7659, England, UK
Journal of Information and Computing Science
Vol. 2, No. 3, 2007, *pp*. 209-214

Eye Location and Eye State Detection in Facial Images with Unconstrained Background *

Qiong Wang +, Jingyu Yang

School of Computer, Nanjing University of Science & Technology, Nanjing 210094, China

(*Received 30 June 2006, accepted 5 October 2006*)

Abstract. This paper presents an efficient approach to achieve fast and accurate eye states detection in still gray level images with unconstrained background. The structure of eye region is used as a robust cue to find eye pair candidates. Eyes are located by eye verification using SVMs. The eye contour information is used to detect whether eyes are open or closed after locating eyes. The proposed method is robust to deal with illumination changes, moderate rotations, glasses wearing and partial face occlusions. The proposed method is evaluated on the BioID face database and experimental results demonstrate its effectiveness.

Keywords: Eye location, Eye state detection, Binary template matching, SVM, Hough transform

1. Introduction

Eyes are the most important features of human face. Eye states detection has received a great deal of attention. There are many applications of the robust eye states extraction. For example, the eye states provide important information for recognizing facial expression, human-computer interface systems and driver fatigue monitoring system.

The eye states can be obtained from the eye features such as the inner corner of eye, the outer corner of eye, iris, and eyelid. There are many methods to detect eye features. Deng and Lai [1] used improved deformable templates to locate eye features. The deformable templates can extract above eye features. However, the deformable templates base on minimizing energy; it will take much time to detect eye states. Tian, Kanade and Cohn [2] developed a dual-state model based system of tracking eye features. They detected eye states by using half iris circle mask. Their approach could distinguish two eye states, open and closed. Bernogger et al. [3] presented an approach to synthesize the eye movement by using the extracted eye features to compute the deformation of the eyes of the 3D model. They used the deformable templates to detect eye features. When they constructed the energy function the color information was used. Liu, Wu and Zha [4] used color information and edge map to detect iris. The saturation of color is used to detect whether the eye is open or closed, then the edge map of the eye image is used to detect the irises.

Many works for eye or iris detection assume either that eye windows have been extracted or rough face regions have been already located. This paper proposes an efficient approach for eye states detection in still gray level images with unconstrained background. Since color images can be turned into gray level ones easily, the proposed method can be applied to color images too. However, color images contain more useful information than gray level ones and can be processed more efficiently. In our method, eye pairs are located before detecting eye states. The structure of eye region is used as a robust cue to locate eyes. Eyes are extracted by using binary template matching and SVMs. After locating eyes, the eye contour information is used to detect dual eye states: open or closed. Our method is robust to illumination changes, partial face occlusions, facial expressions and moderate rotations.

The rest of this paper is organized as follows. The proposed method is given in Section 1. Experimental results are reported in Section 2 and conclusions are drawn in Section 3

2. Eye Location

Before the eye states are detected, eyes must be located and eye images must be obtained. We consider

* This work is supported by the Project of Science and Technology Plan of Jiangsu Province (Grant No. BG2005008).
+ Corresponding author. *E-mail address*: nustdaisy@gmail.com.

that the structure of eye region is a stable and robust feature which can distinguish eye pair from other patterns. The proposed method uses this robust cue to locate eye pairs. Firstly, the facial image is enhanced and binaried to obtain the structure image. We define the structure image as the binary image which contains the structure feature of human face. Then a binary eye pair template is used to find eye pair candidates in the image. All the eye pair candidates are then rescaled to a fixed size and sent to the SVM classier, which can verify the candidates and obtain real eye pairs. Finally, eyes are located according to the verification results.

2.1. Preprocessing

The facial images which will be processed in template matching step are binary images which contain the facial structure information. To obtain good-segmented binary images, the preprocessing consisting of four ordered steps is applied to the input images. In order to compensate the illumination and obtain more image details, the homomorphic filter is used to enhance the brightness and the contrast of the images. Then we use clustering algorithm to divide the facial feature from the skin and background, and the binaried images are obtained through thresholding.

2.1.1 Homomorphic Filtering

Homomorphic filtering is a generalized technique for nonlinear image enhancement and correction. It simultaneously normalizes the brightness across an image and increases contrast.

An image can be expressed as the product of illumination and reflectance:

$$f(x,y) = i(x,y) \cdot r(x,y) \tag{1}$$

When the illumination is uniform, $i(x,y)$ is considered to be a constant and the image is considered to the reflectance of the object. However, the lighting condition is usually unequal. The illumination component tends to vary slowly and its frequency fastens on low part in the frequency domain; the reflectance tends to vary rapidly and its frequency is in high part. If two components can be operated separately, the illumination problem will be solved and the image will be enhanced. Hence, the log transform is used to equation (1):

$$\ln f(x,y) = \ln i(x,y) + \ln r(x,y) \tag{2}$$

Then Fourier transform is used to equation (2), which makes the succedent operation is in frequency domain. The basic homomorphic filtering procedure is as follow:

$$f(x,y) \Rightarrow \ln \Rightarrow FFT \Rightarrow H(u,v)$$
$$\Rightarrow (FFT)^{-1} \Rightarrow exp \Rightarrow g(x,y)$$

the illumination and relectance turn to additive though log transform. Then 2-D Fourier transform is used and the coordinate variables are u and v. $H(u,v)$ is the homomorphic filter function applied to the illumination and reflectance, respectively. After taking the inverse Fourier transform and exponent transform, we get enhanced image $g(x,y)$. $H(u,v)$ used in this paper has the following form:

$$H(u,v) = (H_H - H_L) \cdot (1 - exp(-C \cdot \frac{D}{D_0})) + H_L \tag{3}$$

If the parameters H_L and H_H are chosen to be H_L <1 and H_H >1, then the filter $H(u,v)$ will decrease the contribution of the low frequency (illumination) and amplify the contribution of mid- and high frequencies (reflectance). As shown in Fig.2, Fig.2 (a) is the input image with low contrast due to the illumination, and it can't obtain good segmentation result directly. Fig.2 (b) demonstrates the image enhanced by homomorphic filtering, the contrast is improved and the details in face region are enhanced.

2.1.2 Clustering and Thresholding

We divide the features of interest from the skin and the background by clustering the grey level image into three clusters through the K-Mean Clustering algorithm. The lightest grey level representing the background or other light pixels is set to 255, the intermediate representing the skin and is set to 128, and the darkest representing both the features and other dark pixels of the image (for example the hair and the beard) is set to 0. Fig.1 (c) shows the facial image processes by clustering algorithm.

After clustering, a threshold is set to 128. Then a binary image is obtained, which contains the facial structure obviously. Fig .1 (d) shows the thresholding result. Considering the non-face area may influence the speed and the results of template matching, we eliminate the large black area which is useless in the binary image. Then the final feature image is obtained, as shown in Fig.1 (e).

Fig. 1: An example of preprocessing. (a) Original image. (b) Enhanced image. (c) Clustered image.(d) Binaried image. (e) Structure image.

2.2. Binary template matching

In order to determine the set of rows which contains the eyes, we apply the binary template matching to the structure image in order to search possible eye pairs [5]. The difficulty is that we are not looking for an object with a fixed shape. For this reason we adopt a binary eye pair template which models the two eyes in a very rough way but clearly embody the structure of eye region. A single template has been used for all the images which are of significant different size, thus showing a desirable scale independence property. Eye pairs will be found no matter whether they are open or closed.

Among the positions with the high cross correlation, all the eye pair candidates are extracted. In the proposed method, some in depth rotation of the face depth or rotation on-the-plane of the image are permitted as long as both eyes of the face are visible. Fig. 2 shows some results of eye pair candidates selection.

Fig. 2. Examples of eye pair candidates selection.

In order to detect faces in different scales, the template image is repeatedly scaled by a factor of 1.2. In each scale, all eye pair candidates are extracted and verified by the eyes verifier.

2.3. Eyes Verifier

For the purpose of getting real eye pairs, a simple eyes verifier is applied to all the candidates. We use eyes verifier instead of face verifier proposed in some literatures [6], [7], because the eye pattern is considered reliable enough to confirm the existence of the face.

2.3.1 SVMs

In this paper, we choose the SVM as the classifying function. One distinctive advantage this type of classifier has over traditional neural networks is that SVMs achieve better generalization performance [8].

Support vector machine is a patter classification algorithm developed by V. Vapnik and his team [9]. It is a binary classification method that finds the optimal linear decision surface based on the concept of structural risk minimization. As shown by Vapnik, this maximal margin decision boundary can achieve optimal worst-case generalization performance. Note that SVMs are originally designed to solve problems where data can be separated by a linear decision boundary. By using kernel functions (see Osuna et al. [10] for details), SVMs can be used effectively to deal with problems that are not linearly separable in the original space. Some of the commonly used kernels include Gaussian Radial Basis Functions (RBFs), polynomial functions, and sigmoid polynomials whose decision surfaces are known to have good approximation properties. In this paper, we choose Gaussian radial basis function as the kernel function.

2.3.2 Eye Pair Candidates Verification

All the eye pair candidates (gray level images) are extracted according to the results of binary template matching. Then they are normalized into the size of 25×5 pixels and verified by using SVM to obtain real eye pairs.

The training data used for generating eye verification SVM consists of 400 images of each class (eye-pair and non-eye-pair). Non-eye-pair images are more various than eye pair images, so it is difficult to select standard images. Selection of proper non-eye-pair images is very important to train SVM because performance of SVM is influenced by what kind of non-eye-pair images are used. In the initial stage of training SVM, we use non-eye-pair images similar to eye pair such as eyebrows, nostrils and other eye-pair-like patches as eye pair. And we generate non-eye-pair images using bootstrapping method proposed by Sung and Poggio [11].

After the real eye pairs are obtained, the eye regions and the corresponding face regions will be located according to the position of each eye pair.

3. Eye States Detection

After locating the eyes, the eye states detection is processed in the eye region and the image we use is the structure image which has been described in the preprocessing step. And the eye region image is separated into two parts and detects the states of left eye and right eye respectively.

3.1. Dual Eyes States

There are two default eye states: open and closed. Before detecting the eye states, what open eye is must be defined. If a man can see something, it is considered that his eyes are open. But we cannot take this condition as the computer criterion. Our criterion is that if the iris and the white of eye are visible, the eye is open. Otherwise, the eye is closed. Different states of eye are shown in Fig. 3.

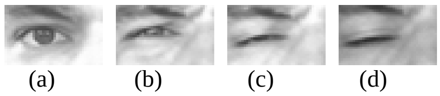

Fig. 3. Dual eye states. (a) and (b) are open eyes. (c) and (d) are close eyes.

It is difficult to detect eye states when eye is nearly closed, as shown in Fig.4(c). The difficulty come from that the eye is actually open, but the iris is not visible. The nearly closed eye is regarded as closed and it is adoptable in many practical applications.

3.2. Eye States Detection by Using Contour Information

The iris can provide important information about the eye state because if the eye is closed the iris will not be visible, but if the eye is open part of the iris will normally be visible. In our method, we use the structure images obtained in preprocessing step to extract outer and inner contour by using chain code tracing. Fig. 4 illustrates the structure images and the contour images getting from the original images in Fig. 3.

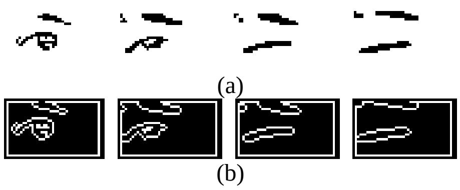

Fig. 4. Eye contour image from different states. (a) Structure image. (b) Contour image.

After obtaining the contour image of eye, we use Hough transform to detect iris circle. Detecting iris is a problem of finding the iris within an eye image. One of possible ways to solve this problem is to move a mask with an appropriate shape and size along the image and look for correlation between the image and the mask [2]. However, detecting iris can be solved successfully by Hough transform. Hough transform is a very powerful technique for curve detection. Exponential growth of the accumulator data structure with the increase of the number of curve parameters restricts its practical usability to curves with few parameters. If the prior information about radius of iris is used, computational complexity can be decreased significantly. In our method, the eye pair candidates are found by using template matching, so the eye size is roughly known according to the template size. And the radius of iris is in proportion to the eye size, for this reason a range of iris radius is set to tackle different iris circle. Then the computational complexity is decreased.

Eye state can be determined by the detection results of iris circle. If the iris circle is detected, the eye is open. Otherwise, the eye is closed. Fig.5 shows some results of detecting iris. The center of iris is marked with red cross and the iris contour is marked with blue circle. We can find from Fig.5 that the wider the eye open, the better the iris contour we can obtain.

Fig. 5. Using the Hough transform on eye contour image.

4. Experimental Results

The proposed method is tested on the BioID face database. The BioID face database consists of 1521 images (384 × 286 pixels, gray level) of 23 different test persons and has been recorded during several sessions at different places. This set features a larger variety of illumination, background and face size. It stresses real world conditions. So it is believed to be more difficult than other dataset containing images with uniform illumination and background.

The eye pair candidates can be selected successfully in most cases, no matter whether face patterns are in different scale, expression, and illumination conditions. So the eye location is successful after eyes verification by using SVM in most case. The eye location rate is 95.6% in BioID database.

Some eye states detection results are shown in Figs. 6. The input images vary greatly in background, scale, expression and illumination, the images also including partial face occlusions and glasses wearing. The iris detected is marked with white cross and the eye is considered to be open. The eye is considered to be closed if the iris circle can't be detected.

Fig. 6. Some eye states detection results.

5. Conclusions and Future Research

In this paper we present an efficient method for locating eyes and detecting eye states in still gray level images with unconstrained background. The structure of eye region is used as a robust cue to find eye pairs. Contour information of the iris is used to detect the eye states. The proposed method can deal with illumination changes, moderate rotations, glasses wearing and partial face occlusions. However, the eye states detection will fail if the reflection of glasses is too strong which leads to lose details of the eye pattern. In our future research, more information will be combined together to detect eye states more efficiently, such as eye corners and eyelids.

6. References

[1] J. Deng and F. Lai. Region-Based Template Deformation and Masking for Eye-feature Extraction and Description. Pattern Recognition. 1997, **30**(3): 403–419.

[2] Y. Tian, T. Kanade, and J. Cohn. Dual-State Parametric Eye Tracking, International Conference on Face and Gesture Recognition. 1999, 110-115.

[3] S. Bernogger, L. Yin, A. Basu and A. Pinz. Eye Tracking and Animation for MPEG-4 Coding. In Proc. IEEE Conf. Pattern Recognition. 1998, **2:** 1281–1284.

[4] H. Liu, Y.W Wu, and H.B Zha. Eye States Detection from Color Facial Image Sequence. *Proceeding of SPIE*. 2002, 4875, 693.

[5] P. Campadelli, R. Lanzarotti. Localization of Facial Features and Fiducial Points. In Processings of the International Conference Visualization, Imaging and Image Processing, Malaga (Spagna). 2002, 491–495.

[6] J.X Wu, Z.H Zhou. Efficient Face Candidates Selector for Face Detection. Pattern Recognition. 2003, **36:** 1175 – 1186.

[7] C.-C. Han, H.-Y.M. Liao, G.-J. Yu, L.-H. Chen. Fast Face Detection via Morphology-based Pre-processing. In Proceedings of the Ninth International Conference on Image Analysis and Processing. 1998, 469–476.

[8] C. A. Waring, X.W Liu. Face Detection Using Spectral Histograms and SVMs. IEEE Trans Systems, Man, and Cybernetics-Part B: C Cybernetics. 2005, **35**(3): 467– 476.

[9] V. Vapnik. The Nature of Statistical Learning Theory. Springer-Verlag: New York. 1995.

[10] E. Osuna, R. Freund, and F. Girosi. Training Support Vector Machines: An Application to Face Detection. In Proc. IEEE Conf. Computer Vision Pattern Recognition. 1997, 130–136.

[11] K.-K. Sung and T. Poggio. Example-based learning for view-based human face detection. IEEE Trans. Pattern Anal. Mach. Intell. 1998. **20**(1): 39–51.