EYE EXTRACTION USING SPATIAL FUZZY CLUSTERING METHOD

K.W. Tse*, W.H. Lau*, SH Leung[†] and Alan WC Liew*

Department of Computer Engineering & Information Technology*

Department of Electronic Engineering[†]

City University of Hong Kong

Tat Chee Avenue, Kowloon, Hong Kong, China

ABSTRACT

Facial features extraction and tracking are crucial steps for many multimedia communication applications such as automated visual interpretation, human face recognition, and development of high quality model-based coding (e.g. MPEG-4) systems. Among different facial features, eyes and lip play an important role in either recognition process or teleconferencing applications. However, automatic human facial features detection is a difficult task due to different scale, rotation and translation of the features.

This paper describes a computational approach for locating the eye position from a given eye window. The proposed algorithm will extract the eye in two stages. The iris will first be extracted using a region-based energy minimization approach from the membership map generated by the spatial fuzzy clustering technique. In second two, control points will be detected to locate the sclera and two parabolas will be used to model the upper and lower eyelids. Satisfactory results have been achieved with the proposed method.

1. INTRODUCTION

A variety of approaches have been proposed for eye extraction and mainly they can be divided in two streams. One stream is to first locate the face region from a cluttered image and then to extract the facial features within the resulted face region. An example-based approach for locating the vertical frontal views of human face in complex scenes has been developed by Sung and Poggio [1]. However, this approach cannot locate the face if its orientation deviates from vertical frontal view. Sobettka et al. [2] detected human facial features for all races by using the color information. However, it only focused on the frontal view face images. Since the sizes of the facial features are relatively small comparing with the entire image, it will be of advantage to limit the search region by identifying the face first. However, the accuracy of the facial features extraction is highly sensitive to the results of the face detection.

Another approach is to extract the eye from an input eye window image. This means that a rough eye region is required for the system to extract the precise location of the eye, therefore, the performance of this kind of systems is generally good. Yuille et al. [3] proposed to use deformable templates composed of parametric curves to describe the eye location and shape. However, this method is computational intensive and hard to be used as an eye tracker for image sequences and it will provide unexpected result with improper selection of weighting factors. The idea of using the eye corners is introduced by Lam et al. [4] to help in template matching. However, quality edge image for template matching may not be obtained if the contrast of eye region is relatively low. Takacs et al. [5] developed iconic filter banks to locate facial landmarks. However, an object model, such as eye and mouth, has to be trained before detection.

In this paper, we will describe an eye extraction technique using spatial fuzzy clustering technique and parametric model. This method assumes that a rough region containing the eye has been already identified.

2. OUTLINE OF THE ALGORITHM

2.1 Spatial Fuzzy Clustering

Eye is an important feature of a facial image and is the most difficult region to extract due to variation in luminance, viewing angles and face orientation. Eye extraction can be viewed as a process of separating the homogenous color regions from a given image. The spatial fuzzy clustering technique [6] has been shown to be an effective way to group homogenous regions according to the pixels' class memberships by considering the influences of neighborhood pixels. Experimental results have shown that regions can be extracted reliably by analyzing the color image in the CIE LAB domain. Hence all the RGB images are transformed to $\{L, a, b\}$ representation for clustering analysis.

The extraction process involves two stages and are described in the following:

2.2 Stage one processing

1. Segment an eye window image into 4 clusters with the fuzzy clustering technique.
2. Determine the cluster belonging to iris from the value of luminance.
3. Use the iris template to locate the eye.

2.3 Stage two processing

4. Extract the control points for the sclera.
5. Use parabolas to fit the upper and lower eyelids.

Detailed operations of the algorithm are described in the following section.

3. EYE EXTRACTION

3.1 Iris Segmentation and Modeling

In the first stage of processing, it is natural to segment the image into 4 regions, one for iris, one for sclera and other for skin region and frame of glasses. The spatial fuzzy clustering technique [6] is used to analyze the eye image and all pixels are assigned with membership values ranged from 0 to 1. Finally, 4 clusters will be generated according to the distribution of the membership values. Fig. 1 shows the clustering results of two images, one with glasses and one without. It can be seen that the segmentation is of good quality and the iris can clearly be seen.

Fig. 1. Clustering results of images for a rough eye region

Since the luminance of the iris is lower than that of the sclera and skin, the eye cluster (C_{eye}) can then be identified by comparing the luminance value of each cluster.

$$C_{eye} = C_i \quad \text{if } L_i < L_j \quad \forall j \leq c \text{ and } j \neq i$$

where L_i is value of luminance component of i th cluster.

Since both the eyebrow and iris are usually of low luminance value, they will appear in the same cluster as shown in Fig. 2.

Fig. 2. Binary images contain the iris and eyebrow.

An iris template with three parameters (r, x_c, y_c), as shown in Fig. 3, is used to describe the boundary between the sclera and the iris.

Fig. 3. The iris Template

The three parameters can be found by maximizing the cost function $F(r, x_c, y_c)$ using the downhill simplex method.

$$F(r, x_c, y_c) = \frac{r_{iris}}{\text{area of iris}} + \frac{r_{background}}{\text{area of background}}$$

where r_{iris} and $r_{background}$ represent the sum of the membership values within the iris area and the background area respectively.

The estimated position and size of the iris is shown in Fig. 4.

Fig. 4. Results of the iris template matching

3.2 Locating the Control points

Having located the iris, we will carry out the stage two processing. The rough eye region is first reduced to avoid unnecessary computations and a refined area is defined as shown in Fig. 5. The spatial fuzzy clustering analysis is then applied to the refined area and the results are given in Fig. 6. The sclera can clearly be seen from the clustering results.

Since the sclera is always brighter than the iris or even the skin, the cluster corresponds to the sclera (C_{sclera}) can be located according to the following:

$$C_{sclera} = C_i \quad \text{if } L_i > L_j \quad \forall j \leq c \text{ and } j \neq i$$

where L_i is value of luminance component of i th cluster.

Having determined the cluster containing the sclera, the image is divided into 2 parts as shown in Fig. 7.

Finally we need to isolate the sclera from the skin region, the variances of each white region are calculated. The region with the minimum variance corresponds to the region of sclera (R_{sclera}).

Fig. 5. The refined area

Fig. 6. The clustering results of the refined area

Fig. 7. (a) Binary images contain the sclera; (b) the binary images is divided into two parts

$$R_{sclera} = R_i \text{ if } \text{var}(R_i) < \text{var}(R_j) \quad \forall j \leq c \text{ and } j \neq i$$

where

$$\text{var}(R_i) = E(p_i^2) - (E(p_i))^2$$

$$= \frac{\sum (x_i - \bar{x}_i)^2 + \sum (y_i - \bar{y}_i)^2}{N_i}$$

with $p_i = (x_i, y_i)$ denotes the location of pixel belong to R_i , N_i represents the number of pixels belong to portion R_i .

The sclera identified is shown in Fig. 8.

Fig. 8. Results of sclera extraction

Having located the sclera, three control points are required to be estimated for modeling the eye. They are the eye corners, the intersection points between the iris and the upper eyelid and the intersection points between the iris and the lower eyelid as shown in Fig. 9. However, by the examining the extracted sclera as shown in Fig. 8, it is observed that we can only estimate one corner point

reliably. For the other eye corner, it can be estimated from Fig. 2 as shown in Fig. 10.

Fig. 9. The locations of the control points

Fig. 10. Localization of left corner point

Fig. 11 shows the locations of the control points identified from the sclera and Fig. 12 shows the results overlaid to the original eye image.

Fig. 11. Results of identified control points

Fig. 12. Control points overlaid on the original images

3.3 Eye Shape Modeling

Once the control points are fixed, the shape of the eye can then be fitted with two parabolas, one representing the upper eyelid and one for the lower eyelid.

The parabola used to define the eye shape is given by:

$$y = ax^2 + bx + c$$

Assuming that the three control points are labeled by (x_1, y_1) , (x_2, y_2) and (x_3, y_3) , we can write

$$y_1 = ax_1^2 + bx_1 + c$$

$$y_2 = ax_2^2 + bx_2 + c$$

$$y_3 = ax_3^2 + bx_3 + c$$

The solution of (a, b, c) can be obtained from the following equations:

$$a = \frac{y_1}{(x_1 - x_2)(x_1 - x_3)} + \frac{y_2}{(x_2 - x_1)(x_2 - x_3)} + \frac{y_3}{(x_3 - x_1)(x_3 - x_2)}$$

$$b = \frac{-y_1(x_2 + x_3)}{(x_1 - x_2)(x_1 - x_3)} + \frac{-y_2(x_1 + x_3)}{(x_2 - x_1)(x_2 - x_3)} + \frac{-y_3(x_1 + x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

$$c = \frac{y_1 x_2 x_3}{(x_1 - x_2)(x_1 - x_3)} + \frac{y_2 x_1 x_3}{(x_2 - x_1)(x_2 - x_3)} + \frac{y_3 x_1 x_2}{(x_3 - x_1)(x_3 - x_2)}$$

It should be noted that the above formulation only involves three control points, however, the actually modeling of the eye shape involves four control points for each parabola as shown in Fig. 9. To find the upper eyelid, we will find two sets of parameters, (a_1, b_1, c_1) is obtained by considering points (P_1, P_2, P_4) , and (a_2, b_2, c_2) is obtained by considering points (P_1, P_3, P_4) . Finally, the parabola parameters (a, b, c) for the upper eyelid is obtained by simply taking the average of the two sets of parameters, i.e.

$$a = \frac{a_1 + a_2}{2}, b = \frac{b_1 + b_2}{2}, c = \frac{c_1 + c_2}{2}$$

Similarly, the lower eyelid can be obtained using the same approach. Fig. 13 shows the extraction results by combining the iris and the upper and lower eyelids.

Fig. 13. The eye extraction results

4. PERFORMANCE

The proposed algorithm has been used to analysis the database from the Purdue Robot Vision Laboratory [7]. 185 images representing 134 individuals have been analyzed. A successful rate of 95.68% for iris extraction and 74.49% for eyelid fitting have been achieved.

It has been observed that the iris and hair will be classified into the same cluster if part of the eye is covered by hair. The extraction of iris will be affected and the size of the

iris will usually be large than the actual size. There are two reasons accounting for the low successful rate for the eyelid fitting. Naturally, a poor iris extraction cannot provide a good eyelid fitting since the identification of control points becomes difficult. On the other hand, the reflection of light on the glasses may also cause undesirable clustering results which result in incorrect controls points identification. In order to improve the performance, we are currently investigating the image quality by overlaying the sclera to the image given in Fig. 2. It is shown that a better eye region can be achieved and improved extraction results are expected in near future.

5. CONCLUSION

An automatic method of eye extraction has been developed. The proposed method, which makes use of deformable template and spatial fuzzy clustering technique to extract and describe the detailed eye features from an color image, can be used as a preprocessing stage of a practical facial feature recognition system or other multimedia applications.

6. REFERENCE

- [1] Sung, K. K. and Poggio, T., "Example-based learning for view-based human face detection", *IEEE Trans. on PAMI*, Vol. 20, No. 1, pp. 39-51, 1998
- [2] Sobottka, K. and Pitas, I. "Extraction of facial regions and features using colour and shape information", *Proceeding of ICIP*, Vol. 3, pp. 483-486, 1996
- [3] Yuille, A. L., Cohen, D. S. and Hallinan, P. W., "Feature extraction from faces using deformable templates", *Proceeding of CVPR '89*, pp. 104-109, 1989
- [4] Lam, K.M. and Yan, H., "Locating and extracting the eye in human face images", *Pattern Recognition*, Vol. 29, No. 5, pp. 771-779, 1996
- [5] Takacs, B. and Wechsler, H. "Detection of faces and facial landmarks using iconic filter banks", *Pattern recognition*, Vol. 30, No. 10, pp. 1623-1636, 1997
- [6] A.W.C. Liew, S.H. Leung and W.H. Lau, "Fuzzy Image Clustering Incorporating Spatial Continuity," *IEE Proc. Visual Image Signal Processing*, Vol. 147, No. 2, pp. 185-192, 2000
- [7] http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html