

利用聽覺與發音參數於多語言語音合成

陳嘉平², 黃奕欽¹, 吳宗憲^{1,*}, 李冠德¹

¹國立成功大學資訊工程學系

²國立中山大學資訊工程學系

chunghsienwu@gmail.com

IEEE/ACM Trans. Audio, Speech, and Language Processing, Vol. 22, No. 10, October 2014, pp.1558-1570.

本研究提出在只使用單一語言的語者語料庫基礎上，應用跨語言音框選擇之方

法，以建構一多語語音合成系統。本方法之基本構想是選擇最優音框，以產生人工平行語料，並使用所產生的語料在別國語言與目標語者語言間作調適，以建立一多語語音合成模型。此意味著我們可應用說話者講不同語言的單語語音數據庫來開發一個多語言合成系統。



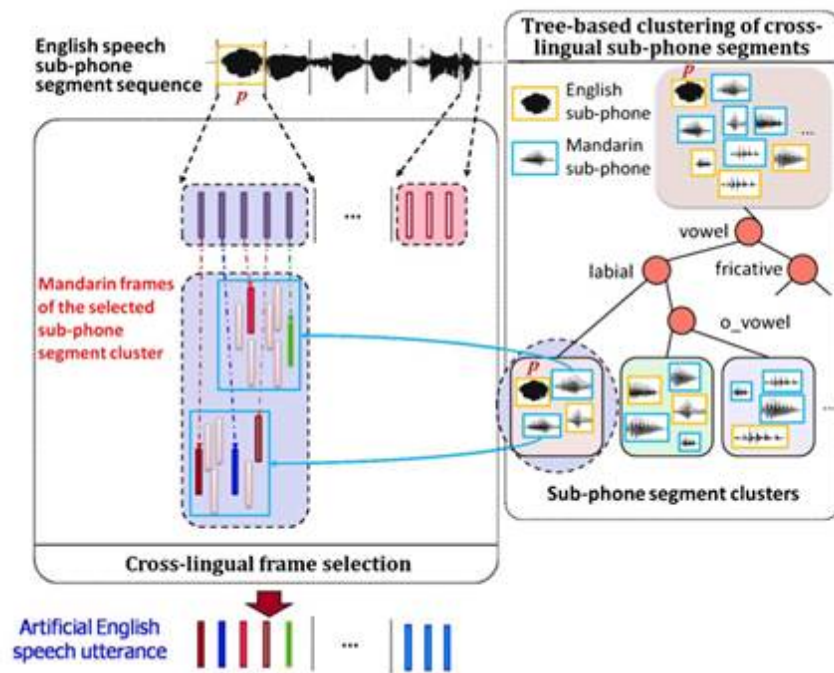
在本方法中，我們試圖找到目標語者語音音框中的主要語言來創建第二語言平行語料。與長度較長的單元如狀態段或音段比較，採用音框是更精細和更精確的方式。所提出的方法之主要構想是將搜索範圍限制於決策樹中的樹狀結構中，以減少搜索成本。此外，其相鄰音框的選擇偏愛發音屬性，所以音框相似度可增強語境之功能。

所提出的方法的一個重要方面是語音音框的表示。頻譜特徵通常忠實傳達發音者身份的信息，我們引入發音屬性為特徵，因為它們是語言和說話者無關的。此外，聽覺特徵被採用以較佳地量化聽覺相似性。

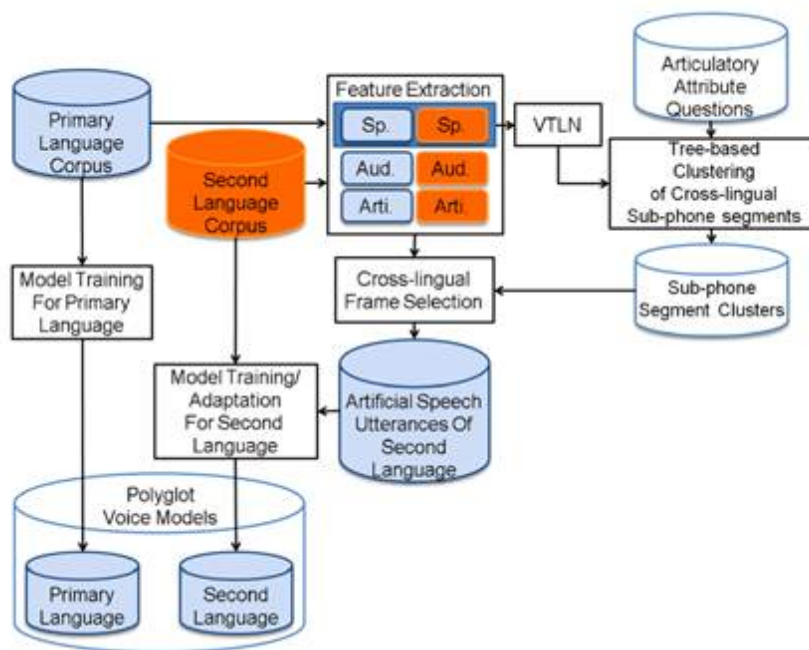
該方法的主要步驟是：

1. 蒐集主要語言（國語）和第二語言（英語）語者的語音數據；
2. 構建用於發音功能偵測的發音屬性檢測器；
3. 將音框段作分類；
4. 在英語語料庫，找到最佳的國語音框序列的每個語音段；
5. 使用英文人工語音以調適成目標語者之英語合成模型；

所提出的想法圖示於圖1。基於目標語者說國語的聲音，人工英語發音則經由音框選擇獲得，並用於調適一個英文合成模型。結合調適之英文合成模型，並訓練國語合成模型，我們實現了一目標語者多語合成系統。圖2表示了整個系統的框架。



圖一 方法之構想



圖二 多語合成系統之架構

為了進行評估國語及英語多語言系統，其中目標語者只說國語。結果顯示，所提出的方法在語者身分和語音質量性能上均有不錯之表現。

參考文獻：

[1] A. Kain and M. Macon, "Spectral voice conversion for text-to-speech synthesis," in Proc. *IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 1998, pp. 285–288.
 [2] C.-H. Wu, C.-C. Hsia, T.-H. Liu, and J.-F. Wang, "Voice conversion using duration-embedded bi-HMMs for expressive speech synthesis," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 4, pp. 1109–1116,

Jul. 2006.

- [3] C.-H. Wu, C.-C. Hsia, C.-H. Lee, and M.-C. Lin, "Hierarchical prosody conversion using regression-based clustering for emotional speech synthesis," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 6, pp. 1394–1405, Aug. 2010.
- [4] D. Erro, A. Moreno, and A. Bonafonte, "INCA algorithm for training voice conversion systems from nonparallel corpora," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 4, pp. 944–953, Jul. 2010.
- [5] A. Mouchtaris, J. Van der Spiegel, and P. Mueller, "Nonparallel training for voice conversion based on a parameter adaptation approach," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 3, pp. 952–963, May 2006.

Copyright 2016 National Cheng Kung University