Timbre Tuning: Variation in Cello Spectrum Across Pitches and Instruments

Robert Pond
rpond@uvic.ca

Alex Klassen
heardxal@uvic.ca

Kirk McNally
kmcnally@uvic.ca

University of Victoria, 3800 Finnerty Road, Victoria, BC V8P 5C2

ABSTRACT

The process of learning to play a string instrument is a notoriously difficult task. A new student to the instrument is faced with mastering multiple, interconnected physical movements in order to become a skillful player. In their development, one measure of a players quality is their tone, which is the result of the combination of the physical characteristics of the instrument and their technique in playing it. This paper describes preliminary research into creating an intuitive, real-time device for evaluating the quality of tone generation on the cello: a "timbre-tuner" to aid cellists evaluate their tone quality. Data for the study was collected from six post-secondary music students, consisting of recordings of scales covering the entire range of the cello. Comprehensive spectral audio analysis was performed on the data set in order to evaluate features suitable to describe tone quality. An inverse relationship was found between the harmonic centroid and pitch played, which became more pronounced when restricted to the A string. In addition, a model for predicting the harmonic centroid at different pitches on the A string was created. Results from informal listening tests support the use of the harmonic centroid as an appropriate measure for tone quality.

CCS Concepts

•Applied computing → Sound and music computing; Performing arts;

1. INTRODUCTION

A difficulty with teaching tone quality to novice cellists is that the language used to describe it is subjective, and the student may not have sufficient content knowledge to self evaluate their tone and determine when it is good, or poor. Where pitch accuracy can be measured by even the beginner cellist, through the use of the ubiquitous digital tuner, a novice cellist may has few resources to turn to when approaching questions of tone generation. For instance, in a recent book meant for autodidactic cellists, the following passage "instructs" the reader on how to find the ideal contact point between bow and string:"optimal placement is approximately three-quarters of an inch to an inch from the bridge ... It produces a sound that is most characteristic of the cello, a sound that engages the emotions of the listener to the often voluptuous musical expressions given to the cellist by the composer" [1]. The danger in such statements is that they rely both on the student's technical skill and their ability to be discerning and objective listeners. In contrast, a passage in the following paragraph discusses how to perform an amplitude change: "An increase of pressure correlated with an increase of speed is the usual way to crescendo on a single note. Actively changing the placement by moving the bow closer to the bridge gradually throughout the stroke assists in the crescendo". The passage is much more concrete in describing the mechanics of the technique, implicitly contradicting the previous passage by directing a movement away from the "optimal placement" of the bow.

1.1 Timbre as a Teaching Aid

In order to identify timbre features that could aid in helping cellists develop their own sound, it is necessary to know which features of the cello's timbre are critical to its quality, and which offer some degree of flexibility. When cello sound is studied for its qualities, it is often done in conjunction with other instruments. This was especially true for early research on measuring timbre, which examined cello sound within the context of the entire orchestra ([9], [8]). More recent research has focused on the specific mechanisms of the cello, determining subtle differences between it and other string instruments ([11], [12]). The most extensive work on cello timbre alone has been done by Magdalena Chudy, who identified various low-level features as being instrument, or player-specific, and then attempted to relate specific performance gestures to those features [6].

The problem of using timbre analysis as a pedagogical tool has been examined on the violin through the work of Jane Charles [4]. She identified several useful features in distinguishing amateur and professional violinists, most of which could readily be applied to the cello. A majority of these features were found to correspond to faults in the fundamentals of bow control. However, Charles' proposed teaching aid was explicitly designed not to give real-time feedback, so that the student could focus entirely on playing their instrument [5]. To avoid duplication, our focus is on cellists who have fulfilled the basics of tone generation and wish to develop their own sound further. Additionally, we choose to work towards a real-time teaching aid due to the perceived subtlety of the movements required to create a consistent tone across the range of the cello versus detecting fundamental mistakes.

2. METHODS

2.1 Recording Procedure

Six student cellists were recorded to create the study data set. All players were performance majors from a post-secondary music program, three being undergraduate students and three graduate students. All students performed

NIME'18, June 3-6, 2018, Blacksburg, Virginia, USA.

on their own instruments for the recordings, and cellos were tuned to A = 440 Hz. The recording was performed in a small room with some acoustic treatment, resulting in almost no echo. An AKG C414 XLS condenser microphone with cardioid polar pattern was used and placed approximately 20 centimeters from one of the f-holes. Each player's scales were recorded with a Zoom H6 Handy Recorder and saved as monophonic 24-bit, 96 kHz WAV files.

Each student was asked to play two four octave major scales, beginning and ending on the lowest available D (D2, $f \approx 73.4$ Hz). The students were told to play each scale slowly and focus on tone generation over pitch accuracy. The students were allowed to use vibrato as desired. Most of the students were instructed to play more slowly when playing the scale a second time. In addition, the players were instructed to play a one octave major scale beginning and ending on D3 as slowly as possible while remaining comfortable. In total, 761 individual notes were recorded. One student played a three octave scale instead of one of the four octave scales due to inexperience, while another student repeated the highest note on both four octave scales.

2.2 Analysis

The recordings were analyzed using a variety of tools in the Essentia analysis library [2]. Note onsets in each scale were marked by hand using both the waveform and spectrogram data in Sonic Visualizer [3] as guides. Some ambiguity between notes occurred when the cellists changed strings due to the previous string resonating. In addition, the last note of each scale was cut off at the point where the upper partials of the note had noticeably decayed on the spectrogram in an attempt to prevent the natural decay of the cello's resonance from altering the results. Most of the analysis techniques used, including a range of spectral measures and MFCCs, were informed by Charles' prior research on the subject. To ensure that any useful measure we found would be computable in real time, we used a relatively small FFT size of 1024 samples (11 ms) for all analysis.

Following the analysis of multiple spectral and cepstral features, we opted to focus on the spectral centroid of each note. Using the *pandas* library in Python [10], the harmonic centroid (HC) was determined for individual notes by dividing the spectral centroid over time by the expected fundamental frequency, then averaging all of the data points. The HC is a dimensionless number that represents the partial content of a pitched sound, with one representing a sine wave and higher numbers indicating louder partials relative to the fundamental. This metric was chosen to ease comparisons between the spectra of different notes. It has been shown that the HC can predict the clarity of string instruments in a listening setting [7].

3. RESULTS

Once an average HC was determined for each note, all instances of each pitch were averaged together. The results are shown in Figure 1. While there is a downward trend in the HC over the entire four octaves, the pitch range from D2 to A3 shows no clear pattern. Careful examination of instances of notes below A3 reveal trends that could explain the values for some pitches but not others. For instance, the extremely high value observed for D2 was found to be due to the fundamental and lower partials being much softer. In several instances, the seventh harmonic was substantially louder than all six harmonics below it. This was mostly likely due to the larger amount of energy required to activate the lower partials. While we anticipated that the pitches of the open strings (G2, D3, A3) would be noticeably different from the surrounding pitches, it was surprising to observe that A2 had a higher HC than G2.

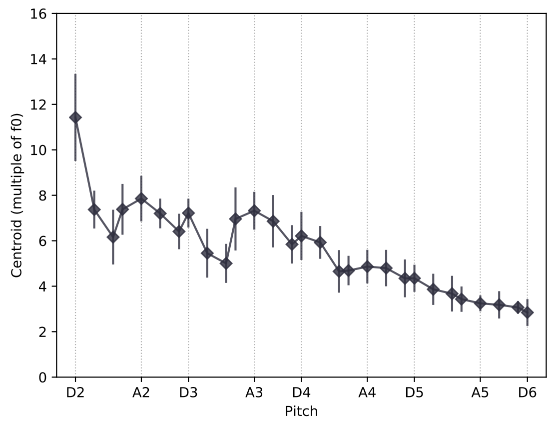

Figure 1: Average harmonic centroid per pitch over the entire four octave range. Error bars represent the standard deviation.

Another observation was made when looking at the HC over time for all notes of one pitch. While most cello notes showed fairly consistent HCs over time after applying a smoothing filter, some of the instances of the D2 and D3 pitches either started with substantially higher values or ended with sudden changes in value. Figure 2 shows the smoothed HC for all players graphed over time. The anomalies were determined to be due to a combination of two factors. First, some of the instances were either the beginning or end of entire scales the cellists were asked to play, meaning the attack and decay portions of those particular instances were not masked by notes played before or after. Thus, these instances would be very quiet at the beginning or end, allowing background noise to substantially influence the value. Comparison to instances of A3, another pitch on an open string, showed that this explained the notes which started or ended with substantially lower HCs. Both of these cases could be compensated for by removing some of the data points at both the beginning and ending of each note, which would allow us to focus in on the notes' fully excited state. However, we found that removing the first and last 10 percent of each note decreased the HC of all of the pitches almost uniformly by three percent, whether or not they were affected by the anomalies listed above.

3.1 Results on the A String

While no significant trend can be seen for the HC over the pitches below A3, the HC on all of the notes played on the A string were almost uniformly decreasing (see Figure 3). An adequate explanation could not be determined for why three of the pitches- C♯4, F♯4, and G4- observed HCs that fell below the general trend. When the notes were separated by cellist and separated by pitch, the harmonic spectrum for each cellist was less uniform but still decreased with pitch. With the exception of one cellist, one of the three pitches highlighted above observed a significant decrease in HC from the previous note, the remaining cellist had an exceptionally low HC on the B3 pitch instead.

After experimenting with a few different models, we found that a power regression over frequency (i.e $f(x) = ax^r$) with a negative exponent, consistently outperformed other models in describing the average HC across all cellists. Obtaining a coefficient of determination $R^2 \geq 0.96$ was done

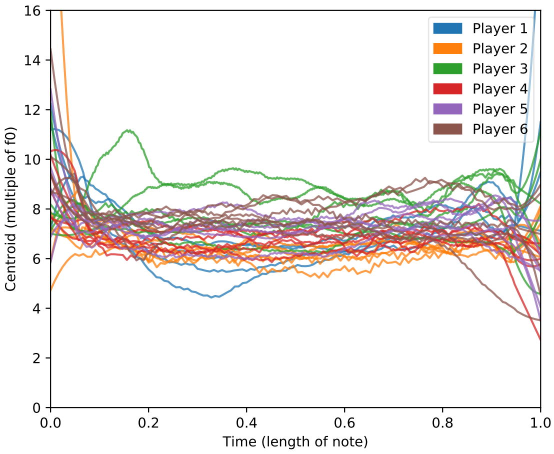

Figure 2: Smoothed harmonic centroids over time for all instances of D3, colored by player.

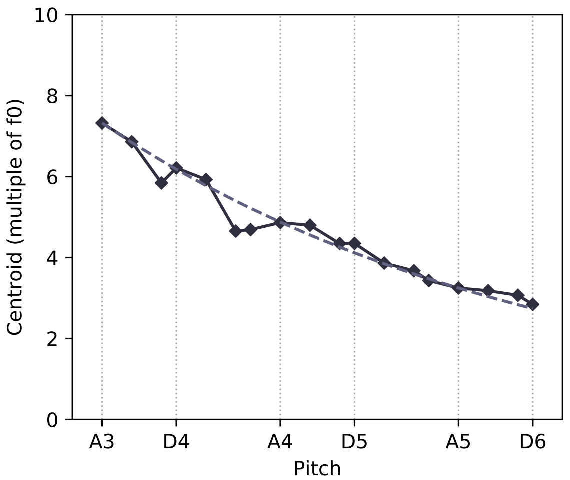

Figure 3: Average harmonic centroid per pitch on the A string for all players, with power regression.

through trial and error. Once this was established, we examined the data to see if we could obtain a useful model for each individual cellist from the actual values at the open string (A3) and one other pitch. If so, then one could theoretically model the ideal spectral content of pitches on the A string of any cello by taking just two samples from it. Ultimately, we achieved the best results by solving $f(x) = ax^r$ with $f(220)$ and $f(880)$ being the observed HCs on the A3 and A5 pitches respectively. This model achieved $R^2 \approx 0.955$ for the entire dataset, with $a \approx 173.1$ and $r \approx -0.586$, and is denoted by the dashed curve in Figure 3. Models and coefficients of determination for each individual cellist are shown in Figure 4. These models consistently outperformed naive linear regressions models, obtained by using the values at the lowest and highest pitches on the A string, both over frequency and MIDI values.

4. DISCUSSION

In order to evaluate the legitimacy of using HC as an indicator of timbre quality, we performed informal listening tests. For each pitch in the dataset, we identified the notes with

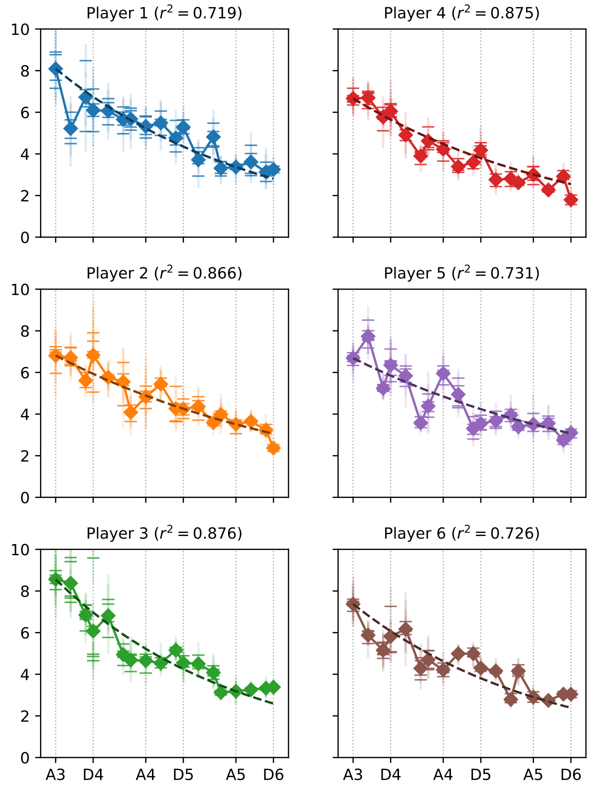

Figure 4: Power regressions for the average harmonic centroid of each cellist. Dashes indicate observed harmonic centroids at each pitch.

the highest and lowest HCs. In addition, we identified one or two notes that were closest to the average HC for each pitch. When grouped together, the note with the highest HC was judged to sound weaker than the other notes, with very few exceptions. However, the note with the lowest HC was considered to be higher quality than the notes closest to the average HC for a significant number of pitches. These listening tests highlight the fact that the HC of a sound can be increased in two different ways: either increasing the amplitudes of higher partials, or decreasing the amplitude of the fundamental. After normalizing to a constant loudness, the notes with the highest HC almost always had quieter fundamental frequencies than any of the other notes, resulting in tones that sounded weak or “nasally”.

On the other hand, the same effect occasionally affected the notes closest to the average HC as well. In some cases where the note with the lowest HC was perceived to be higher quality, it was found that the lowest note actually had more upper partials than one of the aforementioned notes, these notes landing closer to the average because its fundamental was relatively weak. While this was not studied in depth, using the amplitude of the fundamental frequency relative to the other harmonics in conjunction with the HC could produce a reliable ranking of tone quality for all notes.

One clear issue with the analysis was that the fundamental frequency was assumed instead of measured for all of the notes. This became evident when listening to some of the groups of notes described above, where occasionally one or more notes would be audibly sharp or flat. We attempted to use pitch detection algorithms in both Sonic Visualizer and Essentia to determine the actual fundamental frequencies, but the algorithms could not detect this pitch for the lowest notes of the scale, even when larger window sizes were used. In addition, the small window size we used for the analysis produced very chaotic results. This was to be expected, since the lowest note the cellists played has a larger period than the window length. To judge the effect of window size on the models, we recalculated the spectral centroids for all of the cellists using a window size of 4096 samples and hop size of 1024 samples. As expected, this dramatically decreased the variance of the HC calculations for pitches below G3. While R^2 for the power regressions increased for all of the players, the gains were insignificant except for Player 6 (for whom it increased from 0.726 to 0.769).

5. CONCLUSION

The HC appears to be a promising metric for tone quality. Most critically, it can be used to create a model for ideal cello timbre on the most commonly used string that can then be calibrated to individual instruments and players. Further research is required to determine whether the model can be extended to the other three strings of the cello.

Given that the spectral centroid and fundamental frequency of a sound can be computed relatively quickly, it is theoretically possible to compute the HC of cello playing in real time. Future work is necessary to investigate whether cellists could effectively use a "timbre tuner" based on the HC to refine the tone of their playing. Further, our analysis points to pitches along the string that consistently deviate from the best fit. We believe that this might be explained by the modes of each cello or relative strengths of the fingers on each cellist's left hand. A study requiring each cellist to play on the same instrument might reveal whether these outliers are more likely to be instrument or player-specific, and therefore whether they should be accounted for in order to create a more accurate model of desired cello timbre.

6. REFERENCES

[1] E. Benedetti. *Cello, Bow and You: Putting it All Together*. Oxford University Press, Oxford, UK, 2017.

[2] D. Bogdanov, N. Wack, E. Gómez, S. Gulati, P. Herrera, O. Mayor, G. Roma, J. Salamon, J. Zapata, and X. Serra. Essentia: an open-source library for sound and music analysis. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 855–858. ACM, 2013.

[3] C. Cannam, C. Landone, and M. Sandler. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the ACM Multimedia 2010 International Conference*, pages 1467–1468, October 2010.

[4] J. A. Charles. *Playing Technique and Violin Timbre: Detecting Bad Playing*. PhD thesis, Dublin Institute of Technology, 2010.

[5] J. A. Charles et al. Development of a computer based violin teaching aid, vitool. In *118th Convention of the Audio Engineering Society*. AES, May 2005.

[6] M. Chudy et al. On the relation between gesture, tone production and perception in classical cello performance. In *Proceedings of Meetings on Acoustics*, volume 19. ICA, 2013.

[7] K. Hermes et al. The harmonic centroid as a predictor of string instrument timbral clarity. In *140th Convention of the Audio Engineering Society*. AES, May 2016.

[8] D. Luce and M. Clark. Duration of attack transients of nonpercussive orchestral instruments. *Journal of the Audio Engineering Society*, 13(3):194–199, July 1965.

[9] D. A. Luce. Dynamic spectrum changes of orchestral instruments. *Journal of the Audio Engineering Society*, 23(7):565–568, September 1975.

[10] W. McKinney. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.

[11] J. Woodhouse and A. R. Loach. Torsional behaviour of cello strings. *Acta Acustica united with Acustica*, 85(5):734–740, September 1999.

[12] A. Zhang, J. Woodhouse, and G. Stoppani. Motion of the cello bridge. *Journal of the Acoustical Society of America*, 140(4):2636–2645, October 2016.