Musical Intelligence: Progress Update

MAIS 202, McGill University

Tomer Moran
October 14, 2019

1. Project Goal

The goal of this project is to develop a program capable of listening to simple melodies and extracting the musical notes being played. If possible, this program could be extended to recognize the instrument(s) and the beat of the song. For more information, see the project description file.

2. Project Progress

The first steps of this project were to develop a basic codebase to generate and play WAV file; perform Fourier transforms on audio data; encode the data from a file; and other general utilities. This was then used to train a simple Naïve Bayes classifier to recognize single notes being played on three instruments: an upright piano, a Salamander grand piano, and a Spanish classical guitar. The data was obtained from the FreePats project.

3. Preliminary Results

The models trained on individually recorded notes from the 3 instruments are excellent; the model is able to correctly classify nearly all of the test data (see Table 1). An analysis of the mistakes reveals that some parts of the data are not correctly labelled; the model effectively made the correct guess.

Table 1: NB performance on single notes

Instrument	Train Accuracy	Test Accuracy
Upright Piano	100%	99.6%
Grand Piano	99.6%	99.8%
Classical Guitar	100%	97.8%

A challenge the models have to overcome is that musical notes are often accompanied by the resonant frequencies (also called overtones, see Figure 1): integer multiples of the base frequency. As such, many mistakes made by the models are of predicting the correct note but an octave higher or lower, indicating that the model is mixing up the overtones.

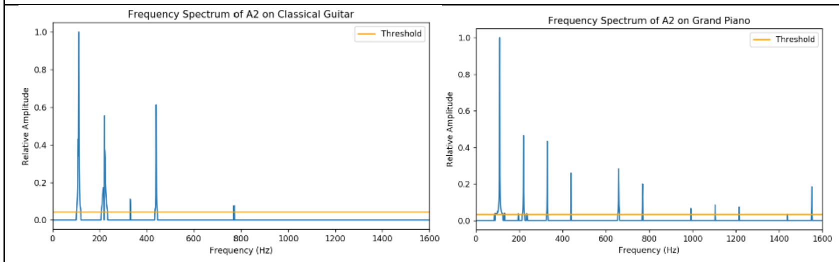
We chose statistical methods for preliminary results for their simplicity and because we suspected they would perform well. Indeed, the dimension of the features does not exceed 5000 (as frequencies over 5000 Hz are uncommon and unnecessary to distinguish notes). To reduce dimensionality, frequencies are grouped into bins, whose current width is 10 Hz. As such the problem at hand is largely to recognize the first large peak and to match it to a set of known frequency spectra. This is exactly what statistical methods excel at, hence the choice to use a Naïve Bayes classifier.

4. Next Steps

Now that the simpler toy problem is solved, we move forward and attack the next step: to recognize simple melodies on an actual instrument.

We will play simple melodies on a piano and collect the output in two different ways: through a MIDI recording, and through an audio recording. We will start by generating audio from the MIDI data and training a model on it. We expect that this will be easier than training on the real audio, as the former is clean of all perturbations while the latter will contain reverb, noise and other artefacts. To start, the melodies will only feature a single note at a time, without the pedal. If all goes well we will then try to play multiples notes at once, and perhaps even integrate chords.

Figure 1. Frequency Spectrum of A2 (110 Hz)



An A2 note played by a classical guitar (left) and grand piano (right). Observe that the grand piano has more overtones than the guitar, making it sound differently.