Abstract

Extracted features which are obtained from a multiview video stream form a special case of a multi-sensor observation sequence. If the sensors are not synchronous, the observed features of views are not aligned together and this makes some difficulties in classification applications. A new architecture for hidden Markov model, namely pyramidal layered hidden Markov model, is proposed in this paper to handle this situation. This is accomplished by means of separate decoding in each view stream in bottom layer and then fusion of the aligned decoded symbols in top layer. Structure and algorithms of the new structure are introduced and are then used for human behaviour recognition in multiview video sequences. Considering collected information from all views of a multiview human action recognition system, one expects the recognition rate to increase and some problems like occlusion to be rectified. Several experiments have been performed in this paper. The experimental results show high performance, about 93.8% in average, in multiview human behavior recognition, as well as accuracy improvement compared to similar methods. The results are also compared with other contributions on three different multiview behavior datasets.


Index Terms

Computer Science

Video Processing

Keywords

Pyramidal Layered HMM  PLHMM learning and decoding  Dynamic Time Warping  Human Behavior Recognition