CHAPTER II

Related Works

Face and facial feature detection has received considerable attention because of its wide range in many applications such as human computer interaction, video surveillance, and face recognition [8–12]. Due to the requirements in many applications, the detection approach must handle variations of face images which may occur in practice. In this chapter, existing face and facial feature detection methods are reviewed. The methods are categorized into five categories; some methods clearly overlap category boundaries.

1. **Geometrical information-based methods.** These methods utilize geometrical information [13–19]. The methods are developed based on the rules derived from the researcher's knowledge of human faces . It is easy to come up with simple rules to describe the features of a face and their relationships. Each feature is demonstrated as a geometrical shape: for instance, the shape of the eye ball is a circle and the eyelid is an ellipse. A standard patterns of face and facial features, which are usually in frontal view, are manually predefined or parameterized by a function. Given an input image, the correlation values with the standard patterns are computed for the face boundary, eyes, nose, and mouth independently. The existence of face and facial features is determined based on the correlation values. The advantage of the methods is being simple to implement. This approach can accurately detect face and facial features, but cannot handle large variations of the face images because they are very sensitive to some factors, which affect the

appearance of image, such as facial expression, occlusion, lighting condition, and noise. One more problem with this approach is the difficulty in translating human knowledge into well-defined rules. If the rules are strict, they may fail to detect faces that do not pass all the rules. If the rules are too general, they may give many false positives. Moreover, it is difficult to extend this approach to detect faces in different poses since it is challenging to enumerate all possible cases.

2. **Color information-based methods.** In recent years, human skin color has been used to be an effective feature in many applications from face and facial feature detection to hand tracking. Although different people have different skin color, several researches have depicted that the major difference lies largely between their intensity rather than their chrominance. Several color spaces have been utilized to label pixels as skin including RGB, HSV, YCrCb, and YIQ. Many methods [20–28] have been proposed to build a skin color model for face and facial feature detection. The simplest method is to define a region of skin tone pixels using Cr, Cb values [29] from samples of skin color pixels. With carefully chosen thresholds, $[Cr_1, Cr_2]$ and $[Cb_1, Cb_2]$, a pixel is classified to have skin tone if its values (Cr, Cb) fall within the ranges, i.e., $Cr_1 \leq Cr \leq Cr_2$ and $Cb_1 \leq Cb \leq Cb_2$. In this category, the methods face difficulties in robustly detecting skin colors in the presence of complex background and different lighting conditions. Moreover, these algorithms are applicable only for color images.

3. **Appearance-based methods.** Contrasted to the Geometrical information-based methods where templates are predefined by experts, the template in appearance-based methods [30–45] are learned from a set of training images. Typically, appearance-based methods use gray value (intensity) as the most important parameter for the detection and rely on techniques from statistical analysis and

machine learning to find the relevant characteristics of face and non-face images. The learned characteristics are in the form of distribution models or discriminant functions that are consequently used for face and facial feature detection. Meanwhile, dimensionality reduction is usually carried out for the sake of computation efficiency and detection efficacy. These methods yield high detection rate, but they are not able to perfectly detect face images with poor intensity, some occlusions, and unnatural intensity.

4. **Motion-based methods.** The locations of face and facial features are continuously estimated from the image sequence in real time [46–54]. The difference between two adjacent frames in image sequence is an auxiliary parameter for the detection algorithms. Using such methods, face and facial features cannot be detected using only a single still image from one view.

5. **Edge-based methods.** These methods [55, 56] detect faces are detected without using information on intensity and motion. The edge information is used as input parameter. The methods can handle large variations of the face image. However, facial feature detection is ignored in these methods.

It is obvious that no method can be applied to all intensity types, poses, image conditions, invariant rotation, and face appearances.

To avoid a limitation of intensity usage in this dissertation, position and face shape information are used as inputs. Processing is divided into two steps: face detection and facial feature localization, that are carried out by constructing three detectors, each of them in charge of one specific view (frontal, left, right). In the first step, the faces are detected from an original image. Canny edge detection [57] is applied to find the edge of the image. A candidate face region can be found from the region having the number of pixels corresponding to average face template. Then, the matching value is calculated and applied to find the actual face. Second, facial feature detection is applied to the face obtained in the previous step. A proposed neural visual model (NVM) [61,62,66] is used to recognize all possibilities of facial feature locations. The input parameters are obtained from the face characteristics and the locations of facial features which are independent of the intensity information. Finally, to improve the results, image dilation [67] is applied to remove some irrelevant feature regions. Additionally, the algorithms can be extended to cover rotational invariance problem by using Radon transform [68] to extract the main angle of the face.