Eye detection using eye filter and minimisation of NMF-based reconstruction error in facial image

C.W. Park, K.T. Park and Y.S. Moon

A novel approach for eye detection is proposed, where an eye filter considering textural characteristics of eye regions and non-negative matrix factorisation (NMF)-based image reconstruction is utilised. The eye filter for detecting pairs of eye candidate points is utilised. To detect the exact eye position, all pairs of the eye candidates are evaluated by a cost function, which minimises the NMF-based image reconstruction error. Experimental results have shown that the proposed method achieves better performance of eye detection than existing methods under various environments such as head pose, facial expression and illumination conditions.

Introduction: Face recognition is one of the most important and interesting research topics in the area of computer vision. For obtaining effective results of face recognition, it is necessary to normalise a face region in terms of image size, position and orientation in the training phase of face images. Such normalisation can be performed by using some facial features. Among these features, eyes can be considered for the normalisation of the face region, because the eyes are more salient and stable features than others. However, it is very difficult to detect the eye position accurately, because of conditions such as illumination and closed eyes.

Many approaches for eye detection have been proposed in the literature. Kawaguchi and Rizon [1] proposed to detect the iris using the intensity and the edge information. Song *et al.* [2] used the binary edge images to detect eyes. Sirohey and Rosenfeld [3] proposed an eye detection algorithm based on linear and nonlinear filters. Zhou and Geng [4] used the hybrid projection function (HPF) to detect eyes. Wang and Yin [5] proposed an eye detection algorithm by terrain feature matching. Kim and Kim [6] and Wang *et al.* [7] proposed eye detection methods using a support vector machine (SVM).

In this Letter, we propose a novel method to detect eyes based on eye filter and NMF-based image reconstruction. The proposed method is composed of three steps. In the first step, illumination normalisation is carried out to reduce the illumination effects for various lighting conditions. In the second step, pairs of eye candidate points are extracted by the proposed eye filter considering characteristics of eye texture. In the final step, a pair of eye positions is detected from the pairs of eye candidate points by using a cost function, which minimises the NMF-based image reconstruction error.

Illumination normalisation: To normalise illumination effects under various lighting conditions, local statistics information of an image is utilised. The illumination normalisation is performed by dividing the intensity value by the mean intensity value of neighbouring pixels, as in

$$N(x, y) = \frac{I(x, y)}{\mu(x, y)} \tag{1}$$

where $N(x, y)$ is the normalised illumination value, $I(x, y)$ is the intensity value, and $\mu(x, y)$ is the mean intensity value of neighbouring pixels in the input facial image.

Extraction of pairs of eye candidate points by eye filter: In this Letter, pairs of eye candidate points are extracted by the proposed eye filter, which considers two characteristics of eyes. The first is that eye regions have horizontal shape, regardless of eyes being open or closed. The other is that the pupils of eyes have circular shape and they are darker than their surroundings.

For effectively detecting eye regions with horizontal shape, the horizontal even Gabor filter is used:

$$g(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \times \cos(2\pi f x) \tag{2}$$

where σ is the standard deviation of a Gaussian envelop and f is the radial frequency of sinusoid.

For detecting pupils of eyes, we define a circular filter:

$$c(x, y) = \left(\frac{2}{1 + \left(\sqrt{(x^2+y^2)/2\alpha^2}\right)^{2n}} - 1 \right) \times \left(\frac{1}{1 + \left(\sqrt{(x^2+y^2)/2\beta^2}\right)^{2m}} \right), \quad \alpha < \beta, n < m \tag{3}$$

where α is the inner cutoff variable, β is the outer cutoff variable, n is the inner order, and m is the outer order.

The proposed eye filter is generated by combining two filters, $\mathcal{J}[g]$ and $\mathcal{J}[c]$, as shown in Fig. 1. $\mathcal{J}[\cdot]$ denotes the Fourier transform of the argument. The response of the proposed eye filter is calculated for detecting candidate points of eyes in the normalised input image. Then, if the distance between two points of remaining eye candidate points is larger than a threshold value, the two points are considered as a pair of exact eye positions.

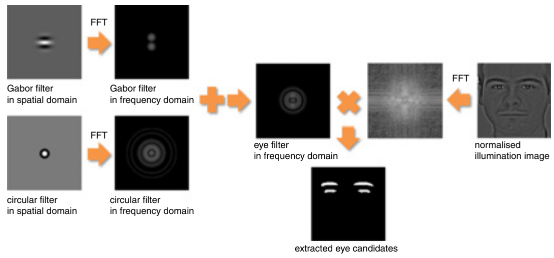

Fig. 1 *Procedure of proposed eye filter*

Detection of eye position by minimising NMF-based image reconstruction error: NMF is distinguished from traditional subspace methods by its use of non-negativity constraints. These constraints lead to a parts-based representation through only additive combinations [8]. The image database is regarded as an $n \times m$ matrix $\boldsymbol{V}$, each column of which contains n non-negative pixel values of one of the m facial images. It is possible to find two new matrices $\boldsymbol{W}$ and $\boldsymbol{H}$ in order to approximate the original matrix $\boldsymbol{V} \simeq \boldsymbol{WH}$. The dimensions of the matrix factors $\boldsymbol{W}$ and $\boldsymbol{H}$ are $n \times r$ and $r \times m$, respectively. Generally, r is chosen so that $(n + m)r < nm$, where r columns of $\boldsymbol{W}$ are called the basis images, and each column of $\boldsymbol{H}$ is called an encoding. To estimate the factorisation matrices, an iterative learning algorithm is used to find the linear decomposition:

$$\boldsymbol{W}_{ia} \leftarrow \boldsymbol{W}_{ia} \sum_{\mu} \frac{\boldsymbol{V}_{i\mu}}{(\boldsymbol{WH})_{i\mu}}, \ \boldsymbol{W}_{ia} \leftarrow \frac{\boldsymbol{W}_{ia}}{\sum_j \boldsymbol{W}_{ja}}, \qquad \boldsymbol{H}_{a\mu} \leftarrow \boldsymbol{H}_{a\mu} \sum_i \boldsymbol{W}_{ia} \frac{\boldsymbol{V}_{i\mu}}{(\boldsymbol{WH})_{i\mu}} \tag{4}$$

Generally, in order to project a new input image onto basis matrix $\boldsymbol{W}$, the iterative algorithm is used without changing the matrix $\boldsymbol{W}$. Because the processing time of the projection method is very expensive, we use a pseudo-inverse matrix $\boldsymbol{W}^{\dagger}$ generated by singular value decomposition (SVD) for projecting a new vector onto NMF:

$$\boldsymbol{h} = \boldsymbol{W}^{\dagger}\boldsymbol{v} \tag{5}$$

Suppose $\boldsymbol{v}$ is an image vector; encoding vector $\boldsymbol{h}$ is obtained by multiplying the pseudo-inverse matrix $\boldsymbol{W}^{\dagger}$ by the column vector $\boldsymbol{v}$. Then, the reconstruction of the obtained encoding vector $\hat{\boldsymbol{v}}$ can be performed by linear combination of $\boldsymbol{W}$ and $\boldsymbol{h}$:

$$\hat{\boldsymbol{v}} = \boldsymbol{Wh} \tag{6}$$

To exactly detect the eye position, we minimise the NMF-based image reconstruction error. In a training face dataset, all face images are normalised by manually selected eye positions. Therefore, if a pair of eye positions is exactly selected from an input facial image, the reconstructed image by NMF is remarkably similar to the input facial image, as shown in Fig. 2. With the characteristics mentioned above, the reconstruction using the pairs with two points among eye candidate points is performed repeatedly. Then the corresponding pair is selected that minimises the sum of pixel difference between the reconstructed image and the aligned input image:

$$d(\boldsymbol{a}, \boldsymbol{b}) = \sum_{k \in N} |\boldsymbol{a}(k) - \boldsymbol{b}(k)|$$
$$eye(i_0, j_0) = \arg \min_{i,j \in C} d(\boldsymbol{h}_{ij}, \hat{\boldsymbol{v}}_{ij}), \; i \neq j, C = \{c_1, c_2, \ldots, c_m\} \tag{7}$$

where $eye(i_0, j_0)$ is the pair selected as the final eye position, C is the eye candidate set, $\boldsymbol{h}_{ij}$ is the aligned input image by the pair of eye candidate points ij, and $\hat{\boldsymbol{v}}_{ij}$ is the reconstructed image of $\boldsymbol{h}_{ij}$. Accordingly, the exact eye position is determined.

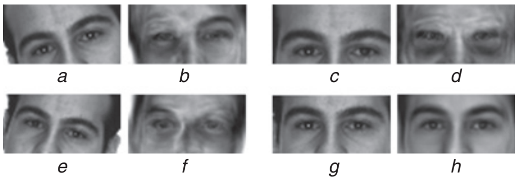

Fig. 2 *Results of NMF-based image reconstruction*

a, *c* and *e* Input images that inexactly select eye positions
b, *d* and *f* Reconstructed images of *a*, *c* and *e*
g Input image that exactly selects eye positions
h Reconstructed image of *g*

Experimental results: To evaluate the performance of the proposed method, AR, Bern, Aberdeen, JAFFE and ORL face databases were used [1–7]. Table 1 shows the performance comparison of various eye detection methods. As shown in Table 1, the recognition rate of the proposed eye detection method has been improved by up to 10% over existing methods.

Table 1: Performance comparisons of various eye detection methods

	AR-63	AR-564	Bern-150	Aberdeen-72	JAFFE-213	ORL-400
Proposed method	96.8%	98.8%	99.3%	100%	100%	98%
Kawaguchi and Rizan [1]	96.8%	-	95.3%	-	-	-
Song *et al.* [2]	96.8%	96.6%	98.7%	-	-	-
Sirohey and Rosenfeld [3]	-	-	-	90%	-	-
Zhou and Geng [4]	-	-	-	-	97.2%	-
Wang and Yin [5]	-	-	-	-	95.8%	-
Kim and Kim [6]	-	-	-	-	-	91.9%
Wang *et al.* [7]	-	-	-	-	99.1%	90.7%

Conclusions: In this Letter, we propose a novel method for detecting the exact eye position. To detect the eye position, the extraction of eye candidates by an eye filter is proposed and the minimisation of NMF-based image reconstruction error is utilised for detecting exact eye positions. From experimental results, it has been shown that the proposed method achieves better performance of eye detection than existing methods.

Acknowledgment: This work was supported by a National Research Foundation of Korea Grant funded by the Korean Government (No. 2009-0077434).

17 November 2009
doi: 10.1049/el.2010.3239
One or more of the Figures in this Letter are available in colour online.

C.W. Park and Y.S. Moon (*Computer Vision and Pattern Recognition Laboratory, Department of Computer Science and Engineering, Hanyang University, Korea*)

E-mail: ysmoon@cse.hanyang.ac.kr

K.T. Park (*Ambient Intelligence Software Team, Institute of Hanyang BK21, Hanyang University, Korea*)

References

1 Kawaguchi, T., and Rizon, M.: ‘Iris detection using intensity and edge information’, *Pattern Recognit.*, 2003, **36**, pp. 549–562
2 Song, J., Chi, Z., and Liu, J.: ‘A robust eye detection method using combined binary edge and intensity information’, *Pattern Recognit.*, 2006, **39**, pp. 1110–1125
3 Sirohey, S.A., and Rosenfeld, A.: ‘Eye detection in a face image using linear and nonlinear filters’, *Pattern Recognit.*, 2001, **34**, pp. 1367–1391
4 Zhou, Z.H., and Geng, X.: ‘Projection functions for eye detection’, *Pattern Recognit.*, 2004, **37**, pp. 1049–1056
5 Wang, J., and Yin, L.: ‘Detecting and tracking eyes through dynamic terrain feature matching’. IEEE Conf. on Computer Vision and Pattern Recognition, San Diego, CA, USA, 2005, pp. 78–85
6 Kim, H.J., and Kim, W.Y.: ‘Eye detection in facial images using Zernike moments with SVM’, *ETRI J.*, 2008, **30**, pp. 335–337
7 Wang, Q., Zhao, C., and Yang, J.: ‘Efficient facial feature detection using entropy and SVM’, *Lect. Notes Comput. Sci.*, 2008, **5358**, pp. 763–771
8 Lee, D.D., and Seung, H.S.: ‘Learning the parts of objects by non-negative matrix factorization’, *Nature*, 1999, **401**, pp. 788–791