MCIpa

Geoffroy.Peeters@ircam.fr, David Fenech, Xavier Rodet
IRCAM – CNRS STMS
M C I p a
a Music Content Information Player and Annotator for discovering music

Geoffroy.Peeters@ircam.fr, David Fenech, Xavier Rodet
IRCAM – CNRS STMS
Introduction

- Many researches on
  - query over large databases
    - by specific music characteristics (melody, genre, mood, tags, chords, ...)
    - by-similarity
- Few works address the problem of using content-information
  - to guide the user during its listening of a track
  - to allow the user have a better understanding of the music content of a track
- Many researches rely on training data or validation data
  - Tools that allow annotating music audio files in terms of specific music characteristics are still missing
- This work:
  - propose a tool for music content visualization and music content annotation
Related works

Existing media players

- iTunes
- Windows Media Player
- RealPlayer
- WinAmp
- ...
Starting point

- Peeters 2002 Music Structure
  - Matlab Interface Prototype
    - [Peeters ISMIR2002, Springer2005]

- Ludovich Gaillard
  - Flash version

- Samuel Goldsmidth
  - Extension of the flash interface
    - Hierarchical structure representation
    - Semantic HIFI remote-controller
    - [Boutard, Goldsmidth, Peeters LSAS2007]

- Semantic HIFI user-testing
  - Users found the interface “interesting, useful and innovating”
  - Some weak point of the interface:
    - No possibility to assign a label to each block of the structure
    - No possibility to exchange annotations among user
    - No possibility to assign a color to each block
# Related works

## Existing media players

- iTunes
- Windows Media Player
- RealPlayer
- WinAmp
- …

## Existing audio annotation tools

- AS Annotation
- Sonic Visualizer
- CLAM annotator / MUSOCA
- Wavesurfer
- Praat
- Acoustmographe
- Transcriber
- Audacity
- …

- Ircam
- QMUL
- IUA-UPF
- KTH
- IPS
- GRM
- DGA
Related works

- AS (Audiosculpt) Annotation
  - IRCAM
  - MacOSX
    - [http://recherche.ircam.fr/equipes/analyse-synthese/ASAnnotation/](http://recherche.ircam.fr/equipes/analyse-synthese/ASAnnotation/)
  - Annotation over the visualization of the spectrogram (note-gram) using markers (midi-notes), integrated algorithms
Related works

- Sonic Visualizer
  - Queen Mary University of London
  - Linux/ MacOSX / Windows/ Open Source
    - [http://www.sonicvisualiser.org/index.html](http://www.sonicvisualiser.org/index.html)
  - Annotation over the waveform/ spectrogram, analysis plugins (Vamp)/ Effect plugin (LADSPA/DSSI), visual masks
Related works

- CLAM annotator / MUSOCA
  - IUA-UPF
  - Linux/ MacOSX/ Windows/ OpenSource
    - [http://clam.iua.upf.edu/](http://clam.iua.upf.edu/)
  - Framework for developing graphical interfaces and signal analysis algorithms;
  - Online system for collaborative global annotation
Related works

• Wavesurfer
  - KTH
  - Linux/ MacOSX/ Windows/ OpenSource
    • http://www.speech.kth.se/wavesurfer/
  - Annotation over waveform/ spectrogram/ pitch/ energy, some plug-ins available for analysis (Video/ Beat/ ...), browsing by content (marker lists)
Related works

- Praat
  - IPS (Institut Phonétique d’Amsterdam)
  - Linux/ MacOSX/ Windows/ OpenSource
    - http://www.fon.hum.uva.nl/praat/
  - Many signal analysis algorithms but mostly dedicated to speech
Related works

- Acousmographe
  - GRM
  - Windows XP
  - Annotation using a wide-range of graphical tools(shape, color, masks)
Related works

- Transcriber
  - DGA
  - Linux/ MacOSX/ Linux/ Open Source
  - Mostly dedicated to the speech-to-text annotation, browsing by markers
Related works

- Audacity
  - Linux/ MacOSX/ Windows/ OpenSource
  - http://audacity.sourceforge.net/
  - Good audio and midi synchronization but no specific tools for annotation
Analysis of related works

- Currently no tools dedicated specifically to the annotation of music in terms of music content (structure, chords, ...)
  - Most of the tool are made for the annotation of generic audio (or speech) over a signal representation (signal, spectrogram)

- Our point of view: annotation of music content is facilitated by visualization of music content
  - Ground visualization of our tool: a similarity matrix

- Good functionalities in existing software
  - Visual masking and transparency system
  - Use of specific color and shape for each type of annotations
  - Separation between the graphical interface and the content-extraction tools
  - Possibly to quickly browse the file by annotation
**MCIpa requirements**

**Generic requirements**
- The tool must be **easy to use**, understand and install
- The **same tool** is used for both annotation and visualization
- Automatic extraction is **not part of the tool**, use your favorite one, or take an existing annotation
  - communication using a set of xml files
- The tool must be **cross-platform**
- The interface should read and play directly the most used music formats, **mp3 files**

**Interface requirements**
- Interface must be **intuitive**.
  - Main paradigm= click and listen to what you have clicked on”
- Action must be **quick** and quickly reachable
  - Keyboard shortcuts are used extensively in order to do that
- It should be easy for the user to **navigate** in the representation and to change the type of representation being displayed
MCIpa: two parts

- Music database browser:
  - Search in a database by
    - Artist
    - Title
    - Music genre
    - Music mood
    - Similarity

- Music track browser
  - Represents graphically the content of a given track
  - Allows navigation inside the track

Query database

Search by similarity

Search results
Music track browser

- Loading, playing functionalities
- Editorial metadata (ID3)
- Time-line, zoom-in/out, auto-scroll, playhead markers

![Diagram of Music track browser features]

- Playing functionalities
- ID3 informations
- Playhead
- Time-line
## Music track browser

**Representation of the content**

- **Music Content Information (MCI) objects**
  - Describes a specific music content
  - Has a distinct visual representation
  - Has a distinct feedback (audio, play-head position)
  - Has a distinct xml representation
  - Has an existing automatic extraction tool (initialize annotation)

<table>
<thead>
<tr>
<th>MCI</th>
<th>Graphical representation</th>
<th>User Interaction / Interface feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity Matrix</td>
<td>As a 2D image on background</td>
<td>click anywhere inside the image starts playing at the given position</td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Similarity Matrix" /></td>
<td>------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Music Structure      | As a part-roll (each type of part is represented on a specific line) | - choose the number of parts used for the subdivision  
- click inside a part starts playing at the part beginning  
- forward-backward by parts          |
|                      | ![Music Structure](image)                                      |------------------------------------------------------------------------------------------------------|
| Chord progression    | As a set of TABs with chord labels                             | click inside a chord 1) starts playing at the chord beginning 2) plays the corresponding chord prototype |
|                      | ![Chord progression](image)                                    |------------------------------------------------------------------------------------------------------|
| Downbeat/beat positions | As a set of vertical lines (thick lines for downbeats, thin lines for beats) | Audio click when the play-head crosses a beat marker                                                 |
|                      | ![Downbeat/beat positions](image)                              |------------------------------------------------------------------------------------------------------|
| Multi-pitch          | As a piano-roll (each note-stream is represented by a specific color) | - choose the displayed note channels  
- click inside a note plays the corresponding note prototype                                           |
|                      | ![Multi-pitch](image)                                           |------------------------------------------------------------------------------------------------------|
| Sound-events         | As a sound-event-roll (each type of sound-event is represented on a specific line) | not yet                                                                                               |
|                      | ![Sound-events](image)                                         |------------------------------------------------------------------------------------------------------|
Music track browser

- Structure
- Structure accuracy
- Similarity Matrix
Music track browser

- Chords
- Drum events
- Measure/Beat
- Multi-Pitches
- Zoom in time line
Annotation

- Enter or correct the displayed content descriptions
  - Annotation is greatly facilitated by the knowledge (visualization) of other types of content description
- Marker system, generate markers
  - on the fly
  - at the mouse position
  - at the beginning of a specific MCI
  - corresponding to a specific marker-type
- Move, Add, Delete markers, edit marker labels
- Generate the corresponding XML code
- Reload the edited description
Overall architecture and file formats

- Reads directly mp3 file
- Descriptions: set of XML files
- Top XML file with pointer to
  - Mp3 file
  - Various XML description files
  - Simple and light XML format (not MPEG-7 based)
Overall architecture and file formats

FILEs
- similarity matrix .sdif
- structure .xml
- beat .xml
- chord.xml
- note.xml
- instrument.xml
- drum.xml

AUTOMATIC estimation
- ircamstructure
- ircambeat
- ircamchord
- ircammultipitch
- ircaminstrument
- ENSTdrum

mp3 file → top xml file

Database (MYSQL)

Search by similarity

PHP

Query by artist
Query by title
Query by genre

<?xml version="1.0" encoding="UTF-8"?>
<mcipa generator="*" version="*" date="*" />
<media>Tracy.mp3</media>
<similaritydescriptionfilemat>Tracy_sim.jpg</similaritydescriptionfilemat>
<structuredescriptionfile>Tracy_struct.xml</structuredescriptionfile>
<beatedescriptionfile>Tracy_beat.xml</beatedescriptionfile>
<chorddescriptionfile>Tracy_chord.xml</chorddescriptionfile>
<invariantdescriptionfile>Tracy_instrument.xml</invariantdescriptionfile>
<drumdescriptionfile>Tracy_drum.xml</drumdescriptionfile>

</mcipadescription>
Overall architecture and file formats

< ?xml version="1.0" encoding="UTF-8" ? >
< mcipadescription generator="*" version="*" date="*" >
  < media>Tracy.mp3</media >
  < description type="structure" nbstate="2" dictionary="statenum" id="*" generator="*" version="*" date="*" >
    < segment start="1.784" stop="20.3861" label="1" id="*" />
    < segment start="20.9674" stop="38.9882" label="2" id="*" />
    < segment start="39.5695" stop="48.2892" label="1" id="*" />
  </description>
  < description type="structure" nbstate="3" >
    < segment start="1.784" stop="20.3861" label="3" id="*" />
    < segment start="20.9674" stop="38.9882" label="2" id="*" />
    < segment start="39.5695" stop="48.2892" label="1" id="*" />
  </description>
</mcipadescription>
Overall architecture and file formats

< ?xml version="1.0" encoding="UTF-8" ? >
< mcipadescription generator="*" version="*" date="*" >
  < media>Tracy.mp3</media>
  < description type="beat" dictionary="beatnum" id="*" generator="*" version="*" date="*" >
    < marker start="0.45855" label="1" id="*" />
    < marker start="1.067100" label="2" id="*" />
    < marker start="1.675720" label="3" id="*" />
    < marker start="1.675720" label="4" id="*" />
    < marker start="1.675720" label="1" id="*" />
  </description>
</mcipadescription>
Overall architecture and file formats

**FILES**
- similarity matrix .sdif
- structure .xml
- beat .xml
- chord .xml
- note .xml
- instrument .xml
- drum .xml

**AUTOMATIC ESTIMATION**
- ircamstructure
- ircambeat
- ircamchord
- ircammultipitch
- ircaminstrument
- ENSTdrum

*mp3 file* → *top xml file*

```xml
<?xml version="1.0" encoding="UTF-8" ?>
<mcipadescription generator="*" version="*" date="*" >
  <media>Tracy.mp3</media>
  <description type="chord" dictionary="free" id="*" generator="*" version="*" date="*" >
    <segment start="0.0000000" stop="1.6671280" label="C" id="*" />
    <segment start="1.6671280" stop="2.6143598" label="CMaj" id="*" />
    <segment start="2.6143598" stop="4.1083732" label="Gmin" id="*" />
  </description>
</mcipadescription>
```
Overall architecture and file formats

FILES
- similarity matrix .sdif
- structure .xml
- beat .xml
- chord .xml
- note .xml
- instrument .xml
- drum .xml

AUTOMATIC ESTIMATION
- ircamstructure
- ircambeat
- ircamchord
- ircammultipitch
- ircaminstrument
- ENSTdrum

< ?xml version="1.0" encoding="UTF-8" ? >
< mcipadescription generator="*" version="*" date="*" >
  < media>Tracy.mp3</media >
  < description type="note" dictionary="midinote" id="*" generator="*" version="*" date="*" >
    < segment start="0.835828" stop="0.928707" label="37" track="1" velocity="76" id="*" />
    < segment start="0.742948" stop="0.928707" label="53" track="1" velocity="51" id="*" />
    < segment start="0.835828" stop="0.928707" label="56" track="2" velocity="77" id="*" />
    < segment start="1.114467" stop="1.207347" label="33" track="2" velocity="90" id="*" />
  </description>
</mcipadescription>
Overall architecture and file formats

< ?xml version="1.0" encoding="UTF-8"?>
< mcipadescription generator="*" version="*" date="*" >
< media>Tracy.mp3</media >
< description type="instrument" dictionary="free" id="*" generator="*" version="*" date="*" >
< segment start="0.522426" stop="1.044875" label="po" id="*" />
< segment start="0.928776" stop="1.277075" label="vc" id="*" />
< segment start="1.149365" stop="2.089773" label="vc" id="*" />
< segment start="2.252313" stop="3.285601" label="vc" id="*" />
< segment start="2.832812" stop="4.504649" label="po" id="*" />
< /description >
< /mcipadescription >
< ?xml version="1.0" encoding="UTF-8"?>
< mcipadescription generator="*" version="*" date="*" >
< media>Tracy.mp3</media >
< description type="drum" dictionary="free" id="*" generator="*" version="*" date="*" >
< marker start="0.189" label="bd" id="*" />
< marker start="0.189" label="hh" id="*" />
< marker start="0.793" label="sd" id="*" />
< /description >
< /mcipadescription >
Development

- Graphical interface
  - Flash 9,
  - Action Script 3
  - Can be run on any platform supporting the Adobe Flash 9 plugin (Linux, Windows, Mac OS-X, portable devices, ....)
  - Standalone version for Windows and Mac OS-X

- Database management
  - PHP scripts to a MySQL database return a list of pointer to top-xml files
  - XAMPP (Apache, mySQL, PHP)
Conclusion

• Future works
  - Paradigm used by MCIpa has been partially tested during the user-testings of the Semantic HIFI system
    - Test the paradigm for annotation purposes
      - Establish an experimental protocol: choice of a set of annotation tasks, set of music items
    - Extending the current architecture to a plug-in architecture in order to be able to display new types of MCI

• MCIpa usages
  - Visual representations can be understand by a large number of people
  - Standard media player,
  - musical education,
  - comparative musicology (cnfr Ircam mediatheque),
  - musician practicing tool (playing over Aebersold records),
  - research purposes (quick visualization of results)
  - annotation

• Available at
  - http://recherche.ircam.fr/equipes/analyse-synthese/peeters/mcipa/