Senior Project – Electrical Engineering – 2020

Audio Detection, Conversion, and Transcription to MIDI & Sheet Music

Ian Krause, Raphael Sebastian II
Advisor – Aussie Schnore

INTRODUCTION

- Music is typically recorded as a digital waveform in the time domain that is 'continuous' in the sense that the notes cannot be differentiated or changed individually once recorded.
- This poses a problem for digital musicians who work with discretized music – MIDI (musical instrument digital interface) – and want more control over recorded audio.
- Additionally, musicians who work digitally are not often familiar with or literate in traditional sheet music, but may want to write a part for a classically trained musician to use.
- Accessible programs that allow collaboration between traditional musicians and digital musicians are usually very expensive and their accuracy leaves room for improvement.

DESIGN REQUIREMENTS

1. The system should accept monophonic (one note at a time) audio input and polyphonic (multiple notes at a time) audio input.
2. Total error must be less than 30% for polyphonic transcription and less than 20% for monophonic transcription, both at 120 BPM recording speed.
3. The system must output MIDI files, sheet music, and the original audio files.
4. Transcription must happen in real-time, or near real-time to so it can be used actively by musicians.

PRELIMINARY RESULTS

- During our fall term we developed a real-time visualizer that displayed the time series audio waveforms and frequency domain spectrum to analyze the waveforms and assess the precision of the data stream.
- We determined that we would be able to discern the notes easily and reduce noise by using a direct injection unit, and sending the balanced signal to low-cost analog-digital converter.
- We also discovered that the data stream speed was fast enough to extract rhythm and frequency data in real-time at 120 BPM.

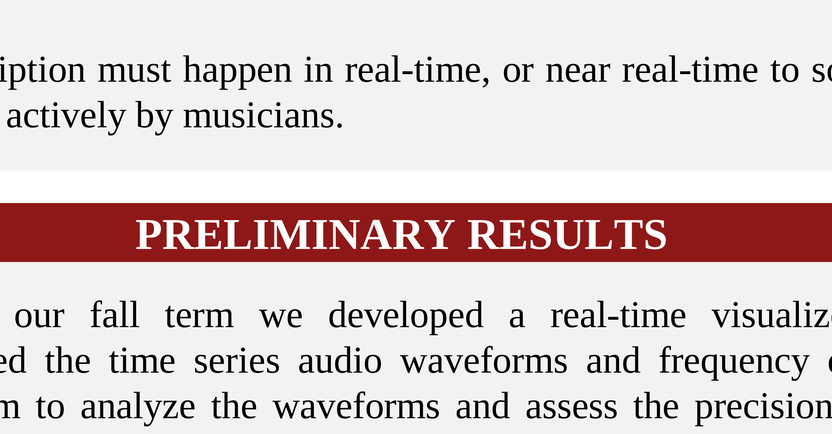

Figure 7: Time Domain (top) and Frequency Domain (bottom) Spectrum of a Whistle Roughly Approximating a Pure Tone

CURRENT DESIGN

- The design we plan to use has been implemented with Python code so that it is easy to tweak and can run on common microcontrollers like the Raspberry Pi.
- The current planned process for monophonic audio detection uses peak detection to find the fundamental signal frequency, while machine learning can be used for detecting polyphonic audio.

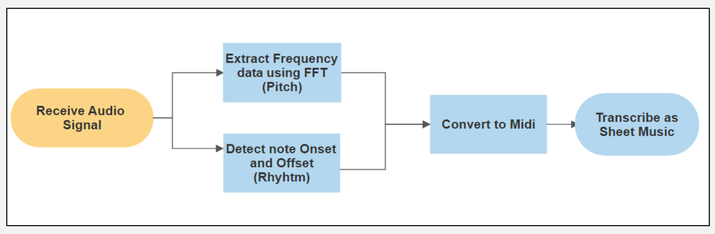

Figure 5: Flowchart of Monophonic Audio Transcription

```
1. For each Frame of input data:
2.    Convert Frame to Frequency Domain using DFT
3.    Find Peak Frequency, P
4.    If P is in the range of guitar frequencies:
5.          convert frequency to midi, store and display
6.          store time data and display
7. Use calculated data to build MIDI file
8. Use MIDI file to build PDF of sheet music
```

Figure 6: Pseudocode of System Algorithm

- In addition, we have a working version of the real-time pitch analysis functional using NumPy and MIDIUtil to simultaneously record time and frequency domain information and create MIDI files with them.
- We have automated the pitch detection to find frequency spikes and decide which musical note the pitch is closest to, using the following equation:

$$f = 440 \cdot 2^{(n-69)/12}$$

- We record the time when the note was detected to record the rhythm and save all the data points to a MIDI file.

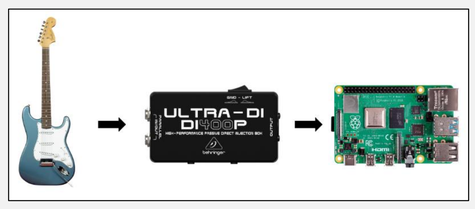

Figure 7: Hardware Setup for System

FUTURE WORK

- Algorithm tuning and scheduling, as well as completing the code for sheet music transcription.
- Create sample data and use it to train machine learning algorithm for polyphonic pitch detection.

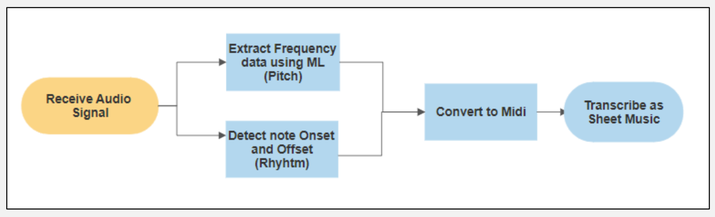

Figure 7: Flowchart of Polyphonic Audio Transcription

REFERENCES & ACKNOWLEDGMENTS

1. Nakamura, E., Benetos, E., Yoshii, K., & Dixon, S. (2018). Towards Complete Polyphonic Music Transcription: Integrating Multi-Pitch Detection and Rhythm Quantization. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). doi:10.1109/icassp.2018.8461914

Thank you to Aussie Schnore for countless inspirations and insights.