

Abstract (Chinese)

英文語音學習系統是結合音訊處理及語音辨識技術的學問，本論文主要論述的重點在於偵測英文母音的發音錯誤。我們提出可以不需要目標語句的錯誤發音偵測與學習的方法。許多語音上的研究都提到共振峰頻率係數的重要性，因此我們研究將此係數與傳統用於語音辨識的梅爾倒頻譜係數整合於隱藏式馬可夫模型(Hidden Markov Model, HMM)，以期提升語音辨識以及錯誤偵測的正確率。另外，我們也提出發音混淆網路的方法，預測發音錯誤，之後再計算錯誤音素的可信度，提升錯誤發音的正確率。共振峰在發音上的特性則是提供回饋訊息的依據。最後我們設計不同的實驗說明各方法的可行性及效能。



Abstract (English)

This thesis investigates the method for detecting error pronunciation of English vowels in utterances spoken by L2 learners, which requires the techniques from digital signal processing and speech recognition. We propose a text independent approach (which does not require the use of a target utterance) for English vowels error detection and learning. Various studies in formant-based speech synthesis have suggested the importance of formant coefficients; this motivates us to investigate pronunciation assessment using formant information instead of MFCC (Mel-frequency cesptrum coefficients) alone.

In particular, we explore the addition of formant information to improve the recognition rates of HMM. Then we propose the use of PCN (pronunciation confusion network) together with a formant-based confidence measure to raise error detection rates. The phonology knowledge about the formant and the articulator is then employing to generate high-level feedbacks to the user. Experimental results demonstrate that automatic generation of reliable pronunciation instruction (without using a target utterance) becomes highly possible.