



Bundesministerium
für Bildung
und Forschung

Freie Universität



Berlin



Digital Engineering · Universität Potsdam



Automation and standardization of semantic video annotations for large-scale empirical film studies

SWIB 2018

Henning Agt-Rickauer / **Christian Hentschel** / Harald Sack
Hasso Plattner Institute, University of Potsdam, Germany

Analyzing Audio-Visual Rhetorics of Affect



- empirical research on audio-visual rhetorics by means of film analysis
 - film scientist from FU Berlin
 - computer scientists from HPI, Université de Nantes
- guiding research question/project goals:
 - How do audio-visual images shape emotional attitudes towards certain topics?
 - identifying an initial set of audio-visual rhetorical figures (typology)
 - developing computational methods for the study of audio-visual rhetorics
- subject matter:
 - feature films, documentaries and tv news reports on the global financial crisis (2007-), total: >100h

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Motivation



- identification, localization and classification of audio-visual staging patterns
- many annotations necessary for a scientific and holistic understanding of a movie
- technological requirements
 - a. consistent data management
 - b. support for semi-automatic annotation data generation

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Linked Open Data - consistent data management

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Chart 4

AdA Ontology - Motivation

eMAEX annotation routine

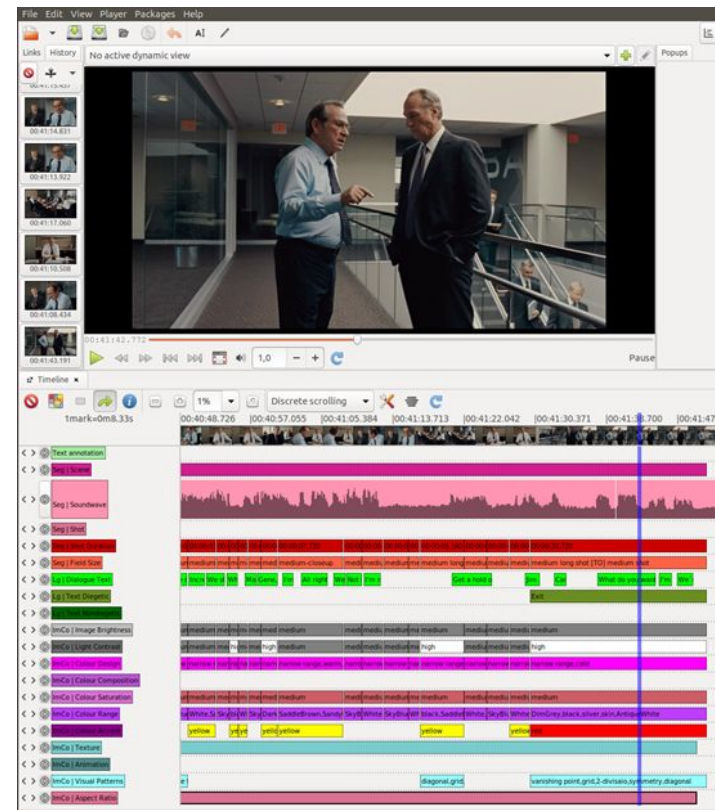
- Film-analytical method
- Systematic: categories, types, values
- ...but not machine-readable

Free annotations

- Natural language
- Typos
- Synonyms (medium shot vs. waist shot)
- Spelling (colour range vs. color range)

Goal

- Reusable, explicit vocabulary with film-analytical concepts, terms and descriptions
- Accessible on the Web
- Integrate into video annotation software Advене

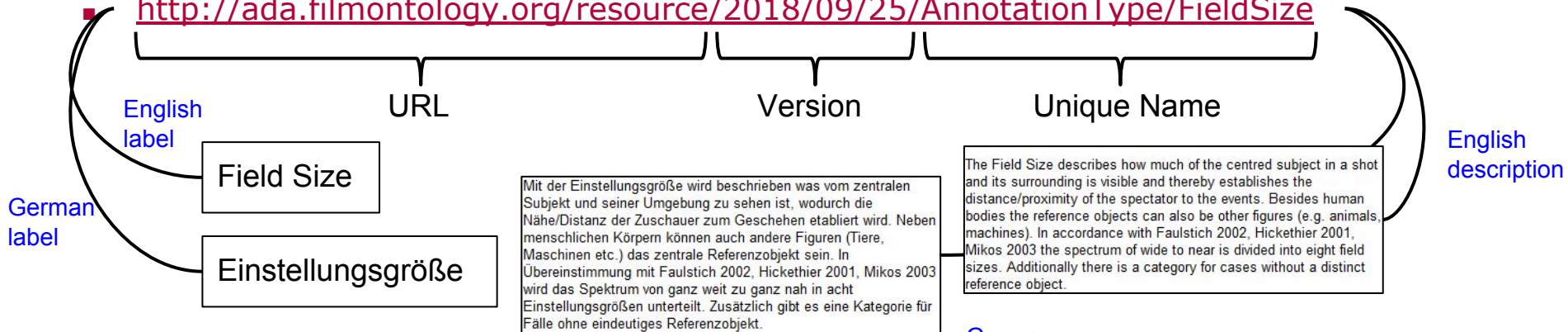


AdA Ontology - Vocabulary

Unique identifiers for domain-specific concepts and terms

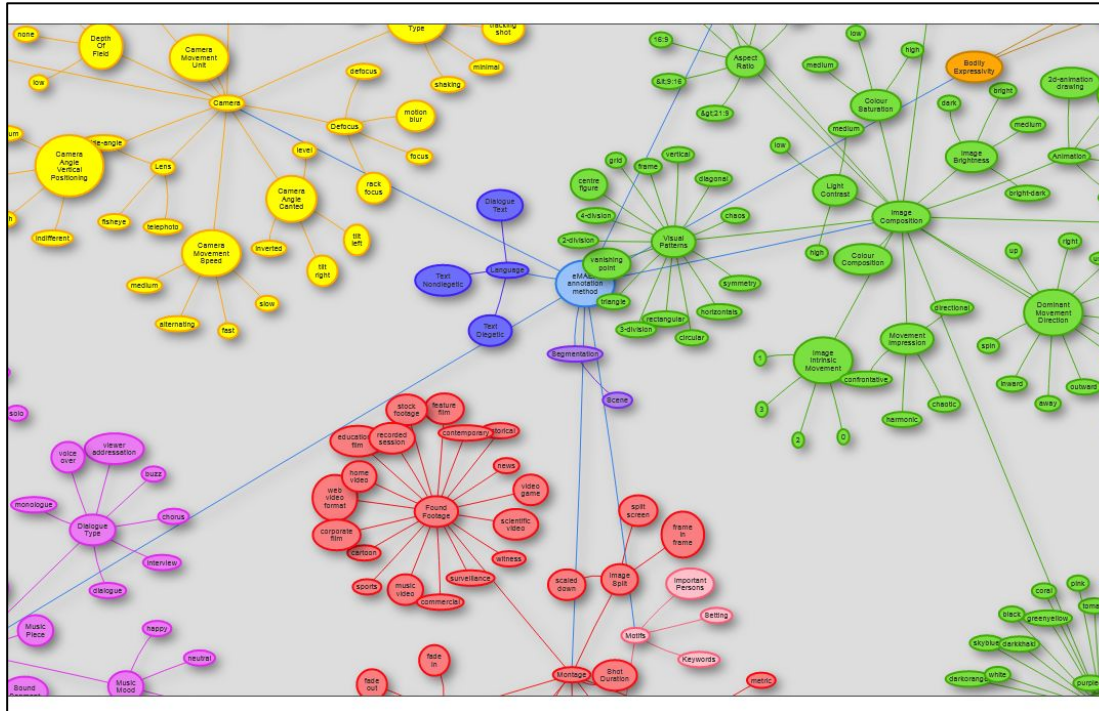
- Uniform Resource Identifier (URI)

<http://ada.filmontology.org/resource/2018/09/25/AnnotationType/FieldSize>



- Store information and make it retrievable
- encoded with RDF

AdA Ontology - Vocabulary Visualization Demo



Annotation Vocabulary

- 9 Annotation Level
- 78 Annotation Types
- 435 Annotation Values

Download at

<https://github.com/ProjectAdA/public>

Automation and standardization
of semantic video
annotations for large-scale
empirical film studies

Christian Hentschel

AdA Ontology - Example Annotation

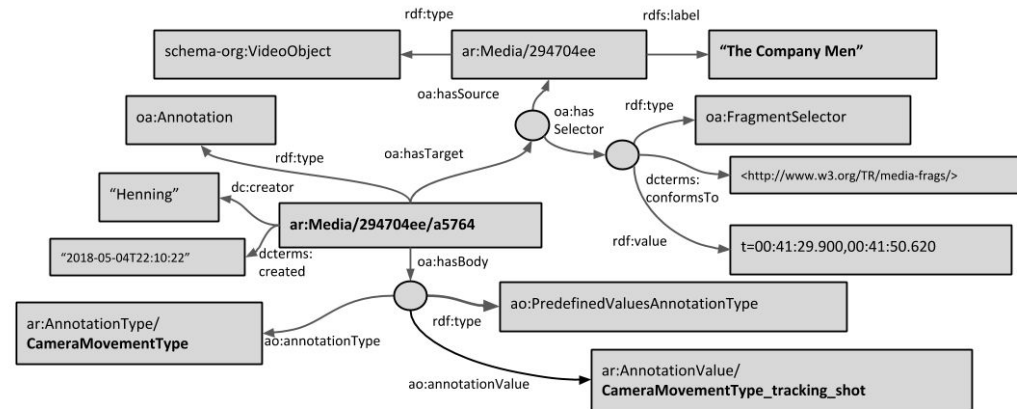
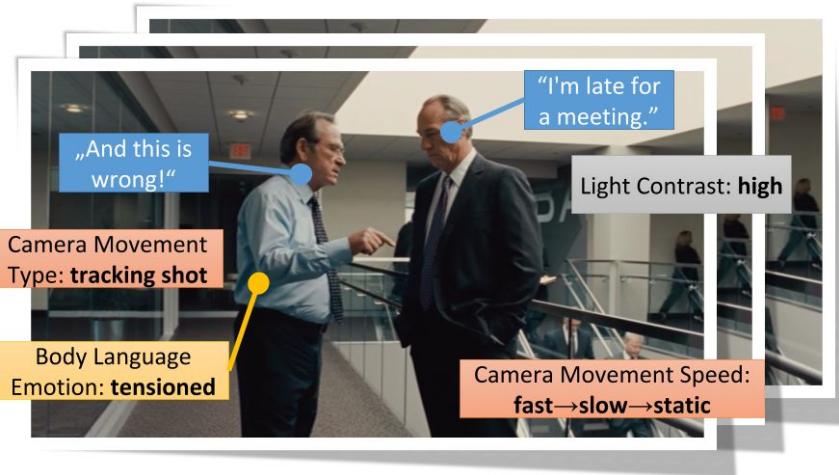


Chart 8

Linked Data Applications

example: Company Men

- More than 24,000 annotations, mostly manual

Goal

- Publish this valuable data by means of Linked Data

How

- Advene RDF Export
- AdA Ontology Data Model
- W3C Web Annotation Standard, Media Fragments URI

Make Linked Data Usable

- Visual Analysis
- Queries

annotype	count
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/ColourRange	3041
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/ImageContent	1383
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/DominantMovementDirection	1288
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/FieldSize	1086
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/CameraAngle	1051
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/CameraMovementType	1019
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/ImageIntrinsicMovement	937
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/CameraAngleVerticalPositioning	881
http://ada.filmontology.org/resource/2018/09/25/AnnotationType/DialogueEmotion	744

...

Automation and standardization of semantic video annotations for large-scale empirical film studies

Christian Hentschel

Annotation Query

Motivation

- Huge amount of annotations
- How to find interesting parts / patterns?

Goals

- Search and retrieve segments with same characteristics
- Within a movie and across movies

Movie 1



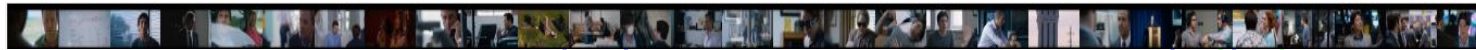
BodyLanguageIntensity: 5
ImageContent: Group

BodyLanguageIntensity: 5
ImageContent: Group

Automation and standardization
of semantic video
annotations for large-scale
empirical film studies

Christian Hentschel

Movie 2



BodyLanguageIntensity: 5
ImageContent: Group

BodyLanguageIntensity: 5
ImageContent: Group

Chart 10

Annotation Query - Demo

Ada Annotation Inspector

[Annotation Timeline](#) [Video Selection](#)

Movie:

The Company Men

Annotation Type:

Motf | Image Content

Annotation Value:

ImageContent_person

Annotation Type:

BodExp | Body Language Emotion

Annotation Value:

BodyLanguageEmotion_surprised

GET VIDEOS

REMOVE LAST SELECTOR

ADD SELECTOR



<http://ada.filmontology.org/annotations/>

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Chart 11

Automated Multimedia Analysis - support for semi-automatic annotation data generation

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Chart 12

Automated Multimedia Analysis

- huge amounts of annotations
 - Company Men: more than 24.000
 - labor intense: 3 mins of video → 10-12h of manual annotation
 - error-prone
- make a computer able to summarize the contents of video
 - (to some syntactical extend)
 - by extracting low-level features
 - increase the speed of video annotation
- two modalities:
 - audio stream
 - video stream

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Automated Multimedia Analysis



- Examples:
 - Montage/ShotDuration
 - ImageComposition/ColourRange
 - Language/DialogueText

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Automated Multimedia Analysis



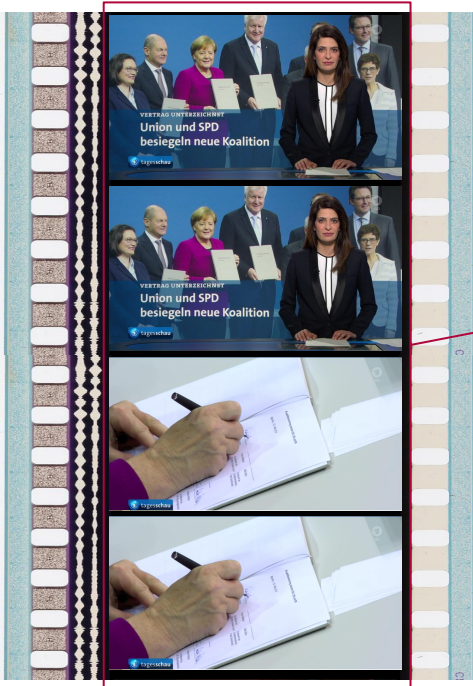
Montage/ShotDuration

Duration of a shot. A Shot of a film is a perceivable continuous image and is bound by a discontinuation of the whole composition.

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

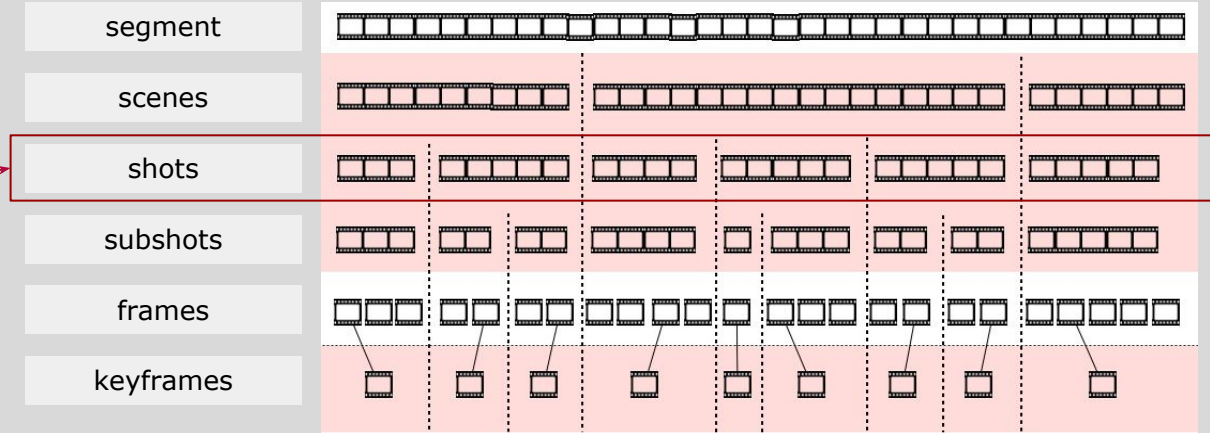
Christian Hentschel

Automated Multimedia Analysis



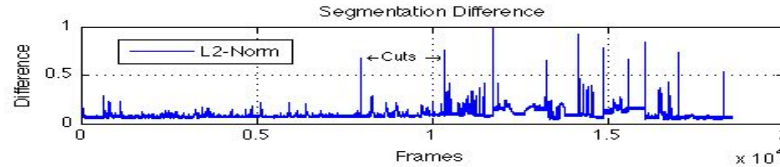
Video

structural segmentation



Automated Multimedia Analysis

- Example: Shot-Detection
- Uses differences in consecutive images to identify discontinuities
 - idea: high visual redundancy in video stream



- Type of cuts:
 - hard-cuts
 - soft-cuts (fade-in, fade-out, wipe)
 - should be robust to artifacts (e.g., dropouts)

Automated Multimedia Analysis



ImageComposition/ColorRange

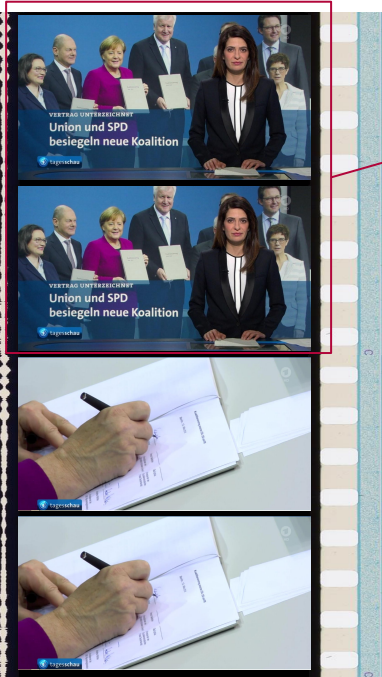
Simplified notation of the color range that is used in a sequence.
For the purpose of comparability colors have to be picked from a reduced set of colors.

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Chart 18

Automated Multimedia Analysis



Video



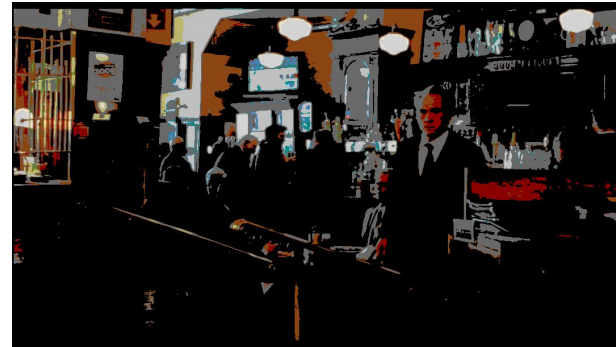
...



- quantize all colors in a shot according to their most similar color from palette
 - compute Euclidean distance between color values of palette and frames
 - find NN

Automated Multimedia Analysis

- quantize according to CIE L*a*b*
 - color model according to human perception
 - separates chroma from lightness
 - Euclidean distance between color values similar to perceived color differences



black, white, wheat1, gold,
saddlebrown, khaki, blue

'black': 0.63,
'dimgrey': 0.21,
'saddlebrown': 0.06,
'silver': 0.05

Automation and standardization
of semantic video
annotations for large-scale
empirical film studies

Christian Hentschel

Chart 20

Automated Multimedia Analysis



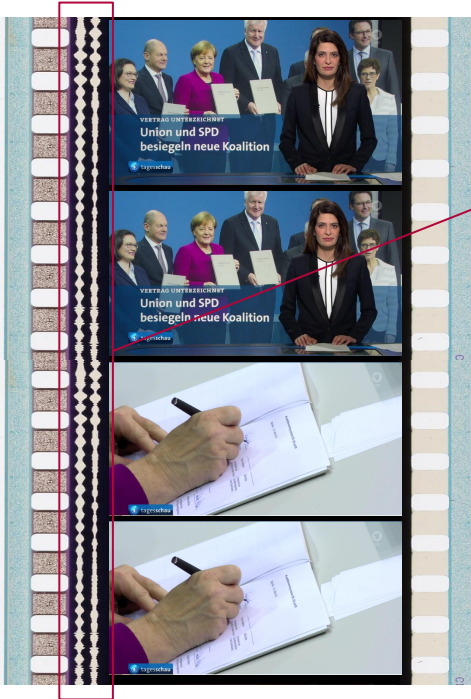
Language/DialogueText

Dialogue is a transcription of understandable, spoken language that is dominant within the film. This is usually dialogue from protagonists, off-commentary, but also chorus. Nonverbal utterances (e.g. laughing, coughing, stuttering) will not be transcribed in this basic version.

**Automation and standardization
of semantic video
annotations for large-scale
empirical film studies**

Christian Hentschel

Automated Multimedia Analysis



Audio



ASR

- Automatic Speech Recognition (ASR)
 - subtitles?

Automation and standardization
of semantic video
annotations for large-scale
empirical film studies

Christian Hentschel

Automated Multimedia Analysis - ASR



- based on supervised machine learning
 - requires (large) corpus of manually transcribed speech

- 2 stage approach
 1. acoustic model
 - convolutional neural network that transcribes utterances to letters
 - trained on ~1000 hours of audiobook recordings (LibriSpeech)

 2. language model
 - domain specific mapping of letters to words
 - based on word/letter co-occurrences

Automation and standardization
of semantic video
annotations for large-scale
empirical film studies

Christian Hentschel

Automated Multimedia Analysis - ASR

example



we review and some five hundred projects and programmes focusing on those with significant marketing opportunities song everything home contribute immediately to profitability selecting thirty set as promising to teach it growth program setting aside the rest for future consideration as taking more than miss mclarry you were talking earlier about fiscal two thousand eleven and a good job of convincing us that the credit markets frozen your sales revenue in two thousand ten love grey overs but can you talk about two thousand eleven what sort of a percentage increase you anticipate talk your people do our people who are you suggesting that you are expecting any gross in your division extent arm suggesting that we face increased foreign competition and a difficult credit market for large capital expenditures like no growth and two thousand eleven i am confident that while shipbuilding will remain challenge

Automation and standardization of semantic video annotations for large-scale empirical film studies

Christian Hentschel

Chart 24



Affektrhetoriken des Audiovisuellen

Thank you
for your attention!

Henning Agt-Rickauer, Christian Hentschel, Harald Sack

Hasso Plattner Institute, University of Potsdam, Germany