

應用辭彙校準及合成片斷產生之最佳化方法於中文轉台灣手語之研究

邱毓賢、吳宗憲*、蘇弘毓、鄭功偉

成功大學 資訊工程系

chwu@csie.ncku.edu.tw

This article is based on the publication: Yu-Hsien Chiu, Chung-Hsien Wu, Hung-Yu Su, and Chih-Jen Cheng, "Joint Optimization of Word Alignment and Epenthesis Generation for Chinese to Taiwanese Sign Synthesis," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 29, No. 1, January 2007, pp.28~39.

手

語為一種視覺化的手勢語言，

也是聽語障人士賴以溝通的主要語言，如同我們所使用的中文、英文同屬於自然語言的範疇之一。聽語障人士生活於一般聽人社會中，往往會遭遇到許多溝通或表達的困難，進而造成在教育、交際以及就業上極大的阻礙。為了改善聽語障人士與常人溝通的能力，替代與增益性溝通輔助科技的

提出，使得聽語障溝通輔具的發展有了長足的進步。近年來，有關機器翻譯技術的發展，對於消弭口語與手語之間的屏障有了合適的解決方法。在本研究首先提出了一個基於統計式機器翻譯架構的中文到手語的翻譯系統，再利用影像串接之技術，將預先拍好的手語詞彙影像片段以動態方式合成轉譯後之手語語句，提供聽語障者一個影像基礎的手語呈現方式。圖一為整個系統之操作介面。

一個平行的雙語對照語料庫為發展一個機器翻譯系統最重要的部份。在基於語料庫的機器翻譯系統中，對照平行語料為訓練翻譯模型的基礎。由於手語語料在記錄上的複雜性，造成了語料收集的困難。在本研究中，對照語料乃收集自台南啟聰國小之國語教材，並由國立中正大學語言所之手語專家做標記及驗證。另外為了串接手語影像之輸出介面，我們設計且拍攝了一個以手語詞彙間之手的位置與移動方向差異為考量的影像資料庫。

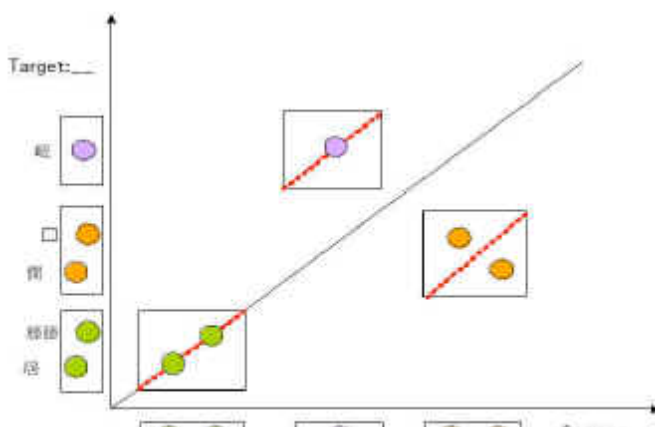
為了減少一般統計式機器翻譯的計算複雜度，我們提出了一個兩階段的翻譯架構。在輸入欲翻譯之中文句後，經由句法及詞彙語意的分析，並將所有詞彙分為兩大類：主單元及次單元。在第一次的順序校準中，決定了主單元與連續次單元組成的主結構的順序。第二次的順序校準則是針對主結構中的次單元來執行。表一列出了經由句法及知網分析後，各類詞彙的分類結果。圖二表示了中文轉譯手語時之搜尋空間。

表一 列出了經由句法分析後，主單元及次單元的分類結果

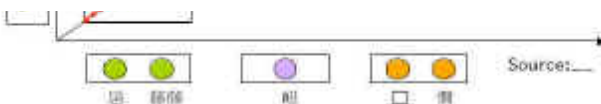
Primary Units (PUs)		Secondary Units (SUs)	
Syntactic Cluster	Examples	Syntactic Cluster	Examples
V	喜歡(like)、買(buy)、有(have)	Agent	爸爸(father)、我(I)
Negative, Positive	無法/不行(cannot)、會/可以(can)	V (adverb)	便宜(cheap)、漂亮(beautiful)
Conjunctions	跟/和(and)、或者(or)、但是/is/可是(but)	Dfr	很/非常(very)、格外(extra)
Nouns (Time, proper Np, Ned)	明天(tomorrow)、昨天(yesterday)、台北(Taipei)、餐館(restaurant)、什麼(what)、哪/哪裡/哪兒(when)	Nouns (common, New)	一(one)、百(hundred)、花(flower)、太陽(sun)、車子(car)



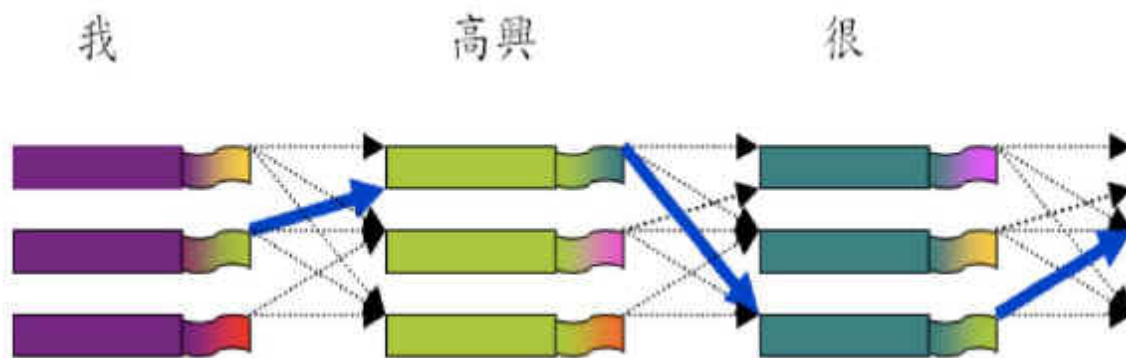
圖一為中文轉台灣手語影像系統之操作介面



經由翻譯之後所產生的手語詞彙序列，會以影像串接方式加以呈現。圖三為一示意圖。在手語影像庫中，各手語詞彙均保留多個取自不同的片段並包涵該詞彙在施打時與其它詞彙的串接可能。因此在串接影像時，影帶與影帶之間的串接點的決定參考了兩個影像框架中，手的位置差異及移動方向的差異，並以動態演算法由圖三的網格中決定差異最小的路徑，來串接出平滑順暢的手語影像輸出。



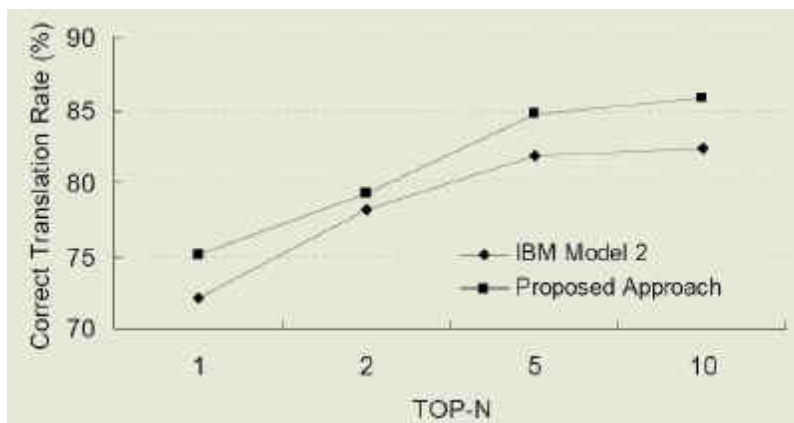
圖二表示了中文轉譯手語時之搜尋空間



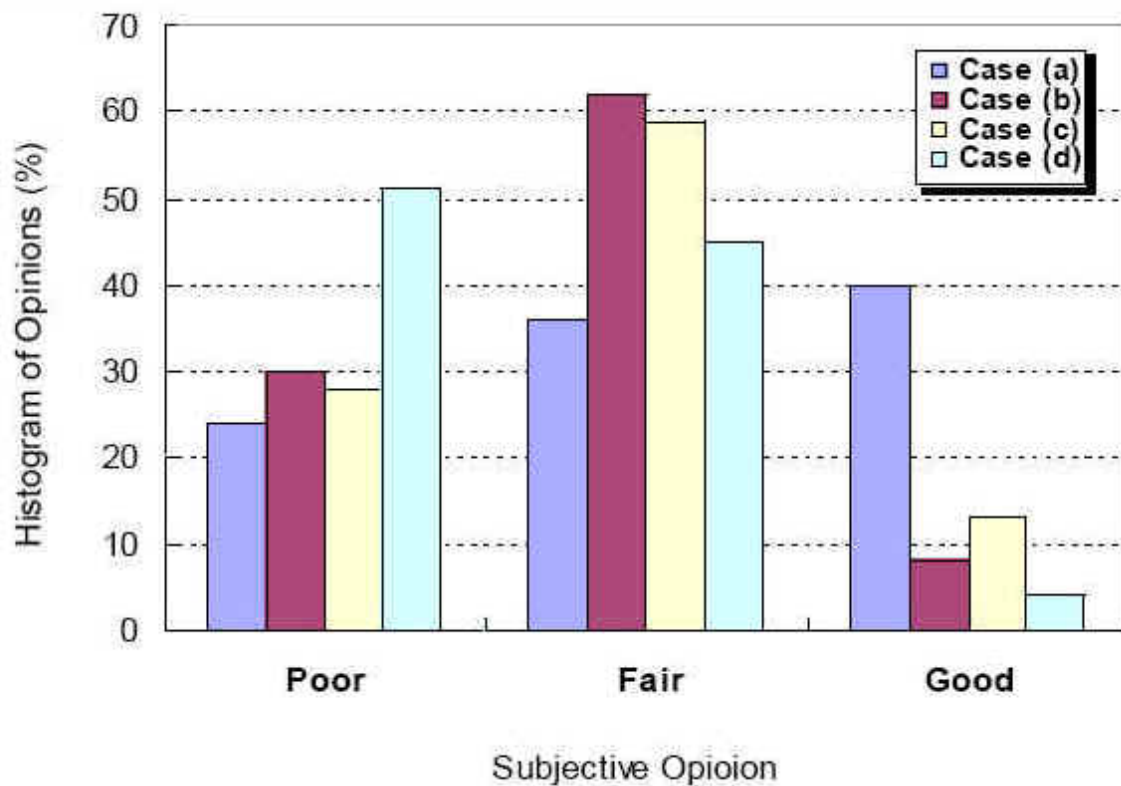
圖三為以動態演算法決定差異最小的路徑來串接出平滑順暢的手語影像輸出

在系統的評量方面，首先是統計式機器翻譯效能的評估。在本論文中，提出的方法與國際商務機器公司所提出的翻譯模型2相比較。圖4評估了在前N個可能的翻譯中正確的比例。從實驗結果可以明顯的看到我們所提出的方法有三到四個百分點的領先。另外，在整體系統的評估上，由於沒有適當的客觀評估方式，所以採用平均意見分數。我們選擇十句中文句經翻譯並合成之手語影像，由手語使用者來觀看，並對串接與翻譯的品質做三等級的評估：完美、一般及不理想。

圖五表示了觀看者對四種串接考量的評分。A為考量手的位置及移動方向，B和C分別為考量移動方向和手的位置，D則是直接串接的結果。



圖四 評估了在前個可能的翻譯中正確的比例



圖五表示了觀看者對四種串接考量的評分

由以上的評估結果，本研究所提出的翻譯模型及影像串接架構已達一可用之成果。本系統預期可應用於電腦輔助介面或是手語電子書的建構。另外在公眾服務方面，本系統可提供各式電腦化互動式之自動交易機器一手語介面、以影像方式引導聽語障人士使用各種設備以及應用於電視手語字幕等方面。