Final project part II

Zhi Jie Huang

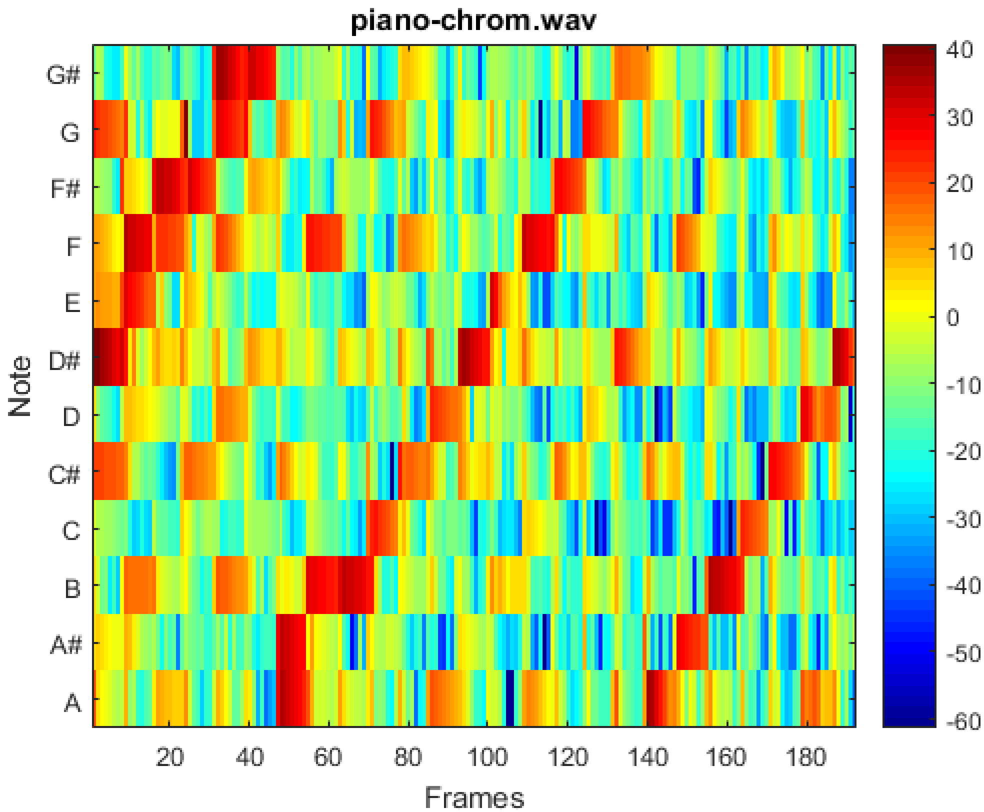
ECEN 4632

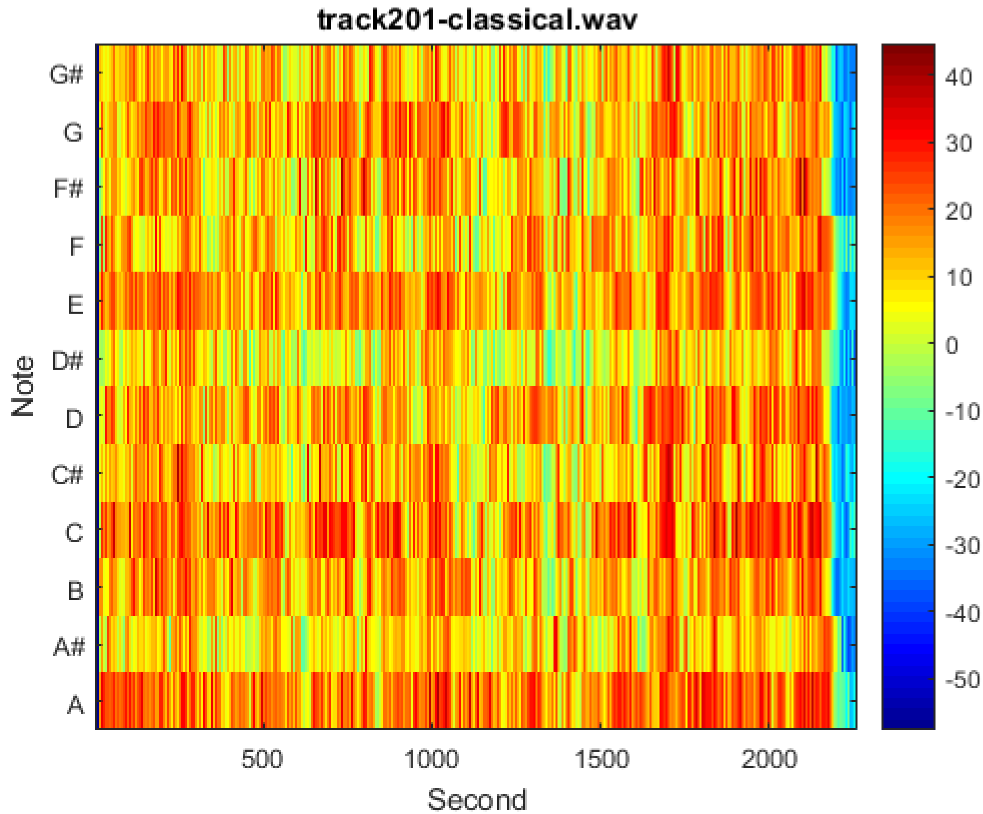
Project partner: Richard Smith

In this part of the final project, I am trying to first find corresponding frequency of the local maxima of the song, bring all of those peak frequencies into the first octave with different semitones, based on the distance the semitones that is away from the octave; I distributed the peak frequencies with the weighted function. And finally I square the peak value of the peak frequencies and sum up the energy of each note.

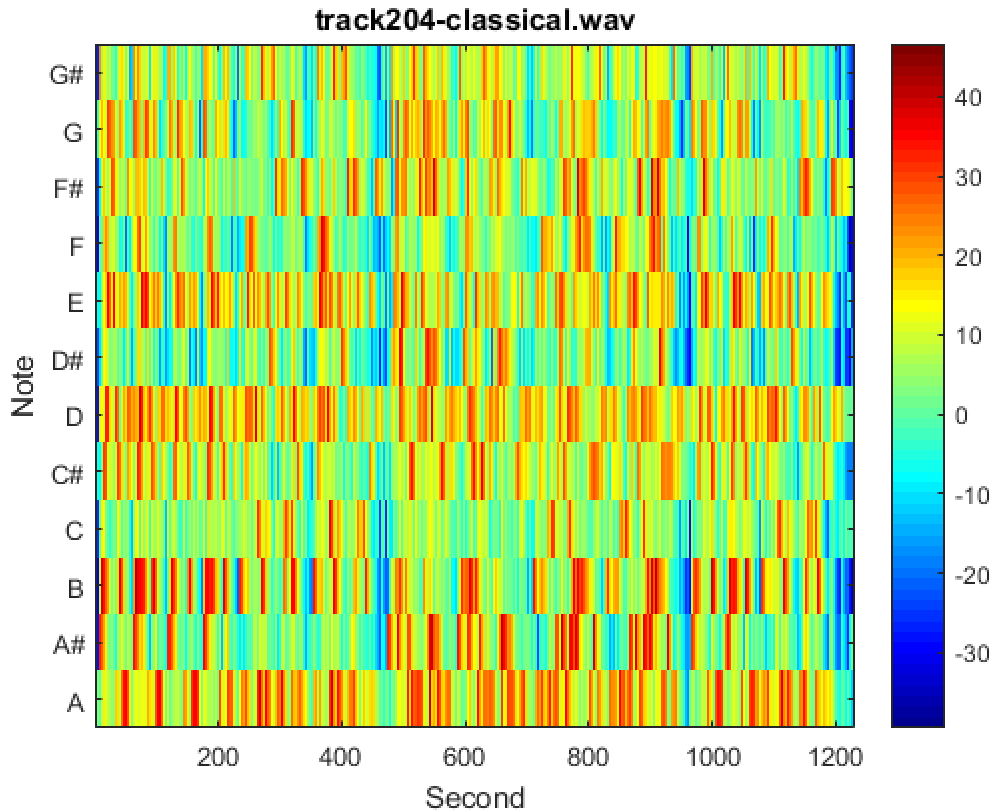
The way my algorithm works is I find the maximum value of the whole song, use a value that is 23 db below my maximum value as a threshold, any value below the threshold get filter out.The way to calculate the NPCP(Normalized Pitch Class Profile) is to pass in 2048 samples at a time into my NPCP.m, in addition to that, I move the frame of the song at half of the frame which it's 1024 samples at a time. Because I want my signal to overlapped in order to prevent the loss of information. I am using the hann window to cookie-cut my signal.I did a Fast Fourier transform on the 1024 samples of my signals multiply by the hann window in order to take the signal into frequency domain. Since the fft will give us half positive frequency and half negative frequency, I truncate the negative frequency of the fft signal by taking the first signal up to its nyquist frequency (i.e. 2048/2+1 = 1025 in this case). With the fourier transform of the signal, since some of the value of the fourier transform can be negative. I took the absolute value of the fourier transform, find the local maxima of the song that is within the range of -23db of the maximum peak of the song. After I have the peak frequency, since the fk is in representation of the frame index, therefore, I did conversion to frequency. Bring all the frequencies to semitones, and then we divide each semitone by 12 and take the remainder of the division. That gives me the chroma of the weighted sum of all the peak frequencies into the first octave. Since matlab have index starting at 1 instead of 0, therefore, I added 1 to all the note. Since while we calculate the semitone, I did some rounding, therefore, since 12*log2(FK/f0) is always between 1 and -1 of the semitones. After I mapped all the peak frequencies into different note, since I don't have time to look at each note one by one, therefore, I sum up the total energy of each note on all of the octave and multiply the peak square by the weighted function.

The scale on the right of the graph represents the decibel of the particular notes being played. When color is represented the amplitude on the graph, it means that the note on a scale of A-G is being played. Most easily followed along with songs that have a single instrument playing a melody, the notes of the music being played can be visually read on the graph. For instance, when I was looking at the piano-chrom.wav, there is only a single note been played at a time, I can see clearly the pianist started with D# and going up in frequency. Since all octave has been bring into the first octave, when the music hit the next octave,it will just start with A again.

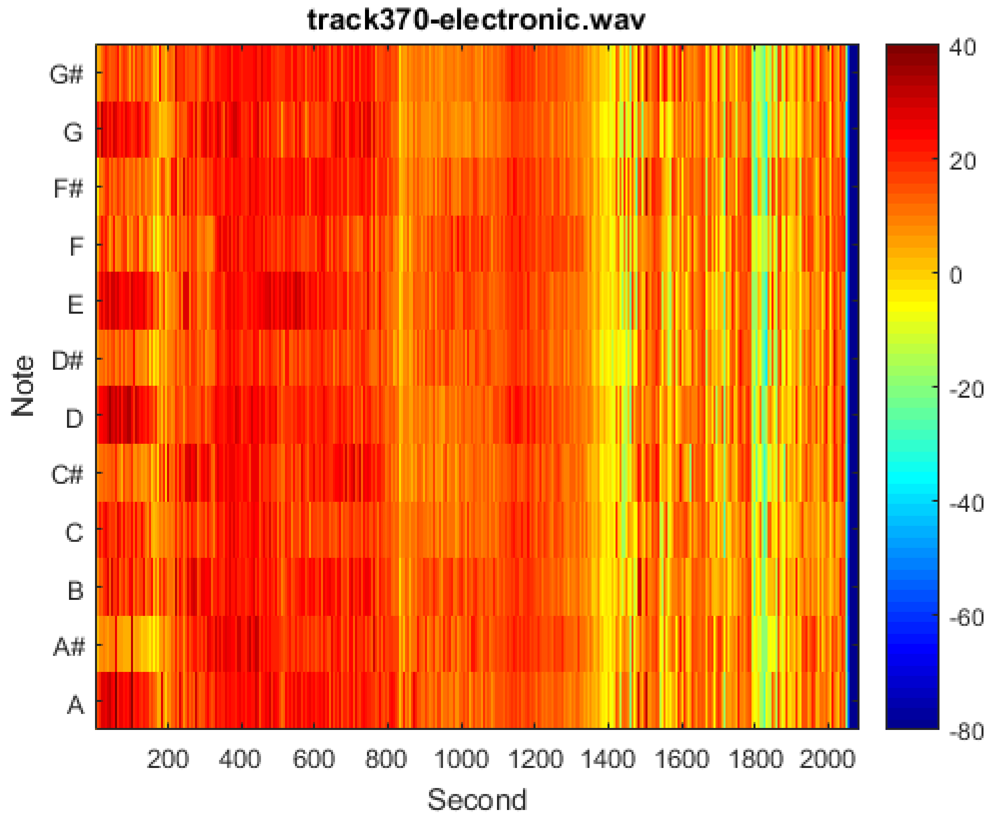




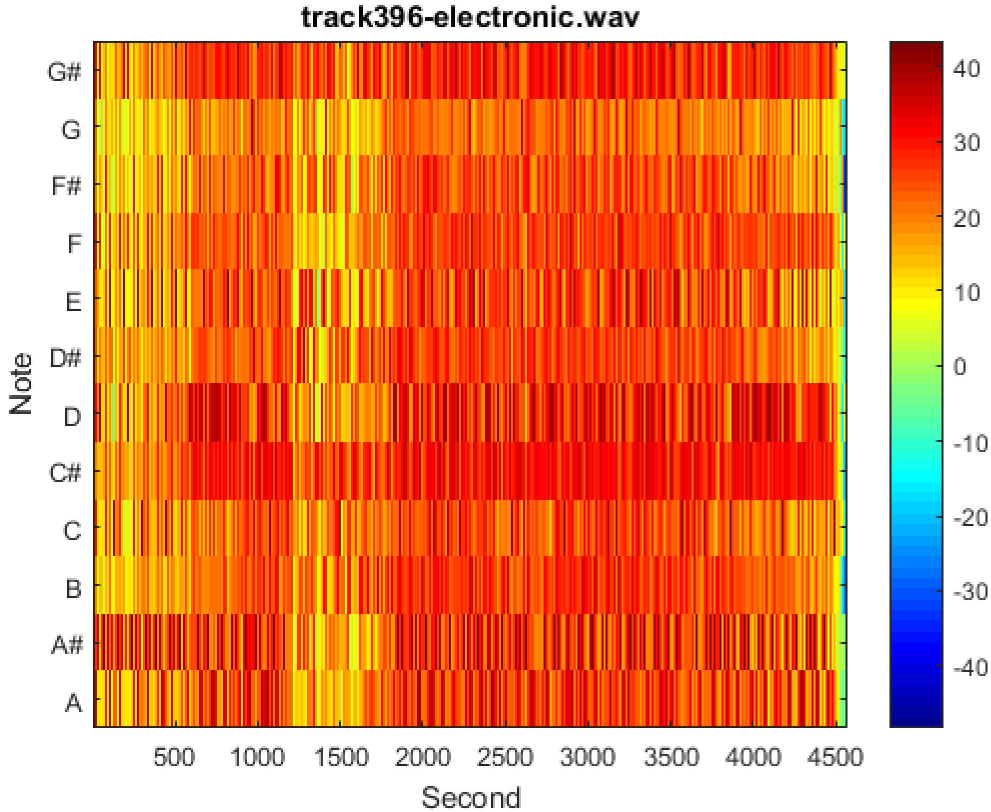
The song Violin Concerto in A Minor is in the key of A minor, which has a tonal center of the most frequently occurring note A. The graph shows A as the most common note in the song. The concerto has multiple instruments playing different parts and notes, therefore the graph corresponds with its representation of many notes being played simultaneously. I can hear a lot of harmony on this track.



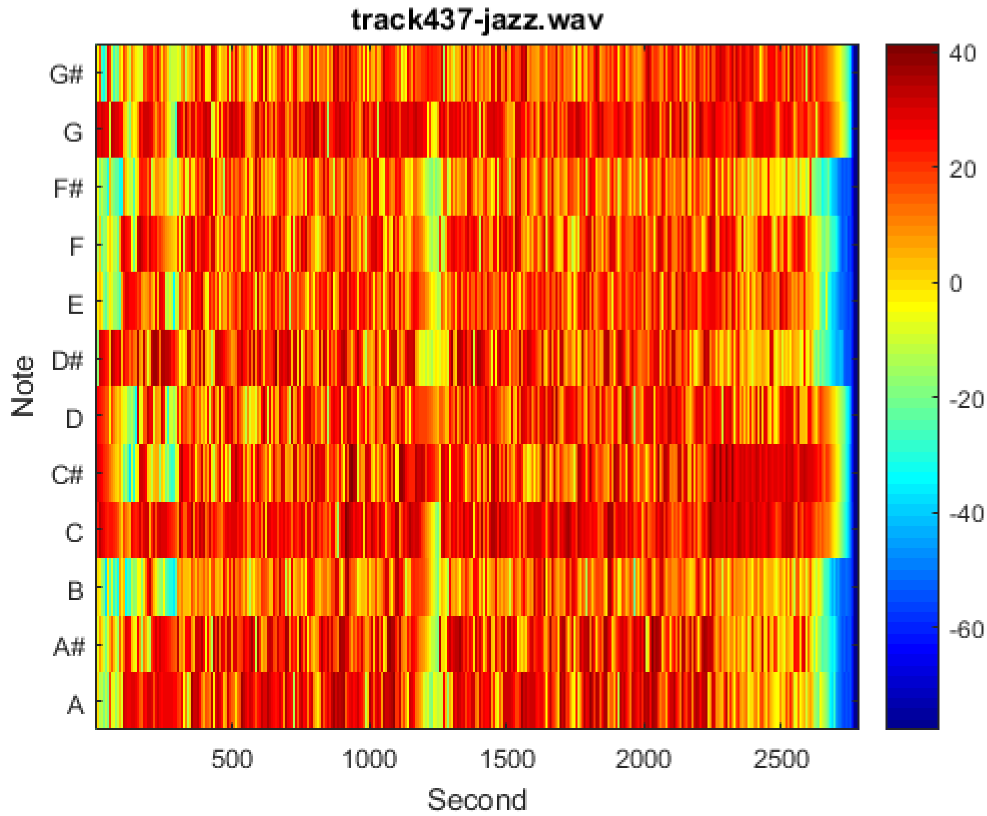
The song Schumann Davidsbundlertanze is simpler to follow than some other examples. The melody from the piano can be followed on the graph as the song's first phrase starts on an A#, rises to a G, then falls back down. So this track has a pattern of starting with low note, go up to high note and then go back down to low note. While I was listening to both of the classical music, I notice there aren't a lot of sharp and flat being play, it is mostly natural notes.



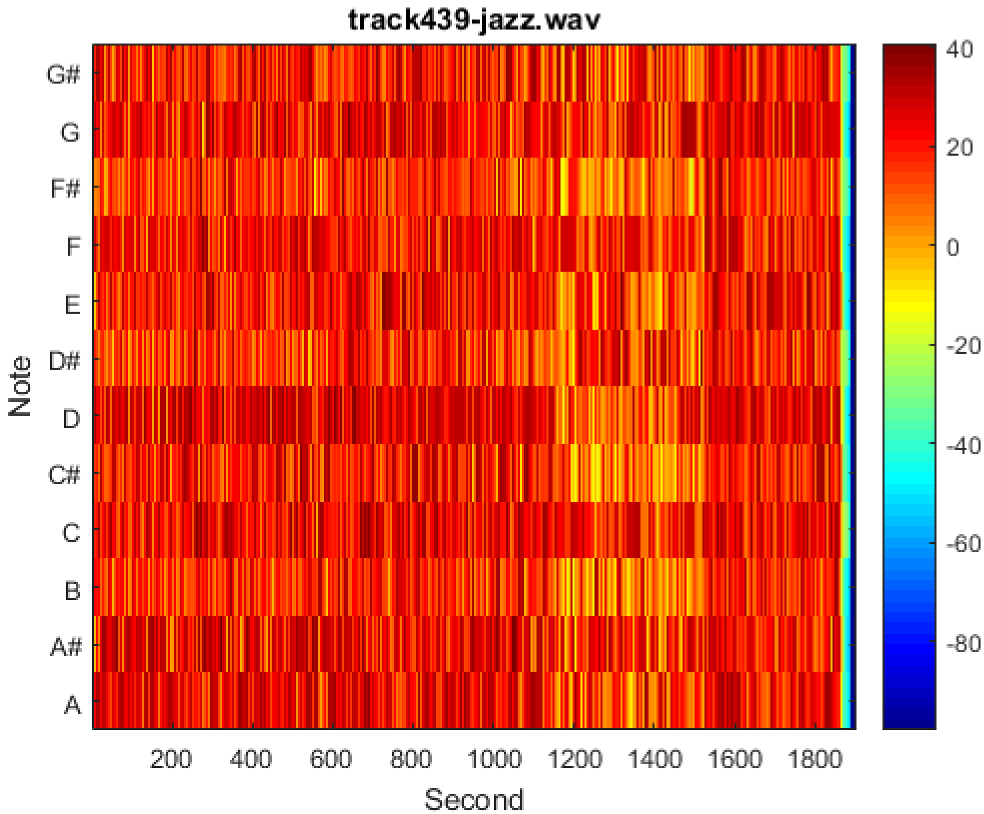
The track 370 is an electronic song and electronic are made with computers, synthesizers and the drum machines and it makes it harder to distinguish from the sound. The song is very rhythmic. From the graph it starts with almost all the notes but more heavily on A,D,E,G because according to the graph it's more heavy marked in red. In the center the notes are rapidly rising and falling, because of the notes are changing rapidly. None of the note has been held for a long time.



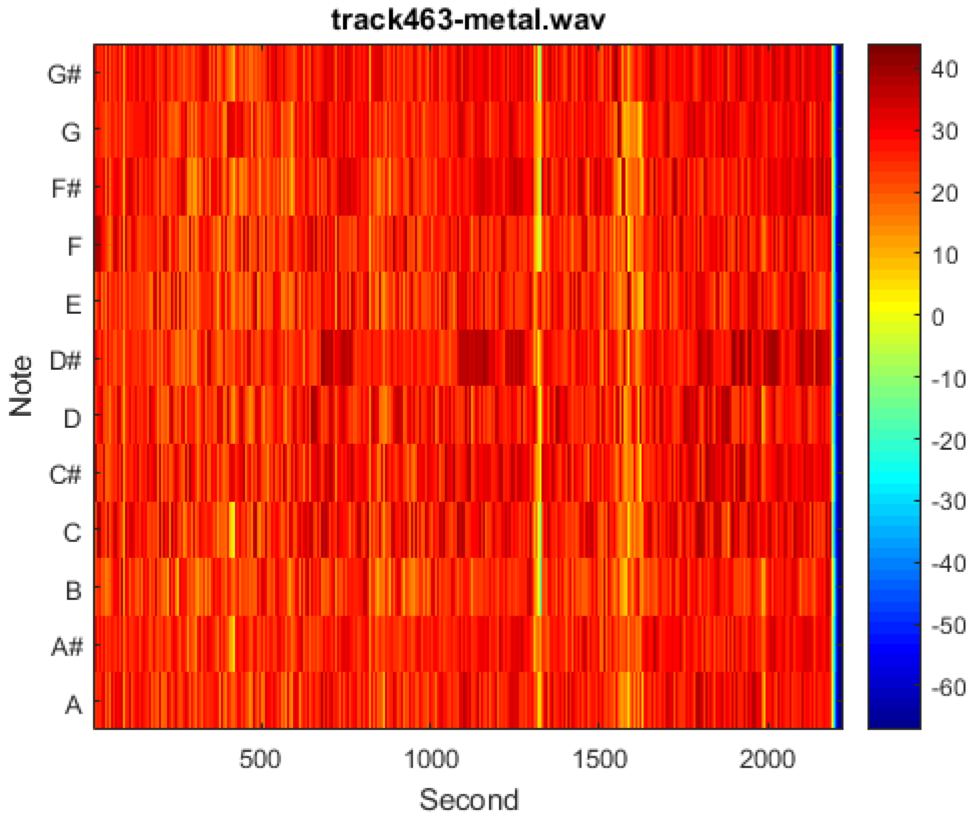
The song track 396 can be classify as Trance Electronic. The Trance more like the classical form of electronic music because Trance style are more tonal but repetitive beat and melodic aesthetics as I can see from the graph the notes are repetitive. I can hear that they are typically faster beats and there is point in graph which it rest between drops, and then it adds vocal into the Trance electronic. So if you can see very fast switch between notes, it's probably electronic music.



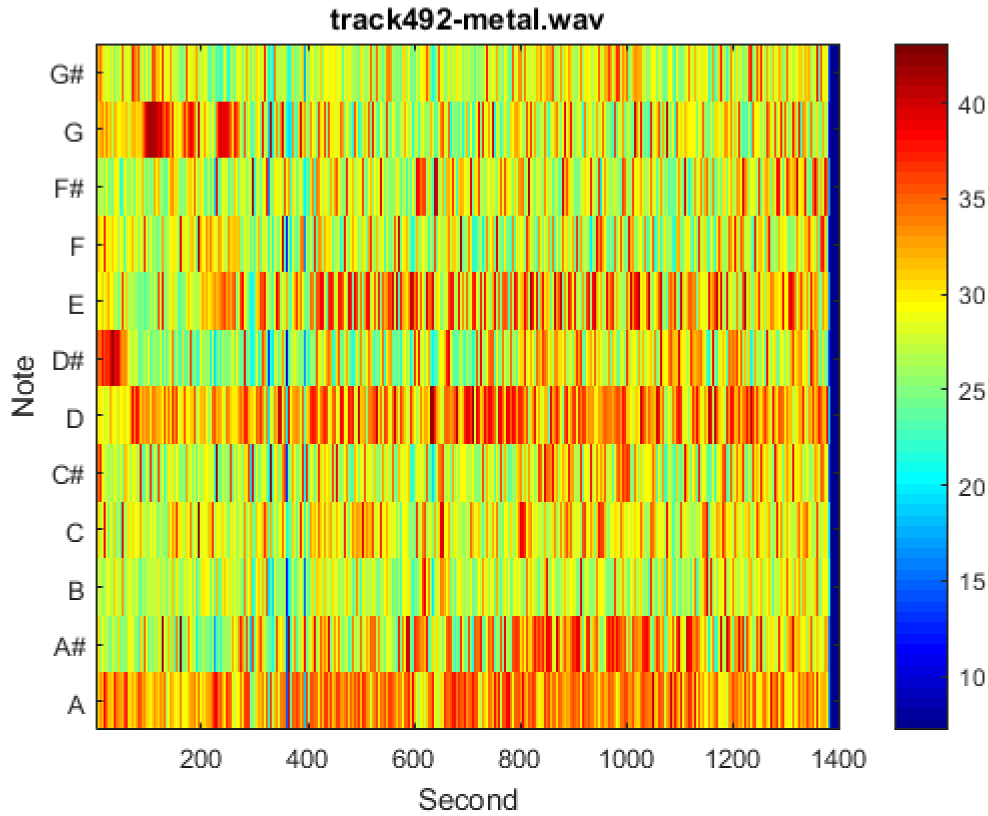
The song is classified as fusion jazz, it can be a combination of two or more genres. It can be mixing of funk and rhythm and blues rhythms. In the beginning of music, it's kind of quiet because the piano was starting off with C, D#, and rise to G notes. Then drums come in and play which you see more decibel on the graphs. The seconds between 100 to 150 there were a rest going on and then both the piano and the drums starts off again that why you see the repetitive of same decibel happen around 50-100 and 130-230. Then piano finished off in the background toward the very end. The fusion jazz can be simple melodic that can be elaborates on the chord progressions.



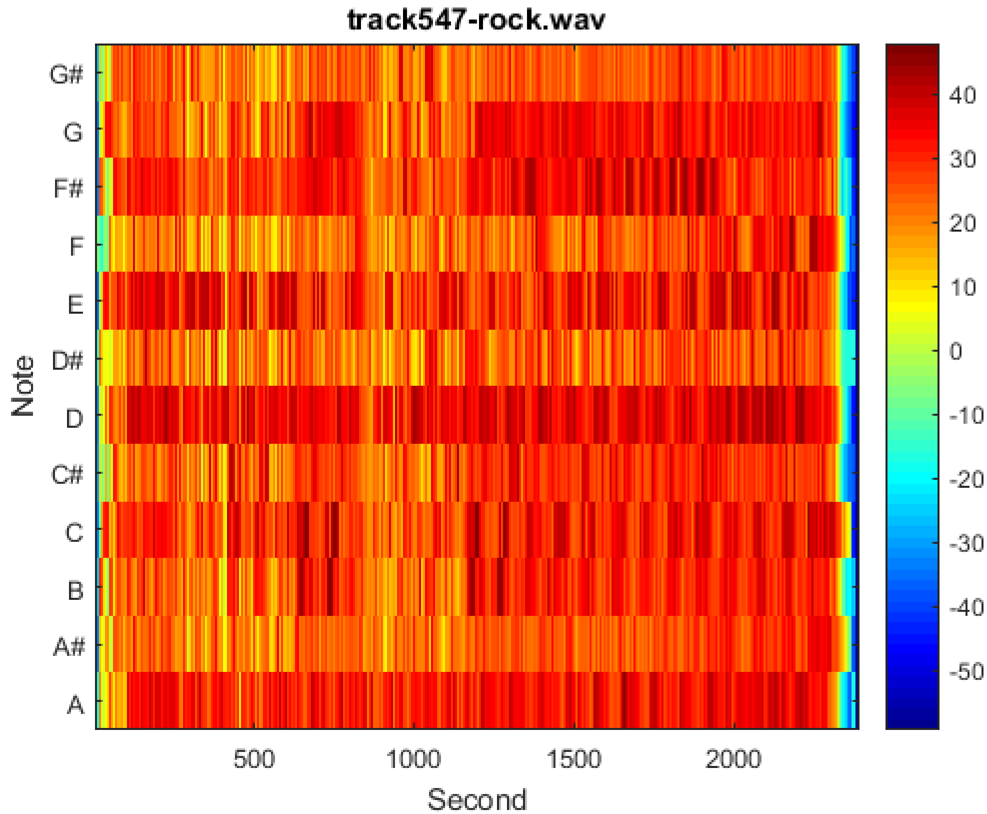
The song wallawalla from Drop Trio starts of heavy with all the instrument at the beginning that's why the graph shows decibel of red and orange tone. During the 20-100 syncopation occurs which you can see in the graph, the decibel is shaded more heavy red, this is the place where it shifts the emphasis of song's rhythm and beat pattern. Then it transitions back to normal rhythm and beat pattern. From 120-140 there is a transition where the keyboard has its own solo parts that why you see more heavy in A#, C, D#, F, G chords. Then it changes back to same repetitive rhythm and beat pattern that why the graph show that 160-180 second is similar to the beginning parts.



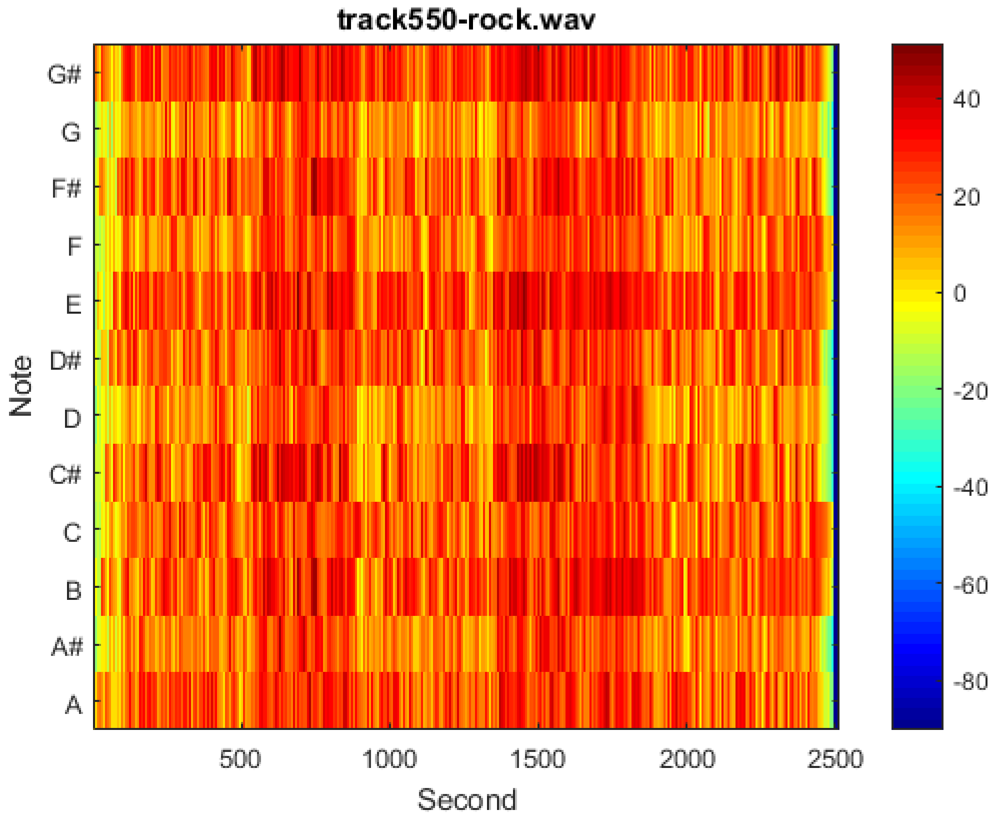
The tempo of this track is extremely fast blast beat, therefore it also makes it seems like you are hitting all of the notes at once, however, you are just hitting short two-note or three-note rhythmic. With the two-note and three-note rhythmic, you can see above the graph is having a lot of hits,so this can be classify as metal.



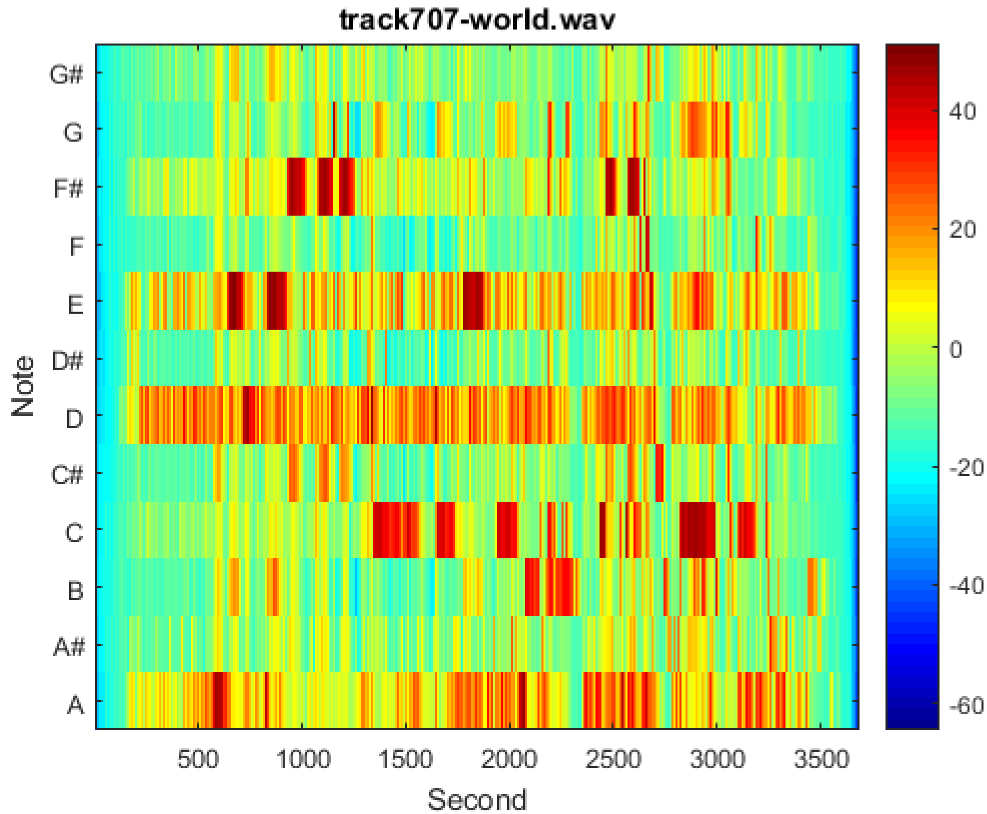
When I am listening to track 492, you can hear the constant electric guitar as a base. The most common note for an electric guitar is tuned E, A, D,G and B. As you can see from the graph above, note A, D, E and G are the most common note for track 492. In this song, you can see the main groove is characterized by short, two-note or three-note rhythmic figures, usually you can classify this as metal.



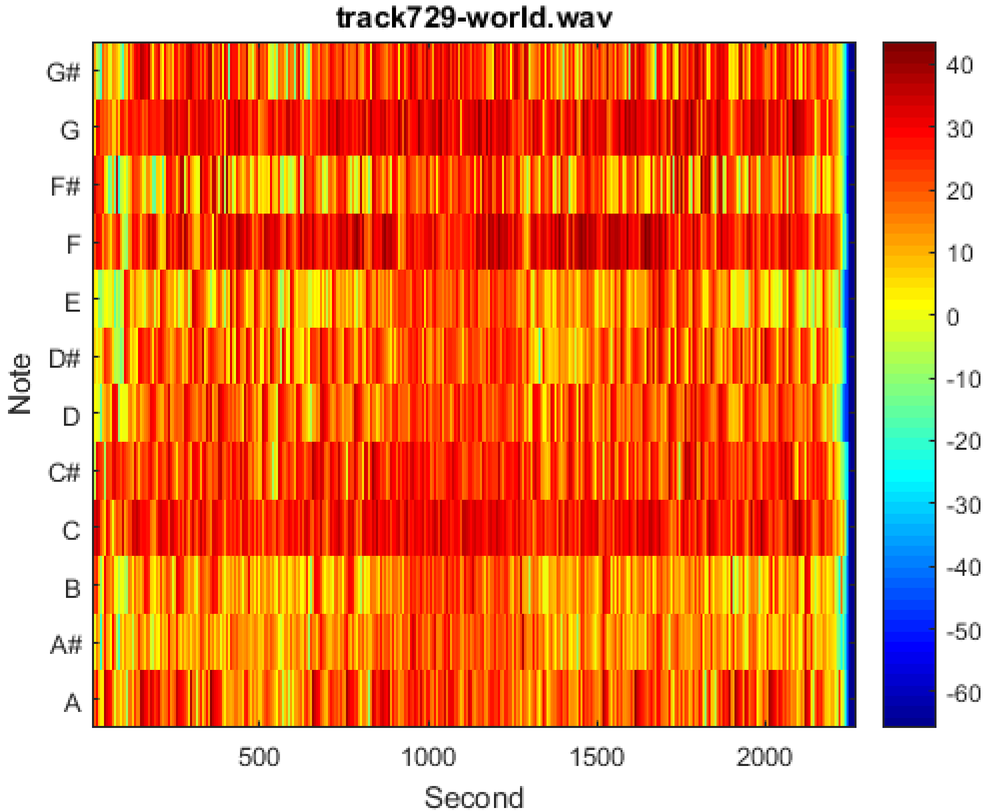
This track 547 rock has the same electric guitar as the base, however, it doesn't have a short two-note or three-note rhythmic, so it looks really like metal, but it's not as note wasn't hitting in a fast blast beat tempo. So, this song can be classify as rock.



In this rock song, the singer is playing with rhythm section instrument such as guitar or bass. At the beginning of graph the bass starts off softly so the decibel is quieter. Later the vocal was added and played with bass. From 50-100 second all the instrument was playing along with vocal. From 100-140 the singer is playing rhythm with guitar and then it repeats back to playing together. The last part of the graph it also portrayed that the singer is playing along with guitar while bass is at the background playing simple notes.



At the beginning of the track 707, it's very quiet. You can see that on the graph above, and then you can hear the flute play D note throughout the whole song, since world music usually have either a relatively quiet start or end, in addition to that, it has been play in a monotonic tone. So, this can be classify as world music.



This song may connect with traditional Japanese instruments such as the Koto and Shamisen, they are under the category of stringed instruments. The main notes are C,C#,F,G they are played throughout the whole song as indicated on the graph. The song was very consistent with the rhythm and repetitive. So, it can be classify as world music.

Appendix:

FinalProjectPartII:

```
clear all;close all;clc;
%%%%%%%%%%%%%Final project part I From the wav file to the psychoacoustic
%%%%%%%%%%%%%features
%%--------------use dir **/*.wav to list all the wav file in the directory
tic
%------------Read in all the filename--------------------------
filename = {'track201-classical.wav','track204-classical.wav',...
   'track370-electronic.wav','track396-electronic.wav',...
   'track437-jazz.wav','track439-jazz.wav',...
   'track463-metal.wav','track492-metal.wav',...
   'track547-rock.wav','track550-rock.wav',...
   'track707-world.wav','track729-world.wav'};

%    filename = {'piano-chrom.wav'};
  for fileIndex = 1:12
     [song,fs] = audioread(char(filename(fileIndex)));
%         [song,fs] = audioread(char(filename));
   %%%% audio files are sampled at fs = 11025 Hz
   %sound(song,fs)

   %%%%%%--------------initialization----------------------
    SongLength = length(song);
%      Mid = floor(SongLength/2);
%-----------Calculate the maximum value of the whole song,use for threshold
%and normalization.
    Threshold = max(max(song));
   %%%%-----------------------Part 1------------------------------
   %% Take the whole song and calculate the
%     xn = song(Mid:Mid+24*fs-1);
   w = hann(2048);
   fftsize = 2048; %%%%Size of fft
   nf = floor((SongLength-2048)/fftsize/2);
```

```
    index = 1;
%Allocate a 12 by number of frames vector
output = zeros(12,nf);
%Passing each frame of the song into the NPCP funciotn
  for n = 1:fftsize/2:floor(length(song)-2048)
      output(:,index) = NPCP(song(n:n+fftsize-1),fs,fftsize,w,Threshold);
      index = index + 1;
  end
  output = 10*log10(output/max(max(song)));
  figure
  imagesc(output);
  title(filename(fileIndex));
  set(gca,'YDir','normal');
  set(gca,'YTick',[1:12]);
  set(gca,'YTickLabel',({'A';'A#';'B';'C';'C#';'D';'D#';'E';'F';'F#';'G';'G#'}))
  xlabel('Second');
  ylabel('Note');
  colormap jet
  colorbar
  end
toc
```

NPCP.m

```
function [ output ] = NPCP( wav,fs,fftSize,window,Threshold)
%
% PART 2 : processing of the audio vector In the Fourier domain.
%_______________________________________________________________
%
  xn = wav(1:fftSize);
  Y = fft(xn.*window);
  K = fftSize/2+1;
%    L = K-1;
  Xn = abs(Y(1:K));
%-------------Plot of the fourier transform-------------------------
%    f = fs*(0:L)/L;
%    plot(f,Xn);
%--------Step 1 pick out the peak frequencies-------
f0 = 27.5;
%------Goal of the threshold:Get rid of the small amplitude note----
[pkt,fk] = findpeaks(Xn,'Threshold',Threshold*1e-6);
% [pkt,fk] = findpeaks(Xn,'NPeaks',20);
FK = round(fk*fs/fftSize);
%---------Step 2: Assignment of peak frequencies to
%----------------------------semitones------------------------
  sm = round(12*log2(FK/f0));
  c = mod(sm,12)+1;
  r = 12*log2(FK/f0)-sm;
  output = zeros(12,1);
  for j = 1:12
      var = ismember(c,j);
      w = (cos(pi/2*r)).^2;
      output(j,1) = sum(var.*(w.*pkt.^2));
  end
end
```