

行政院國家科學委員會補助專題研究計畫  成果報告  
 期中進度報告

基於尺寸不變性之穩健式三維頭部追蹤  
**Robust 3D Tracking of Human Head Based on Invariance of Size**

計畫類別： 個別型計畫  整合型計畫

計畫編號：NSC 98-2221-E-009-124-MY2

執行期間：2009年8月1日至2011年10月31日

計畫主持人：莊仁輝

共同主持人：

計畫參與人員：羅國華、郭育璋、余東諺、陳芝穎、盧沛怡、陳怡廷、  
陳暉晴、馬國濂、周致傑、姚柏安、王之容、廖偉成、黃星陸

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、  
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年  二年後可公開查詢

執行單位：

中 華 民 國 99 年 8 月 30 日

# 行政院國家科學委員會專題研究計畫成果報告

## 基於尺寸不變性之穩健式三維頭部追蹤

### Robust 3D Tracking of Human Head Based on Invariance of Size

計畫編號：NSC 98-2221-E-009-124-MY2

執行期限：2009年8月1日至2011年10月31日

主持人：莊仁輝 國立交通大學資訊工程學系(所)

**中文摘要:**(關鍵字：頭部偵測、人員追蹤、人員定位、快速定位誤差分析、遮掩、多攝影機資訊整合)

「基於尺寸不變性之穩健式三維頭部追蹤」二年計畫的主要目的為，利用高角度的攝影機對監控的場景進行人物頭部偵測，並透過投影幾何的平面轉換對頭部中心進行定位，以建構在單攝影機下之人物定位系統。而當環境中具有多攝影機時，系統可透過選擇性與整合性的多攝影機資訊融合，來整合多攝影機的定位資訊，以達成準確的定位。在第一年的研究計畫中，我們發展一套頭部偵測方法，可穩定地偵測出影像中人物的頭部區域。並且，我們也發展了一套不需要複雜的校正之影像特徵點定位方法，可對人物頭部中心點進行定位。在頭部偵測方面，我們利用兩種主要方法偵測人物頭部，分別為基於髮膚色特徵和前景物輪廓區域最高點之偵測演算法。首先我們利用高斯混合模型分割出人員所在的前景區塊，再藉由偵測各前景區塊中的髮膚色區域，過濾出候選頭部區域。另一方面我們找出前景物輪廓的區域最高點後，由此點附近找出候選頭部區域。最後藉由整合髮膚色特徵和區域最高點所找出之候選頭部區域，即可正確地偵測出場景中的人物頭部中心點位置，並可被作為一影像特徵點進行後續人物定位。在影像特徵點定位的方面，我們可以藉由單一攝影機，在已知特徵點高度以及攝影機位置的條件下，透過投影幾何的平面轉換和三維幾何關係，計算出特徵點在其高度平面上之位置。實驗的結果顯示頭部偵測的結果應用於多人的場景中，仍有良好的正確性，而特徵點定位也能落在一定的誤差範圍之內。其後第二年的研究我們分析特徵點定位過程中產生的誤差，例如徑向畸變誤差、人為測量誤差和取像誤差，並利用誤差分析的結果來整合多攝影機資訊，使得系統的定位結果能夠為穩定、準確。

**英文摘要:** (Keyword: Head Detection, People Tracking, Localization, Rapid Error Analysis for Localization, Occlusion, Multi-Camera Information Fusion)

The goal of the two-year project, "Robust 3D Tracking of Human Head Based on Invariance of Size," is to locate people based on head detection and homographic transformation. Not only can each camera in the monitored environment locate people independently, the localization results of the multiple cameras can also be fused to achieve more accurate localization. In the first year, a robust head detection method is developed to obtain the center of human head in images. The result is then used by the proposed feature point-based people localization scheme without complex calibration of cameras. The above head detection is achieved by two algorithms based on (i) detection of hair/skin color and (ii) identification of local maximum of a contour, respectively, both for a foreground region segmented using the Gaussian mixture model. From the foreground regions, head candidates are first found separately by (i) and (ii). Then, centers of head are found by integrating the two

sets of head candidates, and are regarded as feature points in subsequent people localization. The proposed localization scheme assumes that the height of a feature point is known a priori so that the point location in the plane of the known height can be found by homographic transformation and 3D geometry which only require camera locations. Experimental results show that the proposed head detection is performed satisfactorily even for crowded scene. However, the localization results are not always accurate enough due to errors in the localization process. The errors may arise from radial distortion of cameras, inaccuracy associated with manual measurement, and noises in the imaging process. In the second year, we analyze these errors and use the analytic results to fuse multi-camera information so as to reduce the localization error and improve system stability.

## 簡介

本計畫的主要目的是以高角度的攝影機對場景進行監控，藉由對人物頭部偵測找出人物的頭部中心點後，再以投影幾何的平面轉換對頭部中心進行持續定位。當環境中具有多攝影機時，系統可透過選擇性與整合性的多攝影機資訊融合，來整合多攝影機的定位資訊，以達成準確的定位。此兩年計畫的研究成果，共包含下列三大主題：

### 一、穩健式之人頭偵測研究：

目標是以多種特徵進行人物頭部的偵測，包含髮膚色資訊、前景資訊，以及特徵間的幾何關係進行穩健的頭部偵測，以供後續進行定位與誤差分析。

### 二、基於影像特徵點定位之人員定位：

以頭部中心點，以單一攝影機並利用投影幾何的平面轉換對人物進行定位，並持續產生俯視圖定位結果。

### 三、基於影像特徵點定位之人員定位之誤差分析：

針對單一攝影機所取得的目標點，進行誤差分析以及穩定程度的估計，並利用估計的結果，整合多攝影機的資訊，提升定位的準確性與穩定度。

在第一年的研究計畫中，我們發展一套頭部偵測方法，可穩定地偵測出影像中人物的頭部區域。並且，我們也發展了一套不需要複雜的校正之影像特徵點定位方法，可對人物頭部中心點進行定位。在頭部偵測方面，我們利用兩種主要方法偵測人物頭部，分別為基於髮膚色特徵和前景物輪廓區域最高點之偵測演算法。首先我們利用高斯混合模型分割出人員所在的前景區塊，再藉由偵測各前景區塊中的髮膚色區域，過濾出候選頭部區域。另一方面我們找出前景物輪廓的區域最高點後，由此點附近找出候選頭部區域。最後藉由整合髮膚色特徵和區域最高點所找出之候選頭部區域，即可正確地偵測出場景中的人物頭部中心點位置，並可被作為一影像特徵點進行後續人物定位。在影像特徵點定位的方面，我們可以藉由單一攝影機，在已知特徵點高度以及攝影機位置的條件下，透過投影幾何的平面轉換和三維幾何關係，計算出特徵點在其高度平面上之位置。實驗的結果顯示頭部偵測的結果應用於多人的場景中，仍有良好的正確性，而特徵點定位也能落在一定的誤差範圍之內。

其後第二年的研究我們分析特徵點定位過程中產生的誤差，例如徑向畸變誤差、人為測量誤差和取像誤差，透過影像校正與空間測量點自動修正，以減少後續定位的誤差。後續則依據特徵點位於影像中不同的位置，動態地估計出定位可能產生的動態誤差，再利用這樣的資訊整合多攝影機資訊，使得系統的定位結果能夠穩定、準確。

## 主題一：穩健式之人頭偵測研究

### 1.1 研究目的

在常見視訊監控系統中多採用高角度的監控攝影機進行攝影，原因是即使影像中出現人群遮掩的狀況，大多數頭部區域仍可以被清楚地觀察到。因此，我們希望透過頭部偵測，判定影像中人物的所在位置，使得後續能對頭部進行定位與追蹤。因此本計畫的首要研究主題，即在於能夠有效地找出監控影像中的人物頭部區域。

### 1.2 文獻探討

頭部偵測相關的偵測方法包含臉部區域偵測與頭部區域偵測兩類。在臉部區域偵測中，主要是透過系統化的機器學習方式，找出最具代表性的特徵，再搭配 Boosting 分類器就可以快速有效的偵測出人臉的區域[1]。類似的人臉偵測架構，還可以擴展到人員頭肩部與軀體的偵測[2]，達到在遮蔽的情況下，偵測出人員的效果；而頭部區域偵測的方法，則常透過多種特徵資訊來找出可能的頭部區域，如頭的邊緣形狀、顏色、膚色等，在[3]中，即提出以頭至肩部的外形來作偵測，[4]則再多考慮了人物的頭部頂點作為偵測的線索，而[5, 6]則是利用橢圓形的特性來搜尋或比對出頭部區域，在[7]中，則是透過 motion difference 的找出可能發生移動的圓形頭部區域，並利用 Hough transformation 來定出頭部的位址。在本計畫中，我們採用頭部顏色、形狀等資訊作為偵測的基礎，並整合頭部尺寸大小在時序上有緩慢變化的特點，發展一個快速而有效的頭部偵測方法。

### 1.3 研究方法

我們利用兩種主要的特徵，分別為髮膚色區域與區域最高點。首先我們利用高斯混合模型分割出人員所在的前景區塊，再分別以此兩項特徵各自找出可能的頭部區域，再將結果進行整併，以達成完整的頭部偵測。

#### 1.3.1 髮膚色區域偵測

為了有效降低光線變化所帶來的影響，我們選定 YCbCr 色彩空間來偵測前景區中的膚色的區域，透過多次的實驗，我們選定一個合適的膚色範圍為  $77 \leq Cb \leq 127$ ， $133 \leq Cr \leq 173$ ，而得到的髮色範圍為  $Y \leq 50$ 。針對一室外場景，如圖 1-1(a)所示，進行髮膚色偵測，而所偵測出的髮色與膚色區域顯示於圖 1-1(b)與(c)。

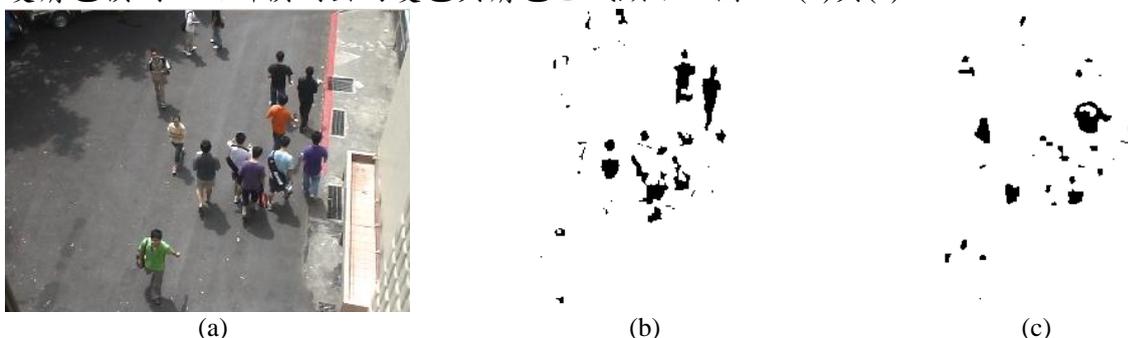


圖 1-1 利用顏色資訊過濾之結果，(a) 輸入影像，(b) 髮色區域，(c) 膚色區域。

接著我們利用連通元件標記法(connected component labeling)，偵測出髮色與膚色區

域中的各個不相連的獨立區域。其結果如圖 1-2 中的 (a)與(b)所示，我們以不同的顏色顯示這些區域。我們可以看到圖中仍有很多非頭部的區域，因此我們將以過濾的方式移除這些區域。



圖 1-2 元件連通標記法結果，(a) 圖 1-1(b)執行後之結果，(b) 圖 1-1(c)執行後之結果。

### 1.3.2 以過濾方式篩選出髮膚色的頭部候選區域

接著我們利用頭部的其他特徵，包含：(1)頭部大小有一定之範圍，(2)頭部形狀近似球形，(3)頭部位於身體最高處，故其下方具有足量之前景區域。針對上述之各項特徵，我們提出了以下幾種過濾方式，針對特徵(1)採用面積大小過濾，針對特徵(2)採用長寬比過濾，針對特徵(3)檢查下方需具足量前景區域，將依序介紹各種過濾的方式如下。

#### 方法 A. 面積大小

針對頭部面積所給予的限制，我們設定了兩個門檻值，來排除掉面積過大或過小的髮膚色區域，透過這樣的方式，影子與黑色衣服等雜訊區域將會被移除。

#### 方法 B. 長寬比

雖然理論上頭部的區域會接近球形，但經過觀察發現我們所偵測出的髮膚色區域中，頭部的區域並非呈現正圓形，因此我們保留長寬比例在一定範圍內的髮膚色區域。我們使用式 1.1 來計算候選區域的長寬比：

$$R^* = H/W \quad (1.1)$$

如果  $R^*$ 值越大，代表此候選區域越細長。本系統中，只要長寬符合  $0.8 \leq R^* \leq 2$ ，我們便保留此候選區域。

#### 方法 C. 下方具足量前景區域

通過上述兩種過濾方式後，我們將更進一步判斷這些區域下方是否有足量之前景區域，藉此判斷其是否為頭部位置所在。我們採用的方法是以區域之質心作為基準點，以垂直方向向下搜尋前景區域的輪廓邊緣，如在一距離內找尋不到輪廓邊緣，代表此區域下方有足量前景區域。如圖 1-3 所示，紫色框線為髮色區域示意圖，我們由紫色框線之質心為基準點，並以黃色箭頭的方向向下搜尋輪廓邊緣，如果搜尋到輪廓邊緣則表示下方前景區域不足，將會被視為非頭部區域而予以刪除。

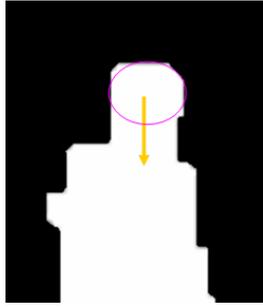


圖 1-3 判斷候選區域下方是否有足量前景示意圖。

利用上述方法 A、B、C 可將大部份之非頭部區域刪除，以圖 1-4 為例，圖 1-4 (a) 與(d)為輸入影像，圖 1-4 (b)與(e)為髮膚色過濾後之前景影像，圖 1-4 (c)與(f)為執行頭部特徵偵測演算法後之結果。我們可以看到一部分的頭部區域已成功偵測出來，但仍有許多未偵測到，因此我們下面將利用區域最高點偵測剩餘的頭部。

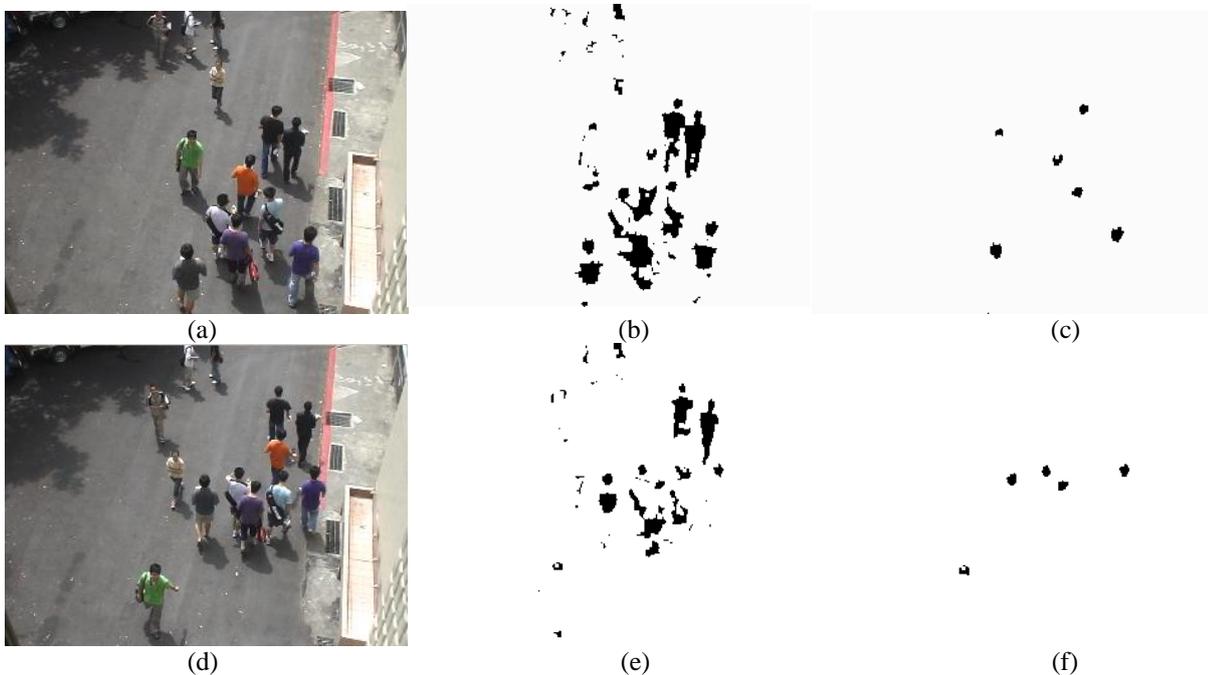


圖 1-4 以髮膚色特徵為基礎之偵測頭部演算法，(a)(d) 輸入影像，(b)(e) 髮膚色過濾後之前景影像，(c)(f) 圖(b)(e)執行頭部特徵偵測演算法後之結果。

### 1.3.3 以區域最高點偵測出頭部候選區域

由於頭部為人體的頂端，因此在許多應用情境中，我們可以假設一個輪廓之最高點將對應到頭部的位置。但因為一個前景區域內部可能包含了多個人，所以我們利用偵測輪廓區域最高點的方式，找出此前景區域輪廓上可能的人物的頂點。我們提出一個前景物輪廓區域最高點之偵測演算法(local maximum of foreground height, LMFH)，具有三大步驟，首先是找出每個前景區域的輪廓點，接著對這些輪廓點，從前景物輪廓的最高點當作起始點往終點執行區域最高點的偵測，最後再反向(由終點往起始點)執行區域最高點之偵測。令座標原點如圖 1-5 所示，我們利用式子 1.2 來偵測出輪廓的區域最高點，並用式子 1.3 來更新累計高度值，

$$P_{t+1}(y) - P_t(y) < 0 \text{ and } FH_t \geq T_y \quad (1.2)$$

$$\begin{aligned} FH_{t+1} &= P_{t+1}(y) - P_t(y) + FH_t, \text{ if } P_{t+1}(y) - P_t(y) \geq 0 \\ FH_{t+1} &= 0, \text{ otherwise} \end{aligned} \quad (1.3)$$

其中  $P_t(y)$  為第  $t$  個輪廓點的  $y$  座標值、 $FH_t$  為累計高度值 (foreground height) 而  $T_y$  為系統設定的門檻值，當式子 1.2 成立時， $P_t(x,y)$  即為區域最高點，系統設定  $T_y=2$ ，先將輪廓上所有區域最高點 (起伏點) 找出，再利用後續演算法排除非頭部位置的區域最高點。

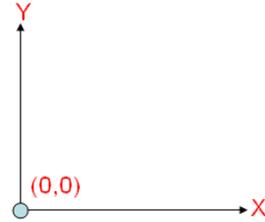


圖 1-5 座標原點示意圖。

找出區域最高點後，我們可以利用一些簡單的判斷排除掉明顯不是頭部位置的區域最高點，如圖 1-6 所示，紅色圓圈所在位置皆符合式 1.2 區域最高點之定義，但是明顯不是頭部位置。我們根據先前假設：頭部為人體之頂端，若區域最高點上方為前景，則可能是輪廓中人員之腋下、胯下 (圖 1-6 紅色圓圈) 或是前景切割結果破碎處，故只保留區域最高點上方為非前景區域者。此外，利用顏色資訊對保留下來的區域最高點再次進過濾，我們偵測此區域最高點下方  $3 \times 3$  區塊之顏色範圍，判斷其是否落在髮色與膚色之範圍，若不符合則進行移除。

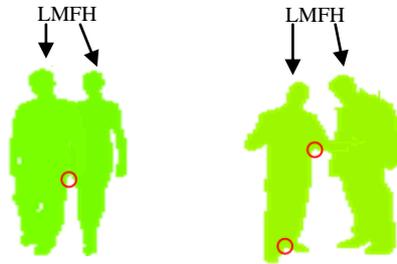


圖 1-6 LMFH 之偵測以及需刪除之位置示意圖。

經由實驗發現，同一輪廓由正向及反向所搜尋到的區域最高點可能會略有差異，以圖 1-7 為例，假設灰色部分為某前景區域，而綠點為輪廓點，若我們由左到右執行 LMFH 演算法，我們會發現找到的累計高度值 (RFH) 都很小以至於無法偵測出區域最高點。反之，若我們由右到左執行 LMFH 演算法則找到的累計高度值 (LFH) 明顯會超過  $T_y$  值，這也是為什麼同一個輪廓要執行兩次 LMFH 演算法的原因。因此，不同方向執行 LMFH 演算法都會產生各自之結果，所以必須檢查兩個不同方向的結果，而檢查的方式如式子 1.4 所示，同一點必須在兩個不同方向都被判定成區域最高點才選出，而像是圖 1-7 中只有單方向被判定成區域最高點的輪廓點將會被濾除。

$$RFH > T_y \text{ and } LFH > T_y \quad (1.4)$$

圖 1-8 為 LMFH 演算法執行結果，圖 1-8(a)為輸入影像，圖 1-8 (b)為執行結果，系統所偵測出之前景物輪廓區域最高點以紅圓圈出。LMFH 演算法之優點為計算量少，缺點為無法偵測到位於多人區塊內部的人物，故前景區域內部的頭部區域可由前述的髮膚色特徵偵測出來。

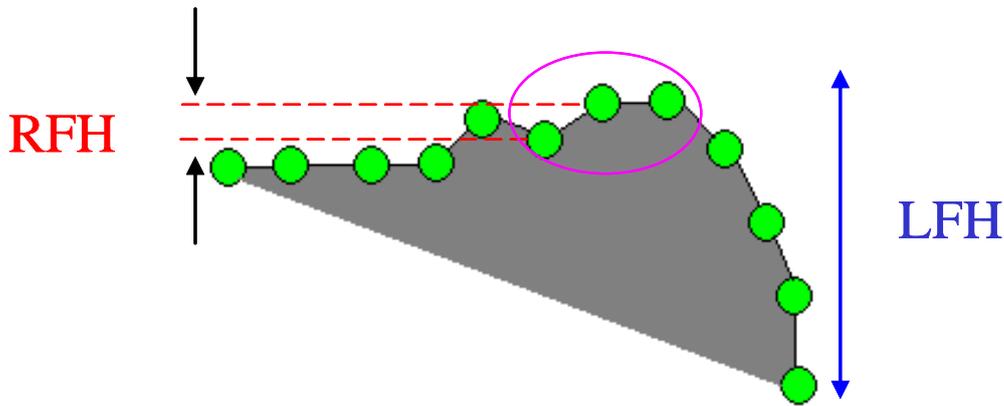


圖 1-7 執行不同方向的 LMFH 演算法之差別。



圖 1-8 前景物輪廓區域最高點之偵測，(a) 輸入影像，(b) LMFH 之偵測結果。

#### 1.3.4 整合髮膚色區域與區域最高點之頭部候選區域

目前我們所偵測出的兩種候選區域如圖 1-9 所示，圖中綠色圈代表髮膚色特徵偵測之結果，而紅色圈代表區域最高點偵測之結果，我們發現圖中有不少地方出現紅色圈與綠色圈重疊之現象，這代表同一頭部被重複地偵測，因此我們的目的就是將紅色圈與綠色圈的資訊進行整合處理。

每個頭部候選區域依照其在影像上之位置會有對應的半徑，我們假設有兩頭部候選區域如圖 1-10 所示，令紅色圈圓心為  $O_1$  且其半徑長度為  $R_1$ ，綠色圈圓心為  $O_2$  且其半徑長度為  $R_2$ ，系統利用式子 1.5 來判斷是否發生重疊之現象。若發生重疊，將原先之頭部候選區域刪除，並取  $O_1$ 、 $O_2$  之中點以及  $R_1$ 、 $R_2$  之平均作為新頭部候選區域的圓心  $O_3$  及半徑  $R_3$ ，以取代原先之重疊頭部候選區域，如圖 1-11 所示。依照上述方法將重疊部分整合後之結果如圖 1-12 所示， $O_1$ 、 $O_2$  和  $O_3$  為找到的頭部中心點，並以此頭部中心點為圓心，對頭部候選區域畫上紅色圓圈標示。

$$\overline{O_1O_2} < R_1 + R_2 \quad (1.5)$$



圖 1-9 頭部候選區域偵測之結果(整合前)。

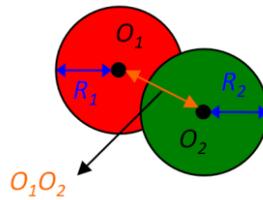


圖 1-10 頭部候選區域示意圖。

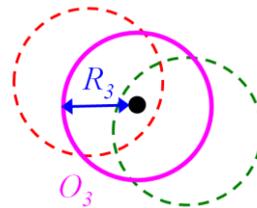


圖 1-11 整合重疊的頭部候選區域示意圖。

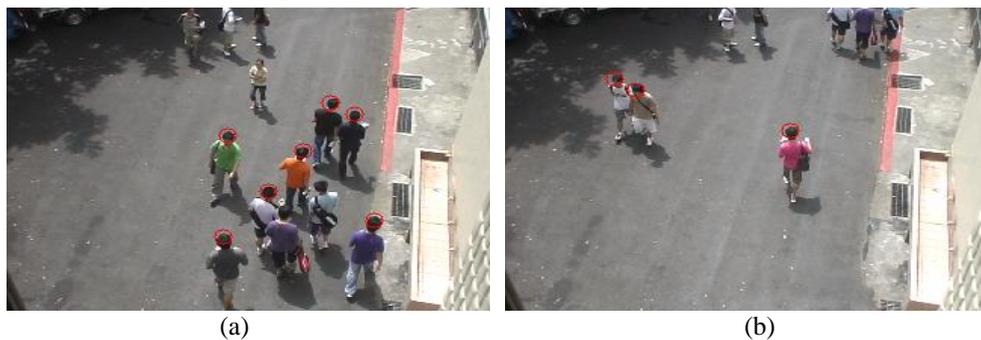


圖 1-12 頭部候選區域偵測之結果(整合後)。

#### 1.4 研究結果與討論

在實驗的部分，我們以三個不同的室外場景進行試驗。圖 1-13 為執行頭部偵測之結果，其中黃色圈為偵測到的頭部位置，而在每張影像的左上角則統計出該張影像的人(頭)數，以圖 1-13 為例，其中圖 1-13(a)、(b)、(c)、(e)、(f)中頭部的偵測都相當正確，僅有圖 1-13(d)、(f)發生了偵測失誤。紅色箭頭所指為未偵測到之人員，由於這些人員所在位置皆於輪廓內部且頭部顏色特徵不明顯，因此無法有效的偵測出頭部區域。而圖 1-13(d)中藍色箭頭所指人員為因為手持雨傘而導致偵測失誤，雨傘內部有兩位人員，但是經由頭部偵測演算法只偵測到一位。

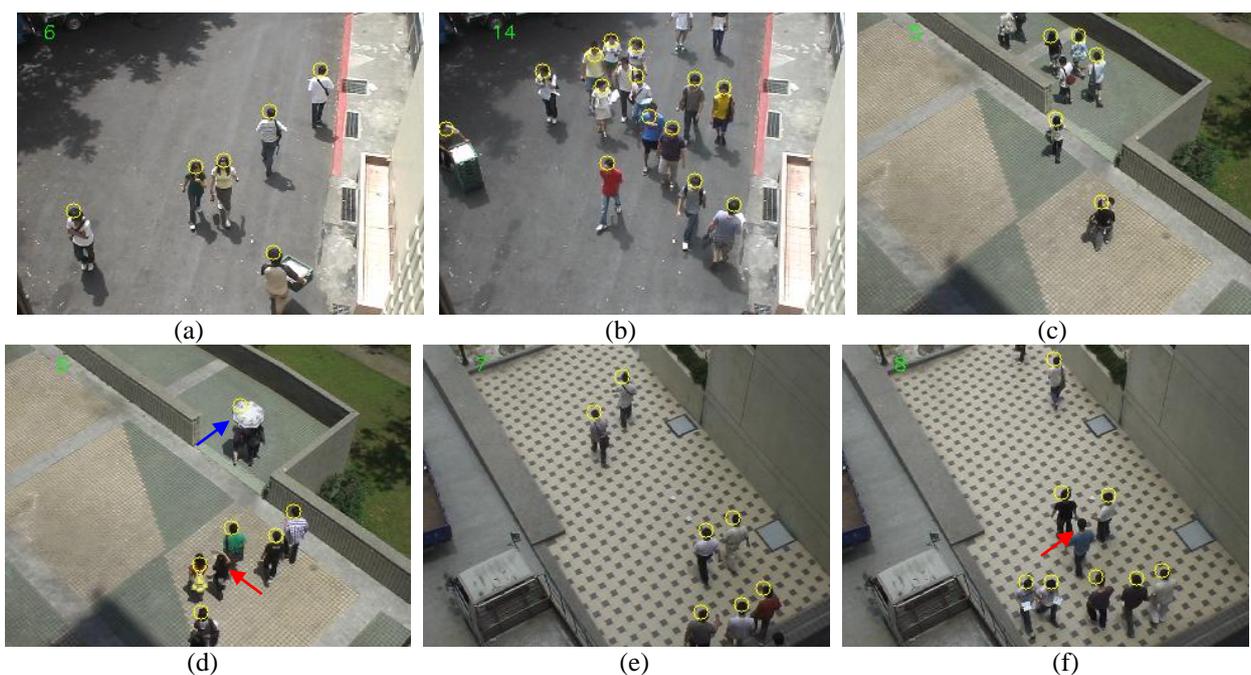


圖 1-13 不同場景中執行頭部偵測之結果。

## 1.5 參考文獻

- [1] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, pp. 137-154, 2004.
- [2] B. Wu and R. Nevatia, "Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors," *International Journal of Computer Vision*, vol. 75, pp. 247-266, 2007.
- [3] H. Yoon, D. Kim, S. Chi, and Y. Cho, "A robust human head detection method for human tracking," *IEEE Conference Intelligent Robots and Systems*, pp. 4558-4563, 2006.
- [4] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," *IEEE Conference Computer Vision and Pattern Recognition*, vol. 2, pp. II - 459-66, 2003.
- [5] S. Birchfield, "An elliptical head tracker," *IEEE Conference Signals, Systems & Computers*, vol. 2, pp. 1710-1714, 1997.
- [6] Z. Zhang, H. Gunes, and M. Piccardi, "An accurate algorithm for head detection based on XYZ and HSV hair and skin color models," *International Conference on Image Processing*, pp. 1644 -1647, 2008.
- [7] S. S. Ghidary, Y. Nakata, T. Takamori, and M. Hattori, "Human detection and localization at indoor environment by homerobot," *IEEE International Conference Systems, Man, and Cybernetics*, vol. 2, pp. 1360-1365, 2000.

## 1.6 計畫成果自評

人物頭部偵測之結果雖然仍非完美，除了因人物頭部被遮蔽在人群內部或手持雨傘等因素而無法成功將人物頭部偵測出來外，基本上已經是可以接受的結果，大致上能初步提供給後續人員定位系統使用，未來可以針對不同的情形設計新的方法，進一步地提升其偵測效率。

## 主題二：基於影像特徵點定位之人員定位

### 2.1 研究目的

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。因此我們實作一套基於電腦視覺的定位系統，該系統的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。由於我們的定位系統是以點做為定位目標(point-based)，因此只要存在目標點即可進行定位。至於目標點的選擇可以是立足點、頭部中心點等。

### 2.2 文獻探討

在[1]中，Mittal 和 Davis 提出以立體視覺(stereo vision)為基礎的多攝影機追蹤模式，用來追蹤人群，透過機率模型決定影像中的每個像素分別是屬於哪一個被追蹤人員；在[2]的方法中，則透過立體視覺資訊，進一步作空間中人物的定位與追蹤；對於[1, 2]的方法，由於需要攝影機之間事先進行校正，故在實際應用上較為不便。在[3, 4]的追蹤研究中，則提出以「頂視監控」的概念來進行人群追蹤；類似的頂視追蹤概念在[5, 6]中亦有探討，構想是基於影像與頂視面的 homographic transformation，在頂視面中估計人物腳點位置並進行多人追蹤的方法；在[7]中，則整合了攝影機資訊與地板感測器，作人員的定位與追蹤。不同於上述[3-7]的追蹤與定位研究，在此計畫中，我們非僅著眼於透過 homographic transformation 轉換，發展頂視平面的定位追蹤。

### 2.3 研究方法

我們進行影像特徵點定位時需要幾個已知條件，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，第三則是預先計算好的單應性轉換矩陣(homographic matrix)用來將影像上的座標轉換到參考平面上。由於在估算真實空間目標點位置的過程中，我們會利用攝影機的中心位置和目標點的高度，因此在實作的方面，攝影機的中心位置與單應性轉換矩陣可藉由事先測量與計算得到；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，在此我們用頭部中心點當作目標點。

有了這些資訊後，我們即可估算出真實空間中目標點的位置。如圖 2-1 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參考平面  $\pi_1$  上，即得到投影點  $P_{RP}$ ，轉換關係如式子 2.1，其中  $P_{IP}=[u, v, 1]^T$ ， $[x, y]^T$  則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三維連線  $L$  和  $\pi_2$  平面之方程式，其中  $\pi_2$  平面與  $\pi_1$  平面平行且平面高度等於目標點的高度。最後再計算出三維連線  $L$  與  $\pi_2$  的交點  $P_{HP}$ ，此交點即為估算的定位結果。

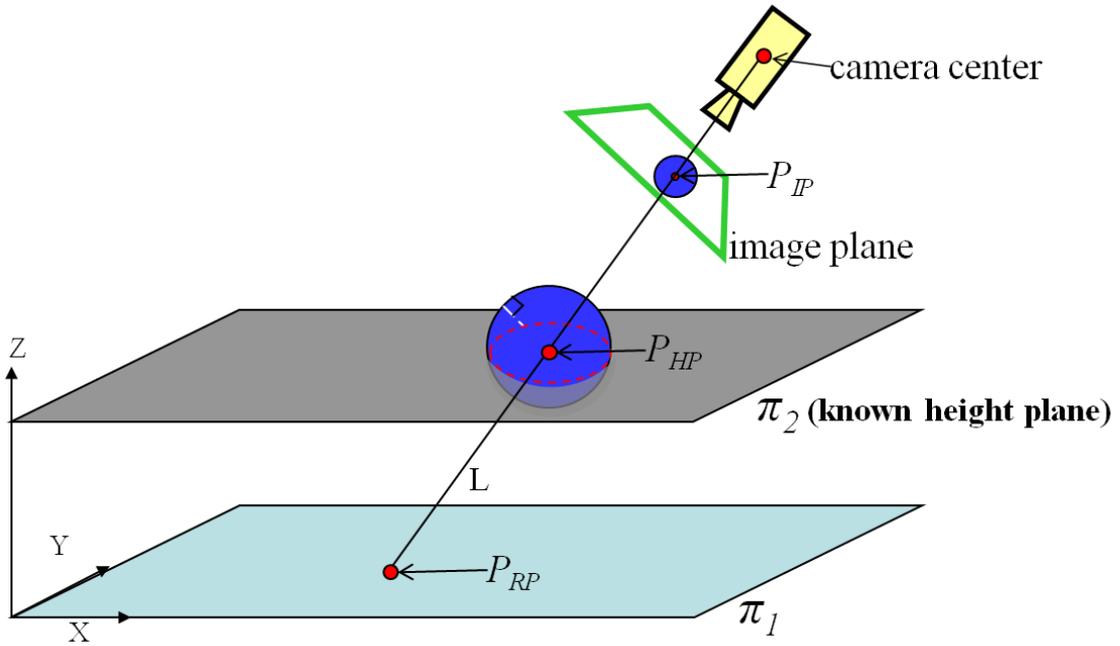
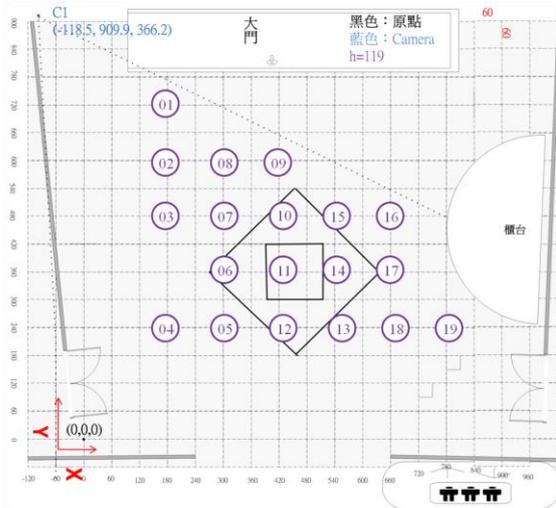


圖 2-1 單攝影機的定位示意圖。

$$H \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} \quad (2.1)$$

## 2.4 研究結果與討論

如圖 2-2 所示，(a)為實驗場景的俯視圖，其中紫色圓圈為用來測量的目標點位置，共有 19 個位置；(b)為場景影像，其中黑色十字對應到(a)中之紫色圓圈內的虛線交叉點，紅色米字為我們選取的參考點，共有 5 個參考點。在此實驗場景中，我們將目標點設在地平面(參考平面)上，即目標點的高度平面等於參考平面。接著利用我們所提出的定位方法，估算出 19 個定位位置。圖 2-3 為定位結果，其中藍色點為估算出來的位置。整體平均誤差為 17.36 公分。



(a) (b)

圖 2-2 實驗場景，(a)場景頂視圖，(b)場景影像。

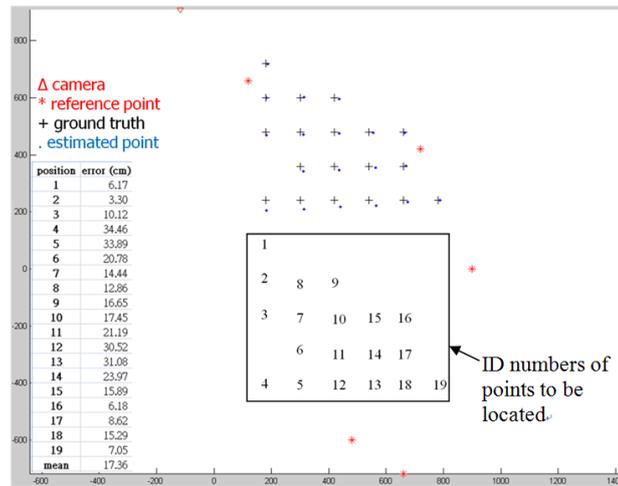


圖 2-3 實驗場景之定位結果。

由於攝影機和目標點皆靜止不動，並且我們設定目標點所在高度平面與參考平面相同，但是卻存在明顯的定位誤差，其誤差來源有徑向畸變誤差(radial distortion)、人為測量誤差和取像誤差，這些誤差將會導致定位的準確度下降。因此我們未來將對定位過程中可能的產生的誤差加以分析與改善，期望減少誤差對於定位的影響。

## 2.5 參考文獻

- [1] A. Mittal and L. S. Davis, "M2Tracker: a multi-view approach to segmenting and tracking people in a cluttered scene," *International Journal of Computer Vision*, vol. 51, no. 3, pp.189-203, 2003.
- [2] S. Bahadori, L. Iocchi, G. R. Leone, D. Nardi, and L. Scozzafava, "Real-time people localization and tracking through fixed stereo vision," *Applied Intelligence*, vol. 26, pp. 83-97, 2007.
- [3] T. Darrell, D. Demirdjian, N. Checka, and P. Felzenszwalb, "Plan-view trajectory estimation with dense stereo background models," *IEEE International Conference Computer Vision*, pp. 628-635, 2001.
- [4] S. V. Martnez, J. F. Knebel, and J. P. Thiran, "Multi-object tracking using the particle filter algorithm on the top-view plan," *European Signal Processing Conference*, 2004.
- [5] S. M. Khan and M. Shah, "A multiview approach to tracking people in crowded scenes using a planar homography constraint," *European Conference Computer Vision*, pp. 133-146, 2006.
- [6] R. Mieziako and D. Pokrajac, "Localization of detected objects in multi-camera network," *15th IEEE International Conference Image Processing*, pp. 2376-2379, 2008.
- [7] C. R. Yu, C. L. Wu, C. H. Lu, and L. C. Fu, "Human location via multi-cameras and floor sensors in smart home," *IEEE International Conference Systems, Man and Cybernetics*, pp. 3822-3827, 2006.

## 2.6 計畫成果自評

由實驗結果可以看出來我們所提出的定位方法確實可行，且只需經由簡單的事先校正就可以達成即時的定位。然而定位的結果仍有一些定位誤差。目前我們已著手分析這些誤差對於定位的影響。未來，我們希望能運用誤差分析的結果，更進一步的提升系統的定位穩定性與準確性。

### 主題三：基於影像特徵點定位之人員定位之誤差分析

#### 3.1 研究目的

定位的過程中存在多種誤差，包含徑向畸變誤差、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。在下面的討論中，我們將其分類為靜態誤差與動態誤差。靜態誤差利用校正的方法有效的減低，而較難處理的動態誤差則藉由即時的誤差分析，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。

#### 3.2 文獻探討

延續我們過去發展的幾何計算誤差分析[1]和空間定位誤差分析[2, 3]研究，在此計畫中，我們將對人物定位的過程進行誤差的分析作為多攝影機資訊整合的重要依據，並期望在多攝影機環境下，能夠有效整合不同攝影機的追蹤、定位差異，以得出更精確的定位、追蹤效果。

#### 3.3 研究方法

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。前兩種誤差，我們以校正的方式處理，接著我們著重於處理動態的取像誤差，以提升定位的準確性。我們首先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。然而由於暴力法需要進行大量的計算，不適用於即時的監控應用，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。

##### 3.3.1 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，特別是魚眼鏡頭情況會更為明顯。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。因此影像中目標點座標，將無法以單應性轉換正確地對應到真實空間中的位置。所以徑向畸變誤差對於定位會產生一定的誤差。除此之外，由於我們所提出來的定位方法，需以真實空間中參考點的位置進行單應性轉換矩陣的計算。由於參考點的位置資訊是經由人為測量所得到，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。而取像誤差，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。在接下來的兩小節中，我們將介紹如何改善靜態誤差和估計動態誤差分佈，以期得到準確的定位結果。

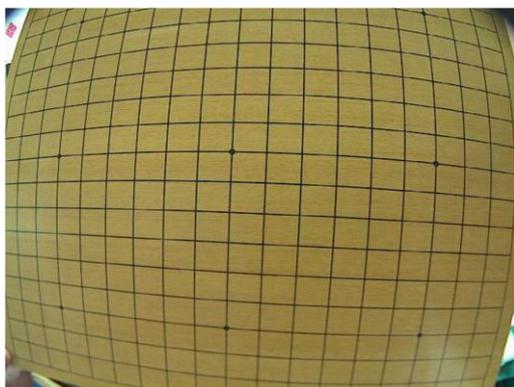


圖 3-1 受到徑向畸變影響的影像。

### 3.3.2 靜態誤差

由於徑向畸變誤差和人為測量誤差定位誤差是不會隨著 Frame 不同而發生改變，因此我們將其定義為靜態誤差。此兩種將可透過影像作校正與微調參考點的測量位置，有效地降低後續對於定位造成的誤差。

#### 3.3.2.1 徑向畸變誤差

受到鏡頭的影響使得影像發生徑向畸變，如圖 3-1 所示。我們可以發現離影像中心點越遠的直線段變成了曲線，這樣將會使得後續使用單應性轉換發生嚴重的誤差。為此，我們將透過參考式(3-1)對影像進行校正，希望使得彎曲的線段變成直線，如圖 3-2 所示。

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

其中  $P = (P_u, P_v)$  為正規化(normalization)後的影像中心點座標， $\|\cdot\|$  為向量的長度； $a_u$  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。更詳細的校正影像的步驟為取得校正特徵點的影像座標，該座標的起始點為影像的中心，並進行正規化，如圖 3-3 (a)與(b)所示。則式子 3-1 會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。

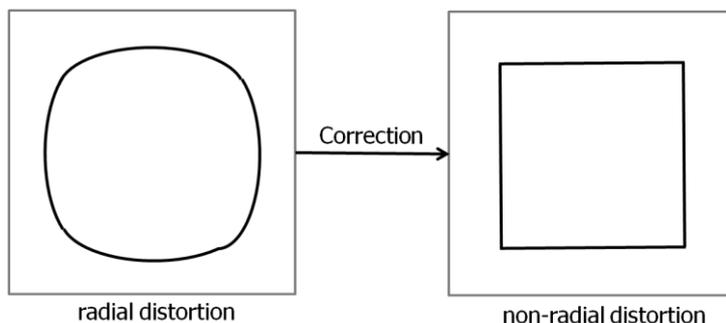


圖 3-2 徑向畸變校正示意圖。

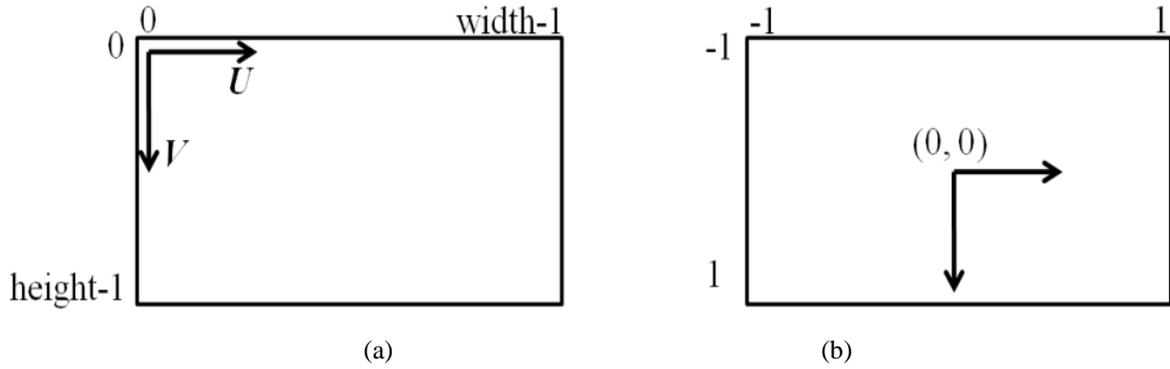


圖 3-3 影像座標正規化。(a)影像座標 $(u, v)$ ，(b)置中且正規化影像座標 $(P_u, P_v)$ 。

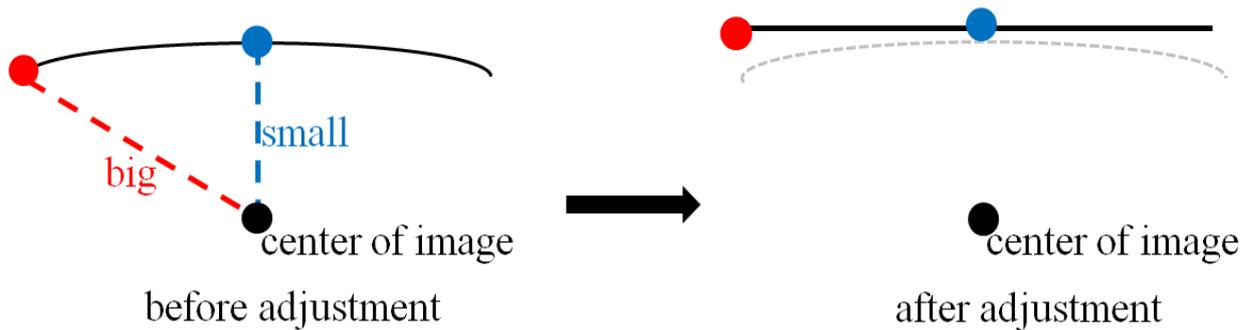


圖 3-4 影像校正時，座標調整示意圖。

在實作的部份，我們利用圖 3-1 中的格子交叉點當驗證點，以找出合適的校正參數  $a_u$  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) 評估這每條直線(共有 38 條)上的驗證點的共直線程度作為誤差的分數，並挑選最低分數者作為較合適的參數，以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。

範例結果如圖 3-5 中(a)與(b)所示，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。且校正後的結果使得圖 2-2 中的 19 個目標點的平均誤差已為校正前的 17.36 公分降到 14.26 公分，大約下降 18%。詳細的差異可參照圖 2-3 與圖 3-6。

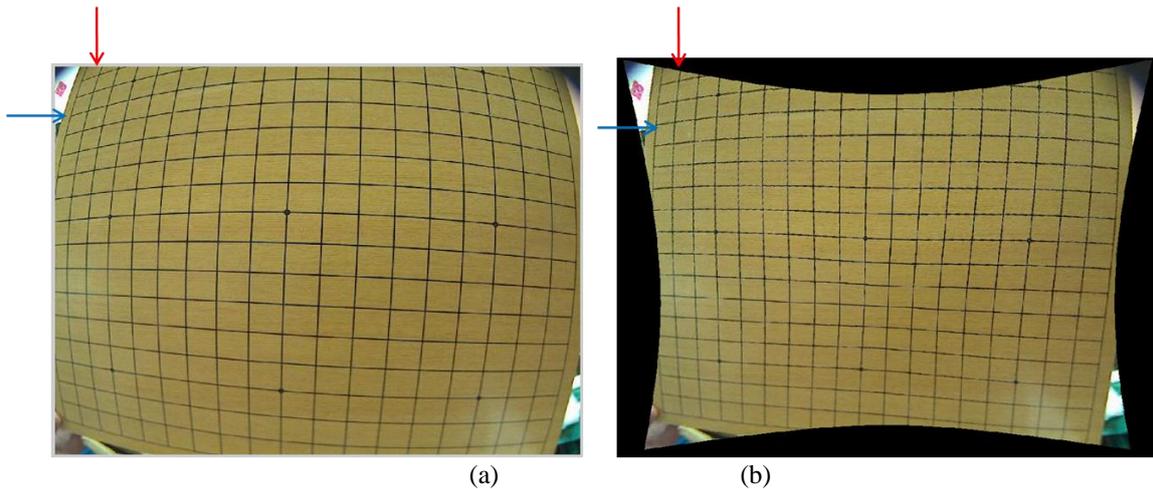


圖 3-5 影像校正。(a) 校正前，(b) 校正後。

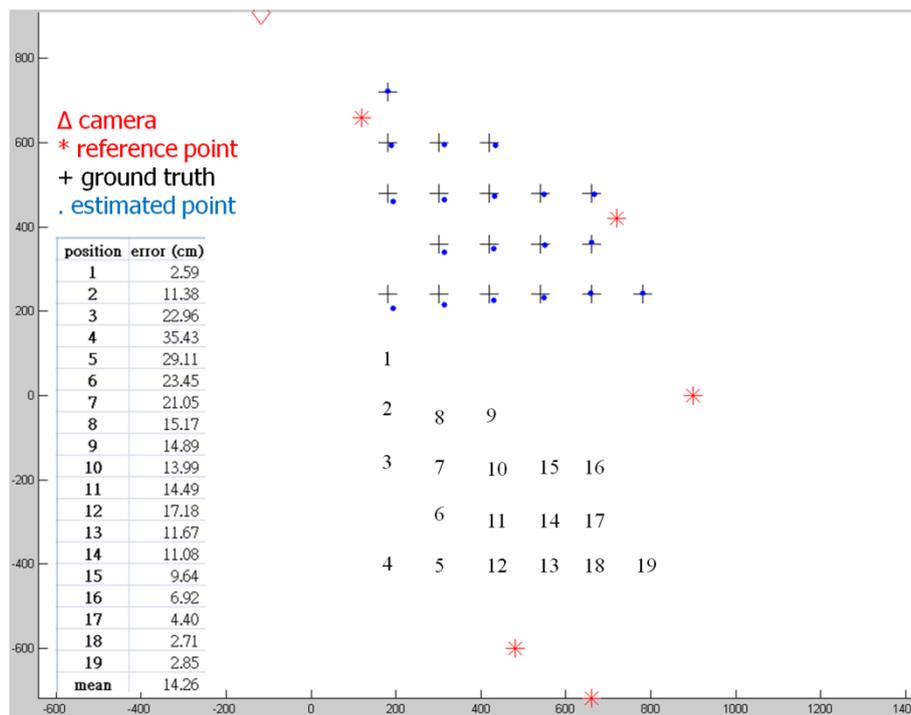


圖 3-6 實驗場景校正後之定位結果。

### 3.3.2.2 人為測量誤差

以手動測量參考點於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，以推估參考點的正確空間位置，如圖 3-7，紅色米字代表測得的 5 個參考點位置。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、左上移動固定的距離，再加上一個不移動則共有 9 種可能。後續再選取一些驗證點，以驗證各種組合估算出來的定位誤差。實驗場景有 5 個參考點，總共具有 59049 ( $9^5$ ) 種可能的組合。圖 3-8 顯示圖 3-7 中參考點微調後的位置，而圖 3-9 顯示影像校正並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大幅改善定位結果。接下來，我們欲討論取像時受雜訊干擾所造成的動態誤差。

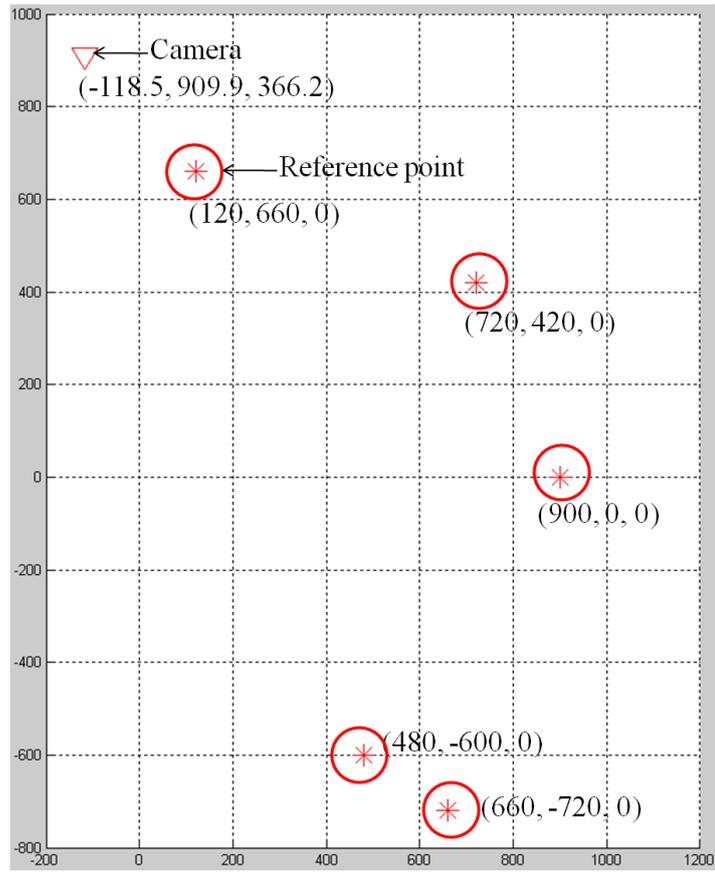


圖 3-7 場景中參考點的位置。

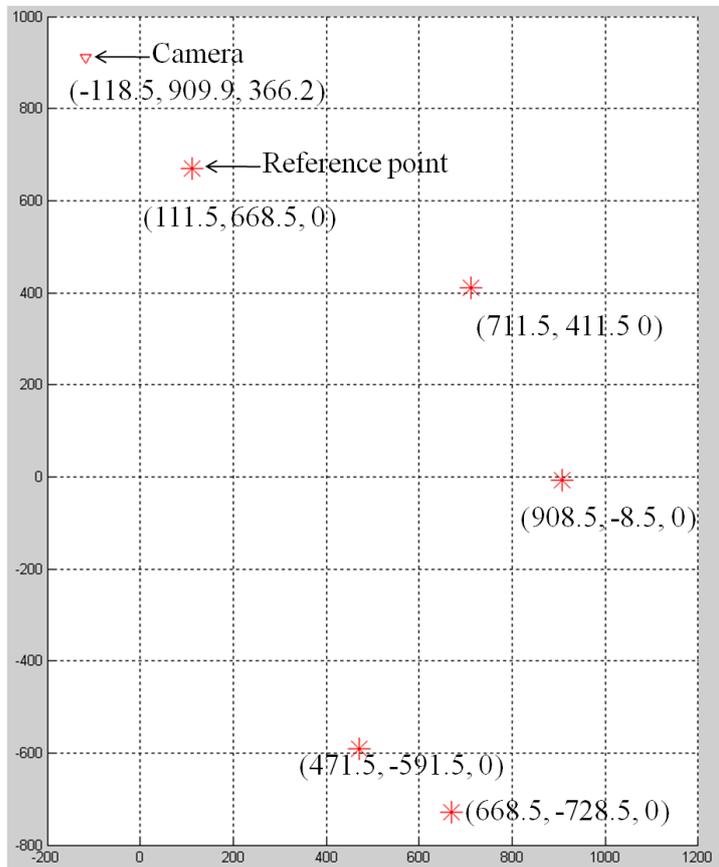


圖 3-8 微調後場景中參考點的位置。

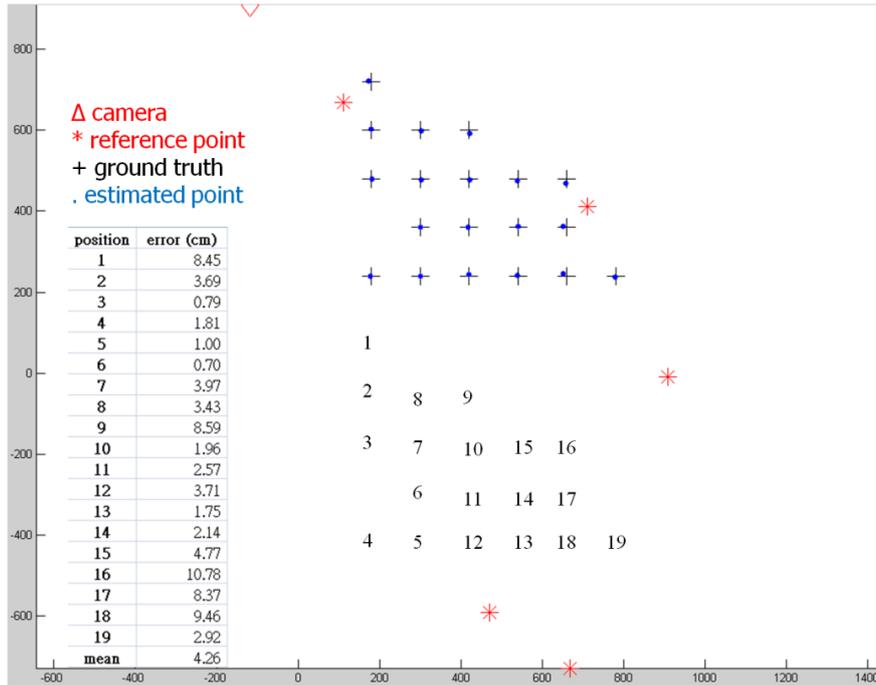


圖 3-9 實驗場景校正和微調參考點後之定位結果。

### 3.3.3 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位置不準確，所產生的誤差我們稱之為動態誤差。為了瞭解動態誤差所造成的影響，我們以模擬的方式進行觀察。首先，我們以高斯雜訊產生 100 個受到干擾後的影像目標點位置。圖 3-10(a)中為原始目標點所在的位置，而藍色點為模擬受高斯雜訊影響的目標點位置。我們可以其分佈的情況大致上呈現圓形分佈；(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，即為受干擾後的定位分佈，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 3-10(b)我們還可以觀察到，誤差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。然而為了獲得此定位分佈，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，最後才可知道定位誤差的分布情形，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布情形，才有機會能應用於實際的即時定位系統中。

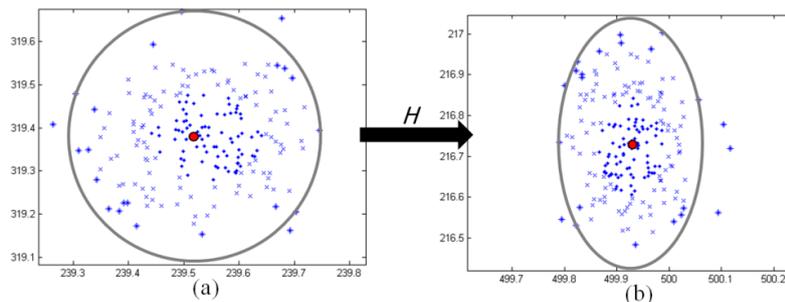


圖 3-10 (a) 原始目標點與模擬受高斯干擾的目標點，(b) 經轉換於參考平面上的目標點分布情形。

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  $\sigma$ ，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 所示而其對應的三維空間中的情形則如圖 3-12 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， $\theta$  為橢圓旋轉的角度。

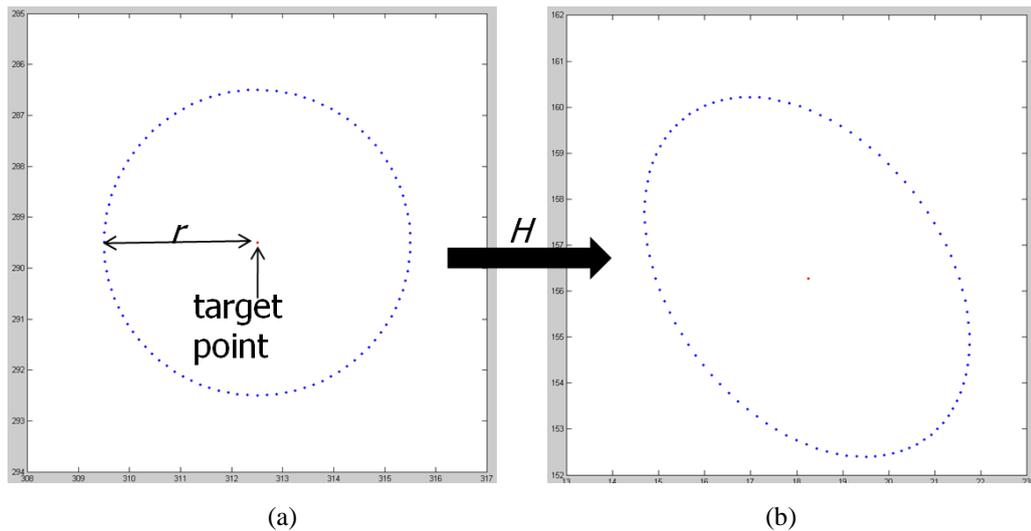


圖 3-11 模擬影像誤差。(a) 影像上的邊界圓，(b) 經由單應性轉換矩陣作用後的邊界圓。

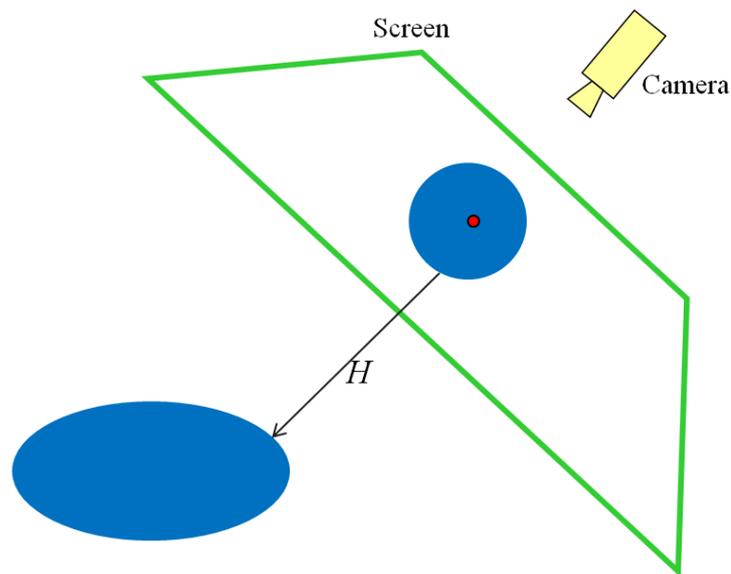


圖 3-12 圓形誤差分佈的定位誤差橢圓。

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)。

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

$$\begin{aligned} Ax^2 + Bxy + Cy^2 + Dx + Ey + F &= 0 \\ A &= a^2 - 2pag + p^2g^2 + d^2 - 2qdg + q^2g^2 - r^2g^2 \\ C &= b^2 - 2pbh + p^2h^2 + e^2 - 2qeh + q^2h^2 - r^2h^2 \\ F &= c^2 - 2pci + p^2i^2 + f^2 - 2qfi + q^2i^2 - r^2i^2 \\ B &= 2ab - 2p(ah+bg) + 2p^2gh + 2de - 2q(dh+eg) + 2q^2gh - 2r^2gh \\ D &= 2ac - 2p(ai+cg) + 2p^2gi + 2df - 2q(di+fg) + 2q^2gi - 2r^2gi \\ E &= 2bc - 2p(bi+ch) + 2p^2hi + 2ef - 2q(ei+hf) + 2q^2hi - 2r^2hi \end{aligned} \quad (3-5)$$

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。

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

$$\gamma = -D\sin\theta + E\cos\theta$$

$$\delta = \lambda^2/4\alpha + \gamma^2/4\beta - F$$

到目前為止，我們已經成功地推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。

### 3.3.4 一對攝影機的定位方法

利用先前提過在單攝影機之定位方法，對目標物進行定位後，我們希望整合一對攝影機的資訊，其方法如圖 3-13 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。

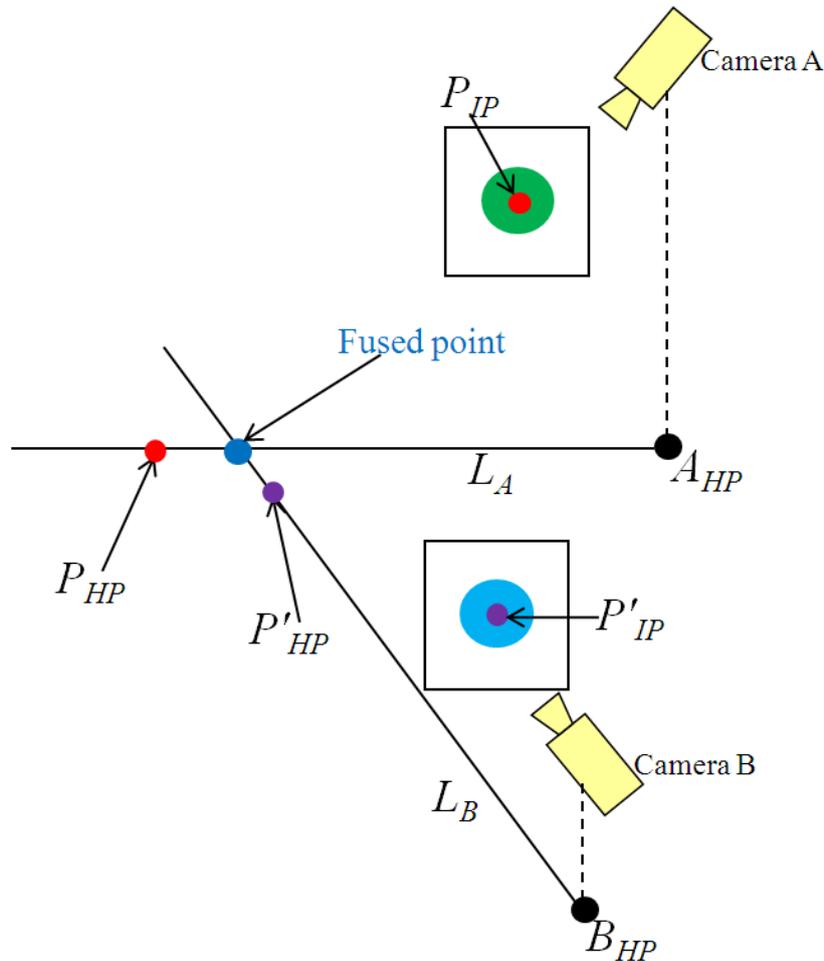


圖 3-13 整合一對攝影機的定位結果。

### 3.3.5 暴力法估算定位誤差分佈

對於此一對攝影機的定位方法，我們透過與上小節相似的模擬動態誤差的方式希望能夠先了解誤差的分佈情形。首先，我們在影像目標點之邊界圓上放置 24 個模擬點，如圖 3-14 所示，接著利用 3.3.4 中整合一對攝影機的定位方法，我們可以暴力法估算所有 FP 共有 576(24 × 24) 個，接著我們採取凸包(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將具有較準確的定位結果。因此，當環境中具有多攝影機時，我們可以從每對攝影機的 CR 結果，挑選出具最小 CR 的一對攝影機，以提供較準確的多攝影機定位結果。

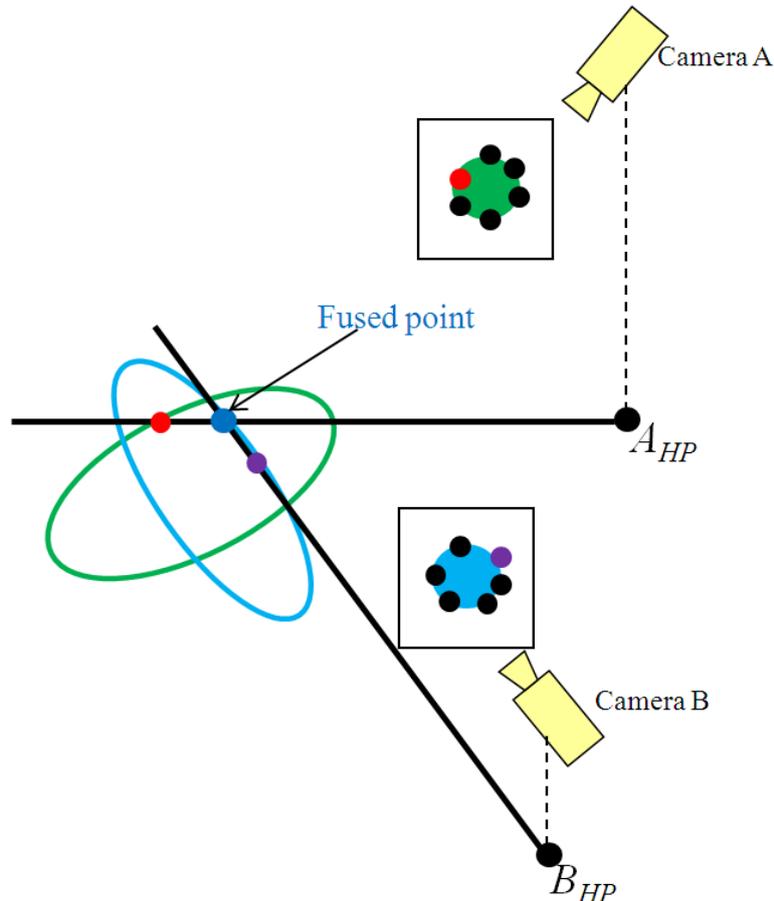


圖 3-14 整合多攝影機間的定位誤差分佈。

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。

### 3.3.6 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，因此我們將利用在 3.3.3 小節中推導的橢圓參數，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 3-15 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 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接近。

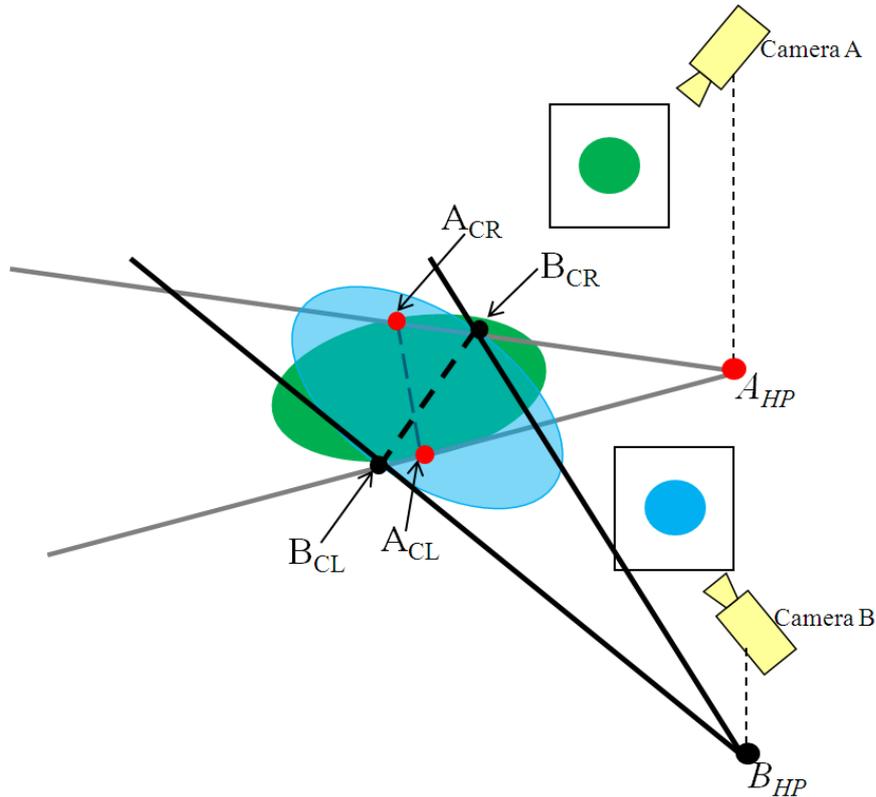


圖 3-15 簡化法估算誤差分佈示意圖。

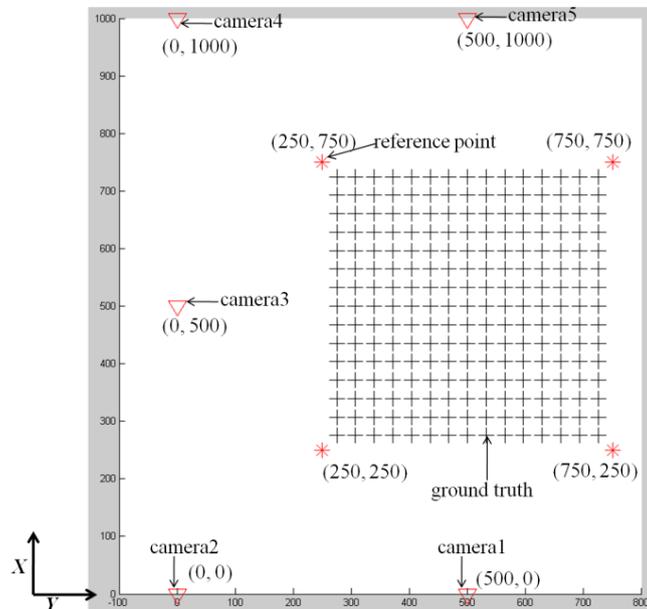
### 3.4 研究結果與討論

在本節中，我們將以模擬的方式測試我們多攝影機整合的方式是否具有成效。接著我們利用真實的場景進行試驗，首先以在地面上的目標點(雷射筆所產生)，沿著特性的軌跡進行移動。接著更進一步我們搭配上簡單的人物頭部偵測以找出頭中心作為目標點，並進行實際的測試。

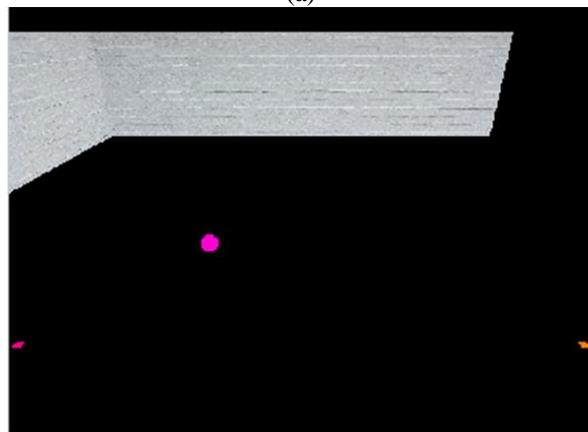
#### 3.4.1 虛擬場景的定位結果

我們建構一個虛擬場景，如圖 3-16 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)為虛擬場景其中一個 frame 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，我們對此目標點多次地加上高斯雜訊，則可產生多個模擬受到雜訊影響而發生偏移的目標點位置。於此實驗中我們產生了 100 個模擬受到雜訊干擾的目標點。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 3.1 為單攝影的定位誤差，由表 3.1 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的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 3.2 為整合多攝影機的定位結果，其中 pair1&2 即為 1 號攝影機和 2 號攝影機的整合定位結果，以此類推。我們可以看到 *Selection* 的結果比起任何一支攝影機都還要來得好，這證明了利用 CR 來挑選可以有效地減低定位的誤差。

接著我們要驗證 3.3.6 節中介绍的 ACR 是否能有效地藉由挑選減低定位的誤差。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 3.2 可看出來，*Selection* 的定位誤差仍然是最小的。另外，在表 3.2 中亦可發現 Pair 1&5 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，這將會使得 3.3.4 節中介绍的整合定位結果較容易變得不穩定，因為兩條接近平行的直線，焦點的位置很容易因雜訊的影響發生大量的偏移。



(a)



(b)

圖 3-16 虛擬場景。(a)場景頂視圖，(b)場景影像。

表 3.1 單攝影機的定位誤差。

Camera	Mean of errors	Variance of errors
Camera 1	2.991	1.429
Camera 2	4.967	2.558
Camera 3	2.989	1.352
Camera 4	4.978	2.649
Camera 5	2.984	1.295
Selection	1.872	0.209

表 3.2 整合多攝影機的定位誤差。

Camera pair	Mean of errors	Variance of errors
Pair 1&2	2.935	1.442
Pair 1&3	1.777	0.225
Pair 1&4	5.507	103.958
Pair 1&5	167.926	879443.584
Pair 2&3	2.912	1.435
Pair 2&4	2.314	0.313
Pair 2&5	5.735	129.731
Pair 3&4	2.845	1.186
Pair 3&5	1.755	0.197
Pair 4&5	2.809	1.212
Selection	1.564	0.134

### 3.4.2 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誤差分析能夠整合多攝影機的定位結果，以取得誤差較佳的定位位置。接下來，我們欲更進一步地以實際場景進行驗證。在實際場景中我們進行兩個實驗，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。

首先在第一個實驗中，我們先測試單攝影機挑選的結果。各個單攝影機在影像中所偵測到的雷射點所形成的軌跡結果顯示於圖 3-17 中(a)-(d)，而俯視圖的定位結果如圖 3-18 中(a)-(d)，而(e)為對於每個目標點位置都參考 CR 來挑選後所得的定位結果。我們可以看到大部分的情況下，挑選的結果都比任一個單攝影機的结果要好。接著我們測試以攝影機配對(camera pair)為單位，並利用 ACR 來進行的挑選。各攝影機配對的定位結果顯示於圖 3-19(a)-(f)中；(g)利用 ACR 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。此外我們也進一步的測試同時考慮單攝影機與攝影機配對的定位誤差估計，即利用 CR 與 ACR 來進行挑選，其結果顯示於圖 3-20，我們可以看到矩形上邊的部分，定位結果變得更靠近黑色舉行邊緣一些，這也證明同時考慮單攝影機與攝影機配對來進行挑選，將可以得到更佳的结果。

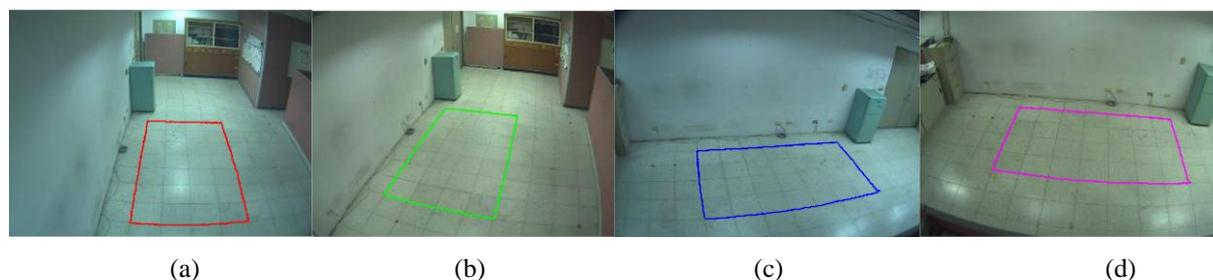
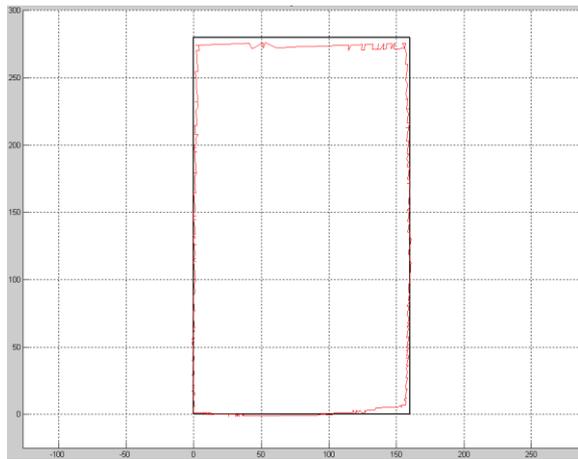
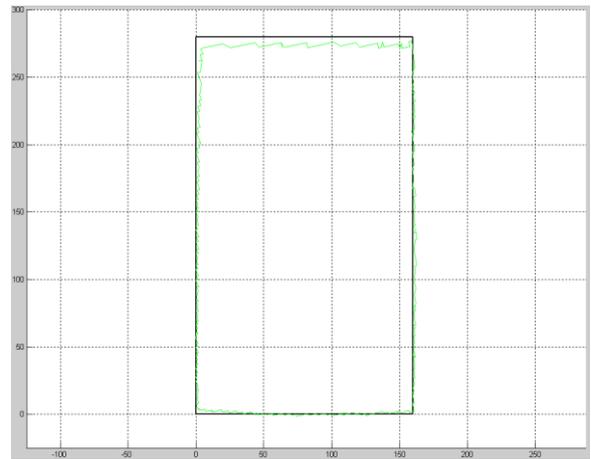


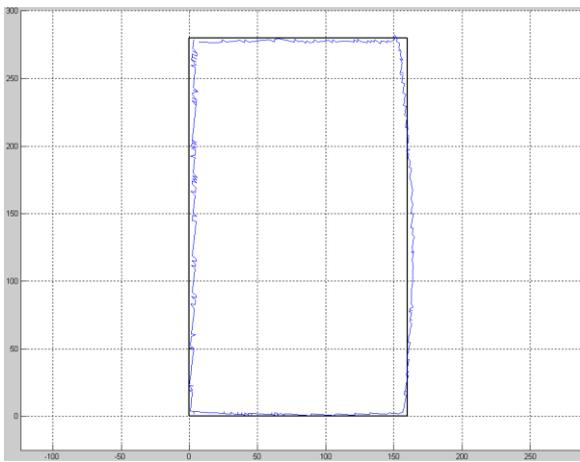
圖 3-17 實際場景—雷射點為目標點。(a) 攝影機 1 的影像，(b) 攝影機 2 的影像，(c) 攝影機 3 的影像，(d) 攝影機 4 的影像。



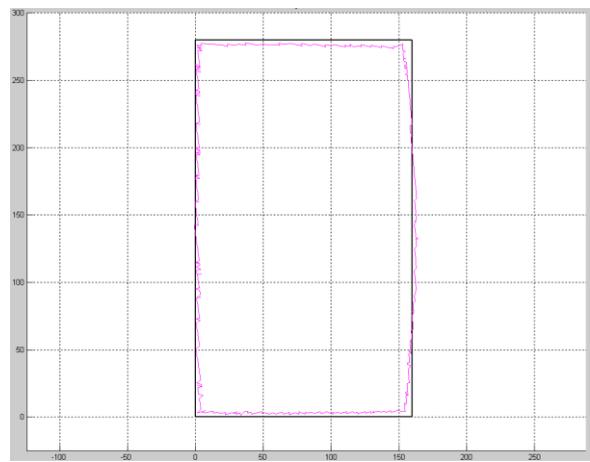
(a)



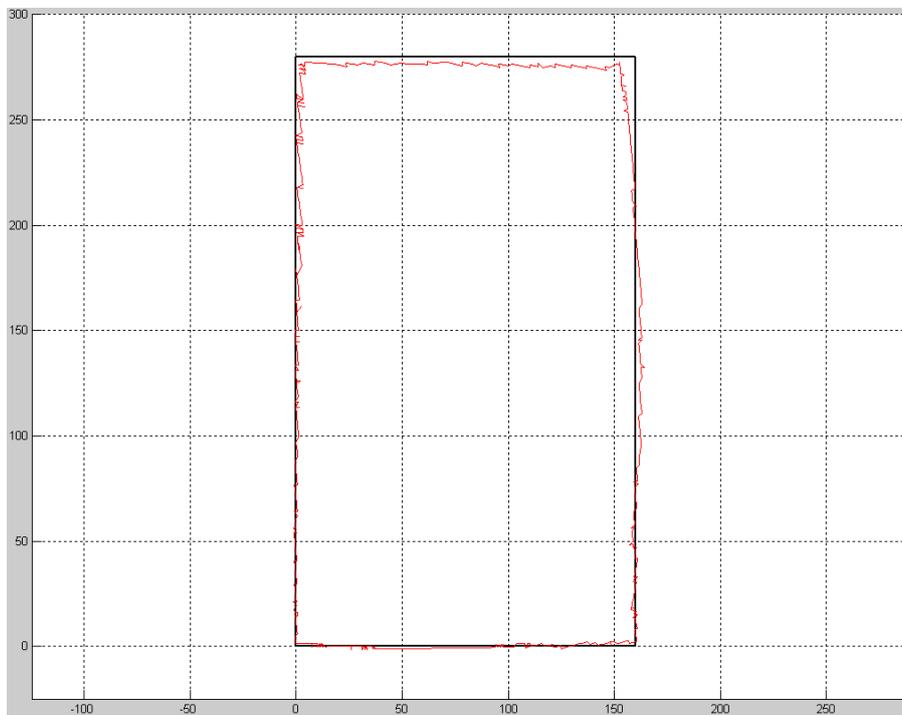
(b)



(c)



(d)



(e)

圖 3-18 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。

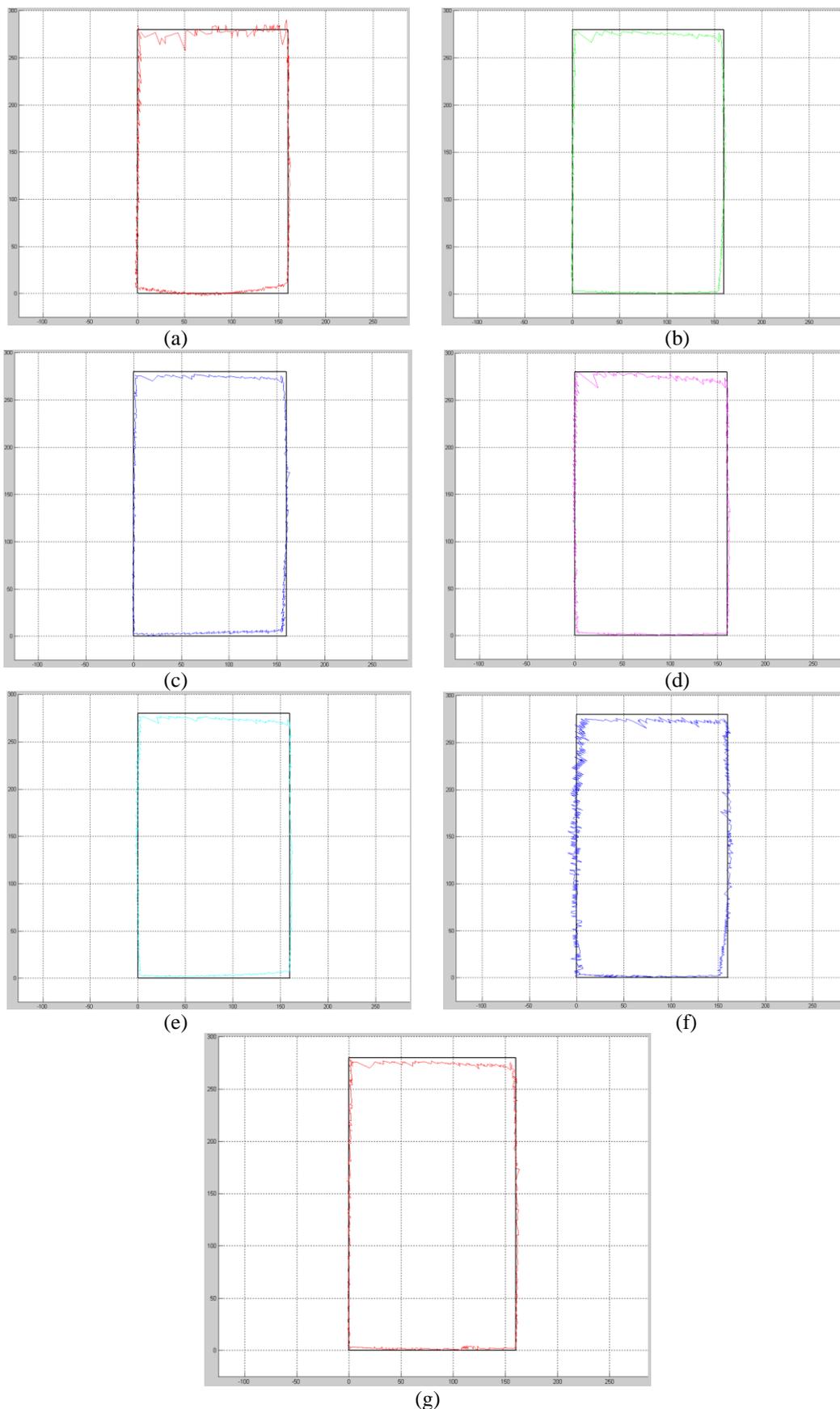


圖 3-19 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。

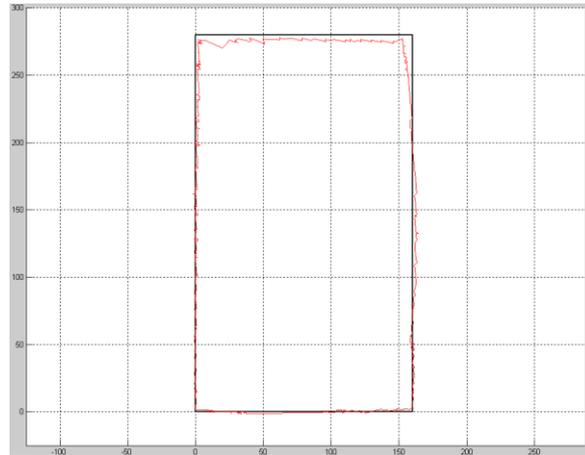


圖 3-20 單多攝影機一起挑選的結果。

接著第二個實驗我們以人物的頭部中心為目標點的定位結果，如圖 3-21 所示，(a)、(c)、(e)和(g)分別為 4 隻攝影機所拍到的影像，(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。定位結果以相似於前面圖 3-18、圖 3-19 與圖 3-20 相似，呈現於圖 3-22、圖 3-23 和圖 3-24。雖然實驗結果較不明顯，但仍可看出，由挑選過後的定位結果都會比沒有挑選的更靠近黑色矩形。

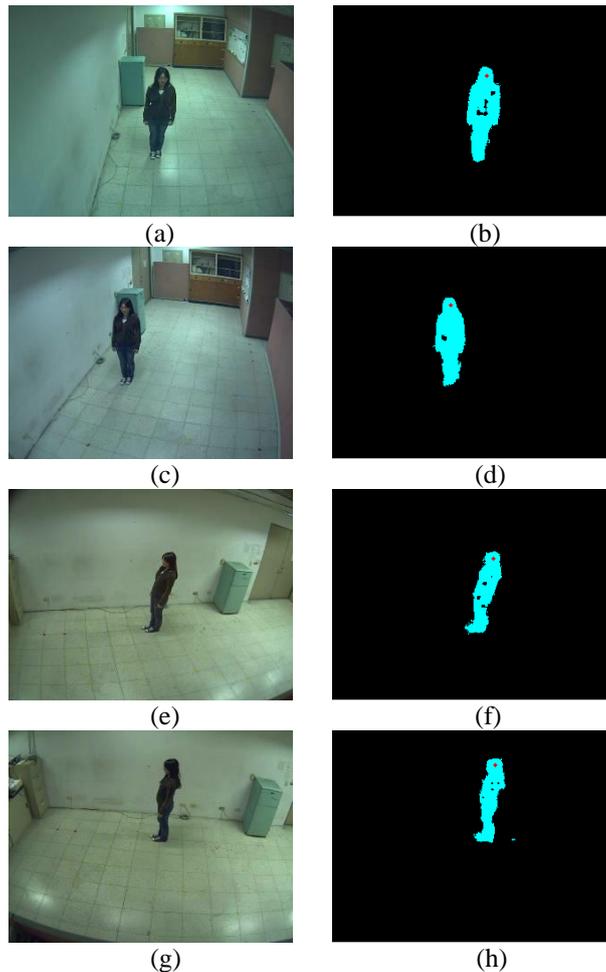


圖 3-21 實際場景—人物。

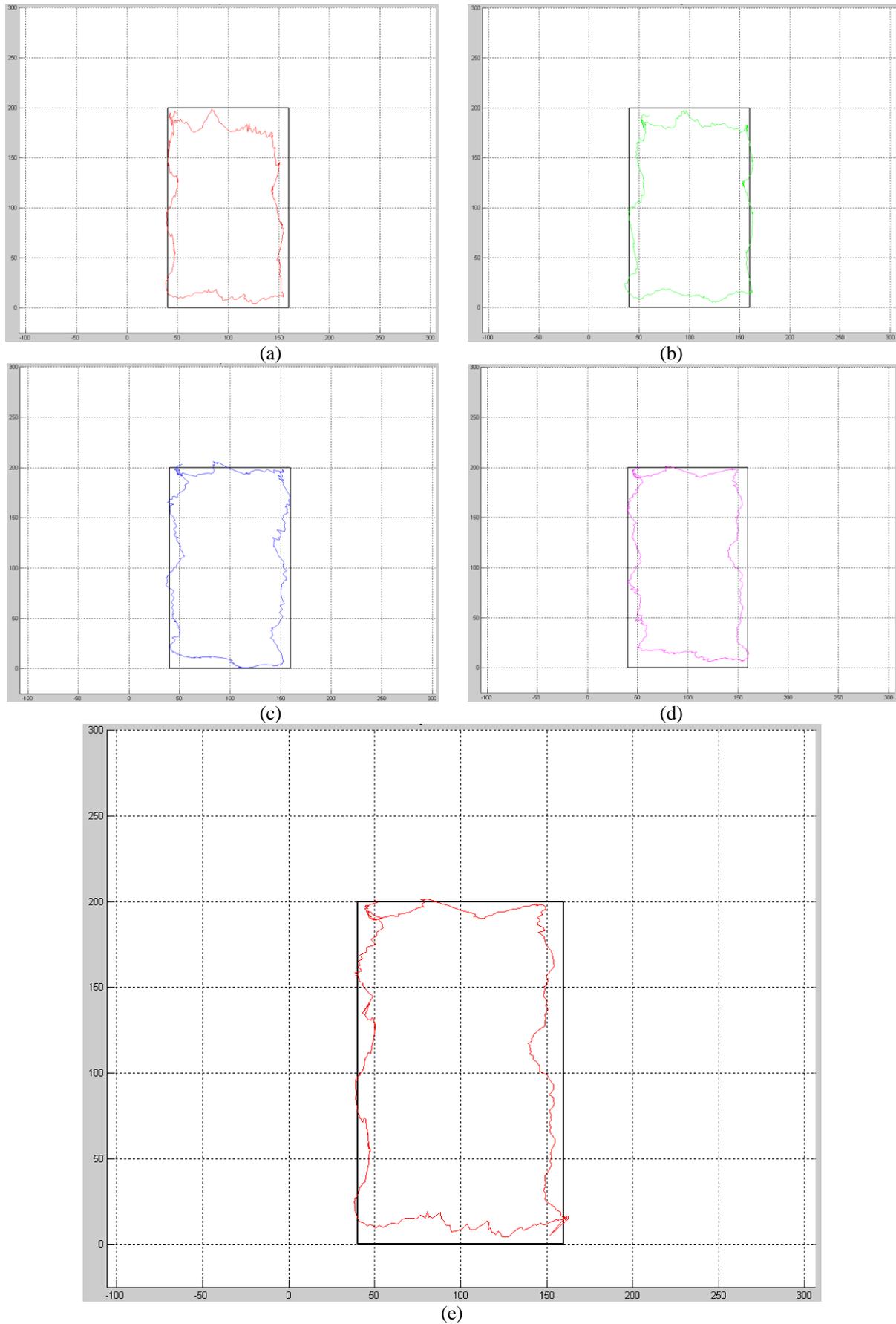


圖 3-22 單攝影機的定位結果。(a) 攝影機 1 的結果，(b) 攝影機 2 的結果，(c) 攝影機 3 的結果，(d) 攝影機 4 的結果，(e) 挑選過後的結果。

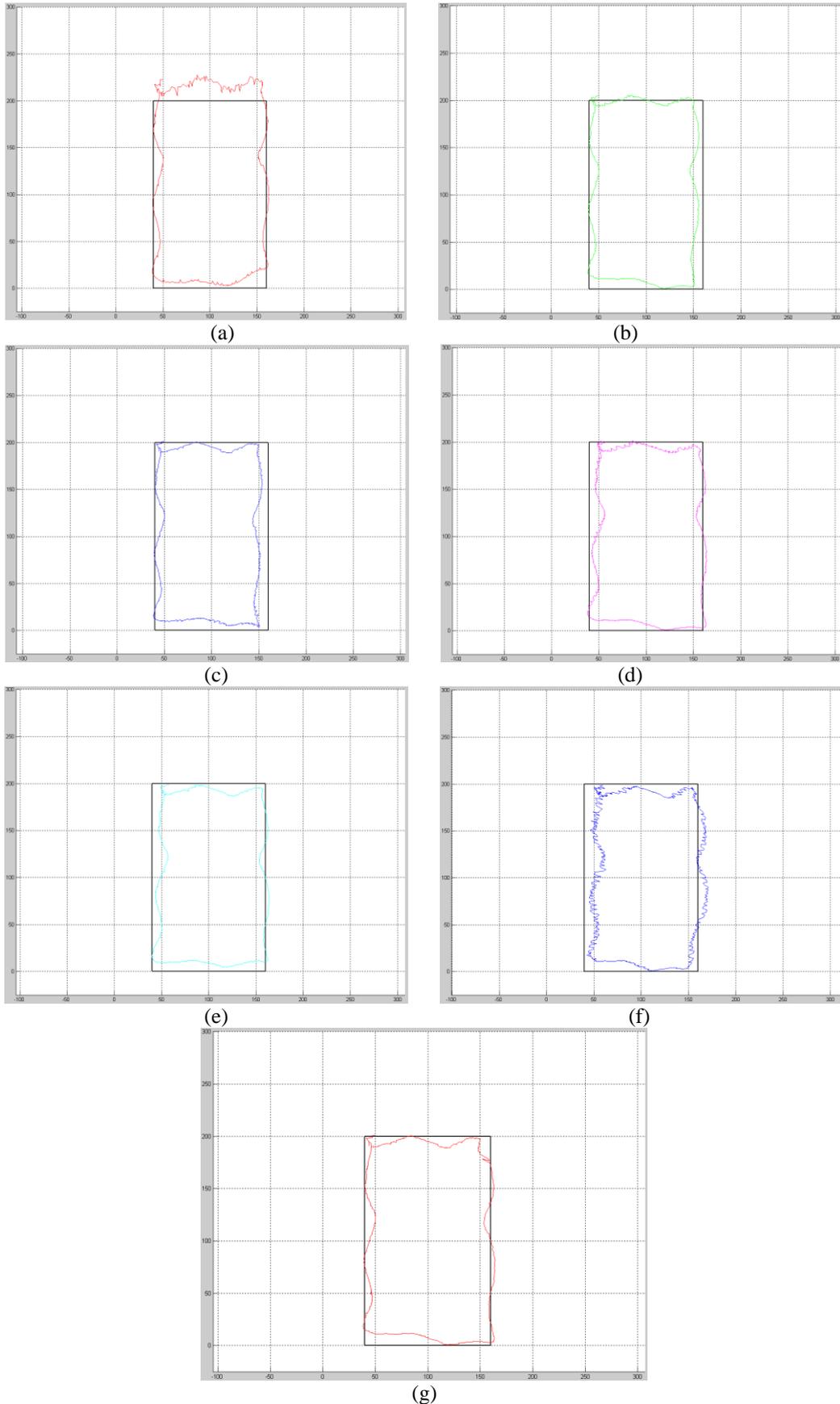


圖 3-23 多攝影機的定位結果。(a) 攝影機 1 和攝影機 2 的結果，(b) 攝影機 1 和攝影機 3 的結果，(c) 攝影機 1 和攝影機 4 的結果，(d) 攝影機 2 和攝影機 3 的結果，(e) 攝影機 2 和攝影機 4 的結果，(f) 攝影機 3 和攝影機 4 的結果，(g) 挑選過後的結果。



圖 3-24 單多攝影機一起挑選的結果。

### 3.5 參考文獻

- [1] J. S. Liu and J. H. Chuang, "A geometry-based error estimation of cross-ratios," *Pattern Recognition*, vol. 35, pp. 155-167, 2002.
- [2] J. H. Chuang, J. H. Kao, H. H. Lin, and Y. T. Chiu, "Practical error analysis of cross-ratio-based planar localization," *IEEE Pacific Rim Symposium Image Video and Technology*, Santiago, Chile, 2007.
- [3] Y. T. Tsai, K. H. Lo, H. L. Huang, and J. H. Chuang, "Error analysis of a real-time vision-based pointing system," *International Computer Symposium*, pp. 202-206, 2008.

### 3.6 計畫成果自評

我們發展了一套特徵點的定位方法及其誤差分析，可估計出影像目標點可能發生的動態誤差，進而推算出對應於地面上的誤差分佈。依據此誤差分析的結果，可提供三種挑選攝影機的模式，第一種為由單攝影機的結果可挑選出(極可能)具最小誤差者。第二種則是以攝影機配對為單位，挑選出(極可能)具最小誤差的攝影機配對。第三種方法則是同時考慮上面兩種方法的結果進行挑選。實驗的結果證明經由挑選整合能夠有效地提升定位結果。

## 國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

### 1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

### 2. 研究成果在學術期刊發表或申請專利等情形：

論文：已發表 未發表之文稿 撰寫中 無

專利：已獲得 申請中 無

技轉：已技轉 洽談中 無

其他：（以 100 字為限）

- Kuo-Hua Lo, Jen-Hui Chuang, Yueh-Hsun Hsieh, and Hon-Yue Chou, "A Point-Based Localization with Error Analysis," *The 2010 International Computer Symposium*, Dec 16-18, 2010, Taiwan.
- Kuo-Hua Lo and Jen-Hui Chuang, "Vanishing Point-Based Line Sampling for Efficient Axis-Based People Localization," *IEEE International Conference on Image Processing*, Belgium, Sep.11-14, 2011.

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以500字為限）

所提出之定位方法是以多種特徵進行人物頭部位置的偵測，因此可以避免人物受遮蔽影響提供準確的頭部偵測結果。這部分的成果將可實用於人物計數的系統，例如賣場或者是捷運相關出入口。而上述的定位結果亦可回饋給其他視覺相關系統，例如視覺追蹤系統，進而提升整體監控的效能。

第二年以特徵點為目標的誤差分析結果，可以動態地整合多攝影機的資訊以提供穩定的定位結果，且此方法僅需要簡單的影像校正，在定位計算上亦具有簡單且快速之特性，故特別適合應用於即時的監控系統。此外，定位精準度的提升也將能擴大應用的層面，使得應用的層面除了一般監控之外，也可以提升至高價物品的監控定位方面，例如由田科技推出的精品珠寶監控系統－「精品特警」。