Chapter 2

DETECTING FACES IN STILL IMAGES

Images containing faces are essential to intelligent vision-based human computer interaction, and research efforts in face processing include face recognition, face tracking, pose estimation, and facial expression recognition. However, many reported methods assume that the faces in an image or an image sequence have been identified and localized. To build fully automated systems that analyze the information contained in face images, robust and efficient face detection algorithms are required. Given a single image, the goal of face detection is to identify all image regions which contain a face regardless of its three-dimensional position, orientation, and the lighting conditions. Such a problem is challenging because faces are non-rigid and have a high degree of variability in size, shape, color, and texture. Numerous techniques have been developed to detect faces in a single image, and the purpose of this chapter is to categorize and evaluate these algorithms. We also discuss relevant issues such as data collection, evaluation metrics, and benchmarking. After analyzing these algorithms and identifying their limitations, we conclude with several promising directions for future research.

1. INTRODUCTION

With the ubiquity of new information technology and media, more effective and friendly methods for human computer interaction (HCI) are being developed which do not rely on traditional devices such as keyboards, mice, and displays. Furthermore, the ever decreasing price/performance ratio of computing coupled with recent decreases in video image acquisition cost imply that computer vision systems can be deployed in desktop and embedded systems [152] [153] [154]. The rapidly expanding research in face processing is based on the premise that information about a user's identity, state, and intent can be extracted from images, and that computers can then react accordingly, e.g., by

M.-H. Yang et al., *Face Detection and Gesture Recognition for Human-Computer Interaction*

observing a person's facial expression. In the last five years, face and facial expression recognition have attracted much attention though they have been studied for more than twenty years by psychophysicists, neuroscientists, and engineers. Many research demonstrations and commercial applications have been developed from these efforts. A first step of any face processing system is detecting the locations in images where faces are present. However, face detection from a single image is a challenging task because of variability in scale, location, orientation (up-right, rotated), and pose (frontal, profile). Facial expression, occlusion, and lighting conditions also change the overall appearance of faces.

We now give a definition of *face detection*: Given an arbitrary image, the goal of face detection is to determine whether or not there are any faces in the image, and if present, return the image location and extent of each face. The challenges associated with face detection can be attributed to the following factors:

- Pose: The images of a face vary due to the relative camera-face pose (frontal, 45 degree, profile, upside down), and some facial features such as an eye or the nose may become partially or wholly occluded.
- Presence or absence of structural components: Facial features such as beards, mustaches, and glasses may or may not be present, and there is a great deal of variability amongst these components including shape, color, and size.
- Facial expression: The appearance of faces are directly affected by a person's facial expression.
- Occlusion: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other faces.
- Image orientation: Face images directly vary for different rotations about the camera's optical axis.
- Imaging conditions: When the image is formed, factors such as lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, lenses) affect the appearance of a face.

There are many closely related problems to face detection. *Face localization* aims to determine the image position of a single face; this is a simplified detection problem with the assumption that an input image contains only one face [119] [142]. The goal of *facial feature detection* is to detect the presence and location of features such as eyes, nose, nostrils, eyebrow, mouth, lips, ears, etc. with the assumption that there is only one face in an image [39] [74]. *Face recognition* or *face identification* compares an input image (probe) against a