TOSCA-MP: TASK-ORIENTED SEARCH AND CONTENT ANNOTATION FOR MEDIA PRODUCTION

Alberto Messina[1], Frank Gläser[2], Eugene Mbanya[3], Carlos Ruiz[4], Werner Bailer[5]

[1]RAI, IT; [2]Techicolor, DE; [3]Fraunhofer HHI, DE; [4]playence, ES; [5]JOANNEUM RESEARCH, AT

ABSTRACT

TOSCA-MP researches and develops tools for content analysis and search in professional media production processes. This paper discusses three core topics: A framework for networked media repositories, called Distributed Repository Framework, adaptive content analysis tools and semantic search.

1. PROJECT OVERVIEW

The TOSCA-MP[1] project aims to develop user-centric content annotation and search tools for professionals in networked media production and archiving (television, radio, online), addressing their specific use cases and workflow requirements. The 30 months project started in Oct. 2011 and brings together 10 partners from 5 European countries including industrial partners providing solutions for the media industry, public service broadcasters as well as their European association, a university and research centers. TOSCA-MP investigates scalable and distributed content processing methods performing advanced multimodal information extraction and semantic enrichment. Other key technology areas include search methods across heterogeneous networked content repositories and novel user interfaces. An open standards based service oriented framework integrates the components of the system. TOSCA-MP enables professionals in media production and archiving to seamlessly access content and indexes from distributed heterogeneous repositories in the network. This is achieved by providing technologies that allow instant access to a large network of distributed multimedia databases, including beyond state-of-the-art metadata linking and alignment. The distributed repositories can be accessed through a single user interface that provides novel methods for result presentation, semi-automatic annotation and means for providing implicit user feedback. The networked approach of TOSCA-MP enables content holders to leverage scalable distributed processing in the network, using both, in-house or external service models. The project develops models of key user tasks in the audiovisual media production workflow. These models are used to adapt the components of the system to the specific and dynamic requirements of real user tasks in the media production domain, and to evaluate the tools in a cost-effective way. In the following, we discuss three core topics of TOSCA-MP: A framework for networked media repositories, called Distributed Repository Framework, adaptive content analysis tools and semantic search.

[1] EU FP7/2007-2013 project n° 287532, for more information see http://www.tosca-mp.eu

2. DISTRIBUTED REPOSITORY FRAMEWORK

The main idea of the Distributed Repository Framework (DRF) is to create a framework for the realization of customized networked distributed repositories for professional environments in the media industry. Based on this framework, a distributed repository for networked distributed information extraction, semantic enrichment and search is being realized for TOSCA-MP.

A DRF-based repository implementation consists substantially of a number of local repositories which are administrated and managed by a central module. The internal communication and synchronization is realized based on a messaging system with different messages for configuration and data handling. The DRF provides general mechanisms for data-format and -model independent data handling in combination with a global interface, which provides access to all data stored in the networked distributed local repositories.

For seamless handling of unstructured and structured data, the DRF provides the basis for realizations of various types of specific local repositories with different sets of functionality and specific local interfaces. Local repositories can consist of various types of internal storage and are able to handle dedicated data formats and data models, depending on the requirements of the associated local user applications. Due to this open architecture, the integration of third party repositories in the form of a local repository is supported well.

3. ADAPTIVE CONTENT ANALYSIS

Adaptivity denotes here the ability to perform content analysis tasks depending on specific properties of the content. In TOSCA-MP the main result in this context is

represented by genre-dependant content processing. The first step of the process consists in automatically identifying the genre of the content (e.g., news, talk shows, commercials) using an approach based on an extended version of [1], which is now capable of online genre calculation on arbitrary portions of the content. This is possible due to a persistence layer that saves low-level features in a key-value store and retrieves the features on a content segment level. Based on genre information, a rule system is applied which at first associates content items to collections depending on their genre, and then applies content processing rules to the obtained collections. This layered approach is flexible to a number of changing conditions in the processing chain, including different processing rules for the same genre in specific application domains, and different processing choices applied to different portions of the same content.

Within this framework, the project consortium works on improving a number of content analysis tools. The approach for detecting concepts in videos is described in [2]. Also, as a result of the fact that automatic techniques for analyzing video content usually suffer from the lack of appropriate and sufficient training data, especially in broad video domains such as general broadcast and web video content, static image content (for which well-defined and extensive datasets exist) is used to augment video frames. This takes into account certain adaptation techniques to reduce the cross media gap between images and video frames. Furthermore, based on the ontology of concepts to be detected in the videos, exclusion rules (rules applied to sets of mutually exclusive occurring concepts e.g. “indoor” and “outdoor”, ensuring only one is present in the video frame) and inference rules (rules applied to sets of categories such as “tree” and “plant”, in order to boost the probability of existence of the inferred category (plant) to the inferring category’s (tree) probability of existence, if it is otherwise smaller) are applied to optimize generated metadata [2].

We also work on tools for assessing the quality of audiovisual content in order to decide about the reusability of archive content. This includes automatic detectors for the common impairments noise and grain, video breakups, sharpness, image dynamics and blocking, as well as user interfaces for the efficient viewing and verification of the automatic results by an operator. Results on quality analysis are described in detail in [3]. The analysis algorithms are integrated into a service oriented architecture, focusing on the recent standardization efforts by EBU and AMWA’s Joint Task Force on a Framework for Interoperability of Media Services in TV Production (FIMS).

4. SEMANTIC SEARCH

TOSCA-MP aims at providing a networked search engine with a seamless and unique interface over distributed repositories, offering the retrieval of assets, frames and segments based on feature and semantic descriptions. Therefore, the search engine indexes the metadata extracted from feature extraction services and exploits the semantics described by multiple domain ontologies that can be deployed in the system in an incremental and independent way.

Three main parts compose the networked search engine: annotation components, which connect different feature and metadata extractors with the domain ontologies and the internal search schema; indexing, which uses all available annotations (linguistics-, semantic-, and feature-based) as well as other relevant information for retrieval (e.g. relevance) to create the index; semantic search, which processes the user’s query, indentifying main concepts and relationships within the domain ontologies, carrying out a semantic extension of the query to improve the accuracy and recall.

Besides the search component, the system exploits different visual paradigms for representing results in order to allow exploratory analysis where, instead of having a one-shot search process, the process of refinement, filtering, and query enhancement is supported by different visual functionalities.

5. DEMO

Concerning the DRF, the demo shows a networked distributed content and metadata repository that supports distributed metadata processing and enrichment. It demonstrates the possibility of seamless handling of structured and unstructured stored semantic metadata in various types of local repositories. Dependent on the different requirements of local applications, support for automatic semantic metadata generation and metadata annotation and enrichment is provided. Adaptive content analysis will be demonstrated through a live demo of online genre calculation and rule-based workflow selection services. The results of the content analysis tools are shown as part of the semantic indexing. The search engine makes use of external metadata sources as well as automatically extracted metadata in order to index rich semantic annotations and uses them for search and filtering.

6. REFERENCES

[1] M. Montagnuolo, A. Messina, “Parallel neural networks for multimodal video genre classification”, *Multimedia Tools and Applications*, 41(1), Jan. 2009, pp 125-159.

[2] E. Mbanya, S. Gerke, C. Hentschel, P. Ndjiki-Nya, “Sample Selection, Category Specific Features and Reasoning”, In *CLEF (Notebook Papers/LABs/Workshops)*, 2011.

[3] H. Fassold, S. Wechtitsch, A. Hofmann, W. Bailer, P. Schallauer, R. Borgotallo, A. Messina, M. Liu, P. Ndjiki-Nya and P. Altendorf, “Automated Visual Quality Analysis for Media Production,” in Proc. IEEE ISM, 2012, pp. 385-394.