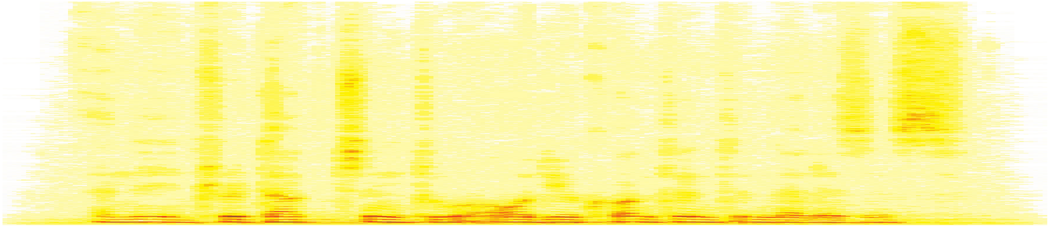
Introduction to Audio Content Analysis

Module 6.1: Onset Detection

alexander lerch



Georgia Tech Center for Music Technology
College of Design

introduction

overview

Georgia Tech Center for Music Technology College of Design

corresponding textbook section

Chapter 5 — Temporal Analysis: pp. 135–139

- **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

introduction
overview

corresponding textbook section

Chapter 5 — Temporal Analysis: pp. 135–139

- **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

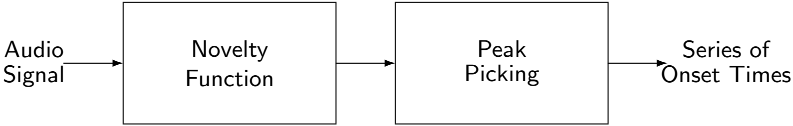
onset detection

problem statement

- **onset**: begin of musical event
- **polyphonic** audio signals:
  - unknown number of voices and events
  - multiple onsets occur at “the same” time
  - onset might be obfuscated by other musical content

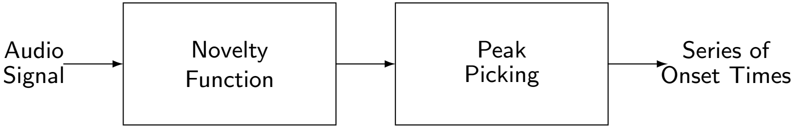
onset detection

overview



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

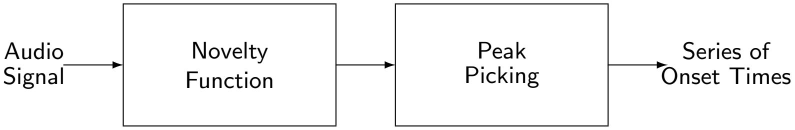
onset detection
overview



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

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

onset detection

novelty function

- **terms**
  - 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

- **terms**
  - 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

- **terms**
  - 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

- **terms**
  - 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

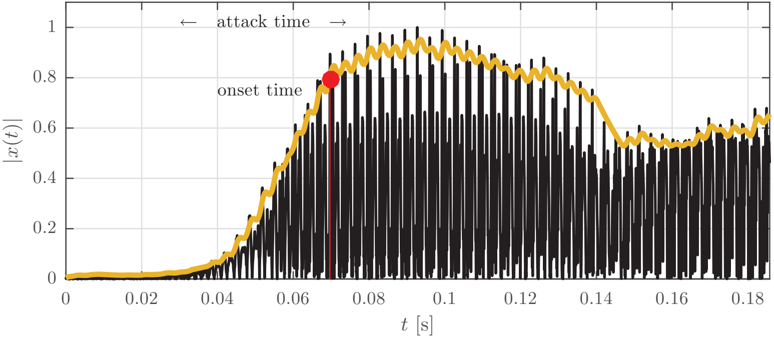
Georgia Tech Center for Music Technology College of Design

- **terms**
  - 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**
   - extract time domain envelope
   - calculate slope



2. **pitch-based**: evaluate pitch changes[1]

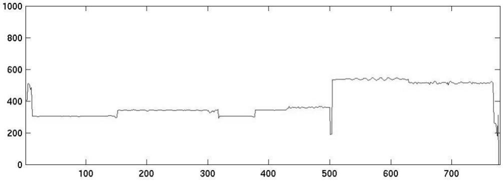
matlab source: matlab/displayOnset.m

onset detection

novelty function examples 1/3

1. **time domain**
   - extract time domain envelope
   - calculate slope

2. **pitch-based**: evaluate pitch changes[1]



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

onset detection

novelty function examples 2/3

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

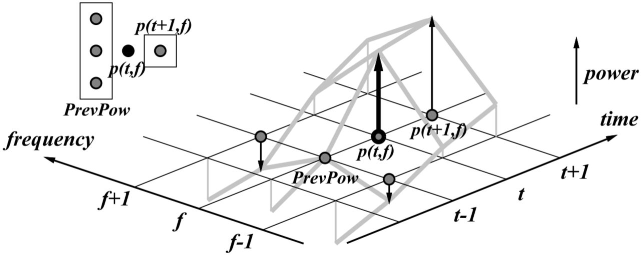
$$d_{\mathrm{dux}}(n) = \sum_{k=0}^{\mathcal{K}/2-1} |X(k,n) - X(k,n-1)|$$

onset detection

novelty function examples 3/3

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
novelty function: variants

- **number of frequency bands**
  - varies: 1, 3, 6, 21, 960, FFT length, ...
  - larger number of bands not necessarily better
    $\rightarrow$ adjust number of bands adaptively?
- **combination of bands**
  - (weight and) add novelty functions per band
  - onset detection per band and combine results

onset detection

novelty function: variants

- **number of frequency bands**
  - varies: 1, 3, 6, 21, 960, FFT length, ...
  - larger number of bands not necessarily better
    $\rightarrow$ adjust number of bands adaptively?
- **combination of bands**
  - (weight and) add novelty functions per band
  - onset detection per band and combine results

onset detection
novelty function: variants

- **number of frequency bands**
  - varies: 1, 3, 6, 21, 960, FFT length, ...
  - larger number of bands not necessarily better
    $\rightarrow$ adjust number of bands adaptively?

- **combination of bands**
  - (weight and) add novelty functions per band
  - onset detection per band and combine results

onset detection

novelty function: variants

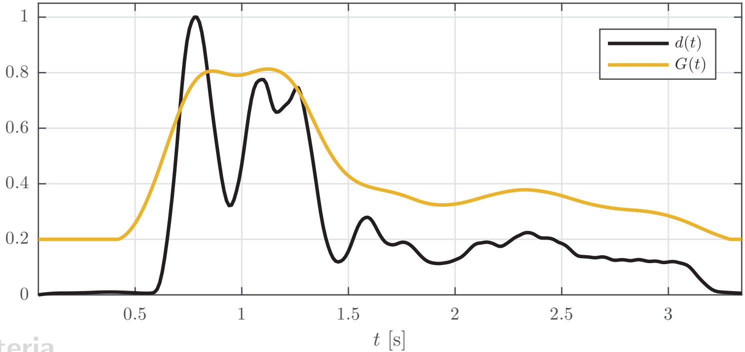
- **number of frequency bands**
  - varies: 1, 3, 6, 21, 960, FFT length, ...
  - larger number of bands not necessarily better
    $\rightarrow$ adjust number of bands adaptively?

- **combination of bands**
  - (weight and) add novelty functions per band
  - onset detection per band and combine results

onset detection

peak picking: introduction

- detect onsets in the smoothed novelty function



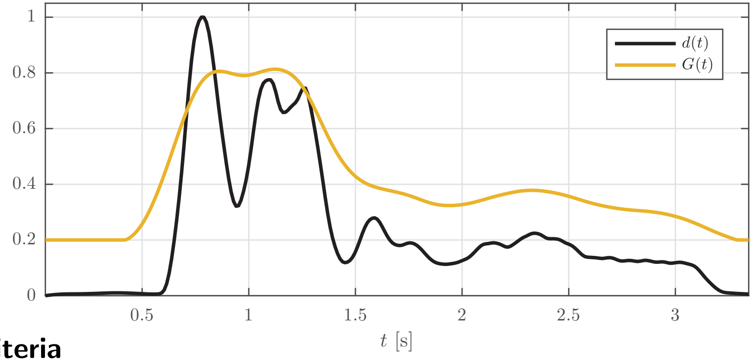
matlab source: matlab/displayNoveltyFunction.m

- 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, ...

onset detection

peak picking: introduction

- detect onsets in the smoothed novelty function



matlab source: matlab/displayNoveltyFunction.m

- 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, ...

onset detection

peak picking: thresholding

- options for thresholding
  - **fixed** threshold

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

  - **smoothed** threshold

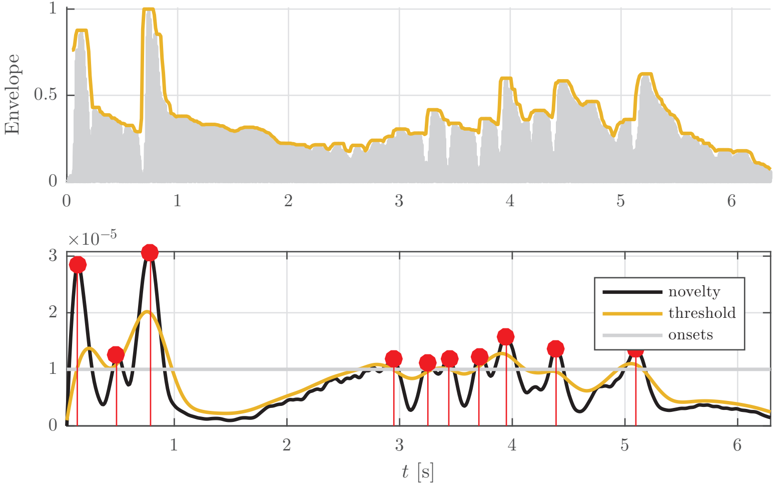
$$G_{d,\mathrm{ma}} = \lambda_2 + \sum_{j=0}^{\mathcal{O}-1} b(j) \cdot d(i-j)$$

  - **median** threshold

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

onset detection

peak picking: thresholding



matlab source: matlab/displayPeakPicking.m

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