

國立交通大學

多媒體工程研究所

碩士論文

基於人物動作分析之影像日誌系統

Image Diary System Based on Human Activity/Movement
Analysis

研究生：陳暉晴

指導教授：莊仁輝 教授

中華民國一百年六月

基於人物動作分析之影像日誌系統
Image Diary System Based on Human Activity/Movement Analysis

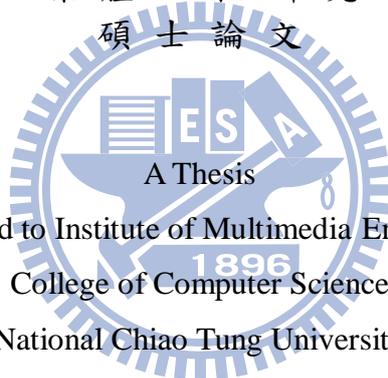
研究生：陳曄晴

Student：Wei-Ching Chen

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



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 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年六月

基於人物動作分析之影像日誌系統

學生：陳 暉 晴

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



在現今工商業社會中，年輕人日間忙於工作無法陪伴家中年邁長輩，為了提供有效的協助，讓家屬能安心工作，本論文利用電腦視覺之技術，透過監視攝影機影像來分析被觀測者的活動狀態、高度變化、行為動作，並識別各肢體部位建立人體姿態模型，據以發展一個具隱私保護、以及兼顧各個使用者對被觀測者關心層面不同之需求的階層式、方便瀏覽的居家看護影像日誌系統。此系統除了可以減輕看護與醫療人員的工作負擔，同時可以讓家屬經由監視畫面及時掌握居家年長者的生活狀況與安全。另外由於監視畫面經隱私保護處理，被觀測者也可以比較自在安心。本階層式系統架構不僅可應用於居家看護系統，同時也可以應用於其他活動，例如球類運動，提供使用者不同的需求。

Image Diary System Based on Human Activity/Movement Analysis

Student : Wei-Ching Chen Advisor : Dr. Jen-Hui Chuang

Institute of Multimedia Engineering

National Chiao Tung University

Abstract

The watermark logo of National Chiao Tung University is circular with a gear-like border. Inside, there is a stylized building with the letters 'ES' on it, a book, and the year '1896' at the bottom.

In today's society, young people are often too busy to accompany their family elders. In order to provide an effective assistance to family members, we propose a surveillance system based on computer vision technology to analyze people's active level, height changes, and action behavior, and recognize each limb to build a human body model for a display with privacy protection. Accordingly, a hierarchical home care image diary system is developed for convenient browsing, and to meet various concerns and demands. For example, the system can reduce nursing and medical staff's workload, while family members can use the monitoring screen to grasp the living conditions and safety of elders. Furthermore, since video images have been processed for privacy protection, people being looked after will feel more comfortable. In addition to home care system, this hierarchical idea of system architecture can be applied to videos of other activities, such as ball games, to serve various needs of the viewers.

誌 謝

本論文得以順利完成，首先要感謝指導教授莊仁輝博士的殷殷指導與諄諄教誨，並不厭其煩的導正觀念與耐心校正。老師不僅教導我獨立思考的能力與解決問題的方法，更隨時以身作則、潛移默化我許多人生哲理，使我在學業上與待人處世方面皆獲益良多，在此獻上最誠摯的敬意與謝意。同時也要感謝三位口試委員，楊谷洋教授、顏嗣鈞教授、賴飛熊教授，在百忙之中抽空提供寶貴的意見，讓本論文能夠更充實、更完整。

另外，感謝實驗室的學長姐們，慷慨分享自己的經驗與知識，啟發我研究上的靈感，特別感謝泓宏學長與國華學長給了我很多實務方面的指導。感謝親愛的同學們，怡廷、致傑、國濂、沛怡、俊憲、智愷、而益、士元，與可愛的學弟妹們，之容、柏安、偉成，給予我實驗上的鼎力相助及生活上的無限歡笑與回憶。感謝我摯愛的家人，你們無私的付出與鼓勵，對我無限的關懷與疼愛，讓我能夠無後顧之憂的專心學習，勇往直前。

最後，由衷的感謝這一路上所有關心我、幫助我的人，謝謝你們！

目 錄

中文摘要.....	i
英文摘要.....	ii
誌謝.....	iii
目錄.....	iv
圖目錄.....	vi
表目錄.....	viii
第一章 緒論	1
1.1 研究動機.....	1
1.2 研究背景與相關研究.....	2
1.2.1 人類行為認知.....	2
1.2.2 監控視訊瀏覽與檢索.....	3
1.3 論文貢獻.....	4
1.4 系統架構.....	5
1.5 章節說明.....	6
第二章 基於掃描線演算法之人體姿態識別.....	7
2.1 人體區域偵測.....	7
2.1.1 人物活動之前景區域偵測.....	8
2.1.2 雜訊濾除及前景中破碎區域填補.....	8
2.2 姿態識別.....	9
2.2.1 基於掃描線演算法之近似中心軸線分析與頭、手、腳定位.....	9
2.2.2 運用頭部位移之合理性判斷以提昇姿態識別之準確性.....	11
2.2.3 實驗結果與討論.....	15
第三章 階層式人物動作分析	19
3.1 活動狀態分析.....	19
3.2 高度變化分析.....	21
3.3 動作分類與行為分析.....	21
3.3.1 動作分類.....	22
3.3.2 行為分析.....	22

第四章 階層式影像日誌系統設計.....	25
4.1 系統需求分析.....	25
4.2 系統功能規劃.....	25
4.3 系統介面使用說明.....	26
第五章 實驗結果與討論.....	38
5.1 活動狀態分析結果.....	38
5.2 高度狀態分析結果.....	38
5.3 動作狀態分析結果.....	39
第六章 階層式人物動作分析之其他應用—籃球一對一鬥牛賽.....	41
6.1 球類活動之前景區域偵測.....	41
6.2 階層式狀態分析.....	42
6.2.1 比賽進行狀態分析.....	42
6.2.2 球高度變化分析.....	44
6.2.3 持球方動作分析.....	44
6.3 實驗結果與討論.....	45
6.3.1 系統介面.....	45
6.3.2 階層式狀態分析結果.....	46
第七章 結論與未來展望.....	49
參考文獻.....	50

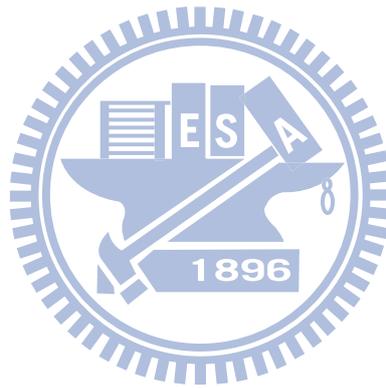
圖目錄

圖 1-1 階層式系統之視訊分析與顯示流程圖.....	5
圖 2-1 人體區域偵測流程圖.....	7
圖 2-2 中紅外線影像利用前景偵測找出之人體區域.....	8
圖 2-3 不同方向掃描的橫截面與中心軸線.....	9
圖 2-4 不同方向肢體可能的近似中心軸線.....	10
圖 2-5 不同方向近似中心軸線之整合.....	10
圖 2-6 肢體數量較多之判斷範例.....	11
圖 2-7 肢體數量較少之判斷範例.....	11
圖 2-8 運用[11]的方法之姿態識別結果.....	12
圖 2-9 頭部位移在合理範圍內之範例.....	12
圖 2-10 利用最近肢體改正頭部定位之範例.....	13
圖 2-11 沿用前一張頭部位置改正頭部定位之範例.....	13
圖 2-12 利用最近肢體新增頭部定位之範例.....	13
圖 2-13 沿用前一張頭部位置新增頭部定位之範例.....	14
圖 2-14 運用頭部位移之合理性改善[11]的頭、手、腳定位結果之流程圖.....	14
圖 2-15 人物走動然後坐下的影片姿態識別結果.....	15
圖 2-16 人物彎腰、蹲下及躺下的影片姿態識別結果.....	17
圖 2-17 少數肢體位置判斷錯誤之例子.....	18
圖 3-1 階層式人物動作分析流程圖.....	19
圖 3-2 人體質心示意圖.....	20
圖 3-3 利用人體外接矩形找出人物影像長度.....	21
圖 3-4 將人物影像長度分成三個區域.....	21
圖 3-5 正常躺下之人物高度變化.....	22
圖 3-6 跌倒情況下之人物高度變化.....	23
圖 3-7 跌倒之人物高度起伏例子.....	23
圖 3-8 將人物影像長度分成四個區域.....	23
圖 3-9 由人物高度變化偵測跌倒之流程圖.....	24
圖 4-1 系統功能架構圖.....	26
圖 4-2 完整系統介面操作流程圖.....	26
圖 4-3 選擇與輸入視訊檔案.....	27
圖 4-4 顯示人物活動狀態層.....	28
圖 4-5 選取人物活動狀態層中要更仔細觀看的時段.....	28
圖 4-6 人物活動狀態層之縮放功能.....	29

圖 4-7 選取要向下展開的高度狀態層時段.....	30
圖 4-8 顯示人物高度狀態層.....	31
圖 4-9 人物高度狀態層之縮放功能.....	31
圖 4-10 選取要向下展開的動作狀態層時段.....	32
圖 4-11 顯示人物動作狀態層.....	33
圖 4-12 人物動作狀態層之縮放功能.....	33
圖 4-13 選取要向下展開的姿態影像層時段.....	34
圖 4-14 顯示人物姿態影像層.....	35
圖 4-15 人物姿態影像播放視窗.....	35
圖 4-16 顯示檢索功能鍵.....	36
圖 4-17 人物活動中之檢索結果視窗.....	36
圖 4-18 人物靜止中之檢索結果視窗.....	36
圖 4-19 人物走路動作之檢索結果視窗.....	37
圖 4-20 人物站立動作之檢索結果視窗.....	37
圖 4-21 人物坐著動作之檢索結果視窗.....	37
圖 4-22 人物躺著動作之檢索結果視窗.....	37
圖 4-23 人物跌倒之檢索結果視窗.....	37
圖 5-1 影片一之活動狀態分析結果.....	38
圖 5-2 影片二之活動狀態分析結果.....	38
圖 5-3 影片一之高度狀態分析結果.....	39
圖 5-4 影片二之高度狀態分析結果.....	39
圖 5-5 影片一之動作狀態分析結果.....	40
圖 5-6 影片二之動作狀態分析結果.....	40
圖 6-1 中紅外線影像利用前景偵測找出之人體區域與球區域.....	42
圖 6-2 洗球範例.....	43
圖 6-3 在球與人分離的連續影像狀況下判斷比賽進行狀態之流程圖.....	43
圖 6-4 人與球之間的高度關係.....	44
圖 6-5 籃球一對一鬥牛賽之階層式動作分析系統介面.....	45
圖 6-6 洗球判斷範圍.....	46
圖 6-7 比賽進行狀態分析結果.....	46
圖 6-8 球高度狀態分析結果.....	47
圖 6-9 球被人遮蔽之例子.....	47
圖 6-10 持球方動作狀態分析結果.....	47
圖 6-11 投籃過程之範例.....	48

表目錄

表 2-1 人物走動然後坐下的影片姿態識別正確率之比較.....	16
表 2-2 人物彎腰、蹲下及躺下的影片姿態識別正確率之比較.....	18
表 5-1 活動狀態分析之正確率.....	38
表 5-2 高度狀態分析之正確率.....	39
表 5-3 動作狀態分析之正確率.....	40



第一章 緒論

1.1 研究動機

由於全民健保的實施及醫療品質的提昇，人類的平均壽命延長，人口結構趨向高齡化。根據行政院經建會的台灣人口推計[1]報告顯示，65 歲以上人口占總人口數比率在 2010 年為 10.7%，未來 10 年將增加為 16.3%，推估 50 年後（2060 年）更將增為 41.6%。更嚴重的是受傳統觀念所致，多數老年人希望子孫承歡膝下，對於進住老人安養機構意願不高。根據行政院內政部的老人狀況調查[2]報告顯示，僅 2.79% 的老年人居住於安養護機構，有高達 68.47% 的年長者選擇與子女同住。而在現今工商業社會中，就業機會增加，年輕人日間忙於工作無法陪伴家中年邁長輩，為了能及時掌握居家年長者的生活狀況與安全，一個可靠、具高親和力介面的居家看護系統將可以提供極大助益。

近年來，居家看護系統的研究多數以醫療照護為主，日常生活照護研究相對較少。因此發展一套以日常生活照護導向，適用於不需特別醫療照護老年人的居家看護系統也是有其需要性。由於各個使用者對被觀測者日常生活關心的主題不盡相同；有的人關心老年人是否有跌倒等危險，有的人關心老年人活動量是否足夠等等。為了達到具隱私保護、以及兼顧各個使用者對被觀測者關心層面不同之需求，本論文利用電腦視覺之技術，透過監視攝影機影像來分析被觀測者的活動狀態、高度變化、行為動作，並識別各肢體部位建立人體姿態模型，據以發展一個階層式、方便瀏覽的居家看護影像日誌系統。

1.2 研究背景與相關研究

居家看護系統的主要目的是辨識被觀測者之行為，並提供使用者有效且可靠的資訊。此外，監視資料量非常龐大，如何有效的管理與應用也是重要的研究課題。因此本論文相關研究主題可分為人類行為認知、監控視訊瀏覽與檢索兩大方向進行討論。

1.2.1 人類行為認知

近年來，有關人類行為認知的應用與研究愈來愈廣泛，發展也很迅速，常見的主題有以下幾種：姿勢與動作分類、異常行為分析、姿態識別等。姿勢與動作分類是利用具學習功能的分類器或若干規則，來分類被觀測者是處於站立、走路、跑步、坐下、蹲下、躺下等姿勢動作中的哪一個。異常行為分析是利用人體特徵或分類器來分辨被觀測者的姿勢動作是屬於正常行為，或是跌倒、徘徊等異常行為。姿態識別是利用人體較為顯著的末端特性，偵測出被觀測者的頭、手、腳、軀幹等位置。

姿勢與動作分類的研究中，有一種方法是利用若干規則進行分類。如文獻[3]將取得的人體輪廓，藉由收集與整理日常生活中常見姿勢的共通點與特徵，設計出一套簡單的參數與規則判斷人體姿勢。又如文獻[4]使用單一攝影機偵測移動物體並產生剪影，利用掃描線找出頭與雙腳端點產生人體骨架，藉由骨架特徵判斷人體動作，再利用若干規則，進一步區別不同週期的類似動作（例如走路、跑步）。另一種方法是利用分類器進行分類。如文獻[5]使用色度（chromaticity）和梯度（gradient）雙層背景減去法取得人體輪廓，並使用星狀骨架作為辨識特徵，再利用支持向量機（Support Vector Machine, SVM）分類器定義動作，藉由字串匹配（string matching）方式分類人體動作。還有一種方法是結合環境位置進行動作推斷。如文獻[6]是利用單一攝影機，劃定出環境中不同位置區塊，再透過人所處的位置，推測人體動作（例如在廚房煮飯、在浴室刷牙）。

異常行為分析的研究中，有一種方法是擷取人體特徵後，以固定規則進行分析。如文獻[7]將人體以一個近似橢圓形表示，結合人體投影直方圖和頭部位置資訊得到偏移角度與長短軸比值，藉由角度與長短軸比值變化判斷人體動作，並進行跌倒偵測。另一種異常行為分析的方法是利用具學習功能的分類器進行分析。如文獻[8]基於非監督式學習（unsupervised learning）建立一個階層式架構，使用背景減去法將人體區域分割出來，經由人體外觀分類與動作序列比對，進行異常行為偵測。

姿態識別的研究中，比較常見的是利用樹狀結構找出人體的頭、軀幹、手、腳等位置。如文獻[9]將人體分成十個區域，利用由軀幹為出發點的樹狀結構、對稱區域顏色相似以及區域間有無連動影響等三種區域關係做為判斷依據，最後將各區域用矩形表示，並涵蓋人體最大可能範圍。姿態識別的另一種方法是擷取人體各部位的特徵資訊與資料庫進行比對。如文獻[10]使用由上而下（up-down）的樹狀結構將人體各區域與資料庫做比對，並以不同顏色的點表示對應的人體部位。還有一種方法是利用頭、手、腳在人體中較為顯著的特性進行識別。如文獻[11]將擷取出的人體區域，使用水平、垂直和斜角等不同方向的掃描線來擷取橫截面，利用適當規則來連接橫截面之中心點，取得該肢體的近似中心軸線，並整合不同方向的近似中心軸線，選出代表肢體末端的近似中心軸線，再利用若干規則識別頭、手、腳。

1.2.2 監控視訊瀏覽與檢索

近年來，監控視訊技術愈來愈被重視，研究與發展也愈來愈廣泛，有以下幾種常見主題：監控視訊播放、監控視訊瀏覽、監控視訊檢索。監控視訊播放技術的發展提供功能更強大的視訊播放器，讓使用者能享受方便、高品質的服務。監控視訊瀏覽技術的發展讓使用者可更快速了解整個影片內容，提昇監控效率。監控視訊檢索技術的發展讓使用者可更快速搜尋到有興趣的部分。

監控視訊播放的研究中，比較常見的是設計使用者介面與提供不同播放功能。如文獻[12]顯示各段影片的關鍵影像（key frame），讓使用者可快速切換至有興趣的段落，省去在播放時間軸拖曳滑鼠的動作，並在影片播放結束時，根據影片關鍵字，由電腦硬碟或網際網路搜尋相關的影片，供使用者參考。

監控視訊瀏覽的研究中，有一種方法是同時顯示不同時間點出現的移動物件。如文獻[13]在單位時間內，將所有移動物件放在同一畫面播放，藉以縮短影片長度，讓使用者可在短時間內瀏覽所有監視影像。另外一種方法是分析移動物件的軌跡。如文獻[14]根據移動物件將影片分段，分析物件的移動軌跡，將各移動物件的關鍵影像與軌跡結合，藉以提供影片摘要。還有一種方法是根據影片內容提供影片大綱的簡單流程圖。如文獻[15]分析影片中的人物動作，再根據不同動作的前後發生關係，畫出影片內容的主軸動作發展流程圖。

監控視訊檢索的研究中，常見的方法是將影片切割以分鏡（shot）為單位作處理，根據每一個分鏡選取適當的關鍵影像作為代表。文獻[16-20]利用影像的顏色、紋理、文字等特徵將影片切割成一個一個分鏡。文獻[21-22]使用階層式架構，將相同背景與物件下的分鏡以場景（scene）為單位進行影片分割。文獻[23-24]分析影片中的人物姿勢與動作，根據不同的姿勢動作將影片分段。另一種監控視訊檢索方法是使用者可輸入關鍵文字或影像，搜尋影片中有興趣的部分。如文獻[25]利用手繪板畫出欲搜尋的物件與物件移動方向，根據物件與移動線索，以內容為基礎進行影片檢索。

1.3 論文貢獻

針對上述人類行為認知以及監控視訊瀏覽與檢索之研究，發現很少有結合兩個領域進行研究，即便有也是單從人類動作的角度進行影片切割。因此，本論文主要研究貢獻在發展一個基於人物動作分析的階層式影像日誌系統，並提供使用者方便操作及快速瀏覽的使用者介面。此階層式系統架構可應用於居家看護系統，除了可以減輕看護與醫療人員的工作負擔，同時可以讓親友經由監視畫面及時掌

握居家年長者的生活狀況與安全。另外由於監視畫面經隱私保護處理，被觀測者也可以比較自在安心。此階層式系統架構也可應用於球類運動訓練與比賽計分之輔助，甚至可應用於購物中心、車站及娛樂場所等之顧客流量分析，以進行收銀與售票櫃檯的人力機動調度。

1.4 系統架構

圖 1-1 為本論文的階層式系統之視訊分析與顯示流程圖，由攝影機取得影像後，偵測人體區域，根據人體區域擷取人體特徵進行階層式動作分析，結合各階層的分析資訊發展影像日誌系統。

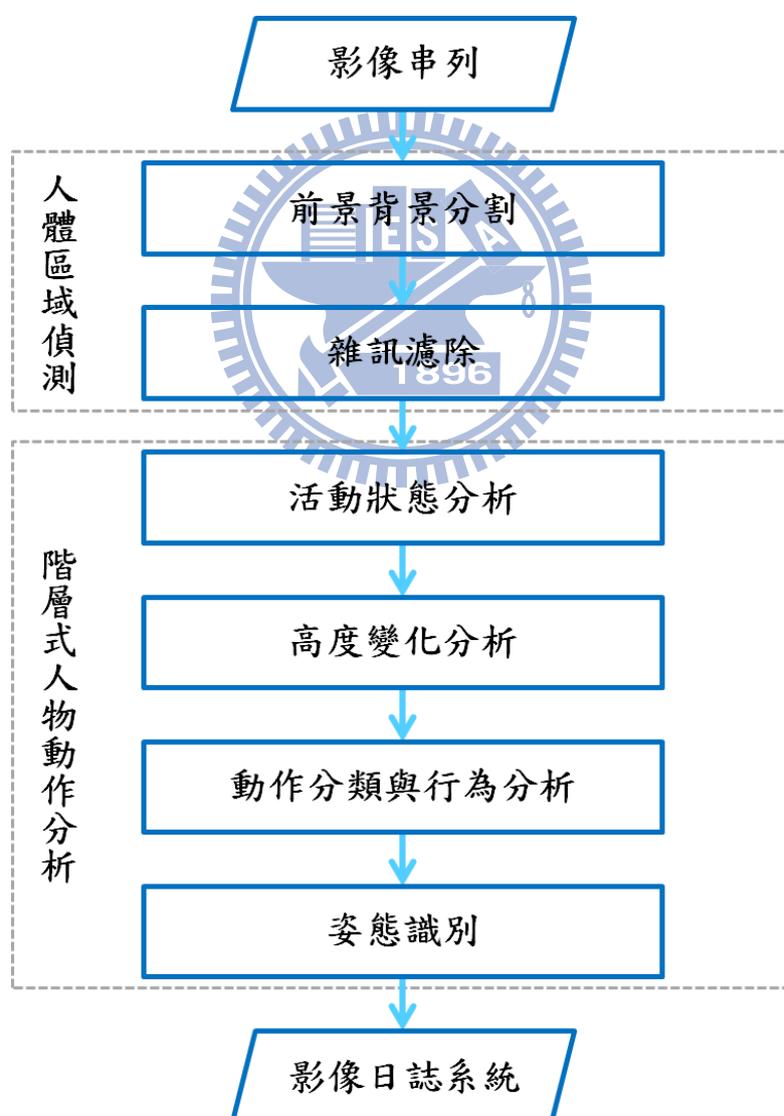
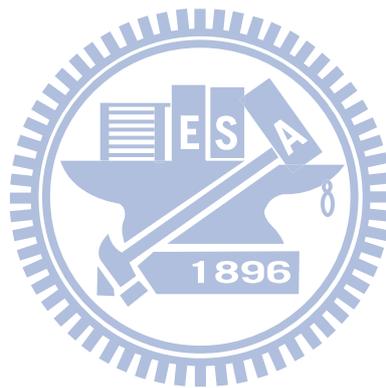


圖 1-1 階層式系統之視訊分析與顯示流程圖。

1.5 章節說明

本論文全文共分成七章，第一章介紹研究動機與目的，並概略介紹所設計之系統架構，第二章使用前景背景分割及雜訊濾除等技巧來擷取人體區域，並利用掃描線演算法識別各肢體部位建立人體姿態模型。第三章說明利用階層式概念擷取人體特徵進行分析，包括人物活動狀態、高度變化、動作分類、行為分析。第四章介紹本論文所設計的階層式影像日誌系統，包含如何操作使用者介面。第五章為實驗結果與討論。第六章利用階層式系統架構於其他應用，驗證整體設計上的可行性。第七章為結論與未來展望。



第二章 基於掃描線演算法之人體姿態 識別

為了能夠對人體區域進行肢體及姿態分析，本章首先在 2.1 節中敘述如何將人體區域從影像中分離出來，然後在 2.2 節中將敘述如何利用 2.1 節中所得到的
人體區域進行人體姿態識別。

2.1 人體區域偵測

常見的影像來源有可見光影像與中紅外線影像兩種，前者的像素值代表的是顏色，而後者的像素值代表的是溫度，所以背景上較前者為單純。本論文的研究重點是在前景偵測之後的分析，所以選用前景偵測較為單純的中紅外線影像作為輸入值，並針對其特性設計一套適合的前景背景分割法。本節旨在敘述如何針對輸入影像之特性，使用適當的前景背景分割法來擷取人體區域，並將雜訊予以濾除。人體區域偵測之詳細流程圖請參考圖 2-1。

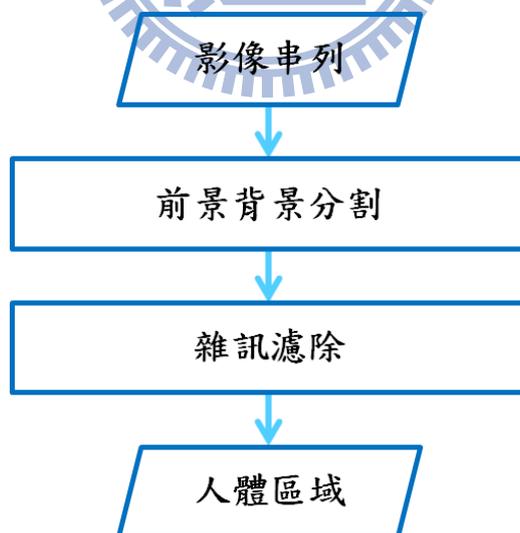


圖 2-1 人體區域偵測流程圖。

2.1.1 人物活動之前景區域偵測

由於中紅外線影像的像素值代表的是溫度，而同一物體的溫度值會非常接近，根據此特性使用式子(2-1)對中紅外線影像做二值化，以擷取出中紅外線影像中屬於較高溫部分的人體區域。

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

其中 $f(x, y)$ 是中紅外線影像中座標 (x, y) 的 16 位元溫度值， TH_{temp} 是二值化的門檻值，經由實驗設定當室內溫度較高時其值為 35000，而溫度較低時其值為 26500。

2.1.2 雜訊濾除及前景中破碎區域填補

由於 2.1.1 節所擷取出的人體區域會有雜訊及破碎的情形發生，如圖 2-2 所示，(b) 為使用 2.1.1 節提及之人體區域擷取方法得到的 (a) 之前景影像，可以看出找到的前景影像中有雜訊及破碎的情形。我們使用連通元件標記法 (connected component labeling) 區分不同區域並計算各區域所涵蓋的面積大小，將面積過小的區域濾除，並利用型態學上的技巧：侵蝕 (erosion) 及膨脹 (dilation) 來消除剩餘之雜訊及填補前景影像中破碎之區域，結果如圖 2-2(c) 所示。可以明顯看出，原本在圖 2-2(b) 存在的細碎區域可以被濾除，而人體區域中的空洞可以被填滿。

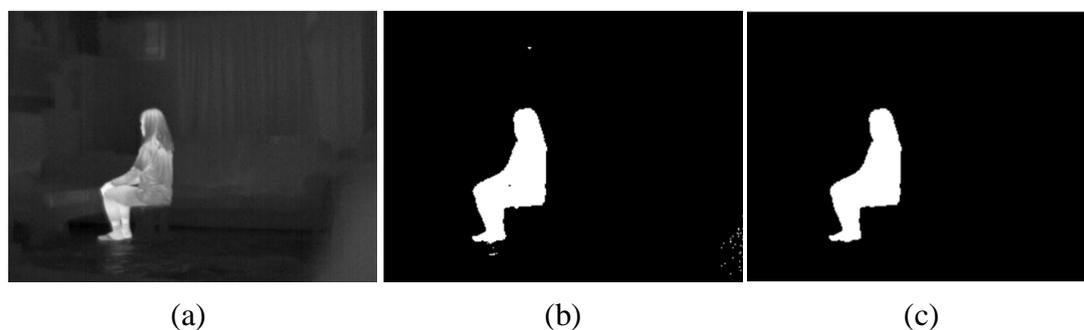


圖 2-2 中紅外線影像利用前景偵測找出之人體區域：(a) 原始影像，(b) 經由取門檻值之二值化前景影像，(c) 濾除過小面積及經過侵蝕、膨脹技巧處理之前景影像。

2.2 姿態識別

首先我們在 2.2.1 節中將 2.1 節所得到的人體區域利用[11]的概念擷取人體肢體的近似中心軸線進行頭、手、腳定位。在 2.2.2 節中我們將根據常理判斷，改進與補強[11]的頭、手、腳定位方法，得到更準確的人體姿態識別模型。在 2.2.3 節中將顯示經 2.2.2 節改善後的姿態識別結果，並與[11]之識別結果進行比較。

2.2.1 基於掃描線演算法之近似中心軸線分析與頭、手、腳定位

本節將簡述[11]的作法，包含以下四個主要步驟，分別為：(1) 肢體近似中心軸線分析，(2) 標示肢體末端，(3) 整合近似中心軸線，(4) 姿態識別—頭、手、腳定位。

(1) 肢體近似中心軸線分析

首先以 0° 、 45° 、 90° 和 135° 四個不同方向掃描人體區域並擷取出其橫截面，如圖 2-3 中的藍色線段所示。由於橫截面中心點連接而成的近似中心軸線可代表局部肢體的特徵，所以將屬於同一部分的肢體橫截面中心點相連，並考慮人體的肢體部分具有一定長度且不太可能呈現過寬或過扁的形狀，將不合理的中心軸線濾除，獲得較適合代表肢體的近似中心軸線，如圖 2-3 中的黑色線段所示。

(2) 標示肢體末端

利用端點的概念，選出更適合代表肢體的近似中心軸線。由圖 2-3 可清楚地看出，接近肢體末端的近似中心軸線，其橫截面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。利用此特性將這一類近似中心軸線之起始（終止）橫截面的中心點標示為末端點，如圖 2-4 中水藍色點所示。

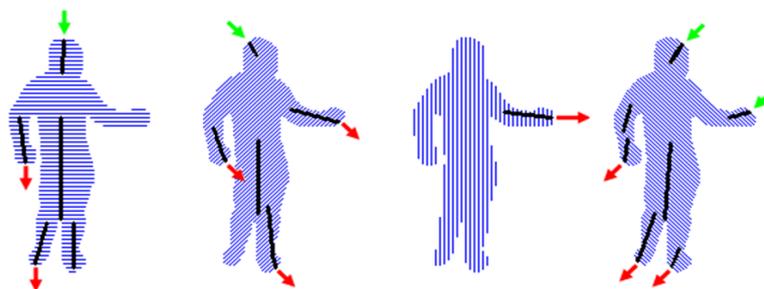


圖 2-3 不同方向掃描的橫截面與中心軸線，由左而右分別為 0° 、 45° 、 90° 、 135° 。

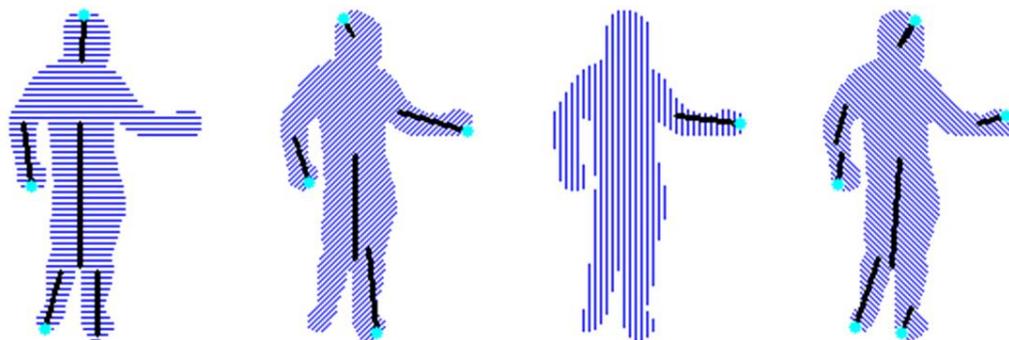


圖 2-4 不同方向肢體可能的近似中心軸線，由左而右分別為 0° 、 45° 、 90° 、 135° 。

(3) 整合近似中心軸線

在取得四個方向的近似中心軸線後，由於同一肢體可能有多個候選的近似中心軸線，所以此步驟將選出一個最合適代表該肢體的近似中心軸線。

首先，考慮空間上的關係，將位置較相近且含有末端點的近似中心軸線分為同一集合，如圖 2-4 會分成頭、右手、右腳、左手、左腳五個集合。由於近似中心軸線和橫截面垂直方向的夾角越小，越能代表該肢體的走向。所以設定最小夾角的容許範圍，並從中挑選最長的近似中心軸線，圖 2-5 為整合結果之示意。

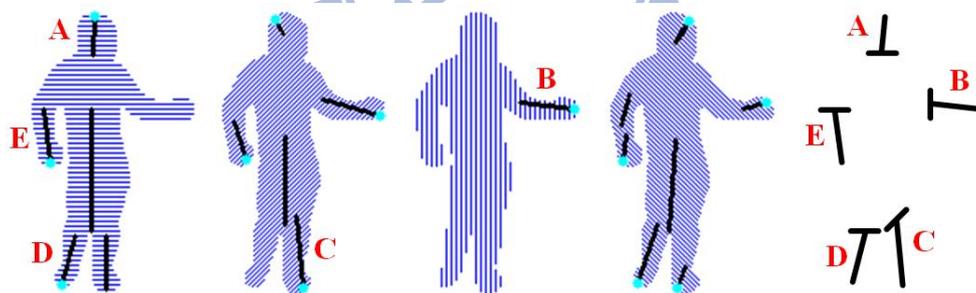


圖 2-5 不同方向近似中心軸線之整合，由左而右分別為 0° 、 45° 、 90° 、 135° 、整合結果。

(4) 姿態識別—頭、手、腳定位

在找出代表各肢體的近似中心軸線後，進一步地利用近似中心軸線的資訊，判斷該近似中心軸線是代表哪一個肢體部位。針對找到的肢體數量多寡進行不同方式之判斷，當肢體數量較多時（4、5 個），利用角度及相對位置判斷頭、手、腳，如圖 2-6 所示由於脖子的長度較肢體短，所以頭的角度會比肢體大，故 A 判斷為頭，距離頭較近的 B、E 判斷為手，而距離頭較遠的 C、D 判斷為腳；當肢

體數量較少時（1、2、3 個），使用區域分佈（region splitting）的方式來找頭、手、腳，如圖 2-7 所示 A 位於手區域故判斷為手，而 B 位於腳區域故判斷為腳。

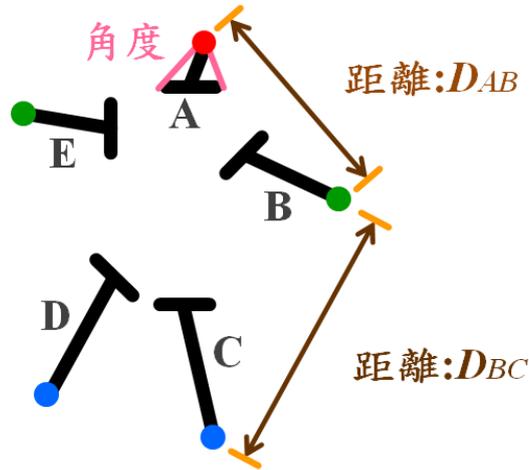


圖 2-6 肢體數量較多之判斷範例。

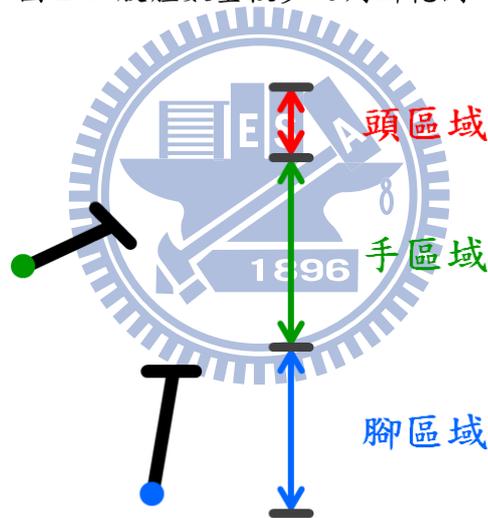


圖 2-7 肢體數量較少之判斷範例。

2.2.2 運用頭部位移之合理性判斷以提昇姿態識別之準確性

由於[11]所採用的頭、手、腳識別方法，僅針對單張影像，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 2-8 為連續三張影像之識別結果，由於手逐漸往身體靠攏，造成圖 2-8(b)原本是手的部分被識別成頭，這顯然是誤判。本節將說明如何運用時間軸的資訊技巧來避免這些錯誤。

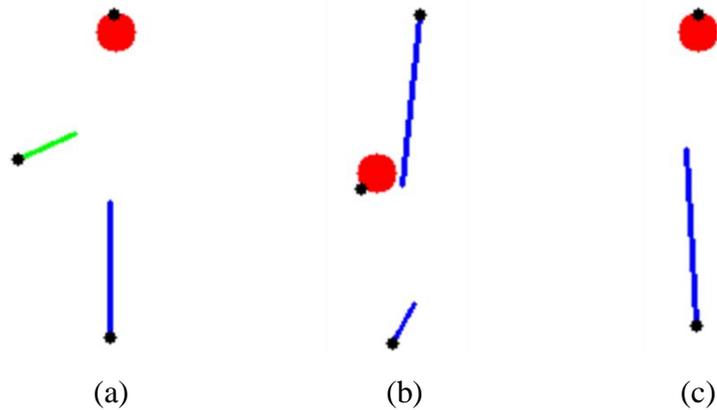


圖 2-8 運用[11]的方法之姿態識別結果:(a) 第 $t-1$ 張影像,(b) 第 t 張影像,(c) 第 $t+1$ 張影像。

我們知道在實際的肢體運動中，單位時間內頭部的移動量相較於手腳的移動量必然小得多。因此我們檢視在 2.2.1 節中運用[11]的方法得到的頭、手、腳定位結果，將頭部沒找到或明顯掉落等不合理的結果再加斟酌，希望有所改善。本論文所採用的改善方法主要是以頭部位移的合理性為依據，目標是讓每一張影像都能夠顯示合理的頭部位置。

我們將[11]之頭、手、腳定位結果，分為有找到頭部與沒有找到頭部兩種情況進行討論。在有找到頭部的情況下，我們首先計算與前一張頭部的位移量，如果位移量在頭部移動的合理範圍內，則判定[11]之定位結果正確，不做任何更改，如圖 2-9 所示。如果位移量不在頭部移動的合理範圍內，則計算各肢體與前一張頭部的最近距離，若最近距離小於門檻值，則表示此肢體位置應該是頭部，進而將其餘肢體重新進行定位，如圖 2-10 所示。若最近肢體與前一張頭部距離大於門檻值，則沿用前一張的頭部位置，並重新定位其餘肢體，如圖 2-11 所示。

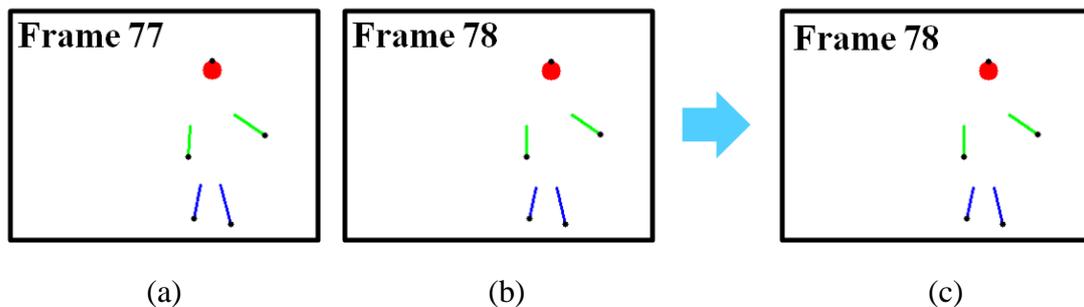


圖 2-9 頭部位移在合理範圍內之範例:(a) 前一張影像的頭、手、腳定位結果，(b) [11]之本張影像定位結果，(c) 本論文之本張影像定位結果。

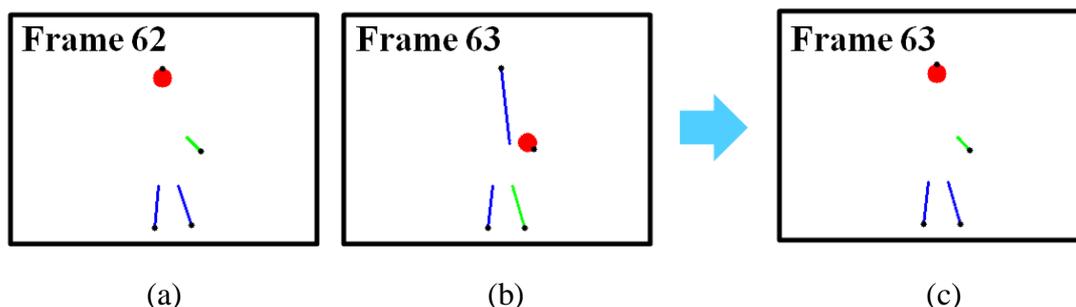


圖 2-10 利用最近肢體改正頭部定位之範例：(a) 前一張影像的頭、手、腳定位結果，(b) [11]之本張影像定位結果，(c) 本論文之本張影像定位結果。

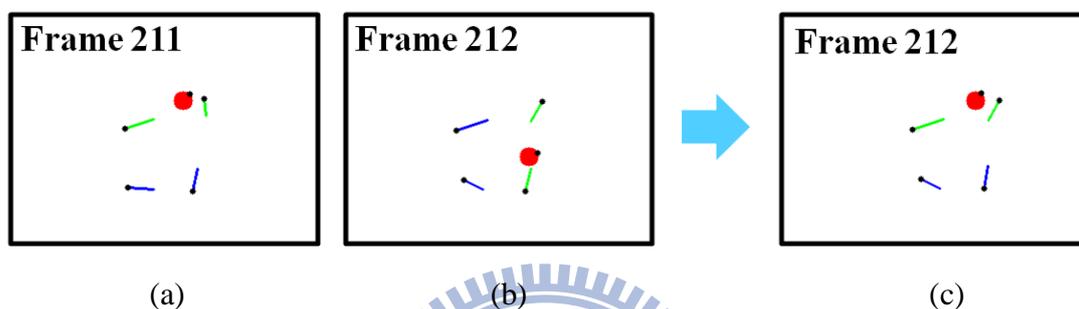


圖 2-11 沿用前一張頭部位置改正頭部定位之範例：(a) 前一張影像的頭、手、腳定位結果，(b) [11]之本張影像定位結果，(c) 本論文之本張影像定位結果。

在沒有找到頭部的情況下，首先計算各肢體與前一張頭部之最近距離，若距離小於門檻值，則將該最近肢體改判成頭部，其餘肢體將重新定位，如圖 2-12 所示。若距離大於門檻值，則沿用前一張的頭部位置新增頭部定位，並重新定位其餘肢體，如圖 2-13 所示。運用頭部位移合理性判斷之流程圖請參考圖 2-14，其中 TH_{head} 和 TH_{limb} 為距離門檻值，經由實驗設定 TH_{head} 值為 25， TH_{limb} 值為 20。

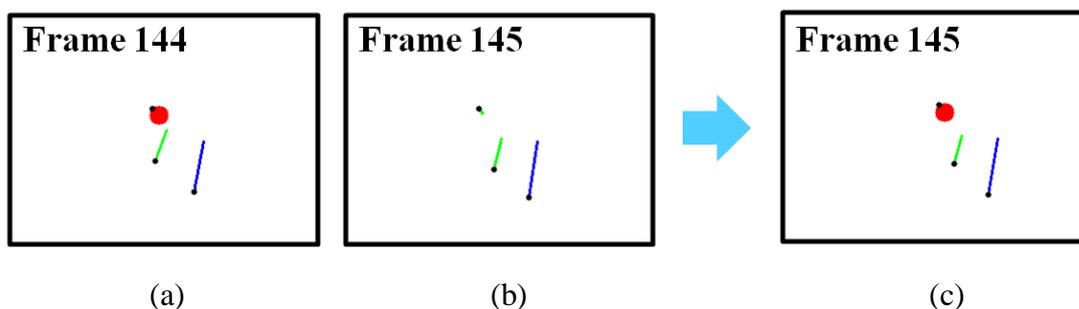


圖 2-12 利用最近肢體新增頭部定位之範例：(a) 前一張影像的頭、手、腳定位結果，(b) [11]之本張影像定位結果，(c) 本論文之本張影像定位結果。

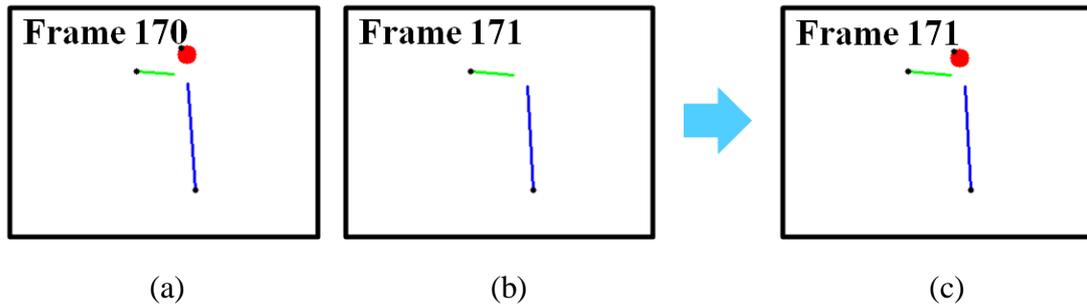


圖 2-13 沿用前一張頭部位置新增頭部定位之範例：(a) 前一張影像的頭、手、腳定位結果，(b) [11]之本張影像定位結果，(c) 本論文之本張影像定位結果。

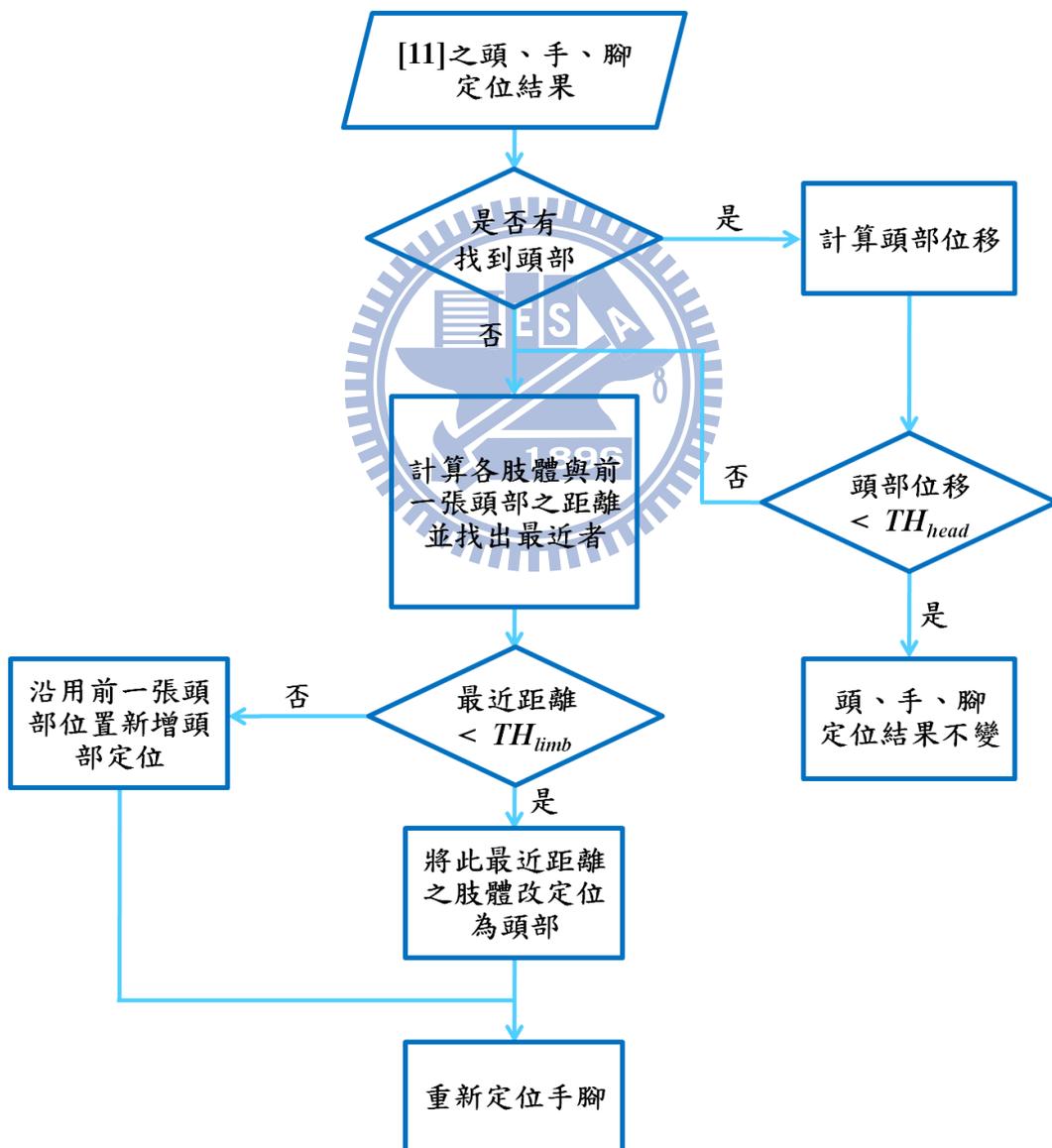


圖 2-14 運用頭部位移之合理性改善[11]的頭、手、腳定位結果之流程圖。

2.2.3 實驗結果與討論

本節將顯示兩段影片之前景圖、[11]之識別結果與本論文之識別結果，並比較兩者的識別正確率。

圖 2-15 是影片內容為一人物走動然後坐下的姿態識別結果，其中紅色代表頭，綠色代表手，藍色代表腳。表 2-1 為此段影片姿態識別結果之正確率比較，其中頭、手、腳標記正確率是指有多少個該部位識別正確/總共多少個該部位，整個影像正確率是指多少張影像中頭、手、腳部位都識別正確/影像總張數。

由於此段影片的人物動作較為單純，加上影像的前景分割結果極佳，所以[11]之姿態識別正確率不錯。但由表 2-1 看出本論文可將[11]之識別正確率進一步向上提升。

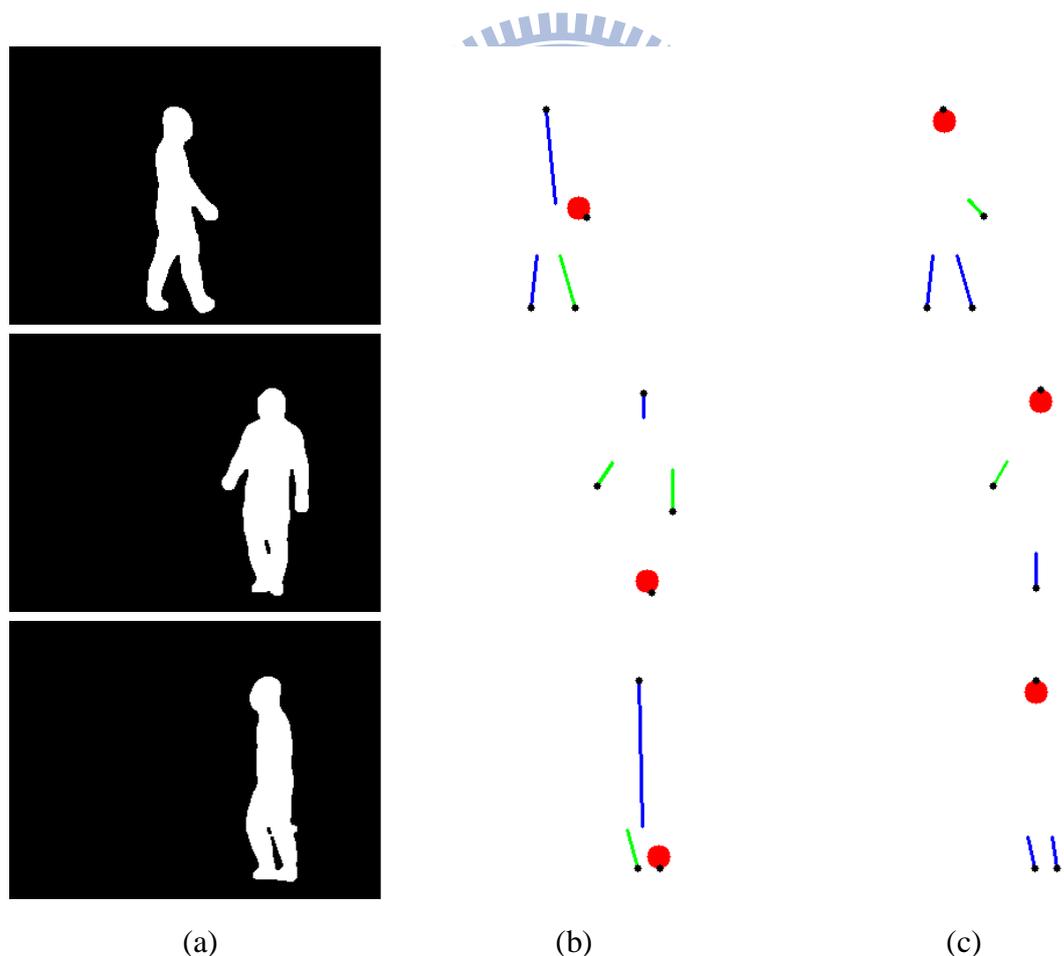


圖 2-15 人物走動然後坐下的影片姿態識別結果：(a) 前景圖，(b) [11]之識別結果，(c) 本論文之識別結果。

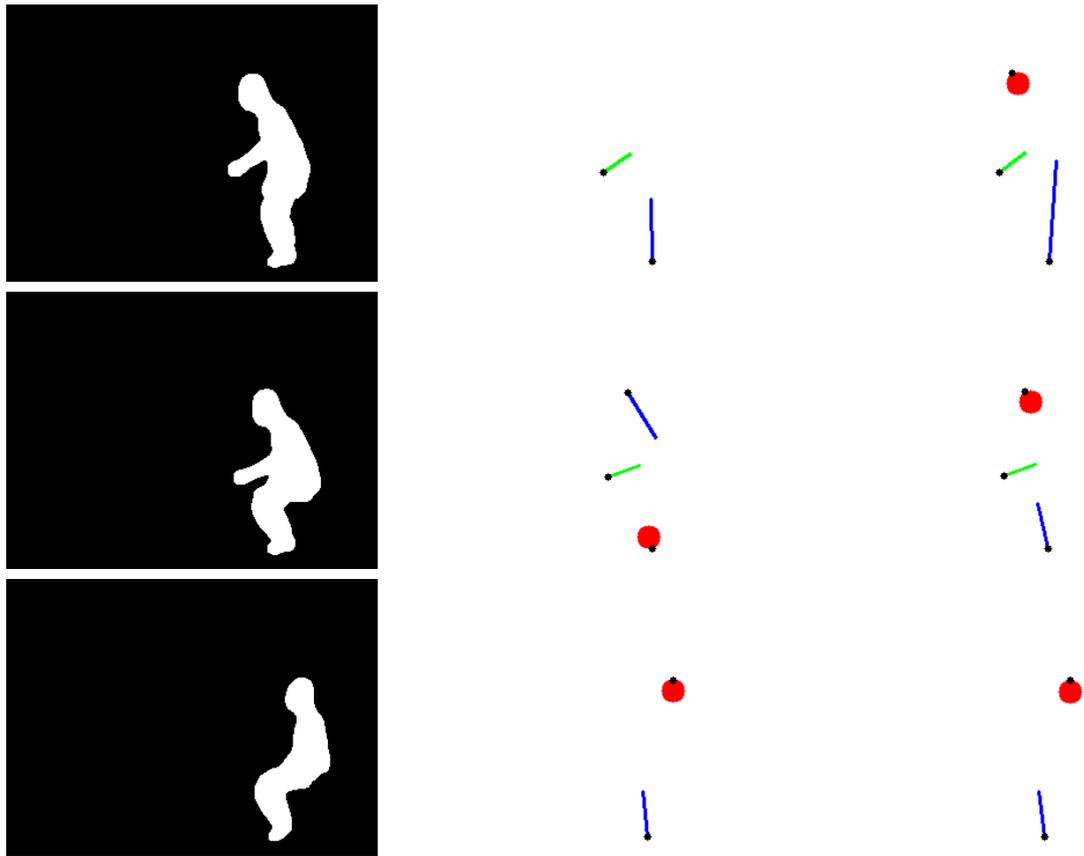


圖 2-15 (續)

表 2-1 人物走動然後坐下的影片姿態識別正確率之比較。

	[11]識別結果	本論文識別結果
頭部標記正確率	243 / 255 = 95.29%	261 / 261 = 100%
手標記正確率	217 / 223 = 97.31%	223 / 223 = 100%
腳標記正確率	389 / 397 = 97.98%	397 / 397 = 100%
整個影像正確率	248 / 261 = 95.01%	261 / 261 = 100 %

圖 2-16 是影片內容為一人物彎腰、蹲下及躺下的姿態識別結果，其中紅色代表頭，綠色代表手，藍色代表腳。表 2-2 為此段影片姿態識別結果之正確率比較，計算方式同表 2-1。由於此段影片的人物動作較為複雜，比較表 2-2 中[11]和本論文之正確率，可發現本論文加入時間軸資訊進行判斷，可使正確率顯著提昇。由於本論文以頭部判斷為主，各肢體位置係沿用[11]之判斷結果進行重新分配，以致少數肢體位置判斷錯誤並未改正，如圖 2-17 所示。

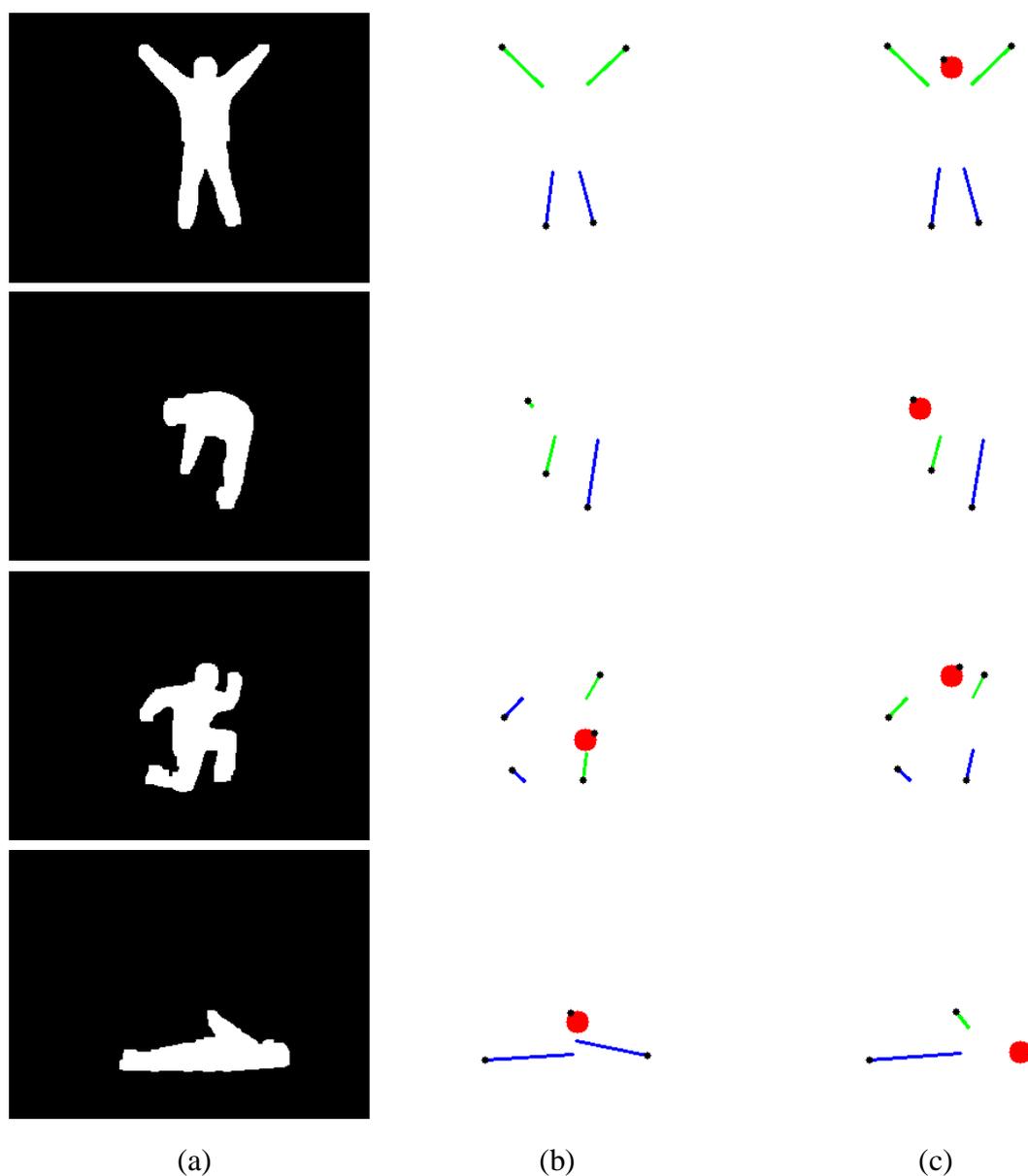


圖 2-16 人物彎腰、蹲下及躺下的影片姿態識別結果：(a) 前景圖，(b) [11]之識別結果，(c) 本論文之識別結果。

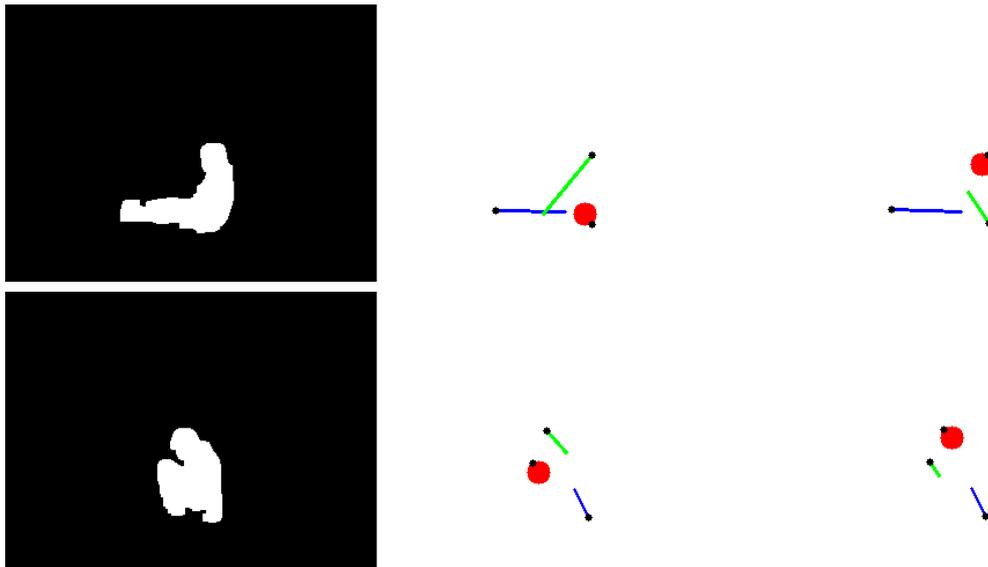


圖 2-17 少數肢體位置判斷錯誤之例子。

表 2-2 人物彎腰、蹲下及躺下的影片姿態識別正確率之比較。

	[11]識別結果	本論文識別結果
頭部標記正確率	$337 / 388 = 86.86\%$	$429 / 437 = 98.17\%$
手標記正確率	$395 / 441 = 89.57\%$	$425 / 441 = 96.17\%$
腳標記正確率	$618 / 669 = 92.38\%$	$657 / 669 = 98.21\%$
整個影像正確率	$359 / 437 = 82.15\%$	$412 / 437 = 94.28\%$

第三章 階層式人物動作分析

由於考慮到各個使用者對被觀測者的關心層面各不相同，除了 2.2 節中介紹的人體姿態識別之外，本章利用階層式概念逐步擷取人體特徵進行分析，作為發展階層式影像日誌系統的基礎。首先在 3.1 節中我們將利用 2.1 節所得到的人體區域計算人體質心位置，以進行活動狀態分析；接著在 3.2 節中利用人體外接矩形進行高度變化分析；結合 3.1 節和 3.2 節的分析結果，在 3.3 節中進行人物動作分類與行為分析。階層式人物動作分析之詳細流程圖請參考圖 3-1。



圖 3-1 階層式人物動作分析流程圖。

3.1 活動狀態分析

為了對人物進行活動狀態分析，首先我們參考[26]的作法，利用式子(3-1)找出影像中人體質心的座標位置，如圖 3-2 所示。

$$x_c(t) = \frac{1}{N} \sum_{i=1}^N x_i, \quad y_c(t) = \frac{1}{N} \sum_{i=1}^N y_i \quad (3-1)$$

其中 $(x_c(t), y_c(t))$ 是第 t 張影像之人體區域內所有像素的平均座標，也就是人體質心位置， (x_i, y_i) 是人體區域內的像素座標，而 N 表示整個人體區域所涵蓋的像素總數。

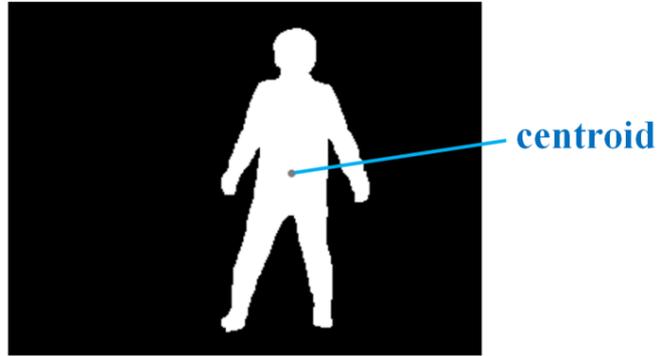


圖 3-2 人體質心示意圖。

在找到每一張影像的人體質心位置後，我們使用式子(3-2)的歐基里德距離公式，計算前後張影像的人體質心位移。由於使用人體質心位移進行人物活動狀態判斷時，易受前景區域分割好壞的影響。為了得到較穩定的判斷結果，我們利用式子(3-3)對人體質心位移進行權重累計，藉以得到較可靠的判斷數據。

$$d_c(t) = \sqrt{(x_c(t) - x_c(t-1))^2 + (y_c(t) - y_c(t-1))^2} \quad (3-2)$$

$$d_{ac}(t) = (1-w) \times d_{ac}(t-1) + w \times d_c(t) \quad (3-3)$$

其中 $d_c(t)$ 是第 t 張與第 $t-1$ 張的人體質心位移， $d_{ac}(t)$ 是第 1 張到第 t 張影像的累計人體質心位移，而 w 是累計的權重值，經由實驗設定其值為 0.5。

根據累計人體質心位移值，我們可以判斷該張影像中的人物是處於活動狀態或是靜止狀態：

$$\begin{cases} \text{Moving: } d_{ac}(t) > TH_{act} \\ \text{Stationary: } d_{ac}(t) \leq TH_{act} \end{cases} \quad (3-4)$$

其中 TH_{act} 是活動量門檻值，當累計人體質心位移值大於門檻值判斷為活動狀態，小於等於門檻值則判斷為靜止狀態，經由實驗設定 TH_{act} 值為 2.5。

3.2 高度變化分析

為了分析人物的高度變化，我們利用人體外接矩形（bounding box）找出影像中人體區域的長度，如圖 3-3 所示。在日常生活中，人物的高度大致可以區分為較具代表性的站、坐、躺三種，如圖 3-4 我們將人物高度分成三種區域，其中 TH_{high} 和 TH_{low} 是人物影像長度門檻值，大於 TH_{high} 為較高區域，小於 TH_{low} 為較低區域，介於兩者之間為中間區域，經由實驗設定 TH_{high} 值為 140， TH_{low} 值為 50。

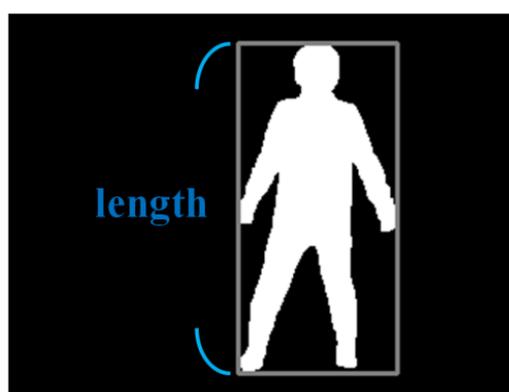


圖 3-3 利用人體外接矩形找出人物影像長度。

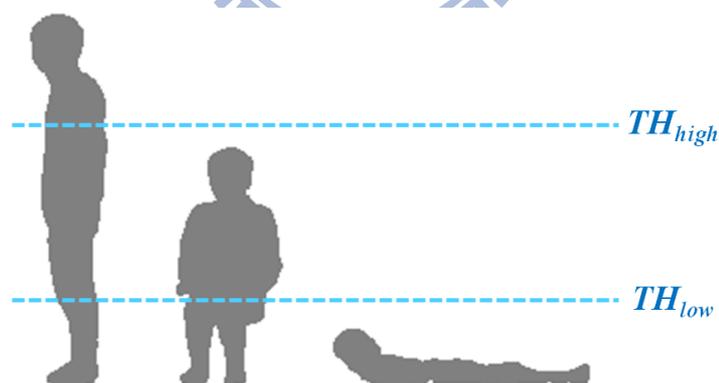


圖 3-4 將人物影像長度分成三個區域。

3.3 動作分類與行為分析

本節針對日常生活中，較常見的基本動作如站立、走路、坐下、躺下等，進行判斷與分類。另外，由於跌倒是家中較常見、也是危險度較高的異常行為，所以在 3.3.2 節中我們將特別針對跌倒之特性，設計一套快速且有效的偵測方法。

3.3.1 動作分類

我們結合 3.1 節和 3.2 節得到的人物活動狀態與高度變化資訊，對人物進行動作分類。首先利用人物影像長度進行判斷：

$$\begin{cases} \text{Lying: } length < TH_{low} \\ \text{Sitting: } TH_{low} \leq length \leq TH_{high} \\ \text{Otherwise: } length > TH_{high} \end{cases} \quad (3-5)$$

依據 3.2 節所區分的三種高度區域，將高度屬於較低區域的人物動作判斷為躺著，而高度屬於中間區域的人物動作判斷為坐著。接著利用人物活動狀態判斷高度屬於較高區域者：

$$\begin{cases} \text{Standing: } length > TH_{high} \text{ and } d_{ac}(t) \leq TH_{act} \\ \text{Walking: } length > TH_{high} \text{ and } d_{ac}(t) > TH_{act} \end{cases} \quad (3-6)$$

利用 3.1 節求得的人物活動狀態，將狀態為活動中的人物動作判斷為走路，而狀態為靜止中的人物動作判斷為站立。

3.3.2 行為分析

為了能找出跌倒行為之特性，我們觀察圖 3-5 正常躺下與圖 3-6 跌倒情況下的人物高度變化。由圖中的藍色虛線可明顯看出正常躺下時，人物高度變化較為和緩，而跌倒情況下人物高度變化較為劇烈。因此，我們可以依據人物高度有無劇烈變化判別是否有跌倒情形發生。

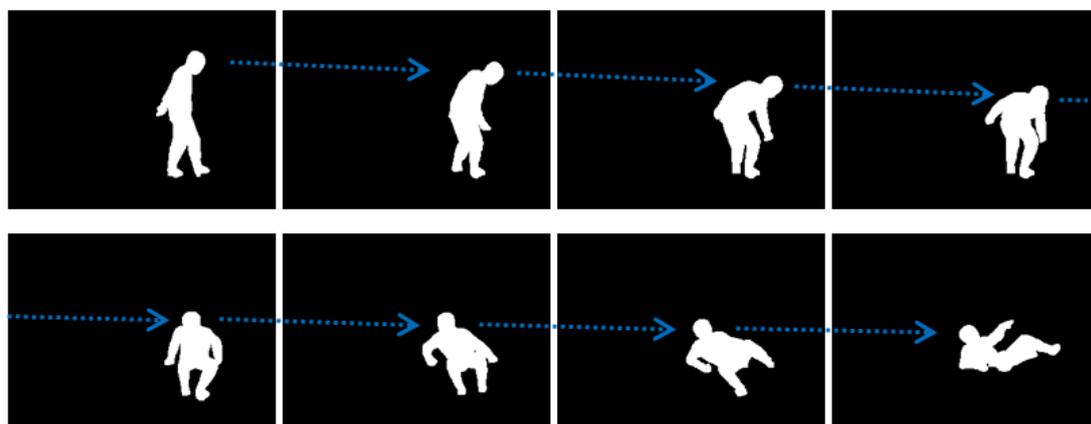


圖 3-5 正常躺下之人物高度變化。

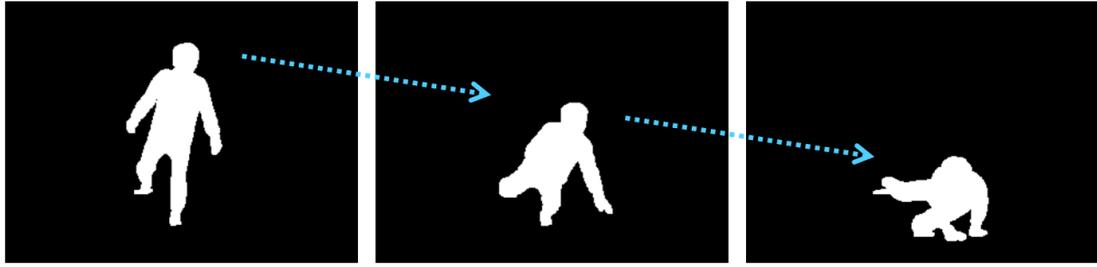


圖 3-6 跌倒情況下之人物高度變化。

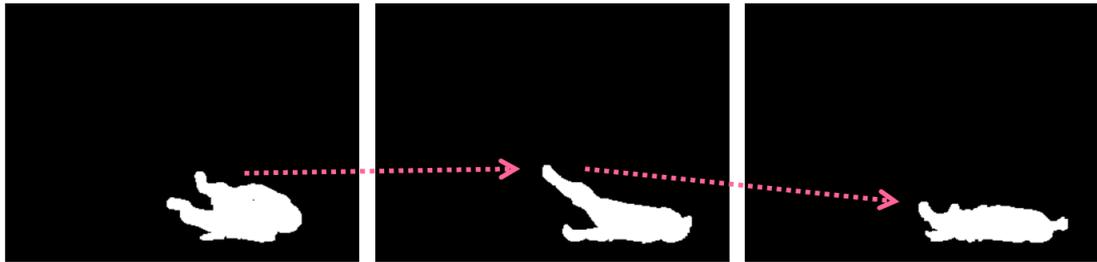


圖 3-7 跌倒之人物高度起伏例子。

此外，在跌倒時人物高度變化由高到低的過程中，可能不會很平順地下降，如圖 3-7 所示，當跌倒落地時，腳可能會有因反彈力而往上抬起的過渡情形。為了因應這種情況，我們依據腳受反彈力抬起的可能高度範圍，將 3.2 節所區分的高度區域由三種增加成四種。如圖 3-8 所示，藉由粉紅色人影示意跌倒腳抬起的可能高度區域，我們將介於 TH_{high} 和 TH_{low} 之間的區域利用 TH_{mid} 門檻值再細分成兩個區域， TH_{high} 和 TH_{mid} 之間是非反彈區域，而 TH_{mid} 和 TH_{low} 之間是反彈區域，經由實驗設定 TH_{mid} 值為 100。

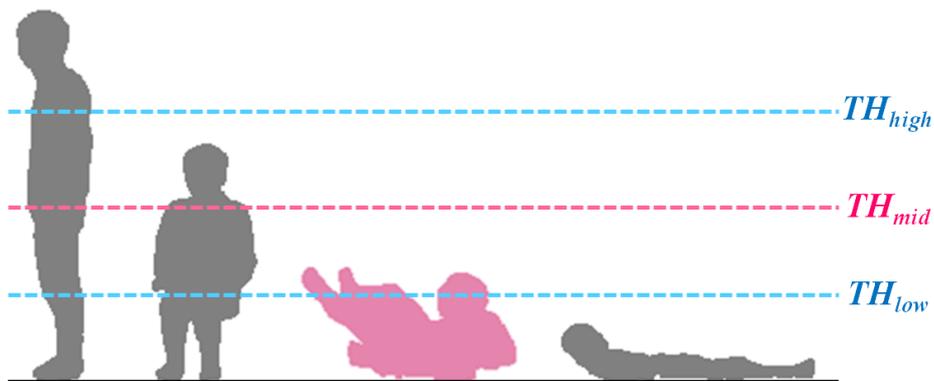


圖 3-8 將人物影像長度分成四個區域。

在將人物高度分成四個區域後，我們即可利用人物高度變化進行跌倒判斷。圖 3-9 為跌倒判斷之流程圖，首先我們觀察如果高度有由高往低的情形，則判斷非反彈區與反彈區的高度變化，如果變化快速則判斷有跌倒情形發生。其中 TH_{n1} 和 TH_{n2} 是影像張數門檻值，經由實驗設定 TH_{n1} 值為 2，而 TH_{n2} 值為 5。

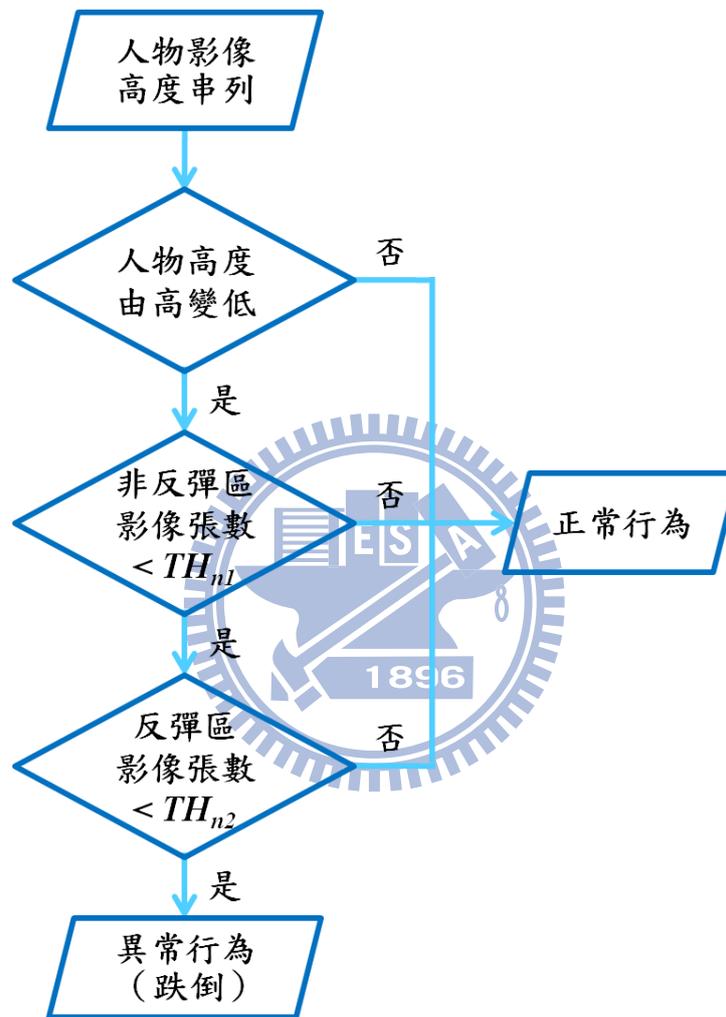


圖 3-9 由人物高度變化偵測跌倒之流程圖。

第四章 階層式影像日誌系統設計

本章將利用第二章與第三章中所得人物活動狀態、高度變化、動作分類、行為分析與姿態識別等技術，設計並建構一個階層式影像日誌系統。首先，在 4.1 節中將進行系統需求分析，接著在 4.2 中根據系統需求進行系統功能規劃，最後於 4.3 節中詳細介紹系統介面與使用說明。

4.1 系統需求分析

本階層式影像日誌系統為了達到必須長時間監看之目的，必須與影像監視儲存裝置結合，取得附有時間標記的視訊影片檔案。使用者可以選擇與檢索所關心時段的視訊影片檔案進行瀏覽。

由於各個使用者對被觀測者日常生活所關心的主題不盡相同；有的人關心被觀測者是否有跌倒危險，有的人關心被觀測者活動量是否足夠，有的人關心被觀測者行為是否反常等等不一而足；此外當然每個人都關心隱私保護。本階層式、具隱私保護的影像日誌系統，係利用第二章與第三章中所得人物活動狀態、高度變化、動作分類、行為分析與姿態識別等電腦視覺技術，所設計建構而成。

4.2 系統功能規劃

依據 4.1 節所分析的系統需求，規劃出本系統之功能架構，如圖 4-1 所示。本系統主要功能有：(1) 選擇與輸入想觀看時段的視訊檔案，(2) 利用階層方式，顯示由第二章與第三章分析技術得到的判斷結果，階層架構由上而下分別為人物活動狀態、高度狀態、動作狀態與姿態識別影像，(3) 能縮放顯示範圍及選取欲進一步觀看下一層狀態的時段，(4) 觀看下層狀態時可顯示上層於該時段的對應位置，(5) 針對影片中人物活動狀態與動作狀態進行檢索。各個主要功能的詳細使用說明將會在 4.3 節中介紹。

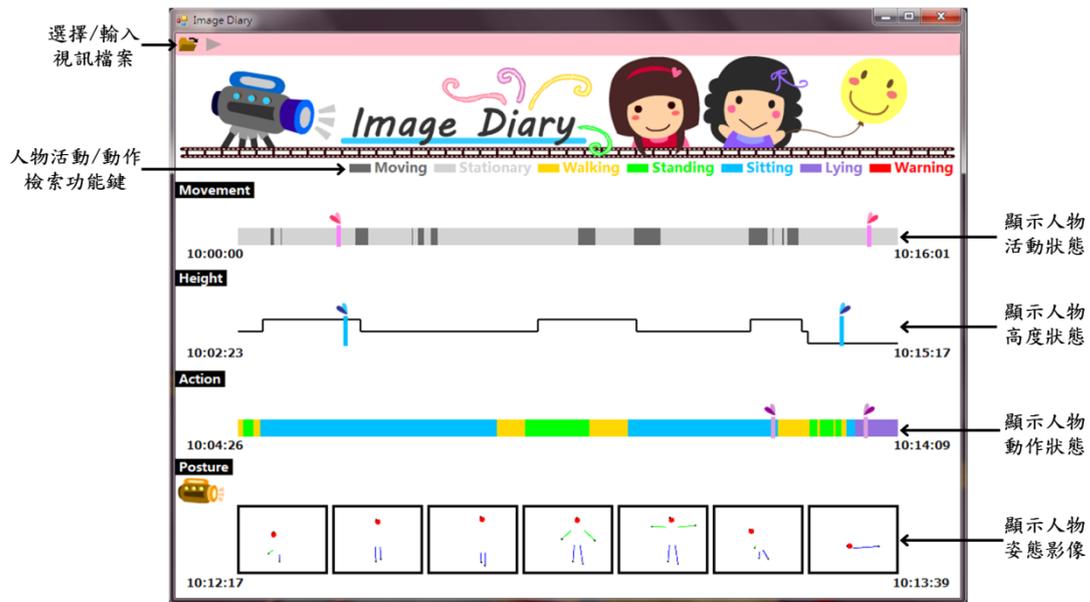


圖 4-1 系統功能架構圖。

4.3 系統介面使用說明

本節將依據圖 4-2 所示的完整系統介面操作流程進行本階層式影像日誌系統使用說明。

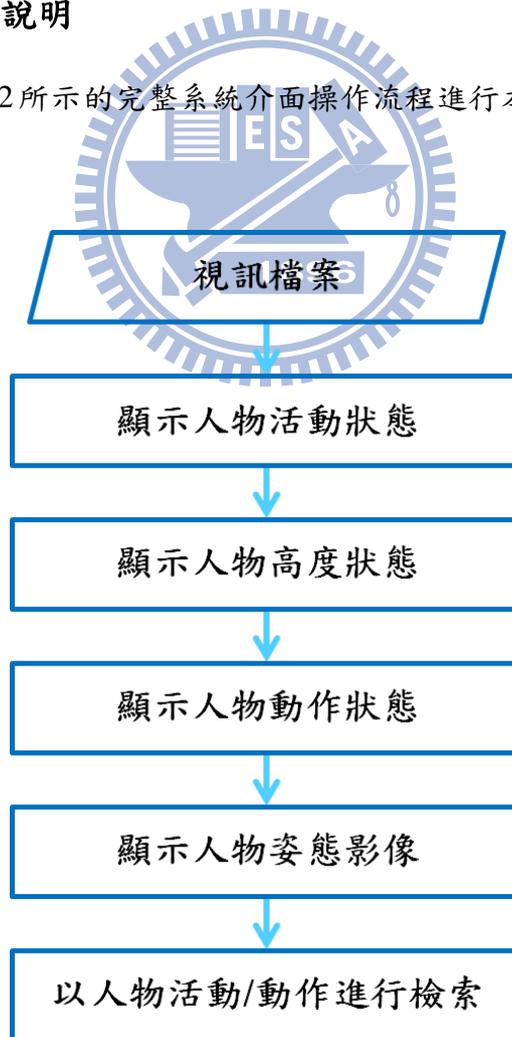


圖 4-2 完整系統介面操作流程圖。

(1) 選擇與輸入視訊檔案

首先使用者可選擇與輸入想觀看時段的視訊檔案，如圖 4-3 所示按下圖中編號 1 的功能鍵時會跳出視訊檔案選取視窗。在選定檔案並輸入後，系統將開始進行狀態分析，待分析完畢時編號 2 的功能鍵會由灰色轉為綠色。

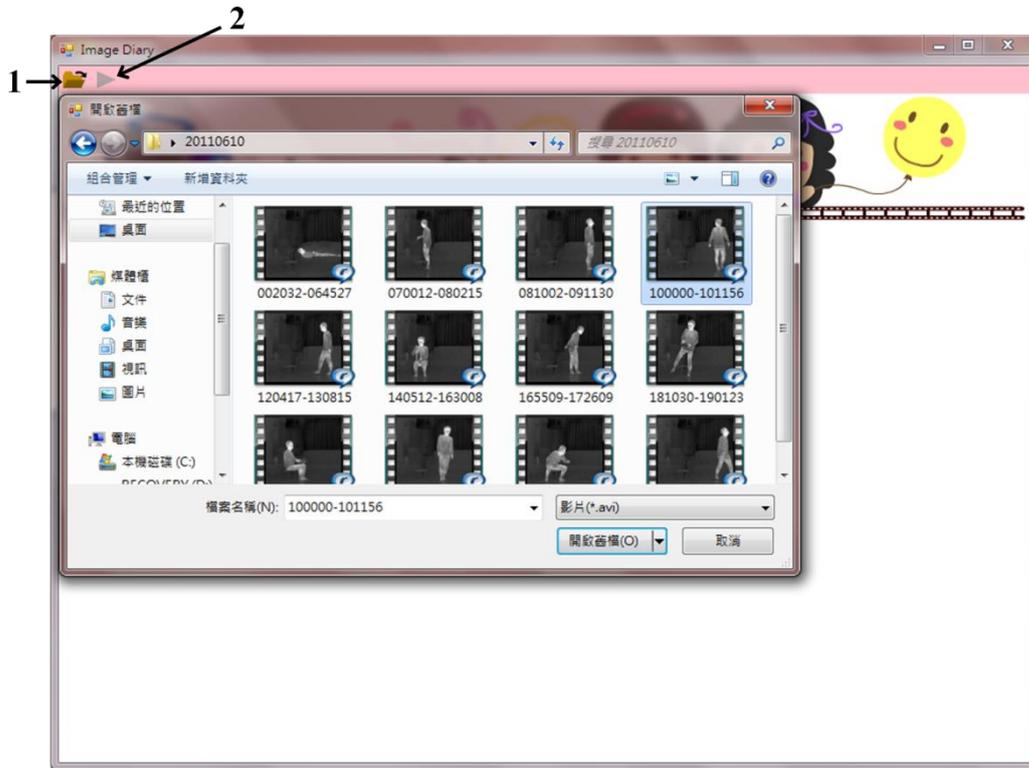


圖 4-3 選擇與輸入視訊檔案。

(2) 顯示人物活動狀態

在系統分析完畢後，我們即可進行階層式人物狀態觀看。首先按下圖 4-3 中已由灰色轉為綠色的編號 2 功能鍵，即會顯示階層式結構中的第一層「人物活動狀態」，如圖 4-4 中編號 3 所示，其中深灰色表示人物活動量大屬於活動狀態，淺灰色表示人物活動量小屬於靜止狀態。在圖 4-4 中的編號 4 與編號 5 顯示此視訊檔案的起始時間與終止時間，而滑鼠在活動狀態列上移動時，會即時顯示該位置的時間如編號 6 所示。當視訊檔案包含的時段較長時，會有資料過於密集的情形發生，這時就可以利用縮放功能，將需要仔細觀看的時段放大，以便能更清楚了解人物活動狀態。



圖 4-4 顯示人物活動狀態層。



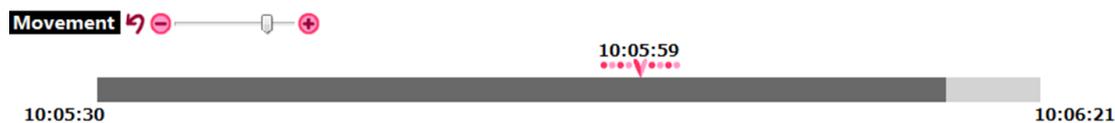
圖 4-5 選取人物活動狀態層中要更仔細觀看的時段。

縮放功能使用說明如下：首先按住滑鼠右鍵並拖曳滑鼠選取要更仔細觀看的時段，圖 4-5 中的編號 7 粉紅色條棒表示按下滑鼠右鍵時所選取的起始時間，編號 8 為經滑鼠拖曳後選取的終止時間。當選定要縮放之時段後放開滑鼠右鍵，即

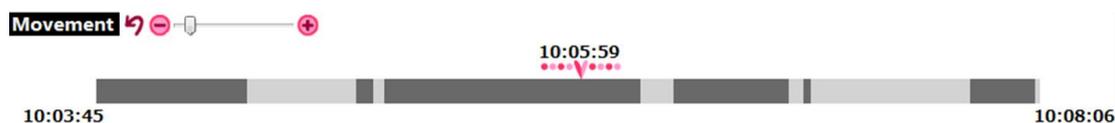
會出現縮放功能鍵如圖 4-6(a)，並將人物活動狀態列改為顯示所選取的時段。圖 4-6(a)中的編號 10 縮放功能鍵是以所選取時段的中心點為基準進行固定差值縮放，圖 4-6(b)與(c)分別為圖 4-6(a)放大與縮小後之結果。而圖 4-6(a)-(c)中的時間點（10:05:59）為便於比較縮放結果所舉的例子。在進行縮放功能後，如果覺得所選取的時段不理想或還想更仔細觀看其他時段，可使用圖 4-6(a)中的編號 9 功能鍵返回圖 4-4 的畫面。



(a)



(b)



(c)

圖 4-6 人物活動狀態層之縮放功能：(a) 欲縮放的時段，(b) 經等差放大三次後的結果，(c) 經等差縮小四次後的結果。

(3) 顯示人物高度狀態

在取得人物活動狀態後，我們可以進一步觀看階層式結構中的第二層「人物高度狀態」。首先在人物活動狀態層上按下滑鼠左鍵並拖曳滑鼠，選取要觀看人物高度狀態的時段，如圖 4-7 所示，編號 11 粉紅色條棒為按下滑鼠左鍵時所選取欲展開的高度狀態層起始時間，編號 12 為經滑鼠拖曳後選取的終止時間。當滑鼠左鍵放開時，已顯示的人物活動狀態層下，會出現經由滑鼠左鍵所選取時段的人物高度狀態，如圖 4-8 中編號 13 所示，以三種不同高度的線段表示人物的高度狀態。圖 4-8 中的編號 14 虛線與編號 15 虛線是為了輔助說明第二層顯示時段是由第一層所選取，實際操作時是不會出現的。



圖 4-7 選取要向下展開的高度狀態層時段。

為了使各階層彼此之間可以互相對照，當滑鼠在高度狀態列上移動時，除了會顯示該層位置的時間，同時也會將上一層中所對應的時間位置標示出來，如圖 4-8 中編號 16 虛線所示，人物高度狀態層上的時間點（10:04:23）係對應到人物活動狀態層上的紅色旗幟處。

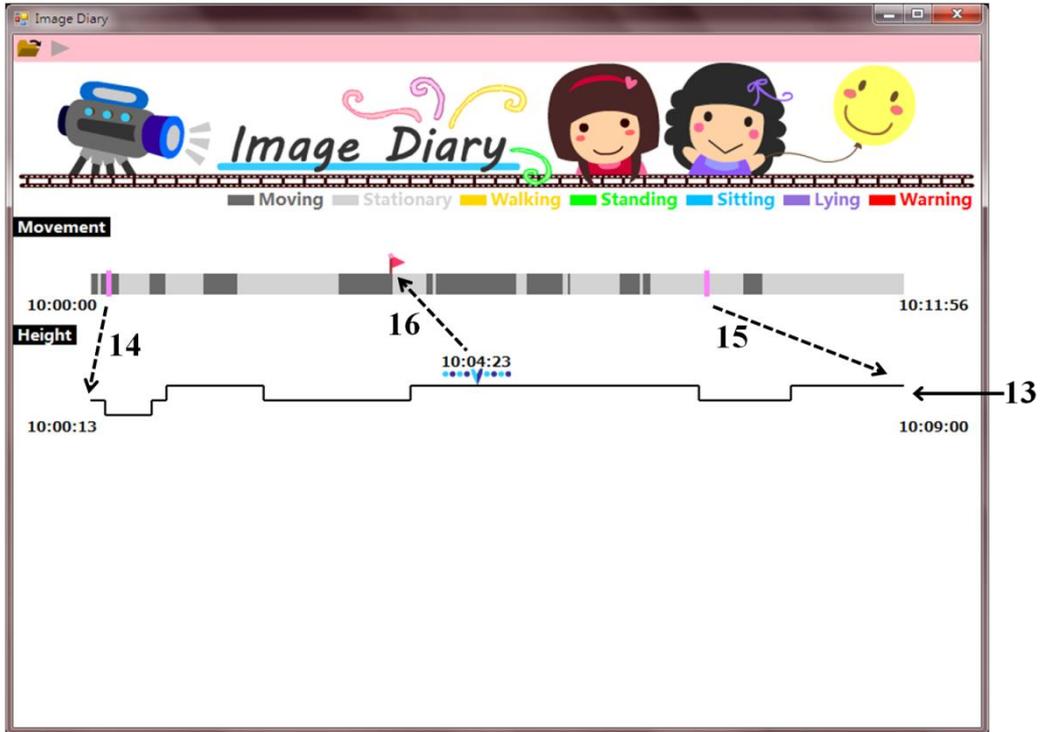


圖 4-8 顯示人物高度狀態層。

在人物高度狀態層中有提供與人物活動狀態層一樣的縮放功能鍵，如圖 4-9 所示，由於操作方式在 (2) 顯示人物活動狀態中已詳細說明，此處不再贅述。

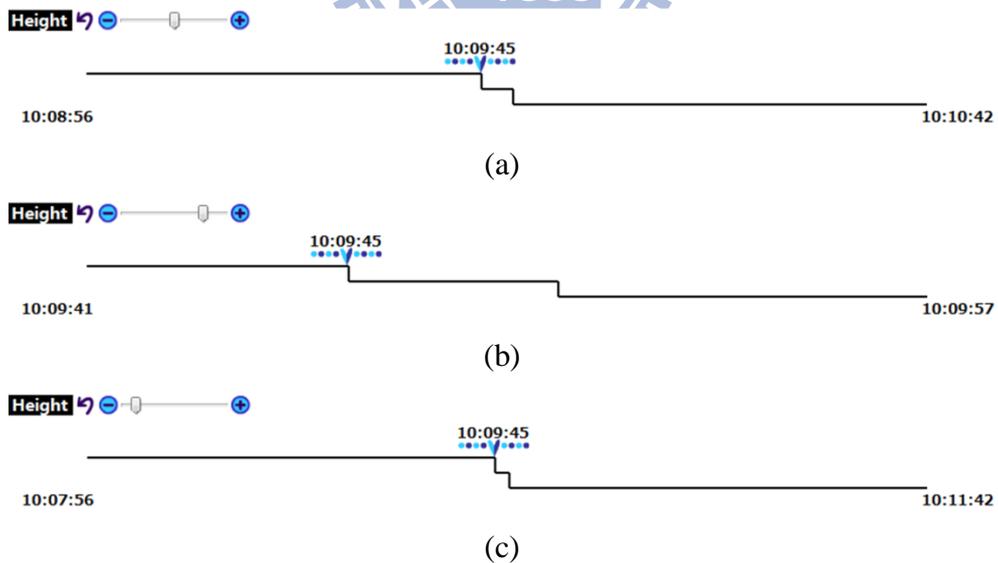


圖 4-9 人物高度狀態層之縮放功能：(a) 欲縮放的時段，(b) 經等差放大三次後的結果，(c) 經等差縮小四次後的結果。

(4) 顯示人物動作狀態

在知道人物活動與高度狀態後，我們可以再進一步觀看階層式結構中的第三層「人物動作狀態」。首先在人物高度狀態層上按下滑鼠左鍵並拖曳滑鼠，選取要觀看的人物動作狀態時段，如圖 4-10 所示，編號 17 藍色條棒標示欲展開的動作狀態層起始時間，編號 18 為欲展開的動作狀態層終止時間。當選好展開時段後，在人物高度狀態層下將會顯示所選取時段的人物動作狀態層，如圖 4-11 編號 19 所示，其中人物走路用黃色表示、站立用綠色表示、坐著用藍色表示、躺著用紫色表示，另外，在圖 4-11 中沒出現的跌倒警告用紅色表示。圖 4-11 中的編號 20 虛線與編號 21 虛線是為了輔助說明第三層顯示時段是由第二層所選取出來，實際操作時不會出現。

另外，為了使各階層之間可以互相對照，當滑鼠在動作狀態列上移動時，除了會顯示該層位置的時間，同時也會將前兩層中所對應的時間位置標示出來，如圖 4-11 中編號 22 虛線所示，人物動作狀態層上的時間點（10:01:59）係對應到人物高度狀態層上的藍色旗幟處以及人物活動狀態層上的紅色旗幟處。

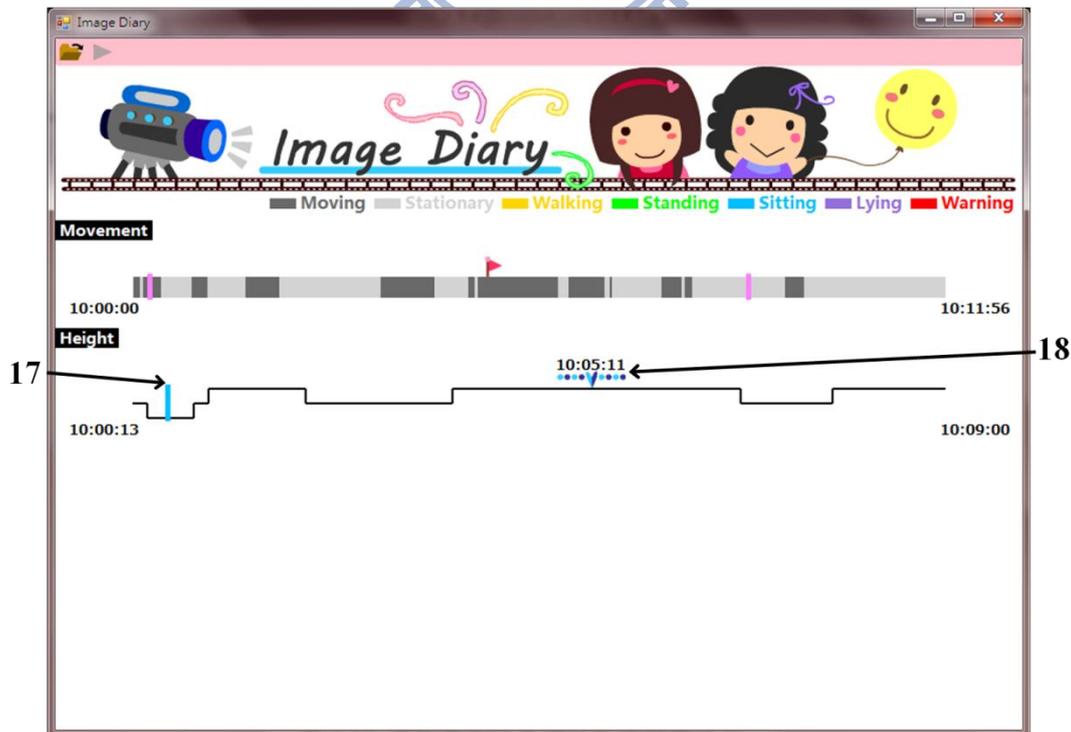


圖 4-10 選取要向下展開的動作狀態層時段。

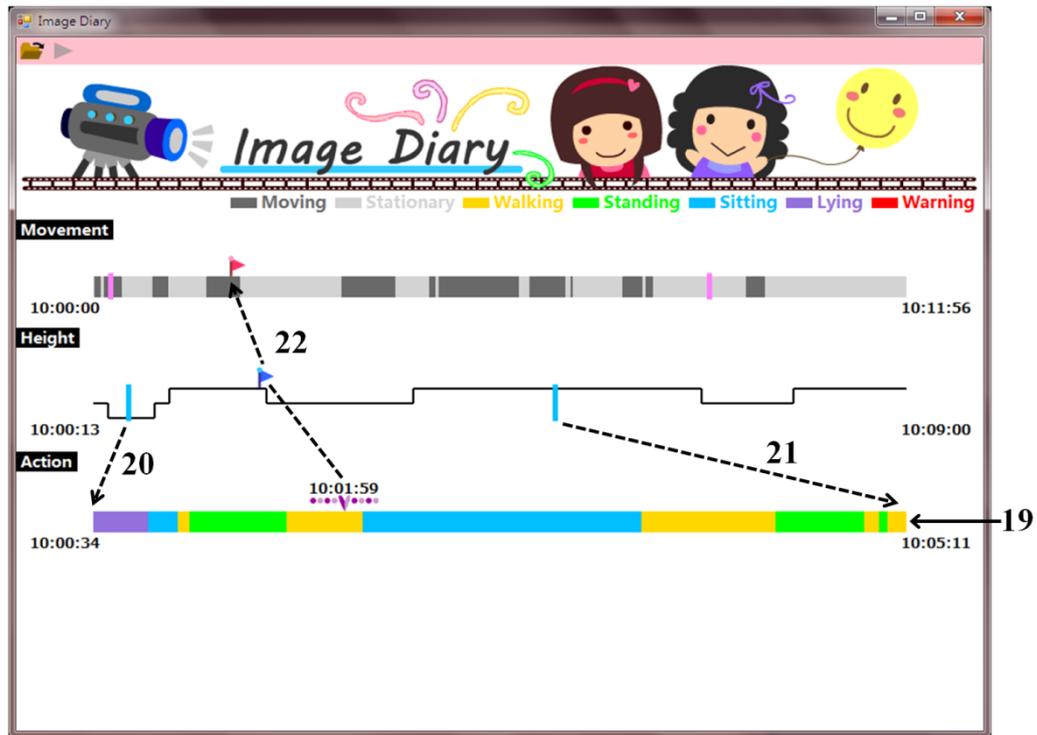


圖 4-11 顯示人物動作狀態層。

在人物動作狀態層中也有提供與人物活動、高度狀態層一樣的縮放功能鍵，如圖 4-12 所示。由於操作方式在 (2) 顯示人物活動狀態中已詳細說明，此處不再贅述。

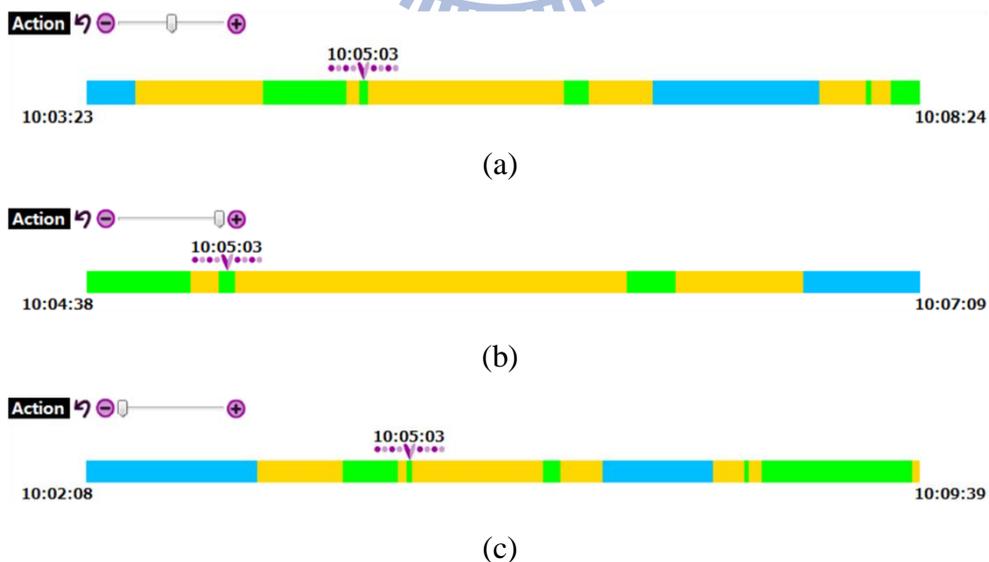


圖 4-12 人物動作狀態層之縮放功能：(a) 欲縮放的時段，(b) 經等差放大五次後的結果，(c) 經等差縮小五次後的結果。

(5) 顯示人物姿態影像

在得到人物活動、高度、動作狀態後，我們可以將進一步觀看階層式結構中的第四層「人物姿態影像」。首先在人物動作狀態層上按下滑鼠左鍵並拖曳，選取要觀看人物姿態影像的時段，如圖 4-13 所示，編號 23 紫色條棒標示欲展開的姿態影像起始時間，編號 24 為欲展開的姿態影像終止時間。當選好展開時段並放開滑鼠左鍵時，在人物動作狀態層下將會顯示選取時段的人物姿態影像，如圖 4-14 編號 25 所示，顯示起始時間與終止時間的姿態影像，並等間隔挑選中間五張姿態影像作為代表。圖 4-14 中的編號 26 虛線與編號 27 虛線是為了輔助說明第四層顯示時段是由第三層所選取出來，實際操作時不會出現。

為了使各階層之間可以互相對照，當滑鼠在姿態影像列上移動時，除了會顯示該層位置的時間，同時也會將前三層中所對應的時間位置標示出來，如圖 4-14 中編號 28 虛線所示，人物姿態影像層上的時間點 (10:04:02) 係對應到人物動作狀態層上的紫色旗幟處、人物高度狀態層上的藍色旗幟處、以及人物活動狀態層上的紅色旗幟處。

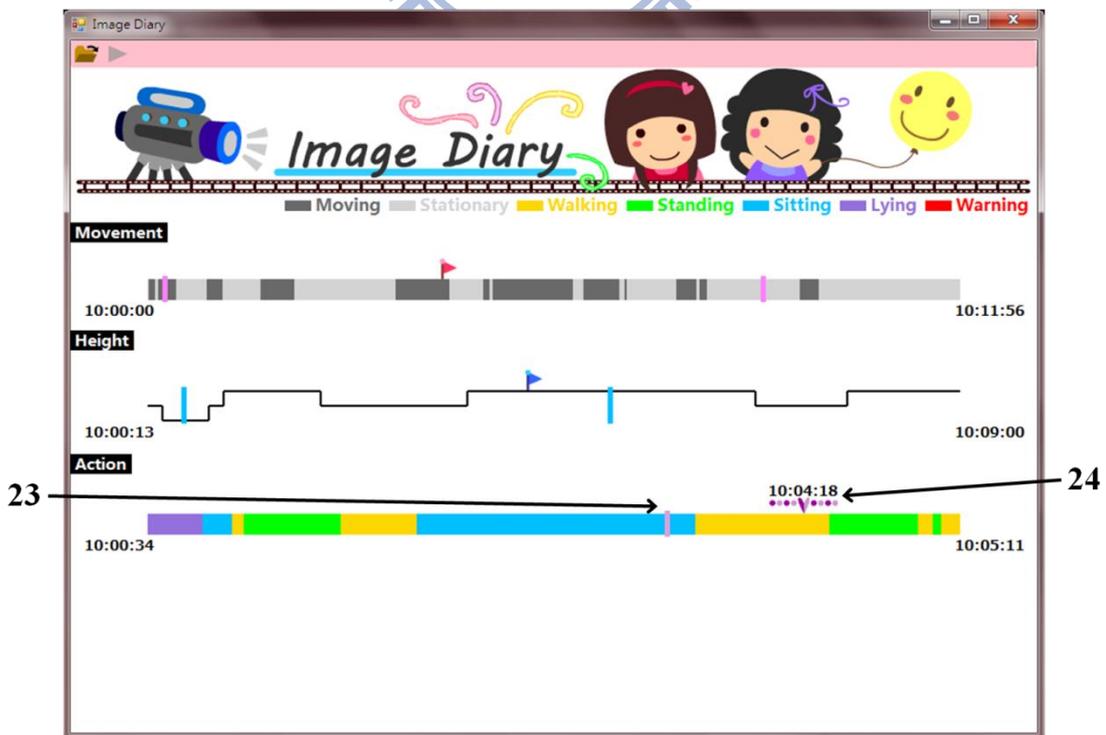


圖 4-13 選取要向下展開的姿態影像層時段。

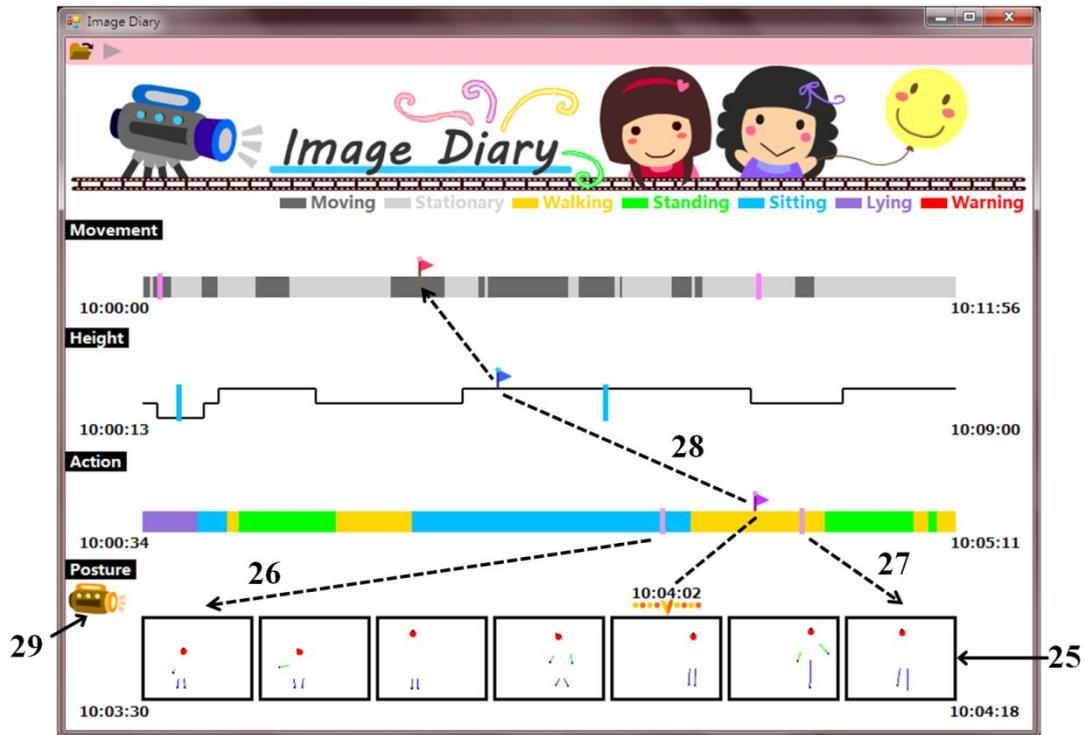


圖 4-14 顯示人物姿態影像層。

由於畫面僅挑選七張人物姿態影像顯示，如果想更詳細觀看該選取時段的所
 有人物姿態影像，可點選圖 4-14 編號 29 功能鍵跳出影片播放視窗。圖 4-15 為
 影片播放視窗，左下角顯示目前影像所屬的時段，位於右排的功能鍵由上而下分
 別為播放、暫停、停止、放大畫面、縮小畫面、加快播放速度、放慢播放速度，
 其中畫面大小共有五種供選擇，分別是：320×240、400×300、480×360、560×420、
 640×480，而播放速度共有五種供選擇，分別是：5 張/秒、10 張/秒、20 張/秒、
 50 張/秒、100 張/秒。

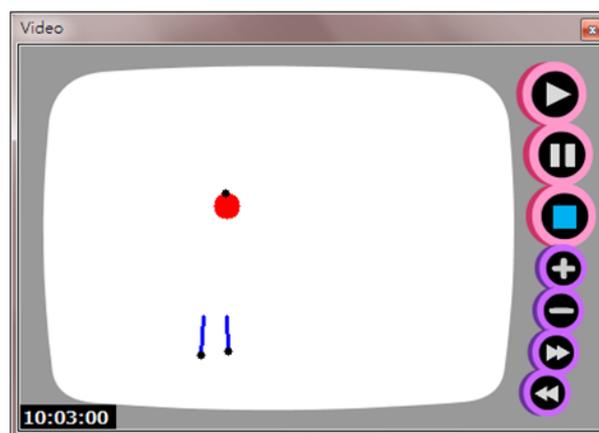


圖 4-15 人物姿態影像播放視窗。

(6) 以人物活動 / 動作進行檢索

為了能快速搜尋到有興趣的人物活動狀態與動作狀態，可點選圖 4-16 所示之人物活動 / 動作檢索功能鍵。其中編號 30 的 Moving、Stationary 為人物活動檢索功能鍵，而編號 31 的 Walking、Standing、Sitting、Lying、Warning 為人物動作檢索功能鍵。

各功能鍵之檢索結果視窗分別如圖 4-17 到圖 4-23 所示。在各個檢索結果視窗右上角會顯示該狀態發生的總數目，並依序顯示每一個項目的關鍵影像，以及各項目發生的起始時間與終止時間。



圖 4-16 顯示檢索功能鍵。

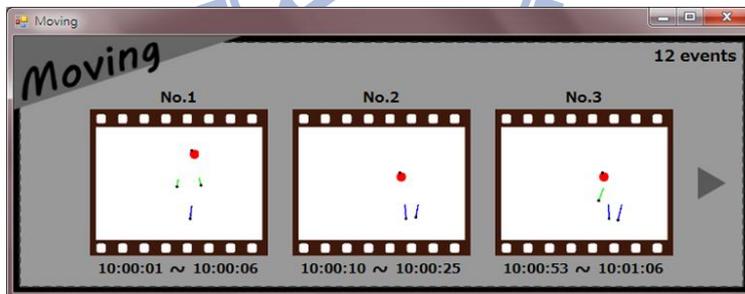


圖 4-17 人物活動中之檢索結果視窗。

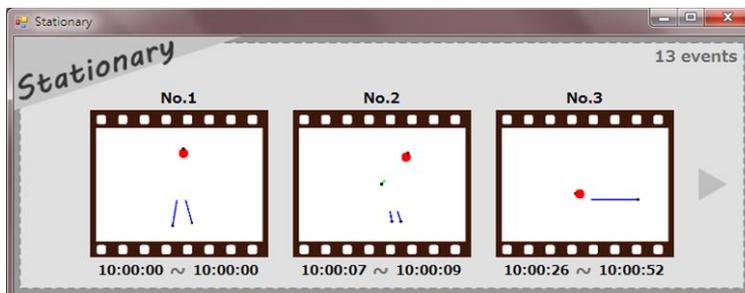


圖 4-18 人物靜止中之檢索結果視窗。

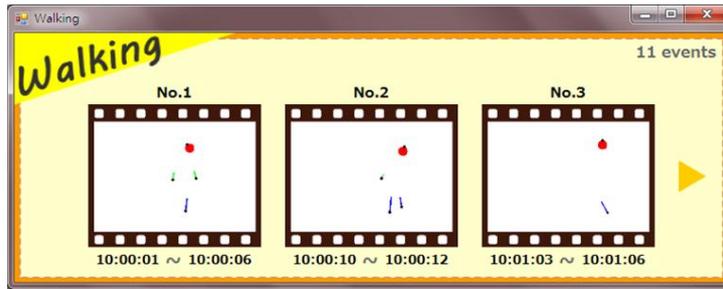


圖 4-19 人物走路動作之檢索結果視窗。

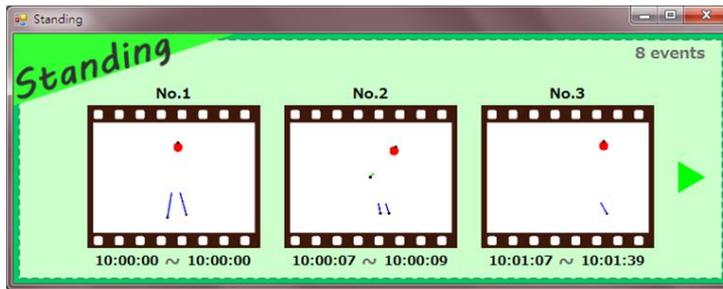


圖 4-20 人物站立動作之檢索結果視窗。

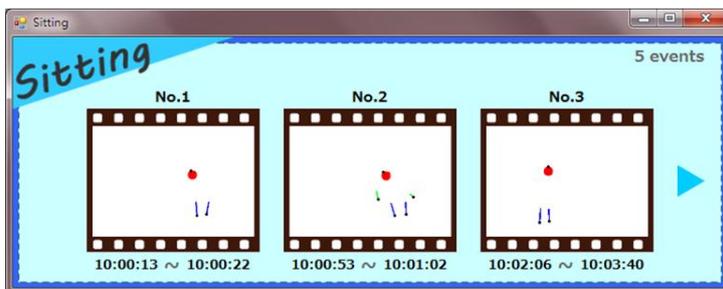


圖 4-21 人物坐著動作之檢索結果視窗。

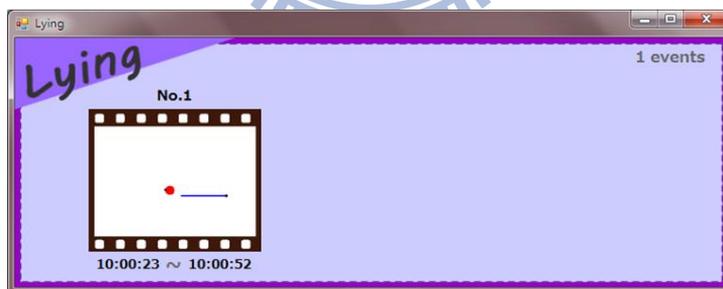


圖 4-22 人物躺著動作之檢索結果視窗。



圖 4-23 人物跌倒之檢索結果視窗。

第五章 實驗結果與討論

本章將針對本論文所發展的階層式人物動作分析系統，以兩段不同被觀測者的實驗影片進行實際驗證，其中實驗影片一長度約一小時，由走路、站立、坐著、躺著等正常動作所組成，而實驗影片二長度約二十分鐘，由前述四個正常動作和跌倒之異常行為所組成。在 5.1 節中將驗證活動狀態分析結果。在 5.2 節中將驗證高度狀態分析結果。在 5.3 節中將驗證動作狀態分析結果。

5.1 活動狀態分析結果

本論文人物活動狀態分析結果如圖 5-1 和圖 5-2 所示，圖中上排為本論文判斷結果，下排為實際狀態（人工判斷）。整體而言，相較於實際狀態，本論文對於活動狀態之判斷結果並沒有遺漏任何重要訊息，只在起始點有些許落差，此乃由於本論文的判斷因考慮系統穩定性，須累積足夠資訊作為依據。表 5-1 為綜整影片一與影片二之活動狀態分析的正確率，顯示各活動狀態判斷正確率都相當高，平均正確率可達 97%。



圖 5-1 影片一之活動狀態分析結果，上排為本論文判斷結果，下排為實際狀態。



圖 5-2 影片二之活動狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

表 5-1 活動狀態分析之正確率。

Est.	Truth	Moving	Stationary
	Moving		1605
Stationary		67	3130
Recognition Rate		96%	98%
Average Recognition Rate		97%	

5.2 高度狀態分析結果

本論文人物高度狀態分析結果如圖 5-3 和圖 5-4 所示，圖中上排為本論文判斷結果，與下排實際狀態（人工判斷）比較可清楚看出，不同高度下之判斷結果與實際狀態幾乎一致。表 5-2 為綜整影片一與影片二之高度狀態分析的正確率，顯示各高度狀態判斷正確率都極高，平均正確率可高達 99.3%。

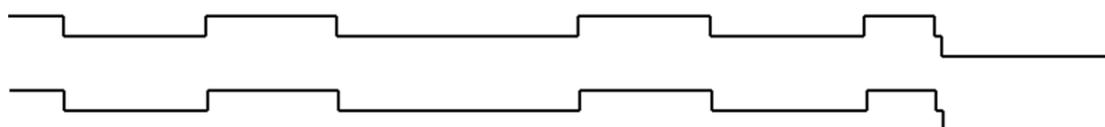


圖 5-3 影片一之高度狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

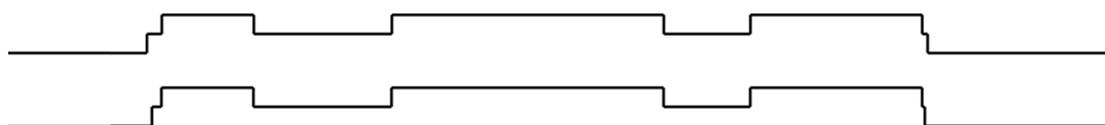


圖 5-4 影片二之高度狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

表 5-2 高度狀態分析之正確率。

Est.	Truth	High	Medium	Low
	High		2115	0
Medium		21	1537	12
Low		0	0	1181
Recognition Rate		99%	100%	99%
Average Recognition Rate		99.3%		

5.3 動作狀態分析結果

本論文人物動作狀態分析結果如圖 5-5 和圖 5-6 所示，圖中上排為本論文判斷結果，與下排實際狀態（人工判斷）比較可清楚看出，不同顏色的人物動作絕大多數都有確實判斷出來，只有在動作間轉換的過渡期會與實際狀態有些微落差，

如圖中下排灰色部分所示。表 5-3 為綜整影片一與影片二之動作狀態分析的正確率，顯示各動作狀態判斷正確率相當高，平均正確率可達 98.4%。



圖 5-5 影片一之動作狀態分析結果，上排為本論文判斷結果，下排為實際狀態。



圖 5-6 影片二之動作狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

表 5-3 動作狀態分析之正確率。

Truth \ Est.	Walking	Standing	Sitting	Lying	Falling down
Walking	977	45	0	0	0
Standing	30	1084	0	0	0
Sitting	0	0	1498	6	0
Lying	0	0	0	637	0
Falling down	0	0	0	0	589
Recognition Rate	97%	96%	100%	99%	100%
Average Recognition Rate	98.4%				

第六章 階層式人物動作分析之其他應用—籃球一對一鬥牛賽

階層式人物動作分析架構的應用範圍非常廣泛，本章將探討將其應用於球類活動分析上，並以籃球一對一鬥牛賽為例。我們探討的主題在於比賽進行狀態分析、球高度變化分析、以及持球方動作分類等。首先在 6.1 節中將影像中的人體區域與球區域偵測出來。在 6.2 節中進行階層式狀態分析，包含比賽進行狀態、球高度狀態、持球方動作狀態。最後於 6.3 節中討論籃球一對一鬥牛賽之階層式動作分析系統的設計規劃結果與驗證。

6.1 球類活動之前景區域偵測

在球類活動進行時，由於球會與人接觸，並與地板摩擦產生熱度，這是中紅外線攝影機可以運用的特性。以籃球一對一鬥牛賽為例，我們利用 2.1.1 節中的式子(2-1)對中紅外線影像做二值化，擷取中紅外線影像中屬於較高溫部分的人體區域與球區域。再利用 2.1.2 節中介紹的連通元件標記法以及型態學上的技巧濾除雜訊及填補前景影像中破碎之區域。

我們考慮到室外球場的外在因素較多，背景較為複雜，為了使問題單純化，所以選擇於室內球場拍攝。然而室內球場也有它的缺點，最明顯的是地板屬於光滑材質，易產生人物倒影，使得經二值化，並填補破碎區域與雜訊濾除後的前景區域結果會包含人物倒影。為了能將人物倒影移除，我們將攝影機架設於貼近地面的高度，使人物倒影與真實人物大小相似，這樣經由二值化所找到的人物倒影比較完整，方便於我們將含有倒影的人體區域依據人物影像長度對切成兩半，將下半部屬於倒影的區域移除。圖 6-1(a)為中紅外線原始影像，圖 6-1(b)為(a)經破碎區域填補與雜訊濾除之二值化前景影像。圖 6-1(c)為(b)將倒影移除後之結果。可以明顯看出，將含有倒影的人體區域依長度對切成兩半，並擷取上半部作為人體區域之做法是可行的。

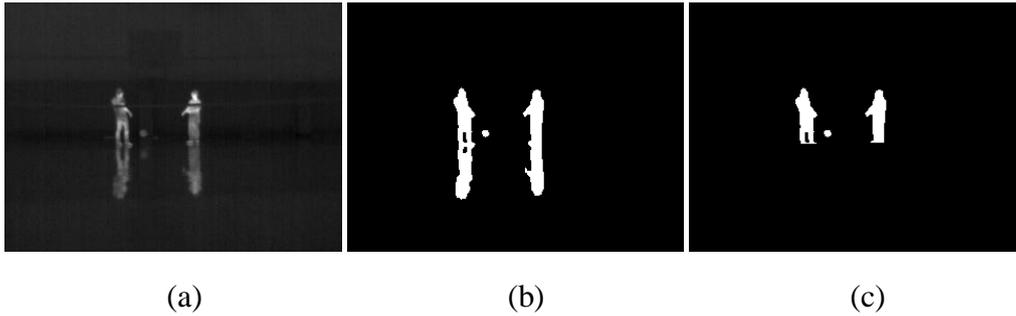


圖 6-1 中紅外線影像利用前景偵測找出之人體區域與球區域：(a) 原始影像，(b) 經破碎區域填補與雜訊濾除之二值化前景影像，(c) 倒影移除後之前景影像。

6.2 階層式狀態分析

本節將利用 6.1 節中得到的人體區域與球區域位置資訊進行階層式狀態分析，在 6.2.1 節中將利用球的水平移動資訊分析比賽進行狀態；接著在 6.2.2 節中利用人與球的高度關係進行球高度變化分析；結合 6.2.1 節和 6.2.2 節的分析結果，在 6.2.3 節中進行持球方動作分類。

6.2.1 比賽進行狀態分析

在球類比賽中，大致可分為比賽開始、比賽進行中、比賽暫停、比賽結束等比賽狀態。在一對一鬥牛比賽中，由於比賽時間短，也沒有換人或討論戰術的情形，比賽狀態較為單純，常見的比賽進行狀態有攻守交換時的洗球以及洗球後的開始鬥牛。因此在分析比賽進行狀態時，我們主要是判斷比賽處於洗球狀態或是鬥牛狀態。

我們知道洗球是指球在雙方手中來回傳遞，由圖 6-2 中可看出當球從一方傳遞到另一方手上時，中間會有連續幾張影像球與人體分離，如圖 6-2(a)中影像編號 2 到 5 和(b)中影像編號 2 到 4 所示。由於洗球時兩人保持一定距離範圍，依據上述特性我們找出人與球分離的連續影像，且影像總張數在合理張數內的影像片段，並利用影像中球的水平座標變化，判斷連續影像中球的移動方向。

在正常洗球情形下，一方接到球後就會回傳給另一方不會持球太久，所以球來回的間隔時間不會很長，加上在籃球一對一鬥牛賽中沒有隊友，不會有傳球情

形發生。因此在合理的時間範圍內和合理的連續張數狀況下，球有來回傳遞的情況即可判斷為洗球狀態，而其餘情況則判斷為鬥牛狀態。詳細流程圖請參考圖 6-3，其中 TH_{min} 和 TH_{max} 為連續影像張數門檻值，而 TH_{num} 為連續影像片段間的間隔門檻值，經由實驗設定 TH_{min} 值為 3， TH_{max} 值為 10，而 TH_{num} 值為 20。

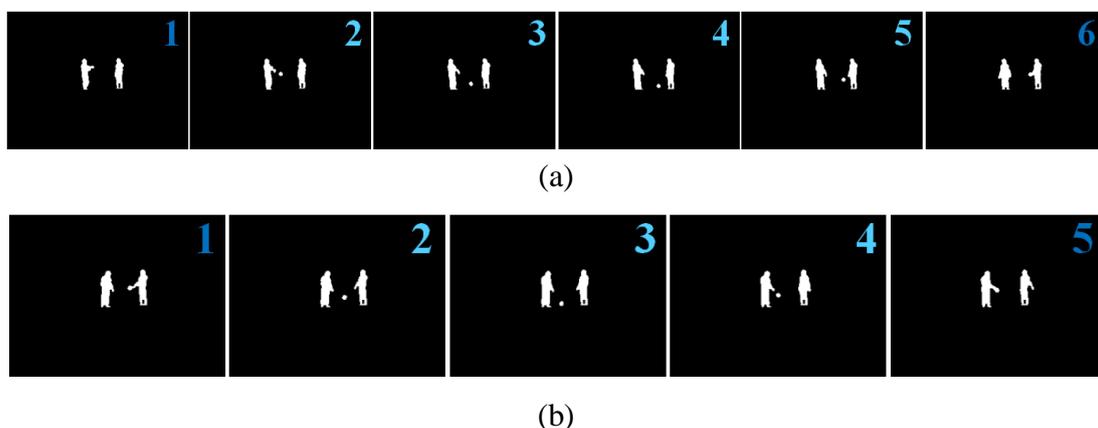


圖 6-2 洗球範例：(a) 球由左往右傳，(b) 球由右往左傳。

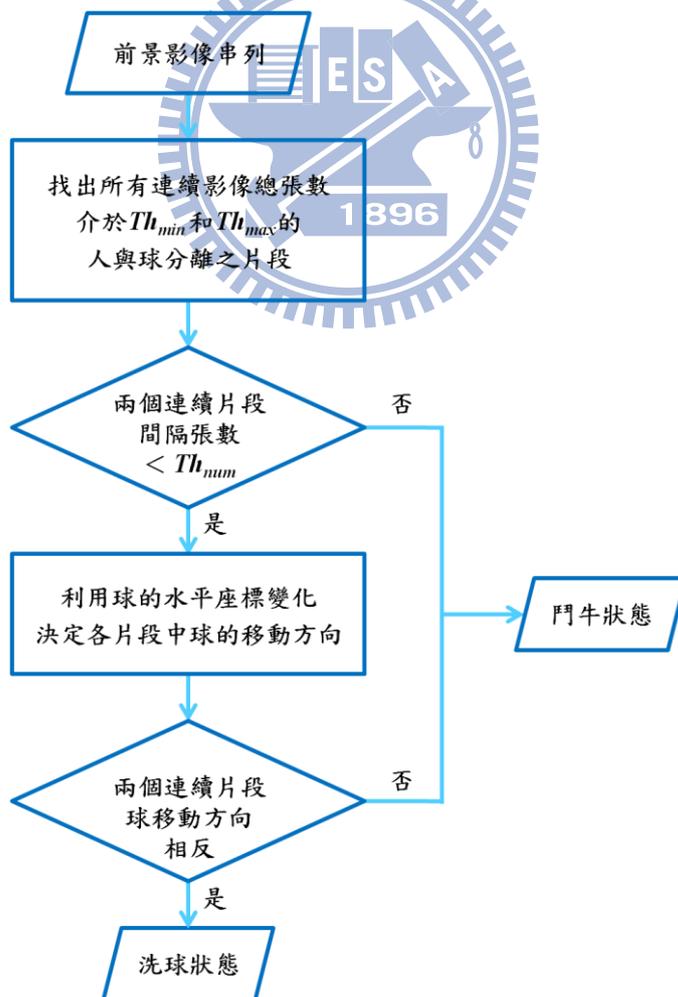


圖 6-3 在球與人分離的連續影像狀況下判斷比賽進行狀態之流程圖。

6.2.2 球高度變化分析

我們知道在籃球比賽中，人與球之間的高度關係隱含很多訊息。例如：球比人低時的可能狀況有運球、持球、抄球、撿球等；球比人高時的可能狀況有投球、跳球、搶籃板等。本節將人與球之間的高度關係分為：球比人高、球比人低，以及球與人連在一起，如圖 6-4 所示。

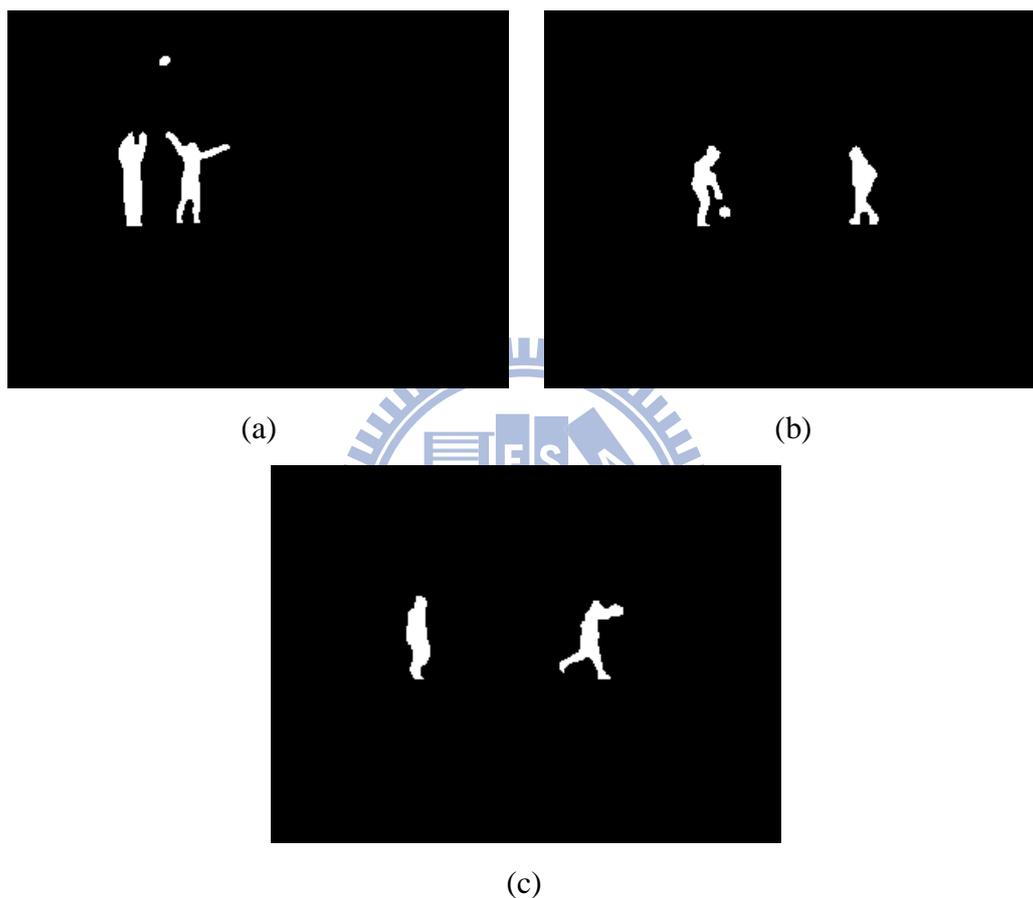


圖 6-4 人與球之間的高度關係：(a) 球比人高，(b) 球比人低，(c) 球與人連在一起的過渡期。

6.2.3 持球方動作分析

本節將結合 6.2.1 節和 6.2.2 節中得到的比賽進行狀態與球高度變化資訊，對持球方進行動作分類。首先我們利用球高度進行判斷，當球比人高時判斷為持球方投籃動作，而當投球後，球由比人高轉為比人低時，在還沒有與人連在一起前，表示還沒有人搶到球，因此這段時間判斷為搶籃板動作。接著我們結合比賽進行狀態進行判斷，洗球後到持球方投籃的這段時間，如果球一直都與人連在一起表

示持球方為洗球後直接投籃，因此這段時間判斷為持球方投籃動作；而在持球方投籃後至洗球前的這段時間將判斷為持球方進球後之撿球動作，其餘的鬥牛狀態部分將判斷為持球方運球動作。

6.3 實驗結果與討論

我們針對籃球一對一鬥牛賽設計一套階層式動作分析系統，在 6.3.1 節中會說明此系統介面，並於 6.3.2 節中進行分析結果討論與驗證。

6.3.1 系統介面

籃球一對一鬥牛賽之階層式動作分析系統介面如圖 6-5 所示，階層式架構由上而下分別為比賽進行狀態、球高度狀態、持球方動作狀態與前景影像。

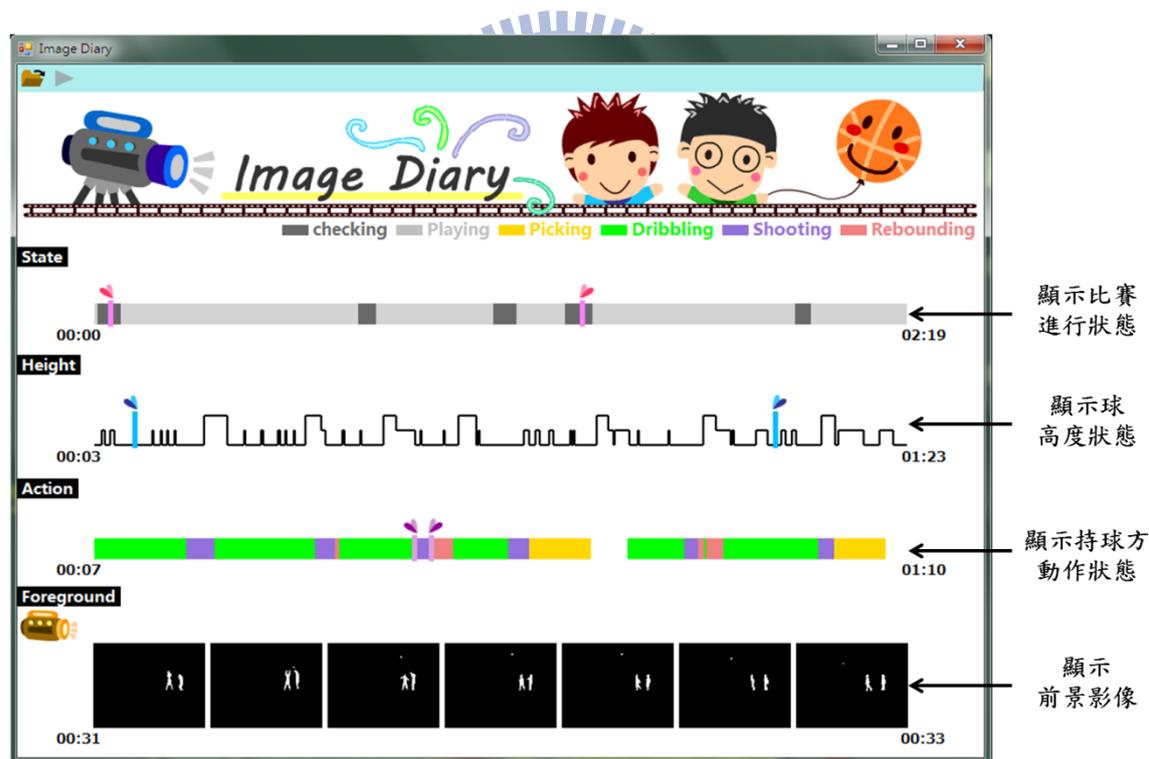


圖 6-5 籃球一對一鬥牛賽之階層式動作分析系統介面。

階層式結構中的第一層「比賽進行狀態」，以兩種不同深度的灰色表示，深灰色為洗球狀態，淺灰色為鬥牛狀態。階層式結構中的第二層「球高度狀態」，以三種不同高度的線段表示，由高而低分別為球比人高、球比人低、球與人連在一起。階層式結構中的第三層「持球方動作狀態」，以不同顏色作為區分，其中

持球方運球用綠色表示、投籃用紫色表示、搶籃板用紅色表示、進球後的撿球用黃色表示，而沒有顯示顏色的部分則為第一層中的洗球狀態。階層式結構中的最後一層「前景影像」，將顯示影像中人與球的前景區域。

6.3.2 階層式狀態分析結果

本節將以一段長約三分鐘的模擬籃球一對一鬥牛賽影片，進行階層式動作分析，並對各層之分析結果進行討論與驗證。

(1) 比賽進行狀態分析結果

由於本論文中判斷洗球的方法是利用人與球分離的連續影像，對球進行移動方向判斷，當連續兩段人與球分離的連續影像，球移動方向有來回的情形發生，則將該兩段連續影像包含中間人與球連在一起的影像，整個判斷為洗球狀態，如圖 6-6 中紅色框起來的部分。由於此判斷方法是由人與球分離的連續影像做為判斷起始與結尾，所以在第一次洗球時，球傳遞出去前停留在人手中，以及最後一次洗球時，球接到後進行下一個動作前，這兩種情形會被誤判成鬥牛狀態。

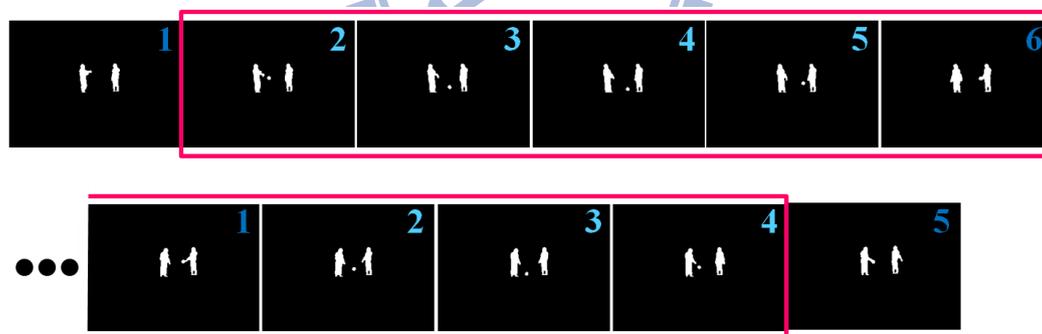


圖 6-6 洗球判斷範圍。



圖 6-7 比賽進行狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

本論文比賽進行狀態分析結果如圖 6-7 所示，其中上排為本論文判斷結果，與下排實際狀態比較可清楚看出，五段深灰色的洗球狀態大致都有正確判斷出來，只有在上述提及的洗球起始點、結束點會與實際狀態有些微的落差。

(2) 球高度狀態分析結果

本論文球高度狀態分析結果如圖 6-8 所示，其中上排為本論文判斷結果，與下排實際狀態比較可清楚看出，當球比人高時，判斷結果與實際狀態相近，而球比人低和球與人連在一起的情況，則可能會因攝影機架設位置的影響，造成球被人遮蔽，使得實際狀態是球比人低的情形，被誤判成球與人連在一起，如圖 6-9 所示。

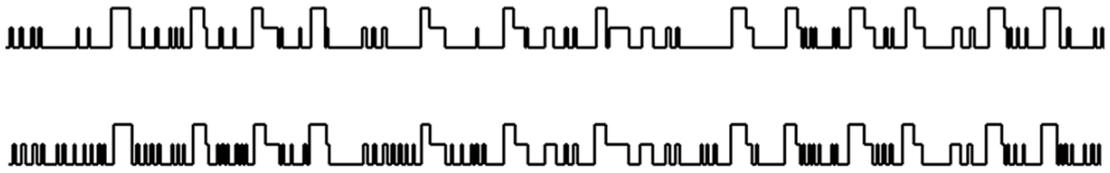


圖 6-8 球高度狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

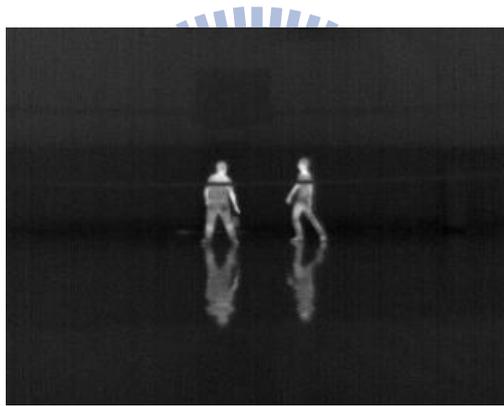


圖 6-9 球被人遮蔽之例子。

(3) 持球方動作狀態分析結果

本論文持球方動作狀態分析結果如圖 6-10 所示，其中上排為本論文判斷結果，與下排實際狀態比較可清楚看出，不同顏色的持球方動作絕大多數都有確實判斷出來，但在各動作起始點與終止點會與實際狀態產生些微差距。

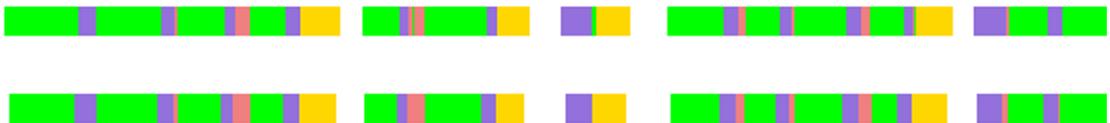


圖 6-10 持球方動作狀態分析結果，上排為本論文判斷結果，下排為實際狀態。

由於本論文持球方投籃動作判斷方法是當球與人分離且比人高時，才會將此判斷為投籃動作，如圖 6-11(b)；而當手舉起來準備投籃時，如圖 6-11(a)，則不會被判斷成投籃動作，因為有可能是假動作，所以仍判斷為運球動作。因此在圖 6-10 之分析結果中，紫色投籃動作的起始時段會比實際狀態在時序上有些微落後。

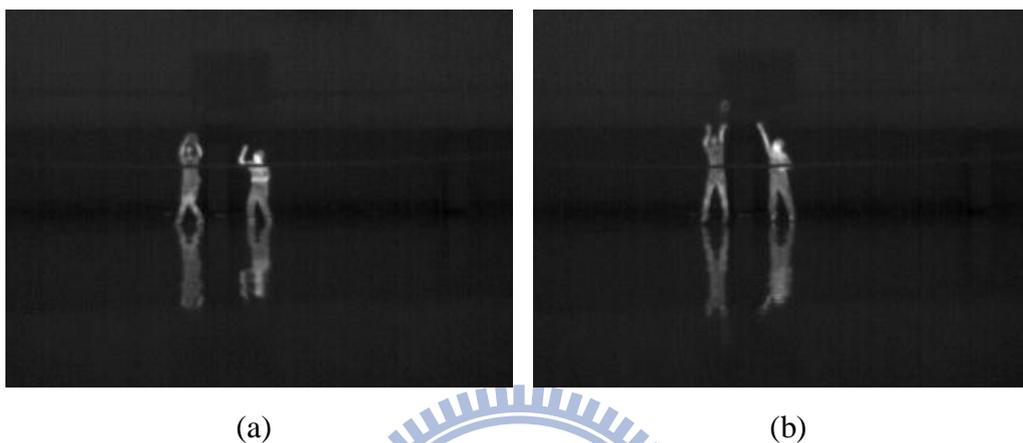


圖 6-11 投籃過程之範例：(a) 準備投籃，(b) 球投出去。



第七章 結論與未來展望

在本篇論文中，我們實作了一個基於人物動作分析的階層式影像日誌系統，可以提供給工作繁忙的家屬及時掌握居家年長者的生活狀況與安全。這個階層式系統結合人類行為認知以及監控視訊瀏覽與檢索兩個研究領域，利用電腦視覺之技術，透過監視攝影機影像來分析被觀測者的活動狀態、高度變化、行為動作，以及識別各肢體部位建立人體姿態模型，並提供使用者方便操作及快速瀏覽的使用者介面。

本階層式人物動作分析系統使用人體質心位置進行人物活動狀態分析，並利用人體外接矩形進行人物高度狀態分析，再結合兩者進行人物動作狀態分析，包含日常生活中，較常見的基本動作如站立、走路、坐下、躺下等，以及針對較常見也是危險度較高的跌倒行為，設計一套快速且有效的偵測方法。最後，改進與補強[11]的頭、手、腳定位方法，得到更準確的人體姿態識別模型。

本階層式影像日誌系統提供一個可靠、具高親和力的使用者介面。介面主要功能有：可選擇與輸入想觀看時段的視訊檔案；利用階層方式，顯示由人物動作分析技術得到的判斷結果，階層架構由上而下分別為人物活動狀態、高度狀態、動作狀態與姿態識別影像；能縮放顯示範圍及選取欲進一步觀看下一層狀態的時段；觀看下層狀態時可顯示上層於該時段的對應位置；針對影片中人物活動狀態與動作狀態進行檢索。

本階層式系統架構除了可應用於居家看護系統，也可以應用於球類運動訓練與比賽計分之輔助，未來甚至可應用於購物中心、車站及娛樂場所等之顧客流量分析，以進行收銀與售票櫃檯的人力機動調度。

參考文獻

- [1] 行政院經濟建設委員會，2010 年至 2060 年臺灣人口推計報告，
<http://www.cepd.gov.tw/m1.aspx?sNo=0000455>，Sep. 2010。
- [2] 行政院內政部統計處，98 年老人狀況調查結果，
<http://sowf.moi.gov.tw/stat/week/list.htm>，March 2011。
- [3] C.-C. Li and Y.-Y. Chen, "Human Posture Recognition by Simple Rules," *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 4, pp. 3237-3240, Oct. 2006.
- [4] B. Hosseini, S. V. Mashak, and S. A. R. Abu-Bakar, "Categorization of Human Movement Based on Rule-Based Classifier," *Fourth Asia International Conference on Mathematical / Analytical Modelling and Computer Simulation*, pp. 341-345, May 2010.
- [5] X. Yuan and X. Yang, "A Robust Human Action Recognition System Using Single Camera," *International Conference on Computational Intelligence and Software Engineering*, pp. 1-4, Dec. 2009.
- [6] Z. Zhou, X. Chen, Y.-C. Chung, Z. He, T. X. Han, and J. M. Keller, "Activity Analysis, Summarization, and Visualization for Indoor Human Activity Monitoring," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, pp. 1489-1498, Nov. 2008.
- [7] H. Foughi, B. S. Aski, and H. Pourreza, "Intelligent Video Surveillance for Monitoring Fall Detection of Elderly in Home Environments," *11th International Conference on Computer and Information Technology*, pp. 219-224, Dec. 2008.
- [8] F. Nater, H. Grabner, and L. V. Gool, "Exploiting Simple Hierarchies for Unsupervised Human Behavior Analysis," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2014-2021, June 2010.
- [9] H. Jiang, "Human Pose Estimation Using Consistent Max-Covering," *IEEE 12th International Conference on Computer Vision*, pp. 1357-1364, Oct. 2009.
- [10] L. Zhu, Y. Chen, Y. Lu, C. Lin, and A. Yuille, "Max Margin AND/OR Graph Learning for Parsing The Human Body," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, June 2008.
- [11] S.-C. Huang, "Human Posture Recognition System Based on Image Scanning," Master's Thesis, Dept. Multimedia and Engineering, Nation Chiao Tung University, June 2009.
- [12] L. Chang, Y. Yang, and X.-S. Hua, "Smart Video Player," *IEEE International Conference on Multimedia and Expo*, pp. 1605-1606, June 2008.
- [13] Y. Pritch, A. Rav-Acha, and S. Peleg, "Nonchronological Video Synopsis and Indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.

- 30, pp. 1971-1984, Nov. 2008.
- [14] Z. Ji, Y. Su, R. Qian, and J. Ma, "Surveillance Video Summarization Based on Moving Object Detection and Trajectory Extraction," *2nd International Conference on Signal Processing Systems*, Vol. 2, pp. V2-250-V2-253, July 2010.
- [15] A. Gupta, P. Srinivasan, J. Shi, and L. S. Davis, "Understanding Videos, Constructing Plots Learning A Visually Grounded Storyline Model from Annotated Videos," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2012-2019, June 2009.
- [16] S. Kurabayashi, T. Ueno, and Y. Kiyoki, "A Context-Based Whole Video Retrieval System with Dynamic Video Stream Analysis Mechanisms," *11th IEEE International Symposium on Multimedia*, pp. 505-510, Dec.2009.
- [17] B. V. Patel, A. V. Deorankar, and B. B. Meshram, "Content Based Video Retrieval Using Entropy, Edge Detection, Black and White Color Features," *2nd International Conference on Computer Engineering and Technology*, Vol. 6, pp. V6-272-V6-276, April 2010.
- [18] S. de Avila, A. da Luz, and A. de A. Araujo, "VSUMM: A Simple and Efficient Approach for Automatic Video Summarization," *15th International Conference on Systems, Signals and Image Processing*, pp. 449-452, June 2008.
- [19] X.-J. Yuan and H.-C. Feng , "An Abrupt Shot Change Detection Algorithm Based on The YUV Space," *2010 International Conference on Electrical and Control Engineering*, pp. 4630-4633, June 2010.
- [20] J. Chang, R.-M. Hu, and H. Zhang, "Summarization Extraction Method for Surveillance Video Based on Color Spatial Distribution Characteristic," *International Conference on Computational Intelligence and Software Engineering*, pp. 1-5, Dec. 2009.
- [21] H. Wang and R. Tong, "Hierarchical Video Data Modeling and Indexing for Virtual Scene Construction," *Second International Conference on Machine Vision*, pp. 69-73, Dec. 2009.
- [22] L. Pan, X. Wu, and X. Shu, "Key Frame Extraction Based on Sub-shot Segmentation and Entropy Computing," *Chinese Conference on Pattern Recognition*, pp. 1-5, Nov. 2009.
- [23] M. Hoai, Z.-Z. Lan, and F. de la Torre, "Joint Segmentation and Classification of Human Actions in Video," *IEEE Conference on Computer Vision and Pattern Recognition*, June 2011.
- [24] K. M. Kitani, T. Okabe, Y. Sato, and A. Sugimoto, "Fast Unsupervised Ego-Action Learning for First-Person Sports Videos," *IEEE Conference on Computer Vision and Pattern Recognition*, June 2011.

- [25] J. P. Collomosse, G. McNeill, and Y. Qian, "Storyboard Sketches for Content Based Video Retrieval," *IEEE 12th International Conference on Computer Vision*, pp. 245-252, Oct. 2009.
- [26] L. Ma and Z. Liu, "Human Motion Recognition in Video," *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, Vol. 4, pp. 49-53, Oct. 2008.

