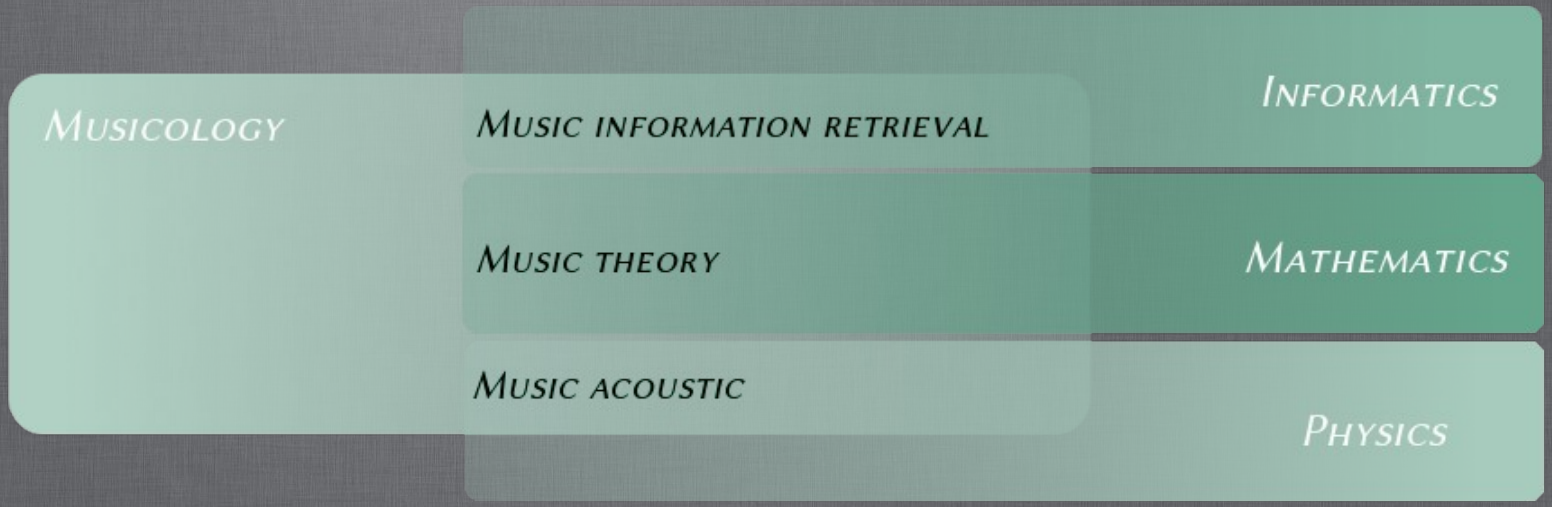
Music Information Retrieval
State-of-the-art techniques

Ladislav Maršík
Charles University, Prague

Music Information Retrieval (MIR)



Applications

Outline

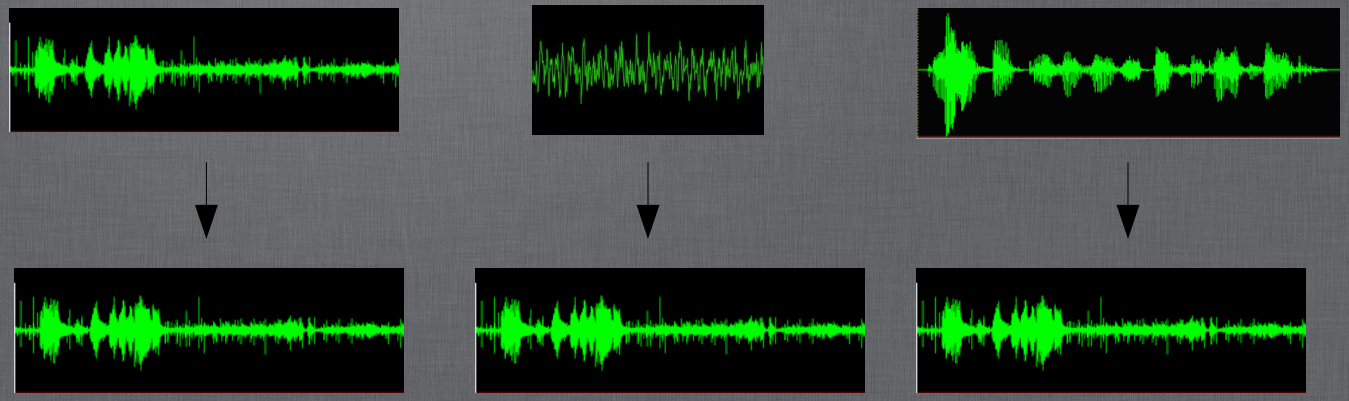
MIR problems (focus: audio query)
with state-of-the-art techniques

Categorization of techniques

MIR problems (audio query)

1. Audio Fingerprinting
2. Whistling and Humming Queries
3. Cover Song Identification
4. Audio similarity (related: music recommendation)

1. 2. 3. and 4.



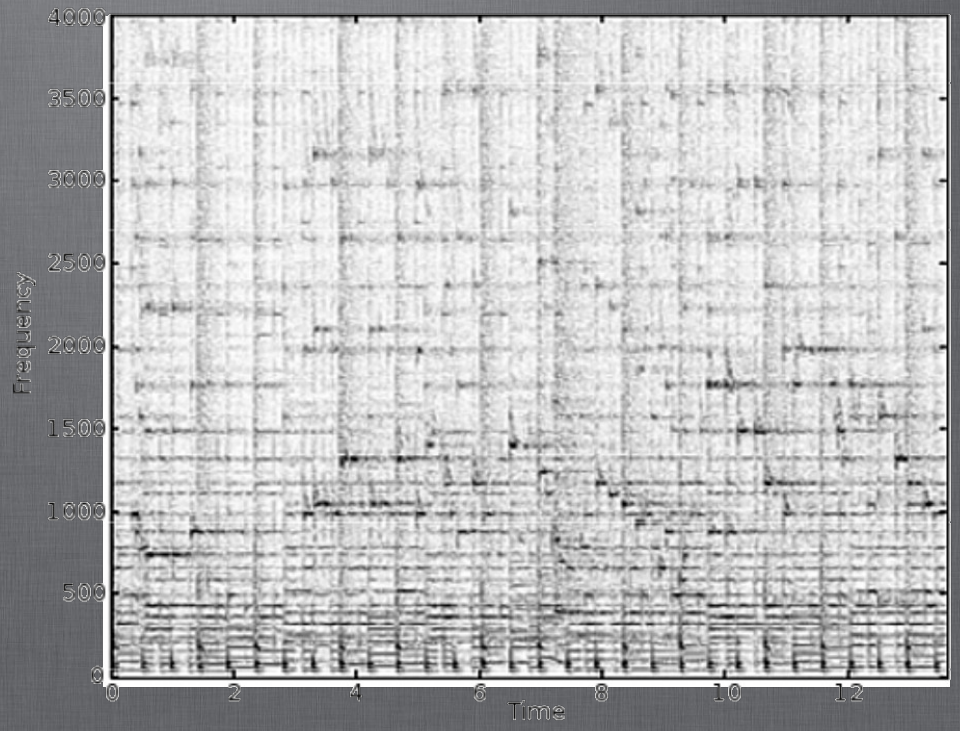
1. Audio Fingerprinting

INPUT: Song recording

OUTPUT: The exact match

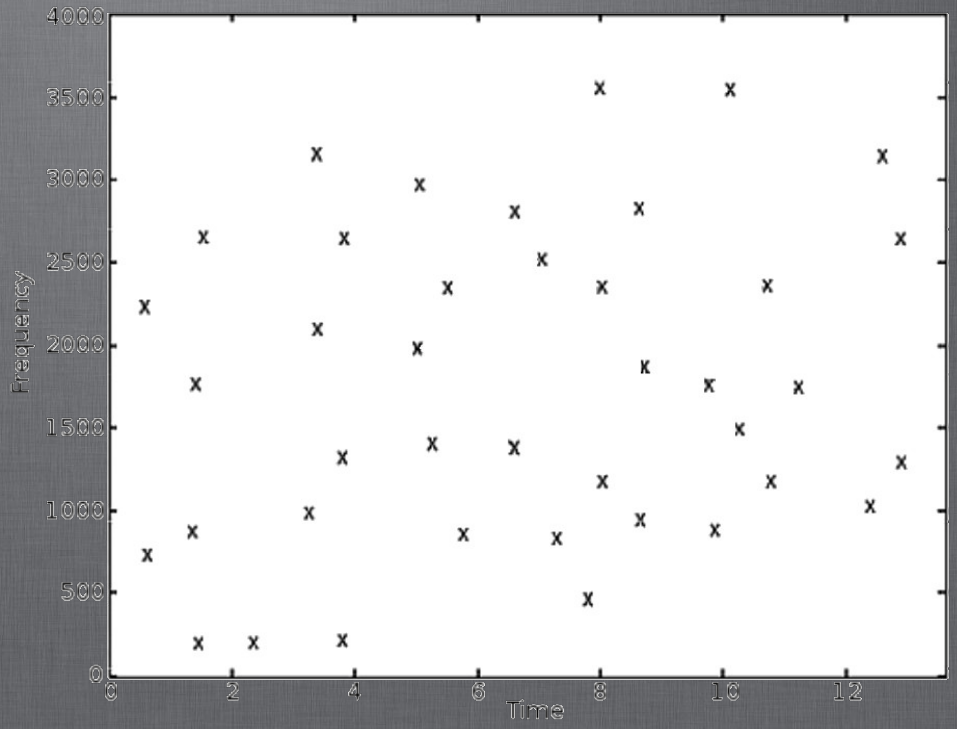
1. Audio Fingerprinting
Wang and Smith: An Industrial-Strength Audio Search Algorithm (2002)

Time-Frequency spectrogram



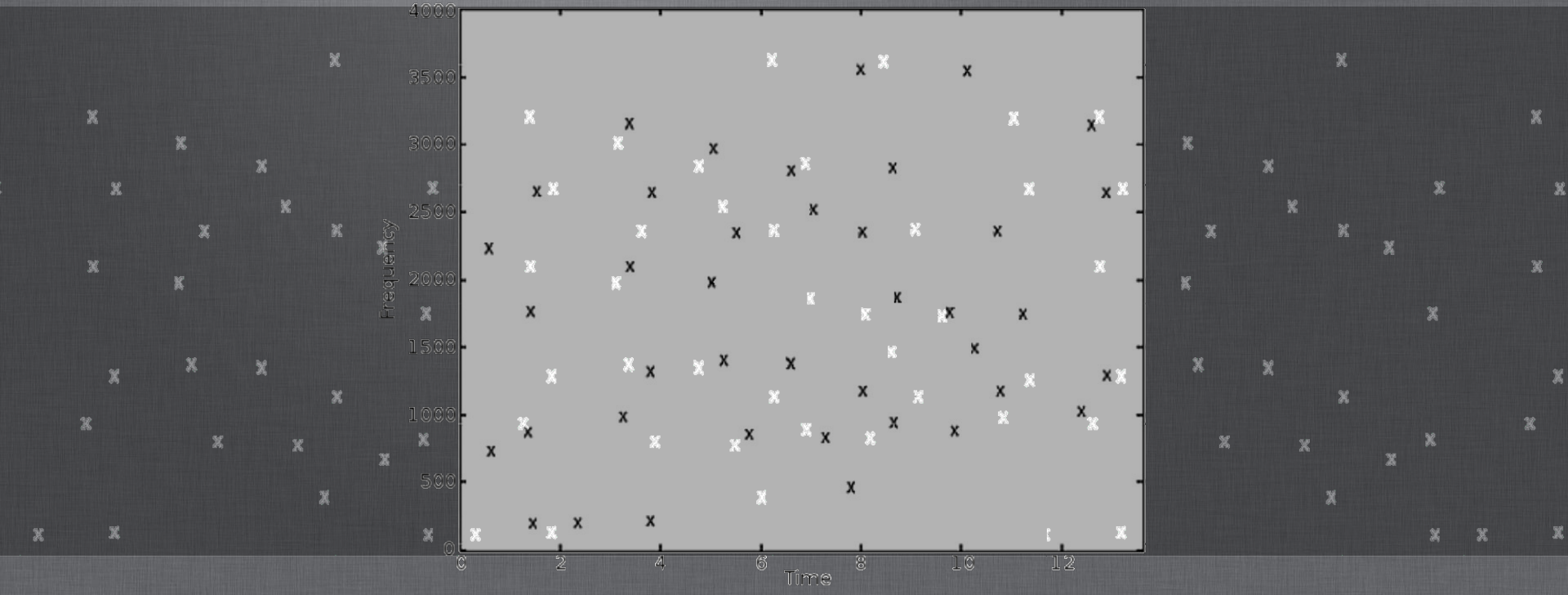
1. Audio Fingerprinting
Wang and Smith: An Industrial-Strength Audio Search Algorithm (2002)

Constellation analysis



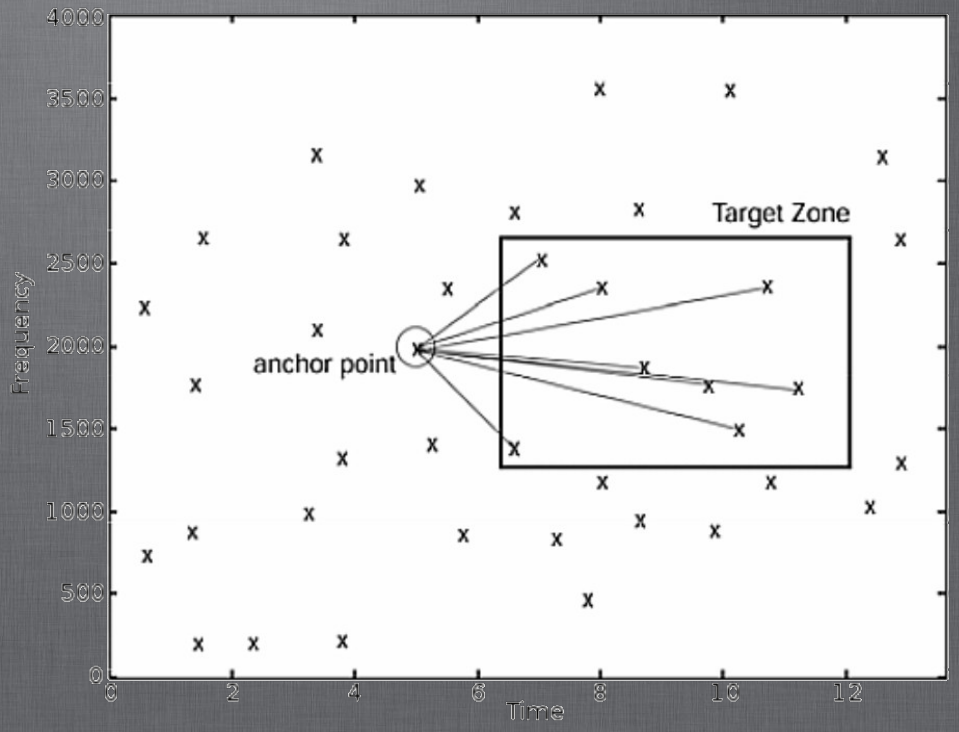
1. Audio Fingerprinting
Wang and Smith: An Industrial-Strength Audio Search Algorithm (2002)

Constellation analysis



1. Audio Fingerprinting
Wang and Smith: An Industrial-Strength Audio Search Algorithm (2002)

Combinatorially hashed $h(f_1,f_2,t_2\text{-}t_1) \mid t_1$



Summary

- Short search time: 5-500 milliseconds / query
- Robust to noisy environment

State-of-the-art

- Various indexing techniques
- Benchmarking: MIREX 2015
- Focus on commercial deployment, advertisment

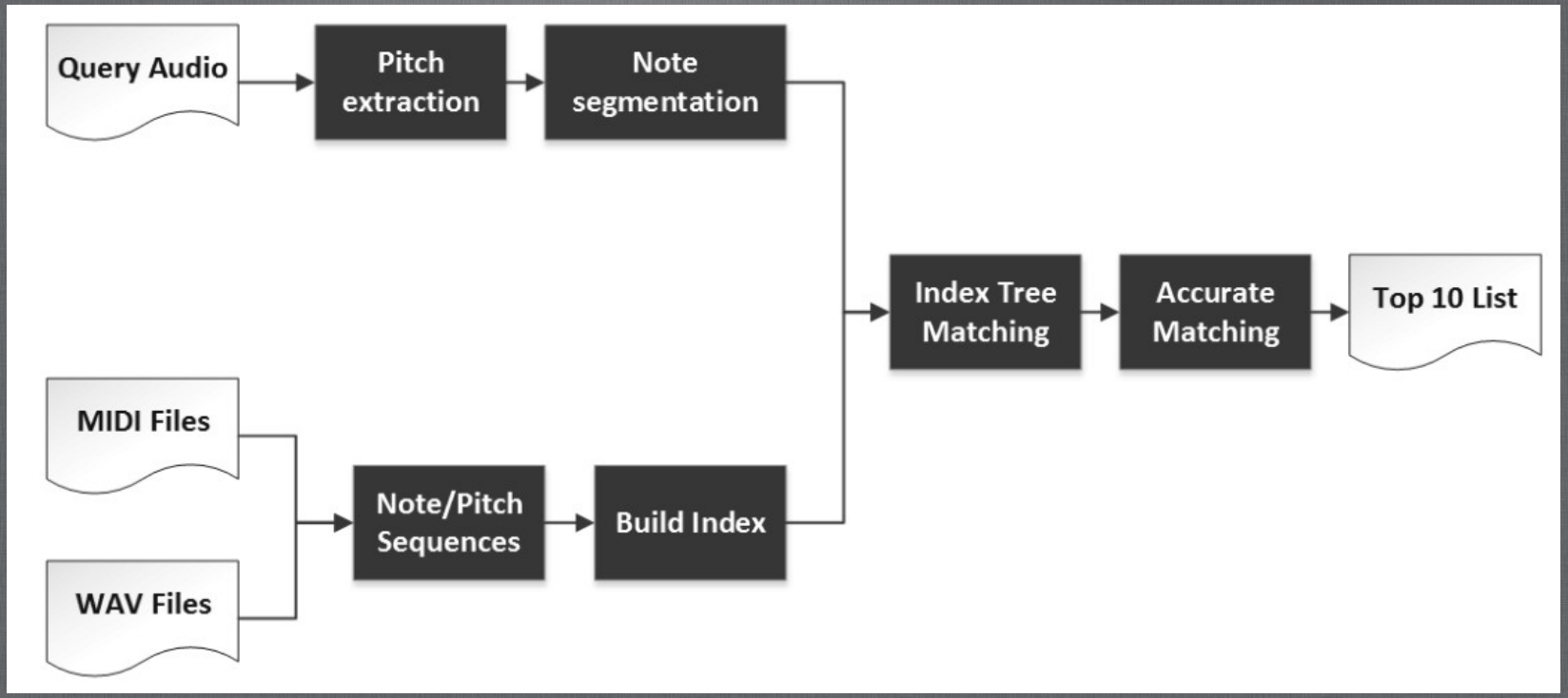
2. Whistling and Humming Queries

INPUT: Whistling or Humming

OUTPUT: Song containing the melody

2. Whistling and Humming Queries
Shen and Lee: Whistle for Music (2007)

- Whistle: 700Hz-2.8KHz
- Translation to MIDI (Query and DB)
- String matching methods



Summary

- Fast & Effective
- False positives

State-of-the-art

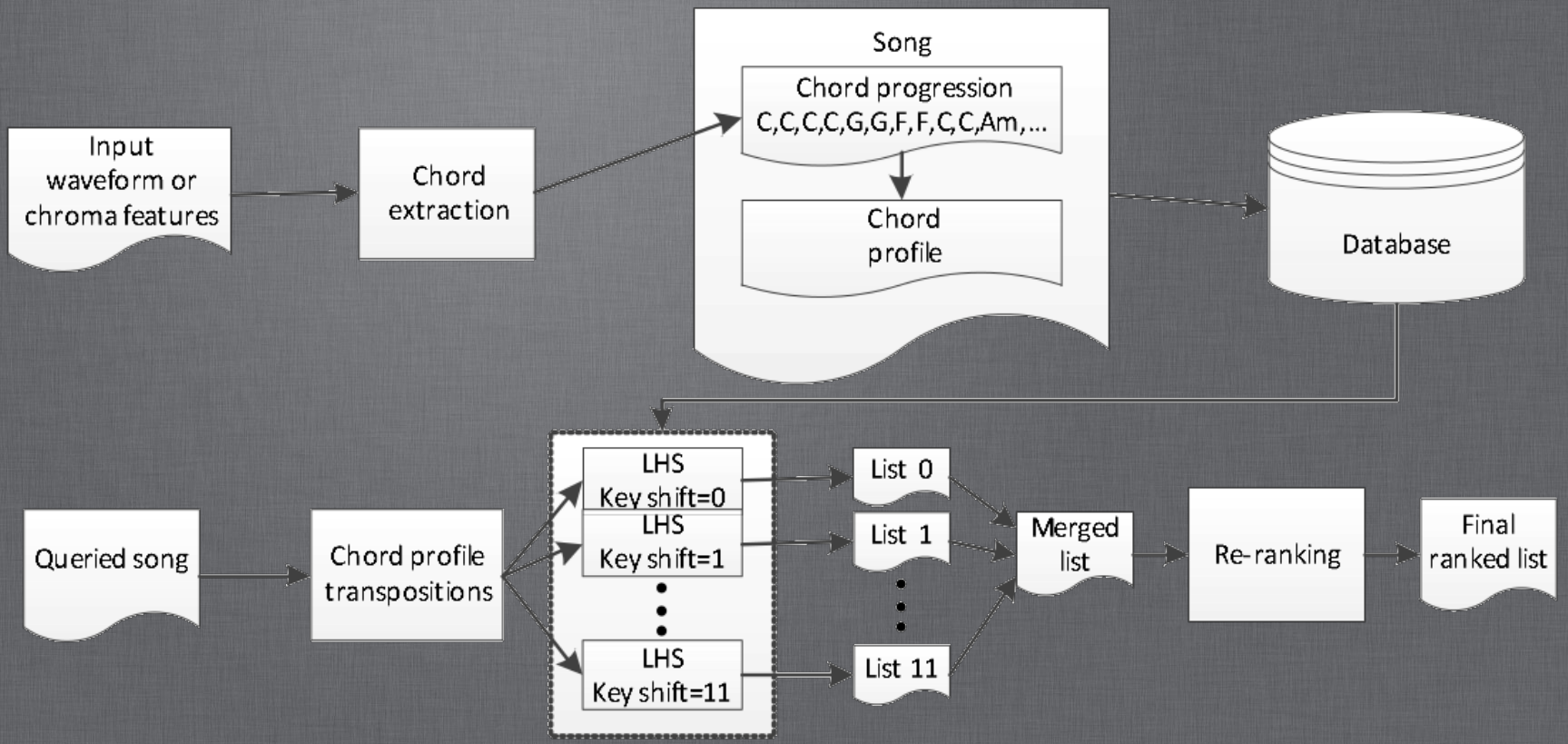
- Hou et al.: Hierarchical K-means tree, dynamic progr.
- MusicRadar
- Benchmarking: MIREX 2015

3. Cover Song Identification

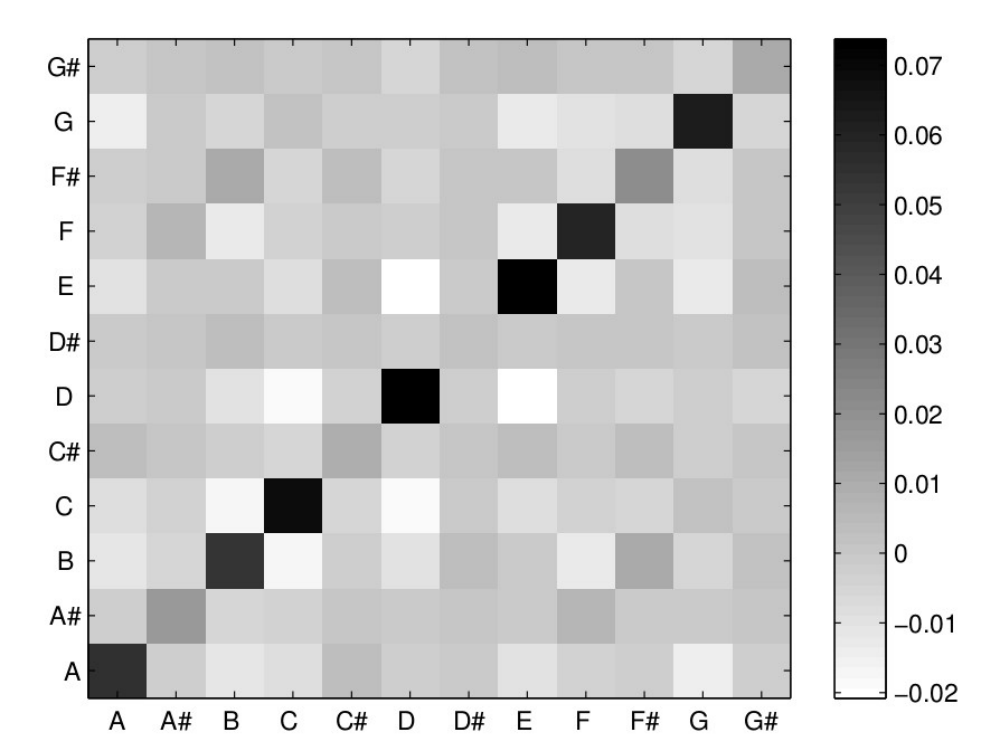
INPUT: Song / Recording

OUTPUT: Cover song / Performances

3. Cover Song Identification
Khadkevich and Omologo: CSI Using Chord Profiles (2013)



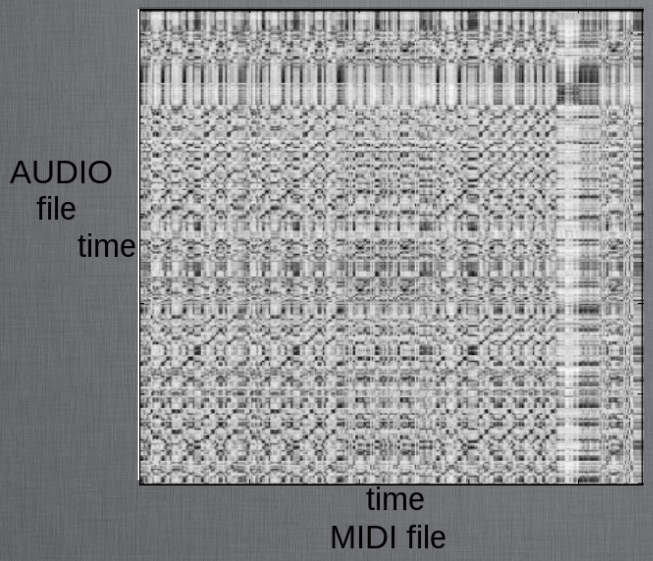
3. Cover Song Identification
Kim et al.: Music Fingerprint Extraction

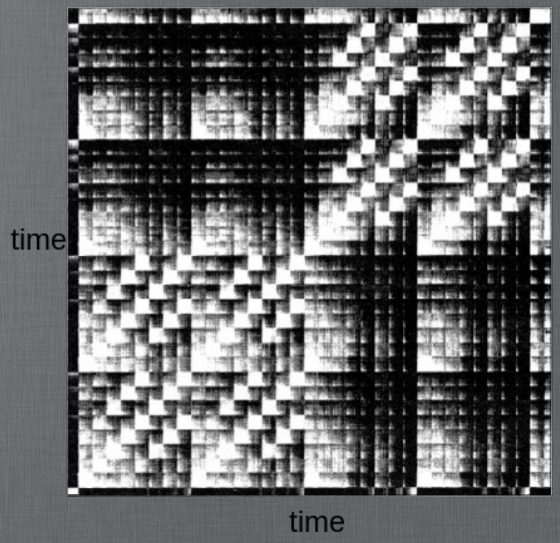


Use of Covariance Matrix Fingerprint, Beat synchronization

3. Cover Song Identification

Cross-Similarity and Self-similarity matrices (Tzanetakis 2003, Foote 1999)

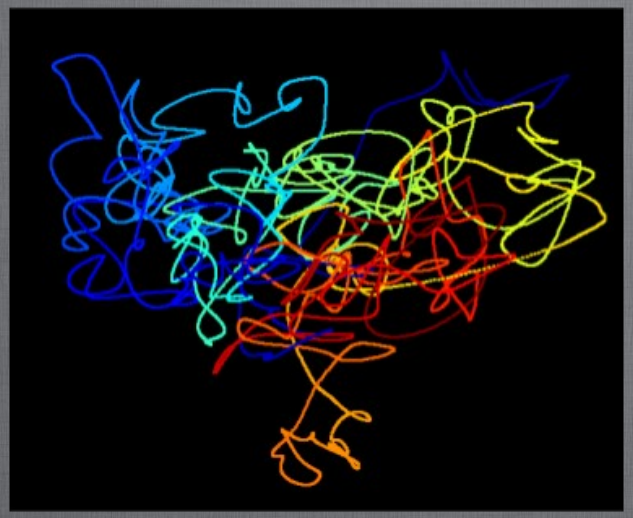


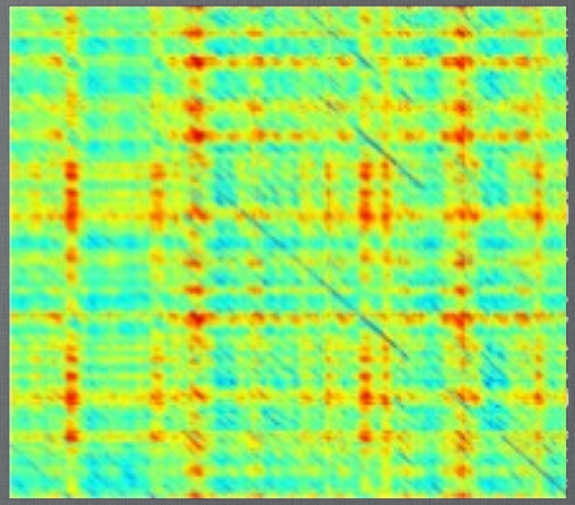


Alignment using: Chromagram, Spectrogram

3. Cover Song Identification

Cross-Similarity using MFCC (Traile, 2015)





Alignment using: MFCC

Summary

- Many various techniques
- Overall 80-90% precision of identifying covers

State-of-the-art

- Benchmarking: MIREX 2015
- Academia Sinica (Tsai, Wang): Melody extraction
- Bordeaux (Hanna): Local alignment of chroma sequences

4. Audio Similarity

INPUT: Song

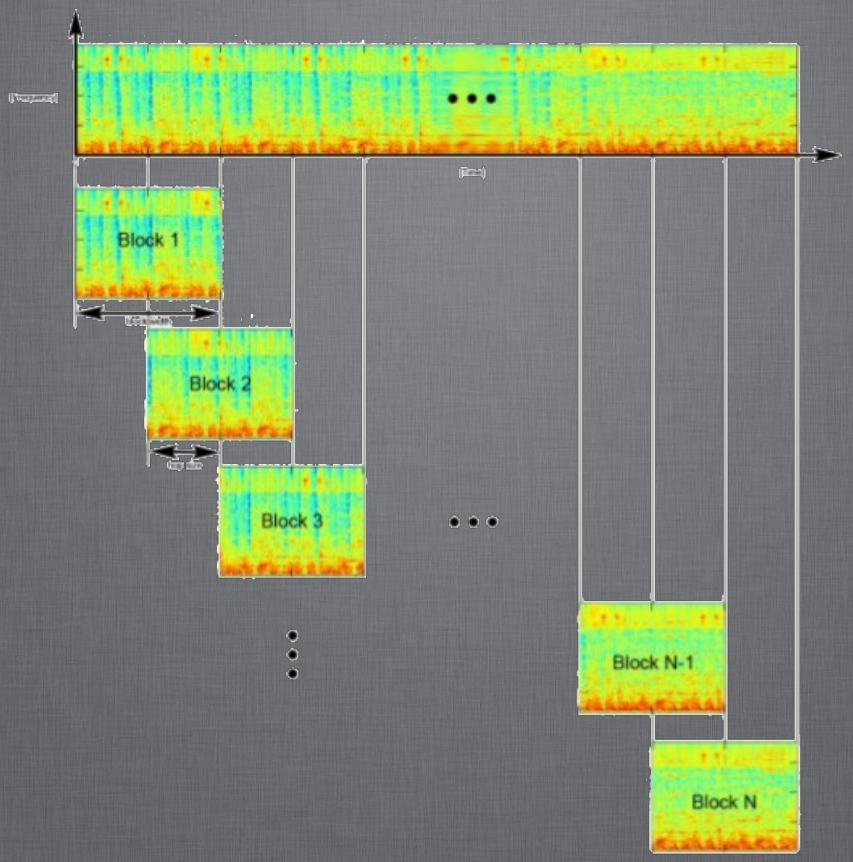
OUTPUT: Similar sounding song

Music recommendation:

OUTPUT: Song that user would like to listen to

4. Audio Similarity

Seyerlehner, Schedl: Block-Level Audio Features (2009)



Audio → blocks

deriving features from blocks

generalizing for the song

Distance measures

Summary

- Many various techniques
- Useful for genre classification / maybe recommentation?

State-of-the-art

- Benchmarking: MIREX 2015

Categorization of techniques

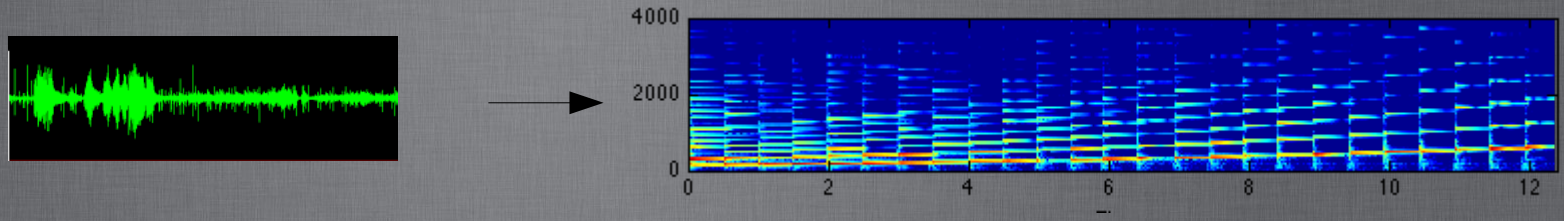
Audio → Spectrogram

Audio → MIDI

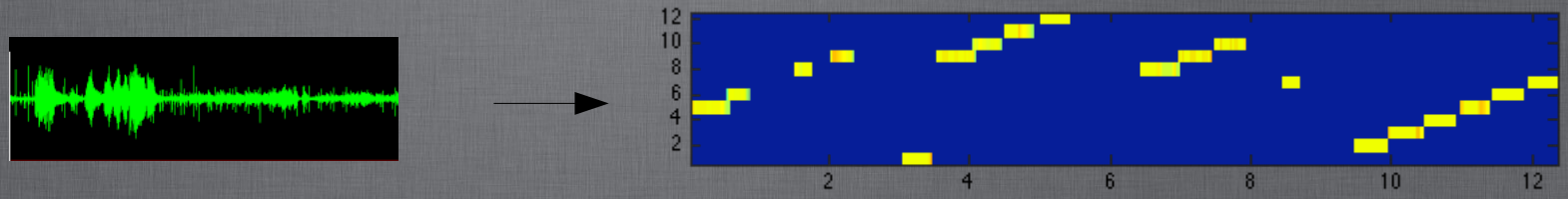
Audio → Chromagram

Categorization of techniques

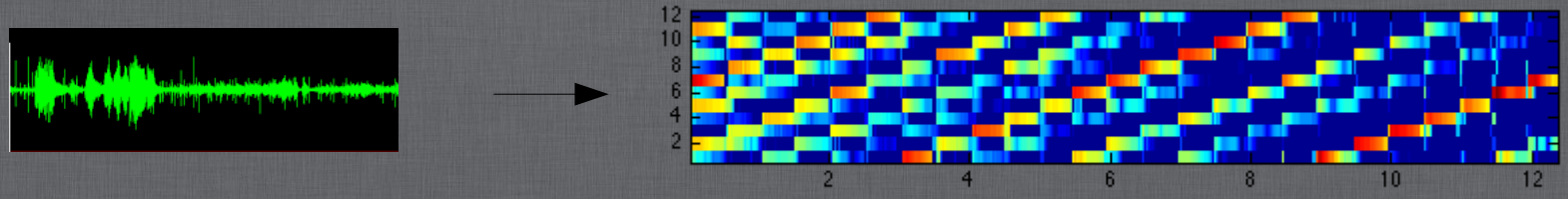
Audio → Spectrogram



Audio → MIDI



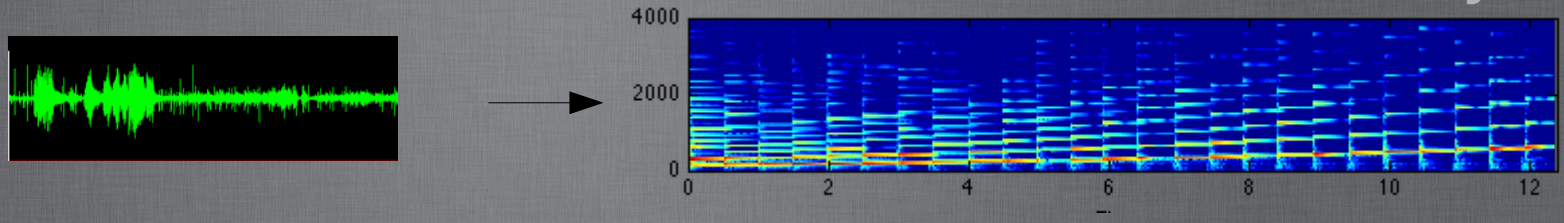
Audio → Chromagram



Categorization of techniques

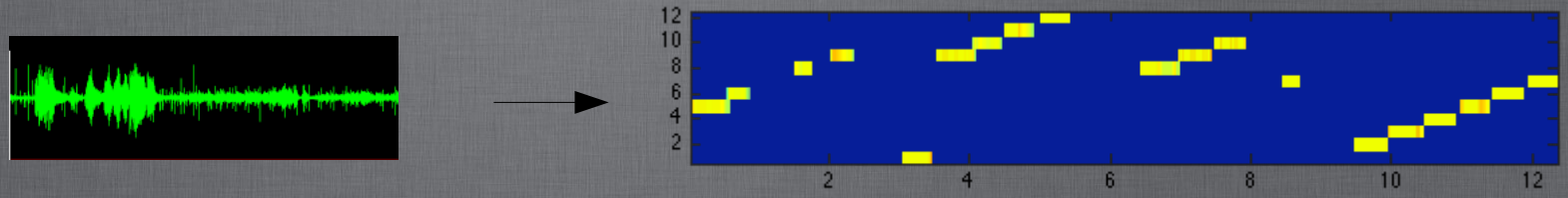
Audio → Spectrogram

1. Audio Fingerprinting
4. Audio Similarity



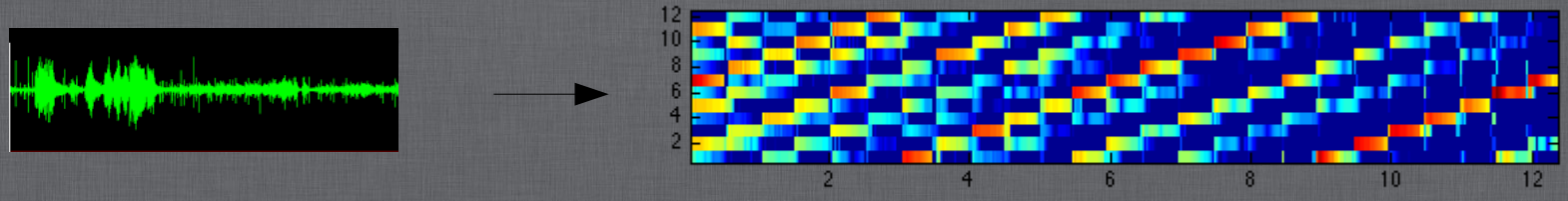
Audio → MIDI

2. Whistle and Humming Queries



Audio → Chromagram

3. Cover song identification
4. Audio Similarity



Thank you for your attention