FaceWidth}}{6}.$$

Assuming that the gradient $\nabla I(\mathbf{x})$ of a given image $I(\mathbf{x})$ is available, then for each radius r and for each pixel $\mathbf{x}$ an affected pixel $\mathbf{x}(r)$ is obtained by translating its coordinates at distance proportional to r along the gradient direction, i.e.:

$$\mathbf{x}(r) = \mathbf{x} + round\Big(\frac{\nabla I(\mathbf{x})}{||\nabla I(\mathbf{x})||_2} r\Big) \tag{9}$$

where $||\mathbf{x}||_2$ denotes the l_2 norm of the vector $\mathbf{x}$ and $round(\alpha)$ denotes the operator that rounds α to the closest integer.
In addition, a ”Magnitude” and an ”Orientation” Projection Image are constructed as follows:

$$M(\mathbf{x}(r)) \quad = \quad M(\mathbf{x}(r)) + ||\nabla I(\mathbf{x})||_2 \tag{10}$$
$$O(\mathbf{x}(r)) \quad = \quad O(\mathbf{x}(r)) + 1. \tag{11}$$

Finally, the contribution of every radius r is convolved with a Gaussian kernel $G_r(\mathbf{x})$ and summed to form the final result:

$$S_{\mathcal{P}}(\mathbf{x}) = \sum_{r \in \mathbb{W}} \mathcal{P}(\mathbf{x}(r)) * G_{\sigma_r}(\mathbf{x}(r)),\ \mathcal{P} \in \{M,\ O\} \tag{12}$$

with the standard deviation σ_r of the Gaussian kernel being controlling the smoothing of the magnitude and orientation component and their normalized counterpart are computed as follow:

$$\bar{S}_{\mathcal{P}}(\mathbf{x}) = \frac{S_{\mathcal{P}}(\mathbf{x})}{\max\limits_{\mathbf{x} \in \mathbb{R}}\{S_{\mathcal{P}}(\mathbf{x})\}}, \mathcal{P} \in \{M,\ O\}. \tag{13}$$

Note that in the proposed Modified RST, the normalization of the RSTs takes place after the summation of the projection images defined in Equation (12) over the set $\mathbb{W}$.

Having defined the normalized projection images, we are going to propose their computation from different modalities of the same original image. This, as we are going to see in the next paragraphs, drastically improves the performance of RST.

2.3.1. Magnitude and Orientation Projection Images

In order to exploit the enhanced contrast existing between the eyes and the skin, the Magnitude projection image is proposed to be computed from the Red component of the original image. The selection of this specific color component is based on the fact that the contribution of the skin pixels to that color component is higher than the other two, thus achieving a greater contrast between the eyes and the skin.
On the other hand, in order to exploit the edge-preserving filtering in distinguishing the eye shape, the orientation component of RST is applied to the SQ Image, defined in (4). This component of the RST selects the pixels that contribute to the circular shape of the eye, ignoring their magnitude and counting only on the orientation of the gradient of the image.

2.3.2. The Optimization Problems

Having computed the above mentioned projection images and by taking into account that their values are proportional to the circular symmetry of the original image, the location of the eye center results from the solution of the following optimization problems:

$$\mathbf{x}_s^\star = \arg\max_{\mathbf{x} \in\ \mathbb{R}_s}\{\bar{S}_M(\mathbf{x}) + \bar{S}_O(\mathbf{x})\},\ s \in \{L,\ R\}. \tag{14}$$

3. EXPERIMENTS

3.1. Experimental Setup

In order to evaluate the performance of the proposed method we have conducted several experiments based on two publicly available face databases. Specifically, the selected databases MUCT [19] and BioID [20] are among the most challenging and characteristic datasets and were widely used in previous eye-center localization techniques. The images where the face detector failed to detect the face due to extreme poses, were excluded for the experiments.
The MUCT face database consists of 3755 low resolution (640×480 pixels) color images of frontal or near frontal faces, containing a wide variety of ages, ethnicities and light conditions. The images were acquired using five webcams from different positions, resulting in pose variations. This, in combination with the occlusions from hair, glasses and reflections, makes the precise localization extremely difficult.
The BioID database consists of 1521 grayscale images of 23 subjects taken at different times of the day in different positions with a low resolution camera (384×288 pixels). The size, the position as well as the pose of the faces varies. Furthermore, many subjects are wearing glasses, while in some

instances the eyes are closed or hidden by strong reflections on glasses. Thus, it is regarded as one of the most challenging databases. For the purpose of eye center localization, 29 images with totally closed eyes were manually removed.
As a performance measure, the following normalized error, quantifying the worst eye center estimation of the two eyes, is adopted from [21]:

$$e = \frac{\max\{||\hat{C}_L - C_L||_2,\ ||\hat{C}_R - C_R||_2\}}{||C_L - C_R||_2} \tag{15}$$

where, $\hat{C}_L,\ \hat{C}_R$ are the localized by the used method left and right eye center coordinates respectively and $C_L,\ C_R$ are the manually labeled corresponding ones. The accuracy of the algorithm is expressed by the ratio between the number of the eye center localizations that fall below an assigned error threshold and the total number of them. Points where $e \leq 0.25$ belong to a disk with its center located to the eye center and its radius equals to the distance of its center to the eye corner, points with $e \leq 0.1$ belong to the disk of the iris while points with $e \leq 0.05$ belong to the pupil area.

3.2. Experimental Results

The evaluation of the proposed method leads us to the conclusion that it constitutes a robust and highly precise localization method. Indeed, the proposed method deals successfully with the most challenging circumstances including shadows, pose variations, occlusions by hair or strong reflections, out-of-plane rotations and presence of glasses (Figure 3). However, it fails to accurately locate the eye centers only in cases when the eyes are totally closed and in extreme cases of irregular illuminations, shadows and occlusions where the eyes can be semi-hidden (Figure 4).

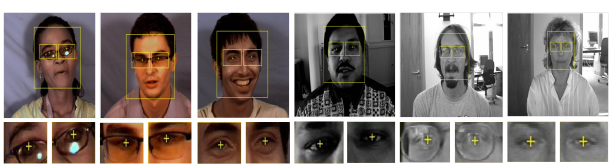

Fig. 3: Precise eye center localization results

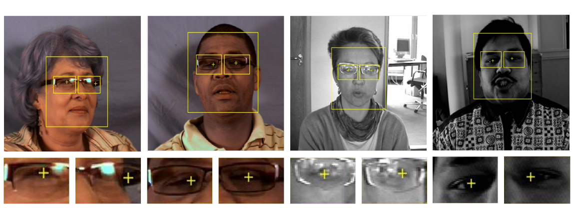

Fig. 4: Examples where our algorithm gave inaccurate results

In this experiment we apply the proposed technique in a number of face images and compare its performance against to the state of the art methods. The results we obtained are contained in Tables 1 and 2. The contents of the tables provide supporting evidence of the superior performance of the proposed method. Table 1 presents a comparison between the performance of the proposed method and the performance of other relative works, tested on the MUCT database. As it is evident, a significant improvement is achieved in all three error categories.

Table 1: Accuracy vs. normalized error in the MUCT database

Method	Accuracy (%)		
	$e \leq 0.05$	$e \leq 0.1$	$e \leq 0.25$
Proposed	**94.43**	**98.53**	**99.62**
Skodras [9]	92.9	97.2	99.0
Timm [22]	78.6	94.9	98.6
Yang [8]	81.6	89.5	94.5
Valenti [5]	63.1	76.7	94.1

The performance of the proposed method against the state-of-the-art methods from their application in the BioID face database are contained in Table 2. In this experiment, since BioID contains only grayscale images, instead of the red image component the original grayscale image was used for the computation of the Magnitude projection image. The superiority of the proposed method is obvious in all error categories resulting a high precision localization.

Table 2: Accuracy vs. normalized error in the BioID database

Method	Accuracy (%)		
	$e \leq 0.05$	$e \leq 0.1$	$e \leq 0.25$
Proposed	**87.10**	**98.00**	**100**
Valenti [6]	86.10	91.67	97.87
Anjith [23]	85.00	94.30	-
Timm [22]	82.50	93.40	98.00
Leo [24]	80.70	87.30	94.00
Campadelli [11]	62.00	85.20	96.10
Cristinacce [25]	57.00	96.00	97.10
Niu [10]	75.00	93.00	97.00
Asadifard [26]	47.00	86.00	96.00

4. CONCLUSION

In this paper, a new, fully automatic eye center localization technique was presented, based on a modified version of Radial Symmetry Transform to precisely detect the eye centers. The proposed method was tested among the most challenging, in terms of degradations, face databases and outperformed in accuracy most of the state of the art techniques.

5. REFERENCES

[1] A. Yuille, P. Hallinan, and D. Cohen, "Feature extraction from faces using deformable templates," *International Journal on Computer Vision*, vol. 8, pp. 99111, Feb 1992.

[2] X. Xie, R. Sudhakar, and H. Zhuang, "On improving eye feature extraction using deformable templates," *Pattern Recognition*, vol. 27, pp. 791799, June 1994.

[3] M. Dobes, J. Martinek, D. Skoupil, Z. Dobesova, and J. Pospisil, "Human eye localization using the modified hough transform," *International Journal on Light Electron Optics*, vol. 117, pp. 468473, October 2006.

[4] N. Van Huan and H. Kim, "A novel circle detection method for iris segmentation," in *Congress on Image and Signal Processing (CISP'08)*. IEEE, 2008, p. 620624.

[5] 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)*. IEEE, 2008, p. 18.

[6] Valenti R. and T. Gevers, "Accurate eye center location through invariant isocentric patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 17851798, Sept 2012.

[7] 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)*. IEEE, 2006, p. 511514.

[8] P. Yang, B. Du, S. Shan, and W. Gao, "A novel pupil localization method based on gabor eye model and radial symmetry operator," in *International Conference on Image Processing (ICIP'04)*. IEEE, 2004, p. 6770.

[9] Skodras E. and N. Fakotakis, "Precise localization of eye centers in low resolution color images," *Image and Vision Computing Journal, Elsevier*, 2015.

[10] Z. Niu, S. Shan, S. Yan, X. Chen, and W. Gao, "2d cascaded adaboost for eye localization," in *International Conference on Pattern Recognition (ICPR'06)*. IEEE, 2006.

[11] P. Campadelli, R. Lanzarotti, and G. Lipori, "Precise eye localization through a general-tospecific model definition," in *Proceedings CVPR'94*. BMVC, 2006, pp. 187–196.

[12] Z. Zhu and Q. Ji, "Eye and gaze tracking for interactive graphic display," *Mach. Vis. Appl*, vol. 15, pp. 139148, March 2004.

[13] H.M. Peixoto, A.M.G. Guerreiro, and A.D. Neto, "Image processing for eye detection and classification of the gaze direction," in *International Joint Conference on Neural Networks*. IEEE, 2009, p. 24752480.

[14] M. Everingham and A. Zisserman, "Regression and classification approaches to eye localization in face images," in *International Conference on Automatic Face and Gesture Recognition*. IEEE, 2006, p. 441446.

[15] F. Samaria and S. Young, "Hmm-based architecture for face identification," *Image and Vision Computing Journal, Elsevier*, vol. 12, pp. 537543, August 1994.

[16] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 959973, August 2003.

[17] P. Viola and M. Jones, "Robust real-time face detection," *Int. Journal on Compututer Vision*, vol. 57, pp. 137–154, February 2004.

[18] H. Wang, S.Z. Li, Y. Wang, and J. Zhang, "Self quotient image for face recognition," in *ICIP*. IEEE, 2004, vol. 2, pp. 1397–1400.

[19] S. Milborrow, J. Morkel, and F. Nicolls, "The muct landmarked face database," *Pattern recognition association of South Africa*, vol. 201, 2010.

[20] "The bioid face database," in *B. T. Research*, 2001.

[21] O. Jesorsky, K. J. Kirchbergand, and R. Frischholz, "Robust face detection using the hausdorff distance," *In Audio and Video Biom. Pers. Authentication*, p. 9095, 1992.

[22] F. Timm and E. Barth, "Accurate eye centre localization by means of gradients," in *VISAPP*, 2011, pp. 125–130.

[23] G. Anjith and A. Routray, "Fast and accurate algorithm for eye localization for gaze tracking in low resolution images," *IET Computer Vision*, 2016.

[24] M. Leo, D. Cazzato, T. De Marco, and C. Distante, "Unsupervised eye pupil localization through differential geometry and local self-similarity matching," in *In Proceedings of PLOS*, 2014.

[25] D. Cristinacce, T. Cootes, and I. Scott, "A multi-stage approach to facial feature detection," in *In Proceedings of BMVC*, 2004, p. 277286.

[26] M. Asadifard and J. Shanbezadeh, "Automatic adaptive center of pupil detection using face detection and cdf analysis," in *In Proceedings of IMECS*, 2010.