Available online at www.sciencedirect.com
SCIENCE DIRECT
Pattern Recognition 39 (2006) 1110–1125
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

A robust eye detection method using combined binary edge and intensity information

Jiatao Song[a, b, c], Zheru Chi[a, *], Jilin Liu[b]

[a]*Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong*
[b]*Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, PR China*
[c]*College of Electronic and Information Engineering, Ningbo University of Technology, Ningbo 315016, PR China*

Received 26 January 2005; received in revised form 10 November 2005; accepted 10 November 2005

Abstract

In this paper, a new eye detection method is presented. The method consists of three steps: (1) extraction of binary edge images (BEIs) from the grayscale face image based on multi-resolution wavelet transform, (2) extraction of eye regions and segments from BEIs and (3) eye localization based on light dots and intensity information. In the paper, an improved face region extraction algorithm and a light dots detection algorithm are proposed for better eye detection performance. Also a multi-level eye detection scheme is adopted. Experimental results show that a correct eye detection rate of 98.7% can be achieved on 150 Bern images with variations in views and gaze directions and 96.6% can be achieved on 564 AR images with different facial expressions and lighting conditions.

Keywords: Eye detection; Binary edge images; Multi-resolution face image analysis; Multi-level eye detection; Light dots

1. Introduction

In the past two decades, automatic human face image analysis and recognition has become one of the most important research topics in computer vision and pattern recognition. Because of tremendous potential applications, topics such as face detection [1,2], face identification and recognition [3–5] and facial expression analysis [6,7] have attracted more and more attention. Among these research topics, one fundamental but very important problem to be solved is automatic eye detection. As eyes are the most salient and stable feature in a human face, extraction of eyes are often regarded as the most important step in many face detection algorithms [2,8,9]. Only those image regions that contain possible eye pairs will be fed into a subsequent face verification system. Localization of eyes is also a necessary step for many face recognition methods. Before two face images can be compared, they should be aligned in orientation and normalized in scale. Since both the locations of two eyes and the interocular distance are relatively constant for most people [10], the eyes are often used for face image normalization. Eye localization also further facilitates the detection of other facial landmarks. Phillips et al. [11] used whether the centers of eyes can be automatically located or not as a criterion to distinguish a fully automatic recognition system from a partially automatic system. In addition, because the eyes often reflect a person's desires, needs and emotions, eye detection, together with eyes tracking is crucial for facial expression analysis [6,7], human computer interaction, and attentive user interfaces.

According to Zhu et al. [12], existing eye detection methods can be broadly classified into two categories: the active infrared (IR)-based approaches [12–14] and the traditional image-based passive approaches. The former is based on the principle of red-eye effect in flash photographs, utilizing a special IR illuminator and an IR-sensitive CCD for image acquisition. This approach is relatively simple and very effective. It can obtain high eye detection and tracking robustness and accuracy, especially indoors. But it requires a special lighting and synchronization scheme. The success of such a system mainly depends on the brightness and size

* Corresponding author. Tel.: +852 2766 6219; fax: +852 2362 8439.
E-mail address: enzheru@inet.polyu.edu.hk (Z. Chi).

0031-3203/$30.00
doi:10.1016/j.patcog.2005.11.015

of pupils, which are often a function of the face orientation, external illumination interference and the distance of the subjects to the camera. In other words, the most significant problem with this approach is that it requires a relatively stable lighting condition and a camera set close to the subject [12].

The image-based passive eye detection approaches often employ the special properties of eyes in intensity or color distribution, shape or appearance. The commonly used approaches in this category include the template matching method [4,15], eigenspace [16] method and Hough transform-based method. In the template matching method, different segments of an input image are compared with those in the template, usually using correlation values to evaluate the similarity of the counterpart. The simple template matching method is not very robust since it cannot deal with eye variations in scale, expression, rotation and illumination. In Brunelli and Poggio's approach [4], multi-scale templates were used, which partially solved the scale problem. Yuille et al. [17] proposed a deformable template for face feature extraction. An eye was described by a parameterized template and an energy function was defined to link edges, peaks, and valleys in the input image to corresponding parameters in the template. The advantage of this approach is that besides the location, more features of an eye, such as its shape and size, can be obtained at the same time. But the approach is time consuming, and its rate of success relies greatly on the initial position of the template. Lam and Yan [18] extended Yuille's work by introducing the concept of eye corners, which proved to be effective in reducing the processing time.

Pentland et al. [16] used an eigenspace method for eye detection. This method can achieve better eye detection performance than a simple template method since training samples have covered different eye variations in appearance, orientation and lighting conditions. One problem with this method is that its performance is closely related to the selection of the training set. Another drawback is that it requires the training and test images to be normalized in size and orientation.

Hough transform is another widely used eye detection method. It is based on the shape feature of an iris and often works on binary valley or edge maps [19,20]. The shortcoming of this approach is that its performance depends on the threshold values selected for the binarization of valley or edge maps.

Besides these three classical approaches, recently many other image-based eye detection techniques have been reported. Feng and Yuen [21] used intensity, the direction of the line joining the centers of the eyes, the response of convolving an eye variance filter with the face image and the variance projection function (VPF) [22] technique to detect eyes. Zhou and Geng [23] extended the idea of VPF to the generalized projection function (GPF) and showed with experimental results that the hybrid projection function (HPF), a special case of GPF, is better than VPF and integral projection function (IPF) for eye detection. Kawaguchi and Rizon [19] located the iris using intensity and edge information. The main techniques included in their algorithm are a feature template, a separability filter, Hough transform and template matching. Sirohey and Rosenfeld [24] proposed an eye detection algorithm based on linear and nonlinear filters. In Huang and Wechsler's method [10], the task of eye location was considered as a test bed for developing navigation routines implemented as visual routines. They used genetic algorithms and built decision trees to detect eyes. For the purpose of face detection, Wu and Zhou [9] employed size and intensity information to find eye-analog segments from a gray scale image, and exploited the special geometrical relationship to filter out the possible eye-analog pairs. Similarly, Han et al. [8] applied such techniques as morphological closing, conditional dilation and labeling process to detect eye-analog segments. Hsu et al. [2] used color information for eye detection.

Although much effort has been spent and some progress has been made, the problem of automatic eye detection is still far from being fully solved owing to its complexity. Many factors, such as facial expression, face rotation in plane and depth, occlusion, lighting conditions, and so on, undoubtedly affect the performance of eye detection algorithms. Unfortunately, most of the existing approaches mainly focus on eye detection from the face images of the frontal view without specifically taking into consideration the above-mentioned factors. Kawaguchi and Rizon [19] tried to develop an eye detection method with good robustness to gaze direction and reflected light dots. They used images from two databases, i.e., the Bern database [25] and the AR database [26], to evaluate the validity of their algorithm, achieving a correct iris detection rate of 95.3% for 150 Bern face images and 96.8% for 63 AR images. But they did not explain how to automatically detect the light dot in the iris.

In this paper, a novel eye detection method is presented. The aim of our work is to address the problem of eye detection from images with variations in pose, gaze direction and lighting condition. Reflected light spots often exist in the face images captured using a camera with flash photography. They are often located near the center of the irises and can be approximately considered as the center of an eye. In this paper, the light dot is viewed as an important cue for eye localization. An automatic light dot detection algorithm is also proposed. The input images in our study are supposed to be head-and-shoulder images with a plain background. The proposed method consists of three main steps, i.e., face region extraction, eye region segmentation and fine eye localization. In order to facilitate the extraction of the face region and eye segments, a robust algorithm for the extraction of the binary edge image (BEI) is developed. The flowchart of our proposed eye detection method is shown in Fig. 1.

The remaining part of this paper is organized as follows. In Section 2, a method for the extraction of BEI based on the multi-resolution property of wavelet transform (WT) is presented. An improved eye detection method is discussed in

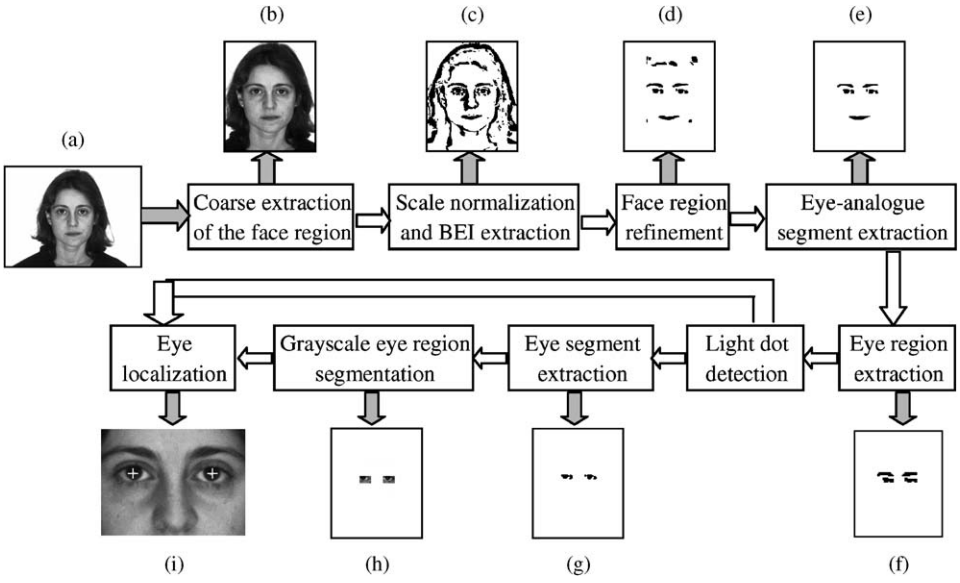


Fig. 1. Flowchart of our proposed eye detection method.

Section 3. In Section 4, experimental results on the Bern [25] and AR [26] face databases are reported. Finally, concluding remarks are drawn in Section 5.

2. Extraction of BEIs

2.1. BEI extraction method

Many operators, such as Prewitt, Sobel, and Canny operators, can be used for the extraction of image edges. But binary edges extracted using these operators normally consist of some contours (often broken) of the objects in the scene. Thus, face edge maps obtained using these operators are not suitable for the segmentation of face components. In our study, an improved wavelet-based edge extraction algorithm is proposed.

WT can be considered as a microscope for signal analysis. Using the multi-resolution property of WT, a signal can be decomposed into a low-frequency component and a number of high-frequency components, i.e.,

$$S = \{A_K, D_K, D_{K-1}, D_{K-2}, \ldots, D_2, D_1\},$$

where S is the original signal, A_K is the low-frequency component at scale K, $D_i (i = 1, 2, \ldots, K)$ are high-frequency components. In our method, the high-frequency components are first employed to reconstruct the face image. The first two rows of Fig. 2 show examples of the high-frequency reconstructed images at scale 0 (original image), 1, 3 and 5 and their corresponding grayscale histograms. The image reconstructed at scale K means that only high-frequency components $D_K, D_{K-1}, \ldots, D_1$ are utilized for image reconstruction, while the low-frequency component A_K is discarded. By observing the high-frequency reconstructed image and its histogram, it was found that the histogram is almost symmetric and the pixels on edges have relatively lower grayscales. Let the high-frequency image reconstructed at scale i be denoted HI_i; a binarization is done by

$$SBEI_i(x, y) = \begin{cases} 1 & \text{if } HI_i(x, y) < \overline{HI_i} + \delta, \\ 0 & \text{others,} \end{cases} \quad (1)$$

where $\overline{HI_i}$ is the mean gray level of HI_i and δ is a small number, such as 5, $1 \leqslant x \leqslant M$, $1 \leqslant y \leqslant N$, where M and N are the height and width of the image, respectively. The binary image produced, denoted by $SBEI_i$, mainly contains edges.

The third row of Fig. 2 shows the binary images obtained at scales 0, 1, 3 and 5. The case of scale 0 corresponds to the binarized image of the original face image using a global threshold. It can be observed from Fig. 2 that when the scale is small, such as 1, edges of face components are thin and often disconnected. With an increased scale, more pixels are segmented as foreground, so the edges of face components become clearer and more connective. But if the scale is too large, some details of the face components may be included, making the useful edges blurred. In order to have a high-quality BEI, an algorithm for the re-extraction of edges was developed. The algorithm includes the following two steps:

(1) Merging all BEIs obtained at every single scale by using simple addition at every pixel, i.e.,

$$X(x, y) = \sum_{i=1}^{LS} SBEI_i(x, y), \quad (2)$$

where LS is the largest scale used and is set to 6 in our study. An example is shown in Fig. 3(a). Obviously, the pixel values of image X are between 0 and LS.

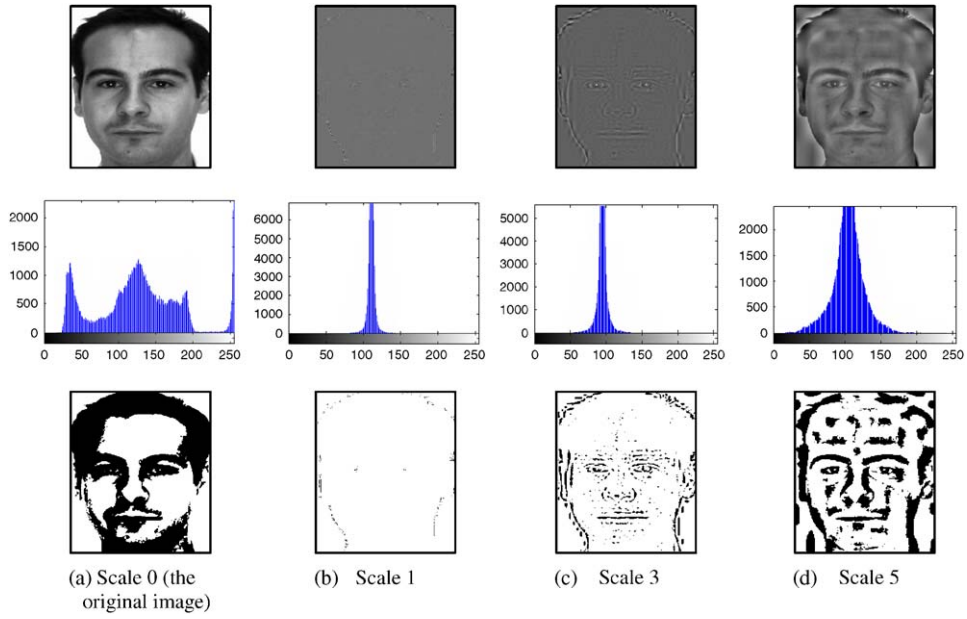


Fig. 2. Images reconstructed with high-frequency components at different scales, their histograms and binary edge images (from upper to lower).

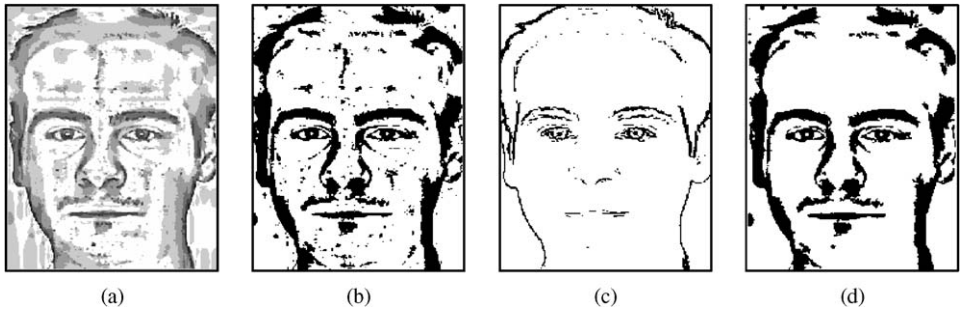


Fig. 3. Extraction of face edges. (a) The merged edge image, (b) BEI after binarization, (c) BEM, (d) BEI after noise removing.

(2) Thresholding X at a threshold T to obtain a BEI, i.e.,

$$BEI(x, y) = \begin{cases} 1 & \text{if } X(x, y) > T, \\ 0 & \text{others,} \end{cases} \tag{3}$$

where the threshold T is determined by

$$T = \arg \min_{t=1,2,\ldots,LS-1} \left\{ \left| \frac{N_{fp}(t)}{M \times N} - 0.2 \right| \right\}, \tag{4}$$

where N_{fp} denotes the number of foreground pixels in the binary image that resulted from thresholding the grayscale image using threshold t.

The achieved BEI after two binarization processes is shown in Fig. 3(b) for the original image shown in Fig. 2(a). It can be observed that the face components including eyebrows, eyes, nose and mouth have been clearly extracted. But unfortunately, there is much noise in it. In order to improve the quality of the BEI, the binary edge map (BEM) of the original grayscale image is used for noise removing. The BEM, as shown in Fig. 3(c), was obtained using the Sobel-based locally adaptive threshold (LAT) algorithm [27]. Compared with BEI, BEM has an advantage of less noise, especially in the region around the face components. Using this noise-reduced BEM as a template, an algorithm for noise removing from BEI is proposed. It includes the following steps:

(1) Multiplying BEI by BEM at each pixel, i.e.,

$$P(x, y) = BEI(x, y) \times BEM(x, y), \tag{5}$$

(2) Finding out all eight-connected components in BEI and labeling the clusters formed with integers $1, 2, \ldots, N_{CC}$, where N_{CC} is the total number of clusters. An L matrix is defined by

$$L(x, y) = \begin{cases} k & \text{if } (x, y) \in \text{cluster } k \\ 0 & \text{if } (x, y) \in \text{background} \end{cases}$$

$$(1 \leqslant k \leqslant N_{CC}). \tag{6}$$

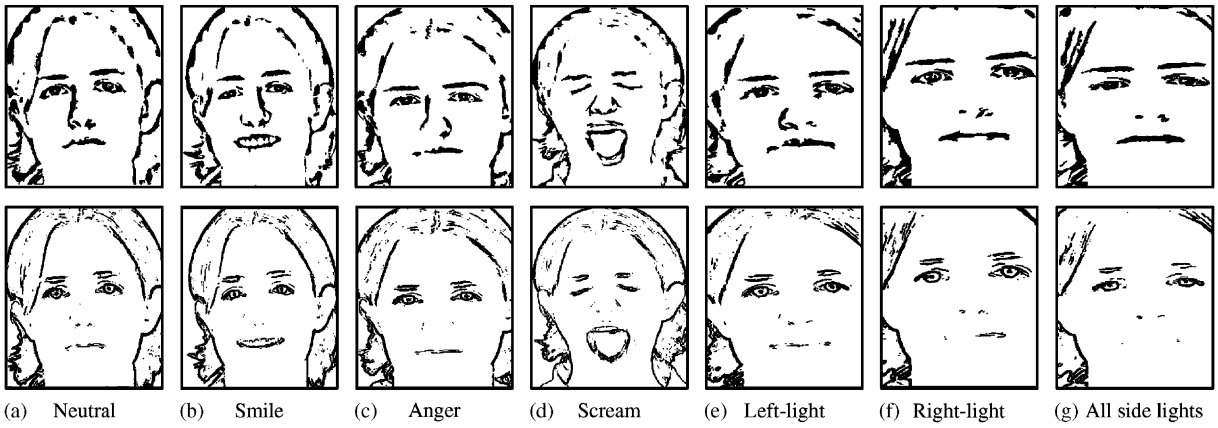


Fig. 4. More examples of BEIs from the AR face database, together with their corresponding BEMs (the first row: BEIs, the second row: BEMs).

(3) Multiplying L by P at each pixel and storing the result into matrix NI, i.e.,

$$NI(x, y) = L(x, y) \times P(x, y). \tag{7}$$

(4) Processing the connected components in BEI individually in the following way: if a nonzero element in L does not exist in NI, it means that the corresponding foreground pixels in BEI are noise and must be removed, i.e.,

$$BEI(x, y) = \begin{cases} 0 & \text{if } L(x, y) \neq 0 \text{ and} \\ & \quad NI(x, y) = 0, \\ BEI(x, y) & \text{others.} \end{cases} \tag{8}$$

The BEI after noise removing is shown in Fig. 3(d). Compared with Fig. 3(b), a large portion of noise in the image has been removed.

2.2. Characteristics of BEI

More examples of BEIs from AR face images with different expressions and lighting conditions are shown in Fig. 4. As a comparison, the corresponding BEMs are also shown. It is observed that BEIs have the following distinct characteristics:

(1) In a BEM, only the contours of face components are extracted and they are usually broken. While in noise-removed BEI, the shape and contour of face components are clearly extracted. Merged face components are rare. Face components, especially eyes and eyebrows, even irises, are clearly extracted. This is very helpful for the segmentation of face components.
(2) The details of face components are clearly seen. For example, the iris can be easily distinguished from the whites of eyes. Even the light dots in the irises can be clearly seen. This may help the detection of face feature points, such as the corners of eyes and mouth, light dots and so on.

Fig. 5. BEIs generated from images with different resolutions (a) 320 × 240; (b) 240 × 180.

(3) Good robustness to illumination changes. For example, when lighting is strong, some face components, such as the mouth and noise, disappear from BEMs, but still exist in BEIs. Experimental results show that the proposed edge detection method can reliably extract face components in most illumination conditions.

One issue that should be pointed out here is that in the proposed eye detection method, before the eyes can be located, eye segments should be extracted from the BEI. In order to facilitate the segmentation of face components, both the connectivity of pixels belonging to one face component and the separability of segments belonging to different face components in BEI should be high. But sometimes the resulting BEI is not very suitable for the segmentation of face components because of the poor quality of the grayscale image. Fortunately, it is observed that image resolution has slight influence on the connectivity of edges. Fig. 5 shows two BEIs, which are obtained by first downsampling the input grayscale image to two different resolutions. It can be seen that in the case of 320 × 240, the separability of different face component segments such as eyebrow segment,

eye segment and nose segment is much better than that in the case of 240×180. Motivated by this, in our proposed eye detection method, two BEIs with different resolutions are used for eye region extraction, light dot detection and eye segments extraction, and a multi-level eye detection scheme is adopted, which will be discussed in detail in Section 3.6.

3. The proposed eye detection method

3.1. Face region extraction

Extraction of the face region from the original input image is the first step in our proposed eye detection method. We first utilize a method proposed by Kawaguchi and Rizon [19] to coarsely detect the face region. The extracted image is shown in Fig. 1(b) for the input image of Fig. 1(a). One drawback of this face detection method is that when there is too much hair on the two sides of the face, the lower boundaries are distorted significantly from the actual face boundaries. Thus, hair, neck and even shoulder parts may be included in the extracted face region, which will undoubtedly increase the difficulty in the subsequent eye detection. So in our system, the face region is further refined.

The face boundaries are refined using the information contained in the BEI. We first define two terms to describe the property of a foreground pixel in the BEI, i.e., *horizontal connection length* (*HCL*) *and vertical connection length* (*VCL*).

The *HCL* of a foreground pixel P in the BEI, denoted by HCL_P, is defined as the length of the horizontal line (the number of pixels) that passed the pixel and crosses the foreground region.

The *VCL* of a foreground pixel P, denoted by VCL_P, is defined similarly except that it is vertically connected.

For example, in Fig. 6, $HCL_P = |\mathrm{AB}|$ and $VCL_P = |\mathrm{CD}|$, where the lengths of |AB| and |CD| are measured in pixels.

Furthermore, we define

$$\lambda_P = \frac{VCL_P}{HCL_P}. \tag{9}$$

It is obvious that in a BEI, the λ value for most pixels in eyebrow segments, eye segments or mouth segment is less than 1, while most pixels on the left and right boundaries of the face have a λ value of greater than 1. Based on this, the left and right face boundaries can be easily removed from a BEI by iteratively removing those pixels which satisfy any of the following conditions until all these pixels have been removed.

$$\lambda(k) = \frac{VCL(k)}{HCL(k)} \geqslant 3 \quad (k = 0, 1, 2, \ldots), \tag{10}$$

$$\frac{HCL(k)}{HCL(0)} < 0.5 \quad (k = 0, 1, 2, \ldots), \tag{11}$$

$$\frac{VCL(k)}{VCL(0)} < 0.5 \quad (k = 0, 1, 2, \ldots), \tag{12}$$

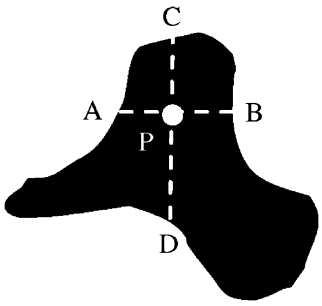


Fig. 6. Illustration of HCL and VCL.

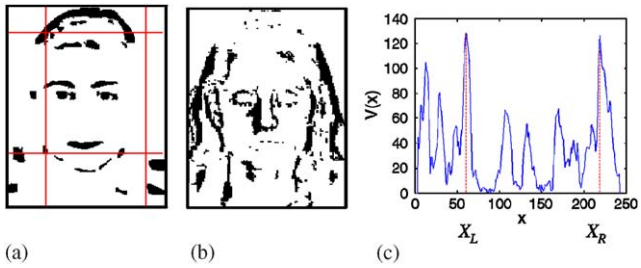


Fig. 7. Refinement of the face region. (a) BEI after the boundary pixel removing operation, (b) removed parts, (c) vertical projection of (b).

$$\sum_{x=1}^{N_{BEI}} BEI_k(x, y) < 10, \tag{13}$$

where $BEI_k (k = 0, 1, 2, \ldots)$ denotes the BEI after the kth removing operation, $HCL(k)$ and $VCL(k)$ $(k = 0, 1, 2, \ldots)$ represent the HCL and VCL of a pixel after the kth removing operation N_{BEI} is the width of the BEI image

Fig. 7(a) shows that the BEI resulted from the above pixel removing operation. Compared with Fig. 1(c), it can be seen that most pixels on the vertical face boundaries have been removed, while those horizontal face components, especially eye and eyebrow segments, show little change. At the same time, another interesting phenomenon appears in the BEI, i.e., most foreground pixels between the eyebrows and eyes, and between the eyes and nose, have almost been removed. This greatly decreases the connectivity between different components and it is very helpful in the segmentation of these regions.

Fig. 7(b) shows the difference BEI before and after the boundary pixel removing operation. It can be seen that the information about the left and right face boundary have been well reflected in it. So, by projecting this difference BEI in the vertical direction, the updated face boundaries can be obtained.

Let $V(x)$ be the vertical projection of the difference BEI, the locations on the X-axis of potentially refined left and right face boundary, denoted by x_1 and x_2, can be obtained with

$$x_1 = \arg \max_{1 \leqslant x \leqslant (N_{BEI}/2)} \{V(x)\} \tag{14}$$

$$x_2 = \arg \max_{(N_{BEI}/2) < x \leqslant N_{BEI}} \{V(x)\}. \quad (15)$$

The left face boundary x_L and right face boundary x_R are updated according to the following rules:

$$x_L = \begin{cases} x_1 & \text{if } V(x_1) \geqslant \dfrac{M_{BEI}}{3}, \\ 1 & \text{others}, \end{cases} \quad (16)$$

$$x_R = \begin{cases} x_2 & \text{if } V(x_2) \geqslant \dfrac{M_{BEI}}{3}, \\ N_{BEI} & \text{others}, \end{cases} \quad (17)$$

where M_{BEI} denotes the height of the BEI.

The upper boundary y_U is given by

$$y_U = \begin{cases} \dfrac{N_{BEI} - (x_R - x_L)}{2} & \text{if } \dfrac{N_{BEI} - (x_R - x_L)}{2} \leqslant \dfrac{M_{BEI}}{5}, \\ \dfrac{M_{BEI}}{5} & \text{others}. \end{cases} \quad (18)$$

The new lower boundary y_L is obtained with

$$y_L = y_U + 1.2 \times (x_R - x_L). \quad (19)$$

By removing those foreground pixels outside the new face boundaries from the BEI, a refined BEI (RBEI), can be obtained (see Fig. 1(d)).

3.2. Extraction of eye-analog segments

Eye segments are extracted from the RBEI. Before this task is carried out, it is helpful to remove those blocks that are unlikely to be eye segments from the RBEI. The blocks to be removed should satisfy at least one of the following criteria:

- area is too small, e.g., less than 20 pixels in the 320×240 resolution case;
- part or all pixels are at the upper $\frac{1}{5}$ or the lower $\frac{1}{9}$ regions of the BEI;
- the width of the minimum bounding rectangle, denoted by w_{BR}, is equal to or greater than half of the width of the RBEI N_{BEI}, i.e.,

$$w_{BR} \geqslant \frac{N_{BEI}}{2}; \quad (20)$$

- the maximum *HCL* is less than or equal to 5, i.e.,

$$\max_{P \in B_i} (HCL_P) \leqslant 5, \quad (21)$$

where B_i denotes block i, $i = 1, 2, \ldots, N_{SRBEI}$; N_{SRBEI} is the number of segments in the RBEI. The BEI after the above block removing operation is called possible eye segments BEI (PESBEI).

In the BEI, eye segments are horizontal blocks, and they are often relative large in size, so the following constraints are used to guide the detection of eye-analog segments from the PESBEI:

- Area is greater than a threshold.
- The ratio of the maximum *HCL* to the maximum *VCL* is greater than or equal to 2, i.e.,

$$\frac{\max_{P \in B_i} (HCL_P)}{\max_{Q \in B_i} (VCL_Q)} \geqslant 2 \quad (i = 1, 2, \ldots, N_{SPESBEI}), \quad (22)$$

  where $N_{SPESBEI}$ is the number of segments in the PESBEI.
- The ratio of the width w_{BR} to the height h_{BR} of its minimum bounding rectangle is greater than 1.5, i.e.,

$$\frac{w_{BR}}{h_{BR}} > 1.5. \quad (23)$$

The resulting BEI, named eye-analog segments BEI (EASBEI), is shown in Fig. 1(e). It mainly contains eye, eyebrow or mouth segments. Other components, such as most of the remaining face boundary blocks, hair blocks and even nose blocks have been removed.

Fig. 8 shows a more detailed example of eye segment extraction. Figs. 8(a) and (b) show the RBEI and PESBEI, respectively. The obtained EASBEI is shown in Fig. 8(c). Fig. 8(d) shows the BEI with eye and eyebrow segments only remained.

3.3. Detection of the eye region

An eye region refers to the area that contains eyes and eyebrows only. It is detected by making use of a prior knowledge about the layout of face components. The knowledge used includes (1) the vertical distance between eyes and nose is normally greater than that between eyes and eyebrows and that between nose and mouth; (2) there are two eyes and two eyebrows, but only one mouth and one nose on the face. Based on these knowledge, the eye region can be extracted using the horizontal and vertical integral projection method.

The BEI with eye and eyebrow segments only left is called the eye region BEI (ERBEI). Examples of ERBEI are shown in Figs. 1(f) and 8(d).

3.4. Detection of eye segments

After the ERBEI is obtained, the next task is to detect reflected light dots from the left and right eye region, or to extract eye segments from them if no light dot is detected. The method for light dot detection will be discussed in Section 3.7. Here, we first describe how to extract eye segments from the ERBEI.

Fig. 9 shows some typical eye regions. In order to detect two eye segments, the ERBEI is first divided into two regions, i.e., the left eye region and the right eye region, using the vertical projection method. For each region, we then

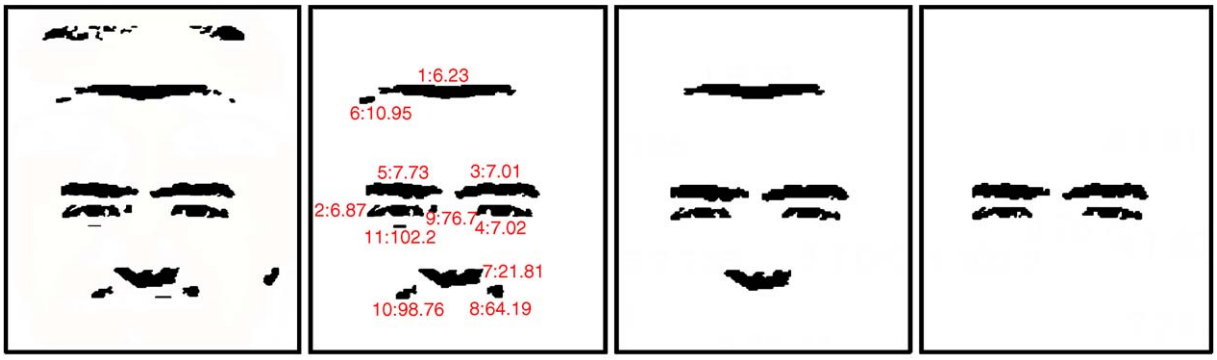


Fig. 8. Detection of eye region. (a) RBEI, (b) PESBEI, (c) EASBEI, and (d) ERBEI.

Fig. 9. Four examples of ERBEI.

project it horizontally. Often, the number of horizontal projection interval (HPI) with non-zero projection value may be 1 (e.g., the left eye region of Fig.9(c)), 2 (e.g., the two parts of Fig. 9(a)), or 3 or more (e.g., the two parts of Fig. 9(b)).

For the cases of two, three and more HPIs, normally the segment corresponding to the lowest HPI is the eye segment. But when the lower eyelid is segmented as a single block, such as the case in Fig. 9(b), this may lead to false detection of eye segments. In order to solve this problem, we use both intensity and location cues for the eye segment detection. Excluding face images of black skin, we verified by experiments that the intensity of the lower eyelid segment is often five times higher than that of the iris segment, i.e., the actual eye segment. The iris segment often has a very low intensity and its intensity is very close to the lowest intensity of the whole face image. So, if we first remove those segments with a relatively high intensity from the left and right eye region BEIs and then select the segment at the lowest position as the eye segments, the false eye segment detection rate will be reduced significantly.

Since the PESBEI is more likely to contain iris segments than the EASBEI and the ERBEI, it is used for segment intensity calculation. The intensity of block i, denoted by $BI(i)(i = 1, 2, \ldots, N_{SPESBEI})$, is defined as

$$BI(i) = \min_{(x,y)\in B_i} \{f(x, y)\} \quad (i=1, 2, \ldots, N_{SPESBEI}), \tag{24}$$

where B_i stands for block i, $f(x, y)$ is the pixel intensity at location (x, y) of the grayscale image, and $N_{SPESBEI}$ is the number of segments in the PESBEI.

For each segment in the PESBEI, we calculate its intensity using Eq. (24). The lowest block intensity in the PESBEI, denoted by *LBI*, is used to approximate the lowest intensity of the whole image. Those segments in the ERBEI that satisfy the following condition are removed:

$$BI(i) \geqslant 5 \times LBI \quad (i = 1, 2, \ldots, N_{SERBEI}), \tag{25}$$

where N_{SERBEI} is the number of segments in the ERBEI.

Fig. 8(b) shows a PESBEI with the intensities and intensity ranks marked near each segment. It can be observed that the *LBI* is 6.23. The left and right eye segments have an intensity of 7.02 and 6.84, which rank No. 4 and No. 2, respectively. These two intensities are very close to the lowest intensity. While the segment below the right eye block has a very high intensity, 102.2. Obviously, it meets the condition set in Eq. (25).

In the case of one HPI, we first horizontally project the left or right eye region BEI, then find the dividing line between eye and eyebrow segments by searching for the valley point from the horizontal projection curve. If the dividing line is detected, the block below the dividing line is considered as the eye segment; otherwise eye segment detection fails.

3.5. Locating eyes using intensity information

The last step of the proposed eye detection method is to locate the eyes using the intensity information. The four boundaries of the detected eye segment are used to guide the extraction of an eye from the original image. The resulting grayscale eye image is shown in Fig. 1(h). Because the intensity of an iris is often lower than that of its surroundings, so from each grayscale eye region, we first find the lowest pixel intensity, denoted by *LPI*, then scan all pixels included in the set *PS* defined by

$$PS = \{(x, y) | f(x, y) \leqslant (\alpha \times LPI), (x, y) \in \text{the left or right eye region}\}, \tag{26}$$

where α is a constant (> 1) determined by experiments. It is used for the improvement of the eye location accuracy. The center of mass of *PS* is used as the final eye position.

3.6. Multi-level eye detection scheme

As shown in Fig. 1, successful eye detection from a face image requires the eye region to be first extracted from the BEI. From the eye region, the light dot may be detected if it exists. If there is no light dot in the left or right eye region, the eye segment should be further segmented. But as pointed out in Section 2.2, to extract the eye region, light dot and eye segment from just one BEI may fail sometimes. By downsampling the grayscale face image, a new BEI with better edge feature may be obtained. By using two BEIs of different resolutions, we can increase the eye detection rate.

The multi-level eye detection scheme is realized as follows. For each extracted face image, two BEIs, one with a resolution of 320×240, the other with 240×180, are produced. And two eye-region BEIs are extracted from them. The eye region from the high-resolution BEI is labeled as HERBEI, the other from the low resolution as LERBEI. Both HERBEI and LERBEI are used for light dot detection. If there is no light dot detected from the left or right eye region, the HERBEI is first utilized for eye segment extraction. If it fails to extract the eye segment, the LERBEI is further utilized for the extraction of the eye segment.

3.7. Detecting reflected light dots from the eye region

Using light spots for eye localization is a main feature of our proposed eye detection method. Light dots are automatically detected from images. Figs. 3 and 5 show that light dots are some small holes, i.e., connected background pixels enclosed by some foreground pixels. These holes can be easily located and labeled.

Unfortunately, besides these light dot-related holes, there are many other possible holes in a BEI. Note that one distinct feature of a light dot is that in the grayscale image the light dot and its surroundings show a striking contrast. For each hole in the BEI, the contrast, denoted by C, can be calculated by

$$C(i) = \frac{\max_{(x,y)\in H_i}\{f(x,y)\}}{AI(i)} \quad (i = 1, 2, \ldots, N_H), \tag{27}$$

where H_i represents hole i, N_H is the number of holes in a BEI, and $AI(i)$ represents the average pixel intensity of a small region around hole i. Because the intensity of the region around a light dot is quite low, the contrast of the light dot-related hole is often higher than that of non-light dot-related holes. So, in our method, those holes with a contrast satisfying the following constraint:

$$C \geqslant \beta \tag{28}$$

are considered as the potential light dots, where β is a threshold determined by experiments.

Light dots belonging to the left and right eye regions are detected separately. The hole in the left or right eye region with the largest contrast of all the potential light dots in the same eye region is labeled as a light dot.

4. Experimental results

4.1. Face databases and methodology

Two face databases used by Kawaguchi and Rizon [19] were employed for our eye detection experiments. They are the Bern database [25] and the AR face database [26]. The images selected from the Bern database include all the 150 face images without spectacles (10 views for each of the 15 people). This database was used to evaluate the influence of gaze directions on our proposed method.

The AR database is stored in eight CD-ROMs. Images feature frontal view faces with different facial expressions, illumination conditions and occlusions. Two image subsets of this database, named AR-63 and AR-564, respectively, were used for our experiments. AR-63 contains 63 images of 21 people (12 men and 9 women) without spectacles, stored in the first CD ROM. They show three expressions (neutral, smile and anger) and were all captured under natural illumination conditions. AR-564 is from the first four CD ROMs, including 564 images of 94 persons (44 men and 50 women) without spectacles. All the images in AR-564 are divided into six groups. The first three groups represent three expressions, i.e., neutral, smile and anger under natural illumination environment, while the other three groups denote three lighting conditions, i.e., left light on, right light on and all side lights on, with neutral expression. The AR database is aimed to evaluate the robustness of our eye detection method to variations in eye appearance and lighting conditions, and to demonstrate the role of reflected light dots in automatic eye localization.

Two indexes are often used for the performance evaluation of an eye detection method. One is the detection rate; the other is the location accuracy. The former refers to the ratio of the number of images for which two eyes are correctly detected to the total number of images tested, while the latter means the disparity between the manually detected eye position and the automatically detected position. Often, the larger of the two eye disparities in a face image is adopted for the accuracy measure of eye detection. One thing relevant to the location accuracy is its representation of the disparity. Some proposed the absolute disparity value in pixels [10,16]. Apparently, this metric relies on the image resolution. Jesorsky et al. [28] proposed *the relative error* to judge the quality of eye detection, which is defined by

$$err = \frac{\max(d_l, d_r)}{d_{lr}}, \tag{29}$$

where d_l is the left eye disparity in pixel, d_r is the right eye disparity, and d_{lr} the Euclidean distance between the manually detected left and right eye locations. Obviously, this metric is resolution independent.

It is clear that based on different precision criteria, different detection rates may be obtained. In Jesorsky et al. [28], Wu and Zhou [9] and Zhou and Geng's work [23], if $err < 0.25$, the detection is considered to be success.

Because d_{lr} roughly equals double of an eye width, the criterion $err < 0.25$ means that the larger of d_l and d_r should be less than half an eye width. Obviously, this criterion is very loose and may not be very suitable when the detected eye positions are used for face normalization. In our experiments, the criterion $err < 0.125$ is adopted. Because the radius of an iris, denoted by r, is about $\frac{1}{4}$ of an eye width, our criterion corresponds to

$$\max(d_l, d_r) < r. \tag{30}$$

The criterion established by Eq. (30) also means that if both the left and right eye positions detected hit irises, the eye detection is considered to be success. This standard is very close to that adopted by Kawaguchi and Rizon [19].

4.2. Determination of parameters α and β

In order to make the proposed method more efficient, the parameter α in Eq. (26) and β in Eq. (28) should be optimized. In our work, they are determined by experiments.

Parameter β in Eq. (28) is used for the selection of potential light dots. If it is too small, for those images without light dots, some other holes, such as those corresponding to whites of eyes may be detected as light dots, which lead to a false detection of light dots. But if β is too large, there may be no light dot candidate at all, so some true light dots may not be extracted. Because in the proposed multi-level eye detection scheme, using the light dots are prior to intensity information for eye localization, only when there is no light dot detected in an eye region, the intensity information is used. This means that false light dots are strictly prohibited, for it may lead to a failure in the eye detection, while failure detection of true light dots is allowed. So, for an image set, suppose that NLD_t and NLD_f, respectively, represent the number of true light dots and that of false light dots detected from it, the criterion for setting β is to maximize the true detection while reducing the false detection to zero, i.e.,

$$\beta^* = \arg\max\{NLD_t(\beta) | NLD_f(\beta) = 0\}. \tag{31}$$

Figs. 10 and 11 show the NLD–β curves for the AR-63 dataset and the Bern image subset, respectively. For AR-63, each AR image has two distinct light dots in its irises, so the actual number of light dots is 126. When β changes from 1.2 to 2.1, both $NLD_t(=123)$ and $NLD_f(=1)$ remain unchanged. When β reaches 2.4, the NLD_f falls to zero; also the NLD_t begins to decrease. For the 150 Bern images, only 77 eyes, mainly in looking-up faces, actually contain light spots and the light dots in Bern images are not as apparent as those in AR images. Fig. 11 shows that the NLD_f for the Bern subset decreases rapidly with the increasing of β. When β reaches 2.4, NLD_f is 3.

From Figs. 10 and 11, we can conclude that the optimal value of parameter β is 2.5. For AR-63, $NLD_t = 122$ and $NLD_f = 0$ when $\beta = 2.5$. For the Bern subset, the corresponding data are 62 and 0, respectively.

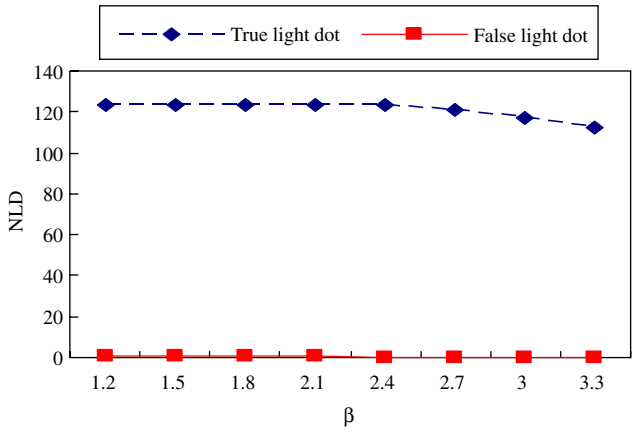


Fig. 10. The number of light dots (NLD) versus β for the AR-63 data set.

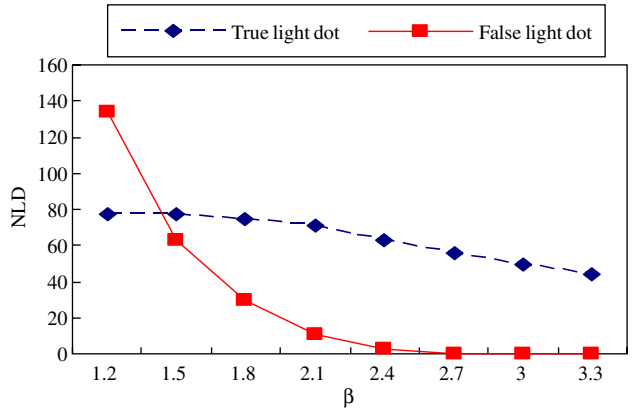


Fig. 11. The number of light dots (NLD) versus β for the Bern image subset.

Parameter α in Eq. (26) may affect the accuracy of eye location. In our study, the average error of eye location for an image set, denoted by $aerr$, is expressed as

$$aerr(\alpha) = \sum_{i=1}^{N} \frac{(d_l(i, \alpha) + d_r(i, \alpha))}{2d_{lr}(i)}, \tag{32}$$

where N is the number of images in the image set. The optimal value of α is determined by

$$\alpha^* = \arg\min(aerr(\alpha)). \tag{33}$$

Without loss of generality, for determining α, we used 128 Bern images in which eyes are localized by using intensity information. Fig. 12 shows the $aerr(\alpha)$–α curve. It shows when $\alpha = 2.0$, the minimum error is obtained. So in our experiments, the α is set to 2.0.

4.3. Experiments using Bern images

Fig. 13 shows some examples of the Bern images for which the proposed method can correctly detect both eyes. Based on the criterion defined by Eq. (30), the correct eye

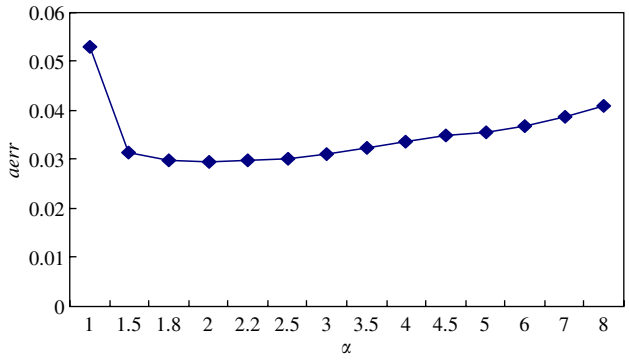


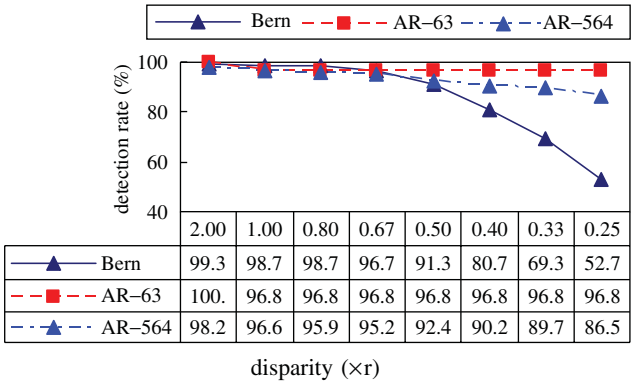
Fig. 12. Eye location error versus parameter α.

Fig. 14. The Bern images for which eyes could not be correctly detected by the proposed method.

detection rate of the proposed algorithm on the 150 Bern images is 98.6%. In other words, of all tested Bern images, there are only two images, shown in Fig. 14, in which eyes cannot be correctly located by our proposed method. Fig. 15 gives a more detailed performance evaluation of the proposed method on the Bern subset. The horizontal axis *disparity* is measured in the radius of an iris, r. Obviously, the less the disparity, the higher the eye location accuracy. Different disparity in Fig. 15 also corresponds to different eye detection criteria similar to that in Eq. (30), i.e.,

$$\max(d_l, d_r) < disparity, \tag{34}$$

where $disparity = 2.0r, 1.0r, 0.8r, 0.67r, 0.50r, 0.40r, 0.33r, 0.25$. From Fig. 15, we can see that of all the 150 Bern images, more than 96% have the eye location error less than $\frac{2}{3}$ of the iris radius, and more than 50% have eye location error



	2.00	1.00	0.80	0.67	0.50	0.40	0.33	0.25
Bern	99.3	98.7	98.7	96.7	91.3	80.7	69.3	52.7
AR–63	100.	96.8	96.8	96.8	96.8	96.8	96.8	96.8
AR–564	98.2	96.6	95.9	95.2	92.4	90.2	89.7	86.5

Fig. 15. Correctly eye detection rate for Bern, AR-63 and AR-564 images according to different criteria.

Fig. 13. Examples of Bern face images for which the eyes are correctly detected by the proposed algorithm.

Table 1
Comparison of the correct eye detection rate of different methods tested using 150 Bern images

Algorithm	Detection rate (%)
Our proposed method[a]	98.7
Kawaguchi and Rizon [19]	95.3
Template matching [19]	77.9
Eigenface method using 50 training samples [19]	90.7
Eigenface method using 100 training samples [19]	93.3

[a]The criterion described in Eq. (30) is adopted.

less than $\frac{1}{4}$ of the iris radius. This indicates that the proposed method achieves not only high eye detection rate, but also high eye location accuracy. Fig. 15 also shows that if the criterion proposed by Jesorsky et al. [28] is adopted, the detection rate can reach 99.3%.

Table 1 shows the performance comparison of our proposed eye detection method with Kawaguchi and Rizon's method [19], the template matching method, and the eigenface method. It is clear that our algorithm obtains a higher correct eye detection rate than other methods. When used to detect eyes from faces looking downwards, the method proposed by Kawaguchi and Rizon [19] deteriorates. This deterioration does not occur in our method. This indicates that our method is more effective when exposed to a variation in views and gaze directions than other methods.

4.4. Experiments using AR images

In order to further compare the performance of our method with that of Kawaguchi and Rizon's method [19], we first use AR-63 for experiments.

Some examples of AR faces for which the proposed method succeeds in eye detection are shown in Fig. 16. Also, four negative samples of AR faces are shown in Fig. 17. The detailed performance of the proposed method for AR-63 is also shown in Fig. 15. It can be seen that the proposed method achieves a success rate of 96.8% based on the criteria defined in Eq. (30). This eye detection rate is the same as that reported by Kawaguchi and Rizon [19]. One important feature of our proposed method is that for AR-63 our method obtains very high location accuracy. From Fig. 15, we can see that even the location precision increases to $\frac{1}{4}$ of the iris radius, the successful eye detection rate can still reach 96.8%.

Because parameters α and β used in the proposed method are determined using the Bern subset and AR-63 images,

Fig. 16. Examples of AR images for which two eyes are correctly detected by the proposed method (two images from each group).

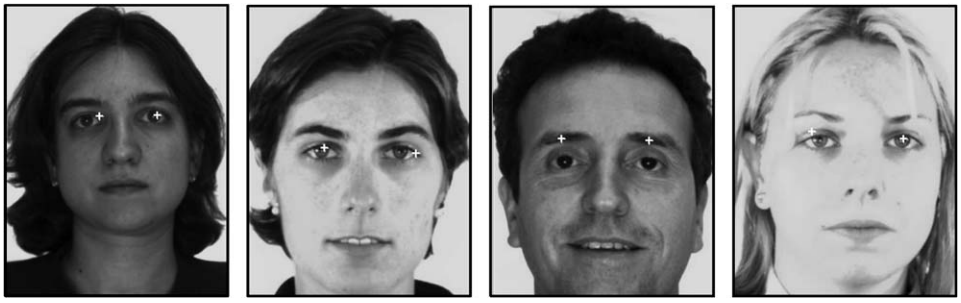


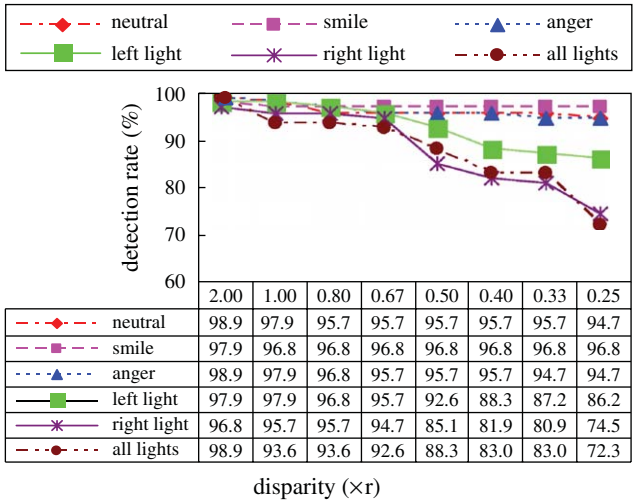
Fig. 17. Examples of AR images for which eyes are incorrectly detected by the proposed method.

the following experiments on AR-564 can be viewed as an extensive performance test of the proposed method. For AR-564, the six groups are examined separately. The detailed results are shown in Fig. 18. The detection rates for the six groups, i.e., neutral, smile, anger, left light on, right light on, all side lights on, are 97.9%, 96.8%, 97.9%, 97.9%, 95.7%, 93.6%, respectively.

It can be concluded from Fig. 18 that for the AR images, (1) the proposed method can achieve both high eye detection rate and high eye location accuracy. Even in the worst situation, i.e., all side lights on, the eye detection rate can still reach 93.6%, and of all the 94 images, more than 72% have an eye location disparity of less than $\frac{1}{4}$ of the iris radius; (2) the influence of facial expressions on the eye detection rate is less than that caused by an illumination change. This can be verified with the fact that compared with that of the "neural"group (97.9%), the detection rate of the "'smile" group (96.8%) and that of the "anger" group (97.9%) change very little, while the detection rate of the "right light on" group (95.4%) and that of the "all side lights on" group (93.6%) show a more obvious change. This trend becomes more distinct when the precision increase to $0.25r$.

The high performance of the proposed method on AR images can be explained with the high detection rate of light dots. As mentioned above, each AR image shows two clear reflected light dots in the irises, and each of these light dots are very close to the center of the iris. Using the light dot detection algorithm described in Section 3.7, the proposed method can reliably detect the light dots from AR images, especially from those images in the first three groups. Table 2 shows the number of images of AR-63 and AR-564 for which both eyes are located by light dots. It indicates that the light dot detection rate for the total 564 AR images in AR-564 is 84.4% and the rate for the first three groups is over 93%.

Table 2 also shows that the light dot detection rates for the "smile" and "anger" group are very close to that for the "neutral" group, while those for the last three groups decrease significantly. This can be used to explain the results that the influence of lighting changes on the proposed eye detection rate is larger than that caused by expression changes, especially when the required precision is high.



	2.00	1.00	0.80	0.67	0.50	0.40	0.33	0.25
neutral	98.9	97.9	95.7	95.7	95.7	95.7	95.7	94.7
smile	97.9	96.8	96.8	96.8	96.8	96.8	96.8	96.8
anger	98.9	97.9	96.8	95.7	95.7	95.7	94.7	94.7
left light	97.9	97.9	96.8	95.7	92.6	88.3	87.2	86.2
right light	96.8	95.7	95.7	94.7	85.1	81.9	80.9	74.5
all lights	98.9	93.6	93.6	92.6	88.3	83.0	83.0	72.3

Fig. 18. Eye detection rate for the AR-564 based on different disparity criteria.

Table 2
Number of AR images for which two eyes are located by light dots

Image subset	AR-63	AR-564						
		Neutral	Smile	Anger	Left light on	Right light on	All side lights on	Total
Number of images	61	88	91	89	79	66	63	476
Ratio to the total images (%)	96.8	93.6	96.8	94.7	84.0	70.2	67.0	84.4

The reduction of the light dot detection rate for the last three groups of AR-564 can be due to the decreasing hole contrast caused by the strong illumination. It is easy to understand that when a face image is illuminated by light, the denominator AI in Eq. (27) may increase, which leads to the contrast C decline. Table 3 shows a typical case of contrast C of light dot-related holes under different illumination conditions. It indicates that the contrast of the light dot in the light illuminated side decreases significantly. When the

Table 3
The contrast C of light dot-related holes versus lighting conditions

Lighting conditions	Left light dot	Right light dot
Natural	5.11	5.22
Left light on	3.64	5.01
Right light on	5.32	3.81
All side lights on	3.64	3.26

contrast reduces to a level less than the threshold β, this light dot may not be detected.

The overall performance of the proposed method on AR-564 is summarized in Fig. 15. An eye detection rate of 96.6% is achieved according to the criterion defined in Eq. (30). The ratio of images with eye detection error less than $\frac{1}{4}$ of the iris radius is 86.5%. This result is promising and indicates that the proposed method is effective when it is exposed to changes in expression and illumination, particularly the former. From Fig. 16, we can still see that by making use of light dots for eye detection, our method can correctly locate those eyes that are partially occluded by hair (see the first and the last image in the first row) and those eyes that are closely connected with the eyebrows (see the second image in the second row).

5. Discussion

Compared with the existing eye detection approaches, the proposed method has made the following improvements:

(1) The problem of eye detection from face images with different views has been carefully studied and the results are reported in this paper. Our method can reliably detect eyes not only from frontal-view face images, but also from face images of other views, such as looking to the left, to the right, downward and upward. The result is satisfactory and is better than that obtained by Kawaguchi and Rizon [19].
(2) The proposed method worked well when it was used for eye detection from images with different facial expressions, e.g., neutral, smile and anger. In particular, because the light dot in the iris is used for eye detection in our method, eyes with different appearances, or partially occluded by hairs can be accurately located. Obviously, these are difficult problems for shape-based eye detection methods, e.g., template matching and Hough transform methods.
(3) The influence of lighting conditions on eye detection was specifically investigated in this paper. Our method can obtain high performance in most illumination environments, making it more applicable than some intensity-based approaches.
(4) A reflected light dot in the iris is used for eye detection and a robust algorithm for the automatic light spot detection is presented. Reflected light dots often appear in irises, especially when a person is facing a brighter front. The existence of a light spot greatly changes the distribution of pixel gray levels in the iris, so it causes difficulty for most existing eye detection methods. Kawaguchi and Rizon [19] tried to remove light dots from images, but they did not provide the details of the automatic light dot detection method used. Making use of the light dot for eye detection, our method can achieve not only a high detection rate, but also a high detection accuracy.
(5) The cues used for eye detection in our proposed method include image intensity, light dot, the shape of eye segment in the BEI and the configuration of face components. Obviously, all these are the intrinsic attributes of a face image and eyes, and have nothing to do with the image size. So, unlike the conventional template matching and Eigenspace approaches, our proposed method requires no normalization to the face image and it can tolerate larger image rotation in plane and depth.

The success of the proposed method is mostly due to high-quality BEIs produced by our proposed method. The improved wavelet-based edge extraction method used in our method includes two binarization steps and a noise removing step. The BEI is robust to lighting conditions. Different components in the resulting BEI are well separated and pixels of the same face component are well connected. This greatly facilitates the segmentation of face components and the detection of some key feature points, such as reflected light spots.

6. Conclusions

Eye detection is a very important step for the establishment of an automatic human face recognition system. It is also a very challenging research topic due to its complexity. Although many efforts have been dedicated to this study, the problem is still far from being solved. In this paper, a novel eye detection method is presented. Our method consists of three steps: (1) binary edge images (BEIs) are firstly extracted from the grayscale image, (2) the eye region and eye segments are extracted from BEIs and (3) the light dot or intensity information is utilized for the eye localization. In order to enhance the eye detection performance, a face region refinement algorithm and a multi-level eye detection scheme have been proposed. Experimental results show that a correct eye detection rate of 98.7% can be achieved for 150 Bern face images without spectacles and 96.6% for 564 AR images.

Based on the BEI and the proposed eye detection method, much more work can be done. For example, automatic location of face landmarks from the BEI is a very meaningful research direction. Another very meaningful research

direction is to study lighting conditions in which reflected light spots may appear, and further develop a light dot-based eye detection system. This system is very similar to an infrared-based system, but needs no special image capturing device and just uses an ordinary camera, so it will be more applicable.

Acknowledgements

The work described in this paper was partially supported by a research grant from the Hong Kong Polytechnic University (G-T768), a grant from Zhejiang Provincial Natural Science Foundation of China (602118), and a grant from the Scientific Research Foundation for the Returned Overseas Chinese Scholars of Zhejiang Province (72).

References

[1] M.-H. Yang, D.J. Kriegman, N. Ahuja, Detecting faces in images: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 24 (1) (2002) 34–58.
[2] R.-L. Hsu, M. Abdel-Mottaleb, A.K. Jain, Face detection in color images, IEEE Trans. Pattern Anal. Mach. Intell. 24 (5) (2002) 696–706.
[3] A. Samal, P.A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: a survey, Pattern Recognition 25 (1) (1992) 65–77.
[4] R. Brunelli, T. Poggio, Face Recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042–1052.
[5] Y. Gao, K.H. Maylor, Face recognition using line edge map, IEEE Trans. Pattern Anal. Mach. Intell. 24 (6) (2002) 764–779.
[6] B. Fasel, J. Luettin, Automatic facial expression analysis: a survey, Pattern Recognition 36 (2003) 259–275.
[7] Y. Tian, T. Kanade, J.F. Cohn, Recognizing action units for facial expression analysis, IEEE Trans. Pattern Anal. Mach. Intell. 23 (2) (2001) 97–115.
[8] C.C. Han, H.Y.M. Liao, G.J. Yu, L.H. Chen, Fast face detection via morphology-based pre-processing, Pattern Recognition 33 (2000) 1701–1712.
[9] J. Wu, Z.H. Zhou, Efficient face candidates selector for face detection, Pattern Recognition 36 (2003) 1175–1186.
[10] J. Huang, H. Wechsler, Visual routines for eye location using learning and evolution, IEEE Trans. Evol. Comput. 4 (1) (2000) 73–82.
[11] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.
[12] Z. Zhu, K. Fujimura, Q. Ji, Real-time eye detection and tracking under various light conditions, in: ACM SIGCHI symposium on eye tracking research and applications, New Orleans, LA, USA, 2002.
[13] A Haro, M. Flickner, I. Essa, Detecting and tracking eyes by using their physiological properties, dynamics, and appearance, in: IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR 2000), South Carolina, USA, 2000, pp. 163–168.
[14] C.H. Morimoto, D. Koons, A. Amir, M. Flickner, Pupil detection and tracking using multiple light sources, Image and Vis. Comput. 18 (2000) 331–335.
[15] D.J. Beymer, Face recognition under varying pose, in: IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle, Washington, USA, 1994, pp. 756–761.
[16] A. Pentland, B. Moghanddam, T. Starner, View-based and modular eigenspaces for face recognition, in: IEEE Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR'94), Seattle, Washington, USA, 1994, pp. 84–91.
[17] A.L. Yuille, P.W. Hallinan, D.S. Cohen, Feature extraction from faces using deformable template, Int. J. Comput. Vis. 8 (2) (1992) 99–111.
[18] K.M. Lam, H. Yan, Locating and extracting the eye in human face images, Pattern Recognition 29 (5) (1996) 771–779.
[19] T. Kawaguchi, M. Rizon, Iris detection using intensity and edge information, Pattern Recognition 36 (2003) 549–562.
[20] G. Chow, X. Li, Towards a system for automatic facial feature detection, Pattern Recognition 26 (1993) 1739–1755.
[21] G.C. Feng, P.C. Yuen, Multi-cues eye detection on gray intensity image, Pattern Recognition 34 (2001) 1033–1046.
[22] G.C. Feng, P.C. Yuen, Variance projection function and its application to eye detection for human face recognition, Pattern Recognition Lett. 19 (1998) 899–906.
[23] Z.H. Zhou, X. Geng, Projection functions for eye detection, Pattern Recognition 37 (5) (2004) 1049–1056.
[24] S.A. Sirohey, A. Rosenfeld, Eye detection in a face image using linear and nonlinear filters, Pattern Recognition 34 (2001) 1367–1391.
[25] B. Achermann. The face database of University of Bern, http://iamwww.unibe.ch/~fkiwww/staff/achermann.html, Institute of Computer Science and Application Mathematics, University of Bern, Switzerland, 1995.
[26] A.M. Martinez, R. Benavente, The AR Face Database, CVC Technical Report #24, June 1998.
[27] S.Y. Lee, Y.K. Ham, R.H. Park, Recognition of human faces using knowledge-based feature extraction and neuro-fuzzy algorithm, Pattern Recognition 29 (11) (1996) 1863–1876.
[28] O. Jesorsky, K.J. Kirchberg, R.W. Frischholz, Robust Face detection using the Hausdorff distance, in: Proceedings of the Third International Conference on Audio- and Video-based Biometric Person Authentication, Halmstad, Sweden, 2001, pp. 90–95.

About the Author—JIATAO SONG received his PhD degree in Information and Communication Engineering from Zhejiang University, China in 2005. From 2003 to 2004, he worked as a Research Assistant at the Hong Kong Polytechnic University. He is currently an Associate Professor in the College of Electronic and Information Engineering, Ningbo University of Technology, China. His research interests include image processing, pattern recognition and computer vision.

About the Author—ZHERU CHI received his BEng and MEng degrees from Zhejiang University in 1982 and 1985, respectively, and his PhD degree from the University of Sydney in March 1994, all in electrical engineering. Between 1985 and 1989, he was on the Faculty of the Department of Scientific Instruments at Zhejiang University. He worked as a Senior Research Assistant/Research Fellow in the Laboratory for Imaging Science and Engineering at the University of Sydney from April 1993 to January 1995. Since February 1995, he has been with the Hong Kong Polytechnic University, where he is now an Associate Professor in the Department of Electronic and Information Engineering. Since 1997, he has served as a Session Organizer/Session Chair/Area Moderator/Program Committee Member for a number of international conferences. He was one of contributors to the Comprehensive Dictionary of Electrical Engineering (CRC Press and IEEE Press, 1999). His research interests include image processing, pattern recognition, example-based machine translation, and computational intelligence. Dr. Chi has authored/co-authored one book and eight book chapters, and published more than 140 technical papers. Dr. Chi is a member of IEEE.

About the Author—JILIN LIU received his BS degree from Tsinghua University in 1970 and MS degree from the Graduate School of Academy Sinica in 1981, respectively, both in Electronic Engineering. He was a Visiting Scholar at Technical University Berlin, Germany from 1985 to 1987 and then at UC Berkeley, USA in 1999. Professor Liu is currently the Director of Institute of Information and Communication Engineering in the Department of Information Science and Electronic Engineering, Zhejiang University. He is an IEEE member and has authored/co-authored over 80 publications. His main research focuses on multimedia, image processing, pattern recognition and intelligent transportation systems, especially on video navigation of intelligent vehicles.