

# Grundlagen des Information Retrieval: Music Information Retrieval

<http://www.ifs.tuwien.ac.at/mir>

**Andreas Rauber**

**Department of Softwaretechnology and  
Interactive Systems**

**Vienna University of Technology**

<http://www.ifs.tuwien.ac.at/~andi>



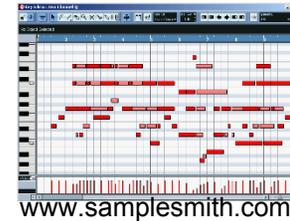
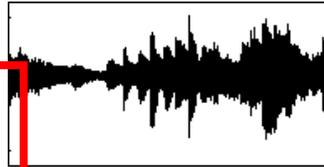
## Chorus

- Lead-in
- **Chorus**
- **Verse 1: What is Music-IR?**
- Verse 2: Audio Features
- Verse 3: Retrieval and other applications
- Fade-out

## What is „Music“?

- Music, of course!

- Audio: wav, au, mp3, ...
- Symbolic: MIDI, mod, ...
- Scores: Scan, MusicXML



www.westminster.gov.uk

- Text

- Song lyrics
- Artis Biographies
- Websites:  
Fanpages, tags,  
Album Reviews,  
Genre descriptions

- Video/Images

- Album covers
- Music videos

- Other data

- Playlists
- Market basket
- Network of  
linked sites,  
events

## Music - Sound

- Sound as acoustic wave
- Characterized by the properties of waves (frequency/wavelength, amplitude)
- Frequency: pitch
  - Humans can hear approx. 20Hz-20kHz
  - speech: 200Hz-8kHz
- Amplitude: Loudness
  - measured as pressure in micropascal  $\mu Pa$
  - hearing threshold: approx.  $20\mu Pa$
  - logarithmic decibel scale  $L_p = 10 \log_{10} \left( \frac{p^2}{p_0^2} \right) = 20 \log_{10} \left( \frac{p}{p_0} \right)$  dB

## Music - Sound

- Different file formats for storing sound:
  - lossless formats
    - WAV (may hold compressed audio, but usually lossless PCM)
    - FLAC, Shorten, Monkey's Audio, ATRAC Advanced Lossless, Apple Lossless, WMA Lossless, TTA
  - lossy formats
    - MP3
    - ATRAC
    - AAC
    - Ogg Vorbis
    - WMA
    - ...

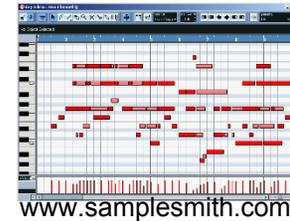
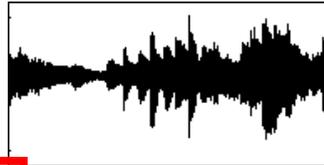
## Features for Sound / PCM signal

- Many different types
  - Timbre
  - Rhythm
  - Pitch, melody, chords, ...
  - Loudness, roughness, ...
- Time domain vs. Frequency domain
- Capturing different aspects of sound
- (New approaches via feature learning / deep learning)
- To be discussed in detail a bit later...

# Music IR – Music?

## What is „Music“?

- Music, of course!
  - Audio: wav, au, mp3, ...
  - Symbolic: MIDI, mod, ...
  - Scores: Scan, MusicXML

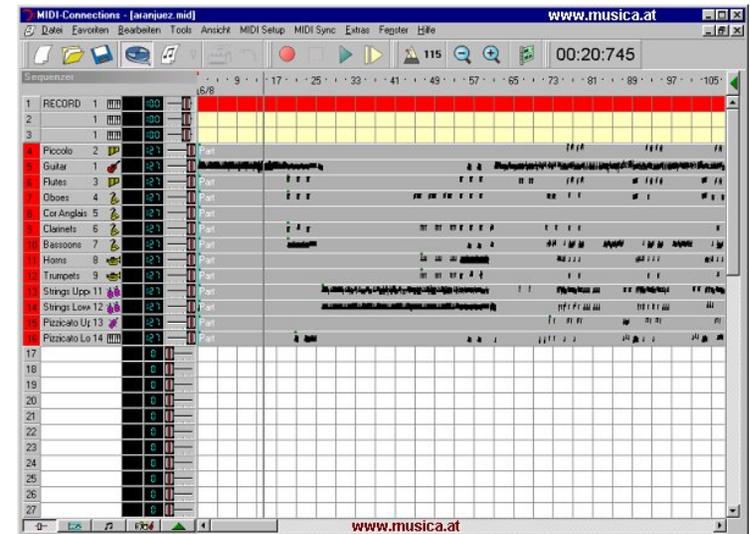


www.westminster.gov.uk

- |   |  |   |
|---|--|---|
| <ul style="list-style-type: none"> <li>■ Text           <ul style="list-style-type: none"> <li>– Song lyrics</li> <li>– Artis Biographies</li> <li>– Websites:<br/>Fanpages, tags,<br/>Album Reviews,<br/>Genre descriptions</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>■ Video/Images           <ul style="list-style-type: none"> <li>– Album covers</li> <li>– Music videos</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>■ Other data           <ul style="list-style-type: none"> <li>– Playlists</li> <li>– Market basket</li> <li>– Network of<br/>linked sites,<br/>events</li> </ul> </li> </ul> |
|---|--|---|

## Musical Instrument Digital Interface - MIDI

- Symbolic Music File Format
- Dave Smith, proposed in 1981
- MIDI specification 1.0 in 1983
- Interacting with keyboard produces messages
  - *Note-On, Aftertouch, and Note-Off*
  - 127 note pitches
- Sequence of control commands



<http://www.musica.at/software/arrange/midic2.htm>

## Music IR – Music?

### Features extractable from MIDI representations

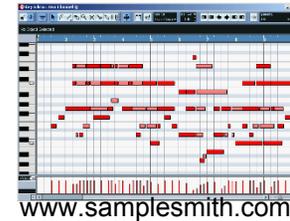
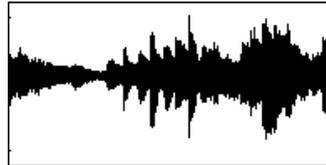
(Tool: e.g. JSymbolic: <http://jmir.sourceforge.net/> )

- Pitch: occurrence rates of diff. notes, pitch classes, ranges, variety
- Rhythm: time intervals, attacks, duration of notes, meters and rhythmic patterns
- Melody: melodic intervals, variation, melodic contours, phrases
- Chords: types of chords, vertical intervals, harmonic movement
- Instrumentation: types of instruments, importance, pitched vs. non-pitched, ...
- Texture: # + rel. importance of independent voices, polyphonic, homophonic
- Dynamics: loudness of notes, variations in dynamics

# Music IR – Music?

## What is „Music“?

- Music, of course!
  - Audio: wav, au, mp3, ...
  - Symbolic: MIDI, mod, ...
  - Scores: Scan, MusicXML



- |   |  |   |
|---|--|---|
| <ul style="list-style-type: none"> <li>■ Text           <ul style="list-style-type: none"> <li>– Song lyrics</li> <li>– Artis Biographies</li> <li>– Websites:<br/>Fanpages, tags,<br/>Album Reviews,<br/>Genre descriptions</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>■ Video/Images           <ul style="list-style-type: none"> <li>– Album covers</li> <li>– Music videos</li> </ul> </li> </ul> | <ul style="list-style-type: none"> <li>■ Other data           <ul style="list-style-type: none"> <li>– Playlists</li> <li>– Market basket</li> <li>– Network of<br/>linked sites,<br/>events</li> </ul> </li> </ul> |
|---|--|---|

## Scores

- Also referred to as „Sheet Music“
- Hand-written or printed form of musical notation
  - Handwritten scores
  - Printed scores
  - Typeset scores
  - MusicXML
- Different IR tasks
  - Scan & Optical Music Recognition (OMR)
  - Score following
  - Melodic retrieval



## Music Typesetting / Scorewriter

- Software used to automate the task of writing and engraving sheet music, aka word processor for text
- Input via text editor or MIDI interface, some support Scan+OMR
- Output: PS/PDF, graphics, MIDI, MusicXML
- Popular programs:
  - GNU LilyPond Software: <http://lilypond.org/>
  - GUIDO Music Notation: <http://www.salieri.org/GUIDO/>
  - Finale: <http://www.finalemusic.com/>
  - Sibelius: <http://www.sibelius.com/>
  - Comprehensive list:  
<http://en.wikipedia.org/wiki/Scorewriter#Scorewriters>

## GNU LilyPond Software

- <http://lilypond.org/>
- Input: UTF-8, no graphical interface  
some graphical editors produce LilyPond output  
(e.g. Rosegarden, NoteEdit, Canorus)
- Output: compiled to PDF, SVG, MIDI, ...
- Notes are entered in note, pitch and length format
- Used by several projects (Mutopia, Musipedia)

## Music IR – Music?

### LilyPond example

(1/5, from [http://en.wikipedia.org/wiki/GNU\\_LilyPond](http://en.wikipedia.org/wiki/GNU_LilyPond))

```
#!lilypond firebreathers.ly -*- coding: utf-8; -*-  
%% Theme to "Fire Breathers", a homebrew NES game perpetually  
%% under development. Composed by Urpo Lankinen.  
%% Note: The composer has made this source code available  
%% to Wikipedia under the GFDL license. Other versions outside  
%% Wikipedia are typically under CC BY-SA license.  
%% This file uses Finnish note names (for example, where  
%% Americans use "F#" and "Bb", Finns use "Fis" and "B").  
%% Dutch note names are used by default.  
\include "suomi.ly"  
%% Optional language upgrade helper.  
\version "2.6.0"
```

## LilyPond example

(2/5, from [http://en.wikipedia.org/wiki/GNU\\_LilyPond](http://en.wikipedia.org/wiki/GNU_LilyPond))

%% The header block defines the titles and texts.

```
\header {  
  title = "Theme to ``Fire Breathers!'"  
  instrument = "For the 2A03 or SID"  
  composer = "Urpo Lankinen"  
  enteredby = "Urpo Lankinen"  
  updatedby = "Jan Nieuwenhuizen"  
  date = "June 2005"  
}
```

## LilyPond example

(3/5, from [http://en.wikipedia.org/wiki/GNU\\_LilyPond](http://en.wikipedia.org/wiki/GNU_LilyPond))

```
Melody = \relative c" {
  \clef treble
  \time 3/4
  \key a \minor
  %% The piece starts with a quarter-note partial bar, "\partial 4"
  %% tells so to LilyPond.
  \partial 4
  a4 | e'4.( d8[ c]) r8 | d4.( c8[ h]) r8 | a2. | e2
  a4 | e'4.( d8[ c]) r8 | d4.( e8[ f]) r8 | e2. | r2
  e4 | f4.( e8[ d]) r8 | d4.( c8[ h]) r8 | a2. | e2
  a4 | e'4.( d8[ c]) r8 | d4.( c8[ h]) r8 | a2. ~ a2 r4 | \bar "." }
  ..... }
```

## LilyPond example

(4/5, from [http://en.wikipedia.org/wiki/GNU\\_LilyPond](http://en.wikipedia.org/wiki/GNU_LilyPond))

%% This is the second voice.

```
SecondVoice = \relative c {
```

```
  \clef bass
```

```
  \time 3/4
```

```
  \key a \minor
```

```
  \partial 4
```

```
  r4 | e2.      | d2.      | a2.      | e2
```

```
  a4 | e'2.     | d2      f4 | e2.     | r2.
```

```
    | f2.      | d2.      | a2.      | e2
```

```
  a4 | e'2.     | d2      h4 | a2.     ~ a2 r4 | \bar "|."
```

```
}
```

.....

# Music IR – Music?

## LilyPond example

(5/5, from [http://en.wikipedia.org/wiki/GNU\\_LilyPond](http://en.wikipedia.org/wiki/GNU_LilyPond))

### Theme to "Fire Breathers!"

For the 2A03 or SID

Urpo Lankinen



The image shows a musical score for the theme to "Fire Breathers!". It consists of three staves. The top staff is in treble clef, and the two bottom staves are in bass clef. The time signature is 3/4. The music is written in a simple, melodic style with a repeating bass line. The score is divided into two systems, with a measure rest at the beginning of the second system.

.....

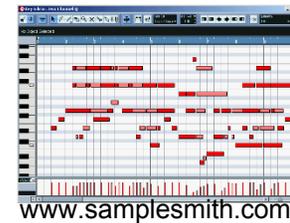
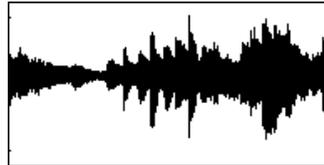
## Features extractable from scores

- Similar to other symbolic notations (MIDI)
- Additionally layout-features, annotations
- Visual features for OMR
- ...

# Music IR – Music?

## What is „Music“?

- Music, of course!
  - Audio: wav, au, mp3, ...
  - Symbolic: MIDI, mod, ...
  - Scores: Scan, MusicXML



### ■ Text

- Song lyrics
- Artis Biographies
- Websites:  
Fanpages, tags,  
Album Reviews,  
Genre descriptions

### ■ Video/Images

- Album covers
- Music videos

### ■ Other data

- Playlists
- Market basket
- Network of  
linked sites,  
events

## Text Data

- Song lyrics
- Artist biographies
- Album reviews
- Tags
- Fan websites
- Genre description sites
- Discussion fora
- ...

## Text: Song lyrics

- Conveys a lot of additional musical information
- Some genres are strongly related with certain texts
- Semantics of music: love songs, christmas songs, ...
- Standard Text-IR: content analysis
- Genre-Analysis: style, rhymes, stop-words, ...
- Lyric portals: plenty of them, some generic, some specialized
  - Lyrics.de
  - lyric.com.ar
  - sing365lyrics.com
  - oldielyrics.com
  - ...

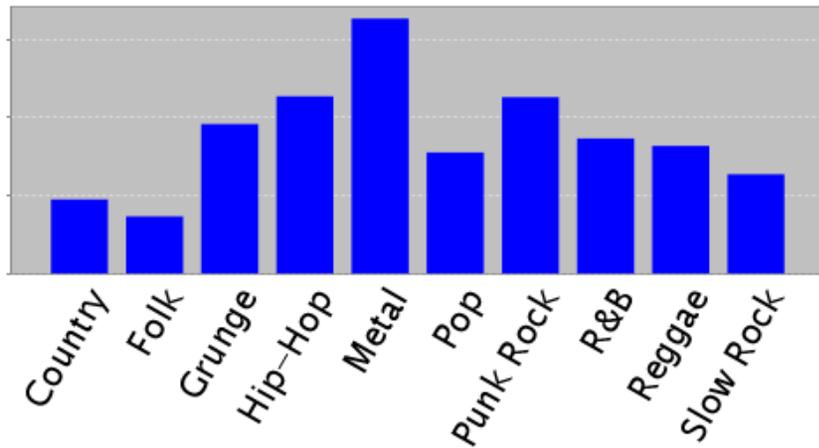
.....

## Features for textual facet of music

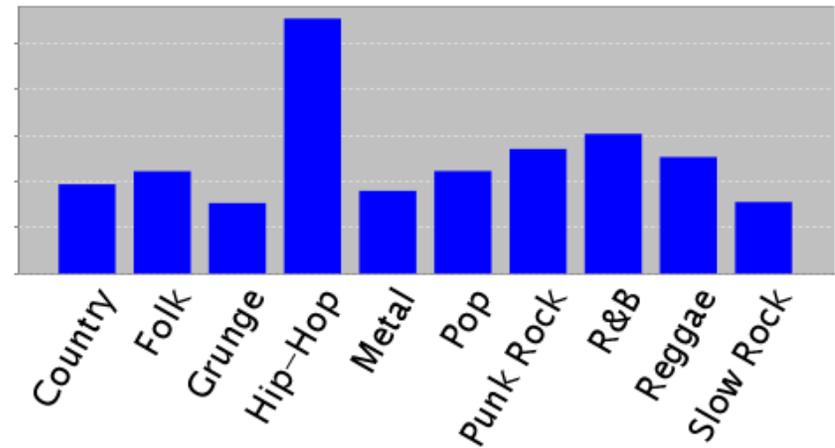
- Standard BOW Features for content analysis
- Specific features for, e.g., lyrics
  - Text Genre Features:
    - ExclamationMark, colon, singleQuote, comma, questionMark, full-stop, hyphen, semicolon
    - Counts of digits d0-d9
    - WordsPerLine, UniqueWordsPerLine, UniqueWordsRatio, CharsPerWord
    - WordsPerMinute
  - PartOfSpeech: nouns, verbs, pronouns, prepositions, adverbs, articles, modals, adjectives
  - Rhyme Features: phoneme transcription + rhyme schemes

## Features for textual facet of music: lyrics features

- Text genre statistics



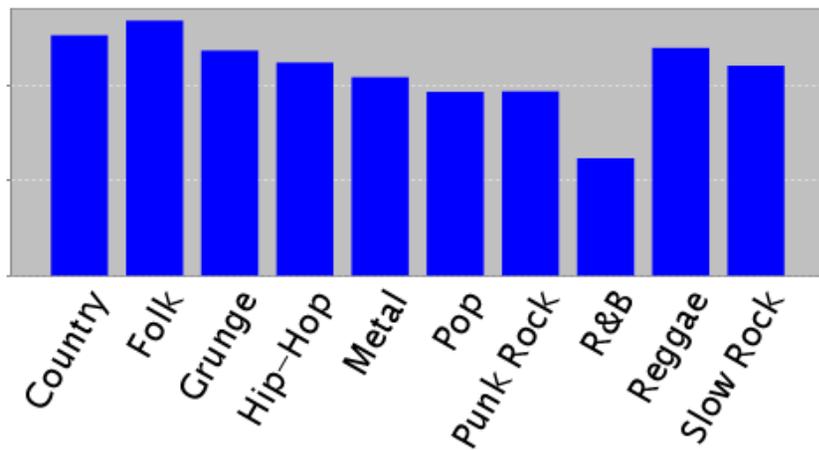
(a) # question marks



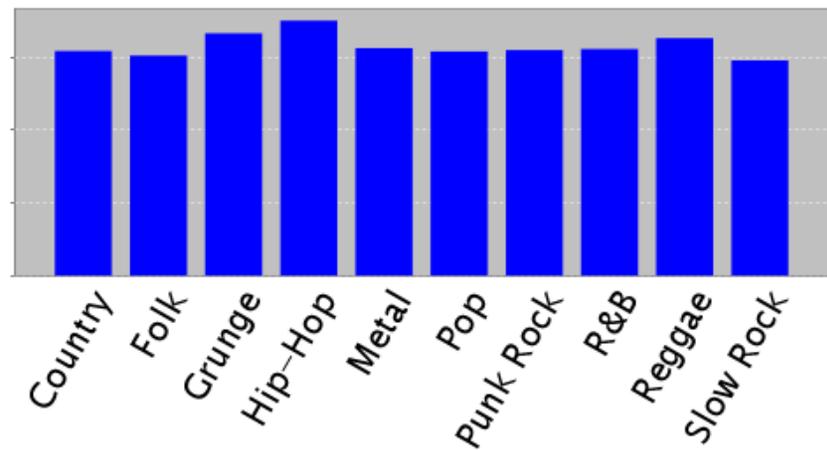
(b) # words per minute

## Features for textual facet of music: lyrics features

- Part-of-Speech statistics



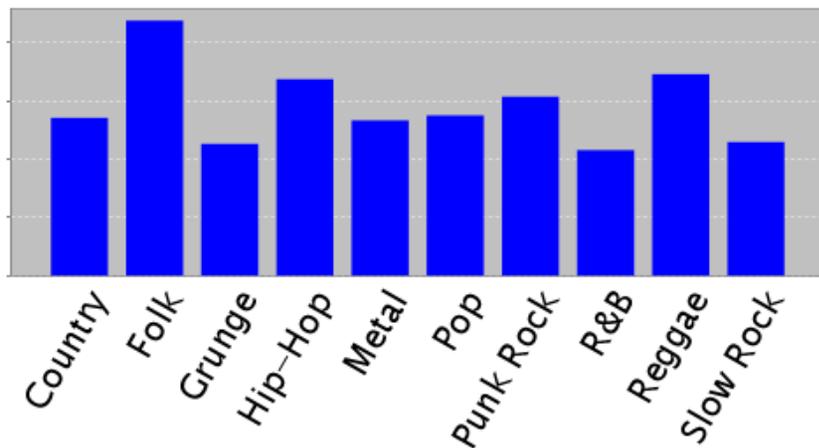
(a) # articles



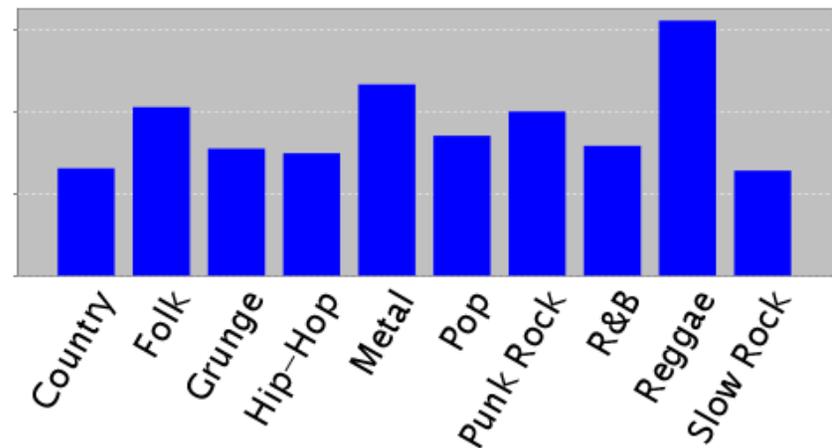
(b) # nouns

# Music IR – Music?

- Features for textual facet of music: lyrics features
- Rhyme feature statistics



(a) # unique rhyme words

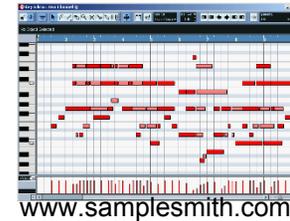
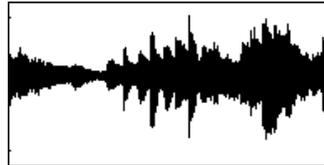


(b) # rhymes AABB

# Music IR – Music?

## What is „Music“?

- Music, of course!
  - Audio: wav, au, mp3, ...
  - Symbolic: MIDI, mod, ...
  - Scores: Scan, MusicXML



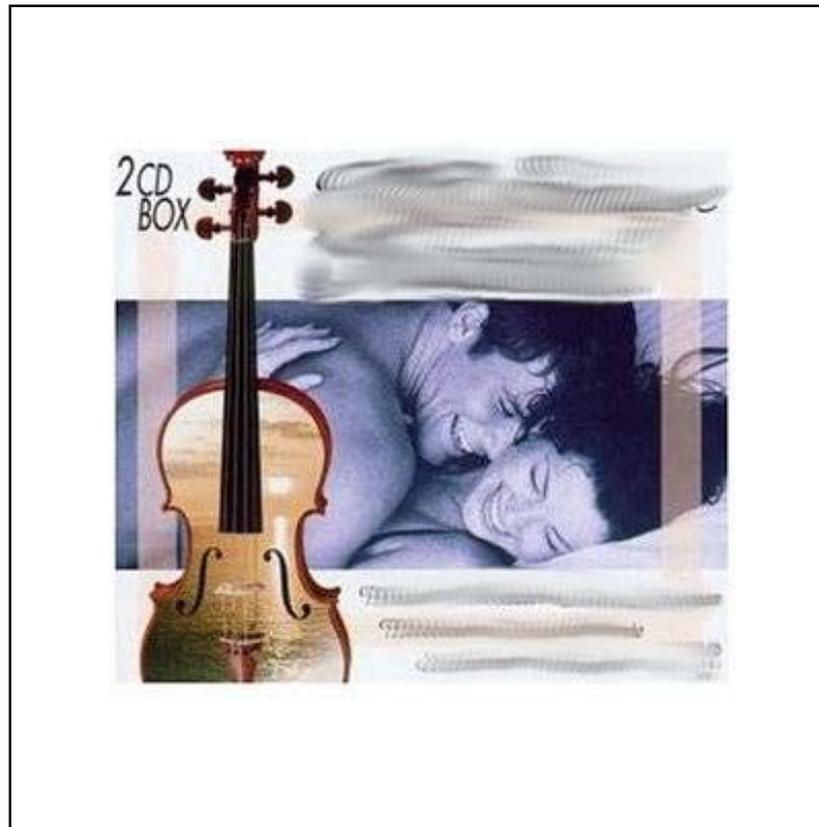
- Text
  - Song lyrics
  - Artis Biographies
  - Websites:  
Fanpages, tags,  
Album Reviews,  
Genre descriptions

- Video/Images
    - Album covers
    - Music videos
    - Artist websites

- Other data
  - Playlists
  - Market basket
  - Network of  
linked sites,  
events

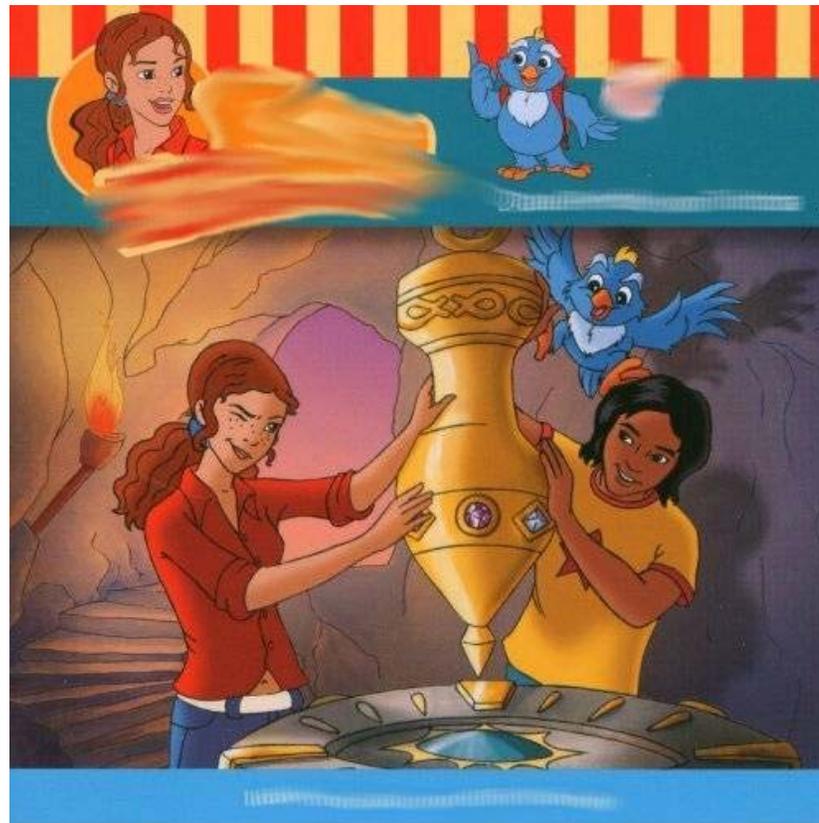
## Music IR – Music?

- There is more to music than sound and text
- Which genre is this album?



## Music IR – Music?

- There is more to music than sound and text
- Which genre is this album?



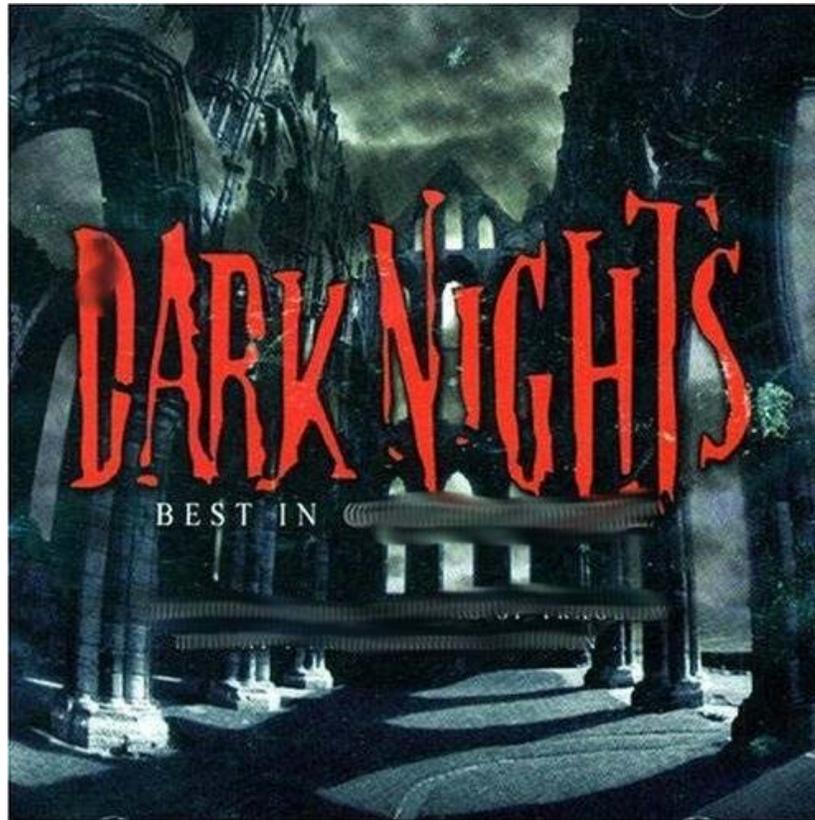
## Music IR – Music?

- There is more to music than sound and text
- Which genre is this album?



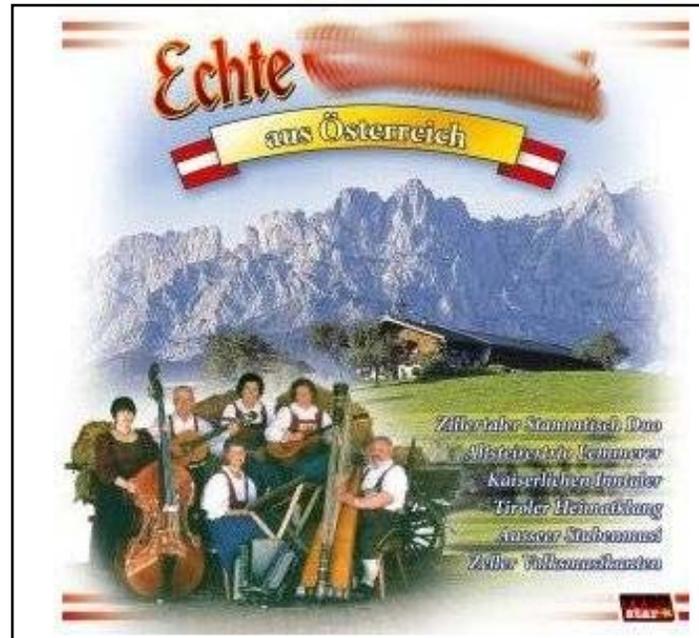
## Music IR – Music?

- There is more to music than sound and text
- Which genre is this album?



## Music IR – Music?

- There is more to music than sound and text
- Which genre is this album?



## Music IR – Music?

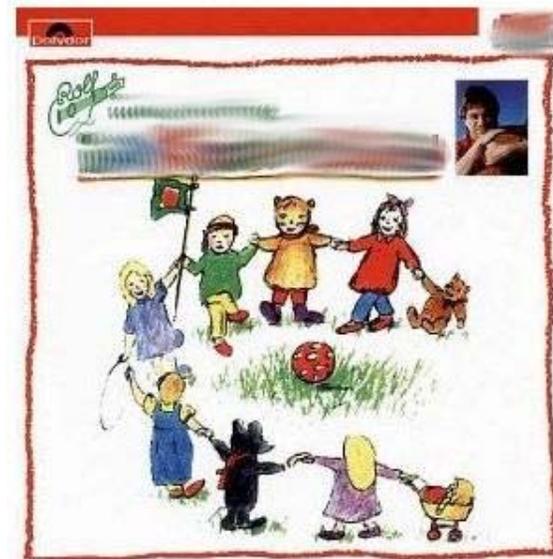
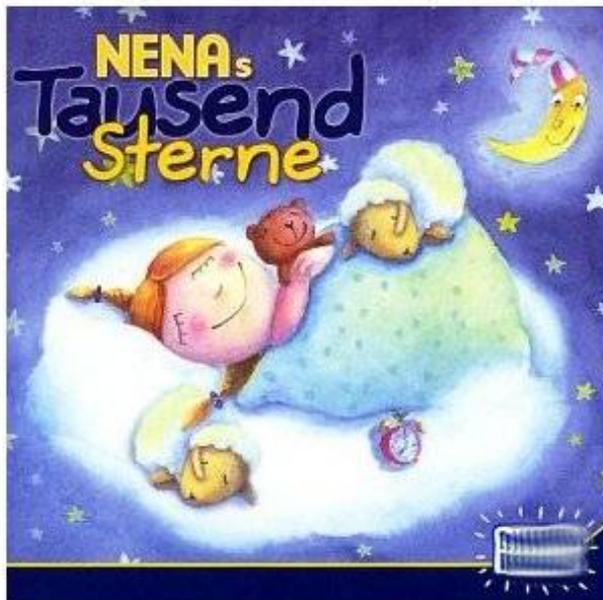
- There is more to music than sound and text
- Which genre is this album?





# Audio, Text and Images

- There is more to music than sound and text
- Which genre is this album?

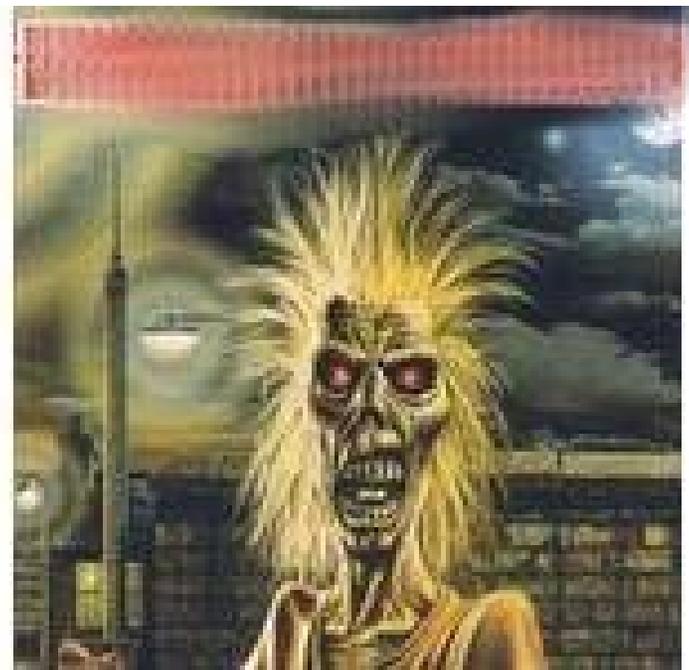
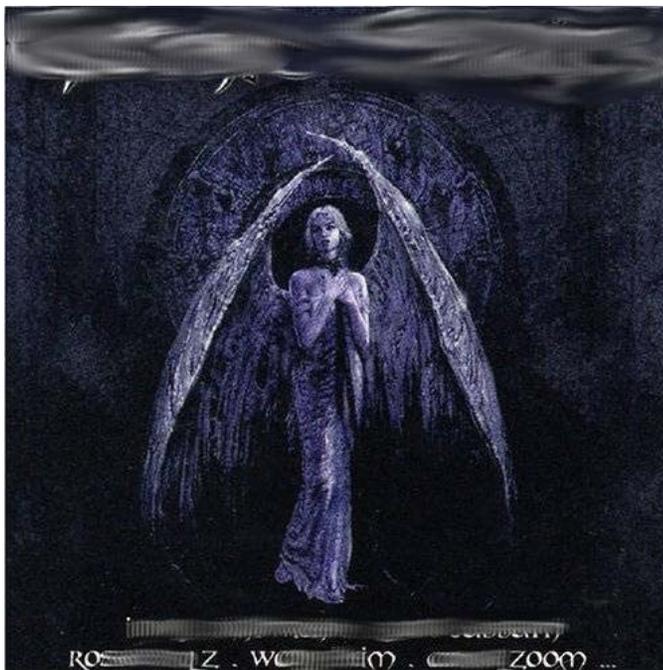


.....



# Audio, Text and Images

- There is more to music than sound and text
- Which genre is this album?



.....

## Image / Video

- Album covers
- Music videos
- Carefully designed to convey a specific information  
Style, Image, Character
- Hardly exploited so far
- Indications, that humans are able to deduce music genre from album covers.

(Sally Jo Cunningham: „What People Do When They Look for Music: Implications for Design of a Music Digital Library“, Proc. Of ICADL 2002)



## Image / Video

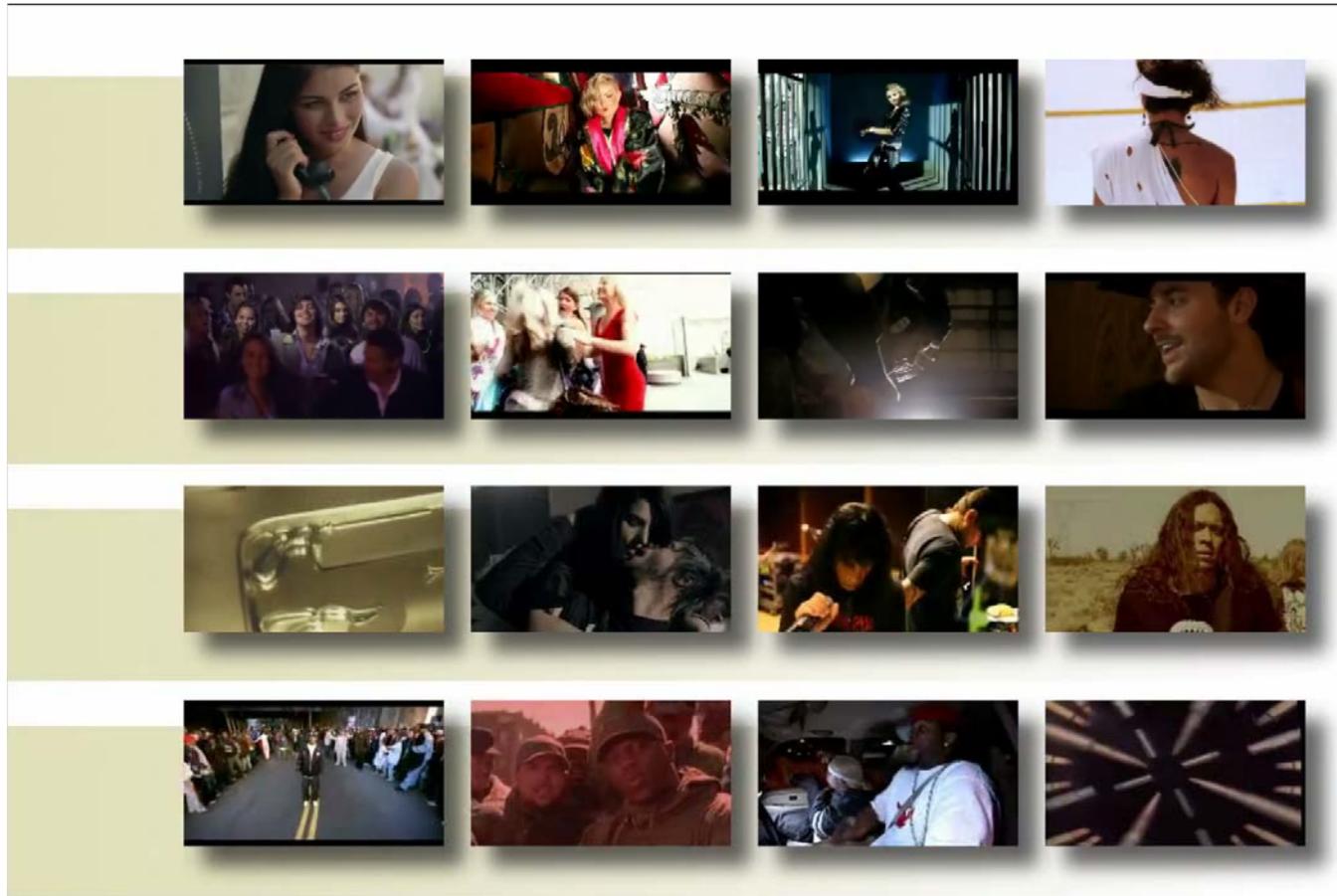
- Analyse correlation between color and emotions
- Analyse dynamics in music videos and music
- Face detection and object recognition
- Alexander Schindler. A picture is worth a thousand songs: Exploring visual aspects of music. In *Proceedings of the 1st International Digital Libraries for Musicology workshop (DLfM 2014)*, London, UK, September 12 2014.
- Alexander Schindler and Andreas Rauber. A music video information retrieval approach to artist identification. In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR2013)*, Marseille, France, October 14-18 2013



<https://prezi.com/v7xa3nrogq1c/visual-music-computing/>

# Music IR – Music?

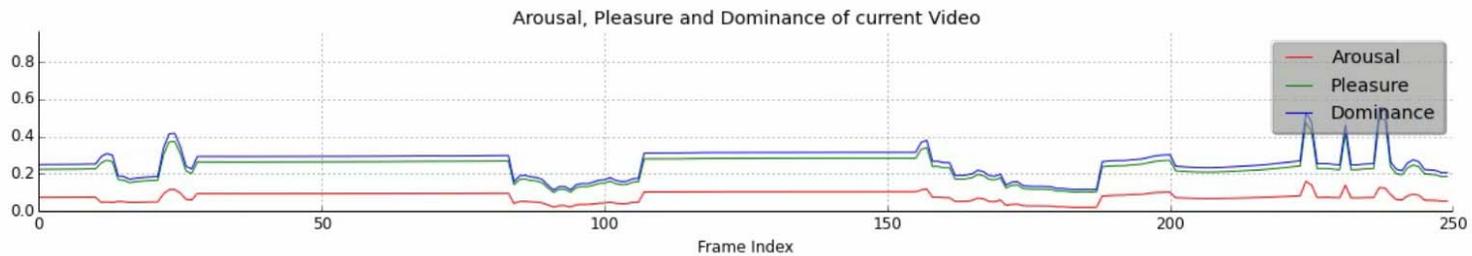
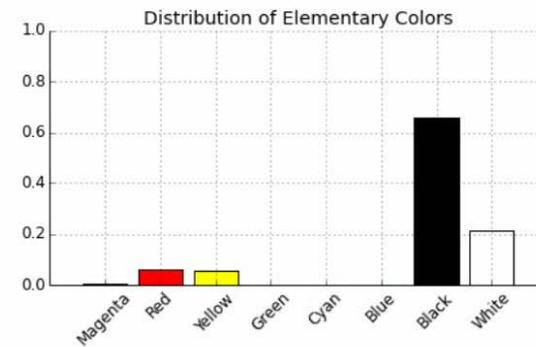
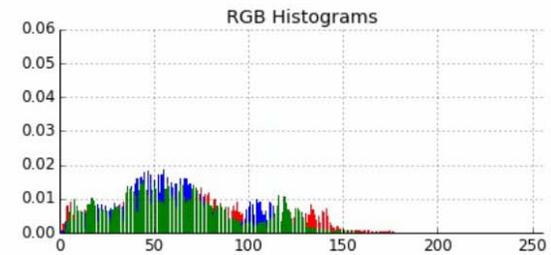
## Image / Video



.....

# Music IR – Music?

## Image / Video



## Music IR – Music?

### Features for visual facet

- Standard Image IR features: color histograms, texture, ...
- Face detection/recognition
- Object detection

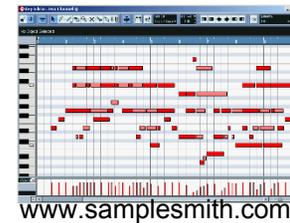
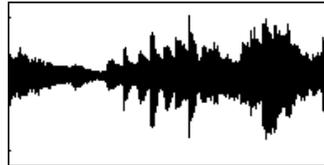
### For music videos

- Cut frequency, type of cuts, camera movement, ...

# Music IR – Music?

## What is „Music“?

- Music, of course!
  - Audio: wav, au, mp3, ...
  - Symbolic: MIDI, mod, ...
  - Scores: Scan, MusicXML



- Text
  - Song lyrics
  - Artis Biographies
  - Websites:  
Fanpages, tags,  
Album Reviews,  
Genre descriptions

- Video/Images
  - Album covers
  - Music videos
  - Artist websites

- Other data
  - Playlists
  - Market basket
  - Network of  
linked sites,  
events

## Other Data

- Playlists as sets or sequences, co-occurrence (market basket)
- Network/Graph analysis of band compositions
- Cover versions
- Technical metadata of recordings
  
- Dedicated feature representations & analysis techniques

### What is Music IR?

- Searching for Music, of course!
  - Searching for music on the Web
  - Query by Humming
  - Similarity Retrieval
  - Identity detecting (fingerprinting)
- Plenty of other tasks!

### What is Music IR? - Other tasks

- Genre classification
- Mood classification
- Artist identification
- Artist similarity
- Cover song detection
- Rhythm and beat detection
- Score following
- Chord detection
- Automatic Composition
- Audio segmentation
- Instrument detection
- Automatic source separation
- Onset detection
- Optical music recognition
- Melody transcription
- Symbolic music similarity

## Music IR material

### ■ Papers

- Stephen Downie: *Music information retrieval*. Annual Review of Information Science and Technology 37: 295-340. 2003.  
[http://www.isrl.uiuc.edu/~music-ir/downie\\_mir\\_arist37.pdf](http://www.isrl.uiuc.edu/~music-ir/downie_mir_arist37.pdf)
- Nicola Orio: *Music Retrieval: A Tutorial and Review*. In: Foundations and Trends in Information Retrieval, Volume 1 Issue 1. 2006.  
<http://www.nowpublishers.com/getpdf.aspx?doi=1500000002&product=INR>

### ■ Conferences

- ISMIR: International Conference on Music Information Retrieval
- DAFx: Conference on Digital Audio Effects
- ICMC: International Computer Music Conference
- other Multimedia, Information Retrieval, and Digital Library Conferences

### ■ Journals

- ICMJ: International Computer Music Journal
- JNMR: Journal on New Music Research



## Music IR material

- On-line class

- Started Jan 6 2014

- <http://marsyas.cs.uvic.ca/mirBook/course/>

- Participation via

- <https://plus.google.com/communities/102479783616723367135>



- Companion Text Book

- George Tzanetakis: Music Information Retrieval

- <http://marsyas.cs.uvic.ca/mirBook/book/>

- (draft version)

## Chorus

- Lead-in
- Verse 1: What is Music-IR?
- **Chorus**
- **Verse 2: Audio Features**
- Verse 3: Retrieval and other applications
- Fade-out

## Audio Features

- So far we talked about music IR...  
... without ever touching upon music itself
- Retrieval based on audio
- Need to calculate features from audio
- Text: bag of words, n-grams, phrases, POS,...
- Music: ???
- What have we got?

Beethoven:  
Per Elisa



Korn:  
Freak on a Leash



## General principles

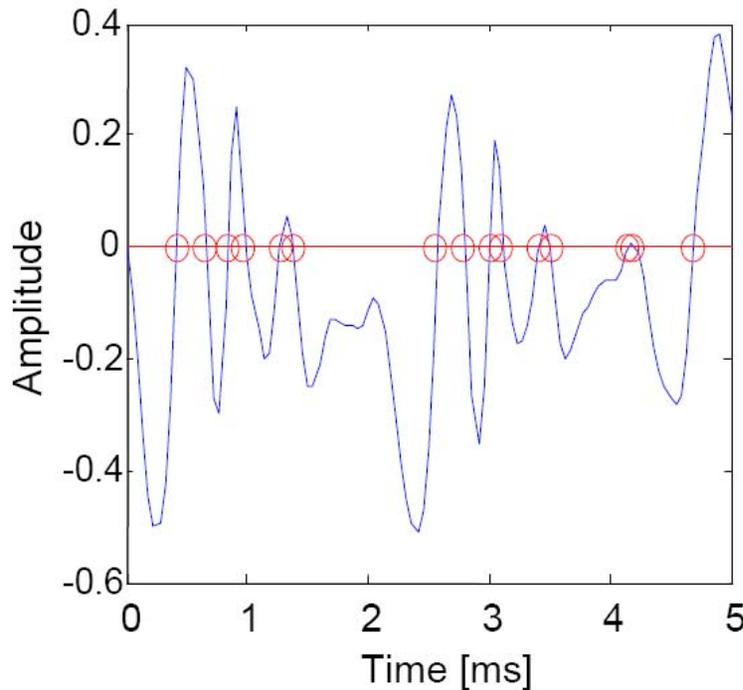
- Extract characteristics from audio signal
- Model certain aspects of what we perceive
  - Melodic
  - Rhythmic
  - Timbre
- Usually represented in high-dim vector space
- Used as basis for computing similarity:
  - Retrieval
  - Classification, clustering,...

## Audio Features

- A number of features can be calculated
  - Zero-crossings
  - Statistical characteristics of the spectrum
  - MFCCs
  - Rhythm Patterns
  - Rhythm Histograms
  - SSDs
- Capture different characteristics of sound
- Have different dimensionality
- Perform differently on different tasks

# Time-Domain Features

Zero Crossing Rate (ZCR) = 3/ms



= 15 / 5ms

measures noisiness

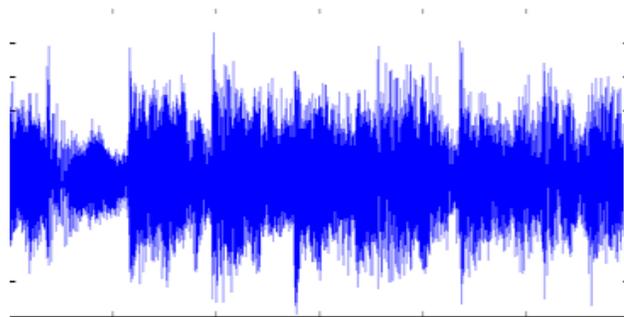
“crossing zero” is defined as:

$(x[n-1] < 0 \text{ and } x[n] > 0)$

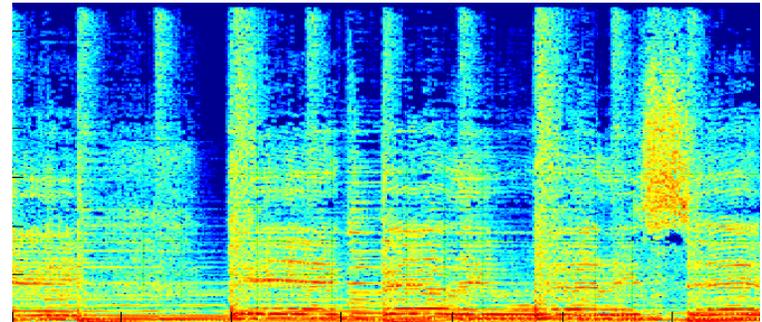
or  $(x[n-1] > 0 \text{ and } x[n] < 0)$

or  $(x[n-1] \neq 0 \text{ and } x[n] = 0)$ .

# Signal Processing



Time Domain  
(„Wave Form“)



Frequency Domain  
(„Spectrum“)

## Time-Frequency Transformation

- Fourier Transform (FFT)
- Discrete Cosine Transform (DCT)
- Wavelet Transform

.....

# Spectral Features

- Spectral Centroid
  - center of gravity (balancing point of the spectrum)
  - gives an indication of how “dark” or “bright” a sound is

$$SC = \frac{\sum_{n=1}^N P_t[n] * n}{\sum_{n=1}^N P_t[n]}$$

$P_t[n]$  ...  $n^{\text{th}}$  frequency bin of **power spectrum** (with N bins)  
t ... timeframe

# Spectral Features

- Spectral Rolloff

- the frequency below which some fraction,  $k$  (typically 0.85, 0.9 or 0.95 percentile), of the cumulative spectral power resides
- measure of the skewness of the spectral shape
- indication of how much energy is in the lower frequencies

$$\sum_{n=1}^{SR_t} P_t[n] = k \sum_{n=1}^N P_t[n]$$

$P_t[n]$  ...  $n^{\text{th}}$  frequency bin of **power spectrum** (with  $N$  bins)  
 $t$  ... timeframe

# Spectral Features

- Spectral Flux
  - squared differences in frequency distribution of two successive time frames
  - measures the rate of local change in the spectrum

$$SF_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2$$

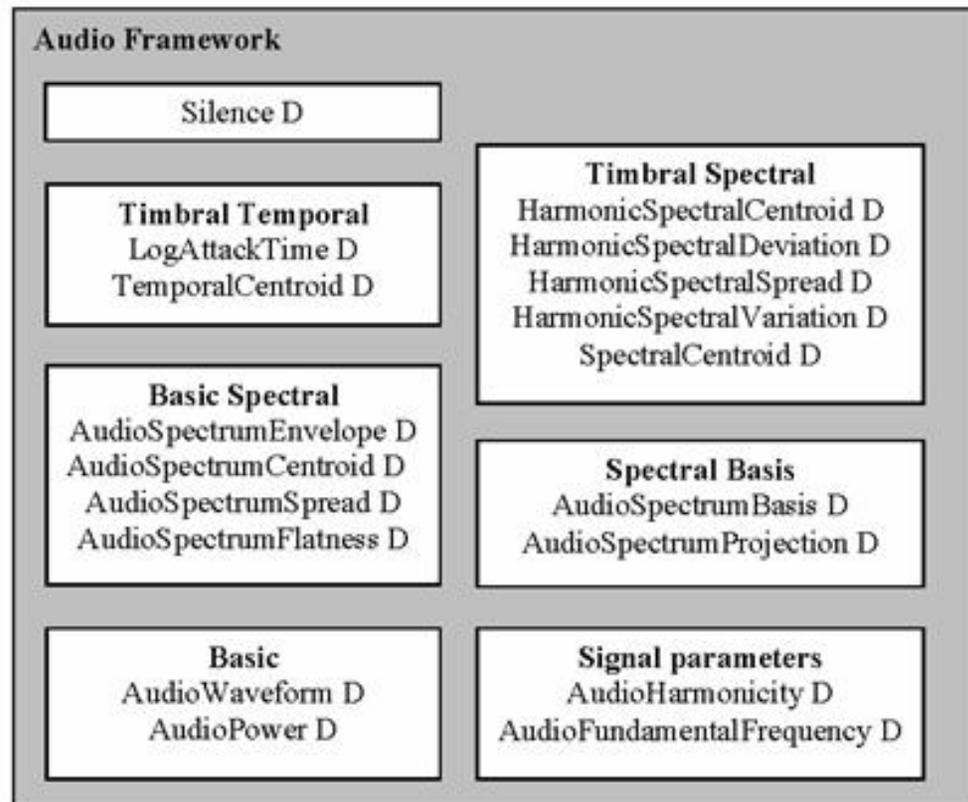
computed from the **normalized magnitude spectrum**  $N_t[n]$

# Spectral Features

- **Spectral Variability:**
  - standard deviation of the bin values of the magnitude spectrum
  - provides an indication of how flat the spectrum is and if some frequency regions are much more prominent than others
- **Strongest Partial:**
  - center frequency of the bin of the magnitude or power spectrum with the greatest strength
  - can provide a primitive form of pitch tracking
- and others ...

# MPEG7 Features

## MPEG7 Features

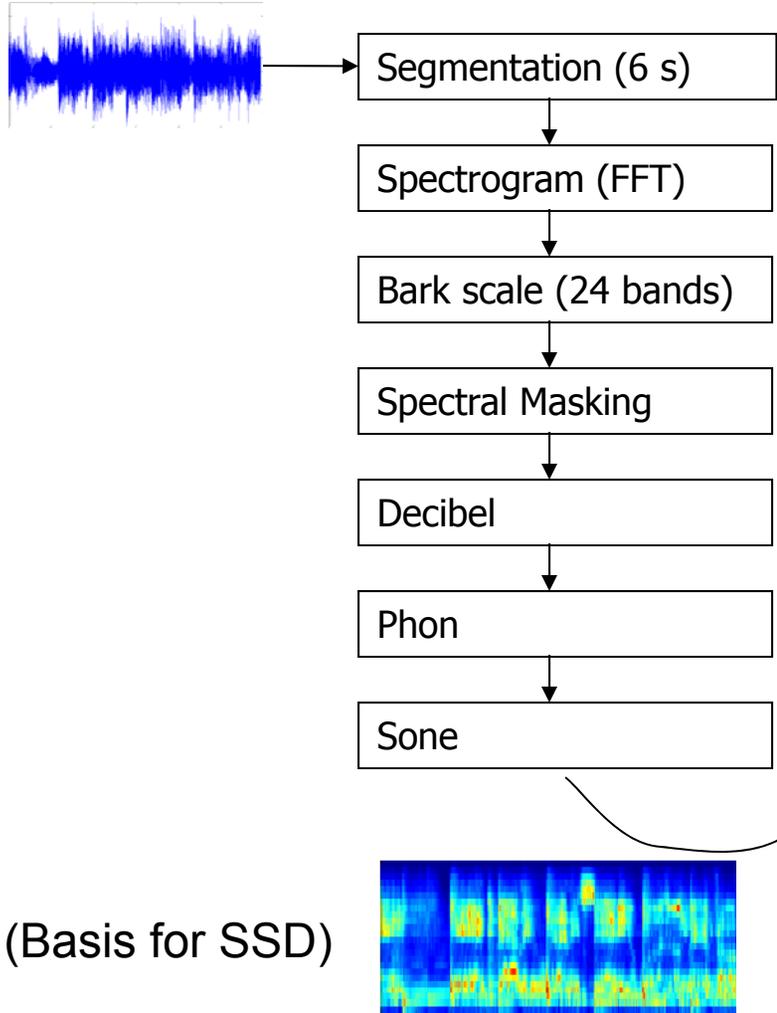


## MFCC Features

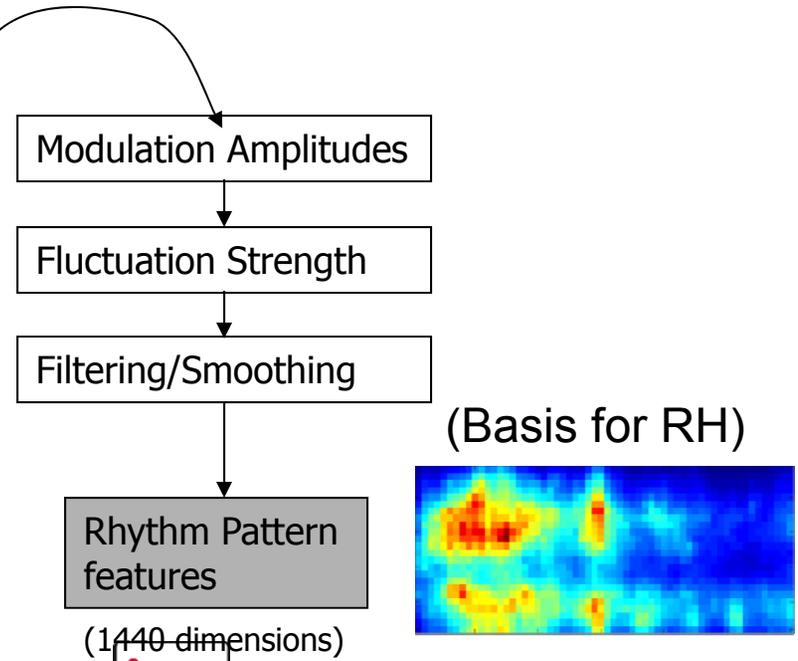
### Mel-Frequency Cepstral Coefficients (MFCC)

- used previously in speech recognition
- model human auditory response (Mel scale)
- „cepstrum“ (s-p-e-c reversed): result of taking the Fourier transform (FFT) of the decibel spectrum as if it were a signal
- show rate of change in the different spectrum bands
- good timbre feature

# Rhythm Pattern (RP)



E. Pampalk, A. Rauber, D. Merkl:  
**Content-based Organization and Visualization of Music Archives**  
 In: Proceedings of ACM Multimedia 2002, pp. 570-579, December 1-6, 2002, Juan-les-Pins, France.

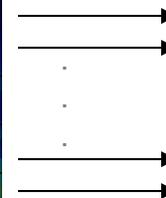
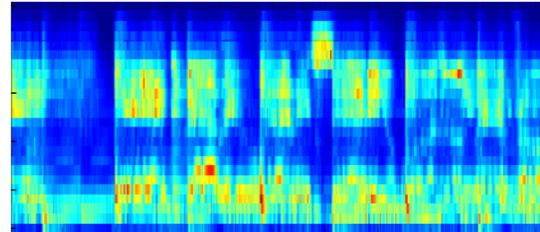


.....

## Statistical Spectrum Descriptor (SSD):

- description of each of the 24 critical bands of the Sonogram by 7 statistical measures
- 168 feature attributes (24x7)

24  
critical  
Bands  
(Bark scale)

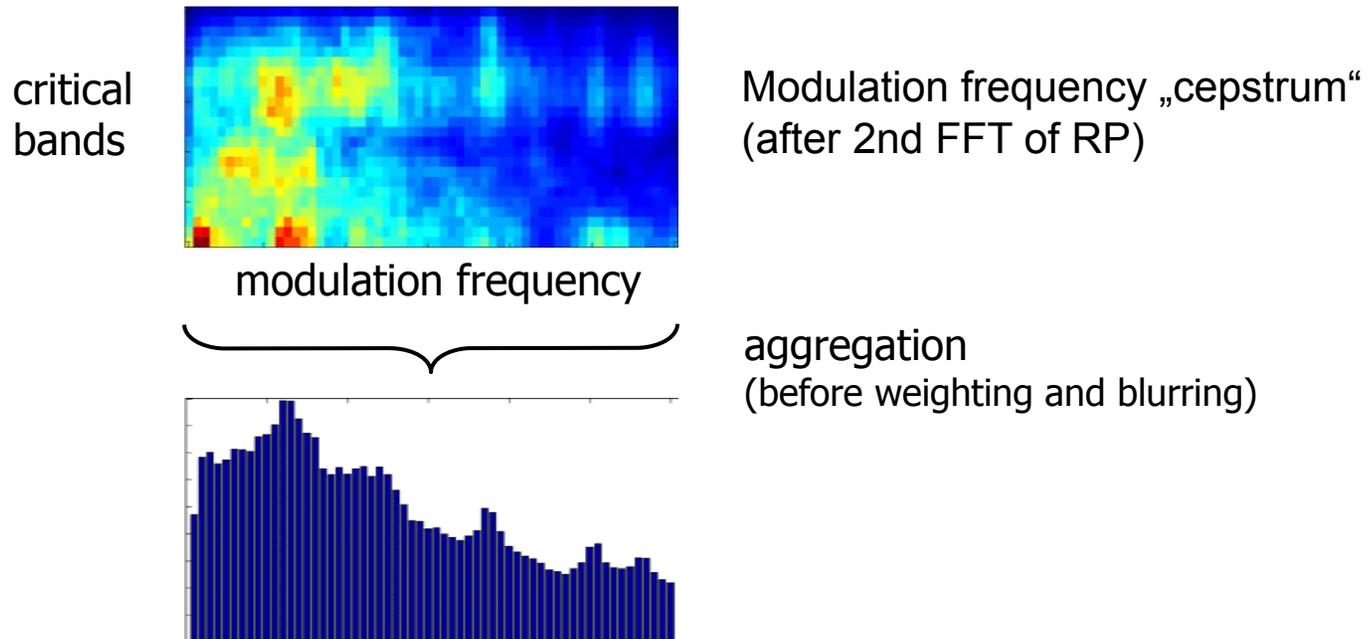


mean  
median  
variance  
skewness  
kurtosis  
min  
max

Bark-scale Sonogram (after Sone Step of RP)

- Thomas Lidy, Andreas Rauber.  
**MIREX 2006: Computing Statistical Spectrum Descriptors for Audio Music Similarity and Retrieval.** [MIREX 2006 - Music Information Retrieval Evaluation eXchange](#), Victoria, Canada, October 8-12, 2006.

## Rhythm Histogram (RH)



- histogram of modulation magnitude per modulation frequency
- 60 bins -> 60 feature attributes

# Indexing Audio

## ■ Software for Audio Features

- RP/RH/SSD extractor (TU Wien, Matlab, Java, Webservice)  
<http://www.ifs.tuwien.ac.at/mir/downloads.html>
- Marsyas Framework (George Tzanetakis, C++)  
<http://sourceforge.net/projects/marsyas>
- BeatRoot (Simon Dixon, Java)  
<http://www.elec.qmul.ac.uk/people/simond/beatroot/index.html>
- jMIR: jAudio, jSymbolic (Cory McKay, Java)  
<http://jmir.sourceforge.net>
- MIRToolbox (Olivier Lartillot, Matlab)  
<http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- CLAM (MTG, UPF Barcelona, C++ Library)  
<http://clam-project.org>
- M2K:Flow-chart based toolbox (Stephen Downie et al., Java)  
<http://www.music-ir.org/evaluation/m2k>
- Echonest (commercial, Web API)  
<http://developer.echonest.com>

## Chorus

- Lead-in
- Verse 1: What is Music-IR?
- Verse 2: Audio Features
- **Chorus**
- **Verse 3: Retrieval and other applications**
- Fade-out



# Retrieval and Classification

## Example 1: classical song

Query Song: classical\_2-fruhlingsnacht.mp3 

Top 5 and #10 similar songs according to different feature sets:

Features: Rank:	RH	SSD	RP	MFCC
1.	 classic	 classic	 classic	 classic
2.	 classic	 classic	 classic	 classic
3.	 classic	 classic	 classic	 classic
4.	 world	 classic	 classic	 classic
5.	 classic	 classic	 classic	 classic
10.	 classic	 classic	 classic	 classic

.....



# Retrieval and Classification

Query Song: rock\_pop\_1-nocturne.mp3 

Top 5 and #10 similar songs according to different feature sets:

Features: Rank:	RH	SSD	RP	MFCC
1.	 rock_pop	 rock_pop	 rock_pop	 rock_pop
2.	 world	 rock_pop	 rock_pop	 electronic
3.	 electronic	 world	 world	 electronic
4.	 jazz_blues	 electronic	 jazz_blues	 electronic
5.	 rock_pop	 electronic	 metal_punk	 electronic
10.	 rock_pop	 rock_pop	 metal_punk	 electronic



# Retrieval and Classification

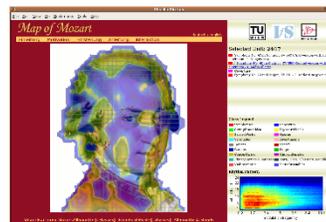
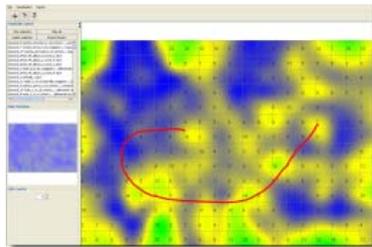
Query Song: electronic\_10-walking\_safely.mp3 

Top 5 and #10 similar songs according to different feature sets:

Features: Rank:	RH	SSD	RP	MFCC
1.	 rock_pop	 electronic	 rock_pop	 rock_pop
2.	 rock_pop	 jazz_blues	 rock_pop	 world
3.	 world	 rock_pop	 world	 electronic
4.	 electronic	 rock_pop	 rock_pop	 rock_pop
5.	 world	 rock_pop	 world	 metal_punk
10.	 electronic	 rock_pop	 metal_punk	 rock_pop

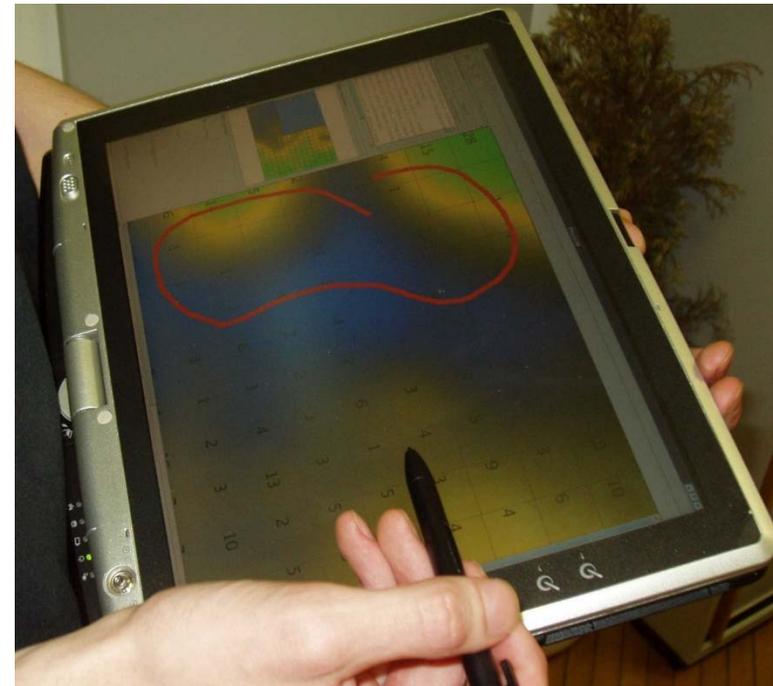
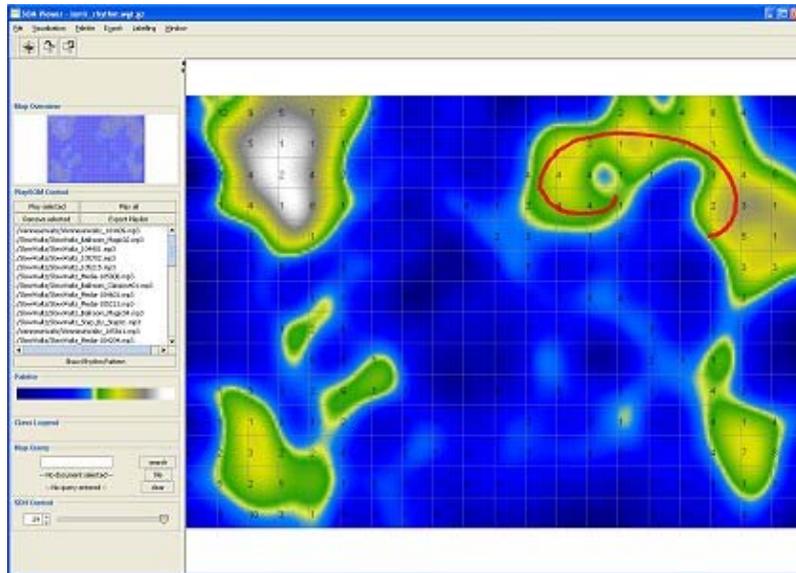
# Clustering & Browsing

- Need new interfaces to access huge music archives
- SOMeJB: SOM-enhanced Jukebox
- Cluster music (by feature sets)
- Based on Self-Organizing Map (SOM)
- Mapping from input- to output space ("2 dim. map")
- Preservation of Neighbourhood relationships
- Map of music space
- PlaySOM and PocketSOM applications



## PlaySOM

- Organizing Music
- Creating Playlists

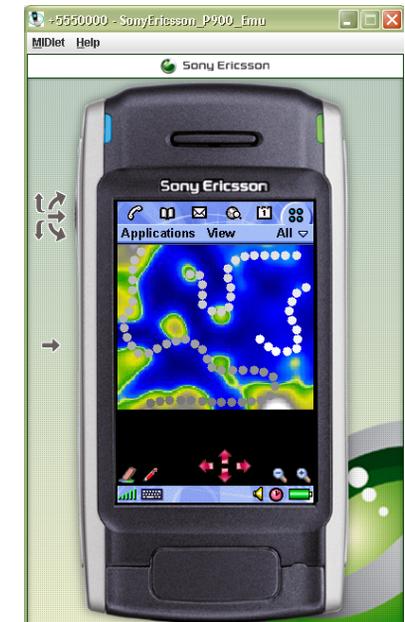


.....

## PocketSOM-Player

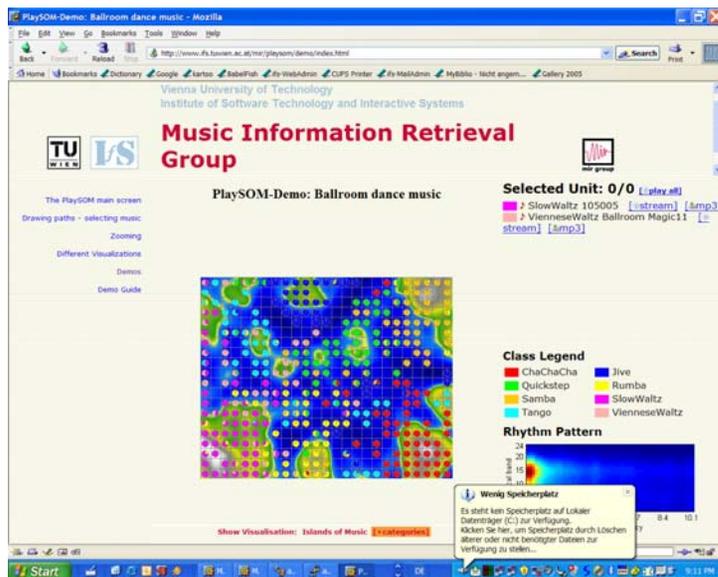
- Application for mobile devices
- Streaming audio
- Remote control

- <http://www.ifs.tuwien.ac.at/mir/pocketsom>

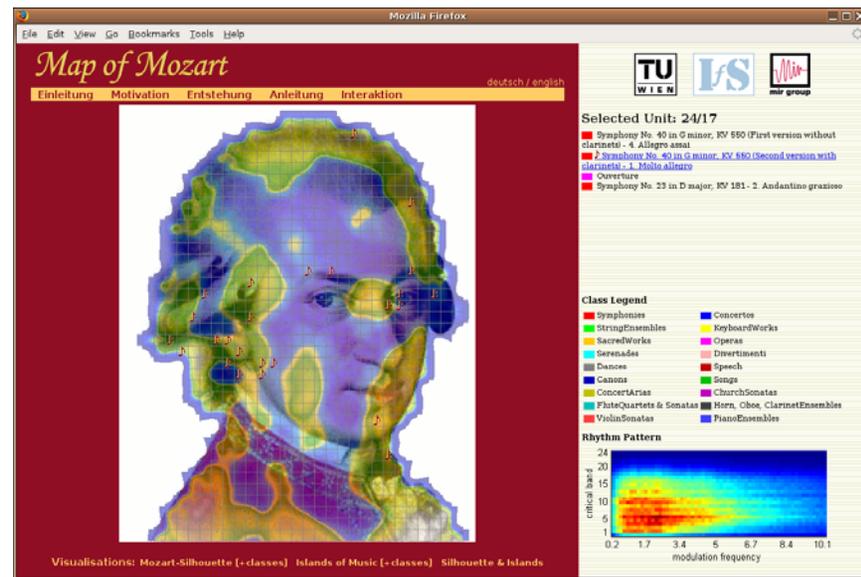


## Web-based Browsing

- Web-based interface
- Reduced functionality



<http://www.ifs.tuwien.ac.at/mir/playsom/demo>



[http://www.ifs.tuwien.ac.at/mir/mozart/index\\_en.html](http://www.ifs.tuwien.ac.at/mir/mozart/index_en.html)

## Summary

- numerous Music IR algorithms exist already
- numerous commercial applications based on Music IR already exist as well
- but still there is a large number of open issues
- benchmarking and evaluation is important, but also faces challenges
- it's a very interesting domain!  
... still a lot of research to do! get involved!

# Chorus

- Lead-in
- Verse 1: What is Music-IR?
- Verse 2: Audio Features
- Verse 3: Retrieval and other applications
- **Chorus**
- **Fade-out**

### You have heard a lot about Music IR

- Different types of music representation
- Different types of musical information
- Features we can compute from audio
- State of the art in retrieval, classification
- Evaluation and benchmarking challenges
- Applications for browsing music collections
- Challenging application scenarios

## Fade-out

### But

- There is a lot more to learn...
- ...and a lot of open problems to solve!
- Music IR is a very young discipline
- Many surprises, unknown territory waiting to be explored
- I hope this presentation has
  - given you some interesting and new information
  - inspired you to pick up challenging research questions in this field



# Fade-out

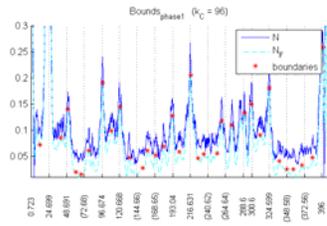


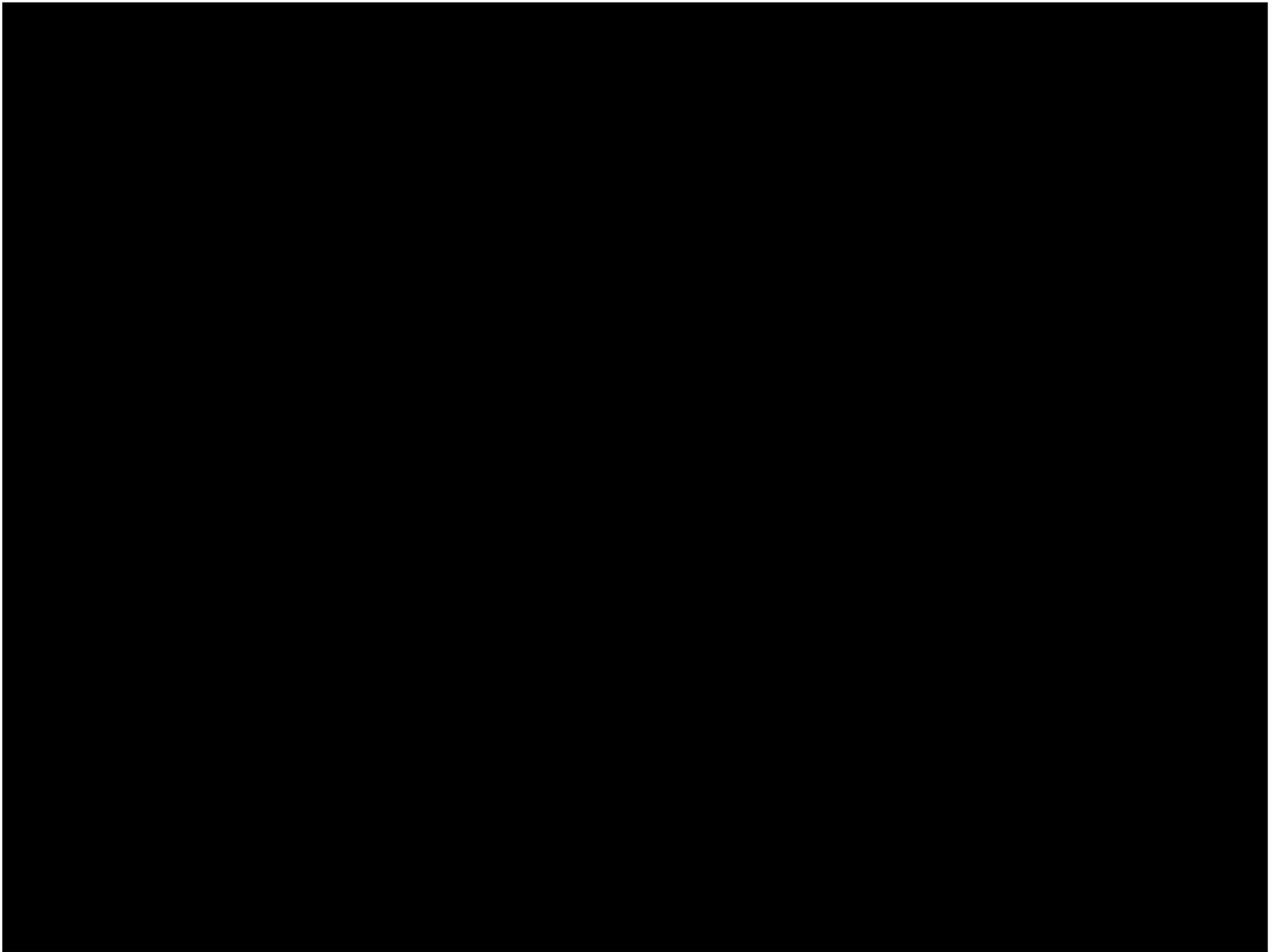
<http://www.ifs.tuwien.ac.at/mir>

.....



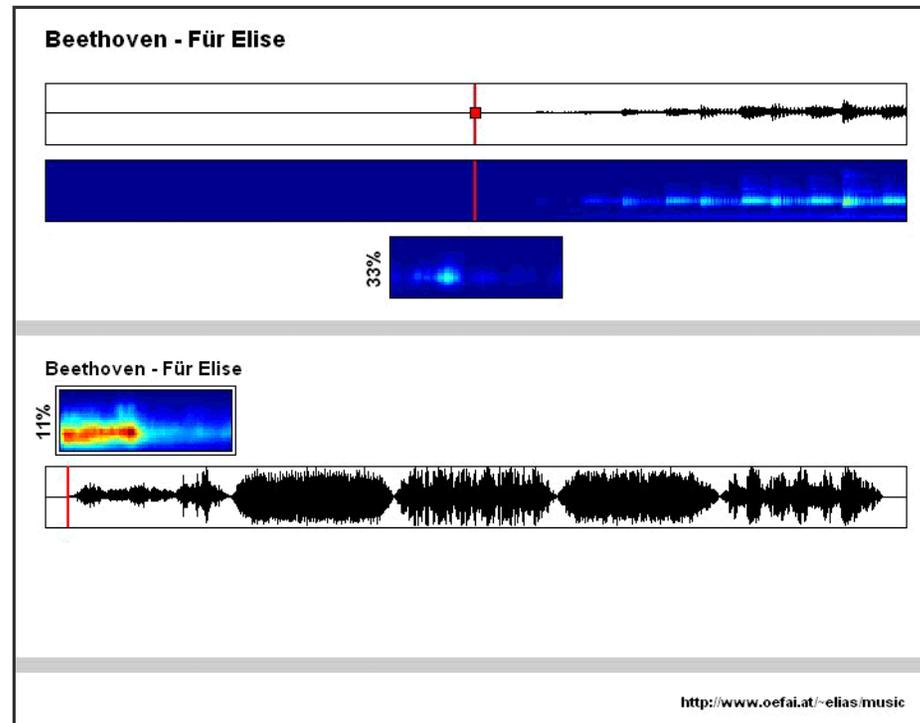
FACULTY OF **INFORMATICS**





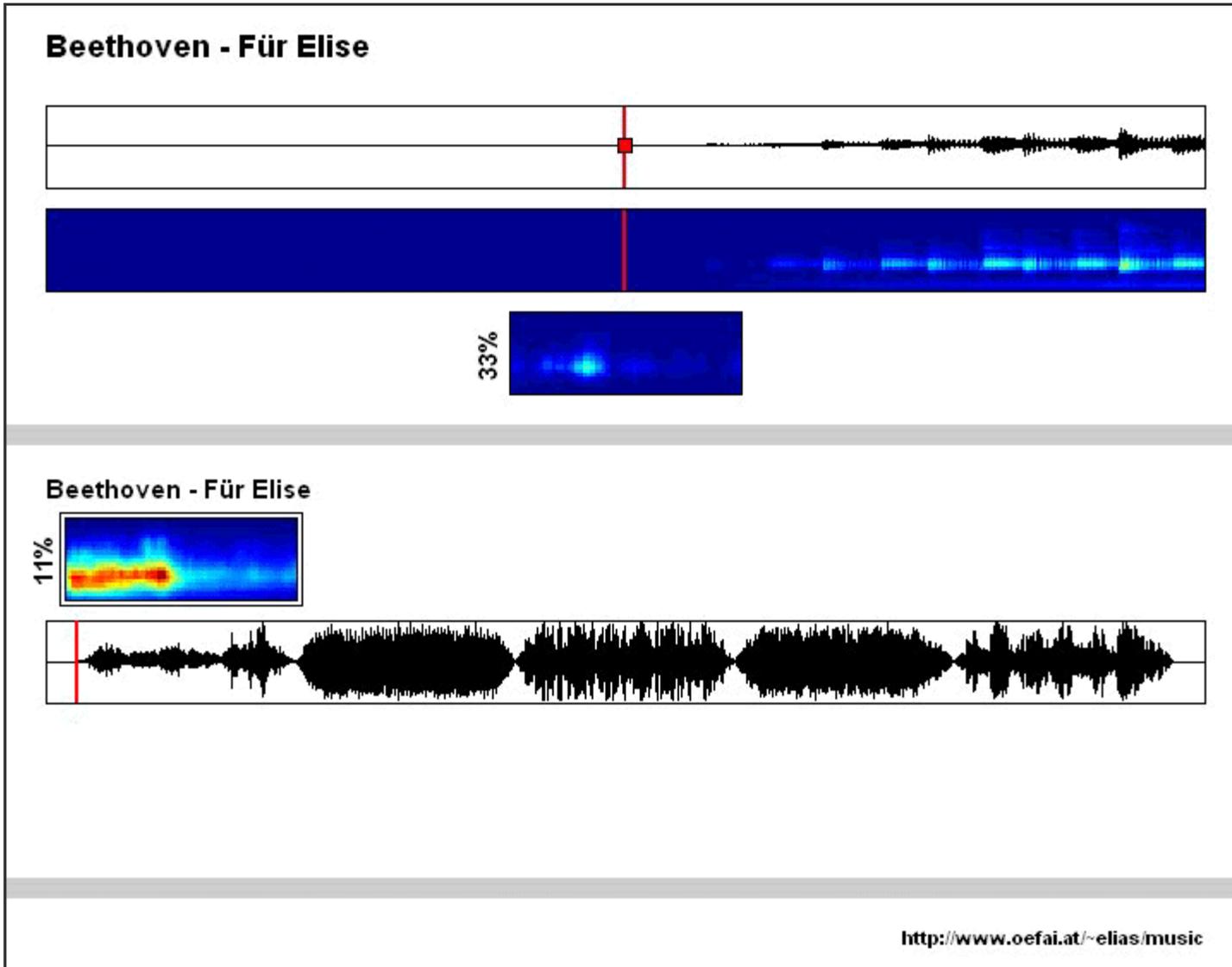
## Rhythm Patterns

- Demo Video  
(prepared by Elias Pampalk)





# Feature Extraction



.....



# Feature Extraction

Beethoven - Für Elise

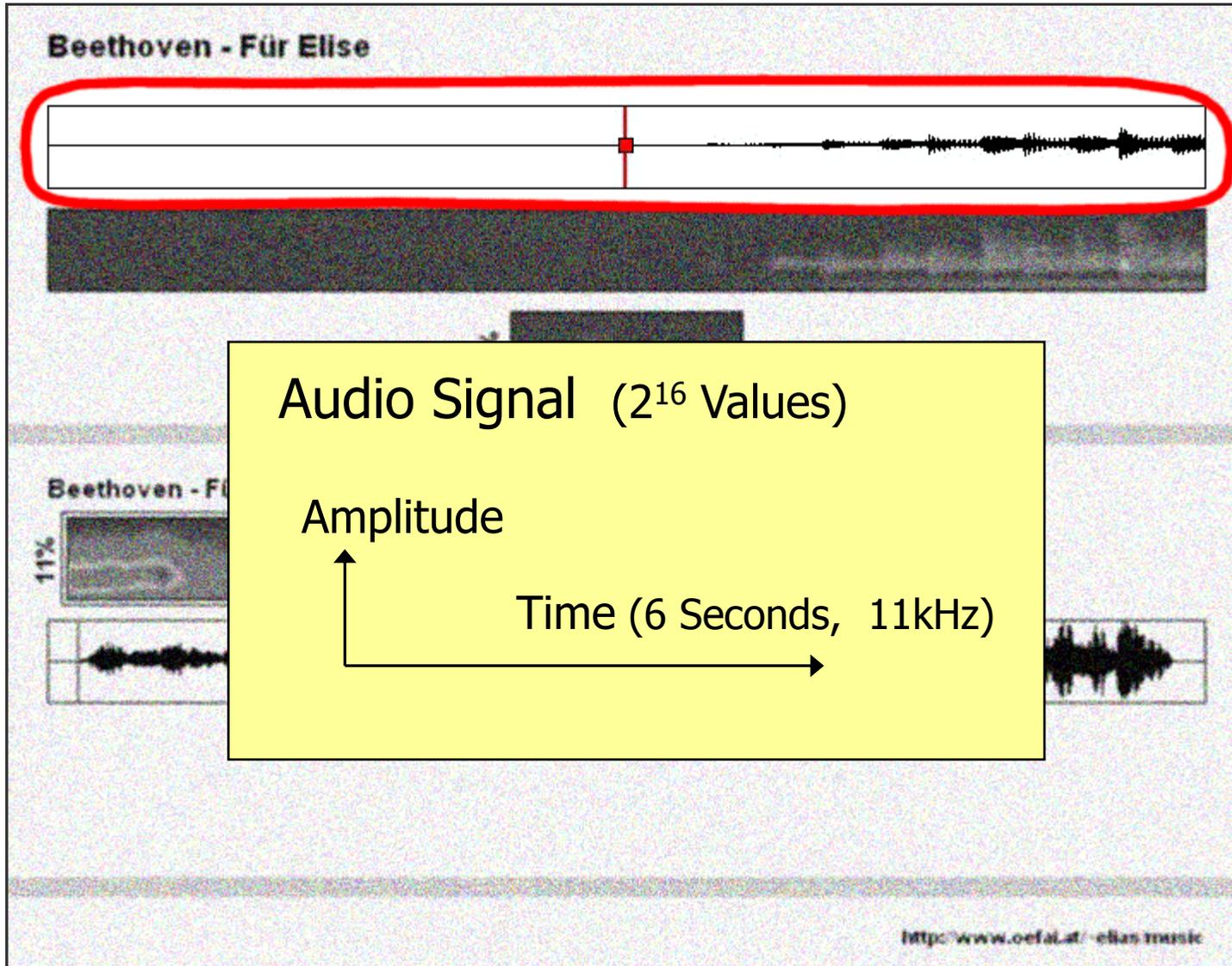
5 x 20 Seconds

- Beethoven – Für Elise
- K's Choice – Not An Addict
- Sublime – Doin' Time
- Korn – Freak On A Leash
- Boomfunk MCs – Freestyler

<http://www.oefal.at/~elias.music>



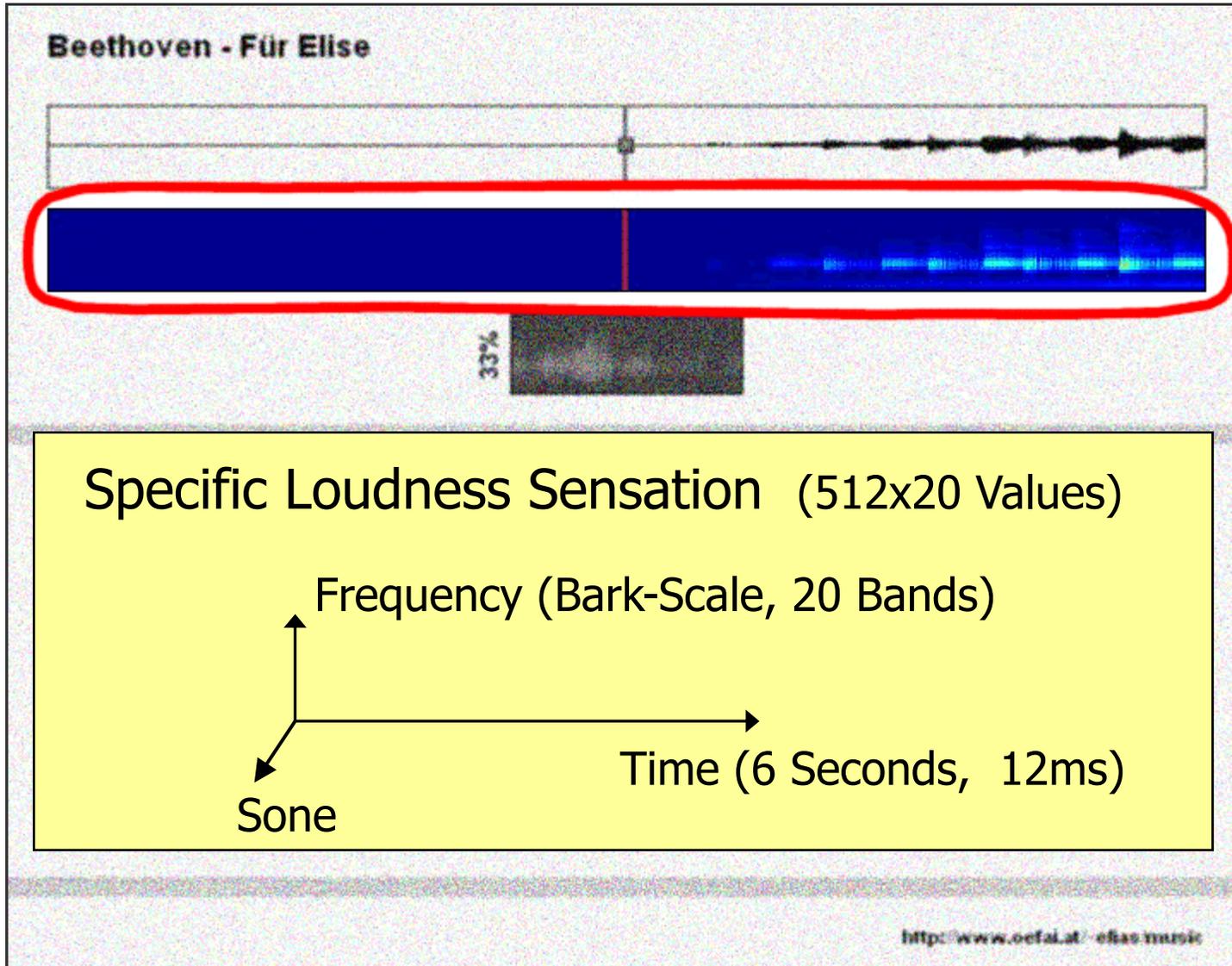
# Feature Extraction



.....



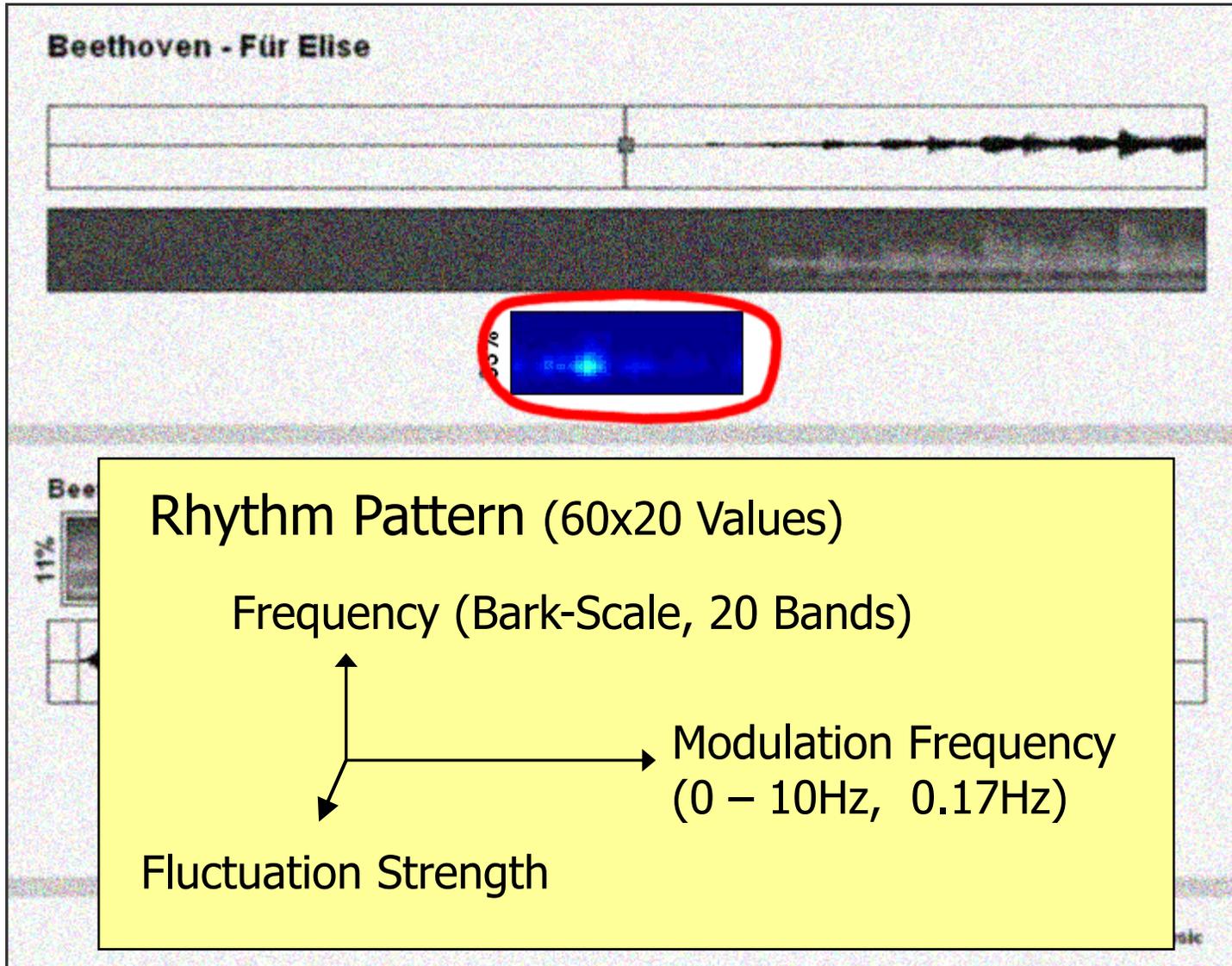
# Feature Extraction



.....

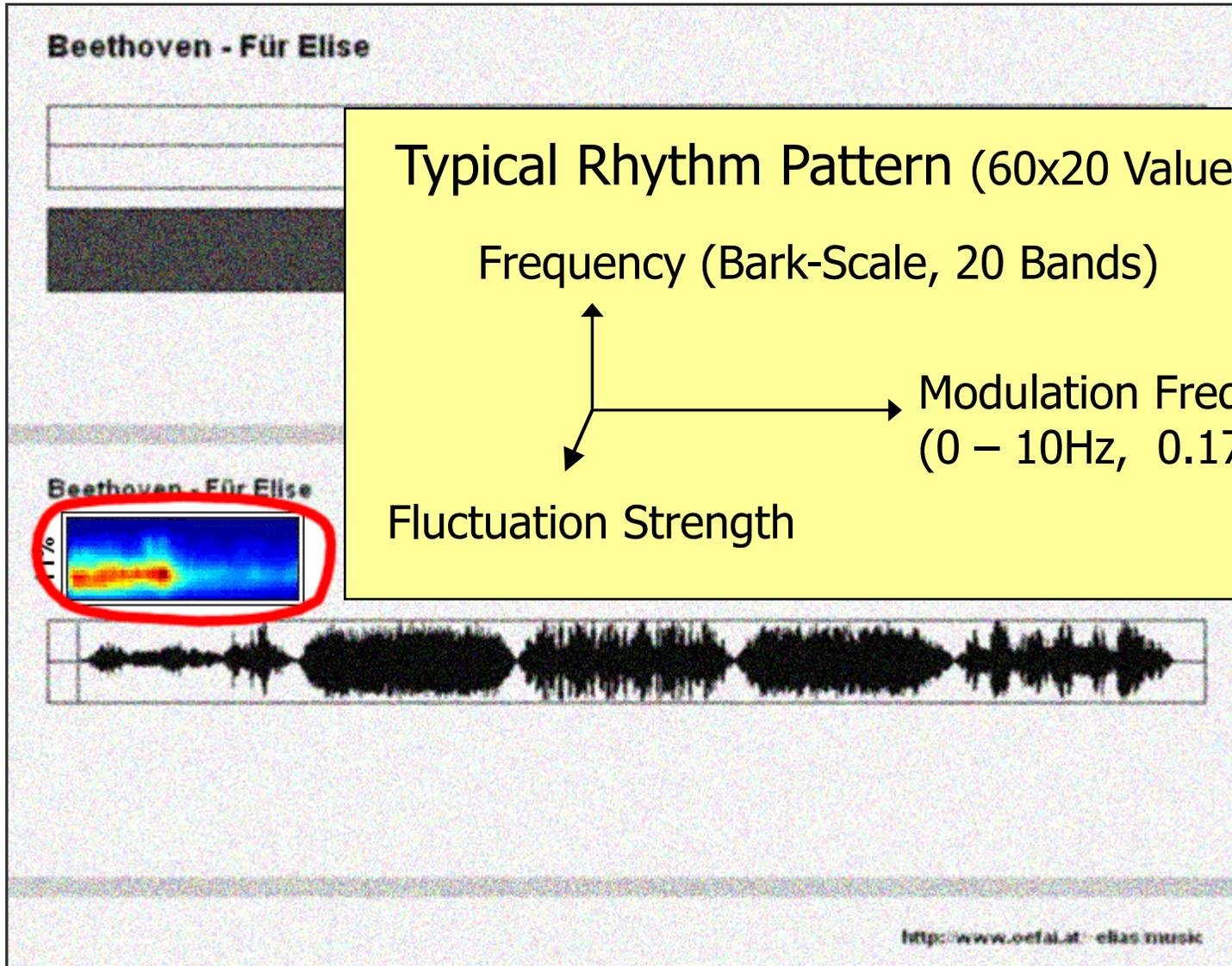


# Feature Extraction





# Feature Extraction





# Feature Extraction

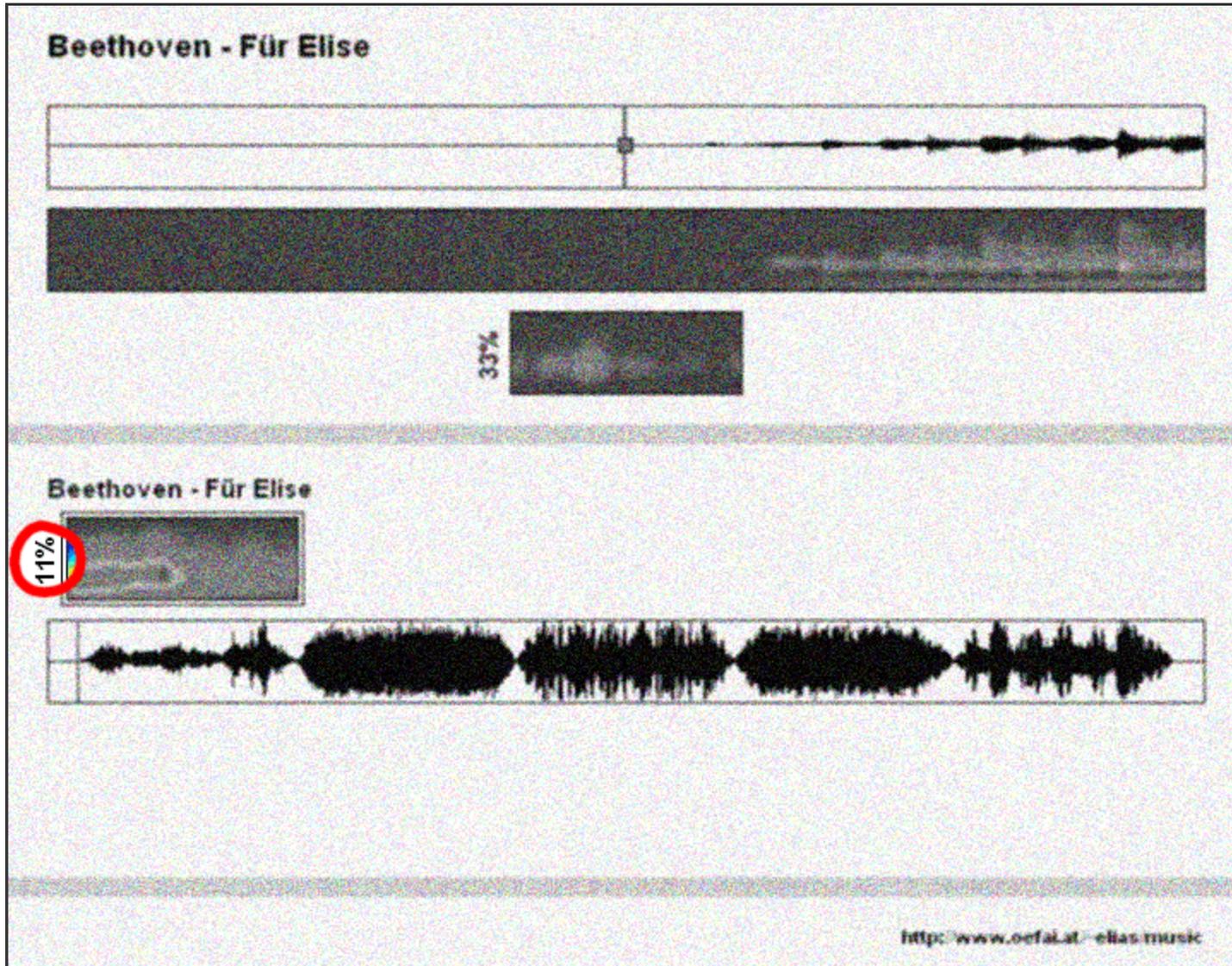


.....





# Feature Extraction



.....

## Rhythm Patterns

- Demo Video  
(prepared by Elias Pampalk)

