REPOVIZZ: A MULTIMODAL ON-LINE DATABASE AND BROWSING TOOL FOR MUSIC PERFORMANCE RESEARCH

Oscar Mayor

Music technology Group,
Universitat Pompeu Fabra,
Barcelona, Spain
oscar.mayor@upf.edu

Jordi Llop

Music technology Group,
Universitat Pompeu Fabra,
Barcelona, Spain
jordi.llop@upf.edu

Esteban Maestre

Music technology Group,
Universitat Pompeu Fabra,
Barcelona, Spain
esteban.maestre@upf.edu

ABSTRACT

RepoVizz is a data repository and visualization tool for structured storage and user-friendly browsing of music performance multi-modal recordings. The primary purpose of RepoVizz is to offer means for the scientific community to gain on-line access to a music performance multi-modal database shared among researchers. RepoVizz is designed to hold synchronized streams of heterogeneous data (audio, video, motion capture, physiological measures, extracted descriptors, etc.), annotations, and musical scores. Data streams are stored as single-channel PCM files. The motion-sequence format MJPEG is chosen for video. Annotations are stored as text files. MusicXML is used for musical scores. Data is structured by customizable XML skeleton files enabling meaningful retrieval. Skeleton files are first created during data gathering, and provide means to (re-)organize acquired data in any desired hierarchical structure, as they only hold pointers to stored data files. Once a dataset is created and uploaded to the server, each skeleton file defines a view. Multi-track data visualization is done remotely via a powerful HTML5-based environment that enables web-driven editing (add annotations, extract descriptors) and downloading of datasets. A first instance of RepoVizz has been created within the EU FET-Open Project SIEMPRE, devoted to studying social interaction in ensemble performance.

1. INTRODUCTION

Multi-modal data capturing, fusion and analysis is a trend that is broadly used in the scientific community but despite that data is "easy" to store and analyze using software tools, this large amounts of data are difficult to exchange broadly and the current existing interfaces to access to multimodal databases are very limited, there are no visual solutions to browse the content of a multimodal recording and there are no standards regarding data formats to store the data in such recordings to easily share them with other researchers [1]. The SIEMPRE project (Social Interaction and Entrainment using Music Performances) [2] funded by the European Union under the FET program aims at developing new theoretical frameworks, computational methods and algorithms for the analysis of creative social behavior within small groups of people focusing the research activities on ensemble musical performances. Concrete objectives are models, techniques and algorithms for the extraction of social features based on the analysis of synchronization processes underlying expressive movements, audio and biometric signals during interpersonal creative communication. To achieve these objectives a series of multimodal recordings have to be done during the project to have meaningful data to analyze and a methodology to prepare/perform these recordings as well as a repository to store/share all the data collected and a tool to visualize the recorded data and make annotations in a user friendly and flexible way is being developed and presented in this paper.

2. DB OVERVIEW AND DATA STRUCTURE

The on-line repository is intended to store multimodal recordings of professional musicians (mainly in a duet or string quartet scenario) and share them with the scientific community. The first step before recording a multimodal performance is creating an XML that defines the experiment using the XML creator tool (figure 1) developed especially for this purpose. This tool creates a "Datapack" that consists in a structured XML with the description of a multimodal recording with all the associated data acquisition devices that will be involved and links to the data files that will be generated during the recording containing streams of formatted data.
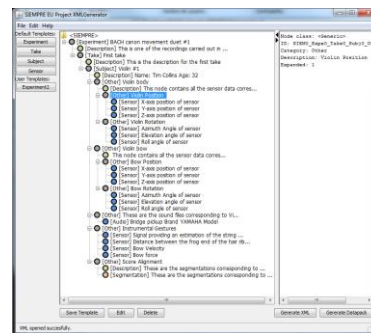

Figure 1. XML creator tool to define a recording setup.

Once a multimodal recording has been done, the upload to the server is restricted to some research groups that have adopted the SIEMPRE protocol for defining the recording setup and for converting all recorded data to make it compatible with the agreed SIEMPRE file formats. After the upload, all data streams of a multimodal recording are stored in the server and are accessed from the client through a front-end web server that connects to a structured MySQL database which contains links to the actual recording data files. The browsing and download of multimodal recordings is open to all the scientific community via a web portal [3] and the recordings' rights are protected under a Creative Commons Non-Commercial license.
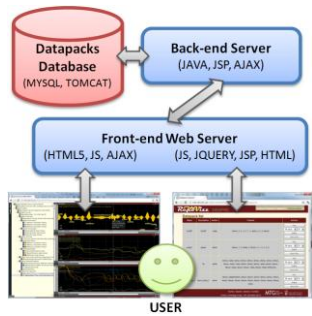

Figure 2. Diagram of the RepoVizz System modules.

3. CLIENT ON-LINE INTERFACE

The Users of the RepoVizz can easily browse, visualize, download, change/edit annotations, extract audio features, change/customize the view, etc. Once a multimodal recording is selected from the database using the client browsing tool (figure 3), then the visualizer tool (figure 4) is executed and allows viewing with detail all associated data acquired during the recording session of the musical performance including position sensors, physiological sensors (EMG, EKG, EMG, EEG), cameras, microphones, MoCap devices, accelerometers and also other extra information like score files, user annotations, derived descriptors, etc.
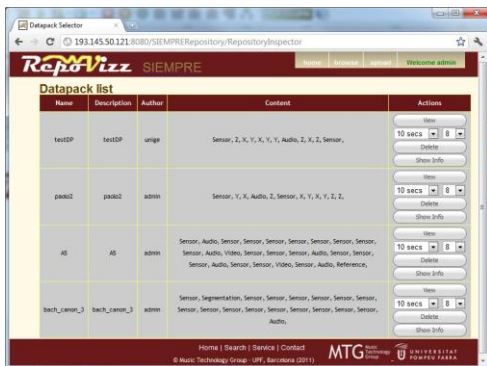

Figure 3. Browsing tool showing available recordings.

The web based visualizer (figure 4) uses modern html5 features for the GUI and allows users to zoom in/out, go forward and backward, select which data streams to view, synchronize between different views, configure the layout/color of each data stream, play audio and video files in sync with other streams, show/edit annotations, etc. All data files remain in the server and are automatically retrieved by the user on demand (avoiding memory and bandwidth problems for large amounts of data) depending on the segment that the user is viewing and the level of zoom for each data stream. The visualizer can also display the XML description of the recording in the left side of the screen, with each node of the XML representing a data stream to help the user select which data streams wants to visualize, just dragging them to the panels in the right side (see figure 4).
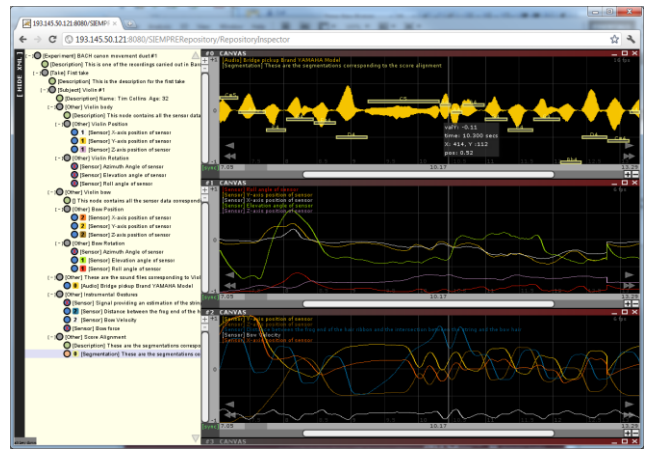

Figure 4. Web based multimodal recording visualizer.

4. CONCLUSIONS

The multimodal on-line database and browsing tool is a very useful and needed tool for sharing and browsing multimodal musical performances and will help many research centers make use of professional recordings that otherwise would not be able to record by themselves because of the need of expensive equipment and the overhead in terms of setup preparation and synchronization between different recording devices.

5. REFERENCES

[1] I. Cossette and M. Wanderley: "The MoCap Data Exchange and the Establishment of a preliminary Database of Music Performances" *workshop organized during the Society for Music Perception and Cognition 2007 conference*, Montreal.

[2] SIEMPRE project. http://www.infomus.org/siempre/

[3] RepoVizz web portal. http://repovizz.upf.edu:8080/