Song Recognition Using Audio Fingerprinting

Varun Khatri, Lukas Dillingham, Ziqi Chen

Department of Electrical and Computer Engineering

Audio Fingerprinting

- **How does it work?**
  - Use semantic features? No
  - Use non-semantic features? Yes
- **Require something that is**
  - Compressed
  - Unique
  - Invariant to degradation
  - Invariant to effects
- **Similar to Human Fingerprint**

Algorithm

Create fingerprint

Convert stereo to mono

Resample the signal to 8192 Hz

Apply hamming window (window length 1024)

STFT hop size of 32

Apply a MAX filter

Apply threshold

Extract peaks

Searching with hashes

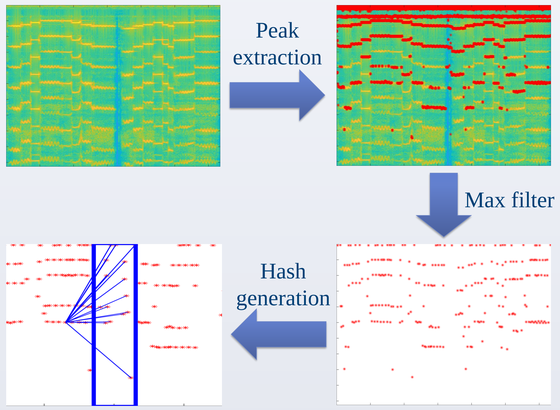
Generate hash tables from peaks

Compare hashes from song clip with database

Count number of matches

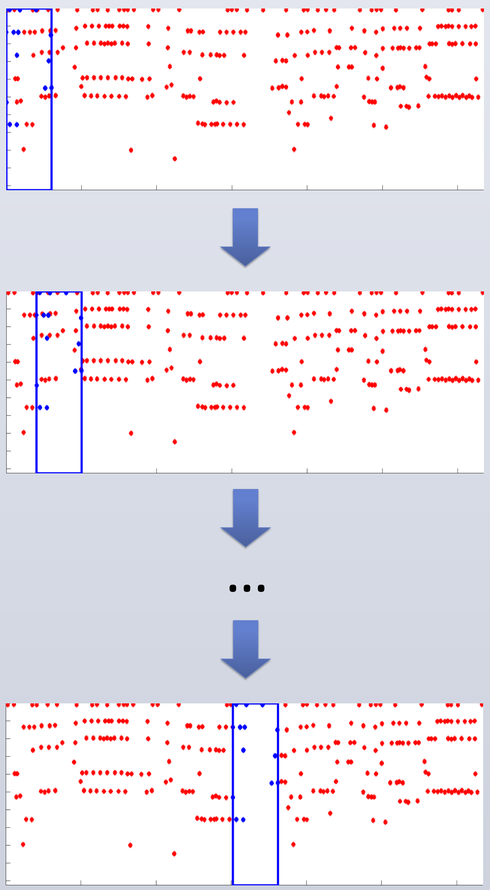
Fingerprinting & Matching

Fingerprinting



Matching

simple slide and compare



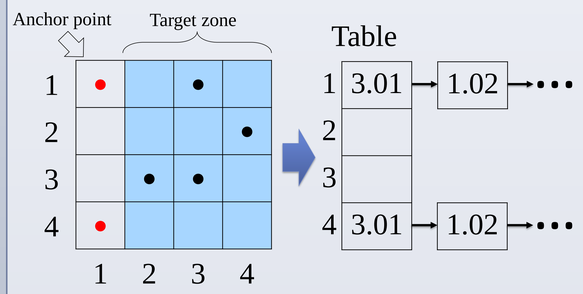
Hash tables for quick searching

hashes consist of :

- *Frequency of the anchor*
- *Frequency of the point*
- *Delta time between the anchor and the point*

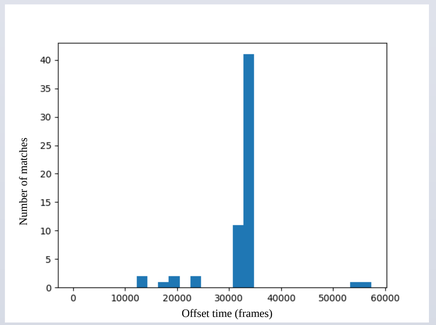
Hash table generation

$$\text{Table}[f_{anchor1}] = f_{target\ zone\ point} + \frac{\Delta t}{100}$$

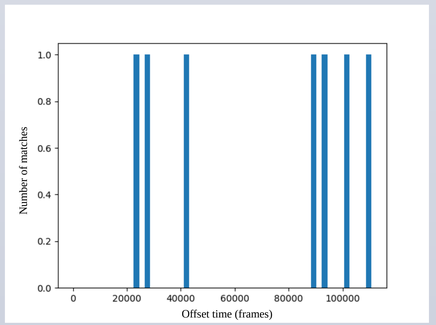


Results

Histograms of offset times for hash matches



(a) Correct song



(b) Incorrect song

Tables of hash search results:

Comparing a clip of "Bistro Fada" with different noise levels to itself and three other songs.

Indexes:

- *Ratio: no. of matches / total no. of hashes in clip;*
- *Histogram Max: max value of histogram for each comparison*

	-3 dB SNR	
Name	Ratio	Histogram Max
Bistro Fada	0.0651	40
Kamaloka	0.0091	6
Goodbye Pork Pie Hat	0.0115	6
The Star Boys	0.0429	22

	1 dB SNR	
Name	Ratio	Histogram Max
Bistro Fada	0.0857	65
Kamaloka	0.0111	8
Goodbye Pork Pie Hat	0.0121	7
The Star Boys	0.0656	26

	7 dB SNR	
Name	Ratio	Histogram Max
Bistro Fada	0.1797	99
Kamaloka	0.0124	5
Goodbye Pork Pie Hat	0.0183	5
The Star Boys	0.1217	37

	15 dB SNR	
Name	Ratio	Histogram Max
Bistro Fada	0.4645	136
Kamaloka	0.0561	6
Goodbye Pork Pie Hat	0.0508	11
The Star Boys	0.2787	40

	Original Song	
Name	Ratio	Histogram Max
Bistro Fada	0.6285	187
Kamaloka	0.0714	7
Goodbye Pork Pie Hat	0.0766	13
The Star Boys	0.3075	36

References

Wang, A. (2003, October). An Industrial Strength Audio Search Algorithm. In *ISMIR* (Vol. 2003, pp. 7-13).

Müller, M. (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*. Springer. pp. 360-370.