A comprehensive framework for audio and music analysis

Olivier Lartillot
Aalborg University, Denmark

MIRtoolbox

MiningSuite
Outline

- Audio and musical feature extraction from recordings
  - MIRtoolbox

- Music analysis from symbolic representation:
  - grouping, reduction, motivic patterns

- MiningSuite
Comprehensive set of audio / musical features

Modular syntactic layer on top of Matlab

- both simple and powerful language

Referential framework in Music Information Retrieval

- 10000s of downloads, 500+ citations in papers
Dynamics: Global energy

Dvorak, Symphony 8 G major, Allegretto grazioso - Molto vivace
Timbre

- Energy decomposed into frequencies (spectrum)

Energy in close frequency range: rough sound

Harmonic sound: regular frequency series vs. noise

Low frequency energy: dark sound

High frequency energy: bright sound
Timbral gesture: brightness

Beethoven, 9th Symphony, Scherzo

Beethoven, 7th Symphony, Allegretto

Harty, Miniature for oboe and piano, Orientale

Pitch gesture

Temporal location of events (in s.)


Harty, Miniature for oboe and piano, Orientale
MIRtoolbox

“Symbolic” level

Audio level

Notes

Dynamics

Timbre

Pitch

Sound
Dynamics: Note attacks

CPE Bach, Cello Concerto in A major, WQ172, 3rd movement
“Symbolic” level

Audio level

MIRtoolbox

Metre

Notes

Pitch

Timbre

Sound

Dynamics

Audio level

“Symbolic” level

MIRtoolbox

Metre

Notes

Pitch

Timbre

Sound

Dynamics

Audio level

“Symbolic” level

MIRtoolbox

Metre

Notes

Pitch

Timbre

Sound

Dynamics

Audio level

“Symbolic” level
Tempo estimation

- `o = mironsets('mysong', 'Detect', 'No')`
- `do = mironsets(o, 'Diff')`
- `ac = mirautocor(do)`
- `pa = mirpeaks(ac, 'Total', 1)`

In short:
- `[t, pa] = mirtempo('mysong')`
- `t = 129.6333` bpm
Tempo estimation

- $o = \text{mironsets}('mysong', 'Detect', 'No')$
- $do = \text{mironsets}(o, 'Diff')$
- $f = \text{mirframe}(do)$
- $ac = \text{mirautocor}(do)$
- $pa = \text{mirpeaks}(ac, 'Total', 1)$

In short:
- $[t, pa] = \text{mirtempo}('mysong', 'Frame')$
MIRtoolbox

"Symbolic" level

Mode

Notes

Dynamics

Timbre

Pitch

Audio level

Sound
* Sykes, Glowinski, Grandjean, Lartillot, Eliard. Is a contemporary listener able to distinguish between the musical emotional figures created by Monteverdi? SysMus 2013

Tonal gesture

Monteverdi, *Hor che'l ciel e la terra*, 1st part *

Beethoven, 9th Symphony, Scherzo

Tiersen, *Comptine d'un autre été : L’après-midi*

Schönberg, *Verklärte Nacht*, Sehr Ruhig
Timbral structure


Brahms, Symphony No.3 in F major, Poco Allegretto
Tonal structure Bazzini, *Dance of the Goblins*

Similarity matrix

Sequential repetition

Novelty curve
Groups
Segments

“Symbolic” level

Dynamics
Notes
Timbre
Pitch

Audio level

Sound

Patterns

MIRtoolbox
MiningSuite

Mode
Metre
Motifs

Groups
Segments

“Symbolic” level

Dynamics
Notes
Timbre
Pitch

Audio level

Sound

Patterns

MIRtoolbox
MiningSuite
Local grouping

Mozart, Variation XI on “Ah, vous dirai-je maman”, K.265/300e
Ornamentation reduction

Mozart, Variation XI on “Ah, vous dirai-je maman”, K.265/300e

- syntagmatic network
L1

M1

U1

L2

U2

M2

U3

L3

J.S. Bach, *Well-Tempered Clavier*, Book II, Fugue XX
Detected subject entries

Lartillot, ISMIR 2014
MiningSuite

code.google.com/p/miningsuite

SIGMINR
signal processing

SEQMINR
sequence processing

AUDMINR
auditory modelling

PATMINR
pattern mining

MUSMINR
music analysis
• New syntactic layer in the source code itself
  • Matlab optimisation & code readability
• Significantly optimised (in speed and memory)
  Fully rewritten using recent Matlab object-oriented programming capabilities
• Open source. Developers’ community
• Tutorial at ISMIR 2014 conference