

# Binary Classifiers and Radial Symmetry Transform for Fast and Accurate Eye Localization

Pei Qin<sup>(✉)</sup>, Junxiong Gao<sup>(✉)</sup>, Shuangshuang Li, Chunyu Ma, Kaijun Yi<sup>(✉)</sup>, and Tomas Fernandes

Wuhan Hongshi Technologies Co., Ltd., Wuhan 430200, China  
{qinpei, junxiong.gao, shuangshuang.li, chunyu.ma, kaijun.yi, fernandes}@hongshi-tech.com

**Abstract.** In order to locate eyes for iris recognition, this paper presents a fast and accurate eye localization algorithm under active infrared (IR) illumination. The algorithm is based on binary classifiers and fast radial symmetry transform. First, eye candidates can be detected by the fast radial symmetry transform in infrared image. Then three-stage binary classifiers are used to eliminate most unreliable eye candidates. Finally, the mean eye template is employed to identify the real eyes from the reliable eye candidates. A large number of tests have been completed to verify the performance of the proposed algorithm. Experimental results demonstrate that the algorithm proposed in this article is robust and efficient.

**Keywords:** Binary classifiers · Fast radial symmetry transform · Mean eye template · Eye localization

## 1 Introduction

At present, iris recognition is regarded as one of the most reliable and accurate biometrics in all biological recognition technologies [1]. It will gain a wide range of application for personal identification in the public security system, military, counter-terrorism and bank payment system etc. The wide application of iris recognition could create a safer and more convenient living environment for society. In order to exploit the capability of iris recognition, it is extremely crucial to develop the eye localization technology with low complexity and high accuracy.

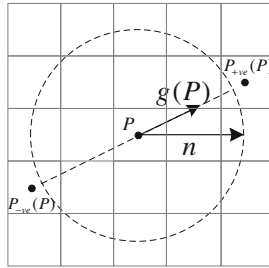
The existing technologies could be classified into three categories based on the information or patterns [2]. One of the three categories is based on measuring eye characteristics [3–6]. This type of method exploits the inherent features of eyes, such as their distinct shapes and strong intensity contrast [2]. This approach is relatively simple, however, its accuracy is highly affected by eyebrow, fringe, glass frame and illumination etc. Another type of method is based on learning statistic appearance model [2], which extracts useful visual features from a large set of training images [7, 8]. The performance of this method is limited by the training sample space and the high complexity is inevitable. The third type of method is based on structural information [9, 10]. This approach

exploits the spatial structure of interior components of eyes or the geometrical regularity between eyes and other facial features in the face context [2]. To overcome the complicated uncontrolled conditions, a statistical eye model is usually integrated into the third type of method.

For our proposed method in this paper, both eye characteristics and structural information are considered. In the next section, the theory of fast radial symmetry transform will be introduced. In Sect. 3, the proposed eye localization method will be described in detail. Assessment method of eye localization precision and experimental results are presented in Sect. 4. The last section concludes this work.

## 2 Fast Radial Symmetry Transform

Fast radial symmetry transform (FRST) [11] is regarded as the extension of general symmetry transform and a gradient-based interest descriptor in space domain. It can be used to detect high radial symmetry objects. Firstly, it is necessary to introduce two important concepts: orientation projection image  $O_n$  and magnitude projection image  $M_n$ . Orientation projection image  $O_n$  denotes the projection count accumulation of each pixel on its gradient orientation under radius  $n$ . Magnitude projection image  $M_n$  denotes the gradient magnitude accumulation of each pixel on its gradient orientation under radius  $n$ .



**Fig. 1.** The locations of positively-affected pixel  $P_{+ve}(P)$  and negatively-affected pixel  $P_{-ve}(P)$  affected by the gradient element  $g(P)$  for a range of  $n = 2$ .

As shown in Fig. 1,  $P_{+ve}(P)$  represents the positively-affected pixel location on the positive gradient direction of pixel  $P$ . Similarly,  $P_{-ve}(P)$  represents the negatively-affected pixel location on the negative gradient direction of pixel  $P$ . The coordinates of  $P_{+ve}(P)$  and  $P_{-ve}(P)$  are given by

$$P_{+ve}(P) = P + \text{round} \left( \frac{g(P)}{\|g(P)\|} n \right), \quad P_{-ve}(P) = P - \text{round} \left( \frac{g(P)}{\|g(P)\|} n \right). \quad (1)$$

Where  $\text{round}(\cdot)$  denotes the operation of getting the nearest integer of each vector element.  $g(P)$  is the gradient vector.

Orientation projection image  $O_n$  and magnitude projection image  $M_n$  should be initialized to zero [11]. For each pair of affected pixels, the values of orientation projection and magnitude projection corresponding points  $P_{+ve}(P)$  and  $P_{-ve}(P)$  are given by

$$O_n(P_{+ve}(P)) = O_n(P_{+ve}(P)) + 1, \quad O_n(P_{-ve}(P)) = O_n(P_{-ve}(P)) - 1, \quad (2)$$

$$M_n(P_{+ve}(P)) = M_n(P_{+ve}(P)) + \|g(P)\|, \quad M_n(P_{-ve}(P)) = M_n(P_{-ve}(P)) - \|g(P)\|. \quad (3)$$

For the radius  $n$ , fast radial symmetry contribution  $S_n$  is defined as the convolution

$$S_n = F_n(P) * A_n. \quad (4)$$

Where

$$F_n(P) = \frac{M_n(P)}{k_n} \left( \frac{|\tilde{O}_n(P)|}{k_n} \right)^\alpha, \quad (5)$$

$$\tilde{O}_n(P) = \begin{cases} O_n(P), & \text{if } O_n(P) < k_n \\ k_n, & \text{otherwise} \end{cases}.$$

$A_n$  is a 2-D Gaussian filter.  $\alpha$  is the radial strictness parameter.  $k_n$  is a scaling factor. The final symmetry transform image  $S$  is determined by the average of all radial symmetry contribution  $S_n$

$$S = \frac{1}{len(N)} \sum_{n \in N} S_n. \quad (6)$$

Where  $N$  is a vector which is composed of radius  $n$  considered,  $len(\cdot)$  calculates the length of the input vector.

## 3 Eye Localization Method

### 3.1 Eye Candidate Detection

Since the excellent performance of FRST on detecting radial symmetry structure, it is exploited to detect the eye candidates in the target image because of the highly symmetry irises and pupils.

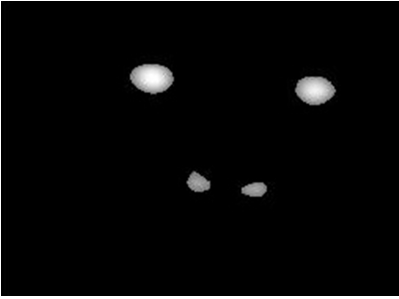
The first step of detecting candidates is to search the regions of interest (ROI) in the target image and scale each ROI to  $256 \times 256$  pixels. A ROI is a relatively small region where contains two eyes. While the visual angle of iris camera is usually very small in recognition products, so the whole target image can be an available ROI as shown in Fig. 2. If the background of the target image is extremely rich, it is necessary to employ face detection technology [15] to search ROIs in the target image.



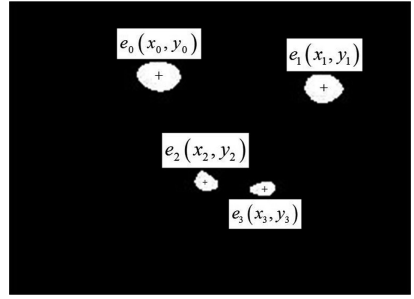
**Fig. 2.** The target image from the iris camera with small visual angle which can be a ROI (Authorized by Institute of Automation of China Academia Sciences, CASIA).



**Fig. 3.** The output image of FRST.



**Fig. 4.** The binarization result of FRST image  $S$ .



**Fig. 5.** The coordinates of eye candidates after region generation process and binarization.

Next, gradient vector  $[g_x(P)g_y(P)]^T$  of each pixel in every ROI is computed by  $3 \times 3$  Sobel operator. Then each ROI can be processed by the FRST according to Sect. 2 and the output image  $S$  is shown in Fig. 3. The considered radii values can be set according to the size of ROI. Here, we set the vector  $N$  of considered radius to be  $\{7, 9, 11, 12\}$ .

In the FRST output image  $S$ , it is obvious that there are some bright regions. These bright regions contain eyes, nostrils, nose pads of glasses etc. In order to eliminate the background from the bright regions, we utilize the binarization method to process the image  $S$  and the processing result is shown in Fig. 4. Then region generation skill is employed to generate the coordinate of the middle pixel of each eye candidate region as shown in Fig. 5. The coordinate vector of eye candidates could be denoted as  $\{e_0(x_0, y_0), \dots, e_{M-1}(x_{M-1}, y_{M-1})\}$ , where  $M$  is the number of eye candidates.

### 3.2 Binary Classifiers

Although it is efficient and reliable to detect eye candidates by FRST, binarization and region generation, it also remains a great challenge to identify the eye(s) accurately from the candidates. Because the candidates may also be nostrils, nevus, nose pads of glasses etc.

Three-stage binary classifiers are designed to further exclude unreliable eye candidates. The binary classifiers are denoted as  $C_k(w_k, h_k)$ ,  $k \in \{0, 1, 2\}$  shown in Fig. 6, where  $w_k$  and  $h_k$  are respectively the width and height of classifier  $C_k$ . Classifier  $C_0$  is only a  $2n_{\max} \times 2n_{\max}$  square and classifier  $C_1$  is composed of three  $2n_{\max} \times 2n_{\max}$  squares, where black color represents “1” and white color represents “0”,  $n_{\max}$  is the maximum element of radius value vector  $N$ . Classifier  $C_2$  is a little bit complex and this paper only describes the black region for simplicity. The black region is composed of a  $2n_{\max} \times 2n_{\max}$  square, two  $n_{\max} \times 0.5n_{\max}$  rectangles and two  $0.25n_{\max} \times 0.25n_{\max}$  squares. Here  $n_{\max}$  is equal to 12.

These binary classifiers are used to classify the binary eye candidates. Then, binarization transform is independently processed with dynamic threshold for each ROI, not for the whole target image. Figure 7 shows the result of ROI binarization transform. Then the eye candidates  $\{e_0(x_0, y_0), \dots, e_{M-1}(x_{M-1}, y_{M-1})\}$  are classified respectively by the three-stage binary classifiers. And the steps of classification is shown below.

Firstly, segment the eye candidate regions  $B_{i,k}(s_{i,k}, t_{i,k}, w_{i,k}, h_{i,k})$ ,  $i \in [0, M-1]$ ,  $k \in \{0, 1, 2\}$  from the ROI binary image, where  $(s_{i,k}, t_{i,k})$ ,  $w_{i,k}$  and  $h_{i,k}$  are respectively the left corner coordinate, width and height of the eye candidate regions. The width  $w_{i,k}$  and height  $h_{i,k}$  are the same as these of classifier  $C_k$ . The left corner coordinates of eye candidate regions  $B_{i,k}$  are given by

$$\begin{aligned} S_{i,0} &= x_i - n_{\max}, & t_{i,0} &= y_i - n_{\max}; \\ S_{i,1} &= x_i - n_{\max}, & t_{i,1} &= y_i - 5n_{\max}; \\ S_{i,2} &= \text{round}(x_i - 1.625n_{\max}), & t_{i,2} &= y_i - n_{\max}. \end{aligned} \quad (7)$$

Where  $1.625n_{\max}$  is a half of the black region width in Classifier  $C_2$  and the mean of  $\text{round}(\cdot)$  is the same with Eq. (1). If  $s_{i,k} < 0$  or  $t_{i,k} < 0$ , the eye candidate  $e_i(x_i, y_i)$  should be rejected.

Secondly, calculate eigenvalue difference  $\Delta v_{k,i}$ ,  $i \in [0, M-1]$ ,  $k \in \{0, 1, 2\}$ . Eigenvalue difference  $\Delta v_{k,i}$  is defined as

$$\Delta v_{k,i} = v_{k,i} - \min\{v_{k,i}\}. \quad (8)$$

Where  $v_{k,i}$  is given by

$$v_{k,i} = C_k \times B_i. \quad (9)$$

Thirdly, compare the eigenvalue difference  $\Delta v_{k,i}$  with the threshold  $\lambda_k$ . If  $\Delta v_{k,i} < \lambda_k$ , it represents the eye candidate  $e_i(x_i, y_i)$  passes the classification of classifier  $C_k$ .

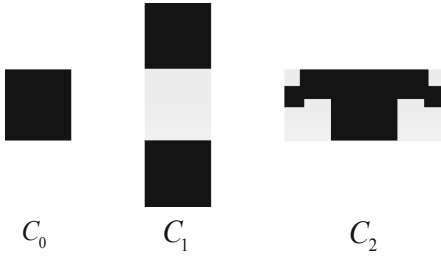


Fig. 6. Three binary classifiers.



Fig. 7. Output of ROI binarization transform.

Suppose there are  $Q$  eye candidates passed the classification of the three-stage classifiers, which are denoted as  $\{\tilde{e}_0(x_0, y_0), \dots, \tilde{e}_{Q-1}(x_{Q-1}, y_{Q-1})\}$ .

### 3.3 Eye Localization by IMED Matching

After three-stage classification, one or more non-eye candidates maybe also exist. To detect the real eyes from  $\{\tilde{e}_0(x_0, y_0), \dots, \tilde{e}_{Q-1}(x_{Q-1}, y_{Q-1})\}$ , template matching is employed which is based on intensity appearance [12]. This approach is relatively reliable and does not require any training. The template is obtained by the average result of 200 eye samples with  $4n_{\max} \times 4n_{\max}$  pixels cut from CASIA-Iirs-Distance database (CASIA-Iirs-Distance database is authorized by CASIA). Suppose  $S_n^e$  is the eye samples cut from ROIs of database images, the mean eye template  $T_e$  is given by

$$T_e = \frac{1}{200} \sum_{n=1}^{200} S_n^e. \quad (10)$$

Finally, the real eyes are identified through the similarity between eye candidates and eye template  $T_e$ . Image Euclidean Distance (IMED) [13] is proved to be an excellent assessment of image similarity. Because it can reduce mismatching caused by eye rotation or deformation. The IMED  $d_r^2(G_e, G_T)$  between eye template and eye candidates is calculated by

$$d_r^2(G_e, G_T) = \frac{1}{2\pi} \sum_{i,j=1}^{WH} \exp\left\{-|P_i - P_j|^2/2\right\} (g_e^i - g_T^i)(g_e^j - g_T^j). \quad (11)$$

Where  $G_T$  and  $G_e$  are the vectors  $\{g_T^1, g_T^2, \dots, g_T^{WH}\}$  and  $\{g_e^1, g_e^2, \dots, g_e^{WH}\}$  which are respectively composed of all the pixels of eye template  $T_e$  and each eye candidate region  $D_r(s, t, w, h)$ ,  $r \in [0, Q - 1]$ .  $s, t, w$  and  $h$  are given by

$$s = x_i - 2 \times n_{\max}, t = y_i - 2 \times n_{\max}, w = 4 \times n_{\max}, h = 4 \times n_{\max}. \quad (12)$$

These eye candidate regions are different from the ones used to be classified.

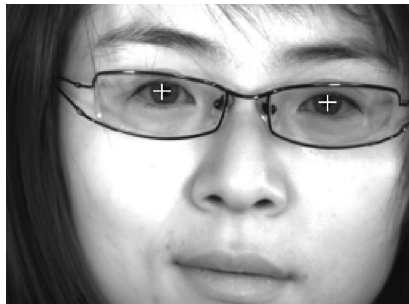
Smaller image Euclidean distance means higher similarity. So it is easy to search two real eyes from the IMED set  $d_r^2$ ,  $r \in [0, Q - 1]$ . Table 1 shows the proposed eye localization algorithm framework. And the result of eyes localization is shown in Fig. 8.

**Table 1.** Eye localization algorithm framework.

---

- Input a target image  $I$  ;
- Initialize  $N = \{7, 9, 11, 12\}$ ,  $[\lambda_0, \lambda_1, \lambda_2] = [35, 70, 85]$ ;
- Search ROIs in the target image  $\{R_0, \dots, R_{p-1}\}$  and scale each ROI to  $256 \times 256$  pixels;
- for  $r = 0, \dots, P-1$  :
  1. Detect the eye candidates  $e_i(x_i, y_i)$ ,  $i \in [0, M-1]$  in the ROI  $R_r$  ;
  2. for  $i = 0, \dots, M-1$  :
    - (1). Binarization transform for ROI  $R_r$  ;
    - (2). Segment the eye candidate regions  $B_{i,k}(s_{i,k}, t_{i,k}, w_{i,k}, h_{i,k})$ , then reject the eye candidate  $e_i(x_i, y_i)$  and jump to next loop if  $s_{i,k} < 0$  or  $t_{i,k} < 0$  ;
    - (3). Calculate eigenvalue difference  $\Delta v_{k,i}$  ;
    - (4). Compare the eigenvalue difference  $\Delta v_{k,i}$  with  $\lambda_k$ , then put the eye candidate  $e_i(x_i, y_i)$  into the eye candidate set of passed classification if  $\Delta v_{k,i} < \lambda_k$  .
  3. Template matching to locate the real eyes.

---



**Fig. 8.** The result of eyes localization.

## 4 Experimental Results

### 4.1 Assessment Method of Eye Localization Precision

In order to assess the precision of the proposed method, the most commonly used measurement proposed in [14] is used in this article, which is given by

$$d_{eye} = \frac{\max(d_l, d_r)}{\|C_l - C_r\|}. \quad (13)$$

Where  $C_l$  and  $C_r$  are the true eye center.  $d_l$  and  $d_r$  are the Euclidean distances between the detected eye center and the true eye center.

### 4.2 Results and Analysis

The proposed method is tested on the JAFFE and CASIA-Iirs-Distance databases for its performance. The database of CASIA-Iirs-Distance is from the Center for Biometrics and Security Research of CASIA. There are 2576 images in the database which are taken under active infrared (IR) illumination. People on the images have different postures and some wear glasses.

**Table 2.** The experimental results on JAFFE database.

| Method                            | Accuracy        |                  |
|-----------------------------------|-----------------|------------------|
|                                   | $d_{eye} < 0.1$ | $d_{eye} < 0.25$ |
| Multi-view eyes localization [16] | 60.9 %          | 98.6 %           |
| AdaBoost + SVM [17]               | 99.5 %          | 100 %            |
| Proposed method                   | 96.82 %         | 98.90 %          |

**Table 3.** The experimental results on CASIA-Iirs-Distance database.

| Method                      | Accuracy                                 |
|-----------------------------|--|
| AdaBoost [15]               | 89.74 % (with glasses + without glasses) |
| SVM [3]                     | 95.50 % (with glasses + without glasses) |
| Appearance model + GST [18] | 99.4 % (without glasses)                 |
|                             | 88.3 % (with glasses)                    |
| Proposed method             | 96.26 % (with glasses + without glasses) |

The experimental results on JAFFE database are shown in the Table 2. The proposed method performance is superior to multi-view eyes localization method [16], and worse than Tang's [17]. Tang's method employs the Adaboost and SVM skills to obtain excellent performance, but it introduces a large amount of computation and complex training. So it is extremely difficult to achieve real-time eye localization.



The methods proposed by Zhu [3] and Zhao [18] can be compared with the method proposed in this article since they also exploit the active infrared illumination. Table 3 shows the eye localization performance of proposed method when comparing with other representative ones. From Table 3, the proposed method is superior to other representative methods except the one proposed by Zhao. But Zhao's method is sensitive to the interference of glasses.

## 5 Conclusions

A novel method based on binary classifiers and fast radial symmetry transform is proposed to improve the accuracy rate of eye localization under active infrared (IR) illumination. It is able to efficiently overcome the detrimental interference from eyebrow, fringe, glass frame and the reflected light from glasses etc. Because besides the geometric structure and intensity information, binary information is also considered in this research. From the experimental results, the proposed method is proved to be superior to most of other states-of-art methods. It is worth mentioning that the complexity degree of our method is highly reduced because the training step is eliminated. Therefore the proposed method could achieve a better tradeoff between eye localization performance and complexity.

## References

1. Sachin, G., Chander, K.: Iris recognition: the safest biometrics. *Int. J. Eng. Sci.* **4**, 265–273 (2011)
2. Song, F., Tan, X., Chen, S., Zhou, Z.H.: A literature survey on robust and efficient eye localization in real-life scenarios. *J. Pattern Recogn.* **46**(12), 3157–3173 (2013)
3. Zhu, Z., Fujimura, K., Ji, Q.: Real-time eye detection and tracking under various light conditions. In: *Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, vol. 25, pp. 139–144 (2002)
4. Yuille, A.L., Cohen, D.S., Hallinan, P.W.: Feature extraction from faces using deformable templates. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 104–109 (1989)
5. Feng, G.C., Yuen, P.C.: Variance projection function and its application to eye detection for human face recognition. *Pattern Recogn. Lett.* **19**(9), 899–906 (1998)
6. Kothari, R., Mitchell, J.L.: Detection of eye locations in unconstrained visual images. In: *Proceedings of International Conference on Image Processing*, vol. 3, pp. 519–522 (1996)
7. Everingham, M.R., Zisserman, A.: Regression and classification approaches to eye localization in face images. In: *Proceedings of IEEE Conference on Automatic Face and Gesture Recognition*, pp. 441–448 (2006)
8. Wang, P., Green, M., Ji, Q., Wayman, J.: Automatic eye detection and its validation. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 3, pp. 164–172 (2005)
9. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. *Comput. Vis. Image Underst.* **61**(1), 38–59 (1995)

10. Tan, X., Song, F., Zhou, Z.H., Chen, S.: Enhanced pictorial structures for precise eye localization under uncontrolled conditions. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1621–1628, 20–25 June 2009
11. Loy, G., Zelinsky, A.: Fast radial symmetry for detecting points of interest. IEEE Trans. PAMI **25**(8), 959–973 (2003). IEEE Press
12. Li, W., Wang, Y., Wang, Y.: Eye location via a novel integral projection function and radial symmetry transform. Int. J. Digital Content Technol. Appl. **5**(8), 70–80 (2011)
13. Wang, L., Zhang, Y., Feng, J.: On the euclidean distance of images. IEEE Trans. PAMI **27**(8), 1334–1339 (2005). IEEE Press
14. Jesorsky, O., Kirchberg, K., Frischholz, R.: Robust face detection using the hausdorff distance. In: Proceedings of International Conference on Audio and Video-Based Biometric Person Authentication, pp. 90–95 (2001)
15. Viola, P., Jones, M.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
16. Fu, Y., Yan, H., Li, J., Xiang, R.: Robust facial features localization on rotation arbitrary multi-view face in complex background. J. Comput. **6**(2), 337–342 (2011)
17. Tang, X., Ou, Z., Su, T., Sun, H., Zhao, P.: Robust precise eye location by AdaBoost and SVM techniques. In: Proceedings of International Conference on Advances in Neural Networks, pp. 93–98 (2005)
18. Zhao, S., Grigat, R.: Robust Eye Detection under active infrared illumination. In: Proceedings of International Conference on Pattern Recognition, vol. 4, pp. 481–484 (2006)