Project Proposal: Harmony Analyser Using RNNs

Adrian McIntosh

Objectives

The primary objective of this project would be to create a model, that takes in an audio recording and classifies the chord structure of the recording.

The project can be done in a scalable way to allow for progress to be made within the time constraints. Additionally, as the project develops limitations and opportunities may present themselves which might change these objectives as experimentation is done. With this in mind the following objectives have been chosen at different levels of complexity:

1. **Single Note Transcription (Piano only):** For this stage, only monotonic piano melodies will be input as the training data, and the goal would be to output the timestamp and note (eg C4, C5 etc.)
2. **Single Note Transcription (Other instruments):** Stage 1 will be expanded to using other instruments as input (for example monotonic vocal, flute and guitar melodies)
3. **Piano Harmony Recognition:** This would expand on the previous stages, now inputting piano harmony pieces and outputting the timestamp and chord and chord quality (first the triad eg C, Dm, Gdim, and then to extended harmonies Am7, CMaj7, C6). This stage could also be scalable in itself, starting out with MIDI produced audio, and moving onto real recordings
4. **General Chord Analyser:** This stage would use not only piano music but various musical recordings (pop songs and full arrangements)

Methodologies

Fourier Analysis and Frequency Domain Features

Using signal processing techniques (like the Fast Fourier Transform), features of the frequency spectrum of segments of an audio recording can be used as input to the system, rather than just the raw recording to remove noise, and better highlight more significant aspects of the harmonic qualities of the recording (more information can be found here: https://epubs.siam.org/doi/pdf/10.1137/S00361445003822).

Common Chord Patterns and Temporal Data for RNN

The chord structures of songs and pieces often follow common patterns or guidelines (like a IV vi ii V I, or I IV V). This would be a useful feature in working on a project using RNNs as based on some preliminary research, they are particularly useful for problems that have a temporal element and depend on time-based structures or patterns.

Training Data

There is a large amount of publicly available music with labelled chord information making this project feasible with many observations to run the model off of. Some examples can be found at the following links: (https://github.com/tmc323/Chord-Annotations, https://www.music-ir.org/mirex/wiki/2021:Audio_Chord_Estimation#Chord_Segmentation). Additionally, MIDI files and digital music notation software like musescore can be used to generate training data.