VAnalyzer: a MPEG-7 based Semantic Video Annotation Tool

Florian Stegmaier, Mario Döller, David Coquil, Vanessa El-Khoury
and Harald Kosch

Chair of Distributed Information Systems, University of Passau, Germany
{forename.surname}@uni-passau.de

1 Introduction

Much recent research deals with the automatic extraction of multimedia metadata. While automatic extraction is already well understood for low-level features (e.g., color), it remains an open issue for high-level (semantic) features [1].

In this context, an evaluation of currently available MPEG-7 annotation tools and frameworks has been performed [2]. Most tools can extract low-level features automatically, and some also provide functionalities to describe high-level features manually. This paper highlights VAnalyzer[1], an MPEG-7 based semantic video annotation tool, which is also able to extract certain semantic annotation automatically. Here, we only introduce the features of the VAnalyzer without discussing its architecture. Finally, the demonstration procedure will be highlighted.

2 Main features of VAnalyzer

VAnalyzer is implemented in Java and based on the NetBeans Platform[2], a generic framework for Swing applications. The application offers regular video player functions (e.g., stop or play) and produces valid MPEG-7 metadata descriptions. The raw (decoded) video frames are accessed through the Java Media Framework (JMF)[3] and its plug-in FOBS4JMF[4]. This FFmpeg wrapper plug-in adds a broad support of media formats in addition to the few that JMF can natively process.

Figure 1 outlines VAnalyzer's interface: (a) shows the main menu of the application; (b) illustrates the user interface to (de-) activate or configure the metadata extraction algorithms, completed by a visualisation of the algorithm processing.

The actual content of the extracted metadata depend on the selected algorithms. The following algorithms have been integrated in the VAnalyzer:

- *MPEG-7 low-level features extraction*[5] processes for Color Layout, Color Structure, Dominant Color and Edge Histogram.
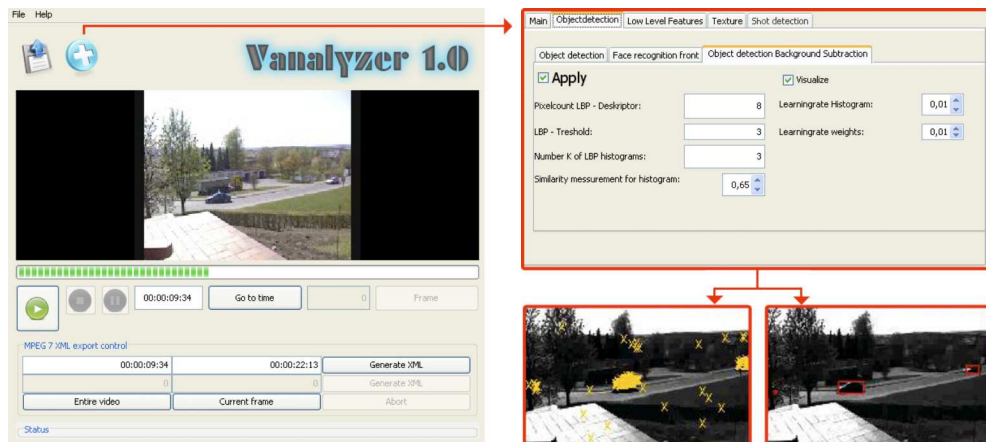
---

[1] http://www.dimis.fim.uni-passau.de/iris/index.php?view=vanalyzer
[2] http://platform.netbeans.org
[3] http://java.sun.com/javase/technologies/desktop/media/jmf
[4] http://fobs.sourceforge.net
[5] Adapted from the Caliph & Emir project: http://www.semanticmetadata.net

(a) Main menu & player functionalities   (b) Algorithm configuration & visualisation

Fig. 1. User interfaces and process visualisation of the VAnalyzer

– Two approaches for *object recognition and tracking*. The first approach makes use of the OpenCV[6] project for object (e.g., faces) detection using Haar cascades. The second follows a modified approach of Heikkilä et al. [3], which detects objects using background substraction and background models.
– A *shot detection* algorithm based on a modified approach described in [4]. Our improvements concentrate on an enhanced dissolve and fade detection through a new one sided feature approach and efficient block matching technique.

3   Demonstration procedure

The demonstration will contain an (optional) talk regarding the architectural facets of the application, a computer-assisted presentation of the central features and the validation of the extracted metadata. Thereby, the used data set will consist of a video surveillance footage.

References

1. J. S. Hare, P. A. S. Sinclair, P. H. Lewis, K. Martinez, P. G. B. Enser, and C. J. Sandom, "Bridging the Semantic Gap in Multimedia Information Retrieval: Top-down and Bottom-up approaches," in *Proceedings of the 1st International Workshop on Semantic Web Annotations for Multimedia*, (Edinburgh, Scotland), 2006.
2. M. Döller and N. Lefin, "Evaluation of available MPEG-7 Annotation Tools," in *Proceedings of I-MEDIA/I-SEMANTICS '07, 1st International Conference on New Media Technology*, (Graz, Austria), September 2007.
3. M. Heikkilä and M. Pietikäinen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects," *IEEE TPAMI*, vol. 28, no. 4, pp. 657–662, 2006.
4. H. Fang, J. Jiang, and Y. Feng, "A Fuzzy Logic Approach for Detection of Video Shot Boundaries," *Pattern Recogn.*, vol. 39, no. 11, pp. 2092–2100, 2006.

---

[6] http://opencv.willowgarage.com