Segmentation and Recognition of Continuous Human Activity

Anjum Ali

Abstract

The aim of this research is to present a methodology for the automatic segmentation and recognition of human activity. Human activity consists of a continuous flow of discrete human actions in succession. An example of such an activity is a sequence in which a subject enters a room, sits down, then stands up, walks ahead, bends down to pick up something and then gets up and walks away. Having captured such an activity the aim is to first identify breakpoints between the different action components and identify each action in the sequence. Each of the components such as walking, sitting down, standing up, bending down and getting up are discrete action primitives. The automatic interpretation of such a sequence is an uphill task primarily because of its inherent continuity. When humans move from one action to another they do so smoothly and transitions between actions are not well defined. Further since human movement invloves non rigid motion, the amount of deformation that is possible is immense. Some of the earlier work that has been done in this field is cited in [3]. Madabhushi and Aggarwal in [3] successfully classified 12 different classes of actions in the lateral or frontal view. However all the sequences consisted of discrete actions. Davis and Haritaoglu in [2] present a system that recognizes postures associated with actions. They recognize postures from frames that have been manually picked from a sequence. Another work was presented by Campbell and Bobick in [1] where they segment and recognize nine fundamental steps of classical ballet. We presently are looking at continuous sequences of more than four actions in succession. Further we believe that continuous human motion recognition poses a greater challenge than structured movement like ballet which follows well defined patterns.

The subject is tracked over a number of frames in which he/she executes a series of actions in the lateral view. There are no distinct breaks or pauses between the execution of different actions. We have no prior knowledge about the commencement or termination of each action. The aim is to segment the sequence of human activity into separate actions and correctly identify each action. We compute the angles subtended by three major segments of the body with the vertical axis namely the torso, the upper segment of the leg and the lower segment of the leg. This information is used first to detect 'key frames'. Once these are detected a forward and backward search is applied to recognize breakpoint and non break point frames. These breakpoints indicate a particular action's commencement or termination. They are recognized using the angles subtended by the the torso, the upper leg and the lower leg with the vertical axis and applying the nearest neighbor classifier to the feature vector that results.

Training sequences are discrete action sequences. The test sequences, on the other hand, are continuous sequences of human activity that consist of three or more actions in succession. The system will detect the breakpoints between actions of walking, sitting down, standing up, bending, getting up, squatting and rising, and classify each action.

References

[1] Lee Campbell and Aaron Bobick. The recognition of human body motion using phase space constraints. *In Proceedings Fifth International Conference on Computer Vision*, pages 624–630, 1995.

[2] Larry Davis and Ismail Haritaoglu. Ghost: A human body part labeling system using silhouettes. *In Proceedings Fourteenth International Conference on Pattern Recognition,Brisbane, Australia*, pages 77–82, 1998.

[3] Anant Madabhushi and J. K. Aggarwal. Using head movement to recognize human activity. *International Conference on Pattern Recogntion, Baracelona, Spain*, 2000.