

臺南市政府 106 年度自行研究報告書

利用電腦視覺及攝影測量建立三維點坐標
以協助指界

臺南市永康地政事務所

劉琇琳 著

中華民國 106 年 10 月 1 日

摘要

地政事務所大多圖資皆屬平面資料，平面圖資雖易保管且保存容量小，但卻無法直觀的提供立體視覺。當今空間資訊發展領域中呈現出許多建物、土地、環境等等的三維模型，且應用的領域相當廣泛且受歡迎，尤其是對於一般民眾，模型比平面資料更淺顯易懂，因此這樣的三維圖資儼然成為空間資訊的主流，重建物體三維模型可能是對於事務所的創新議題。在地政事務所的土地糾紛案件中，我們可以將建物及土地的模型重建展現給民眾以協助指界，除了免去現場查看的人力和時間耗費，也可以用較具有可靠性的尺寸來取得民眾的信任。

傳統測量可以應用共面式以及光速法平差等演算法求出三維點雲，但缺點是耗費時間長且初始值估計不易。然而現今的電腦視覺領域將新穎的觀念以及演算法引入，提供本質矩陣(Essential Matrix)、基礎矩陣(Fundamental Matrix)、影像對極幾何(Epipolar Geometry)、特徵匹配(Feature Based Matching)、基於運動重建(Structure From Motion)等觀念來快速重建相機相對方位以及三維點雲座標。因此本研究選取相機和欲建模建物，對房屋拍攝一個立體像對，希望透過電腦視覺中的上述演算法來重建該永康區建物的模型，並分析演算法的可行性以及三維點雲的精準度。

關鍵詞：電腦視覺、特徵匹配、基於運動重建、基礎矩陣、三維點雲

目錄

摘要.....	I
目錄.....	II
圖目錄.....	V
表目錄.....	VII
第一章 介紹.....	1
第一節、研究動機與目標	1
第二節、文獻回顧	3
第三節、研究流程	4
第四節、章節架構	4
第二章 相機率定.....	6
第三章 特徵式匹配.....	13
第一節、特徵偵測 (Feature detection).....	13
一、積分影像 (Integral image).....	14
二、盒型空間 (Box-space).....	14
三、海森矩陣 (Hessian Matrix).....	16
第二節、特徵描述	17

一、特徵點主方向 (Dominate orientation)	17
二、特徵點描述子 (Descriptor)	18
第三節、特徵匹配	19
一、特徵比對	19
二、匹配策略	20
第四章 相對方位以及物點三維座標	21
第一節、共面式	21
第二節、電腦視覺	23
第三節、物體三維空間資訊	25
第五章 實驗成果與分析	26
第一節、實驗一：近景攝影之小型物體模型重建	26
一、實驗對象與研究基本資訊	26
二、回復相對方位	28
三、物體模型重建	30
第二節、實驗二：近景攝影之大型物體模型重建	32
一、實驗對象與研究基本資訊	32
二、特徵匹配成果	34

三、檢核對極幾何	35
四、回復相對方位	39
五、 物體模型重建	41
六、 使用商用軟體 VisualSFM 進行物體模型重建	44
第六章 結論及建議	47
第一節、結論	47
第二節、建議	48
第七章 參考文獻	49

圖目錄

圖 1. 成功大學光復校區砲台模型重建	2
圖 2. 成功大學相機圓盤率定場	7
圖 3. 相機率定報告書 1	10
圖 4. 透鏡輻射畸變圖	11
圖 5. 相機率定報告書 2	12
圖 6. 積分影像捲積運算	14
圖 7. 一階與二階盒型濾波器	15
圖 8. 區域性非最大值抑制	17
圖 9. 特徵點決定主方向示意圖	18
圖 10. 特徵點描述子	19
圖 11. 最鄰近區域距離比率匹配策略	20
圖 12. 共面條件	22
圖 13. 交會模糊性(共 4 解)	24
圖 14. 具有量測雜訊的前方交會	25
圖 15. 實驗一的立體像對(紅標為共軛點)	27
圖 16. 三維模型重建(具交會模糊性)	31
圖 17. 真實影像與模型重建成果比較圖	31
圖 18. 實驗二立體像對	33

圖 19. 使用 SURF 的匹配成果	34
圖 20. 使用 SURF+RANSAC 的匹配成果	34
圖 21. 檢核影像 2(特徵匹配)的對極幾何	36
圖 22. 檢核影像 1(特徵匹配)的對極幾何	37
圖 23. 檢核影像 2(人工匹配)的對極幾何	38
圖 24. 檢核影像 1(人工匹配)的對極幾何	38
圖 25. 建物模型重建圖角度 1(特徵匹配).....	42
圖 26. 建物模型重建放大圖 1.....	42
圖 27. 建物模型重建放大圖 2.....	43
圖 28. 建物模型重建圖角度 2(特徵匹配).....	43
圖 29. 建物模型重建圖(人工匹配).....	44
圖 30. VisualSFM 軟體介面	45
圖 31. VisualSFM 軟體操作過程.....	46
圖 32. VisualSFM 軟體房屋建模成果	47

表目錄

表 1. 使用相機規格介紹.....	6
表 2. 相機率定參數表	8
表 3. 相機率定參數顯著性測試表	9
表 4. 盒型空間參數表	15
表 5. 基礎矩陣及本質矩陣介紹表	23
表 6. 實驗一立體像對的基礎及本質矩陣	28
表 7. 實驗一求解之相對平移及旋轉矩陣	29
表 8. 實驗一電腦視覺方式求出的相機座標和旋轉角度	29
表 9. 實驗一共面式方式求出的相機座標和旋轉角度	30
表 10. 實驗一兩方式所求出之相對方位的差值.....	30
表 11. 實驗二研究對象立體像對基本資訊	32
表 12. 使用 SURF 和 SURF+RANSAC 的成果比對表.....	35
表 13. 實驗二(特徵匹配)的基礎矩陣及本質矩陣	36
表 14. 實驗二(人工匹配)的基礎矩陣及本質矩陣	39
表 15. 實驗二(特徵匹配)求解之相對平移及旋轉矩陣	39
表 16. 實驗二(特徵匹配)電腦視覺求出的相機座標和旋轉角度	40
表 17. 實驗二(人工匹配)求解之相對平移及旋轉矩陣	40
表 18. 實驗二(人工匹配)電腦視覺求出的相機座標和旋轉角度	41

第一章 介紹

第一節、研究動機與目標

近年來二維的影像資訊已經不足以用來展示現實環境中的情況了，為求身臨其境的感覺，逐漸開發出許多三維空間模型的成果以及相關應用，例如：城市模型圖以及災區模型重建等，城市模型圖讓人可不須真實到達目的地也能看見該地方的環境狀況，而災區模型重建讓救災人員可以不用再一次實際冒險去了解災區的情況，而使用模型來做出合理且安全的災後分析，抑或是最近較夯的虛擬實境遊戲(VR)，透過點雲呈現現實物體，並賦予物體屬性資料，讓使用者除了可以感覺到視覺、重量以及空間的概念等，並真實體現身臨其境四個字。

因此若是能夠利用攝影測量的方式求得三維物空間座標，那我們可以省去許多時間及人力，例如：當我們正進行土地糾紛調解時，為了讓所有人能了解該筆土地現地狀況，我們常以照片或是直接去現地勘查的方式進行，若是今天我們以模型來展現該筆土地的現況，那我們進行調解時不但方便說明，對於現地欲了解之長寬尺寸也能直接透過模型測得。而由於僅是對於有糾紛土地建置三維模型，因此也不耗資源。圖 1. 展現了使用多角度相機拍攝和商用軟體 Photo Scan 而建成的成功大學光復校區的砲台模型，模型的擬真度相當高，除了可以看出外型，甚至還有紋路和顏色，未來也可以賦予其他數性資料，讓

模型的利用價值更高。

雖然利用近景攝影建置三維空間資料的軟體市面上已經存在，但是由於購置軟體需要經費，而在考量不花費金錢以及程式複雜度的條件下，於是本研究對於其中演算法來進行探討，期望有一套簡易且不需花費的程式可以運行。因此，本次研究是希望透過自行拍攝的立體相對來重建物體的點雲三維空間資料。從真實世界的物點三維資訊可以透過簡單的投影方法獲得在投影平面的影像座標，但是相反的，要從平面座標轉換物點的三維座標卻是難以進行，因此需要透過影像間的共軛點搭配前方交會以及尺度等概念，才可以重建物點的三維資訊。在傳統攝影測量中，會透過區域性匹配、共面式求相對方位以及光束法平差來獲得大量的點雲。但是現今成果除了講求真實和精度，尚還追求產製時間，因此本研究主要透過電腦視覺的多個演算法來體現三維空間資訊，並分析其相關條件以及成果差異。



圖 1. 成功大學光復校區砲台模型重建

第二節、文獻回顧

SURF (Bay et al., 2008) 以及 SIFT (Lowe, 2004) 皆是電腦視覺中的特徵匹配，而 SURF 是 SIFT 的改良版，除了加速匹配速度，也增加匹配穩健性，SURF 之所以可以加速偵測，是因為積分影像的概念(Viola, 2001)和盒型濾波器(Bottou, et al., 1999)的應用。在 SURF 偵測的特徵點條件中是根據海森矩陣(Lindeberg, 1998)的行列式值來決定候選的特徵點。最後特徵點相鄰區域的梯度統計來給予 64 個描述子(Bay et al., 2008)，並根據歐幾里得距離來決定最佳的匹配點。

電腦視覺中有學多學者提出概念來描述一個立體像對的關係，其中有以基礎矩陣和本質矩陣(Hartley, 1997)的概念最為重要，內容以八點法來求基礎矩陣，對於基礎矩陣求解相對方位，則是以 SVD 方法(Chien, J, 2016)最為普及，接著使用區域性座標性統可以將相對方位(Li, C. and Y. Zhao, 2012)轉換成絕對方位。最後利用以求得的相機位置和旋轉角度搭配前方交會(Longuet-Higgins, H. C., 1981.)就可以獲得三維點雲。

第三節、研究流程

1. 選擇適合的相機以及欲建模的物體。
2. 進行相機率定工作。
3. 對物體進行近景拍攝，維持 70%以上的覆蓋率。
4. 使用人工或是自動化的特徵點偵測找尋兩張相片的共軛點。
5. 將共軛點對座標丟進電腦視覺演算法中恢復拍攝時的兩相機的相對方位。
6. 建立區域性座標系統(以第一台相機為原點)，並引入尺度概念，確立兩相機的位置以及旋轉角度。
7. 使用兩相機以及影相共軛點的前方交會求出三維物空間座標。

第四節、章節架構

第一章強調研究動機和目的，並提出相關的研究論文，再給予計畫的研究流程。第二章主要探討相機的率定工作、程序以及率定成果，以建立相機內部參數。因為該研究都是從立體像對出發，影像是相當重要的角色，因此我們必須相探討影像的成像誤差，而成像誤差部分來自相機和透鏡的誤差，包含像主點平移以及透鏡畸變參數等。

第三章陳述的是特徵式匹配。找尋共軛點的方式有許多種，本研究採用 SURF 特徵式匹配，本章第一節描述 SURF 是如何選取特徵點，第二節介紹 SURF 對於特徵點的 64 個描述子，第三節講解如何將兩張

立體像對中的特徵點進行成功匹配。

第四章是介紹回復相對方位的方法。在進行模型重建前，必須恢復相機在拍攝瞬間的位置和旋轉角度，所以第四章介紹二種回復相對方位的方法。本章第一節是傳統攝影測量的共面式，雖然精度高但是迭代計算相當耗費時間，因此第二節介紹了電腦視覺方式，電腦視覺中會使用基礎矩陣和本質矩陣來描述兩相機的相對方位。當完成相對方位後，第三節是說明如何使用三角測量(前方交會)來重建物點的三維空間資訊。

第五章設計兩個實驗來重建物體模型。本章第一節第一點介紹第一個實驗對象(小型模型-立體黑白版)以及人工產生的共軛點，第二點說明透過電腦視覺方法恢復相對方位，並使用共面式的方式檢核相對方位成果，第三點使用上節的相對方位成果搭配三角測量重建出小型模型，並檢核模型的正確性。本章第二節第一點介紹第二個實驗對象(大型物體-房子)以及相機相關資訊，第二點呈現特徵式匹配得到的共軛點成果以及人工匹配的共軛點成果。第三點計算基礎和本質矩陣，並檢核其對極幾何以確定本質矩陣的正確性。第四點回復拍照瞬間的相機相對方位，本節會比較分別使用特徵是匹配和人工匹配所產生的成果差異。第五點展現使用三角測量得到的房屋模型。第六章為結論和建議。第七章為參考文獻。

第二章 相機率定

因為所拍攝的相片用於量測作業及重建三維點雲，而任何相機和鏡頭都會因為製作過程或是設計產生大小不一的偏差量，其中包含像主點偏移、透鏡輻射畸變以及透鏡偏心畸變等，因此需對相機進行率定作業。使用相機為「SONY NEX-5N」，下表 1. 為該相機的規格介紹。

表 1. 使用相機規格介紹

外觀	
型號	SONY NEX-5N
感測器大小	23.4 x 15.6 mm
拍攝模式	選擇手動模式
焦距	16mm
光圈孔徑	f/4.5
曝光時間	1/125
ISO 速度	100
感光元素像素	1610 萬像素
最大解像度	4912 × 3264

這次研究使用國立成功大學的圓盤率定場，如下圖 2。率定作業程序如下，在圓盤率定場，從兩個角度拍攝。先站在圓盤右前方拍攝兩張照片(正像和翻轉 90 度)，接著轉動圓盤 45 度再拍攝兩張，以此類推，直到拍攝完 360 度後，換站到圓盤正前方，拍攝整個圓盤，也是正像和翻轉 90 度兩張相片，因此總拍攝相片數應為 $\frac{360}{45} \times 2 + 2 = 18$ 張影像。在進行拍攝時需要注意以下兩點，第一點是盡量讓標佈滿整張影像，第二點是要讓白色的標和背景的黑產生較大的對比。針對這 18 張影像，我們使用 Australis 軟體來進行該相機的率定工作，因為圓盤上的標的物是可以由電腦自動辨認，因此使用自動偵測標的物以及匹配選項，自動進行率定工作。

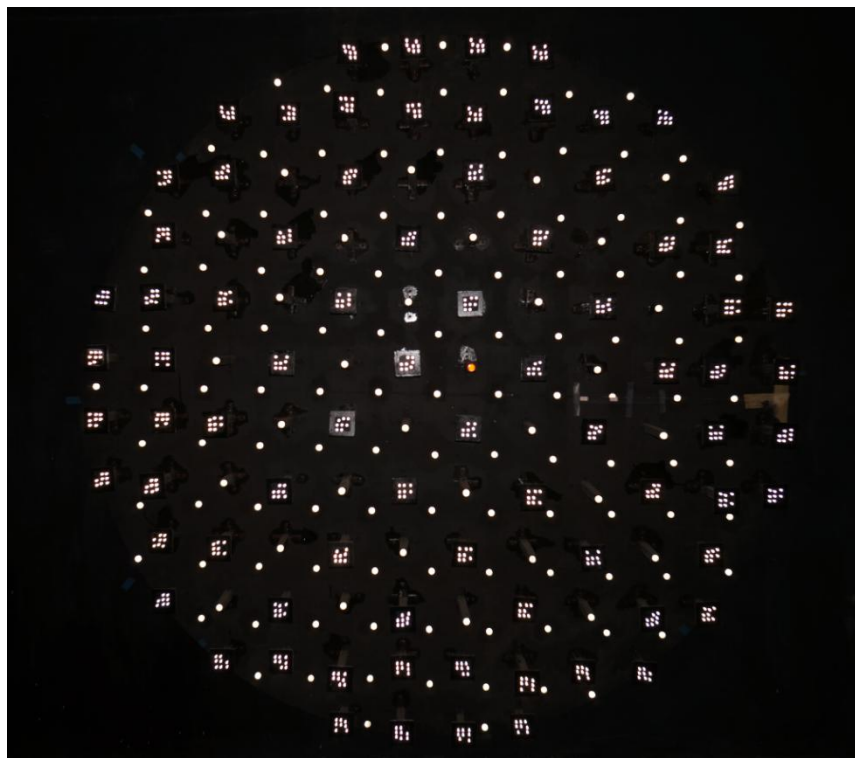


圖 2. 成功大學相機圓盤率定場

經過標準的率定程序後，可以得到一份相機率定報告，如下圖 3。

報告中包含率定後焦距、像主點偏移、透鏡輻射畸變參數以及偏心畸變參數等，將報告數中的參數整理成表 2。但並非所有參數都需要，需透過顯著性測試來決定留下來的率定參數，顯著性表示該參數對率定作業的影響大小，若參數顯著性高留下，相反地，如果參數顯著性低則捨去，顯著性測試如下表 3，經過顯著性測試後，僅留下焦距 c 、像主點偏移 x_p 和 y_p 、輻射畸變參數 $K1K2$ 以及偏心畸變參數。

表 2. 相機率定參數表

參數	數值	標準偏差
c	15.9586mm	< 0.001mm
x_p	-0.1588mm	< 0.001mm
y_p	0.3158mm	< 0.001mm
$K1$	2.57106e-004	7.2637e-007
$K2$	-1.52722e-006	8.6319e-009
$K3$	1.52622e-009	3.2495e-011
$P1$	-4.6321e-005	6.876e-007
$P2$	-1.3778e-004	6.263e-007
$B1$	0.0000e+000	5.094e-010
$B2$	0.0000e+000	5.094e-010

表 3. 相機率定參數顯著性測試表

參數	精度(pixel 大小)	下降程度	顯著性測試
c	2.33		重要參數需保留
xp yp	2.02	下降 0.31	顯著→保留
K1	1.40	下降 0.62	顯著→保留
K2	0.84	下降 0.56	顯著→保留
K3	0.84	下降 0.00	不顯著→設為零並且 fix
P1	0.82	下降 0.02	不顯著→雖不顯著，但加入 p2 後一起率定則為顯著，故保留
P2	0.75	下降 0.07	顯著→保留

CAMERA CALIBRATION REPORT

PROJECT DETAILS

Camera: SONY NEX-5N

Filename: D:\20170707newcalibrationfield\20170714.aus

Calibration Date: 15/07/2017 15:44pm

METRIC CALIBRATION PARAMETERS

Resolution = 4912 x 3264 pixels

Pixel width = 0.0048mm, Pixel height = 0.0048mm

	VALUE	STANDARD ERROR
Principal distance	c = 15.9856mm	< 0.001mm
Principal point offset in x-image coordinate	xp = -0.1588mm	< 0.001mm
Principal point offset in y-image coordinate	yp = 0.3158mm	< 0.001mm
3rd-order term of radial distortion correction	K1 = 2.57106e-004	1.2481e-007
5th-order term of radial distortion correction	K2 = -1.52722e-006	1.6364e-009
7th-order term of radial distortion correction	K3 = 1.52622e-010	6.4427e-012
Coefficient of decentering distortion	P1 = -4.6321e-005	1.522e-007
Coefficient of decentering distortion	P2 = -1.3778e-004	1.762e-007
No significant differential scaling present	B1 = 0.0000e+000	5.286e-020
No significant non-orthogonality present	B2 = 0.0000e+000	5.286e-020
9th-order term of radial distortion correction	K4 = 0.00000e+000	5.2856e-036
11th-order term of radial distortion correction	K5 = 0.00000e+000	5.2856e-040

STANDARD CORRECTION EQUATION

The corrected image coordinates $x(\text{corr})$ & $y(\text{corr})$ can be calculated from the measured coordinates $x(\text{meas})$ & $y(\text{meas})$ by using the formulas:

$$x = x(\text{meas}) - xp$$

$$y = y(\text{meas}) - yp$$

x and y are now with respect to the principal point,

$$r^2 = x^2 + y^2$$

$$dr = K1 \cdot r^3 + K2 \cdot r^5 + K3 \cdot r^7 + K4 \cdot r^9 + K5 \cdot r^{11}$$

$$x(\text{corr}) = x(\text{meas}) - xp + x \cdot dr/r + P1 \cdot (r^2 + 2x^2) + 2 \cdot P2 \cdot x \cdot y$$

$$y(\text{corr}) = y(\text{meas}) - yp + y \cdot dr/r + P2 \cdot (r^2 + 2y^2) + 2 \cdot P1 \cdot x \cdot y$$

Camera self-calibration determined in a network of 241 images and 7607 points, to an image measurement accuracy (RMS 1-sigma) of 0.26 pixels or 1.23 μm , and of 1.0.

Produced by Australis from Photometrix - <http://www.photometrix.com.au>

PAGE 1 of 5

圖 3. 相機率定報告書 1

透鏡畸變是指物體在成像平面成像後的實際影像位置和理想的成像位置有偏移，大致可以分為輻射畸變以及偏心畸變，其中又以透鏡輻射畸變最為普遍，輻射畸變是指像點依像主點呈現輻射狀的偏移，通常以像外為正向輻射畸變，而向內為負向輻射畸變，如圖 4. 展現取樣像點因為透鏡輻射畸變所產生的位移，越向外圍偏移量越大，因此若研究所需量測精度高，則需做相機率定工作，抑或是捨去周圍影像資訊，也可參考圖 5. 的透鏡畸變數值圖。

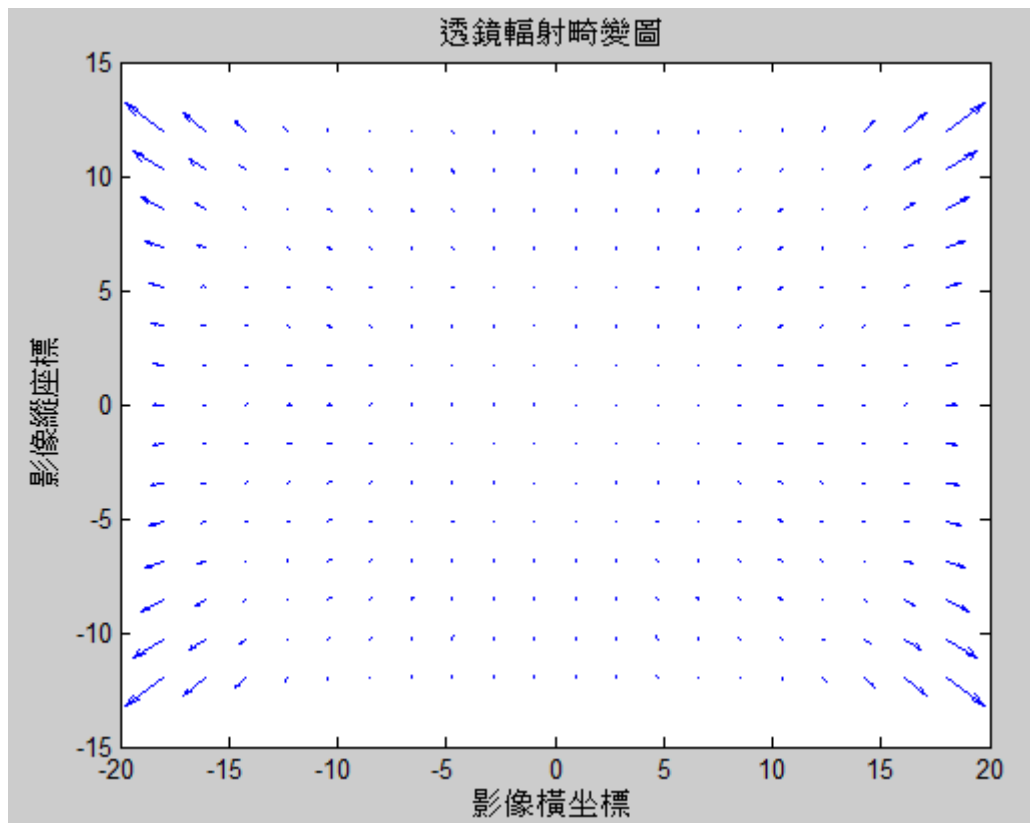


圖 4. 透鏡輻射畸變圖

CAMERA CALIBRATION REPORT

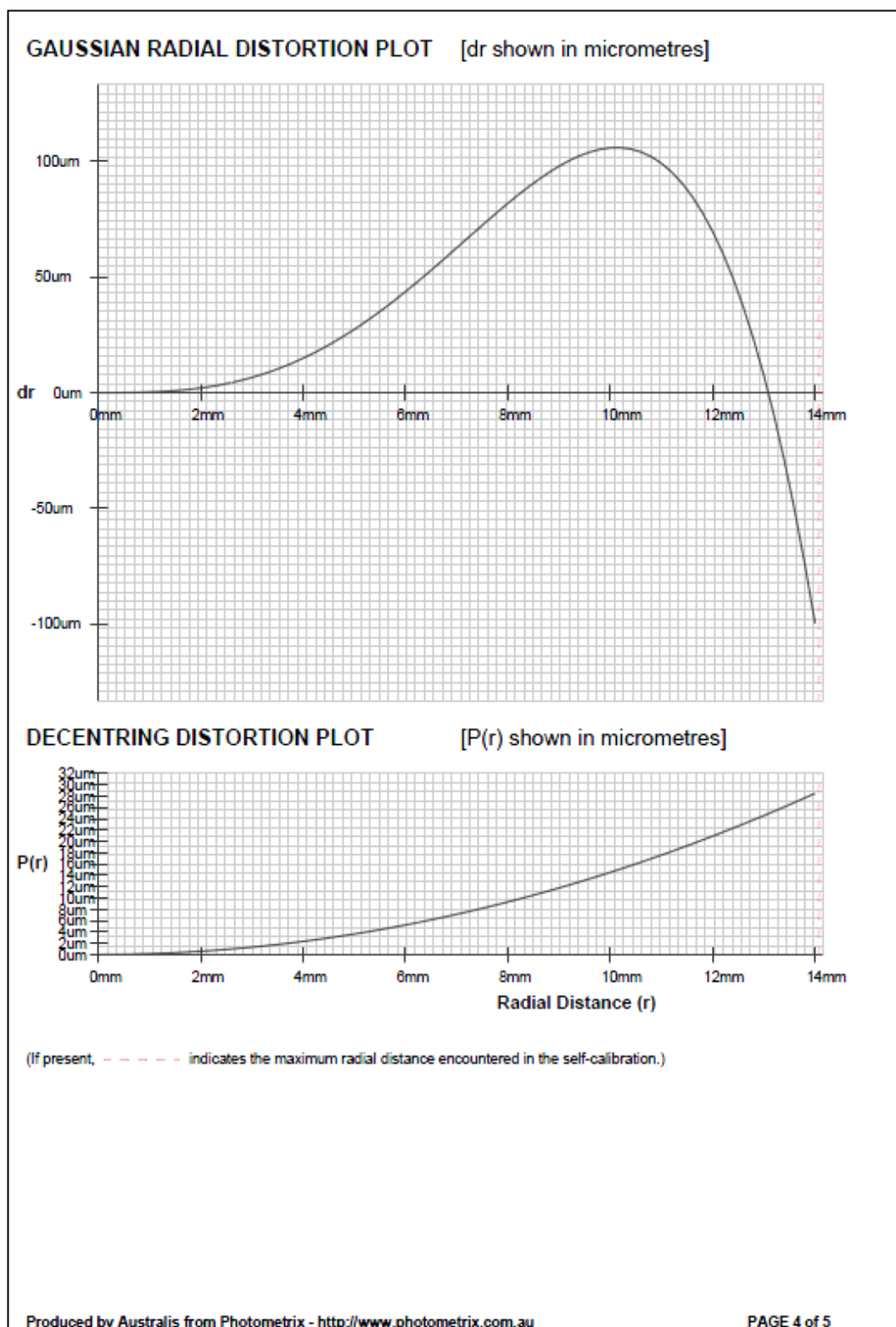


圖 5. 相機率定報告書 2

第三章 特徵式匹配

當完成兩張立體相對後，需要找尋影像中的共軛點。若使用肉眼偵測雖然可以找尋較準確以及欲重建的物點，但是所花費的時間過長，且不易獲得大量共軛點。因此本章介紹 SURF 特徵式匹配，SURF 全名為 Speeded Up Robust Features，可以穩健且迅速獲得兩張立體相對中的共軛點。其主要由三個連續步驟組成，分別是特徵偵測、特徵描述及特徵匹配。第一部分的特徵偵測會敘述積分影像的應用、盒形空間海森矩陣的行列式值計算以及特徵選取的基準，而第二部分的特徵描述則會介紹偵測點的空間尺度、特徵點的主方向決定以及特徵點的 64 個描述子，最後部分的特徵匹配則會著墨在匹配策略以及除錯方法。

第一節、特徵偵測 (Feature detection)

特徵偵測主要建立在「積分影像」以及「盒型空間」兩個新穎的概念上，接著將積分影像與盒型濾波器捲積後可以得到「海森矩陣」的元素值並計算其行列式值，最後透過特徵選取的相關策略，萃取出每個盒型空間尺度下相對應的偵測特徵點。SURF 演算法須達到特徵點的相似不變性，其中相似不變性包含了平移、尺度以及旋轉不變性，若在盒型空間下，特徵點具有尺度和平移不變性。

一、積分影像 (Integral image)

SURF 演算法有個相較於其他特徵匹配的優點是計算效能好及運算時間短，其最主要原因在於 SURF 使用積分影像來進行捲積運算 (convolution)。積分影像的計算方式如下式。透過積分影像的捲機計算，可以將原本的大量運算子以及記憶體所減至 3 個運算元以及 4 個記憶體空間配置，積分影像的計算方式及過程可參考圖 6。

$$U(x,y) := \sum_{0 \leq i \leq x} \sum_{0 \leq j \leq y} u(i,j)$$

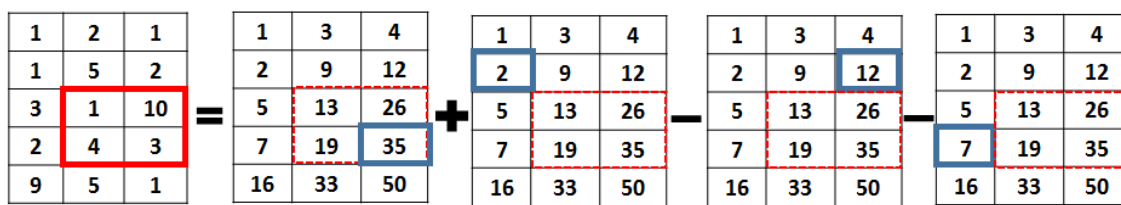


圖 6. 積分影像捲積運算

二、盒型空間 (Box-space)

盒型濾波器是建構在盒型空間下，盒型空間簡單說明就是將高斯空間採近似運算獲得的空間模型。盒型空間由「音階(Octave)」以及「層(Layer)」組成，某一個音階的某一層由一個尺度參數(L)來描述，L 參數決定了一二階的盒型濾波器尺寸、特徵點的所在尺度、對應的高斯空間尺度等。在盒型空間中，可以讓特徵點達到尺度以及平移不變性。下表 4.是針對盒型空間的各參數以及數值加以呈述。

表 4. 盒型空間參數表

o	i	$\sigma(L)$	σ_k	L	$3L$	$l(L)$	$w(L)$
1	1	1.2	1	3	9	2	0.9129
	2	2.0	2	5	15	4	0.9487
	3	2.8	3	7	21	6	0.9636
	4	3.6	4	9	27	7	0.9718
2	1	2.0	2	5	15	4	0.9487
	2	3.6	4	9	27	7	0.9718
	3	5.2	5	13	39	10	0.9806
	4	6.8	7	17	51	14	0.9852
3	1	3.6	4	9	27	7	0.9718
	2	6.8	7	17	51	14	0.9852
	3	10	10	25	75	20	0.9900
	4	13.2	13	33	99	26	0.9924
4	1	6.8	7	17	51	14	0.9852
	2	13.2	13	33	99	26	0.9924
	3	19.6	20	49	147	39	0.9949
	4	26.0	26	65	195	52	0.9962

一階的盒型濾波器(類似哈爾小波濾波器)則是用於特徵點的鄰近區域梯度運算以及給予特徵點描述子。而二階盒型濾波器則是和積分影像進行捲積計算來間接求得每個像素的海森矩陣行列式值。下圖 7. 是一階盒型濾波器和二階盒型濾波器的示意圖。

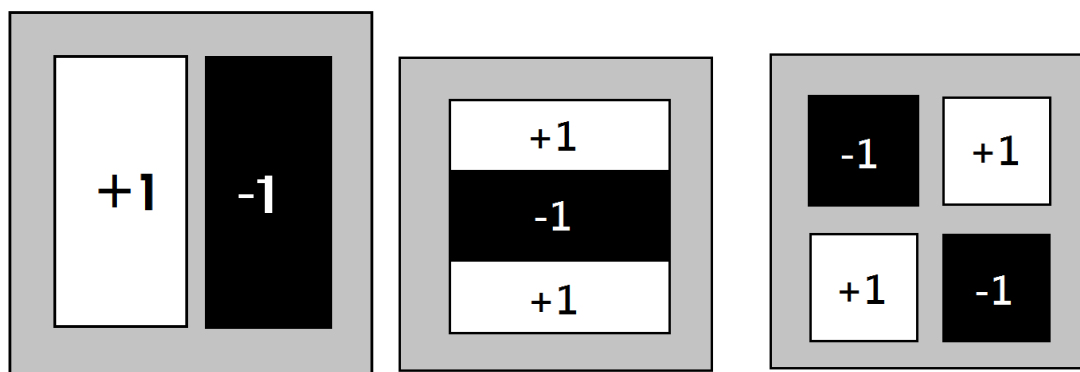


圖 7. 一階與二階盒型濾波器

三、海森矩陣 (Hessian Matrix)

海森矩陣是一個多變量函數的二階偏導數組成的方塊矩陣，經過積分影像與各個濾波器捲積後可分別得到海森矩陣的元素值如下式。SURF 演算法就是計算每個像素的海森矩陣行列式值如下式，因為海森矩陣中的元素是由盒型濾波器(近似濾波器)而來的，因此為保有近似濾波器以及原高斯濾波器的差異，需引入權值概念來減少行列式值誤差量，其權值約為 0.912。

$$H(x, \sigma) = \begin{bmatrix} D_{xx}(x, \sigma) & D_{xy}(x, \sigma) \\ D_{xy}(x, \sigma) & D_{yy}(x, \sigma) \end{bmatrix}$$
$$DoH^L(u) := \frac{1}{L^4} (D_{xx}^L u \cdot D_{yy}^L u - (w D_{xy}^L u)^2)$$

計算行列式值後，將透過設定門檻值以及區域非最大抑制的方法來萃取各尺度下的偵測特徵點。因為上述的行列式值式經過歸一化，因此各尺度的門檻值皆設定在 1000~50000 之間。接著需使用區域非最大值抑制來比對該像素與鄰近的 26 個像素行列式值大小，其中 26 個像素是該層的周圍 8 個像素和上下層的 18 個像素如下圖 8.，若為區域最大值才可被視為偵測特徵點。

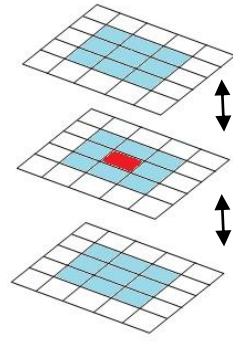


圖 8. 區域性非最大值抑制

第二節、特徵描述

在特徵描述段落中，必須準確指出每一個所偵測到的特徵點所在的空間尺度為何，並使用不同的圓盤狀範圍來偵測特徵點的主方向以達到特徵點的旋轉不變性，最後根據主方向建立 64 個數值來當作該特徵點的描述子。

一、特徵點主方向 (Dominate orientation)

特徵點需確定被偵測出的空間尺度(σ_k)，依據尺度設立半徑為 $6\sigma_k$ 的圓盤狀偵測區，並以特偵點為中心，計算範圍內 $6\sigma_k$ 的所有一階盒型濾波器響應值(計算各像素梯度值)，為減少距離的影響，尚須乘上尺度為 $2\sigma_k$ 高斯函數的權重，如下式：

$$\begin{aligned} \forall (x, y) \in \Omega \cap \beta_{6\sigma_k}(x_k, y_k), \phi_k(x_k, y_k) \\ := \begin{pmatrix} D_x^{L_k} \\ D_y^{L_k} \end{pmatrix} \circ u(x, y) \cdot G_1\left(\frac{x - x_k}{2\sigma_k}, \frac{y - y_k}{2\sigma_k}\right) \end{aligned}$$

接著將圓盤劃分出 40 個主方向($k = 0, 1, 2 \dots 39$)，並搭配扇形掃描範圍 $\pm 30^\circ$ ，依序計算每個 60° 的掃描區域當中的濾波器響應值。在總共 40 個扇形掃描區域中，找出其響應值最大者即為該特徵點的主方向，決定主方向操作示意圖參考下圖 9。

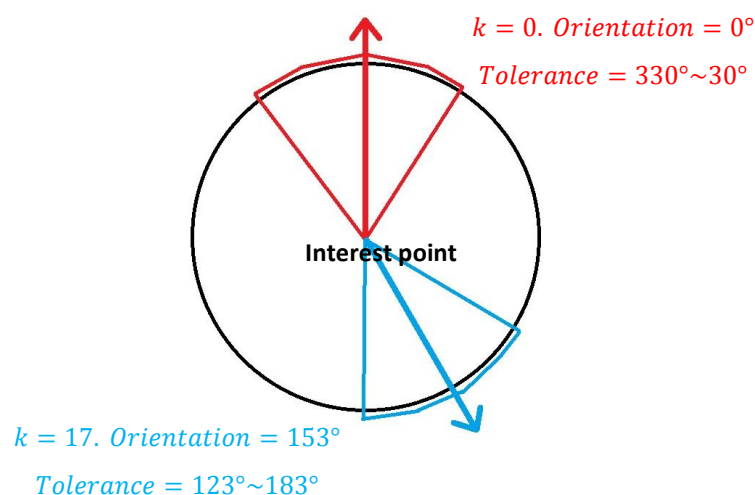
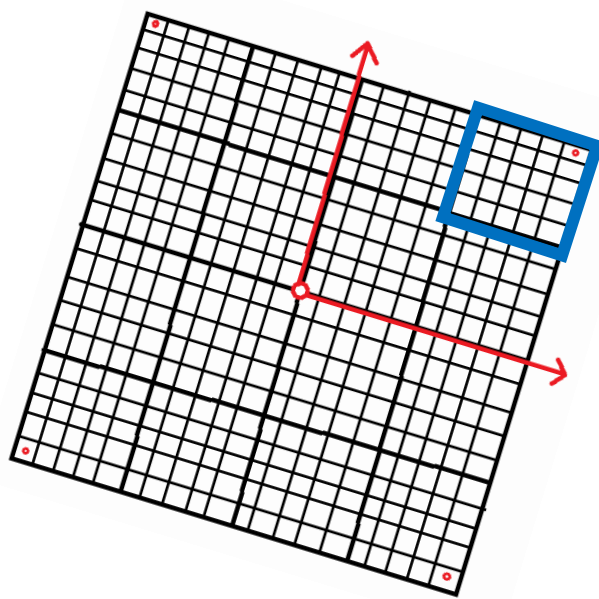


圖 9. 特徵點決定主方向示意圖

二、特徵點描述子 (Descriptor)

決定主方向後，將特徵點原本的影像座標系轉換至以主方向為主的座標系統，並建立一個 4×4 的描述區域，每一個子區域是由 25 個邊長為 σ_k 的方格所組成。這 25 個方格會決定每一個子區域 4 個描述量，分別為 $d_x, d_y, |d_x|, |d_y|$ ，而全部的 16 個子區域又會決定該特徵點的 64 個描述量，稱為該特徵點的向量描述子，如下圖 10。



中心為特徵點，建立一組 4×4 個子區域，其中藍色框為一個子區域，子區域內又包含 5×5 個方格，每個方格邊長為 σ_k 。

圖 10. 特徵點描述子

第三節、特徵匹配

每個被偵測出的特徵點都擁有一個 16×4 的描述子，現有二張待匹配影像，比對兩影像中特徵點的歐幾里得距離的最小差值，即可能為最適合的匹配點，但匹配成果尚須其他匹配策略或是除錯程式來剔除錯誤的匹配點，例如: NN-DR 和 RANSAC。

一、特徵比對

經過偵測和描述後，兩張影像各有數百個特徵點，計算兩張影像的兩個點間的歐幾里得距離如下式，將距離最小視為暫時的最佳匹配點。但是數百點對數百點的匹配運算量過大，所以需先計算各點的拉

普拉斯正負號如下式，若兩點的正負號相反，則有很大的機率這兩點並非適合的匹配點，可事先去除以減少計算量。

$$d_{k,l}^2 = \|X_k - Y_l\|_2^2 = 2(1 - X_k^T \cdot Y_l)$$

$$\Delta^L(u) := \frac{1}{L^2} (D_{xx}^L + D_{yy}^L)u$$

二、匹配策略

歐幾里得距離最小並非準確的就是最佳匹配點，故採用「最鄰近區域距離比例(NN-DR)」匹配策略來檢核匹配成果。作法為選出匹配的歐幾里得距離最小值以及第二小的距離，求出的兩值比例若小於某個門檻值($t = 0.8$)，則可以決定為最佳匹配點，操作示意圖如下圖 11。

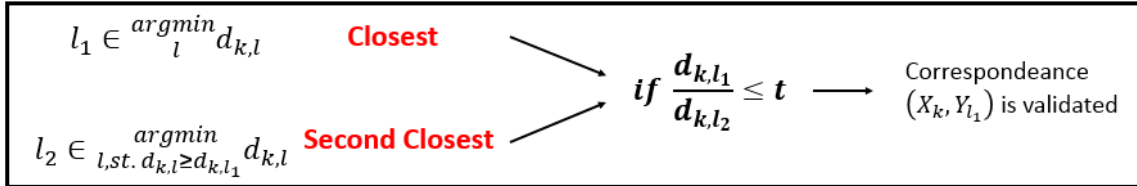


圖 11. 最鄰近區域距離比率匹配策略

第四章 相對方位以及物點三維座標

有了兩立體像對地共軛點就可以使用其座標回復兩相機拍攝瞬間的相對方位，在此將求相對方位的方法分為兩種，一為傳統攝影測量的共面式，另一種則是電腦視覺技術。共面式是大家熟知的方法，但共面式屬非線性方程式，所以必須線性化，過程需要匯入待求參數的初始值並進行迭代計算，若無良好的初始值，就容易得到發散解。所以現在較常用電腦視覺技術來求解相對方位參數，而主要是利用其中的本質矩陣(Essential Matrix)以及基礎矩陣(Fundamental Matrix)，透過演算法將矩陣進行分解，得兩相機間的相對旋轉以及相對距離矩陣。最後，在區域性座標系統性，將兩相機位置以及旋轉使用前方交會，求得物點的三維空間資訊。

第一節、共面式

共面式是指三個向量在同一面的幾何關係如圖 12.，而這三向量分別為基線(兩相機間的向量)、第一相機到物點的射線以及第二相機到物點的射線，這三向量因為在同一平面，故其外積為零。因為共面式原先為非線性方程式，需予以線性化，可得到最後的線性觀測方程式如下式。觀測方程式中有兩個係數，分別為共面條件係數 Q 以及投影係數 N ，其中共面條件係數 Q 為兩條射線的 Y 方向視差，因此當完成共面條件的建立時，其 Q 係數會相當接近零，因此我們會把 Q

當作共面條件的觀測量，倘若沒有建立良好的共面條件，Q 觀測量會偏離零，導致無法獲得正確相對方位。

$$Q = B_x d\mu - B_x \frac{Y_2}{Z_2} dv - \left(Z_2 + \frac{Y_2^2}{Z_2} \right) N_2 d\omega - \frac{Y_2 X_2}{Z_2} N_2 d\varphi + X_2 N_2 dk$$

因為轉換為線性方程式，再進行迭代計算之前必須輸入良好的近似值，但是對於近景攝影測量，初始值的估計較不易且不方便，因此除了拍攝前的規畫，尚可以使用電腦視覺的 Structure from motion 技術回復拍攝瞬間的相機外方位參數概略值，將概略值當作方程式的初始估計值代入迭代計算，若為良好初始值，可在低迭代次數得到收斂的外方位參數；若為偏差量大的初始值，則會得到發散解。一對共軛點可以獲得一個觀測方程式，因為相對方位共有 6 個待求參數，故最少需要 6 個共軛點求解，若大於 6 個則可使用最小二乘法求得精度較好的 6 個相對方位參數。

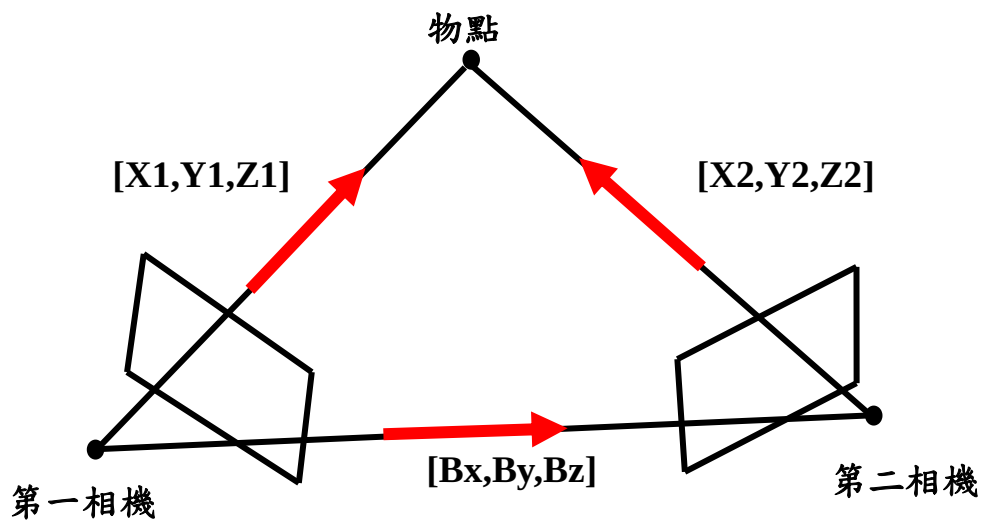


圖 12. 共面條件

第二節、電腦視覺

電腦視覺中必須先介紹本質矩陣以及基礎矩陣。本質矩陣是一個 3×3 的矩陣，用來描述兩相機的相對方位，而基礎矩陣也是一個 3×3 的矩陣，用來描述相機的相對方位以及相機內方位參數，因此可以發現兩矩陣是可以透過相機內部參數互相轉換的，本質矩陣以及基礎矩陣如下表 5.。

表 5. 基礎矩陣及本質矩陣介紹表

	基礎矩陣	本質矩陣
定義	描述兩立體像對間的相對方位以及相機的內部參數。	描述兩立體像對中的相對方位。
公式	$F = K'^{-T}[t]_{\times}RK^{-1}$	$E = [t]_{\times}R$
形式	$\begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{32} \\ f_{31} & f_{23} & f_{33} \end{bmatrix}$	$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$

如何求得本質矩陣或是基礎矩陣？我們會使用八點法如下式。在八點法中，我們通常會輸入立體像對影像坐標到八點法，因為待求參數共有 8 個(基礎矩陣的元素值)，所以我們至少需要 8 個共軛點才能求的唯一解，若有更多的共軛點，則可利用最小二乘法獲得精度較高

的解。針對八點法，若輸入的參數是影像座標，則獲得的元素值就是基礎矩陣的元素值；但若輸入元素為像點的相機座標系，則獲得的元素值為本質矩陣的 9 個元素值。

$$uu'f_{11} + vu'f_{12} + u'f_{13} + uv'f_{21} + vv'f_{22} + v'f_{23} + uf_{31} + vf_{32} + f_{33} = 1$$

有了基礎矩陣後，可以使用相機內部參數轉換成本質矩陣，此時就會擁有一個 3×3 的本質矩陣，但因為矩陣元素間的相依性質很高，所以必須透過奇異值分解如下式來拆解成相對平移以及相對旋轉兩個矩陣。在奇異值分解後，共可以得到四種不同的解，我們稱為交會模糊性如圖 13.，在四種解中只有一組的真正的前方交會以及正確位置，其餘都存在後方交會或是相機位置錯誤。因此我們可以透過一些方法，例如測試相機和物體深度或是檢查相機方向和物體間的幾何關係來確認何者為正確解，即獲得正確的相對方位以及相對旋轉。

$$E = U \Sigma V^T$$

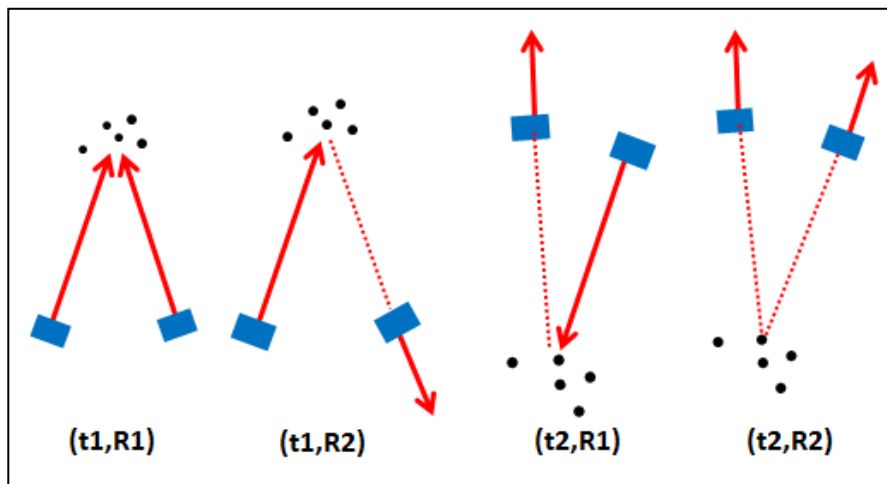


圖 13. 交會模糊性(共 4 解)

第三節、物體三維空間資訊

使用電腦視覺方法可以獲得相對旋轉矩陣 R 以及相對距離矩陣 T ，將 R 和 T 矩陣依 $P \sim K[R \ T]$ 結合，稱為投影矩陣 P ，用來描述兩相機的相對外方位參數。有了投影矩陣 P ，物點三維座標可以由量測的共軛點對透過兩條光束的後方交會獲得，參考圖 14.。但是因為量測的雜訊，所以後方交會的兩條光線並不會真實的交在一起，因此在假設雜訊誤差量是屬於高斯分布的情況下，需要選擇像點的理想位置以及實際位置的最小方差，如下式。對於兩張影像的模型重建，可以使用非迭代的求解系統獲得三維座標，但是對於 2 張影像以上的模型重建，則須倚靠迭代的非線性求解系統。

$$X = \operatorname{argmin} \sum_i \|u_i - \hat{u}_i(P_i, X)\|^2$$

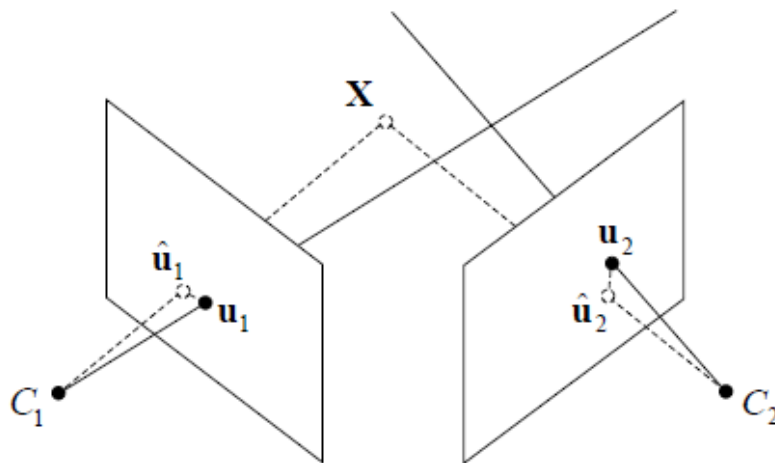


圖 14. 具有量測雜訊的前方交會

第五章 實驗成果與分析

第一節、實驗一：近景攝影之小型物體模型重建

在重建大型三維模型前，先檢核上述演算法的實行程度以及正確性，因此選取體型較小且具有規則排列的平版模型進行建模。除了驗證演算法的可行性，此外也驗證選取的共軛點的方法是否會影響建模成功，因此實驗一採取人工選點，而實驗二採用自動化特徵匹配法 SURF 找尋共軛點。若實驗一的建模成果完成度高，再依照類似的步驟進行大型物件的建模。

一、實驗對象與研究基本資訊

實驗一針對一個立體的平板組進行二個角度拍攝如圖 15. 平板中的黑白間隔長度相等，便於模型重建後檢查成果。模型重建之初需要在影像中找尋共軛點，因為該影像物體是具有穩定且規則的對比的紋路和特徵，因此採用人工選取特徵點找尋 32 個特徵點，平均分布於黑白間隔中，如圖 15. 中的紅色十字。

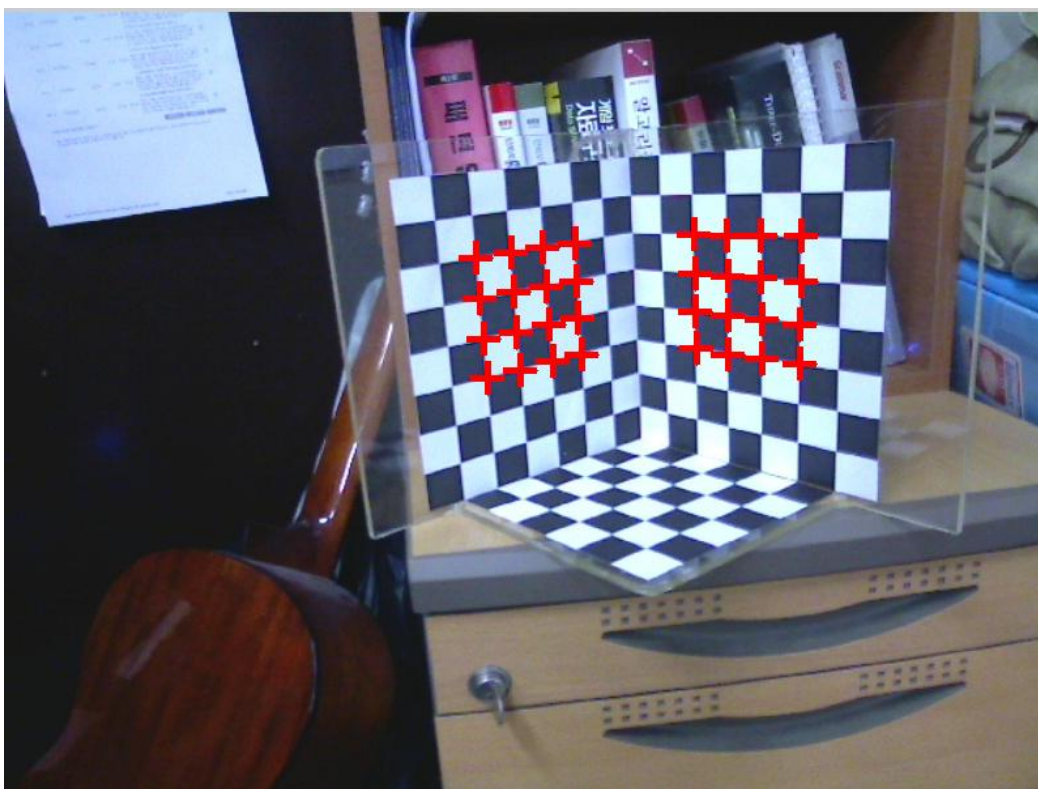


圖 15. 實驗一的立體像對(紅標為共軛點)

二、回復相對方位

回復拍攝瞬間兩相機的相對方位，有上述的電腦視覺以及傳統的共面式攝影測量。首先針對電腦視覺方法進行實驗，將 32 個共軛點的影像座標輸入八點法演算法後並透過相機參數轉換後，可以得到下表 6. 的基礎矩陣以及本質矩陣。因為在本質矩陣間的 9 個元素間彼此存在高相依性，因此並無法直接從數值間看出相機相對方位的幾何特性，所以使用奇異值分解演算法後可以得到四組方向不同的相對平移和旋轉的矩陣如表 7.。得到四組解後須檢查交會模糊性，目的為在四組解當中需選出一個正解，該正解為相機方向前方交會於物空間的物點上，當確認何種為正解後，就可以在區域性座標系統下計算出第一及第二相機的相對位置和旋轉如下表 8.。

表 6. 實驗一立體像對的基礎及本質矩陣

	基礎矩陣	本質矩陣
數 值	$\begin{bmatrix} -1.7718 \times 10^{-7} & 2.5560 \times 10^{-7} & 3.4718 \times 10^{-4} \\ -1.3424 \times 10^{-6} & -1.6680 \times 10^{-7} & 0.0028 \\ 1.6667 \times 10^{-4} & -0.0024 & -0.1535 \end{bmatrix}$	$\begin{bmatrix} -0.0837 & 0.1127 & 0.2471 \\ -0.5917 & -0.0686 & 1.4394 \\ -0.2159 & -1.5108 & -0.0042 \end{bmatrix}$

表 7. 實驗一求解之相對平移及旋轉矩陣

	平移矩陣	旋轉矩陣
第一組解	$T = \begin{bmatrix} -0.9824 \\ 0.1684 \\ -0.0809 \end{bmatrix}$	$R = \begin{bmatrix} 0.9539 & 0.0367 & 0.2979 \\ -0.0356 & 0.9993 & -0.0093 \\ -0.2980 & -0.0017 & 0.9545 \end{bmatrix}$
第二組解	$T = \begin{bmatrix} 0.9824 \\ -0.1684 \\ 0.0809 \end{bmatrix}$	$R = \begin{bmatrix} 0.9539 & 0.0367 & 0.2979 \\ -0.0356 & 0.9993 & -0.0093 \\ -0.0017 & -0.0017 & 0.9545 \end{bmatrix}$
第三組解	$T = \begin{bmatrix} -0.9824 \\ 0.1684 \\ -0.0809 \end{bmatrix}$	$R = \begin{bmatrix} 0.8517 & -0.2968 & 0.4319 \\ -0.2740 & -0.9547 & -0.1158 \\ 0.4467 & -0.0197 & -0.8944 \end{bmatrix}$
第四組解	$T = \begin{bmatrix} 0.9824 \\ -0.1684 \\ 0.0809 \end{bmatrix}$	$R = \begin{bmatrix} 0.8517 & -0.2968 & 0.4319 \\ -0.2740 & -0.9547 & -0.1158 \\ 0.4467 & -0.0197 & -0.8944 \end{bmatrix}$

表 8. 實驗一電腦視覺方式求出的相機座標和旋轉角度

	XL1	YL1	ZL1	Omega1	Phi1	Kappa1
第一相機	0	0	0	0	0	0
第二相機	0.91897	-0.13237	0.3714 5	0.10384	-17.34030	2.1368 0

因為要確認使用電腦視覺方式求出的相對方位是否正確所以再使用共面式來檢核，表格 9. 是建立在區域性座標性統下的第一及第二相對方位。而表格 10. 是計算兩個方法所求出來的相對方位差值，從差值可以觀察出距離的最大差異值為 0.00001，而角度最大差異值為 0.0007，兩差值皆在容許誤差範圍內，故可證明兩個方法皆可應用在攝影測量的相對方位求解中。

表 9. 實驗一共面式方式求出的相機座標和旋轉角度

	XL2	YL2	ZL2	Omega2	Phi2	Kappa2
第一相機	0	0	0	0	0	0
第二相機	0.91897	-0.1323 6	0.37144	0.10396	-17.34100	2.1368 0

表 10. 實驗一兩方式所求出之相對方位的差值

	dXL	dYL	dZL	dOmega	dPhi	dKappa
差值	0	0.00001	-0.00001	0.00012	-0.00070	0

三、物體模型重建

上一段落有提到使用 SVD 後會有四個解需要檢核，我們直接將四組平移和旋轉都拿來重建三維模型，檢核出哪一個才是正確解以及真正的三維模型如圖 16.。圖 16. 中的箭頭為相機的拍攝方向(鏡頭方向)，經比較後可以得知第一組解為正確的前方交會幾何，所以我們拿第一組解當作真實的 3D 模型重建圖並和原始立體像對比較，發現目標物的幾何關係和位置都相當符合如 17.，證明 3D 模型重建成功。

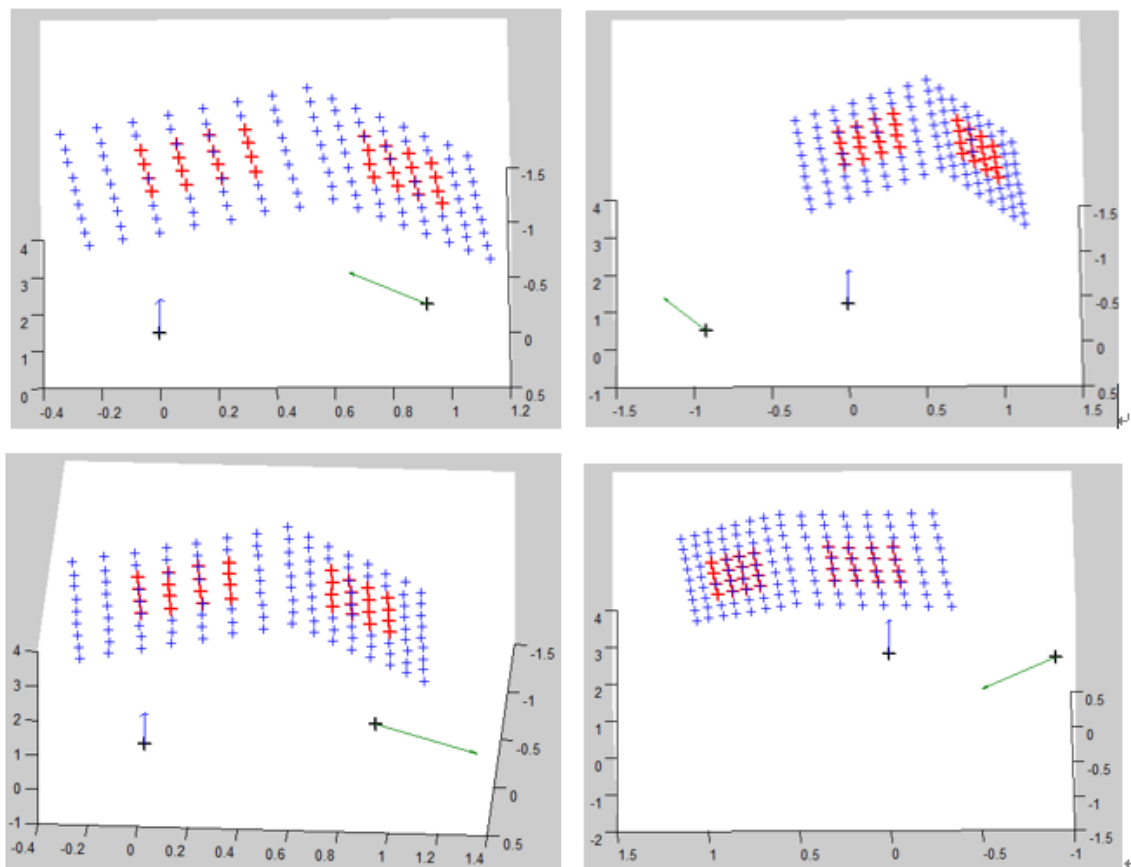


圖 16. 三維模型重建(具交會模糊性)



圖 17. 真實影像與模型重建成果比較圖

第二節、實驗二：近景攝影之大型物體模型重建

實驗一已確認研究方法以及相關法的可行性，因此利用類似的步驟來進行大型物體的模型重建，實驗中尚會比較人工選點以及自動特徵匹配所產生的成果有何不同，並進一步分析並討論模型重建的完成度、錯誤共軛點、基礎矩陣以及錯誤的三維點雲座標。

一、實驗對象與研究基本資訊

以研究目標為主的建物模型重建來選取永康一棟建物拍攝兩張像片如圖 18.，採用的相機是 SONY NEX-5N，表 11.描述這立體像對的基本資訊。而本實驗的特徵點選取方式會採用自動化特徵匹配及人工偵測兩種，如此可以比較兩個方法所重建的模型差異、成果好壞以及處理時間。

表 11. 實驗二研究對象立體像對基本資訊

影像尺寸	4912 × 3264
感測器尺寸	23.4mm × 15.6mm
水平解析度	350dpi
垂直解析度	350dpi
焦距	16mm

相機矩陣	$\begin{bmatrix} 3358.6 & 0 & 2456 \\ 0 & 2040 & 1632 \\ 0 & 0 & 1 \end{bmatrix}$
------	---



圖 18. 實驗二立體像對

二、特徵匹配成果

本研究使用的特徵匹配是 SURF，因為立體像對屬於高解析度影像，SURF 特徵匹配可以有效地縮短匹配時間。下圖 19.是原 SURF 的匹配成果，而下圖 20.過 RANSAC 除錯的匹配成果，因為如果只使用 SURF 的匹配成果中會存在部分的錯誤匹配點，錯誤匹配點在計算立體像對的基礎矩陣時會有影響，因此使用 RANSAC 來移除錯誤的匹配點，下表 12.展現是否經過 RANSAC 除錯的前後特徵點比對。

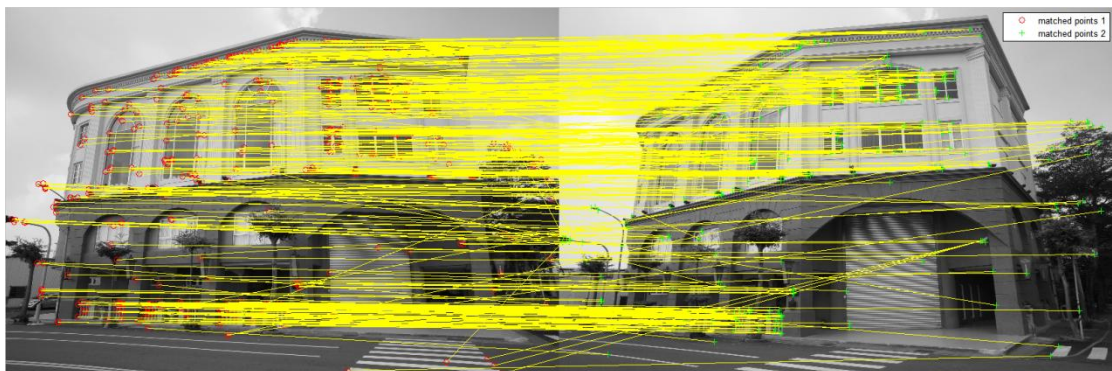


圖 19. 使用 SURF 的匹配成果

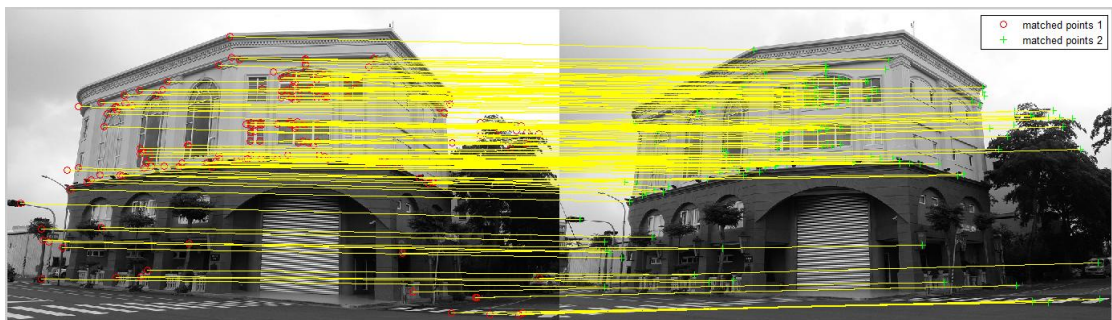


圖 20. 使用 SURF+RANSAC 的匹配成果

表 12. 使用 SURF 和 SURF+RANSAC 的成果比對表

	SURF	SURF+RANSAC
條件	X	迭代次數 10，門檻值 150
特徵點數量	409	99
匹配成果	存在錯誤匹配點	成功移除部分錯誤匹配點

三、檢核對極幾何

使用共軛點計算立體像對的基礎矩陣和本質矩陣後，並無法從簡單的陣列中獲得幾何資訊，因此為檢核本質矩陣的正確性，我們透過對極幾何的關係來完成。在對極幾何中，若影像 1 和影像 2 是研究的立體像對，那影像 1 的共軛點和基礎矩陣相乘後會得到在影像 2 的一條極線，而影像 1 點位在影像 2 中的對應點會位在極線上，但因為量測誤差或是匹配的些微偏差，可能會讓對應點坐落在極線附近。因此我們透過計算共軛點的極線，如圖 21.和圖 22.，可以看到共軛點對在兩張影像當中都位在極線的上方或是附近，因此可以確認相機參數、特徵點匹配以及基礎矩陣的計算都具有正確性。表 13.列出該立體像對的基礎矩陣以及本質矩陣數值。

表 13. 實驗二(特徵匹配)的基礎矩陣及本質矩陣

	數值
基礎矩陣	$\begin{bmatrix} -3.0756 \times 10^{-10} & 6.5362 \times 10^{-8} & -1.7900 \times 10^{-4} \\ -3.8136 \times 10^{-8} & 3.0578 \times 10^{-9} & -2.9737 \times 10^{-4} \\ 1.0648 \times 10^{-4} & 2.5319 \times 10^{-4} & 0.1229 \end{bmatrix}$
本質矩陣	$\begin{bmatrix} -0.0035 & 0.4478 & -0.2455 \\ -0.2613 & 0.0127 & -0.7875 \\ 0.1461 & 0.8542 & -0.0119 \end{bmatrix}$



圖 21. 檢核影像 2(特徵匹配)的對極幾何



圖 22. 檢核影像 1(特徵匹配)的對極幾何

另外，本實驗也採用人工選點，總共選取 30 個共軛點。人工選點的工作雖然相當耗費時間，而且數量少，但是可以有效的降低錯誤共軛點的狀況，而且在模型重建的步驟中，可以只重建選取的共軛點對，讓模型具有簡單的特徵。下圖 23. 和圖 24. 是檢核對極幾何的立體像對，很明顯的共軛點對也都位在極線上或是附近。表 14. 列出該立體像對的基礎矩陣以及本質矩陣數值。

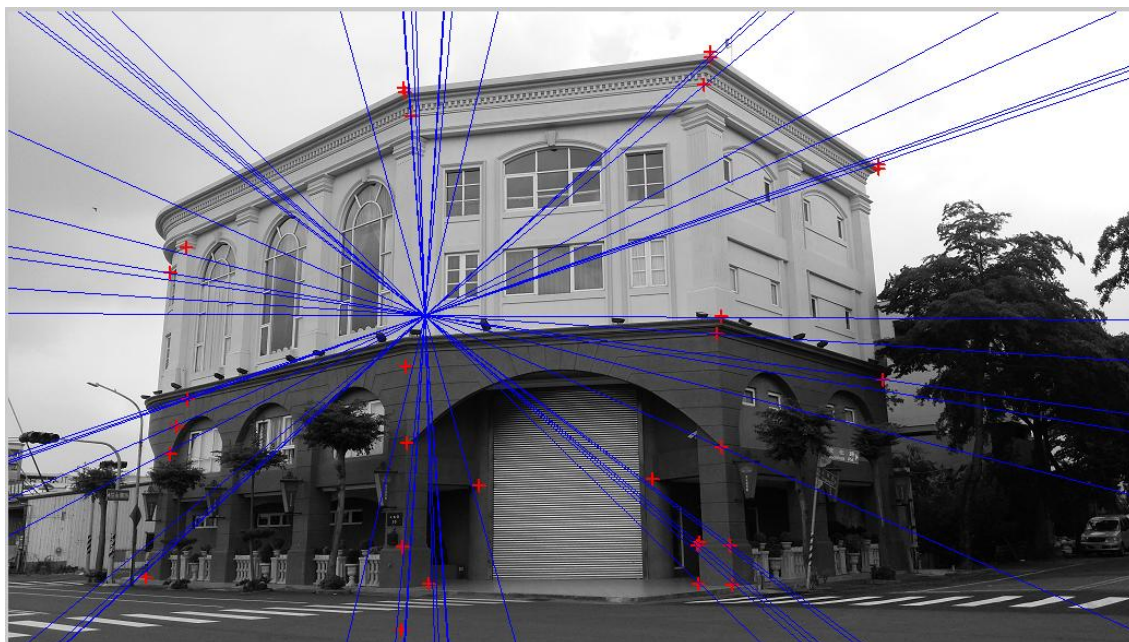


圖 23. 檢核影像 2(人工匹配)的對極幾何

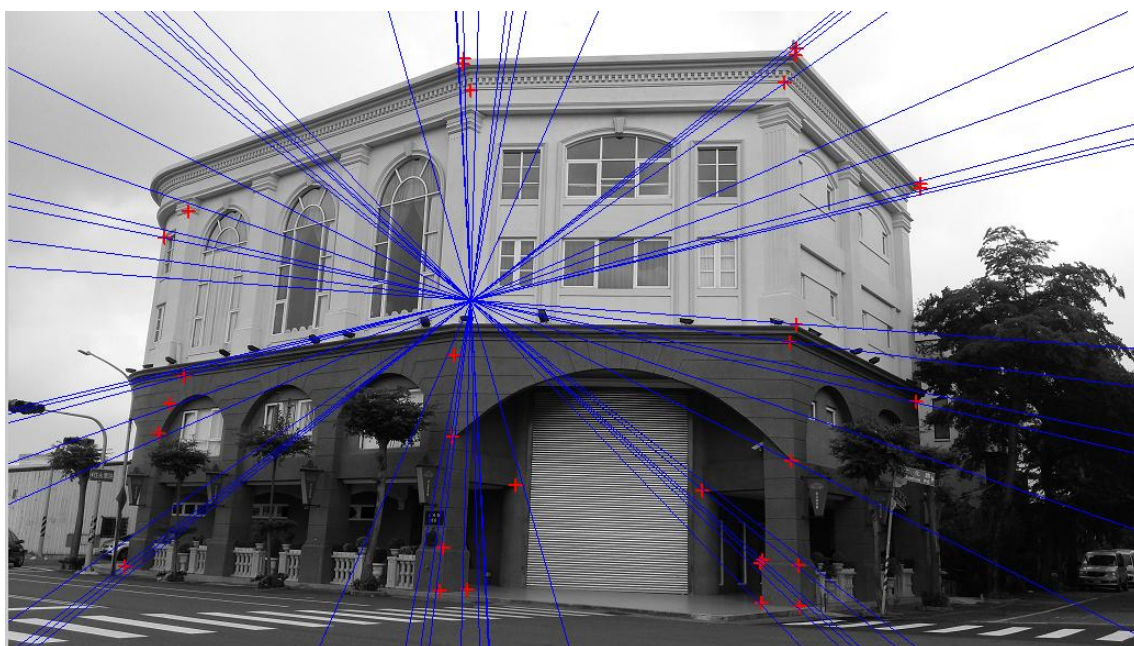


圖 24. 檢核影像 1(人工匹配)的對極幾何

表 14. 實驗二(人工匹配)的基礎矩陣及本質矩陣

	數值
基礎矩陣	$\begin{bmatrix} 3.2310 \times 10^{-7} & 4.6897 \times 10^{-6} & -0.0011 \\ -4.4689 \times 10^{-6} & -1.9694 \times 10^{-7} & 0.0019 \\ 0.0014 & -0.0017 & -0.1203 \end{bmatrix}$
本質矩陣	$\begin{bmatrix} -0.0552 & 0.8014 & 0.0382 \\ -0.7637 & 0.00337 & -0.1755 \\ -0.0380 & 0.2586 & 0.0077 \end{bmatrix}$

四、回復相對方位

先將使用特徵匹配的成果輸入八點法再使用奇異值分解獲得表 15. 的 4 組平移和旋轉矩陣，最後選出一組正確的相對方位，透過區域性座標系統(以第一相機位置為原點，第一相機的鏡頭方向為 z 軸)，算出第二台相機的位置和旋轉角度，如表 16. 。

表 15. 實驗二(特徵匹配)求解之相對平移及旋轉矩陣

	平移矩陣	旋轉矩陣
第一組解	$T = \begin{bmatrix} -0.8567 \\ 0.2603 \\ 0.4453 \end{bmatrix}$	$R = \begin{bmatrix} 0.6192 & -0.3842 & -0.6848 \\ -0.3822 & -0.9091 & 0.1634 \\ -0.6853 & 0.1613 & -0.7101 \end{bmatrix}$
第二組解	$T = \begin{bmatrix} 0.8567 \\ -0.2603 \\ -0.4453 \end{bmatrix}$	$R = \begin{bmatrix} 0.6192 & -0.3842 & -0.6848 \\ -0.3822 & -0.9091 & 0.1634 \\ -0.6853 & 0.1613 & -0.7101 \end{bmatrix}$

第三組解	$T = \begin{bmatrix} -0.8567 \\ 0.2603 \\ 0.4453 \end{bmatrix}$	$R = \begin{bmatrix} 0.9836 & 0.1026 & 0.1485 \\ -0.1037 & 0.9946 & -0.0004 \\ 0.1477 & -0.0150 & 0.9889 \end{bmatrix}$
第四組解	$T = \begin{bmatrix} 0.8567 \\ -0.2603 \\ -0.4453 \end{bmatrix}$	$R = \begin{bmatrix} 0.9836 & 0.1026 & 0.1485 \\ -0.1037 & 0.9946 & -0.0004 \\ 0.1477 & -0.0150 & 0.9889 \end{bmatrix}$

表 16. 實驗二(特徵匹配)電腦視覺求出的相機座標和旋轉角度

	XL1	YL1	ZL1	Omega1	Phi1	Kappa1
第一相機	0	0	0	0	0	0
第二相機	0.9354	-0.1643	-0.3130	0.8666	-8.4965	6.0171

再來使用人工匹配的成果輸入八點法再使用奇異值分解獲得表 17. 的 4 組平移和旋轉矩陣，最後選出一組正確的相對方位，透過區域性座標系統(以第一相機位置為原點，第一相機的鏡頭方向為 z 軸)，算出第二台相機的位置和旋轉角度，如表 18.。

表 17. 實驗二(人工匹配)求解之相對平移及旋轉矩陣

	平移矩陣	旋轉矩陣
第一組解	$T = \begin{bmatrix} -0.3079 \\ -0.0251 \\ 0.9511 \end{bmatrix}$	$R = \begin{bmatrix} 0.9943 & 0.0624 & -0.0867 \\ -0.0606 & 0.9979 & 0.0237 \\ 0.0880 & -0.0183 & 0.9960 \end{bmatrix}$
第二組解	$T = \begin{bmatrix} 0.3079 \\ 0.0251 \\ -0.9511 \end{bmatrix}$	$R = \begin{bmatrix} 0.9943 & 0.0624 & -0.0867 \\ -0.0606 & 0.9979 & 0.0237 \\ 0.0880 & -0.0183 & 0.9960 \end{bmatrix}$

第三組解	$T = \begin{bmatrix} -0.3079 \\ -0.0251 \\ 0.9511 \end{bmatrix}$	$R = \begin{bmatrix} -0.8582 & -0.0245 & -0.5127 \\ 0.0717 & -0.9948 & -0.0725 \\ -0.5083 & -0.0990 & 0.8555 \end{bmatrix}$
第四組解	$T = \begin{bmatrix} 0.3079 \\ 0.0251 \\ -0.9511 \end{bmatrix}$	$R = \begin{bmatrix} -0.8582 & -0.0245 & -0.5127 \\ 0.0717 & -0.9948 & -0.0725 \\ -0.5083 & -0.0990 & 0.8555 \end{bmatrix}$

表 18. 實驗二(人工匹配)電腦視覺求出的相機座標和旋轉角度

	XL1	YL1	ZL1	Omega1	Phi1	Kappa1
第一相機	0	0	0	0	0	0
第二相機	0.2210	-0.0617	-0.9733	1.0515	5.0468	3.4874

五、物體模型重建

使用電腦視覺中的三角測量可以獲得共軛點對所對應的三維點雲，圖 25.是將所有由 SURF+RANSAC 所匹配的共軛點對去建立點雲資料，由圖中可以明顯看出少數的點雲偏離點群，這些點把它視為錯誤點，其原因有二，第一它可能是錯誤匹配點，但是沒被 RANSAC 偵測出來，第二可能是因為線性三角測量的轉換並不適用於所有的幾何交會。圖 26.是圖 25.的局部放大，主要檢視建物點群的重建情形，由點群分布來說，此實驗並無法產製相當良好的建物點雲圖，因為從點雲中並無法看到房屋的幾何形狀，而且比對房屋相片和三維點雲圖後，並沒有得到直接相關聯的幾何關係。但是依舊可以觀察到點群有集中的情形，符合相片中所匹配成功的共軛點對分布。圖 27.中的紅十字和箭

頭是分別表示相機位置以及相機方向，從位置來看，相機距離是符合當時拍攝的距離比例，而且相機的箭頭是朝向建物拍攝，所以方向也是正確的。圖 28.是將點雲資料翻轉，換個角度觀察模型重建成果。

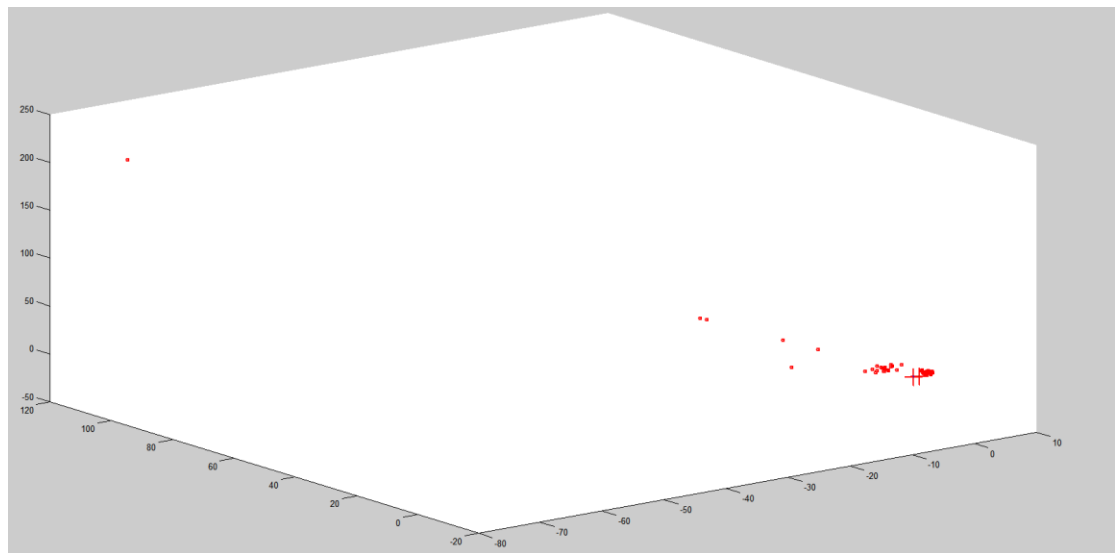


圖 25. 建物模型重建圖角度 1(特徵匹配)

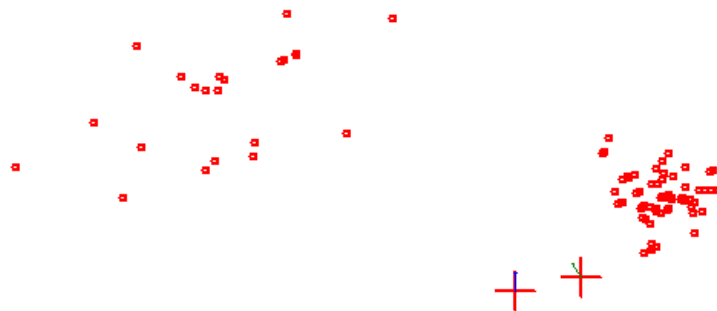


圖 26. 建物模型重建放大圖 1

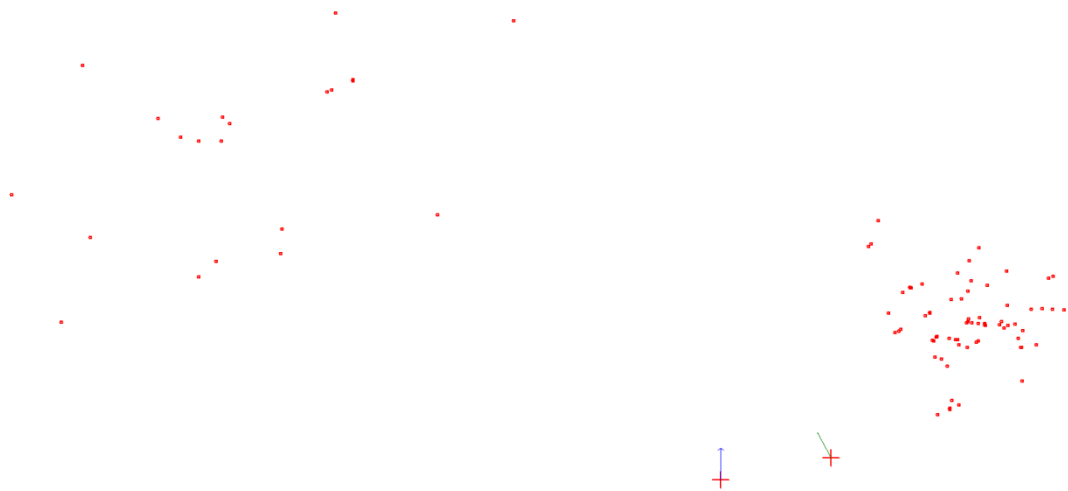


圖 27. 建物模型重建放大圖 2

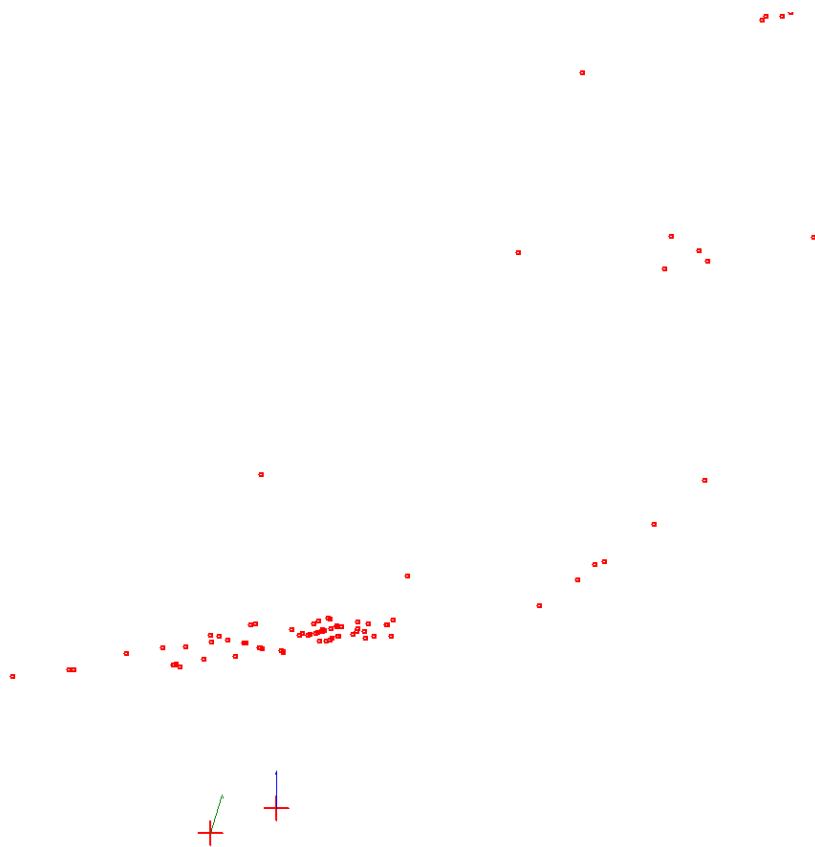


圖 28. 建物模型重建圖角度 2(特徵匹配)

接著根據人工選取匹配點的成果進行模型重建，得到的點雲資料如圖 29。此成果依舊沒有很好，而且尚存部分雜點，推測原因有二，第一無法獲得良好的相機內方位參數矩陣，第二線性三角測量的演算法存在部分限制。導致雖能呈現點雲，但幾何關係和位置無法完全正確。

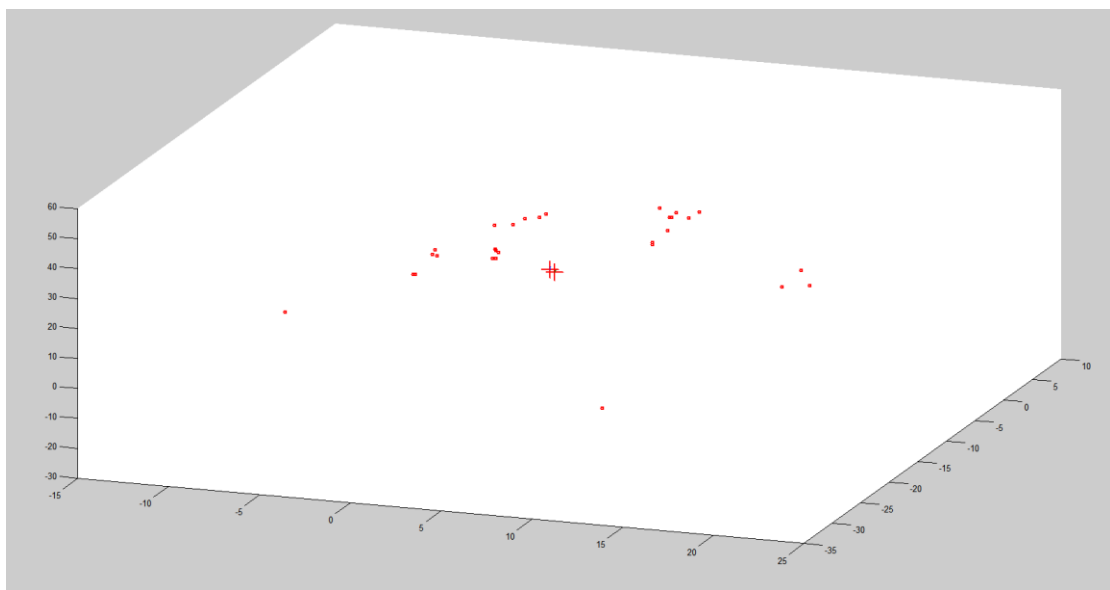


圖 29. 建物模型重建圖(人工匹配)

六、 使用商用軟體 VisualSFM 進行物體模型重建

上述成果皆是用 Matlab 自行撰寫程式所執行的建模成果，但因無法獲得良好的模型，所以嘗試使用現有建模軟體 VisualSFM，此簡易軟體可由以下網址下載 <http://ccwu.me/vsfm/install>，屬於開放資料，並且提供多個版本可以適用於眾多電腦，雖然該軟體屬較陽春的建模軟體，但是操作上顯得簡易而對於初步建模是綽綽有餘，軟體介面可

參考圖 30。該軟體我們主要使用到三個功能，第一是匯入具有高覆蓋率的影像群(二張以上)，第二是計算影像間的投影矩陣，在計算投影矩陣前該軟體也會進行特徵匹配，但軟體使用的是 SIFT 特徵匹配，優點是偵測點多，但是缺點是偵測速度慢，第三是物體建模。

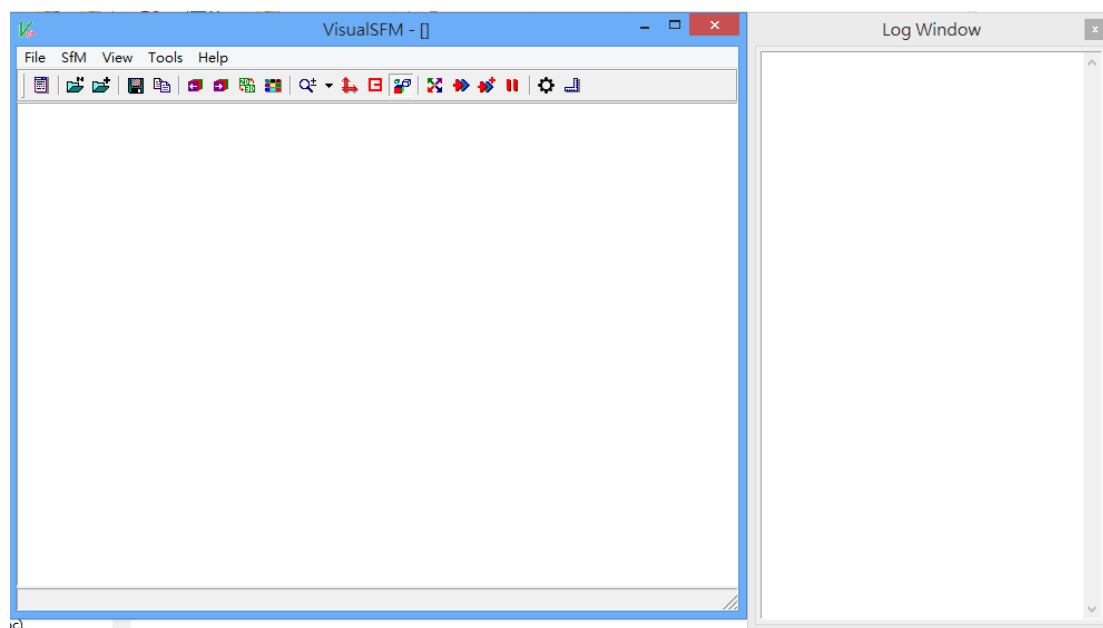


圖 30. VisualSFM 軟體介面

匯入影像後可以在右邊的 Log Window 看見影像資訊，例如影像尺寸和焦距等。接著按下 Compute Missing Matches，就會開始進行影像匹配，並計算個影像間的投影矩陣，當計算完成後就可以按下 Compute 3D Reconstruction 進行模型重建，操作過程的介面可以參考圖 31。該軟體可以處理多張影像，但該研究主要以二張的立體像對作為研究對象，故處理時間較快。

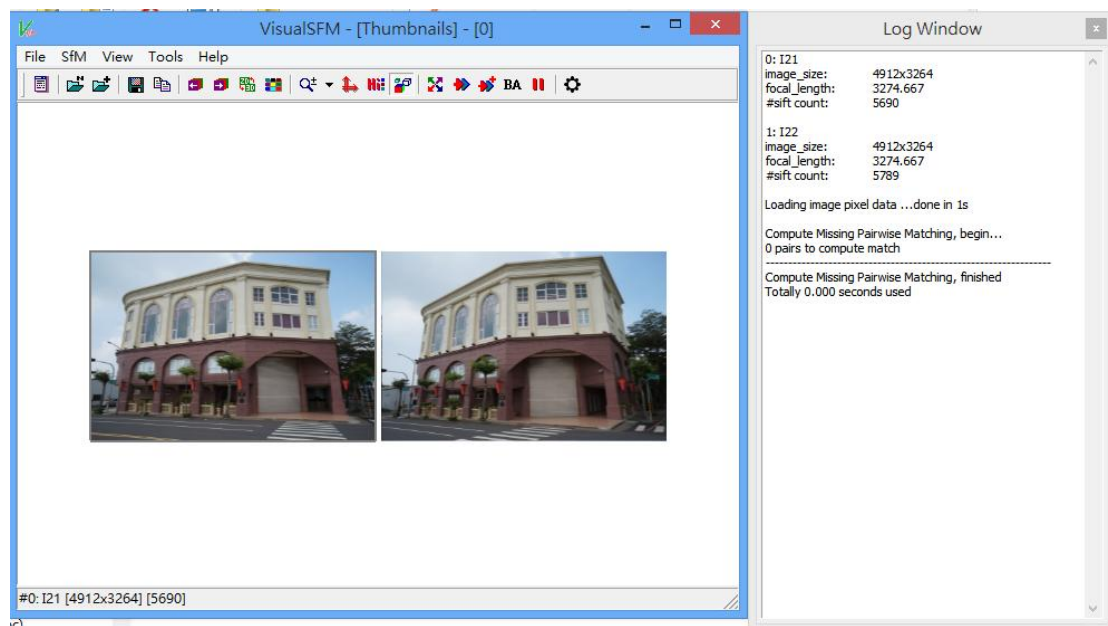


圖 31. VisualSFM 軟體操作過程

經過軟體建模後，圖 32.展現其成果。可以觀察到以下不同，第一點雲數量明顯較多，這應該是使用 SURF 和 SIFT 的主要差別。第二點雲的位置明顯較正確，因為我們兩個拍攝的相片是針對單一建物的一面牆，所以模型重建後應該眾多點雲會呈現在一平面上，比對成果圖發現是相當符合的，而且就點雲的分布可以看出建物單一面牆的外型，若要完整的呈現建物的空間感，就需要環繞建物拍取多張影像，影像間需具有高重疊率。根據該建模的成果，可以推斷若使用商業軟體是可以成功進行房屋點雲資料的建立。

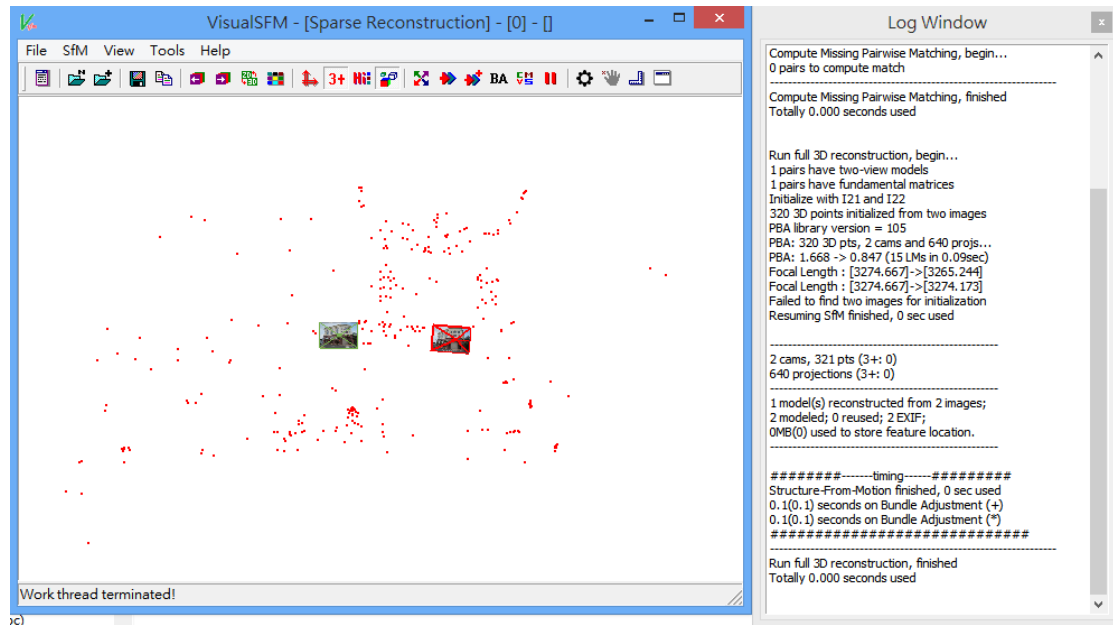


圖 32. VisualSFM 軟體房屋建模成果

第六章 結論及建議

第一節、結論

實驗一的建模成果相當良好，所有的黑板板間的間隔都相當一致，而且板子的外型和相機的拍攝位置以及方向也都符合拍攝情況。這樣的成果可以佐證在共軛點對的選取以及電腦視覺的演算法應用上都是成功的，在這個成功的實驗當中，觀察到共軛點的選取需要足夠量（至少 8 個）以及好的點分布，因為八點法中的矩陣為病態矩陣，若點之間不具良好的幾何關係，例如一直線或是集中於某半側等，容易產生錯誤的基礎矩陣。故根據實驗一可以確認第一章第三節的實驗流程是正確的。

但實驗二所針對的大型物件建模成果就不盡理想，可以明顯發現重建的點雲資料有明顯變形和偏移的現象，導致房屋形狀的變形。此外少數點雲偏移量過大，為明顯大錯，這些必須加以剔除。對於房屋模型點雲資料不佳的狀況有以下四個推測原因，第一是相機內方位參數矩陣(K)不佳，導致描述兩張影像的本質矩陣(E)有錯，第二是部分的錯誤匹配點讓計算後的投影矩陣(P)有錯，第三是使用的電腦視覺直接線性三角測量所適用的三維座標轉換有限制條件或是需具有特定的良好幾何關係，第四是相機間的距離和點雲間的距離尺度不盡相同。另外，在特徵匹配當中發現當高解析度影像進行找尋共軛點時，會產生多個錯誤的匹配點，若沒有搭配 RANSAC 除錯，很容易影響基礎矩陣(F)的計算。若不使用 RANSAC，我們尚可以透過降解析度方式，將較低解析度影像進行 SURF 匹配可以獲得較好的共軛點成果。

第二節、建議

電腦視覺技術確實可以快速且方便的透過影像來產製三維模型資料，不但省時間省人力，也可以讓民眾更加信服地政事務所的相關服務，而且成果也淺顯易懂，所以電腦視覺技術是值得專研的領域之一。在本次的研究中，模型重建並沒有做的相當完整，實驗二中的房屋建模會有變形及偏移的情形，未來可以針對該問題進行探討和解決，讓

初步模型建模成果可以更完整。此外，本研究只針對二張影像進行模型重建，但在電腦視覺中是可以適用於多張影像進行模型重建的，若只有二張影像通常只能產製單一面或是單一視角的點雲資料，對於一個建物的模型是相當缺乏的，因此希望未來能針對多張影像的描述和模型重建有更深入的探討。並嘗試使用更多商業軟體進行物體模型的重建，再試著比較其建模成果。

第七章 參考文獻

（一）期刊論文

Adelson E. H., C. H. Anderson, J. R. Bergen, P. J. Burt and J. M. Ogden, 1984. Pyramid methods in image processing.

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool, 2008. Speeded-up robust features (SURF), Computer Vision and Image Understanding, 110, pp. 346–359.

Hartley, R.I., 1997. "In Defence of the 8-point Algorithm"

Li, C. and Y. Zhao, 2012. "Approach of Camera Relative Pose Estimation Based on Epipolar Geometry" Information Technology Journal, 11: 1202-1210.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints, cascade filtering approach, IJCV, 60(2):91-110.

Lindeberg, T, 1998. Feature detection with automatic scale selection, IJCV, 30(2): 79-116

Longuet-Higgins, H. C., 1981. "A computer algorithm for reconstructing a

scene from two projections” Journal of Nature Vol. 293.

Oyallon, E. and J. Rabin, 2015. An Analysis of the SURF Method, Image Processing On Line.

Scaramuzza, D. and F. Fraundorfer, 2011. “Visual Odometry.Part I: The First 30 Years and Fundamentals”

Simard, Patrice Y, L. Bottou, P. Haffner, Y. LeCun, 1999. Boxlets: a fast convolution algorithm for signal processing and neural networks, Advances in Neural Information Processing Systems, pp. 571–577. ISBN:0-262-11245-0.

Viola, P.A. and M.J. Jones, 2001. Rapid object detection using a boosted cascade of simple features. In CVPR (1), page 511-518.

Xu, A., G. Namit, 2008. SURF: Speeded-Up Robust Features, COMP 558 – Project Report, McGill University.

(二)書籍

Bradski, G. and A. Kaehler, 2008. Learning OpenCV: Computer vision with the OpenCV library, O’Reilly Media, ISBN:978-0-596-51613-0.

(三)網路資料

Chien, J, 2016. “Beginner’s Guide to Fundamental Matrix, Essential Matrix and Camera Motion Recovery”

Mare’s Computer Visual Studying, 2011. “8 point algorithm (Matlab source code) / The method to get the Fundamental Matrix and the Essential matrix”