

# The PrestoSpace Metadata Access and Delivery Platform

L. Boch      G. Dimino      A. Messina  
RAI - Centre for Research and Technological Innovation  
Corso Giambone 68, I-10135 Turin, Italy  
{a.messina, l.boch, g.dimino}@rai.it

R. Basili  
Università Roma Torvergata  
Via del Politecnico 1, I-00133 Roma, Italy  
basili@info.uniroma2.it

W. Bailer  
JOANNEUM Research  
Steyrergasse 17, 8010 Graz, Austria  
werner.bailer@joanneum.at

W. Allasia      M. Vigilante  
euriX s.r.l.  
c.so Tortona 17, I-10153 Turin, Italy  
{allasia,vigilante}@eurix.it

Christoph Bauer  
ORF  
Würzburggasse 30, A-1136 Vienna, Austria  
christoph.bauer@orf.at

## 1. Introduction

The PrestoSpace Integrated Project was launched in 2004 under the Information Society Technologies priority of the Sixth Framework Programme of the European Community (IST FP6 507336). Several European broadcasters and audiovisual archive owners, universities and research centres and industry representatives are part of the PrestoSpace consortium. The project website is [www.prestospace.org](http://www.prestospace.org). The objective of the project is to provide technical devices and systems for digital preservation of all types of audio-visual collections, by building up preservation factories providing affordable digitisation, management and distribution services. This demonstration concerns the Metadata Access and Delivery subsystem of the PrestoSpace infrastructure, which aims at collecting descriptive metadata from the numerical analysis of audiovisual material, and providing a search and retrieve interface for archivists to this information [9]. The former task is performed by the Documentation Platform. The latter task is the function of the Publication Platform, a Web application collecting the extracted metadata in an organised and ergonomic way, for a fast and efficient browsing of the documented material. The Publication Platform is the object of the proposed demonstration.

## 2. Summary of the metadata extraction tools

Metadata extraction tools operate in the metadata collection phase. There are two families of tools: content analysis

tools and semantic analysis tools. The PrestoSpace Documentation Platform includes the following audiovisual content analysis tools:

- Shot boundary detection. The shot boundary detection tool segments a video to its primary building blocks, i.e. its shots, and is capable of detecting both abrupt (cuts) and gradual transitions (such as dissolves, fades, wipes, etc.) [1].
- Key frame and stripe image extraction. The key frame detector extracts a number of key frames per shot, depending on the amount of visual change. Stripe images are spatiotemporal representations of the visual essence, created from the content of a fixed or moving column of the visual essence over time.
- Camera motion detection. The camera motion detector analytically describes four basic types of camera motion in the content (pan, tilt, zoom, roll), a rough quantisation of the amount of motion, and the length of the segments in which they appear [2].
- Speech to text transcription. An automatic speech-to-text engine is used, developed by ITC-IRST [5], capable of extracting text from English and Italian spoken content.
- Audio structuring and segmentation. This analysis consists in classifying segments of audio in four principal categories (silence, music, speech, noise).

- Editorial parts segmentation. Editorial parts are the constituent parts of the programme from the editorial point of view. The Documentation Platform uses an automatic editorial segmentation technique in the news domain, choosing a multi-layer approach that merges video and audio information.

The PrestoSpace Documentation Platform includes the following semantic analysis tools:

- Linguistic Processing. The linguistic processing is carried out by a Natural language parser called CHAOS [3][4], which includes several language processing components used to extract semantic entities from text.
- News Categorization. Semantic categories are automatically assigned by a text classifier based on a traditional supervised machine learning model. We used an extended version of the profile-based classifier, known as the Rocchio model.
- Web Alignment. A spidering process is employed to retrieve all the documents published in a target temporal window in the main news websites. This is centred on the broadcasting day by adopting a symmetric span.
- Ontological integration. In MAD, the KIM platform [8] is in charge of making available extensive ontological knowledge about the news domain, and supporting indexing and navigation functionalities. It provides novel Knowledge and Information Management infrastructure and services for automatic semantic annotation, indexing, and retrieval.

### 3. Architecture of the Publication Platform

The extracted metadata are represented by a single XML-based document format, taking the best from each of two metadata standards natively orientated to the description of audiovisual objects, MPEG-7 [7] and P\_META [6]. P\_META was adopted due to its complete set of information structures for the identification, classification and publication-related features of a programme, while MPEG-7 standard was adopted due to its powerful temporal segmentation tools and for its comprehensive set of standard audiovisual descriptors. The rich variety of information extracted by different analysis modules poses several requirements to the Information Retrieval functionalities in the publication phase. The user interface should model access methods according to different (and integrated) capabilities: a) full text search as usually applied by mostly popular search engines; b) Natural Languages Questions; c) Semantic browsing as navigation through concepts, relations and instances of the ontology. The Publication Platform architecture is based on a Web application as user interface,

a DBMS storing the available information related to programmes, and the KIM indexing and search engine. The search interface supports the various retrieval approaches. The user can choose the target of his/her search (e.g. a programme or a news item), which can be filtered by title, broadcast date and service, contributions (e.g. authors, journalists, directors), classification (topics, categories), text of description. The browsing interface is made up of four frames: a video preview, the editorial parts tree, the key frames, and an extensible multi-tab frame, each of which is representing a specific elaboration result. The content of all the frames is synchronised during user interaction.

### 4. Demonstration contents

The Publication Platform has been extensively tested during the project lifetime, through the organisation of structured users workshops in which all tested tools got an overall positive result. The demonstration is based on a collection of several hours of audiovisual material, taken from RAI, BBC and ORF archives. Search and retrieve functionalities and browsing capabilities are illustrated in detail during the demonstration, with queries covering the full set of the available modalities.

### References

- [1] W. Bailer, H. Mayer, H. Neuschmied, W. Haas, M. Lux, and W. Klieber. Content-based video retrieval and summarization using MPEG-7. In *In Proc. of Internet Imaging V*, 2004.
- [2] W. Bailer, P. Schallauer, and G. Thallinger. Joanneum research at TRECVID 2005 - camera motion detection. In *In Proc. of TRECVID Workshop*, 2005.
- [3] R. Basili, M. T. Pazienza, and F. M. Zanzotto. Efficient parsing for information extraction. In *In Proc. of European Conference on Artificial Intelligence (ECAI98)*, 1998.
- [4] R. Basili and F. Zanzotto. Parsing engineering and empirical robustness. *Journal of Language Engineering*, 8(2/3):97–120, 2002.
- [5] F. Brugnara, M. Cettolo, M. Federico, and D. Giuliani. A system for the segmentation and transcription of italian radio news. In *In Proc. of RIAO, Content-Based Multimedia Information Access*, 2000.
- [6] EBU. P\_META 2.0 metadata library. Technical specification, European Broadcasting Union. Metadata Advisory Group, 2007.
- [7] ISO/IEC 15938. Multimedia content description interface. Iso/iec international standard.
- [8] A. Kiryakov, B. Popov, D. Ognyanoff, D. Manov, A. Kiriliov, and M. Goranov. Semantic annotation, indexing, and retrieval. *Elsevier's Journal of Web Semantics*, 2(1), 2005.
- [9] A. Messina, L. Boch, G. Dimino, R. Basili, W. Bailer, P. Schallauer, W. Allasia, M. Vigilante, and M. Groppo. Creating rich metadata in the TV broadcast archives environment: the PrestoSpace project. In *Proc. of AXMEDIS06 Conference*, 2006.