多媒體作業2程式碼說明

F74104040\_資訊系\_黃芊

Problems:

1. 利用python librosa 套件將.mp3檔讀進來，會得到audio signal 和 sample rate，再利用matplotlib 將圖畫出來
2. 利用python librosa 套件中的librosa.feature.rms 得到root mean square的array，也就是平均平方數，，energy 的概念是隨著時間將一定長度的frame中的值作平方並相加，得到能量「大小」，沒有正負，而這邊我用的方法是用root mean square 的方式，講義上則是取平方相加而已，曲線較不圓滑，但還是有做呈現。
3. 利用python librosa 套件中的librosa.feature.zero\_crossing\_rate 得到zero crossing rate 的 array，它的原理一樣是audio source在一定frame的長度中隨著時間去計算和0交集的比率，這邊我用zcr這個變數去存成一個array並畫出來。
4. 做end-point detection 時，我利用我取每10ms當frame length，一樣計算energy(array)，然後設一個energy threshold，如果能量強度大於臨界值，將這個時間的記下來，然後遇到下次低於臨界值的值時，將這段時間標記為一段。
5. 利用python librosa 套件中的librosa.pyin可以得到音檔中人聲音高的估計值，這個套件所運用的是YIN演算法，此演算法是基於autocorrelation的方法去偵測音高，最後我將此函數第一個回傳值(time series of fundamental frequencies in Hertz)存成array畫出來。
6. 利用librosa.stft的函式計算Short-time Fourier transform，再將amplitude轉換成db用spectrogram的方式畫出來。