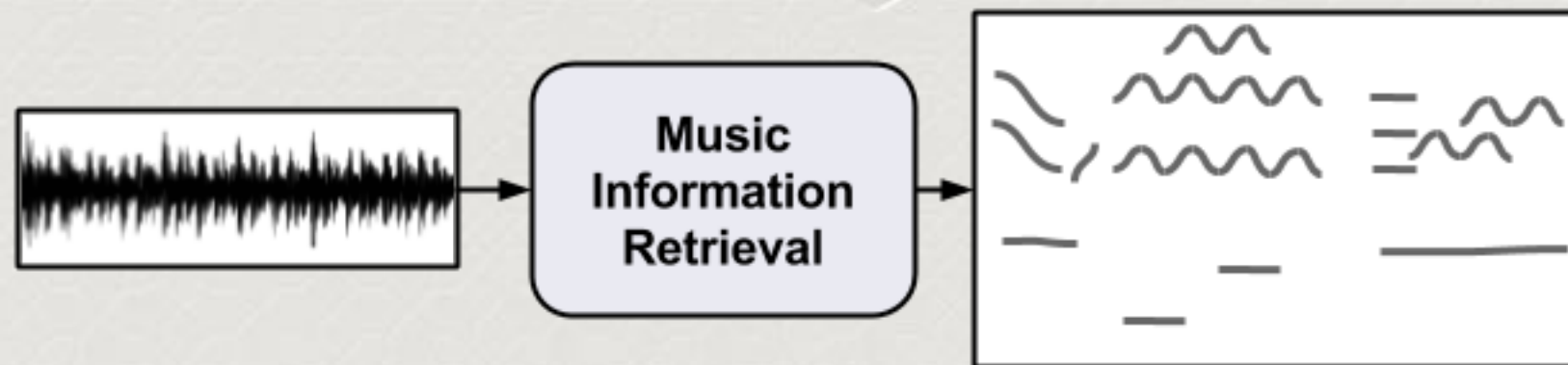


Features for Tracking Events in Musical Audio

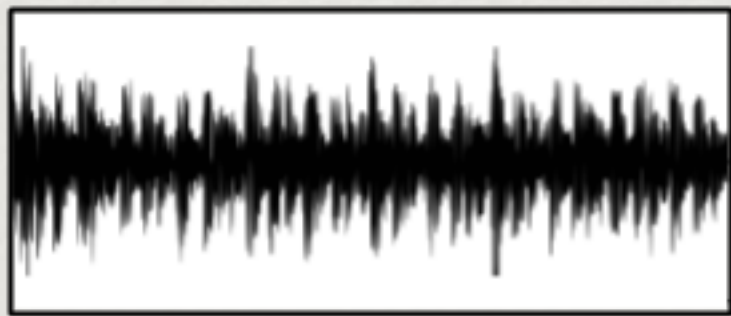
Georgi Dzhambazov
Music Technology Group (MTG),
Universitat Pompeu Fabra, Barcelona, Spain

Music Information Retrieval (MIR)

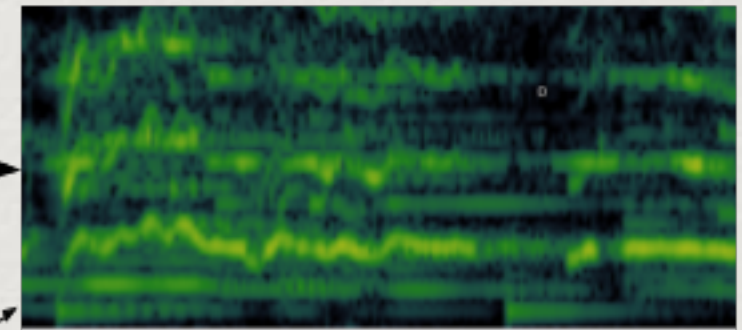


- ✦ *MIR = signal processing + machine learning + music theory*

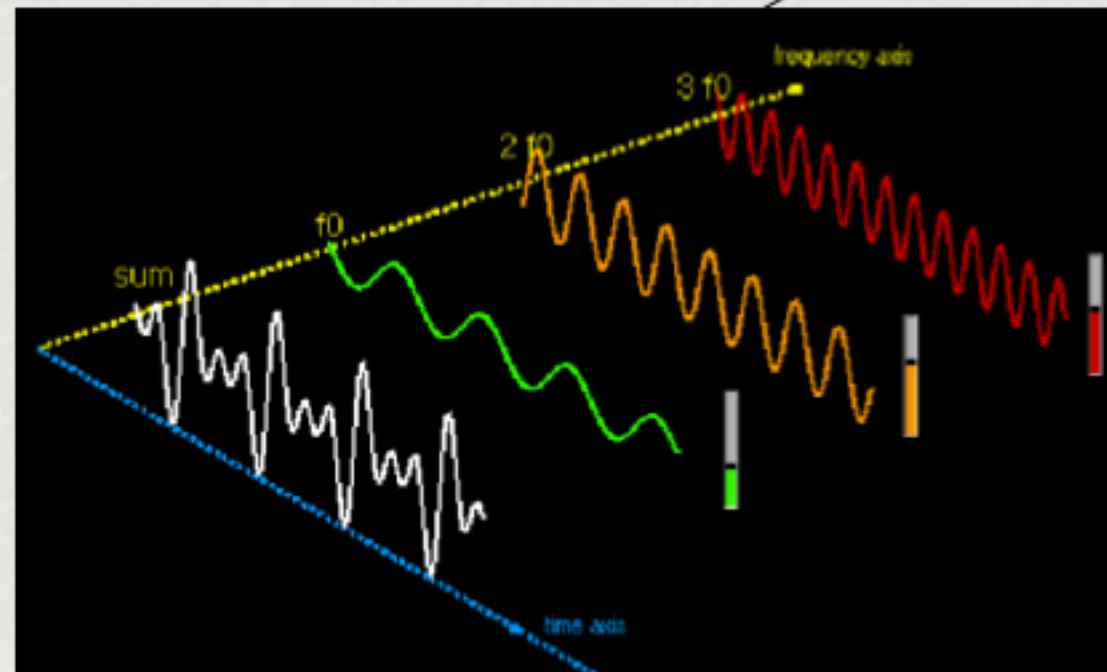
Sound content: waveform and spectrograms



waveform



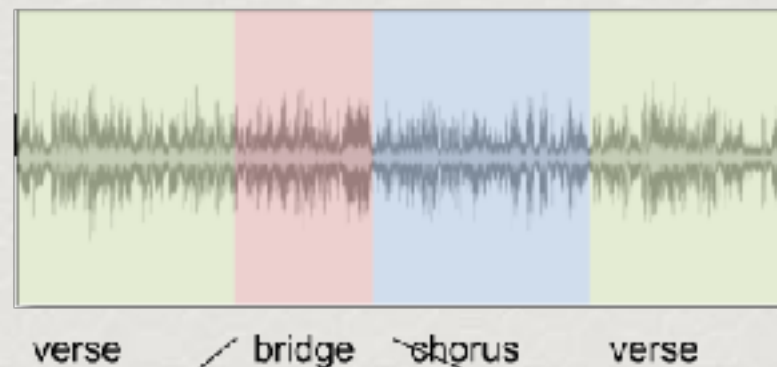
spectrogram



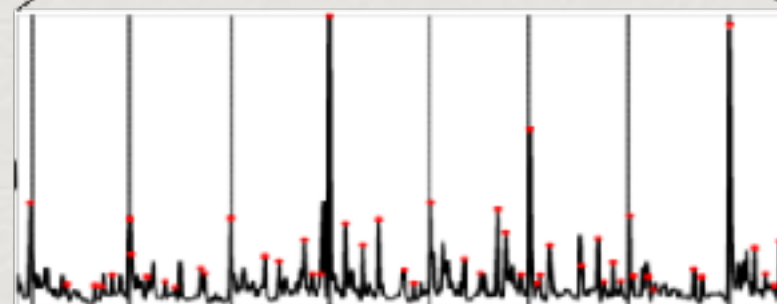
mixture of simple waves with different frequencies

Anatomy of music audio

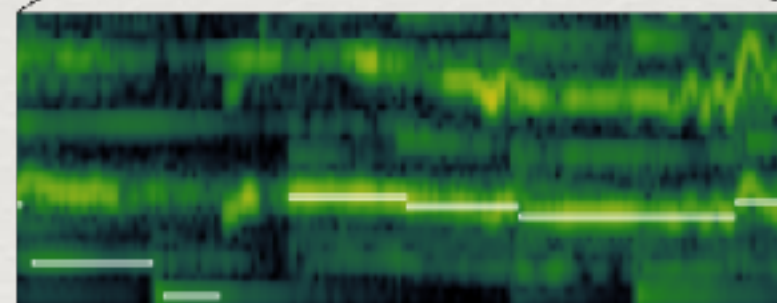
structure



rhythmic
measures



musical
events



semantic structure

repeating building block with
relatively similar pattern
silences

Music events

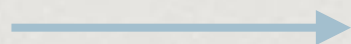
- ✿ *beats*
- ✿ *chords*
- ✿ *notes*
- ✿ *lyrics*

Acoustic features

✧ *beats*

✧ *chords*

✧ *notes*



✧ *percussive onsets*

✧ *chroma scale*

✧ *note attack*

Automatic feature extraction

Essentia: Collection of audio analysis algorithms

essential spectrogram

then beats and chords

- ✦ *Open source C++ library*
- ✦ *python wrapper*
- ✦ *realtime processing as external for Pd and Max/MSP*

ESSENTIA

- HOME
- DOCUMENTATION
- DOWNLOAD
- CONTRIBUTE
- APPLICATIONS
- NEWS
- ABOUT

Essentia is an open-source C++ library for audio analysis and audio-based music information retrieval released under the [Affero GPLv3 license](#) (also available under proprietary license upon request). It contains an extensive collection of reusable algorithms which implement audio input/output functionality, standard digital signal processing blocks, statistical characterization of data, and a large set of spectral, temporal, tonal and high-level music descriptors. In addition, Essentia can be complemented with Gaia, a C++ library with python bindings which implement similarity measures and classifications on the results of audio analysis, and generate classification models that Essentia can use to compute high-level description of music (same license terms apply).

Essentia is not a framework, but rather a collection of algorithms (plus some infrastructure) wrapped in a library. It doesn't provide common high-level logic for descriptor computation (so you aren't locked into a certain way of doing things). It rather focuses on the robustness, performance and optimality of the provided algorithms, as well as ease of use. The flow of the analysis is decided and implemented by the user, while Essentia is taking care of the implementation details of the algorithms being used. A number of predefined executable extractors for the available music descriptors are provided with the library as examples, however they should not be considered as the only correct way of doing things.

The diagram illustrates the workflow of the Essentia library. It is divided into three main sections: Blockchain, TimeExtractor, and TimeOutput. The Blockchain section includes AudioLoader, Resampler, and Analsmp. The TimeExtractor section includes Spectrum, Windowing, SpectralFrame, MFCC, Pitch, and Feature. The TimeOutput section includes TimeOutput. Arrows indicate the flow of data from Blockchain to TimeExtractor and then to TimeOutput. The diagram is labeled 'Streaming Mode' and 'Normal Mode'.

News

- Updates to cepstral features (MFCC and GFCC)**
30/12/16
- Essentia 2.1 beta3 released**
17/10/16
- Static binaries for command-line extractors for music descriptors**
22/12/16

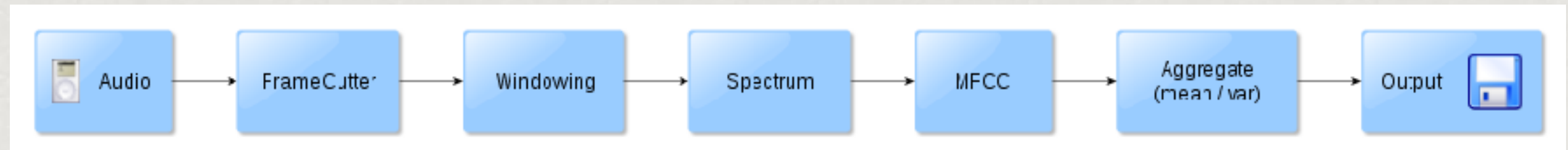
[more](#)

Automatic feature extraction

Essentia: Collection of audio analysis a

essential spectrogram
then beats and chords

✦ *Feature extractor steps:*



https://github.com/georgid/essentia_hacks/

Automatic Feature Extraction

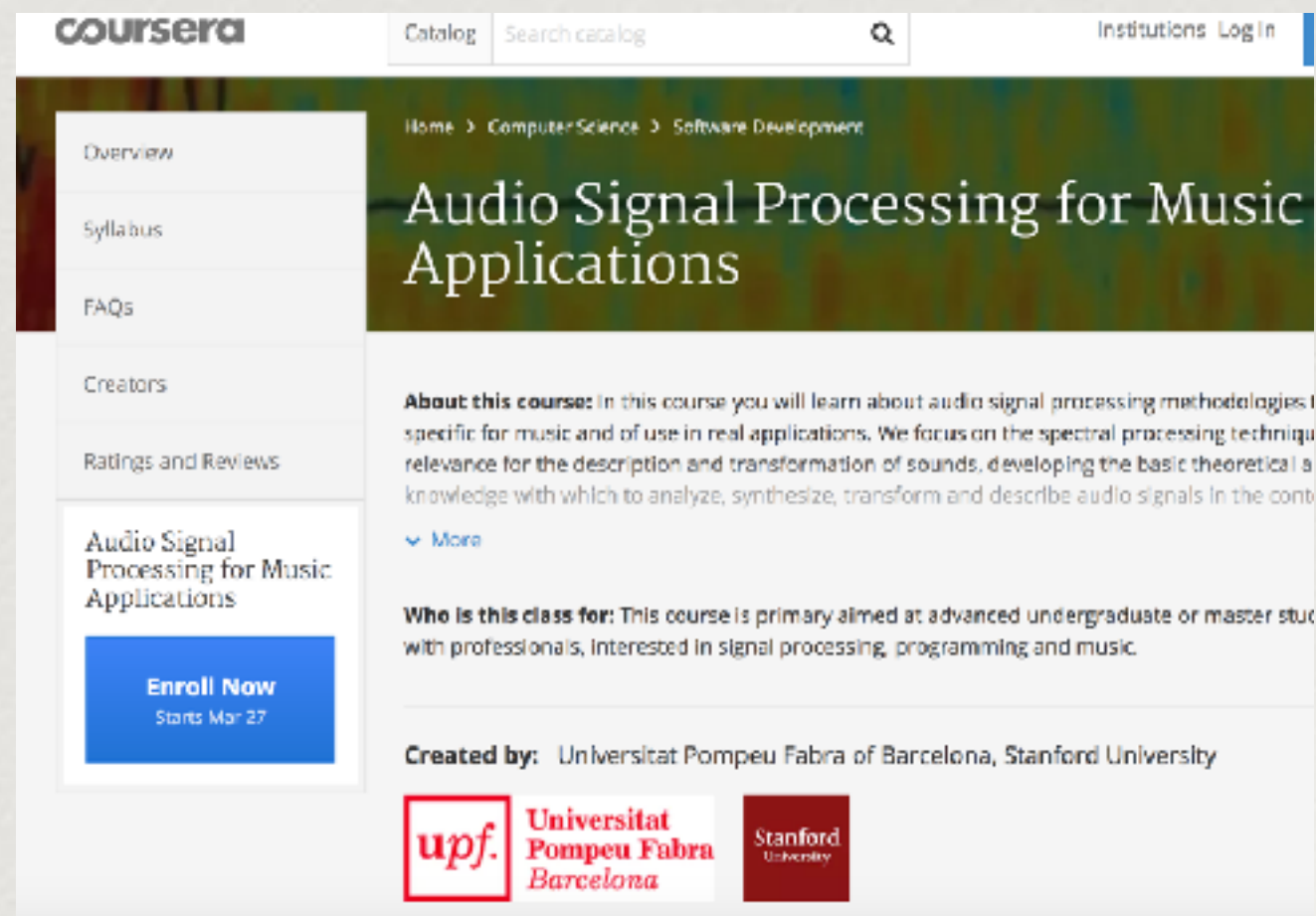
Other Tools

- ✦ *Sonic Visualiser VAMP (C++)*
- ✦ *Music Bricks (misc.)*
- ✦ *librosa (python)*
- ✦ ...

Automatic Feature Extraction: Precomputed Features

- ✿ *Acoustic Brainz*
- ✿ *Spotify API*

Music Technology Group *Courses*



The screenshot shows the Coursera website interface for the course 'Audio Signal Processing for Music Applications'. The page layout includes a top navigation bar with the Coursera logo, a search bar, and links for 'Catalog', 'Search catalog', 'Institutions', and 'Log In'. A breadcrumb trail indicates the path: 'Home > Computer Science > Software Development'. On the left, a sidebar menu lists 'Overview' (selected), 'Syllabus', 'FAQs', 'Creators', and 'Ratings and Reviews'. Below this menu, the course title 'Audio Signal Processing for Music Applications' is displayed, followed by a blue 'Enroll Now' button and the text 'Starts Mar 27'. The main content area features a large header image with the course title, followed by an 'About this course' section. This section describes the course's focus on audio signal processing methodologies for music and real-world applications, emphasizing spectral processing techniques. A 'Who is this class for?' section states that the course is aimed at advanced undergraduate or master students with professional interests in signal processing, programming, and music. At the bottom, the 'Created by' section lists 'Universitat Pompeu Fabra of Barcelona' and 'Stanford University', each accompanied by its respective logo.

coursera Catalog Search catalog Institutions Log In

Home > Computer Science > Software Development

Audio Signal Processing for Music Applications

About this course: In this course you will learn about audio signal processing methodologies specific for music and of use in real applications. We focus on the spectral processing technique relevance for the description and transformation of sounds, developing the basic theoretical knowledge with which to analyze, synthesize, transform and describe audio signals in the context of music.

▼ More

Who is this class for: This course is primarily aimed at advanced undergraduate or master students with professional interests in signal processing, programming and music.

Created by: Universitat Pompeu Fabra of Barcelona, Stanford University

upf. Universitat Pompeu Fabra Barcelona

Stanford University

Music Technology Group

Spinoffs

MTG
Music Technology
Group

RESEARCH

TECH TRANSFER

DOWN

Spinoffs

The MTG has created three spinoffs.



BMAT is an audio technology company that provides music worldwide to provide BMAT counts on a young, creative music and technology as a monitoring service position the one who serves the largest Organizations in the world, in Latin America and the Caribbean.




Reactable Systems' mission is Human-Computer Interaction, interactive products and software for entertainment, education, mediation of culture and knowledge, with emphasis on musical creation and commercial launch of the Reactable.

Music Technology Group *Hackathons*

innovation challenge

[home](#) [news](#) [challenges](#) [practical info](#) [program](#)

Sonar Innovation Challenge



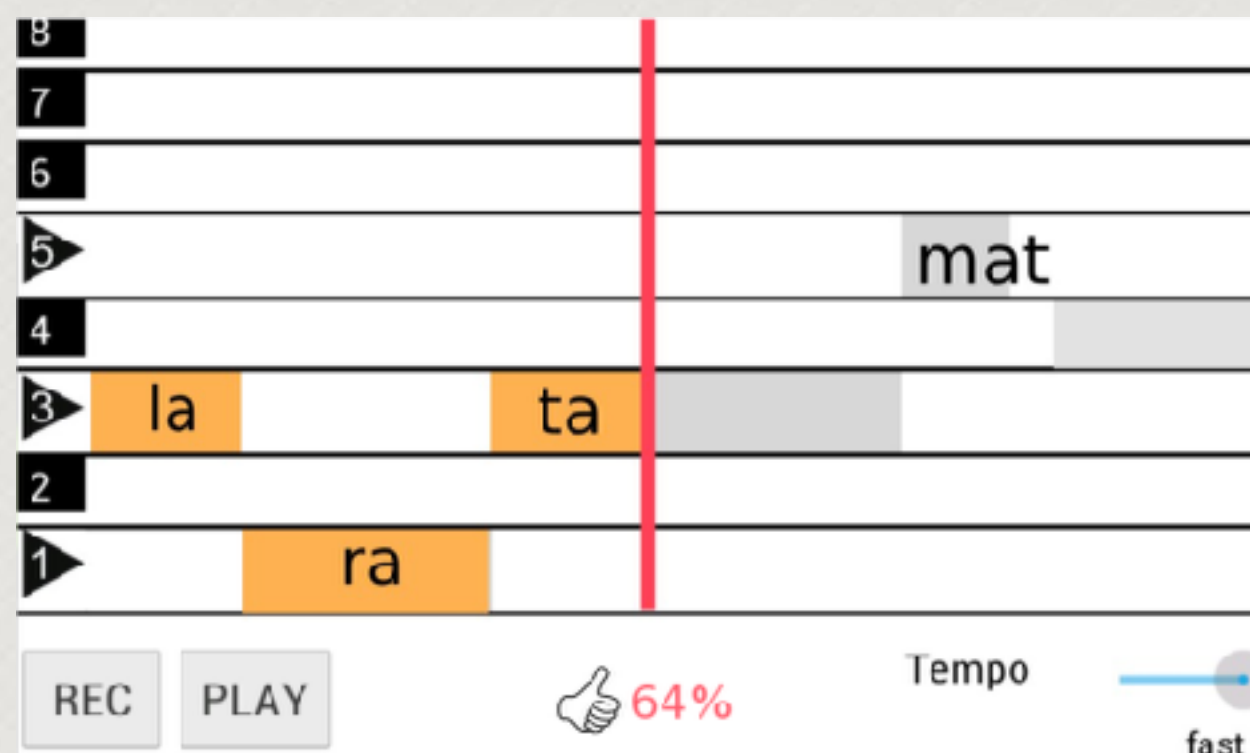
what is the sónar innovation challenge? how does it work?

My research:

Automatic lyrics-to-audio alignment

- ✿ *inputs: waveform + lyrics*
- ✿ *output: timestamps of words*

Sing Master: Karaoke game



Q & A

