Introduction to **Audio Content Analysis**

module 9.1: introduction to tempo & rhythm terminology

alexander lerch

Georgia Tech | Center for Music Technology
College of Design

introduction
overview

corresponding textbook section

sections 9.1 & 9.2

- **lecture content**
  - terminology for rhythm detection
  - perceptually motivated rhythm accuracy

- **learning objectives**
  - describe the terms onset, tempo, meter, bar, and rhythm
  - give two examples of typical onset times for musical instruments

introduction

overview

corresponding textbook section

sections 9.1 & 9.2

- **lecture content**
  - terminology for rhythm detection
  - perceptually motivated rhythm accuracy

- **learning objectives**
  - describe the terms onset, tempo, meter, bar, and rhythm
  - give two examples of typical onset times for musical instruments

temporal events
introduction

- **categorization of temporal parameters**:
  - *score* parameters:
    structure, time signature, rhythm, ...
  - *performance* parameters:
    tempo, timing, ...

- **perception of temporal parameters**:
  - audio signal/stream is segmented into distinct events $\Rightarrow$ *onsets* (segment start)
  - humans *structure and group* these events due to position, salience, ...

temporal events
introduction

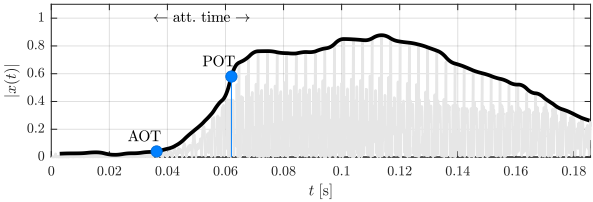
- **categorization of temporal parameters**:
  - *score* parameters:
    structure, time signature, rhythm, ...
  - *performance* parameters:
    tempo, timing, ...

- **perception of temporal parameters**:
  - audio signal/stream is segmented into distinct events $\Rightarrow$ *onsets* (segment start)
  - humans *structure and group* these events due to position, salience, ...

human perception of temporal events

introduction to onsets

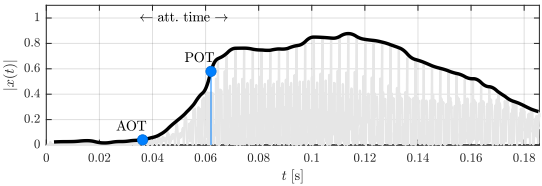
- **onset** is start of a musical event
- **properties**:
  - position
  - strength, salience
  - length?



matlab source: plotOnset.m

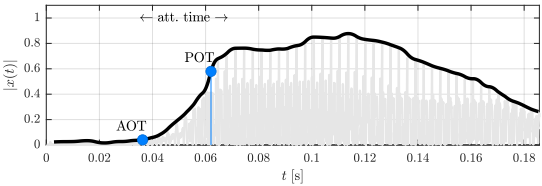
human perception of temporal events
initial transients

- percussive instruments:
  - 3-20 ms
- woodwind instruments:
  - up to 300 ms
- typical range for majority of instruments:
  - 15–50 ms



human perception of temporal events
initial transients

- percussive instruments:
  - 3-20 ms
- woodwind instruments:
  - up to 300 ms
- typical range for majority of instruments:
  - 15–50 ms



human perception of temporal events

human detection accuracy

- *detection & discrimination* of 2 subsequent onsets
  - detection $\Delta t > 2\,\mathrm{ms}$, discrimination $\Delta t > 20\,\mathrm{ms}$[1]
- *prediction* of looped monophonic instrument onsets
  - IOI 600 ms: $\sigma = 12\,\mathrm{ms}$
  - IOI < 240 ms: $\sigma = 10\,\mathrm{ms}$
- manual onset time *annotation*
  - piano: mean abs. error: 4.3 ms, max: 35 ms
  - various: mean abs. error: 10 ms, max: 30 ms
- ensemble performance
  - string & woodwind: deviations up to 30-50 ms
  - piano: $\sigma = 14 - 38\,\mathrm{ms}$

[1] I. J. Hirsh, "Auditory Perception of Temporal Order," *JASA*, vol. 31, no. 6, p. 759, 1959.

human perception of temporal events

human detection accuracy

- *detection & discrimination* of 2 subsequent onsets
  - detection $\Delta t > 2\,\mathrm{ms}$, discrimination $\Delta t > 20\,\mathrm{ms}$
- *prediction* of looped monophonic instrument onsets
  - IOI 600 ms: $\sigma = 12\,\mathrm{ms}$[1]
  - IOI < 240 ms: $\sigma = 10\,\mathrm{ms}$[2]
- manual onset time *annotation*
  - piano: mean abs. error: 4.3 ms, max: 35 ms
  - various: mean abs. error: 10 ms, max: 30 ms
- ensemble performance
  - string & woodwind: deviations up to 30-50 ms
  - piano: $\sigma = 14 - 38\,\mathrm{ms}$

[1] J. W. Gordon, "Perception of Attack Transients in Musical Tones," Dissertation, Stanford University, Center for Computer Research in Music and Acoustics (CCRMA), Stanford, 1984.

[2] A. Friberg and J. Sundberg, "Perception of just noticeable time displacement of a tone presented in a Metrical Sequence at Different Tempos," *STL-QPSR*, vol. 33, no. 4, pp. 97–108, 1992.

human perception of temporal events

human detection accuracy

- *detection & discrimination* of 2 subsequent onsets
  - detection $\Delta t > 2\,\mathrm{ms}$, discrimination $\Delta t > 20\,\mathrm{ms}$
- *prediction* of looped monophonic instrument onsets
  - IOI 600 ms: $\sigma = 12\,\mathrm{ms}$
  - IOI < 240 ms: $\sigma = 10\,\mathrm{ms}$
- manual onset time *annotation*
  - piano: mean abs. error: 4.3 ms, max: 35 ms[1]
  - various: mean abs. error: 10 ms, max: 30 ms[2]
- ensemble performance
  - string & woodwind: deviations up to 30-50 ms
  - piano: $\sigma = 14 - 38\,\mathrm{ms}$

[1] B. H. Repp, "Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's 'Träumerei'," *JASA*, vol. 92, no. 5, pp. 2546–2568, 1992.

[2] P. Leveau, L. Daudet, and G. Richard, "Methodology and Tools for the Evaluation of Automatic Onset Detection Algorithms in Music," in *ISMIR*, Barcelona, 2004.

human perception of temporal events

human detection accuracy

- *detection & discrimination* of 2 subsequent onsets
  - detection $\Delta t > 2\,\mathrm{ms}$, discrimination $\Delta t > 20\,\mathrm{ms}$
- *prediction* of looped monophonic instrument onsets
  - IOI 600 ms: $\sigma = 12\,\mathrm{ms}$
  - IOI < 240 ms: $\sigma = 10\,\mathrm{ms}$
- manual onset time *annotation*
  - piano: mean abs. error: 4.3 ms, max: 35 ms
  - various: mean abs. error: 10 ms, max: 30 ms
- ensemble performance
  - string & woodwind: deviations up to 30-50 ms[1]
  - piano: $\sigma = 14 - 38\,\mathrm{ms}$[2]

[1] R. A. Rasch, "Synchronization in Performed Ensemble Music," *Acustica*, vol. 43, pp. 121–131, 1979.

[2] L. H. Shaffer, "Timing in Solo and Duet Piano Performances," *Quarterly Journal of Experimental Psychology*, vol. 36A, pp. 577–595, 1984.

human perception of temporal events

offsets

what about offsets/end of notes

human perception of temporal events

offsets

what about offsets/end of notes

- **perceptually not as important** as an onset
  - offset are often ignored in rhythm perception
- **systematic difficulties**: when does a note end?
  - performer stops excitation
  - instrument stops oscillation
  - listener cannot recognize it anymore
- **practical difficulties**: hard to detect
  - low volume
  - reverberation
  - masking

human perception of temporal events

tempo, meter & rhythm

- **tempo**: perceived equal duration pulses at a "natural" rate — tactus
  - constant tempo

$$\mathfrak{T} = \frac{\mathcal{B} \cdot 60\,\mathrm{s}}{\Delta t_\mathrm{s}} \ [\mathrm{BPM}]$$

  - dynamic tempo

$$\mathfrak{T}_\mathrm{local}(j) = \frac{60\,\mathrm{s}}{t_\mathrm{b}(j+1) - t_\mathrm{b}(j)} \ [\mathrm{BPM}]$$

  - perceived overall tempo?
    - average, main, mode, ...
- **meter**
  - group of strong and weak musical elements/beats
  - typically 3 to 7 beats (app. 5 s)
- **rhythm**
  - group length 1–8 beats
  - defined by accents and time intervals

human perception of temporal events

tempo, meter & rhythm

- **tempo**: perceived equal duration pulses at a "natural" rate — tactus
  - constant tempo

$$\mathfrak{T} = \frac{\mathcal{B} \cdot 60\,\mathrm{s}}{\Delta t_\mathrm{s}} \ [\mathrm{BPM}]$$

  - dynamic tempo

$$\mathfrak{T}_\mathrm{local}(j) = \frac{60\,\mathrm{s}}{t_\mathrm{b}(j+1) - t_\mathrm{b}(j)} \ [\mathrm{BPM}]$$

  - perceived overall tempo?
    - average, main, mode, ...
- **meter**
  - group of strong and weak musical elements/beats
  - typically 3 to 7 beats (app. 5 s)
- **rhythm**
  - group length 1–8 beats
  - defined by accents and time intervals

human perception of temporal events

tempo, meter & rhythm

- **tempo**: perceived equal duration pulses at a "natural" rate — tactus
  - constant tempo

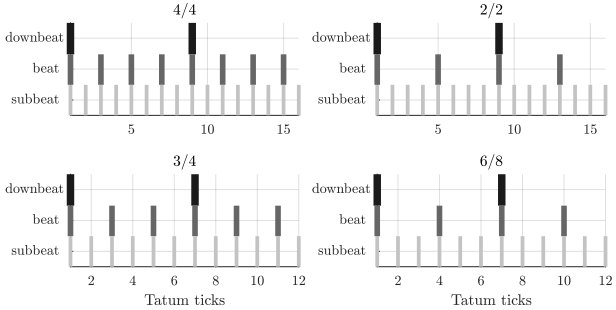
$$\mathfrak{T} = \frac{\mathcal{B} \cdot 60\,\mathrm{s}}{\Delta t_\mathrm{s}} \ \ [\mathrm{BPM}]$$

  - dynamic tempo

$$\mathfrak{T}_\mathrm{local}(j) = \frac{60\,\mathrm{s}}{t_\mathrm{b}(j+1) - t_\mathrm{b}(j)} \ \ [\mathrm{BPM}]$$

  - perceived overall tempo?
    - average, main, mode, ...
- **meter**
  - group of strong and weak musical elements/beats
  - typically 3 to 7 beats (app. 5 s)
- **rhythm**
  - group length 1–8 beats
  - defined by accents and time intervals

temporal events
hierarchical structure



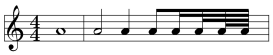
matlab source: plotBeatHierarchy.m

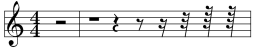
musical notation of temporal events

tempo, time signature, bar & note value

- **tempo**
  - *Largo*, *Adagio*, *Andante*, *Moderato*, *Allegro*, *Presto*
  - *ritardando*, *accelerando*, . . .
  - modern scores: sometimes overall tempo in BPM
- **bar**
  - score equivalent of perceptual meter
  - begin of bar is marked by a vertical line
- **time signature**
  - conveys length of bar
- **note value**

$\frac{4}{4}$ $\frac{3}{4}$ $\frac{2}{4}$ $\frac{2}{2}$





summary
lecture content

- **perceptual terms**
  - onset, tempo, meter, rhythm
- **musical terms**
  - tempo, bar, time signature, note value, rhythm
- **accuracy range of interest**
  - 2–300 ms