Introduction to Audio Content Analysis

Module 9.3: Onset Detection

alexander lerch

Georgia Tech Center for Music Technology
College of Design

corresponding textbook section

Section 9.3

- **lecture content**
  - detection of the start of musical events
  - fundamental methods for generating a novelty function
  - fundamental methods for peak picking
- **learning objectives**
  - describe the term onset
  - implement an automatic onset detection system

corresponding textbook section

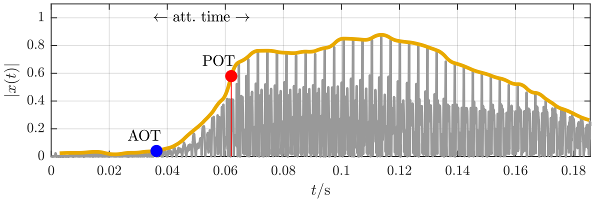
Section 9.3

- **lecture content**
  - detection of the start of musical events
  - fundamental methods for generating a novelty function
  - fundamental methods for peak picking
- **learning objectives**
  - describe the term onset
  - implement an automatic onset detection system

onset detection
problem statement

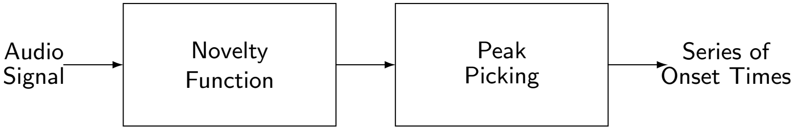
- **onset**: begin of musical event
- **goal**: detect the point in time of an onset
- **challenges**:
  - which time stamp of the initial attack time actually marks the onset time?
  - *polyphonic* audio signals:
    - unknown number of voices and events
    - multiple onsets occur at “the same” time
    - onset might be obfuscated by other musical content

- **note onset time**:
  - time the instrument is *triggered*
- **acoustic onset time**:
  - time of first *measurable* instrument output
- **perceptual onset time**:
  - time the event is *perceived* by listener

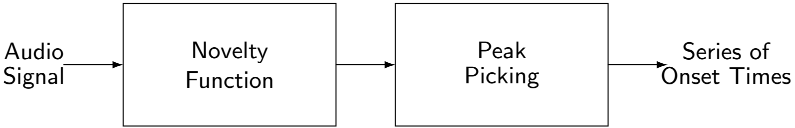


matlab source: plotOnset.m

onset detection
overview



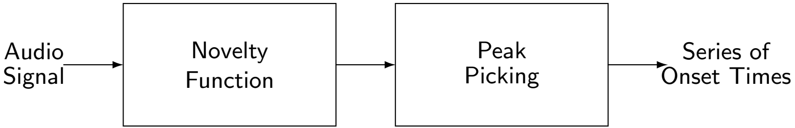
1. **novelty function**
   - measure of probability for new events/signal change over time
2. **peak picking**
   - identify the most likely locations for onsets



1. **novelty function**
    - measure of probability for new events/signal change over time
2. **peak picking**
    - identify the most likely locations for onsets

onset detection

overview



1. **novelty function**
   - measure of probability for new events/signal change over time
2. **peak picking**
   - identify the most likely locations for onsets

- alternative **terms** for *novelty function*
  - detection function
  - difference function

- **processing steps**
  1. extract features
  2. compute derivative
  3. smooth result
  4. apply Half-Wave-Rectification (HWR)

- alternative **terms** for *novelty function*
  - detection function
  - difference function

- **processing steps**
  1. extract features
  2. compute derivative
  3. smooth result
  4. apply Half-Wave-Rectification (HWR)

onset detection
novelty function

- alternative **terms** for *novelty function*
  - detection function
  - difference function

- **processing steps**
  1. extract features
  2. compute derivative
  3. smooth result
  4. apply Half-Wave-Rectification (HWR)

onset detection

novelty function

- alternative **terms** for *novelty function*
  - detection function
  - difference function

- **processing steps**
  1. extract features
  2. compute derivative
  3. smooth result
  4. apply Half-Wave-Rectification (HWR)

onset detection
novelty function

- alternative **terms** for *novelty function*
  - detection function
  - difference function

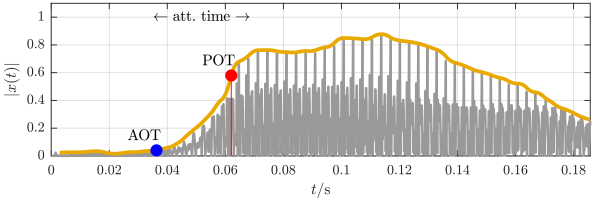
- **processing steps**
  1. extract features
  2. compute derivative
  3. smooth result
  4. apply Half-Wave-Rectification (HWR)

onset detection

novelty function examples 1/3

1 time domain example

- *feature*: time domain envelope
- *derivative*: slope of envelope
- *HWR*: only interested in onsets, not offsets



2 pitch domain:

- *feature*: pitch contour
- *derivative*: changes in pitch

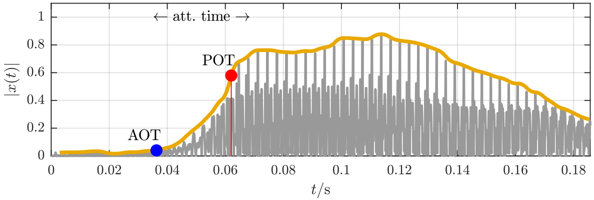
matlab source: plotOnset.m

onset detection

novelty function examples 1/3

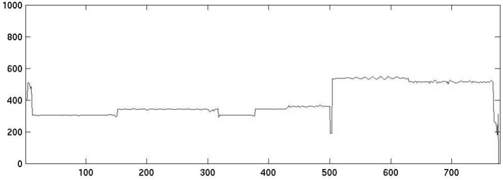
1 **time domain** example

- *feature*: time domain envelope
- *derivative*: slope of envelope
- *HWR*: only interested in onsets, not offsets



2 **pitch domain**:

- *feature*: pitch contour
- *derivative*: changes in pitch



matlab source: plotOnset.m

[1] N. Collins, "Using a pitch detector for onset detection," in *ISMIR*, 2005, pp. 100–106.

onset detection
novelty function examples 2/3

3. **STFT-based**: compute block difference
   - *flux*
     - $d_{\mathrm{hai}}(n) = \sum\limits_{k=0}^{\mathcal{K}/2-1} \log_2\left(\frac{|X(k,n)|}{|X(k,n-1)|}\right)$
     - $d_{\mathrm{lar}}(n) = \sum\limits_{k=k(f_{\mathrm{min}})}^{k(f_{\mathrm{max}})} \sqrt{|X(k,n)|} - \sqrt{|X(k,n-1)|}$
   - *cosine distance*
   - *complex*

onset detection
novelty function examples 2/3

3 **STFT-based**: compute block difference

- *flux*
  - $$d_{\mathrm{hai}}(n) = \sum_{k=0}^{\mathcal{K}/2-1} \log_2\left(\frac{|X(k,n)|}{|X(k,n-1)|}\right)$$
  - $$d_{\mathrm{lar}}(n) = \sum_{k=k(f_{\mathrm{max}})}^{k(f_{\mathrm{min}})} \sqrt{|X(k,n)|} - \sqrt{|X(k,n-1)|}$$
- *cosine distance*
  - $$d_{\mathrm{foo}}(n) = 1 - \frac{\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n)|\cdot|X(k,n-1)|}{\sqrt{\left(\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n)|^2\right)\cdot\left(\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n-1)|^2\right)}}$$
- *complex*

onset detection

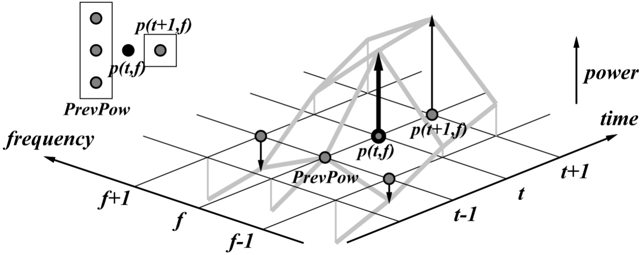
novelty function examples 2/3

3 STFT-based: compute block difference

- *flux*
  - $$d_{\mathrm{hai}}(n) = \sum_{k=0}^{\mathcal{K}/2-1} \log_2\left(\frac{|X(k,n)|}{|X(k,n-1)|}\right)$$
  - $$d_{\mathrm{lar}}(n) = \sum_{k=k(f_{\mathrm{min}})}^{k(f_{\mathrm{max}})} \sqrt{|X(k,n)|} - \sqrt{|X(k,n-1)|}$$
- *cosine distance*
  - $$d_{\mathrm{foo}}(n) = 1 - \frac{\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n)|\cdot|X(k,n-1)|}{\sqrt{\left(\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n)|^2\right)\cdot\left(\sum\limits_{k=0}^{\mathcal{K}/2-1} |X(k,n-1)|^2\right)}}$$
- *complex*
  - $$d_{\mathrm{dux}}(n) = \sum_{k=0}^{\mathcal{K}/2-1} |X(k,n) - X(k,n-1)|$$

3 **STFT-based** cont'd

- *Goto*-distance[2]
  - higher power than closest preceding and following bins

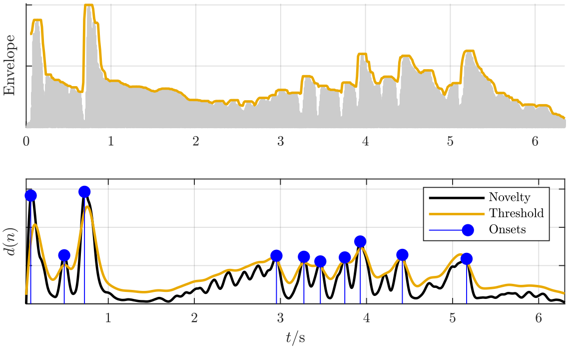


[2] M. Goto and Y. Muraoka, "Music Understanding At The Beat Level – Real-time Beat Tracking For Audio Signals," in *Proceedings of the Workshop on Computational Auditory Scene Analysis (IJCAI)*, Aug. 1995.

onset detection

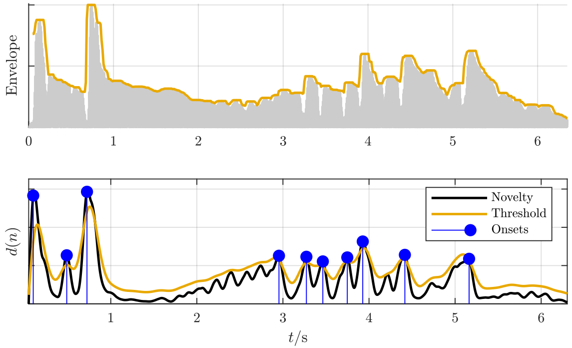
peak picking: introduction

- detect onsets in the smoothed novelty function
- typical **criteria**
  - local maximum & salient peak
  - higher than minimum likelihood
  - not too close to maxima with higher likelihood
  - other options: high attack slope, distance to prev. min, ...



matlab source: plotNovelty.m

- detect onsets in the smoothed novelty function
- typical **criteria**
  - local maximum & salient peak
  - higher than minimum likelihood
  - not too close to maxima with higher likelihood
  - other options: high attack slope, distance to prev. min, ...



matlab source: plotNovelty.m

onset detection
peak picking: thresholding

Georgia Tech Center for Music Technology College of Design

- options for thresholding
  - **fixed** threshold

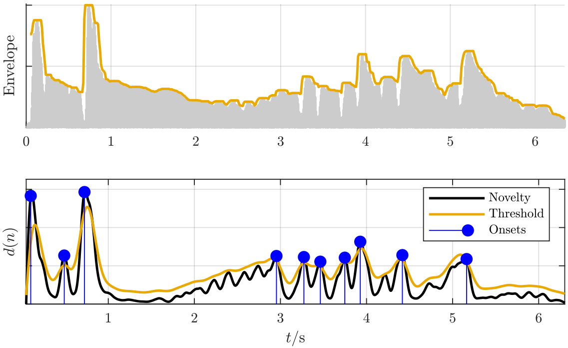
$$G_{d,\mathrm{c}} = \lambda_1$$

  - **smoothed** threshold

$$G_{d,\mathrm{ma}} = \lambda_2 + \sum_{j=0}^{\mathcal{O}-1} b(j) \cdot d(i-j)$$

  - **median** threshold

$$G_{d,\mathrm{me}} = \lambda_2 + \hat{Q}_d(0.5)$$



matlab source: plotNovelty.m

- **goal**
  - compare a series of ground truth onset time stamps with a series of predicted time stamps
- **ground truth annotation** problems
  - deviations between annotators
  - how to annotate quasi-synchronous onsets
- **metrics**
  - measure TP with tolerance range $\Rightarrow$ TP, FN, FP (TN only implicitly)
  - Precision, Recall, F-Measure
  - other metrics
    - mean (absolute) deviation
    - standard deviation
    - max deviation

- **goal**
  - compare a series of ground truth onset time stamps with a series of predicted time stamps
- **ground truth annotation** problems
  - deviations between annotators
  - how to annotate quasi-synchronous onsets
- **metrics**
  - measure TP with tolerance range $\Rightarrow$ TP, FN, FP (TN only implicitly)
  - Precision, Recall, F-Measure
  - other metrics
    - mean (absolute) deviation
    - standard deviation
    - max deviation

Georgia Tech Center for Music Technology
College of Design

- **goal**
  - compare a series of ground truth onset time stamps with a series of predicted time stamps
- **ground truth annotation** problems
  - deviations between annotators
  - how to annotate quasi-synchronous onsets
- **metrics**
  - measure TP with tolerance range $\Rightarrow$ TP, FN, FP (TN only implicitly)
  - Precision, Recall, F-Measure
  - other metrics
    - mean (absolute) deviation
    - standard deviation
    - max deviation

summary
lecture content

- **novelty function**
  - measure of unexpectedness - likelihood of an event
    - often a measure similar to flux
- **peak picking**
  - detecting peaks (onsets) in the novelty function
  - usually done by smoothing and adaptive thresholding