# Exercise 2

## Task 1

| | Meyda Audio Features | Justification |
|---|---|---|
| Sound 1 - Ex2_sound1.wav | Spectral Centroid | Spectral Centroid provides an indicator of the "brightness" of a sound. This can be detected in this audio as when the audio changes its tune, from a soft to a loud one, it is louder, and thus 'brighter'. By checking the console, this can be seen in the sudden increase in the value of the Spectral Centroid, from around 7 to 17. (The overall range fluctuates from 6 to 17. The higher the number, the brighter). From my p5.js application itself, this can be seen from the increase of width of my rectangle. However, this audio still has a relatively low Spectral Centroid. |
| Sound 1 - Ex2_sound1.wav | ZCR (Zero Crossing Rate) | ZCR tells us the number of times the signal crosses the buffer's zero value. It aids in the distinction between percussive and pitched sounds. Percussive sounds have a variable ZCR across buffers, whereas pitched sounds have a more consistent value. This audio has a percussive sound as checking the console, the ZCR value ranges from as low as 6 all the way to 30. From my p5.js application itself, this can be seen in a sudden increase of width of my rectangle then going back to a slimmer profile |

| | | |
|---|---|---|
| | | (smaller width), cementing the fact that it has a wide range of values. |
| Sound 1 - Ex2_sound1.wav | Energy | The Energy is the infinite integral of the squared signal. This is yet another indicator of the signal's loudness. By checking the console, this can be seen in the sudden increase in the value of the Energy from around 0.18 to 2.05. From my p5.js application itself, this can be seen from the very slight increase of width of my rectangle before returning to its original width. Not a noticeable increase, but from hearing the audio itself, it is easy to tell that the loudness changes from rather soft to loud then going back to soft again. |

| | **Meyda Audio Features** | **Justification** |
|---|---|---|
| Sound 2 - Ex2_sound2.wav | Energy | The Energy is the infinite integral of the squared signal. This is yet another indicator of the signal's loudness. As we hear the sound, it starts rather softly before the loudness increases to the maximum and afterwards going back to a relatively quieter tone. By checking the console, this can be seen a slow rise of an energy value 0.21 to a maximum of 28.11 before dropping back down again to 0.467. From my p5.js application itself, this can be seen from the gradual increase of width of my |

| | | |
|---|---|---|
| | | rectangle to its maximum before returning to a shorter width as the loudness drops. |
| Sound 2 - Ex2_sound2.wav | ZCR (Zero Crossing Rate) | ZCR tells us the number of times the signal crosses the buffer's zero value. It aids in the distinction between percussive and pitched sounds. Percussive sounds have a variable ZCR across buffers, whereas pitched sounds have a more consistent value. Overall, I feel this audio file produces a percussive sound at first, and a pitched sound towards the end. By checking the console, this can be seen a relatively constant value of 0 to 8. And then it increases all the way to 262 before dropping back down again to 28. |
| Sound 2 - Ex2_sound2.wav | Spectral Flatness | Spectral Flatness measures the flatness of the spectrum. It is calculated by dividing the geometric and arithmetic means. It determines how noisy a sound is. A pure sine wave, for example, has a flatness of approaching 0.0, while white noise has a flatness that approaches 1.0.<br>This audio file itself is not flat. This can be seen from the console, it is not flat as the value ranges at around 0.006 to 0.01, which is approaching zero. However, the value increases to as high as 0.62 as the sound file approaches the end before dropping back to 0.32. What we can conclude |

| | | |
|---|---|---|
| | | from this is that at the start, it is not flat as the value is near to 0.0. However, the sound progresses, it gets noisier, and starts increasing and therefore it is becoming quite flat, before going back to not flat as the value drops back. |

| | **Meyda Audio Features** | **Justification** |
|---|---|---|
| Sound 3 - Ex2_sound3.wav | Spectral Centroid | Spectral Centroid provides an indicator of the "brightness" of a sound. The range of values is relatively quite small, fluctuating from 31 to 45. Overall has a low Spectral Centroid. |
| Sound 3 - Ex2_sound3.wav | ZCR (Zero Crossing Rate) | ZCR tells us the number of times the signal crosses the buffer's zero value. It aids in the distinction between percussive and pitched sounds. Percussive sounds have a variable ZCR across buffers, whereas pitched sounds have a more consistent value. Overall, I feel this audio file produces a pitched sound. By checking the console, this can be seen a relatively constant value which fluctuates from 50 to around 80. (The default ZCR range is 0-255) |
| Sound 3 - Ex2_sound3.wav | Spectral Flatness | Spectral Flatness measures the flatness of the spectrum. It is calculated by dividing the geometric and arithmetic means. It determines how noisy a sound is. A pure sine wave, for example, has a flatness of approaching 0.0, while |

| | | |
|---|---|---|
| | | white noise has a flatness that approaches 1.0. This audio file itself overall is not flat. This can be seen from the console, it is not flat as the value ranges at around 0.002 to 0.052, which is approaching zero. |

# Task 2

*Spectral Flatness*
Spectral Flatness measures the flatness of the spectrum. It is calculated by dividing the geometric and arithmetic means. It determines how noisy a sound is. A pure sine wave, for example, has a flatness of approaching 0.0, while white noise has a flatness that approaches 1.0.
(Range 0.0 – 1.0; where 0.0 is not flat, meanwhile 1.0 is very flat)

This audio file itself overall is not flat. This can be seen from the console, it is not flat as the value ranges at around 0.005 to 0.072, which is approaching and closer to zero. There is however instances where the values go above 0.1 and hits a high spike of value 0.135 but for the most part the song is in the range of 0.005 to 0.072. Overall, I can conclude that this audio file is not flat.

This can be seen from my p5.js application, this audio feature is represented by a black rectangle on the centre of my canvas. It stayed mostly black as the values remain constant from the range of 0.005 to 0.072 as mentioned above. However, there are instances where it hits a spike of value more than 0.1, and hence making it turn green in colour.
I have also multiplied the value of Spectral Flatness by 999 so the changes in values can be visually seen and represented by the black and flashes of green.

*ZCR – Zero Crossing Rate*
ZCR tells us the number of times the signal crosses the buffer's zero value. It aids in the distinction between percussive and pitched sounds. Percussive sounds have a variable ZCR across buffers, whereas pitched sounds have a more consistent value.
(default ZCR range is 0 - 255)

This audio file itself is deemed to have a pitched sound. This can be seen from the console, as the values ranges consistently from 4 to 28 for most part of the file. However, there are some occasions where it can reach to a high value of 132.

This can be seen from my p5.js application, as this audio feature is represented by the 12 white squares orbiting the rest of the shapes in the centre. For this, I will be adjusting the rotation of the square. I have put a negative sign so it will be rotated to the left. I have divided it by 100 so the rotations will be a lot more visible to the human eyes. If I did not, the squares will keep on spinning non-stop, which defeats the purpose of having a visual effect linked to the audio feature, which is ZCR in this case. As the song plays, you can see that the squares have been rotated for only quite a small angle, which represents the constant values ranging from 4 to 28. However, as mentioned above, we can observe that there are instances where it hits a spike of value more than 100. That translates to a bigger angle rotation made by all the squares.

A signal's ZCR can be viewed as a measure of its noisiness. In the event of noisy signals, for example, it frequently has greater values. It is also known to reflect the spectral features of a signal in a pretty coarse manner.

*Spectral Kurtosis*

Spectral Kurtosis is a measure of how pointy the spectrum is. It can be deemed as the opposite of Spectral Flatness. It is often used to describe a sound's "pitchiness / tonality." Ranged from 0.0 to 1.0, where 1.0 being very tonal whereas 0.0 being not tonal.

This can be seen from the console, where the values ranges widely from as low as 6 to as high as 162. This indicates a huge spread of value. However, there are some occasions where it can reach to a negative value of as low as -38.

This can be seen from my p5.js application, as this audio feature is represented the floating red square. For this, I will be adjusting the colour of the square. I have multiplied it by 25 to amplify the effect so it will be easier to visualise as the range of the values are so spread out in this case. As the values are widely spread out, whenever the square is red in colour, it means that it is displaying a positive value. However, on the other hand, if it turns black, the values go to the negative region. There are instances where it turns black to red and black again repeatedly, meaning that the values fluctuate from negative to a positive and back to being negative.

*Root Mean Square (RMS)*

RMS is the waveform's root mean square, which corresponds to its loudness. It helps to get a basic notion and rough idea of how loud a signal is.

From the console, the values has a wide range. Starting at a very low number of 0.0012 to as high as 0.302.

This can be seen from my p5.js application, as this audio feature is represented by the blue square at the centre of the canvas. For this, I will be adjusting the size of the square. I have multiplied the value of RMS by 1000 as the values are quite small. This is to aid in bringing the visual to the table, so we can see clearly observe the audio feature being displayed here.

As the audio file starts from being soft, with a low value of 0.0012, the size of square will remain rather small. As the beat changes and it gets louder, the size of square pops out and gets larger as the values goes to more than 0.1 and hits a high of 0.302. The song has parts to it where it is soft while other parts which are loud, which explains the fluctuating square size.

## Further Development

I have implemented one of the further development ideas. I have added p5.speech to my application, so it will be able to recognise my voice. Whenever I say a colour, it will detect what I said and it will change the colour of the background accordingly.
Colours include: White, Red, Orange, Yellow, Green, Blue, Purple, Pink

I have also implemented another feature whereby the original shapes are able to transform to other shapes just by us saying it. Shapes included are: Circle, Pentagon, Square and Triangle

**Coursera generated link:**
https://hub.labs.coursera.org:443/connect/sharedszmlboub?forceRefresh=false&path=%2FvMBxTlyvUcM5DXb37Upsye6EhaJPW4z7oTLHiHdIrBJcLsAIkbpg8LFmKA6qdMB0%2F

However, the generated link above does not work. Need to trim URL, remove everything else after false.

This is the working link:
https://hub.labs.coursera.org:443/connect/sharedszmlboub?forceRefresh=false