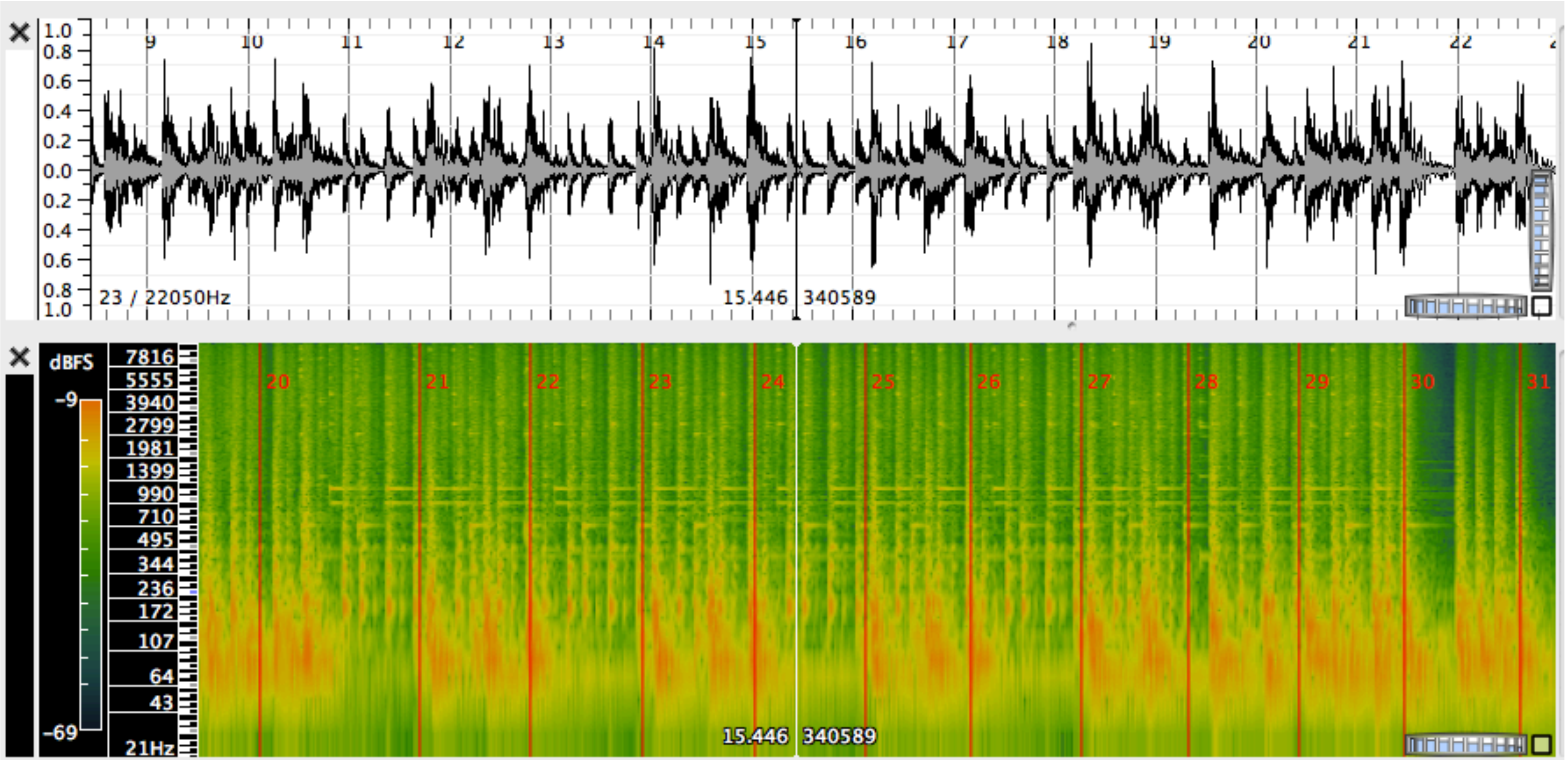
Lecture 10: Beat Tracking

1. Rhythm Perception
2. Onset Extraction
3. Beat Tracking
4. Dynamic Programming

Dan Ellis
Dept. Electrical Engineering, Columbia University
dpwe@ee.columbia.edu   http://www.ee.columbia.edu/~dpwe/e4896/

1. Rhythm Perception

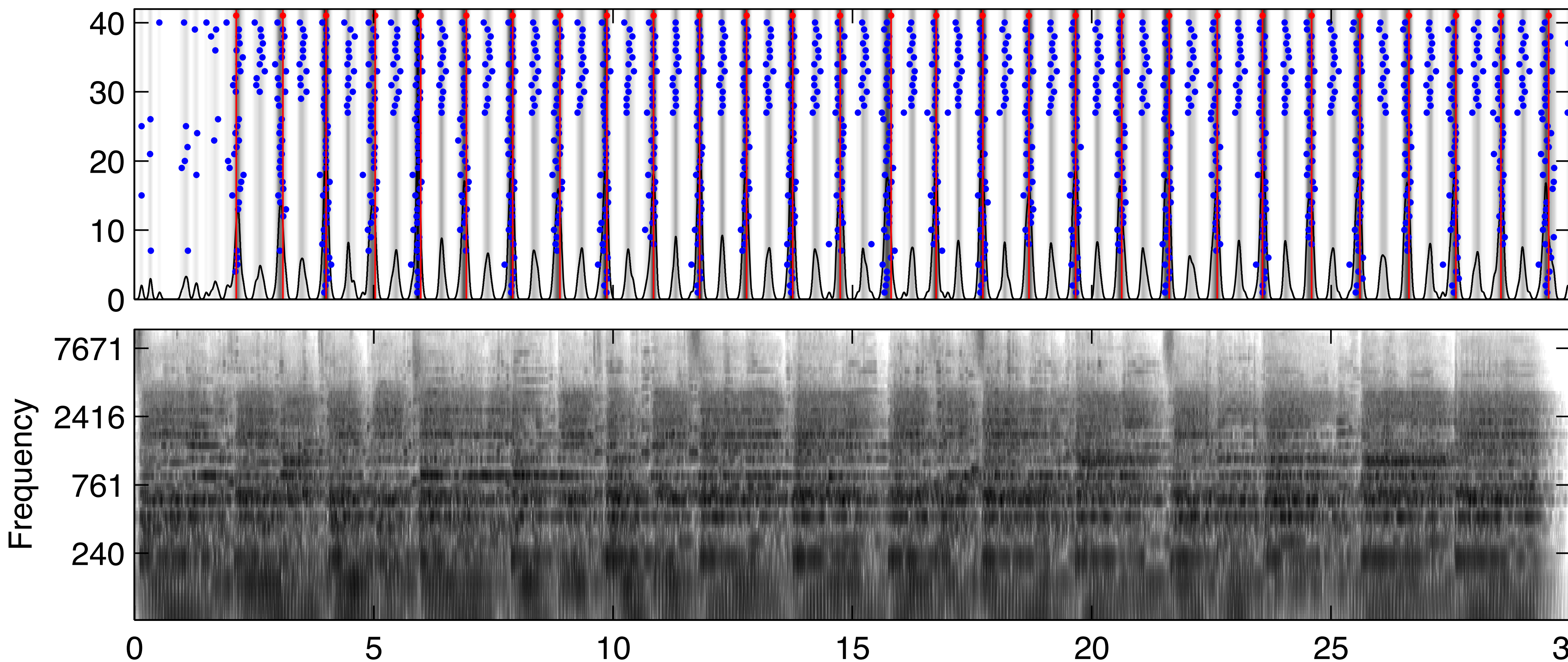
- What is rhythm?
  - aspects, origin?



Rhythm Perception Experiments

McKinney & Moelants 2006

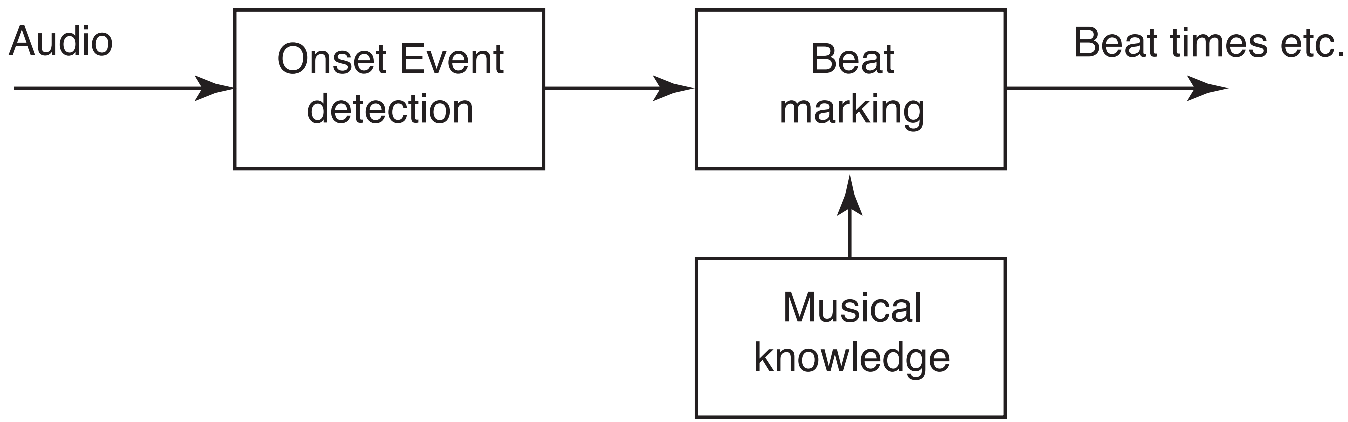
- Tapping experiment
    - ambiguous; hierarchy



mirex06/train10 Harnoncourt

Rhythm Tracking Systems

- Two main components
  - front end: extract 'events' from audio
  - back end: find plausible beat sequence to match



- Other outputs
  - tempo
  - time signature
  - metrical level(s)

2. Onset detection

Bello et al. 2005

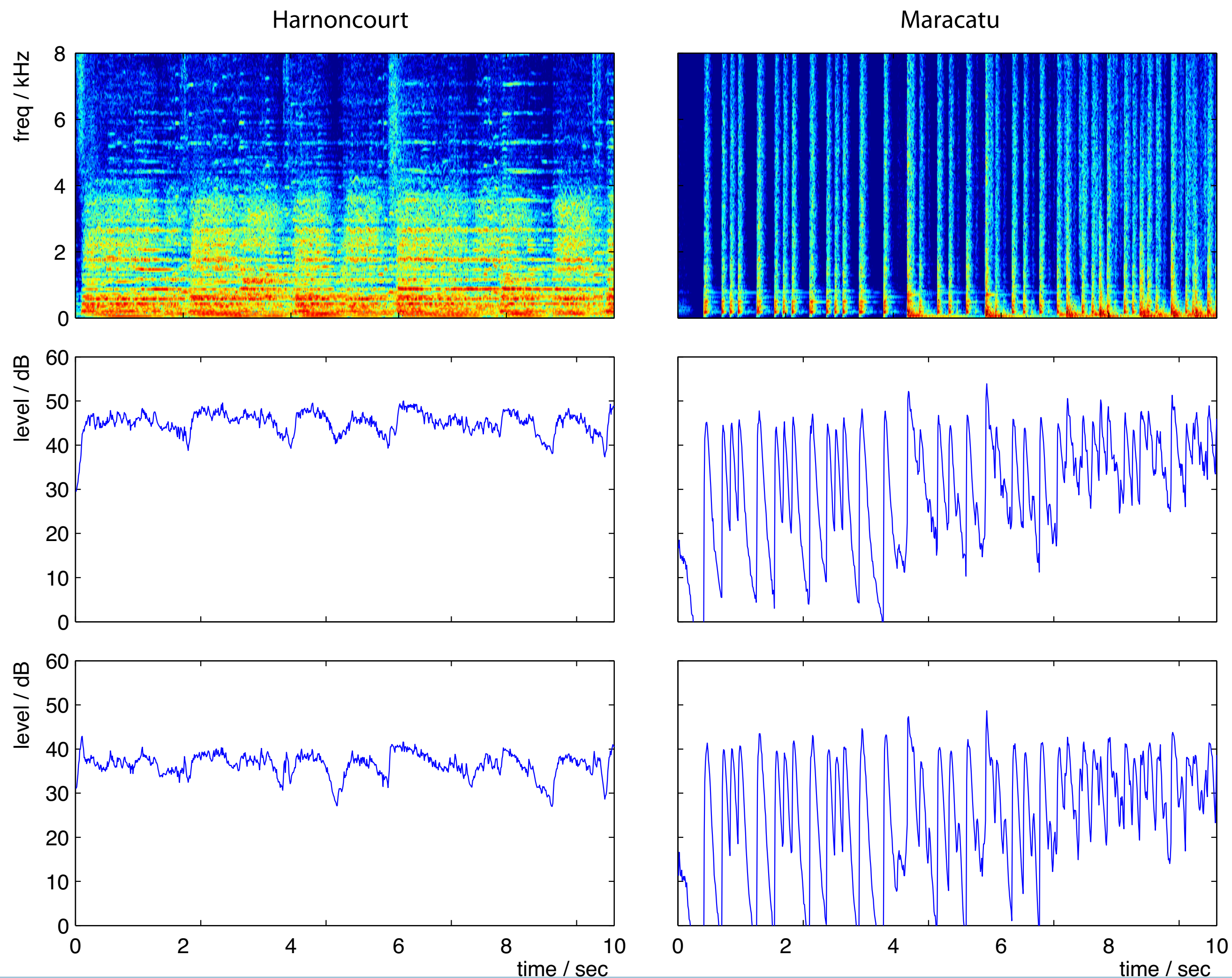
- Simplest thing is **energy envelope**

$$e(n_0) = \sum_{n=-W/2}^{W/2} w[n]\,|x(n+n_0)|^2$$

- emphasis on high frequencies?

$$\sum_f |X(f,t)|$$

$$\sum_f f \cdot |X(f,t)|$$



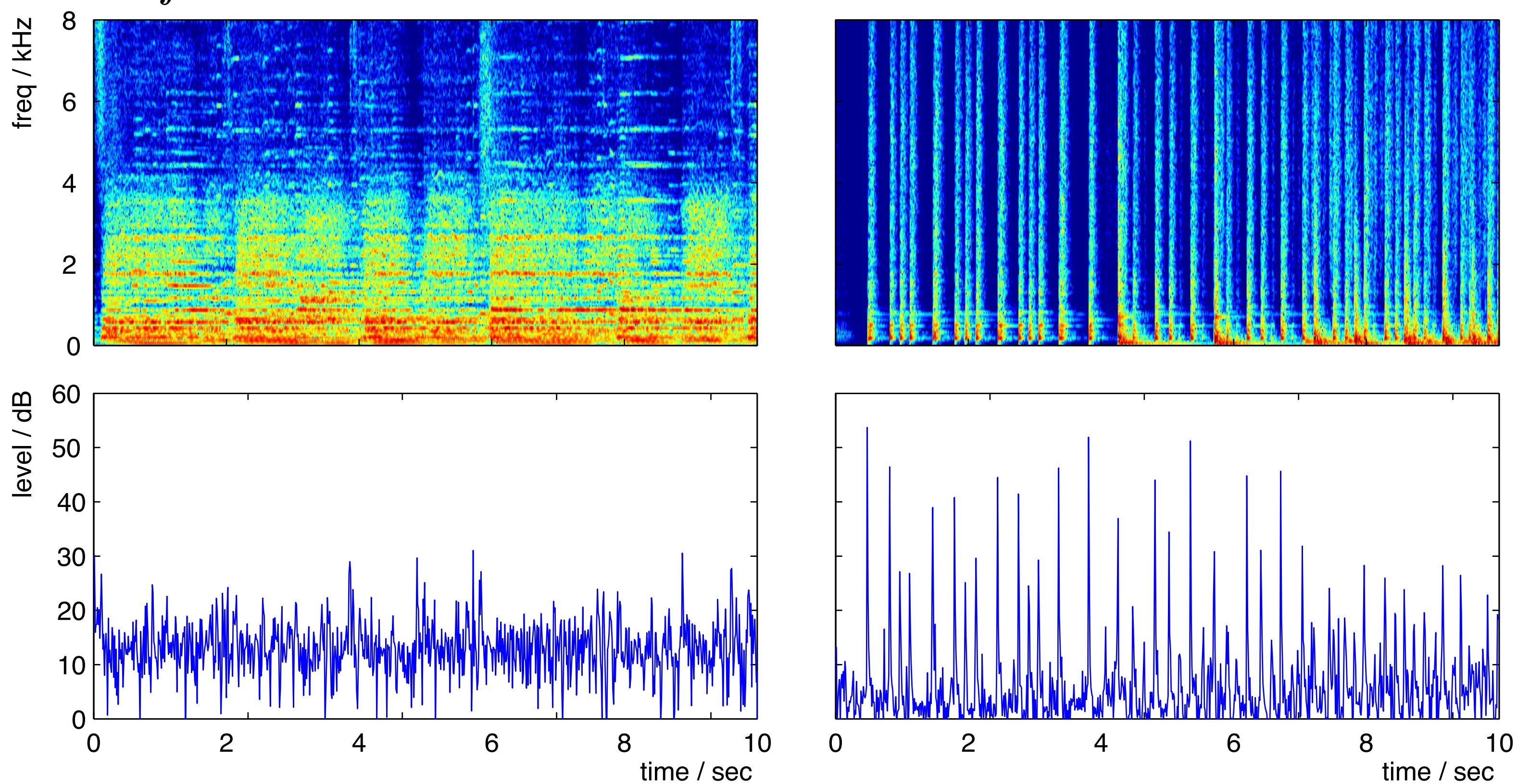
Multiband Derivative

Puckette et. al 1998

- Sometimes energy just "shifts"
  - calculate & sum onset in multiple bands
  - use ratio instead of difference - normalize energy

$$o(t) = \sum_f W(f)\max(0, \frac{|X(f,t)|}{|X(f,t-1)|} - 1)$$

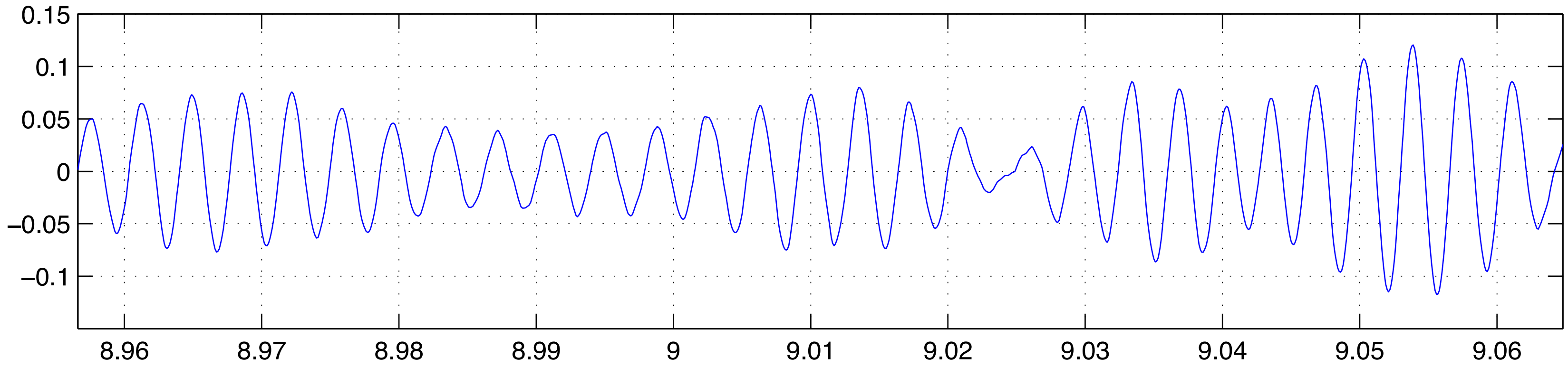
bonk~



Phase Deviation

Bello et al. 2005

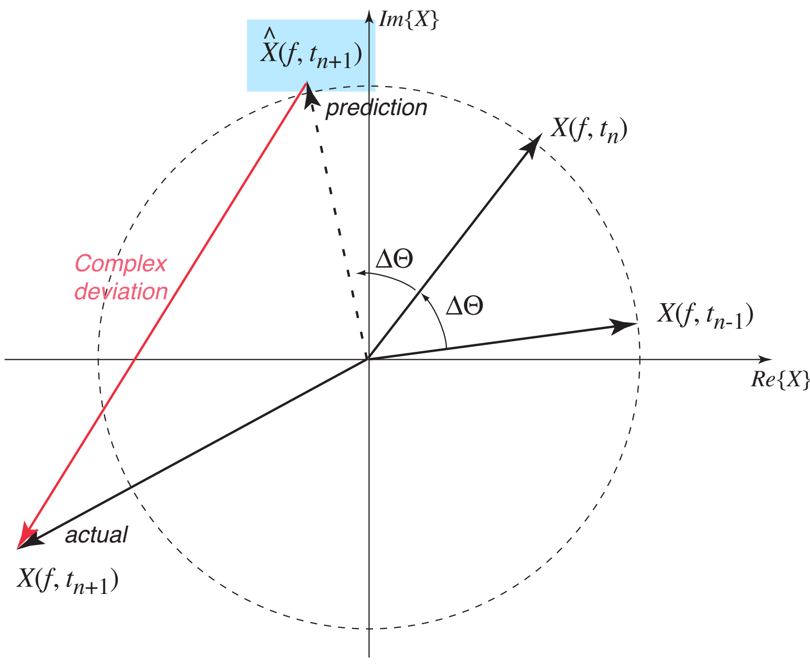
- When amplitudes don't change much, phase discontinuity may signal new note



- Can detect by comparing actual phase with extrapolation from past

$$\hat{X}(f, t_{n+1}) = X(f, t_n)\frac{X(f, t_n)}{X(f, t_{n-1})}$$

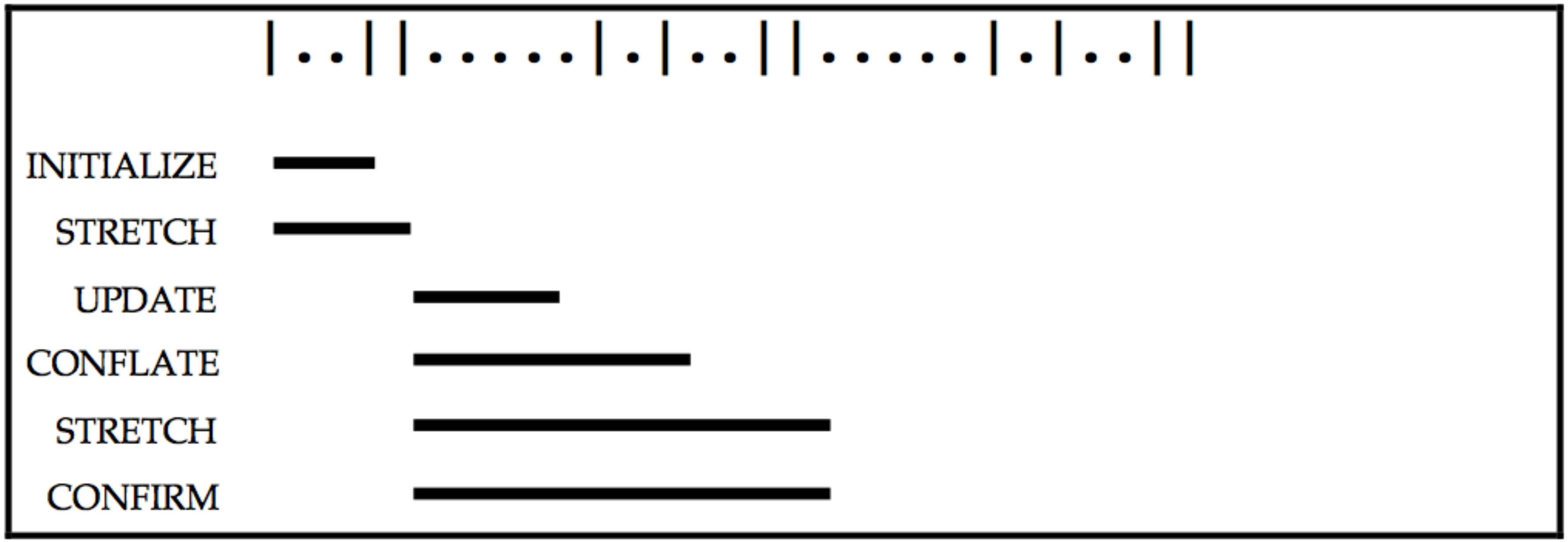
  - combine with amplitude?



3. Rhythm Tracking

Desain & Honing 1999

- **Earliest systems were rule based**
  - based on musicology *Longuet-Higgins and Lee, 1982*
  - inspired by linguistic grammars - Chomsky



  - input: event sequence (MIDI)
  - output: quarter notes, downbeats

Resonators

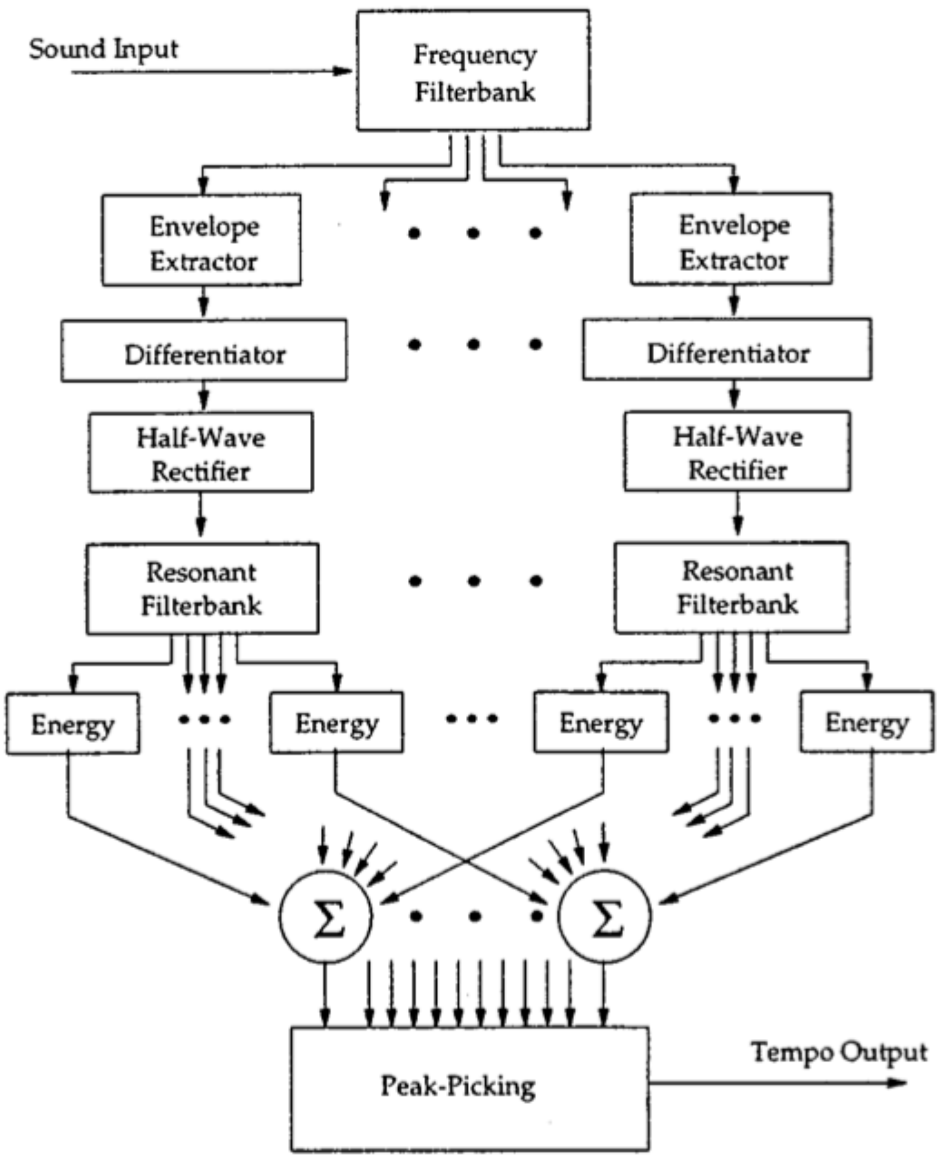
Scheirer 1998

- How to address:
  - build-up of rhythmic evidence
  - "ghost events"
  - (audio input)

- Seems more like a comb filter...
  - resonant filterbank of

$$y(t) = \alpha y(t - T) + (1 - \alpha)x(t)$$

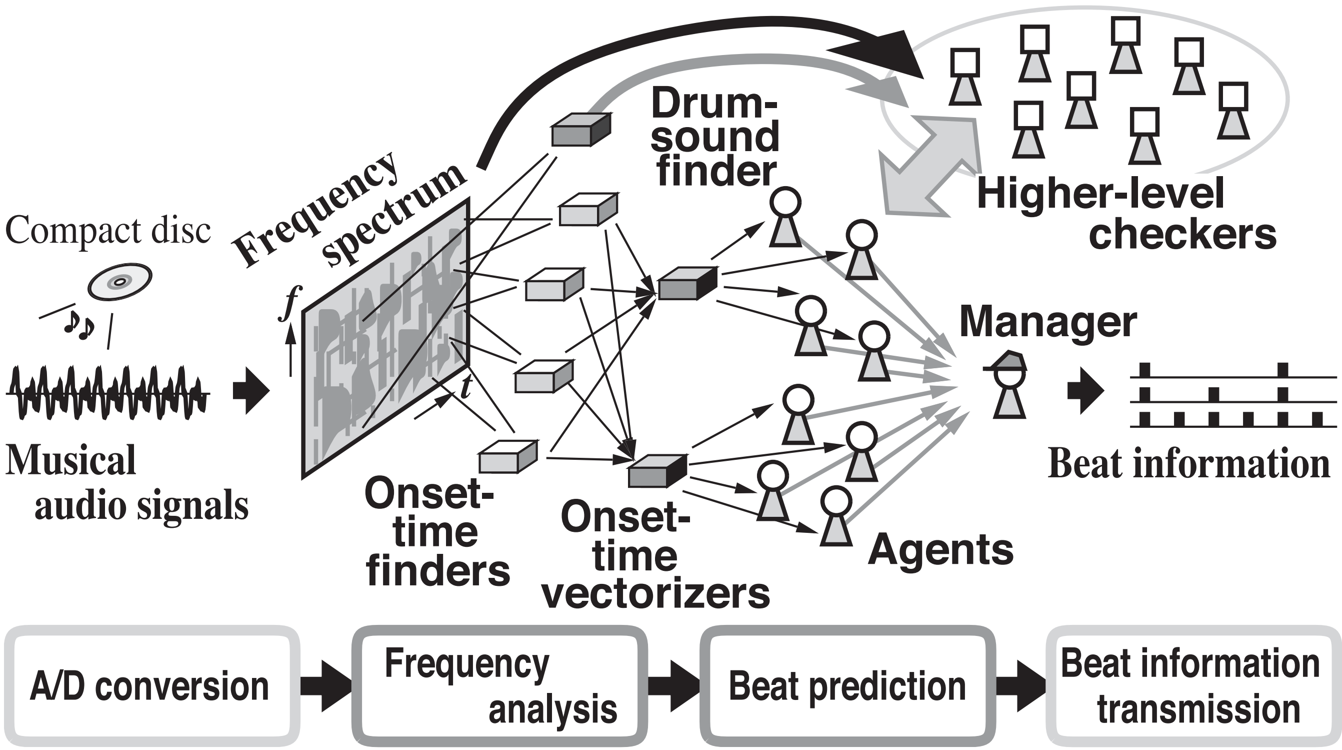
for all possible T



Multi-Hypothesis Systems

Goto & Muraoka 1994
Goto 2001
Dixon 2001

- Beat is ambiguous
  → develop several alternatives



- inputs: music audio
- outputs: beat times, downbeats, BD/SD patterns...

4. Dynamic Programming

Ellis 2007

- Re-cast beat tracking as optimization:
  Find beat times $\{t_i\}$ to maximize

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

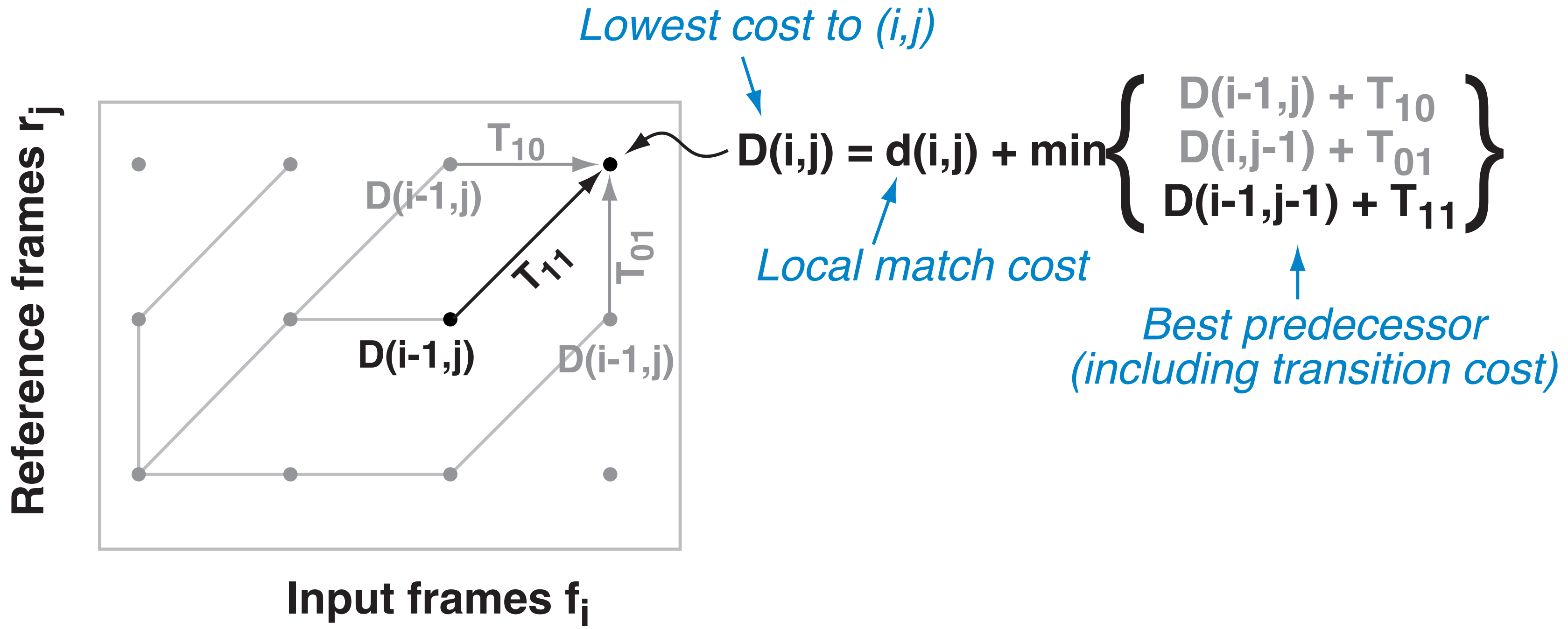
  - $O(t)$ is onset strength function
  - $F(\Delta t, \tau)$ is tempo consistency score e.g.

$$F(\Delta t, \tau) = -\left(\log \frac{\Delta t}{\tau}\right)^2$$

- Looks like an exponential search over all $\{t_i\}$
  - ... but Dynamic Programming saves us

Dynamic Programming (DP)

- DP is a general algorithm for optimizing "optimal substructure" problems
  - i.e. where optimal total solution can be built from optimal partial solutions
- e.g. best path through cost matrix

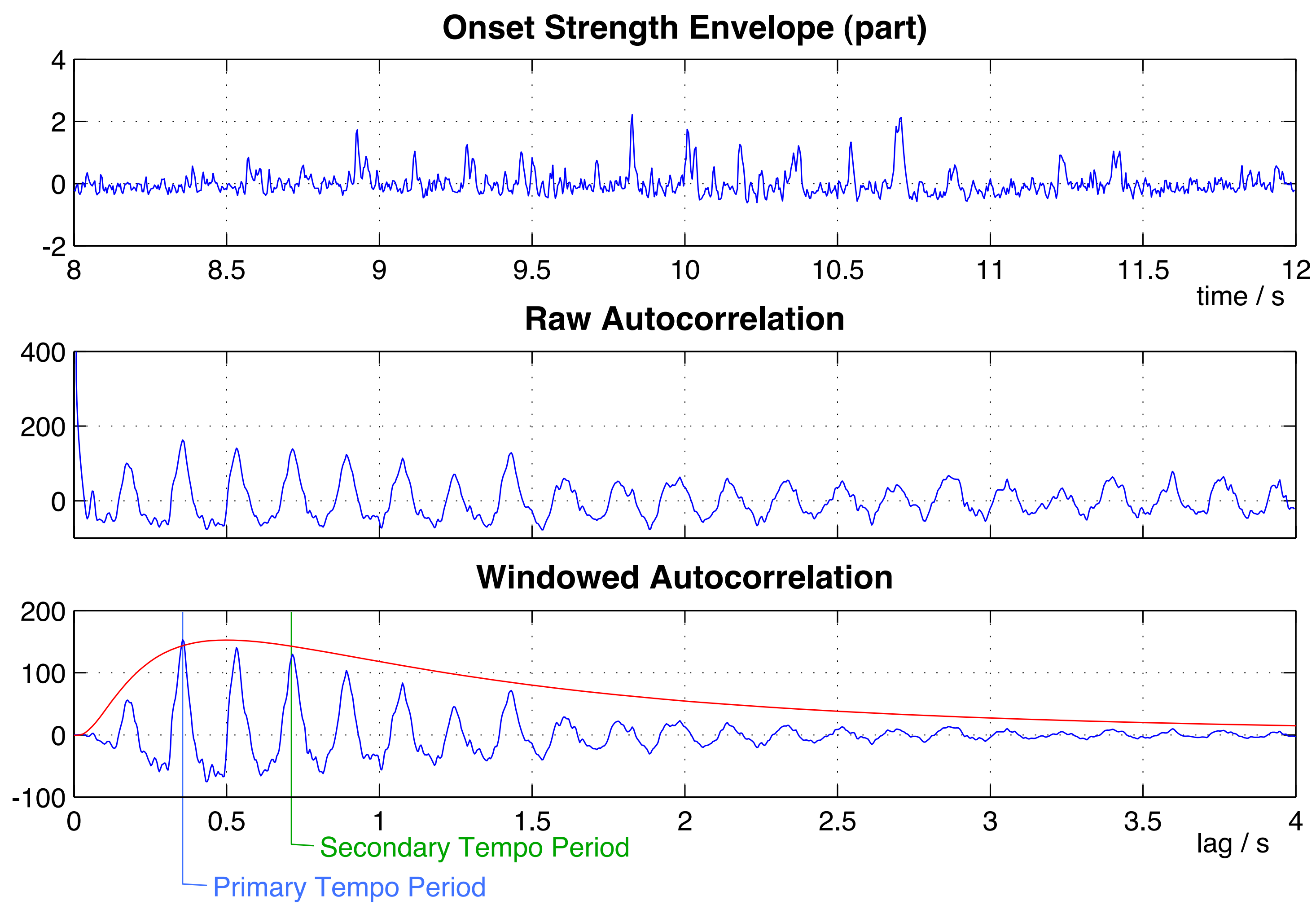


  - path after (i,j) is independent of how we got there

Tempo Estimation

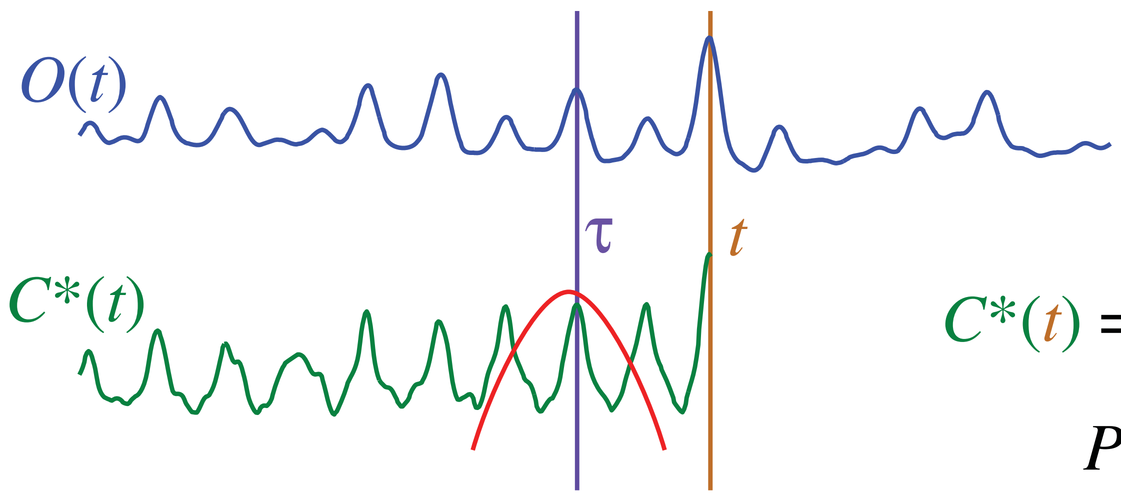
- Algorithm needs global tempo period τ
  - otherwise problem is not "optimal substructure"

- Pick peak in onset envelope autocorrelation
  - after applying "human preference" window
  - check for subbeat



Beat Tracking by DP

- To optimize $C(\{t_i\}) = \sum_{i=1}^{N} O(t_i) + \alpha \sum_{i=2}^{N} F(t_i - t_{i-1}, \tau_p)$
  - define $C^*(t)$ as best score up to time t
  - then build up recursively (with traceback $P(t)$)



$$C^*(t) = O(t) + \max_{\tau}\{\alpha F(t - \tau, \tau_p) + C^*(\tau)\}$$

$$P(t) = \operatorname*{argmax}_{\tau}\{\alpha F(t - \tau, \tau_p) + C^*(\tau)\}$$

  - final beat sequence $\{t_i\}$ is best C^* + back-trace

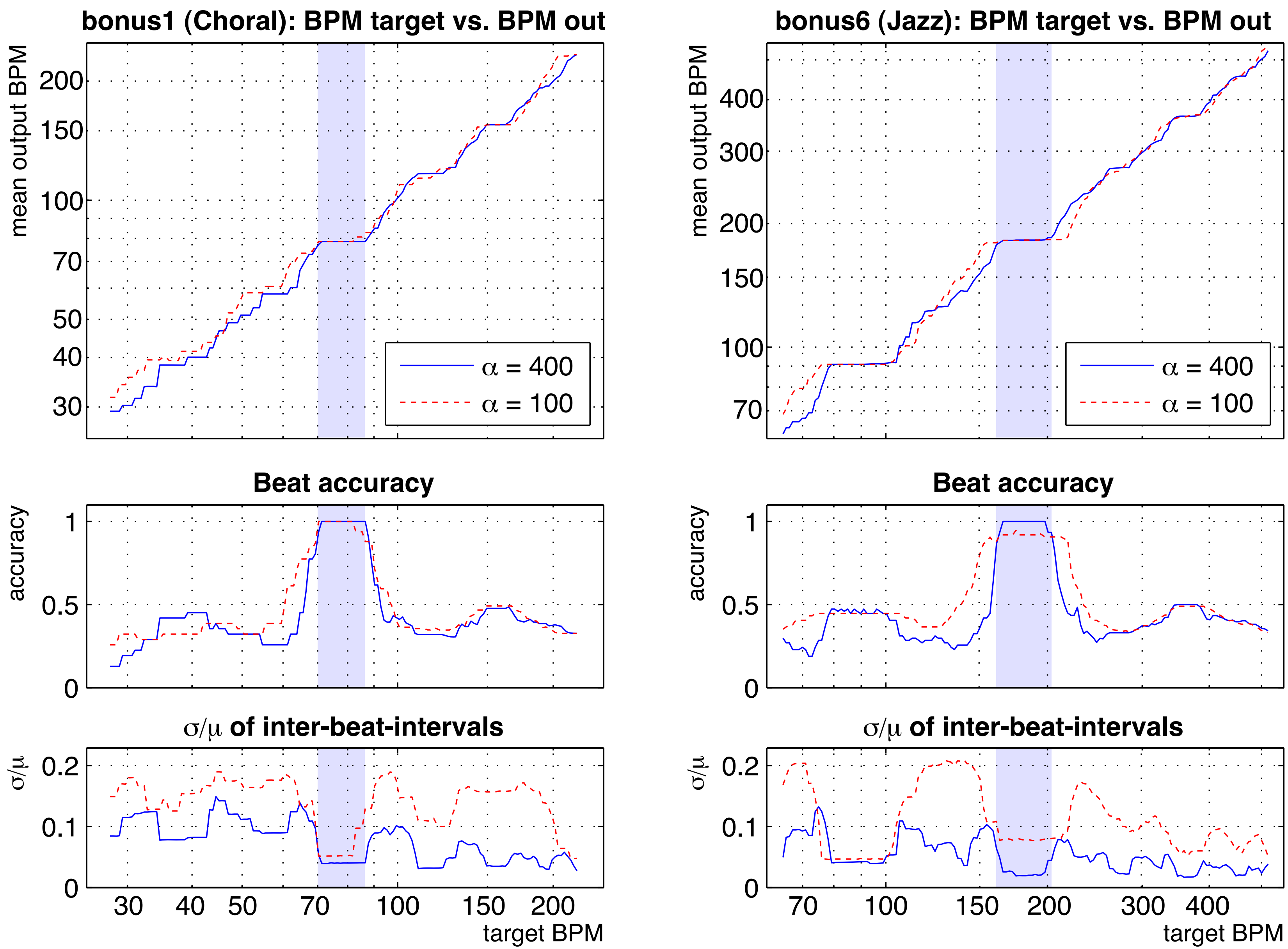
beatsimple

- Beat tracking in 15 lines of Matlab

```
function beats = beat_simple(onset, osr, tempo,
alpha)
% beats = beat_simple(onset, osr, tempo, alpha)
%   Core of the DP-based beat tracker
%   <onset> is the onset strength envelope at
frame rate <osr>
%   <tempo> is the target tempo (in BPM)
%   <alpha> is weight applied to transition cost
%   <beats> returns the chosen beat sample times
(in sec).
% 2007-06-19 Dan Ellis dpwe@ee.columbia.edu

if nargin < 4; alpha = 100; end

% backlink(time) is best predecessor for this
point
% cumscore(time) is total cumulated score to this
point
localscore = onset;
backlink = -ones(1,length(localscore));
cumscore = zeros(1,length(localscore));

% convert bpm to samples
period = (60/tempo)*osr;

% Search range for previous beat
prange = round(-2*period):-round(period/2);
% Log-gaussian window over that range
txwt = (-alpha*abs((log(prange/-period)).^2));

for i = max(-prange + 1):length(localscore)

  timerange = i + prange;

  % Search over all possible predecessors
  % and apply transition weighting
  scorecands = txwt + cumscore(timerange);
  % Find best predecessor beat
  [vv,xx] = max(scorecands);
  % Add on local score
  cumscore(i) = vv + localscore(i);
  % Store backtrace
  backlink(i) = timerange(xx);

end

% Start backtrace from best cumulated score
[vv,beats] = max(cumscore);
% .. then find all its predecessors
while backlink(beats(1)) > 0
  beats = [backlink(beats(1)),beats];
end

% convert to seconds
beats = (beats-1)/osr;
```

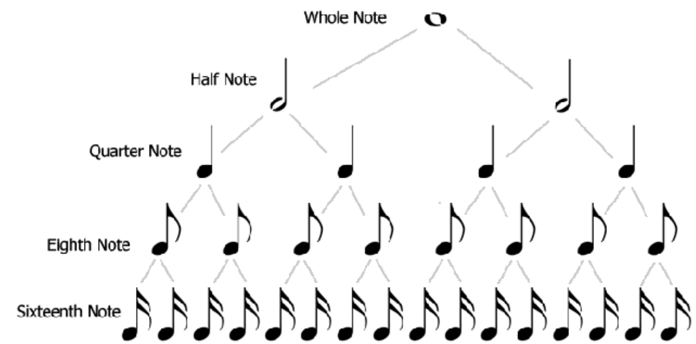
Results

- Verify against human tapping data
    - vary tradeoff weight α
    - vary tempo estimate τ

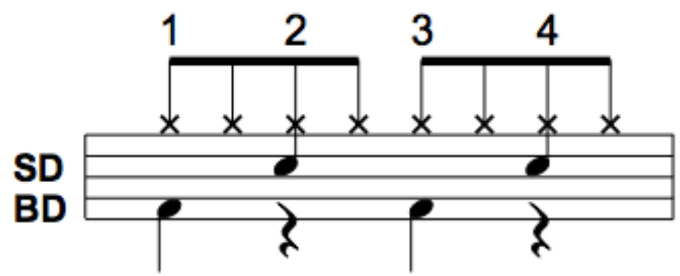


Downbeat Detection

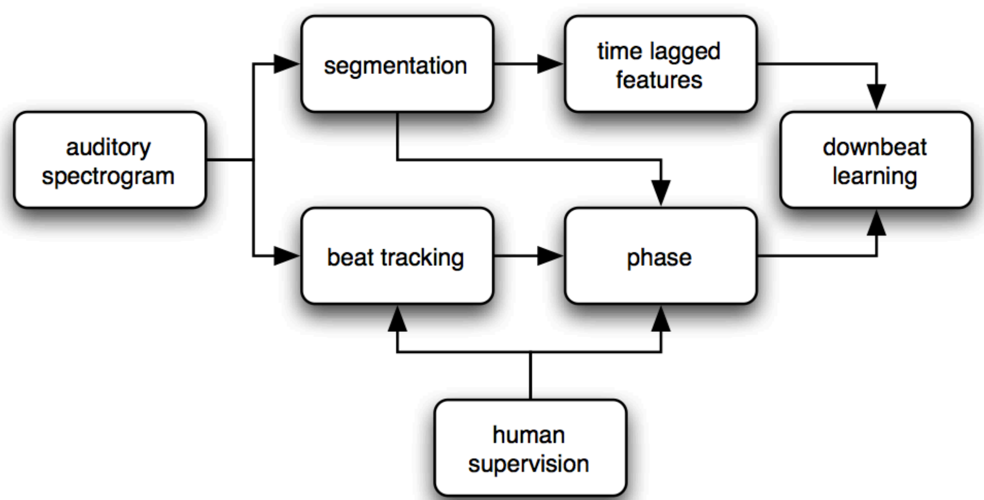
- **Downbeat = start of "bar"**
  - one level up in metrical hierarchy



- **Approaches**
  - Goto'94 BTS:
    Pop music
    SD/BD template



  - Jehan'05:
    Trained classifier



Summary

- Rhythm perception
  Innate and strong, hierarchic
- Beat tracking models
  Need to account for buildup & persistence
- Dynamic Programming
  Neat way to maintain multiple hypotheses

References

- J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, M. B. Sandler, "A Tutorial on Onset Detection in Music Signals," *IEEE Tr. Speech and Audio Proc.*, vol.13, no. 5, pp. 1035-1047, September 2005.
- P. Desain & H. Honing, "Computational models of beat induction: The rule-based approach," *J. New Music Research*, vol. 28 no. 1, pp. 29-42, 1999.
- Simon Dixon, "Automatic extraction of tempo and beat from expressive performances," *J. New Music Research*, vol. 30 no. 1, pp. 39-58, 2001.
- D. Ellis, "Beat Tracking by Dynamic Programming," *J. New Music Research*, vol. 36 no. 1, pp. 51-60, March 2007.
- Tristan Jehan, "Creating Music By Listening," Ph.D Thesis, MIT Media Lab, 2005.
- Masataka Goto & Yoichi Muraoka "A Beat Tracking System for Acoustic Signals of Music," *ACM Multimedia*, pp.365-372, October 1994.
- Masataka Goto, "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds," *J. New Music Research*, vol. 30 no. 2, pp.159-171, June 2001.
- M. F. McKinney and D. Moelants, "Ambiguity in tempo perception: What draws listeners to different metrical levels?" *Music Perception*, vol. 24 no. 2 pp. 155-166, December 2006.
- M. Puckette, T. Apel, D. Zicarelli, "Real-time audio analysis tools for Pd and MSP," *Proc. Int. Comp. Music Conf.*, Ann Arbor, pp. 109–112, October 1998.
- Eric. D. Scheirer, "Tempo and beat analysis of acoustic musical signals," *J. Acoust. Soc. Am.*, vol. 103, pp. 588-601, 1998.