

行政院國家科學委員會專題研究計畫 期中進度報告

智慧型居家看護影像監控系統(1/3)

計畫類別：個別型計畫

計畫編號：NSC93-2218-E-002-104-

執行期間：93年08月01日至94年07月31日

執行單位：國立臺灣大學電機工程學系暨研究所

計畫主持人：陳永耀

計畫參與人員：李冠德、邱保盛、柯怡賢、許國基、李乾丞、黃正聖、劉建玟、
蘇義盛

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 94 年 6 月 1 日

行政院國家科學委員會專題研究計畫成果報告

智慧型居家看護影像監控系統(1/3)

計畫編號: NSC93-2218-E-002-104

執行日期: 93 年 8 月 1 日至 94 年 7 月 31 日

主持人: 陳永耀

台大電機系教授

計畫參與人員: 李冠德、邱保盛、柯怡賢、許國基、李乾丞、黃正聖、劉建玫

一、摘要

中文摘要

本計畫主要目標為建構一套「智慧型影像居家看護系統」，當家中的個人發生危險時，可以經由攝影機畫面的取得以及智慧型人體動作辨識軟體的操作，發出警訊至醫院或是通知正在工作的其他家人，避免延遲送醫的憾事發生。有關於這一方面探討的文章很多，發展也很迅速。本年度計畫，首先探討人體動作之智慧型辨識方法，主要先利用由動畫軟體取得的人體輪廓以及關節角度值，探討以人體特定關節角度值判斷人體動作的智慧型方法。目前計畫成果已完成可以辨識特定人體動作之辨識系統，例如正在跑步或是正在走路，此外利用姿勢轉換的辨識結果，例如由站姿很快的變成躺姿且一直維持躺姿，可以判斷出危險狀況的發生。

Abstract

The primary goal of this project is on the development of intelligent home care system. Human action identifications enable us to identify human actions so that proper warnings can be issued when dangerous situations occur, especially in domestic environment. The goal of the project in the first year is focused on the intelligent human action identifications

based on the given human contour and human angle parameters. Graphical software "Life Forms" is firstly adopted to initially provide the human joint angles for the human action identifications. Fuzzy action classifications and rule-base for identification are designed and constructed such that actions such as standing, walking, running, sitting, and lying can be identified. The change of actions from standing to lying indicates the possibility of dangerous situations and special warning is triggered.

二、計畫緣由及目的

多年來人們一直嘗試於製造智慧型系統以增加生產力及改善生活。這種智慧型系統必須具備有感知人類活動與提供自然回饋方式的能力。由於電腦視覺是一種非接觸性的感測方式，比起其他可用來感測人類活動的方法，電腦視覺對於人類而言是一種更自然、更便利的感測方式。因此發展和研究電腦視覺技術來感知人類的活動是必需且重要的。本計畫之目的即在於利用影像視覺辨識人體姿勢，歸納人體行為，並以 ARM 為發展平台，架構嵌入式之即時影像辨識系統。

本計畫規劃分成影像處理、人體動作與行為辨識、及硬體實現三個主題分

年進行研究。影像處理方面，主要的目的為研究已擷取到的影像資訊建立三維人體模型(Human Model)，藉由此人體模型上的各種參數來代表受測者肢體、關節等各部分位置、方向、角度等運動參數。將這些參數交由本計劃之智慧型行為辨識系統來做為辨識受測者的姿勢、行為的依據。

在硬體實現上，主要目的在於建立影像處理的硬體平台，為影像辨別方面建立起完整的系統架構，最後希望完成的方式是以嵌入式系統架構為主，在此系統架構之下，可將影像處理的演算法則以硬體的方式呈現，並能解決在硬體設計方面所會面臨的問題，預期系統能夠將本主題所描述的設計流程，如人體動作之辨別與影像擷取的功能實現在所架構的平台之上。

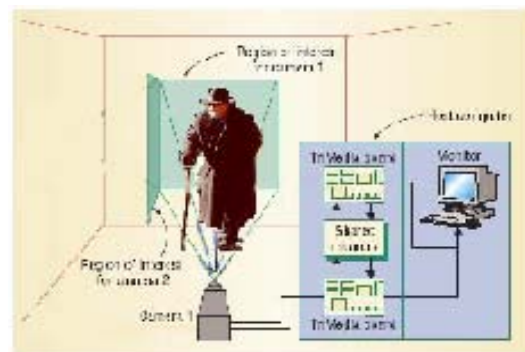
三、居家看護以及人體動作辨識之文獻探討

以往的居家看護，往往是請一個傭人在家裡照料。但是隨著時代的進步，人力的使用也越來越不符合時代潮流，再配合上電腦處理速度越來越快速，因此越來越多人朝著利用影像處理的技術來建構居家看護的系統。

一般而言，在不同的場景中，人眼都有辦法輕易且清楚分辨出背景與人體，也可以很準確的判斷出人體的動作是站著或是坐著，不可能會把坐著的人誤認為一張椅子。甚至於他在喝水或是喝不是透明飲料都可以很輕易的分辨出來。至於動態的行為，例如跑步，也絕對不會誤認為是在走路。

人的影像辨識行為非常複雜，人在判斷行為的過程當中，大腦的認知也佔了一個很必要的因素。而且人類也會學

習，所以並不是單單的因為人有眼睛就可以判斷週遭的動作行為。因此，就有人開始著手研究人類如何用眼睛感知週遭的事物以及如何判斷人體的動作姿態[7]。想藉由一些目前可以達成的技術再配合軟體，來達成判斷人體姿態的目的[10]、[11]，進而可以運用在居家看護上面。最方便取得的硬體且可以用來代替眼睛的工具，就是 CCD 攝影機，再藉由影像處理方法可以分離背景與人體[1]、[2]、[6]。並且建立人體 2D 或是 3D 模型[3]、[4]、[5]、[9]，再將這些取出來的輪廓與建立的模型做比對[12]、[13]，經由比對的結果希望可以達到辨識人體動作的功能[6]、[8]、[14]、[15]。



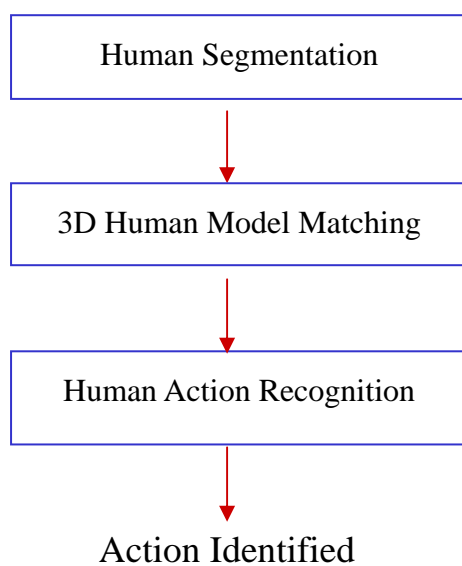
圖一 人體動作辨識系統

四、研究內容

(一) 人體動作辨識之基本概念

一般判斷人體動作的流程圖如圖二所示。首先要先把從複雜的背景中分離出有關人體的資訊，有許多不同的做法，最常被使用的方法稱為背景註冊法 (Background Registration)，使用於靜止的攝影機，意即靜止的場景。利用兩張背景相同的圖片，將會動的物體分離出來，此時會再搭配橢圓形找頭法，判斷所分離出來的物體是否為人體。若是分

離出來的物體在未來一段時間之內都靜止不動，則在這段時間之後便會把這件物體也視為背景，意即背景會一直被更新。接著利用自建的 3D 模型與分離出來的物體做匹配 (3D Human Model Matching)，而常用的人體模型又可分為圓柱狀以及棍棒狀，從人體模型所得到的參數，用來做為辨識人體動作，(Human Activity Recognition) 之用。

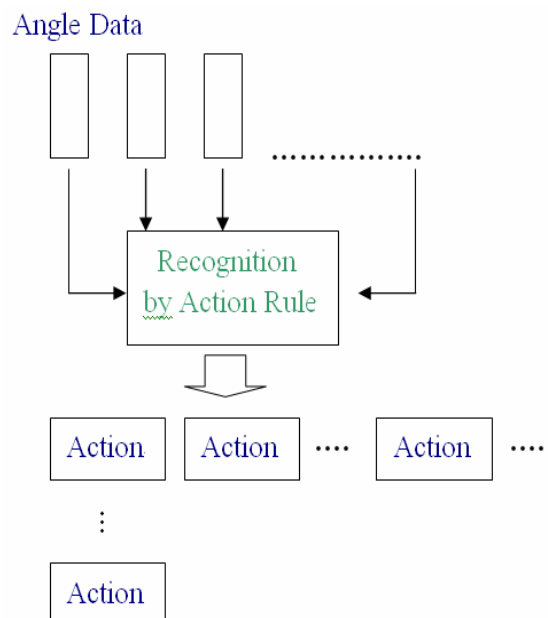


圖二 居家看護整體流程圖

本計畫在本年度首先進行人體動作辨識方面之探討。在辨識人體動作這一方面也有很多相關的討論，最多人使用的方法稱為隱藏式馬可夫陣列[16]，是屬於一種機率的方法。也有人使用基因演算法來辨識姿態[6]。但是本計劃採用不同的方法來辨識人體動作與姿態[17][18]。由於必須能夠先取得人體動作之相關參數，本計劃先使用「Life Forms」軟體，以這套軟體模擬 3D 人體模型並取得 3D 模型的特定關節角度值，之後再將這些關節角度值轉換成 BVH [19](Biovision Hierarchical) 檔案。BVH 其實只是那些

關節角度的列表，只是要利用 BVH 檔案來當作「Life Forms」軟體的輸入值。

至於為何要利用關節角度值當作辨別人體姿態的依據呢？由於當人眼辨識動作時，通常是依據動作的型態來做辨識，例如：走路是觀察到一個人在慢速移動而且雙腳會作前後的交替規律運動。這些規則就可以做為判斷動作的依據。由於動作常常會有其模糊的地帶，因此本計劃採用模糊理論做為當動作出現模糊地帶時的判斷依據，而選擇權重最大的那個姿勢作為最終的姿勢。動作判斷流程如圖三所示：



圖三 動作判斷流程圖

當取得角度值之後，將資料傳送到模糊判斷機制，利用模糊規則辨識輸入資料是否符合動作判斷規則，一旦符合，系統即會輸出所判斷之動作名稱。

(二) 人體動作辨識之設計

通常人類在判斷動作時，並不是由於腳彎曲幾度，或是手抬高多少來做判斷。而是依據動作的型態來判斷動作。

根據這個想法，本計畫設計了一套辨識動作的方法。

人體動作型態之辨識

通常人類在判斷人體動作時只會注意到人體上某些重要的特徵，並依據這些特徵來做判斷。例如：如果有人人在牆壁後面只露出頭並且頭一直在往前移動，則我們會覺得他在走路。這些特徵才是人類真正用來判斷動作的準則。

一般而言，一個動作不會只有單一特徵，而是由許多特徵共同組成。例如：走路時，雙腳會做前後規律的擺動，並且移動速度不快，身體擺幅不大，雙手也會做前後規律的擺動，這些都是用來形容走路時候的特徵。甚至於動作與動作間也會存些一些模糊地帶，例如一個人從走路然後到跑步，這之間會有一些動作很難歸類是走路還是跑步。

當人類要判斷走路還是跑步時，會認為這兩個動作都有可能，此時會產生判斷的困難，此時人類利用其他的特徵加以輔助，例如腳的擺動的速度，或是身體移動的速度，此時，頭是否有左右搖擺就不會被用來作為判斷的特徵。因此，在判斷動作時，眾多特徵之間會有其重要的先後順序，這是影響最後判斷結果的重要因素。

人體動作之分類

為了防止動作之間的混雜與增進系統效率，對動作的分類是必要的。首先，把動作粗分為四大類，如圖四所示。

第一類動作稱為 Basic Actions，主要在描述人體整體的動作，例如站、坐、躺及走路、跑步等。

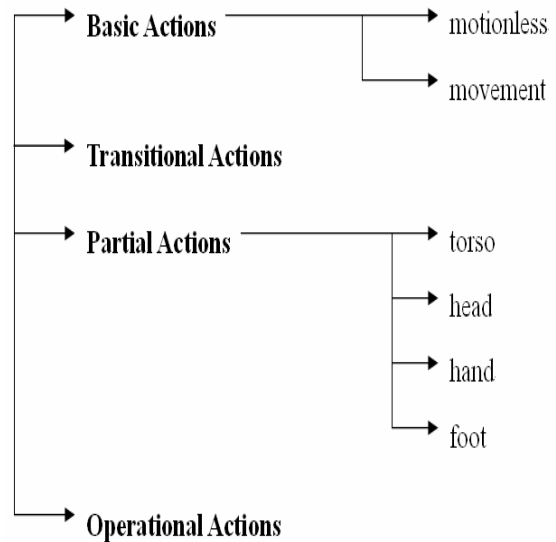
第二類動作稱為 Transitional Actions，只要是第一類動作之間的轉

換，通通會被歸成這一類。如：從坐著慢慢站起來，由於最後的動作為站著，因此會被歸類成“站起來”。

第三類動作稱為 Partial Actions，與第一類動作相反，是描述人體小幅度的動作，例如：手臂的動作、或是頭部的動作、或是腳部的動作都會被歸成這一類。

第四類動作稱為 Operational Actions，只要描述不是以上三類的動作。如：拿杯子喝水等等。

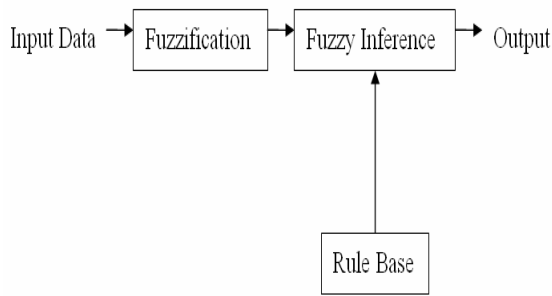
接著再把常見的動作一一歸類到這四大類動作中。如附錄一所示。本計劃目前僅以合理之範圍進行人體動作之分類，但並不嘗試涵蓋所有的動作。



圖四 人類動作的分類

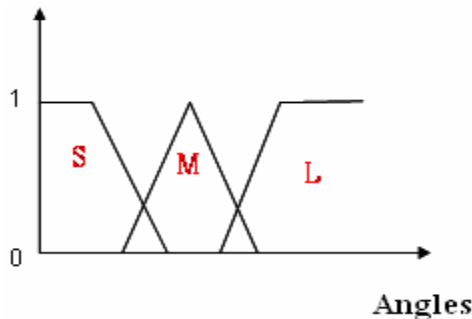
模糊動作之判定

模糊動作的判定主要以模糊規則庫的形式加以建構，單一人體動作可能以不同之權重值對應至不同的動作判定，因此權重是決定動作的重要參數。完整的架構圖如圖五所示。



圖五 模糊動作判斷的架構圖

在圖五的 **Fuzzification** 部分，主要將人體關節角度分為：角度大、角度中、角度小。如下圖所示：



圖六 模糊動作判斷的歸屬函數(membership function)

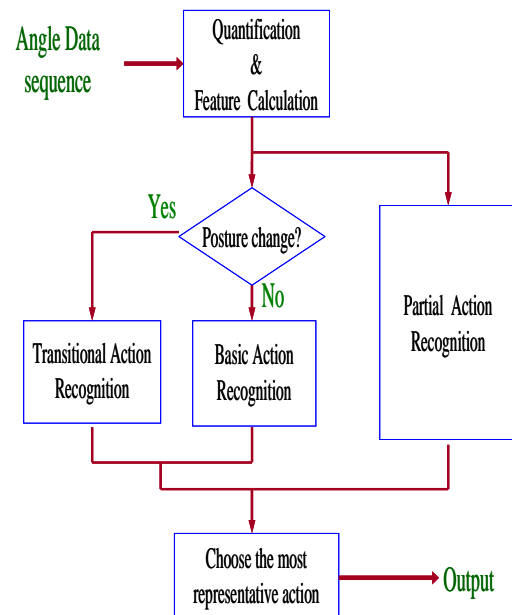
而在圖五之 Fuzzy Inference Engine 則選擇最小推論方法(Minimum Inference Engine)。至於 Rule Base，可依據附錄二的規則表來決定動作該屬於哪一條規則。

(三) 整體架構流程

經由上述的模糊規則與規則庫以及動作的定義之後，接下來要介紹整個系統的流程。圖七為系統的整體架構。

在取得關節角度值之後，先排除對於判斷沒有用處的角度，其餘角度則運用 membership function 來進行角度之描述，再將結果加上 Rule Base 裡的規則判斷動作。最後選擇最有可能的動作，即完成判斷。基本上每張影像進來可以馬上判斷出動作為何。至於前端 Angle

Data 的輸入，下面章節會再多加以說明。



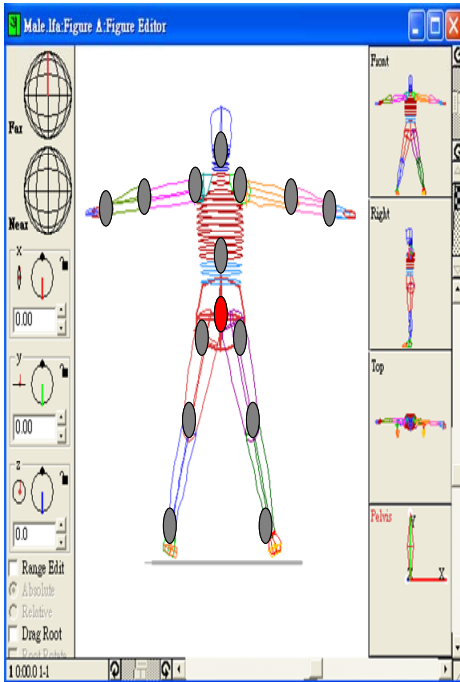
圖七 動作判斷之架構圖

五、 模擬及研究結果

(一) 模擬軟體

由於本計劃首先進行人體動作辨識規則之建立，而在影像處理以及人體參數之取得方面仍待下面年度之研究進行，尚未能經由比對人體 3D 模型而得到所需的關節角度。

因此本計劃目前使用前端模擬軟體「Life Forms」，以此軟體來獲得與輸入影片相同的模型，並且得到所需的關節角度值，作為人體動作辨識系統的輸入。圖八是軟體操作介面的範例。經由這套軟體，可以模擬出各種人體姿勢以及產生所需的關節角度。這是一套 3D 軟體，可以針對某個關節對 X 軸或 Y 軸或 Z 軸旋轉，並且可以得知旋轉完之後的角度值，因此很適合作為人體動作辨識系統的前端輸入。所需的關節以及其角度值如表一及表二所示



圖八 Life Forms 的人體模型

Basic angle	X	-180 ~ 180
	Y	-180 ~ 180
	Z	-180 ~ 180
Chest	X	-77 ~ 32
	Y	-10 ~ 11
	Z	-32 ~ 36
Head	X	-78 ~ 54
	Y	-70 ~ 70
	Z	35 ~ 35
Shoulder	X	-59 ~ 167
	Y	-52 ~ 54
	Z	-48 ~ 180
Elbow		0 ~ 152

表一 所需關節的角度值

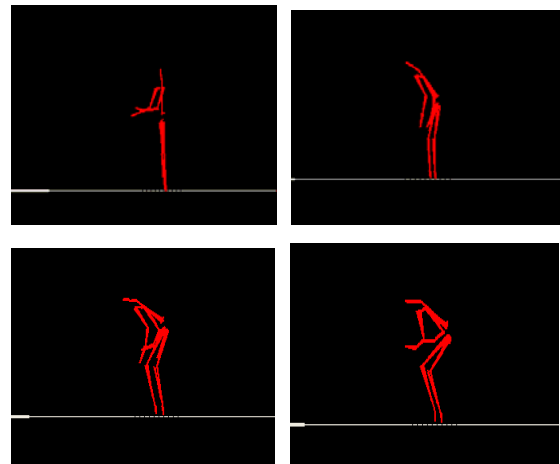
Hand	X	-90 ~ 90
	Y	-18 ~ 20
	Z	-61 ~ 61
Thigh	X	-31 ~ 120
	Y	-54 ~ 54
	Z	-45 ~ 54
Knee		-135 ~ 15
foot	X	-115 ~ 48
	Y	-21 ~ 22
	Z	-27 ~ 32

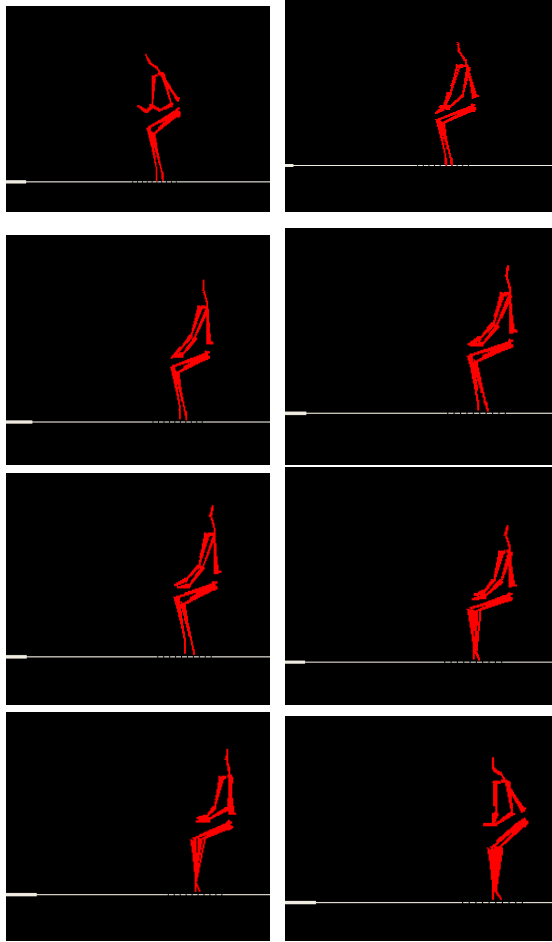
表二 所需關節的角度值

(二) 模擬結果

為測試目前完成之人體動作辨識系統效能，同時配合目前以 Life Forms 軟體取得人體關節角度參數之限制，以下採用一段影像當中的 4 個片段做為測試範例。同時，由於居家看護系統之主要目的在辨識是否有危險動作之存在，因此我們特別選取了以下的片段：一個人從走路到不小心跌倒。經過 Life Forms 繪圖之後的結果如圖九所示。

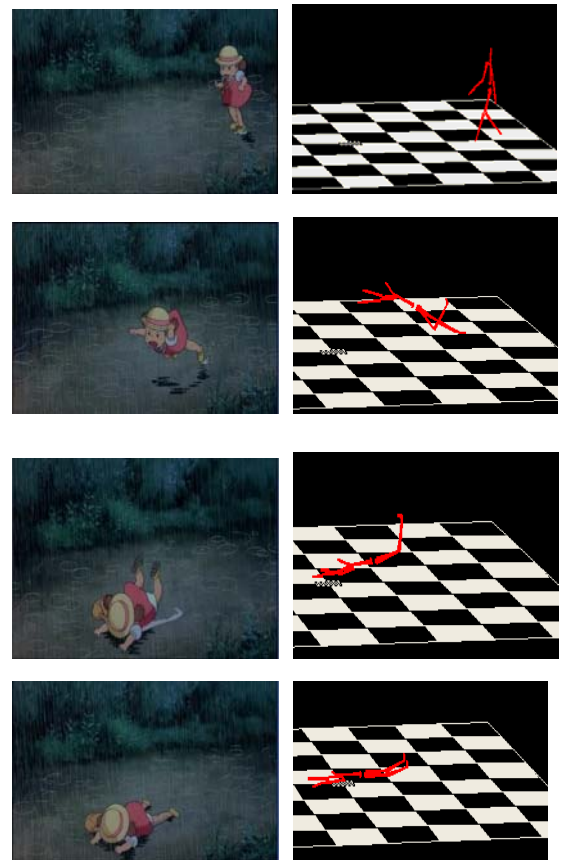
在經由 Life Forms 軟體建構出與影片人物相同動作的模型後，這些模型以人體動作辨識系統判斷其動作。模擬結果如圖九及表三所示。





圖九 模擬結果(一)

長的人更是如此。以下便針對「跌倒」做了模擬。如圖十及表四所示。



圖十 模擬結果(二)

Frame Number	Simulation Output	Frame Number	Simulation Output
1	standing	7	sitting
2	standing	8	sitting
3	standing	9	sitting
4	standing	10	sitting
5	sit down	11	sitting
6	sitting	12	stand up

表三 畫面數目以及輸出的動作名稱

圖九為模擬站起來並拉出椅子，然後坐下，最後動作為站起來。表三亦很精確的判斷出動作。

另一個則是針對「跌倒」做模擬。跌倒是很危險的動作，尤其對於年紀較

Frame Number	Simulation Output
1	standing
2	Standing & raising a hand
3	trip down
4	trip down

表四 畫面數目以及輸出的動作名稱

圖十為依據電影動畫所模擬之動作，表四為其所對應之動作判斷。在畫面 2 因為尚未達到絆倒的姿態，因此仍判斷為站立，且由於舉手的動作符合姿態辨別規則，因此予以判斷顯示。

六、 結論

本計畫目的為判斷出是否發生危險動作以及發生危險動作時應及時發出警告。針對本計畫前端人體特徵的取出，如人體輪廓取得、環境參數描述，雖然在此著墨不多，不過目前已經著手研究，亦有初步的了解，

至於取得人體相關資訊後，動作辨識的部份，由模擬結果顯示，對於日常生活中常見的動作，如走路突然跌倒，或是由站姿變成坐姿等等，已經可以根據人體特定關節角度值準確的判斷。同時對於人類動作的特徵以及人類判斷人體動作的程序亦有深入的了解。

七、 計劃成果自評

本計劃此年度完成了根據人體特定關節角度值進而判斷人體動作。若未來再加上前端發展出可以由複雜的背景中取出人體以及算出所需的關節角度值（不是利用軟體所得的關節角度值），便可以建構一套完整的系統。

不過，現在對於行為上的判斷仍有不足的地方，這方面需要得知週遭環境的參數，例如可以得知這是一罐農藥或是一瓶水。如此一來便可以更加完整的判斷出行為，例如知道人在喝水或是喝危險的液體。如此一來，便可以進一步預防危險的發生，而不是像現在只能被動的辨識危險是否已經發生，爭取寶貴的時間，此部分仍須多加努力。

本計劃未來預計完成的工作為，必須完成前端人體特徵的取出，如人體輪廓取出，以便未來可以直接由攝影機中擷取人體並完成動作判斷。

八、 參考文獻

- [1]. S. Y. Chien, S. Y. Ma, and L. G. Chen, "Efficient moving object segmentation algorithm using background registration technique," *IEEE Trans. On Circuits and Systems for Video Technology*, vol.12, no7, pp.577-586, July 2002.
- [2]. R. Jain and H. H. Nagel, "On the analysis of accumulative difference pictures from image sequences of real scenes".*IEEE Trans. On PAMI*, **1**(2):206-214, 1979.
- [3]. D. Hogg, "Model-based vision: A program to see a walking person", *Image and Vision Computing* **1**(1), 1983, 5-20.
- [4]. H. J. Lee and Z. Chen, "Determination of 3D human body posture from a single view", *Comp. Vision, Graphics, Image Process.***30** 1985, 148-168.
- [5]. M. K. Leung and Y. H. Yang, "First sight: A human body Outline labeling system", *Trans. Pattern Anal. Mach. Intelligence* **17**(4),1995,359-377
- [6]. C. Hu, Q. Tu, T. Li, and S. Ma, "Extraction of parameter human model for posture recognition using genetic algorithm", in *Proc. Fourth International Conference on Automatic Face and Gesture Recognition*, pp.518-523,2000.
- [7]. G. Johansson, "Visual perception of biological motion and a model for its analysis", *Perception Psychophys.* **14**(2), 1973, 201-211.
- [8]. N. Badler and C. Phillips, and B. Webber, "Simulating humans", Oxford Univ. Press, Oxford, 1993.
- [9]. J. Wang, G. Lorette, and P. Bouthemy, "Analysis of human motion: a model-based approach", in *Scandinavian Conference on Image Analysis*, 1991.
- [10]. D. M. Gavrilu, "The visual analysis of human movement:A survey, *Computer Vision and Image Understanding*, **73**(1), pp.82-98,1999.
- [11]. J. K. Aggarwal and Q. Cai, "Human motion analysis: a review", *Computer Vision and Image Understanding* **73**(3), 1999
- [12]. K. Rohr, "Human Movement Analysis

Based on Explicit Motion Models”, chap8, pp. 171-198, Kluwer Academic, Dordrecht/Boston, 1997.

- [13]. Y. Guo, G. Xu, and S. Tsuji, “Tracking human body motion based on a stick figure model”, *J. Visual Comm. Image Representation* **5**, 1994, 1-9.
- [14]. C. R. Wren and A. P.Pentland, “Understanding purposeful human motion”, in *International Workshop on Modeling People at ICCV’99, Corfu, Greece, September 1999*.
- [15]. F.Lerasle, G. Rives, and M.Dhome, “Tracking of human limbs by multiocular vision”, *Comp. Vision Image Understanding* **75**, 1999, 299-246
- [16]. Jie Yang, Yangsheng Xu, and Chiou S. Chen. “Human action learning via hidden markov model”. *IEEE Trans. on Systems, Man and Cybernetics*, A:34-44, 1997
- [17]. Mori, T.; Tsujioka, K.; Sato, T. “Human-like action recognition system on whole body motion-capture file”. *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on Volume 4, 29 Oct.-3 Nov. 2001*
Page(s):2006-2073 vol.4
- [18]. Mori, T.; Tsujioka, K.; Shimosaka, M.; Sato, T. “Human-like action recognition system using features extracted by human”. *Intelligent Robots and System, 2002.IEEE/RSJ International Conference on Volume 2, 30 Sept.-5 Oct. 2002*
Page(s):1214-1220 vol.2
- [19]. <http://www.metamotion.com/>

附錄一：動作分類表之細分

Basic Actions	
motionless	movement
Standing	Walking
Sitting	Running
Lying	
Squatting	
Kneeling	

Transitional Actions	
Stand up	
Sit down	
Lie down	
Fall down	Slip down
	Trip down
Kneel down	
Squat down	
Get up	

Partial Actions	
hand	head
Raising a hand	Shaking head

附錄二: Rule Base

Lie(posture)		Kneel(posture)	
1	Body is approximately horizontal	1	Body is approximately vertical
Stand(posture)		2	Feet don't bend simultaneously
1	Body is approximately vertical	3	Knees bend simultaneously
2	Feet don't bend simultaneously	Squat(posture)	
3	Knee don't bend simultaneously	1	Body is approximately vertical
Sit(posture)		2	Feet bend simultaneously
1	Body is approximately vertical	3	Knees bend greatly simultaneously
2	Feet bends simultaneously	Fall down	
3	Knees don't bend greatly simultaneously	1	Other posture to lie
Walking		2	Velocity is fast
1	Posture is stand	Trip down	
2	Feet back and forth	1	fall down
3	Velocity isn't very fast	2	knees bend > threshold
Running		Slip down	
1	Posture is stand	1	fall down
2	Feet back and forth	2	knees don't bend
3	Velocity is fast	Raising a hand	
Shaking head			The Hand is height
1	Neck rotated from side to side		A hand was waved up
Lie down		Get up	
1	Other posture to lie	1	From kneel to stand
2	Velocity is slow	2	From squat to stand
Standing		3	From lie to stand, sit, kneel and squat
1	Stand posture keep one second	Sitting	
2	Feet isn't back and forth		Sit posture keep one second
lying		Kneeling	
1	Lie posture keep one second	1	Kneel posture keep one second
Squatting		Stand up	
1	Squat posture keep one second	1	From sit to stand
Sit down		Kneel down	
1	From stand to sit	1	From stand or sit to kneel
Squat down			
1	From stand or sit to squat		