

Analyzing Timbres of Various Musical Instruments Using FFT and Spectral Analysis

Ellen Jannereth¹ and Lisa Esch¹

¹Carrollwood Day School, Tampa, FL, USA

ABSTRACT

A sound's unique timbre is based on the various harmonic frequencies present within its waveform. Through Fast Fourier Transform software, waveforms can be easily decomposed into their component frequencies and a spectral analysis of frequency can be conducted as a method of quantitatively describing timbral characteristics of a sound. In this investigation, the range of frequencies present in a spectrum as well as the average intensity of the first 10 overtones in a sound will be used to classify the timbres of various instruments relative to one another. This will be done by generating a Range-Intensity graph of harmonic frequencies present in sound samples of each instrument.

The results of this investigation reveal that it is not only possible to quantitatively analyze instrumental timbre by generating and mapping out the harmonic frequency data of a specific sound, but that such a quantitative analysis is also incredibly useful. Unlike the traditional, qualitative method of describing timbre, a quantitative analysis would allow for timbral qualities to be transformed into information that can be understood by computers. Today, timbral classification and the decomposition of waveforms has many applications in science and sound engineering. By refining methods for quantitative timbral analysis, it becomes possible to further enhance timbre recognition software and apply such methods to a wider range of technological developments.

Background

While rhythm, beats, and pitch are musical terms that are typically described quantitatively, timbre is an aspect of music often associated with a qualitative description. For instance, the sound that a violin produces varies greatly from that of a clarinet, even if played at the same pitch and volume. However, by analyzing the different frequencies present in a signal, a quantitative approach to examining timbre not only becomes possible, but also incredibly practical in understanding and manipulating various sounds.

An instrument's unique timbre is generated by the overtones present in the sound that it produces. Each note played on an instrument has a specific fundamental frequency, which is the lowest frequency present in a signal. On top of this fundamental frequency are layers of harmonic frequencies, which are integer multiples of the fundamental. This can be visualized in Figure 1.

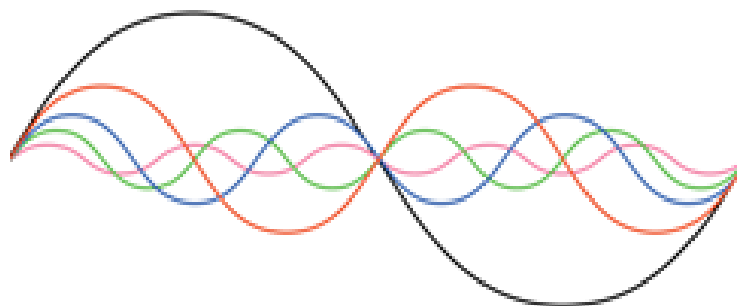


Figure 1: Fundamental and harmonic frequencies. The fundamental frequency is in black. (“Harmonics”).

Essentially, a note played on an instrument is comprised of not just one, but a series of different frequencies of sine waves that combine to form a unique waveform, as illustrated in Figure 2.

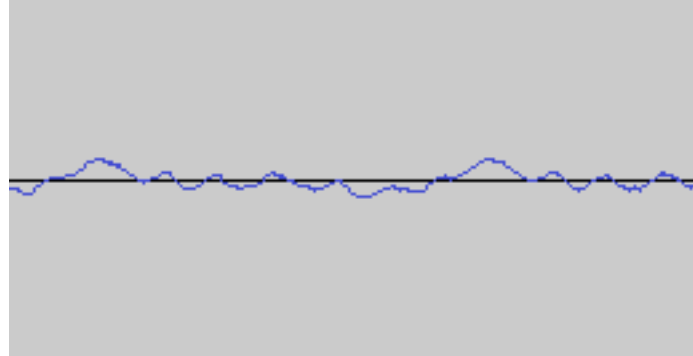


Figure 2: Waveform of a violin (A=440 Hz), a result of the summation of various sine waves of different frequencies present in the sound.

It is this summation of different frequencies of sine waves within a waveform that lends to a sound’s distinct timbre as perceived by the human ear. To analyze timbre, one can examine the frequency spectrum of a sound wave, which displays the harmonic frequencies that are present in a sound and the corresponding intensity of each frequency in decibels. From this spectrum, the most dominant frequencies in a specific waveform can be determined and compared to other waveforms, producing quantitative data that can be used to describe the timbres of various instruments.

To generate a harmonic spectrum of a waveform, data must be transformed from the time domain of the sound wave to the frequency domain. This will essentially decompose the sound wave into its elemental harmonic frequencies and can be accomplished using the Forward Fourier Transform (FFT). Much of the mathematics behind the Fourier transform is quite complex and extends beyond the scope of this investigation, however, a brief overview of the Fourier transform will be discussed before addressing the Fast Fourier Transform, an algorithm used in computer programs that automatically generates frequency spectrums from the data present in a waveform.

To generate a harmonic spectrum, calculations must be conducted to determine how much of a certain frequency is present in a given time domain (“Introduction to Fourier Transform”). The Forward Fourier Transform is represented in the complex plane, where sinusoids of any frequency within the time domain can be expressed in terms of sine and cosine functions using Euler’s Formula. In the complex plane, the amplitude of the sinusoid is the modulus, and its phase (θ) as measured from the positive real axis is the argument. Essentially, the sinusoid of a frequency in the time domain of a sound wave can be expressed as $e^{i2\pi\nu t}$, where ν is the frequency (“Frequency Domain and Fourier Transforms”). The following represents the sinusoids of the frequencies present in a signal within a given time domain:

$$x(t)e^{-i2\pi\nu t}$$

Summing all frequencies across all time:

$$x(\nu) = \int_{-\infty}^{\infty} x(t) e^{-i2\pi\nu t} dt$$

The result is the Forward Fourier Transform (“Frequency Domain”).

It is important to note that although the integral goes from $-\infty$ to ∞ , the data collected from a signal will always be contained within a finite time domain (Jacobsen). The Forward Fourier Transform produces discrete frequency data from a finite sample of time. In this investigation, discrete samples of sound waves will be analyzed. To do so, the Discrete Fourier Transform (shown below) becomes more ideal.

$$x_k = \sum_{n=1}^{N-1} x_n e^{-\frac{2\pi kn}{N}}$$

This replaces the integral from $-\infty$ to ∞ with a summation from the first sample n to the $N-1$ sample of data. k replaces v as the frequencies present in a signal are now confined to the samples taken within the finite time domain (“Frequency Domain”).

Fast Fourier Transform (FFT)

The Fourier Transform is incredibly complex and tedious to carry out by hand. For this reason, algorithms have been created to make the application of the Fourier Transform significantly easier and faster. Implemented in countless areas of study and modern technologies, the Fast Fourier Transform (FFT) algorithm essentially computes the same results as the Discrete Fourier Transform (DFT), transforming data in a signal from the time domain to the frequency domain (“FFT Waveform Analysis”). The software that will be used in this investigation to obtain the frequency spectrums of various instruments is Audacity® (“Audacity Team”), which implements the FFT algorithm to compute frequency data from a finite sample of sound, generating a frequency spectrum. This frequency spectrum plot will allow for the analysis of timbre as it clearly displays the constituent frequencies present in a given sound sample. In this investigation, Audacity’s FFT software will be used to generate these frequency spectrums in the logarithmic scale.

This investigation aims to use this FFT software as a method of quantitatively analyzing the timbre of a musical instrument. The implications of such a quantitative analysis of timbre are extensive. Unlike the traditional, qualitative method of describing timbre, a quantitative analysis would allow for timbral qualities to be transformed into information that can be understood by computers. In this way, timbral analysis can be automated and applied more effectively in sound engineering and music producing technologies. Using FFT software, the timbre of a specific instrument can be digitally recreated by combining digital signals of the various frequencies present in an instrument’s waveform. Spectral analysis techniques are widely used in sound engineering, where frequencies can be eliminated, altered, or isolated in order to manipulate the timbre of a specific sound. Timbre classification software through FFT is also used in song recognition technologies and applications such as Shazam (Jovanovic). Furthermore, quantitative descriptions of timbre can become especially useful in the development of technologies for the deaf or hard of hearing. Timbre identification through FFT and spectral analysis software can be used to convert timbral elements into alternative sensory signals that can allow an individual to understand the nature of a sound without hearing it directly.

Investigation

In this investigation, the timbres of the sounds produced by a violin, a ukulele, a piano, a flute, and a French horn will be analyzed. These instruments were chosen based on a traditional, qualitative analysis of timbre, as each sound distinct from one another. The aim was to choose instruments with a wide range of timbral qualities so as to investigate whether or not a quantitative analysis would effectively reflect the qualitative timbral differences. For instance, one could observe, using qualitative descriptors, that a violin sounds much “brighter” than a French horn. By choosing these instruments, the investigation will reveal whether or not this qualitative difference in timbre will be detected in a quantitative analysis. It is important to note that the perception of timbre is a largely subjective experience. Moreover,

instruments of the same type can often have slightly different timbres, and the timbral qualities of a single instrument can also change depending on variables like playing technique. This investigation will merely focus on examining the timbral differences of the above instruments relative to one another. The sound samples used to generate each instrument's frequency spectrum were intended to represent the most "typical" sounds of each type of instrument. To achieve this, no pitch-augmenting or ornamentation techniques such as vibrato were used in any of the samples. Recordings of these samples were taken and inputted into Audacity. The pitch A=440 Hz was used for all instrument samples. The intensity of each instrument sound was not kept completely constant throughout this process, which may have affected the results of this investigation. For each instrument, the entire recorded waveform was selected and run through Audacity's FFT software, generating frequency spectrums for each instrument. Below are the spectrums generated by Audacity's FFT algorithm, where the x-axis displays the frequencies present, and the y-axis displays the intensity of the frequencies in decibels. Each frequency spectrum displays all overtones present within the entire duration of the sound sample. The peaks indicate the most prominent harmonic frequencies present in each instrument.

Firstly, we notice that each spectrum has a large peak (ranging from -28 to -36 dB) at the fundamental frequency (440 Hz). The differences in shape and spread of each spectrum is also quite noticeable, which explains the distinct timbres of each instrument. From the graphs, we can see that the violin, piano, and flute also have prominent peaks at the first harmonic (880 Hz). In the flute, the first harmonic appears to be even more intense than the fundamental. This greater intensity of higher harmonics may be attributed to the brighter sounding timbres of these instruments. In the violin and ukulele for instance, we can see that there is a much greater presence of higher harmonic frequencies than can be found in the other instruments. The spectrums for the violin and the ukulele are similar in range and intensity, suggesting similar timbres. This may be due to the comparable shapes of both instruments, as well as the fact that they are both chordophones. On the other hand, the French horn has the smallest range of harmonics present (from 200 Hz to slightly less than 8000 Hz). This could explain its smooth timbre, a characteristic that allows its sound to blend easily with other instruments.

We can begin to organize the timbres of each instrument based on two main characteristics displayed in the harmonic spectrums: the spread of the harmonic frequencies present and the intensity of higher harmonic frequencies. For this investigation, the average intensity of the first 10 overtones (frequency peaks higher than the fundamental frequency) will be examined. Each peak can be regarded as a data value point that expresses frequency and the corresponding intensity. We will begin by calculating the ranges (R) of each harmonic spectrum as follows:

$$R = f_{max} - f_{min}$$

where the maximum frequency is the highest recorded peak, and the minimum frequency is the lowest recorded peak. If subharmonic frequencies, or frequencies below the fundamental, are present in a sample, those will be considered to be the minimum frequency. These frequency values were obtained more precisely through the spectrum graphing program displayed directly on Audacity, as opposed to the graphs shown in Fig. 3 of this paper. By running the cursor over any given peak, the corresponding frequency and intensity at that location was displayed to the nearest Hz. The frequency ranges for each instrument sample are displayed in Appendix A. The data for the intensities of each peak were obtained for the first 10 overtones and the averages of these first 10 overtones were calculated, allowing for a quantitative assessment of the brightness of timbres (more intense higher harmonics would indicate a brighter timbre). Appendix B displays the average intensities of the first 10 overtones for each instrument.

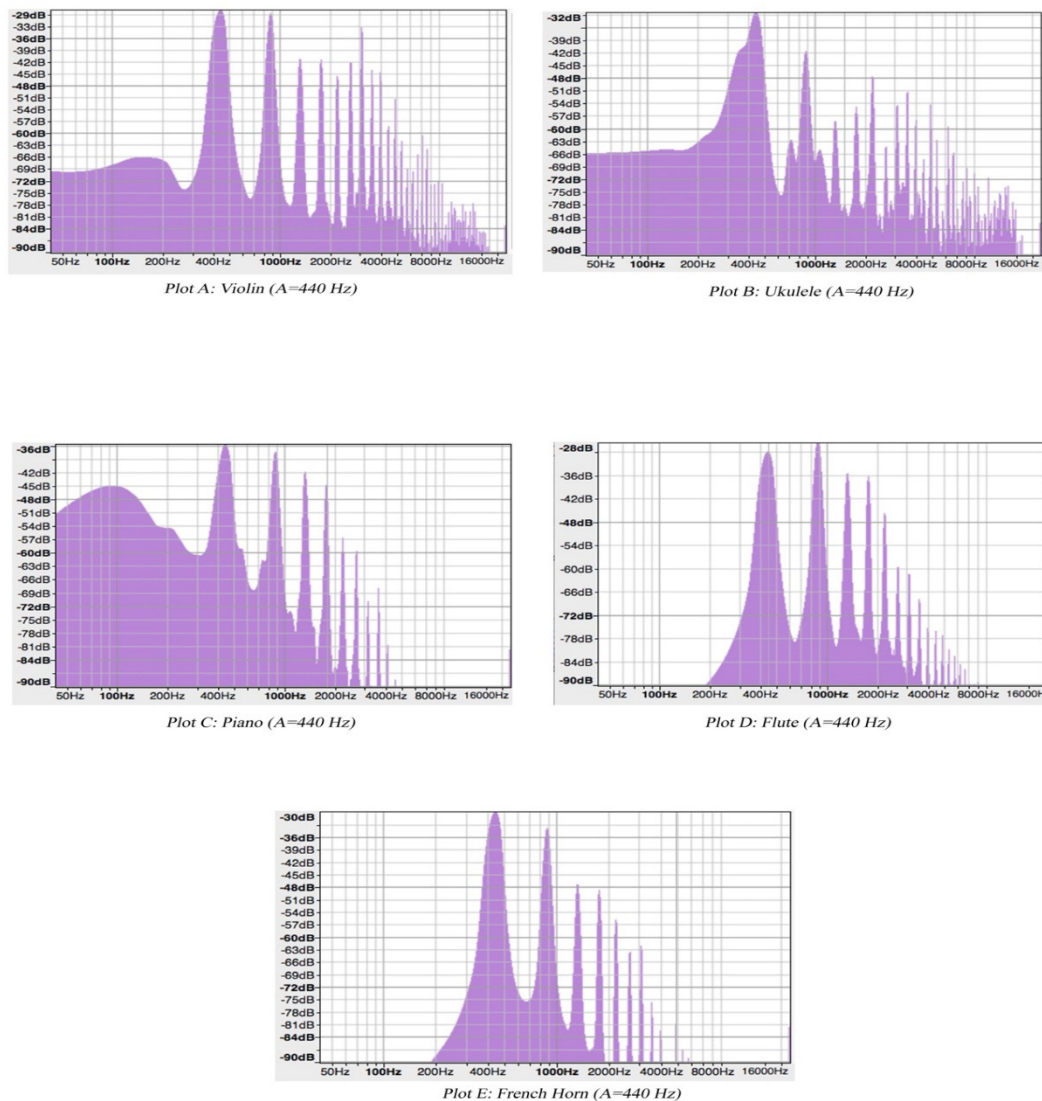
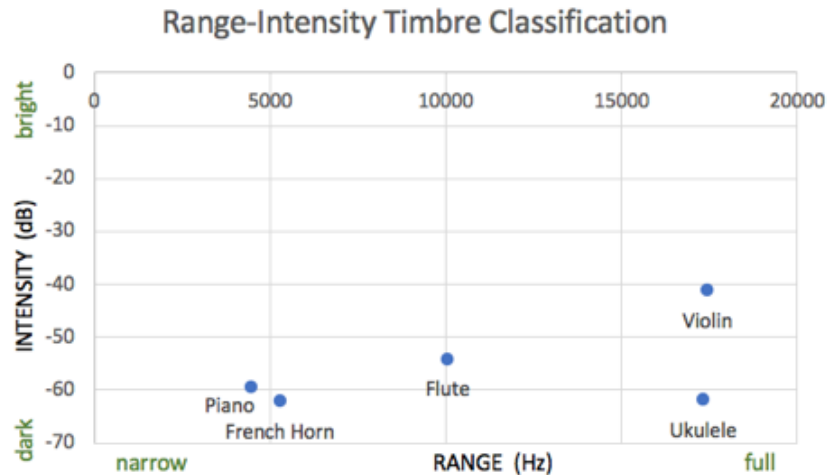


Figure 3: Frequency spectrums of each instrument generated from Audacity’s FFT software. These will be referenced in the analysis below.

In music, there are many terms in the English language that can be used to describe the timbral qualities of a sound. Using language to illustrate timbre is a largely subjective process, as each individual perceives sounds differently and often associates sounds with different descriptors. The “proper” words that should be used to describe timbre continue to be debated in the world of musical semantics. This investigation attempts to link quantitative analyses of timbre to qualitative descriptions in a way that allows for the quantitative description of the timbral qualities of an instrument relative to other instruments. According to a study conducted at the Academy of Performing Arts in Prague, the words “dark,” “bright,” “narrow,” and “full” were among the words considered to be the most “suitable” in describing timbre (O. Moravec, J. Stepanek). These words will be used in this paper as a reference point in describing timbre based on the quantities obtained for range and intensity from the frequency spectrums. Instruments with greater average overtone intensities will be considered “brighter,” while those with smaller overtone intensities will

considered “darker.” Similarly, greater range of harmonic frequencies displayed in a harmonic spectrum would indicate a “full” sounding timbre, whereas a smaller range of frequencies would imply a “narrower” timbral quality. It is important, however, to acknowledge that these words themselves are largely subjective and can easily be substituted with other descriptors depending on one’s personal perception of a sound. These terms will be used in this paper only in an effort to link the quantitative qualities of a sound (obtained through spectral analysis) to familiar psychological perceptions in order to compare the timbres of various instruments with respect to one another.

Below is a graphical representation of the timbres of different instruments based on the range of frequencies present in each harmonic spectrum, as well as the intensities of the first 10 overtones in each sound.



Note: Error bars are present but are not visible due to the scale of the graph.

Graph 1: Range-Intensity Timbre Classification.

In this way, the data collected through the FFT software helps to create a clear visualization of the timbral differences of each instrument. The violin, with a broad spectral range of approximately 17000 Hz and relatively intense higher harmonics (a -41.06 dB average intensity over the first 10 overtones) is shown to have a full and bright timbre in comparison to the other instruments. The “clear” sounds of the piano and French horn are reflected in their “narrow” and “dark” timbres as a result of their relatively low ranges and overtone frequency intensities.

We can also compare our results in the range-intensity graph to the spectrograms of each instrument sound sample (shown below). The spectrogram view of the sounds allows us to see how the intensity of certain frequencies changes over time.

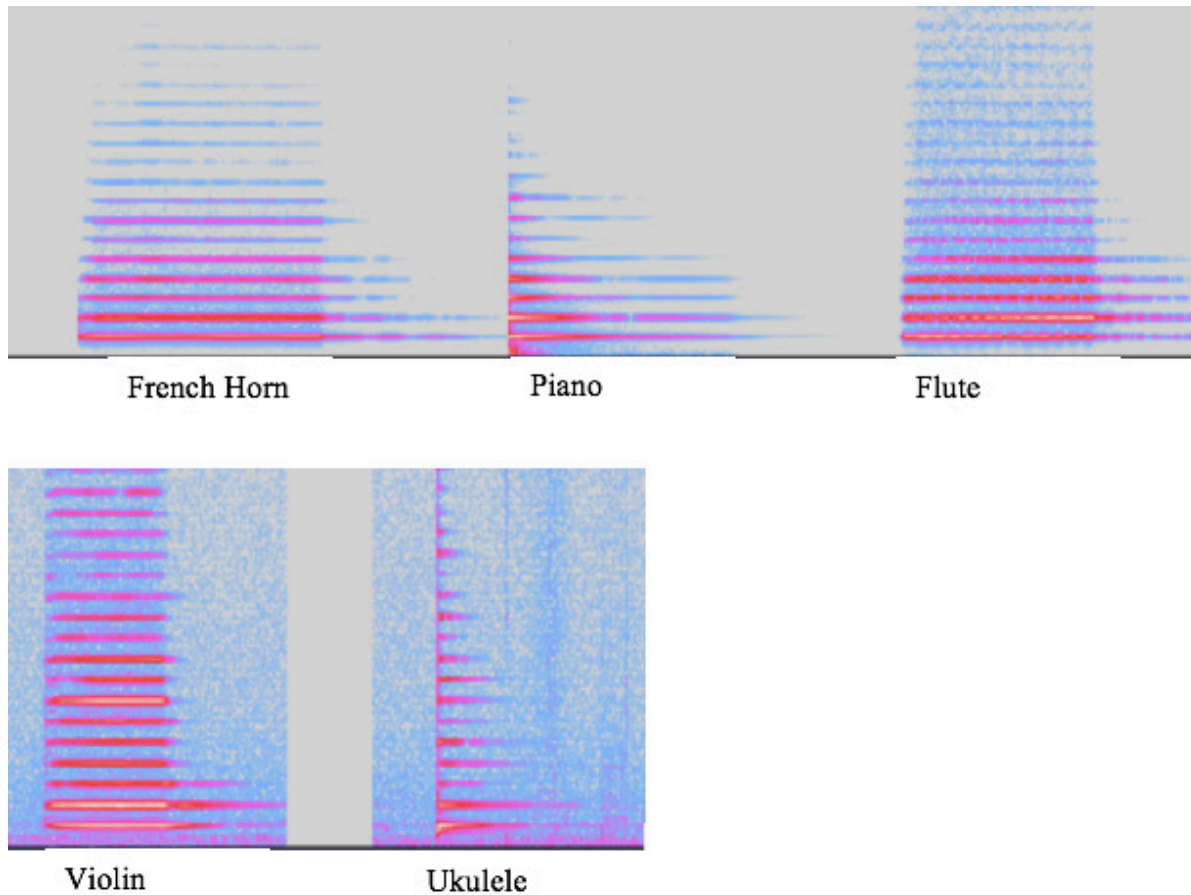


Figure 4: Spectrograms of each instrument, illustrating the frequencies present as represented by each horizontal line and the decay of each frequency with time. The length of each line reflects how quickly certain frequency lines decay in comparison to other frequencies present in each instrument sample.

In the violin spectrogram, it is easy to notice a wide range of frequencies that last for nearly the entire duration of playing. The intensity of higher frequencies sees relatively little decay in comparison to the other instruments, which explains its “brighter” sounding timbre. Another interesting observation that can be made by examining the spectrograms is that of the rapidly decaying higher frequencies of the ukulele, which appear to reinforce the supposed “darker” timbral quality suggested by the lack of intensity of the higher harmonics. The relatively large initial range of frequencies also indicates that the act of plucking the ukulele string may contribute to a brighter and fuller sounding timbre at the moment of attack without lasting for the entire duration of the sound. In the French horn, the rapidly decaying higher harmonic frequencies reflect the “darker” timbral quality illustrated in Graph 1. The flute spectrogram shows more intense frequencies than the French horn, indicating a brighter timbre. However, the higher frequencies decay relatively rapidly which may indicate that the timbre of a flute may be “narrower” than implied by its harmonic spectrum. The piano spectrogram appears to have the smallest range, in accordance with its harmonic spectrum, as its higher harmonics decay the most rapidly in comparison to the other instruments.

By comparing the spectrograms of each instrument with the range-intensity graph of timbre classification, we can see that a combination of both quantitative analysis of harmonic spectrums, as well as qualitative assessment of the spectrograms can reveal much about the timbral elements of various instruments. Implications of these results will be further discussed in the following conclusion in Section IV.

Conclusion

From this investigation, it can be concluded that spectral analysis is useful in extracting information that can be used to compare the timbral qualities of different instruments. According to the results, the violin has the “brightest” and “fullest” timbre with a range of 17451 ± 2 Hz and a -41.06 ± 0.1 dB average intensity over the first 10 overtones. The piano has the “narrowest” timbral quality, with a range of 4469 ± 2 Hz, while the French horn has the “darkest” sound with an average intensity of -61.93 ± 0.1 dB. Organizing the timbres of each instrument based on the frequency range and the average intensity of the first 10 overtones can be useful in gauging how the timbral qualities of each instrument relate to one another. For instance, the results of this investigation show that the range of frequencies present in the harmonic spectrum of a sound produced by a violin is greater than that of a flute, which suggests a relatively brighter timbral quality in the violin.

This investigation is limiting in various aspects. Perhaps the most prominent limitation is the subjectivity of timbral perception. Each individual perceives sounds differently. When it comes to using language to describe those sounds, the subjective aspects of timbral perception become increasingly apparent. In this essay, a more objective approach is attempted through the effort to categorize timbre by means of a spectral analysis, using words like “bright” or “dark” only as a reference point in describing differences in the timbres of various instruments. However, timbral perception is a complex process and the data collected through spectral analysis may not be sufficient in fully describing an instrument’s timbre as it is perceived by the human ear.

With regards to the method of timbral analysis used in this investigation, there are several points of improvement that can be made to obtain more precise and accurate results. Specifically, finding the average intensity over all present overtones rather than the first 10 would likely yield a more accurate assessment of the “brightness” of each instrument’s timbre. During recording, any background noise would have been run through Audacity’s FFT software, possibly affecting the resulting frequency spectrums. Moreover, the frequency peaks displayed on the spectrums generated by Audacity’s FFT software are limited in their precision since each peak is approximated as a curve. The possibility of computational errors in the software must also be acknowledged. Furthermore, the instrument samples may not have been perfectly tuned to $A=440$ Hz, possibly affecting the generated frequency spectrums. Lastly, the intensity of each sound sample was not kept constant, thereby influencing the timbral descriptions based on the intensities of the first 10 overtones.

As discussed in Section II, a quantitative analysis of timbre in the form of spectral analysis has many applications in music, science, and technology. By developing a method to quantitatively analyze timbre, it becomes possible to accomplish what traditional qualitative timbral analysis cannot. Through a quantitative analysis, it becomes possible to translate timbral qualities so often tied to emotion or feeling into data that computers can understand and work with. Perhaps the most interesting application of spectral analysis is using FFT software to convert timbral qualities into alternative sensory signals. Today, there are many studies that have shown it is possible to translate sound into vibrations or visually stimulating patterns so as to allow a deaf person to perceive sound through a different sensory experience. Quantitative timbral analysis would allow for this technology to be further enhanced, even allowing a deaf person to experience the full emotional breadth of a symphony. By quantizing the information contained within a sound’s timbre, one can transform this data to generate an alternative sensory output while preserving the same emotions and meaning carried by the timbral qualities of the sound itself.

Acknowledgments

I would like to thank Mrs. Esch for her support throughout these last two years of my high school physics career, and for her guidance in making this project a reality.

References

- “Audacity Team (2020).” Audacity®: Free Audio Editor and Recorder (Computer application). Version 2.2.2. <https://audacityteam.org>. Accessed July 2020.
- “FFT (Fast Fourier Transform) Waveform Analysis.” *DATAQ Instruments*. www.dataq.com/data-acquisition/general-education-tutorials/fft-fast-fourier-transform-waveform-analysis.html. Accessed July 2020.
- “Frequency Domain and Fourier Transforms.” *Princeton*, https://www.princeton.edu/~cuff/ele201/kulkarni_text/frequency.pdf. Accessed July 2020.
- “Harmonics.” *TheDAWstudio*, 13 May 2016, www.thedawstudio.com/harmonics/. Accessed July 2020.
- “Introduction to Fourier Transform.” *Swarthmore*, <https://lpsa.swarthmore.edu/Fourier/Xforms/FXformIntro.html>. Accessed July 2020.
- Jacobsen, Eric. “Time-Domain Periodicity and the Discrete Fourier Transform.” *DSP Related*. <https://www.dsprelated.com/showarticle/174.php>. Accessed July 2020.
- Jovanovic, Jovan. “How does Shazam work? Music Recognition Algorithms, Fingerprinting, and Processing.” *Toptal*, <https://www.toptal.com/algorithms/shazam-it-music-processing-fingerprinting-and-recognition>. Accessed August 2020.
- O. Moravec, J. Stepanek. “Verbal Descriptions of Musical Sound Timbre and Musician’s opinion of their Usage.” *Academy of Performing Arts in Prague*, <https://pdfs.semanticscholar.org/7308/060d353cd0fa36ec7c5d6b502ce6e63a2954.pdf>. Accessed July 2020.