

A Comparative Review of Recent Kinect-based Action Recognition Algorithms

Lei Wang, Du Q. Huynh, Piotr Koniusz

Abstract— Video-based human action recognition is currently one of the most active research areas in computer vision. Various research studies indicate that the performance of action recognition is highly dependent on the type of features being extracted and how the actions are represented. Since the release of the Kinect camera, a large number of Kinect-based human action recognition techniques have been proposed in the literature. However, there still does not exist a thorough comparison of these Kinect-based techniques under the grouping of feature types, such as handcrafted versus deep learning features and depth-based versus skeleton-based features. In this paper, we analyze and compare ten recent Kinect-based algorithms for both cross-subject action recognition and cross-view action recognition using six benchmark datasets. In addition, we have implemented and improved some of these techniques and included their variants in the comparison. Our experiments show that the majority of methods perform better on cross-subject action recognition than cross-view action recognition, that skeleton-based features are more robust for cross-view recognition than depth-based features, and that deep learning features are suitable for large datasets.

For the published version of record document, go to:

<http://dx.doi.org/10.1109/TIP.2019.2925285>