

國語關鍵詞擷取計劃

Mandarin Keyword Spotting

計畫編號: NSC 88-2213-E009-064

計畫期限: 87/8/1 - 88/7/31

主持人: 劉啟民

交通大學資訊工程系副教授

一、中文摘要 (關鍵詞: 關鍵詞擷取, 中文詞語辨識)

關鍵詞擷取的目標是由一段輸入語音中擷取出裡面的關鍵詞。由於現階段語音辨識技術, 詞語 (word) 辨識要比句子辨識成熟; 因此, 如何由一句子或一段話中擷取出關鍵詞語是語音辨識使用普及的一重要技術。近年來, 隨著辨識技術在商品應用的逐漸成熟, 關鍵字擷取技術也成為語音界一主要探討的主題 [1-7]。本計畫目標是要探討並提出適合於國語的關鍵詞辨識技術, 並整合此技術於本實驗室過去三年來所架設的詞語辨識系統上。

本計畫主要探討主題有四: (一) 具應用調整性的非關鍵詞和關鍵詞模擬技術, (二) 關鍵詞存在與否判斷技術, (三) 最可能關鍵詞尋找技術, 和 (四) 展示系統設計。

英文摘要 (Keywords: Speaker Recognition, Hidden Markov Model, Speech Processing)

The objective of keyword spotters is to spot keywords embedded in extraneous speech. Since keyword spotting technique can provide user-friendly interfaces for speech recognition systems, the research on spotting techniques is receiving more attentions in recent years. This project considers the techniques for Mandarin keyword spotting, and integrates the spotting technique with the Mandarin speech recognition systems we established in the last three years. This projects consists of mainly four topics:

1. the application-flexible modeling technique for the keywords and the nonkeywords,
2. the existence detection technique of the keywords,
3. the keyword searching techniques,
4. the integration with the word recognition system.

二、計畫之緣由與目的

就一實用語音辨識系統來說, 如何由一句子或一段話中擷取出關鍵詞語是語音辨識使用普及的一重要技術。尤其本計畫的展示系統

以整合完成高效能詞語辨識技術和雜訊容忍性, 若能進一步整合關鍵詞擷取技術將對國語語音辨識技術得推展有相當大貢獻。因此, 本計畫對累積此方面技術有其未來市場潛力及前瞻性。

文獻上 (請參考文獻 [1-8]), 傳統關鍵詞擷取技術是根據一應用所用關鍵詞與非關鍵詞, 來模擬非關鍵詞和關鍵詞模擬單位, 因此, 本計畫希望能夠研究出適當的非關鍵詞和關鍵詞模擬方式, 並配合高效率的關鍵詞尋找技術, 來達到擷取關鍵詞的目的。而在關鍵詞存在與否判斷技術方面, 我們希望能夠根據發音方式, 非關鍵詞與關鍵詞長度比例, 關鍵詞發音內容等提出更精確的關鍵詞存在判斷技術。

三、研究方法與成果

我們根據文獻所提的關鍵詞擷取架構 [1] 建立系統, 請參考 fig 1。用填充模型 (filler model) 來模擬關鍵詞之外的所有詞語, 並整合樹狀搜尋演算法 (tree-trellis search algorithm) 找出關鍵詞。

根據中文的特性, 我們設計了三種不同的填充架構: 「音節」、「聲母/韻母」、「聲母-韻母」。「音節」式的填充架構是指 fig 1 上的每一個填充模型 (filler model) 均是一個音節。「聲母/韻母」式的填充架構是指每一個填充模型是聲母或是韻母。而「聲母-韻母」式的填充架構則是指每一個填充模型均是先聲母後韻母的組合。所採用的聲學模型可分為 CI (context-independent) 以及 RCD (right context-dependent) 兩種。

由實驗結果發現, 在「聲母/韻母」或是「聲母-韻母」的填充架構之下, RCD 這種較精確的聲學模型不見得會比 CI 種較粗糙的聲學模型好。這是因為在這兩種填充架構之下, 使用 RCD 模型會造成過多填充模型的可能組合, 而其中有許多實際上在中文當中並不是合法的組合。因此, 就造成填充模型和關鍵詞之間分界點的誤判, 而進一步影響了辨識率。

我們提出兩種方法解決此一問題, 第一種是針對填充模型的分數隨著時間而加以抑制, 藉以削弱過多填充模型的可能組合造成的影響。第二種方式是針對填充模型以及關鍵詞分別以不同精確度的聲學模型來模擬: 以 CI 來模擬填充模型, 而以 RCD 來模擬關鍵詞的部分。這是因為在「聲母/韻母」以及「聲母-

韻母」的填充架構之下，若是採用 RCD 模型作為填充的聲學模型，會產生過多不存在的填充組合。

我們針對 500、5000、25000 的人名作為關鍵詞的詞庫大小，綜合使用這兩種方法，可以分別得到 5.8%、9.3%、8.7% 的辨識率改善。而辨識時間即使在 25000 詞庫之下，也可以在 2 秒之內完成。

四、結論與未來展望

本計畫成果已達成計畫書所提預期成果：

1. 具應用調整性的非關鍵詞和關鍵詞模擬技術。
2. 關鍵詞存在與否判斷技術。
3. 最可能關鍵詞尋找技術。
4. 展示系統設計。

在技術方面，本計畫提出的兩種方法，在辨識率以及辨識時間上均有相當不錯的表現，相當具有實用性。

而在未來展望方面，將來希望能夠整合雜訊環境，在必須抗雜訊的系統要求下，更進一步探討關鍵詞擷取的相關技術。

本計畫已在最著名的語音期刊 IEEE Tran. On Speech and Audio Processing[9] 發表。

五、參考文獻

- [1] E.F. Huang, H.C. Wang, and F.K. Soong, "A fast algorithm for large vocabulary keyword spotting application," IEEE Transactions on Speech and Audio Processing, pp. 449-452, vol. 2, no. 3, July, 1994.
- [2] J. C. Junqua and J. P. Haton, "Robustness in automatic speech recognition," Chapter 10, "Word-spotting and rejection," Kluwer Academic Publishers, 1996.
- [3] C. J. Lee and E. F. Huang, "A large vocabulary keyword spotter for automated Mandarin phone directory assistant services," *ISMIP*, Dec. 1996, pp. 577-582.
- [4] R. Rose and D. Paul, "A hidden Markov model based keyword recognition system," *ICASSP*, 1990, pp. 129-132.

- [5] A. Manos and V. Zue, "A segment-based wordspotter using phonetic filler models," *ICASSP*, 1997, pp. 899-902.
- [6] H. Boulard, B. D'joore, and J. Boite, "Optimizing recognition and rejection performance in word spotting systems," *ICASSP*, 1994, vol. I, pp. 373-375.
- [7] R. Rose, "Automatic speech and speaker recognition," Chap13, "Word spotting-extracting partial information from continuous utterances," Kluwer Academic Publishers, 1996.
- [8] E.F. Huang, F.K. Soong, and H.C. Wang, "The use of tree-trellis search for large-vocabulary Mandarin polysyllabic word speech recognition," *Computer Speech and Language*, 8, pp. 39-50, 1994.
- [9] Chi-Min Liu, Chin-Chih Chiu, and Hung-Yuan Chang, "Design of Vocabulary-Independent Mandarin Keyword Spotters," to appear in *IEEE Trans. on Speech and Audio Processing*, 1999

五、附圖

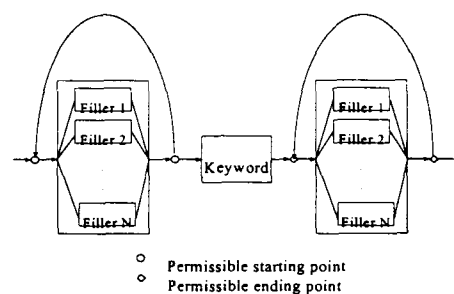


fig 1. 關鍵詞擷取架構