upf. Universitat Pompeu Fabra Barcelona

MTG Music Technology Group

Music Feature Analysis with ESSENTIA

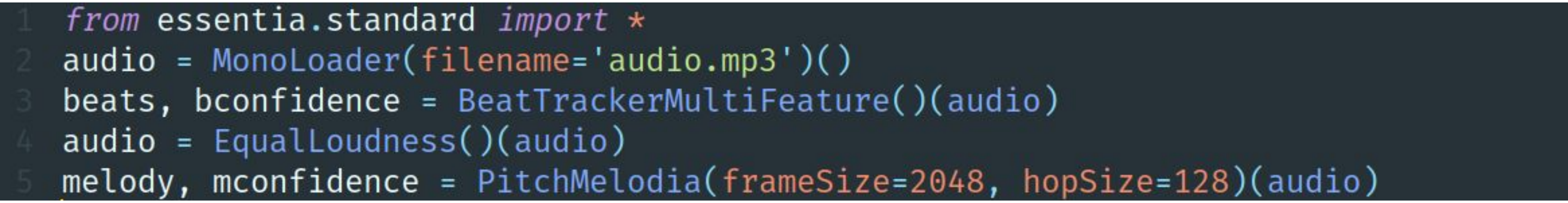
Dmitry Bogdanov

23.05.2019

Essentia http://essentia.upf.edu

Open-source library and tools for audio and music analysis, description and synthesis

- Extensive collection of reusable algorithms
- Written in C++ and optimized for computational speed
- Python bindings for fast prototyping
- Feature extractors for **large-scale audio analysis**
- **Cross-platform** (Linux, Mac OS X, Windows, iOS, Android, and JavaScript)
- Support for mobile platforms and **real-time** processing



```python
from essentia.standard import *
audio = MonoLoader(filename='audio.mp3')()
beats, bconfidence = BeatTrackerMultiFeature()(audio)
audio = EqualLoudness()(audio)
melody, mconfidence = PitchMelodia(frameSize=2048, hopSize=128)(audio)
```

Bogdanov, D., Wack N., Gómez E., Gulati S., Herrera P., Mayor O., et al. (2013). ESSENTIA: an Audio Analysis Library for Music Information Retrieval. International Society for Music Information Retrieval Conference (ISMIR'13). 493-498.

Essentia

License: Affero GPLv3 + available as a commercial license

Documentation: http://essentia.upf.edu/documentation

Code, issues, feature requests: http://github.com/MTG/essentia

Blog: https://mtg.github.io/essentia-labs/

Email: dmitry.bogdanov@upf.edu

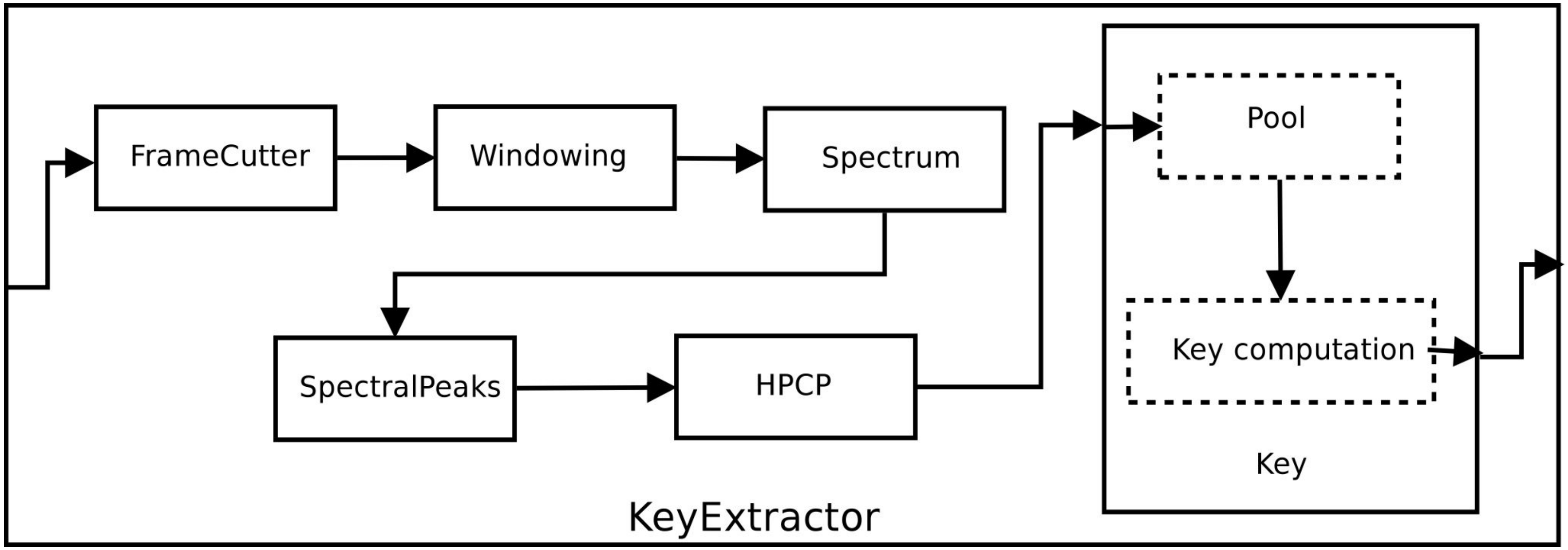
Essentia

Over 200 algorithms for audio signal processing and analysis, and sound and music description, developed at MTG

- Standard audio IO & DSP
- Sound and music descriptors
  - spectral features
  - rhythm and tempo
  - tonality, pitch and melody
  - loudness/dynamics
  - sound envelope
  - audio segmentation
  - fingerprinting
- Machine-learning based descriptors
  - genres, moods, instrumentation, …
  - SVM classifiers
  - TensorFlow deep learning models

Extractors

Users can build their own extractors for the descriptors they want to compute in a “data-flow” manner



How it compares?

- Many algorithms (http://essentia.upf.edu/documentation/algorithms_reference.html)
- Semantically meaningful descriptors
- Successfully used in academia and industry
- Open-source
- Cross-platform (Linux, OSX, Win, iOS, Android, JavaScript)
- Optimized for computational time
- Real-time applications

Applications

http://essentia.upf.edu/documentation/applications.html

- Visualization, interaction and music creation
- Sound and music indexing
- Music classification and auto-tagging
    - Genre, mood, rhythm, instrumentation
- Music similarity, recommendation, playlist generation
- Music identification
- Real-time music audio analysis
- Digital audio instruments
- Audio analysis for games
- Audio quality analysis

Over 100 academic papers are using Essentia

http://essentia.upf.edu/documentation/research_papers.html

Industrial applications

freesound

Collaborative database of Creative Commons Licensed sounds

AcousticBrainz

Open platform for gathering music information from audio

KORG

Interactive iPhone tuner app for learning to play music instruments

CRYPT OF THE NECRODANCER

Award winning independent rhythm game using Essentia's beat detection

fretello

An iOS/Android app that helps you learn guitar

reactable

Cutting-edge interactive products for musical creation

bmat MUSIC INNOVATORS

Worldwide music identification and monitoring service using fingerprinting technology

MusicMuni Labs

Technologies for music education and infotainment

TELE MUSIC

50 years of library music for movies, adverts, TV and radio

La Cupula Music

Tech solutions for digital music distribution, catalogue management and royalties collection

WAVES

Audio plugins for mixing, mastering, and recording

Tunebat

Key & BPM database and music finder

Industrial R&D collaborations

Flits Music GmbH

Solutions for live concert audio stream identification.

- Robust **cover song identification**
- Music fingerprinting

La Cúpula

Solutions for the music distribution industry

- Automatic **audio quality** control

La Cupula
Music

Alonso-Jiménez,, P., Joglar-Ongay, L., Serra X., & Bogdanov D. Automatic Detection of Audio Problems for Quality Control in Digital Music Distribution. AES 146th International Convention, 2019.

Use-cases

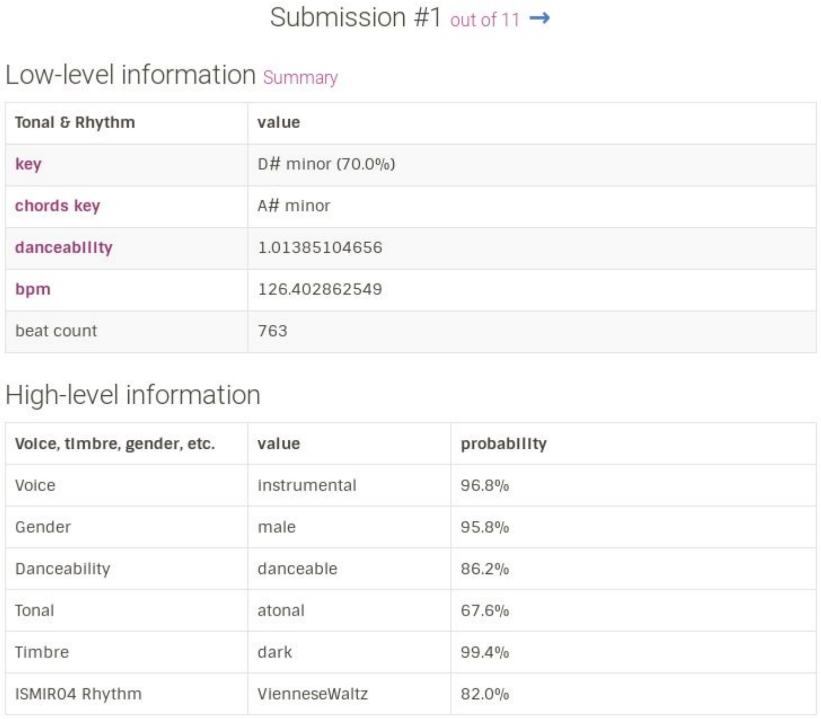
- Research and rapid prototyping
  - Python extension, Vamp plugin for visualization
- Real-time analysis
  - EssentiaRT~: external for PureData and Max/MSP
  - Cortosia: interactive tuner iPhone app
- Offline (large-scale) analysis
  - Optimized C++ extractors
  - AcousticBrainz (over 11 million analyzed tracks)
  - Freesound
- Embedded systems
- Audio synthesis, digital audio instruments

Large-scale music analysis with AcousticBrainz

A community platform for gathering music information obtained from audio

https://acousticbrainz.org

- Open data computed by open algorithms (Essentia)
- Built on **submissions from the community**
- Over 11 million analyzed tracks
- ~3,000 statistical **music features**
- High-level annotations (genres, moods, instrumentation)
- **Linked to metadata** in MusicBrainz
- Working on adding frame data for deep learning models



Submission #1 out of 11 →

Low-level information Summary

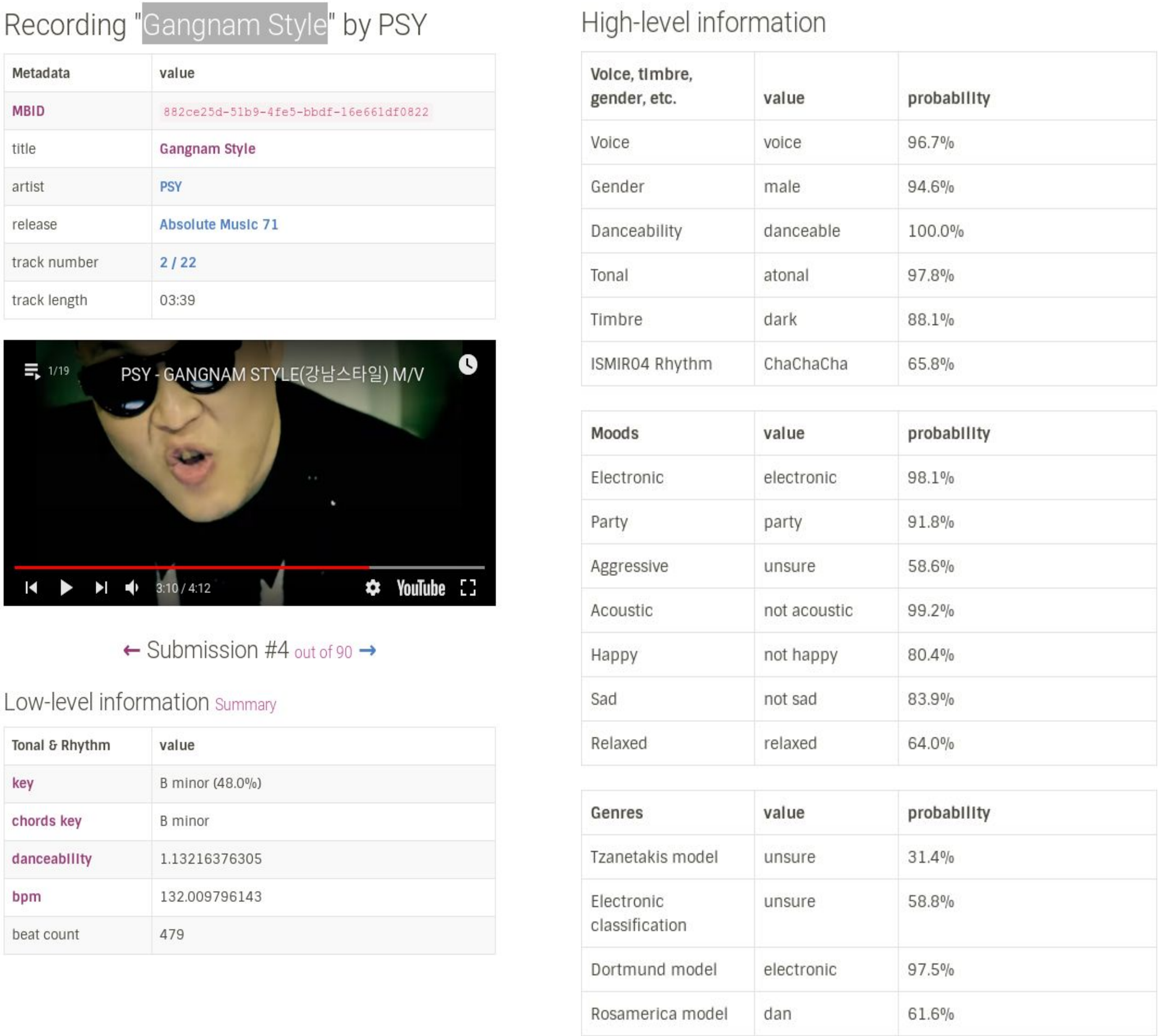
Tonal & Rhythm	value
key	D# minor (70.0%)
chords key	A# minor
danceability	1.01385104656
bpm	126.402862549
beat count	763

High-level information

Voice, timbre, gender, etc.	value	probability
Voice	instrumental	96.8%
Gender	male	95.8%
Danceability	danceable	86.2%
Tonal	atonal	67.6%
Timbre	dark	99.4%
ISMIR04 Rhythm	VienneseWaltz	82.0%

Porter, A., Bogdanov D., Kaye R., Tsukanov R., & Serra X. (2015). AcousticBrainz: a community platform for gathering music information obtained from audio. 16th International Society for Music Information Retrieval Conference (ISMIR 2015). 786-792.

http://acousticbrainz.org/882ce25d-51b9-4fe5-bbdf-16e661df0822?n=3



Recording "Gangnam Style" by PSY

Metadata	value
MBID	882ce25d-51b9-4fe5-bbdf-16e661df0822
title	Gangnam Style
artist	PSY
release	Absolute Music 71
track number	2 / 22
track length	03:39

← Submission #4 out of 90 →

Low-level information Summary

Tonal & Rhythm	value
key	B minor (48.0%)
chords key	B minor
danceability	1.13216376305
bpm	132.009796143
beat count	479

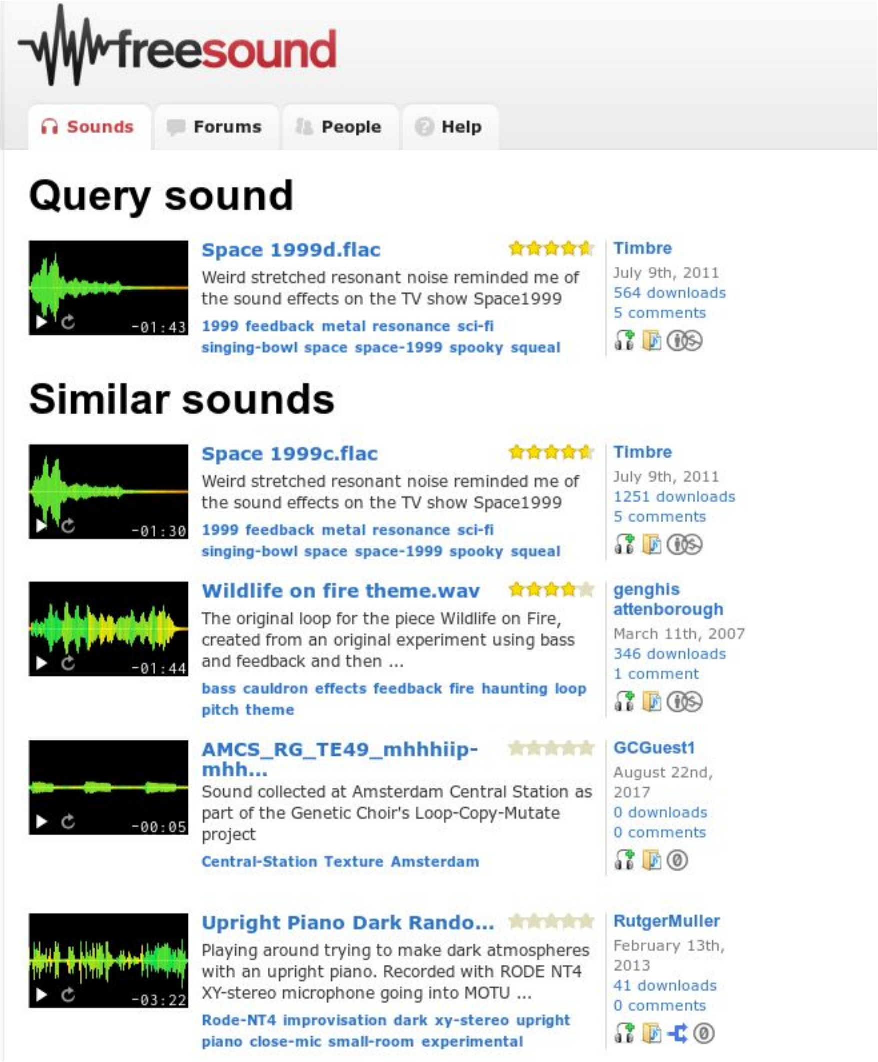
High-level information

Voice, timbre, gender, etc.	value	probability
Voice	voice	96.7%
Gender	male	94.6%
Danceability	danceable	100.0%
Tonal	atonal	97.8%
Timbre	dark	88.1%
ISMIR04 Rhythm	ChaChaCha	65.8%

Moods	value	probability
Electronic	electronic	98.1%
Party	party	91.8%
Aggressive	unsure	58.6%
Acoustic	not acoustic	99.2%
Happy	not happy	80.4%
Sad	not sad	83.9%
Relaxed	relaxed	64.0%

Genres	value	probability
Tzanetakis model	unsure	31.4%
Electronic classification	unsure	58.8%
Dortmund model	electronic	97.5%
Rosamerica model	dan	61.6%

Similarity search in Freesound https://freesound.org



Query sound

Space 1999d.flac

Weird stretched resonant noise reminded me of the sound effects on the TV show Space1999

1999 feedback metal resonance sci-fi singing-bowl space space-1999 spooky squeal

Timbre
July 9th, 2011
564 downloads
5 comments

Similar sounds

Space 1999c.flac

Weird stretched resonant noise reminded me of the sound effects on the TV show Space1999

1999 feedback metal resonance sci-fi singing-bowl space space-1999 spooky squeal

Timbre
July 9th, 2011
1251 downloads
5 comments

Wildlife on fire theme.wav

The original loop for the piece Wildlife on Fire, created from an original experiment using bass and feedback and then ...

bass cauldron effects feedback fire haunting loop pitch theme

genghis attenborough
March 11th, 2007
346 downloads
1 comment

AMCS_RG_TE49_mhhhiip-mhh...

Sound collected at Amsterdam Central Station as part of the Genetic Choir's Loop-Copy-Mutate project

Central-Station Texture Amsterdam

GCGuest1
August 22nd, 2017
0 downloads
0 comments

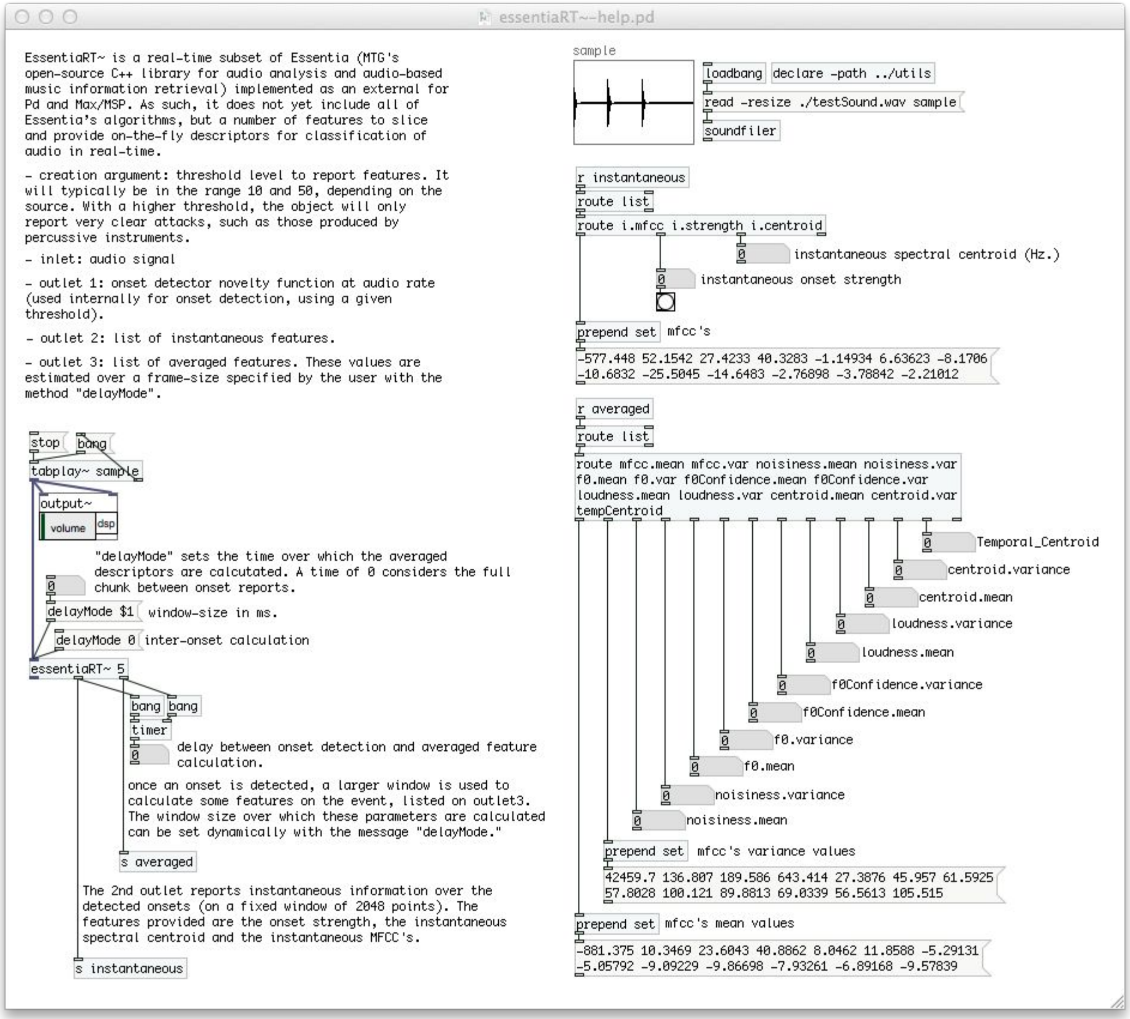
Upright Piano Dark Rando...

Playing around trying to make dark atmospheres with an upright piano. Recorded with RODE NT4 XY-stereo microphone going into MOTU ...

Rode-NT4 improvisation dark xy-stereo upright piano close-mic small-room experimental

RutgerMuller
February 13th, 2013
41 downloads
0 comments

EssentiaRT~ for Pd and Max/MSP
https://www.upf.edu/web/mtg/essentiart



EssentiaRT~ is a real-time subset of Essentia (MTG's open-source C++ library for audio analysis and audio-based music information retrieval) implemented as an external for Pd and Max/MSP. As such, it does not yet include all of Essentia's algorithms, but a number of features to slice and provide on-the-fly descriptors for classification of audio in real-time.

- creation argument: threshold level to report features. It will typically be in the range 10 and 50, depending on the source. With a higher threshold, the object will only report very clear attacks, such as those produced by percussive instruments.

- inlet: audio signal

- outlet 1: onset detector novelty function at audio rate (used internally for onset detection, using a given threshold).

- outlet 2: list of instantaneous features.

- outlet 3: list of averaged features. These values are estimated over a frame-size specified by the user with the method "delayMode".

"delayMode" sets the time over which the averaged descriptors are calcutated. A time of 0 considers the full chunk between onset reports.

0 delayMode $1 window-size in ms.

delayMode 0 inter-onset calculation

0 delay between onset detection and averaged feature calculation.

once an onset is detected, a larger window is used to calculate some features on the event, listed on outlet3. The window size over which these parameters are calculated can be set dynamically with the message "delayMode."

The 2nd outlet reports instantaneous information over the detected onsets (on a fixed window of 2048 points). The features provided are the onset strength, the instantaneous spectral centroid and the instantaneous MFCC's.

0 instantaneous spectral centroid (Hz.)

0 instantaneous onset strength

mfcc's

```
-577.448 52.1542 27.4233 40.3283 -1.14934 6.63623 -8.1706
-10.6832 -25.5045 -14.6483 -2.76898 -3.78842 -2.21012
```

0 Temporal_Centroid

0 centroid.variance

0 centroid.mean

0 loudness.variance

0 loudness.mean

0 f0Confidence.variance

0 f0Confidence.mean

0 f0.variance

0 f0.mean

0 noisiness.variance

0 noisiness.mean

mfcc's variance values

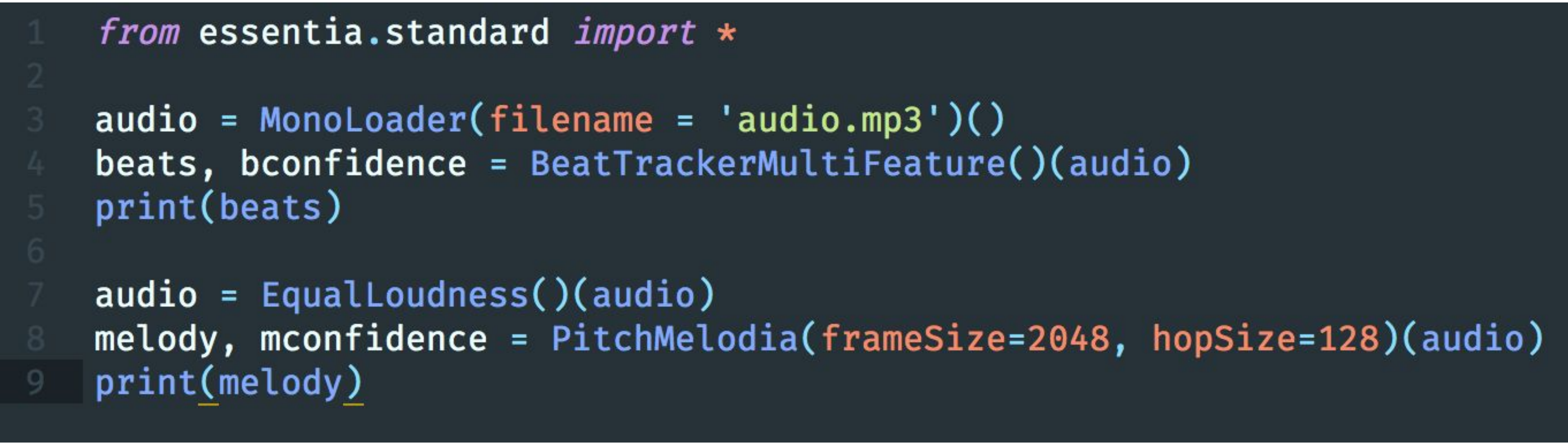
```
42459.7 136.807 189.586 643.414 27.3876 45.957 61.5925
57.8028 100.121 89.8813 69.0339 56.5613 105.515
```

mfcc's mean values

```
-881.375 10.3469 23.6043 40.8862 8.0462 11.8588 -5.29131
-5.05792 -9.09229 -9.86698 -7.93261 -6.89168 -9.57839
```

A quick taste of Essentia

Compute melody and extract beat positions:



```python
from essentia.standard import *

audio = MonoLoader(filename = 'audio.mp3')()
beats, bconfidence = BeatTrackerMultiFeature()(audio)
print(beats)

audio = EqualLoudness()(audio)
melody, mconfidence = PitchMelodia(frameSize=2048, hopSize=128)(audio)
print(melody)
```

Python tutorial

http://essentia.upf.edu/documentation/essentia_python_tutorial.html

A quick taste of Essentia

Algorithms

http://essentia.upf.edu/documentation/algorithms_reference.html

Code examples

- Python https://github.com/MTG/essentia/tree/master/src/examples/tutorial
- C++ https://github.com/MTG/essentia/tree/master/src/examples

Installing Essentia

Docker images (cross-platform)

- Essentia https://github.com/MTG/essentia-docker
    - Official Essentia docker image
    - Essentia Python extension, command-line extractors, Vamp plugins
- MIR toolbox https://github.com/MTG/MIR-toolbox-docker
    - Many useful Python tools for MIR including Essentia
    - Jupyter notebook running in a docker container
- DL4MIR https://minzwon.github.io/dl4mir
    - A docker image for the deep learning including Essentia

Installing Essentia for Python

- MacOS (needs some extra steps)
  - Install Command Line Tools for Xcode: xcode-select --install
  - Install Homebrew: http://brew.sh

    ```
    brew tap MTG/essentia
    brew install essentia --HEAD
    ```

- Windows (needs some extra step)
  - Install Bash on Ubuntu
  - Run **pip install essentia** from bash

Building Essentia from source

http://essentia.upf.edu/documentation/installing.html

1) Install dependencies
2) Build Essentia (Linux):

```
./waf configure --with-python --with-examples
./waf
sudo ./waf install
```

```
python
>>> import essentia
```

Recommendations for Python scientific environment

- Use docker or **pip install essentia**
- Do not compile Essentia from source for a Conda environment (incompatible dependencies)
- Use Python 3 (because it’s 2019!)
- Use jupyter notebooks: https://jupyter.org/
  - *python3 -m pip install jupyter*
- There are many useful Python packages out there!
  - **Matplotlib** or seaborn - for plots and data visualization
  - **Scikit-learn** - for basic machine learning
  - **Pandas** - data tables analysis

Essentia in Python

Basic tutorial:

http://essentia.upf.edu/documentation/essentia_python_tutorial.html

https://github.com/MTG/essentia/blob/master/src/examples/tutorial/essentia_python_tutorial.ipynb

Examples of using Essentia

https://essentia.upf.edu/documentation/essentia_python_examples.html

Command-line feature extractors

- Executable feature extractors (no need to use Python)
- Available in our Docker images
- Build from source: (./waf configure --with-examples)
- Binaries online: http://essentia.upf.edu/documentation/extractors/

Documentation

- http://essentia.upf.edu/documentation/extractors_out_of_box.html
- http://essentia.upf.edu/documentation/streaming_extractor_music.html
- http://essentia.upf.edu/documentation/freesound_extractor.html

Essentia tutorial notebooks

```
git clone https://github.com/MTG/essentia-tutorial.git
docker pull mtgupf/mir-toolbox
cd essentia-tutorial
./run.sh
```

Go to http://localhost:8888/
Password: mir

Audio Chords Estimation

http://essentia.upf.edu/documentation/algorithms_reference.html#tonal

- ChordsDetection
- ChordsDetectionBeats

Audio Melody Extraction and Multiple Pitch Estimation & Tracking

http://essentia.upf.edu/documentation/algorithms_reference.html#pitch

- PredominantPitchMelodia
- MultiPitchKlapuri
- MultiPitchMelodia

https://github.com/MTG/essentia/blob/master/src/examples/python/musicbricks-tutorials/5-melody_analysis.ipynb

Universitat Pompeu Fabra
Barcelona

MTG
Music Technology
Group

Q&A

Dmitry Bogdanov
@di_bogdanov
dmitry.bogdanov@upf.edu