Stochastic Representation and Recognition of High-level Group Activities: Describing Structural Uncertainties in Human Activities

M. S. Ryoo and J. K. Aggarwal

Computer & Vision Research Center / Department of ECE

The University of Texas at Austin

{mryoo|aggarwaljk}@mail.utexas.edu

Abstract

A significant amount of computer vision research has addressed the recognition of human activities recently. Researchers are particularly successful in recognizing the activities of one individual or between two individuals, such as pushing and fighting. Notably, in our previous work [1], we have presented a representation syntax to describe high-level human-human interactions based on their sub-events, and proposed a hierarchical algorithm to recognize represented interactions probabilistically. Not only simple interactions such as punching, kicking, and shaking hands are recognized, but also recursive interactions like 'fighting' between two persons are recognized with our previous framework. In this paper, we take our next evolutionary step in human activity recognition: recognition of group activities.

Group activities are the activities that can be characterized by movements of members who belong to one or more conceptual groups [2]. Recognition of groups and their activities will make the detection of high-level events possible. Especially, when such events are semantically meaningful in terms of overall actions of multiple persons considered jointly but not when they are considered individually. Automated recognition of suspicious groups and their activities such as 'a group of thieves robbing the bank' are essential for the construction of high-level surveillance systems. The analysis of movements and plays in team sports also becomes possible with the group activity recognition system. The semantic understanding of military operations and joint works is another application of group activity recognition.

This paper describes a stochastic methodology for the recognition of various types of high-level group activities. Our system maintains the probabilistic representation of a group activity, describing how individual activities of its group members must be organized temporally, spatially, and logically. In order to recognize each of the represented group activities, the system searches for a set of group members that has the maximum posterior probability while satisfying its representation. A hierarchical recognition algorithm utilizing a Markov chain Monte Carlo (MCMC)-based probability distribution sampling has been designed to detect group activities and find the acting groups simultaneously. The system is developed to recognize complex activities such as 'two groups fighting', 'a group of thieves stealing an object from another group', and 'a group assaulting a person'. Videos downloaded from YouTube as well as videos that we have taken are tested. Experimental results shows that our system recognizes complicated group activities, and it does it more reliably and accurately compared to previous approaches by analyzing them probabilistically.

References

[1] M. S. Ryoo and J. K. Aggarwal, "Semantic Representation and Recognition of Continued and Recursive Human Activities", International Journal of Computer Vision (IJCV), 2009 82:1-24.

[2] M. S. Ryoo and J. K. Aggarwal, "Recognition of High-level Group Activities Based on Activities of Individual Members", Proceedings of IEEE Workshop on Motion and Video Computing (WMVC), Copper Mountain, CO, January 2008.