

Automation and Standardization of Semantic Video Annotations for Large-Scale Empirical Film Studies

Henning Agt-Rickauer / Christian Hentschel / Harald Sack

Hasso Plattner Institute for IT Systems Engineering, University of Potsdam, Potsdam, Germany / FIZ Karlsruhe & Karlsruhe Institute of Technology Karlsruhe, Karlsruhe, Germany

The study of audio-visual rhetorics of affect scientifically analyses the impact of auditory and visual staging patterns on the perception of media productions as well as the conveyed emotions. By large-scale corpus analysis of TV reports, documentaries and genre-films of the topos “political crisis”, film scientists aim to follow the hypothesis of TV reports drawing on audio-visual patterns in cinematographic productions to emotionally affect viewers. However, localization and description of these patterns is currently limited to micro-studies due to the involved extremely high manual annotation effort. The [AdA Project](#) presented here, therefore, pursues two main objectives: 1) creation of a standardized annotation ontology based on Linked Open Data principles and 2) semi-automatic classification of audio-visual patterns. Linked Open Data annotations enable the publication, reuse, retrieval, and visualization of data from film studies based on standardized vocabularies and Semantic Web technology. Furthermore, automatic analysis of video streams allows to speed up the process of extracting audio-visual patterns. Temporal video segmentation, visual concept detection and audio event classification are examples for the application of computer vision and machine learning technologies within this project. The ontology as well as the created semantic annotations of audio-visual patterns are published as [Linked Open Data](#) in order to enable reuse and extension by other researchers. The annotation software as well as the extensions for automatic video analysis developed and integrated by the project are [published as open source](#) as we envision these tools to be useful for general deep semantic analysis of audio-visual archives.