Grant Bouvier

Computational Musicology

Final Project - Write Up

The Chroma Analysis of Metallica's *Master of Puppets*: A Discussion

Overview:

The ultimate objective of this project is to use chroma analysis methods on Metallica's 1986 thrash metal magnum opus *Master of Puppets*. The data gathered from the chroma analysis of each track will then be plotted in graphs to better visualize the pitched content of the recorded audio. Chroma analysis involves using complex digital signal processing algorithms to divide an audio file into a set of choma feature bins which roughly approximate musical pitches that we hear as notes. Chroma is typically set to twelve bins, which equates to the standard even tempered chromatic scale in western music. The data acquired through this analysis can then be used to plot the relative energy present in a note bin across all registers at a given point in an audio file. Valuable practical and theoretical knowledge can be derived from chroma feature analysis and plots, and since this type of analysis directly relates to the music characteristics of an audio file, it is particularly handy for the purpose of this project.

The primary hypothesis of this project is that Metallica's *Master of Puppets* will contain a significant concentration of tonal information around E and F notes across all octaves. This hypothesis is drawn from the author's own experience with the musical material, and the idea

that the design and tuning of the modern electric guitar lends itself to writing rock and metal compositions that are centered around E and F keys.

The author is somewhat new to programming in python, so a secondary objective of this project is to properly implement chroma analysis via the Librosa library. Both a constant-Q transform and an short-time fourier transform are used to generate chroma graphs for visual analysis. A hop size of 512 samples was used (for both stft and cqt) with a fft window size of 2048 samples. The signal was divided into 12 pitch bins (for a western chromatic distribution).

A Summary of Past and Relevant Work:

Chroma features analysis has been a topic of research and interest in signal processing for a significant period of time. A wealth of published research exist on using chroma features for music analysis and identification, ranging in general complexity from introductory to advanced. Two seminal work that seems to be heavily referenced are "Musical Similarity Analysis based on Chroma Features and Text Retrieval Methods" (Englmeier, Hubig, Goebl, Bohm) and "Audio Thumbnailing of Popular Music Using Chroma-Based Representations" (Bartsch and Wakefield). Both effectively describe algorithms most effective for pulling musically relevant harmonic and melodic signifiers from signal.

Unfortunately there has been fairly limited spectral research into heavy metal music specifically. The only significant and reliable work that has any relevance to this project is "Automatic Classification of Heavy Metal Music" (Mulder). This research does consider chroma vectors as a useful element in metal classification and analysis. Other than this work, the field appears to be somewhat thin on heavy metal signal processing and analysis.

Setup and Technical Details:

This project involved curating MP3 files of the original *Master of Puppets* record, which consists of eight tracks in total. These MP3 files were then put in the same file directory as a python script (including the librosa and numpy libraries). The sample rate is 44100 and the hop length is set at 512 samples with an fft window of 2048 samples. These setting specify how the audio is windowed, or broken up for processing. A function is then used to apply the chroma_cqt and chroma_stft analysis to the desired audio file (in this case a full length track from the record). Both of these chroma analysis provide different but useful data, so both were used. This function was then run for each audio file to provide chroma feature plots for every song on *Master of Puppets*. The tracks themselves are sequential in the code, but are also labeled via comments. 3 of the plots are discussed below in the analysis section, with the other viewable in the original code file.

Constant Q Transform vs. Short-Time Fourier Transform:

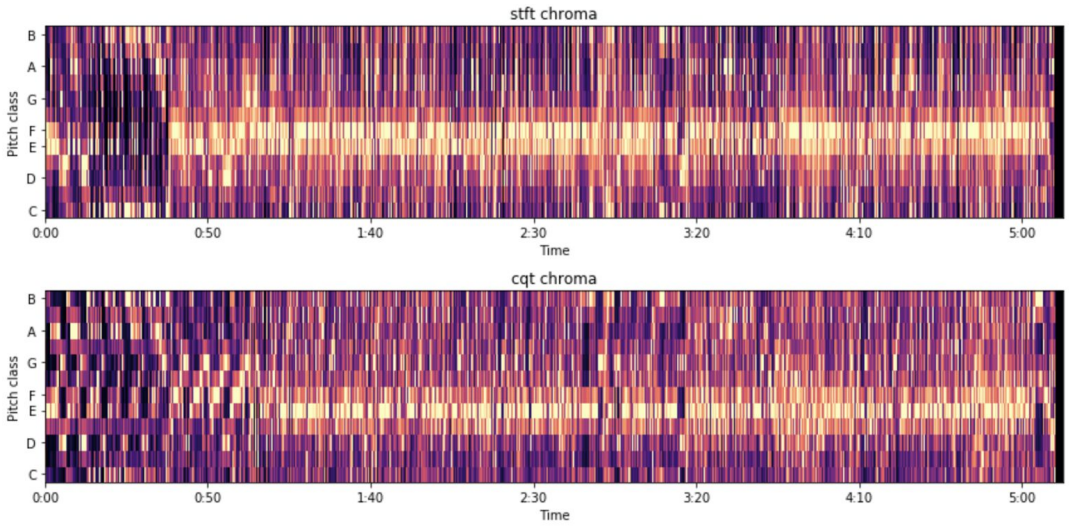
This project involved analyzing choma features of full length audio tracks using both constant Q transform (CTQ) and short-time fourier transform (STFT). The main difference between these two transforms is that CTQ utilizes logarithmically spaced filters to divide the signal into frequency bins while STSF uses linearly spaced filters. Since CQT bins are logarithmic, they more accurately represent how the human ear perceives frequency as pitch. CQT also has the added benefit of requiring fewer bins. These characteristics of the CQT have given it a reputation as being more useful for musical signals and music related analysis.

The STFT is still a useful and effective method for analyzing chroma features however. The main advantage of the STFT is that it is less complex to implement and therefore is more easily utilized with limited computational resources. Additionally, STFT has a higher accuracy in the upper frequency ranges due to its linear segmentation and filtering of audio. Since there are useful applications for these two transforms, both are implemented in the chroma function devised for this project.

Analysis:

The following written analysis discusses the plotted chroma data related to the audio tracks themselves. The tracks featured in this analysis were chosen to highlight specific points of interest related to the hypothesis, and those plots not featured can be views in the original code.

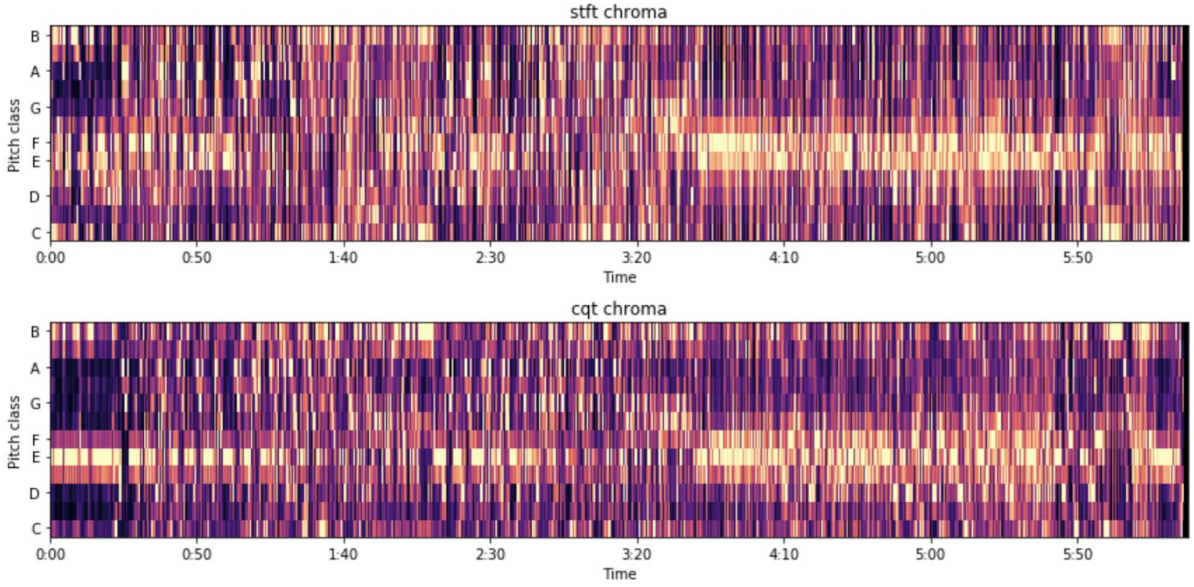
Track 1 - Battery



The first track on the record, entitled “Battery”, features a clear band of concentration in the E and F note bins throughout almost the entire track (the first minute is less concentrated because of an intro that differs both stylistically and compositionally from the rest of the song). Battery is

driven by rapidly played E and F power chords throughout the song with only short breaks for solos and bridges. In the instance of Battery, there seems to be no major deviation from the CQT and STFT data sets, although it does seem that the CQT set more accurately depicts the song's intro section. This could be because toward the end of the intro there is a significant dynamic increase and timbral shift that is more effectively compensated for by the CQT.

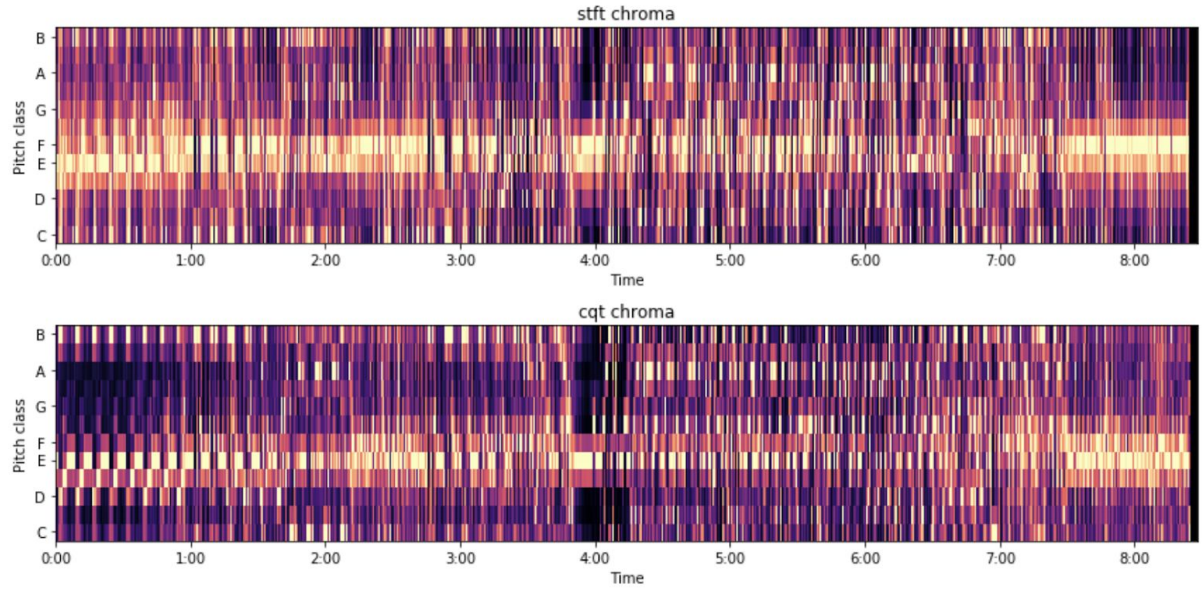
Track 4 - Welcome Home



"Welcome Home" is the fourth track on the record and features another intro section that differs from the rest of the song musically and stylistically. The track itself does not seem to display a clear concentration in the E and F bins until about the 3:30 mark. This point in the song sees a shift in the composition to rapid E and F power chords, similar to what is performed in "Battery". The remainder of the song features a clear band of concentration. It is again interesting to note that the CQT data again appears to do a better job handling the intro section,

which like “Battery” suddenly transitions in a moment of significant timberal and dynamic change.

Track 7 - Orion



“Orion” is the seventh track off the record and is specifically interesting to the project because it is an instrumental track. Although the chroma algorithms employed in the main function features some filter parameters to focus in on the harmonic content of the audio, the fact that this track has no vocals makes possible for the analysis to produce even more accurate data. This track again displays the predicted band of concentration in the E and F bins, but this band seems to be less concentrated overall than other tracks. That could be a result of the composition itself, but one could argue that it is related to the reduction in overall sonic material in the audio without vocals.

Results:

Based on the plot sets derived from the chroma analysis of *Master of Puppets*, it seems as though the original hypothesis holds up. The record displays a clear compositional tendency for the songs to center around E and F with minimal deviation. There was an attempt to apply smoothing to the chroma analysis which was ultimately unsuccefuly, however the unrefined chroma data on its own illustrates a significant band of concentration in the chromatic range predicted.

Historically this record is considered among the most relevant and important works of the era, going on to be very influential in the genre and rock music at large. It is fair to say that this record therefore represents a somewhat accurate depiction of metal music in the mid 1980s. It is impossible to determine if the tuning and design of the modern electric guitar itself played a role in the development of the songwriting tendencies of this genre at that time, but one could make a case for the correlation. The lowest tuned string of the guitar in standard tuning is E, which allows for chords based on the low E to dominate faster paced music as it can be played quickly with minimal hand movement. Further research and testing would be needed to prove this correlation.

As for the differences in the constant Q and short-time fourier transforms, it seems as though both are adequate for analysis in the scope of this project and that either one could stand on its own for the purposes outlined here. The CQT did seem to provide more clear and accurate data (compared to the actual audio itself) in moments where the composition made sudden or rapid changes in dynamics or timbre. Ultimately, there is some credence to the argument that CQT is, in fact, more “musical” than the STFT. Furthermore, the logarithmic filtering used in

CQT has a smoothing type effect on the data when plotted, which was helpful since an actual smoothing algorithm was not successfully applied in this project.

Conclusion:

This project was a success in that the original hypothesis was met and that the author was able to effectively implement chroma analysis with librosa in python. For future projects, it could be advantageous to integrate chroma smoothing to the plots, but even the raw chroma data yielded valuable information that could be used to support the hypothesis.

Additional Research - Beat Synchronous Chroma:

Following the completion of the initial research for this project, the author revisited the hypothesis using a beat synchronous chroma algorithm. Beat synchronous chroma analysis essentially divides the audio file into beats by detecting steady rhythmic events, and then calculates the chroma features at those points. This is seen as a somewhat more efficient means of analyzing chroma, since significant musical information usually occurs on the beats of the rhythm (in western popular music, at least). Upon viewing the plotted results, E and F note bins once again have a clear concentration against other note bins during the main portions of the compositions (excluding intros, solos, outros, etc.). There is an overall reduction in chroma feature energy across all bins as a result of the data being calculated only on detected beats.

While the beat synchronous chroma analysis is certainly a useful method for retrieving useful chroma data, overall it is not significantly more useful for the purpose of this project. The plots themselves are somewhat easier to look at because of the overall reduction of energy, and it

also makes identifying solo and intro sections of tracks much easier since those portions of the compositions typically lack clear percussive rhythms. The code and plots for the beat synchronous chroma analysis can be viewed with the original code: https://github.com/gbouvier1/Computational-Musicology-Fall-2018-/tree/master/finalProject/code