

# MUS4831 - Lydanalyse

Feature extraction & music information retrieval

3 oktober 2013

## Feature Extraction: Ausgangspunkt

På norsk: *Egenskapsuttrekking*.

*Feature* kalles også *descriptor* (Se Peeters et al. 2011).

- Så langt i kurset har vi hovedsaklig sett på hvordan en lyd kan representeres i både tidsdomenet og frekvensdomenet
- Selv om vi kan få mye informasjon ut av et spektrogram kan det være vanskelig å gjøre en presis sammenlikning av to spektrogrammer. Det samme gjelder for sammenlikning av to bølgeformer.
- Vi kan trekke ut *features* som beskriver isolerte aspekter ved en lyd.

# Feature Extraction: Noen eksempler

- Brightness
- Loudness
- Attack
- Decay
- Noisiness
- Spectral spread
- Spectral centroid
- Temporal centroid
- Zero-crossing rate

# Feature Extraction: Noen eksempler

- Brightness
- Loudness
- Attack
- Decay
- Noisiness
- Spectral spread
- Spectral centroid
- Temporal centroid
- Zero-crossing rate

Global descriptors

# Feature Extraction: Noen eksempler

- Brightness
- Loudness
- Attack
- Decay
- Noisiness
- Spectral spread
- Spectral centroid
- Temporal centroid
- Zero-crossing rate

Time-varying descriptors

# Et eksempel på feature: Root mean square (RMS)

**Root mean square (RMS):** Energien i en waveform.

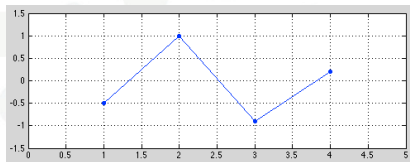
Selve navnet “Root mean square” beskriver utregningen:

- 1 hvert sample opphøyd i andre
- 2 gjennomsnittet
- 3 kvadratroten av det hele

$$\sqrt{\frac{s_1^2 + s_2^2 + s_3^2 + \cdots + s_n^2}{n}}$$

# RMS eksempel

$$\sqrt{\frac{s_1^2 + s_2^2 + s_3^2 + \dots + s_n^2}{n}}$$



Eksempel: et signal med 4 samples: -0.5 , 1 , -0.9 , 0.2

Vi setter det inn i formelen:

$$\sqrt{\frac{(-0.5)^2 + 1^2 + (-0.9)^2 + 0.2^2}{4}} = \sqrt{\frac{0.25 + 1 + 0.81 + 0.04}{4}}$$

Regner vi ut telleren får vi

$$\sqrt{\frac{2.1}{4}} = \sqrt{0.525} = 0.7246$$

# RMS av pianolyd

Vanligvis bruker vi ferdige script til å regne ut features, siden manuell utregning tar lang tid (pianolyden har 44100 samples). I Matlab kan man bruke funksjonen “rms”.

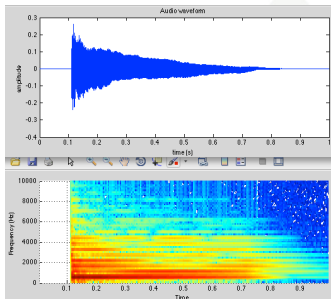
les inn lyden:

```
piano = wavread('piano.wav');
```

Beregn root mean square:

```
rms(piano)
```

Svaret du får er: 0.0375



# RMS av pianolyd

Vanligvis bruker vi ferdige script til å regne ut features, siden manuell utregning tar lang tid (pianolyden har 44100 samples). I Matlab kan man bruke funksjonen “rms”.

les inn lyden:

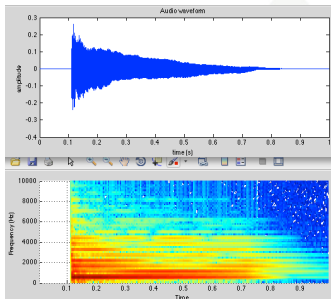
```
piano = wavread('piano.wav');
```

Beregn root mean square:

```
rms(piano)
```

Svaret du får er: 0.0375

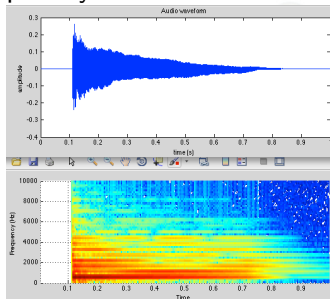
Er dette en global eller tidsvarierende deskriptor?



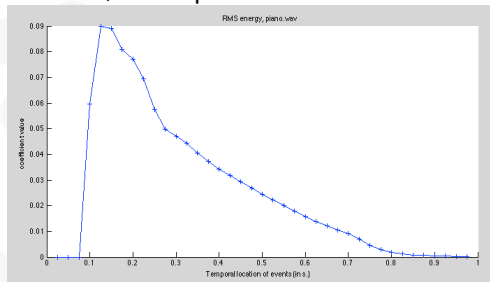
# Vinduer og RMS

Hvis vi bruker vinduer (slik vi gjør for spektrogrammer) kan vi få RMS som en tidsvarierende deskriptor:

pianolyd



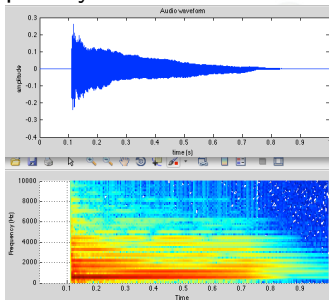
Root mean square, med en vindusstørrelse på 0.025 sekunder



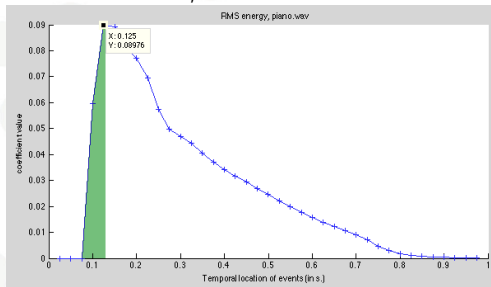
# Vinduer og RMS

Med tidsvarierende RMS kan vi regne ut andre features, som for eksempel attack-tid.

pianolyd

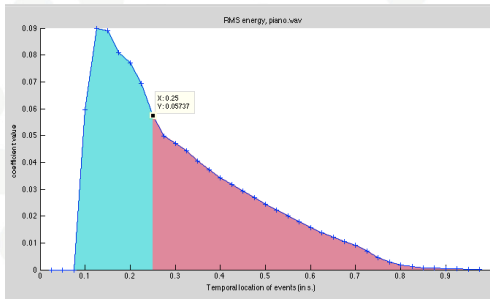
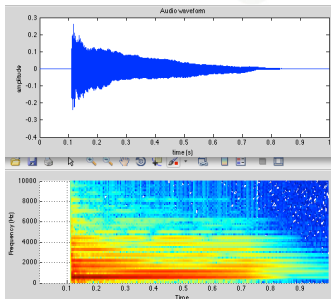


Attacktiden her er fra 0.075 sekunder til 0.125 sekunder, dvs 50 ms.



# Temporal centroide

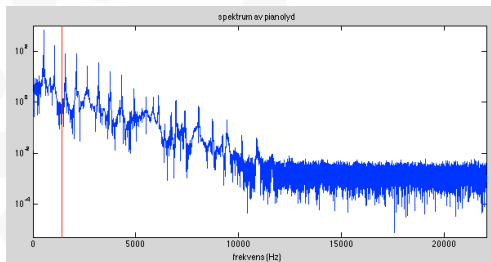
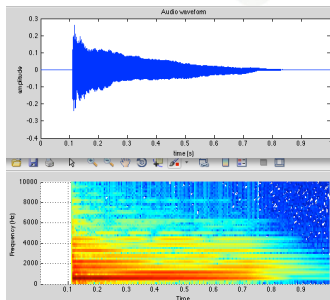
Temporal centroid: Midtpunktet (i tid) for energien i lydfilen



- Temporal centroid for denne lyden: 0.25 sekunder
- Hvis Temporal Centroide er under 0.5 betyr det at det meste av lydens energi befinner seg i første halvdel av lydfilen

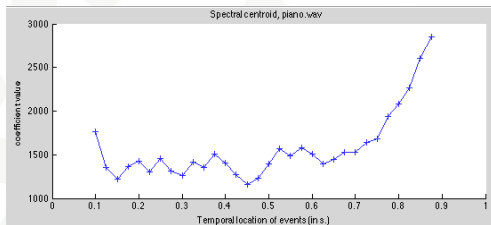
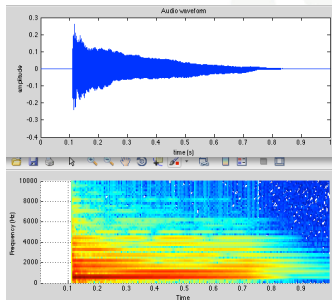
# Spektral centroide

Spectral centroid: Midtpunktet i lydfilens spektrum.



- Spektral centroide for denne lyden: 1391 Hz
- Vi bruker som regel vinduer når vi regner ut spektral centroide, slik at vi kan se hvordan den endrer seg over tid.

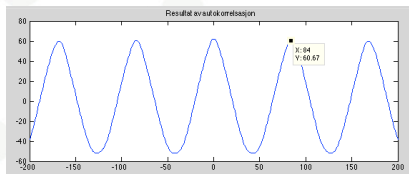
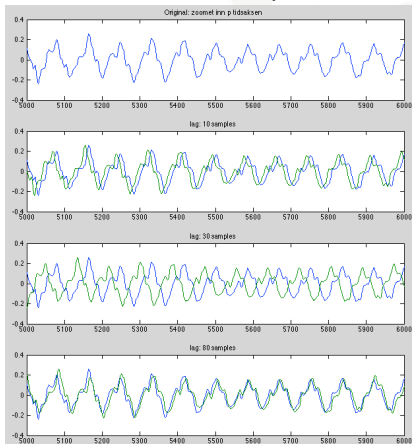
# Spektral centroide



- Med en vindusstørrelse på 25 millisekunder ser vi at spektral centroide utvikler seg over tid.

# Autocorrelation

Ved å sammenlikne et signal med en tidsforskjøvet versjon av seg selv kan vi finne ut periodisiteter i signalet.



Her har vi en høy korrelasjonskoeffisient ved 84 samples.

Siden vi har en samplingsrate på 44100 finner vi frekvensen ved  $44100/84 = 523 \text{ Hz}$ .

# Toolboxer og kilder for beregning av features

## Litteratur:

- Collinsboka
- Peeters (2004): "A large set of audio features for sound description in the CUIDADO project" Tech. report. IRCAM. (tilgjengelig på Fronter)
- Peeters et al. (2011): "The Timbre Toolbox: Extracting audio descriptors from musical signals" i *The Journal of the Acoustical Society of America*, 130(5), s. 2902–2916 (tilgjengelig på Fronter)

## Toolbokser og annen programvare:

- MATLAB: MIR toolbox, fra universitetet i Jyväskylä.  
<https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- MATLAB: Timbre toolbox, fra Peeters et al.  
<http://www.cirmmt.mcgill.ca/l/research-tools/timbretoolbox/timbretoolbox.zip>
- Praat: <http://www.fon.hum.uva.nl/praat/>
- Max / PD (for å trekke ut lydfeatures i sanntid)

# Toolboxer og kilder for beregning av features

## Litteratur:

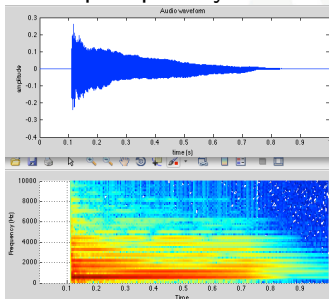
- Collinsboka
- Peeters (2004): "A large set of audio features for sound description in the CUIDADO project" Tech. report. IRCAM. (tilgjengelig på Fronter)
- Peeters et al. (2011): "The Timbre Toolbox: Extracting audio descriptors from musical signals" i *The Journal of the Acoustical Society of America*, 130(5), s. 2902–2916 (tilgjengelig på Fronter)

## Toolbokser og annen programvare:

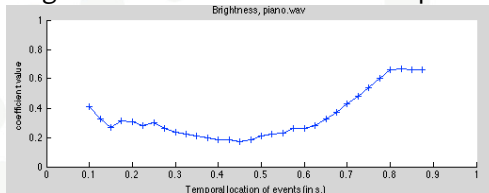
- **MATLAB: MIR toolbox, fra universitetet i Jyväskylä.**  
<https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- **MATLAB: Timbre toolbox, fra Peeters et al.**  
<http://www.cirmmt.mcgill.ca/l/research-tools/timbretoolbox/timbretoolbox.zip>
- Praat: <http://www.fon.hum.uva.nl/praat/>
- Max / PD (for å trekke ut lydfeatures i sanntid)

# Statistiske beregninger på tidsvariable deskriptorer

## Eksempel: pianolyd

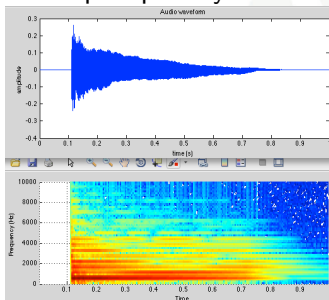


## *Brightness* som tidsvariabel deskriptor:

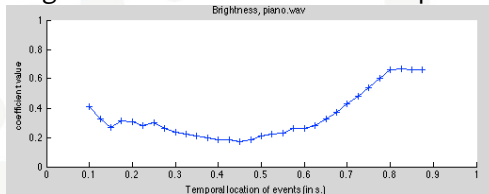


# Statistiske beregninger på tidsvariable deskriptorer

## Eksempel: pianolyd



## Brightness som tidsvariabel deskriptor:

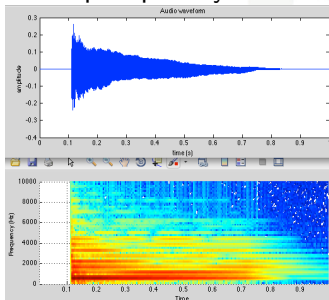


## Brightness som global deskriptor:

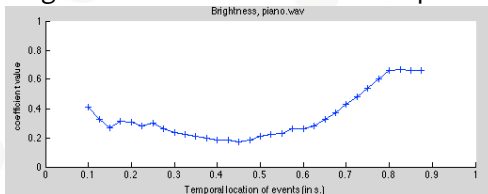
- Gjennomsnitt av brightness = 0.34
- Maksimumverdi av brightness = 0.66
- Minimumverdi av brightness = 0.17
- Varians av brightness = 0.026

# Statistiske beregninger på tidsvariable deskriptorer

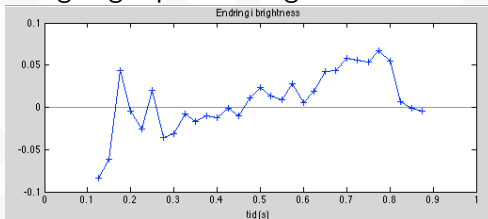
## Eksempel: pianolyd



## Brightness som tidsvariabel deskriptor:



Vi kan også se på hvordan *brightness* endres over tid, og gjøre samme beregninger på dette signalet:



# Music Information Retrieval (MIR)

- Forskningsfelt som forsker på å trekke ut egenskaper og klassifisere musikk.
- Bruker maskinlæringsteknikker og mønstergjenkjenningsmetoder for eksempel for å:
  - Kjenne igjen en spesifikk sjanger
  - Finne lydfiler som ligner på hverandre (“query by example”)
  - Finne lydfiler basert på nynning (“query by humming”)
  - Finne lydfiler basert på bevegelse (“query by motion”)

- Toolbox for matlab som tar for seg mye innenfor MIR
- Bruker en egen type datavariabel
- Ulike nivåer når man jobber med MIR: Fra enkeltlyder til motiver, fraser og fullstendige verk
- Noen mirfunksjoner er nyttige på alle disse nivåene, mens andre er mer spesialiserte for ett nivå (f.eks. sjangerklassifisering eller estimering av tonehøyde).
- MIR Toolbox manualen er ganske oversiktlig og det er lurt å ta en titt der for å få oversikt over funksjonene.
- Sjekk også ut hjelpefilene til hver funksjon (F1-tast)

# MIR Toolbox - noen funksjoner til detaljnivå lydanalyse

Les inn en fil i variabelen `minlyd`:  
`minlyd = miraudio('minlydfil.wav')`

Spill av filen:  
`mirplay(minlyd)`

Gi meg "alle" features:  
`mirfeatures(minlyd)`

Beregn RMS:  
`mirrms(minlyd)`

Tidsvarierende RMS:  
`mirrms(minlyd, 'frame')`

Beregn envelope i variabelen `Xenv`  
`Xenv = mirenvelope(minlyd)`

Spektrum  
`mirspectrum(minlyd)`

Spektrum i desibel  
`mirspectrum(minlyd, 'dB')`

Spektrogram  
`mirspectrum(minlyd, 'frame')`

Spektrogram i *Bark*  
`mirspectrum(minlyd, 'Bark', 'frame')`

Demonstrasjon av flere ulike funksjoner:  
`mirdemo`

Konverter `mirdata` til vanlig matlab variabel:  
`data = mirgetdata(Xrms);`

- Bruk et semikolon på slutten av kommandoen for å ikke vise figur.

# MIR Toolbox - noen funksjoner til detaljnivå lydanalyse

|   |  |
|---|--|
| Autokorrelasjon                         | <code>mirautocor(minlyd)</code>  |
| Autokorrelasjon<br>min 1ms, max 5ms     | <code>mirautocor(minlyd,'Min',0.001,'Max',0.005)</code>                        |
| Samme som<br>ovenfor, over tid          | <code>mirautocor(minlyd,'Min',0.001,'Max',...<br/>0.005,'frame')</code>        |
| Samme, plottet på<br>en frekvensakse    | <code>mirautocor(minlyd,'Min',0.001,'Max',...<br/>0.005,'frame','freq')</code> |
| Pitch                                   | <code>mirpitch(minlyd,'frame')</code>  |
| Flux: Endring i et<br>spektrum over tid | <code>s = mirspectrum(minlyd,'frame');</code><br><code>mirflux(s)</code>       |
| cepstrum<br>spektrum av spektrum        | <code>mircepstrum(minlyd)</code>   |

- `mirpitch`-funksjonen er basert på `mirautocor`, men `mirautocor` kan også brukes til andre ting som for eksempel analyse av tempo/rytme