

LIVE: Semantic-based Multi-Stream Broadcasting of Media Events

Tobias Bürger, Rupert Westenthaler, Christian Eckes, Felix Zielke, Janez Zetelji, and Richard Wages

Abstract—Broadcasting of media events is a real-time action demanding reliable just in time decisions based on the current content of incoming video streams and the availability of background material. Multi-stream broadcasting of this type of event thus demand monitoring of multiple streams and background material. Due to the potentially large amount of streams and other available material, manual monitoring is likely to fail on the long term. We therefore developed an indexing pipeline based on semantic technologies that enables real-time analysis of broadcasted streams and reliable content recommendations of streams and background material based on formal machine understandable descriptions of content. Our approach enables real-time interpretation of broadcasted streams and thus establishes a bridge over the “Semantic Gap” in video analysis.

Index Terms—Interactive TV, real time video analysis, semi-automatic annotation, semantic indexing.

I. INTRODUCTION

The integrated project “LIVE Staging of Media Events” aims at the creation of novel intelligent content production methods and tools for broadcasters to stage live media events in the area of sports, such as the 2008 Olympic Games. In the terminology of the project, “staging live media events” is a notion for the real-time creation of a non-linear multi-stream video show, which changes due to the interests of the consumer (end user). LIVE currently develops a knowledge structure and a framework, including an ontology landscape and a service oriented architecture for an intelligent live content production process. This process amongst others includes semi-automated real-time annotation and dynamic human annotation to realize novel iTV video formats for live media events. The main tasks of the LIVE production system are to support staging professionals and to enable personal tailored broadcasts of huge media events. Problems and questions that arise in this setting are how (i) to manually handle all the incoming information (eg. video streams, archive material, and event information), (ii) to combine parts of this incoming information, and finally (iii) to adapt outgoing information to the preferences of the watching consumer. In the remainder of this paper we first present the scenario of staging and the novel iTV formats that are the basis for broadcasted media events, introduce the Intelligent Media Framework (IMF) and finally describe how

Tobias Bürger and Rupert Westenthaler are with Salzburg Research Forschungsgesellschaft mbH, Salzburg, Austria. Email: {tobias.buerger, rupert.westenthaler}@salzburgresearch.at

Christian Eckes and Felix Zielke are with Fraunhofer IAIS, Sankt Augustin, Germany. EMail: {christian.eckes, felix.zielke}@iais.fraunhofer.de

Janez Zetelji is with the University of Lubljana, Lubljana, Slovenia. EMail: janez.zetelji@ldos.fe.uni-lj.si

Richard Wages is with the Cologne University of Applied Sciences, Cologne, Germany. EMail: richard.wages@fh-koeln.de

real-time (semantic) indexing and content recommendation are enabled through the IMF in LIVE.

II. SCENARIO: LIVE STAGING OF MEDIA EVENTS

The aim of the LIVE project is to support the parallel production of multiple channels presenting different aspects of an event to consumers with different interests and preferences. This process consists of combining and interweaving these streams in real-time to produce a coherent collection of audiovisual (AV) streams which can be received by consumers. What distinguishes our approach from other approaches for interactive TV is, that we do not only want to produce several live AV streams instead of a single one, but in addition will also create transitions (switching points), where the consumer is invited to switch to another sub-channel. Thus our aim is to prevent a mere “zapping” between channels and instead guide the consumer to navigate through the content of a live show of interlinked streams with a coherent overall dramaturgy. The concept of interlinked non-linear multi-channel content formats imposes several conceptual and technical challenges on the development of a real-time staging support system which are outlined in [1].

III. THE INTELLIGENT MEDIA FRAMEWORK

The *Intelligent Media Framework (IMF)* is the core component of the LIVE production system. In the production process it has to accept and handle partial information about particular media items, to add semantic information to the items and to infer and attach contextual knowledge to the items.

The IMF provides services to access the vocabularies and terms of the controlled vocabulary related to the current context of a stream to guarantee the unambiguousness of the terms used, to access information specific to the staged event and it supports knowledge based queries for available content. In addition to these pull-type services, the IMF includes a messaging system to explicitly support the real-time aspect of the staging process, i.e. to trigger meta data generation or announce newly created knowledge in the system.

The IMF and its underlying knowledge structure which models the environment of the LIVE staging domain is outlined in [2].

IV. REAL TIME CONTENT PRODUCTION AND INTERPRETATION

For real-time production the IMF integrates the following components of the *LIVE production support system*:

- 1) The *Intelligent Media Asset Information System*: providing access to services for the storage of media, knowledge models and metadata relevant for the live staging process and providing services for the creation and management and delivery of intelligent media assets.
- 2) The *Metadata Generation System*, dealing with automatic and real-time detection, and extraction of metadata
- 3) The *Human Annotation Tool* allowing manual annotation of the output of (2)
- 4) The *Recommender System*, that recommends content to the professional user based on the staging profile of the show, the personal profile of the video conductor and current and previous consumer feedback.
- 5) The *Video Conducting System*, dealing with supporting the editor and the video conductor by the real-time staging of a live event.

The semantic enrichment process of content in the production chain is based on a semantic indexing pipeline consisting of the components (1) - (3). Based on the output of this pipeline, messages are sent to (4) and (5). (5) sends content requests to (4) which in turn sends recommendations based on incoming messages.

A. Automatic Content Annotation

The Metadata Generation System is the first component in the video indexing pipeline in LIVE. It detects close-ups, shots, faces, camera-motion, color schemes, scenes and artists in the broadcasted streams. The output of the component is later on enriched in a manual step through the human annotation tool and in the IMF that has knowledge about the context of the analyzed media item. Due to the real-time aspects of live media events, knowledge must be available within roughly 20 ms for each video frame in order to avoid congestion of successive information systems. We attack this real-time challenge by using analysis within the compressed MPEG domain, by exploiting Machine Learning Research in cascade-based pattern recognition algorithms and by using efficient algorithms on multi-thread multi core CPU as well as GPU-hardware. (e.g. [4], [5], [6]).

B. Manual Annotation with the Human Annotation Tool

The *Human Annotation Tool* is the second component in the LIVE indexing pipeline, which further enriches the detected information of the *Metadata Generation System*. It can be used to assign terms coming from controlled vocabularies to the low-level information that was extracted in the basic analysis step. These terms are then used by the IMF to attach more semantic information of the current action or event to the media items that is possibly inferred by the current event schedule or other particular information that was detected in the course of this event. The tool therefore offers a controlled vocabulary, a standardized common metadata model, and supports free text annotations. The interface of the Human Annotation Tool includes a collection of clickable buttons, which might include a starter list (*Who?*), adapted action terms for the type of event (*What?*) or geographical places around the events venue (*Where?*). To ease the annotation process, the

user interface which is handled via a touch screen helps the user to be as fast as possible.

C. Semantic Based Content Recommendation

Based upon the enriched information coming from the semantic indexing pipeline, the goal of the *Content Recommender System* is to provide automatic selection of suitable content from the pool of available live or archive content. The content selection procedure is primarily focused on the selection of semantically annotated audio-visual materials from the TV archives or from live sources according to the preferences of the target audience. To realize that, the IMF, provides context information and the controlled vocabulary for the annotation process and propagates the detected *meaning* to the recommender system for the professional user. On the basis of these recommendations, the video conductor decides which live streams and switching possibilities are offered to the consumers [7].

V. CONCLUSION AND FUTURE WORK

In LIVE we implemented a basic semantic indexing pipeline to support the real-time analysis of broadcasted video streams and its semantic enrichment. In the implementation of our first prototype we have shown, that combining several automatic and manual steps is one way towards achieving a reliable solution for video analysis. In future work we will put a special focus on the user interfaces to support the professional user to handle the huge amount of incoming information during a broadcast.

ACKNOWLEDGMENTS

This work was partially funded by the European Commission within the 6th Framework of the IST programme under grant number FP6-27312 "LIVE Staging of Media Events".

REFERENCES

- [1] S. M. Grünvogel, R. Wages, T. Bürger, J. Zaletelj *A Novel System for Interactive Live TV*. Proceedings of the ICEC 2007: Sixth International Conference on Entertainment Computing, 15-17 September 2007, Shanghai, P.R. China, 2007.
- [2] G. Güntner, T. Bürger, R. Westenthaler, and D. Glachs *Basic Specification of the Intelligent Media Framework: Public Synopsis*. LIVE Deliverable D7.4, 2007. Available online at <http://www.ist-live.org>
- [3] T. Bürger and G. Güntner, *Smart Content Factory – Semantic Knowledge Based Indexing of Audiovisual Archives*. Proceedings of the 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies (EWIMT '05), 2005.
- [4] H. A. Rowley, S. Baluja, and T. Kanade *Neural network-based face detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, No. 1, Vol. 20, pp. 23–38, 1998.
- [5] P. Viola, M. J. Jones, and D. Snow *Detecting pedestrians using patterns of motion and appearance*. Proceedings of the 9th International Conference on Computer Vision, pp. 734–741, 2003.
- [6] S. Eickeler, S. Müller, and G. Rigoll *Recognition of JPEG compressed face images based on statistical methods*. Image and Vision Computing, Vol. 18, pp. 279–287, 2000.
- [7] J. Zaletelj, R. Wages, T. Bürger, and S. M. Grünvogel *Content Recommendation System in the Production of Multi-Channel TV Programs*. accepted at Axmedis 2007, November 28-30, Barcelona (Spain) (forthcoming), 2007.