Politechnika Rzeszowska im. Ignacego Łukasiewicza,
Discipline of Information and Communication Technology

Rzeszow, 2023

Summary of the doctoral dissertation Recognition of human activities using point clouds

Author: Kamil Sidor, MSc
Supervisor: Marian Wysocki,PhD, DSc
Key words: activity recognition, point clouds, VFH descriptor, DTW method, BiLSTM neural networks, classifier fusion

The aim of the dissertation was to develop a methodology for recognizing human activities based on sequences of point cloud data obtained from depth images captured by a 3D camera. The considered approach does not require specific clothes of the observed person, facilitates the separation of the person's silhouette from the background, and surface shape representation regardless of lighting conditions.

The scope of the work includes the essential elements of the activity recognition process, especially silhouette extraction, depth map conversion to point clouds, point cloud filtering, feature extraction and selection, as well as the development of classification methodology details. The feature vector is based on the VFH descriptor of appropriately decomposed point clouds to enhance distinctiveness. The classification methods employed are: (1) the nearest neighbors method with distance determined by dynamic time warping algorithm DTW, (2) BiLSTM neural networks, and (3) fusion of neural classifiers.

The research was conducted using three datasets recorded by a single camera (UTD-MHAD and MSR-Action3D with typical activities, and SyKoMi where a deaf person performs sentences in Polish sign language) and one dataset with views from five cameras (UTD Multi-View Action Dataset). All datasets include multiple repetitions of activities by various individuals.

The experiments focused on methods of point cloud decomposition through segmentation of the enclosing bounding box, variants of the feature vector, parameters, effectiveness, and speed of classification algorithms, as well as classifier fusion by voting, Bayesian averaging, Dempster-Shafer theory, and fuzzy integral. For the classification of images from five cameras, a two-criteria approach maximizing classification effectiveness and diversity of the ensemble together was also used.