

Thesis advisor(s)
Timothy Baldwin
Manfred Pinkal

Author
Matěj Korvas

Extracting Characteristics of Human-produced Video Descriptions

Abstract

This thesis contributes to the SMILE project, aiming for video understanding. We focus on the final stage of the project where information extracted from a video should be transformed into a natural language description. Working with a corpus of human-made video descriptions, we examine it to find patterns in the descriptions. We develop a machine-learning procedure for finding statistical dependencies between linguistic features of the descriptions. Evaluating its results when run on a small sample of data, we conclude that it can be successfully extended to larger datasets. The method is generally applicable for finding dependencies in data, and extends methods for association rule mining for the option to specify distributions of features. We show future directions which, if followed, will lead to extracting a specification of common sentence patterns of video descriptions. This would allow for generating naturally sounding descriptions from the video understanding software.