Violin Timbre Analysis with Mel-Frequency Cepstral Coefficients

Victor Ronchetti

University of Rochester ECE 272/472

ABSTRACT

Timbre Analysis serves to quantify the subtle color changes that make for an effective musical performance. What the human ear distinguishes easily requires some manipulation in the digital world. This paper examines the abilities of the Mel-Frequency Cepstral Coefficient (MFCC) to distinguish between the timbre of two different violins and the timbre of three different types of playing on one violin.

OBJECTIVES

The objective of this project is threefold:

• First, to demonstrate the MFCC as capable of distinguishing different instruments by timbre (As in HW 4).

• Second, to show the same method is possible for two different violins.

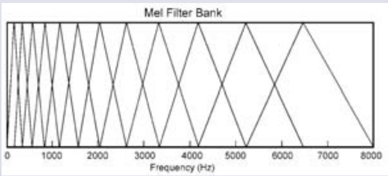
• Finally, to explore the effect of different types of violin playing on MFCC timbre calculations. Can poor playing reflect in a consistent MFCC grouping for a musical scale as compared with "correct" playing?

THEORY

• The MFCC is based upon the raw Cepstrum mathematical approach:

$$X[q] = IFFT(log|abs(FFT[x[n]])|) \quad (1)$$

• The MFCC involves filtering the magnitude spectrum through a set of overlapping triangular filters based upon the mel scale:



• To convert to/from the mel scale from the frequency scale:

$$m = 2595\log_{10}\left(1+\frac{f}{700}\right) \quad (2)$$

• The Mel filter bank more realistically resembles the real-life filtering of the human ear than the full Cepstrum spectrum.

• The magnitude spectrum in filtered, converted to the logarithmic scale, and, finally, converted to Mel-frequency Cepstral Coefficients through the Discrete Cosine Transform:

METHOD

1. MFCC

• Hamming windowed
• Window length: 2048 (~46 ms)
• sampled at 44,100 Hz
• 15 MFCCs calculated

```
function [mfcc] = my_mfcc(audio)

fs = 44100; %sampling rate
window_length = 2048; %window length
num_mfccs = 15; %number of mfccs to calculate
rows = 2048;
nfft = 4096;
offset = 45000; %offset to choose from middle of audio
f = (0:rows-1)*fs/nfft; %frequency vector

window = hamming(nfft/2); %hamming window
windowed_signal = audio(offset+1:offset+window_length).*window;

mag_spec = abs(fft(windowed_signal)); %mag spec
[mel_filter] = melfilter(num_mfccs, f);  %find filter from freq vector

mP = mel_filter * mag_spec; %filter mag spectrum with mel filterbank
mfcc = dct(log10(mP)); %db and dct

end
```

my_mfcc.m

2. SETUP

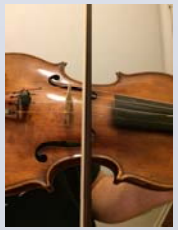
• **Test A: Compare three different Instruments**
 - Flute, Clarinet and Trumpet audio was used

• **Test B: Compare two different violins**
 - Two different scales recorded with similar vibrato and approach
 - (DEMO)

• **Test C: Compare 3 different tonal approaches on one violin**
 - (DEMO)

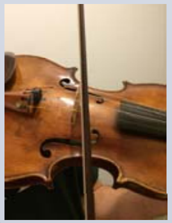
Violin Technique Background:

There is a general consensus among violin pedagogues that the proper contact point for the bow on the strings should fall roughly near the middle of the bridge and the fingerboard as shown below. The bottom two pictures represent what could be a beginners technique – closer to the bridge producing an unpleasant metallic noise and towards the fingerboard, producing a tone with less core.

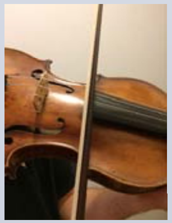


Correct

Towards Bridge

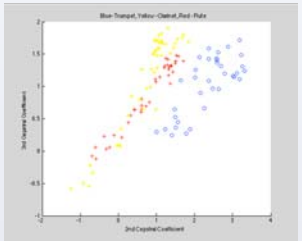


Towards Fingerboard



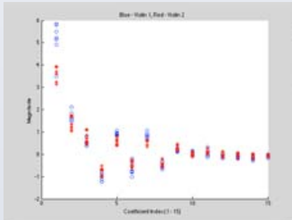
RESULTS

Test A:

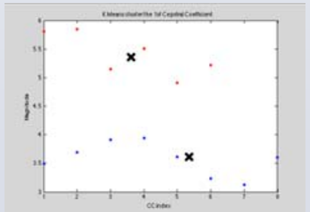


2^{nd} vs 3^{rd} MFCC

Test B:

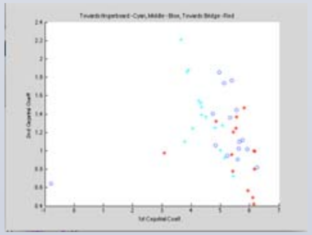


Comparing 15 MFCC Coefficients

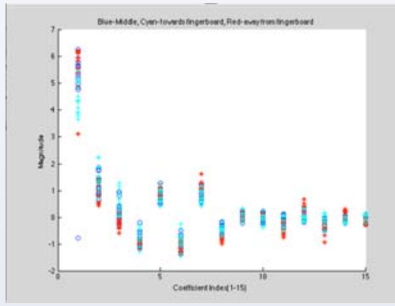


K means cluster 1^{st} MFCC

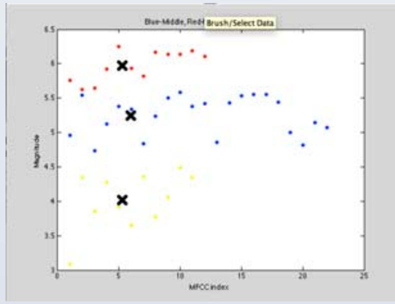
Test C:



1^{st} vs 2^{nd} MFCC



Comparing 15 MFCC Coefficients



K means cluster 1^{st} MFCC

CONCLUSIONS

• MFCC is a useful tool for detecting timbre differences for different instruments, different makes of the same instrument, and different playing techniques on the same instrument.

• K means statistics show logical groupings that should extrapolate for larger data samples.

FUTURE WORK/APPLICATIONS

• A larger sample size would always be helpful. If these models hold up to a large scale, there is merit for the MFCC as a violin pedagogical tool.

• Producing good tone in scales is an integral part of a musicians practice. Post-Analysis of a recorded scale against a database of Professionally recorded and processed scales could help the beginner student when his teacher is not around.