Rhythm and meter

Part 1 & Part 2:

Snippet	Perceived (bpm)	Calculated (bpm)
`dream_theater.wav`	82	97.5
`michael_jackson.wav`	130-170	185.3
`mozart.wav`	112-140	140.8
`queen.wav`	66	105.3
`taylor_swft.wav`	40	51.5

- Except for one (‘*queen.wav’*), for all other tracks, the tempo is very much similar. I do not personally see a trend in the discrepancies, but what I feel is the tempo I perceive seems to be a bit less than the actual calculated one. I believe the small discrepancy in general could very well be due to the difference in tapping while listening to music. For *‘queen.wav’* I think, the ups and downs in the wave are not perceivable but are sufficient to show change in tempo, and hence possibly the difference. Find the plots below.

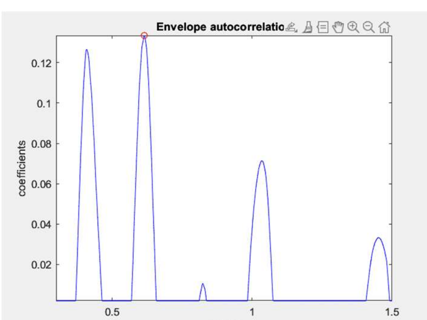


Figure 3 Dream Theater

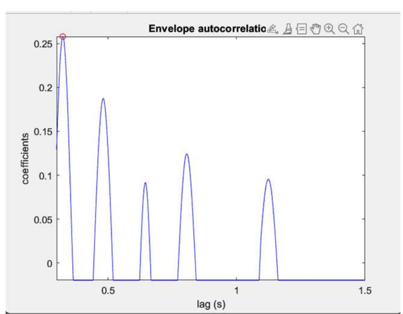


Figure 1 Michael Jackson

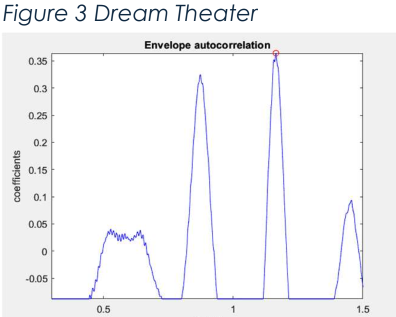


Figure 2 Mozart

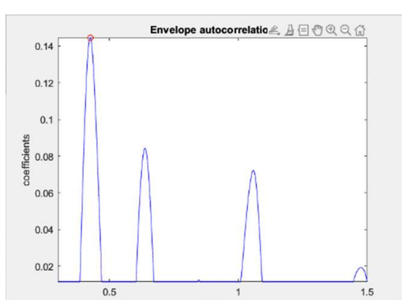


Figure 5 Queen

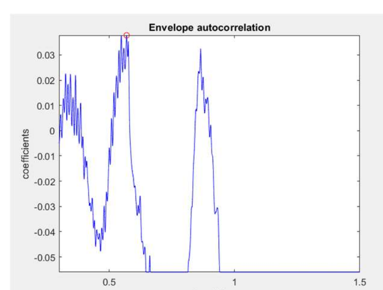
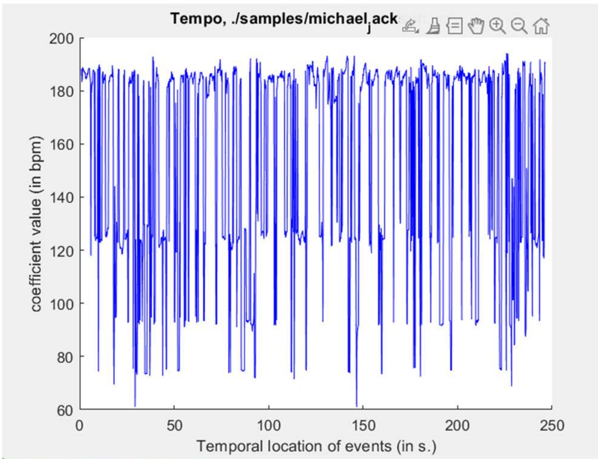
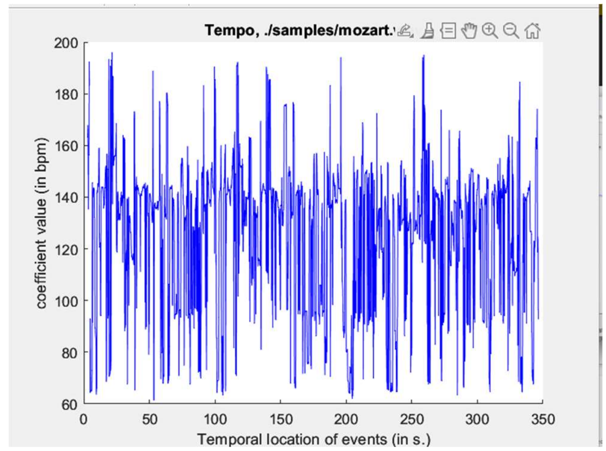


Figure 4 Talyor

Part-3 :





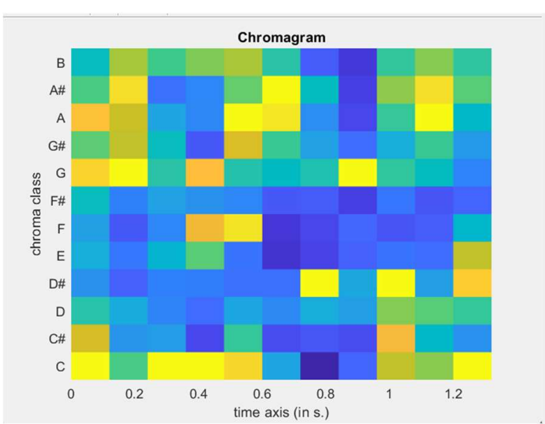
For '*michael.wav*', the actual range of tempo seems to be between 80-185, and some of points are least most of them are in range of 120-180 which is I perceived. 'Mozart.wav', we can see the range in general to be between 60-200, and the major region is 90 to 140 which is also close to my estimate.

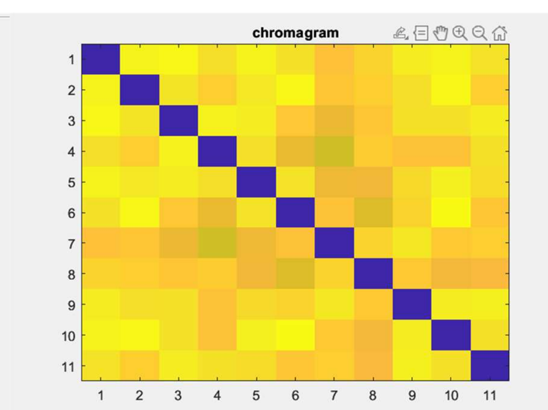
Repetition in Music:

X = rollnumber % 7= 2020102047 % 7= 5

So, analyse *'05.wav'*.

Part 1 & Part 2: chromagram, Similarity matrix using *mirmatrix*



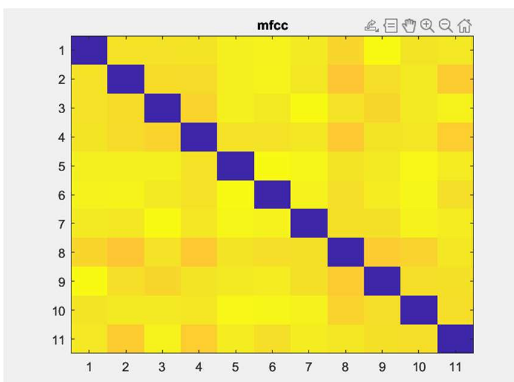


The similarity matrix for a music file shows how similar different segments of the music are to each other. The checkered rectangles in the matrix represent adjacent segments that are likely to be similar due to their temporal proximity. The lines in the matrix represent the boundaries between different musical sections or phrases. By analyzing the patterns in the similarity matrix, we can identify the different musical sections in a piece of music and understand how they relate to each other. In simpler terms, the similarity matrix for a music file shows how similar different parts of the music are to each other. The checkered rectangles represent similar adjacent parts of the music, while the lines represent the boundaries between different parts. By looking at the patterns in the matrix, we can identify different musical sections and understand how they are connected.

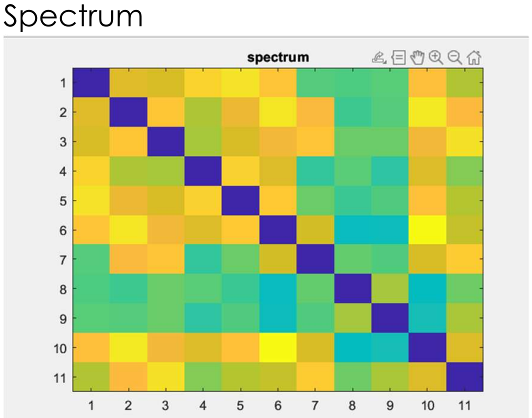
Part 3: change parameter

a) the audio features

➔ mfcc

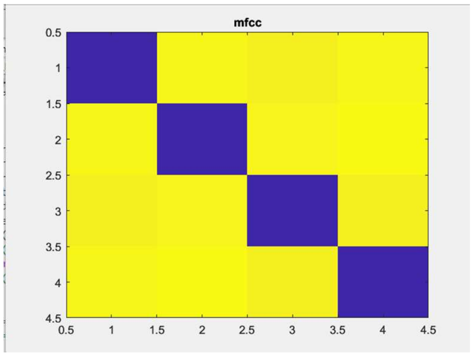


➔ Spectrum

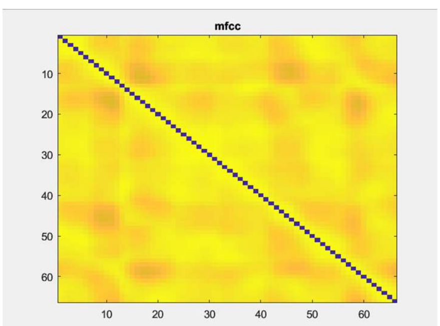


B) the frame length and hop factor,

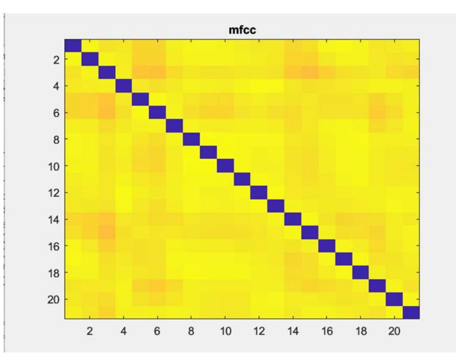
- l = 0.2, h=0.6 are presented above
- l=0.5, h=0.6



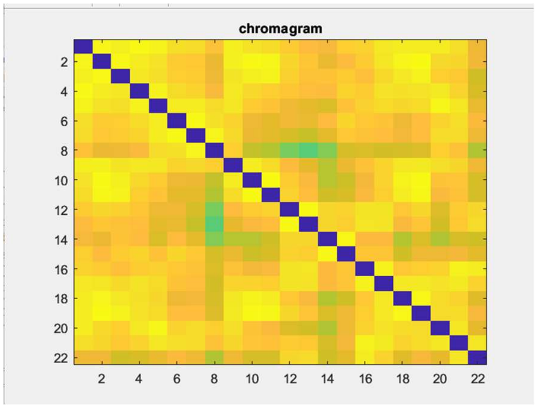
- l=0.2, h = 0.01



- l = 0.3, h = 0.2



- l = 0.2, h = 0.3

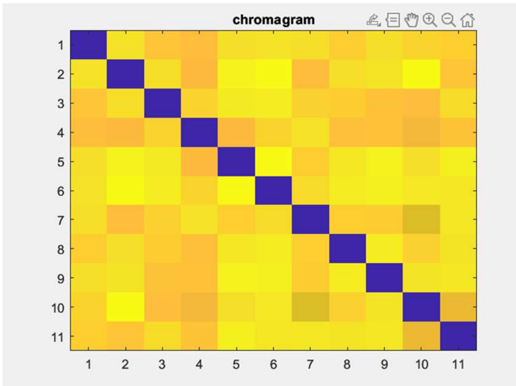


Part 4: which is best?

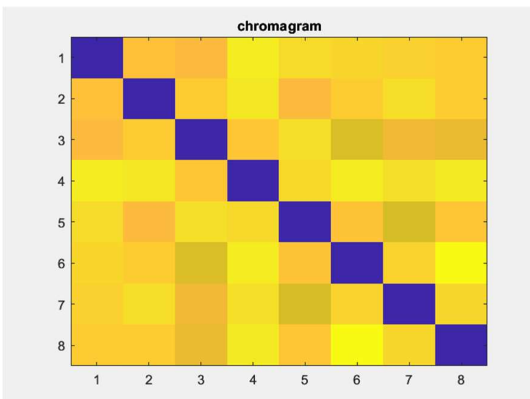
- Changing model parameters can significantly impact analysis results. For example, changing the audio feature from chromagram to mfcc or spectrum will produce a different representation of the musical content, potentially affecting the similarity matrix's accuracy. Similarly, adjusting the frame length and hop factor can have an impact on the results. A longer frame length provides a more global representation of the musical content, whereas a shorter frame length captures more local details. The hop factor regulates the overlap between adjacent frames, and increasing it results in a smoother and more consistent representation of the musical content. To determine the best combination of parameters for a specific analysis task, it is important to experiment with different parameter values and evaluate their impact on the results. This can be done through a process of trial and error, or by using optimization techniques such as grid search or random search to systematically explore the parameter space.
- Perceptual segmentation and repetition in music are complex phenomena with multiple representations. However, the chromagram feature is widely regarded as one of the most effective representations of these concepts. The chromagram is a representation of the pitch content of the music, with each bin representing a specific pitch class. This makes it especially useful for analysing melodic and harmonic patterns in music, which are frequently required for perceiving repetition and segmentation. Furthermore, the chromagram has been demonstrated to be effective at capturing higher-level musical structures such as chords, keys, and tonal centres. These structures are frequently used to generate musical phrases and sections, which are essential for perceptual segmentation. Overall, while other features such as MFCC and spectrum can also be useful for analysing perceptual segmentation and repetition in music, the chromagram is a commonly used and effective representation for these concepts.

➔ Trying on '03.wav'

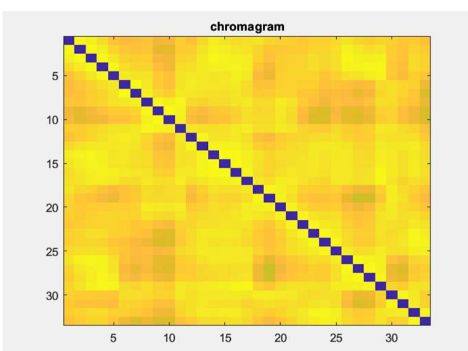
- l = 0.2, h=0.6



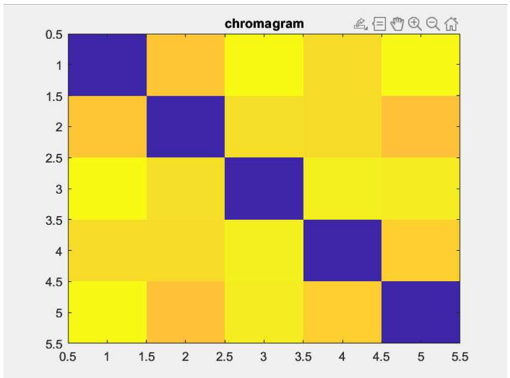
- l = 0.2, h = 0.9



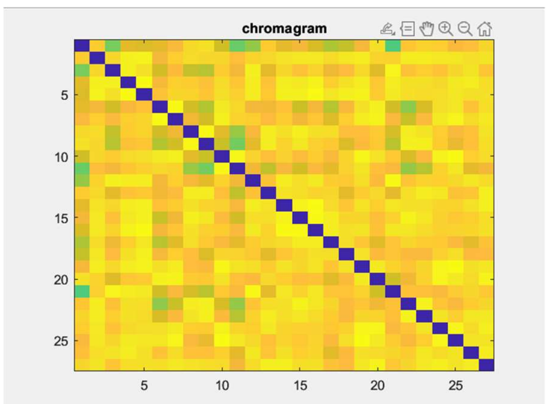
- l = 0.2, h = 0.2



- l = 0.4, h = 0.6



- l = 0.09, h = 0.6



<u>Conclusion:</u> The parameters in the Chromagram are greatly enhanced by frame length, which appears to be an important factor in the supplied parameters. Similarly, for audio features, chromagram and mfcc provide excellent feature descriptions in shorter frame lengths. The top left box in the image of the mfcc and Chromagram appears to be well separated from the chromagram, giving the impression that Chromagram is the best of them all. Shortening the frame allows us to see more and more local details in comparison to the overall picture. Values are found using hit and trial searches.

Harmonics :

Part-a,b,c:

As we implemented three melodies. Comparing part-b melody to part-a it is deep to melody. As part-c is lighter comparing to part-a.

Virtual pitch:

Part-a,b,c,d:

As in part-a it is same, but as we removed fundamentals in part-b I can hear the melody at higher strength. As per part-c I can't hear anything those are fundamentals and it is in very low audible frequency and it so called virtual pitches.