

# 第1章 緒論

## 1.1 研究主題

本論文的研究主題是「英文語音評分」，包含「說話驗證」、「英文語音訊號切割」以及「英文語音評分」三個部份，希望融合目前音訊處理和語音辨識的技術，對不同母語的使用者進行公正的語音評分。

由於語音辨識系統不是百分之百的可靠，如果能得到識別結果的可靠性並對識別結果的正確性加以驗證即可以減少辨識的錯誤，因此在語音評分系統中，如果能先擋下內容和標準語音完全不同的評分語音，可以使整個語音評分系統變得更具公信力。本論文運用了一些語音辨識處理的技術來達成說話驗證(Utterance Verification)。

確保了評分語音內容的正確後，對於英文母語及非英文母語的語音，我們使用不同的聲學模型來對其切割出每個音素(Phoneme)的時間區段。使用不同聲學模型可確保評分系統切割出來的音素區段有一定的可信度及正確率。

在英文語音評分部份，我們利用標準語音資料來進行一種較為主觀的評分方式，主要使用圖樣比對(Pattern Matching)的方法，將評分語音和標準語音的資料逐音素地來做比較，以期找出評分語音和標準語音的差異程度，並藉此評分。

在本論文中我們使用以下四個特徵參數：音量強度曲線(Magnitude)、基頻軌跡曲線(Pitch Contour)、發聲急緩變化(Rhythm)以及 HMM 對數機率差異(HMM Log Probability Difference)。音量強度曲線代表聲音音量大小的變化趨勢；基頻軌跡曲線代表聲音音高的起伏；發聲急緩變化代表說話的快慢節拍；HMM 對數機率差異則是將語音訊號經由辨識後得出的對數機率表示語音的內容，藉由比較其差距來計算出兩語音間的差異程度。

## 1.2 英文語音評分系統簡介

本論文提到了英文語音評分的三個主體，分別是：「說話驗證」、「英文語音訊號切割」及「英文語音評分」，根據這三個部份我們實作了一個英文語音評分系統。

首先以說話驗證做為英文語音評分系統的第一道檢視關卡，爾後以聲學模型來對標準語音及評分語音切割音素的時間區段，再將這些資訊送至英文語音評分系統的核心，利用音量強度曲線、基頻軌跡曲線、發聲急緩變化以及 HMM 對數機率差異當作特徵參數，逐音素地比較評分語音和標準語音的差異程度，並配合評分機制對兩者的差異程度進行評分。圖 1-1 為本論文英文語音評分系統的流程圖：

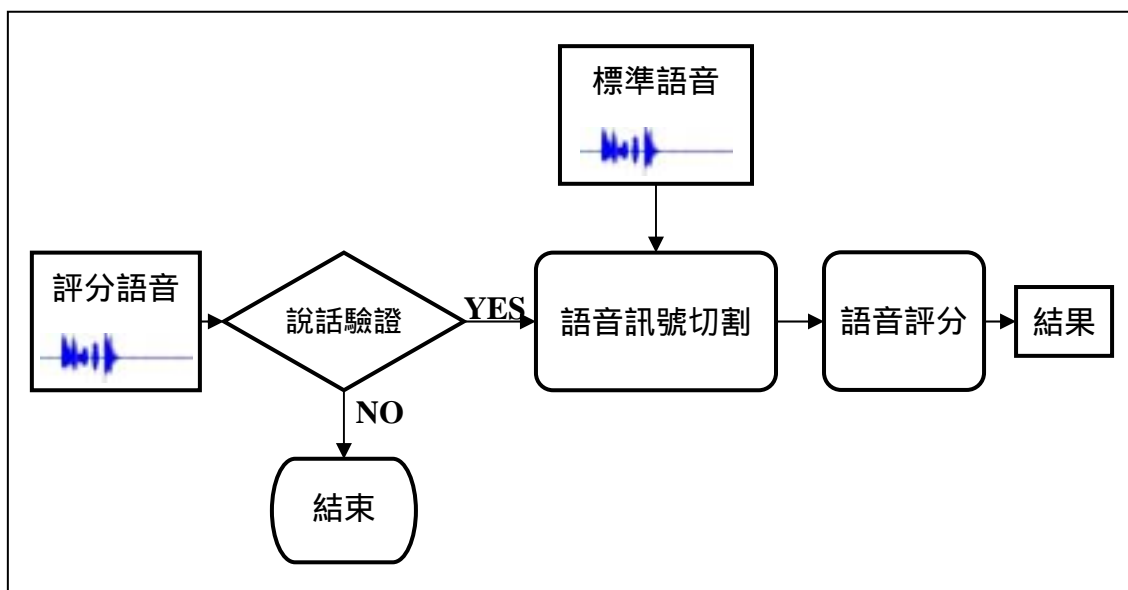


圖 1-1 英文語音評分系統流程圖

### 1.3 本論文研究方向和主要成果

本論文的研究方向為探討英文語音的評分方法，從定義評分系統的各個部份開始，依英文發音的特有性，設計合理的評分流程，並且針對每個部份設計了對應的實驗，期許建立一套合理的英文語音評分系統。

本論文主要的成果為整合目前許多音訊處理及語音辨識相關的技術，運用在英文的語音評分上，用以比較評分語音與標準語音之間的相似程度，並且建立合理的評分機制以及實作出一個完整的英文語音評分系統。

## 1.4 前人相關研究

1996 年 L. Neumeyer 和 H. Franco 等以法文為主，定義 HMM Log-Likelihood、Segment Classification、Segment duration 和 Timing 等四項特徵參數，並經由語音實驗分析觀察後得到了一項結論：語言專家們對於語音評分最注重的一項特徵為 Segment Duration 【18】；而在 1997 年時，C. Cucchiaroni、H. Strik 及 L. Boves 以荷蘭語為主，定義了 Total Duration of Speech no/plus Pause、Mean Segment Duration、Rate of Speech 以及 Global Log-Likelihood，經由類似的實驗後得出 Global Log-Likelihood 對於人類主觀評分占較重的比重 【19】；1999 年 L. Neumeyer、H. Franco、V. Digalakis 和 M. Weintraub 再一次以法文當做實驗的語料庫，採用 HMM Log-Likelihood、Normalized Acoustic、Segment classification、Segment Duration、Timing 當作其實驗的特徵參數，經由實驗後得出了 Normalized Acoustic 在評分系統和語言專家給予的分數中，其相關性高於 Segment Duration 【10】；至於華語的語音評分，2002 年清華大學的李俊毅以實驗得到 MFCC 對於華語語音評分的重要性最大，此處的 MFCC 即為梅爾倒頻譜參數，在其論文中代表語音的內容，而 Magnitude 及 Pitch 的影響較小，另外他也將各個特徵的差異程度轉換成分數，以回饋給使用者參考 【15】。

## 1.5 章節概要

本論文第二章將介紹「說話驗證」，提出我們實作說話驗證時所使用的方法，並且將此單元所用到的技術於各小節中一一介紹，包含聲學模型相似度排名、驗證系統的建立及驗證系統的可靠性等。

第三章介紹「英文語音訊號切割」，這一章包含了許多語音辨識的技巧，諸如隱藏式馬可夫模型(Hidden Markov Model)、維特比演算法(Viterbi Algorithm)、語音辨識前處理與流程以及語音訊號的切割等；另外也會針對語音訊號切割的正確性及聲學模型的辨識率提出說明。

第四章則是介紹整篇論文的核心 - 「英文語音評分」，其中提到了關於特徵參數的擷取、特徵參數正規化、圖樣比對流程、評分機制的建立等；並設計實驗以求出各個特徵參數對於英文語音評分的權重。

第五章為整篇論文的總結，並提出對今後研究工作的展望。