

孩子的大玩偶

智慧型互動陪伴機器人

南台科技大學 機器人研究中心 謝銘原

找不出更多時間來陪伴小孩，已漸漸成為許多先進國家忙碌爸媽的難題。在台灣，越來越多的孩子，放學後得參加各式各樣的安親班、課輔班；或回到空蕩蕩的家。許多家長，因工作忙碌，甚至得利用晚上或假日加班，同時托育衍生的教養、安全及經濟負擔等問題，更成為許多夫妻不願生育或少子的原因之一。

日本的機器人研發腳步，往往領先其他各國，在家庭陪伴照料上，也有成熟的機器人產品問世。其中，以 NEC 公司研發之祿母機器人—PaPeRo（如圖一）最受矚目。PaPeRo 在 2005 年愛知世博會發表時，相當受到歡迎，他可以一問一答，會繞口令，可同時分辨 5 個人的聲音、方位及交談，還會追隨人，辨識主人，稱讚他「卡哇伊」，它還會臉紅害羞。同時可讓家長透過手機操控 PaPeRo 機器人，來監看家中狀況，或透過簡訊給 PaPeRo，機器人可以『講』給小孩聽。



圖一、日本 NEC 公司開發的機器人“PaPeRo”

類似這樣的『孩子的大玩偶』，該需要哪些技術，考慮哪些問題，才能實現祿母、陪伴、甚至教養的功能呢？以下將就國內外研究之相關實例，以及南台科大機器人研究中心開發之『Fairy』機器人，分成幾個部分來討論。

機器人該長怎麼樣？

機器人的身體及身形（機構及配置），跟任何人一樣，往往是視需要而定。在家與孩童互動，太高大或笨重的機器人，會給孩童壓迫、危險的感覺；太小或過輕的機器人，又恐讓頑皮的孩童『欺負』。以 3~6 歲的孩童身高約為 90~120cm 來看，坐下或蹲下時，視線高度約在 40~60cm，因此，機器人的身高也應在此範圍為佳。以 PaPeRo 為例，他的身高為 38.5cm，重 5kg，與孩童互動時會略微仰頭，讓孩童覺得相當安心。

為察覺環境及與人互動，視覺、聽覺、語音及觸覺感知系統是機器人不可或缺的部分。此外，由於機器人的安全考量往往高於功能考量，因此如何設計讓機器人不會撞上物體及人們，可考慮超音波感測、紅外線感測、或雷射偵測等，來輔助機器人視覺、觸覺等系統之不足。

機器人運動的方式，與其工作或服務的場所環境有關。以平坦地面居多的居家環境，可考慮採用較易操控的輪式運動方式。而如需克服高低地面、爬樓梯或多障礙環境之場所，則應考慮使用足形（人形）運動的模式。不過，實際應用的觀點來看，輪形運動具有較佳的操

控性能、速度及穩定性，往往是祿母機器人採用的運動方式。

不像人的機器人？

至於機器人的外型，從過去許多機器人的研究來看，很多機器人（Robot）並不像人。如何研發外觀和功能都與人相同或相似的機器人，一直是科學家努力不懈的目標。此外，另一種機器人—機器寵物，則以親近人類生活為目標，設計成像動物或可愛玩偶之外型，但同時也具有表達能力，可融入居家生活的一環。日本 SONY 公司推出的機器狗愛寶（AIBO），就是一個受歡迎的機器寵物，具有人工智慧，能進行簡單的動作及功能，由於外形與一般的小狗極為相似，行為也栩栩如生，所以推出後受到全世界的矚目。

祿母機器人的角色，著重看護陪伴老人或孩童，其外型與功能類似寵物機器人，多半需為可愛，讓人容易親近。唯如何增加其思考、對話、情感、互動及安全保護之能力，則是設計技術需突破之處。名古屋市商業設計研究所開發的『IFBOT』機器人，就是一個能夠理解人的感情並與人溝通的機器人（如圖二），他體內存儲了數萬種對話場景和 40 種面部表情，不僅能夠識別交談者的語言，還能判斷對方的感情，並通過“眨眼睛”、“轉動眼球”等方式表達自己的“感情”。與“IFBOT”說話，老人大腦的思維能力得以激活，從而緩解憂鬱孤獨並避免患健忘症。



圖二、機器人“IFBOT”與人語音交談之情形

國內的研究，以國立交通大學的研究較為成熟，其中，他們在『人臉辨識及其在家用機器人與人互動之應用』之研究中，開發適用於特定人搜尋之人臉辨識系統。此系統提供每個家庭成員擁有專屬辨識自己臉的類神經網路，每個類神經網路僅需具備及負責分辨該所屬家庭成員之能力即可，這樣可以使辨識時間不會隨著家庭成員的增加而增加，進而提高辨識的正確性及效率。

由於看護陪伴的機器人通常不像工作機器人強調工作及性能之精確度，它著重的重點在於提供看護，強調互動關係、自動化之輔助能力。在科技日新月異的發展下，讓互動式看護機器人成為家庭的一員，是未來的趨勢；這樣不僅可減少人力的負擔，也會使生活更加便捷安全。

Fairy 機器人

南台科技大學機器人研究中心最近開發了一個名為『Fairy』的居家看護機器人，該機器人不僅具有容易親近、可愛靈活的特性，也展現可與人互動、服務及監控之能力，包括能與家庭成員互動、溝通、娛樂及保全監看等。機器人系統的配置，亦以操作方便性及實用性來設計，並結合豐富的語音聲光動作，來提升使用意願及興趣。

機器人的系統配置，包括影像處理系統、馬達驅動控制系統、遠端監控伺服系統、語音辨識理解系統、避障偵測控制系統、以及電力系統等六部分，圖三分別表示機器人外型設計及實體外觀圖。機器人之影像處理系統僅使用低價且介面簡單的羅技 Webcam，以 Microsoft 視窗程式設計影像處理介面。當 Webcam 擷取動態影像後，機器人（影像處理系統）可進行動態影像邊緣偵測、顏色判別、物件判斷及追蹤等影像處理，來控制執行避障、互動追隨及互動表情呈現等動作。



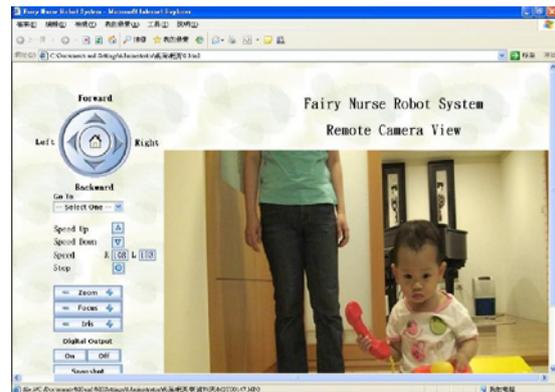
圖三、機器人『Fairy』之外觀

機器人之聲音辨識系統，係採用 VCMM 語音聲控模組。該模組將語音辨識結果透過 RS232 傳至機器人內部的主控電腦。機器人由語音及應對決策系統，控制其表情呈現模式與語音回應等工作。此外，機器人結合紅外線及雷射掃瞄儀組成避障偵測及控制系統，將其偵測環境之物件相對配置資訊，提供給主控程式決定避障運動之動作模式。而機器人的電源系統，則由兩組蓄電池來提供各子系統所需之電力。

Fairy 機器人除了可與家中的小孩與老人進行互動陪伴的工作外，於居家安全防護上，也是一個很好的幫手。為使 Fairy 機器人具有緊急應變之能力，除重要事件提醒、遠端監視、遠端操控（如圖四）等功能外，未來希望將設計使機器人能與主人手機、警局或醫院連線，當家庭成員不小心跌倒、昏倒或任何突發狀況時，機器人透過影像處理判斷確認後，可先傳送即時現場影像資訊給家中其他成員的手機，同時也發送訊息、電話到最近的警局或醫院，讓醫護救援工作能在第一時間做出最適當的應變措施。此外，未來也將增強防護措施，例如以聲音做出警告、與保全系同連線等功能，來增強保全功能的效率。

我們所發展的居家看護機器人『Fairy』，目前完成多種互動功能之設計，如跟隨、避

障、追球、互動呈現（如圖五）、簡單語言交談、問候等，未來仍有許多需要加強改進的地方，以期符合實際的需求。語音辨識與分析部份，未來希望可以結合類神經學習網路，來提高語音辨識與語意分析的辨識率。另外未來的語意資料庫也希望可以增加至 1~2 萬組，讓對話的模式可以多樣化，就像與真人溝通一樣。



圖四、遠端監控時之介面畫面（可看到機器人前方之畫面）



圖五、Fairy 與家中成員互動之情形

中文語音辨識及理解的困難

中文語詞辨識的困難在於中文語言結構不易斷詞，與英文每個詞（word）間會有空白間隔之斷詞容易相比，中文斷詞的不易也造成語意理解及分析的困難。中文自然語言處理（理解及意義）之最小的基本單位是「詞」；而在中文語言結構上，每個最小的單位是

「字」。因此在中文文章中只有字的界線而無詞的界線。但唯有將文章正確斷詞，才能進行自然語言分析（關鍵詞組、句法、語義等）。

由此可知，中文斷詞是自然語言處理之基礎且重要的工作。如果可以將文章（或語音段落）中特定詞句（人名，物名，時間，事件，數字，組織名稱，動作，對象）辨識出來，將會有效提升斷詞之準確度。當然中文斷詞與特定詞辨識的成效，將影響許多自然語言處理（NLP）的應用成敗，例如，文件管理，資訊檢索、語音辨識、機器翻譯等。

機器人語音辨識的另一困難，則是語音擷取及抗雜訊的問題。一般而言，透過指向性高的麥克風，可以取得音訊品質較佳的聲音，並有利提升語音辨識的正確率。唯這類的麥克風，大多需要配置在發話者的嘴巴附近（通常是耳掛式聽筒及麥克風）。這樣的配備相當普遍，但缺點是不方便使用，特別是於家庭日常生活中使用。因此，機器人之聽聲辨位系統之設計，可以讓機器人朝主要語音的方向收音；或利用語音訊號強化技術（抗回音、噪音及雜訊），提升語音擷取的正確性；或加強特定語者之語音特徵學習演算，利用反覆聆聽特定語者語音，來提升語音辨識的辨識率。

影像處理於人臉、表情及動作之辨識

利用電腦或微處理器系統來辨識人臉 (Face Recognition)，到目前為止都還是一件困難的工作。雖然目前影像處理技術已進步到相當好的境界，但該如何有效辨識人臉卻還不是一件容易的事。由於影像取得之人像在大小、方向、光線及背景均存在極大的變異性，因此造成人臉辨識上的困難。如何找到足以代表不同人臉的特徵，來開發有效的辨識方法，一直是國內外相關研究學者所努力的目標。

近幾年，陸續有研究提出一些實用的方法，其中包括以類神經網路架構進行主要元素

分析 (Principal Component Analysis, PCA)。但該如何去找出（或定義）人臉的特徵 (features)，是此問題的主軸。透過特徵臉 (Eigenface) 的歸納及建立，可建置人臉辨識系統。亦即，可使用 PCA 歸納出相似臉的集合，再以此訓練類神經網路來區分相似臉。麻省理工學院 Pentland 等人，提出以影像水平投影平均值做為類神經網路的輸入特徵，這是因為從觀察人像角度的變化，可發現主要與水平變化有關，也就是說以人臉的水平投影來當特徵，會比由垂直投影來當特徵，有較小的變異性。

然而在即時人臉辨識的研究上，目前最大的困難仍是即時環境變化的複雜性。人臉辨識的第一步，是如何將人臉正確地定位，如果人臉定位不準確，辨識誤判的機率將大增。一般的作法是利用膚色過濾，來偵測可能為人臉的區塊，同時減少影像處理運算的資訊量。但往往在實際環境下，因光線變化太快使得影像曝光變化，造成膚色因光線太亮而偏白，或因光線太暗而偏黑。此時，可利用 $YCbCr$ 色彩模型來進行亮度補償修正，來解決光線的影響。

此外，在人臉辨識的部份，一般的作法係將偵測到的人臉區域以方形框萃取其影像特徵，唯實際應用發現，使用太大的方形框會將背景框了進來；反之太小的框則會失去部分的人臉資訊。由於人的臉型多近似橢圓形，因此如果以橢圓形框取得人臉影像，則可有效去除背景。特別是在動態擷取時，環境可能隨時在變，若將背景也框了進來，很容易造成人臉辨識特徵的誤判。

追蹤人臉，改變機器人位置使人臉移到影像畫面之中央，並調整相對距離使人臉大小變化至適當大小，是有效解決動態辨識人臉的方法之一。這樣的想法，已成功應用在居家保全機器人 PaPeRo 上。當 PaPeRo 與交談對象距離太遠時，往往會利用影像追蹤及相對運動，讓交談對象的人臉影像維持在畫面中央，並移

動靠近對象，使人臉變大至適合的大小。一方面使人臉（特別是表情）辨識的正確性提高，另一方面也可提高交談及互動的親切感。

此外，人臉辨識的主要應用，即是表情辨識。許多研究指出，人的臉部表情變化與眉宇、嘴型變化息息相關。利用笑（微笑、高興、興奮、冷笑）、哭（悲傷、痛苦）、怒（憤怒）、害怕、驚訝…等，各種表情之各部特徵，建立不同表情之特徵模型。當機器人（系統）可辨別互動對象之表情情緒，自然可延伸產生適切的表情或應對。此外，以視覺監控對象之手勢及人體動作，可提升機器人看護及協助居家環境監控的能力，也是降低居家意外的重要課題。然而由於某些手勢、動作（昏倒、睡覺）具有相似性，區別錯誤將使得應對動作的誤判，因此必須經由學習機制的建立，讓機器人（系統）反覆訓練，才能正確判斷，這或許也是這類研究的困難點。