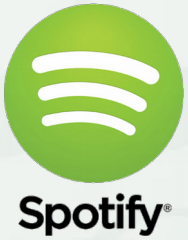
A Survey on Music Retrieval Systems Using Microphone Input

Ladislav Maršík[1], Jaroslav Pokorný[1], Martin Ilčík[2]
[1] Charles University, Prague
[2] Vienna University of Technology, Vienna

Music Information Retrieval (MIR)

It has many applications

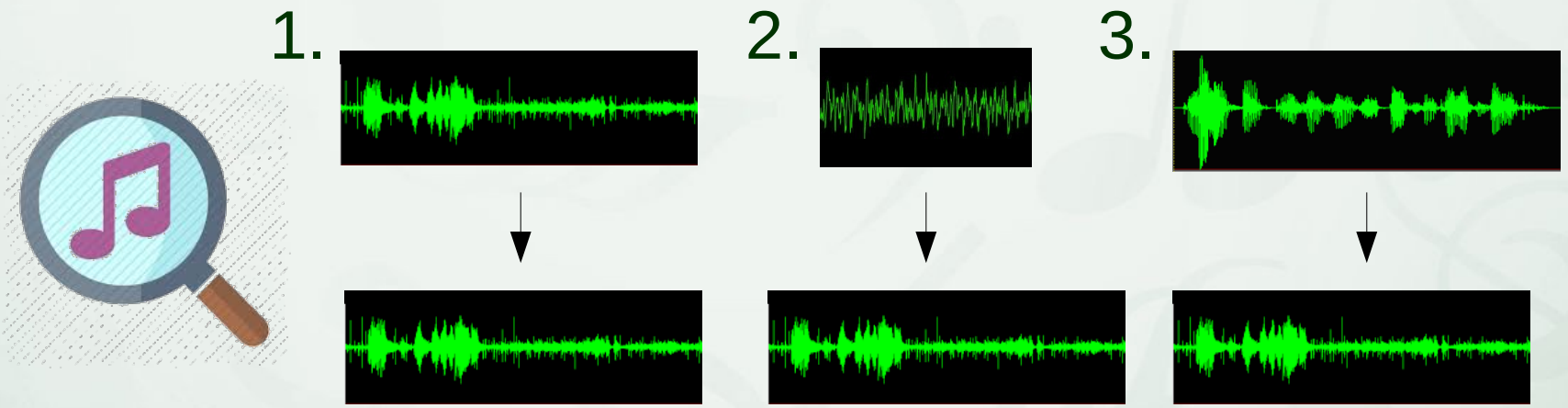


Motivation

- Understand recent MIR systems
- Find out where we can make improvements
  - Recognizing
  - Segmenting
  - Annotating
  - Recommending
  - Retrieving
  - Composing
  - Notation
  - Storing
  - Playback
  - Understanding

Music Retrieval

1. Audio Fingerprinting
2. Whistling and Humming Queries
3. Cover Song Identification



Audio Fingerprinting

INPUT: Song recording

Audio Fingerprinting

INPUT: Song recording

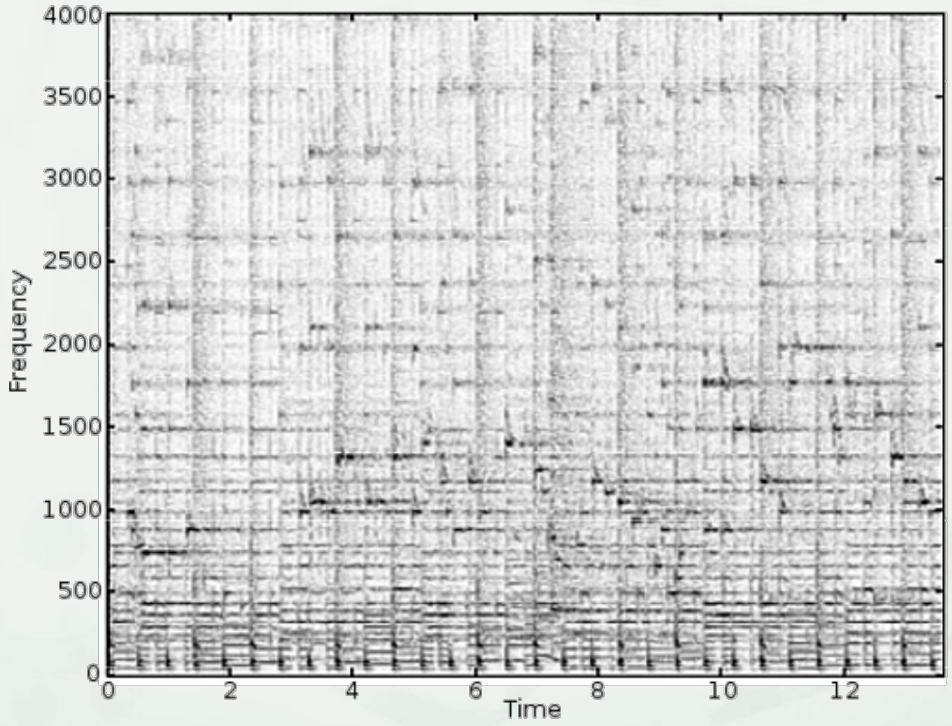
OUTPUT: The exact match

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

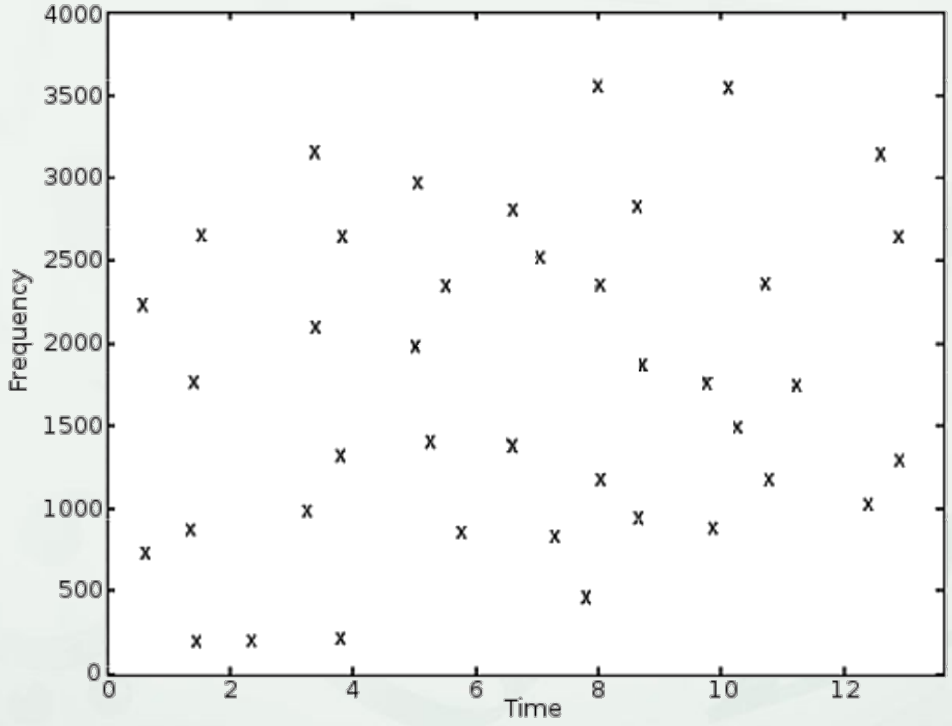
“Combinatorially hashed time-frequency constellation analysis”

- Time-Frequency
- Constellation analysis
- Combinatorially hashed

Time-Frequency spectrogram

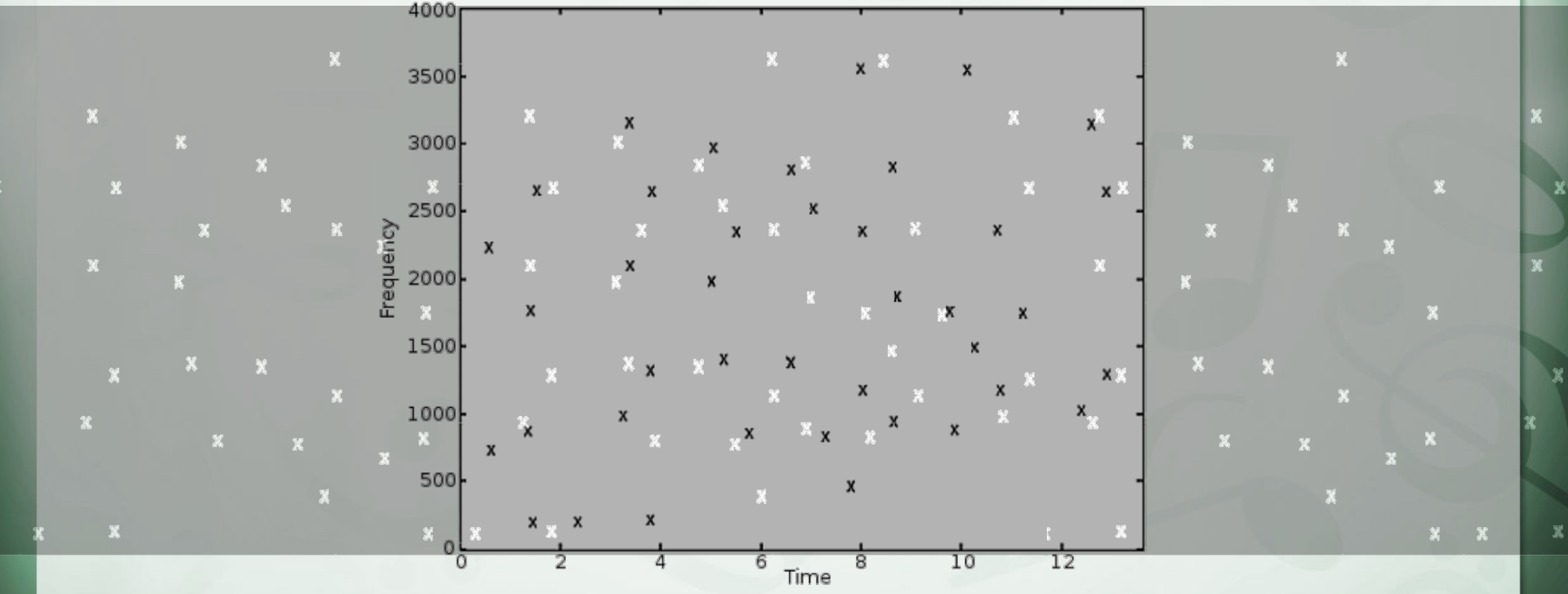


Constellation analysis



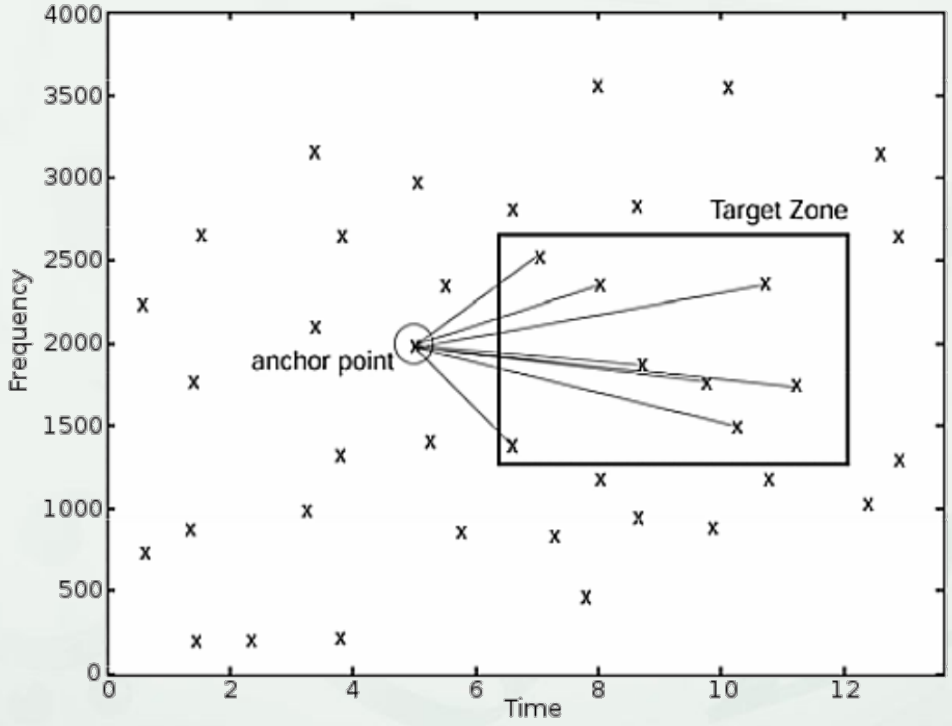
Audio Fingerprinting
Wang and Smith: An Industrial-Strength Audio Search Algorithm (2002)

Constellation analysis



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

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



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

Summary

- Short search time: 5-500 milliseconds / query
- Robust to noisy environment
- Possible extension to abstract from tonality

- Only exact match results

Audio Fingerprinting
State-of-the-art

- No benchmarking until recently

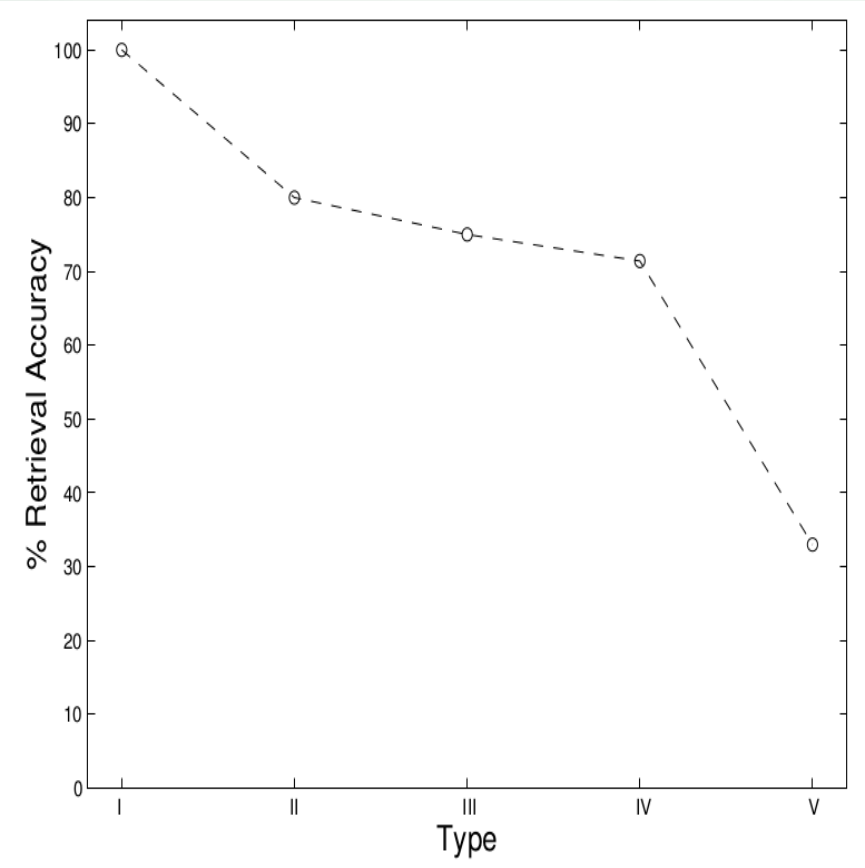
  (focus on commercial deployment)

Audio Fingerprinting
State-of-the-art

- No benchmarking until recently
  (focus on commercial deployment)
- Various indexing techniques

and peeks comparison algorithms

Audio Fingerprinting
State-of-the-art



Yang (2001)

Peek sequence: $P_1P_2P_3$ ...

Euclidean distance

- Type I: Identical digital copy
- Type II: Same analog source, different digital copies, possibly with noise
- Type III: Same instrumental performance, different vocal components
- Type IV: Same score, different performances (possibly at different tempo)
- Type V: Same underlying melody, different otherwise, with possible transposition

Audio Fingerprinting
State-of-the-art

- No benchmarking until recently
  (focus on commercial deployment)
- Various indexing techniques
and peeks comparison algorithms
- New use cases: Advertisement, TV program

Whistling and Humming Queries

INPUT: Whistling or Humming

Whistling and Humming Queries

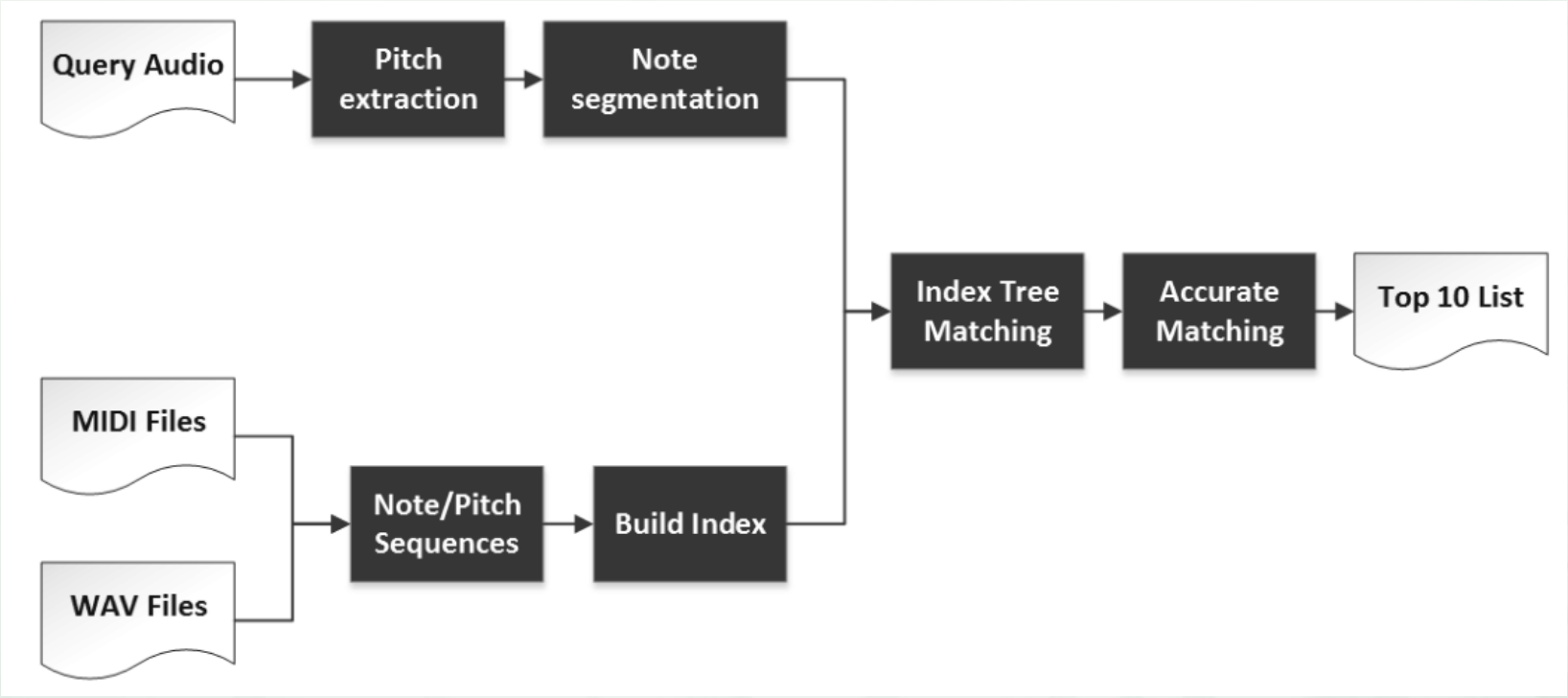
INPUT: Whistling or Humming

OUTPUT: Song containing the melody

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

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

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



Whistling and Humming Queries
Unal et al.: Query by Humming Systems (2008)

- Use of fingerprinting (relative pitch movement)

Benchmarking: MIREX 2014

(Music Information Retrieval Evaluation Exchange)

http://www.music-ir.org/mirex/wiki/MIREX HOME

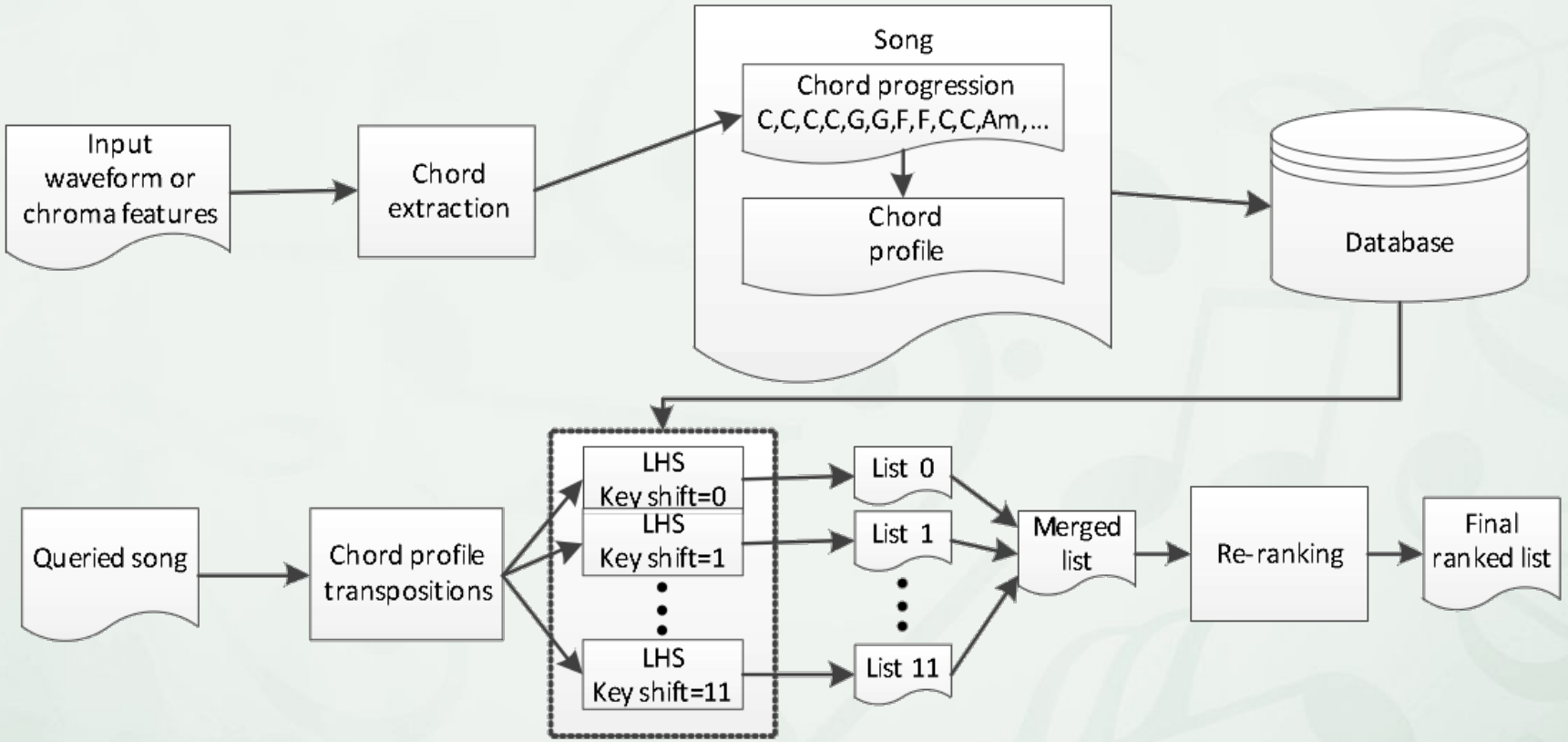
- Hou et al.: Hierarchical K-means tree, dynamic progr.
- MusicRadar

Cover Song Identification

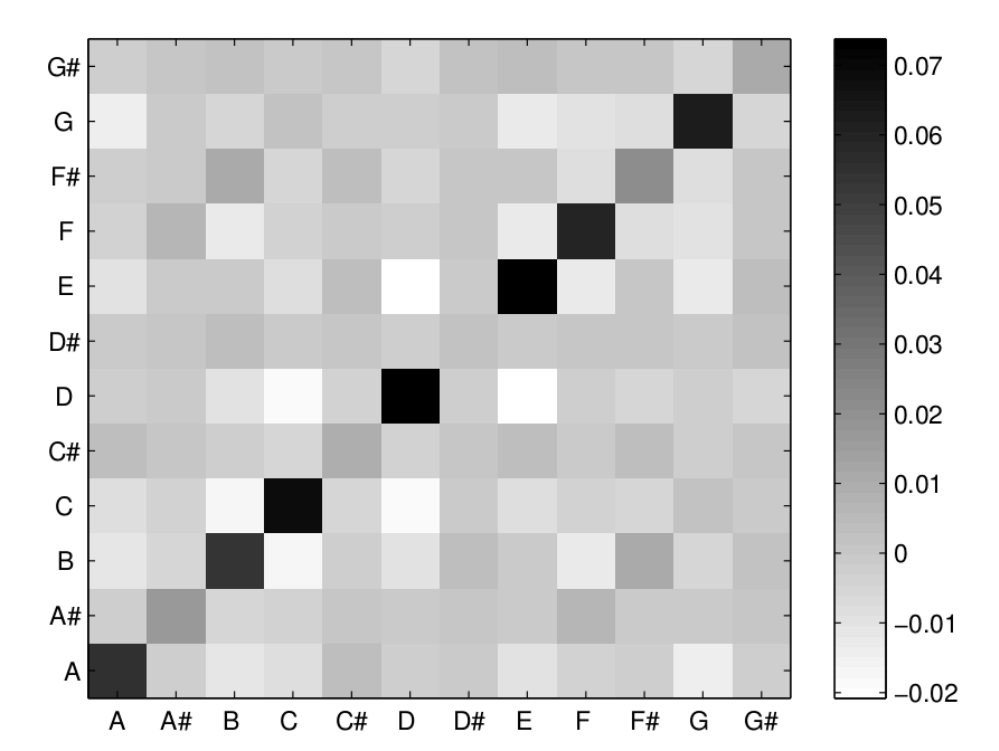
INPUT: Song / Recording

OUTPUT: Cover song / Performances

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



Cover Song Identification
Kim et al.: Music Fingerprint Extraction



Use of Covariance Matrix Fingerprint, Beat synchronization

Cover Song Identification
State-of-the-art

Benchmarking: MIREX 2014

(Music Information Retrieval Evaluation Exchange)

http://www.music-ir.org/mirex/wiki/MIREX HOME

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

Overall 80-90% precision of identifying covers

Proposals for improvements

- Low-level vs. High-level techniques
- Melody, Harmony, Tonality, Rhythm, Tempo
- Stabilize descriptors and use DTW to find similarities
- Combine Cover Song Identification with Microphone input methods

Summary

Survey on Music Retrieval Systems:

- Audio Fingerprinting
- Whistle and Humming Queries
- Cover Song Identification

- Proposal for improvements

Thank you for your attention