Music, Mind, and Technology

Assignment 2 & 3

Question A

1. Harmonics

a) Check the file named “BabySharkMelody.wav” in the repo
b) Check the files named “BabySharkMelodyOdd.wav” and “BabySharkMelodyEven.wav” in the repo
c) Perceptual Differences:
  - ❖ First Ten Harmonics: produce a full, rich tone. It enhances the melody's harmonic completeness and complexity.
  - ❖ Odd Harmonics: In comparison to the complete harmonic version, it sounds more hollow, there is a lack of depth in music.
  - ❖ Even Harmonics: In comparison to the odd harmonics version, this version has the same sharpness but feels high pitched, there is also the lack of completeness.

2. Virtual Pitch

a) Check the file named “BabySharkMelody.wav” in the repo
b) Check the file named “BabySharkMelodyWithoutFundamentals.wav” in the repo
c) Check the file named “BabySharkMelodyWithoutFirstAndSecondHarmonics.wav” in the repo
d) Perceptual Differences:
  - ❖ First Ten Harmonics: produce a full, rich tone. It enhances the melody's harmonic completeness and complexity.
  - ❖ Without Fundamental Frequencies: The intended pitch using the harmonic series may seem slightly thinner or less grounded without the fundamental frequency, but the pitch remains recognisable.
  - ❖ Without First and Second Harmonics: Pitch is shaky, absence of these lower harmonics might slightly challenge pitch perception.

Question B

1. Rhythm & Meter

Estimated Tempos

Michael Jackson.mp3 : 185.3
Dream_theater.mp3 : 97.5
Mozart.mp3 : 140.87
Queen.mp3 : 109.4
Taylor_swft.mp3 : 51

Comparing Computational and Perceptual Estimates

- **Tempo Variability :** Algorithms adjust tempo estimates based on different criteria, such as the most memorable parts of a piece. Depending on their design, they may calculate an average tempo, identify the most prominent tempo, or offer a range of possibilities, resulting in varying estimates.
- **Subdivisions of the Beat :** Sometimes, there can be a discrepancy in identifying the beat's subdivision. While people may tap along with quarter notes, algorithms might detect the primary tempo at the eighth-note level, effectively doubling the perceived Beats Per Minute (BPM).
- **Perception and Analysis :** Individuals often base their tempo perception on the main rhythm they notice, such as vocals or bass lines. In contrast, algorithms analyse factors like overall audio energy or specific features, which may not always align perfectly with human perception.

Frame Based Tempo Analysis

Queen

```
Frame-based tempo estimation for queen.mp3:
  Columns 1 through 16

       NaN       NaN  165.7389   89.1906   72.4651  102.0794  100.6696  180.2921  140.2827  105.8523  185.8005  154.1060  173.2089  179.5041  105.8165  169.9539

  Columns 17 through 32

  109.1302   93.0570   67.8579  104.9087  129.6747  129.7685  189.2938  109.2828  119.5291   88.5289  151.1776  175.5485  113.6084  138.4067  115.8942  175.3542

  Columns 33 through 48

  184.3866  185.6197   64.4737  116.3070  143.3825  107.4693   85.5410  102.0041  140.1241   68.4190  104.3304  158.5252  187.5736  167.9467   84.0403  154.6358

  Columns 49 through 54

       NaN   64.0301   69.9854   67.0569   70.7021  124.0763

Tempo range for queen.mp3: 64.03 BPM to 189.29 BPM
```

Dream Theatre

```
Frame-based tempo estimation for dream_theater.mp3:
  Columns 1 through 13

   71.1799  110.9390   78.1314  125.4365  106.5530   84.1882   69.0160   92.1836  146.2523  138.5265  153.1815  150.2326  141.9632

  Columns 14 through 26

  150.0248   76.6531   95.7456   88.5942   81.5238  120.6125  126.3113   68.8161   72.8922  137.7237   96.8539   82.6311  145.9193

  Columns 27 through 39

  155.0477  154.5767  131.0708  146.6213  140.5465   98.0181  100.7735   95.9487   94.9041  142.5918  152.5514  151.3096  111.2088

  Columns 40 through 52

   99.1476   97.1013  117.1472  140.4897  101.8472   73.4856  147.3464   98.1267   94.2490  120.7373  149.5203  147.3038   76.4575

  Columns 53 through 65

  144.3756  120.2658  101.2323  138.2124  156.5538  149.4374   76.0106   98.0782   96.9944  145.5715  130.7484  103.1375   76.2867

  Columns 66 through 78

  126.1050  100.3990  106.6444  101.1405  119.3586  140.7051   99.3145  154.2828  134.5052  141.3732  154.4601   76.2022   92.3088

  Column 79

  126.7454

Tempo range for dream_theater.mp3: 68.82 BPM to 156.55 BPM
```

Mozart

```
Frame-based tempo estimation for mozart.mp3:
  Columns 1 through 16

       NaN   92.8471   87.0615   94.1806  152.1106  126.6265  121.6295  139.0951  139.3442  116.4351  144.4288  144.1041  132.5894  133.6264  138.1661   70.3972

  Columns 17 through 32

  105.4952  132.8464   96.4018  127.7956   87.4697  134.9029   70.8715  137.6001  139.6443  115.6584   92.5978  141.5199  147.4749  190.5861  131.2717  139.6546

  Columns 33 through 48

  108.3870  117.5749  144.2105  142.3128  138.0282  129.4539  175.2860  148.4586   67.5594  138.6758   95.5983   74.8396   95.5138   98.1503   76.0074  137.5794

  Columns 49 through 64

  137.2103  107.4915   81.2016  162.2156   87.8882   69.5607   90.6560  135.4164  135.4264  103.2424  110.1417  123.4372  126.8864   92.1762  134.0219  127.6197

  Columns 65 through 80

  133.0472  158.5390  130.9809  141.8747  113.0017  145.7180  125.3478  113.5459  123.4421  137.8957  139.6044  146.9578   91.0654   68.6793  142.9082   95.3737

  Columns 81 through 87

  142.9102  113.8671  110.9945   68.2532  139.2242  126.5932  162.9815

Tempo range for mozart.mp3: 67.56 BPM to 190.59 BPM
```

Taylor Swift

```
Frame-based tempo estimation for taylor_swft.mp3:
  Columns 1 through 16

   68.6681   68.7111   69.3922   68.7252   71.0158   69.2951   69.5159   68.6865   68.6668   68.7161   68.6846   68.7085   69.4361  196.5211  194.1755  106.7896

  Columns 17 through 32

   69.4382   68.6700   68.6555   68.6444  194.1277  159.0854   67.9638   69.9369  163.3715  100.7321   68.6609   68.7137   69.4576   68.7207   67.9485   68.6944

  Columns 33 through 48

   68.6382   68.7047   72.1150   68.6405  111.0608  106.5146   68.0496   68.5219   68.7207   68.6869   68.7831  164.4226   69.7604   91.3298   67.9952  104.8730

  Columns 49 through 64

  194.0904  159.1185  103.3390   78.4435  106.9749  108.7866  108.8555  123.3659  173.4872  104.0484  162.3167   75.8094  163.6690  130.8344   68.5017  115.3152

  Columns 65 through 70

  147.8089   68.4923   69.4226   68.6783   70.2083   87.9778

Tempo range for taylor_swft.mp3: 67.95 BPM to 196.52 BPM
```

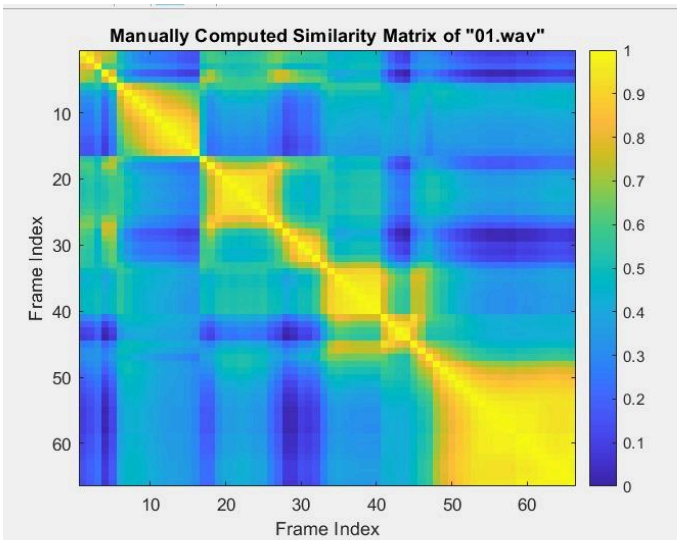
Michael Jackson

```
Frame-based tempo estimation for michael_jackson.mp3:
  Columns 1 through 13

  188.3125  188.4588  124.8305  184.8877  185.9030  123.0075  125.4767  187.7798  122.6227  122.9047  185.5106  123.8951  184.5043

  Columns 14 through 26

  129.2957  185.2995  187.0300  185.6472  186.1400  187.9981  124.4553  121.4391  182.5347   92.6473   71.9876  127.2715  184.1347

  Columns 27 through 39

  185.2511  124.8060  183.9211  124.6954  188.1954  182.5607  187.5344  172.8571  185.2812  188.0857  188.7543  180.4594  186.8263

  Columns 40 through 52

  188.7673  124.2171  187.8702  186.8575  187.8991  181.1357   91.8867  181.4710  184.8656  187.4574  124.3154  184.8817  185.5789

  Columns 53 through 62

  186.3772  183.3683  185.8956  122.5644  190.2065  147.0908  186.6637  124.2986  184.5280  185.6615

Tempo range for michael_jackson.mp3: 71.99 BPM to 190.21 BPM
```

The ranges of variation of tempi observed in frame-based analysis offer a detailed view like around 90 to 200 of how tempo evolves within each piece, highlighting both expected and unexpected variations. These ranges and our initial estimates vary, emphasizing the value of detailed tempo analysis for an understanding of rhythm dynamics in music.

2. Repetition in Music

Similarity Matrix for “01.wav”



Link Between Lines and Checkered Rectangles in Similarity Matrix:

- Lines in the similarity matrix represent segments with similar chroma features throughout the song, such as recurring musical phrases, identical melodic or harmonic patterns, or sections with similar chord progressions.
- Checkered rectangles in the similarity matrix indicate self-similarity within a segment, suggesting a section that repeats itself, like a repeated melodic motif or a rhythmic pattern that loops.

Impact of Model Parameters on Results:

Chromagram: Useful for identifying tonal and harmonic repetitions as it focuses on the harmonic content of the music.

MFCC (Mel-frequency cepstral coefficients): Might be more suitable for capturing rhythmic and timbre similarities due to its focus on spectral shape.

Spectrum: Less abstract than chroma or MFCC, it might be less suitable for repetition detection because raw frequency content can be more sensitive to variations.

Choosing the Best Feature:

The best audio feature depends on your specific definition of repetition:

- Tonal and Harmonic Repetitions: Chromagram is a good choice due to its focus on harmonic content.

- Rhythmic or Timbral Repetitions: Experiment with MFCCs to capture spectral shape changes related to rhythm and timbre.

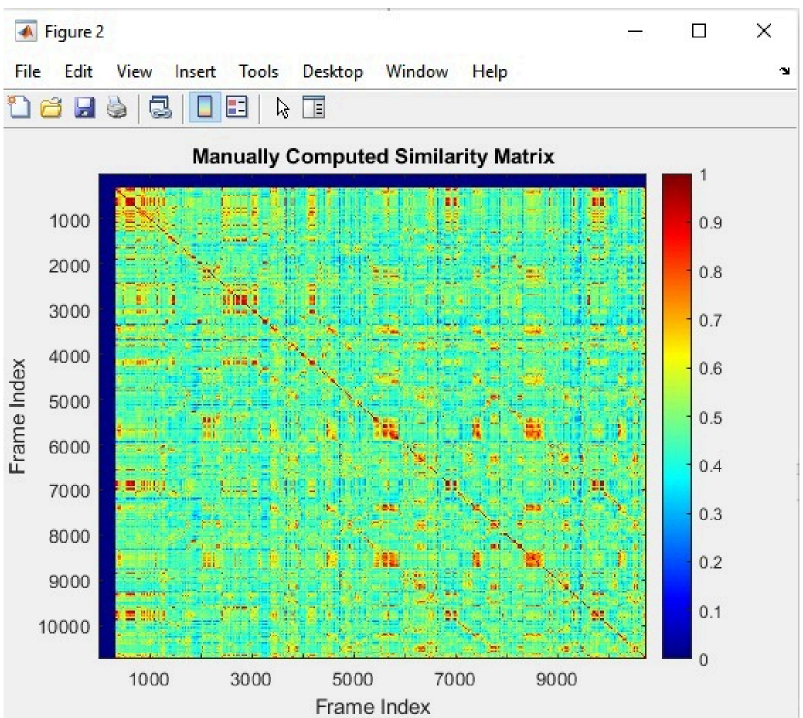
Overall:

Mel-frequency cepstral coefficients, or MFCCs, are coefficients that are used to analyze the audio signal's short-term power spectrum in order to capture the timbre aspects of sound. They act as a model for the spectral envelope.

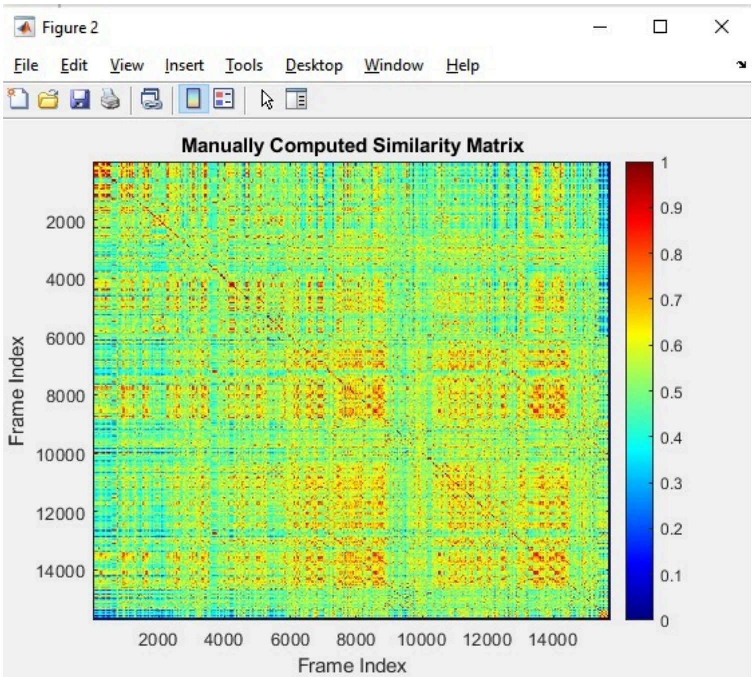
Repetition: MFCCs are especially good at identifying timbral repeats and variations; they are also good at identifying situations in which the same sounds return, regardless of changes in melody.

Similarity Matrices For Sample Files

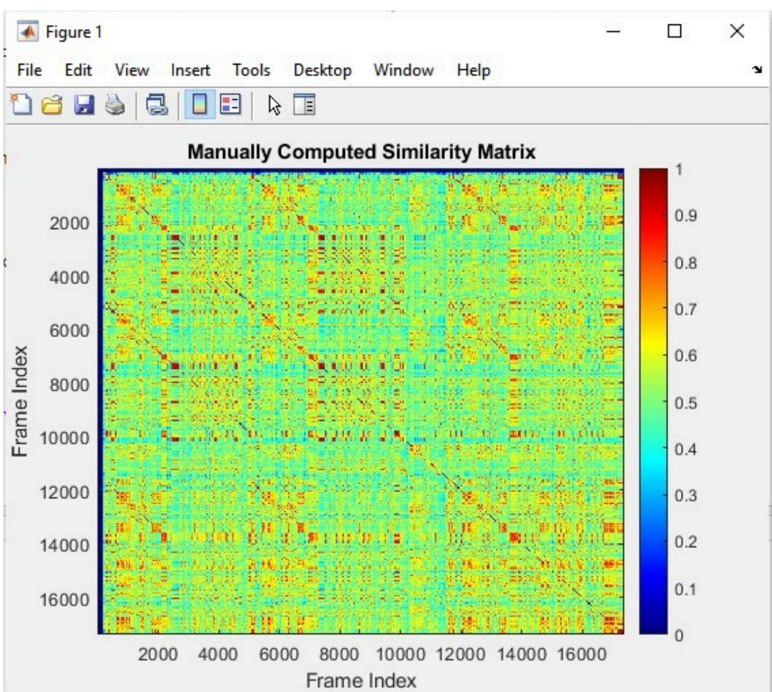
Queen



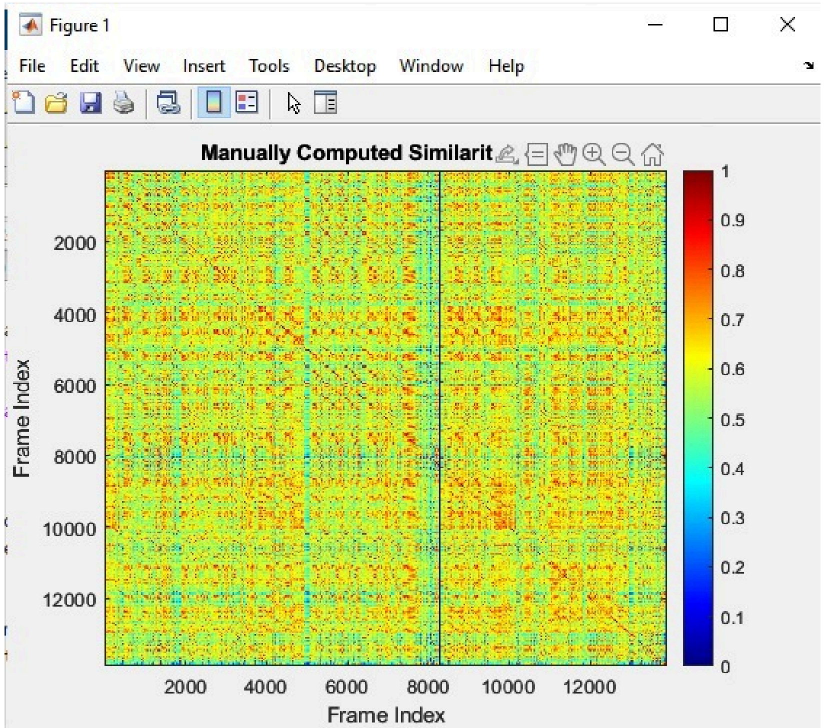
Dream Theatre



Mozart



Taylor Swift



Michael Jackson