

# 國立交通大學

多媒體工程研究所

碩士論文

基於影像掃描之人體姿態識別系統



Human Posture Recognition System Based on Image

Scanning

研究生：黃聖中

指導教授：莊仁輝 教授

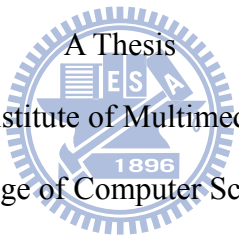
中華民國九十八年六月

基於影像掃描之人體姿態識別系統  
Human Posture Recognition System Based on Image Scanning

研究生：黃聖中                      Student：Sheng-Chung Hunag

指導教授：莊仁輝                    Advisor：Jen-Hui Chuang

國立交通大學  
多媒體工程研究所  
碩士論文



A Thesis  
Submitted to Institute of Multimedia Engineering  
College of Computer Science  
National Chiao Tung University  
in partial Fulfillment of the Requirements  
for the Degree of  
Master  
in  
Computer Science

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

# 基於影像掃描之人體姿態識別系統

學生：黃聖中

指導教授：莊仁輝 博士

國立交通大學

多媒體工程研究所碩士班

## 摘 要

近年來，電腦輔助之人類行為認知應用越來越廣泛，例如行為分析可對於異常事件(如跌倒)進行偵測與預警；姿勢分類和姿態識別則常見於人機互動，利用不同的姿勢或姿態來表示不同的命令或是指示；活動表示則用來代表人體之活動，同時也可以人體以卡通化之方式做一呈現。但是多數之研究常需使用昂貴之設備，或是要求被觀測者使用較不方便的穿戴式感應器，使得成本偏高或是方便性不足。為了改善上述情形且加上即時性及隱私性之考量，本論文利用電腦視覺的方法，透過監視攝影機來建立一個注重隱私且具有較少資料量及運算量的簡化人體模型（僅識別人體影像之突出部分）。本研究成果期望提供給姿勢分類或是行為分析之研究者做更細部動作之分析。

# Human Posture Recognition System Based on Image Scanning

Student : Sheng-Chung Huang      Advisor : Dr. Jen-Hui Chuang

Institute of Multimedia Engineering

National Chiao Tung University

## Abstract



In recent year, computer-aided human behavior understanding becomes important in a variety of applications. For example, behavior analysis can assist the detection and warning of a falling event. Posture classification and posture estimation can be used to generate commands in human-machine interactions. Human activity representation helps animate human actions. Most research efforts use expensive devices or wearable sensors, which may increase the cost and making people uncomfortable. In this paper, we propose a surveillance system based on computer vision to establish human gesture representations which can preserve personal privacy efficiently in terms of data size and computation cost. To that end, an approach to automatic recognition of rough human posture, through identification of extruding parts of human body, is developed. The identification results can serve as a base for future research works in behavior and posture analysis which need to analyze more detailed human actions.

# 致 謝

首先，由衷地感謝我的指導教授莊仁輝博士，如果不是老師的辛勤指導，也無法讓我這個不知影像處理為何物的懵懂之人，對此領域產生興趣及熱誠。也因為有老師的教誨，才讓我可以完成此篇研究。

同時感謝三位口試委員楊谷洋教授、雷欽隆教授、賴飛龍教授，因為有三位口試委員的意見及指教，讓我的研究更趨於完善，也讓我對電腦視覺這個領域有更進一步的認知。

感謝實驗室的學長羅國華的指教，因為有跟學長討論才能讓我的研究進度不致於落後太多，同時感謝實驗室的夥伴們永昌、尚一、宣良、邦展、世旻，要不是有你們的一路扶持、一起歡笑、一起共患難，也不會造就今天的我，這兩年在交大的回憶也因為有你們而顯得燦爛異常。也感謝我的女朋友小奴，有妳的一路扶持，陪伴我這兩年的辛勞，有過許多歡笑及淚水更顯得彼此的珍貴。

最後要感謝的就是我的家人們，讓我在疲憊之餘有一個悠閒的避風港，因為有你們的辛苦栽培及關愛，才会有現今的我，感謝大家。

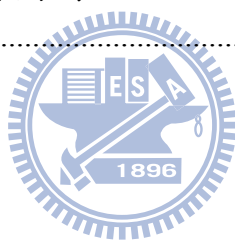
# 目 錄

中文摘要.....	i
英文摘要.....	ii
致謝.....	iii
目錄.....	iv
圖目錄.....	v
表目錄.....	vii
一、緒論.....	1
1.1 研究動機.....	1
1.2 相關研究.....	1
1.3 論文架構.....	7
二、人體區域偵測.....	10
2.1 可見光影像人體區域擷取.....	10
2.2 中紅外線影像人體區域擷取.....	12
2.3 消除雜訊及填補人體中破碎之區域.....	12
三、基於掃描線演算法之肢體近似中心軸線分析.....	14
3.1 掃描線演算法.....	14
3.2 掃描線橫截面之近似中軸線段.....	19
3.3 標示頭及肢體之末端.....	21
四、基於頭、手、腳定位的姿態識別.....	24
4.1 整合近似中軸線段.....	24
4.1.1 夾角的角度容許範圍之定義及選取.....	28
4.1.2 兩近似中軸線段是否面積重疊之判斷.....	29
4.1.3 多節肢體中肢體末端之選取.....	32
4.2 基於肢體末端數目的頭、手、腳定位.....	33
4.2.1 基於 5 個肢體末端條件下的作法.....	38
4.2.2 基於 4 個肢體末端條件下的作法.....	39
4.2.3 基於 3 個肢體末端條件下的作法.....	41
4.2.4 基於 2、1 個肢體末端條件下的作法.....	44
4.3 基於頭、手、腳定位之人體模型表示.....	45
五、實驗結果及討論.....	47
六、結論及未來工作.....	60
參考文獻.....	61
附錄 A、連續橫截面的中心點是否相連.....	63
附錄 B、判斷線段是否相交.....	65

# 圖目錄

圖 1-1 從人體區域中擷取水平及垂直的投影直方圖的示意圖.....	3
圖 1-2 人體區域做三角化的示意圖.....	4
圖 1-3 人體區域的外接矩形依高度之比例分做頭、手、腳三個區域的示意圖.....	5
圖 1-4 依佔人體區域外接矩形的高度比例分做頭、身體、腳三個區域的示意圖.....	5
圖 1-5 人體區域輪廓圖及輪廓和質心距離的曲線圖.....	6
圖 1-6 系統流程圖.....	9
圖 2-1 人體區域偵測流程圖.....	10
圖 2-2 可見光影像經由前景偵測找出之人體區域.....	12
圖 2-3 可見光影像經由前景偵測找出之人體區域.....	13
圖 2-4 中紅外線影像經由前景偵測找出之人體區域.....	13
圖 3-1 擷取肢體近似中心軸線流程圖.....	14
圖 3-2 以四個不同方向的掃描線掃描人體區域後所得之橫截面，圖中每條線段各代表一個橫截面.....	15
圖 3-3 將圖 3-2 中寬度過短之橫截面濾除的結果.....	16
圖 3-4 空洞造成橫截面擷取不佳的情形.....	17
圖 3-5 掃描線演算法中的連續兩橫截面關係圖.....	17
圖 3-6 改良掃描線演算法所得之橫截面.....	18
圖 3-7 連續橫截面中心點連線與橫截面之垂直方向的夾角的關係圖.....	19
圖 3-8 四個方向掃描線的近似中軸線段.....	20
圖 3-9 移除圖 3-8 中過寬或過短的近似中軸線段.....	21
圖 3-10 圖 3-9 右腳部份及整合後的近似中軸線段.....	22
圖 3-11 找近似中軸線段末端方法的示意圖.....	22
圖 3-12 四個方向掃描線可以代表肢體末端的近似中軸線段.....	23
圖 4-1 姿態識別流程圖.....	24
圖 4-2 頭部之橫截面分布及含末端特性的近似中軸線.....	25
圖 4-3 足以代表頭部的所有近似中軸線段.....	26
圖 4-4 近似中軸線段的整合.....	27
圖 4-5 近似中軸線段與橫截面垂直方向的夾角的關係圖.....	28
圖 4-6 同一腳部區域之近似中軸線段.....	29
圖 4-7 利用梯形代表肢體部位的形狀.....	29
圖 4-8 兩梯形重疊的特殊例子.....	30
圖 4-9 兩梯形與外接矩形之關係.....	32
圖 4-10 多節肢體中末端選取之方式.....	33
圖 4-11 從近似中軸線段找到頭、手、腳定位點的例子.....	33
圖 4-12 近似中軸線段的資訊.....	34

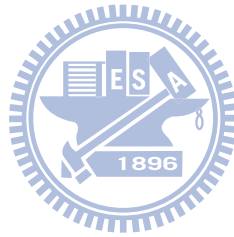
圖 4-13 近似中軸線段的肢體部位資訊.....	35
圖 4-14 兩隻腳合併成一隻腳的角度大於頭部的例子.....	38
圖 4-15 5 個肢體的識別結果.....	39
圖 4-16 4 個肢體的識別結果.....	40
圖 4-17 3 個肢體的識別結果.....	43
圖 4-18 2 個肢體的識別結果.....	44
圖 4-19 人體模型表示結果.....	46
圖 5-1 中紅外線影像姿態識別結果.....	47
圖 5-2 可見光影像姿態識別結果.....	47
圖 5-3 人像走動並坐下的連續中紅外線影像及姿態識別後所產生的人體模型..	48
圖 5-4 人像呈大字形走動的連續可見光影像及姿態識別後所產生的人體模型..	52
圖 5-5 人像彎腰、蹲下及躺下的連續影像及姿態識別後所產生的人體模型.....	55
圖 6-1 連續影像的姿態識別結果.....	60
圖A-1 人體腰部至腿部的局部橫截面示意圖.....	63
圖A-2 人體頭部至軀幹的局部橫截面示意圖.....	64
圖 B-1 兩線段是否相交的例子.....	65
圖 B-2 點在線段的延伸線的哪個方向.....	66
圖 B-3 兩線段是否相交的情形.....	67





# 表目錄

表 4-1 姿態識別中所使用到的參數一覽.....	45
表 4-2 人體模型的隱密性及傳輸資料比較.....	46
表 5-1 人像走動並坐下的連續中紅外線影像的姿態識別正確率.....	51
表 5-2 人像呈大字形走動的連續可見光影像的姿態識別正確率.....	54
表 5-3 人像彎腰、蹲下及躺下的連續影像的姿態識別正確率.....	58
表 5-4 識別結果正確率.....	59
表 5-5 本論文與[11]、[12]差異之比較.....	59



# 一、緒論

## 1.1 研究動機

近年來，人類行為認知的應用越來越廣泛，例如在行為分析方面可藉由科技的輔助（安全監控系統、感應裝置）來偵測跌倒等異常事件，並在意外發生時發出警報，使得醫療人員及看護的負擔得以減輕，並有效地減少醫療人力成本之浪費，為日趨老化的社會[1]帶來極大的貢獻；姿勢分類和姿態識別則常見於人機互動，利用不同的姿勢或姿態來表示不同的命令或是指示，例如舉手表示開冷氣、躺著代表關燈等；活動表示則用來表示人體之活動，同時也可以將人體之活動以卡通化之方式做一呈現，藉以保護被觀測者之隱私。

但是多數之人類行為認知的研究常需使用昂貴之設備，或是要求被觀測者使用較不方便的穿戴式感應器，使得成本偏高或是方便性不足。同時，被觀測者通常不希望自己的真實影像被記錄和檢視，因此需注重保護被觀測者之隱私。為了達到成本低、保護隱私及方便性較高之需求，本論文利用電腦視覺的方法，透過監視攝影機擷取之影像來識別各個人體肢體部位，並儲存肢體部位之形狀與角度等資訊，藉以建立一個注重隱私且具有較少資料量及運算量之人體模型，並利用此人體模型來分析被觀測者的狀態，期望提供給姿勢分類或是行為分析之研究者做更深入之分析。

## 1.2 相關研究

在人類行為認知之相關研究中，有以下常見的主題：行為分析、姿勢分類、姿態識別、活動表示等。行為分析[2-4]是指利用若干方法來分辨被觀測者是屬於走路、跑步、坐下、躺下等正常行為或是跌倒、爬牆等異常行為。姿勢分類[5-7]則是利用分類器或是若干規則來分類出被觀測者是站、坐、躺、蹲下、彎曲等姿

勢中的哪一個。姿態識別[8-10]則是偵測被觀測者的頭、手、腳等位置並標示出來。活動表示[11-12]則是將被觀測者用模型之方式呈現，其目的在於保護被觀測者之隱私權。

行為分析有一種方法是擷取人體特徵並利用分類器來進行分析，如[2]是先利用背景註冊法（background registration）自圖片中切割出人體區域，再從人體區域中擷取水平及垂直的投影直方圖（projection histogram）。得到水平及垂直的投影直方圖後，再各自進行 DFT（Discrete Fourier Transform）轉換，得到水平及垂直的 DFT 參數，之後擷取各自扣除第 1 個以外的前 20 個 DFT 參數並作正規化，如圖 1-1 為從人體區域中擷取水平及垂直的投影直方圖的示意圖，右圖是水平方向的投影直方圖統計，下圖則是垂直方向的投影直方圖統計。接著將這 40 個參數以及人體區域之長寬比作為輸入，藉由 neural fuzzy network 分類器得到站立、彎曲、坐以及躺等四個姿勢，最後利用跌倒是在短時間內姿勢由站立變成彎曲最後變成躺的變化規則來判斷是否有跌倒。

另一種行為分析的方法是擷取人體特徵後，再以固定規則來進行判斷。如[3]是先將影像用 time/background frame difference 將前景分割出來後，接著利用 Bi-modal thresholding 將人體分割出來並用一個常態分布函數來代表人體，最後計算其標準差是否大於門檻值，意即以身體寬度是否過寬來判斷是否為跌倒。[4]的方法則是先將影像用影像相減法（frame difference）將移動區域切割出來，再去算出它的移動並建立 motion history image，接著利用背景減去法（background subtraction）將人體區域切割出來後，將人體以一個橢圓形逼近，並得到其偏移角度  $\theta$  及長短軸比  $\rho$ ，藉由移動是否過大及  $\theta$ 、 $\rho$  是否變化過大，以及一段時間內移動是否降低至門檻值以下，及  $\theta$ 、 $\rho$  變化是否趨於平緩等條件來做為是否跌倒之依據。

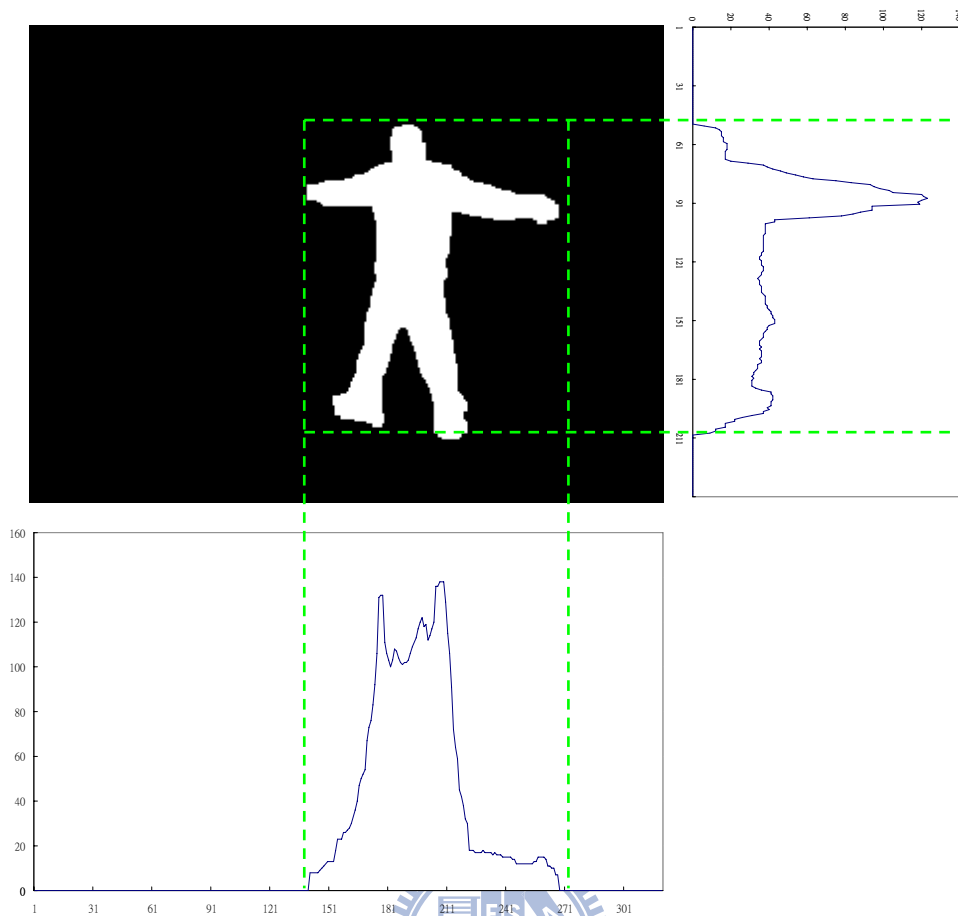


圖 1-1 從人體區域中擷取水平及垂直的投影直方圖的示意圖。

姿勢分類的研究中有一種是擷取人體特徵後，再以若干判斷方式來進行分類。如[5]則是使用背景減去法將前景切割出來後，利用侵蝕（erosion）的觀念配合  $3 \times 3$  矩陣，將人體區域以細化（thinning）的方式擷取出人體近似骨架，再分析每個肢體部位的骨架數量及狀態來識別人體的姿勢。

姿勢分類還有一種方法是擷取人體特徵資訊與資料庫的姿勢進行比對。如[6]是先用背景減去法將人體區域切割出來後，將人體區域做三角化並找出各個三角形之重心，再將各個重心予以連接，取得該人體姿勢的近似骨架，如圖 1-2 為一人體區域做三角化的示意圖，其中綠色部分為三角化後的三角形，紅色點代表三角形之重心，紅色連線代表人體近似骨架。接著使用 string matching 的方式，將該人體近似骨架與資料庫內的人體姿勢模型進行比對識別，一旦該姿勢與資料庫內之正常行為姿勢無一吻合，即判斷該姿勢為異常行為姿勢。而[7]則是先將 RGB

色彩空間轉至 AB 色彩空間後，接著用背景減去法將人體區域切割出來，再將人體區域做垂直方向的投影直方圖，最後與資料庫內預先建立的人體姿勢作比對，藉以分析人體的姿勢。

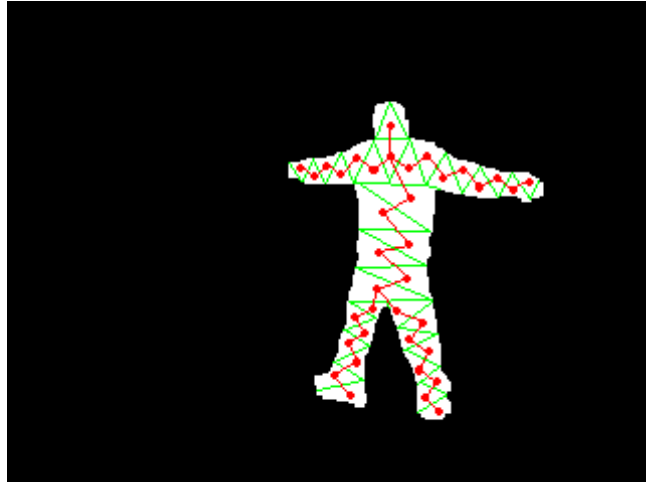


圖 1-2 人體區域做三角化的示意圖。

姿態識別的研究中，比較常見的是利用區域分割 (region splitting) 的方式來找頭、手、腳，如[8]的作法是先利用背景減去法將人體區域切割出來，接著找到人體區域的外接矩形 (bounding box)，將外接矩形依照高度之比例分為三個區域，如圖 1-3 所示，最下面的 33% 是腳部區域，腳部區域的上面 37% 是手部區域，最上面的 20% 是頭部區域。接著針對腳部區域，利用掃描線的方式尋找是否具有 FBF (foreground-background-foreground) 的特性，若有即代表具有兩隻腳，若無則代表兩隻腳合併成一隻腳，接著利用由下而上的方式找到面積佔腳部區域中前景面積的 25%，將該面積所佔之區域的質心當做腳點，最後利用 nearest point algorithm 的追蹤技巧來分類兩腳。而針對手之識別，則是在手部區域中找到具有膚色特徵即代表手，再將膚色區域以一個橢圓形逼近，取長軸與短軸的端點中位於最下面的點當做手點，而分類兩手的方式則是使用 bounding box overlapping algorithm 的追蹤技巧來完成。最後在頭部區域中，利用由上而下的方式找到面積佔頭部區域中前景面積的 66%，將該面積所佔之區域的質心當做頭點。[9]的作

法則是先利用 $W^4$  (Who, When, Where, What) 將人體區域擷取出來，接著利用Cardboard Model將人體區域的外接矩形分成頭、身體和腳三個部份，如圖 1-4 所示，外接矩形高度為 $h$ ，頭佔 $h/5$ ，身體和腳各佔剩餘部份的 $h/2$ ，其中身體和腳是允許局部重疊的。接下來針對各個部份，先計算該區域的前景平均寬度來做為頭、身體和腳的寬度，並計算其逼近橢圓的主軸來判定頭、身體和腳的呈現方向。而手則是經由身體定位後，找到連接身體但在身體外側的部份即代表手。最後追蹤各個部份來使其識別更加準確。

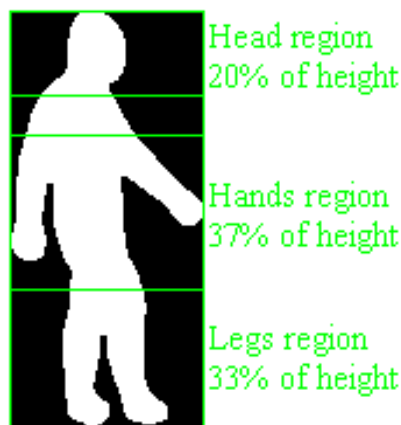


圖 1-3 人體區域的外接矩形依高度之比例分做頭、手、腳三個區域的示意圖。

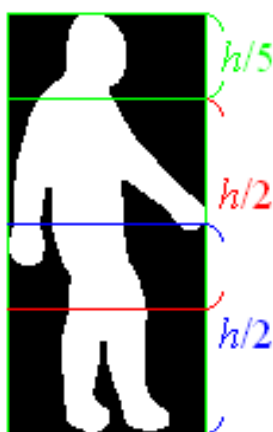


圖 1-4 依估人體區域外接矩形的高度比例分做頭、身體、腳三個區域的示意圖。

姿態識別的研究中還有一種是利用頭、手和腳在人體中較為顯著的特性來做識別的，如[10]的作法是先使用影像相減法來擷取人體區域並更新背景模型，當背景模型經過一段時間可供參考時，就使用背景註冊法來擷取人體區域。接著計算人體區域之質心與人體區域的輪廓各點的距離形成一條曲線，擷取此曲線中的區域最大值來當作特徵點並計算該夾角，藉此進一步分析頭、手和腳。如圖 1-5，左圖為人體區域之輪廓圖，五個框起來的點是右邊曲線圖中找到之區域最大值，右上角的數字代表該點夾角，右圖則是計算人體區域之質心與人體區域的輪廓各點的距離形成的曲線圖，紅點代表曲線中的區域最大值。接著將人體區域以一個橢圓逼近，並擷取長、短軸，從短軸上半部找度數在  $50^\circ$  到  $150^\circ$  之間的夾角，如果只有一個那就是頭，如果有兩個以上，則找距離長軸端點最近的當作頭，如果沒有一個符合，就將最靠近長軸端點的膚色區域當作頭，如果連膚色區域都沒有，那就不用長軸端點當作頭。接下來從短軸下半部開始找夾角度數在  $100^\circ$  以下的，如果有兩個以下那就是腳，如果超過兩個，就將距離頭最遠的兩個當作腳，如果沒有任何一個符合，就將長軸的端點當作腳。最後從短軸端點兩側找度數在  $100^\circ$  以下的夾角，如果有兩個以下那就是手，如果超過兩個，就將距離短軸端點最近的兩個當作手，如果沒有任何一個符合，則選擇最靠近短軸端點的膚色區域當作手，如果連膚色區域都沒有，那就不用短軸兩端點當作手。

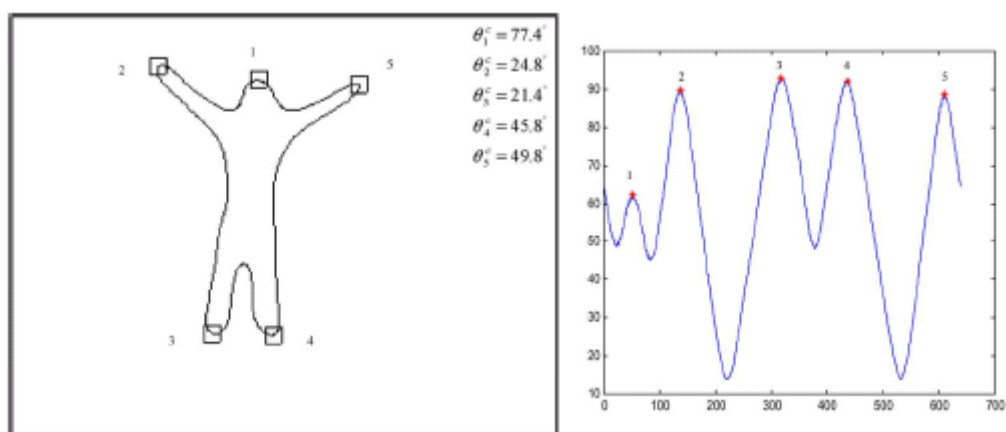


圖 1-5 人體區域輪廓圖及輪廓和質心距離的曲線圖（擷取自[10]）。



活動表示之研究中，[11]的作法是先將輸入之紅外線影像，用 thresholding 及等高線演算法擷取出人體區域，再利用  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$  和  $135^\circ$  的掃描線來擷取橫截面，接著利用適當的規則來連接橫截面之中心點，得到該肢體或軀幹的近似中軸線段。在排除過小或過寬之近似中軸線段後，將四個方向的近似中軸線段做整合，得到表示肢體及軀幹的若干個最佳的近似中軸線段，接著選出代表肢體末端的近似中軸線段。最後將代表肢體末端的近似中軸線段利用若干規則找出頭、手、腳，並利用橢圓及平行四邊形代表頭、手、腳來表示人體模型。而[12]的作法與[11]類似，不過輸入來源是可見光影像，因此它使用高斯混合模型 (Gaussian mixture models, GMM) 來擷取人體區域。[12]與[11]最大的差異是從四個方向的近似中軸線段做整合這個階段開始，[11]的方式是挑出代表該肢體或軀幹的一個最佳之近似中軸線段，而[12]則是考慮到手、腳有關節存在，無法單純只由一個近似中軸線段來代表，於是允許局部重疊的情況。[12]接著找出代表末端之近似中軸線段，省略了[11]的頭、手、腳規則判斷，直接用梯形代表肢體並建立人體模型。

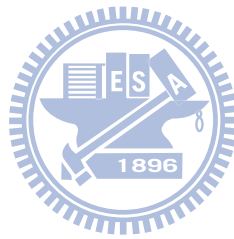


### 1.3 論文架構

針對上述行為分析或姿勢分類之研究，我們可以發現他們都無法分析像揮手、抬腳、投擲、拾物等較為局部之動作，原因在於他們缺少了頭、手、腳等的位置資訊。因此，本論文旨在由人體影像分析完成姿態識別，期望提供給姿勢分類或是行為分析之研究者做更細部動作之分析。據此，本論文首先將[11]、[12]的研究成果做一整合，並提出新的演算法加以改良，再加以延伸至[11]、[12]所沒有做到的頭手腳定位，期望藉由人體模型的肢體位置與關係來做為姿態識別之依據。圖 1-6 為本論文的系統流程圖，一開始利用前背景分離及雜訊濾除等技巧來擷取人體區域，這部份將在第二章做詳細的介紹。接著以[13]的近似中軸線概念為基礎，將原先四個方向的掃描線演算法作優化來得到橫截面，並利用更嚴謹



之規則來得到該肢體或軀幹的近似中軸線段。與[11]、[12]做完整後才找末端不同，我們在這裡就先找出有可能是肢體末端的近似中軸線段，所以我們提出一個有別與[11]、[12]的找末端演算法，詳細的做法會在第三章做介紹。由於我們在上一步驟就可以得到肢體末端的近似中軸線段，所以我們提出新的整合方式來產生人體模型，並且從觀察及實驗中得知一些可作為姿態識別之規則，利用這些規則配合人體模型做更進一步的延伸來區別頭、手、腳，第四章會將這些做一個完整的呈現。第五章是本論文的實驗結果及討論。第六章則是為本論文做一結論及探討未來工作。



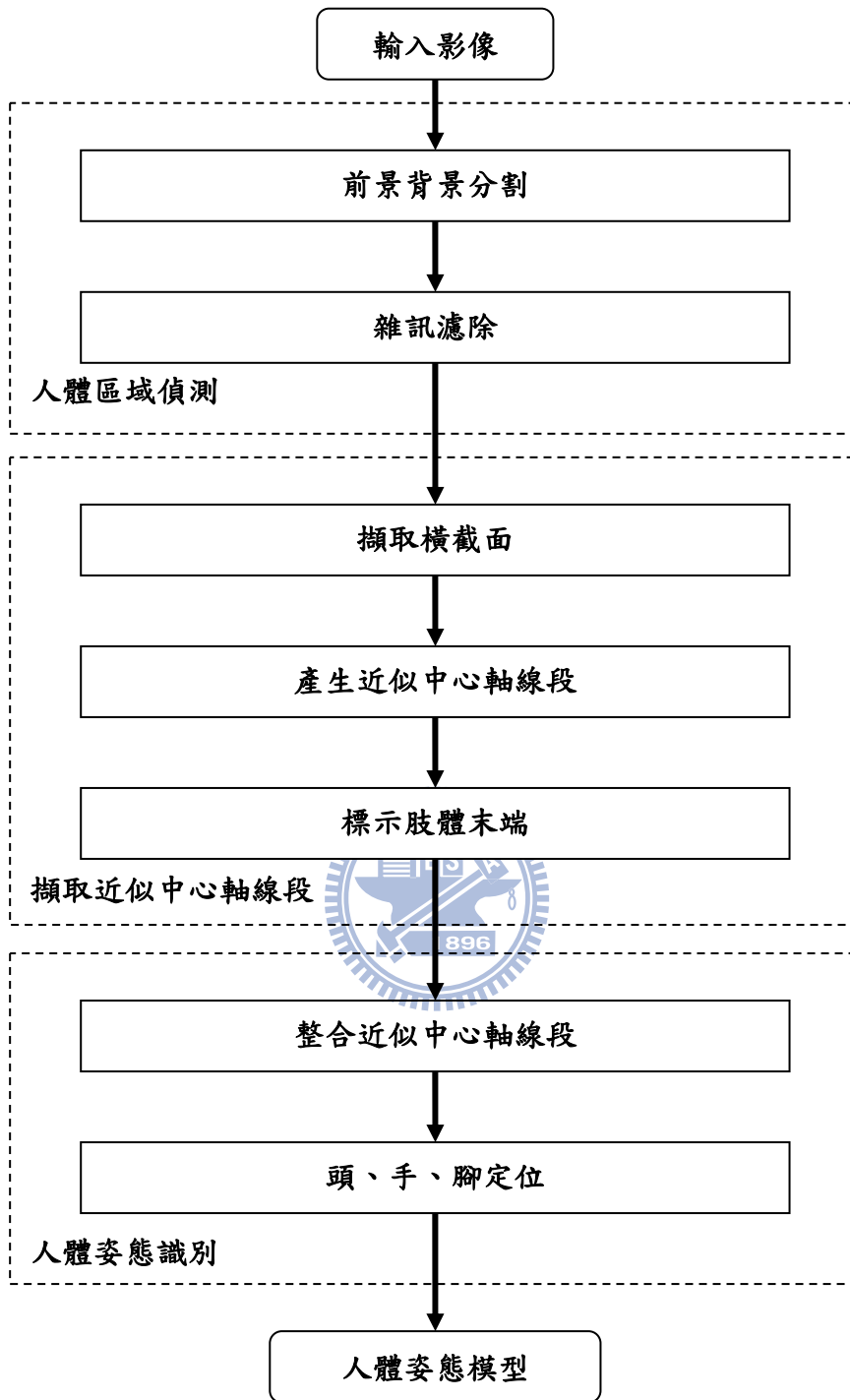


圖 1-6 系統流程圖。

## 二、人體區域偵測

為了能夠對人體區域進行肢體及姿態分析，本論文首先需將人體區域從背景影像中分離出來。因為可見光影像的像素值表示的是顏色，而中紅外線影像的像素值所代表的是溫度，所以必須針對兩者的特性分別設計出適合的前背景分割法。所以本章節旨在敘述針對不同特性之輸入影像，使用適合之前背景分割法來擷取人體區域，並將雜訊予以濾除。詳細流程圖請參考圖 2-1，接下來本論文會在各個小節將每個步驟做逐一介紹。

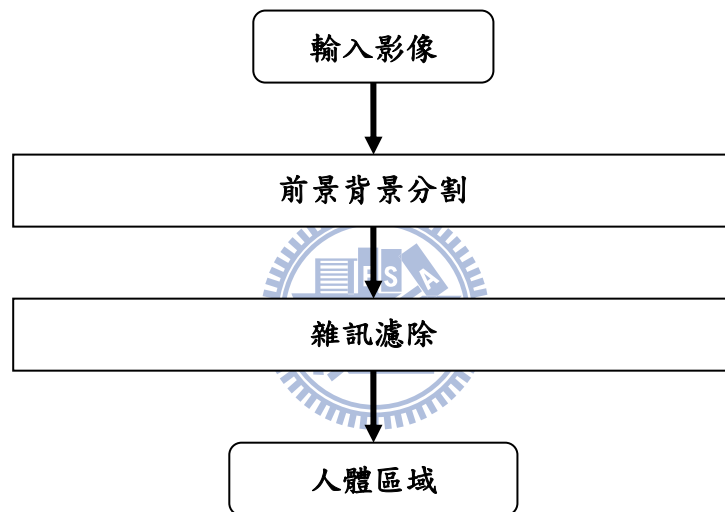


圖 2-1 人體區域偵測流程圖。

### 2.1 可見光影像人體區域擷取

針對可見光影像本論文使用背景減去法（background subtraction）來擷取人體區域，此方法的特點在於運算速度較快，且可適應光影變化問題，較適用於室內環境。其主要步驟如下：

- (1) 將擷取來的第一張影像當作背景模型。
- (2) 對每一張新進影像的每一個像素的強度分別和背景模型中同位置的像素強度相減，若其絕對值小於門檻值 $T_{BG}$ 即是背景，否則將之歸類為前景。

(3) 若一個像素在步驟(2)被判斷是背景，則利用式子(2-1)更新該像素位置的背景模型。

$$\mu_t = \alpha I_t + (1 - \alpha) \mu_{t-1} \quad (2-1)$$

其中 $t$ 代表新進來的影像時間， $\mu$ 代表背景模型的強度， $I$ 代表新進影像的強度，而 $\alpha$ 代表給定之學習速率，代表新進來的像素佔多大的比重。若 $\alpha=0$ ，則不論新進影像之像素為前景或是背景都不會更新背景模型的資訊，也就是將第一張影像視為背景模型後則永遠不再更新。若 $\alpha=1$ ，則代表直接取代舊有的模型資訊，也就是將最新一張影像視為背景模型更新的唯一來源。經過多次實驗，本論文將 $T_{BG}$ 設為 20， $\alpha$ 設為 0.025。

圖 2-2(a)為一可見光影像，圖 2-2(b)是圖 2-2(a)做完背景減去法的前背景分割圖，從圖中可以清楚地看到由於影子的關係，造成腳的下方有連接成一塊的情形，此情形會增加我們後面分析之難度。為了消除影子所造成的影響，本論文參考[14]的作法，將 RGB 色彩空間轉至 HSV 色彩空間，並用下列條件判斷是否為影子，如果下列條件均成立就被判定為影子。

$$(a) \quad \alpha \leq \frac{P^V(x, y)}{B^V(x, y)} \leq \beta$$

$$(b) \quad |P^S(x, y) - B^S(x, y)| \leq T_S$$

$$(c) \quad |P^H(x, y) - B^H(x, y)| \leq T_H$$

其中 $P^H(x, y)$ 、 $P^S(x, y)$ 、 $P^V(x, y)$ 分別為現在影像座標 $(x, y)$ 的HSV強度值， $B^H(x, y)$ 、 $B^S(x, y)$ 、 $B^V(x, y)$ 則分別為背景模型座標 $(x, y)$ 的HSV強度值， $\alpha$ 、 $\beta$ 、 $T_H$ 、 $T_S$ 是手動指定之門檻值，本論文經由實驗設定其值分別為 0.45、1.0、32 和 32。

圖 2-2(c)是將圖 2-2(b)做消除影子後的結果，我們可以清楚地看到原本腳下面因為影子的關係而連在一起的部份，得以正確地被分開。為了加快運算速度，本系統只針對背景減去法後被判斷成是前景的像素作消除影子的動作。

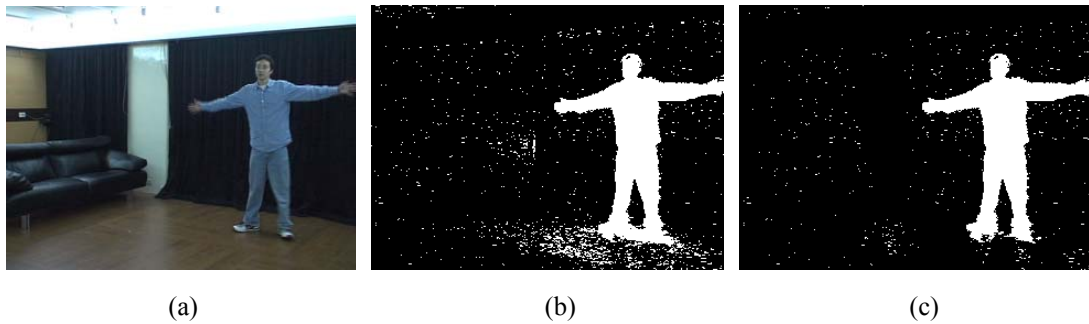


圖 2-2 可見光影像經由前景偵測找出之人體區域：(a) 原始影像，(b) 經背景減去法後得到之前景影像，(c) 經過消除影子後之前景影像。

## 2.2 中紅外線影像人體區域擷取

在[15]的研究中提到，作者經由實驗發現，中紅外線影像中若屬於同一物體的相鄰區域，則其溫度數值會非常接近。所以一個簡單擷取中紅外線影像中人體區域的方法就是直接利用式子(2-2)對中紅外線影像做二值化

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) > T_{temp} \\ 0, & \text{otherwise} \end{cases} \quad (2-2)$$

其中 $f(x, y)$ 代表中紅外線原始影像中座標 $(x, y)$ 的溫度值， $T_{temp}$ 是門檻值，本論文經由實驗設定其值為 2980，式子(2-2)的目的是將中紅外線影像中屬於較高溫部分的人體區域給偵測出來。

## 2.3 消除雜訊及填補人體中破碎之區域

由於在 2.1 節及 2.2 節中所擷取出之人體區域會有雜訊及破碎的情形發生，如圖 2-3(a)和 2-4(a)分別為輸入之可見光影像及中紅外線影像，圖 2-3(b)和 2-4(b)為使用 2.1 節及 2.2 節中提及的人體區域擷取方法所得到的前景影像，我們可以看到圖 2-3(b)和 2-4(b)中都有一些雜訊及破碎之情形。所以我們使用連通元件標記法 (connected component labeling) 來區分不同區域並計算其區域大小面積，將面積過小的區域予以濾除，並利用型態學上的技巧：侵蝕 (erosion) 以及擴散

(dilation) 來消除剩餘之雜訊及填補人體中破碎之區域。由圖 2-3(c)和 2-4(c)可以清楚地看到，原本在圖 2-3(b)和 2-4(b)存在的細碎區域及空洞都可以被濾除或是填滿。

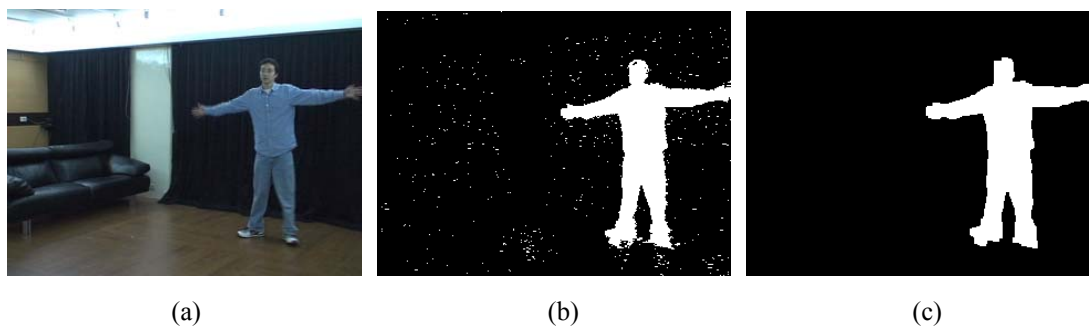


圖 2-3 可見光影像經由前景偵測找出之人體區域：(a) 原始影像，(b) 經背景減去法及消除影子後得到之前景影像，(c) 濾除過小面積及透過侵蝕、擴散後之前景影像。

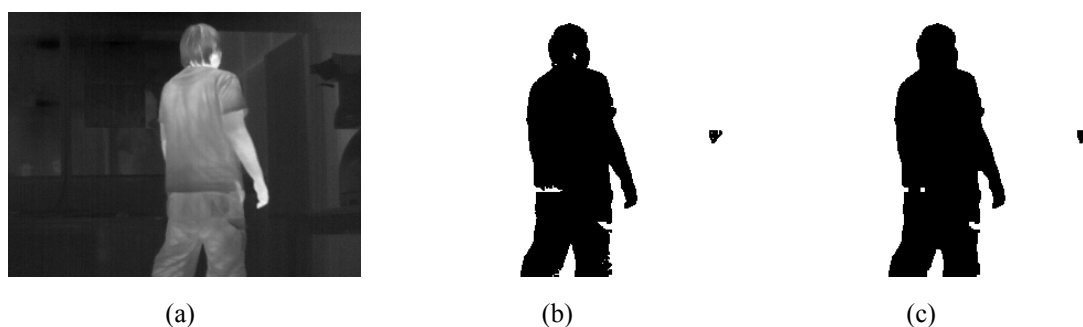


圖 2-4 中紅外線影像經由前景偵測找出之人體區域：(a) 原始影像，(b) 經取門檻值後得到之前景影像，(c) 濾除過小面積及透過侵蝕、擴散後之前景影像。

## 三、基於掃描線演算法之肢體近似中心軸線分析

本章節是利用[11]、[12]的概念來擷取肢體的近似中心軸線。首先我們將第二章中所得的人體區域，利用四個方向以及固定間隔的掃描線方法來掃描前景物以得到橫截面，並將過短的橫截面剔除，再利用 3.2 節的規則來建立近似中軸線段，並找出有可能是肢體末端之近似中軸線段以擷取建立模型所需之資訊。詳細流程圖請參考圖 3-1，接下來我們會在各個小節將每個步驟做逐一介紹。

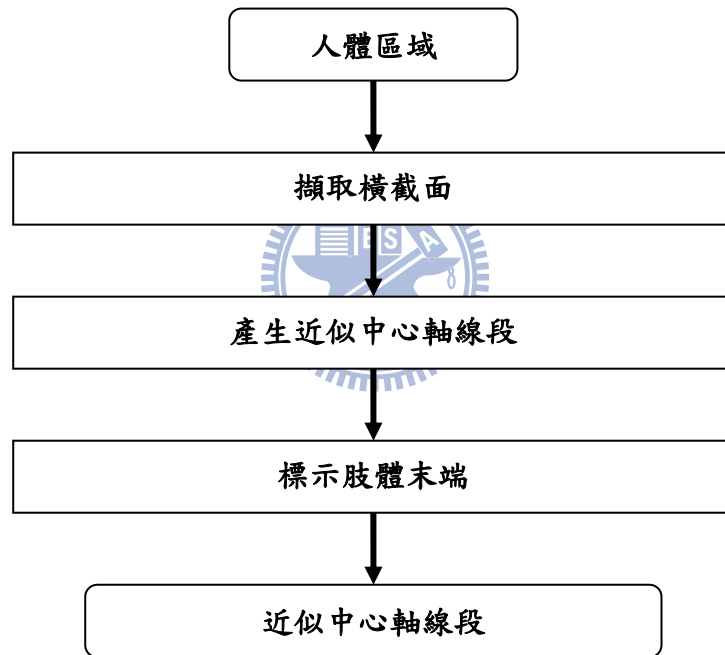


圖 3-1 擷取肢體近似中心軸線流程圖。

### 3.1 掃描線演算法

在找出肢體的近似中心軸線過程中，首先參考[13]的作法，以四種不同的方向來掃描人體區域，分別是  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$  和  $135^\circ$ ，藉此擷取人體區域之橫截面。掃描線與掃描線之間隔以參數  $\delta$  表示， $\delta$  愈小，則產生之肢體模型會更接近影像中的人體區域，但也會造成計算量提高及產生肢體模型時所需參考的資訊量變大。

在本論文中，我們假設肢體橫截面有一定的寬度，故將寬度過短的橫截面視為雜訊予以刪除，不考慮這些寬度過短的橫截面。經過數次實驗後，我們發現 $\delta$ 設定為 3 以及最小橫截面寬度的門檻值 $T_{CS}$ 設定為 5 是較合適的選擇。圖 3-2 顯示了設定 $\delta=3$  以及利用四個方向的掃描線掃描前景物所得到的橫截面，紅色的圈代表寬度過短之橫截面。圖 3-3 則是將圖 3-2 中寬度過短之橫截面刪除後的結果，紅色的圈代表已刪除的寬度過短之橫截面。

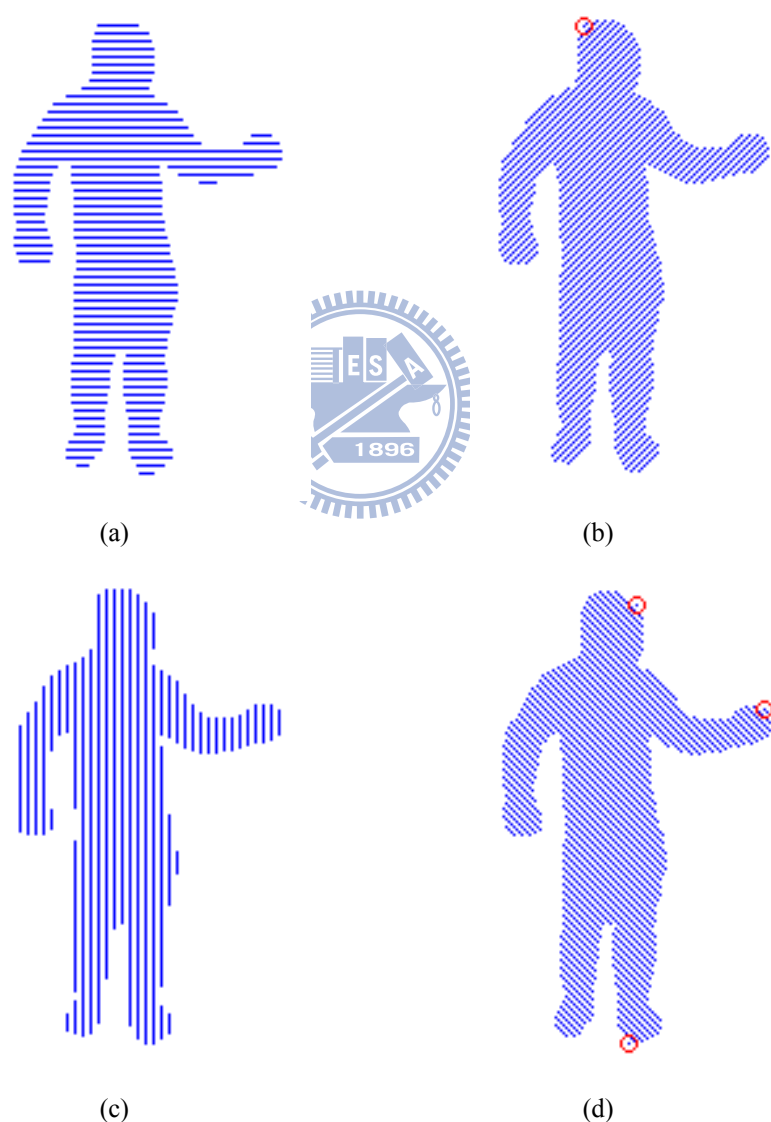


圖 3-2 以四個不同方向的掃描線掃描人體區域後所得之橫截面，圖中每條線段各代表一個橫截面： $(a) 0^\circ$ ， $(b) 45^\circ$ ， $(c) 90^\circ$ ， $(d) 135^\circ$ 。



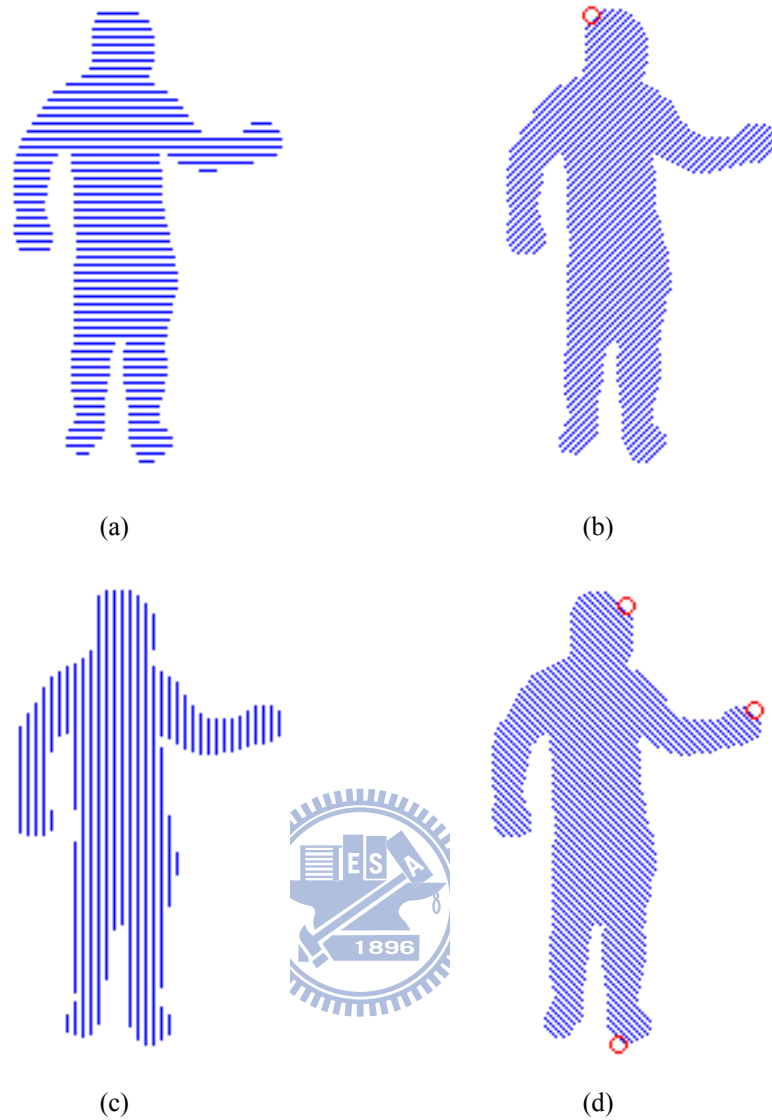


圖 3-3 將圖 3-2 中寬度過短之橫截面濾除的結果：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $90^\circ$ ，(d)  $135^\circ$ 。

在研究[11]、[12]中，都是先利用掃描線演算法擷取人體區域之橫截面，再將寬度過短的橫截面視為雜訊予以剔除。本論文為了加快運算速度，提出一個新的掃描線演算法，可於擷取橫截面時一併將過短的橫截面予以濾除，同時考慮到因為背景減去法所擷取的前景不佳產生過小之空洞，造成後面分析困難，如圖 3-4(a)為一前景影像，紅色所圈的空洞造成了圖 3-4(b)和(c)的橫截面不連續，這有可能會造成我們後面分析找出來的手位置不對或是變成有兩隻手，所以必須將這種過小的空洞給補起來。

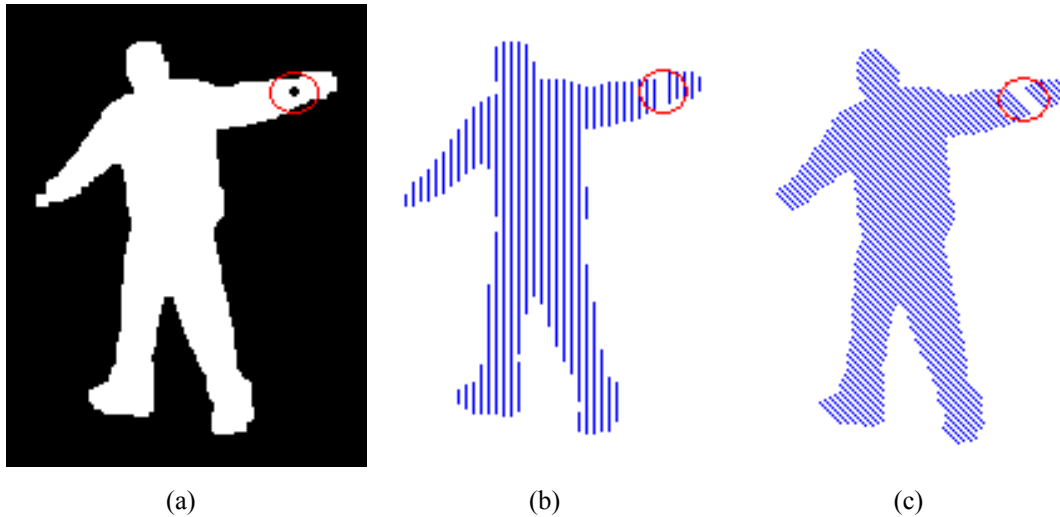


圖 3-4 空洞造成橫截面擷取不佳的情形：(a) 前景圖，(b)  $90^\circ$  濾除寬度過短之橫截面圖，(c)  $135^\circ$  濾除寬度過短之橫截面圖。

如圖 3-5 所示，掃描線演算法的做法如下：

- (1) 令掃描線的第一個橫截面的起點座標為  $A$ ，終點座標為  $B$ 。
- (2) 找到沿著掃描方向的下一個橫截面，令其起點座標為  $A'$ ，終點座標為  $B'$ 。
- (3) 判斷  $A'$  跟  $B$  的距離是否小於  $T_{noise}$ ，如果是即代表中間的空洞過小需要補滿，因此我們將  $B'$  紀錄為  $B$  並跳至步驟(2)繼續判斷。如果不是即表示這個空洞過大不屬於雜訊，跳至步驟(4)判斷橫截面  $AB$  是否過短予以刪除。其中門檻值  $T_{noise}$  代表其值多小視為空洞，本論文經由實驗設定其值為 3。
- (4) 判斷橫截面  $AB$  的距離是否小於  $T_{CS}$ ，如果是則視為雜訊不予以輸出，如果不是則表示非雜訊予以輸出，並將  $A'$  紀錄為  $A$ ， $B'$  紀錄為  $B$ ，並跳至步驟(2)繼續判斷。

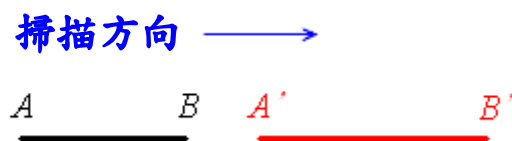


圖 3-5 掃描線演算法中的連續兩橫截面關係圖。

圖 3-6(a)、(d)顯示了  $90^\circ$  和  $135^\circ$  原來的橫截面圖，綠色的圈指的是寬度過短的橫截面，紅色的圈代表的是會造成後面分析誤判之空洞。圖 3-6(b)、(e)是經由 [11]、[12] 的掃描線演算法處理後之橫截面圖。圖 3-6(c)、(f)則是經過本論文的掃描線演算法處理後之橫截面圖。我們可以清楚地看到 [11]、[12] 的掃描線演算法雖然可以濾除寬度過短的橫截面（綠色的圈），但卻無法將會造成後面分析誤判之空洞（紅色的圈）補滿。而本論文的掃描線演算法不僅可以消除掉雜訊（綠色的圈），更可以進一步地將會造成後面分析誤判之空洞（紅色的圈）補滿。

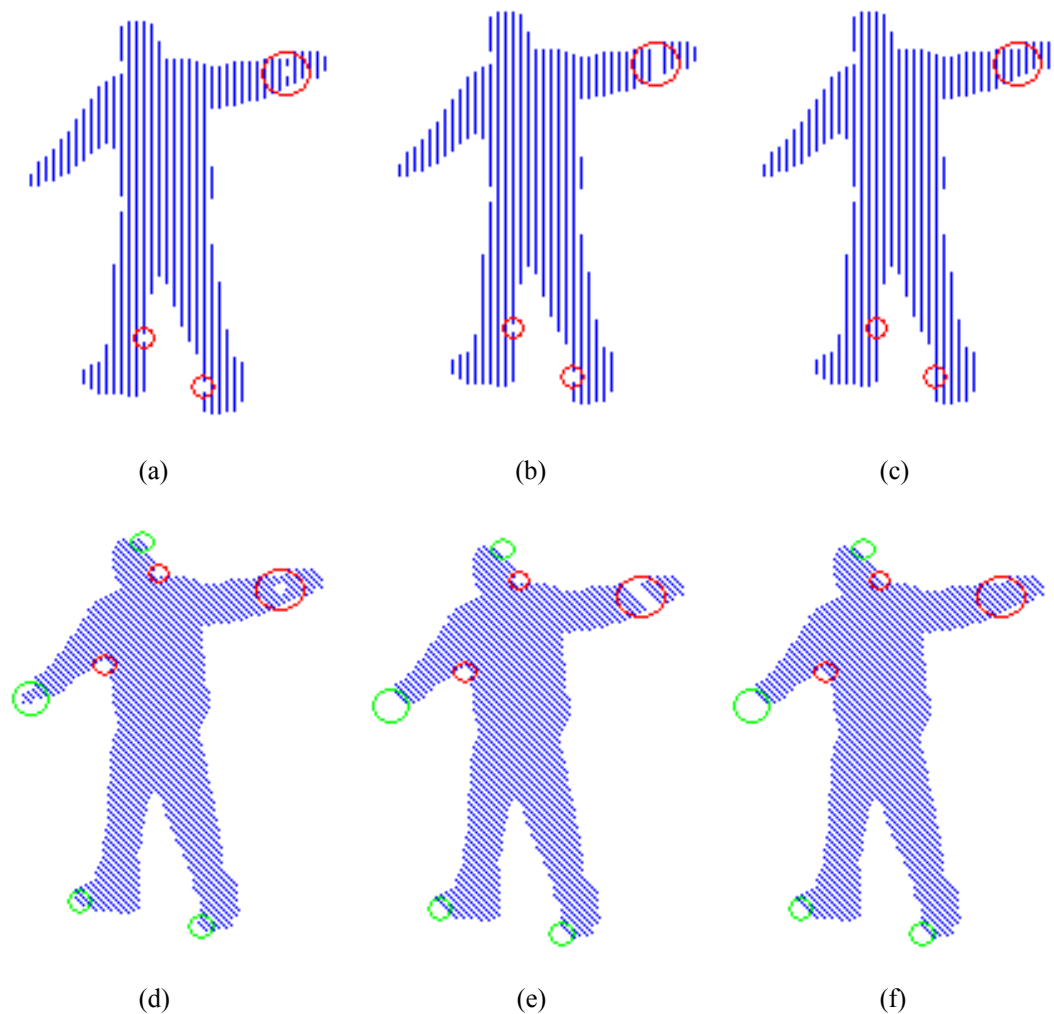


圖 3-6 改良掃描線演算法所得之橫截面：(a)  $90^\circ$  所有橫截面，(b) 濾除(a)中寬度過短之橫截面，(c) 濾除(a)中寬度過短之橫截面且補滿過小之空洞，(d)  $135^\circ$  所有橫截面，(e) 濾除(d)中寬度過短之橫截面，(f) 濾除(d)中寬度過短之橫截面且補滿過小之空洞。

### 3.2 掃描線橫截面之近似中軸線段

由 [11]、[12] 的研究可得知，橫截面中心點所連接而成的近似中軸線段可代表局部肢體的特徵，所以本章節旨在敘述利用上一節所找出的掃描線橫截面來建立近似中軸線段。由於這個部份的分析是引用 [11]、[12] 的研究成果，所以詳細的證明及推導過程本論文在此不多加冗述，有興趣的讀者可參考 [11]、[12] 的研究或是詳見附錄 A。[11]、[12] 提出，決定兩個連續掃描線的中心點是否相連，所考慮的條件為：

- (a) 兩中心點連線  $Q$  與橫截面之垂直方向  $R$  的夾角  $\theta$  是否小於等於  $45^\circ$ 。
- (b) 兩相鄰橫截面的長度差值是否小於等於 5。

一旦符合上述兩個條件，即將該橫截面視為同肢體的一部份並將中心點相連，如此我們便可獲得許多近似中軸線段。圖 3-7 為兩連續橫截面中心點連線與橫截面之垂直方向的夾角的關係圖，其中紅點為橫截面之中心點，兩中心點連線為  $Q$ ，橫截面之垂直方向為  $R$ ， $\theta$  為  $Q$  和  $R$  的夾角。

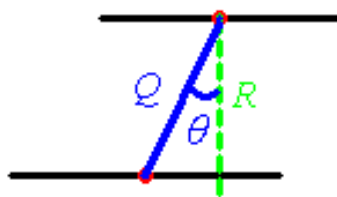


圖 3-7 連續橫截面中心點連線與橫截面之垂直方向的夾角的關係圖。

[11]、[12] 提出，由於人的肢體部份具有一定長度且不太可能是呈現過寬或過扁的形狀，如圖 3-8 所示，藍線代表四個方向的掃描線，黑線表示找出來的近似中軸線段，紅色圓圈代表的即是過寬或過扁的形狀，我們可以清楚地看到它們並不適合代表人的肢體。

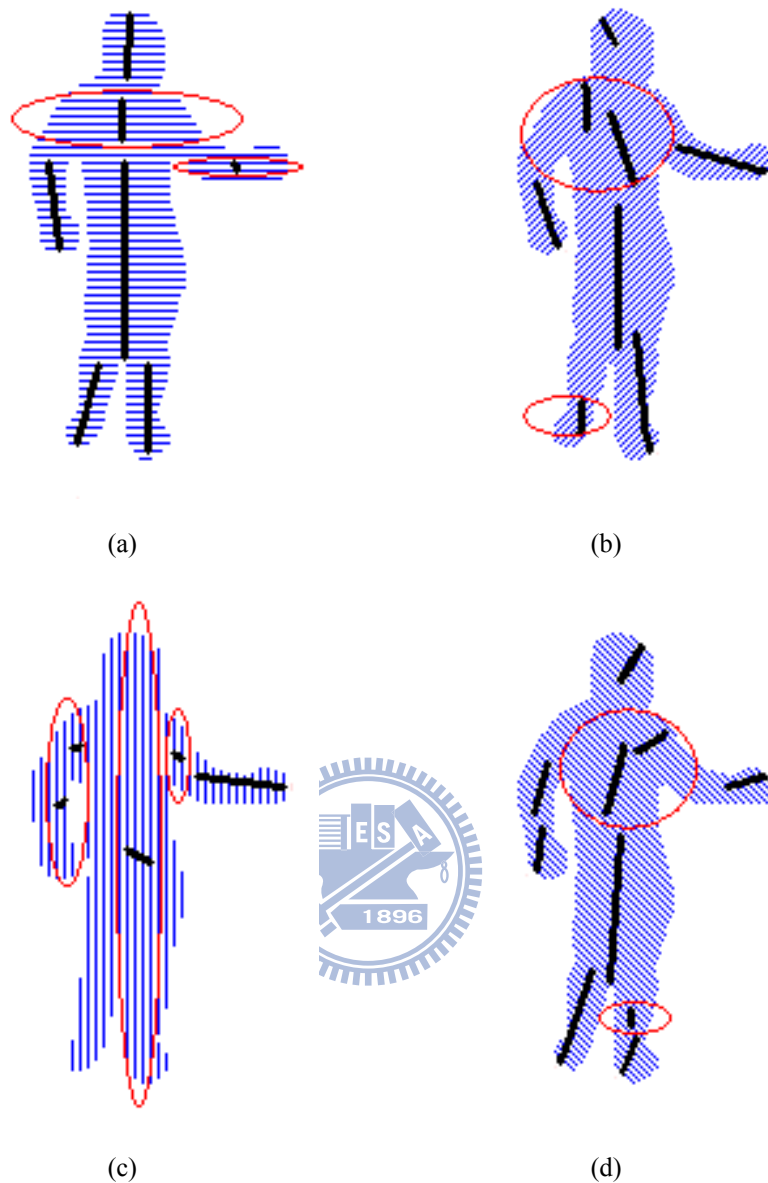


圖 3-8 四個方向掃描線的近似中軸線段：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $90^\circ$ ，(d)  $135^\circ$ 。

因此[11]、[12]進一步地利用下列條件來過濾較不合適的中心軸線段：

- (a) 該中心軸線段是否由三個以上的橫截面所構成。
- (b) 中心軸線段的平均橫截面寬度是否小於等於該中心軸線段之長度。

圖 3-9 是移除圖 3-8 中過寬或過扁的近似中軸線段後的圖，由圖中我們可以清楚地看到，經過上述的條件的過濾後，我們可以獲得較適合代表肢體的近似中軸線段。

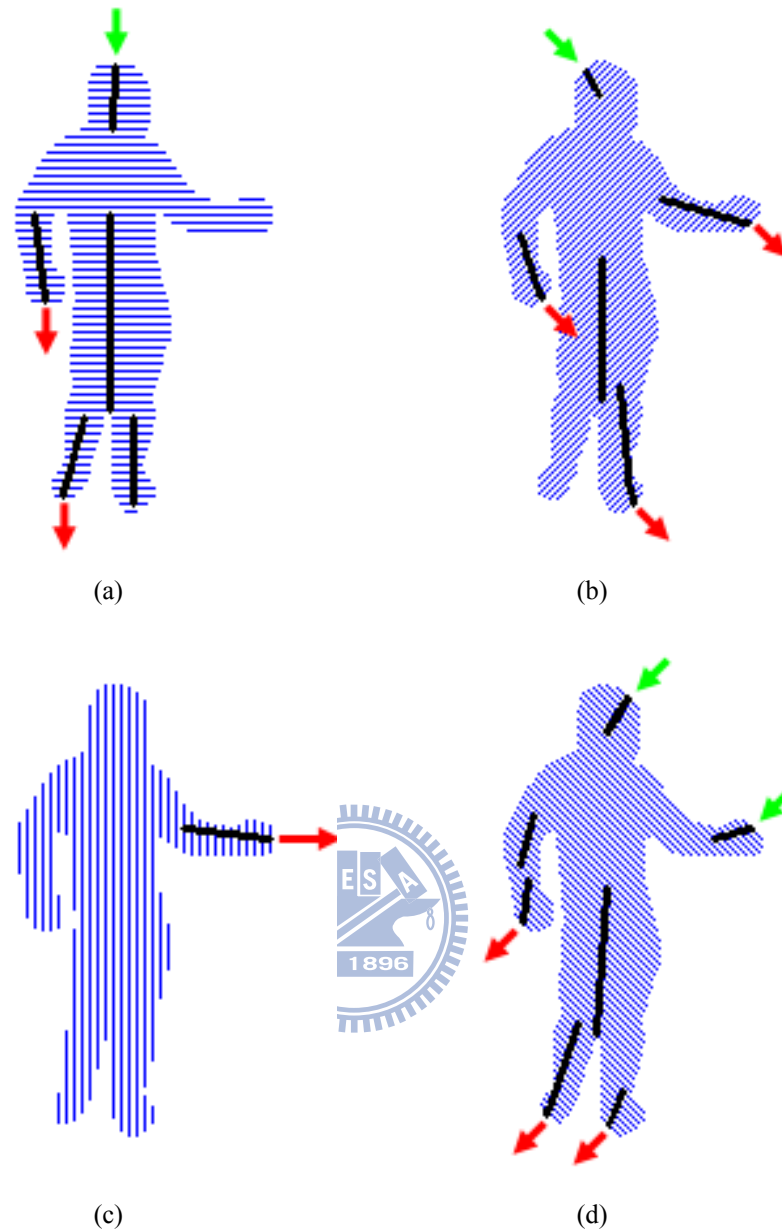


圖 3-9 移除圖 3-8 中過寬或過短的近似中軸線段：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $90^\circ$ ，(d)  $135^\circ$ 。

### 3.3 標示頭及肢體之末端

接下來，不同於[11]、[12]的做法，我們在此階段先不進行四個方向掃描線的整合，而是先找出有可能是肢體末端的近似中心軸線段。這是由於若先進行整合的話，原先代表肢體末端之近似中軸線段有可能因為長度或角度不佳而遭到取代，使得代表肢體末端的近似中軸線段變少或不見，會更進一步造成我們所建立出來的人體模型不符合該人物的真實姿態。如圖 3-10 為圖 3-9 中取右腳部份，

選取角度較佳的做為代表該部份之近似中軸線段，圖 3-10(d)中可以清楚地看到，因為圖 3-10(a)的角度較佳而被選取，但是它卻不具有末端特性(稍待會詳述何謂末端特性)，這會造成我們找不到右腳。

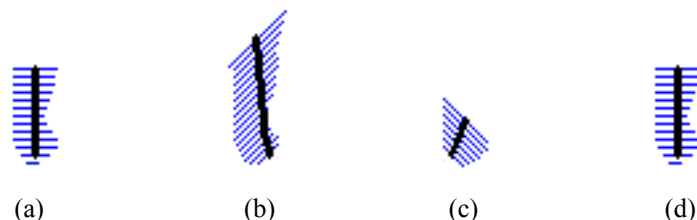


圖 3-10 圖 3-9 右腳部份及整合後的近似中軸線段：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $135^\circ$ ，(d) 整合後。

由圖 3-9 中我們可清楚地看出，可以代表肢體末端的近似中軸線段，其橫截面的分布必定會出現由無到有（綠色箭頭）或者由有到無（紅色箭頭）的情形。因此，我們根據這個末端特性來判斷近似中軸線段的起始（終止）橫截面，其行進方向上（下）一段掃描線的相對位置不應存在橫截面。如圖 3-11， $L_1$ 為近似中軸線段的起始橫截面， $L'_1$ 為 $L_1$ 在行進方向上一段掃描線上的相對位置， $L_2$ 為近似中軸線段的終止橫截面， $L'_2$ 為 $L_2$ 在行進方向下一段掃描線上的相對位置。我們可以清楚地看到，在 $L'_1$ 上並不存在任何橫截面，因此我們可以說 $L_1$ 的中心點足以代表肢體末端，反之在 $L'_2$ 上存在一橫截面，因此 $L_2$ 的中心點不足以代表肢體末端。圖 3-12 為圖 3-9 經過找末端方式後所得到的肢體末端圖，水藍色的點即代表具有末端特性的肢體末端，我們稱其為末端特徵。

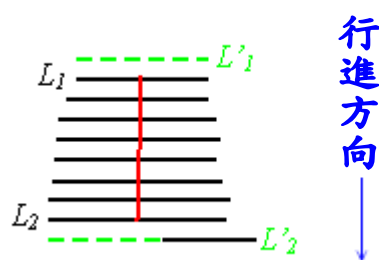


圖 3-11 找近似中軸線段末端方法的示意圖。

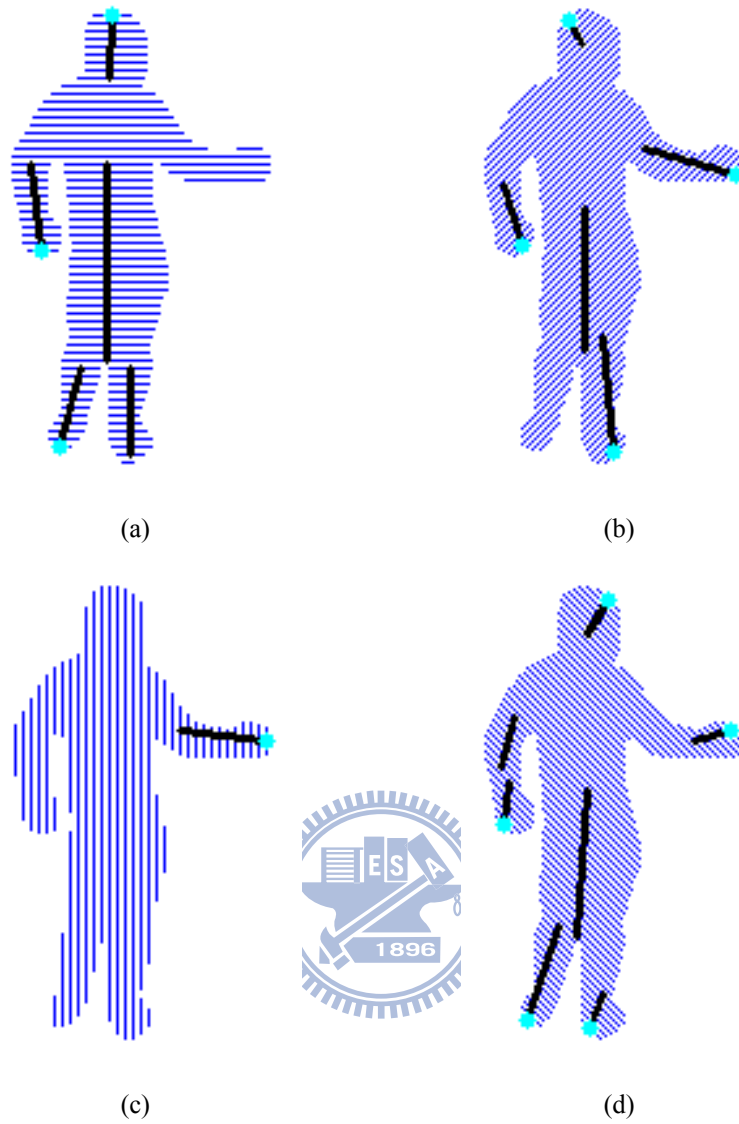


圖 3-12 四個方向掃描線可以代表肢體末端的近似中軸線段：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $90^\circ$ ，(d)  $135^\circ$ 。



## 四、基於頭、手、腳定位的姿態識別

為了得到人體肢體之活動表示並進行姿態識別，因此本章節首先將第三章中所得到的四個方向的近似中軸線段做整合，選出最適合表示該肢體形狀之近似中軸線段，並找出具有末端特徵的肢體，接著利用 4.2 節的規則對其肢體進行頭、手、腳之定位。為了便於觀察，我們利用一人體模型來顯示姿態識別後的結果，這方面將於 4.3 節作介紹。詳細流程圖請參考圖 4-1，接下來我們會在各個小節對於每個步驟做逐一介紹。

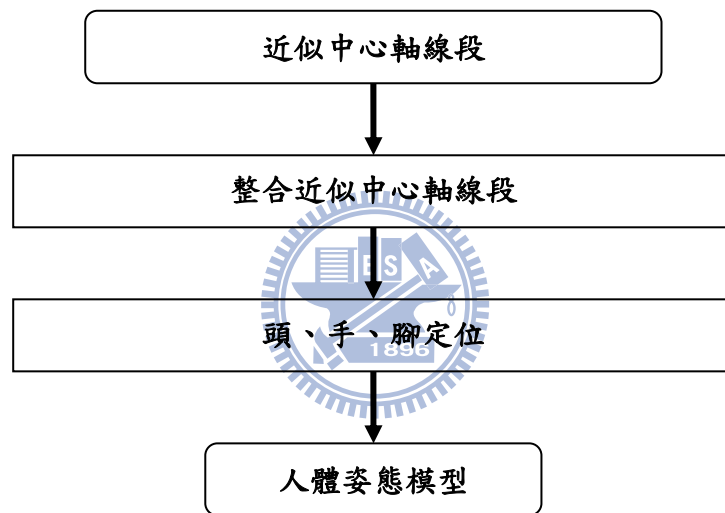


圖 4-1 姿態識別流程圖。

### 4.1 整合近似中軸線段

為了提供一人體模型來進行姿態識別，我們必須將第三章中所得到的四個方向的近似中軸線段整合成一個人體模型骨架。首先，我們利用第三章所得到的末端資訊來幫助我們做不同肢體之分類，之所以不直接利用末端做整合是因為擁有末端特徵並不代表其形狀就足以代表某一肢體之末端。如圖 4-2 為頭部之橫截面分布（藍線）及含末端特徵（水藍色點）的近似中軸線段（黑線），90°方向掃描線因為沒有頭部之近似中軸線段，所以此處不予列出，由圖中我們可以看出，如

果我們直接選擇利用末端特徵來進行整合，則代表頭部的近似中軸線段不是  $45^\circ$  方向（圖 4-2(b)）就是  $135^\circ$  方向（圖 4-2(c)），可是實際上選擇  $0^\circ$  方向（圖 4-2(a)）才是最接近頭部形狀及走向的近似中軸線段，因此末端特徵不適用於直接做整合。同樣地，如果選取圖 4-2(a)代表頭部，會因為它不具有末端特徵，使得我們找不到頭部之定位點，所以我們必須利用圖 4-2(b)和 4-2(c)的末端特徵來輔助圖 4-2(a)，輔助的方式會在稍後的部分提到。

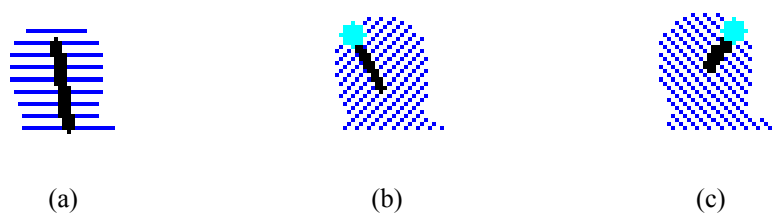


圖 4-2 頭部之橫截面分布及含末端特性的近似中軸線：(a)  $0^\circ$ ，(b)  $45^\circ$ ，(c)  $135^\circ$ 。

當我們在整合頭部的近似中軸線段時，很顯然地我們並不會考慮到手的近似中軸線段，因為頭和手是屬於不同的肢體。因此，我們首先須將所有的近似中軸線段分類成對應於不同肢體部位的集合。例如我們最後分類成 5 個集合，這 5 個集合可能代表的是頭、兩隻手和兩隻腳。而未屬於任一集合中的近似中軸線段意即無法代表肢體末端，一般來說就是軀幹部份。所以我們挑選所有具有末端特徵的近似中軸線段，對其所對應的每一條近似中軸線段做下列條件之判斷：

- (a) 起始或是終止橫截面的中心點與該末端距離小於  $T_{tip}$ 。
- (b) (a)找到的中心點與該末端的連線中沒有連出去到背景。

其中，本論文經由實驗設定  $T_{tip}$  為 20。我們將符合上述兩個條件的近似中軸線段歸類後，便表示同一集合內的近似中軸線段皆足以代表某一肢體。如圖 4-3，黑色部分是背景，白色部分是前景，紅線、藍線和綠線分別是  $0^\circ$ 、 $45^\circ$  和  $135^\circ$  掃描線所獲得的近似中軸線段，水藍色點代表末端特徵，由圖中我們可以清楚地看到藍線的末端特徵點和紅線的起點  $A$  及綠線的末端特徵點的距離夠近，且點的連線

中沒有連出去到背景，所以我們可以知道這三條近似中軸線段都是可以代表某一肢體的，因此就會將這三條近似中軸線段歸類在同一集合內。

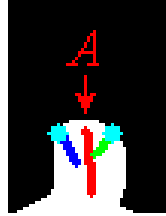


圖 4-3 足以代表頭部的所有近似中軸線段。

接著，我們對於每個集合，挑選出一個最適合代表該肢體末端部位的近似中軸線段，挑選步驟如下：

- 一、將集合中近似中軸線段和橫截面垂直方向的夾角最小的挑選出來，以其夾角作為考量的原因是當夾角越小越能代表該近似中軸線段越貼近肢體的走向。
- 二、再從最小的夾角角度的容許範圍內，挑選長度最長的近似中軸線段，如果該近似中軸線段沒有末端特徵，則我們在最接近末端的起始或終止橫截面中心點上添加末端特徵。如圖 4-3，假設我們選取到紅線所代表的近似中軸線段，因為其不具有末端特徵，因此我們找到距離藍線或綠線所代表的近似中軸線段的末端特徵中最接近的點  $A$  當作紅線的末端特徵。
- 三、將找到的最長的近似中軸線段跟集合中每個近似中軸線段做面積重疊的比較，將有重疊的近似中軸線段予以濾除，如果集合中還有近似中軸線段存在，則回到步驟(1)繼續判斷。
- 四、如果找到的近似中軸線段只有一條，則用該近似中軸線段代表那一肢體末端部位。如果找到的近似中軸線段不只一條，即代表該肢體屬於多節肢體，此時我們挑選出最適合的近似中軸線段來代表肢體末端部位。

以上為我們挑選肢體末端的主要步驟，步驟二～四的詳細做法我們將於接下來的段落做介紹。如圖 4-4(a)、4-4 (b)、4-4 (c)、和 4-4 (d)分別代表  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$  和

135°的橫截面分布及近似中軸線段，圖 4-4(e)則是由這四個角度作整合後的近似中軸線圖，圖中的標記代表著該近似中軸線段是由哪個角度所獲得的，如圖 4-4(e)中標記 C 的近似中軸線段就是從 0°方向（圖 4-4(a)）所獲得的。為了不要每次都標記像圖 4-4 的關係圖，所以之後我們都像圖 4-4(e)所示，在非末端特徵點畫出橫截面來示意該近似中軸線是從哪個方向所獲得的。

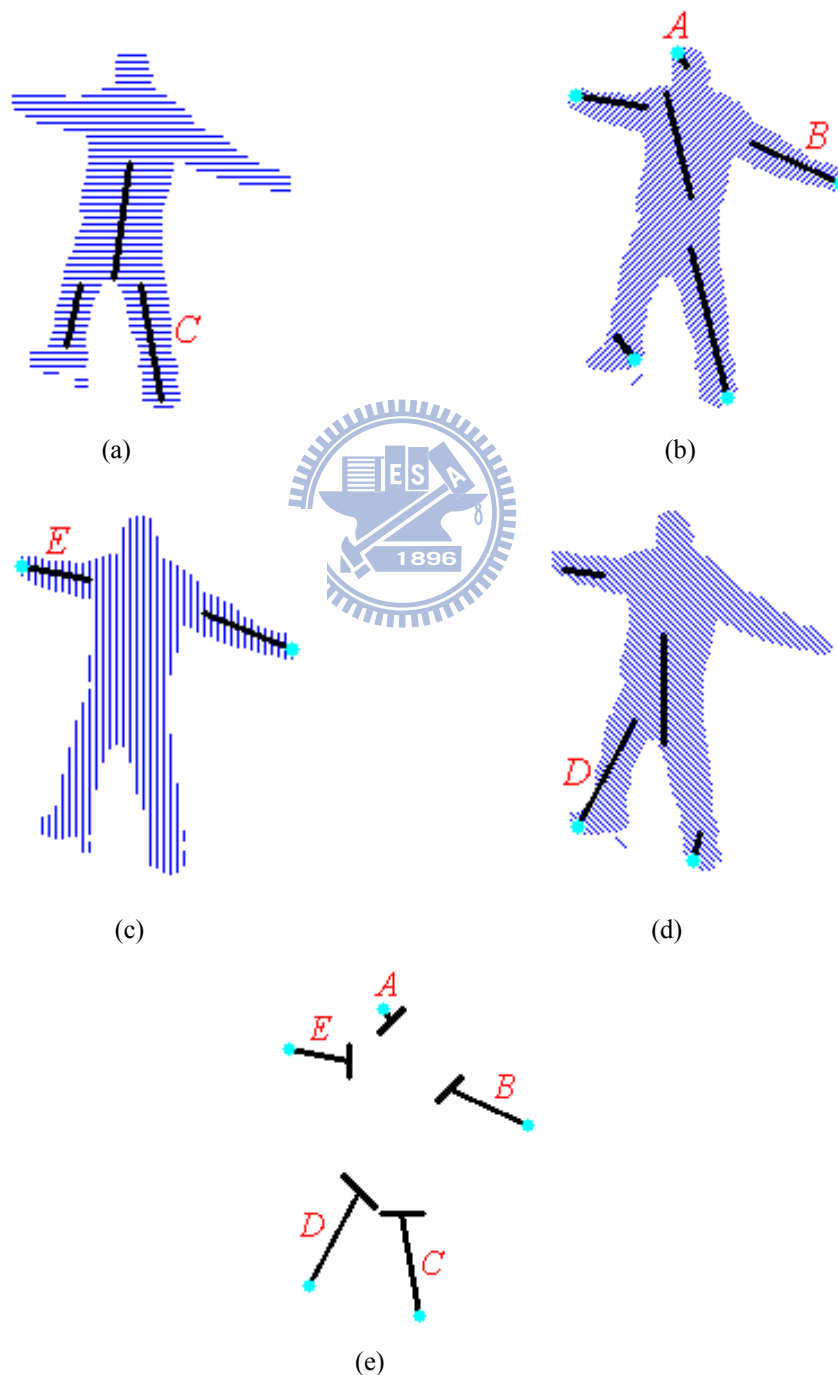


圖 4-4 近似中軸線段的整合：(a) 0°，(b) 45°，(c) 90°，(d) 135°，(e) 整合結果。

### 4.1.1 夾角的角度容許範圍之定義及選取

在步驟二中，因為我們考慮到同時使用近似中軸線段和橫截面垂直方向的夾角，以及其長度來做為我們挑選最適合表示肢體末端的依據，因此在我們找出夾角最小的近似中軸線段後，再從最小的夾角角度的容許範圍中，挑選長度最長的近似中軸線段。如圖 4-5 為一近似中軸線段（以下用  $Q$  代表）與橫截面垂直方向（以下用  $R$  代表）的夾角（以下用  $\theta$  代表）。

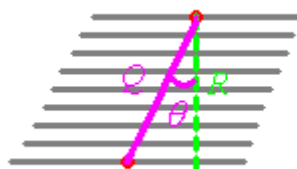


圖 4-5 近似中軸線段與橫截面垂直方向的夾角的關係圖。

從 3.2 節的連接兩連續橫截面中心點其夾角必須在  $45^\circ$  以內的條件，我們可以知道  $\theta$  最大會是  $45^\circ$ ，也就是說  $\theta$  的範圍會在  $0^\circ \sim 45^\circ$ 。同時，由附錄 A 之討論，我們可以知道  $\theta$  在  $22.5^\circ$  內表示  $Q$  的走向用  $R$  代表是最佳的。因此，我們將角度分為  $25^\circ$ （最好）、 $30^\circ$ （好）、 $35^\circ$ （普通）、 $40^\circ$ （不好）和  $45^\circ$ （最不好）這五種等級。 $0^\circ \sim 25^\circ$ 、 $25^\circ \sim 30^\circ$ 、 $30^\circ \sim 35^\circ$ 、 $35^\circ \sim 40^\circ$ 、 $40^\circ \sim 45^\circ$  分別屬於  $25^\circ$ 、 $30^\circ$ 、 $35^\circ$ 、 $40^\circ$ 、 $45^\circ$  這一類。而角度的容許範圍就是比目前等級次一級的我們也加入長度的考慮，如圖 4-6 為腳部近似中軸線段，其中近似中軸線段的顏色表示的就是角度目前等級，黑色表示  $25^\circ$ （最好），紅色表示  $30^\circ$ （好），綠色表示  $35^\circ$ （普通），藍色表示  $40^\circ$ （不好），紫紅色表示  $45^\circ$ （最不好），之後的圖示只要沒有特別提及，則近似中軸線段的顏色代表的就是  $\theta$  的等級。在圖 4-6 中我們選到的  $\theta$  等級會是  $25^\circ$ （最好），同時我們會將次一等級，也就是  $30^\circ$ （好）也一併納入長度之選取，也就是黑線跟紅線都會拿來做長度之選取，最後我們會選取到最長的紅線來代表腳。



圖 4-6 同一腳部區域之近似中軸線段：(a)  $0^\circ$ ，(b)  $135^\circ$ 。

#### 4.1.2 兩近似中軸線段是否面積重疊之判斷

在步驟三中，我們為了將在同一肢體部位的近似中軸線斷集中跟選取到的近似中軸線段有交集的部分去掉，所以我們必須進行面積是否重疊之判斷。由於近似中軸線段代表的其實是該肢體部位的骨架，為了呈現近似該肢體部位的形狀，於是我們將近中軸線段兩端的橫截面當作梯形的上、下底，利用此梯形模型來代表肢體部位的形狀。如圖 4-7 所示，綠線代表兩端的橫截面，也就是梯形上下底，紅線則是連接上下底的端點，也就是梯形的側邊。

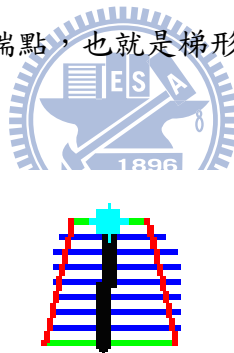


圖 4-7 利用梯形代表肢體部位的形狀。

為了判斷兩梯形面積是否有重疊，[12]的作法是藉由判斷兩梯形的四個邊及一條對角線是否有相交，來判斷梯形是否重疊。此方法的優點是速度較快，但是卻無法排除有較大的梯形包住較小的梯形或是只有一點或是一邊重疊的情況。因此，本論文對此方法做一修改，令其能將上述情形也加以排除。由於前面部分是引用[12]的作法來判斷線段是否相交，在此就不多加冗述，有興趣的讀者可參考附錄 B。接下來我們便延伸線段是否相交的判斷方式來進行兩梯形是否重疊的判斷，其做法如下：





經過上面的步驟後，我們就可以知道兩梯形是否有重疊，然而我們也可清楚地知道如果梯形沒有重疊，就必須做  $4 \times 4 = 16$  次的線段比較以及 8 次點是否在梯形內部的判斷，而 1 次的線段比較需要 4 次的 CCW 運算，1 次點是否在梯形內部的判斷也需要 4 次的 CCW 運算，所以共需要 96 次的 CCW 運算才能確定兩梯形沒有重疊，加上經過我們觀察兩梯形不重疊的機率大於重疊的機率，因此在做上述步驟之前我們先利用下列簡單的方式做篩檢，以減少判斷兩梯形沒有重疊所花的計算時間：

- (1) 將欲判斷是否重疊之兩梯形，找到各自的外接矩形 (bounding box)，令其中一外接矩形的端點為  $(P_1, P_2)$ 、 $(P_3, P_4)$ ，另一外接矩形的端點為  $(P_5, P_6)$ 、 $(P_7, P_8)$ 。
- (2) 如果  $|P_1 - P_7|$  小於等於  $(|P_1 - P_3| + |P_5 - P_7|)$  且  $|P_2 - P_8|$  小於等於  $(|P_2 - P_4| + |P_6 - P_8|)$ ，則代表兩梯形有可能重疊，繼續做 CCW 之判斷。如果條件不符，即代表兩梯形絕不可能重疊。

如圖 4-9(a) 呈現的是兩個梯形 ( $T_1, T_2$ ) 沒有重疊其外接矩形 (綠色虛線) 也沒有重疊，圖 4-9(b) 呈現的是兩個梯形 ( $T_1, T_2$ ) 有重疊其外接矩形 (綠色虛線) 也有重疊，圖 4-9(c) 中雖然兩個梯形 ( $T_1, T_2$ ) 沒有重疊但是外接矩形 (綠色虛線) 卻有重疊。因此，此方法雖然可以減少沒有重疊的梯形的計算時間，但同時會增加有重疊的梯形的計算時間。本論文經過統計加入此方法可有效減少 10% 計算梯形是否重疊的時間。



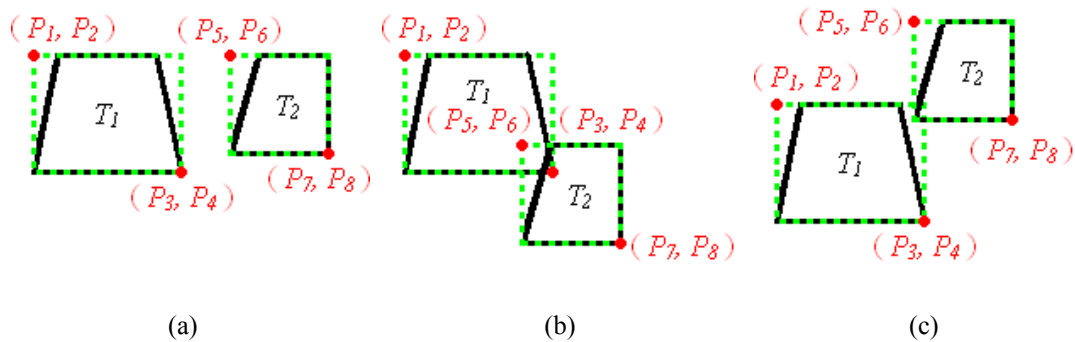


圖 4-9 兩梯形與外接矩形之關係：(a) 梯形與外接矩形皆不重疊，(b) 梯形與外接矩形皆重疊，(c) 梯形不重疊但外接矩形卻重疊。

### 4.1.3 多節肢體中肢體末端之選取

在步驟四中，如果我們只是選一個較佳的近似中軸線段來代表肢體末端，這有可能造成選到的近似中軸線段的末端特徵離該肢體的真實末端過於遙遠的窘境。如圖 4-10(a)為一尚未經過步驟四所判斷的人體模型骨架，我們可以清楚地看到，右邊手部（藍色圈）存在一多節肢體的情形，如果我們只是選擇一個較佳的近似中軸線段來表示此手部，則選到的人體模型表示如圖 4-10(b)所示，這是因為黑色的角度及長度都比紅色的來的較好，可是考慮到為了呈現真實肢體末端，應該是選擇紅色的來代表手部肢體較為恰當，因為其末端特徵離手部真實的末端較近。因此，本論文利用末端特徵到另一近似中軸線段的非末端特徵的橫截面中點的距離來選取適當的肢體末端，如圖 4-10(a)的黑色的末端特徵  $A$  到紅色非末端特徵  $B'$  的距離小於紅色的末端特徵  $A'$  到黑色非末端特徵  $B$  的距離，也就是  $AB' < A'B$ 。因此，我們可以斷定紅色是屬於多節肢體中較為外側的，較適合表示肢體末端部位，如圖 4-10(c)。

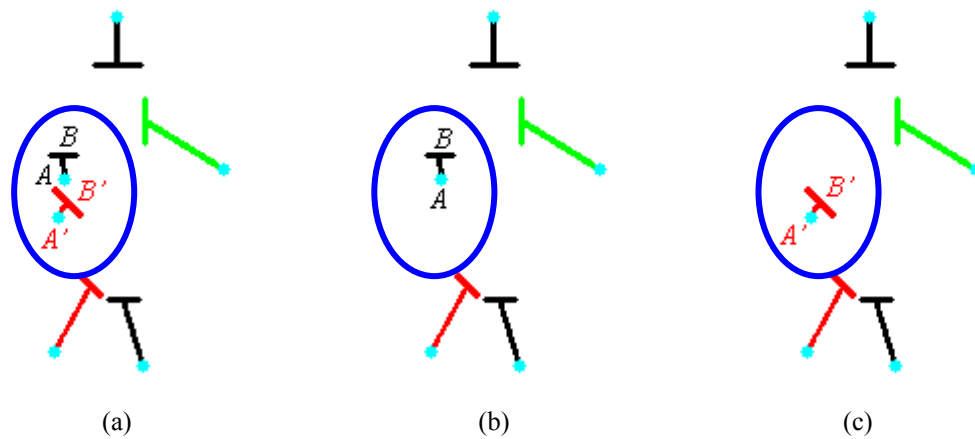


圖 4-10 多節肢體中末端選取之方式：(a) 都選取，(b) 只選取角度及長度較佳的，(c) 只選取多節表示中與外側較接近的。

## 4.2 基於肢體末端數目的頭、手、腳定位

接下來我們將介紹[11]、[12]所沒有做到的頭、手、腳定位。經過上述的處理後，我們可以得到足以代表人體模型骨架的近似中軸線段及其末端特徵。經由觀察我們發現這些末端特徵都足以當作頭、手、腳的定位點，如圖 4-11 為處理後所獲得的近似中軸線段，我們會認為 1 代表頭部定位點，2、3 代表手部定位點，4、5 代表腳部定位點。基於以上的觀察，這節的主要目的就是將上一節所得到的近似中軸線段做進一步地識別，判斷該近似中軸線是代表哪個肢體部位。

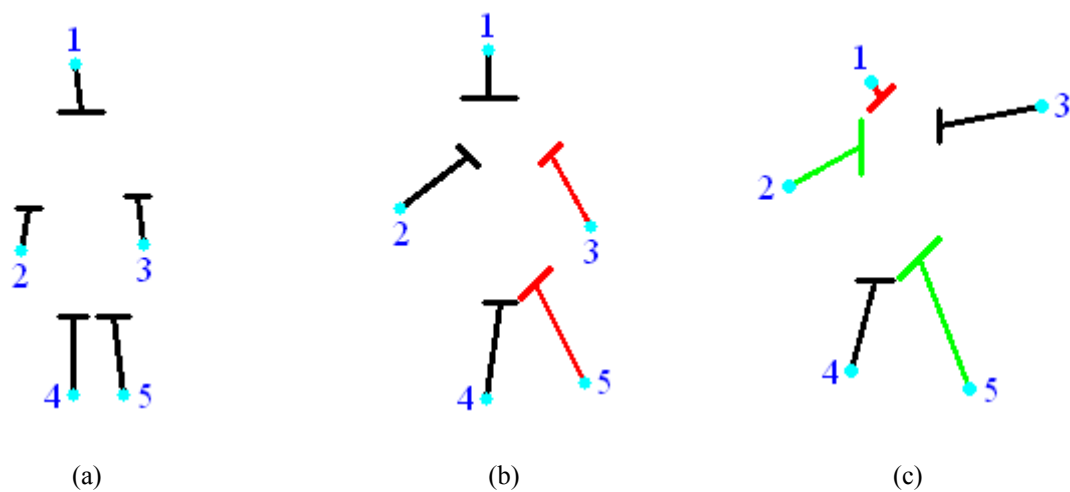


圖 4-11 從近似中軸線段找到頭、手、腳定位點的例子：(a) 例子 1，(b) 例子 2，(c) 例子 3。

為了識別出近似中軸線段所代表的肢體部位是頭、手還是腳，我們利用近似中軸線的一些資訊來進行判斷。如圖 4-12 所示，近似中軸線段所擁有的資訊有：

- (a) 角度：連接末端特徵點及非末端特徵的橫截面起點與終點來計算其夾角，如圖 4-12 藍色部分所示。
- (b) 長度：末端特徵點到非末端特徵的橫截面中點的距離，如圖 4-12 紅色部分所示。
- (c) 寬度：末端特徵點和非末端特徵的橫截面平均寬度，如圖 4-12 綠色部分所示的寬度平均。
- (d) 長寬比：末端特徵到非末端特徵的橫截面的垂直距離及寬度的比例，如圖 4-12 紫紅色部分所示的垂直距離長度除以寬度平均。

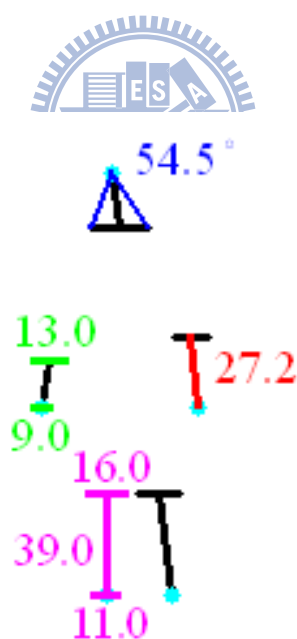
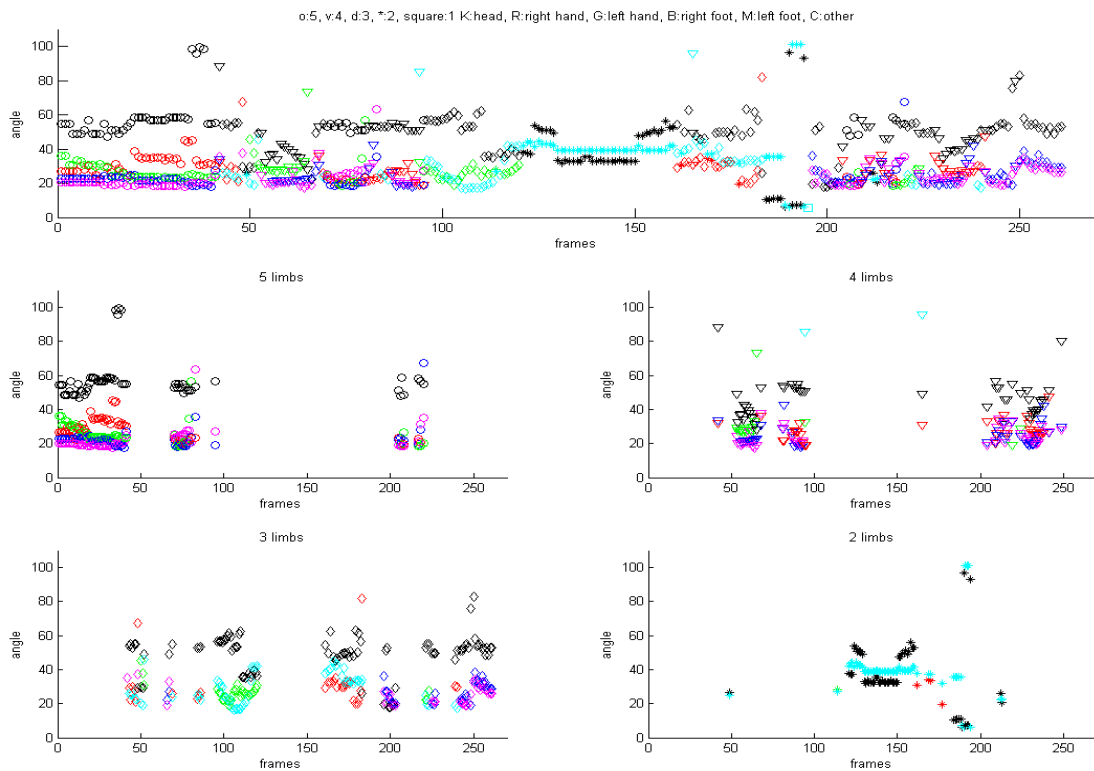


圖 4-12 近似中軸線段的資訊。

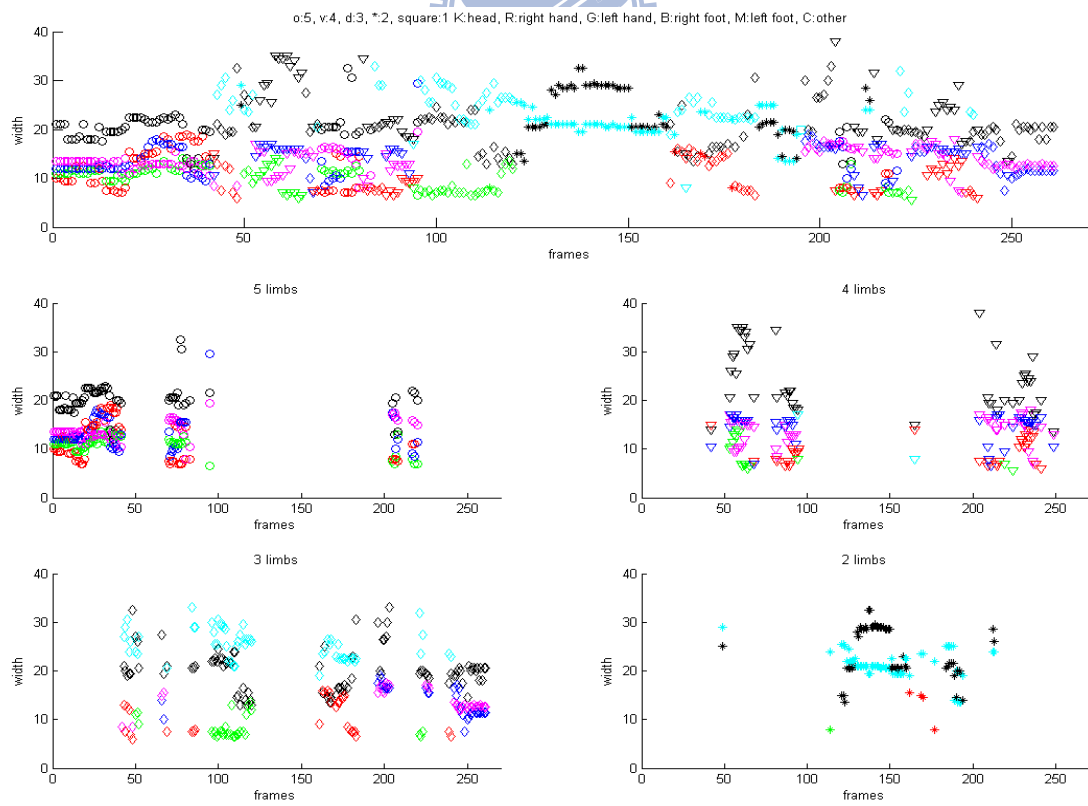
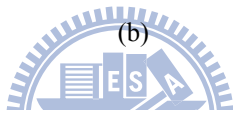
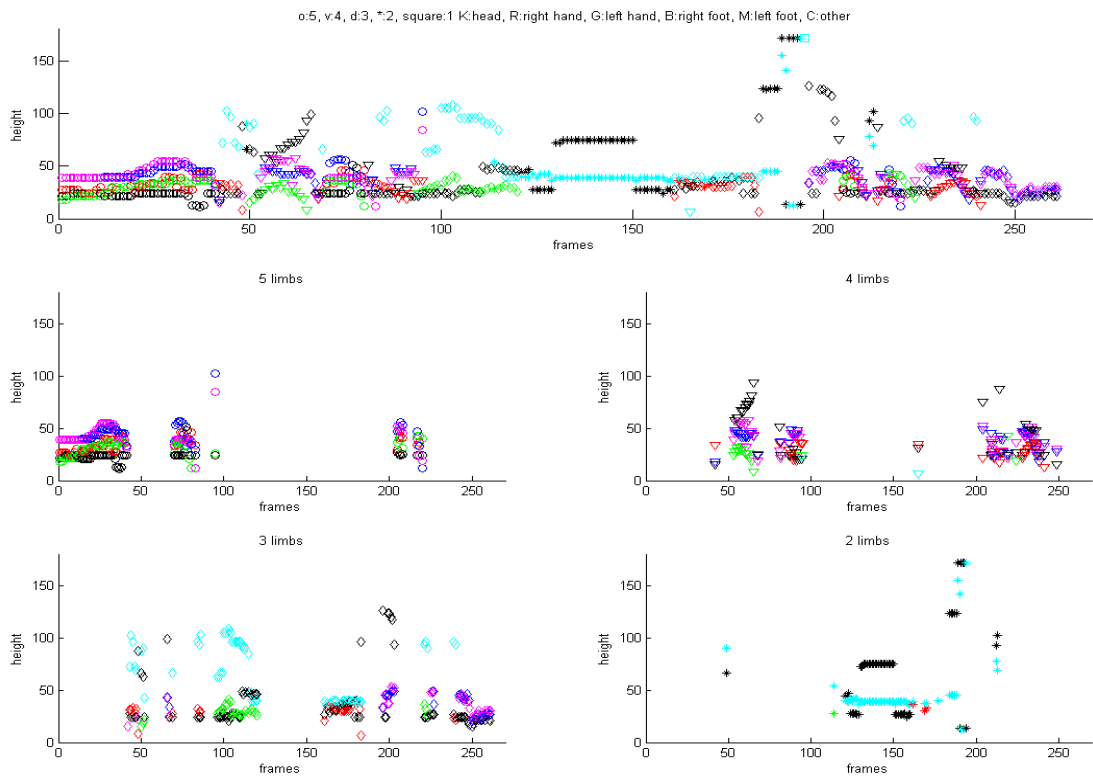
為了清楚地知道上述資訊所能帶給我們識別頭、手、腳的可能性，我們利用一段影片（268 frames）擷取出近似中軸線段及其資訊後，將資訊的數值畫在以 frame 編號為橫軸的座標上，並以人工方式進行分類，希望藉由分布圖來找到識別頭、手、腳之特徵。圖 4-13 為頭、手、腳的資訊分布圖，圖 4-13(a)、4-13(b)、

4-13(c)、4-13(d)分別是角度、長度、寬度、長寬比的分布圖。每個圖中的顏色代表不同之肢體部位，黑色代表頭部，紅色、綠色分別代表右手和左手，藍色、紫紅色分別代表右腳和左腳，水藍色代表兩隻腳合併成一隻腳。形狀則是代表這張影像中有幾個肢體部位，圓圈代表五個，三角形代表四個，菱形代表三個，星星代表二個，正方形代表一個。為了清楚地看到不同肢體的頭、手、腳分布，我們在每個圖中分別放入四個子圖，左中的子圖代表著五個肢體的頭、手、腳分布，右中的子圖代表著四個肢體的頭、手、腳分布，左下的子圖代表著三個肢體的頭、手、腳分布，右下的子圖代表著二、一個肢體的頭、手、腳分布。



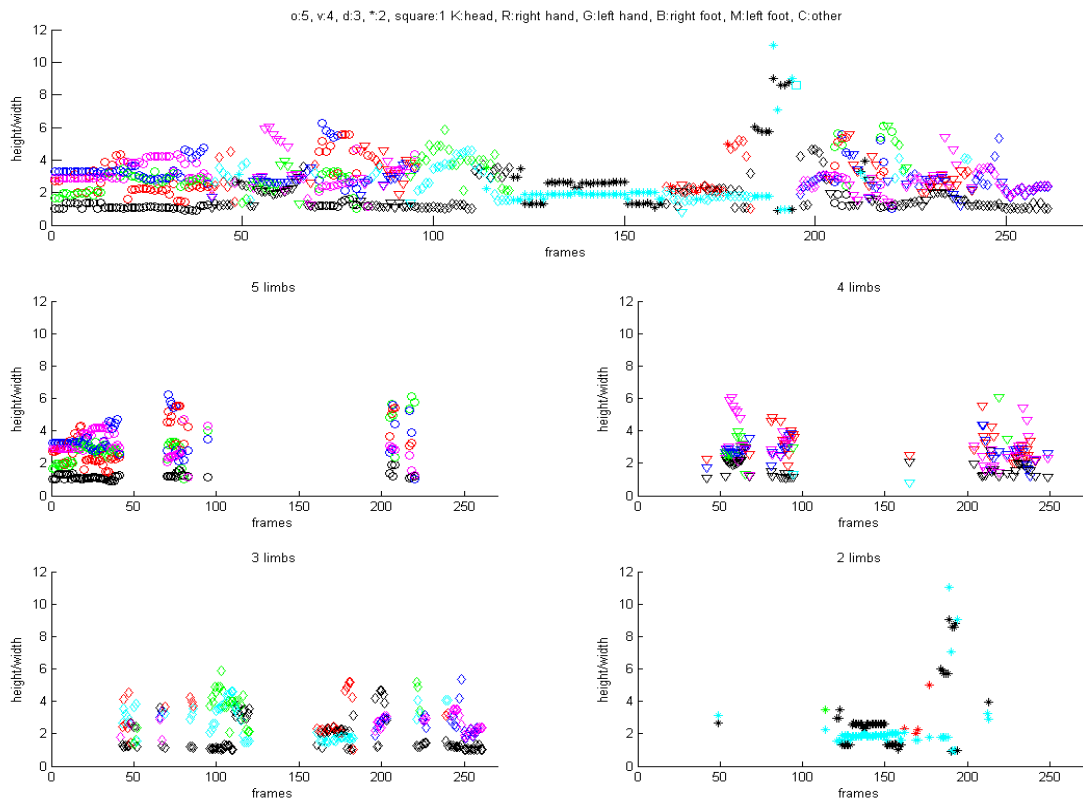
(a)

圖 4-13 近似中軸線段的肢體部位資訊：(a) 角度分布圖，(b) 長度分布圖，(c) 寬度分布圖，(d) 長寬比分布圖。



(c)

圖 4-13 (續)



(d)  
圖 4-13 (續)

從圖中我們可以看出角度是比較可以區分出頭的一個特徵，但是我們也可以看到在 1、2 個肢體部分的分布圖中反而不好利用角度來區分出頭，一是因為肢體少不好做比較，二則是因為當兩隻腳合併成一隻的時候會造成其角度大於頭部。如圖 4-14 皆是兩隻腳合併成一隻腳的角度大於頭部的例子，紅色的數字代表該近似中軸線段的的角度。因此，我們必須針對肢體數量的多寡來進行不同方式之判斷，當肢體數量較多時，利用角度及其他特徵來找頭、手、腳，當肢體數量較少時（1、2 個），則使用區域分割（region splitting）的方式來找頭、手、腳。接下來各個小節會說明針對不同肢體數量的詳細做法。

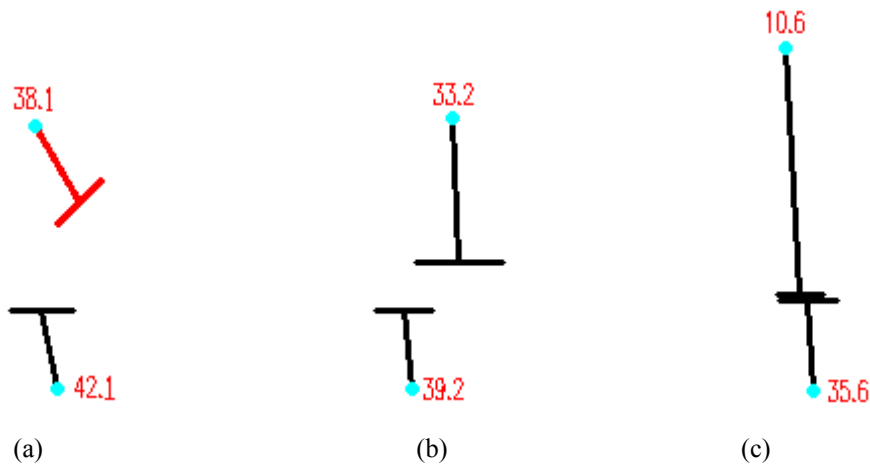


圖 4-14 兩隻腳合併成一隻腳的角度大於頭部的例子：(a) 例子 1，(b) 例子 2，(c) 例子 3。

#### 4.2.1 基於 5 個肢體末端條件下的作法

由於 5 個肢體的數量較多，因此我們直接選取角度最大的近似中軸線當作頭，而從圖 4-13 我們可以清楚地看到手和腳無法使用這些特徵來做明顯區分，因此我們使用身體部位的相對關係來做手、腳之判斷。經由觀察我們得知，腳和頭的距離遠大於手和頭的距離，因此我們可利用肢體和頭的距離來區分手腳，如圖 4-15(a)就是利用此方法做識別，圖中數字代表該肢體的角度。但是如果遇到像是 4.1.3 節所提及的多節肢體，就有可能因為多節肢體被分成數個部份，使得其角度大於頭部，造成識別錯誤，如圖 4-15(b)。經由觀察，我們可以知道多節肢體會發生在頭部的情況極少，因此我們針對這種被辨識為多節肢體的近似中軸線，不將其進行角度最大的挑選，如圖 4-15(c)即是不將多節肢體進行挑選最大角度的近似中軸線的識別結果，藍色的環代表的是該肢體為多節肢體。以下將我們的做法以條列式敘述：

- (1) 找出 5 個肢體中，不含多節肢體中角度最大的即是頭。
- (2) 找到距離頭最遠的 2 個肢體即是腳。
- (3) 剩下的 2 個肢體就是手。

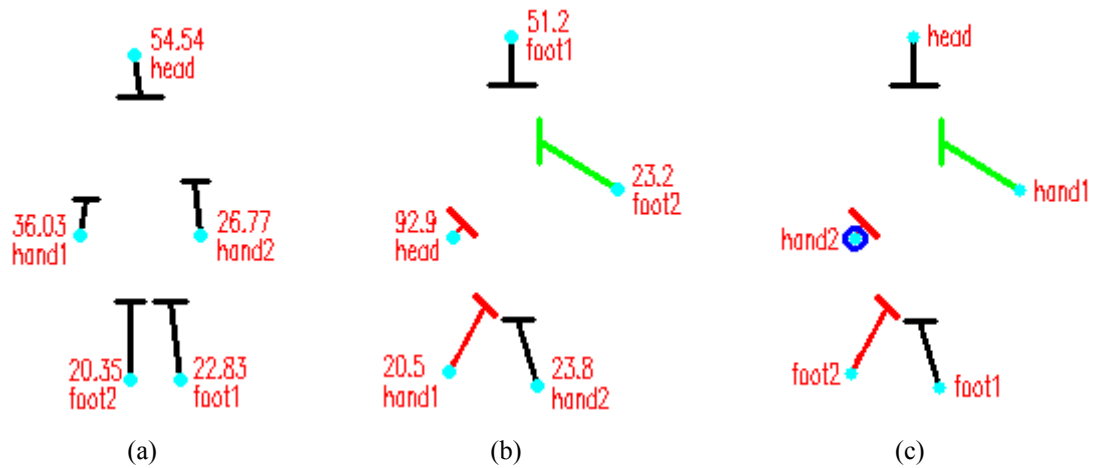


圖 4-15 5 個肢體的識別結果：(a) 不具多節肢體的例子，(b) 有多節肢體的例子，(c) 將(b)加入多節肢體不是頭的識別結果。

#### 4.2.2 基於 4 個肢體末端條件下的作法

由於 4 個肢體所呈現的有可能是 1 頭 1 手 2 腳、1 頭 2 手 1 腳、0 頭 2 手 2 腳這三種情形，因此我們不能像 4.2.1 節一樣，直接把角度最大的近似中軸線當作頭，因為會讓 0 頭 2 手 2 腳這種情形識別錯誤，如圖 4-16(a)。我們首先須將有頭和沒有頭的情形做分類，經由觀察我們發現，如果是 0 頭 2 手 2 腳這種情形，所有的肢體角度都會非常相近，所以我們藉由判斷最大角度和最小角度的差值是否小於門檻值  $T_{headless}$ ，如果是即代表是 0 頭 2 手 2 腳這種情形，經由實驗我們設定  $T_{headless}$  為  $10^\circ$ 。當我們分類完有頭跟沒有頭的情形後，接下來就是識別出 0 頭 2 手 2 腳的手和腳了，由於 4.2.1 節是利用頭當參考點找出其相對距離來識別手腳，因此我們必須在沒有頭的情形下找出一個參考點來識別手腳，我們利用人體區域的長寬比來判斷現在人體的姿勢是屬於直立的還是躺著的，如果長寬比大於門檻值  $T_{lie}$  就是直立的，此時我們利用  $0^\circ$  掃描線最上面的橫截面中點來當作我們的參考點，如圖 4-16(b)。如果長寬比小於等於門檻值  $T_{lie}$  就是躺著的，此時因為我們無法識別出人頭是朝左還是朝右，因此在此階段我們會輸出它是躺著的標記，並先當作人頭是朝右，等到之後如果能做更精確的判斷時再予以修正，此時我們利用  $90^\circ$  掃描線最右邊的橫截面中點來當作我們的參考點。經由實驗我們設



定 $T_{lie}$ 為 0.7。而針對 1 頭 2 手 1 腳和 1 頭 1 手 2 腳的情形，我們只需識別出它是 1 腳還是 2 腳就可以知道是哪種情形了，由觀察得知頭分別到兩腳的距離其實相差不多，所以我們利用距離頭第二遠跟最遠的比值是否大於門檻值 $T_{foot}$ ，如果是就是 1 頭 1 手 2 腳的情形，如圖 4-16(c)。否則就是 1 頭 2 手 1 腳的情形，如圖 4-16(d)。經由實驗我們設定 $T_{foot}$ 為 0.8。以下將我們的做法以條列式敘述：

- (1) 找出 4 個肢體中，不含多節肢體中角度最大和最小的近似中軸線，如果最大和最小的角度差值小於 $T_{headless}$ ，則是 0 頭 2 手 2 腳的情形，跳至步驟(2)，否則最大角度的就是頭，跳至步驟(4)。
- (2) 判斷人體區域的長寬比是否大於 $T_{lie}$ ，如果是則使用  $0^\circ$  掃描線最上面的橫截面中點來當作參考點，否則使用  $90^\circ$  掃描線最右邊的橫截面中點來當作參考點。
- (3) 找到距離參考點最遠的 2 個肢體即是腳，剩下的 2 個肢體就是手。
- (4) 找到距離頭最遠的 2 個肢體，如果第二遠跟最遠的比值大於門檻值 $T_{foot}$ 就二個都是腳，否則只有最遠的一個是腳。而剩下的肢體就是手。

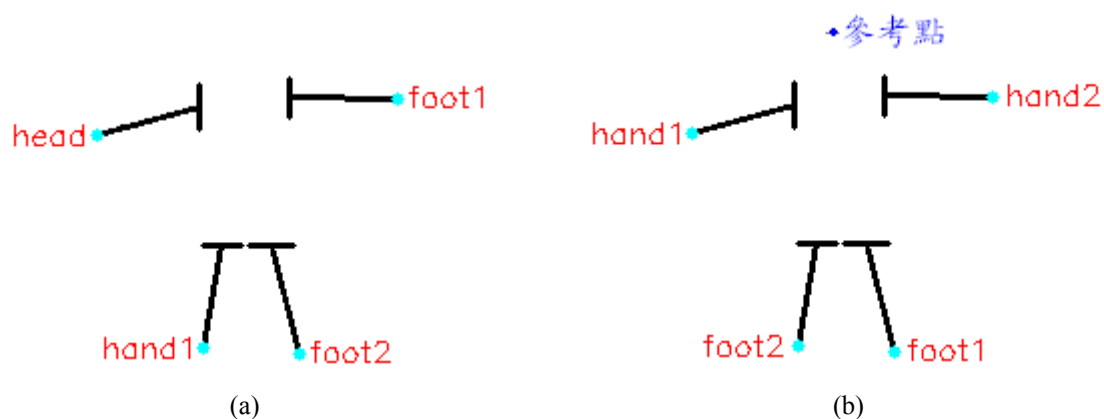


圖 4-16 4 個肢體的識別結果：(a) 未經過判斷有沒有頭的 0 頭 2 手 2 腳的識別結果，(b) (a) 經過判斷有沒有頭的識別結果，(c) 1 頭 1 手 2 腳的識別結果，(d) 1 頭 2 手 1 腳的識別結果。

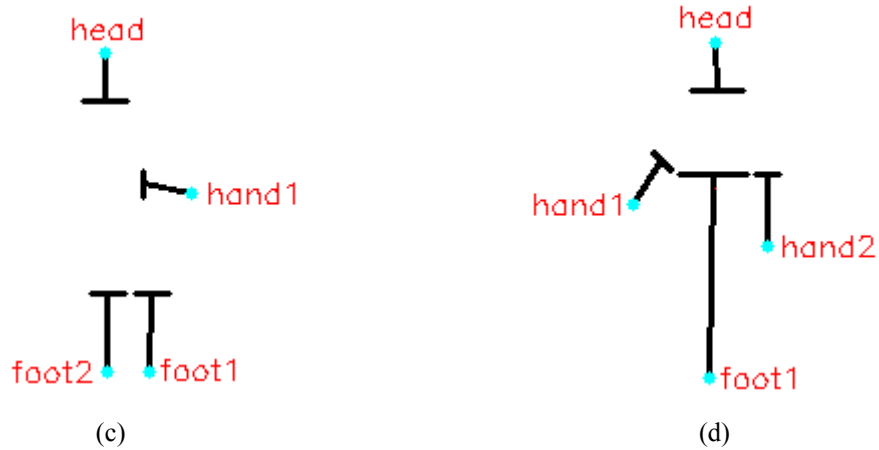


圖 4-16 (續)

### 4.2.3 基於 3 個肢體末端條件下的作法

由於 3 個肢體和 4 個肢體一樣，都會分做有沒有頭的情形，因此我們利用 4.2.2 節的方法將其分類，不過由於 1 頭 0 手 2 腳的情形有時候會發生在雙手向身體靠攏，使得頭部所找出來的角度和腳的角度相異不大，如圖 4-17(a)，因此實際會有(1)有頭的情況，包含 1 頭 1 手 1 腳、1 頭 2 手 0 腳、1 頭 0 手 2 腳和(2)沒有頭的情況，包含 1 頭（連軀幹）0 手 2 腳、0 頭 2 手 1 腳、0 頭 1 手 2 腳，這兩種情形。

#### (1)有頭的情況

在這一類中我們利用 4.2.2 節分類 1 頭 2 手 1 腳和 1 頭 1 手 2 腳的判斷方式就可以識別出 1 頭 1 手 1 腳的情形，如圖 4-17(b)。因此接著就是識別出我們所找出的 2 個肢體究竟是手還是腳就可以了，而我們識別的方法就是利用頭到腳的距離和人體區域的長度是差不多的來進行識別。首先利用 4.2.2 節中的方式先判斷現在是直立還是躺的姿勢，如果是直立的姿勢那就使用  $0^\circ$  掃描線最上面的橫截面中點到最下面的橫截面中點的距離來當作參考距離，如果是躺的姿勢那就使用  $90^\circ$  掃描線最右邊的橫截面中點到最左邊的橫截面中點的距離來當作參考距離，最後判斷頭到肢體的最遠距離和參考距離比值是否大於門檻值  $T_{foot}$ ，如果是就是 1 頭 0 手 2 腳，如圖 4-17(c)。否則就是 1 頭 2 手 0 腳，如圖 4-17(d)。

## (2)沒有頭的情況

而第(2)類因為會有 1 頭（連軀幹）0 手 2 腳的情形出現，所以我們使用區域分割來識別頭。同樣地我們先區分它是直立還是躺的姿勢，如果是直立的姿勢，則使用  $0^\circ$  掃描線最上面的橫截面  $y$  值到最下面的橫截面  $y$  值的距離來當作參考距離，如果末端特徵的  $y$  值在最上面的橫截面  $y$  值的  $1/10$  參考距離範圍內就是頭。如果是躺的姿勢那就使用  $90^\circ$  掃描線最右邊的橫截面  $x$  值到最左邊的橫截面  $x$  值的距離來當作參考距離，如果末端特徵的  $x$  值在最右邊的橫截面  $x$  值的  $1/10$  參考距離範圍內就是頭。如果在此範圍內有找到頭，就是 1 頭 0 手 2 腳的情形，如圖 4-17(e)。如果是 0 頭 2 手 1 腳、0 頭 1 手 2 腳就利用 4.2.2 節判斷 0 頭 2 手 2 腳的方式及分類 1 頭 2 手 1 腳和 1 頭 1 手 2 腳的方式來判斷，在此不多加敘述，如圖 4-17(f)。以下將我們的做法以條列式敘述：

- (1) 找出 3 個肢體中，不含多節肢體中角度最大和最小的近似中軸線，如果最大和最小的角度差值小於  $T_{headless}$ ，則跳至步驟(2)，否則最大角度的就是頭，跳至步驟(4)。
- (2) 判斷人體區域的長寬比是否大於  $T_{lie}$ ，如果是則使用  $0^\circ$  掃描線最上面的橫截面  $y$  值到最下面的橫截面  $y$  值的距離來當作參考距離，如果末端特徵的  $y$  值在最上面的橫截面  $y$  值的  $1/10$  參考距離範圍內就是頭。否則使用  $90^\circ$  掃描線最右邊的橫截面  $x$  值到最左邊的橫截面  $x$  值的距離來當作參考距離，如果末端特徵的  $x$  值在最右邊的橫截面  $x$  值的  $1/10$  參考距離範圍內就是頭。如果有找到頭，則剩下兩個肢體就是腳。
- (3) 如果人體區域的長寬比大於  $T_{lie}$ ，則使用  $0^\circ$  掃描線最上面的橫截面中點來當作參考點，否則使用  $90^\circ$  掃描線最右邊的橫截面中點來當作參考點。找到距離參考點最遠的 2 個肢體，如果第二遠跟最遠的比值大於門檻值  $T_{foot}$  就兩個都是腳，否則只有最遠的一個是腳。而剩下的肢體就是手。
- (4) 判斷頭到兩個肢體的距離比值是否小於  $T_{foot}$ ，如果是則最遠的一個是腳，另

一個就是手。否則如果人體區域的長寬比大於 $T_{lie}$ ，使用 $0^\circ$ 掃描線最上面的橫截面中點到最下面的橫截面中點的距離來當作參考距離，否則使用 $90^\circ$ 掃描線最右邊的橫截面中點到最左邊的橫截面中點的距離來當作參考距離，最後判斷頭到肢體的最遠距離和參考距離比值是否大於門檻值 $T_{foot}$ ，如果是就判斷兩個都是腳，否則兩個就是手。

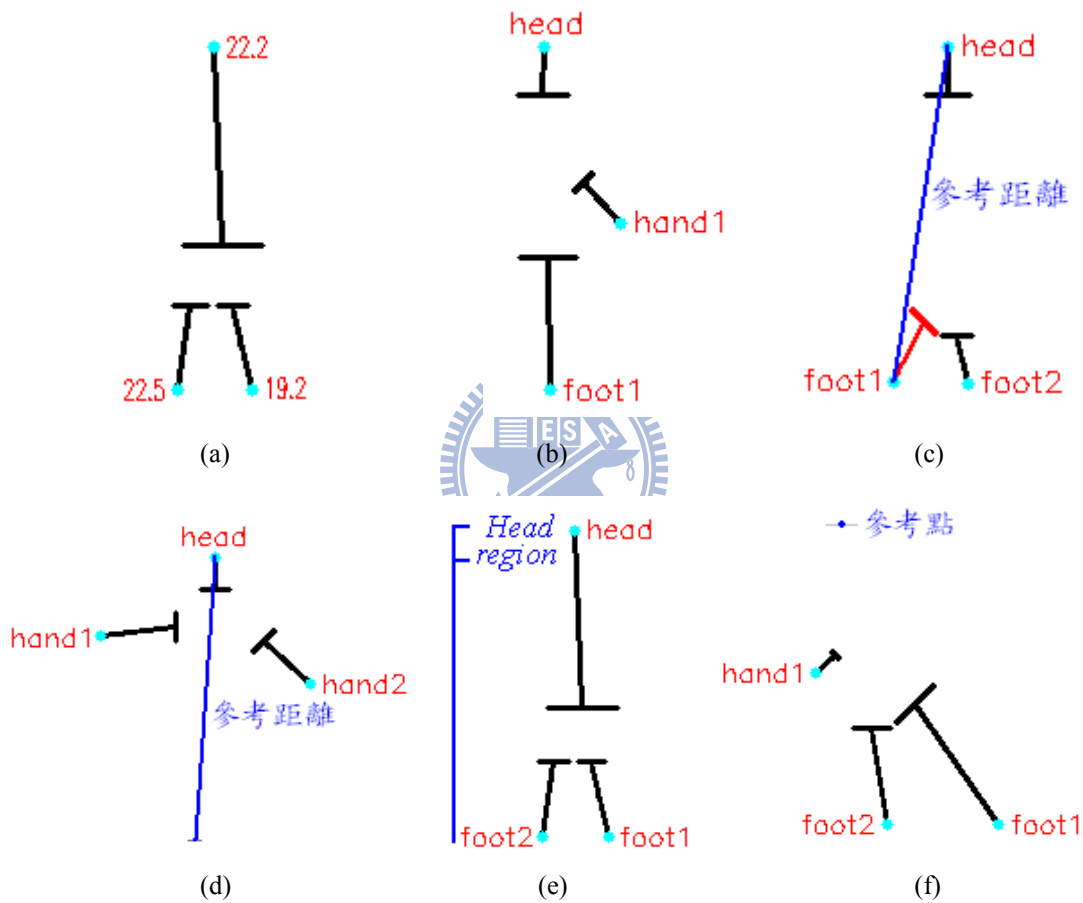


圖 4-17 3 個肢體的識別結果：(a) 1 頭 2 手 2 腳的角度過於接近的例子，(b) 1 頭 1 手 1 腳的識別結果，(c) 1 頭 0 手 2 腳的識別結果，(d) 1 頭 2 手 0 腳的識別結果，(e) (a) 的識別結果，(f) 0 頭 1 手 2 腳的識別結果。

#### 4.2.4 基於 2、1 個肢體末端條件下的作法

由於 2、1 個肢體資訊量太少，因此此階段我們使用像 4.2.3 節中的區域分割技巧來識別頭、手、腳，圖 4-18 為其識別結果。同時，為了方便起見，我們將姿態識別過程中所使用到的參數顯示在表 4-1。以下將我們的做法以條列式敘述：

- (1) 判斷人體區域的長寬比是否大於  $T_{lie}$ ，如果是則代表是直立的，跳至步驟(2)。表示躺著的，跳至步驟(3)。
- (2) 使用  $0^\circ$  掃描線最上面的橫截面  $y$  值到最下面的橫截面  $y$  值的距離來當作參考距離，如果末端特徵的  $y$  值在最上面的橫截面  $y$  值的  $1/10$  參考距離範圍內就是頭。如果在最下面的橫截面  $y$  值的  $3/10$  參考距離範圍內就是腳，剩下就是手。
- (3) 使用  $90^\circ$  掃描線最右邊的橫截面  $x$  值到最左邊的橫截面  $x$  值的距離來當作參考距離，如果末端特徵的  $x$  值在最右邊的橫截面  $x$  值的  $1/10$  參考距離範圍內就是頭。如果在最左邊的橫截面  $x$  值的  $3/10$  參考距離範圍內就是腳，剩下就是手。

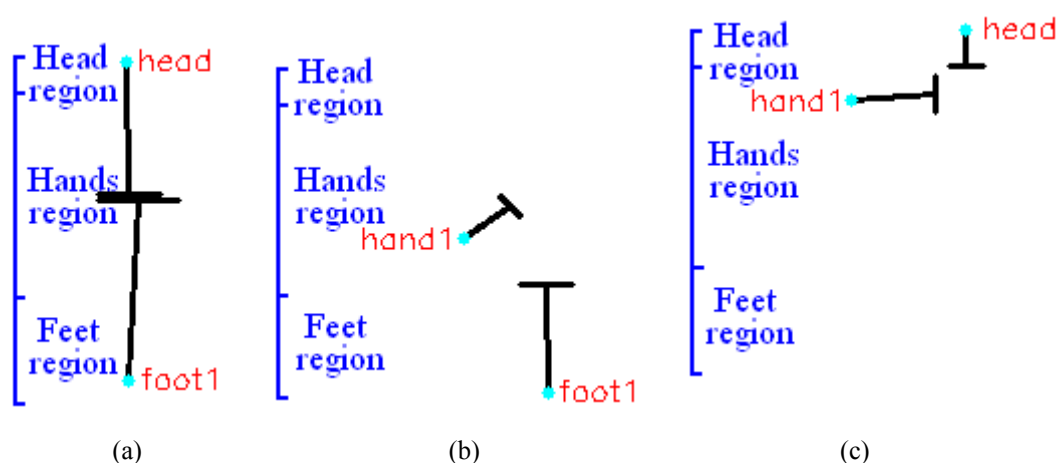


圖 4-18 2 個肢體的識別結果：(a) 1 頭 0 手 1 腳的識別結果，(b) 0 頭 1 手 1 腳的識別結果，(c) 1 頭 1 手 0 腳的識別結果。

表 4-1 姿態識別中所使用到的參數一覽。

參數	說明	應用時機(肢體末端數量)	數值
$T_{headless}$	分辨是否有頭存在	4、3	$10^\circ$
$T_{lie}$	分辨現在是直立的還是躺著的	4、3、2、1	0.7
$T_{foot}$	分辨是 1 隻腳還是 2 隻腳的情形	4、3	0.8

### 4.3 基於頭、手、腳定位之人體模型表示

在[11]、[12]的研究中，分別使用了橢圓、矩形和梯形來表示人體模型，由於居家看護系統中，看護對象往往不希望自己的真實型態被加以記錄，且處理端跟顯示端通常會設置在不同地方，例如在房間裡偵測到人體並加以分析，把結果傳到一個中央控制室做顯示，因此傳輸的資料要越少越好，這會造成傳輸速度較快。因此本論文為了隱密性較高、傳輸較少資料，使用了骨架式的人體模型來代表。圖 4-19(a)前背景圖，圖 4-19(b)為[11]的人體模型，圖 4-19(c)為[12]的人體模型，圖 4-19(d)為本論文所使用的人體模型，紅色代表頭，綠色代表手，藍色代表腳，黑色的點是定位點。就隱密性來說，[12]的方法最差，因為它幾乎會完全呈現出肢體寬度。其次是[11]的方法，因為它所畫的肢體寬度是選取橫截面中最小的，雖然不是完全反應肢體寬度，但是還是容易看出有多瘦。本論文的方法是隱密性最高的，因為它只是顯示骨架，完全看不出這個人的肢體是胖還是瘦。

至於在傳輸資料的多寡方面，因為定位點是每個都有傳的所以不將其加入下面的討論，以 OPENCV 提供的繪圖函數來說，[12]的方法傳輸資料最多，因為一個梯形需要 4 個點才能重繪，1 個點有 2 個值，共 5 個梯形所以有 40 個值。其次是[11]，因為一個橢圓需要中心點、長軸、短軸、旋轉角度，1 個中心點有 2 個值，共 3 個橢圓所以有 15 個值，而矩形只需要 2 個點，1 個點有 2 個值，共 3 個矩形所以有 12 個值，全部加起來是 27 個值。而本論文的方法傳輸的資料最少，因為一個圓需要中心點、半徑，1 個中心點有 2 個值，所以一個圓需要 3 個值，而一條線只要 2 個點，1 個點有 2 個值，共 4 條線所以有 16 個值，全部加起來是 19 個值。表 4-2 顯示了隱密性高低及傳輸資料多寡的比較。

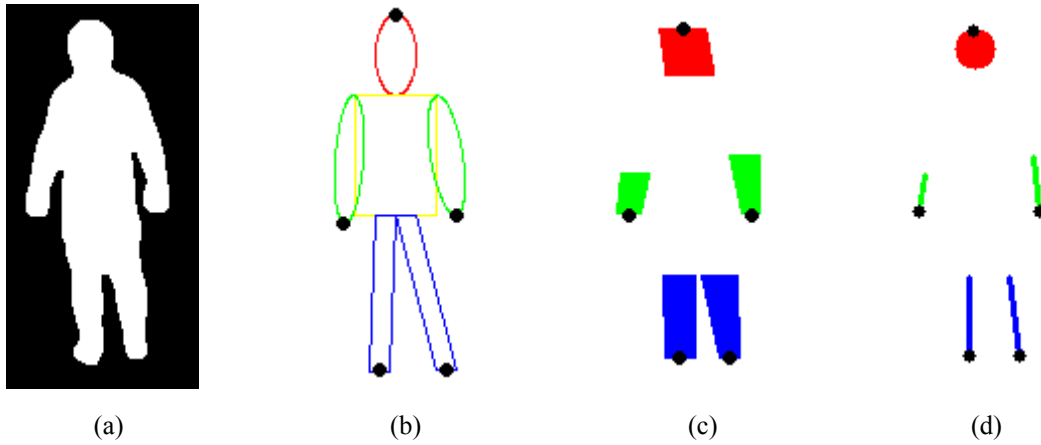
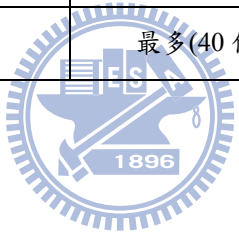


圖 4-19 人體模型表示結果：(a) 前背景圖，(b) [11]的人體模型表示，(c) [12]的人體模型表示，(d) 本論文的人體模型表示。

表 4-2 人體模型的隱密性及傳輸資料比較。

	[11]的人體模型	[12]的人體模型	本論文的人體模型
隱密性	中	低	高
傳輸資料	中(27 個值)	最多(40 個值)	最少(19 個值)





## 五、實驗結果及討論

在本章節中，我們利用實作出來的姿態識別系統來進行實際影像畫面的實驗，並證明本論文的系統具有較佳的正確性與實用性。整個實驗的硬體環境為 Pentium 4 3.40GHz，作業系統為 Windows XP，程式作業環境為 Microsoft Visual Studio 2005 的平台，在這個平台下所能達到的平均效能為每秒 13 張影像。輸入來源可以是影片、影像或是 raw 檔，解析度為 320x240。如圖 5-1 和圖 5-2 分別是中紅外線影像和可見光影像的識別結果。為了能多顯示一些實驗結果，我們接下來只在一開始說明該段影片之背景環境，之後就只顯示前背景圖及識別結果，同時為了顯示每段的識別正確率，我們會在最後顯示出該段影片的識別正確率。

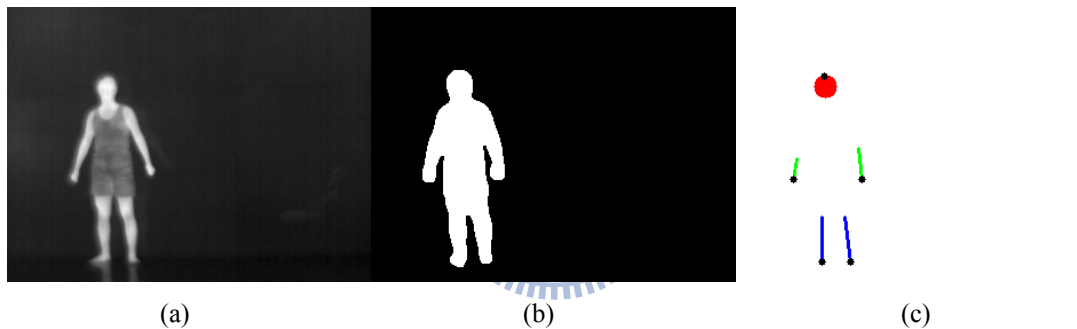


圖 5-1 中紅外線影像姿態識別結果：(a) 中紅外線影像，(b) 前背景圖，(c) 姿態識別的結果。

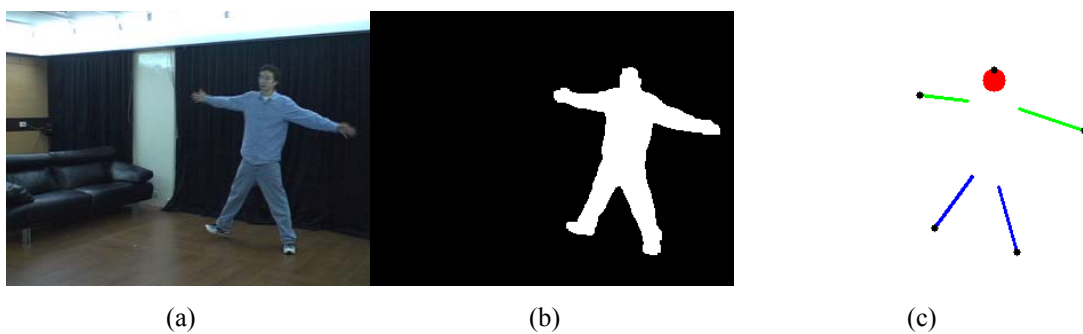
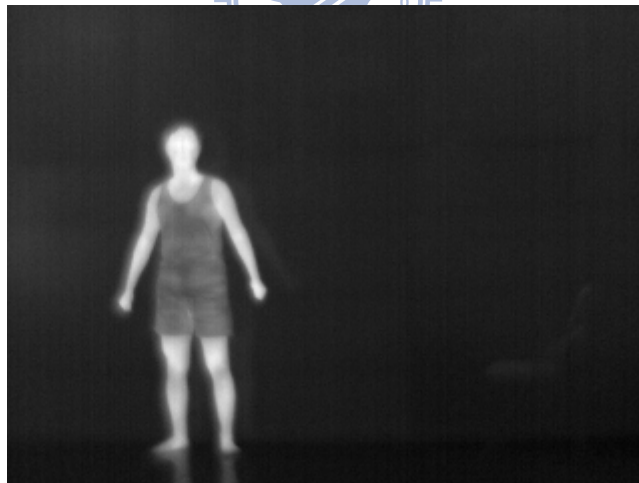


圖 5-2 可見光影像姿態識別結果：(a) 可見光影像，(b) 前背景圖，(c) 姿態識別的結果。

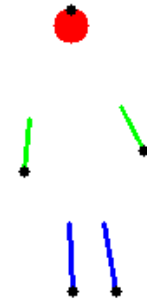
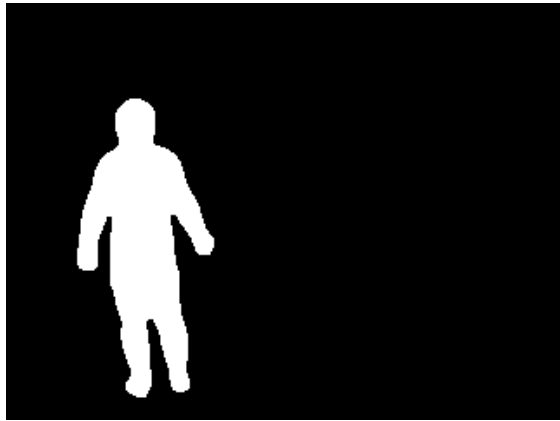


如圖 5-3 為一人像走動並坐下的連續中紅外線影像及姿態識別後所產生的人體模型，圖 5-3(a)為其背景環境，從圖中我們可以看到由於中紅外線顯示的是溫度，所以背景大致上單純，也因此造成這段影像中的前背景分割結果都極佳，另外由圖中我們也可以看出這一連續的影像都能對姿態有不錯的識別，所以我們可以說，在前背景分割佳的條件下，本論文的姿態識別能有較不錯的識別度。表 5-1 為這段影像的識別結果之正確率，其中肢體種類指的是可能出現的情況，而後面的數字代表的是一多少張影像是正確的/多少張這種肢體種類的影像，各個肢體標記正確率指的是各個肢體的正確率，紅色的代表頭，綠色的代表手，藍色的代表腳，黑色的是三個的總和，而後面的數字代表的是一多少個部位是正確的/多少個這種部位，整個影像的正確率指的是，一張影像中全部的肢體部位都識別正確的有多少，而後面的數字代表的是一多少張影像是完全識別正確的/多少張影像是屬於這一類的。

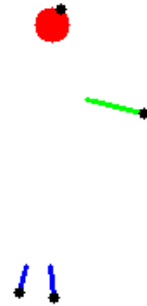


(a)

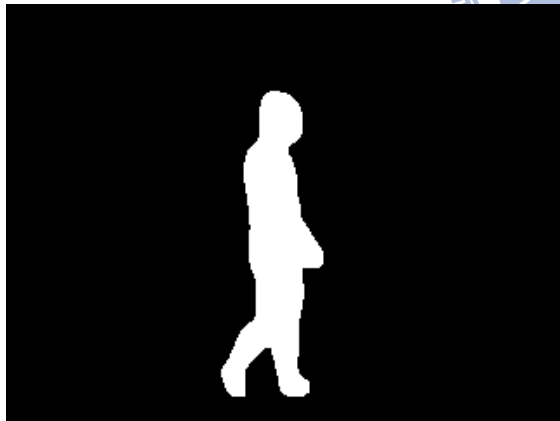
圖 5-3 人像走動並坐下的連續中紅外線影像及姿態識別後所產生的人體模型。



(b)



(c)



(d)

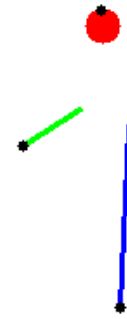
圖 5-3 (續)



(e)



(f)



(g)



圖 5-3 (續)

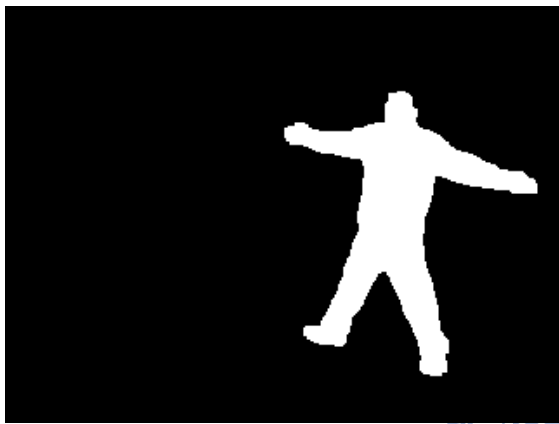
表 5-1 人像走動並坐下的連續中紅外線影像的姿態識別正確率。

	肢體種類	各個肢體標記正確率	整個影像正確率
5 個肢體	1 頭 2 手 2 腳 58/58	58/58 = 100% 116/116 = 100% 116/116 = 100% 290/290 = 100%	58/58 = 100%
4 個肢體	1 頭 1 手 2 腳 43/47 1 頭 2 手 1 腳 1/2 0 頭 2 手 2 腳 0/0	45/49 = 91.83% 49/51 = 96.07% 92/96 = 95.83% 186/196 = 94.89%	44/49 = 89.79%
3 個肢體	1 頭 0 手 2 腳 57/61 1 頭 1 手 1 腳 29/33 1 頭 2 手 0 腳 0/0 0 頭 1 手 2 腳 0/0 0 頭 2 手 1 腳 0/0	86/94 = 91.48% 32/33 = 96.96% 147/155 = 96.15% 265/282 = 93.07%	86/94 = 91.48%
2、1 個肢體	1 頭 1 手 0 腳 0/0 1 頭 0 手 1 腳 54/54 0 頭 1 手 1 腳 5/5 0 頭 2 手 0 腳 0/0 0 頭 0 手 2 腳 0/0 1 頭 0 手 0 腳 0/0 0 頭 1 手 0 腳 0/0 0 頭 0 手 1 腳 0/1	54/54 = 100% 5/5 = 100% 60/60 = 100% 119/119 = 100%	60/60 = 100%
<b>Total</b>		243/255 = 95.29% 202/205 = 98.53% 415/427 = 97.18% 860/887 = 96.95%	248/261 = 95.01%

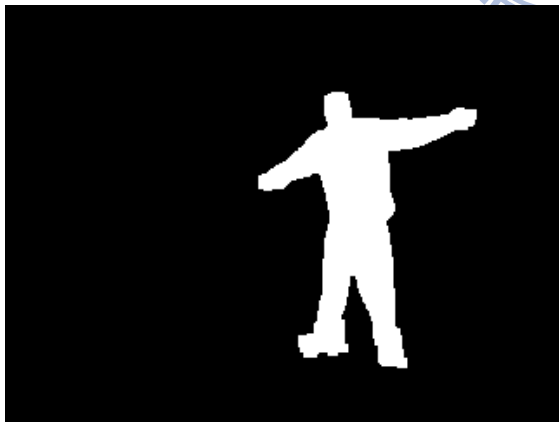
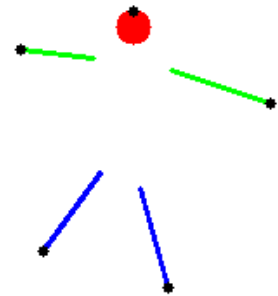
圖 5-4 則是一人像呈現大字形走動的可見光影像之影片及姿態識別後所產生的人體模型，圖 5-4(a) 為其背景環境，由圖中我們可以看到背景和人物之衣服顏色相異較易區分，唯一造成分割困擾的是腳部的影子有時候會過深以至於無法濾除，使得影響到辨別結果。由圖中可以看出就算前背景分割結果尚可，但是肢體部位絕大部分的有出現的情況下，也就是在 4、5 個肢體末端的姿態識別中，本論文的效果極佳。表 5-2 為這段影像的識別結果之正確率。



(a)



(b)



(c)

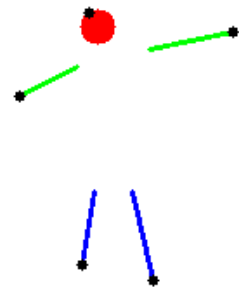
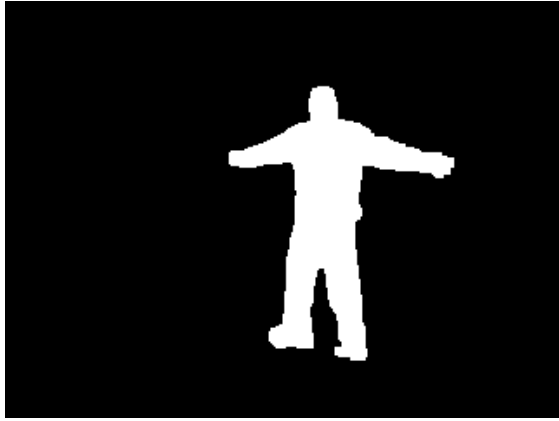
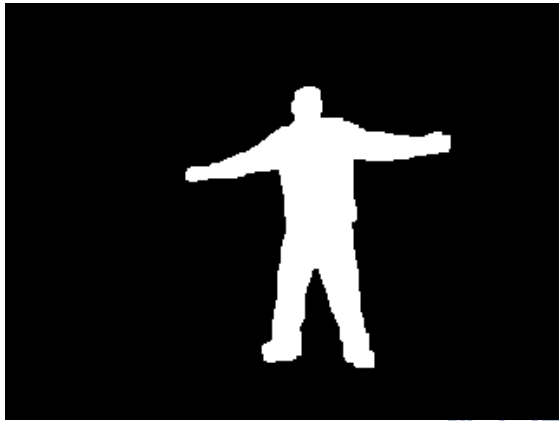
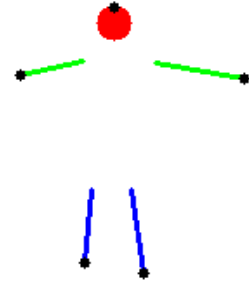


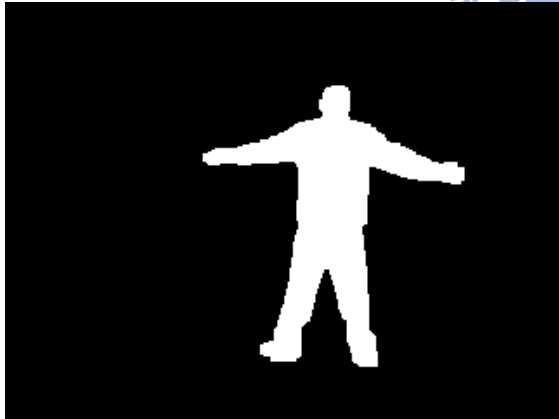
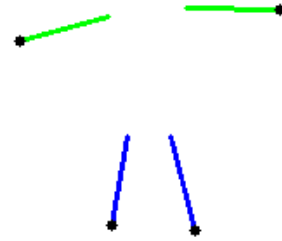
圖 5-4 人像呈大字形走動的連續可見光影像及姿態識別後所產生的人體模型。



(d)



(e)



(f)

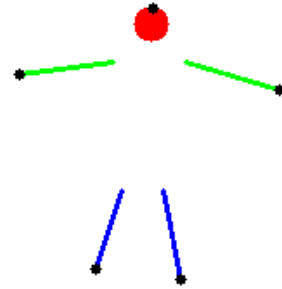
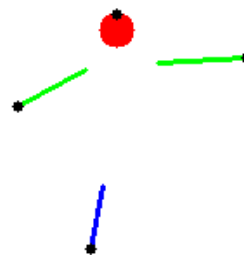
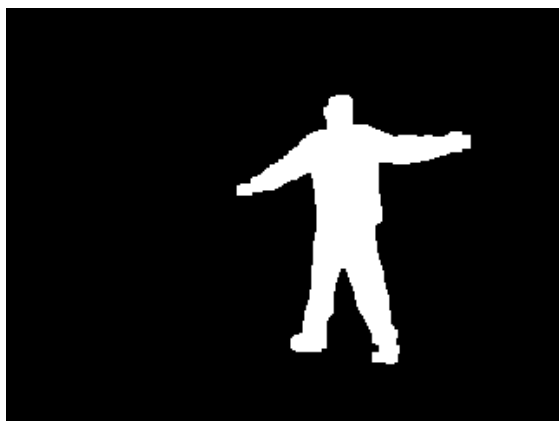


圖 5-4 (續)



(f)

圖 5-4 (續)

表 5-2 人像呈大字形走動的連續可見光影像的姿態識別正確率。

	肢體種類	各個肢體標記正確率	整個影像正確率
5 個肢體	1 頭 2 手 2 腳 51/51	$51/51 = 100\%$ $102/102 = 100\%$ $102/102 = 100\%$ $255/255 = 100\%$	$51/51 = 100\%$
4 個肢體	1 頭 1 手 2 腳 0/0 1 頭 2 手 1 腳 2/2 0 頭 2 手 2 腳 3/3	$2/2 = 100\%$ $10/10 = 100\%$ $8/8 = 100\%$ $20/20 = 100\%$	$5/5 = 100\%$
3 個肢體	1 頭 0 手 2 腳 0/0 1 頭 1 手 1 腳 0/0 1 頭 2 手 0 腳 0/0 0 頭 1 手 2 腳 0/0 0 頭 2 手 1 腳 0/0	$271/312 = 86.85\%$ $148/200 = 74\%$ $462/541 = 85.39\%$ $881/1053 = 83.66\%$	$0/0 = 0\%$
2、1 個肢體	1 頭 1 手 0 腳 0/0 1 頭 0 手 1 腳 0/0 0 頭 1 手 1 腳 0/0 0 頭 2 手 0 腳 0/0 0 頭 0 手 2 腳 0/0 1 頭 0 手 0 腳 0/0 0 頭 1 手 0 腳 0/0 0 頭 0 手 1 腳 0/0	$0/0 = 0\%$ $0/0 = 0\%$ $0/0 = 0\%$ $0/0 = 0\%$	$0/0 = 0\%$
<b>Total</b>		$53/53 = 100\%$ $112/112 = 100\%$ $110/110 = 100\%$ $275/275 = 100\%$	$56/56 = 100\%$

圖 5-5 則是一人像彎腰、蹲下及躺下的連續影像及姿態識別後所產生的人體模型，圖 5-5(a)為其背景環境，由圖中我們可以看到背景和人物之衣服顏色相異較易區分，但是影子有時候會過深以至於無法濾除，且加上鞋子跟地板之顏色過近還有地板會反光等影響，都會使得辨別結果不佳。由圖中可以看出前背景分割較不佳且產生較多遮蔽的情況下，本論文的識別度較沒有前兩個高，但仍然有不錯的表現。表 5-3 為這段影像的識別結果之正確率。表 5-4 為測試五段(包含上述三段)性質相異、不同姿態的影片所得到的識別結果的正確率。由表中我們知道本系統的平均正確率高達八成，因此可以稱的上是一個有效的姿態識別系統。表 5-5 為本論文與[11]、[12]的差異比較，由表中我們可以看出本論文不僅將[11]、[12]的掃描線演算法做一改良及優化，也提供了傳輸資料量較少及隱密性較高之人體模型，更重要的是可以將頭、手、腳定位，這是[11]、[12]所做不到的。雖然比[11]、[12]多完成了頭、手、腳定位，然而本論文的執行速度還是最快的。



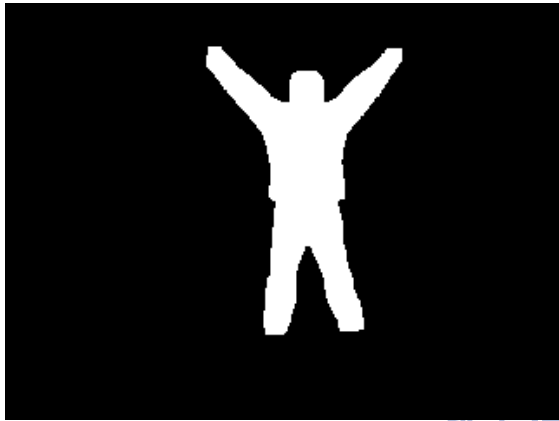
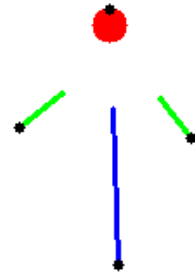
(a)

圖 5-5 人像彎腰、蹲下及躺下的連續影像及姿態識別後所產生的人體模型。

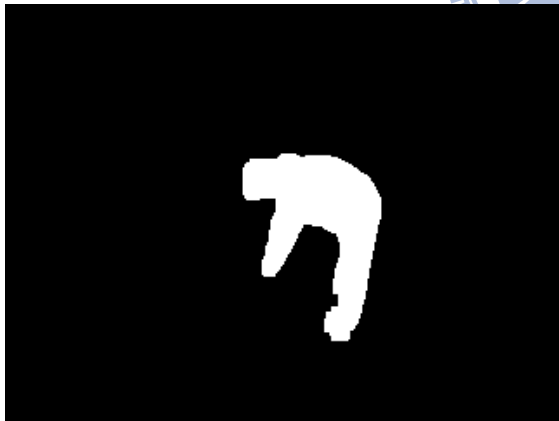
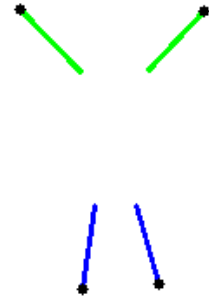




(b)



(b)



(c)

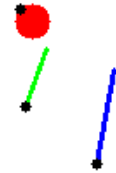
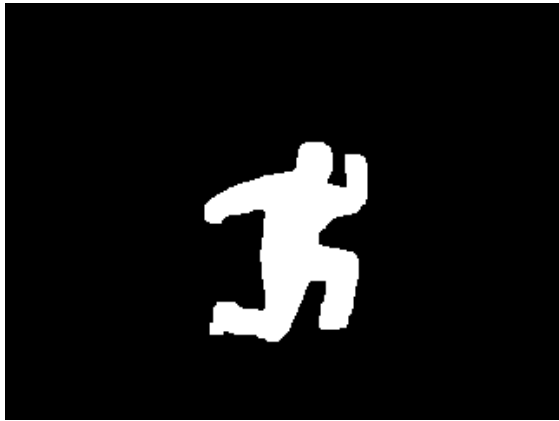
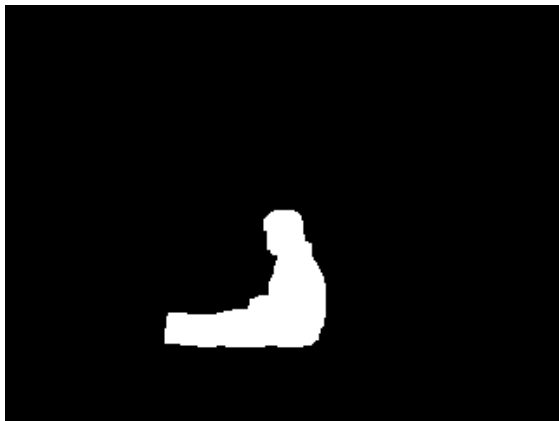


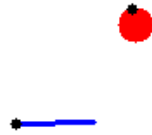
圖 5-5 (續)



(d)



(e)



(f)



圖 5-5 (續)

表 5-3 人像彎腰、蹲下及躺下的連續影像的姿態識別正確率。

	肢體種類	各個肢體標記正確率	整個影像正確率
5 個肢體	1 頭 2 手 2 腳 115/141	$124/141 = 87.94\%$ $233/282 = 82.62\%$ $236/282 = 83.68\%$ $593/705 = 84.11\%$	$115/141 = 81.56\%$
4 個肢體	1 頭 1 手 2 腳 51/60 1 頭 2 手 1 腳 6/9 0 頭 2 手 2 腳 4/4	$59/73 = 80.82\%$ $69/86 = 80.23\%$ $122/137 = 89.05\%$ $250/296 = 84.45\%$	$61/73 = 83.56\%$
3 個肢體	1 頭 0 手 2 腳 41/48 1 頭 1 手 1 腳 62/70 1 頭 2 手 0 腳 0/0 0 頭 1 手 2 腳 0/0 0 頭 2 手 1 腳 0/4	$103/118 = 87.28\%$ $70/78 = 89.74\%$ $147/170 = 86.47\%$ $320/366 = 87.43\%$	$103/122 = 84.42\%$
2、1 個肢體	1 頭 1 手 0 腳 0/1 1 頭 0 手 1 腳 46/72 0 頭 1 手 1 腳 5/5 0 頭 2 手 0 腳 0/2 0 頭 0 手 2 腳 0/0 1 頭 0 手 0 腳 7/7 0 頭 1 手 0 腳 0/0 0 頭 0 手 1 腳 18/18	$56/80 = 70\%$ $4/10 = 40\%$ $93/95 = 97.89\%$ $153/185 = 82.7\%$	$76/105 = 72.38\%$
<b>Total</b>		$342/412 = 83\%$ $376/456 = 82.45\%$ $598/684 = 87.42\%$ $1316/1552 = 84.79\%$	$355/441 = 80.49\%$

表 5-4 識別結果正確率。

	肢體種類	各個肢體標記正確率	整個影像正確率
5 個肢體	1 頭 2 手 2 腳 345/386	$354/386 = 89.37\%$ $701/772 = 90.8\%$ $706/772 = 91.45\%$ $1761/1930 = 91.24\%$	$345/386 = 89.37\%$
4 個肢體	1 頭 1 手 2 腳 132/153 1 頭 2 手 1 腳 19/30 0 頭 2 手 2 腳 21/43	$152/180 = 84.44\%$ $228/299 = 76.25\%$ $354/422 = 81.75\%$ $734/901 = 81.46\%$	$172/226 = 76.1\%$
3 個肢體	1 頭 0 手 2 腳 137/155 1 頭 1 手 1 腳 133/157 1 頭 2 手 0 腳 0/0 0 頭 1 手 2 腳 1/35 0 頭 2 手 1 腳 0/4	$271/312 = 86.85\%$ $148/200 = 74\%$ $462/541 = 85.39\%$ $881/1053 = 83.66\%$	$271/351 = 77.2\%$
2、1 個肢體	1 頭 1 手 0 腳 0/1 1 頭 0 手 1 腳 133/160 0 頭 1 手 1 腳 13/17 0 頭 2 手 0 腳 0/2 0 頭 0 手 2 腳 2/2 1 頭 0 手 0 腳 8/9 0 頭 1 手 0 腳 0/0 0 頭 0 手 1 腳 23/23	$147/170 = 84.7\%$ $1/22 = 54.54\%$ $201/204 = 98.52\%$ $357/396 = 90.15\%$	$179/214 = 83.64\%$
<b>Total</b>		$921/1048 = 87.88\%$ $1089/1293 = 84.22\%$ $1723/1939 = 88.86\%$ $3733/4280 = 87.21\%$	$967/1177 = 82.15\%$

表 5-5 本論文與[11]、[12]差異之比較。

	本論文	[11]	[12]
掃描線演算法	能同時濾除橫截面寬度過小之雜訊且將有空洞之部份補起來	僅能濾除橫截面寬度過小之雜訊	僅能濾除橫截面寬度過小之雜訊
人體模型	傳輸資料量最少且隱密性最高	傳輸資料量次少且隱密性次高	傳輸資料量最多且隱密性最低
姿態識別	不僅能表示人體模型，還能將頭、手、腳定位	僅能表示人體模型	僅能表示人體模型
處理速度	最快，13 frames/s	次快，11 frames/s	最慢，9 frames/s

## 六、結論及未來工作

在本篇論文中，我們將[11]、[12]的研究做一整合及優化，使得我們可以用隱密性高及較少傳輸資料的人體模型來表示，並更進一步地利用肢體的資訊來識別出頭、手、腳，找到頭、手、腳的定位點，以提供給行為分析和姿勢分類的研究者做更細部動作之分析。

本篇論文所採用的識別方式，僅是針對單張影像，因此如果有一些肢體遮蔽或是前景形狀不完整的情形發生時，就會造成識別錯誤，如圖 6-1 為連續四張影像之識別結果，我們可以清楚地看到因為手逐漸往身體靠攏，造成原本圖 6-1(c) 是手的被識別成頭，這顯然是不對的，因此我們可以加上時間軸的資訊，例如 tracking 等技巧來避免這些錯誤。

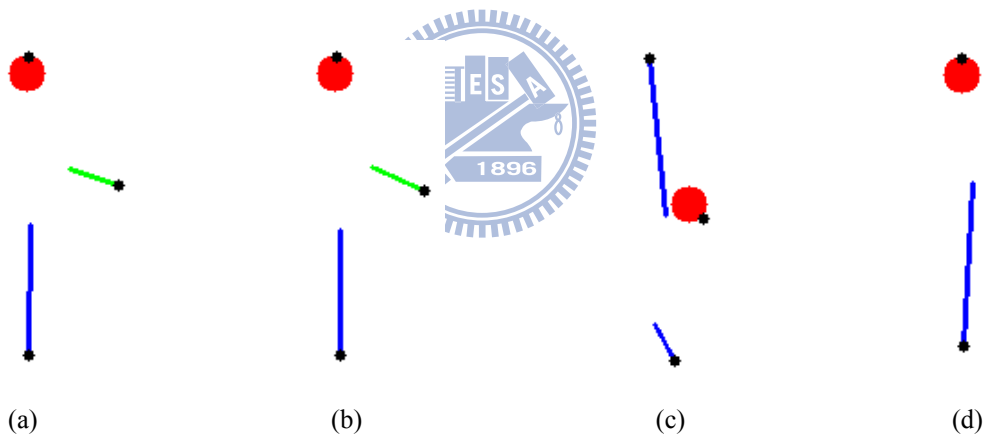


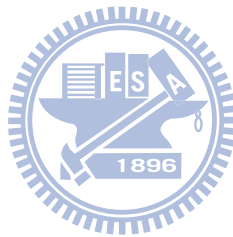
圖 6-1 連續影像的姿態識別結果：(a) 第  $t-2$  張影像，(b) 第  $t-1$  張影像，(c) 第  $t$  張影像，(d) 第  $t+1$  張影像。

同時，本論文目前之研究方向皆是在 2D 影像，造成有些遮蔽現象無法更清楚的識別到，因此藉由使用具有深度資訊的攝影機或是多架攝影機來延伸至 3D 影像也是我們可以持續研究的課題。而藉由姿態識別我們也可以延伸應用到姿勢分類及行為分析，甚者可以做更精細的姿態識別，例如手腕、手肘、肩膀、髖骨、膝蓋、腳踝等關節點之定位等，也是未來我們可以挑戰的研究方向。

## 參考文獻

- [1] 行政院衛生署，92 年人口統計資料，<http://www.doh.gov.tw/statistic/生命統計/92.htm>，Dec. 2004.
- [2] Chia-Feng Juang and Chia-Ming Chang, "Human Body Posture Classification by a Neural Fuzzy Network and Home Care System Application," *IEEE Transactions on Systems, Man and Cybernetics, Part A*, Volume 37, Issue 6, pp. 984-994, Nov. 2007.
- [3] Chen-Chiung Hsieh and Shu-Shuo Hsu, "A Simple and Fast Surveillance System for Human Tracking and Behavior Analysis," *Third International IEEE Conference on Signal-Image Technologies and Internet-Based System*, pp. 812-818, Dec. 2007.
- [4] Caroline Rougier, Jean Meunier, Alain St-Arnaud, and Jacqueline Rousseau, "Fall Detection from Human Shape and Motion History Using Video Surveillance," *21st International Conference on Advanced Information Networking and Applications Workshops*, Volume 2, pp. 875-880, May 2007.
- [5] Ciro Castiello, Tiziana D'Orazio, Anna Maria Fanelli, Paolo Spagnolo, and Maria Alessandra Torsello, "A Model-free Approach for Posture Classification," *IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 276-281, Sept. 2005.
- [6] Jun-Wei Hsieh, Yung-Tai Hsu, Liao, Hong-Yuan Mark Liao, and Chih-Chiang Chen, "Video-based Human Movement Analysis and Its Application to Surveillance Systems," *IEEE Transactions on Multimedia*, Volume 10, Issue 3, pp. 372-384, April 2008.
- [7] Ping Guo and Zhen-Jiang Miao, "Projection Histogram Based Human Posture Recognition," *The 8th International Conference on Signal Processing*, Volume 2, pp. 16-20, 2006.
- [8] Frederic Jean, Robert Bergevin, and Alexandra Branzan Albu, "Body Tracking in Human Walk from Monocular Video Sequences," *The 2nd Canadian Conference on Computer and Robot Vision*, pp. 144-151, May 2005.
- [9] Ismail Haritaoglu, David Harwood, and Larry Davis, "W<sup>4</sup>: A Real Time System for Detecting and Tracking People," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 962-962, June 1998.
- [10] Chia-Feng Juang, Chia-Ming Chang, Jiu-Hou Wu, and Demei Lee, "Computer Vision-based Human Body Segmentation and Posture Estimation," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, Volume 39, Issue 1, pp. 119-133, Jan. 2009.
- [11] Jen-Hui Chuang, Chun-Wei Lee, and Kuo-Hua Lo, "Human Activity Analysis

- Based on a Torso-less Representation,” *19th International Conference on Pattern Recognition*, pp. 1-4, Dec. 2008.
- [12] Ning-Hsuan Chu, “A Torso-less Representation of Human Activity with Multi-joint Trapezoids,” Master’s Thesis, Dept. Computer Science and Engineering, Nation Chiao Tung University, June 2008.
- [13] Jau-Hong Kao, Jen-Hui Chuang, and Tsaipei Wang, “Chromosome Classification Based on the Band Profile Similarity along Approximate Medial Axis,” *Pattern Recognition*, Volume 41, pp. 77-89, Jan. 2008.
- [14] Rita Cucchiara, Costantino Grana, Massimo Piccardi, Andrea Prati, and Stefano Sirotti, “Improving Shadow Suppression in Moving Object Detection with HSV Color Information,” *Proceedings of IEEE Intelligent Transportation Systems*, pp. 334-339, Aug. 2001.
- [15] Yu-Hsin Chen, “Visual Surveillance Using Infrared Images,” Master’s Thesis, Dept. Computer Science and Engineering, Nation Chiao Tung University, June 2004.



## 附錄 A、連續橫截面的中心點是否相連

本附錄旨在論述 3.2 節中決定兩個連續橫截面的中心點是否相連的條件。首先，因為人體會有呈現分岔處的情形，因此我們不能直接將兩個連續橫截面予以連接，如圖 A-1 為一人體腰部至腿部的局部橫截面示意圖，就是呈現分岔處的例子。如果我們直接將兩連續橫截面予以連接，會造成代表軀幹的褐色近似中軸線跟代表腿部的紅色或粉紅色近似中軸線相連，這顯然是不正確的，因為軀幹和腿並不屬於同一肢體部位，所以我們必須經由一些條件來將其做明顯的區分。從圖中我們可以清楚地看到，決定兩橫截面是否連接的條件跟  $\theta$  有關，當  $\theta$  越大表示兩橫截面的距離越遠，因此  $\theta$  的選取就是我們接下來要討論的議題。 $\theta$  代表的是掃描線垂直方向與兩橫截面中心點連線的夾角，在理想的情況下，也就是寬度完全均勻的長條狀物體在掃描時，最佳掃描方向將滿足  $0^\circ \leq \theta \leq 22.5^\circ$ 。然而，由於人體各個肢體皆非完全均勻的長條狀物體，加上有可能因為雜訊而造成橫截面連接不如預期，因此我們為  $\theta$  設定一個容許範圍，讓別的掃描方向也有機會將我們因為雜訊所失去的資訊予以補足，所以我們將  $\theta$  設定在  $45^\circ$  以內。經由上述的條件限制後，圖 A-1 的三條近似中軸線就會被區分為不同的肢體部位了。

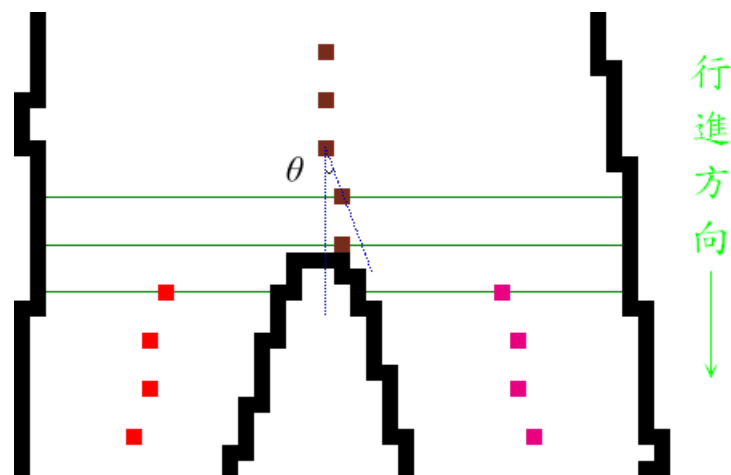


圖 A-1 人體腰部至腿部的局部橫截面示意圖。



另外，經由觀察，在頭部與軀幹交接處，由頭部所產生的橫截面寬度與軀幹所產生的橫截面寬度將會產生極大的差值，如圖 A-2 所示，因此利用橫面寬度的差值，可以用來判斷中心點是否隸屬於同一肢體部位。經由實驗，我們發現差值只要大於 5 就可視為是不同的肢體部位，如圖 A-2 經過此條件篩選後，就會將頭部的近似中軸線和軀幹的近似中軸線分類為不同的肢體部位。

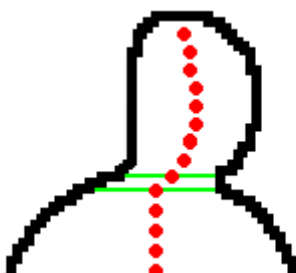


圖 A-2 人體頭部至軀幹的局部橫截面示意圖。



## 附錄 B、判斷線段是否相交

本附錄旨在論述 4.1.2 節中如何利用線段相交的方式來達到判斷兩近似中軸線段是否面積重疊之判斷。首先，我們藉由相交的線段其端點互為相異方向的特性來判斷線段是否相交，如圖 B-1(a)，點  $A$  和點  $B$  分別在線段  $CD$  的延伸線（綠線）的順時針及逆時針方向，但是點  $C$  和點  $D$  都在線段  $AB$  的延伸線（綠線）的順時針方向，所以線段  $AB$  和線段  $CD$  沒有相交，而圖 B-1(b)，不僅點  $A'$  和點  $B'$  分別在線段  $C'D'$  的延伸線（綠線）的順時針及逆時針方向，同時點  $C'$  和點  $D'$  也分別在線段  $A'B'$  的延伸線（綠線）的逆時針及順時針方向，所以線段  $A'B'$  和線段  $C'D'$  有相交。

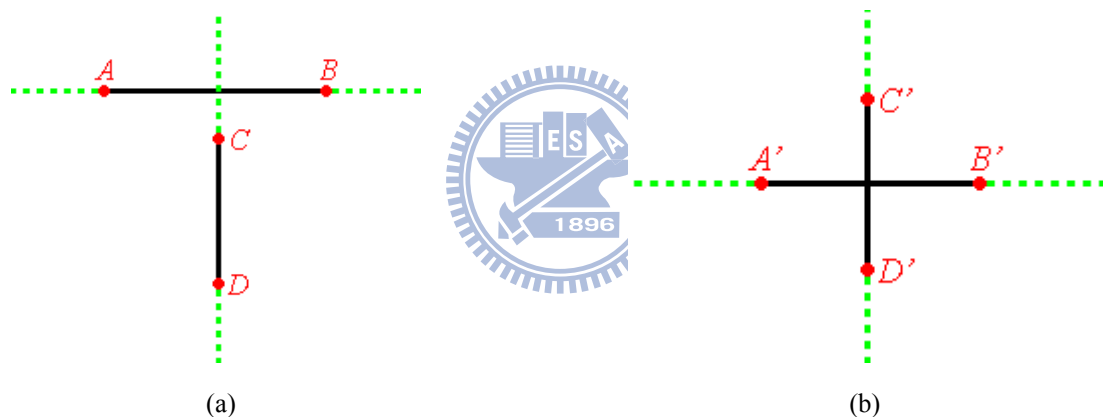


圖 B-1 兩線段是否相交的例子：(a) 沒有相交，(b) 有相交。

首先，我們利用 CCW (Counter Clockwise Turn) 演算法來判斷點是在線段的延伸線的哪一個方向，假設一條線段  $P_1P_2$  及一點  $P_3$ ，主要做法如下：

- (1) 定義向量  $V_1$  為  $P_2 - P_1$ ，向量  $V_2$  為  $P_3 - P_1$ 。
- (2) 計算向量  $V_1$ 、 $V_2$  之外積值。
- (3) 如果外積值為正，即代表點  $P_3$  在線段  $P_1P_2$  的延伸線的逆時針方向，如圖 B-2(a) 所示。
- (4) 如果外積值為負，即代表點  $P_3$  在線段  $P_1P_2$  的延伸線的順時針方向，如圖 B-2(b)

所示。

- (5) 如果外積值為零，即代表點 $P_3$ 在線段 $P_1P_2$ 的延伸線上，如圖B-2(c)所示。

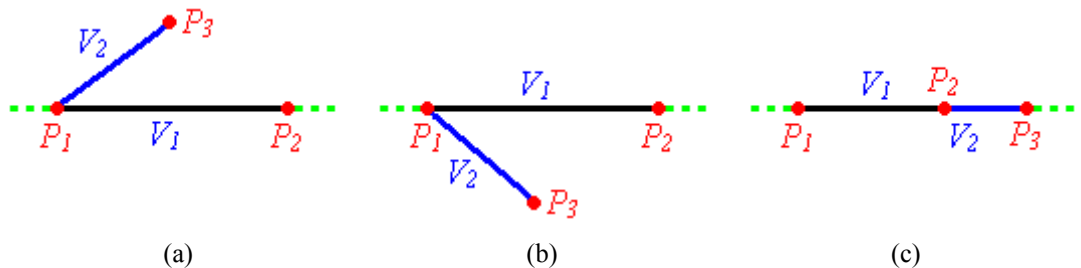


圖 B-2 點在線段的延伸線的哪個方向：(a) 逆時針方向，(b) 順時針方向，(c) 在線段的延伸線上。

接下來我們利用CCW來判斷兩線段的端點是否互為相異方向，假設兩線段 $P_1P_2$ 、 $P_3P_4$ ，主要做法如下：

- (1) 利用CCW判斷點 $P_3$ 在線段 $P_1P_2$ 的延伸線的哪個方向，接著再判斷點 $P_4$ 在線段 $P_1P_2$ 的延伸線的哪個方向，如果點 $P_3$ 和 $P_4$ 在線段 $P_1P_2$ 的延伸線的相異方向，跳至步驟(2)判斷是圖B-3(a)還是圖B-3(b)的情況。如果點 $P_3$ 和 $P_4$ 在線段 $P_1P_2$ 的延伸線的不同方向，則代表線段 $P_1P_2$ 、 $P_3P_4$ 沒有相交，如圖B-3(c)。如果點 $P_3$ 和 $P_4$ 有一點在線段 $P_1P_2$ 的延伸線上，則跳至步驟(2)判斷是圖B-3(d)還是圖B-3(e)的情況。
- (2) 利用CCW判斷點 $P_1$ 在線段 $P_3P_4$ 的延伸線的哪個方向，接著再判斷點 $P_2$ 在線段 $P_3P_4$ 的延伸線的哪個方向，如果點 $P_1$ 和 $P_2$ 在線段 $P_3P_4$ 的延伸線的相異方向，即代表線段 $P_1P_2$ 、 $P_3P_4$ 相交，如圖B-3(a)、B-3(e)。如果點 $P_1$ 和 $P_2$ 在線段 $P_3P_4$ 的延伸線的不同方向，則代表線段 $P_1P_2$ 、 $P_3P_4$ 沒有相交，如圖B-3(b)、B-3(d)。如果點 $P_1$ 和 $P_2$ 有一點在線段 $P_3P_4$ 的延伸線上，則跳至步驟(3)。
- (3) 當執行至此步驟即代表我們無法利用CCW來準確判斷線段 $P_1P_2$ 、 $P_3P_4$ 是否相交，因為有一點在某一線段的延伸線上，可是我們卻無法判斷它是屬於圖B-3(f)、B-3(g)何種情形。因此，此階段我們先當作兩線段不相交，讓梯形

的其他邊來判斷是何種情形。

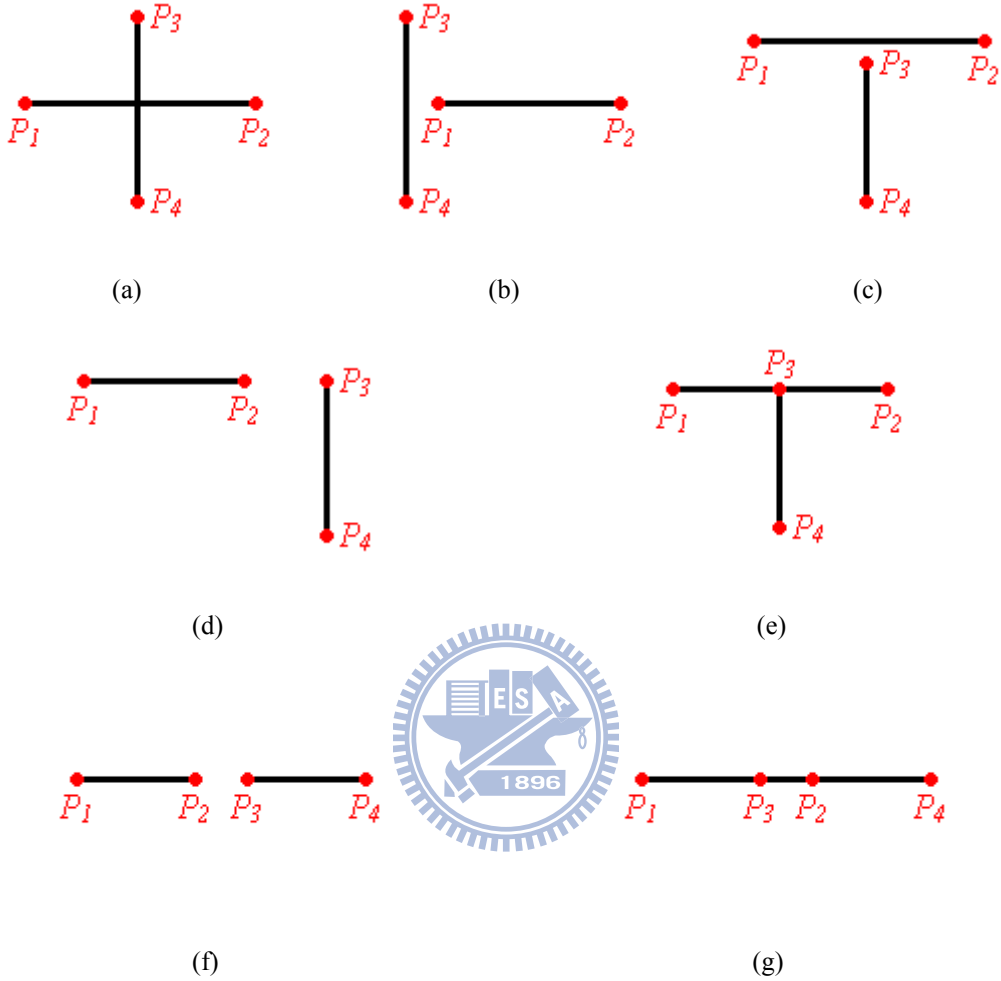


圖 B-3 兩線段是否相交的情形：(a) 相交，(b) 沒有相交，(c) 沒有相交，(d) 沒有相交，(e) 相交，(f) 沒有相交，(g) 相交。