

國立交通大學

資訊學院資訊科技(IT)產業研發碩士班

碩士論文

一個基於多節梯形的無軀幹人像活動表示方法

A Torso-less Representation of Human Activity with Multi-joint
Trapezoids

研究生：朱寧玄

指導教授：莊仁輝 教授

中華民國九十七年六月

一個基於多節梯形的無軀幹人像活動表示方法

A Torso-less Representation of Human Activity with Multi-joint
Trapezoids

研究生：朱寧玄

Student : Ning-Hsuan Chu

指導教授：莊仁輝

Advisor : Jen-Hui Chuang

國立交通大學

資訊學院資訊科技(IT)產業研發碩士班

碩士論文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Industrial Technology R&D Master Program on Computer Science and Engineering

June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

一個基於多節梯形的無軀幹人像活動表示方法

學生：朱寧玄

指導教授：莊仁輝 教授

國立交通大學資訊學院產業研發碩士班

摘 要

由於人類的平均年齡逐漸上升，不管在安養中心或是醫院，都需要大量的看護與醫療資源。因此我們提出一個基於電腦視覺的監控系統，幫助醫療人員或家屬能夠透過遠端即時的觀察被觀測者的活動情況。醫療監控系統不同於一般的安全監控系統，由於被觀測者通常不希望自己的真實影像被記錄和檢視，因此須注重保護被觀測者的隱私。由於被觀測者發生的情況經常需要緊急的處理，醫療監控系統同時也需要提供醫療人員即時性的資料。因此建立一個注重隱私，且具有較少資料量以及運算量的人體模型，是本篇論文的研究主題。在本篇論文中，我們提出一個快速且簡單的演算法，藉由監視攝影機，獲得人類影像中的肢體部位。最終提供一個基於多節梯形的無軀幹模型，並利用此無軀幹的肢體模型達成人像活動分析的目的。

A Torso-less Representation of Human Activity with Multi-joint Trapezoids

Student : Ling-Shuang Chu

Advisor : Jen-Hui Chuang

Industrial Technology R&D Master Program of Computer Science College
National Chiao Tung University

ABSTRACT

Because of the increase of average human age over the years, plenty of medical and nursing resources are needed, either in a nursing center or a hospital. To that end, we propose a surveillance system based on computer vision to help medical staffs or family members to monitor patient activities via remote access. Different from the surveillance for security purposes, it is usually desirable that original videos of patients' daily life are not available to the public. Therefore, the main feature of the proposed system is that the personal privacy is always ensured by using only a synthetic yet realistic human pattern to represent a patient. Since there are urgent situations in which instant care of patients is needed, a simple torso-less human pattern of multi-joint trapezoids is proposed in this paper, which can be generated near real-time through a computationally efficient analysis of a surveillance video. Experimental results show that one can indeed comprehend human activities reasonably well by examining the proposed torso-less human pattern.

誌 謝

轉眼間兩年的研究生涯就結束了，在這邊真的過得很快樂，每天都學習到很多不一樣的東西。而現在也能有一點小小的研究成果，最感謝的要算是老師和國華學長，雖然老師總是非常忙碌，但一路走來給了我很多寶貴的建議，也很辛苦的批改我寫的詞不達意的論文。沒有學長的話，今天這個系統也不會誕生出來了，如果說老師這篇論文的導演，學長就是這個系統的技術顧問了！很感謝學長總是不厭其煩的解答程式和演算法上的問題。再來要感謝的是我的家人，有了你們我才沒有後顧之憂，尤其我腸胃這麼差，以前總是要煩惱要吃什麼才不會拉肚子，住家裡不管是吃的還是睡覺都非常令人安心，不管我情緒不好還是狀況不佳，回家只要看到你們和咪咪（我們家的貓），煩惱都會迎刃而解。最後要感謝我的同學們和學弟和兩年多來一起遊山玩水的朋友們和球友，你們讓我的生活有更多感動和回憶，這樣年紀老了以後才有東西可以拿來說嘴啊！！總之要感謝的人太多了，我想就感謝天吧！

目 錄

目 錄.....	VI
圖目錄.....	VII
一、簡介	1
1.1 研究動機.....	1
1.2 相關研究.....	2
1.3 系統流程.....	3
二、前景偵測	4
2.1 傳統前景物的追蹤與偵測.....	5
2.2 高斯混合模型的定義.....	6
2.3 背景建立與前景偵測.....	7
三、掃瞄線演算法	8
3.1 掃瞄前景物.....	9
3.2 掃瞄線線段之近似中軸線.....	10
3.3 總結.....	13
四、人體的簡化模型	14
4.1 近似中軸線群組的對應梯形區塊.....	15
4.2 整合各群梯形區塊.....	15
4.2.1 相同截面的梯形區塊整合.....	16
4.2.2 不同方向截面的梯形區塊整合.....	18
4.3 搜尋肢體末端部位.....	23
4.4 肢體追蹤.....	25
4.5 總結.....	26
五、實驗結果與討論	27
5.1 辨識末端肢體的實驗結果.....	27
5.2 辨識末端肢體的實驗結果.....	28
5.3 無軀幹人像模型的實驗結果.....	30
5.4 計算旋轉矩型面積以及高度的實驗結果.....	34
5.5 使用者辨識模型的實驗結果.....	36
六、結論與未來工作	37
參考文獻	38

圖目錄

圖 1-1 系統流程圖	3
圖 2-1 建立高斯混合模型的簡略流程圖	4
圖 2-2 根據混合高斯模型所找出的前景	7
圖 3-1 掃瞄線演算法流程圖	8
圖 3-2 利用四種斜率的掃瞄線掃瞄前景區的結果	9
圖 3-3 淘汰圖 3-2 較長的橫截面後的結果	10
圖 3-4 角度與長度關係圖	11
圖 3-5 人類頭部與頸部部位放大圖	11
圖 3-6 近似中軸線段群組	12
圖 3-7 相同肢體部位利用角度不同掃瞄線產生的群組	13
圖 3-8 水平掃瞄線所獲得的近似中軸線	13
圖 4-1 人類梯形模型系統流程圖	14
圖 4-2 對應近似中軸線群組的梯形區塊	15
圖 4-3 梯形區塊不符合 (4.1) 式的例子	16
圖 4-4 經由 45° 掃瞄線產生的梯形區塊	17
圖 4-5 系統整合梯形區塊的流程圖	19
圖 4-6 考慮長度與角度進行整合	21
圖 4-7 考慮長度與角度進行整合	22
圖 4-8 整合時允許部分梯形重疊	23
圖 4-9 被誤判為肢體部分的身體軀幹	23
圖 4-10 搜尋肢體末端方法示意圖	24
圖 4-11 利用搜尋肢體末端方法改善肢體模型	24
圖 4-12 連續影像中的肢體追蹤	25
圖 5-1 利用辨認肢體的方法去除非肢體的區域	28
圖 5-2 利用辨認肢體的方法去除非肢體的區域	29
圖 5-3 人像雙手平舉的實驗結果	30
圖 5-4 人像雙手高舉的實驗結果	31
圖 5-5 人像向右行走的實驗結果	31
圖 5-6 人像半蹲的實驗結果	32
圖 5-7 人像在地面上滑倒的實驗結果	32
圖 5-8 人像坐下的實驗結果	33
圖 5-9 掃瞄旋轉的矩型區域	34
圖 5-10 掃瞄旋轉的矩型面積正確率	35
圖 5-11 掃瞄旋轉的矩型長度正確率	35
圖 5-12 使用者辨識模型的正確率	36

一、簡介

1.1 研究動機

近來電腦視覺的監控系統逐漸受到廣泛的注意，主要原因為電腦能過濾大量的資訊，並能幫助人類進行初步的影像內容分析。監控系統的應用主要可以分為兩大類，一類為保全監控系統，另外一類為醫療監控系統。不管是哪一類系統，電腦都可以幫助人類過濾和辨識有用的資訊。在保全系統中，電腦可以幫助警示環境中的入侵者，進而記錄其侵入的路徑，分析其肢體動作等。在醫療系統之中，電腦能夠幫助醫療人員分析及記錄病患的日常行為。

本研究著重於基於電腦視覺的監控系統在醫療監控系統之中的應用。醫療監控系統不同於一般的監控系統，由於被觀測者往往不希望自己的真實影像被記錄，因此醫療監控系統必須更注重被觀測者的隱私。另外由於被觀測者發生的情況往往需要緊急的處理，醫療監控系統需要提供醫療人員更即時性的資料與行為分析結果，來掌握被觀測者的日常行為。因此建立一個沒有被觀測者的原始影像，並且具有較少資料量以及運算量的人體模型，進而利用此人體模型進行行為分析是本篇論文的研究主題。

在基於電腦視覺的行為分析之中，如果建立越完整的人體模型，行為分析的結果也會越準確，但越完整的人體模型不管是資料量或計算量也會越大。為了在模型精細程度與計算量之間求取平衡，我們利用建立人體的肢體部位的模型，來達成我們的需求。我們提出一個快速而有效率的演算法，搜尋人類影像中的肢體部位，並儲存肢體部位形狀與角度的資訊，最後提供一個基於多節梯形的無軀幹模型來代表人體，並進而進行人像活動的分析。

接下來的小節中我們將介紹相關人類行為分析的研究，以及其他利用行為分析的成果來發展的應用系統。

1.2 相關研究

在人類的行為分析的相關研究中，一種作法是利用完整的人體模型來進行分析，如[1][2][3][4]，但將會使資料量以及電腦的運算量大幅提高，因此有些人利用人體的局部資訊，例如[5][6]使用人類的影像輪廓完成行走姿態的分析，[7]則是使用projection histogram來分析人類姿態。此外，也有許多研究是利用近似人類的骨架或是近似中軸線來對人類進行姿態辨識。利用近似骨架來進行姿態辨識的原因是因為在人體的結構中有許多長條形的部分，其骨架的角度與形狀是一個重要的特徵，可利用其來辨識人類的行為。

根據上述的理論，[8] 利用尋找物體之中的近似中軸線，將影像自含有人體影像的圖片中切割出來。 [9] 是尋找近似人體骨架的另一個方法，作者首先把人體影像三角化，找出三角形的重心，再利用深先搜尋（depth-first）連接各重心，取得該人體姿態的近似骨架，再與資料庫中的人體姿態模型進行比對辨識。從實際的實驗結果可以觀察到，近似骨架與人體姿態有相當程度的符合。另外[1]提出一個取出完整近似骨架的方法，作者將人類影像分為兩群：長形部位，例如四肢，以及關節部位，並利用worm和circle模型來組合成一個近似骨架。Worm模型用來描述長條形物體，circle模型則用來描述關節部位，作者利用能量函數來搜尋人像輪廓內部適合的模型，來獲得近似的骨架。

[10]的方法是將近似中軸線的概念應用在辨識染色體(DNA)，由於染色體具有長條且寬度恆定的特性，尋找該染色體的近似中軸線能代表該染色體的一項重要特徵。作者利用四個方向的掃瞄線來掃瞄染色體，並藉此獲得染色體的橫截面，再挑選適當長度的橫截面，整合後將可獲得近似於染色體本身形狀的近似中軸線。

我們的方法是將[10]的演算法加以改進，但由於人類影像比單一染色體多了關節部位，部分橫截面的變化也更加劇烈，因此我們經由四個掃瞄方向所獲得的近似中軸線，需要更複雜的整合，最後才能獲得有意義的人體模型。

1.3 系統流程

本論文中的系統主要分為兩大步驟，流程圖如圖1-1所示。第一步驟是以掃瞄線演算法為基礎，我們首先利用一台攝影機拍攝視訊畫面，接著使用高斯混合模型GMM(Gaussian mixture models, GMM)，找出拍攝影像中的前景物，再透過掃瞄線演算法找到肢體部位的近似中軸線群組。第二步驟，是利用第一步驟所得到的近似中軸線群組，產生多節梯形，並分析其角度與位置，再對於所有的多節梯形作進一步的整合，最後得到一個基於多節梯形的無軀幹人體模型。

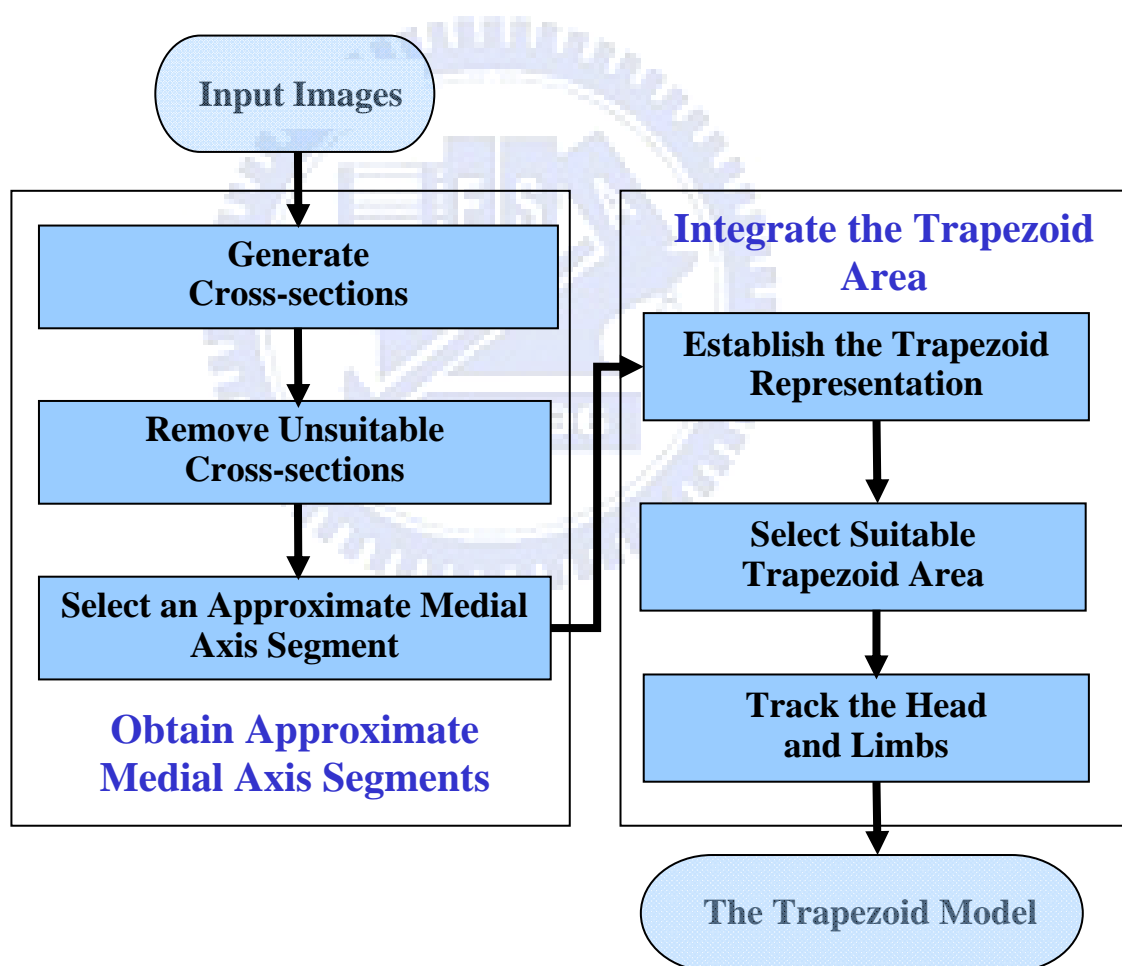


圖 1-1 系統流程圖

二、前景偵測

本系統是從攝影機拍攝的影像中，找出人體的影像區域，分析其肢體形狀以及建立人體模型，目前僅針對單一的人類影像來進行處理。系統首先需從拍攝的影像之中將人體區域自背景影像中分離出來，我們採用高斯混合模型(Gaussian mixture models, GMM)來建立背景影像，此方法可自動建立背景模型，並且能適應背景的光影變化。建立背景影像後，我們利用背景影像相減的原理即可找出影像中的人體區域，上述自動建立背景模型方法的主要步驟如下：

- (1) 對影像的每一個像素，觀察其色彩強度發生的機率值，並建立數個高斯分佈模型，且給予不同的權重值。
- (2) 每一張新進的影像的每一個像素分別和背景模型比對，符合的為背景、否則為前景物體。
- (3) 根據比對影像的數值，動態調整步驟(1)中所得高斯模型的權重。

圖 2-1 為建立高斯混合模型的簡略流程圖，我們將在接下來的各節中逐一介紹傳統偵測前景物的作法以及介紹我們使用的方法與原理。

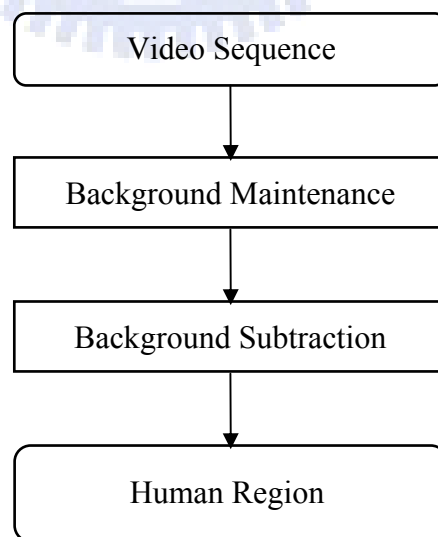


圖 2-1 建立高斯混合模型的簡略流程圖

2.1 傳統前景物的追蹤與偵測

目標物的偵測與追蹤所使用的演算法，往往決定視訊監控系統效果的好壞。在傳統上偵測移動目標物，有下列兩種常用的方式：

- (1) 連續影像相減法 (temporal difference)：連續影像相減法無需須事先建立背景模型，只需利用連續的畫面作像素亮度的相減，即可判斷目標物的位置。連續影像相減法的優點在對亮度等環境上的變化有較佳的適應性，但缺點是偵測移動物體易發生影像破碎的情形，且無法找出靜止一段時間或移動距離非常小的物體。另一方面，若攝影機不固定，由於前景背景都在移動，會使得前景與背景皆有差異而難以分離。另一個與連續影像相減法類似的作法為光流偵測法 (optical flow)，透過複雜的運算求出連續影像光流量的變化來偵測物體。此方法當攝影機跟隨物體移動時，仍有不錯的效果，但由於龐大的計算量因此不適合應用於即時系統。
- (2) 背景相減法 (background subtraction)：背景相減法的原理是事先建立背景模型，再將新影像與背景影像的像素強度值做相減，以偵測出移動的物體區域。背景相減法較不會有前景物破碎的問題，但其缺點在於對光線的變化較為敏感，當背景與目前影像的亮度值差異過大時，會造成無法正確的偵測出前景物。但基於運算量低的理由，背景相減法是目前目標物偵測中最常使用的方法。

在背景相減法的基礎底下，有許多類似的研究，在產生以及更新背景影像上做出改進，其方法約略又分為三個種類：1. 靜態攝影機之非迭代式方法 (Non-recursive approaches with stationary camera)、2. 靜態攝影機之迭代式方法 (Recursive approaches with stationary camera)、3. 非靜態攝影機的全景背景模型 (Panorama background image with non-stationary camera)。在這裡僅針對我們系統採用的靜態攝影機之迭代式的方法進行介紹。

靜態攝影機之迭代式的方法不需儲存一段影像的像素值，而是以迭代的方式在每個頁面更新背景模型。傳統的作法在出現背景誤判時，將導致錯誤持續一段

時間。目前普遍改進的作法是採用指數型的權重來去除過去的資訊，並選用正確的背景像素進行參數的更新。我們所採用的高斯混合模型即是屬於此類方法。高斯混合模型可以建立並同時追蹤數個高斯分佈，且由於高斯混合模型是參數式模型。不需額外的空間儲存一段影像資訊，可以動態的更新背景模型，下一節中我們將簡單介紹高斯混合模型的定義，更詳細的說明請參考 [16]。

2.2 高斯混合模型的定義

混合高斯模型是由不同權重的高斯函數所組成的機率密度函數。一個混合 M 個高斯模型的高斯混合模型可由下列公式代表：

$$A = \{w_k, \mu_k, c_k\}, k = 1, \dots, M \quad (2.1)$$

其中 μ_k 為高斯混合機率密度平均向量(mean vector)、 c_k 為共變異矩陣(covariance matrix)及 w_k 為其相對應之混合權重。而單一高斯函數維度，當維度為 D 時其形式如下列所示：

$$c_k(x) = \frac{1}{(2\pi)^{D/2} |C_k|^{1/2}} \exp\left\{-\frac{1}{2}(x - \mu_k)^T C_k^{-1} (x - \mu_k)\right\} \quad (2.2)$$

其中 x 為 D 維之向量，也是我們欲求得像素之發生機率值，我們再將單一的高斯函數維度推廣到混合機率密度函數，其形式為：

$$p(x|\Lambda) = \sum_{k=1}^M w_k c_k(x) \quad (2.3)$$

在這裡 $b_k(x)$, $k = 1, \dots, M$ 為混合 M 個的機率密度函數並有著不同權重，

w_k , $k = 1, \dots, M$ 為其相對應之權重，並有 $\sum_{k=1}^M w_k = 1$ 的限制。我們將在下節中說明我們如何利用上述的高斯混合模型之定義來建立背景的背景模型，以及偵測前景的目標物。

2.3 背景建立與前景偵測

圖 2-2 為我們根據混合高斯模型的定義，找到背景影像後與原始影像做像素相減以及二值化後的前景影像。我們使用下列步驟來建立背景模型：

- (1) 首先利用輸入的影片取初始的幾個影格建立每個像素亮度及顏色發生機率的數個高斯函數模型。
- (2) 利用每個新進的影格，更新每個像素的亮度以及顏色的權重值。
- (3) 每個新進的影格，當像素符合既有的高斯函數，將此像素加入這個高斯函數，並更新 μ_k 、 c_k 。如果沒有符合的高斯函數，原先像素發生機率最低的一個高斯函數將被此新加入的像素產生的高斯模型所取代，並給予一個最低的權重。
- (4) 所有的高斯函數模型皆利用其權重來排序其發生的機率，機率最高的高斯模型將被選為背景。



圖 2-2 根據混合高斯模型所找出的前景(a) 輸入影像，(b) 二值化的前景影像。

三、掃瞄線演算法

本論文是經由分析影像中的人體區域來產生模型。因此如果能夠取得越多影像中的資訊，人體模型也會越符合影像中的人體區域。但為了能在模型的精細程度與計算量之間取得平衡，我們提出利用四種方向以及一定間隔的掃瞄線方法來掃瞄前景物，以獲得建立模型所需的資訊。我們首先將處理過的前景物利用 0° , 45° , 90° , 135° 四種斜率的掃瞄線，過濾出橫截面，並將長度、角度過大的橫截面除去，再透過橫截面之近似中軸線，將長度與角度相近之橫截面分群。詳細演算法的流程圖請參考圖 3-1，我們將在接下來的各小節逐一介紹每個步驟的詳細作法與原理。下一章中我們將會利用此橫截面群組整合成代表人體四肢、頭部的梯形模型。

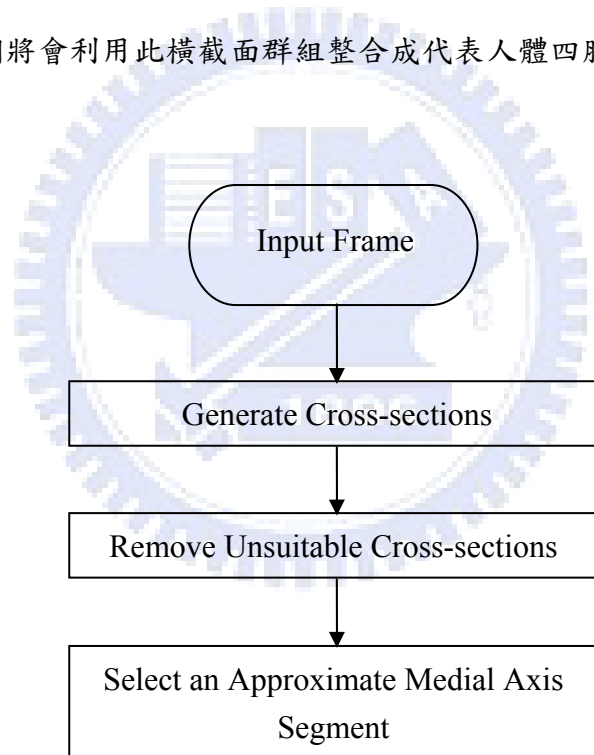


圖 3-1 掃瞄線演算法流程圖

3.1 掃瞄前景物

在利用影像中的人體區域來建立模型的時候，我們首先利用角度分別為 0° , 45° , 90° , 135° 掃瞄線，掃瞄前景物區域，藉此獲得前景物的橫截面。掃瞄線間隔以參數 δ 來表示，參數 δ 越小，產生肢體模型時所能參考的資訊量越大，產生的肢體模型會更接近影像中的人體區域，但計算量相對也會提高，關於如何決定參數 δ ，我們將在第五章討論。圖 3-2 分別是利用 $\delta=6$ （單位為像素）以及四種斜率掃瞄線掃瞄前景區，所得的結果即為四種方向的前景物橫截面。

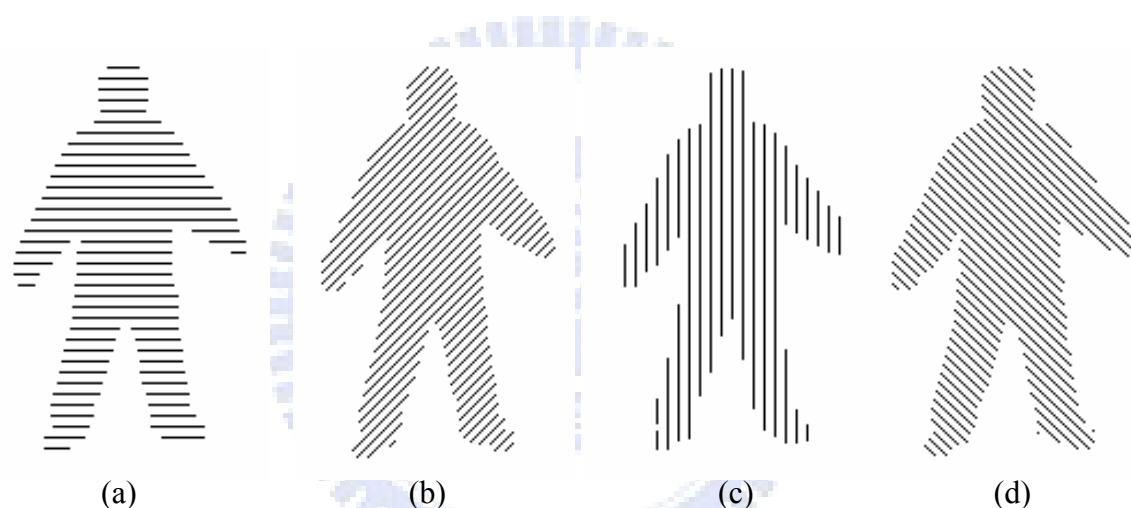


圖 3-2 利用四種斜率的掃瞄線掃瞄前景區的結果：(a) 0° ，(b) 45° ，(c) 90° ，(d) 135° 。

本論文研究的目標為建立人體肢體的模型，人類肢體的寬度一般佔身體寬度的一半以下，因此我們希望將超過身體寬度一半以上的橫截面去除，初步過濾非肢體的區域。考慮雜訊以及前景物移動等因素，我們以各方向掃瞄線所能掃瞄的最大長度 L 的三分之一為臨界值，去除長度超過之橫截面。圖 3-3 顯示淘汰圖 3-2 較長的橫截面後的結果。在下一小節中，我們將選擇長度與角度近似之橫截面並產生其所對應的近似中軸線，進而將橫截面分類以搜尋肢體的區域。

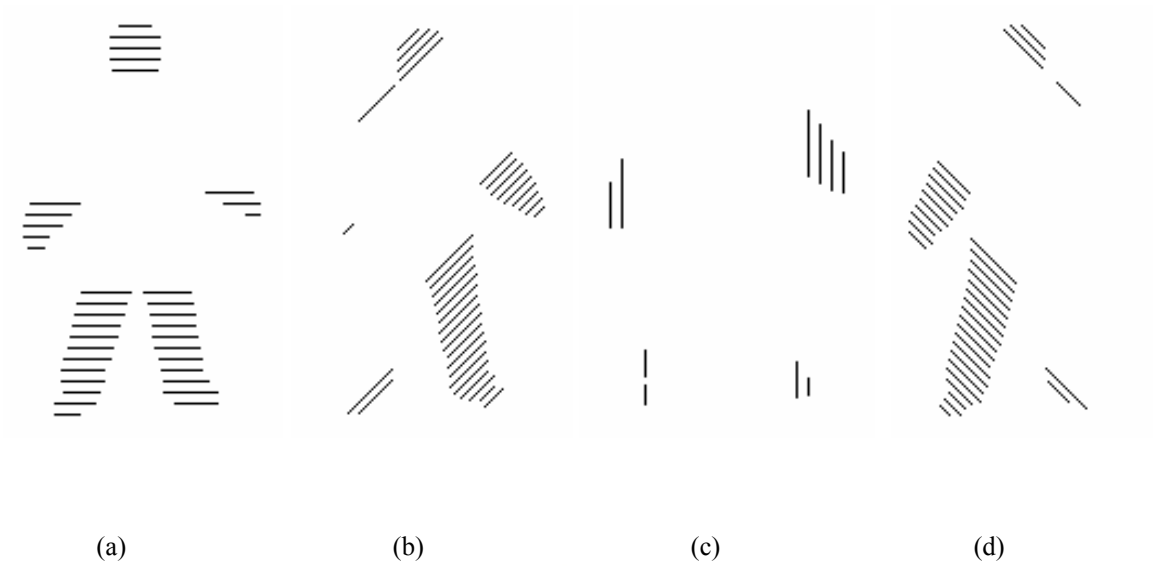


圖 3-3 淘汰圖 3-2 較長的橫截面後的結果：(a) 0° ，(b) 45° ，(c) 90° ，(d) 135° 。

3.2 掃瞄線線段之近似中軸線

在人類的肢體中，骨架往往和肢體的橫截面垂直，因此搜尋與橫截面垂直的近似中軸線，可以代表局部肢體的特徵。根據上述的理由，我們的方法為搜尋肢體的近似中軸線，且近似中軸線段與橫截面越接近九十度越理想。我們首先在同一角度掃瞄線產生的橫截面中，取兩個連續橫截面的中點，如果兩個中點的距離為少於臨界值 D ，且橫截面的長度變化量少於 T ，則將此二者的橫截面中點相連成為近似的中軸線段。長度 D 如何訂定與我們允許兩中點連線與中垂線的夾角大小有關，當 D 越趨近兩橫截面的間隔距離 δ ，近似中軸線與截面中垂線夾角將會越小。如圖 3-4 所示，我們首先定義兩中點連線距離為 D_1 ，紅色端點為我們假設的橫截面中點，實線 Q 為兩橫截面中點的連線，虛線 R 為兩橫截面的中垂線， θ 為 Q 與 R 的夾角。當 θ 越趨近於零度時，代表近似中軸線段 Q 長度會越接近中垂線 R 。故我們可以得知當兩中點連線越偏離中垂線，則 D_1 也會越長。因此當我們控制 D_1 ，即可控制近似中軸線段越接近中垂線。

但在真實影像中，由於有雜訊的干擾與攝影機角度等問題， D 不宜訂定太小，我們將在第五章討論改變 D 之下實驗結果的差異。

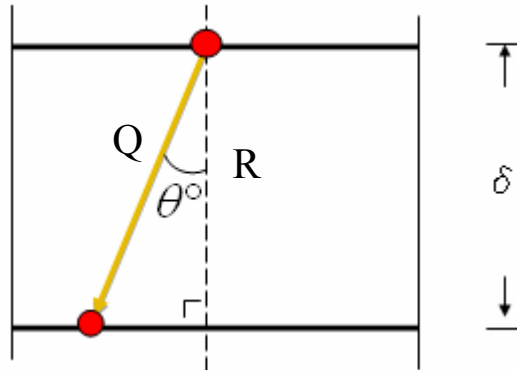


圖 3-4 角度與長度關係圖

至於我們希望控制橫截面的長度變化量的原因是由於在肢體與身體，或是頭部與頸部的交界處，橫截面的長度變化量通常非常劇烈，我們可以觀察3-5圖，如果我們這時候將頸部部位的兩橫截面的中點相連，則會造成將頭部與頸部無法區分，導致之後的肢體模型出現錯誤。因此當橫截面的長度變化量大於 T ，即使近似中軸線與橫截面垂直，如圖中紅色的部分，這兩個橫截面的中點仍不會連成近似中軸線段。

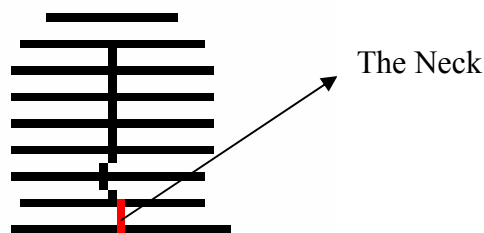


圖 3-5 人類頭部與頸部部位放大圖

人類的身體結構中，組成身體之各部位皆有其骨骼結構，例如大小腿部位，且骨骼與該部位的橫截面大致呈現垂直。利用這個特徵，我們可以利用相似橫截面的近似中軸線來區分不同肢體，即每一條近似中軸線代表每一個肢體部位。在

求得各橫截面的近似中軸線段後，我們可將近似中軸線段相連的橫截面分為同一群。如圖3-6 (b)，為我們將3-6 (a)的左腳與右腳透過相同的近似中軸線段，連結成相同的群組之放大圖。

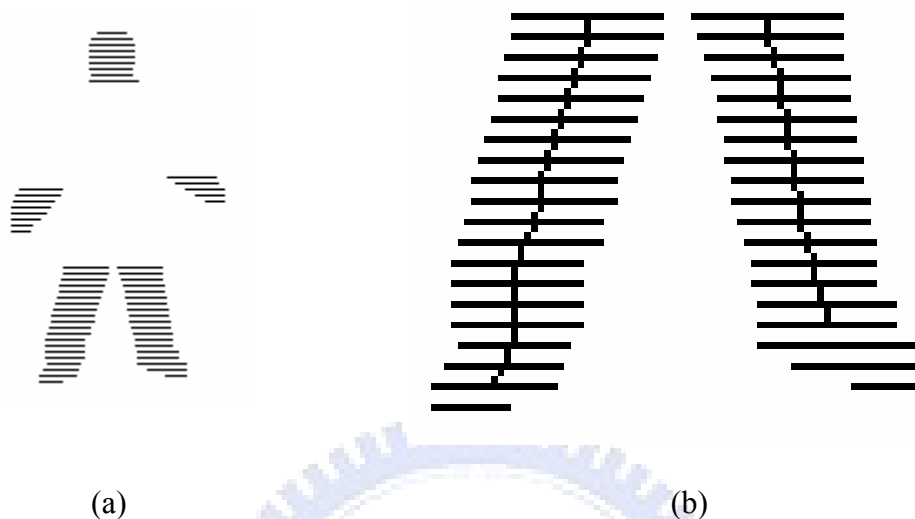


圖 3-6 近似中軸線段群組 (a) 利用 0° 掃瞄線掃瞄後的橫截面，(b) 腳部近似中軸線群組放大圖。

如本小節一開始所述，我們希望近似中軸線能夠儘量垂直於橫截面。但即使控制近似中軸線段偏移角度在一定的範圍之內，在產生近似中軸線段群組後，此偏移角度仍可能逐漸累積。圖 3-7 為相同肢體部位利用不同角度掃瞄肢體部位所得的結果，如圖 3-7(a)所示，群組中的近似中軸線較圖 3-7(b)偏移量較大($\theta_1 > \theta_2$)，當群組中的近似中軸線較偏離橫截面的中垂線時，代表利用這個方向掃瞄所得的橫截面群組不理想，故我們認為圖 3-7(a)掃瞄出來的橫截面群組較圖 3-7(b)的橫截面群組較不能代表該肢體部位。我們會記錄每一個橫截面群組偏移中垂線的資訊，以供梯形模型整合時使用，我們會在第四章詳細介紹。

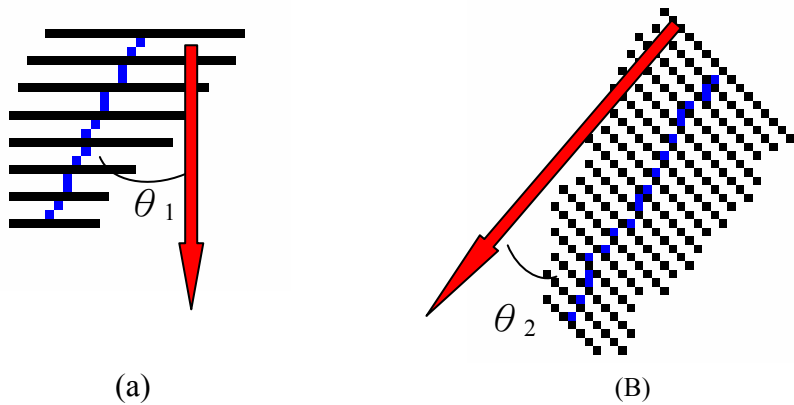


圖 3-7 相同肢體部位利用角度不同掃瞄線產生的群組 (a) 135° 度的掃瞄線，(b) 0° 的掃瞄線。

3.3 總結

在本章中，我們利用掃瞄線演算法，將人體四肢利用近似中軸線代表其特徵，並加以區分。我們首先利用各種不同斜率以及一定間隔的掃瞄線，掃瞄前景影像中的人體區域，過濾出位於人體內的橫截面，並將長度超過肢體寬度的橫截面去除，再利用連接橫截面找出代表四肢之近似中軸線，接著將相同角度以及長度的橫截面分群。如圖3-8所示，不同的顏色線段代表的是連結不同橫截面群組的近似中軸線。

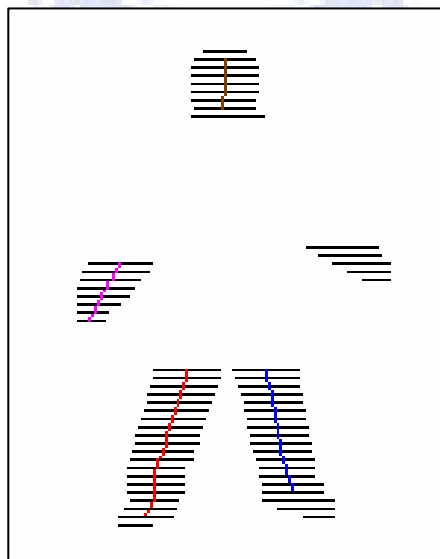


圖 3-8 水平掃瞄線所獲得的近似中軸線。

四、人體的簡化模型

在本章節中我們將對於第三章的結果，進一步的分析與整合，最終提供一個簡單的多節梯形模型來代替人體影像。由於本系統建構在一個即時系統上，所以模型的資料量必須儘量減少，以利於傳輸。我們使用的梯形模型來表示人類肢體，只需記錄梯形的頂點座標，即有上述資料量少的優點。當取得橫截面群組後，我們只用一個梯形來代表包含該群橫截面的群組，接著再整合四個方向的掃描所得到的梯形區塊，最後建立包含能夠代表人體肢體的梯形模型，系統流程如圖4-1所示。

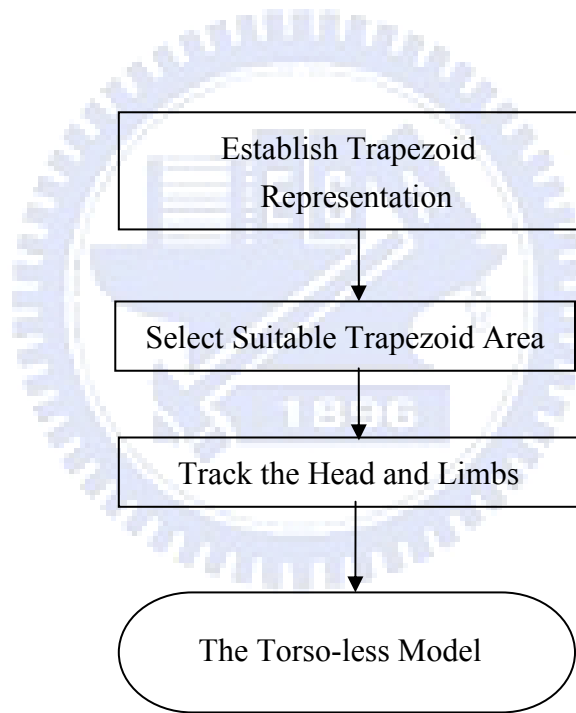


圖 4-1 人類梯形模型系統流程圖

4.1 近似中軸線群組的對應梯形區塊

為了能夠讓人體模型的資料量儘量減少，我們決定使用梯形模型來表示人類肢體之各部位。首先我們將相同橫截面群組的第一個橫截面與最後一個橫截面的四個端點，定義為該肢體部位所代表梯形的端點，形成包含該群橫截面線段的梯形區塊。利用此一方法，我們可以將圖3-8的近似中軸線群組對應至不同的梯形區塊，如圖4-2所示。

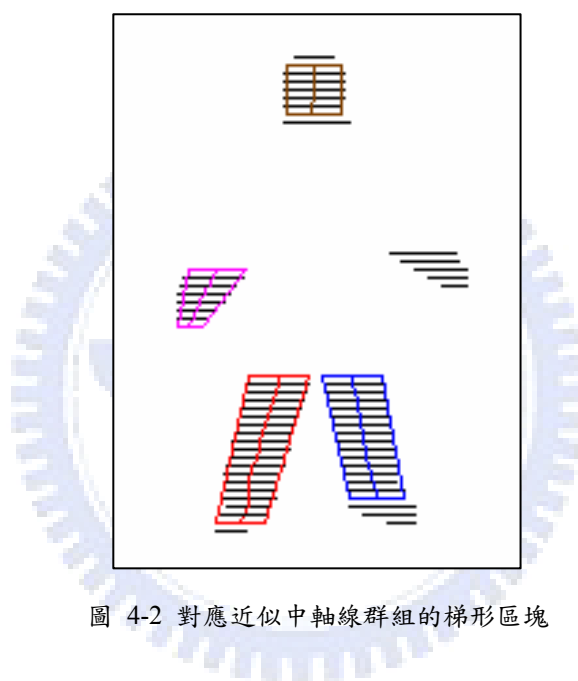


圖 4-2 對應近似中軸線群組的梯形區塊

4.2 整合各群梯形區塊

由於我們最終提供一個多節梯形模型來代表人體，所以我們必須將四種不同的近似中軸線段群組整合成一個能代表肢體的多梯形模型。首先我們先將相同斜率掃描線所產生的梯形區塊，不適合的部分加以淘汰，再將不同斜率掃描線所產生的梯形區塊，取長度與角度適合的部分整合，最後產生我們的梯形模型。

4.2.1 相同截面的梯形區塊整合

為了能在整合時減少計算量以及減少錯誤率，我們將先過濾每個斜率掃瞄線所產生的梯形區塊。由於我們使用的是一般的光學攝影機，在前景偵測中，會有許多的細碎影像雜訊，雜訊仍然會產生梯形區塊，但往往面積不大，因此過小的梯形區塊我們將視為雜訊予以去除。首先我們定義 H 為梯形區塊包含多少橫截面，在本系統中我們設定 H 必須大於 3，滿足這個條件的梯形區塊才保留下來。

人體實際的肢體區域，寬度變化不大，因此我們只保留橫截面變化不大的梯形區塊。我們在第三章的掃瞄線演算法產生近似中軸線線群組中，已設定橫截面變化必須小於一個臨界值 T ，但在產生其對應的梯形區塊後，其橫截面變化仍可能逐漸累積。我們定義梯形的上底長度為 T_a ，下底長度 T_b ，我們利用 (4.1) 式計算 T_a 與 T_b 的變化量：

$$\frac{|T_a - T_b|}{\text{Min}(T_a, T_b)} < 4 \quad (4.1)$$

在 (4.1) 式中 T_a 與 T_b 的變化量，會受到影片的解析度以及人像在影片中的比例影響，在本系統中，當差距達到四倍以上，我們即視其由上而下橫截面所累積的變化量過大，而不保留這個梯形區塊。圖 4-3 中紅色框線中，即為近似中軸線線群組不符合 (4.1) 式的例子。

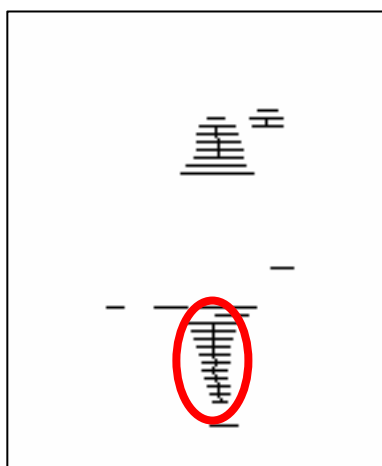


圖 4-3 梯形區塊不符合 (4.1) 式的例子

人體實際的肢體區域，其形狀多為細長形，因此在這個階段我們希望排除較扁的梯形區塊。我們首先定義上底長度為 T_a ，下底長度 T_b ， H 為上底與下底的中點連線距離，並使用 (4.2) 來計算梯形區塊的高寬比 ($Ratio$)：

$$Ratio = H / W \quad (4.2)$$

其中， W 表示 T_a 與 T_b 的平均長度。利用 (4.2) 算出來的結果，如果值越大，代表此梯形區塊越細長。本系統中，我們設定高寬比必須大於 2.2，只要高寬比超過 2.2，我們就判定梯形區塊不會太扁，才保留此梯形區塊。如圖 4-4 (a) 所示，在經過高寬比過濾後，我們去除紅色標示區域的較扁梯形，留下的梯形區塊在這個掃瞄階段，較能代表比較完整的肢體區塊。

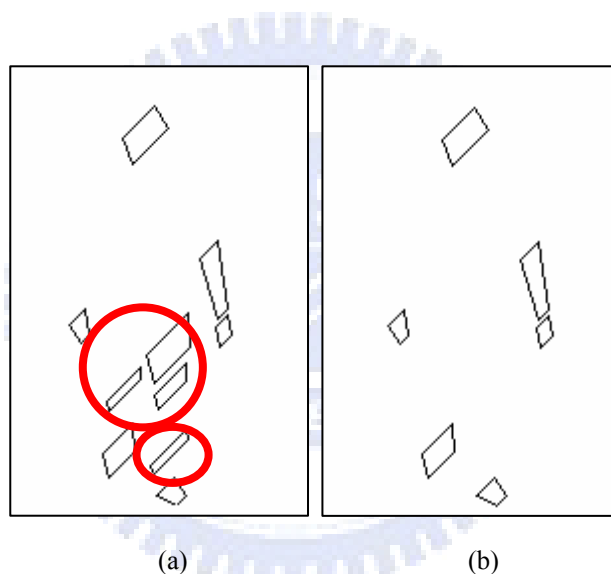


圖 4-4 經由 45° 掃瞄線產生的梯形區塊。(a) 未經高寬比 2.2 淘汰的梯形區塊，(b) 經高寬比 2.2 淘汰後的梯形區塊。

4.2.2 不同方向截面的梯形區塊整合

在這個階段我們挑選四個方向掃描線所產生的梯形區塊來進行整合。首先梯形區塊經由 4.2.1 節所利用的方法來去除不適合的部分後，剩下的梯形會產生許多重疊的情形，而有待整合。為了方便說明整合梯形的流程，在此我們先定義 P 、 Q 為兩個待整合的梯形區塊， H 為兩梯形上底與下底中點連線長度， Z 為兩梯形較短的 H 佔較長 H 的百分比， THs 是一個篩選重疊面積大小的臨界值，以下為選擇適合的梯形區塊的流程：

- (1) 輸入兩個欲比較的梯形區塊 P, Q 。
- (2) 判斷是否重疊，若不重疊則輸出 P, Q 。
- (3) 計算兩個梯形的 H ，以及 Z 是否達到 90% 以上，若有則跳到步驟 (6)。
- (4) 比較兩個梯形重疊的面積是否大於臨界值 THs ，若否則輸出 P, Q 。
- (5) 分別計算梯形區塊的 H 長度，並輸出 H 較長的梯形區塊，捨棄 H 較短的區塊。
- (6) 分別計算梯形區塊的近似中軸線偏離中垂線的平均長度，並輸出偏離長度較小的梯形區塊，捨棄偏離長度較大的梯形區塊。

圖 4-6 為系統篩選適合梯形區塊的流程圖，我們將在接下來的段落中詳細介紹各個步驟的詳細作法與原理。

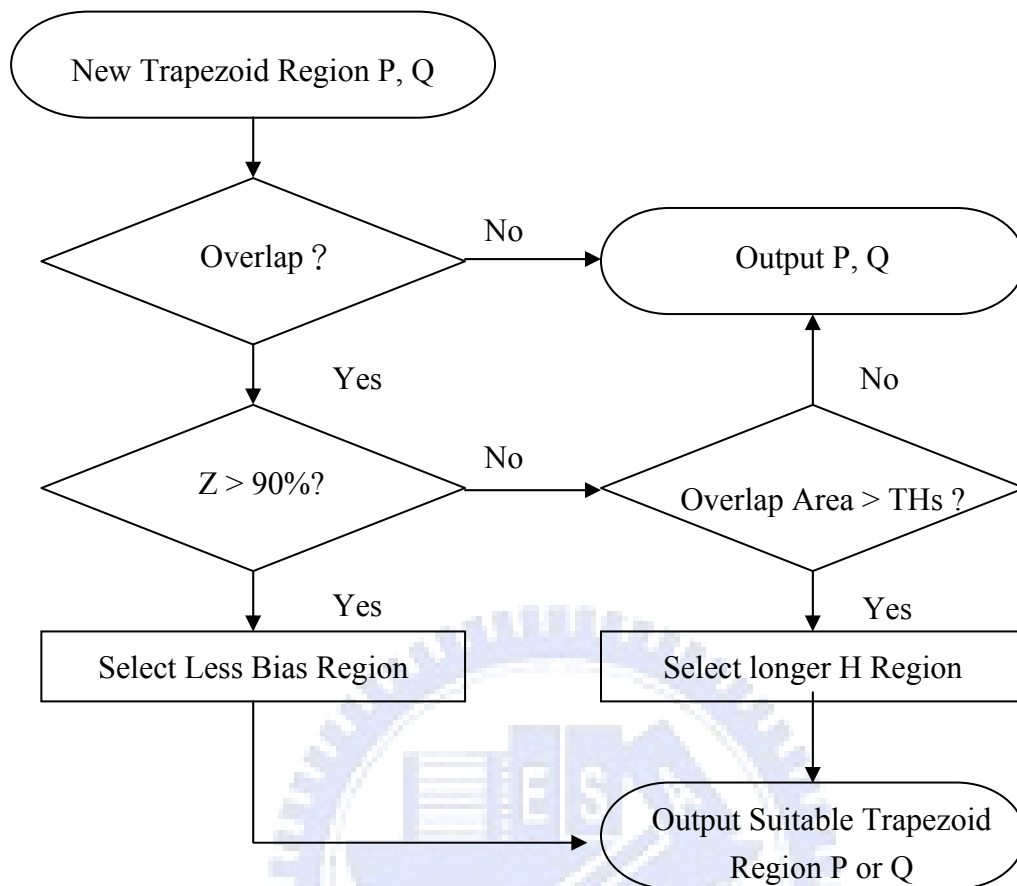


圖 4-5 系統整合梯形區塊的流程圖

在整合不同橫截面產生的梯形區塊時，我們首先必須先判斷梯形區塊是否重疊。為了計算速率與效果考量，我們採用 CCW (Counter Clock Turn) 演算法來判斷邊與邊是否有重疊。我們首先分別取梯形的四個邊以及一條 135° 的對角線，再利用運算四次 CCW 來判斷是否有任何一條邊或對角線重疊，即可判斷梯形區塊是否重疊。我們首先假設一個點 p_3 並且 p_3 不在 p_1 、 p_2 所構成的線段之上，CCW 演算法的流程如下：

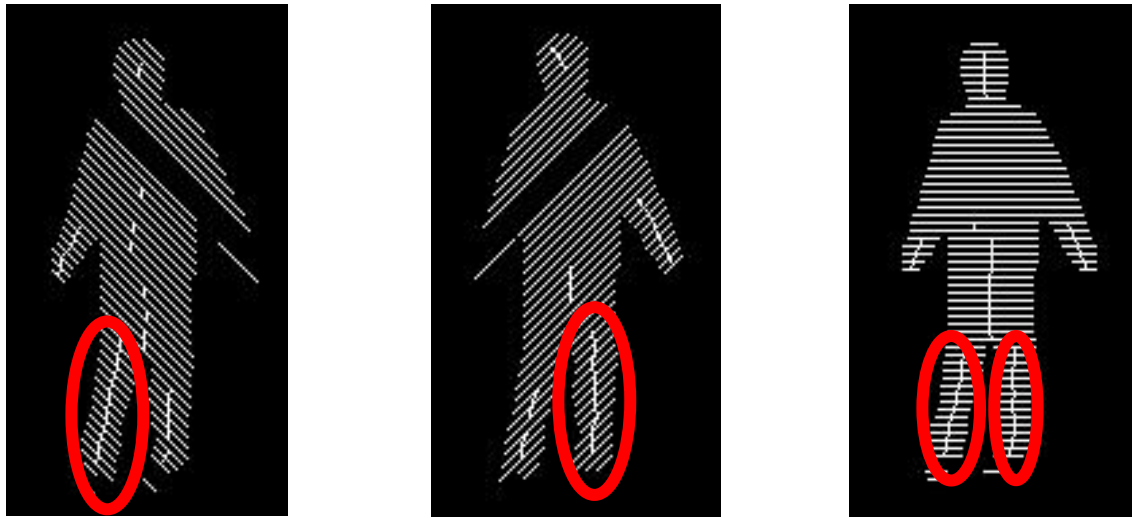
- (1) 定義向量 A 為 $p_2 - p_1$ ，向量 B 為 $p_3 - p_1$ 。
- (2) 計算向量 A 、 B 之內積。
- (3) 如果內積值為正，則代表點 p_3 在向量 A 的順時針方向，輸出 +1。
- (4) 如果內積值為負，則代表點 p_3 在向量 A 的逆時針方向，輸出 -1。

然而利用CCW演算法只能判斷一個點是否在一條線段的順時針或逆時針方向，如果要判斷線段有無交錯，總共必須做四次CCW，我們首先定義兩交錯線段A、B的兩端點 (p_1, p_2) (p_3, p_4) ，演算法的流程如下：

- (1) 判斷點 p_3 以及點 p_4 是否在線段A的兩側，我們先做一次CCW判斷 p_3 是否為線段A的順時針方向，再做一次CCW判斷 p_4 是否為線段A的順時針方向，如果 p_3 及 p_4 是在線段A的兩側，則將兩次CCW的結果相乘應為-1。
- (2) 判斷點 p_1 以及點 p_2 是否在線段B的兩側，首先做一次CCW判斷 p_1 是否為線段B的順時針方向，再做一次CCW判斷 p_2 是否為線段B的順時針方向，如果 p_1 及 p_2 是在線段B的兩側，則將兩次CCW的結果相乘應為-1。
- (3) 根據(1)(2)的結果，如果兩者的輸出皆為-1，則判斷線段A、B確為交錯，反之兩線段並無交錯。

使用CCW演算法判斷梯形區塊是否重疊並不能保證梯形沒有重疊，例如一個較大梯形包含一個較小的梯形，即有可能在沒有任何邊交錯的情況下，梯形仍重疊。但這樣的情形較少發生在我們的系統中，考量此作法在判斷梯形重疊上極有效率，我們仍決定採用CCW演算法來判斷梯形是否重疊。

在留下各橫截面產生較佳的梯形區塊後，對於沒有重疊的梯形區塊，我們即認為其能夠充分代表該肢體部分的區域，因此將予以完全保留。



(a) (b) (c)

圖 4-6 考慮長度與角度進行整合 (a) (b) 考慮長度，(c) 考慮長度與角度。

接著我們繼續處理重疊的梯形區塊，我們首先定義兩個重疊部分梯形區塊上底與下底中點的連線長度為 H 。由於在相近區域之中，較長的梯形代表該角度的掃描線所產生的橫截面與此區域的中軸線較為垂直，且橫截面變化也較不劇烈，因此在相同區域中，如果有梯形區塊重疊的情形，我們會優先保留 H 較長的梯形。

但在長度相近的時候，只考慮長度的條件不一定能挑選到最理想的梯形。在第三章提到近似中軸線段群組偏離其中垂線的情形，這種情形違反我們希望近似中軸線儘量垂直於橫截面的條件。因此在整合梯形時，如果 H 相近時，除了考慮長度的條件，我們將一併考慮梯形的角度。如圖4-6所示，如果只考慮 H 的條件，我們會選擇圖4-6 (a)的左腳以及圖4-6 (b)的右腳作為整合後梯形的左、右腳，但是這兩者的近似中軸線段群組都已經涵蓋到軀幹的部分，如果考量角度與 H 的條件，我們會選擇圖4-6 (c)的左、右腳，這兩者的近似中軸線段群組較為近似原影像中的腿部的肢體部分。在本論文中，當兩個重疊梯形的 H 差在90%以內，我們便會以偏離中垂線的角度來選擇適合的梯形進行整合。結果如圖4-7所示，在圖4-7 (a)中可以見到只考慮長度時，腿部的梯形區域已經涵蓋身體。圖4-7 (b)中可以見到依據長度與角度進行整合後，腿部的梯形區域較接近真實人體影像中的腿部區域。

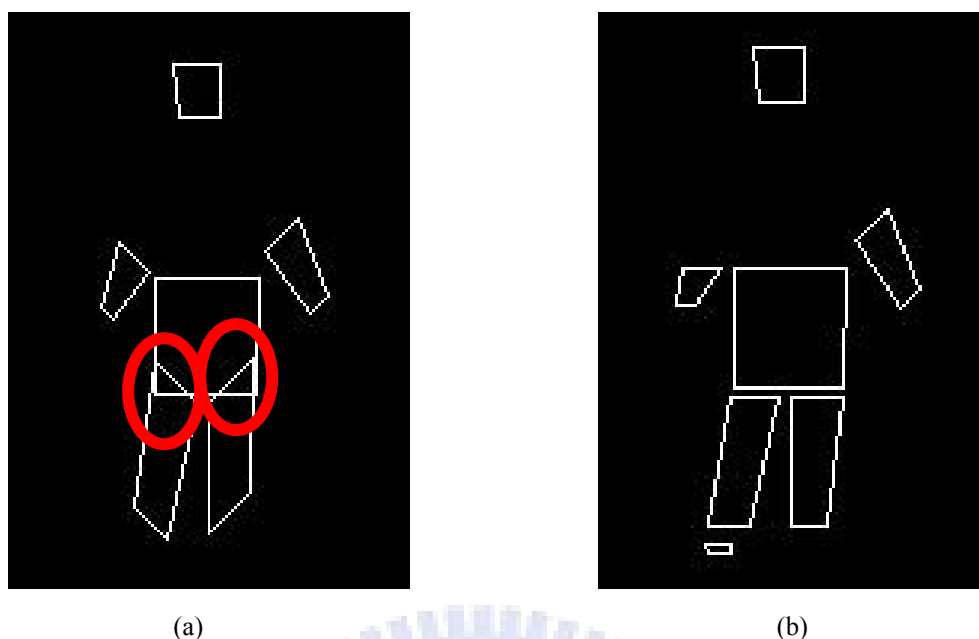


圖 4-7 考慮長度與角度進行整合 (a) 只考慮長度，(b) 兩者皆考慮。

在人類肢體部位，擁許多關節以及轉折處，在關節處的梯形區域極可能會互相重疊，此時不管考慮長度或是角度，若選擇其中一個梯形來進行整合，都有可能遺漏另外一個梯形的資訊。因此我們在整合梯形時，允許梯形能夠有部分重疊，但重疊的面積不宜太大，以免失去整合的意義。我們首先假設重疊的梯形區域面積分別為 A 、 B ，在本篇論文中設定重疊的區域的門檻值不能超過

$\frac{1}{3} \text{Min}(A, B)$ ，超過門檻值的兩個梯形區域還是會依照長度以及角度的資訊，選擇其中一個梯形來進行整合。如圖4-8 (a) 為梯形重疊，整合時未考慮允許部分區域重疊的情形，我們將會漏失許多肢體部位的資訊。圖4-8 (b) 為容許部分梯形區域重疊，我們將可獲得較完整的肢體資訊，紅色的圓圈標示出梯形重疊的部分區域，其中編號1、3、4的圓圈中有重疊的梯形區域，但部分較扁的梯形已經利用相同橫截面中的長寬比公式進行濾除，而編號2的紅圈即是利用此允許梯形重疊法留下的兩個梯形區域。

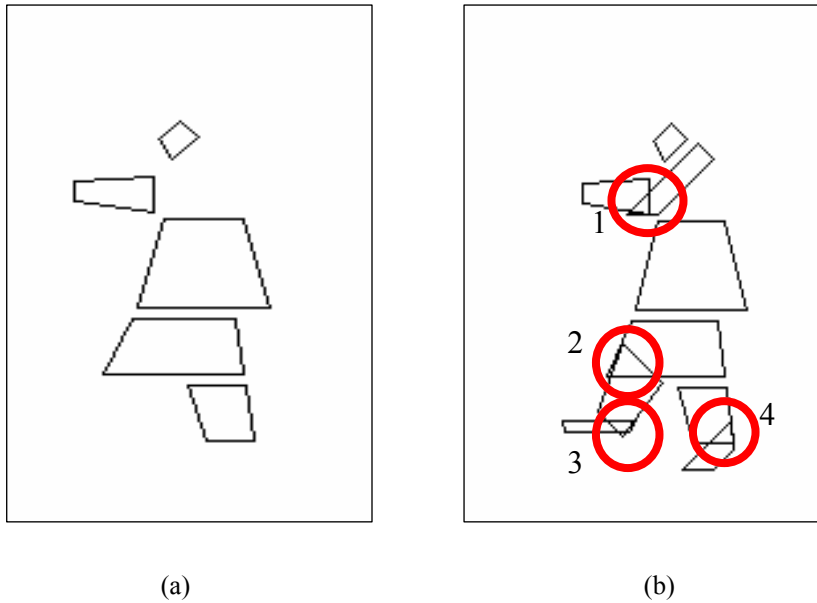


圖 4-8 整合時允許部分梯形重疊 (a) 梯形完全未重疊，(b) 允許部分梯形重疊。

4.3 搜尋肢體末端部位

我們的演算法著重於分析肢體部位，故希望能排除非肢體部位的資訊以減少資料與計算量。但在少數情況下，系統會把較細的身體部位誤判為肢體部位，例如對象以側面面對攝影機，或將身體向上延展等。如圖4-9之紅圈即代表被誤判的身體軀幹部分。由於身體軀幹部位並非肢體的末端部位，因此可以利用搜尋肢體的末端部位，進一步排除非肢體的部位（例如身體軀幹）。

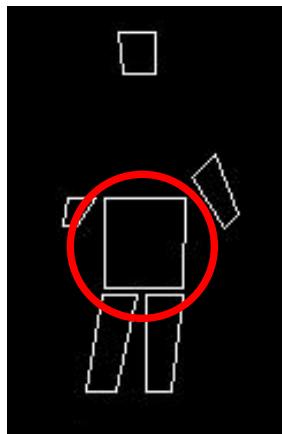


圖 4-9 被誤判為肢體部分的身體軀幹

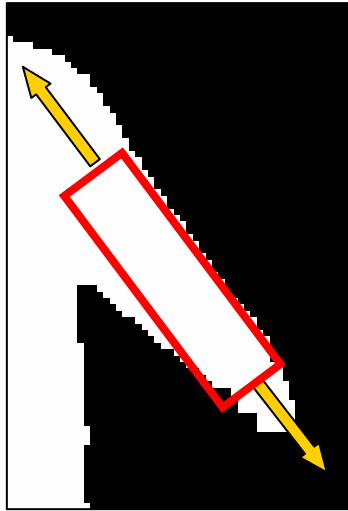


圖 4-10 搜尋肢體末端方法示意圖

我們採用的方法是以梯形區域的上底與下底的中間點為基準點，以垂直於上底與下底的方向，同時向外搜尋人體的輪廓邊緣，如在一定距離內找尋不到輪廓的邊緣，代表此梯形區域非肢體末端，藉此可以排除如身體軀幹等部位。如圖4-10所示，紅色框線為肢體的末端梯形示意圖，我們由梯形的上底與下底為基準點，以黃色箭頭的方向向外搜尋輪廓，而較下方的黃色箭頭將會迅速找到身體輪廓邊緣，只要兩個方向的其中之一搜尋到輪廓邊緣即視為肢體末端。結果如圖4-11所示，利用此方法我們可以獲得較為理想的肢體模型。

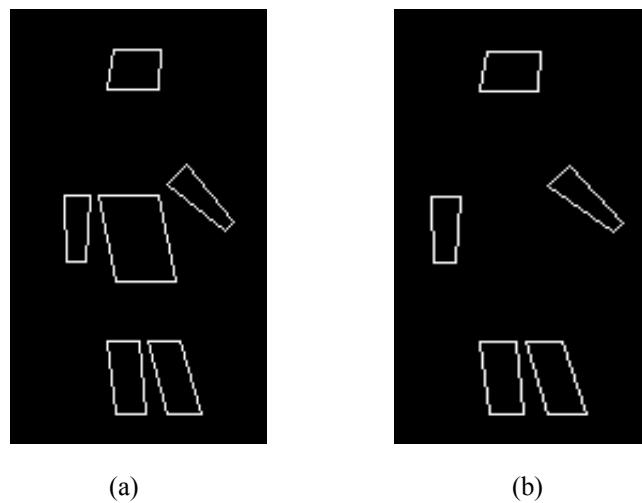


圖 4-11 利用搜尋肢體末端方法改善肢體模型 (a) 未使用肢體末端資訊，(b) 改善後的肢體模型。

4.4 肢體追蹤

由於人體的姿態會隨著連續影像的進行不斷的改變，單一影格中的人類影像資訊往往不夠我們分析人體的姿態與動作。因此我們利用分析前後影格中的肢體模型，來進行人體的肢體追蹤，作為分析影像中的人像活動之用。考慮在連續影像中，人類肢體的位置與長度變化有限，以及在計算量與正確性之間求取平衡，我們比對肢體模型的依據為長度以及重心的相對位置。首先我們將第一個搜尋到人體影像的每一個肢體模型都分別予以不同的編號，之後則會對於當前影像與上一張影像中的肢體模型予以交叉比對，當長度與重心位置在門檻值之內，則將編號設為相同，否則給予一個新的編號。在本論文之中，長度以及位置比對的門檻值分別為長度的變化量不能超過前一個模型的70%，以及重心位置的移動量小於20個像素值。最後我們將每一個模型的編號對應不同的顏色顯示，因此當追蹤到相同的肢體部位時，我們會給予當前梯形模型與上一張影像中比對相同的模型一樣的編號以及顏色，結果如圖4-12所示。

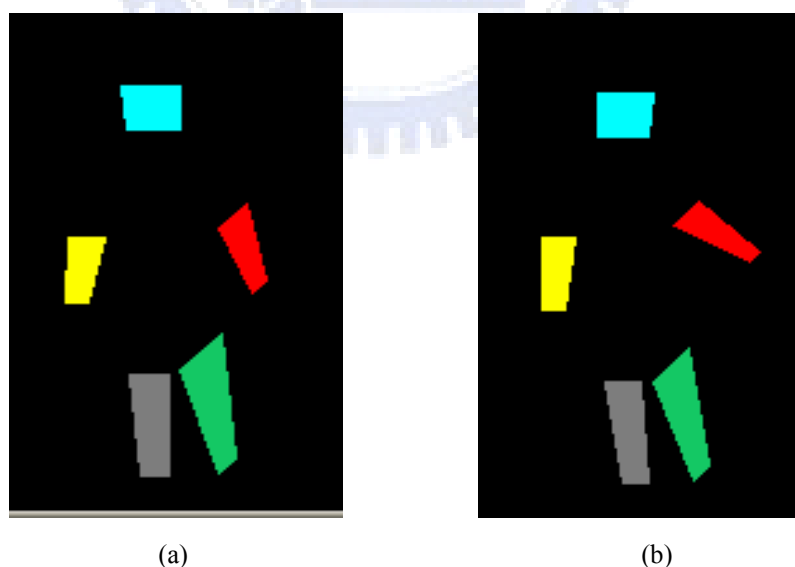


圖 4-12 連續影像中的肢體追蹤 (a) 前一張影像，(b) 後一張影像的肢體追蹤。

4.5 總結

在本章中，我們利用梯形表示法來代表掃瞄線演算法所產生的近似中軸線群組，並將四個方向掃瞄線所獲得的梯形予以整合，產生一個多節的梯形模型，再利用辨認末端肢體的方法，濾除非肢體的部位，建立一個無軀幹的多節梯形模型。最後利用肢體追蹤的方法，將不同的肢體部位使用不同的顏色模型來代表，以便我們瞭解並分析模型在連續影像中的活動情況。



五、實驗結果與討論

在本章節中，我們利用實作出來的系統來進行實際影像畫面的實驗，並證明掃瞄線演算法的確具有一定的正確性與實用性。本章分為五個小節，首先在第一小節我們討論前述章節中的系統參數對於我們產生的人體模型的影響，接著藉由第二小節提供一個多節梯形的無軀幹人像模型之肢體追蹤結果，並在第三小節之中顯示辨識末端肢體方法的正確性。之後透過第四小節的實驗來驗證本演算法不會因為偵測區域的旋轉影像其面積與長度的判斷。最後在第五小節由30位使用者根據我們的實驗結果，嘗試辨認被觀測者的行為，並將結果與實際被觀測者的行為互相對照，顯示本系統在醫療監控系統上具有一定的實用性。整個實驗的硬體環境為Pentium Core II 2.13 GHz，作業系統為Windows XP，程式作業環境為Visual C++ 6.0的平台。以下分別針對此五小節作詳細的實驗與分析探討。

5.1 辨識末端肢體的實驗結果

在掃瞄線演算法中，掃瞄線間隔距離 δ 直接影響對於原始影像的取樣精細程度， δ 越小模型的精細程度越佳，也越接近原始影像。如何讓計算量與模型的精準度之間求取一個平衡，是這個系統的重要課題。

我們的系統所使用的實驗影片的解析度為320X240，在這個解析度之下，所拍攝的對象肢體部位的寬度約為10個像素。由於人類肢體部位的形狀與染色體相似，根據[10]中所述，掃瞄線間隔距離為染色體寬度的1/2至1/3之間的結果較佳，故我們選擇以肢體寬度的1/3，也就是 $\delta=3$ 來進行掃瞄，以達到在運算量與模型精細程度上較為令人滿意的結果。 δ 也需視環境中被觀察對象離攝影機的遠近來進行調整，當被觀察者離攝影機越遠，就更需要間隔距離越短的掃瞄線來對原始影像進行取樣。

除了 δ 以外，系統參數尚有第二章所提到的過濾過長之掃瞄線，以及求取近似中軸線群組中時所使用的中點連線距離門檻值等共五個系統參數，這些參數都不需要使用者來進行設定，使用者只需在攝影機環境改變時，根據當前攝影影像的解析度，

以及攝影機與被觀察者之間的距離這兩個因素，設定掃瞄線の間隔距離。其他的系統參數都可以透過掃瞄線の間隔距離這項參數，動態的進行調整。

5.2 辨識末端肢體的實驗結果

在第四章的第三小節中我們曾提到辨識末端肢體並排除非肢體部分的梯形區域，在這個小節中我們以實際的影像結果來展示這個方法的實用性以及正確性，如圖5-1所示。圖5-1 (a) 我們首先輸入一段真實的影像畫面，並根據高斯混合模型計算出前景影像，獲得圖5-1 (b) 的結果，再透過第三章以及第四章的前兩小節的方法我們獲得四個方向整合後的梯形區域，如圖5-1 (c) 所示。最後透過辨識末端肢體的方法我們將非肢體區域的部分去除，得到我們的無軀幹多節梯形模型，如圖5-1 (d) 所示。圖5-2透過如圖5-1的處理流程，我們透過辨識末端肢體的方法，去除圖5-2 (c) 中與身體軀幹重疊的部分，獲得一個較為理想的無軀幹多節梯形模型，如圖5-2 (d) 所示。

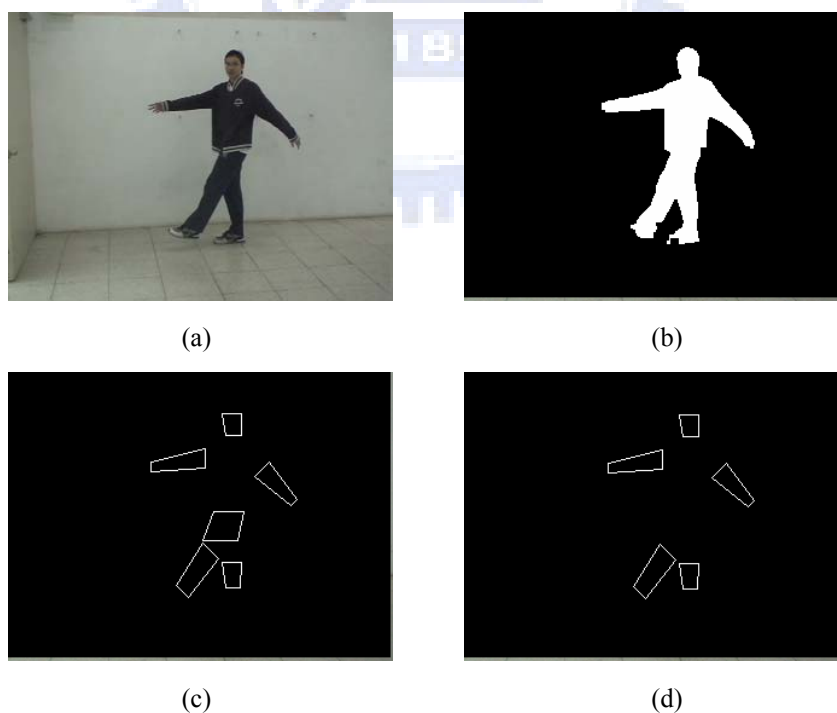


圖 5-1 利用辨認肢體的方法去除非肢體的區域 (a) 原始影像，(b) 二值化前景影像，(c) 未消除前，(d) 去除後之人體肢體的模型。

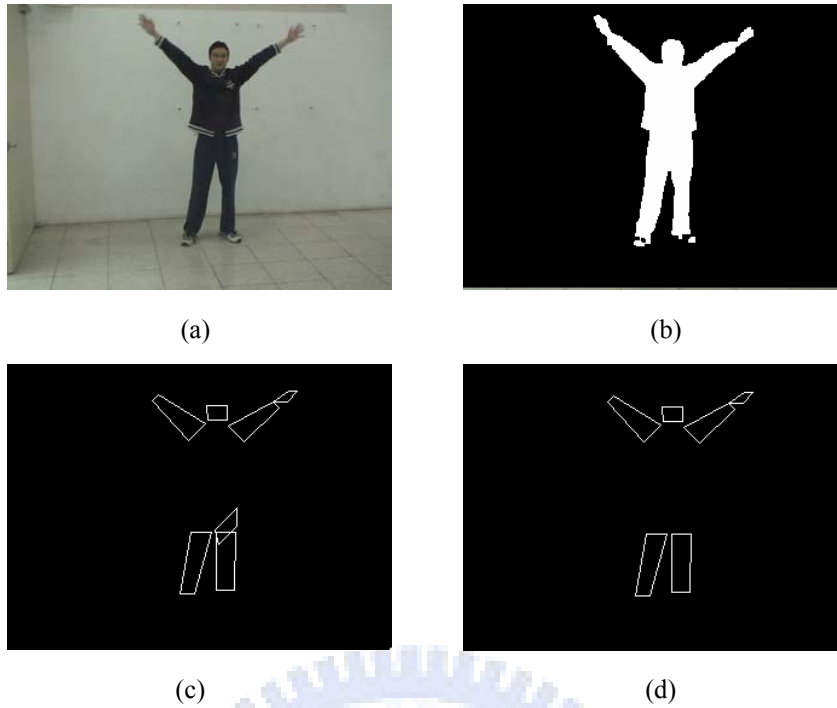


圖 5-2 利用辨認肢體的方法去除非肢體的區域 (a) 原始影像，(b) 二值化前景影像，(c) 未消除前，(d) 去除後之人體肢體的模型。

5.3 無軀幹人像模型的實驗結果

我們在這個小節除了展示無軀幹的人像模型的實驗結果之外，也進一步的利用第四章第四節的肢體追蹤方法，將初始建立的梯形分別予以編號，並以不同的顏色模型來代表不同的肢體模型。如果追蹤的方法辨識到當前影格中與前一個影格有相符的肢體部位，則會將當前影格中的肢體部位對應前一個影格中的肢體顏色模型，這個部分也將提供實驗的影像結果。如圖 5-3 所示，這是一段雙手由放下到雙手平舉的一段連續影片，圖 5-3 (a)以及圖 5-3 (c)分別為原始的影像中的第 1 個 frame 和第 10 個 frame，在經過高斯混合模型計算出前景影像後再予以二值化所得。圖 5-3 (b)和圖 5-3 (d)分別為圖 5-3 (a)和圖 5-3 (c)所產生的無軀幹模型，其中圖 5-3 (d)中與圖 5-3 (b) 顏色相同的區域代表追蹤到的相同部位的肢體模型。另外，圖 5-4 中人像活動的情形是由雙手平舉不斷抬高直到雙手高舉。而圖 5-5 則為人像側面面對拍攝鏡頭並朝右邊行走的實驗結果。

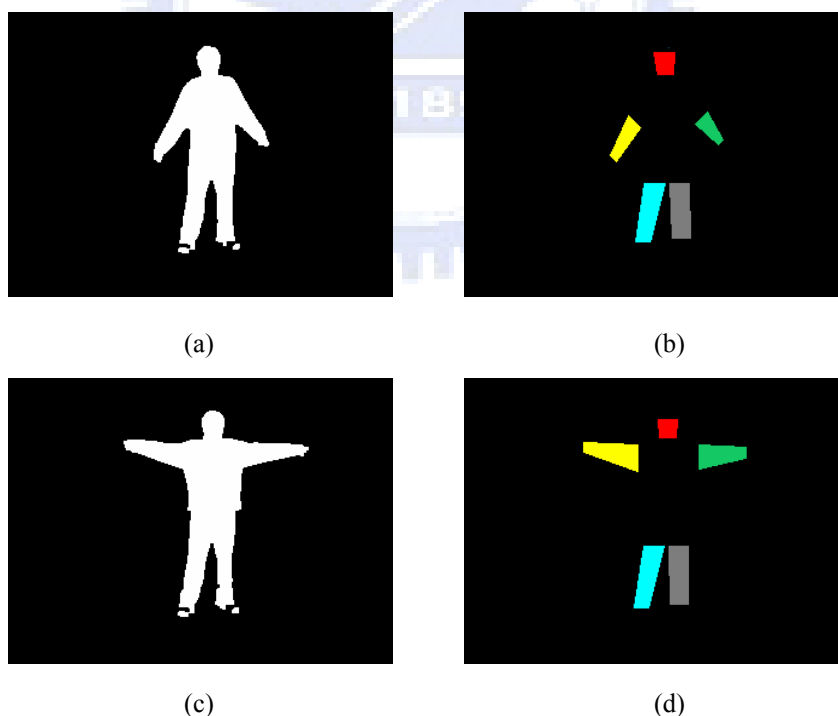


圖 5-3 人像雙手平舉的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無軀幹模型，(c)影片中第 10 個 frame，(d)產生的的無軀幹模型。

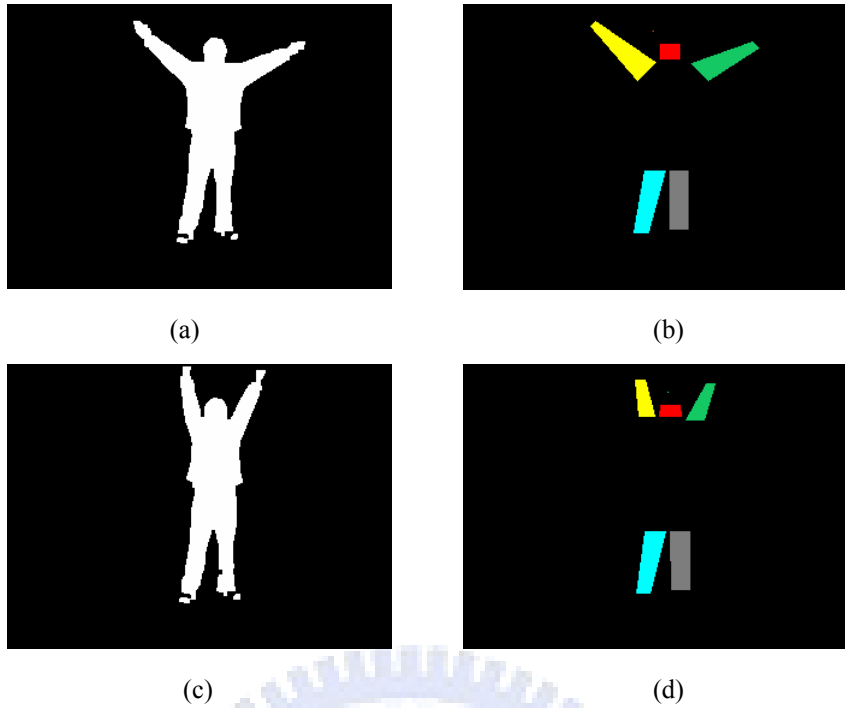


圖 5-4 人像雙手高舉的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無驅幹模型，(c)影片中第 24 個 frame，(d)產生的的無軀幹模型。

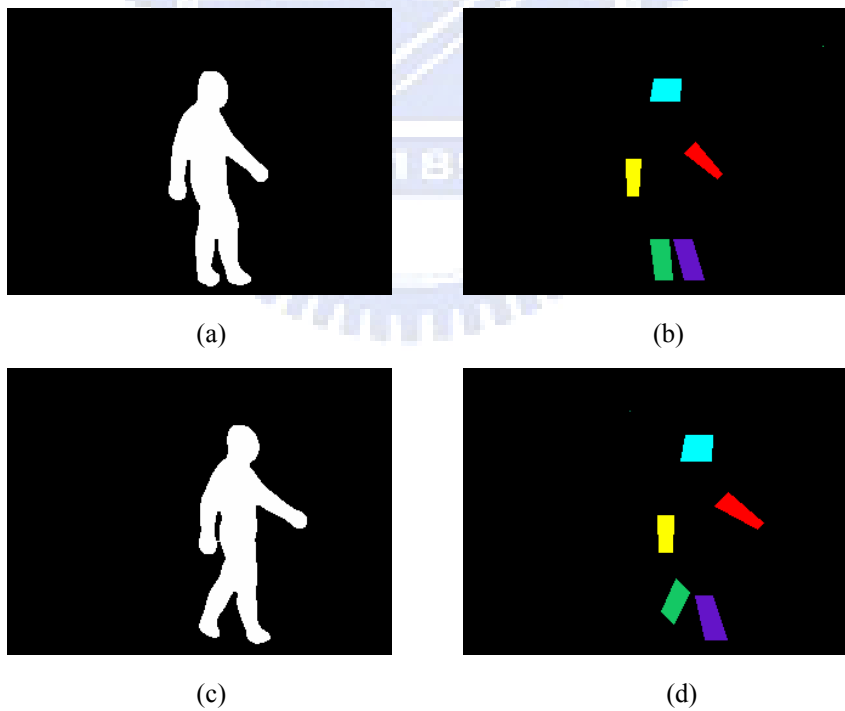


圖 5-5 人像向右行走的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無驅幹模型，(c)影片中第 10 個 frame，(d)產生的的無軀幹模型。

接著我們來觀察較不同的實驗結果，如圖 5-6 的人像在連續的影像中進行半蹲的動作，而圖 5-7 的人像則是在濕滑的地面上滑倒。根據我們的演算法，即使人像活動非立姿，仍可還算正確的判斷出人像中肢體的部位。

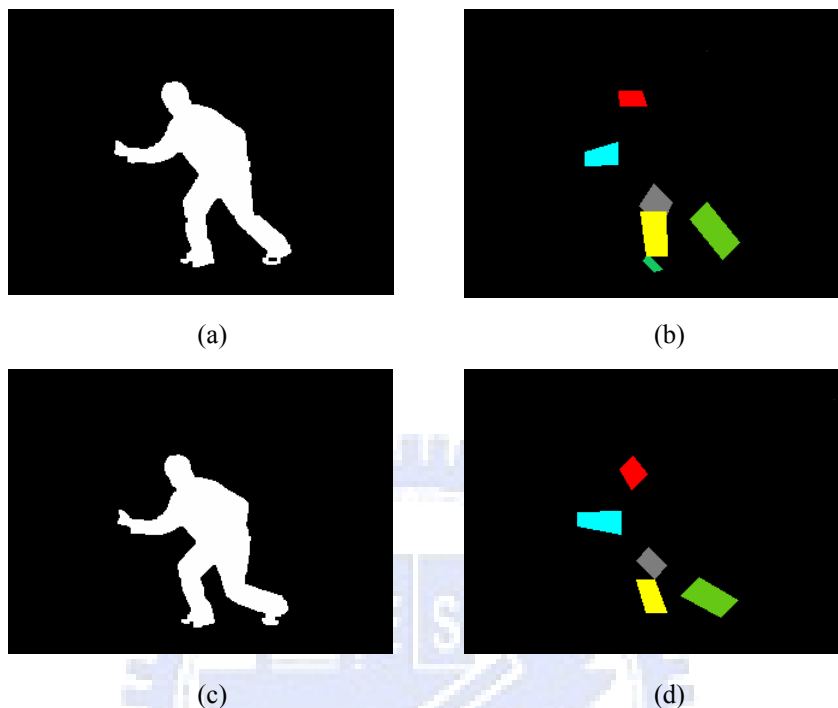


圖 5-6 人像半蹲的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無軀幹模型，(c)影片中第 4 個 frame，(d)產生的的無軀幹模型。

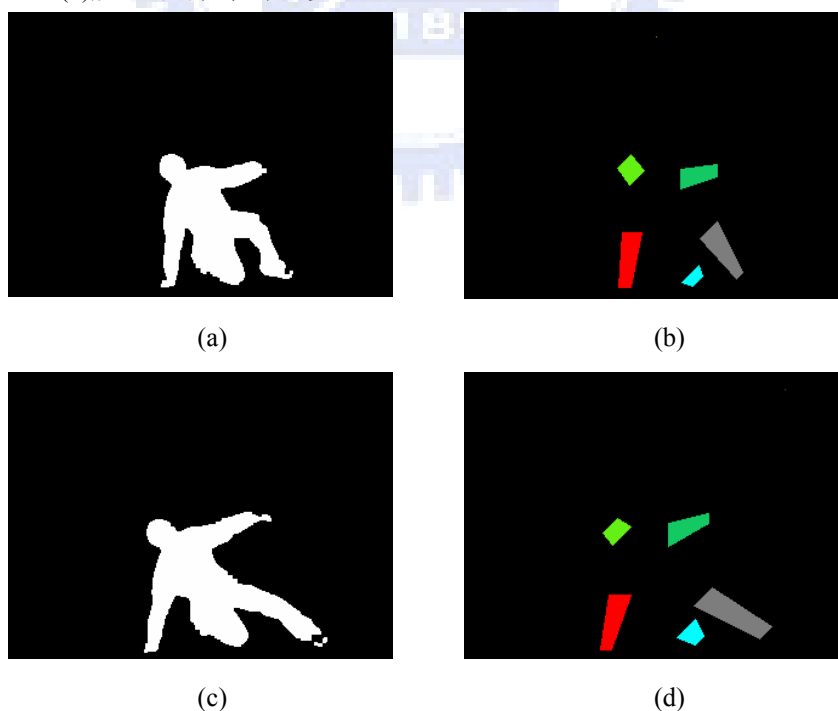


圖 5-7 人像在地面上滑倒的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無軀幹模型，(c)影片中第 9 個 frame，(d)產生的的無軀幹模型。

有些情況下，人像活動時肢體部位會被身體軀幹所遮蔽，由於我們的演算法一開始假設的情況是人像皆具有肢體部位，如果被遮蔽，在當下這個影格將無法搜尋到肢體部位。根據實際拍攝的影像我們發現，一個正常的人像活動不會長時間遮蔽其身體的任何肢體部位。因此當我們的系統應用於醫療監控系統之中，如果影像中的人像肢體部位一直被遮蔽，系統在一段時間內皆無法搜尋到任何肢體部位，我們也將視其為一個警告的資訊，提供給醫療人員參考。在圖 5-8 之中，人像側面朝向攝影機，並被身體軀幹遮蔽了其雙手，但因為雙腳以及頭部我們可偵測其區域，因此仍可持續追蹤這個人像的活動情況。

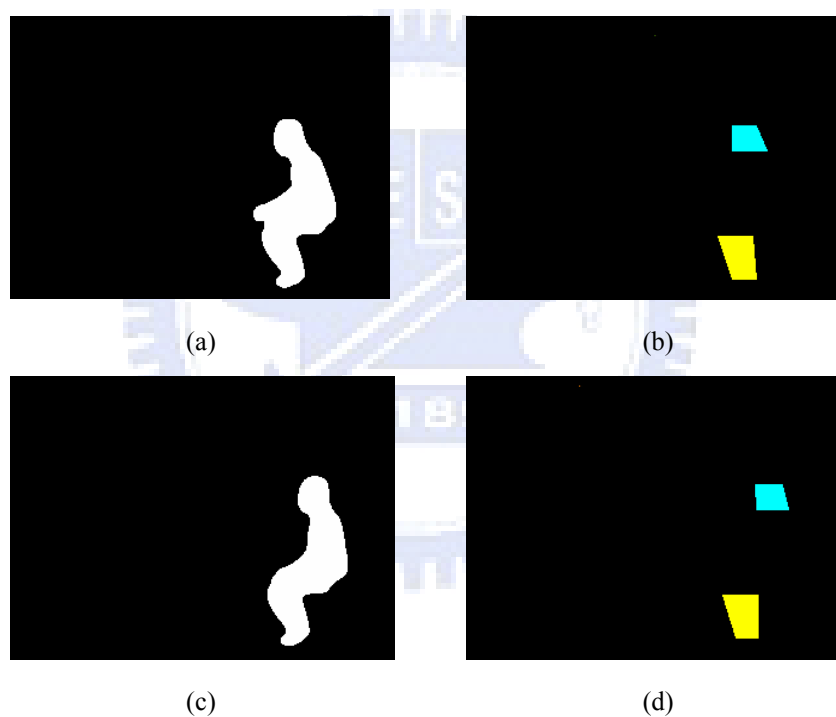


圖 5-8 人像坐下的實驗結果 (a)影片中第 1 個 frame，(b) 產生的無軀幹模型，(c)影片中第 7 個 frame，(d)產生的無軀幹模型。

5.4 計算旋轉矩型面積以及高度的實驗結果

由於在我們的系統中，掃瞄線的掃瞄角度有限，演算法中若僅使用四個方向的掃瞄線，是否會因為肢體區域的旋轉，而改變多節梯形區域的大小，造成人體模型的不準確，將在本小節中考慮。我們是利用計算單一矩形在旋轉過程中，其面積以及長度的變化情形，來評估因為只使用四個方向的掃瞄線，所造成的計算誤差。本實驗使用的矩型寬度為 20，長度為 70，以水平的方向開始作旋轉。由於我們的系統使用四種不同角度的掃瞄線，因此面積以及長度的變化量是以每 45 度為一個循環，故我們將矩型每次旋轉一度，總共旋轉 45 度，觀察一個循環之中的面積以及長度的變化情形，如圖 5-9 所示。

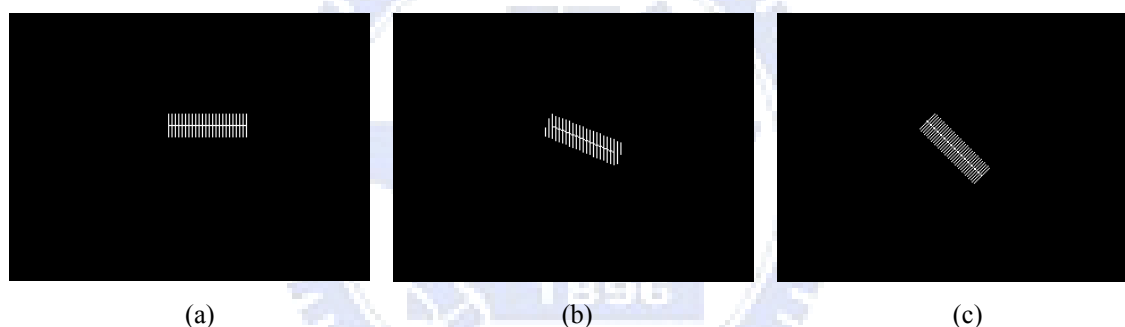


圖 5-9 掃瞄旋轉的矩型區域 (a) 未旋轉，(b) 旋轉 22 度，(c) 旋轉 45 度的矩型掃瞄結果。

我們可以發現當矩型旋轉至掃瞄線角度的交接處，掃瞄出的矩型區域將比實際上的矩型面積較小（越不準確）。不管是面積或是長度與原始矩型比較均有較大之誤差。由於面積的計算考慮到長與寬之間的變化，因此當矩型旋轉至掃瞄線角度的交界處附近（22.5 度），面積的正確率並不是最低，且呈現較不規則的變化，如圖 5-10 所示。但如果只觀察矩型長度的變化，我們可以發現當矩型旋轉至在掃瞄線角度的交界處，掃瞄出來的長度最短，反之則較為符合原本矩型的長度。在 1 到 45 度的旋轉過程中，長度大致上呈現一個週期性的變化，如圖 5-12 所示。在本實驗中，面積的整體正確率為 91.1%，長度的整體正確率為 92.8%，可以證明不論矩型旋轉的角度為何，掃瞄線演算法都有一定的準確度。

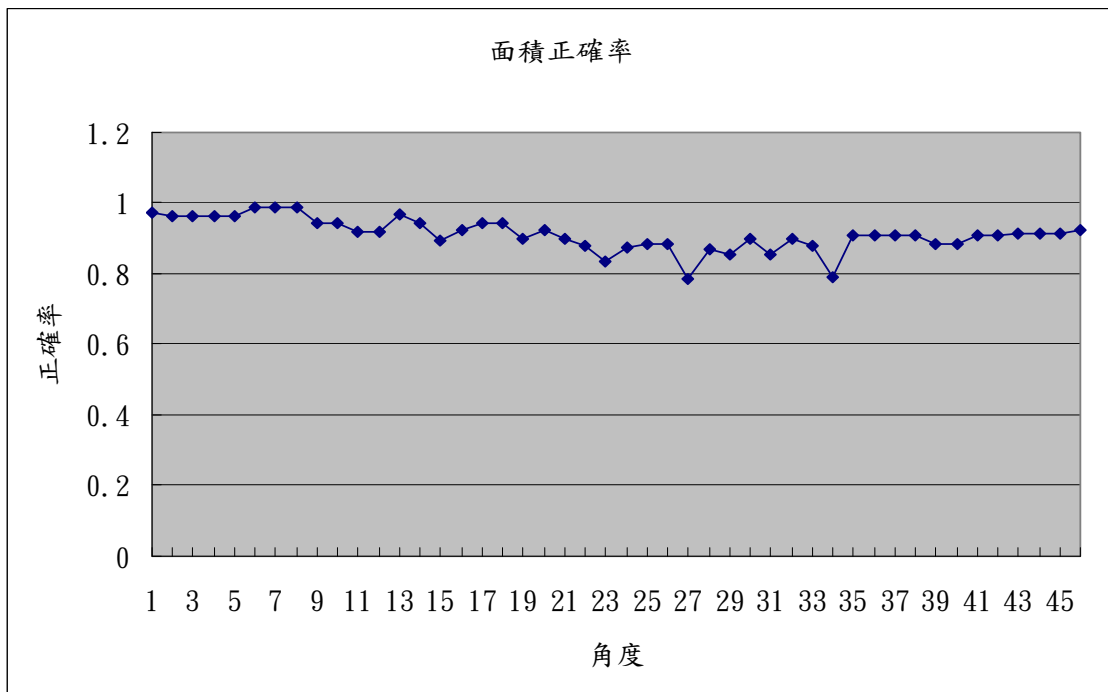


圖 5-10 掃瞄旋轉的矩型面積正確率

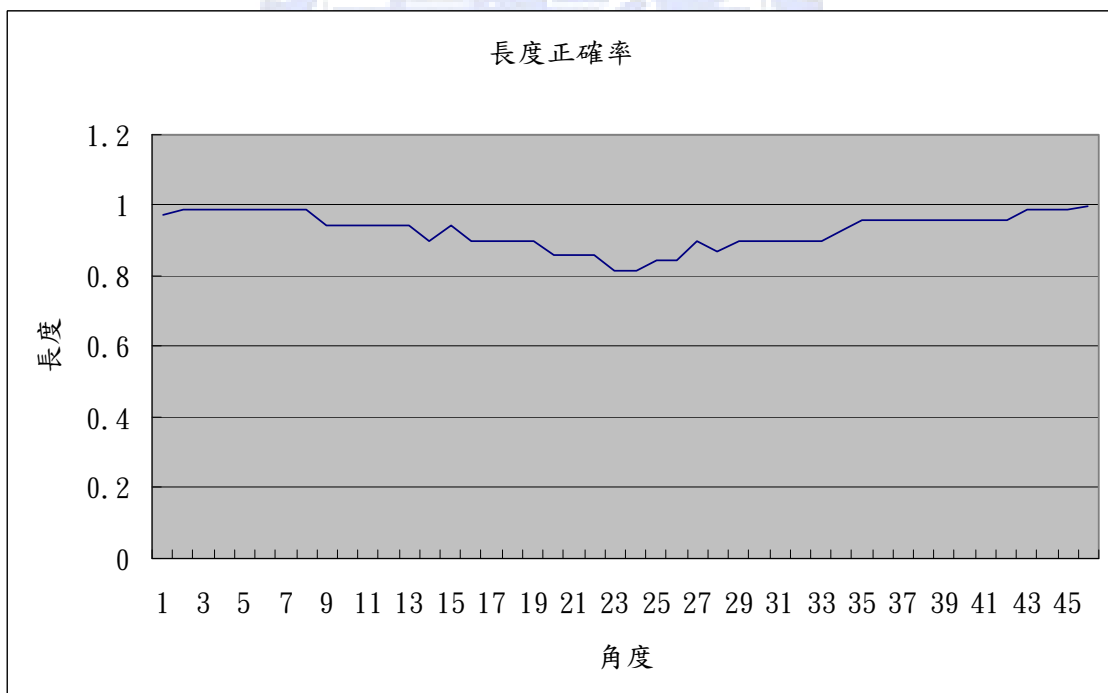


圖 5-11 掃瞄旋轉的矩型長度正確率

5.5 使用者辨識模型的實驗結果

由於本論文主要應用的領域在監視系統，因此使用者是否能辨識我們系統中被觀察對象的行為，是一項重要的課題。故我們設計了一套問卷，使用 5.3 小節的實驗結果，由首次接觸本系統模型的使用者來進行測驗，並記錄其辨識成功率，證明本系統的人體模型具有一定的實用性。本實驗共計有 30 位不同的實驗對象，男性對像為 14 位，女性為 16 位，平均年齡為 24 歲，實驗進行方式為使用者觀看一段類似 5.3 節的實驗結果影片（未包含原始影片），再從我們提供的 5 個動作選項中挑選答案作答，結果如圖 5-13 所示。由實驗結果可以發現，多數使用者大致上可辨認被觀測者的行為動作，整體正確率為 91.9%。使用者多數易辨認錯誤的行為是坐下這個動作，這是由於背景已被去除，在首次使用本系統的使用者易與其他姿勢（如蹲下）混淆，這是後續可以在進行改進的部分。此外較有趣的部分是男性辨認成功率較女性高出 5 個百分點，理工領域實驗對象辨識成功率更比文組對象高出 11 個百分點，是否使用者族群的不同也會影響辨識的成功率，這點相當值得後續的探討。

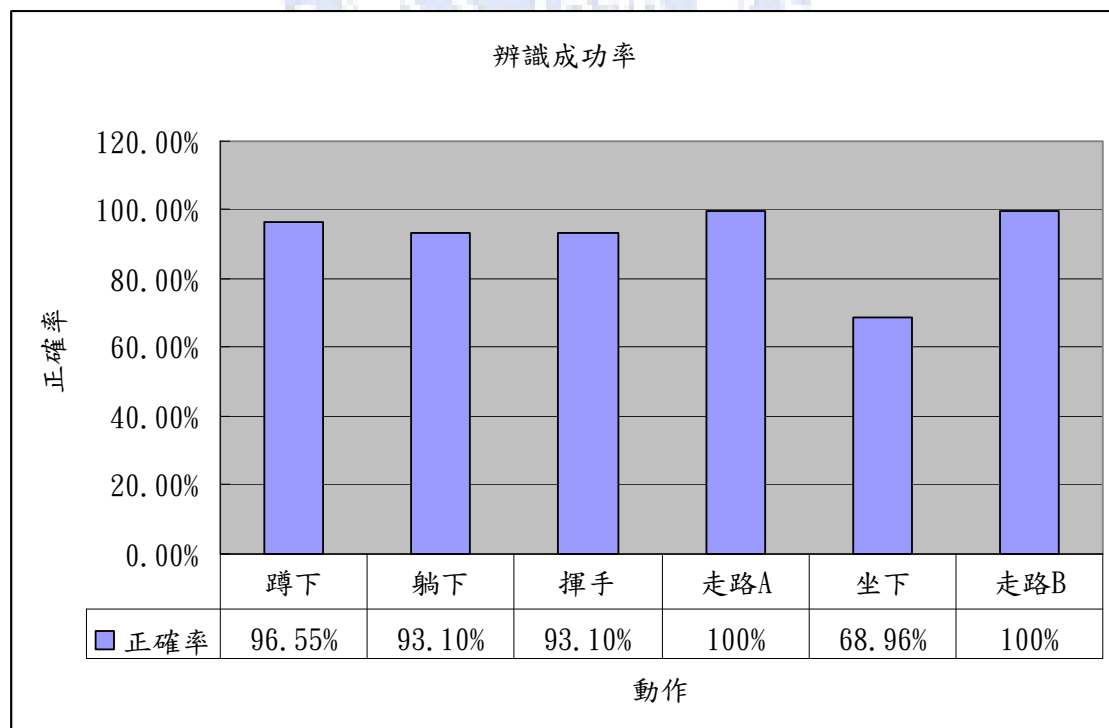


圖 5-12 使用者辨識模型的正確率

六、結論與未來工作

在本篇論文中，我們提出一個快速而有效率的演算法，搜尋人類影像中的肢體部位，儲存肢體部位形狀與角度的資訊，並提供一個基於多節梯形的無軀幹模型來代表人體，最後透過前後影像之間的關係，持續的追蹤相同部位肢體模型。由於本系統的人體模型已去除被觀測者的私人特徵，因此適合應用在注重被觀測者的隱私的醫療監控系統。此外，在醫療監控系統中被觀測者發生的情況往往需要緊急的處理，因而需要提供醫療人員更即時性的資料與行為分析結果，來掌握被觀測者的日常行為，透過我們的系統能建立具有較少資料量以及運算量的人體模型，並能兼顧符合原始輸入的人類影像的模型。值得注意的是，由於本系統利用單一的攝影機來產生人體模型，並未具有景深的資訊，此外也無法偵測到受到遮蔽的肢體部位。因此未來可發展多攝影機的系統來偵測被身體軀幹遮蔽的肢體部分，相信對於建立肢體模型會更為準確。除此之外，我們也希望可以利用建立的人體肢體模型以及肢體追蹤的結果來進行人像活動的分析，進一步的幫助我們自動化的辨認人像的各項動作。

參考文獻

- [1] B. Fan and Z. -F. Wang, "Pose Estimation of Human Body Based on Silhouette Images," *IEEE Int'l Conference on Information Acquisition*, pp. 296-300, June 2004.
- [2] M. K. Leung and Y. -H. Yang, "First Sight: A Human Body Outline Labeling System," *IEEE Transactions Pattern Analysis and Machine Intelligence*, Volume 17, Issue 4, pp. 359-377, April 1995.
- [3] C. Hu, Q. Yu, Y. Li, and S. Ma, "Extraction of Parametric Human Model for Posture Recognition Using Genetic Algorithm," *Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 518-523, March 2000.
- [4] S. Kim, C.-B. Park, and S.-W. Lee, "Tracking 3D Human Body Using Particle Filter in Moving Monocular Camera," *IEEE International Conference on Pattern Recognition*, Volume 4, pp. 805-808, 2006.
- [5] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette Analysis-Based Gait Recognition for Human Identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 25, Issue 12, pp. 1505-1518, Dec. 2003.
- [6] H. Fujiyoshi and A. J. Lipton, "Real-time Human Motion Analysis by Image Skeletonization," *Proceedings of 4th IEEE Workshop on Applications of Computer Vision*, pp. 15-21, Oct. 1998.
- [7] P. Guo and Z. Miao, "Projection Histogram Based Human Posture Recognition," *IEEE International Conference on Signal Processing Proceeding*, Volume 2, 2006.

- [8] Tyng-Luh Liu, Davi Geiger, Alan L. Yuille, "Segmenting by Seeking the Symmetry Axis," *Proceedings of 14th IEEE International Conference on Pattern Recognition*, Volume 2, pp. 994-998, Aug. 1998.
- [9] C.-C. Chen, J.-W. Hsieh, and Y.-Y. Huang, "Segmentation of Human Body Parts Using Deformable Triangulation," *IEEE International Conference on Pattern Recognition*, Volume 1, pp. 355-358, 2006.
- [10] Y. Yuan, Z. Miao, and S. Hu, "Real-Time Human Behavior Recognition in Intelligent Environment," *IEEE International Conference on Signal Processing*, Volume 3, 2006.
- [11] S. Dasgupta, "Learning Mixtures of Gaussians," *40th Annual Symposium on Foundations of Computer Science*, pp. 634-644, Oct 2002.
- [12] J.-H. Kao, J.-H. Chuang, and T. Wang, "Chromosome Classification Based on The Band Profile Similarity along Approximate Medial Axis," *Pattern Recognition*, Volume 41, pp. 77-89, Jan 2008.
- [13] Y.-T. Hsu, and J.-W. Hsieh, "Boosted String Representation and Its Application to Video Surveillance," *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 339-342, Dec. 2006.
- [14] G. Mori and J. Malik, "Estimating Human Body Configurations Using Shape Context Matching," *Proceeding of European Conference on Computer Vision*, Volume 3, pp. 666-680, 2002.
- [15] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder:Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 19, Issue 7, pp. 780-785, July 1997.