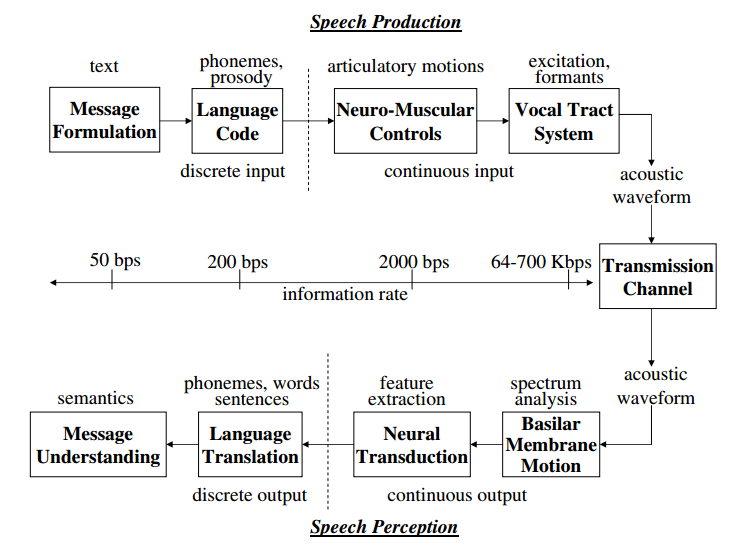
語音處理期末考題2012/12/25

題目共5頁，10題。

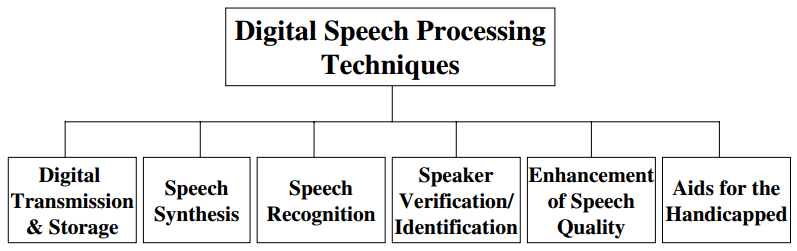
1…

語音(speech)是人類互相溝通的最佳媒介，下圖為「語音鍊」(speech chain)的概念圖，請用以中文(可以穿插英文專有名詞)寫一段小文章，來介紹它，盡量詳細。



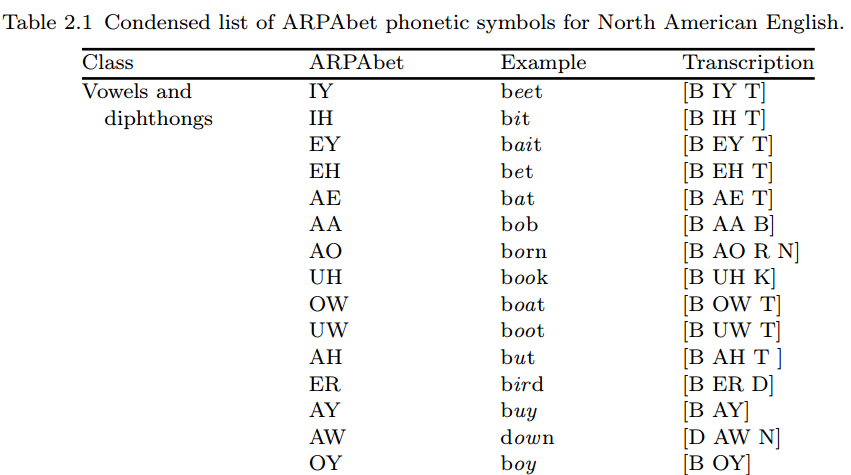
2…

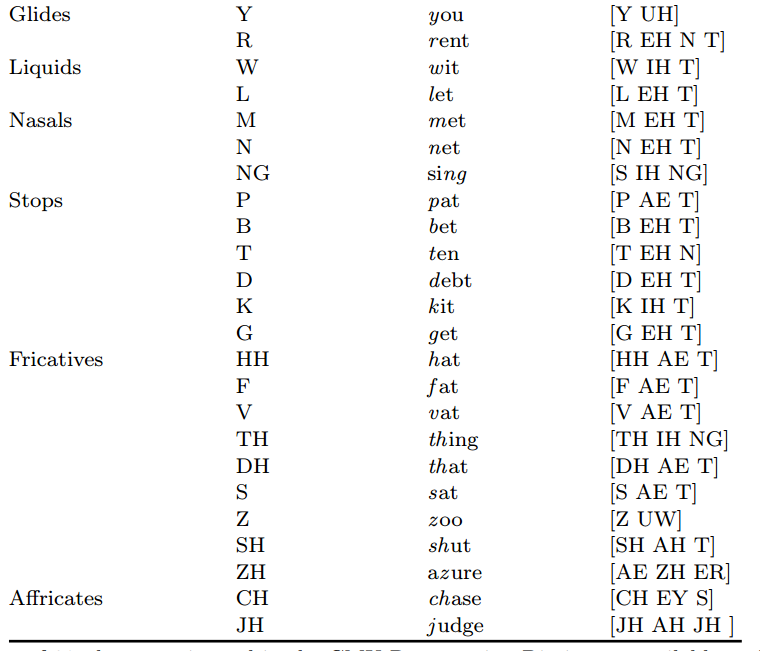
現代語音處理技術有許多分支，如下圖所示，請寫一段小文章解釋每個分支，並提出可能的應用情境。



3…

研究語音處理時，運用一套「音標系統」來紀錄語音是很重要的，對英語而言，下表是一個很流行的音標系統，稱為ARPAbet。





(3.1) 請以其為標準，把下段英文文字的音標寫出來。

“Most of the adventures recorded in this book really occurred; one or two were experiences of my own, the rest those of boys who were schoolmates of mine.”

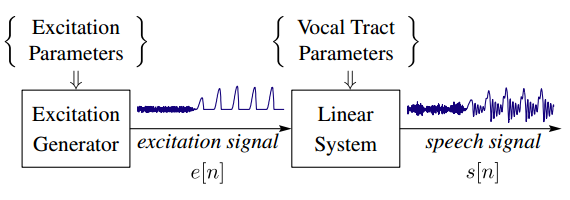
(3.2) 把這段英文文字翻譯成中文。

(3.3) 把譯成的中文用漢語拼音寫出來。

## (本題可以上網運用The CMU Pronouncing Dictionary以及Google Translator。)

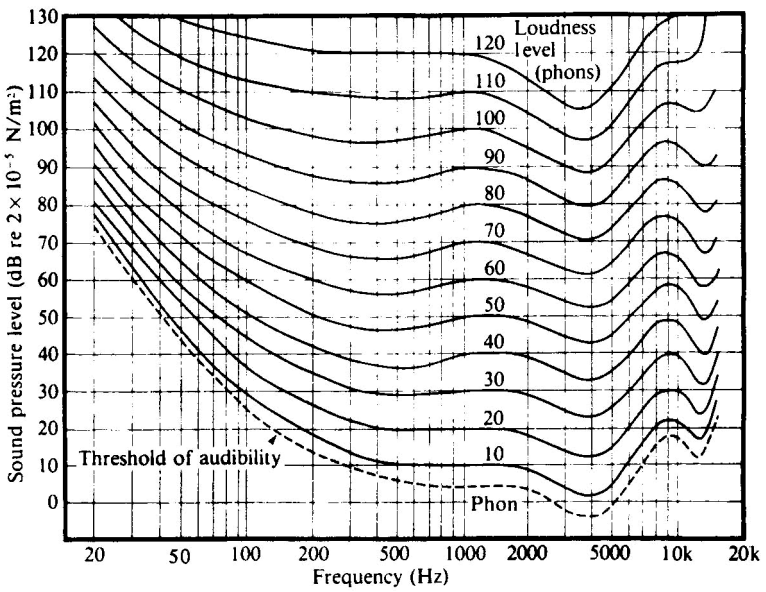
4…

下圖為產生語音訊號的數位系統模型，請你就其中出現的所有專業術語做些解釋(不僅是英文翻譯成中文而已，要做解釋。)特別要說明e[n] 以及 s[n] 的關係，若能以數學方程式描述更好。



5…

下圖為有關人類對聲音的聽覺響度(Loudness)與聲音的聲壓(Sound Pressure)及頻率(Frequency)的關係圖。



請問以下問題：

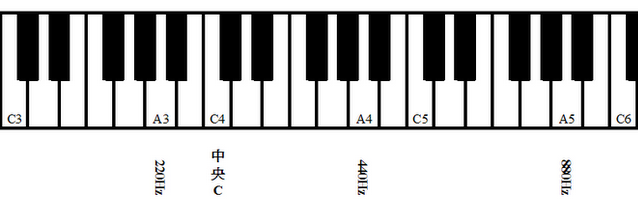
(5.1) 針對頻率為1,000 Hz的聲音，聲壓必須高過哪個數值人類才聽得到(audible)？

(5.2) 聲音A= (10 dB, 1000Hz) ; 聲音B = (20 dB, 100Hz); 聲音C = (10 dB, 4000Hz)，這3個聲音的響度的大小關係為何？

(5.3) 「聲壓比較大的聲音其聽覺響度也比較大。」這句話正確性如何？請解釋你的答案。

6…

人類耳朵對聲音中的「音高」(pitch)的感受雖然與聲音的基本頻率(fundamental frequency)有關，但兩者並不是線性關係。一般而言，鋼琴鍵盤中央的A音訂其基本頻率為440Hz，高八度的A音其頻率為2倍，亦即880Hz；而低八度的A音其頻率為1/2亦即220Hz。任2相鄰琴鍵之音高之基本頻率成等比級數分布。下圖為鋼琴鍵盤中央附近的部分圖示。



(6.1) 請估算出圖中所有C音的基本頻率。

(6.2) 請畫出橫軸(x軸)代表基本頻率(frequency)，縱軸(y軸)代表琴鍵的音高(pitch)，以音名(note name)命名(如C, C#, D, D#, E, F, F#, G, G#, A, A#, B, C,…)的函數關係圖。

(6.3) 如果把音高(pitch)用比音名更精細的「梅爾」(mel)為單位，則音高(pitch)與基本頻率(frequency, 以Hz為單位)有一個數學公式可以表達，將它找出來並寫下來，順道畫出函數關係圖。

(6.4) 根據 (6.3)，頻率為1000Hz, 2000Hz, 4000Hz, 8000Hz各對應到的音高為多少mel?

7…

語音訊號處理很重視「短時」 (short-time)的觀念，有「短時能量」(short-time energy)、「短時過零率」(short-time zero crossing rate)、「短時頻譜」(short-time spectrum)，若給你一段 1 秒鐘長度的語音訊號 x[n]，取樣頻率Fs = 10,000 Hz，短時音框寬度定為0.01秒，短時音框每次移動0.005秒，請你據此寫出「短時能量」訊號E[m]的演算法。(用數學公式表示即可，可加上圖形說明。)

8…

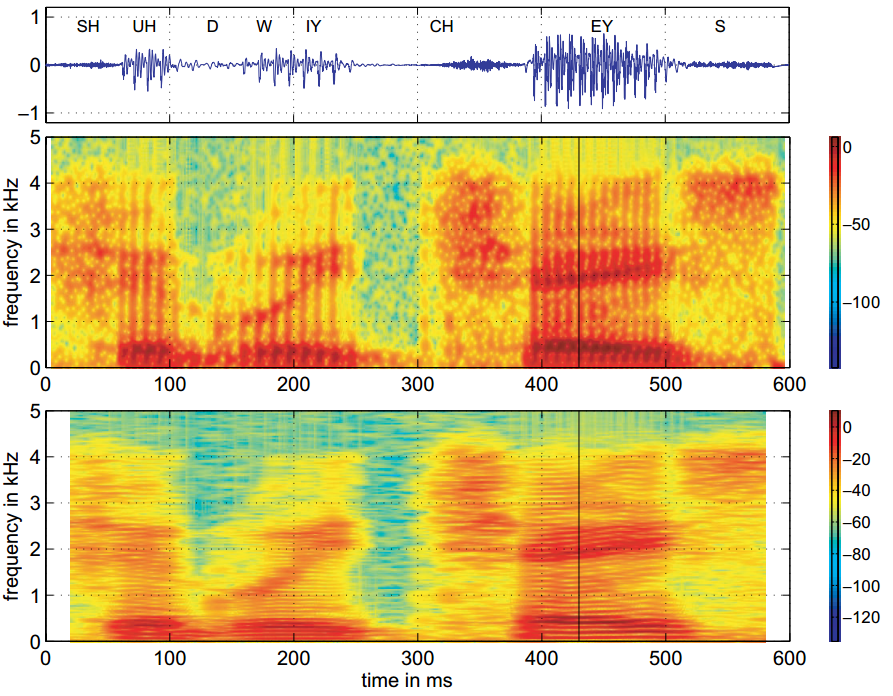
倒頻譜 (Cepstrum) 顧名思義，就是將[頻譜](http://zh.wikipedia.org/wiki/%E9%A0%BB%E8%AD%9C)（spectrum）的英文前四個字母反過來寫。常做為語音訊號的重要特徵被擷取。給你一個音框的語音訊號x[n]，如何求出其倒頻譜？你能否說明一下倒頻譜有什麼用途，能從中找出何種語音訊號的特徵？

9…

自動語音辨識(Automatic Speech Recognition , ASR),是語音處理的一個重要分支，描述一個ASR系統的效能常常要先確定語音的型態 (type, style)，詞彙量的大小(Vocabulary size)，如此一來，提及系統的效能(通常是字詞的錯誤率, word error rate)才有意義。你知道目前針對朗讀式(Read text)的英語語音(English speech)，詞彙量大約60,000左右的語音辨識系統，其字詞錯誤率大約多少？(請指出你的答案的來源出處)

10…

聲譜圖(spectrogram)是語音分析中很常用的工具，如下圖示一段語音的2種聲譜圖。請解釋它們的不同點在哪裡？分別可以抓出所分析的語音的何種特徵？



ITM121 語音暨影像處理 ] [ 開課序號：96382

問卷調查

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 非常同意 | 同意 | 不同意 | 非常不同意 |
| 你是否在本課程中學到以下語音處理以及影像處理的基本觀念？ |  |  |  |  |
|  |  |  |  |  |
| 1. Introduction to Speech Signal |  |  |  |  |
| 2. Hearing and Auditory Perception |  |  |  |  |
| 3. Short-Time Analysis of Speech |  |  |  |  |
| 4. Homomorphic Speech Analysis (Cepstrum) |  |  |  |  |
| 5. Linear Predictive Analysis |  |  |  |  |
| 6. Digital Speech Coding |  |  |  |  |
| 7. Text-to-Speech Synthesis Methods (TTS) |  |  |  |  |
| 8. Automatic Speech Recognition (ASR) |  |  |  |  |
| 9. Introduction to Python Programming Language |  |  |  |  |
| 10. Introduction to Image Processing |  |  |  |  |
|  |  |  |  |  |
| 你是否在本課程中學到很多用Python程式語言來做語音處理以及影像處理的實際經驗？ |  |  |  |  |
|  |  |  |  |  |
| 你是否覺得本課程實用？ |  |  |  |  |
|  |  |  |  |  |
| 你是否覺得本課程有趣？ |  |  |  |  |
|  |  |  |  |  |
| 你是否會推薦其他同學來修習本課程？ |  |  |  |  |
|  |  |  |  |  |
| 請隨意為本課程提供建議。 |  |  |  |  |