

行政院國家科學委員會專題研究計畫 成果報告

人性化電腦整合智慧型環境(3/3) 研究成果報告(完整版)

計畫類別：個別型
計畫編號：NSC 95-2221-E-002-235-
執行期間：95年08月01日至96年07月31日
執行單位：國立臺灣大學資訊工程學系暨研究所

計畫主持人：傅立成

計畫參與人員：博士班研究生-兼任助理：廖峻鋒、陸敬互、詹益銘

處理方式：本計畫可公開查詢

中 華 民 國 96 年 09 月 21 日

行政院國家科學委員會補助專題研究計畫成果報告

人性化電腦整合智慧型環境

計畫類別：☒ 個別型計畫 ☐ 整合型計畫

計畫編號：NSC95-2221-E-002-235

執行期間：95 年 08 月 01 日至 96 年 07 月 31 日

計畫主持人：傅立成 教授

共同主持人：

計畫參與人員：廖峻鋒、陸敬互、詹益銘

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

本成果報告包括以下應繳交之附件：

- ☐ 赴國外出差或研習心得報告一份
- ☐ 赴大陸地區出差或研習心得報告一份
- ☐ 出席國際學術會議心得報告及發表之論文各一份
- ☐ 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、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

詢

執行單位：國立臺灣大學資訊工程學系暨研究所

中華民國：96 年 7 月 31 日

中文摘要

為了讓電腦藉由各項數位服務提供人類更好的生活品質，本計畫將研究如何建構一個具備多重感測能力及多重互動方式的智慧型環境作為提供數位服務的平台，並在平台上發展一個具備高度整合能力的系統，此即導致建構一「人性化電腦整合智慧型環境」的終極目標。

為達此目標，須先達成數項技術指標，首先，以電腦視覺為基礎之人形偵測系統，係涵蓋所有智慧型環境空間，以達到多目標無間隙追蹤(Seamless Tracking)與定位(Localization)為其目的；於此部分透過貝氏網路整合背景模型(Background Modeling)與運動資訊(Motion Information)，有效在不受光線變化影響下，成功將影像前景切割出來。接著透過階層式人形資料庫的架構，經由 Hausdorff 樣板比對的技術，智慧型環境中的人可以被快速且成功偵測出來，最後，我們採用貝氏濾波技術(Bayesian Filtering)為基礎的多目標物追蹤技術，並且在量測影像相似度上考慮接合影像相似度(Joint Image Likelihood)，作為解決目標在影像觀測上重疊時的判斷，以達成所謂多目標物之追蹤。

其次，對於空間中的人體姿態的視覺辨識，係提供使用者友善的「人機互動介面」與「行為分析」兩課題裡的重要一環，即輔助電腦對使用者行為作適當的理解與描述另一方面，透過 RJMCMC 的搜尋技術以及資料驅動的方式，有效率達到估算人體姿態，做為分析與理解使用者於智慧型環境之行為模式。另一方面，為求姿態估算之精準度、效能以及減少遮蔽現象之影響，我們更進一步整合多台相機，透過整合運動資訊以及多樣板的技術來達成姿態估算，以提供系統做為提供服務的重要依據。首先，藉著比較從兩台相機中擷取之的剪影(Silhouette)和資料庫的範例(Examples)來做靜態辨識，資料庫儲存多樣化的人物動作組態和相對於相機的 2D 人物外型，比對範例並選取最相似的範例作為答案候選人。接著，透過推算出運動資訊，然後這些運動資訊被用來推算哪一個結果比較好。

再者，為了達到人性化之目的，本計畫亦將建立有效的多種感測資料收集系統、使用者喜好模型、即時運算使用者目前狀況並事先提供服務、多人環境偵測、動態調整使用者喜好模型，經由所提出了三階層的模型(Model)來達成推論出應該提供給整個群體的服務，而其中第一層負責處理從感測器傳來的原始資料，第二層負責每個居住者個體的喜好模式，最高的第三層負責處理多位使用者之間的關係。最後，除了建置智慧型環境、取得環境資訊之外，我們針對上述問題探討智慧型環境中人機互動服務的機制，並提出一個整合數位服務的平台，讓系統除了能夠觀察環境中的資訊，適應使用者的行為，還可以依據不同的需求及特性來組織系統的資訊呈現系統的狀態，針對環境資源的變化動態調整自己的架構，進而將最佳的互動服務提供給使用者，在此我們依據 OSGi 架構設計出一套以服務為導向(Service-Oriented)之系統來達成上述的功能。

關鍵詞：無間隙追蹤、多感測器、背景模型，運動資訊，RJMCMC，姿態辨識、服務客制化、智慧型環境、智慧型空間。

英文摘要

The purpose of technology is to upgrade the quality of human life. In recent years, the advancement of computer technology brings computer into human daily activities, and makes computer an essential part of modern life. Digital services are also more and more popular. The appearance of the embedded system, information appliances, and home control network has brought various kinds of computer applications into human life.

For this purpose to provide better life quality through computer integrated digital services, we propose a multi-sensing multimodal intelligent environment as the platform to provide digital services. One important task of the vision system in e-space application is to keep track of the trajectory and position of the human for further processing. Under this consideration, we aimed at developing a multi-person tracking system to detect and track indoor persons seamlessly. First, a framework integrating background model and motion information is used to perform the foreground segmentation and motion estimation simultaneously. Then, a hierarchical template database is utilized to detect the human in the obtained foreground mask. Finally, an approach based on Bayesian filtering is to track multiple targets with occlusion effect. Human pose estimation is an important topic for providing friendly user interface. Here, the RJMCMC mechanism with data-driven algorithm is utilized to search the solution efficiently and effectively. In order to make the system robust to different view of angle, we propose an approach using multiple cameras with templates to achieve effective and efficient pose estimation.

To achieve context-aware, a multi-agent architecture is proposed here to robustly perform many tasks and integrate many sensors including camera, light, and pressure floor sensors in intelligent space. Furthermore, a Bayesian network is used to infer the user preference by gathering the information from these sensors to provide more suitable and comfortable services. Here, we develop a system to learn a multiple users' preference model that represents relationships among users as well as dependency between services and sensor observations. Thus, the service can be inferred based on the learnt model. To achieve this, we propose a three-layer model in our work. At the first layer, raw data from sensors are interpreted as context information after noise removal. The second layer is dynamic Bayesian networks which model the observation sequences including inhabitants' location and electrical appliance (EA) information. At the highest layer, we integrate second layer's environment information and the relations between inhabitants to provide the service to inhabitants. Therefore, the system can infer appropriate services to inhabitants at right time and right place, and let them feel comfortable. In experiments, we show our model can provide reliable and appreciate services to inhabitants in a smart home environment.

To achieve context-aware, a multi-agent architecture is proposed here to robustly perform many tasks and integrate many sensors including camera, light, and pressure floor sensors in intelligent space. The architecture of a conventional smart home is usually server-centric and thus

causes many problems. Mobile devices and dynamic services cause the dynamically changing environment, which can result in very difficult interaction. In addition, how to provide services efficiently and appropriately is always an important issue for a smart home. To solve the problems caused by traditional architectures, to deal with the dynamic environment, and to provide appropriate service, we propose a service-oriented architecture for smart home environments, based on OSGi and mobile agent technology. This architecture is a Peer-to-Peer model based on multiple OSGi platforms, in which service-oriented mechanisms are used for system components to interact with one another, and mobile agent technology is applied to augment the interaction mechanisms.

Keywords: Tracking, Bayesian Filtering, RJMCMC, Posture Estimation, Intelligent Environment, Bayesian Network, OSGi.

目 錄

<u>摘要</u>	II
<u>第一章、計畫緣由</u>	1
<u>第一節 研究背景</u>	1
<u>第二節 研究目的</u>	7
<u>第二章、研究方法與執行成果</u>	10
<u>第一節 無間隙人形追蹤系統</u>	10
<u>第二節 人體姿態辨識</u>	14
<u>第三節 使用者喜好與多感測器之客制化服務</u>	19
<u>第四節 智慧型環境建制與整合</u>	22
<u>第三章、結果與討論</u>	24
<u>第四章、計畫成果自評</u>	25
<u>參考文獻</u>	26

第一章、計畫緣由

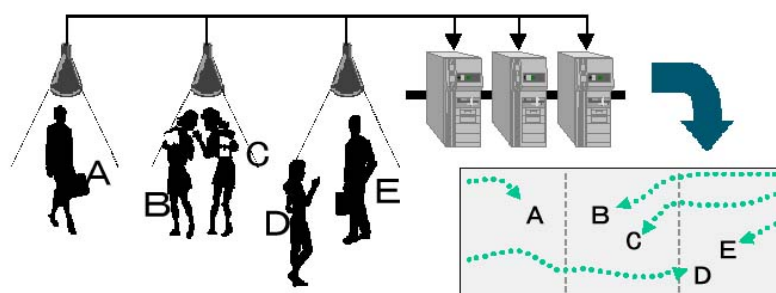
「人性化電腦整合智慧型環境」是當前人類利用電腦整合目前先進光、電、通訊、資訊等科技以提昇所處環境的智慧性之主要研究課題，不僅為現今研究主流，預期成果對人類生活的影響亦相當的大，本計劃在研究如何建構一個具備多重感測能力及多重戶動方式之智慧型環境，以作為數位服務提供平台。而隨著電腦影像技術之進步，我們期望透過電腦視覺達到涵蓋智慧型環境空間之人形偵測與追蹤，以達到多目標無間隙追蹤與定位，並進一步空間中之人形進行姿態的視覺辨識，提供使用者友善之人機互動介面及行為分析。

再者，為達到人性化之目的，本計劃亦須建立有效的多種感測資料收集系統、使用者喜好模型、即時運算使用者目前狀況並事先提供服務、多人環境偵測、動態調整使用者喜好模型；最後，除了建置智慧型環境、取得環境資訊之外，我們針對上述問題探討智慧型環境中人機互動服務的機制，並提出一個整合數位服務的平台。綜言之，本計畫主要針對以下四項重點子題之相關背景及目的進一步詳細說明，此四個子題主要包含：(1)無間隙人形追蹤系統；(2)人體姿態辨識；(3) 使用者喜好與多感測器之客制化服務；(4) 智慧型環境建制與整合。

第一節、研究背景

(一) 無間隙人形追蹤系統

無間隙追蹤系統可同時整合了多個相機的觀測範圍，形成一組在廣大區域的目標偵測系統，系統擴大的偵測範圍取決於同時設置並啟用的相機個數，相機個數越多可以創造出越大的偵測範圍。無間隙追蹤系統的好處在於，應用此種影像系統在人性化電腦整合智慧型環境中，可使系統偵測到環境中人形目標物的確切位置、姿態甚至運動軌跡，另外當目標進行大範圍的運動，即使由一個相機的視野切換至另一個相機的視野當中，該系統仍然可以透過辨識目標物特徵的機制，判定相鄰兩相機視野中的目標為同一個目標物，一般而言，廣大區域的目標追蹤會有多個目標物需同時追蹤，無可避免的，更由於目標物間彼此會有交互重疊的情形，因此問題的複雜度比單一目標物的追蹤要來的困難，需要發展不同於單一目標追蹤的演算法方能達此目的。影像系統與其他感測系統最大的差異在於，能夠完整偵測人形在環境中之位置與行為姿態，在多目標無間隙人形追蹤系統中，我們預期利用多個影像設備，以涵蓋所有智慧型環境空間，並進一步有效偵測與追蹤多個人形在環境中之位置與運動軌跡，整個系統架構如圖一所示。

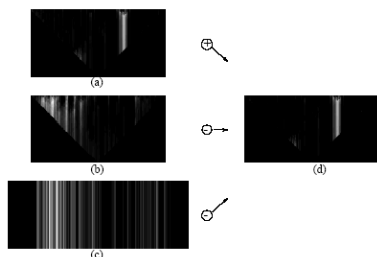


圖一：多相機無間隙追蹤系統示意圖

多相機的影像系統中，最大的挑戰在於如何找出各相機影像之相對應關係，以重建出目標物在運動空間中的影像。常用的技術大致上可分為兩類，一為傳統之校正技術，主要是計算出相機之間的轉換關係[1]，這種技術最大的缺點在於校正的過程十分繁雜。另一類的技術主要是利用最佳化的方式，找出兩個影像之間的對應關係，這類的方法，主要被應用在全景影像的重建技術上[2]。一般而言，用於辨識人形之電腦影像技術依據所採用的影像資訊不同，可區為以外觀為主之人形偵測(Appearance-Based Approach)以及以運動為主之人形偵測(Motion-Based Approach)。

1.外觀為主之人形偵測

以外觀為主之人形偵測主要是利用人在單一影像中的外觀特徵，以偵測影像中人形的存在，此類人形偵測技術主要透過一些影像特徵點(Low-Level Features)來表示人之可能外觀，經由標準圖形辨識(Pattern Recognition)的技術來達到人形偵測目的。此方法中，最直接的方式即透過搜尋影像中人形特徵點，Broggi *et al.* [3] 利用人形在影像中所呈現之垂直邊對稱性，來偵測影像中人形的位置與大小，圖二為所提出之垂直邊對稱性之計算過程。Hayfron [4] 則是偵測空間與時間軸上的對稱性(Spatio-Temporal Symmetry)(如圖三所示)來偵測人形的存在。Wu 和 Yu [5] 提出一個結合 Boltzman 模型與 Markov 模型的二階統計模型(Two-Layer Statistical Field Model)，可對人體不規則的外型作特徵的描述，由於引入機率分佈去衡量比較的相似性，當人體某部份重疊或遮蔽時，仍然可以成功地偵測到人體，除外，二階統計模型可以對影像中所觀察到的資料做更有彈性的描述。



圖二：垂直邊對稱性計算流程圖

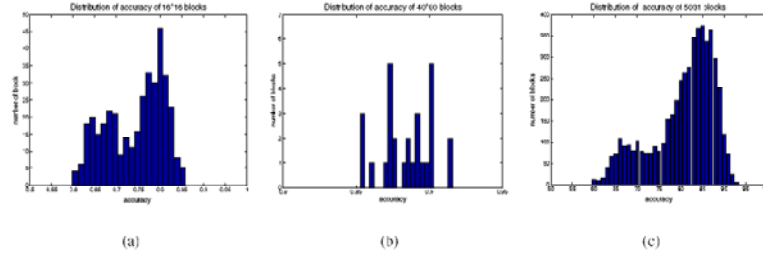


圖三：空間與時間軸上的對稱性示意圖

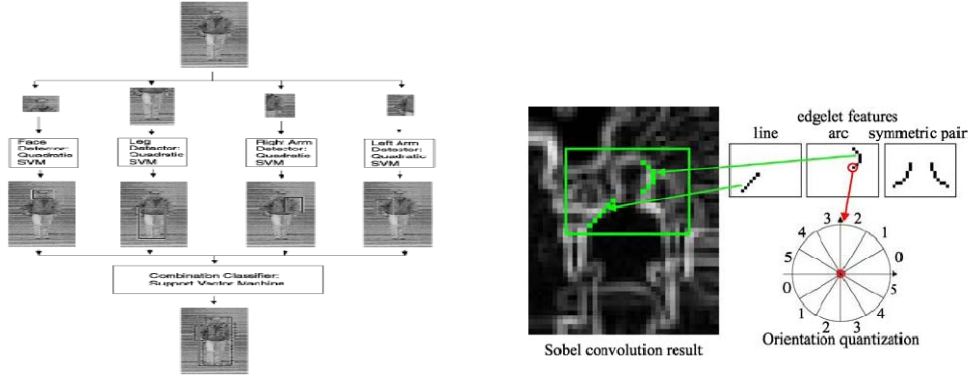
另一種方式則是將人形偵測的問題，轉換成樣板比對的問題，首先建構不同角度以及姿勢的人形，接著透過樣板比對的方式來偵測人形存在，在人形外觀特徵的表示上，Gavrila *et al.*[6][7] 以及 Liu *et al.* [8] 利用人體的輪廓(Silhouette)或邊緣影像(Edge Image)來表示人形，人形樣板皆被轉換成 DT(Distance Transform)影像。Oren [9]為了更有效克服物體位移(Translation)、比例(Scale)與旋轉(Orientation)變化，採用 Harr Vertical 與 Horizontal wavelets 計算出微波係數(Wavelet Coefficients)的人形特徵圖(如圖四所示)。於[10][11]研究中，旋轉強度之統計長條圖(Histogram of Oriented Gradients)(如圖五所示)被用來表示人形之特徵，透過 SVM(Supported Vector Machine)機器學習的方式，所得之分類器(Classifier)可有效的表示此類特徵，並作為影像中之人形偵測。



圖四：微波係數特徵圖



圖五：旋轉強度之統計長條圖



圖六:Mohan *et al.* 人形偵測架構圖

圖七:Edgelet 特徵點

然而上述之技術，主要想法是將人於影像中所呈現外觀視為一完整物體(Object)，經由比對的方式來達到人形偵測的目的。但這種方式的主要缺點在，其無法克服人形與影像中彼此遮蔽(Occlusion)的問題，因此目前有許多研究提出基於透過辨識人體各個部份肢體(Body Part)的方式，最後根據人體幾何結構性(Geometric Property)或其他特性，來達到人形辨識的目的。Mohan A. *et al.* [12] 提出了一個階層式架構分類器，Adaptive Combination of Classifiers(ACC)，首先在影像中偵測身體各部位，最後再用一個 SVM 結合各個部位以辨識出人形，圖六為其人形偵測之架構圖。D. Ramanan *et al.* [13] 提出一個以人體各部位為基礎的姿態模型，再針對各種人體姿態在影像中比對是否有符合此人體姿態的人形，再藉由偵測每一張影像中的模型做到追蹤人形的功能。Wu 與 Nevatia [14]使用所謂的 Edgelet(如圖七所示)去表示人形部分外觀，在偵測每個人體部分後，採用貝氏機率架構結合各分類器所偵測的結果。Leibe *et al.* [15] 則採用一個 Implicit Shape Model (ISM)去描述人形不同肢體與人體中心點的關係，透過投票的演算法來偵測人形。

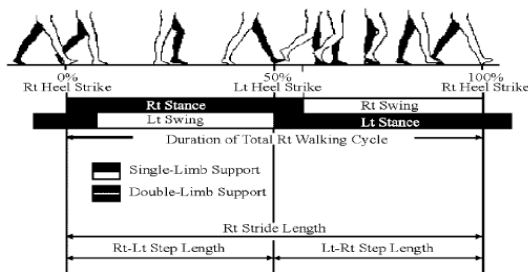
2. 以運動為主之人形偵測

第二類為以運動為主之人形偵測，其主要運用原理是根據觀察人於影像中通常會造成影像變化。此類技術最直接的方式是去學習人形運動所造成之影像圖形(Pattern)，Viola *et al.* [16] 採用連續兩張影像之間上下左右位移影像差，作為人移動時之特徵，其計算方式如下方程式所示，而學習的過程則採用 Adaboost。

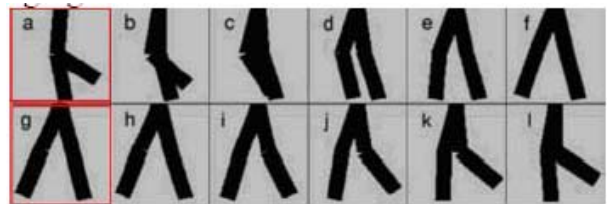
$$\begin{aligned} \Delta &= \text{abs}(I_t - I_{t+1}) & U &= \text{abs}(I_t - I_{t+1} \uparrow) \\ L &= \text{abs}(I_t - I_{t+1} \leftarrow) & R &= \text{abs}(I_t - I_{t+1} \rightarrow) \\ D &= \text{abs}(I_t - I_{t+1} \downarrow) \end{aligned}$$

其中 I_t 與 I_{t+1} 分別表示於時間 t 與 $t+1$ 連續兩張影像，而符號 $\uparrow\downarrow\rightarrow\leftarrow$ 則表示往上下左右四各方向位移的運算元。Little 與 Boyd [17] 則是經由計算連續兩張影像的光流(Optical Flow)，作為人運動時之特徵點，再經由運動週期性的分析來確定影像中人形之存在。於[18]所提出之方法中，則透過區域追蹤(Region Tracking)的方式，來找出不同時間點之對應區域，並採用時間延遲類神經網路(Time-Delay Neural Network)，來偵測所觀察區域之寬度是否呈現週期變化，進而達到人形偵測的目的。

根據影像觀察得知，人於運動過程中的步伐(Gait)有很明顯的特徵，也因此於目前研究中被廣泛應用來作為人之存在偵測。Cunado *et al.* [19] 對於步伐有詳細的描述與定義(如圖八所示)，並提出一個鐘擺的模型(Pendulum Model)來描述人行走的過程。Niyogi 與 Adelson [20] 計算人行輪廓於 XYT 軸上的變化，進而分析出人行走時의 步伐圖形，以作為步伐的偵測。Y. Ran *et al.* [21][22] 根據所提出雙鐘擺模型(Twin-Pendulum Model)(如圖九所示)去表示行人步伐，接著透過影像處理的技術，找尋雙腳分開角度極大與極小的圖形，分別對應至圖九中標示紅色之兩各步伐圖示，作為分析週期之特徵。



圖八：行走步伐描述



圖九：雙鐘擺模型

(二) 人體姿態辨識

人體姿態的視覺辨識，是我們在規劃「人性化電腦整合智慧型環境」裡提供使用者友善的「人機互動介面」與「行為分析」兩主題中最重要的一環。長久以來其也一直是電腦科學中活躍的研究主題之一。由於視覺感測器於現今電腦上的普及，也由於在日常生活中，人與人之間的溝通即主要依靠著語言、文字並適當結合手勢、身體姿勢與面部表情等來傳遞，因此視覺訊息在於我們的系統，可說是傳統語音與自然語言理解之外，最能有效補充人機介面效能的新方向。而另一方面，在智慧型環境中，人體姿態的辨識其最終的目標在輔助電腦對使用者行為作適當的理解與描述，乃至於近一步的預測與學習，故其更是全自動化服務系統中不可或缺的技术。隨著電腦影像技術的日趨成熟，投注於人體姿態辨識之研究也與日俱增，人體姿態辨識主要的目的在於估算出各別身體軀幹於影像中的狀態，其中包含位置與角度，以做為進一步分析行為動作感測的依據。關於分類與概述人體姿態辨識之技術發表於[23][24][25]，姿態辨識的技術依據使用人體模型的差異，大致上可分為三類：

1、沒有明確形狀模型的二維方法：

此類方法通常沒有模型的觀念，僅直接對影像中的像素作低階影像處理的演算，進而擷取出諸如：外觀、輪廓、顏色或其他重要肢體特徵等參考樣本，來判斷人體姿態或手勢，他們通常無法很精準描述我們期望的結果，但能迅速有效率的偵測出近似值。此類方法變化性繁複沒有一特性的共通點，因此較著名的一些影像方法：如 SNAKE、AAM (ACTIVE APPEARANCE MODEL)、SVM、PCA (PRINCIPLE COMPONENT ANALYSIS) 或 WAVELET

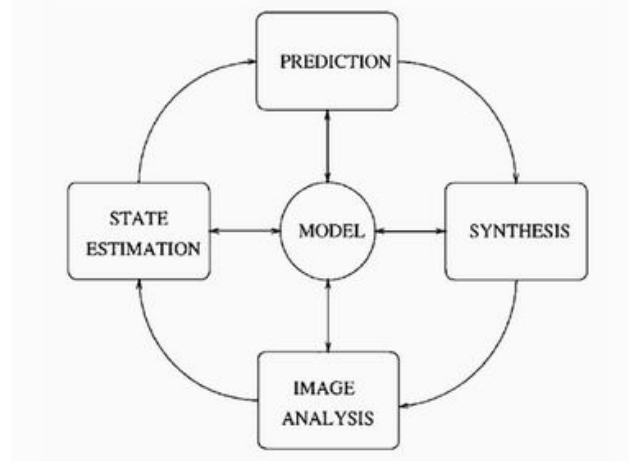
等都曾被使用於此範疇。

2、擁有明確形狀模型的二維方法：

為克服精準描述的上述問題，很多研究者傾向提供一明確人體模型來作為偵測的輔助工具，這方面的研究主要在追蹤身體每個區段於影像中的位置。利用事先形狀模型的知識可以提供我們人體運動上的限制資訊，減低搜尋空間與錯誤判斷率。在此人體模型通常藉由一些簡單的幾何形狀來表現人體各個部位，也可能僅使用一些簡單的線段或曲線來表達，然遮蔽的問題是其先天無法克服的缺陷。此類方法能於 [23] 中找到更詳細的說明。

3、三維方法：

三維的方法本質上類似於上個類別，需要使用者提供一明確三維模型來輔助辨識，程式透過模型 3D 投影與所擷取的影像來擷取匹配的人體或手勢姿態。然在此方法中利用多台相機追蹤人體，可以賦予事先我們對動作的知識。因此這類的演算法大部分包含幾項重要的步驟：人體模型的建立、特徵點對應的尋找與三維人體姿態的攫取與持續追蹤。O'Rourke [26] 與 Badler 給予這項研究一個主要的方法，他們將辨識系統分為一些重要的元件，如圖十所示。



圖十：三維演算法架構圖

(三) 使用者喜好與多感測器之客制化服務

『智慧型家庭系統』的研究領域包括系統架構設計、資訊家電、家庭看護等等，但是的目前的這些研究都是屬於被動的提供服務而不是主動的提供服務，並且目前這些研究的領域都過於專注於對於服務的提供，但卻沒有進一步的了解使用者為什麼想要這個服務。於是系統中便少了人機之間的互動，也少『智慧型家庭系統』中最重要的部分，就是如何在適當的時間提供適當的服務給使用者。最重要的就是要讓『智慧型家庭系統』充分了解使用者的使用習性，並且在了解使用者的喜好習慣後可以進一步的事先提供服務給他。

Mozer [27]基於類神經網路的技術，建構一個類神經網路家庭(Neural Network Home)用以控制空調、燈光以及熱水。另一個由 Georgia Institute of Technology 所建構的 Aware Home [28]，經由感測居住者的喜好來達到輔助居住者的目的。麻省理工學院(MIT)提出所謂 House_n 的構想[29][30]，著重在於如何整合現有之技術，使系統能夠依據環境變化而具有動態(Dynamic)以及逐步調整(Evolving)的能力。MavHome 於[31]所提出，主要整合多種不同的感測器，以提供居住者適當的服務。雖然這個觀念是智慧型家庭系統的中心思想，目前其他國家也有相關

研究在進行中，但是卻都只停留在資料收集[33][35][36]，系統架構[34]，以及偵測不尋常行為模式[32]等，沒有提供一個有效的辦法去找出使用者的使用習性。

(四) 智慧型環境建制與整合

隨著科技的進步，電腦開始進入人類生活的每一個部分，嵌入式裝置、資訊家電、及家庭控制網路的出現讓人類生活中存在各種不同類型的電腦提供應用，人類跟電腦之間的互動不再侷限於坐在電腦前用螢幕、鍵盤、滑鼠，而是可以擴展到更多樣化的方式，也開啟了智慧型環境的研究及應用，不論是辦公室、家庭、教室、會議、健康照顧、生活協助，這些都是智慧型環境可以應用的領域。智慧型環境是一個極度重視環境導向（context-aware）的使用者介面，它是一個裝置了感測器（Sensor）的空間，讓感測器可以觀察並瞭解環境中的活動，並在得到這些資訊後，經過推論，對某些事件做出適當的反應。智慧型環境可以被視為一個連接著各式各樣的設備（Appliances）的人工活體（Artificial Organism），一個不會移動的機器人，它無間斷地（Seamlessly）環繞著使用者，將使用者置於它的空間之中，並提供使用者智慧型的協助及服務。

在建構智慧型環境時，首先就是提供各類不同的資訊輸入，其中包含：(1)藉由攝影機得到環境中的影像資訊；(2)通過麥克風擷取環境中的聲音作為事件觸發的判斷；(3)設置壓力感測器偵測環境中使用者的行為；(4)從使用者與系統間人機介面的互動來瞭解使用者的目的；(5)透過使用者的互動及注意力集中的焦點來取得環境中的事件；(6)藉由直接與使用者身體接觸獲得資訊。近幾年來，一些比較著名的智慧型環境研究計畫包括：日本東京大學（University of Tokyo）的 Intelligent Space[37]及 Robotic Room[38]，IBM 的 BlueSpace[39]，法國國家電腦科學及控制研究學會（INRIA）的 SmartOffice[40]，Microsoft 的 EasyLiving[41-42]，美國加州大學（University of California）用 AVIARY[43]發展出來的 Intelligent Meeting Room[44]，美國科羅拉多大學（University of Colorado）的 ACHE[46]，以及瑞士 ETHZ（Swiss Federal Institute of Technology Zurich）的 Ada[45]。

除了建置智慧型環境、取得環境資訊之外，一個整合數位服務的系統架構也是很重要的。有了這樣一個平台，可以有效地為使用者提供服務，可以為系統中的資源提供溝通的機制，也可以讓系統具有更大的發展空間。因為在智慧型環境與使用者的互動之中，會產生相當多的變化。系統的服務會因應使用者的需求、系統升級等等原因而新增或移除服務，或是因為外在環境改變而必須動態地改變系統本身的架構。各項系統中提供服務的資源可能動態地進行各種組合，原本存在於環境中的硬體資源可能因為失效而無法繼續提供服務，甚至是因為意外而讓服務停止。相對地，若是系統中有新增加的資源較為適合提供服務，系統也可能會因應做出處理。而若是使用者本身改變位置、身份等個人資訊而影響到系統中對人機互動的設定，系統也得做出必須調整服務的判斷。除此之外，使用者與電腦所需要的資訊呈現方式並不相同，甚至系統中每項資源的需求都不一樣。我們將針對上述問題探討智慧型環境中人機互動服務的機制，並提出一個整合數位服務的平台，讓系統除了能夠觀察環境中的資訊，適應使用者的行為，還可以依據不同的需求及特性來組織系統的資訊呈現系統的狀態，針對環境資源的變化動態調整自己的架構，進而將最佳的互動服務提供給使用者。

第二節、研究目的

顧名思義，「人性化電腦整合智慧型環境」，即是希望能塑造一個環境，使用者無論在環境中的何處「彈指之間」及可取得所需要之服務，達到“服務無所不在”(Ubiquitous Service)之境界，為達此，須首先解決四項子題之研究工作，分別為：『多目標無間隙人形追蹤系統』、『人體姿態辨識』、『使用者喜好與多感測器之客制化服務』、『智慧型環境建制與整合』，以下即針對各項子題之目的作一詳細之描述：

(一) 無間隙人形追蹤系統

多目標無間隙人形追蹤系統的主要在利用多個影像設備，以涵蓋所有人形之運動空間，並進一步有效偵測與追蹤多個人形在智慧型環境的位置與運動軌跡。綜言之，在此計畫中，人形追蹤系統主在到以下目標：

- **即時 (Real-time) 監控與追蹤：**影像系統能夠快速的分析影像，使系統不至於出現延遲與停頓的現象。
- **多人追蹤：**在一般的應用環境中，通常不侷限於單一使用者，在這樣的考量下，此系統需能夠在同一時間同時追蹤多個使用者。
- **多台相機追蹤：**為追蹤使用者在系統所預先設定運動空間之位置，相機需能涵蓋所有使用者之運動空間。
- **人形 (Identity) 身分的識別：**在多相機的室內環境中，進行多人之偵測與監控，系統需能夠標示出不同人形的身份，並保持此目標物身份的前後一致性 (Consistence)。當有新的人出現在運動空間中，系統需動態的識別並透過多相機對其識別與追蹤。
- **能在辨識與追蹤受到部分遮蔽 (Partial Occlusion) 與不同姿態 (Posture) 的人形：**當人在運動的過程中，難免會與其他的人產生彼此互動的行為，此會導致影像中遮蔽現象的產生。另一方面，人形在影像中所呈現的姿態並非固定不動的，系統需能夠容忍這兩種人形變動的因素，以有效的偵測並追蹤人形的軌跡與位置。

(二) 人體姿態辨識：

人體姿態估算可以被應用來感測使用者的意圖(Intension)，然而由於人形外形與姿態的變化非常複雜，因此如何有效的偵測出人形是一件不易的工作，因此如何克服下列所列之主要挑戰(Challenges)為此計畫之主要目的。

- **衣服穿著(Clothing):** 行人所穿著的顏色與樣式非常多變，因此透過建構顏色模型來達到偵測行人的目的。
- **遮蔽(Occlusion):** 由於人體有許多的自由度(Degree of Freedom)，因此會造成身體有些部分無法成像於影像上。
- **複雜背景(Complex Background):** 在一般自然行車場景中，影像通常會有著複雜的背景，提高行人偵測的複雜度。
- **動作複雜度(Articulated-Deformation):** 由於人體的四肢有諸多自由度，因此會產生許多姿態。

(三) 使用者喜好與多感測器之客制化服務

「透過多感測器檢知使用者喜好，進而達到服務客制化」(Service Customization through Multi-Sensor Based Preference Perception)主要是利用多種感測裝置偵測使用者在家庭使用設備（一般家庭電器、燈光、空調等設備）的偏好，並進一步建立使用者的喜好模型，而最終的目的就是做到『根據該使用者以往的使用偏好而動態提供適合的服務』以達到『服務客制化』的目標。在此主題中，我們希望能達到以下目標：

■ 建立有效的多種感測資料收集系統

建立一個可以有效收集智慧型家庭內的各種感測器資料的系統，透過不同的感應裝置收集不同的資料，收集的資料包括有家電控制（電視、燈光、冷氣等家電）、使用者移動位置、環境變數（溫度、溼度、亮度等）、時間日期等，並且將不種感測器的資料群組，並且去除多餘的資訊，以提供更精確有效的資訊給『服務客制化系統』分析。

■ 建立使用者喜好模型

將收集到的資料利用隱藏式馬可夫模型（HMM、Hidden Markov Model）建立起使用者喜好模型，並且動態的調整模型架構。

■ 即時運算使用者目前狀況並事先提供服務

因為人的行為會一直在改變，因而導致各種感測資料的變化，所以必須在實際的測試環境中即時擷取目前的各種感測器的資料，並且將資料丟入模型計算，算出使用者目前想要的服務，並且進一步的先提供改服務給使用者。

■ 多人環境偵測

在一般的應用環境中，通常使用者並不只一人，所以在這樣的考量下，此系統需能夠在同一時間收集不同使用者的資料並建立不同使用者的喜好模型。

■ 動態調整使用者喜好模型

系統必須提供一個錯誤回朔機制，並且透過錯誤的經驗動態調整使用者喜好模型，讓系統更穩固。

(四) 智慧型環境建制與整合

所架構的智慧型環境在面對使用者一般行為時，必須裝備傳統的感測器來偵測，並針對非傳統的人機互動方式採用一些新型的感測系統，從多方面廣泛收集環境中的資訊。而若要監控系統中的資源，除了嵌入式系統及資訊家電可以藉由與電腦連結的能力直接與系統溝通之外，傳統的設備必須以家庭控制網路及閘道器（gateway）與系統連接，才能讓這些設備成為智慧型環境中可用的資源。關於所建構的智慧型環境，我們希望能夠達到以下目標：

■ 多重感測：

首先，智慧型環境必須具備接收外界資訊的能力，才有據以操作系統的參考資訊。可以參考的資訊愈多，環境能夠做出的互動行為也愈豐富。

■ 多樣互動（Multimodal Interaction）：

使用者除了使用滑鼠、鍵盤、顯示器等傳統方式與一般電腦進行溝通之外，智慧型

環境必須讓使用者的溝通方式不必受限於電腦前，提供可以讓使用者在環境中移動並同時溝通的方式，甚至以一般的自然模式與系統溝通。

■ **監控環境資源：**

智慧型環境除了能夠控制環境中的資源與使用者進行溝通之外，還必須能夠知道系統中這些提供服務的各項資源狀況，供智慧型環境對本身系統進行評估與分析，作為系統調整資源分配的參考資料。

■ **學習使用者資訊：**

除了與使用者進行互動之外，智慧型環境還會觀察使用者與環境資源的互動行為，記錄使用者在系統中的狀態，評估使用者與系統的互動情形，作為改進服務的依據。

■ **動態調整架構**

因應環境中的外在因素，系統中內部資源的改變，或是適應使用者的行為，智慧型環境將動態調整本身的系統架構，提供最適合目前狀態的服務。

第二章、研究方法與執行成果

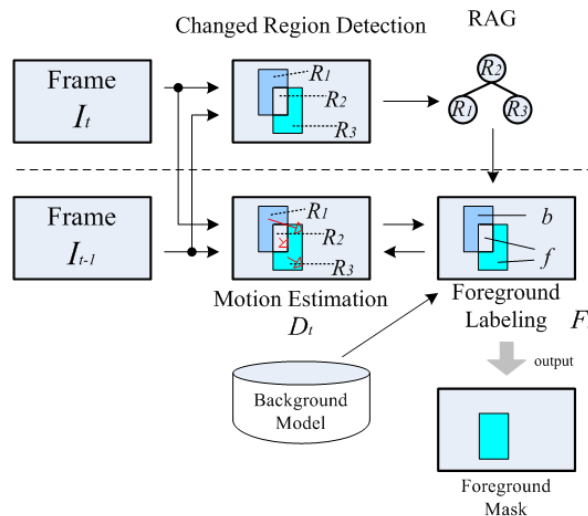
第一節、無間隙人形追蹤系統

影像系統與其他感測系統最大的差異在於能夠完整偵測人形在環境中之位置與行為姿態，而在此系統中，我們預期利用多個影像設備覆蓋所有智慧型環境空間，並進一步有效偵測與追蹤多個人形在環境中之位置與運動軌跡。就技術方面而言，以影像為基礎之室內多相機人形追蹤系統主要涵蓋：前景切割 (Foreground Segmentation)、人形辨識 (Human Recognition and Classification) 以及追蹤 (Human Tracking)。於此計畫報告中，我們將詳細說明於單一相機下所採用之前景偵測與人形辨識的技術。

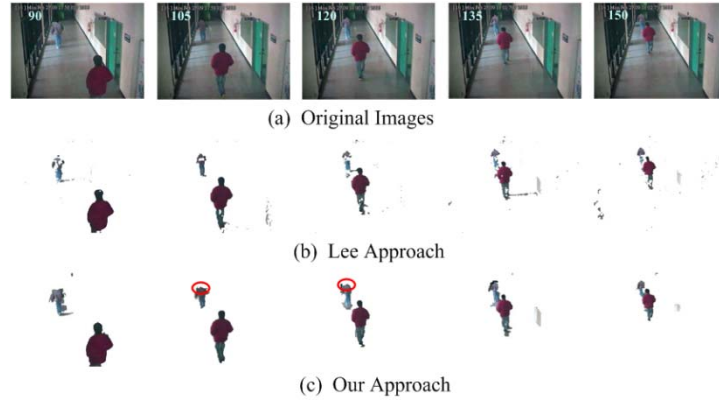
(一)前景切割

前景偵測的主要目的在於將所要分析的影像自背景中切割出來。由於光線波動的影響，相機所感測之影像亮度與顏色並非固定不變；另一方面，背景亦可能隨著時間而呈現不同的景緻。大致上來說影響背景變動的因素可分為：漸進光線變化 (Gradual Illumination Variation)、劇烈光線變化 (Sudden Illumination Variation) 以及陰影 (Shadow)。目前用來進行前景切割的技術主要可以分為兩類：背景相減法 (Background Subtraction) 以及根據運動切割法 (Motion-Based Segmentation)。背景相減法意即透過背景模型建構之方式將背景影像記錄下來，經由與背景模型對照相減後，將前景擷取出來，然而這類主要是在像素層次 (Pixel-Level) 進行前景切割，因此容易受到光線變化的影響。而經由運動切割法的技術，主要根據前景的出現皆會伴隨著運動的概念，經由運動切割的方式來切割出前景區域，然而如何有效估算運動 (Motion Estimation) 仍是非常複雜的問題。

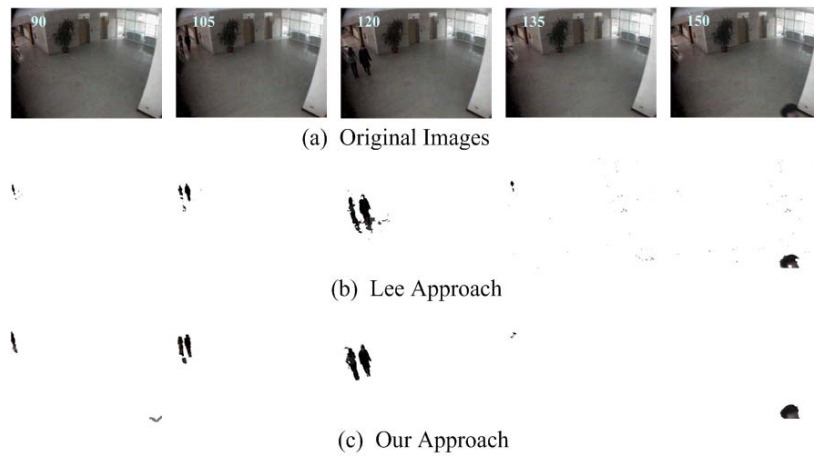
有別於第一年所提出之背景相減法，我們提出一個架構 (如圖十所示) 整合了背景相減以及運動切割兩種方式來達成。主要的貢獻在於利用運動資訊將前景偵測從像素層次 (Pixel-Level) 提升至區域層次 (Region-Level)，另外根據貝氏網路我們整合背景模型以及運動資訊，同時達到前景偵測以及運動估算。圖十一與十二為我們針對 IPPR 測試影像與 Lee 所提出之方法所獲得之結果。



圖十：前景切割架構圖



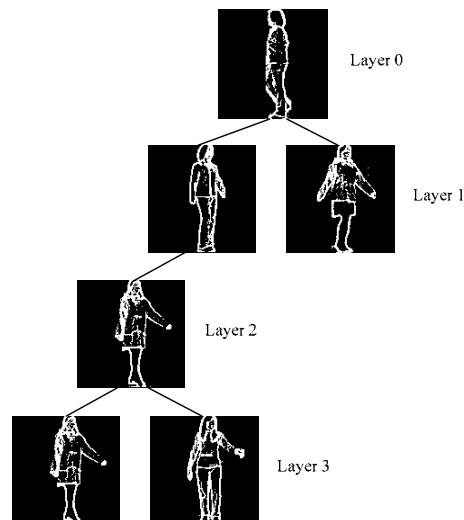
圖十一：IPPR 測試影像一



圖十二：IPPR 測試影像二

(二)人形辨識

在這個計畫中，主要是利用人形樣版來偵測人形於影像中之存在，由於人形影像隨著其姿的不同而會有不同的影像外觀，因此我們需要建構一個人形樣版資料庫作為人形比對之用。而為使人形之比對更有效率，使人形辨識能夠達到即時的效果，我們將資料庫之人形樣版做有系統的建構，所採用的方法為基因演算法（GA Algorithm），圖十三為所建構後之部分人形樣版資料庫。



圖十三：人形階層資料庫

由於人形並非剛體，因此我們採用 Hausdorff Distance 來進行人形樣板比對以達到人形辨識的目的。Hausdorff Distance $H(.,.)$ 的定義為，對於兩個集合 $A=\{a_1, a_2, ..., a_{N_A}\}$ ， $B=\{b_1, b_2, ..., b_{N_B}\}$ ，

$$\begin{aligned} H(A, B) &= \max(h(A, B), h(B, A)) \\ h(A, B) &= \max_{a \in A} \{D(a, B)\} \\ \forall a \in A, D(a, B) &= \min_{b \in B} \{d(a, b)\} \end{aligned}$$

其中 $d(.,.)$ 為 Euclidean Distance。然而基本的 Hausdorff Distance 其比對的效果易受的真實環境中的雜訊影像。在此計畫中，我們採用了 M-HD 做為樣板相似度 (Similarity) 比對的度量，主要的原因在於，M-HD 在避免雜訊干擾以及遮蔽現象上有較好的成效及結果。M-HD 與原始 Hausdorff Distance 最大的差異在於，其導入了所謂的 Cost Function ρ ，定義如下：

$$\rho(x) = \begin{cases} |x|, & |x| \leq \tau \\ \tau, & |x| > \tau \end{cases}$$

其中 τ 為預設的門檻值。換句話說，在結合了 Cost Function 與 Directed Hausdorff Distance，M-HD 的函數可定義如下，將所擷取之影像與人形資料庫比對後，若其相似度大於所設定之門檻值，即意味所偵測之前景為人形。

$$h_{M-HD}(A, B) = \frac{1}{N_A} \sum_{a \in A} \rho(D(a, B))$$

(三)物體追蹤

於多目標物追蹤技術，我們使用了以貝氏濾波技術(Bayesian Filtering)[47]為基礎的，並且在量測影像相似度上考慮接合影像相似度(Joint Image Likelihood)[48]，作為解決目標在影像觀測上重疊時的判斷。

目標物的追蹤即為一種狀態的估測(State Estimation)，我們先由考慮單一目標的追蹤開始，採用粒子濾波技術(Particle Filtering)為基礎的 Sequential Monte Carlo 方法來有效取得每一時刻之事後機率的機率密度函數(pdf, Probability Density Function)。為了減少濾波時耗費的計算時間以及增加估測準確度，我們採用了重要性取樣(Importance Sampling)的方法[49][50]，將對目標物的已知資訊整合進影像追蹤當中，整合重要性取樣的粒子濾波技術可以用下式表示之

$$p(\mathbf{x}_t | \mathbf{z}_{0:t}) \approx \alpha \sum_{i=1}^{N_t} w_t^{(i)} K(\mathbf{x}_t - \mathbf{x}_t^{(i), q(\mathbf{x})}),$$

其中 t 代表時刻， $q(\mathbf{x})$ 為重要性函數(Importance Function)，為利用對目標物的輔助資訊(Auxiliary Knowledge)所得到目標物狀態的可能區域，而 $K(\cdot)$ 為核函數(Kernel Function)用作 KDE (Kernel Density Estimation)， α 為正規化常數， $w_t^{(i)}$ 為重要性取樣修正權重(Corrected Weightings)。

$$w_t^{(i)} = \frac{f(\mathbf{x}_t^{(i), q(\mathbf{x})})}{q(\mathbf{x}_t^{(i), q(\mathbf{x})})} p(\mathbf{z}_t | \mathbf{x}_t = \mathbf{x}_t^{(i), q(\mathbf{x})})$$

而在多重目標物的追蹤可以考慮為多目標物結合狀態的事後機率(Posterior Probability)：

$$p(X_t | \mathbf{z}_{0:t}) \cong \alpha p(\mathbf{z}_t | X_t) \cdot \int p(X_t | X_{t-1}) p(X_{t-1} | \mathbf{z}_{0:t-1}) dX_{t-1}$$

其中 $X_t = \{\mathbf{x}_{1,t}, ..., \mathbf{x}_{M,t}\}$ 代表 M 個目標物的狀態向量(State Vector)集合，在這裡我們考慮此 M 個

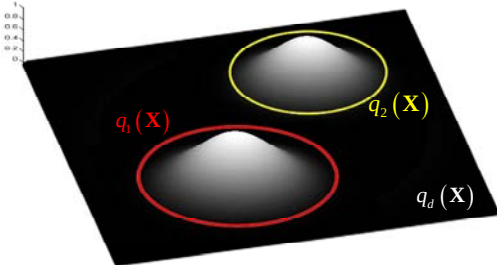
外觀相同的目標物， $\mathbf{z}_t$ 代表 t 時刻的觀測影像，而 $\mathbf{z}_{0:t}$ 則代表由初始時刻累積至 t 時刻的觀測影像歷史。 $p(\mathbf{z}_t | X_t)$ 為結合相似度(Joint Likelihood)，當目標物在影像當中沒有互相重疊之時，此式可以拆解成為

$$p(\mathbf{z}_t | X_t) = p(\mathbf{z}_t | \mathbf{x}_{1,t}) \cdots p(\mathbf{z}_t | \mathbf{x}_{M,t}),$$

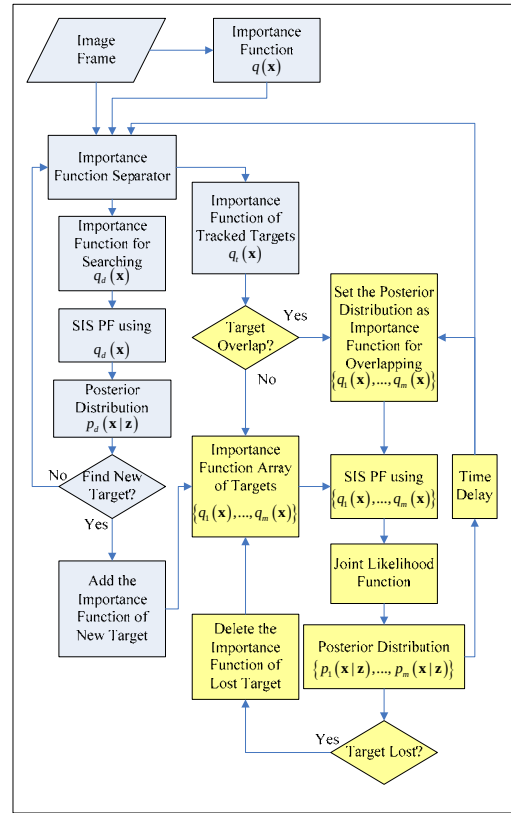
並且套用分隔式重要性取樣(Separated Importance Sampling)的方法，將對此外觀之目標物所求得之重要性函數 $q(\mathbf{x})$ ，利用上一刻的追蹤結果，分隔此重要性函數為

$$q(\mathbf{x}) = q_d(\mathbf{x}) + \sum_{m=1}^M q_m(\mathbf{x})$$

其中 $\{q_m(\mathbf{x}), m=1, \dots, M\}$ 為對第 m 個目標物的重要性函數，而 $q_d(\mathbf{x})$ 為扣除 $q_m(\mathbf{x})$ 之區域所構成之目標物偵測重要性函數，如下圖所示。



圖十四：分隔式重要性函數示意圖



圖十五：多目標物追蹤系統架構

套用分隔式重要性取樣，我們可以得到下式

$$p(X_t | \mathbf{z}_{0:t}) \cong \alpha \prod_{m=1}^M \sum_{i=1}^{N_m} w_{m,t}^{(i)} K(\mathbf{x}_{m,t} - \mathbf{x}_{m,t}^{(i), q_m(\mathbf{x})})$$

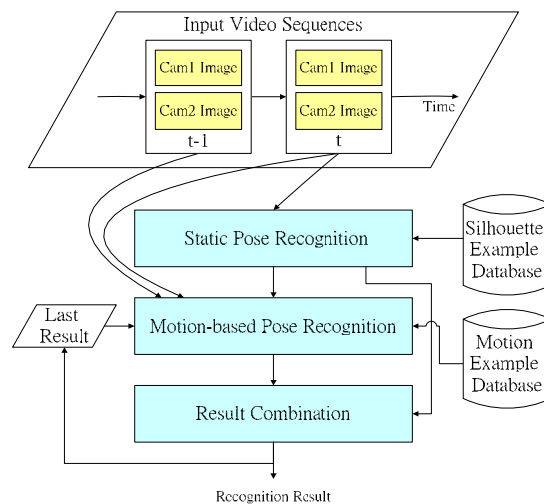
然而，上述式子所代表的意義即為分別對個別目標物量測影像相似度，但此情形只限於目標物無相互接觸、重疊或阻擋(Occlusion)的情況下，換言之，當目標物相互阻擋時，結合相似度便無法如上式拆解，不同目標物相互重疊時相似度必須結合考慮，因此在這裡我們採用一判斷機制，判斷目標物間的重要性函數是否重疊，藉此判斷是否需切換至結合影像相似度的量測，因此系統的總體架構如圖十五所示。

第二節、人體姿態辨識

建構於前項研究的基礎上，我們可以將「人體姿態辨識」視為一種人機互動介面。以電腦視覺為基礎的人機互動介面，可以讓使用者在無需利用傳統輸入裝置的情況下，自然且充分地使用肢體動作下達指令。我們設定使用者操作此介面時，以正面面對相機，視覺系統會偵測出人臉以及手臂在 2D 影像中的位置與姿態。於去年的研究中，我們利用單一相機來達成人體姿態的估算，但為了提升人體姿態估算的準確度，並達到可以即時應用的效能，今年我們採用兩台相機之影像來並採用以多樣本為主之演算法達成人體姿態估算的演算法。

(一) 系統概述

於人體姿態估算技術，我們提出一個結合靜態辨識和動作資訊以範例為基礎之方法，其中主要包含三個主要的部份，包括靜態姿勢辨識、動作為基礎的姿態辨識和結果整合(如圖十六所示)。採用兩台相機所擷取之影像來估算人體姿態，其主要的優點在於可以避免一些遮蔽(Occlusion)，和在單一相機中常會遭遇到辨識結果模稜兩可(Ambiguity)的問題。首先，藉著比較從兩台相機中擷取之的剪影(Silhouette)和資料庫的範例(Examples)來做靜態辨識，資料庫儲存多樣化的人物動作組態和相對於相機的 2D 人物外型，在這裡我們利用 shape context 來描述人物外型。我們比對範例並且找到最相似的當作答案候選人。接著，透過推算出運動資訊，然後這些運動資訊被用來推算哪一個結果比較好。這個方法使用顏色強度的變化來推算運動資訊，可以和以剪影為基礎的辨識系統互補，因為剪影捨去了光影強度資訊。而這個方法也克服了在傳統上用預測方法所造成的錯誤預測(Error Prediction)問題。最後我們結合在兩台相機上靜態姿態辨識和以運動為基礎的姿態辨識的結果，我們提出一個權重和的方式來作結果整合，最整合所得的結果就是我們方法最後所辨識出的結果。



圖十六：人體估算系統架構圖

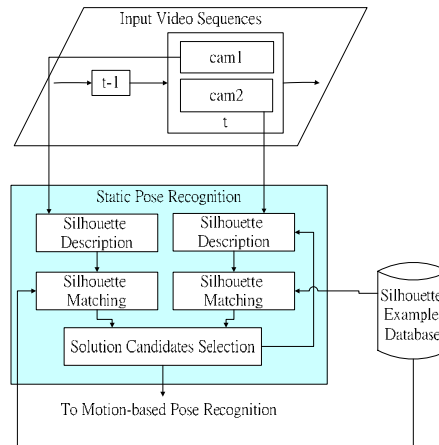
(二) 靜態姿態辨識

靜態姿態辨識就是只用目前的影像來辨認現在的姿勢而不會用到過去的資訊，我們提出利用剪影以範例為基礎的方法。剪影是一種常被用來辨識重建 3D 人物的方法，因為它包含豐富的人物的姿態資訊，而且不會被人物衣服的顏色和紋路影響。我們利用觀察到的剪影去跟資料庫裡合成的剪影做比較，我們將最相似的剪影所對應的姿態當作我們所判斷的。

在剪影為基礎的方法中，我們使用兩台相機當作輸入。原因是因為剪影只有一個物體 2D 外型的資訊而沒有了 3D 結構的資訊。為此我們利用不同角度的相機去觀察使用者，如此使用者的姿態較容易被辨認出。例如圖十七，在圖十七(a)中我們很難利用剪影去判斷人左臂的位子，但是如果我們有另一個在不同角度相機所觀察到的剪影如圖十七(b)，我們可以很容易知道左臂是否有伸直。但是，當手臂靠近身體的時候，有可能兩台相機的視野皆被遮蔽到，這是仍然很難判斷位置，我們將利用動作的資訊去解決這個問題。圖十八為我們方法的流程。首先描述剪影的階段，我們從輸入的影像擷取出剪影，剪影的外型再經由 shape contexts 來描述。然後，我們在利用輸入的剪影去跟範例做剪影比對。為了節省時間，我們在選取答案候選人時使用串接方式。一開始我們比對第一台相機，刪去比對分數太高的範例。之後再比對第二台相機，一樣刪去分數太高的範例。最後剩下的範例當作答案候選人供之後估算。



圖十七：不同視角之兩台相機影像擷取圖

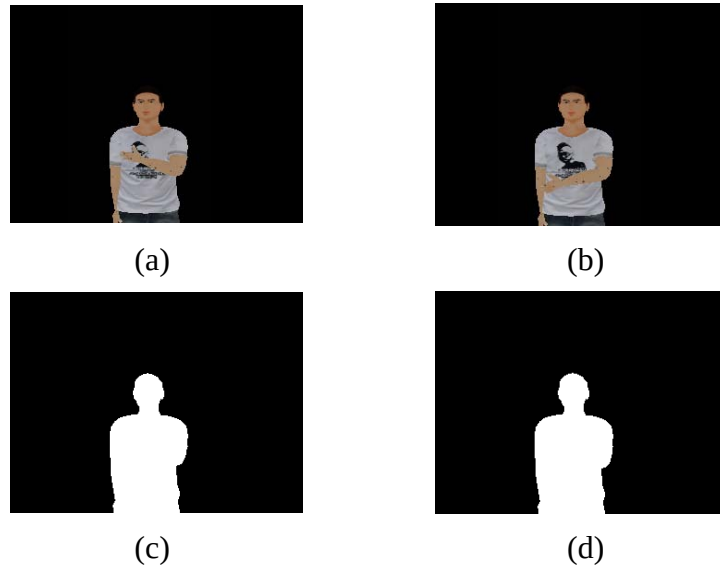


圖十八：靜態辨識的流程圖

(三) 以運動為基礎的辨識方法

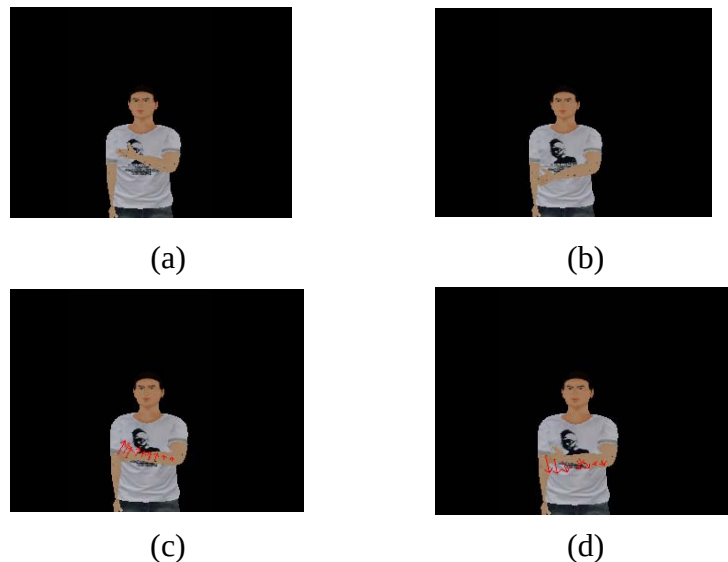
雖然剪影對於人類動作的辨認提供豐富和直覺的資訊，並已經被廣泛應用在人體姿態估算技術上，然而其卻忽略了剪影內部的資訊(如圖十九所示)，在影像中左前臂在身體區域，雖然在兩張影像中左前臂的位置是不一樣的，但是再剪影中卻看起來一樣。因為有許多候選人有相似的剪影而他們似乎都是適合的答案，這個問題將會造成以剪影為基礎地辨識方法不能夠找到真正的姿勢。基於上述考量，我們提出一個方法利用運動資訊的演算法來克服這個問題，如圖十九所示，僅使用剪影，前臂的位置是無法被正確估算出來。然而，假設我們知道從上一張影像到目前這張影像的動作，我們可以使用這些資訊來估計現在的動作，例如圖二十，我們可以很容易區分出有相同剪影的兩個不同動作。假設姿勢在前一張影像是一樣的，而之後兩個姿勢的動作不同，因為動作的資訊是從影像顏色強度改變所，而這部份被剪影捨

棄，因此我們的方法可以跟以剪影為基礎的辨識方法有好的互補。



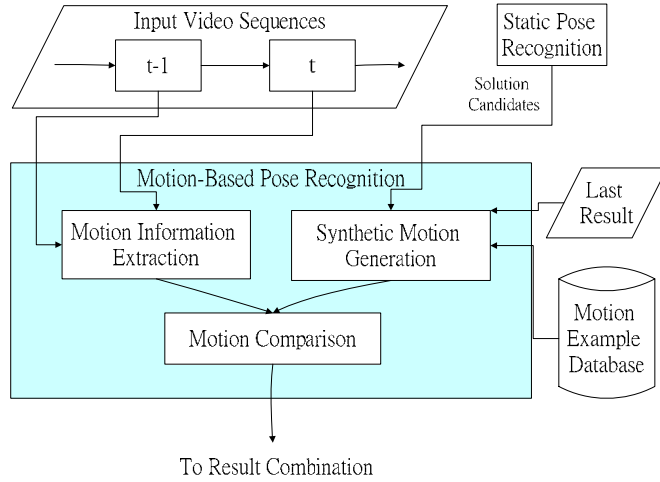
圖十九：模稜兩可的問題這個範例表示使用剪影產生。(a, b)表示兩個不同的姿勢，而(d, e)為對應的剪影。

我們很難分辨兩個剪影的不同，所以我們不能僅僅使用剪影來分辨姿勢。



圖二十：利用動作資訊來辨視動作。(a, b)為我們不能用剪影分的動作。(c, d)表示動作資訊可以幫助我們辨識出動作(a, b)，在這裡箭頭表示動作方向和大小。

以運動資訊為主之姿態估算流程圖如圖二十一所示，產生以動作為基礎辨識方式的範例，為節省時間，我們只比較在靜態姿勢辨識時所選出的可能答案。我利用前一個辨識結果和範例來估計身體部位的移動，產生每個可能範例的合成動作。我們利用在輸入影片中上一張影像 $t-1$ 和目前影像 t 利用 optical flow 的技術擷取出真實動作資訊。因為輸入影像人物的大小和位置也許和範例中人物不一樣，因此我們利用在 shape context 比對所得到的剪影輪廓的點的對應關係，估測特徵點位置在輸入影像和範例的轉換，藉著使用轉換，我們可以校正真實運動和合成運動，藉著計算輸入影像和每個範例合成動作的距離，我們做運動的比較。



圖二十一：以運動為基礎辨識方式的流程。

(四) 姿態估算結果整合

在獲得靜態辨識和以動作為基礎的辨識之後，下一個步驟極為如何從辨識結果中找最適合現在姿勢的答案。最直覺的方法就是使用以動作為基礎辨識系統的具有最小動作距離的範例。因為以動作為基礎的辨識只比較利用靜態辨視所選出的答案候選人，因此在動作為基礎的辨識方式中最好的答案，應該在靜態辨識中也是好的答案。在某些情形下，若只使用以動作為基礎的辨識方式，不能夠辨識出正確的姿勢，例如：當超過一個範例有相同且最小的動作距離，因此我們想要使用靜態辨識的結果來幫助找到較好的答案，亦或輸入影像所觀測到的動作資訊不足，導致無法辨識姿勢，而且太少的動作資訊很容易雜訊影響，導致上一張辨識結果有一些錯誤發生，那麼將會影響姿態估算之正確性，因此結合靜態和以動作為基礎的兩種辨識會是比較好的方式。一種簡單的整合方法，是假設兩類辨識結果是獨立的，在這種假設下，我們可以兩類辨識所得到的距離加起來，數值最小的當作我們的結果。然而，並非所有情形下，兩種辨識系統重要性皆相同。因此我們給予不同辨識方法和相機的結果不同權重，使用權重和的方式整合結果。

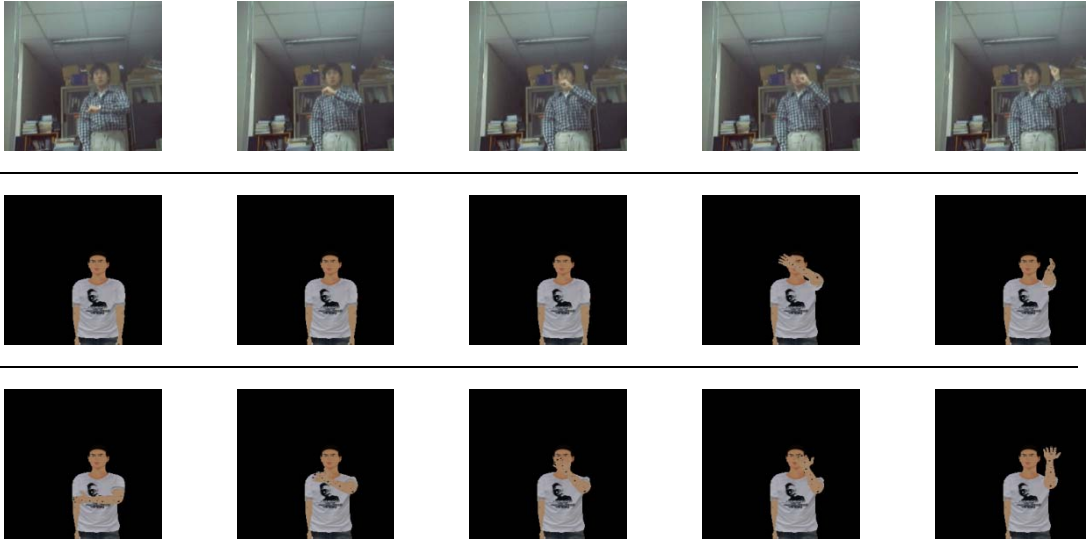
為計算目前影像中一種辨識方法在一台相機上的結果有多可靠，我們引入一個評量標準稱為可靠度(Reliability)。可靠度值介於 0 到 1 之間，他表示使用這個辨識方法的辨識結果為真實姿勢的機率。既然我們有兩種方法和兩台相機，對每一張影像可以得到四個辨識結果 r_1 、 r_2 、 r_3 和 r_4 。給定在現在時間影像的辨識結果 r_k ，我們想要估測可靠度 R_k 。我們的想法是最好的結果範例和輸入影像的距離可以表示辨識結果的可靠度。例如，假設在靜態辨識中，若最好的辨識結果 r_A 的範例跟輸入剪影有很小的剪影距離 10，和另一個最好的辨識結果 r_B 的範例跟輸入剪影有很大的剪影距離 200，我們認為 r_A 較為可靠，因為它跟輸入的剪影很相似。如果最好範例跟 r_B 一樣距離很大，這種情形代表在資料庫中沒有跟輸入相似的範例，有可能是的剪影擷取或者是某些身體部位有遮蔽的情形，因此我們用在結果中最小的距離來估計靜態辨識的可靠度。我們想要估計的可靠度函數可以表示成 $\Gamma_{static}(d_{min})$ ， d_{min} 為在 r_k 中 $d_{k,j}$ 最小的值。估算以動態為基礎的辨識方式地結果可靠度，不能僅僅考慮最小的動作距離，還要考慮從輸入影像中擷取出來的動作向量個數。這是因為動作資訊多寡可能會影響動態辨識結

果的可靠度，假設動作向量數目很少，代表我們只有一點點動作資訊，所以我們可能會找到許多和觀測動作相似的範例，因此可靠度函數可以表示成 $\Gamma_{motion}(d_{min}^k, N_k)$ ， d_{min}^k 為在 r_k 中 $d_{k,j}$ 最小的值， N_k 表示在以動作為基礎的辨識 k 中動作向量的個數。

估算完每一個結果的可靠度值後，我們利用權重和的方式來整合不同方法和相機的結果。因為剪影比對距離和動作距離的單位是不一樣的，所以必須先正規化結果。因為最好的範例的距離比最差的有意義，因此我們以最好的範例距離為基礎正規化辨識結果。再每個結果中每一個範例 j 在正規化完後的分數為 $c_{k,j}$ ，分數 $c_{k,j}$ 的範位介於 0 到 1 之間，而在每一個結果中最好的範例數值為 1。令 $d_{k,j}$ 為每一個範例 j 在結果 r_k 中的距離，而 d_{min}^k 是所有 $d_{k,j}$ 中的數值，因此分數為 $c_{k,j} = \frac{d_{min}^k}{d_{k,j}}$ 。用不同方法所得結果權重，由他們的可靠度決定。理所當然給予較高可靠度的結果較高的權重，因此對每個結果 r_k ，他們的權重 α_k 計算如下

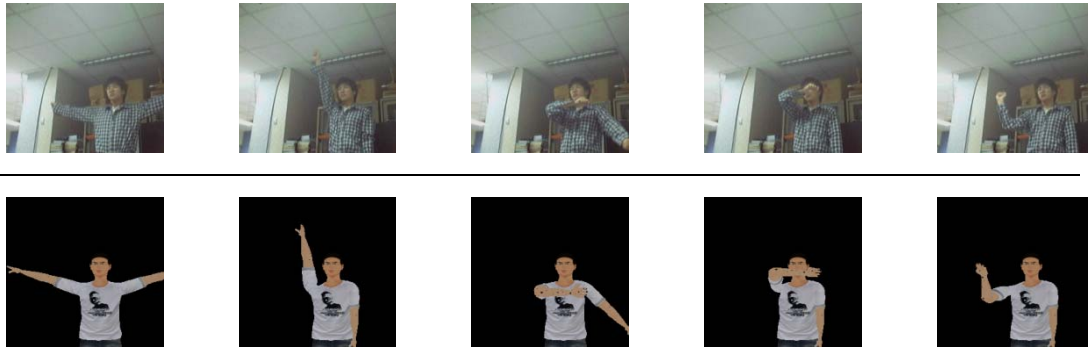
$$\alpha_k = \frac{R_k}{\sum_{i=1}^4 R_i}$$

之後對於每一個範例 j ，他的最後分數 ρ_j 為 $\rho_j = \sum_{k=1}^4 \alpha_k c_{k,j}$ ，經由比較最後分數 ρ_j ，我們選出分數最高的範例當作我們方法在這一張影像上最後辨識出的結果。下圖為利用單一相機與兩台相機所獲得之姿態估算之結果。



圖二十二：單一相機的結果。第一列為輸入影像，第二列為指使用靜態辨識結果，而第三列為加入動作資訊辨識結果



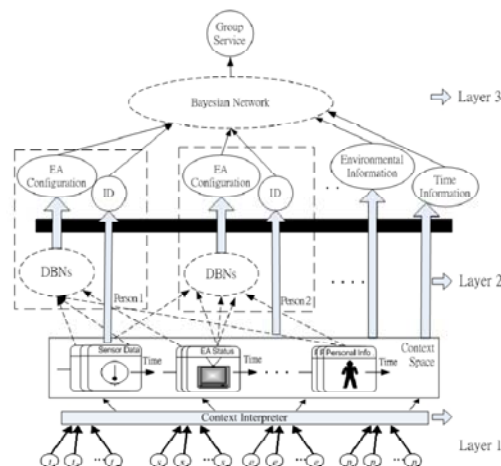


圖二十三： 我們方法的結果，這張圖顯示兩台相機的輸入和辨識的結果

第三節、使用者喜好與多感測器之客制化服務

「自動化個人喜好學習系統」的主要目的，是找出使用者的「個人喜好資料檔」，並建立「個人喜好模型」，學習預測多居住者的喜好並且提供服務的建議給他們，使他們感到舒適與放鬆，但是在家中，多位居住者和各種環境資訊間有很複雜的關係，環境資訊包含了人、事、時、位置和物件等等。居住者的習慣也有時間上的特性，所以，要恰當的學習出居住者的喜好必須將整個連續的互動加入考慮。

為達成上述目標，我們提出了三階層的模型(Model)，第一層負責處理從感測器傳來的原始資料，第二層負責每個居住者個體的喜好模式，最高的第三層負責處理多位使用者之間的關係，而推論出應該提供給整個群體的服務。使用階層架構的好處主要在於：1)各層的學習和推論可以分開處理，若環境的佈置改變了，我們只需重新訓練最下面那層，不用改變其他層的喜好模型。例如，客廳的佈置改變了，居住者的看電視行為的活動路徑也跟著改變了，但是多居住者之間的互動並沒有改變，所以只需重新訓練最下那層，其餘不用改變。2)階層式的架構將多使用者的喜好模型切成多個部份，較下面的階層處理各子問題，而最上面的階層統合下層所推論的結果。在本研究中，我們先將多位居住者的整體喜好問題切成各個居住者獨立的喜好問題，然後在最高的階層使用一個貝氏網路(Bayesian Network)來統合每個居住者喜好模型所獨立推論出來的結果，整個系統架構圖如下圖所示。



圖二十四： 三階層推論系統模型

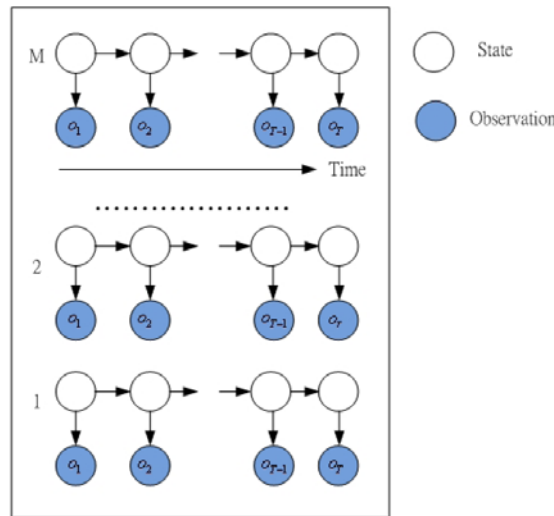
(一) 資訊轉譯(Context Interpreter)

第一層稱為資訊轉譯(Context Interpreter)，它負責將感測器所輸出的原始資料轉換成較高階的資訊，換言之，其賦予低階原始資料實質語意上的意義，它所轉換的資訊包含了時間資訊、感測器資訊、家電狀態和使用者資訊，另外，不使用低階的資料而使用高階的資訊的背後目的是降低建置我們所提的階層模型的複雜度。其主要可分為兩個部分：

- **群聚**：我們使用 K-means 演算法來把原始的資料聚成 k 群， k 是根據本領域的一般知識和實驗的結果來決定的。
- **量化**：根據群聚的結果及本領域的一般知識，我們可以給予每個群一個語意上的意義。

(二) 家電控制模型(Electrical Appliance Controller Model)

透過辨識居住者的日常生活習慣模式從而學習他們的喜好來個人化服務的提供，一般來說，上述所提的習慣模式是人的喜好和家電之間的相互關係。在本研究中，我們在第二層中用一個動態貝氏網路(DBN) λ^m 根據居住者的活動路徑和感測器與家電的資訊來預測一個家電控制器(Electrical Appliance Controller) E^m 的狀態，下圖為家電控制器的動態貝氏網路(DBN)圖，每一個家電控制器會對應到一個動態貝氏網路(DBN)，若環境中有 M 個家電，那就有 M 個控制模型。

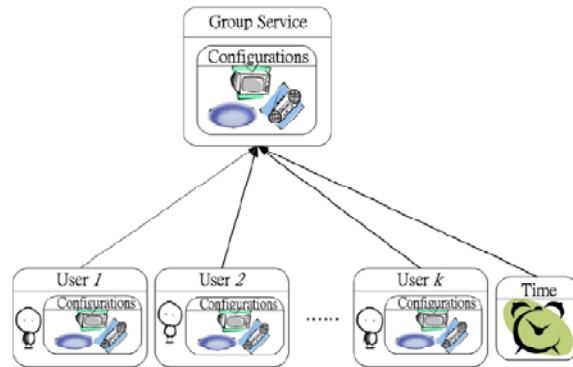


圖二十五：家電控制模型貝氏網路

每個動態貝氏網路(DBN)中的狀態 $X_t = ON, OFF, \dots$ 代表一個家電在時間 t 的狀態， $O_t = \{P_t, E_t, S_t\}$ 為所觀察到的資訊，其中：

- P_t ：表示居住者在時間 t 所在的位置。
- $E_t = \{E_t^1, E_t^2, E_t^3, \dots, E_t^M\}$ ：表示家電的資訊。
- M ：表示環境中家電的數目
- $S_t = \{S_t^1, S_t^2, S_t^3, \dots, S_t^R\}$ ：表示在時間 t 感測器所量測到的資訊。
- R ：表示環境中感測器的數目

(三) 多使用者互動模型(Multiple User Interaction Model)

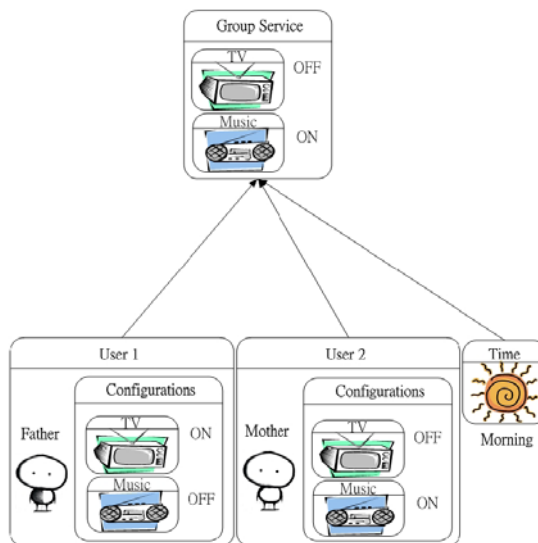


圖二十六： 互動模型系統概觀圖

上圖為多使用者互動模型(Multiple User Interaction Model)的概觀，在這層中，我們使用貝氏網路(BNs)來整合所有居住者所喜好的服務而滿足所有人的需求，我們把這個問題想成是學習並預測 K 個居住者的喜好，我們假設位在不同房間的居住者並不會互相影響並且我們都可以得知每個居住者的身分，感測器所量測到的資料中，若無法區分是屬於哪個居住者的，則屬於同房間的所有居住者，例如，聲音和溫度資料是屬於同房間的所有居住者的。我們認為群體的服務和每個居住者個別的服務、環境資訊及時間資訊有關，所以使用貝氏網路(BNs)把這些因素關聯起來，節點為那些影響因素，節點之間的相連是以本領域的一般知識來決定的，本貝氏網路(BNs)所包含的有：

- G ：表示所要推論的群體服務。
- $ID_1, ID_2, \dots, ID_K$ 為 K 個居住的的身分識別。
- $C^{(1)}, C^{(2)}, \dots, C^{(K)}$ 是第二層分別為 K 個使用者所推論出的喜好。
- T ：表示時間資訊。

如圖二十七，環境中有兩位居住者，父親和母親，時間是早上，父親想要打開電視和電燈，母親想要開音樂和電燈，根據以往資料所學習出的他們兩個的互動模式，系統推論出音樂打開而電視關掉是比較適合整體的服務。



圖二十七： 多人互動模型推論範例

第四節、智慧型環境建制與整合

所設計之「智慧型環境」是建構在多代理人程式（Multi-Agent）的概念上，而此系統架構中，有多個「代理人程式平台」，並有多種「行動式代理人程式」在「代理人程式平台」上進行移動、收集資訊、或執行工作。

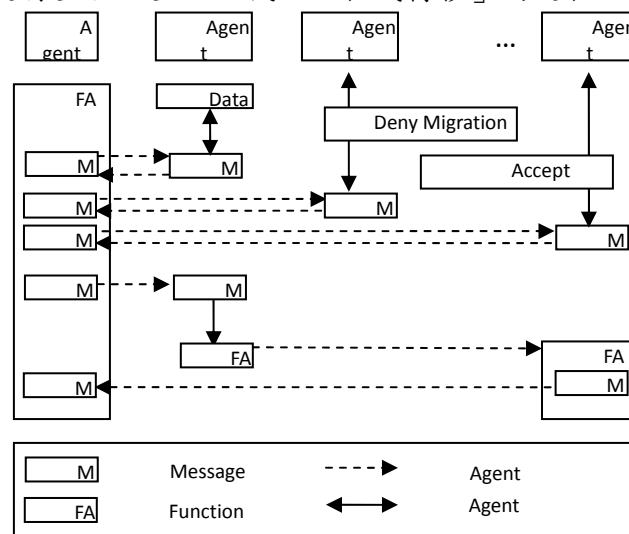
（一）降低網路頻寬及計算負載

在傳統的智慧型環境系統中，其管理機制通常是由一個中央伺服器進行整個系統的運作，因此系統中來往的訊息量龐大，計算量負載也很重。但是在以「行動式代理人」為基礎的系統架構上，各項管理功能是由派遣到各「代理人程式平台」的「行動代理人」進行，因此大多數的計算及溝通是藉由「裝置管理代理人」來完成，並不需要與系統中其他的元件進行通訊，也不需要耗費系統其他部分的資源。

另一項特點，則是藉由「裝置管理代理人」將系統中的環境資訊非同步地儲存在資料庫中，所以系統不需要藉由「輪詢（polling）」的方式來偵測環境中的狀態，也就避免了對那些很少發生的事件進行非必要的偵測，也避免了對於某些事件的反應來不及進行處理。藉由這種只在必要時才傳遞訊息的方式，就可以有效地降低網路頻寬。

（二）經由「代理人程式轉移」讓系統負載達到動態分擔

我們的智慧型環境架構的另一項重要特點，則是動態分擔系統負載。當系統中某項管理功能需要大量的計算而需要相當的計算資源時，則這項計算工作的負載就可以藉由「行動式代理人程式」的移動而被分散到其他的「代理人程式平台」去執行。首先，「行動式代理人程式」會從系統資料庫中尋找適合的「代理人程式平台」，然後再以 KQML[13] (Knowledge Query and Manipulation Language，一種 Agent Communication Language) 與其進行溝通協調。若是「代理人程式平台」需要保留其運算資源，如：播放 MPEG4 檔案或是執行重要程式，則這項「代理人程式轉移」就會被拒絕。而若是「代理人程式平台」願意接受這項工作，則「行動式代理人程式」就會從程式資料庫建立另一個「行動式代理人程式」，並轉移至願意提供服務的「程式代理人平台」進行被指派的工作。而在該工作執行完畢後，計算的結果就會被送回給原來的「行動式代理人程式」進行整合。整個「代理人程式轉移」的過程如下圖所示。



圖二十八：代理人程式轉移的流程圖

(三) 資源的彈性配置及卸載

在智慧型環境中，為了使用者的便利性及維持系統的運作，資源需要與系統架構保持彈性。這些資源通常是那些會偶爾離開系統，又會在某個其他地方或某個時間重新連上系統的裝置。而藉由「行動代理人程式」的能力，智慧型環境就可以彈性地對這些資源進行配置及卸載，而不會產生不好的影響。舉例而言，智慧型系統中的某個「代理人程式平台」是一部筆記型電腦，且其主要工作是為系統提供其運算資源。則這部筆記型電腦會使用 UPnP[14]向系統資料庫註冊自己的資源，以讓系統中其他的「代理人程式平台」可以從資料庫發現其存在並向其尋求運算資源上的協助。接著，考慮使用者將這部筆記型電腦從智慧型環境中帶離的狀況，則執行在這部筆記型電腦上的「行動代理人程式」就會暫停工作，然後轉移到其他的「代理人程式平台」繼續其工作。若是智慧型環境並非架構於「行動代理人程式」，則執行在筆記型電腦上的未完成工作就會因筆記型電腦離開智慧型環境的架構而遺失其運算結果。

考慮另一個例子，當智慧型環境中的可攜式電腦周邊或可攜式資訊家電在環境中的不同空間中移動的時候。當某個空間缺乏與 Jini 裝置溝通的能力時，則當某個 Jini 裝置轉移到該空間時，智慧型環境將無法控制轉移至該空間的 Jini 裝置，甚至無法發現該裝置存在於該空間。然而，在我們的架構中，可以藉由從系統的程式資料庫中啟動具備 Jini 溝通能力的「裝置管理代理人程式」，然後轉移至該空間對該 Jini 裝置進行管理，讓智慧型環境可以對該裝置繼續進行相關管理控制。這樣做的優點是「代理人程式平台」可以在有需要時動態地下載「裝置管理代理人」來應付空間中的新資源，又可以在不需要這份支援時將其移除，藉此減輕「代理人程式平台」的儲存空間負擔，而這一點對於執行「代理人程式平台」的嵌入式系統是一件很重要的事情，因為在資源相對缺乏的嵌入式系統中，記憶體儲存空間是很重要的。

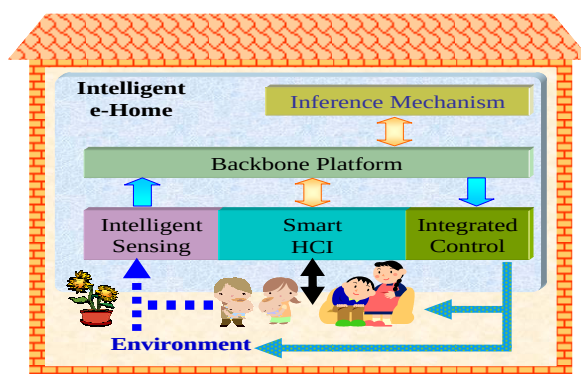
(四) 高容錯性

除了提供資源的彈性配置，系統資料庫還在「容錯機制」中扮演了一個相當重要的角色。在我們的智慧型環境架構中，只要各代理人程式運作正常，就會將自己的狀態以及系統的狀態做非同步的更新。因此，一旦有某個代理人程式因為硬體錯誤而運作失效，則當相關硬體修復後，代理人程式就可以根據自己先前儲存於系統資料庫的狀態進行回復。而若是運作失效的代理人程式所執行的工作與硬體無關，如：計算相關的工作，則這個因為代理人程式失效而中斷的工作並不需要等待硬體修復，其他的「代理人程式平台」就可以啟動另一個代理人程式，並根據失效的代理人程式先前儲存於系統資料庫的運算結果而繼續這個運算工作。此外，若是代理人程式的失效與「代理人程式平台」無關，則「代理人程式平台」就會重新啟動一個代理人程式繼續其中斷的工作。至於代理人程式儲存於系統資料庫的內容，則沒有一定的內容，而與代理人程式各自的需求有關。由於智慧型環境中所有的系統狀態，代理人程式狀態，以及代理人程式的程式碼，都儲存在系統資料庫中，因此智慧型環境的確會因為資料庫的失效而遭受到嚴重的破壞。然而，在代理人程式的架構下，智慧型環境的各部分功能仍然能夠正常運作及反應，只是無法產生新的支援功能。而且這項缺點可以藉由資料庫方面的技巧予以改進，像是在系統中進行多個資料庫的互相支援及備份。相對於須對整個系統進行備份以達到容錯，以「行動式代理人程式」為基礎的架構對於智慧型系統是相當適合的。

第三章、結果與討論

在此計畫報告中，我們詳述如何建構一個具備多重感測能力及多重互動方式的智慧型環境作為提供數位服務的平台，並在平台上發展一個具備高度整合能力的系統，此即導致建構一「人性化電腦整合智慧型環境」的終極目標，使系統能夠經由人體姿態估算與諸多感測裝置，進一步推論使用者需求而提供適當的各項數位服務，以提升更好的生活品質。

於今年的計劃中，我們針對系統各關鍵技術，提出創新之研究方法，並將其實作在所提出之智慧型環境系統，其中包含分為四個部份-無間隙追蹤、人體姿態辨識、使用者喜好與多感測器之客制化服務、智慧型環境建制與整合(如下圖所示)。



於無間隙追蹤方面，在整合背景模型與運動資訊下，使前景偵測更為準確且不受雜訊影響，且多目標物追蹤架構的提出，使追蹤系統能適用於有遮蔽的狀態下。人體姿態辨識方面，我們提出一個快速的人體姿態估算演算法，並整合多相機之剪影與運動資訊。為了能使系統實用於一般家庭環境，我們考量並建構多人之間的關係，提出一套三層之喜好推論系統，以提供適當之服務。最後，我們提出一個整合數位服務的平台，讓系統除了能夠觀察環境中的資訊，適應使用者的行為，還可以依據不同的需求及特性來組織系統的資訊呈現系統的狀態，針對環境資源的變化動態調整自己的架構，進而將最佳的互動服務提供給使用者。

第四章、計畫成果自評

於電腦影像應用方面，我們提出一套有效之演算法以達遮蔽狀況下之前景切割、人體追蹤與適用於不同服式顏色之姿態估算，且有多篇論文發表於相關國際會議與期刊上[51, 52, 53]。於系統平台與喜好推論方面，我們設計出一套彈性且方面之系統平台，且能夠正確分析所獲得之資料與喜好推論，所獲得之成果，也發表於國際著名期刊上[54,55]。

參考文獻

- [1] H. Hashimoto, J. H. Lee, and N. Ando, "Self-Identification of Distributed Intelligent Networked Device in Intelligent Space" *IEEE International Conference on Robotics and Automation*, 2003.
- [2] J. K. Horng, and M. Ouhyoung, "Photo VR Editor: A Panoramic and Spherical Environment Map Authoring Tool for Image-Based VR Browsers," *Communications and Multimedia Lab. Dept. of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.*, 2003.
- [3] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-Based Pedestrian Detection," *IEEE Intelligent Vehicles Symposium*, pp. 215—219, 2000.
- [4] D. M. Gavrila, "Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle," *International Journal of Computer Vision*, 2007.
- [5] Y. Wu and T. Yu, "A Field Model for Human Detection and Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 753-765, 2006.
- [6] D. M. Gavrila, "Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle," *International Journal of Computer Vision*, 2007.
- [7] D. M. Gavrila, "Pedestrian Detection from a Moving Vehicle," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2000.
- [8] Cheng-Yi Liu and Li-Chen Fu, "Computer Vision Based Object Detection and Recognition for Vehicle Driving," *IEEE International Conference on Robotics and Automation*, 2001.
- [9] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio. "Pedestrian detection using wavelet templates". *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1997.
- [10] Q. Zhu, S. Avidan, M. C. Yeh, and K. T. Cheng, "Fast Human Detection Using a Cascade of Histograms of Oriented Gradients," *IEEE Conference on Computer Vision and Pattern Recognition*, 2006.
- [11] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [12] A. Mohan, C. Papageorgiou, and T. Poggio, Member,, "Example-Based Object Detection in Images by Components," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 349-361, 2001.

-
- [13] D. Ramanan, D. A. Forsyth, and A. Zisserman, "Tracking People by Learning Their Appearance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, pp. 65—81, 2007.
- [14] B. Wu and R. Revatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Based Part Detector," *International Journal of Computer Vision*, 2007.
- [15] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian Detection in Crowded Scenes," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 878—885, 2005.
- [16] P. Viola, M. J. Jones, and D. Snow, "Detecting Pedestrians Using Patterns of Motion and Appearance," *IEEE International Conference on Computer Vision*, 2003.
- [17] J. J. Little and J. E. Boyd, "Recognizing People by Their Gait: The Shape of Motion," *Journal of Computer Vision Research*, pp. 1—32, 1998.
- [18] B. Heisele and C. Wohler, "Motion-Based Recognition of Pedestrians," *IEEE International Conference on Pattern Recognition*, pp. 1325—1330, 1998.
- [19] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic Extraction and Description of Human Gait Models for Recognition Purposes," *Computer Vision and Image Understanding*, vol. 90, pp. 1—41, 2003.
- [20] S. A. Niyogi and E. H. Adelson, "Analyzing and Recognizing Walking Figures inXYT," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 469—474, 1994.
- [21] Y. Ran, Q. Zheng, I. Weiss, L. S. Davis, W. A. Almageed, and L. Zhao, "Pedestrian Classification From Moving Platforms Using Cyclic Motion Pattern," *IEEE International Conference on Image Processing*, vol. 2, pp. 854—857, 2005.
- [22] Y. Ran, I. Weiss, Q. ZHeng, and L. S. Davis, "Pedestrian Detection via Periodic Motion Analysis," *International Journal of Computer Vision*, vol. 71, no. 2, pp. 143—160, 2007.
- [23] D. M. Gavrila, "The Visual Analysis of Human Movement: A Survey," *Computer Vision and Image Understanding*, vol. 73, no. 1, pp. 82—89, 1999.
- [24] J. K. Aggarwal and Q. Cai, "Human Motion Analysis: A Review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428—440, 1999.
- [25] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *Computer Vision and Image Understanding*, vol. 83, pp. 231—268, 2001.
- [26] J. O'Rourke and N. Badler, "Model-Based Image Analysis of Human Motion Using Constraint Propagation," *IEEE Trans. Pattern Anal. Mach. Intell.* 2(6), 1980, 522—536.

-
- [27] M. C. Mozer, "The Neural Network House: An Environment that Adapts to Its Inhabitants," *Proceedings of the American Association for Artificial Intelligence Spring Symposium on Intelligent Environment*, pp. 110—114, 1998.
- [28] I. A. Essa, "Ubiquitous Sensing for Smart and Aware Environment: Technologies Towards the Building of an Aware Home," *DARP/NSF/NIST Workshop on Smart Environment*, 1999.
- [29] S. Intille and K. Larson, J. S. Beaudin, J. Nawyn, E. M. Tapia, and P. Kaushik, "A Living Laboratory for the Design of Ubiquitous Computing Technologies," *Conference on Human Factors in Computing Systems*, 2005.
- [30] S. S. Intille and K. Larson, "Designing and Evaluating Supportive Technology for Homes," *Proceedings of the IEEE/ASME Intl. Conf. on Advanced Intelligent Mechatronics*, 2003.
- [31] G. M. Youngblood, L. B. Holder, and D. J. Cook, "Managing Adaptive Versatile Environments," *IEEE Conference on Pervasive Computing and Communication*, pp.351—360, 2005.
- [32] Hara, K.; Omori, T.; Ueno, R. "Detection of Unusual Human Behavior in Intelligent House" *Neural Networks for Signal Processing*, 2002.
- [33] Noguchi, H.; Mori, T.; Sato, T.; "Construction of Data Accumulation System for Human Behavior Information in room" *Intelligent Robots and System*, 2002.
- [34] Mori, T.; Sato, T.; Asaki, K.; Yoshimoto, Y.; Kishimoto, Y.; "One-room-type sensing system for recognition and accumulation of human behavior" *Intelligent Robots and Systems*, 2000
- [35] Mori, T.; Asaki, K.; Noguchi, H.; Sato, T.; "Accumulation and summarization of human daily action data in one-room-type sensing system" *Intelligent Robots and Systems, 2001. Proceedings*.
- [36] Matsuoka, K.; "Measuring behavior in a house" *SICE 2002. Proceedings of the 41st SICE Annual Conference*, 2000.
- [37] Appenzeller, G.; Joo-Ho Lee; Hashimoto, H.; "Building topological maps by looking at people: an example of cooperation between intelligent spaces and robots", *Proceedings of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1997.
- [38] Sato, T.; "Robotic room: Human behavior measurement, behavior accumulation and personal/behavioral adaptation by intelligent environment", *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 2003.
- [39] Yoshihama, S.; Chou, P.; Wong, D.; "Managing behavior of intelligent environments", *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, 2003.

-
- [40] Le Gal, C.; Martin, J.; Lux, A.; Crowley, J.L.; “SmartOffice: design of an intelligent environment”, *Intelligent Systems, IEEE* , Volume: 16 Issue: 4 , July-Aug. 2001.
- [41] Krumm, J.; Harris, S.; Meyers, B.; Brumitt, B.; Hale, M.; Shafer, S.; “Multi-camera multi-person tracking for EasyLiving”, *Visual Surveillance*, 2000. *Proceedings. Third IEEE International Workshop on* , 1 July 2000
- [42] Hedberg, S.R.; “After desktop computing: a progress report on smart environments research”, *Intelligent Systems, IEEE [see also IEEE Expert]* , Volume: 15 Issue: 5 , Sept.-Oct. 2000
- [43] Trivedi, M.; Huang Kohsia; Mikic, I.; “Intelligent environments and active camera networks”, *Systems, Man, and Cybernetics*, 2000 *IEEE International Conference on* , Volume: 2 , 8-11 Oct. 2000.
- [44] Mikic, I.; Huang, K.; Trivedi, M.; “Activity monitoring and summarization for an intelligent meeting room”, *Human Motion*, 2000. *Proceedings. Workshop on* , 7-8 Dec. 2000
- [45] Kynan Eng; Bdbler, A.; Bernardet, U.; Blanchard, M.; Costa, M.; Delbrilck, T.; Douglas, R.J.; Hepp, K.; Klein, D.; Manzolli, J.; Mintz, M.; Roth, F.; Rutishauser, U.; Wasserman, K.; Whatley, A.M.; Wittmann, A.; Wyss, R.; Verschure, P.F.M., “Ada - intelligent space: an artificial creature for the swiss Expo.02”, *Robotics and Automation*, 2003. *Proceedings. ICRA '03. IEEE International Conference on* , Volume: 3 , Sept.14-19, 2003.
- [46] Mozer, M.C.; “An Intelligent Environment Must Be Adaptive”, *Intelligent Systems, IEEE [see also IEEE Expert]* , Volume: 14 Issue: 2 , March-April 1999.
- [47] Z. Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond," <http://soma.crl.mcmaster.ca/homepage.htm>, 2003.
- [48] C. Rasmussen, "Joint Likelihood Methods for Mitigating Visual Tracking Disturbances," *IEEE International Conference on Computer Vision and Pattern Recognition*, 2001.
- [49] P. F. Fezenszwalb, “Pictorial Structures for Object Recognition”, *International Journal of Computer Vision*, 61(1):55—79, 2005.
- [50] A. Agarwal and B. Triggs, “3D Human Pose Estimation from Silhouettes by Relevance Vector Regression”, *IEEE Computer Society Conference on CVPR*, 2:882—888, 2004.
- [51] Shih-Shinh Huang, Li-Chen Fu, and Pei-Yung Hsiao, “Region-Level Motion-Based Background Modeling and Subtraction Using MRFs,” *IEEE Trans. on Image Processing (SCI)* Vol. 16, No. 5, pp. 1446-1456, May 2007.
- [52] Shih-Shinh Huang, Li-Chen Fu, and Pei-Yung Hsiao, “A Bayesian Network for Foreground Segmentation in Region Level,” *Asian Conference on Computer Vision*, 2007.

-
- [53] Shih-Shinh Huang, Cheng-Hsiung Chuang, Li-Chen Fu, and Pei-Yung Hsiao, "Region-Level Foreground Segmentation Using Bayesian Network-An Application to Human Detection," *Computer Vision and Graphics and Image Processing*, pp. 180-185, Miaoli, Taiwan, 2007.
- [54] Wen-Hau Liao, Chao-Lin Wu, and , Li-Chen Fu, "Inhabitants Tracking System in a Cluttered Home Environment via Floor Load Sensors," *Accepted, IEEE Transactions on Automation Science and Engineering*, (SCI) 2007.
- [55] Chao-Lin Wu, C.-F. Liao, and Li-Chen Fu, "Service-Oriented Smart Home Architecture based on OSGi and Mobile Agent Technology", *IEEE Transactions on Systems, Man, and Cybernetics--Part C : Applications and Reviews*, Vol. 37, No. 2, pp. 193-205 (SCI), 2007.