



**姓名：簡仁宗**

**學歷：**

國立清華大學電機博士

**現職及經歷：**

國立成功大學資訊系副教授

中央研究院資訊所訪問學人

中華民國計算語言學學會理事

日本東京工業大學資訊系客座研究員

國立成功大學資訊系助理教授

美國 Panasonic Speech Technology Lab 客座研究員



**著作名稱：**

1. Jen-Tzung Chien, "Online hierarchical transformation of hidden Markov models for speech recognition", *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 6, pp. 656-667, November 1999.

2. Jen-Tzung Chien, "Quasi-Bayes linear regression for sequential learning of hidden Markov models", *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 268-278, July 2002.

3. Jen-Tzung Chien, "Linear regression based Bayesian predictive classification for speech recognition", *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 1, pp. 70-79, January 2003.

**中文簡介：**

簡仁宗副教授的三篇得獎著作皆發表在語音處理最具代表性的國際期刊 *IEEE Transactions on Speech and Audio Processing*，為一系列解決在惡劣周遭環境下仍能達到強健語音辨識(Robust Speech Recognition)的

關鍵性研究。一般而言，自動語音辨識技術是建立人與電腦友善人機界面的關鍵技術；然而，語音辨識實際應用受環境存在著不同程度雜訊、通道及語者差異的影響，諸如冷氣風扇噪音、汽車噪音、音樂噪音、室內回音、旁人講話雜音等，不同麥克風、電話話機、手機以及不同的有線 / 無線語音傳輸之通道效應，及語者因性別、腔調、口音、說話速度、情緒、年齡以及身體狀況等之差異，產生不可避免的不匹配現象於訓練語料庫與測試語句之間而導致語音辨識率下降，如果我們能使用少量調整語料將原始語音隱藏式馬可夫模型(Hidden Markov Model, HMM)調整到適合測試之環境，那麼語音辨識效果將能有效的提昇。在大字彙連續語音辨識系統中，我們常採用轉換式調整(Transformation-Based Adaptation)，該方法的優點在於能從少量調整語料估測出轉換函數來調整龐大的語音模型參數，使調整過的參數達到較高的語音辨識率。然而，傳統的調整策略(Adaptation Strategy)多屬於批次調整(Batch Adaptation)，它需要儲存所有的調整語料後才能做參數調整，其所花的時間及計算成本都較高，且測試環境的聲學特性常會隨著時間演進而改變，致使用批次調整無法真實追蹤到非固定式(Non-stationary)的測試環境。因此，為了提昇大字彙連續語音辨識系統的調整效率，第一篇著作即提出一套線上 / 累進調整(Online / Incremental Adaptation)演算法，它可以使用累進觀測到的調整語料不斷的將語音模型調整到最新環境特性上，使記憶體及計算量的需求可以大幅下降。另外，考慮到調整語料不易取得的問題，我們也可結合轉換式調整技術，使其

在少量調整語料情況下，也能將所有聲音單元的模型參數都調整到。本論文主要貢獻在於提出一套完整的累進調整理論基礎，將統計學上的近似貝氏估測(Quasi-Bayes Estimate)法則應用於求取語音模型的最佳轉換參數，使找出來的參數具有封閉解(Closed-Form Solution)以達到快速調整的目的，同時也利用累積出來的參數事後(Posterior)機率與事前(Prior)機率屬於相同機率分布家族的特性，發展出參數事前統計量(Prior Statistics)累進更新的機制，使語音辨識系統可以持續追蹤到最新的環境而不斷提昇不利環境(Adverse Environment)下的語音辨識效果。第二篇則是將這種線上轉換式調整技術延伸至統計學上的線性回歸(Linear Regression)模型上，其應用近似貝氏線性回歸(Quasi-Bayes Linear Regression, QBLR)演算法進行語音模型的線上累進式調整，由於採用高效能的線性回歸調整，效果比起第一篇的方法更佳，且在語者調適(Speaker Adaptation)的實驗中已獲得不錯的驗證結果。本篇論文是相當完整且一般化的，文獻上相當有名的最佳相似線性回歸(Maximum Likelihood Linear Regression, MLLR)及最佳事後機率線性回歸(Maximum a Posteriori Linear Regression, MAPLR)即為本方法的兩種簡化後的演算法。另外，考慮到在線性回歸調整的架構中，常會因為調整語句有雜訊或線性回歸轉換有對應不良(Ill-Posed)的現象，造成估測出來的線性回歸參數有失真且不確定性(Uncertainty)的問題，第三篇則發展出一套強健性決策法則稱為線性回歸為主之貝氏預估分類法(Linear Regression Based Bayesian Predictive Classification, LRBPC)，它可以

有效的將回歸參數估測的不確定性注入於語音辨識的決策法則中，較傳統語音模型調整假設估測出來的回歸參數是真實的(True)所設計出來的決策法則有較佳的語音辨識效果，且其封閉解，可以實現快速且強健性的決策。由於我們有將轉換參數作累進式調整並且將參數的不確定性適當的反應在決策法則上，使得語音辨識器不只具有累進學習能力，其決策法則也較具強健性，能有效增進語音辨識系統的實用性。

三篇著作提出的線上語音參數調整及強健性分類法則技術較傳統方法具紮實的理論基礎，在統計學上的貝氏理論(Bayesian Theory)、估測理論(Estimation Theory)及預估分類法則(Predictive Classification)上有一定程度的貢獻，且在大字彙語音辨識少量調整語料情況下，每個聲音單元的語音模型都能精準的進行累進式調整及強健性預估分類，這對語音辨識系統的實用性提昇有具體幫助。這種語音訊號處理的技術並不僅限於中文語音辨識系統，亦可以廣泛應用在英文、日文、拉丁文、法文、德文等其他國語言的語音辨識系統上。在過去幾年語音相關領域的國際期刊及國際會議上，已有多篇論文參考本方法做延伸性之研究。這一系列方法除了應用在無線電話及PDA聲控系統、汽車環境下的人機介面系統，在電話環境下的自動總機、資訊查詢及遠端交易等應用上亦扮演著關鍵性技術的角色。此外，這套技術還能延伸應用到一般的圖形識別(Pattern Recognition)上，例如光學文字辨識、人臉辨識、文件分類等。

## 評審簡評：

簡仁宗先生的三篇代表作是一系列解決在惡劣周遭環境下仍能達到強健語音辨識的基礎研究。簡先生使用其深厚的統計及訊號處理基礎，提出一系列具有創意的演算法則去解決上述的難題。簡先生在語言辨識領域的研究無疑在國內是數一數二，甚至在國際上亦佔有一席之地。簡先生的大部份期刊論文均刊登在最頂級的期刊，且三篇代表作均為Single Author，這代表簡先生已是一個成熟的研究人才。評審對簡先生能與世界各地不同研發團隊均能成功地合作於以高度肯定。評審認為簡先生在5到10年內能成為語音辨識領域的世界級人物，是屬於明日之星級的學者。