WEBCHM: AN ONLINE TOOL FOR MUSIC ANALYSIS, TRANSCRIPTION AND ANNOTATION

Matevž Pesek, Luka Zakrajšek, Matija Marolt
University of Ljubljana
Faculty of computer
and information science
{*matevz.pesek,matija.marolt*}*@fri.uni-lj.si*
luka.zakrajsek@lgm.fri.uni-lj.si

ABSTRACT

We present the WebCHM — a web based tool for music analysis, transcription and annotation. The tool is based on our compositional hierarchical model (CHM) for music modeling [2] and has several goals: to provide a free and accessible way for researchers to evaluate the CHM on their datasets; to provide a web-based service for MIR-related music analysis; and last to offer a free tool for music transcription, which could enable expansion of the number of annotated recordings usable for a variety of MIR tasks.

1. WEBCHM: THE ANALYSIS

WebCHM is a web-based add-on for our compositional hierarchical model (CHM) for music modeling [2]. It is built on top of our existing software code (written in C#) and provides the following features:

- uploading new audio,
- training new models,
- testing new models,
- model analysis and graphical representation,
- piano roll visualization of model's activations (Fig. 1),
- multiple layers which can be used for comparison, analysis (Fig. 2),
- MIDI overlay and playback (Fig. 2),
- additional layers for human annotation (e.g. grouping activations into notes, adding/erasine events...) with support for adding beat boundaries (Fig. 3)

Like many other services, for example Songle by Goto et al [1], the web interface presents a visualization of the observed music piece. It includes editing features like copying and drawing, and offers multiple layers (in a photoshop-like manner) for comparison of different algorithms and

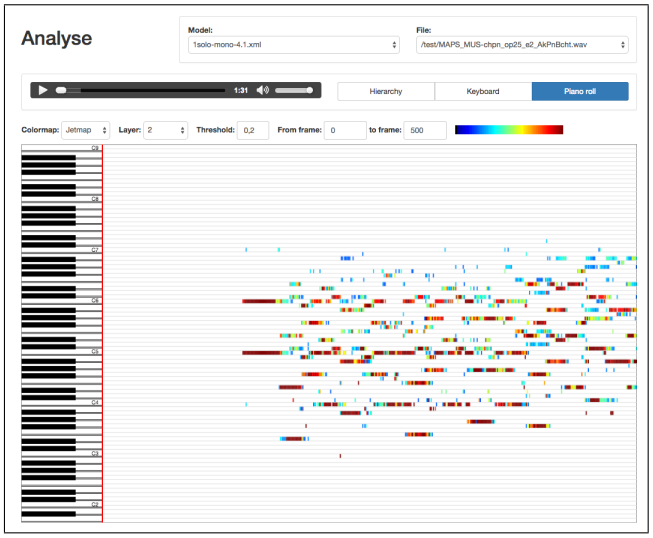

Figure 1. Piano roll visualization with integrate audio player. The type of observed features, threshold and colormap can be dynamically modified.

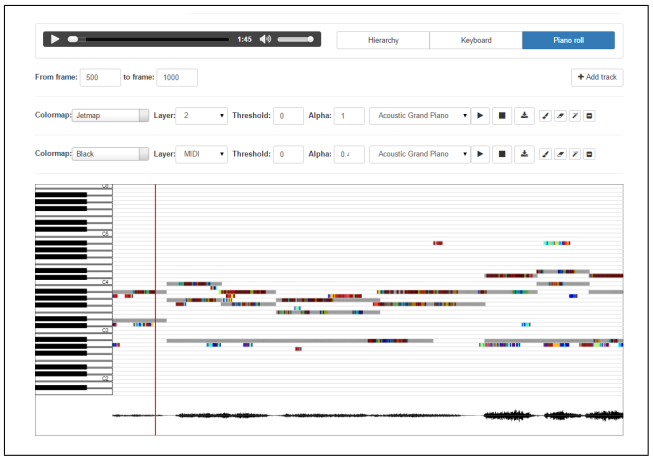

Figure 2. Piano roll visualization with two layers. The first layer displays output of the CHM (jet color map), the second layer the MIDI transcription (transparent black). Each layer can be synthesized independently with different sound fonts.

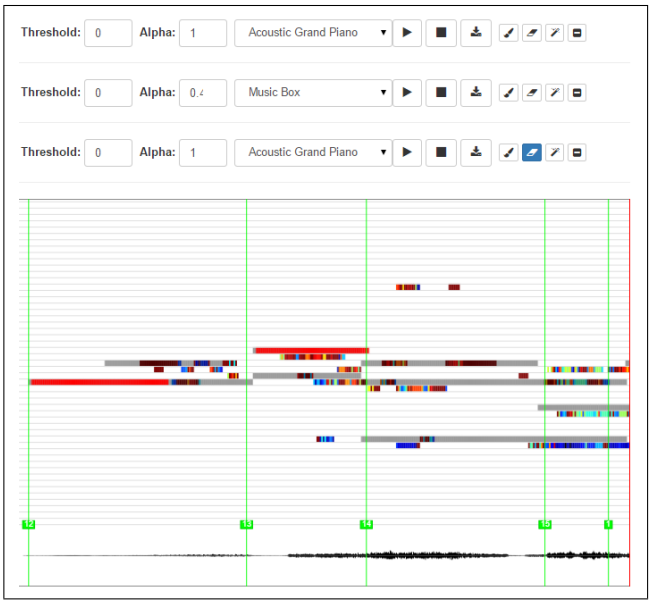

Figure 3. A zoomed-in piano roll visualization with three layers. The first two layers represent the CHM output and MIDI output. The third layer (events of that layer are marked in red) is used for human annotation. The vertical green bars represent beat boundaries. After editing any layer, annotators can export their work to a MIDI file by clicking on the download icon next to any layer (top-right corner).

human annotations with free-draw feature. Output can be exported to MIDI format at any stage of analysis. Each layer can be played back independently with a variety of sound fonts. A suite of other features is also available, for example adjustment of playback speed (by 1% and 5%, up to half / double speed), a magic wand tool for performing fast copying between layers, eraser for simultaneous deletion of events on multiple layers, etc.

The tool can also be used as a sound synthesis tool where parts act as sound generators, thus comprising a virtual electric organ-like instrument, where the interface serves for manipulation. It offers a drawbar-like interface with unlimited amount of drawbars at any desired location, also enabling for inharmonic sounds (depending on the harmonicity of part positions) as seen in Fig. 4. The interface supports any physically connected MIDI device and also enables the use of a virtual keyboard.

The WebCHM tool provides a web service for inclusion of additional annotation algorithms through a REST API. Although currently limited to the CHM, we plan to include other publicly available algorithms into the framework.

2. FUTURE WORK

The current platform offers a flexible tool for music annotation. Its current biggest drawback to large scale deployment is in computing power,as most computation takes place on the backend. We are planning on partially rewriting the model to work fully in the browser (JavaScript), thus eliminating the need for powerful and expensive hard-

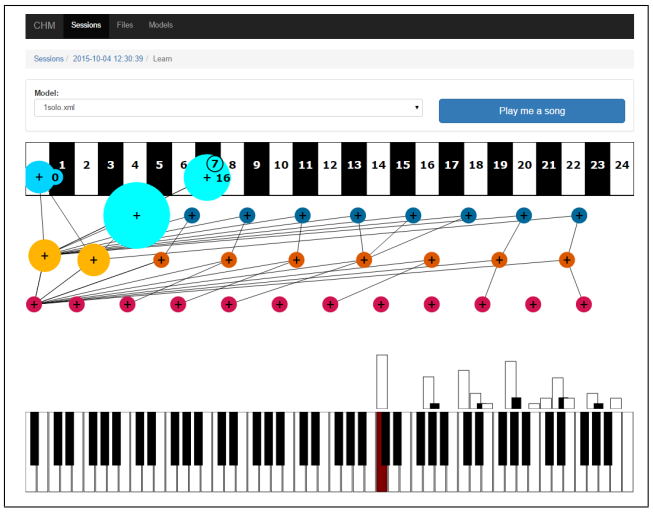

Figure 4. WebCHM as a sound generator. The keys pressed a physical keyboard are displayed in red. The histogram representation marks the cumulative magnitudes of frequencies of the synthesized sound at corresponding positions. The size of each CHM part represents the modifiable magnitude.

ware on the backend.

We are currently developing a new interface that will enable researchers to upload their own algorithms and datasets and evaluate their algorithm, as well as compare them to others in a single procedure. We aim to grow the service into a web-based repository of results, with a growing number of available algorithms. Future development includes export of annotations in MusicXML format and a web-based sheet music rendering API to eliminate the need for additional desktop applications.

3. REFERENCES

[1] Masataka Goto, Jun Ogata, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano. PodCastle and songle. In *Proceedings of the ACM multimedia 2012 workshop on Crowdsourcing for multimedia - CrowdMM '12*, page 1, New York, New York, USA, October 2012. ACM Press.

[2] Matevž Pesek, Aleš Leonardis, and Matija Marolt. A compositional hierarchical model for music information retrieval. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 131–136, Taipei, 2014.