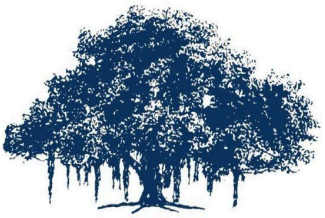

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY

HYDERABAD

Assignments 2 and 3

Music, Mind & Technology

Raj - 2020101049

Question A:

1. Harmonics:

What are the perceptual differences between all 3 versions?

Harmonics - give a lush and complete sound. They add depth and richness to the melody, making it more intricate and fully harmonized.

Odd Harmonics - This rendition seems emptier in comparison to the complete harmonic version.

Even Harmonics - This variant sounds more distinct but lacks the cozy warmth of the odd harmonics version.

Output audio files generated are melody.wav, evenmelody.wav, oddmelody.wav

2. Virtual Pitch:

What are the perceptual difference between all 3 melodies?

Melody with All Harmonics: offers a lush and complete sound experience.

Without the Fundamental Frequency: The intended note using the harmonic series may sound a bit thinner or less stable without the fundamental frequency, but the note remains recognizable.

Without both Fundamental and Second Harmonics: Reduces the melody's richness, making it even more airy. The absence of these lower harmonics could slightly affect how we hear the pitch.

Output audio files generated are melodywithoutboth.wav, melodywithoutfund.wav

Question B:

1. Rhythm and Meter:

Tempos estimated:

mozart.mp3	140.7
queen.mp3	108.9
taylor_swft.mp3	52
dream_theater.mp3	98.1
michael_jackson.mp3	185.8

Compare the computational estimates with your perceptual estimates. To what extent do they agree? What are the typical discrepancies? Can you explain the reason for them?

Variability in Tempo - Adjusting tempo assessments based on the most memorable segments of a composition is common practice. Depending on their design, algorithms may compute tempo by averaging across the entire piece, selecting the most prominent tempo, or presenting a range, resulting in varying estimates.

Beat Subdivisions - Discrepancies can arise in identifying the primary beat level being tracked. While individuals may tap along with quarter notes, algorithms might recognize eighth notes as the predominant tempo, effectively doubling the perceived beats per minute (BPM).

Perception versus Analysis - Human tempo estimations often stem from the main rhythm they perceive, such as vocals or bass lines. In contrast, algorithms analyze overall audio signal energy or specific features, which may not always align with human perception.

Frame-based Tempo Analysis:

All the results are in the “results” folder with their respective audio file namings.

What are the ranges of variation of tempi? Do they correspond to your estimates?

Frame-based analysis reveals a comprehensive spectrum of tempo variation, typically spanning from around 90 to 200, showcasing the intricate evolution of tempo within each musical piece. This analysis brings to light both anticipated fluctuations and unforeseen deviations, underscoring the significance of meticulous tempo examination in comprehending the rhythmic dynamics of music.

2. Repetition in Music:

Similarity matrix for “01.wav” is in the plots folder with the name : simmatrix01

Try to understand the link between the lines in the similarity matrix and the checkered rectangles.

Rows in the similarity matrix depict segments characterized by consistent chroma features across the entirety of the song, including recurring musical phrases, identical melodic or harmonic patterns, or portions with comparable chord progressions. Checkered rectangles within the similarity matrix denote self-replication within a segment, implying a passage that iterates, such as a recurring melodic motif or a rhythmic pattern that loops.

Try to see the impact of any change of the different parameters of the model in the final results.

Chromagram: Valuable for detecting tonal and harmonic repetitions, as it emphasizes the harmonic aspects of the music.

MFCC (Mel-frequency cepstral coefficients): Potentially more adept at capturing rhythmic and timbral similarities owing to its emphasis on spectral shape.

Spectrum: While less abstract compared to chroma or MFCC, it might be less effective for repetition detection due to its sensitivity to variations in raw frequency content.

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

The optimal audio feature varies depending on your precise interpretation of repetition:
Tonal and Harmonic Repetitions: Chromagram stands out as a favorable option because of its emphasis on harmonic content.
Rhythmic or Timbral Repetitions: Consider exploring MFCCs to effectively capture spectral shape alterations associated with rhythm and timbre.

Mel-frequency cepstral coefficients, or MFCCs, serve as coefficients utilized for analyzing the short-term power spectrum of an audio signal, with the aim of capturing the timbral characteristics of sound. They function as a representation of the spectral envelope.In terms of repetition, MFCCs excel at identifying both timbral repetitions and variations. They are proficient at recognizing instances where identical sounds recur, irrespective of alterations in melody.

For the Sample Files, the Similarity Matrices computed are given in the “plots” folder named according to their respective file names.All the codes are present in the “codes” folder.