

語音應用於行動路徑導航之研究

研究生：王欣逸

指導教授：王傳芳

黃台生

國立交通大學交通運輸研究所

摘 要

國內外發展汽車導航從最初期的筆記型電腦配置電子地圖光碟與導航軟體，發展至今已有成熟的車用導航產品，並且逐漸成為車輛標準配備。隨著地理資訊系統技術的演進，目前車輛導航系統可顯示詳盡的圖形資料於其配備之顯示器上，但圖形的複雜度直接關係到駕駛者將注意力分散到顯示器所引發的行車安全問題。另外，駕駛者在行車途中也可能引發操作導航軟體的需求，目前車用導航系統以按鍵操作軟體的方式，對於行車安全仍然造成影響。

本研究思考以語音作為汽車導航系統之人機介面，並實驗性地將語音辨識與語音提示功能加入其中，以達到提升行車安全的目的。在本研究中所使用的訊號處理技術，均是語音辨識領域中所慣用的，或者是計算量較少的，字詞範圍為特定詞辨識。從讀入錄音檔案開始，本研究將理論公式化為程式實作，達到使用者下達地圖操作指令後，軟體出現應對動作的目標，以及在轉彎路口前，以語音提示駕駛人詳細的轉彎資訊，包括路口距離、轉向角度、轉向後之路名。另外，亦具有設施提示功能，以語音告知駕駛者道路周邊之停車場、飯店、加油站等設施資訊。

將語音辨識應用於個人數位助理器上容易發生執行效能不佳的情況，尤其是語音辨識需要對語音訊號作一連串的處理，並比較各語音模型間，語音特徵值的失真量，使得辨識速度與硬體有直接關連。而本研究之語音提示使用文字轉語音技術，雖然能清楚達到轉向提示與設施提示兩大目的，但在遭遇破音字的情況下，有可能會產生不正確的發音，應是此系統未來應著重的問題。

關鍵詞：汽車導航系統、地理資訊系統、語音辨識、個人數位助理器

The Study of Voice Application for Route

Navigation on Mobile Devices

Student : Hsin-I Wang

Advisor : Chuan-Fang Wang

Tai-Sheng Huang

Institute of Traffic and Transportation

National Chiao Tung University

ABSTRACT

Vehicle navigation system has been developed from in-car notebook with digital map on CD-ROM to the advanced devices nowadays. At the same time, the progress of geographic information system enables the graphics on the screen to show details. However, the complexity of the digital map also draws driver's attention as result of safety concerned. On the other hand, the manual operation of the navigation devices drives the same issue.

The study takes the voice as a human-machine interface to improve the driving safety with speech recognition and voice response functions programmed on the personal digital assistant. The algorithms used in the study are referring to the method and theory in the key word speech recognition field. The navigation software built by this study has realized the speech recognition on the PDA to operate map controls, and the voice responses include the corner distance, angle, next route, the information of parking lot, hotel, and gas station.

The major obstacle applying speech recognition is the efficiency of the program, a series process handling a voice file which slows down the speed of recognition. Although the users can get the voice responses clearly, the multiple pronunciations of a Chinese word may decrease the accuracy of it, and should be an issue in the future study.

Key Words : Vehicle Navigation 、 Geographic Information System 、 Speech Recognition 、 Personal Digital Assistant

誌 謝

本篇論文能順利完成，承蒙恩師王傳芳老師以及黃台生老師不厭其煩地指導。黃老師總是能掌握論文的整體架構與重點，當我鑽研於旁枝末節的小問題之際，黃老師總能適時給予提醒與指導，使我在研究過程中不致偏離主題。而指導學業的同時，可以感受到黃老師謙卑為懷的處世態度，這是我所敬佩的，也使得我對未來感到困惑時，能與老師如同朋友般談笑。智者知日：明白事理；慧者掃心：掃除私心，這是我由兩位老師身上學習到的態度。

論文口試期間，承蒙羅濟群教授、曾平毅教授熱心指正與建議，並惠賜寶貴之見解，深表由衷感謝。論文研討期間，黃承傳教授、藍武王教授、徐淵靜教授，以及王明智學姊、黃昱凱學長、黃益三學長惠予寶貴意見與看法，特此致謝。

語音辨識較偏向電機資訊學院的領域，需要理論的認識以及程式設計的基礎，對於以交大運管系為背景的我，算得上是跨領域的嘗試。誠如牛頓所說：我站在巨人的肩膀上。感謝參考文獻中所有的作者，使得我對這門學問有所認識，以及台大電信所語音實驗室的高中同學孫良哲百忙之中抽空向我講解語音辨識的理論與方法。

研究所兩年中，佳彥是好友也是讀書良伴，明穎、文聖、雲萍、老吳、宏儒、智淵、佳宇、漢銘、哲強、彥琪、姿琦等好友常給予我幫助與支持，所辦洪璿瑛小姐、柳美智小姐對我的照顧，還有所有同窗好友相互的砥礪與鼓勵，這些都令我永生難忘。女友素芬的體貼，也令我感動。

僅以本論文獻給最偉大的父母，二十多年以來的辛勤栽培，感激之情，無以言表。

目 錄

	頁次
中文摘要·····	i
英文摘要·····	ii
誌謝·····	iii
目錄·····	iv
圖目錄·····	vi
表目錄·····	viii
第一章 緒論·····	1
1.1 研究背景與動機·····	1
1.2 研究目的與課題·····	1
1.3 研究架構·····	2
1.4 研究範圍·····	3
1.5 研究內容與方法·····	4
1.6 研究流程·····	5
第二章 文獻回顧·····	6
2.1 簡圖顯示之路徑導航方面文獻·····	6
2.2 語音辨識方面文獻·····	10
2.3 汽車導航系統之原理與發展·····	12
第三章 路徑導引之語音功能需求·····	25
3.1 路徑導引之功能與架構·····	25
3.1.1 語音辨識·····	26
3.1.2 轉向提示·····	27
3.1.3 設施資訊·····	30
3.2 駕駛人下達之關鍵詞指令·····	31
3.3 語音提示內容·····	35
3.3.1 轉向提示·····	35
3.3.2 設施資訊·····	37
第四章 建立語音辨識與提示功能·····	38
4.1 開發工具·····	38
4.2 辨識原理·····	39
4.2.1 類比／數位轉換·····	39
4.2.2 語音訊號前置處理·····	41
4.2.3 計算音框特徵參數·····	45
4.2.4 動態時間校準法·····	52
4.3 文字轉語音技術·····	58
第五章 導航系統展現·····	60
5.1 地圖操作·····	60

5.1.1	讀入路徑與設施座標檔案	60
5.1.2	放大與縮小功能	61
5.1.3	地圖平移功能	62
5.2	錄音操作	63
5.3	轉向與設施提示功能	64
5.3.1	轉向提示	64
5.3.2	設施提示	65
5.4	語音辨識結果顯示	66
第六章	結論與建議	67
6.1	結論	67
6.2	建議	68
	參考文獻	70

圖 目 錄

	頁次
圖 1-1 研究架構圖	2
圖 1-2 本研究之導航系統架構圖	3
圖 1-3 研究流程圖	5
圖 2-1 路徑訊息認知圖	7
圖 2-2 J2ME 程式編輯流程	8
圖 2-3 Palm 導航資訊簡圖	8
圖 2-4：路徑角度關係	9
圖 2-5 車用導航系統架構圖	13
圖 2-6 人類資訊處理裝置	21
圖 2-7 資訊呈現數量與資訊處理數量關係圖	22
圖 3-1 路徑導引系統功能架構圖	25
圖 3-2 語音辨識功能的操作流程	27
圖 3-3 轉向提示的輸入輸出資訊	28
圖 3-4 分別代表四個象限的節線 a, b, c, d	29
圖 3-5 計算節線角度之條件判斷式	29
圖 3-6 節線夾角之 12 時方位	30
圖 3-7 語音提示之設施資訊流程	31
圖 3-8 林詠傑建立之簡圖導航軟體系統架構圖	33
圖 3-9 駕駛者下達之關鍵詞指令樹狀圖	34
圖 3-10：轉向提示可能產生問題的情況	36
圖 3-11 轉向提示句型結構	36
圖 4-1 eMbedded Visual Basic 開發環境	38
圖 4-2 Pocket PC 2002 模擬器執行畫面	39
圖 4-3 實機錄製「放大」一詞之取樣值	41
圖 4-4 實機錄製「地圖」一詞之取樣值	41
圖 4-5 漢明窗函數時域圖	44
圖 4-6 漢明窗頻域圖	45
圖 4-7 特徵參數計算步驟	45
圖 4-8 求取自相關係數程式流程圖	47
圖 4-9 求取線性預估係數之程式流程圖	48
圖 4-10 計算倒頻譜參數之程式流程圖	49
圖 4-11 「放大」一詞之倒頻譜參數	50
圖 4-12 計算轉移倒頻譜參數之程式流程圖	51
圖 4-13 理想狀況之動態校準法路徑	52
圖 4-14 整體搜尋路徑限制之平行四邊形	53

圖 4-15 決定 $D_{\min}(m,n)$ 時搜尋三節點示意圖·····	54
圖 4-16 在解空間中可能出現的搜尋路徑示意圖·····	55
圖 4-17 動態校準法程式流程圖·····	56
圖 4-18 「放大」一詞之測試與參考樣本差異之立體曲面圖·····	57
圖 4-19 「放大」、「縮小」二詞參考樣本差異之立體曲面圖·····	58
圖 4-20 文字轉語音程式流程·····	59
圖 5-1 讀入導航路徑簡圖·····	60
圖 5-2 放大 32 倍之後顯示文字·····	61
圖 5-3 向上平移後的結果·····	62
圖 5-4 錄音控制列·····	63
圖 5-5 轉向提示·····	64
圖 5-6 設施提示·····	65
圖 5-7 語音辨識後的結果·····	66

表 目 錄

	頁次
表 2-1 敘述性文字語意詞彙轉換值表.....	10
表 2-2 聲音輸出輸入介面 API 功能敘述.....	11
表 3-1 中文聲母與韻母內容.....	32
表 3-2 MAT 語料庫統計表.....	32
表 3-3 設施資訊內容.....	37
表 4-1 取樣資料範圍表.....	40
表 4-2 每八十筆取樣資料存入陣列.....	42
表 4-3 Sample 陣列與 Frame 陣列之關係.....	44