

行政院國家科學委員會專題研究計畫成果報告

子計劃一：家用機器人系統之資訊與網路平台研發

計畫編號：92-2213-E-009-004-

執行期限：92年8月1日至93年7月31日

主持人：胡竹生 國立交通大學電機與控制工程學系

計畫參與人員：蘇宗敏、吳立偉、周春成、康創閔 國立交通大學電機與控制工程學系

一、中文摘要

本計畫將針對家用機器人之資訊系統軟硬體平台進行研發，建立一個機器人的視覺與聽覺平台做為追蹤和定位，而其重點在於省電高效能以及具備現代化資訊整合與連結能力的機器人發展平台。此外，希望機器人能與人類一樣具有視覺與聽覺判斷環境的能力，以挑戰在高度未知與動態環境中達成任務的目標。在系統架構上，本計畫提出三層式控制架構改善機器人在處理工作時硬體結構與軟體規劃的問題。此三層式架構以資料處理硬體的角度分為：1.即時控制層、2.閘道伺服器層、3.網際網路層。而以學習與智慧成長之軟體角度可分為：1.反射動作層、2.單一智慧層、3.合作式學習層。在實作上，本計畫預計使用嵌入式技術來建立 Robot 上的中央控制單元，並採用 Embedded Linux 為其作業系統。在視覺與聽覺訊號前處理上，本計畫研究如何萃取有意義的資訊，亦即移除背景或雜訊。在聽覺方面本計畫實作麥克風陣列以純化語音訊號，而視覺方面則採用統計的方法作強健型的背景濾除。這些成果搭配前述的三層式架構，將建構一 Robot Thin Client，透過 Robot House PC 的閘道功能，可以運用網際網路上無限多的訊息資訊提供給機器人系統，並可運用網路上無限多的計算資源建構一個複雜且龐大的學習與合作的機制。

關鍵詞：機器人，家用機器人，嵌入式系統，即時多工控制，麥克風陣列，影像背景濾除

Abstract

The goal of this project is to develop an

information platform for home robot systems and build a vision hardware platform of robot to tracking and position, with special emphasis on power saving and modern information and connection capability. Besides, the goal is to enable the robots with human-like vision capability to estimation the environments. By doing so, the robots can complete missions in highly unknown and dynamic environment, i.e., to build a real-time tracking system to handle the obtained frame in 1/15 second. Using embedded system design technology, the project is going to construct a development environment based on real-time multi-tasking control architecture. A three-layer control structure is proposed to define the hardware and software technology involved in the system. From hardware viewpoint, we have: 1. Real-time Control Layer; 2. Gateway and Server Layer; 3. Internet Layer. On the other hand, the software layers are defined as: 1. Reflective Action Layer; 2. Unitary Intelligence Layer; 3. Cooperative Learning Layer. The project will use RISC CPU to design the central control unit for the robot. Embedded Linux will be used as the operating system. On the vision and hearing research, this project studied the methods of extracting meaningful information from image and sound signals. For hearing, a microphone array was designed to purify speech. For vision, a robust background removal algorithm is proposed based on statistical methods. These results combined with the three-tier architecture form a thin client robot. Via Internet connection, the robot can own an infinite resource and computing power to establish a huge and complex learning and cooperative mechanism.

Keywords: Robotics, Home Robot, Embedded System, Real-time Multi-tasking Control, Microphone Array, Image Background

Removal

二、緣由與目的

機器人之研究與應用在近年來已從工廠自動化，漸漸走入一般人的生活。日本本田汽車公司於 1996 年底開發完成的本田機器人 [1]。本田機器人可以依據地形狀況決定自己如何動作，因此能夠於上下樓梯及斜坡進行時保持身體直立。早稻田大學最早於 1970 年代開始研究二足步行機器人、1992 年成立人形機器人研究中心。另一方面他們開發的 Handaly-2 機器人同時進行模仿人類頭部與雙眼方面的研究 [2]。除了本田公司的本田機器人外，SONY 公司已完成商品化產品 AIBO 犬型寵物機器人，並可用來參加 ROBOCUP 機器人足球競賽。其他學術機構或公司所提供的人型機器人更是所在多有 [3-8]。

可移動式的機器人走入家庭似乎已經是一個重要趨勢，然而現階段具備高度智慧的移動式機器人大都屬於較龐大的設備。對於民生化、家庭化的需求仍有一段距離。一個家用機器人系統，所應有的要求應該是一個體積小、省電、運作速度快、網際網路資訊介面寬廣，具有影像與聲音系統的處理環境，擁有對外在環境之感測介面，以及互動性和家電系統有一定之整合能力的設備。甚至擁有一定程度的人工智慧與擴充能力。反觀現階段環境，要符合以上要求的彈性發展系統並不多。其次，大部分與網路結合的機器人，以 tele-operation 或遠端控制居多 [9-11]。就一個具備網路能力的 home robot 而言，其所需具備的網路功能，似乎應再加上現代化資訊平台如 Multi-agent 等技術 [12]。

考量未來整合應用的趨勢，一個 Home Robot 系統及可能成為家用資訊與控制之出入口，由於其兼具移動式能力，為目前所有方案(如一般 Gateway)等所不能及的。其次，Home Robot 結合網路代表虛擬世界 (Virtual

World)與實體世界 (Real World)的整合。許多例子均顯示，這兩個世界的結合將延伸出無限的想像與應用的空間。近十年來在 Robotics 中熱門的研究題目之一為多機器人之間的合作問題，其相關的人工智慧研究稱為分散式人工智慧 (Distributed Artificial Intelligence, DAI)。DAI 的進展目前以模擬居多，也就是大部份均在虛擬世界中進行。與實體結合時均礙於計算平台不足或系統架構過於侷限(如通訊協定不夠開放)而無法發揮。近年來網際網路的成長為這個合作的研究提供了愈來愈好的環境，而 Home Robot 則將因為這些研究的成果應用，對人類產生更大的價值。Home Robot 由於其互動性高於其他電子裝置，將成為未來世界裡一個重要的 agent。雖然 multi-agent 這個觀念早已被提出並有大量的研究，但以 Home Robot 為主的技術研發並不多見。其中很明顯的是缺乏合適的軟硬體平台。

此計畫已完整落實解決了資訊，控制，電機機械與系統整合的問題，讓人工智慧與學習理論相關學術研究擁有一個很好的橋樑。更可進一步驗證多層式即時多工控制理論架構的可行性，並能夠給予產業學術界注入一股新的契機與力量。

三、結果與討論

本計畫 92 年度主要完成四項結果。1. 四足機器人之雛型；2. OMAP dual-core processor 硬體平台與 embedded linux 建立；3. 聽覺系統之語音純化介面；4. 影像背景濾除，其內容簡述如下：

1. 四足機器人之雛型：

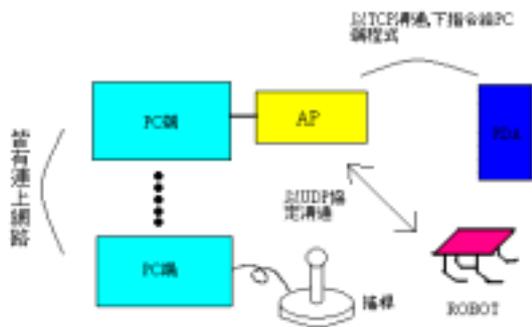
在 92 年度的研究中，我們已經建立一四隻腳，16 個自由度的四足機器人的機構及控制系統。使用 Servo Motor 作為其驅動源。利用單晶片給 Servo Motor 控制訊號，來控制機器狗的動作。其次，利用無線網路模組，我們

可以建立網路上電腦與其通訊介面，透過網路溝通，使四足機器人具備強大的運算能力及網路資源。四足機器人外觀圖如下：



圖一、四足機器人外觀圖

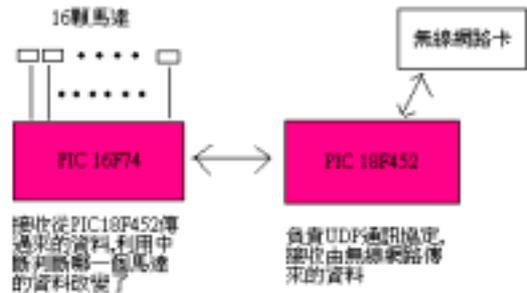
其系統架構圖如下：



圖二、四足機器人系統架構圖

機器人是透過一無線網路模組連上無線網路，所以所有有連上網路的機器都能夠透過網路來控制他。在機器人上用了兩個單晶片，PIC 18F452 負責無線網路模組部份，負責接收傳送由網路傳送而來的資料。另一個 PIC 16F74 負責控制 16 顆馬達的動作，根據由 PIC 18F452 傳過來的資料下指令動作。無線網路模組的通訊協定定為 UDP，所以只要撰寫上層 USER 端的程式，也是以 UDP 為通訊協定，透過 AP，兩端就可以連上傳送資料，再根據已經訂好的資料格式，透過馬達控制器(即 PIC 16F74)處理，送出控制訊號。

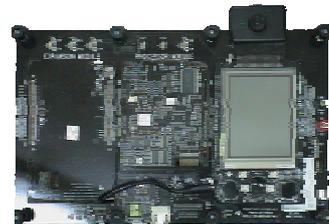
連接每一個 Servo Motor 有三條線，其中兩條為 VCC 跟 GND，第三條為訊號線。Servo Motor 旋轉的角度，是由輸入訊號線之訊號的 Duty Cycle 來決定，而這個 Duty Cycle 就由 PIC 16F74 來產生。其控制架構圖如下：



圖三、四足機器人控制架構圖

2. OMAP dual-core processor 硬體平台與 embedded linux 建立：

OMAP 處理器包含 RISC(ARM)與 DSP 雙核心處理器，透過這兩種類型的處理器的合作，讓一些耗時間的演算法得以快速執行。本計劃已開發在 OMAP 開發硬體(Innovator)上的 Embedded Linux 平台。如圖四所示，Innovator 平台包含許多週邊設備如網路介面、音效輸入與輸出、LCD 與 CMOS camera 等週邊。



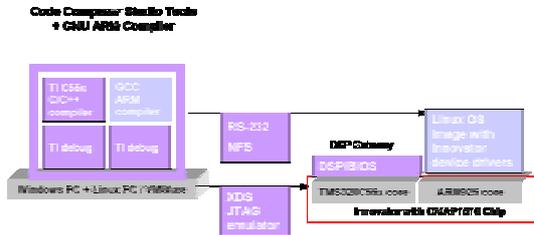
圖四、OMAP 開發平台

本計劃研究之 Embedded Linux 平台與測試環境包含：

- VMware
- ARM linux patch[13]
- Montavista OMAP Processor and Innovator Peripherals patch[14]
- DSP Gateway OMAP DSP patch[15]
- ARM 編譯器
- DSP Gateway

OMAP 處理器包含 ARM925T 處理器與 TMS320C55x 數位訊號處理器 DSP Gateway 則是讓載 ARM 處理器上執行的 Linux 作業系統可以與 DSP 通訊的一套軟體。DSP Gateway 包含兩部份,ARM 上所執行的 Linux 裝置驅動程式與 DSP 上的函式庫，靠著這兩者合作達到通訊的功能。DSP Gateway 支援五種裝置驅動程式如下：

- ◇ DSP task devices
 - ◇ DSP task watch device
 - ◇ DSP control device
 - ◇ DSP watchdog device
 - ◇ DSP exmem device
- 程式的開發環境的示意圖如下



圖五、OMAP 開發環境

3. 聽覺系統之語音純化介面：

➢ 陣列訊號處理

陣列式訊號，是指數個相同的感應器排成特定的形狀，接收空間中傳遞的訊號，經過處理之後，達到空間濾波(Spatial Filter)的功能。空間濾波是指利用輸入訊號具有空間上的資訊，在不同的方向給予不同的增益，如此當輸入訊號經過處理之後，只會保留特定方向訊號，其他方向的訊號變小，輸入輸出的 SNR 即可大為提高。但語音輸入訊號屬於近場球形波，原有的陣列式訊號模型無法適用，因此搭配適應性訊號處理的演算法使用，產生合適的空間濾波器。

➢ 設計適應性空間濾波器

引用 Mattias Dahl [16] 提出的架構，將適應性訊號處理與陣列訊號處理整合在一起。可以得到以下的設計適應性空間濾波器的方法。陣列式訊號搭配適應性演算法系統結構如圖六所示。當使用者不講話的時候，持續做 NLMS，產生最適合目前環境的空間濾波器，當使用者講話的時候，則停止 NLMS，將輸入訊號通過空間濾波器做處理，空間濾波器會將使用者方向的增益維持一定大小，而干擾源訊號增益很小，如此即可去除干擾源，純化語音訊號。

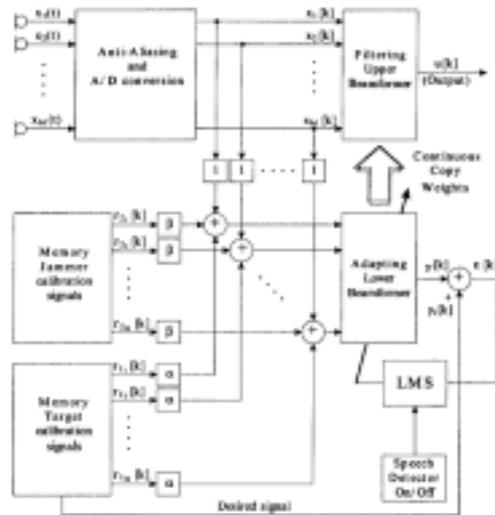
➢ 實驗結果

實驗環境：

使用者：與麥克風陣列中心距離 30 公分，夾角 90 度

雜訊源 1：與麥克風陣列中心距離 37 公分，夾角 30 度

麥克風陣列距離桌面 27 公分 麥克風陣列距離反射面 60 公分。表一為實驗數據整理，處理前後能量的 SNR 可以提升約 10dB。



圖六、適應性陣列式麥克風系統結構

Mic. No.	Order	Signal (dB)	Noise (dB)	SNR (dB)	Improve SNR (dB)
1	1	-18.8012	-30.2613	11.4601	0
8	10	-20.9207	-40.4597	19.5390	8.0790
4	10	-20.6620	-38.8524	18.1903	6.7303
2	10	-21.6847	-38.4107	16.7260	5.2659
8	5	-21.7027	-37.9313	16.2286	4.7685
4	5	-21.4106	-37.4978	16.0872	4.6272
2	5	-21.8621	-37.8480	15.9859	4.5258
8	1	-23.0120	-39.5797	16.5676	5.1076
4	1	-23.6932	-38.0371	14.3438	2.8838
2	1	-26.5263	-38.5457	12.0194	0.5593

表一、實驗數據整理

4. 影像背景濾除：

影像背景濾除基本的概念在最早期 Leung et al [17] 已經明白的說明，一般來說，可以大略分成三個部份：

- 背景模型化(Background model): 如何及時且有效地建立一個強韌的模型，儲存且不定時的更新來當作背景，是背景濾除研究領域中最核心的所在。
- 像素分類(pixel classification)或相減運算(Subtraction operation): 將目前的影像與

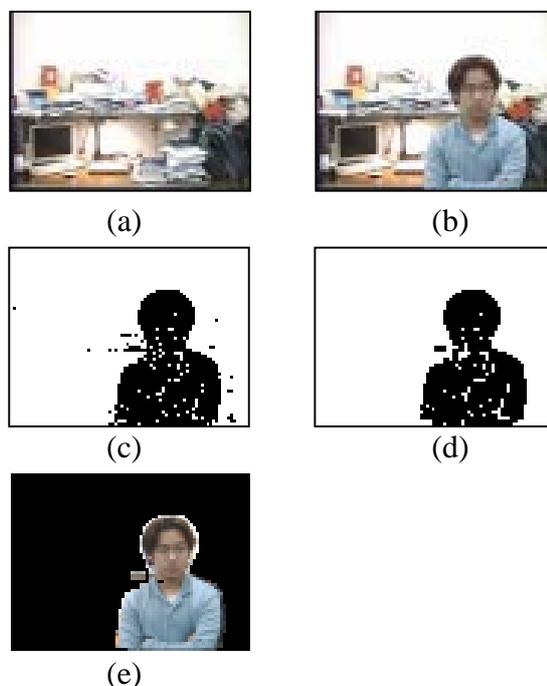
之前建立的背景模型經過此步驟的運算,將移動的物體(可能是前景)與背景的部分分開。

- 臨界點選擇(Threshold selection):最後再對相減運算後的影像通過臨界點的二值化,偵測出我們所要的前景,在這部份,如何找到最合適的臨界點,使得系統能精確的偵測出前景,也是重要的目的之一。

在 92 年度的研究中,我們針對背景模型化的方法進行探討,利用該類方法中的混合高斯分布模型(Mixture Gaussian Model, GMM)來進行影像背景濾除之研究,GMM 是根據所收到的影像資訊,以及其估測的參數,使用「特定」的統計分佈函數來近似其真正的分布。

在許多論文中,像 Stauffer and Friedman[18]、Grimson [19]、...等,都採取 GMM 的方式來建立背景的模型,除此之外,Gao et.al [20]比較單一(single)與混合(mixture)高斯分佈的背景模型,證明出混合模型確實比單一模型在效能上來的好,即使在靜態的場景中也是如此,所以目前幾乎沒有人在拿單一的高斯來做背景模型。在 GMM 的方法中,有一個很重要且困難的問題,就是多少個高斯分布函數的組合才能最近似真實的分布情況,這部分是屬於分類的問題,一般可以使用的�方法有 Forgy's algorithm、K-mean algorithm 等等...。而一旦選定 GMM 的基礎高斯分布函數個數之後,如何組合這些基礎高斯函數組合來近似真實分布,則需要一種機制來尋找並最佳化它,一般是用 EM (expectation maximization) algorithm 來進行。

圖七是在室內環境利用 SONY EVI-D30 攝影機進行影像擷取(擷取速度 30 frames/sec),採用 950 張影像利用 GMM 建立背景模型,由於利用 GMM 所建立之背景模型可以有效的克服背景像素因為燈光閃爍所造成的像素值飄移問題,因此當有前景進入之後,可以有效的分割出前景物體,以便於後端更進一步的應用。



圖七：(a)建立背景模型之影像序列資料 (frame 120),(b)測試背景濾除之影像 (frame 1024),(c) 影像背景濾除之結果,(d) 經過 Morphological filter 之結果,(e)前景影像。

四、計畫成果自評

項目	完成情況
與原計畫相符程度	100%
達成預期目標	90%
研究成果學術價值	高
研究成果應用價值	具實用性
學術期刊發表合適否	是
申請專利合適否	是
主要發現或其他價值	1. 四足機器人之雛型； 2. OMAP dual-core processor 硬體平台與 embedded linux 建立； 3. 聽覺系統之語音純化介面； 4. 影像背景濾除

五、參考文獻

[1] K. Hirai, M. Hirose, Y. Hirose, Y. Haikawa, and T. Takenaka, "The development of Honda humanoid robot," Proceedings IEEE International Conference on Robotics and Automation, pp. 1321 – 1326, 1998.

- [2] S. Hashimoto, S. Narita, H. Ksahara, A. Takanishi, S. Sugano, K. Shirai, T. Hoashi, "Humanoid robot-development of an information assistant robot Hadaly, Proceedings IEEE International Workshop on Robot and Human Communication, pp.106 – 111, 1997.
- [3] T.Ikeda, C..Miyazi, "Swimmable hummanoid robot, " Proceedings IEEE of the 24th Annual Conference of the Industrial Electronics Society, pp1164 –1169, 1998.
- [4] J. A. Driscoll, R. A. Peters, and K. R. Cave, "A visual attention network for a humanoid robot , " Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.96 – 101, 1998.
- [5] S. Hashimoto, "Vision system for humanoid robot-toward emotional man-machine-environment interaction, " Proceeding KES '98 Second International Conference on Knowledge-Based Intelligent Electronic Systems, pp. 269 –273, 1998.
- [6] K. Nishiwaki, A. Ionno, K. Nagashima, M. Inaba, and H. Inoue, " The humanoid Saika that catches a thrown ball. " Proceedings IEEE International Workshop on Robot and Human Communication, pp. 94 – 99, 1997.
- [7] R. A. Brooks, "Behavior-based humanoid robotics, "Proceedings of the IEEE/RSJ International Conference on Intelligent Robotics and Systems, pp. 1-8, 1996.
- [8] S. A. Setieawan, H. H. Sang, J. Yamaguchi, and A. Takanishi, "Physical Interaction between Human and a Bipedal Humanoid Robot –Realization of Human – follow Walking, "Proceedings IEEE International Conference on Robotics and Automation, pp. 361 –367, 1999.
- [9] S. A. Setieawan, H. H. Sang, J. Yamaguchi, and A. Takanishi, "Physical Interaction between Human and a Bipedal Humanoid Robot –Realization of Human – follow Walking, "Proceedings IEEE International Conference on Robotics and Automation, pp. 361 –367, 1999.
- [10] S. Lee D.S. Choi, M. Kim, C. W. Lee, and J. B. Song, "An unified approach to teleoperation: human and robot integration, " Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.261 –266, 1998.
- [11] S. Lee. D. S. Choi, M. Kim, and C. W. Lee, "The Distriuted Controller Architecture for a Masterarm and its Application to Teleoperation with Force Feedback, " Proceedings IEEE International Conference on Robotics and Automation, pp. 375 – 380, 1999.
- [12] Nishiyama, H.; Yamazaki, W.; Mizoguchi, F., "Negotiation protocol for proof of realization of cooperative task in multi-agent robot systems," Systems, Man, and Cybernetics, 2000 IEEE International Conference on , Volume: 3 , 2000, Page(s): 1685.
- [13] The ARM Linux Project, <http://www.arm.linux.org.uk/>.
- [14] Montavista OMAP Innovator Linux Patch, <http://www.mvista.com/>.
- [15] DSP Gateway for Linux, <http://dspgateway.sourceforge.net/>.
- [16] Mattias Dahl, Ingvar Claesson, Sven Nordholm, and Sven Nordebo, "Acoustic Echo and Noise Cancelling Using Micphone Arrays," in *Proc. ISSPA, 1996*, pp. 379-382.
- [17] M.K. Leung and Y.H. Yang " Human body motion segmentation in a complex scene", *Pattern Recognition*, 20:55-64, 1987
- [18] Nir Friedman, Stuart Russell, "Image Segmentation in Video Sequences." *In Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Providence, Rhode Island: Morgan Kaufmann, 1997.*
- [19] C.Stauffer and W.E.L. Grimson, "Learning patterns of activity using real-time tracking", *IEEE Trans. on PAMI*,22(8):747-757,Aug 2000.
- [20] X. Gao, T.E. Boulton, F. Coetzee, and V. Ramesh, " Error analysis of background subtraction", *In Proceedings of International Conference on Computer Vision and Pattern Recognition, 2000.*