Chapter VIII
Human Motion Analysis

Human motion analysis (Moeslund et. al., 2006; Wang et. al., 2003) is currently one of the most active research areas in computer vision due both to the number of potential applications and its inherent complexity. This high interest is driven by many applications in many areas such as surveillance, virtual reality, perceptual, control applications or analysis of human behaviors. However, the research area contains a number of difficult, such as ill-posed problem. So, many researchers have investigated these problems. Human motion analysis is generally composed of three major parts: human detection, tracking and the behavior understandings.

Human detection finds the segmentation regions corresponding to human and people in an image sequence. It is a very important part in human motion analysis because it greatly affects the subsequent processes such as the tracking and behavior understandings. There are many approaches for finding the human including: background subtraction (Kristensen et. al., 2006; Piccardi, 2004), motion-based segmentation (Sidenbladh, 2004; Gonzalez et. al., 2003; Sangi et. al., 2001), appearance-based segmentation (Mohan et. al., 2001; Utsumi and Tetsutani, 2002; Kang et. al., 2005), shaped-based segmentation (Zhao and Thorpe, 2002; Haritaoglu et. al., 1998), depth-based segmentation (Haritaoglu et. al., 2002; Hayashi et. al., 2004; Li et. al., 2004). Currently many researchers focuse on segmenting human or people in only a still image(Tuzel et. al., 2008, Navneet et.al., 2005, Zhu et. al., 2006). It is quite a difficult problem because we do not know prior information such as motion cues and background cues. The key framework of their method is that a certain feature extractor such as covariance descriptor, histogram of gradient is used for representing each human and then train classifier such as SVM or applying boosting method. And efficient detection method with variously scaled image is crucial to detect humans.

Human tracking involves matching objects between consecutive frames with respect to position, velocity, shape, texture and color using features such as points, lines or blobs.
The tracking methods in essence assume temporal relations among continuous frames. This issue is also important as a means for the pose estimation and gesture recognition. In most of the approaches for tracking, some mathematical techniques are used: Kalman filter (Kalman, 1960; Welch and Bishop, 2001), Condensation algorithm (Isard and Blake, 1998) or Particle filter and Dynamic Bayesian Network (Pavlovic, 1999). Even though these algorithms give optimal estimation, there may be no way to track a human or humans in the case of partially occluded situations. In order to solve this problem, many researchers currently study tracking a human or humans in partial occlusion. Part based human representation with edgelet (Wu and Nevatia, 2005) give a chance to avoid partial occlusion, because we do not track the whole body of a human but only a limb such as an arm, head, or leg. The part based representation is not a state of the art idea: however, a combination of both part detector and priorily known articulated human structure can solve the problem. By using these ideas, we can boost up not only detection but also tracking performance.

Understanding human behavior is to analyze and recognize human motion patterns. In this area, there are many investigations reflected by a large number of different ideas and approaches. The approaches relay on the goal of applications or researchers. But generally, because the behavior understanding problems handles the time-varying data, the general methods for handling time-varying data are proposed: dynamic time warping (Myers and Rabinier, 1980; Bobick and Wilson, 1995; Takahashi and Seki, 1994), hidden Markov models (Poritz, 1988; Brand et. al., 1997; Galata et. al., 2001), neural network (Guo et. al., 1994; Rosenblum et. al., 1994).

This chapter is organized into two parts. The first part reivews some preliminary background, which are the scale adaptive filters, the self-organizing feature map (SOM), and the iterative closest point (ICP) algorithm. The scale adaptive filters (SAF) are specially designed 3D shape filters for human detection using disparity map data. Because the disparity map contains geometric range data, applying the SAF to the disparity map is more robust than applying similar shape filters to image data. Self organizing maps (SOM), also known as Kohonen feature maps or topology-preserving maps, uses a competition-based network paradigm for data clustering (Jang et. al., 1997; Kohonen, 1990; Fausett, 1999). The ICP is a kind of registration algorithm that finds an optimal tranformation that maps given two sets of point data.

The second part introduces five recently developed human motion analysis algorithms. The first algorithm is a pose-robust human detection algorithm based on the SAFs. The second is a hand gesture recognition algorithm and its application to smart home environment systems, where the electronic devices in the smart home are controlled by hand gestures. The third is another example of the hand gesture recognition algorithm that recognizes musical conducting actions. The fourth is a human body gesture recognition algorithm using multiple cameras and silhouette analysis technique. The fifth is another approach to human body gesture recognition using a 3D articulated human body model.

8.1 SCALE ADAPTIVE FILTERS

Human detection is an essential task for many applications such as human robot interaction, video surveillance, human motion tracking, gesture recognition, human behavior analysis, etc. Among many applications, we are interested in the human detection in the field of hu-
Related Content

Curve Fitting Methods: A Survey
[www.igi-global.com/article/curve-fitting-methods/180665?camid=4v1a](www.igi-global.com/article/curve-fitting-methods/180665?camid=4v1a)

Human Motion Tracking in Video: A Practical Approach
[www.igi-global.com/chapter/human-motion-tracking-video/39335?camid=4v1a](www.igi-global.com/chapter/human-motion-tracking-video/39335?camid=4v1a)

QuEST for Information Fusion in Multimedia Reports
[www.igi-global.com/article/quest-for-information-fusion-in-multimedia-reports/130618?camid=4v1a](www.igi-global.com/article/quest-for-information-fusion-in-multimedia-reports/130618?camid=4v1a)
Investigation of Human Monitoring Capabilities for Multiple Watch Windows
www.igi-global.com/article/investigation-of-human-monitoring-capabilities-for-multiple-watch-windows/177209?camid=4v1a