

An Approach for Browsing Video Collections in Media Production

Werner Bailer, Wolfgang Weiss, Christian Schober, Georg Thallinger

JOANNEUM RESEARCH Forschungsgesellschaft mbH
DIGITAL – Institute for Information and Communication Technologies
Steyrergasse 17, 8010 Graz, Austria
{firstname.lastname}@joanneum.at

Abstract. This paper describes a video browsing tool for media (post-) production, enabling users to efficiently find relevant media items for redundant and sparsely annotated content collections. Users can iteratively cluster the content set by different features, and restrict the content set by selecting a subset of clusters. In addition, similarity search by different features is supported. Desktop and Web-based variants of the user interface, including temporal preview functionality, are available.

1 Introduction

The proposed video browsing tool attended the Video Browser Showdown (VBS) 2012 [3]. The tool has been designed for applications in (post-) production phase of movie and broadcast production. In this application scenario, users typically deal with large amounts of audiovisual material with a high degree of redundancy and need to select a small subset for use in a production. Newly shot material is typically sparsely annotated, thus the browsing tool has to rely on automatically extracted features. The content sets in this application are typically larger than those used in the Video Browser Showdown, reaching about 100 hours.

Automatic content analysis is performed during the ingest of content. Currently, camera motion estimation, visual activity estimation, extraction of global color features and estimation of object trajectories are performed. The extracted features are represented using the MPEG-7 Audiovisual Description Profile [1] and indexed in an SQLite database.

In order to select content, the user follows an iterative selection process, consisting of alternating steps of clustering and selecting subsets of the current data set. Users cluster content by one of the automatically extracted features and can then select relevant clusters to reduce the content set. Further clustering by the same or other features can then be applied to the reduced set. In addition, items similar to one of the items in the cluster can be retrieved. A more detailed description of the tool and the browsing process can be found in [2].

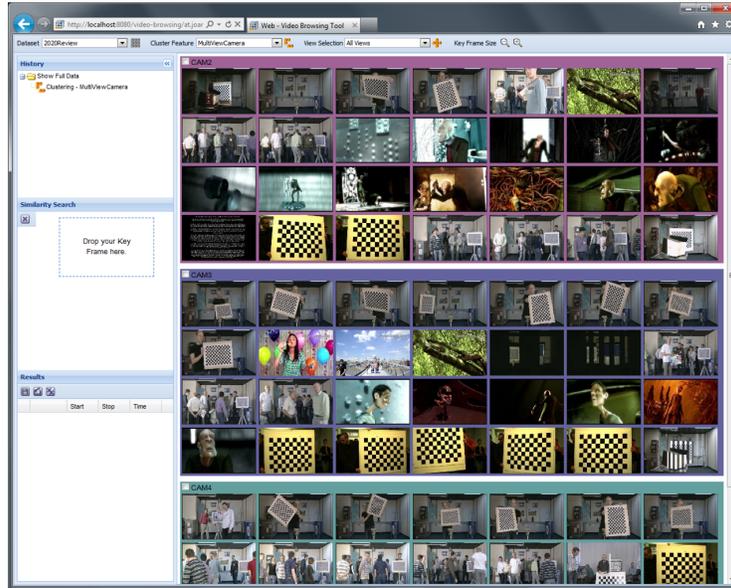


Fig. 1. Screenshot of the Web-based video browsing tool

2 Browsing User Interface

The central component of the video browsing tool's user interface is a light table (cf. Figure 1). The light table shows the current content set and cluster structure using a number of representative key frames for each of the clusters. The clusters are visualized by colored areas around the images. The size of the images in the light table view can be changed dynamically so that the user can choose between the level of detail and the number of visible images without scrolling. By clicking on a key frame in the light table view, a video player is opened and plays the segment of the video that is represented by that image. The temporal context of a key frame is shown by a time line of temporally adjacent key frames that appears when the user moves the mouse over a frame. This time line shows one line of key frames which is limited by the width of the screen. If the user wants to get a broader range of temporally adjacent key frames, it is possible to zoom in. Therefore, a larger list of key frames is shown in the light table.

The tool provides support for performing similarity search based on following features: camera motion, motion activity, color layout, multi-view media item and multi-view camera. To execute a similarity search, the user drags a key frame into the similarity search area (see also Figure 2), selects a similarity search option and executes the search. In addition, segments from the temporal proximity of a result segment can be retrieved. The latter supports the user in cases where a presented item is topically similar to the wanted item and might thus be nearby in the programme, but is not similar in terms of visual features.

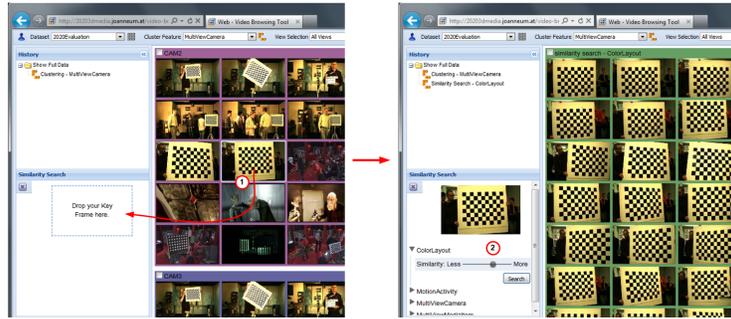


Fig. 2. Executing the similarity search

On the left side of the application window the history and the result list are displayed. The history window automatically records all clustering and selection actions done by the user. By clicking on one of the entries in the history, the user can go back to a previous point. Then users can choose to discard the subsequent steps and use other cluster/selection operations, or to branch the browsing history and explore the content using alternative cluster features. The result list can be used to memorize video segments and to extract segments of videos for further video editing, e.g. as edit decision list (EDL). Users can drag relevant key frames into the result list at any time, thus adding the corresponding segment of the content to it.

The user interface is available both as desktop and as a Web-based version, offering the same functionality. Both use the same backend implementation, which is accessible as a SOAP Web service for the Web-based client.

Acknowledgements. The research leading to these results has received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreements n° 215475, “2020 3D Media – Spatial Sound and Vision” (<http://www.20203dmedia.eu/>) and n° 287532, “TOSCA-MP - Task-oriented search and content annotation for media production” (<http://www.tosca-mp.eu>).

References

1. Information technology - multimedia content description interface - part 9: Profiles and levels, amendment 1: Extensions to profiles and levels. ISO/IEC 15938-9:2005/Amd1:2012, 2012.
2. Werner Bailer, Wolfgang Weiss, Gert Kienast, Georg Thallinger, and Werner Haas. A video browsing tool for content management in post-production. *International Journal of Digital Multimedia Broadcasting*, 2010, Mar. 2010.
3. Werner Bailer, Wolfgang Weiss, Christian Schober, and Georg Thallinger. A Video Browsing Tool for Content Management in Media Post-Production. In *18th International Conference on Advances in Multimedia Modeling*, volume 7131 of *Lecture Notes in Computer Science*, pages 658–659. Springer Berlin, Heidelberg, January 2012.