Project-- Speech Processing

1. Please use the extracted audio file from Project 1 for this project.

2. Extract the speech features in Time Domain

(1) Waveform

(2) Energy contour

(3) Zero-crossing rate contour

(4) End point detection

(5) Pitch contour

3. Calculate the spectrogram of the extracted audio data in Frequency Domain

4. Present the program code, the methods you used and the results

5. Submission due: Please upload your program and the results to the ftp site within two weeks after the date of project assignment. The ftp site

can be found at the course website.

系級：資訊系114. 學號：H34101096 姓名：陳廷笙

**2.(1) Waveform**

**Code:**

一張含有 文字, 字型, 螢幕擷取畫面 的圖片

自動產生的描述

先透過audioread函式讀取Project1的audio file，檔名為Audio.mp3，讀取後透過audioread函式return sample data並儲存到y，return sample rate並儲存到Fs。

接著計算time，把各個y的值，都和Fs相除。

並用figure把圖表畫出，x軸為Time，單位為秒(s)，y軸為Amplitude。

**Output:**

一張含有 螢幕擷取畫面, 繪圖, 文字, 圖表 的圖片

自動產生的描述

x軸為Time，單位為秒(s)

y軸為Amplitude

**2.(2) Energy contour**

**Code:**

**一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述**

將一個window的time設定成0.02秒，並計算window\_length代表了一個window的length，window\_num代表的是可以把音檔分成幾個window。

並透過for迴圈，從第一個window跑到最後一個window，start代表每個window的第一個位置，stop代表每個window的最後一個位置，並以教授上課ppt的公式先將x(n)乘上h(n)後，再將結果平方。

將該window內所包含的平方結果加總起來，就可以得到該window所包含的energy了。

並用figure把圖表畫出，x軸為Time，單位為秒(s)，y軸為Energy。

**Output:**

**一張含有 文字, 螢幕擷取畫面, 繪圖, 圖表 的圖片

自動產生的描述**

x軸為Time，單位為秒(s)

y軸為Energy

**2.(3) Zero-crossing rate contour**

**Code:**

**一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述**

第一個for迴圈將從第一個window跑到最後一個window，start代表每個window的第一個位置，stop代表每個window的最後一個位置。

並依序計算各個window的zero crossing值，計算完zero crossing值後，再將zero crossing除window length。

最後再將每個window儲存至zero crossing rate的matrix中。

x軸為Time，單位為秒(s)，y軸為zero crossing rate。

**Output:**

一張含有 文字, 螢幕擷取畫面, 繪圖, 行 的圖片

自動產生的描述一張含有 文字, 螢幕擷取畫面, 繪圖, 行 的圖片

自動產生的描述

x軸為Time，單位為秒(s)

y軸為zero crossing rate

**2.(4) End point detection**

**Code:**

**一張含有 文字, 螢幕擷取畫面, 字型 的圖片

自動產生的描述**

計算出ITU後，利用start來儲存音檔中第一個window energy超過ITU的位置，利用stop來儲存音檔中最後一個window energy超過ITU的位置。

並在start和stop的對應時間畫線。

x軸為Time，單位為秒(s)，y軸為Amplitude。

**Output:**

**一張含有 文字, 螢幕擷取畫面, 繪圖 的圖片

自動產生的描述**

x軸為Time，單位為秒(s)

圖中較左的線為start的時間

圖中較右的線為stop的時間

**2.(5) Pitch contour**

**Code:**

**一張含有 文字, 字型, 白色 的圖片

自動產生的描述**

先用mean將多channel的音檔合併成單一channel，並存到y\_pitch，然後利用pitch函式產生pitch圖，x軸代表時間。

**Output:**

一張含有 文字, 螢幕擷取畫面, 繪圖, 字型 的圖片

自動產生的描述

x軸為Time，單位為秒(s)

y軸為f0，單位為Hz

**3. Calculate the spectrogram of the extracted audio data in Frequency Domain Code:**

**一張含有 文字, 字型, 螢幕擷取畫面, 代數 的圖片

自動產生的描述**

採用音檔的第一channel，設定overlap\_ratio為0。

然後使用spectrogram函數，並畫出spectrogram。

x軸為Time，單位為秒(s)，y軸為Frequency。

**Output:**

一張含有 文字, 螢幕擷取畫面, 鮮豔, 繪圖 的圖片

自動產生的描述

x軸為Time，單位為秒(s)

y軸為Frequency

圖中的顏色也代表不同的數值。