A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles

Kun Peng, Liming Chen,
LIRIS, Département MI, Ecole Centrale de Lyon
36 avenue Guy de Collongue, BP 163, 69131 Ecully, France

Su Ruan,
Equipe Image, CReSTIC, Département GE&II, IUT de Troyes
9 rue de Quebec, 10026 Troyes, France

and

Georgy Kukharev
Faculty of Computer Science and Information Technology, Technical University of Szczecin
Zolnierska 49, 71-210 Szczecin, Poland

perkpeng@hotmail.com

ABSTRACT

This paper presents a robust eye detection algorithm for gray intensity images. The idea of our method is to combine the respective advantages of two existing techniques, feature based method and template based method, and to overcome their shortcomings. Firstly, after the location of face region is detected, a feature based method will be used to detect two rough regions of both eyes on the face. Then an accurate detection of iris centers will be continued by applying a template based method in these two rough regions. Results of experiments to the faces without spectacles show that the proposed approach is not only robust but also quite efficient.

Keywords: Eye detection, Face detection, Face recognition, Image processing, Pattern recognition

1. INTRODUCTION

As one of the salient features of the human face, human eyes play an important role in face recognition and facial expression analysis. In fact, the eyes can be considered salient and relatively stable feature on the face in comparison with other facial features. Therefore, when we detect facial features, it is advantageous to detect eyes before the detection of other facial features. The position of other facial features can be estimated using the eye position [1]. In addition, the size, the location and the image-plane rotation of face in the image can be normalized by only the position of both eyes.

Eye detection is divided into eye position detection [1, 2] and eye contour detection [3, 15, 16]. (The second plays an important role in applications such as video conferencing and vision assisted user interface [2]). However, most algorithms for eye contour detection, which use the deformable template proposed by Yuille et al. [3], require the detection of eye positions to initialize eye templates. Thus, eye position detection is important not only for face recognition and facial expression analysis but also for eye contour detection. In this paper eye detection means eye position detection.

Related work

The existing work in eye position detection can be classified into two categories: active infrared (IR) based approaches and image-based passive approaches. Eye detection based on active remote IR illumination is a simple yet effective approach. But they all rely on an active IR light source to produce the dark or bright pupil effects. In other words, these methods can only be applied to the IR illuminated eye images. It's certain that these methods would not be widely used, because in many real applications the face images are not IR illuminated.

Thus this paper only focuses on the image-based passive methods, which can be broadly classified into three categories: template based methods [3-6], appearance based methods [7-9] and feature based methods [10-14]. In the template based methods, a generic eye model, based on the eye shape, is designed firstly. Template matching is then used to search the image for the eyes. While these methods can detect eyes accurately, they are normally time-consuming.

The appearance based methods [7-9] detect eyes based on their photometric appearance. These methods usually need to collect a large amount of training data, representing the eyes of different subjects, under different face orientations, and under different illumination conditions. These data are used to train a classifier such as a neural network or the support vector machine and detection is achieved via classification.

Feature based methods explore the characteristics (such as edge and intensity of iris, the color distributions of the sclera and the flesh) of the eyes to identify some distinctive features around the eyes. Although these methods are usually efficient, they lack accuracy for the images which have not high contrast. For example, these techniques may mistake eyebrows for eyes.

In summary, the image-based eye detection approaches locate the eyes by exploiting eyes differences in appearance and shape from the rest of the face. The special characteristics of the eye such as dark pupil, white sclera, circular iris, eye corners, eye shape, etc. are utilized to distinguish the human eyes from other objects. However, these approaches lack either efficiency or accuracy, and they are not ideal for some real applications.

Principle of our method

In this paper, we propose a robust algorithm for eye detection on gray intensity face, based on combining the feature base methods and template based approaches. Combining the respective strengths of different complementary techniques and overcoming their shortcomings, the proposed method uses firstly the feature based method to find out broadly the two regions of eyes in a face, and the template based method is then used to locate the center of iris accurately.

The template based approaches are usually time-consuming. Its inefficiency comes from two main factors. Firstly, in order to improve the accuracy, these methods have to match the whole face with an eye template pixel by pixel. Secondly, as we don't know the size of eyes for an input face image, we need to repeat the matching process with eye templates of different sizes. That is to say, we have to perform the template matching several times.

So the solution to improve the efficiency of this algorithm focuses on two points: reducing the area in the face image for template matching and cutting down the times of this type of matching. In fact, our method firstly detects the two rough regions of eyes in the face using a feature based method. Thus the following template matching will be performed only in these two regions which are much smaller than the whole face. In addition, we can evaluate the size of eye template according to the size of these two regions. In other words, profiting from possibility of evaluating the size of eyes, our algorithm performs the template matching just once. Altogether, the proposed method combines the accuracy of template based methods and the efficiency of feature based methods.

Outline of the paper

The remainder of the paper is organized as follows: The details of eyes detection algorithm are described in Section 2. Section 3 is devoted to the experiments on the ORL face database. The discussion and conclusion are given in the last section.

2. PROPOSED METHOD

Architecture

Currently, there are a lot of promising face detection methods [17-19]. This paper therefore assumes that (1) a rough face region has been located or the image consists of only one face, and (2) eyes in face image can be seen.

The architecture of the proposed approach is shown in Fig. 1. When a face image is presented to the system, face detection will be firstly performed to locate the rough face region. The second step, which uses an efficient feature based method, is to locate two rough regions of eyes in the face. In the same time, on the basis of these two regions, the sizes of two eyes will be evaluated, and the templates of eyes will be created according to the estimated sizes. Finally, the precise locations of the two centers of iris will be found out after template matching is applied in these two rough regions.

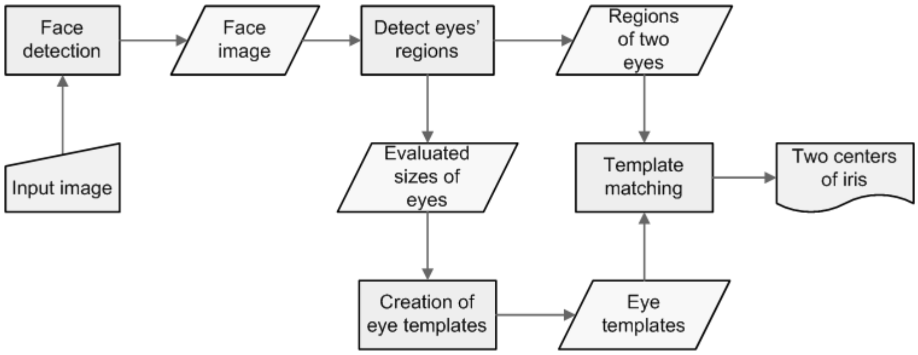

Fig. 1. Flowchart of proposed method

Detection of eyes' regions

When the rough face region is detected, as we have said, an efficient feature based method will be sequentially applied to locate the rough regions of both eyes which will be used to the following affining detection. Fig. 2 shows the processes of the proposed method:

The first step is to calculate the gradient image (b) of the rough face region image (a). Then we apply a horizontal projection to this gradient image. As we know that the eyes locate in the upper part of the face and that the pixels near the eyes are more changeful in value comparing with the other parts of face, it is obvious that the peak of this horizontal projection in the upper part can give us the horizontal position of eyes. According to this horizontal position and the total height of the face, we can easily line out a horizontal region (c) in which the eyes locate.

And then we perform a vertical projection to all pixels in this horizontal region of image (c), and a peak of this projection can be found near the vertical center of face image. In fact, the position of this vertical peak can be treated as the position of vertical center of face (d), because the area between both eyes is most bright in the horizontal region.

In the same time, a vertical projection will be done to the gradient image (b). There are two peaks of projection near the right and left boundary of face image which correspond to right and left limit of the face (e). In addition, from these two vertical limit lines, the width of face can be easily estimated.

Combining all results from (c), (d) and (e), we can get an image segmented like (f). Finally, based on the result of (f) and the estimated width of face, the regions of both eyes can be lined out (g).

Creation of eye templates

After the two rough regions of eyes are detected, template matching will be used to locate the precise positions of iris centers in these regions. Because the matching region reduces from the whole face to the two rough regions, the efficiency of algorithm is well improved.

Obviously, the first obligatory step for a template matching is to create a template. It's easy to find out eye templates which can be obtained from a real face image. But the template can't be directly used for matching, because the size of the eye in the template is not same as that in the input image. A simple solution for this problem is to perform the process of matching several times, and each time we will use the template with different size. But this method is very ineffective.

Concerning our algorithm, in order to improve the efficiency, the size of the eyes will be estimated automatically. Thus the process of matching can be only performed just once. As we have said, the width of face is already estimated (see Fig. 2), and the size of eye template can be easily decided according to the width of face and the geometric structure of human face. The last image (g) in Fig. 2 shows two eye templates (at two top corners) created basing on the estimated eye sizes.

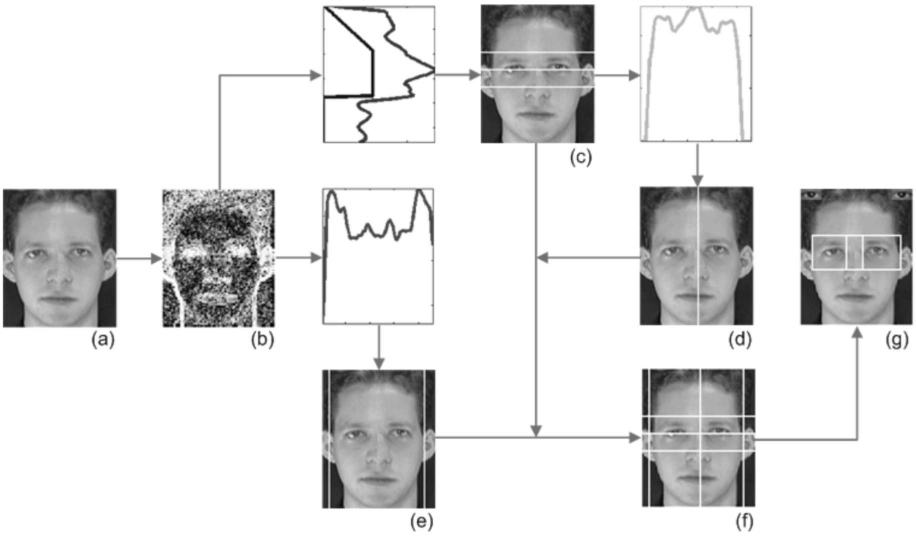

Fig. 2. Detection of eyes' regions

Localization of iris centers

Suppose that we have a template $g[i, j]$ and we wish to detect its instances in an image $f[i, j]$. An obvious thing to do is to place the template at a location in an image and to detect its presence at that point by comparing intensity values in the template with the corresponding values in the image. Since it is rare that intensity values will match exactly, we require a measure of dissimilarity between the intensity values of the template and the corresponding values of the image. Several measures may be defined:

$$\max_{[i,j]\in R} |f-g|,$$

$$\sum_{[i,j]\in R} |f-g|,$$

$$\sum_{[i,j]\in R} (f-g)^2,$$

where R is the region of the template.

The sum of the squared errors is the most popular measure. In the case of template matching, this measure can be computed indirectly and computational cost can be reduced. We can simplify:

$$\sum_{[i,j]\in R} (f-g)^2 = \sum_{[i,j]\in R} f^2 + \sum_{[i,j]\in R} g^2 - 2\sum_{[i,j]\in R} fg \, .$$

Now if we assume that f and g are fixed, then $\sum fg$ gives a measure of mismatch. A reasonable strategy for obtaining all locations and instances of the template is to shift the template and use the match measure at every point in the image. Thus, for an $m \times n$ template, we compute:

$$M[i, j] = \sum_{k=1}^{m}\sum_{l=1}^{n} g[k,l]\, f[i+k, j+l],$$

where k and l are the displacements with respect to the template in the image. This operation is called the *cross-correlation* between f and g.

Our aim will be to find the locations that are local maxima or are above a certain threshold value. However, a minor problem in the above computation was introduced when we assumed that f and g are constant. When applying this computation to images, the template g is constant, but the value of f will be varying. The value of M will then depend on f and hence will not give a correct indication of the match at different locations. This problem can be solved by using normalized cross-correlation. The match measure M then can be computed using:

$$C_{fg}[i, j] = \sum_{k=1}^{m}\sum_{l=1}^{n} g[k,l]\, f[i+k, j+l]$$

$$M[i, j] = \frac{C_{fg}[i, j]}{\left\{\sum_{k=1}^{m}\sum_{l=1}^{n} f^2[i+k, j+l]\right\}^{1/2}} \, .$$

Fig. 3 shows an example of using template matching to locate the iris centers. In left side, the first line displays two templates of eyes which are created according to the sizes estimated. The images in the second line are the rough eyes' regions in which the template matching will be applied. The result of template matching using cross-correlation is shown by the images of the third line. And the image in left side shows the final result of eye position detection.

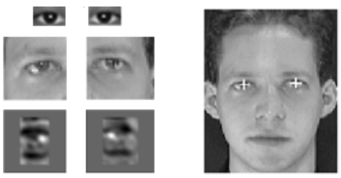

Fig. 3. Template matching

3. EXPERIMENTAL RESULTS

In this section, we present the experimental results of our algorithms. We use the images in the ORL database, a well-known free database of faces, to do our experiments. In this database, there are completely photographs of 40 persons, of which each one has 10 various views. The 10 views of the same person include faces looking to the right, to the left, downward and upward (see the first line of Fig. 4). All faces in this database are presented by images in gray-level with the size of 92×112.

We made experiments using all faces without spectacles which concerns 227 face images and 29 persons. The success rate of proposed algorithm for all 227 faces is 95.2%. Fig. 4 shows examples of the images for which the proposed algorithm could correctly detect the irises of both eyes. In the first line, there are five face views of the same person. And the images in the second line are faces of five different persons.

Fig. 4. Examples for faces without spectacles

The execution time of the proposed algorithm is about 0.982 second on average by a PC whose CPU is Pentium IV, 1.8 GHz. It's remarkable that this execution time is reckoned for a program written in Matlab. Obviously, the execution time would be reduced a lot if the program is transplanted from Matlab to C or C++.

We also made experiments to the faces with spectacles, but the results are very unsatisfied. Fig. 5 shows some examples of these experiments. The images in the first line are five examples for faces with spectacles for which the proposed algorithm could correctly detect the iris centers of both eyes. And the second line gives five faces for which the proposed algorithm failed the correct detection of the iris centers. After comparing and analyzing the detection results, we found out that the false detection is mainly due to the reflection of the spectacles. That is to say, the reflection of the spectacles changes the intensity values of pixels around eyes, which leads to a false template matching.

Fig. 5. Examples for faces with spectacles

4. CONCLUSION

A robust eye detection method for gray intensity faces is reported in this paper. The proposed algorithm combines two existing techniques: feature base method and template based method. The proposed algorithm firstly makes use of feature based methods to detect two rough regions of eye. The precise locations of iris centers are then detected by performing template matching in these two regions.

The proposed method has been tested by images from ORL face database. Experimental results show that this method works well with the faces without spectacles. For 227 faces without spectacles, the detection accuracy is 95.2%. In addition, the average execution time of proposed algorithm shows that this approach is also quite efficient.

However, the proposed method doesn't work so well for the faces with spectacles. Experimental results show that the false detection is mainly due to the reflection of spectacle. In view of above limitation, the future work will be concentrated on improving the detection accuracy for the faces with spectacles by reducing the effect of reflection.

REFERENCES

[1] F. R. Brunelli, T. Poggio, Face recognition: features versus templates, IEEE Trans. Pattern Anal. Mach. Intell. 15 (10) (1993) 1042–1052.

[2] Y. Tian, T. Kanade, J.F. Cohn, Recognizing upper face action units for facial expression analysis, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, 2000, Vol. 1, pp. 294–301.

[3] A.L. Yuille, P.W. Hallinan, D.S. Cohen, Feature extraction from faces using deformable templates, Int. J. Comput. Vision 8 (2) (1992) 99–111.

[4] X. Xie, R. Sudhakar, H. Zhuang, On improving eye feature extraction using deformable templates, Pattern Recognit. 27 (1994) 791–799.

[5] K.M. Lam, H. Yan, Locating and extracting the eye in human face images, Pattern Recognit. 29 (1996) 771–779.

[6] M. Nixon, Eye spacing measurement for facial recognition, in: Proc. of the Society of Photo-Optical Instrument Engineers, 1985.

[7] A. Pentland, B. Moghaddam, T. Starner, View-based and modular eigenspaces for face recognition, in: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94), Seattle, WA, 1994.

[8] W. min Huang and R. Mariani, Face detection and precise eyes location, in: Proc. Int. Conf. on Pattern Recognition (ICPR'00), 2000.

[9] J. Huang, H. Wechsler, Eye detection using optimal wavelet packets and radial basis functions (rbfs), Int. J. Pattern Recognit. Artif. Intell. 13 (7) (1999) 1009–1025.

[10] G.C. Feng, P.C. Yuen, Variance projection function and its application to eye detection for human face recognition, Int. J. Comput. Vis. 19 (1998) 899–906.

[11] G.C. Feng, P.C. Yuen, Multi-cues eye detection on gray intensity image, Pattern Recognit. 34 (2001) 1033–1046.

[12] S. Kawato, J. Ohya, Real-time detection of nodding and head-shaking by directly detecting and tracking the between-eyes, in: Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2000, pp. 40–45.

[13] Y. Tian, T. Kanade, J.F. Cohn, Dual-state parametric eye tracking, in: Proc. 4th IEEE Int. Conf. On Automatic Face and Gesture Recognition, 2000.

[14] S.A. Sirohey, A. Rosenfeld, Eye detection in a face image using linear and nonlinear filters, Pattern Recognit. 34 (2001) 1367–1391.

[15] G. Chow, X. Li, Toward a system for automatic facial feature detection, Pattern Recognition 26 (1993) 1739–1755.

[16] K.M. Lam, H. Yan, Locating and extracting the eye in human face images, Pattern Recognition 29 (5) (1996) 771–779.

[17] H.A. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1) (1998) 23–38.

[18] K.K. Sung, T. Poggio, Example-based learning for viewbase human face detection, IEEE Trans, Pattern Anal. Mach. Intell. 20 (1) (1998) 39–51.

[19] P.C. Yuen, G.C. Feng, J.P. Zhou, A contour detection method: initialization and contour model, Pattern Recognition Lett. 20 (2) (1999) 141–148.

Received: Mar. 2005. Accepted: May 2005.