A collaborative web platform for sound archives management and analysis

Thomas Fillon
Parisson
Paris, France
thomas@parisson.com

Guillaume Pellerin
Parisson / IRCAM
Paris, France
guillaume.pellerin@parisson.com

ABSTRACT

In the context of digital sound archives, an innovative web framework for automatic analysis and manual annotation of audio files has been developed. This web framework, is called Timeside and is available under an open-source license.

The TimeSide framework associates an audio processing engine, an audio database, a web API and a client-side multimedia player.

The audio processing engine is written in Python language and has been designed for speech and audio signal analysis and Music Information Retrieval (MIR) tasks. It includes a set of audio analysis plugins and additionally wraps several state-of-the-art audio features extraction libraries to provide automatic annotation, segmentation and Music Information Retrieval analysis. It also provides decoding and encoding methods for most common multimedia formats.

The audio database application is handled through Django (Python) and is interfaced with the audio processing engine.

The web API component provides these functionalities over the web to enable web client to run analysis on the sounds in the audio database. Last but not least, the multimedia player provides an web player associated with several sound and analysis visualizations together with an annotations editor through a multi-tracks display.

The TimeSide platform is available as an open-source project at the following addresses:

TimeSide: https://github.com/Parisson/TimeSide

1. INTRODUCTION

In the context of digital sound archives management, researchers and archivists have expressed the need for computational tools to help them manage and analyze their archives.

For this purpose, an innovative web framework for automatic analysis and manual annotation of audio files has been developed. This web framework, is called Timeside and is developed under an open-source license in a joint collaboration between MIR & Humanities researchers, software developers and archivists.

The TimeSide framework associates four main components : an audio processing engine, an audio database management system, a web API and a client-side multimedia player.

2. AUDIO PROCESSING ENGINE

The audio processing engine is written in Python language and has been designed for speech and audio signal analysis and Music Information Retrieval (MIR) tasks. It includes a set of audio analysis plugins and additionally wraps several state-of-the-art audio features extraction libraries to provide automatic annotation, segmentation and Music Information Retrieval analysis. It also provides decoding and encoding methods for most common multimedia formats. It includes most commonly used python packages for data science and provides convenient Jupyter (ipython) notebook interface for developers and computer science researchers.

The TimeSide engine architecture is composed of several modules and makes it easy to develop and add new plugins.

3. AUDIO DATABASE

The audio database application is handled through Django (Python) and is interfaced with the audio processing engine. It enable to deal with the basic metadata associated with multimedia files and also is used to store the information about audio analysis (state, parameters, versioning). Several analysis can be grouped and organized in experiments and be run over a selection of items (corpus). The management of the audio corpus together with associated metadata available through the same database or linked through another database containing further metadata (e.g. cultural or musical metadata) enables to extract or manage semantic information both from automatic content analysis and from contextual metadata. The database models also provide tools for time-segment annotation of the multimedia items. Both database objects and analysis methods can be managed together through Timeside.

4. WEB API AND MULTIMEDIA PLAYER

The audio engine and the database functionalities are fully accessible through a Web API. This API is associated with an embeddable web multimedia player. This multimedia player is fully compatible with HTML5 and use web audio API components. It can display several sound visualizations such as waveform or time-frequency representations. It can also simultaneously displays results of automatic analysis

Web Audio Conference WAC-2017, August 21–23, 2017, London, UK.

and further enable to tune the analysis parameters. Analysis are displayed in a multitracks fashion together with manual annotations. Indeed, the web player also acts as an annotation editor and gives the possibility for user to collaboratively annotate and share time-segments piece of multimedia files.

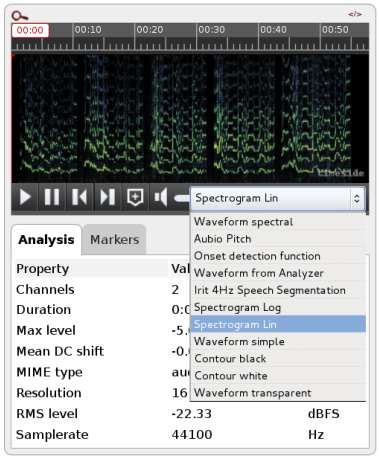

Figure 1: Timeside player with various sound representations.

5. APPLICATIONS IN THE CASE OF ETHNOMUSICOLOGICAL SOUND ARCHIVES

The Timeside framework is a part of a larger Digital assets management framework for sound archives, the Telemeta framework, also available under an open-source license. Telemeta is a web audio platform for the management and the access to digital sound archives and audio metadata. It is developed since 2007 in the context of a collaboration between Parisson, a company, and the Center of Research in EthnoMusicology (CREM) in France.

Telemeta focuses on the enhanced and collaborative user-experience in accessing audio items and their associated metadata and on the possibility for the expert users to further enrich those metadata though hierarchical and structured fields, thesaurus and ontologies. This platform has been deployed since 2011 in the context of ethnomusicological archives and hold the archives of the CREM[1]. The platform is fully operational and is now used on a daily basis by researchers, teachers and archivists in the fields of ethnomusicology, anthropology, linguistics and acoustics.

Telemeta relies on TimeSide for audio decoding, encoding and streaming methods. Also it enables the users of the CREM sound archives to have access to many different automatic audio analysis (e.g. pitch, speech segmentation, music segmentation, polyphony detection, start of recording session for tape recorder, ...).

Through collaboration with academic research labs in computer science, speech processing and music information retrieval, new automatic analysis functionalities are brought to the platform regularly.

[1]http://archives.crem-cnrs.fr

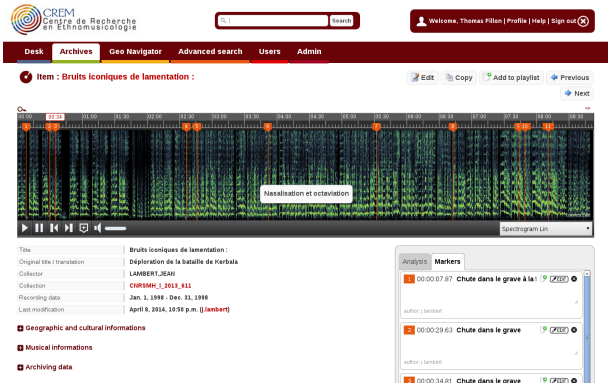

Figure 2: The Telemeta platform deployed for the CNRS-CREM sound archives

The Telemeta and TimeSide platform are available as open-source projects at the following addresses:

- Telemeta: https://github.com/Parisson/Telemeta
- TimeSide: https://github.com/Parisson/TimeSide