Advanced Course Computer Science

Music Processing

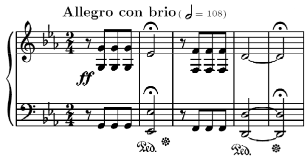
Summer Term 2010

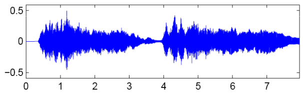
Meinard Müller

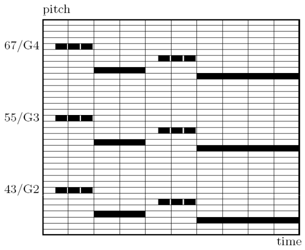
Saarland University and MPI Informatik
meinard@mpi-inf.mpg.de

Music Synchronization

Music Data







Music Data

Various interpretations – Beethoven‘s Fifth

Bernstein

Karajan

Scherbakov (piano)

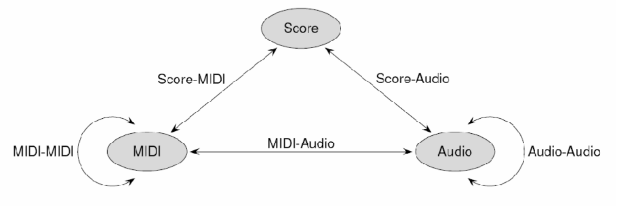
MIDI (piano)

General Goals

- Automated organization of complex and inhomogeneous music collections
- Generation of annotations and cross-links
- Tools and methods for multimodal search, navigation and interaction

Music Information Retrieval (MIR)

Music Synchronization



Schematic view of various synchronization tasks

Music Synchronization (Audio Alignment)

- Turetsky/Ellis (ISMIR 2003)
- Soulez/Rodet/Schwarz (ISMIR 2003)
- Arifi/Clausen/Kurth/Müller (ISMIR 2003)
- Hu/Dannenberg/Tzanetakis (WASPAA 2003)
- Müller/Kurth/Röder (ISMIR 2004)
- Raphael (ISMIR 2004)
- Dixon/Widmer (ISMIR 2005)
- Müller/Mattes/Kurth (ISMIR 2006)
- Dannenberg /Raphael (Special Issue ACM 2006)
- Kurth/Müller/Fremerey/Chang/Clausen (ISMIR 2007)
- Fujihara/Goto (ICASSP 2008)
- Wang/Iskandar/New/Shenoy (IEEE-TASLP 2008)
- Ewert/Müller/Grosche (ICASSP 2009)

Music Synchronization: Audio-Audio

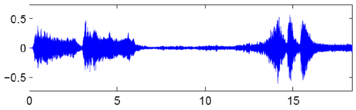
Given: Two different audio recordings of the same underlying piece of music.

Goal: Find for each position in one audio recording the musically corresponding position in the other audio recording.

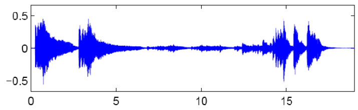
Music Synchronization: Audio-Audio

Beethoven‘s Fifth

Karajan ▶



Scherbakov ▶

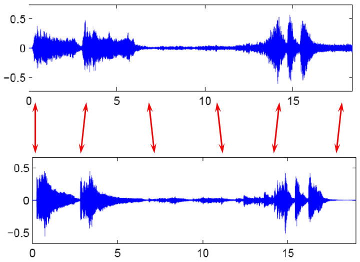


Music Synchronization: Audio-Audio

Beethoven‘s Fifth

Karajan ▶

Scherbakov ▶

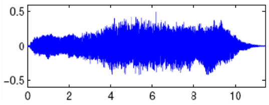


Synchronization: Karajan → Scherbakov ▶

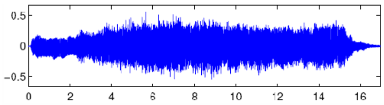
Music Synchronization: Audio-Audio

Bach Toccata

Koopman ▶



Ruebsam ▶

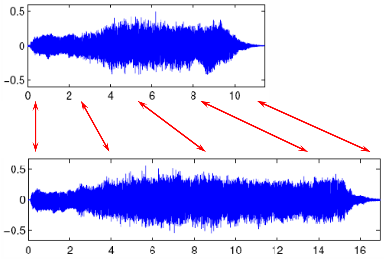


Music Synchronization: Audio-Audio

Bach Toccata

Koopman ▶

Ruebsam ▶



Synchronization: Koopman → Ruebsam ▶

Music Synchronization: Audio-Audio

- Transformation of audio recordings into sequences of feature vectors

  $\rightsquigarrow \; V := (v^1, v^2, \ldots, v^N)$

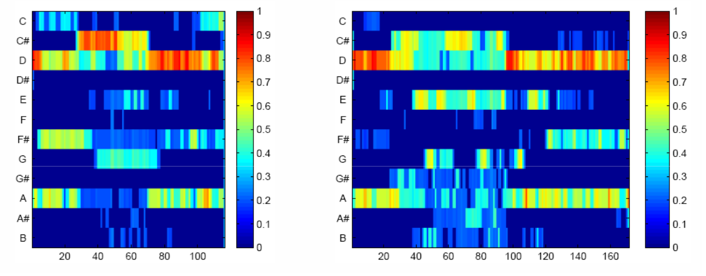
  $\rightsquigarrow \; W := (w^1, w^2, \ldots, w^M)$

- Fix cost measure c on the feature space
- Compute $N \times M$ cost matrix $C(n,m) := c(v^n, w^m)$
- Compute cost-minimizing warping path from C

Chroma Features

Example: Bach Toccata

Koopman ▶ ▶ Ruebsam ▶ ▶

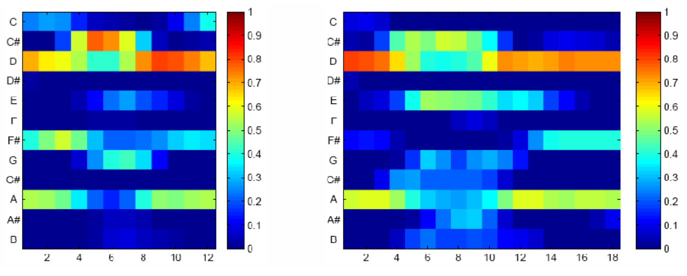


Feature resolution: 10 Hz

Chroma Features

Example: Bach Toccata

Koopman ▶ ▶ Ruebsam ▶ ▶



Feature resolution: 1 Hz

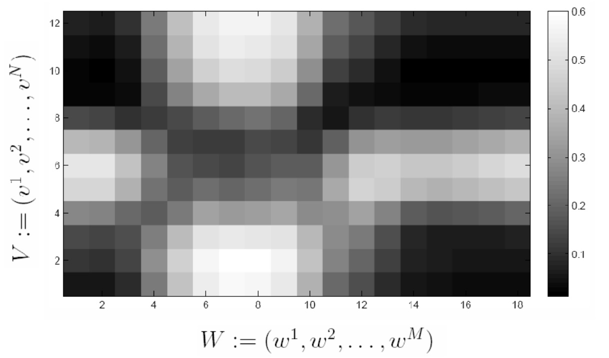
Music Synchronization: Audio-Audio

- Koopman ⇝ $V := (v^1, v^2, \ldots, v^N)$ $N = 12$

  Ruebsam ⇝ $W := (w^1, w^2, \ldots, w^M)$ $M = 18$
- v^n, w^m = 12-dimensional normalized chroma vectors
- Local cost measure $c : \mathbb{R}^{12} \times \mathbb{R}^{12} \to \mathbb{R}$

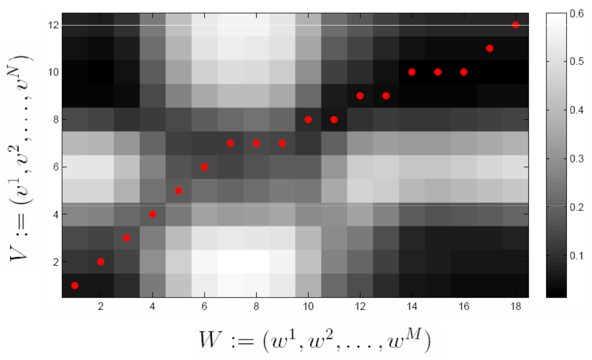
  $$c(v^n, w^m) := 1 - \langle v^n, w^m \rangle$$
- $N \times M$ cost matrix $C(n,m) := c(v^n, w^m)$

Music Synchronization: Audio-Audio



Music Synchronization: Audio-Audio

Cost-minimizing warping path



Cost-Minimizing Warping Path

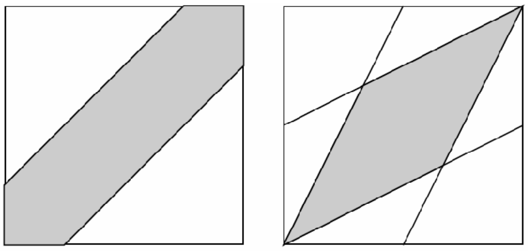
- Computation via dynamic programming

  ⇝ Dynamic Time Warping (DTW)
- Memory requirements and running time: *O*(*NM*)
- Problem: Infeasible for large *N* and *M*
- Example: Feature resolution 10 Hz, pieces 15 min

  ⇒ *N, M* ~ 10,000

  ⇒ *N · M* ~ *100,000,000*

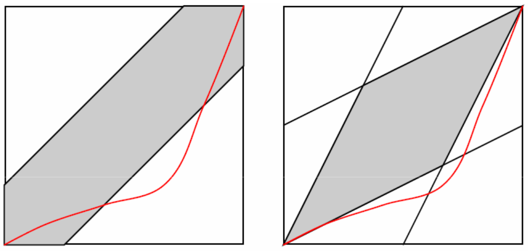
Strategy: Global Constraints



Sakoe-Chiba band

Itakura parallelogram

Strategy: Global Constraints

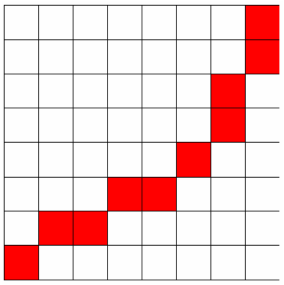


Sakoe-Chiba band

Itakura parallelogram

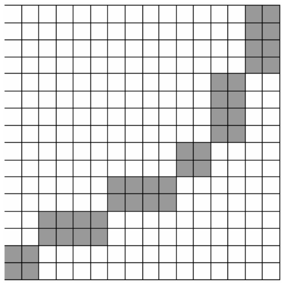
Problem: Optimal warping path not in constraint region

Strategy: Multiscale Approach



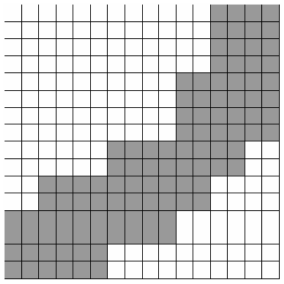
Compute optimal warping path on coarse level

Strategy: Multiscale Approach



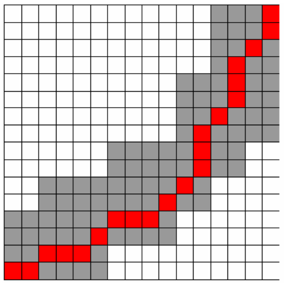
Project on fine level

Strategy: Multiscale Approach



Specify constraint region

Strategy: Multiscale Approach



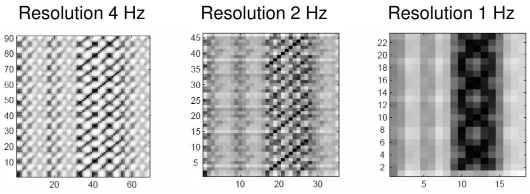
Compute *constrained* optimal warping path

Strategy: Multiscale Approach

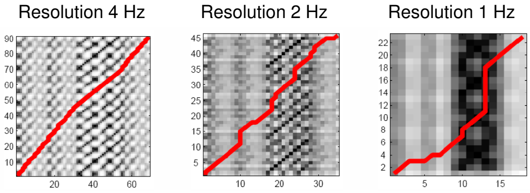
- Suitable features?
- Suitable resolution levels?
- Size of constraint regions?

Good trade-off between efficiency and robustness?

Strategy: Multiscale Approach



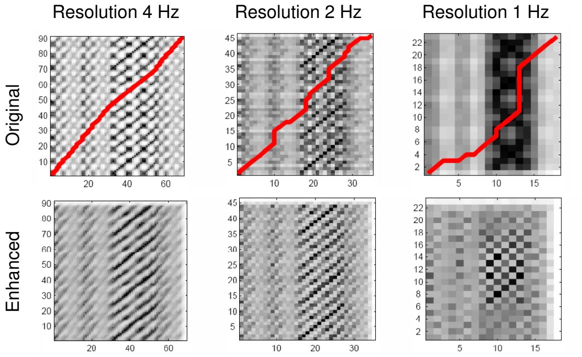
Strategy: Multiscale Approach



Problem: Cost matrix may degenerate
⇝ useless warping path

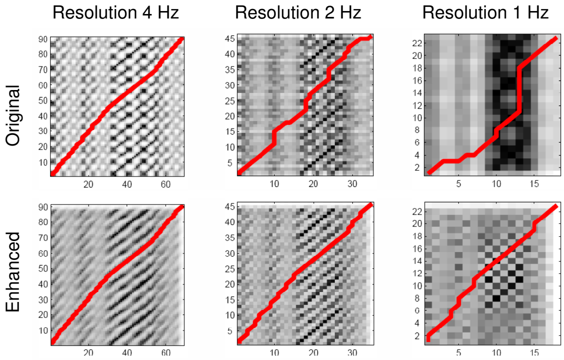
Strategy: Multiscale Approach

Improve robustness by enhancing cost matrix



Strategy: Multiscale Approach

Improve robustness by enhancing cost matrix



Strategy: Multiscale Approach

Chroma features at three levels: 0.33 Hz / 1 Hz / 10 Hz

Recording 1	length [sec]	Recording 2	length [sec]	t_{DTW} [sec]	t_{MsDTW} [sec]	[%]
Beet9Bern	1144.9	Beet9Kar	1054.8	31.18	1.08	3.46

Strategy: Multiscale Approach

Chroma features at three levels: 0.33 Hz / 1 Hz / 10 Hz

Recording 1	length [sec]	Recording 2	length [sec]	t_{DTW} [sec]	t_{MsDTW} [sec]	[%]
Beet9Bern	1144.9	Beet9Kar	1054.8	31.18	1.08	3.46

Number of matrix entries needed for DTW and MsDTW:

	DTW	MsDTW	%
Level 1	120,808,050	2,117,929	1.75
Level 2	1,209,030	17,657	1.46
Level 3	134,464	134,464	100

Music Synchronization: Audio-Audio

Conclusions

- Chroma features

  ⇝ suited for harmony-based music

- Relatively coarse but good global alignments

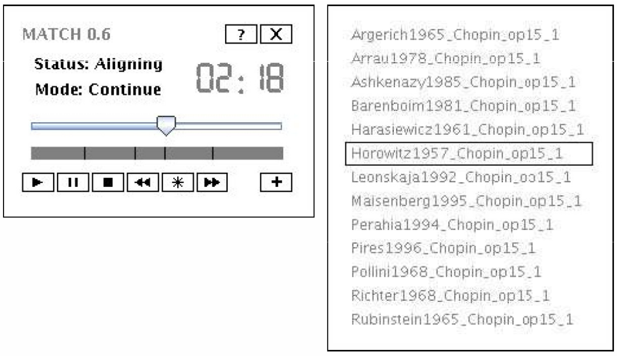
- Multiscale approach: simple, robust, fast

Music Synchronization: Audio-Audio

Applications

- Efficient music browsing
- Blending from one interpretation to another one
- Mixing and morphing different interpretations
- Tempo studies

System: Match (Dixon)



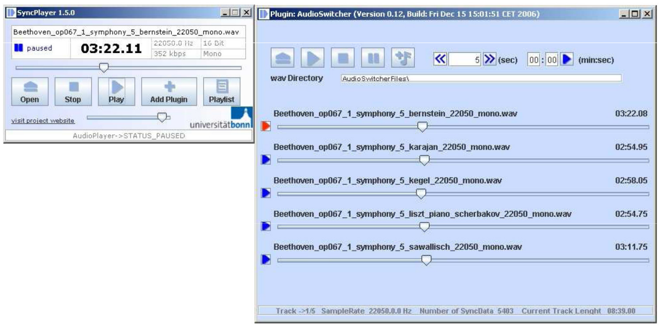
MATCH 0.6

Status: Aligning
Mode: Continue

02:18

Argerich1965_Chopin_op15_1
Arrau1978_Chopin_op15_1
Ashkenazy1985_Chopin_op15_1
Barenboim1981_Chopin_op15_1
Harasiewicz1961_Chopin_op15_1
Horowitz1957_Chopin_op15_1
Leonskaja1992_Chopin_op15_1
Maisenberg1995_Chopin_op15_1
Perahia1994_Chopin_op15_1
Pires1996_Chopin_op15_1
Pollini1968_Chopin_op15_1
Richter1968_Chopin_op15_1
Rubinstein1965_Chopin_op15_1

System: SyncPlayer/AudioSwitcher



SyncPlayer 1.5.0

Beethoven_op067_1_symphony_5_bernstein_22050_mono.wav

paused 03:22.11 22050.0 Hz 16 Bit 352 kbps Mono

visit project website

universitätbonn

AudioPlayer->STATUS_PAUSED

Plugin: AudioSwitcher (Version 0.12, Build: Fri Dec 15 15:01:51 CET 2006)

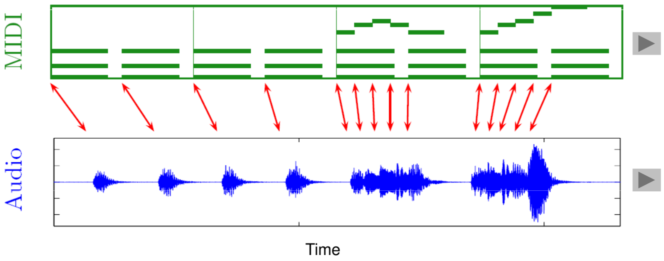
5 (sec) 00 : 00 (min:sec)

wav Directory AudioSwitcherFiles\

Beethoven_op067_1_symphony_5_bernstein_22050_mono.wav	03:22.08
Beethoven_op067_1_symphony_5_karajan_22050_mono.wav	02:54.95
Beethoven_op067_1_symphony_5_kegel_22050_mono.wav	02:58.05
Beethoven_op067_1_symphony_5_liszt_piano_scherbakov_22050_mono.wav	02:54.75
Beethoven_op067_1_symphony_5_sawallisch_22050_mono.wav	03:11.75

Track ->1/5 SampleRate 22050.0.0 Hz Number of SyncData 5403 Current Track Lenght 08:39.00

Music Synchronization: MIDI-Audio



Music Synchronization: MIDI-Audio

MIDI = meta data

Automated annotation

Audio recording

Sonification of annotations ▶ ▶

Music Synchronization: MIDI-Audio

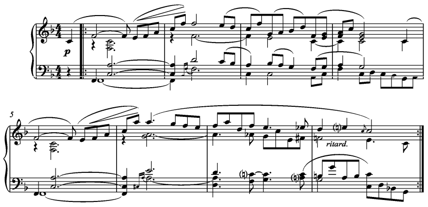
MIDI = reference (score)

Tempo information

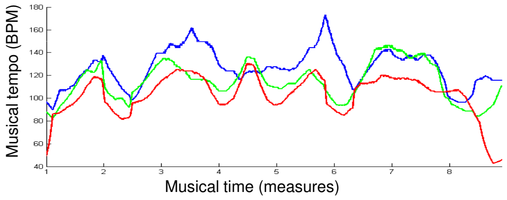
Audio recording

Performance Analysis: Tempo Curves

Schumann: Träumerei



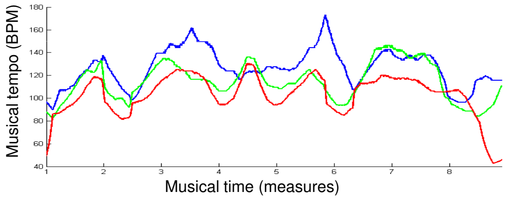
▶
▶
▶



Performance Analysis: Tempo Curves

What can be done if no reference is available?

▶
▶
▶

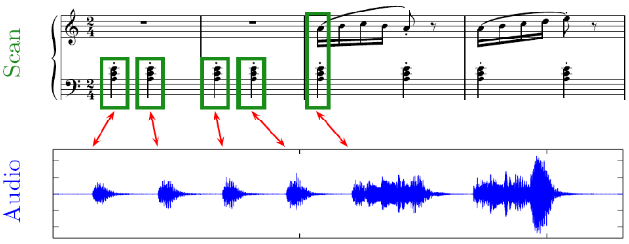


Music Synchronization: MIDI-Audio

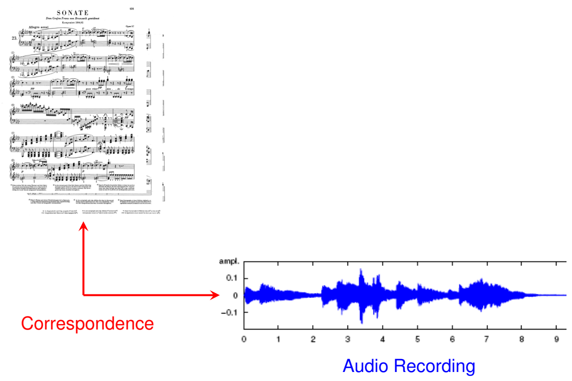
Applications

- Automated audio annotation
- Accurate audio access after MIDI-based retrieval
- Automated tracking of MIDI note parameters during audio playback
- Performance Analysis

Music Synchronization: Scan-Audio



Music Synchronization: Scan-Audio

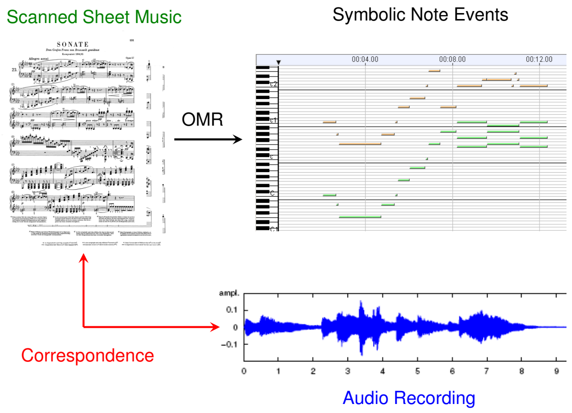


Scanned Sheet Music

Correspondence

Audio Recording

Music Synchronization: Scan-Audio



Scanned Sheet Music

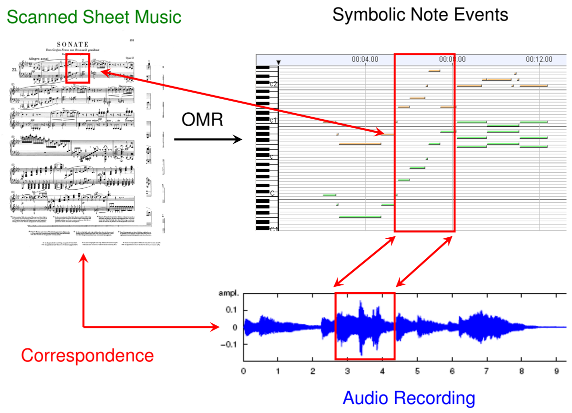
Symbolic Note Events

OMR

Correspondence

Audio Recording

Music Synchronization: Scan-Audio



Scanned Sheet Music

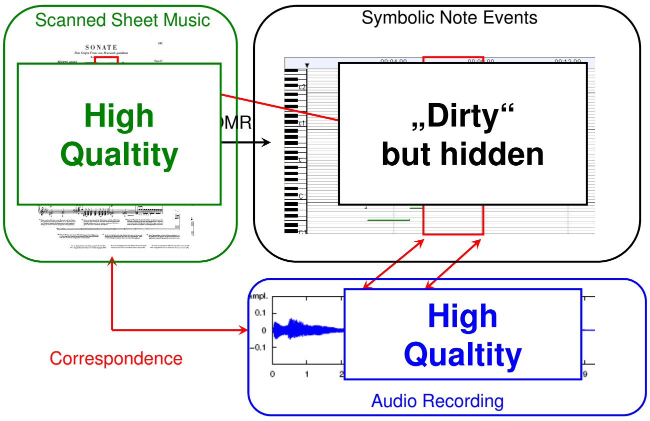
Symbolic Note Events

OMR

Correspondence

Audio Recording

Music Synchronization: Scan-Audio



Scanned Sheet Music

Symbolic Note Events

High Qualtity

OMR

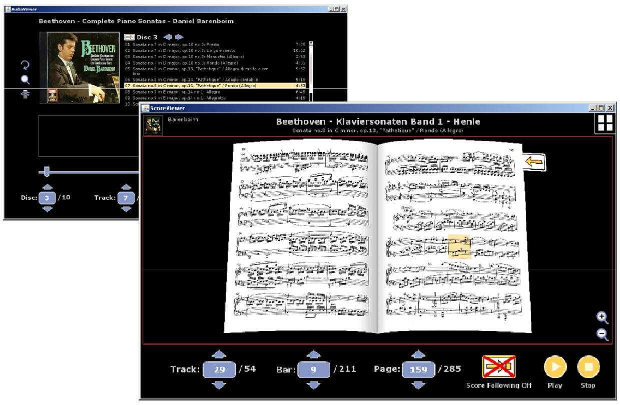
„Dirty“ but hidden

Correspondence

High Qualtity

Audio Recording

Application: Score Viewer



AudioViewer

Beethoven - Complete Piano Sonatas - Daniel Barenboim

Disc 3

01 Sonata no.7 in D major, op.10 no.3: Presto 7:08
02 Sonata no.7 in D major, op.10 no.3: Largo e mesto 10:02
03 Sonata no.7 in D major, op.10 no.3: Menuetto (Allegro) 2:53
04 Sonata no.7 in D major, op.10 no.3: Rondo (Allegro) 4:05
05 Sonata no.8 in C minor, op.13, "Pathetique" / Allegro di molto e con brio 9:32
06 Sonata no.8 in C minor, op.13, "Pathetique" / Adagio cantabile 5:19
07 Sonata no.8 in C minor, op.13, "Pathetique" / Rondo (Allegro) 4:53
08 Sonata no.9 in E major, op.14 no.1: Allegro 6:48
09 Sonata no.9 in E major, op.14 no.1: Allegretto 4:16

Disc: 3 /10 Track: 7

ScoreViewer

Barenboim

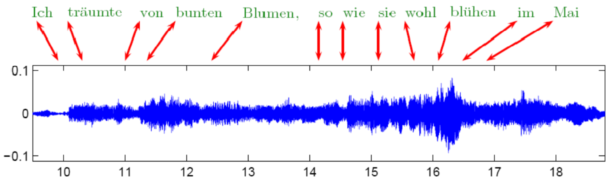
Beethoven - Klaviersonaten Band 1 - Henle

Sonata no.8 in C minor, op.13, "Pathetique" / Rondo (Allegro)

Track: 29 /54 Bar: 9 /211 Page: 159 /285

[ECDL 08, ICMI 08]

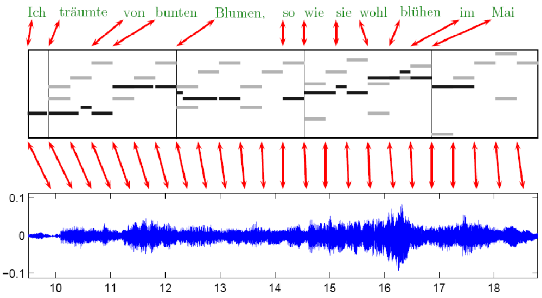
Music Synchronization: Lyrics-Audio



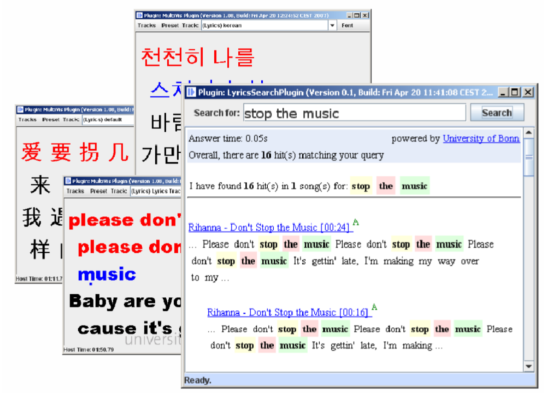
Difficult task!

Music Synchronization: Lyrics-Audio

Lyrics-Audio → Lyrics-MIDI + MIDI-Audio



System: SyncPlayer/LyricsSeeker


Plugin: MultiVis Plugin (Version 1.08, Build: Fri Apr 20 12:24:52 CEST 2007)
Tracks Preset Track: (Lyrics) korean
Font
천천히 나를
바
가만
Plugin: MultiVis Plugin (Version 1.08, Build:
Tracks Preset Track: (Lyrics) default
爱 要 拐 几
来
我 进
样
Host Time: 01:11.7
Plugin: MultiVis Plugin (Version 1.08, Build:
Tracks Preset Track: (Lyrics) Lyrics Trac
please don'
please don
music
Baby are yo
cause it's
Host Time: 01:50.79
Plugin: LyricsSearchPlugin (Version 0.1, Build: Fri Apr 20 11:41:08 CEST 2...
Search for:
Search
Ready.

Search for: stop the music

Answer time: 0.05s powered by University of Bonn

Overall, there are **16** hit(s) matching your query

I have found **16** hit(s) in **1** song(s) for: **stop** **the** **music**

Rihanna - Don't Stop the Music [00:24] A

... Please don't **stop** **the** **music** Please don't **stop** **the** **music** Please don't **stop** **the** **music** It's gettin' late, I'm making my way over to my ...

Rihanna - Don't Stop the Music [00:16] A

... Please don't **stop** **the** **music** Please don't **stop** **the** **music** Please don't **stop** **the** **music** It's gettin' late, I'm making ...

High-Resolution Music Synchronization

- Normalized chroma features
  → robust to changes in instrumentation and dynamics
  → robust synchronization of reasonable overall quality
- Drawback: low temporal alignment accuracy
- Idea: Integration of note onset information

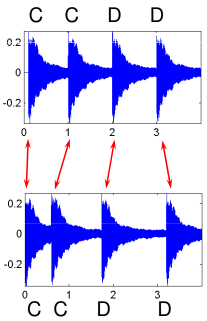
High-Resolution Music Synchronization

- Normalized chroma features
  → robust to changes in instrumentation and dynamics
  → robust synchronization of reasonable overall quality
- Drawback: low temporal alignment accuracy
- Idea: Integration of note onset information
- Example: MIDI-Audio synchronization

  Chroma-Chroma:
  Chroma-Chroma + onset information:

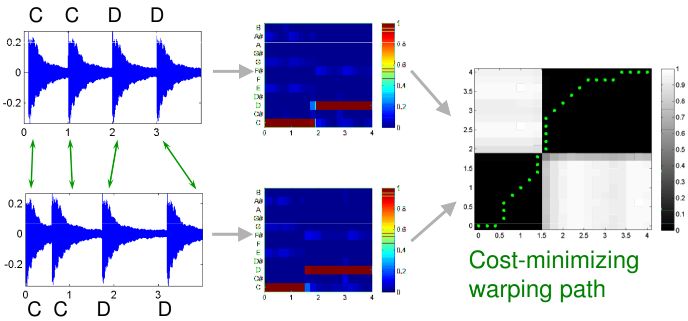
High-Resolution Music Synchronization

Example: C – C – D – D



High-Resolution Music Synchronization

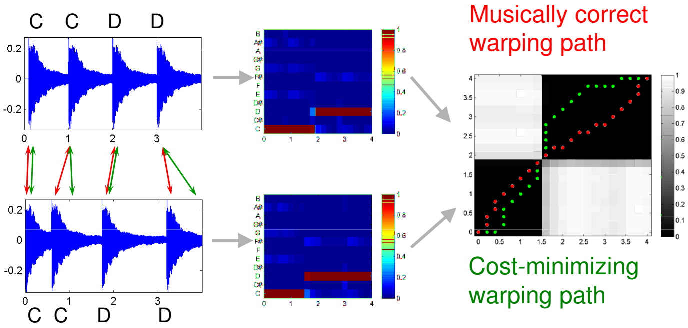
Example: C – C – D – D



Cost-minimizing warping path

High-Resolution Music Synchronization

Example: C – C – D – D

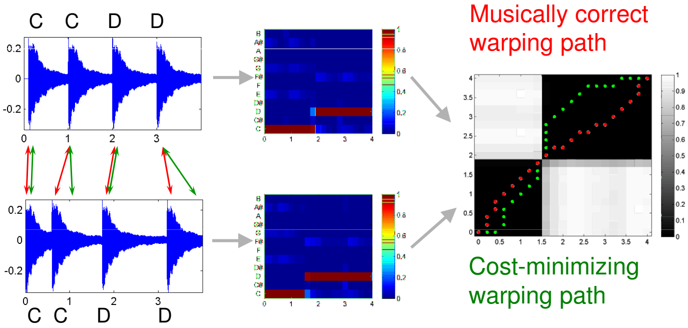


Musically correct warping path

Cost-minimizing warping path

High-Resolution Music Synchronization

Example: C – C – D – D



Musically correct warping path

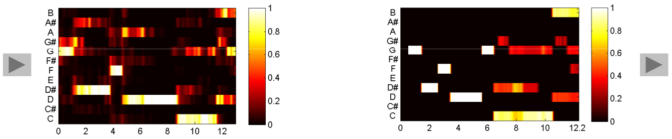
Cost-minimizing warping path

Problem: note onsets are not captured in feature representation

High-Resolution Music Synchronization

Example: Beethoven's Fifth

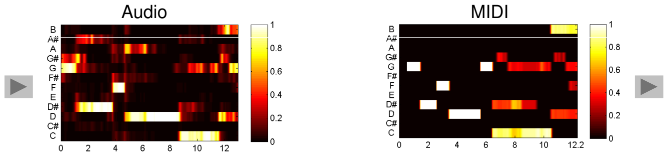
Chroma representations



Problem: note onsets are not captured in feature representation

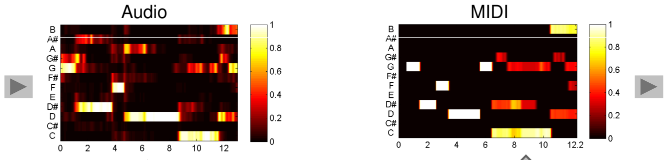
High-Resolution Music Synchronization

Example: Beethoven's Fifth

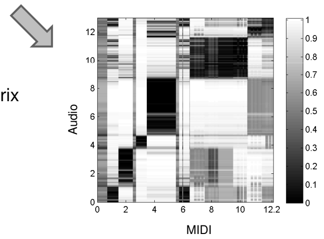


High-Resolution Music Synchronization

Example: Beethoven's Fifth

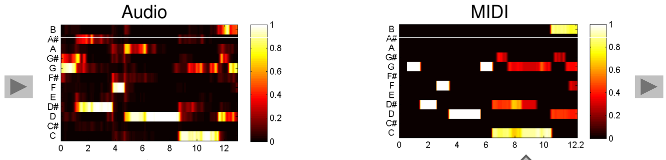


Cost matrix

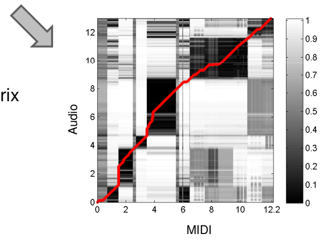


High-Resolution Music Synchronization

Example: Beethoven's Fifth



Cost matrix

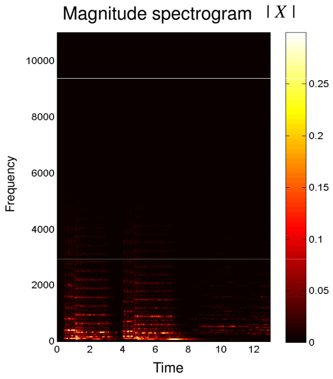


Warping path of poor local quality

Onset Detection

- General goal: Detection of onsets of musical notes
- Typical signal properties at note onset positions:
  - increase in energy
  - change of pitch
  - change of spectral content
  - high frequency content
- Idea: locate note onset candidates by measuring changes in spectral content

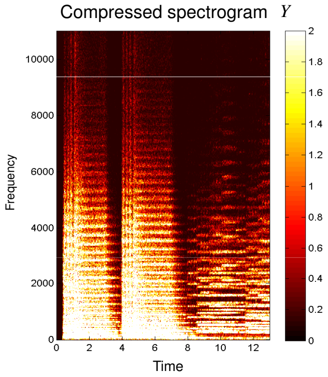
Onset Detection



Steps:

1. Spectrogram

Onset Detection



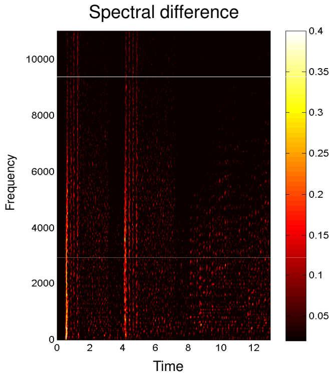
Steps:

1. Spectrogram
2. Logarithmic compression

$$Y = \log(1 + C \cdot | X |)$$

- *human sensation*
- *enhances low intensity values*
- *high frequency content*
- *reduces influence of amplitude modulation*

Onset Detection

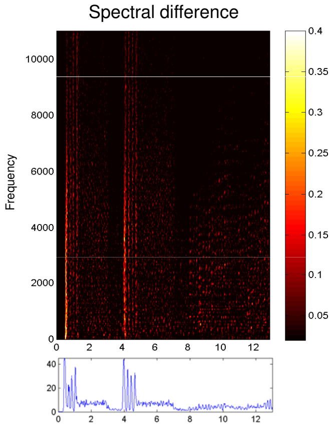


Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation

- *energy increase to be captured*
- *only positive values considered*

Onset Detection



Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation

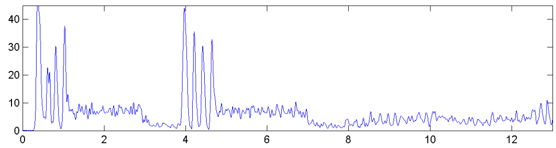
Novelty Curve

Onset Detection

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation

Novelty Curve



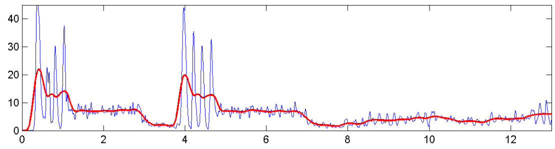
Onset Detection

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization

Novelty Curve

Substraction of local average

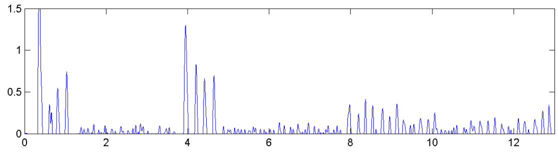


Onset Detection

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization

Normalized novelty curve

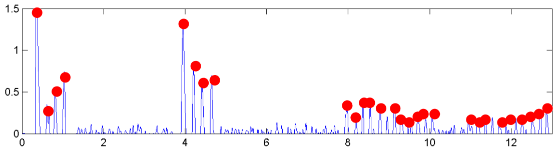


Onset Detection

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization
6. Peak picking

Normalized novelty curve

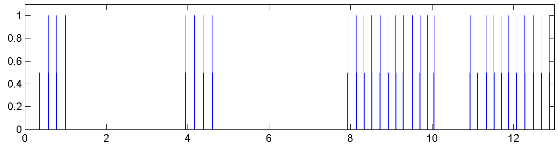


Onset Detection

Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization
6. Peak picking

Impulses

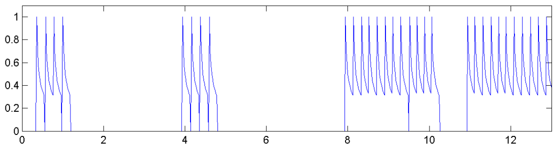


Onset Detection

Steps:

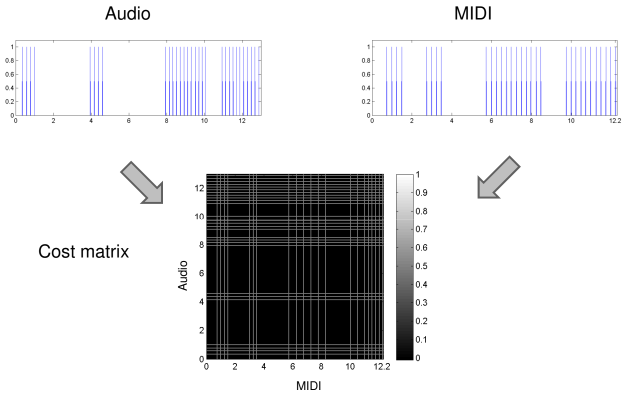
1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization
6. Peak picking
7. Decay Filter

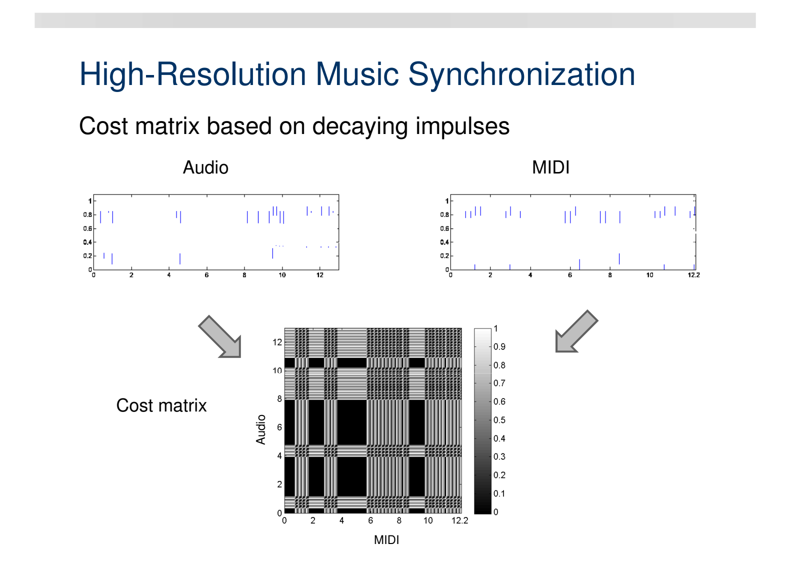
Decaying impulses



High-Resolution Music Synchronization

Cost matrix based on impulses





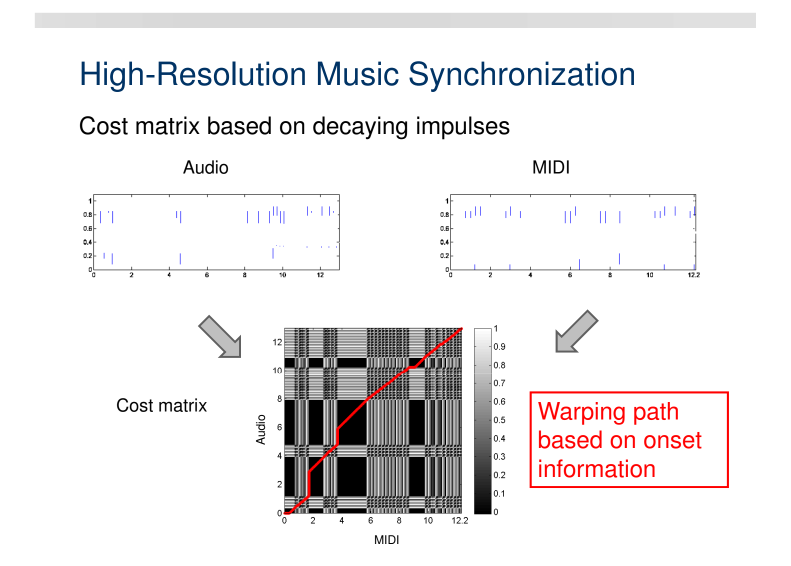
High-Resolution Music Synchronization

Cost matrix based on decaying impulses

Audio

MIDI

Cost matrix



High-Resolution Music Synchronization

Cost matrix based on decaying impulses

Audio

MIDI

Cost matrix

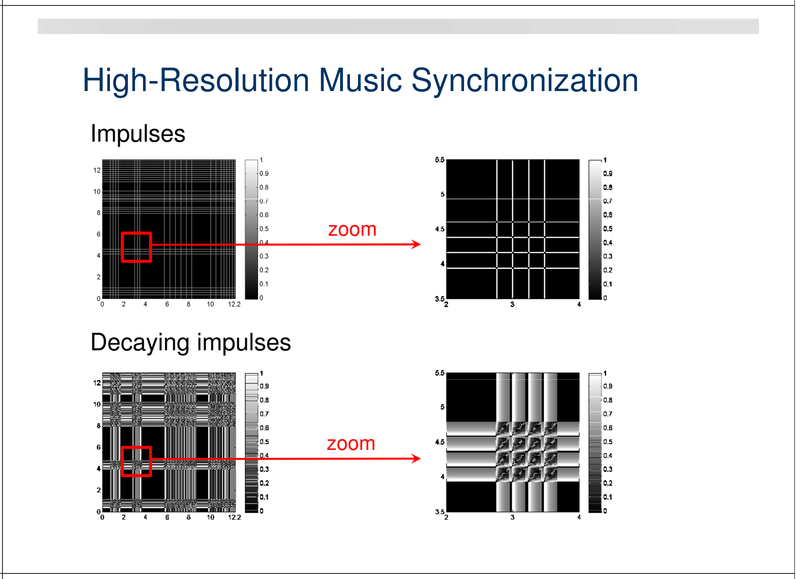
Warping path based on onset information

High-Resolution Music Synchronization

Ideas:

- Build up cost matrix with corridors of low cost
- Decaying strategy enforce corridor structure
- Each corridor corresponds to MIDI-audio pair of note onset candidates
- Warping path tends to run through corridors of low cost

→ note onset positions are likely to be aligned



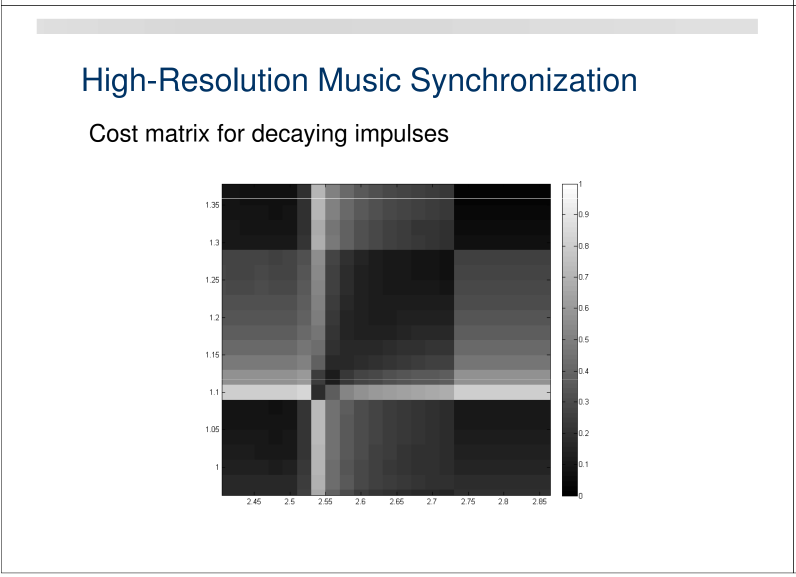
High-Resolution Music Synchronization

Impulses

zoom

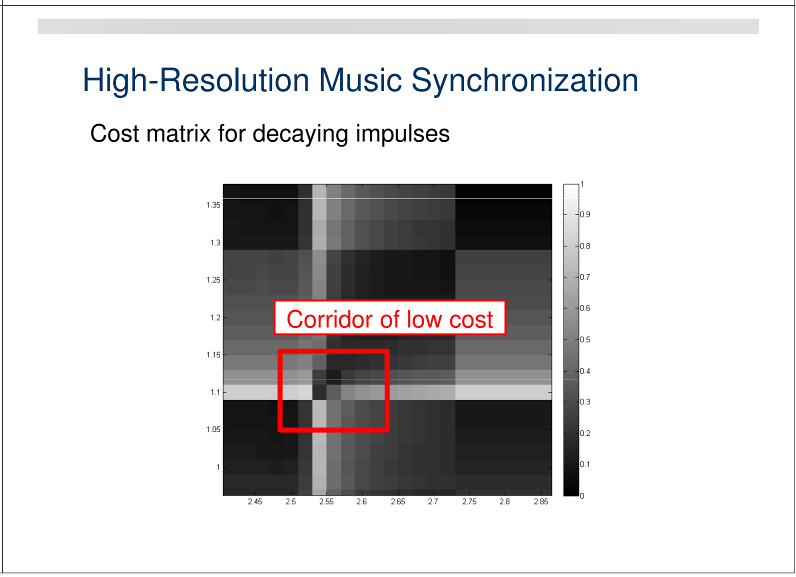
Decaying impulses

zoom



High-Resolution Music Synchronization

Cost matrix for decaying impulses



High-Resolution Music Synchronization

Cost matrix for decaying impulses

Corridor of low cost

High-Resolution Music Synchronization

Combination of two different types of cost matrices:

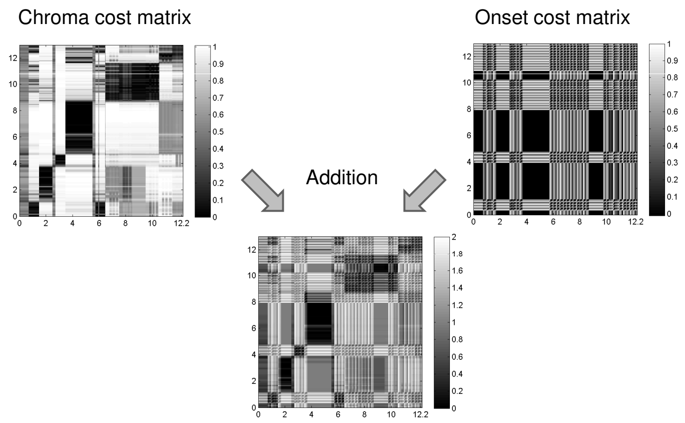
- Cost matrix obtained from chroma features controls the global course of warping path

  → robust synchronization

- Cost matrix obtained from onset information controls the local course of warping path

  → accurate alignment

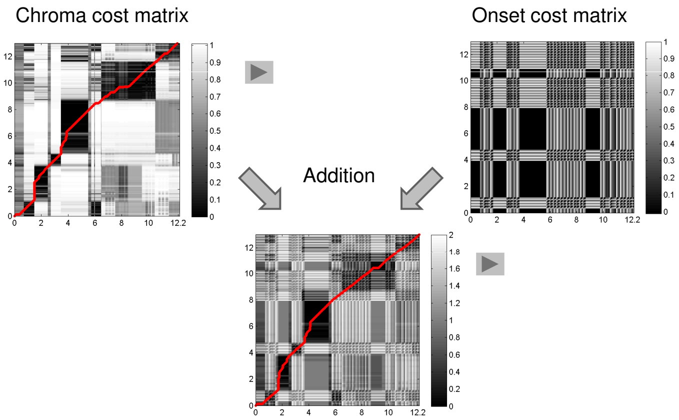
High-Resolution Music Synchronization


Chroma cost matrix

Onset cost matrix

Addition

High-Resolution Music Synchronization


Chroma cost matrix

Onset cost matrix

Addition

Conclusions: Music Synchronization

Various requirements

- Efficiency
- Robustness
- Accuracy
- Variablity of music

Conclusions: Music Synchronization

Combination of various strategies

- Feature level
- Local cost measure level
- Global alignment level
- Evidence pooling using competing strategies

Conclusions: Music Synchronization

Offline vs. Online

- Online version: Dixon/Widmer (ISMIR 2005)
- Hidden Markov Models: Raphael (ISMIR 2004)
- Score-following
- Automatic accompaniment

Conclusions: Music Synchronization

Presence of variations

- Instrumentation
- Musical structure
- Polyphony
- Musical key
- …