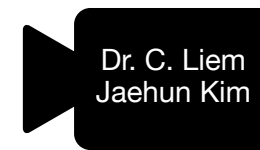


Open-source audio fingerprinting for identification of music in movies

CSE3000: Midterm Progress Poster



AUDIO FINGERPRINTING



MOVIE MUSIC IDENTIFICATION



Could audio fingerprinting help?

1. Standard for music identification
2. Uses perceptual characteristics
3. Configurable parameters

2 Fingerprinting with *dejavu*

Research questions

*How do insights from academic research literature connect to *dejavu*?*

*How does *dejavu* perform in practice in music identification in movies based on the collective benchmark?*

Starring *dejavu*

An open-source audio fingerprinting framework written in Python

Important findings

- Fingerprints with computer vision
- Sensitivity to pitch shifts and tempo differences
- Good performance with reasonable amounts of noise

Method

1. Literature study and framework analysis

- A. Research the algorithm of *dejavu*
- B. Analyse the implementation and identify configurable parameters

2. Empirical study

- A. **Data collection**
 1. Determine common distortion-over-music categories
 2. Extract audio from suitable movie clips
 3. **Synthesise data based on pre-determined categories**
- B. Test and adjust *dejavu*
 1. Explore capabilities of *dejavu* using the data
 2. Adjust the configurable parameters in *dejavu* and observe performance

Distortion categories

1. Speech
2. Ambient noise
3. Clean recordings

Structural categories

1. Pitch shifting
2. Tempo changes
3. Low amplitude of music
4. Compression

Bonus content

Researching *dejavu* performance gaps and fixing low points, such as pitch shift and tempo sensitivity

DONE

DOING

TODO

1 The Collective Benchmark

How do we determine a collective benchmark to evaluate the performance of audio fingerprinting algorithms for music identification in movies?

Description

What is it? A set of universal metrics to measure performance of audio fingerprinting frameworks in movie music identification.

Why do we need it? To compare different frameworks fairly.

How do we get it? The benchmark was determined with from combination of research and justification of significant properties.

Criteria

Robustness

How often does the framework fail to identify the track?

$$\frac{FN}{FN + TP}$$

Reliability

How often does the framework identify the track incorrectly?

$$\frac{FP}{FP + TP}$$

Match speed & scalability

How does the framework perform with a large database of fingerprints to compare?

F = False, T = True; P = Positive, N = negative