CCRMA MIR Workshop 2014
Beat-finding and Rhythm Analysis

Leigh M. Smith
Humtap Inc.
leigh@humtap.com

Outline

- Modelling Rhythm Cognition.
- Onset-detection.
- Beat-tracking & Tempo-derivation.
  - Autocorrelation.
  - Beat Spectral approaches.
  - Histogram models.
- Meter determination.
- Applications, Exercises

Basic system overview

Beat-finding and Tempo Derivation

- Why?
  - Tempo and Beat are strong discriminators in judgements of music similarity, and even genre (Tzanetakis & Cooke 2002, Dixon et. al 2004).
  - Understanding the beat facilitates understanding the importance other musical elements:
    - Relative importance of tonal features.
    - Diatonic or chromatic character of a piece.
    - Complexity of a piece.
  - Applications: musicology & ethnomusicology, automatic DJing, query by example, composition tools.

Example: Foot-tapping to singing

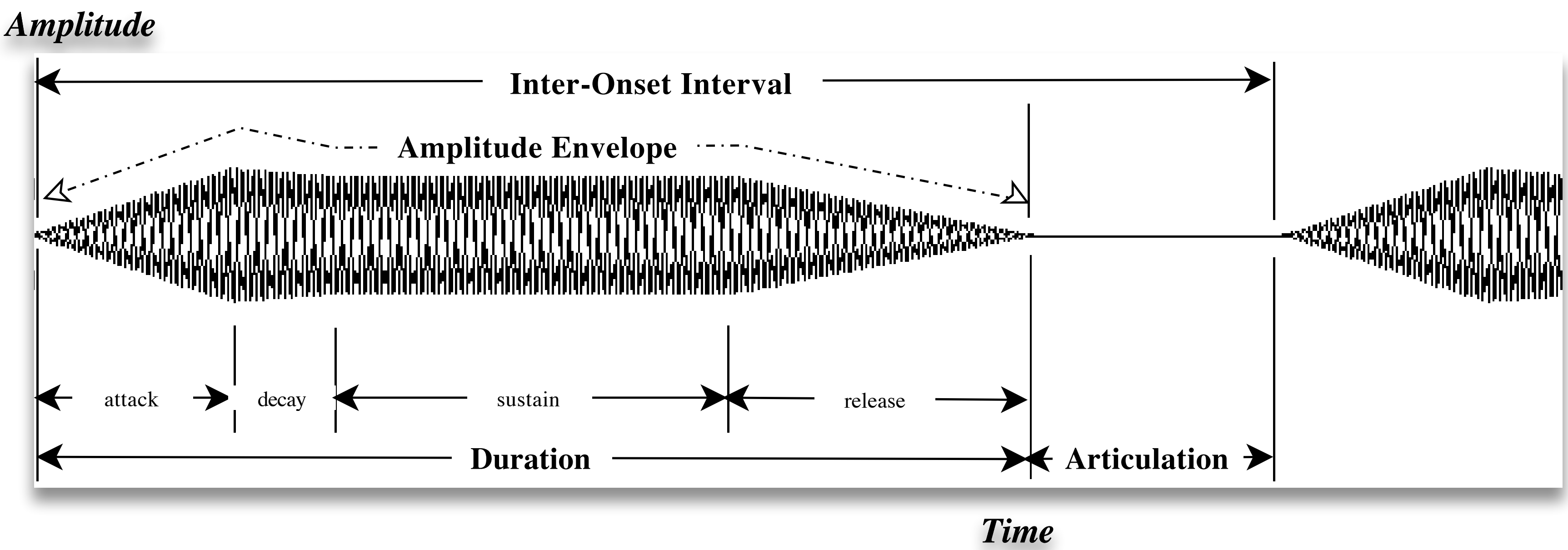
- Singing examples of Dutch folk songs from the "Onder de Groene Linde" collection (Meertens Institute).
- Uses continuous wavelet transform of rhythmic signals (Smith 1996, Smith & Honing 2008) to derive tactus:
- Example 1: Original... + Accompaniment.
- Example 2: ...Original + Accompaniment.

Modelling Rhythm

- “...the *systematic* patterning of sound in terms of timing, accent, and grouping.” (Patel 2008 p.96)
  - (Not always periodic patterns)
- Accent sources include: dynamics, melody, harmony, articulation, timbre, onset asynchrony etc.
- Consists of *hierarchical* and *figural* (proximal) temporal structures.

Measuring Beat

- Inter-Onset Intervals (IOI)
- Inter-Beat Interval (IBI)
- Tempo: frequency of the beat (BPM) = 1/IBI

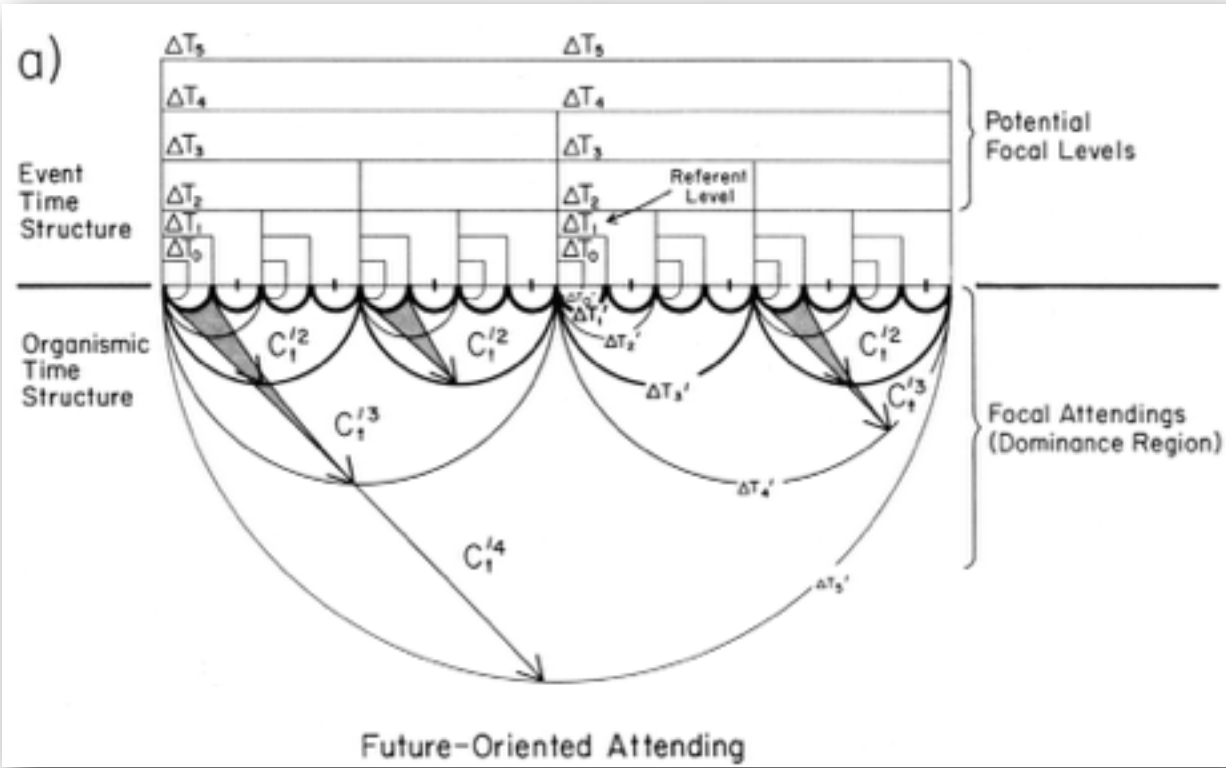


Theory and Perception of Musical Time

- Multiple simultaneous levels of musical time
  - Tactus: the foot-tapping rate.
  - Tempo: estimated from tactus, typically median IBI.
  - Meter: Periodic perceived accentuation of beats.
  - Tatum: Shortest interval between events.
- Rubato – change in tempo during performance to emphasise structure.

Rhythmic Strata

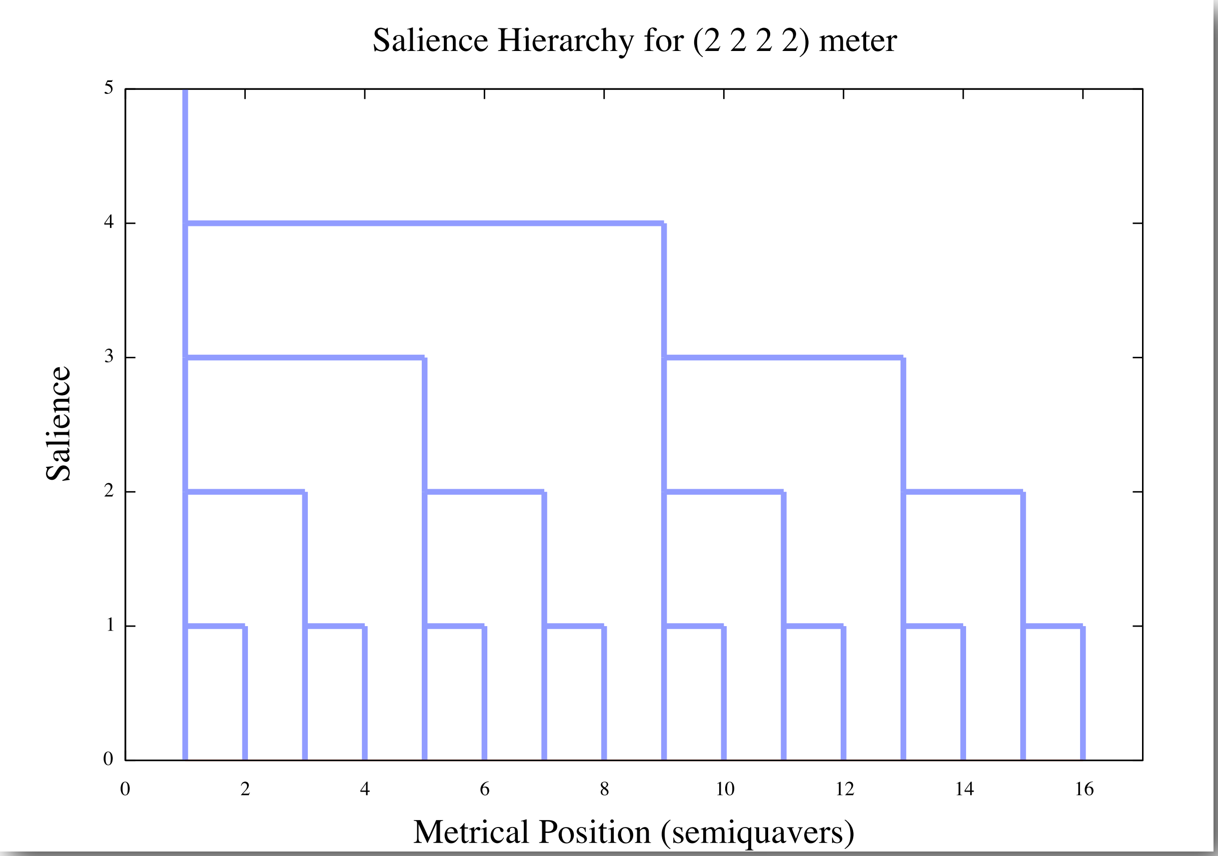
- Musical rhythm can be considered as composed of a hierarchy of temporal levels or strata (Yeston 1976, Lerdahl & Jackendoff 1983, Clarke 1987, Jones & Boltz 1989).



From Jones & Boltz '89

Metrical Structure

- Meter is expressed as a hierarchical grouping in time. e.g Subdivision of 4/4 (4 beats to the bar):

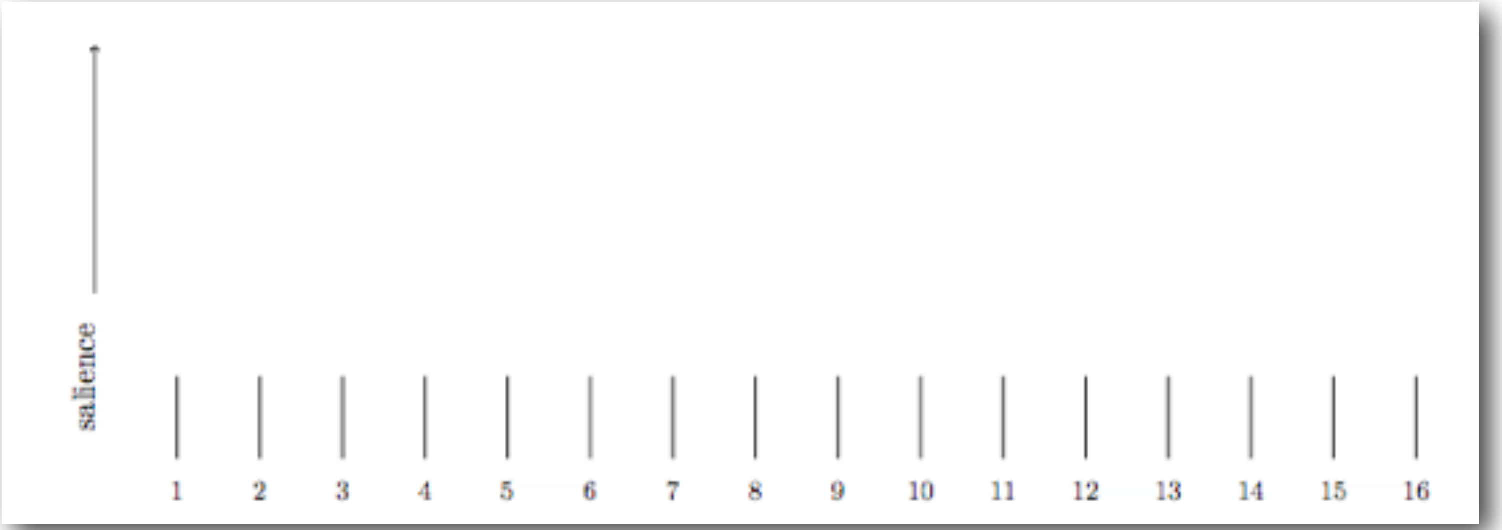


Martin 1972, Longuet-Higgins & Lee 1982, Honing 2002

Meter

- Meter is expressed in Western music as time-signatures (4/4, 3/4 etc).

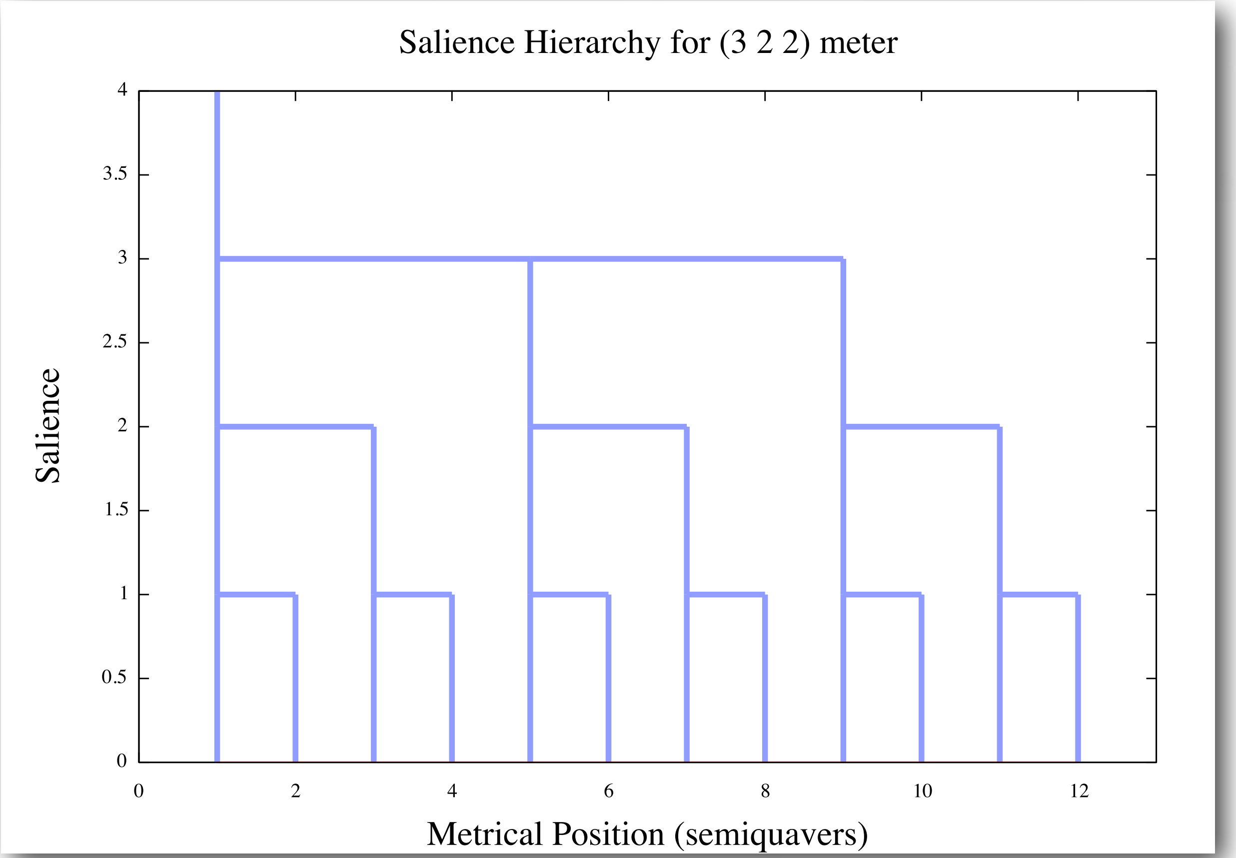
Subdivision of 4/4 (4 beats to the bar):



(Courtesy Olivia Ladinig)

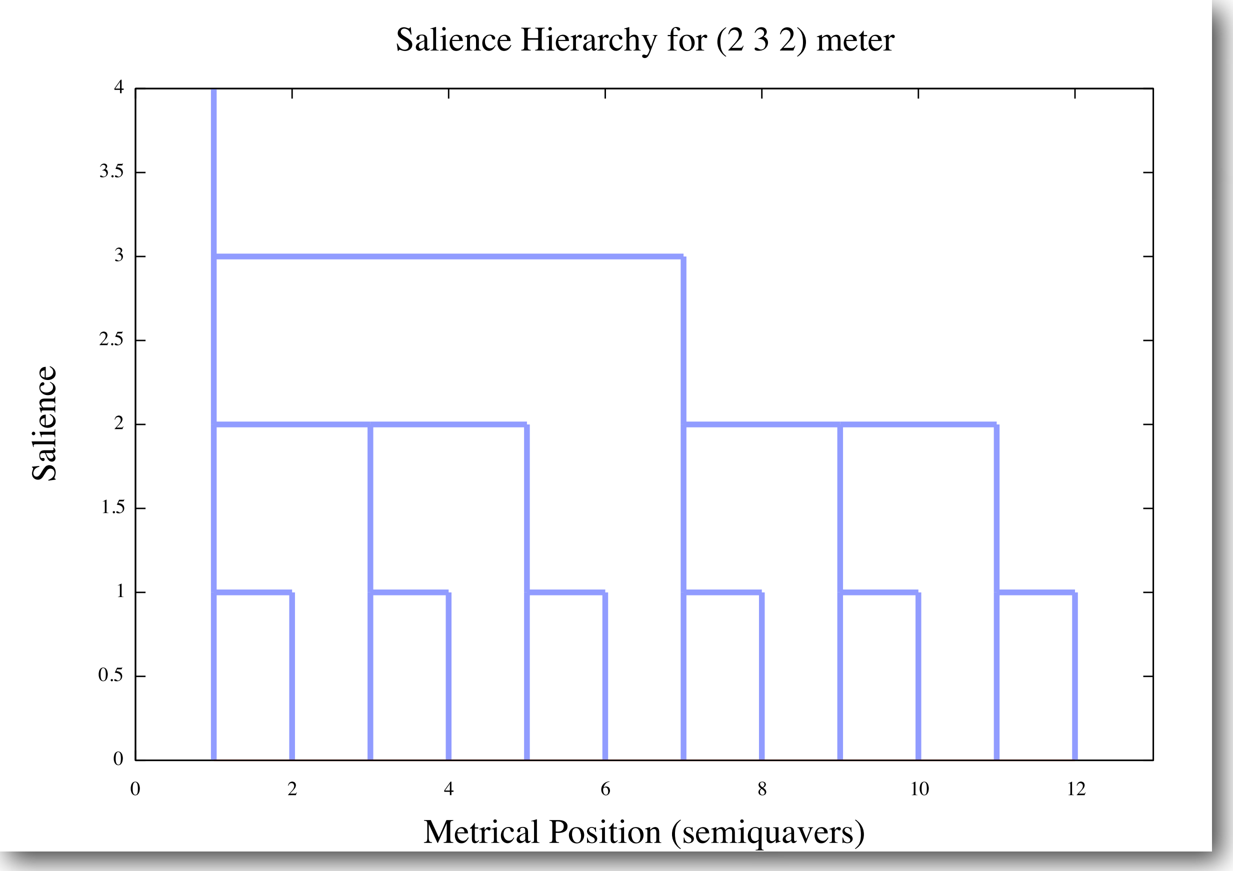
Meter

- Subdivision of 3/4 (3 beats to the bar):



Meter

- Subdivision of 6/8:



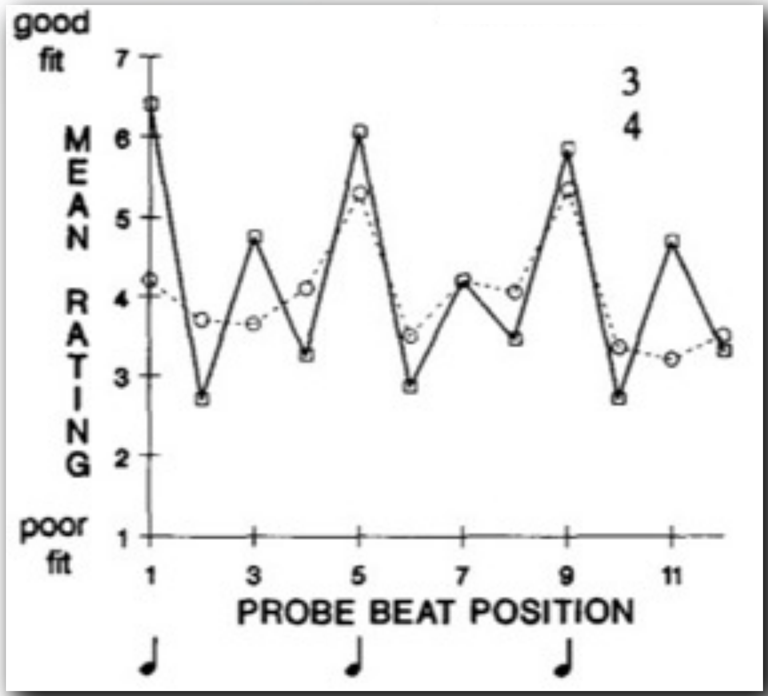
Hierarchical Grouping: Meter

- Meters are argued to arise from the interaction between temporal levels (Yeston 1976).
  - Therefore a meter implies two frequencies: the pulse rate and the measure (“bar”) rate.
- The tactus is considered as the most salient hierarchical level, consistent with the notated meter, or the foot tapping rate (Desain & Honing 1994).

Mental schemas for Meter

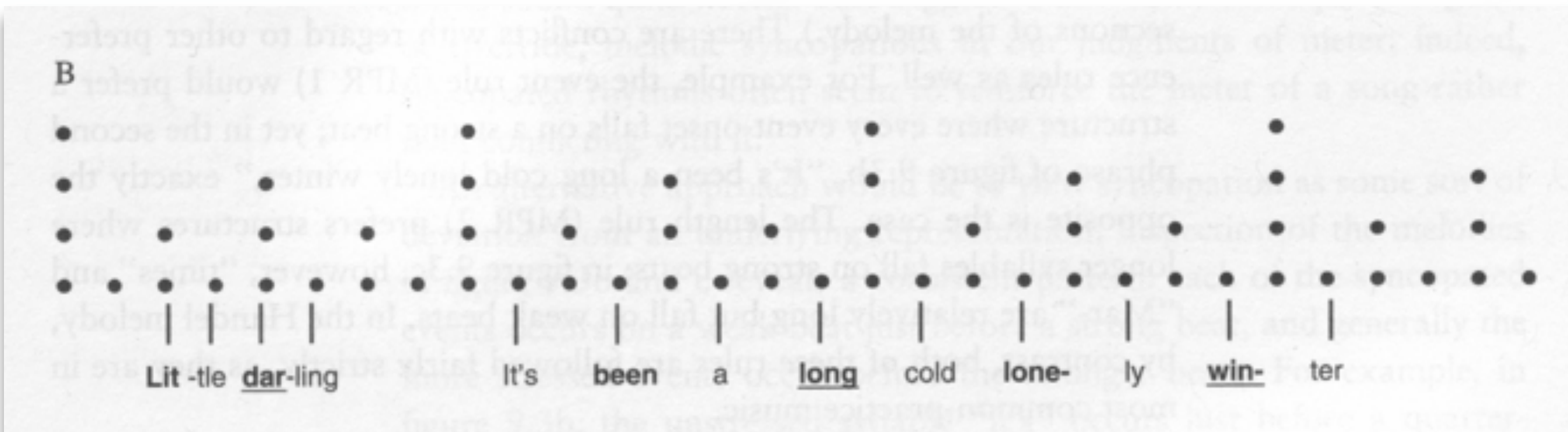
- Metrical Profiles (Palmer & Krumhansl 1990)
  - Pre-established mental frameworks (“schemas”) for musical meter are used during listening.

From Palmer & Krumhansl (1990). Mean goodness-of-fit ratings for musicians (solid line) and nonmusicians (dashed line).



Syncopation

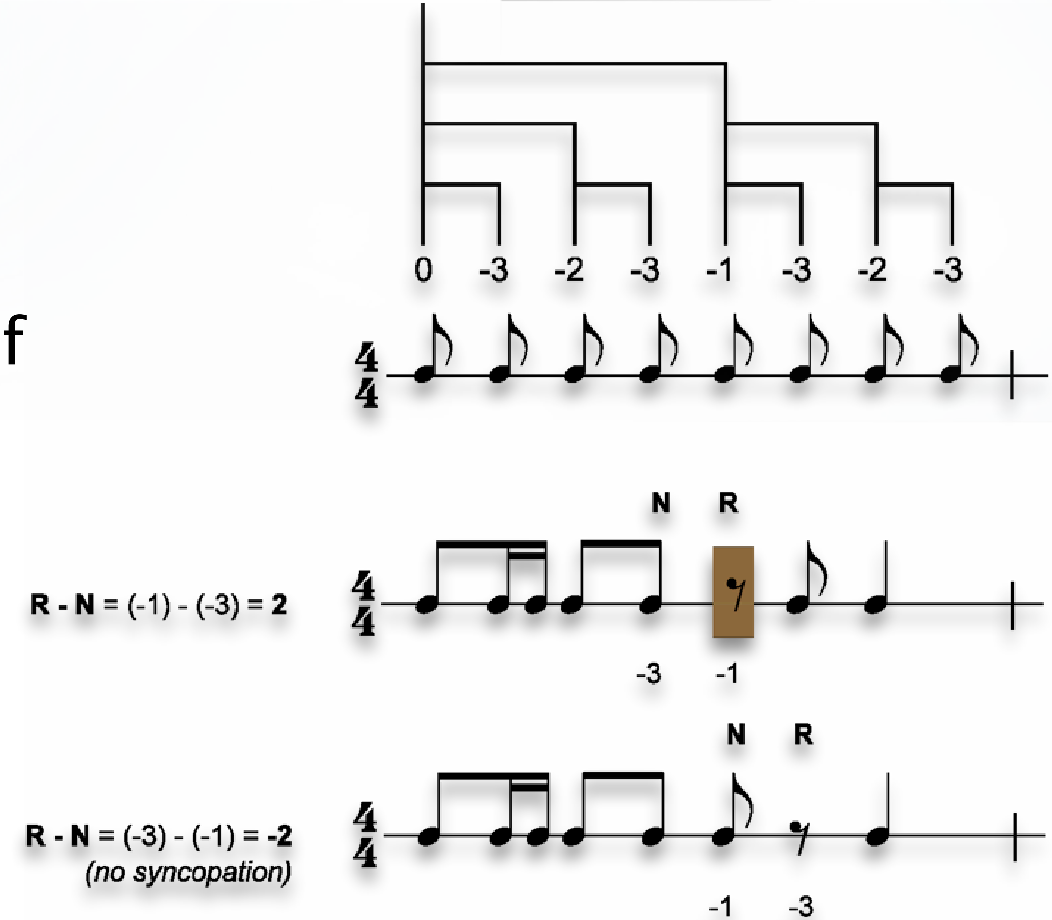
- “...some kind of conflict between [phenomenal] accents and meter” (Temperely 2001, p.239).



From Temperley D., “The Cognition of Basic Musical Structures” MIT Press, 2001, p.241

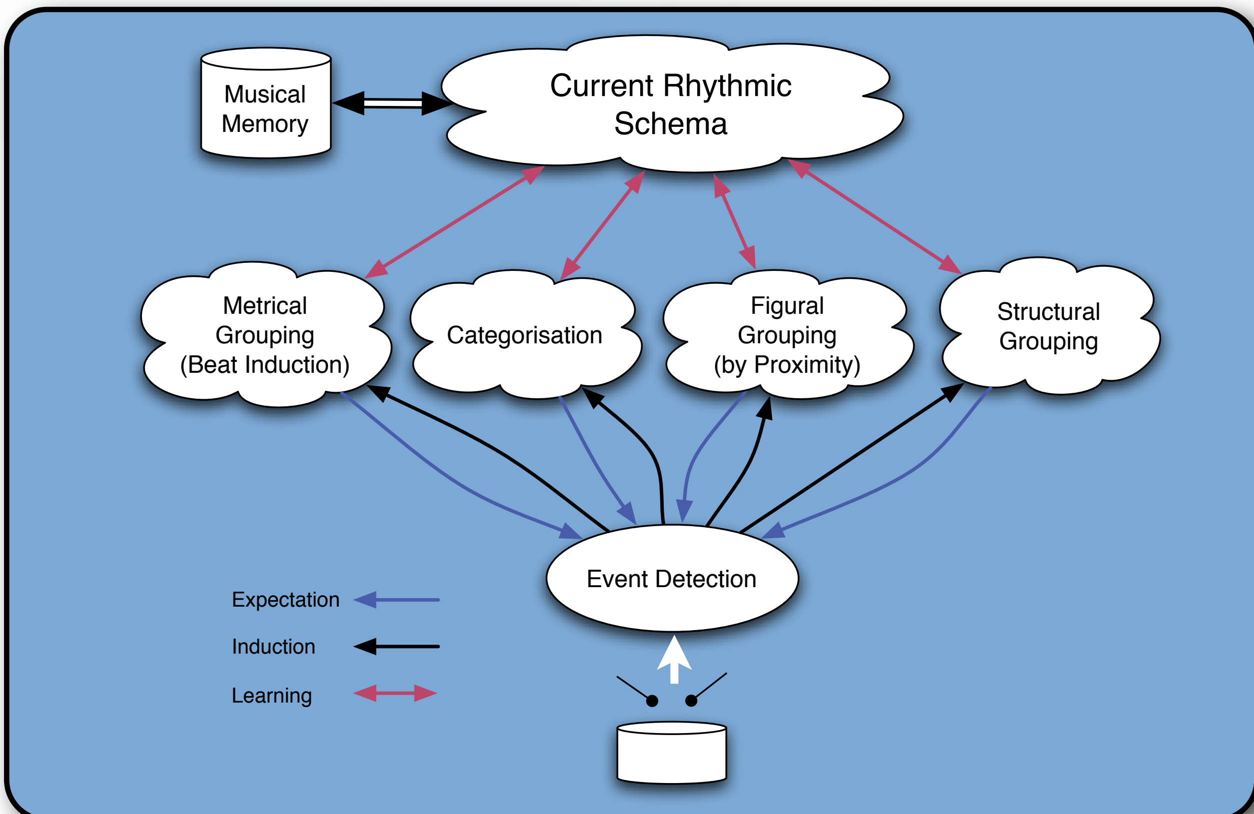
Syncopation

- Listener judgements of musical complexity are correlated with degree of syncopation (i.e. note location within the beat) (Shmulevich & Povel 2000, Smith & Honing 2006).
- Compared judgements against formal model of syncopation (Longuet-Higgins & Lee 1984).



Active Rhythm Perception

- Viewed as a resonance between top down and bottom-up processes (see e.g Desain & Honing 2001):



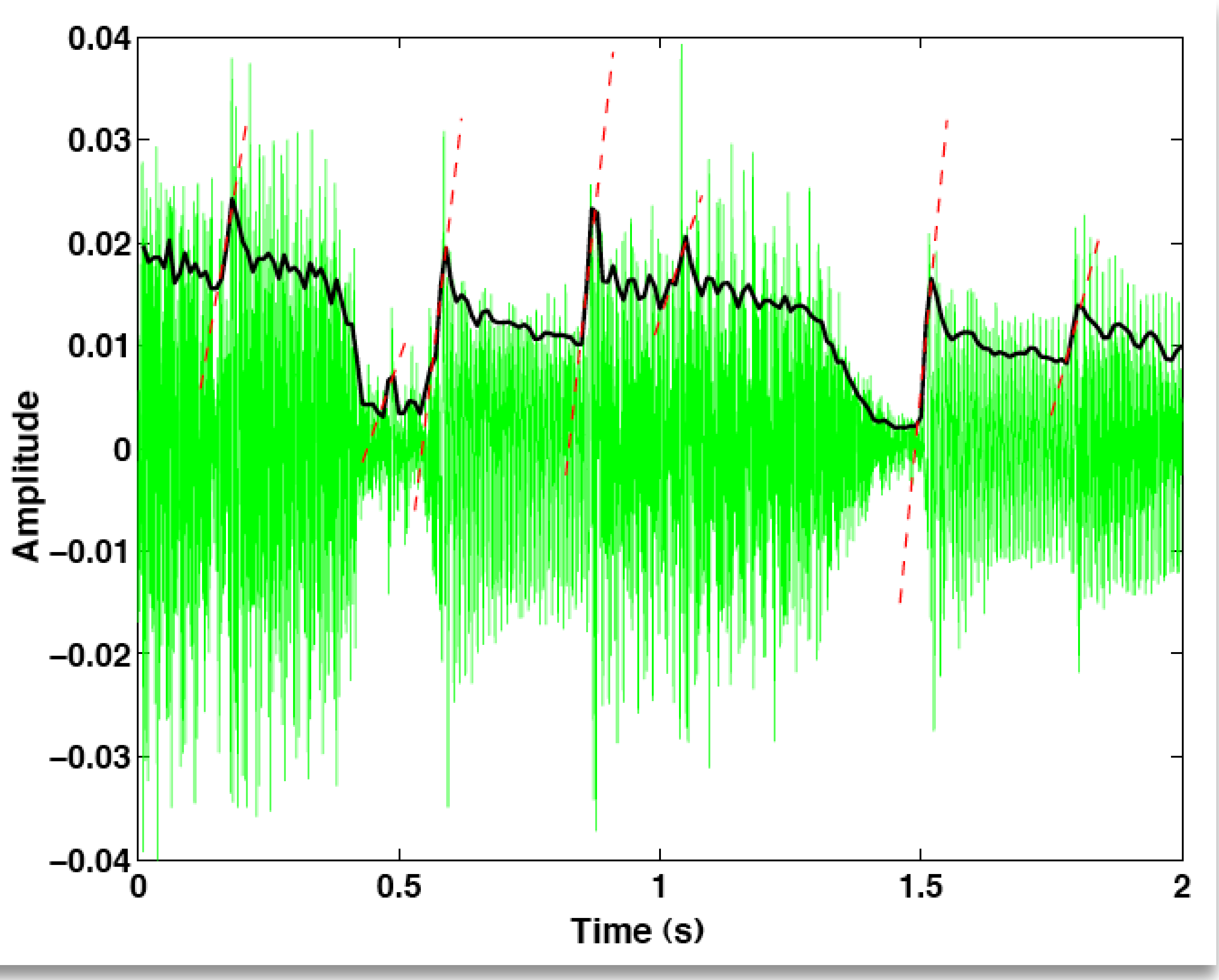
Onset-detection vs. Beat-detection

- Traditionally beat detection relied on accurate onset detection.
  - i.e from MIDI data for **Score Following** (Dannenberg 1991, Cont 2009).
- This can be difficult for MIR from polyphonic audio recordings.
  - A higher freq. **Onset Detection Function** from the entire audio signal can be used for beat tracking without all onsets being detected (Schloss 1985, Goto & Muraoka 1994, Scheirer 1998).

The Onset Detection Function

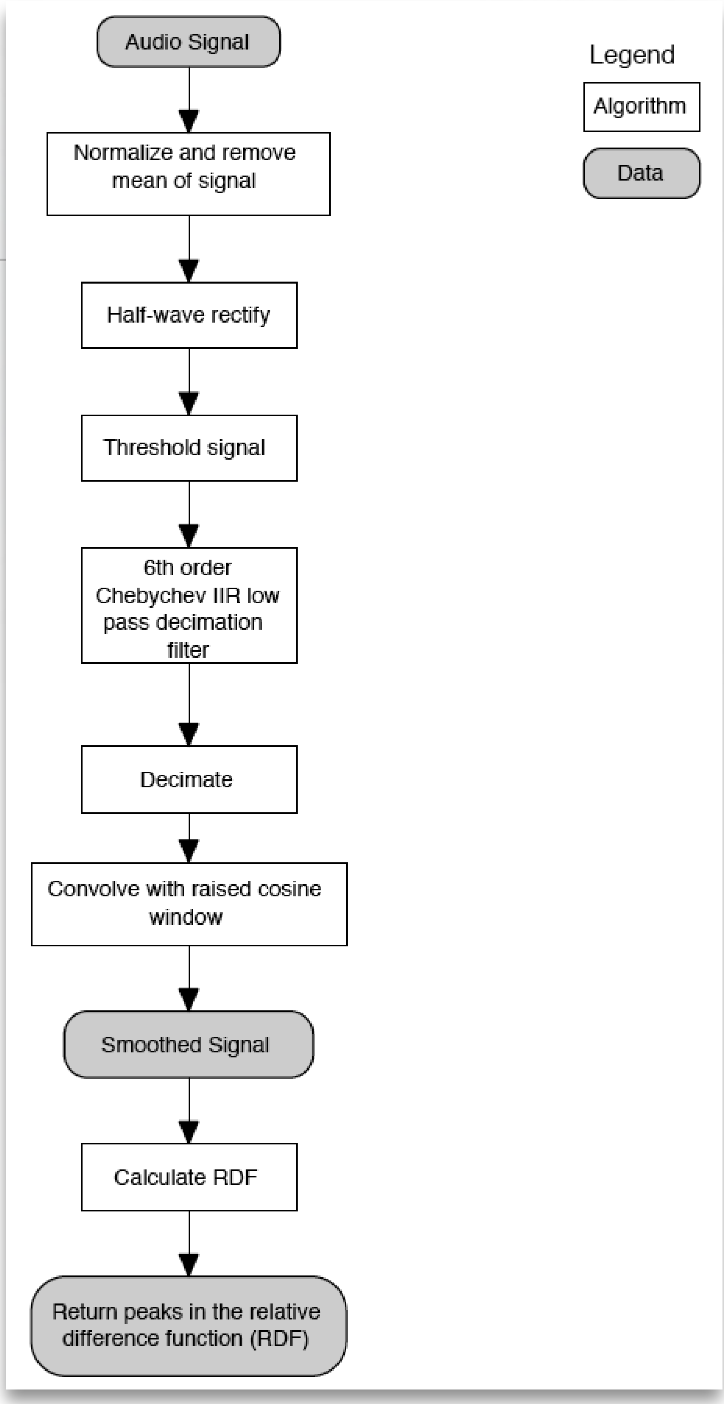
- Represents:
  - Ideal: Each note that contributes to the beat.
  - Practice: Combined envelopes of all notes.
- Tends to emphasise:
  - strong transients (i.e. impulsive sounds)
  - loud notes
  - bass notes
  - wide-band spectrum events (e.g. snare drums).

Dixon's Envelope Onset Detection



Example Onset Detection

- Pre-processing
- Filtering
- Down-sampling
- Difference function



Common ODF methods

- e.g (Bello et. al 2005, Dixon 2007, Peeters 2007)

- Optional pre-rectification filtering.
- Envelope mixture from rectification/energy.
- Smoothing of envelope (LP filter).
- Down-sampling for data reduction.
- d(log E)/dt highlights perceived impulses.
- Weighting higher frequencies captures wide-band events.
- Spectral difference between STFT frames.

Existing Beat tracking Models

- Parsing metrical grammars (Longuet-Higgins and Lee 1982).
- Forward projection of likelihood (Desain 1992).
- Autocorrelation (Desain & Vos 1990, Brown 1993, Eck 2006).
- Oscillator bank entrainment (Toiviainen 1998, Large & Kolen 1994, Ohya 1994, Miller, Scarborough & Jones 1989).
- Frequency of Onset Function: (Scheirer 1998, Klapuri et al. 2006, Peeters 2007, Davies & Plumbley 2007).
- Dynamic time warping of beat interval (Dixon 2001, Ellis 2007).
- Multiresolution Approaches (Todd 1994, Todd, O'Boyle & Lee 1999, Smith & Honing 2008).

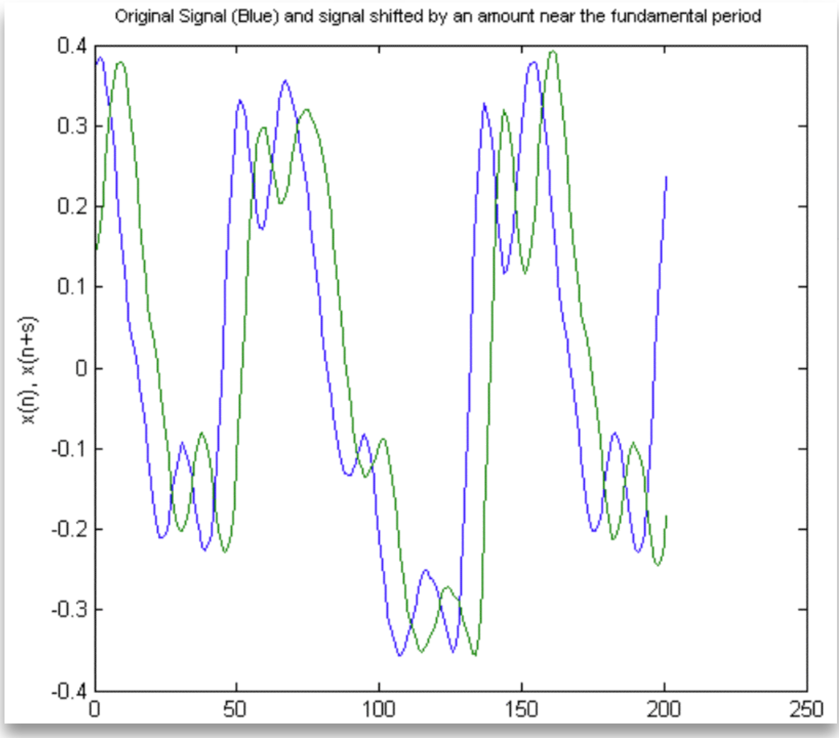
Approaches to beat tracking considered

- Autocorrelation
  - Finding Periodicity in the ODF.
- Beat Spectrum approaches:
  - Spectrum of the ODF.
  - Multi-resolution representation of ODF.
- Dynamic Programming approaches.
  - Efficient selection of correct beat interval.

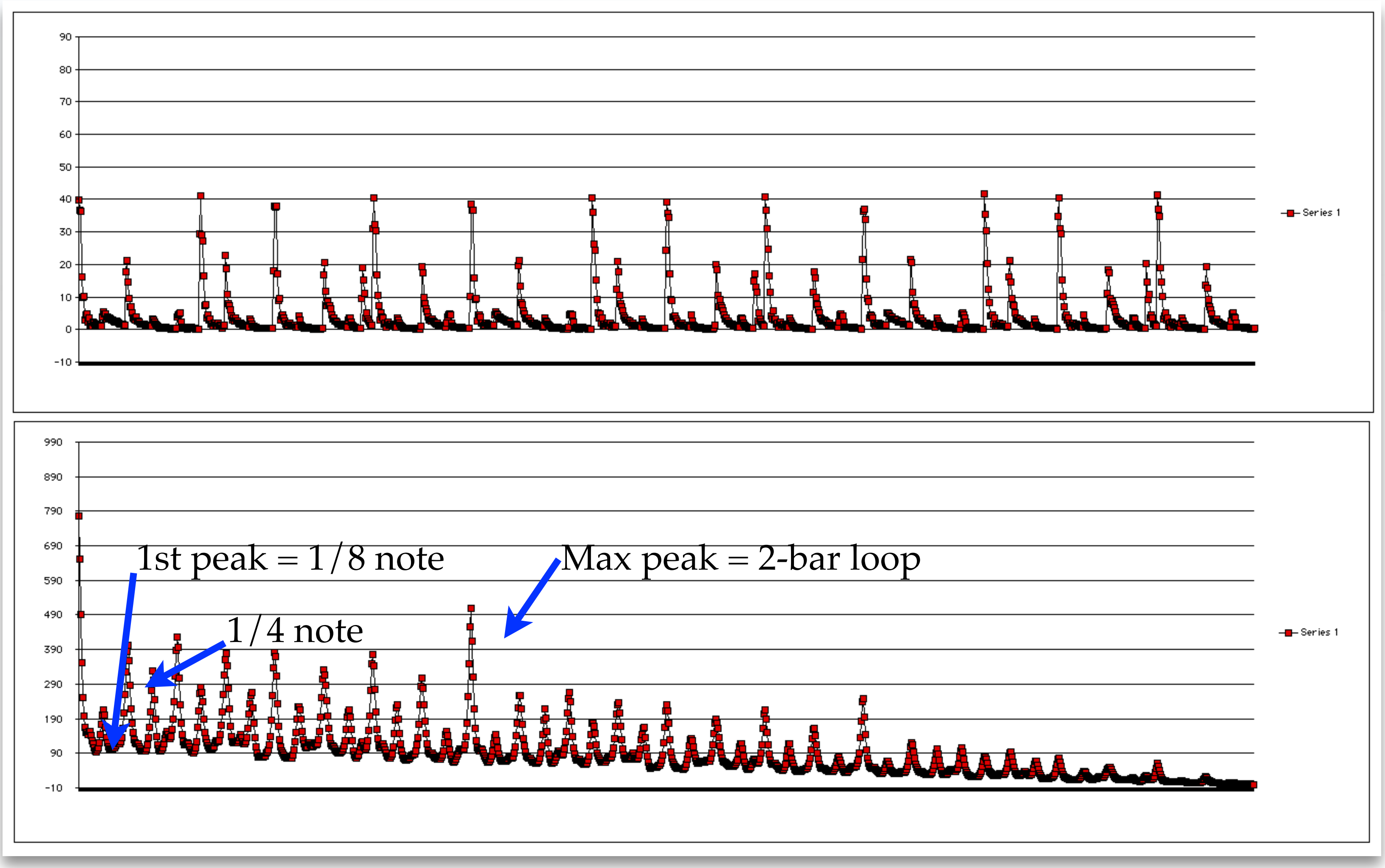
Autocorrelation of ODF

- AC peaks ⇒ time lags

  where signal is most similar to itself.
- Captures periodicities of ODF.
- Does not capture rubato well.
- OK for metronomic music, not for those with variation in tempo.



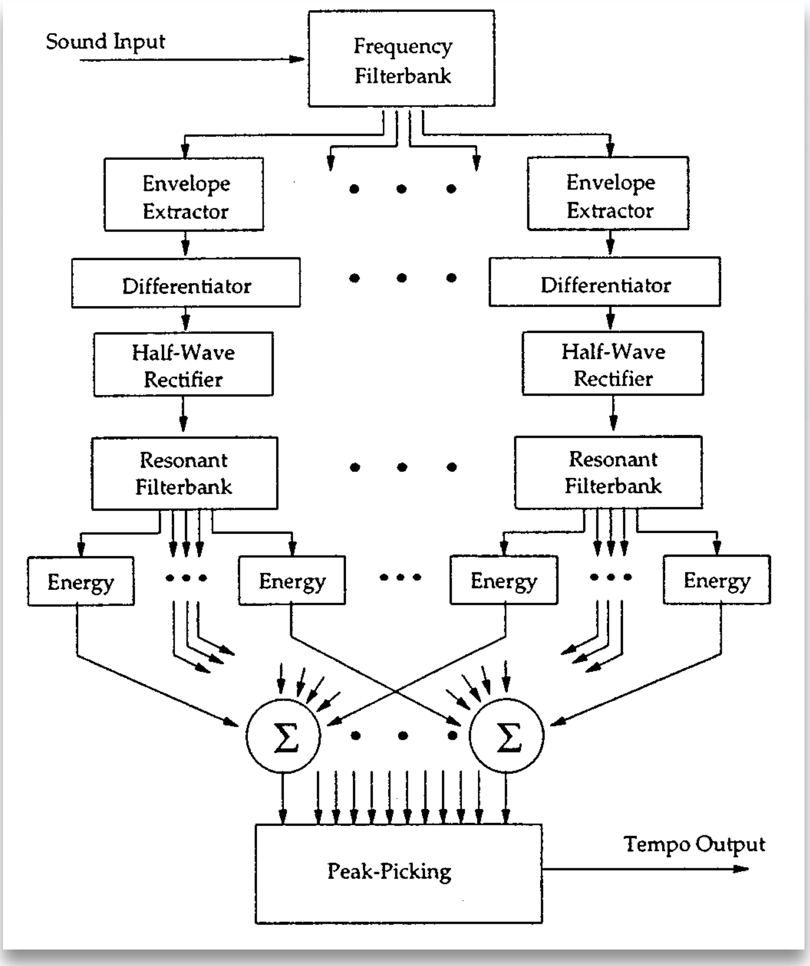
Windowed RMS and its Autocorrelation (for drum loop)



Beat spectrum methods (Scheirer 1998)

- Filterbanks of tuned resonators (i.e. “rhythmic reverb”) of the ODF.
- Resonator whose resonant F matches rate of ODF modulation will phase-lock.
- Resonator outputs of common freq summed across subbands:

$$T = \arg\max_r \sum_s^S F_{rs}$$



Beat Tracking by Peeters (2007)

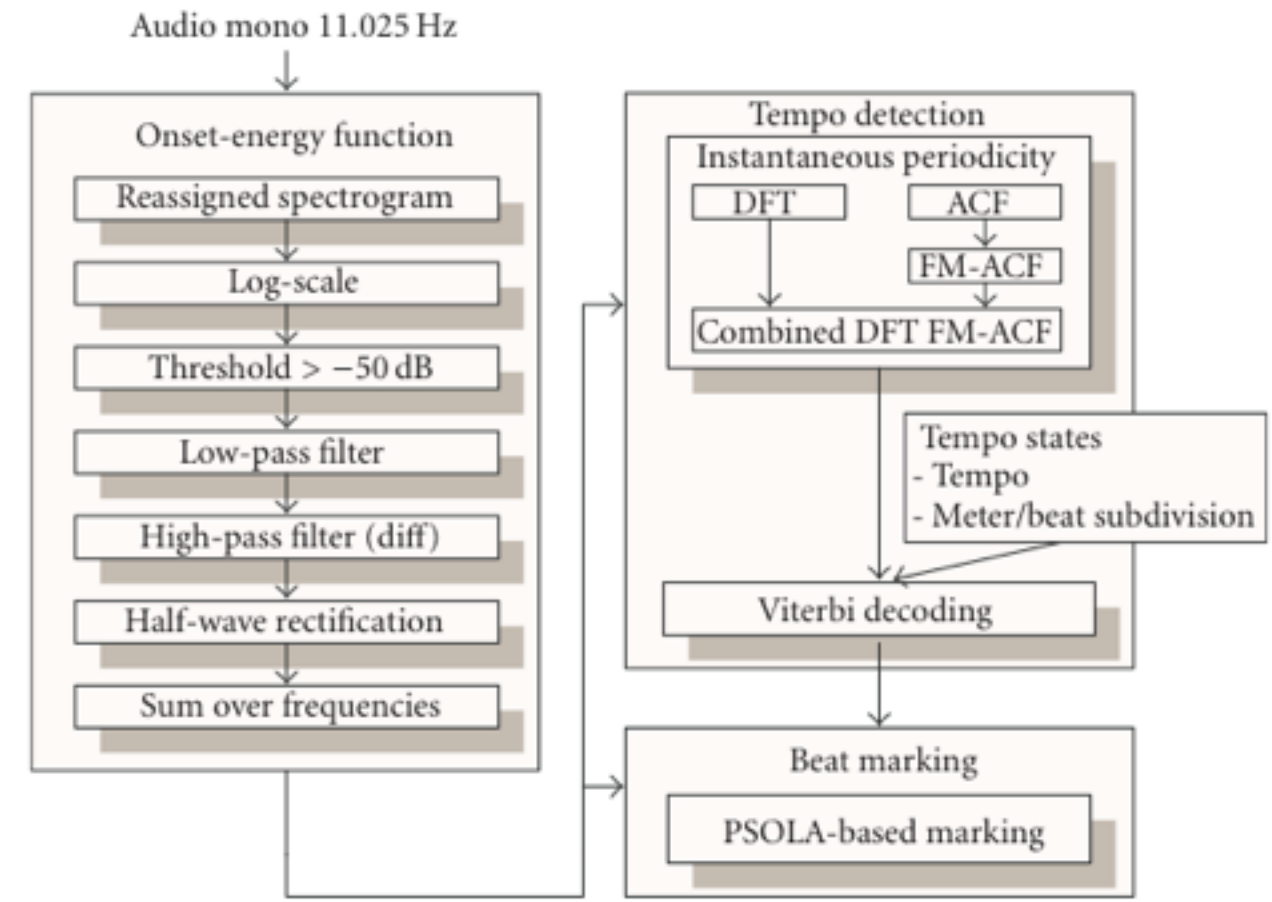


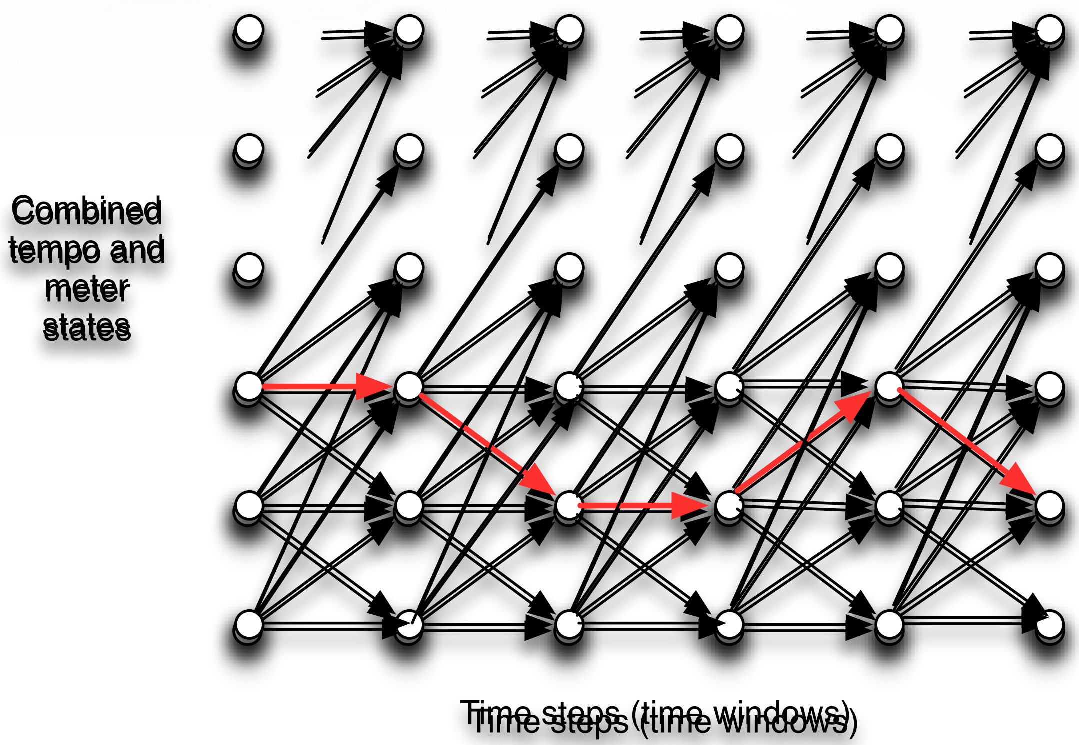
FIGURE 1: Flowchart of our system for tempo, meter estimation, and beat marking.

Peeters 2007

- Filtered, rectified spectral energy envelope
  - Onset detection function.
- Combined Fourier & autocorrelation analysis
  - DFT of ODF, ACF of ODF
  - ACF result mapped into Fourier domain.
  - DFT * Freq(ACF) – disambiguates periodicities.
  - Octave errors occur in two different domains.

Peeters 2007

- Viterbi decoding of joint estimates of meter and tempo.



Dynamic Programming (Ellis 2007)

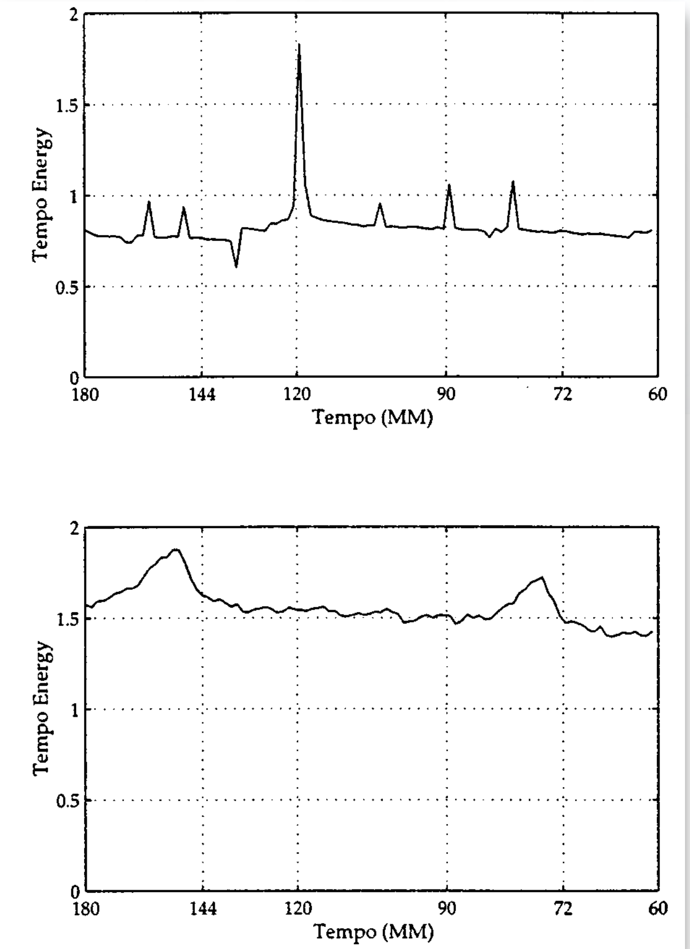
- Goal to generate beat times that match onsets and have near constant IBI.

$$C(\{t_i\}) = \sum_{i=1}^{N} O(t_i) + \alpha \sum_{i=2}^{N} F(t_i - t_{i-1}, \tau_p)$$

- F(Δt, τ) = – log(actual IBI/ideal IBI)^2.
- Ideal IBI from tempo estimation from weighted autocorrelation.
- Recursively calculates max C*(t) starting from t=0–2τ, finding times of max(F + C*(τ)).
- Chooses final max C*(t) from last interval, backtraces the saved times.

Beat Histograms

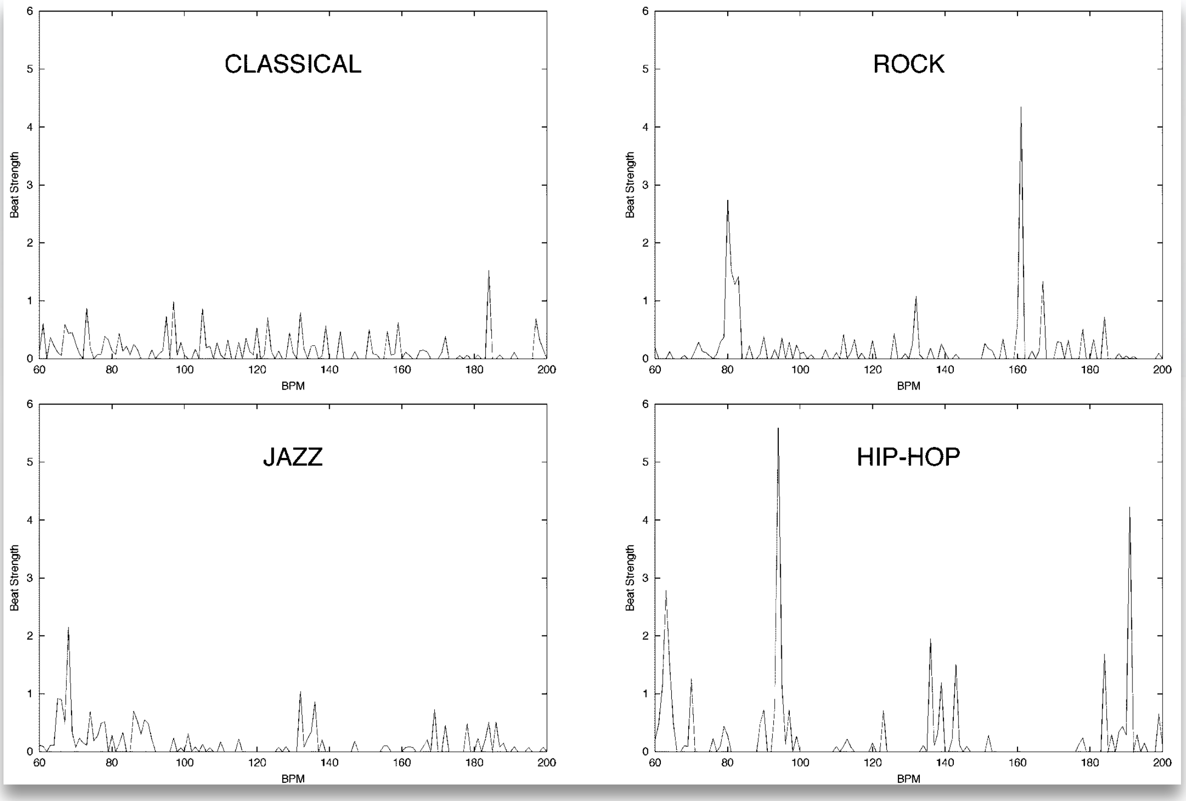
- Summarises rhythmic behaviour of a piece for similarity measures, classification etc.
- Pampalk, Dixon & Widmer (2003)
  - Uses summation of comb filters of Scheirer, not just argmax, for comparison.
  - Tempo histogram is weighted using a preference model (van Noorden & Moelants 1999).
  - PCA used to reduce 2000 ➪ 60 dimensions for matching.



(from Scheirer 1998)

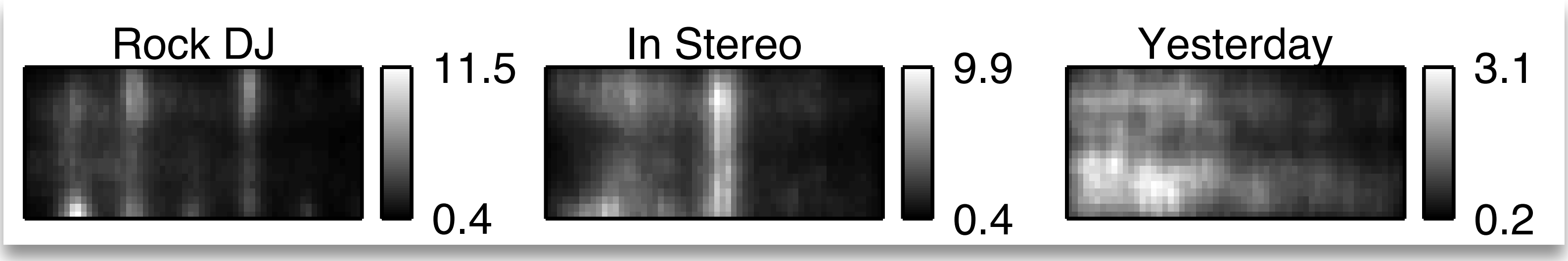
Beat Histograms (Tzanetakis and Cook, 2002)

- Similar approach using Autocorrelation.
- Add the amplitudes of the top 3 AC peaks to histogram at each frame.
- Beat histograms are reducible to single features including sum and peak/mean.



Fluctuation Patterns

- Also summarises rhythmic behaviour.
- FFT of envelope: the fluctuation (AM) frequency of the perceived loudness of critical bands (log spectral) (represented on the Bark scale).
- 20 Bark x 60 BF matrix ➪ PCA for matching



Median of the fluctuation patterns of examples of (L-R) Heavy Metal, Dance and Pop. Y axis shows critical bands (Bark 1-20), X axis shows beat frequencies 0-10Hz (0-600BPM) From Pampalk, Rauber & Merkl, (2002)

Meter estimation

- Requires measure (“bar”) period and phase (downbeat) identification.
- Measure period reasonably successful, albeit with octave errors.
- Downbeat identification much harder!
- Genre dependent.

Meter Estimation Systems

c. 2006

Table 1. Characteristics of some meter estimation systems.

Reference	Input	Output	Technique
Temperley & Sleator (1999)	MIDI	Meter, time quantization	Rule-based approach; implementation of the preference rules in (Lerdahl et al.,1983)
Dixon (2001)	MIDI, audio	Tactus	First find periods using IOI histogram, then phases using multiple agents
Raphael (2001)	MIDI, audio	Tactus, time quantization	Probabilistic generative model for onset times; MAP estimation (Viterbi)
Cemgil & Kappen (2003)	MIDI	Tactus, time quantization	Probabilistic generative model for onset times; sequential Monte Carlo methods
Goto & Mur-aoka (1995, 1997)	Audio	Meter	Extract onset components; IOI histogram; multiple tracking agents
Scheirer (1998)	Audio	Tactus	Bank of comb filters to analyze periodicity of power envelopes at six subbands
Laroche (2001)	Audio	Tactus, swing	Extract discrete onsets; maximum-likelihood estimation
Sethares & Staley (2001)	Audio	Meter	Calculate RMS-energies at 1/3-octave subbands; apply a periodicity transform
Gouyon et al. (2002)	Audio	Tatum	First find periods (IOI histogram), then phases by matching isochronous pattern
Klapuri et al. (to appear)	Audio	Meter	Measure degree of accentuation; bank of comb filters; probabilistic model

Joint estimation of chord change and downbeat (Papadopoulos & Peeters 2008)

- Hidden Markov Model:
  - States: 24 Major & Minor triads * 4 positions within the Measure (pim) for (4/4 time signature).
  - Computes chroma features at each beat.
  - Assumes independence between beat position and chord type: P(O|s) = P(O|c) P(O|pim)
  - Transition probabilities enforce sequential beats & likelihood of chord transitions.
- Optimal state determined by Viterbi decoding.
  - Chord progression detection improved using metrical knowledge.
  - Identification of downbeats aided by harmonic information.

Review

- Modeling rhythm requires representing perception.
- Onset detection functions capture significant events.
- Multiple approaches to beat-tracking represent competing perceptual models.
- Beat-tracking enables higher-level rhythmic features (FP, BH).
- Beat-tracking enables multi-modal estimation (e.g., down-beat from chords).
- References: http://ccrma.stanford.edu/workshops/mir2011/BeatReferences.pdf