

ENABLING ACCESS TO SOUND ARCHIVES THROUGH INTEGRATION, ENRICHMENT AND RETRIEVAL

Ivan Damnjanovic

Queen Mary, University of London
Mile End Road, E1 4NS, London
United Kingdom
ivan.damnjanovic@elec.qmul.ac.uk

Dan Barry

Dublin Institute of Technologies
Kevin Street, Dublin 8
Ireland
dbarry@dit.ie

Josh Reiss

Queen Mary, University of London
Mile End Road, E1 4NS, London
United Kingdom
josh.reiss@elec.qmul.ac.uk

ABSTRACT

Many digital sound archives still suffer from tremendous problems concerning access. Materials are often in different formats, with related media in separate collections, and with non-standard, specialist, incomplete or even erroneous metadata. Thus, the end user is unable to discover the full value of the archived material. EASAIER addresses these issues with the development of an innovative remote access system which extends beyond standard content management and retrieval systems. The EASAIER system has been designed with sound archives, libraries, museums, broadcast archives, and music schools in mind. However, the tools may be used by anyone interested in accessing archived material; amateur or professional, regardless of the material involved. Furthermore, it enriches the access experience enabling the user to experiment with the materials in exciting new ways. The system features; enhanced cross media retrieval functionality, multi-media synchronisation, audio and video processing, analysis and visualisation tools, all combined within in a single user configurable interface.

Index Terms— Sound Archives, Multimedia Retrieval, Music Ontology, marking, looping, time-scaling.

1. INTRODUCTION

EASAIER address several key areas that still lack a deep, systematic, and focused approach: multi and cross- media retrieval, interactivity tools, integration of speech and music processing methods, and systemic archive analysis. In order to cope with these kinds of problems, innovative audio processing, data mining, and visualization techniques, alongside proper user needs and evaluation studies, are being developed and integrated into prototypes. These will be deployed in several sound archives in order to demonstrate a qualitative jump in usability, effectiveness and accessibility. EASAIER aims to improve and implement the separation and representation of sound objects from audio signals by:

- Establishing a *common set of metadata* and provide a mapping for various existing archive ontologies.
- Developing and improving *segmentation and source separation techniques*.
- *Sound object recognition tools* which aim to identify the sources of the sounds in archived recordings.

2. ENHANCED CONTENT RETRIEVAL

The EASAIER system provides multiple online retrieval systems, allowing for searching of content and metadata using multiple techniques and modalities:

- *Speech retrieval* - The development of *speech/music separation and segmentation* technology will reduce the overhead of indexing and will supply metadata on the audio content.
- *Music retrieval* - Music retrieval involves searching and organising audio collections according to their relevance to music-related queries. This process consists of the generation of compact representations for both the query and the collection and the search for similarities between these representations.
- *Cross-media retrieval* – This allows the user to search media in various formats (audio recordings, video recordings, notated scores, images etc...) and find related material across different media. For instance, a search for similar media to a piece of music could result in musically similar pieces as well as relevant text and video linked to the song or performer.

Figure 1 illustrates the query interface for the EASAIER web based client. It provides a convenient method for users to generate textual and musical query combinations

3. FEATURE EXTRACTION

In addition, the EASAIER client enhances the interaction capabilities of an end-user with sound archives through the exploitation of the content within. The client application enables different representations of an audio stream, such as waveforms, spectrograms and several other visualizations

useful for education and analysis. Available metadata is increased by the extraction of various low, mid and high level features which enables greater search functionality. This enables searches according to musical features such as key and time signature, tempo and melody. Additional feature extraction allows for beat tracking, structural segmentation and various other automatic markups which can be used for easy navigation.



Figure 1: EASAIER Query Interface (web client)

Figure 2 shows a screen shot of the client interface with audio content in both the time and frequency domain along with related media such as pictures and text.

4. ENRICHED INTERACTIVITY

Tools to allow time aligned textual markup are provided to the user. Sections of audio can be selected manually or automatically and looped seamlessly for learning or analysis purposes. Advanced audio signal processing tools allow the user to specify how they listen to and interact with the media content. A source separation tool allows a user to listen to individual instruments within the piece of music while a noise reduction tool can be used to eliminate unwanted artefacts.

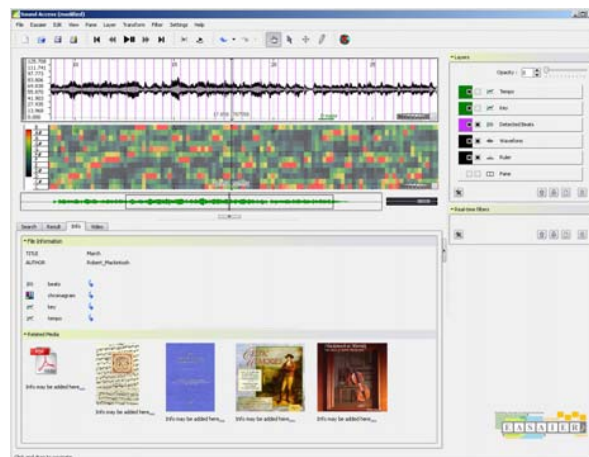


Figure 2: EASAIER Client Interface

Time-stretching allows a user to slow down (or speed up) recordings, without modifying the pitch in real-time. This enables a music student or musicologist for example, to easily learn or analyze a piece of music. The ability to speed up the audio content also gives the user the ability to browse long segments of audio rapidly. The same technology also allows for pitch-shifting of the audio without affecting the time scale. A key innovation also allows the video stream to be synchronised with the audio during time and pitch scaling. It is also possible to zoom the video content, allowing closer inspection of an instrumentalist's particular technique. This suite of enriched access tools is presented in real-time with all functionalities accessible simultaneously.

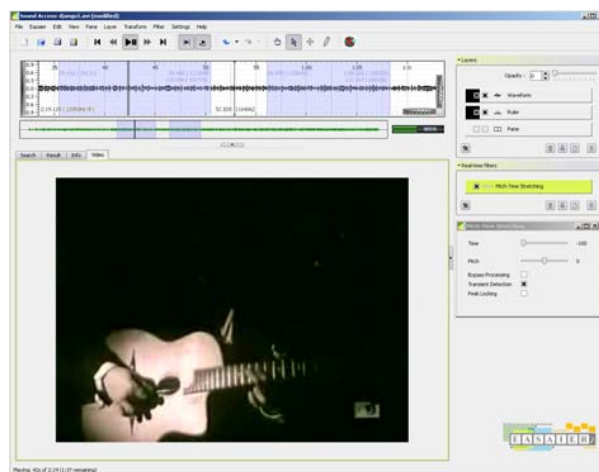


Figure 3: EASAIER Client Interface with video time scaling

ACKNOWLEDGMENTS

This work has been partially supported by the European Community 6th Framework Programme project EASAIER (IST-033902): www.easaier.org.