Abstract

One interesting field of research that has gained much attention in recent times is Gesture Recognition. In the context of human-computer non-verbal communication, the visual interface becomes important in establishing communication via understanding human intention from their behaviour, such as facial expressions, hand gestures, etc. With the increasing interest in human-computer interaction (HCI), there has been rapid growth in studies related to vision-based gesture recognition in recent years. Hand gesture recognition from visual images finds applications in areas like human-computer interaction, machine vision, sign language, virtual reality, augmented reality, and so on.

Gesture recognition may be accomplished either by capturing gestural motion using sensor devices or by analyzing gesture images/videos using computer vision techniques. The accurate detection of gestures is significantly affected by physical movement, variations in illumination and shadows, presence of skin-like colors in background, occlusion, background complexity, and different other factors. The complex articulated shape of the hand makes it harder to model the hand appearance for both static and dynamic gestures. Variation of gesture parameters due to spatio-temporal variance in dynamic gesture makes the recognition process more difficult. Different classifiers with the existing features used for vision-based gesture recognition may not be capable of simultaneously handling all the gesture classification problems. Each one has some drawbacks limiting the overall performance. Various researchers have developed different methods to overcome such problems. However, there are still many shortcomings with the algorithms developed so far. This has motivated the candidate to carry out research in this field.

In ***Chapter 1***, along with the introduction various challenges faced by the computer vision community in recognizing hand gestures are mentioned. These challenges include segmentation-related problems, tracking related problems and constraints with extracted features etc. Segmentation-related problems arise due to the presence of skin-like colors in the background, illumination variation, background complexity, occlusion etc. ***Chapter 2*** reviews several existing methods for hand gesture recognition under various conditions. In this research work, multiple frameworks have been proposed to overcome some of the above-mentioned challenges in hand gesture recognition. The main goal of this dissertation is to detect and recognize hand gestures which can be applied for different HCI applications. In the first proposed method presented in ***Chapter 3***, a gesture trajectory tracking method has been proposed through a two-level segmentation process by integrating motion and color information, and a double-tracking system. In the preprocessing step, a skin segmentation framework suppressing the luminance component to compensate for the illumination variation is used. A double-tracking system incorporating mean-shift and particle filters with occlusion handling ability is applied for tracking the gesture trajectory. But tracking the physical movement of the hand is quite challenging due to the varied size, shape and color of the hand. So, in the second work presented in ***Chapter 4***, a motion template guided by optical flow (OFMT) is proposed which can track the movement of the hand irrespective of the shape, size and color of the hand. It is used in a 2D-CNN network along with a 3D-CNN layer in a fusion scheme to enhance the recognition performance. Most hand-crafted features require prior

knowledge which can be avoided by deep learning methods that have a robust and effective feature learning capability directly from the raw data. That is why, deep learning methods have been used for both feature extraction as well as classification purposes. In ***Chapter 5***, a deep learning-based method has been tried incorporating attention-based semantic segmentation to achieve improved results in the recognition of both static and dynamic hand gestures. Experimental results have shown that the proposed trajectory extraction methods have achieved better performance compared to state-of-the-art methods in hand gesture recognition. Finally, conclusions are drawn in ***Chapter 6*** by highlighting the strengths and shortcomings of the schemes and outlining possible extensions as future scope of work.