Continuous Hand Gesture Recognition using Deep Coarse and Fine Hand Features

Hazem Wannous* and Jean-Philippe Vandeborre
IMT Nord Europe, Univ. Lille, CNRS, UMR 9189 - CRISTAL, F-59000 Lille, France
*Hazem.wannous@imt-nord-europe.fr

Introduction

Context and Issues:
- Using hand gestures as an HCI modality introduces intuitive and easy-to-use interfaces for a wide range of applications.
- The hand is an object with a high number of degrees of freedom and with high similarities derived from the heterogeneities of possible gestures.
- Feature learning has to learn mutually spatial and temporal information, because gestures can be defined both by the shape variations and movements of the hand.
- The computation complexity has to be small enough so that the algorithm can predict an incoming gesture in real time

Contributions:
- We propose an end-to-end architecture based separately on learned temporal variation of coarse and fine features extracted from a CNN trained on depth sequences.
- Both features fed to two RNNs in order to model the temporal aspect of the hand poses and the shape variations over the time.
- We introduce of a new dataset of heterogeneous gestures recorded in an online scenario by a depth camera.
- We design a light efficient approach for online recognition of hand gestures. Simplicity and lightness is one of our goals for HCI applications.

Approach

Datasets:
- **Online DHG**: 280 sequences of 10 continuous gestures, 14 categories: fine and coarse gestures (14 and 28 classes). Pre-segmented sequences of this dataset constitutes the SHREK17 dataset of 2800 gesture sequences [1].
- **NVIDIA** [2]: 1532 sequences of 25 gestures captured following a HCI based on hand gestures in a car scenario.

Metrics:
- Receiver Operating Characteristic (ROC) curve
- Normalized Time to Detect (NTID) [3]
- Accuracy

Results

Conclusion

References