Violin Timbre Analysis with Mel-Frequency Cepstral Coefficients

Victor Ronchetti

ECE 272 Final Project

vronchet@u.rochester.edu

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 different instruments, different violins, and three different types of violin playing. K means clustering is used to sort the resulting data.

I. Introduction

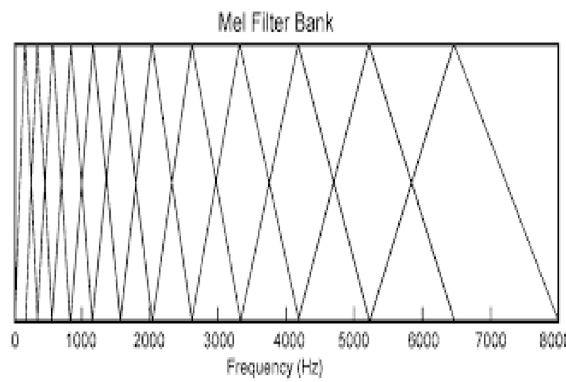
The objective of this project is threefold. First, to verify that MFCC as capable of distinguishing different instruments by timbre. Second, to show the same ability is possible for two different violins. Finally, to explore the effect of different types of violin playing on MFCC timbre calculations to see if poor technique can reflect in a consistent MFCC grouping for a musical scale as compared with correct technique.

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:

Figure 1: Mel Filterbank



The general form for coverting between the frequency and mel scale:

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

The magnitude spectrum in filtered, converted to the logarithmic scale, and, finally, converted to Mel-frequency Cepstral Coefficients through the Discrete Cosine Transform. The Mel filter bank more realistically resembles the real-life filtering of the human ear than the full Cepstrum spectrum.[1] For this reason, MFCC is used in cutting-edge speech processing and

[1] Walker, J, The Use of Mel-frequency Cepstral Coefficients in Musical Instrument Identification. University of Limerick, Ireland. 2013. pg 2

voice recognition applications. An instrument within the range of the human voice, the violin should be a good candidate for MFCC analysis.

II. Method

This project used parameters widely believed to be optimal for MFCC analysis: Hamming windowed, frame length less than 100ms, sampled at 44,100kHz, around 14 coefficients calculated.[2]

Three tests were prepared. First, three different instruments were compared using MFCC analysis: a trumpet, a clarinet and a flute. A wide range scale was used for each instrument. Second, two different violins were compared. Scales were recorded with vibrato and a *tenuto* bowing approach. Finally, one violin was recorded playing on three different locations on the string- the middle, towards the fingerboard, and towards the bridge.

Figure 3: Towards Bridge

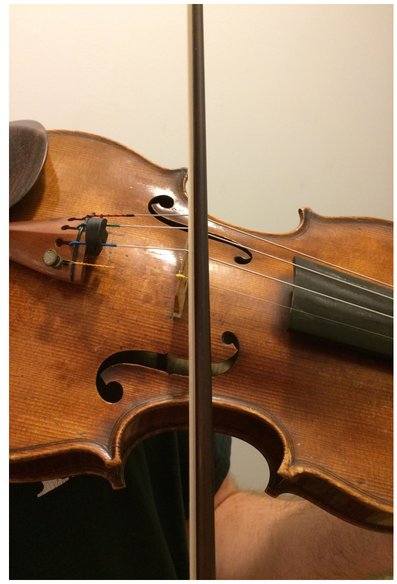


Figure 2: Correct approach

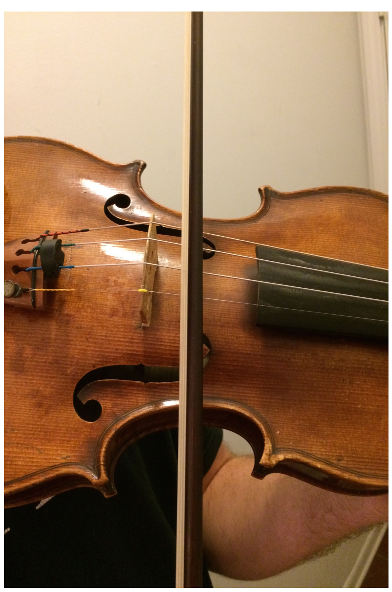


Figure 4: Towards Fingerboard

[2]Lukasik, E. Long Term Cepstral Coefficients for violin identification. Poznan University of Technology, Institute of Computing Science, Poznan. 2010. pg 1

III. Results

Figure 5: ***2nd vs. 3rd MFCC***: *Blue - Trumpet, Yellow - Clarinet, Red - Flute*

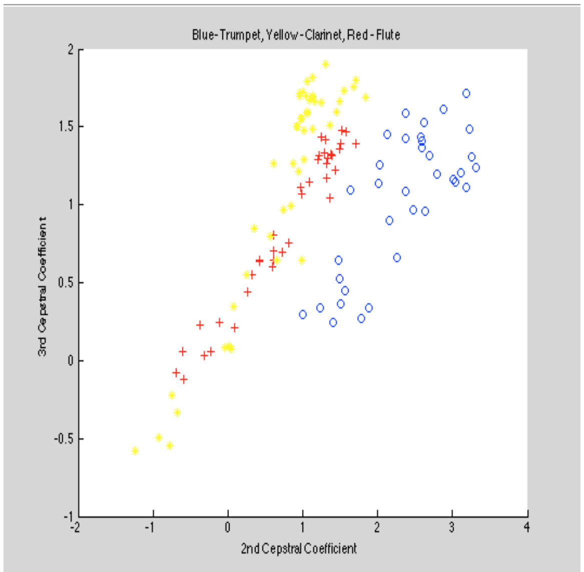


Figure 7: ***K means clustering MFCC 1st*** *blue- Violin 1, red - Violin 2*

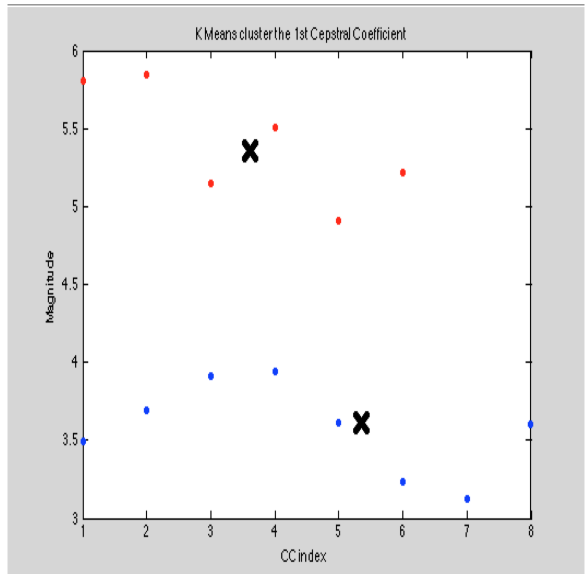


Figure 6: ***15 MFCC*** *blue- Violin 1, red - Violin 2*

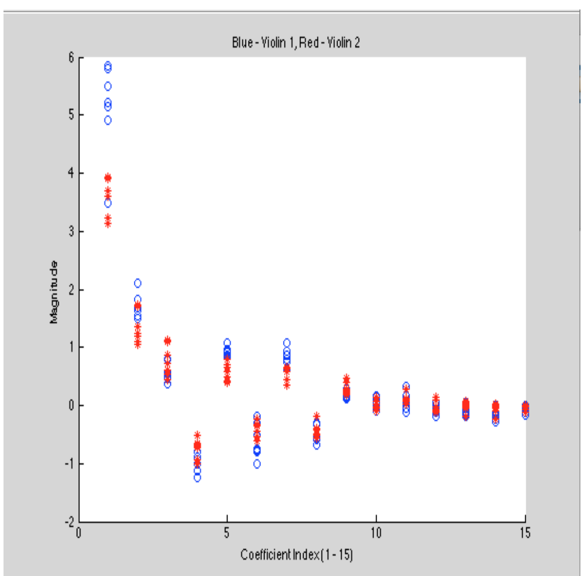


Figure 8: ***15 MFCC*** *blue- middle, red - towards bridge, cyan - towards fingerboard*

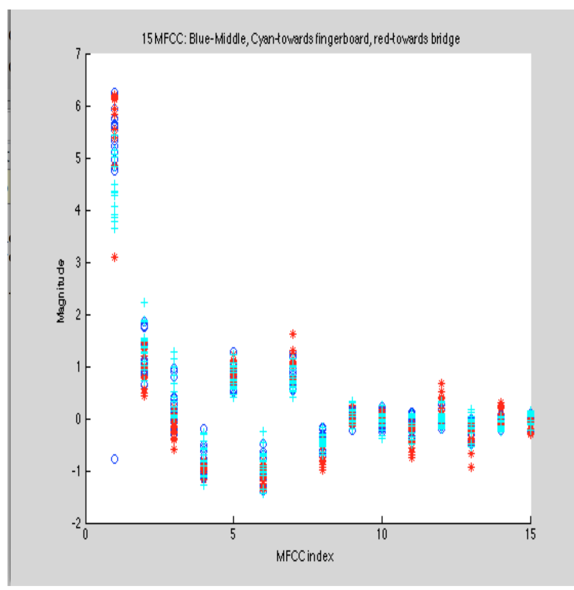
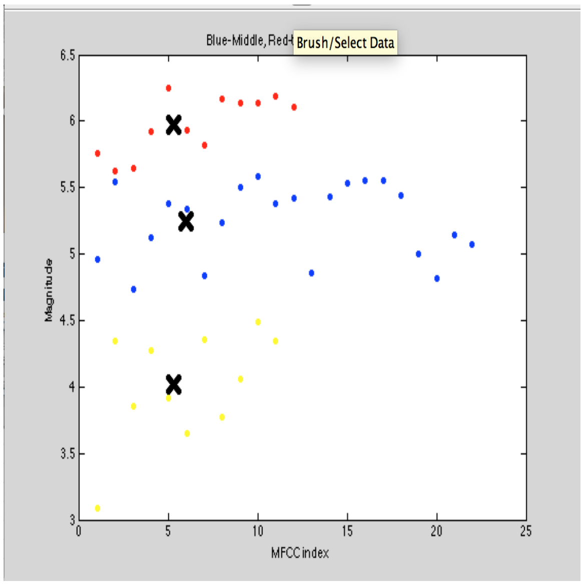


Figure 9: ***K means clustering MFCC 1st*** *blue-middle red - towards bridge yellow - away from bridge*



IV. Discussion

The MFCC was a useful tool when comparing different instruments by timbre. As Figure 5 shows, the three instruments occupy almost their own area of the plot. It was expected that measurements of the same type of instrument (two different violins) would be less distinct than the first test. However, certain characteristics consistently occurred. Figure 6 shows each of the 15 MFCC coefficients for every note recorded on the two violins. At the coefficient index gets higher, the two values blend, but the first seven coefficients indicate distinct differences - particularly the first coefficient. K means clustering partitions data into distinct groups, called clusters. When a two-cluster, k means cluster algorithm is performed on the 2-violin 1st MFCC data (Figure 7), the two instruments partition virtually cleanly. Only 1 data point is incorrectly grouped.

Analyzing the timbre differences of different technique on one violin proved even more subtle than Test 2. As Figure 8 demonstrates, the MFCCs of the 3 playing techniques exhibit significant overlap. However, the first few CC's appear to show some distinction. Once again K means clustering confirms this. Figure 9 shows three bands of a data with three centroids marking the mean of the clusters. Here, 4 data points (out of 18) have been incorrectly clustered.

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. Future work could include a real time MFCC analysis so the playing would not have to be post-processed.

References

[1] Jacqueline Walker, *The Use of Mel-frequency Cepstral Coefficients in Musical Instrument Identification*, Science Federation Ireland, 2013.

[2] Ewa Lukasik, *Long Term Cepstral Coefficients for violin identification*, Poznan University of Technology, 2010.

[3] Jane Charles, *Playing Technique and Violin Timbre:Detecting Bad Playing*, Dublin Institute of Technology, 2010.

[4] William Brent, *Cepstral Analysis Tools for Percussive Timbre Identification*, University of California, San Diego 2010.