

國立交通大學
資訊科學與工程研究所

碩士論文



立體視覺指向下系統之準確性分析

Accuracy Analysis of a Stereo Vision-Based Pointing System

研究 生：黃仁信

指導教授：莊仁輝 教授

中華民國 九十五 年六月

立體視覺指向系統之準確性分析
Accuracy Analysis of a Stereo Vision-Based Pointing System

研究 生：黃仁信

Student : Jen-Shing Huang

指 導 教 授：莊 仁 輝

Advisor : Jen-Hui Chuang

國 立 交 通 大 學
資 訊 科 學 與 工 程 研 究 所
碩 士 論 文

A Thesis
Submitted to Institute of Computer Science and Engineering
College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年六月

立體視覺指向系統之準確性分析

學生：黃仁信

指導教授：莊仁輝 博士

國立交通大學

資訊科學與工程研究所

摘要

本論文中實作了一個即時的指向系統，不同於其他相關的研究中使用傳統三維重建的方法來找出指向點，我們沒有使用到相機的內部參數以及朝向等資訊，而是利用平面投影轉換以及一些三維空間內的平面關係作為系統實作的基礎，概念上較為直覺而且也相對地容易理解。而本系統與傳統使用三維重建的方法一樣都存在有誤差的問題，目前我們已經利用平滑過的指向點軌跡當作給予電腦的指令輸入，然而，如果能有更精確的系統，則我們將能研發出更多的應用。因此在建構出此指向系統後，我們亦針對系統的誤差進行分析，檢視誤差的來源並且使用適當的方法降低誤差。

Accuracy Analysis of a Stereo Vision-Based Pointing System

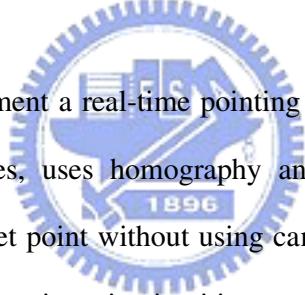
Student : Jen-Shing Huang

Advisor : Dr. Jen-Hui Chuang

Institute of Computer Science and Engineering

National Chiao Tung University

ABSTRACT



In this thesis, we implement a real-time pointing system which, unlike ordinary 3D reconstruction approaches, uses homography and some known 3D reference planes to reconstruct the target point without using camera's intrinsic parameters and orientation. The basic concept is quite intuitive and easy to realize. Despite the existence of reconstruction errors, like ordinary 3D reconstruction approaches, we have developed an application to use the smoothed trajectory of target point as a command to the computer. However, more applications can be considered if we could develop a more accurate system. To that end, we also try to identify the sources of the reconstruction errors and suggest possible ways to reduce them.

致謝

本論文得以完成，首先必須感謝我的指導教授莊仁輝老師的耐心教導，老師熱心的指導讓我在研究方面事半功倍；而老師嚴謹的研究態度，更讓我在研究跟生活上的態度上皆有所啟發。此外，更要感謝三位口試委員，陳玲慧教授、雷欽隆教授以及顏嗣鈞教授的意見及指教，讓本篇論文的內容得以更加充實且完善，在此衷心感謝他們。

我還要感謝智慧型系統實驗室的學長、同學、學弟妹們，感謝高肇宏學長跟林泓宏學長在課業及研究上的指導；感謝庭瑋、開國、建良跟其他碩二的同學們，在那段努力寫論文、拼畢業的日子裡，是靠你們的陪伴與扶持我才能走過來；感謝學弟妹們，尤其是思慧、光兆和郁婷，在忙昏頭、心情低落的時候，你們的鼓勵與打氣給了我很大的力量；還要感謝從大學就一起跟我奮鬥到現在的資科 93 同學們。兩年的碩士生活裡，因為有你們的陪伴，讓我過得很充實也很愉快，真的非常謝謝你們。

最後，感謝一直以來支持我、為我操心的家人，以及所有幫助過我的朋友們，僅以此篇表達我誠摯的謝意。

目錄

| | |
|---|----|
| 第一章 簡介 | 1 |
| 1.1 相關研究 | 2 |
| 1.1.1 使用傳統三維重建方法找出指向 | 3 |
| 1.1.2 使用特殊的影像分析與重建找出指向 | 4 |
| 1.2 系統流程 | 5 |
| 1.3 本論文各章節介紹 | 7 |
| 第二章 雷射筆影像追蹤之探討 | 8 |
| 2.1 利用物體顏色作為限制條件 | 9 |
| 2.1.1 Color Models: RGB & HSI..... | 9 |
| 2.1.2 選取指向物色彩範圍及結果 | 10 |
| 2.2 利用動態物體偵測作為限制條件..... | 13 |
| 2.3 雷射筆之線形分析 | 15 |
| 2.3.1 Linear Least Square Fit..... | 15 |
| 2.3.2 Principal Components Analysis | 17 |
| 2.4 總結 | 18 |
| 第三章 基本投影幾何 | 19 |
| 3.1 齊次座標系表示法 | 19 |
| 3.2 平面投影轉換 | 20 |
| 3.2.1 最小對應點數 | 20 |
| 3.2.2 Over-determined 情形 | 21 |
| 3.2.3 Homography Matrix 的算法 | 21 |
| 3.2.4 使用平面投影轉換取得指向物的三維空間座標 | 23 |
| 3.3 三維空間中的相關數學運算 | 24 |
| 3.3.1 三維空間中三點求解平面方程式 | 24 |
| 3.3.2 三維空間中三平面求解交點 | 25 |
| 3.3.3 系統實作用到的數學運算 | 26 |
| 3.4 總結 | 27 |
| 第四章 準確性估測的實驗結果 | 28 |
| 4.1 系統的穩定性分析 | 31 |
| 4.2 系統的準確性分析 | 39 |
| 4.3 總結 | 42 |
| 第五章 結論與未來工作 | 44 |
| 參考文獻 | 45 |

圖目錄

| | | |
|--------|---|----|
| 圖 1.1 | 使用三維重建為基礎的指向系統示意圖 | 3 |
| 圖 1.2 | [4] 中使用到的幾何關係示意圖 | 4 |
| 圖 1.3 | 本實驗系統的場景示意圖 | 6 |
| 圖 1.4 | 本實驗系統的流程圖 | 7 |
| 圖 2.1 | 雷射筆示意圖 | 8 |
| 圖 2.2 | 追蹤雷射筆流程圖 | 9 |
| 圖 2.3 | RGB Color Model 示意圖 | 10 |
| 圖 2.4 | 左右兩台攝影機拍攝後擷取的畫面 | 11 |
| 圖 2.5 | (a) Hue 值色域示意圖、(b) 本實驗中 HSI 色彩過濾器所使用的 Hue 值範圍示意圖 | 12 |
| 圖 2.6 | (a) 攝影機輸入畫面、(b) 經過 HSI 過濾器處理後的結果 | 12 |
| 圖 2.7 | 連續時序影像經由色彩過濾後的結果 | 13 |
| 圖 2.8 | (a) 連續影像相減之後的結果、(b) 相減影像再經由門檻值 (50) 設定後的二值化影像 | 14 |
| 圖 2.9 | 將圖 2.8(b) 對應到目前時序影像所得的結果 | 14 |
| 圖 2.10 | 框出只包含雷射筆像素的 Bounding Box 示意圖 | 15 |
| 圖 3.1 | 在三維空間選取的標記點示意圖 | 23 |
| 圖 3.2 | 影像平面與參考平面轉換示意圖 | 23 |
| 圖 3.3 | 同平面上三點示意圖 | 24 |
| 圖 4.1 | 指向系統實驗所使用到的 PTZ 攝影機 | 28 |
| 圖 4.2 | 實驗場景俯視示意圖 | 29 |
| 圖 4.3 | 使用者使用指向系統側照圖 | 30 |
| 圖 4.4 | 指向系統結果示意圖 | 31 |
| 圖 4.5 | 亮紅色包裝紙長度較長的雷射筆 | 32 |
| 圖 4.6 | 雷射筆沿平行地面方向來回移動之資料點的 PCA fitting 誤差 (a) 使用原本長度的雷射筆、(b) 使用加長長度的雷射筆 | 33 |
| 圖 4.7 | 原本長度的雷射筆指向物沿平行地面方向來回移動之指向物線段斜率變化 (a) 測試片段 1、(b) 測試片段 2 | 35 |
| 圖 4.8 | 較長的雷射筆指向物沿平行地面方向來回移動之指向物線段斜率變化 (a) 測試片段 1、(b) 測試片段 2 | 36 |
| 圖 4.9 | 原本長度的雷射筆指向物沿垂直地面方向來回移動之指向物線段斜率變化 (a) 測試片段 1、(b) 測試片段 2 | 37 |
| 圖 4.10 | 較長的雷射筆指向物沿垂直地面方向來回移動之指向物線段斜率變化 (a) 測試片段 1、(b) 測試片段 2 | 38 |
| 圖 4.11 | 理想的指向系統之誤差結果，(a) 僅有平移的情況、(b) 有平移及扭曲的 | |

| | |
|---|----|
| 情形..... | 40 |
| 圖 4.12 實際的指向系統之誤差實驗結果 | 41 |
| 圖 4.13 實際的指向系統之誤差結果，(a)~(f)各次測試過的結果 | 42 |



表目錄

表 2.1 攝影機位於不同光源下所拍攝的亮紅色包裝紙色彩比較 11



第一章 簡介

目前市面上提供給使用者用以與電腦互動的設備，例如滑鼠、鍵盤、手寫板，均仰賴外加的設備以及一些特定的輸入方式才能與電腦互動，而且在使用的距離上也受制於設備的使用範圍，這些設施在提供人們使用上仍然缺乏了直覺性以及便利性。如果我們能直接利用一些隨手可得的物體，例如筆，或者人的手臂，這些取得跟使用上比較便利的物體，來直接跟電腦互動，相信對於改進人機互動有相當大的助益，而且對於使用者在操作上也會更為方便。

目前已經有相當多改善人機介面的研究在進行中，我們針對其中的「指向系統」作為我們的研究主題。所謂的「指向系統」，是讓使用者直接利用手指、手臂、筆等輕易取得的物體，直接對實際的物品或平面做指向的動作，而系統會分析這個指向的動作，並且找出使用者指向的方向跟目標，接著再對指向的目標進行相對應的指令。

指向系統的應用相當地廣，例如 (1) 簡報系統：人們在進行簡報的時候，如果要對投影片下達換頁的動作，只能直接到電腦旁按鍵下指令，若是報告者跟電腦有一段距離的話，就需要不斷來回的走動，而現有的簡報筆只能對電腦下投影片換頁的指令，缺乏圈出重點部分讓聽講者注意的功能。如果能利用指向系統，就可以直接利用指向對電腦下換頁的指令，同時也可以操控滑鼠，讓滑鼠可以停留在希望聽講者注意的部分。(2) 家電操控：現今的家庭裡，有太多遙控器以致於在使用上會造成困擾，如果可以直接用手指指向家電，然後指到的電器便會做出相對應的動作，例如讓冷氣開機、讓電視轉台或調整音量等，如此可以省去找尋跟使用遙控器的麻煩，在操縱家電上也會變得更為簡易一點。(3) 三維空間建模：如果可以直接利用指向系統指出那些位置有什麼物品、大小尺寸為何，這樣便可以建出一個三維空間的簡易模型，之後可以再對於此簡易模型做精確化的動作，如此就可以省去實際場景量測的苦工。(4) 射擊遊戲：現在的電動遊戲

中，如果需要玩射擊遊戲的話，玩家需要使用搖桿、手把或是槍枝樣式的外加設備，如果可以直接用手指做射擊動作的指向，如此玩家在進行射擊遊戲的時候不但不需要使用額外的設備而且也可以跟遊戲有比較直接的互動。由於還有其他相當多可以使用指向系統來做的應用，因此研發出一個穩定的指向系統是有相當大的可用性及應用性的。

值得注意的一點是，根據不同的應用目的，對於指向系統的精確度也會有不同的要求，若我們將指向系統當成是射擊遊戲中的瞄準系統，那麼對於指向系統精確度會有較高的要求；若是將指向系統應用在一般的簡報系統中，那麼對於指向系統的精確度要求就不會太高，只要所指向的目標點會隨著我們手指移動的方向而做出正確的移動，我們就能利用使用者回饋（feedback）的方式指到想指的目標點。因此除了研究指向系統的應用性之外，分析指向系統的準確度也是一個重要的課題，而在本篇論文中，除了介紹我們所提出的指向系統之實作方法，同時也將針對我們的指向系統之準確度做探討。



1.1 相關研究

指向系統的相關論文裡，在找出使用者指向的方法中，大致上可以分為兩種：(1) 使用傳統三維重建的方法，(2) 使用特殊的影像分析與重建。在第一種方法中，系統至少需要兩台以上的攝影機，而且每台相機都經過校正獲得相機的內外部參數[1]，接著再以內外部參數作為系統的基礎運算找出指向點。而第二種方法裡，則是使用一些特殊影像處理的方法，利用深度影像的資料或是利用影像對應點的關係對應，從單張或多張影像中求得指向物的三維座標，再計算出指向點的座標。

在本節中，我們將選出幾篇與指向系統相關的研究論文作介紹，藉由這些介紹我們可以瞭解到其他作者使用到什麼技術來找出指向點，以及他們將指向系統做了哪些應用。

1.1.1 使用傳統三維重建方法找出指向

使用這種方法找出指向的系統，都至少需要兩台以上已經校正好的相機，以取得其內部參數跟外部參數。而這種方法中，使用相機的內外部參數來實作系統的方法也有許多種，例如直接使用內外部參數重建指向物的三維座標、使用內外部參數作為數學運算式推導過程的數值等。因此，我們選了三篇使用內外部參數但是實作方法不同的論文在這一小節內作介紹。

例如 Y. P. Hung et al.[2]便利用 IIS (Institute of Information Science) head 這個已經校正好的立體視覺系統，使用三維重建的方式達到重建指向點的目的。在這個系統中，他們提供(1) 指尖跟手掌中心的連線以及(2) 指尖跟眼睛的連線，這兩種方法給使用者進行指向的動作。圖 1.1 是他們系統的示意圖，其中顯示了系統將使用者的指向點繪在投影布幕上（布幕上 3D 下的一個紅點）的結果。我們也可以同時由圖中發現，左側用來拍攝使用者指向的 IIS head 的體積龐大，幾乎與人身高相仿，而且機器必須很靠近使用者用以得到清晰的手指影像，因此在使用上仍會有所不便。



圖 1.1 使用三維重建為基礎的指向系統示意圖

Y. Yamamoto et al.[3]則是在房間中天花板的四個角落架設照相機，拍攝俯視角度的影像後，再使用一套已經研發完成並且校正過的 Ubiquitous Stereo Vision (USV) 系統，從這些俯視角度的影像中，利用一些三維的資訊求出使用者手臂及頭部的方向、重心位置、朝向角度等資訊，接著再利用這些資訊算出使用者手

臂的水平延伸方向、垂直延伸方向，用此兩個方向求出使用者指向的方向。而他們也利用這個指向系統完成電視控制的應用，可以以手臂的指向來控制電視轉台、音量調節、開啟關閉。

而 C. Colombo et al.[4]也使用了兩台以上校正過的相機來建構一個指向系統，圖 1.2 是他們系統的示意圖， π_s 為指向平面、 L 為指向線。他們推導出每個影像平面 π_i 上的指向線： l_i ，跟指向點 P_s 都具有一個限制條件： $l_i^T H P_s = 0$ ， H 為 π_i 與 π_s 間的平面投影矩陣，而藉由這個限制條件以及已知的內外部參數，可以推導出 P_s 的座標，也就得到了指向點的位置。

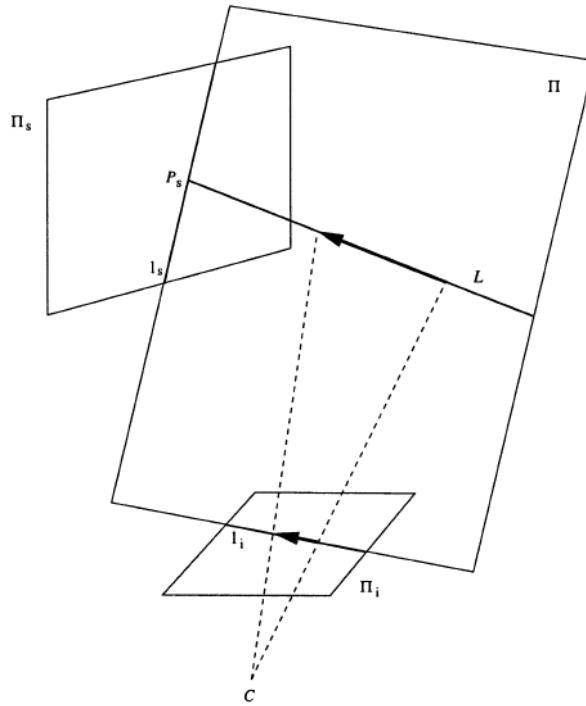


圖 1.2 [4]中使用到的幾何關係示意圖

1.1.2 使用特殊的影像分析與重建找出指向

這類方法不同於使用傳統三維重建的方法，僅需要從多張影像的資訊中，利用一些影像處理的技巧，擷取、計算出使用者的指向物的三維資訊，便可以再進一步計算出使用者的指向。

例如 E. Hosoyam et al.[5]，他們提出了一個在立體空間中，指向空間裡的機器會執行相對應指令的指向系統，他們不是利用三維重建的方式來達成這個目的，而是利用深度影像（depth image）的輔助，從二維影像中的座標，反推得到使用者手臂重心跟肩膀的三維座標，再以此三維座標求出使用者是指向哪個方向、哪個機器，繼而做出相對應的指令。這個系統需要先記錄哪些位置有什麼機器，而且同一個區域裡不能有太多的機器、機器也不能重疊，否則會讓系統無法找出使用者指向的目標機器，這部分是提供給使用者應用上比較受限制的地方。

在 M. S. Lee et al.[6]的提出的指向系統中，他們使用兩台相機，並且提出了兩種方法給使用者進行指向給使用者進行指向：(1) Peek Thru Configuration、(2) Direct View Configuration。第一種方法是將兩台相機架設在目標平面（target plane）上，直接拍攝使用者正面的影像；第二種方法則是將相機架設在使用者左右兩側的斜後方，各自拍攝使用者側面以及目標平面的影像。第一種方法可以提供使用者操控電視轉台的應用，第二種方法則可以實作簡報系統。而這兩種方法都使用了 homography 的方式，利用影像平面跟目標平面上的四個特徵點對應關係，從影像平面上使用者的眼睛以及指尖位置計算出眼睛以及指尖在目標平面上的位置，進一步找出指向點。這個系統比較困難的部分在於使用者眼睛跟指尖的追蹤。這篇論文在使用到 homography 這點上與本論文實作的系統相同，相異處除了指向物的選擇不同外，更重要的是兩者選用的參考平面不一樣，這篇論文直接使用目標平面作為影像平面的參考平面，而我們則是利用三維空間的另一個平面作為參考平面；同時本篇論文只利用了在參考平面上線的交點當作指向點，我們還引用了三維空間平面相交的概念來求出指向點。因此，本篇論文與我們的系統在實作上還是有很大的差異存在。

1.2 系統流程

圖 1.3 為本實驗的場景示意圖，本系統中使用了兩台安裝在天花板中的 PTZ

攝影機，使用者正前方距離約 3~4 公尺遠的平面為我們用來指向的指向平面 (target plane)。在系統的實驗階段中，提供給使用者用來指向的物品為圖中使用者手裡握著的雷射筆，使用者利用此雷射筆在指向平面進行指向，而本系統會偵測出使用者所指向的指向點 (target point)。在地上所標記的使用者移動區域 (user moving area) 則是為了讓兩台攝影機可以同時追蹤到使用者手中的雷射筆而繪出的限制區域。

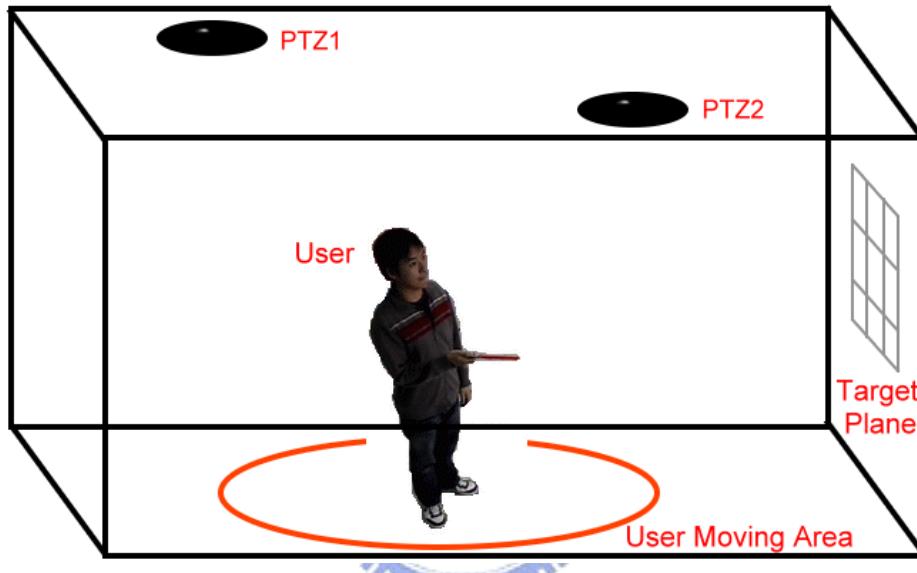


圖 1.3 本實驗系統的場景示意圖

圖 1.4 是本實驗的系統流程圖，由流程圖中可以先瞭解到我們的系統實作大致上可以先分為兩部分：(1) 在兩台攝影機的畫面中，追蹤指向物的二維空間位置，(2) 量測實驗場景的三維空間座標。再利用這兩部分得到的資訊，使用平面投影轉換 (homography) 的技巧，推算出使用者的指向點位置為何。而詳細的系統實作方法以及實驗的環境，我們將於接下來的章節做詳細的介紹。

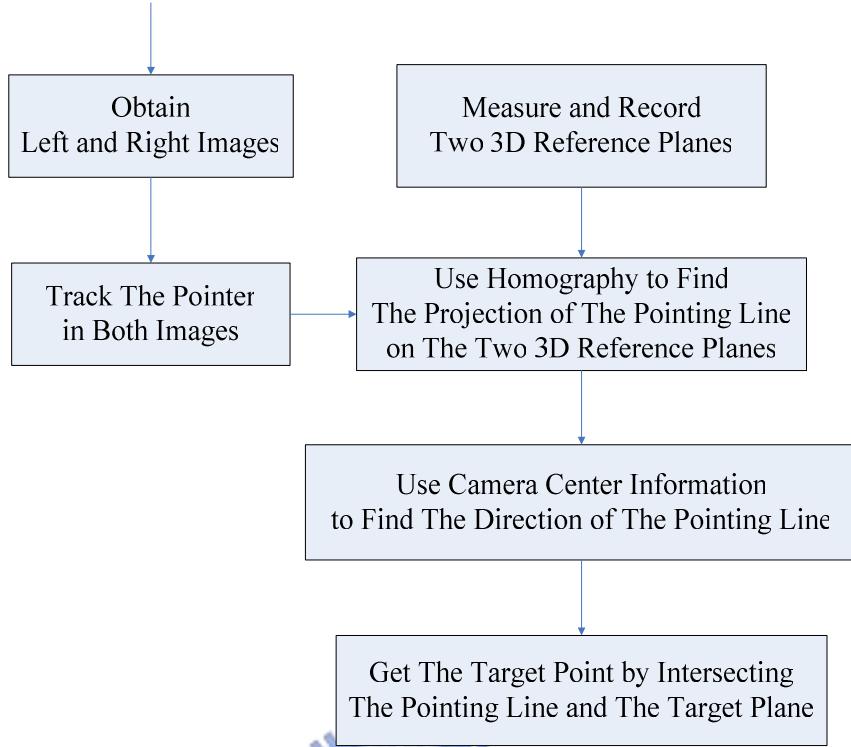


圖 1.4 本實驗系統的流程圖



1.3 本論文各章節介紹

本篇論文全文共分為五章，除了本章為緒論外，第二章介紹追蹤指向物的方法，內容包含色彩限制條件、動態物體偵測以及求解線形的方法介紹。第三章介紹基本的投影幾何以及求解指向點所需的相關的數學運算。第四章介紹指向系統的一些穩定性的分析及精確度估計，以及使用不同方法的結果比較。第五章為結論並說明本研究未來的一些研究方向。

第二章 雷射筆影像追蹤之探討

本篇論文中，我們實作了一個即時指向的系統，在系統開發及實驗的階段，我們提供給使用者進行指向動作的指向物是一隻具有特殊顏色的雷射筆，使用雷射筆的好處是：(1) 雷射筆的形狀及其包覆的特殊顏色紙張可讓我們在進行追蹤的時候更加簡便及快速，(2) 可利用雷射筆所射出的紅色標記點跟系統重建的指向點做精確度的比較。同時我們也預期在本系統的穩定度以及準確度較為完善之後，可以再繼續研究讓使用者直接利用手指或其他物體進行指向的動作。

圖 2.1 為我們使用的雷射筆之示意圖，我們用一層亮紅色的包裝紙覆蓋在雷射筆的外部，使得在進行指向物追蹤時，目標物可以更為明確，同時也可以增進追蹤指向物的速度及準確度。



圖 2.1 雷射筆示意圖

本章中，我們將針對如何追蹤雷射筆此問題作簡略的探討，追蹤雷射筆的步驟主要包含了使用物體的顏色跟運動作為兩項限制條件。通過這兩部分的限制條件之後的像素，我們先利用雷射筆長條狀的形狀及雷射筆為連接物的特性來刪除雜訊點，最後再對篩選出來的雷射筆像素點進行線形的分析，用以求出雷射筆在影像中的位置所在，圖 2.2 為追蹤雷射筆此部分的流程圖。

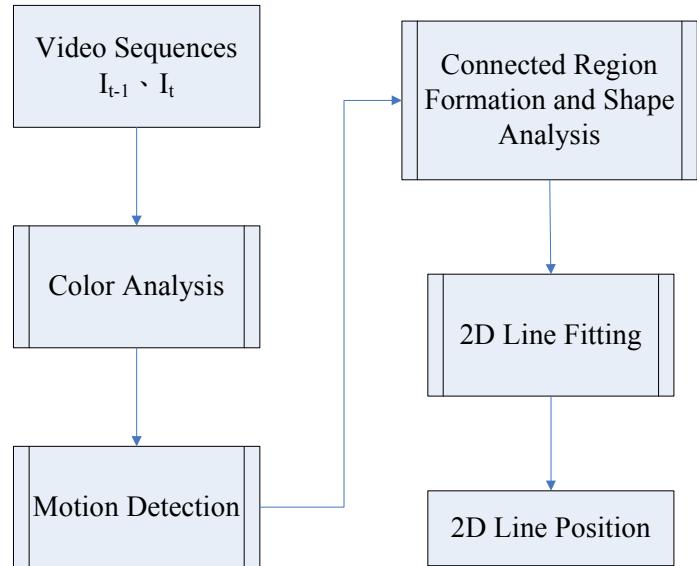


圖 2.2 追蹤雷射筆流程圖

2.1 利用物體顏色作為限制條件

在設計追蹤雷射筆的方法前，我們先觀察用來指向的雷射筆外觀的特性，期望使用雷射筆外觀的一些特性來增進追蹤的速度跟準確度。而由圖 2.1 可以發現，我們使用了在一般場景內出現機率較低的亮紅色包裝紙包覆在雷射筆外層，因此在追蹤雷射筆時先利用顏色作為過濾的限制條件是一種很直覺而且可以增進速度的方法。

因為從攝影機中擷取出來的影像為 RGB24 的彩色影像，所以使用最常用的 RGB 色彩模型（Color Model）來表示顏色似乎是最簡單的方法，然而由於場景內光線變化對於以 RGB 色彩模型表示的顏色影響非常大，會導致在追蹤雷射筆時濾掉了屬於雷射筆的部分。因此我們進一步使用另一種常用的 HSI 色彩模型來表示顏色，以獲得抑制光線變化影響的目的。本節將針對 RGB 及 HSI 兩種色彩模型做個簡略的介紹，並且展示我們選取的色彩範圍跟結果。

2.1.1 Color Models: RGB & HSI

RGB Model 是最常使用的色彩模型，概念上是用 Red、Green、Blue 三個主色的組合來表示各個色彩，因此我們可以利用一個 cube 表示 RGB Color Model，如圖 2.3 所示，而在此 cube 中的任何一點就代表一種顏色。

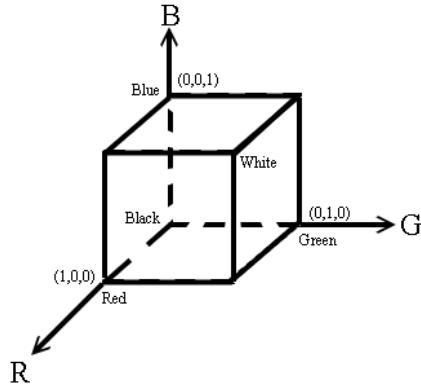


圖 2.3 RGB Color Model 示意圖

除了 RGB 色彩模型之外，還有其他概念跟 RGB 色彩模型不同的色彩模型，例如 HSI 便是經由一些數學式子的運算，將色彩由 RGB 空間轉換到 HSI 空間，轉換的公式如下：

$$H = \begin{cases} \theta & , \text{if } B \leq G \\ 360 - \theta & , \text{otherwise} \end{cases} \quad (2.1)$$

$$S = 1 - \frac{3}{(R + G + B)} * [\min(R, G, B)] \quad (2.2)$$

$$I = \frac{1}{3} * (R + G + B) \quad (2.3)$$

其中 $\theta = \left\{ \frac{0.5 * [(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{0.5}} \right\}$ 。我們由式 2.1～式 2.3 可以觀察得知，HSI 是以色調（Hue）、飽和度（Saturation）、亮度（Intensity）三者作為組合色彩的元件。使用 HSI 用這種色彩表示法，是將色彩的色調、飽和度、亮度分離成三個元素，各自獨立去觀察，不同於 RGB 色彩模型將色彩的色調、飽和度、亮度等資訊均混合在一起，因此對於色彩的變化情形可以表現得更為清楚。

2.1.2 選取指向物色彩範圍及結果

在介紹完 RGB 跟 HSI 兩種不同的色彩模型後，我們必須要選擇符合我們需

求的色彩模型並且選取適當的過濾範圍來得到亮紅色雷射筆的像素。圖 2.4 為我們在實驗場景中左右兩台攝影機所拍攝的影像，我們在兩張影像中各自追蹤雷射筆的位置，可以發現畫面中的亮紅色雷射筆線段相當醒目，而且場景內均沒有出現類似的顏色。



圖 2.4 左右兩台攝影機拍攝後擷取的畫面

表 2.1 攝影機位於不同光源下所拍攝的亮紅色包裝紙色彩比較

| | 圖(2.4) (a) | 圖(2.4) (b) |
|---------|-----------------|----------------|
| RGB 平均值 | (210, 110, 110) | (100, 50, 50) |
| HSI 平均值 | (0, 0.5, 0.7) | (0, 0.9, 0.35) |

我們將圖 2.4 中，包覆在雷射筆上的亮紅色包裝紙上的 RGB 跟 HSI 色彩平均值整理在表 2.1 中。由表 2.1 的數據可以發現，亮紅色紙張的顏色在不同強度的光源下，如果使用 RGB 色彩模型的話，紅色範圍約為 100~210、綠色跟藍色範圍約為 50~110，R、G、B 的散佈範圍相當的廣，在設定色彩過濾的範圍時會有所不便。相對之下，使用 HSI 色彩模型，H 的值幾乎都在 0 左右。

因此，本系統最終選取的是 HSI 色彩模型，而由表 2.1 可以得知，當我們在設定 H、S、I 的過濾範圍時，只要將 H 的範圍設窄一點而把 S 跟 I 的範圍設寬

一點，如此便可以濾掉非雷射筆部分的像素同時也可以抑制光線變化的影響。在本系統中，經過多次實驗之後將 HSI 的範圍設定為色調： $340^\circ \sim 20^\circ$ 、飽和度： $0.15 \sim 1.0$ 、亮度： $0.1 \sim 1.0$ 。圖 2.5 為色彩過濾器的示意圖，圖 2.5(a)為 Hue 值色域示意圖。圖 2.5(b)為本實驗中 HSI 色彩過濾器所使用的 Hue 值範圍示意圖。

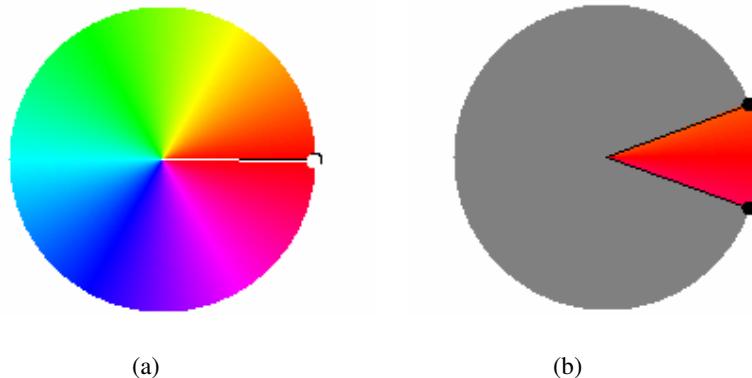


圖 2.5 (a) Hue 值色域示意圖、(b) 本實驗中 HSI 色彩過濾器所使用的 Hue 值範圍示意圖



圖 2.6 (a) 攝影機輸入畫面、(b) 經過 HSI 過濾器處理後的結果

圖 2.6(a)為圖 2.4(a)攝影機擷取的影像，圖 2.6(b)為經過 HSI 色彩過濾器處理之後的結果，可以發現當使用色彩為限制條件之後，場景內大部分的像素值均已被過濾掉不計，而接下來的步驟也只針對色彩符合限制的像素做處理，因此可以大幅加快處理速度。

2.2 利用動態物體偵測作為限制條件

在使用顏色作為過濾條件之後，我們再使用動態物體偵測作為第二個限制的條件。我們假設場景中除了使用者跟用來指向的雷射筆之外的物體都是靜止不動的，因此使用動態物體偵測可以有效的把不會移動的背景部分去除。而為了要完成動態物體偵測的目的，系統需要暫存一張上一個時序的影像，用來與現在這個時序的影像做相減的動作，以數學式子表示為：

$$I_{dif} = I_{t-1} - I_t \quad (2.4)$$

I_{t-1} 、 I_t 分別為時序 $t-1$ 及時序 t 的兩張畫面，而 I_{dif} 為相減後的結果。經過運算得到 I_{dif} 之後，我們再利用一個門檻值（Threshold）來抑制雜訊，避免產生過多誤判區域的情況，大於門檻值的部分我們才視為有 motion 的部分，反之則當作雜訊處理。

圖 2.7 為連續時序所擷取出來的兩個畫面在經過色彩的限制條件後的結果。

圖 2.8(a)為圖 2.7 中兩張影像相減的結果，圖 2.8(b)則是再藉由門檻值的設定，所得到的二值化影像，而圖 2.9 是將圖 2.8(b)對應到目前時序影像所得的結果。我們從圖 2.9 中可以發現畫面中只剩下雷射筆的像素及一些小雜訊，而其他如椅子、牆壁之類的背景部分，在經過動態物體偵測的限制條件之後，就都已經被過濾移除了。



圖 2.7 連續時序影像經由色彩過濾後的結果

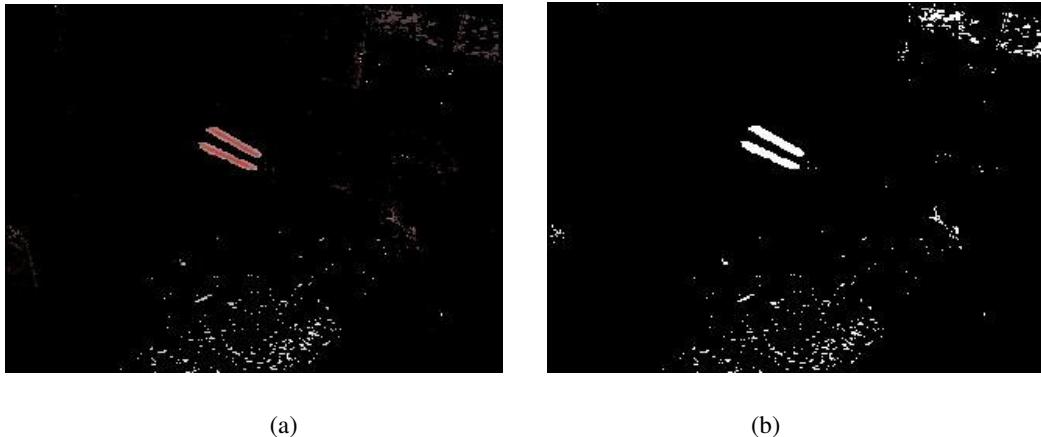


圖 2.8 (a) 連續影像相減之後的結果、(b) 相減影像再經由門檻值 (50) 設定後的二值化影像

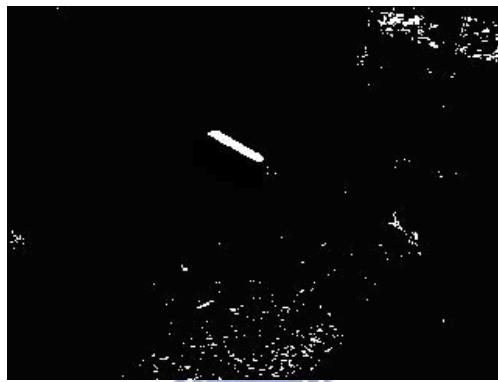


圖 2.9 將圖 2.8(b)對應到目前時序影像所得的結果

在使用過色彩跟動態物體偵測的限制條件之後，影像裡僅剩下雷射筆跟一些雜訊點的像素，接著我們再利用雷射筆的兩個特性：(1) 雷射筆為連接物體，(2) 雷射筆的形狀為長條形，對畫面中的像素做篩選的動作。藉由雷射筆的特性篩選後，我們可以在圖 2.9 中框出一個只包含雷射筆像素的 Bounding Box，如圖 2.10 所示。經由色彩、動態物體偵測、連通物體、形狀的判別後，我們可從影像中得到屬於雷射筆的像素，而接下來的工作，便是找出最能表示這些像素點的直線方程式，藉此直線方程式來找出雷射筆在影像中的起點、終點座標，作為重建使用者指向點的資訊。

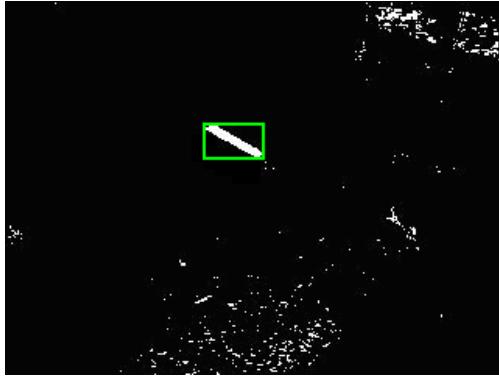


圖 2.10 框出只包含雷射筆像素的 Bounding Box 示意圖

2.3 雷射筆之線形分析

在這節我們要介紹如何由符合限制條件的像素點，來找出代表雷射筆在影像上的直線方程式，這個問題相當於：給予一些座標點，要找出跟這些座標點誤差最小的一條直線。而這個問題有很多數學上的解法，在本節我們將討論兩種求解的方法，分別是最小平方差調適法（Least Square Fitting）與主成分分析法（Principal Components Analysis）。



2.3.1 Linear Least Square Fit

線性最小平方差調適法（Linear Least Square Fitting）之目標是找尋一條直線使得各資料點到此線的距離誤差最小。而各點跟此直線的距離有兩種計算方式：
(1) vertical offsets 及(2) perpendicular offsets，由圖 2.11 可看出兩種方式的不同，vertical offsets 是計算點與線在座標軸垂直方向的距離誤差，而 perpendicular offsets 則是計算點與線垂直的距離誤差。實作上將 vertical offsets 的總和最小化是計算較簡單的方式。以下便針對 linear least square fitting 的計算流程做個介紹。

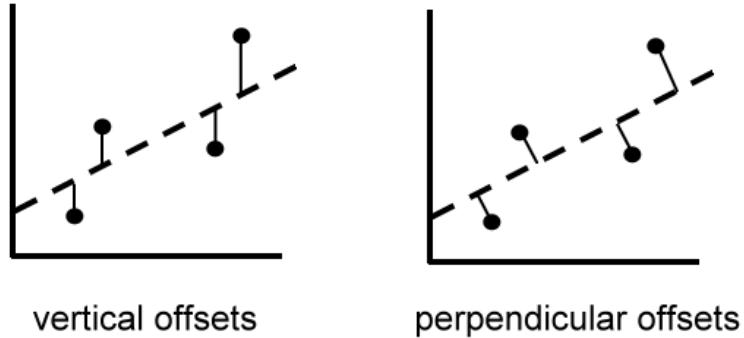


圖 2.11 資料點與線的誤差距離示意圖

已知有 n 個二維空間的點， (x_i, y_i) ， $1 \leq i \leq n$ ，欲使用最小平方差調適法找出一條直線 l ，使各點到 l 的距離平方和最小。假設 l 的方程式為： $y = a + bx$ ，各點到此直線的垂直距離平方和可寫為：

$$R^2(a, b) = \sum_{i=1}^n [y_i - (a + bx_i)]^2 \quad (2.5)$$

$R^2(a, b)$ 即為各點到 l 的距離平方和，欲使 $R^2(a, b)$ 最小化，我們對 $R^2(a, b)$ 的兩個變數 a 、 b 作偏微分，求解最小值發生的地方：

$$\frac{\partial(R^2)}{\partial a} = -2 \sum_{i=1}^n [y_i - (a + bx_i)] = 0 \quad (2.6)$$

$$\frac{\partial(R^2)}{\partial b} = -2 \sum_{i=1}^n [y_i - (a + bx_i)]x_i = 0 \quad (2.7)$$

由式 2.6、式 2.7 可以推得兩個等式，進一步推導便可求得 a 、 b 的值，求解 a 、 b 的算式如下：

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \left[\begin{array}{c} \sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i \\ n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \end{array} \right] \quad (2.8)$$

由式 2.8 求出 a 、 b 後，便可求得直線 l 的方程式，用以表達雷射筆的方向。

在這裡需要注意的是，由於我們在計算最小平方差調適法時，是利用資料點與朝向線的 y 座標垂直距離之平方和當作其誤差。因此，若是資料點的分佈是平行 x 軸的方向，則我們可以找出一條誤差值較小而且能完美代表這些資料點的朝向線；然而，若是資料點的分佈是垂直 x 軸的方向，則計算出來的誤差值將會相

當得大，而且也無法求得一條能代表資料點的朝向線，會導致計算出的雷射筆方程式與實際的雷射筆有嚴重的誤差。由於最小平方差調適法有這樣的特殊例外狀況存在，會影響到我們在實驗測試時的準確度，因此，我們再選取了將 perpendicular offsets 最小化的方法來實作，也就是下一節要介紹的 PCA。

2.3.2 Principal Components Analysis

PCA (Principal Components Analysis) [7]已經在人臉辨識、影像壓縮等領域被廣泛地使用。這是一種將資訊從高維度降低至低維度的方法，此技巧利用線性轉換將資料點轉換到一新座標系，在此新座標系內，所有原座標系內的點座標，其最大變異量會沿著於第一象限軸方向（稱為第一主成分），而次大變異量會沿著第二象限軸方向（稱為第二主成分），依此類推。在將資訊從高維度降到低維度的過程中，PCA 會保留原座標系資訊的特性，以利用維度較低的特徵來充分代表原始高維度資訊的特徵，如此可以加快計算速度並且不至於喪失準確性。以下便針對 PCA 的計算流程做個介紹。

已知有 n 個二維空間的點， (x_i, y_i) ， $1 \leq i \leq n$ ，欲使用 PCA 找出最能表示這些點朝向的線。我們先計算出這些點的重心座標：

$$(\bar{x}, \bar{y}) = \left(\frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n} \right) \quad (2.9)$$

再將每個資料點的座標減去重心座標，做一個將座標系原點平移的動作：

$$(x'_i, y'_i) = (x_i - \bar{x}, y_i - \bar{y}) , \quad 1 \leq i \leq n \quad (2.10)$$

得到了這些微調過的點座標後，我們再利用這些點座標計算共變異數矩陣

(Covariance Matrix)，此矩陣的計算方式如下：

$$C = \begin{pmatrix} \text{cov}(X, X) & \text{cov}(X, Y) \\ \text{cov}(Y, X) & \text{cov}(Y, Y) \end{pmatrix} \quad (2.11)$$

式 2.11 中的 X 及 Y 分別表示 x_i 的座標集合及 y_i 的座標集合，而 covariance 的算法為：

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)} = \frac{\sum_{i=1}^n x_i' y_i'}{(n-1)} \quad (2.12)$$

至此，我們已經由原始資料點算出代表其間相互變異關係的共變異數矩陣，下一步就是對 C 做特徵向量分解 (eigenvector decomposition)，可求得兩個特徵值 (eigenvalue) 及其對應的特徵向量 (eigenvector)。由於選取越大的特徵值可以得到越大的變異量 [7]，所以我們選取數值最大的特徵值所對應的特徵向量作為這些資料點的朝向線。

假設我們選取的特徵向量為 $V = [v_1 \ v_2]^T$ ，則此向量即可表示為這些資料點的朝向。若需進一步求出此朝向線的直線方程式 $l: y = ax + b$ ，則將 v_2 、 v_1 相除可得此線的斜率，又知式 2.9 中求得的資料點的重心必在此線上，因此有了斜率及直線通過的點便可以算出此線的方程式，如下式：

$$l: y = \frac{v_2}{v_1} x + [\bar{y} - (\frac{v_2}{v_1})\bar{x}] \quad (2.13)$$

我們可以用此方程式表示為追蹤到的雷射筆線段。



2.4 總結

在本章中我們介紹了在追蹤指向物所使用的方法，由於本系統在實驗的時候是利用雷射筆作為指向物，因此本章著重的追蹤方法均是針對雷射筆而設計的，但是，如果將來系統提供給使用者用其他的物體來指向，例如以人的手指或人的手臂來指向，我們仍然可以延伸本章的概念，以顏色、Motion、形狀等特徵，來作為追蹤各種指向物的方法。利用了本章所介紹的追蹤方法後，我們可以在影像中追蹤到雷射筆的位置，接著便是使用雷射筆在影像上的座標加上一些數學運算，來求得雷射筆在三維空間的座標，進一步算出指向點的位置。

第三章 基本投影幾何

在本篇論文當中，我們會使用到一些電腦視覺系統相關的觀念。在本章裡，我們將針對各個使用到的觀念做簡單的介紹，其中包含了齊次座標系表示法、投影轉換、三維空間中的平面方程式計算及平面交點求解。此外，由於電腦視覺方面的研究已有相當長遠的一段時間，在本章裡僅選出與此系統較相關的部分，若想要深入探討本章沒有提到的部分，請參考電腦視覺相關書籍[1][8]。

3.1 齊次座標系表示法

在代數學中，我們可以用等式 $ax+by+c=0$ 來表示平面上的任意一條直線，不同的直線會有不同的 a 、 b 、 c 參數組合，因此使用向量的形式 $(a,b,c)^T$ 來表示平面上的一條直線是一種相當方便的選擇。如果一個點 $p = (x_0, y_0)^T$ 在一條直線 $l = (a,b,c)^T$ 上，則將點座標 (x_0, y_0) 代入便可以得到 $ax_0 + by_0 + c = 0$ 這個式子。為了能使用向量內積來簡化這個式子，我們可將 p 點的座標多加上一個 1，重新寫成 $p = (x_0, y_0, 1)^T$ 的形式，則 $ax_0 + by_0 + c = (x_0, y_0, 1)(a, b, c)^T = x^T \cdot l = 0$ ，其中 $(x_0, y_0, 1)^T$ 和 $(a, b, c)^T$ 就是平面上點 x 和直線 l 的齊次座標表示法。

值得注意的是 $ax_0 + by_0 + c = 0$ 和 $kax_0 + kby_0 + kc = 0$ 代表的是同一條直線，因此若在 $k \neq 0$ 的情況下，由向量形式表示的 $(a, b, c)^T$ 和 $(ka, kb, kc)^T$ ，代表相同的一條直線。同樣的情形之下， $ka(x_0) + kb(y_0) + kc = 0$ 和 $a(kx_0) + b(ky_0) + c(k) = 0$ 的意義也是相同的，因此在 $k \neq 0$ 的情況下， $(x_0, y_0, 1)^T$ 和 $(kx_0, ky_0, k)^T$ 表示的是同一個點。也就是說，在 $x_3 \neq 0$ 的情況下， $(x_1, x_2, x_3)^T$ 可以對應到我們熟悉的歐幾里德平面上的 $(x_1/x_3, x_2/x_3)$ 這一點。除此之外，在齊次座標表示法中， $(0, 0, 0)^T$ 這個向量是沒有給予定義的。

齊次座標系對於我們在代數的運算式化簡有相當大的幫助，例如在一個二維空間裡，兩個座標系間平移、旋轉的對應關係，使用了齊次座標系的概念，便可將繁瑣的方程式化簡成一個矩陣相乘的式子，在書寫及運算上均簡潔許多。而我們下一節所要介紹的：平面投影轉換（homography），同樣引用了齊次座標系的概念來簡化計算式。

3.2 平面投影轉換

根據[8]，平面投影轉換（homography mapping）的定義為：一群位於同一平面的點集合 x_i ，在經過投影轉換的處理後，得到一群位於另一平面的對應點集合 x'_i ，其間存在一個投影轉換矩陣，可將 x_i 中的每個點轉換為 x'_i 中對應的點。也就是說，對於點的對應， $x_i \leftrightarrow x'_i$ ，如果考慮齊次座標的表示方式，則可求得一個 3×3 的矩陣 H ，使得兩組對應點和 H 的關係表示如下：

$$Hx_i = \lambda x'_i \quad (3.1)$$

其中， λ 是一個比例常數，而 H 則是個 3×3 的可逆（non-singular）矩陣。

3.2.1 最小對應點數

對於求解投影轉換矩陣 H ，我們遇到的第一個問題是，要有多少個 $x_i \leftrightarrow x'_i$ 的對應點，才能夠求出 H 。這問題可以透過自由度(Degree of Freedom)和限制數(Number of Constraint)的計算，得到對應點數目的下限值(Lower Bound)。首先， H 為一 3×3 的矩陣，共有9個變數，但是因為矩陣內加上了一個縮放係數(up to scale)才會需要9個變數，因此實際上 H 的自由度為8。再者，2D平面上的每個點都有其 x 座標和 y 座標，所以每一組對應的點，就會有2個限制數(分別是 x 座標的限制和 y 座標的限制)。在經由自由度和限制數的計算後，我們可以得知最少要有4組對應點才能求出 H 。因此如果我們有4個點的對應關係，則可求得一個 H ，使之完全符合這4個點的對應關係。

3.2.2 Over-determined情形

接下來考慮更複雜的情形，如果對應點的數目大於4，會有Over-determined的情形產生。若是多出來的對應點跟原來的4個點相容，則不會有任何影響；但如果多出來的對應點跟原來的4個點不相容，則無法找到一個可以把所有點都完美地從 x_i 對應到 x_i' 的投影矩陣 H 。然而這並不代表無法求得一個轉換矩陣 H ，我們還是可以求出一個有些微誤差的 H ，只是我們需要一個額外的cost function，來使這些誤差降到最小，進而得到一個近似的平面投影轉換矩陣(homography matrix)。

3.2.3 Homography Matrix的算法

給定一個2D平面上點的對應集合， $x_i \leftrightarrow x_i'$ ，理想情況下，是期望求出一個轉換矩陣 H ，滿足 $Hx_i = x_i'$ 。另一方面，如果允許 Hx_i 不完全等於 x_i' ，而是具有相同方向的向量，則存在一個不為零的比例係數，亦即 $Hx_i = \lambda x_i'$ ， λ 是一個不為零的實數。以外積的形式來代表 $Hx_i = \lambda x_i'$ ，則我們可以將此式改寫成：

$$x_i \times Hx_i = 0, \text{ with } H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \quad (3.2)$$

我們以 h^{iT} 來表示 H 的第 j 列，如此便可以得到：

$$Hx_i = \begin{bmatrix} h^{1T} x_i \\ h^{2T} x_i \\ h^{3T} x_i \end{bmatrix} \quad (3.3)$$

令 $x_i' = (u_i', v_i', w_i')^T$ ，則可以推導出式3.2中外積的結果為：

$$x_i \times Hx_i = \begin{bmatrix} v_i' h^{3T} x_i - w_i' h^{2T} x_i \\ w_i' h^{1T} x_i - u_i' h^{3T} x_i \\ u_i' h^{2T} x_i - v_i' h^{1T} x_i \end{bmatrix} \quad (3.4)$$

因為矩陣運算的關係，所以 $h^{jT} x_i = x_i^T h^j$ ， $1 \leq j \leq 3$ ，式3.4藉由此關係式化簡後，

可推導出一個以 H 列向量為變數的三個方程式，用矩陣的方法表示如下：

$$\begin{bmatrix} 0^T & -w_i^T x_i^T & v_i^T x_i^T \\ w_i^T x_i^T & 0^T & -u_i^T x_i^T \\ -v_i^T x_i^T & u_i^T x_i^T & 0^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (3.5)$$

有一點要注意，雖然式3.5有三個方程式，但是只有前兩個方程式是線性獨立的(linear independent)，第三個方程式其實是由前兩個方程式所產生，所以我們只需要用到前兩個方程式，把式3.5改寫後得到式3.6：

$$\begin{bmatrix} 0^T & -w_i^T x_i^T & v_i^T x_i^T \\ w_i^T x_i^T & 0^T & -u_i^T x_i^T \end{bmatrix} \begin{bmatrix} h^1 \\ h^2 \\ h^3 \end{bmatrix} = 0 \quad (3.6)$$

由3.2.1節可知，我們只要選取兩個對應點集合中的四組對應點 x_i 、 x_i' ， $1 \leq i \leq 4$ 套入式3.6中便可以產生八個方程式，進而求解出 H ，而求解 H 的演算法詳述如下：

給定 n 個對應點 $x_i \leftrightarrow x_i'$ ， $n \geq 4$ ，求一個矩陣 H ，使得 $Hx_i = \lambda x_i'$ 。

- (1) 對於 $x_i \leftrightarrow x_i'$ 中每組對應的點，可從式3.6得到一個 2×9 的矩陣 A_i 。
- (2) 有 n 組對應點，則可得到 n 個 2×9 的矩陣 A_i ，將這些 A_i 結合成一個 $2n \times 9$ 的矩陣 A 。
- (3) 將 A 做SVD分解，得到 $A = UDV^T$ ，其中 D 為 A 的singular value所構成的對角矩陣(diagonal matrix)、 V 為 singular vector所構成的正交矩陣(orthogonal matrix)。而我們所要求的 H 就是最小的 singular value 所對應的 singular vector，一般而言，SVD分解後的 D 會將singular value由大到小排列，也就是說 V 中第9行的九個元素即為構成 H 的元素。

- (4) V 的第9行為： $[h_1 \ h_2 \ h_3 \ h_4 \ h_5 \ h_6 \ h_7 \ h_8 \ h_9]^T$ ，透過式(3.2)，可以從 V 的第9行得到 H 。

3.2.4 使用平面投影轉換取得指向物的三維空間座標

在第二章結束的時候，我們已經取得影像上雷射筆線段的二維座標，為了要求取指向點，我們必須要將雷射筆的二維座標轉換到三維空間來，而其間的轉換方式，便是利用這節所介紹的平面投影轉換。因為攝影機拍攝出的畫面，與實際的三維空間有著投影轉換的關係，因此，我們只要選取了影像平面與三維空間某個參考平面上共同出現的四個標記點，就可以算出影像平面跟參考平面之間的平面投影轉換矩陣。如圖3.1所示，地面上圈出的四個紅色圓圈便是我們所選取的標記點，而這些標記點的三維座標我們都已經事先量測好。

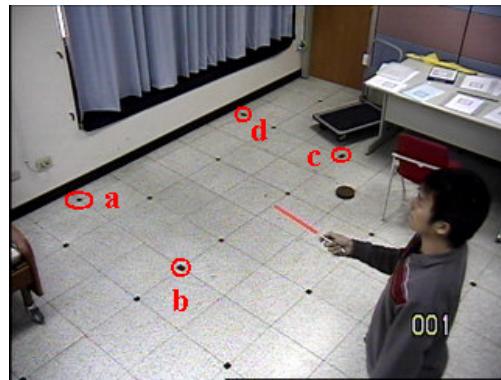


圖 3.1 在三維空間選取的標記點示意圖

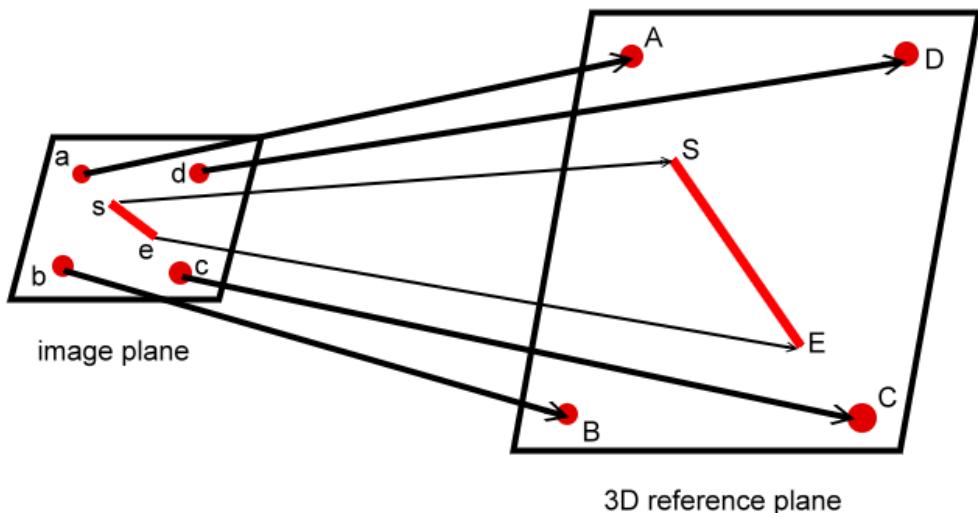


圖 3.2 影像平面與參考平面轉換示意圖

圖3.2為影像平面跟參考平面轉換的示意圖， a 、 b 、 c 、 d 為影像平面上的標記點， A 、 B 、 C 、 D 為三維參考平面的標記點，經由這四點的對應，我們可

以算出兩平面間的投影轉換矩陣 H ，再利用 H 跟影像平面上追蹤到的雷射筆端點 s 、 e 算出在三維參考平面的 S 、 E ，如此我們便完成了使用平面投影轉換求得雷射筆三維座標的目的。

由於左右攝影機所採取的三維參考平面不同，所以會求解出兩組雷射筆端點在參考平面上的座標，我們假設這兩組座標為 $\{S_L, E_L\}$ 、 $\{S_R, E_R\}$ ，而左右攝影機的相機中心座標為 $Camera_L$ 、 $Camera_R$ ，則我們可用 $\{S_L, E_L\}$ 與 $Camera_L$ 交出一個平面 π_L ；同理， $\{S_R, E_R\}$ 、 $Camera_R$ 也可交出另一個平面 π_R 。 π_L 與 π_R 會交出一條為雷射筆方向的指向線（pointing line），接著我們便可利用這條指向線與指向平面的交點求出使用者的指向點。而這些相關的數學運算，將在下一節做詳細的介紹。

3.3 三維空間中的相關數學運算

在本論文中，我們在求解指向點的三維空間座標時，需要利用到一些三維空間裡的代數運算，例如有三個點座標，求此三點所構成的平面方程式；有三個平面方程式，求這些平面的交點座標。因此，在本節中，我們將介紹如何利用相關數學運算求得我們所需的結果。

3.3.1 三維空間中三點求解平面方程式

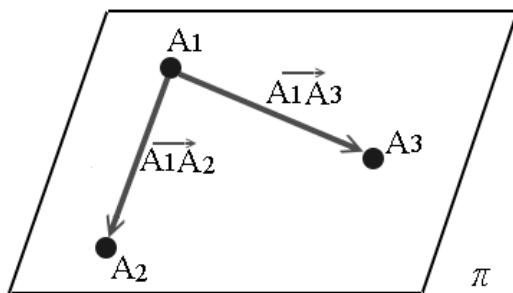


圖 3.3 同平面上三點示意圖

在三維空間中給予三點座標為 $A_i = (x_i, y_i, z_i)$ ， $1 \leq i \leq 3$ ，欲求解此三點所構

成的平面 π 之方程式，點與平面的關係如圖3.3所示。首先我們知道一個平面方程式的表示法為： $ax + by + cz = d$ ， $\vec{N} = (a, b, c)$ 為平面之法向量，又 \vec{N} 與平面上的任意向量均垂直，因此我們可以利用已知的三點構成 $\overrightarrow{A_1A_2}$ 、 $\overrightarrow{A_1A_3}$ 兩向量，再以此兩向量的外積來求得 \vec{N} ，算式如下：

$$\vec{N} = (a, b, c) = \alpha(\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3}) \quad (3.7)$$

α 是一個比例常數，因為 $\overrightarrow{A_1A_2} \times \overrightarrow{A_1A_3}$ 的結果為 \vec{N} 相同方向的向量，但數值未必會剛好等同於 \vec{N} ，在一般情況我們可以將 α 視為1而忽略不計。求得 \vec{N} 之後，只要再任意選取一點座標即可算出 d 的值，以 A_1 為例：

$$d = ax_1 + by_1 + cz_1 \quad (3.8)$$

至此，我們已經推導出如何由空間中任意三點求得所構成的平面之方程式，接下來我們要介紹如何由三平面求其交點的方法。



3.3.2 三維空間中三平面求解交點

空間中任意三平面方向之間的相互關係有很多種，例如：

- (1) 三個平面互相平行
- (2) 兩平面平行，第三平面與此兩平面各交出一條直線
- (3) 三平面交於一條直線
- (4) 三平面兩兩交於一條直線，而這三條直線互相平行
- (5) 三平面交於一點

在本篇論文當中，我們在求解指向點的時候，會先由左右兩個不平行的平面交出雷射筆的指向線，再以此指向線與指向平面相交求指向點，因此我們可以假設只會出現上述的第五種情形，而將(1)~(4)種情形以例外情況處理而不探討之，因此以下所討論的方法為求解三平面交於一點的方法。

已知三平面 π_1 、 π_2 、 π_3 ，其平面方程式為： $a_i x + b_i y + c_i z = d_i$ ， $1 \leq i \leq 3$ ，利用[10]中所介紹的Cramer's Rule來求此線性方程組之唯一解。Cramer's Rule需先計算四個行列式(determinant)的值，如下所示：

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \Delta_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, \Delta_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, \Delta_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} \quad (3.9)$$

由平面方程式之係數求解出 Δ 、 Δ_x 、 Δ_y 、 Δ_z 的值後，套用Cramer's Rule即可求出三個平面交點的座標：

$$(x, y, z) = \left(\frac{\Delta_x}{\Delta}, \frac{\Delta_y}{\Delta}, \frac{\Delta_z}{\Delta} \right) \quad (3.10)$$

在此需要注意的是，因為我們只探討上述的第五種情形，所以計算出的 Δ 並不會等於0，因此我們必然可以透過Cramer's Rule求出此三平面之交點。



3.3.3 系統實作用到的數學運算

這上節中，我們介紹了一些數學上求平面跟求交點的運算，現在則要講解我們在哪些實作的步驟時會用到這些數學運算。延續3.2.4節最後一段的敘述，我們利用平面投影轉換得到雷射筆在不同參考平面的三維座標 $\{S_L, E_L\}$ 、 $\{S_R, E_R\}$ ，而且事先量測了左右攝影機的相機中心座標，假設為 $Camera_L$ 、 $Camera_R$ 。

我們可使用3.3.1節的三點求平面的方法，利用 S_L 、 E_L 、 $Camera_L$ 三點，算出其構成的平面 π_L ；同理，使用 S_R 、 E_R 、 $Camera_R$ 三點也可以求出另一個平面 π_R 。假設指向平面為 $\pi_{Projection}$ ，則 $\pi_{Projection}$ 、 π_L 與 π_R 三平面所交的點即為使用者的指向點，在求解指向點的運算便可利用3.3.2節介紹的三平面求交點的方法來完成。因此，我們利用介紹的相關數學運算便可以推導出使用者指向點的三維空間座標為何，至此便完成重建指向點的目標。

3.4 總結

在本章介紹了一些我們有使用到的電腦視覺系統的概念以及相關的三維空間數學的計算，對於了解整個系統的流程及實作將會有所助益。



第四章 準確性估測的實驗結果

基於前面章節的介紹，我們已經開發出一套可以提供給使用者進行即時指向的指向系統，然而此系統之穩定性並不如我們所預期，為了要讓系統在使用上更為方便、可靠，我們需分析導致系統不穩定的因素並且嘗試解決之；同時，除了穩定性的問題之外，系統的精確度也是一個我們需考量的主題。我們將針對系統既有的誤差建立一個誤差模型，並且分析導致系統誤差的來源，期望能使用後置的處理，將此誤差降低至可容忍的範圍。系統開發的環境為 Windows XP 作業系統之個人電腦，CPU 為 P4 3.2G，2G 記憶體，程式平台為 Microsoft Visual Studio 2005，程式語言為 C#，用來拍攝使用者指向畫面的攝影機為圖 4.1 所示型號 D7722 的 PTZ 攝影機。



圖 4.1 指向系統實驗所使用到的 PTZ 攝影機

本系統的實驗流程包括了在影像中追蹤雷射筆的位置、實際量測三維空間的座標、計算影像平面與三維參考平面間的平面投影轉換矩陣、計算指向點位置。圖 4.2 是實驗場景的俯視示意圖，圖中標示了我們所量測的三維空間的座標，圖中的座標均以公分為單位。我們設定地面為 XY 平面，垂直 XY 平面則設為 Z 軸。圖中兩個黑色圓圈為安裝在天花板的 PTZ 攝影機、八個紅色的點則是用來計算左右攝影機平面投影矩陣的標記點。使用者站在場景中央的位置，左攝影機 C_1 由使用者的左後方拍攝、右攝影機 C_2 則由使用者的右前方拍攝，如此計算出的兩個平面之交線即為包含雷射筆的指向線。我們使用壁報紙黏貼在場景內的窗戶上作為我們的指向平面，而這個指向平面的方程式為： $Y = -20$ ，如此我們可以利

用壁報紙作為投影機的投影布幕，可直接把結果利用投影機投射出來利於使用者觀察。

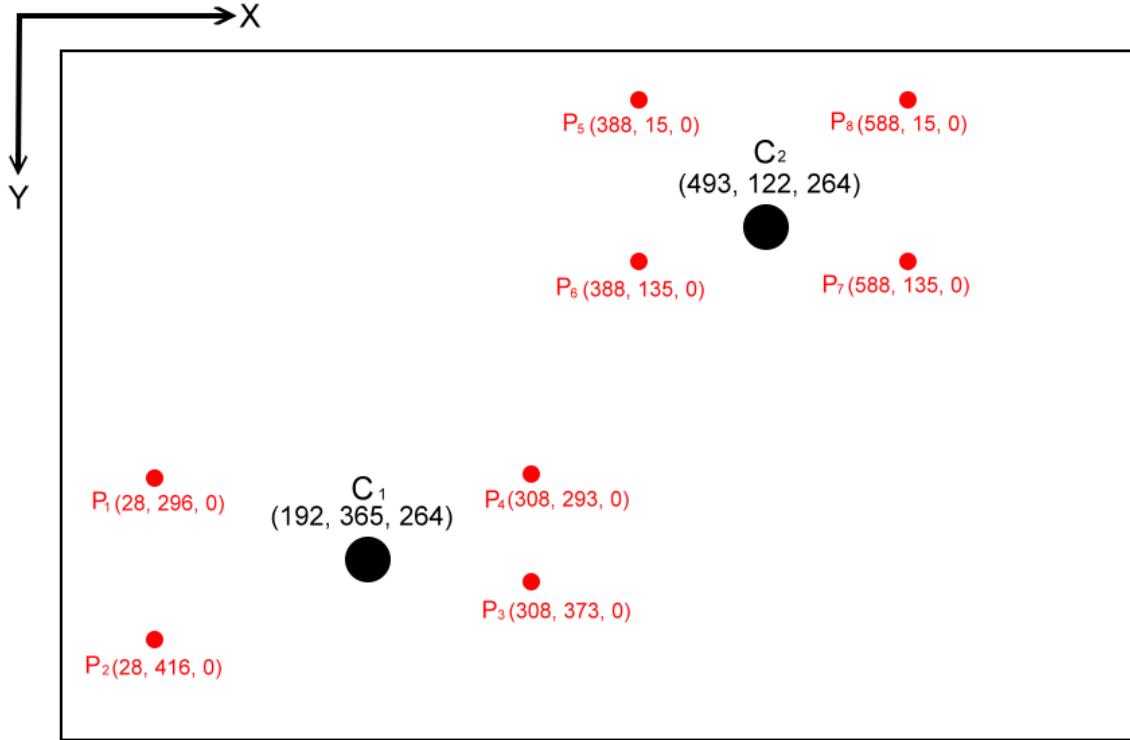


圖 4.2 實驗場景俯視示意圖

由於左右攝影機拍攝到的畫面無法照到相同的參考平面，因此左右攝影機的平面投影矩陣是使用不同的參考點各別計算出來的，而左攝影機 C_1 所使用的標記點為圖 4.2 中的 $P_5 \sim P_8$ ，右攝影機 C_2 使用的標記點為 $P_1 \sim P_4$ ，我們計算後取得的兩個平面投影轉換矩陣如下， H_L 、 H_R 各別為 C_1 、 C_2 使用的矩陣：

$$H_L = \begin{bmatrix} -0.00231184626042929 & -0.00137555370383911 & -0.340319115902899 \\ -0.00176139008075235 & -0.00516421929625191 & 0.94029008261871 \\ 6.8813341479990 * 10^{-7} & -1.0782541575954 * 10^{-5} & -0.000671808121441 \end{bmatrix}$$

$$H_R = \begin{bmatrix} -0.00178461532197717 & 0.00479595448795168 & 0.105894274445466 \\ -0.00186676553362327 & -0.00022812316168039 & 0.994362050347551 \\ 1.53311725714096 * 10^{-7} & 7.3934444947710 * 10^{-6} & 0.000890240732828 \end{bmatrix}$$



圖 4.3 使用者使用指向系統側視圖

圖 4.3 為使用者在使用本系統進行指向時的側視圖，使用者站在離指向平面約 4 公尺遠的地方進行指向的動作，圖中可觀察到一台安裝在天花板的 PTZ 在使用者的斜後方偵測使用者的動作，同時，我們利用投影機將系統運算出的指向點展示在指向平面上。圖中的紅點為雷射筆打出的紅光，而我們操控滑鼠指標來表示系統運算出的指向點。圖 4.4 則為指向系統使用中的畫面，紅點是雷射筆所打出的紅光在指向平面上的位置，同時我們利用系統運算出的指向點來控制滑鼠的位置，由圖可以發現紅點跟滑鼠箭頭相對的距離關係，也可由此觀測我們的系統的準確度。

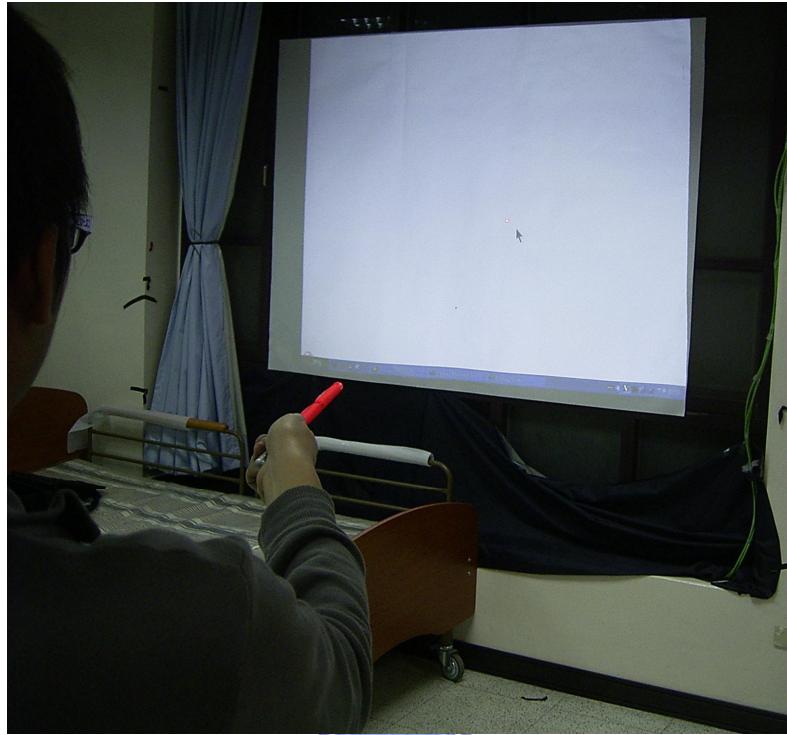


圖 4.4 指向系統結果示意圖

在介紹完整個實驗系統的場景、設備之後，我們接著要針對本系統的穩定性以及準確度進行分析。我們將會探討導致本系統不穩定的因素，並且嘗試一些實驗來使得系統穩定度提高；同時我們分析系統的誤差情況，並且期望使用後置的修正來使指向點的重建結果更為準確。

4.1 系統的穩定性分析

本系統實作完成之後，在實驗的階段裡，我們發現系統重建出的指向點能夠跟著使用者指向的移動而作正確方向的移動，然而，指向點跳動的情形卻是相當地嚴重，常常會發生跳動到不正確的位置然後再跳回正確的位置的情形。同時，當使用者靜止的時候，重建出的指向點也不會固定停止在某個位置，而是會持續地跳動。因此，在分析本系統的準確度之前，我們必須先找出導致系統不穩定的原因，並且嘗試改進系統的穩定度。我們分析實作步驟中可能會導致系統不穩定的步驟，列舉於下：

- (1) 追蹤雷射筆：雷射筆的長度太短、提供給系統追蹤的像素太少，導致使用者些許地移動會造成雷射筆線段的變化過於劇烈，如此便會影響到系統的穩定度。
- (2) 使用平面投影轉換將點投射到三維參考平面：若是投影矩陣有誤差，導致點的對應有錯，可能會進而影響系統的穩定度。

首先分析第二點，我們在計算影像平面跟參考平面的平面投影矩陣時，無論是在量測三維空間的座標或者是在點選影像上的標記點時都仔細操作，注意不要造成太大的偏差。而且若是兩個平面間的投影矩陣有誤差，結果只會使對應到參考平面上的座標有誤差，進而對系統的準確度有影響，並不會導致系統產生指向點跳動不穩定的現象，因此我們認為第二點並非是導致系統不穩定的原因。至於第一點，我們認為不穩定的雷射筆線段追蹤，會造成推算指向點時的誤差被放大許多，因此很有可能影像上的雷射筆線段有些許變化就造成指向點有很大的差異，而在視覺上就呈現指向點一直跳動的情況。為了要證實我們的推論，我們嘗試延伸雷射筆上的亮紅色包裝紙長度，從原本的 20 公分延伸為 50 公分，圖 4.5 即為使用比較長的亮紅色包裝紙包覆雷射筆的示意圖。我們增加亮紅色包裝紙在畫面中所佔有的像素數量，用來測試不穩定的因素是否來自於此部分。



圖 4.5 亮紅色包裝紙長度較長的雷射筆

在開始測試之前，我們先將在求解 PCA 步驟時的誤差輸出，觀察 PCA 求解出的直線方程式對於資料點 fitting 的誤差情況如何。實驗的結果誤差均呈現在一個相當固定的範圍內，我們選擇一些實驗結果整理於圖 4.6 中，圖 4.6 中橫軸為畫面序號，每點的間距為 0.1 秒；縱軸為誤差結果的變化，以像素為誤差單位；而圖 4.6 中(a)、(b)各約為 200 個畫面的結果。由圖 4.6(a)可以發現誤差幾乎都在 0.6~1.5 之間異動、圖 4.6(b)的誤差則約在 0.9~1.5 之間變化，因此無論是原本長度或是加長長度的雷射筆線段，經由 PCA 計算後的誤差均在可容忍的範圍內。

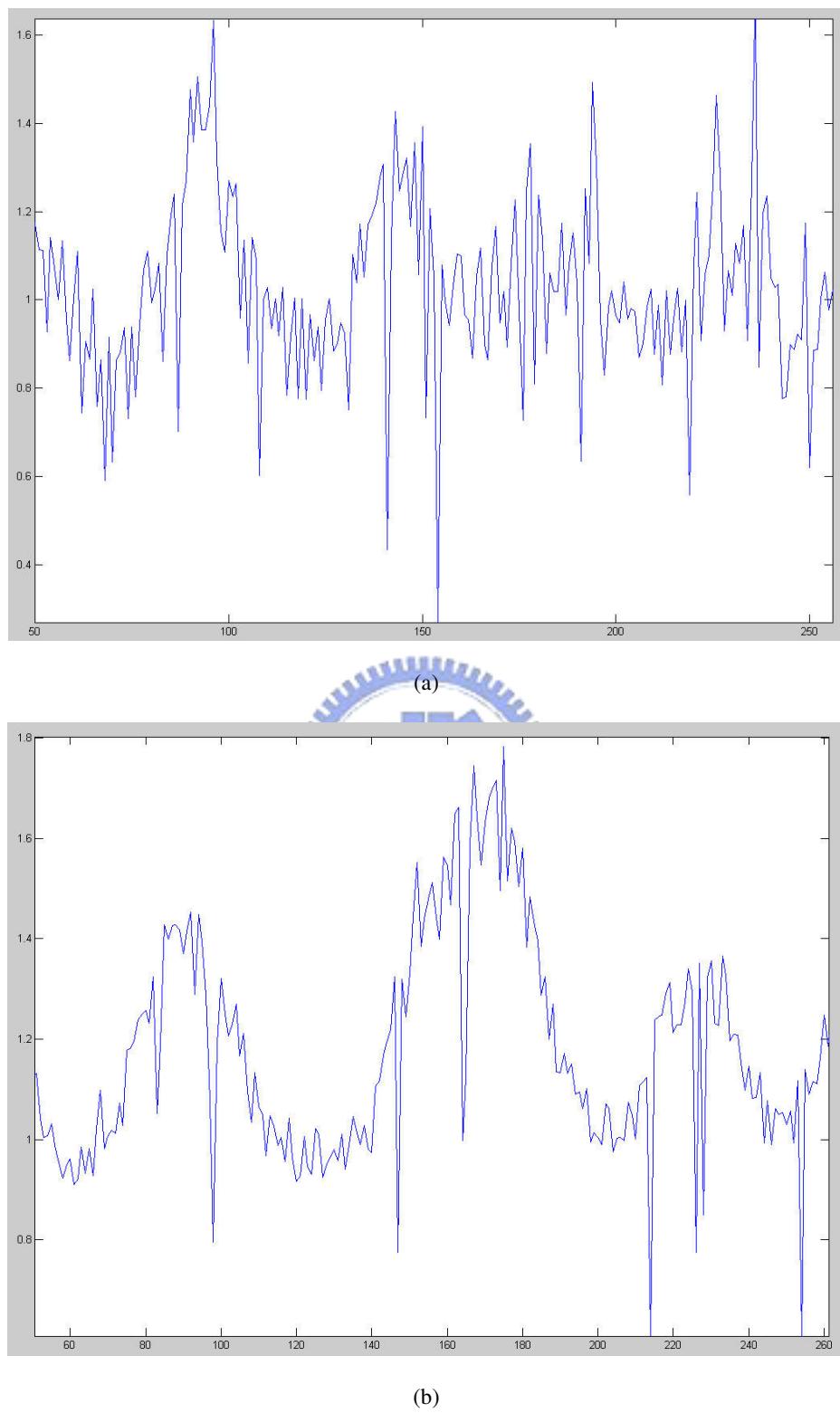


圖 4.6 雷射筆沿平行地面方向來回移動之資料點的 PCA fitting 誤差

(a) 使用原本長度的雷射筆、(b) 使用加長長度的雷射筆

我們為了再仔細觀察不同長度的亮紅色包裝紙對穩定度的影響，除了觀察 PCA fitting 資料點的誤差之外，再進一步將程式中計算出的雷射筆直線方程式： $y = ax + b$ 裡的 a （斜率）輸出，觀察直線方程式的斜率是否為平滑的變化或者是跳動的很厲害，企圖利用斜率這個數值來觀察在影像上所追蹤到的雷射筆直線之變化是否為不穩定的。

在實驗時，我們將原本長度及加長長度的雷射筆作平行和垂直地面的兩個方向之平滑的週期性移動，因此在影像中的雷射筆線段之斜率會呈現一個週期性的變化，會有遞增、遞減、再遞增的反覆週期性變化。圖 4.7～圖 4.10 的橫軸為畫面序號，每點的間距為 0.1 秒，縱軸為直線方程式斜率的變化，(a)、(b)則各表示兩次不同的測試片段且每次測試片段約為 200 個畫面的結果。

圖 4.7 是使用原本長度的雷射筆線段作平行地面方向做來回移動之斜率變化，由圖中我們可以發現線段斜率大致上存在規律的變化，不過部份時間點之間存在有雜訊造成的劇烈變化，同時斜率跳動的情況也是相當地嚴重。而圖 4.8 是使用較長的雷射筆線段作平行地面方向做來回移動之斜率變化，觀察圖 4.8 我們可以發現與圖 4.7 的差異，圖 4.8 中線段的斜率同樣呈現規律的變化，雖然還是有些突然隆起的數值出現，不過斜率跳動的情況並不多，同時每個時間點之間的變化數值也不會差異太多，整體而言數值是呈現相當平滑地變化。

圖 4.9 與圖 4.10 則是將雷射筆沿著垂直地面的方向做來回移動，我們可以發現使用原長度的雷射筆線段的斜率跳動得更加嚴重，雖然整體的趨勢還是存在，不過每個時間點之間的斜率幾乎沒有平順地變化，呈現一個相當不穩定的情況。相對的，使用較長的雷射筆線段所得到的斜率變化就平滑了許多，斜率跳動的情形也幾乎沒有出現。

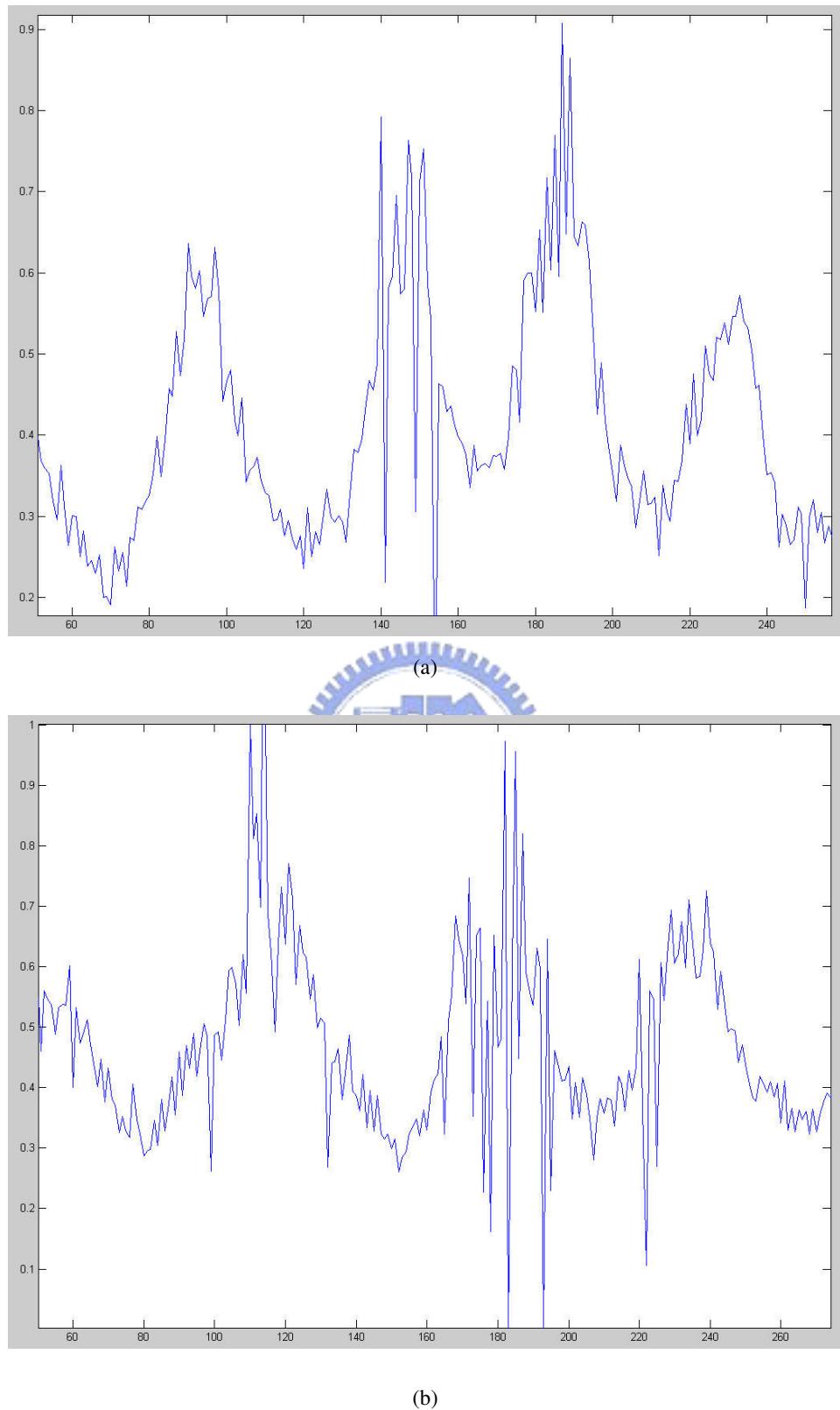


圖 4.7 原本長度的雷射筆指向物沿平行地面方向來回移動之指向物線段斜率變化

(a) 測試片段 1、(b) 測試片段 2

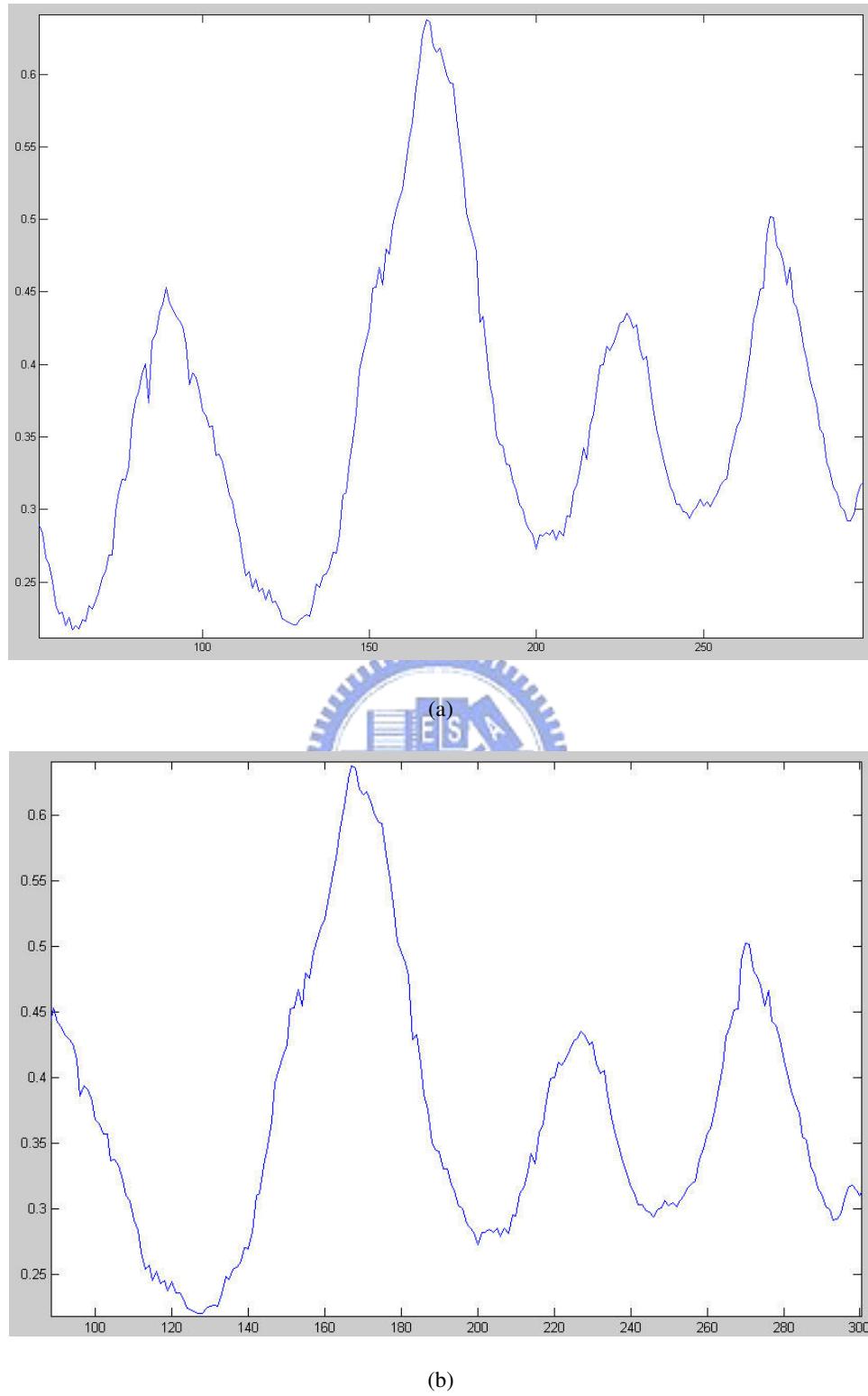


圖 4.8 較長的雷射筆指向物沿平行地面方向來回移動之指向物線段斜率變化

(a) 測試片段 1、(b) 測試片段 2

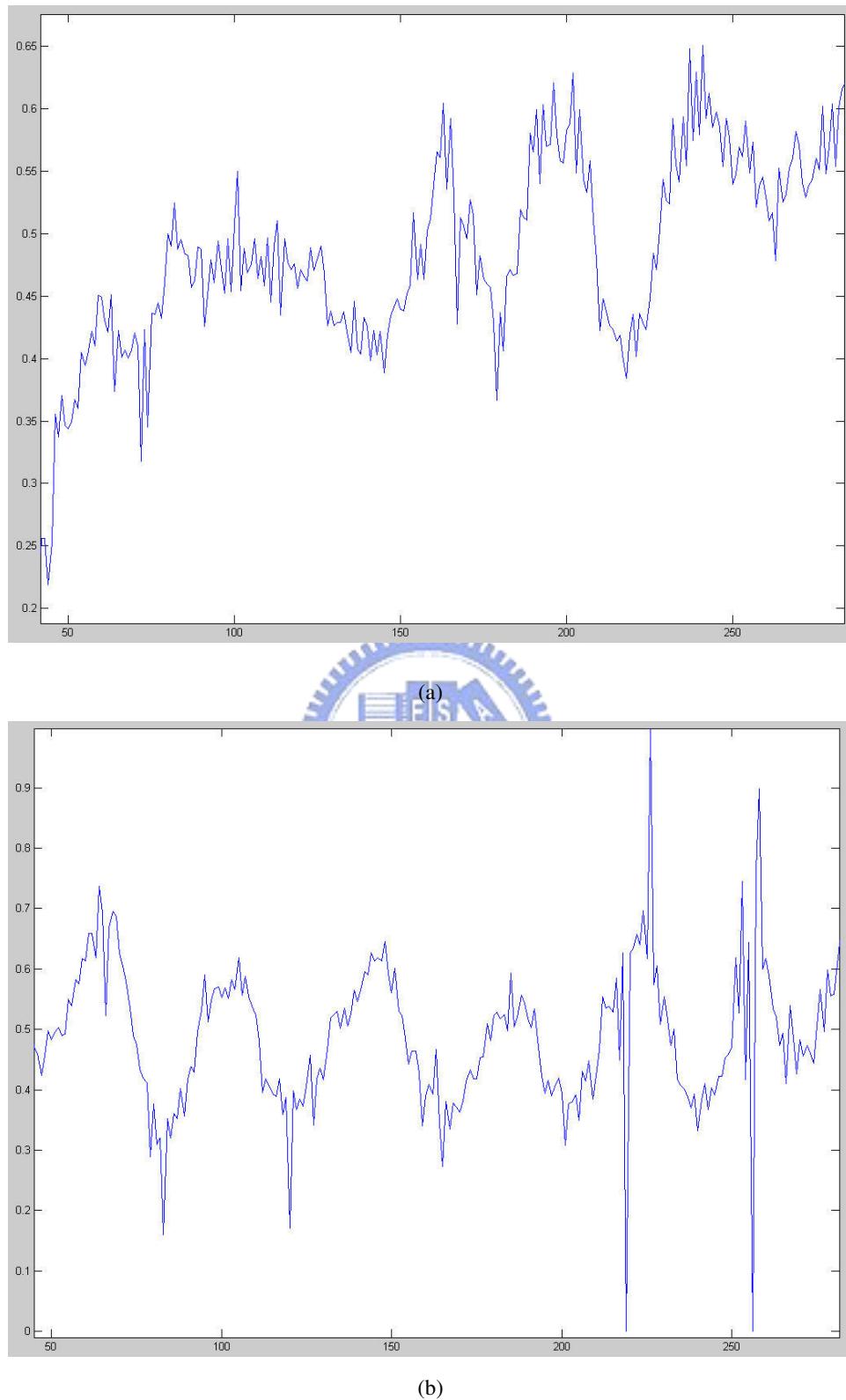
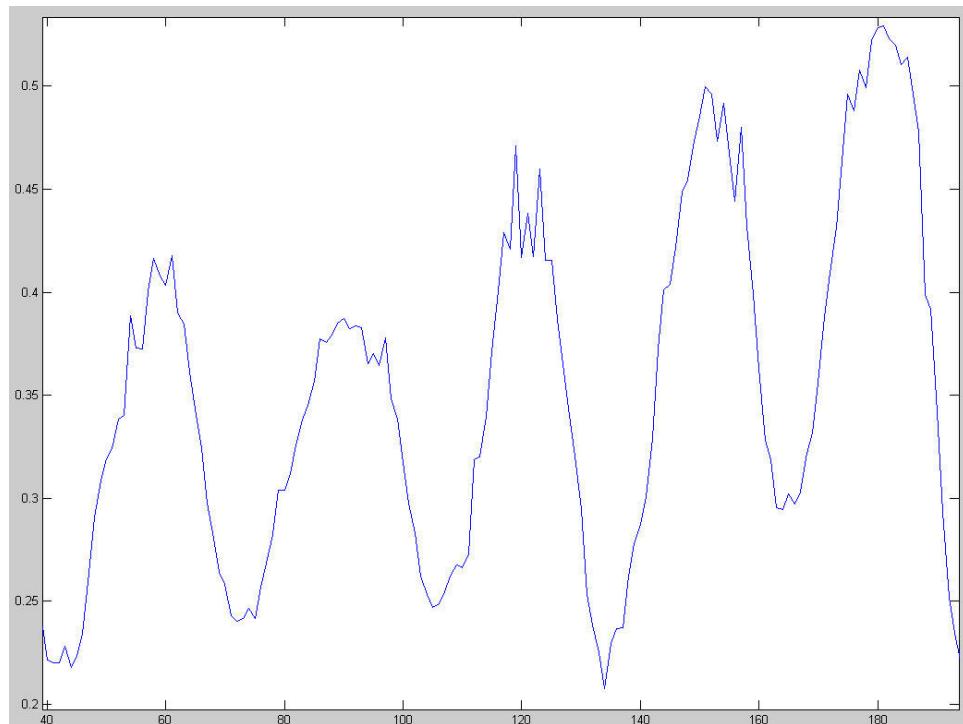
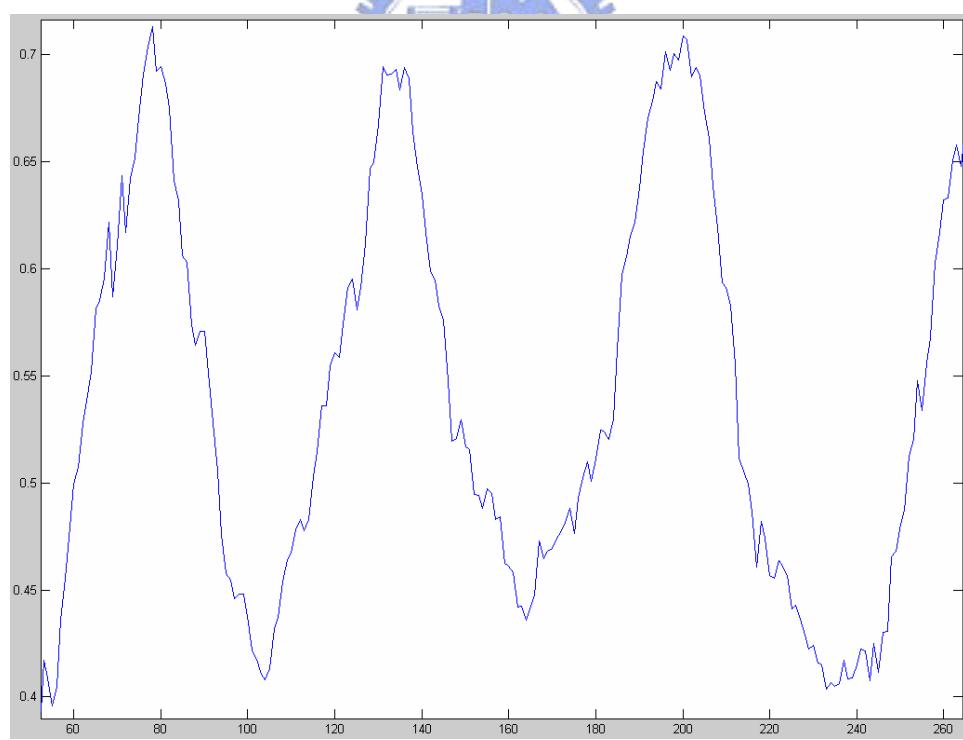


圖 4.9 原本長度的雷射筆指向物沿垂直地面方向來回移動之指向物線段斜率變化

(a) 測試片段 1、(b) 測試片段 2



(a)



(b)

圖 4.10 較長的雷射筆指向物沿垂直地面方向來回移動之指向物線段斜率變化

(a) 測試片段 1、(b) 測試片段 2

在本節的討論裡，我們發現了一個導致系統不穩定的因素，並且藉由實際的實驗數據驗證了我們的推論：包覆雷射筆的亮紅色包裝紙尺寸對於系統重建指向點的穩定度有絕對性的影響，藉由延伸紙張的長度、增加亮紅色包裝紙在畫面中所佔有的像素數量的確可以提高系統的穩定度。然而，在實際使用了較長的亮紅色包裝紙實驗之後，本系統重建的指向點仍然會有些許不穩定的跳動，雖然跳動的情況已經較原本的情況改善許多，不過若是想要再進一步增進系統的穩定性，則需要更多的分析及研究輔助。

4.2 系統的準確性分析

在針對系統的穩定性做了分析與探討之後，我們將針對較為穩定的系統來分析其準確性，我們在實驗時利用包裝紙長度較長的雷射筆（見圖 4.5）來做準確性的實驗，企圖利用穩定性較高的實驗器材分析系統的準確度。一般而言，可能導致系統不準確的一些誤差來源，包括：(1) 三維空間的座標量測不準確，包含用來計算平面投影轉換的點之座標、相機中心的座標、投影平面的座標、(2) 投影平面並不是完全平坦，而且傾斜的情況、(3) 量測得到的相機中心的位置並不是實際的相機中心位置，因為我們不能準確地取得相機中心的座標，而僅能用估測的方式得到大概的相機中心位置。以上提到的這些誤差來源均是難以避免而僅能仔細量測跟配置來避免誤差過於龐大，因此我們希望利用一些後置的處理，來修正必然存在的誤差來源所導致的誤差，期望能用簡單的方法來提高系統的準確度。

在實際使用我們開發出的指向系統時，可以由滑鼠的游標跟雷射筆紅點的相對位置看出系統的誤差情況，而我們在使用中觀測到滑鼠的游標與雷射筆的紅點相對位置大約為游標在紅點左下角的位置，因此我們預期誤差的情況會以一個固定的形式出現。例如圖 4.11 中的黑色長方形是指向平面上之投影布幕的示意圖，而紅色的四邊形則是系統可能的誤差情況，圖 4.11(a)表示系統重建出的投影

布幕之四個角點的位置僅是原本的四個角點位置往左下角平移而沒有任何形變扭曲的情況，圖 4.11(b)的情況則是四個角點的位置不僅往左下角平移而且每一個角點平移的距離不相同，造成長方形形變為不規則的四邊形之結果。然而不論是僅有平移的情況或是還加上形變的情況，在理想的狀況下，只要誤差的情況是以一個固定的形式出現，我們均可利用四個角點的對應關係，使用 3.2 節介紹的平面投影轉換方式來修正之。

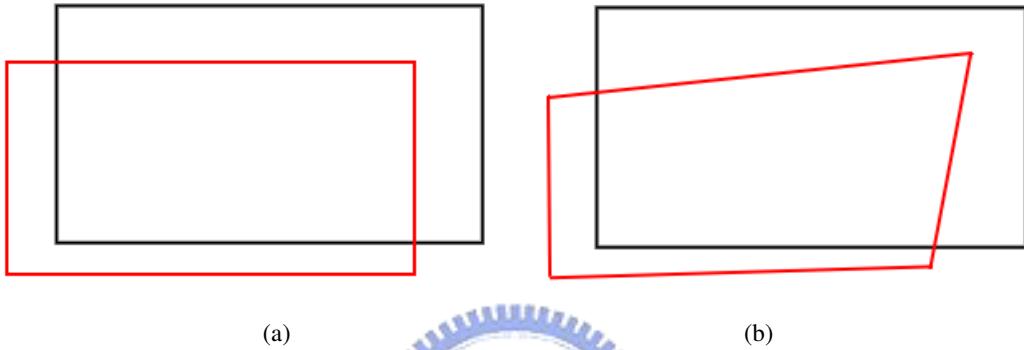


圖 4.11 理想的指向系統之誤差結果，(a) 僅有平移的情況、(b) 有平移及扭曲的情形

為了檢驗系統的誤差情況是否為固定形式出現的情形，我們使用雷射筆沿著投影布幕的四個邊來回的移動，用以得到系統重建出的投影布幕的四個邊，同時為了要減少使用者使用上的誤差，我們將雷射筆放在一個平面板子上，讓雷射筆在同一個平面上移動，因此所有的誤差來源均是系統既有的誤差情況而與使用者的使用無關。圖 4.12 是其中一次實驗結果的資料，圖中左側的青色點是我們使用雷射筆對著沿著投影布幕左側的邊來回移動所重建出的指向點，而右側的黃色點、上方的綠色點、下方的藍色點也是沿著投影布幕各個邊來回移動所得到的結果。由於系統尚不穩定以及一些誤差的影響，造成這些指向點是在投影布幕邊線的附近移動，因此我們使用 PCA 來對這些點做 fitting 的動作，而四個紅色的邊就是我們使用 PCA 所得到的 fitting line。我們利用 PCA 運算得到的四個邊線所建構出的四邊形來當作系統誤差的觀察準則，把這個紅色的四邊形跟投影布幕的矩形來比較，用以觀察系統的誤差情況為何。

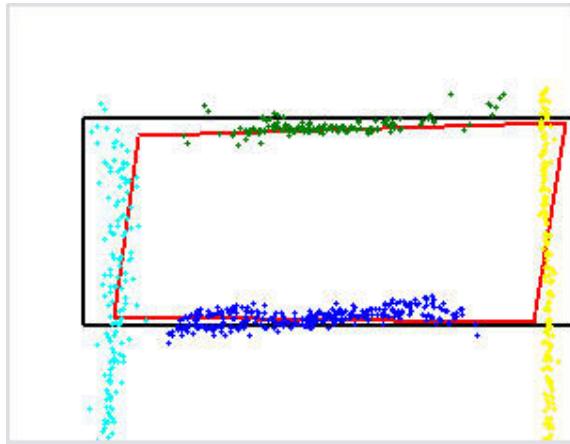


圖 4.12 實際的指向系統之誤差實驗結果

在經過多次的實驗之後，我們發現系統的誤差情況並不如我們所預期的理想。圖 4.13 中我們列舉了幾次實驗的結果，我們觀察到了僅有四邊形左上角的角點在各次實驗裡均是位於原先角點的右下角，而其他三個角點，則有時在原先角點的左下角或右上角，並沒有一致的模式。我們猜測這是由於系統的不穩定所導致的情況，由於系統還不夠穩定，因此重建點的位置跳動、偏差的情況時常會出現，造成我們將雷射筆沿著投影布幕每個邊移動的時候，重建點跳動的情形不一致，而所得的誤差結果也就都不一樣。因此在這種情況下，我們無法如預期的想法，使用平面投影轉換的方式來修正誤差，因為可能會導致重建點在某些角點的誤差會更為嚴重。

然而即使系統的準確性尚未完美，我們仍然能利用此系統開發一些應用。例如[10]便是利用此系統開發出一個人機互動的使用介面，利用使用者在指向平面上繪出的軌跡來當作控制電腦的指令，例如使用者繪出一個 O 的軌跡便代表開啟檔案的指令。而我們也可利用此系統開發其他的應用，例如在指向平面上放置幾個按鍵，當使用者指到某個按鍵時，便會通知電腦進行相對應的指令，例如指到『開啟瀏覽器』的按鍵，電腦便會自動開啟瀏覽器，不過這種不需要使用者繪出軌跡而僅需停止動作指在某個位置的靜態之應用，在系統準確度上的限制就嚴格得多，因為越精確的系統，可以在指向平面上安置越多的按鍵供使用者使用，相反的，較不精確的系統，就只能提供少量的按鍵功能給使用者選擇。

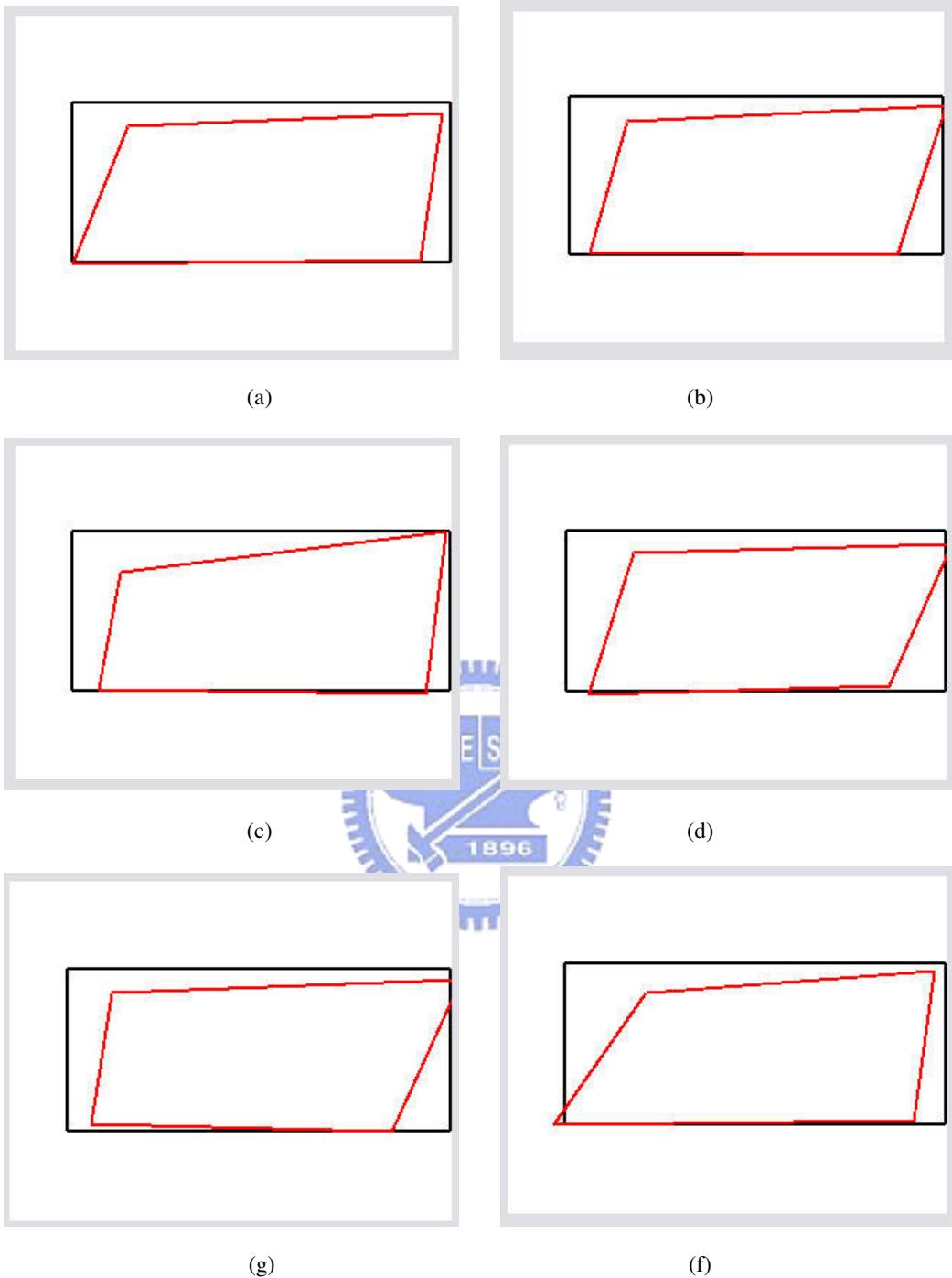


圖 4.13 實際的指向系統之誤差結果，(a)~(f)各次測試過的結果

4.3 總結

在本章裡，我們針對已經建構出的即時指向系統做了穩定性及準確性的分

析並且嘗試解決穩定性及準確性的問題。在實驗中，我們發現的二維影像上的雷射筆線段追蹤對於穩定性很重要，而增進穩定追蹤的方法之一便是增加雷射筆所佔有的像素，就如同我們實驗時，將亮紅色的包裝紙延伸的結果，實驗也證明了如此的確可以增進系統的穩定性。準確度方面，我們試圖使用後置的平面投影轉換來修正系統既有的誤差，然而由於系統的不穩定導致誤差的結果並不一致，而無法使用預期的方法來修正。我們針對穩定性及準確度做了初步的分析與改進，然而如何將系統改進至可實際應用的階段，仍是一個相當重要的課題。



第五章 結論與未來工作

在本篇論文當中，我們提出了一個較為簡單的方法來實作指向系統，我們利用了影像平面與三維參考平面間的平面投影轉換關係，求得指向物在三維參考平面上的座標，再利用相機中心位置的資訊求得兩個平面以交出指向線，最後再利用此指向線求出與指向平面的交點，即為指向點。系統實作上僅需要兩台攝影機偵測使用者動向即可，我們的系統可以讓使用者在距離指向平面約3~4公尺的距離進行跟指向的動作，對於改善使用者跟電腦的互動有很大的幫助。

本系統在實驗階段時，提供使用者使用的指向物為一支雷射筆，而在攝影機畫面中雷射筆所佔有的像素數量並不多，因此些許追蹤的誤差就會對指向點重建的結果有很大的影響，在接下來的工作，可以嘗試使用人的手臂或其他體積較大的物體來提供使用者進行指向的動作，用以增進系統的穩定度。此外，由於系統會具有一些不可避免的誤差，例如量測場景、計算的誤差，將來可以探討更多的後置處理，進一步降低誤差來增進系統的準確度。最後，使用此系統來開發一些應用也是可以進行的工作，例如簡報系統、空間物體指向的確認等，均可利用此系統作為開發的基礎。

參考文獻

- [1] D. Forsyth and J. Ponce, *Computer Vision A Modern Approach*, Prentice Hall, 2002.
- [2] Y. P. Hung, Y. S. Yang, Y. S. Chen, I. B. Hsieh, C. S. Fuh, "Free-Hand Pointer by Use of an Active Stereo Vision System," *Proceedings of 14th International Conference on Pattern Recognition*, pp. 1244-1246, 1998.
- [3] Y. Yamamoto, U. Yoda, and K. Sakaue, "Arm-pointing Gesture Interface Using Surrounded Stereo Cameras System," *Proceedings of 17th International Conference on Pattern Recognition*, Vol. 4, pp. 965-970, 2004.
- [4] C. Colombo, A. D. Bimbo, and A. Valli, "Visual Capture and Understanding of Hand Pointing Actions in a 3-D Environment," *IEEE Transactions on System, Man, and Cybernetic-Part B: Cybernetics*, Vol. 33, No. 4, pp. 677-686, 2003.
- [5] E. Hosoya, H. Sato, M. Kitabata, I. Harada, H. Nojima, and A. Onozawa, "Arm-Pointer: 3D Pointing Interface for Real-World Interaction," *European Conference on Computer Vision Workshop on Human Computer Interaction*, pp. 72-82, 2004.
- [6] M. S. Lee, D. Weinshall, E. Cohen-Solal, A. Colmenarez and D. Lyons, "A Computer Vision System for On-Screen Item Selection by Finger Pointing," *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, pp. 1026-1033, 2001.
- [7] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Elsevier Academic Press, 2003.
- [8] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2003.
- [9] R. Larson and B. H. Edwards, *Elementary Linear Algebra*, Houghton Mifflin Company, 2000.
- [10] J. C. Hsu, "An Implement of a Real-time Human Computer Interaction

Application under Stereo Vision-based Environment," Master Thesis, National
Chiao Tung University, 2006.

