MIR Project 1 Description and Scope

Team: Donghyun Park, Josh Miao

Task Scope: Recreate MIREX 2019 Audio Cover Song Identification

Baseline Algorithm: CNN

In contrast to the classical method of encoding and computing Euclidean distance in the latent space, we will implement cover detection with CNN. CQTNET takes a CQT as an input(as the name would suggest) and learns to extract key and tempo invariant representations as a music similarity metric. The CQT input passes through convolutional layers for feature extraction and long-term melodic structure, undergoes adaptive global pooling, produces a 300-dimensional embedding as the song representation and finally passes through a softmax for classification.

Dataset:

- Second Hand Songs 100K
    - Most likely end up using a small subset as downloading 100k sound files would be difficult to store.

Instructions for Dataset Download:

- Metadata from GitHub
- Actual songs ripped from youtube, using metadata

Relevant Links:

- MIREX 2019 Submission
- Full Implementation Paper
- CNN Source code