

國立交通大學

多媒體工程研究所

碩士論文

多攝影機資訊整合之人員定位與誤差分析

People Localization and Error Analysis based on Integration
of Multi-Camera Information

研 究 生：郭育瑋

指導教授：莊仁輝 教授

中 華 民 國 九 十 九 年 六 月

多攝影機資訊整合之人員定位與誤差分析

People Localization and Error Analysis based on Integration of
Multi-Camera Information

研 究 生：郭育瑋

Student：Yu-Wei Guo

指導教授：莊仁輝

Advisor：Jen-Hui Chuang

國 立 交 通 大 學

多 媒 體 工 程 研 究 所

碩 士 論 文



A Thesis

Submitted to Institute of Multimedia Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

多攝影機資訊整合之人員定位與誤差分析

學生：郭育璋

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



在本論文中，我們實作一套基於電腦視覺的定位系統，該系統運用到平面投影轉換關係和三維幾何關係，達成對影像中的目標物進行定位。然而實際定位的過程中仍將存在著某些誤差，例如徑向畸變誤差、人為測量誤差和取像誤差，這些誤差將會使得定位的準確度下降。因此我們對於定位過程中可能產生的誤差加以分析與改善，減少這些誤差對於定位的影響，我們所採取的方法包含對影像作校正和微調測量的位置。另一方面，我們可以利用分析的結果進一步地由多攝影機中挑選出，具有較低定位誤差之攝影機配對，以獲得較佳的定位結果。由虛擬場景和實際場景的實驗結果可看出，利用誤差分析的結果能使我們成功地整合多攝影機資訊，並取得較為準確、穩定之定位結果。

People Localization and Error Analysis based on Integration of Multi-Camera Information

Student : Yu-Wei Guo Advisor : Dr. Jen-Hui Chuang

Institute of Multimedia Engineering

National Chiao Tung University



Abstract

In this thesis, we implement a vision-based localization system that can locate objects in the scene from their images by homography and 3D geometry. In reality, the object localization may not be always accurate due to errors arising from radial distortion of cameras, noises in the imaging process, and errors associated with manually measurement, etc. We analyze these errors and try various ways, including simple camera calibration and fine-tuning the measurement data, to reduce the degradation in localization accuracy. On the other hand, according to these analytic results, the system can suggest an appropriate pair of cameras in a multi-camera environment to achieve more accurate localization. Both synthetic and real scene data are used to verify the implemented localization system. Experimental results show that the proposed approach can indeed reduce the localization error and improve system stability.

致謝

本論文得以順利完成，首先要感謝的是我的指導教授莊仁輝老師。由於老師的耐心和熱心的教導，以及嚴謹的研究態度，讓我在研究方面可以事半功倍，也讓我對研究領域更加了解。除了課業之外，平時老師也關心我的生活，使我在生活上的態度能有所啟發和成長。另外，也要感謝三位口試委員，賴飛熊教授、顏嗣鈞教授和王聖智教授的指導與建議，讓本論文的內容可以更加完整與充實，在此衷心的感謝他們。

接著我要感謝實驗室的所有伙伴。感謝 FB、胖虎和胖嘟嘟，一起熬夜趕虛弱報告，一起熬夜看日出，一起努力寫論文和修改論文，一起拼畢業，一起吃美食，一起運動瘦身(雖然越運動越胖)，一起熬夜玩 Wii(只為了訂回家車票)，好多好多好多的一起，有了你們的陪伴與扶持，我才能走過來。感謝國華學長、泓宏學長和星陸在課業及研究上不吝嗇地指導，讓我在研究上可以更上一層樓。感謝小屁股、槓柑、小馬、周傑與小貝，由於你們帶給大家歡樂，讓實驗室的氣氛一起十分和樂。還有大家一起去台南高雄之旅和青草湖夜市爭霸戰，更是讓我留下難忘和美好的回憶，有你們真好。

也謝謝從高中一直以來的好友，阿達、大泡、阿今和瓜仔等，再謝謝打工所認識的黃小姐、淑華、銘佩和趴趴，謝謝耀萱在業務和口試時的大力相助，感謝你們的關心和協助，所有的事情都點滴在心頭，認識你們真的很開心。

最後，感謝一直以來支持我和為我操心的家人，有了他們讓我可以沒有後顧之憂的取得碩士學位。也要感謝所有幫助過我的朋友們，僅以此篇表達我誠摯的謝意，どうもありがとう，thank you sincerely。要謝的人那麼多，那就謝天吧，thanks god。

目錄

中文摘要	i
英文摘要	ii
致謝	iii
目錄	iv
圖目錄	v
表目錄	vii
第一章 緒論	1
1.1 研究動機.....	1
1.2 相關研究.....	2
1.2.1 利用單攝影機追蹤人物	2
1.2.2 利用多攝影機追蹤人物	3
1.3 系統流程.....	8
1.4 各章簡介.....	9
第二章 基於單應性轉換之定位方法	10
2.1 單應性轉換矩陣	10
2.2 定位方法.....	13
2.3 初步定位結果	14
第三章 誤差來源之探討與改善	17
3.1 誤差的來源	17
3.2 靜態誤差.....	18
3.2.1 徑向畸變誤差	18
3.2.2 人為測量誤差	22
3.3 動態誤差.....	24
第四章 整合多攝影機之定位結果	28
4.1 整合一對攝影機的定位方法	28
4.2 暴力法估算定位誤差分佈	29
4.3 簡化法估算定位誤差分佈	30
第五章 實驗結果	32
5.1 虛擬場景的定位結果	32
5.2 真實場景的定位結果	34
第六章 結論與未來展望	47
6.1 結論.....	47
6.2 未來展望.....	47

圖目錄

圖 1-1 兩主軸之間的關係。	4
圖 1-2 立足點的二元影像。第一列為前三個視野的前景圖，第二列為 View1、View2 和 View3 投影到參考視野 View4 的影像，由第二列的影像得到第三列左邊的協同增益圖，對協同增益圖取門檻值得到第三列右邊的二元影像。	5
圖 1-3 立足點的軌跡圖。(a)每張影格找到的二元位置圖，(b)二元位置圖依圖形分割演算法，即可得到軌跡圖，每種顏色分別代表一個人物的軌跡。	6
圖 1-4 將右上角的立足點(紅色、藍色和綠色點)投影到左上角的視野中，可看出其中一個藍色矩形中存在兩個立足點，代表遮蔽發生。	7
圖 1-5 高度平面為 178 公分的頭部候選區域。(a), (b)兩台攝影機的影像，(c)影像(b)投影到高度為 178 公分的平面，(d)影像(c)疊到影像(a)上，(e)其他攝影機影像經由投影後，疊到影像(a)上，(f)影像(e)的變異量，紅色代表變異量低。	8
圖 1-6 系統流程圖。	9
圖 2-1 場景中參考點的配置。(a) 影像中的參考點，(b) 真實空間中的參考點。	12
圖 2-2 單攝影機的定位示意圖。	13
圖 2-3 實驗場景 1。(a)場景頂視圖，(b)場景影像。	15
圖 2-4 實驗場景 1 之定位結果。	16
圖 3-1 受到徑向畸變影響的影像。	18
圖 3-2 徑向畸變校正示意圖。	19
圖 3-3 影像座標正規化。(a)影像座標(u, v)，(b)置中且正規化影像座標(P_u, P_v)。	19
圖 3-4 影像校正時，座標調整示意圖。	19
圖 3-5 影像校正。(a) 校正前，(b) 校正後。	20
圖 3-5 (續)。	21
圖 3-6 實驗場景 1 校正後之定位結果。	21
圖 3-7 微調後真實空間中的參考點位置。	23
圖 3-8 實驗場景 1 校正和微調參考點後之定位結果。	23
圖 3-9 模擬取像誤差。(a) 影像中加入 100 個高斯雜訊，(b) 經由單應性轉換矩陣作用後的 100 個雜訊。	24
圖 3-10 模擬影像誤差。(a) 影像上的邊界圓，(b) 經由單應性轉換矩陣作用後的邊界圓。	25
圖 3-11 圓形誤差分佈的定位誤差橢圓。	25
圖 4-1 整合一對攝影機的定位結果。	29
圖 4-2 整合多攝影機間的定位誤差分佈。	30
圖 4-3 簡化法估算誤差分佈示意圖。	31

圖 5-1 虛擬場景。(a)場景頂視圖，(b)場景影像。	33
圖 5-2 實際場景—雷射點為目標點。(a) 攝影機 1 的影像，(b) 攝影機 2 的影像，(c) 攝影機 3 的影像，(d) 攝影機 4 的影像。	35
圖 5-2 (續)。	36
圖 5-3 單攝影機的定位結果。(a) 攝影機 1 的結果，(b) 攝影機 2 的結果，(c) 攝影機 3 的結果，(d) 攝影機 4 的結果，(e) 挑選過後的結果。	36
圖 5-3 (續)。	37
圖 5-4 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。	38
圖 5-4 (續)。	39
圖 5-4 (續)。	40
圖 5-5 單多攝影機一起挑選的結果。	40
圖 5-6 實際場景—人物。	41
圖 5-7 單攝影機的定位結果。(a) 攝影機 1 的結果，(b) 攝影機 2 的結果，(c) 攝影機 3 的結果，(d) 攝影機 4 的結果，(e) 挑選過後的結果。	42
圖 5-7 (續)。	43
圖 5-8 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。	43
圖 5-8 (續)。	44
圖 5-8 (續)。	45
圖 5-9 單多攝影機一起挑選的結果。	46

表目錄

表 5.1 單攝影機的定位誤差。	33
表 5.2 整合多攝影機的定位誤差。	34



第一章 緒論

1.1 研究動機

在人們的生活週遭中隨處都可以看到攝影機，這些攝影機常被用於視覺監控。隨著安裝場所的不同，視覺監控可應用的範圍相當廣泛[1]，例如重要建築和社區的安全守衛、特定區域的進出管制和身份認定、人潮的統計與分析、異常偵測和警報等，這也是為何視覺監控引起越來越多人注意的原因。

一般而言，視覺監控系統主要包含四個部份：建立環境的模型、偵測運動物體、分類移動物體和追蹤物體，其中我們感興趣的部份在於追蹤物體方面。因為在電腦視覺領域中，以攝影機為基礎做物體追蹤與定位一直是個典型且重要的研究議題。此研究議題之所以典型在於物體追蹤與定位技術，在簡單的環境設定中，已有許多效能良好的技術被提出，例如靜態而簡單的背景、單一追蹤目標物、使用已校正的攝影機等。而此研究議題之所以重要在於一旦攝影機拍攝的是真實且複雜的場景，則諸多的變因，像是光線變化、人群遮蔽、場景變動快速等，將使得物體追蹤與定位問題變得複雜而難解。因此近年來視覺監控追蹤與定位的研究發展，多朝向將原本適用於簡單場景的技術，逐步擴展應用到真實而複雜環境中。

在追蹤的過程中常常會需要利用位置資訊，所以準確的定位可以幫助追蹤。因此在本篇論文中，我們實作一套基於電腦視覺的定位系統，該系統利用平面投影轉換關係和三維幾何關係，達成對影像中的目標物進行定位。然而定位的過程中仍存在著某些誤差，例如徑向畸變誤差、人為測量誤差和取像誤差，這些誤差將會使得定位的準確度下降。因此我們對於定位過程中可能發生的誤差加以分析與改善，並減少這些誤差對於定位的影響。接著，我們利用分析的結果進一步地由多攝影機中挑選出，具有較低定位誤差之攝影機配對，以獲得較佳的定位結果。

1.2 相關研究

近年來有許多的視訊追蹤方法被提出來，依據目標物的種類，大致可分成追蹤人物、動物和車輛。我們考慮以人物為主的情況下，依攝影機的多寡分成單攝影機的追蹤與多攝影機的追蹤。以單攝影機為主的追蹤方法可能會遇到遮蔽的問題，尤其在場景人物眾多的情況下。因此藉由使用多攝影機為主的追蹤方法，除了可能解決遮蔽的問題，更重要的是使用多攝影機可以取得更多的場景資訊，以便有機會得到更準確的追蹤和定位結果。在此章節中，我們將介紹幾篇與追蹤方法相關的研究論文。

1.2.1 利用單攝影機追蹤人物

文獻[2]提出一個以單攝影機為基礎的即時人物追蹤系統，它利用多種特徵的統計模型來偵測和追蹤人物。首先參考多種特徵利用單高斯模型建立背景模型，接著依據最大可能出現的機率來偵測和追蹤物體。此系統雖然可以即時作用，但是它假設背景中只存在一個物體，在多個物體的情況下可能無法得到良好的結果。文獻[3]提出一個適用於擁擠場合的追蹤演算法。在學習的過程中，建立影像間的親和模型(affinity model)，其可判定任兩個部分的軌跡(tracklet)是否屬於同一個目標物，然後參考親和模型便可將部份的軌跡串聯起來得到完整的軌跡。

文獻[4]結合空間和時間的資訊，並利用馬可夫鏈蒙地卡羅(Markov chain Monte Carlo)之方法找出目標物之間的對應關係，依據該對應關係即可得到追蹤的結果。通常解決對應性的問題會假設一個前景區域只對應到一個目標物，即一對一對應。在遮蔽發生的情況下，一個前景區域不一定保證只存在一個目標物，因此一對一對應不成立。然而此系統沒有假設一對一對應，而是使用馬可夫鏈蒙地卡羅的方法，因此即使遮蔽的情況發生，仍然可以得到正確的追蹤結果。文獻[5]提出基於隱馬可夫模型(hidden Markov model)的演算法追蹤多個物體。首先建立一個觀察模型，其包含原始影像、前景遮罩和偵測物體的資訊。原始影像可提供目標物的資訊，前景遮罩是由高斯混合模型產生，偵測物體的資訊則由物體偵

測器得到，此資訊記錄著像素間(pixel-wise)的物體偵測分數。接著使用基於隱馬可夫模型的方法估算出狀態序列和觀察序列的聯合機率最大值，此最大值對應的狀態序列即包含軌跡資訊。

雖然已經有許多效能良好的演算法被提出來，但是這些方法都只適用於單攝影機的場景，一旦發生長時間遮蔽的情況，基於單攝影機的演算法可能無法得到良好的追蹤效果。所以接下來我們將探討以多攝影機為主的方法。

1.2.2 利用多攝影機追蹤人物

利用多攝影機的監控系統，可拍攝到較大範圍且來自不同視野的資訊，在許多情況下相當的有幫助。例如在多攝影機的場景中，若有一台攝影機的視野，看到的人物被場景中的其他人遮蔽，此時有可能存在某一台攝影機，其視野中的此人物不被其他人遮蔽，因此後者的資訊比前者的資訊有用。將多台攝影機所拍攝到的畫面進行分析與整合，可提升偵測和追蹤的穩定度。然而使用多攝影機的監控系統額外衍生出一些問題，像是在場景中如何擺放多攝影機[6]、多攝影機的校正問題[7, 8, 9, 10]、解決多攝影機之間目標物的對應關係[11]等，其中文獻[9]提出校正大範圍多攝影機的方法。他們在多個位置之間移動 LED，藉此獲得多個視野之間的對應性，以減少算錯對應性所帶來的誤差。他們的方法可以計算出投影矩陣和徑向畸變的參數，其中計算參數時不需用到非線性極小化(non-linear minimization)。2 維物體的攝影機校正[10]使用已知的平面樣本而不是用複雜的 3 維物體作校正，使得校正的過程更佳具有彈性。使用者需要將已知的平面樣本貼在板子上，藉由在攝影機前多次(至少兩次)移動板子以校正攝影機。這項技術背後的觀念是利用特徵點在平面樣本上和影像上之間的單應性轉換以算出攝影機的參數。即使是簡單的準備過程，依然可以藉由 2 維物體的校正得到足夠準確的校正結果。接著我們介紹一些以多攝影機為主的追蹤方法。

文獻[12]利用非參數高斯模型(non-parameter Gaussian model)建立色彩模型並切割出前景區域。當遮蔽發生時，用貝氏分類器(Bayesian classifier)依機率判

斷遮蔽區域中屬於各個人物的像素。緊接著結合對極幾何(epipolar geometry)，將前景區域投影到地平面上，找出人物的位置，並同時建立可能性圖(likelihood map)，進而利用卡爾曼濾波器(Kalman filter)進行追蹤。雖然使用顏色資訊找出不同攝影機之間的對應關係是一個直覺的方法，但是若存在多個人物穿著一樣顏色的衣服，則無法正確找出對應關係。

文獻[13]結合主軸(principal axis)和單應性轉換關係(homography)進行人物追蹤。首先找出某攝影機中代表前景人物的主軸 L_s^i ，藉由單應性矩陣將此 L_s^i 轉換到另一個攝影機的畫面中形成 L_s^{ij} ， L_s^{ij} 與該畫面中的某一前景人物的主軸 L_k^j 交出一點 P ，如圖 1-1 所示。從前景外接矩形(bounding box)可得到地上點 X_s^i 和 X_k^j ，即前景主軸與外接矩形的最低交點，當遮蔽發生時，此地上點用預測的方式估計出來。然後將交點 P 與地上點 X_k^j 作比較，即可用距離判斷出這兩個人物是否相互對應，若是的話就更新地上點。最後將找出來的地上點套到卡爾曼濾波器進行追蹤。文獻[14]則是利用主軸關係和粒子濾波器(particle filtering)進行追蹤。

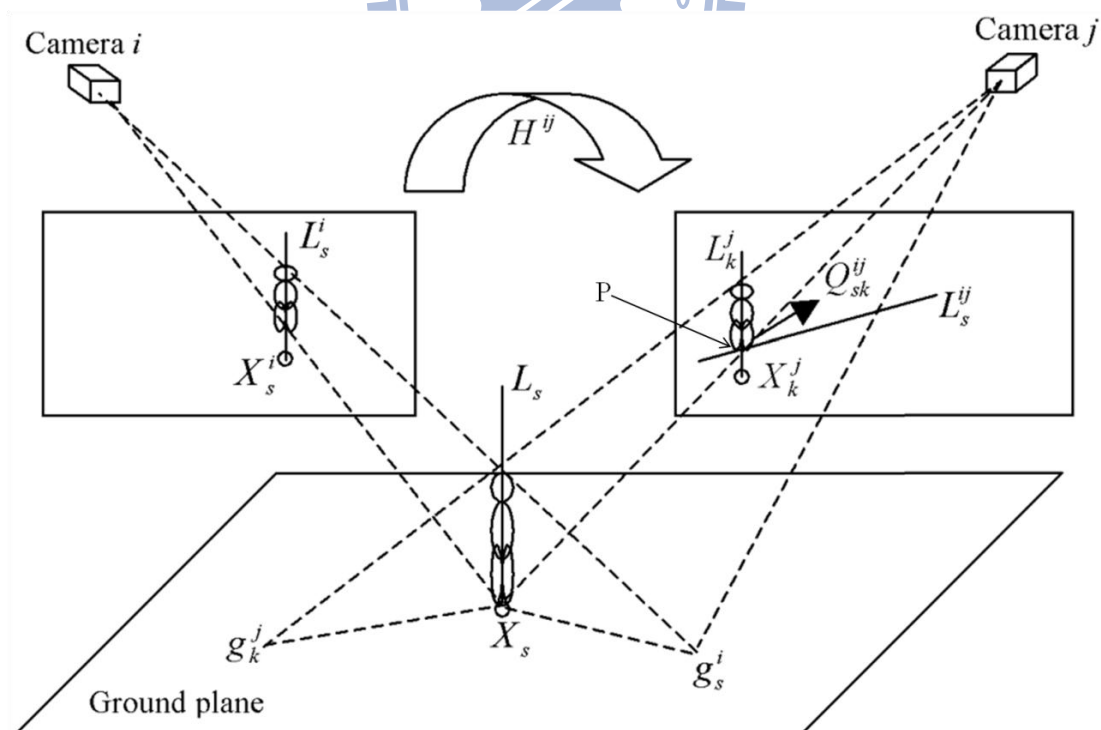


圖 1-1 兩主軸之間的關係。

文獻[15]利用混合高斯模型得到各個攝影機視野的前景圖，再將找到的前景圖經由單應性轉換矩陣投影到某一參考攝影機視野中，並累乘所有的結果得到一個協同增益圖(synergy map)，最後對協同增益圖取門值(threshold)，可得到代表一個人的立足點位置，如圖 1-2 所示，右下角的圖中之每一個白色點的中心，即為一個人的立足點位置。將每張影格得到的位置二元影像依圖形分割演算法(graph-cut algorithm)，即可得到立足點的軌跡圖，如圖 1-3 所示。

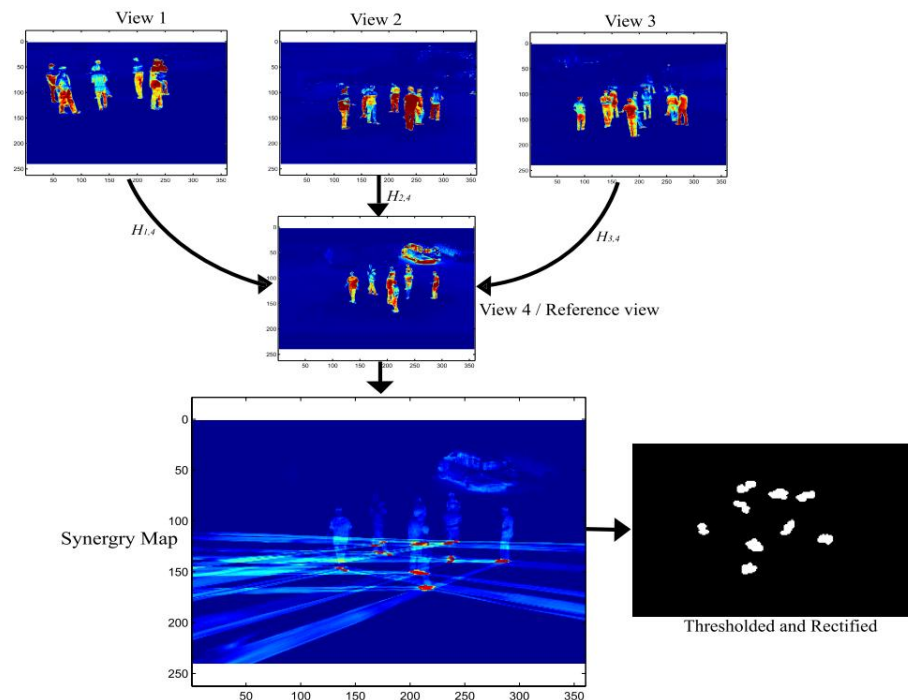


圖 1-2 立足點的二元影像。第一列為前三個視野的前景圖，第二列為 View1、View2 和 View3 投影到參考視野 View4 的影像，由第二列的影像得到第三列左邊的協同增益圖，對協同增益圖取門檻值得到第三列右邊的二元影像。

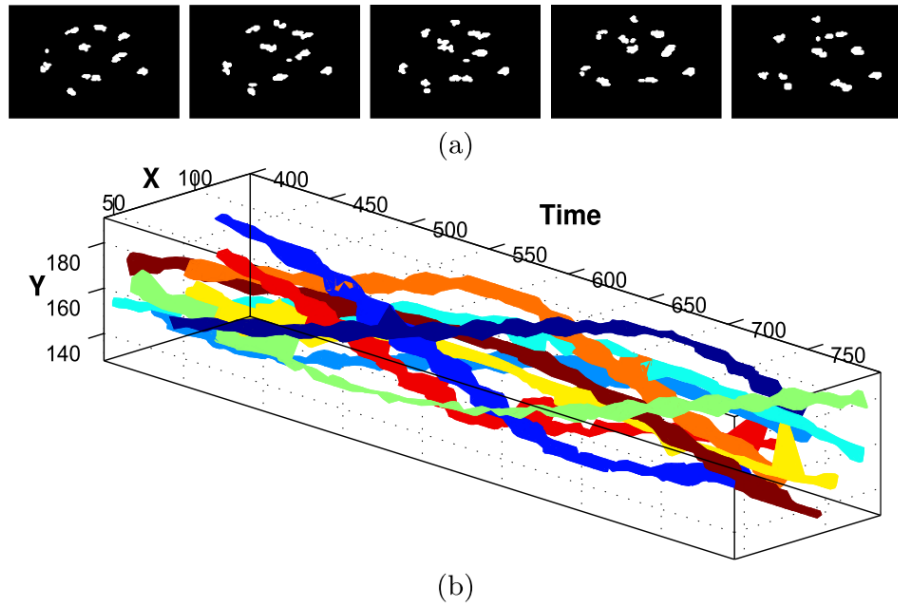


圖 1-3 立足點的軌跡圖。(a)每張影格找到的二元位置圖，(b)二元位置圖依圖形分割演算法，即可得到軌跡圖，每種顏色分別代表一個人物的軌跡。

文獻[16]提出一個基於物體位移的機率之攝影機視野切換技術(view-hopping technique)以達到系統自動選取較適合的攝影機視野觀測物體。首先使用簡單的高斯模型找出前景和其立足點位置，此立足點為切割出之前景外接矩形的最低點。由於人物的正面具有較多的資訊，此系統利用影像中 Y 軸的變化算出人物面向和背向攝影機的機率，即人物靠近時 Y 軸的變化為正值，離開時 Y 軸的變化為負值。一旦發生遮蔽時，即可利用此機率值選擇合適的攝影機視野作切換。至於遮蔽偵測，則是利用找出的前景外接矩形，將矩形中的影像立足點，利用單應性轉換矩陣投影到另外一個攝影機視野中，一旦一個外接矩形中存在二個以上的影像立足點，即代表發生遮蔽情形，如圖 1-4 所示。

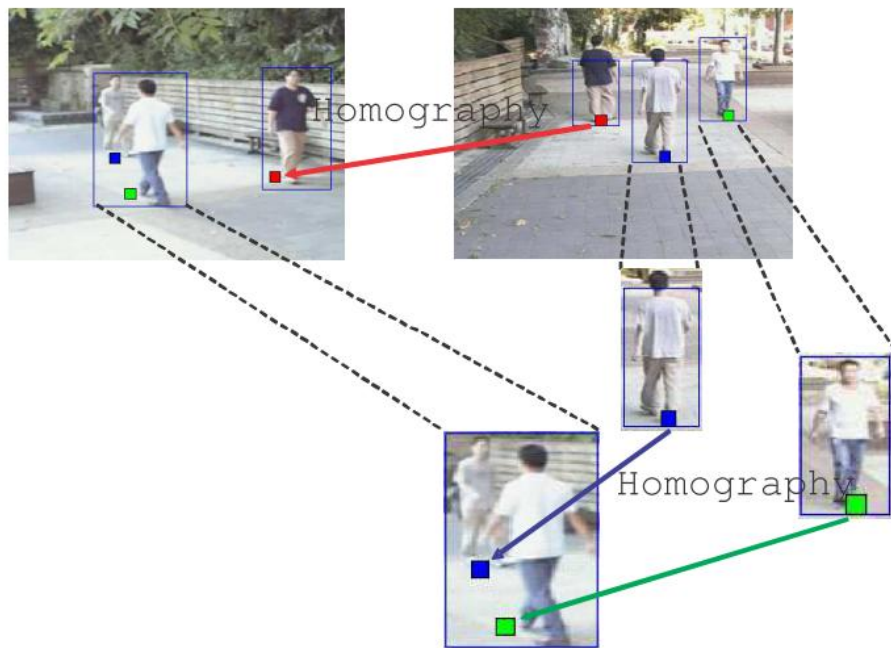


圖 1-4 將右上角的立足點(紅色、藍色和綠色點)投影到左上角的視野中，可看出其中一個藍色矩形中存在兩個立足點，代表遮蔽發生。

不同於前面提到找立足點的方法，文獻[17]提出以頭部為主的追蹤方法，因為頭部的遮蔽現象較少發生。首先使用簡單的背景模型找出前景區域，再利用不同高度間的單應性轉換矩陣將前景區域投影到對應的高度平面。以某一高度平面為例，將所有攝影機投影過後的結果疊到原始影像上，利用強度值找出變異量較小的地方即為此高度的頭部候選區域，如圖 1-5 所示。將不同高度的頭部候選區域投影到地平面即可找到該人物的位置，有了位置資訊便可拿來追蹤。追蹤的過程可分成二個步驟，先是利用運動方向和速度預測此物體在下一個時間點的位置，經由此步驟，我們可以得到許多軌跡，其中包含正確的軌跡或錯誤的軌跡。再來利用可能性關係，結合之前得到的所有軌跡，即可得到正確的完整軌跡。

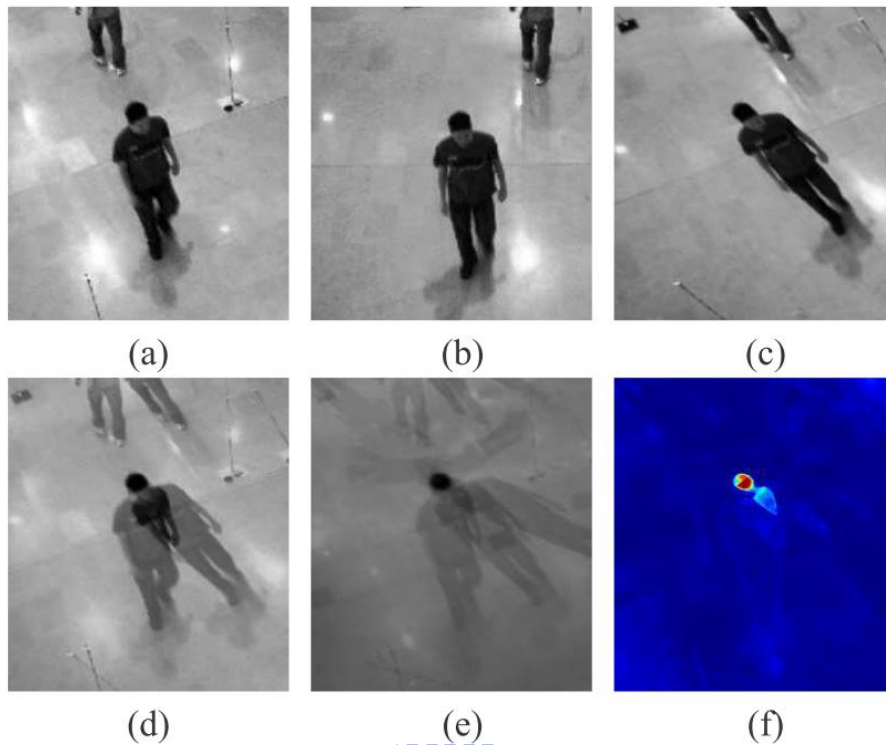


圖 1-5 高度平面為 178 公分的頭部候選區域。(a), (b)兩台攝影機的影像，(c)影像(b)投影到高度為 178 公分的平面，(d)影像(c)疊到影像(a)上，(e)其他攝影機影像經由投影後，疊到影像(a)上，(f)影像(e)的變異量，紅色代表變異量低。

1.3 系統流程

本論文中所提出的定位方法是以單應性轉換關係和三維幾何關係為基礎，我們以一個點當作定位的目標點，像是影像立足點或是影像頭部中心點，希望能找出目標點在真實空間中的位置。由於立足點常常受到遮蔽的影響，加上頭部中心具有不變性，即從攝影機從不同的角度拍攝，影像中頭部中心的位置皆相去不遠，因此可拿頭部中心當作我們的目標點。系統流程圖如圖 1-6 所示，首先我們在取得的影像中找出目標點，接著利用單應性轉換關係和三維幾何關係，估算出目標點在真實世界中的位置。由於系統會有一些誤差產生，因此接下來我們將分析可能造成定位誤差的誤差來源。最後我們將利用先前分析的結果，整合多攝影機的定位結果，期望得到較小的定位誤差。

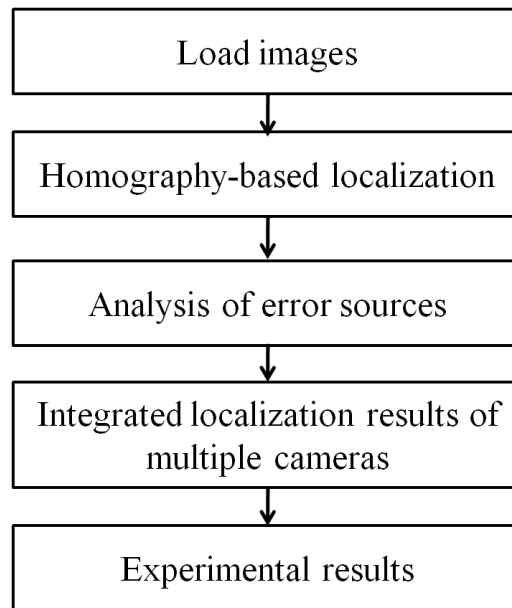


圖 1-6 系統流程圖。

1.4 各章簡介

本論文總共分為六章，除了本章為緒論外，第二章將介紹我們所使用的定位方法，並給初步定位結果。由於定位的結果仍存在明顯的誤差，因此在第三章中我們將探討造成定位誤差的原因，並設法加以分析與改善。第四章將介紹如何利用第二章提及的定位方法，整合多攝影機的定位，期望可以藉由第三章分析誤差的結果，獲得到較佳的定位結果。在第五章中則是我們的實驗結果，最後一章則為本論文作個總結與本研究可應用的方向。

第二章 基於單應性轉換之定位方法

在本論文中，我們使用不同平面間投影轉換的觀念，以單應性轉(homography)為基礎，當取得影像中的目標點後，透過單應性轉換矩陣將目標點轉換到真實空間中的參考平面上，以獲得目標點在該平面上的定位結果。由於在定位的過程中，我們使用了單應性轉換矩陣，因此在 2.1 節中，我們對單應性轉換矩陣的數學理論作簡單的介紹，接著描述如何利用參考點以計算單應性轉換矩陣；算出單應性轉換矩陣後，在 2.2 節中，我們介紹如何利用單應性轉換矩陣以估算在單攝影機情況下之真實空間中參考平面上的定位結果；在 2.3 節中，我們將估算出場景中的定位結果。

2.1 單應性轉換矩陣

單應性轉換的相關數學理論

假設具有兩個平面 π_A 和 π_B ，若在兩平面上找到多個對應的參考點，則兩平面間存在一個 3×3 的單應性轉換矩陣 H ，可將 π_A 平面上的點 P_A ，經由透視投影轉換到另一 π_B 平面上的點 P_B 。若以齊次座標表示法表示，則 $P_A = [x_A, y_A, 1]^T$ 、 $P_B = [x_B, y_B, 1]^T$ ， H 、 P_A 和 P_B 的轉換關係如式(2-1)。我們用外積形式改寫式(2-1)，結果如式(2-2)。用 H^j 表示 H 的第 j 列，對式(2-2)作展開，推導出式(2-3)。

$$HP_A = P_B \quad (2-1)$$

$$P_B \times HP_A = 0, H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (2-2)$$

$$P_B \times HP_A = \begin{bmatrix} y_B h^3 P_A & - & h^2 P_A \\ h^1 P_A & - & x_B h^3 P_A \\ x_B h^2 P_A & - & y_B h^3 P_A \end{bmatrix} \quad (2-3)$$

最後再將式(2-3)化簡成式(2-4)。由於式(2-4)中矩陣的第三列為第一列和第二列的線性組合，因此對式(2-4)求解，等同於對式(2-5)求解。至此已經推導出以 h_i 為變數的方程式，接著只要解出 h_i ，其中 $1 \leq i \leq 9$ ，我們便可算出 H 。

$$\begin{bmatrix} 0^T & -P_A^T & y_B P_A^T \\ P_A^T & 0^T & -x_B P_A^T \\ -y_B P_A^T & x_B P_A^T & 0^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (2-4)$$

$$\begin{bmatrix} 0^T & -P_A^T & y_B P_A^T \\ P_A^T & 0^T & -x_B P_A^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (2-5)$$

雖然單應性轉換矩陣 H 有 9 個未知數，但因為其中 1 個為縮放係數(up to scale)，所以單應性轉換矩陣 H 的自由度為 8，即至少需要 8 條方程式才能解出 H 。由式(2-5)得知，給定一組對應點 $[x_A, y_A]^T$ 和 $[x_B, y_B]^T$ 可以獲得兩條方程式，故至少需要 4 組對應點，方能算出單應性轉換矩陣 H 。而求解 H 的方法不只一種，我們只介紹在本論文中的作法。

單應性轉換矩陣之計算

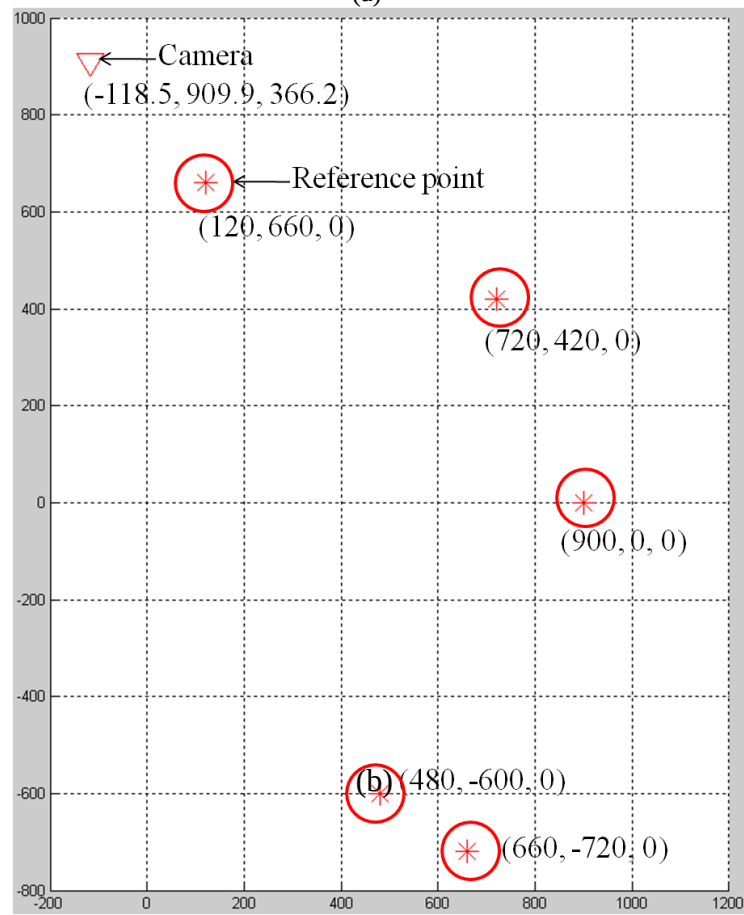
- (1) 由式(2-5)，給定一組 A 與 B 的參考點，可求得一個 2×9 的矩陣 M_i 。
- (2) 我們選擇 5 組參考點，如圖 2-1 中紅色圓圈所示，因此可得到 5 個 2×9 矩陣 M_i ，將這些 M_i 結合成一個 10×9 的矩陣 M 。
- (3) 將 M 做奇異值分解(singular value decomposition)，得到 $M = UDV^T$ ，其中 D 為 M 的奇異值所構成的對角矩陣(diagonal matrix)、 V 為奇異值所構成的正交矩陣(orthogonal matrix)，而我們欲求的 H 是最小的奇異值所對應的奇異向量。一般而言，奇異值分解後的對角矩陣的對角項會由大到小排列，也就是正交矩陣的第九行為構成 H 的 9 個元素，即

$$H = [h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9]^T。$$

- (4) 得到構成 H 的 9 個元素後，根據式(2-2) H 的定義，即能求出 H 。



(a)



(b)

圖 2-1 場景中參考點的配置。(a) 影像中的參考點，(b) 真實空間中的參考點。

2.2 定位方法

在介紹定位方法之前，我們需要兩個先決條件，第一是知道真實空間中攝影機的中心位置資訊，第二是知道真實空間中目標點的高度資訊。由於在估算真實空間中的目標點位置的過程中，我們會利用到真實空間中攝影機的中心位置和目標點的高度位置，因此我們假設此兩項資訊為已知。在實作的方面，攝影機的中心位置資訊可藉由事先測量得到；目標物的高度資訊，則可假設場景中有一道閘門，閘門上有高度資訊的刻度，當目標點進入場景時會先通過這道閘門，因此我們事先得到此目標點的高度資訊。

到目前為止，我們已經知道的資訊有單應性轉換矩陣、攝影機的中心位置資訊和目標點的高度資訊，有了這些資訊後，我們將介紹如何利用這些資訊估算出真實空間中目標點的位置。如圖 2-2 所示，首先我們在影像平面上取出目標點 P_{IP} ，再藉由單應性轉換矩陣 H ，將影像目標點投影到 π_1 平面上，即參考平面投影點 P_{RP} ，轉換關係如式(2-6)，其中 $P_{IP}=[u, v, 1]^T$ ， $[x, y]^T$ 為真實空間中參考平面上的 X 座標和 Y 座標。接著，算出參考平面投影點和攝影機中心點的 3 維連線 L 和 π_2 平面之方程式，其中 π_2 平面與 π_1 平面平行且平面高度等於目標點的高度。最後，算出 3 維連線 L 與 π_1 的交點 P_{HP} ，此交點即為估算的定位結果。

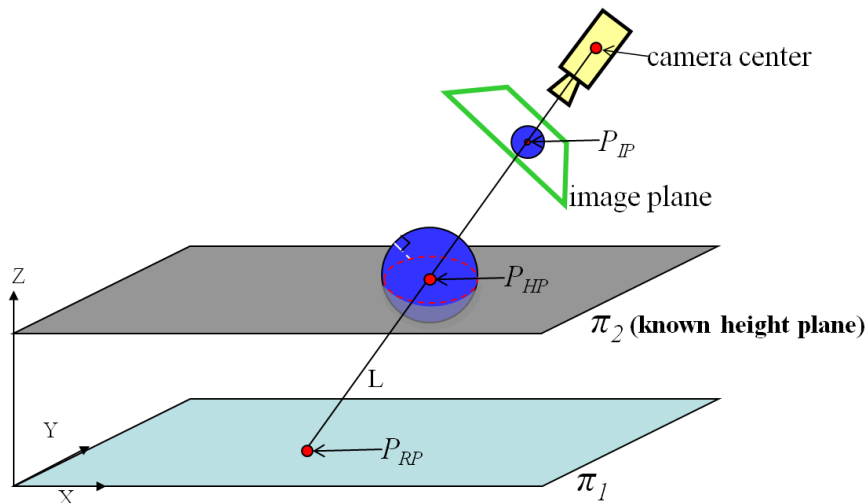


圖 2-2 單攝影機的定位示意圖。

$$H \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} \quad (2-6)$$

因為單應性轉換為透視投影的關係，一旦我們取得目標點所在平面的多組參考點後，即可透過單應性轉換矩陣直接找出目標點在真實空間中的位置，不必再透過直線與平面的交點這道程序。我們可進一步簡化作法，假設 π_1 平面等於 π_2 平面，即只存在一個參考平面，則此作法可省去計算平面與直線交點的步驟。接下來的討論中，我們皆只探討一個參考平面的情況。

2.3 初步定位結果

在本實驗中，我們具有兩個實驗場景，場景 1 的環境中，只存在一台攝影機，而場景 2 則存在多台攝影機。接下來我們將討論場景 1 的定位結果，在場景 1 中，我們設定參考平面為地板平面，且目標點亦位於地板平面。如圖 2-3 所示，(a)為場景 1 的俯視圖，其中紫色圓圈為用來測量的目標點位置，共有 19 個位置；(b)為取得的場景影像，其中黑色十字對應到(a)中的紫色圓圈中的交叉點，紅色米字為我們選取的參考點，共有 5 個參考點。圖 2-4 為場景 1 的定位結果，其中藍色點為估計出來的位置，整體誤差的平均為 17.36 公分。由於攝影機和目標點皆靜止不動，但是卻存在明顯的定位誤差，因此接下來我們欲探討造成定位誤差的原因。

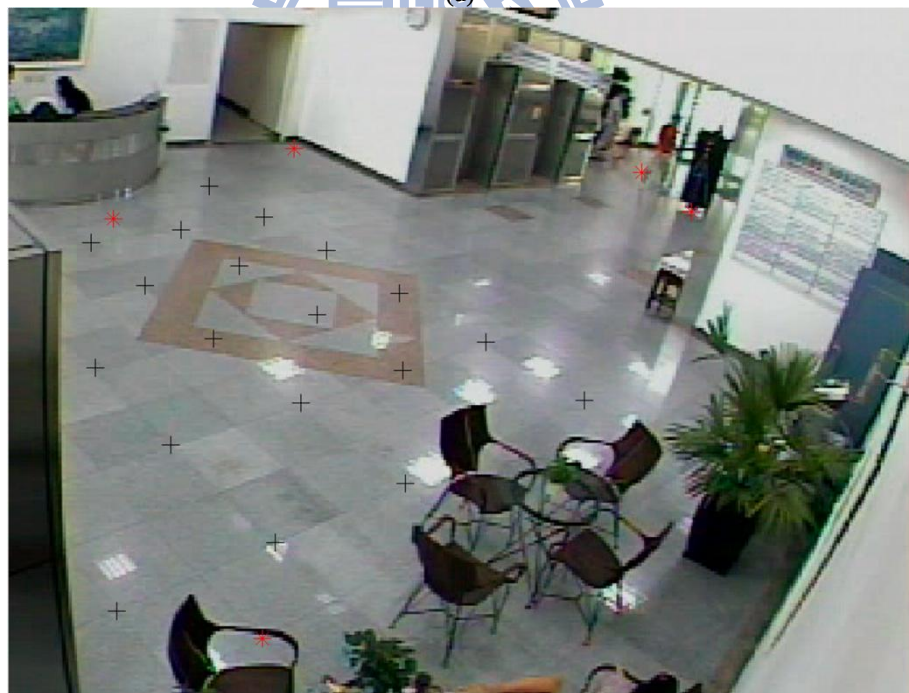
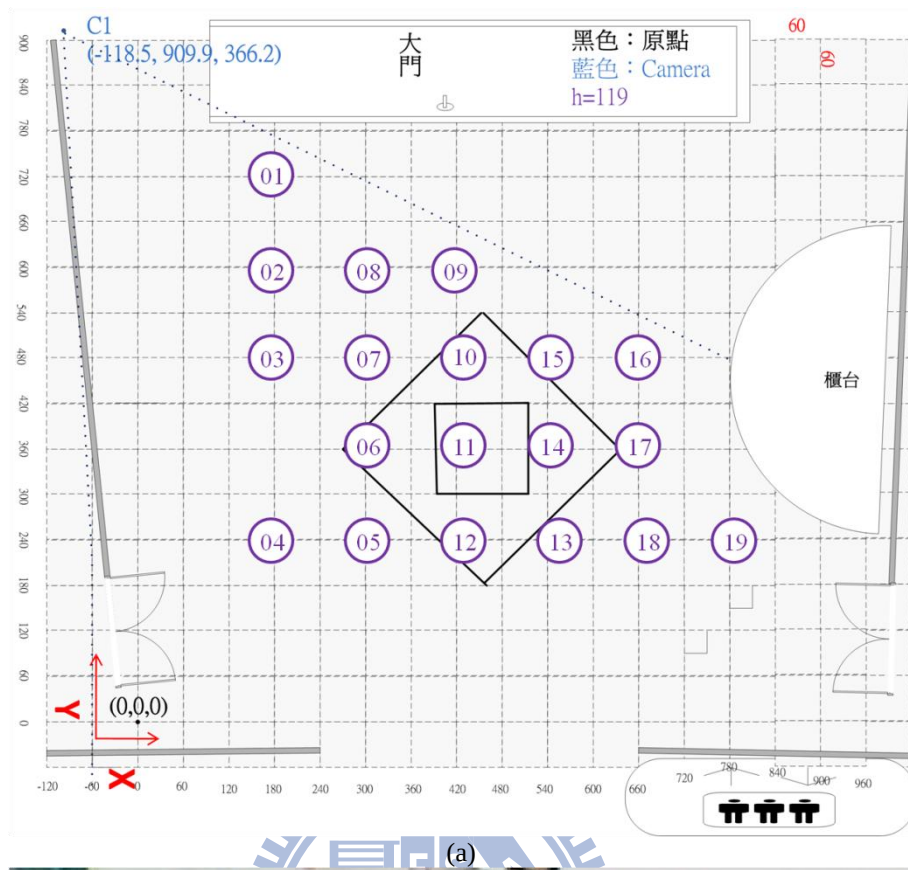


圖 2-3 實驗場景 1。(a)場景頂視圖，(b)場景影像。

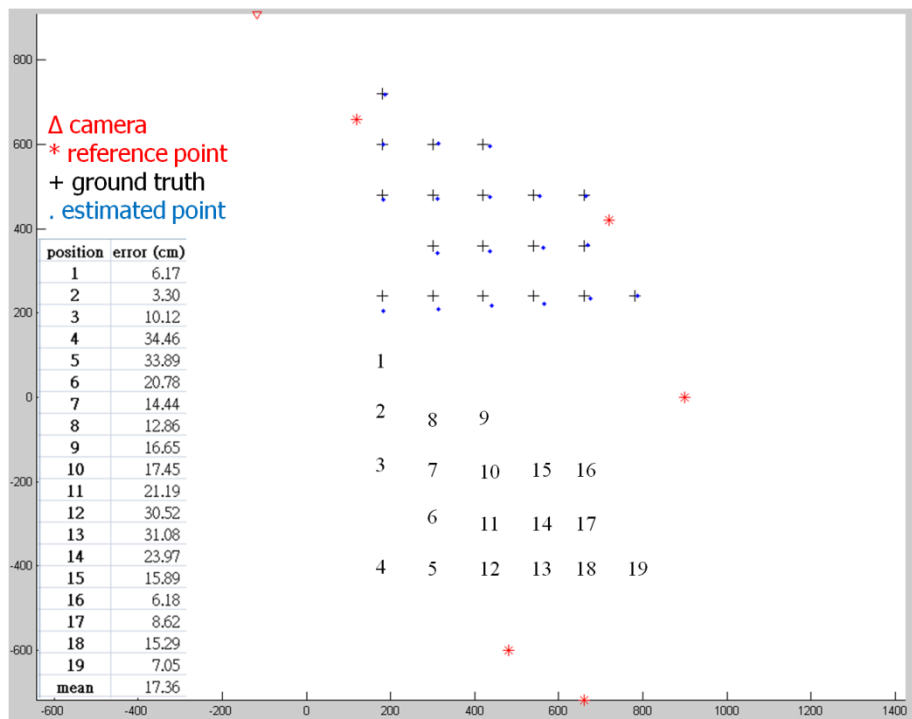


圖 2-4 實驗場景 1 之定位結果。



第三章 誤差來源之探討與改善

由 2.3 節可得知，在我們實驗之後發現，仍存在一些明顯的定位誤差。因此在本章中，我們欲探討可能造成定位誤差的原因，主要的原因有三種，分別是靜態的徑向畸變誤差(radial distortion)和人為測量誤差，以及動態的取像誤差，接著對這些誤差來源加以分析和改善，以獲得較佳的定位結果。

3.1 誤差的來源

參考圖 3-1，很明顯地在影像中我們可以看出，本來應該是水平/鉛直的直線因受到攝影機鏡頭的影響，產生徑向畸變，特別是魚眼鏡頭影響更大。徑向畸變使得原本應該是直的線段變成彎曲的線段，距離影像中心越遠，彎曲的程度越大。由於這個原因，使得在影像中找到的目標點座標，無法以單應性轉換正確地對應到我們在真實空間中所經過的位置。所以徑向畸變誤差會是造成定位誤差的來源之一。除此之外，我們所提出來的定位方法，需要用到真實空間中參考點的位置以計算單應性轉換矩陣。由於參考點的位置資訊是經由人為測量所得，因此勢必存在測量誤差，間接地反應到定位誤差上。所以人為測量誤差亦是造成定位誤差的來源之一。最後，由於攝影機在取像的過程中，經常受到光影等干擾，使得我們取得的目標點資訊會受到雜訊干擾，造成最後定位出來的結果不如預期。因此，取像誤差也是造成定位誤差的來源之一。經由以上的討論，我們將徑向畸變誤差和人為測量誤差視作靜態誤差，而取像誤差視作動態誤差。由於徑向畸變和人為測量誤差在所造成的定位誤差是固定不變的，因此稱作靜態誤差。不像取像誤差，可能因為光影等因素的影響，造成每次的定位誤差皆不同，因此稱作動態誤差。在接下來的兩節中，我們將介紹如何改善靜態誤差和模擬動態誤差加以分析定位誤差的分佈，以期得到更佳的定位結果。

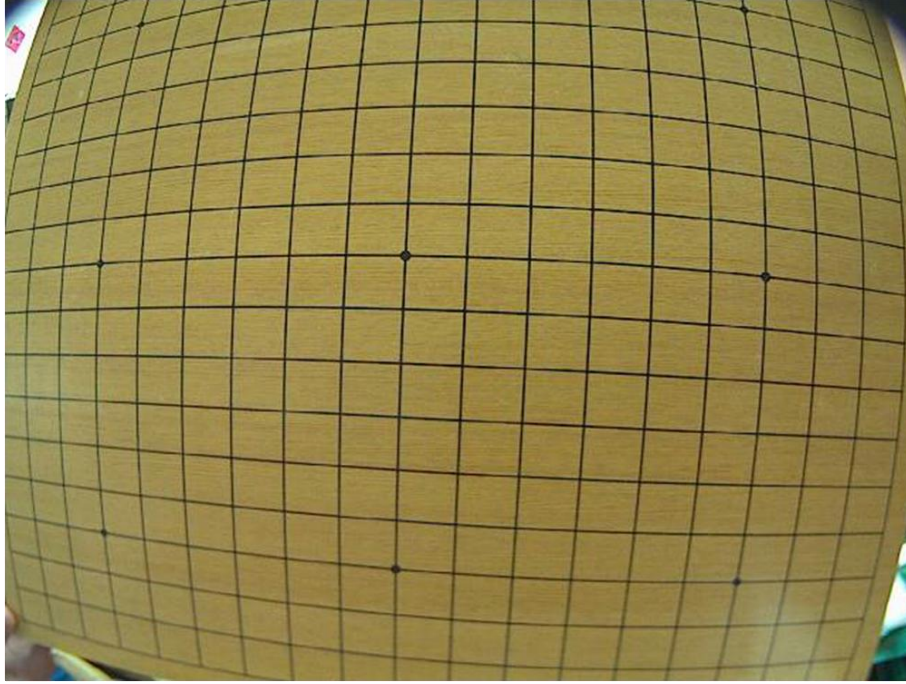


圖 3-1 受到徑向畸變影響的影像。

3.2 靜態誤差

在本節中我們將分別描述如何改善徑向畸變誤差和人為測量誤差。我們所採取的方式為對影像作校正，做完校正後再微調參考點的測量位置，希望藉由這些方式，可以減少靜態誤差對定位造成的影響。

3.2.1 徑向畸變誤差

在徑向畸變的方面，我們想要做的事情為，如何把彎曲的線段變成直線，如圖 3-2 所示。我們參考式(3-1)對影像進行校正

$$\begin{aligned} P'_u &= P_u(1 - a_u \|P\|^2) \\ P'_v &= P_v(1 - a_v \|P\|^2) \end{aligned} \quad (3-1)$$

其中 $P = (P_u, P_v)$ 為置中且正規化(normalization)的影像座標， $\|\cdot\|$ 為座標向量的長度； a_u 和 a_v 為校正參數，其數值皆小於 0。參數 a_u 為調整影像 U 座標的比值，參數 a_v 為調整影像 V 座標的比值，當兩個值皆為 0 時，即不調整影像座標。校正影像的運作原理為，先將影像座標作置中且正規化，如圖 3-3 所示。再依據正規化後的影像座標與影像中心點的距離作調整，當距離較大時，往外推的幅度較大，當距離較小時，往外推的幅度較小。舉例來說，假設目前要調整藍色點和紅

色點的影像 V 座標，如圖 3-4 所示，由於藍色點與影像中心的距離較小，因此調整的距離較小；反觀紅色點與影像中心的距離較大，因此調整的距離也較大。

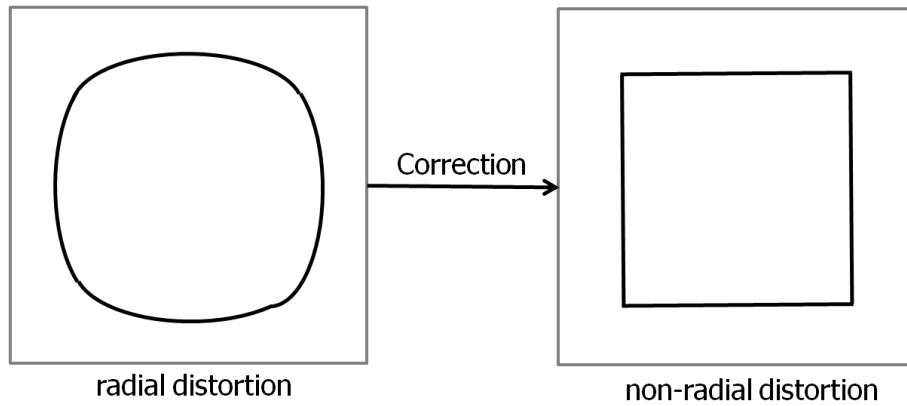


圖 3-2 徑向畸變校正示意圖。

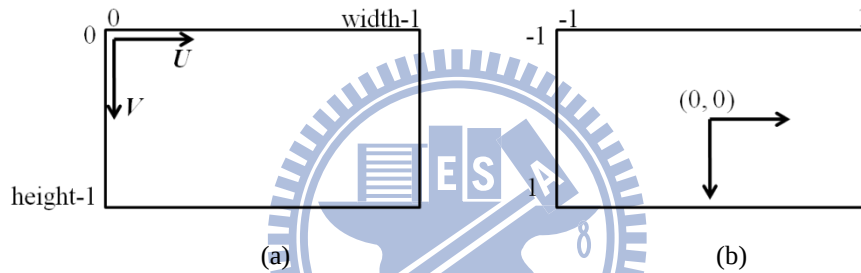


圖 3-3 影像座標正規化。(a)影像座標 (u, v) ，(b)置中且正規化影像座標 (P_u, P_v) 。

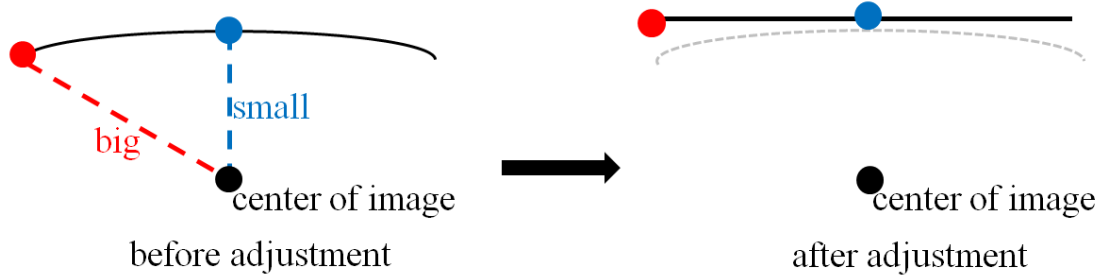


圖 3-4 影像校正時，座標調整示意圖。

在實作的部份，我們利用圖 3-1 中的格子交叉點當驗證點，以找出校正參數 a_u 和 a_v 。所有的驗證點可以形成 16 條水平線和 22 條垂直線。我們的目的是希望能將影像中彎曲的水平線和垂直線，經校正後能變直線。因此我們的做法為在一定範圍內選取校正參數 a_u 和 a_v ，透過所有驗證點，套用式(3-1)，以主成分分析 (principal component analysis, PCA) 當依據，找出範圍內校正效果較佳的參數 a_u

和 a_v ，校正參數詳細的作法如下：

- (1) 亂數選取一定範圍內的校正參數 a_u 和 a_v 。
- (2) 對每一條線(含水平和垂直)，我們選取線上的驗證點，利用公式(3-1)加以校正，共有 38 條線，其中 16 條為水平線，22 條為垂直線。
- (3) 將校正後的驗證點，利用主成分分析算出最小特徵值 (eigenvalue)，我們取所有特徵值的總合作為此校正參數的分數。
- (4) 試過範圍內的所有校正參數後，挑選一組分數最小的校正校數，此組參數即可推得較佳的校正結果。

如圖 3-5 所示，(a)為未校正的影像，(b)為校正後的影像。從圖中可以看出，藍色箭頭指的水平線和紅色箭頭指的垂直線，校正後有一定幅度的改善。在找出校正參數後，我們對場景 1 的 19 個目標點作校正，並估算出其定位結果，如圖 3-6 所示。19 個位置的平均誤差已由 17.36 公分降到 14.26 公分，大約下降 18%。

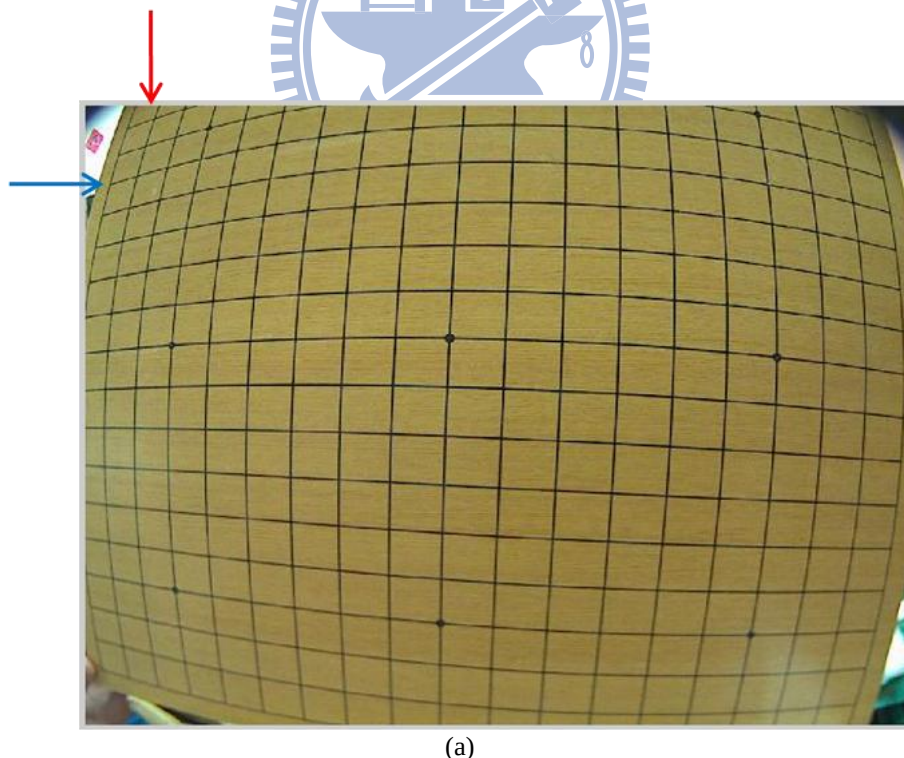
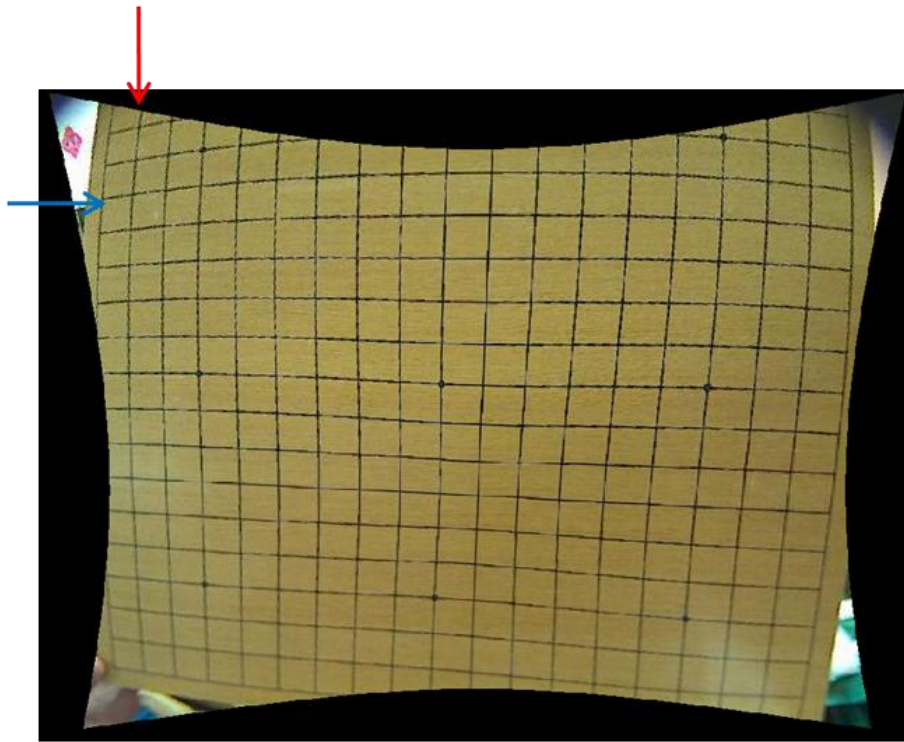


圖 3-5 影像校正。(a) 校正前，(b) 校正後。



(b)
圖 3-5 (續)。

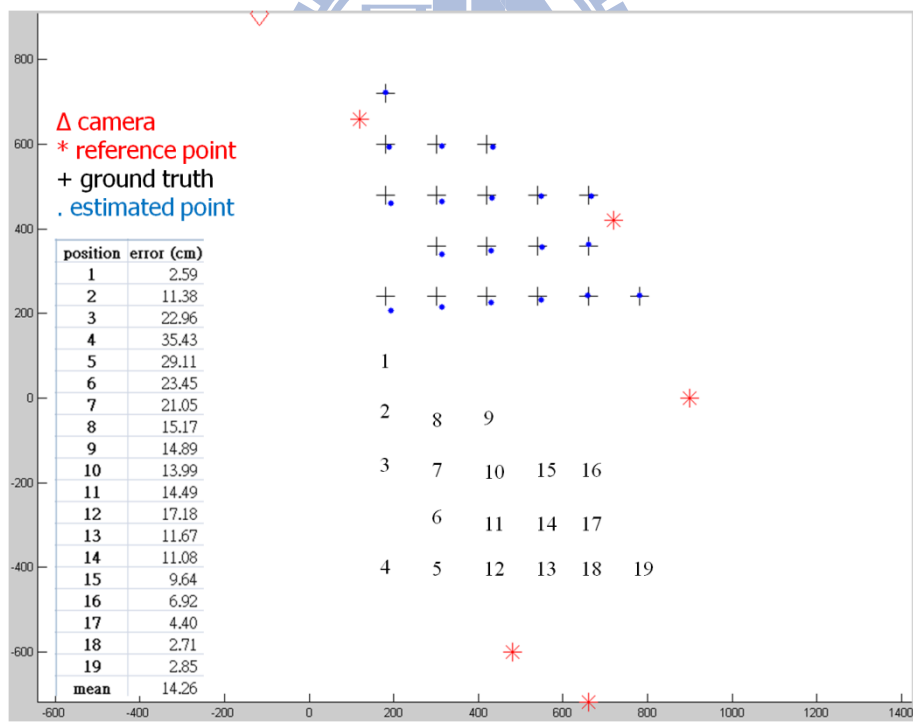


圖 3-6 實驗場景 1 校正後之定位結果。

3.2.2 人為測量誤差

由於人為測量參考點在真實空間中的位置時，勢必存在測量誤差，因此我們希望透過微調真實空間中參考點的測量位置，以找出可能是正確的真實空間中的參考點位置，如圖 2-1(b)，紅色米字即為我們測量出來的 5 個真實空間中的位置。我們微調的方式為，針對一個參考點而言，對其作上、右上、右、右下、下、左下、左、左上等 8 個方向的微調，微調固定的距離，再選取一些驗證點，以驗證每種組合估算出來的定位誤差。實驗場景 1 有 5 個參考點，每個參考點微調的 8 個方向，加上本身未移動的情況，總共具有 59049 (9^5) 種可能的組合。經由實驗後，微調的距離為 8.5 公分時，所得到的定位誤差最小。如圖 3-7 所示，紅色米字代表各個參考點微調後真實空間的位置。如圖 3-8 所示，其為校正加上微調參考點位置後的定位結果。我們可以看出校正加上微調參考點位置後，定位誤差已由 17.36 公分降到 4.26 公分，下降幅度達 75%。到目前為止，經由校正影像和微調參考點後，已能得到較佳的定位結果。接下來，我們欲討論取像時受雜訊干擾所造成的動態誤差。



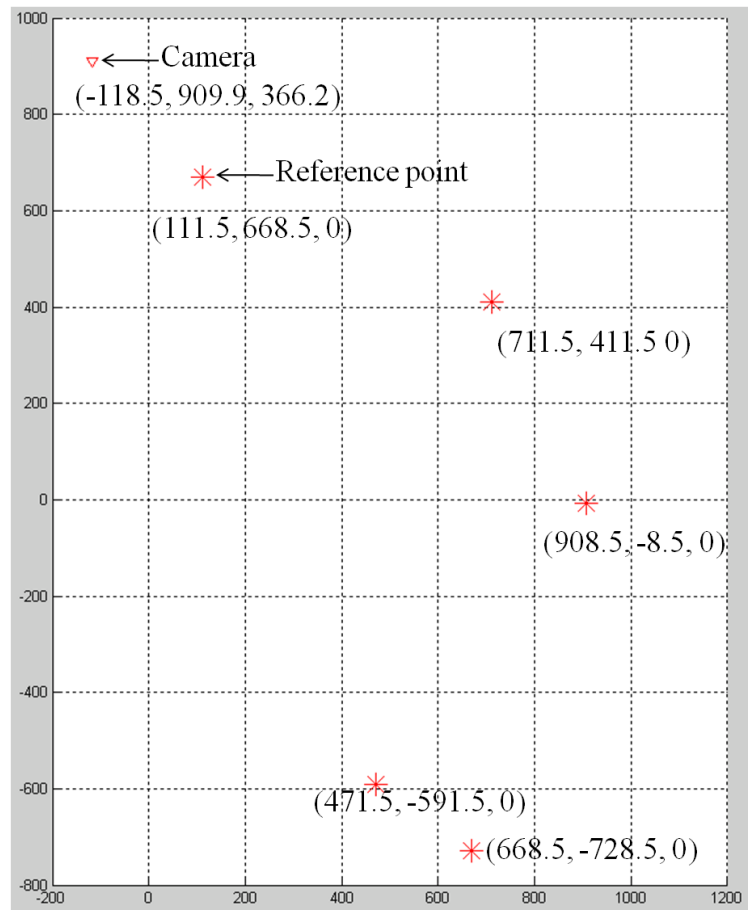


圖 3-7 微調後真實空間中的參考點位置。

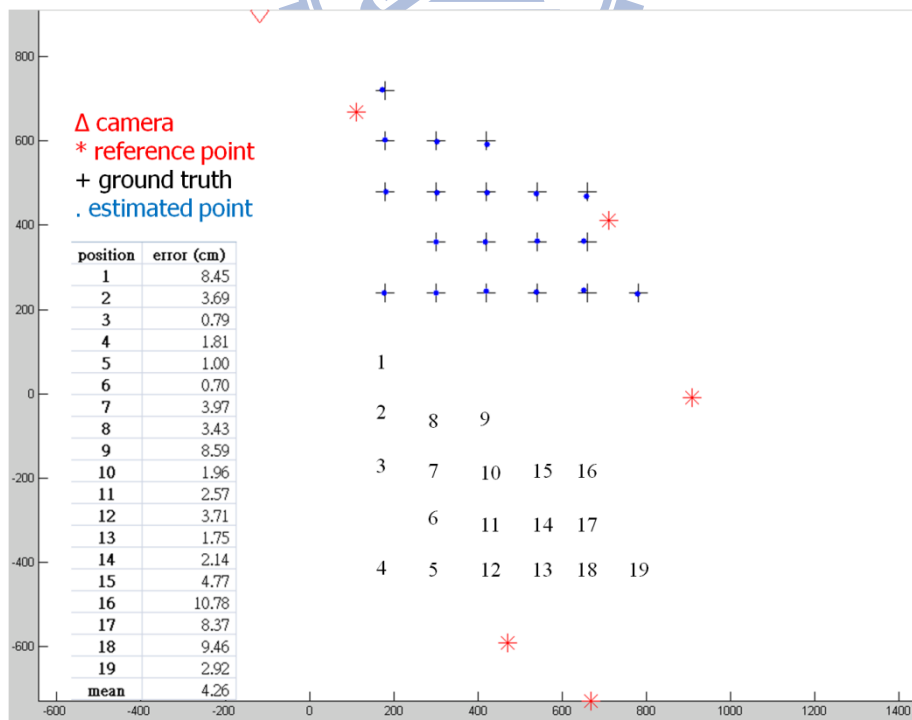


圖 3-8 實驗場景 1 校正和微調參考點後之定位結果。

3.3 動態誤差

由於攝影機在取得影像的過程中，常常會受到光影變化等雜訊干擾，使得在影像上取得目標點的結果不穩定，我們稱之為動態誤差。為了分析動態誤差所造成的影響，我們欲模擬取像的雜訊。首先，我們在影像目標點周圍加上 100 個高斯雜訊點，以模擬影像目標點受到雜訊干擾後的位置。如圖 3-9 所示，在(a)中，我們可以看到影像上加入雜訊後，其分佈的情況呈現圓形分佈，其中紅色點為影像目標點；(b)為將(a)中的雜訊經由單應性轉換矩陣作用後，投影到參考平面上的情形，即為我們的定位誤差分佈，可看得出來其呈現橢圓形的分佈。由此可知，在影像上受到雜訊干擾的分佈，經過投影後會有不同的結果，即不同的攝影機拍攝角度與不同的距離將會造成不同的分佈結果。由圖 3-9(b)我們還可以觀察到，誤差的分佈在橢圓長軸的方向與短軸方向之大小亦不同，在橢圓長軸的方向誤差分佈較大，反觀在橢圓短軸方向則誤差分佈較小，此代表著定位結果在橢圓短軸的方向具有較高的穩定性，而在橢圓長軸的方向則穩定性較差，因此可藉由此特性幫助我們評估定位結果的可靠度。由於為了看定位誤差分佈的情形，每次都要對影像目標點模擬多個高斯雜訊點，再經由單應性轉換矩陣將各個雜訊點投影到參考平面上，最後才可知定位誤差的分佈，此過程太耗時間。因此，若能較方便且快速地找出定位誤差的分佈情形，則可提昇效率。

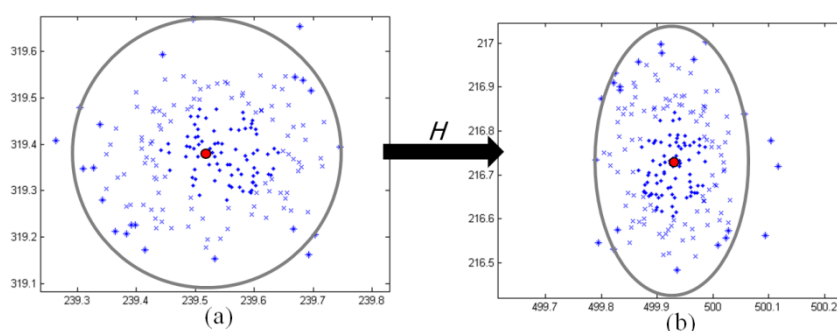


圖 3-9 模擬取像誤差。(a) 影像中加入 100 個高斯雜訊，(b) 經由單應性轉換矩陣作用後的 100 個雜訊。

為了達到這個目的，在影像上我們設定了一個圓形的邊界，此邊界可以將模擬產生的高斯雜訊點圍住。而圓形的半徑則可依據攝影機拍攝角度與不同的距離

等情形來設定，例如估計出高斯雜訊分佈之標準差為 σ ，則在影像上以目標點為中心的圓形邊界，其半徑可被設定成高斯雜訊 k 倍的標準差，即 $r = k\sigma$ ，其中 k 為一個常數，如圖 3-10 所示，其中半徑 $r = 3$ 個像素。接下來，我們欲利用影像目標點在影像中目標點的座標、邊界圓的半徑和單應性轉換矩陣 H 找出定位誤差分佈呈現的橢圓參數，藉由參數估算定位誤差分佈的情形，如圖 3-11 所示。其中 $[x, y]$ 為橢圓中心點位置， $[a, b]$ 為橢圓的長軸和短軸， θ 為橢圓旋轉的角度。

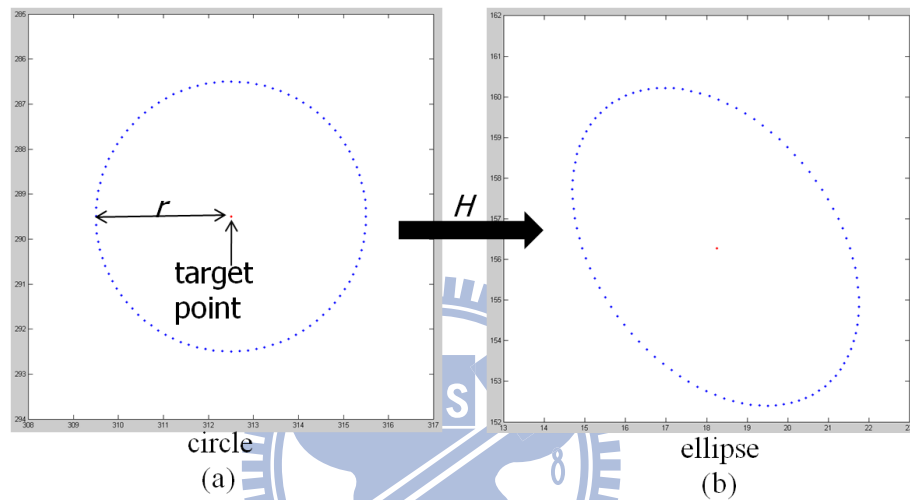


圖 3-10 模擬影像誤差。(a) 影像上的邊界圓，(b) 經由單應性轉換矩陣作用後的邊界圓。

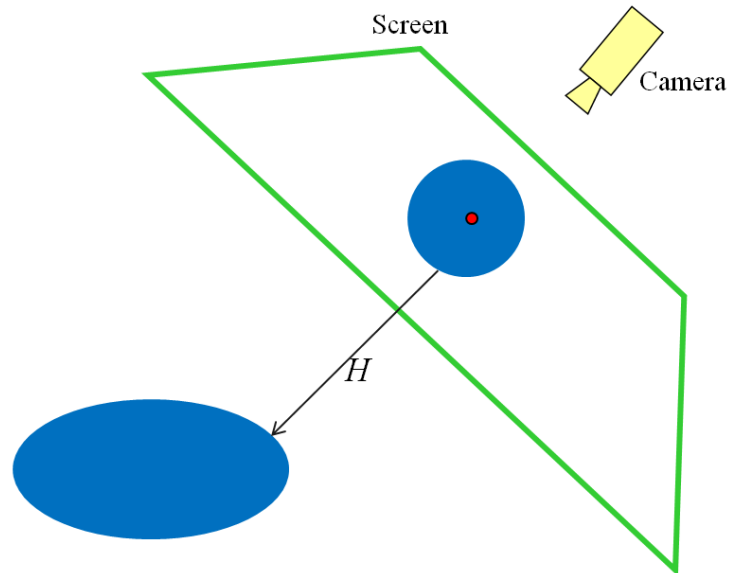


圖 3-11 圓形誤差分佈的定位誤差橢圓。

假設邊界圓的方程式為 $(u-p)^2 + (v-q)^2 = r^2$ ，其中 p 和 q 為影像目標點的座標。再進一步推導出參考平面上橢圓的參數，即定位誤差分佈的橢圓。 H 為先

前算出來的單應性轉換矩陣，可用式(3-3)將參考平面上的點轉換到影像上，即 $[x, y, 1]^T$ 轉換後變 $[u, v, 1]^T$ 。

$$H = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad (3-2)$$

$$H \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} ax + by + c \\ dx + ey + f \\ gx + hy + i \end{bmatrix} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad (3-3)$$

將 u 和 v 代入邊界圓方程式中得到式(3-4)，再將式(3-4)整理成式(3-5)。

$$((ax + by + c) / (gx + hy + i) - p)^2 + ((dx + ey + f) / (gx + hy + i) - q)^2 = r^2 \quad (3-4)$$

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

$$A = a^2 - 2pag + p^2g^2 + d^2 - 2qdg + q^2g^2 - r^2g^2$$

$$C = b^2 - 2pbh + p^2h^2 + e^2 - 2qeh + q^2h^2 - r^2h^2$$

$$F = c^2 - 2pci + p^2i^2 + f^2 - 2qfi + q^2i^2 - r^2i^2 \quad (3-5)$$

$$B = 2ab - 2p(ah + bg) + 2p^2gh + 2de - 2q(dh + eg) + 2q^2gh - 2r^2gh$$

$$D = 2ac - 2p(ai + cg) + 2p^2gi + 2df - 2q(di + fg) + 2q^2gi - 2r^2gi$$

$$E = 2bc - 2p(bi + ch) + 2p^2hi + 2ef - 2q(ei + hf) + 2q^2hi - 2r^2hi$$

此(斜)橢圓又可以表示為式(3-6)，其中 $[j, k]$ 為橢圓的中心，展開後為式(3-7)。比較式(3-5)和(3-7)後，可得出 $D = -2Aj - Bk$ 和 $E = -Bj - 2Ck$ 。

$$A(x - j)^2 + B(x - j)(y - k) + C(y - k)^2 + f = 0 \quad (3-6)$$

$$Ax^2 + Bxy + Cy^2 + (-2Aj - Bk)x + (-Bj - 2Ck)y + (Aj^2 + Bjk + Ck^2 + f) = 0 \quad (3-7)$$

因此，橢圓的中心點 $[j, k]$ 可用式(3-8)算出；橢圓的旋轉角度可由 $\tan(2\theta) = B/(A-C)$ 算出；令旋轉前的橢圓為 $[X, Y]^T$ ，旋轉後的橢圓為 $[x, y]^T$ ，則兩者的關係如(3-9)。

$$\begin{bmatrix} j \\ k \end{bmatrix} = \begin{bmatrix} -2A & -B \\ -B & -2C \end{bmatrix} \begin{bmatrix} D \\ E \end{bmatrix} \quad (3-8)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \quad (3-9)$$

至於橢圓的長軸與短軸可由式(3-10)算出。

$$\begin{aligned}
 axis_a &= \sqrt{\frac{\delta}{\alpha}}, \quad axis_b = \sqrt{\frac{\delta}{\beta}} \\
 \alpha &= A\cos^2\theta + B\sin\theta\cos\theta + C\sin^2\theta \\
 \beta &= A\sin^2\theta - B\sin\theta\cos\theta + C\cos^2\theta \\
 \lambda &= D\cos\theta + E\sin\theta \\
 \gamma &= -D\sin\theta + E\cos\theta \\
 \delta &= \lambda^2/4\alpha + \gamma^2/4\beta - F
 \end{aligned} \tag{3-10}$$

到目前為止，我們已經成功地導出用影像目標點的座標、邊界圓的半徑和單應性轉換矩陣 H 找到參考平面上定位誤差分佈的橢圓參數，參數包含橢圓中心的座標、橢圓的長短軸和橢圓的旋轉角度。利用橢圓參數將可使我們估算定位誤差分佈的時間大為縮短，有效提昇誤差分析的效率。另外，橢圓內的區域可被視為信心區域(confidence region, CR)，即定位點有極高的機率會落在此區域內。這樣的資訊可以提供我們挑選合適的攝影機之定位結果，例如，當橢圓面積很小時，即代表該攝影機的定位結果具有較高的穩定性，此結果可被用來當作較佳的定位結果。此外由橢圓的參數我們可以看出，雖然在影像中受到同樣程度的干擾，然而定位誤差分佈在橢圓短軸的方向偏移較小，而在橢圓長軸的方向偏移較大，這也能作為挑選攝影機的一種參考。

第四章 整合多攝影機之定位結果

在 3.1 節和 3.2 節中，我們討論了造成定位誤差的原因，並且加以改善。在 3.3 節中，我們探討取像誤差和如何分析其造成的定位誤差分佈。在本章中，我們欲討論當場景中具有多個攝影機時，如何利用第二章與第三章討論的方法，整合多攝影機的定位結果。我們提出一個暴力法以分析多攝影機之定位誤差分佈，期望分析的結果可以作為挑選較準確的攝影機配對之依據。由於暴力法需要較多的計算，因此我們再提出簡化法，僅利用橢圓參數即可估算定位誤差分佈，藉此減少計算量，加以提昇誤差分析的效率。在 4.1 節中，我們討論如何將兩台攝影機的定位結果整合在一起；在 4.2 節中，我們討論如何利用暴力法估算多攝影機的定位誤差分佈，依據定位誤差的分佈找出整合多攝影機定位結果的信心區域，挑選合適的攝影機配對；在 4.3 節中，我們討論如何利用簡化法估算多攝影機的定位誤差分佈，為拿此分佈當作挑選合適攝影機配對的依據。

4.1 整合一對攝影機的定位方法

由 2.2 節得知，在單攝影機情況下之定位方法，我們想要整合一對攝影機的定位結果，我們所使用的方法如圖 4-1 所示。首先，在攝影機 A 的影像中找到目標點 P_{IP} ，利用單應性轉換矩陣將其投影到參考平面上，得到攝影機 A 的定位點 P_{HP} 。接著，將攝影機 A 的中心點垂直投影到參考平面上，即 A_{HP} 。再算出 P_{HP} 和 A_{HP} 在 3 維中的連線，即 L_A 。同理，對攝影機 B 亦能算出 3 維中的連線 L_B 。最後，找出 L_A 和 L_B 的交點(fused point, FP)，此交點即為整合一對攝影機之定位點。

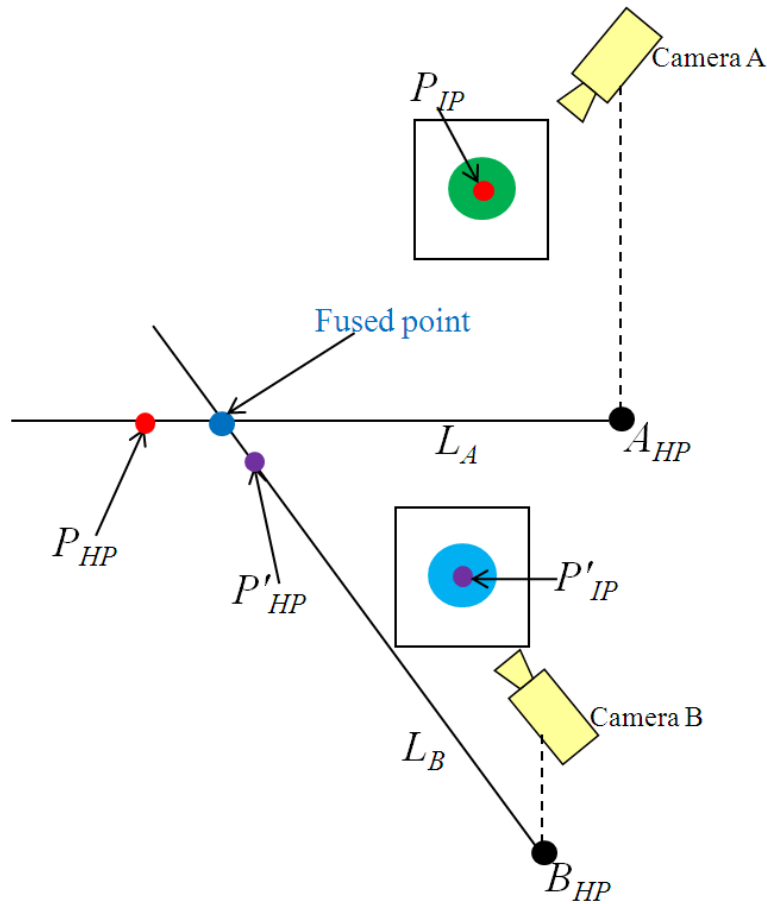


圖 4-1 整合一對攝影機的定位結果。

4.2 暴力法估算定位誤差分佈

在此節中，我們欲利用 4.1 節提出的整合方法，加上模擬動態誤差，以分析定位誤差的分佈。一對攝影機之定位結果會因攝影機各別受到雜訊干擾而變得較不穩定，為了模擬這樣的情形，我們在影像目標點之邊界圓上放置 24 個模擬點，如圖 4-2 所示，利用 4.1 節提出的整合一對攝影機的定位方法，我們可以找出 576(24x24)個 FP。接下來，我們欲用暴力法估算所有 FP 的分佈。我們採取凸包(convex hull)演算法將所有的點圍繞起來，形成一個四邊形的區域，此區域即可被視為一對攝影機之定位的 CR。當 CR 的面積較小時，我們可以判斷該對攝影機具有較準確的定位結果。因此，我們可以整合多攝影機的定位結果，並依據每種種組合的 CR 挑選出合適的攝影機配對，以達到較準確的定位結果。

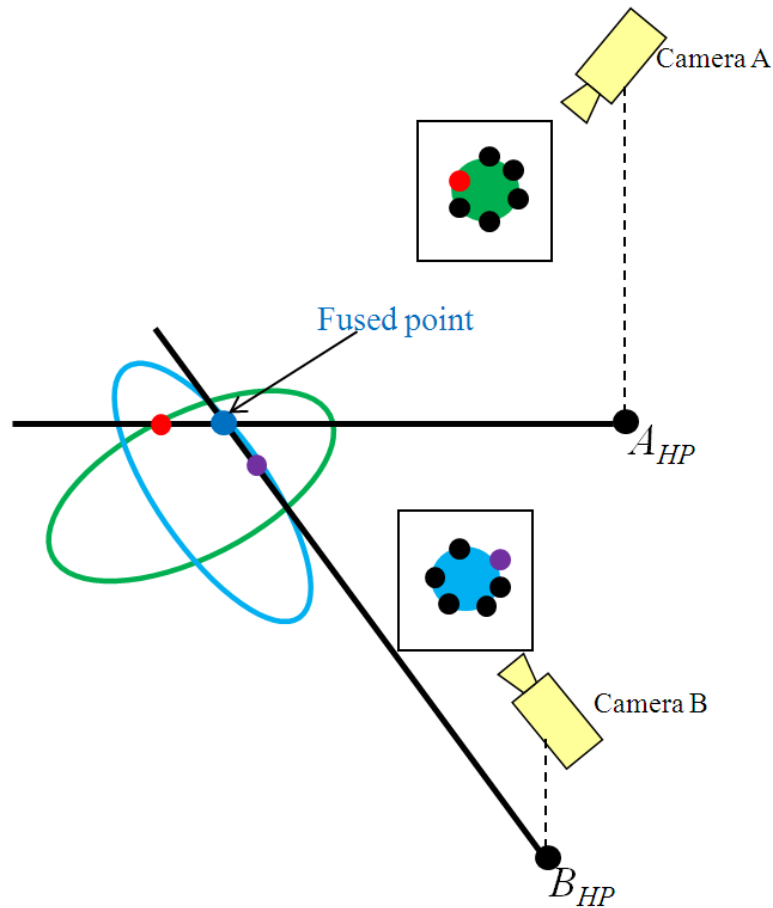


圖 4-2 整合多攝影機間的定位誤差分佈。

為了準確地估算 CR，邊界圓上放置的模擬點個數不宜太少。但是隨著模擬點數的增加，暴力法估算誤差分佈的計算量也會呈現平方的速度增加。因此，下面我們提出一個較為簡化的方法，利用之前找出來的橢圓參數，期望能更有效率地估算出誤差分佈。

4.3 簡化法估算定位誤差分佈

由於暴力法估算 CR 所需之計算量很高，因此我們利用在 3.3 節中所得到的定位誤差分佈的橢圓參數，快速地估算出一個大概的信心區域(approximate confidence region, ACR)以有效地減少計算量。如圖 4-3 所示，我們可以由先前攝影機 A 求出的橢圓參數分別算出橢圓短軸上的兩個端點 A_{CR} 和 A_{CL} ，再各自算出兩端點與 A_{HP} 的兩條直線。同理，利用攝影機 B 求出的橢圓參數，我們可以找出另外兩條直線，則此四條直線的交點所圍成之四邊形即為我們所求的 ACR。

由於攝影機架設都具有一定的高度，即與物體相距一定的距離，因此 ACR 的面積將會與用暴力法估算出來的定位誤差面積極為相似。

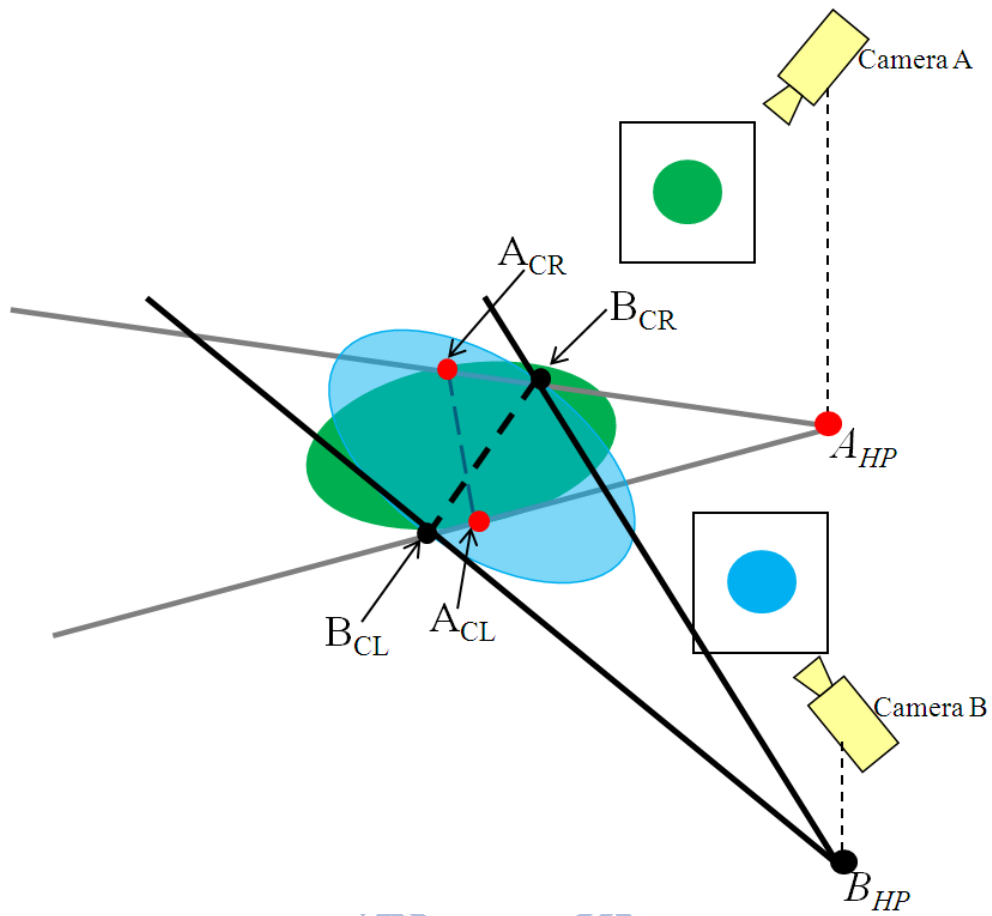


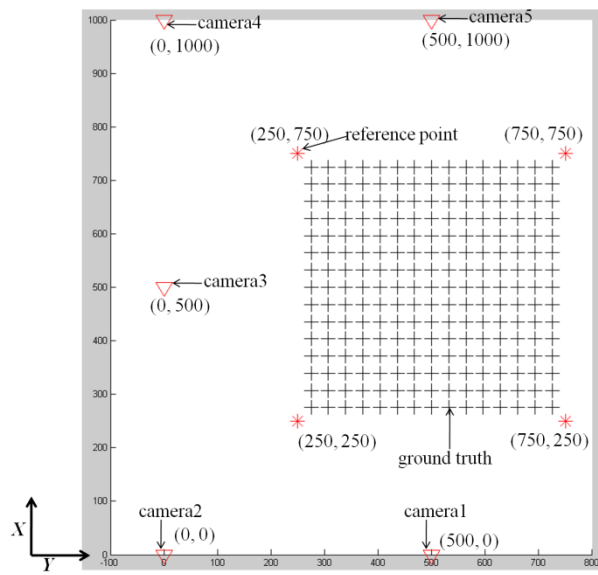
圖 4-3 簡化法估算誤差分佈示意圖。

第五章 實驗結果

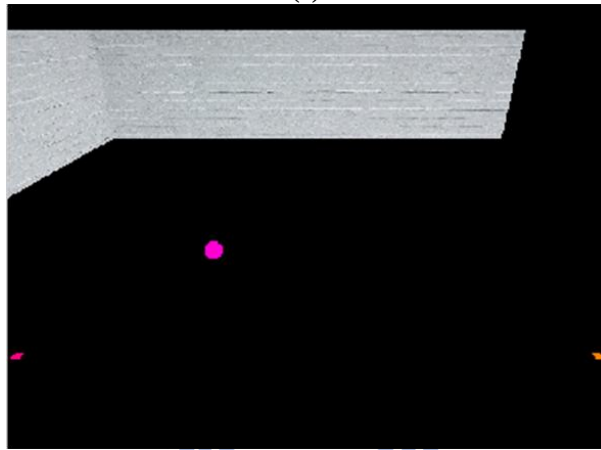
我們所提出來的基於單應該轉換之定位方法，在本章節可以看到較完整的實驗結果，證明我們所提出來的方法，確實可以準確地定位。我們使用虛擬的場景和實際場景以驗證定位方法的正確性。

5.1 虛擬場景的定位結果

我們利用程式建構一個虛擬場景，如圖 5-1 所示，(a)為虛擬場景的頂視圖，其中紅色三角形為攝影機，共有 5 台攝影機；紅色米字為參考點，共有 4 個參考點；黑色十字為目標點所經過的位置，共有 225 個位置。(b)為虛擬場景的影像，其中的粉紅色球體即為目標物，我們取目標物的中心當目標點。取得影像目標點後，我們還加上 100 個模擬點，以模擬受到雜訊干擾。接著由計算出各攝影機的橢圓參數，以估算信心區域，作為挑選合適攝影機配對的依據。表 5.1 為單攝影的定位誤差，由表 5.1 可看出，經由挑選攝影機後的平均定位誤差，明顯地比其他未經挑選攝影機之平均定位誤差來的小。在單攝影機情況下，挑選攝影機的依據為橢圓的面積，即橢圓面積為 CR。未挑選的結果來自所有位置都使用同一台攝影機作定位；至於挑選的結果，則是對目標點經過每個位置的作定位時，皆挑選橢圓面積最小的攝影機之定位結果。表 5.2 為整合多攝影機的定位結果，其中 pair12 即為 1 號攝影機和 2 號攝影機的整合定位結果，以此類推。由於場景中有 5 台攝影機，每次選兩台攝影機配對的定位結果，共有 10 種可能。從表 5.2 可看出來，經由攝選後的定位誤差仍然是最小的，挑選的依據則是 4.3 節所提到的 ACR。另外，在表 5.2 中亦可發現 pair15 的誤差很大，其原因為在所有的位置上，1 號攝影機和 5 號攝影機皆對看，因此無法彌補對方的不足，所以定位結果不好。較佳的情形為兩台攝影機的視線接近垂直，其具有較佳的定位結果。



(a)



(b)

圖 5-1 虛擬場景。(a)場景頂視圖，(b)場景影像。

表 5.1 單攝影機的定位誤差。

Error	mean	variance
camera1	2.991	1.429
camera2	4.967	2.558
camera3	2.989	1.352
camera4	4.978	2.649
camera5	2.984	1.295
Selection	1.872	0.209

表 5.2 整合多攝影機的定位誤差。

Error	mean	variance
pair12	2.935	1.442
pair13	1.777	0.225
pair14	5.507	103.958
pair15	167.926	879443.584
pair23	2.912	1.435
pair24	2.314	0.313
pair25	5.735	129.731
pair34	2.845	1.186
pair35	1.755	0.197
pair45	2.809	1.212
Selection	1.564	0.134

5.2 真實場景的定位結果

由虛擬場景的實驗結果可以看出，我們所提出來的定位方法是準確的。接下來，我們欲驗證實際場景的準確性。在實際場景中我們分成兩種情形，第一是以雷射點為目標點，雷射點在地上所經過的位置形成一個矩形，欲計算雷射點的軌跡；第二是以人物的頭部中心當目標點，欲計算人物所經過的軌跡。如圖 5-2 所示，(a)-(d)為我們在影像中所偵測到的雷射點所形成的軌跡。利用我們所提出的定位方法，定位結果如圖 5-3、圖 5-4 和圖 5-5 所示，其中圖 5-3 的(a)-(d)為只考慮單攝影機的定位結果，(e)為參考單攝影機的誤差分析，經由挑選後的定位結果，其中黑色的矩形為雷射點所經過的真實位置。圖 5-4 的(a)-(f)為考慮多攝影機(攝影機配對)的定位結果，(g)為參考多攝影機的誤差分析，經由挑選後的定位結果。圖 5-5 則為參考單攝影機和多攝影機的誤差分析，一起挑選後的定位結果。由實驗結果可以看出，經由挑選後的定位結果比未經挑選過的定位結果來得好一點。



(a)



(b)

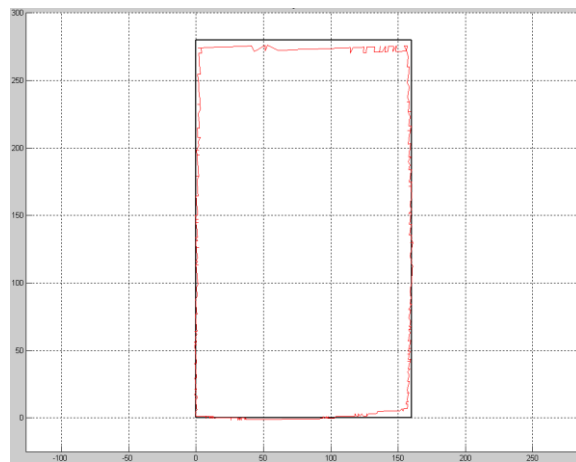


(c)

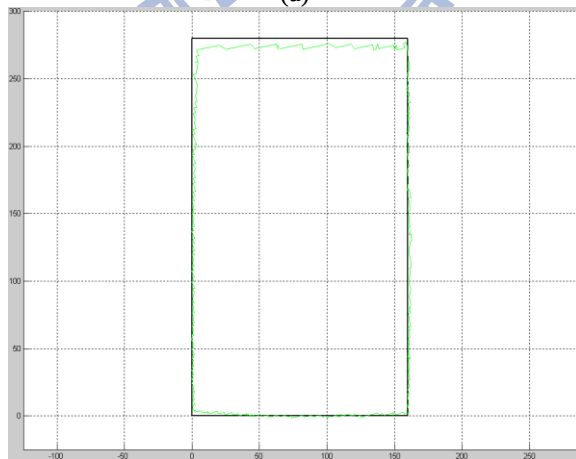
圖 5-2 實際場景—雷射點為目標點。(a) 攝影機 1 的影像，(b) 攝影機 2 的影像，(c) 攝影機 3 的影像，(d) 攝影機 4 的影像。



(d)
圖 5-2 (續)。

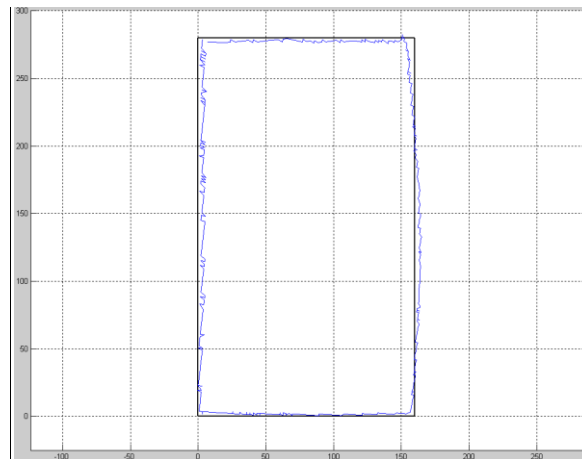


(a)

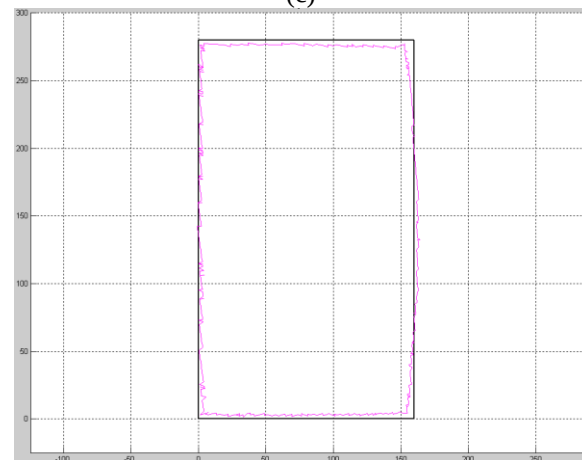


(b)

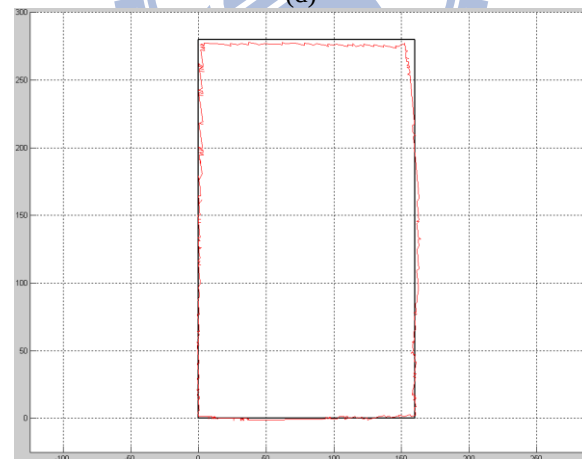
圖 5-3 單攝影機的定位結果。(a) 攝影機 1 的結果，(b) 攝影機 2 的結果，(c) 攝影機 3 的結果，
(d) 攝影機 4 的結果，(e) 挑選過後的結果。



(c)

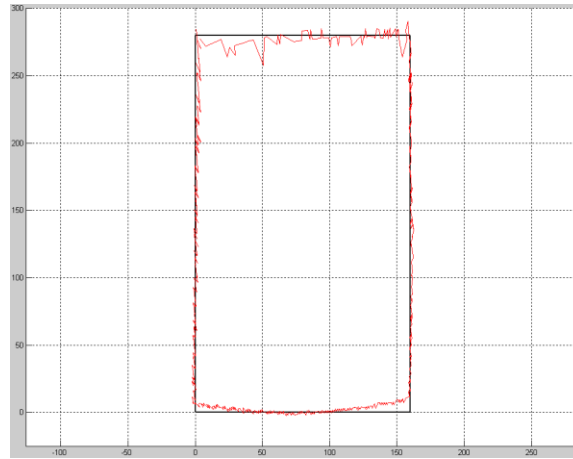


(d)

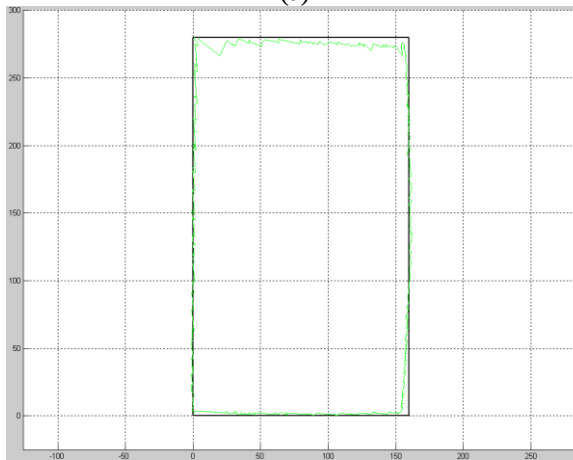


(e)

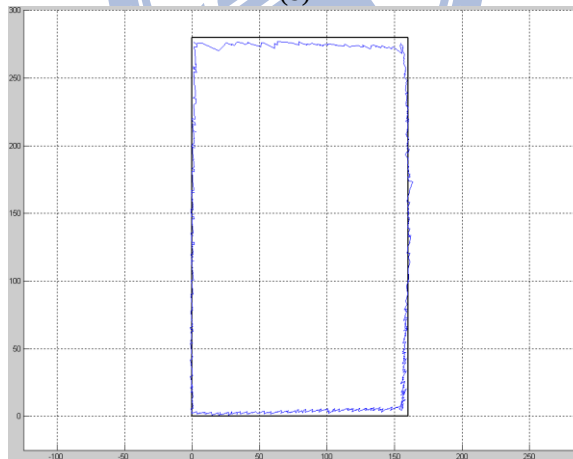
圖 5-3 (續)。



(a)

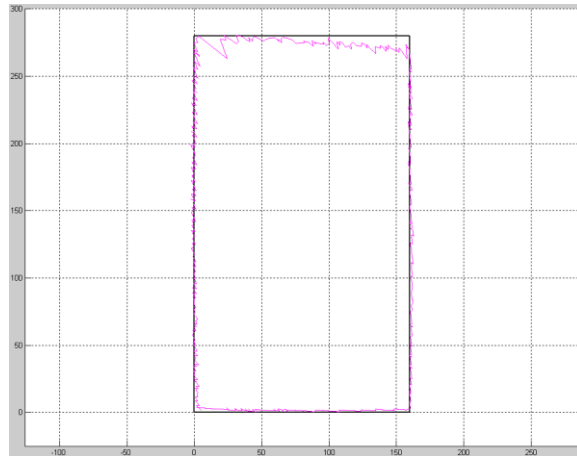


(b)

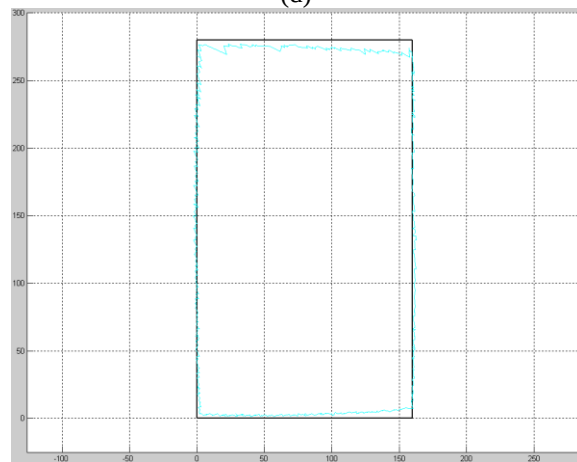


(c)

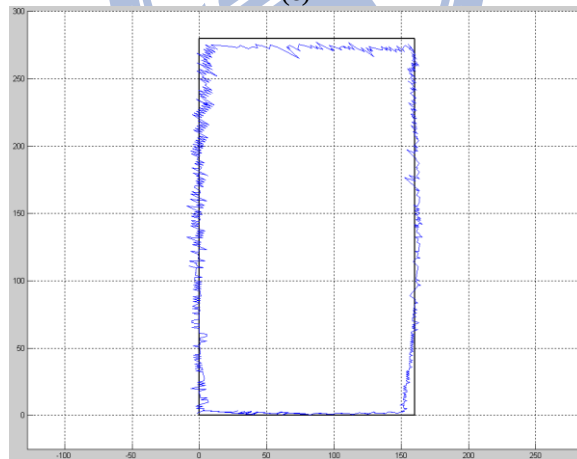
圖 5-4 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。



(d)

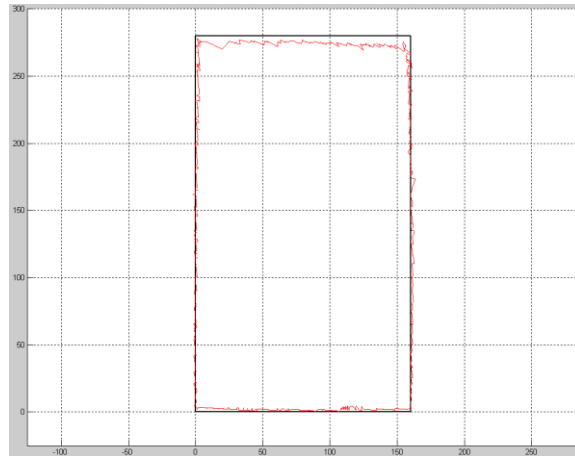


(e)



(f)

圖 5-4 (續)。



(g)
圖 5-4 (續)。

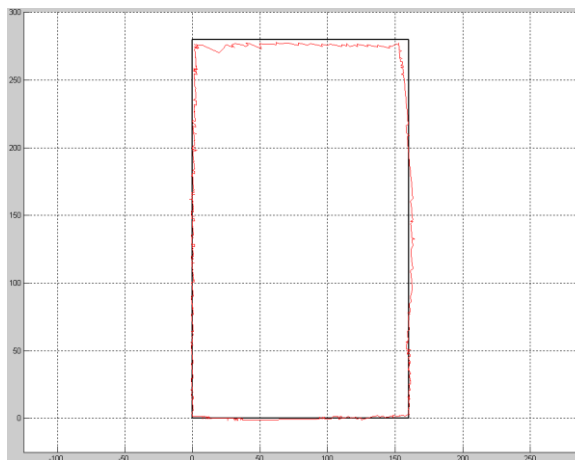


圖 5-5 單多攝影機一起挑選的結果。

接著我們考慮以人物的頭部中心為目標點的定位結果，如圖 5-6 所示，(a)、(c)、(e)和(g)分別為 4 隻攝影機所拍到的影像之一，(b)、(d)、(f)和(h)分別為偵測出來的前景區域，其中紅色點代表人物的頭部中心點。定位結果如圖 5-7、圖 5-8 和圖 5-9 所示，其中圖 5-7 的(a)-(d)為只考慮單攝影機的定位結果，(e)為參考單攝影機的誤差分析，經由挑選後的定位結果。圖 5-8 的(a)-(f)為考慮多攝影機(攝影機配對)的定位結果，(g)為參考多攝影機的誤差分析，經由挑選後的定位結果。圖 5-9 則為參考單攝影機和多攝影機的誤差分析，一起挑選後的定位結果。由實驗結果可看出，整合多攝影機且經由挑選過後的定位結果比單攝影機的定位結果準確。

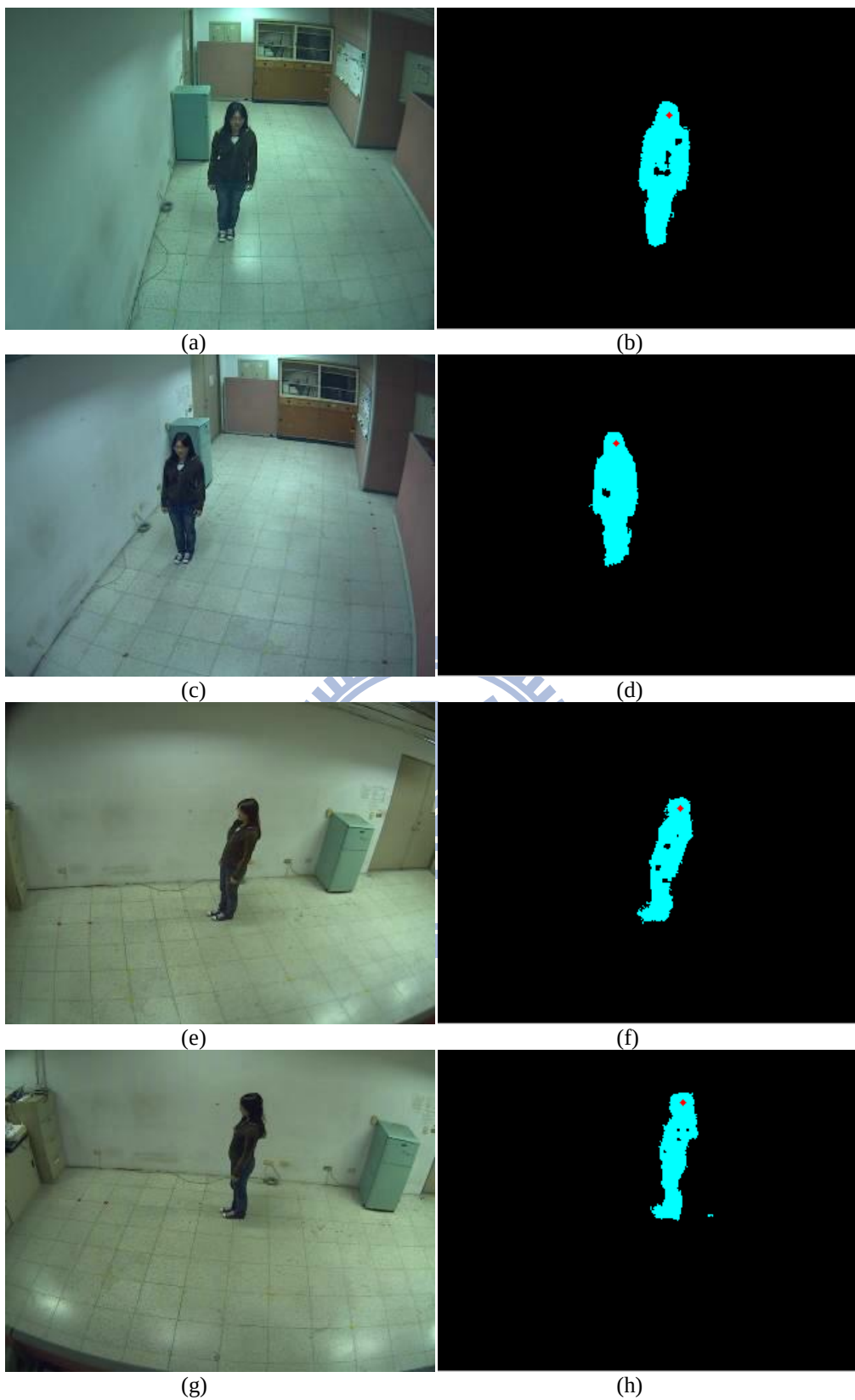
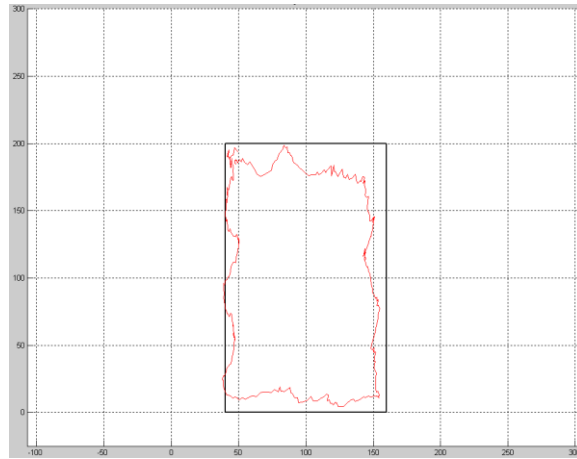
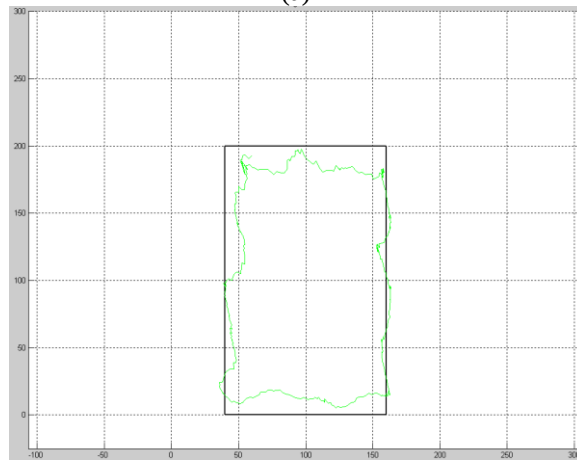


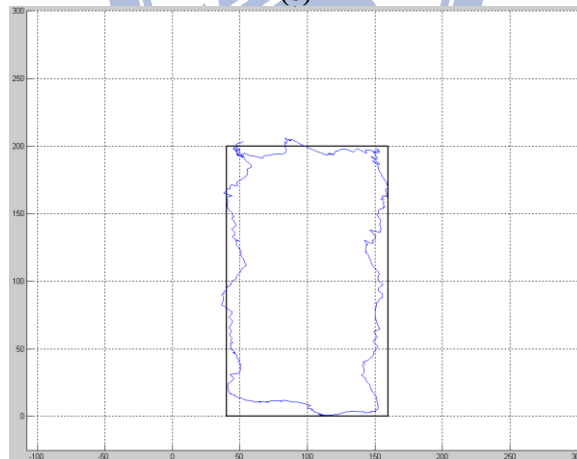
圖 5-6 實際場景—人物。



(a)

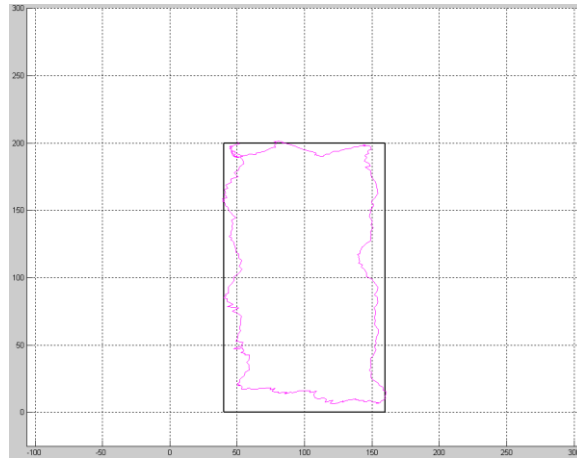


(b)

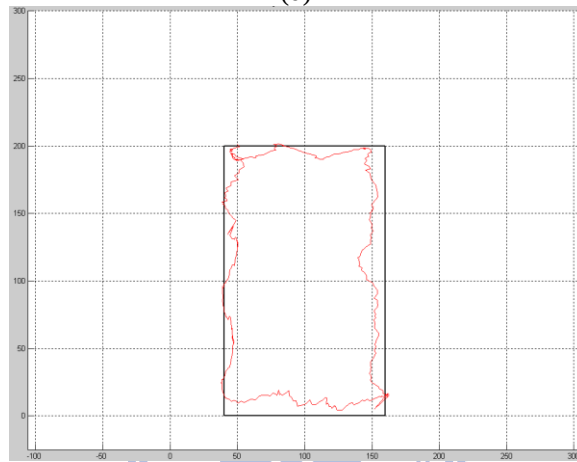


(c)

圖 5-7 單攝影機的定位結果。(a) 攝影機 1 的結果，(b) 攝影機 2 的結果，(c) 攝影機 3 的結果，
(d) 攝影機 4 的結果，(e) 挑選過後的結果。

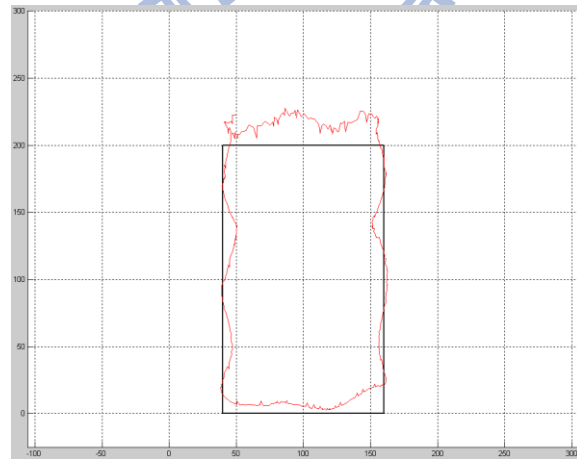


(d)



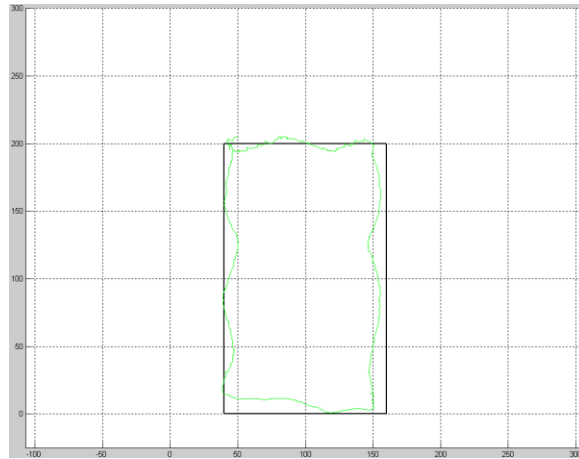
(e)

圖 5-7 (續)。

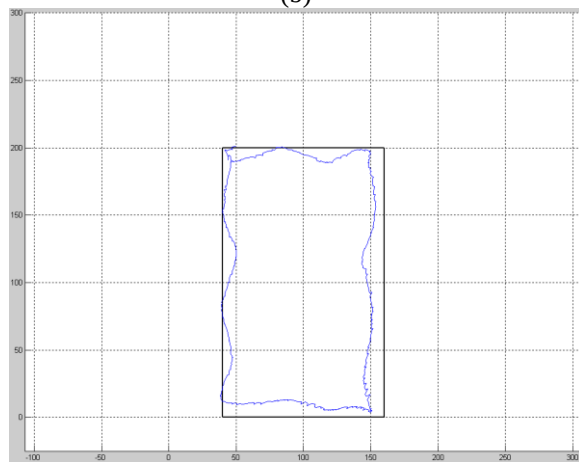


(a)

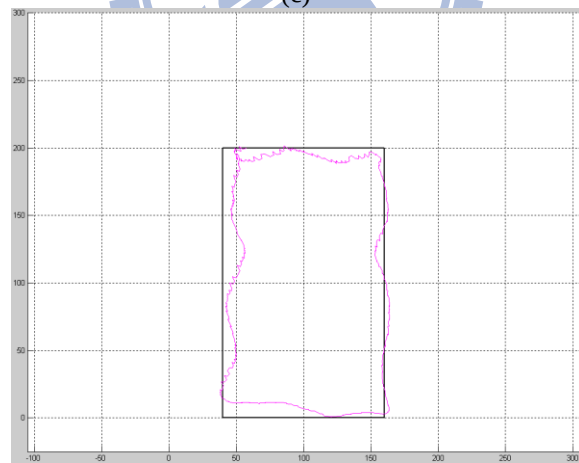
圖 5-8 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。



(b)

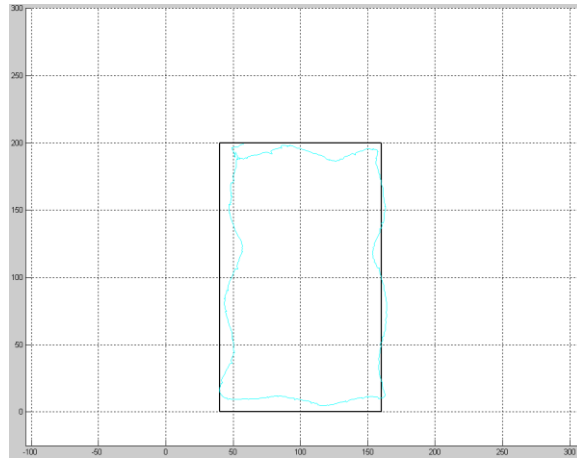


(c)

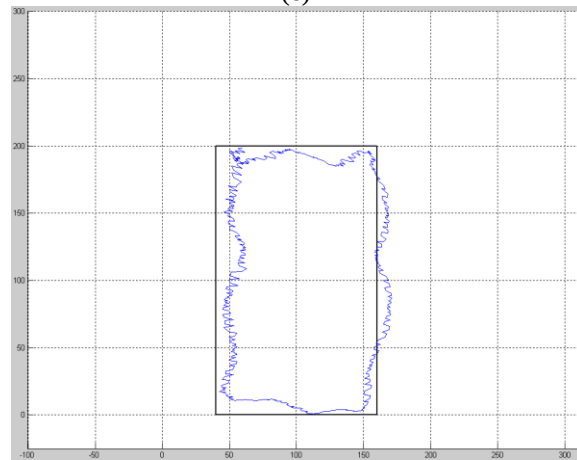


(d)

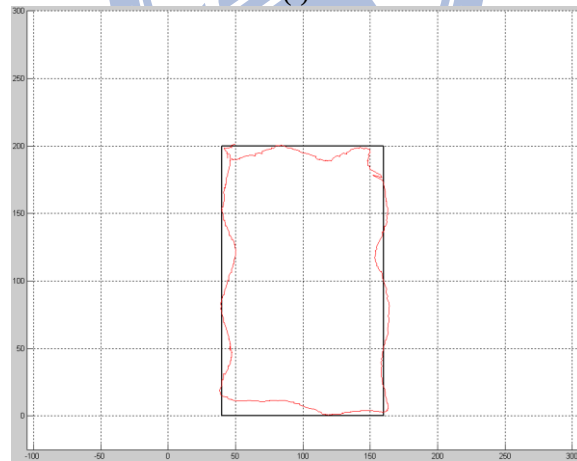
圖 5-8 (續)。



(e)



(f)



(g)

圖 5-8 (續)。

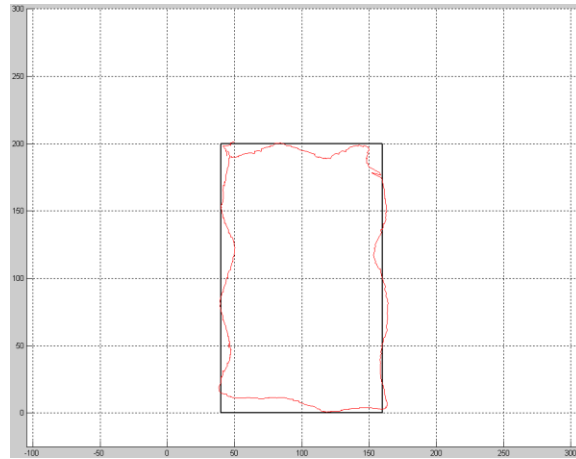
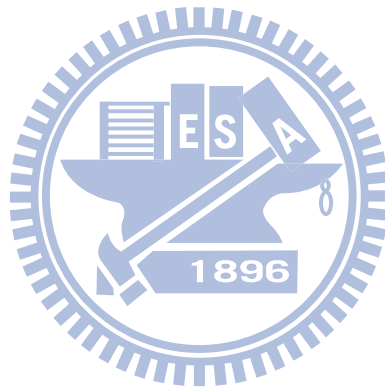


圖 5-9 單多攝影機一起挑選的結果。



第六章 結論與未來展望

6.1 結論

在本論文中，我們提出一套定位方法。基於電腦視覺，先在影像上擷取目標點，以單應性轉換關係，計算出目標點在真實世界中的位置，即只要取得目標點，即可作定位。在實驗的過程中，我們發現，當攝影機和目標點皆靜止不動時，定位的結果仍與實際位置有所誤差。因此，我們探討可能造成定位誤差的來源，並加以改善。除此之外，我們也模擬取像產生的雜訊，並分析其對定位結果造成的定位誤差分佈，依此當作信心區域。當我們欲整合多攝影機的定位結果時，根據信心區域，我們可以挑合適的攝影機配對，以得到較佳的定位結果。由虛擬場景和實際場景的實驗結果可看出，利用誤差分析的結果能使我們成功地整合多攝影機資訊，並取得較為準確且穩定之定位結果。

6.2 未來展望

本論文特色在於只要取得目標點，即可對其作定位。也因為這個原因，當我們的無法取得正確的目標點時，定位的結果勢必也會不準確。另外，我們所使用的定位方法中，有用到兩個限制，一個是需要知道攝影機的 3 維中心位置，另一個是需要知道目標點的高度資訊。當這兩個假設其中一個不成立時，我們的定位方法便無法正確執行。因此，若能自動算出攝影機的 3 維中心位置和目標點的高度資訊，搭配上我們提出的定位方法，便可得到準確的定位結果。

參考文獻

- [1] W. Hu, T. Tan, L. Wang, and S. Maybank, “A Survey on Visual Surveillance of Object Motion and Behaviors,” IEEE Transactions on Systems, Volume 34, Issue 3, pp. 334-352, 2004.
- [2] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, “Pfinder: Real-Time Tracking of the Human Body,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 19, Issue 7, pp. 780-785, 1997.
- [3] Y. Li, C. Huang, and R. Nevatia, “Learning to Associate: HybridBoosted Multi-Target Tracker for Crowded Scene,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 2953-2960, 2009.
- [4] Q. Yu, G. Medioni, and I. Cohen, “Multiple Target Tracking Using Spatio-Temporal Markov Chain Monte Carlo Data Association,” IEEE transactions on Pattern Analysis and machine Intelligence, Volume 31, Issue 12, pp. 2196-2210, 2009.
- [5] M. Han, W. Xu, H. Tao, and Y. Gong, “An Algorithm for Multiple Object Trajectory Tracking,” IEEE Conference on Computer Vision and Pattern Recognition, Volume 1, pp. 864-871, 2004.
- [6] I. Pavlidis, V. Morellas, and S. Harp, “Urban Surveillance Systems: From the Laboratory to the Commercial World,” Proceedings of the IEEE, Volume 89, Issue 10, pp. 1478-1497, 2001.
- [7] L. Lee, R. Romano, and G. Stein, “Monitoring Activities From Multiple Video Streams : Establishing a Common Coordinate Frame,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, Issue 8, pp. 758-767, 2000.
- [8] G. P. Stein, “Tracking from Multiple View Points: Self-Calibration of Space and

- Time,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 1, pp. 521-527, 1999.
- [9] A. C. Sankaranarayanan and R. Chellappa, “Optimal Multi-View Fusion of Object Locations,” IEEE Workshop on Motion and video Computing, pp. 1-8, 2008.
- [10] J. P. Barreto and K. Daniilidis, “Wide Area Multiple Camera Calibration and Estimation of Radial Distortion,” In Proceedings of the Workshop on Omnidirectional Vision and Camera Networks, pp. 1-5, 2004.
- [11] Z. Zhang, “A Flexible New Technique for Camera Calibration,” Technical Report, Microsoft Research, 1998.
- [12] A. Mittal and L. S. Davis, “M2Tracker: A Multi-View Approach to Segmenting and Tracking People in a Cluttered Scene Using Region-Based Stereo,” Proceedings of European Conference on Computer Vision, Volume 2359, pp. 18-33, 2002.
- [13] W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, “Principal Axis-Based Correspondence Between Multiple Cameras for People Tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 28, Issue 4, pp. 663-671, 2006.
- [14] K. Kim and L. S. Davis, “Multi-camera Tracking and Segmentation of Occluded People on Ground Plane Using Search-Guided Particle Filtering,” Proceedings of European Conference on Computer Vision, Volume 3953, pp. 98-109, 2006.
- [15] S. M. Khan and M. Shah, “A Multiview Approach to Tracking People in Crowded Scenes using a Planar Homography Constraint,” Proceedings of European Conference on Computer Vision, Volume 4, pp. 133-146, 2006.
- [16] S. W. Sun, H. Y. Lo, H. J. Lin, Y. S. Chen, F. Huang, and H. Y. M. Liao, “A

Multi-camera Tracking System That Can always Select A Better View to Perform Tracking,” Annual Summit and Conference on Asia-Pacific Signal and Information Processing Association, pp. 373-379, 2009.

- [17] R. Eshel and Y. Moses, “Homography Based Multiple Camera Detection and Tracking of People in a Dense Crowd,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2008.

