Human Action Recognition via Skeletal and Depth based Feature Fusion

**ABSTRACT**

This paper addresses the problem of recognizing human actions captured with depth cameras. Human action recognition is a challenging task as the articulated action data is high dimensional in both spatial and temporal domains. An effective approach to handle this complexity is to divide human body into different body parts according to human skeletal joint positions, and performs recognition based on these part-based feature descriptors. Since different types of features could share some similar hidden structures, and different actions may be well characterized by properties common to all features (sharable structure) and those specific to a feature (specific structure), we propose a joint group sparse regression-based learning method to model each action. Our method can mine the sharable and specific structures among its part-based multiple features meanwhile imposing the importance of these part-based feature structures by joint group sparse regularization, in favor of discriminative part-based feature structure selection. To represent the dynamics and appearance of the human body parts, we employ part-based multiple features extracted from skeleton and depth data respectively. Then, using the group sparse regularization techniques, we have derived an algorithm for mining the key part-based features in the proposed learning framework. The resulting features derived from the learnt weight matrices are more discriminative for multi-task classification. Through extensive experiments on three public datasets, we demonstrate that our approach outperforms existing methods.

This paper will be presented at The 9th International Conference on Motion in Games 2016 San Francisco, October 10-12, 2016, San Francisco, USA.

Supervisor: Dr LEUNG, Wing Ho Howard

Research Interests: Computer Vision; Multimedia and Image Processing; 3D Human Motion Analysis and Retrieval

All are welcome!