

MPEG-7 Video Annotation Tool

Helmut Neuschmied

(JOANNEUM RESEARCH, Austria
helmut.neuschmied@joanneum.at)

Abstract: This paper presents a video annotation system which enables efficiently to annotate video footage. It provides automatic feature extraction methods which support the user in navigation through and structuring of the video content. An annotation at image region level is supported by object redetection and tracking functionalities. The result content description of the footage is saved in MPEG-7.

Keywords: MPEG-7, Video Annotation Tool, Automatic Feature Extraction, Metadata

Categories: H.3, I.4

1 Introduction

A detailed description of the video content can be used for different purposes. It is especially required for content retrieval, film analysis, and as input data for generating interactive video content. The annotation of video footage is a time-consuming task. We have developed annotation tools which enables the efficient annotation in combination with a set of automatic annotation plug-ins. All metadata of the tools is stored in the ISO standard MPEG-7 [ISO/IEC, 01] [MPEG-7 Lib, 07]. The annotation workflow can be seen in *Figure 1*. The pre-processing tool extracts metadata automatically from the video. After that this data can be edited and extended semi-automatically by the MPEG-7 Video Annotation Tool.

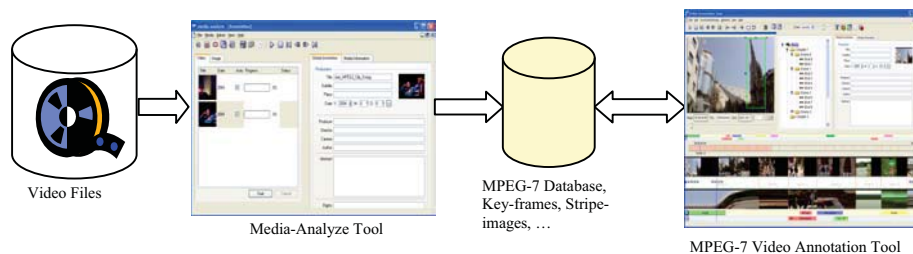


Figure 1: Annotation workflow

2 Media Analyze – Pre-processing Tool

In a first preparation step of the annotation workflow the global description of the source video data has to be specified and an automatic analysis process has to be

started. For this a separate tool, the Media Analyze Tool, has been developed. During the automatic content analysis the characteristic camera motion, shot boundaries including dissolves, relevant key-frames, several image similarity features, and the visual features which are required for searching for similar image regions in the video are extracted. The result is a first metadata description of the video which is stored in MPEG-7. The automatic analysis can be started for several videos and then the process can run over night.

The core part of this program is a module framework. The framework serves as an execution environment for analysis modules. The modules are interconnected by the framework to constitute a so-called module graph (see Figure 2). The module graph is defined by a XML file.

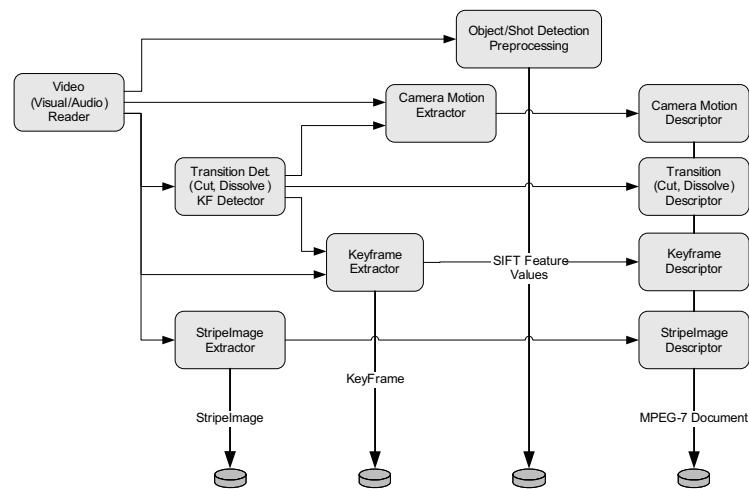


Figure 2: Analyses module graph of the MediaAnalyze Tool

3 MPEG-7 Video Annotation Tool

As soon as the automatic analysis of a video is completed the produced metadata description can be displayed, edited and extended by the MPEG-7 Video Annotation Tool (see Figure 3). This tool has a number of views which enables fast and easy navigation in the video. Through the key frames and the stripe image one gets a quick overview of the video content. There are two time lines, one for the whole video time and one which shows only a selected time period (time zoom). In the time lines the shot boundaries and the dissolves are displayed and they can also be edited. There is also the possibility to structure the video depending on the video content. For example shots can be grouped to scenes; scenes can be combined to build chapters and so on. This structure yields a kind of table of contents and is displayed by a separate view. Depending on the selected structural element different textual annotations are possible. These are for example the title of the structural element, content description, remarks, and specifications about time, location, and persons. At the shot level shooting settings like camera motion, camera angle, or view size can be documented.

The integrated video player has drawing functionalities for the annotation of regions (objects). Image regions can be specified by drawing a rectangle or a polygon. Once an image region is specified (see green rectangle in Figure 3) it is possible to start an automatic search for a similar region in other shots of the video. The result of the object search is displayed in a separate key-frame view. The object redetection functionality is very fast because it uses the pre-processed visual feature values [OBER, 05]. By using the object search result textual annotations can be assigned simultaneously to several objects or shots.

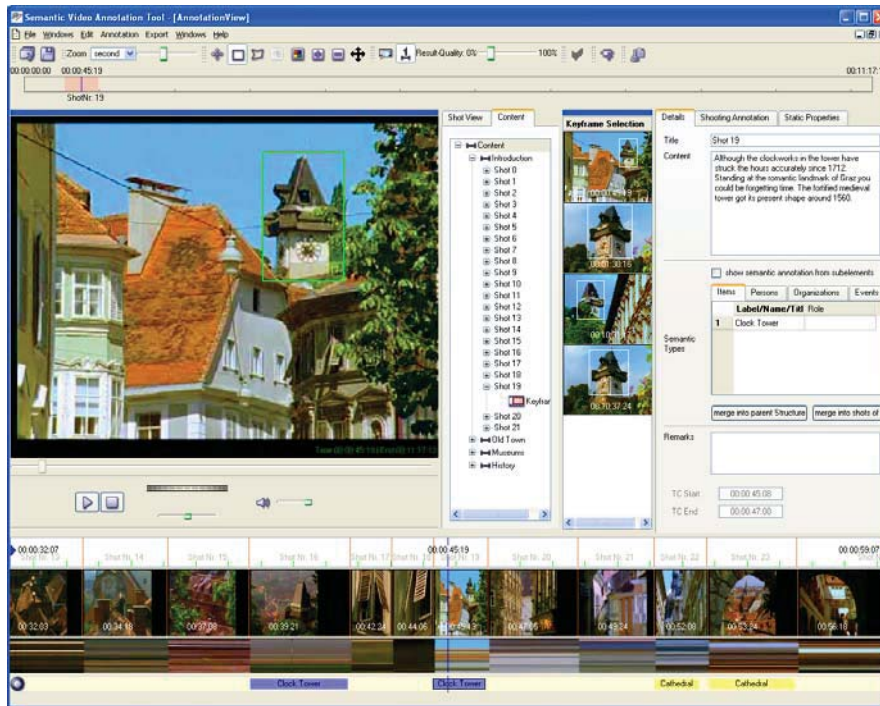


Figure 3: User interface of the MPEG-7 Video Annotation Tool

4 Conclusions

A video annotation system has been presented which allows creating a MPEG-7 content description in different levels of detail (e.g. scene, shot, key-frame, object).

References

[MPEG-7 Lib, 07] MPEG-7 library: <http://mpeg-7.joanneum.at>

[ISO/IEC, 01] ISO/IEC 15938:2001

[OBER, 05] Ober Sandra, Neuschmied Helmut, Bischof Horst, Object Recognition and Video Indexing for Media Analysis, OEGAI Journal, 24 (1), pp. 4-10, 2005