Multimedia Systems

III. Sound and Music Computing

3.2. Sound and Music Description (II)

António Sá Pinto

FEUP

Agenda (Last Lecture)

- The need for Sound and Music Description
- The Semantic Gap
- Standardisation
- Classification of Music Descriptors
- Computation of Low-Level Descriptors

Agenda

- Practical Issues
- Examples of Common Descriptors
- Applications
- Resources

Practical Issues I

No standardization on Sound and Music Descriptors

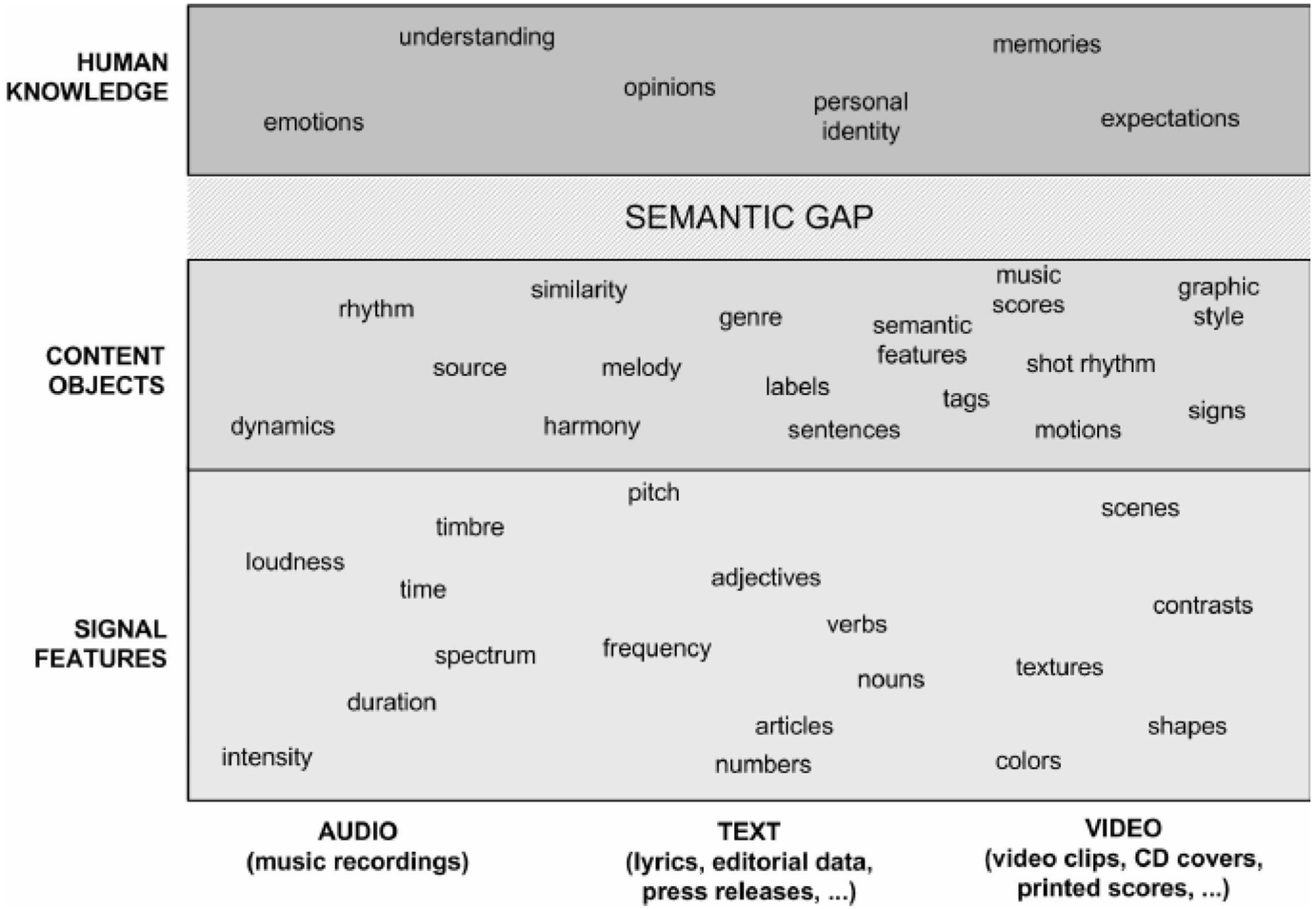
- Application-Dependent
- System-Wide Coherence
- New/Different features are developed per need
- Before Machine Learning (Feature Engineering vs DL)

Practical Issues II

Critical Analysis of Computation

- Application-Dependent
- Specific Descriptors are meaningful for specific sounds
- *Eg. Log-attack time in sounds with different envelopes; sustain in a non-sustained sound; pitch in white-noise, etc.*

Semantic Gap

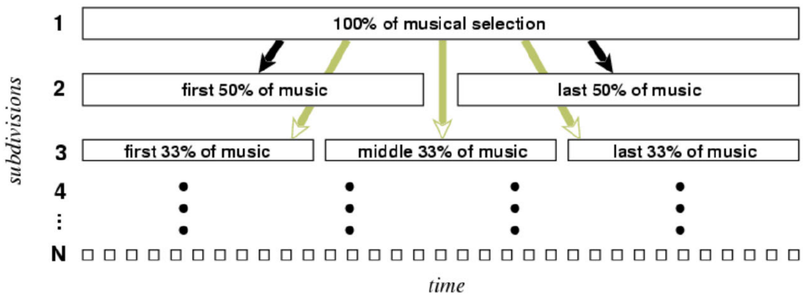


Music Descriptors

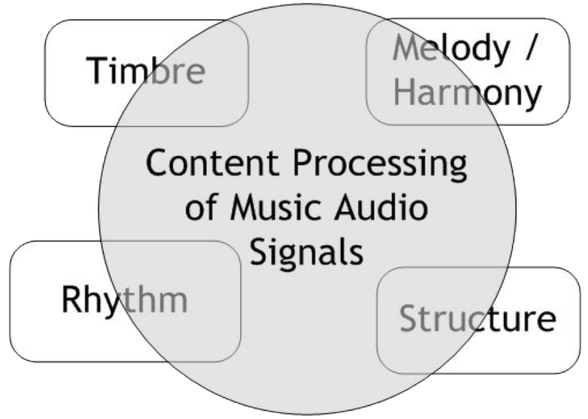
I. Abstraction

STRUCTURE		CONCEPT LEVEL		MUSICAL CONTENT CATEGORIES AND FEATURES				
CONTEXTUAL	GLOBAL DESCRIPTORS	HIGH II	EXPRESSIVE	expression				
				affect experiance				
		HIGH I	STRUCTURAL	melody	harmony	rhythm	source	dynamics
				key profile	tonality cadence	patterns tempo	instrument voice	trajectory articulation
		MID	PERCEPTUAL	successive intervallic pattern	simultane intervallic pattern	beat i o i	spectral envolopo	dynamic range sound level
NON-CONTEXTUAL	LOCAL DESCRIPTORS	LOW II	SENSORIAL	pitch		time	timbre	loudness
		LOW I	ACOUSTICAL	periodicity pich pitch deviations fundamental frequency		note-duration onset offset	roughness spectral flux spectral-centroid	peak neural-energy
				frequency		duration	spectrum	intensity

II. Temporal Scope



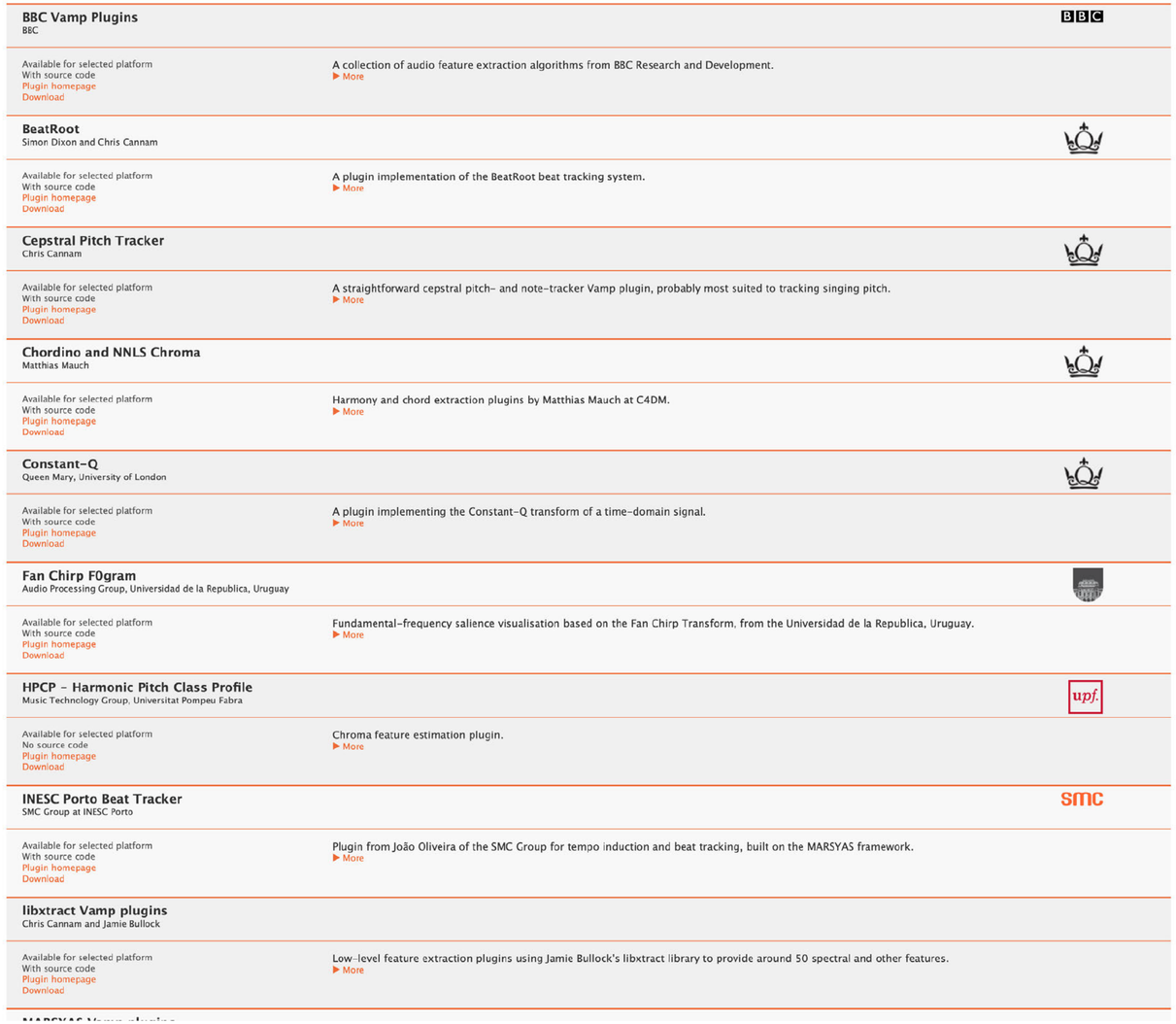
III. Musical Facet



Resources

Sonic Visualiser (Vamp Plugins)

- MIR.EDU, libxtract, aubio, etc.



BBC Vamp Plugins
BBC

Available for selected platform
With source code
Plugin homepage
Download

A collection of audio feature extraction algorithms from BBC Research and Development.
▶ More

BeatRoot
Simon Dixon and Chris Cannam

Available for selected platform
With source code
Plugin homepage
Download

A plugin implementation of the BeatRoot beat tracking system.
▶ More

Cepstral Pitch Tracker
Chris Cannam

Available for selected platform
With source code
Plugin homepage
Download

A straightforward cepstral pitch- and note-tracker Vamp plugin, probably most suited to tracking singing pitch.
▶ More

Chordino and NNLS Chroma
Matthias Mauch

Available for selected platform
With source code
Plugin homepage
Download

Harmony and chord extraction plugins by Matthias Mauch at C4DM.
▶ More

Constant-Q
Queen Mary, University of London

Available for selected platform
With source code
Plugin homepage
Download

A plugin implementing the Constant-Q transform of a time-domain signal.
▶ More

Fan Chirp F0gram
Audio Processing Group, Universidad de la Republica, Uruguay

Available for selected platform
With source code
Plugin homepage
Download

Fundamental-frequency salience visualisation based on the Fan Chirp Transform, from the Universidad de la Republica, Uruguay.
▶ More

HPCP - Harmonic Pitch Class Profile
Music Technology Group, Universitat Pompeu Fabra

Available for selected platform
No source code
Plugin homepage
Download

Chroma feature estimation plugin.
▶ More

INESC Porto Beat Tracker
SMC Group at INESC Porto

Available for selected platform
With source code
Plugin homepage
Download

Plugin from João Oliveira of the SMC Group for tempo induction and beat tracking, built on the MARSYAS framework.
▶ More

libxtract Vamp plugins
Chris Cannam and Jamie Bullock

Available for selected platform
With source code
Plugin homepage
Download

Low-level feature extraction plugins using Jamie Bullock's libxtract library to provide around 50 spectral and other features.
▶ More

Resources

Python

Essentia

- Many audio and music descriptors
- Common audio analysis processes
- Flexible and easily extendable algorithms
- “Difficult” Learning Curve
- https://essentia.upf.edu
- Bogdanov, D., Wack, N., Gómez, E., Gulati, S., Herrera, P., Mayor, O., … Serra, X. (2013). **ESSENTIA: an Audio Analysis Library for Music Information Retrieval.** In International Society for Music Information Retrieval Conference (ISMIR 2013) (pp. 493–498). Curitiba, Brasil.

Resources

Python

Librosa

- General purpose
- It implements few sound descriptors
- Very simple to learn (e.g. pip install)
- https://librosa.github.io
- McFee, B, Raffel, C., Liang, D., Ellis, D., …. **librosa: Audio and Music Signal Analysis in Python.** In Proceedings of the 14th Python in Science Conference (SCIPY2015). Curitiba, Brasil.
- https://www.youtube.com/watch?v=MhOdbtPhbLU (SCIPY Presentatiion)

Resources

Matlab

Timbre Toolbox

- Good for use on large collections
- Difficult to change parameters (hard-coded)
- https://github.com/mondaugen/timbretoolbox
- Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). **The Timbre Toolbox: Extracting audio descriptors from musical signals**. The Journal of the Acoustical Society of America, 130(5), 2902. https://doi.org/10.1121/1.3642604

Resources

Matlab

MIR Toolbox

- User-friendly
- Well documented
- https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox
- Lartillot, O., Toiviainen, P., & Eerola, T. (2007). **A Matlab Toolbox for Music Information Retrieval**. Data Analysis, Machine Learning and Applications, 261–268. https://doi.org/10.1007/978-3-540-78246-9_31

Low-level descriptors

1. Windowing (frame)

- Size of window,
- Overlap factor (hop size);
- Shape of window.

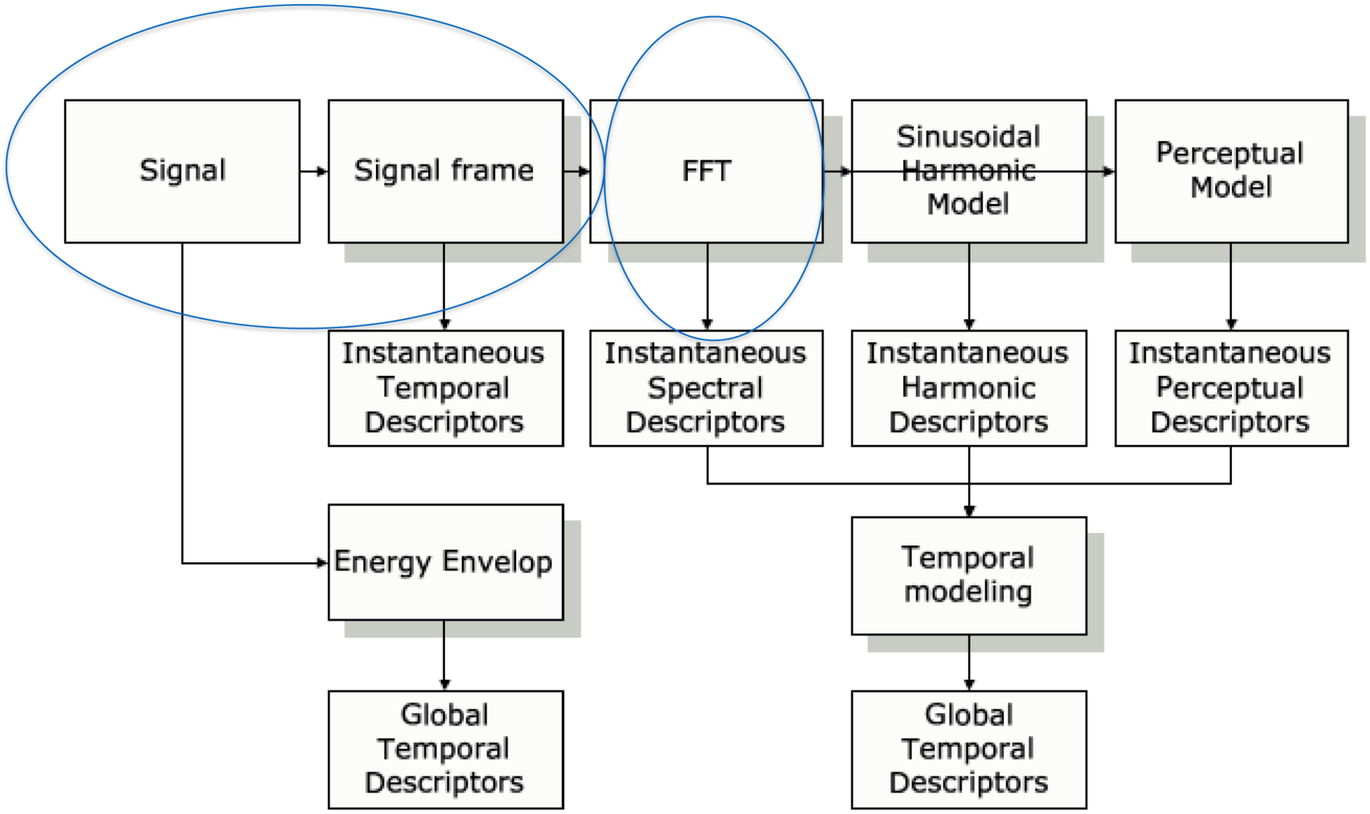
Influence in the analysis!!!

3. Spectral Analysis:

DFT(FFT) → STFT

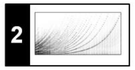
FFT size? Zero-Padding

Influence in the analysis!!!



Fourier Analysis

Chapter 2: Fourier Analysis of Signals



Topic	Relation to [Müller, FMP, Springer 2015] & Description	HTML	IPYNB
Complex Numbers	[Section 2.2.3.1, Exercise 2.1] Absolute value; angle; polar representation; conjugation; inverse	[html]	[ipynb]
Exponential Function	[Section 2.3.2.1, Exercise 2.15] Power series; Euler's formula; root of unity	[html]	[ipynb]
Discrete Fourier Transform (DFT)	[Section 2.1.3, Section 2.4] Inner product; DFT matrix; fast Fourier transform (FFT)	[html]	[ipynb]
DFT: Phase	[Section 2.3.2] Exponential function; polar coordinates; complex Fourier coefficient; phase	[html]	[ipynb]
Discrete Short-Time Fourier Transform (STFT)	[Section 2.1.4, 2.5.3] Time localization; spectrogram; physical interpretation;	[html]	[ipynb]
STFT: Influence of Window Function	[Section 2.5] Window type; window size	[html]	[ipynb]
STFT: Padding	[Section 2.1.4, Section 2.5.3] Padding variants; edge phenomena	[html]	[ipynb]
STFT: Frequency Grid Density	[Section 2.1.4] DFT frequency grid; zero padding; increasing density	[html]	[ipynb]
STFT: Frequency Interpolation	[Section 2.1.4, Section 3.1.1] Linear interpolation; cubic interpolation; log-frequency STFT; cents	[html]	[ipynb]
STFT: Inverse	[Section 2.4.2, Section 8.1.2.1] DFT; inverse DFT; window function; overlap–add technique; artition of unity	[html]	[ipynb]
STFT: Conventions and Implementations	[Section 2.1.4, 2.5.3] Time axis convention; centered windowing; frequency conversion; implementation issues	[html]	[ipynb]
Digital Signals: Sampling	[Section 2.2.2.1, Exercise 2.28] Equidistant sampling; aliasing; sampling theorem	[html]	[ipynb]
Digital Signals: Quantization	[Section 2.2.2.2, Exercise 2.9] Uniform quantization; quantization error	[html]	[ipynb]
Interference and Beating	[Section 2.3.3.1, Exercise 2.19] Constructive interference; destructive interference; chirp; sweep	[html]	[ipynb]