A novel algorithm for automatic localization of human eyes

Liang Tao (陶 亮)[1,2], Juanjuan Gu (顾涓涓)[3], and Zhenquan Zhuang (庄镇泉)[2]

[1] *Department of Electronic Engineering and Information Science, Anhui University, Hefei 230039*

[2] *Department of Electronic Science and Technology, University of Science and Technology of China, Hefei 230026*

[3] *Department of Computer and Information Engineering, Hefei Association University, Hefei 230022*

Received July 22, 2003

Based on geometrical facial features and image segmentation, we present a novel algorithm for automatic localization of human eyes in grayscale or color still images with complex background. Firstly, a determination criterion of eye location is established by the prior knowledge of geometrical facial features. Secondly, a range of threshold values that would separate eye blocks from others in a segmented face image (i.e., a binary image) are estimated. Thirdly, with the progressive increase of the threshold by an appropriate step in that range, once two eye blocks appear from the segmented image, they will be detected by the determination criterion of eye location. Finally, the 2D correlation coefficient is used as a symmetry similarity measure to check the factuality of the two detected eyes. To avoid the background interference, skin color segmentation can be applied in order to enhance the accuracy of eye detection. The experimental results demonstrate the high efficiency of the algorithm and correct localization rate.

OCIS codes: 100.5010, 330.1880, 100.2000.

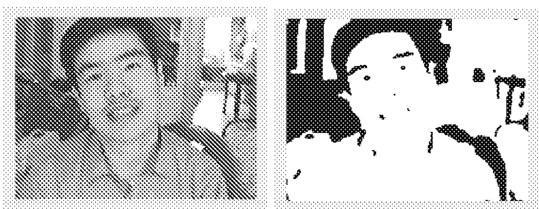
In an automatic face recognition system, the features or representation of a face are extracted automatically from an input face image and then compared in a matching process[1,2]. Since face recognition is often based on the normalized face images and normalizing input face images mainly depend on eye location, the localization of human eyes in face images is a fundamental step in the process of face recognition[3]. A common method used for locating eyes is the deformable templates, proposed by Yuille *et al.*[4] and improved by Xie *et al.*[5]. This method makes use of global information and hence improves the reliability of locating the contour of the eye. However, the template scheme is associated with problems such as slow convergence and lengthy processing time. In addition, for almost all the existing algorithms for locating eyes, face location must be known before locating eyes[6]. In this paper, we present a novel algorithm to achieve fast accurate eye detection that is robust to changes in illumination and background. There is no need to know the face location before the proposed algorithm can be applied to locating eyes. The proposed algorithm is based on geometrical facial features and image segmentation and works on color or grayscale still images (256 gray levels). To avoid the background interference, skin color segmentation can be applied in order to enhance the accuracy of eye detection.

To detecting and locating human eyes automatically and efficiently, suppose that face images are taken under the following reasonable conditions: only one face appears in each image, and the face takes up 15% to 50% of the whole pixels in the image. Images are taken against plain or complex background with faces in quasi-frontal position (with tolerance for some side movement). Eyes in face images should be visible and clear. Wearing glasses is allowed, but no reflection of light and no black frame of glasses are expected because the reflection and black frame of glasses could seriously affect the results of detecting and locating eyes.

Human faces have a special pattern that is usually different from the patterns of background objects in face images. The grayscales of the pupil and the iris of an eye are usually lower than those of the skin near the eye and those of the white of the eye; therefore, if we can find an appropriate threshold value to segment a face image, the eyes can be separated from other facial features and background objects in the segmented face image (i.e., the binary image, where, if the grayscale of a pixel is more than or equal to the threshold value, the normalized grayscale of the pixel will be set to be 1 (white pixel); otherwise, it will be set to be 0 (black pixel)). An example is shown in Fig. 1.

The connected components (black pixels) in the segmented face image are called a block. The morphological operations of "majority"[7] need to be performed on the segmented face image in order to remove some tiny blocks. All the blocks are labeled. The number and the position of each block are recorded. To locate eyes from an appropriately segmented face image, a determination criterion of eye location needs to be established by the prior knowledge of geometrical facial features as follows.

- The distance between the geometrical centers of the two eye blocks should be within a certain range of pixel number such as from 15 to 45 pixels in a face image with size 120×160 pixels.
- There are no other blocks in a certain area under each eye.



(a) (b)

Fig. 1. (a) An grayscale face image and (b) its segmented face image with an appropriate threshold value 0.31.

• The vertical distance difference between the geometrical centers of the two eye blocks is not more than a certain number of pixels.

• The size (the pixel number) in each eye block is limited in a certain range.

• There is no other block between the two eye blocks.

• The proportion of height to length in the rectangular bounding box around each eye block is limited in a certain range.

• Any block connected with or very close to the four edges of face images is not an eye block.

Next we need to find out the so-called optimal threshold value, which is an appropriate threshold value used to segment a face image so that the two eyes can be separated from other facial features and background objects. However, it is very difficult to estimate the optimal threshold value directly, instead, we suppose that the optimal threshold value is in a range from T_0 to $T_{\max}$. Statistics on more than 1000 face images shows that in most clear face images with plain or complex background, T_0 and $T_{\max}$ can be set to be 0.1 and 0.6, respectively. When images are taken in a fixed background and an unchanged illumination condition, the range $(T_0, T_{\max})$ can be limited in an even smaller range.

After the range $(T_0, T_{\max})$ is determined, a novel algorithm for detecting and locating eyes is presented as follows. With the progressive increase of the threshold value by a small step T_{step} in the range from T_0 to $T_{\max}$, we can see that in a segmented face image the size of the existing blocks will expand, some existing blocks will merge

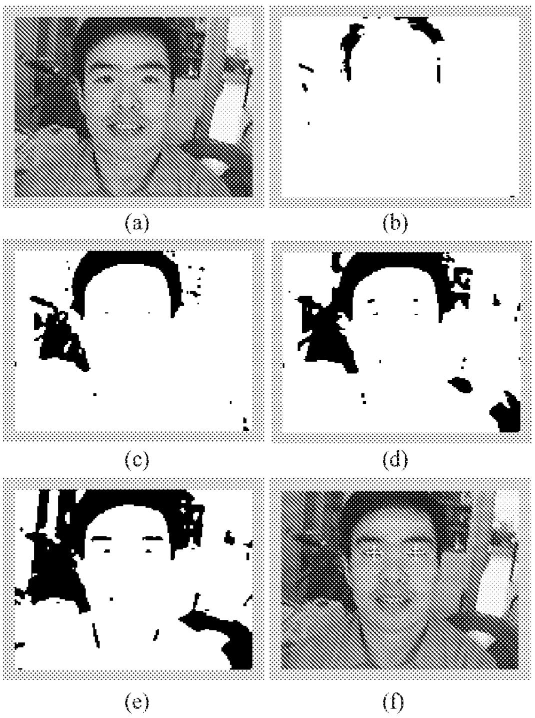


Fig. 2. Detecting and locating human eyes in segmented face images based on automatic search for the optimal threshold value. (a) Original facial image, (b)∼(e) segmented face images with threshold values from 0.12 to 0.3, and the threshold step is 0.06. When the threshold value becomes 0.3 in image (e), eye blocks are detected. (f) The result of locating eyes (displayed by white crosses).

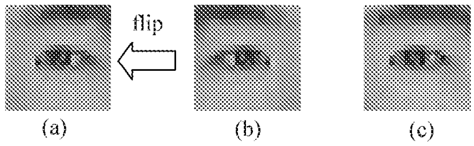


Fig. 3. (a) The flipped right eye B_R, (b) the right eye A_R, and (c) the left eye A_L.

into one block, and some new blocks will emerge in the segmented face image (Fig. 2). Once two eye blocks appear from the segmented image and spread to a certain size, they will be detected by the determination criterion of eye location. However, sometimes, we found that in the detection process the two detected eyes might not be real eyes because of the interference of complex image background. Therefore, the 2D correlation coefficient is used as a similarity measure to check the factuality of the two detected eyes.

First of all, taking out a small sub-image with the size of $M \times N$ (such as 21×21, see Fig. 3), denoted with A_L and A_R respectively for the left eye and the right eye, from the original grayscale face image at the geometrical center of the each detected eye block. Secondly, flipping the matrix A_R in left direction to get a new matrix B_R. Finally, computing the 2D correlation coefficient r as follows

$$r = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[A_L(m,n)-\overline{A}_L][B_R(m,n)-\overline{B}_R]}{\sqrt{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[A_L(m,n)-\overline{A}_L]^2\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}[B_R(m,n)-\overline{B}_R]^2}},$$

where $\overline{A}_L$ and $\overline{B}_R$ are respectively the mean of the elements of A_R and B_R.

At the beginning of the process of the progressive increase of threshold value for seeking the optimal threshold, r is set to be zero. During the process, if finding $r \geq 0.5$, the detected two eyes are usually true. If finding $r < 0.5$, the process will continue until finding $r \geq 0.5$. If finding $0 < r < 0.5$ always in the whole process, the maximum of r, denoted with $r_{\max}$, will be selected and the two detected eyes corresponding to $r_{\max}$ will be taken as true eyes. If finding $r \leq 0$ always in the whole searching process, no true eyes are detected. In this way, the optimal threshold can be automatically found such that eyes can be accurately located. The procedure of the algorithm is shown in Fig. 4.

We used a face database (50 people, 250 quasi-frontal face images taken in our laboratory using a web camera) to evaluate the performance of the proposed algorithm. The algorithm was programmed by the code of Matlab 6.1 and worked on a Pentium III/450M personal computer. The average runtime is about 1 second per image and the correct localization rate is close to 98%. The main reason for detection failure is that there are some patterns in image backgrounds that are similar to human face patterns. We also tested 28 non-face images. The proposed algorithm gave only two false alarms also because there was a false facial pattern in each of the two non-face image backgrounds. Figure 5 displays some correct examples of locating eyes using the proposed

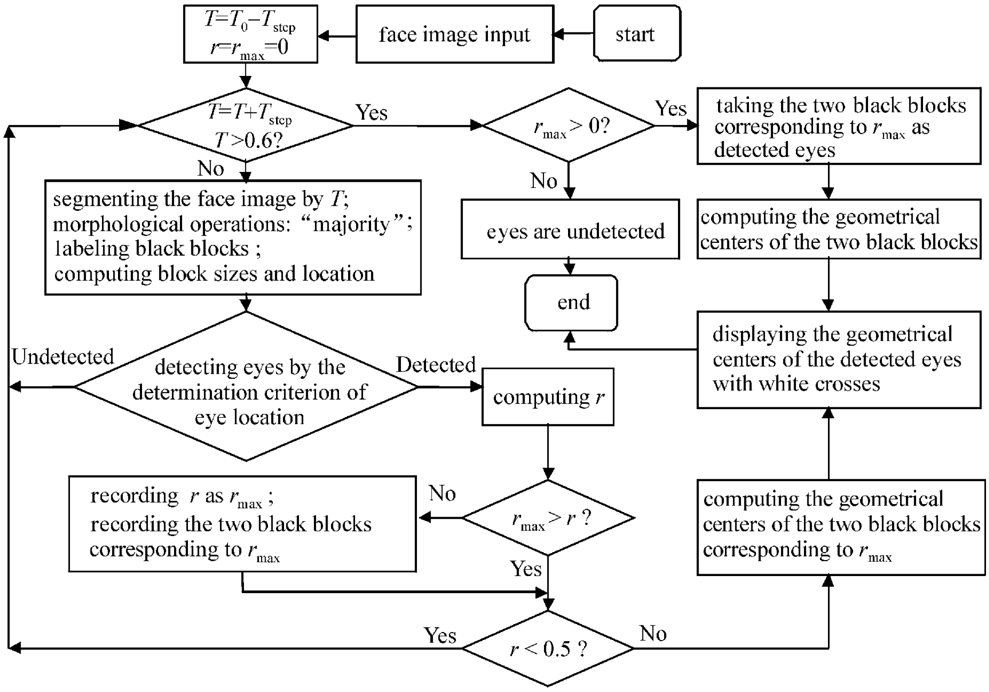


Fig. 4. The procedure of the proposed algorithm.

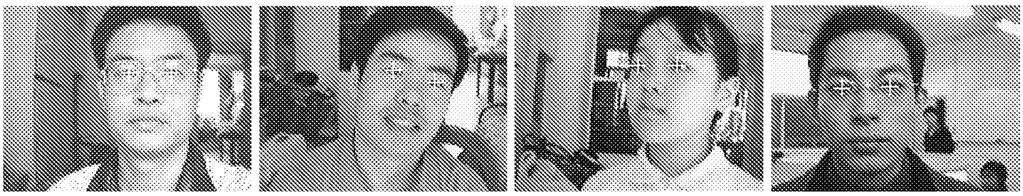


Fig. 5. Several examples for locating eyes (Eye positions are denoted with white crosses).

algorithm. More examples can be found in our published conference paper[8].

To avoid the background interference, skin color segmentation can be applied here. Compared to the skin color models commonly used in the current face detection algorithms[6], the skin color model we used is not for detecting face but for reducing the range of detecting eyes in order to enhance the accuracy of eye detection. Therefore, we do not need to build a rigorous skin color model; in other words, the model we used can be more robust to illumination variations.

Let a color face image I_c with RGB values be normalized to $r = R/(R+G+B)$, and $g = G/(R+G+B)$ at each pixel (x, y), we determine the skin tone color space by manually selecting skin regions in 100 color face images and histogramming their r and g values, giving $H_i(r, g)$, where subscript i is the image number. The average histogram H is

$$H = \frac{1}{100} \sum_{i=1}^{100} H_i(r, g), \tag{1}$$

it is shown as a 2D scatter plot in Fig. 6, where the skin-tone color space takes up a very small region. A simple rectangular bounding box around the region is taken for

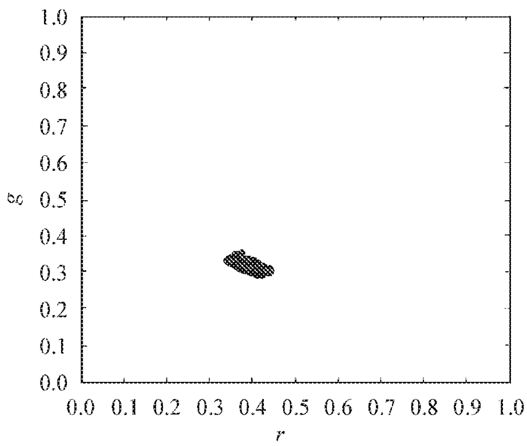


Fig. 6. Average scatter plot of skin tone, H.

a skin color space D_{skin}, which is used to segment the face image I_c, giving a binary image, called a mask image I_{mask}

$$I_{\text{mask}} = \begin{cases} 1 & \text{if at the pixel } (x, y) \text{ of } I_c, \ (r, g) \in D_{\text{skin}} \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

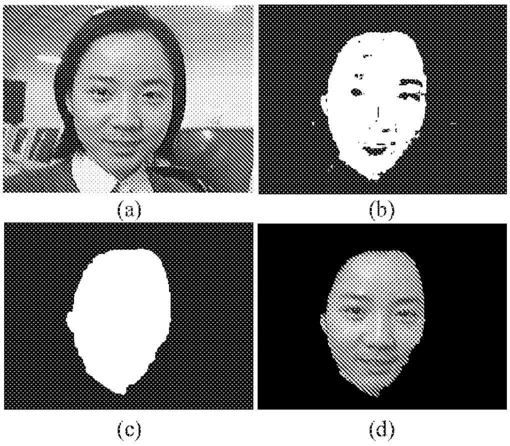


Fig. 7. (a) Original color face image, (b) its mask image with small black blocks in the white region, (c) its new mask image with the black blocks removed, and (d) the segmented grayscale face image I.

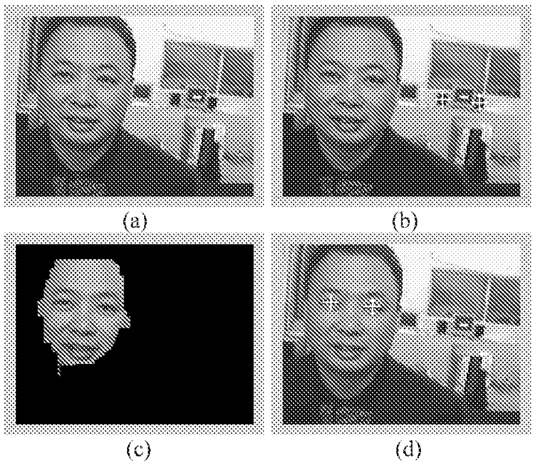


Fig. 8. (a) Original color face image, (b) the eye localization failure due to the background interference, (c) the grayscale face image I obtain by using the skin color segmentation, and (d) the correct eye localization by using Fig. 8(c).

Figure 7 displays a face image and its mask images. In the mask image Fig. 7(b) we can see that there are some black blocks in the white region that are not expected to appear. To remove them, the morphologic operations of both "dilate" and "erode"[7] on Fig. 7(b) can be applied for 10 or 15 times, giving its new mask image I_{mask} with the black blocks removed (Fig. 7(c)). Now we can get the segmented face image I_0 by $I_0(x,y) = I_{\text{mask}}(x,y) \times I_c(x,y)$, and then transfer the color image I_0 into a grayscale image I (Fig. 7(d)), which is ready for eye detection and localization. Figure 8 gives an experimental example displaying the efficiency of the skin color segmentation method, by which we reduce the range of detecting eyes to the skin region and avoid the background interference with eye localization in the grayscale image (Fig. 8(b)). The runtime in this experiment is about 1.3 seconds for the correct eye localization (Fig. 8(d)).

Unfortunately, so far, no international face database has been available for eye detection and localization, which impedes the direct comparison of our proposed algorithm with other algorithms.

In summary, by automatic search for the optimal threshold value, a novel algorithm is developed in this paper for locating human eyes based on the determination criterion of eye location and the similarity measure of two eyes. The skin color segmentation is also applied to remove the affection of false facial patterns in backgrounds. The experimental results demonstrated the efficiency of the algorithm and correct localization rate.

This research was supported by the Excellent Young Teachers Program of the Ministry of Education, P. R. China, the National Natural Science Foundation of China (No. 60375010), and the Excellent College Teachers Program of the Education Committee in Anhui Province. L. Tao's e-mail address is taoliang@ustc.edu.

References

1. A. Samal and P. A. Iyengar, Pattern Recognition **25,** 65 (1992).
2. R. Chellappa, C. L. Wilson, and S. Sirohey, Proceedings of IEEE **83,** 705 (1995).
3. S. G. Shan, W. Gao, and X. L. Chen, J. Software (in Chinese) **12,** 570 (2001).
4. A. L. Yuille, D. Cohen, and P. Hallinan, in *Proceedings of IEEE CVPR'89* 104 (1989).
5. X. Xie, R. Sudhakar, and H. Zhuang, Pattern Recognition **27,** 791 (1994).
6. E. Hjelmås, Computer Vision and Image Understanding **83,** 236 (2001).
7. Y. J. Zhang, *Image Segmentation* (in Chinese) (Science in China Press, Beijing, 2001).
8. L. Tao, J. J. Gu, Q. W. Gao, and Z. Q. Zhuang, Proc. SPIE **4875,** 710 (2002).