

國立交通大學

多媒體工程系

碩士論文

利用多攝影機在人群中決定人物之位置與身高

People Localization and Height Determination for Dense Crowds
Using Multiple Cameras

研究生：黃星陸

指導教授：莊仁輝 教授

中華民國一〇〇年六月

利用多攝影機在人群中決定人物之位置與身高
People Localization and Height Determination for Dense Crowds
Using Multiple Cameras

研究生：黃星陸

Student : Hsing-Lu Huang

指導教授：莊仁輝

Advisor : Jen-Hui Chuang



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 and Information Science

June 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇〇年六月

利用多攝影機在人群中決定人物之位置與身高

學生：黃星陸

指導教授：莊仁輝 博士

國立交通大學多媒體工程學系（研究所）碩士班

摘 要

近幾年來以多攝影機進行多人物定位與追蹤的研究越來越受到重視，其中大部分的方法都需要大量的計算才能夠處理嚴重遮蔽的問題，也因此往往需要倚賴特殊硬體才能達成即時的定位與追蹤。不同於這些研究，本論文主要特色在於以取樣過的前景樣本，投影於地面以限縮人物可能出現之區域，再利用人物所在區域會被較多投影的前景樣本覆蓋之特性，設計一隨機演算法能於限縮的區域中快速地尋找出人物的數目及位置。由於所需的搜尋空間以及次數都大量的被減低，使得所需的計算量也大幅的被減少，因此我們的方法能夠提供即時的人物定位結果。此外我們利用了人物於空間中佔有一定的體積之觀念，以等間隔的地面取樣來排除因人物間彼此遮蔽所產生的錯誤偵測。進而達成在嚴重遮蔽情況下仍能正確地偵測出人物所在的位置，並提供準確的定位結果。

People Localization and Height Determination for Dense Crowds Using Multiple Cameras

Student : Hsing-Lu Huang

Advisor : Dr. Jen-Hui Chuang

Institute of Multimedia Engineering

National Chiao Tung University

ABSTRACT

In recent years, researchers have been paying much attention on people tracking and localization using multiple cameras. Most of methods require a large number of computations to cope with serious occlusions, and need to rely on special hardware to achieve real-time locating and tracking. Unlike these studies, we use two dimension line samplings of foregrounds to restrict regions of possible locations of people. According to the nature of occupancy constraint that possible locations of people should be covered by more projected foreground pixels, we propose a random algorithm to efficiently find the locations and number of people. Our methods can provide real-time location results because of the smaller of searching space from line sampling and less computing time due to the randomness nature. Besides, we exploit the concept of *people volume* so that uniformly sampling on the ground plane can prevent error from occlusions. Experimental results show that our approach provides real-time and accurate people localization results under serious occlusions.

誌 謝

首先我必須誠摯地感謝指導教授莊仁輝博士，老師在電腦視覺領域淵博的學問並悉心指導，在我研究的路上不斷給予我建議與鼓勵，讓我得以完成這篇論文，老師嚴謹的研究態度也是我效法的典範。感謝三位口試委員王聖智教授、黃仲陵教授以及賴飛熊教授對本論文寶貴的建議，能夠讓本論文更加充實完整。

感謝實驗室博士班泓宏學長與國華學長，研究所以來提供我許多建議及指導，讓我在實驗過程中能夠更加順遂。感謝同儕們涵恩、育璋、東諺及芝穎，在我研究所生涯中，與你們共同修課、討論、打打鬧鬧以及打電動玩樂等一點一滴都是我永遠無法忘記的回憶。感謝實驗室宣良、永昌、尚一、聖中、世旻、維辰及邦展學長們，還有致傑、怡廷、國濂、暉晴、沛怡學弟妹們，我的研究所生涯也因為有你們而更加多采多姿。感謝維桓學長，在我需要幫忙時，每次都盡心盡力地予我幫忙並且開導我許多人生的觀念。還有感謝在我需要鼓勵與支持時給予我動力的咻 D，也提供我許多更多元化的想法與作法，加上與妳時常地聊天說笑，讓我能夠卸下心中所有的煩惱，能更加努力地學習與成長，讓我感到非常溫暖，亦為我增添許多歡樂與上進的向心力。

我必須感謝我親愛的父母親，長年來供應我衣食無缺的生活，並提供我一個溫暖的避風港。在長達多年的新竹生活中，持續支持著相信著我能夠完成所想要做的事情，真的非常感謝如此地照顧與愛護我。感謝姊姊和姊夫，每次當我回花蓮的時候，都邀請我到家裡玩，還請我大啖美食與看電影，身心感覺真的放鬆許多。

感謝所有一路走來陪伴我的人，沒有你們，就沒有今天的我，更不會有這篇論文的完成，感恩！

目 錄

中文摘要.....	iii
英文摘要.....	iii
誌謝.....	iii
目錄.....	iii
圖目錄.....	v
表目錄.....	vi
一、 緒論.....	1
研究動機.....	1
論文架構.....	2
二、 相關研究.....	3
三、 產生人物候選位置.....	9
3.1. 計算影像中之消失點.....	10
3.2. 以消失點為基礎對前景進行線取樣.....	11
3.3. 人物候選區域之地面均勻取樣.....	12
四、 決定人數及候選位置.....	15
4.1. 隨機取樣演算法.....	15
4.1.1. 建立三維軸線（3D major axis, 3DMA）.....	16
4.1.2. 隨機取樣.....	16
4.1.3. 初步驗證機制.....	17
4.2. 體積概念與人物過近移除假象人物.....	18
4.3. 重新取樣之範圍.....	22
五、 實驗結果與討論.....	26
5.1. 體積機制濾除人物假象之實驗結果.....	26
5.2. 人物過近機制濾除人物假象.....	28
六、 結論.....	32
參考文獻.....	33

圖目錄

圖 2-1 藍線交點為人物立足點.....	4
圖 2-2 綠色人物在 View 1 中被遮蔽，但仍能透過其他 View 而找出綠色交點.....	4
圖 2-3 從 從協同圖中可看出較高區域（紅色）為人物可能的立足區域.....	5
圖 2-4 藍軸交點為人物立足點，其中 C_g 為攝影機與地平面之垂足.....	5
圖 2-5 紅點位置為人物假象之位置.....	6
圖 2-6 左圖：極座標圖和可能射線；右圖：統計值較低的藍色區域被判斷為假象區域.....	6
圖 2-7 以消失點為基礎的前景線取樣.....	7
圖 2-8 紅色線段為利用前景像素及主成份分析法計算出的二維軸線.....	8
圖 2-9 綠色最長軸為計算不同高度平面的 L_1 與 L_2 的投影交點所產生的一個三維軸線.....	8
圖 3-1 同一場景的五個攝影機原始影像。左圖：原始影像；右圖：對應左圖之前景影像.....	9
圖 3-2 V_p 為對某個畫面所算出之消失點.....	11
圖 3-3 (a)-(e)分別為圖 3-1 中五個多視角影像之二維線取樣.....	12
圖 3-4 地平面上等間隔之樣本點.....	13
圖 3-5 左圖：二維線取樣；右圖：對應左圖二維線取樣之 USGP.....	13
圖 3-6 五組一維 USGP 在地面上的分佈情形.....	14
圖 4-1 圓中心為多組 1D-USGP 經隨機取樣演算法分群後後之群中心，此結果為分成九群.....	17
圖 4-2 非圓心的粗點為步驟 b-II. 所提到距離圓心 50 cm 內能取代群中心的 USGP.....	18
圖 4-3 群 7（圖 7）的中心經權重移位後由 A 移至 B 點.....	18
圖 4-4 由第 4-1 節之隨機取樣所得到的人物初步位置.....	19
圖 4-5 桃紅色區域為圖 4-4 中以初步位置為中心，周圍 60 cm 所測出來的 IFR 高度圖.....	20
圖 4-6 根據圖 4-5 中的桃紅區域中依 IFR 大小經權重移位後中心位置由白點移至黃點.....	20
圖 4-7 根據圖 4-6 的中心位置與人物高度描繪於現實場景中之結果.....	21
圖 4-8 圖 9 為假象人物合併之結果.....	21
圖 4-9 圖 6 右上方的群因為 Volume 不足而被移除.....	22

圖 4-10 場景 1 中 TH_IFR = 0.80 之紮實頂視圖	23
圖 4-11 (a)-(b)分別為場景 2 及場景 3 之紮實頂視圖	23
圖 4-12 (a)-(d)為場景 4 第 601-631 幀中 TH_IFR = 0.80 之紮實頂視圖.....	24
圖 4-13 (a)-(b) 分別為場景 4 與場景 5 中 TH_IFR = 0.65 之紮實頂視圖	25
圖 5-1 左圖：第 4.1 節之方法得出之初步人物位置；右圖：經過體積濾除假象人物之結果	26
圖 5-2 (a)-(c)為圖 5-1 之實際定位結果	27
圖 5-3 (a)黃圈為無假象之定位結果。(b)-(d)中黃圈顯示修正過後人物均能被穩定定位	29
圖 5-4 (a)-(d)分別為圖 5-3 (a)-(d)之實際定位結果.....	30
圖 5-5 場景 4 中第 609 幀的結果.....	31

表 目 錄

表格 1 場景 1 至場景 5 中的正確率與執行速度之量化結果。.....	31
---------------------------------------	----



一、緒論

研究動機

以視覺進行人物追蹤與定位的相關研究在近幾年越來越受到重視，尤其是真實的應用環境下，人物往往是以多人且群聚地方式出現，因此真正能實用的系統必須要能夠克服在複雜多人的情況下估算出正確的位置。在單一攝影機監控系統中，若所追蹤的人物在影像上與其他人是分離的，會因為目標視覺特徵（如：顏色、形狀及運動方向等）大致上變化不大，可以利用此特性來分析追蹤。而在擁擠的場景裡，由於人群密度高容易造成遮蔽現象，可能會有前景像素叢集成一團，但卻不保證該叢集的前景像素是屬於單獨一個人或多個人，甚至會發生一個人被其他人給完全遮蔽的情形。

而由近年來的相關研究我們也發現到這類的研究已經由使用單攝影機趨向於使用多攝影。主要的原因在於使用單一攝影機時因為人物間遮蔽之影響，所獲得的資訊相對於多攝影來說較為不足；相對地，多攝影機所提供的多視角影像通常能有較多的資訊。在多攝影機的相關研究中，近年來有些研究可以提供在多人擁擠的環境下正確地定位〔1, 2, 3〕，但往往需要較長的時間，而無法提供即時的定位與追蹤結果。本研究提出以多攝影機進行快速且精確的多人物定位方法，透過對於前景影像的取樣，有效地限縮人物在三維位置中可能的位置，另外再利用三維取樣及驗證機制搭配多畫面中的前景區域，有效率地估計出人物的人數以及位置與高度。

本研究的主要貢獻有三項。首先，我們能在嚴重遮蔽的情況下正確地偵測出人物所在的位置。其二，該演算法能夠被即時地執行，可提供立即的人物定位結果，這樣的特色將可滿足各種需要即時系統運作的應用。實驗結果顯示，我們的方法可以比文獻〔3〕作者所提出的方法快 2 倍以上。

論文架構

本論文在第二章相關研究中將常見的三維定位與追蹤分成三類，並且說明其所使用的分析方法。第三章中，將介紹我們的方法，該方法將前／背景分割後的前景影像當成輸入，搭配已知垂直於地平面的消失點對前景進行線取樣，並以平面投影幾何轉換（homography）的方法將相關樣本轉換至參考平面，有效地限縮人物在三維中可能的位置。第四章介紹從人物可能的三維位置中，依體積的概念來決定人物的位置與身高。第五章為利用第三、四章之方法所計算出之人物定位結果。



二、相關研究

本研究採用多攝影機所提供的多視角影像來提供只用單一攝影機取得的不足資訊，彌補單一影像在人物間產生遮蔽時所提供模稜兩可的資訊。近年來，常見的三維定位與追蹤，可區分成三類：（1）分析人物的特徵點、（2）使用完整的前景資訊、（3）僅使用部分的前景資訊。

第一類的方法通常是利用粒子群演算法（particle filter 來作定位與追蹤），由於多攝影機可取得多視角影像，因此能獲得豐富的人物外觀資訊，因此形狀（如：頭、軀幹與下肢）與顏色可以有效地被應用於人物的偵測與追蹤，如文獻[4,5]。在文獻[4]中，使用了人物上半身之三維模型及利用似擬函數（likelihood function）來將前景中符合此三維模型的區域當成是人物前景。三維模型的好處在於不需要訓練顏色及形狀等特性，其中利用改良粒子群演算法（modified particle filter）計算出似擬函數最大值之作法，可用於多人的追蹤，以及使用數量較少的粒子群來找出人物在影像上的位置，其中計算前景裡何處較符合該三維模型（可旋轉），並以該符合之區域為人物所在的前景區域。文獻[6]利用標點過程（marked point process）來描述擁擠的人群，將標點過程聯結到空間隨機過程（spatial stochastic process）來管理場景中人物的數量與位置，並且替每個被標點的人物加註身體大小、形狀及運動方向之特性，以便於讓這些特性在追蹤時進行學習與更新。

第二類的方法將各畫面中的人物利用投影幾何轉換投影至參考平面（地平面），可觀察屬於同一人物的立足點被投影在相近的位置（如圖 2-1）。這樣的想法被普遍應用於許多相關的文獻中[1,7,8,9,10]，並且能解決人物遮蔽情形（如圖 2-2）。

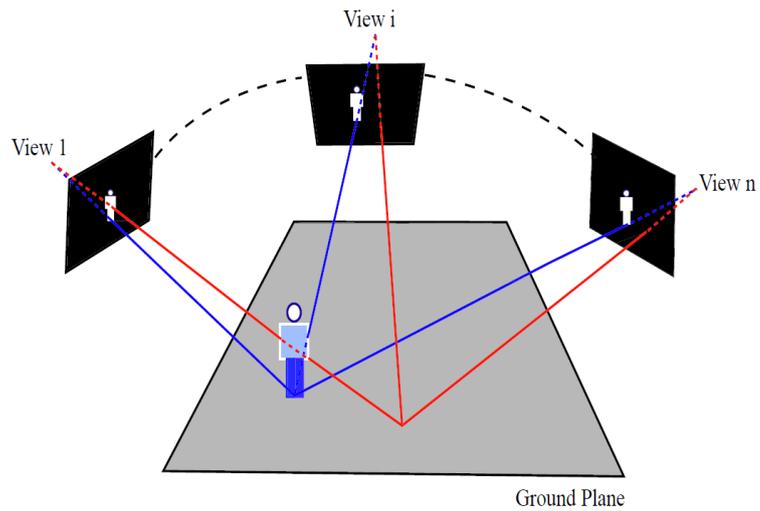


圖 2-1 藍線交點為人物立足點

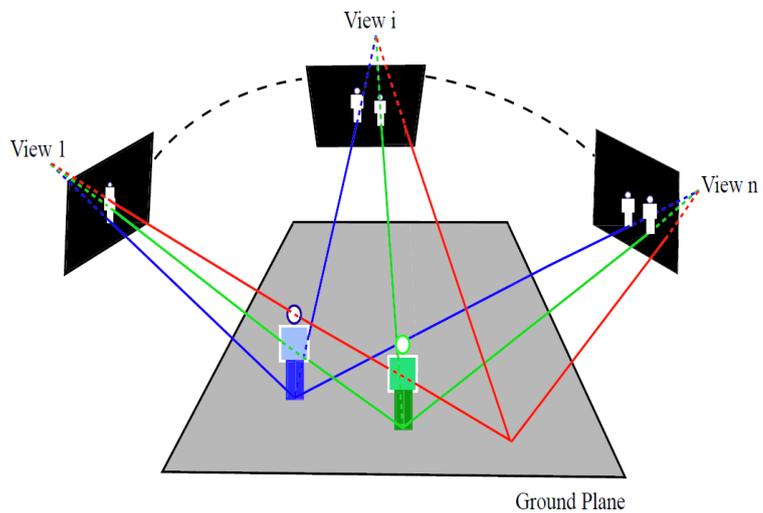


圖 2-2 綠色人物在 View 1 中被遮蔽，但仍能透過其他 View 而找出綠色交點

在文獻 [9] 中，是將人物利用投影幾何轉換將各視角影像，再利用高斯背景模型來做前／背景分割，計算每個前景像素所可能為前景的機率值，形成前景可能圖（foreground likelihood maps）；而後再利用投影幾何轉換將各前景可能圖投影至某一固定參考影像中並將前景可能值累加，該固定參考影像則成為協同圖（synergy map）。在協同圖中，前景可能值較高的區域，即形成人物的佔地區域（如圖 2-3）。

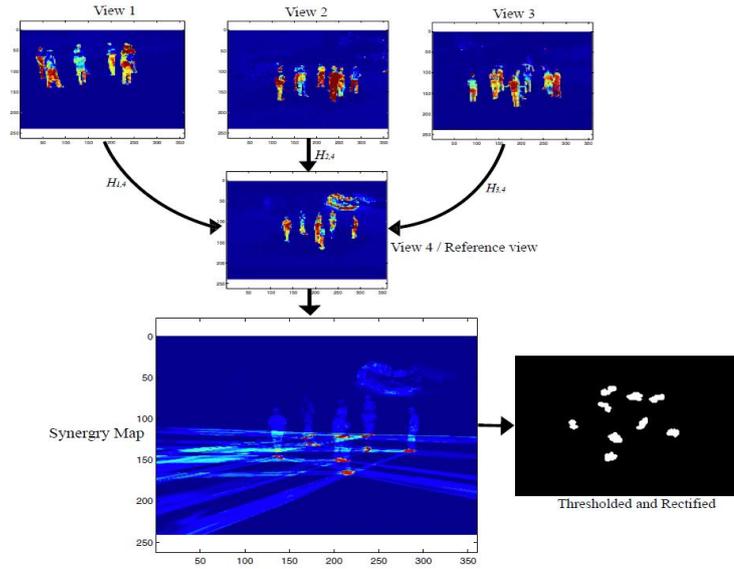


圖 2-3 從 從協同圖中可看出較高區域（紅色）為人物可能的立足區域

文獻 [1] 將前景投影至地平面，利用攝影機交於地平面的垂足，使用最小平方（least square）法找出前景在地平面通過攝影機腳點之軸線，所取得之各軸線交點即為人物立足點（如圖 2-4）。在使用將多人之前景投影至地平面找人物佔地區域時，因為前景多部位的相疊可能造成在無人的位置卻呈現有人的假象（如圖 2-5）。

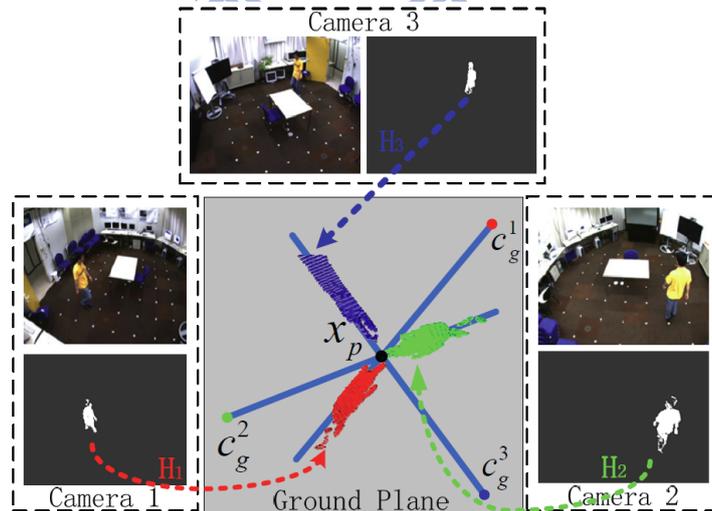


圖 2-4 藍軸交點為人物立足點，其中 C_g 為攝影機與地平面之垂足

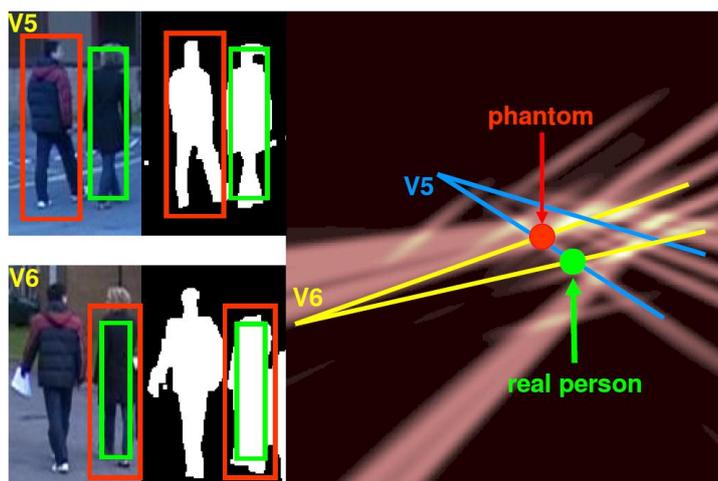


圖 2-5 紅點位置為人物假象之位置

文獻 [11] 則是提出支持積分法 (support integration) 來產生類似協同圖的積分圖，將積分圖中具有較高積分值的區域當成是人物的佔地區域。文獻 [7] 為了濾過假象區域，在投影至地平面之協同圖中，以各攝影機腳點為極坐標原點並調整角度而產生的線段為可能射線 (likelihood rays)，沿著可能射線統計所經各點的佔地可能性來濾除假象的區域 (如圖 2-6 藍色區域)。

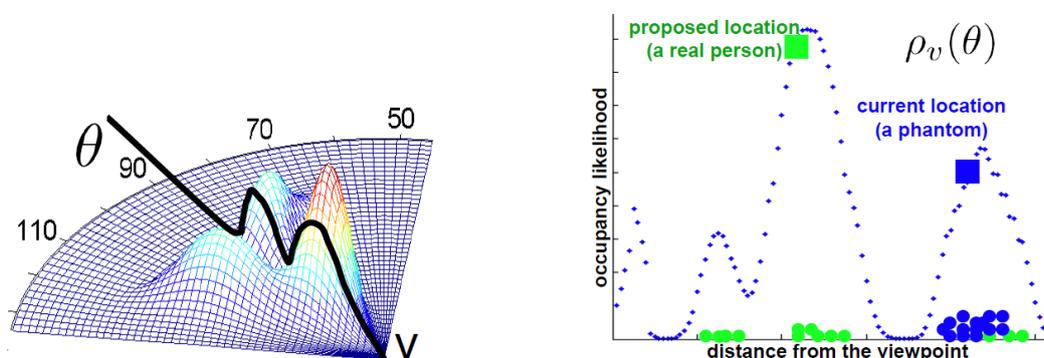


圖 2-6 左圖：極坐標圖和可能射線；右圖：統計值較低的藍色區域被判斷為假象區域

文獻 [12] 提出若將頭點使用某一高度之投影幾何轉換將各畫面中人物前景投影至某一影像中，並將其前景像素值累加，則頭點部份會形成具有較高的累加值。由於頭點部份在所屬之高度會有較高的累加值，因此可決定累加值高的區域為人物頭點之位置，利用頭點位置來追蹤可避免計算因遮蔽而看不到的立足點位置，其中所使用之投影幾何轉換之高度為人物身高。

接著介紹第三類的方法：人物在影像中可能會處於行走或站立的狀態，但在行走時立足點可能為前腳、後腳或中間。文獻 [13] 提出了使用人物軸線為基礎

的方式，以人物軸線代表前景來進行多人追蹤。由於攝影機架設具有一定的高度，使得垂直於地平面的竿子在影像中，隨著擺放的位置而在影像中非垂直地站立著並通過消失點。因此在文獻〔3〕中，作者提出以軸線的概念來代表前景，有別於許多先前介紹的方法需要投影所有的前景像素至參考平面，取而代之的是以消失點為基礎的線取樣來代表前景（如圖 2-7），當人群產生遮蔽時所產生的前景，僅利用單一軸線代表該前景較不適合。因此線取樣的概念能比單一軸線更具前景代表性，又能比投影全部前景像素減少大量的計算花費而使得人物定位更加有效率。作者提出以軸線為基礎的人群定位，是以影像中二維的線取樣來重建出空間中的三維軸線叢集（軸線高度可能為人物高度），再利用體積概念濾掉空間中不適合的軸線，所剩下的軸線叢集分佈，即為人物可能的所在區域。

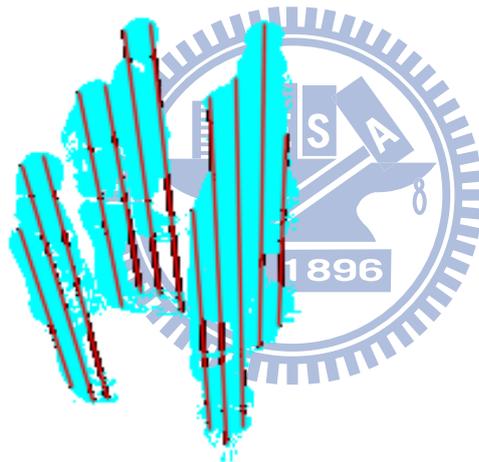


圖 2-7 以消失點為基礎的前景線取樣

為了用簡單的例子來介紹如何利用軸線來定位，以下先以單軸的單人定位來瞭解概念。圖 2-8 表示前景像素及利用主成份分析法（principal component analysis, PCA）所計算出的二維軸線。以圖 2-9 中的綠色軸為例，在畫面 2 中的 L_1 及畫面 2 中的 L_2 均代表同一位人物，而 L_1 與 L_2 分別能用投影幾何轉換矩陣 $H_{1\pi}$ 與 $H_{2\pi}$ 投影至參考平面 π 上並相交於 $P_{12\pi}$ 。隨著計算投影在不同高度平面的 L_1 與 L_2 的投影交點，將產生一個三維軸線（如圖 2-9 中之綠軸），此軸線將代表三維中的人物定位。接著將單人軸線觀念延伸至多人三維軸線取樣，經過分群後，將

分佈在同一群的三維軸線取樣之位置與高度做加權平均可以合成出一根三維軸線，此軸線之位置與高度，即代表人物之位置與高度。



圖 2-8 紅色線段為利用前景像素及主成份分析法計算出的二維軸線

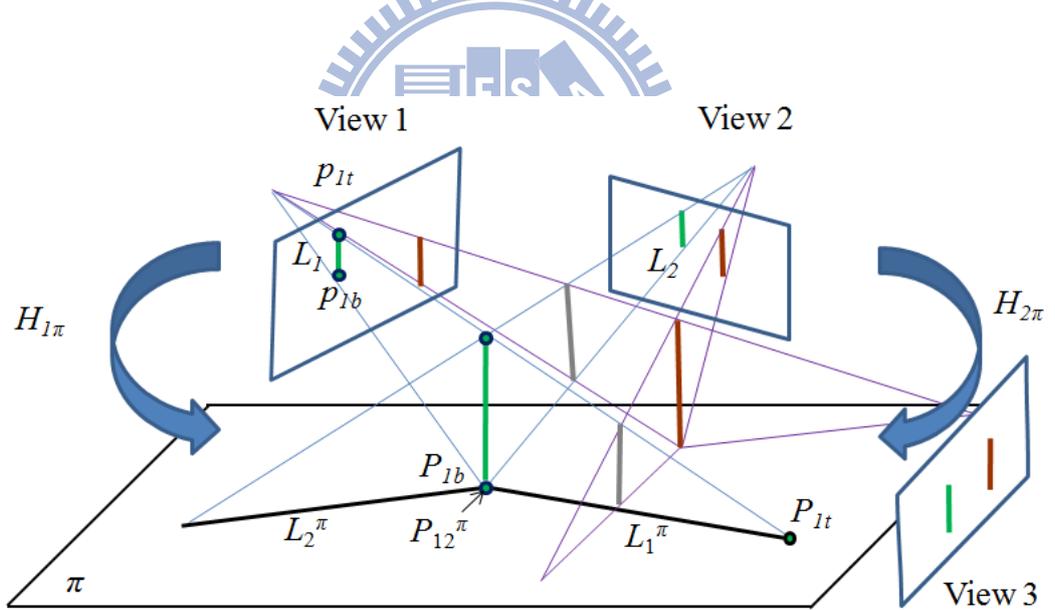


圖 2-9 綠色最長軸為計算不同高度平面的 L_1 與 L_2 的投影交點所產生的一個三維軸線

下一章將介紹本論文利用文獻〔3〕之線取樣、從二維影像轉換至三維參考平面以及不同高度之投影平面等概念，提出在參考平面上的隨機取樣找出人物之定位作法，再搭配多重高度之投影幾何轉換而計算出人物身高之方法。

三、產生人物候選位置

由於利用前景影像中二維線取樣來計算三維空間中的交點，其計算花費取決於二維線取樣之多寡，若需在二維線取樣中兩兩計算交點，將有許多時間花費在計算交點上。本章將介紹在不利用前景影像中二維線取樣於三維空間中計算交點的方式，來產生出人物可能立足的候選位置。第 3.1 節介紹影像中消失點之計算方式，以提供第 3.2 節使用來對前景做線取樣。第 3.2 節說明僅利用前／背景分割後的前景影像當成輸入搭配消失點的位置來做二維線取樣。第 3.3 節介紹將線取樣以平面投影幾何轉換(homography)的方法將前景影像二維線取樣轉換至參考平面，並在三維空間中做一維均勻取樣，得以有效地限縮在三維空間(地平面)中人物可能所在位置樣本數。

當監控的區域廣大時，要搜尋人物所在的位置往往需要耗費較多的時間，因此我們利用前景取樣的方式，有效率的限縮人物在影像中可能的所在位置。本論文利用文獻〔14〕的前景切割演算法取得前景影像(如圖 3-1)，搭配垂直於地平面直線的消失點來對前景區域進行線取樣，並將線取樣轉換至參考平面，再透過簡單的內差，即可找到人物可能所在的地面位置。

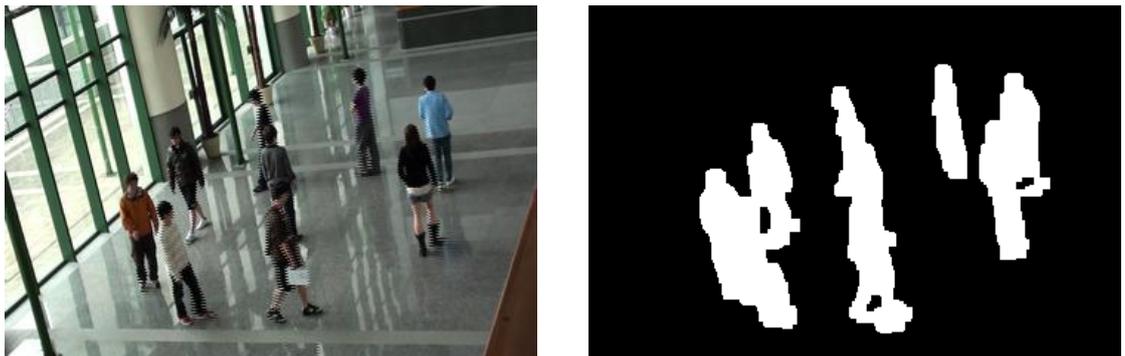


圖 3-1 同一場景的五個攝影機原始影像。左圖：原始影像；右圖：對應左圖之前景影像

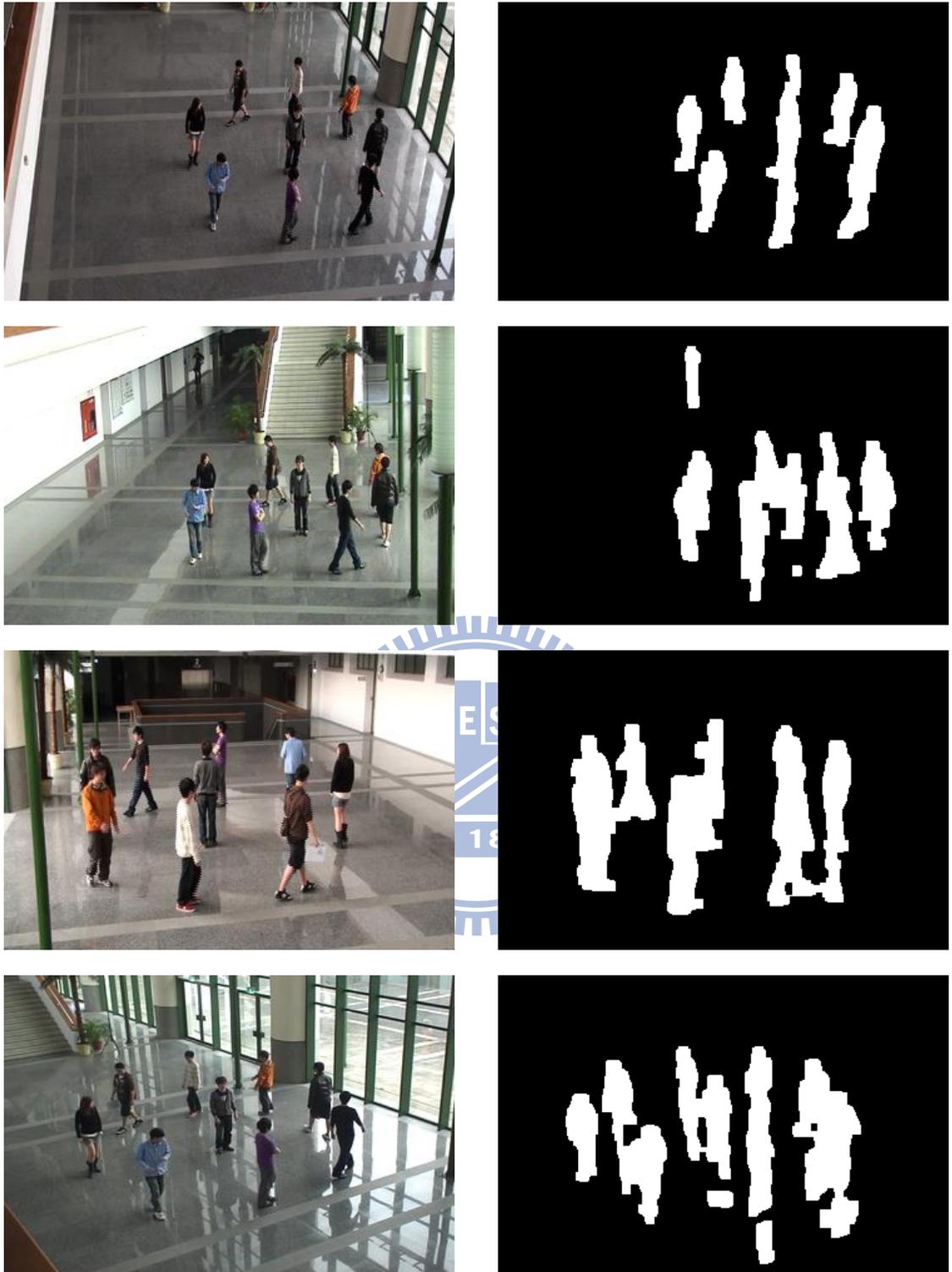


圖 3-1 (續)

3.1. 計算影像中之消失點

在三維場景中之平行線於二維影像上會交匯於一點，稱此交點為消失點（如

圖 3-2)。為了計算垂直地面直線在影像中的消失點，我們在場景中架立四根垂直於地平面的標記竿，利用在影像上的八個紅色標記，可計算出四條線，此四條線之交點即為消失點 V_p 。透過這樣的方式，我們可以為每一個畫面建立出垂直地面直線的消失點。

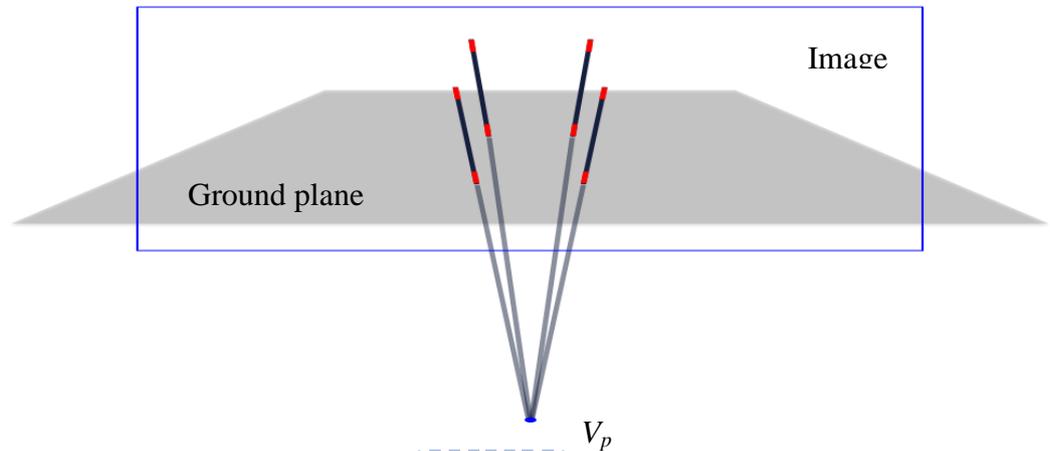


圖 3-2 V_p 為對某個畫面所算出之消失點

3.2. 以消失點為基礎對前景進行線取樣

本節介紹如何利用上一節所求得的消失點來對前景進行線取樣。我們以計算出的消失點為出發點對前景像素進行二維線取樣（如圖 3-3 (a)-(e)），其中每個相鄰的線取樣在影像的下邊界中具有相等的間隔。對於包含前景不足的線取樣將予以捨棄，因為它們可能遠離主要軸線（在前景邊緣處），而無法提供足夠的貢獻。對於包含足夠前景的線取樣，也必須能容忍前景分割的錯誤，如破洞及陰影（圖 3-3 (a)）。這樣的好處是取樣速度相當快速且後續產生候選位置是透過這些二維線取樣將可省去大量的時間。

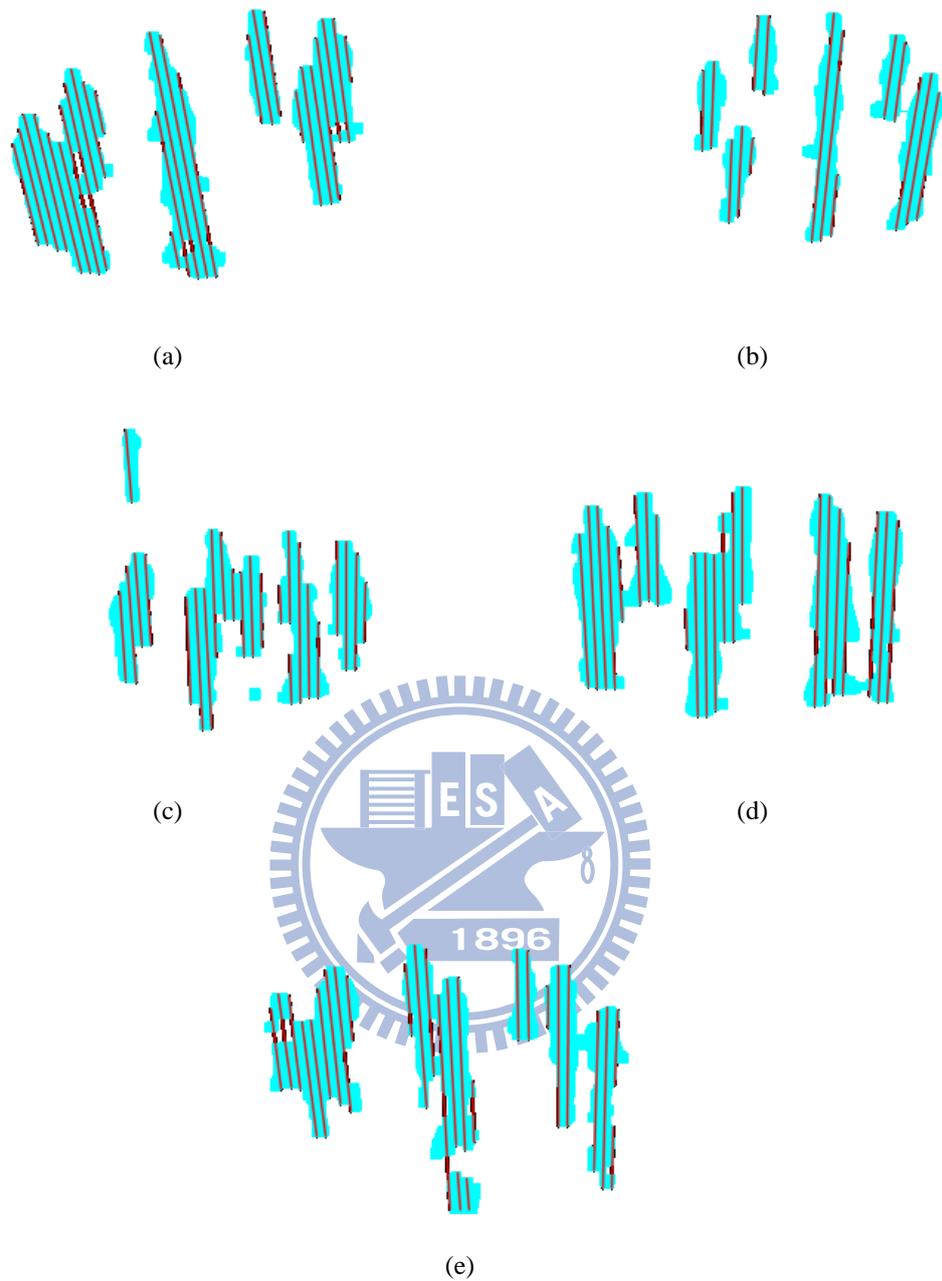


圖 3-3 (a)-(e)分別為圖 3-1 中五個多視角影像之二維線取樣

3.3. 人物候選區域之地面均勻取樣

將第 3.2 節中取得的二維線取樣投影至地平面，則在有較多二維線取樣經過的區域則較有可能是人物所在之位置，相似的觀念也被應用在 [3] [9] 中。然而在 [3] 中，將二維線取樣投影至地平面與計算線段交點坐標的計算花費依然偏高，我們將不計算線段的交點，而是利用線段交點的區域具有較多二維線取樣經過的

特性，於下一節介紹一套演算法期望能夠快速的找出此種區域（即人物可能的所在區域）。為此，我們先將線段兩端點投影至地平面，在投影後的兩點間僅使用簡單的內插建立一維等間隔之樣本點（uniform sampling of ground plane, USGP）如圖 3-4。圖 3-5 分別為五個畫面所建出的地面均勻取樣；圖 3-6 為合併五組畫面之一維 USGP，不難看出在地面上的 USGP 分佈有密疏的差異。

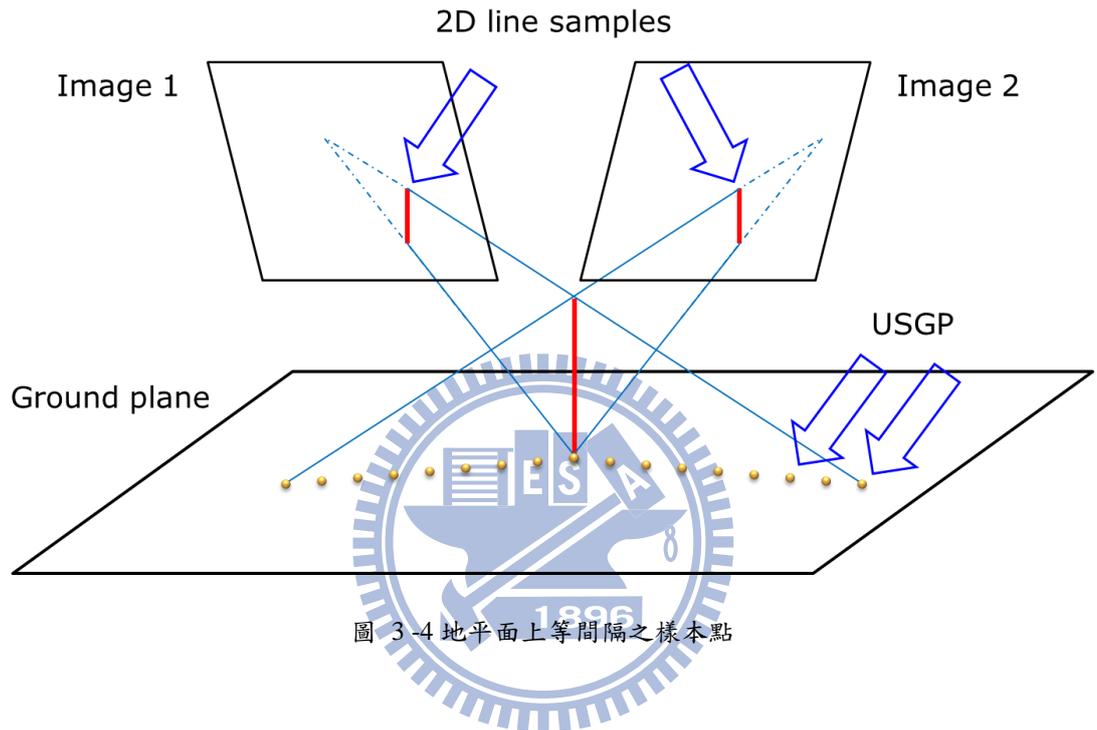


圖 3-4 地平面上等間隔之樣本點

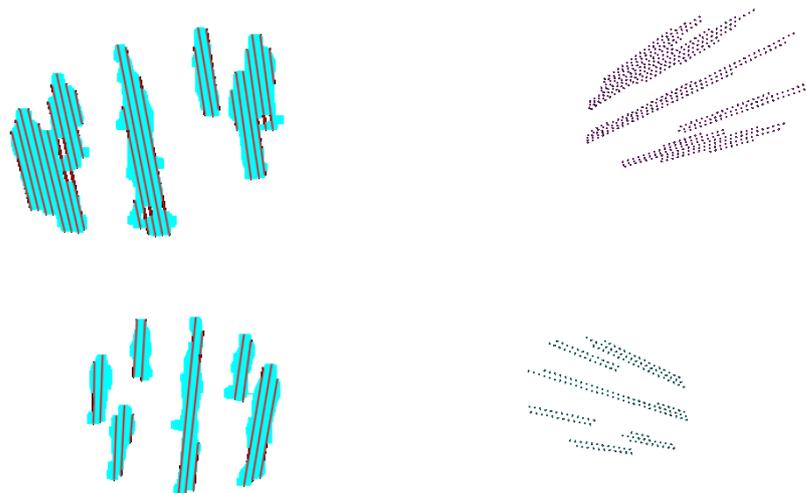


圖 3-5 左圖：二維線取樣；右圖：對應左圖二維線取樣之 USGP



圖 3-6 五組一維 USGP 在地面上的分佈情形

四、決定人數及候選位置

在文獻 [3] 中，作者以每對二維線取樣建立出可能的三維線取樣，再對這些三維線取樣對於所有畫面的前景進行平均前景覆蓋率 (average foreground coverage rate, AFGR) 驗證，這樣的作法在建立三維線取樣時會需要較大的計算量。不同於他們的方法，我們以 USGP 作為人物可能的位置，因此省去了在三維空間中算交點來建出三維線取樣所需的時間，同時我們可以利用 USGP 群聚處為人物所在的可能性較高的概念，有效地減低需檢驗的樣本數目，並同時完成分群 (clustering)，以決定人物的數目以及候選的位置。本章接下來所討論到的均為指一維 USGP，在下一章將會再提到二維 USGP。首先，本演算法中所使用的符號定義如下：

INTERVAL：4 (實際場景中一維 USGP 的間隔為 $200\text{cm} / 50 = 4\text{ cm}$)。

$Layer_i$ ：平面高度。其中 $Layer_0$ 為地平面， $Layer_{50}$ 高度為 200 cm。

ListSample：存放 USGP 的串列。

$USGP_i$ ：ListSample 中的第 i 個 USGP。

IFR_i ： $USGP_i$ 之影像前景率。

$3DMA_i$ ：以 AFGR 概念與 $USGP_i$ 建立出來的三維軸線。

ListCenter：儲放群中心的串列。

$Center_j$ ：ListCenter 中的第 j 個群中心。

DIS：兩個 USGP 間的距離 (10 cm)。

TH_IFR：IFR 閾值 0.75。

TH_RADIUS：半徑閾值 100 (實際場景中為 100 cm)。

TH_HALF：近距閾值 50 (實際場景中為 50 cm)。

4.1. 隨機取樣演算法

4.1.1. 建立三維軸線 (3D major axis, 3DMA)

本小節在介紹如何利用 AFCR 概念建立 3DMA。對於在 $Layer_0$ 上的每個 $USGP_i$ (實數座標)，將其反投影回各畫面之前景影像，若所對應的像素灰階值為 255，則 IFR_i 累加 1，否則不累加。當每個 $USGP$ 從 $Layer_0$ 至 $Layer_{50}$ 都累加完之後，對 IFR_i 做正規化的動作。因此 IFR_i 最小值為 0；最大值為 1。此時產生出 $3DMA_i$ ，其座標為 $USGP_i$ 之座標，高度比率為 IFR_i 。

4.1.2. 隨機取樣

上一節所介紹計算 IFR 的方法，若需要為每個 USGP 都計算 IFR，則耗費掉許多計算時間花費。由於在第三章介紹的方法所建立出來之人物候選位置，不必知道攝影機在地平面中的可視範圍，因此將候選位置儲存在一個一維陣列，不需額外空間儲存於二維平面中。再則，每個候選位置當中彼此並沒有上下左右的關聯性，但利用人物的可能所在位置會有較多 USGP 分佈的特性，可使用隨機取樣的方式篩選人物候選位置 (取樣位置在有較多 USGP 分佈的區域有較高被選取的可能性)。此節在說明利用隨機取樣演算法來建立 3DMA，並且將其 IFR 大於 TH_IFR 的 3DMA 依位置做分群。當樣本逐一分群的同時，群的中心位置同步更新為該群中具最高 IFR 的位置。最後，群裡中心的位置與個數分別代表人物的初始位置與人數。隨機取樣演算法如下：首先，對 $ListSample$ 做亂序的動作，接著循序存取 $ListSample$ 內的 USGP。在存取 $USGP_i$ 時有下列需要判斷：

- (a) 當 $|ListCenter|=0$ 時：需要計算 IFR_i 。
若 $IFR_i \geq TH_IFR$ ，則 $USGP_i$ 為群中心，將 $USGP_i$ 加入 $ListCenter$ 而且 $USGP_i$ 獨自成為一群 (如圖 4-1 圓中心)。否則 $USGP_i$ 不歸群。
- (b) 當 $|ListCenter| \neq 0$ 時： $USGP_i$ 需要逐一與 $ListCenter$ 中所有 $Center_j$ 計算 DIS，其中 j 從 1 到 $|ListCenter|$ 。
 - I. 若 $DIS > TH_RADIUS$ (100 cm)：需要計算 IFR_i 。
若 $IFR_i \geq TH_IFR$ ，則 $USGP_i$ 為群中心，將 $USGP_i$ 加入 $ListCenter$ 而且

$USGP_i$ 獨自成為一群(如圖 4-1 中各圓中心至少相距 100 cm)。否則 $USGP_i$ 不歸群。

II. 若 $DIS \leq TH_HALF$ (50 cm) : 需要計算 IFR_i 。

若 $IFR_i \geq IFR_j$ ，則 $USGP_i$ 取代 $Center_j$ 成為 $Center_j$ 所屬群之中心(如圖 4-2 非圓心的粗點)。否則 $USGP_i$ 歸屬於 $Center_j$ 所在的群。

III. 若 $DIS > TH_HALF$ 而且 $DIS \leq TH_RADIUS$: 不需要計算 IFR_i 。 $USGP_i$ 歸屬於 $Center_j$ 所在的群。

當 $ListSample$ 中所有的樣本經過 (a) 和 (b) 步驟處理後, $ListCenter$ 裡中心的位置與個數分別為人物可能的初始位置與人數。

4.1.3. 初步驗證機制

在第 4.1.2 節所提及到的群裡頭, 所有 3DMA 擁有的 IFR 必定都大於等於 TH_IFR , 但並不一定能代表群中具最高 IFR 的位置即為初始位置, 因為群裡頭的 IFR 都很高。所以我們利用群裡 3DMA 的 IFR 來當權重, 再將群的中位置做權重移位(如圖 4-2 非圓心的粗點與圖 4-3 黃色圓心)。

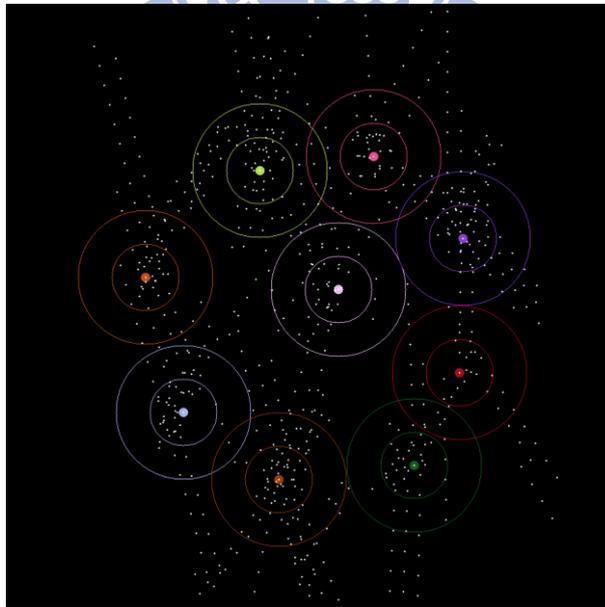


圖 4-1 圓中心為多組 1D-USGP 經隨機取樣演算法分群後後之群中心, 此結果為分成九群

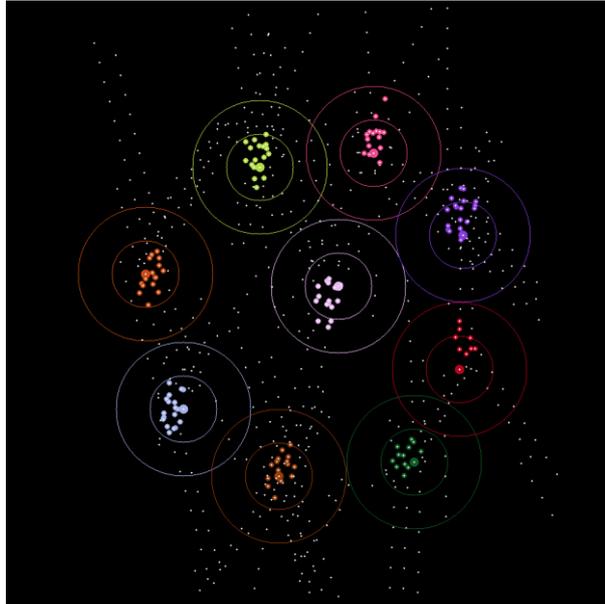


圖 4-2 非圓心的粗點為步驟 b-II. 所提到距離圓心 50 cm 內能取代群中心的 USGP

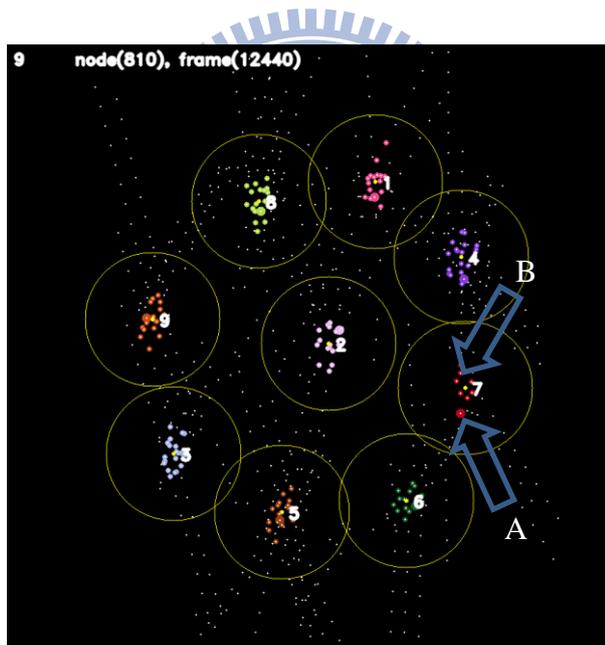


圖 4-3 群 7 (圖 7) 的中心經權重移位後由 A 移至 B 點

4.2. 體積概念與人物過近移除假象人物

第 4.1 節的方法是在限制的樣本空間中決定了初始化的人物位置與人數，但在實際場景中人物分佈的情形，卻並不為限制性的樣本空間。因此本節介紹基於 4.1.3

節之初步人物可能的位置，利用體積概念來做候選人的篩選。

在三維參考平面上，以人物位置為中心（如圖 4-4），周圍半徑為 60 cm 內，每隔 5 cm 來做二維 USGP（圖 4-5 中水藍色的點）並且計算 IFR。假若 IFR_i 大於等於 TH_IFR（圖 4-5 中桃紅色區域），則 IFR_i 累加入體積 i ($Volume_i$)。當 $Volume_i$ 大於等於體積閾值（TH_VOLUME）時，判斷成該位置可能有人物存在，若小於 TH_VOLUME，則該人物為假象人物。當判斷為有人物存在時，以圓內桃紅色區域樣本之 IFR 以權重來重新調整位置（圖 4-6 中圓 3 的中心位置）。在人物高度決定方面，採用桃紅色區域樣本所在位置中，取在何種投影層高度時使得其 IFR 為 1.0，而最高的投影層高度即決定為該人物之身高（如圖 4-7 中黃色數字）。此外，若兩人物位於半徑 60 cm 內，在本論文中則合併為同一人（如圖 4-8）。圖 4-9 為假象人物因 $Volume$ 不足而被移除的例子。

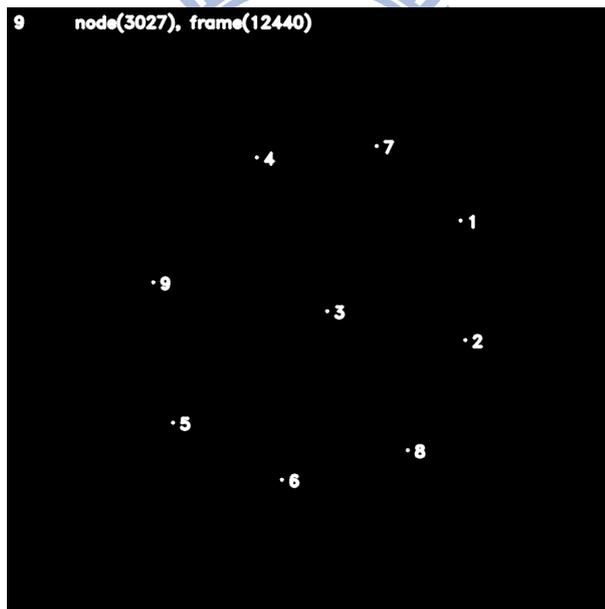


圖 4-4 由第 4-1 節之隨機取樣所得到的人物初步位置

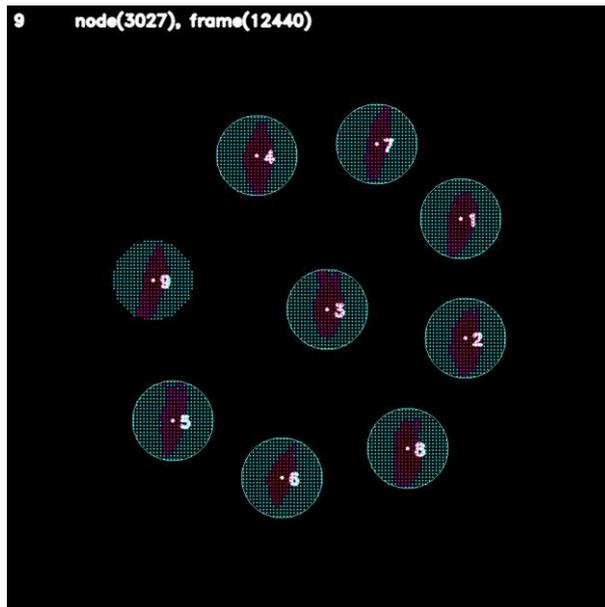


圖 4-5 桃紅色區域為圖 4-4 中以初步位置為中心，周圍 60 cm 所測出來的 IFR 高度圖

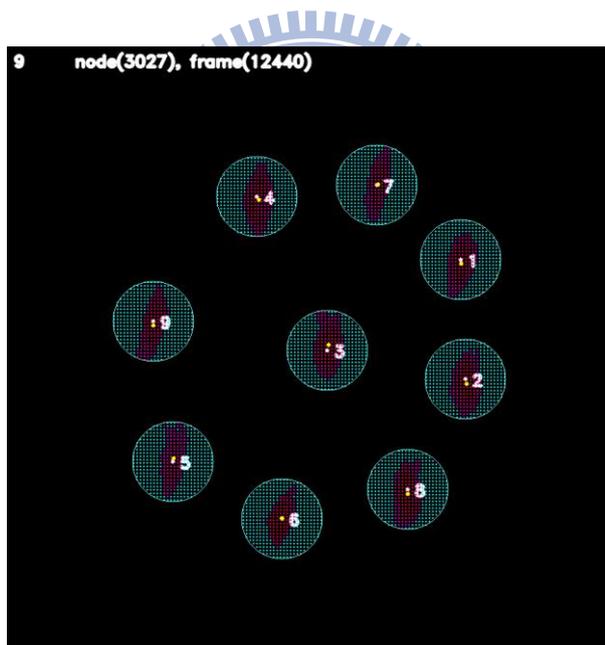


圖 4-6 根據圖 4-5 中的桃紅區域中依 IFR 大小經權重移位後中心位置由白點移至黃點

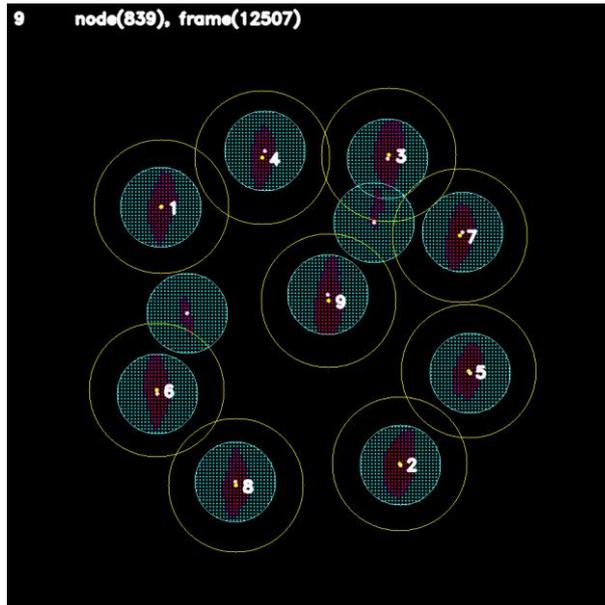


圖 4-9 圖 6 右上方的群因為 *Volume* 不足而被移除

4.3. 重新取樣之範圍

第 4.2 節使用圓形來做 *Volume* 計算，並且可看出在人物佔地區域（桃紅色區域）外有許多的二維 USGP 不為佔地位置，或許會浪費許多計算花費在計算 IFR 於非佔地位置上。圖 4-10 表示場景 1 中 $TH_IFR = 0.80$ 的紮實頂視圖，不難看出若是利用綠色矩形來做 *Volume* 計算，將可節省許多計算 IFR 的花費並且能增加準確度。圖 4-11(a)-(b)分別代表場景 2 及場景 3 之 $TH_IFR = 0.80$ 之紮實頂視圖和矩形範圍。

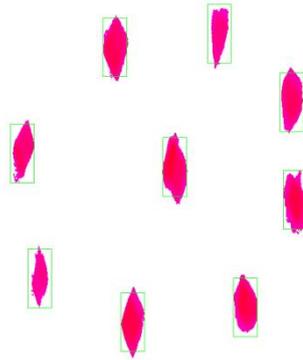


圖 4-10 場景 1 中 $TH_IFR = 0.80$ 之紮實頂視圖

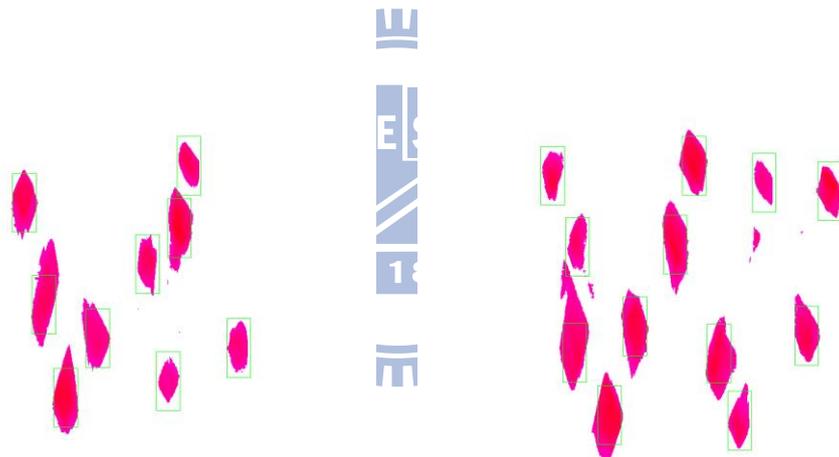


圖 4-11 (a)-(b)分別為場景 2 及場景 3 之紮實頂視圖

場景 1~3 為室內場景，其中場景 1 為八人圍繞中心一人行走、場景 2 為九人隨機行走及場景 3 為 11 人圍繞中心一人行走；場景 4~5 為室外場景（餐廳外），人數約七至九人。場景 1~3 中因為攝影機架設的關係而導致人物在參考平面上的佔地分佈呈現長形狀，因而可使用長形矩形來計算 *Volume*。但由於室外場景 4~5 與場景 1~3 的攝影機架設不同，因而呈現在場景 4 中使用 $TH_IFR = 0.80$ 時，造成人物佔地容易消失再出現的情形（圖 4-12(a)-(d)），因此經由挑選後，針對場景 4~5 使用 $TH_IFR = 0.65$ 則較能擁有足夠的資訊來做人物定位（如圖 4-13(a)-(b)），並且

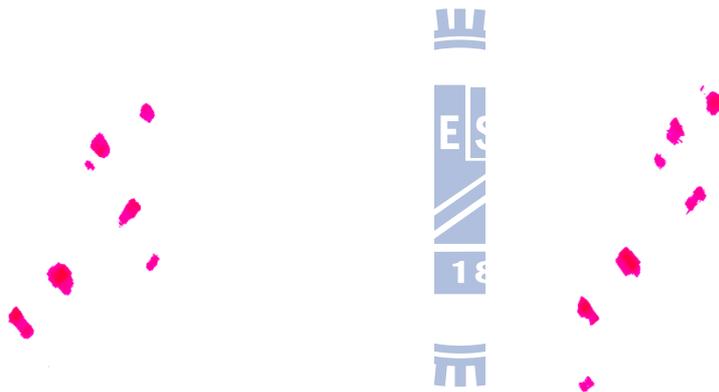
使用與場景 1~3 不同的矩形來計算 *Volume*。



(a) 第 601 幀



(b) 第 611 幀



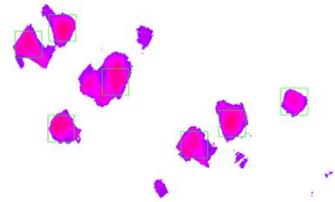
(c) 第 621 幀

(d) 第 631 幀

圖 4-12 (a)-(d)為場景 4 第 601-631 幀中 $TH_IFR = 0.80$ 之紮實頂視圖



(a)



(b)

圖 4-13 (a)-(b) 分別為場景 4 與場景 5 中 $TH_IFR = 0.65$ 之紫實頂視圖



五、實驗結果與討論

本章內容將依照前兩章所提到的取樣與驗證機制方法，利用程式進行實驗並分析。此實驗場景為九人之實驗場景，利用四個攝影機拍攝場景中的九人，其中八人圍繞著中心的人物行走。接下來將討論本論文所提出之隨機取樣方法與驗證機制，雖然在初始人物數量決定上或許不是實際場景中的人數，但是經過體積濾除與人物過近的機制下，依然能夠將場景中的人物準確地定位與估算身高。

5.1. 體積機制濾除人物假象之實驗結果

從第 12519 幀到第 12521 幀中，不難看出在人物的初始數量與位置決定已超過九人（如圖 5-1 左）。在經過體積分析的濾除機制下，能保持九個人物在場景中的結果，在經過體積機制的權重衡量後依然能準確定位（如圖 5-1 右）。

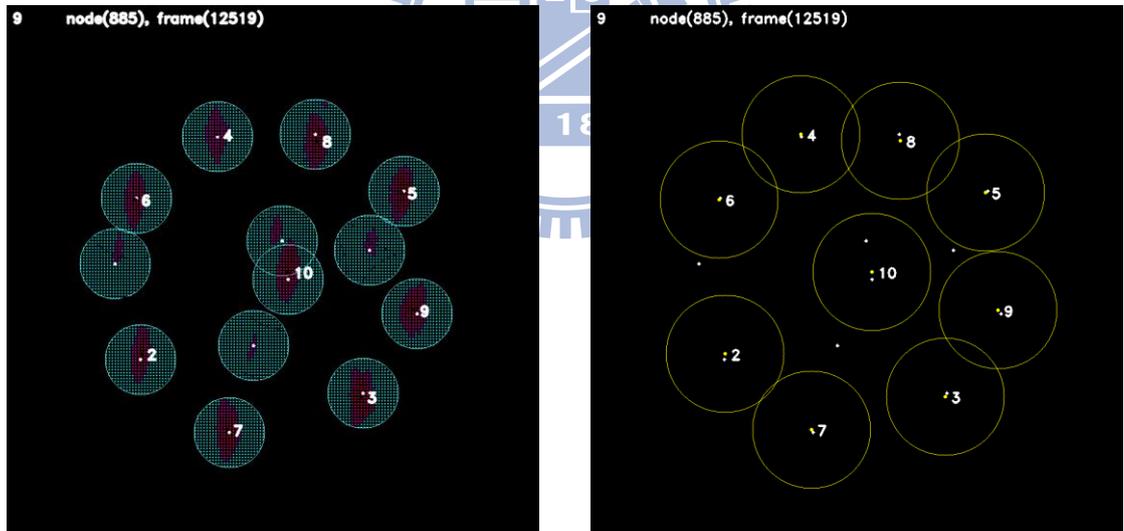


圖 5-1 左圖：第 4.1 節之方法得出之初步人物位置；右圖：經過體積濾除假象人物之結果

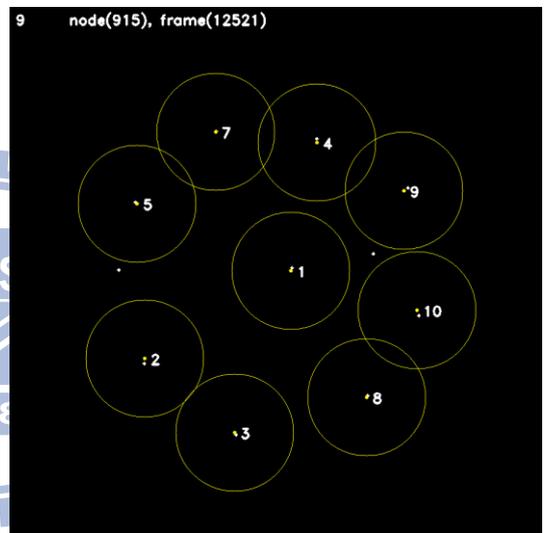
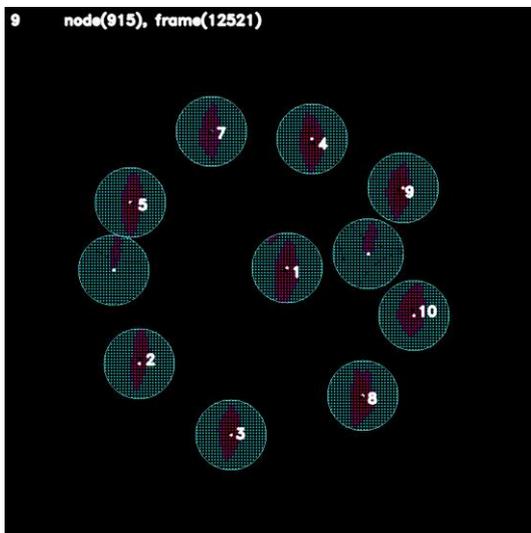
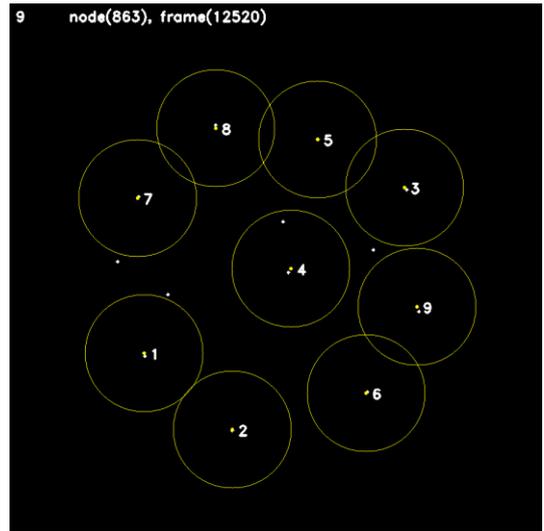
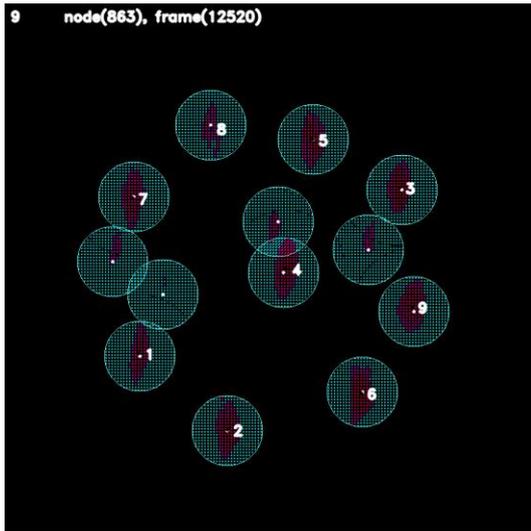
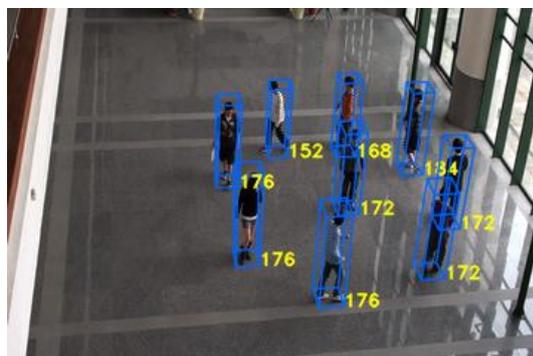
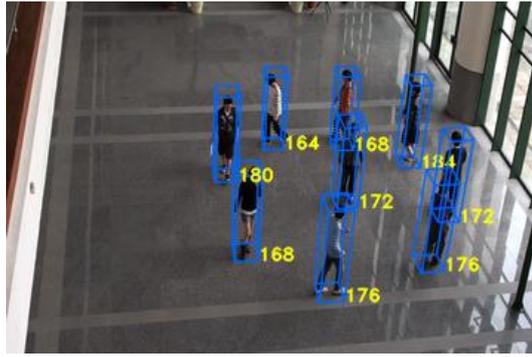


圖 5-1 (續)

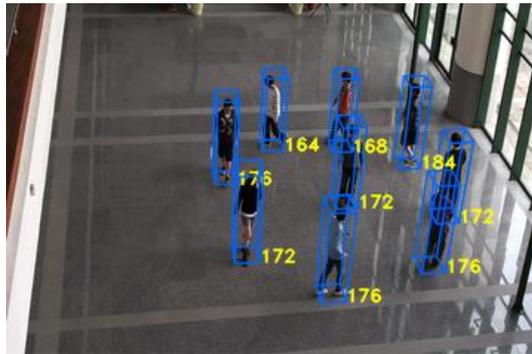


(a) 第 12519 幀

圖 5-2 (a)-(c) 為圖 5-1 之實際定位結果



(b) 第 12520 幀

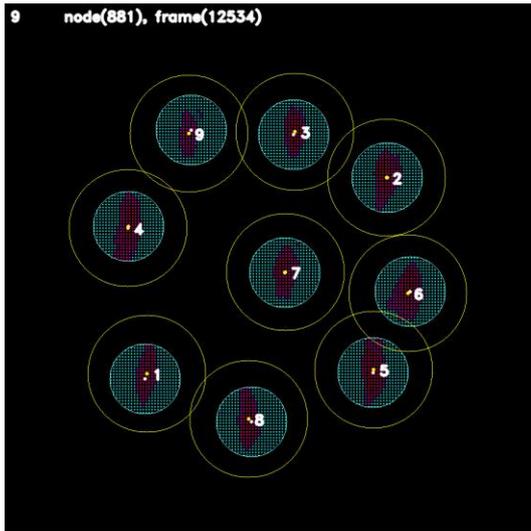


(c) 第 12521 幀

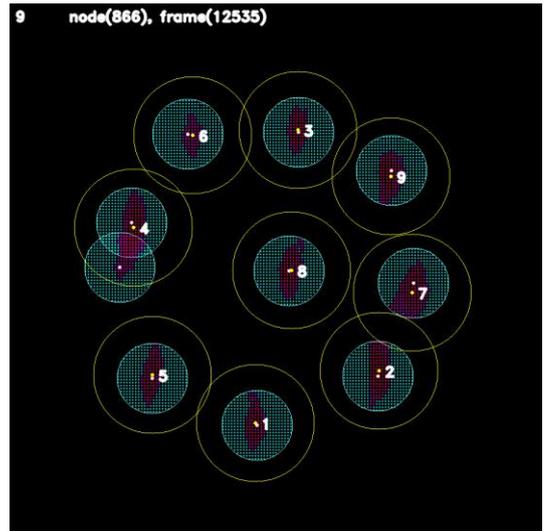
圖 5-2 (續)

5.2. 人物過近機制濾除人物假象

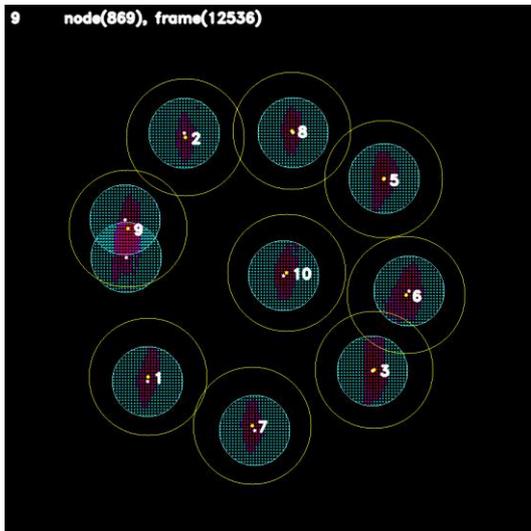
觀察從第 12534 幀（無假象人物，如圖 5-3(a)）到第 12537 幀的結果，其中人物之初始數量雖然在後三個幀中有出現第十個人物(假象)的情形(如圖 5-3(b)-(d))，但第十人透過人物過近機制被濾除掉，並且經過體積的判斷與權重調整位置，可看出在場景中九人之相關位置依然能被準確定位（如圖 5-4(a)-(d)黃圈之結果）。詳細數據請見表格 1。



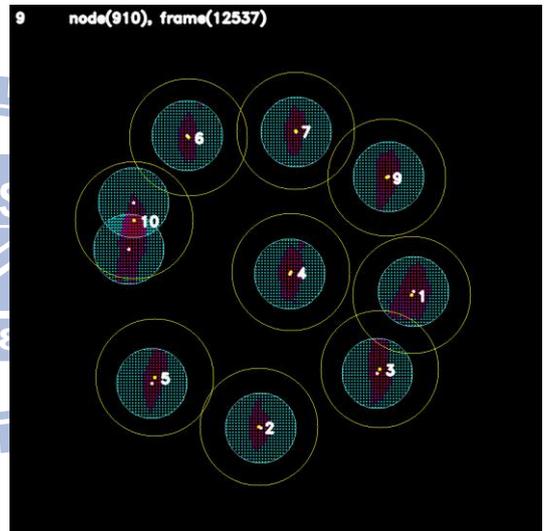
(a)



(b)

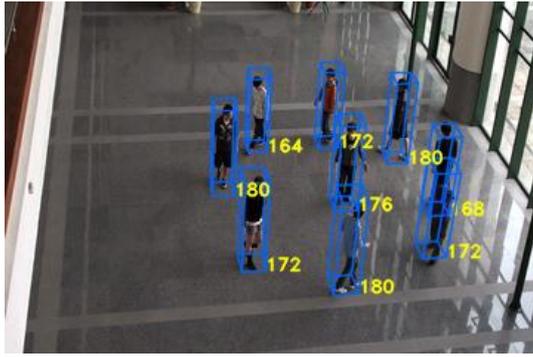


(c)

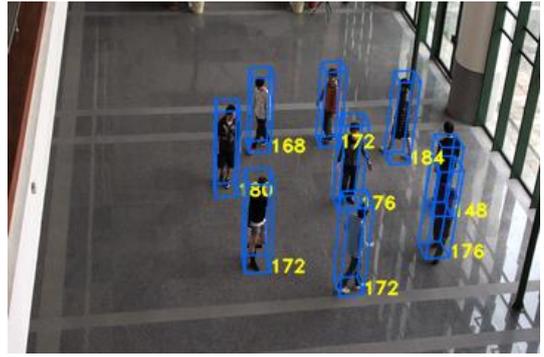


(d)

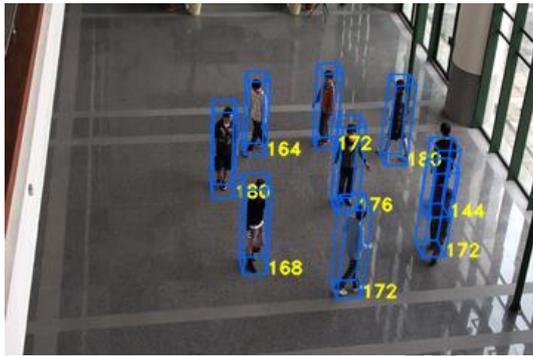
圖 5-3 (a)黃圈為無假象之定位結果。(b)-(d)中黃圈顯示修正過後人物均能被穩定定位



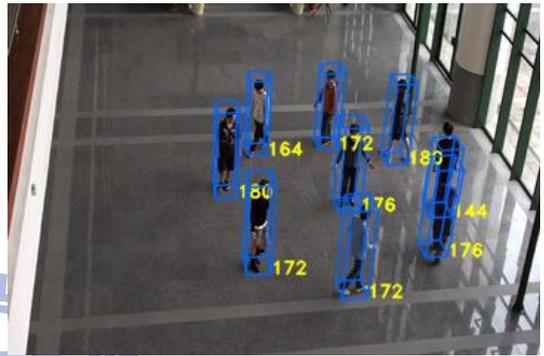
(a)



(b)



(c)



(d)

圖 5-4 (a)-(d)分別為圖 5-3 (a)-(d)之實際定位結果



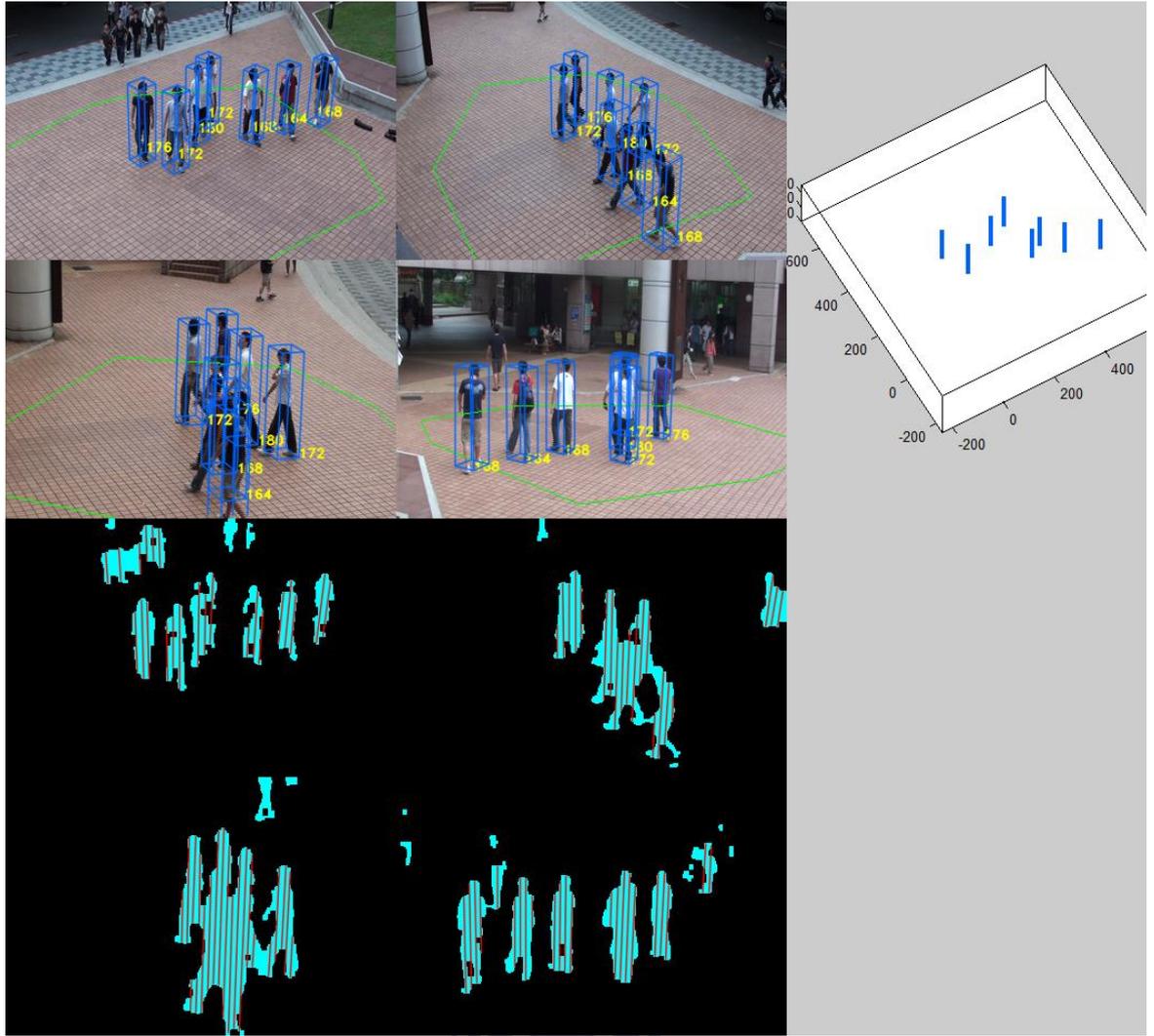


圖 5-5 場景 4 中第 609 幀的結果

表格 1 場景 1 至場景 5 中的正確率與執行速度之量化結果。

	Frames 數量	正確偵測	Miss detection	False alarm	Recall	Precision	Error Avg. (cm)	Error Std. (cm)	FPS Avg.
S1	690	5967	252	274	0.959479	0.956097	11.55673	39.55099	19.24
S2	775	5764	320	408	0.947403	0.933895	11.19735	41.03482	18.31
S3	270	3032	220	345	0.932349	0.897838	11.78517	42.07284	12.31
S4	70	420	51	45	0.89172	0.903226	10.38212	32.62293	10.31
S5	40	234	46	20	0.835714	0.92126	13.08267	61.93023	10.83

六、結論

本篇論文利用人物佔地的概念來進行研究，其中僅使用了前景資訊，在不分析人物特徵點的情況下，除了利用軸線概念外，還使用了（1）人物候選區域之地面均勻取樣（一維 USGP）方式，來減少在投影軸線上的計算成本；（2）在一維 USGP 上隨機取樣產生人物的初始位置與人數；在人物候選的位置取樣上，使用隨機取樣與檢查距離的方式大量地減少計算 IFR 的開銷；（3）體積概念依權重來定位，以獲得穩定的人物定位；經過隨機取樣得到初始位置與人數後，重新區域取樣（二維 USGP）使得體積與人物過近機制能將假象人物妥善地被濾除；（4）體積權重來重新衡量人物位置，能保持人物定位的準確度。本研究有效率地大量降低系統的計算成本，符合即時運算的需求，並且在人物靠近的情形下（大於 60 cm），依然能不失定位的精準度。



參考文獻

- [1] Luo Sun, Huijun Di, Linmi Tao, Guangyou Xu, "A Robust Approach for Person Localization in Multi-Camera Environment, " in *Proceedings of International Conference on Pattern Recognition*, pp. 4036-4039, 2010.
- [2] Murtaza Taj, Andrea Cavallaro, "Multi-View Multi-Object Detection and Tracking," in *Collection of Computer Vision: Detection, Recognition and Reconstruction*, pp. 263-280, 2010.
- [3] Kuo-Hua Lo and Jen-Hui Chuang, "Vanishing Point-Based Line Sampling for Efficient Axis-Based People Localization," *IEEE International Conference on Image Processing*, pp. 11-14, Sep. 2011.
- [4] Jamal Saboune and Robert Laganier, "People Detection and Tracking Using the Explorative Particle Filtering," *IEEE International Conference on Computer Vision*, pp.1298-1305, 2009.
- [5] K. Kim and L. Davis, "Multi-Camera Tracking and Segmentation of Occluded People on Ground Plane Using Search-Guided Particle Filtering," in *Proceeding of European Conference on Computer Vision*, pp. 98-109, 2006.
- [6] Weina Ge and Robert T. Collins, "Crowd Density Analysis with Marked Point Processes," *IEEE Signal Processing Magazine*, pp. 107-111 and pp. 123, Sep. 2010.
- [7] Weina Ge and Robert T. Collins, "Crowd Detection with a Multiview Sampler," *European Conference on Computer Vision*, pp. 324-337, 2010.
- [8] László Havasi and Zoltán Szilávik, "Using Location and Motion Statistics for the Localization of Moving Objects in Multiple Camera Surveillance Videos," *IEEE International Workshop on Visual Surveillance*, pp. 1275-1281, 2009.
- [9] S.M. Khan and M. Shah, "A Multiview Approach to Tracking People in Crowded Scenes Using a Planar Homography Constraint," in *Proceeding of European*

Conference on Computer Vision, pp.133-146, 2006.

- [10] Anastasios L. Kesidis, Dimitrios I. Kosmopoulos, "Robust Occlusion Handling with Multiple Cameras Using a Homography Constraint," in *Proceedings of International Conference on Computer Vision Theory and Applications*, pp. 560-565, 2009.
- [11] Santos, T. T. and Morimoto, C. H., "Multiple Camera People Detection and Tracking Using Support Integration," *Pattern Recognition Letters*, vol. 32, no. 1, pp. 47-55, 2011.
- [12] R. Eshel and Y. Moses, "Homography Based Multiple Camera Detection and Tracking of People in a Dense Crowd," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1-8, 2008.
- [13] Weiming Hu, Min Hu, Xue Zhou, Tieniu Tan, Jianguang Lou and Steve Maybank, "Principal Axis-Based Correspondence between Multiple Cameras for People Tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 663-671, Apr. 2006.
- [14] Horng-Horng Lin, Jen-Hui Chuang and Tyng-Luh Liu, " Regularized Background Adaptation: A Novel Learning Rate Control Scheme for Gaussian Mixture Modeling," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 822-836, Mar. 2011.