MIR DEMO/EXCERCISE

Rhythm And Meter (Part 1, 2, 3)

1. Ragtime.wav
    a. Perceived: 132
    b. Calculated: 129.4694
2. vivaldi.wav
    a. Perceived: 192
    b. Calculated: 182.4657
3. valse_triste_happy.wav
    a. Perceived: 168
    b. Calculated: 121.5778
4. laksin.wav
    a. Perceived: 112-192
    b. Calculated: 83.729 - 176.556
5. czardas.wav
    a. Perceived: 132-208
    b. Calculated: 104.9 - 170.753

The perceptual estimates and the computational estimates are the not entirely the same but we can see that they are not very far away. Some values don't match mainly because there are mixed beats in the song and not a single beat across the music.

Part 4:

Song Name	Perceived	Calculated
Derezzed	124	119.9937
Despacito	208	88.9931
Makeba	120	115.998
Rite_of_spring	173	112.0576
Stream_of_Consciousness	129	127.126

Perceived: Makeba < Derezzed < Stream_of_Consciousness < Rite_of_spring < Despacito
Calculated: Despacito < Rite_of_spring < Makeba < Derezzed < Stream_of_Consciousness

We can see that almost all the perceived estimates are equal to the calculated estimates except 'Despacito' and 'Rite_of_spring'. This is because the beats in these songs are variable across the music.

Pulse Clarity:
Derezzed: 0.82191
Despacito: 0.47385

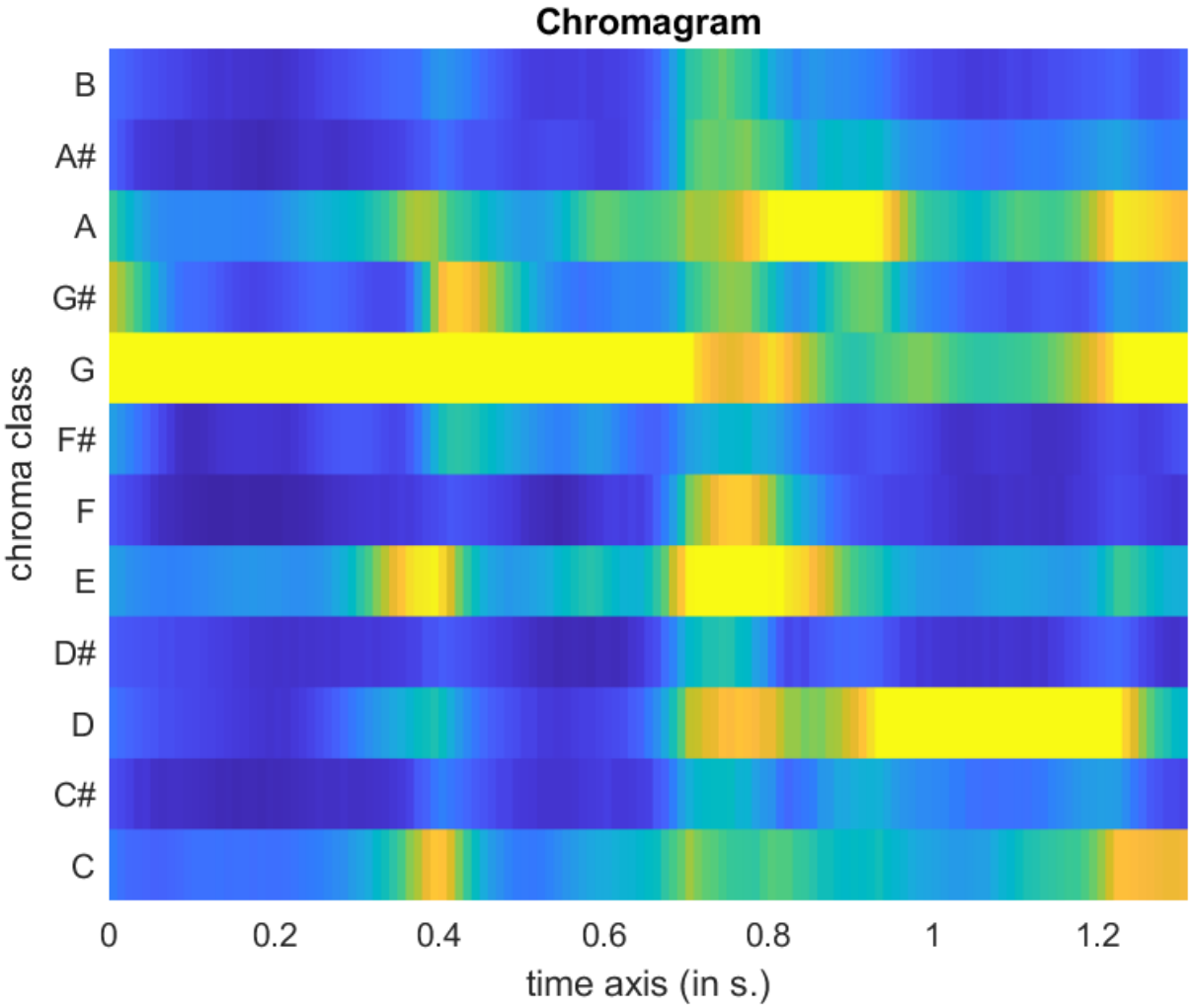
Makeba: 0.75561
Rite_of_spring: 0.052759
Stream_of_Consciousness: 0.10328

The order in terms of the tempo is given below:
Perceived: Rite_of_spring < Stream_of_Consciousness < Despacito < Makeba < Derezzed
Calculated: Rite_of_spring < Stream_of_Consciousness < Despacito < Makeba < Derezzed

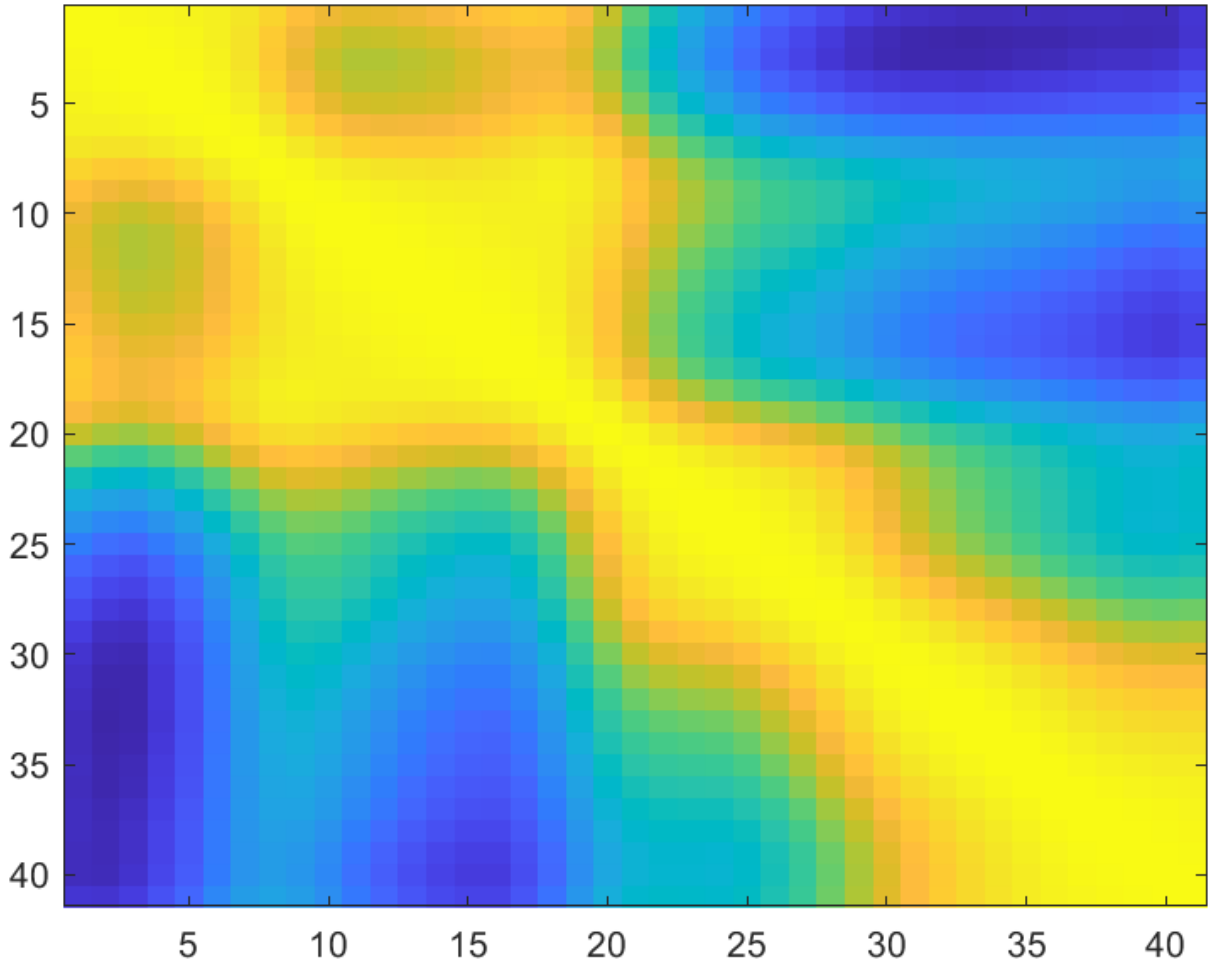
Steadier the tempo more will be the Pulse Clarity. I completely agree with the order of the files.

Repetition in Music

Chromagram related to 02.wav, of sampling rate 44100 Hz.

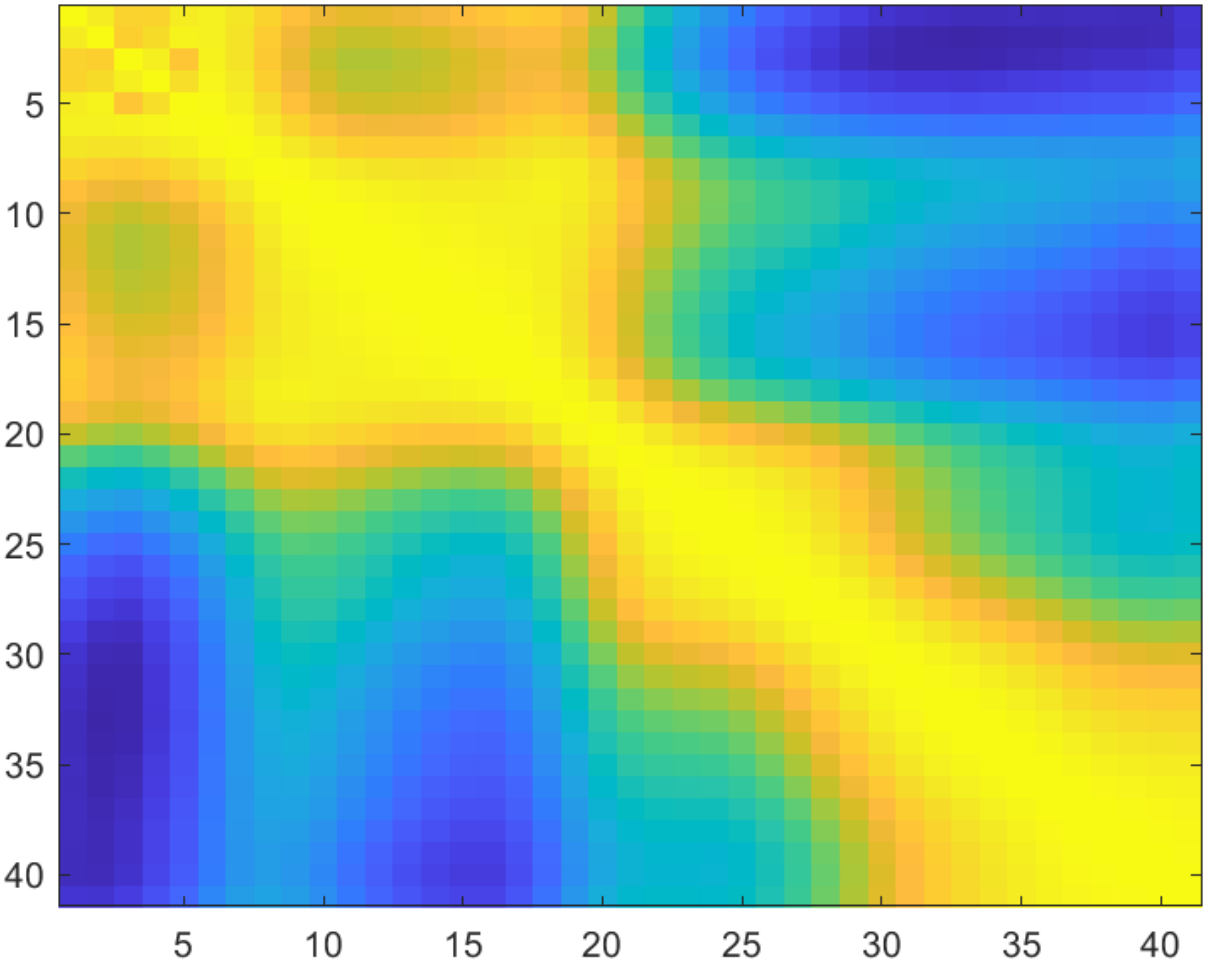


The similarity matrix of the sequence of chromagram is:

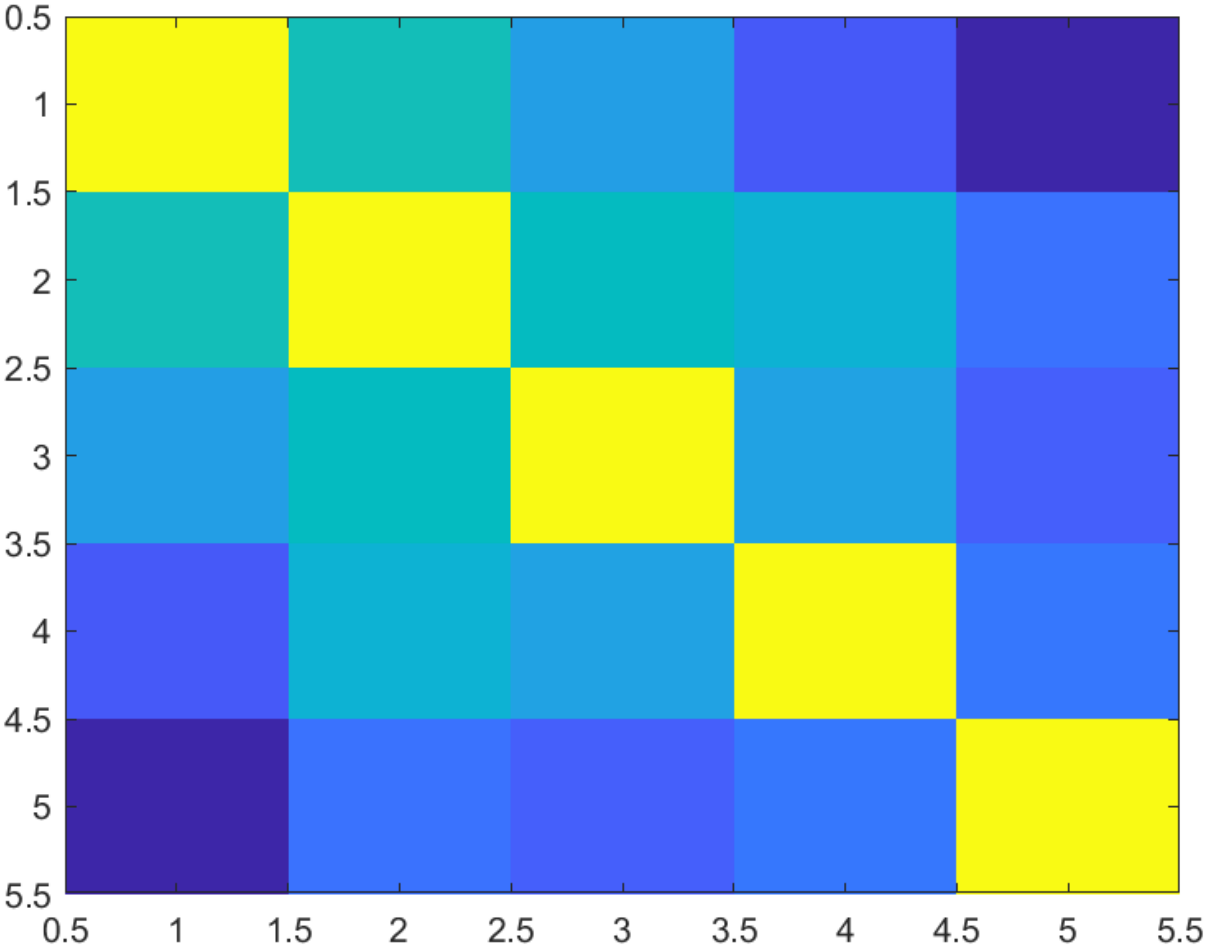


The lines in a similarity matrix represent the values of similarity or dissimilarity between pairs of objects or variables, while the checkered rectangles represent the cells in the matrix that correspond to comparisons between two different objects or variables.

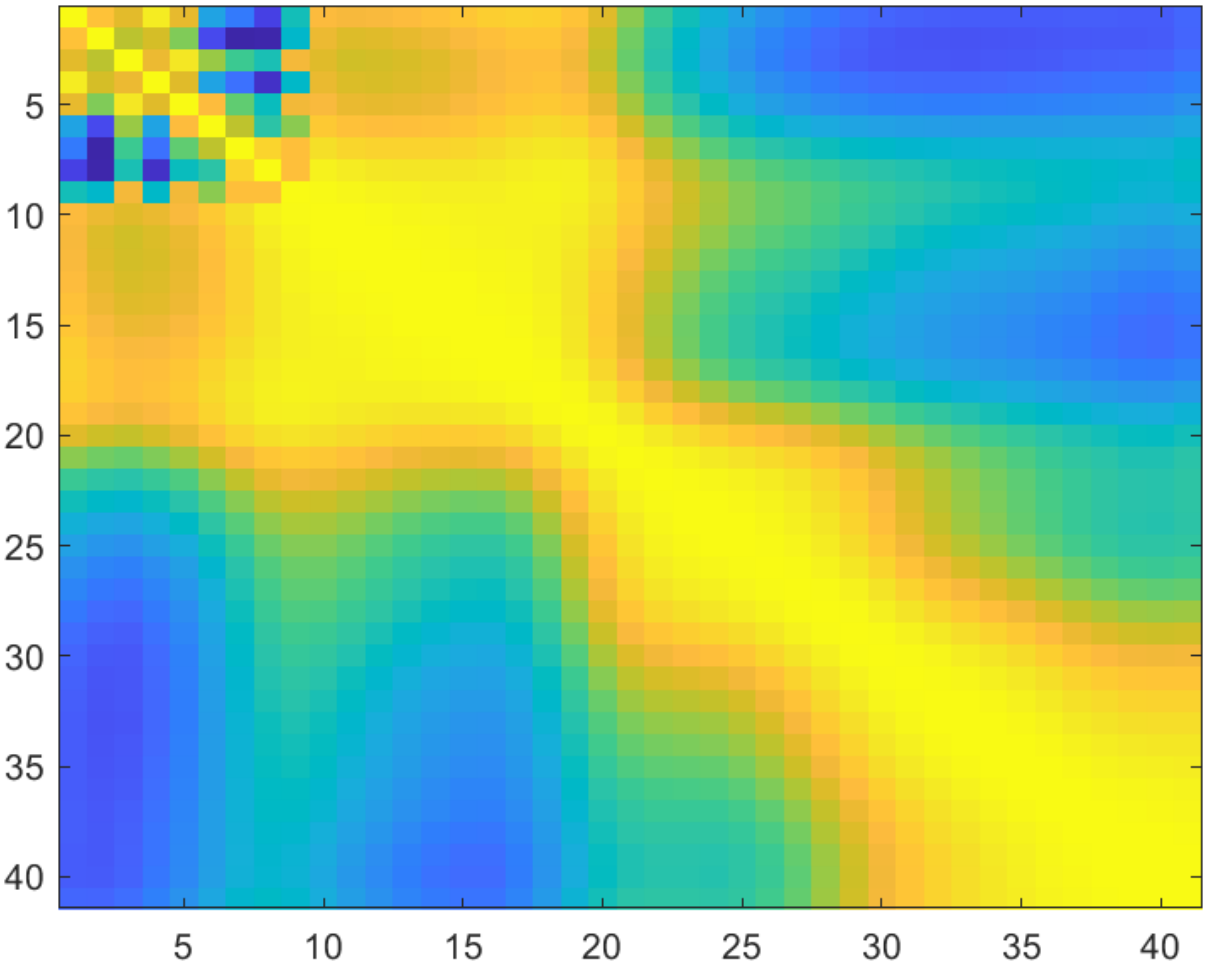
The similarity matrix of the sequence of the mfcc is:



The similarity matrix of the sequence of the spectrum is:



The similarity matrix of the sequence of chromagram is with frame length 0.3 and hop length 0.5:



Which of the features best represents your notion of perceptual segmentation and repetition?

Since perceptual segmentation involves dividing an audio signal into meaningful parts based on perceptual cues such as pitch, rhythm, and timbre, the use of appropriate **frame length and hop factor** can aid in identifying these cues by capturing temporal variations in the signal.

Finding acoustic correlation of perceptual constructs

Perceptual Construct	Acoustic Correlates
Joy	High pitch, fast tempo, loudness, and variability in pitch and loudness
Sadness	Low pitch, slow tempo, softness, and monotonicity in pitch and loudness
Moved	High pitch, slow tempo, softness, and variability in pitch and loudness
Connection	Moderate pitch, moderate tempo, medium loudness, and stability in pitch and loudness
Warmth	Moderate pitch, slow tempo, medium loudness, and stability in pitch and loudness
Chills	High pitch, slow tempo, softness, and sudden changes in pitch and loudness
Beauty	High pitch, slow tempo, softness, and smoothness in pitch and loudness

Table 1: Acoustic correlates of perceptual constructs

As for timbral features, Low energy, spectral centroid, spectral rolloff, spectral irregularity, spectral entropy, and MFCC are all features that can be used to analyze the timbre of a song.

1. Low Energy: It is a measure of the overall loudness of a signal. Low energy values can indicate softer or more muted sounds, while high energy values can indicate louder or more intense sounds.
2. Spectral Centroid: It is a measure of the "brightness" or "darkness" of a sound. Low values correspond to darker or "bassier" sounds, while high values correspond to brighter or "treblier" sounds.
3. Spectral Rolloff: It is a measure of the frequency below which a certain percentage of the energy in the signal falls. Low values correspond to sounds with more energy at lower frequencies, while high values correspond to sounds with more energy at higher frequencies.
4. Spectral Irregularity: It is a measure of the "roughness" or "complexity" of a sound. Low values correspond to smoother or more regular sounds, while high values correspond to more complex or irregular sounds.
5. Spectral Entropy: It is a measure of the amount of information or "disorder" in a sound. Low values correspond to more ordered or predictable sounds, while high values correspond to more disordered or unpredictable sounds.
6. MFCCs: Mel Frequency Cepstral Coefficients are a set of features that represent the spectral envelope of a sound. They are often used as a way to capture the timbre or "tone color" of a sound. Low values of MFCCs can correspond to smoother or more "mellow" sounds, while high values can correspond to more "sharp" or "edgy" sounds.