

具辭書式情境感知之語音導引機器人

王宗一 蔡昆樺 邱迪凱 張書龍 謝東成 蘇建元

國立成功大學工程科學系

wti535@mail.ncku.edu.tw

摘要

一般大型公共場所或會場的服務人員，常見的服務是導引訪客到想去之目的地，而除了服務人員的協助外，也有單位提供電腦輔助查詢的方式（如政府或醫療單位）以降低人力成本。雖然電腦可以提供輔助導引的功能，但所提供的資訊均為固定，不會隨情境變換而更新，例如：此單位未到開放時間等。本研究製作一具有情境感知之智慧型語音導引機器人，使用非特定語者的語音辨識與訪客對話互動，並告知來訪者目的地情境之即時資訊，讓來訪者決定是否要導引至目的地，減輕來訪者的搜尋成本；除此之外，亦系統能讓使用者易於置換地圖和情境，以滿足使用者不同之需求。實驗結果顯示，該具辭書式情境感知之語音導引機器人在非特定語者之辨識率達到 73.37% 且可依據情境的變換提供即時資訊，有效的引導使用者至目的地。

- 導引訪客至想去的地方
- 閃避障礙物
- 路徑規劃
- 導引地圖使用者操作介面

但這些功能不足以提供完善的服務，因為如果只引導訪客至目的地，但訪客並不知道目的地的目前的情境，例如：道路不通、此單位未到開放時間等情境，將導致訪客可能白跑一趟或在引導的過程中碰到無法行走的情況。因此，如果在導引機器人的架構上加入情境感知，可使得機器人的服務更完整且讓人們具有便利。

因此，本研究實作一具辭書式情境感知之導引機器人，本研究使用非特定語者的語音與訪客對話互動，並導引訪客至目的地或顯示環境地圖提供訪客參考，更加入主動告知來訪者目的地情境之即時資訊，讓來訪者決定是否要導引至目的地，減輕訪客的搜尋成本。

一、前言

現今科技快速的發展和自動化的應用越來越普遍，許多國家都投入智慧型機器人之相關研究，目的在希望提高工作或服務之效能，降低人力成本，讓生活能夠更便利，於是開始發展各式各樣用途的機器人。例如：日本愛知博覽會[1]，這裡展示各式各樣的機器人，這股熱潮一直延燒到台灣及世界各地，智慧型機器人的研究已不再遙遠，同時也意味著智慧型機器人世代即將到來。

其中，引起大家關注的機器人，又以智慧型服務型機器人最為顯著，例如愛知博覽會中許多招待型導引機器人，引起人們廣泛興趣及注意。智慧型導引機器人最主要的功能：

- 與人進行對話、溝通及協助
- 提供相關資訊查詢

二、系統架構

本研究提出具情境感知之導引機器人，不但能與人互動並導引，也可定義相關的情境訊息來提供資訊給來訪者，並提供使用單位輕易更改地圖及情境，來達到降低機器人與場所之耦合性和更有善的服務。如圖 1 所示，為本系統的架構圖。

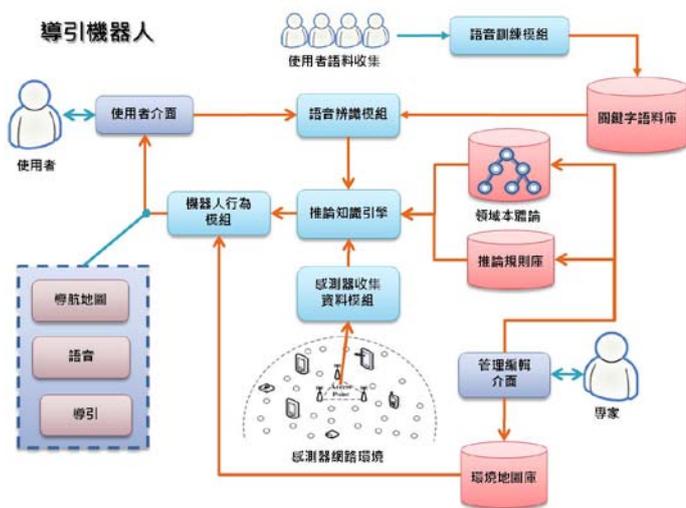


圖 1 系統架構

(一) 語音訓練/辨識模組

若要訓練導引機器人的非特定語者各個字 (Word) 的語音模型，就可用訓練 (參數重估) 模型的流程來實現，如圖 2 所示。首先使用者透過語音人機互動介面，將關鍵詞語音輸入至麥克風。語音訊號經由音效處理晶片 (音效卡)，實作取樣、量化和編碼，將訊號由類別轉換成數位。

本研究使用能量端點偵測的方法，先找出有效語音段，可增加運算上的速度。接著，將語音訊號切割成多個獨立音框 (Frame)，各別進行前處理 (Preprocessing)。

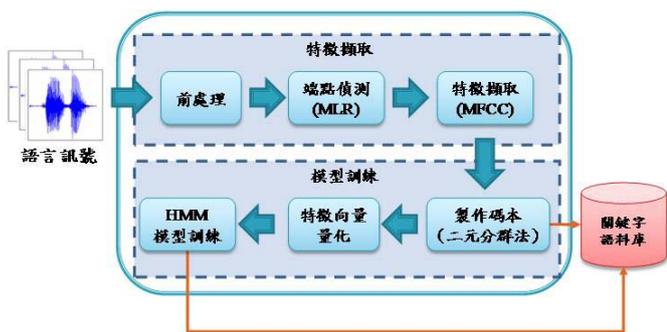


圖 2 語音訓練模型流程

(二) 情境感知推論

推論知識引擎如圖 3 所示，如果語音辨識模組所辨識出來的關鍵字為目的地，推論知識引擎會至感測器收集資料模組讀取目的地所有感測器的資訊，然後再依據推論規則庫所提供之規則，將領域辭書所定義的概念和利用收集資料模組所收集的感測資訊進行知識推論，並可得知目的地有什麼事件發生。

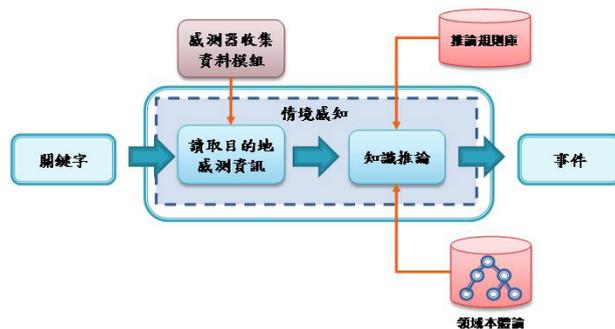


圖 3 推論知識引擎

本系統利用領域辭書來建立目的地情境知識的建構，如圖 4 所示，在領域知識辭書 (Domain Ontology) 架構[3]中，包括 Domain Layer、Category Layer 及 Class Layer 等三層。其中，Domain Layer 代表一個 Ontology 的 Domain Name，它是由多個 Categories 所組合而成的；每個 Categories 是由多個於 Class Layer 內的 Concept 所組成的；而每個 Concept 又包括多個屬性名稱和屬性值，用來敘述這個 Concept。

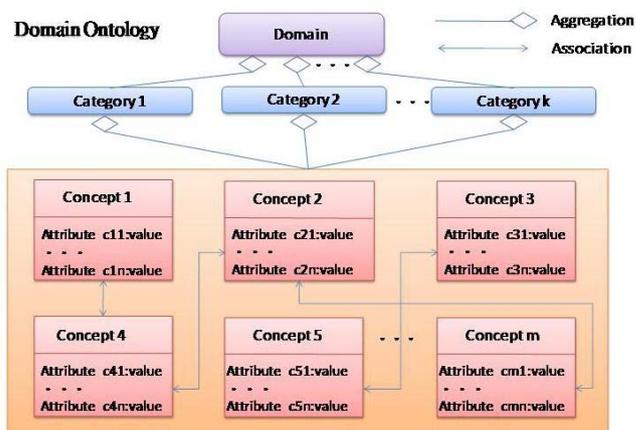
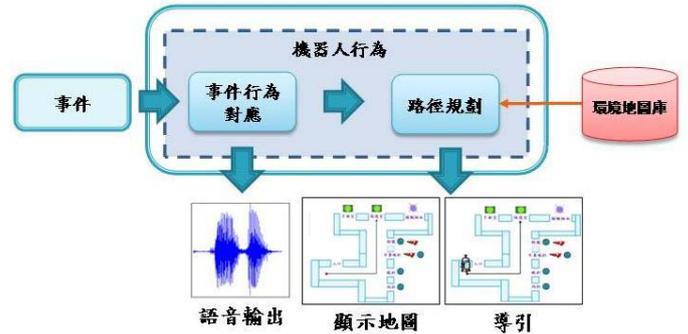


圖 4 Domain Ontology 架構圖

除此之外，我們所定義的 Ontology 有 2 種關係 (Relations) 分別是聚合 (Aggregation) 及結合 (Association)。連接一個 Domain 和它對應的 Category 是聚合 (Aggregation) 的關係，而 Concept 跟 Concept 之間屬於結合 (Association) 關係。

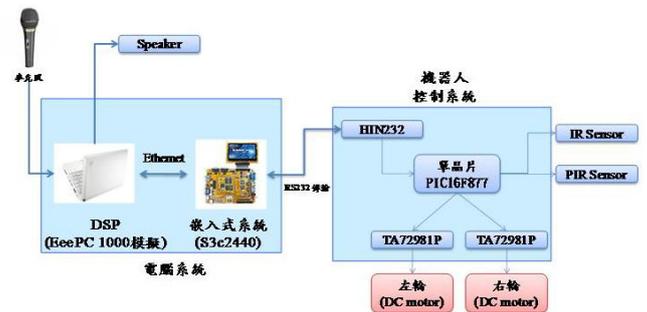
(三) 機器人行為模組

機器人行為模組是負責將行為透過各種方式回應給使用者，例如：語音、顯示地圖和行走等，如圖 10 所示，為機器人行為模組細部架構。本模組將知識推論出來的事件，透過事件與行為對應的方式，把發生的事件轉換成所定義的語音回應給來訪者，告知來訪者目前目的地的資訊。倘若事件為正常，行為模組就會直接將機器人目前所在位置與目的地作最短路徑規劃，然後開始進行導引的工作。如果從推論知識引擎接收到是來訪者要求看地圖，行為模組可直接將規劃好的路徑顯示在畫面中。機器人行為模型提供語音輸出與來訪者進行互動，也能依據推論所得到的行為，回應資訊給來訪者。對話的互動語音可由專家設計，主要是依據語音辨識和語音合成技術來與訪客進行對話。在本研究的環境設定中，機器人要能夠讀取環境地圖資訊，並且知道自己目前所在環境位置，就能夠自動進行至目的地位置的路徑規劃，達成點對點的路徑規劃。路徑規劃的方式有許多，如：距離轉換法[9]、路徑轉換法[7,8]等，本研究中所使用的機器人路徑規劃演算方法參考[4,5]，利用路徑轉換法來進行機器人的路徑規劃。



三、系統實作

如圖 6 所示，導引機器人整個系統程式開發主要是由嵌入式系統 (Samsung S3c2440)、數位信號處理器 (由 EeePC 小筆電模擬)、單晶片 (PIC) 電路和直流馬達電路所組成。



本系統提供語音介面，讓來訪者可決定要導引或顯示地圖和想去那個目的地，並讓來訪者決定是否要前往目的地。而導引過程中也會以廣播說明機器人要去那個目的地，如果其他人也想去這個目的地，也可以跟隨機器人的腳步前往，如圖 7 所示，為導引機器人的實體圖。



圖 7 機器人的實體圖

本系統採用焦熱式紅外線感測器來偵測有沒有人接近機器人。當來訪者接近時，機器人能主動問來訪者需要什麼服務。當來訪者只想觀看環境地圖時，機器人會顯示路徑規劃後的路徑給來訪者觀看，提供如何自行前往的路徑給來訪者。本系統也提供介面給專家來客製環境地圖、情境事件建立和修改，使用簡易的介面來置換地圖、辭書和推論規則，使導引機器人可降低機器人與場所之耦合性。專家或使用者只要透過環境地圖編輯介面，就可修改或更新所需要的內部環境地圖。

四、實驗結果

建立非特定語者語音模型，須將不同的語者錄製語料混合訓練，但是收集多人語料極為困難，因此本研究的語料來源，主要是收集國立成功大學 20 位學生的語音資料，以建立不特定語者模型。要以建立非特定語者語音辨識模型，可以用語詞(Phrase)、音素(phone)、韻(syllable)和次音節(subsyllable)為單位來建立模型，在本研究中採取以次音節為單位。

(一) 實驗一

實驗對象為男性和女性各 10 位，並且建立男性與女性的語音模型。但由於男性與女生語音頻率不同，如果混合在一起訓練會降低辨識率，因此本實驗將分開建立男/女性模型。

實驗總共有 30 個關鍵字，每字收集 10 次語音語料，測試方法也是使用 K 次交叉驗證(K-fold cross-validation)來計算出辨識率。另外，我們考慮每個實驗者的口音。實驗中，每個人需錄製 30 個關鍵字語音，總共會建立 30 個語音模型，實驗結果顯示，男性的語音辨識率為 74.15%，女性的語音辨識為 72.58%

實驗結果

本研究非特定語者的辨識率是 73.37%，雖然看是不高，但是與[2]相比較起來，如果將實驗人數設定相同，男生 6 個人和女生 2 個，則辨識率就提高至 86.24%，與[2]辨識率差不多。在此實驗中，當越多人的語料一起訓練，辨識率就會下降，主要是因為將不同語者的語料混合訓練，雖然裡面還是包含不同語者的語音特性，但也同時降低了每一位語者在模型中所佔的比例。另外，辨識率會下降的可能原因是每位語者的錄音環境沒有完全相同，可能因為錄製時背景環境噪音干擾的關係，使得彼此差異性變大。

在過去的研究指出[2,6]，建立非特定語者模型通常辨識率會比特定語者模型還要低，主要是因為訓練語料內不包含不同語者之間的差異性，這是非特定語者模型所面臨的問題。

(二) 實驗二

在場地租借不易及經費不足的情況下，本實驗是以成功大學工科系五樓為實驗環境，來模擬實際醫院地圖，如圖 8 所示，而每個目的地上面的情境狀態，定義於圖 9。

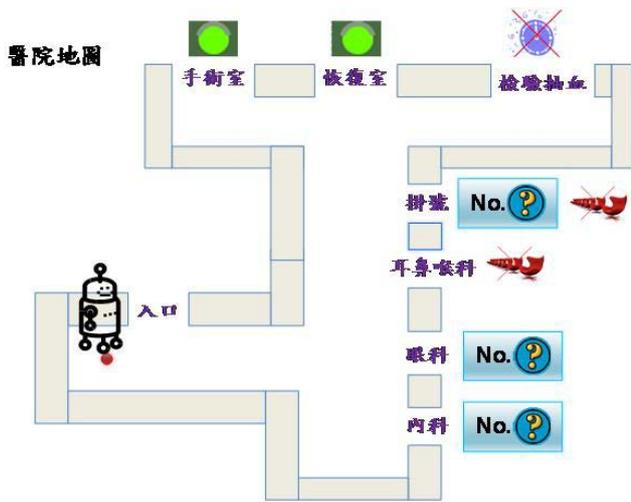


圖 8 醫院情境與環境地圖

實驗過程是當來訪者接近機器人時，機器人會主動詢問來訪者是否需要服務，來訪者可要求導引或顯示地圖，而機器人會提供情境感知之資訊，讓來訪者決定是否要前往目的地，如果回答是，則會引導來訪者至目的地

情境	圖案
未到開放時間	
正常	
目前叫號	
椅子沒位置可休息	
緊急情況	
火災	

圖 9 定義情境感知之事件

實驗結果

當 2 位訪客接近導引機器人的時候，導引機器人會說「你好!很高興為你服務，請問要幫您導引或顯示地圖」，然後第 1 位女性訪客跟他說「導引」，然後機器人接著問「請問您要去那裡?」第 1 位訪客回答說「手術室」，機器人會在問訪客是否要去「手術室」，但 2 位訪客突然改變心意，於是第 1 位女性訪客回答「不是」，接著機器人會再詢問一次「請問您要去那裡?」，

2 人決定到恢復室，於是第 2 位男性訪客回答說「恢復室」，機器人會再次確認是否要去「恢復室」。直到第 1 位或第 2 位訪客回答是之後，機器人會經由推論機制來取得恢復室的目前情境正常，並無任何訊息要給這 2 位來訪者知道，於是機器會直接開始進行導引，帶他們至目的地。導引的過程中機器人會一直重覆的說「要去恢復室請跟我來」，如果當其他路過的人聽到這個訊息，剛好也是想去恢復室的人，也可以跟隨機器人至目的地。

五、結論

本研究提出一具情境感知之導引機器人的架構，溝通介面採用非特定語者語音辨識技術，提高機器人與人的互動性，以符合實際應用情境。導入情境感知與機器人做結合，使機器人更具有人工智慧，並且提出情境感知結合辭書架構。這個架構不只可應用在醫院，還可應用於研討會或公家機關等活動上，而且可依需使用者需求更換地圖或定義情境事件。

實驗證明，本研究所提出的具情境感知之導引機器人的架構方法是可行的。本機器人所使用的非特定者的語音辨識率為 73.37%，足夠讓機器人與人用語音進行互動，而且本系統可透過感測器和領域辭書得到知識推論資訊提供給來訪者，並讓機器人使用單位可編輯環境地圖與情境事件之更新，以滿足他們不同之需求。

六、致謝

本研究承蒙國科會計畫 NSC95-2221-E-006-158-MY3 經費部分補助，特此感謝。

七、參考文獻

- [1] 日本愛知博覽會 .
from <http://www.expo2005.or.jp/tcn/robot/>
- [2] 李佳慧，不特定語者國語語音字詞辨識系統研究，中原大學資工所碩士論文，2001.
- [3] 李健興，適用於家庭應用之推論引擎技術研究，工研院期中報告，2006.
- [4] 林文景，克服樓層障礙之智慧型機械人，國立成功大學工科所碩士論文，2006.
- [5] 徐世昌，郵件機器人之路徑規劃與實現，國立成功大學工科所碩士論文，2007.
- [6] 楊永泰，隱藏式馬可夫模型應用於中文語音辨識之研究，中原大學資工所碩士論文，2000.
- [7] Zelinsky, A., Using path transforms to guide the search for findpath in 2D, International Journal of Robotics Research, Vol. 13, 315-325, 1994.
- [8] Zelinsky, A., Jarvis, R. A., Byarne, J. C., & Yuta, S., Planning Paths of Complete Coverage of an Unstructured Environment by a Mobile Robot, Proceedings of the IEEE International Conference on Robotics Automation, 1993.
- [9] Zelinsky, A., & Yuta, S., Reactive Planning for Multiple Robots Using Numerical Potential Fields, In Proceedings of Intelligent Autonomous Systems IAS-3, 1993.