Robust Eye Detection under Active Infrared Illumination

Shuyan Zhao, Rolf-Rainer Grigat
Vision Systems, 4-08/1
Hamburg University of Technology
Harburger Schloßstr. 20, 21079 Hamburg, Germany
s.zhao@tu-harburg.de, grigat@tu-harburg.de

Abstract

Eye detection is very important for automatic face recognition and gaze tracking. In this paper we propose an algorithm for eye detection under active infrared (IR) illumination. A simple hardware enables us to make use of a physiological property of the eyes. A new thresholding method is introduced in order to effectively search the regions of interest (ROI). An appearance model is then used to verify the pupil candidates. However, the existence of eyeglasses has a negative effect on selection of candidates. Regarding this the generalized symmetry transform (GST) is exploited. By using a simplified distance weight, we reduce the computational cost of the original transform. Experimental results demonstrate the effectiveness of the proposed eye detection method.

1 Introduction

Eye detection has a wide variety of applications in the field of computer vision. For example, eye detection can be used for gaze tracking, which can be further used as a Human Computer Interface (HCI). To achieve an automatic face recognition system, eye detection is a significant step, because eye positions are generally used to align and normalize the face. Existing technologies of eye detection include deformable templates [10], evolutionary computation [5], eigenfeatures [6], etc. However, computational efficiency remains a challenge.

When an infrared lighting source is placed close to the axis of the camera, the pupils usually look unnaturally bright because of the reflection of the blood-rich retina. This is the well known *bright pupil effect*, which has been used for eye detection and tracking [4]. We make use of this physiological property of eyes also. This kind of approaches needs a bright image and a dark image, and carries out pupil thresholding on the difference image. With the aid of the Euler number, our thresholding method makes a good trade-off of accuracy and computational complexity. An appearance model is then used to verify the candidates, which outperforms the approach of template matching [11]. Nevertheless, when a person wears glasses, the frame of glasses often looks even brighter than the pupils. In this case the true eyes may be excluded from the candidates. In [12, 9], methods for detecting and removing glasses have been proposed, however, they assume that the face has been detected and the eye positions are known, which does not match our intention, i. e. eye detection robust against glasses. We then apply the generalized symmetry transform [8] to search for ROI, which requires no a priori knowledge. We improve it by using a simplified distance weight.

Our algorithm is shown in Fig. 1. In the next section, a new thresholding method will be introduced. In Section 3, the appearance-based verification approach will be described. The generalized symmetry transform will be presented in Section 4. Experimental results are shown in Section 5, finally the last section concludes this work.

2 Thresholding

The existing approach [4] used an interlacing camera equipped with two lighting sources, one on the axis and the other off the axis of the camera. The camera generated images with bright and dark pupils alternatively by switching on and off the on-axis and off-axis lighting sources. In contrast, we use a progressive camera and only one on-axis lighting source [11]. In order to obtain stable illumination conditions, an IR filter is used to block the visible light. Our camera always produces images with bright pupils (shown in Fig. 2-(a)). By applying morphological opening operation [7], the bright spots on the pupils are removed (shown in Fig. 2-(b)). The difference image of the input and the image after the opening operation gives clues about possible eye positions (shown in Fig. 2-(c)). After thresholding the difference image, we take the connected components in the resulting binary image as pupil candidates (see Fig. 2-(d)).

The thresholding algorithm directly determines the num-

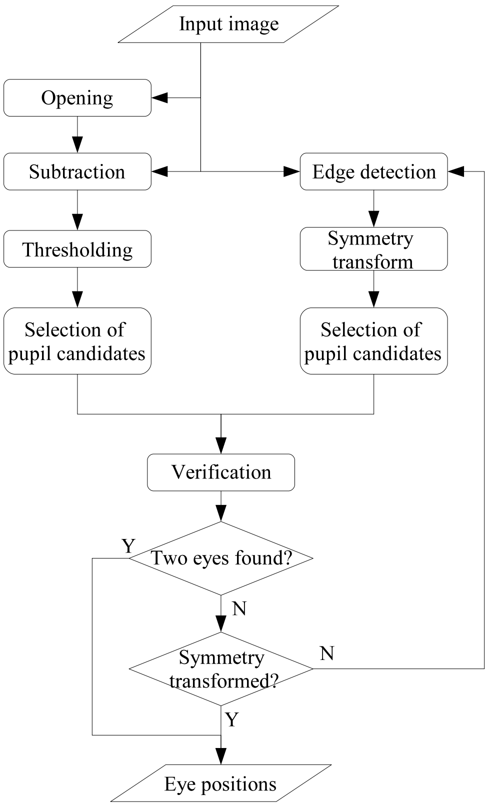
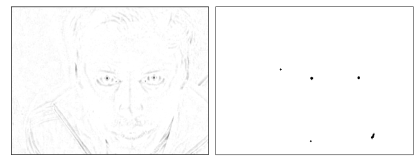


Figure 1. The presented eye detection approach

(a) Input image (b) Opening operation



(c) Difference image (inverted and contrast stretched) (d) Binary image (inverted)

Figure 2. Bright pupil effect

ber of pupil candidates. Too few candidates might result in one or two pupils missed, while too many candidates lead to high computational cost for the following process. Euler number is a topological feature, which is defined as the number of objects in the image minus the number of holes in those objects. It can be calculated very efficiently with $O(N)$ [1]. Our thresholding method integrates Euler number calculation with simple global thresholding:

1. Select an initial threshold T, $T = s \cdot L_{D_{\max}}$, where $L_{D_{\max}}$ is the maximum intensity of the difference image, $0 \leq s \leq 1$ is a scaling factor.
2. Threshold the image using T
3. Calculate the Euler number n_e of the binary image
    - If $s < 0$ or $s > 1$, stop
    - If $n_{e_{\min}} \leq n_e \leq n_{e_{\max}}$, stop
    - If $n_e < n_{e_{\min}}$, then $s \Leftarrow s - \Delta s$, go to Step 2
    - If $n_e > n_{e_{\max}}$, then $s \Leftarrow s + \Delta s$, go to Step 2

The advantage of our method is that the number of pupil candidates can be controlled by parameters $n_{e_{\min}}$ and $n_{e_{\max}}$. Thanks to the Euler number, a simple global thresholding strategy is sufficient.

3 Verification

In [11] we used circle matching to verify pupil candidates, which has difficulties in distinguishing pupils from the interior corner of the eye and some background noise. Seeing that in this paper we utilize an appearance model, which is created by using Principal Component Analysis (PCA) [6] and Support Vector Machines (SVM) [2].

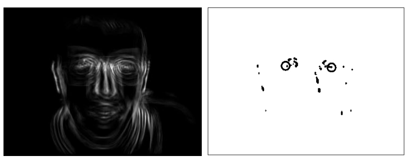
So far we know there are no public infrared face sets available, thus we have to collect the training data ourselves. We captured 850 eye images under IR illumination in our laboratory. The data pool consisted of the original images and their mirrored copies. Firstly 800 eye images were used to train an eye space based on PCA. We then manipulated the other 900 eye images by rotating them by $\pm 5°$, $\pm 10°$ and scaling them to 110%. The resulting 5400 eye images and a non-eye set including 800 images were employed to train a SVM model with a RBF kernel. The 800 representative non-eye examples were obtained by bootstrapping [3].

With the appearance model, multiscale verification is performed. In order to further reduce false positives, template matching and distance checking are exploited.

4 Selection of Pupil Candidates Based on Symmetry Transform

Although the *bright pupil effect* is a powerful feature for pupil candidate selection, it becomes less dominant when the subject wears glasses (see Fig. 3), because the reflection of glasses and frames is even brighter. Generalized symmetry transform [8] can be used to detect ROI without any a priori knowledge. Nevertheless, due to its context-free characteristic, the search for symmetry points is time consuming. Thus only when the so-far mentioned approach fails, this module will be invoked (see Fig.1). In order to reduce the computational complexity, we use the Euclidean distance and replace the original Gaussian weight with a dual weight $D_d(i, j)$ (see Eq. 2).

(a) Input image (b) Pupil candidates with thresholding



(c) Symmetry map (d) Pupil candidates with ST

Figure 3. Selection of pupil candidates (correct pupils are circled)

The symmetry transform is applied after edge detection [7]. Using the nomenclature from [8], we have $p_k = (x_k, y_k) \in I_E$, where I_E is the edge map of the original image. Suppose ∇p_k is the gradient of the intensity at p_k, $r_k = \log(1 + \|\nabla p_k\|)$ and $\theta_k = \arctan(\frac{\partial p_k}{\partial_y} / \frac{\partial p_k}{\partial_x})$. Let $\alpha_{i,j}$ be the angle between the horizontal axis and the line formed by p_i and p_j, we have the following definitions:

$$\Gamma(p) = \{(i,j)|p = \frac{p_i + p_j}{2},\ p_i \in I_E,\ p_j \in I_E\} \tag{1}$$

$$D_d(i,j) = \begin{cases} 1 & \|p_i - p_j\| \leq d \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$P(i,j) = (1 - \cos(\theta_i + \theta_j - 2\alpha_{ij}))(1 - \cos(\theta_i - \theta_j)) \tag{3}$$

The contribution of points p_i and p_j is

$$C(i,j) = D_d(i,j)P(i,j)r_i r_j \tag{4}$$

We consider only the symmetry magnitude $M_d(p)$ of p

$$M_d(p) = \sum_{(i,j)\in\Gamma(p)} C(i,j) \tag{5}$$

The points that have high symmetry magnitude are possible pupil candidates. By binarizing the symmetry map and filtering out the large components, we obtain the pupil candidates shown in Fig. 3-(d).

5 Experiments

In order to evaluate the presented algorithm, we collected three data sets, shown in Table 1. The validation sets Set1 and Set2 contain images without and with eyeglasses respectively. The test set Set3 includes images of different subjects, which were captured under a different IR lighting condition compared to Set1 and Set2. 98 images of Set3 contain glasses, 234 images contain no glasses. All images in the three sets have resolution of 320×240. Figure 4 illustrates some examples from the various data sets.

Table 1. Data sets

Data set	Size	Function
Set1	400	Validation without glasses
Set2	150	Validation with glasses
Set3	332	Test

Figure 4. Examples from different data sets. Row 1 from Set1, Row 2 from Set2, and Row 3 from Set3.

First we analyze our thresholding algorithm on Set1 and Set2 in terms of accuracy and the number of iterations, where accuracy means two true pupils are included in the pupil candidates. Suppose that the initial value of s is $s_0 = 0.4$, $\Delta s = 0.05$, and $n_{e_{\min}} = 2$. Figure 5 and 6 show the performance on Set1 and Set2 respectively. $n_{e_{\max}}$ should be larger for images with eyeglasses than images without glasses. By taking into account the accuracy and the number of iterations, we set $n_{e_{\max}} = 12$ for the following experiments. With $n_{e_{\max}} = 12$, less than 3 iterations at average are needed, and since the Euler number can be calculated with $O(n)$, thresholding is a very fast process.

Without utilizing the generalized symmetry transform, we then compare the two verification approaches, circle matching [11] and appearance model. The evaluation is performed on Set3 in terms of hit rate and the number of false acceptances (FA). Table 2 shows that appearance-based verification outperforms the circle matching on the subset without glasses. But their hit rates on the images with glasses are

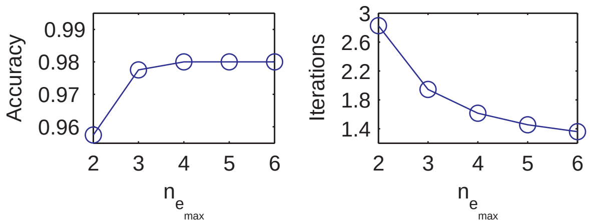


Figure 5. Performance of Thresholding on Set1 (without glasses)

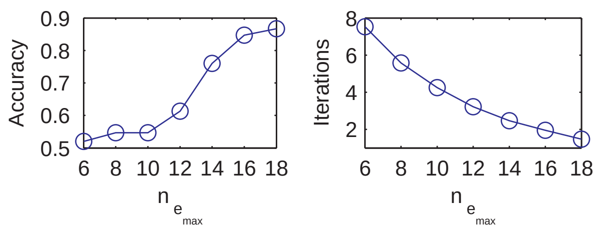


Figure 6. Performance of Thresholding on Set2 (with glasses)

similarly unsatisfying. After integrating the GST with the appearance-based verification as shown in Fig. 1, the performance is greatly improved, see Table 3.

Table 2. Comparison of circle matching and appearance model on Set3

Subsets of Set3	Methods	Hit	FA
Without glasses	Circle matching	84.4 %	71
Without glasses	Appearance model	99.4 %	0
With glasses	Circle matching	50.5 %	89
With glasses	Appearance model	54.1 %	3

6 Conclusion

In this paper we propose an eye detection approach working under active IR illumination. Based on a sophisticated thresholding algorithm, pupil candidates can be effectively selected. An appearance model, trained using PCA and SVM, is exploited to verify the pupil candidates, which performs better than our previous approach of circle matching. To increase the robustness against eyeglasses, the generalized symmetry transform is incorporated, which is improved by using a simplified distance weight. The experimental results demonstrate the effectiveness of the presented eye detection approach.

Table 3. The effect of GST

Subsets of Set3	Methods	Hit	FA
Without glasses	Appearance model	99.4 %	0
Without glasses	Appearance model + GST	99.4 %	0
With glasses	Appearance model	54.1 %	3
With glasses	Appearance model + GST	88.3 %	14

Acknowledgement

This work is part of the project *KATO – Innovative Cabin Technologies* funded by the German Federal Ministry of Economics and Technology.

References

[1] J. Arvo. *Graphics Gems II*. Boston: Academic Press, 1991.

[2] C. C. Chang and C. J. Lin. libsvm, 2005. `http://www.csie.ntu.edu.tw/~cjlin/libsvm/`.

[3] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. A Wiley-Interscience Publication, second edition, 2001.

[4] A. Haro, M. Flickner, and I. Essa. Detecting and tracking eyes by using their physiological properties, dynamics, and appearance. In *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 2000.

[5] J. Huang and H. Wechsler. Visual routines for eye location using learning and evolution. *IEEE Transactions on Evolutionary Computation*, 4(1):73–82, 1999.

[6] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings of International Conference on Computer Vision and Pattern Recognition*, 1994.

[7] W. K. Pratt. *Digital Image Processing: PIKS inside*. A Wiley-Interscience publication, third edition, 2001.

[8] D. Reisfeld and H. W. und Y. Yeshurun. Context free attentional operators: the generalized symmetry transform. *International Journal of Computer Vision*, 14:119–130, 1995.

[9] C. Wu, C. Liu, H.-Y. Shum, Y.-Q. Xu, and Z. Zhang. Automatic eyeglasses removal from face images. *IEEE Transaction on Pattern Recognition and Machine Intelligence*, 26(3):322–336, 2004.

[10] A. Yuille, D. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 104–109, New York, 1989.

[11] S. Zhao and R. Grigat. An automatic face recognition system in the near infrared spectrum. In *Proceedings of International Conference on Machine Learning and Data Mining*, pages 437–444, Leipzig, Germany, 2005.

[12] J. Zhong and R. Mariani. Glasses detection and extraction by deformable contour. In *Proceedings of Internatinal Conference on Pattern Recognition*, 2000.