Chapter 3: Facial Feature Localization

3.1 INTRODUCTION

The eyes can be considered salient and relatively stable features on the human 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 positions of other facial features can be estimated using the eye positions. In addition, the size, location and image-plane rotation of the face in the image can be normalized by only the positions of both eyes. Eye detection plays an important role in applications such as video conferencing and vision assisted user interfaces.

Most of iris detection algorithms previously reported use template matching, eigenspace method, or Hough transform. However, template matching and eigenspace method require the normalization of the image face in its size and orientation. And, these algorithms can detect the irises only from faces whose eye patterns are similar to sample eye images used as eye models. In addition, iris detection algorithms using Hough transform need to estimate the searching windows for the irises in the face region and so they require complete face region detection.

In this chapter we propose a new algorithm to extract the irises of both eyes from a face image [35]. The algorithm first detects the face region in the image and then extracts intensity valleys from the face region. Next, the algorithm extracts iris candidates from the valleys using the feature template of Lin and Wu [14] and the separability filter of Fukui *et al* [23]. Finally, using a cost of each iris candidate-pair, the algorithm selects a pair of iris candidates corresponding to the irises.

The iris detection algorithm proposed in this chapter only requires a rough estimation of size and orientation of the face in the image. And, although the proposed algorithm uses an eye template cut off from a face image, the performance of the proposed algorithm is not so sensitive to the variation of the template. In addition, the algorithm does not require complete face region detection. For example, the face regions detected by the edge-based face detection algorithm include the hair as shown in Fig.12. The detected face

regions may contain the necks or may lose the mouths. The proposed iris detection algorithm can correctly detect the irises even from such incomplete face regions.

3.2 ALGORITHMS OF LIN, FUKUI AND CHOW FOR FACIAL FEATURE LOCALIZATION

In this section we show three eye detection algorithms closely connected to the proposed iris detection algorithm, which are the ones proposed by Lin and Wu [14] and Fukui *et al.* [23] and Chow *et al.* [28], respectively.

The algorithm of Lin and Wu [14] first extracts the face region from the input image using a region-growing method. Next, for each pixel (x,y) in the face region, the algorithm computes a cost C(x, y) by,

$$C(x,y) = C1(x,y) + C2(x,y) \qquad (3.1)$$

Let S(x,y) denote the square region with center (x,y) and side-length d. Then, C1(x,y) is given by,

$$C_1(x,y) = \sum_{j=y-d/2}^{y+d/2} V_r(j) + \sum_{i=x-d/2}^{x+d/2} V_c(i) \qquad (3.2)$$

Vr(j) and Vc(i) are the mean crossing numbers of row j and column i which are defined as follows. Let I(i,j) denote the intensity values of pixels (i,j) in the image. Then, for each row j, Vr(j) represents the number of pixels (i,j), $x-d/2 \le i \le x+d/2$, such that one of the $I(i-1,j)$ and I(i,j) is greater than $\mu+K$ and the other is smaller than μ minus K where μ is the average intensity of pixels (i,j), $x-d/2 \le i \le x+d/2$ and K is a constant. Vc(i) denotes the number of pixels (i,j), $y-d/2 \le j \le y+d/2$, such that one of the $I(i,j-1)$ and I(i,j) is greater than $\mu+K$ and the other is smaller than (μ-K. C2(x,y)) is a function which evaluates the intensity difference between the central part and the boundary parts of S(x,y). After partitioning the face region into the left eye subregion, the right eye subregion and the mouth subregion, the algorithm selects three pixels with the largest costs, one from each subregion, and determines these pixels to be the positions of the left eye, the right eye and the mouth.