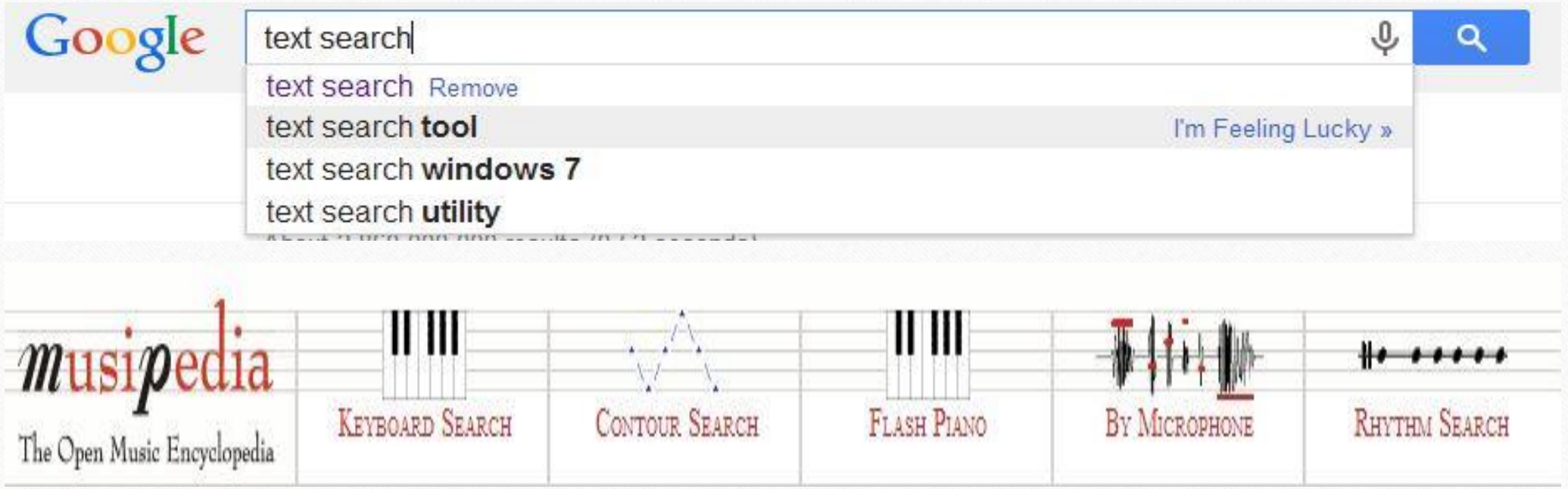
Music Information Retrieval Past, Present and Future

Kaavya Srinivasan

When did all this start?

- The term music information retrieval was heard in 1960's.
- MIR research was first started by various universities in early 1990's.
- The first international conference was started in 2000, ISMIR by Stephen Downie, Donald Byrd, Simon Dixon.

What is MIR?


Google
text search
text search Remove
text search tool
I'm Feeling Lucky »
text search windows 7
text search utility
musipedia
The Open Music Encyclopedia
Keyboard Search
Contour Search
Flash Piano
By Microphone
Rhythm Search

Initial works in MIR

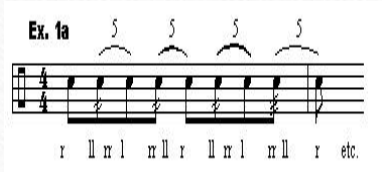
- Text Based
  - A meta-data of music.
  - Categories for describing music.
  - Artist information, Album information, genre, album title, artist title.
  - Any text based retrieval approaches can then be used.

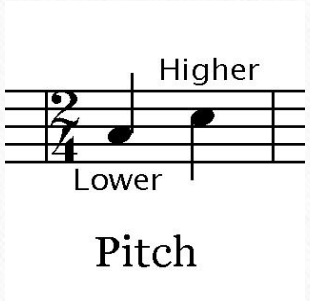
Current work in MIR

- Text based features are still in use.
- Research focus has moved on to finding similarity between music files.
- Difficult as music files have variety of features.
- Recommendations systems that are suggestion based are gaining interests.

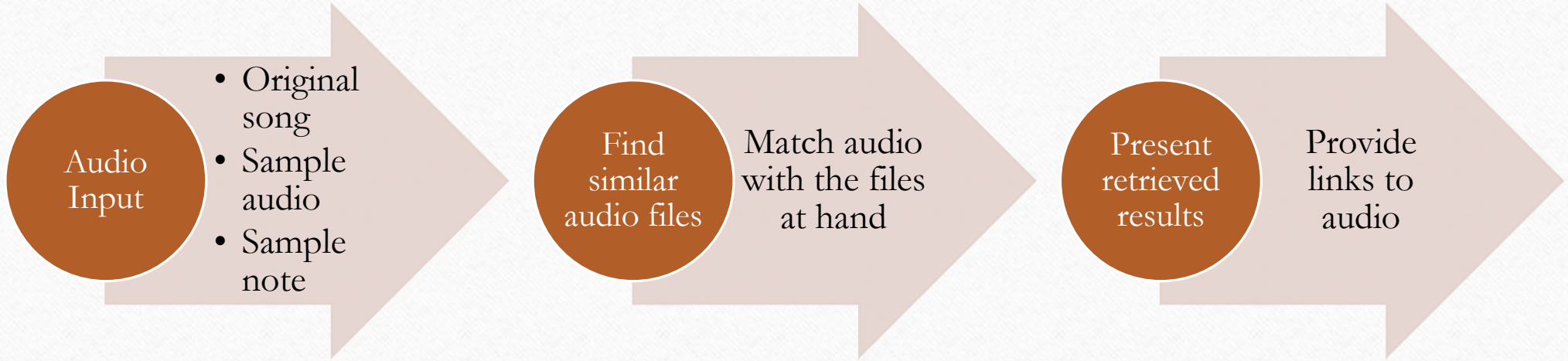
How is it done?

- Content Based
  - More difficult than the text based.
  - Collection of raw audio files.
  - Pre process the data by extraction of pitch, melody, rhythm.





Query Process



How to find similar audio files?

- Monophonic tunes are easier to compare than polyphonic tunes
- Monophonic tunes have only single pitch. Audio feature extraction becomes easier.
- Polyphonic tunes have multiple pitches and thus comparing becomes difficult.

Similarity music files

- Feature vector formation. Similar to vector formation in text documents.
- Audio extraction uses digital signal processing techniques.
- Each track is represented as Mel Frequency Cepstral Coefficients (MFCCs)
- MFCC's are obtained as a result of FFT, DCT techniques.
- Songs are represented as MFCC's in the feature dimension.
- Dimension reduction is done to match only specific features.
- Importance of a feature is calculated by chi-square test.
- When a new track is presented, a proximity feature is used to find similar tracks.

Common features

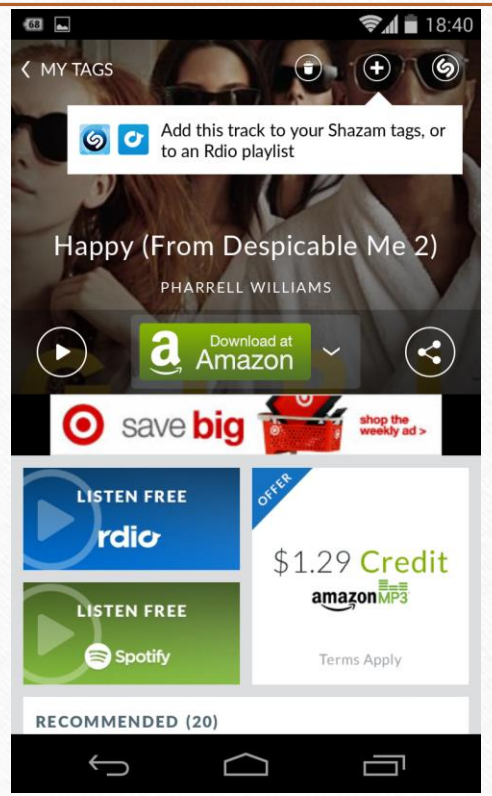
- Frequency, time, pitch (Also as intensity) are the most common features.
- Sometimes, temporal factors are ignored to find similarity.
- If an exact match is to be obtained, temporal features become crucial.

The problem

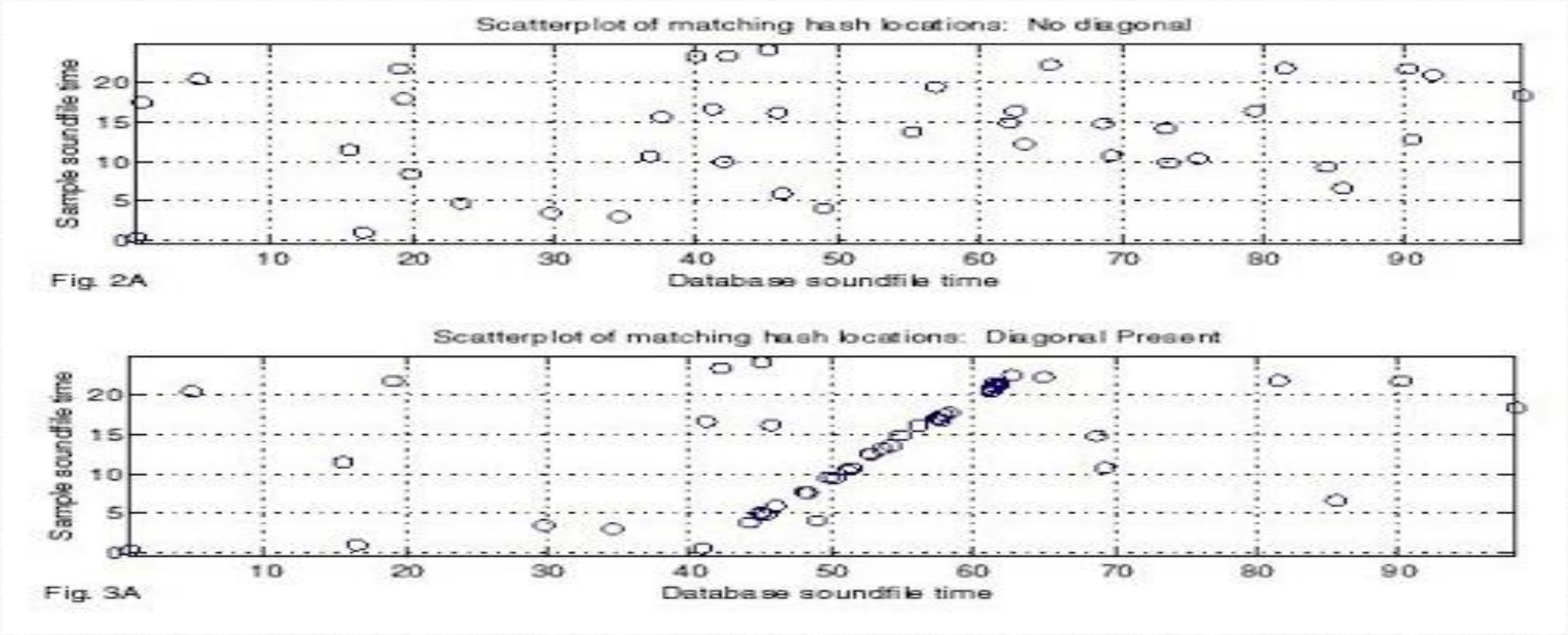
- Easier said than done.
- Humans can easily identify tunes, pitch, rhythm, beats.
- Songs evoke emotions.
- Integrating the features together in an almost perfect way is an ongoing research in MIR.
- There is no clear definition yet for polyphonic sounds.

An Example

- Shazam – Phone based music identification service.
- Identifies songs based on audio samples.
- Uses acoustic fingerprinting to find similarity between songs

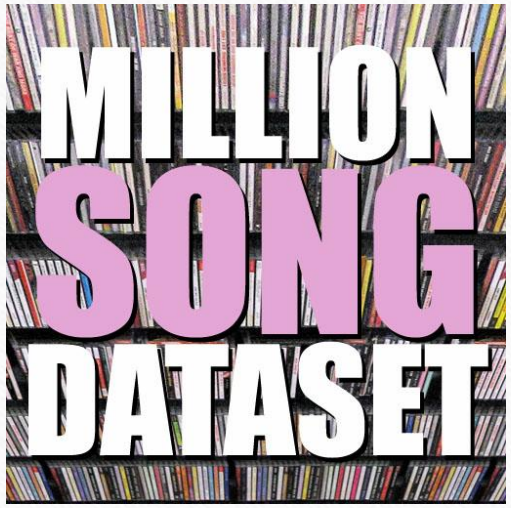


Fingerprint comparison



Applications of MIR

Few available databases



the echonest

last.fm

Future Work

- Good Recommender Systems are yet to come by.
- Extraction of polyphonic audio pitches.
- Providing similar audio files based on social media.

References

[1] Casey, M.: Audio Tools for Music Discovery. [http://www.methodsnetwork.ac.uk/redist/pdf/casey.pdf]

[2] Problems of Music Information Retrieval in the Real World, Donald Byrd University of Massachusetts Amherst

[3] An Industrial-Strength Audio Search Algorithm, Avery Li-Chun Wang, Shazam Entertainment, Ltd.

[4] Cheng Yang, "MACS: Music Audio Characteristic Sequence Indexing For Similarity Retrieval", in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2001

[5] Can Humans Benefit from Music Information Retrieval? Frans Wiering

[6] Million Song Dataset, Proceedings of the 12th International conference on music information retrieval, 2011