Computational Music Analysis Proposal

Automatic Note Extraction from a monophonic piece and Graphical Representation of note sequences

Thiago R Roque, Rhythm Jain, Noel Alben

September 2021

1 Problem Statement

We present a system that is capable of extracting note level information from a given audio signal of a monophonic melodic performance. We intend to use the extracted information to generate graphs that represent the relationship and frequency of the notes performed in the given piece.

2 Motivation

This project will serve as a precursor for future applications in:

- Cross cultural or cross genre comparison of melodic trends
- The relationship established between the notes can help train future generative models
- It can be used in the field of education as the graphical output will serve as an intuitive representation for note-note relationships and frequency of occurrence within a piece.

3 Deliverables

We aim to present a web interface that accepts a monophonic audio performance as input and generates a transcription along with a graphical representation of the note sequences.
A potential user can improvise a monophonic piece and gain insight on their melodic traversal throughout the performance.

4 Measure Of Success

The ground truth is obtained from the annotations in the data sets below:

- DAMP:`https://ccrma.stanford.edu/damp/`
- NSynth:`https://magenta.tensorflow.org/datasets/nsynth`
- `https://www.upf.edu/web/mtg/news/-/asset_publisher/WM181VyAQipW/content/id/218066353/maximized#.YT1jlp5KiCi`

The annotations follow the MIREX multiple-f0 estimation (frame-basis) format. This format is also support by mir_eval. For Validation of our note and onset estimations, they are compared with the annotated time stamps and f0 information from the above datasets.

5 Method

5.1 Audio transcription (MIR)

Based on our initial literature survey, we will have to complete the following tasks:

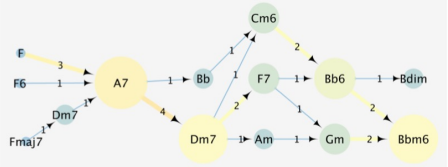
- Convert the frame level audio to spectogram representation
- Extract Pitch information from the spectogram
- Derive onsets and offsets from the given audio sample
- Use the Onsets and pitch information between inter onset intervals for note estimation and transcription [1]
  We are going to look at multiple signal processing approaches that will provide us with the best results for our goals.

5.2 Graph visualization

The second part of our project will:

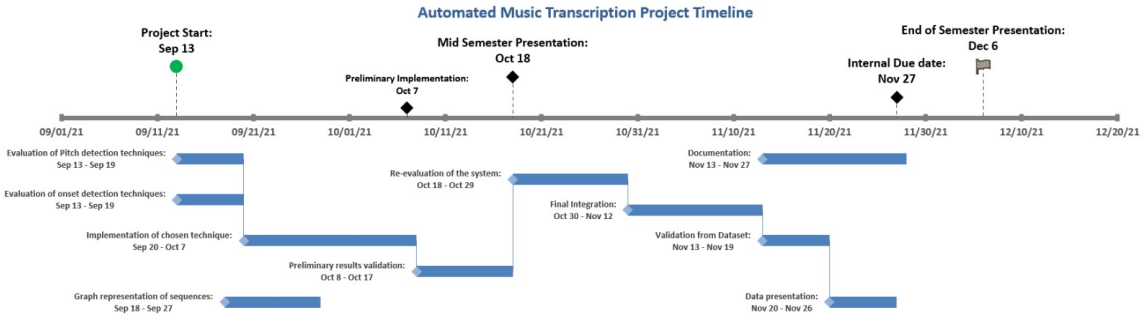
- Use graph theory to represent melodic sequences and the frequency of note occurrence within the given audio signal using the estimated note information

Figure 1: Example of a graphical representation from [2]



The above figure represents Harmonic Sequences using nodes and their relationship, we plan to implement this at the note level. [2].

6 Project Timeline



- Week 1: Evaluation of State-of-Art techniques
- Week 2: Implementation of chosen technique
- Week 3: Preliminary results validation
- Week 4: Present Note level information extracted from an audio signal and be able to represent graphs for given data. [Mid Sem Presentation Prelim Results]
- Week 5: Re-evaluation of the system
- Week 6,7: Integration of our system
- Week 8: Validation from Dataset
- Week 9: Data presentation
- Week 10: Documentation

References

[1] Emmanouil Benetos, Simon Dixon, Zhiyao Duan, and Sebastian Ewert. Automatic music transcription: An overview. *IEEE Signal Processing Magazine*, 36(1):20–30, 2019.

[2] Jeff Miller, Vincenzo Nicosia, and Mark Sandler. Discovering common practice: Using graph theory to compare harmonic sequences in musical audio collections. New York, NY, USA, 2021. Association for Computing Machinery.