

An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds

Published by Masataka Goto
Journal of New Music Research
2001, Vol. 30, No. 2, pp.159-171

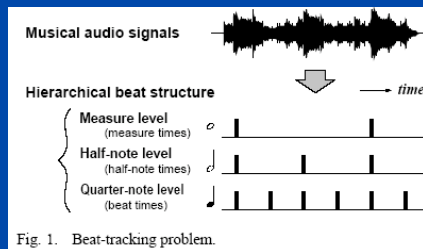
Presented by Edward Stein
ISE575: Feb 3, 2005

Types of Beat Detection

- Beats per Minute (BPM)
 - A measure of tempo with no universal base.
 1. Quarter-notes per minute (MM)
 2. Number of Drum hits per minute
 3. What counts as a Drum hit?
- Beat Time and Phase
 - Predicts inter-beat intervals (IOI) or the time expected between beat events.
 - Phase is a measure of the difference between the implied beat and what is expressed in performance

Types of Beat Detection (2)

- Hierarchical Beat Structure Detection
 - Captures both what is implied and what is explicitly performed by an artist in a song



Issues in Recognizing Beat Structure

1. Detecting Beat-tracking Cues in Audio Signals
 - Onset times
 - Chord Changes
 - Drum Patterns
2. Interpreting the Cues to Infer the Beat Structure
3. Dealing with the Ambiguities of Interpretations

Bottom-Up Vs. Top Down

- Bottom Up
 1. The classification of audio events especially in a highly polyphonic piece is difficult and computationally expensive.
 2. Errors in interpretation can guide you in the wrong direction.
- Top Down
 1. Starts with multiple hypotheses of musical structure and finds reinforcing characteristics.
 2. Requires defined knowledge of musical structure matching type or genre of input data.

The Rules

- Time Signature of 4/4
- 61-185 M.M. With Drums (61-120 M.M. w/o)
- Drums
 - (a-1) "A frequent inter-onset interval is likely to be the inter-beat interval"
 - (a-2) "Onset times tend to coincide with beat times (i.e. sounds are likely to occur on beats)."
 - (c-1) "The beginning of the input drum pattern indicates a half-note time"
 - (c-2) "The input drum pattern has the appropriate inter-beat interval."
 - (x-1) Assumes the snare (if present) will occur on the 2nd and 4th quarter-note in a measure.

Note: M.M. stands for Mälzel's Metronome: the number of quarter notes per minute

The Rules (2)

- Chords
 - (b-1) "Chords are more likely to change on beat times than on other positions"
 - (b-2) "Chords are more likely to change on half-note times than on other positions of beat times"
 - (b-3) "Chords are more likely to change at the beginning of measures than at other positions of half-note times"

- Summary

Table 1. Musical knowledge selection for music with drum-sounds and music without drum-sounds.

Beat structure	Without drums	With drums
Measure level	quarter-note chord-change possibility (knowledge (b-3))	quarter-note chord-change possibility (knowledge (b-3))
Half-note level	quarter-note chord-change possibility (knowledge (b-2))	drum pattern (knowledge (c-1))
Quarter-note level	eighth-note chord-change possibility (knowledge (b-1))	drum pattern (knowledge (c-2))

Onset Vectors

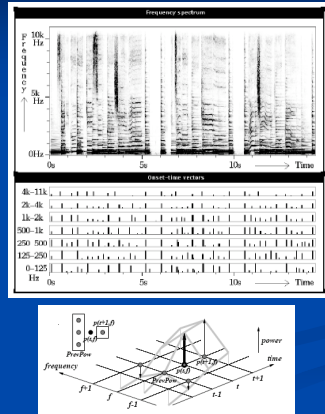
- Contain onset information over 7 frequency bins from 0 to 11KHz
- Considers the rapidity of an increase in power per bin per frame time using a "degree-of-onset function" $D(t) = \sum d(t, f)$

$$d(t, f) = \begin{cases} \max(p(t, f), p(t+1, f)) - PrevPow & (\min(p(t, f), p(t+1, f)) > PrevPow), \\ 0 & (\text{otherwise}), \end{cases} \quad (1)$$

$$PrevPow = \max(p(t-1, f), p(t-1, f \pm 1)). \quad (2)$$

Note: $p(t, f)$ is the power of the spectrum of frequency f at time t .

Onset Vectors (2)



Drums Sub-System

- Follows only a bass drum and a snare by checking for frequency signatures.
- Determines if a drum beat exists by an autocorrelation of snare-onset times in the 2nd and 4th quarter note (rule x-1)
- Uses 8 preset drum patterns or building blocks to attempt construction of the best match to the input audio data.

Drum Sub-System (2)

- Detecting a Bass Drum
 - Frequency signature is unknown. The $d(t,f)$ term from eq. 1 is used.
 - A finder peaks through the histogram and judges that a BD has sounded when an onset's peak frequency coincides with the characteristic frequency that is given by the lowest-frequency peak of the histogram.
- Detecting a Snare Drum
 - More complicated (see paper for equations)
 - Snare drum is treated as a noise distribution between 1.4 and 7.5KHz. The technique is essentially to compare the spectral power in this range to the power in the high and low ranges.
 - If power is more uniformly distributed, it is considered general noise. If instead it is focused, a finder locates an onset time similar to the standard onset detection method described on the previous slides.

Chord Change Detection

- Possibilities of chord changes in a frequency spectrum are examined without identifying musical notes or chords by name.
- Dominant frequency signatures are tracked through neighboring frames of data for significant changes.
- Frames are assumed to be at the quarter-note level, evaluated, then the frame is divided and recalculated at the eighth-note level. The scale at which chords are changing can give more clues as to the actual beat and tempo.

Complete System

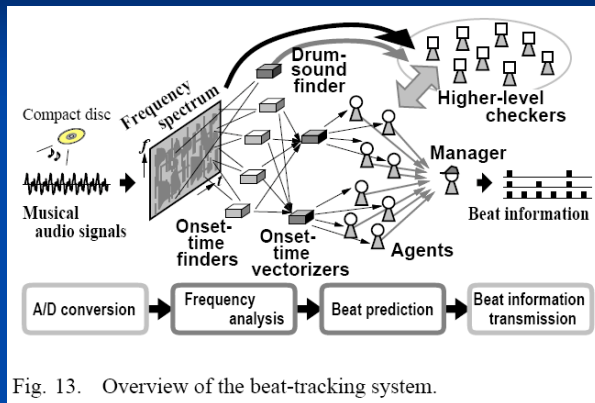


Fig. 13. Overview of the beat-tracking system.

Results

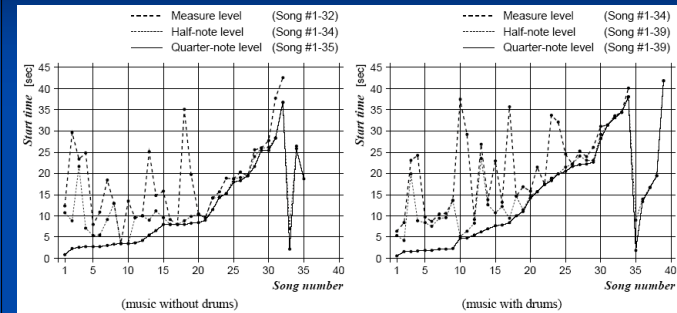
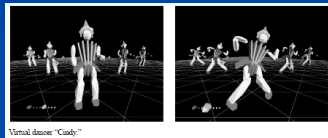


Fig. 15. Start time of tracking the correct beat structure.

Applications

- Beat-driven real-time computer graphics



- Intelligent audio effects processing
 - [sample.wav](#)
 - [StatyVerb.wav](#)
 - [IntellaVerb.wav](#)

Questions

- Grass is green because it is filled with chlorophyll that is a green pigment that reflects green light from the sun and absorbs the other colors
- Sunlight collides with air, scatters blue wavelengths
- The meaning of life is to give and receive love and work at something you feel passionate about. Or not.

Questions and Answers provided by
www.why-is-the-sky-blue.tv