Politechnika Śląska
Wydział Automatyki, Elektroniki i Informatyki

Rozprawa doktorska

mgr inż. Magdalena Pawlyta

Klasyfikacja zachowań postaci ludzkiej z wykorzystaniem uczenia maszynowego na podstawie trajektorii punktów charakterystycznych w reprezentacji 3D i 2D

Promotor: prof. dr hab. inż. Konrad Wojciechowski,
Promotor pomocniczy: dr inż. Przemysław Skurowski,

Gliwice 2024

ABSTRACT

This thesis seeks to explore the extent to which the ability to recognize human actions/behavior in vision systems depends on the representation of tracked points, their number, and the position and orientation of the camera in the case of 2D representation.

For this purpose, a 3D motion database containing 9 unique simple actions was created using Motion Capture technology. Using this technology allowed for obtaining a reference, undistorted data. To avoid bias in the classification results, an algorithm for dividing the data into any number of subsets was developed.

A 2D database was created based on the basis of the 3D database. For this purpose, a virtual camera model was developed, with both internal and external parameters modifiable. This virtual camera was set on a virtual scene in various locations, and then a series of perspective projections was performed. The results of these projections constituted the source of 2D data.

Several classification experiments were carried out in both 3D and 2D space. Since the main focus of this thesis was not to develop an ideal model for classifying human behavior but to examine aspects affecting the quality of classification, two quite generic architectures of deep neural networks were used for experiments - the CNN network and the LSTM network.

This work shows the impact of changing data representation from 3D to 2D on classification accuracy. It has also been shown that reducing, the dimensionality of the input space improves the classification quality to some extent. The level of reduction depends on the model used and the data space - different in 3D space and different in 2D space.

The influence of perspective projection on the overall classification accuracy and the recognition of individual behaviors/actions was also examined. For this purpose, two additional measures were introduced, related to the specificity of the created database. New methods of visualizing the results in the form of special domes were also developed. The obtained results were analyzed in detail and discussed separately for each action. This allowed for demonstrating the impact on the quality of classification, although the results obtained in some cases are not intuitive.

Keywords: classification of human behavior, machine learning, deep neural networks