Eye Detection in Facial Images Using Zernike Moments with SVM

Hyoung-Joon Kim and Whoi-Yul Kim

ABSTRACT—An eye detection method for facial images using Zernike moments with a support vector machine (SVM) is proposed. Eye/non-eye patterns are represented in terms of the magnitude of Zernike moments and then classified by the SVM. Due to the rotation-invariant characteristics of the magnitude of Zernike moments, the method is robust against rotation, which is demonstrated using rotated images from the ORL database. Experiments with TV drama videos showed that the proposed method achieved a 94.6% detection rate, which is a higher performance level than that achievable by the method that uses gray values with an SVM.

Keywords—Eye detection, Zernike moments, support vector machine (SVM).

I. Introduction

Eye detection has been widely studied for use in diverse applications including drowsiness detection for intelligent vehicle systems, gaze detection for human-machine interfaces, and face normalization (or alignment) for automatic face recognition systems [1]. A support vector machine (SVM), which is a binary classification method, has been successfully applied to the detection and verification of human eyes [2]-[5]. The SVM determines the presence/absence of eyes using an input vector consisting of gray values in a moving window. In the case of rotated face images, the method often fails to detect eyes because such images are inconsistent with the training image set. This does not present a problem for authentication applications with the constraint that a human face must be in an upright position. However, in order to be widely applicable to photo albums [6] and automatic video management systems [7], eye detection methods must be able to detect human eyes even in rotated faces.

This paper presents an eye detection method for facial images using Zernike moments [8] and the SVM. Specifically, the method exploits the rotation-invariant characteristics of the magnitude of Zernike moments in order to represent eye/non-eye patterns. This rotation-invariant feature allows the proposed method to detect eyes even if a face is rotated. The magnitudes of Zernike moments, extracted from an eye/non-eye training dataset, are learned by the SVM in order to successfully determine the presence/absence of eyes in a facial image.

II. Eye Detection

1. Eye-Candidate Detection

An efficient way to detect an eye region in a facial image is to find the dark regions. To find these regions, a 7×7 minimum filter is first applied to the facial image. The filtered image is then binarized using half of the average intensity of the filtered image as a threshold. Figure 1 shows the process of detecting eye-candidate regions. In order to reduce the undesirable effect from hair or background, only pixels of the facial image corresponding to black regions within a circle (as shown in Fig. 1(c)) are considered as eye candidates.

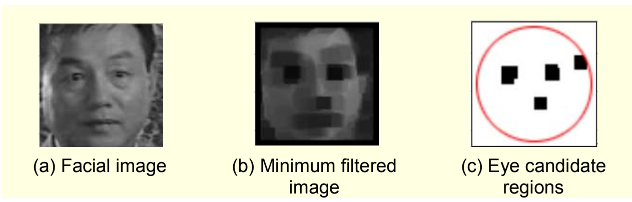

Fig. 1. Process of eye-candidate detection.

Manuscript received July 8, 2007; revised Jan. 12, 2008.
Hyoung-Joon Kim (phone: + 82 10 9270 8308, email: khjoon@gmail.com) was with the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Rep. of Korea and is with the Samsung Electronics Co., Ltd., Gyeonggi-do, Rep. of Korea.
Whoi-Yul Kim (email: wykim@hanyang.ac.kr) is with the Department of Electronics and Computer Engineering, Hanyang University, Seoul, Rep. of Korea.

2. Eye Detection

The magnitude of Zernike moments is used as a rotation-invariant feature to represent eye/non-eye patterns (see [8] for further explanation of Zernike moments). Zernike moments are complex numbers by which an image is mapped onto a set of two-dimensional complex Zernike polynomials. The orthogonal property of Zernike polynomials enables the contribution of each moment to be unique and independent of the information in an image.

To cope with faces of different sizes in an image, a hierarchical approach has been adopted. That is, the facial image is scaled down by a factor of 1.2 consecutively. Then, Zernike moments are computed on a 15×15 window centered at each eye-candidate pixel from the scaled down images. For fast computation of Zernike moments, a lookup table of Zernike basis functions is created beforehand utilizing their symmetric/anti-symmetric characteristics [8]. The extracted Zernike moments are then normalized by (1) to render them less sensitive to illumination changes [9]:

$$X_{n,m} = \frac{|Z_{n,m}|}{|Z_{0,0}|}, \qquad (1)$$

where $|Z_{n,m}|$ is the magnitude of the Zernike moment of order n with repetition m satisfying n-$|m|$=(even) and $|m| \leq n$. The highest order n is empirically set to 15, and 72 Zernike moments are used.

Using the normalized moments as input vectors, excluding $X_{0,0}$, the SVM determines whether each window is an eye. Let $\mathbf{x}_{k,x_k,y_k}$ be the column vector consisting of the normalized moments extracted from the window centered at (x_k, y_k) in the k-th scaled image. According to the trained SVM, which uses the radial basis function kernel [2], the vector $\mathbf{x}_{k,x_k,y_k}$ is classified as an eye if the output of the SVM $f(\mathbf{x}_{k,x_k,y_k})$ is positive.

3. Determination of Eye Positions

It is necessary to precisely determine the location of an eye because often more than one eye region is detected for each real eye, as shown in Fig. 2(a). To determine eye positions, an eye probability map (EPM) is employed. The EPM p_k for the k-th scaled image is the sum of the Gaussian function weighted by $f(\mathbf{x}_{k,x_k,y_k})$:

$$p_k(x,y) = \sum_{f(\mathbf{x}_{k,x_k,y_k})>0} f(\mathbf{x}_{k,x_k,y_k}) \exp\left(-\frac{(x-\hat{x}_k)^2}{\sigma_k^2} - \frac{(y-\hat{y}_k)^2}{\sigma_k^2} \right), \qquad (2)$$

where $\hat{x}_k = 1.2^k x_k$, $\hat{y}_k = 1.2^k y_k$, and $\sigma_k = 1.2^k \alpha$, and α is empirically set to 4. Then, the EPM p for the given facial image is the sum of all p_ks, that is, $p(x,y) = \sum_k p_k(x,y)$.

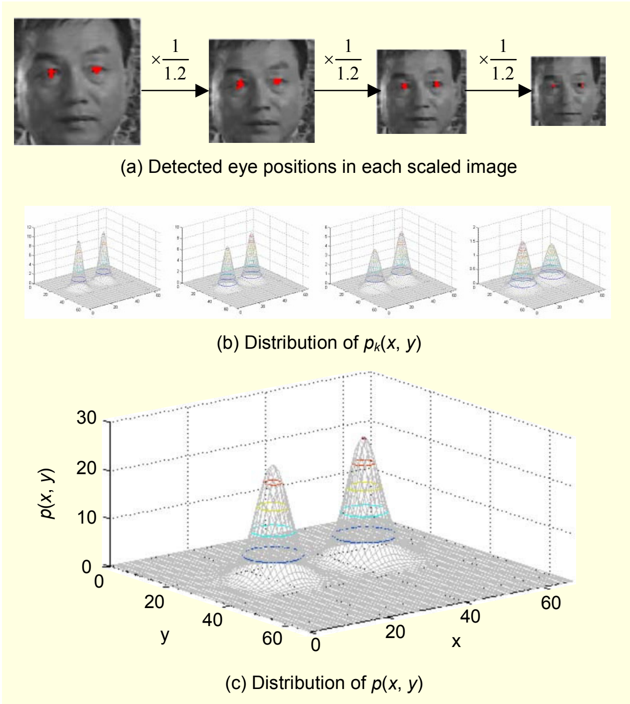

Fig. 2. Eye probability map.

An example of the EPM p is shown in Fig. 2(c). From the EPM p, all of the local maximums are found by gradient descent, and finally, the positions of the left and right eye are determined by selecting the two positions having the highest values among the found local maximums.

III. Experimental Results

In training the SVM, 1,382 eye images and 1,314 non-eye images of 15×15 pixels were used, as shown in Fig. 3. For the eye image set, the eye regions were cropped from 691 faces in 317 images gathered from the Internet. Bootstrapping was applied to build the non-eye image set [2].

First, the performance of the method was evaluated using 400 92×112 facial images contained in the ORL database. To demonstrate the method's robustness to rotation, four sets of images, rotated by 10, 30, 60 and 90 degrees, were prepared. Table 1 shows the performance of the method (denoted by

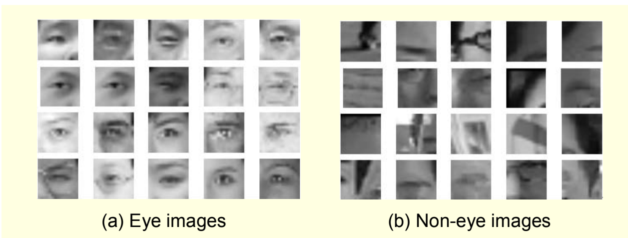

Fig. 3. Examples of eye and non-eye images.

Table 1. Eye detection results for images from ORL database.

Rotation angle	Gray/SVM		Zernike/SVM	
	Detection rate	False alarm	Detection rate	False alarm
0°	91.00% (2.14)	0.38%	91.88% (2.32)	0.88%
10°	79.00% (2.64)	1.13%	89.50% (3.25)	1.38%
30°	40.38% (3.21)	0.75%	78.25% (3.37)	1.13%
60°	30.00% (3.94)	2.25%	78.00% (3.22)	1.13%
90°	48.50% (4.25)	2.5%	91.25% (2.86)	0.63%

Table 2. Eye detection results on TV drama videos.

Method	Detection rate	False alarm	Time (P-IV 3.4 GHz)
Gray/SVM	82.87% (1.84)	3.69%	710 ms
Zernike/SVM	94.61% (1.94)	4.28%	187 ms


Fig. 4. Eye detection results.

Zernike/SVM) compared with the method that uses 225 (=15×15) gray values as the input vector for the SVM (denoted by Gray/SVM) [5]. The numbers in parentheses indicate the average error between the detected eye position and the ground truth. Due to the rotation-invariant characteristics of Zernike moments, the Zernike/SVM method showed better performance at all angles than the Gray/SVM method, but with a slightly larger error with some rotated-image sets.

Second, the method was tested on 606 352×240 frame images captured from TV drama videos. Of these images, 677 faces were detected by Viola's face detection method [10] implemented in Open CV [11]. Although Viola's method was trained only for frontal face detection, it was able to detect slightly rotated faces. The Zernike/SVM eye detection method was applied to the same 677 detected faces. Table 2 shows that the proposed method achieved a higher detection rate with a faster processing time. The detected results are shown in Fig. 4, where the crosses and rectangles indicate the location of the detected eyes and faces, respectively.

IV. Conclusion

An eye detection method exploiting Zernike moments with an SVM was presented. The method utilizes the magnitude of Zernike moments to represent eye/non-eye patterns, employing the SVM as a classifier. Due to the rotation-invariant characteristics of the magnitude of Zernike moments, the proposed method can detect eyes well, even if a face has been rotated. The experimental results confirm that the proposed method, in comparison with the method using gray values with an SVM, offers superior performance.

References

[1] P. Wang and Q. Ji, "Multi-view Face and Eye Detection Using Discriminant Features," *Computer Vision and Image Understanding*, vol. 105, no. 2, 2007, pp. 99-111.

[2] E. Osuna, R. Freund, and F. Girosit, "Training Support Vector Machines: An Application to Face Detection," *Proc. CVPR*, 1997, pp. 130-136.

[3] H. Jee, K. Lee, and S. Pan, "Eye and Face Detection Using SVM," *Proc. ISSNIP*, 2004, pp. 577-580.

[4] L. Jin, X. Yuan, S. Satoh, J. Li, and L. Xia, "A Hybrid Classifier for Precise and Robust Eye Detection," *Proc. ICPR*, 2006, pp. 731-735.

[5] G. Pan, W. Lin, Z. Wu, Y. Yang, "An Eye Detection System Based on SVM Filter," *Proc. SPIE*, vol. 4925, 2002, pp. 326-331.

[6] S. Yang, K. Seo, Y. Ro, S.-K. Kim, J.-Y. Kim, and Y. Seo, "User-Centric Digital Home Photo Album," *Proc. ISCE*, 2005, pp. 226-229.

[7] J.-H. Lee, "Automatic Video Management System Using Face Recognition and MPEG-7 Visual Descriptors," *ETRI Journal*, vol. 27, no. 6, Dec. 2005, pp. 806-809.

[8] S.-K. Hwang and W.-Y. Kim, "A Novel Approach to the Fast Computation of Zernike Moments," *Pattern Recognition*, vol. 39, no. 11, 2006, pp. 2065-2076.

[9] S. Kim, I. Kweon, and I. Kim, "Robust Model-Based 3D Object Recognition by Combining Feature Matching with Tracking," *Proc. ICRA*, vol. 2, 2003, pp. 2123-2128.

[10] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," *Proc. CVPR*, vol. 1, 2001, pp. 511-518.

[11] http://www.intel.com/technology/computing/opencv/.