CM3065 Intelligent Signal Processing

Exercise 2 Report

LIM WEE KIAT

Task 1

	Meyda audio features	Justification
Sound 1		
	RMS	- Root mean square(RMS) is one of the time-domain features for Meyda audio feature extraction. - The root mean square of a waveform is calculated in the time domain to indicate its loudness. - For the sound 1, we can use RMS to have a rough idea of the loudness of the signal.
	ZCR	- The zero-crossing rate (ZCR) is the rate at which a signal changes from positive to zero to negative or from negative to zero to positive. - For the sound 1, ZCR can help classify percussive and pitched sounds. - Percussive sounds will have random ZCRs in the buffer, where pitched sounds will return a more constant value.
	Energy	- The energy of a signal corresponds to the total magnitude of the signal. - It is an infinite integral of a squared signal. - For the sound 1, energy is another indicator of signal loudness.
Sound 2		
	RMS	- The root mean square of a waveform is calculated in the time domain to indicate its loudness. - For the sound 2, we can get a rough idea of the loudness of the signal.
	Loudness	- Loudness is the subjective perception of sound pressure. - This is because humans’ perception of frequency is non-linear. - Loudness is one of the perceptual features. - For the sound 2, we can use loudness features to build filters that better match human perception of loudness.

	Energy	- The energy of a signal corresponds to the total magnitude of the signal. - It is an infinite integral of a squared signal. - For the sound 2, energy is another indicator of signal loudness that roughly corresponds to how loud the signal is.
Sound 3		
	ZCR	- The zero-crossing rate is the rate of significant changes along with a signal, i.e., the rate at which the signal changes from positive to negative or back. - This feature has been used heavily in both speech recognition and music information retrieval. - For the sound 3, ZCR can help classify percussive and pitched sounds.
	Spectral Kurtosis	- Spectral kurtosis is a parameter in the frequency domain indicating how the impulsiveness of a signal varies with frequency - It is an indicator of how sharp the spectrum is. - For the sound 3, we can use spectral kurtosis to denote the ‘pitch/tone’ of a sound.
	Spectral Centroid	- It indicates where the ” center of mass” for a sound is located and is calculated as the weighted mean of the frequencies present in the sound. - If the frequencies in music are same throughout then spectral centroid would be around a centre and if there are high frequencies at the end of sound then the centroid would be towards its end. - For the sound 3, we can use spectral centroid to quantify the “brightness” of a sound.

Task 2

For the audio visualization, I had chosen some meyda extraction features to make the audio visualization more dynamic. Firstly, the audio source is a remix song, and it will be loudness, brightness, and so on. In this case, I had used the meyda features extration such as ‘zcr’, ‘rms’, ‘energy’, ‘spectralCentroid’, and ‘spectralSpread’ in order to make the audio visualisation of the sound source as shown in figure 1.0.1.

```
function setup() {
    createCanvas(1000, 700);
    background(180);
    angleMode(DEGREES);

    playStopButton = createButton('play');
    playStopButton.position(200, 20);
    playStopButton.mousePressed(playStopSound);

    fft = new p5.FFT(0.2, 2048);

    if (typeof Meyda === "undefined"){
        console.log("Meyda could not be found!");
    }else{
        analyzer = Meyda.createMeydaAnalyzer({
            "audioContext":getAudioContext(),
            "source":mySound,
            "bufferSize":512,
            "featureExtractors":["rms","zcr","energy","spectralCentroid","spectralSpread"],
            "callback":features =>{
                console.log(features);
                borderSize = features.rms*100;
                rectColour = features.zcr*3,
                rectBorderColour = features.energy*20,
                opacity = features.spectralCentroid,
                rectSize = features.spectralSpread;
            }
        })
    }
}
```

Figure 1.0.1 Meyda features extration(setup function)

Audio Features	Justification
RMS	Corresponds to the remix song’s loudness.
ZCR	Helps differentiating between percussive and pitched sounds.
Energy	Another indicator of the loudness of the remix song.
Spectral centroid	Quantifying the “brightness” of the remix song.
Spectral spread	Used to differentiate between noisy (high spectral spread) and pitched sounds (low spectral spread).

For the loundness, RMS and energy is the best meyda features extration to done the audio visualisation. For the pitch, ZCR can help distinguish percussive and pitched sounds. Percussion sounds will have random ZCRs in the buffer, where pitched sounds will return a more constant value.

Output of audio visualization:

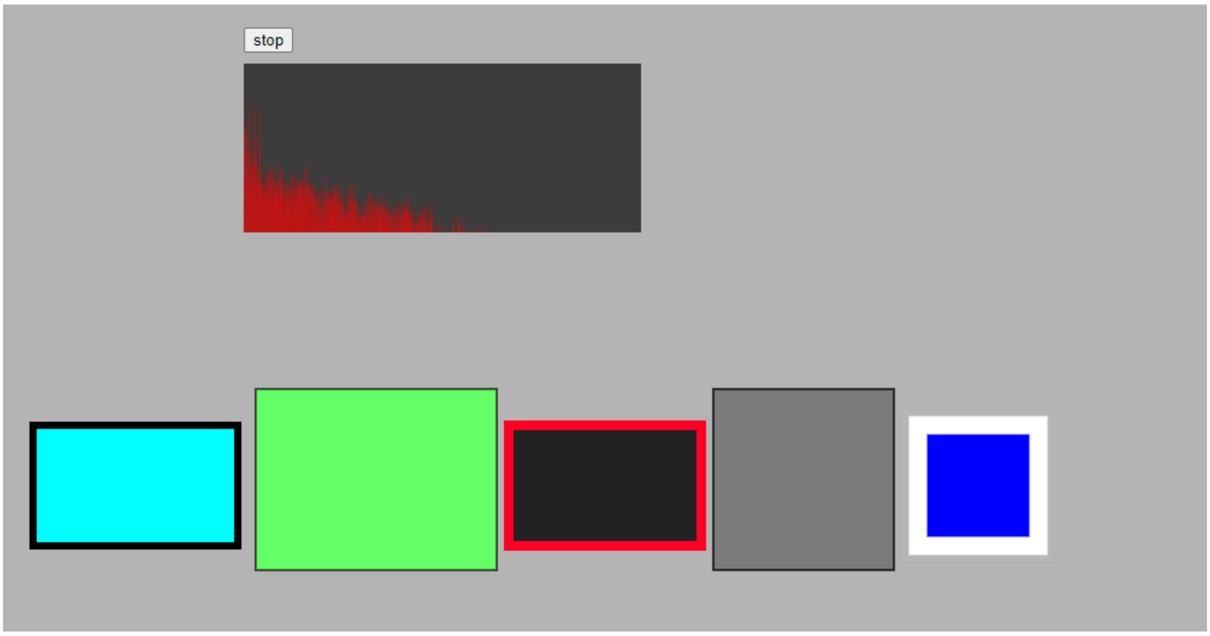

Figure 1.0.2 Output of audio visualisations

For the rectangle 1, it represents the ‘rms’ of meyda features extractions. If the song have high loudness, the border size of the rectangle 1 will become thick.

For the rectangle 2, it represents the ‘zcr’ of meyda features extractions. If the song has a high pitch, the rectangle colour will be changed with the pitch.

For the rectangle 3, it represents the ‘energy’ of meyda features extractions. If the song has high loudness, the border colour of the rectangle 3 will be changed with the loudness.

For the rectangle 4, it represents the ‘spectralCentroid’ of meyda features extractions. If the ‘brightness’ of the song changes over time, the rectangle fill colour opacity will be affected.

For the rectangle 5, it represents the ‘spectralSpread’ of meyda features extractions. The border size of rectangle 5 will be changed with the song’s noisy and pitched sounds.