

行政院國家科學委員會專題研究計畫 成果報告

個人助理機器人關鍵視覺功能之建立與應用(I)

計畫類別：個別型計畫

計畫編號：NSC93-2218-E-009-030-

執行期間：93年08月01日至94年07月31日

執行單位：國立交通大學資訊科學學系(所)

計畫主持人：莊仁輝

共同主持人：連震杰

計畫參與人員：高肇宏，張柏仁，鄭兆良

報告類型：精簡報告

報告附件：出席國際會議研究心得報告及發表論文

處理方式：本計畫可公開查詢

中 華 民 國 94 年 10 月 27 日

中文摘要

本計劃之目的在於利用電腦視覺技術，提供「個人助理機器人」智慧中之視覺功能，為先前的「先進家用機器人之研製」三年計畫所發展的電腦視覺技術之延伸。在先前的計畫中，我們完成了機器人視覺系統關鍵技術 - 「藉由二維影像重建三維資訊」，包括：1. 影像特徵點之擷取、對應與追蹤；2. 三維景物之重建。在本計畫中，我們對於更貼近個人之個人助理機器人系統，進行更深入務實之研發；同時，也探討、開拓其應用方向，如個人學習與娛樂。我們選擇一個基於此電腦視覺技術的智慧型指向系統，來發展一個實務應用。目前有關於指向系統的研究，大部分都是分析手勢的影像，來計算所指的方向，但這類的方法精確度不高。本計畫首先進行場景特徵點的三維重建，保留這個步驟所計算出來的基本矩陣。而後續的計算，則直接用所保留的基本矩陣去重建指向物體。本方法同時具有計算時間短且有較高的精確度的優點。實驗方面，我們分別探討指向物體不同、基本矩陣不同、解析度不同，以及指向方向不同的情況下，所重建出來的三維場景及指向位置，並列出幾點可能影響指向精確度的因素，以供日後改進做參考。

英文摘要

The goal of this project is to provide a personal assistant robotic system with visual functions by using computer vision techniques. This project is an extension of the previous 3-year project "Research and Development of Computer Vision System of Home Robot" wherein we developed several key techniques for the main functions of the robot vision system including: 1. the extraction, matching, and tracking of image feature points; and 2. the reconstruction of 3D scenes. In the proposed project, we investigate key functions for a robot working more closely to human. In particular, we conduct a vision-based intelligent pointer system as a practical application.

Most of recent studies on pointer systems calculate the pointing direction and position by analyzing the images of hand gestures, which is not very accurate. In this research, we carry out the 3D reconstruction of the feature points in the scene and calculate the fundamental matrix. The obtained fundamental matrix is kept and then used to perform the reconstruction of the pointer. The advantages of this method include efficient computation and accurate reconstruction results. In the experiments, the accuracy of our 3D scene reconstruction and the obtained target positions using different pointers, fundamental matrices, image resolutions, and pointing directions are analyzed. Several future research directions that can improve the developed pointer system are also listed in this report.

一、前言

本計劃利用先前「先進家用機器人系統研發」三年計畫所發展的個人助理機器人系統視覺模組，發展實際的延伸應用系統。其目的在於完整化個人助理機器人的視覺功能，使其能透過重建所處環境的三維場景，進行各種智慧型應用。

在先前的「先進家用機器人系統研發」計畫中，我們對於機器人視覺系統之最主要功能 - 藉由二維影像重建三維資訊，完成了關鍵技術之發展，包括：(1)影像特徵點之擷取、對應與追蹤，(2)三維景物之重建。在本計畫中，我們對於更為貼近個人系統之視覺應用，進行更深入，務實之研發。我們選擇一個基於電腦視覺技術的智慧型指向系統，做為一個實務專案。

本計劃由投影幾何出發，逐步探究各項電腦視覺關鍵技術，包含基本矩陣求解、相機校正、三維重建步驟、指向系統之設置，最後探討指向系統之精確度分析。我們以各種不同的條件，來分析我們的指向系統，並提出了一些可能的改進方式。下一節中，我們詳細說明我們的研究目的。緊接著是相關研究的介紹。指向系統的研究方法與發展細節，將在第四節中說明，最後是我們計畫執行的成果與討論。

二、研究目的

如前所述，我們選擇一個基於電腦視覺技術的智慧型指向系統，做為一個應用個人助理機器人視覺系統的實務專案。在現有的指向系統中，有許多不同的實作方式，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。本計劃利用三維重建，取得指向物體的三維座標，進而計算出所指向的目標點。我們在判斷指向方向時，只需重建指向物體的兩個特徵點，不必再重建整個場景，也不需要再算一次基本矩陣。如此，可達到加速運算，簡化重建所需的步驟的目的。

另外，我們也依據各種不同的假設與條件進行了完整的實驗，包括基本矩陣不同(參考點不同)，相機解析度不同，指向方向不同以及指向物體不同等因素。藉由如此的指向系統效能分析，我們得以推論並規劃對應的改進方法，以為後續進階深入研究的參考。

三、文獻探討

指向系統所應用的技術，是屬於多視角立體重建的範疇。這種重建的好處是利用多台相機同時拍攝影像，在一個特定的時刻就可以取得足夠的資訊。Narayanan 等人 [1] 利用了一個由 51 台同步且完全校正 (fully calibrated) 的相機所結合而成的半球形結構。然而使用高達 51 台相機，在應用上的彈性將會受到限制，而相機之間的同步技術也是高難度的挑戰 [2]。

除了多視角立體重建外，也有利用單相機，進行多視角重建的系統。Pollefeys 等人 [3] 便實作了這樣的一個系統，其由一連串同一個場景的影像，經過特徵點擷取步驟，進行相似度比較找出其對應關係。透過了 RANSAC (RANdom Sampling Consensus) 的方式，算出許多組可能的代表每兩張影像之中所選各點的轉換關係的基本矩陣，再使用數值最穩定的基本矩陣，進行投影重建。

在現有的指向系統當中，絕大多數都是利用手勢來判斷指向的方向，而用來判斷指向的兩個點，通常都是使用指尖與指根，也有使用指尖與手腕或是使用指尖與眼睛的做法。這些方法中，影響指向精確度最大的因素就是如何準確取得兩個決定指向的特徵點。Shin 等人 [4] 所提出的系統，即是利用指尖與指根來判斷指向的方向。他們利用一個模板去近

似手指位置，找到指尖與指根之後就可以決定指向的方向，進而利用三維重建去算出指向的位置。這個方法中，只用了九個模板方向來尋找指尖，但是手指能指出的方向卻是無限多個，所以這個方法只能簡略地近似我們所要指向的方向。

而在 Hung 等人 [5] 所提出的系統中，是嘗試使用指尖與眼睛的連線，當成指向的向量，再計算此向量與已知的目標平面的交點，求得指向點。此方法在使用上必須持續追蹤人的眼睛，因此相機視角範圍必須有所限制方可得到穩定的計算結果。

除了以手指做為指向媒介，我們也可以使用輔助的工具，當作指向設備。在一般的會議中，最常使用雷射筆來當作我們的指向工具，用雷射筆所打出的光點資訊，如果可以經由視覺模組來追蹤辨認，並配合場景重建，即可以做許多豐富的應用。例如，我們可經由雷射筆做一些輸入指令的功能，使雷射筆成為使用者與個人助理機器人系統之間溝通的媒介。Rahul 等人 [6] 所提出的相關研究主題，即為一個雷射筆簡報系統。圖 1 為雷射筆簡報系統的示意圖。

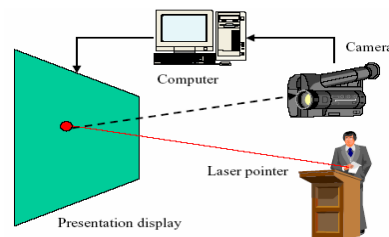


圖1 雷射筆簡報系統示意圖。

本計劃使用一般的雷射筆，來當作我們的指向設備。不過，雷射光點在本計劃中僅用於驗證我們的重建結果，並不用於與指向重建相關的計算。下一節將詳細介紹我們的研究方法。

四、研究方法

本計劃的研究，是由投影幾何出發，逐步探究各項電腦視覺關鍵技術，最後實作一基於電腦視覺的指向系統，及其穩定性的分析。研究步驟包含探討一般化相機模型的投影幾何、基本矩陣的求解、由相機校正結果來求得其投影矩陣、包含指向系統的歐式三維重建。至於指向系統之精確度分析，則將在下一節中透過實驗來做探討。

1. 相機模型與投影幾何

首先，我們透過數學算式的推導，介紹本系統所使用的一般化相機模型[7]。考慮一個針孔相機的模型如圖 2(a)，圖中 C 是相機光學中心、 P 是 principal point、 f 是焦距 (focal length)。空間中的點 $X = (X, Y, Z)^T$ 在影像平面上的對應點，即 X 和相機中心 C 所形成的直線和影像平面的交點 x 。用相似三角的幾何關係，參考圖 2(b)，我們可以知道 $x = (fX/Z, fY/Z, f)^T$ 。

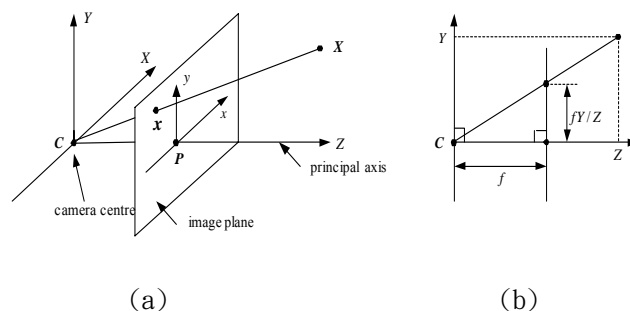


圖2 針孔相機幾何 (a)針孔相機模型 (b)影像點的座標。

假設二維的點以 $m=[u, v]^T$ 來表示，而一個三維的點以 $M=[X, Y, Z]^T$ 來表示。我們令他們的齊次座標表示法分別為 $\tilde{m}=[u, v, 1]^T$ 和 $\tilde{M}=[X, Y, Z, 1]^T$ ，那麼，一個三維的點 M 和該點投影在影像上的二維點座標 m 就有以下的關係存在：

$$s\tilde{m} = A[R | t]\tilde{M} \quad (1)$$

其中 s 是一個比例常數， $[R | t]$ 是由世界座標系統轉換到相機座標系統的旋轉矩陣和平移矩陣， A 則是相機的內部參數矩陣。

另外，在相同三維場景中，所取得兩張不同視角的影像之間，存在一個很重要的限制關係，稱為極線限制 (epipolar constraint)，如圖 3 所示。假設 M 是三維空間中的一個點，而它在第一張影像所成像的影像點是 m ，在第二張影像的影像點是 m' ，則 M 和影像點 m 、 m' 以及相機的光學中心 C 和 C' 是共平面的，稱為極線平面 (epipolar plane) π 。另外， C 和 C' 的連線分別與第一和第二個影像平面相交於 e 和 e' ，我們稱這兩個點為極點 (epipole)。由上述各點之幾何關係，我們可以發現一個重要特性，第一張影像上的任何一點和其光學中心所形成的射線，投影在第二個影像平面上，將成像為一條線 (極點除外)，即所謂的極線 (epipolar line)，如圖 3 中的 l' 。同時，所有的極線都會通過極點。極線限制可以將影像對應點的搜尋範圍，從整張影像減少到極線附近。而極線皆會交於極點這個性質，也可以用來當作一個誤差評量的標準。

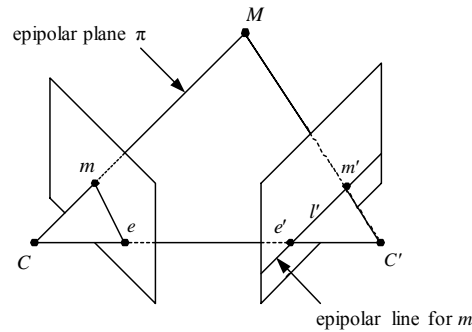


圖3 極線幾何。

2. 基本矩陣求解

接著，我們說明進行三維重建所需的重要計算，亦即基本矩陣的求解。基本矩陣是用來描述在極線幾何中，任一點與其極線之間的對應關係。由於在第一張影像中的任一點 m ，在第二張影像會有一條對應的極線 l' ，二者可以一個齊次座標中的線性關係來表示：

$$l' = F m. \quad (2)$$

因為 m 在第二張影像上的對應點 m' 會位在 l' 這條線上，所以 m' 的座標和 l' 的向量內積為零：

$$m'^T l' = m'^T F m = 0. \quad (3)$$

同理，

$$m^T F^T m' = 0. \quad (4)$$

根據式(3)及(4)，每一組對應點可以產生一個線性方程式，由於基本矩陣 F 扣除了一個縮放的比例常數之後，會剩下八個未知數，因此只需要八組對應點，即可以線性方法解出一個允許改變大小比例的基本矩陣，這個方法稱為八點演算法 (8-point algorithm) [8]。然而，考慮到數位相機在成像時可能的取樣誤差，再加上影像本身會有不穩定的雜訊，甚至在特徵點的擷取與對應上發生錯誤，只利用八個點來計算基本矩陣，並不保證能得到正確的結果。因此，有很多方法被提出來改進基本矩陣的計算。比如使用超過八組對應點計算出多組基本矩陣，再透過最小平方 (least-square) 法來求出最佳的解，以增加基本矩陣的可靠度。有關基本矩陣的最佳化方法比較，可以參考[9][10][11]。

3. 投影矩陣計算

利用前述之方法得到基本矩陣之後，第二個步驟我們要算出兩台相機各自的投影矩陣。由已知的兩台已校正相機的內部參數矩陣 A 和 A' ，加上前述算出的基本矩陣 F ，透過 $E = A'^T F A$ ，我們便可以算出必要矩陣 E 。必要矩陣有以下的兩個性質：

- i. 3×3 矩陣 $E = [t]_{\times} R$ 為一必要矩陣若且唯若 E 的兩個奇異值相等且第三個為零[12]。因為必要矩陣 E 可分解為 $[t]_{\times} R$ ，其中 $[t]_{\times}$ 為斜對稱矩陣，考慮

$$W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{且} \quad Z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5)$$

則 3×3 斜對稱矩陣 $[t]_{\times}$ 可分解為 $[t]_{\times} = kUZU^T$ ， U 為正交矩陣，而 $Z = \text{diag}(1,1,0)W$ ，此即具有兩個相等奇異值之 E 的奇異值分解 (SVD)。

- ii. 假設 E 的 SVD 為 $E = U \text{diag}(1,1,0) V^T$ ，則使用上述的表示法， $E = [t]_{\times} R$ 有以下兩種可能的分解方式：

$$[t]_{\times} = UZU^T, \quad R = UWV^T \quad \text{或} \quad R = UW^T V^T \quad (6)$$

上式中的 R 即為相機的旋轉矩陣，而平移向量 t 則包含在 $[t]_{\times}$ 裡。由於 $[t]_{\times}$ 的富氏長度為 2，表示若 $[t]_{\times}$ 包含比例，則 $\|t\| = 1$ 。因為 $[t]_{\times} t = 0$ ，可知

$$t = U(0,0,1)^T = u_3。$$

由以上兩個性質，我們可以找出可能的 R 和 t 。若第一部相機的投影矩陣為 $P = [I | 0]$ ，則第二部相機的投影矩陣 P' 會有四種可能的選擇，分別為：

$$P' = [UWV^T | u_3] \quad \text{或} \quad [UWV^T | -u_3] \quad \text{或} \quad [UW^T V^T | u_3] \quad \text{或} \quad [UW^T V^T | -u_3]。 \quad (7)$$

4. 歐式三維重建

在取得兩部相機的投影矩陣 P 和 P' 之後，接下來便可進行點的歐式三維重建。若我們將影像點改寫成齊次座標系，並將投影矩陣以列向量表示，可得到：

$$\begin{bmatrix} up_3^T - wp_1^T \\ vp_3^T - wp_2^T \end{bmatrix}_{2 \times 4} M_{4 \times 1} = 0 \quad (8)$$

令 $B_i = \begin{bmatrix} up_3^T - wp_1^T \\ vp_3^T - wp_2^T \end{bmatrix}_{2 \times 4}$ ，則對於每一台相機，皆有 $B_i M = 0$ 的關係存在。每一個三維點

M ，皆有一組影像對應點，就可以獲得兩個如式(8)的關係式，將兩式合而為一，便為一滿足 $BX=0$ 的 4×4 矩陣 B 。欲求得三維點座標 M ，理論上可利用特徵值分解之對應特徵值為零的特徵向量，可是當有雜訊時，則需由對應於最小特徵值的特徵向量加以求得，求得的方法通常是利用矩陣的 SVD 分解，來得到三維點 M 。經過了這些步驟，便可重建出特徵點的三維座標，到了這個階段，我們得到的是一個與歐式重建具有縮放比例關係的公制重建 (metric reconstruction)。接著，利用場景中任何一個已知的尺寸，便可以將這個比例固定下來，得到真正的歐式重建，也就是與實物等尺寸的三維重建。

另一方面，在歐式重建的過程裡，我們可以從基本矩陣中直接抽取出兩台相機間的旋轉矩陣及平移向量，因此我們可以很直接地估算出每台相機相對於重建場景的位置。實際上在計算的時候，由(7)式可觀察到，由於使用了 SVD 分解法，因此所分解出來的旋轉矩陣和平移向量，對於正負符號的決定會有著不確定性，導致相機位置無法求出，也無法檢查場景是否在兩台相機前方。為了解決這個問題，我們可以利用一些額外的限制，來找出正確的正負號。對於旋轉矩陣，我們可以利用旋轉矩陣一個重要的特性，也就是行列式值 (determinant) 等於 1，來決定出它的正負號。而平移向量則可從已知的相機位移方向決定，比如事先知道相機是向右移動，則可知 x 方向的平移為正。確定了旋轉矩陣及平移向量的正負號之後，我們也就可以檢查四種可能的三維重建結果，哪一個是在兩台相機的前方，來決定第二部相機的位置。

對於所重建出的三維點，若它們的座標系不同，我們也可以利用這些已知的相機座標系關係，將重建點轉換為以同一個座標系為基礎。如圖 4， M_1 代表重建物體於 C_1 座標系下的座標向量，若我們要計算同一個物體相對於 C_2 座標系下的座標向量 M_2 ，相當於算出 M_1 經過與相機座標系相反的旋轉及平移之後，所得到相對於 C_1 座標系的點座標，也就是圖中的 M_1' 。而由 M_1 到 M_1' 的這個過程，剛好是和相機座標系的旋轉和平移相反，如圖 4 所示， M_1 先向左平移了 $-t_{12}$ 之後，再順時針旋轉 R_{12}^T ，得到 M_1' 。因此，我們可以將這樣的轉換寫成以下的關係式：

$$M_2 = M_1' = R_{12}^T (M_1 - t_{12}). \quad (9)$$

其中， M_1 為原物體相對於 C_1 座標系下的座標向量， M_2 為原物體相對於 C_2 座標系下的座標向量，而 M_1' 則為原物體經過與相機座標系相反的旋轉及平移之後，所得到相對於 C_1 座標系下的座標向量。反之，若我們已知 M_2 而欲求 M_1 ，則可透過相反的旋轉及平移過程，關係式如下：

$$M_1 = R_{12} M_2 + t_{12} \quad (10)$$

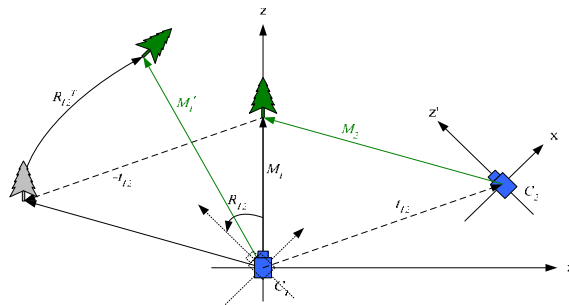


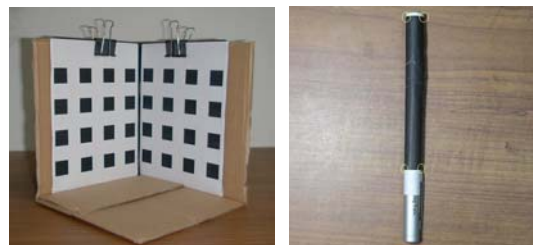
圖4 相機座標與點座標的旋轉及平移關係。

本計劃延伸之前利用上述理論基礎所發展的視覺系統 [13]，除了增進其穩定性及效率，並在不同的場景條件下，進行三維重建，並分析指向系統之精確度。所實作的智慧型指向系統，一開始先將場景的三維模型重建好，並且保留計算好的基本矩陣，之後當我們在判斷指向方向時，只需重建指向物體的兩個特徵點，不必再重建整個場景，也不需再算一次基本矩陣，可以加速運算，簡化重建所需的步驟。實驗流程共分為相機校正、特徵點的選取、特徵點對應、三維重建、計算指向物體所指的目標點定位等五個步驟。實驗結果於下一節說明。

五、結果與討論

我們以數位相機來模擬個人助理機器人的視覺取像設備。所用的數位相機，右影像為 Sony P100 左影像為 Sony P10。圖 5 (a)為實驗場景所使用的由兩個夾角約為 90 度的平面所構成的正交方塊圖板，每個平面上有 16 個方塊。我們針對這些方塊圖案平面進行三維重建，並且藉由指向物體來判斷所指向的方塊。

在實驗結果的分析方面，我們分別探討基本矩陣不同，解析度不同，指向方向不同以及指向物體不同的情況下，所重建出來的三維場景及指向位置。分別說明如下：



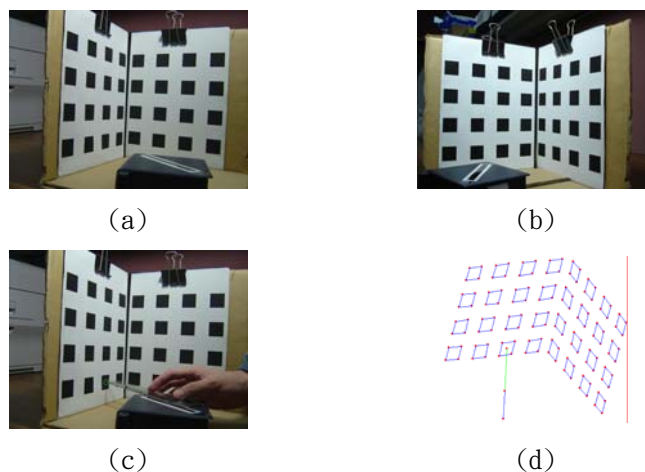
(a)

(b)

圖5 實驗場景。(a) 正交方塊圖板 (b) 指向設備所使用之雷射筆。

1. 使用不同的指向物體進行三維重建

一個理想的指向物體，就是具有可以很準確被我們取得的二維指向特徵點，且僅需極少量的計算。圖 5(b)中的指向雷射筆為本次實驗所採用的指向物體，其優點為不管在任何的角度都能準確的取得四個角點 (如圖中黃色圓圈所標示)，上方兩個角點的中點與下方兩個角點的中點之延伸即為指向方向。



(a)

(b)

(c)

(d)

圖6 使用指向長方條做指向的實驗結果 (a)重建場景的左影像 (b)重建場景的右影像 (c)利用尺的延伸去標示指向的目標點 (d)重建之後的目標點。

圖 6 為使用指向長方條做指向的實驗結果，由結果發現指向精確度很高，與我們所預期的位置非常接近，圖 7 為使用指向雷射筆做指向的實驗結果，由結果發現指向精確度也很高，與我們所預期的位置也非常接近。由於指向雷射筆可以射出雷射光打在立體校正板上，方便我們使用肉眼判斷指向效果的準確性，因此，在接下來的實驗中，皆使用指向雷射筆當作我們的指向物體。

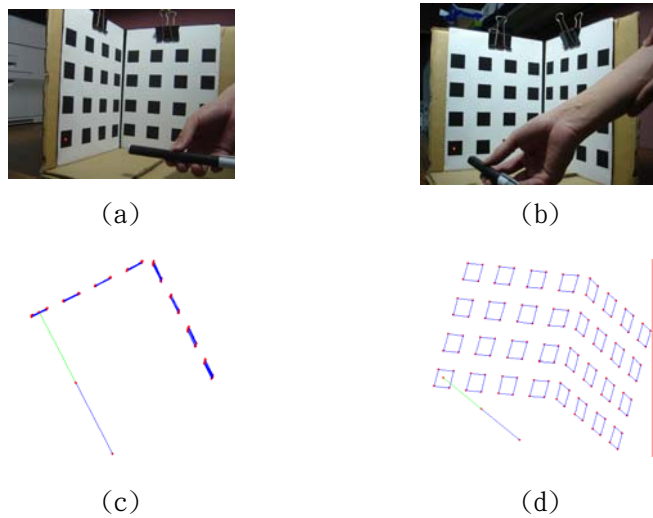


圖7 使用指向雷射筆做指向的實驗結果 (a)重建場景的左影像 (b)重建場景的右影像 (c)重建場景的俯視影像 (d)重建之後的目標點。

2. 使用不同的基本矩陣進行三維重建

在三維重建中，基本矩陣的正確性，對於重建的結果有很大的影響，我們把正交方塊圖板上 128 個特徵點加以編號，再擷取不同的特徵點計算基本矩陣，被選到之特徵點將於圖中以圓圈做記號。由實驗結果，我們可以探討不同特徵點取法所算出的基本矩陣，對於三維重建以及指向結果的影響。圖 8 為相同的重建場景但取不同的八個特徵點去算基本矩陣所重建出來的結果。

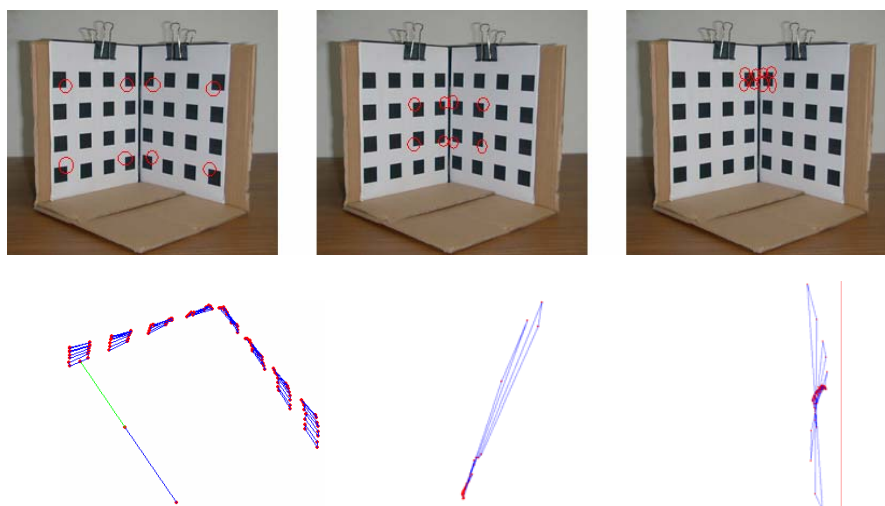


圖8 取不同的八個特徵點去算基本矩陣所重建出來的結果。(上方的圖顯示所選取之特徵點，下方的圖為其重建結果)

比較取編號 1 至 128 的特徵點，與取編號 1 至 128 的特徵點再加指向設備上的兩個特徵點，所算出來的基本矩陣去做三維重建的結果，如圖 9 所示，我們發現用 128 個特徵點

與 130 個特徵點所算出來的基本矩陣重建效果都很好，目標點的計算也十分的接近，而這兩個方式所算出來的基本矩陣數值相差也不會太大。兩個基本矩陣對應元素的實際差值為

$$|F_1 - F_2| = \begin{bmatrix} 0.00000014833681 & 0.00000012489978 & 0.00015852627085 \\ 0.00000025063960 & 0.00000000842801 & 0.00014887054605 \\ 0.00009814191172 & 0.00017814062252 & 0.00000237283885 \end{bmatrix}$$

由本階段實驗結果可以得知，當我們在取特徵點時，只要取到足夠多以及對應穩定的特徵點，即可得到很好的重建結果。

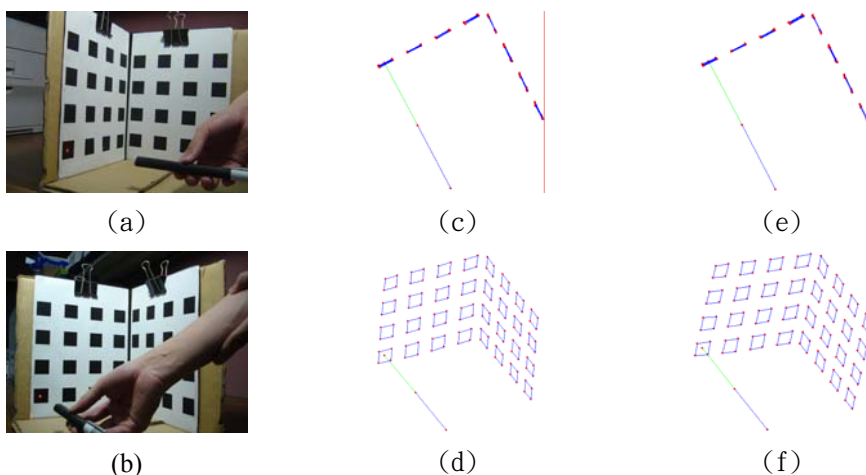


圖9 (a)重建場景的左影像 (b)重建場景的右影像 (c)取128個特徵點所算出的基本矩陣之重建結果俯視圖 (d)取128個特徵點所算出的基本矩陣之目標點重建結果 (e)取130個特徵點所算出的基本矩陣之重建結果俯視圖 (f)取130個特徵點所算出的基本矩陣之目標點重建結果。

3. 使用不同的解析度進行三維重建

我們以 640 x 480 和 1280 x 960 兩種不同的解析度去做三維重建，並分析其指向精確度。同時，我們也測試特徵點擷取準確性對於重建結果的影響。我們於隔板加上較大的特徵點紙板，如圖 10 所示，並稱這四個方塊的特徵點為輔助點。我們利用這些輔助點計算基本矩陣，再把特徵點紙板移除，然後以這個基本矩陣去重建我們所要重建的物體。

圖 11 為 640 x 480 影像之三維重建之結果，由圖中我們可以發現，場景重建結果還不錯，只是方塊的形狀有些許的變形，而關於目標點位置的計算，大約向左偏移了半個方塊邊長的距離。而圖 12 也是使用上述的方法但以 1280 x 960 的影像做三維重建，由圖中我們可以發現，場景重建的結果很好，而目標點位置的計算也只向上偏移了一點。



圖10 加入場景的特徵點紙板。其上的特徵角點稱為輔助點。

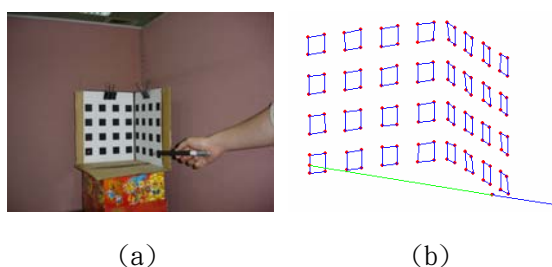
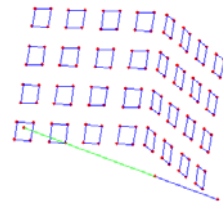


圖11 使用640 x 480的影像做三維重建 (a)場景的右影像 (b)重建的目標點。



(a)



(b)

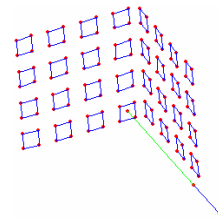
圖12 使用1280 x 960的影像做三維重建 (a)場景的右影像 (b)重建的目標點。

4. 不同的指向方向進行三維重建

在本階段的實驗中，我們測試水平與垂直方向上的移動。我們以圖 12 為基準，若將指向物體向右移動三方塊大小，則結果如圖 13，若將指向物體向上移動三方塊大小，則結果如圖 14。



(a)



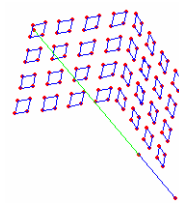
(b)

圖13 指向方向往右移動：(a)重建場景的右影像 (b)重建之後的目標點。

實驗結果發現，本計劃所發展的指向系統，不管是在水平或者是垂直方向上的移動，其所指的方向都會相當正確的隨著指向物體的移動而移動。表 1 為使用輔助點去算基本矩陣與未用輔助點去算基本矩陣（即以立體方塊圖版上 128 個點去算基本矩陣）的誤差比較表，由表中我們可以發現使用輔助點可以使我們的誤差減少三分之二，達成改善指向精確度的目標。



(a)



(b)

圖14 指向方向往上移動：(a)重建場景的右影像 (b)重建之後的目標點。

表1 使用輔助點與未使用輔助點的誤差比較表(單位:公分)

	圖 12	圖 13	圖 14	平均誤差
使用輔助點算基本矩陣，目標點的誤差	0.7	0.2	0.9	0.6
未使用輔助點算基本矩陣，目標點的誤差	1.4	3	1	1.8

在現有的指向系統中，大部分都是利用手勢來判斷指向方向，需要經過模板的比對，才能計算手指頭所指的方向。本計劃以發展於個人助理機器人的電腦視覺系統，作一實際的應用。透過指向系統，來驗證及改進我們的系統穩定性及效率。我們所提出的方法，減少了基本矩陣的計算量，而在三維場景重建的部份，我們也提出數種方法，用以改善三維場景重建的效果。

在本論文的實驗與分析裡，為了避免被特徵點擷取產生的誤差所影響，因此在特徵點的選擇與對應，是以人工介入的方式來達成，未來可進一步與前端的特徵點自動擷取系統結合，並實作出即時的系統，分析重建與指向的結果，並從影像品質、特徵點的選取與對應、視角與視角差之決定等方面設法改善。此外，在本計劃後續的研究方向上，我們亦探討利用其他指向物體特徵，如邊緣，來建立指向及決定目標物的可行性及其實用性，以滿足實際應用的進一步要求。

參考文獻

- [1] P. Narayanan, P. Rander, and T. Kanade, "Constructing Virtual Worlds Using Dense Stereo," *Proc. International Conference on Computer Vision*, 1998.
- [2] M. Okutomi and T. Kanade, "A Multiple-baseline Stereo," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 4, pp. 353-363, April 1993.
- [3] M. Pollefeys and L. Van Gool, "From Images to 3D Models," *Communications of the ACM*, vol. 45, no. 7, pp.50-55, July 2002.
- [4] S. Sato and S. Sakane, "A Human-Robot Interface Using an Interactive Hand Pointer that Projects a Mark in the Real Work Space," *IEEE International Conference on Robotics & Automation*, April 2000.
- [5] 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," *Proc. 14th International Conference on Pattern Recognition*, pp. 1244-1246, Brisbane, August 1998.
- [6] R. Sukthankar, R. G. Stockton and M. D. Mullin, "Self-Calibrating Camera-Assisted Presentation Interface," *Proc. International Conference on Control, Automation, Robotics and Vision*, 2000.
- [7] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2001.
- [8] H.C. Longuet-Higgins, "A Computer Algorithm for Reconstructing A Scene from Two Projections," *Nature*, vol. 293, pp. 133-135, 1981.
- [9] K.-L. Lee, "Optimization of Fundamental Matrix and Euclidean Reconstruction," *Master Thesis*, National Chiao Tung University, 2002.
- [10] R. Hartley, "In Defense of the Eight-point Algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580-593, June 1997.
- [11] Z. Zhang, "Determining the Epipolar Geometry and Its Uncertainty: A Review," *Research Report, No. 2927, INRIA Sophia-Antipolis*, July 1996.

- [12] T. Huang and O. Faugeras, "Some Properties of the E Matrix in Two-View Motion Estimation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 11, no. 12, pp. 1310-1312, Dec. 1989.
- [13] L.-C. Guo, "Robot Vision for 3D Scene Reconstruction," *Master Thesis*, National Chiao Tung University, 2003.

六、計畫成果自評

1. 研究內容與原計畫相符程度

本計畫的研究內容，包含了：1. 影像特徵點之擷取、對應與追蹤；2. 三維景物之重建，也對於更貼近個人之個人助理機器人系統，進行更深入務實之研發，並探討、開拓其應用方向，如個人學習與娛樂。因此，符合原計畫書所列之各研究主題。

2. 達成預期目標情況

本計畫做了豐富的實驗，分別探討指向物體不同、基本矩陣不同、解析度不同，以及指向方向不同的情況下，所重建出來的三維場景及指向位置，並列出幾點可能影響指向精確度的因素，以供日後改進做參考。計畫書所列的預期目標，皆有相當程度的達成，並能夠作為進一步研究的參考。

3. 研究成果之學術或應用價值

本計畫的成果，已用於我們更先進的指向系統的先期研究之重要參考。對於該系統的研發，提供了紮實的理論基礎，及豐富的實作經驗。因此，本計畫研究成果，有良好的學術和應用價值。

4. 是否適合在學術期刊發表或申請專利

本計畫的研究內容，為國際學術領域中，重要的研究方向之一。我們的成果將整理發表於國際會議及期刊論文。

5. 主要發現或其他有關價值

我們的研究成果，可提供機器人視覺系統一個感知環境的重要應用。