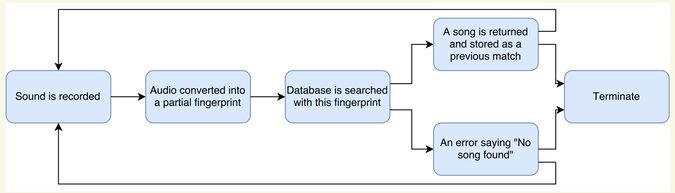
Stelhound: A Music Identification Tool

Introduction

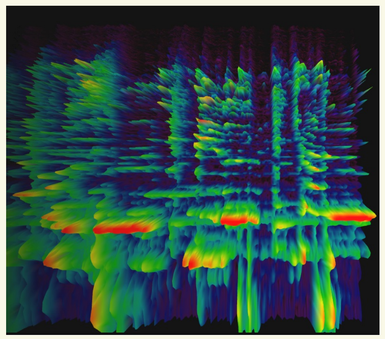
Stelhound is an open source music identification tool which records sound through a microphone and, if there is music playing, displays its information.

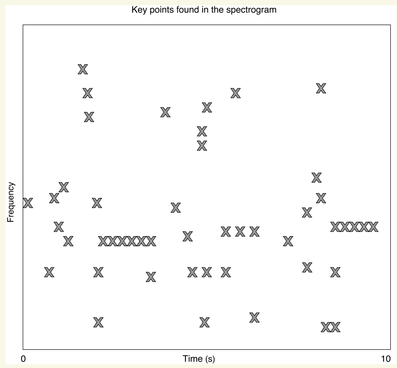
Pipeline



Uniqueness in music & fingerprint generation

What makes a sound in a song unique is a combination of its pitch, its amplitude and its timbre. Extracting this information is done with Fourier transforms.
A Fourier transform done on a song many times per second can be used to produce a spectrogram, a graph with time, frequency, and amplitude as its axes. Each different piece of music will have a unique spectrogram, so if a recording has a similar spectrogram to another piece of music it is likely that they represent the same song.





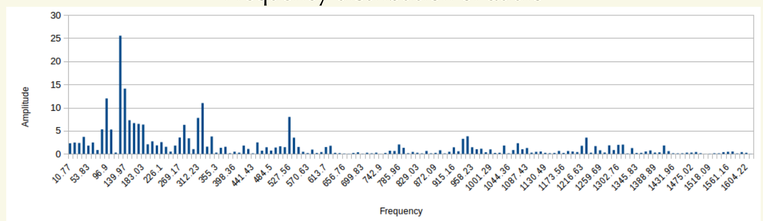
Data storage

This program's database stores information about each track along with its generated fingerprint. The information found in the header of .mp3 files is used in conjunction with the fingerprints to populate the database automatically. The database stores each song's Title, Artist, Duration, Album art and fingerprint. If any of these fields are absent they are given a null value. A NoSQL database was used for this program as the information is most efficiently stored in a non-relational way.

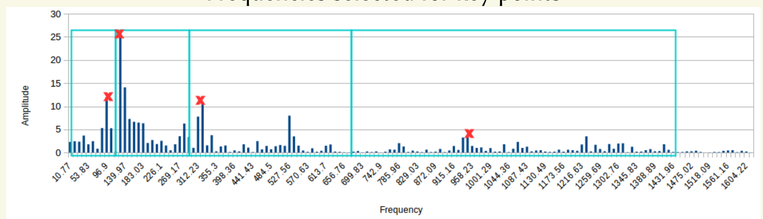
Fingerprinting

Instead of directly comparing the two spectrograms, every 10th of a second the dominant frequencies are chosen from each. The information is then converted into a *fingerprint fraction* and all the fractions together create the *fingerprint*. Once all the fingerprints have been compared with the fingerprint generated from the recording, the program decides whether there is a match or not.

Frequency distribution of audio

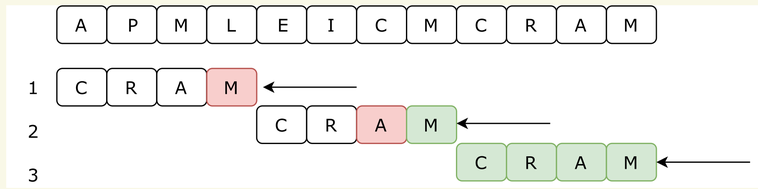


Frequencies selected for key points



Searching

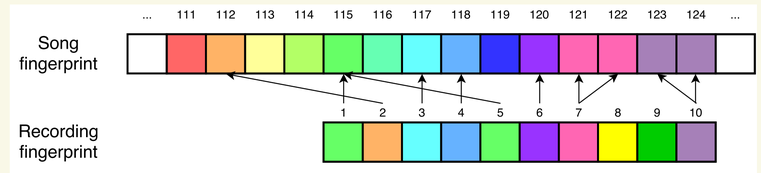
The program uses the Boyer-Moore substring matching algorithm to search through each fingerprint in the database. For every match the algorithm finds, it records the offset. This offset is then used in the matching process.



Boyer-Moore has a time complexity of $O(nm)$ in the worst case and $\Omega(n/m)$ in the best case, where m is the length of the fingerprint and n is the length of the fraction. The algorithm is optimised for this application as it specialises in finding small sequences in much larger ones, making it the best choice for the program.

Matching

In order to correctly match songs to recording's time information needs to be included in the frequency data. To do this an offset matching algorithm is used which does a side by side comparison between a song's fingerprint and the recording's fingerprint. An example is given below.



Resulting map:

110: 2

113: 1

114: 6

115: 1

1. A search for the fraction at position 1 in the recording is conducted through the song's fingerprint.
2. A match for this fraction is found in the song's fingerprint at position 115.
3. The offset value is calculated ($115 - 1 = 114$).
4. There are currently no stored values so a new `HashMap` entry is made with 114 as the `key` and 1 as the `value`.
5. The rest of the fingerprint is searched for another instance of the fraction, but none is found. The algorithm then resets and does the same process used on the fraction at position 2.
6. A match is found at position 112.
7. The offset is now calculated as $112 - 2 = 110$ and stored.
8. This algorithm continues until the end of recording and produces the resulting map which shows 114 to have the highest counter.

Results

The final program was tested with different tracks from different genres and gave positive results, rarely finding a false positive or negative. This result is acceptable and shows that the underlying premise of the program works.

The program's main issue is that the searching solution does not scale well. This is an issue that could be addressed in future work on this project.

Matthew K. Jenje
18516009@sun.ac.za
Computer science department, Stellenbosch University, Western Cape, South Africa

UNIVERSITEIT STELLENBOSCH UNIVERSITY