Basic system overview

Segmentation
(Frames, Onsets, Beats, Bars, Chord Changes, etc)

Feature Extraction
(Time-based, spectral energy, MFCC, etc)

Analysis / Decision Making
(Classification, Clustering, etc)
Short Term Fourier Transform

Time-Frequency Grid of the STFT

Minimum Time Resolution

Minimum Discriminable Frequency
Wavelet time–frequency analysis

Continuous wavelet transform (CWT) decomposes \textbf{(invertibly)} a signal onto scaled and translated instances of a finite time “mother function” or “basis”.

\[
W_s(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(\tau) \cdot \overline{g}\left(\frac{\tau - b}{a}\right) d\tau, \quad a > 0
\]

\[g(t) = e^{-t^2/2} \cdot e^{i\omega_0 t}\]

\[
g(t) = e^{-t^2/2} \cdot e^{i\omega_0 t}
\]
Example: Sinusoidal Signal
Example: Sinusoidal Signal

Phase Mapping:

\[ \pi \]
\[ \pi/2 \]
\[ 0 \]
\[ 3\pi/2 \]
Example: Simple Rhythm

Scaleogram and Phaseogram of an isochronous pulse rhythmic signal:
Implementation

- Implemented as a set of complex value bandpass filters in Fourier domain.
- Scaling produces a “zooming” time window for each frequency “scale”.
- Creates simultaneous time and frequency localisation close to the Heisenberg inequality.
Wavelet Time-Frequency Resolution from Dilation ("Zooming")

Scale as IOI Range in Samples

Time in Samples
Wavelets for Rhythm (Smith & Honing 2008)

- The CWT enables representation of temporal structure in terms of time varying rhythmic frequencies.
  - Produces magnitude and phase measures which reveal time–frequency ridges indicating the frequencies present in the input rhythm signal (collectively a skeleton, Tchamitchian & Torrésani ’92).
Musical Example

- The rhythm of “Greensleeves”...
Greensleeves

Magnitude of greensleeves-perform-medium

Phase of greensleeves-perform-medium
Memory Based Tactus

Wavelet rhythm analysis is also applicable to continuous onset salience traces from auditory models (Coath, et. al 2009).
Memory Based Tactus

- Uses lossy windowed integrator to amass tactus likelihood.
- Suppress all but the magnitude coefficients of the extracted tactus ridge.
- Invert the extracted tactus ridge and original phase plane back to the time domain. Creates a single beat oscillation.
- Nominating a starting beat and noting its phase, all other foot-taps are generated for the same phase value.
Reconstructed Phase

Foot tap phase of hum-68_cropped_lognovelty
Example: Foot-tapping to singing

- Singing examples of Dutch folk songs from the "Onder de Groene Linde" collection (Meertens Institute).
- Uses continuous wavelet transform of rhythmic signals (Smith 1996, Smith & Honing 2008) to derive tactus:
- Example 1: Original... + Accompaniment.
- Example 2: ...Original + Accompaniment.