Formal Grammars in Recognition of Activities in Visual Data

Jan Tinka

itinka@fit.vutbr.cz

November 26, 2018

Visual data contains a lot of information, hence it is no wonder a whole field of computer science called Computer Vision exists. Main focus of computer vision is to gather and understand high-level information that is encoded in visual data. Visual data is processed in a myriad of different ways depending on what part of the processing pipeline we look at, and what kind of information is being gathered. Examples of such information could be volume of traffic on the streets, anomalies in a workplace or movement of people.

This presentation discusses a family of techniques for activity detection and recognition using stochastic context-free grammars, leveraging the hierarchical logical structure and predictability of some types of activities. This is demonstrated on two case studies. The first one shows an elementary use of a SCFG for hand-gesture recognition. The second case study then demonstrates the use of a SCFG for recognition of a simple square gesture and music beat pattern from the hand movement of a musical conductor. Before the actual examples an Early-based context-free parsing algorithm, used as a parser in these examples, is introduced.

At the end a Haar-like feature classifier cascade and Hidden Markov Models, which were used to generate the terminal symbols in the use cases, are very briefly discussed.

[1] Bobick, A. F.; Ivanov, Y. A.: Action Recognition using Probabilistic Parsing. 1998.

[2] Chen, Q.; Georganas, N. D.; Petriu, E. M.: M.: Hand gesture recognition using Haarlike features and a stochastic context-free grammar. *IEEE Transactions on Instrumentation and Measurement*. 2008: pp. 1562–1571.