

# 美和學校財團法人美和科技大學

## 99 年度教師產學合作計畫

### 結案報告書

計畫名稱：使用影像結構地圖與網格型類神經網路於照護機器人視覺定位與引導之研究

計畫編號：NSC 99-2221-E-276-005

計畫期間：民國 99 年 08 月 01 日起至 100 年 07 月 31 日止

計畫主持人：呂全斌

共同主持人：廖俊鑑

研究助理：周美君、李瑞彬、劉偉羿、莊明輝

經費總額： 459,000 元

經費來源：行政院國家科學委員

行政院國家科學委員會補助專題研究計畫  成果報告  
 期中進度報告

使用影像結構地圖與網格型類神經網路於照護機器人視覺定位與引導之研究

計畫類別： 個別型計畫  整合型計畫

計畫編號： NSC 99-2221-E-276-005

執行期間： 99 年 08 月 01 日至 100 年 07 月 31 日

執行機構及系所：美和科技大學 健康產業資訊科技系

計畫主持人：呂全斌

共同主持人：廖俊鑑

計畫參與人員：周美君、李瑞彬、劉偉羿、莊明輝

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

本計畫除繳交成果報告外，另須繳交以下出國心得報告：

赴國外出差或研習心得報告

赴大陸地區出差或研習心得報告

出席國際學術會議心得報告

國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年  二年後可公開查詢

中 華 民 國 101 年 08 月 01 日

成果報告：

- ◆ 呂全斌, 莊明輝, 李瑞彬, 劉偉羿, 2011, "照護機器人視覺定位與引導方法之研究", 2011 南台灣健康照護暨健康產業國際學術研討會, pp. 244-251.

## 照護機器人視覺定位與引導方法之研究 The study of vision-based location and guidance for mobile robots

呂全斌<sup>1</sup> 莊明輝<sup>2</sup> 李瑞彬<sup>3</sup> 劉偉羿<sup>4</sup>

<sup>1</sup> 美和科技大學 健康產業資訊科技系 助理教授

<sup>2 3 4</sup> 美和科技大學 健康產業資訊科技系 學士生

<sup>1</sup> [cplu@meiho.edu.tw](mailto:cplu@meiho.edu.tw)    <sup>2</sup> [xbunnerairyx@gmail.com](mailto:xbunnerairyx@gmail.com)

<sup>3</sup> [tachikoma.lin@gmail.com](mailto:tachikoma.lin@gmail.com)    <sup>4</sup> [trail3424@gmail.com](mailto:trail3424@gmail.com)

### 摘要

本研究是進行一項以電腦視覺對行動式機器人進行視覺定位與引導之處理方法找尋，此系列研究將應用於遠距式居家照護系統之照護機器人的定位與引導上，讓遠距照護者可以透過手持裝置下達目的地指令，照護機器人能自主避開移動路徑上的障礙物，朝著目的地前進，進而執行照護的工作。目前以視覺為基礎的定位方法有兩種，一是以機器人裝置上的攝影機視訊為視野，透過室內擺設物品的影像特徵來建立數據地圖或拓撲地圖，然而室內物品的擺設的變化，容易影響到地圖的可信度；另外方法則是以室內高處架設的攝影機視訊為視野，透過全方向大範圍的視訊畫面來定位，並配合最佳化演算法來規劃移動路徑，該方法的優點是不受室內物品的擺設變動的影響，直接由當下的狀況判定機器人的移動方向，如何以全方向的二維影像來引導三維空間的機器人還尚未有文獻提出；因此，在本研究中，我們以此方式為基礎，嘗試找尋兩項問題的解決方法。

**關鍵字：**遠距照護系統、視覺定位與引導、照護機器人、影像處理

### ABSTRACT

In this study, we search a method of the vision-based location and guidance for mobile robots. This method will be applied to care robots of tele-healthcare system. We model indoor environments with PTZ cameras using the following visual cues: local features for moving regions recognition and the discrete cell regions for real time motion planning. The methods have been considered a solution for image processing in autonomous mobile robots location and guidance. Two main types of vision-based robot location are available: one is image information is acquired form camera on robot; the other is image information is acquired form camera on ceiling. Our method is based up on the second type, and used for mobile robot location. In this study, we will propose an image structure maps with cellular neural networks (CNNs) for motion planning to reach a fixed target in an environment with obstacles. The image structure map also can be used for another motion planning algorithms, like as genetic algorithms motion planning. Lastly, we will show that the proposed method successful locate and guide the mobile robot by experimental results. And, some recommendations and relevant discussions are given. In the future, the method will be extended to other applications, such as security service robots and automatic poster robots.

**Keyword:** Tele-healthcare System, Mobile Robot, Vision-based Location and Guidance, Digital Image Processing

## 前言

經過行政院經建會的調查[1]，2008年國人超過65歲的人口比例為10.4%，預估在2019年時，將會超過14.7%，之後快速上升，台灣已經邁入高齡化的社會。然而，高齡化的社會伴隨而來的是人口快速老化、疾病型態的慢性化與身心障礙老衰人口的增加，國人對健康照護的需求日益增加，而國內培養的護理相關人才畢業後大都投入醫療單位，經過幾年辛苦的臨床工作，部分護理人員即轉任其他職場，繼續從事護理工作的人員也會隨著其他因素(如工作福利與待遇的不佳)逐年減少，我國護理人力不足的現象早已經出現多年，更遑論能夠投入老人照護工作；除此之外，根據行政院衛生署的分析調查，60歲以上的高齡者，平均的醫療支出是其他年齡層的4.5倍，有鑑於此，為了減少健保的負擔與提供國人足夠的需求，行政院衛生署已經在進行一連串的遠距照護相關計畫(行政院衛生署，2008[2])，期望藉由資訊科技的導入，使得照護服務便利化、科技化與自動化，並且降低照護人力的需求。目前國內逐漸形成的老人照護的產業，目前正積極發展各式各樣的遠距醫療與照護相關產品，如血壓、血糖、心跳、體溫遠距管理與醫療照護通訊等生理資訊管理平台；此外在輔助獨立生活感測器[3]則有智慧型藥盒、遠距視訊攝影機、跌倒偵測器、離床偵測器、PIR動作偵測器、個人隨身壓扣發報器等等；還有尚在發展的多功能照護機器人。其中以為緊急應變機制功能為主的輔助獨立生活感測器最為重要，一旦有狀況發生，受護者則可以透過這些感測器傳送訊息至通報平台，通知鄰近社福機構或醫療單位到家處理；然而，穿戴式感測器存在著一些常見的問題，如容易受到其他電子器材的影響而產生電磁干擾現象，而導致感測器資料流失(ITIS智網-產業評析，2008)[4]，以及感測器因受護者穿戴方式不正確而掉落、電池沒電、忘記佩帶或故障等等，一旦有跌倒或者是急性病症情況發生，照護者就無法獲知受護者的狀況，因此照護者必須透過遠距視訊攝影機，來隨時注意受護者是否有狀況發生；然而，現行的遠距照護系統所使用的攝影機大多為安裝在固定位置上的攝影機，由於攝影機無法移動，只要受護者離開攝影機的視野範圍，遠距照護者則無法取得受護者在家中的影像，為了慎重起見照護者只能通知鄰近社福機構或醫療單位到府探視與處理，這類的情況經常發生，也是現今遠距照護工作面臨的一項難題。尚在發展的多功能照護機器人(Care Robot)則可以協助解決這樣的問題，照護機器人可以裝設許多感測器，如超音波、紅外線、攝影機與溫度等等感測器，由於照護機器人具有行動能力，因此可以移動到室內中的任何一

個位置，協助遠距照護者解決上述的問題；然而，照護機器人必須由照護者以遠操控的方式進行移動控制，在操控的過程中必須讓照護機器人避開路徑中的障礙物(傢俱)，並且找尋適當的路徑讓機器人順利的移動到目標地。為了讓具行動能力的機器人更智慧化，因此機器人自主性行動的研究開始有了許多的探討，這些研究的最終目標是操控者只要給予目標地的指令，機器人則就能夠自主性的找尋最佳路徑與避開障礙物移動至目標地。而這一點對照護機器人相當重要，如此一來才能真正用於居家照護的工作上；我們在此研究中進行機器人自主性行動方法上的探討，並且以影像資訊為主進行機器人定位與路徑規劃方法的研究。

## 文獻探討

資訊技術在醫療與照護的相關研究在近幾年也受到相當的重視，在相關學術文獻的探討如下：Brender於1998[5]提出資訊技術應用在醫療照護上的作法與建議；Schelkens(1999)[6]、Caldelli(2006)[7]與Singh(2007)[8]將影像處理技術應用在遠距醫療與照護上，而這些文獻主要是著重在醫療資訊收集與辨識方法的探討。在照護機器人方面的研究則有，Song等人(2007~2009)[9~12]所發展的家用機器人與保全機器人，主要的研究是著重在機器人於保全或家用的特定功能上，如老人跌倒偵測、機器人臉部控制、機器人追蹤控制、機器人情緒識別、保全警示、行走輔助等；行動式機器人(Mobile Robot)定位的相關研究也在許多的文獻中提出，如Wang(2009)[13]、Cho(2009)[14]、Kim(2007)[15]與Han(2007)[16]等人的研究，這些研究著重在使用各類型的感測器(如Ultrasonic、RFID、Infrared Rays與GPS)來進行機器人的定位，在這些研究的方法上，主要是透過座標的轉換將地圖上感興趣的資訊轉換至機器人自身的座標上，進而控制機器人移動到目的地。定位可以提供機器人具有自主能力的一種重要的基礎，所以定位的研究是一項重要的議題，自主性的移動能力對照護機器人尤其顯得相當重要，這是因為居家環境中會有許多傢俱與複雜裝潢，這些將會嚴重考驗照護機器人的移動能力。相較於其他類型的感測器而言，光感應式的攝影機具有幾項不可獲缺的特性，如攝影機可以取得大範圍的環境資訊與可以使用直覺式的技巧來進行定位控制(就如同人一樣)。使用視覺影像為基礎(Vision-based)的定位方法有兩種：其一是以機器人裝置上的攝影機視訊為視野，透過室內擺設物品的影像特徵來建立數據地圖(Metric Maps)、拓樸地圖(Topological Maps)或認知地圖(Cognitive Maps)。數據地圖[17~19]的方法是使用幾何特徵來描述環境，並且儲存環境中所有的細節，並且設定單張影像參考畫面(Reference

Frame)來做為環境的代表特徵，並以此特徵為標準進行機器人位置定位；拓撲地圖[20,21]則是相反的使用影像參考畫面中相鄰地標間相對空間的位置，並且將相鄰的地標建立連結，依據這樣的基礎與其他感測器所擷取到的特徵，使用識別法則來分析出機器人的位置；認知地圖[22,23]的方式是由人事先對環境直覺來進行定義，定義後的認知地圖則直接給予機器人參考使用，Kuipers[24]在研究中提出五種不同的資訊要素來建置認知地圖，分別是觀測特徵、路線描述、固定特徵、拓撲關係與數據之間關係；然而室內部分結構會有經常性的變化，如家具的移動或變更，參考畫面不易設定，如此一來容易影響到地圖的可信度，此外尚有兩項因素不利於此方法的定位，一是空間背景紋理也不能過於複雜，複雜紋理背景容易影響到物品特徵偵測的準確度，另一項因素則是機器人裝置上攝影機的高度決定環境的視野範圍，高度太低會因環境的物品太多而無法有效偵測到具代表性的物件，高度太高可偵測到的物品可能太少而失去地圖的精確度。另外的定位方法，則是以室內架設高處的攝影機視訊為視野，透過全方向大範圍的視訊畫面來定位與引導機器人，該方法的優點是不受室內物品的擺設變動的影響，由當下的狀況直接判定機器人的移動方向，此方法必須在每條走道上架設攝影機，而每部攝影機畫面必須有小部分重疊，如此一來才能取得完整的環境影像資訊。此方法必須要使用多部攝影機，看來似乎是一項缺點，但現今辦公大樓與居家的保全系統也都設置了許多的攝影機來記錄人員的進出狀況，恰好可以符合此方法的需求。此方法的相關文獻有 Paulraj 等人(2008)的研究[25]，該文獻中提出以兩組攝影機的影像畫面，在單純的環境下配合類神經網路方式來進行機器人的引導，但文章中所使用的攝影機的架設角度必須垂直地面，因此一個環境必須架設須多部攝影機才能取得完整二維影像的視野，此外所使用的影像演算法太過於簡單，潛藏著許多的問題；Cheng 等人(2009)則是使用可角度變化的攝影機取得不同角度的視訊畫面，透過二維影像來計算機器人在實際三維空間的位置，進行環境中物體的定位，之後再將定位後的結果回饋至機器人，這樣的作法必須計算大量的影像資料，才能確實描述機器人於三維空間的位置，運算成本相當高；Gavrilit 等人(2006, 2007)[26,27]曾提出以網格型類神經網路(Cellular Neural Network / CNN)來進行機器人路徑規劃，透過 CNN 的分析讓機器人可以找出最佳路徑前進，CNN 方法[28]可以是路徑規劃的一種較佳的解決方案；然而在 Gavrilit 等人的方法中，僅僅是以二維的虛擬方格圖來進行路徑規劃模擬分析，而全方向影像要如何轉成 CNN 能使用的細胞陣列，以及如何以全方向的二維影像來引

導三維空間的機器人是該方法實際的難題，目前還尚未有文獻提出解決的方法。在上述兩種定位方法中，由於第一種方法下的機器人無法觀察到自身的位置，因此在尋找路徑上的效能較低落；反之，我們認為第二種方法使用外部攝影機，來取得機器人的位置與大範圍環境影像，在機器人的定位與引導效能上較佳，同時較適合應用在動態的居家環境下，這樣的方法能有更佳的彈性與更快速的環境適應性。因此，在此研究中，我們嘗試找出適用此方法來做為照護機器人的定位基礎。

## 研究方法

在本研究的方法上，我們嘗試使用多部 PTZ 快速球型攝影機來擷取全方向大範圍的視訊畫面，配合聲波感測器、電子羅盤的使用。PTZ 攝影機架設在各走道的中間高處，於機器人首次定位之前，先行旋轉一圈取得多方向影像畫面並進行影像接合，之後建立大範圍環境影像結構地圖，以機器人起始位置的鄰近影像區域(地板)作為樣板進行識別，找出相似區域也就是第一次估測的可移動區域，並作為首次 CNN 路徑規劃的區域，之後以循序估測的方式，更新影像結構地圖的正確性，PTZ 攝影機也會隨著機器人的移動而改變取像角度，機器人移動過程中配合聲波感測器、電子羅盤實際方位的訊號進行可移動區域的修正、再識別與移動路徑規劃，直到機器人移動到目標地，同時完成該攝影機所屬精確影像結構地圖的建置，接著由下一部攝影機影像畫面接手引導並重複上述步驟，而機器人返回原地則以先前紀錄的方位角度來移動，直到居家環境有變化而影像結構地圖必須更新，影像結構的重新識別建立與路徑規劃。過程中我們使用邊緣檢測、影像接合分析、Bayes 分類器 [29]、移動物體偵測 [30]、快速標記處理 [31]、網格型類神經網路 [28]等等方法。

此外，我們採 Client-Server 架構，設定負責控制機器人的工作站為 Server 端，行動式機器人為 Client 端，Server 端透過無線通訊 2.4GHz Zigbee 方式下指令，並接收 Client 端感測器回饋訊號，使用 PTZ 快速球型攝影機來取得環境的影像與機器人的位置資訊，藉由環境影像與其他感測器所建置出的影像結構地圖，來達到環境感知的目的，過程中以循序估測的方法更新影像結構地圖資訊，採用 SLAM(Simultaneous Localization and Mapping) 的作法—定位的過程中同時建立地圖，因此，影像結構地圖資訊將會隨著機器人走過的路徑越多，越能精確描述環境資訊。在我們的方法中包含了幾個階段的運算(如圖一所示)，如影像結構初始化分析、行動式機器人定位、影像結構地圖循序估測建置、CNN 路徑規劃、影像地圖結構更新與儲存。

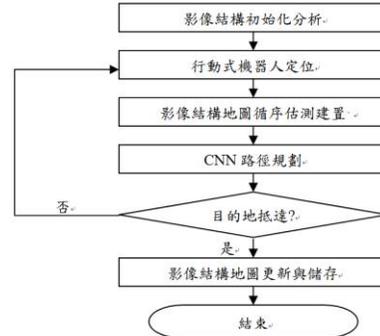
影像結構初始化分析進行初步估測環境影像中的結構，這些結構將做進一步處理並作為 CNN 的細胞；由於影像資訊對電腦而言僅是光感應變化下的資料，影像擷取下來後必須針對具有相同結構的像素以群組方式定義與識別，我們假設機器人初始位置為可移動區域(如地板)，並以機器人相鄰的區域影像作為樣本，透過 Bayes 分類器配合顏色的特徵進行識別分類，將影像區域分為可移動區域與未知區域兩類，並將未知區域設定為障礙物。在定位方法部分，機器人從預設起始點開始移動，過程中將會經過許多中途點，直到最後目的地，我們透過影像來計算機器人從起始到中途點的方向，透過電子羅盤來獲取機器人實際的移動方位，工作站透過無線通訊來指引機器人往中途點方向移動。在影像資訊分析上，我們使用移動物體偵測方法來取得機器人在影像中的位置，定位出機器人的位置。由於室內環境的光線明暗度不均一，相同的物質在影像中呈現出的色彩也會有所不同，這一點對地板影像識別會有相當程度的影響，但有一項規則是可觀察到的，就是地板影像亮度會逐漸變亮或逐漸變暗，因此當機器人移動中途點時，我們再次擷取鄰近的地板影像並做為識別樣本，使用 Bayes 分類器下一階段的可移動區域，並與初始階段的可移動區域合併，採循序估測方式進行；當機器人完整的走過所有路徑一次之後，我們即可取得精確且完整的影像結構地圖，而機器人之後的移動則只要使用該影像地圖來進行路徑規劃與移動即可，直到環境部分場景被改變，該重複上述的作法並更新該部分的地圖資訊。獲取影像結構地圖後，我們使用網格型類神經網路來進行路徑規劃，每次規劃皆以可移動區域與目的地方向作為目標，過程中假設未知區域為障礙物；移動過程中若聲波感測器回傳可移動區域為障礙物，則該機器人停止移動，並重新進行可疑區域識別與路徑規劃。

機器人定位、影像結構地圖循序估測建置、CNN 路徑規劃步驟重複進行直到機器人移動到目的地；返回原地則以先前紀錄路徑移動至初始點。以下則是上述步驟流程中所使用的方法：

### 1. 影像結構初始化分析

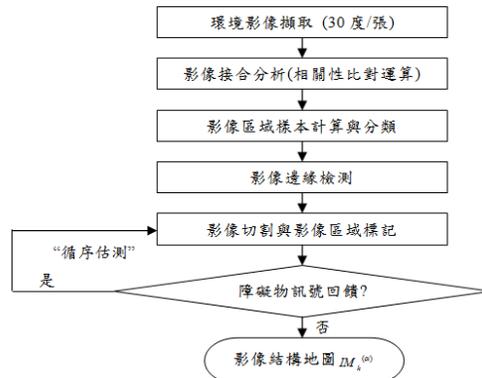
影像結構地圖建立步驟如圖二所示，在進行機器人定位之前，我們以影像結構初始化步驟來估測環境影像中的結構，為了要取得整個環境的影像，我們使用 PTZ 快速球型攝影機，以 15 度為間隔轉動進行影像擷取，若取像環境為走道則設定 180 度為攝影機旋轉角度，取像環境為家中客廳，則可設定為 360 度，以擷取足夠的環境資訊為原則；接著使用相關性比對運算方式來計算出每張影像重疊之處，並以重疊之處做為接合點，獲得環境影像後，進行影像邊緣檢測，由於邊緣線可以展現物體之間

在初始化步驟並不讀取聲波感測器回饋訊號，直到機器人移動時才分析回饋訊號，若聲波感測器讀取



圖一、方法架構流程圖

的差異，因此利用物件於影像中的邊緣線作為切割線；之後我們使用 Lifeng 等人[31]在 2009 年提出的非遞迴式的快速標記法來進行區域標記；以機器人起始位置的影像區域作為樣本，透過 Bayes 分類器進行分類，將影像區域分為可移動區域與未知區域。



圖二、影像結構地圖建立步驟

距離小於設定距離，我們則認定前方區域是為障礙物，並停止移動機器人，標記前方影像區域為障礙物，同時擷取機器人所在鄰近影像做為新樣本，以循序估測方式再次進行區域分類識別，分類的區域則不包含先前已移動影像區域。每次所估測出的結果為影像結構地圖  $IM^{(\alpha)}$  ( $\alpha$  為估測次數)，當值越大，也會越精確。接下來，我們使用相關性比對方法來分析影像連接之處，方法如下所述：

### 1. 相關性比對 [29]

相關匹配的原理是利用兩張影像函數進行乘積運算(如  $f \cdot \omega$ )，這樣的乘積運算可以是交叉相關或自相關，所乘積出來的值越高表示兩實數函數的相關性越高，基於此相關性定義，我們可從多個實數函數的相關性乘積值進一步進行比對的分析，這樣的比對分析亦稱為匹配。匹配的作法中， $f$  為包含物體或區域的一個影像，若我們想要分析  $f$  的資訊中是否有我們感興趣的特定物體或區域資訊，我們假設特定物體或區域資訊為  $\omega$ ， $\omega$  影像也稱為樣版

(Template)。若  $f$  確實包含  $\omega$  影像資訊，這兩個函數的相關會在  $f$  的某個區域相對應處找到最大值，我們以圖三來做一相關匹配方法的說明。於圖三中的  $f$  為一  $M \times N$  ( $M$ ：高度， $N$ ：寬度) 大小的影像，我們在  $f$  中找尋樣版影像  $\omega$  尺寸為  $m \times n$  的資訊，其中假設  $M \geq m$  與  $N \geq n$ ，兩實數函數進行乘積運算得到相關性  $cor$  為

$$cor(x, y) = \sum_s \sum_t f(s, t) \omega(x+s, y+t), \quad (1)$$

其中

$$x = (m-1)/2, (m-1)/2+1, (m-1)/2+2, \dots, M-(m-1)/2$$

$$y = (n-1)/2, (n-1)/2+1, (n-1)/2+2, \dots, N-(n-1)/2$$

其相關性必須經過  $f$  與  $\omega$  重疊的影像區域上進行。整個相關性匹配的計算從  $f$  影像的左上角附近的 Origin Point 開始進行， $\omega$  的原點位於它的中心， $\omega$  由 Origin Point 開始進行相關性計算直到座標  $(M-m-(1)/2, N-n-(1)/2)$  停止，因此會有  $M-(m-1) \times N-(n-1)$  個相關性  $c(x, y)$ ，然而相關性  $c$  對於  $f$  與  $\omega$  像素的變化會較為敏感。為了克服這一點，此研究採用相關係數 (Correlation Coefficient)  $\gamma(x, y)$  來進行匹配，其相關係數  $\gamma(x, y)$  的定義為

$$\gamma(x, y) = \frac{\sum_s \sum_t [\omega(s, t) - \bar{\omega}] \sum_s \sum_t [f(x+s, y+t) - \bar{f}(x+s, y+t)]}{\left\{ \sum_s \sum_t [\omega(s, t) - \bar{\omega}]^2 \sum_s \sum_t [f(x+s, y+t) - \bar{f}(x+s, y+t)]^2 \right\}^{0.5}} \quad (2)$$

其中  $x, y$  為

$$x = (m-1)/2, (m-1)/2+1, (m-1)/2+2, \dots, M-(m-1)/2$$

$$y = (n-1)/2, (n-1)/2+1, (n-1)/2+2, \dots, N-(n-1)/2$$

$\bar{\omega}$  為  $\omega$  影像區域的像素灰階平均值，而  $\bar{f}$  為  $\omega$  目前位置下重合的區域  $f$  中的像素灰階平均值，而加總是在  $f$  與  $\omega$  共同座標上進行的，而  $\gamma(x, y)$  值的範圍則落在  $(-1, 1)$  之間。一張影像結構地圖是由多個子地圖所組成，組成的方法如下：首先就以接合的方法做說明並以其中兩張圖 ( $g$  與  $f$ ) 為例，我們從  $g$  中間位置 ( $N/2$ ) 擷取樣版影像  $\omega$  ( $M \times 10$ )，計算  $\omega$  與  $f$  之間的相關係數，為了節省時間相關係數計算的範圍  $R$  僅設定為  $N/2$  寬度，並找出相關係數最大值  $\gamma_{\max}(\bar{x}, \bar{y})$ ，其  $\bar{y}$  即為  $\omega$  在  $f$  影像中的位置，而在影像  $g$  的  $0 \leq y \leq (N/2 - \bar{y})$  即為影像  $g$  於影像  $f$  中無重覆之處，這部份我們則做為影像結構地圖使用，並定義  $IM_g^{(\omega)}$ ；影像結構地圖範圍的計算，採任意前後兩張影像來進行，並且只在影像結構初始化階段計算一次，最後一張影像 (如  $IM_g^{(\omega)}$ ) 則以全部的範圍來做為影像結構地圖使用。

## 2. 影像區域樣本計算與分類

首先擷取機器人初始位置上的影像區域作為樣本，並假設該區域是可移動的平面，而影像灰階值的機率密度分佈為 Gaussian 函數，透過線性統計式分類器 Bayes 分類器來對所有的區域進行識別，Bayes 分類器是以高斯機率密度函數為基礎分類器，於此我們假設有兩類高斯密度分佈的兩個圖樣類別：一為可移動區域，另外則為未知區域 (包含障礙物與其他)，假設兩個圖樣類別的平均值分別為  $m_1$  與  $m_2$ ，而標準差為  $\sigma_1$  與  $\sigma_2$ ，Bayes 決策函數  $d$  的定義如下：

$$d_j(g) = p(g/\omega_j)P(\omega_j) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(g-m_j)^2}{2\sigma_j^2}} P(\omega_j) \quad (3)$$

其中  $p(g/\omega_j)$  是圖樣灰階值  $g$  來自  $\omega_j$  類的機率密度函數， $P(\omega_j)$  是  $\omega_j$  發生的機率。在研究中我們將使用 24-bit 的彩色影像來進行區域識別，因此，我們把灰階值  $g$  擴展到 3 維  $\mathbf{g} = [R, G, B]^T$ ，第  $j$  個的圖樣類別的圖樣向量 Gaussian 密度函數則表示為

$$p(\mathbf{g}/\omega_j) = \frac{1}{(2\pi)^{3/2} |\mathbf{C}_j|^{1/2}} e^{-\frac{1}{2}(\mathbf{g}-\mathbf{m}_j)^T \mathbf{C}_j^{-1} (\mathbf{g}-\mathbf{m}_j)} \quad (4)$$

其中  $\mathbf{m}_j$  是為樣本區域影像灰階平均值向量， $\mathbf{C}_j$  是為其他區域影像灰階值與樣本的共變異數矩陣，其定義如下

$$\mathbf{m}_j = E_j\{\mathbf{g}\}, \quad \mathbf{C}_j = E_j\{(\mathbf{g}-\mathbf{m}_j)(\mathbf{g}-\mathbf{m}_j)^T\} \quad (5)$$

其中  $E_j\{\cdot\}$  表示引數在  $\omega_j$  類圖樣上的期望值， $T$  則為轉置處理，為了方便決策函數  $d$  的計算，我們透過自然對數 (單調遞增函數) 處理來等效於指數形式的決策函數，因此我們將  $p(\mathbf{g}/\omega_j)$  乘上一自然對數，並且透過移項處理與公式的展開，我們可得到以下的決策函數：

$$d_j(\mathbf{g}) = \ln P(\omega_j) + \mathbf{g}^T \mathbf{C}_j^{-1} \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{C}_j^{-1} \mathbf{m}_j \quad (6)$$

此決策函數是為最小距離分類器的決策函數，經展開後我們可以得到以下公式

$$d_j(\mathbf{g}) = a_1 R + a_2 G + a_3 B + a_4 \quad (7)$$

由於我們的方法只取可移動區域一類的樣本，因此我們必須設定決策函數數值範圍  $\lambda$ ，若  $d_j(\mathbf{g}) \leq \lambda$ ，則該區域與樣本同一類型；

## 3. 影像邊緣檢測

由於影像數據對電腦而言僅是光感應變化下的資料，影像擷取下來後必須針對具有相同結構的像素以群組 (區域) 方式標記並賦予意義，數據才能成為有用的資訊；為了擷取能代表環境特徵的物件影像區域，我們以物件於影像中呈現出來的邊緣線作為切割線，因此我們需要使用高通濾波器並配合一些影像處理演算法來檢測出邊緣，所需的演算法與步驟：色彩模式

轉換→Gaussian模糊化處理→Sobel濾波邊緣檢測→影像二值化處理→中值濾波→細線化處理。

### I. 色彩模式轉換：

首先將24-bit彩色影像轉換至8-bit的灰階影像可降低演算的複雜度，轉換公式可由下式來獲得(Gonzalez, et al., 2002) (其中  $R$ 、 $G$ 、 $B$  分別為彩色影像中三成份的灰階值，而  $f$  為轉換為8 Bits影像的灰階值)。

$$f = 0.299R + 0.587G + 0.114B \quad (8)$$

### II. Gaussian模糊化處理：

接著，為了讓邊緣資訊的效果提昇，必須降低影像雜點資訊，才能保留物件清楚的邊緣線，所以透過Gaussian濾波模糊化處理( $f_G$ )來降低影像雜點的干擾，二維的Gaussian函數如下(Gonzalez, et al., 2002)：

$$f_G = f * e^{-\frac{x^2+y^2}{\sigma}} \quad (9)$$

其中  $x$  與  $y$  為影像水平軸與垂直軸座標符號， $\sigma$  為標準差，符號  $*$  為旋積運算(Convolution Operation)處理。

### III. Sobel濾波邊緣檢測：

邊緣檢測演算法有幾種方式可以實現，如Roberts梯度運算子、Prewitt運算子、Sobel運算子、Laplacian運算子等等，這些方法都可以獲得影像邊緣資訊，而效果會隨著不同的應用而有所不同，在此我們選擇能抑制雜訊特性的Sobel運算子來尋找影像中的邊緣。Sobel運算子可由以下梯度公式  $\nabla f_G$  來獲得(Gonzalez, et al., 2002)

$$\begin{aligned} \nabla f_G &= \text{mag}(\nabla f_G) = [D_x^2 + D_y^2]^{0.5}, \\ \nabla f_G &= [D_x \quad D_y]^T = \begin{bmatrix} \frac{\partial f_G}{\partial x} & \frac{\partial f_G}{\partial y} \end{bmatrix}^T, \\ D_x &= (s_7 + 2s_8 + s_9) - (s_1 + 2s_2 + s_3), \\ D_y &= (s_3 + 2s_6 + s_9) - (s_1 + 2s_4 + s_7). \end{aligned} \quad (10)$$

取得影像的邊緣資訊  $\nabla f_G$  後，接著進行二值化處理，將邊緣強度較高的部份擷取出來做為區塊邊界，讓影像內部多個能代表特殊屬性的區域分割出來。

### IV. 影像二值化處理：

二值化的處理重要關鍵在於灰階閾值，於此我們使用 Otsu 學者所提出(Otsu, 1978)的統計式閾值決定法，此方法的概念為假設有一最佳閾值  $\eta$ ， $\eta$  可將影像中的像素分為兩區( $O_1$  與  $O_2$ )，其  $\eta$  的決定必須滿足兩個條件其中之一即可。

條件一：使得  $O_1$  與  $O_2$  之間的變異數最大。

條件二：使得  $O_1$  與  $O_2$  內在的變異數總和最小。

$$\beta = \begin{cases} 255, & \text{if } \nabla f_G > \eta \\ 0, & \text{if } \nabla f_G \leq \eta \end{cases} \quad (11)$$

此方法詳細說明可參考Otsu學者於1978的文獻，定義符號  $\beta$  為二值化後的影像。

### V. 中值濾波：

使用中值濾波是為了要去除二值化後，邊緣影像上的一些小雜點，本研究使用Yang 與Toh所提出的適應性模糊多層中值濾波器(Adaptive Fuzzy Multilevel Median Filter/AFMMF)來去除影像中的雜訊，此方法是結合模糊關聯記憶法(Fuzzy Associative Memory)，所提出的修正式的多層中值濾波器，該濾波器克服了傳統多層中值濾波器(A Multilevel Median Filter/MLMF)的一些缺點，此方法於保留影像內容物的邊緣資訊能有更好的效果。

### VI. 細線化處理：

透過細線化處理將邊緣線的寬度最小化為一個像素的寬度，如此一來，這樣的格線就可以作為網格的邊界，接著再透過切割與標記處理，每個區域網格之間也有相對應的交互關係，這樣一來我們就可以使用網格型類神經網路於具網格型態的影像結構地圖上。細線化的定義是假設有兩個於集合  $A$  與  $B$ ， $A$  被一個結構元素  $B$  細線化，可表示成使用  $A \otimes B$ ，我們可以使用一個結構元素序列來定義細線化(Gonzalez, et al., 2002)：

$$A \otimes \{B\} = ((\dots(A \otimes B^1) \otimes B^2 \dots) \otimes B^n) \quad (12)$$

### 4. 影像切割與影像區域標記

取得細線化後的邊緣線作為影像區域切割的基準，定義邊緣線為1，其餘部分為0，在邊緣線之間經常存在著少許的縫隙，為了讓邊緣線更完整，我們使用型態處理方法(Morphology)中的膨脹運算(Gonzalez, et al., 2002)來進行填補的動作。膨脹的定義為假設有兩個於影像函數  $Z^2$  中像素的集合  $A$  與  $B$ ，使用  $B$  對  $A$  來進行膨脹處理，是為  $z$  所有位移(Displacement)的集合，這些位移可以使得  $\hat{B}$  和  $A$  之間至少有一個元素的重疊，將處理結果標記為  $A \oplus B$ ，公式的定義如下：

$$\bar{\beta} = A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\}. \quad (13)$$

每個點  $z \in B$ ，這些點對  $A$  來進行平移，並將全部的平移結果取聯集(OR)處理。我們以邊緣線為切割線  $\bar{\beta}$ ，其餘部分則為切割後的影像區域  $\tilde{\beta} = 1 - \bar{\beta}$ ， $\tilde{\beta}$  影像區域標記處理，是使用 Lifeng 等人所提出的非遞迴式(Non-recursive)的快速標記法來進行區域標記，他們的方法是一種 Two-scan 標記演算法，演算效率比傳統遞迴式標記法、Multi-scan 標記演算法、Hybrid 標記演算與 Tracing-type 標記演算來的更快速，非常適合使用在影像結構地圖即時定位

上，在他們的方法上，使用 8 連通的遮罩來分析像素的鄰近位置，並且定義了 16 種像素間可能的連結型態，配合 Karnaugh 地圖來決定該像素歸屬的標籤我們把先前影像區域識別的結果與標記後的區域用聯集方式結合起來，即是我們需要的影像結構地圖。

### 5. 移動物體偵測

我們透過影像處理的方式來進行行動式機器人位置的偵測，並且設計動作識別行為能夠在多移動物體出現時進行辨識；為了能夠偵測到機器人在畫面中的位置，我們使用先前研究(Chiu, Lu and Wen, 2006[30])所提出的移動物體偵測方法來進行機器人移動偵測，該方法的優點是能自動隨著室內光線變化來學習背景資訊，不需事先定義背景資訊，並且可抗光干擾的能力較佳。該方法定義了三個移動模式，分別是當前移動模式(Current Motion Model)  $\tilde{\Psi}_c(t)$ 、背景移動模式(Background Motion Model)  $\tilde{\Psi}_b(t)$ 、主要移動模式(Main Motion Model)  $\tilde{\Psi}_m(t)$ ，定義  $\Psi(t)$  為“邊緣影像”與“灰階影像”的組合，如下式所示：

$$\Psi(t) = w_1 \text{edge}(f_t) + w_2 \text{gray}(f_t)$$

$$\Psi(t) = w_1 \text{edge}(f_t) + w_2 \text{gray}(f_t) \quad (14)$$

其中  $w_1$  與  $w_2$  是為權重係數並且符合  $w_1 + w_2 = 1$  的條件， $f_t$  是於  $t$  時間下的影像畫面， $\text{edge}(\cdot)$  為 Sobel 邊緣檢測， $\text{gray}(\cdot)$  為  $f_t$  灰階影像。物體移動偵測藉由兩個規則來進行判斷，其規則如下所示：

規則1:

$$\tilde{\Psi}_c(t) > (\tilde{\Psi}_m(t) + \Psi_T) \text{ 或 } \tilde{\Psi}_c(t) < (\tilde{\Psi}_m(t) - \Psi_T) \quad (15)$$

規則2:

$$\tilde{\Psi}_c(t) > (\tilde{\Psi}_b(t) + \Psi_T) \text{ 或 } \tilde{\Psi}_c(t) < (\tilde{\Psi}_b(t) - \Psi_T) \quad (16)$$

其中  $\Psi_T$  為敏感參數，當  $\tilde{\Psi}_c(t)$  符合這兩項規則時，則表示有人員移動的情況。

### 6. Cellular Neural Networks 路徑規劃

網格式類神經網路 CNN [31] 是一種類比、非線性、動態式與多維度的迴路，同時具有區域週期性的拓撲結構，基本的迴路單位稱為細胞或人工神經元，而這些細胞僅與鄰近的細胞單元連結，基本的網格式類神經網路有二維的方陣架構，其組成是由一群同性質以非線性迴路方式排列的細胞，而這些細胞都內含非線性的動態行為。由於神經元之間是為區域性連結，這樣的場域可以由許多的晶片透過電線來組成實現，並且實現在現今的超大型積體電路技術上；CNN 可以透過一組非線性的微分方程來表示，也可將其看成迴路單元，單一細胞  $\mathcal{G}(I, J)$  的狀態方程式的數學模式可以表示如下：

$$\dot{X} = \frac{dX_{IJ}}{dt} = -X_{IJ} + \sum_{\mathcal{G}_{KL} \in \Omega_r} A_{IJ, KL} Y_{KL} + \sum_{\mathcal{G}_{KL} \in \Omega_r} B_{IJ, KL} U_{KL} + Z_{IJ} \quad (17)$$

其中  $X_{IJ}$  定義了細胞  $\mathcal{G}_{IJ}$  的狀態， $Y_{KL}$  與  $U_{KL}$  個別定義了細胞  $\mathcal{G}_{IJ}$  的輸出與輸入，與以細胞  $\mathcal{G}_{IJ}$  為中心的球狀影響範圍半徑  $r$  的影響力  $\Omega_r$  ( $\mathcal{G}_{KL} \in \Omega_r$ )； $A(IJ, KL)$  與  $B(IJ, KL)$  則分別是回饋與控制樣板， $Z_{IJ}$  則是位移項(Bias Term)。另外，表示細胞  $\mathcal{G}_{IJ}$  的輸出值於下述的關係式中：

$$Y_{IJ} = F(X_{IJ}) = \frac{1}{2} [|X_{IJ} + 1| - |X_{IJ} - 1|] \quad (18)$$

其中  $Y_{IJ}$  則定義是  $\mathcal{G}_{IJ}$  的輸出值。

### 結論與討論

本研究主要進行一項以電腦視覺對行動式機器人進行視覺定位與引導之處理方法找尋，此系列研究主要是應用於遠距式居家照護系統之照護機器人的定位與引導上，如何以全方向的二維影像來引導三維空間的機器人是我們此系列研究目標；在未來的工作上，我們將進一步展現本研究方法的性能。

### 致謝

本研究是由 NSC 國家科學委員會的支持，99 年度計畫編號：NSC 99-2221-E-276-005。

### 參考文獻

- [1] 行政院經建會, (2008). 「中華民國臺灣97年至145年人口推計」, <http://www.cepd.gov.tw/>.
- [2] 行政院衛生署, (2008). 「遠距照護試辦計畫」, <http://doh.telecare.com.tw/>.
- [3] 康妥舒, <http://www.tunstall.com.tw>.
- [4] ITIS智網-產業評析, <http://www.itis.org.tw/>
- [5] Brender, Jytte, (1998). Trends in Assessment of IT-based Solutions in Healthcare and Recommendations for the Future. *International Journal of Medical Informatics*, 52(1): 217-227.
- [6] Schelkens, Peter, Munteanu, Adrian, Cornelis, Jan. (1999). Wavelet-based Compression of Medical Images: Protocols to Improve Resolution and Quality Scalability and Region-of-interest Coding. *Future Generation Computer Systems*, 15(2): 171-184.
- [7] Caldelli, Roberto, Filippini, Francesco, Barni, Mauro, (2006) Joint Near-lossless Compression and Watermarking of Still Images for Authentication and Tamper Localization. *Signal Processing: Image Communication*, 21(10): 890-903.
- [8] Singh, Sukhwinder, Kumar, Vinod, Verma, H.K. (2007). Adaptive Threshold-based Block

- Classification in Medical Image Compression for Teleradiology. *Computers in Biology and Medicine*, 37(6): 811-819.
- [9] 宋開泰, (2008). Rola 家用型機器人, <http://isci.cn.nctu.edu.tw/index.html>.
- [10] Chia-How Lin, Su-Hen Yang, Hong-Tze Chen and Kai-Tai Song, (2008). Mobile Robot Intruder Detection Based on a Zigbee Sensor Network, in *Proc. of 2008 IEEE International Conference on Systems, Man and Cybernetics (SMC 2008)*, Singapore, Oct. 12-15: 2786-2791.
- [11] Fu-Sheng Huang and Kai-Tai Song, (2008). Vision SLAM Using Omni-Directional Visual Scan Matching, in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IEEE/RSJ IROS 2008)*, Nice, France, Sep. 22-26:1588-1593.
- [12] Meng-Ju Han, Jing-Huai Hsu, Kai-Tai Song and Fuh-Yu Chang, (2007). A New Information Fusion Method for SVM-Based Robotic Audio-Visual Emotion Recognition, in *Proc. of IEEE International Conference on Systems, Man and Cybernetics*, Montreal, Canada, pp. 2656-2661.
- [13] M. L. Wang, Y. C. Chung, and H. Y. Lin, (2009). A Self-Localization Technique for Mobile Robots using Image-Based Ground Plane Detection, *Proceedings of the 2009 International Conference on Service and Interactive Robotics*
- [14] S.-K. Cho, T. Yang, M.-G. Choi; J.-M. Lee, (2009). Localization of a High-Speed Mobile Robot Using Global Features, *ICARA 2009. 4th International Conference on Autonomous Robots and Agents*, pp. 138-142.
- [15] S.-H. Kim, C.-W. Roh, S.-C. Kang, M.-Y. Park, (2007). Outdoor Navigation of a Mobile Robot Using Differential GPS and Curb Detection, *2007 IEEE International Conference on Robotics and Automation*, pp. 3414-3419.
- [16] S. S. Han, D. K. Kim, J. M. Lee, (2007). A New Tag Arrangement Pattern for a Differential Driving Mobile Robot Based on RFID System, *ICCAS '07-International Conference on Control, Automation and Systems*, pp 1228-1233.
- [17] I.J. Cox, J.J. Leonard, (1994). Modeling a Dynamic Environment Using a Bayesian Multiple Hypothesis Approach, *Artificial Intelligence* 66 (1) : 311-344.
- [18] A. Elfes, (1989). Occupancy Grids: A Probabilistic Framework for Robot Perception and Navigation, Ph.D. Dissertation, *Department of Electrical and Computer Engineering, Carnegie Mellon University*.
- [19] F. Lu, E. Millos, (1997). Robot Pose Estimation in Unknown Environment by Matching 2D Range Scans, *Journal of Intelligent and Robotics Systems* 18 (3):249-275.
- [20] H. Choset, K. Nagatani, (2001). Topological Simultaneous Localization and Mapping (SLAM): Toward Exact Localization Without Explicit Localization, *IEEE Transactions on Robotics and Automation* 17 (2): 125-137.
- [21] D. Kortenkamp, T. Weymouth, (1994). Topological Mapping for Mobile Robots Using a Combination of Sonar and Vision Sensing, in: *Proceedings of the 12th National Conference on Artificial Intelligence*, pp. 979-984.
- [22] H. Frezza-Buet, F. Alexandre, (2002). From a Biological to a Computational Model for the Autonomous Behavior of an Animation, *Information Sciences* 144 (1-4): 1-43.
- [23] M.E. Jefferies, W.-K. Yeap (Eds.), *Robotics and Cognitive Approaches to Spatial Mapping, Springer Tracts in Advanced Robotics*, vol. 38, Springer, 2008.
- [24] B. Kuipers, The cognitive map: could it have been any other way?, in: H.L. Pick, L.P. Acredolo (Eds.), *Spatial Orientation: Theory Research and Application, Plenum Press*, New York, 1983.
- [25] M. P. Paulraj, R.B. Ahmad, C.R. Hema, F. Hashim, S. Yusoff, (2008). Active Stereo Vision Based System for Estimation of Mobile Robot Orientation Using Affine Moment Invariants, *ICED 2008 International Conference on Electronic Design*, pp 1-7.
- [26] I. Gavrilut, A. Gacsadi, C. Grava, V. m. Tiponut, , (2006). Vision Based Algorithm for Path Planning of a Mobile Robot by Using Cellular Neural Networks, *Automation, Quality and Testing, Robotics, 2006 IEEE International Conference on Vol. 2, 25-28:306-311*.
- [27] I. Gavrilut, V. Tiponut, A. Gacsadi, (2006). Path Planning of Mobile Robots by Using Cellular Neural Networks, *Cellular Neural Networks and Their Applications*, 2006. CNNA '06. 10th International Workshop on, 28-30:1 - 6
- [28] L. O. Chua, and L. Yang, (1988). Cellular Neural Networks: Theory and Application, *IEEE Trans. On Circuits and Systems, (CAS)*, Vol. 35:1257-1290.
- [29] Gonzalez, Rafael C., Woods, Richard E. 2002. *Digital Image Processing*. Second Edition, Prentice-Hall.
- [30] Chiu, Shih-Hsuan, Lu, Chuan-Pin, Wen, Che-Yen, (2006). A Motion Detection Based Framework for Improving Image Quality of CCTV Security Systems. *Journal of Forensic Sciences*, 51(5): 1115-1119.
- [31] B. Siemiatkowska, A. Dubrawski, (1998). Cellular Neural Networks for Navigation of a Mobile Robot, *Springer-Verlag Berlin Heidelberg*, pp. 147-154.